GANITA, Vol.75(1), 2025, 259-267

A NEW CLASS OF RATIO ESTIMATORS FOR POPULATION MEAN USING AUXILIARY PARAMETERS: AN APPLICATION TO REAL-WORLD DATA

GAGAN KUMAR, DUSHYANT TYAGI, GAUTAM GUPTA and SURENDRA KUMAR

Abstract

This study introduces a naive general family of ratio estimators that leverage various auxiliary measures to estimate the population mean of a study variable under a simple random sampling without replacement framework. The paper examines specific cases that incorporate auxiliary information such as the coefficient of variation, median, and quartile deviation. The bias and Mean Square Error (MSE) of the introduced estimators are retained to a first-order approximation. Furthermore, theoretical conditions for comparing the efficiency of the introduced estimators with existing ones are provided, and their performance is validated using real-world data. Numerical analysis demonstrates that the proposed estimators are more efficient than other ratio-based estimators, appealing for the use in various areas of applications of the real world, which includes agriculture, biological sciences, defence, economics, mathematical sciences. management, medical sciences, social sciences etc.

2010 *Mathematics subject classification:* 62D05; 94A20.

Keywords and phrases: Study Variable, Auxiliary Parameter, Median, Bias, Mean Square Error.

1. Introduction

An alternative to thorough enumeration that saves time, money, and labour is sampling. The sample mean $\bar{y}$ is most appropriate for the population mean $\bar{Y}$ since the associated statistics is the best estimator for the parameter being studied. When estimating the parameters of the primary variable Y, auxiliary variable X is quite important. For higher population mean estimate, ratio type estimators are employed if it has a positive correlation with the primary variable. In contrast, when X and Y have a negative correlation, product type estimators are employed to improve the estimation of $\bar{Y}$. Sampling strategies are frequently used when working with huge populations or when it is impractical to obtain data from every unit. By looking at a portion of the population, these techniques offer a useful way to investigate the properties of a target variable. Population parameters are usually estimated using sample statistics; for instance, the sample mean $\bar{y}$ is a commonly used estimator for the population mean $\bar{Y}$ due to its objectivity, despite the possibility of substantial variance. Much effort has gone into creating more effective estimators for $\bar{Y}$ by enhancing current techniques. Utilizing auxiliary data to improve the accuracy of sample-based population parameter estimations is the main goal of many of these initiatives. Many authors from all over the globe used the auxiliary information in different forms for enhanced estimation of $\bar{Y}$ and suggested various efficient estimators of $\bar{Y}$.

Cochran [1] introduced the classical ratio estimator, utilizing information on X that is highly positively correlated with Y. Srivastava [2] proposed an improved estimator for by incorporating X information in sample surveys. Sisodia $\bar{Y}$ and Dwivedi [3] developed a modified ratio estimator for $\bar{Y}$ using the known coefficient of variation (CV) of X. Rao [4] explored methods to efficient estimation of $\bar{Y}$ through ratio and regression estimators. Singh et al. [5] proposed a new estimator for $\bar{Y}$ by utilizing the coefficient of kurtosis, while Upadhyaya and Singh [6] worked on an elevated estimator for $\bar{Y}$ by combining the known CV and coefficient of kurtosis of X. Singh and Tailor [7] enhanced the ratio estimator $\bar{Y}$ for using the known correlation coefficient, Kadilar and Cingi [8] focused on developing improved ratio estimators for better estimation of $\bar{Y}$. In double sampling, Singh and Vishwakarma [9] and [10] created an effective variation of the product and ratio estimators and in sample surveys, certain estimators of the finite population mean use auxiliary data. Yadav et al. [11] introduced modified ratio estimators leveraging non-traditional auxiliary parameters for efficient estimation of $\bar{Y}$. Khoshnevisan et al. [12] proposed a general family of ratio estimators for $\bar{Y}$, identifying various estimators as members of their family. Koyuncu and Kadilar [13] designed efficient estimators for $\bar{Y}$, while Yan and Tian [14] improved the estimation of $\bar{Y}$ using the known coefficient of skewness of X. Subramani and Kumarapandiyan [15] proposed an efficient estimator for $\bar{Y}$ by utilizing known quartile information of X. Jeelani et al. [16] developed new ratio estimators for Y, combining the coefficient of skewness and quartile deviation of X. Finally, Subramani [17] introduced a general family of estimators for based on auxiliary parameters.

