RATIO ESTIMATOR OF POPULATION VARIANCE USING COEFFICIENT OF VARIATION OF AUXILLIARY VARIABLE IN SIMPLE RANDOM SAMPLING

Saudat Ali Adamu*, Umar Kabir Abdullahi, Tasi'u Musa, Yahaya Zakari, Jamila Abdullahi, Hauwa'u Yusuf and Hafsat Abdulkarim

Department of Statistics, Ahmadu Bello University, Zaria, Nigeria.
*Corresponding author E-mail: saudatali@gmail.com

ABSTRACT

The estimation of population variance is one of the challenging aspects in sampling theory and Population study. Much efforts has been employed to improve the precision of estimates. The existing estimators are still lagging behind in terms of precision. In this study, a ratio estimator of population variance using coefficient of variation of auxiliary variable was developed, under simple random sampling scheme. The bias and MSE of the developed estimator were derived up to the first order approximation, along with its optimal value, δ. Theoretical and empirical comparison of the developed estimator with some other ratio estimators have been established. It was found that the developed estimator is better than other ratio estimators compared with in the study in terms of having minimum MSE of 149646605 and 2898713 and higher PRE of 447.98 and 249.61 under population 1 and 2 respectively, thus most preferred over the existing estimators for the use in practical application.

Keywords: Ratio Estimators, Population Variance, Coefficient of Variation, Bias and Mean Square Error.

1. INTRODUCTION

One of the major objectives in sampling theory is to estimate the parameter of interest with more precision. Statisticians are incorporating the auxiliary information to achieve these objectives and have recognized that use of these information helps to get improvement in the gain of precision. Estimators that are good in this context are Ratio, Regression and Product estimators. In survey sampling, utilizing auxiliary information has been addressed during the estimation process to enhance the accuracy of estimations. The auxiliary information in sampling surveys is used at estimation and selection to obtain more efficient estimators and also improve the design. In a class of efficient estimators, a most efficient is said to have a minimum variance; its description can be expressed by the extent of its bias. An estimator that exhibits the lowest variance is considered superior to other estimators within the same group (Rajesh *et al*, 2011).

Coefficient of Variation (CV) is a statistical measure of dispersion that is used to determine the relative variability of a data set. It is expressed as the standard deviation of the population values in respect to the mean. The role of CV in statistics is to provide a standardized way of comparing the variability of datasets especially when the means of the datasets are different,

Ratio estimator is a statistical method used to estimate a population parameter (such as mean or total) based on two variables in a sample or ratio. Ratio estimators are often used for more accurate and efficient estimates than other sampling methods especially when the auxiliary variable, is strongly correlated with the main variable. The ratio estimator is said to be sensitive to the accuracy of estimates of the auxiliary variable, therefore it is important to ensure that these estimates are reliable before using the ratio estimator.

Variance is a statistical measure of dispersion used to describe the spread or dispersion of a population's datum around its mean. The best method to estimate the population parameter in question is by using the equivalent estimator derived from the sample, which is referred to as a statistic.

2. BACKGROUND OF THE STUDY

Consider a finite population, $U=(U_1,U_2,...,U_N)$, of N units; let a sample be drawn using simple random sampling without replacement (SRSWOR); let Y and X represent the study variable and auxiliary variable respectively. For a sample survey problem, if the interest is to estimate the population variance, S_y^2, when the population and sample coefficient of variation of auxiliary variable is known, one should assume that a sample of size, *n*, is to be selected using simple random sampling scheme. Each unit of the population is identifiable by means of assigning numbers to the population units from 1 to *N*, where the numbers assigned are of nominal scale.

Some notations used in the study are given below;

- N – population size.
- n – sample size.
- $f = n/N$ – sampling fraction.
- Y – study variable.
- X – auxiliary variable.
- $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the mean per unit estimator of *X*.
- $\bar{Y}$, $\bar{X}$ – population means of *Y* and *X*, respectively.
- $\bar{y}$, $\bar{x}$ – sample means of *Y* and *X*, respectively.
- s_y^2, s_x^2 – sample variance of *Y* and *X*, respectively.
- $Bias(.)$ – Bias of the estimator.
- MSE– Mean Squared Error of the estimator.
- PRE– Percentage Relative Efficiency.
- S_Y^2 – Population variance of *Y*.
- S_X^2 – Population variance of *X*.
- ρ – correlation coefficient between, *Y* and *X*.
- C_Y – population coefficient of variation of *Y*.
- C_x – population coefficient of variation of *X*.
- c_x – sample coefficient of variation of *X*.

- l – preliminary large sample size.
- β_1 – coefficient of skewness of X
- β_2 – coefficient of kurtosis of X
- Q_1 – First quartile of the sample
- Q_2 – Second quartile of the sample
- M_d – Median

2.1. Ratio Method of Estimation

Consider the case where the variable of interest Y can be assumed to be approximately proportional to a known auxiliary variable X (i.e. $y = bx$), where b is the constant of proportionality. In such instance, the concept of ratio method of estimation uses the sample to estimate the change in aggregate values of y relative to the aggregate value of x. Isaki (1983) defined ratio-type estimator as;

$$s_I^2 = s_y^2\left(\frac{S_x^2}{s_x^2}\right) . \tag{1}$$

where I means Isaki's estimator

The bias and mean squared error of s_I^2 using first degree of approximation are, respectively, given as;

$$bias\left(s_I^2\right) = \left(\frac{1-f}{n}\right)S_y^2\left\{(\theta_{04}-1)-(\theta_{22}-1)\right\} \tag{2}$$

$$MSE\left(s_I^2\right) = \left(\frac{1-f}{n}\right)S_y^4\left\{(\theta_{40}-1)+(\theta_{04}-1)-2(\theta_{22}-1)\right\} . \tag{3}$$

