International Journal of Probability and Statistics 2012, 1(3): 62-66
DOI: 10.5923/j.ijps.20120103.02

Variance Estimation Using Median of the Auxiliary Variable

J. Subramani*, G. Kumarapandiyan

Department of Statistics, Ramanujan School of Mathematical Sciences, Pondicherry University, R V Nagar, Kalapet, 605014, Puducherry
kumarstat88@gmail.com

Abstract The present paper deals with a modified ratio type variance estimator for estimation of population variance of the study variable, when the population median of the auxiliary variable is known. The bias and the mean squared error of the proposed estimator are obtained and also derived the conditions for which the proposed estimator performs better than the traditional ratio type variance estimator suggested by Isaki[10] and the modified ratio type variance estimators suggested by Kadilar and Cingi[11]. Further we have compared the efficiencies of the proposed estimator with that of traditional ratio type variance estimator and existing modified ratio type variance estimators for certain known populations. From the numerical study it is observed that the proposed estimator performs better than the traditional ratio type variance estimator and existing modified ratio type variance estimators.

Keywords Bias, Mean Squared Error, Natural Populations, Simple Random Sampling

1. Introduction

Consider a finite population $U = \{U_1, U_2, \ldots, U_N\}$ of N distinct and identifiable units. Let Y be a real variable with value Y_i measured on $U_i, i = 1,2,3,\ldots,N$ giving a vector $Y = \{Y_1, Y_2, \ldots, Y_N\}$. The problem is to estimate the population mean $\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ on the basis of a random sample selected from the population U and / or its variance $S_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(Y_i - \overline{Y})^2$. When there is no additional information on the auxiliary variable available, the simplest estimator of population mean is the simple random sample mean without replacement. However if an auxiliary variable X closely related to the study variable Y is available then one can use Ratio or Regression estimators to improve the performance of the estimator of the study variable. In this paper, we consider the problem of estimation of the population variance and use the auxiliary information to improve the efficiency of the estimator of population variance $S_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(Y_i - \overline{Y})^2$.

Estimation of population variance is considered by Isaki[10] where ratio and regression estimators are proposed. Prasad and Singh[14] have considered a ratio type estimator for estimation of population variance by improving Isaki's estimator[10] with respect to bias and precision. Arcos et al.[14] have introduced another ratio type estimator, whichhas also improved the Isaki's estimator[10], which is almost unbiased and more precise than the other estimators.

Before discussing further about the traditional ratio type variance estimator, modified ratio type variance estimators and the proposed modified ratio type variance estimator, the notations to be used in this paper are described below:

$N -$ Population size
$n -$ Sample size
$\gamma = 1/n$
$Y -$ Study variable
$X -$ Auxiliary variable
$\overline{X}, \overline{Y} -$ Population means
$\overline{x}, \overline{y} -$ Sample means
$S_y^2, S_x^2 -$ Population variances
$s_y^2, s_x^2 -$ Sample variances
$C_X, C_y -$ Coefficient of variations
$\rho -$ Coefficient of correlation
$B(.) -$ Bias of the estimator
$MSE(.) -$ Mean squared error of the estimator
$\hat{S}_R^2 -$ Traditional Ratio type variance estimator of S_y^2
$\hat{S}_{KCi}^2 -$ Existing modified ratio type variance estimator of S_y^2
$\hat{S}_{SK}^2 -$ Proposed modified ratio type variance estimator of S_y^2

Isaki[10] 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 together with its bias and mean squared error as given below:

$$\hat{S}_R^2 = s_y^2 \frac{S_x^2}{s_x^2} \tag{1}$$

* Corresponding author:
drjsubramani@yahoo.co.in (Jambulingam Subramani)
Published online at http://journal.sapub.org/IJPS

Table 1. Existing modified ratio type variance estimators with their biases, mean squared errors

