GENERALISED EXPONENTIAL RATIO-CUM-PRODUCT ESTIMATOR FOR ESTIMATING POPULATION VARIANCE IN SIMPLE RANDOM SAMPLING

Rafia Jan[1], T. R. Jan[2] and Faizan Danish[3*]

•

[1]Department of Statistics, Government Degree College Bejbehara Anantnag, J&K, India
[2] Department of Statistics, University of Kashmir, J&K, India
[3]Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu, Beside AP Secretariat, Amravati, Andhra Pradesh-522237, India.
Email: 1:rafiajan836@gmail.com
2:drtrjan@gmail.com
3.danishstat@gmail.com

Abstract

This study presents a comprehensive investigation into the estimation of population variance for a study variable using Simple Random Sampling, with the incorporation of two auxiliary variables. To address the challenges of variance estimation in complex scenarios, a novel approach termed the "Proposed Generalized Exponential Ratio-cum-Product Estimator" is introduced. This innovative estimator belongs to a class of estimators that rely on exponential functions of the auxiliary variables, providing enhanced precision and efficiency in variance estimation. To thoroughly assess the performance of the proposed estimators, the research develops equations for both Mean Square Errors and Biases, unveiling their statistical properties. The study systematically explores the conditions under which these estimators demonstrate superior efficiency compared to traditional alternative estimators, thereby enabling researchers to identify contexts where their utilization is most beneficial. The empirical aspect of the research constitutes a significant contribution to the study's validity. Through empirical analysis, the proposed estimators are directly compared against the conventional Unbiased Sample Variance Estimator, showcasing their clear superiority in terms of efficiency. Furthermore, Mean Square Errors and Percent Relative Efficiency are calculated for all estimators and subjected to theoretical and empirical comparisons with existing estimation methods. These findings corroborate the advantageous attributes of the proposed estimator in real-world scenarios, reinforcing its practicality and reliability in various research domains. Beyond methodological developments, this study also delves into the real-world implications and applications of the proposed estimators. It highlights the potential benefits of utilizing these estimators in situations where study variables exhibit intricate relationships with auxiliary variables, offering valuable insights into multifaceted data sets and multidimensional factors. Additionally, a comprehensive sensitivity analysis is undertaken to assess the robustness of the proposed estimators under varying assumptions and sampling schemes. The researchers' meticulous evaluation enhances the credibility of the proposed estimators and ensures their adaptability across diverse practical scenarios. Overall, this study's significance extends beyond statistical theory, presenting valuable practical implications for researchers and practitioners across different fields. Improved population variance estimation leads to enhanced decision-making, optimal resource allocation, and deeper insights into underlying phenomena. By introducing the proposed estimator and thoroughly examining its performance through rigorous theoretical and empirical analyses, this research lays a solid foundation for more robust and efficient variance estimation techniques. The insights gained from this study can reshape statistical practices, paving the way for advancements in diverse scientific disciplines and inspiring further knowledge exploration.

Keywords: Exponential estimator, Auxiliary variable, Mean Square Error and Percent Relative Efficiency

I. Introduction

In recent years, the use of sample surveys has gained popularity due to the practicality of overcoming logistical challenges associated with conducting comprehensive census surveys. This trend has led to the widespread adoption of estimators like the ratio, product, and regression estimators for efficiently estimating population parameters, particularly the mean of the variable of interest. These estimators capitalize on the inherent correlation between the study variable and auxiliary variables, either during the survey design or at the estimation stage, to yield accurate results while optimizing resources. The central focus of this research is to develop a novel modified exponential ratio estimator for the population mean. This estimator aims to address potential limitations of existing estimators and enhance the precision of estimates, as evaluated through mean squared error comparisons. By exploring alternative approaches and incorporating adjustments, the researchers anticipate achieving more reliable and efficient estimates of the population mean.

