

## Variance Estimator Using Tri-mean and Inter Quartile Range of Auxiliary Variable

S.K. Yadav <sup>1</sup>, Sheela Misra <sup>2</sup>, S. S. Mishra <sup>3\*</sup> and Shankar Prasad Khanal <sup>4</sup>

*Submitted: 12 July 2017; Accepted: 30 July 2017*

---

### ABSTRACT

**Background:** Whenever the population is large and it is very time taking and costly to take observation on each unit of the population then sampling is the only way to get the appropriate estimate of the population parameter under consideration. Many authors have given many estimators for estimating population variance with greater efficiency.

**Objective:** The objective of the study is to search for more efficient estimator than the competing estimators of population variance of study variable.

**Materials and Methods:** The estimator utilizing information on tri-mean and inter quartile range of auxiliary variable has been developed. The expressions for the bias and mean squared error (MSE) of the proposed estimator have been derived up to the first order of approximation. A theoretical comparison of the proposed estimator has been made with the competing estimators of population variance.

**Results:** The theoretical findings have been justified with the help of numerical example from some natural populations. It has been found that the proposed estimator is best among the competing estimators of population variance as it has least mean squared error among them.

**Conclusion:** Since the proposed estimator is best among the competing estimator of the population variance, therefore it must be used for the improved estimation of population variance.

**Keywords:** Auxiliary variable, bias, efficiency, mean squared error, study variable,

---

**Address correspondence to the authors:** Department of Mathematics and Statistics (A Centre of Excellence), Dr RML Avadh University, Faizabad-224001, U.P., India<sup>1</sup>,

Email: drskystats@gmail.com; Department of Statistics, University of Lucknow, Lucknow-226007, U.P., India<sup>2</sup>, Email: profsheelamisra@gmail.com; Department of Mathematics and Statistics (A

Centre of Excellence), Dr RML Avadh University, Faizabad-224001, U.P., India<sup>3</sup>,

Email: profsantmishra@gmail.com (\* Corresponding author email); Central Department of Statistics, Tribhuvan University, Kirtipur, Nepal<sup>4</sup>, Email: drshankarcds@gmail.com

---

## INTRODUCTION

The natural estimator for any parameter is the corresponding statistics. Thus the most appropriate estimator for the population variance is the sample variance of the study variable. Auxiliary information is supplied by the auxiliary variable which is highly associated (positively or negatively) with the study variable. The auxiliary information is in use almost from the use of sampling itself. It is used in various forms and at various stages of sampling in the form of population parameters of auxiliary variable for more efficiently estimating population parameters of the study variable under investigation. The auxiliary information is used in two different stages of designing and estimation. In the present study, it has been used at estimation stage only. Whenever the study variable ( $Y$ ) is highly positively correlated with the study variable and the line of regression of  $Y$  on  $X$  passes through origin, the ratio type estimators are used for improvement in the estimating population parameters while product type estimators are used when  $X$  and  $Y$  are negatively correlated. On the other hand regression estimators are used when the line of regression does not pass through the origin.

Let we have  $N$  distinct and identifiable units in the finite population under consideration and let  $(x_i, y_i), i = 1, 2, \dots, n$  be a finite sample of size  $n$  taken on bivariate paradigm  $(X, Y)$  using a *SRSWOR* scheme. Let  $\bar{X}$  be the population mean of the auxiliary variable and  $\bar{Y}$  the population mean of study variable respectively. Let  $\bar{x}$  and  $\bar{y}$  be the corresponding sample means of auxiliary and study variables respectively and are unbiased estimators of  $\bar{X}$  and  $\bar{Y}$  respectively. Let  $\rho$  be the correlation coefficient between  $X$  and  $Y$  and  $Q_r = Q_3 - Q_1$  be the inter quartile range, semi-quartile range  $Q_d = \frac{Q_3 - Q_1}{2}$ , quartile average  $Q_a = \frac{Q_3 + Q_1}{2}$  of the auxiliary variable  $X$ . We propose a ratio type estimator of  $\bar{Y}$  using tri-mean  $TM = \frac{(Q_1 + 2Q_2 + Q_3)}{4}$  and  $Q_r$ .

