Journal of Modern Applied Statistical Methods

Volume 12 | Issue 1 Article 11

5-1-2013

Estimation of Variance Using Known Coefficient of Variation and Median of an Auxiliary Variable

J. Subramani
Pondicherry University

G. Kumarapandiyan
Pondicherry University

Estimation of Variance Using Known Coefficient of Variation and Median of an Auxiliary Variable

Cover Page Footnote
The authors wish to express their gratitude and thanks for financial assistance received from UGC-MRP and the Department of Science and Technology (DST), New Delhi, India.

Journal of Modern Applied Statistical Methods
May 2013, Vol. 12, No. 1, 58-64

1538 – 9472/13/$95.00

Estimation of Variance Using Known Coefficient of Variation and Median of an Auxiliary Variable

J. Subramani G. Kumarapandiyan
Pondicherry University
R V Nagar, Kalapet, Puducherry

A modified ratio type variance estimator for estimating population variance of a study variable when the population median and coefficient of variation of an auxiliary variable are known is proposed. The bias and mean squared error of the proposed estimator are derived and conditions under which the proposed estimator performs better than the traditional ratio type variance estimators and modified ratio type variance estimators are obtained. Using a numerical study results show that the proposed estimator performs better than the traditional ratio type variance estimator and existing modified ratio type variance estimators.

Key words: Bias, mean squared error**,** natural populations, ratio type estimators, simple random sampling.

Introduction

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 mean, $\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ or its variance, $S_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(Y_i - \overline{Y})^2$ on the basis of a random sample selected from a population, U. When no additional information is available on an auxiliary variable, the simplest estimator of population mean is a simple random sample mean without replacement. However if an auxiliary variable X is closely related to a study variable Y then ratio and regression estimators can be used to improve the performance of the estimator of the study variable. This article considers the problem of estimating a population variance and uses auxiliary information to improve the efficiency of the population variance estimator

$$S_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(Y_i - \overline{Y})^2.$$

Population variance estimation is considered by Isaki (1983) who proposed ratio and regression estimators. Prasad & Singh (1990) considered a ratio type estimator for estimating population variance by improving Isaki's estimator (1983) with respect to bias and precision. Arcos, et al. (2005) introduced another ratio type estimator that also improved the Isaki's estimator (1983) and is less biased and more precise than other estimators. Notations used throughout this article are:

N: population size

n: sample size

γ: 1/n

Y: study variable

X: auxiliary variable

$\overline{X}, \overline{Y}$: population means

J. Subramani is an Associate Professor and Head in the Department of Statistics in the Ramanujan School of Mathematical Sciences. Email him at: drjsubramani@yahoo.co.in.
G. Kumarapandiyan is a Research Scholar in the Department of Statistics in the Ramanujan School of Mathematical Sciences. Email him at: kumarstat88@gmail.com.

$\bar{x}, \bar{y}$:	sample means
S_y^2, S_x^2 :	population variances
s_y^2, s_x^2 :	sample variances
C_x, C_y :	coefficients of variation
ρ :	correlation coefficient
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
$\hat{S}_{JG}^2$:	proposed modified ratio type variance estimator of S_y^2

Isaki (1983) suggested a ratio type variance estimator for population variance S_y^2 when the population variance S_x^2 of an auxiliary variable X is known together with its bias and mean squared error as:

$$\hat{S}_R^2 = s_y^2 \frac{S_x^2}{s_x^2} \quad (1)$$

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

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

where

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

and

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

The ratio type variance estimator in (1) is used to improve the precision of a population variance estimate compared to simple random sampling when a positive correlation exists between X and Y. Further improvements are achieved on the classical ratio type variance estimator by introducing a number of modified ratio type variance estimators using known parameters such as the coefficients of variation and kurtosis. The problem of constructing efficient estimators for a population variance has been discussed by many authors including: Agarwal & Sithapit (1995), Ahmed, et al. (2000), Al-Jararha & Al-Haj Ebrahem (2012), Arcos, et al.(2005), Bhushan (2012), Das & Tripathi (1978), Garcia &Cebrain (1997), Gupta & Shabbir (2008), Isaki (1983), Kadilar & Cingi (2006a, 2006b), Prasad & Singh (1990), Reddy (1974), Singh & Chaudhary (1986), Singh, et al. (1988, 2003), Subramani & Kumarapandiyan (2012b, 2012c), Upadhyaya & Singh (2006) and Wolter (1985)

Based on works by Sisodia & Dwivedi (1981), Upadhyaya & Singh (1999) and Singh, et al. (2004), Kadilar & Cingi (2006a) suggested four ratio type variance estimators using known values of coefficient of variation C_x and coefficient of kurtosis $\beta_{2(x)}$ of an auxiliary variable X together with their biases and mean squared errors (see Table 1).

