Kurdish Studies
September 2024
Volume: 12, No: 5, pp 1312-1319
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.53555/ks.v12i5.3474

Comprehensive Class Of Estimators For Estimation Of Population Mean Under Stratified Sampling: Application With Real Data Sets And Simulation Analysis

Sajid Khan[1*], Muhammad Farooq[1], Sardar Hussain[2],Sohaib Ahmad[3], Muhammad Atif[4], Muhammad Ilyas[4]

[1]Department of Statistics, University of Peshawer, Peshawar, Pakistan
[2]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan
[3]Department of Statistics, Abdul Wali Khan University, Mardan, Pakistan
[4]Department of Statistics, University of Malakand, Chakdara, Pakistan
E-mails:sajidkhan2022@uop.edu.pk,m.farooq@uop.edu.pk,shussain@stat.qau.edu.pk,sohaib_ahmad@awkum.edu.pk, m.atif@uop.eu.pk

***Corresponding author:** Sajid Khan
*Department of Statistics, University of Peshawer, Peshawar, Pakistan

Abstract
Survey sampling focuses significantly on auxiliary information to provide precise parameter estimates (mean, variance, distribution function, etc.) in order ensure the best possible result. To find the mean value of an investigated variable in the population, this study makes use of additional information. The primary objective of this research is to create an improved estimator of a finite population's mean by using information from an auxiliary variable in stratified random sampling. Up to the first order of approximation, the bias and mean square error (MSE) expressions of the suggested estimator are inferred. We show that these proposed estimators perform well during the estimation process by doing a thorough analysis utilizing criteria like percentage relative efficiency and mean square error. In addition, we conduct a thorough simulation study to demonstrate that our suggested estimate exceeds other estimators that have been addressed in the literature, including conventional unbiased estimators and traditional regression estimators. Our proposed estimator appears to be the best option when compared to numerous other methods. In addition to making significant advancements in the area of survey sampling methods, the study findings give crucial information for predictive modeling on real data sets.

Keywords: Auxiliary varaible, bias, MSE, PRE, stratified random sampling.

Mathematical subject classification: 62D05

1.Introduction

In sampling theory, improving the accuracy of population mean estimation has been a subject of discussion. Despite the variety of estimators that have been offered for improving estimation of the population mean in stratified random sampling, a more exact mean estimate is still required. Sampling is an easy and reliable way to get statistically significant data from huge populations, which can help with making decisions based on characteristics shared by the whole. Conclusions can be drawn from the sample since it provides the most accurate representation of the population at large. Developing an estimator or estimators with auxiliary data centered on real population characteristics and strongly correlated with the study variable is the aim of this research. Estimators like these find usage in a wide range of fields, including sociology, marketing, agriculture, economics, industry, and medical. Various sampling processes can make use of auxiliary variables. Some examples of these methods include stratified sampling, cluster sampling, simple random sampling, multi-stage sampling, and two-stage sampling. Using an auxiliary variable that correlates positively with the study variable enhances the effectiveness of the ratio type estimators, as has been well-established in the literature for some time. In order to get more representative samples from various populations using appropriate sampling methods, the idea of stratification is used. Stratified sampling allows us to divide the total population into several groups, all of which will remain very similar to each other. It is common practice to sample each stratum using Simple Random Sampling.

In order to conduct surveys, the stratified random sample technique's mean estimate approach is necessary. It comprises dividing the population into similar groups, or strata, defined by shared characteristics. In order to provide reliable estimates of population metrics, this technique is employed. Each stratum is sampled separately using random sampling techniques. When compared to the traditional random sampling method, this one provides more accurate findings with a less unpredictable sample procedure. With stratified sampling, estimates are improved because each stratum of the population is considered for their unique characteristics. To achieve this, samples are collected with a major emphasis on each stratum. Think of this made-up city as having a middle class, a rich class, and a destitute class with different levels of income. Our understanding of the subject will be enhanced by this. In order to get a better look at the average family's income, statisticians use stratification and sample allocation procedures after the population is sorted into income-based categories.

Researchers and policymakers might perhaps make better use of their resources and come up with more precise estimations if they follow this approach. Knowing one's demographic categories is crucial in situations like these. With auxiliary data playing such a pivotal role in survey sampling, this work aims to investigate its potential applications in enhancing population parameter estimations. We aim to fill this knowledge gap by using a stratified random sample framework with auxiliary data to improve the efficiency and accuracy of the estimate. The accuracy of the mean estimate can be enhanced with the help of auxiliary data when dealing with complex sample schemes. When estimating population parameters, auxiliary information refers to any extra data that can make estimators more accurate and efficient. Researchers like [7] were among the first to apply ratio estimators and make use of auxiliary data. Stratified random sample with auxiliary data for population mean estimation is discussed in several notable publications, such as [1-6], [9], and [13-26].

The rest of the paper follows this format: Section 2 consists of notations and symbols. In Section 3, we discuss existing estimators in the context of stratified random sampling. Section 4 presents the estimators that are suggested. Numerical investigations are covered in Section 5. The simulation study is included in Section 6. A detailed analysis of the numerical outcomes is presented in Section 7. The paper concludes with its final conclusions in Section 8.

2. Methodology

Consider a finite population of distinct and identifiable units, $\mathcal{L} = \{\S_1, \S_2, \dots, \S_N\}$ of size N, are separated into L homogeneous strata of size N_h(h=1, 2,…, L), such that $\sum_{h=1}^{L} N_h = N$. A simple random sample of size n_h is drawn without replacement from the h^{th} stratum such that $\sum_{h=1}^{L} n_h = n$. Let Y and X be the study and auxiliary variables respectively, assuming values y_{h_i} and x_{h_i} for the i^{th} unit in the h^{th} stratum.

$\bar{Y}_h = \sum_{i=1}^{N_h} \frac{y_{h_i}}{N_h}$ and $\bar{X}_h = \sum_{i=1}^{N_h} \frac{x_{h_i}}{N_h}$

$\bar{y}_{st} = \sum_{h=1}^{L} w_h \bar{y}_h$ and $\bar{x}_{st} = \sum_{h=1}^{L} w_h \bar{x}_h$

$\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$ and $\bar{x}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{hi}$

$w_h = \frac{N_h}{N}$

$\bar{Y} = \sum_{h=1}^{L} w_h \bar{Y}_h$ and $\bar{X} = \sum_{h=1}^{L} w_h \bar{X}_h$

Let $\varepsilon_{0st} = \frac{\bar{y}_{st} - \bar{Y}}{\bar{Y}}$, $\varepsilon_{1st} = \frac{\bar{x}_{st} - \bar{X}}{\bar{X}}$,

$E(\varepsilon_i) = 0$, for $(i = 0,1)$,

$\omega_{rs} = \sum_{h=1}^{L} w_h^{r+s} \frac{E[(\bar{y}_h - \bar{Y})^r (\bar{x}_h - \bar{X})^s]}{\bar{Y}^r \bar{X}^s}$

$E(\varepsilon_{0st}^2) = \frac{\sum_{h=1}^{l} w_h^2 \lambda_h C_{yh}^2}{\bar{Y}^2} = \omega_{200}$, $E(\varepsilon_{1st}^2) = \frac{\sum_{h=1}^{l} w_h^2 \lambda_h C_{xh}^2}{\bar{X}^2} = \omega_{020}$, $E(\varepsilon_{0st}\varepsilon_{1st}) = \frac{\sum_{h=1}^{l} w_h^2 \lambda_h S_{yxh}}{\bar{Y}\bar{X}} = \omega_{110}$,

$\lambda_h = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)$.

