Some Improved Estimators of Population Mean Under Systematic Sampling

Shagufta Mehnaz¹ and Shakeel Ahmed*²
¹Department of Statistics, university of Lahore, Sargodha Pakistan
²Department of Mathematics and Statistics, Institute of Business Management Karachi Pakistan
*Corresponding Author: shakeel.ahmed@iobm.edu.pk

Abstract
Auxiliary information is very important in constructing estimators for the population parameters for increasing the efficiency different sampling schemes. In this paper, we consider the problem of estimation of population mean using information on auxiliary variables in systematic sampling. We derive the expressions for the bias and mean squared error (MSE) of the suggested estimators up to the 1st degree of approximation. Proposed estimators are compared with usual mean estimator in systematic sampling scheme theoretically as well as empirically.

Keywords: Bias, Mean squared error, Efficiency, Auxiliary information, Systematic sampling.

Introduction
Utilizing information about the auxiliary variables that are highly correlated with the variable of interest is a common practice for efficiency improvement in estimation of population parameters. Many authors including Cochran (1977), Singh et al. (2004), Khan et al. (2015), Singh et al. (2011), Suberamani and Kumarpandiyan (2013), Sisodia and Dwivedi (1981), Yan and Tian (2010), Upadhyaya and Singh (1999) and Kadilar and Cingi (2004 and 2006) have worked on estimation of population mean using auxiliary information under simple random sampling.

Systematic sampling is preferred over simple random sampling because of two advantages i.e. (1) its simplicity and (2) it give the assurance that the population will be evenly sampled. There may be a chance in simple random sampling that allows of selecting clusters of units in simple random sampling. Systematic sampling overcomes this deficiency by systematically selecting unit from the population. Keeping these advantages in consideration

Let $Y$ be the study variable and $X$ and $Z$ be the auxiliary variables observed on a finite Population $U=\{U_1, U_2, \ldots U_N\}$ of size $N$ units numbered from $1$ to $N$. To take a systematic sample of size $n$, we divide the population into $k$ mutually exclusive groups or intervals, where $k= N/n$. Then we select $i^{th}$ unit from first interval randomly and select $i+k, i+2k, \ldots, i+(n-1)k$ unit from $2^{nd}, 3^{rd}, \ldots, k^{th}$ interval respectively. Let $(y_{ij}, x_{ij}, z_{ij})$ for $(i=1,2,\ldots,k; j=1,2,\ldots,n)$ be the value of $i^{th}$ unit from $j^{th}$ interval on variables $Y, X$ and $Z$ respectively. Some important notations are given by

\[
\begin{align*}
\bar{X} &= \frac{1}{n} \sum_{i=1}^{n} X_{ij}, \quad \bar{Y} = \frac{1}{n} \sum_{j=1}^{n} y_{ij}, \quad \bar{Z} = \frac{1}{n} \sum_{i=1}^{n} z_{ij}, \\
\bar{x}_y &= \frac{1}{n} \sum_{i=1}^{n} X_{ij}, \quad \bar{y}_y = \frac{1}{n} \sum_{j=1}^{n} y_{ij}, \\
\bar{y}_x &= \frac{1}{n} \sum_{j=1}^{n} Z_{ij}, \quad S_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{X})^2}, \quad S_y = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2}, \quad S_z = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (z_i - \bar{Z})^2}, \\
\rho_{xy} &= \frac{S_{xy}}{S_x S_y}, \quad \rho_{xz} = \frac{S_{xz}}{S_x S_z}, \quad \rho_{yz} = \frac{S_{yz}}{S_y S_z}, \quad \beta_{1(\cdot)} = \frac{\mu_x}{\mu_y}, \quad \beta_{2(\cdot)} = \frac{\mu_x}{\mu_y}^2, \quad \text{and} \quad R = \frac{Y}{X},
\end{align*}
\]

Further, we define $M_{d(X)}$ and $M_{d(Z)}$ as the population median of the auxiliary variables $X$ and $Z$ respectively.

