Improved Estimators of Population Coefficient of Variation under Simple Random Sampling

Rajesh Singh\textsuperscript{a} and Anamika Kumari\textsuperscript{a*}

\textsuperscript{a} Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi, India.

Authors’ contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2022/v19i430474

Abstract

In this article, we suggest some novel estimators of population Coefficient of Variation (CV) of the study variable using the known information on an auxiliary variable like population mean and population variance. Up to the first order of approximation, formulas for the bias and Mean squared Errors (MSE) of the proposed estimators are obtained. The efficiencies of proposed and competing estimators are evaluated by comparing their MSEs. A real and two simulated data sets are used to verify the efficiency conditions. The results showed that the proposed estimators were more efficient than the other existing estimators considered in the study.

Keywords: Study variable; auxiliary variable; bias; mean square error; and coefficient of variation.

1 Introduction

In theory of Sampling Survey, use of auxiliary information improves the efficiency of the estimator. Use of auxiliary information can be done at various stages, it helps in improving the precision of the estimator. Cochran (1940) was the first to introduce a ratio estimator of Population Mean using auxiliary information. Shabbir and Gupta [1], Singh et al. (2007), Koyuncu and Kadilar [2] and Chaudhary et al. [3] have considered the problem of estimating population mean taking into consideration information on auxiliary variable.
The Coefficient of Variation (CV) is a well-known measure of dispersion, which is defined as the ratio of the standard deviation to the mean of the characteristic under study. It is used to compare variability in populations or samples with different units of measurement. For example in the investment, the CV allows us to determine how much risk one is assuming in comparison to the amount of return one can expect from ones investment. The lower the ratio of standard deviation to mean return, the better risk-return trade-off. Whenever the population is very large, the complete enumeration is very time consuming and costly then the population CV is estimated through sample CV by using auxiliary information as it improves precision. McKay [4] was the first to estimate population CV. Archana & Rao [5] gave some new estimator of CV for the enhancement of the estimation of CV. Shabbir and Gupta [6] used two auxiliary variables to improve the estimation of population CV in simple and stratified random sampling under a two-phase sampling technique. Singh et al. [7] proposed various improved and more enhanced estimators based on the arithmetic mean, geometric mean, and harmonic mean of these estimators. Singh and Mishra [8] proposed estimating the population CV using a single auxiliary variable. Audu et al. [9] proposed difference cum ratio type estimators for estimating population CV under SRS and demonstrated that their estimators are more efficient than the existing estimators. More on the estimation of the population coefficient of variation for, we refer to Abu-Shaweish et al. [10], Banik and Kibria [11], Ahmed et al. [12] and Gulhar et al. [13] among others.

To estimate any parameter under study we want to have efficient estimators. In search of efficient estimators, we proposed new estimators of population CV under SRS using known auxiliary parameters. These new estimators are expected to give a precise and efficient estimate of the population CV than the existing estimators considered in this paper.

Let us consider a finite population \( U = (U_1, U_2, \ldots, U_N) \) of size ‘N’ consisting of distinct and identifiable units. Let \( Y \) and \( X \) denotes the study and auxiliary variables and let \( Y_i \) and \( X_i \) be their values corresponding to \( i^{th} \) unit in the population \( (i = 1, 2, \ldots, N) \). For the population observations, we define:

\[
\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i \quad \text{and} \quad \bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i
\]

as the population means for the study and Auxiliary Variables.

\[
S_{Y}^2 = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})^2 \quad \text{and} \quad S_{X}^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2
\]

as the population mean squares for the study and auxiliary variables.

\[
S_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})(X_i - \bar{X})\]

as the population covariance.

Let us consider that a sample of size ‘n’ has been drawn from this population of size ‘N’ units using SRSWOR. For this sample let \( y_i \) and \( x_i \) denote the value of the \( i^{th} \) sample unit corresponding to study variable \( Y \) and auxiliary variable \( X \). For the sample observations we have:

\[
\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \quad \text{and} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i
\]

as the sample means for the study and Auxiliary Variables.

\[
s_{y}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2 \quad \text{and} \quad s_{x}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2
\]

as the sample mean squares for the study and auxiliary variables.
\[s_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})\] as the sample covariance.

Now let us define
\[\varepsilon_0 = \frac{\bar{y}}{\bar{Y}} - 1, \varepsilon_1 = \frac{\bar{x}}{\bar{X}} - 1, \varepsilon_2 = \frac{s_y^2}{s^2} - 1 \text{ and } \varepsilon_3 = \frac{s_x^2}{S^2} - 1\]

Such that
\[
E(\varepsilon_0) = E(\varepsilon_1) = E(\varepsilon_2) = E(\varepsilon_3) = 0 \\
E(\varepsilon_0^2) = \gamma C_y^2, E(\varepsilon_1^2) = \gamma C_x^2, E(\varepsilon_2^2) = \gamma(\lambda_{40} - 1), E(\varepsilon_3^2) = \gamma(\lambda_{04} - 1) \\
E(\varepsilon_0\varepsilon_1) = \gamma C_{yx}, E(\varepsilon_0\varepsilon_2) = \gamma C_y\lambda_{30}, E(\varepsilon_0\varepsilon_3) = \gamma C_y\lambda_{12}, E(\varepsilon_1\varepsilon_2) = \gamma C_x\lambda_{21} \\
E(\varepsilon_1\varepsilon_3) = \gamma C_x\lambda_{03}, E(\varepsilon_2\varepsilon_3) = \gamma(\lambda_{22} - 1)
\]

Where
\[
\gamma = \left(\frac{1}{n} - \frac{1}{N}\right), C_y = \frac{s_y}{\bar{Y}} \text{ and } C_x = \frac{s_x}{\bar{X}} \text{ are the population coefficient of variation for the study variable Y and auxiliary variable X. Also } \rho_{yx} \text{ denotes the correlation coefficient between X and Y.}
\]

\[\lambda_{rs} = \frac{\mu_{rs}}{\mu_{r20}^{1/2} \mu_{o2}^{1/2}}, \mu_{rs} = \frac{1}{N-1} \sum_{i=1}^{n} (y_i - \bar{Y}^r)(X_i - \bar{X}^s)\]

