

## Original Research article

### New modified efficient Variance Estimators for the estimation of Population Variance in Survey Sampling

#### Abstract:

For efficient and precise estimate of population variance, we have proposed some new ratio type variance estimators in the current research work by incorporating the linear combination of conventional and non-conventional parameters as auxiliary information. Bias and mean square error has been computed up to the first order of approximation. Numerical illustration of both existing and proposed estimators has been demonstrated to test the efficacy of proposed estimators.

**Keywords**’; Sample Random Sampling, Bias, MSE, HL, MR, and efficiency.

- 1. Introduction:** Population variance is one of the key indicators to provide the knowledge about the variations. As the variations occurs naturally almost in every field of science, even in our day to day life. To have the accuracy about the levels of these variations, we need precise and efficient estimators for the estimation of finite population variance. Especially when the estimates are being operated in terms of ratios, factors like the differences in ratios, skewness and presence of outliers cause biases in the estimation. As it is well known fact that accurate and precise estimates help us in decision making, cost savings as well as in policy making. Thus in order to have the reliable and precise estimates of population parameters, finding an efficient estimator is a common practice of authors to address the issue of precision. Using auxiliary data, Isaki [1] introduced the ratio and regression estimator. Later, Kadilar & Cingi [2] proposed ratio type variance estimator by utilizing coefficient of skewness as supplementary information to enhance the efficiency of estimator. Similarly, Upadhaya and Singh [3] have introduced ratio type variance estimators by incorporating the coefficient of kurtosis as auxiliary information to improve the efficiency of estimators over existing estimators. A Class of estimators was also proposed by Subramani and Kumarapandiyan who have modified estimators using quartiles and quartile deviation as supplementary information to improve the efficiency of estimators. By employing the modification technique to achieve the goal of efficiency, we have developed some new efficient estimators for the estimation of population variance.

#### 2. Review of estimators in literature:

Let the finite population under survey be  $U = \{U_1, U_2, \dots, U_N\}$ , consists of  $N$  distinct and identifiable units. Let  $Y$  be a real variable with value  $Y_i$  measured on  $U_i, i = 1, 2, 3, \dots, N$ , giving a vector  $Y = \{y_1, y_2, \dots, y_N\}$ . The goal is to estimate the populations mean  $\bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i$  and its variance  $S_Y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2$  on the basis of random sample selected from a finite population ‘U’.

#### 3. The notations used for the present research work are as follows:

$N =$  Population size,

$n =$  sample size .

$$\gamma = \frac{1}{n}$$

$Y =$  Study Variable,

$X =$  Auxiliary Variable,

$\bar{X}, \bar{Y} =$  Population means for auxiliary and study variable ,

$\bar{x}, \bar{y} =$  Sample means for auxiliary and study variable ,

$S_Y^2, S_x^2 =$  population Variances for study and auxiliary variable ,

$s_y^2, s_x^2 =$  Sample variances for study and auxiliary variable

$C_x, C_y =$  Coefficient of variations,

$\rho =$  Coefficient of Correlation between auxiliary and study variable ,

$\beta_{2(x)} =$  Coefficient of kurtosis for auxiliary variable ,

$\beta_{2(y)} =$  Coefficient of kurtosis for Study Variable

$HL =$  Hodg's - Lehmanne,  $MR =$  Mid - range for auxiliary variable

$$\lambda_{rs} = \frac{\mu_{rs}}{\mu_{rs}^2}, \mu_{rs} = \frac{1}{N-1} \sum_{i=1}^N (Y - \bar{Y})^r (X - \bar{X})^s$$

### 3.1. Existing Estimators from the literature;

#### 3.2 Ratio type Variance estimator proposed by Isaki [1]:

$$\hat{S}_R^2 = s_y^2 \frac{S_x^2}{s_x^2}$$

$$\text{Bias} ((\hat{S}_R^2)) = \gamma S_y^2 \left[ (\beta_{2(x)} - 1) - (\lambda_{22} - 1) \right]$$

$$\text{MSE} ((\hat{S}_R^2)) = \gamma S_y^4 \left[ (\beta_{2(y)} - 1) + (\beta_{2(x)} - 1) - 2(\lambda_{22} - 1) \right]$$