We also define following error terms

\[
\begin{align*}
\varepsilon_o &= \frac{(\bar{Y} - \bar{Y})}{\bar{Y}}, \quad \varepsilon_1 = \frac{(\bar{x}_y - \bar{X})}{\bar{X}}, \quad \varepsilon_2 = \frac{(\bar{y}_x - \bar{Z})}{\bar{Z}},
\end{align*}
\]

such that $E(\varepsilon_i) = 0$ for $i=0, 1, 2$ and

\[
\begin{align*}
E(\varepsilon_0^2) &= \lambda p_x^2 C_y^2, \quad E(\varepsilon_1^2) = \lambda p_x^2 C_z^2, \quad E(\varepsilon_2^2) = \lambda p_z^2 C_y^2 \\
E(\varepsilon_0 \varepsilon_1) &= \lambda k C_x \sqrt{p_x p_y}, \quad E(\varepsilon_0 \varepsilon_2) = \lambda k C_x \sqrt{p_y p_z}, \quad E(\varepsilon_1 \varepsilon_2) = \lambda k C_x \sqrt{p_z p_x}.
\end{align*}
\]
where \( \rho_{**} = \frac{\rho_{yx}^*}{\rho_{zx}^*}, \rho_{x}^* = \frac{\rho_{yx}^*}{\rho_{zx}^*}, k = \frac{\rho_{yx}^* C_y}{C_z}, k^* = \frac{\rho_{yx}^* C_y}{C_z}, \rho_{x}^* = \frac{1}{\rho_{zx}^*}, \rho_{y}^* = \frac{1}{\rho_{zx}^*}, \rho_{z}^* = \frac{1}{\rho_{zx}^*} \) and \( \lambda = \frac{N-1}{nN} \).

We use \( \rho_{x}, \rho_{y} \) and \( \rho_{z} \) to denote the correlation among the pair units in a systematic sample for the variable \( x, y \) and \( z \) respectively.

**Extension of some well-known estimators in Systematic sampling**

The simple mean estimator under systematic random sampling is given by

\[
\hat{Y}_0 = \frac{1}{n} \sum_{j=1}^{n} y_{ij} = \bar{y}_i (i=1,2,3..k).
\]  

(2.1)

It is easy to prove that it is an unbiased estimator of population mean with variance

\[
V(\hat{Y}_0) = \lambda \overline{Y}^2 \rho_{y}^* C_y^2.
\]  

(2.2)

For efficiency improvement in estimation of mean in systematic sampling we extend Singh et al. (2004), Upadhyaya and Singh (1999), Kadilar and Cigni (2004), Yan and Tian (2010), Sisodia and Dwivedi (1981), Suberamani and Kumarpandiyan (2013) and Tailor et al. (2013) estimators in systematic sampling

\[
\hat{Y}_{g}(xy) = \bar{y}_{sy} \left[ \frac{aX + b}{a\bar{x}_{sy} + b} \right]^q \left[ \frac{c\bar{Z} + d}{c\bar{z}_{sy} + d} \right]^\rho,
\]  

(2.3)

\[
\hat{Y}_{1}(sy) = \bar{y}_{sy} \left[ \frac{X + C_x}{\bar{x}_{sy} + C_x} \right],
\]  

(2.4)

\[
\hat{Y}_{2}(sy) = \bar{y}_{sy} \left[ \frac{X + \beta_{1(sy)}}{\bar{x}_{sy} + \beta_{1(sy)}} \right],
\]  

(2.5)

\[
\hat{Y}_{3}(sy) = \bar{y}_{sy} \left[ \frac{X + \beta_{1(s)}^*}{\bar{x}_{sy} + \beta_{1(s)}^*} \right],
\]  

(2.6)

\[
\hat{Y}_{4}(xy) = \bar{y}_{xy} \left[ \frac{X + \rho_{yx}^*}{\bar{x}_{xy} + \rho_{yx}^*} \right],
\]  

(2.7)

\[
\hat{Y}_{5}(sy) = \bar{y}_{sy} \left[ \frac{XC_x + \beta_{2(s)}}{\bar{x}_{sy} C_x + \beta_{2(s)}} \right],
\]  

(2.8)

\[
\hat{Y}_{6}(xy) = \bar{y}_{xy} \left[ \frac{XC_x + \beta_{2(s)}}{\bar{x}_{xy} C_x + \beta_{2(s)}} \right].
\]  