Let $C_{yh} = \frac{S_{yh}}{\bar{Y}}$ and $C_{xh} = \frac{S_{xh}}{\bar{X}}$,

$s_{yh}^2 = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{y}_h)^2}{N_h - 1}$ and $s_{xh}^2 = \frac{\sum_{i=1}^{N_h} (x_{hi} - \bar{x}_h)^2}{N_h - 1}$, $s_{yh} = \sqrt{\frac{\sum_{i=1}^{N} (y_{hi} - \bar{y}_h)^2}{N_h - 1}}$ and $s_{xh} = \sqrt{\frac{\sum_{i=1}^{N_h} (x_{hi} - \bar{x}_h)^2}{N_h - 1}}$,

$s_{yxh} = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{y}_h)(x_{hi} - \bar{x}_h)}{N_h - 1}$.

3. Literature review

This section include some adopted existing estimators under stratified random sampling, which are given by:

(i) The usual estimator is:

$$\hat{\bar{Y}}_{(ST)} = \sum_{h=1}^{L} w_h \bar{y}_h.$$

The variance of $\hat{\bar{Y}}_{st}$, is given by

$$\text{Var}(\hat{\bar{Y}}_{(ST)}) = \bar{Y}^2 \omega_2 \quad (1)$$

(ii) [7] suggested the ratio estimator:

$$\hat{\bar{Y}}_{R(ST)} = \bar{y}_{(ST)} \left(\frac{\bar{X}}{\bar{x}_{(ST)}}\right) \quad (2)$$

The bias and MSE of $\hat{\bar{Y}}_R$, are given by:

$$\text{Bias}(\hat{\bar{Y}}_{R(ST)}) = \bar{Y}(\omega_{020} - \omega_{110}),$$

and

$$\text{MSE}\left(\hat{\bar{Y}}_{R(ST)}\right) \cong \bar{Y}^2(\omega_{200} + \omega_{020} - 2\omega_{110}) \quad (3)$$

(iii) The usual product estimator $\hat{\bar{Y}}_{P(ST)}$, which is given by:

$$\hat{\bar{Y}}_{P(ST)} = \bar{y}_{(ST)} \left(\frac{\bar{x}_{(ST)}}{\bar{X}}\right) \quad (4)$$

The bias and MSE of $\hat{\bar{Y}}_{P(ST)}$, are given by:

$$\text{Bias}(\hat{\bar{Y}}_{P(ST)}) = \bar{Y}\omega_{11},$$

and

$$\text{MSE}\left(\hat{\bar{Y}}_{P(ST)}\right) \cong \bar{Y}^2(\omega_{200} + \omega_{020} + 2\omega_{110}) \quad (5)$$

(iv) The difference estimator $\hat{\bar{Y}}_{dif}$, is given by:

$$\hat{\bar{Y}}_{dif(ST)} = \bar{y}_{(ST)} + \text{d}\left(\bar{X} - \bar{x}_{(ST)}\right), \quad (6)$$

$$d_{opt} = \frac{\bar{Y}\omega_{110}}{\bar{X}\omega_{020}},$$

$$\text{Var}(\hat{\bar{Y}}_{dif(ST)})_{min} = \frac{\bar{Y}^2(\omega_{200}\omega_{020} - \omega_{110})}{\omega_{020}}, \quad (7)$$

(v) [10] proposed the following estimator:

$$\hat{\bar{Y}}_{R,D(ST)} = Q_1\,\bar{y}_{(ST)} + Q_2\left(\bar{X} - \bar{x}_{(ST)}\right), \quad (8)$$

where Q_1 and Q_2 are constants.

The bias and MSE of $\hat{\bar{Y}}_{R,D(ST)}$, given by:

$$\text{Bias}\left(\hat{\bar{Y}}_{R,D(ST)}\right) = \bar{Y}\,(Q_1 - 1),$$

and

$$\text{MSE}\left(\hat{\bar{Y}}_{R,D(ST)}\right) = \bar{Y}^2 - 2\,Q_1\bar{Y}^2 + Q_1^2\bar{Y}^2 + Q_1^2\bar{Y}^2\omega_{200} - 2Q_1Q_2\,\bar{Y}\bar{X}\,\omega_{110} + Q_2^2\bar{X}^2\omega_{020}.$$

The optimum values of Q_1 and Q_2 are given by:

$$Q_{1opt} = \frac{\omega_{020}}{(\omega_{020}\omega_{200} - \omega_{110}^2 + \omega_{020})},$$

and

$$Q_{2opt} = \frac{\bar{Y}\omega_{110}}{\bar{X}(\omega_{200}\omega_{020} - \omega_{110}^2 + \omega_{020})}.$$

The minimum MSE of $\hat{\bar{Y}}_{R,D(ST)}$ is given by:

$$\text{MSE}\left(\hat{\bar{Y}}_{R,D(ST)}\right)_{min} = \frac{\bar{Y}^2(\omega_{200}\omega_{020} - \omega_{101}^2)}{(\omega_{200}\omega_{020} - \omega_{110}^2 + \omega_{020})}. \quad (9)$$

(vi) [11] suggested the following exponential type estimator, which is given by:

$$\hat{\bar{Y}}_{Singh(ST)} = \bar{y}_{(ST)}\exp\left(\frac{a(\bar{X} - \bar{x}_{(ST)})}{a(\bar{X} + \bar{x}_{(ST)}) + 2b}\right), \quad (10)$$

The bias and MSE of $\hat{\bar{Y}}_{Singh}$, is given by:

$$\text{Bias}\left(\hat{\bar{Y}}_{Singh(ST)}\right) = \bar{Y}\left(\frac{3}{8}\,\theta^2\omega_{020} - \frac{1}{2}\theta\omega_{110}\right),$$

and

$$\text{MSE}\left(\hat{\bar{Y}}_{Singh(ST)}\right) \cong \frac{\bar{Y}^2}{4}\,(4\omega_{200} + \theta^2\omega_{020} - 4\theta\omega_{110}). \quad (11)$$

