Journal of Scientific Research & Reports

18(6): 1-6, 2018; Article no.JSRR.40222
ISSN: 2320-0227

A Robust Approach to Survey Estimation Using Non-conventional Auxiliary Population Parameters

M. A. Bhat[1*], S. Maqbool[1], S. A. Mir[1] and Tehleel[1]

[1]Division of Agricultural Statistics, Sher-e-Kashmir University of Agricultural Sciences and Technology, Kashmir, India.

Authors' contributions

This work was carried out in collaboration between all authors. Author MAB designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Authors SM and SAM managed the analyses of the study. Author Tehleel managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JSRR/2018/40222
Editor(s):
(1) Shunpu Zhang, Department of Statistics, University of Nebraska – Lincoln, USA.
Reviewers:
(1) Angela Shirley, The University of the West Indies, Trinidad and Tobago.
(2) Gargi Tyagi, Banasthali Vidyapith, India.
Complete Peer review History: http://www.sciencedomain.org/review-history/24246

Original Research Article

Received 31st January 2018
Accepted 10th April 2018
Published 20th April 2018

ABSTRACT

In the present study, we have utilized the tool of modification technique for the precision and improvement of estimators to estimate the finite population variance, when the survey under investigation does not account for some of the important conditional distribution properties in its nature. The comparison and efficiency conditions of existing estimators with proposed estimators have been made through empirical study to seek the efficiency of modified and suggested estimators over existing estimators. Expressions for bias and mean square error have been derived up to the first order of approximation. The efficiency of modified and developed estimators is clearly based on the lesser mean square error of proposed estimators.

Keywords: Simple random sampling; bias; mean square error; standard deviation; tri-mean and deciles; efficiency.

**Corresponding author: E-mail: mabhat.1500@gmail.com;*

1. INTRODUCTION

The methodology of developing the efficient estimators has been widely discussed by the researchers in order to contribute their modern and creative approach to the theory of survey estimation. Some of them who are most popular in the area of survey estimation are: - Isaki [1], who developed the ratio and regression estimator under the utilization of known auxiliary information. Upadhyaya and Singh [2], where the authors have developed the estimators by using the coefficient of kurtosis as an auxiliary variable for its improvement. Kadilar and Cingi [3], proposed improved ratio estimators through utilization of auxiliary information as the coefficient of skewness. Subramani, J and Kumarapandiyan.G [4], developed the estimators through the technique of using quartiles as auxiliary variables to enhance the efficiency of estimators. On the same guidelines, several authors have developed different estimators by using this auxiliary information in different forms to obtain the efficient and precise estimators for the estimation of finite population variance in a sample survey. Some researchers utilize their efforts to contribute their creative ideas to the variance estimation and have used different population parameters as auxiliary variables to improve the precision and efficiency of variance estimators. Similarly, Bhat et al. have used linear combination of skewness and quartiles as auxiliary information to obtain the precision of estimators [5-11].

Let the finite population under survey be $U = \{U_1, U_2, ..., U_N\}$, consists of N distinct and identifiable units. Let Y be a real variable with value Y_i measured on $U_i, i = 1,2,3.......N$, giving a vector $Y = \{y_1, y_2, ..., y_N\}$. The goal is to estimate the populations mean $\bar{Y} = \frac{1}{N}\sum_{I=1}^{N} y_i$ or its variance $S_Y^2 = \frac{1}{N-1}\sum_{I=1}^{N}(y_i - \bar{y})^2$ on the basis of the random sample selected from a population U. In this paper, our aim is to provide the robust measures for the estimation of finite population variance when the population under investigation is non-normal or badly skewed as non-conventional parameters are robust measures against skewed populations.

2. MATERIALS 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, $\rho =$ 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, M_d= Median of the auxiliary variable, B(.)=Bias of the estimator, MSE(.)= Mean square error, $\hat{S}_R^2 =$ Ratio type variance estimator $\hat{S}_{Kc1}^2, \hat{S}_{jG}^2, =$ Existing modified ratio estimators, TR=Tri-mean, D_i =Deciles, i=1,2,3…10.