Over the years, several scholars have made significant contributions to the field of survey estimation. Various authors have made numerous work for the estimation of population variance from time to time including [14],[9] , [13], [8] [1], [5], [11],[12],[15] and [10] have made important studies on this topic in the literature. Notably, [17] made pioneering strides by explicitly utilizing auxiliary information for estimation purposes, laying the foundation for the ratio estimator. Subsequently, [18] further advanced this concept by employing auxiliary information to refine estimations.

When dealing with scenarios where the coefficient of correlation is negative between the study variable and auxiliary variables, [19] introduced the product-type estimator, which has proven to be valuable in specific contexts. Additionally, [20] proposed an innovative approach by combining multiple ratio estimators based on individual auxiliary variables positively correlated with the study variable. This technique allowed for greater accuracy in estimation. The product estimator was formalized by [21], providing a well-defined framework for its application. Furthermore, [22] delved into the complexities of ratio estimators involving two or more correlated variables, shedding light on new possibilities for refining estimation methods. The exponential type estimators of population mean were thoroughly investigated by [23] using auxiliary data, resulting in a comprehensive analysis of their performance and potential improvements. [24] took a unique approach by incorporating transformed auxiliary variables, which led to promising results in estimating the mean of the study character. The literature offers an array of other contributions in this area, including the works of [25], [26], [27], and [28], who introduced their respective estimators and demonstrated their efficacy in diverse sampling scenarios. Moreover, [29] and [30] took on the challenge of developing superior exponential type estimators by considering information from two altered auxiliary variables, further expanding the range of available estimation techniques. To gain a more comprehensive understanding of this topic, interested readers can refer to [31], which offers an in-depth exploration of various aspects of survey estimation. In recent times, [32], [33], and [34] have made notable contributions to this area of study, introducing novel ideas and methodologies that hold promise for advancing the field of survey estimation even further.

In conclusion, this research endeavors to create a Generalized Ratio-cum-product estimator of population variance that builds upon the knowledge and advancements made by previous scholars. By harnessing the power of auxiliary information and exploring innovative avenues, the researchers aim to provide an enhanced and efficient approach to estimating the population mean and contributing to the growing body of knowledge in survey estimation techniques.

.

Consider a population of size N. Let y and x be the variable of interest and the Auxiliary variables respectively.

The usual unbiased variance estimator is,

$$t_0 = s_y^2, \tag{1}$$

With variance

$$V(t_0) = \gamma S_y^4 [\lambda_{40} - 1], \tag{2}$$

where, $S_y^{\ 2} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$ and $S_x^{\ 2} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$ are the population variances of y and x respectively.

[6] proposed the ratio type estimator of S_y^2

$$t_1 = s_y^{\ 2}\frac{S_x^{\ 2}}{s_x^{\ 2}}, \tag{3}$$

and the MSE is

$$MSE(t_1) = \gamma S_y^{\ 4}[(\lambda_{40} - 1) + (\lambda_{04} - 1) - 2(\lambda_{22} - 1)], \tag{4}$$

where,

$\lambda_{22} = \frac{\mu_{11}}{\sqrt{\mu_{02}\,\mu_{20}}}$ = covariance between $S_y^{\ 2}$ and $S_x^{\ 2}$.

$\lambda_{40} = \beta_2(y) = \frac{\mu_{40}}{\mu_{20}^{\ 2}}$ = kurtosis for population of Y.

$\lambda_{04} = \beta_2(x) = \frac{\mu_{04}}{\mu_{02}^{\ 2}}$ = kurtosis for population of X.

$\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^r (x_i - \bar{X})^s$ and $\gamma = \frac{1}{n}$.

