Vol.6.Issue.2.2018 (April-June)

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

http://www.bomsr.com

Email:editorbomsr@gmail.com

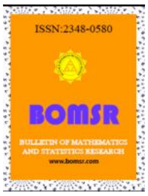
INTERNATIONAL STANDARD SERIAL NUMBER 2348-0580

RESEARCH ARTICLE

GENERALIZED RATIO TYPE EXPONENTIAL ESTIMATOR OF FINITE POPULATION VARIANCE IN SAMPLE SURVEYS

SITANSHU SEKHAR MISHRA

Lecturer in Statistics, Department of Statistics, Salipur (Autonomous) College, Salipur, Cuttack
Odisha, India
E-mail: sitanshumishra0@gmail.com



ABSTRACT

In this paper a generalized class of ratio-type exponential estimators to estimate the finite population variance is proposed and is compared with a generalized class of power transformation ratio type estimator suggested by Swain (2014) as regards their biases and mean square errors both theoreticals and with numerical illustrations.

Keywords : Auxiliary variable, Ratio estimator, Power transformation and exponential ratio type estimator, Finite population variance, Bias, mean square error

AMS Classification : 62 D 05

1. Introduction

The use of auxiliary information in sample surveys is often resorted to find more efficient estimators of the population parametric function such as population mean, population variance etc compared to means without use of auxiliary information.

As the problem of estimation of finite population variance of the study variable is of importance in sample surveys, many attempts have been made in sampling theory literature to improve the precision of the estimates of the finite population variance.

Ratio and regression methods of estimation are two classical methods which make use of auxiliary information to derive improved estimators of the population mean or total of the main variable under study (Cochran, 1953). Although the estimation of finite population variance arises while estimating the variance arises while estimating the variance of the estimates of finite population mean or total, it was only through the work of Evans (1951) and Liu (1974) separate attention was given to the estimation of finite population variance. Liu (1974) considered the problem of estimating variance in a general set up and presented a general class of quadratic functions and obtained a class of unbiased estimators under some conditions. Isaki(1983) proposed the ratio method of estimation to estimate the finite population variance. Some early work on the estimation of finite population variance are due to Das and Tripathy (1978) and Srivastava and Jhajj (1980) among others.

Singh and Solanki (2009-2010), Tailor and Sharma (2012), Solanki and Singh (2013), Singh and Solanki (2013a,b), Subramani and Kumara Pandiyan (2013a,b) and Yadav and Kadilar (2013a,b) have paid their attention towards the improved estimator of population variance of the study variable *y* using information on the known parameters of the auxiliary variable *x* such as mean, variance, coefficient of skewness, coefficient of kurtosis, correlation coefficient between the study variable *y* and the auxiliary variable *x* etc.

Swain (2015) studied the estimation of variance by using information on both the mean $\bar{X}$ and variance S_x^2 for an auxiliary variable *x*.

In this paper a generalized ratio type exponential estimator is suggested for the estimation of finite population variance. Some special cases of this generalized estimator are compared with those of generalized class of estimators proposed by Swain (2015) as regards bias and mean square error.

Define $\bar{Y}$ and $\bar{X}$ as the mean of *y* and *x* respectively.

$$\bar{Y}=\frac{1}{N}\sum_{i=1}^{N} y_i\ ,\quad \bar{X}=\frac{1}{N}\sum_{i=1}^{N} x_i\ .$$

The finite population variance of *y* and *x* are defined by

$$S_y^2=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{Y})^2\ ,\qquad S_x^2=\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{X})^2\ .$$

and finite population covariance as

$$S_{yx}=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{Y})(x_i-\bar{X}).$$

The correlation coefficient *y* and *x* is defined as $\rho=\frac{S_{yx}}{S_y S_x}$.