Abid et al. [18] proposed enhanced ratio estimators for $\bar{Y}$ by employing unconventional measures of dispersion, while Subzar et al.[19] developed an efficient class of ratio estimators for $\bar{Y}$ using known auxiliary parameters. Unal and Kadilar [20], Singh and Usman [21] introduced improved estimators for $\bar{Y}$ that accounted for non-response scenarios by incorporating auxiliary parameters. Priam [22] proposed a generalized family of estimators for the efficient estimation of $\bar{Y}$. Yadav et al. [23] focused on efficient estimation of average paddy production using robust measures, and Yadav et al.[24] introduced a naive class of estimators aimed at improving the estimation of average peppermint yield through known auxiliary information (X). Zaman et al. [25] explored robust ratio-type estimators for $\bar{Y}$ in simple random sampling (SRS) using simulation studies. Ali et al. [26] proposed an advanced class of estimators for $\bar{Y}$ leveraging known auxiliary parameters, while Singh et al.[27] introduced a refined family of estimators utilizing X information for improved estimation of $\bar{Y}$. Vishwakarma [28] Weighted technique for estimating the finite population mean in the presence of auxiliary data. Numerous other researchers have also contributed to the development of enhanced estimators for $\bar{Y}$ by utilizing known auxiliary parameters. The remainder of this paper is organized into various sections.

The remainder of the work is divided into many areas, such as the theoretical efficiency comparison, the proposed estimator, the review of estimators, the numerical analysis, and the comparison of the results.

2. Review of Estimators

Cochran [1] utilized the auxiliary information and suggested the usual ratio estimator of $\bar{Y}$ as,

$$t_0 = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)$$

The MSE of t_R for an approximation of order one is given by,

$$MSE(t_R) = \lambda \bar{Y}^2(C_y^2 + R^2 C_x^2 - 2RC_{yx}) \tag{2.1}$$

Where,

$\lambda = \frac{1}{n} - \frac{1}{N}$, $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$, $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$, $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$,

$S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$, $S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})(X_i - \bar{X})$

Sisodia and Dwivedi [3] used the CV of X and suggested the following estimator of $\bar{Y}$ as,

$$t_1 = \bar{y}\left(\frac{\bar{X} + C_x}{\bar{x} + C_x}\right)$$

The MSE of t_1 is given by,

$$MSE(t_1) = \lambda \bar{Y}^2(C_y^2 + R_1^2 C_x^2 - 2R_1 C_{yx}) \tag{2.2}$$

Where, $R_1 = \frac{\bar{X}}{\bar{X}+C_x}$

Singh at al. [5] used coefficient of kurtosis of X and given an estimator of $\bar{Y}$ as,

$$t_2 = \bar{y}\left(\frac{\bar{X} + \beta_2}{\bar{x} + \beta_2}\right)$$

$$MSE(t_2) = \lambda \bar{Y}^2(C_y^2 + R_2^2 C_x^2 - 2R_2 C_{yx}) \tag{2.3}$$

Where, $R_2 = \frac{\bar{X}}{\bar{X}+\beta_2}$

Upadhyaya and Singh [6] proposed two estimators of $\bar{Y}$ by using C_x and β_2 as,

$$t_3 = \bar{y}\left(\frac{\bar{X}\beta_2 + C_x}{\bar{x}\beta_2 + C_x}\right), \quad t_4 = \bar{y}\left(\frac{\bar{X}C_x + \beta_2}{\bar{x}C_x + \beta_2}\right)$$

$$MSE(t_i) = \lambda \bar{Y}^2(C_y^2 + R_i^2 C_x^2 - 2R_i C_{yx}), \quad i = 3, 4 \tag{2.4}$$

Where, $R_3 = \frac{\bar{X}\beta_2}{\bar{X}\beta_2+C_x}$, $R_4 = \frac{\bar{X}C_x}{\bar{X}C_x+\beta_2}$

Zaman at el.[29] introduced a class of estimators of $\bar{Y}$ as,

$$t_5 = \bar{y}\left(\frac{\bar{X} + \beta_2/G(x)}{\bar{x} + \beta_2/G(x)}\right)$$

Where, $G(x)$ is the function of X.

Some of the special cases of t_5 are,

$$t_{5(1)} = \bar{y}\left(\frac{\bar{X} + \beta_2/C_x}{\bar{x} + \beta_2/C_x}\right), \quad t_{5(2)} = \bar{y}\left(\frac{\bar{X} + \beta_2/M_d}{\bar{x} + \beta_2/M_d}\right), \quad t_{5(3)} = \bar{y}\left(\frac{\bar{X} + \beta_2/QD}{\bar{x} + \beta_2/QD}\right)$$

Where M_d is the median and QD is the quartile deviation of X.

