Thailand Statistician
April 2024; 22(2): 363-373
http://statassoc.or.th
Contributed paper

Calibrated Estimator for Sensitive Variables under Stratified Random Sampling

Riffat Jabeen*[a], Muhammad Kamran Aslam [a], Aamir Sanaullah [a] and Azam Zaka [b]
[a] Department of Statistics, COMSATS University Islamabad, Lahore Campus, Lahore, Pakistan.
[b] Department of Statistics, Government Graduate College of Science, Wahdat Road, Lahore, Pakistan.
*Corresponding author; e-mail: drriffatjabeen@cuilahore.edu.pk

Received: 15 January 2021
Revised: 25 April 2022
Accepted: 20 May 2022

Abstract

Calibration sampling is a general tool to adjust the sampling weights and enhance the precision of the estimates. This technique is also helpful to reduce the non-response errors. In order to remove or minimize the biases produced by non-response errors and variances, the calibration technique is utilized. In this paper, Calibration technique is used to reduce the distance between the calibrated weights and the given distance measure. We propose new calibration estimators for estimating the population of a sensitive variable based on scrambled responses collected using some improved random response device and auxiliary information. This study is to propose some improved calibrated generalized estimators for estimation of population mean of a quantitative sensitive variable. The results show that the proposed estimator having an extra calibration constraint is more efficient.

Keywords: Auxiliary information, calibration, scrambled randomized response technique, stratified RR technique.

1. Introduction

In stratified random sampling, the problem of estimating the mean using the additional information has been well documented in the survey sampling. The scheme of stratified random sampling involves the division of a population into homogeneous subgroups called strata, and then from each stratum through simple random sampling we select the sample.

The randomized response technique is used in complex survey related to the sensitive issues such as use of drug, criminal record of the people or tax evocation to get the response from the respondents. Hansen and Hurwitz (1946) developed classical ratio and regression estimators for stratified sampling. Further, Kadilar and Cingi (2003) presented some estimators for stratified random sampling and Shabbir and Gupta (2005) adapted some estimators and improved them under stratified sampling. Proportion allocation is widely used for the selection of sample size for each stratum made, if costs and variances are about equal for each stratum. If the costs and variances differ across strata, then disproportionate stratification is preferred.

Calibration is a procedure to incorporating auxiliary information to adjust the sampling weights known as calibration weights that make the estimates agree with known totals. These weight resulted design consistent estimator and are more efficient than Horvitz-Thompson estimator. When too many auxiliary variables are involved in estimation, then calibration approach can produce extreme and negative weights. As calibration technique produces only a single weight for corresponding study variable value and this weight satisfy calibration to all benchmark constrains.

An alternative technique was introduced by Eichhorn and Hayre (1983) suggested that required to ask a sensitive question, the respondent reply quantitative form, this scheme called Scrambled Randomized Response (SRR) and the respondent themselves following some provided mechanism. SRR method is a special case of Pollock and Bek (1976). Mangat (1994) modified, Mangat and Singh (1990) under conditions that are obtained under which the suggested strategy is improved than those of Warner (1965) or Mangat and Singh (1990). Zaman (2019) suggested new ratio estimators in stratified random sampling using the information of an auxiliary attribute. Zaman and Bulut (2020) gave a new idea to propose new regression-type estimators utilizing Tukey-M, Hampel M, Huber MM, LTS, LMS and LAD robust methods and MCD and MVE robust covariance matrices in stratified sampling. Zaman and Kadilar (2020) proposed exponential estimator taking auxiliary attribute and compared it with some existing estimators. Zaman and Kadilar (2021) proposed exponential ratio estimators in the stratified two-phase sampling using an auxiliary attribute. Afterwards Zaman (2021) suggested estimator under stratified random sampling to estimate the population mean. Jabeen et al. (2021) provide calibration estimators using different calibration constraints and distance measures and proved that selection of calibration constraints effects the efficiency of the estimators.

Calibration estimators proposed by different survey statisticians for estimating population mean has been discussed and a brief discussion on some existing estimators to estimate population mean for sensitive study variable has been delivered along with the bias and the mean squared error. To deal with the situation of sensitive study variables some randomized response techniques, the work on the estimation of sensitive variable of interest using different response models are given with the bias and the mean square error.

