Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.12785/jsap/020314

A RatioType Estimator for the Estimation of Population Variance using Quartiles of an Auxiliary Variable

Mursala Khan* and Javid Shabbir
Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan
**E-mail: mursala.khan@yahoo.com*

Received: 17 Jan. 2013, Revised: 20 Aug. 2013, Accepted: 17 Sep. 2013
Published online: 1 Nov. 2013

Abstract: In this paper we have suggested aratio type estimator for the estimation of population variance of the study variable using quartiles and its functions as an auxiliary variable. The expression of bias and mean squared error of the proposed estimator is derived up to first order approximation.The proposed estimator is compared with the other existing estimators and its efficiency condition is carried out. An empirical study is carried out with the help of four natural populations to judge the merits of the suggested estimator over other existing estimators practically.

Keywords: Auxiliary Variable, Ratio Estimator, Simple Random Sampling, Bias, Mean Squared Error, Efficiency.

1 Introduction

Consider a finite population $U=\{U_1, U_2, U_3....,U_N\}$ of size N units. Let y and x be the real valued functions defined on a finite population U.In order to estimate the unknown population variance we take a sample of size n units from the population U by using simple random sample with replacement. In this paper we use the knowledge of quartiles and functions of quartiles as an variable to estimate the population variance of the study variable.

2 Notations

The different symbols and notations used in this paper are given by.

$\bar{Y}, \bar{X}$ Population means of the study and auxiliary variable respectively

$\bar{y}, \bar{x}$ Sample means of the study and auxiliary variable respectively

S_y^2, S_x^2 Population variances of the study and auxiliary variable respectively

C_y, C_x Population coefficient of variation of the study and auxiliary variable respectively

ρ Coefficient correlation

Q_1 First quartile of the auxiliary variable

Q_3 Third quartile of the auxiliary variable

$Q_r = Q_3 - Q_1$ Inter quartile range

$Q_a = \left(\frac{Q_3 + Q_1}{2}\right)$ Quartiles average

$\beta_1 = \frac{N\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^3}{(N-1)(N-2)S^2}$ Coefficient of skewness of the auxiliary variable

$\beta_2 = \frac{N(N+1)\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^4}{(N-1)(N-2)(N-3)S^4} - \frac{3(N-1)^2}{(N-2)(N-3)}$ Coefficient of kurtosis of the auxiliary variable.

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

$\hat{S}_R^2$ The usual ratio estimator of S_y^2

$\hat{S}_k^2$ Existing modified ratio type variance estimators of S_y^2

$\hat{S}_p^2$ Proposed ratio type variance estimator of S_y^2

The usual ratio estimator for estimating the population variance of the study variable suggested by Isaki [4] is given by

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

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

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

where $\beta_{2(x)} = \frac{\mu_{04}}{\mu_{02}^2}$, $\beta_{2(y)} = \frac{\mu_{40}}{\mu_{20}^2}$, $\lambda_{22} = \frac{\mu_{22}}{\mu_{02}\mu_{20}}$ and $\mu_{rs} = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^r\left(X_i - \bar{X}\right)^s$

A large number of estimators have been suggested bystatisticians for improving the efficiency of the ratio estimators by utilizing some known parameters or function of parameters of the auxiliary variable, such as median, correlation coefficient, quartiles, coefficient of variation, coefficient of kurtosis, Coefficient of skewness, etc. The problem for estimating the unknown population variance of the study variable using knowledge of the auxiliary variable has been discussed by various authors such as Das and Triphati [2], Srivvastava and Jhajj [10], Upadhyaya and Singh[12], etc. Agarwal [1], Kadilar and Cingi [6]suggested various ratio type estimators in simple random and stratified random sampling to improve the efficiency of the estimators by utilizing the known information of the auxiliary variable.Singh et al. [9] improved an estimator of the population mean using power transformation, Gupta and Shabbir[3] proposed a class of hybrid variance estimators in which the efficiency of the ratio estimators is increases as the traditional ratio and some other existing ratio estimators.

Recently Subramani and Kumarapandiyan[11] proposed different class of modified ratio type estimators for the estimation of finite population variance of the study variable.

A list of different modified ratio type estimators are presented in Table 1.

