# Heuristical Approach for Optimizing Population Mean Using Ratio Estimator in Stratified Random Sampling

G. R. V. Triveni and Faizan Danish*

*Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu, Beside AP Secretariat, Amaravati AP-522237, India*
*E-mail: danishstat@gmail.com*
**Corresponding Author*

Received 05 April 2023; Accepted 20 October 2023;
Publication 20 November 2023

## Abstract

In this study, we have developed a Ratio type estimator in Stratified sampling to estimate the population average of study variable by using the information of a concomitant variable. By utilizing Taylor's series, we have derived the expressions for Bias and MSE upto first degree of approximation. In numerical illustration, employing a real data set, we have demonstrated that the proposed estimator has highest Percentage relative efficiency when compared to the considered existing estimatrs. Furthermore, we have demonstrated that Separate ratio type estimators have the highest relative efficiency when contrasted with the Combined ratio type estimators.

**Keywords:** Relative effciency, Bias, mean square error, stratified sampling, Taylors series, ratio estimators.

*Journal of Reliability and Statistical Studies, Vol. 16, Issue 1 (2023),* 137–152.
doi: 10.13052/jrss0974-8024.1617

## 1 Introduction

In contemporary survey methodologies, the practice of stratification serves as a strategic tool aimed at strengthening the accuracy of estimations. In this approach, a diverse population is partitioned into distinct strata with the objective of achieving homogeneity within each stratum. Subsequently, a sample is drawn from each stratum utilizing various sampling designs. This method proves particularly effective when the selected units are not substantial in number, and the population demonstrates significant variability. The act of segmenting the population into strata offers the potential to yield more precise estimates of population parameters while concurrently reducing variance within each stratum.

Optimal stratification emerges as a technique that yields the minimal possible variance in this process. Ratio estimation constitutes a central endeavour, aiming to estimate the same ratio within the sample by utilizing the population total of a variable of interest and the population size. Ratio estimation stands out as a robust method for estimating population parameters, and its utility becomes especially apparent in situations where the population exhibits high levels of variability or when the selected units are limited in number. One of the key advantages of employing ratio estimation within the framework of stratified random sampling lies in its capacity to reduce estimator variance, thereby enhancing the precision of estimates.

The study of estimators has been supported by numerous researchers. Cochran (1971) gave a brief summary of the stratified sampling process for estimating the estimators using various sampling designs, whereas Sarjinder Singh (2003) spoke extensively about the evolution of estimators in general. A class of product cum ratio type estimators was presented by Sisodia and Dwivedi (1981). Ratio and Product type exponential estimators were shown to have higher precision by Bahl and Tuteja (1991) after Prasad (1989) worked on a class of ratio estimators. Upadhyaya and Singh (1999) provided a new approach by transforming the auxiliary variable by assuming auxiliary variable is known and they concluded that proposed estimator has lower MSE and it performs better than estimators in literature. In 2003, Kadilar and Cingi assessed a variety of ratio-type estimators before going on to create a new estimator. They expanded their study and contributed significantly to the field of survey sampling in 2005 by introducing a novel estimation method. In order to enhance several different forms of ratio estimators, Shabbir and Gupta (2005) worked on them. Later, Koyuncu and

Kadilar (2009) provided effective estimators for the population mean and demonstrated the effectiveness of their proposal by numerical illustration. Yadav et al. (2011) promoted an enhanced separate ratio exponential estimator and demonstrated the superiority of the estimator by utilizing theoretical and numerical comparisons.

Additionally, Tailor and Chouhan and Tailor and Lone (2014) researched the characteristics of distinct ratio-type estimators and worked on ratio cum product type exponential estimators. Yadav et al. (2014) developed ratio and product exponential estimators and evaluated its efficiency by using an empirical study. Later, using information from two auxiliary variables and a difference cum exponential ratio type estimator, Shabbir and Gupta (2017) were able to use stratified random sampling as well as basic random sampling. The efficiency of the novel estimators was supported by theoretical properties and numerical findings provided by Verma et al. (2017) in their proposal of a distinct class of estimators. A number of estimators in various sample designs were created by Koyuncu et al. (2018), Zaman (2019), Luengo et al. (2019), Kumar and Vishwakarma (2020), Yadav and Tailor (2020), Dahiru et al. (2021), and Khare et al. (2022). Rather and Kadilar (2022) demonstrated an increase in the proposed estimator's efficiency by contrasting it with an unbiased estimator. Additionally, Kumar et al. (2023) created a group of estimators using stratified random sampling, compared those estimators to others already in use, and came to the conclusion that the new class of estimators outperformed the older ones. In this paper, we have developed an estimator in Section 4 and proved the efficiency of proposed and considered estimators in Section 6.

## 2 Terminology

- $\mathrm{N}$ = no. of units in population
- $\mathrm{n}$ = the no. of units in the sample
- $\mathrm{P_h} = \frac{\mathrm{N_h}}{\mathrm{N}}$ is stratum weight
- k = no. of strata
- $\mathrm{N_h}$ = population units in $\mathrm{h^{th}}$ stratum
- $\mathrm{n_h}$ = sample units in $\mathrm{h^{th}}$ stratum
- $\overline{\mathrm{Z}}_1 = \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P_h}\, \overline{\mathrm{z}}_{1h}$ is the average of study variable
- $\overline{\mathrm{Z}}_2 = \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P}_h\, \overline{\mathrm{z}}_{2h}$ is the average of auxiliary variable

