

# Estimation of Population Variance Using the Coefficient of Kurtosis and Median of an Auxiliary Variable under Simple Random Sampling

Tonui Kiplangat Milton<sup>1\*</sup>, Romanus Otieno Odhiambo<sup>2</sup>, George Otieno Orwa<sup>2</sup>

<sup>1</sup>Pan African University, Institute for Basic Sciences, Technology and Innovation (PAUISTI), Nairobi, Kenya

<sup>2</sup>Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Email: \*tonuimilton@gmail.com

**How to cite this paper:** Milton, T.K., Odhiambo, R.O. and Orwa, G.O. (2017) Estimation of Population Variance Using the Coefficient of Kurtosis and Median of an Auxiliary Variable under Simple Random Sampling. *Open Journal of Statistics*, 7, 944-955.

<https://doi.org/10.4236/ojs.2017.76066>

**Received:** September 24, 2017

**Accepted:** November 29, 2017

**Published:** December 2, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In this study we have proposed a modified ratio type estimator for population variance of the study variable  $y$  under simple random sampling without replacement making use of coefficient of kurtosis and median of an auxiliary variable  $x$ . The estimator's properties have been derived up to first order of Taylor's series expansion. The efficiency conditions derived theoretically under which the proposed estimator performs better than existing estimators. Empirical studies have been done using real populations to demonstrate the performance of the developed estimator in comparison with the existing estimators. The proposed estimator as illustrated by the empirical studies performs better than the existing estimators under some specified conditions *i.e.* it has the smallest Mean Squared Error and the highest Percentage Relative Efficiency. The developed estimator therefore is suitable to be applied to situations in which the variable of interest has a positive correlation with the auxiliary variable.

## Keywords

Modified Ratio Type Variance Estimator, Study Variable, Auxiliary Variable, Kurtosis, Median, Bias, Mean Squared Error (MSE), Percentage Relative Efficiency (PRE), Simple Random Sampling

---

## 1. Introduction

It is notable that the appropriate use of auxiliary information in probability sampling designs yields considerable reduction in the variance of the estimators

of population parameters namely, population mean, median, variance, regression coefficient and population correlation coefficient. [1] was the first to show the contribution of known auxiliary information in improving the efficiency of the estimator of the population mean  $\bar{Y}$  in survey sampling.

Survey samplings now touch almost every field of scientific study, including demography, education, energy, transportation, health care, economics, forestry, sociology, politics and so on. In fact it is not an exaggeration to say that much of the data that are statistically analyzed are collected in surveys. It is imperative to note that as the demand in use of surveys increase, the need for more effective methods of analyzing and interpreting the resulting data is inevitable. Measure of precision being a prime requirement of good surveys and appear now in most analysis hence the need to be obtained for almost each estimate derived from the survey data.

On regular instances we encounter surveys in which an auxiliary variable  $x$  is relatively cheap (with regard to time and money) to monitor than the study variable  $y$ . Use of auxiliary information can increase the precision of an estimator when the study variable  $y$  is highly correlated with auxiliary variable  $x$ . In reality such situations do occur when information is available in the form of auxiliary variable, which is highly correlated with study variable, for example, number of trees in an orchard and the yield of fruits.

The most common and widely used measure of precision is the variance of the survey estimator. In reality population variances are always not known but must be estimated from the survey data themselves. In this study we are interested in the estimation of population variance using known auxiliary information under simple random sampling without replacement (SRSWOR) sampling scheme. The precision of estimators under this situation is always increased, the ratio, product and regression estimators gives better outcome than those of simple random samplings.

Consider a finite population  $V = \{V_1, V_2, V_3, \dots, V_N\}$  of  $N$  distinct identifiable units. Let  $Y$  be our study variable and  $X$  be its corresponding auxiliary variable. Suppose we take a random sample of size  $n$  from this bivariate population  $(Y, X)$  that is  $(y_i, x_i)$ , for  $i = 1, 2, 3, \dots, n$  using a Simple Random Sampling Without Replacement (SRSWOR) method. Let  $\bar{Y}$  and  $\bar{X}$  be the population means of the study and auxiliary variable respectively and their corresponding sample means be  $\bar{y}$  and  $\bar{x}$ .

This study focuses on improving the efficiency in the estimation of

$$S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2 \quad (1)$$

using the coefficient of kurtosis and median.