Table 1: Existing ratio type estimators with their biases, mean square errors and their constants

Estimator	$Bias(\hat{S}_k^2)$	$MSE(\hat{S}_k^2)$	Constant R_k
$\hat{S}_1^2 = s_y^2\left[\frac{S_x^2+C_x}{s_x^2+C_x}\right]$ Kadilar and Cingi[5]	$\lambda S_y^2 R_1\left[R_1(\beta_{2(x)}-1) - (\lambda_{22}-1)\right]$	$\lambda S_y^4\left[(\beta_{2(y)}-1)+R_1^2(\beta_{2(x)}-1) - 2R_1(\lambda_{22}-1)\right]$	$R_1 = \frac{S_x^2}{S_x^2+C_x}$
$\hat{S}_2^2 = s_y^2\left[\frac{S_x^2+\beta_{2(x)}}{s_x^2+\beta_{2(x)}}\right]$ Kadilar and Cingi[5]	$\lambda S_y^2 R_2\left[R_2(\beta_{2(x)}-1) - (\lambda_{22}-1)\right]$	$\lambda S_y^4\left[(\beta_{2(y)}-1)+R_2^2(\beta_{2(x)}-1) - 2R_2(\lambda_{22}-1)\right]$	$R_2 = \frac{S_x^2}{S_x^2+\beta_{2(x)}}$
$\hat{S}_3^2 = s_y^2\left[\frac{S_x^2\beta_{2(x)}+C_x}{s_x^2\beta_{2(x)}+C_x}\right]$ Kadilar and Cingi[5]	$\lambda S_y^2 R_3\left[R_3(\beta_{2(x)}-1) - (\lambda_{22}-1)\right]$	$\lambda S_y^4\left[(\beta_{2(y)}-1)+R_3^2(\beta_{2(x)}-1) - 2R_3(\lambda_{22}-1)\right]$	$R_3 = \frac{S_x^2\beta_{2(x)}}{S_x^2\beta_{2(x)}+C_x}$
$\hat{S}_4^2 = s_y^2\left[\frac{S_x^2C_x+\beta_{2(x)}}{s_x^2C_x+\beta_{2(x)}}\right]$ Kadilar and Cingi[5]	$\lambda S_y^2 R_4\left[R_4(\beta_{2(x)}-1) - (\lambda_{22}-1)\right]$	$\lambda S_y^4\left[(\beta_{2(y)}-1)+R_4^2(\beta_{2(x)}-1) - 2R_4(\lambda_{22}-1)\right]$	$R_4 = \frac{S_x^2C_x}{S_x^2C_x+\beta_{2(x)}}$
$\hat{S}_5^2 = s_y^2\left[\frac{S_x^2+Q_1}{s_x^2+Q_1}\right]$ Subramani and Kumarapandiyan[11]	$\lambda S_y^2 R_5\left[R_5(\beta_{2(x)}-1) - (\lambda_{22}-1)\right]$	$\lambda S_y^4\left[(\beta_{2(y)}-1)+R_5^2(\beta_{2(x)}-1) - 2R_5(\lambda_{22}-1)\right]$	$R_5 = \frac{S_x^2}{S_x^2+Q_1}$
$\hat{S}_6^2 = s_y^2\left[\frac{S_x^2+Q_3}{s_x^2+Q_3}\right]$ Subramani and Kumarapandiyan[11]	$\lambda S_y^2 R_6\left[R_6(\beta_{2(x)}-1) - (\lambda_{22}-1)\right]$	$\lambda S_y^4\left[(\beta_{2(y)}-1)+R_6^2(\beta_{2(x)}-1) - 2R_6(\lambda_{22}-1)\right]$	$R_6 = \frac{S_x^2}{S_x^2+Q_3}$
$\hat{S}_7^2 = s_y^2\left[\frac{S_x^2+Q_r}{s_x^2+Q_r}\right]$ Subramani and Kumarapandiyan[11]	$\lambda S_y^2 R_7\left[R_7(\beta_{2(x)}-1) - (\lambda_{22}-1)\right]$	$\lambda S_y^4\left[(\beta_{2(y)}-1)+R_7^2(\beta_{2(x)}-1) - 2R_7(\lambda_{22}-1)\right]$	$R_7 = \frac{S_x^2}{S_x^2+Q_r}$
$\hat{S}_8^2 = s_y^2\left[\frac{S_x^2+Q_d}{s_x^2+Q_d}\right]$ Subramani and Kumarapandiyan[11]	$\lambda S_y^2 R_8\left[R_8(\beta_{2(x)}-1) - (\lambda_{22}-1)\right]$	$\lambda S_y^4\left[(\beta_{2(y)}-1)+R_8^2(\beta_{2(x)}-1) - 2R_8(\lambda_{22}-1)\right]$	$R_8 = \frac{S_x^2}{S_x^2+Q_d}$
$\hat{S}_9^2 = s_y^2\left[\frac{S_x^2+Q_a}{s_x^2+Q_a}\right]$ Subramani and Kumarapandiyan[11]	$\lambda S_y^2 R_9\left[R_9(\beta_{2(x)}-1) - (\lambda_{22}-1)\right]$	$\lambda S_y^4\left[(\beta_{2(y)}-1)+R_9^2(\beta_{2(x)}-1) - 2R_9(\lambda_{22}-1)\right]$	$R_9 = \frac{S_x^2}{S_x^2+Q_a}$

