# A FAMILY OF LOGARITHMIC ESTIMATORS FOR POPULATION VARIANCE UNDER DOUBLE SAMPLING

Chandni Kumari and Ratan Kumar Thakur
Department of Statistics, Babasaheb Bhimrao
Ambedkar University, Lucknow-226025, India.

**Abstract**

In this paper, an improved estimator for variance has been proposed to improvise the log-type estimators proposed by Kumari (2019). These classes of estimators provide a better alternative to the classes of estimators provided by Kumari (2019) as well as to some other commonly used estimators available in literature. A numerical study is included to support the use of the suggested classes of estimators.

## 1 Introduction

In sample surveys, auxiliary information always plays a vital role in better estimation of the parameter under investigation. This information can be used to improve the precision of the estimators. The suitable utilization of this auxiliary information can reduce the MSE of the sample mean, thus resulting in more efficient estimators. This paper, is an attempt to extend the powerful Searls approach to the traditional estimators using auxiliary information regarding to variables in simple random sampling. Many authors like, Singh et al. (1973), Das and Tripathi (1978), Sisodia and Dwivedi (1981), Isaki (1983), Bahl and Tuteja (1991), Prasad and Singh (1992), Swain (1994), Garcia and Cebrian (1996), Upadhaya and Singh (2001), Kalidar and Cingi (2006a, 2006b); Gupta and Shabbir (2006, 2007), Yadav and Kadilar (2013, 2014) had proposed an improved ratio estimators using Searls type estimators.
Recently, Kumari et al. (2017, 2018) among others; Bhushan et al. (2015) have made the use of logarithmic relationship between the auxiliary variable and study variable as logarithm function which is very common in various branches of science as well as non-science disciplines. In this paper, some improved logarithmic estimators are proposed for improving the efficiency of the Bhushan and Kumari (2019) estimators as these classes of estimators are expected to improve the mean squared error. The proposed estimators would work considerably well in case when the study variable is logarithmically related to the auxiliary variable.

Consider a finite population $U = U_1, U_2, ..., U_N$ of size $N$. From this population we draw a large sample of size $n'$ to estimate the population parameters. Again from $n'$ a sample of size $n$ is drawn according to simple random sampling without replacement (SRSWOR) under double sampling. Let $y_i$ and $x_i$ denotes the values of the study and auxiliary variables for the $i$th unit $(i = 1, 2, ..., N)$, of the population. Further, let $\bar{x'}$ be the mean of larger sample, $\bar{y}$ and $\bar{x}$ be the sample means and $s_x^{2'} = \sum_{i=1}^{n} (x_i - \bar{x'})^2/(n-1)$ be the larger sample variance and $s_y^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2/(n-1)$ and$s_x^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2/(n-1)$ be the sample variance of the study and auxiliary variables respectively.

## 2 The proposed estimators

We propose the following new classes of log-type estimators for the population variance $S_y^2$ as:

$$T_1^* = w_1\, s_y^2 \left[1 + log\left(\frac{s_x^{'2}}{s_x^2}\right)\right]^{a_1} \quad (2.1)$$

$$T_2^* = w_2\, s_y^2 \left[1 + a_2\, log\left(\frac{s_x^{'2}}{s_x^2}\right)\right] \quad (2.2)$$

$$T_3^* = w_3\, s_y^2 \left[1 + log\left(\frac{s_x^{'2}}{s_x^{*2}}\right)\right]^{a_3} \quad (2.3)$$

$$T_4^* = w_4\, s_y^2 \left[1 + a_4\, log\left(\frac{s_x^{'2}}{s_x^{*2}}\right)\right] \quad (2.4)$$

where

$$s_x^{*2} = as_x^2 + b$$
$$s_x^{'*2} = as_x^{'2} + b$$

such that $a(\neq 0)$, $b$ are either real numbers or functions of the known parameters of the auxiliary variables $x$ such as the standard deviations $S_x$ , coefficient of variation $C_x$, coefficient of kurtosis $b_{2x}$, coefficient of skewness $b_{1x}$ and correlation coefficient $\rho$ of the population.
If $a_i = 0$, then the proposed estimator becomes the usual per unit variance estimator $s_y^2$ . If $a_i = +1$, then the proposed estimator works as ratio type estimator and if $a_i = -1$, then the proposed estimator works as product type estimator having efficiency conditions equivalent to that of generalized product and ratio estimators respectively.