2.2 Existing Estimators from the Literature

2.2.1 Ratio type variance estimator proposed by Isaki [1]

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

Bias and mean square error are given as

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

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

2.2.2 Ratio type variance estimator proposed by Kadilar and Cingi [3]

$$\hat{S}_{kc1}^2 = s_y^2 \left[\frac{S_x^2 + C_x}{s_x^2 + C_x}\right]$$

Expressions for bias and mean square error are given as

Bias ($(\hat{S}^2_{kc1})$) =

$$\gamma S_y^2 A_1 \left[A_1\left(\beta_{2(x)}-1\right)-(\lambda_{22}-1) \right]$$

MSE ($(\hat{S}^2_{kc1})$) =

$$\gamma S_y^4 \left[\left(\beta_{2(y)}-1\right)+A_1^2 \left(\beta_{2(x)}-1\right)- 2A_1(\lambda_{22}-1) \right]$$

2.3 Recent Developments

2.3.1 Ratio type variance estimator proposed by Subramani. J and Kumarapandiyan, G [4]

$$\hat{S}^2_{jG} = s_y^2 \left[\frac{S_x^2 + \alpha w_i}{s_x^2 + \alpha w_i} \right]$$

Bias ($(\hat{S}^2_{jG})$) =

$$\gamma S_y^2 A_{jG} \left[A_{jG}\left(\beta_{2(x)}-1\right)-(\lambda_{22}-1) \right]$$

MSE ($(\hat{S}^2_{jG})$) =

$$\gamma S_y^4 \left[\left(\beta_{2(y)}-1\right)+A_{jG}^2 \left(\beta_{2(x)}-1\right)- 2A_{jG}(\lambda_{22}-1) \right]$$

where α=1 and W_i= Q_i, i=1, 2, 3, Qa, Qd, Qr.

3. MODIFIED AND DEVELOPED ESTIMATORS

$$\hat{S}^2_{MS1} = s_y^2 \left[\frac{S_x^2 + (TM + D_1)}{s_x^2 + (TM + D_1)} \right]$$

$$\hat{S}^2_{MS2} = s_y^2 \left[\frac{S_x^2 + (TM + D_2)}{s_x^2 + (TM + D_2)} \right]$$

$$\hat{S}^2_{MS3} = s_y^2 \left[\frac{S_x^2 + (TM + D_3)}{s_x^2 + (TM + D_3)} \right]$$

$$\hat{S}^2_{MS4} = s_y^2 \left[\frac{S_x^2 + (TM + D_4)}{s_x^2 + (TM + D_4)} \right]$$

$$\hat{S}^2_{MS5} = s_y^2 \left[\frac{S_x^2 + (TM + D_5)}{s_x^2 + (TM + D_5)} \right]$$

$$\hat{S}^2_{MS6} = s_y^2 \left[\frac{S_x^2 + (TM + D_6)}{s_x^2 + (TM + D_6)} \right]$$

$$\hat{S}^2_{MS7} = s_y^2 \left[\frac{S_x^2 + (TM + D_7)}{s_x^2 + (TM + D_7)} \right]$$

$$\hat{S}^2_{MS8} = s_y^2 \left[\frac{S_x^2 + (TM + D_8)}{s_x^2 + (TM + D_8)} \right]$$

$$\hat{S}^2_{MS9} = s_y^2 \left[\frac{S_x^2 + (TM + D_9)}{s_x^2 + (TM + D_9)} \right]$$

$$\hat{S}^2_{MS10} = s_y^2 \left[\frac{S_x^2 + (TM + D_{10})}{s_x^2 + (TM + D_{10})} \right]$$

3.1 We Have Derived Here the Bias and Mean Square Error of the Proposed Estimator

$\hat{S}^2_{MSi};\ i = 1,2,...10$ up 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 $s_y^2 = S_y^2(1+e_0)$ $s_x^2 = S_x^2(1+e_0)$ and from the definition of e_0 and e_1 we obtain:

$$E[e_0] = E[e_1] = 0 \qquad E[e_0^2] = \frac{1-f}{n}(\beta_{2(y)}-1)$$

$$E[e_1^2] = \frac{1-f}{n}(\beta_{2(x)}-1)$$

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

The proposed estimator $\hat{S}^2_{MSi};\ i = 1,2,3,...10$ is given below:

$$\hat{S}^2_{MSi} = s_y^2 \left[\frac{S_x^2 + \alpha a_i}{s_x^2 + \alpha a_i} \right] \tag{1}$$

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

$$\Rightarrow \quad \hat{S}^2_{MSi} = \frac{S_y^2(1+e_0)}{(1+A_{MSi}e_1)}$$

where $A_{MSi} = \frac{S_x^2}{S_x^2 + \alpha a_i}$

$$a_i = (TM + D_i);\ i = 1,2,3,\ldots,10$$

$$\Rightarrow \quad \hat{S}^2_{MSi} = S^2_y (1+e_0)(1+A_{MSi}e_1)^{-1} \tag{2}$$

$$\Rightarrow \hat{S}^2_{MSi} = S^2_y (1+e_0)(1 - A_{MSi}e_1 + A^2_{MSi}e_1^2 - A^3_{MSi}e_1^3 + \ldots..) \tag{3}$$