The MSE of $t_{5(1)}$ are given by,

$$MSE(t_{5(i)}) = \lambda \bar{Y}^2(C_y^2 + R_{5(i)}^2 C_x^2 - 2R_{5(i)} C_{yx}), \quad i = 1,\ 2,\ 3 \tag{2.5}$$

Where, $R_{5(1)} = \frac{\bar{X}}{\bar{X}+\beta_2/C_x}$, $R_{5(2)} = \frac{\bar{X}}{\bar{X}+\beta_2/M_d}$, $R_{5(3)} = \frac{\bar{X}}{\bar{X}+\beta_2/QD}$,

3. Proposed estimator

Motivated by Zaman et al. [29], many other authors in the literature and thinking of that many estimators will be the particular cases of the proposed class of estimator, we have introduced a generalized estimator for estimation of $\bar{Y}$ as,

$$t_p = \bar{y}\left(\frac{\bar{X} + \beta_2/G(x)}{\bar{x} + \beta_2/G(x)}\right)^{\alpha} \tag{3.1}$$

Where, α is the scalar to be obtained such that MSE of t_p is least.
To study the large sampling property of t_p , the following approximations are used,
$\bar{y} = \bar{Y}(1 + e_0)$, $\bar{x} = \bar{X}(1 + e_1)$, with $E(e_0) = E(e_1) = 0$ and $E(e_0^2) = \lambda C_y^2$
$E(e_1^2) = \lambda C_x^2$, $E(e_0 e_1) = \lambda C_{yx}$
Representing t_p pusing e_0 and e_1, we obtain,

$$t_p = \bar{Y}(1 + e_0)\left(\frac{\bar{X} + \beta_2/G(x)}{\bar{X}(1 + e_1) + \beta_2/G(x)}\right)^{\alpha}$$

$$= \bar{Y}(1 + e_0)\left(\frac{\bar{X} + \beta_2/G(x)}{\bar{X}e_1 + \bar{X} + \beta_2/G(x)}\right)^{\alpha}$$

$$= \bar{Y}(1 + e_0)\left(\frac{1}{\frac{\bar{X}e_1}{\bar{X}+\beta_2/G(x)} + 1}\right)^{\alpha}$$

$$= \bar{Y}(1 + e_0)\,(1 + \theta e_1)^{-\alpha}$$

where, $\theta = \frac{\bar{X}}{\bar{X}+\beta_2/G(x)}$

$$= \bar{Y}(1 + e_0)\left(1 - \alpha\theta e_1 + \frac{\alpha(1 + \alpha)}{2} e_1^2 - ...\right)$$

$$t_p = \bar{Y}\left[1 + e_0 - \alpha\theta e_1 - \alpha\theta e_0 e_1 + \frac{\alpha(1 + \alpha)}{2} e_1^2 - ...\right]$$

Subtracting $\bar{Y}$ on both sides of above equation, we have,

$$t_p - \bar{Y} = \bar{Y}\left[e_0 - \alpha\theta e_1 - \alpha\theta e_0 e_1 + \frac{\alpha(1 + \alpha)}{2} e_1^2 - ...\right] \tag{3.2}$$

We have bias of t_p as,

$$Bias(t_p) = \lambda\bar{Y}\left[\frac{\alpha(1 + \alpha)}{2} C_x^2 - \alpha\theta C_{yx}\right] \tag{3.3}$$

Squaring on both sides of (3.2) and for order one, we have,

$$MSE(t_p) = \bar{Y}^2 E\left[e_0^2 + \alpha^2\theta^2 e_1^2 - 2\alpha\theta e_0 e_1 + ...\right] \tag{3.4}$$

Putting values of different expectation, we get the MSE of t_p, as,

$$MSE(t_p) = \lambda \bar{Y}^2 \left[C_y^2 + \alpha^2\theta^2 C_x^2 - 2\alpha\theta C_{yx}\right] \tag{3.5}$$

Differentiating $MSE(t_p)$ with respect to α and putting it equal to zero, we get,

$$\frac{\partial}{\partial\alpha} MSE(t_p) = 0$$

which gives,

$$\alpha = \frac{C_{yx}}{\theta C_x^2} = \alpha_{opt} \tag{3.6}$$

Putting the value of α_{opt} in (3.5), we get the minimum value of $MSE(t_p)$ as,

$$MSE_{min}(t_p) = \lambda \bar{Y}^2 \left[C_y^2 - \frac{C_{yx}^2}{C_x^2}\right] \tag{3.7}$$