## 3 Properties of proposed estimators

In this paper, the mean square error(s) of all the estimators are considered up to the terms of order $n^{-1}$.

**Theorem 1.** *MSE of $T_1^*$ are given by*

$$MSE\,(T_1^*) = S_y^4\, w_1^2 \left[1 + I b_{2y}^* + (I - I')(2a_1^2 b_{2x}^* + 4a_1 I_{22}^* - 2a_1 b_{2x}^*)\right] - 2S_y^4 w_1 \left[1 + (I - I')(a_1 I_{22}^* - a_1 b_{2x}^* + \frac{a_1^2 b_{2x}^*}{2})\right] + S_y^4$$

*Proof.* Consider the estimator,

$$\begin{aligned}
T_1^* &= w_1\, s_y^2 \left[1 + log\left(\frac{s_x'^2}{s_x^2}\right)\right]^{a_1} \\
&= w_1 S_y^2\,(1+\epsilon_0)\left[1 + log\,(1+\epsilon_1')\,(1+\epsilon_1)^{-1}\right]^{a_1} \\
&= w_1\; S_y^2 \left[1 + \epsilon_0 + a_1\left(\epsilon_1' - \epsilon_1 + \epsilon_1^2 - \epsilon_1\epsilon_1'\right) - \frac{a_1\,(\epsilon_1' - \epsilon_1)^2}{2} + \frac{a_1\,(a_1-1)}{2}\,(\epsilon_1' - \epsilon_1)^2 \right. \\
&\quad \left. + a_1\,(\epsilon_0\epsilon_1' - \epsilon_0\epsilon_1)\right] \\
&= (w_1 - 1)\, S_y^2 + w_1 S_y^2 \left[\epsilon_0 + a_1\left(\epsilon_1' - \epsilon_1 + \epsilon_1^2 - \epsilon_1\epsilon_1'\right) - \frac{a_1\,(\epsilon_1' - \epsilon_1)^2}{2} + \frac{a_1\,(a_1-1)}{2}\,(\epsilon_1' - \epsilon_1)^2 \right. \\
&\quad \left. + a_1\,(\epsilon_0\epsilon_1' - \epsilon_0\epsilon_1)\right]
\end{aligned}$$

Squaring on both the sides and then taking expectation over the proposed estimator, we have

$$= S_y^4\,(w_1-1)^2 + w_1^2\, S_y^4\,\left[E(\epsilon_0)^2 + a_1^2\, E(\epsilon_1 - \epsilon_1')^2 + 2\,a_1\, E(\epsilon_0\,\epsilon_1 - \epsilon_0\epsilon_1')\right] + 2S_y^4 w_1\,(w_1-1)\left[a_1 E(\epsilon_0\epsilon_1' - \epsilon_0\epsilon_1) - a_1 E(\epsilon_1 - \epsilon_1')^2 + \frac{a_1^2}{2}\,E(\epsilon_1 - \epsilon_1'))^2\right] \tag{3.1}$$

Using the results from Sukhatme and Sukhatme, we have
$E(\epsilon_0) = 0 = E(\epsilon_1)$, $E(\epsilon_0)^2 = I b_{2y}^*$, $E(\epsilon_1)^2 = I b_{2x}^*$, $E(\epsilon_{,1})^2 = I' b_{2x}^*$, $E(\epsilon_0\epsilon_1) = I I_{22}^*$, $E(\epsilon_0\epsilon_1') = I' I_{22}^*$ where $b_{2x}^* = b_{2x} - 1$, $b_{2y}^* = b_{2y} - 1$ and $I_{22}^* = I_{22} - 1$; $I_{pq} = m_{pq}/m_{20}^{\frac{p}{2}}\, m_{02}^{\frac{q}{2}}$, $m_{pq} = \sum_{i=1}^{N}(Y_i - \bar{Y})^p (Y_i - \bar{Y})^q/N$, $I = 1/N$, $I = 1/N'$ , $b_{2y} = m_{40}/m_{20}^2$, $b_{2x} = m_{04}/m_{02}^2$ are the coefficient of kurtosis of $y$ and $x$ respectively.
Now, substituting the above results in (4.1), we get,