(2.9)
\[ \hat{Y}_{6(y)} = \overline{y}_{xy} \left[ \frac{\overline{X} + M_{d(x)}}{\overline{x}_{xy} + M_{d(x)}} \right], \]  
\[ \hat{Y}_{7(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{\overline{x}_{xy}} \overline{X}, \]  
\[ \hat{Y}_{8(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{(\overline{x}_{xy} + C_x)} (\overline{X} + C_x), \]  
\[ \hat{Y}_{9(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{(\overline{x}_{xy} + \beta_{2(1)})} (\overline{X} + \beta_{2(1)}), \]  
\[ \hat{Y}_{10(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{(\overline{x}_{xy} C_x + \beta_{2(1)})} (\overline{X} C_x + \beta_{2(1)}), \]  
\[ \hat{Y}_{11(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{(\overline{x}_{xy} + \beta_{1(1)})} (\overline{X} + \beta_{1(1)}), \]  
\[ \hat{Y}_{12(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{(\overline{x}_{xy} \beta_{1(1)} + \beta_{2(1)})} (\overline{X} \beta_{1(1)} + \beta_{2(1)}), \]  
\[ \hat{Y}_{13(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{(\overline{x}_{xy} + \rho_{yx})} (\overline{X} + \rho_{yx}), \]  
\[ \hat{Y}_{14(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{(\overline{x}_{xy} C_x + \rho_{yx})} (\overline{X} C_x + \rho_{yx}), \]  
\[ \hat{Y}_{15(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{(\overline{x}_{xy} + \rho_{yx})} (\overline{X} + \rho_{yx}), \]

and

\[ \hat{Y}_{16(y)} = \frac{\overline{y}_{xy} + b_{yx} (\overline{X} - \overline{x}_{xy})}{(\overline{x}_{xy} + \beta_{2(1)})} (\overline{X} + \beta_{2(1)}). \]

To incorporate information about two auxiliary variables, Tailor et al. (2013) proposed a ratio-cum-product estimator in systematic sampling which is given by
Some Improved Estimators of Population Mean Under Systematic Sampling

\[
\hat{Y}_{17(xy)} = \bar{y}_y \left( \frac{\bar{X}}{X_{xy}} \right) \left( \frac{Z}{z_{xy}} \right)
\]  (2.19)

The bias of estimators given in Equations (2.17-2.3) are approximated up to 1st order and the final expressions are given by

\[
\text{Bias}(\hat{Y}_{1(xy)}) \cong \bar{Y} \rho \lambda^2 \left[ \theta^2 - k \theta \sqrt{\rho^*} \right],
\]  (2.20)

\[
\text{Bias}(\hat{Y}_{3(xy)}) \cong \bar{Y} \rho \lambda^2 \left[ \theta^2 - k \theta \sqrt{\rho^*} \right],
\]  (2.21)

\[
\text{Bias}(\hat{Y}_{4(xy)}) \cong \bar{Y} \rho \lambda^2 \left[ \theta^2 - k \theta \sqrt{\rho^*} \right],
\]  (2.22)

\[
\text{Bias}(\hat{Y}_{5(xy)}) \cong \bar{Y} \rho \lambda^2 \left[ \theta^2 - k \theta \sqrt{\rho^*} \right],
\]  (2.23)

\[
\text{Bias}(\hat{Y}_{6(xy)}) = \bar{Y} \rho \lambda^2 \left[ \theta^2 - k \theta \sqrt{\rho^*} \right],
\]  (2.24)

\[
\text{Bias}(\hat{Y}_{7(xy)}) \cong R \lambda \rho \lambda^2 \left[ 1 + \frac{\bar{X}}{\bar{Y}} b_{xy} - k \sqrt{\rho^*} \right],
\]  (2.25)

\[
\text{Bias}(\hat{Y}_{8(xy)}) \cong R \lambda \rho \lambda^2 \left[ 1 + \frac{\bar{X}}{\bar{Y}} b_{xy} - k \sqrt{\rho^*} \right],
\]  (2.26)

\[
\text{Bias}(\hat{Y}_{9(xy)}) \cong R \lambda \rho \lambda^2 \left[ 1 + \frac{\bar{X}}{\bar{Y}} b_{xy} - k \sqrt{\rho^*} \right],
\]  (2.27)

\[
\text{Bias}(\hat{Y}_{10(xy)}) = R \lambda \rho \lambda^2 \left[ 1 + \frac{\bar{X}}{\bar{Y}} b_{xy} - k \sqrt{\rho^*} \right],
\]  (2.28)

\[
\text{Bias}(\hat{Y}_{11(xy)}) \cong R \lambda \rho \lambda^2 \left[ 1 + \frac{\bar{X}}{\bar{Y}} b_{xy} - k \sqrt{\rho^*} \right],
\]  (2.29)