- $\overline{\mathrm{z}}_{1h} = \frac{1}{\mathrm{n}_h}\sum_{\mathrm{j}=1}^{\mathrm{n_h}} \mathrm{z}_{1hj}$ is the study variable's average for the sample at $\mathrm{h}^{\mathrm{th}}$ stratum
- $\overline{\mathrm{z}}_{2h} = \frac{1}{\mathrm{n}_h}\sum_{\mathrm{j}=1}^{\mathrm{n_h}} \mathrm{z}_{2hj}$is the auxiliary variable's average for the sample at $\mathrm{h}^{\mathrm{th}}$ stratum
- $\mathrm{S}^2_{\mathrm{hz_1}} = \frac{1}{\mathrm{N_h}-1}\sum_{\mathrm{j}=1}^{\mathrm{n}_h} (\mathrm{z}_{1hj} - \overline{\mathrm{Z}}_{1h})^2$ is the study variable's variance at $\mathrm{h}^{\mathrm{th}}$ stratum
- $\mathrm{S}^2_{\mathrm{hz_2}} = \frac{1}{\mathrm{N_h}-1}\sum_{\mathrm{j}=1}^{\mathrm{n}_h} (\mathrm{z}_{2hj} - \overline{\mathrm{Z}}_{2h})^2$ is the auxiliary variable's variance at $\mathrm{h}^{\mathrm{th}}$ stratum
- $\mathrm{S}_{\mathrm{hz_1z_2}} = \frac{1}{\mathrm{N}_h-1}\sum_{\mathrm{j}=1}^{\mathrm{n}_h} (\mathrm{z}_{1hj} - \overline{\mathrm{Z}}_{1h})(\mathrm{z}_{2hj} - \overline{\mathrm{Z}}_{2h})$ is the covariance between variables at $\mathrm{h}^{\mathrm{th}}$ stratum
- $\rho_{\mathrm{hz_1z_2}} = \frac{\mathrm{S_{hz_1z_2}}}{\mathrm{S_{hz_1}S_{hz_2}}}$ is correlation coefficient between variables at $\mathrm{h}^{\mathrm{th}}$ stratum
- $\mathrm{C_{hz_1}} = \frac{\mathrm{S_{hz_1}}}{\overline{\mathrm{Z}}_{1h}}$ is Coefficient of variation for study variable at $\mathrm{h}^{\mathrm{th}}$ stratum
- $\mathrm{C_{hz_2}} = \frac{\mathrm{S_{hz_2}}}{\overline{\mathrm{Z}}_{2h}}$ is Coefficient of variation for auxiliary variable at $\mathrm{h}^{\mathrm{th}}$ stratum

A population 'U' made up of units $\mathrm{U_1, U_2, U_3 \ldots U_N}$ of size N. Let $\overline{\mathrm{Z}}_1$ and $\overline{\mathrm{Z}}_2$ represents the study and concomitant variable. If population U is separated into K homogeneous strata of sizes $\mathrm{n}_h(\mathrm{h} = 1, 2, 3, \ldots \mathrm{k})$ and $\mathrm{n}_h$ represents the sample size drawn from $\mathrm{h}^{\mathrm{th}}$ stratum. The population mean $\overline{\mathrm{Z}}_1$ for Separate Ratio estimator is expressed as

$$\overline{\mathrm{Z}}_{1SR} = \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P_h}\, \overline{\mathrm{z}}_{1h} \left[\frac{\overline{\mathrm{Z}}_{2h}}{\overline{\mathrm{z}}_{2h}}\right]$$

## 3 Estimators in Literature

(1) The Traditional Separate ratio estimator is

$$\begin{aligned}\overline{\mathrm{Z}}_{1\mathrm{SR}} &= \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P}_h \overline{\mathrm{Z}}_{1hR} \\ &= \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P_h} \left[\frac{\overline{\mathrm{z}}_{1h}}{\overline{\mathrm{z}}_{2h}} \overline{\mathrm{Z}}_{2h}\right]\end{aligned}$$

Where, $\overline{Z}_{1hR}$ is the ratio estimator of $h^{th}$ stratum population mean $\overline{Z}_{1h}$ and $\frac{\overline{z}_{1h}}{\overline{Z}_{2h}}$ is the proportion of sample means of the study variable $z_1$ and auxiliary variable $z_2$ of $h^{th}$ stratum. The Bias and MSE are as follows:

$$\mathrm{Bias}(\overline{Z}_{1SR}) = \sum_{h=1}^{k} P_h \overline{Z}_{1h} \left(\frac{1-f_h}{n_h}\right) [C^2_{hz_2} - \rho_{hz_1z_2} C_{hz_1} C_{hz_2}]$$

$$\mathrm{MSE}(\overline{Z}_{1SR}) = \sum_{h=1}^{k} P_h^2 \overline{Z}_{1h}^{\,2} \left(\frac{1-f_h}{n_h}\right) \times [C^2_{hz_1} + C^2_{hz_2} - 2\rho_{hz_1z_2} C_{hz_1} C_{hz_2}] \quad (1)$$

Rajesh and Lone (2014) modified the following estimators into Separate ratio estimators as:

(2) Separate ratio type estimator utilizing $C_{z_2}$ of auxiliary variate in $h^{th}$ stratum as

$$\overline{Z}_{1RS}^{SD} = \sum_{h=1}^{k} P_h \overline{Z}_{1h} \left[\frac{\overline{z}_{2h} + C_{z_{2h}}}{\overline{z}_{2h} + C_{z_{2h}}}\right]$$

where $C_{z_{2h}}$ is the auxiliary variate's coefficient of variation in $h^{th}$ stratum

$$\mathrm{Bias}(\overline{Z}_{1RS}^{SD}) = \sum_{h=1}^{k} P_h \overline{Z}_{1h} \left(\frac{1-f_h}{n_h}\right) [\lambda_{1h}^2 C^2_{hz_2} - \rho_{hz_1z_2} \lambda_{1h} C_{hz_1} C_{hz_2}]$$

$$\mathrm{MSE}(\overline{Z}_{1RS}^{SD}) = \sum_{h=1}^{k} P_h^2 \overline{Z}_{1h}^{\,2} \left(\frac{1-f_h}{n_h}\right) \times [C^2_{hz_1} + \lambda_{1h}^2 C^2_{hz_2} - 2\rho_{hz_1z_2} \lambda_{1h} C_{hz_1} C_{hz_2}] \quad (2)$$