(vii) The suggested estimator of [12], is given as:

$$\hat{\bar{Y}}_{Gk(ST)} = \left\{Z_1\bar{y}_{(ST)} + Z_2\left(\bar{X} - \bar{x}_{(ST)}\right)\right\}\exp\left(\frac{a(\bar{X} - \bar{x}_{(ST)})}{a(\bar{X} + \bar{x}_{(ST)}) + 2b}\right). \quad (12)$$

The bias and MSE of $\hat{\bar{Y}}_{Gk(ST)}$ are given by:

$$\text{Bias}\left(\hat{\bar{Y}}_{Gk(ST)}\right) = \bar{Y}(Z_1 - 1) + \frac{3}{8}\,\theta^2Z_1\bar{Y} + \frac{\theta}{2}\,Z_2\bar{X}\,\omega_{020} - \frac{\theta}{2}\,\bar{Y}\,\omega_{110}Z_1,$$

and

$$\text{MSE}\left(\hat{\bar{Y}}_{Gk(ST)}\right)_{min} \cong Z_2^2\bar{X}^2\omega_{020} + Z_1^2\,\bar{Y}^2\omega_{200} + 2\theta Z_1Z_2\,\bar{Y}\bar{X}\,\omega_{020} - 2Z_1Z_2\bar{Y}\bar{X}\omega_{110} + \bar{Y}^2 - 2\,Z_1\bar{Y}^2 + \theta^2Z_1^2\bar{Y}^2 2\,Z_1\bar{Y}^2\omega_{110} - \theta Z_2\bar{Y}\bar{X}\omega_{020} - \theta Z_1^2\bar{Y}^2\omega_{110} - \frac{3\theta^2}{4}\,Z_1\,\bar{Y}^2\omega_{020} + \theta^2Z_1^2\,\bar{Y}^2\omega_{020}. \quad (13)$$

The optimum values of Z_1 and Z_2 are given as:

$$Z_{1(opt)} = \frac{\omega_{020}(\theta^2\omega_{020} - 8)}{8(-\omega_{200}\omega_{020} - \omega_{110}^2 - \omega_{020})},$$

$$Z_{2(opt)} = \frac{\bar{Y}(\theta^2\omega_{020}^2 - \theta^2\omega_{020}\omega_{020} + 4\theta\omega_{200}\omega_{020} - 4\theta\omega_{110}^2 - 4\theta\omega_{020} + 8\omega_{110})}{8\bar{X}(\omega_{200}\omega_{020} - \omega_{110}^2 + \omega_{020})},$$

The minimal MSE of $\hat{\bar{Y}}_{Gk(ST)}$, are:

$$\hat{\bar{Y}}_{GK(ST)_{min}} = \frac{\bar{Y}^2}{64}\left(64 - 16\theta^2\omega_{020} - \frac{\omega_{020}\,(-8 + \theta^2\omega_{020})^2}{\omega_{020}(1 + \omega_{200}) - \omega_{110}^2}\right) \quad (14)$$

4. Proposed Estimator

Efficiency gains for estimators are possible with the use of auxiliary variables introduced during the design or estimate stages. When the study and the auxiliary variable are correlated, ratio type estimators produce strong results. Alternatively, when a negative result occurs among the study and auxiliary variables, the product type estimators provide efficient results. Following the work of [1], we introduced a novel improved class of estimators for population mean under stratified random sampling by combining product and ratio exponential type estimators. More efficient outcomes, such as higher PREs and minimal MSEs, are produced by this class of estimators. The suggested class of estimators is given by:

$$\hat{\bar{T}}^{*}_{prop}=\left[\Omega_1\bar{y}_{(ST)}\left\{\frac{1}{4}\left(\frac{\bar{X}}{\bar{x}_{(ST)}}+\frac{\bar{x}_{(ST)}}{\bar{X}}\right)\left(\exp\left(\frac{\bar{X}-\bar{x}_{(ST)}}{\bar{X}+\bar{x}_{(ST)}}\right)+\exp\left(\frac{\bar{x}_{(ST)}-\bar{X}}{\bar{x}_{(ST)}+\bar{X}}\right)\right)\right\}+\Omega_2\left(\bar{X}-\bar{x}_{(ST)}\right)\right]\exp\left[\frac{a(\bar{X}-\bar{x}_{(ST)})}{a(\bar{X}-\bar{x}_{(ST)})+2b}\right]$$

(15)

Where Ω_1 and Ω_2 are constants.

$$\hat{\bar{T}}^{*}_{prop}=\left[\Omega_1\bar{Y}(1+\varepsilon_0)\left(1+\frac{5}{8}\varepsilon_1^2\right)-\Omega_2\bar{X}\varepsilon_1\right]\left[1-\frac{1}{2}\theta\varepsilon_1+\frac{3}{8}\theta^2\varepsilon_1^2\right] \quad (16)$$

By expressing (16), we have

$$\hat{\bar{T}}^{*}_{prop}-\bar{Y}=\bar{Y}+\bar{Y}\left[(\Omega_1-1)+\Omega_1\left\{\varepsilon_0-\frac{1}{2}\theta\varepsilon_1-\frac{1}{2}\theta\varepsilon_0\varepsilon_1+\frac{1}{8}(5+3\theta^2)\varepsilon_1^2-\Omega_2 R\left(\varepsilon_1-\frac{1}{2}\theta\varepsilon_1^2\right)\right\}\right] \quad (17)$$

Table 1: Some members of the proposed and existing estimators for different choices of *a* and *b*

a	B	$\hat{\bar{T}}_{Singh}$	$\hat{\bar{T}}_{Grover}$	$\hat{\bar{T}}_{Proposed}$
1	$C_{x(ST)}$	$\hat{\bar{T}}_{Singh1}$	$\hat{\bar{T}}_{Grover1}$	$\hat{\bar{T}}_{Proposed1}$
1	$\beta_{x(ST)}$	$\hat{\bar{T}}_{Singh2}$	$\hat{\bar{T}}_{Grover2}$	$\hat{\bar{T}}_{Proposed2}$
$\beta_{x(ST)}$	$C_{x(ST)}$	$\hat{\bar{T}}_{Singh3}$	$\hat{\bar{T}}_{Grover3}$	$\hat{\bar{T}}_{Proposed3}$
$C_{x(ST)}$	$\beta_{x(ST)}$	$\hat{\bar{T}}_{Singh4}$	$\hat{\bar{T}}_{Grover4}$	$\hat{\bar{T}}_{Proposed4}$
1	$\rho_{yx(ST)}$	$\hat{\bar{T}}_{Singh5}$	$\hat{\bar{T}}_{Grover5}$	$\hat{\bar{T}}_{Proposed5}$
$C_{x(ST)}$	$\rho_{yx(ST)}$	$\hat{\bar{T}}_{Singh6}$	$\hat{\bar{T}}_{Grover6}$	$\hat{\bar{T}}_{Proposed6}$
$\rho_{yx(ST)}$	$C_{x(ST)}$	$\hat{\bar{T}}_{Singh7}$	$\hat{\bar{T}}_{Grover7}$	$\hat{\bar{T}}_{Proposed7}$
$\beta_{x(ST)}$	$\rho_{yx(ST)}$	$\hat{\bar{T}}_{Singh8}$	$\hat{\bar{T}}_{Grover8}$	$\hat{\bar{T}}_{Proposed8}$
$\rho_{yx(ST)}$	$\beta_{x(ST)}$	$\hat{\bar{T}}_{Singh9}$	$\hat{\bar{T}}_{Grover9}$	$\hat{\bar{T}}_{Proposed9}$
1	$N\bar{X}(ST)$	$\hat{\bar{T}}_{Singh10}$	$\hat{\bar{T}}_{Grover10}$	$\hat{\bar{T}}_{Proposed10}$