We define the following notations that we will use throughout the article. For the population observations we have;

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i, \quad \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i, \quad S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2,$$

$$S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2, \quad S_{xy} = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})(X_i - \bar{X}).$$

Also we define the following from the sample observations:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2,$$

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x}).$$

In general, we define the following parameters:

$$\mu_{rs} = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^r (x_i - \bar{x})^s \tag{2}$$

$$\lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{\frac{r}{2}} \mu_{02}^{\frac{s}{2}}} \tag{3}$$

Thus we note the following;

$$\mu_{20} = S_y^2, \quad \mu_{02} = S_x^2, \quad \text{and} \quad \mu_{11} = S_{xy}; \quad \lambda_{22} = \frac{\mu_{22}}{\mu_{20}\mu_{02}}, \quad \lambda_{21} = \frac{\mu_{21}}{\mu_{20}^{\frac{1}{2}}}$$

$C_y = \frac{S_y^2}{\bar{Y}^2} = \frac{\mu_{20}}{\bar{Y}^2}$  is the coefficient of variation for the study variable  $y$ ,

$C_x = \frac{S_x^2}{\bar{X}^2} = \frac{\mu_{02}}{\bar{X}^2}$  is the coefficient of variation for the auxiliary variable  $x$  and

$\rho_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{\mu_{11}}{\sqrt{\mu_{20}} \sqrt{\mu_{02}}}$  coefficient of correlation between  $x$  and  $y$ ,

$\kappa_{(y)} = \lambda_{40} = \frac{\mu_{40}}{\mu_{20}^2}$  coefficient of kurtosis for the study variable,  $\kappa_{(x)} = \lambda_{04} = \frac{\mu_{04}}{\mu_{02}^2}$

coefficient of kurtosis for the auxiliary variable and  $M_x$  population median of the auxiliary variable.

Many authors have come up with more precise estimators by employing prior knowledge of certain population parameter(s). [2] for example attempted use of the coefficient of variation of study variable but prove inadequate for in practice, this parameter is unknown. Motivated by [2] work, [3] [4] and [5] used the known coefficient of variation but now that of the auxiliary variable for estimating population mean of study variable. Reasoning along the same path [6] used the prior value of coefficient of kurtosis of an auxiliary variable in estimating the population variance of the study variable  $y$ .

Kurtosis in most cases is not reported or used in many research articles, in spite of the fact that fundamentally speaking every statistical package provides a measure of kurtosis. This maybe attributed to the likelihood that kurtosis is not well understood or its importance in various aspects of statistical analysis has not been explored fully. Kurtosis can simply be expressed as

$$\kappa = \frac{E(x - \mu)^4}{(E(x - \mu)^2)^2} = \frac{\mu^4}{\sigma^4} \tag{4}$$

where  $E$  —the expectation operator,  $\mu$  —the mean,  $\mu^4$  —the fourth moment about the mean and  $\sigma$  —the standard deviation.

Median being the middlemost value in a distribution (when the values are arranged in ascending or descending order) has the advantage of being less affected by the outliers and skewed data, thus is preferred to the mean especially when the distribution is not symmetrical. We can therefore utilize the median and the coefficient of kurtosis of the auxiliary variable to derive a more precise ratio type population variance.

## 2. Existing Population Variance Estimators

In this section we have reviewed some finite population variance estimators existing in literature which will help in the construction and development of the proposed estimator. Notably, when auxiliary information is not available the usual unbiased estimator to the population variance is

$$t_1 = s_y^2 \quad (5)$$

The bias and MSE of  $t_1$

$$Bias(t_1) = \frac{1-f}{n} S_y^2 \left\{ (\kappa_x - 1) \Psi_1 \left( \Psi_1 - \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right\} = 0 \quad (6)$$

$$\begin{aligned} MSE(t_1) &= Var(t_1) = \frac{1-f}{n} S_y^4 \left\{ (\kappa_y - 1) + (\kappa_x - 1) \Psi_1 \left( \Psi_1 - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \\ &= \frac{1-f}{n} S_y^4 (\kappa_y - 1) \end{aligned} \quad (7)$$

where  $\Psi_1 = 0$

Population variance,  $S_y^2$  estimation using auxiliary information was considered by [7], and proposed ratio type population variance estimator, given by