### 2 Existing Estimators

- The usual unbiased estimator to estimate the population coefficient of variation is given by:

\[t_0 = \hat{C}_y = \frac{s_y}{\bar{Y}} \tag{2.1}\]

The mean square error (MSE) expression of the estimator \(t_0\) is given by:

\[MSE(t_0) = \gamma C_y^2 \left(C_y^2 + 0.25(\lambda_{04} - 1) - C_y\lambda_{30}\right) \tag{2.2}\]

- Archana & Rao [14] proposed ratio type estimator for the population coefficient of variation is given by:

\[t_{AR} = \hat{C}_y \left(\frac{X}{\bar{X}}\right) \tag{2.3}\]

The mean square error (MSE) expression of the estimator \(t_{AR}\) is given by:

\[MSE(t_{AR}) = \gamma C_y^2 \left(C_y^2 + 0.25(\lambda_{04} - 1) - C_y\lambda_{30} - C_x\lambda_{21} + 0.25(\lambda_{40} - 1)\right) \tag{2.4}\]

- Singh et al. [7] proposed ratio-type, exponential ratio-type and difference-type estimators for coefficient of variation of the study variable Y using mean of auxiliary variable and are given below with their MSEs as

\[t_1 = \hat{C}_y \left(\frac{X}{\bar{X}}\right)^a \tag{2.5}\]

\[t_2 = \hat{C}_x \exp\left\{\beta \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right)\right\} \tag{2.6}\]
where

\[ \alpha = \frac{\lambda_{21} - 2\rho_{yx}C_y}{2C_x}, \quad \beta = \frac{\lambda_{21} - 2\rho_{yx}C_y}{C_x}, \quad d_1 = \frac{C_y\lambda_{21} - 2\rho_{yx}C_y^2}{4XC_x} \]

- Singh et al. [7] proposed arithmetic, geometric and harmonic mean estimators (AM, GM, HM) based on \( t_0 \) and \( t_1 \) estimators for estimating coefficient of variation of the study variable \( Y \) and are given below with their MSEs as

\[
t_{S}^{AM} = \frac{\hat{C}_y}{2} \left[ 1 + \left( \frac{\bar{X}}{\bar{X}} \right)^{\alpha} \right] \tag{2.11}
\]

\[
t_{S}^{GM} = \frac{\hat{C}_y}{2} \left( \frac{\bar{X}}{\bar{X}} \right) \tag{2.12}
\]

\[
t_{S}^{HM} = 2\hat{C}_y \left[ 1 + \left( \frac{\bar{X}}{\bar{X}} \right)^{\alpha} \right]^{-1} \tag{2.13}
\]

\[
MSE(t_{S}^{j}) = \gamma C_y^2 \left[ C_y^2 + \frac{\lambda_{40} - 1}{4} + \frac{\alpha^2C_x^2}{4} - C_y\lambda_{30} + \alpha\rho_{yx}C_yC_x - \frac{\alpha}{2} C_x\lambda_{21} \right] \tag{2.14}
\]

where \( \alpha = \frac{\lambda_{21} - 2\rho_{yx}C_y}{2C_x}, \quad j = AM, GM, HM \)

- Singh et al. [7] proposed arithmetic, geometric and harmonic mean estimators (AM, GM, HM) based on \( t_0 \) and \( t_2 \) estimators for estimating coefficient of variation of the study variable \( Y \) and are given below with their MSEs as

\[
t_{S}^{AM} = \frac{\hat{C}_y}{2} \left[ 1 + \exp \left\{ \beta \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \right\} \right] \tag{2.15}
\]

\[
t_{S}^{GM} = \frac{\hat{C}_y}{2} \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \tag{2.16}
\]

\[
t_{S}^{HM} = 2\hat{C}_y \left[ 1 + \exp \left\{ \beta \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \right\} \right]^{-1} \tag{2.17}
\]

\[
MSE(t_{S}^{j}) = \gamma C_y^2 \left[ C_y^2 + \frac{\lambda_{40} - 1}{4} + \frac{\beta^2C_x^2}{16} - C_y\lambda_{30} + \frac{\beta}{2} \rho_{yx}C_yC_x - \frac{\beta}{4} C_x\lambda_{21} \right] \tag{2.18}
\]

where \( \beta = \frac{2(\lambda_{31} - 2\rho_{yx}C_y)}{C_x} \)
Singh et al. [7] proposed arithmetic, geometric and harmonic mean estimators (AM, GM, HM) based on $t_1$ and $t_2$ estimators for estimating coefficient of variation of the study variable $Y$ and are given below with their MSEs as

$$t_{a1}^{AM} = \frac{\hat{C}_y}{2} \left[ \frac{\bar{X}}{\bar{Y}} \right]^a + \exp \left( \beta \left( \frac{\bar{X} - \bar{Y}}{\bar{X} + \bar{Y}} \right) \right)$$

(2.19)

$$t_{a1}^{GM} = \hat{C}_y \left( \frac{\bar{X}}{\bar{Y}} \right)^{\frac{a}{2}} \exp \left( \beta \left( \frac{\bar{X} - \bar{Y}}{\bar{X} + \bar{Y}} \right) \right)$$

(2.20)

$$t_{a1}^{HM} = 2\hat{C}_y \left( \frac{\bar{X}}{\bar{Y}} \right)^{\frac{a}{2}} \exp \left( -\beta \left( \frac{\bar{X} - \bar{Y}}{\bar{X} + \bar{Y}} \right) \right)^{-1}$$

(2.21)

$$\text{MSE} (t_{a1}) = \gamma C_y^2 \left[ C_x^2 + \frac{\lambda_{40} - 1}{4} \left( \alpha + \frac{\beta}{2} \right) C_x^2 - C_y \lambda_{30} + \left( \alpha + \frac{\beta}{2} \right) \rho_{yx} C_y C_x - \frac{1}{4} \left( \alpha + \frac{\beta}{2} \right) C_x \lambda_{21} \right]$$