\[
\text{Bias}(\hat{Y}_{12(xy)}) \cong R \lambda \rho \lambda^2 \left[ 1 + \frac{\bar{X}}{\bar{Y}} b_{xy} - k \sqrt{\rho^*} \right],
\]  (2.30)

\[
\text{Bias}(\hat{Y}_{13(xy)}) \cong R \lambda \rho \lambda^2 \left[ 1 + \frac{\bar{X}}{\bar{Y}} b_{xy} - k \sqrt{\rho^*} \right],
\]  (2.31)

\[
\text{Bias}(\hat{Y}_{14(xy)}) \cong R \lambda \rho \lambda^2 \left[ 1 + \frac{\bar{X}}{\bar{Y}} b_{xy} - k \sqrt{\rho^*} \right],
\]  (2.32)
The Mean Squared Errors of different estimators are approximated up to terms containing square of errors

\[
MSE(\hat{Y}_{1(sy)}) \approx Y^2 \rho_y^* \lambda \left[ \rho_{yy}^* C_y^2 + \theta_y^2 C_x^2 - 2C_y^2 k\theta_1 \sqrt{\rho_{yy}^*} \right],
\]

\[
MSE(\hat{Y}_{2(sy)}) \approx \rho_x^* \lambda Y^2 \left[ \rho_{xx}^* C_y^2 + \theta_x^2 C_x^2 - 2C_y^2 k\theta_2 \sqrt{\rho_{xx}^*} \right],
\]

\[
MSE(\hat{Y}_{3(sy)}) \approx \bar{Y}^2 \rho_x^* \lambda \left[ \rho_{xx}^* C_y^2 + \theta_x^2 C_x^2 - 2C_y^2 k\theta_3 \sqrt{\rho_{xx}^*} \right],
\]

\[
MSE(\hat{Y}_{4(sy)}) \approx \bar{Y}^2 \rho_x^* \lambda \left[ \rho_{xx}^* C_y^2 + \theta_x^2 C_x^2 - 2C_y^2 k\theta_4 \sqrt{\rho_{xx}^*} \right],
\]

\[
MSE(\hat{Y}_{5(sy)}) \approx \bar{Y}^2 \rho_x^* \lambda \left[ \rho_{xx}^* C_y^2 + \theta_x^2 C_x^2 - 2C_y^2 k\theta_5 \sqrt{\rho_{xx}^*} \right],
\]

\[
MSE(\hat{Y}_{6(sy)}) \approx \bar{Y}^2 \rho_x^* \lambda \left[ \rho_{xx}^* C_y^2 + \theta_x^2 C_x^2 - 2C_y^2 k\theta_6 \sqrt{\rho_{xx}^*} \right],
\]

\[
MSE(\hat{Y}_{7(sy)}) \approx \rho_x^* \lambda \left[ \rho_{xx}^* S_y^2 + \rho_{xx}^* S_x^2 + \frac{S_{yy}^2}{S_y^2} + 2RS_{yx} - 2R^2 S_x^2 k\sqrt{\rho_{xx}^*} - 2RS_{yy} \right],
\]

\[
MSE(\hat{Y}_{8(sy)}) \approx \rho_x^* \lambda \left[ \rho_{xx}^* S_y^2 + \rho_{xx}^* S_x^2 + \frac{S_{yy}^2}{S_y^2} + 2RS_{yx} - 2R^2 S_x^2 k\sqrt{\rho_{xx}^*} - 2RS_{yy} \right],
\]

\[
MSE(\hat{Y}_{9(sy)}) \approx \rho_x^* \lambda \left[ \rho_{xx}^* S_y^2 + \rho_{xx}^* S_x^2 + \frac{S_{yy}^2}{S_y^2} + 2RS_{yx} - 2R^2 S_x^2 k\sqrt{\rho_{xx}^*} - 2RS_{yy} \right],
\]

\[
MSE(\hat{Y}_{10(sy)}) \approx \rho_x^* \lambda \left[ \rho_{xx}^* S_y^2 + \rho_{xx}^* S_x^2 + \frac{S_{yy}^2}{S_y^2} + 2RS_{yx} - 2R^2 S_x^2 k\sqrt{\rho_{xx}^*} - 2RS_{yy} \right],
\]
Some Improved Estimators of Population Mean Under Systematic Sampling

\[ \text{MSE}(\hat{\mu}_{1(s)}) = \rho \lambda \left[ \rho^2 \frac{S_y^2}{S_x^2} + R_{11}^2 \frac{S_x^2}{S_x^2} + 2R_{11}S_{xy} - 2R_{11}S_{xy} \sqrt{\rho^2} \right] \],

