Original Research Article

Almost unbiased estimators for population coefficient of variation using auxiliary information

.

ABSTRACT

In this paper we have proposed an almost unbiased ratio estimator for the finite coefficient of variation(CV). It has been shown that Archana and Rao (2014) ratio and product type coefficient of variation estimators are particular members of the proposed estimator. Also, we have proposed an exponential ratio type and log ratio type estimators for estimating population coefficient of variation.

Commented [WU1]: The abstract is not comprehensive and thus it should contain the following;
1.Objectives of the paper,
2.The datasets used in assessing the perfomances of the estimators,
3. The criteria adopted in assessing the perfomances of the estimators,
4. The conclusion drawn from the findings of the study and
5. Recommendations and suggestion.

Keywords: auxiliary information; bias; mean squared error; coefficient of variation; log type estimator.

1. INTRODUCTION

Research of Cochran (1977) is generally associated with the idea of incorporating auxiliary information to improve estimator's efficiency. The method basically consists of incorporating more information to the sampled data beyond what can be obtained from it individually for the purpose of increasing the accuracy or effectiveness of estimators. Researchers might be enabled to reduce variability in samples and provide more precise estimates of population parameters by utilizing auxiliary information. The basic principles for this methodology were established by Cochran (1977) and it is now commonly employed in many different kinds of domains, such as survey sampling, econometrics, and statistics.

Commented [WU2]: Re-cast the sentence to bring out the actual meaning.

Ratio and product estimators are widely used in survey sampling and other fields where auxiliary information is available and can be utilized to improve the accuracy and efficiency of estimators. When there is a positive correlation between an auxiliary variable and the study variable of interest, Cochran(1940) established the concept of ratio estimators as an approach to utilize auxiliary information. Ratio estimators calculate ratios among the study variable's sample means or totals and the auxiliary variables, taking into consideration any known population variables. The above technique makes utilizes the connection between the auxiliary and study variables in order to improve estimation efficiency. Ratio estimators enable us to analyze an estimated value with other known information to estimate a value more accurately. To get more accurate estimates, they consider the correlation between various variables. For example, we may accurately estimate the total income of a neighborhood if we know the average income of that neighborhood and the population of the entire city. This can be done by using a ratio estimator. By utilizing more information, this approach increases the precision of our estimations.

Commented [WU3]: Re-cast the sentence to bring out the actual meaning.

On the other hand, Robson(1957) and Murthy(1964) proposed the product estimator, which is another method for incorporating auxiliary information into estimation. The product estimator involves forming the product of the study variable and the auxiliary variable and then using this product as the basis for estimation. Similar to ratio estimators, the product estimator seeks to capitalize on the association between the auxiliary variable and the study variable to enhance the precision of the estimates. Number of authors, including Solanki et al. (2012), Ray and Sahai(1980), and Srivastava and Jhajj (1981), have made significant contributions to the utilization of auxiliary information for estimating population parameters such as the population mean, variance, standard deviation, and other related statistics. Some important works illustrating use of auxiliary information at estimation stage are Singh et al.(2018), Singh and Kumar(2011), Malik and Singh(2013) etc.

Evaluating the coefficient of variation hasn't received as much attention from researchers in the past. Some authors have recently started to address this problem. In another context, researchers are concentrating more on determining a dataset's variability in relation to its mean, which is a fundamental statistical measure. For example, Das and Tripathi(1992) were the first to suggest an estimator for the coefficient of variation when samples were chosen using simple random sampling without replacement (SRSWOR). Other researchers, such as Patel and Rina(2009), have also explored into this area. Breunig(2001) suggested an almost unbiased estimator of the coefficient of variation. Additionally, Rajyaguru and Gupta(2005) explored estimating the coefficient of variation under different sampling methods like simple random sampling and stratified random sampling in 2002 and 2003, .

Commented [WU4]: Authors have not established a gap in extant literature that necessitates this study. Objectives of the study have not been stated.

In this paper, we have proposed an almost unbiased estimator for estimation of population coefficient of variation utilizing information on a single auxiliary variable in SRSWOR.

Commented [WU5]: The authors needs to state the significance or contribution of this study.

Let's consider a finite population P = (P_1, P_2, P_3, ..., P_N) of size `N` and each unit are uniquely defined. Let Y and X defined as study and auxiliary variable and Y_i and Xi are the values corresponding their unit i (i = 1, 2, 3, ..., N).