The modified ratio type variance estimators discussed are biased but have minimal mean squared errors compared to traditional ratio type variance estimators. The list of estimators in Table 1 uses known values of the parameters such as S_x^2, C_x, β_2 and their linear combinations. Subramani & Kumarapandiyan (2012a) used the linear combination of a known value of the population median M_d and Coefficient of variation C_x of an auxiliary variable to improve the ratio estimators in estimating a population mean. Based on Subramani & Kumarapandiyan (2012a), a modified ratio type variance estimator using a linear combination of known value of a population median and coefficient of variation of an auxiliary variable is proposed. Results show that the proposed estimator performs better than the traditional ratio type and existing modified ratio type variance estimators.

Table 1: Existing Modified Ratio Type Variance Estimators with their Biases and Mean Squared Errors
(Source: Kadilar & Cingi, 2006a)

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

Notes: $A_1 = \frac{S^2_x}{S^2_x + C_x}$, $A_2 = \frac{S^2_x}{S^2_x + \beta_{2(x)}}$, $A_3 = \frac{S^2_x\beta_{2(x)}}{S^2_x\beta_{2(x)} + C_x}$, $A_4 = \frac{S^2_x C_x}{S^2_x C_x + \beta_{2(x)}}$

Proposed Estimator

The performance of the estimator of the study variable can be improved by using known population parameters of an auxiliary variable, which are positively correlated with a study variable. A modified ratio type variance estimator using the linear combination of known value of the population median M_d and coefficient of variation C_x of an auxiliary variable is proposed; this modified ratio type variance estimator for population variance S^2_y is defined as

$$\hat{S}^2_{JG} = s^2_y\left[\frac{C_x S^2_x + M_d}{C_x s^2_x + M_d}\right]. \tag{2}$$

The bias and mean squared error of $\hat{S}^2_{JG}$ to the first degree of approximation are:

$$B\left(\hat{S}^2_{IG}\right) = \gamma S^2_y A_{JG}\left[A_{JG}\left(\beta_{2(x)} - 1\right) - \left(\lambda_{22} - 1\right)\right] \tag{3}$$

$$MSE\left(\hat{S}^2_{JG}\right) = \gamma S^4_y\left[\begin{array}{l}\left(\beta_{2(y)} - 1\right) + A^2_{JG}\left(\beta_{2(x)} - 1\right) \\ -2A_{JG}\left(\lambda_{22} - 1\right)\end{array}\right] \tag{4}$$

where $A_{JG} = \frac{S^2_x}{S^2_x + M_d}$

Efficiency Comparison of the Proposed Estimator

The bias and mean squared error of the traditional ratio type variance estimator are:

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

$$MSE\left(\hat{S}_R^2\right)=\gamma S_y^4\begin{bmatrix}\left(\beta_{2(y)}-1\right)+\left(\beta_{2(x)}-1\right)\\-2\left(\lambda_{22}-1\right)\end{bmatrix} \quad (6)$$

For convenience and for the ease of comparisons, the biases mean squared errors and constants of the modified ratio type variance estimators shown in Table 1 are represented in a single class as:

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

$$MSE\left(\hat{S}_{KCi}^2\right)=\gamma S_y^4\begin{bmatrix}\left(\beta_{2(y)}-1\right)+A_i^2\left(\beta_{2(x)}-1\right)\\-2A_i\left(\lambda_{22}-1\right)\end{bmatrix},$$
$$i = 1, 2, 3, 4 \quad (8)$$

where

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

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

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

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

where $A_{JG}=\frac{S_x^2C_x}{S_x^2C_x+M_d}$

The condition for which the proposed estimator $\hat{S}_{JG}^2$ is more efficient than the traditional ratio type variance estimator derived from (6) and (10) is:

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

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

$$MSE\left(\hat{S}_{JG}^2\right)<MSE\left(\hat{S}_{KCi}^2\right)$$
$$\text{if } \lambda>1+\frac{\left(A_{JG}+A_i\right)\left(\beta_{2(x)}-1\right)}{2};$$
$$i = 1, 2, 3 \text{ and } 4 \quad (12)$$

Numerical Study

The performance of the proposed modified ratio type variance estimator was assessed and compared with a traditional ratio type and existing modified ratio type variance estimators (see Table 1) for natural populations. Populations 1 and 2 are a real data set taken from the Italian Bureau for Environment Protection (APAT) 2004 Report on Waste 2004. (Data and reports are available at: http://www.osservatorionazionalerifiuti.it.) For

Table 2: Parameters and Constants of the Populations

Parameters	Population1	Population2	Population3	Population4
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_{JG}	0.9996	0.9990	0.8763	0.9942

each Italian province, 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). Population 3 is from Murthy (1967, p. 228) and population 4 is from Cochran (1977, p. 152). The population parameters and constants computed from these populations are shown in Table 2. The biases and mean squared errors for the existing and proposed modified ratio type variance estimator for the populations are shown in Tables 3 and 4.

Table 3 shows that the proposed modified ratio type variance estimator has less bias compared to the biases of the traditional and existing modified ratio type variance estimators. It is observed that the mean squared error of the proposed modified ratio type variance estimator is lower than the mean squared errors of the traditional and existing modified ratio type variance estimators.

