

Assessing the Efficiencies of Different Ratio Estimators for Estimating Average Production of Sugarcane Yield

Surendra Kumar

Department of Mathematics

Mahamaya Govt. Degree College, Mahona, Lucknow (U.P.), India

Email: surendra.kumar776@gmail.com

Mehdi Ali*, Subhash Kumar Yadav

Department of Statistics

Babasaheb Bhimrao Ambedkar University, Lucknow (U.P.), India

Email: alimehdi244@gmail.com, drskystats@gmail.com

***Corresponding author**

Abstract

In this research paper, we conducted a thorough examination of various existing ratio estimators utilizing a single auxiliary variable by analyzing real population data. The study systematically compares their efficiencies, employing a robust methodology. The experimental setup involves the selection of a sample from a real population, utilizing the simple random sampling without replacement method. Unlike simulations, our approach involves direct application to actual population data, enhancing the external validity of the findings. We evaluate the performance of different ratio estimators concerning bias, mean square error, providing valuable insights into their real-world applicability. The results revealed that, the estimator t_{13} is the most efficient and t_2 is the least efficient estimator of population mean. This comprehensive analysis contributes valuable insights into the comparative efficiency and performance characteristics of various ratio estimators under the given conditions, providing a detailed understanding of their applicability in estimating population parameters.

Keywords: Sampling, Main Variable, Auxiliary Variable, Bias, MSE.

Introduction

Sample surveys serve as a widely employed and cost-effective means of gathering data to draw valid inferences about the population parameters. Within the realm of sample surveys, it becomes feasible to not only measure the specific characteristic under study but also to assess other relevant attributes that exhibit a high correlation with the study variable. This supplementary data, known as auxiliary information, contributes valuable insights to the overall understanding of the subject matter. Numerous sample surveys have been conducted both in India and internationally, leveraging auxiliary information that demonstrates a significant correlation with the variable of interest. This strategic utilization of additional information enhances the precision and comprehensiveness of the findings, enabling researchers and policymakers to derive more nuanced insights into the intricacies of the studied phenomena. Consequently, the incorporation of auxiliary information in sample surveys becomes a private tool in enhancing the accuracy and reliability of the conclusions drawn from the collected data.

Tripathi (1978) employed auxiliary information from one or more variables in sample surveys through three fundamental approaches Muhammad *et al.* (2019), Kumar and Kumar (2020), Ahuja *et al.* (2021) and others have extensively employed auxiliary information in various forms. This application of auxiliary information aims to enhance the performance of estimators for the study variable. Particularly, the ratio estimator tends to exhibit robust performance when a positive correlation exists between the study variable and auxiliary variables.

The effectiveness of the ratio method of estimation has been further refined by Sisodia and Dwivedi (1981), Bahl and Tuteja (1991), Upadhyaya and Singh (1991), and Kadilar and Cingi (2004). Noteworthy contributions to different aspects of the ratio method of estimation include works by Subramani and Kumarapandiyan (2012), Abid *et al.* (2016), Kanwai *et al.* (2016), Singh *et al.* (2019), Yadav *et al.* (2020), Baghel and Yadav (2020), Singh and Yadav (2020), Tiwari *et al.* (2021) Nderitu *et al.* (2022), Ali *et al.* (2023) and Adejumobi *et al.* (2023). These studies collectively contribute to the ongoing development and refinement of ratio-based estimation methods in sample surveys. In statistical research, obtaining authentic real-world data sets for the purpose of comparing estimators under practical conditions can be a complex endeavour. Recognizing the inherent challenges in sourcing genuine data, this study deviates from conventional approaches and relies on an actual data set for the assessment of estimators. Unlike simulation-based methods, which generate synthetic data, the methodology employed here involves scrutinizing the properties of estimators using a bona fide data set.

Drawing inspiration from the need for empirical relevance, this research endeavours to contribute insights based on the analysis of estimators applied to a real data set. This departure from simulation-based studies ensures that the findings are grounded in the complexities and nuances present in the actual data, providing a more authentic basis for assessing the performance of the estimators under examination.