Let us consider a SRS of size n drawn from the given population of `N` units and corresponding unit y_i and x_i.

Commented [WU6]: delete

Let, s

$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ and $\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ are the population means of the study and auxiliary variables Y and X,

$S_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(Y_i - \overline{Y})^2$ is the population variance of the study variable Y,

$S_x^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(X_i - \overline{X})^2$ is the population variance of the auxiliary variable X,

$S_{xy} = \frac{1}{(N-1)}\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})$ is the population covariance of the auxiliary and study variable Y and X,

$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ - are the sample mean of the study and auxiliary variables Y and X,

Commented [WU7]: add 'respectively'

$s_y^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \overline{y})^2$ - is the sample variance of the study variable y,

$s_x^2 = \dfrac{1}{(n-1)}\sum_{i=1}^{n}(x_i-\bar{x})^2$ - is the sample variance of the auxiliary variable x.

Let us define sampling errors for both mean and variance of study and auxiliary variables as-

$e_0 = \dfrac{\bar{y}-\bar{Y}}{\bar{Y}}$, $e_1 = \dfrac{\bar{x}-\bar{X}}{\bar{X}}$, $e_2 = \dfrac{(s_y^2 - S_y^2)}{S_y^2}$, $e_3 = \dfrac{(s_x^2 - S_x^2)}{S_x^2}$ such that

$\bar{y} = \bar{Y}(1+e_0)$, $\bar{x} = \bar{X}(1+e_1)$, $s_y^2 = S_y^2(1+e_2)$, $s_y^2 = S_y^2(1+e_2)$, $s_x^2 = S_x^2(1+e_3)$

$E(e_0) = E(e_1) = E(e_2) = E(e_3) = 0$,

$E(e_0^2) = \gamma C_y^2$, $E(e_1^2) = \gamma C_x^2$, $E(e_2^2) = \gamma(\lambda_{40}-1)$, $E(e_3^2) = \gamma(\lambda_{04}-1)$,

$E(e_0 e_1) = \gamma\rho C_y C_x$, $E(e_0 e_2) = \gamma C_y \lambda_{30}$, $E(e_0 e_3) = \gamma C_y \lambda_{12}$,

$E(e_1 e_2) = \gamma C_x \lambda_{21}$, $E(e_1 e_3) = \gamma C_x \lambda_{03}$, $E(e_2 e_3) = \gamma(\lambda_{22}-1)$.

Here, $\gamma = \dfrac{1}{n}(1-f)$, $f = \dfrac{n}{N}$, f is known as sampling fraction, C_y and C_x are the population coefficient of variations of study variable Y and auxiliary variable X respectively and defined as $C_y = \dfrac{S_y}{\bar{Y}}$ and $C_x = \dfrac{S_x}{\bar{X}}$. ρ is the correlation coefficient between X and Y.

Commented [WU8]: insert comma

Commented [WU9]: insert comma

In general form,

$$\mu_{rs} = \frac{\sum_{i=1}^{N}(y_i-\bar{y})^r(x_i-\bar{x})^s}{(\mathrm{N}-1)} \quad \text{and} \quad \lambda_{rs} = \frac{\mu_{rs}}{(\mu_{20}^{r/2}\mu_{02}^{s/2})}.$$

2. EXISTING ESTIMATORS

Usual estimator t_0 for estimating Cy is given by

$$t_0 = \hat{C}_y = \frac{s_y}{\bar{y}} \tag{1}$$

The bias of the estimator t_0 is given by:

$$Bias(t_0) = C_y\gamma\left(C_y^2 - \frac{1}{8}(\lambda_{40}-1) - \frac{1}{2}C_y\lambda_{30}\right) \tag{2}$$

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

$$MSE(t_0) = C_y^2\gamma(C_y^2 + \frac{1}{4}(\lambda_{40}-1) - C_y\lambda_{30}) \tag{3}$$