Where $\lambda_{1h} = \frac{\overline{Z}_{2h}}{\overline{Z}_{2h} + C_{z_{2h}}}$

(3) Separate Ratio type estimator by utilizing kurtosis $\beta_2(z_2)$ of $z_2$ in $h^{th}$ stratum as

$$\overline{Z}_{1RS}^{SE} = \sum_{h=1}^{k} P_h \overline{z}_{1h} \left[\frac{\overline{Z}_{2h} + \beta_{2h}(z_2)}{\overline{z}_{2h} + \beta_{2h}(z_2)}\right]$$

Where $\beta_2(z_2)$ is the auxiliary variate's coefficient of kurtosis in $h^{\mathrm{th}}$ stratum.

$$\mathrm{Bias}(\overline{Z}_{1\mathrm{RS}}^{\mathrm{SE}}) = \sum_{h=1}^{k} P_h \overline{Z}_{1h}\left(\frac{1-f_h}{n_h}\right)[\lambda_{2h}^2 C_{hz_2}^2 - \rho_{hz_1z_2}\lambda_{2h}C_{hz_1}C_{hz_2}]$$

$$\mathrm{MSE}(\overline{Z}_{1\mathrm{RS}}^{\mathrm{SE}}) = \sum_{h=1}^{k} P_h^2 \overline{Z}_{1h}^2\left(\frac{1-f_h}{n_h}\right) \times [C_{hz_1}^2 + \lambda_{2h}^2 C_{hz_2}^2 - 2\rho_{hz_1z_2}\lambda_{2h}C_{hz_1}C_{hz_2}] \tag{3}$$

where $\lambda_{2h} = \frac{\overline{Z}_{2h}}{\overline{z}_{2h}+\beta_{2h}(z_2)}$

(4) By utilizing coefficients of kurtosis and variation in $h^{\mathrm{th}}$ stratum, preferred a separate ratio-type estimator as

$$\overline{Z}_{1\mathrm{RS}}^{\mathrm{US1}} = \sum_{h=1}^{k} P_h \overline{Z}_{1h}\left[\frac{\overline{z}_{2h}\beta_{2h}(z_2) + C_{z_{2h}}}{\overline{z}_{2h}\beta_{2h}(z_2) + C_{z_{2h}}}\right]$$

Where $\beta_2(z_2)$ is the coefficient of kurtosis and $C_{z_{2h}}$ is the auxiliary variate's coefficient of variation in $h^{\mathrm{th}}$ stratum.

$$\mathrm{Bias}(\overline{Z}_{1\mathrm{RS}}^{\mathrm{US1}}) = \sum_{h=1}^{k} P_h \overline{Z}_{1h}\left(\frac{1-f_h}{n_h}\right)[\lambda_{3h}^2 C_{hz_2}^2 - \rho_{hz_1z_2}\lambda_{3h}C_{hz_1}C_{hz_2}]$$

$$\mathrm{MSE}(\overline{Z}_{1\mathrm{RS}}^{\mathrm{US1}}) = \sum_{h=1}^{k} P_h^2 \overline{Z}_{1h}^2\left(\frac{1-f_h}{n_h}\right) \times [C_{hz_1}^2 + \lambda_{3h}^2 C_{hz_2}^2 - 2\rho_{hz_1z_2}\lambda_{3h}C_{hz_1}C_{hz_2}] \tag{4}$$

where $\lambda_{3h} = \frac{\overline{z}_{2h}\beta_{2h}(z_2)}{\overline{z}_{2h}\beta_{2h}(z_2)+C_{z_{2h}}}$

(5) By using coefficients of variation and kurtosis in $h^{\mathrm{th}}$ stratum, preferred a separate ratio-type estimator as

$$\overline{Z}_{1\mathrm{RS}}^{\mathrm{US2}} = \sum_{h=1}^{k} P_h \overline{Z}_{1h}\left[\frac{\overline{z}_{2h}C_{z_{2h}} + \beta_{2h}(z_2)}{\overline{z}_{2h}C_{z_{2h}} + \beta_{2h}(z_2)}\right]$$

$$\mathrm{Bias}(\overline{\mathrm{Z}}_{1\mathrm{RS}}^{\mathrm{US2}}) = \sum_{\mathrm{h=1}}^{\mathrm{k}} \mathrm{P_h} \overline{\mathrm{Z}}_{1h} \left(\frac{1-\mathrm{f}_h}{\mathrm{n_h}}\right) [\lambda_{\mathrm{4h}}^2 \mathrm{C}_{\mathrm{hz_2}}^2 - \rho_{\mathrm{hz_1z_2}} \lambda_{\mathrm{4h}} \mathrm{C}_{\mathrm{hz_1}} \mathrm{C}_{\mathrm{hz_2}}]$$

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{RS}}^{\mathrm{US2}}) = \sum_{\mathrm{h=1}}^{\mathrm{k}} \mathrm{P}_h^2 \overline{\mathrm{Z}}_{1h}^{\;2} \left(\frac{1-\mathrm{f_h}}{\mathrm{n}_h}\right) \times [\mathrm{C}_{\mathrm{hz_1}}^2 + \lambda_{\mathrm{4h}}^2 \mathrm{C}_{\mathrm{hz_2}}^2 - 2\rho_{\mathrm{hz_1z_2}} \lambda_{\mathrm{4h}} \mathrm{C}_{\mathrm{hz_1}} \mathrm{C}_{\mathrm{hz_2}}] \quad (5)$$

where $\lambda_{\mathrm{4h}} = \frac{\overline{\mathrm{z}}_{2h} \mathrm{C}_{\mathrm{z_{2h}}}}{\overline{\mathrm{z}}_{2h} \mathrm{C}_{\mathrm{z_{2h}}} + \beta_{\mathrm{2h}}(\mathrm{z_2})}$

