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 (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 \( U = \{U_1, U_2, ..., U_n, ..., N\} \) 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 \( (x_{ij}, z_{ij}) \) \( (i = 1, 2, 3, ..., k; j = 1, 2, 3, ..., n) \) be the observed values of the study variable \( y \) and the auxiliary variables \((x,z)\), respectively for the \( j^{th} \) unit in the \( i^{th} \) possible
Let \( \bar{y}^{(o)} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} y_{ij} \), \( \bar{x}^{(o)} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} \) and \( \bar{z}^{(o)} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} z_{ij} \) be the systematic sample means corresponding to population means \( \bar{y} = \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{n} y_{ij} \), \( \bar{x} = \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} \) and \( \bar{z} = \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{n} z_{ij} \), respectively.

Let \( \rho_y = 1 + (n - 1) \rho_{y} \), \( \rho_x = 1 + (n - 1) \rho_{x} \) and \( \rho_z = 1 + (n - 1) \rho_{z} \), where \( \rho_{y}, \rho_{x}, \rho_{z} \) are the systematic sample means of \( y, x, z \), respectively. Let \( \rho_{xy}, \rho_{xz}, \rho_{yz} \) be the population correlation coefficients between \( y, x \), \( y, z \) 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.

The intra-class correlation coefficients between pairs of units within the same systematic sample for \( y, x \) and \( z \), respectively. Let \( \zeta_{y} \), \( \zeta_{x} \) and \( \zeta_{z} \) be the usual standard deviations for \( y, x \) and \( z \), respectively, where

\[
\zeta_y = \sqrt{\frac{1}{N-1} \sum_{i=1}^{n} \sum_{j=1}^{n} (y_{ij} - \bar{y})^2}, \quad \zeta_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_{ij} - \bar{x})^2} \]

and

\[
\zeta_z = \sqrt{\frac{1}{N-1} \sum_{i=1}^{n} \sum_{j=1}^{n} (z_{ij} - \bar{z})^2}
\]

We also define the following error terms.

Let \( \zeta_{0} = \frac{\bar{y}^{(o)}}{\bar{y}} - 1 \), \( \zeta_{1} = \frac{\bar{x}^{(o)}}{\bar{x}} - 1 \), \( \zeta_{2} = \frac{\bar{z}^{(o)}}{\bar{z}} - 1 \), such that \( E(\zeta_{i}) = 0 \) for \( (i = 0, 1, 2) \). Also \( E(\zeta_{0}) = \Phi \rho_y \), \( E(\zeta_{1}) = \Phi \rho_x \), \( E(\zeta_{2}) = \Phi \rho_z \), \( E(\zeta_{0} \zeta_{1}) = \Phi C_y \sqrt{\rho_y \rho_x} \), \( E(\zeta_{0} \zeta_{2}) = \Phi C_y \sqrt{\rho_y \rho_z} \) and \( E(\zeta_{1} \zeta_{2}) = \Phi C_x \sqrt{\rho_x \rho_z} \), where \( C_y = \rho_{xy} C_x C_z, \quad C_x = \rho_{xz} C_x C_z, \quad C_z = \rho_{yz} C_x C_z \) and \( \Phi = \frac{N-1}{nN} \).

Now we discuss some of the existing estimators.

The usual sample mean estimator is given by

\[ \hat{y}_0 = \bar{y}^{(o)} \]  

The variance of \( \hat{y}_0 \) is given by

\[ Var(\hat{y}_0) = \Phi \Gamma \rho_y C_y \]  

Single auxiliary variable

(i) The traditional ratio estimator is given by

\[ \hat{y}_R = \bar{y}^{(o)} \left( \frac{\bar{x}}{\bar{y}^{(o)}} \right) \]  

The bias and MSE of \( \hat{y}_R \) to first order of approximation are respectively given by

\[ B(\hat{y}_R) \equiv \Phi \Gamma \left\{ \rho_y C_y^2 - C_x \sqrt{\rho_y \rho_x} \right\} \]  

and

\[ MSE(\hat{y}_R) \equiv \Phi \Gamma^2 \left\{ \rho_y C_y^2 + \rho_x C_x^2 - 2C_y \sqrt{\rho_y \rho_x} \right\} \]  

(ii) The usual exponential ratio type estimator is given by

\[ \hat{y}_{RE} = \bar{y}^{(o)} \exp \left( \frac{\bar{X} - \bar{Y}^{(o)}}{\bar{X} + \bar{Y}^{(o)}} \right) \]  

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

\[ B(\hat{y}_{RE}) \equiv \Phi \Gamma \left\{ \frac{2}{8} \rho_y C_y^2 - \frac{1}{2} C_y \sqrt{\rho_y \rho_x} \right\} \]  

and

\[ MSE(\hat{y}_{RE}) \equiv \Phi \Gamma^2 \left\{ \rho_y C_y^2 + \frac{1}{4} \rho_x C_x^2 - C_y \sqrt{\rho_y \rho_x} \right\} \]  

(iii) The usual difference estimator is given by

\[ \hat{y}_D = \bar{y}^{(o)} + d \left( \bar{X} - \bar{Y}^{(o)} \right) \]  

where \( d \) is the constant.