From (17), the bias of $\hat{\bar{T}}^{*}_{prop}$ is given by:

$$\text{Bias}\left(\hat{\bar{T}}^{*}_{prop}\right)=\bar{Y}\left[(\Omega_1-1)+\Omega_1\left\{\frac{1}{8}(5+3\theta^2)\omega_{02}-\frac{1}{2}\theta\omega_{11}+\frac{1}{2}\Omega_2 R_{(ST)}\omega_{02}\right\}\right]$$

Squaring and taking expectation of (17), we have:

$$\text{MSE}\left(\hat{\bar{T}}^{*}_{prop}\right)=\bar{Y}^2\left[1+\Omega_1^2A_{1(ST)}+\Omega_2^2B_{1(ST)}-2\Omega_1C_{1(ST)}-2\Omega_2D_{1(ST)}+2\Omega_1\Omega_2E_{1(ST)}\right] \quad (18)$$

where

$$A_{1(ST)}=1+\left[\omega_{20}+\left(\frac{5}{4}+\theta^2\right)\omega_{02}-2\theta\omega_{11}\right],$$

$$B_{1(ST)}={R_{(ST)}}^2\omega_{02}, \text{ where } R_{(ST)}=\frac{\bar{Y}_{(ST)}}{\bar{X}_{(ST)}},$$

$$C_{1(ST)}=1+\left(\frac{5+3\theta^2}{8}\right)\omega_{02}-\frac{1}{2}\theta\omega_{11},$$

$$D_1=R_{(ST)}\left(\frac{\theta\omega_{02}}{2}\right), E_1=R_{(ST)}(\theta\omega_{02}-\omega_{11}).$$

Differentiate (18) w.r.t Ω_1 and Ω_1, we got the minimum MSEs, which are given by:

$$\Omega_{1(opt)}=\frac{B_{1(ST)}C_{1(ST)}-D_{1(ST)}E_{1(ST)}}{A_{1(ST)}B_{1(ST)}-E_{1(ST)}^2}, \Omega_{2(opt)}=\frac{A_{1(ST)}D_{1(ST)}-C_{1(ST)}E_{1(ST)}}{A_{1(ST)}B_{1(ST)}-E_{1(ST)}^2}$$

Putting the values of $\Omega_{1(opt)}$ and $\Omega_{1(opt)}$ in (18), we get the minimum MSEs, which are given by:

$$MSE\left(\hat{\bar{T}}^{*}_{prop}\right)_{min}=\bar{Y}^2\left[1-\frac{A_1D_1^2+B_1C_1^2-2C_1D_1E_1}{A_1B_1-E_1^2}\right] \quad (19)$$

5. Numerical study

This section numerically examines the performance of the suggested and modified mean estimators. Two populations are considered in the present circumstance. Tables 2 and 3 display the summary statistics for these populations.

Population-I: [**Source:** [8]]

Y is the no. of instructors and

X is the no. of the trainees in 2007 for 923 districts in six regions.

Population-II: [**Source:** [8]]

Y is the no. of instructors and

X is the no. of the schools in 2007 for 923 districts in six regions.

Table 2: Summary statistics of Population-I

H	N_h	n_h	W_h	λ_h		$\bar{Y}_h$	$\bar{X}_h$	S_{yh}	S_{xh}	*Ryxh*		*Ryzh*	
1	127	31	0.1375	0.0244	0.0390	704	20805	883.83	486.75	0.9366	0.9956	0.8239	0.6584
2	117	21	0.1267	0.0248	0.0204	413	9212	644.92	5180.77	0.9937	0.9834	0.6337	0.6360

3	103	29	0.1115	0.0406	74	14309	1033.46	27549.7	0.9893	0.6595
4	170	38	0.1841	0.0207	425	9479	810.58	8218.93	0.9651	0.5863
5	205	22	0.2221		267	5570	403.65	8497.77		
6	201	39	0.2177		394	12998	711.72	3094.14		

Table 3: Summary statistics of Population-II

h	N_h	n_h	W_h	λ_h	$\bar{Y}_h$	$\bar{X}_h$	S_{yh}	S_{xh}	$Ryxh$	$Ryzh$
1	127	31	0.1375	0.0244	704	498	883.83	55.58 65.45	0.9366 0.9956	0.8239 0.6584
2	117	21	0.1267	0.0391	413	318	644.92	12.95	0.9937	0.6337
3	103	29	0.1115	0.0248	574	431	1033.46	458.02	0.9834	0.6360
4	170	38	0.1841	0.0204 0.0406	425	311	810.58	60.85	0.9893	0.6595
5	205 201	22	0.2221	0.0207	267	227	410.65	97.05	0.9651	0.5863
6		39	0.2177		394	314	711.72			

Table 4: MSE using Population I

Estimators	MSEs	Estimators	MSEs	Estimators	MSEs	Estimators	MSEs
$\hat{\bar{Y}}_{(ST)}$	2229.266	$\hat{\bar{Y}}_{Singh1}$	602.6083	$\hat{\bar{T}}_{Grover1}$	192.9409	$\hat{\bar{Y}}_{Proposed1}$	172.4814
$\hat{\bar{Y}}_{R(ST)}$	216.4183	$\hat{\bar{Y}}_{Singh2}$	604.3442	$\hat{\bar{Y}}_{Grover2}$	192.9546	$\hat{\bar{Y}}_{Proposed2}$	172.5040
$\hat{\bar{Y}}_{P(ST)}$	9205.298	$\hat{\bar{Y}}_{Singh3}$	602.4519	$\hat{\bar{Y}}_{Grover3}$	192.9485	$\hat{\bar{Y}}_{Proposed3}$	172.4794
$\hat{\bar{Y}}_{dif(ST)}$	194.2832	$\hat{\bar{Y}}_{Singh4}$	603.4636	$\hat{\bar{Y}}_{Grover4}$	192.9518	$\hat{\bar{Y}}_{Proposed4}$	172.4926
$\hat{\bar{Y}}_{R,D(ST)}$	194.0853	$\hat{\bar{Y}}_{Singh5}$	602.5282	$\hat{\bar{Y}}_{Grover5}$	192.9487	$\hat{\bar{Y}}_{Proposed5}$	172.4804
		$\hat{\bar{Y}}_{Singh6}$	602.4893	$\hat{\bar{Y}}_{Grover6}$	192.9486	$\hat{\bar{Y}}_{Proposed6}$	172.4798
		$\hat{\bar{Y}}_{Singh7}$	602.6161	$\hat{\bar{Y}}_{Grover7}$	192.949	$\hat{\bar{Y}}_{Proposed7}$	172.4815
		$\hat{\bar{Y}}_{Singh8}$	602.4481	$\hat{\bar{Y}}_{Grover8}$	192.9485	$\hat{\bar{Y}}_{Proposed8}$	172.4793
		$\hat{\bar{Y}}_{Singh9}$	604.4351	$\hat{\bar{Y}}_{Grover9}$	192.9549	$\hat{\bar{Y}}_{Proposed9}$	172.5052
		$\hat{\bar{Y}}_{Singh10}$	2226.8349	$\hat{\bar{Y}}_{Grover10}$	194.0853	$\hat{\bar{Y}}_{Proposed10}$	178.5637