(6) The traditional Combined Ratio estimate is

$$\overline{Z}_{1_{RC}} = \frac{\overline{\mathrm{z}}_{1st}}{\overline{\mathrm{z}}_{2st}} \overline{\mathrm{Z}}_2$$

Its MSE is given by

$$\mathrm{MSE}(\overline{Z}_{1_{RC}}) = \sum_{h=1}^{k} \mathrm{P}_h^2 \left(\frac{1-\mathrm{f_h}}{\mathrm{n}_h}\right) [\mathrm{S}_{\mathrm{hz_1}}^2 + R^2 \mathrm{S}_{\mathrm{hz_2}}^2 - 2\mathrm{R} S_{hz_1z_2}] \quad (6)$$

(7) Kadilar and Cingi (2005) proposed a new Ratio estimator as

$$\overline{Z}_1^{KC} = m^* \overline{Z}_{1_{RC}}$$

$$\mathrm{MSE}(\overline{Z}_1^{KC}) = m^* \sum_{h=1}^{k} \mathrm{P}_h^2 \left(\frac{1-\mathrm{f_h}}{\mathrm{n}_h}\right) [\mathrm{S}_{\mathrm{hz_1}}^2 + R^2 \mathrm{S}_{\mathrm{hz_2}}^2 - 2\mathrm{R} S_{hz_1z_2}] + (m^*-1)^2 \overline{Z}_1^2 \quad (7)$$

Where $m^* = \frac{\overline{Z}_1^2}{\overline{Z}_1^2 + \sum_{h=1}^{k} \mathrm{P}_h^2 \left(\frac{1-\mathrm{f_h}}{\mathrm{n}_h}\right) [\mathrm{S}_{\mathrm{hz_1}}^2 + R^2 \mathrm{S}_{\mathrm{hz_2}}^2 - 2\mathrm{R} S_{hz_1z_2}]}$

## 4 Proposed Estimator

In Stratified random Sampling, we propose that

$$\overline{\mathrm{Z}}_{\mathrm{1st}} = \frac{t^2}{2} \sum_{\mathrm{h=1}}^{\mathrm{k}} \mathrm{P}_h \left[\frac{\overline{\mathrm{Z}}_{1h}}{\overline{\mathrm{z}}_{2h}} \overline{\mathrm{Z}}_{2h}\right] \quad (8)$$

Proposed Estimator's bias can be obtained as

$$\mathrm{E}(\overline{\mathrm{Z}}_{1\mathrm{st}}-\overline{\mathrm{Z}}_1)=\mathrm{E}\left(\frac{t^2}{2}\sum_{\mathrm{h}=1}^{\mathrm{k}}\mathrm{P}_h\left[\frac{\overline{\mathrm{Z}}_{1h}}{\overline{\mathrm{z}}_{2h}}\overline{\mathrm{Z}}_{2h}\right]-\overline{\mathrm{Z}}_1\right)$$

$$=\overline{\mathrm{Z}}_2\mathrm{E}\left[\frac{\frac{t^2}{2}\sum_{\mathrm{h}=1}^{\mathrm{k}}\mathrm{P}_{\mathrm{h}}\overline{\mathrm{Z}}_{1h}-\left(\frac{\overline{\mathrm{Z}}_1\overline{\mathrm{z}}_{2h}}{\overline{\mathrm{Z}}_2}\right)}{\overline{\mathrm{z}}_{2h}}\right] \tag{9}$$

We have,

$$\frac{1}{\overline{\mathrm{z}}_{2\mathrm{h}}}=\frac{1}{\overline{\mathrm{Z}}_2}\left(1+\frac{\overline{\mathrm{z}}_{2h}-\overline{\mathrm{Z}}_2}{\overline{\mathrm{Z}}_2}\right)^{-1}$$

By using Taylor Series expansion, we get

$$\frac{1}{\overline{\mathrm{z}}_{2\mathrm{h}}}\cong\frac{1}{\overline{\mathrm{Z}}_2}\left(1-\frac{\overline{\mathrm{z}}_{2h}-\overline{\mathrm{Z}}_2}{\overline{\mathrm{Z}}_2}\right) \tag{10}$$

Substituting Equation (10) in Equation (9), we get

$$\mathrm{E}(\overline{\mathrm{Z}}_{1\mathrm{st}}-\overline{\mathrm{Z}}_1)=\overline{\mathrm{Z}}_2\mathrm{E}\left\{\left(\frac{t^2}{2}\sum_{\mathrm{h}=1}^{\mathrm{k}}\mathrm{P}_{\mathrm{h}}\overline{\mathrm{Z}}_{1h}-\frac{\overline{\mathrm{Z}}_1\overline{\mathrm{z}}_{2h}}{\overline{\mathrm{Z}}_2}\right)\frac{1}{\overline{\mathrm{Z}}_2}\left(1-\frac{\overline{\mathrm{z}}_{2h}-\overline{\mathrm{Z}}_2}{\overline{\mathrm{Z}}_2}\right)\right\}$$

$$=\left[\frac{t^2}{2}-1\right]\overline{\mathrm{Z}}_1+\frac{1}{\overline{\mathrm{Z}}_2}\sum_{\mathrm{h}=1}^{\mathrm{k}}\mathrm{P}_{\mathrm{h}}\overline{\mathrm{Z}}_{1h}\left(\frac{1-\mathrm{f}_{\mathrm{h}}}{\mathrm{n}_{\mathrm{h}}}\right)$$

$$\times\left[\mathrm{C}^2_{\mathrm{hz}_2}-\mathrm{t}\rho_{\mathrm{hz}_1\mathrm{z}_2}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}\right]$$