(2.22)

where, $\beta = 2 \left( \frac{\lambda_{21} - 2 \rho_{yx} C_y}{C_x} - a \right)$

Audu et al. [9] suggested the following two difference-cum-ratio type estimators of $C_y$ utilizing the known $\bar{X}$ as,

$$t_{a1} = \left[ \frac{\hat{C}_y}{2} \left( \bar{X} + \frac{\bar{Y}}{\bar{X}} \right) + w_1 (\bar{X} - \bar{Y}) + w_2 \hat{C}_y \right] \left( \frac{\bar{X}}{\bar{Y}} \right)$$

(2.23)

$$\text{MSE} (t_{a1}) = C_y^2 (a + w_1^2 b + w_2^2 c + 2w_1 d - 2w_2 e - 2w_1 w_2 f)$$

(2.24)

Where

$$a = \gamma \left( C_x^2 + C_y^2 + 2 \rho_{yx} C_x C_y - C_x \lambda_{21} - C_y \lambda_{30} + \frac{(\lambda_{40} - 1)}{4} \right), \quad b = \gamma \delta^2 C_x^2, \quad \delta = \frac{\bar{X}}{C_y}$$

$$c = 1 + \gamma \left( 3C_x^2 + 3C_y^2 + 4 \rho_{yx} C_x C_y - 2C_x \lambda_{21} - 2C_y \lambda_{30} \right), \quad d = \gamma \delta \left( C_x^2 + \rho_{yx} C_x C_y - \frac{C_x \lambda_{21}}{2} \right)$$

$$e = \gamma \left( \frac{3C_x \lambda_{21}}{2} - 3 \rho_{yx} C_x C_y - \frac{5C_x^2}{2} - 2C_y^2 + \frac{3C_y \lambda_{30}}{2} + \frac{(\lambda_{40} - 1)}{8} \right)$$

$$f = \gamma \delta \left( \frac{C_x \lambda_{21}}{2} - \rho_{yx} C_x C_y - 2C_x^2 \right)$$

Adichwala et al. [15] suggested the following estimator for estimating $C_y$ using the known $S_x^2$ as,

$$t_\gamma = \delta_1 \left[ \frac{(1 - \eta)S_x^2 + \eta S_y^2}{\eta S_x^2 + (1 - \eta)S_y^2} \right] \hat{C}_y + (1 - \delta_1) \left[ \eta S_y^2 \left( \frac{1}{(1 - \eta)S_x^2 + \eta S_y^2} \right) \right]$$

(2.25)

Where $\delta_1$ and $\eta$ are the characterizing constants to be determined such that the MSEs of the estimators $t_\gamma$ is least.

The minimum MSEs of the estimator $t_\gamma$ for the optimum values of these constants is,

$$\text{MSE} (t_\gamma) = \text{MSE} (t_0) - \frac{1}{4} \gamma \left[ \frac{(\lambda_{22} - 1) - 2C_y \lambda_{12}}{(\lambda_{04} - 1)} \right] C_y^4$$

(2.26)
Singh et al. [7] proposed ratio type, exponential ratio-type and difference-type estimators for estimating coefficient of variation of the study variable Y using variances of the auxiliary variables and are given below:

\[ t_8 = \hat{C}_y \left( \frac{s^2_y}{s^2_x} \right)^{\alpha} \]  \hspace{1cm} (2.27)

\[ t_9 = \hat{C}_y \exp \left\{ \beta \left( \frac{s^2_x - s^2_y}{s^2_x + s^2_y} \right) \right\} \]  \hspace{1cm} (2.28)

\[ t_{10} = \hat{C}_y + d_x (S_x^2 - s_x^2) \]  \hspace{1cm} (2.29)

\[ MSE(t_8) = \gamma C_y^2 \left[ C_y^2 + \frac{\lambda_{41} - 1}{4} + \alpha^2 (\lambda_{04} - 1) - C_y \lambda_{30} - \alpha (\lambda_{22} - 1) + 2 \alpha C_y \lambda_{12} \right] \]  \hspace{1cm} (2.30)

\[ MSE(t_9) = \gamma C_y^2 \left[ C_y^2 + \frac{\lambda_{41} - 1}{4} + \beta^2 (\lambda_{04} - 1) \right] - C_y \lambda_{30} + \beta C_y \lambda_{12} - \frac{\beta}{2} (\lambda_{22} - 1) \]  \hspace{1cm} (2.31)

\[ MSE(t_{10}) = \gamma \left[ C_y^2 (C_y^2 - C_y \lambda_{30} + \frac{\lambda_{40} - 1}{4}) + d_x^2 S_x^2 (\lambda_{04} - 1) + 2 C_y^2 d_x S_x^2 \lambda_{12} - d_x S_x^2 \lambda_{12} \right] \]  \hspace{1cm} (2.32)

where

\[ \alpha = \frac{\lambda_{22} - 1 - 2C_y \lambda_{21}}{2(\lambda_{04} - 1)}, \beta = \frac{\lambda_{22} - 1 - 2C_y \lambda_{21}}{\lambda_{04} - 1}, d_x = \frac{C_y (\lambda_{22} - 1) - 2C_y^2 \lambda_{12}}{2 S_x^2 (\lambda_{04} - 1)} \]