Materials and Method

Table 1: List of different ratio estimators of population mean with their bias and mean squared errors

Estimator	Bias	Mean Square Error
$t_r = \bar{y} \left(\frac{C_y}{C_x} \right)$ Cochran (1940)	$\frac{(1-f)}{n} Y [C_x^2 - \rho C_y C_x]$	$\frac{(1-f)}{n} Y^2 [C_y^2 + C_x^2 - 2\rho C_y C_x]$
$t_1 = \bar{y} \left(\frac{C_y}{\bar{x} + C_x} \right)$ Sisodia and Dwivedi (1981)	$\frac{(1-f)}{n} Y [\theta^2 C_x^2 - \theta_1 \rho C_y C_x]$	$\frac{(1-f)}{n} Y^2 [C_y^2 + \theta^2 C_x^2 - 2\theta_1 \rho C_y C_x]$

$t_2 = \bar{y} \left(\frac{X_x + \beta_2}{\bar{x}C_x + \beta_2} \right)$ Upadhyaya and Singh (1999)	$\frac{(1-f)}{n} \bar{y} [\theta_2^2 C_x^2 - \theta_2 \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_2^2 C_x^2 - 2\theta_2 \rho C_y C_x]$
$t_3 = \bar{y} \left(\frac{X_x + C_x}{\bar{x}\beta_2 + C_x} \right)$ Upadhyaya and Singh (1999)	$\frac{(1-f)}{n} \bar{y} [\theta_3^2 C_x^2 - \theta_3 \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_3^2 C_x^2 - 2\theta_3 \rho C_y C_x]$
$t_4 = \bar{y} \left(\frac{X + \rho}{\bar{x} + \rho} \right)$ Singh and Tailor (2003)	$\frac{(1-f)}{n} \bar{y} [\theta_4^2 C_x^2 - \theta_4 \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_4^2 C_x^2 - 2\theta_4 \rho C_y C_x]$
$t_5 = \bar{y} \left(\frac{X + \beta_2}{\bar{x} + \beta_2} \right)$ Singh <i>et al.</i> (2004)	$\frac{(1-f)}{n} \bar{y} [\theta_5^2 C_x^2 - \theta_5 \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_5^2 C_x^2 - 2\theta_5 \rho C_y C_x]$
$t_6 = \bar{y} \left(\frac{X + \beta_1}{\bar{x} + \beta_1} \right)$ Yan and Tian (2010)	$\frac{(1-f)}{n} \bar{y} [\theta_6^2 C_x^2 - \theta_6 \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_6^2 C_x^2 - 2\theta_6 \rho C_y C_x]$
$t_7 = \bar{y} \left(\frac{X_x + \beta_1}{\bar{x}C_x + \beta_1} \right)$ Yan and Tian (2010)	$\frac{(1-f)}{n} \bar{y} [\theta_7^2 C_x^2 - \theta_7 \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_7^2 C_x^2 - 2\theta_7 \rho C_y C_x]$
$t_8 = \bar{y} \left(\frac{X_x + M_d}{\bar{x}C_x + M_d} \right)$ Subramani and Kumarpandiyam (2012)	$\frac{(1-f)}{n} \bar{y} [\theta_8^2 C_x^2 - \theta_8 \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_8^2 C_x^2 - 2\theta_8 \rho C_y C_x]$
$t_9 = \bar{y} \left(\frac{X_1 + Q_d}{\bar{x}\beta_1 + Q_d} \right)$ Jeelani <i>et al.</i> (2013)	$\frac{(1-f)}{n} \bar{y} [\theta_9^2 C_x^2 - \theta_9 \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_9^2 C_x^2 - 2\theta_9 \rho C_y C_x]$
$t_{10} = \bar{y} \left(\frac{X + n}{\bar{x} + n} \right)$ Jerajuddin and Kishun (2016)	$\frac{(1-f)}{n} \bar{y} [\theta_{10}^2 C_x^2 - \theta_{10} \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_{10}^2 C_x^2 - 2\theta_{10} \rho C_y C_x]$