Therefore,

$$\mathrm{Bias}(\overline{\mathrm{Z}}_{1\mathrm{st}})=(\mathrm{t}-1)\overline{\mathrm{Z}}_1+\frac{1}{\overline{\mathrm{Z}}_2}\sum_{\mathrm{h}=1}^{\mathrm{k}}\mathrm{P}_{\mathrm{h}}\overline{\mathrm{Z}}_{1h}\left(\frac{1-\mathrm{f}_{\mathrm{h}}}{\mathrm{n}_{\mathrm{h}}}\right)$$

$$\times\left[\mathrm{C}^2_{\mathrm{hz}_2}-\frac{t^2}{2}\rho_{\mathrm{hz}_1\mathrm{z}_2}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}\right]$$

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}})=\frac{t^4}{4}\mathrm{var}(\overline{\mathrm{Z}}_{1SR})+\overline{\mathrm{Z}}_1{}^2\left[\frac{t^2}{2}-1\right]^2 \tag{11}$$

Thus, the MSE of the proposed estimator will take the following form as:

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}}) = \frac{t^4}{4}\sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^2 \overline{\mathrm{Z}}_{1h}^{\,2}\left(\frac{1-\mathrm{f}_h}{\mathrm{n}_\mathrm{h}}\right) \times [\mathrm{C}_{\mathrm{hz}_1}^2 + \mathrm{C}_{\mathrm{hz}_2}^2 - 2\rho_{\mathrm{hz}_1\mathrm{z}_2}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}] + \overline{\mathrm{Z}}_1^{\,2}\left[\frac{t^2}{2}-1\right]^2 \quad (12)$$

For minimizing MSE differentiate above equation with respect to 't' and equating to zero, we get

$$\frac{\mathrm{dMSE}(\overline{\mathrm{Z}}_{1\mathrm{st}})}{\mathrm{dt}} = 0$$

$$t^2\left\{\sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^2 \overline{\mathrm{Z}}_{1h}^{\,2}\left(\frac{1-\mathrm{f}_\mathrm{h}}{\mathrm{n}_\mathrm{h}}\right)[\mathrm{C}_{\mathrm{hz}_1}^2 + \mathrm{C}_{\mathrm{hz}_2}^2 - 2\rho_{\mathrm{hz}_1\mathrm{z}_2}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}] + \overline{\mathrm{Z}}_1^{\,2}\right\} = 2\overline{\mathrm{Z}}_1^{\,2}$$

We obtained "$t^2$" from the above equation as follows:

$$t^2 = \frac{2\overline{\mathrm{Z}}_1^{\,2}}{\overline{Z}_1^2 + \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^2 \overline{\mathrm{Z}}_{1h}^{\,2}\left(\frac{1-\mathrm{f}_\mathrm{h}}{\mathrm{n}_h}\right)[\mathrm{C}_{\mathrm{hz}_1}^2 + \mathrm{C}_{\mathrm{hz}_2}^2 - 2\rho_{\mathrm{hz}_1\mathrm{z}_2}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}]}$$

where, t lies between 0 and 1.

## 5 Empirical Studies

In the empirical study, we compare our proposed estimator with traditional Separate ratio estimator, Combined Ratio estimator and the other estimators listed below. Additionally, we examine various conditions as outlined below:

From Equations (1) and (12), we get

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}}) < \mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{SR}})$$

$$\overline{\mathrm{Z}}_1^{\,2}\left[\frac{t^2}{2}-1\right]^2 < \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^2 \overline{\mathrm{Z}}_{1h}^{\,2}\left(\frac{1-\mathrm{f}_h}{\mathrm{n}_\mathrm{h}}\right) \times [\mathrm{C}_{\mathrm{hz}_1}^2 + \mathrm{C}_{\mathrm{hz}_2}^2 - 2\rho_{\mathrm{hz}_1\mathrm{z}_2}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}]\left(1-\frac{t^4}{4}\right) \quad (13)$$

From (2) and (12), we have the condition as

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}}) < \mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{RS}}^{\ \mathrm{SD}})$$

$$\overline{\mathrm{Z}}_1^{\ 2}\left[\frac{t^2}{2}-1\right]^2 < \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^2 \overline{\mathrm{Z}}_{1h}^{\ 2}\left(\frac{1-\mathrm{f}_h}{\mathrm{n}_{\mathrm{h}}}\right)$$
$$\left\{[\mathrm{C}_{h\mathrm{z}_1}^2+\lambda_{1\mathrm{h}}^2\mathrm{C}_{\mathrm{hz}_2}^2-2\rho_{\mathrm{hz}_1\mathrm{z}_2}\lambda_{1\mathrm{h}}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}]\right.$$
$$\left.-\frac{t^4}{4}[\mathrm{C}_{\mathrm{hz}_1}^2+\mathrm{C}_{\mathrm{hz}_2}^2-2\rho_{\mathrm{hz}_1\mathrm{z}_2}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}]\right\} \quad (14)$$