4. Efficiency Comparison

The estimator t_p is better than t_R under the condition if,

$MSE_{min}(t_p) - MSE(t_R) > 0$, or,

$\left[C_y^2 - \frac{C_{yx}^2}{C_x^2}\right] - \left[R^2 C_x^2 - 2RC_{yx}\right] > 0$

The t_p is more efficient than[3] t_1 if,

$MSE_{min}(t_p) - MSE(t_1) > 0$, or,

$\left[C_y^2 - \frac{C_{yx}^2}{C_x^2}\right] - \left[R_1^2 C_x^2 - 2R_1 C_{yx}\right] > 0$

The t_p is better than Singh et al. [5] estimator t_2 under the condition if,

$MSE_{min}(t_p) - MSE(t_2) > 0$, or,

$\left[C_y^2 - \frac{C_{yx}^2}{C_x^2}\right] - \left[R_2^2 C_x^2 - 2R_2 C_{yx}\right] > 0$

The t_p is more efficient than Upadhyaya and Singh [6] estimator t_i, $i = 3, 4$, if,

$MSE_{min}(t_p) - MSE(t_i) > 0$, $i = 3, 4$ or,

$\left[C_y^2 - \frac{C_{yx}^2}{C_x^2}\right] - \left[R_i^2 C_x^2 - 2R_i C_{yx}\right] > 0$

The t_p is better than Zaman et al. [29] estimator $t_i, i = 1, 2, 3$, under the condition if,

$MSE_{min}(t_p) - MSE(t_i) > 0,\ i = 1, 2, 3$ or,

$$\left[C_y^2 - \frac{C_{yx}^2}{C_x^2}\right] - \left[R_{5(i)}^2 C_x^2 - 2R_{5(i)} C_{yx}\right] > 0$$

5. Numerical Study

To verify the theoretical results, we have considered the following population, given in Kadilar and Cingi [8]. The parameters of the above population are presented in Table-1. The MSE and Percentage Relative Efficiency (PRE) of the estimators with respect to t_R are presented in Table-2. The graphs and MSE and PRE are presented in Figure-1 and Figure-2 respectively.

TABLE 1. Parameters of the population

$N = 80$	n=20	$\bar{Y}$=51.8264	$\bar{X} = 11.2646$
$\rho = 0.9413$	$S_y = 18.3569$	$C_x = 0.7500$	$S_x = 8.4542$
$C_y = 0.3542$	$\beta_2 = 2.866$	$\beta_1 = 1.05$	$TM = 9.318$

TABLE 2. MSE and PRE of the estimators

S. No.	Estimator	MSE	PRE
1.	t_r	2565510.488	100.00
2.	t_1	2551970.93	100.53
3.	t_2	2554353.47	100.44
4.	t_3	2551970.931	100.53
5.	t_4	2552963.345	100.49
6.	$t_{5(1)}$	2544185.217	100.84
7.	$t_{5(2)}$	2542989.056	100.89
8.	$t_{5(3)}$	2542988.851	100.89
9.	t_{opt}	**1458856.246**	**175.86**

6. Results

For better population mean estimate, we presented a generalized class of ratio estimators in this work. To a first-order approximation, the bias and MSE are obtained. Efficiency requirements are determined after a theoretical comparison of the introduced estimators' performance with that of the current estimators. A real dataset is used to validate these theoretical conclusions. It is evident from Table-2 that the MSE of the estimators in competition lie in the interval [2542988.851, 2565510.488] while the MSE of introduced estimator is 1458856.246 and the PRE of competing estimators lie in the interval [100.00, 100.89] while the PRE of the proposed estimator is 175.86. These results have also been shown in the form of graphs in Figure-1 and Figure-2 respectively.