For a simple random sample without replacement of size *n*, define the sample means $\bar{y}$ and $\bar{x}$ of *y* and *x* respectively as

$$\bar{y}=\frac{1}{n}\sum_{i=1}^{n} y_i\ ,\ \bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i\ .$$

Define $\theta=\frac{1}{n}-\frac{1}{N}=\frac{1-f}{n}\ ,\ f=\frac{n}{N}$.

Further, define $s_y^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2\ ,\ s_x^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\ ,$

$$s_{yx}=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})\ (y_i-\bar{y})^2$$

as the sample variances of study and auxiliary variables and covariance respectively.

A ratio estimator of the finite population variance suggested by Isaki (1983) is given by

$$t_R=\frac{s_y^2}{s_x^2}\cdot S_x^2$$

To $O(\frac{1}{n})$,

$$MSE(t_R)=\theta S_y^4\left[(\lambda_{40}-1)+(\lambda_{04}-1)-2(\lambda_{22}-1)\right]$$

$$B(t_R)=\theta S_y^2\left[(\lambda_{04}-1)-(\lambda_{22}-1)\right]$$

Swain (2014) consider a generalized class of estimators of finite population variance.

$$t_g=s_y^2\left[\alpha\left(\frac{S_x^2}{s_x^2}\right)^g+(1-\alpha)\left(\frac{s_x^2}{S_x^2}\right)^h\right]^\delta$$

where α, g, h and δ are real and free parameters to be chosen suitably. The mean square error of t_g to $O(\frac{1}{n})$ is given by

$$MSE(t_g)$$

$$=\theta S_y^4\left[(\lambda_{40}-1)+\delta^2(h-\alpha(g+h))^2(\lambda_{04}-1)+2\delta(h-\alpha(g+h))(\lambda_{22}-1)\right]$$

Minimizing with respect to α, we have

$$\alpha_{opt}=\frac{h}{h+g}+\frac{1}{\delta(h+g)}\left(\frac{\lambda_{22}-1}{\lambda_{04}-1}\right)$$

$$K=\frac{\lambda_{22}-1}{\lambda_{04}-1}$$

i.e. $$\alpha_{opt}=\frac{\delta h+K}{\delta(g+h)}$$

Substituting the optimum value of α in $MSE(t_g)$, the optimum mean square error of t_g is given by

$$MSE(t_g)_{opt}=\theta S_y^4\left[(\lambda_{40}-1)-\frac{(\lambda_{22}-1)^2}{\lambda_{04}-1}\right]$$

To $O(\frac{1}{n})$,

$$B(t_g)=\theta S_y^2\left[\delta\{h-\alpha(g+h)\}(\lambda_{22}-1)+\delta\left\{\frac{h(h-1)}{2}+\alpha\left(\frac{g(g+1)}{2}-\frac{h(h-1)}{2}\right)\right\}(\lambda_{04}-1)\right.$$

$$\left.+\frac{\delta(\delta-1)}{2}\{h-\alpha(g+h)\}^2(\lambda_{04}-1)\right]$$

Substituting the optimum value of α in the approximate expression of bias to $O\left(\frac{1}{n}\right)$, we have

$$B(t_g)_{opt}$$

$$=\theta S_y^2\left[\delta\{h-\alpha_{opt}(g+h)\}(\lambda_{22}-1)+\delta\left\{\frac{h(h-1)}{2}+\alpha_{opt}\left(\frac{g(g+1)}{2}-\frac{h(h-1)}{2}\right)\right\}(\lambda_{04}-1)\right.$$

$$\left.+\frac{\delta(\delta-1)}{2}\{h-\alpha_{opt}(g+h)\}^2(\lambda_{04}-1)\right]$$

2. Proposed Estimator :

Following Swain (2013), we propose a generalized ratio type exponential estimator for the finite population variance as

$$t_g^* = s_y^2 \left[\alpha \cdot e^{g\left(\frac{S_x^2 - s_x^2}{S_x^2}\right)} + (1-\alpha) \cdot e^{h\left(\frac{s_x^2 - S_x^2}{S_x^2}\right)} \right]^{\delta} \qquad (2.1)$$

where α, g, h and δ are free real constants to be suitably chosen.

