Modified Ratio Estimator in Simple Random Sampling using Auxiliary Information

Sumaira Ajmal Khan Hafsa Abbas Mehran Faiz Lubna Shaheen

ABSTRACT
In this study, we offer a new estimator for population variance in simple random sampling based on auxiliary information. We calculated the proposed estimator's bias and MSE equations and compared them to the bias and MSE of existing estimators, demonstrating that the new estimator is more efficient than the existing estimators proposed by different authors. With the aid of numerical example, we can support this theoretical result.

Keywords
Variance estimator, bias, MSE, simple random sampling, auxiliary information, efficiency

1. INTRODUCTION

Auxiliary information has been used extensively in estimation of parameters like mean and variance since several decades. Isaki (1983) [1] got inspiration from ratio estimator of finite population mean and proposed a ratio estimator, usually called classical estimator, of finite population variance. But Singh and Solanki (2013a) claimed that Isaki (1983) [1] ratio estimator is the member of the class of estimators developed by Das and Tripathi (1978) [3]. Arcos and Rueda (1997) [4] suggested multivariate ratio estimator for population variance. Ahmed et al. (2000) [5] criticized the claim of Arcos and Rueda (1997) [4], Kadilar and Cingi (2006a) [6] developed an estimator, ratio-type estimator of the mean of population, Kadilar and Cingi (2006b) [7] extended the idea of Isaki (1983) [1] ratio estimator for population variance. Kadilar and Cingi (2006b) [7] involved the information available about coefficient of variation and coefficient of kurtosis of the auxiliary variable to generate these estimators under simple random sampling as well. Gupta and Shabbir (2008) [8] gave a hybrid class of variance estimators of population mean. Subramani and Kumarapandiyan (2012a) [9] modified the usual ratio-type estimator of Kadilar and Cingi (2006b) [7] for population variance using population median obtained from auxiliary variable. Subramani and Kumarapandiyan (2012b) [10] further modified the usual ratio-type variance estimators using lower and upper quartiles, inter-quartile range, quartile deviation and quartile average of the auxiliary variable. Subramani and Kumarapandiyan (2013) [11] developed another more efficient modified ratio-type estimator using median and coefficient of variation of the auxiliary variable.

2. NOTATION

Before delving deeper into the classical ratio type variance estimator, modified ratio type variance estimators, and proposed modified ratio type variance estimator, the notations that will be used in this paper are listed below:

1. Population size
$$N$$
2. Sample size
$$n$$
3. Population mean of the study variable y
$$\bar{Y} = \frac{\sum_{i=1}^{N} y_i}{N}$$
4. Population mean of the auxiliary variable x
$$\bar{X} = \frac{\sum_{i=1}^{N} x_i}{N}$$
5. Sample mean of the study variable y
$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$
6. Sample mean of the auxiliary variable x
$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$
7. Population variance of the study variable y
$$S_Y^2 = \frac{\sum_{i=1}^{N} (y_i - \bar{Y})^2}{N-1}$$
8. Population variance of the auxiliary variable x
$$S_X^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{X})^2}{N-1}$$
9. Sample variance of the study variable y
$$S_y^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1}$$
10. Sample variance of the auxiliary variable x
$$S_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$
11. Coefficient of variation
$$C_x = \frac{S_x}{\bar{X}}, \quad C_y = \frac{S_y}{\bar{Y}}$$
12. Skewness of the auxiliary variable
$$\beta_{1(x)} = \frac{\mu_{03}^2}{\mu_{02}^2}$$
13. Kurtosis of the auxiliary variable
$$\beta_{2(x)} = \frac{\mu_{04}^2}{\mu_{02}^2}$$
14. Kurtosis of the study variable
$$\beta_{2(y)} = \frac{\mu_{40}^2}{\mu_{20}^2}$$
15. $\gamma = \frac{(1-f)}{n}$

3. SOMEEXISTINGESTIMATORS

1. Isaki (1983)

Motivated by the estimator of the population mean

$$\hat{\mu}_{yR} = \frac{\bar{y}}{\bar{x}}\bar{X}$$

Isaki (1983) [1] suggested the ratio estimators of the variance population using auxiliary information.

