International Journal of Engineering Sciences &Research Technology

(A Peer Reviewed Online Journal)

Impact Factor: 5.164

IJESRT

IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

EFFICIENT CLASS OF ESTIMATORS FOR ESTIMATING THE POPULATION VARIANCE USING AUXILIARY VARIABLE AND ATTRIBUTE

Chandni Kumari[1] & Ratan Kumar Thakur*[2]
Department of Statistics, Babasaheb Bhimrao Ambedkar University
(A Central University), Lucknow-226025, India

DOI: 10.5281/zenodo.3828598

ABSTRACT

A family of log-type estimators using information on auxiliary information has been proposed for estimating the population variance of the study variable. It has been shown that these families of log-type estimators have lesser mean square error under the optimum values of the characterizing scalars as compared to some of the commonly used estimators available in the literature. Further, a comparative study is performed to judge the efficiency of proposed estimator. A numerical study is included as an illustration for the proposed work.

KEYWORDS: Ratio type estimator, bias, mean squared error, percent relative efficiency.

1. INTRODUCTION

Estimators obtained using auxiliary information are supposed to be more efficient than the estimators obtained without using auxiliary information. The ratio, regression, product and difference methods take advantage of the auxiliary information at the estimation stage. Many authors like, Pandey and Dubey (1988), Upadhyaya and Singh (1999) , Kadilar and Cingi (2003), Singh and Taylor (2003), Singh (2003), Sisodia and Dwivedi (1981), Koyuncu and Kadilar, Kumari et al. (2018a, 2018b, 2018c, 2018d, 2018e) along with many others have proposed various estimators using auxiliary information on various population parameters like coefficient of skewness, kurtosis, variation, standard deviation, correlation coefficient, etc. The literature deals with a wide range of ratio, product, difference and exponential estimators proposed by various renowned authors using multiple auxiliary information (Olkin (1958), Raj (1965), Singh (1967),Shukla (1966), etc.). Recently, Kumari and Thakur. (2019, 2020a, 2020b, 2020c, 2020d, 2020e, 2020f) had made the use of logarithmic relationship between the study variable and auxiliary information in form of variable and atrribute. The proposed estimators would work in case when the study variable is logarithmically related to the auxiliary variable and attribute.

Consider a finite population U = (U_1,U_2,...,U_N) of size N from which a sample of size n is drawn according to simple random sampling without replacement (SRSWOR). Let y_i , x_i and f_i denotes the values of the study variable, auxiliary variable and auxiliary attribute for the ith unit (i = 1,2,...,N), of the population. Further, let $\bar{y}$, $\bar{x}$ and $\bar{f}$ be the sample means and $s_y^2=\frac{\sum_{i=1}^{n}(yi-\bar{y})^2}{(n-1)}$, $s_x^2=\frac{\sum_{i=1}^{n}(xi-\bar{x})^2}{(n-1)}$ and $s_f^2=\frac{\sum_{i=1}^{n}(f-\bar{f})^2}{(n-1)}$ be the sample variance of the study variable, auxiliary variable and auxiliary attribute respectively.

2. THE SUGGESTED GENERALIZED CLASS OF LOG-TYPE ESTIMATORS

In this section we propose the following class of estimators for population variance using auxiliary information in form of both attribute as well as variable.

$$.T_c=w_1 s_y^2\left[1+\log\left(\frac{S_x^2}{s_x^2}\right)\right]^{a_1}\left[1+\log\left(\frac{S_f^2}{s_f^2}\right)\right]^{a_2}$$

such that a_1 and $a_2$are either are either real numbers or functions of the known parameters of the auxiliary variable x and auxiliary attribute f such as the standard deviations S_x, S_f, coefficient of variation C_x, C_f , coefficient of kurtosis b_{2x}, b_{2f} , coefficient of skewness b_{1x} , b_{1f} and correlation coefficient r of the population.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