2.2. Some Existing Estimators Relevant to the Work with their Biases and Mean Squared Errors

Solanki *et al*(2015) developed a ratio type estimator for population variance given as;

$$\hat{S}_{S1}^2 = s_y^2\left[\frac{\beta_{2(y)}S_y^2 + Q_2^2}{\beta_{2(y)}s_y^2 + Q_2^2}\right] \tag{4}$$

Where $\hat{S}_{S1}^2$ is a notation for Solanki's estimator.

The bias and mean squared error of $\hat{S}_{S1}^2$ using first degree of approximation are, respectively, given as;

$$bias = \left(\frac{1-f}{n}\right)S_y^2 R_{S1}\left[R_{S1}(\theta_{04}-1)-(\theta_{22}-1)\right] \tag{5}$$

$$MSE = \left(\frac{1-f}{n}\right)S_y^4\left[(\theta_{40}-1)+R_{S1}^2(\theta_{04}-1)-2R_{S1}(\theta_{22}-1)\right] \tag{6}$$

Kadilar and Cingi (2006) proposed a ratio estimator for population variance as;

$$\hat{S}^2_{KC1} = s^2_y \left[\frac{S^2_x + \beta_{2(x)}}{s^2_x + \beta_{2(x)}} \right] \tag{7}$$

where $\hat{S}^2_{KC1}$ is a notation for Kadilar and Cingi (2006) estimator.

The bias and MSE $\hat{S}^2_{KC1}$ of are respectively given as;

$$bias = \left(\frac{1-f}{n}\right) S^2_y R_{KC1} \left[R_{KC1} \left(\theta_{04} - 1\right) - \left(\theta_{22} - 1\right) \right] \tag{8}$$

$$MSE = \left(\frac{1-f}{n}\right) S^4_y \left[\left(\theta_{40} - 1\right) + R^2_{KC1} \left(\theta_{04} - 1\right) - 2R_{KC1} \left(\theta_{22} - 1\right) \right] \tag{9}$$

where $R_{KC1} = \dfrac{S^2_x}{S^2_x + \beta_{2(x)}}$

Singh *et al*. (2013) also introduced a ratio-type estimator for the population variance, which can be described as follows:

$$\hat{S}^2_{S3} = s^2_y \left[\frac{S^2_x + Q^2_1}{s^2_x + Q^2_1} \right] \tag{10}$$

where S3 is a notation for Singh's estimator.

The bias and MSE of $\hat{S}^2_{S3}$ are respectively given as;

$$bias = \left(\frac{1-f}{n}\right) S^2_y R_{S3} \left[R_{S3} \left(\theta_{04} - 1\right) - \left(\theta_{22} - 1\right) \right] \tag{11}$$

$$MSE = \left(\frac{1-f}{n}\right) S^4_y \left[\left(\theta_{40} - 1\right) + R^2_{S3} \left(\theta_{04} - 1\right) - 2R_{S3} \left(\theta_{22} - 1\right) \right] \tag{12}$$

where $R_{S3} = \dfrac{S^2_x}{S^2_x + Q^2_1}$

The population variance's ratio estimator proposed by Kadilar and Cingi's (2006) is also given by;

$$S^2_{KC2} = s^2_y \left[\frac{S^2_x + C_x}{s^2_x + C_x} \right] \tag{13}$$

where $\hat{S}^2_{KC2}$ is a notation for Kadilar and Cingi's estimator.

The bias and MSE of $\hat{S}^2_{KC2}$ are respectively given as; $bias = \left(\frac{1-f}{n}\right) S^2_y R_{KC2} \left[R_{KC2} \left(\theta_{04} - 1\right) - \left(\theta_{22} - 1\right) \right]$

(14)

$$MSE = \left(\frac{1-f}{n}\right) S^4_y \left[\left(\theta_{40} - 1\right) + R^2_{KC2} \left(\theta_{04} - 1\right) - 2R_{KC2} \left(\theta_{22} - 1\right) \right] \tag{15}$$

where $R_{KC2} = \frac{S_x^2}{S_x^2 + C_x}$

Another ratio estimator for population variance was proposed by Subramani and Kumarapandiyan (2012) which is defined;

$$\hat{S}_{SK}^2 = s_y^2 \left[\frac{S_x^2 + M_d}{s_x^2 + M_d} \right] \tag{16}$$

where $\hat{S}_{SK}^2$ is a notation for Subramani and Kumarapandiyan (2012) estimator.