Singh et al. [7] proposed arithmetic, geometric and harmonic mean estimators (AM, GM, HM) based on \(t_9\) and \(t_8\) estimators for estimating coefficient of variation of the study variable Y and are given below with their MSEs as

\[ t_{11}^{AM} = \hat{C}_y \left[ 1 + \left( \frac{s^2_y}{s^2_x} \right)^{\alpha/2} \right] \]  \hspace{1cm} (2.33)

\[ t_{11}^{GM} = \hat{C}_y \left( \frac{s^2_y}{s^2_x} \right)^{\alpha/2} \]  \hspace{1cm} (2.34)

\[ t_{11}^{HM} = 2 \hat{C}_y \left[ 1 + \left( \frac{s^2_y}{s^2_x} \right)^{\alpha/2} \right]^{-1} \]  \hspace{1cm} (2.35)

\[ MSE(t_{11}^j) = \gamma C_y^2 \left[ C_y^2 + \frac{\lambda_{40} - 1}{4} + \alpha^2 (\lambda_{04} - 1) \right] - C_y \lambda_{30} + \alpha C_y \lambda_{12} - \frac{\alpha}{2} (\lambda_{22} - 1) \]  \hspace{1cm} (2.36)

where \( \alpha = \frac{\lambda_{22} - 1 - 2C_y \lambda_{21}}{\lambda_{04} - 1}, \ j = AM, GM, HM \)

Rajyaguru et al. [16] proposed arithmetic, geometric and harmonic mean estimators (AM, GM, HM) based on \(t_9\) and \(t_8\) estimators for estimating coefficient of variation of the study variable Y and are given below with their MSEs as

\[ t_{12}^{AM} = \frac{\hat{C}_y}{2} \left[ 1 + \exp \left\{ \beta \left( \frac{s^2_x - s^2_y}{s^2_x + s^2_y} \right) \right\} \right] \]  \hspace{1cm} (2.37)
Singh et al. [7] proposed arithmetic, geometric and harmonic mean estimators (AM, GM, HM) based on $t_9$ and $t_{10}$ estimators for estimating coefficient of variation of the study variable $Y$ and are given below with their MSEs as

\begin{align*}
\hat{t}_{12}^{AM} &= \hat{C}_y \exp \left( \frac{\beta}{2} \left( \frac{S_y^2 - s_x^2}{S_y^2 + s_x^2} \right) \right) \\
\hat{t}_{12}^{GM} &= 2\hat{C}_y \left[ 1 + \exp \left( -\beta \left( \frac{S_y^2 - s_x^2}{S_y^2 + s_x^2} \right) \right) \right]^{-1} \\
MSE(t_{12}) &= \gamma C_y \left[ C_y + \frac{\lambda_{40} - 1}{4} + \frac{\beta^2(\lambda_{40} - 1)}{16} - \frac{C_y \lambda_{13} + \frac{\beta}{2} C_y \lambda_{12} - \frac{\beta}{4} (\lambda_{22} - 1)}{C_y \lambda_{13}} \right]
\end{align*}

where $\beta = \frac{2(\lambda_{22} - 1 - 4C_y \lambda_{21})}{(\lambda_{40} - 1)}$, $j = AM, GM, HM$

- Audu et al. [9] suggested the following two difference-cum-ratio type estimators of $C_y$ utilizing the known $S_x^2$ as,

\begin{align*}
t_{a2} &= \frac{\hat{C}_y}{2} \left( \frac{S_x^2}{S_y^2} + \frac{S_y^2}{S_x^2} \right) + w_3 (S_x^2 - s_x^2) + w_4 \hat{C}_y \left( \frac{S_x^2}{S_y^2} \right) \\
MSE(t_{a2}) &= C_y^2 (a_1 + w_3^2 b_1 + w_4^2 c_1 + 2w_3 d_1 - 2w_4 e_1 - 2w_3 w_4 f_1)
\end{align*}

Where

\begin{align*}
a_1 &= \gamma \left( (\lambda_{40} - 1) + 2C_y \lambda_{12} - C_y \lambda_{30} - (\lambda_{22} - 1) + \frac{(\lambda_{40} - 1)}{4} \right), b_1 = \gamma \delta_2^2 C_y^2, \\
c_1 &= 1 + \gamma (3(\lambda_{40} - 1) + 3C_y \lambda_{12} - 2(\lambda_{22} - 1) - 2C_y \lambda_{30}) \\
d_1 &= \gamma \delta_1 \left( (\lambda_{40} - 1) + \frac{C_y \lambda_{12} - (\lambda_{22} - 1)}{2} \right) \\
e_1 &= \gamma \left( \frac{3(\lambda_{22} - 1)}{2} - 3C_y \lambda_{12} - \frac{5(\lambda_{40} - 1)}{2} - 2C_y^2 + \frac{3C_y \lambda_{40} - (\lambda_{40} - 1)}{8} \right) \\
f_1 &= \gamma \delta_1 \left( \frac{\lambda_{22} - 1}{2} - C_y \lambda_{12} - (\lambda_{40} - 1) \right)
\end{align*}

- Yunusa et al. [17] suggested the following log type ratio estimator of $C_y$ using the known $S_x^2$ as,
The MSE of the estimator \( t_{14} \), up to the first order of approximation is,

\[
\text{MSE}(t_{14}) = \gamma C_y^2 \left[ C_y^2 + \frac{\lambda_{40} - 1}{4} + \frac{\lambda_{40} - 1}{(\ln(S_x^2))^2} \right] - C_y \lambda_{12}
\]

(2.48)

3 Proposed Estimators

Having studied the estimators in section 2, we proposed four estimators for coefficient of variation based on information on a single auxiliary variable.

\[
t_{p1} = \left[ \frac{C_y}{2} \left( \frac{\bar{x}}{\bar{x} + \bar{x}} \right) + k_1(\bar{x} - \bar{x}) + k_2 C_y \right] \left( 2 - \left( \frac{\bar{x}}{\bar{x} + \bar{x}} \right) \exp \left( \frac{\bar{x} - \bar{x}}{\bar{x} + \bar{x}} \right) \right)
\]

(3.1)

\[
t_{p2} = \left[ \frac{C_y}{2} \left( \frac{S_x^2}{S_x^2 + S_x^2} \right) + k_3 S_x^2 + k_4 C_y \right] \left( 2 - \frac{s_x^2}{S_x^2} \right)
\]