Conclusion

This article proposed a modified ratio type variance estimator using a known value of median of an auxiliary variable. The bias and mean squared error of the proposed estimator were obtained and compared with the traditional ratio type and existing modified ratio type variance estimators. Further the conditions for which the proposed estimator is more efficient than the traditional and existing estimators were derived. The performance of the proposed estimator was assessed using four known populations. Results show 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 the known populations considered. Based on results, the proposed modified ratio type variance estimator may be preferred over traditional ratio type and existing modified ratio type variance estimators for the use in practical applications.

Table 3: Biases of the Existing and Proposed Modified Ratio Type Variance Estimators

Source	Estimator	Bias B(.)			
		Population 1	Population 2	Population 3	Population 4
Isaki (1983)	$\hat{S}^2_R$	2422.3488	135.9935	10.8762	630.0302
Kadilar & Cingi (2006a)	$\hat{S}^2_{KC1}$	2420.6810	135.9827	10.4399	629.7285
	$\hat{S}^2_{KC2}$	2379.9609	135.8179	9.2918	628.3006
	$\hat{S}^2_{KC3}$	2422.3041	135.9929	10.7222	629.9798
	$\hat{S}^2_{KC4}$	2393.4791	135.8334	8.8117	628.3727
Proposed Estimator	$\hat{S}^2_{JG}$	2184.1910	132.6505	6.1235	612.4793

Table 4: Mean Squared Error of Existing and Proposed Modified Ratio Type Variance Estimators

Source	Estimator	Bias B(.)			
		Population 1	Population 2	Population 3	Population 4
Isaki (1983)	$\hat{S}^2_R$	670393270	35796612	3925.1627	7235508
Kadilar & Cingi (2006a)	$\hat{S}^2_{KC1}$	670384403	35796605	3850.1552	7234298
	$\hat{S}^2_{KC2}$	670169790	35796503	3658.4051	7228570
	$\hat{S}^2_{KC3}$	670393032	35796611	3898.5560	7235306
	$\hat{S}^2_{KC4}$	670240637	35796512	3580.8342	7228859
Proposed Estimator	$\hat{S}^2_{JG}$	669188376	35794559	3180.7740	7165517

Acknowledgements

The authors wish to express their gratitude and thanks for financial assistance received from UGC-MRP and the Department of Science and Technology (DST), New Delhi, India.

References

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

Ahmed, M. S., Raman, M. S., & Hossain, M. I. (2000).Some competitive estimators of finite population variance Multivariate Auxiliary Information. *Information and Management Sciences*, *11*(*1*), 49-54.

Al-Jararha, J., & Al-Haj Ebrahem, M. (2012).A ratio estimator under general sampling design. *Austrian Journal of Statistics*, *41*(*2*), 105-115.

Arcos, A., et al. (2005). Incorporating the auxiliary information available in variance estimation. *Applied Mathematics and Computation*, *160*, 387-399.

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

Cochran, W. G. (1977). *Sampling techniques*, *3rd Edition*. New York, NY: John Wiley & Sons.

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

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

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

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

Kadilar, C., & Cingi, H. (2006a). Improvement in variance estimation using auxiliary information. *Hacettepe Journal of Mathematics and Statistics*, *35*(*1*), 111-115.

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

Murthy, M. N. (1967*). Sampling theory and methods*. Calcutta, India: Statistical Publishing Society.

Prasad, B., & 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.

Reddy, V. N. (1974).On a transformed ratio method of estimation*. Sankhya*, *C*, *36*, 59-70.

Singh, D., & Chaudhary, F. S. (1986). *Theory and analysis of sample survey designs*. New Delhi: New Age International Limited.

Singh, H. P., Chandra, P., & Singh, S. (2003). Variance estimation using multi-auxiliary information for random non-response in survey sampling. *STATISTICA*, *LXIII*, *1*, 23-40.

Singh, H. P., Tailor, R., Tailor, R., & 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.

Singh, H. P., Upadhyaya, U. D., &Namjoshi, U. D. (1988).Estimation of finite population variance. *Current Science*, *57*, 1331-1334.

Sisodia, B. V. S., & 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.

Subramani, J., & Kumarapandiyan, G. (2012a). Estimation of population mean using coefficient of variation and median of an auxiliary variable. *International Journal of Probability and Statistics*, *1*(*4*), 111-118.

Subramani, J., & Kumarapandiyan, G. (2012b). Variance estimation using median of the auxiliary variable.*International Journal of Probability and Statistics*, *1*(*3*), 36-40.

Subramani, J., & Kumarapandiyan, G. (2012c). Variance estimation using quartiles and their functions of the auxiliary variable. *International Journal of Statistics and Applications*, *2*(*5*), 67-72.

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

Upadhyaya, L. N., & 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.

Wolter, K. M. (1985). *Introduction to variance estimation*. New York, NY: Springer-Verlag.