### Review of variance estimators

For estimating population variance, the most appropriate and often used estimator is the sample variance given by:

$$t_0 = s_y^2, \quad (1)$$

$$\text{where } s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

It is an unbiased estimator of population variance of study variable and up to the first order of approximation, its variance is:

$$V(t_0) = \lambda S_y^4 (\lambda_{40} - 1) \quad (2)$$

where  $\lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{r/2} \mu_{02}^{s/2}}$ ,  $\mu_{rs} = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^r (X_i - \bar{X})^s$  ( $r = 0, 1, \dots, 4$   $s = 0, 1, \dots, 4$ ),

$$\lambda = \frac{(1-f)}{n} \text{ and } f = \frac{n}{N}.$$

Isaki (1983) used the positively correlated auxiliary information and proposed the following usual ratio estimator of population variance of the study variable as:

$$t_R = s_y^2 \left[ \frac{S_x^2}{s_x^2} \right] \tag{3}$$

where  $s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ ,  $S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2$ ,  $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$ ,  $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$ ,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

The Bias and Mean Square Error of above estimator, up to the first order of approximation, are respectively given by:

$$B(t_R) = \lambda S_y^2 [(\lambda_{04} - 1) - (\lambda_{22} - 1)], \tag{4}$$

$$MSE(t_R) = \lambda S_y^4 [(\lambda_{40} - 1) + (\lambda_{04} - 1) - 2(\lambda_{22} - 1)] \tag{5}$$

Upadhyaya and Singh (1999) used the information on population coefficient of kurtosis of auxiliary variable and proposed the following ratio type estimator of population variance as:

$$t_1 = s_y^2 \left[ \frac{S_x^2 + \beta_{2(x)}}{s_x^2 + \beta_{2(x)}} \right] \tag{6}$$

The Bias and Mean Square Error of above estimator, up to the first order of approximation, are respectively given by:

$$B(t_1) = \lambda S_y^2 R_1 [R_1 (\lambda_{04} - 1) - (\lambda_{22} - 1)] \tag{7}$$

$$MSE(t_1) = \lambda S_y^4 [(\lambda_{40} - 1) + R_1^2 (\lambda_{04} - 1) - 2R_1 (\lambda_{22} - 1)] \tag{8}$$

where,  $R_1 = \frac{S_x^2}{S_x^2 + \beta_{2(x)}}$

Kadilar and Cingi (2006) proposed the following three ratio type estimators of population variance using population coefficient of variation and population coefficient of kurtosis of auxiliary variable as:

$$t_2 = s_y^2 \left[ \frac{S_x^2 + C_x}{s_x^2 + C_x} \right] \tag{9}$$

$$t_3 = s_y^2 \left[ \frac{S_x^2 \beta_{2(x)} + C_x}{S_x^2 \beta_{2(x)} + C_x} \right] \tag{10}$$

$$t_4 = s_y^2 \left[ \frac{S_x^2 C_x + \beta_{2(x)}}{S_x^2 C_x + \beta_{2(x)}} \right] \tag{11}$$

The Biases and Mean Square Errors of above estimators, up to the first order of approximation, are respectively given by:

$$B(t_i) = \lambda S_y^2 R_i [R_i (\lambda_{04} - 1) - (\lambda_{22} - 1)], \quad (i = 2, 3, 4) \tag{12}$$

$$MSE(t_i) = \lambda S_y^4 [(\lambda_{40} - 1) + R_i^2 (\lambda_{04} - 1) - 2R_i (\lambda_{22} - 1)], \quad (i = 2, 3, 4) \tag{13}$$

where,  $R_2 = \frac{S_x^2}{S_x^2 + C_x}$ ,  $R_3 = \frac{S_x^2 \beta_{2(x)}}{S_x^2 \beta_{2(x)} + C_x}$ ,  $R_4 = \frac{S_x^2 C_x}{S_x^2 C_x + \beta_{2(x)}}$