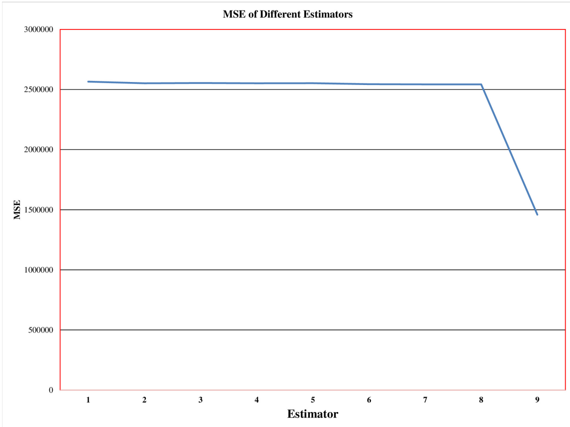

Figure 1. MSE of different estimators

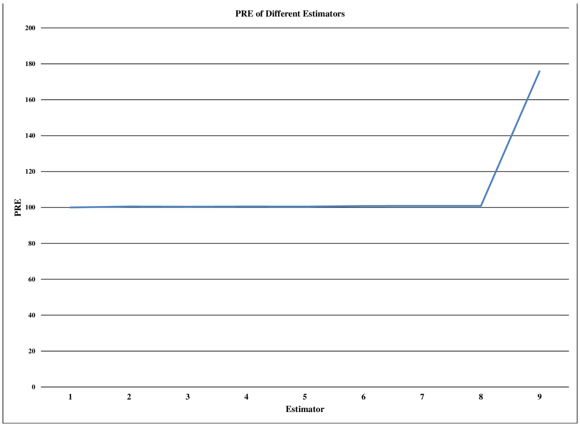

Figure 2. PRE of the estimators with respect to t_R

7. Conclusion

The suggested estimator is the most efficient of the compared estimators, as indicated in Table 2, since it obtains the lowest MSE and the highest PRE. As a result, the introduced estimator can be used successfully for precise estimation in a variety of applications. Further it is also to mention that in future direction, some more modified estimators with greater efficiency may be created and may be applied in various application areas. For instance, the average production may be estimated having information on the area, the factory average production may be estimated using units as auxiliary information, the current year average may be estimated by using previous year information etc. .

Acknowledgement

The authors are thankful to the learned referee for valuable suggestions, leading to improvement of the paper.

References

[1] Cochran, W.G., *Sampling Techniques*, Third Edition, Wiley Eastern Limited. (1940).
[2] Srivastava, S.K., *An estimator using auxiliary information in sample surveys*, Calcutta Statistical Association Bulletin, **16**, (1967), 121-132.
[3] Sisodia, B.S. and Dwivedi, V.K., *A modified ratio estimator using coefficient of variation of auxiliary variable*, Journal of Indian Society Agricultural Statistics, **33**, (1981), 13–18.
[4] Rao, T. J., *On certain methods of improving ratio and regression estimators*, Communication in Statistics Theory and Method, **20**, (1991), 3325-3340.
[5] Singh, H.P. Tailor, R., Tailor, R. and Kakran, M.S., *An improved estimator of population mean using power transformation*, Journal of the Indian society of agricultural statistics, **58(2)**, (2004), 223-230.
[6] Upadhyaya, L.N. and Singh, H.P., *Use of transformed auxiliary variable in estimating the finite population mean*, Biometrical Journal, **41 (5)**, (1999), 627–636.
[7] Singh, H.P., and Tailor. R., *Use of known correlation coefficient in estimating the finite population means*, Statistics in Transition, **6(4)**, (2003), 555–560.
[8] Kadilar, C. and Cingi, H., *Ratio estimators in simple random sampling*, Applied mathematics and computation, **151(3)**, (2004), 893-902.
[9] Singh, H.P. and Vishwakarma, G.K., *An efficient variant of the product and ratio estimators in double sampling*, Model Assisted Statistics and Applications, **1**, 3, (2005), 155-165,
[10] Singh, H.P. and Vishwakarma, G.K., *Some estimators of finite population mean using auxiliary information in sample surveys*, Journal of Applied Statistical Science, **16**, 4, (2008), 11-22.
[11] Yadav, S.K., Mishra, S.S., Shukla, A.K., Kumar and Singh, R.S., *Use of Non-Conventional Measures of Dispersion for Improved Estimation of Population Mean*, American Journal of Operational Research, **6(3)**, (2016), 69-75.
[12] Khoshnevisan. M., Singh, R., Chauhan, P., Sawan, N. and Smarandache, F.A., *general family of estimators for estimating population mean using known value of some population parameter(s)*, Far East J. Theor. Statist, **22**, (2007), 181-191.
[13] Koyuncu. N. and Kadilar. C., *Efficient estimators for the population mean*, Hacettepe Journal of Mathematics and Statistics, **38(2)**, (2009), 217-225.
[14] Yan, Z., and Tian, B., *Ratio method to the mean estimation using coefficient of skewness of auxiliary variable*, ICICA 2010, Part II, CCIS, **106**, (2010), 103–110.
[15] Subramani, J. and Kumarapandiyan, G., *Modified Ratio Estimators for Population Mean Using Function of Quartiles of Auxiliary Variable*, Bonfring International Journal of Industrial Engineering and Management Science, **2(2)**, (2012), 19-23.
[16] Jeelani, M.I., Maqbool, S. and Mir. S.A., *Modified ratio estimators of population mean using linear combination of coefficient of skewness and quartile deviation*, International Journal of Modern Mathematical Sciences, **(6)**, (2013), 174–183.
[17] Subramani, J., *Generalized Modified Ratio Estimator for Estimation of Finite Population Mean*, Journal of Modern Applied Statistical Methods, **12(2)**, (2013), 121-155.
[18] Abid, M., Abbas, N., Sherwani, R.A.K and Nazir, H.Z., *Improved Ratio Estimators for the Population Mean Using Non-Conventional Measures of Dispersion*, Pakistan Journal of Statistics and Operations Research, **12(2)**, (2016), 353-367.
[19] Subzar, M., Maqbool, S. and Raja, T.A., *A class of efficient Ratio type estimators for the Estimation of Population Mean Using the auxiliary information in survey sampling*, IJEDR, **5(2)**, (2017), 1756-1761.