The minimum variance of \( \hat{y}_D \) at optimum value of \( d \)

\[ d_{opt} = \frac{\bar{y} \rho_{y} C_y \sqrt{\rho_y \rho_x}}{XC_y}, \]  

is given by

\[ Var(\hat{y}_{D})_{min} \equiv \Phi \Gamma^2 \rho_y C_y^2 \left( 1 - \rho_{xy}^2 \right) = MSE(\hat{y}_{D})_{min} \]  

(iv) Kocyigit and Cingi (2017) suggested a new class of
unbiased linear estimator which is given by

\[ \hat{Y}_{KC} = Y^{(sys)} \left\{ q_1 + q_2 \left( \frac{aX + b}{aX + b} \right)^{\alpha} + q_3 \left( \frac{aX + b}{aX + b} \right)^{\delta} \right\}, \]  

\[ ...(11) \]

where \((\alpha, \delta) \in R \) and \( q_i (i = 1, 2, 3) \) 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 kurtosis of the auxiliary variable \( x \).

The minimum \( MSE \) of \( \hat{Y}_{KC} \) at optimum value of \( A_y = (q_2\alpha + q_3\delta) \) i.e. \( A_{y(ops)} = -\frac{\rho_{ys}C_y\sqrt{\rho_s}}{gC_y\sqrt{\rho_s}} \), where \( g = \frac{aX}{aX + b} \), is given by

\[ MSE(\hat{Y}_{KC})_{\text{min}} = \Phi \hat{Y}^{2} \rho_{ys}^{2} C_{ys}^{2} (1 - \rho_{ys}^{2}), \]  

\[ ...(12) \]

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:

\[ \hat{Y}_{RDS} = \left\{ t_1 Y^{(sys)} + t_2 \left( X - X^{(sys)} \right) \right\} \exp \left\{ \frac{X - X^{(sys)}}{X + X^{(sys)}} \right\}, \]  

\[ ...(13) \]

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

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

\[ \hat{Y}_{RDS} = \left\{ t_1 Y^{(sys)} + t_2 \left( X - X^{(sys)} \right) \right\} \exp \left\{ \frac{X - X^{(sys)}}{X + X^{(sys)}} \right\}, \]  

\[ ...(13) \]

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

\[ B(\hat{Y}_{RDS}) \equiv (t_1 - 1) Y + t_1 \Phi Y \left\{ \frac{3}{8} \rho_{ys}^{2} C_{ys}^{2} - \frac{1}{2} C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} \right\} \]
\[ + \frac{1}{2} t_1 \Phi X \rho_{ys}^{2} C_{ys}^{2}, \]  

\[ ...(14) \]

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

\[ MSE(\hat{Y}_{RDS}) \equiv (t_1 - 1)^2 Y^2 + t_1^2 Y^2 A + t_1^2 X^2 B - t_1 Y^2 C \]
\[ - t_1 Y X B + 2t_1 t_2 D, \]  

where \( A = \Phi \left( \rho_{ys}^{2} C_{ys}^{2} + \rho_{ys}^{2} C_{ys}^{2} - 2 C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} \right), \]
\[ B = \Phi \rho_{ys}^{2} C_{ys}^{2}, \]
\[ C = \Phi \left( \frac{3}{4} \rho_{ys}^{2} C_{ys}^{2} - C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} \right), \]
\[ D = \Phi \left( \rho_{ys}^{2} C_{ys}^{2} - C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} \right). \]

The minimum MSE of \( \hat{Y}_{RDS} \) at optimum values of \( t_i (i = 1, 2) \) i.e. \( t_{i(ops)} = \frac{B(C - D + 2)}{2(AB - D^2 + B)} \)
\[ t_{2(ops)} = \frac{\bar{Y}(AB - CD + B - 2D)}{2\bar{X}(AB - D^2 + B)} \] is given by

\[ MSE(\hat{Y}_{RDS})_{\text{min}} = \Phi \left( 1 - \frac{1}{4} \left( \frac{AB^2 + BC^2 - 2BCD + B^2 + 4BC - 4BD + 4B}{AB - D^2 + B} \right) \right) \]  

\[ ...(15) \]

Two auxiliary variables

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

\[ \hat{Y}_{RR} = Y^{(sys)} \left( \frac{X}{X^{(sys)}} \right) \left( \frac{Z}{Z^{(sys)}} \right) \]  

\[ ...(16) \]

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

\[ B(\hat{Y}_{RR}) \equiv \Phi \left( \rho_{ys}^{2} C_{ys}^{2} + \rho_{ys}^{2} C_{ys}^{2} + C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} - C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} - C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} \right) \]  

\[ ...(17) \]

and

\[ MSE(\hat{Y}_{RR}) \equiv \Phi \left( \rho_{ys}^{2} C_{ys}^{2} + \rho_{ys}^{2} C_{ys}^{2} + \rho_{ys}^{2} C_{ys}^{2} + 2C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} - 2C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} - 2C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} \right) \]
\[ - 2C_{ys} \sqrt{\rho_{ys}^{2} \rho_s^{2}} \]  

\[ ...(18) \]