The bias and MSE of $\hat{S}_{SK}^2$ are respectively given as;

$$bias = \left(\frac{1-f}{n} \right) S_y^2 R_{SK} \left[R_{SK} \left(\theta_{04} - 1 \right) - \left(\theta_{22} - 1 \right) \right] \tag{17}$$

$$MSE = \left(\frac{1-f}{n} \right) S_y^4 \left[\left(\theta_{40} - 1 \right) + R_{SK}^2 \left(\theta_{04} - 1 \right) - 2R_{SK} \left(\theta_{22} - 1 \right) \right] \tag{18}$$

where, $R_{SK} = \frac{S_x^2}{S_x^2 + M_d}$

Also, Solanki *et al.* (2012) proposed an estimator for the population variance which is defined as;

$$S_{S(\alpha,\delta)}^2 = s_y^2 \left[2 - \left(\frac{s_x^2}{S_x^2} \right)^{\alpha} \exp\left\{ \frac{\delta \left(s_x^2 - S_x^2 \right)}{\left(s_x^2 + S_x^2 \right)} \right\} \right]. \tag{19}$$

The bias and MSE of $\hat{S}_{(\alpha,\delta)}^2$ are respectively given as;

$$bias = \left(\frac{1-f}{n} \right) S_y^2 \left[-\frac{\left(2\alpha - \delta \right)}{2} \left(\theta_{22} - 1 \right) - \frac{\left(2\alpha + \delta \right)\left(2\alpha + \delta - 2 \right)}{8} \left(\theta_{04} - 1 \right) \right] \tag{20}$$

$$MSE = \left(\frac{1-f}{n} \right) S_y^4 \left[\left(\theta_{40} - 1 \right) + \frac{\left(2\alpha + \delta \right)^2}{4} \left(\theta_{04} - 1 \right) - (2\alpha + \delta)\left(\theta_{22} - 1 \right) \right] \tag{21}$$

where; $\left(\alpha, \delta \right)$ are suitable chosen scalars.

3. METHODS OF EFFICIENCY COMPARISON OF THE PROPOSED ESTIMATORS WITH SOME EXISTING ESTIMATORS

3.1. Methods for Analytical Study

The study adopts analytical method. Mathematical and statistical operations involving parameters in linear regression and ratio methods were used to develop the estimators for population variance. The concepts of first order approximation, expectation and method of partial differential equation were used to derive the bias and Mean Square Error (MSE) of the proposed estimators. The concept of inequality was also used to obtain the efficiency condition of the proposed estimators.

A new estimator say s_{Pi}^2 will be more efficient than any estimator say s_T^2 if,

$$\text{MSE}\left(S_{pi}^2\right) < \text{MSE}\left(S_T^2\right) \tag{22}$$

3.2. Methods for Empirical Study

To empirically evaluate the efficiency of the new parameter-based estimators, two life data sets are used and simulation studies was also carried out. R statistical package (version 3.5.3) was employed in analyzing the data.

The Percentage Relative Efficiency (PRE) was employed as performance measure for determining the best among the competing estimators.

The PRE of different estimators, S_T^2 with respect to S_I^2 is defined as,

$$\mathrm{PRE}\left(S_I^2, S_T^2\right) = \frac{MSE(\mathrm{S}_I^2)}{MSE(\mathrm{S}_T^2)} \times 100 \tag{23}$$

3.3. Proposed Coefficient of Variation Based Ratio Estimator

Motivated by Srivastava and Jhajj (1981) variance based estimator, we define the general form for coefficient of variation based estimator of population variance using Simple Random Sampling scheme as

$$\hat{S}_{CV}^2 = s_y^2 \left[\frac{C_x}{c_x}\right]^{\delta} \tag{24}$$

where, s_y is sample mean of the study variable, C_x is population coefficient of variation of the auxiliary variable, c_x is the sample coefficient of variation of the auxiliary variable and δ is a real value chosen to minimize the MSE of $\hat{S}_{CV}^2$.

From equation (24), if δ is positive, the estimator is in the form of ratio and can be expressed as;

$$\hat{S}_{CVR}^2 = s_y^2 \left(\frac{C_x}{c_x}\right)^{\delta} = s_y^2 \left(\frac{\bar{x}}{\bar{X}}\right)^{\delta} \left(\frac{S_x^2}{s_x^2}\right)^{-\frac{\delta}{2}}$$

3.4. Bias and Mean Square Error of the Proposed Ratio Estimator

To obtain the bias and Mean Square Error (MSE) of the proposed estimator, let

$$\Omega_{sx} = \frac{\left(s_x^2 - S_x^2\right)}{S_x^2},\ \Omega_{sy} = \frac{\left(s_y^2 - S_y^2\right)}{S_y^2} \text{ and } \Omega_{\bar{x}} = \frac{\left(\bar{x} - \bar{X}\right)}{\bar{X}}$$

Such that $E\left(\Omega_{sx}\right) = E\left(\Omega_{sy}\right) = E\left(\Omega_{\bar{x}}\right) = 0$.

Then, $s_x^2 = S_x^2\left(1+\Omega_{sx}\right)$ and $s_y^2 = S_y^2\left(1+\Omega_{sy}\right)$,

$$E\left(\Omega_{sx}^2\right) = \frac{Var\left(s_x^2\right)}{\left(S_x^2\right)^2} = \left(\frac{1-f}{n}\right)\left(\theta_{04}-1\right) = C_{ssx}^{!},\ E\left(\Omega_{sy}^2\right) = \frac{Var\left(s_x^2\right)}{\left(S_x^2\right)^2} = \left(\frac{1-f}{n}\right)\left(\theta_{40}-1\right) = C_{ssy}^{!} \text{ and}$$

$$E\left(\Omega_{sx}\Omega_{sy}\right) = \frac{1-f}{n}\left(\theta_{22}-1\right)$$

where, $E\left(\Omega_{\bar{y}}\Omega_{sy}\right) = \left(\frac{1-f}{n}\right) C_y \theta_{12} = C_{ys}^{!}$ and $E\left(\Omega_{\bar{x}}\Omega_{sx}\right) = \left(\frac{1-f}{n}\right) C_x \theta_{03} = C_{xs}^{!}$

where, $\alpha_{lm} = (N-1)^{-1} \sum_{i=1}^{N} \left(X_i - \bar{X}\right)^m \left(Y_i - \bar{Y}\right)^l$, $\beta_{lm} = (N-1)^{-1} \sum_{i=1}^{N} \left(s_x^2 - S_x^2\right)^l \left(s_y^2 - S_y^2\right)^m$