3 Proposed Ratio TypeEstimator

In this section, we have proposed a ratio type estimator for the estimation of population variance. The expression for bias and *MSE* is derived up to first order of approximation. The bias, *MSE* and the constant of the proposed estimator is given by.

Table 2: The Proposed Estimator

Estimator	$Bias(\hat{S}_p^2)$	$MSE(\hat{S}_p^2)$	Constant R_p
$\hat{S}_p^2 = s_y^2\left[\dfrac{S_x^2\rho+Q_3}{s_x^2\rho+Q_3}\right]$	$\lambda S_y^2 R_p\left[\begin{array}{c}R_p\left(\beta_{2(x)}-1\right)\\-\left(\lambda_{22}-1\right)\end{array}\right]$	$\lambda S_y^4\left[\begin{array}{r}\left(\beta_{2(y)}-1\right)+R_p^2\left(\beta_{2(x)}-1\right)\\-2R_p\left(\lambda_{22}-1\right)\end{array}\right]$	$R_p=\dfrac{S_x^2\rho}{S_x^2\rho+Q_3}$

4 Efficiency Comparison

In this section, we compare the proposed estimator with the other existing estimators and some efficiency comparison condition is carry out under which the proposed ratio type estimator is better than $\hat{S}_k^2\ (k=1,2,3,\ldots\ldots,9)$.

The biases and *MSE's* of these existing ratio type estimators up to first order of approximation are given by.

$$Bias\left(\hat{S}_k^2\right)=\lambda S_y^2 R_k\left[R_k\left(\beta_{2(x)}-1\right)-\left(\lambda_{22}-1\right)\right]$$
$$MSE\left(\hat{S}_k^2\right)=\lambda S_y^4\left[\left(\beta_{2(y)}-1\right)+R_k^2\left(\beta_{2(x)}-1\right)-2R_k\left(\lambda_{22}-1\right)\right]$$
$$k=1,2,3,\ldots\ldots,9 \tag{2}$$

where $R_1=\dfrac{S_x^2}{S_x^2+C_x}$, $R_2=\dfrac{S_x^2}{S_x^2+\beta_{2(x)}}$, $R_3=\dfrac{S_x^2\beta_{2(x)}}{S_x^2\beta_{2(x)}+C_x}$, $R_4=\dfrac{S_x^2C_x}{S_x^2C_x+\beta_{2(x)}}$, $R_5=\dfrac{S_x^2}{S_x^2+Q_1}$, $R_6=\dfrac{S_x^2}{S_x^2+Q_3}$, $R_7=\dfrac{S_x^2}{S_x^2+Q_r}$, $R_8=\dfrac{S_x^2}{S_x^2+Q_d}$ and $R_9=\dfrac{S_x^2}{S_x^2+Q_a}$.

Now the bias, *MSE's* and constant of the proposed estimator is given by.

$$Bias\left(\hat{S}_p^2\right)=\lambda S_y^2 R_p\left[R_p\left(\beta_{2(x)}-1\right)-\left(\lambda_{22}-1\right)\right]$$
$$MSE\left(\hat{S}_p^2\right)=\lambda S_y^4\left[\left(\beta_{2(y)}-1\right)+R_p^2\left(\beta_{2(x)}-1\right)-2R_p\left(\lambda_{22}-1\right)\right]$$
$$R_p=\frac{S_x^2\rho}{S_x^2\rho+Q_3} \tag{3}$$

From equation (2) and (3) we have, derived the efficiency comparison condition under which the proposed estimatorperform better than the other existing estimators as given by.

$$MSE\left(\hat{S}_p^2\right)\le MSE\left(\hat{S}_k^2\right) \text{ if } \lambda_{22}>1+\frac{\left(R_p+R_k\right)\left(\beta_{2(x)}-1\right)}{2} \tag{4}$$
$$k=1,2,3,\ldots\ldots,9.$$

5 Numerical Study

To illustrate the performance of the proposed ratio type and the modified ratio type estimators of population variance S^2 we have considered four natural populations.