(ii) Tailor et al. (2013) suggested the ratio-product estimator which is given by
(iii) The bias and MSE of \( \hat{\theta}_{RP} \) to first order of approximation are given by

\[
B(\hat{\theta}_{RP}) \equiv \Phi \Phi \left\{ \rho_r^2 C_i^2 - C_n \sqrt{\rho_r^2 \rho_i^2} - C_n \sqrt{\rho_r^2 \rho_i^2} + C_n \sqrt{\rho_r^2 \rho_i^2} \right\}
\]

and

\[
MSE(\hat{\theta}_{RP}) \equiv \Phi \Phi \left\{ \rho_r^2 C_i^2 + \rho_r^2 C_i^2 + \rho_r^2 C_i^2 - 2 C_n \sqrt{\rho_r^2 \rho_i^2} - 2 C_n \sqrt{\rho_r^2 \rho_i^2} + 2 C_n \sqrt{\rho_r^2 \rho_i^2} \right\}
\]

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

\[
\hat{\theta}_{RRE} = \frac{X}{Z} \exp \left\{ \frac{X - \bar{X}}{X + \bar{X}} \right\} \exp \left\{ \frac{Z - \bar{Z}}{Z + \bar{Z}} \right\}
\]

The bias and MSE of \( \hat{\theta}_{RRE} \) to first order of approximation are given by

\[
B(\hat{\theta}_{RRE}) \equiv \Phi \Phi \left\{ \frac{3}{8} \rho_r^2 C_i^2 + \frac{1}{4} C_n \sqrt{\rho_r^2 \rho_i^2} - \frac{1}{2} C_n \sqrt{\rho_r^2 \rho_i^2} \right\}
\]

and

\[
MSE(\hat{\theta}_{RRE}) \equiv \Phi \Phi \left\{ \rho_r^2 C_i^2 + \rho_r^2 C_i^2 + \rho_r^2 C_i^2 - \rho_r^2 C_i^2 - C_n \sqrt{\rho_r^2 \rho_i^2} \right\}
\]

(v) Tailor and Mishra (2018) suggested the exponential ratio-product type estimator, which is given by

\[
\hat{\theta}_{RPE} = \frac{X}{Z} \exp \left\{ \frac{X - \bar{X}}{X + \bar{X}} \right\} \exp \left\{ \frac{Z - \bar{Z}}{Z + \bar{Z}} \right\}
\]

The bias and MSE of \( \hat{\theta}_{RPE} \) to first order of approximation are given by

\[
B(\hat{\theta}_{RPE}) \equiv \Phi \Phi \left\{ \frac{3}{8} \rho_r^2 C_i^2 - \rho_r^2 C_i^2 - \rho_r^2 C_i^2 - \rho_r^2 C_i^2 + \frac{1}{4} C_n \sqrt{\rho_r^2 \rho_i^2} \right\}
\]

and

\[
MSE(\hat{\theta}_{RPE}) \equiv \Phi \Phi \left\{ \rho_r^2 C_i^2 + \rho_r^2 C_i^2 + \rho_r^2 C_i^2 - \rho_r^2 C_i^2 + \frac{1}{4} C_n \sqrt{\rho_r^2 \rho_i^2} \right\}
\]

(vi) The traditional difference-difference type estimator is given by

\[
\hat{\theta}_{DD} = \frac{X}{Z} + d_1 (\bar{X} - \bar{X}) + d_2 (\bar{Z} - \bar{Z})
\]

where \( d_1 \) and \( d_2 \) are constants.

The minimum variance of \( \hat{\theta}_{DD} \), at optimum values of \( d_1 \) and \( d_2 \) i.e. \( d_1(\text{opt}) = \frac{YC}{X} \sqrt{\rho_r^2 (\rho_a - \rho_a \rho_a) \rho_a} \) and \( d_2(\text{opt}) = \frac{YC}{\rho_r^2 (1 - \rho_a^2)} \), is given by

\[
\text{Var}(\hat{\theta}_{DD})_{\text{min}} \equiv \Phi \Phi \left\{ \rho_r^2 C_i^2 (1 - R_{x,y}) \right\} = MSE(\hat{\theta}_{DD})_{\text{min}},
\]

where \( R_{x,y}^2 = \rho_x^2 + \rho_x^2 - 2 \rho_x \rho_y \rho_y \) is the multiple correlation coefficient of \( y \) on \( x \) and \( z \).

(vii) Khan and Singh (2015) and Khan (2016) suggested the following estimators respectively:

\[
\hat{\theta}_K = \frac{X}{Z} \exp \left\{ \frac{f(X - \bar{X})}{X + b_y \bar{X}} \right\} \exp \left\{ \frac{h(Z - \bar{Z})}{Z + (g-1) \bar{Z}} \right\}
\]

where \( -\infty < f < \infty, -\infty < h < \infty, g > 0 \) and \( \eta > 0 \).