\[ \text{MSE}(\hat{\mu}_{12(s)}) = \rho^2 \lambda \left[ \rho^2 \frac{S_y^2}{S_x^2} + R_{12}^2 \frac{S_x^2}{S_x^2} + 2R_{12}S_{xy} - 2R_{12}S_{xy} \sqrt{\rho^2} \right] \],

\[ \text{MSE}(\hat{\mu}_{13(s)}) = \rho \lambda \left[ \rho^2 \frac{S_y^2}{S_x^2} + R_{13}^2 \frac{S_x^2}{S_x^2} + 2R_{13}S_{xy} - 2R_{13}S_{xy} \sqrt{\rho^2} \right] \],

\[ \text{MSE}(\hat{\mu}_{14(s)}) = \rho^2 \lambda \left[ \rho^2 \frac{S_y^2}{S_x^2} + R_{14}^2 \frac{S_x^2}{S_x^2} + 2R_{14}S_{xy} - 2R_{14}S_{xy} \sqrt{\rho^2} \right] \],

\[ \text{MSE}(\hat{\mu}_{15(s)}) = \rho \lambda \left[ \rho^2 \frac{S_y^2}{S_x^2} + R_{15}^2 \frac{S_x^2}{S_x^2} + 2R_{15}S_{xy} - 2R_{15}S_{xy} \sqrt{\rho^2} \right] \],

\[ \text{MSE}(\hat{\mu}_{16(s)}) = \rho \lambda \left[ \rho^2 \frac{S_y^2}{S_x^2} + R_{16}^2 \frac{S_x^2}{S_x^2} + 2R_{16}S_{xy} - 2R_{16}S_{xy} \sqrt{\rho^2} \right] \],

\[ \text{MSE}(\hat{\mu}_{17(s)}) = \lambda \bar{Y}^2 \left[ \rho^2 \frac{S_y^2}{S_x^2} + \frac{\rho^2}{4} \left( 1 - 2K \sqrt{\rho^2} \right) + \rho^2 \frac{C^2}{C^2} \right] \],

where \( \theta_1 = \frac{\bar{X}}{\bar{X} + C_x} \), \( \theta_2 = \frac{\bar{X}}{\bar{X} + \beta_{1(x)}} \), \( \theta_3 = \frac{\bar{X}}{\bar{X} + \rho_{xy}} \), \( \theta_4 = \frac{\bar{X}}{\bar{X} + \beta_{2(x)}} \), \( \theta_5 = \frac{\bar{X}C_x}{\bar{X} + \beta_{2(x)}} \),

\( \theta_6 = \frac{\bar{X}}{\bar{X} + M_{d(x)}} \), \( R_7 = \frac{\bar{Y}}{\bar{X}} \), \( R_8 = \frac{\bar{Y}}{\bar{X} + C_x} \), \( R_9 = \frac{\bar{Y}}{\bar{X} + \beta_{2(x)}} \), \( R_{10} = \frac{\bar{Y}C_x}{\bar{X} + \beta_{2(x)}} \), \( R_{11} = \frac{\bar{Y}}{\bar{X} + \beta_{1(x)}} \).

\( R_{12} = \frac{\bar{X} \beta_{1(x)} + \beta_{2(x)}}{\bar{X} + \beta_{1(x)}} \), \( R_{13} = \frac{\bar{Y}}{\bar{X} + \rho_{xy}} \), \( R_{14} = \frac{\bar{Y}C_x}{\bar{X} + \rho_{xy}} \), \( R_{15} = \frac{\bar{Y} \rho_{xy}}{\bar{X} \rho_{xy} + C_x} \),

and \( R_{16} = \frac{\bar{Y} \rho_{xy}}{\bar{X} \rho_{xy} + \beta_{2(x)}} \).