The exponential ratio type and product type estimators for S_y^2

$$t_2 = s_y^2 \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right), \tag{5}$$

$$t_3 = s_y^2 \exp\left(\frac{s_x^2 - S_x^2}{s_x^2 + S_x^2}\right), \tag{6}$$

And the MSE is

$$MSE(t_2) = \gamma S_y^{\ 4}\left[(\lambda_{40} - 1) + \frac{(\lambda_{04} - 1)}{4} - (\lambda_{22} - 1)\right], \tag{7}$$

$$MSE(t_3) = \gamma S_y^{\ 4}\left[(\lambda_{40} - 1) + \frac{(\lambda_{04} - 1)}{4} + (\lambda_{22} - 1)\right]. \tag{8}$$

II. Proposed Estimator

We propose the generalized exponential ratio-cum product estimator following [16],

$$t^* = s_y^2 \exp\left[\frac{\left(\frac{S_x^2}{s_x^2}\right)^{\alpha} - 1}{\left(\frac{S_x^2}{s_x^2}\right)^{\alpha} + 1}\right], \tag{9}$$

where α is a suitable constant to minimize the $MSE(t^*)$.

Special cases:

i. If $\alpha = 1$, we get exponential ratio type estimator.

ii. If $\alpha = -1$, we get exponential product type estimator.

To get the bias and MSE

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

such that

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

and

$E(e_0{}^2) = \gamma(\beta_2(y)-1) = \gamma(\lambda_{40}-1)$,

$E(e_1{}^2) = \gamma(\beta_2(x)-1) = \gamma(\lambda_{04}-1)$,

$E(e_0 e_1) = \gamma(\lambda_{22}-1)$.

Expressing (9) in e's we have

$$t^* = S_y^2(1+e_0)\exp\left[\frac{\left(\frac{S_x^2}{S_x^2(1+e_1)}\right)^{\alpha}-1}{\left(\frac{S_x^2}{S_x^2(1+e_1)}\right)^{\alpha}+1}\right]$$

After simplification we have,

$$t^* = S_y^2(1+e_0)\exp\left[-\frac{\alpha e_1}{2}\left(1-\frac{\alpha e_1}{2}\right)^{-1}\right]$$

After solving we get the MSE given as

$$MSE(t_2) = \gamma S_y{}^4\left[(\lambda_{40}-1)+\frac{\alpha^2(\lambda_{04}-1)}{4}-\alpha(\lambda_{22}-1)\right] \tag{10}$$

To get minimum $MSE(t^*)$ differentiating above equation with respect to α,

$$\alpha_{opt} = 2\frac{(\lambda_{22}-1)}{(\lambda_{04}-1)}. \tag{11}$$

Using the value of α_{opt}, we get the $MSE_{\min}(t^*)$ given as

$$MSE_{\min}(t^*) = \gamma S_y^4\left[(\lambda_{04}-1)-\frac{(\lambda_{22}-1)^2}{(\lambda_{04}-1)}\right] \tag{12}$$

Which coincides with the MSE of the regression estimator.

III. Efficiency Comparison

$$MSE(t_0) - MSE_{\min}(t^*) = \frac{S_y^4}{n}\left[\frac{(\lambda_{22}-1)^2}{(\lambda_{04}-1)}\right] \geq 0$$

$$MSE(t_1) - MSE_{\min}(t^*) = \frac{S_y^4}{n}\left[\frac{(\lambda_{04}-\lambda_{22})^2}{(\lambda_{04}-1)}\right] \geq 0$$

$$MSE(t_2) - MSE_{\min}(t^*) = \frac{S_y^4}{n}\left[\frac{(\lambda_{04}-2\lambda_{22}+1)^2}{(\lambda_{04}-1)}\right] \geq 0$$

All the above conditions are always satisfied.

IV. Numerical Illustration

For empirical study three data sets are given as:
The Population I has been taken from [7] , Population II from [3] and Population III from [4].