#### 3.3 Ratio type Variance estimator proposed by Upadhaya (1999) [2]:

$$\hat{S}_{US}^2 = s_y^2 \left[ \frac{S_x^2 + \beta_{2x}}{s_x^2 + \beta_{2x}} \right]$$

$$\text{Bias} ((\hat{S}_{US}^2)) = \gamma S_y^2 A_{US} \left[ A_{US} (\beta_{2(x)} - 1) - (\lambda_{22} - 1) \right]$$

$$\text{MSE} ((\hat{S}_{US}^2)) = \gamma S_y^4 \left[ (\beta_{2(y)} - 1) + A_{US}^2 (\beta_{2(x)} - 1) - 2A_{US} (\lambda_{22} - 1) \right]$$

#### 3.4 Ratio type Variance estimator proposed by Kadilar and Cingi [3]:

$$\hat{S}_{kc1}^2 = s_y^2 \left[ \frac{S_x^2 + C_x}{s_x^2 + C_x} \right]$$

$$\text{Bias} ((\hat{S}_{kc1}^2)) = \gamma S_y^2 A_1 [A_1 (\beta_{2(x)} - 1) - (\lambda_{22} - 1)]$$

$$\text{MSE} ((\hat{S}_{kc1}^2)) = \gamma S_y^4 [(\beta_{2(y)} - 1) + A_1^2 (\beta_{2(x)} - 1) - 2A_1 (\lambda_{22} - 1)]$$

### 3.5 Ratio type Variance estimator proposed by Subramani and Kumarapandiyan [4]:

$$\hat{S}_{kc1}^2 = s_y^2 \left[ \frac{S_x^2 + Q_d}{s_x^2 + Q_d} \right]$$

$$\text{Bias} ((\hat{S}_{kc1}^2)) = \gamma S_y^2 A_1 [A_1 (\beta_{2(x)} - 1) - (\lambda_{22} - 1)]$$

$$\text{MSE} ((\hat{S}_{kc1}^2)) = \gamma S_y^4 [(\beta_{2(y)} - 1) + A_1^2 (\beta_{2(x)} - 1) - 2A_1 (\lambda_{22} - 1)]$$

### 4. Proposed estimators

$$1. \hat{S}_{MS1}^2 = s_y^2 \left[ \frac{S_x^2 + MR + \left\{ \frac{S_x + C_x}{\beta_{1x}} \right\}}{s_x^2 + MR + \left\{ \frac{S_x + C_x}{\beta_{1x}} \right\}} \right]$$

$$2. \hat{S}_{MS1}^2 = s_y^2 \left[ \frac{S_x^2 + MR + \left\{ \frac{S_x + \beta_{2x}}{\beta_{1x}} \right\}}{s_x^2 + MR + \left\{ \frac{S_x + \beta_{2x}}{\beta_{1x}} \right\}} \right]$$

$$3. \hat{S}_{MS1}^2 = s_y^2 \left[ \frac{S_x^2 + MR + \left\{ \frac{S_x + HL}{\beta_{1x}} \right\}}{s_x^2 + MR + \left\{ \frac{S_x + HL}{\beta_{1x}} \right\}} \right]$$

### 5. We have derived here the bias and mean square error of the proposed estimator

$\hat{S}_{MSi}^2$ ;  $i = 1, 2$  to first order of approximation as given below:

Let  $e_0 = \frac{s_y^2 - S_y^2}{S_y^2}$  and  $e_1 = \frac{s_x^2 - S_x^2}{S_x^2}$ . Further we can write  $s_y^2 = S_y^2(1 + e_0)$  and  $s_x^2 = S_x^2(1 + e_1)$

and from the definition of  $e_0$  and  $e_1$  we obtain:

$$E[e_0] = E[e_1] = 0, \quad E[e_0^2] = \frac{1-f}{n} (\beta_{2(y)} - 1), \quad E[e_1^2] = \frac{1-f}{n} (\beta_{2(x)} - 1),$$

$$E[e_0 e_1] = \frac{1-f}{n} (\lambda_{21} - 1)$$

The proposed estimator  $\hat{S}_{MSi}^2$ ;  $i = 1, 2, 3$  is given below:

$$\hat{S}_{MSi}^2 = s_y^2 \left[ \frac{S_x^2 + \alpha a_i}{s_x^2 + \alpha a_i} \right] \quad (5.1)$$

$$\Rightarrow \hat{S}_{MSi}^2 = s_y^2 (1 + e_0) \left[ \frac{S_x^2 + \alpha a_i}{s_x^2 + e_1 S_x^2 + \alpha a_i} \right] \Rightarrow \hat{S}_{MSi}^2 = \frac{S_y^2 (1 + e_0)}{(1 + A_{MSi} e_1)} \text{ Where } A_{MSi} = \frac{S_x^2}{S_x^2 + \alpha a_i},$$

$$a_i = \left[ \left( \frac{S_x + C_x}{\beta_{1x}} \right) \right], \left[ \left( \frac{S_x + \beta_{2x}}{\beta_{1x}} \right) \right], \left[ \left( \frac{S_x + HL}{\beta_{1x}} \right) \right]; i = 1, 2, 3$$

$$\Rightarrow \hat{S}_{MSi}^2 = S_y^2 (1 + e_0) (1 + A_{MSi} e_1)^{-1} \quad (5.2)$$

$$\Rightarrow \hat{S}_{MSi}^2 = S_y^2 (1 + e_0) (1 - A_{MSi} e_1 + A_{MSi}^2 e_1^2 - A_{MSi}^3 e_1^3 + \dots) \quad (5.3)$$

Expanding and neglecting the terms more than 3<sup>rd</sup> order, we get

$$\hat{S}_{MSi}^2 = S_y^2 + S_y^2 e_0 - S_y^2 A_{MSi} e_1 - S_y^2 A_{MSi} e_0 e_1 + S_y^2 A_{MSi}^2 e_1^2 \quad (5.4)$$

$$\hat{S}_{MSi}^2 - S_y^2 = S_y^2 e_0 - S_y^2 A_{MSi} e_1 - S_y^2 A_{MSi} e_0 e_1 + S_y^2 A_{MSi}^2 e_1^2 \quad (5.5)$$

By taking expectation on both sides of (5.5), we get

$$E(\hat{S}_{MSi}^2 - S_y^2) = S_y^2 E(e_0) - S_y^2 A_{MSi} E(e_1) - S_y^2 A_{MSi} E(e_0 e_1) + S_y^2 A_{MSi}^2 E(e_1^2) \quad (5.6)$$

$$Bias(\hat{S}_{MSi}^2) = S_y^2 A_{MSi}^2 E(e_1^2) - S_y^2 A_{MSi} E(e_0 e_1) \quad (5.7)$$

$$Bias(\hat{S}_{MSi}^2) = \gamma S_y^2 A_{MSi} [A_{MSi} (\beta_{2(x)} - 1) - (\lambda_{21} - 1)] \quad (5.8)$$

Squaring both sides of (5.5) and neglecting the terms more than 2<sup>nd</sup> order and taking expectation, we get

$$E(\hat{S}_{MSi}^2 - S_y^2)^2 = S_y^4 E(e_0^2) + S_y^4 A_{MSi}^2 E(e_1^2) - 2S_y^4 A_{MSi} E(e_0 e_1) \quad (5.9)$$

$$MSE(\hat{S}_{MSi}^2) = \gamma S_y^4 [(\beta_{2(y)} - 1) + A_{MSi}^2 (\beta_{2(x)} - 1) - 2A_{MSi} (\lambda_{21} - 1)] \quad (5.10)$$

**6. Efficiency conditions:** Here, we have derived the efficiency conditions of proposed estimators with other existing estimators under which proposed estimators have performed better than the existing estimators