We have proposed calibration estimators when the study variable is of sensitive nature. The modifications on Tracy (2003) and Eichhorn and Hayre (1983) is done using different calibration constraints and we see that the calibration constraints play a vital role in the efficiency of the estimators.

2. Methodology

2.1. Notations in calibration estimator

Consider a finite population consists of "N" units $U=\{U_1,U_2,\ldots,U_N\}$. Let y be the sensitive variable under study with population mean and variance respectively as $\overline{Y}=\frac{1}{N}\sum_{i=1}^{N}Y_i$

and $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2$. Let X_i be the i^{th} $(i = 1, 2, ..., q)$ sensitive auxiliary variable having population mean and variance as $\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ and $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2$.

The size of population, N, is stratified into K^{th} strata with J^{th} stratum containing "N_j" units, where $(j = 1, 2, ..., K)$. Such that $\sum_{j=1}^{k} N_j$. Let a sample of size n_1 is selected from the population through simple random sampling without replacement (SRSWOR) and consider $\overline{y} = \frac{1}{n_1}\sum_{i=1}^{n_1} y_i$ and $\overline{x} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i$ be the sample means for "y" and "x", respectively. The classical unbiased estimator of the population mean is given by $\overline{y}_{st} = \frac{1}{N}\sum_{J=1}^{N} W_j \overline{y}_j$, where $W_j = \frac{N_j}{N}$ is the stratum weight and $\overline{y}_j$ is stratum mean for j^{th} strata.

The conventional estimator under stratified random sampling is given as

$$\overline{y}_{st} = \sum_{j=1}^{K} w_j \overline{y}_j.$$

In calibration approach, the aim is to reduce the weights (w_j) to increase the efficiency of the estimators used under stratified random sampling. The calibration weights are produced by minimizing the commonly used chi-square distance measure using different constraints. The chi-square distance measure is given here as

$$\sum_{j=1}^{k} \frac{\left(\Omega_j - W_j\right)^2}{Q_j W_j},$$

where Q_j are the weights which determine about the form of the estimator and Ω_j are the weights minimize distance measure. The auxiliary information is efficiently used to increase the precision of the estimators.

2.2. Existing estimators in literature

Tracy et al. (2003) proposed the calibration estimator for estimating population mean as

$$\overline{y}_{st} = \sum_{j=1}^{K} w_j^{\,S} \overline{y}_j.$$

Here, $w_j^{\,S}$ are the calibrated weights which are chosen by minimizing the chi-square distance measure and using the following constraints

$$\sum_{j=1}^{k} w_j^s \overline{x}_j = \overline{X}, \quad \sum_{j=1}^{k} w_j^s = 1.$$

The optimum weights are given below

$$w_j^{\ S}=w_j+\frac{\left(w_jQ_j\bar{x}_j\right)\left(\sum_{j=1}^{k}w_jQ_j\right)-w_jQ_j\left(\sum_{j=1}^{k}w_jQ_j\bar{x}_j\right)}{\left(\sum_{j=1}^{k}w_jQ_j\right)\left(\sum_{j=1}^{k}w_jQ_j\bar{x}_j^{\ 2}\right)-\left(\sum_{j=1}^{k}w_jQ_j\bar{x}_j\right)^2}\left(\bar{X}-\sum_{j=1}^{k}w_j\bar{x}_j\right).$$

The calibrated estimator by Tracy et al. (2003) may be written as

$$\bar{y}_{st}=\sum_{j=1}^{k}w_j\bar{y}_j+\hat{\beta}\left(\bar{X}-\sum_{j=1}^{k}w_j\bar{x}_j\right).$$

Tracy et al. (2003) also proposed the calibration estimator for estimating the population mean under stratified sampling scheme as

$$\bar{y}_{st}=\sum_{j=1}^{k}w_j\bar{y}_j+\hat{\beta}\left(\bar{X}-\sum_{j=1}^{k}w_j\bar{x}_j\right),$$

by using the following constraints

$$\sum_{j=1}^{k}w_j^{\ TSA}\bar{x}_j=\bar{X},\quad \sum_{j=1}^{k}w_j^{\ TSA}s_{xj}^2=\sum_{j=1}^{k}w_jS_{xj}^2,$$

and the minimization of chi-square distance measure subject to the constraints will provide us with the following calibration weights