$$t_{Isaki} = s_y^2 \frac{S_x^2}{s_x^2}$$

The MSE of the above estimator using first order approximation is

$$MSE(t_{Isaki}) \cong \lambda' S_y^4\{\beta_2(y) + \beta_2(x) - 2h\}$$

Or

$$MSE(t_{Isaki}) \cong \lambda' S_y^4\{\beta_2'(y) + \beta_2'(x) - 2h'\}$$

And the bias is

$$Bias(t_{Isaki}) \cong \lambda' S_y^2\{\beta_2(x) - h\}$$

Or

$$Bias(t_{Isaki}) \cong \lambda' S_y^2\{\beta_2'(x) - h'\}$$

Or

$$Bias(t_{Isaki}) \cong -Cov\left(\frac{s_y^2}{s_x^2}, S_x^2\right)$$

2. Kadilar and Cingi (2005)

Kadilar and Cingi(2005) [2] proposed the following class of estimators for population variance:

$$t_{CH1} = s_y^2[s_x^2 - \beta_2(x)]^{-1}[S_x^2 - \beta_2(x)]$$

$$t_{CH2} = s_y^2[s_x^2 C_x - \beta_2(x)]^{-1}[S_x^2 C_x - \beta_2(x)]$$

$$t_{CH3} = s_y^2(s_x^2 - C_x)^{-1}(S_x^2 - C_x)$$

$$t_{CH4} = s_y^2[s_x^2\beta_2(x) - C_x]^{-1}[S_x^2\beta_2(x) - C_x]$$

The MSE's of the above estimators are given below respectively:

$$MSE(t_{CH1}) \cong \lambda' S_y^4\left[\beta_2'(y) - 2\frac{S_x^2}{S_x^2 - \beta_2(x)}h' + \left\{\frac{S_x^2}{S_x^2 - \beta_2(x)}\right\}^2 \beta_2'(x)\right]$$

$$MSE(t_{CH2}) \cong \lambda' S_y^4\left[\beta_2'(y) - 2\frac{S_x^2}{S_x^2 C_x - \beta_2(x)}h' + \left\{\frac{S_x^2}{S_x^2 C_x - \beta_2(x)}\right\}^2 \beta_2'(x)\right]$$

$$MSE(t_{CH3}) \cong \lambda' S_y^4\left[\beta_2'(y) - 2\frac{S_x^2}{S_x^2 - C_x}h' + \left\{\frac{S_x^2}{S_x^2 - C_x}\right\}^2 \beta_2'(x)\right]$$

$$MSE(t_{CH4}) \cong \lambda' S_y^4\left[\beta_2'(y) - 2\frac{S_x^2}{S_x^2\beta_2(x) - C_x}h' + \left\{\frac{S_x^2}{S_x^2\beta_2(x) - C_x}\right\}^2 \beta_2'(x)\right]$$

The Biases of above estimators are given below respectively:

$$Bias(t_{CH1}) \cong \lambda' S_y^2 \frac{S_x^2}{S_x^2 - \beta_2(x)}\left[\frac{S_x^2}{S_x^2 - \beta_2(x)}\beta_2'(x) - h'\right]$$

$$Bias(t_{CH2}) \cong \lambda' S_y^2 \frac{S_x^2}{S_x^2 C_x - \beta_2(x)}\left[\frac{S_x^2}{S_x^2 C_x - \beta_2(x)}\beta_2'(x) - h'\right]$$

$$Bias(t_{CH3}) \cong \lambda' S_y^2 \frac{S_x^2}{S_x^2 - C_x}\left[\frac{S_x^2}{S_x^2 - C_x}\beta_2'(x) - h'\right]$$

$$Bias(t_{CH4}) \cong \lambda' S_y^2 \frac{S_x^2}{S_x^2\beta_2(x) - C_x}\left[\frac{S_x^2}{S_x^2\beta_2(x) - C_x}\beta_2'(x) - h'\right]$$