Archana & Rao (2014) introduced estimators t_1 and t_2 for calculating the finite population coefficient of variation as follows:

$$t_1 = C_y\left(\frac{S_x^2}{s_x^2}\right) \tag{4}$$

$$t_2 = C_y\left(\frac{s_x^2}{S_x^2}\right) \tag{5}$$

The bias of the estimators t_1 and t_2 are respectively given as-

Commented [WU10]: insert comma

Commented [WU11]: insert comma

Commented [WU12]: delete

$$Bias(t_1) = C_y\gamma\left(C_y^2 - \frac{1}{8}(\lambda_{40}-1) - \frac{1}{2}C_y\lambda_{30} + C_y\lambda_{12} + (\lambda_{04}-1) - \frac{1}{2}(\lambda_{22}-1)\right) \tag{6}$$

$$Bias(t_2) = C_y\gamma\left(C_y^2 - \frac{1}{8}(\lambda_{40}-1) - \frac{1}{2}C_y\lambda_{30} + \frac{1}{2}(\lambda_{22}-1) - C_y\lambda_{12}\right) \tag{7}$$

MSE of the estimators t_1 and t_2 are respectively given as-

Commented [WU13]: insert comma

Commented [WU14]: insert comma

Commented [WU15]: delete

$$MSE(t_1) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40}-1) - C_y\lambda_{30} + 2C_y\lambda_{12} + (\lambda_{04}-1) - (\lambda_{22}-1)\right) \tag{8}$$

$$MSE(t_2) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40}-1) - C_y\lambda_{30} - 2C_y\lambda_{12} + (\lambda_{04}-1) + (\lambda_{22}-1)\right) \tag{9}$$

3. Proposed almost unbiased estimator

Commented [WU16]: Capilalizrd each word for manuscript headings

Let,

$$t_0 = C_y, t_1 = C_y\left(\frac{S_x^2}{s_x^2}\right), t_2 = C_y\left(\frac{s_x^2}{S_x^2}\right) \tag{10}$$

such that t_0 , t_1 , $t_2 \in L$, where L denotes the set of all possible estimators for estimating the population coefficient of variation C_y .

By definition, the set L is a linear variety if

$$t_g = \sum_{i=0}^{2} g_i t_i \in L \tag{11}$$

$$t_g = g_0 C_y + g_1 C_y\left(\frac{S_x^2}{s_x^2}\right) + g_2 C_y\left(\frac{s_x^2}{S_x^2}\right)$$

For $\sum_{i=0}^{2} g_i = 1, g_i \in R$ (12)

where g_i (i = 0, 1, 2) denotes the statistical constants and R denotes the set of real numbers.

Commented [WU17]: Delete spaces

Table 1. Members of the proposed family t_g of estimators

g_0	g_1	g_2	**Estimators**
1	0	0	$\hat{C}_y$
0	1	0	$C_y\left(\frac{S_x^2}{s_x^2}\right)$
0	0	1	$C_y\left(\frac{s_x^2}{S_x^2}\right)$

To obtain the bias and MSE of the estimator t_g, we write t_g in the form of error terms as-

$$t_g = C_y\left(1+e_2\right)^{1/2}\left(1+e_0\right)^{-1}\left[g_0 + g_1\left(\frac{1}{(1+e_3)}\right) + g_2\left(1+e_3\right)\right] \quad (13)$$

Expanding the right hand side of equation (13) and retaining terms up to second powers of e's, we have

$$t_g = C_y\begin{bmatrix} 1-e_0+e_0^2+\frac{1}{2}e_2-\frac{1}{2}e_0e_2-\frac{1}{8}e_2^2-(g_1-g_2)e_3 \\ +(g_1-g_2)e_0e_3-(g_1-g_2)\frac{1}{2}e_2e_3+g_1e_3^2 \end{bmatrix} \quad (14)$$

Subtracting C_y and then taking expectation both sides, we get the bias of the estimator t_g, up to the first order of approximation as-

$$Bias(t_g) = C_y\gamma\begin{pmatrix} C_y^2-\frac{1}{8}(\lambda_{40}-1)-\frac{1}{2}C_y\lambda_{30}+(g_1-g_2)C_y\lambda_{12} \\ +g_1(\lambda_{04}-1)-(g_1-g_2)\frac{1}{2}(\lambda_{22}-1) \end{pmatrix} \quad (15)$$

From equation (15),
We have

$$\left(t_g - C_y\right) \cong C_y\left[\frac{1}{2}e_2 - e_0 - (g_1-g_2)e_3\right] \quad (16)$$

where,
$(g_1 - g_2)$=H. (17)