Table 1: *Descriptive Statistics of the populations*

Population I	Population II	Population III
$N=106$	$N=100$	$N=278$
$n=20$	$n=10$	$n=30$
$\rho=0.82$	$\rho=0.6500$	$\rho=0.7213$
$S_y=64.25$	$S_y=14.6595$	$S_y=56.4571$
$S_x=491.89$	$S_x=7.5323$	$S_x=40.6747$
$\beta_{2(x)}=25.71$	$\beta_{2(x)}=2.2387$	$\beta_{2(x)}=38.8898$
$\beta_{2(y)}=80.13$	$\beta_{2(y)}=2.3523$	$\beta_{2(y)}=25.8969$
$\lambda_{22}=33.30$	$\lambda_{22}=1.5432$	$\lambda_{22}=26.8142$

Table 2: *Results obtained for different estimators using above populations*

Population	Estimators	MSE	PRE
Population I	t_0	67,423,045.30	100
	t_1	33,431,595.28	201.66
	t_2	45164775.33	149.28
	t^*	314489.80	214.28
Population II	t_0	6245.40	100
	t_1	6948.36	89.87
	t_2	5166.55	120.88
	t^*	5145.31	121
Population III	t_0	8431372.53	100
	t_1	82660.52	102
	t_2	68547.74	123
	t^*	35575.41	237

IV. Conclusion

In the realm of survey sampling, a novel approach called the Generalized Exponential Ratio-Cum-Product Estimator (GERPE) for sampling variance in Simple Random Sampling (SRS) has been introduced. This method offers a more robust and efficient means of estimating sampling variance as compared to existing techniques discussed in the literature. The crux of GERPE lies in its ability to strike a balance between ratio and product estimators, leading to enhanced precision in variance estimation. Remarkably, when GERPE is operating at its optimal settings, its Mean Squared Error (MSE) attains the same value as that of the conventional regression estimator. This is a significant advantage because the usual regression estimator has been proven to exhibit lower MSE compared to other existing estimators. Consequently, the application of GERPE not only outperforms the methods previously documented but also aligns itself with the performance level of the superior regression estimator, thus promising more accurate and reliable variance estimates in SRS scenarios. This innovative approach opens up new avenues for improving the precision and efficiency of sampling variance estimation in the field of survey sampling and can be a valuable addition to the statistical toolbox of researchers and practitioners alike.

Acknowledgement: We are thankful to Editor-in-chief for his suggestions that improved the manuscript.

References

[1] Agrawal, M. C., Sthapit, A. B. (1995). Unbiased ratio-type variance estimation. *Stat. Prob. Lett.*, 25:361–364.

[2] Bahl, S. & Tuteja, R. K. (1991). Ratio and product type exponential estimator. *Journal of Information and Optimization Sciences*, 12(1): 159 163.

[3] Cochran, W. G. Sampling techniques (3rd ed.). Wiley Eastern Limited,1977.

[4] Das, A. K. (1988). Contribution to the theory of sampling strategies based on auxiliary information. Ph.D. thesis submitted to BCKV, Mohanpur, Nadia, West-Bengal, India.

[5] Gupta, S., Shabbir, J. (2008). Variance estimation in simple random sampling using auxiliary information. *Hacett. J. Math. Stat.*, 371:57–67.

[6] Isaki, C.T. (1983). Variance estimation using auxiliary information. J. *Am. Stat. Assoc.*, 78381:117–123.

[7] Kadilar, C., Cingi, H. (2007). Improvement in variance estimation in simple random sampling. *Commun. Stat. Theo. Meth.*, 36:2075–2081.

[8] Prasad, B., Singh, H. P. (1990). Some improved ratio-type estimators of finite population variance in sample surveys. *Commun. Stat. Theo. Meth.*, 193:1127–1139.

[9] Singh, H.P. (1986). A note on the estimation of variance of sample mean using the knowledge of coefficients of variation in normal population. *Commun. Stat. Theo. Meth.,* 1512:3737–3746.

[10] Singh, H. P., Pal, S. K. (2016). An efficient class of estimators of finite population variance using quartiles. *J. Appl. Stat.*, 10.1080/02664763.2015.1125865

[11] Singh, H. P., Solanki, R. S. (2013a). A new procedure for variance estimation in simple random sampling using auxiliary information. *Statistical Papers,* 542:479–497.

[12] Singh, H. P., Solanki, R. S. (2013b). Improved estimation of finite population variance using auxiliary information. *Commun. Stat. Theo. Meth*., 215:2718–2730.