Proposed Estimators in Systematic Random Sampling

On the basis of the fact that median of the auxiliary variable is easy to obtain from the population being studied, we suggest some more efficient estimators in systematic sampling by following Suberamani and Kumarpandiyan (2013) as follows

\[ \hat{Y}_{18P(sy)} = \bar{y}_{sy} \left[ \frac{\bar{X} \beta_{1(x)} + M_{d(x)}}{\bar{X} \beta_{1(x)} + M_{d(x)}} \right] \],

\[ \hat{Y}_{19P(sy)} = \bar{y}_{sy} \left[ \frac{\bar{X} \beta_{1(x)} + M_{d(x)}}{\bar{X} \beta_{1(x)} + M_{d(x)}} \right] \].
\[
\hat{Y}_{20p(xy)} = \bar{y}_{xy} \left[ \frac{\bar{X}M_{d(x)} + \beta_{2x}}{x_{xy}M_{d(x)} + \beta_{2x}} \right],
\]  
(3.3)

\[
\hat{Y}_{21p(xy)} = \bar{y}_{xy} + b_{yx} \left( \frac{\bar{X} - x_{xy}}{x_{xy} + M_{d(x)}} \right) (\bar{X} + M_{d(x)}).
\]  
(3.4)

Taking inspiration from Tailor et al. (2013), we proposed the following ratio and product type estimator, for using median along with mean two auxiliary variable.

\[
\hat{Y}_{22p(xy)} = \bar{y}_{xy} \left[ \frac{\bar{X} + M_{d(x)}}{x_{xy} + M_{d(x)}} \right] \left[ \frac{\bar{Z} + M_{d(z)}}{z_{xy} + M_{d(z)}} \right],
\]  
(3.5)

where \( M_{d(X)} \) and \( M_{d(Z)} \) are the population median of auxiliary variable \( X \) and \( Z \) respectively.

The bias of estimators given in Equations (3.1-5) are given by

\[
Bias(\hat{Y}_{18(xy)}) = \bar{y} \rho^*_x C_x^2 \lambda \left[ \theta_{18} - k\theta_{18} \sqrt{\rho^{**}} \right],
\]  
(3.6)

\[
Bias(\hat{Y}_{19(xy)}) = \bar{y} \rho^*_x C_x^2 \lambda \left[ \theta_{19}^2 - k\theta_{19} \sqrt{\rho^{**}} \right],
\]  
(3.7)

\[
Bias(\hat{Y}_{20(xy)}) = \bar{y} \rho^*_x C_x^2 \lambda \left[ \theta_{20}^2 - k\theta_{20} \sqrt{\rho^{**}} \right],
\]  
(3.8)

\[
Bias(\hat{Y}_{21(xy)}) = R_{21} \lambda \rho^*_x S_x^2 \left[ \frac{1}{\bar{X}} + \frac{\bar{X}}{\bar{Y}} b_{yx} - k\rho^{**} \right]
\]  
(3.9)

and

\[
Bias(\hat{Y}_{22(xy)}) = \bar{y} \lambda \left[ b_{22}^2 \rho^*_x C_y^2 + \rho^*_x C_x^2 (a_{22}^2 - a_{22} k\bar{X}) + \rho^*_z C_z^2 (a_{22} b_{22} k^{**} \sqrt{\rho^{**}} - b_{22} k^{**} \sqrt{\rho^{**}}) \right].
\]  
(3.10)

The Mean Squared Error of proposed estimators given in Equations (3.1-5) are expressed as follow

\[
MSE(\hat{Y}_{18(xy)}) = \bar{y}^2 \rho^*_x \lambda \left[ \rho^{**} C_y^2 + \theta_{18}^2 C_x^2 - 2C_x^2 k\theta_{18} \sqrt{\rho^{**}} \right],
\]  
(3.11)

\[
MSE(\hat{Y}_{19(xy)}) = \bar{y}^2 \rho^*_x \lambda \left[ \rho^{**} C_y^2 + \theta_{19}^2 C_x^2 - 2C_x^2 k\theta_{19} \sqrt{\rho^{**}} \right],
\]  
(3.12)

\[
MSE(\hat{Y}_{20(xy)}) = \bar{y}^2 \rho^*_x \lambda \left[ \rho^{**} C_y^2 + \theta_{20}^2 C_x^2 - 2C_x^2 k\theta_{20} \sqrt{\rho^{**}} \right],
\]  
(3.13)
Some Improved Estimators of Population Mean Under Systematic Sampling

\[
MSE(\hat{Y}_{21(\alpha)}) = \rho_x^* \lambda \left[ \rho_x^* S_x^2 + R_{21}^2 S_x^2 + \frac{S_{yx}^2}{S_x^2} + 2R_{21} S_{yx} - 2R_{21}^2 S_x^2 \right] (3.14)
\]