$$t_2 = s_y^2 \frac{S_x^2}{s_x^2} \quad (8)$$

The bias and Mean Squared Error of Isaki's estimator,

$$Bias(t_2) = \frac{1-f}{n} S_y^2 \left\{ (\kappa_x - 1) \Psi_2 \left( \Psi_2 - \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right\} = \frac{1-f}{n} S_y^2 [(\kappa_x - 1) - (\lambda_{22} - 1)] \quad (9)$$

$$\begin{aligned} MSE(t_2) &= \frac{1-f}{n} S_y^4 \left\{ (\kappa_y - 1) + (\kappa_x - 1) \Psi_2 \left( \Psi_2 - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \\ &= \frac{1-f}{n} S_y^4 [(\kappa_y - 1) + (\kappa_x - 1) - 2(\lambda_{22} - 1)] \end{aligned} \quad (10)$$

where  $\Psi_2 = 1$

[6] initiated the use of coefficient of kurtosis in estimating population variance of a study variable  $y$ . Later, the coefficient of kurtosis was used by [3] [5] [8] in the estimating the population mean.

[9] using the known information on both  $S_x^2$  and  $\kappa_x$  suggested modified ratio type population variance estimator for  $S_y^2$  as

$$t_3 = s_y^2 \left[ \frac{S_x^2 + \kappa_x}{s_x^2 + \kappa_x} \right] \tag{11}$$

The estimator,  $t_3$  bias and MSE obtained as

$$Bias(t_3) = \frac{1-f}{n} S_y^2 \left[ \left( \{\kappa_x - 1\} \Psi_3 \left( \Psi_3 - \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right] \tag{12}$$

$$MSE(t_3) = \frac{1-f}{n} S_y^4 \left[ \left\{ \kappa_y - 1 \right\} + \left\{ \kappa_x - 1 \right\} \Psi_3 \left( \Psi_3 - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right] \tag{13}$$

where  $\Psi_3 = \frac{S_x^2}{S_x^2 + \kappa_x}$

[10] suggested four modified ratio type variance estimators using known values of  $C_x$  and  $\kappa_x$ ,

$$t_4 = s_y^2 \left\{ \frac{S_x^2 - C_x}{s_x^2 - C_x} \right\} \tag{14}$$

$$t_5 = s_y^2 \left\{ \frac{S_x^2 - \kappa_x}{s_x^2 - \kappa_x} \right\} \tag{15}$$

$$t_6 = s_y^2 \left\{ \frac{S_x^2 \kappa_x - C_x}{s_x^2 \kappa_x - C_x} \right\} \tag{16}$$

$$t_7 = s_y^2 \left\{ \frac{S_x^2 C_x - \kappa_x}{s_x^2 C_x - \kappa_x} \right\} \tag{17}$$

The biases and MSE of their estimators,

$$Bias(t_4) = \frac{1-f}{n} S_y^2 (\kappa_x - 1) \left\{ \Psi_4 \left( \Psi_4 - \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right\} \tag{18}$$

$$MSE(t_4) = \frac{1-f}{n} S_y^4 \left\{ \left( \kappa_y - 1 \right) + \Psi_4 (\kappa_x - 1) \left( \Psi_4 - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \tag{19}$$

$$Bias(t_5) = \frac{1-f}{n} S_y^2 (\kappa_x - 1) \left\{ \Psi_5 \left( \Psi_5 - \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \tag{20}$$

$$MSE(t_5) = \frac{1-f}{n} S_y^4 \left\{ \left( \kappa_y - 1 \right) + \Psi_5 (\kappa_x - 1) \left( \Psi_5 - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \tag{21}$$

$$Bias(t_6) = \frac{1-f}{n} S_y^2 (\kappa_x - 1) \left\{ \Psi_6 \left( \Psi_6 - \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \tag{22}$$

$$MSE(t_6) = \frac{1-f}{n} S_y^4 \left\{ \left( \kappa_y - 1 \right) + \Psi_6 (\kappa_x - 1) \left( \Psi_6 - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \tag{23}$$

$$Bias(t_7) = \frac{1-f}{n} S_y^2 (\kappa_x - 1) \left\{ \Psi_7 \left( \Psi_7 - \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \tag{24}$$

$$MSE(t_7) = \frac{1-f}{n} S_y^4 \left\{ \left( \kappa_y - 1 \right) + \Psi_7 (\kappa_x - 1) \left( \Psi_7 - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \tag{25}$$

where;

$$\Psi_4 = \frac{S_x^2}{S_x^2 - C_x}; \quad \Psi_5 = \frac{S_x^2}{S_x^2 - \kappa_x}; \quad \Psi_6 = \frac{S_x^2 \kappa_x}{S_x^2 \kappa_x - C_x}; \quad \Psi_7 = \frac{S_x^2 C_x}{S_x^2 C_x - \kappa_x}.$$

