

# **Separate Ratio Estimator Using Calibration Approach for the Population mean in Stratified Random Sampling**

## **Abstract**

In this paper we investigate the problem of estimation of finite population mean using auxiliary information. The aim of this paper is to modify a separate ratio estimator by using calibration approach for the population mean using stratified random sampling. We use the calibration approach to improve the precision of estimate of the population parameter by incorporating auxiliary information. We have used the technique of chi-square distance function to minimize the distance between the original design weight and the new design weight. We used two constraints called calibration equations to find the new calibration weights. We use the Taylor's linearization technique to find the variance or mean squared error of the proposed estimators up to the first degree of approximation. We conducted a simulation study to assess the efficiency of the proposed calibration estimator based on an artificially generated data set. It has been shown that the proposed estimator is more efficient than Sinha et al. (2016) and the usual separate ratio estimator. To judge the merits of the proposed estimators, an empirical study was also conducted which shows that the proposed estimator performs better than the relevant estimators.

*Key words: Calibration estimator; stratified random sampling; separate ratio estimator; Taylor's linearization technique.*

## **1. Introduction**

In survey sampling, we use auxiliary information in a number of ways to improve our estimator. With the help of auxiliary information, we have some well-known techniques such as ratio-type estimation method and regression-type estimation method. Deville and Sarndal (1992) first introduced the calibration approach for the population total. The calibration approach is a method that uses auxiliary information to adjust the original design weights to improve the accuracy of the estimator of the population parameter. Calibration is used to improve the estimator by modifying the sample weights, and we choose calibration weights to minimize a given distance measure while satisfying the constraints on the auxiliary variable. Some research using the calibration approach was done by Wu and Sitter (2001). Tracy et al. (2003) introduced a pair of constraints that use first and second-order moments of the auxiliary variables to obtain calibration weights for estimating the population mean in stratified sampling. Rarrel and Singh (2005), Arnab and Singh (2005), and Estavo and Sarndal (2006) have also given some calibration estimators. Kim et al. (2007) proposed a calibration approach for the ratio estimators in stratified sampling design. Koyuncu and Kadilar (2013) proposed estimators using the population mean calibration approach with stratified random sampling. Sinha et al. (2016) introduced the calibration approach estimation of mean in stratified sampling. In this paper, we proposed a calibration approach for separate ratio estimators that uses stratified sampling in the case of a known value of the population mean of the auxiliary character in each stratum. The properties of the proposed estimator have been obtained and a

comparative study of the proposed estimator with relevant estimators have been performed using a simulation study and also for a real data set.

## 2. Procedures and Notations

Consider a finite population  $U:\{1,2,\dots,N\}$  of size  $N$ . Let  $Y$  and  $X$  be the study variable and auxiliary variable. Let the population of size  $N$  be divided into  $L$  homogeneous groups called strata such that the  $h^{th}$  stratum contains of  $N_h$  units and a sample of size  $n_h$  is drawn using simple random sampling

without replacement (SRSWOR) method such that  $\sum_{h=1}^L N_h = N$  and  $\sum_{h=1}^L n_h = n$ , where  $n$  is the total

sample size. Let  $y_{hi}$  be the value of  $i^{th}$  unit of the study variable in the  $h^{th}$  stratum, where  $i=1,2,\dots,n_h$ . Let  $X_{hi}$  be the value of  $i^{th}$  unit of the auxiliary variable in the  $h^{th}$  stratum. For the  $h^{th}$  stratum,  $W_h = N_h / N$  be the stratum weight.

Under the stratified sampling scheme, the unbiased estimator of the population mean is given by

$$\bar{y}_{st} = \sum_{h=1}^H W_h \bar{y}_h$$

(1)

where  $\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$  is the sample mean of study variable  $Y$  in the  $h^{th}$  stratum.

The variance of the estimator  $\bar{y}_{st}$  is given by

$$V(\bar{y}_{st}) = \sum_{h=1}^H W_h^2 \left( \frac{1-f_h}{n_h} \right) S_{hy}^2$$

(2)

where  $S_{hy}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h)^2$  is population mean square for  $h^{th}$  stratum.

