

Original Research Paper

A New Modified approach for the improvement of new estimator using known value of Downton's method as auxiliary information for estimating the population variance

***M A Bhat, S A Mir, S Maqbool, T.A.Raja and Immad.A.Shah.
Division of Agricultural Statistics SKUAST-Kashmir (190025)**

Abstract: In the present study we have suggested a newly modified ratio estimator for estimating the population variance by utilizing the auxiliary information of Downton's method. Properties associated with proposed estimators are assessed by mean square error and bias through numerical demonstration. We have also provided the theoretical efficiency comparison of the present study.

Keywords: Ratio estimator, Downton's Method; SRSWOR, MSE, Bias and Efficiency.

1. 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 $U_i, i = 1, 2, 3, \dots, N$ given a vector $[Y_1, Y_2, Y_3, \dots, Y_N]$. Sometimes in sample surveys information on auxiliary variable X correlated with study variable Y is available can be utilized to obtain the efficient estimator for the estimation of Population Variance. "Here in this

modification, we aim to estimate the population variance $S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2$ on the

bases of the random sample selected from population U " to obtain the improved and precise estimates. Various authors have proposed different ratio estimators for finite population variance by utilizing the ancillary information of different parameters of the population such as Isaki [2], Kadilar and Cingi [3], Subramani and Kumarapandiyan [5] and Bhat *et al.* [1]. According to WEIR *et al.*, [6] the random processes implicated in gamete creation help to increase variation that is basically away from control to accurate the increased sample sizes. Irrespective of the census (survey study of the entire population), their results are still affected by "sampling" in that the particular residents sampled is but one of the many possible replicates that could have arisen under the same conditions. Moustafa *et al.* [7] suggest that their proposed methods have useful properties for heavy-tailed distributions and moderate sample sizes and they can perform better than the

control charts based on range, especially for heavy-tailed distributions. In their Study, Garcia, et. Al.[8] also propose a new unbiased estimator for the population variance infinite population sample surveys using auxiliary information. Singh. et al.,[9] concluded that out of many ratios, product and regression methods of estimation are good examples for estimating the finite population mean.

2. Material and Methods:

2.1 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. $B(\cdot)$ = Bias of the estimator.

$MSE(\cdot)$ = Mean square error. $\hat{S}_R^2$ = Ratio type variance estimator. D = Downtown's method

3. Suggested Estimator:

$$\hat{S}_{MS1}^2 = s_y^2 \left[\frac{S_x^2 + D}{s_x^2 + D} \right]$$

3.1 We have derived here the bias and mean square error of the proposed estimator to the 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

$s_x^2 = S_x^2(1 + e_1)$ from the definition of e_0 and e_1 we obtain:

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

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

The proposed estimator is given below:

$$\hat{S}_p^2 = s_y^2 \left[\frac{S_x^2 + \alpha D}{s_x^2 + \alpha D} \right]$$

$$\Rightarrow \hat{S}_p^2 = s_y^2(1 + e_0) \left[\frac{S_x^2 + \alpha D}{s_x^2 + e_1 S_x^2 + \alpha D} \right]$$

$$\Rightarrow \hat{S}_p^2 = \frac{S_y^2(1+e_0)}{(1+A_p e_1)} \text{ where } A_p = \frac{S_x^2}{(S_x^2 + \alpha D)}$$

$$\Rightarrow \hat{S}_p^2 = S_y^2(1+e_0)(1+A_p e_1)^{-1}$$

$$\Rightarrow \hat{S}_p^2 = S_y^2(1+e_0)(1-A_p e_1 + A_p^2 e_1^2 - A_p^3 e_1^3 + \dots)$$

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

$$\hat{S}_p^2 = S_y^2 + S_y^2 e_0 - S_y^2 A_p e_1 - S_y^2 A_p e_0 e_1 + S_y^2 A_p^2 e_1^2$$

$$\Rightarrow \hat{S}_p^2 - S_y^2 = S_y^2 e_0 - S_y^2 A_p e_1 - S_y^2 A_p e_0 e_1 + S_y^2 A_p^2 e_1^2$$