Table 5: MSE using Population II

Estimators	MSEs	Estimators	MSEs	Estimators	MSEs	Estimators	MSEs
$\hat{\bar{Y}}_{(ST)}$	2229.266	$\hat{\bar{Y}}_{Singh1}$	881.4006	$\hat{\bar{T}}_{Grover1}$	101.1277	$\hat{\bar{Y}}_{Proposed1}$	95.18579
$\hat{\bar{Y}}_{R(ST)}$	193.2885	$\hat{\bar{Y}}_{Singh2}$	920.0258	$\hat{\bar{Y}}_{Grover2}$	101.1611	$\hat{\bar{Y}}_{Proposed2}$	95.32489
$\hat{\bar{Y}}_{P(ST)}$	6936.636	$\hat{\bar{Y}}_{Singh3}$	877.6357	$\hat{\bar{Y}}_{Grover3}$	101.1242	$\hat{\bar{Y}}_{Proposed3}$	95.17167
$\hat{\bar{Y}}_{dif(ST)}$	101.5021	$\hat{\bar{Y}}_{Singh4}$	909.7116	$\hat{\bar{Y}}_{Grover4}$	101.1525	$\hat{\bar{Y}}_{Proposed4}$	95.28874
$\hat{\bar{Y}}_{R,D(ST)}$	101.4481	$\hat{\bar{Y}}_{Singh5}$	880.3402	$\hat{\bar{Y}}_{Grover5}$	101.1267	$\hat{\bar{Y}}_{Proposed5}$	95.18182
		$\hat{\bar{Y}}_{Singh6}$	879.6034	$\hat{\bar{Y}}_{Grover6}$	101.126	$\hat{\bar{Y}}_{Proposed6}$	95.17906
		$\hat{\bar{Y}}_{Singh7}$	881.4854	$\hat{\bar{Y}}_{Grover7}$	101.1278	$\hat{\bar{Y}}_{Proposed7}$	95.18611
		$\hat{\bar{Y}}_{Singh8}$	877.5616	$\hat{\bar{Y}}_{Grover8}$	101.1242	$\hat{\bar{Y}}_{Proposed8}$	95.17139
		$\hat{\bar{Y}}_{Singh9}$	920.8959	$\hat{\bar{Y}}_{Grover9}$	101.1618	$\hat{\bar{Y}}_{Proposed9}$	95.32791
		$\hat{\bar{Y}}_{Singh10}$	2227.442	$\hat{\bar{Y}}_{Grover10}$	101.4481	$\hat{\bar{Y}}_{Proposed10}$	96.94195

Table 6: PRE using Population I

Estimators	PREs	Estimators	PREs	Estimators	PREs	Estimators	PREs
$\hat{\bar{Y}}_{(ST)}$	100	$\hat{\bar{Y}}_{Singh1}$	369.9362	$\hat{\bar{T}}_{Grover1}$	1155.3655	$\hat{\bar{Y}}_{Proposed1}$	1292.4676
$\hat{\bar{Y}}_{R(ST)}$	1030.073	$\hat{\bar{Y}}_{Singh2}$	368.8736	$\hat{\bar{Y}}_{Grover2}$	1155.3318	$\hat{\bar{Y}}_{Proposed2}$	1292.2979
$\hat{\bar{Y}}_{P(ST)}$	24.21721	$\hat{\bar{Y}}_{Singh3}$	370.0323	$\hat{\bar{Y}}_{Grover3}$	1155.3686	$\hat{\bar{Y}}_{Proposed3}$	1292.4829
$\hat{\bar{Y}}_{dif(ST)}$	1147.431	$\hat{\bar{Y}}_{Singh4}$	369.4119	$\hat{\bar{Y}}_{Grover4}$	1155.3489	$\hat{\bar{Y}}_{Proposed4}$	1292.3839
$\hat{\bar{Y}}_{R,D(ST)}$	1148.601	$\hat{\bar{Y}}_{Singh5}$	369.9854	$\hat{\bar{Y}}_{Grover5}$	1155.3671	$\hat{\bar{Y}}_{Proposed5}$	1292.4754
		$\hat{\bar{Y}}_{Singh6}$	370.0093	$\hat{\bar{Y}}_{Grover6}$	1155.3679	$\hat{\bar{Y}}_{Proposed6}$	1292.4792
		$\hat{\bar{Y}}_{Singh7}$	369.9314	$\hat{\bar{Y}}_{Grover7}$	1155.3654	$\hat{\bar{Y}}_{Proposed7}$	1292.4668
		$\hat{\bar{Y}}_{Singh8}$	370.0345	$\hat{\bar{Y}}_{Grover8}$	1155.3687	$\hat{\bar{Y}}_{Proposed8}$	1292.4833
		$\hat{\bar{Y}}_{Singh9}$	368.8181	$\hat{\bar{Y}}_{Grover9}$	1155.303	$\hat{\bar{Y}}_{Proposed9}$	1292.2891
		$\hat{\bar{Y}}_{Singh10}$	100.1092	$\hat{\bar{Y}}_{Grover10}$	1148.6014	$\hat{\bar{Y}}_{Proposed10}$	1248.4431