Note that minimum MSEs of \( \hat{\theta}_M \) and \( \hat{\theta}_K \) are exactly equal to the minimum variance of \( \hat{\theta}_{DD} \) to first order of approximation but the estimator \( \hat{\theta}_{DD} \) is preferable because of unbiasedness.

(viii) Qureshi et al. (2018) introduced the following estimator for population mean
\[ \hat{Y}_{QKH} = \frac{\bar{Y}(\text{sys})}{\bar{X}(\text{sys})} \left( \frac{X}{X} \right)^{v_1} \exp \left( \alpha \frac{Z - \bar{z}(\text{sys})}{Z + (v_2 - 1)z^{(\text{sys})}} \right), \] ...(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

\[ B(\hat{Y}_{QKH}) \equiv \Phi \left[ \left( v_1 C_{my} \sqrt{\rho_j \rho_z} - \frac{\alpha}{v_2} C_{uy} \sqrt{\rho_y \rho_z} \right) \right. \]
\[ + \left. \left( \frac{v_1}{v_2} \right) C_{uy} \sqrt{\rho_y \rho_z} + \frac{1}{2} v_1 (v_1 + 1) \rho_j C_{yj} \right] \]
\[ + \left( \frac{\alpha}{v_2} \right) C_{uy} \sqrt{\rho_y \rho_z} \]
\[ = \frac{\alpha}{v_2} \frac{1}{2} v_1 (v_1 + 1) \rho_j C_{yj} \]
\[ \text{and} \]
\[ \alpha = \sqrt{\rho_y H_{sy} - H_{sy} \rho_y \left( 1 - \rho_y \right)^2} \]

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:

\[ B(\hat{Y}_{QKH(c)}) \equiv \Phi \left[ \left( v_1 C_{my} \sqrt{\rho_j \rho_z} - \frac{\alpha}{v_2} C_{uy} \sqrt{\rho_y \rho_z} \right) \right. \]
\[ + \left. \left( \frac{v_1}{v_2} \right) C_{uy} \sqrt{\rho_y \rho_z} + \frac{1}{2} v_1 (v_1 + 1) \rho_j C_{yj} \right] \]
\[ + \left( \frac{\alpha}{v_2} \right) C_{uy} \sqrt{\rho_y \rho_z} \]
\[ = \frac{\alpha}{v_2} \frac{1}{2} v_1 (v_1 + 1) \rho_j C_{yj} \]
\[ \text{and} \]
\[ \alpha = \sqrt{\rho_y H_{sy} - H_{sy} \rho_y \left( 1 - \rho_y \right)^2} \]

The minimum MSE of \( \hat{Y}_{QKH} \) at optimum values of \( v_1 \)
and \( \frac{\alpha}{v_2} \) i.e. \( (v_1)_{\text{opt}} = \frac{C_{uy} \sqrt{\rho_j \rho_z}}{C_{uy} \sqrt{\rho_j \rho_z} - \rho_j \rho_z} \)
and \( (\alpha/v_2)_{\text{opt}} = \frac{C_{uy} \sqrt{\rho_j \rho_z}}{C_{uy} \sqrt{\rho_j \rho_z} - \rho_j \rho_z} \) is given by

\[ \text{MSE}(\hat{Y}_{QKH})_{\text{min}} \equiv \Phi \left[ \rho_j C_{yj}^2 \left( 1 - R_{x:\text{sys}}^2 \right) \right] \] ...(32)

(ix) Mishra et al. (2018) suggested the following class of estimators:

\[ \hat{Y}_{MSS} = \hat{Y} + \hat{Y}_{RP} + \hat{Y}_{RPE}, \] ...(33)

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

Mishra et al. (2018) derived the bias and minimum MSE of \( \hat{Y}_{MSS} \) up to first order of approximation as:

\[ B(\hat{Y}_{MSS}) \equiv \Phi \left[ \left( \delta_1 + \frac{\delta_2}{2} \right) \left( \rho_j \sqrt{\rho_j \rho_z} - \rho_j \sqrt{\rho_j \rho_z} \right) \right. \]
\[ + \left. \left( \frac{\delta_1}{4} + \frac{\delta_2}{8} \right) \rho_j \sqrt{\rho_j \rho_z} \right] \] ...

\[ \text{and} \]
\[ \text{MSE}(\hat{Y}_{MSS})_{\text{min}} \equiv \Phi \left[ \rho_j C_{yj}^2 + \frac{3}{8} \rho_j \sqrt{\rho_j \rho_z} \rho_j \sqrt{\rho_j \rho_z} \right. \]
\[ - \frac{\delta_1}{8} \rho_j \sqrt{\rho_j \rho_z} \] ...(37)

The optimum value of \( \left( \delta_1 + \frac{\delta_2}{2} \right) \) is:

\[ \left( \delta_1 + \frac{\delta_2}{2} \right)_{\text{opt}} = \frac{\rho_j \sqrt{\rho_j \rho_z} - \rho_j \sqrt{\rho_j \rho_z}}{\rho_j \sqrt{\rho_j \rho_z} + \rho_j \sqrt{\rho_j \rho_z} + 2 \rho_j \sqrt{\rho_j \rho_z}} \] ...(38)