In this regard we may fix, *g*, *h* and *δ* and minimize the approximate mean square error of t_g^* to $O(\frac{1}{n})$ with respect to *α*. Substituting the optimum value of *α* in the expression for mean square error of t_g^* we obtain the optimum mean square error of t_g^* and also the optimum approximate bias of t_g^*.

Let $e_0 = \frac{s_y^2 - S_y^2}{S_y^2}$ and $e_1 = \frac{s_x^2 - S_x^2}{S_x^2}$.

Thus, we write

$$s_y^2 = S_y^2(1+e_0) \ , \ s_x^2 = S_x^2(1+e_1),$$

with $E(e_0) = E(e_1) = 0$,

$$E(e_0^2) = V(e_0) = \frac{1-f}{n}(\lambda_{40} - 1) = \theta(\lambda_{40} - 1)$$

$$E(e_1^2) = V(e_1) = \frac{1-f}{n}(\lambda_{04} - 1) = \theta(\lambda_{04} - 1)$$

$$E(e_0 e_1) = \text{cov}(e_0 e_1) = \frac{1-f}{n}(\lambda_{22} - 1) = \theta(\lambda_{22} - 1).$$

$$\lambda_{22} = \frac{\mu_{22}}{\mu_{20}\mu_{02}}$$

$$\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^r (X_i - \bar{X})^s.$$

Expanding t_g^* in power series and keeping upto second degree terms we have

$$t_g^* = S_y^2(1+e_0)\left[\alpha\left(1 - ge_1 + g^2\frac{e_1^2}{2} + \cdots\right) + (1-\alpha)(1 + he_1 + h^2\frac{e_1^2}{2} +) \right]^{\delta}$$

$$= S_y^2\Big[1 + \delta\{h - \alpha(g+h)\}e_1 + \delta\left\{\frac{h^2}{2} + \alpha\left(\frac{g^2 - h^2}{2}\right)\right\}e_1^2$$

$$+ \frac{\delta(\delta-1)}{2}\{h - \alpha(g+h)\}^2 e_1^2 + e_0 + \delta\{h - \alpha(g+h)\}e_0 e_1\Big]$$

Thus the bias of t_g^* to $O(\frac{1}{n})$ is given by

$$B(t_g^*)=\theta S_y^2\left[\delta\{h-\alpha(g+h)\}(\lambda_{22}-1)+\delta\left\{\frac{h^2}{2}+\alpha\left(\frac{g^2-h^2}{2}\right)\right\}(\lambda_{04}-1)\right.$$
$$\left.+\frac{\delta(\delta-1)}{2}\{h-\alpha(g+h)\}^2(\lambda_{04}-1)\right] \quad (2.2)$$

The mean square error (MSE) of t_g^* to $O(\frac{1}{n})$ is given by

$$MSE(t_g^*)=\theta S_y^4\left[(\lambda_{40}-1)+\delta^2\{h-\alpha(g+h)\}^2(\lambda_{04}-1)\right.$$
$$\left.+2\delta\{h-\alpha(g+h)\}(\lambda_{22}-1)\right] \quad (2.3)$$

Minimizing with respect to α, we have

$$\alpha_{opt}=\frac{h}{h+g}+\frac{1}{\delta(h+g)}\left(\frac{\lambda_{22}-1}{\lambda_{04}-1}\right)$$

i.e. $$\alpha_{opt}=\frac{K+\delta h}{\delta(g+h)}$$

Substituting the optimum value of α in $MSE(t_g^*)$, the optimum mean square error of t_g^* is given by

$$MSE(t_g^*)_{opt}=\theta S_y^4\left[(\lambda_{40}-1)-\frac{(\lambda_{22}-1)}{\lambda_{04}-1}\right]$$
$$=MSE(t_g)_{opt}. \quad (2.4)$$