(3.2)

\[
t_{p3} = \left[ \frac{C_y}{2} \left( \frac{S_x^2}{S_x^2 + S_x^2} \right) + k_5(\bar{x}^2 - s_x^2) \right] \left( 2 - \frac{s_x^2}{S_x^2} \right)
\]

(3.3)

\[
t_{p4} = \left[ \frac{C_y}{2} \left( \frac{S_x^2 - S_x^2}{S_x^2 + S_x^2} \right) + k_3 S_x^2 + k_4 C_y \right] \left( 2 - \frac{s_x^2}{S_x^2} \right)
\]

(3.4)

Expressing the estimators \( t_j, j = p1, p2, p3, p4 \) in terms of \( \epsilon_i, i = 0, 1, 2, 3 \) and simplifying respectively, we have

\[
t_{p1} = \left[ \frac{S_y(1 + \epsilon_2)\bar{x}}{2Y(1 + \epsilon_0)} \left( \frac{\bar{x}(1 + \epsilon_1)}{\bar{x}(1 + \epsilon_1)} + \frac{\bar{x}(1 + \epsilon_1)}{\bar{x}} \right) \right] \left[ 2 - \frac{\bar{x}(1 + \epsilon_1)\bar{x}}{\bar{x}(1 + \epsilon_1)} \exp \left( \frac{\bar{x}(1 + \epsilon_1) - \bar{x}}{\bar{x}(1 + \epsilon_1) + \bar{x}} \right) \right]
\]

(3.5)

\[
t_{p2} = \left[ \frac{S_y(1 + \epsilon_2)\bar{x}}{2Y(1 + \epsilon_0)} \left( \frac{\bar{x}(1 + \epsilon_1) - \bar{x}}{\bar{x}(1 + \epsilon_1) + \bar{x}} \right) \right] \left[ 2 - \frac{\bar{x}(1 + \epsilon_1)\bar{x}}{\bar{x}(1 + \epsilon_1) + \bar{x}} \exp \left( \frac{\bar{x}(1 + \epsilon_1) - \bar{x}}{\bar{x}(1 + \epsilon_1) + \bar{x}} \right) \right]
\]

(3.6)

\[
t_{p3} = \left[ \frac{S_y(1 + \epsilon_2)\bar{x}}{2Y(1 + \epsilon_0)} \left( \frac{S_x^2}{S_x^2(1 + \epsilon_1)} + \frac{s_x^2}{S_x^2} \right) \right] \left[ 2 - \frac{s_x^2(1 + \epsilon_1)}{S_x^2} \right]
\]

(3.7)
Subtracting from all above four equations and taking expectations on both sides and putting values of different expectations, we get the biases of $t_{p1}, t_{p2}, t_{p3}$ and $t_{p4}$ up to the approximation of order one respectively as,

$$
Bias(t_{p1}) = \left[ y \left( C_y^2 - \frac{C_y A_{30}}{2} - \frac{(\lambda_{40} - 1)}{8} + \frac{C_y^2}{2} + \frac{3C_y A_{21}}{4} - \frac{3C_y A_{21}}{2} \right) - k_1 \frac{\bar{X}^2}{C_y} \right]
$$

$$
+ k_2 \left( 1 + \gamma \left( C_y^2 - \frac{C_y A_{30}}{2} - \frac{(\lambda_{40} - 1)}{8} + \frac{C_y^2}{2} + \frac{3C_y A_{21}}{4} - \frac{3C_y A_{21}}{2} \right) \right)
$$

$$
Bias(t_{p2}) = \left[ y \left( C_y^2 - \frac{C_y A_{30}}{2} - \frac{(\lambda_{40} - 1)}{8} + \frac{C_y^2}{2} + \frac{3C_y A_{21}}{4} - \frac{3C_y A_{21}}{2} \right) - k_3 \frac{\bar{X}^2}{C_y} \right]
$$

$$
+ k_4 \left( 1 + \gamma \left( C_y^2 - \frac{C_y A_{30}}{2} - \frac{(\lambda_{40} - 1)}{8} + \frac{C_y^2}{2} + \frac{3C_y A_{21}}{4} - \frac{3C_y A_{21}}{2} \right) \right)
$$

$$
Bias(t_{p3}) = \left[ y \left( C_y^2 - \frac{C_y A_{30}}{2} - \frac{(\lambda_{40} - 1)}{8} + \frac{C_y A_{12}}{2} - \frac{(\lambda_{22} - 1)}{2} \right) - k_5 \frac{S_y^2}{C_y} \right]
$$

$$
+ k_6 \left( 1 + \gamma \left( C_y^2 - \frac{C_y A_{30}}{2} - \frac{(\lambda_{40} - 1)}{8} + \frac{C_y A_{12}}{2} - \frac{(\lambda_{22} - 1)}{2} \right) \right)
$$

$$
Bias(t_{p4}) = \left[ y \left( C_y^2 - \frac{C_y A_{30}}{2} - \frac{(\lambda_{40} - 1)}{8} + \frac{C_y A_{12}}{2} - \frac{(\lambda_{22} - 1)}{2} \right) - k_7 \frac{S_y^2}{C_y} \right]
$$

$$
+ k_8 \left( 1 + \gamma \left( C_y^2 - \frac{C_y A_{30}}{2} - \frac{(\lambda_{40} - 1)}{8} + \frac{C_y A_{12}}{2} - \frac{(\lambda_{22} - 1)}{2} \right) \right)
$$
Subtracting $C_y$ from (3.9), (3.10), (3.11), & (3.12), squaring and taking expectation, we get the MSEs of the suggested estimators as

\[
MSE (\hat{p}_{1}) = C_v^2 (A_1 + k^2 B_1 + k^2 C_1 + 2 k_1 D_1 - 2 k_2 E_1 - 2 k_4 F_1) \tag{3.17}
\]

\[
MSE (\hat{p}_{2}) = C_v^2 (A_2 + k^2 B_2 + k^2 C_2 + 2 k_2 D_2 - 2 k_4 E_2 - 2 k_4 F_2) \tag{3.18}
\]

\[
MSE (\hat{p}_{3}) = C_v^2 (A_3 + k^2 B_3 + k^2 C_3 + 2 k_3 D_3 - 2 k_6 E_3 - 2 k_6 F_3) \tag{3.19}
\]

\[
MSE (\hat{p}_{4}) = C_v^2 (A_4 + k^2 B_4 + k^2 C_4 + 2 k_7 D_4 - 2 k_8 E_4 - 2 k_8 F_4) \tag{3.20}
\]