Estimator	Bias - $\mathbf{B}(.)$	Mean squared error $\mathbf{MSE}(.)$
$\hat{S}^2_{KC1} = s^2_y \left[\frac{S^2_x + C_x}{s^2_x + C_x}\right]$ Kadilar and Cingi[11]	$\gamma S^2_y A_1 \left[A_1\left(\beta_{2(x)} - 1\right) - (\lambda_{22} - 1)\right]$	$\gamma S^4_y \left[\left(\beta_{2(y)} - 1\right) + A^2_1\left(\beta_{2(x)} - 1\right) - 2A_1(\lambda_{22} - 1)\right]$
$\hat{S}^2_{KC2} = s^2_y \left[\frac{S^2_x + \beta_{2(x)}}{s^2_x + \beta_{2(x)}}\right]$ Kadilar and Cingi[11]	$\gamma S^2_y A_2 \left[A_2\left(\beta_{2(x)} - 1\right) - (\lambda_{22} - 1)\right]$	$\gamma S^4_y \left[\left(\beta_{2(y)} - 1\right) + A^2_2\left(\beta_{2(x)} - 1\right) - 2A_2(\lambda_{22} - 1)\right]$
$\hat{S}^2_{KC3} = s^2_y \left[\frac{S^2_x\beta_{2(x)} + C_x}{s^2_x\beta_{2(x)} + C_x}\right]$ Kadilar and Cingi[11]	$\gamma S^2_y A_3 \left[A_3\left(\beta_{2(x)} - 1\right) - (\lambda_{22} - 1)\right]$	$\gamma S^4_y \left[\left(\beta_{2(y)} - 1\right) + A^2_3\left(\beta_{2(x)} - 1\right) - 2A_3(\lambda_{22} - 1)\right]$
$\hat{S}^2_{KC4} = s^2_y \left[\frac{S^2_x C_x + \beta_{2(x)}}{s^2_x C_x + \beta_{2(x)}}\right]$ Kadilar and Cingi[11]	$\gamma S^2_y A_4 \left[A_4(\beta_{2(x)} - 1) - (\lambda_{22} - 1)\right]$	$\gamma S^4_y \left[\left(\beta_{2(y)} - 1\right) + A^2_4\left(\beta_{2(x)} - 1\right) - 2A_4(\lambda_{22} - 1)\right]$

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

where $\beta_{2(y)} = \frac{\mu_{40}}{\mu^2_{20}}, \beta_{2(x)} = \frac{\mu_{04}}{\mu^2_{02}}, \lambda_{22} = \frac{\mu_{22}}{\mu_{20}\mu_{02}}$ and

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

The Ratio type variance estimator given in (1) is used to improve the precision of the estimate of the population variance compared to simple random sampling when there exists a positive correlation between X and Y. Further improvements are also achieved on the classical ratio estimator by introducing a number of modified ratio estimators with the use of known parameters like, Co-efficient of Variation and Co-efficient of Kurtosis. The problem of constructing efficient estimators for the population variance has been widely discussed by various authors such as Das and Tripathi[7], Isaki[10], Singh et al.[17,19], Agarwal and Sithapit[1], Garcia and Cebrain[8], Arcos et al.[4], Ahmed et al.[2], Al-Jararha and Al-Haj Ebrahem[3], Bhushan[5], Prasad and Singh[14], Reddy[15], Singh and Chaudhary[16], Upadhyaya and Singh[23], Wolter[24], Kadilar and Cingi[11,12] and Gupta and Shabbir[9].

Motivated by Sisodia and Dwivedi[20], Singh et al.[18] and Upadhyaya and Singh[22], Kadilar and Cingi[11] suggested four ratio type variance estimators using known values of Co-efficient of variation C_X and Co-efficient of Kurtosis $\beta_{2(x)}$ of an auxiliary variable X together with their biases and mean squared errors as given in the Table 1:

$$\text{where } A_1 = \frac{S^2_x}{S^2_x + C_x}, A_2 = \frac{S^2_x}{S^2_x + \beta_{2(x)}}, A_3 = \frac{S^2_x\beta_{2(x)}}{S^2_x\beta_{2(x)} + C_x} \text{ and}$$

$$A_4 = \frac{S^2_x C_x}{S^2_x C_x + \beta_{2(x)}}$$

The modified ratio type variance estimators discussed above are biased but have minimum mean squared errors compared to the traditional ratio type variance estimator. The list of estimators given in Table 1 uses the known values of the parameters like S^2_x, C_x, β_2 and their linear combinations. Subramani and Kumarapandiyan[21] used the known value of the population median M_d of the auxiliary variable to improve the ratio estimators in estimation of population mean. Further we know that the value of median is unaffected and robustness by the extreme values or the presence of outliers in the population values unlike the other parameters like the variance, coefficient of variation and coefficient of kurtosis. The above discussed points have motivated us to introduce a modified ratio type variance estimator using the known value of the population median of the auxiliary variable. As a result, it is observed that the proposed estimator performs better than the traditional ratio type variance estimator as well as the existing modified ratio type variance estimators listed in Table 1. The materials of the present study are arranged as given below. The proposed estimator with known population median is presented in section 2 where as the conditions in which the proposed estimator performs better than the existing estimators are derived in section 3. The performances of the proposed and the existing estimators are assessed for certain natural populations in section 4 and the conclusion is presented in section 5.