Squaring both sides of equation (16) and then taking expectations , we get MSE of the estimator t_g, up to the first order of approximation, as-

$$MSE(t_g) = C_y^2\gamma\left(C_y^2+\frac{1}{4}(\lambda_{40}-1)-C_y\lambda_{30}+2HC_y\lambda_{12}+H^2(\lambda_{04}-1)-H(\lambda_{22}-1)\right) \quad (18)$$

Which is minimum when

$$H = \frac{1}{2}\frac{\left(\lambda_{22}-1\right)}{\left(\lambda_{04}-1\right)} - \frac{C_y\lambda_{12}}{\left(\lambda_{04}-1\right)}. \quad (19)$$

Putting this value of $H=\frac{1}{2}\frac{(\lambda_{22}-1)}{(\lambda_{04}-1)}-\frac{C_y\lambda_{12}}{(\lambda_{04}-1)}$ in equation (2.18) we get the Min. MSE of the estimator t_g as-

$$Min.MSE(t_g)=C_y^2\gamma\left(C_y^2+\frac{1}{4}(\lambda_{40}-1)-C_y\lambda_{30}+2HC_y\lambda_{12}+H^2(\lambda_{04}-1)-H(_{22}-1)\right) \qquad (20)$$

From equation (17) and (19)
we have,

$$(g_1-g_2)=H=\frac{1}{2}\frac{(\lambda_{22}-1)}{(\lambda_{04}-1)}-\frac{C_y\lambda_{12}}{(\lambda_{04}-1)} \qquad (21)$$

From equation (12) and (17), we have only two equations in three unknowns. It is not possible to find the unique values for g_i's , (i = 0, 1, 2). In order to get unique values of g_i's , we shall impose the linear restriction.

$$\sum_{i=0}^{2}g_iB(t_i)=0 \qquad (22)$$

Such that

$$g_0B(t_0)+g_1B(t_1)+g_2B(t_2)=0 \qquad (23)$$

where $B(t_i)$ denotes the bias in the i^{th} estimator.
Equations (2.12), (2.17) and (2.23) can be written in the matrix form as-

$$\begin{bmatrix}1 & 1 & 1\\ 0 & 1 & -1\\ B(t_0) & B(t_1) & B(t_2)\end{bmatrix}\begin{bmatrix}g_0\\ g_1\\ g_2\end{bmatrix}=\begin{bmatrix}1\\ H\\ 0\end{bmatrix} \qquad (24)$$

From the system of equation (2.24) , we get the unique values of g_i's (i=0, 1, 2) as-

$$g_0=\frac{B(t_2)+B(t_1)-HB(t_2)-HB(t_1)}{B(t_2)+B(t_1)-2B(t_0)} \qquad (25)$$

$$g_1=\frac{HB(t_2)-B(t_0)-HB(t_0)}{B(t_2)+B(t_1)-2B(t_0)} \qquad (26)$$

$$g_2=\frac{HB(t_0)-B(t_0)-HB(t_1)}{B(t_2)+B(t_1)-2B(t_0)} \qquad (27)$$

such that

$$g_0+g_1+g_2=1 \qquad (28)$$

Use of these g_i's (i=0, 1, 2) remove the bias up to terms of order $o(n^{-1})$

4. Another almost unbiased estimator

Commented [WU18]: This should be under section 3, and empeirical study should be under section 4.

In this section we propose another almost unbiased estimator t_{g1} for coefficient of variation. For this we have taken three estimators m_0 , m_1 and m_2 which are defined as

$$m_0=C_y=(t_0) \qquad (29)$$

The bias of the estimator m_0 is given by:

$$Bias(m_0) = C_y \gamma \left(C_y^2 - \frac{1}{8}(\lambda_{40} - 1) - \frac{1}{2} C_y \lambda_{30} \right) \quad (30)$$

The Mean square error (MSE) expression of the estimator m_0 is given by:

$$MSE(m_0) = C_y^2 \gamma (C_y^2 + \frac{1}{4}(\lambda_{40} - 1) - C_y \lambda_{30}) \quad (31)$$

The exponential and logarithmic estimator for estimating population coefficient of variation is given as follows-

$$m_1 = C_y \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2} \right) \quad (32)$$

$$m_2 = C_y \left\{ 1 + \log\left(\frac{s_x^2}{S_x^2} \right) \right\} \quad (33)$$

