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 of auxiliary information. Bias and mean square error have been computed up to the first order of approximation. Numerical demonstration has been carried out to test the efficiency of new proposed estimators against existing 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 estimatorsforthe 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 skewnes 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 bySubramani J. and Kumarapandiyan (201b)[4]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 $U = \{U_1, U_2, ..., U_N\}$, consists of N distinct and identifiable units. Let Y be a real variable with value Y_i measured on$U_i, i = 1,2,3........N$, giving a vector $Y = \{y_1, y_2, ..., y_N\}$. The goal is to estimate the populations mean $\bar{Y} = \frac{1}{N}\sum_{I=1}^{N} y_i$ and its variance $S_Y^2 = \frac{1}{N-1}\sum_{I=1}^{N}\left(y_i - \bar{y}\right)^2$ on the basis of random sample selected from a finite population 'U'.

$N = Population\ size$,

$n = sample \quad size$.

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

$Y = Study$ 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 = *Cofficient of variations*,

ρ = *Cofficient of Correlation between auxiliary and study variable*,

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

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

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

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

3.0 Existing Estimators from the literature;

3.1 Ratio type Variance estimator proposed by Isaki [1]:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4. Proposedestimators: 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:-

1. $\hat{S}_{MS_1}^2 = s_y^2 \left[\dfrac{S_x^2 + MR + \{\dfrac{S_x + C_x}{\beta_{1x}}\}}{s_x^2 + MR + \{\dfrac{S_x + C_x}{\beta_{1x}}\}} \right]$

2. $\hat{S}_{MS_1}^2 = s_y^2 \left[\dfrac{S_x^2 + MR + \{\dfrac{S_x + \beta_{2x}}{\beta_{1x}}\}}{s_x^2 + MR + \{\dfrac{S_x + \beta_{2x}}{\beta_{1x}}\}} \right]$

3. $\hat{S}_{MS_1}^2 = s_y^2 \left[\dfrac{S_x^2 + MR + \{\dfrac{S_x + HL}{\beta_{1x}}\}}{s_x^2 + MR + \{\dfrac{S_x + HL}{\beta_{1x}}\}} \right]$

Let $e_0 = \dfrac{s_y^2 - S_y^2}{S_y^2}$ and $e_1 = \dfrac{s_x^2 - S_x^2}{S_x^2}$. 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:

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

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

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

$$\hat{S}^2_{MSi} = s^2_y\left[\frac{S^2_x + \alpha a_i}{s^2_x + \alpha a_i}\right] \tag{4.1}$$

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

$$a_i = [(\frac{S_x + C_x}{\beta_{1x}})],[(\frac{S_x + \beta_{2x}}{\beta_{1x}})],[(\frac{S_x + HL}{\beta_{1x}})];\ i = 1,23$$

$$\Rightarrow \hat{S}^2_{MSi} = S^2_y(1+e_0)(1+A_{MSi}e_1)^{-1} \tag{4.2}$$

$$\Rightarrow \hat{S}^2_{MSi} = S^2_y(1+e_0)(1 - A_{MSi}e_1 + A^2_{MSi}e^2_1 - A^3_{MSi}e^3_1 + ...) \tag{4.3}$$

Expanding and neglecting the terms more than 3rd order, we get

$$\hat{S}^2_{MSi} = S^2_y + S^2_y e_0 - S^2_y A_{MSi}e_1 - S^2_y A_{MSi}e_0 e_1 + S^2_y A^2_{MSi}e^2_1 \tag{4.4}$$

$$\hat{S}^2_{MSi} - S^2_y = S^2_y e_0 - S^2_y A_{MSi}e_1 - S^2_y A_{MSi}e_0 e_1 + S^2_y A^2_{MSi}e^2_1 \tag{4.5}$$