[11] utilizing population median  $M_x$  came up with a modified ratio type population variance estimator as

$$t_8 = s_y^2 \left\{ \frac{S_x^2 + M_x}{s_x^2 + M_x} \right\} \quad (26)$$

The bias and MSE of their estimator  $t_8$ ,

$$\text{Bias}(t_8) = \frac{1-f}{n} S_y^2 (\kappa_x - 1) \left\{ \Psi_8 \left( \Psi_8 - \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \quad (27)$$

$$\text{MSE}(t_8) = \frac{1-f}{n} S_y^4 \left\{ (\kappa_y - 1) + \Psi_8 (\kappa_x - 1) \left( \Psi_8 - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \quad (28)$$

where,  $\Psi_8 = \frac{S_x^2}{S_x^2 + M_x}$ .

[12] using the known quartiles (upper and lower quartile  $Q_3$  and  $Q_1$  respectively) of the auxiliary variable  $x$  suggested

$$t_9 = s_y^2 \left\{ \frac{S_x^2 + Q_1}{s_x^2 + Q_1} \right\} \quad (29)$$

$$t_{10} = s_y^2 \left\{ \frac{S_x^2 + Q_3}{s_x^2 + Q_3} \right\} \quad (30)$$

The biases and MSE of their estimators  $t_9$  and  $t_{10}$  as follows

$$\text{Bias}(t_9) = \frac{1-f}{n} S_y^2 (\kappa_x - 1) \left\{ \Psi_9 \left( \Psi_9 - \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \quad (31)$$

$$\text{MSE}(t_9) = \frac{1-f}{n} S_y^4 \left\{ (\kappa_y - 1) + \Psi_9 (\kappa_x - 1) \left( \Psi_9 - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \quad (32)$$

$$\text{Bias}(t_{10}) = \frac{1-f}{n} S_y^2 (\kappa_x - 1) \left\{ \Psi_{10} \left( \Psi_{10} - \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \quad (33)$$

$$\text{MSE}(t_{10}) = \frac{1-f}{n} S_y^4 \left\{ (\kappa_y - 1) + \Psi_{10} (\kappa_x - 1) \left( \Psi_{10} - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \quad (34)$$

where  $\Psi_9 = \frac{S_x^2}{S_x^2 + Q_1}$  and  $\Psi_{10} = \frac{S_x^2}{S_x^2 + Q_3}$ . Motivated by [10] and [11] [13]

considered the estimation of finite population variance using known coefficient of variation and median of an auxiliary variable, proposed an estimator.

$$t_{11} = s_y^2 \left[ \frac{C_x S_x^2 + M_x}{C_x s_x^2 + M_x} \right] \quad (35)$$

The bias and MSE obtained to be,

$$Bias(t_{11}) = \frac{1-f}{n} S_y^2 (\kappa_x - 1) \left\{ \Psi_{11} \left( \Psi_{11} - \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \quad (36)$$

$$MSE(t_{11}) = \frac{1-f}{n} S_y^4 \left\{ (\kappa_y - 1) + \Psi_{11} (\kappa_x - 1) \left( \Psi_{11} - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \quad (37)$$

where  $\Psi_{11} = \frac{C_x S_x^2}{C_x S_x^2 + M_x}$ .

### 3. Proposed Estimator

Motivated by the works of [14] [9] [15] [13] [10] and [16] in the improvement of the performance of the population variance estimator of the study variable using known population parameters of an auxiliary variable. We propose the following modified ratio type population variance estimator using a known value of population coefficient of kurtosis  $\kappa_x$  and median  $M_x$  of an auxiliary variable.

$$\hat{S}_{PM}^2 = s_y^2 \left\{ \frac{S_x^2 \kappa_x + M_x^2}{s_x^2 \kappa_x + M_x^2} \right\} \quad (38)$$

To calculate the bias and the MSE of  $\hat{S}_{PM}^2$ ,

We let  $s_y^2 = S_y^2 (1 + \xi_0)$  and  $s_x^2 = S_x^2 (1 + \xi_1)$  or  $\xi_0 = \frac{s_y^2}{S_y^2} - 1$  and  $\xi_1 = \frac{s_x^2}{S_x^2} - 1$

so that  $E(\xi_0) = E(\xi_1) = 0$  and to the first degree of approximations

$$E(\xi_0^2) = \frac{1-f}{n} (\lambda_{40} - 1) \quad (39)$$

$$E(\xi_1^2) = \frac{1-f}{n} (\lambda_{04} - 1) \quad (40)$$

$$E(\xi_0 \xi_1) = \frac{1-f}{n} (\lambda_{22} - 1) \quad (41)$$

The expectations are obtained following the works of [17] [18] [19] and [20].