$$W_j^{\ TSA}=W_j+Q_jW_j\bar{x}_j\frac{\left(\sum_{j=1}^{k}Q_jW_js_{jx_j}^4\right)\left(\sum_{j=1}^{k}W_j\left(\bar{X}_j-\bar{x}_j\right)\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right)\left(\sum_{j=1}^{k}Wj\left(S_{jx}^2-s_{jx}^2\right)\right)}{\left(\sum_{j=1}^{k}Q_jW_js_{jx_j}^4\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j^2\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right)^2}$$

$$-Q_jW_js_{jx}^2\frac{\left(\sum_{j=1}^{k}W_j\left(S_{jx}^2-s_{jx}^2\right)\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j^2\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right)\left(\sum_{j=1}^{k}W_j\left(\bar{X}_j-\bar{x}_j\right)\right)}{\left(\sum_{j=1}^{k}Q_jW_js_{jx}^4\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j^2\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right)^2}.$$

Tracy et al. (2003) calibration estimator is given as

$$\bar{y}_{st}=\sum_{j=1}^{k}W_j\left(\bar{y}_j-\bar{s}\right)_j+2\hat{\beta}_1\sum_{j=1}^{k}W_j\left(\bar{X}_j-\bar{x}_j\right)+2\hat{\beta}_2\sum_{j=1}^{k}W_j\left(S_{jx}^2-s_{jx}^2\right).$$

Koyuncu and Kadilar (2013) proposed the calibration estimator for estimation of population mean as

$$\bar{y}_{st}=\sum_{j=1}^{k}w_j^{\ KK}\bar{y}_j,$$

by using following constraints

$$\sum_{j=1}^{k}w_j^{\ TSA}\bar{x}_j=\sum_{j=1}^{k}w_j\bar{X}_j,\quad \sum_{j=1}^{k}w_j^{\ TSA}s_{xj}^2=\sum_{j=1}^{k}w_jS_{xj}^2,\quad \sum_{j=1}^{k}w_j^{\ TSA}=\sum_{j=1}^{k}w_j.$$

2.3. Proposed estimator-I

Consider new estimator is following Tracy et al. (2003) and Eichhorn and Hayre (1983)

$$\overline{y}_{st(prop1)} = \sum_{j=1}^{k} \Omega_j \bar{Z}_j, \tag{1}$$

$Z_j = S + Y_j$, where Y_j is the real value of the sensitive quantitative variable and S is the scrambling variables whose distribution is assumed to be known. Where Ω_j are the weights minimize distance measure,

$$\sum_{j=1}^{k} \frac{(\Omega_j - W_j)^2}{Q_j W_j}, \tag{2}$$

and satisfying the calibration constraints:

$$\sum_{j=1}^{k} \Omega_j \bar{x}_j = \sum_{j=1}^{k} W_j \bar{X}_j, \tag{3}$$

$$\sum_{j=1}^{k} \Omega_j s_{jx}^2 = \sum_{j=1}^{k} W_j S_{jx}^2. \tag{4}$$

The sensitive variable under auxiliary variable with population mean and variance of the j^{th} stratum respective as $\bar{x}_j = \frac{1}{n_j}\sum_{j=1}^{n_j} x_{ji}$ and $S_{jx}^2 = \frac{1}{N_j - 1}\sum_{i=1}^{N_j}(x_{ji} - \bar{x}_j)^2$. The sample mean and variance estimator's under auxiliary variable of the j^{th} stratum. Consider $\bar{x}_j = \frac{1}{n_{\text{j}}}\sum_{j=1}^{n_j} x_{ji}$ and $S^2{}_{jx} = \frac{1}{n_{\text{j}} - 1}\sum_{i=1}^{n_{\text{j}}}\left(x_{\text{ji}} - \bar{x}_{\text{j}}\right)^2$.

Minimizing chi-square distance measure, given in (2) to used (3) and (4), the Lagrange function is given by

$$D = \frac{\left(\Omega_j - W_j\right)^2}{Q_j W_j} - 2\lambda_1\left(\Omega_j \bar{x}_j - W_j \bar{X}_j\right) - 2\lambda_2\left(\Omega_j s_{jx}^2 - W_j S_{jx}^2\right),$$

where λ_1 and λ_2 are Lagrange multipliers**.** Setting $\frac{\partial D}{\partial \Omega_J} = 0,$ we have

$$\Omega_j = W_j + Q_j W_j\left(\lambda_1 \bar{x}_j + \lambda_2 s_{jx}^2\right), \tag{5}$$