3. PROPERTIES OF THE SUGGESTED CLASSES OF LOG-TYPE ESTIMATORS

In order to obtain the bias and mean square error (MSE), let us consider

$E(\epsilon_0)=E(\epsilon_1)=E(\epsilon_2)=0,\quad E(\epsilon_0)^2=I\,b_{2y}^*,\quad E(\epsilon_1)^2=I\,b_{2x}^*,\quad E(\epsilon_2)^2=I\,b_{2f}^*,\quad E(\epsilon_0\epsilon_1)=I\,I_{22yx}^*,$
$E(\epsilon_0\epsilon_2)=I\,I_{22yf}^*,\ E(\epsilon_1\epsilon_2)=I\,I_{22xf}^*,$

Where $b_{2x}^*=b_{2x}-1,\ b_{2f}^*=b_{2f}-1,\ b_{2y}^*=b_{2y}-1$ and $I_{22yf}^*=I_{22fy}-1, I_{22yx}^*=I_{22xy}-1,\ I_{22xf}^*=I_{22xf}-1,\quad I=\frac{1}{N},\ I_{pq}=m_{pq}\Big/ m_{20}^{\frac{p}{2}}m_{02}^{\frac{q}{2}}$

Theorem 1. The bias and mean squared error of the proposed estimators are given by

$$Bias(T_c)=S_y^2\left[w_1\left\{1+I\left(-a_1r_{yx}\sqrt{b_{2y}^*b_{2x}^*}-a_2r_{yf}\sqrt{b_{2y}^*b_{2f}^*}+a_1a_2r_{xf}\sqrt{b_{2f}^*b_{2x}^*}+\frac{a_2^2}{2}b_{2f}^*+\frac{a_1^2}{2}b_{2x}^*\right)\right\}-1\right]$$

$$MSE(T_c)=S_y^4+w_1^2S_y^4\left[1+I\left\{b_{2y}^*+2\,a_1^2b_{2x}^*+2\,a_2^2b_{2f}^*-4\,a_1r_{yx}\sqrt{b_{2y}^*b_{2x}^*}-4\,a_2r_{yf}\sqrt{b_{2y}^*b_{2f}^*}+4\,a_1r_{xf}\sqrt{b_{2x}^*b_{2f}^*}\right\}\right]-2w_1S_y^4\left[1+I\left\{\frac{a_1^2}{2}b_{2x}^*+a_1a_2r_{xf}\sqrt{b_{2f}^*b_{2x}^*}+\frac{a_2^2}{2}b_{2f}^*-a_1r_{yx}\sqrt{b_{2y}^*b_{2x}^*}-a_2r_{yf}\sqrt{b_{2y}^*b_{2f}^*}\right\}\right]$$

where

$$r_{yx}=\frac{I_{22yx}^*}{\sqrt{b_{2y}^*b_{2x}^*}},\quad r_{yf}=\frac{I_{22yf}^*}{\sqrt{b_{2y}^*b_{2f}^*}},\quad r_{xf}=\frac{I_{22xf}^*}{\sqrt{b_{2f}^*b_{2x}^*}}$$

respectively.

Proof. Consider the estimator,

$$.T_c=w_1s_y^2\left[1+\log\left(\frac{S_x^2}{s_x^2}\right)\right]^{a_1}\left[1+\log\left(\frac{S_f^2}{s_f^2}\right)\right]^{a_2}$$
$$=w_1S_y^2(1+\epsilon_0)[1+\log(1+\epsilon_1)^{-1}]^{a_1}[1+\log(1+\epsilon_2)^{-1}]^{a_2}$$

where $s_y^2=S_y^2(1+\epsilon_0),\ s_x^2=S_x^2(1+\epsilon_1), s_f^2=S_f^2(1+\epsilon_2),$

$$T_c=w_1S_y^2(1+\epsilon_0)[1+\log(1-\epsilon_1+\epsilon_1^2)]^{a_1}[1+\log(1-\epsilon_2+\epsilon_2^2)]^{a_2}$$
$$=w_1S_y^2(1+\epsilon_0)\left[1+a_1(\epsilon_1^2-\epsilon_1)-\frac{a_1}{2}(\epsilon_1^2-\epsilon_1)^2+\frac{a_1(a_1-1)}{2}(\epsilon_1^2-\epsilon_1)^2\right]\left[1+a_2(\epsilon_2^2-\epsilon_2)-\frac{a_2}{2}(\epsilon_2^2-\epsilon_2)^2+\frac{a_2(a_2-1)}{2}(\epsilon_2^2-\epsilon_2)^2\right]$$