Table 7: PRE using Population II

Estimators	PREs	Estimators	PREs	Estimators	PREs	Estimators	PREs
$\hat{\bar{Y}}_{(ST)}$	100	$\hat{\bar{Y}}_{Singh1}$	252.9232	$\hat{\bar{T}}_{Grover1}$	2204.408	$\hat{\bar{Y}}_{Proposed1}$	2342.016
$\hat{\bar{Y}}_{R(ST)}$	1153.336	$\hat{\bar{Y}}_{Singh2}$	242.3048	$\hat{\bar{Y}}_{Grover2}$	2203.679	$\hat{\bar{Y}}_{Proposed2}$	2338.598
$\hat{\bar{Y}}_{P(ST)}$	32.13757	$\hat{\bar{Y}}_{Singh3}$	254.0082	$\hat{\bar{Y}}_{Grover3}$	2204.483	$\hat{\bar{Y}}_{Proposed3}$	2342.363
$\hat{\bar{Y}}_{dif(ST)}$	2196.275	$\hat{\bar{Y}}_{Singh4}$	245.052	$\hat{\bar{Y}}_{Grover4}$	2203.866	$\hat{\bar{Y}}_{Proposed4}$	2339.486
$\hat{\bar{Y}}_{R,D(ST)}$	2197.445	$\hat{\bar{Y}}_{Singh5}$	253.2278	$\hat{\bar{Y}}_{Grover5}$	2204.429	$\hat{\bar{Y}}_{Proposed5}$	2342.113
		$\hat{\bar{Y}}_{Singh6}$	253.4399	$\hat{\bar{Y}}_{Grover6}$	2204.443	$\hat{\bar{Y}}_{Proposed6}$	2342.181
		$\hat{\bar{Y}}_{Singh7}$	252.8988	$\hat{\bar{Y}}_{Grover7}$	2204.406	$\hat{\bar{Y}}_{Proposed7}$	2342.008
		$\hat{\bar{Y}}_{Singh8}$	254.0296	$\hat{\bar{Y}}_{Grover8}$	2204.484	$\hat{\bar{Y}}_{Proposed8}$	2342.37
		$\hat{\bar{Y}}_{Singh9}$	242.0758	$\hat{\bar{Y}}_{Grover9}$	2203.664	$\hat{\bar{Y}}_{Proposed9}$	2338.524
		$\hat{\bar{Y}}_{Singh10}$	100.0819	$\hat{\bar{Y}}_{Grover10}$	2197.445	$\hat{\bar{Y}}_{Proposed10}$	2299.589

6. Simulation study

We have produced two populations of sizes 1200 and 2000 respectively, from normal distribution. The population I consists of six strata of equal sizes and population II consist of four strata of unequal sizes. The characteristic of both populations are given below.

Population I

X = $N(5, 10)$, $Y = X + N(0, 1)$, $N = 1200$, $n=300$

N_1=200, N_2=200, N_3=200, N_4=200, N_5=200, N_6=200

n_1 =50, n_2=50, n_3=50, n_4=50, n_5=50, n_6=50

Population II

$X = N(3, 7)$, Y = X + $N(0, 1)$, $N = 2000$, $n=600$

N_1=500, N_2=500, N_3=500, N_4=500

n_1=200, n_2=150, n_3=150, n_4=100

Following populations are used in simulation study algorithm to investigate the performance of the proposed estimator over the existing estimators

Table 5: MSE using simulated Population I

Estimators	MSEs	Estimators	MSEs	Estimators	MSEs	Estimators	MSEs
$\hat{\bar{Y}}_{(ST)}$	0.2175383	$\hat{\bar{Y}}_{Singh1}$	0.90536901	$\hat{\bar{T}}_{Grover1}$	0.002497143	$\hat{\bar{Y}}_{Proposed1}$	0.001667290
$\hat{\bar{Y}}_{R(ST)}$	0.002521325	$\hat{\bar{Y}}_{Singh2}$	0.104679918	$\hat{\bar{Y}}_{Grover2}$	0.002501332	$\hat{\bar{Y}}_{Proposed2}$	0.001707072
$\hat{\bar{Y}}_{P(ST)}$	0.87469	$\hat{\bar{Y}}_{Singh3}$	0.070982010	$\hat{\bar{Y}}_{Grover3}$	0.002487206	$\hat{\bar{Y}}_{Proposed3}$	0.001593614
$\hat{\bar{Y}}_{dif(ST)}$	0.002479927	$\hat{\bar{Y}}_{Singh4}$	0.087056073	$\hat{\bar{Y}}_{Grover4}$	0.002495792	$\hat{\bar{Y}}_{Proposed4}$	0.001655956
$\hat{\bar{Y}}_{R,D(ST)}$	0.002479663	$\hat{\bar{Y}}_{Singh5}$	0.077467466	$\hat{\bar{Y}}_{Grover5}$	0.002491206	$\hat{\bar{Y}}_{Proposed5}$	0.001620945
		$\hat{\bar{Y}}_{Singh6}$	0.069437500	$\hat{\bar{Y}}_{Grover6}$	0.002486123	$\hat{\bar{Y}}_{Proposed6}$	0.001586607
		$\hat{\bar{Y}}_{Singh7}$	0.090666975	$\hat{\bar{Y}}_{Grover7}$	0.002497190	$\hat{\bar{Y}}_{Proposed7}$	0.001667701
		$\hat{\bar{Y}}_{Singh8}$	0.065584647	$\hat{\bar{Y}}_{Grover8}$	0.002483172	$\hat{\bar{Y}}_{Proposed8}$	0.001568204
		$\hat{\bar{Y}}_{Singh9}$	0.104848904	$\hat{\bar{Y}}_{Grover9}$	0.002501371	$\hat{\bar{Y}}_{Proposed9}$	0.001707493
		$\hat{\bar{Y}}_{Singh10}$	0.227887105	$\hat{\bar{Y}}_{Grover10}$	0.002508170	$\hat{\bar{Y}}_{Proposed10}$	0.001826122

Table 6: MSE using simulated Population II

Estimators	MSEs	Estimators	MSEs	Estimators	MSEs	Estimators	MSEs
$\hat{\bar{Y}}_{(ST)}$	0.1832313	$\hat{\bar{Y}}_{Singh1}$	0.12011631	$\hat{\bar{T}}_{Grover1}$	0.07712684	$\hat{\bar{Y}}_{Proposed1}$	0.07651846
$\hat{\bar{Y}}_{R(ST)}$	0.07741117	$\hat{\bar{Y}}_{Singh2}$	0.12645562	$\hat{\bar{Y}}_{Grover2}$	0.07713740	$\hat{\bar{Y}}_{Proposed2}$	0.07653746
$\hat{\bar{Y}}_{P(ST)}$	0.5050382	$\hat{\bar{Y}}_{Singh3}$	0.10970820	$\hat{\bar{Y}}_{Grover3}$	0.07710331	$\hat{\bar{Y}}_{Proposed3}$	0.07647670
$\hat{\bar{Y}}_{dif(ST)}$	0.07740024	$\hat{\bar{Y}}_{Singh4}$	0.11678600	$\hat{\bar{Y}}_{Grover4}$	0.07712026	$\hat{\bar{Y}}_{Proposed4}$	0.07650672
$\hat{\bar{Y}}_{R,D(ST)}$	0.07676882	$\hat{\bar{Y}}_{Singh5}$	0.11263376	$\hat{\bar{Y}}_{Grover5}$	0.07712697	$\hat{\bar{Y}}_{Proposed5}$	0.07651869
		$\hat{\bar{Y}}_{Singh6}$	0.10663461	$\hat{\bar{Y}}_{Grover6}$	0.07709437	$\hat{\bar{Y}}_{Proposed6}$	0.07646958
		$\hat{\bar{Y}}_{Singh7}$	0.12018505	$\hat{\bar{Y}}_{Grover7}$	0.07712697	$\hat{\bar{Y}}_{Proposed7}$	0.07651869