Where

\[
A_1 = \gamma \left( C_v^2 + \frac{9}{4} C_v^2 + \frac{(\lambda_{40} - 1)}{4} - C_y A_{30} + 3 C_{xy} - \frac{3 C_y A_{21}}{2} \right), \quad B_1 = \gamma g_2 C_v^2, \quad g = \frac{\bar{x}}{C_y}
\]

\[
C_1 = 1 + \gamma \left( 3 C_v^2 + 3 C_v^2 + 6 C_{xy} - 2 C_y A_{30} - 3 C_y A_{21} \right), \quad D_1 = \gamma g_1 \left( \frac{3}{2} C_v^2 + C_{xy} - \frac{C_y A_{21}}{2} \right)
\]

\[
E_1 = \gamma \left( \frac{9 C_y A_{21}}{4} - \frac{9 C_{xy}}{2} - \frac{19 C_v^2}{8} - 2 C_v^2 - \frac{3 C_y A_{30}}{3} - \frac{(\lambda_{40} - 1)}{8} \right), \quad F_1 = \gamma g_1 \left( \frac{C_y A_{21}}{2} - C_{xy} - 3 C_v^2 \right)
\]

\[
A_2 = \gamma \left( C_v^2 + \frac{9}{4} C_v^2 + \frac{(\lambda_{40} - 1)}{4} - C_y A_{30} + 3 C_{xy} - \frac{3 C_y A_{21}}{2} \right), \quad B_2 = \gamma g_2 C_v^2, \quad g = \frac{\bar{x}}{C_y}
\]

\[
C_2 = 1 + \gamma \left( 3 C_v^2 + 3 C_v^2 + 6 C_{xy} - 2 C_y A_{30} - 3 C_y A_{21} \right), \quad D_2 = \gamma g_1 \left( \frac{3}{2} C_v^2 + C_{xy} - \frac{C_y A_{21}}{2} \right)
\]

\[
E_2 = \gamma \left( \frac{9 C_y A_{21}}{4} - \frac{9 C_{xy}}{2} - 2 C_v^2 - 2 C_v^2 + \frac{3 C_y A_{30}}{3} - \frac{(\lambda_{40} - 1)}{8} \right), \quad F_2 = \gamma g_1 \left( \frac{C_y A_{21}}{2} - C_{xy} - 3 C_v^2 \right)
\]

\[
A_3 = \gamma \left( C_v^2 + (\lambda_{40} - 1) + \frac{(\lambda_{40} - 1)}{4} - C_y A_{30} + 2 C_y A_{12} - (\lambda_{22} - 1) \right), \quad B_3 = \gamma g_1^2 (\lambda_{40} - 1),
\]

\[
C_3 = 1 + \gamma \left( 3 C_v^2 + (\lambda_{40} - 1) + 4 C_y A_{12} - 2 C_y A_{30} - \frac{3 (\lambda_{22} - 1)}{2} \right),
\]

\[
D_3 = \gamma g_1 \left( (\lambda_{40} - 1) + C_y A_{12} - \frac{(\lambda_{22} - 1)}{2} \right), \quad g_1 = \frac{S_y^2}{C_y}
\]

\[
E_3 = \gamma \left( \frac{3 (\lambda_{22} - 1)}{2} - \frac{3 (\lambda_{40} - 1)}{2} - 3 C_y A_{12} - 2 C_v^2 - \frac{3 C_y A_{30}}{2} - \frac{(\lambda_{40} - 1)}{8} \right),
\]

\[
F_3 = \gamma g_1 \left( \frac{(\lambda_{22} - 1)}{2} - C_y A_{12} - 2(\lambda_{40} - 1) \right)
\]

\[
A_4 = \gamma \left( C_v^2 + (\lambda_{40} - 1) + \frac{(\lambda_{40} - 1)}{4} - C_y A_{30} + 2 C_y A_{12} - (\lambda_{22} - 1) \right), \quad B_4 = \gamma g_1^2 (\lambda_{40} - 1),
\]

\[
C_4 = 1 + \gamma \left( 3 C_v^2 + (\lambda_{40} - 1) + 4 C_y A_{12} - 2 C_y A_{30} - \frac{3 (\lambda_{22} - 1)}{2} \right),
\]

\[
D_4 = \gamma g_1 \left( (\lambda_{40} - 1) + C_y A_{12} - \frac{(\lambda_{22} - 1)}{2} \right), \quad g_1 = \frac{S_y^2}{C_y}
\]
Differentiating (3.17) partially with respect to \( k_1 \) and \( k_2 \), equate to zero and solve for \( k_1 \) and \( k_2 \) simultaneously, we obtained
\[ k_1 = \frac{C_1 D_1 - E_1 F_1}{F_1^2 - B_1 C_1} \quad \text{and} \quad k_2 = D_1 F_1 - B_1 E_1 F_1^2 - B_1 C_1. \]
Substituting the results in (3.17), we obtained the minimum mean square error of \( t_{p1} \) denoted by \( MSE(t_{p1})_{\text{min}} \)
\[ MSE(t_{p1})_{\text{min}} = C_1^2 \left[ A_1 + \frac{(C_1 D_1^2 + B_1 E_1^2 - 2D_1 E_1 F_1)}{F_1^2 - B_1 C_1} \right] \]