From (3) and (12),

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}}) < \mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{RS}}^{\ \mathrm{SE}})$$

$$\overline{\mathrm{Z}}_1^{\ 2}\left[\frac{t^2}{2}-1\right]^2 < \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^2 \overline{\mathrm{Z}}_{1h}^{\ 2}\left(\frac{1-\mathrm{f}_h}{\mathrm{n}_{\mathrm{h}}}\right)$$
$$\times\left\{[\mathrm{C}_{h\mathrm{z}_1}^2+\lambda_{2\mathrm{h}}^2\mathrm{C}_{\mathrm{hz}_2}^2-2\rho_{h\mathrm{z}_1\mathrm{z}_2}\lambda_{2\mathrm{h}}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}]\right.$$
$$\left.-\frac{t^4}{4}[\mathrm{C}_{\mathrm{hz}_1}^2+\mathrm{C}_{\mathrm{hz}_2}^2-2\rho_{\mathrm{hz}_1\mathrm{z}_2}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}]\right\} \quad (15)$$

From Equations (4) and (12)

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}}) < \mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{RS}}^{\ \mathrm{US1}})$$

$$\overline{\mathrm{Z}}_1^{\ 2}\left[\frac{t^2}{2}-1\right]^2 < \sum_{\mathrm{h}=1}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^2 \overline{\mathrm{Z}}_{1h}^{\ 2}\left(\frac{1-\mathrm{f}_h}{\mathrm{n}_{\mathrm{h}}}\right)$$
$$\times\left\{[\mathrm{C}_{\mathrm{hz}_1}^2+\lambda_{3\mathrm{h}}^2\mathrm{C}_{\mathrm{hz}_2}^2-2\rho_{h\mathrm{z}_1\mathrm{z}_2}\lambda_{3\mathrm{h}}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{h\mathrm{z}_2}]\right.$$
$$\left.-\frac{t^4}{4}[\mathrm{C}_{\mathrm{hz}_1}^2+\mathrm{C}_{\mathrm{hz}_2}^2-2\rho_{\mathrm{hz}_1\mathrm{z}_2}\mathrm{C}_{\mathrm{hz}_1}\mathrm{C}_{\mathrm{hz}_2}]\right\} \quad (16)$$

From (5) and (12), we have condition as

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}}) < \mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{RS}}^{\mathrm{US2}})$$

$$\begin{aligned}\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}}) &= \sum_{\mathrm{h=1}}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^{2} \overline{\mathrm{Z}}_{1h}^{\,2} \left(\frac{1-\mathrm{f}_h}{\mathrm{n_h}}\right) \\ &\quad \times \Big\{ [\mathrm{C}_{\mathrm{hz}_1}^2 + \lambda_{4\mathrm{h}}^2 \mathrm{C}_{\mathrm{hz}_2}^2 - 2\rho_{\mathrm{hz}_1\mathrm{z}_2} \lambda_{4\mathrm{h}} \mathrm{C}_{\mathrm{hz}_1} \mathrm{C}_{\mathrm{hz}_2}] \\ &\quad - \frac{t^4}{4} [\mathrm{C}_{\mathrm{hz}_1}^2 + \mathrm{C}_{\mathrm{hz}_2}^2 - 2\rho_{\mathrm{hz}_1\mathrm{z}_2} \mathrm{C}_{\mathrm{hz}_1} \mathrm{C}_{\mathrm{hz}_2}] \Big\} \end{aligned} \tag{17}$$

From (6) and (12), we have condition as

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}}) < \mathrm{MSE}(\overline{Z}_{1_{RC}})$$

$$\begin{aligned} &\frac{t^4}{4} \sum_{\mathrm{h=1}}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^{2} \overline{\mathrm{Z}}_{1h}^{\,2} \left(\frac{1-\mathrm{f}_h}{\mathrm{n_h}}\right) [\mathrm{C}_{\mathrm{hz}_1}^2 + \mathrm{C}_{\mathrm{hz}_2}^2 - 2\rho_{\mathrm{hz}_1\mathrm{z}_2} \mathrm{C}_{\mathrm{hz}_1} \mathrm{C}_{\mathrm{hz}_2}] \\ &+ \overline{\mathrm{Z}}_1^{\,2} \left[\frac{t^2}{2} - 1\right]^2 \\ &< \sum_{h=1}^{k} \mathrm{P}_h^2 \left(\frac{1-\mathrm{f_h}}{\mathrm{n}_h}\right) [\mathrm{S}_{\mathrm{hz}_1}^2 + R^2 \mathrm{S}_{\mathrm{hz}_2}^2 - 2\mathrm{R} S_{hz_1z_2}] \end{aligned} \tag{18}$$

From Equations (7) and (12), we have

$$\mathrm{MSE}(\overline{\mathrm{Z}}_{1\mathrm{st}}) < \mathrm{MSE}(\overline{Z}_1^{KC})$$

$$\begin{aligned} &\frac{t^4}{4} \sum_{\mathrm{h=1}}^{\mathrm{k}} \mathrm{P}_{\mathrm{h}}^{2} \overline{\mathrm{Z}}_{1h}^{\,2} \left(\frac{1-\mathrm{f}_h}{\mathrm{n_h}}\right) [\mathrm{C}_{\mathrm{hz}_1}^2 + \mathrm{C}_{\mathrm{hz}_2}^2 - 2\rho_{\mathrm{hz}_1\mathrm{z}_2} \mathrm{C}_{\mathrm{hz}_1} \mathrm{C}_{\mathrm{hz}_2}] \\ &+ \overline{\mathrm{Z}}_1^{\,2} \left[\frac{t^2}{2} - 1\right]^2 \\ &< m^* \sum_{h=1}^{k} \mathrm{P}_h^2 \left(\frac{1-\mathrm{f_h}}{\mathrm{n}_h}\right) [\mathrm{S}_{\mathrm{hz}_1}^2 + R^2 \mathrm{S}_{\mathrm{hz}_2}^2 - 2\mathrm{R} S_{hz_1z_2}] \\ &+ (m^* - 1)^2 \overline{Z}_1^2 \end{aligned} \tag{19}$$

## 6 Application of Real-Life Data Set

We used ratio estimators to analyze data from the number of apple trees and the amount of apple output in 854 villages throughout Turkey in 1999 (Source: Institute of Statistics, Republic of Turkey). First, we divided the data into six strata based on the various areas of Turkey, and then we randomly chose samples (villages) from each stratum (region). Neyman allocation (Cochran, 1977) is used

$$n_h = n\frac{N_h S_h}{\sum_{h=1}^{k} N_h S_h}$$

and a total of 140 units selected as sample from a population of 854, the data represented in Table 1.