Now expressing  $\hat{S}_{PM}^2$  in terms of  $\xi_i$ 's we have

$$\begin{aligned} \hat{S}_{PM}^2 &= S_y^2 (1 + \xi_0) \left\{ \frac{\kappa_x S_x^2 + M_x^2}{\kappa_x S_x^2 (1 + \xi_1) + M_x^2} \right\} \\ &= S_y^2 (1 + \xi_0) (1 + \rho^* \xi_1)^{-1} \end{aligned} \quad (42)$$

where  $\rho^* = \kappa_x S_x^2 (\kappa_x S_x^2 + M_x^2)^{-1}$ , we assume that  $|\rho^* \xi_1| < 1$  so that  $(1 + \rho^* \xi_1)^{-1}$  is expandable.

Expanding the right hand side of (42) and multiplying out we have

$$\begin{aligned} \hat{S}_{PM}^2 &= S_y^2 (1 + \xi_0) (1 - \rho^* \xi_1 + \rho^{*2} \xi_1^2 + \dots) \\ &= S_y^2 (1 + \xi_0 - \rho^* \xi_1 - \rho^* \xi_0 \xi_1 + \rho^{*2} \xi_1^2 + \rho^{*2} \xi_0 \xi_1^2 - \dots) \end{aligned} \quad (43)$$

Neglecting terms of  $\xi_i$ 's having power greater than two we have

$$\begin{aligned} \hat{S}_{PM}^2 &\cong S_y^2 (1 + \xi_0 - \rho^* \xi_1 - \rho^* \xi_0 \xi_1 + \rho^{*2} \xi_1^2) \\ \text{or } \hat{S}_{PM}^2 - S_y^2 &\cong S_y^2 (\xi_0 - \rho^* \xi_1 - \rho^* \xi_0 \xi_1 + \rho^{*2} \xi_1^2) \end{aligned} \quad (44)$$

Taking the expectation on both sides of (44)

$$E\left(\hat{S}_{PM}^2 - S_y^2\right) \cong E\left(S_y^2\left(\xi_0 - \rho^* \xi_1 - \rho^* \xi_0 \xi_1 + \rho^{*2} \xi_1^2\right)\right) \quad (45)$$

we have our bias

$$\text{Bias}\left(\hat{S}_{PM}^2\right) = \frac{1-f}{n} S_y^2 (\kappa_x - 1) \rho^* \left\{ \rho^* - \frac{\lambda_{22} - 1}{\kappa_x - 1} \right\} \quad (46)$$

Squaring both sides of (44) and neglecting terms of  $\xi$ 's having power greater than two we have

$$\left(\hat{S}_{PM}^2 - S_y^2\right)^2 \cong S_y^4 \left(\xi_0^2 + \rho^{*2} \xi_1^2 - 2\rho^* \xi_0 \xi_1\right) \quad (47)$$

Taking the expectation on both sides of (47)

$$E\left(\left(\hat{S}_{PM}^2 - S_y^2\right)^2\right) \cong E\left(S_y^4 \left(\xi_0^2 + \rho^{*2} \xi_1^2 - 2\rho^* \xi_0 \xi_1\right)\right) \quad (48)$$

We get the  $\left(\hat{S}_{PM}^2\right)$  estimator's Mean Squared Error as

$$\text{MSE}\left(\hat{S}_{PM}^2\right) = \frac{1-f}{n} S_y^4 \left\{ (\kappa_y - 1) + \rho^* (\kappa_x - 1) \left( \rho^* - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\} \quad (49)$$