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:

\[ B(\hat{Y}_{MSS(c)}) \equiv \Phi \left[ \left( \delta_1 + \frac{\delta_2}{2} \right) \left( \rho_j \sqrt{\rho_j \rho_z} - \rho_j \sqrt{\rho_j \rho_z} \right) \right. \]
\[ - \frac{\delta_1}{8} \rho_j \sqrt{\rho_j \rho_z} \] ...(39)

The correct optimum value of \( \left( \delta_1 + \frac{\delta_2}{2} \right) \) is:

\[ \left( \delta_1 + \frac{\delta_2}{2} \right)_{\text{opt}} = \left( \delta_1 + \frac{\delta_2}{2} \right)_{\text{opt}} \] ...(40)

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

\[
B(\hat{Y}_{pr}) \equiv (A - 1) \hat{Y} + A_1 \hat{Y} \Phi \left\{ \frac{3}{8} \rho \rho' C_i^2 - \frac{1}{2} C_i \sqrt{\rho' \rho} \right\},
\]

\[
+ \frac{1}{2} A_2 \hat{Y} \rho \rho' C_i^2 + A_3 \hat{Y} \rho \rho' C_i^2 - A_1 \hat{Y} \Phi \left\{ \frac{3}{4} \rho \rho' C_i^2
\]

\[
= \frac{1}{2} \rho' \rho - A_1 \hat{Y} \rho \rho' C_i^2 - A_3 \hat{Y} \rho \rho' C_i^2 - A_1 \hat{Y} \Phi \left\{ \frac{3}{4} \rho \rho' C_i^2
\]

\[
+ 2 A_4 \hat{Y} \rho \rho' C_i^2 - C_i \sqrt{\rho' \rho} \right\} + 2 A_4 \hat{Y} \rho \rho' C_i^2 - C_i \sqrt{\rho' \rho} \right\} + 2 A_4 \hat{Y} \rho \rho' C_i^2
\]

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

\[
A_{(opt)} = \frac{1}{8} \Phi \rho \rho' C_i
\]

\[
A_{2(opt)} = \frac{1}{8} \Phi \rho \rho' C_i \left( 1 - R_{\rho, \rho} \right),
\]

\[
A_{3(opt)} = \frac{1}{8} \Phi \rho \rho' C_i \left[ \frac{1}{2} C_i \sqrt{R_{\rho, \rho} \left( 1 - R_{\rho, \rho} \right) - \frac{1}{16} \Phi \rho \rho' C_i^2 \right] + A_i
\]

\[
= \frac{1}{8} \Phi \rho \rho' C_i \left[ \frac{1}{2} C_i \sqrt{R_{\rho, \rho} \left( 1 - R_{\rho, \rho} \right) - \frac{1}{16} \Phi \rho \rho' C_i^2 \right] + A_i
\]

\[
= \frac{1}{8} \Phi \rho \rho' C_i \left( 1 - R_{\rho, \rho} \right).
\]

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

\[
\text{MSE}(\hat{Y}_{pr})_{min} = \frac{1}{8} \Phi \rho \rho' C_i \left( 1 - R_{\rho, \rho} \right).
\]

### Comparison of estimators

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

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

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

\[
\frac{1}{\alpha_i} \left[ \Phi^2 \rho \rho' C_i^2 \left( 1 - R_{\rho, \rho} \right) + \Phi \rho \rho' C_i^2 \right] > 0,
\]

where \( \alpha_i = 1 + \Phi \rho \rho' C_i \left( 1 - R_{\rho, \rho} \right) \) and

\[
\alpha_5 = \frac{1}{4} \Phi^2 \rho \rho' C_i^2 \left( 1 - R_{\rho, \rho} \right).
\]

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

\[
\frac{1}{\alpha_i} \left[ \Phi \rho \rho' C_i^2 \left( 1 - R_{\rho, \rho} \right) \right] > 0,
\]

where \( \alpha_3 = \Phi \rho \rho' C_i \left( 1 - R_{\rho, \rho} \right) \left( 1 - R_{\rho, \rho} \right) \) and

\[
\alpha_4 = \frac{1}{\alpha_i} \Phi \rho \rho' C_i \left( 1 - R_{\rho, \rho} \right).
\]
(iii) By equations (8) and (43), \( \text{MSE}(\hat{Y}_{nt})_{\text{min}} < \text{MSE}(\hat{Y}_{ne}) \) if
\[
\frac{F^2}{\alpha_t} \left[ \Phi \left( \frac{1}{2} \rho^2_{x} C_y - \rho^2_{y} C_{x}, \rho_{xy} \right)^2 + \alpha_2 + \alpha_4 + \alpha_6 \right] > 0,
\]
where \( \alpha_4 = \frac{\text{MSE}(\hat{Y}_{nt})}{F^2} \Phi \rho^2_{x} C_y \left( 1 - R^2_{y,x} \right). \)