On simplification, taking expectation on both the sides, we get the required expression of bias

$$Bias(T_c)=S_y^2\left[w_1\left\{1+I\left(-a_1r_{yx}\sqrt{b_{2y}^*b_{2x}^*}-a_2r_{yf}\sqrt{b_{2y}^*b_{2f}^*}+a_1a_2r_{xf}\sqrt{b_{2f}^*b_{2x}^*}+\frac{a_2^2}{2}b_{2f}^*+\frac{a_1^2}{2}b_{2x}^*\right)\right\}-1\right]$$

Squaring on both the sides of equation (2) and then taking expectation on both the sides, we get

$$MSE(T_c)=S_y^4+w_1^2S_y^4\left[1+I\left\{b_{2y}^*+2\,a_1^2b_{2x}^*+2\,a_2^2b_{2f}^*-4\,a_1r_{yx}\sqrt{b_{2y}^*b_{2x}^*}-4\,a_2r_{yf}\sqrt{b_{2y}^*b_{2f}^*}+4\,a_1r_{xf}\sqrt{b_{2x}^*b_{2f}^*}\right\}\right]-2w_1S_y^4\left[1+I\left\{\frac{a_1^2}{2}b_{2x}^*+a_1\right.\right.$$

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

$$a_2 r_{xf}\sqrt{b_{2f}^* b_{2x}^*} + \frac{a_2^2}{2} b_{2f}^* - a_1 r_{yx}\sqrt{b_{2y}^* b_{2x}^*} - a_2 r_{yf}\sqrt{b_{2y}^* b_{2f}^*}\Big\}\Big]$$

Corollary 1. The optimum value of constant is given by

$$w_{1_{opt}} = \frac{B}{A}$$

$$where\ A = \Big[1 + I\Big\{b_{2y}^* + 2\,a_1^2 b_{2x}^* + 2\,a_2^2 b_{2f}^* - 4\,a_1 r_{yx}\sqrt{b_{2y}^* b_{2x}^*} - 4\,a_2 r_{yf}$$

$$\sqrt{b_{2y}^* b_{2f}^*} + 4\,a_1 r_{xf}\sqrt{b_{2x}^* b_{2f}^*}\Big\}\Big]$$

$$B = \Big[1 + I\Big\{\frac{a_1^2}{2} b_{2x}^* + a_1 a_2 r_{xf}\sqrt{b_{2f}^* b_{2x}^*} + \frac{a_2^2}{2} b_{2f}^* - a_1 r_{yx}\sqrt{b_{2y}^* b_{2x}^*}$$

$$-a_2 r_{yf}\sqrt{b_{2y}^* b_{2f}^*}\Big\}\Big]$$

The optimum mean squared error is given by

$$MSE(T_c)_{opt} = S_y^4\left[1 - \frac{B^2}{A}\right]$$

4. COMPARISON OF ESTIMATORS

In this section, we compare the proposed classes of estimators with some important estimators. The comparison will be in terms of their MSEs up to the order of n^{-1}.

4.1 General estimator of population variance

$$t_o = s_y^2$$

$$MSE(t_o) = I\, S_y^4 b_{2y}^* > MSE(T_c)_{opt}$$

4.2 Ratio-type variance estimator

$$t_1 = s_y^2\left[\frac{S_x^2}{s_x^2}\right]\left[\frac{S_f^2}{s_f^2}\right]$$

$$MSE(t_1) = I\, S_y^4\left[b_{2y}^* + b_{2x}^* + b_{2f}^* - 2I_{22yx}^* - 2I_{22yf}^* + 2I_{22xf}^*\right] > MSE(T_c)_{opt}$$

4.3 Product-type variance estimator

$$t_2 = s_y^2\left[\frac{s_x^2}{S_x^2}\right]\left[\frac{s_f^2}{S_f^2}\right]$$

$$MSE(t_2) = I\, S_y^4\left[b_{2y}^* + b_{2x}^* + b_{2f}^* + 2I_{22yx}^* + 2I_{22yf}^* + 2I_{22xf}^*\right] > MSE(T_c)_{opt}$$