value of Ω_j putting in (3) and (4), we get

$$\lambda_1 \sum_{j=1}^{k} Q_j W_j \bar{x}_j^2 + \lambda_2 \sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2 = \sum_{j=1}^{k} W_j \bar{X}_j - \sum_{j=1}^{k} W_j \bar{x}_j, \tag{6}$$

$$\lambda_1 \sum_{j=1}^{k} Q_j W_j \bar{x}_j^2 s_{jx}^2 + \lambda_2 \sum_{j=1}^{k} Q_j W_j s_{jx}^4 = \sum_{j=1}^{k} W_j S_{jx}^2 - \sum_{j=1}^{k} W_j s_{jx}^2. \tag{7}$$

Solving (6) and (7),

$$\begin{pmatrix} \sum_{j=1}^{k} Q_j W_j \bar{x}_j^2 & \sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2 \\ \sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2 & \sum_{j=1}^{k} Q_j W_j s_{jx_j}^4 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{k} W_j \bar{X}_j - \sum_{j=1}^{k} W_j \bar{x}_j \\ \sum_{j=1}^{k} W_j S_{jx}^2 - \sum_{j=1}^{k} W_j s_{jx}^2 \end{pmatrix},$$

$$\lambda_1 = \frac{\left(\sum_{j=1}^{k} Q_j W_j s_{jx_j}^4\right)\left(\sum_{j=1}^{k} W_j\left(\bar{X}_j - \bar{x}_j\right)\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)\left(\sum_{j=1}^{k} W_j\left(S_{jx}^2 - s_{jx}^2\right)\right)}{\left(\sum_{j=1}^{k} Q_j W_j s_{j_x}^4\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)^2}, \tag{8}$$

$$\lambda_2 = \frac{\left(\sum_{j=1}^{k} W_j\left(S_{jx}^2 - s_{jx}^2\right)\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)\left(\sum_{j=1}^{k} W_j\left(\bar{X}_j - \bar{x}_j\right)\right)}{\left(\sum_{j=1}^{k} Q_j W_j s_{j_x}^4\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)^2}, \tag{9}$$

$$\begin{aligned}\Omega_j = W_j + Q_j W_j \bar{x}_j & \frac{\left(\sum_{j=1}^{k} Q_j W_j s_{jx_j}^4\right)\left(\sum_{j=1}^{k} W_j\left(\bar{X}_j - \bar{x}_j\right)\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)\left(\sum_{j=1}^{k} W_j\left(S_{jx}^2 - s_{jx}^2\right)\right)}{\left(\sum_{j=1}^{k} Q_j W_j s_{jx_j}^4\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)^2} \\ - Q_j W_j s_{jx}^2 & \frac{\left(\sum_{j=1}^{k} W_j\left(S_{jx}^2 - s_{jx}^2\right)\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)\left(\sum_{j=1}^{k} W_j\left(\bar{X}_j - \bar{x}_j\right)\right)}{\left(\sum_{j=1}^{k} Q_j W_j s_{jx}^4\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)^2},\end{aligned} \tag{10}$$

$$\bar{y}_{st(prop1)} = \sum_{j=1}^{k} W_j\left(\bar{y}_j - \bar{s}\right)_j + 2\hat{\beta}_{1(\text{prop1})}\sum_{j=1}^{k} W_j\left(\bar{X}_j - \bar{x}_j\right) + 2\hat{\beta}_{2(\text{prop1})}\sum_{j=1}^{k} W_j\left(S_{jx}^2 - s_{jx}^2\right), \tag{11}$$

$$\hat{\beta}_1 = \frac{\left(\sum_{j=1}^{k} Q_j W_j s_{jx}^4\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{y}_j s_{jx}^2\right)}{\left(\sum_{j=1}^{k} Q_j W_j s_{j_x}^4\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)^2}, \tag{12}$$

$$\hat{\beta}_2 = \frac{\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{y}_j \bar{x}_j\right)}{\left(\sum_{j=1}^{k} Q_j W_j s_{jx}^4\right)\left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j^2\right) - \left(\sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2\right)^2}.$$