Subramani and Kumarpandiyan (2012) proposed the following ratio type estimators of population variance using quartiles and functions of quartiles of auxiliary variable as:

$$t_5 = s_y^2 \left[ \frac{S_x^2 + Q_1}{S_x^2 + Q_1} \right] \tag{14}$$

$$t_6 = s_y^2 \left[ \frac{S_x^2 + Q_3}{S_x^2 + Q_3} \right] \tag{15}$$

$$t_7 = s_y^2 \left[ \frac{S_x^2 + Q_r}{S_x^2 + Q_r} \right] \tag{16}$$

$$t_8 = s_y^2 \left[ \frac{S_x^2 + Q_d}{S_x^2 + Q_d} \right] \tag{17}$$

$$t_9 = s_y^2 \left[ \frac{S_x^2 + Q_a}{S_x^2 + Q_a} \right] \tag{18}$$

The Biases and Mean Square Errors of above estimators, up to the first order of approximation, are respectively given by:

$$B(t_i) = \lambda S_y^2 R_i [R_i (\lambda_{04} - 1) - (\lambda_{22} - 1)], \quad (i = 5, 6, 7, 8, 9) \tag{19}$$

$$MSE(t_i) = \lambda S_y^4 [(\lambda_{40} - 1) + R_i^2 (\lambda_{04} - 1) - 2R_i (\lambda_{22} - 1)], \quad (i = 5, 6, 7, 8, 9) \tag{20}$$

$$R_5 = \frac{S_x^2}{S_x^2 + Q_1}, R_6 = \frac{S_x^2}{S_x^2 + Q_3}, R_7 = \frac{S_x^2}{S_x^2 + Q_r}, R_8 = \frac{S_x^2}{S_x^2 + Q_d}, R_9 = \frac{S_x^2}{S_x^2 + Q_a}.$$

Khan and Shabbir (2013) proposed the following estimator using correlation coefficient between study and auxiliary variable and the third quartile of the auxiliary variable as:

$$t_{10} = s_y^2 \left[ \frac{S_x^2 \rho + Q_3}{s_x^2 \rho + Q_3} \right] \tag{21}$$

The *Bias* and Mean Square Error of above estimator, up to the first order of approximation, are respectively given by:

$$B(t_{10}) = \lambda S_y^2 R_{10} [R_{10}(\lambda_{04} - 1) - (\lambda_{22} - 1)] \tag{22}$$

$$MSE(t_{10}) = \lambda S_y^4 [(\lambda_{40} - 1) + R_{10}^2(\lambda_{04} - 1) - 2R_{10}(\lambda_{22} - 1)] \tag{23}$$

$$R_{10} = \frac{S_x^2 \rho}{S_x^2 \rho + Q_3}$$

Maqbool and Javaid (2017) proposed the following estimator using tri-mean and semi quartile average auxiliary variable as:

$$t_{11} = s_y^2 \left[ \frac{S_x^2 + (TM + Q_a)}{s_x^2 + (TM + Q_a)} \right] \tag{24}$$

The *Bias* and Mean Square Error of above estimator, up to the first order of approximation, are respectively given by:

$$B(t_{11}) = \lambda S_y^2 R_{11} [R_{11}(\lambda_{04} - 1) - (\lambda_{22} - 1)] \tag{25}$$

$$MSE(t_{11}) = \lambda S_y^4 [(\lambda_{40} - 1) + R_{11}^2(\lambda_{04} - 1) - 2R_{11}(\lambda_{22} - 1)] \tag{26}$$

$$R_{11} = \frac{S_x^2}{S_x^2 + (TM + Q_a)}$$

In the series of improvement, latest reference can be made of Yadav, Mishra, Kumar and Kadlilar (2016), and Yadav, Mishra and Shukla (2016).

## MATERIALS AND METHODS

### *Proposed estimator*

On the lines of Maqbool and Javaid (2017), we propose an improved ratio estimator of population variance using tri-mean and inter quartile range of auxiliary variable as:

$$t_p = s_y^2 \left[ \frac{S_x^2 (TM + Q_r)}{s_x^2 (TM + Q_r)} \right] \tag{27}$$

The idea behind proposing new estimator of population variance using tri-mean and the inter quartile range is that  $Q_a$  alone should have no relationship with the variance but inter quartile range  $Q_r$  should better reflect on the variance and this indeed is the case as shown in results.