$$MSE\,(T_1^*) = S_y^4\, w_1^2 \left[1 + I b_{2y}^* + (I - I')(2a_1^2 b_{2x}^* + 4a_1 I_{22}^* - 2a_1 b_{2x}^*)\right] - 2S_y^4 w_1 \left[1 + (I - I')(a_1 I_{22}^* - a_1 b_{2x}^* + \frac{a_1^2 b_{2x}^*}{2})\right] + S_y^4$$

□

**Corollary 2.** *The optimum value of $a_1$ and $w_1$ are*

$$a_1 = \frac{I_{22}^*}{b_{2x}}$$

$$w_1 = \frac{B_1}{A_1}$$

*where*

$$A_1 = 1 + Ib_{2y}^* + (I - I')(2a_1^2 b_{2x}^* + 4a_1 I_{22}^* - 2a_1 b_{2x}^*)$$

$$B_1 = 1 + (I - I')(a_1 I_{22}^* - a_1 b_{2x}^* + \frac{a_1^2 b_{2x}^*}{2})$$

*respectively. Also, the minimum mean squared error of* $T_1^*$ *is*

$$MSE\,(T_1^*)_{opt} = S_y^4\, I \left[1 - \frac{B_1^2}{A_1}\right]$$

# 4 Efficiency comparison

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 in case of SRSWOR

$$MSE(t_1) > MSE(T_1)_{opt}$$

## 4.2 The ratio type variance estimator

$$MSE(t_2) > MSE(T_1)_{opt}$$

## 4.3 Naik and Gupta (1996) ratio estimator using auxiliary attribute

$$MSE(t_3) > MSE(T_1)_{opt}$$

## 4.4 The product estimator using auxiliary variables

$$MSE(t_4) > MSE(T_1)_{opt}$$

Thus, it can be easily seen and verified that the proposed class of log-type estimators is far better than these above mentioned estimators, available in sampling literature.

# 5 Empirical study

The data on which we performed the numerical calculation is taken from some natural populations. The summary and the percent relative efficiency of the following estimators are as follows:

Table 1: *Parameters of the Data*

| Parameters | Population 1 | Population 2 | Population 3 | Population 4 |
|---|---|---|---|---|
| $n'$ | 33 | 36 | 22 | 34 |
| $n$ | 11 | 11 | 8 | 10 |
| $b^*_{2y}$ | 4.032 | 2.632 | 9.433 | 2.725 |
| $b^*_{2x}$ | 1.388 | 2.402 | 7.105 | 12.366 |
| $b^*_{2f}$ | 1.143 | 2.345 | 7.766 | 1.912 |
| $I^*_{22x}$ | 0.305 | 1.835 | 8.140 | 0.224 |
| $I^*_{22f}$ | 1.155 | 2.014 | 8.538 | 2.104 |
| $I^*_{22xf}$ | 0.492 | 2.182 | 7.423 | 0.152 |

Table 2: *PRE of the estimators*

| Estimator | Population 1 | Population 2 | Population 3 | Population 1 |
|---|---|---|---|---|
| $t_1$ | 100 | 100 | 100 | 100 |
| $t_2$ | 141.940 | 277.598 | 574168.1 | 637.142 |
| $t_3$ | 38.525 | 13.593 | 13.009 | 12.407 |
| $t_4$ | 116.860 | 248.212 | 5107.79 | 67.4228 |
| $T_{iopt}$ | 142.235 | 293.078 | 605514.2 | 666.034 |

# 6 Conclusion

The present study extends the idea of Kumari et al. (2019) regarding the effective use of auxiliary information if the relationship between the study variable and the auxiliary variable is of logarithmic type. Further, the efficiency of the proposed estimators are compared with some conventional estimators and some recent estimators. The proposed estimator is most efficient than all the estimators. This study is also supported through an empirical sudy and the result of this study is quite encouraging.