Population-1: *Italian bureau for the environment protection-APATWaste 2004* [13]
Y=Total amount (tons)of recyclable-waste collection in Italy in 2003 and
X=amount (tons)of recyclable-waste collection in Italy in 2002

$N=103$, $n=40$, $\bar{Y}=626.2123$, $\bar{X}=557.1909$, $\rho=0.9936$, $S_y=913.5498$, $C_y=1.4588$, $S_x=818.1117$, $C_x=1.4683$, $B_{2(x)}=37.3216$, $B_{2(y)}=37.1279$, $\lambda_{22}=37.2055$, $\lambda_{22}=37.2055$, $Q_1=142.9950$, $Q_3=665.6250$, $Q_r=522.6300$, $Q_d=261.3150$, $Q_a=404.3100$.

Population-2: *Italian bureau for the environment protection-APAT Waste 2004*[13]
Y=Total amount (tons)of recyclable-waste collection in Italy in 2003 and
X=Number of inhabitants in 2003

$N=103$, $n=40$, $\bar{Y}=62.6212$, $\bar{X}=556.5541$, $\rho=0.7298$, $S_y=91.3549$, $C_y=1.4588$, $S_x=610.1643$, $C_x=1.0963$, $B_{2(x)}=17.8738$, $B_{2(y)}=37.1279$, $\lambda_{22}=17.2220$, $\lambda_{22}=17.2220$, $Q_1=259.3830$, $Q_3=628.0235$, $Q_r=368.6405$, $Q_d=184.3293$, $Q_a=443.7033$.

Population-3: *Murthy* [7]
X=Fixed Capital and Y=Output for 80 factories in a region

$N=80$, $n=20$, $\bar{Y}=51.8264$, $\bar{X}=11.2646$, $\rho=0.9413$, $S_y=18.3569$, $C_y=0.3542$, $S_x=8.4563$, $C_x=0.7507$, $B_{2(x)}=2.8664$, $B_{2(y)}=2.2667$, $\lambda_{22}=2.2209$, $Q_1=5.1500$, $Q_3=16.975$, $Q_r=11.825$, $Q_d=5.9125$, $Q_a=11.0625$.

Population-4: *Singh and Chaudhary*[8]

$N=70$, $n=25$, $\bar{Y}=96.7000$, $\bar{X}=175.2671$, $\rho=.7293$, $S_y=60.7140$, $C_y=0.6254$, $S_x=140.8572$, $C_x=0.8037$, $B_{2(x)}=7.0952$, $B_{2(y)}=4.7596$, $\lambda_{22}=4.6038$, $Q_1=80.1500$, $Q_3=225.0250$, $Q_r=144.8750$, $Q_d=72.4375$, $Q_a=152.5875$.

Table 3: The constants of the competing and proposed ratio estimators

Estimator	Constants R_k			
	Population 1	Population 2	Population 3	Population 4
$\hat{S}_1^2$ Kadilar and Cingi[5]	0.9999	0.9999	0.9896	0.9999
$\hat{S}_2^2$ Kadilar and Cingi[5]	0.9999	0.9999	0.9615	0.9996
$\hat{S}_3^2$ Kadilar and Cingi[5]	0.9999	0.9999	0.9964	0.9999
$\hat{S}_4^2$ Kadilar and Cingi[5]	0.9999	0.9999	0.9493	0.9996
Estimator	Constants R_k			

	Population 1	Population 2	Population 3	Population 4
$\hat{S}_5^2$ Subramani and Kumarapandiyan[11]	0.9997	0.9994	0.9328	0.9960
$\hat{S}_6^2$ Subramani and Kumarapandiyan[11]	0.99901	0.9983	0.8082	0.9888
$\hat{S}_7^2$ Subramani and Kumarapandiyan[11]	0.9992	0.9990	0.8581	0.9928
$\hat{S}_8^2$ Subramani and Kumarapandiyan[11]	0.9996	0.9995	0.9236	0.9964
$\hat{S}_9^2$ Subramani and Kumarapandiyan[11]	0.9996	0.9988	0.8660	0.9924
$\hat{S}_p^2$ (**Proposed estimator**)	0.9990	0.9977	0.7986	0.9847