2.4. Proposed estimator-II

By following Eichhorn and Hayre (1983), Koyuncu and Kadilar (2013) and Jabeen et al. (2021), we may write as

$$\bar{y}_{st(propII)} = \sum_{j=1}^{k} \Psi_j \bar{Z}_j. \tag{13}$$

$Z_j = S + Y_j$, where Y_j is the real value of the sensitive quantitative variable and S is the scrambling variables whose distribution is assumed to be known. Where Ψ_j are the weights minimize distance measure,

$$\sum_{j=1}^{k}\frac{\left(\Psi_j - W_j\right)^2}{Q_j W_j}, \tag{14}$$

and satisfying the calibration constraints

$$\sum_{j=1}^{k}\Psi_j \bar{x}_j = \sum_{j=1}^{k} W_j \bar{X}_j, \tag{15}$$

$$\sum_{j=1}^{k}\Psi_j s_{jx}^2 = \sum_{j=1}^{k} W_j S_{jx}^2, \tag{16}$$

$$\sum_{j=1}^{k}\Psi_j = \sum_{j=1}^{k} W_j. \tag{17}$$

Minimize distance measure, given in (18) by using (15), (16) and (17), the Lagrange function is given by

$$D = \frac{\left(\Psi_j - W_j\right)^2}{Q_j W_j} - 2\lambda_1\left(\Psi_j \bar{x}_j - W_j \bar{X}_j\right) - 2\lambda_2\left(\Psi_j s_{jx}^2 - W_j S_{jx}^2\right) - 2\lambda_3\left(\Psi_j - W_j\right), \tag{18}$$

setting $\frac{\partial D}{\partial \Psi_j} = 0,$ we have $\frac{\partial D}{\partial \Psi_j} = \frac{2\left(\Psi_j - W_j\right)}{Q_j W_j} - 2\lambda_1 \bar{x}_j - 2\lambda_2 s_{jx}^2 - 2\lambda_3 = 0,$

$$\Psi_j = W_j + Q_j W_j\left(\lambda_1 \bar{x}_j + \lambda_2 s_{jx}^2 + \lambda_3\right). \tag{19}$$

Putting the value of Ψ_j in (15), (16) and (17), we get

$$\begin{pmatrix} \sum_{j=1}^{k} Q_j W_j \bar{x}_j^2 & \sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2 & \sum_{j=1}^{k} Q_j W_j \bar{x}_j \\ \sum_{j=1}^{k} Q_j W_j \bar{x}_j s_{jx}^2 & \sum_{j=1}^{k} Q_j W_j s_{j_x}^4 & \sum_{j=1}^{k} Q_j W_j s_{jx}^2 \\ \sum_{j=1}^{k} Q_j W_j \bar{x}_j & \sum_{j=1}^{k} Q_j W_j s_{jx}^2 & \sum_{j=1}^{k} Q_j W_j \end{pmatrix}\begin{pmatrix}\lambda_1 \\ \lambda_2 \\ \lambda_3\end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{k} W_j\left(\bar{X}_j - \bar{x}_j\right) \\ \sum_{j=1}^{k} W_j\left(S_{jx}^2 - s_{jx}^2\right) \\ 0 \end{pmatrix}.$$

Solving the system of equation, we obtain

$$\lambda_1 = \frac{A_1}{B}, \lambda_2 = \frac{A_2}{B}, \lambda_1 = \frac{A_3}{B}. \tag{20}$$

$$A_1 = \left(\sum_{j=1}^{k} W_j\left(\bar{X}_j - \bar{x}_j\right)\right)\left(\left(\sum_{j=1}^{k} Q_j W_j\right)\left(\sum_{j=1}^{k} Q_j W_j s_{jx}^4\right) - \left(\sum_{j=1}^{k} Q_j W_j s_{jx}^2\right)^2\right) + \left(\sum_{j=1}^{k} W_j (S_j^2 - s_j^2)\right),$$

$$\mathrm{A}_2 = \left(\sum_{j=1}^{k} \mathrm{W_j}\left(\mathrm{S_j^2} - \mathrm{s_j^2}\right)\right)\left(\left(\sum_{j=1}^{k} \mathrm{Q_j W_j}\right)\left(\sum_{j=1}^{k} \mathrm{Q_j W_j \bar{x}_j^2}\right) - \left(\sum_{j=1}^{k} \mathrm{Q_j W_j \bar{x}_j}\right)^2\right)$$