**Table 1** Data statistics

| | | | | | |
|---|---|---|---|---|---|
| $N_1 = 106$ | $N_2 = 106$ | $N_3 = 94$ | $N_4 = 171$ | $N_5 = 204$ | $N_6 = 173$ |
| $n_1 = 9$ | $n_2 = 17$ | $n_3 = 38$ | $n_4 = 67$ | $n_5 = 7$ | $n_6 = 2$ |
| $\overline{Z}_{11} = 1536$ | $\overline{Z}_{12} = 2212$ | $\overline{Z}_{13} = 9384$ | $\overline{Z}_{14} = 5588$ | $\overline{Z}_{15} = 967$ | $\overline{Z}_{16} = 404$ |
| $\overline{Z}_{21} = 24375$ | $\overline{Z}_{22} = 27421$ | $\overline{Z}_{23} = 72409$ | $\overline{Z}_{24} = 74365$ | $\overline{Z}_{25} = 26441$ | $\overline{Z}_{26} = 9844$ |
| $\beta_{2z_{11}} = 25.7$ | $\beta_{2z_{12}} = 34.6$ | $\beta_{2z_{13}} = 26.1$ | $\beta_{2z_{14}} = 97.6$ | $\beta_{2z_{15}} = 27.5$ | $\beta_{2z_{16}} = 28.1$ |
| $C_{1z_1} = 4.18$ | $C_{2z_1} = 5.22$ | $C_{3z_1} = 3.19$ | $C_{4z_1} = 5.13$ | $C_{5z_1} = 2.47$ | $C_{6z_1} = 2.34$ |
| $C_{1z_2} = 2.02$ | $C_{2z_2} = 2.1$ | $C_{3z_2} = 2.22$ | $C_{4z_2} = 3.84$ | $C_{5z_2} = 1.72$ | $C_{6z_2} = 1.91$ |
| $S_{1z_1} = 6425$ | $S_{2z_1} = 11552$ | $S_{3z_1} = 29907$ | $S_{4z_1} = 28643$ | $S_{5z_1} = 2390$ | $S_{6z_1} = 946$ |
| $S_{1z_2} = 49189$ | $S_{2z_2} = 57461$ | $S_{3z_2} = 160757$ | $S_{4z_2} = 285603$ | $S_{5z_2} = 45403$ | $S_{6z_2} = 18794$ |
| $\rho_{1z_1z_2} = 0.82$ | $\rho_{2z_1z_2} = 0.86$ | $\rho_{3z_1z_2} = 0.9$ | $\rho_{4z_1z_2} = 0.99$ | $\rho_{5z_1z_2} = .71$ | $\rho_{6z_1z_2} = 0.89$ |
| N = 854 | n = 140 | $\overline{Z}_1 = 2930$ | $\overline{Z}_2 = 37600$ | $S_{z_1} = 17106$ | $S_{z_2} = 144794$ |

The suggested estimator outperforms the separate ratio, combined ratio, and taken into account existing estimators in terms of percentage relative efficiency (PRE). Table 2 shows that the proposed estimator's MSE value is significantly lower than that of the other estimators. Further, It can be concluded that, in Table 2, the estimators 6 & 7 are combined ratio type estimators. It is observed that, they have highest PRE values than others.

**Table 2** MSE values for the considered and proposed estimators

| S. No. | Estimator | MSE | PRE |
|---|---|---|---|
| **1.** | $\overline{Z}_{1SR}$ | 165108.3 | 100 |
| **2.** | $\overline{Z}_{1RS}^{SD}$ | 165112.2 | 99.99 |
| **3.** | $\overline{Z}_{1RS}^{SE}$ | 165335.7 | 99.86 |
| **4.** | $\overline{Z}_{1RS}^{US1}$ | 165100.4 | 100.00 |

(*Continued*)

**Table 2** Continued

| S. No. | Estimator | MSE | PRE |
|---|---|---|---|
| **5.** | $\overline{\mathrm{Z}}_{1\mathrm{RS}}^{\mathrm{US2}}$ | 165182.1 | 99.95 |
| **6.** | $\overline{Z}_{1_{RC}}$ | 225224.1 | 73.31 |
| **7.** | $\overline{Z}_1^{KC}$ | 219466.4 | 75.23 |
| **8.** | $\overline{\mathrm{Z}}_{1\mathrm{st}}$ | 161992.8 | 101.92 |

## 7 Conclusion

In this research, we have introduced a novel separate-type ratio estimator designed for application within stratified random sampling. Our investigation involved a comprehensive comparative analysis, contrasting this newly developed estimator with traditional separate ratio estimators, combined ratio estimators, and certain existing estimators that were previously modified by Rajesh and Lone in 2014. Utilizing a real-world dataset, we empirically validated the efficiency and effectiveness of our proposed estimator.

The findings, as presented in Table 2, prominently demonstrate the superiority of the proposed estimator. Specifically, our estimator consistently yielded lower Mean Squared Error (MSE) values, affirming its ability in achieving higher precision in estimation. Moreover, these results unequivocally underscore the dominance of separate ratio estimators over their combined ratio counterparts, reinforcing the significance of this method in optimizing accuracy within the realm of stratified random sampling.