Table 4: The biases of the competing and the proposed ratio type estimator

Estimator		*Bias* (.)			
		Population 1	Population 2	Population 3	Population 4
Existing	$\hat{S}_1^2$	2420.6810	135.9827	10.4399	364.3702
	$\hat{S}_2^2$	2379.9609	135.8179	9.2918	363.9722
	$\hat{S}_3^2$	2422.3041	135.9929	10.7222	364.4139
	$\hat{S}_4^2$	2393.4791	135.8334	8.8117	363.8627
	$\hat{S}_5^2$	2259.9938	133.4494	8.1749	359.3822
	$\hat{S}_6^2$	1667.7818	129.8456	3.9142	350.4482
	$\hat{S}_7^2$	1829.6315	132.3799	5.5038	355.3634
	$\hat{S}_8^2$	2125.7591	134.1848	7.8275	359.8641
	$\hat{S}_9^2$	1963.6570	131.6458	5.7705	354.8875
Proposed	$\hat{S}_p^2$	1663.3086	127.6040	3.6276	348.1975

Table 5: The *MSE's* of the competing and the proposedratio type estimator

Estimator		*MSE* (.)			
		Population 1	Population 2	Population 3	Population 4
Existing	$\hat{S}_1^2$	67038384403	35796605	3850.1552	1415839
	$\hat{S}_2^2$	670169790	35796503	3658.4051	1414994
	$\hat{S}_3^2$	670393032	35796611	3898.5560	1415931
	$\hat{S}_4^2$	670240637	35796512	3580.8342	1414762

	$\hat{S}_5^2$	669558483	35795045	3480.5516	1427990
	$\hat{S}_6^2$	667000531	35792955	2908.6518	1408858
	$\hat{S}_7^2$	667623576	35794395	3098.4067	1419946
	$\hat{S}_8^2$	668911625	35795495	3427.1850	1429077
	$\hat{S}_9^2$	668182833	35793951	3133.3256	1418424
Proposed	$\hat{S}_p^2$	666910707	35791562	2878.5603	1398150

Table 4 and Table 5 clearly indicates that the bias and the mean squared error of the proposedratio type estimator is less than the biases and mean squared errors of the existing $\hat{S}_k^2\ (k=1, 2, 3, \ldots\ldots, 9)$ ratio estimators. Thus the use of proposed ratio type estimator performed better than other competing estimators discussed in the literature.

6 Conclusions

In this paper,we have proposed a ratio type estimator for estimating the population variance, which is found to be more efficient than the usual estimator and the other existing ratio type estimators. The proposed estimator has been compared with other competing estimators discussed in the literature. We have also derived the condition under which the proposed estimator performs better than the other existing modified ratio estimators. Empirical investigations are carried out by using four natural populations which shows that the efficiency of the proposed estimatoris more than the usual ratio and the other existing ratio type estimators.

References

[1] Agarwal, S. K., Two auxiliary variates in ratio method of estimation. *Biometrical Journal*, **22**, 569-573 (1980).
[2] Das, A. K. and Tripathi, T. P., Use of auxiliary information in estimating the finite po-pulation variance, Sankhya, C, **40**, 139-148 (1978).
[3] Gupta, S and Shabbir, J., Variance estimation in simple random sampling using auxiliary information, *Hacettepe Journal of mathematics and Statistics,* **37**, 57-67 (2008).
[4] Isaki, C, T., Variance estimation using auxiliary information, *Journal of American Statistical Association,* **78**, 117-123 (1983).
[5] Kadilar, C. and Cingi, H., Improvement in variance estimation using auxiliary information, *Hacettepe Journal of mathematics and Statistics,* **35**, 111-115 (2006).
[6] Kadilar, C. and Cingi, H., Ratio estimators for population variance in simple and Stratified sampling, *Applied Mathematics and Computation,* **173**, 1047-1058 (2006).
[7] Murthy, M. N., Sampling Theory and Methods, Statistical Publishing Society Calcutta, India, (1967).
[8] Singh, D. and Chaudhary, F. S., Theory and analysis of sample survey designs, New-Age International Publisher, (1986).
[9] Singh, H. P., Tailor, R. and Kakran, M. S., An improved estimator of population mean using power transformation, *Journal of the Indian Society of Agricultural Statistics,* **58**, 223-230 (2004).
[10] Srivastava, S. K. and Jhajj, H. S., A class of estimators using auxiliary information for estimating finite population variance, Sankhya, *C*, **42**, 87-96 (1980).
[11] Subramani, J. and Kumarapandiyan, G., Variance estimation using quartiles and their functions of an auxiliary variable, *International Journal of Statistics and Applications*, **2**, 67-72 (2012).
[12] Upadhyaya, L. N. and Singh, H. P., Use of auxiliary information in the estimation of population variance, mathematical forum, **4**, 33-36 (1983).
[13] http://www.osservatorionalerifiuti.it/ElencoDocPub.asp?A_TipoDoc=6