and

\[
MSE(\hat{Y}_{22(\alpha)}) = \bar{y}^2 \theta \left[ \rho_y^* C_y^2 + \beta_x C_x (a^2 - 2a_{22}k \sqrt{\rho_x^*}) + \beta_z C_z (b^2 + 2a_{22}b_{22}k^* \sqrt{\rho_z^*}) \right], (3.15)
\]

where

\[
\rho_{\alpha} = \frac{\bar{X}_\alpha \beta_{\alpha} + M_{d(\alpha)}}{\bar{X}_\alpha \beta_{\alpha}}, \quad \theta_{19} = \frac{\bar{X} M_{d(x)}}{\bar{X} M_{d(x)} + \beta_{l(x)}}, \quad \theta_{20} = \frac{\bar{X} M_{d(x)}}{\bar{X} M_{d(x)} + \beta_{2(x)}},
\]

\[
R_{21} = \left[ \frac{\bar{Y}}{\bar{X} + M_{d(x)}} \right], \quad a_{22} = \frac{\bar{X}}{\bar{X} + M_{d(x)}} \quad \text{and} \quad b_{22} = \frac{\bar{Z}}{\bar{Z} + M_{d(x)}}.
\]

Numerical Study

A real life data set is used, to exhibit the performance of the proposed estimators.

Data set: (Source: 1)

We use District Wise Maize Data of Punjab the province of Pakistan for year 2014, where

Y = Production of Maize for year 2007-08 (in tons),

X = Area of Maize for year 2007-08 (in hectares sq),

Z = Area of Maize for year 2008-09 (in hectares sq).

The values of population parameters and the constants obtained from the above data are given below:

\[ N = 25, \lambda = 4.8, n = 5, k = 1.218752, K^* = 1.24581, K^{**} = 1.019864, \beta_x(\xi) = 0.0018, \]

\[ \beta_{2}(x) = 4.2872, \bar{X} = 19.192, C_x = 1.390563, S_x^2 = 712.2324, \rho_x^* = 0.205311, \bar{Y} = 97.016, \]

\[ C_y = 1.795732, S_y^2 = 30350.79, \rho_y^* = 0.258816, \bar{Z} = 18.86, C_z = 1.360231, S_z^2 = 658.1258, \]

\[ \rho_z^* = 0.204505, \rho_y z = 0.94367, \rho_x^* = 1.260606, S_{yx}^2 = 4387.935, M_{d(x)} = 4.7, \rho_{zx} = 0.94367, \]

\[ \rho_{x z} = 1.003939, S_{xz} = 683.0143, M_{d(z)} = 5.1, \rho_{x z} = 0.99762, \rho_y z = 1.265572, S_{yz} = 4217.565. \]

The biases, mean squared errors and percentage relative efficiencies (PREs) of different estimator in systematic sampling are given in Table (4.1). Using the following formula;

\[ PRE(\hat{Y}_{a(\alpha)}, \hat{Y}_{0(\alpha)}) = \frac{Var(\hat{Y}_{0(\alpha)})}{MSE(\hat{Y}_{a(\alpha)})} \times 100 \quad \text{for} \; a = 0, 1, 2, \ldots, 22. \]

where \[ PRE(\hat{Y}_{a(\alpha)}, \hat{Y}_{0(\alpha)}) \] denotes the percentage relative efficiency of \( a^{th} \) estimator.
Table 1: The Absolute Biases, MSE’S and PRE’S of Different Estimators under Systematic Random Sampling