The bias and Mean square error of existing ratio type estimators up to the first order of approximation is given by

$$Bias(\hat{S}_K^2) = \gamma S_y^2 \phi_K [\phi_K (\beta_{2x} - 1) - (\lambda_{21} - 1)] \quad (6.1)$$

$$MSE(\hat{S}_K^2) = \gamma S_y^4 [(\beta_{2y} - 1) + \phi_K^2 (\beta_{2x} - 1) - 2\phi_K (\lambda_{21} - 1)] \quad (6.2)$$

$\phi_K = \text{Existing constant}$

,  $K = 1, 2, 3, 4, \dots$

Bias, MSE and constant of proposed estimators is given by

$$Bias(\hat{S}_P^2) = \gamma S_y^2 \phi_P [\phi_P (\beta_{2x} - 1) - (\lambda_{21} - 1)] \quad (6.3)$$

$$MSE(\hat{S}_P^2) = \gamma S_y^4 [(\beta_{2y} - 1) + \phi_P^2(\beta_{2x} - 1) - 2\phi_P(\lambda_{21} - 1)] \quad (6.4)$$

$\phi_P = \text{proposed constant}$

$P = 1, 2, 3, \dots$

From Equation (6.2) and (6.3), we have

$$MSE(\hat{S}_P^2) \leq MSE(\hat{S}_K^2) \text{ if } \lambda_{21} \geq 1 + \frac{(\phi_P + \phi_K)(\beta_{2x} - 1)}{2}$$

$$MSE(\hat{S}_P^2) \leq MSE(\hat{S}_K^2)$$

$$\gamma S_y^4 [(\beta_{2y} - 1) + \phi_P^2(\beta_{2x} - 1) - 2\phi_P(\lambda_{21} - 1)] \leq \gamma S_y^4 [(\beta_{2y} - 1) + \phi_K^2(\beta_{2x} - 1) - 2\phi_K(\lambda_{21} - 1)] \quad (6.5)$$

$$\Rightarrow [(\beta_{2y} - 1) + \phi_P^2(\beta_{2x} - 1) - 2\phi_P(\lambda_{21} - 1)] \leq [(\beta_{2y} - 1) + \phi_K^2(\beta_{2x} - 1) - 2\phi_K(\lambda_{21} - 1)] \quad (6.6)$$

$$\Rightarrow [\phi_P^2(\beta_{2x} - 1) - 2\phi_P(\lambda_{21} - 1)] \leq [\phi_K^2(\beta_{2x} - 1) - 2\phi_K(\lambda_{21} - 1)] \quad (6.7)$$

$$\Rightarrow (\beta_{2x} - 1)(\phi_P^2 - \phi_K^2) [-2\phi_P(\lambda_{21} - 1)] \leq [-2\phi_K(\lambda_{21} - 1)] \quad (6.8)$$

$$\Rightarrow (\beta_{2x} - 1)(\phi_P^2 - \phi_K^2) [-2(\lambda_{21} - 1)(\phi_P - \phi_K)] \leq 0 \quad (6.9)$$

$$\Rightarrow (\beta_{2x} - 1)(\phi_P^2 - \phi_K^2) \leq [2(\lambda_{21} - 1)(\phi_P - \phi_K)] \quad (6.10)$$

$$\Rightarrow (\beta_{2x} - 1) \leq \frac{2(\lambda_{21} - 1)(\phi_P - \phi_K)}{(\phi_P^2 - \phi_K^2)} \quad (6.11)$$

$$\Rightarrow (\beta_{2x} - 1) \leq \frac{2(\lambda_{21} - 1)(\phi_P - \phi_K)}{(\phi_P - \phi_K)(\phi_P + \phi_K)} \quad (6.12)$$

$$\Rightarrow (\beta_{2x} - 1)(\phi_P + \phi_K) \leq 2(\lambda_{21} - 1) \quad (6.13)$$