		$\hat{\bar{Y}}_{Singh8}$	0.10663461	$\hat{\bar{Y}}_{Grover8}$	0.07709437	$\hat{\bar{Y}}_{Proposed8}$	0.07646104
		$\hat{\bar{Y}}_{Singh9}$	0.12654091	$\hat{\bar{Y}}_{Grover9}$	0.07713753	$\hat{\bar{Y}}_{Proposed9}$	0.07653769
		$\hat{\bar{Y}}_{Singh10}$	0.18317782	$\hat{\bar{Y}}_{Grover10}$	0.07717051	$\hat{\bar{Y}}_{Proposed10}$	0.07659820

Table 7: PRE using Simulated Population I

Estimators	PREs	Estimators	PREs	Estimators	PREs	Estimators	PREs
$\hat{\bar{Y}}_{(ST)}$	100	$\hat{\bar{Y}}_{Singh1}$	251.9141	$\hat{\bar{T}}_{Grover1}$	9133.4481	$\hat{\bar{Y}}_{Proposed1}$	13679.3934
$\hat{\bar{Y}}_{R(ST)}$	8627.936	$\hat{\bar{Y}}_{Singh2}$	217.8787	$\hat{\bar{Y}}_{Grover2}$	9118.1513	$\hat{\bar{Y}}_{Proposed2}$	13360.6095
$\hat{\bar{Y}}_{P(ST)}$	24.87033	$\hat{\bar{Y}}_{Singh3}$	321.3141	$\hat{\bar{Y}}_{Grover3}$	9169.9387	$\hat{\bar{Y}}_{Proposed3}$	14311.8196
$\hat{\bar{Y}}_{dif(ST)}$	8771.963	$\hat{\bar{Y}}_{Singh4}$	261.9866	$\hat{\bar{Y}}_{Grover4}$	9138.9866	$\hat{\bar{Y}}_{Proposed4}$	13773.0235
$\hat{\bar{Y}}_{R,D(ST)}$	8772.897	$\hat{\bar{Y}}_{Singh5}$	294.4142	$\hat{\bar{Y}}_{Grover5}$	9155.2150	$\hat{\bar{Y}}_{Proposed5}$	14070.5123
		$\hat{\bar{Y}}_{Singh6}$	328.4612	$\hat{\bar{Y}}_{Grover6}$	9173.9314	$\hat{\bar{Y}}_{Proposed6}$	14375.0256
		$\hat{\bar{Y}}_{Singh7}$	251.5527	$\hat{\bar{Y}}_{Grover7}$	9133.2738	$\hat{\bar{Y}}_{Proposed7}$	13676.0252
		$\hat{\bar{Y}}_{Singh8}$	347.7570	$\hat{\bar{Y}}_{Grover8}$	9184.8323	$\hat{\bar{Y}}_{Proposed8}$	14543.7189
		$\hat{\bar{Y}}_{Singh9}$	217.5275	$\hat{\bar{Y}}_{Grover9}$	9118.0069	$\hat{\bar{Y}}_{Proposed9}$	13357.3154
		$\hat{\bar{Y}}_{Singh10}$	100.0825	$\hat{\bar{Y}}_{Grover10}$	9093.2909	$\hat{\bar{Y}}_{Proposed10}$	12489.5967

Table 8: PRE using simulated Population II

Estimators	PREs	Estimators	PREs	Estimators	PREs	Estimators	PREs
$\hat{\bar{Y}}_{(ST)}$	100	$\hat{\bar{Y}}_{Singh1}$	152.5449	$\hat{\bar{T}}_{Grover1}$	237.5714	$\hat{\bar{Y}}_{Proposed1}$	239.4602
$\hat{\bar{Y}}_{R(ST)}$	236.6987	$\hat{\bar{Y}}_{Singh2}$	144.8977	$\hat{\bar{Y}}_{Grover2}$	237.5388	$\hat{\bar{Y}}_{Proposed2}$	239.4008
$\hat{\bar{Y}}_{P(ST)}$	36.28067	$\hat{\bar{Y}}_{Singh3}$	167.0169	$\hat{\bar{Y}}_{Grover3}$	237.6439	$\hat{\bar{Y}}_{Proposed3}$	239.5910
$\hat{\bar{Y}}_{dif(ST)}$	236.7322	$\hat{\bar{Y}}_{Singh4}$	156.8949	$\hat{\bar{Y}}_{Grover4}$	237.5916	$\hat{\bar{Y}}_{Proposed4}$	239.4970
$\hat{\bar{Y}}_{R,D(ST)}$	238.6793	$\hat{\bar{Y}}_{Singh5}$	162.6788	$\hat{\bar{Y}}_{Grover5}$	237.6206	$\hat{\bar{Y}}_{Proposed5}$	239.5492
		$\hat{\bar{Y}}_{Singh6}$	169.2426	$\hat{\bar{Y}}_{Grover6}$	237.6564	$\hat{\bar{Y}}_{Proposed6}$	239.6133
		$\hat{\bar{Y}}_{Singh7}$	152.4576	$\hat{\bar{Y}}_{Grover7}$	237.5710	$\hat{\bar{Y}}_{Proposed7}$	239.4595
		$\hat{\bar{Y}}_{Singh8}$	171.8310	$\hat{\bar{Y}}_{Grover8}$	237.6714	$\hat{\bar{Y}}_{Proposed8}$	239.6401
		$\hat{\bar{Y}}_{Singh9}$	144.8000	$\hat{\bar{Y}}_{Grover9}$	237.5384	$\hat{\bar{Y}}_{Proposed9}$	239.4001
		$\hat{\bar{Y}}_{Singh10}$	100.0292	$\hat{\bar{Y}}_{Grover10}$	237.4369	$\hat{\bar{Y}}_{Proposed10}$	239.2109

7. Discussion

We used two real data sets and a simulated study to evaluate the estimators suggested for the stratified sampling strategy for population mean estimation with the help of auxiliary variable information. We also considered the fact that the populations in Scenario I and Scenario II had different sample sizes. We compared the modified stratified estimators to the proposed family of estimators in terms of mean square error and percentage relative efficiency. The minimum MSE of the proposed class of estimator is shown in Equation 19. Table 2 and Table 3 show the summary statistics of the mentioned population. The numerical findings can be found in Tables 4–7, which demonstrate that the MSEs and PREs of various estimators are changes by the choices of *a* and *b*. Furthermore, it is seen that the proposed estimators perform better than the existing ones on both actual data sets and simulated studies when considering MSEs and percentage relative efficiency.

