



International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)

(Open Access, Double Blind Peer-reviewed, Refereed and Indexed Journal)

www.iasir.net

Variance Estimation using Linear Combination of Hodges Lemma and Deciles

M. A. Bhat, S. Maqbool, T. A. Raja, and Mir Subzar
Division of Agricultural Statistics,

Sher-e-Kashmir University of Agricultural Sciences and Technology (SKUAST), Kashmir, INDIA.

Abstract: In this paper, we have proposed a class of modified ratio type variance estimator for estimation of population variance of the study variable, when Hodges Lehmann and Deciles of the auxiliary variable are known. The bias and mean square error (MSE) of the proposed estimator are obtained. From the numerical study it is observed that the proposed estimator performs better than the existing estimators in the literature.

Key words: Simple random sampling, bias, mean square error, hedges lemma and deciles, efficiency.

I. Introduction

Here we consider a finite population $U = \{U_1, U_2, \dots, U_N\}$ 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 population means $\bar{Y} = \frac{1}{N} \sum_{I=1}^N y_I$ 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 . Sometimes in sample surveys along with the study variable Y , information on auxiliary variable X , which is positively correlated with Y , is also available. The information on auxiliary variable X , may be utilized to obtain a more efficient estimator of the population. In this paper, our aim is to estimate the population variance on the basis of a random sample of size n selected from the population U . The problem of constructing efficient estimators for the population variance has been widely discussed by various authors such as Isaki (1983), who proposed ratio and regression estimators. Latter various authors such as Kadilar & Cingi (2006), Subramani & Kumarapandiyan (2015) improved the already existing estimators.

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. ρ = 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. QD = quartile deviation, G = Gini's Mean Difference, D = Downtown's method and S_{pw} = probability Weighted Moments, D_i , $i=1,2,\dots,10$ = Deciles.

II. Recent Developments

Subramani and Kumarapandiyan (2015) estimators where the authors used median, quartiles and Deciles of an auxiliary variable.

$$\hat{S}_{jG}^2 = s_y^2 \left[\frac{S_x^2 + \alpha w_i}{s_x^2 + \alpha w_i} \right] \quad (1)$$

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

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

III. Proposed estimator

We have proposed a new modified ratio type variance estimator of the auxiliary variable by using linear combination of hedges lemma and deciles.

$$\begin{aligned}\hat{S}_{MS1}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_1)}{s_x^2 + (HL + D_1)} \right] & \hat{S}_{MS2}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_2)}{s_x^2 + (HL + D_2)} \right] & \hat{S}_{MS3}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_3)}{s_x^2 + (HL + D_3)} \right] \\ \hat{S}_{MS4}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_4)}{s_x^2 + (HL + D_4)} \right] & \hat{S}_{MS5}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_5)}{s_x^2 + (HL + D_5)} \right] \\ \hat{S}_{MS6}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_6)}{s_x^2 + (HL + D_6)} \right] & \hat{S}_{MS7}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_7)}{s_x^2 + (HL + D_7)} \right] & \hat{S}_{MS8}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_8)}{s_x^2 + (HL + D_8)} \right] \\ \hat{S}_{MS9}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_9)}{s_x^2 + (HL + D_9)} \right] & \hat{S}_{MS10}^2 &= s_y^2 \left[\frac{S_x^2 + (HL + D_{10})}{s_x^2 + (HL + D_{10})} \right]\end{aligned}\quad (2)$$

We have derived here the bias and mean square error of the proposed estimator $\hat{S}_{MSi}^2$; $i = 1, 2, \dots, 10$ 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 $s_x^2 = S_x^2(1 + e_0)$ and

from the definition of e_0 and e_1 we obtain:

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

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

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

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

where $A_{MSi} = \frac{S_x^2}{S_x^2 + \alpha a_i}$ $a_i = (HL + D_i)$; $i = 1, 2, 3, \dots, 10$