(1)

By taking expectation on both sides of (1), we get

$$E(\hat{S}_p^2 - S_y^2) = S_y^2 E(e_0) - S_y^2 A_p E(e_1) - S_y^2 A_p E(e_0 e_1) + S_y^2 A_p^2 E(e_1^2)$$

$$BiasE(\hat{S}_p^2) = S_y^2 A_p^2 E(e_1^2) - S_y^2 A_p E(e_0 e_1)$$

$$BiasE(\hat{S}_p^2) = \gamma S_y^2 A_p [A_p (\beta_{2(x)} - 1) - (\lambda_{22} - 1)]$$

(2)

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

$$E(\hat{S}_p^2 - S_y^2)^2 = S_y^4 E(e_0^2) - S_y^4 A_p^2 E(e_1^2) - 2S_y^4 A_p E(e_0 e_1)$$

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

4. Efficiency conditions: We have derived the efficiency conditions of proposed estimators with other existing estimators under which recommended estimator have performed better than the existing estimators

The bias and Mean square error of existing ratio type estimators up to the first order of approximation is given by

$$BiasE(\hat{S}_K^2) = \gamma S_y^2 R_K [R_K (\beta_{2(x)} - 1) - (\lambda_{22} - 1)]$$

(3)

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

(4)

Where $R_K = \text{Constant}$, $K = 1, 2, 3$

Bias, MSE and constant of proposed estimators is given by

$$Bias(\hat{S}_p^2) = \gamma S_y^2 R_p [R_p (\beta_{2x} - 1) - (\lambda_{22} - 1)]$$

(5)

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

(6)

$R_p = \text{Constant}$, $p = 1$

From Equation (4) and (6), we have

$$MSE(\hat{S}_p^2) \leq MSE(\hat{S}_K^2) \text{ if } \lambda_{22} \geq 1 + \frac{(R_p + R_K)(\beta_{2x} - 1)}{2}$$

$$MSE(\hat{S}_p^2) \leq MSE(\hat{S}_K^2)$$

$$\gamma S_y^4 [(\beta_{2y} - 1) + R_p^2 (\beta_{2x} - 1) - 2R_p (\lambda_{22} - 1)] \leq \gamma S_y^4 [(\beta_{2y} - 1) + R_K^2 (\beta_{2x} - 1) - 2R_K (\lambda_{22} - 1)]$$

(7)

$$\Rightarrow [(\beta_{2y} - 1) + R_p^2(\beta_{2x} - 1) - 2R_p(\lambda_{22} - 1)] \leq [(\beta_{2y} - 1) + R_K^2(\beta_{2x} - 1) - 2R_K(\lambda_{22} - 1)]$$

(8)

$$\Rightarrow [1 + R_p^2(\beta_{2x} - 1) - 2R_p(\lambda_{22} - 1)] \leq [1 + R_K^2(\beta_{2x} - 1) - 2R_K(\lambda_{22} - 1)]$$

(9)

$$\Rightarrow (\beta_{2x} - 1)(R_p^2 - R_K^2) [-2R_p(\lambda_{22} - 1)] \leq [-2R_K(\lambda_{22} - 1)]$$

(10)

$$\Rightarrow (\beta_{2x} - 1)(R_p^2 - R_K^2) [-2(\lambda_{22} - 1)(R_p - R_K)] \leq 0$$

(11)

$$\Rightarrow (\beta_{2x} - 1)(R_p^2 - R_K^2) \leq [2(\lambda_{22} - 1)(R_p - R_K)]$$

(12)

$$\Rightarrow (\beta_{2x} - 1) \leq \frac{2(\lambda_{22} - 1)(R_p - R_K)}{(R_p^2 - R_K^2)}$$

(13)

$$\Rightarrow (\beta_{2x} - 1) \leq \frac{2(\lambda_{22} - 1)(R_p - R_K)}{(R_p - R_K)(R_p + R_K)}$$