Differentiating (3.18) partially with respect to \( k_3 \) and \( k_4 \), equate to zero and solve for \( k_3 \) and \( k_4 \) simultaneously, we obtained
\[ k_3 = \frac{C_2 D_2 - E_2 F_2}{F_2^2 - B_2 C_2} \quad \text{and} \quad k_4 = \frac{D_2 F_2 - B_2 E_2}{F_2^2 - B_2 C_2}. \]
Substituting the results in (3.18), we obtained the minimum mean square error of \( t_{p2} \) denoted by \( MSE(t_{p2})_{\text{min}} \)
\[ MSE(t_{p2})_{\text{min}} = C_2^2 \left[ A_2 + \frac{(C_2 D_2^2 + B_2 E_2^2 - 2D_2 E_2 F_2)}{F_2^2 - B_2 C_2} \right] \]

Differentiating (3.19) partially with respect to \( k_5 \) and \( k_6 \), equate to zero and solve for \( k_5 \) and \( k_6 \) simultaneously, we obtained
\[ k_5 = \frac{C_3 D_3 - E_3 F_3}{F_3^2 - B_3 C_3} \quad \text{and} \quad k_6 = \frac{D_3 F_3 - B_3 E_3}{F_3^2 - B_3 C_3}. \]
Substituting the results in (3.19), we obtained the minimum mean square error of \( t_{p3} \) denoted by \( MSE(t_{p3})_{\text{min}} \)
\[ MSE(t_{p3})_{\text{min}} = C_3^2 \left[ A_3 + \frac{(C_3 D_3^2 + B_3 E_3^2 - 2D_3 E_3 F_3)}{F_3^2 - B_3 C_3} \right] \]

Differentiating (3.20) partially with respect to \( k_7 \) and \( k_8 \), equate to zero and solve for \( k_7 \) and \( k_8 \) simultaneously, we obtained
\[ k_7 = \frac{C_4 D_4 - E_4 F_4}{F_4^2 - B_4 C_4} \quad \text{and} \quad k_8 = \frac{D_4 F_4 - B_4 E_4}{F_4^2 - B_4 C_4}. \]
Substituting the results in (3.20), we obtained the minimum mean square error of \( t_{p4} \) denoted by \( MSE(t_{p4})_{\text{min}} \)
\[ MSE(t_{p4})_{\text{min}} = C_4^2 \left[ A_4 + \frac{(C_4 D_4^2 + B_4 E_4^2 - 2D_4 E_4 F_4)}{F_4^2 - B_4 C_4} \right] \]

4 Empirical Study

In this section, we carry out an empirical study to elucidate the performance of our proposed estimators with respect to some existing related estimators using data set below.

Population: [Murthy [18] p.399]

X: Area under wheat in 1963, Y: Area under wheat in 1964

\[ N=34, n=15, \bar{X} = 208.88, \bar{Y} = 199.44, C_X = 0.72, C_Y = 0.75, \rho = 0.98, \lambda_{21} = 1.0045, \lambda_{12} = 0.9406, \lambda_{40} = 3.6161, \lambda_{04} = 2.8266, \lambda_{30} = 0.9206, \lambda_{03} = 2.52, \lambda_{22} = 3.0133 \]
Table 1. MSEs and PREs of proposed and other estimators in the study

<table>
<thead>
<tr>
<th>Estimators</th>
<th>MSE</th>
<th>PRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_0$</td>
<td>0.008003575</td>
<td>100.00</td>
</tr>
<tr>
<td>$t_{ar}$</td>
<td>0.02715658</td>
<td>29.47</td>
</tr>
<tr>
<td>$t_1$</td>
<td>0.006868341</td>
<td>116.53</td>
</tr>
<tr>
<td>$t_2$</td>
<td>0.006868341</td>
<td>116.53</td>
</tr>
<tr>
<td>$t_3$</td>
<td>0.006868341</td>
<td>116.53</td>
</tr>
<tr>
<td>$t_4$</td>
<td>0.006868341</td>
<td>116.53</td>
</tr>
<tr>
<td>$t_5$</td>
<td>0.006868341</td>
<td>116.53</td>
</tr>
<tr>
<td>$t_6$</td>
<td>0.006868341</td>
<td>116.53</td>
</tr>
<tr>
<td>$t_{a1}$</td>
<td>0.006737495</td>
<td>118.79</td>
</tr>
<tr>
<td>$t_{p1}$</td>
<td>0.006033</td>
<td>132.66</td>
</tr>
<tr>
<td>$t_{p2}$</td>
<td>0.005659</td>
<td>141.43</td>
</tr>
</tbody>
</table>

Auxiliary Information : $\bar{x}$, $\bar{y}$

<table>
<thead>
<tr>
<th>Estimators</th>
<th>MSE</th>
<th>PRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_7$</td>
<td>0.00696301</td>
<td>114.94</td>
</tr>
<tr>
<td>$t_8$</td>
<td>0.006962763</td>
<td>114.95</td>
</tr>
<tr>
<td>$t_9$</td>
<td>0.006962763</td>
<td>114.95</td>
</tr>
<tr>
<td>$t_{10}$</td>
<td>0.006962763</td>
<td>114.95</td>
</tr>
<tr>
<td>$t_{11}$</td>
<td>0.006962763</td>
<td>114.95</td>
</tr>
<tr>
<td>$t_{12}$</td>
<td>0.006962763</td>
<td>114.95</td>
</tr>
<tr>
<td>$t_{13}$</td>
<td>0.006962763</td>
<td>114.95</td>
</tr>
<tr>
<td>$t_{14}$</td>
<td>0.00712551</td>
<td>112.32</td>
</tr>
<tr>
<td>$t_{a2}$</td>
<td>0.006013652</td>
<td>133.09</td>
</tr>
<tr>
<td>$t_{p3}$</td>
<td>0.006417</td>
<td>124.72</td>
</tr>
<tr>
<td>$t_{p4}$</td>
<td>0.004996</td>
<td>160.19</td>
</tr>
</tbody>
</table>

Auxiliary Information : $S^2_{X}, S^2_{Y}$

The formula for Percent Relative Efficiency (PRE) is

$$\text{PRE (estimators)} = \frac{\text{MSE}(\text{estimator})}{\text{MSE}(\text{estimator})} \times 100$$

Table 1 shows the MSEs and PREs of the proposed and other estimators considered in the study. Results in Table 1 revealed that proposed estimators has minimum MSEs and higher PREs compared to other competing existing estimators.