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

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

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

$$\begin{aligned}\hat{S}_{MSi}^2 &= S_y^2 + S_y^2 e_0 - S_y^2 A_{MSi} e_1 - S_y^2 A_{MSi} e_0 e_1 + S_y^2 A_{MSi}^2 e_1^2 \\ \Rightarrow \hat{S}_{MSi}^2 - S_y^2 &= S_y^2 e_0 - S_y^2 A_{MSi} e_1 - S_y^2 A_{MSi} e_0 e_1 + S_y^2 A_{MSi}^2 e_1^2\end{aligned}\quad (3)$$

By taking expectation on both sides of (3), we get

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

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

$$Bias(\hat{S}_{MSi}^2) = \gamma S_y^2 A_{MSi} [A_{MSi} (\beta_{2(x)} - 1) - (\lambda_{22} - 1)] \quad (4)$$

Squaring both sides of (3) and (4), neglecting the terms more than 2nd order and taking expectation, we get

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

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

IV. 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:

$N=80$, $S_x=8.4542$, $n=20$, $C_x=0.7507$, $\bar{X} = 11.2624$, $\beta_{2(x)} = 2.8664$, $\bar{Y} = 51.8264$,
 $\beta_{2(y)} = 2.2667$, $\rho = 0.9413$, $\beta_{1(x)} = 1.05$, $\lambda_{22} = 2.2209$, $S_y=18.3569$, $D= 8.0138$, $D_1= 3.6$,
 $D_2 = 4.6$, $D_3= 5.9$, $D_4 = 6.7$, $D_5 = 7.5$, $D_6 = 8.5$, $D_7 = 14.8$, $D_8 = 18.1$, $D_9 = 25$, $D_{10} = 34.8$, $HL= 10.40$.

Bias and Mean Square Error of the existing and the proposed estimators

Estimators	Bias	Mean Square Error
Isaki (1983)	10.8762	3925.1622
Kadilar&Cingi(2006) 1	10.4399	3850.1552
Subramani&Kumarapandiyam(2015)	6.1235	3180.7740
Proposed (MS1)	4.7670	3007.9951
Proposed (MS2)	4.4580	2969.9951
Proposed (MS3)	4.0870	2927.9400
Proposed (MS4)	3.8370	2901.8230
Proposed (MS5)	3.6280	2874.5710
Proposed (MS6)	3.3660	2850.1577
Proposed (MS7)	1.9650	2734.9031
Proposed (MS8)	1.3650	2694.5913
Proposed (MS9)	0.2810	2655.9832
Proposed (MS10)	-0.8239	2611.6983

Percent relative efficiency of proposed estimators with existing estimators

Estimators	Isaki (1983)	Kadilar&Cingi(2006) 1	Subramani&Kumarapandiyam(2015)
P1	130.4909	129.6064	105.7479
P2	132.1625	131.2644	107.1010
P3	134.0588	133.1366	108.6359
P4	135.2653	134.3485	108.6700
P5	136.5477	135.6221	109.6136
P6	137.7173	136.7700	110.6903
P7	143.5210	142.5482	116.3404
P8	145.6681	144.6807	118.0807
P9	147.7856	146.7690	119.7972
P10	150.2915	149.2577	121.8285

V. Conclusion

From the empirical study we found that our proposed estimators are more efficient than existing estimators as bias and mean square error are 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.

References:

- Cochran, W. G. (1977). *Sampling Techniques*. Third Edition, Wiley Eastern limited.
- Isaki, C.T. (1983). Variance estimation using auxiliary information. *Journal of the American Statistical Association*, 78, 117-123.
- Kadilar, C. & Cingi, H. (2006). Improvement in Variance estimation using auxiliary information. *Haceteppe Journal of mathematics and Statistics*, 35(1), 117-115.
- Murthy, M. N. (1967). *Sampling theory and methods*. Calcutta Statistical Publishing House, India.
- Sumramani, J. and Kumarapandiyam, G. (2015). Generalized modified ratio type estimator for estimation of population variance. *Sri-Lankan journal of applied Statistics*, vol 16-1, 69-90.
- Wolter, K. M. (1985). *Introduction to variance estimation*. Springer- Verlag.