#### 4. Theoretical Comparison

The theoretical conditions under which the proposed modified ratio type estimators  $\hat{S}_{PM}^2$  is more efficient than the other existing estimators  $t_j, j = 0, 1, 2, \dots, 11$ , from MSE of  $t_j, j = 0, 1, 2, \dots, 11$  given to the first degree of approximation in general as

$$\text{MSE}(t_j) = \frac{1-f}{n} S_y^4 \left[ (\kappa_y - 1) + \Psi_j (\kappa_x - 1) \left( \Psi_j - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right] \quad (50)$$

Using Equations ((49) and (50)) we have that  $\text{MSE}\left(\hat{S}_{PM}^2\right) < \text{MSE}(t_j)$ , if

$$\rho^* \left( \rho^* - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) < \Psi_j \left( \Psi_j - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \quad (51)$$

#### 5. Empirical Studies

Using the data from Population I (Source:[21], 228]) and Population II (source: [22]). We assess the performance of the proposed estimator when simple random sampling without replacement (SRSWOR) scheme is used with that of sample variance and existing estimators. We apply the proposed and existing estimators to this data set and the data statistics are given below:

##### Population I

$X$  = Fixed capital

$Y$  = output of 80 factories

$N = 80, n = 20, \bar{X} = 11.265, \bar{Y} = 51.826,$

$S_x^2 = 71.504, S_y^2 = 336.979, S_{xy} = 146.068,$

$\lambda_{04} = \kappa_x = 2.866, \lambda_{40} = \kappa_y = 2.267, \lambda_{22} = 2.221,$



$$\rho_{xy} = 0.941, C_y = 0.354, C_x = 0.751$$

$$M_x = 10.300, Q_1 = 5.150, Q_3 = 16.975$$

**Population II**

$X$  = acreage under wheat crop in 1973

$Y$  = acreage under wheat crop in 1974,

$$N = 70, n = 25, \bar{X} = 175.2671, \bar{Y} = 96.700,$$

$$S_x^2 = 19840.7508, S_y^2 = 3686.1898,$$

$$\lambda_{04} = \kappa_x = 7.0952, \lambda_{40} = \kappa_y = 4.7596, \lambda_{22} = 4.6038,$$

$$\rho_{xy} = 0.7293, C_y = 0.6254, C_x = 0.8037$$

$$M_x = 72.4375, Q_1 = 80.1500, Q_3 = 225.0250.$$

Using the above summary values we have the results in **Table 1** below. From the table Mean Squared Errors it is clear that our proposed modified ratio type population variance estimator  $\hat{S}_{PM}^2$  has the least Mean Squared Error (MSE).

The efficiency of our proposed estimator  $\hat{S}_{PM}^2$  is examined numerically by its Percentage Relative Efficiency (PRE(s)) in comparison with those of existing

**Table 1.** Bias and Mean Squared Errors (MSE).

ESTIMATOR	Population I		Population II	
	BIAS(.)	MSE	BIAS(.)	MSE
$t_1 = s_y^2$	0	5395.289	0	1,313,625.261
$t_2 = s_y^2 \left( \frac{S_x^2}{s_x^2} \right)$	8.151	3276.421	236.154	924,946.481
$t_3 = s_y^2 \left( \frac{S_x^2 + \kappa_x}{s_x^2 + \kappa_x} \right)$	6.956	2740.349	235.656	924,324.375
$t_4 = s_y^2 \left( \frac{S_x^2 - C_x}{s_x^2 - C_x} \right)$	8.512	3006.373	236.187	925,017.011
$t_5 = s_y^2 \left( \frac{S_x^2 - \kappa_x}{s_x^2 - \kappa_x} \right)$	9.518	3186.399	236.445	925,569.577
$t_6 = s_y^2 \left( \frac{S_x^2 \kappa_x - C_x}{s_x^2 \kappa_x - C_x} \right)$	8.279	2965.067	236.159	924,956.421
$t_7 = s_y^2 \left( \frac{S_x^2 C_x - \kappa_x}{s_x^2 C_x - \kappa_x} \right)$	10.002	3275.722	236.517	925,721.916
$t_8 = s_y^2 \left( \frac{S_x^2 + M_x}{s_x^2 + M_x} \right)$	4.530	2377.418	233.201	918,641.426
$t_9 = s_y^2 \left( \frac{S_x^2 + Q_1}{s_x^2 + Q_1} \right)$	6.126	2609.91	232.889	917,976.121
$t_{10} = s_y^2 \left( \frac{S_x^2 + Q_3}{s_x^2 + Q_3} \right)$	2.934	2181.488	227.099	905,689.896
$t_{11} = s_y^2 \left( \frac{S_x^2 C_x + M_x}{s_x^2 C_x + M_x} \right)$	3.656	2314.033	232.485	917,116.922
$\hat{S}_{PM}^2 = s_y^2 \left\{ \frac{S_x^2 \kappa_x + M_x^2}{s_x^2 \kappa_x + M_x^2} \right\}$	0.708	1993.270	207.653	865,134.030

estimators using real populations from [[21], p.228] and [22].