$t_{11} = \bar{y} \left(\frac{Xn + \rho}{\bar{x}n + \rho} \right)$ <p>Yadav <i>et al.</i> (2019)</p>	$\frac{(1-f)}{n} \bar{y} [\theta_{11}^2 C_x^2 - \theta_{11} \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_{11}^2 C_x^2 - 2\theta_{11} \rho C_y C_x]$
$t_{12} = \bar{y} \left(\frac{Xn + C_x}{\bar{x}n + C_x} \right)$ <p>Yadav <i>et al.</i> (2019)</p>	$\frac{(1-f)}{n} \bar{y} [\theta_{12}^2 C_x^2 - \theta_{12} \rho C_y C_x]$	$\frac{(1-f)}{n} \bar{y}^2 [C_y^2 + \theta_{12}^2 C_x^2 - 2\theta_{12} \rho C_y C_x]$
$t_{13} = \bar{y} \left(\frac{\bar{x}}{X} \right)^\alpha \left[\beta \exp \left\{ \frac{(a\bar{X}+b) - (a\bar{x}+b)}{(a\bar{X}+b) + (a\bar{x}+b)} \right\} + (1-\beta) \left\{ 1 + \log \frac{(a\bar{x}+b)}{(a\bar{X}+b)} \right\} \right]$ <p>Ali <i>et al.</i> (2023)</p>	$\frac{(1-f)}{n} \bar{y} \left[-\frac{7}{8} \beta \theta^2 C_x^2 - \frac{1}{2} \theta^2 C_x^2 - \frac{3}{2} \alpha \beta \theta C_x^2 + \alpha \theta C_x^2 + \frac{\alpha(\alpha-1)}{2!} C_x^2 - \frac{3}{2} \beta \theta C_{yx} + \theta C_{yx} + \alpha C_{yx} \right]$	$\bar{y}^2 (A + \alpha^2 B + \alpha C + \beta^2 D - \beta F - \alpha \beta F)$

Where,

$$A = \lambda(\theta^2 C_x^2 + C_x^2 + 2\theta C_{yx}), B = \lambda C_x^2$$

$$C = 2\lambda(\theta C_x^2 + C_{yx}^2)$$

$$D = -\frac{9}{4} \lambda \theta^2 C_x^2$$

$$F = \frac{4}{3} \lambda \theta^2 C_x^2 + 3\lambda \theta C_{yx}$$

$$G = 3\lambda \theta C_x^2$$

and the optimum values of α and β respectively are,

$$\alpha = \frac{(FG-2DC)}{(4BD-G^2)}, \beta = \frac{(2BF-CG)}{(4BD-G^2)}$$

$$\theta_1 = \frac{X}{X+C_x}, \theta_2 = \frac{XC_x}{XC_x+\beta_2}, \theta_3 = \frac{X\beta_2}{X\beta_2+C_x}, \theta_4 = \frac{X}{X+\rho}, \theta_5 = \frac{X}{X+\beta_2}, \theta_6 = \frac{X}{X+\beta_1}$$

$$\theta_7 = \frac{XC_x}{XC_x+\beta_1}, \theta_8 = \frac{XC_x}{XC_x+M_d}, \theta_9 = \frac{X\beta_1}{X\beta_1+Q_d}, \theta_{10} = \frac{X}{X+n}, \theta_{11} = \frac{Xn}{Xn+\rho}, \theta_{12} = \frac{Xn}{Xn+C_x}$$

$\theta = \frac{aX}{aX+b}$, where, a and b can be either a constant or any of the auxiliary variable's parameters,

such as $n, C_x, \beta_1, \beta_2, \rho, Q_d, MR, M_d$, and fourth and sixth decile of the secondary variable or quartiles of auxiliary variable but here we have considered $a = n, b = \beta_1$.

Numerical Study

In this segment, comparison of various estimators has been made using a real population dataset.

The parameters and constants pertinent to the study are elucidated in Table 2.