[20] Unal, C. and Kadilar, C., *Improved family of estimators using exponential function for the population mean in the presence of non-response*, Communications in Statistics-Theory and Methods, **50(1)**, (2019), 237-248.

[21] Singh, G.N. and Usman, M., *Efficient combination of various estimators in the presence of non-response*, Communications in Statistics-Simulation and Computation, **50(1)**, (2019), 1-35.

[22] Priam, R., *Visualization of generalized mean estimators using auxiliary information in survey sampling*, Communications in Statistics-Theory and Methods, **49(18)**, (2019), 4468–4489.

[23] Yadav, S.K., Baghel, S., Saxena, S. and Singh, A.K., *Estimation of Average Paddy Production of Pira Nagar Village at Barabanki District in India*, Journal of Reliability and Statistical Studies, **13(1)**, (2020), 127-148.

[24] Yadav, S.K., Sharma, D.K., and Brown, K., *New Class of Estimators for Enhanced Estimation of Average Production of Peppermint Yield Utilizing Known Auxiliary Variable*, International Journal of Mathematics in Operational Research, **20(2)**, (2021), 281-295.

[25] Zaman, T., Bulut, H. and Yadav, S.K., *Robust ratio-type estimators for finite population mean in simple random sampling: A simulation study*, Concurrency and Computation: Practice and Experience, **34(25)**, (2022), e7273.

[26] Ali, M., Yadav, S.K., *Gupta, R.K., Kumar, S. and Singh, L., Enhancement of Class of Estimators for Estimating Population Mean using Known Auxiliary Parameters*, International Journal of Agricultural and Statistical Sciences, **19(2)**, (2023), 527-534.

[27] Singh, L., Malik, D., Kumar, M. and Yadav, S.K., *Refined Batch of Estimators for Estimating Population Mean with the help of known Auxiliary Parameters*, International Journal of Agricultural and Statistical Sciences, **20(1)**, (2024), 299-307.

[28] Vishwakarma, G.K., *Weighted strategy for estimation of finite population mean in presence of auxiliary information*, Measurement, Interdisciplinary Research and Perspectives, **22**, (2024), 1–11,

[29] Zaman, Q., Haider, B., Shah, Q., Anwar, M., Kaosar, S., Ali, F. and Ayub, G.A. *Novel Class of Ratio Estimators Incorporating Auxiliary Variables in Simple Random Sampling for Estimating Population Mean*, An Application to Real-Life Data, Power System Technology, **48(3)**, (2024), 609-621.

Gagan Kumar, Department of Economics, Govt. PG College Musafirkhana, Amethi 227813, India
e-mail: gkb.gdc@gmail.com

Dushyant Tyagi, Department of Economics, Rajdhani College (University of Delhi), New Delhi 110015, India
e-mail: dtyagidu@rediffmail.com

Gautam Gupta, Department of Sociology, Dr. Ambedkar Govt. PG College Unchahar, Raibareli 229406, India,
e-mail: ggspmgdc@gmail.com

Surendra Kumar, Department of Mathematics, Mahamaya Govt. Degree College Mahona, Lucknow 226203 India
e-mail: surendra.kumar776@gmail.com