We have computed the PRE(s) of the estimators  $t_j, j=1,2,\dots,11$  using the formulae

$$PRE(t_j, s_y^2) = \frac{MSE(s_y^2)}{MSE(t_j)} \times 100 \tag{52}$$

$$= \frac{\frac{1-f}{n} S_y^4 (\kappa_y - 1)}{\frac{1-f}{n} S_y^4 \left[ \{\kappa_y - 1\} + \{\kappa_x - 1\} \Psi_j \left( \Psi_j - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right]} \times 100 \tag{53}$$

$$= \frac{\kappa_y - 1}{\{\kappa_y - 1\} + \{\kappa_x - 1\} \Psi_j \left( \Psi_j - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right)} \times 100 \tag{54}$$

Then PRE for our proposed estimator is subsequently,

$$PRE(\hat{S}_{PM}^2, s_y^2) = \frac{MSE(s_y^2)}{MSE(\hat{S}_{PM}^2)} \times 100 \tag{55}$$

$$= \frac{\frac{1-f}{n} S_y^4 (\kappa_y - 1)}{\frac{1-f}{n} S_y^4 \left\{ (\kappa_y - 1) + \rho^* (\kappa_x - 1) \left( \rho^* - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right) \right\}} \times 100 \tag{56}$$

$$= \frac{\kappa_y - 1}{(\kappa_y - 1) + \rho^* (\kappa_x - 1) \left( \rho^* - 2 \left( \frac{\lambda_{22} - 1}{\kappa_x - 1} \right) \right)} \times 100 \tag{57}$$

Using formula (54) and (57) we compute the Percent Relative Efficiencies and tabulate the results in **Table 2**.

Percentage Relative efficiency being a robust statistical tool that is used to

**Table 2.** Percent Relative Efficiencies (PRE).

ESTIMATOR	POPULATION I	POPULATION II
$t_1$	100	100
$t_2$	164.67	142.02
$t_3$	196.88	142.12
$t_4$	179.46	142.01
$t_5$	169.32	141.93
$t_6$	181.96	142.02
$t_7$	164.71	141.90
$t_8$	226.94	143.00
$t_9$	206.72	143.10
$t_{10}$	247.32	145.04
$t_{11}$	233.16	143.23
$\hat{S}_{PM}^2$	270.68	151.84

measure and ascertain the efficiency of one estimator over another. From the findings summarized in the table above it is clear that our proposed estimator  $\hat{S}_{PM}^2$  performed best, that is it has the highest PRE among all the other estimators. This therefore implies that we can apply our proposed estimator to appropriate practical situations and obtain better and more efficient results than the traditional and other existing population variance estimators.

## 6. Conclusions

In this study we have proposed a modified ratio type population variance estimator using known population parameters the coefficient of kurtosis and the median of the auxiliary variable  $x$ .

We have analyzed the performance of our proposed estimator against the usual unbiased variance estimator and existing estimators using two natural populations by comparing their PRE(s).

Based on the results of our studies, it is evidenced that our proposed estimator works better than the other existing estimators having the highest Percentage Relative Efficiency hence can be applied to practical applications, where knowledge of population parameters of auxiliary variable is available. We also recommend that our proposed estimator can be further improved by extending the number of Taylor's series terms to be more than two.

## Acknowledgements

We give much appreciation to the authors for the numerous and valuable contribution to this work.

## Conflicts of Interest

The authors declare that there is no conflict of interest in the publication of this paper.

## References

- [1] Cochran, W. (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**, 262-275. <https://doi.org/10.1017/S0021859600048012>
- [2] Searls, D.T. (1964) The Utilization of a Known Coefficient of Variation in the Estimation Procedure. *Journal of the American Statistical Association*, **59**, 1225-1226. <https://doi.org/10.1080/01621459.1964.10480765>
- [3] Sen, A.R. (1978) Estimation of the Population Mean When the Coefficient of variation Is Known. *Communications in Statistics—Theory and Methods*, **7**, 657-672. <https://doi.org/10.1080/03610927808827656>
- [4] Sisodia, B. and Dwivedi, V. (1981) Modified Ratio Estimator Using Coefficient of Variation of Auxiliary Variable. *Journal-Indian Society of Agricultural Statistics*, **33**, 13-18.
- [5] Upadhyaya, L.N. and Singh, H.P. (1984) On the Estimation of the Population Mean With Known Coefficient of Variation. *Biometrical Journal*, **26**, 915-922.