2. Proposed Estimator

As we stated earlier one can always improve the performance of the estimator of the study variable by using the known population parameters of the auxiliary variable, which are positively correlated with that of study variable. In this section we have suggested a modified ratio type variance estimator using the population median of the auxiliary variable.

The proposed modified ratio type variance estimator for population variance S_y^2 is defined as

$$\hat{S}_{SK}^2 = s_y^2 \left[\frac{S_x^2+M_d}{s_x^2+M_d}\right] \quad (2)$$

where M_d is the population median of the auxiliary variable X.

The bias and mean squared error of $\hat{S}_{SK}^2$ =to the first degree of approximation are derived and given below:

$$B(\hat{S}_{SK}^2) = \gamma S_y^2 A_{SK}\left[A_{SK}(\beta_{2(x)}-1)-(\lambda_{22}-1)\right] \quad (3)$$

$$MSE(\hat{S}_{SK}^2) = \gamma S_y^4 \begin{bmatrix} (\beta_{2(y)}-1) \\ +A_{SK}^2(\beta_{2(x)}-1)-2A_{SK}(\lambda_{22}-1) \end{bmatrix}$$

$$\text{where } A_{SK} = \frac{S_x^2}{S_x^2+M_d} \quad (4)$$

3. Efficiency Comparison of Proposed Estimator

As we mentioned earlier the bias and mean squared error of the traditional ratio type variance estimator are given below:

$$B(\hat{S}_R^2) = \gamma S_y^2[(\beta_{2(x)}-1)-(\lambda_{22}-1)] \quad (5)$$

$$MSE(\hat{S}_R^2) = \gamma S_y^4[(\beta_{2(y)}-1)+(\beta_{2(x)}-1)-2(\lambda_{22}-1)] \quad (6)$$

For want of space; for the sake of convenience to the readers and for the ease of comparisons, the biases, the mean squared errors and the constants of the modified ratio type variance estimators given in Table 1 are represented in single class as given below:

$$B(\hat{S}_{KCi}^2) = \gamma S_y^2 A_i[A_i(\beta_{2(x)}-1)-(\lambda_{22}-1)]; \quad i = 1,2,3 \text{ and } 4 \quad (7)$$

$$MSE(\hat{S}_{KCi}^2) = \gamma S_y^4 \begin{bmatrix} (\beta_{2(y)}-1)+A_i^2(\beta_{2(x)}-1) \\ -2A_i(\lambda_{22}-1) \end{bmatrix} \quad (8)$$

$$i = 1,2,3 \text{ and } 4$$

$$\text{where } A_1 = \frac{S_x^2}{S_x^2+C_x}, A_2 = \frac{S_x^2}{S_x^2+\beta_{2(x)}}, A_3 = \frac{S_x^2\beta_{2(x)}}{S_x^2\beta_{2(x)}+C_x} \text{ and } A_4 = \frac{S_x^2 C_x}{S_x^2 C_x+\beta_{2(x)}}$$

The bias and mean squared error of the proposed modified ratio type variance estimator are given below:

$$B(\hat{S}_{SK}^2) = \gamma S_y^2 A_{SK}\left[A_{SK}(\beta_{2(x)}-1)-(\lambda_{22}-1)\right] \quad (9)$$

$$MSE(\hat{S}_{SK}^2) = \gamma S_y^4 \begin{bmatrix} (\beta_{2(y)}-1)+A_{SK}^2(\beta_{2(x)}-1) \\ -2A_{SK}(\lambda_{22}-1) \end{bmatrix} \quad (10)$$

where

$$A_{SK} = \frac{S_x^2}{S_x^2+M_d}$$

From the expressions given in (6) and (10) we have derived the condition for which the proposed estimator $\hat{S}_{SK}^2$ is more efficient than the traditional ratio type variance estimator and it is given below:

$$MSE(\hat{S}_{SK}^2) < MSE(\hat{S}_R^2) \text{ if } \lambda > 1 + \frac{(A_{SK}+1)(\beta_{2(x)}-1)}{2} \quad (11)$$