And $\theta_{lm}=\dfrac{\alpha_{lm}}{\alpha_{20}^{\frac{l}{2}}\alpha_{02}^{\frac{m}{2}}}$

$E\left(\Omega_{\bar{y}}^{2}\right)=\dfrac{Var\left(\bar{y}\right)}{\bar{Y}^{2}}=\left(\dfrac{1-f}{n}\right)C_{y}^{2}=C_{yy}^{!}$ and $E\left(\Omega_{\bar{x}}^{2}\right)=\dfrac{Var\left(\bar{x}\right)}{\bar{X}^{2}}=\left(\dfrac{1-f}{n}\right)C_{x}^{2}=C_{xx}^{!}$

$E\left(\Omega_{\bar{x}}\Omega_{\bar{y}}\right)=\dfrac{Cov\left(\bar{x},\bar{y}\right)}{\bar{X}\bar{Y}}=\left(\dfrac{1-f}{n}\right)\rho C_{x}C_{y}=C_{sy}^{!}$

$\left(1+\Omega_{\bar{x}}\right)=1-\alpha\Omega_{\bar{x}}+\alpha\left(\alpha+1\right)\dfrac{\Omega_{\bar{x}}^{2}}{2}$

Expanding in terms of $'\Omega$'s, we have

$$\hat{S}_{CVR}^{2}=S_{y}^{2}\left(1+\Omega_{sy}\right)\left(1+\Omega_{\bar{x}}\right)^{\delta}\left(1+\Omega_{sx}\right)^{-\frac{\delta}{2}} \tag{25}$$

Note that;

$$\left(1+\Omega_{s}\right)^{-\frac{\delta}{2}}=1-\frac{\delta\Omega_{s}}{2}+\frac{\delta\left(\delta-2\right)\Omega_{s}^{2}}{8}\ldots$$

$$\left(1+\Omega_{\bar{x}}\right)^{\delta}=1+\delta\Omega_{\bar{x}}+\frac{\delta\left(\delta+1\right)\Omega_{\bar{x}}^{2}}{2}\ldots$$

We further assume that the contribution of terms involving powers in Ω_{sx},Ω_{sy} and $\Omega_{\bar{x}}$ higher than the second is negligible. Thus, from the above expression we write to the first order of approximation. From (25), we have;

$$\hat{S}_{CVR}^{2}=S_{y}^{2}\left(1+\Omega_{sy}\right)\left(1+\delta\Omega_{\bar{x}}+\frac{\delta\left(\delta-1\right)\Omega_{\bar{x}}^{2}}{2}\right)\left(1-\frac{\delta\Omega_{sx}}{2}+\frac{\delta\left(\delta+2\right)\Omega_{sx}^{2}}{8}\right) \tag{26}$$

$$=S_{y}^{2}\left(1+\delta\Omega_{\bar{x}}+\frac{\delta\left(\delta-1\right)\Omega_{\bar{x}}^{2}}{2}+\Omega_{sy}+\delta\Omega_{sy}\Omega_{\bar{x}}+\frac{\delta\left(\delta-1\right)\Omega_{sy}\Omega_{\bar{x}}^{2}}{2}\right)\left(1-\frac{\delta\Omega_{sx}}{2}+\frac{\delta\left(\delta+2\right)\Omega_{sx}^{2}}{8}\right) \tag{27}$$

$$=S_{y}^{2}\left(\begin{array}{l}\left(1+\delta\Omega_{\bar{x}}+\dfrac{\delta\left(\delta-1\right)\Omega_{\bar{x}}^{2}}{2}+\Omega_{sy}+\delta\Omega_{sy}\Omega_{\bar{x}}+\dfrac{\delta\left(\delta-1\right)\Omega_{sy}\Omega_{\bar{x}}^{2}}{2}\right)\\ -\left(\dfrac{\delta\Omega_{sx}}{2}+\delta\Omega_{\bar{x}}\dfrac{\delta\Omega_{sx}}{2}+\dfrac{\delta\Omega_{sx}}{2}\dfrac{\delta\left(\delta-1\right)\Omega_{\bar{x}}^{2}}{2}+\dfrac{\delta\Omega_{sx}}{2}\Omega_{sy}+\dfrac{\delta\Omega_{sx}}{2}\delta\Omega_{sy}\Omega_{\bar{x}}+\dfrac{\delta\Omega_{sx}}{2}\dfrac{\delta\left(\delta-1\right)\Omega_{sy}\Omega_{\bar{x}}^{2}}{2}\right)\\ +\left(\begin{array}{l}\dfrac{\delta\left(\delta+2\right)\Omega_{sx}^{2}}{8}+\dfrac{\delta\left(\delta+2\right)\Omega_{sx}^{2}}{8}\delta\Omega_{\bar{x}}+\dfrac{\delta\left(\delta+2\right)\Omega_{sx}^{2}}{8}\dfrac{\delta\left(\delta-1\right)\Omega_{\bar{x}}^{2}}{2}+\dfrac{\delta\left(\delta+2\right)\Omega_{sx}^{2}}{8}\Omega_{sy}+\\ \dfrac{\delta\left(\delta+2\right)\Omega_{sx}^{2}}{8}\delta\Omega_{sy}\Omega_{\bar{x}}+\dfrac{\delta\left(\delta+2\right)\Omega_{sx}^{2}}{8}\dfrac{\delta\left(\delta-1\right)\Omega_{sy}\Omega_{\bar{x}}^{2}}{2}\end{array}\right)\end{array}\right)$$