(iv) By equations (10), (12) and (43), \( \text{MSE}(\hat{Y}_{nt})_{\text{min}} < \) \( \text{MSE}(\hat{Y}_{ke})_{\text{min}} \) if
\[
\frac{F^2}{\alpha_t} \left[ \alpha_2 + \alpha_6 \right] > 0,
\]
where \( \alpha_6 = \frac{\text{MSE}(\hat{Y}_{nt})}{F^2} \Phi \rho^2_{x} C_y \left( 1 - R^2_{y,x} \right). \)

(v) By equations (15) and (43), \( \text{MSE}(\hat{Y}_{nt})_{\text{min}} < \text{MSE}(\hat{Y}_{reg})_{\text{min}} \) if
\[
\frac{F^2}{\alpha_t} \left[ \left( 1 + \Phi \rho^2_{x} C_y / 8 \right) + \frac{1}{2} \Phi C_y \sqrt{\rho^2_{x} C_y} \right] \left( 1 - R^2_{y,x} \right) \frac{A_6}{4 A_t} > 0,
\]
where \( A_t = \left( AB^2 + BC^2 - 2 BCD + B^2 + 4BC - 4BD + 4B \right) \) and \( A_t = AB - D^2 + B. \)

(vi) By equations (18) and (43), \( \text{MSE}(\hat{Y}_{nt})_{\text{min}} < \text{MSE}(\hat{Y}_{re}) \) if
\[
\frac{F^2}{\alpha_t} \left[ \Phi \left( \sqrt{\rho^2_{x} C_y} - \sqrt{\rho^2_{y} C_{x}, \rho_{xy}} \right)^2 + \Phi \left( \sqrt{\rho^2_{x} C_y} - \sqrt{\rho^2_{x} C_{x}, \rho_{xy}} \right)^2 \right] > 0,
\]
where \( \alpha_4 = \frac{\text{MSE}(\hat{Y}_{nt})}{F^2} \Phi \rho^2_{x} C_y \left( 1 - R^2_{y,x} \right) \) and \( \alpha_6 = \Phi \rho^2_{x} C_y \left( R^2_{y,x} - \rho^2_{y,x} - \rho^2_{y,x} \right) + 2 \Phi C_y \sqrt{\rho^2_{x} C_y} \)

(vii) By equations (21) and (43), \( \text{MSE}(\hat{Y}_{nt})_{\text{min}} < \text{MSE}(\hat{Y}_{pe}) \) if
\[
\frac{F^2}{\alpha_t} \left[ \left( 1 + \Phi \rho^2_{x} C_y / 8 \right) + \frac{1}{2} \Phi C_y \sqrt{\rho^2_{x} C_y} \right] \left( 1 - R^2_{y,x} \right) \frac{A_6}{4 A_t} > 0,
\]
where \( \alpha_6 = \frac{\text{MSE}(\hat{Y}_{nt})}{F^2} \Phi \rho^2_{x} C_y \left( 1 - R^2_{y,x} \right). \)

(viii) By equations (24) and (43), \( \text{MSE}(\hat{Y}_{nt})_{\text{min}} < \text{MSE}(\hat{Y}_{re}) \) if
\[
\frac{F^2}{\alpha_t} \left[ \Phi \left( \frac{1}{2} \rho^2_{x} C_y - \sqrt{\rho^2_{y} C_{x}, \rho_{xy}} \right)^2 + \Phi \left( \frac{1}{2} \rho^2_{x} C_y - \sqrt{\rho^2_{y} C_{x}, \rho_{xy}} \right)^2 \right] > 0,
\]
where \( \alpha_6 = \frac{\text{MSE}(\hat{Y}_{nt})}{F^2} \Phi \rho^2_{x} C_y \left( 1 - R^2_{y,x} \right). \)

(ix) By equations (27) and (43), \( \text{MSE}(\hat{Y}_{nt})_{\text{min}} < \text{MSE}(\hat{Y}_{pe}) \) if
\[
\frac{F^2}{\alpha_t} \left[ \left( 1 + \left( \Phi \rho^2_{x} C_y \right) / 8 \right) + \frac{1}{2} \Phi C_y \sqrt{\rho^2_{x} C_y} \right] \left( 1 - R^2_{y,x} \right) \frac{A_6}{4 A_t} > 0,
\]
where \( \alpha_6 = \frac{\text{MSE}(\hat{Y}_{nt})}{F^2} \Phi \rho^2_{x} C_y \left( 1 - R^2_{y,x} \right). \)

(x) By equations (29), (32) and (43), \( \text{MSE}(\hat{Y}_{nt})_{\text{min}} < \) \( \text{MSE}(\hat{Y}_{ee})_{\text{min}} \) if
\[
\frac{F^2}{\alpha_t} \left[ \left( 1 + \left( \Phi \rho^2_{x} C_y \right) / 8 \right) + \frac{1}{2} \Phi C_y \sqrt{\rho^2_{x} C_y} \right] \left( 1 - R^2_{y,x} \right) \frac{A_6}{4 A_t} > 0,
\]
where \( \alpha_6 = \frac{\text{MSE}(\hat{Y}_{nt})}{F^2} \Phi \rho^2_{x} C_y \left( 1 - R^2_{y,x} \right). \)