[13] Singh, H. P., Upadhyaya, L. N., Namjoshi, U. D. (1988). Estimation of finite population variance. *Curr. Sci.,* 5724:1331–1334.

[14] Srivastava, S. K., Jhajj, H. S. (1980). A class of estimators using auxiliary information for estimating finite population variance. *Sankhya, C*, 421(2):87–96.

[15] Upadhyaya, L. N., Singh, H. P. (1986). On a dual to ratio estimator for estimating finite population variance. *Nepal Math. Sci. Rep*., 111:37–42.

[16] Vishwakarma, G. K., Zeeshan, S. M. and Bouza, C. N. (2017). Ratio and product type exponential estimators for population mean using ranked set sampling. *Revista Investigacion Operacional*, 38(3): 266-271.

[18] J. Neyman (1934). On two different aspects of the representative methods, the method of stratified sampling and the method of purposive selection, *Journal of. Royal Statistical. Society*, 97: 558-606.

[19] Cochran, W. G. Ratio Estimators, In Sampling Techniques, 3rd ed., India: Wiley Eastern Limited.1940.

[20] Murthy, M. N. (1964). Product Method of Estimation, *Sankhya*., 26(A): 69–74.

[21] Olkin , I. (1958). Multivariate ratio estimation for finite populations, *Biometrika*., 45(1-2), 154–165.

[22] Singh, M. P. (1967) . Ratio-cum-product method of estimation, *Metrika*., 12: 34–72, .

[23] Chand, L. (1975). Some ratio type estimators based on two or more auxiliary variables,

Unpublished Ph.D. dissertation, Iowa State University, Iowa.

[24] Bahl, S. and Tuteja, R. K. (1991). Ratio and product type exponential estimator, *Journal of Information and Optimization Sciences*., 12: 159–163.

[25] Upadhyaya, L. N. and Singh, H. P. (1999). Use of Transformed Auxiliary Variable in Estimating the Finite Population Mean, *Biometrical Journal*., 41: 627–636.

[26] Abu-Dayeh, W. A., Ahmed, M. S. , Ahmed, R. A. and Muttlak, H. A. (2003). Some Estimators of a Finite Population Mean Using Auxiliary Information, *Applied Mathematics and Computation*., 139: 287–298.

[27] Kadilar, C. and Cingi, H. (2005). A new estimator using two auxiliary variables, *Applied Mathematics and Computation*., 162: 901– 908.

[28] Diana, G. and Perri, P. F. (2007). Estimation of finite population mean using multi-auxiliary information, *Metron*., 65(1): 99–112.

[29] Shabbir, J. and Gupta, S. (2011). On estimating finite population mean in simple and stratified random sampling, *Communications in Statistics- Theory & Methods*., 40: 199–212.

[30] Yasmeen, U. , Noor ul Amin, M. and Hanif , M. (2016). Exponential ratio and product type estimators of finite population mean, *Journal of Statistics and Management Systems*., 19: 55–71.

[31] Singh, S. Advanced Sampling Theory with Applications-How Michael 'selected' Amy, Vol-II. London: Kluwer Academic Publishers. 2003

[32] Gupta, R. K. and Yadav, S. K. (2017). New Efficient Estimators of Population Mean Using Non-Traditional Measures of Dispersion, *Open Journal of Statistics*., 7: 394-404.

[33] Muili, J. O. , Agwamba, E. N. , Erinola, Y. A. , Yunusa, M. A . , Audu, A. and Hamza, M. A. (2020). Modified Ratio-Cum-Product Estimators of Population Mean Using Two Auxiliary Variables, *Asian Journal of Research in Computer Science*., 6(1): 55-65.

[34] Singh, H. P. and Yadav, A. (2020). A New Exponential Approach for Reducing the Mean Squared Errors of the Estimators of Population Mean Using Conventional and Non-Conventional Location Parameters, *Journal of Modern Applied Statistical Methods*., 18(1).