The bias of the estimators m_1 and m_2 are respectively given as-

$$Bias(m_1) = C_y \gamma \left(C_y^2 - \frac{1}{8}(\lambda_{40} - 1) - \frac{1}{2} C_y \lambda_{30} + \frac{1}{2} C_y \lambda_{12} + \frac{3}{8}(\lambda_{04} - 1) - \frac{1}{4}(\lambda_{22} - 1) \right) \quad (34)$$

$$Bias(m_2) = C_y \gamma \left(C_y^2 - \frac{1}{8}(\lambda_{40} - 1) - \frac{1}{2} C_y \lambda_{30} + \frac{1}{2}(\lambda_{22} - 1) - C_y \lambda_{12} - \frac{1}{2}(\lambda_{04} - 1) \right) \quad (35)$$

MSE of the estimators m_1 and m_2 are respectively given as-

$$MSE(m_1) = C_y^2 \gamma \left(C_y^2 + \frac{1}{4}(\lambda_{40} - 1) - C_y \lambda_{30} + C_y \lambda_{12} + \frac{1}{4}(\lambda_{04} - 1) - \frac{1}{2}(\lambda_{22} - 1) \right) \quad (36)$$

$$MSE(m_2) = C_y^2 \gamma \left(C_y^2 + \frac{1}{4}(\lambda_{40} - 1) - C_y \lambda_{30} - 2 C_y \lambda_{12} + (\lambda_{04} - 1) + (\lambda_{22} - 1) \right) \quad (37)$$

m_0, m_1 and $m_2 \in L$, where L denotes the set of all possible estimators for estimating the population coefficient of variation C_y.

By definition, the set L is a linear variety if

$$t_{g1} = \sum_{i=0}^{2} l_i m_i \in L \quad (38)$$

$$t_{g1} = l_0 m_0 + l_1 m_1 + l_2 m_2 \quad (39)$$

$$t_{g1} = l_0 C_y + l_1 C_y \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2} \right) + l_2 C_y \left\{ 1 + \log\left(\frac{s_x^2}{S_x^2} \right) \right\}$$

For $\sum_{i=0}^{2} l_i = 1$, $l_i \in R$ (40)

where l_i (i = 0, 1, 2) denotes the statistical constants and R denotes the set of real numbers.

Table 2. Members of the proposed family t_{g1} of estimators

l_0	l_1	l_2	**Estimators**
1	0	0	C_y
0	1	0	$C_y \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)$
0	0	1	$C_y \left\{1 + \log\left(\frac{s_x^2}{S_x^2}\right)\right\}$

To obtain the bias and MSE of the t_{g1}, we write t_{g1} in the form of error terms as-

$$t_{g1} = C_y\left(1+e_2\right)^{1/2}\left(1+e_0\right)^{-1}\left[l_0 + l_1 \exp\left(\frac{-e_3}{2+e_3}\right) + l_2\left\{1+\log(1+e_3)\right\}\right] \quad (41)$$

Expanding the right hand side of (41) and retaining terms up to second powers of e's we have

$$t_{g1} = C_y\begin{bmatrix} 1 - e_0 + e_0^2 + \frac{1}{2}e_2 - \frac{1}{2}e_0e_2 - \frac{1}{8}e_2^2 - (\frac{1}{2}l_1 - l_2)e_3 \\ +(\frac{3}{8}l_1 - \frac{1}{2}l_2)e_3^2 + (\frac{1}{2}l_1 - l_2)e_0e_3 - (\frac{1}{4}l_1 - \frac{1}{2}l_2)e_2e_3 \end{bmatrix} \quad (42)$$

Subtracting C_y from both sides of equation (42) and then taking expectations , we get the bias of the estimator t_{g1}, up to the first order of approximation as-

$$Bias(t_{g1}) = C_y\gamma\begin{pmatrix} C_y^2 - \frac{1}{8}(\lambda_{40}-1) - \frac{1}{2}C_y\lambda_{30} + (\frac{1}{2}l_1 - l_2)C_y\lambda_{12} \\ +(\frac{3}{8}l_1 - \frac{1}{2}l_2)(\lambda_{04}-1) - (\frac{1}{4}l_1 - \frac{1}{2}l_2)\frac{1}{2}(\lambda_{22}-1) \end{pmatrix} \quad (43)$$