Again substituting the optimum value of α in the approximate expression of bias to, $O(\frac{1}{n})$ we have

$$B(t_g^*)_{opt}=\theta S_y^2\left[\delta\{h-\alpha_{opt}(g+h)\}(\lambda_{22}-1)+\delta\left\{\frac{h^2}{2}+\alpha_{opt}\left(\frac{g^2-h^2}{2}\right)\right\}(\lambda_{04}-1)\right.$$
$$+\frac{\delta(\delta-1)}{2}\{h-\alpha_{opt}(g+h)\}^2(\lambda_{04}-1) \quad (2.5)$$

3. Comparison of Biases and Mean square errors of t_g and t_g^*

$$B(t_g)_{opt}=\theta S_y^2\left[\delta\{h-\alpha_{opt}(g+h)\}(\lambda_{22}-1)\right.$$
$$+\delta\left\{\frac{h(h-1)}{2}+\alpha_{opt}\left(\frac{g(g+1)}{2}-\frac{h(h-1)}{2}\right)\right\}(\lambda_{04}-1)$$
$$+\frac{\delta(\delta-1)}{2}\{h-\alpha_{opt}(g+h)\}^2(\lambda_{04}-1)$$

$$B(t_g^*)_{opt}=\theta S_y^2\left[\delta\{h-\alpha_{opt}(g+h)\}(\lambda_{22}-1)\right.$$
$$+\delta\left\{\frac{h^2}{2}+\alpha_{opt}\left(\frac{g^2-h^2}{2}\right)\right\}(\lambda_{04}-1)$$
$$+\frac{\delta(\delta-1)}{2}\{h-\alpha_{opt}(g+h)\}^2(\lambda_{04}-1)$$

$$MSE(t_g)_{opt}=MSE(t_g^*)_{opt}=\theta S_y^4\left[(\lambda_{40}-1)-\frac{(\lambda_{22}-1)^2}{\lambda_{04}-1}\right]$$

Thus, to $O\left(\frac{1}{n}\right), t_g$ and t_g^* are equally efficient.

The biases of t_g and t_g^* are functions of *g*, *h* and *δ* and as such vary with different values of *g*, *h* and *δ*. The best choices of these free parameters are those for which the first order biases vanish. For some specified values for *g, h* and*δ* the biases of t_g and t_g^* are tabulated in Table 1.

Table 1: Comparison of Biases without the common multiplier θS_y^2

Sl. No.	*g*	*h*	*δ*	α_{opt}	$B(t_g)$	$B(t_g^*)$
1	1	1	1	$\frac{K+1}{2}$	$-K(\lambda_{22}-1)+\frac{K+1}{2}(\lambda_{04}-1)$	$-K(\lambda_{22}-1)+\frac{1}{2}(\lambda_{04}-1)$
2	1	0	1	*K*	$-K(\lambda_{22}-1)+K(\lambda_{04}-1)$	$-K(\lambda_{22}-1)+\frac{K}{2}(\lambda_{04}-1)$
3	0	1	1	1+*K*	$-K(\lambda_{22}-1)$	$-K(\lambda_{22}-1)-\frac{K}{2}(\lambda_{04}-1)$
4	1	1	-1	$\frac{1-K}{2}$	$-K(\lambda_{22}-1)+\left(\frac{K-1}{2}+K^2\right)(\lambda_{04}-1)$	$-K(\lambda_{22}-1)+\left(K^2-\frac{1}{2}\right)(\lambda_{04}-1)$
5	1	0	-1	-*K*	$-K(\lambda_{22}-1)+K(K+1)(\lambda_{04}-1)$	$-K(\lambda_{22}-1)+\left(\frac{K}{2}+K^2\right)(\lambda_{04}-1)$
6	0	1	-1	1-*K*	$-K(\lambda_{22}-1)+K^2(\lambda_{04}-1)$	$-K(\lambda_{22}-1)+\left(K^2-\frac{K}{2}\right)(\lambda_{04}-1)$

4. Numerical Illustrations :

Consider 5 natural populations described in Table-2 to compare the biases power transformation ratio type estimator with exponential ratio type estimator as regards their biases.