$$- \left(\sum_{j=1}^{k} \mathrm{W_j}\left(\mathrm{\bar{X}_j} - \mathrm{\bar{x}_j}\right)\right)\left(\left(\sum_{j=1}^{k} \mathrm{Q_j W_j \bar{x}_j s_{jx}^2}\right)\left(\sum_{j=1}^{k} \mathrm{Q_j W_j}\right) - \left(\sum_{j=1}^{k} \mathrm{Q_j W_j \bar{x}_j}\right)\left(\sum_{j=1}^{k} \mathrm{Q_j W_j s_{jx}^2}\right)\right),$$

$$A_3=\left(\sum_{j=1}^{k}W_j\left(\bar{X}_j-\bar{x}_j\right)\right)\left(\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\right)\left(\sum_{j=1}^{k}Q_jW_js_{j_x}^4\bar{x}\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j\right)\left(\sum_{j=1}^{k}Q_jW_js_{j_x}^4\right)\right)$$
$$+\left(\sum_{j=1}^{k}W_j\left(S_j^2-s_j^2\right)\right)\left(\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j^2\right)\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\right)\right),$$

$$B=\left(\sum_{j=1}^{k}Q_jW_j\right)\left(\sum_{j=1}^{k}Q_jW_js_{j_x}^4\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j^2\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j\right)^2\left(\sum_{j=1}^{k}Q_jW_js_{jx}^4\right)$$
$$-\left(\sum_{j=1}^{k}Q_jW_j\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right)^2-\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\right)^2\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j^2\right)$$
$$+2\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j\right)\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right).$$

The value of λ_1, λ_2 and λ_3 putting in (20), we get

$$\bar{y}_{st(propII)}=\sum_{j=1}^{k}W_j\left(\bar{y}_j-\bar{s}_j\right)+2\hat{\beta}_{1(propII)}\sum_{j=1}^{k}W_j\left(\bar{X}_j-\bar{x}_j\right)+2\hat{\beta}_{2(propII)}\sum_{j=1}^{k}W_j\left(S_{jx}^2-s_{jx}^2\right), \quad (21)$$

$$\hat{\beta}_{1(propII)}=\frac{A_4}{B},\hat{\beta}_{2(propII)}=\frac{A_5}{B},$$

$$A_4=\left(\sum_{j=1}^{k}Q_jW_j\bar{y}_j\bar{x}_j\right)\left(\left(\sum_{j=1}^{k}Q_jW_j\right)\left(\sum_{j=1}^{k}Q_jW_js_{jx}^4\right)-\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\right)^2\right)$$
$$-\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\bar{y}_j\right)\left(\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right)\left(\sum_{j=1}^{k}Q_jW_j\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j\right)\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\right)\right)$$
$$+\left(\sum_{j=1}^{k}Q_jW_j\bar{y}_j\right)\left(\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j\right)\left(\sum_{j=1}^{k}Q_jW_js_{jx}^4\right)\right),$$

$$A_5=\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\bar{y}_j\right)\left(\left(\sum_{j=1}^{k}Q_jW_j\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j^2\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j\right)^2\right)$$
$$+\left(\sum_{j=1}^{k}Q_jW_j\bar{y}_j\bar{x}_j\right)\left(\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j\right)\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\right)-\left(\sum_{j=1}^{k}Q_jW_j\right)\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_js_{jx}^2\right)\right)$$
$$+\left(\sum_{j=1}^{k}Q_jW_j\bar{y}_j\bar{x}_j\right)\left(\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j\right)\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\bar{y}_j\right)-\left(\sum_{j=1}^{k}Q_jW_j\bar{x}_j^2\right)\left(\sum_{j=1}^{k}Q_jW_js_{jx}^2\right)\right).$$

3. Results

To examine the performance of the suggestion estimator we produce eight distinctive simulated population where x_{ji}^{*} and y_{hi}^{*} values are form various distribution, as given in Tables 1 and 2. To get different level of relationship among investigation and helping variable we apply few transformations given Table 3. Mean square errors and relative efficiencies are given in Tables 4 and 5. Each populace comprises of three strata having 5 units. We choose n_j= 2, 3, 4 units from every stratum separately along these line, respectively, we get $\binom{5}{2}\binom{5}{3}\binom{5}{4}=500$ samples. We have used $\rho_{xy1}=0.5,\ \rho_{xy2}=0.7,\ \rho_{xy3}=0.9.$ We computed empirical mean square error and relative efficiency using following formulas:

$$MSE\ \left(\bar{y}_{st}(\alpha)\right)=\frac{\sum_{k=1}^{\binom{N}{n}}\left(\bar{y}st(\alpha)-\bar{Y}\right)^2}{\binom{N}{n}},\ \alpha=\text{prop-I, prop-II},\ \ PRE\ =\frac{MSE\left(\bar{y}_{st}(\text{prop-I})\right)}{MSE\left(\bar{y}_{st}(\text{prop-II})\right)}\times 100.$$