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

$$\hat{S}^2_{MSi} = S^2_y + S^2_y e_0 - S^2_y A_{MSi} e_1 - S^2_y A_{MSi} e_0 e_1 + S^2_y A^2_{MSi} e_1^2 \tag{4}$$

$$\Rightarrow \hat{S}^2_{MSi} - S^2_y = S^2_y e_0 - S^2_y A_{MSi} e_1 - S^2_y A_{MSi} e_0 e_1 + S^2_y A^2_{MSi} e_1^2 \tag{5}$$

By taking expectation on both sides of (5), we get

$$E(\hat{S}^2_{MSi} - S^2_y) = S^2_y E(e_0) - S^2_y A_{MSi} E(e_1) - S^2_y A_{MSi} E(e_0 e_1) + S^2_y A^2_{MSi} E(e_1^2) \tag{6}$$

$$Bias(\hat{S}^2_{MSi}) = S^2_y A^2_{MSi} E(e_1^2) - S^2_y A_{MSi} E(e_0 e_1) \tag{7}$$

$$Bias(\hat{S}^2_{MSi}) = \gamma S^2_y A_{MSi} [A_{MSi}(\beta_{2(x)} - 1) - (\lambda_{22} - 1)] \tag{8}$$

Squaring both sides of (5) and (6), neglecting the terms more than 2nd order and taking expectation, we get

$$E(\hat{S}^2_{MSi} - S^2_y)^2 = S^4_y E(e_0^2) + S^4_y A^2_{MSi} E(e_1^2) - 2 S^4_y A_{MSi} E(e_0 e_1)$$

$$MSE(\hat{S}^2_{MSi}) = \gamma S^4_y [(\beta_{2(y)} - 1) + A^2_{MSi}(\beta_{2(x)} - 1) - 2A_{MSi}(\lambda_{22} - 1)]$$

3.2 Efficiency Conditions

We have derived the efficiency conditions of proposed estimators with other existing estimators under which proposed estimators are performing 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

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

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

$$R_K = Existing.cons\tan t$$
$$, K = 1,2,3,4\ldots\ldots..$$

Where

Bias, MSE and constant of proposed estimators is given by

$$Bias(\hat{S}^2_P) = \gamma S^2_y R_P [R_P(\beta_{2x} - 1) - (\lambda_{22} - 1)] \tag{3}$$

$$MSE(\hat{S}^2_P) = \gamma S^4_y [(\beta_{2y} - 1) + R^2_P(\beta_{2x} - 1) - 2R_P(\lambda_{22} - 1)] \tag{4}$$

$$R_P = proposed.cons\tan t$$
$$P = 1,2,3\ldots\ldots\ldots\ldots$$

From Equation (2) and (3), we have

$$MSE(\hat{S}^2_P) \le MSE(\hat{S}^2_K) if \lambda_{22} \ge 1 + \frac{(R_P + R_K)(\beta_{2x} - 1)}{2}$$

$$MSE(\hat{S}^2_P) \le MSE(\hat{S}^2_K)$$
$$\gamma S^4_y [(\beta_{2y} - 1) + R^2_P(\beta_{2x} - 1) - 2R_P(\lambda_{22} - 1)] \le \gamma S^4_y [(\beta_{2y} - 1) + R^2_K(\beta_{2x} - 1) - 2R_K(\lambda_{22} - 1)] \tag{5}$$

$$\Rightarrow [(\beta_{2y} - 1) + R^2_P(\beta_{2x} - 1) - 2R_P(\lambda_{22} - 1)] \le [(\beta_{2y} - 1) + R^2_K(\beta_{2x} - 1) - 2R_K(\lambda_{22} - 1)] \tag{6}$$

$$\Rightarrow [1 + R^2_P(\beta_{2x} - 1) - 2R_P(\lambda_{22} - 1)] \le [1 + R^2_K(\beta_{2x} - 1) - 2R_K(\lambda_{22} - 1)] \tag{7}$$

$$\Rightarrow (\beta_{2x} - 1)(R^2_P - R^2_K)\ [-2R_P(\lambda_{22} - 1)] \le [-2R_K(\lambda_{22} - 1)] \tag{8}$$

$$\Rightarrow (\beta_{2x} - 1)(R^2_P - R^2_K)\ [-2(\lambda_{22} - 1)(R_P - R_K)] \le 0 \tag{9}$$

$$\Rightarrow \quad (\beta_{2x} - 1)(R^2_P - R^2_K) \quad \le [2(\lambda_{22} - 1)(R_P - R_K)] \tag{10}$$

