Original Research paper

Robust Estimators for estimation of population variance using non-conventional population Parameters.

ABSTRACT. We have suggested improved and robust estimators for the estimation of finite population variance using non-conventional population parameters as auxiliary variables to enhance the efficiency of proposed estimators. A comparison between suggested estimators and existing estimators has been made through a numerical illustration to seek the efficiency of proposed estimators over existing estimators. The expression for bias and mean square error has been derived up to the first order of approximation. The improvement of proposed estimators over existing estimators shown is clearly based on the lesser mean square error of proposed estimators.

Key words: Simple random sampling, bias, mean square error, Downton's method, Deciles and efficiency.

1. Introduction; The improvement of the estimators through proper utilization of auxiliary information has been widely discussed by the Staticians when there exists a close association between auxiliary variable (X) and Study variable (Y). Some of them from the literature are Isaki [1], who proposed ratio and regression estimators. Kadilar & Cingi, (H.2006a) [2] utilized the coefficient of skewness C_x as auxiliary population parameter to enhance the efficiency of estimator. Subramani. j and Kumarapandiyan .G [3] used quartiles as auxiliary information to improve the efficiency of modified estimators over existing estimators. On the same lines, Singh. D, and Chaudhary, F.S [4], M .Murthy [5], Arcos. A . M. Rueda, M. D, Martinez. S . Gonzalez and Y. Roman [6], have utilized this auxiliary information in different forms to enhance the precision and efficiency of proposed estimators. Recently, Subhash Kumar Yadav [7], Khan. M and Shabbir. J [8], Jeelani. Iqbal and Maqbool. S [9], have used different population parameters as auxiliary variables to improve the precision and efficiency of variance estimators. Similarly Bhat et al. (2018) [10] have used linear combination of skewness and quartiles as auxiliary information to obtain the precision of estimators

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 $\bar{Y}=\frac{1}{N}\sum_{I=1}^{N} y_i$ or 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 population U... In this paper, our aim is to estimate the precise and reliable estimates for finite population variance when the population under investigation is non-normal or badly skewed as non-conventional parameters are robust measures for skewed populations.

2. Materials and methods:

2.1 Notations: $N=$ Population size. $n=$Sample size. $\gamma=\frac{1}{n}$, Y= study variable. X= Auxiliary variable. $\bar{X}, \bar{Y}=$ Population means. $\bar{x}, \bar{y}=$Sample means. $S_Y^2, S_x^2=$ population variances. $s_y^2, s_x^2=$ sample variances. $C_x, C_y=$Coefficient of variation. $\rho=$ Correlation coefficient. $\beta_{1(x)}=$ Skewness of the auxiliary variable. $\beta_{2(x)}=$ Kurtosis of the auxiliary variable. $\beta_{2(y)}=$ Kurtosis of the study variable. M_d= Median of the auxiliary variable. B(.)=Bias of the estimator. MSE(.)= Mean square error. $\hat{S}_R^2=$ Ratio type variance estimator. $\hat{S}_{Kc1}^2, \hat{S}_{jG}^2$,= Existing modified ratio estimators. *D=Downton;s method.*, D_i, i =1,2,...10 = Deciles.

2.2 Existing Estimators from the Literature

2.2.1. Ratio type Variance estimator proposed by Isaki [1]:

Isaki suggested a ratio type variance estimator for the population variance S_Y^2 when the population variance S_x^2 of the auxiliary variable X is known. Expressions for bias and mean square error are given as

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

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

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

2.2.2 Ratio type Variance estimator proposed by Kadilar and Cingi [2]:

In this existing estimator authors have suggested ratio type variance estimators where they have used known values of C.V and coefficient of kurtosis as auxiliary variables to improve the efficiency of estimators Expressions for bias and mean square error are given as

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

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

2.2.3 Ratio type Variance estimator proposed by Subramani.J and Kumarapandiyan.G [3], where the authors have used median, quartiles and Deciles as auxiliary variables to seek the precision and efficiency of estimators.

$$\hat{S}_{jG}^2 = s_y^2 \left[\frac{S_x^2 + \alpha w_i}{s_x^2 + \alpha w_i} \right]$$ Where α=1 and $w_i = D + D_i, i = 1,2,3............10$

.Expressions for bias and mean square error are given as

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

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

4.32 Modified and suggested estimators:

$$\hat{S}^2_{MS1} = s_y^2\left[\frac{S_x^2 + (D + D_1)}{s_x^2 + (D + D_1)}\right] \quad \hat{S}^2_{MS2} = s_y^2\left[\frac{S_x^2 + (D + D_2)}{s_x^2 + (D + D_2)}\right] \quad \hat{S}^2_{MS3} = s_y^2\left[\frac{S_x^2 + (D + D_3)}{s_x^2 + (D + D_3)}\right]$$