5 Simulation Study

We perform some simulation experiments to find the relative efficiency (RE) of the proposed estimator compared with the existing estimators.

The following steps have been used for the simulation:

1. We generated bivariate random observations of size $N=1000$ units from a bivariate normal distribution with parameters $\mu_x = 100$, $\sigma_x = 11$, and $\mu_y = 120$, $\sigma_y = 14$ and $p = 0.9$.
2. Similarly, generate the population-II with the parameters $\mu_x = 3$, $\sigma_x = 2$, $\mu_y = 5$ and $\sigma_y = 3$.
3. A sample of size $n = 20$ has been selected from this simulated population.
4. Sample statistics that is the sample mean, sample variance, and the values of the suggested and competing estimators of population CV are calculated for this sample.
5. Steps (3) and (4) are repeated $m = 10,000$ times.
6. The MSE of every estimator $t_i$ is calculated through the formula, $\text{MSE}(t_i) = \frac{1}{m} \sum_{j=1}^{m} (t_{ij} - \bar{t}_{ij})^2$.

The formula for Percent Relative Efficiency (PRE) is

$$\text{PRE (estimators)} = \frac{\text{MSE}(\text{estimator})}{\text{MSE}(\text{estimator})} \times 100$$

Table 1 shows the MSEs and PREs of the proposed and other estimators considered in the study. Results in Table 1 revealed that proposed estimators has minimum MSEs and higher PREs compared to other competing existing estimators.
7. The Percentage Relative Efficiency (PRE) of the estimator $t_i$ over the mean per unit estimator $t_0$ given by,

$$\text{PRE}(t_i) = \frac{\sqrt{\text{MSE}(t_0)}}{\text{MSE}(t_i)} \times 100 \quad i = r, 1, 2, \ldots p^4$$

Table 2. MSE values of competing and suggested estimators and PRE with respect to $\hat{C}_x$ for symmetric simulated population

<table>
<thead>
<tr>
<th>Estimators</th>
<th>PRE for population 1</th>
<th>PRE for population 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_0$</td>
<td>100</td>
<td>100.00</td>
</tr>
<tr>
<td>$t_{ar}$</td>
<td>101.02</td>
<td>99.44</td>
</tr>
<tr>
<td>$t_1$</td>
<td>115.63</td>
<td>118.98</td>
</tr>
<tr>
<td>$t_2$</td>
<td>115.56</td>
<td>118.76</td>
</tr>
<tr>
<td>$t_3$</td>
<td>115.54</td>
<td>118.77</td>
</tr>
<tr>
<td>$t_4^1$</td>
<td>115.63</td>
<td>118.36</td>
</tr>
<tr>
<td>$t_5^1$</td>
<td>115.34</td>
<td>118.76</td>
</tr>
<tr>
<td>$t_6^1$</td>
<td>115.83</td>
<td>118.35</td>
</tr>
<tr>
<td>$t_{a1}$</td>
<td>133.43</td>
<td>136.67</td>
</tr>
<tr>
<td>$t_{p1}$</td>
<td>141.67</td>
<td>161.29</td>
</tr>
<tr>
<td>$t_{p2}$</td>
<td>172.43</td>
<td>174.76</td>
</tr>
<tr>
<td>Auxiliary Information : $\bar{X}, \bar{x}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_7$</td>
<td>114.16</td>
<td>116.62</td>
</tr>
<tr>
<td>$t_8$</td>
<td>114.56</td>
<td>116.95</td>
</tr>
<tr>
<td>$t_9$</td>
<td>114.64</td>
<td>116.95</td>
</tr>
<tr>
<td>$t_{10}$</td>
<td>114.54</td>
<td>116.95</td>
</tr>
<tr>
<td>$t_{11}$</td>
<td>114.45</td>
<td>116.95</td>
</tr>
<tr>
<td>$t_{12}$</td>
<td>114.67</td>
<td>116.95</td>
</tr>
<tr>
<td>$t_{13}$</td>
<td>114.33</td>
<td>116.95</td>
</tr>
<tr>
<td>$t_{14}$</td>
<td>113.22</td>
<td>115.32</td>
</tr>
<tr>
<td>$t_{a2}$</td>
<td>135.29</td>
<td>133.09</td>
</tr>
<tr>
<td>$t_{p3}$</td>
<td>133.09</td>
<td>145.72</td>
</tr>
<tr>
<td>$t_{p4}$</td>
<td>175.99</td>
<td>180.19</td>
</tr>
<tr>
<td>Auxiliary Information : $S_x^2, s_x^2$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 shows that our proposed estimators perform better than the existing estimators as our proposed estimators have greater PRE.

6 Conclusion

In this article, we have proposed estimators for the coefficient of Variation (CV) using the information of auxiliary variables. The expressions for Bias and MSE of the suggested estimators have been derived up to the first order of approximation. An empirical approach and simulation study for comparing the efficiency of the proposed estimators with other estimators have been used. The results have been shown in Tables 1 & 2. The Tables show that the proposed estimators turn out to be more efficient as compared to the other estimators for both populations. The proposed estimators are found to be rather improved in terms of lesser MSE and greater PRE as compared to the existing estimators in both real and simulated data sets [19-22]. Based on our empirical study and simulation study, we can conclude that our proposed estimators can be preferred over the other estimators taken in this paper in several real situations. Hence we recommend that the our proposed estimator should be used in both theoretical and real life applications like agriculture sciences, mathematical sciences, biological sciences, poultry, business, economics, commerce, social sciences etc.
Competing Interests

Authors have declared that no competing interests exist.

References


DOI: 10.1080/03610926.2022.2091784

© 2022 Singh and Kumari; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/91098