# References

[1] Bhushan, S. and Kumari, C. (2018). A new log type estimators for estimating the population variance, International Journal of Computational and Applied Mathematics (ISSN 1819-4966), 13 (1), 43-54.

[2] Bhushan, S. and Kumari, C. (2018). 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.

[3] Bhushan, S. and Kumari, C. (2018). 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.

[4] Bhushan, S. and Kumari, C. (2018). Estimation of Variance of Finite Population Using Double Sampling Scheme, International Journal of Scientific and Engineering Research, Volume 9, Issue 8, pg 1893-1901.

[5] Kumari, C., Bhushan, S. and Thakur, R. K. (2018). Modied Ratio 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 1884-1892.

[6] 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.

[7] Bhushan, S. and Kumari, C (2019). Double Sampling Log-type Estimator Using Auxiliary Attribute for Population Variance, Journal of Statistics Applications & Probability, 6, No. 3, pg 1-6

[8] Kadilar C. and Cingi H. (2006a). Improvement in variance estimation using auxiliary information, *Hacet. J. Math. Stat.*, 1(35), 111-115.

[9] Kadilar C. and Cingi H. (2006b). Ratio estimators for population variance in simple and stratied sampling, *App. Math. Comp.*, 1(73), 1047-1058.

[10] Kadilar C. and Cingi H. (2014). A two parameter variance estimator using auxiliary information, *App. Math. Comp.*, 117-122.

[11] Hidiroglou M. A. and Sarndal C. E. (1998). Use of auxiliary information for two-phase sampling , *Surv. Methodol.*, 24(1), 11-20.

[12] Swain A. K. P. C. and Mishra G. (1994). Estimation of population variance under unequal probability sampling , *Sankhya*, B (56), 374-384.

[13] Bahl S. and Tuteja R. K. (1991). Ratio and Product type exponential estimator, *Info. Optim. Sci.*, XII(I), 159-163.

[14] Sukhatme P. V., Sukhatme B. V., Sukhatme S. and Ashok C. (1984). Sampling Theory of Surveys with Applications , *Iowa State Univ. Press*, Ames.

[15] Isaki C. T. (1983). Variance estimation using Auxiliary Information , *J. Amer. Stat. Asso.*, 78, 117-123.

[16] Chaudhury A. (1978). On estimating the variance of a finite population. *Metrika*, 25, 66-67.

[17] Das A. K. and Tripathi T. P. (1978). Use of auxiliary information in estimating the nite population variance. *Sankhya*, C(4), 139-148.

[18] Cochran W. G. (1963). Sampling Techniques , *Wiley Eastern Private Limited*, New Delhi.

[19] Neyman J. (1938). Contribution to the theory of sampling human populations, *J. Amer. Stat. Asso.*, 33, 101-116.

[20] Cebrian A. A. and Garcia R. M. (1997). Variance estimation using auxiliary information : an almost multivariate ratio estimator, *Metrika*, 45, 171-178.

[21] Garcia R. M. and Cebrian A. A. (1996). Repeat substitution method: the ratio estimator for population variance, *Metrika*, 43, 101-105.

[22] Upadhaya L. N. and Singh H. P. (2001). Estimation of the population standard deviation using auxiliary information, *Ame. J. Math. Manag. Sci.*, 21, 345-385.

[23] Searls D. T. (1964). Utilization of known coefficient of kurtosis in the estimation procedure of variance, *J. Amer. Stat. Assoc.*, 59, 1125-1226.

[24] Prasad B. and Singh H. P. (1992).Unbiased estimators of finite population variance using auxiliary information sample surveys, *Commun. Stat. : Theory and Methods*, 21(5), 1367-1376.