Table 2: Description of Populations

Popn. No.	Description	*N*	*n*	*y*	*x*	λ_{40}	λ_{04}	λ_{22}	ρ_{yx}	K
1	Cochran, 1977	196	49	Number of inhabitants in 1930	Number of inhabitants in 1920	8.5362	7.3617	7.8780	0.9820	1.0811
2	Sukhatme and Sukhatme, 1970	170	10	Wheat acreage in 1937	Wheat acreage in 1936	3.1842	2.2030	2.5597	0.9770	1.2965
3	Singh et al., 1988	278	30	The number of agriculture labourers for 1971	The number of agriculture labourers for 1961	24.8969	37.8898	25.8142	0.7273	0.6726
4	Singh,	50	08	Amount (in	Amount in	3.5822	4.5247	2.8411	0.8038	0.5223

Popn. No.	Description	N	n	y	x	λ_{40}	λ_{04}	λ_{22}	ρ_{yx}	K
	2003			$ 1000) of real estate farm loans in different states during, 1997	$ 1000 of non realestates farm loans in different states during 1997					
5	Kadilar and Cingi, 2006	106	20	The level of apple production (1 unit = 100 tonns)	The number of apple trees (1 unit = 100 trees)	80.13	25.71	33.30	0.8200	1.3071

Table 3 : Absolute biases of estimators without constant multiplier θS_y^2

Bias \ Population	1	2	3	4	5
$B(t_{g1})$	0.8162	0.6409	14.1609	1.721	13.7164
$B(t_{g2}^*)$	4.255	1.4206	1.7549	0.8007	29.8640
$B(t_{g2})$	0.5582	0.4625	8.122	0.8793	9.9209
$B(t_{g2}^*)$	3.9974	1.2424	+4.284	0.0414	26.0714
$B(t_{g3})$	7.4358	2.0221	16.6900	0.9616	42.2193
$B(t_{g3}^*)$	10.8742	2.8018	29.0960	1.8818	16.1479
$B(t_{g4})$	0.2567	0.1783	6.0437	0.8422	3.7906
$B(t_{g4}^*)$	3.1819	0.6015	18.4496	1.7627	12.3573
$B(t_{g5})$	6.8767	1.5596	24.8072	1.8408	32.2961
$B(t_{g5}^*)$	3.4376	0.7798	12.4012	0.9198	16.1457
$B(t_{g6})$	0.0009	0.0000	0.0048	0.0005	0.0023
$B(t_{g6}^*)$	3.4394	0.7798	12.4108	0.9208	16.1503
$B(\bar{y}_R)$	0.5163	0.3567	12.0756	1.6836	7.590

Comments :

(i) t_{g_1} is less biased than $t_{g_1}^*$ for population 1,2,5.

(ii) t_{g_2} is less biased than $t_{g_2}^*$ for population 1,2,5.

(iii) t_{g_3} is less biased than $t_{g_3}^*$ for population 1,2,3,4.

(iv) t_{g_4} is less biased than $t^*_{g_4}$ for population 1,2,3,4,5.

(v) $t^*_{g_5}$ is less biased than t_{g_5} for population 1,2,3,4,5.

(vi) t_{g_6} is less biased than $t^*_{g_6}$ for population 1,2,3,4,5.

5. Conclusion

t_g and t^*_g are equally efficient in large samples. But biases of these classes vary for different values of *g*, *h* and *δ*. The choices of *g*, *h* and *δ* are arbitrary and should be so chosen to make first order bias zero. For the numerical illustrations under considerations t_g happens to be less biased than t^*_g for some specified values of *g*, *h* and *δ*. The special cases of the generalized classes of ratio-type estimators discussed in this paper can be distinguished and compared on the basis of biases of these estimators depending on the values of the free parameters involved in their construction.