From (42), we have

$$\left(t_{g1} - C_y\right) \cong C_y\left[\frac{1}{2}e_2 - e_0 - \left(\frac{1}{2}l_1 - l_2\right)e_3\right] \quad (44)$$

where,

$$\frac{1}{2}l_1 - l_2 = H_1 \quad (45)$$

Squaring both sides of equation (43) and then taking expectations , we get MSE of the estimator t_{g1}, up to the first order of approximation, as-

$$MSE(t_{g1}) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40}-1) - C_y\lambda_{30} + 2H_1C_y\lambda_{12} + {H_1}^2(\lambda_{04}-1) - H_1(\lambda_{22}-1)\right) \quad (46)$$

which is minimum when

$$H_1 = \frac{1}{2}\frac{(\lambda_{22}-1)}{(\lambda_{04}-1)} - \frac{C_y\lambda_{12}}{(\lambda_{04}-1)}. \quad (47)$$

Putting the value of $H_1 = \frac{1}{2}\frac{(\lambda_{22}-1)}{(\lambda_{04}-1)} - \frac{C_y\lambda_{12}}{(\lambda_{04}-1)}$ in equation (46) the minimum MSE value of the estimator t_{g1} is given by

$$Min.MSE(t_{g1}) = C_y^2\gamma\left(C_y^2 + \frac{1}{4}(\lambda_{40}-1) - C_y\lambda_{30} + 2H_1C_y\lambda_{12} + H_1^2(\lambda_{04}-1) - H_1(\lambda_{22}-1)\right) \quad (48)$$

From equations (45) and (47), we have

$$\left(\frac{1}{2}l_1 - l_2\right) = H_1 = \frac{1}{2}\frac{(\lambda_{22}-1)}{(\lambda_{04}-1)} - \frac{C_y\lambda_{12}}{(\lambda_{04}-1)} \quad (49)$$

From equation (38) and (45), we have only two equations in three unknowns. It is not possible to find the unique values for l_i's , i = 0, 1, 2. In order to get unique values of l_i's , we shall impose the linear restriction.

$$\sum_{i=0}^{2} l_i B(m_i) = 0 \quad (50)$$

such that

$$l_0B(m_0) + l_1B(m_1) + l_2B(m_2) = 0 \quad (51)$$

here $B(m_i)$ denotes the bias in the i^{th} (i=0,1,2) estimator.

Equations (40), (45) and (51) can be written in the matrix form

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & \frac{1}{2} & -1 \\ B(m_0) & B(m_1) & B(m_2) \end{bmatrix} \begin{bmatrix} l_0 \\ l_1 \\ l_2 \end{bmatrix} = \begin{bmatrix} 1 \\ H_1 \\ 0 \end{bmatrix} \quad (52)$$

Where,

$B(m_0)$, $B(m_1)$ and $B(m_2)$ are defined in equation (31), (34) and (35).

From the system of equation (52) , we get the unique values of l_i's (i=0, 1, 2) respectively as-

$$l_0 = \frac{\frac{1}{2}B(m_2) + B(m_1) - H_1B(m_2) + H_1B(m_1)}{\frac{1}{2}B(m_2) + B(m_1) - \frac{3}{2}B(m_0)} \quad (53)$$

$$l_1 = \frac{H_1B(m_2) - B(m_0) - H_1B(m_0)}{\frac{1}{2}B(m_2) + B(m_1) - \frac{3}{2}B(m_0)} \quad (54)$$

$$l_2 = \frac{H_1B(m_0) - H_1B(m_1) - \frac{1}{2}B(m_0)}{\frac{1}{2}B(m_2) + B(m_1) - \frac{3}{2}B(m_0)} \quad (55)$$

Where,

$$l_0 + l_1 + l_2 = 1 \quad (56)$$

Use of these l_i's (i=0, 1, 2) remove the bias up to terms of order $o(n^{-1})$.

Commented [WU19]: A new section called 'Theoretical Efficiency Comparison' should be provided. Under this section the authors should clearly obtain the theoretical condition under which the proposed estimators are efficient that the existinf estimators considered.

5. Empirical Study

In this section, we will carry out empirical study to demonstrate the performance of the proposed estimator over existing ones using two real data sets.