8. Conclusion

The primary objective of this research is to introduce a new method for estimating population means in heterogeneous settings by using data from auxiliary sources. By deriving and comparing the bias and MSE of the novel estimator/method with those of existing estimators, the mathematical effectiveness of the study can be evaluated. Applying the proposed estimator to real datasets proves its practicality. Furthermore, we constructed a simulation to determine how well our proposed estimator worked with different artificially generated data sets. We have calculated the bias and mean squared errors (MSEs) both the empirical and theoretical of all the estimators. Through the numerical assessments, we show that the recommended estimator perform well as compared to existing estimators.

This estimator is highly recommended for use in sample surveys carried out in many different domains, such as agriculture, e ducation, health sciences, and fisheries. In addition to ratio, product, difference and exponential estimators, we suggest including them in studies that use stratified random sampling. An extensive number of additional information operating under various sampling frameworks can be incorporated into the approach using an extension.

Data availability

All the data are available within the manuscript.

Conflict of interest
The authors declare no conflict of interest

References

1. Ahmad, S., Shabbir, J., Emam, W., Zahid, E., Aamir, M., Khalid, M., & Anas, M. M. (2024). An improved class of estimators for estimation of population distribution functions under stratified random sampling. *Heliyon*, *10*(7).
2. Ahmad, S., Hussain, S., Shabbir, J., Zahid, E., Aamir, M., & Onyango, R. (2022). Improved estimation of finite population variance using dual supplementary information under stratified random sampling. *Mathematical Problems in Engineering*, *2022*(1), 3813952.
3. Bhushan, S., Kumar, A., Shahab, S., Lone, S. A., & Akhtar, M. T. (2022). On efficient estimation of the population mean under stratified ranked set sampling. *Journal of Mathematics*, *2022*(1), 6196142.
4. Bhushan, S., Kumar, A., & Singh, S. (2023). Some efficient classes of estimators under stratified sampling. *Communications in Statistics-Theory and Methods*, *52*(6), 1767-1796.
5. Bhushan, S., Kumar, A., Lone, S. A., Anwar, S., & Gunaime, N. M. (2023). An efficient class of estimators in stratified random sampling with an application to real data. *Axioms*, *12*(6), 576.
6. Bhushan, S., Kumar, A., Tyagi, D., & Singh, S. (2022). On some improved classes of estimators under stratified sampling using attribute. *Journal of Reliability and Statistical Studies*, 187-210.
7. Cochran, W. G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *The journal of agricultural science*, *30*(2), 262-275.
8. Koyuncu, N., & Kadilar, C. (2009). Ratio and product estimators in stratified random sampling. *Journal of statistical planning and inference*, *139*(8), 2552-2558.
9. Mradula, Yadav, S. K., Varshney, R., & Dube, M. (2021). Efficient estimation of population mean under stratified random sampling with linear cost function. *Communications in Statistics-Simulation and Computation*, *50*(12), 4364-4387.
10. Rao, J. N. K. (1994). Estimating totals and distribution functions using auxiliary information at the estimation stage. *Journal of official statistics*, *10*(2), 153.
11. Singh, R., Chauhan, P., Sawan, N., & Smarandache, F. (2007). Improvement in estimating the population mean using exponential estimator in simple random sampling. *Auxiliary Information and a priori Values in Construction of Improved Estimators*, *33*.
12. Grover, L. K., & Kaur, P. (2014). A generalized class of ratio type exponential estimators of population mean under linear transformation of auxiliary variable. *Communications in Statistics-Simulation and Computation*, *43*(7), 1552-1574.
13. Triveni, G. R. V., & Danish, F. (2024). Efficient population mean estimation via stratified sampling with dual auxiliary information: A real estate perspective. *Alexandria Engineering Journal*, *104*, 680-687.
14. Pachori, M., & Garg, N. (2024). Ratio-type estimator of the population mean in stratified sampling based on the calibration approach. *Statistics in Transition*, *25*(2), 39-55.
15. Pal, A., Varshney, R., Yadav, S. K., & Zaman, T. (2024). Improved memory-type ratio estimator for population mean in stratified random sampling under linear and non-linear cost functions. *Soft Computing*, 1-16.
16. Ahmad, S., Hussain, S., Al Mutairi, A., Kamal, M., Rehman, M. U., & Mustafa, M. S. (2024). Improved estimation of population distribution function using twofold auxiliary information under simple random sampling. *Heliyon*, *10*(2).
17. Yadav, S. K., Kumar Verma, M., & Varshney, R. (2024). Optimal strategy for elevated estimation of population mean in stratified random sampling under linear cost function. *Annals of Data Science*, 1-22.
18. Zaagan, A. A., Verma, M. K., Mahnashi, A. M., Yadav, S. K., Ahmadini, A. A. H., Meetei, M. Z., & Varshney, R. (2024). An effective and economic estimation of population mean in stratified random sampling using a linear cost function. *Heliyon*, *10*(10).
19. Zaagan, A. A., Meetei, M. Z., Singh, S., Gupta, R., Yadav, S. K., & Verma, M. K. (2024). Computing the population mean in stratified sampling using an auxiliary attribute. *Journal of Autonomous Intelligence*, *7*(5), 12-21.
20. Zakari, Y., & Muhammad, I. (2023). Modified estimator of finite population variance under stratified random sampling. *Engineering Proceedings*, *56*(1), 177.
21. Zaman, T., & Kadilar, C. (2021). Exponential ratio and product type estimators of the mean in stratified two-phase sampling. *AIMS Mathematics*, *6*(5), 4265-4279.
22. Zaman, T., & Bulut, H. (2020). Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling. *Communications in Statistics-theory and Methods*, *49*(14), 3407-3420.
23. Zaman, T., & Kadilar, C. (2020). On estimating the population mean using auxiliary character in stratified random sampling. *Journal of Statistics and Management Systems*, *23*(8), 1415-1426.
24. Hussain, S., Ahmad, S., Saleem, M., & Akhtar, S. (2020). Finite population distribution function estimation with dual use of auxiliary information under simple and stratified random sampling. *Plos one*, *15*(9), e0239098.
25. Hussain, S., Ahmad, S., Akhtar, S., Javed, A., & Yasmeen, U. (2020). Estimation of finite population distribution function with dual use of auxiliary information under non-response. *Plos one*, *15*(12), e0243584.
26. Ahmad, S., Hussain, S., Aamir, M., Khan, F., Alshahrani, M. N., & Alqawba, M. (2022). Estimation of finite population mean using dual auxiliary variable for non-response using simple random sampling. *Aims Mathematics*, *7*(3), 4592-4613.