4. PROPOSEDESTIMATOR

Using the known value of the auxiliary variable's population mean, we introduced a new modified ratio type variance estimator.For population variance, the modified ratio type variance estimator is defined as follows:

$$t_{new} = s_y^2 \frac{\rho^2 S_x^2 + \bar{X}^2}{\rho^2 s_x^2 + \bar{X}^2}$$

Bias of the proposed estimator

$$t_{new} = S_y^2(1 + e_0)\frac{\rho^2 S_x^2 + \bar{X}^2}{\rho^2 S_x^2(1 + e_1) + \bar{X}^2}$$

$$t_{new} = S_y^2(1 + e_0)\{1 + \Omega e_1\}^{-1}$$

Where

$$\Omega = \frac{\rho^2 S_x^2}{\rho^2 S_x^2 + \bar{X}^2}$$

Expanding the expression by using Taylor's series

$$t_{new} = S_y^2(1 + e_0)\{1 - \Omega e_1 + (\Omega e_1)^2 - (\Omega e_1)^3 + \cdots\}$$
$$t_{new} = S_y^2(1 + e_0)\{1 - \Omega e_1 + (\Omega e_1)^2\}$$

$$t_{new} - S_y^2 = S_y^2\{e_0 - \Omega e_1 + (\Omega e_1)^2 - \Omega e_0 e_1\}$$

$$E(t_{new} - S_y^2) = E[S_y^2\{e_0 - \Omega e_1 + (\Omega e_1)^2 - \Omega e_0 e_1\}]$$

$$B(t_{new}) = \Omega\lambda' S_y^2\{\Omega\beta_{2(x)}' - h'\}$$

MSE of the proposed estimator

Now comes to MSE of the estimator

$$MSE(t_{new}) \quad = \{S_y^2(e_0 - \Omega e_1)\}^2$$

(ignore higher order)

$$MSE(t_{new}) \quad = S_y^4 E\{(e_0^2) + \Omega^2(e_1^2) - 2\Omega(e_0 e_1)\}$$

$$= S_y^4\{\lambda' \beta'_{2(y)} + \Omega^2 \lambda' \beta'_{2(x)} - 2\Omega\lambda' h'\}$$

$$= \lambda' S_y^4 \left\{\beta'_{2(y)} + \beta'_{2(x)}\left(\Omega^2 - 2\Omega\frac{h'}{\beta'_{2(x)}}\right)\right\}$$

$$= \lambda' S_y^4\{\beta'_{2(y)} + \beta'_{2(x)}(\Omega^2 - 2\Omega v')\} \quad \rightarrow (i)$$

Where

$$v' = \frac{h'}{\beta'_{2(x)}}$$

In order to minimize MSE, Differentiate (i) partially w.r.t Ω and equating to zero.

$$\frac{\partial}{\partial\Omega} MSE(t_{new}) = 0$$

$$\lambda' S_y^4 \frac{\partial}{\partial\Omega}\{\beta'_{2(y)} + \beta'_{2(x)}(\Omega^2 - 2\Omega v')\} = 0$$

$$\Omega = v'$$

So,

$$= \lambda' S_y^4\{\beta'_{2(y)} - \beta'_{2(x)} v'^2\}$$

$$= \lambda' S_y^4 \beta'_{2(y)} \left\{1 - v'^2 \frac{\beta'_{2(x)}}{\beta'_{2(y)}}\right\}$$

$$= \lambda' S_y^4 \beta'_{2(y)} \left\{1 - \left(\frac{h'}{\sqrt{\beta'_{2(x)}\beta'_{2(y)}}}\right)^2\right\} MSE(t_{new})$$

$$= \lambda' S_y^4 \beta'_{2(y)}\{1 - (\rho')^2\}$$

Where

$$\rho' = \frac{h'}{\sqrt{\beta'_{2(x)}\beta'_{2(y)}}}$$

5. NUMERICAL STUDY

The execution of the proposed modified ratio type variance estimator is surveyed with that of traditional ratio type estimator and existing modified ratio type variance estimators recorded in table 1 for certain common populations.The population parameters and the constants computed are given below:

Table 1: Parameters and Constants of Populations

	Data 1	Data 2	Data 3	Data 4
N	278	103	80	70
n	30	40	20	25
$\bar{Y}$	39.068	62.6212	51.8264	96.7000
$\bar{X}$	25.111	556.5541	11.2646	175.2671
S_y	56.457167	91.3549	18.3569	60.714
S_x	40.674797	610.1643	8.4563	140.8572
$\beta_2(y)$	25.8969	37.1279	2.2667	4.7596
$\beta_2(x)$	38.8898	17.8738	2.8664	7.0952
h	26.8142	17.2220	2.2209	4.6038
ρ	0.7213	0.7298	0.9413	0.7293

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 2: Biases of the existing and proposed modified ratio type variance estimator

Estimators	Bias			
	Data 1	Data 2	Data 3	Data 4
Isaki[1]	1282.99696	135.993471	10.8758932	367.35093
Kadilar and Cingi[2]	1413.12116	136.169037	12.6980437	367.803965
Kadilar and Cingi[2]	-138.3148	-157.98237	33.3263251	731.217151
Kadilar and Cingi[2]	1288.20345	136.004239	11.328445	367.402218
Kadilar and Cingi[2]	-67.864364	-178.34129	-3.3473142	-57.03954
Proposed Estimator	-241.92096	-784.70582	-3.362946	-77.110599

Table 3: MSE of the existing and proposed modified ratio type variance estimators

Estimators	Mean Squared Error(MSE)			
	Data 1	Data 2	Data 3	Data 4
Isaki[1]	3778793.028	35796611.55	3924.948201	1438805.637
Kadilar and Cingi[2]	3983111.977	35796720.6	4249.507513	1439774.898
Kadilar and Cingi[2]	2514073.351	35823876.09	8667.305927	2300502.295
Kadilar and Cingi[2]	3786820.836	35796618.23	4003.90533	1438915.348
Kadilar and Cingi[2]	7990304.812	59839871.24	3636.702702	1557091.884
Proposed Estimator	2614574.728	45334825.03	3750.092953	1258334.941

From Table 2, 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 Table 3, 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.

6. CONCULSION

In this article we have proposed a modified ratio type variance estimator using known value of population mean of 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. 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 is less than the biases and mean squared errors of the traditional and existing estimators for certain known populations.

7. REFERENCES

[1] Isaki, C. T. (1983). Variance estimation using auxiliary information. *Journal of the American Statistical Association*, *78*(381), 117-123.

[2] Kadilar, C., & Cingi, H. (2005). A new ratio estimator in stratified random sampling. *Communications in Statistics—Theory and Methods*, *34*(3), 597-602.

[3] Das, A. K. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya, c*, *40*, 139-148.

[4] Cebrián, A. A., & García, M. R. (1997). Variance estimation using auxiliary information: An almost unbiased multivariate ratio estimator. *Metrika*, *45*(1), 171-178.

[5] Ahmed, M. S., Raman, M. S., & Hossain, M. I. (2000). Some competitive estimators of finite population variance using multivariate auxiliary information. *International journal of information and management sciences*, *11*(1), 49-54.

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

[7] Kadilar, C., & Cingi, H. (2006b). Ratio estimators for the population variance in simple and stratified random sampling. *Applied Mathematics and Computation*, *173*(2), 1047-1059.

[8] Gupta, S., & Shabbir, J. (2008). Variance estimation in simple random sampling using auxiliary information. *Hacettepe Journal of mathematics and Statistics*, *37*(1), 57-67.

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

[10] Subramani, J., & Kumarapandiyan, G. (2012b). Variance Estimation Using Median of the Auxiliary Variable. *International Journal of Probability and Statistics*, *1*(3), 36–40.

[11] Subramani, J., & Kumarapandiyan, G. (2013). Estimation of variance using known coefficient of variation and median of an auxiliary variable. *Journal of Modern Applied Statistical Methods*, *12*(1), 11.