$$E\left(\hat{S}_{CVR}^{2}-S_{y}^{2}\right)=S_{y}^{2}E\left(\frac{\delta\left(\delta-1\right)\Omega_{\bar{x}}^{2}}{2}+\delta\Omega_{sy}\Omega_{\bar{x}}-\frac{\delta^{2}\Omega_{\bar{x}}\Omega_{sx}}{2}-\frac{\delta\Omega_{sx}\Omega_{sy}}{2}+\frac{\delta\left(\delta+2\right)\Omega_{sx}^{2}}{8}\right) \tag{28}$$

$$Bias\left(\hat{S}_{CVR}^{2}\right)=S_{y}^{2}\left(\frac{1-f}{n}\right)\left(\frac{\delta(\delta-1)}{2}C_{x}^{2}+\delta\rho C_{x}\theta_{30}-\frac{\delta^{2}C_{x}\theta_{03}}{2}-\frac{\delta\left(\theta_{22}-1\right)}{2}+\frac{\delta(\delta+2)\left(\theta_{04}-1\right)}{8}\right) \tag{29}$$

Taking only the leading terms from (26), the first order approximation to the MSE of the proposed estimator is given as,

$$\left(\hat{S}_{CVR}^{2}-S_{y}^{2}\right)^{2}=\left(S_{y}^{2}\right)^{2}\left(\Omega_{sy}+\delta\Omega_{\bar{x}}-\frac{1}{2}\delta\Omega_{sx}\right)^{2} \tag{30}$$

$$=\left(S_{y}^{2}\right)^{2}\left\{\left(\Omega_{sy}^{2}+\delta^{2}\Omega_{\bar{x}}^{2}+\frac{\delta^{2}\Omega_{sx}^{2}}{4}+2\delta\Omega_{sy}\Omega_{\bar{x}}-\delta^{2}\Omega_{sx}\Omega_{\bar{x}}-\delta\Omega_{sx}\Omega_{sy}\right)\right\} \tag{31}$$

$$MSE\left(\hat{S}_{CVR}^{2}\right)=S_{y}^{4}\left(\frac{1-f}{n}\right)\left\{\left(\left(\theta_{40}-1\right)+\delta^{2}C_{x}^{2}+\frac{\delta^{2}\left(\theta_{04}-1\right)}{4}+2\delta\rho C_{x}\theta_{30}-\delta^{2}C_{x}\theta_{03}-\delta\left(\theta_{22}-1\right)\right)\right\} \tag{32}$$

$$=S_{y}^{4}\left(\frac{1-f}{n}\right)\left(\left(\theta_{40-1}\right)+\delta^{2}\left(C_{x}^{2}+\frac{\left(\theta_{04}-1\right)}{4}-C_{x}\theta_{03}\right)+\delta\left(2\rho C_{x}\theta_{30}-\left(\theta_{22}-1\right)\right)\right) \tag{33}$$

3.5. Optimal MSE of the Proposed Coefficient of Variation based Ratio Estimator

To obtain the optimal MSE of $\hat{S}_{CVR}^{2}$, we take the partial derivative of MSE ($\hat{S}_{CVR}^{2}$) with respect to δ, and equate to 0.

$$\frac{\partial MSE\left(\hat{S}_{CVR}^{2}\right)}{\partial\delta}\left[S_{y}^{4}\left(\frac{1-f}{n}\right)\left\{\left(\theta_{40}-1\right)+\delta^{2}C_{x}^{2}+\frac{\delta^{2}\left(\theta_{04}-1\right)}{4}+2\delta\rho C_{x}\theta_{30}-\delta^{2}C_{x}\theta_{03}-\delta\left(\theta_{22}-1\right)\right\}\right]=0 \tag{33}$$

$$2\delta C_{x}^{2}+\frac{\delta\left(\theta_{04}-1\right)}{2}+2\rho C_{x}\theta_{30}-2\delta C_{x}\theta_{30}-\left(\theta_{22}-1\right)=0 \tag{35}$$

$$4\delta C_{x}^{2}+\delta\left(\theta_{04}-1\right)-4\delta C_{x}\theta_{03}=\left(\theta_{22}-1\right)-2\rho C_{x}\theta_{30} \tag{36}$$

$$\delta=\frac{\theta_{22}-2\rho C_{x}\theta_{30}-1}{4C_{x}^{2}+\theta_{04}-4C_{x}\theta_{03}-1} \tag{37}$$