Population 1: [Source: Murthy (1967), p.399]

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

N=34, n=15 , C_x = 0.72, C_y =0.75, ρ =0.98, λ_{21} =1.0045, λ_{12} =0.9406, λ_{40} =3.6161, λ_{04} =2.8266, λ_{30} =1.1128, λ_{03} =0.9206, λ_{22} =3.01133, $\bar{Y}$ =199.44, $\bar{X}$ =208.88

Table 3. The MSE and PRE of the existing and the proposed estimators

Estimators	MSE	PRE
t_0	0.03808827	100.00
t_{AR}	0.188603	20.1948
t_{AR_1}	0.2261359	16.84297
t_1	0.071025	53.62631
t_2	0.226136	16.84309
t_g	**0.037568**	**101.3844**
t_{g1}	**0.037568**	**101.3844**

Commented [WU20]: The results are appropriately presented, but not appropriately described.
The empirical biases of the proposed and existing estimators needs to be computed.

Population 2: [Source: (S. Singh, 2003), p.1116]

Table 4. The MSE and PRE of the existing and the proposed estimators
X: Number of fish caught in year 1993, Y: Number of fish caught in year 1995
N = 69, n=40 , C_x = 1.38, C_y =1.35, ρ =0.96, λ_{21} =2.19, λ_{12} =2.3, λ_{40} =7.66, λ_{04} = 9.84, λ_{30} =1.11, λ_{03} =2.52, λ_{22} =8.19, $\bar{Y}$ =4514.89, $\bar{X}$ =4591.07

Table5. Values of g_i's (i=0,1,2)

S.N.	Scalars	Population I	Population II
1	g_0	**3.23672**	-0.01106
2	g_1	**-1.03618**	0.533243
3	g_2	**-1.20054**	0.477814

Estimators	MSE	PRE
t_0	0.008003	100.00
t_{AR}	0.03365	23.78054
t_{AR_1}	0.05890	13.58789
t_1	0.01128	70.94231
t_2	0.05886	13.59669
t_g	**0.00697**	**114.8289**
t_{g1}	**0.00697**	**114.8289**

Commented [WU21]: The results are appropriately presented, but not appropriately described.
In addition, the empirical biases of the proposed and existing estimators needs to be computed.

Using these values of g_i's (i=0,1,2) given in the Table 3, one can reduce the bias to the order $o(n^{-1})$ in the estimator t_g.

Table 6. Values of l_i's (i =0,1,2)

S.N.	Scalars	Population I	Population II
1.	l_0	**0.631909**	1.677949
2.	l_1	**0.354966**	-0.41501
3.	l_2	**0.013126**	-0.26294

Using these values of l_i's (i=0,1,2) given in the Table 4, one can reduce the bias to the order $o(n^{-1})$ in the estimator t_{g_1}.

6. Simulation study

In this section we have done simulation analysis.
The following steps have been used for the simulation:
1. We have generated bivariate random observations of size N=1000 units from a bivariate normal distribution with parameters μ_x =3, σ_x =4, and μ_y =5, σ_y = 9 and ρ = 0.95.
2. Sample of sizes n = 150, 200 and 250 have been selected from this simulated population.
3. Sample statistics that is the sample mean, sample variance, and the values of the suggested and existing estimators of population CV are calculated for these samples.

4.Steps (3) and (4) are repeated m=10,000 times.

$$PRE(estimator) = \frac{MSE(t_0)}{MSE(estimator)} * 100$$

Table 7. MSE and PRE values of existing and proposed estimators.

Estimators	N=1000, n=150		N=1000, n=200		N=1000, n=250	
	MSE	PRE	MSE	PRE	MSE	PRE
t_0	0.001688	100.00	0.001355	100.00	0.000793	100.00
t_{AR}	0.002342	72.08397	0.001386	97.73829	0.000992	79.97066
t_{AR_1}	0.00932	18.11233	0.007178	18.87754	0.004357	18.20195
t_1	0.000979	172.381	0.000639	212.0795	0.000422	187.9839
t_2	0.00932	18.11233	0.007178	18.87754	0.004357	18.20195
t_g	**0.000953**	**177.043**	**0.000639**	**212.1073**	**0.000417**	**190.3548**
t_{g1}	**0.000953**	**177.043**	**0.000639**	**212.1073**	**0.000417**	**190.3548**

Commented [WU22]: The results are appropriately presented, but not appropriately described.
In addition, biases of the proposed and existing estimators needs to be computed using the simulated data.