4.4 Das and Tripathi (1978) type variance estimator

$$t_3 = s_y^2\left[\frac{S_x^2}{S_x^2 + \alpha_1(s_x^2 - S_x^2)}\right]\left[\frac{S_f^2}{S_f^2 + \alpha_1(s_f^2 - S_f^2)}\right]$$

$$MSE(t_3) > MSE(T_c)_{opt}$$

4.5 Isaki (1983) variance estimator

$$t_4 = w_1\left[\frac{S_y^2}{S_x^2}\right] S_x^2 + w_2\left[\frac{S_y^2}{S_f^2}\right] S_f^2$$

$$MSE(t_4) > MSE(T_c)_{opt}$$

4.6 Singh, Chauhan, Sawan and Smarandache (2011) type variance estimator

$$t_5 = s_y^2 exp\left[\frac{S_x^2 - s_x^2}{S_x^2 + S_x^2}\right] exp\left[\frac{S_f^2 - s_f^2}{S_f^2 + S_f^2}\right]$$

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

$MSE(t_5) > MSE(T_c)_{opt}$

4.7 Olufadi and Kalidar (2014) variance estimator

$$t_6 = s_y^2 \left[\frac{S_x^2}{s_x^2}\right]^{a_1} \left[\frac{S_f^2}{s_f^2}\right]^{a_2}$$

$MSE(t_6) > MSE(T_c)_{opt}$

5. EMPIRICAL STUDY

The data on which we performed the numerical calculation is taken from some natural populations. The source of the data is as follows:

Population 1. (Singh S., Pg. no. 1114).
The data concerns Apples commercial crop, season average price (in $) per pound, by states 1994-1996.
y : season average price (in $) per pound in 1996
f : season average price (in $) per pound in 1995
x: season average price (in $) per pound in 1994.

Population 2. (Singh S., Pg. no. 1123).
The data concerns age specific death rates from 1990-2065.
y : per 100,000 births in 2040
x : per 100,000 births in 1990
f : per 100,000 births in 2000

Population 3. (Choudhary F. S., Pg. no. 117).
y : area under wheat (in acres) in 1974
f : area under wheat (in acres) in 1971
x : area under wheat (in acres) in 1973
The summary and the percent relative efficiency of the following estimators are as follows:

Table 1: Parameters of the data

Parameter	Population 1	Population 2	Population 3
N	36	22	34
n	11	8	10
b_{2y}^*	2.632	9.433	2.632
b_{2x}^*	0.114	7.105	0.114
b_{2f}^*	2.345	0.033	2.345
I_{22yx}^*	0.174	8.5711	0.174
I_{22yf}^*	2.014	-0.133	2.014
I_{22xf}^*	0.146	-0.126	0.146

Table 2: PRE of the estimators

Estimator	Population 1	Population 2	Population 3
t_o	100	100	100
t_1	261.515	5694.435	29.846
t_2	26.964	27.813	15.041
t_3	295.975	19336.11	652.430
t_4	240.421	322.055	86.972
t_5	299.317	20158.8	670.916
t_6	299.317	20158.8	670.916
T_c	299.315	20158.8	670.916

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

6. CONCLUSION

The present study extends the idea regarding the effective use of auxiliary information if the relationship between the study variable and the auxiliary variable is of logarithmic type. The present study provides some novel estimators that may be used when such auxiliary information is available. An empirical study shows that the proposed estimators are better than conventional estimators and equally efficient with some recent estimators.