Substituting (37) in (33),

$$MSE\left(\hat{S}_{CVR}^{2}\right)_{opt}=S_{y}^{4}\left(\frac{1-f}{n}\right)\left(\begin{array}{l}\left(\theta_{40}-1\right)+\left(\dfrac{\theta_{22}-2\rho C_{x}\theta_{30}-1}{4C_{x}^{2}+\theta_{04}-4C_{x}\theta_{03}-1}\right)^{2}\left(C_{x}^{2}+\dfrac{\left(\theta_{04}-1\right)}{4}-C_{x}\theta_{03}\right)\\+\left(\dfrac{\theta_{22}-2\rho C_{x}\theta_{30}-1}{4C_{x}^{2}+\theta_{04}-4C_{x}\theta_{03}-1}\right)\left(2\rho C_{x}\theta_{30}-\left(\theta_{22}-1\right)\right)\end{array}\right) \tag{38}$$

$$=S_{y}^{4}\left(\frac{1-f}{n}\right)\left(\begin{array}{l}\dfrac{\left(\theta_{40}-1\right)\left(4C_{x}^{2}+\theta_{04}-4C_{x}\theta_{03}-1\right)^{2}}{\left(4C_{x}^{2}+\theta_{04}-4C_{x}\theta_{03}-1\right)^{2}}+\dfrac{\left(\theta_{22}-2\rho C_{x}\theta_{30}-1\right)^{2}\left(4C_{x}^{2}+\theta_{04}-4C_{x}\theta_{03}-1\right)}{\left(4C_{x}^{2}+\theta_{04}-4C_{x}\theta_{03}-1\right)^{2}}\\+\dfrac{\left(\theta_{22}-2\rho C_{x}\theta_{30}-1\right)\left(2\rho C_{x}\theta_{30}-\theta_{22}-1\right)\left(4C_{x}^{2}+\theta_{04}-4C_{x}\theta_{03}-1\right)}{\left(4C_{x}^{2}+\theta_{04}-4C_{x}\theta_{03}-1\right)^{2}}\end{array}\right) \tag{39}$$

$$=S_y^4\left(\frac{1-f}{n}\right)\left(\begin{array}{l}(\theta_{40}-1)+\left(\dfrac{\theta_{22}^2-4\rho C_x\theta_{30}\theta_{22}-2\theta_{22}+4\rho^2C_x^2\theta_{30}^2+4\rho C_x\theta_{30}+1}{\left(4C_x^2+\theta_{04}-4C_x\theta_{03}-1\right)}\right)\\+\left(\dfrac{4\rho C_x\theta_{30}\theta_{22}-\theta_{22}^2-4\rho^2C_x^2\theta_{30}^2+1}{\left(4C_x^2+\theta_{04}-4C_x\theta_{03}-1\right)}\right)\end{array}\right) \tag{40}$$

$$=S_y^4\left(\frac{1-f}{n}\right)\left(\left(\theta_{40}-1\right)+\frac{4\rho C_x\theta_{30}-2\theta_{22}+2}{4C_x^2+\theta_{04}-4C_x\theta_{03}-1}\right) \tag{41}$$

3.6. The efficiency condition of the proposed ratio estimator to be better than existing estimators considered

The proposed estimator is more efficient than existing estimators considered if;

1) The MSE of the proposed estimator $MSE\left(\hat{S}_{CVR}^2\right)$ is less than the MSE of the Isaki (1983) ratio estimator $MSE\left(\hat{S}_I^2\right)$ in equation (3) i.e

$\bullet MSE\left(\hat{S}_{CVR}^2\right)\leq MSE\left(\hat{S}_I^2\right)$

$$\Rightarrow\frac{1-f}{n}S_y^4\left[\left(\theta_{40}-1\right)+\delta\left(C_x^2+\frac{\left(\theta_{04}-1\right)}{4}-C_x\theta_{03}\right)+\delta\left(2\rho C_x\theta_{30}-\left(\theta_{22}-1\right)\right)\right]\leq\frac{1-f}{n}S_y^4\left\{\left(\theta_{04}-1\right)+\left(\theta_{04}-1\right)-2\left(\theta_{22}-1\right)\right\}$$

$$MSE\left(\hat{S}_I^2\right)-MSE\left(\hat{S}_{CVR}^2\right)=\frac{1-f}{n}S_y^4\left[\delta\left(C_x^2-C_x\theta_{03}\right)+2\delta\rho C_x\theta_{30}+\frac{(\delta-4)\left(\theta_{04}-1\right)}{4}-\left(\delta+2\right)\left(\theta_{22}-1\right)\right]>0$$

2) The MSE of the proposed estimator $MSE\left(\hat{S}_{CVR}^2\right)$ is less than the MSE of the Kadilar and Cingi (2006) ratio estimator $MSE\left(\hat{S}_{KC1}^2\right)$ in equation (6) i.e

$\bullet MSE\left(\hat{S}_{CVR}^2\right)\leq MSE\left(\hat{S}_{KC1}^2\right)$

$$\Rightarrow\frac{1-f}{n}S_y^4\left[\left(\theta_{40}-1\right)+\delta^2\left(C_x^2+\frac{\left(\theta_{04}-1\right)}{4}-C_x\theta_{03}\right)+\delta\left(2\rho C_x\theta_{30}-\left(\theta_{22}-1\right)\right)\right]\leq\frac{1-f}{n}S_y^4\left\{\left(\theta_{04}-1\right)+R_{KC1}^2\left(\theta_{04}-1\right)-2R_{KC1}\left(\theta_{22}-1\right)\right\}$$

$$MSE\left(\hat{S}_{KC1}^2\right)-MSE\left(\hat{S}_{CVR}^2\right)=\frac{1-f}{n}S_y^4\left[\delta\left(C_x^2-C_x\theta_{03}\right)+2\delta\rho C_x\theta_{30}+\frac{\left(\delta-4R_{KC1}\right)\left(\theta_{04}-1\right)}{4}-\left(\delta+2R_{KC1}\right)\left(\theta_{22}-1\right)\right]>0$$