Table 1 Parameters and distribution of study, auxiliary and scrambler variables

No.	Parameters and distribution of study variable	Parameters and distribution of auxiliary variable	Parameters and distribution of Scrambler variable
1	$f\left(y^*_{ji}\right)=\frac{1}{\Gamma(1.5)}y^{*1.5-1}_{ji}\exp^{-y^*_{ji}}$	$f\left(x^*_{ji}\right)=\frac{1}{\Gamma(0.3)}y^{*0.3-1}_{ji}\exp^{-x^*_{ji}}$	$f\left(s^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-s^*_{ji}}{2}}$
2	$f\left(y^*_{ji}\right)=\frac{1}{\Gamma(0.3)}y^{*0.3-1}_{ji}\exp^{-y^*_{ji}}$	$f\left(x^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-x^{*2}_{ji}}{2}}$	$f\left(s^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-s^*_{ji}}{2}}$
3	$f\left(y^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-y^*_{ji}}{2}}$	$f\left(x^*_{ji}\right)=\frac{1}{\Gamma(0.3)}y^{*0.3-1}_{ji}\exp^{-x^*_{ji}}$	$f\left(s^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-s^{*2}_{ji}}{2}}$
4	$f\left(y^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-y^{*2}_{ji}}{2}}$	$f\left(x^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-x^*_{ji}}{2}}$	$f\left(s^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-s^{*2}_{ji}}{2}}$

Table 2 Parameters and distribution of study, auxiliary and scrambler variables

No.	Parameters and distribution of study variable	Parameters and distribution of auxiliary variable	Parameters and distribution of Scrambler variable
1	$f\left(y^*_{ji}\right)=\frac{1}{\Gamma(1.5)}y^{*1.5-1}_{ji}\exp^{-y^*_{j}}$	$f\left(x^*_{ji}\right)=\frac{1}{\Gamma(0.3)}y^{*0.3-1}_{ji}\exp^{-x^*_{ji}}$	$f\left(s^*_{ji}\right)=\frac{1}{\Gamma(0.5)}s^{*0.5-1}_{ji}\exp^{-s^*_{ji}}$
2	$f\left(y^*_{ji}\right)=\frac{1}{\Gamma(0.3)}y^{*0.3-1}_{ji}\exp^{-y}$	$f\left(x^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-x^{*2}_{ji}}{2}}$	$f\left(s^*_{ji}\right)=\frac{1}{\Gamma(0.5)}s^{*0.5-1}_{ji}\exp^{-s^*_{ji}}$
3	$f\left(y^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-y^*_{ji}}{2}}$	$f\left(x^*_{ji}\right)=\frac{1}{\Gamma(0.3)}y^{*0.3-1}_{ji}\exp^{-x^*_{ji}}$	$f\left(s^*_{ji}\right)=\frac{1}{\Gamma(1.5)}s^{*1.5-1}_{ji}\exp^{-s^*_{ji}}$
4	$f\left(y^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-y^{*2}_{ji}}{2}}$	$f\left(x^*_{ji}\right)=\frac{1}{\sqrt{2\Pi}}\exp^{\frac{-x^*_{ji}}{2}}$	$f\left(s^*_{ji}\right)=\frac{1}{\Gamma(1.5)}s^{*1.5-1}_{ji}\exp^{-s^*_{ji}}$

Table 3 Properties of j^{th} stratum

Strata	Study variable	Scrambler variable	Auxiliary variable
1. Stratum	$y_{1i} = 50 + y_{1i}^*$	$s_{1i} = 50 + s_{1i}^*$	$x_{1i} = 50 + \sqrt{\left(1-\rho_{xy1}^2\right)x_{1i}^*} + \rho_{xy1}\frac{s_{1x}}{s_{1y}}y_{1i}^*$
2. Stratum	$y_{2i} = 150 + y_{2i}^*$	$s_{2i} = 150 + s_{2i}^*$	$x_{2i} = 100 + \sqrt{\left(1-\rho_{xy2}^2\right)x_{2i}^*} + \rho_{xy2}\frac{s_{2x}}{s_{2y}}y_{2i}^*$
3. Stratum	$y_{3i} = 100 + y_{3i}^*$	$s_{3i} = 100 + s_{3i}^*$	$x_{3i} = 300 + \sqrt{\left(1-\rho_{xy3}^2\right)x_{3i}^*} + \rho_{xy3}\frac{s_{3x}}{s_{3y}}y_{3i}^*$