Table 2: Parameters of population under inspection

$N = 181$	$n = 25$	$S_x = 0.519118$	$S_y = 453.9736$
$C_x = 0.526141$	$C_y = 0.500406$	$\bar{Y} = 907.210$	$\bar{X} = 0.987$
$D_6 = 0.961$	$\beta_1 = 3.4641$	$\beta_2 = 19.49481$	$M_d = 0.890$
$QD = 0.204$	$Q_1 = 0.71$	$Q_3 = 1.118$	$\rho = 0.923328$

Table 3: Bias and MSEs of different estimators

Estimator	Bias	MSE	PRE
t_0	1.055	1164.341	1
t_1	-1.276	1448.507	0.8038
t_2	-0.191	6752.394	0.1724
t_3	-1.640	2190.552	0.5315
t_4	-1.617	2074.212	0.5613
t_5	-0.346	6458.505	0.1803
t_6	-1.260	4432.24	0.2627
t_7	-0.844	5440.093	0.2140
t_8	-1.626	3088.21	0.3770
t_9	-0.582	1104.108	1.0546
t_{10}	-0.276	6592.426	0.1766
t_{11}	0.716	1105.525	1.0532
t_{12}	0.856	1127.806	1.0324
t_{13}	-0.0096	1047.762	1.1113

The results in Table 3 have also been presented in the form of graph in Figure 1.

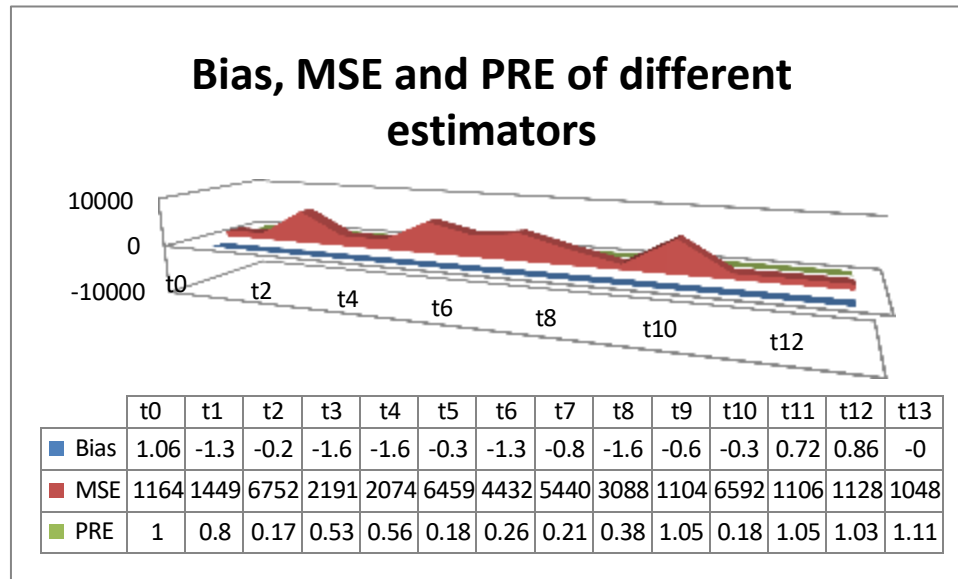


Figure 1: Bias, MSE and PRE of different estimators

Table 3 provides the bias, MSE and PRE of different estimators. Correlation coefficient between the auxiliary variable and main variable was 0.9233 and the sample size taken was 25. It is concluded from the table that the estimator by Ali *et al.* (t_{13}) was the best estimator and estimator by Upadhyaya and Singh (t_2) was the least performing estimator.

Result and Conclusion

Table 3 presents the bias, Mean Squared Error (MSE), and Percent Relative Efficiencies (PRE) values for various estimators. With a correlation coefficient of 0.9233 between the auxiliary and main variables, and a sample size of 25, the analysis indicates that the estimator proposed by Ali *et al.* (t_{13}) emerged as the most effective, while the estimator by Upadhyaya and Singh (t_2) demonstrated the lowest performance among the considered estimators. This thorough examination enhances our understanding of the relative efficiency and performance attributes of diverse ratio estimators across varying conditions. The findings offer valuable insights, shedding light on the practical suitability of these estimators for accurately estimating population parameters with real-world data.