The separate ratio estimator  $\bar{y}_{sr}$  of population mean  $\bar{Y}$  is given by

$$\bar{y}_{sr} = \sum_{h=1}^H W_h \left( \frac{\bar{y}_h}{\bar{x}_h} \right) \bar{X}_h$$

(3)

where  $\bar{x}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{hi}$  is the  $h^{th}$  sample mean of auxiliary character  $X$  in the  $h^{th}$  stratum and

$\bar{X}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} X_{hi}$  is the population mean for the  $h^{th}$  stratum.

The variance of  $\bar{y}_{sr}$  to the first order of approximation is given by

$$V(\bar{y}_{sr}) = \sum_{h=1}^H W_h^2 \left( \frac{1-f_h}{n_h} \right) \bar{Y}_h^2 [C_{hy}^2 + C_{hx}^2 - 2\rho_{hxy} C_{hx} C_{hy}]$$

(4)

where  $f_h = n_h / N_h$ ,  $\bar{Y}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{hi}$ ,  $C_{hx}^2 = S_{hx}^2 / \bar{X}_h^2$ ,  $C_{hy}^2 = S_{hy}^2 / \bar{Y}_h^2$ , and  $\rho_{hxy} = S_{hxy} / S_{hx} S_{hy}$  and  $S_{hx}^2$ ,  $S_{hy}^2$ ,  $S_{hxy}$ ,  $\rho_{hxy}$  denote the population mean square of y, x, covariance of y and x, and population correlation coefficient between y and x respectively in the  $h^{th}$  stratum.

An estimator for the variance of the  $\bar{y}_{sr}$  is given by

$$\hat{v}(\bar{y}_{sr}) = \sum_{h=1}^H W_h^2 \left( \frac{1-f_h}{n_h} \right) [s_{hy}^2 + r_h^2 s_{hx}^2 - 2r_h s_{hxy}] \quad (5)$$

where  $r_h = \frac{\bar{y}_h}{\bar{x}_h}$ ,  $s_{hy}^2 = \frac{1}{n_h-1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$ ,  $s_{hx}^2 = \frac{1}{n_h-1} \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2$ , and

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

and  $s_{hx}^2$ ,  $s_{hy}^2$ ,  $s_{hxy}$ ,  $r_h$  denotes the sample mean square of x, y, covariance of y and x, and the ratio of two sample mean respectively in the  $h^{th}$  stratum.

The calibration estimator defined by Sinha et al. (2016), is given by

$$\bar{y}_{sr}(n) = \sum_{h=1}^H \Omega_h \bar{y}_h \quad (6)$$

where  $\Omega_h$  are calibrated weights used to minimize the chi-square distance function, given as

$$\sum_{h=1}^H \frac{(\Omega_h - W_h)^2}{W_h Q_h}$$

where  $Q_h$  are suitably chosen positive quantity which determines the type of the estimator, subject to the following constraints

$$\sum_{h=1}^H \Omega_h \bar{x}_h = \sum_{h=1}^H W_h \bar{X}_h \quad (7)$$

and

$$\sum_{h=1}^H \Omega_h = 1 \quad (8)$$

The Lagrange function utilizing calibration constraints and chi-square distance function is given by

$$\Phi = \sum_{h=1}^H \frac{(\Omega_h - W_h)^2}{W_h Q_h} - 2\lambda_1 \left( \sum_{h=1}^H \Omega_h \bar{x}_h - \sum_{h=1}^H W_h \bar{X}_h \right) - 2\lambda_2 \left( \sum_{h=1}^H \Omega_h - 1 \right) \quad (9)$$

where  $\lambda_1$  and  $\lambda_2$  are Lagrange multipliers.