From the expressions given in (8) and (10) we have derived the conditions for which the proposed estimator $\hat{S}_{SK}^2$ is more efficient than the existing modified ratio type variance estimators given in Table 1, $\hat{S}_{KCi}^2$; $i = 1, 2, 3$ and 4 and are given below:

$$MSE(\hat{S}_{SK}^2) < MSE(\hat{S}_{KCi}^2) \text{ if } \lambda > 1 + \frac{(A_{SK}+A_i)(\beta_{2(x)}-1)}{2}; \quad (12)$$

$$i = 1, 2, 3 \text{ and } 4$$

4. Numerical Study

The performance of the proposed modified ratio type variance estimator is assessed with that of traditional ratio type estimator and existing modified ratio type variance estimators listed in Table 1 for certain natural populations. The populations 1 and 2 are the real data set taken from the Report on Waste 2004 drew up by the Italian bureau for the environment protection-APAT. Data and reports are available in the following website address http://www.osservatorionazionalerifiuti.it[25]. In the data set, for each of the Italian provinces, three variables are considered: the total amount (tons) of recyclable-waste collection in Italy in 2003 (Y) , the total amount of recyclable-waste collection in Italy in 2002 (X_1) and the number of inhabitants in 2003 (X_2). The population 3 is taken from Murthy[13] given in page 228 and population 4 is taken from Cochran[6] given in page 152. The population parameters and the constants computed from the above populations are given below:

Table 2. Parameters and Constants of the Populations

Parameters	Population 1	Population 2	Population 3	Population 4
N	103	103	80	49
n	40	40	20	20
$\bar{Y}$	626.2123	62.6212	51.8264	116.1633
$\bar{X}$	557.1909	556.5541	11.2646	98.6765
ρ	0.9936	0.7298	0.9413	0.6904
S_y	913.5498	91.3549	18.3569	98.8286
C_y	1.4588	1.4588	0.3542	0.8508
S_x	818.1117	610.1643	8.4563	102.9709
C_x	1.4683	1.0963	0.7507	1.0435
$\beta_{2(x)}$	37.3216	17.8738	2.8664	5.9878
$\beta_{2(y)}$	37.1279	37.1279	2.2667	4.9245
λ_{22}	37.2055	17.2220	2.2209	4.6977
M_d	308.0500	373.820	7.5750	64.0000
A_1	0.9999	0.9999	0.9896	0.9999
A_2	0.9999	0.9999	0.9615	0.9994
A_3	0.9999	0.9999	0.9964	1.0000
A_4	0.9999	0.9999	0.9493	0.9995
A_{SK}	0.9995	0.9989	0.9042	0.9940

The biases and mean squared errors of the existing and proposed modified ratio type variance estimator for the populations given above are given in the following Tables:

Table 3. Biases of the existing and proposed modified ratio type variance estimators

Estimator	Bias $\mathbf{B}(.)$			
	Population 1	Population 2	Population 3	Population 4
$\hat{S}^2_R$ Isaki[10]	2422.3488	135.9935	10.8762	630.0302
$\hat{S}^2_{KC1}$ Kadilar and Cingi[11]	2420.6810	135.9827	10.4399	629.7285
$\hat{S}^2_{KC2}$ Kadilar and Cingi[11]	2379.9609	135.8179	9.2918	628.3006
$\hat{S}^2_{KC3}$ Kadilar and Cingi[11]	2422.3041	135.9929	10.7222	629.9798
$\hat{S}^2_{KC4}$ Kadilar and Cingi[11]	2393.4791	135.8334	8.8117	628.3727
$\hat{S}^2_{SK}$ Proposed Estimator *	2072.7641	132.3292	7.1109	611.7234

Table 4. Mean squared error of the existing and proposed modified ratio type variance estimators

Estimator	Mean Squared Error $\mathbf{MSE}(.)$			
	Population 1	Population 2	Population 3	Population 4
$\hat{S}^2_R$ Isaki[10]	670393270	35796612	3925.1627	7235508
$\hat{S}^2_{KC1}$ Kadilar and Cingi[11]	670384403	35796605	3850.1552	7234298
$\hat{S}^2_{KC2}$ Kadilar and Cingi[11]	670169790	35796503	3658.4051	7228570
$\hat{S}^2_{KC3}$ Kadilar and Cingi[11]	670393032	35796611	3898.5560	7235306
$\hat{S}^2_{KC4}$ Kadilar and Cingi[11]	670240637	35796512	3580.8342	7228859
$\hat{S}^2_{SK}$ Proposed Estimator *	668667061	35794364	3320.2815	7162524

From the values of Table 3, it is observed that the bias of the proposed modified ratio type variance estimator is less than the biases of the traditional and existing modified ratio type variance estimators. Similarly from the values of Table 4, it is observed that the mean squared error of the proposed modified ratio type variance estimator is less than the mean squared errors of the traditional and existing modified ratio type variance estimators.