Acknowledgement

The authors Surendra Kumar and Subhash Kumar Yadav express their gratitude to Government of Uttar Pradesh for financial assistance of this research work through project no. 46/2021/603/Sattar-4-2021-4(56)/2020 dated 30.03.2021.

References

Abid, M., Abbas, N., Sherwani, R.A.K. and Nazir, H.Z. (2016) ‘Improved ratio estimators for the population mean using non-conventional measure of dispersion’, *Pakistan Journal of Statistics and Operations Research*, Vol. 11, No. 2, pp.353–367.

- Ali, M., Yadav, S.K., Gupta, R.K., Kumar, S., Singh, L. (2023) 'Enhancement of Class of Estimators for Estimating Population Mean Using Known Auxiliary Parameters', *International Journal of Agricultural and Statistical Sciences*. DOI: <https://doi.org/10.59467/IJASS.2023.19.527>
- Baghel, S. and Yadav, S.K. (2020) 'Restructured class of estimators for population mean using an auxiliary variable under simple random sampling scheme', *JAMSI*, Vol. 16, No. 1, pp.61–75.
- Cochran, W.G. (1940) 'The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce', *The Journal of Agricultural Science*, Vol. 30, No. 2, pp.262–275.
- Gupta, R.K. and Yadav, S.K. (2017) 'New efficient estimators of population mean using non-traditional measures of dispersion', *Open Journal of Statistics*, Vol. 7, No. 3, pp.394–404.
- Gupta, R.K. and Yadav, S.K. (2018) 'Improved estimation of population mean using information on size of the sample', *American Journal of Mathematics and Statistics*, Vol. 8, No. 2, pp.27–35.
- Ijaz, M. and Ali, H. (2018) 'Some improved ratio estimators for estimating mean of finite population', *Research & Reviews: Journal of Statistics and Mathematical Sciences*, Vol. 4, No. 2, pp.18–23.
- Jeelani, M.I., Maqbool, S. and Mir, S.A. (2013) 'Modified ratio estimators of population mean using linear combination of coefficient of skewness and quartile deviation', *International Journal of Modern Mathematical Sciences*, Vol. 6, No. 3, pp.174–183.
- Jerajuddin, M. and Kishun, J. (2016) 'Modified ratio estimators for population mean using size of the sample, selected from population', *IJSRSET*, Vol. 2, No. 2, pp.10–16.
- Kadilar, C. and Cingi, H. (2006) 'An improvement in estimating the population mean by using the correlation coefficient', *Hacettepe Journal of Mathematics and Statistics*, Vol. 35, No. 1, pp.103–109.
- Sharma, P. and Singh, R. (2013) 'Improved estimators for simple random sampling and stratified random sampling under second order of approximation', *Statistics in Transition – New Series*, Vol. 14, No. 3, pp.379–390.
- Singh, H.P. and Tailor, R. (2003) 'Use of known correlation co-efficient in estimating the finite population means', *Statistics in Transition*, Vol. 6, No. 4, pp.555–560.
- Singh, H.P., Tailor, R., Tailor, R. and Kakran, M.S. (2004) 'An improved estimator of population mean using power transformation', *Journal of the Indian Society of Agricultural Statistics*, Vol. 58, No. 2, pp.223–230.
- Sisodia, B.V.S. and Dwivedi, V.K. (1981) 'A modified ratio estimator using co-efficient of variation of auxiliary variable', *Journal of the Indian Society of Agricultural Statistics*, Vol. 33, No. 1, pp.13–18.
- Srija, R. and Subramani, J. (2018) 'Modified ratio cum product estimators for finite population mean with known median and mean', *SSRG International Journal of Economics Management Studies, Special Issue ICRTECITA*, April, pp.1–9.