Now differentiate (9) with respect to  $\Omega_h$  and equating to zero, we get the calibrated weight given as

$$\Omega_h = W_h + W_h Q_h \bar{x}_h \left[ \frac{\sum_{h=1}^H W_h (\bar{X}_h - \bar{x}_h) \left( \sum_{h=1}^H W_h Q_h \right)}{\left( \sum_{h=1}^H W_h Q_h \bar{x}_h^2 \right) \left( \sum_{h=1}^H W_h Q_h \right) - \left( \sum_{h=1}^H W_h Q_h \bar{x}_h \right)^2} \right] ; \text{ for } h=1,2,\dots,L$$

$$+ W_h Q_h \left[ \frac{-\sum_{h=1}^H W_h (\bar{X}_h - \bar{x}_h) \left( \sum_{h=1}^H W_h Q_h \bar{x}_h \right)}{\left( \sum_{h=1}^H W_h Q_h \bar{x}_h^2 \right) \left( \sum_{h=1}^H W_h Q_h \right) - \left( \sum_{h=1}^H W_h Q_h \bar{x}_h \right)^2} \right]$$

(10)

The calibration approach provides an estimator of variance given by

$$\hat{V}[\bar{y}_{sr}(n)] = \sum_{h=1}^H \Omega_h^2 \left( \frac{1-f_h}{n_h} \right) \frac{1}{(n_h-1)} \sum_{i=1}^{n_h} \hat{e}_{hi}^2 \quad (11)$$

where  $\hat{e}_{hi} = (y_{hi} - \bar{y}_h) - \hat{\beta}_h (x_{hi} - \bar{x}_h)$  denotes the residual term in the  $h^{th}$  stratum.

### 3. Proposed estimators

We propose separate ratio estimator using calibration approach for the population mean using stratified random sampling which is given by

$$\bar{y}_{sr}(c) = \sum_{h=1}^H W_h^* \left( \frac{\bar{y}_h}{\bar{x}_h} \right) \bar{X}_h$$

(12)

where the new weights  $W_h^*$  are chosen such that the chi-square distance function

$$\sum_{h=1}^H \frac{(W_h^* - W_h)^2}{W_h Q_h}$$

is minimum depending on the condition

$$\sum_{h=1}^H W_h^* \bar{x}_h = \sum_{h=1}^H W_h \bar{X}_h \quad \text{and} \quad \sum_{h=1}^H W_h^* = 1$$

(13)

The choice of quantity  $Q_h$  determines the form of the estimator.

Now, let us consider the Lagrange function given by

$$\Psi = \sum_{h=1}^H \frac{(W_h^* - W_h)^2}{W_h Q_h} - 2\nu_1 \left( \sum_{h=1}^H W_h^* \bar{x}_h - \sum_{h=1}^H W_h \bar{X}_h \right) - 2\nu_2 \left( \sum_{h=1}^H \Omega_h = 1 \right)$$

(14)

where  $\nu_1$  and  $\nu_2$  are Lagrange's multipliers. Now differentiating (14) with respect to  $W_h^*$  and equal to zero, we get the new calibration weight as

$$W_h^* = W_h + W_h Q_h \bar{x}_h \left[ \frac{\sum_{h=1}^H W_h (\bar{X}_h - \bar{x}_h) \left( \sum_{h=1}^H W_h Q_h \right)}{\left( \sum_{h=1}^H W_h Q_h \bar{x}_h^2 \right) \left( \sum_{h=1}^H W_h Q_h \right) - \left( \sum_{h=1}^H W_h Q_h \bar{x}_h \right)^2} \right]$$

$$+ W_h Q_h \left[ \frac{-\sum_{h=1}^H W_h (\bar{X}_h - \bar{x}_h) \left( \sum_{h=1}^H W_h Q_h \bar{x}_h \right)}{\left( \sum_{h=1}^H W_h Q_h \bar{x}_h^2 \right) \left( \sum_{h=1}^H W_h Q_h \right) - \left( \sum_{h=1}^H W_h Q_h \bar{x}_h \right)^2} \right]$$

(15)

Now substituting the value of  $W_h^*$  in the equation (12), we get the calibration estimator as

$$\bar{y}_{sr}(c) = \sum_{h=1}^H W_h \left( \frac{\bar{y}_h}{\bar{x}_h} \right) \bar{X}_h + \hat{\beta}_{st}^* \left( \bar{X} - \sum_{h=1}^H W_h \bar{x}_h \right)$$

(16)