3) The MSE of the proposed estimator $MSE\left(\hat{S}_{CVR}^2\right)$ is less than the MSE of the Kadilar and Cingi (2006) ratio estimator $MSE\left(\hat{S}_{KC2}^2\right)$ in equation (15) i.e

$\bullet MSE\left(\hat{S}^2_{CVR}\right) \le MSE\left(\hat{S}^2_{KC2}\right)$

$$\Rightarrow \frac{1-f}{n} S_y^4 \left[\left(\theta_{40}-1\right) + \delta^2 \left(C_x^2 + \frac{\left(\theta_{04}-1\right)}{4} - C_x\theta_{03} \right) + \delta\left(2\rho C_x\theta_{30} - \left(\theta_{22}-1\right)\right) \right] \le \frac{1-f}{n} S_y^4 \left\{ \left(\theta_{04}-1\right) + R_{KC2}^2\left(\theta_{04}-1\right) - 2R_{KC2}\left(\theta_{22}-1\right) \right\}$$

$$MSE\left(\hat{S}^2_{KC2}\right) - MSE\left(\hat{S}^2_{CVR}\right) = \frac{1-f}{n} S_y^4 \left[\delta\left(C_x^2 - C_x\theta_{03}\right) + 2\delta\rho C_x\theta_{30} + \frac{\left(\delta - 4R_{KC2}\right)\left(\theta_{04}-1\right)}{4} - \left(\delta + 2R_{KC2}\right)\left(\theta_{22}-1\right) \right] > 0$$

4) The MSE of the proposed estimator $MSE\left(\hat{S}^2_{CVR}\right)$ is less than the MSE of the Subramani and Kumarapandiyan (2012) ratio estimator $MSE\left(\hat{S}^2_{SK}\right)$ in equation (18) i.e

$\bullet MSE\left(\hat{S}^2_{CVR}\right) \le MSE\left(\hat{S}^2_{SK}\right)$

$$\Rightarrow \frac{1-f}{n} S_y^4 \left[\left(\theta_{40}-1\right) + \delta^2 \left(C_x^2 + \frac{\left(\theta_{04}-1\right)}{4} - C_x\theta_{03} \right) + \delta\left(2\rho C_x\theta_{30} - \left(\theta_{22}-1\right)\right) \right] \le \frac{1-f}{n} S_y^4 \left\{ \left(\theta_{04}-1\right) + R_{SK}^2\left(\theta_{04}-1\right) - 2R_{SK}\left(\theta_{22}-1\right) \right\}$$

$$MSE\left(\hat{S}^2_{SK}\right) - MSE\left(\hat{S}^2_{CVR}\right) = \frac{1-f}{n} S_y^4 \left[\delta\left(C_x^2 - C_x\theta_{03}\right) + 2\delta\rho C_x\theta_{30} + \frac{\left(\delta - 4R_{SK}\right)\left(\theta_{04}-1\right)}{4} - \left(\delta + 2R_{SK}\right)\left(\theta_{22}-1\right) \right] > 0$$

3.7. Assumptions of the proposed Estimator

i. There is always need to ensure that the auxiliary variable is highly correlated with the study variable.
ii. The population under consideration have to be homogeneously distributed.
simulations were replicated 10000 times and averaged over.

4. DATA PRESENTATION

Dataset: The populations considered in this study are natural population datasets. The population is taken from Subramani and Kumarapandiyan (2012) given in page 150. Table 1 below summarized the dataset for the study.

Table 1: Summary of the Dataset

Parameters	Population 1:	Population 2:
N	103	49
n	40	20
$\bar{Y}$	626.2123	116.1633
$\bar{X}$	557.1909	98.6765
ρ	0.9936	0.6904
C_y	1.4588	0.8508
C_x	1.4683	1.0435
S_x	81.8111	102.9709
S_y	91.3549	98.8286
θ_{04}	37.3216	5.9878
θ_{40}	37.1279	4.9245
θ_{03}	17.4683	1.5200
θ_{30}	16.9128	1.1128
θ_{22}	37.2055	4.6977
M_d	308.0500	64.0000
δ	0.2280	0.6984
γ	0.7959	0.6267
λ_1	1.4856	1.2692

Therefore, we see the merit of the proposed estimator $\hat{S}^2_{CVR}$ over some existing Ratio estimators $\hat{S}^2_I$ (Isaki's ratio estimator), $\hat{S}^2_{KC1}$ (kadilar and Cingi's ratio estimator), $\hat{S}^2_{KC2}$ (Kadilar and Cingi's ratio estimator) and $\hat{S}^2_{SK}$ (Subramani and Kumarapandiyan's ratio estimator) for simple random sampling in Table 2.