REFERENCES

[1] Kumari, C and Thakur, R. K. (2020a). "A Family Of Logarithmic Estimators For Population Variance Under Double Sampling", International Journal of Mathematics Trends and Technology, 66 (4), 99-105
[2] Kumari, C and Thakur, R. K. (2020b). "Improved Ratio type Estimators Using Auxiliary attribute For Population Variance" , International Journal of Science and Research (IJSR) , Volume 9, Issue 4, pg 1491-1505 ,
DOI: 10.21275/SR20426180318 1491 ISSN: 2319-7064
[3] Kumari, C and Thakur, R. K. (2020c). "Family of Estimators for Population Variance using Two Auxiliary Information", International Journal for Research in Applied Science & Engineering Technology, Volume 8, Issue 5, pg 310-315
http://doi.org/10.22214/ijraset.2020.5051
[4] Kumari, C and Thakur, R. K. (2020d). "An Advanced Class of Log type estimators for Population Variance Using an Attribute and a Variable", International Journal of Industrial Engineering Research and Development, Volume 11, Issue 1, pg. 1-7
[5] Kumari, C and Thakur, R. K. (2020e). "An Improved Estimation of Population Variance Using the Coefficient of Kurtosis and Median of an Auxiliary Variable", International Journal of Engineering Sciences & Research Technology, Volume 9, Issue 4, pg. 168-175
[6] Bhushan, S., Kumari, C and Thakur, R. K. (2020f). "A Generalized Class of Ratio Type Estimator for Population Variance under Midzuno-Sen Type Sampling Scheme", Journal of Engineering Research and Application, Volume 10, Issue 4, pg. 55-60. DOI: 10.9790/9622-1004065560
[7] Kumari, C., Bhushan, S. And Thakur, R. K. (2019). Optimal Two Parameter Logarithmic Estimators for Estimating the Population Variance, Global Journal of Pure and Applied Mathematics, Volume 15, Issue 5, pg 527-237.
[8] Bhushan, S. and Kumari, C. (2018a). A new log type estimators for estimating the population variance, International Journal of Computational and Applied Mathematics (ISSN 1819-4966), 13 (1), 43-54.
[9] Bhushan, S. and Kumari, C. (2018b). A Class of Double Sampling Log-Type Estimators for Population Variance Using Two Auxiliary Variable, International Journal of Applied Engineering Research, Volume 13, Number (13) 2018, pg 11151-11155.
[10] Bhushan, S. and Kumari, C. (2018c). Some Classes of Log Type Estimators Using Auxiliary Attribute for Population Variance, International Journal of Scientific and Engineering Research, Volume 9, Issue 6, pg 1823-1832.
[11] Bhushan, S. and Kumari, C. (2018d). Estimation of Variance of Finite Population Using Double Sampling Scheme, International Journal of Scientific and Engineering Research, Volume 9, Issue 8, pg 1893-1901.
[12] Kumari, C., Bhushan, S. and Thakur, R. K. (2018e). ModifiedRatio Estimators Using Two Auxiliary Information for Estimating Population Variance in Two-Phase Sampling, International Journal of Scientific and Engineering Research, Volume 9, Issue 8, pg 18841892.
[13] Hidiroglou M. A. and Sarndal C. E. (1998). Use of auxiliary information for two-phase sampling , Survey Methodology, 24(1), 11-20.
[14] Neyman J. (1938). Contribution to the theory of sampling human populations, J. Amer. Stat. Asso., 33, 101-116.
[15] Cochran W. G. (1963). Sampling Techniques , Wiley Eastern Private Limited, New Delhi, 307-310.
[16] Chaudhury A. (1978). On estimating the variance of a finite population. Metrika, 25, 66-67.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[17] Das A. K. and Tripathi T. P. (1978). Use of auxiliary information in estimating the finite population variance. Sankhya, C(4), 139-148.
[18] Gupta S. and Shabbir J. (2008). Variance estimation in simple random sampling using auxiliary information, Hacettepe Journal of Mathematics and Statistics, 37, 57-67.
[19] Isaki C. T. (1983). Variance estimation using Auxiliary Information , Jour. Amer. Statist. Asssoct, 78, 117-123.
[20] Kadilar C. and Cingi H. (2006)a. Improvement in variance estimation using auxiliary information, Hacettepe Journal of Mathematics and Statistics, 1(35), 111-115.
[21] Kadilar C. and Cingi H. (2006)b. Ratio estimators for population variance in simple and stratied sampling, Applied Mathematics and Computation, 1(73), 1047-1058.
[22] Sukhatme P. V., Sukhatme B. V., Sukhatme S. and Ashok C. (1984). Sampling Theory of Surveys with Applications , Iowa State University Press, Ams.
[23] Swain A. K. P. C. and Mishra G. (1994). Estimation of population variance under unequal probability sampling ,Sankhya, B (56), 374-384.
[24] Singh, R., Chauhan, P., Sawan, N. &Smarandache, F. (2011), Improved exponential estimator for population variance using two auxiliary variables, Ital. J. Pure Appl. Math.s 28, 101108.