- <https://doi.org/10.1002/bimj.4710260814>
- [6] Hirano, K., Pandey, B.N. and Singh, J. (1973) On the Utilization of a Known Coefficient of Kurtosis in the Estimation Procedure of Variance. *Annals of the Institute of Statistical Mathematics*, **25**, 51-55. <https://doi.org/10.1007/BF02479358>
- [7] Isaki, C.T. (1983) Variance Estimation Using Auxiliary Information. *Journal of the American Statistical Association*, **78**, 117-123. <https://doi.org/10.1080/01621459.1983.10477939>
- [8] Searls, D.T. and Intarapanich, P. (1990) A Note on an Estimator for the Variance That Utilizes The kurtosis. *The American Statistician*, **44**, 295-296.
- [9] Upadhyaya, L. and Singh, H. (1999) An Estimator for Population Variance That Utilizes the Kurtosis of an Auxiliary Variable in Sample Surveys. *Vikram Mathematical Journal*, **19**, 14-17.
- [10] Kadilar, C. and Cingi, H. (2006) Ratio Estimators for the Population Variance in Simple and Stratified Random Sampling. *Applied Mathematics and Computation*, **173**, 1047-1059. <https://doi.org/10.1016/j.amc.2005.04.032>
- [11] Subramani, J. and Kumarapandiyan, G. (2012) Variance Estimation Using Median of the Auxiliary Variable. *International Journal of Probability and Statistics*, **1**, 6-40. <https://doi.org/10.5923/j.ijps.20120103.02>
- [12] Subramani, J. and Kumarapandiyan, G. (2012) Variance Estimation Using Quartiles and Their Functions of an Auxiliary Variable. *International Journal of Statistics and Applications*, **2**, 67-72. <https://doi.org/10.5923/j.statistics.20120205.04>
- [13] Subramani, J. and Kumarapandiyan, G. (2013) Estimation of Variance Using Known Coefficient of Variation and Median of an Auxiliary Variable. *Journal of Modern Applied Statistical Methods*, **12**, 11. <https://doi.org/10.22237/jmasm/1367381400>
- [14] Khan, M. and Shabbir, J. (2013) A Ratio Type Estimator for the Estimation of Population Variance Using Quartiles of an Auxiliary Variable. *Journal of Statistics Applications and Probability*, **2**, 157-162. <https://doi.org/10.12785/jsap/020314>
- [15] Singh, H.P., Tailor, R., Tailor, R. and Kakran, M. (2004) An Improved Estimator of Population Mean Using Power Transformation. *Journal of the Indian Society of Agricultural Statistics*, **58**, 223-230.
- [16] Yadav, S.K., Misra, S. and Mishra, S. (2016) Efficient Estimator for Population Variance Using Auxiliary Variable. *American Journal of Operational Research*, **6**, 9-15.
- [17] Sukhatme, P. (1944) Moments and Product Moments of Moment-Statistics for Samples of the Finite and Infinite Populations. *Sankhya: The Indian Journal of Statistics*, **6**, 363-382.
- [18] Sukhatme, P. and Sukhatme, B. (1970) Sampling Theory of Surveys with Applications. Asia Publishing House, Bombay.
- [19] Srivastava, S.K. and Jhaji, H.S. (1981) A Class of Estimators of the Population Mean in Survey Sampling Using Auxiliary Information. *Biometrika*, **68**, 341-343. <https://doi.org/10.1093/biomet/68.1.341>
- [20] Tracy, D.S. (1984) Moments of Sample Moments. *Communications in Statistics-Theory and Methods*, **13**, 553-562. <https://doi.org/10.1080/03610928408828700>
- [21] Murthy, M.N. (1967) Sampling Theory and Methods. Statistical Publishing Society, Barrackpore.
- [22] Daroga, S. and Chaudhary, F. (1986) Theory and Analysis of Sample Survey Designs. Wiley Eastern Limited, New York.