DISCUSSION

We have proposed two almost unbiased estimators t_g and t_{g_1} for the estimation of population coefficient of variation utilizing information on a single auxiliary variable in srswor and compared them with some existing estimators.

Commented [WU23]: This sould be Section 6 whent the appropriate numbering is followed.

Commented [WU24]: This is not sufficent

Conclusion

In this paper we have proposed almost unbiased estimators for C_y. With the help of two real data sets and simulation study we have shown that our proposed estimators have minimum variance Table3, Table 4 and Table7 clearly shows that the suggested estimators t_g and t_{g_1} under optimum conditions are better than usual estimator t_0 (= C_y), Archana & Rao (2014) estimators t_{AR} and t_{AR_1} and exponential ratio type t_1 and improved log ratio type estimator t_2.

Commented [WU25]: Reframe this to capture the bias results recommended in the result section.

References

Archana, V., & Rao, A. (2014). Some improved estimators of co-efficient of variation from bi-variate normal distribution: A Monte Carlo comparison. *Pakistan Journal of Statistics and Operation Research*, 87–105.

Breunig, R. (2001). An almost unbiased estimator of the coefficient of variation. *Economics Letters*, *70*(1), 15–19. https://doi.org/10.1016/S0165-1765(00)00351-7

Cochran, W. G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural Science*, *30*(2), 262–275.

Cochran, W. G. (1977). *Sampling techniques*. john wiley & sons. https://books.google.com/books?hl=en&lr=&id=xbNn41DUrNwC&oi=fnd&pg=PA1&dq=Sampling+Techniques+Book+by+William+Cochran&ots=TYrUiAwmkZ&sig=V1VYvy7LsjGVVBuUtFfbRsczbOo

Das, A. K., & Tripathi, T. P. (1992). Use of auxiliary information in estimating the coefficient of variation. *Alig. J. of. Statist*, *12*, 51–58.

Malik, S., & Singh, R. (2013). An improved estimator using two auxiliary attributes. *Applied Mathematics and Computation*, *219*(23), 10983–10986. https://doi.org/10.1016/j.amc.2013.05.014

Murthy, M. N. (1964). Product method of estimation. *Sankhyā: The Indian Journal of Statistics, Series A*, 69–74.

Patel, P. A., & Rina, S. (2009). A Monte Carlo comparison of some suggested estimators of Co-efficient of variation in finite population. *Journal of Statistics Sciences*, *1*(2), 137–147.

Rajyaguru, A., & Gupta, P. C. (2005). On the estimation of the coefficient of variation from finite population-II. *Model Assisted Statistics and Applications*, *1*(1), 57–66. https://doi.org/10.3233/MAS-2006-1110

Ray, S. K., & Sahai, A. (1980). *Efficient Families of Ratio and Product-Type Estimators*.

Robson, D. S. (1957). Applications of Multivariate Polykays to the Theory of Unbiased Ratio-Type Estimation. *Journal of the American Statistical Association*, *52*(280), 511–522. https://doi.org/10.1080/01621459.1957.10501407

Sing, R., & Kumar, M. (2011). A note on transformations on auxiliary variable in survey sampling. *Model Assisted Statistics and Applications*, *6*(1), 17–19. https://doi.org/10.3233/MAS-2011-0154

Singh, R., Mishra, M., Singh, B. P., Singh, P., & Adichwal, N. K. (2018). Improved estimators for population coefficient of variation using auxiliary variable. *Journal of Statistics and Management Systems*, *21*(7), 1335–1355. https://doi.org/10.1080/09720510.2018.1503405

Singh, S. (2003). *Advanced Sampling Theory With Applications: How Michael"" Selected"" Amy* (Vol. 2). Springer Science & Business Media.

Solanki, R. S., Singh, H. P., & Rathour, A. (2012). An Alternative Estimator for Estimating the Finite Population Mean Using Auxiliary Information in Sample Surveys. *ISRN Probability and Statistics*, *2012*, 1–14. https://doi.org/10.5402/2012/657682

Srivastava, S. K., & Jhajj, H. S. (1981). *A class of estimators of the population mean in survey sampling using auxiliary information*.