Table 2: The variance and MSE of $\hat{S}^2_I$, $\hat{S}^2_{KC1}$, $\hat{S}^2_{KC2}$ and $\hat{S}^2_{SK}$ and the proposed estimator $\hat{S}^2_{CVR}$

		Populations	
		1	2
Estimators	$\hat{S}^2_I$	670393270	7235508
	$\hat{S}^2_{KC1}$	670169790	7228570
	$\hat{S}^2_{KC2}$	670384403	7234298
	$\hat{S}^2_{SK}$	668667061	7228859
	$\hat{S}^2_{CVR}$	149646605	2898713

Table 3: Percentage relative efficiency of different estimators with respect to $\hat{S}_I^2$

		Populations	
		1	2
Estimators	$\hat{S}_I^2$	100	100
	$\hat{S}_{KC1}^2$	100.0334	100.0950
	$\hat{S}_{KC2}^2$	100.0013	101.0167
	$\hat{S}_{SK}^2$	100.2582	100.0920
	$\hat{S}_{CVR}^2$	447.9843	249.6111

5. DISCUSSION OF RESULT

The main focus of this study was to propose more efficient estimators based on simple random sampling. We have proposed a ratio estimator and obtained its asymptotically optimum estimator (AOE) with its approximate MSE formula for the proposed estimators using the coefficient of variation of the auxiliary variable X in simple random sampling.

Theoretically, we have demonstrated that the proposed estimator was more efficient than other estimators, $\hat{S}_I^2$, $\hat{S}_{KC1}^2$, $\hat{S}_{KC2}^2$ and $\hat{S}_{SK}^2$ under their optimum value.

In addition, we support these theoretical results numerically using the data sets as shown in Table 1. Table 2 provides the MSE of the proposed estimator ($\hat{S}_{CVR}^2$) and some existing estimators, $\hat{S}_I^2$, $\hat{S}_{KC1}^2$, $\hat{S}_{KC2}^2$ and $\hat{S}_{SK}^2$ as far as mean squared error criterion is concerned. Table 3 provides that there is a considerable gain in efficiency by using proposed estimator $\hat{S}_{CVR}^2$ over the estimators $\hat{S}_I^2$, $\hat{S}_{KC1}^2$, $\hat{S}_{KC2}^2$ and $\hat{S}_{SK}^2$. This shows that even if the scalar δ deviates from its optimum values, the suggested estimator $\hat{S}_{CVR}^2$ will yield better esticmates than, $\hat{S}_I^2$, $\hat{S}_{KC1}^2$, $\hat{S}_{KC2}^2$ and $\hat{S}_{SK}^2$.

5.1. Summary

The study incorporated in this dissertation is aimed to propose more efficient estimators of population variance of the auxiliary variable using coefficient of variation of the auxiliary variable and to explore the need for considering coefficient of variation of the auxiliary variable in ratio method of estimation of the auxiliary variable in sampling theory and population study. We obtained the bias and mean square error of the proposed estimator and also obtained the optimum value of the proposed estimator.

The study considered two natural populations in simple random sampling. The auxiliary variable and study variable in simple random sampling are respectively given as in population (1) The total number of recyclable-waste collection in Italy in 2002 and total amount (tons) of recyclable-waste collection in Italy in 2003 and population (2) is taken from Cochran (1977) given in page 152.

It was observed from the analysis that in simple random sampling scheme, the proposed estimator have the minimum MSE compared to some ratio estimators in existence, and the proposed estimator attain its minimum MSE at its optimum value.

6. CONCLUSION

The main focus of this study was to propose more efficient estimators based on simple random sampling. Evidence from the study revealed that the proposed estimator was more efficient than the already existing ratio type estimators based on some certain efficiency conditions.
Hence, we conclude that the proposed estimator was more efficient than the other estimators in case of their optimality in simple random sampling. Thus, it is preferred to use the proposed estimator in practice over the existing estimators considered. The proposed estimator is more efficient than the other existing estimators in terms of its optimality.

REFERENCES

Abdullahi, U. K., & Yahaya, A. (2017). A modified Ratio-Product Estimator of the Population Mean using some known Parameters of the Auxiliary Variable. *Bayero Journal of Pure Applied Science*, **10**(1):128-137

Housila, P.S. and Neha-Agnihotri (2008). A general procedure of estimating population mean using auxiliary information in sample surveys. *Statistics in Transition-new series*, **9**(1):71-87.

Isaki, C.T. (1983): Variance Estimation using Auxiliary Information. *Journal of American Statistical Association* **78**(381):117–123.

Kadilar, C. & Cingi, H. (2006). Improvement in Variance Estimation using Auxiliary Information. *Hacettepe Journal of Mathematics and Statistics*, **35**(1): 111-115

Rajesh, T., Rajesh, P. & Jong-Mind, K. (2011). Ratio-cum-product Estimators of Population Mean using known Parameters of Auxiliary Variables. *Communications of the Korean Statistical Society*, **18**(2):155– 164.

Singh, H.P., Pal, S.K. & Solanki, R.S. (2013). Improved Estimation of Finite Population Variance using Quartiles, *ISTATISTIK* **6** (3), 116-121.

Solanki, R. S., Singh, H. P. & Pal, S. K. (2015). Improved Ratio-type Estimators of Finite Population Variance using Quartiles. *Hacettepe Journal of Mathematics and Statistics* **44**(3):747-754.

Solanki, R.S., Singh, H. P. & Rathour, A. (2012). An Alternative Estimator for Estimating the Finite Population Mean using Auxiliary Information in Sample Surveys. *International Scholarly Research Network ISRN Probability and Statistics*, 1-14.

Subramani, J. and Kumarapandiyan, G. (2012).Estimation of population mean using co-efficient of variation and median of the auxiliary variable. *International Journal of Probability Statistics*, **1**(4): 111-118, DOI 10.5923/j. ijps.20120104.04.