By taking expectation on both sides of (4.5), we get

$$E(\hat{S}^2_{MSi} - S^2_y) = S^2_y E(e_0) - S^2_y A_{MSi}E(e_1) - S^2_y A_{MSi}E(e_0 e_1) + S^2_y A^2_{MSi}E(e^2_1) \tag{4.6}$$

$$Bias(\hat{S}^2_{MSi}) = S^2_y A^2_{MSi}E(e^2_1) - S^2_y A_{MSi}E(e_0 e_1) \tag{4.7}$$

$$Bias(\hat{S}^2_{MSi}) = \gamma\, S^2_y A_{MSi}[A_{MSi}(\beta_{2(x)} - 1) - (\lambda_{21} - 1)] \tag{4.8}$$

Squaring both sides of (4.6) and neglecting the terms more than 2nd order and taking expectation, we get

$$E(\hat{S}^2_{MSi} - S^2_y)^2 = S^4_y E(e^2_0) + S^4_y A^2_{MSi}E(e^2_1) - 2S^4_y A_{MSi}E(e_0 e_1) \tag{4.9}$$

$$MSE(\hat{S}^2_{MSi}) = \gamma\, S^4_y[(\beta_{2(y)} - 1) + A^2_{MSi}(\beta_{2(x)} - 1) - 2A_{MSi}(\lambda_{21} - 1)] \tag{4.10}$$

5. 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

$$Bias\left(\hat{S}^2_K\right) = \gamma S^2_y \phi_K\left[\phi_K(\beta_{2x} - 1) - (\lambda_{21} - 1)\right] \tag{5.1}$$

$$MSE\left(\hat{S}^2_K\right) = \gamma S^4_y\left[(\beta_{2y} - 1) + \phi^2_K(\beta_{2x} - 1) - 2\phi_K(\lambda_{21} - 1)\right] \tag{5.2}$$

$\phi_K = Existing.cons\tan t$

$, K = 1,2,3,4........$

Bias, MSE and constant of proposed estimators is given by

$$Bias\left(\hat{S}_P^2\right)=\gamma S_y^2\phi_P\left[\phi_P\left(\beta_{2x}-1\right)-\left(\lambda_{21}-1\right)\right] \quad (5.3)$$

$$MSE\left(\hat{S}_P^2\right)=\gamma S_y^4\left[\left(\beta_{2y}-1\right)+\phi_P^2\left(\beta_{2x}-1\right)-2\phi_P\left(\lambda_{21}-1\right)\right] \quad (5.4)$$

$\phi_P = proposed.cons\tan t$

$P = 1,2,3............$

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

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

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

$$\gamma S_y^4\left[\left(\beta_{2y}-1\right)+\phi_P^2\left(\beta_{2x}-1\right)-2\phi_P\left(\lambda_{21}-1\right)\right]\leq \gamma S_y^4\left[\left(\beta_{2y}-1\right)+\phi_K^2\left(\beta_{2x}-1\right)-2\phi_K\left(\lambda_{21}-1\right)\right] \quad (5.5)$$

$$\Rightarrow \left[\left(\beta_{2y}-1\right)+\phi_P^2\left(\beta_{2x}-1\right)-2\phi_P\left(\lambda_{21}-1\right)\right]\leq \left[\left(\beta_{2y}-1\right)+\phi_K^2\left(\beta_{2x}-1\right)-2\phi_K\left(\lambda_{21}-1\right)\right] \quad (5.6)$$

$$\Rightarrow \left[\phi_P^2\left(\beta_{2x}-1\right)-2\phi_P\left(\lambda_{21}-1\right)\right]\leq \left[\phi_K^2\left(\beta_{2x}-1\right)-2\phi_K\left(\lambda_{21}-1\right)\right] \quad (5.7)$$

$$\Rightarrow \left(\beta_{2x}-1\right)\left(\phi_P^2-\phi_K^2\right)\left[-2\phi_P\left(\lambda_{21}-1\right)\right]\leq \left[-2\phi_K\left(\lambda_{21}-1\right)\right] \quad (5.8)$$