$$\hat{S}^2_{MS4} = s_y^2\left[\frac{S_x^2 + (D + D_4)}{s_x^2 + (D + D_4)}\right] \quad \hat{S}^2_{MS5} = s_y^2\left[\frac{S_x^2 + (D + D_5)}{s_x^2 + (D + D_5)}\right] \quad \hat{S}^2_{MS6} = s_y^2\left[\frac{S_x^2 + (D + D_6)}{s_x^2 + (D + D_6)}\right]$$

$$\hat{S}^2_{MS7} = s_y^2\left[\frac{S_x^2 + (D + D_7)}{s_x^2 + (D + D_7)}\right] \quad \hat{S}^2_{MS8} = s_y^2\left[\frac{S_x^2 + (D + D_8)}{s_x^2 + (D + D_8)}\right] \quad \hat{S}^2_{MS9} = s_y^2\left[\frac{S_x^2 + (D + D_9)}{s_x^2 + (D + D_9)}\right]$$

$$\hat{S}^2_{MS10} = s_y^2\left[\frac{S_x^2 + (D + D_{10})}{s_x^2 + (D + D_{10})}\right]$$

We have derived the bias and mean square error of proposed estimators $\hat{S}^2_{MSi}$; $i = 1,2,....,10$ up to the 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 $s_x^2 = S_x^2(1 + e_0)$ and from the definition of e_0 and e_1 we obtain:

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

$$E[e_0 e_1] = \frac{1-f}{n}(\lambda_{22} - 1)$$

The proposed estimator $\hat{S}^2_{MSi}$; $i = 1,2,3,....,10$ is given below:

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

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

$$A_{MSi} = \frac{S_x^2}{S_x^2 + \alpha a_i} \qquad a_i = (D + D_i);\ i = 1,2,3,...,10$$

$$and, \alpha = 1$$

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

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

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

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

$$\Rightarrow \quad \hat{S}^2_{MSi} - S^2_y = S^2_y e_0 - S^2_y A_{MSi}e_1 - S^2_y A_{MSi}e_0 e_1 + S^2_y A^2_{MSi}e^2_1 \qquad (5)$$

By taking expectation on both sides of (5), we get

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

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

$$Bias(\hat{S}^2_{MSi}) = \gamma\, S^2_y A_{MSi}[A_{MSi}(\beta_{2(x)}-1) - (\lambda_{22}-1)] \qquad (6)$$

Squaring both sides of (5) and (6), neglecting the terms more than 2nd order and taking expectation, we get

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

$$MSE(\hat{S}^2_{MSi}) = \gamma\, S^4_y[(\beta_{2(y)}-1) + A^2_{MSi}(\beta_{2(x)}-1) - 2A_{MSi}(\lambda_{22}-1)]$$

3. Efficiency conditions: we have derived the efficiency conditions of proposed estimators with other existing estimators under which proposed estimators $\hat{S}^2_P(P=1,2,3........)$ are performing better than the existing estimators $\hat{S}^2_K(K=1,2,3........)$

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

$$Bias(\hat{S}^2_K) = \gamma S^2_y R_K\left[R_K(\beta_{2x}-1) - (\lambda_{22}-1)\right] \qquad (1)$$

$$MSE(\hat{S}^2_K) = \gamma S^4_y\left[(\beta_{2y}-1) + R^2_K(\beta_{2x}-1) - 2R_K(\lambda_{22}-1)\right] \qquad (2)$$

Where $R_K = Existing.cons\tan t$, $K = 1,2,3,4........$

Bias, MSE and constant of proposed estimators is given by

$$Bias\left(\hat{S}_P^2\right) = \gamma S_y^2 R_P\left[R_P\left(\beta_{2x}-1\right)-\left(\lambda_{22}-1\right)\right] \quad (3)$$

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

$R_P = proposed.cons\tan t$

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

From Equation (2) and (3), we have

$$MSE\left(\hat{S}_P^2\right) \le MSE\left(\hat{S}_K^2\right) if \lambda_{22} \ge 1+\frac{\left(R_P+R_K\right)\left(\beta_{2x}-1\right)}{2}$$

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

$$\gamma S_y^4\left[\left(\beta_{2y}-1\right)+R_P^2\left(\beta_{2x}-1\right)-2R_P\left(\lambda_{22}-1\right)\right] \le \gamma S_y^4\left[\left(\beta_{2y}-1\right)+R_K^2\left(\beta_{2x}-1\right)-2R_K\left(\lambda_{22}-1\right)\right] \quad (5)$$

$$\Rightarrow \left[\left(\beta_{2y}-1\right)+R_P^2\left(\beta_{2x}-1\right)-2R_P\left(\lambda_{22}-1\right)\right] \le \left[\left(\beta_{2y}-1\right)+R_K^2\left(\beta_{2x}-1\right)-2R_K\left(\lambda_{22}-1\right)\right] \quad (6)$$

$$\Rightarrow \left[1+R_P^2\left(\beta_{2x}-1\right)-2R_P\left(\lambda_{22}-1\right)\right] \le \left[1+R_K^2\left(\beta_{2x}-1\right)-2R_K\left(\lambda_{22}-1\right)\right] \quad (7)$$

$$\Rightarrow \left(\beta_{2x}-1\right)\left(R_P^2-R_K^2\right)\left[-2R_P\left(\lambda_{22}-1\right)\right] \le \left[-2R_K\left(\lambda_{22}-1\right)\right] \quad (8)$$

$$\Rightarrow \left(\beta_{2x}-1\right)\left(R_P^2-R_K^2\right)\left[-2\left(\lambda_{22}-1\right)\left(R_P-R_K\right)\right] \le 0 \quad (9)$$

$$\Rightarrow \left(\beta_{2x}-1\right)\left(R_P^2-R_K^2\right) \le \left[2\left(\lambda_{22}-1\right)\left(R_P-R_K\right)\right]$$