- Subramani, J. (2016) 'A new median based ratio estimator for estimation of the finite population mean', *Statistics in Transition New Series*, Vol. 17, No. 4, pp.1–14.
- Subramani, J. and Ajith, S. (2017) 'Improved ratio cum product estimators for finite population mean with known quartiles and their functions', *SMJ Biometrics Biostat.*, Vol. 2, No. 1, pp.1008.
- Subramani, J. and Kumarapandiyan, G. (2012a) 'Estimation of population mean using coefficient of variation and median of an auxiliary variable', *International Journal of Probability and Statistics*, Vol. 1, No. 4, pp.111–118.
- Subramani, J. and Kumarapandiyan, G. (2012b) 'Estimation of population mean using known median and co-efficient of skewness', *American Journal of Mathematics and Statistics*, Vol. 2, No. 5, pp.101–107.
- Subramani, J. and Kumarapandiyan, G. (2013) 'A new modified ratio estimator of population mean when median of the auxiliary variable is known', *Pakistan Journal of Statistics and Operation Research*, Vol. 9, No. 2, pp.137–145.
- Subzar, M., Maqbool, S., Raja, T.A., Mir, S.A., Jeelani, M.I. and Bhat, M.A. (2017) 'Improved family of ratio type estimators for estimating population mean using conventional and non conventional location parameters', *Revista Investigacion Operacional*, Vol. 38, No. 5, pp.510–524.
- Subzar, M., Maqbool, S., Raja, T.A., Pal, S.K. and Sharma, P. (2018) 'Efficient estimators of population mean using auxiliary information under simple random sampling', *Statistics in Transition – New Series*, Vol. 19, No. 2, pp.219–238.
- Sukhatme, P.V. (1954) *Sampling Theory of Surveys with Applications*, Indian Society of Agricultural Statistics.
- Upadhyaya, L.N. and Singh, H.P. (1999) 'Use of transformed auxiliary variable in estimating the finite population mean', *Biometrical Journal*, Vol. 41, No. 5, pp.627–636.
- Wolter, K.M. (1985) *Introduction to Variance Estimation*, Springer-Verlag, New York.
- Yadav, S.K. and Kadilar, C. (2013a) 'Improved class of ratio and product estimators', *Applied Mathematics and Computation*, Vol. 219, No. 22, pp.10726–10731.
- Yadav, S.K. and Kadilar, C. (2013b) 'Efficient family of exponential estimator for population mean', *Hacettepe Journal of Mathematics and Statistics*, Vol. 42, No. 6, pp.671–677.
- Yadav, S.K. and Mishra, S.S. (2015) 'Developing improved predictive estimator for finite population mean using auxiliary information', *Statistika*, Vol. 95, No. 1, pp.76–85.
- Yadav, S.K., Dixit, M.K., Dungana, H.N. and Mishra, S.S. (2019a) 'Improved estimators for estimating average yield using auxiliary variable', *International Journal of Mathematical, Engineering and Management Sciences*, Vol. 4, No. 5, pp.1228–1238.
- Yadav, S.K., Sharma, D.K. and Brown, K. (2019b) 'Estimating peppermint oil yields with auxiliary variable information', *International Journal of Operational Research*, Accepted.
- Yadav, S.K., Gupta, S., Mishra, S.S. and Shukla, A.K. (2016) 'Modified ratio and product estimators for estimating population mean in two-phase sampling', *American Journal of Operational Research*, Vol. 6, No. 3, pp.61–68.

- Yadav, S.K., Sharma, D.K. and Brown, K. (2020) 'New class of estimators of the population mean using the known population median of the study variable', *International Journal of Mathematics in Operational Research*, Vol. 16, No. 2, pp.179–201.
- Yadav, S.K., Sharma, D.K., Mishra, S.S. and Shukla, A.K. (2018) 'Use of auxiliary variables in searching efficient estimator of population mean', *International Journal of Multivariate Data Analysis*, Vol. 1, No. 3, pp.230–244.
- Yan, Z. and Tian, B. (2010) 'Ratio method to the mean estimation using co-efficient of skewness of auxiliary variable', *ICICA, Part II, CCIS*, Vol. 106, pp.103–110.