Acknowledgement

The author is grateful to Prof. A.K.P.C. Swain, retired Professor of Statistics, Utkal University for his valuable suggestions and criticisms, which improved the presentation of results incorporated in this paper.

References

1. Bahl, S. and Tuteja, R.K. (1991). Ratio and product type exponential estimator. Information and Optimization Science, 12, 159-163.
2. Cochran, W.G. (1953). Sampling Techniques, New York, John Wiley and Sons.
3. Das, A.K. and Tripathy, T.P. (1978). Use of auxiliary information in estimating the finite population variance, Sankhya, C, 40, 139-148.
4. Evans, D. (1951). On the variance estimators of SD and Variance, Journ. Amer. Stat. Assoc. 46, 220-224.
5. Isaki, C.T. (1983). Variance estimation using auxiliary information.Journ. Amer. Stat. Assoc., 78, 117-123.
6. Liu, T.P. (1974). A general unbiased estimator for the variance of a finite population, Sankhya, C, 36, 23-32.
7. Singh, H.P., Upadhyaya, L.N. and Namjoshi, U.D. (1988). Estimation of finite population variance.Curr. Sc., 57, 1331-1334.
8. Singh, H .P. and Solanki, R.S. (2013a). A new procedure for variance estimation in simple random sampling using auxiliary information statistical papers, 54(2), 479-497.
9. Singh, H.P. and Solanki, R.S. (2009-2010). Estimation of finite population variance using auxiliary information in presence of random non-response. Gujarat Statistical Review, 36 & 37 (1 & 2), 46-58.
10. Singh, H.P. and Solanki, R.S. (2013b). Improved estimation of finite population variance using auxiliary information. Communications in Statistics Theory and Methods, 42 (15), 2718-2730.

11. Solanki, R.S. and Singh, H.P. (2013). An improved class of estimators for the population variance. Model Assisted Statistics and Applications 8(3), 229-238.
12. Srivastava, S.K. and Jhajj, H.S. (1980). A class of estimators using auxiliary information for estimating finite population variance. Sankhya, C, 42, 87-96.
13. Subramani, J. and Kumarapandiyan, G. (2013a). Estimation of variance using known coeffient of variation and median of an auxiliary variables. Journal of Modern Applied Statistical Methods, 12(1), 58-64.
14. Subramani, J. and Kumarapandiyan, G. (2013b). Estimation of variance using decilesof an auxiliary variables. Proceedings of International conference on Frontiers of statistics and its Applications, Bonfring Publishers, 143-149.
15. Sukhatme, P.V., Sukhatme, B.V. (1970). Sampling Theory of Surveys with Application, Asia Publishing House, New Delhi.
16. Swain, A.K.P.C. (2013): An alternative ratio type exponential estimator. Unpublished communication.
17. Swain, A.K.P.C. (2014): On an improved ratio type estimator of finite population mean sample surveys. Revista Investigation Operational, 35(1), 49-57.
18. Swain, A.K.P.C. and Mishra, S.S. (2016): On comparison of two generalized classes of Ratio-type estimators in sample surveys. Revista Investigation Operational, 37(3), 272-280.
19. Swain, A.K.P.C. (2015). Generalized Estimator of finite population variance. Journal of Statistical Theory and Applications, Vol.14(1), 000-000.
20. Tailor, R. and Sharma, B. (2012). Modified estimators of population variance in presence of auxiliary information, Statistics in Transition, Vol.13, No.1, 37-46.
21. Yadav, S.K. and Kadilar, C. (2013a). A class of ratio-cum-dual to ratio estimator of population variance. Journal of Reliability and Statistical Studies, 6(1), 29-34.
22. Yadav, S.K. and Kadilar, C. (2013b). Improved Exponential type ratio estimator of population variance.Colomb