Table 4 Mean Square Error for Proposed estimator I and II

	Population No.	No. Proposed estimator I MSE $\bar{y}_{st(propI)}$	No. Proposed estimator II MSE $\bar{y}_{st(propII)}$
Using Normal Distribution	1	253952085	180068134
	2	711627033	675652741
	3	57244142	51072864
	4	630224426	595668390
Using Gamma Distribution	5	130347682	127762234
	6	4949554	4067087
	7	54572441	48257210
	8	791389114	725066450

Table 5 Relative efficiency for the proposed estimator I and II

	Population No.	PRE
Normal distribution	1	141.03
	2	105.32
	3	112.05
	4	105.80
Gamma distribution	5	102.07
	6	121.67
	7	113.00
	8	109.17

4. Discussion

We produce eight distinctive simulated populations as given in Table 1, Table 2 and Table 3. The mean square errors are presented in Table 4 and their percentage relative efficiency is given in Table 5. We have compared both proposed estimators under randomized response technique. From Table 5, we see that proposed estimator-II is more efficient than proposed estimator-I.

5. Conclusion

We have proposed two calibration estimators using Tracy et al. (2003), Eichhorn and Hayre (1983) and Jabeen et al. (2021). We have used different calibration contraints and same distance measures in order to compare their efficiency. We conclude that estimator II performs better than estimator I as it utilize an extra calibration constraint $\sum_{j=1}^{k}\Psi_j = \sum_{j=1}^{k} W_j$ which helps to increae the efficiency of the calibration estimator.

References

Eichhorn BH, Hayre LS. Scrambled randomized response methods for obtaining sensitive quantitative data. J Stat Plan Inference. 1983; 7(4): 307-316.

Hansen MH, Hurwitz WN. The problem of non-response in sample surveys. J Am Stat Assoc. 1946; 41(236): 517-529.

Jabeen R, Sanaullah A, Hanif M, Zaka A. Two-exponential estimators for estimating population mean. Aims Math. 2021; 6(1): 737-753.

Kadilar C, Cingi H. Ratio estimators in stratified random sampling. Biometrical J. 2003; 45(2): 218-225.

Koyuncu N, Kadilar C. Calibration estimator using different distance measures in stratified random sampling. Int J Mod Eng Res. 2013; 3(1): 415-419.

Mangat NS, Singh R. An alternative randomized response procedure. Biometrika. 1990; 77(2): 439-442.

Mangat NS. An improved randomized response strategy. J R Stat Soc Series B Stat Methodol. 1994; 56(1): 93-95.

Pollock KH, Bek Y. A comparison of three randomized response models for quantitative data. J Am Stat Assoc. 1976; 71(356): 884-886.

Shabbir J, Gupta S. Improved ratio estimators in stratified sampling. AM J Math-S. 2005; 25(3-4): 293-311.

Tracy D, Singh S, Arnab R. Note on calibration in stratified and double sampling. Surv Meth. 2003; 29(1): 99-104.

Warner SL. Randomized response: A survey technique for eliminating evasive answer bias. J Am Stat Assoc. 1965; 60: 63-69.

Zaman T. Efficient estimators of population mean using auxiliary attribute in stratified random sampling. Adv Appl Stat. 2019; 56(2): 153-171.

Zaman T. An efficient exponential estimator of the mean under stratified random sampling. Math Pop Stud. 2021; 28(2): 104-121.

Zaman T, Bulut H. Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling. Commun Stat-Theory Methods. 2020; 49(14): 3407-3420.

Zaman T, Kadilar C. On estimating the population mean using auxiliary character in stratified random sampling. J Stat Manag Syst. 2020; 23(8): 1415-1426.

Zaman T, Kadilar C. Exponential ratio and product type estimators of the mean in stratified two-phase sampling. AIMS Math. 2021; 6(5): 4265-4279.