5. Conclusions

In this paper we have proposed a modified ratio type variance estimator using known value of Median of the auxiliary variable. The bias and mean squared error of the proposed modified ratio type variance estimator are obtained and compared with that of traditional ratio type variance estimator and existing modified ratio type variance estimators. Further we have derived the conditions for which the proposed estimator is more efficient than the traditional and existing estimators. We have also assessed the performances of the proposed estimator for some known populations. It is observed that the bias and mean squared error of the proposed estimator are less than the biases and mean squared errors of the traditional and existing estimators for certain known populations. Hence we strongly recommend that the proposed modified ratio type variance estimator may be preferred over the traditional ratio type variance estimator and existing modified ratio type variance estimators for the use of practical applications.

ACKNOWLEDGEMENTS

The authors record their gratitude to the editor and the reviewer for their constructive comments, which have improved the presentation of the paper. The second author wishes to record his gratitude and thanks to the Vice Chancellor, Pondicherry University for the financial assistance to carry out this research work through the University Fellowship.

REFERENCES

[1] Agarwal, M.C. and Sithapit, A.B. (1995). Unbiased ratio type estimation, Statistics and Probability Letters 25, 361-364

[2] Ahmed, M.S., Raman, M.S. and Hossain, M.I. (2000). Some competitive estimators of finite population variance Multivariate Auxiliary Information, Information and Management Sciences, Volume11 (1), 49-54

[3] Al-Jararha, J. and Al-Haj Ebrahem, M. (2012). A ratio estimator under general sampling design, Austrian Journal of Statistics, Volume 41(2), 105-115

[4] Arcos, A., Rueda, M., Martinez, M.D., Gonzalez, S., Roman, Y. (2005). Incorporating the auxiliary information available in variance estimation, Applied Mathematics and Computation 160, 387-399

[5] Bhushan, S. (2012). Some efficient sampling strategies based on ratio type, estimator, Electronic Journal of Applied Statistical Analysis, Volume 5(1), 74-88

[6] Cochran, W. G. (1977). Sampling techniques, Third Edition,

Wiley Eastern Limited

[7] Das, A.K. and Tripathi, T.P. (1978). Use of auxiliary information in estimating the finite population variance, Sankhya 40, 139-148

[8] Garcia, M.K. and Cebrain, A.A. (1997). Variance estimation using auxiliary information: An almost unbiased multivariate ratio estimator, Metrika 45, 171-178

[9] Gupta, S. and Shabbir, J. (2008). Variance estimation in simple random sampling using auxiliary information, Hacettepe Journal of Mathematics and Statistics, Volume 37, 57-67

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

[11] Kadilar, C. and Cingi, H. (2006). Improvement in variance estimation using auxiliary information, Hacettepe Journal of Mathematics and Statistics Volume 35 (1), 111-115

[12] Kadilar, C. and Cingi, H. (2006). Ratio estimators for population variance in simple and stratified sampling, Applied Mathematics and Computation 173, 1047-1058

[13] Murthy, M.N. (1967). Sampling theory and methods, Statistical Publishing Society, Calcutta, India

[14] 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 19, 1127-1139

[15] Reddy, V.N. (1974). On a transformed ratio method of estimation, Sankhya, Volume C36, 59-70

[16] Singh, D. and Chaudhary, F.S. (1986). Theory and analysis of sample survey designs, New Age International Publisher

[17] Singh, H.P., Chandra, P. and Singh, S. (2003). Variance estimation using multi-auxiliary information for random non-response in survey sampling, STATISTICA, anno LXIII, n. 1, 23-40

[18] Singh, H.P., Tailor, R., Tailor, R. and Kakran, M.S. (2004). An improved estimator of population mean using power transformation, Journal of the Indian Society of Agricultural Statistics 58(2), 223-230

[19] Singh, H.P., Upadhyaya, U.D. and Namjoshi, U.D. (1988). Estimation of finite population variance, Current Science 57, 1331-1334

[20] Sisodia, B.V.S. and Dwivedi, V.K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable, Journal of the Indian Society of Agricultural Statistics 33(1), 13-18

[21] Subramani, J. and Kumarapandiyan, G. (2012). Modified ratio estimator for population mean using median of the auxiliary variable, Proceedings of National Conference on Recent developments in the Applications of Reliability Theory and Survival Analysis held on 2nd and 3rd February 2012 at the Department of Statistics, Pondicherry University

[22] Upadhyaya, L.N. and Singh, H.P. (1999). Use of transformed auxiliary variable in estimating the finite population mean, Biometrical Journal 41 (5), 627-636

[23] Upadhyaya, L. N. and Singh, H. P. (2006). Almost unbiased ratio and product-type estimators of finite population variance in sample surveys, Statistics in Transition 7 (5), 1087–1096

[24] Wolter, K.M. (1985). Introduction to Variance Estimation, Springer-Verlag

[25] http://www.osservatorionazionalerifiuti.it/ElencoDocPub.asp?A_TipoDoc=6