Where

$$\hat{\beta}_{st}^* = \frac{\left( \sum_{h=1}^H Q_h W_h \right) \left( \sum_{h=1}^H Q_h W_h \bar{y}_h \bar{X}_h \right) - \left( \sum_{h=1}^H Q_h W_h \left( \frac{\bar{y}_h}{\bar{x}_h} \right) \bar{X}_h \right) \left( \sum_{h=1}^H Q_h W_h \bar{x}_h \right)}{\left( \sum_{h=1}^H Q_h W_h \right) \left( \sum_{h=1}^H Q_h W_h \bar{x}_h^2 \right) - \left( \sum_{h=1}^H Q_h W_h \bar{x}_h \right)^2}$$

(17)

### 3.1 MSE or Variance of the proposed estimator

The variance estimate is necessary to estimate the accuracy of the survey estimates. In most scenarios, the unbiased estimates of the nonlinear parametric functions are not available. So we use the Taylor linearization method to find the variance of nonlinear estimators that can be used for complicated survey designs. To find the variance of estimators, we define the following notations and for simplicity we take  $Q_h=1$ . Let us assume

$$t_1 = \sum_{h=1}^H W_h \left( \frac{\bar{y}_h}{\bar{x}_h} \right) \bar{X}_h, \quad t_2 = \sum_{h=1}^H Q_h W_h \bar{y}_h \bar{X}_h, \quad t_3 = \sum_{h=1}^H Q_h W_h \left( \frac{\bar{y}_h}{\bar{x}_h} \right) \bar{X}_h,$$

$$t_4 = \sum_{h=1}^H Q_h W_h \bar{x}_h, \quad t_5 = \sum_{h=1}^H Q_h W_h \bar{x}_h^2, \quad \text{and} \quad t_6 = \sum_{h=1}^H W_h \bar{x}_h$$

(18)

By using the values of  $t_1, t_2, t_3, t_4, t_5$ , and  $t_6$  from (18), equation (16) becomes

$$\bar{y}_{sr}(c) = t_1 + \left( \frac{t_2 - t_3 t_4}{t_5 - t_4^2} \right) (\bar{X} - t_6)$$

(19)

Let,

$$\bar{y}_{sr}(c) = T = F(t_1, t_2, t_3, t_4, t_5, t_6)$$

(20)

For further calculation, we use some approximations for the  $h^{th}$  stratum. For large sample approximation, we assume

$$\bar{y}_h = \bar{Y}_h(1 + \varepsilon_{h0}) \text{ and } \bar{x}_h = \bar{X}_h(1 + \varepsilon_{h1}) \text{ such that } E(\varepsilon_{h0}) = E(\varepsilon_{h1}) = 0$$

$$\text{Then } E(\varepsilon_{h0}^2) = \left( \frac{1 - f_h}{n_h} \right) C_{hy}^2, E(\varepsilon_{h1}^2) = \left( \frac{1 - f_h}{n_h} \right) C_{hx}^2, E(\varepsilon_{h0}\varepsilon_{h1}) = \left( \frac{1 - f_h}{n_h} \right) \rho_{hxy} C_{hx} C_{hy} \text{ and } f_h = \frac{1}{n_h} - \frac{1}{N_h}$$

(21)

Now we have,

$$\begin{aligned} E(t_1) &= \sum_{h=1}^H W_h \bar{Y}_h (1 + E(\varepsilon_1^2) - E(\varepsilon_0\varepsilon_1)) = \theta_1, E(t_2) = \sum_{h=1}^H W_h Q_h \bar{Y}_h \bar{X}_h = \theta_2 \\ E(t_3) &= \sum_{h=1}^H W_h Q_h \bar{Y}_h (1 + E(\varepsilon_1^2) - E(\varepsilon_0\varepsilon_1)) = \theta_3, E(t_4) = \sum_{h=1}^H W_h Q_h \bar{X}_h = \theta_4 \\ E(t_5) &= \sum_{h=1}^H W_h Q_h \bar{X}_h^2 (1 + E(\varepsilon_1^2)) = \theta_5, E(t_6) = \sum_{h=1}^H W_h \bar{X}_h = \theta_6 \end{aligned}$$

(22)