| Estimator | $|Bias(\hat{Y}_a)|$ | $MSE(\hat{Y}_a)$ | $PRE(\hat{Y}_a, \hat{Y}_o)$ |
|-----------|-----------------|-----------------|-----------------|
| $\hat{Y}_{0(x)}$ | 0.000 | 37705.26 | 100 |
| $\hat{Y}_{1(x)}$ | 75.1484 | 7529.906 | 500.7401 |
| $\hat{Y}_{2(x)}$ | 68.1141 | 6556.526 | 575.0798 |
| $\hat{Y}_{3(x)}$ | 83.2613 | 8591.218 | 438.8815 |
| $\hat{Y}_{4(x)}$ | 73.1701 | 7214.039 | 522.6651 |
| $\hat{Y}_{5(x)}$ | 68.111 | 6556.178 | 575.1104 |
| $\hat{Y}_{6(x)}$ | 83.9199 | 9848.845 | 382.8394 |
| $\hat{Y}_{7(x)}$ | 157.2131 | 17091.51 | 220.6081 |
| $\hat{Y}_{8(x)}$ | 172.3844 | 22246.38 | 169.4894 |
| $\hat{Y}_{9(x)}$ | 157.1984 | 17098.87 | 220.5132 |
| $\hat{Y}_{10(x)}$ | 157.2025 | 17096.8 | 220.5398 |
| $\hat{Y}_{11(x)}$ | 128.5061 | 26246.71 | 143.6571 |
| $\hat{Y}_{12(x)}$ | 157.2097 | 17093.23 | 220.5859 |
| $\hat{Y}_{13(x)}$ | 149.8445 | 20697.92 | 182.1693 |
| $\hat{Y}_{14(x)}$ | 151.8434 | 19733.13 | 191.0759 |
| $\hat{Y}_{15(x)}$ | 146.004 | 22523.33 | 167.4054 |
| $\hat{Y}_{16(x)}$ | 157.1975 | 17099.3 | 220.5076 |
| $\hat{Y}_{17(x)}$ | 1.21033 | 36629.17 | 102.9378 |
| $\hat{Y}_{18(x)}$ | 73.873 | 7321.673 | 514.9815 |
| $\hat{Y}_{19(x)}$ | 73.0208 | 7191.781 | 524.2826 |
| $\hat{Y}_{20(x)}$ | 68.10311 | 6555.287 | 575.1886 |
| $\hat{Y}_{21(x)}$ | 2.106374 | 138.7300 | 210.4772 |
| $\hat{Y}_{22(x)}$ | 77.83677 | 4834.807 | 779.8711 |
Some Improved Estimators of Population Mean Under Systematic Sampling

From Table 1, it can be seen that proposed estimators ($\hat{Y}_{18(p)}$, $\hat{Y}_{19(p)}$, $\hat{Y}_{20(p)}$, and $\hat{Y}_{21(p)}$) have greater relative efficiency than systematic version of Suberamani and Kumarpandiyan (2013) estimator ($\hat{Y}_{6(p)}$), Yan and Tian (2010) estimator ($\hat{Y}_{3(p)}$), Singh et al (2004) estimator ($\hat{Y}_{2(p)}$), Kadilar and Cigni (2004) estimator ($\hat{Y}_{8(p)}$) respectively. We can also observed that proposed estimators ($\hat{Y}_{18(p)}$, $\hat{Y}_{19(p)}$, $\hat{Y}_{20(p)}$, and $\hat{Y}_{21(p)}$) have smaller biases than $\hat{Y}_{6(p)}$, $\hat{Y}_{3(p)}$, $\hat{Y}_{2(p)}$, and $\hat{Y}_{8(p)}$ respectively. Hence, it is clear that the proposed estimators in systematic sampling are more efficient and have smaller bias than corresponding parent estimators (estimator from which they are derived). Furthermore, the proposed ratio–cum- product type estimator ($\hat{Y}_{22(p)}$) that utilize a linear combination of mean and median of two auxiliary variables in systematic sampling are more efficient than systematic version of Suberamani and Kumarpandiyan (2013) estimator ( $\hat{Y}_{6(p)}$) and Tailor et al. (2013) estimator ($\hat{Y}_{17(p)}$) that utilize median of only one auxiliary variable. It can also be inferred from the table that the proposed estimators ($\hat{Y}_{20(p)}$) and ($\hat{Y}_{22(p)}$) have a highest percentage relative efficiency, i.e. 575.1886 and 779.8711, as compare to other estimators discussed in this chapter. At same time, we can also observed that proposed estimator ($\hat{Y}_{22(p)}$) that utilize a linear combination of mean and median of two auxiliary variables in systematic sampling are less bias than Suberamani and Kumarpandiyan (2013) estimator ( $\hat{Y}_{6(p)}$) that utilize median of only one auxiliary variable. So it can be inferred from table the proposed estimator $\hat{Y}_{21(p)}$ have a less bias 2.106374 as compare to other estimator discuss in this chapter.

Conclusion

The paper covered utilization of different known parameters of auxiliary variable (i.e. mean, median, coefficient of variation, coefficient of kurtosis and skewness etc) in constructing estimators for population mean under systematic sampling. We also proposed some new ratio and product type estimator, for using median along with mean of two auxiliary variables in systematic random sampling to improve efficiency. These proposed estimators perform better than existing estimator in the term of percent relative efficiency. So we recommend to use these estimators for future research when these information are easily obtainable.

Acknowledgement

Authors are thankful to the editorial team for their support and guidance.

References