By solving equation (6.13), we get

$$MSE(\hat{S}_P^2) \leq MSE(\hat{S}_K^2) \text{ if } \lambda_{21} \geq 1 + \frac{(\phi_P + \phi_K)(\beta_{2x} - 1)}{2}$$

## 7. Numerical Illustration:

We use the data set of Murthy [13] page 228 in which fixed capital is denoted by X (auxiliary variable) and output of 80 factories is denoted by Y (study variable). The Population parameters are given below

$$N = 80, n = 20, \rho = 0.9413, \bar{X} = 11.2646, \bar{Y} = 51.8264, S_y = 18.3549, S_x = 8.4563, C_x = 0.7507$$

$$C_y = 0.3542, \beta_{1x} = 1.05, \beta_{2x} = 2.8664, \beta_{2y} = 2.2667, Q_1 = 5.1500, Q_2 = 7.5750, Q_3 = 16.975, Q_d = 5.9125,$$

$$Q_r = 11.825, Q_a = 11.0625, HL = 10.405, MR = 17.955, \lambda_{22} = 2.2209$$

Table-1- Bias and mean square error of existing and proposed estimators

| Estimators                   | Population-1 |         |
|------------------------------|--------------|---------|
|                              | Bias         | MSE     |
| Existing Estimators          |              |         |
| Isaki (1983)                 | 14.92        | 7705.22 |
| Upadhyaya and Singh (1999)   | 13.50        | 6327.13 |
| Kadilar and Cingi (2006)     | 14.79        | 7665.78 |
| Subramani and Kumarapandiyan | 12.41        | 6207.29 |
| Proposed Estimators          | Bias         | MSE     |
| MS-1                         | 1.64248      | 2700.69 |
| MS-2                         | 1.27416      | 2648.04 |
| MS-3                         | 0.18925      | 2625.53 |

Table-2-Percent Relative efficiency of proposed estimators over existing estimators

| MSE of Existing Estimators | MSE of Proposed Estimators | PRE%   |
|----------------------------|----------------------------|--------|
| 7705.22                    | 2700.69                    | 285.30 |
| 7705.22                    | 2648.04                    | 290.97 |
| 7705.22                    | 2625.53                    | 293.47 |
| 6327.13                    | 2700.69                    | 234.27 |
| 6327.13                    | 2648.04                    | 238.93 |
| 6327.13                    | 2625.53                    | 240.98 |
| 7665.78                    | 2700.69                    | 283.84 |
| 7665.78                    | 2648.04                    | 289.48 |
| 7665.78                    | 2625.53                    | 291.97 |
| 6207.29                    | 2700.69                    | 229.84 |
| 6207.29                    | 2648.04                    | 234.41 |
| 6207.29                    | 2625.53                    | 236.42 |
|                            |                            |        |

### Discussion:

The new proposed efficient estimators modified by incorporating the linear combination of conventional and non conventional parameters as auxiliary information have performed better than the existing estimators. The improvement can be easily accessed from the table-1 and table-2, by comparing the bias and MSE and PRE of existing and new modified estimators.

Hence new modified estimators may be preferred over existing estimators for use in practical applications.