$$\Rightarrow (\beta_{2x}-1) \le \frac{2(\lambda_{22}-1)(R_P-R_K)}{(R_P^2-R_K^2)} \quad (11)$$

$$\Rightarrow (\beta_{2x}-1) \le \frac{2(\lambda_{22}-1)(R_P-R_K)}{(R_P-R_K)(R_P+R_K)} \quad (12)$$

$$\Rightarrow (\beta_{2x}-1)\,(R_P+R_K) \le 2(\lambda_{22}-1) \quad (13)$$

By solving equation (13), we get

$$MSE(\hat{S}_P^2) \le MSE(\hat{S}_K^2)\, if\, \lambda_{22} \ge 1 + \frac{(R_P+R_K)(\beta_{2x}-1)}{2}$$

4. NUMERICAL ILLUSTRATION

We use the data of Murthy (1967) 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, Sx=8.4542, n=20, Cx=0.7507, $\overline{X} =$ 11.2624, $\beta_{2(x)} =$ 2.8664, $\overline{Y} =$ 51.8264, $\beta_{2(y)} =$ 2.2667, $\rho =$ 0.9413, $\beta_{1(x)} =$ 1.05, $\lambda_{22} =$ 2.2209, S_y=18.3569, D= 8.0138, D_1= 3.6, D_2 = 4.6, D_3= 5.9, D_4 = 6.7, D_5 =7.5, D_6 = 8.5, D_7 =14.8, D_8 = 18.1, D_9 = 25, D_{10} =34.8, TM= 9.318.

5. DISCUSSION AND CONCLUSION

In this manuscript, modified and developed variance estimators clearly reveals that proposed estimators are more efficient than the other existing estimators, as the bias, mean square error and percent relative efficiency of proposed estimators is less than the other existing estimators which can be easily seen from the Tables 1 and 2 respectively.

Table 1. Bias and mean square error of the existing and the proposed estimators

Estimators	Bias	Mean square error
Isaki [1]	10.8762	3925.1622
Kadilar&Cingi [3]	10.4399	3850.1552
Subramani&Kumarapandiyan [4]	6.1235	3180.7740
Proposed (MS1)	5.0470	2960.9612
Proposed (MS2)	4.5460	2947.9201
Proposed (MS3)	4.3950	2895.5767
Proposed (MS4)	3.9870	2952.3546
Proposed (MS5)	3.9330	2901.2567
Proposed (MS6)	3.6590	2838.8022
Proposed (MS7)	2.1250	2747.3930
Proposed (MS8)	1.5360	2712.1913
Proposed (MS9)	0.4430	2657.1109
Proposed (MS10)	-0.6916	2656.5055

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

Estimators	Isaki [1]	Kadilar & Cingi [3]	Subramani & Kumarapandiyan [4]
P1	132.5637	131.6519	107.4243
P2	133.1502	132.2343	107.9335
P3	135.5571	132.2476	109.9061
P4	132.9502	132.0649	107.7504
P5	135.2917	134.3611	109.6350
P6	138.2682	137.3171	112.0470
P7	142.8686	141.8858	115.7749
P8	144.7229	143.7274	117.2776
P9	147.7229	146.7216	119.7087
P10	147.7565	146.7550	119.7360

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Isaki CT. Variance estimation using auxiliary information. Journal of the American Statistical Association. 1983;78: 117-123.
2. Upadhyaya LN, Singh HP. Use of auxiliary variable in the estimation of population variance. Mathematical Forum. 1999;4:33-36.
3. Kadilar C, Cingi H. Improvement in variance estimation using auxiliary information. Hacettepe Journal of Mathematics and Statistics. 2006a;35(1): 117-115.
4. Subramani J, Kumarapandiyan G. Generalized modified ratio type estimator for estimation of population variance. Sri-Lankan Journal of Applied Statistics. 2015;16(1):69-90.
5. Singh D, Chaudhary FS. Theory and analysis of sample survey designs. New Age Publishers; 1986.
6. Murthy MN. Sampling theory. Theory and Methods, Statistical Publishing Society, Calcutta; 1967.
7. Arcos A, Rueda M, Martinez MD, Gonzalez S, Roman Y. Incorporatin the auxiliary information available in variance estimation. Applied Mathematical and Computation. 2005;160:387-399.
8. Subhash Kumar Yadav, Sheela Misra, Mishra SS. American Journal of Operational Research. 2016;6(1):9-15.
9. Bhat MA, Maqbool S, Saraf SA, Rouf AB, Malik SH. Journal of Advances in Research. 2018;13(2):1-6. Article No: 37321. ISSN: 2348-0394.
NLMID: 101666096

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history/24246