(14)

$$\Rightarrow (\beta_{2x} - 1)(R_p + R_K) \leq 2(\lambda_{22} - 1)$$

(15)

By solving equation (15), we get

$$MSE(\hat{S}_p^2) \leq MSE(\hat{S}_K^2) \text{ if } \lambda_{22} \geq 1 + \frac{(R_p + R_K)(\beta_{2x} - 1)}{2}$$

5. Numerical Illustration:

We use the data of Murthy [4] 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:

$N=80$, $S_x=8.4542$, $n=20$, $C_x=0.7507$, $\bar{X} = 11.2624$, $\beta_{2(x)} = 2.8664$, $\bar{Y} = 51.8264$,

$\beta_{2(y)} = 2.8664$, $\rho = 0.9413$, $\beta_{1(x)} = 1.05$, $\lambda_{22} = 2.2209$, $S_y=18.3569$, $Md=7.5750$,

$Q1=9.318$, $C_y=0.3542$, $G = 9.0408$, $D = 8.0138$, $S_{pw} = 7.9136$

Table 1: Bias and Mean Square Error of the existing and the proposed estimators

Estimators	Bias	Mean Square Error
Isaki	10.8762	3925.1622
Kadilar&Cingi	10.4399	3850.1552
Subramani&Kumarapandiyan	6.1235	3180.7740
Proposed	5.1964	2469.7695

Table 2: Percent relative efficiency of proposed estimators with existing estimators

Estimators	Isaki	Kadilar and Cingi	Subramani and Kumarapandiyan
Proposed	158.9282	155.8912	128.7882

Conclusion

From the above study, we revealed that our suggested estimator performs better than the existing estimators regarding bias mean square error since they are significantly lower than the corresponding values of the current estimators.

References:

- [1]. Bhat M. A., Maqbool, S., Saraf, S. A., Rouf A. and Malik, S. H. Variance estimation using linear combination of skewness and quartiles. *Journal of Advances in Research* 13(2): 1-6 , 2018 Article no.37321. ISSN: 2348-0394 , NLMID: 101666096
- [2]. Isaki, C.T. (1983). Variance estimation using auxiliary information. *Journal of the American Statistical Association*,78,117-123.
- [3]. Kadilar, C. & Cingi, H.(2006a).Improvement in Variance estimation using auxiliary information *Hacettepe Journal of Mathematics and Statistics*, 35(1).117-115.
- [4]. M. N Murthy, Sampling theory, *Theory and Methods*, Statistical Publishing Society, Calcutta,1967.
- [5]. Subramani, J. and Kumarapandiyan, G. (2015). Generalized modified ratio type estimator for estimation of population variance. *Sri-Lankan journal of applied Statistics*,vol16-1,69-90.
- [6] WEIR, B. S, COCKERHAM, C. C. ESTIMATING F-statistics FOR THE ANALYSIS OF POPULATION STRUCTURE; *Evolution*. 38(6). 1984, pp. 1358-1370
- [7]Moustafa O. Shawiesh A, Mokhtar B. Abdullah (2000) ESTIMATING THE PROCESS STANDARD DEVIATION BASED ON DOWNTON'S ESTIMATOR, *Quality Engineering*, 12:3, 357-363.
- [8] Garcia, M. R, Cebrian, A. A., Repeated substitution method: The ratio estimator for the population variance; *Metrika* (1996) 43: 101.
- [9] Singh, P. H, Vishwakarma, G.K; Modified Exponential Ratio and Product Estimators for Finite Population Mean in Double Sampling. *AUSTRIAN JOURNAL OF STATISTICS*: Vol 36 (2007), Number 3, 217-225
- [10]. M. A. Bhat, S. Maqbool, T. A. Raja, and Mir Subzar. Variance Estimation using Linear Combination of Hodges Lemma and Deciles . *IJETCAS* 17-216; © 2017. ISSN (Print): 2279-0047 ISSN (Online): 2279-0055