In order to study the sampling properties of the proposed estimator,  $t_p$ , we define  $s_y^2 = S_y^2(1 + \varepsilon_0)$

and  $s_x^2 = S_x^2(1 + \varepsilon_1)$  such that  $E(\varepsilon_i) = 0$  for  $(i = 0, 1)$  and  $E(\varepsilon_0^2) = \frac{1-f}{n}(\lambda_{40} - 1)$ ,

$$E(\varepsilon_1^2) = \frac{1-f}{n}(\lambda_{04} - 1), E(\varepsilon_0\varepsilon_1) = \frac{1-f}{n}(\lambda_{22} - 1).$$

Expressing  $t_p$  in terms of  $\varepsilon_i$ 's ( $i = 0, 1$ ), we have

$$t_p = S_y^2(1 + e_0)(1 + R_p e_1)^{-1}, \text{ where } R_p = \frac{S_x^2}{S_x^2 + (TM + Q_r)}$$

After simplification and retaining the terms up to the first order of approximation, we have

$$t_p = S_y^2(1 + e_0 - R_p e_1 - R_p e_0 e_1 + R_p^2 e_1^2).$$

Subtracting  $S_y^2$  on both the sides, we get

$$t_p - S_y^2 = S_y^2(e_0 - R_p e_1 - R_p e_0 e_1 + R_p^2 e_1^2). \tag{28}$$

Taking expectation on both sides of (28), we have the *Bias* of proposed estimator  $t_p$  as:

$$B(t_p) = \lambda S_y^2 [R_p^2(\lambda_{04} - 1) - R_p(\lambda_{22} - 1)] \tag{29}$$

Squaring on both sides of (28), after simplifying terms up to the first order of approximation and taking expectation on both sides, we have the *MSE* of the proposed estimator as:

$$MSE(t_p) = \lambda S_y^4 [(\lambda_{40} - 1) + R_p^2(\lambda_{04} - 1) - 2R_p(\lambda_{22} - 1)] \tag{30}$$

## RESULTS

### Efficiency comparison

The proposed estimator in (27) performs better than the usual mean per unit estimator in (1) under the condition:

$$MSE(t_p) - V(t_0) < 0, \text{ if } R_p(\lambda_{04} - 1) < 2(\lambda_{22} - 1) \tag{31}$$

The proposed estimator in (27) used the information on tri-mean and the inter quartile range of auxiliary variable is better than the Isaki (1983) estimator under the condition that

$$MSE(t_p) - MSE(t_R) < 0, \text{ if } (R_p + 1)(\lambda_{04} - 1) < 2(\lambda_{22} - 1). \tag{32}$$

The proposed estimator in (27) is more efficient than the other competing estimators utilizing auxiliary information in the form of different parameters of auxiliary under the condition:

$$MSE(t_p) - MSE(t_i) < 0, (i = 1, 2, \dots, 10) \text{ if, } (R_p + R_i)(\lambda_{04} - 1) < 2(\lambda_{22} - 1) \tag{33}$$

**Numerical illustration**

The performance of the proposed estimator along with the other competing estimators under simple random sampling without replacement (SRSWOR) has been judged using the data of Murthy (1967). In the above example fixed capital is the auxiliary variable  $X$  and output of 80 factories is the study variable  $Y$ . The biases and the mean squared errors for the proposed and existing estimators have been computed for this data. The parameters for this data are given below.

$$N = 80, n = 20, \bar{Y} = 51.8264, \bar{X} = 11.2646, \rho = 0.9413, S_y = 18.3549, C_y = 0.3542, S_x = 8.4563, C_x = 0.7507, \lambda_{04} = 2.8664, \lambda_{40} = 2.2667, \lambda_{22} = 2.2209, Q_1 = 5.1500, Q_3 = 16.975, Q_r = 11.825, Q_d = 5.9125, Q_a = 11.0625.$$

The results of for the proposed estimator along with the competing estimators for above example are represented in following Table-1 in the form of bias and mean squared error.