(10)

$$\Rightarrow \left(\beta_{2x}-1\right) \le \frac{2\left(\lambda_{22}-1\right)\left(R_P-R_K\right)}{\left(R_P^2-R_K^2\right)}$$

(11)

$$\Rightarrow (\beta_{2x}-1) \leq \frac{2(\lambda_{22}-1)(R_P-R_K)}{(R_P-R_K)(R_P+R_K)} \qquad (12)$$

$$\Rightarrow (\beta_{2x}-1)\ (R_P+R_K) \leq 2(\lambda_{22}-1) \qquad (13)$$

By solving equation (13), we get

$$MSE(\hat{S}_P^2) \leq MSE(\hat{S}_K^2)\, if\, \lambda_{22} \geq 1+\frac{(R_P+R_K)(\beta_{2x}-1)}{2}$$

4. Numerical Illustration:

We use the data of Murthy (1967) page 228 in which fixed capital is denoted by X(auxiliary variable) and output of 80 factories are denoted by Y(study variable).we apply the proposed and existing estimators to this data set and the data statistics is given below:

N=80,Sx=8.4542, n=20, Cx=0.7507, $\overline{X}$ = 11.2624, $\beta_{2(x)}$ = 2.8664, $\overline{Y}$ = 51.8264, $\beta_{2(y)}$ = 2.2667, , ρ = 0.9413 , $\beta_{1(x)}$ = 1.05, λ_{22} = 2.2209, S_y=18.3569, D= 8.0138, D_1= 3.6, D_2 = 4.6,D_3= 5.9, D_4 = 6.7,

D_5 =7.5, D_6 = 8.5, D_7 =14.8, D_8 = 18.1, D_9 = 25,

D_{10} =34.8.

Table-1 Bias and Mean Square Error of existing and proposed estimators

Estimators	Bias	Mean Square Error
Isaki [1]	10.8762	3925.1622
Kadilar&Cingi [2]	10.4399	3850.1552
Subramani&Kumarapandiyan [3]	6.1235	3180.7740

Proposed (MS1)	4.1794	2330.0997
Proposed (MS2)	3.9258	2297.7372
Proposed (MS3)	3.6113	2258.5616
Proposed (MS4)	3.4257	2236.8446
Proposed (MS5)	3.2468	2215.5535
Proposed (MS6)	3.0290	2191.7075
Proposed (MS7)	1.8584	2078.8646
Proposed (MS8)	1.3512	2041.8180
Proposed (MS9)	0.4832	1999.2338
Proposed (MS10)	0.4143	1999.2353

Table-2 Percent relative efficiency of proposed estimators with existing estimators

Estimators	Isaki [1]	Kadilar&Cingi [2]	Subramani&Kumarapandiyan [3]
P1	168.4543	165.2356	136.5089
P2	170.8272	167.5629	138.4315
P3	173.7929	170.4693	140.8327
P4	175.4776	172.1243	142.2000
P5	177.1639	173.7784	143.5665
P6	179.0915	175.6692	145.1285

P7	188.8127	185.2047	153.0060
P8	192.2385	188.5650	155.7824
P9	196.3331	192.5813	159.1004
P10	196.2593	192.5089	159.0407

Decision and conclusion:

In this manuscript, empirical study clearly reveals that our proposed estimators are more efficient than existing estimators which can be seen from tables viz; table-1, and table-2 as bias and mean square error of suggested estimators is less than the already existing estimators in the literature. and also by the percentage relative efficiency criterion. Hence the proposed estimator may be preferred over existing estimators for use in practical applications.

Literature cited:

[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].Subramani, J. and Kumarapandiyan, G. (2015). Generalized modified ratio type estimator for estimation of population variance. *Sri-Lankan journal of applied Statistics,vol16-1,69-90.*

[5]. Singh, D. and Chaudhary, F. S. *Theory and analysis of sample survey designs. New age publishers 1986*

[6].M. N Murthy, Sampling theory, *Theory and Methods, Statistical publishing Society, Calcutta,1967.*

[7].Arcos, A., M. Rueda, M. D. Martinez, S. Gonzalez and Y. Roman. 2005. Incorporatin the auxiliary information available in variance estimation. *Applied Mathematical and Computation,* Vol **160**: 387-399.

[8]. Subhash Kumar Yadav, Sheela Misra and S.S Mishra . *American Journal of operational Research, 2016,6(1):9-15*

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