$$\Rightarrow \left(\beta_{2x}-1\right)\left(\phi_P^2-\phi_K^2\right)\left[-2\left(\lambda_{21}-1\right)\left(\phi_P-\phi_K\right)\right]\leq 0 \quad (5.9)$$

$$\Rightarrow \left(\beta_{2x}-1\right)\left(\phi_P^2-\phi_K^2\right) \leq \left[2\left(\lambda_{21}-1\right)\left(\phi_P-\phi_K\right)\right] \quad (5.10)$$

$$\Rightarrow \left(\beta_{2x}-1\right)\leq \frac{2\left(\lambda_{21}-1\right)\left(\phi_P-\phi_K\right)}{\left(\phi_P^2-\phi_K^2\right)} \quad (5.11)$$

$$\Rightarrow \left(\beta_{2x}-1\right)\leq \frac{2\left(\lambda_{21}-1\right)\left(\phi_P-\phi_K\right)}{\left(\phi_P-\phi_K\right)\left(\phi_P+\phi_K\right)} \quad (5.12)$$

$$\Rightarrow \left(\beta_{2x}-1\right)\left(\phi_P+\phi_K\right)\leq 2\left(\lambda_{21}-1\right) \quad (5.13)$$

By solving equation (5.13), we get

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

6. Numerical Illustration:

We use the data set of Murthy [13] page228 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

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

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

$Q_r=11.825,\ Q_a=11.0625\ , HL=10.405, MR=17{,}955,\ \lambda_{22}=2.2209$

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

Estimators	Population-1		PRE%
Existing Estimators	Bias	MSE	PRE%
Isaki (1983)	14.92	7705.22	285.30
Upadhyaya and Singh (1999)	13.50	6327.13	290.97
Kadilar and Cingi (2006)	14.79	7665.78	293.47
Subramani and Kumarapandiyan	12.41	6207.29	234.27
Proposed Estimators	Bias	MSE	238,93
Proposed Estimator-1	1.64248	2700.69	240.98
Proposed Estimator-2	1.27416	2648.04	283.84
Proposed Estimator-3	0.18925	2625.53	289.48

Conclusion**:**

The new proposed efficient estimatorsmodified 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.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1.

2.

3.

References

[1].Isaki, C.T. (1983).Variance estimation using auxiliary information. *Journal of the American Statistical Association,78,117-123.*

[2]. Upadhaya, 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 informationHacettepe*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 Kumarapandiyan(201b), Variance estimation using quartiles and their functions of an auxiliary variable, International journal of Statistics and applications,2012,Vol.2(5),67-42

[6]. Prasad,B.andSingh,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

[7] M .a .Bhat[*1], S. Maqbool[2], S .A .Saraf[2], Ab. Rouf[2]and S .H Malik[2], NLMID: 101666096. *Journal of Advances in Research 13(2): 1-6, 2018 Article no.37321. ISSN: 2348-0394*

[8]. Sarginder Singh; *Advance Sampling theory with Applications (2003), page (1135-1136). Publisher; Sprissnger-Science, business.*

[9]. 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[1*], S.Maqbool[1],Ab. Rauf[1], S.A.Saraf[1]and M.A.Malik[1], Variance estimation Using Linear Combination of Standard Deviation and Deciles. Current journal of Applied Science and Tecnology 25(6),2017; Article No. CJAST.37323

[12]. M .A.Bhat[1*], S.Maqbool[1], S.A.Saraf[1],S.H..[Malik2#],Yasmeena Ismail[3] and Mir Subzar[3] Improved and Robust Estimators for Finite Population Variance Using Linear Combination of Probability Weighted Moment and Quartiles as Auxiliary Information. Journaal of Scientific Research & Reports; 18(4);1-6,2018; Article No. JSRR.40116. ISSN: 2320-0227

[13].M .A.Bhat[1*], S.Maqbool[1], S.A.Mir[1],N.A.Sofi[1],Ab.Rauf[1] , Immad .A. Shah[1] and Mir Subzar[1]. 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

[14]. M.A.Bhat[*] ,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