**Table 1.** Bias and MSE of proposed and competing estimators of population variance.

Estimator	Bias	MSE
Mean per unit estimator $t_0$	0	5393.89
Isaki (1983) estimator $t_R$	10.87	3925.16
Upadhyaya and Singh (1999) $t_1$	9.29	3658.41
Kadilar and Cingi (2006) $t_2$	10.44	3850.16
Kadilar and Cingi (2006) $t_3$	10.72	3898.56
Kadilar and Cingi (2006) $t_4$	8.81	3580.83
Subramani and Kumarpandiyan (2012) $t_5$	8.17	3480.55
Subramani and Kumarpandiyan (2012) $t_6$	3.91	2908.65
Subramani and Kumarpandiyan (2012) $t_7$	5.50	3098.41
Subramani and Kumarpandiyan (2012) $t_8$	7.82	3427.19
Subramani and Kumarpandiyan (2012) $t_9$	5.77	3133.33
Khan and Shabbir (2013) $t_{10}$	3.62	2878.56
Maqbool and Javaid (2017) $t_{11}$	3.03	2820.06
Yadav et al. (Proposed) $t_p$	2.14	2103.01

From above Table 1, we can see that the proposed estimator has minimum bias 2.14 and mean squared error 2103.01 among the class of all competing estimators of population variance.

## CONCLUSION

In present work, we have proposed a ratio type estimator of population variance of study variable using information on tri-mean and the inter quartile range of auxiliary variable. The bias and mean squared error of the proposed estimator have been obtained up to the first order of approximation. The proposed estimator is compared theoretically as well as numerically with the competing estimators of population variance. From table-1, it is clear that proposed estimator is best among all estimators of population variance as it has least bias and mean squared error. Thus it is recommended to use the proposed estimator for improved estimation of population variance of study variable using auxiliary information.

## CONFLICT OF INTEREST

None of the authors have any proprietary interests or conflicts of interest related to this submission.

## ACKNOWLEDGEMENTS

The authors are very much thankful to the editorial board of NJS and the anonymous referees for the critically examining the manuscript which improved the earlier draft.

## REFERENCES

- Isaki, C. T. (1983). Variance estimation using auxiliary information. *Journal of American Statistical Association*, 78, 117- 123.
- Kadilar, C., & Cingi, H. (2006). Improvement in variance estimation using auxiliary information. *Hacetatepe Journal of Mathematics and Statistics*, 35, 111-15.
- Khan, M., & Shabbir, J. (2013). A Ratio Type Estimator for the Estimation of Population Variance using Quartiles of an Auxiliary Variable. *Journal of Statistics Applications & Probability*, 2, 3, 319-325.
- Subramani, J., & Kumarapandiyan, G. (2012). Variance estimation using quartiles and their functions of an auxiliary variable. *International Journal of Statistics and Applications*, 2, 67-72.
- Maqbool, S., & Jaaid, S. (2017). Variance estimation using linear combination of tri-mean and quartile average, *American Journal of Biological and Environmental Statistics*, 3(1), 5-9.
- Murthy, M. N. (1967). *Sampling Theory and Methods*, Statistical Publishing Society Calcutta, India.



Upadhyaya, L. N., & Singh, H. P. (1999). Use of auxiliary information in the estimation of population Variance. *Mathematical Forum*, 4, 33-36.

Yadav, S.K., Mishra, S.S., Kumar, S., & Kadilar, C. (2016). A new improved class of estimators for the population variance, *Journal of Statistics Applications and Probability*, 5(3), 385-392.

Yadav, S.K., Mishra, S. S., & Shukla, A. K. (2016). Use of Correlation Coefficient and Quartiles of Auxiliary Variable for Improved Estimation of Population Variance, *American Journal of Operational Research*, 6(2), 33-39.

---

**Reference** to this paper should be made as follows:

Yadav, S.K, Misra, S., Mishra, S. S., & Khanal, S. P. (2017). Variance estimator using tri-mean and inter quartile range of auxiliary variable. *Nep. J. Stat.*, 1, 83-91.

---