## References

Bahl, S., and Tuteja, R. (1991). Ratio and product type exponential estimators. *Journal of information and optimization sciences*, 12(1), 159–164.

Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons.

Haruna, N. F., Kumar Singh, R. V., and Dahiru, S. (2021). A modified ratio type estimator of finite population mean under Stratified random sampling scheme. *Indian Journal of Applied Research*, 11, 58–60.

Kadilar, C., and Cingi, H. (2003). Ratio estimators in stratified random sampling. *Biometrical journal*, 45(2), 218–225.

Kadilar, C., and Cingi, H. (2005). A new ratio estimator in stratified random sampling. *Communications in Statistics - Theory and Methods*, 34(3), 597–602.

Kumar, M., and Vishwakarma, G. K. (2020). Generalized classes of regression-cum-ratio estimators of population mean in stratified random sampling. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, *90*, 933–939.

Lone, H. A., Tailor, R., and Verma, M. R. (2017). Efficient separate class of estimators of population mean in stratified random sampling. *Communications in Statistics – Theory and Methods*, 46(2), 554–573.

Mishra, M., Singh, S., and Khare, B. B. (2022). Separate Ratio Estimator Using Calibration Approach for the Population mean in Stratified Random Sampling. *Asian Journal of Probability and Statistics*, 20(3), 64–73.

Prasad, B. (1989). Some improved ratio type estimators of population mean and ratio in finite population sample surveys. *Communications in Statistics-Theory and Methods*, 18(1), 379–392.

Rather, K. U. I., and Kadilar, C. (2022). Dual to Ratio cum Product Type of Exponential Estimator for Population Mean in Stratified Random Sampling. *Journal of Statistics Applications & Probability Letters*, 9(1), 43–48.

Sisodia, B. V. S., and Dwivedi, V. K. (1981). Modified ratio estimator using coefficient of variation of auxiliary variable. *Journal-Indian Society of Agricultural Statistics*, 33(2), 13–18.

Singh, H. P., and Kakran, M. S. (1993). A modified ratio estimator using known coefficient of kurtosis of an auxiliary character. *Journal of the Indian Society of Agricultural Statistics*, 45(2), 141–148.

Shabbir, J., and Gupta, S. (2005). Improved ratio estimators in stratified sampling. *American Journal of Mathematical and Management Sciences*, 25(3–4), 293–311.

Shabbir, J., and Gupta, S. (2017). Estimation of finite population mean in simple and stratified random sampling using two auxiliary variables. *Communications in Statistics-Theory and Methods*, 46(20), 10135–10148.

Shahzad, U., Hanif, M., and Koyuncu, N. (2018). A new estimator for mean under stratified random sampling. *Mathematical Sciences*, 12, 163–169.

Shahzad, U., Hanif, M., Koyuncu, N., and Luengo, A. G. (2019). A family of ratio estimators in stratified random sampling utilizing auxiliary attribute alongside the nonresponse issue. *Journal of Statistical Theory and Applications*, 18(1), 12–25.

Tailor, R., and Chouhan, S. (2014). Ratio-cum-product type exponential estimator of finite population mean in stratified random sampling. *Communications in Statistics-Theory and Methods*, 43(2), 343–354.

Tailor, R., and Lone, H. A. (2014). Separate ratio-type estimators of population mean in stratified random sampling. *Journal of Modern Applied Statistical Methods*, 13(1), 14.

Tiwari, K. K., Bhougal, S., and Kumar, S. (2023). A general class of estimators in stratified random sampling. *Communications in Statistics-Simulation and Computation*, 52(2), 442–452.

Upadhyaya, L. N., and Singh, H. P. (1999). Use of transformed auxiliary variable in estimating the finite population mean. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 41(5), 627–636.

United States. National Agricultural Statistics Service. (1999). 1997 Census of Agriculture. US Department of Agriculture, National Agricultural Statistics Service.

Yadav, R., Upadhyaya, L. N., Singh, H. P., and Chatterjee, S. (2011). Improved separate ratio exponential estimator for population mean using auxiliary information. *Statistics in Transition new series*, 12(2), 401–412.

Yadav, R., and Tailor, R. (2020). Estimation of finite population mean using two auxiliary variables under stratified random sampling. *Statistics in Transition new series*, 21(1), 1–12.

Yadav, R., Upadhyaya, L. N., Singh, H. P., and Chatterjee, S. (2014). Improved ratio and product exponential type estimators for finite population mean in stratified random sampling. *Communications in Statistics-Theory and Methods*, 43(15), 3269–3285.

Zaman, T. (2019). Efficient estimators of population mean using auxiliary attribute in stratified random sampling. *Advances and Applications in Statistics*, 56(2), 153–171.

## Biographies

**G. R. V. Triveni** currently a Ph.D Research Scholar in the Department of Mathematics, School of Advanced Sciences, VIT-AP University, Amaravati, Andhra Pradesh, India. Her Research area is Sampling theory. She had done M.Sc from Andhra University, Andhra Pradesh. She possessed a limited knowledge of R software.

**Faizan Danish** is currently working as an Assistant Professor (Statistics) in VIT-AP University, Vijayawada Andhra Pradesh, India and has done Ph.D. from Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, J&K, India. He has worked as Postdoctoral Fellow in New York University, New York, USA. Dr. Danish has around 3 years of Teaching experience and published around 25 research papers in reputed journals. Further, he has worked as Biostatistician under Research Consultation Services, Doha Qatar. He has research expertise in Sampling Theory, Mathematical, Applied Statistics and Biostatistics. Dr. Danish is well versed with Statistical Software's: R, STATA, SPSS, Fortran 77, O.P STAT, WINDOSTAT, Mathematica.