(xi) By equations (38) and (43), \( \text{MSE}(\hat{Y}_{nt})_{\text{min}} < \text{MSE}(\hat{Y}_{ele}) \) if
\[
\frac{F^2}{\alpha_t} \left[ \left( 1 + \left( \Phi \rho^2_{x} C_y \right) / 8 \right) + \frac{1}{2} \Phi C_y \sqrt{\rho^2_{x} C_y} \right] \left( 1 - R^2_{y,x} \right) \frac{A_6}{4 A_t} > 0,
\]
where \( \alpha_6 = \frac{\text{MSE}(\hat{Y}_{nt})}{F^2} \Phi \rho^2_{x} C_y \left( 1 - R^2_{y,x} \right). \)

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:

\[ N = 40, \quad n = 10, \quad \bar{Y} = 5078.575, \]
\[ \bar{X} = 230.325, \quad \bar{Z} = 945.425, \quad C_y = 0.3295, \]
\[ C_x = 0.8465, \quad C_z = 0.6755, \quad \rho_{xy} = 0.80056, \]
\[ \rho_{xz} = 0.8349, \quad \rho_{zx} = 0.9816, \quad \rho_{yy} = 0.23, \]
\[ \rho_{xx} = 0.086, \quad \rho_{zz} = 0.16. \]

**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:

\[ N = 81, \quad n = 21, \quad \bar{Y} = 33.84, \]
\[ \bar{X} = 112.4568, \quad \bar{Z} = 117.4691, \quad C_y = 0.29719, \]
\[ C_x = 0.12555, \quad C_z = 0.48620, \quad \rho_{xy} = -0.6907, \]
\[ \rho_{xz} = -0.7944, \quad \rho_{zx} = 0.9068, \quad \rho_{yy} = 0.56, \]
\[ \rho_{xx} = 0.35, \quad \rho_{zz} = 0.52. \]

**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:

\[ N = 34, \quad n = 9, \quad \bar{Y} = 199.4412, \]
\[ \bar{X} = 856.411, \quad \bar{Z} = 208.8824, \quad C_y = 0.7531, \]
\[ C_x = 0.8560, \quad C_z = 0.7205, \quad \rho_{xy} = 0.98, \]
\[ \rho_{xz} = 0.4491, \quad \rho_{zx} = 0.11, \quad \rho_{yy} = 0.0091, \]
\[ \rho_{xx} = 0.043. \]

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

The hypothetical data statistics are:

\[ N = 15, \quad n = 3, \quad \bar{Y} = 80, \]
\[ \bar{X} = 44.47, \quad \bar{Z} = 48.40, \quad C_y = 0.56, \]

**Table 1:** Percent relative efficiency (PRE) of different estimators with respect to \( \hat{Y}_0 \)

| Estimator | Data 1 | Data 2 | Data 3 | Data 4 |
|-----------|--------|--------|--------|--------|
| \( \hat{Y}_0 \) | 100.000 | 100.000 | 100.000 | 100.000 |
| One auxiliary variable | | | | |
| \( \hat{Y}_R \) | 59.277 | 62.909 | 102.823 | 389.621 |
| \( \hat{Y}_{RE} \) | 256.384 | 79.017 | 124.694 | 177.435 |
| \( \hat{Y}_{DS}, \hat{Y}_{KC} \) | 278.470 | 191.230 | 124.730 | 3314.670 |
| \( \hat{Y}_{RDS} \) | 297.030 | 196.590 | 139.360 | 3414.630 |
| Two auxiliary variables | | | | |
| \( \hat{Y}_{RR} \) | 11.133 | 13.185 | 150.318 | 66.547 |
| \( \hat{Y}_{RP} \) | 93.280 | 167.560 | 38.630 | 1390.340 |
| \( \hat{Y}_{RE} \) | 68.918 | 29.478 | 304.159 | 81.262 |
| \( \hat{Y}_{RP} \) | 100.074 | 256.942 | 62.409 | 617.606 |
| \( \hat{Y}_{QKH} \) | 341.220 | 274.740 | 2527.400 | 6729.210 |
| \( \hat{Y}_{MSS(c)} \) | 100.970 | 274.450 | 128.100 | 5556.860 |
| \( \hat{Y}_{Prem} \) | 360.730 | 280.020 | 2614.410 | 6962.980 |
C_s = 0.28, \quad C_s = 0.43, \quad \rho_{sy} = 0.9848, \\
\rho_{sx} = -0.9760, \quad \rho_{sz} = -0.953, \quad \rho_{sz} = 0.6652, \\
\rho_{ss} = 0.707, \quad \rho_{is} = 0.5487.