## References

- [1]. Isaki, C.T. (1983). Variance estimation using auxiliary information. *Journal of the American Statistical Association*, 78, 117-123.
- [2]. Upadhyaya, L. N. and Singh, H. P, (1999): use of auxiliary variable in the estimation of population variance, *Mathematical forum*, 4, 33-36 (1936).
- [3]. Kadilar, C. & Cingi, H. (2006a). Improvement in Variance estimation using auxiliary information Hacettepe *Journal of mathematics and Statistics*, 35(1). 117-115.
- [4]. Cochran, W.G (1977), Sampling Techniques. New-York, John Wiley and Sons.
- [5]. Subramani, j. and Kumarapandiyam (2016), Variance estimation using quartiles and their functions of an auxiliary variable, *International journal of Statistics and applications*, 2012, Vol.2(5), 67-42
- [7]. Prasad, B. and Singh, H.P. (1990) 'Some improved ratio type estimators of finite population variance in sample surveys', *Communication in Statistics: Theory and methods*, Vol. 19, pp.1127-1139
- [8] M .a .Bhat<sup>\*1</sup>, S. Maqbool<sup>2</sup>, S .A .Saraf<sup>2</sup>, Ab. Rouf<sup>2</sup> and S .H Malik<sup>2</sup>, NLMID: 101666096. *Journal of Advances in Research* 13(2): 1-6, 2018 Article no.37321. ISSN: 2348-0394
- [9]. Sarginder Singh; *Advance Sampling theory with Applications* (2003), page (1135-1136). Publisher; Sprissnger-Science, business.
- Khan. and Shabir, j (2013) 'A ratio type estimator for the estimation of population variance using quartiles of an auxiliary variable', *Journal of Statistics applications & Probability*, Vol.2, No. 3, pp.319-325
- [10]. Singh, H.P., Tailor, R., Tailor, R. and Kakran, M.S. An improved estimator of population mean using power transformation, *Journal of the Indian Society of Agricultural Statistics* 33(2), 13-18, 1981
- [11]. M.A.Bhat<sup>1\*</sup>, S.Maqbool<sup>1</sup>, Ab. Rauf<sup>1</sup>, S.A.Saraf<sup>1</sup> and M.A.Malik<sup>1</sup>, Variance estimation Using Linear Combination of Standard Deviation and Deciles. *Current journal of Applied Science and Tecnology* 25(6), 2017; Article No. CCAST.37323
- [12]. M .A.Bhat<sup>1\*</sup>, S.Maqbool<sup>1</sup>, S.A.Saraf<sup>1</sup>, S.H. Malik<sup>2#</sup>, Yasmeeana Ismail<sup>3</sup> and Mir Subzar<sup>3</sup> Improved and Robust Estimators for Finite Population Variance Using Linear Combination of Probability Weighted Moment and Quartiles as Auxiliary Information. *Journal of Scientific Research & Reports*; 18(4); 1-6, 2018; Article No. JSRR.40116. ISSN: 2320-0227
- [13]. M .A.Bhat<sup>1\*</sup>, S.Maqbool<sup>1</sup>, S.A.Mir<sup>1</sup>, N.A.Sofi<sup>1</sup>, Ab.Rauf<sup>1</sup>, Immad .A. Shah<sup>1</sup> and Mir Subzar<sup>1</sup>. An Improvement of Robust Estimator Using Known Values of Probability weighted moment for Finite Population Variance. *Asian Journal of Agricultural Extension, Economics & Sociology*. 25(2): 1-5, 2018; Article no. AJAEES.41961. ISSN: 2320-7027
- [13]. M.A.Bhat<sup>\*</sup>, S.Maqbool and Mir Subzar. An Improvement in Variance Estimator for the Estimation of Population Variance, Using Known Values of Auxiliary Information. *International Journal of Pure & Applied Bioscience*. Int.J.Pure App. Biosci. 6(5): 135-138 (2018). ISSN: 2320-7051
- [14]. M.A.Bhat, S.Maqbool, T.A.Raja and N.A. Ganai<sup>1</sup>. Int.J. VARIANCE ESTIMATION USING LINEAR COMBINATION OF GINI'S MEAN DIFFERENCE AND DECILES. *Int.J.Agricult.Stat.Sci*. Vol.13, No.2, pp.527-530, 2017; ISSN: 0973-1903.

[15]. M .A.Bhat<sup>1\*</sup>,T.A. Raja<sup>1</sup>, S.Maqbool<sup>1</sup>,N.A.Sofi<sup>1</sup>,Ab.Rauf<sup>1</sup> ,S.H.Baba<sup>1</sup>and Immad .A. Shah<sup>1</sup>. Robust Estimators for Estimation of Population Variance Using Linear Combination of Downton's Method and Deciles as Auxiliary iInformation. Advances in Research; 15(20: 1-7,2018; Article no. AIR.41956. ISSN:2348-0394,NLM ID 101666096

UNDER PEER REVIEW