The following expression is used to obtain the percent relative efficiency (PRE):

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

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

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 (\(\hat{y}_p\)) is the best among all the estimators considered here. The estimators \(\hat{y}_e, \hat{y}_{se}, \hat{y}_{rh}, \hat{y}_{sp}, \hat{y}_{sre}\) and \(\hat{y}_{sre}\) 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 (\(\hat{y}_d\)) and Koycyigit and Cingi (2017) estimator (\(\hat{y}_{dc}\)) are equal when using single auxiliary variable. The improvement in Riaz et al. (2017) estimator (\(\hat{y}_{rds}\)) is more than (\(\hat{y}_d\)) and (\(\hat{y}_{dc}\)) estimators. Also difference estimators (\(\hat{y}_{sre}, \hat{y}_{srhu}\)) are equally efficient but (\(\hat{y}_{sre}\)) 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 (\(\hat{y}_p\)) is more efficient than the usual sample mean estimator (\(\hat{y}_d\)), traditional ratio estimators (\(\hat{y}_e, \hat{y}_{se}\)), exponential-ratio estimators (\(\hat{y}_{sp}, \hat{y}_{sre}\)), ratio-product estimators (\(\hat{y}_{sre}, \hat{y}_{sre}\)), difference estimators (\(\hat{y}_{sre}, \hat{y}_{srhu}\)) and the other estimators considered here.

REFERENCES

Banarasi Kushwaha S.N.S. & Kushwaha K.S. (1993). A class of ratio, product and difference estimators in systematic sampling. Microelectronic Reliability 33(4): 455–457.

Cochran W.G. (1946). Relative accuracy of systematic and stratified random samples for a certain class of populations. The Annals of Mathematical Statistics 17: 164–177. 

DOI: https://doi.org/10.1214/aoms/1177730978

Gautschi W. (1957). Some remarks on systematic sampling. The Annals of Mathematical Statistics 17: 164–177. 

DOI: https://doi.org/10.1214/aoms/1177706966

Gujrati D.N. (2012). Basic Econometrics, 4th edition. McGraw Hill, New York, USA.

Khan M. (2016). A generalized class of exponential type estimators for population mean under systematic sampling using two auxiliary variables. Journal of Probability and Statistics 2016: Article ID 2374837. 

DOI: https://doi.org/10.1155/2016/2374837

Khan M. & Singh R. (2015). Estimation of population mean in chain ratio-type estimator under systematic sampling. Journal of Probability and Statistics 2015: Article ID 248374. 

DOI: https://doi.org/10.1155/2015/248374

Kocyigit E.G. & Cingi H. (2017). A new class of unbiased linear estimators in systematic sampling. Hacettepe Journal of Mathematics and Statistics 46(2): 315–323. 

DOI: https://doi.org/10.15672/HJMS.2017.413

Kushwaha K.S. & Singh H.P. (1989). Class of almost unbiased ratio and product estimators in systematic sampling. Journal of the Indian Society of Agriculture Statistics 41(2): 193–205.

Madow W.G. & Madow L.H. (1944). On the theory of systematic sampling. Annals of Mathematical Statistics 15: 1–24. 

DOI: https://doi.org/10.1214/aoms/1177731312

Mishra M., Singh B.P. & Singh R. (2018). Estimation of population mean using two auxiliary variables in systematic sampling. Journal of Scientific Research 62: 203–212.

Mukhopadhyay P. (1998). Theory and Methods of Survey Sampling. Prentice-Hall of India, New Delhi, India.

Murthy M. (1967). Sampling Theory and Methods. Statistical Publishing Society, Calcutta, India.

Pal S.K. & Singh H.P. (2017). Estimation of finite population mean using auxiliary information in systematic sampling. International Journal of System Assurance Engineering and Management 8(2): 1392–1398. 

DOI: https://doi.org/10.1007/s13198-017-0609-5

Qureshi M.N., Khalil S. & Hanif M. (2018). Generalized semi exponential type estimator under systematic sampling. Journal of Statistical Theory and Applications 17(2): 283–290. 

DOI: https://doi.org/10.2991/jsta.2018.17.2.8
Riaz S., Diana G. & Shabbir J. (2017). Modified classes of estimators in circular systematic sampling. *Hacettepe Journal of Mathematics and Statistics* 46(4): 743–765.

Singh H.P. & Jatwa N.K. (2012). A class of exponential type estimators in systematic sampling. *Economic Quality Control* 27(2): 195–208.

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

Singh H.P. & Singh R. (1998). Almost unbiased ratio and product type estimators in systematic sampling. *Questio* 22(3): 403–416.

Singh H.P., Tailor R. & Jatwa N.K. (2011). Modified ratio and product estimators for population mean in systematic sampling. *Journal of Modern Applied Statistical Methods* 10: 424–435.

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

Swain A.K.P.C. (1964). The use of systematic sampling ratio estimate. *Journal of the Indian Statistical Association* 2: 160–164.

Tailor R. & Mishra A. (2018). An exponential strategy for estimation of population mean in systematic sampling. *International Journal of Scientific Research in Mathematical and Statistical Sciences* 5(3): 73–78.

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

Tailor Z.T., Jatwa N. & Singh H.P. (2013). A ratio-cum-product estimator of finite population mean in systematic sampling. *Statistics in Transition* 14(3): 391–398.