Now differentiating the function  $\bar{y}_{sr}(c) = F(t_1, t_2, t_3, t_4, t_5, t_6)$  with respect to  $t_1, t_2, t_3, t_4, t_5$ , and  $t_6$ , putting their expected value in the function, we get

$$\begin{aligned} g_1 &= \frac{\partial}{\partial t_1} F(t_1, t_2, t_3, t_4, t_5, t_6) \Big|_{\hat{\theta}=\theta} = 1, g_2 = \frac{\partial}{\partial t_2} F(t_1, t_2, t_3, t_4, t_5, t_6) \Big|_{\hat{\theta}=\theta} = 0 \\ g_3 &= \frac{\partial}{\partial t_3} F(t_1, t_2, t_3, t_4, t_5, t_6) \Big|_{\hat{\theta}=\theta} = 0, g_4 = \frac{\partial}{\partial t_4} F(t_1, t_2, t_3, t_4, t_5, t_6) \Big|_{\hat{\theta}=\theta} = 0 \\ g_5 &= \frac{\partial}{\partial t_5} F(t_1, t_2, t_3, t_4, t_5, t_6) \Big|_{\hat{\theta}=\theta} = 0, g_6 = \frac{\partial}{\partial t_6} F(t_1, t_2, t_3, t_4, t_5, t_6) \Big|_{\hat{\theta}=\theta} = (-1) \left( \frac{t_2 - t_3 t_4}{t_5 - t_4^2} \right) \Big|_{\hat{\theta}=\theta} \end{aligned}$$

(23)

By using (23), we can write equation (19) as

$$T = \bar{Y} + g_1(t_1 - \theta_1) + g_2(t_2 - \theta_2) + g_3(t_3 - \theta_3) + g_4(t_4 - \theta_4) + g_5(t_5 - \theta_5) + g_6(t_6 - \theta_6)$$

(24)

Putting the values of calculated  $g_1, g_2, g_3, g_4, g_5$ , and  $g_6$  in equation (24), we get

$$T = \bar{Y} + g_1(t_1 - \theta_1) + g_6(t_6 - \theta_6)$$

Now variance of  $T$  is given as

$$\begin{aligned} V(T) &= V(\bar{Y}) + V(g_1(t_1 - \theta_1)) + V(g_6(t_6 - \theta_6)), \\ V(T) &= V(\bar{Y}) + g_1^2 V(t_1 - \theta_1) + g_6^2 V(t_6 - \theta_6), \\ V(T) &= V(t_1) + g_6^2 V(t_6) + 2g_6 \text{Cov}(t_1, t_6) \end{aligned}$$

(25)

where  $V(t_1) = \sum_{h=1}^H W_h^2 \left( \frac{1-f_h}{n_h} \right) [S_{hy}^2 + R_h^2 S_{hx}^2 - 2R_h S_{hxy}]$ ,  $V(t_6) = \sum_{h=1}^H W_h^2 \left( \frac{1-f_h}{n_h} \right) S_{hx}^2$ , and

$$\text{Cov}(t_1, t_6) = \sum_{h=1}^H W_h^2 \bar{Y}_h \bar{X}_h [E(\varepsilon_0 \varepsilon_1) - E(\varepsilon_1^2)]$$

After substituting the values of the above expressions into an equation (25), we get the approximate MSE or Variance of the proposed calibrated estimator. We take  $Q_h=1$  throughout our study.

#### 4. Simulation Study

In the sampling theory, researchers typically check the estimator's performance by comparing its MSE or variance with respect to the other estimators. The MSE or variance of the estimators relies entirely on the population parameters assumed to be known. But in the practical situation, parametric values are rarely available to be known. To avoid such practical difficulties, one can perform a simulation analysis based on the actual samples drawn from the population. In simulation analysis, a large number of independent samples are drawn from the population and the mean MSE or variance is calculated to compare the efficiency of the estimators. We therefore attempted to conduct a simulation study to verify the performance of the proposed calibration estimator  $\bar{y}_{sr}(c)$ . We perform a simulation study by generating four artificial population from normal distribution with different assumed means and standard deviations. To get different level of correlations between study variables and auxiliary variable, we use transformation from Reddy et al. (2010). Each population consists of four strata. We select a sample from each stratum by using simple random sampling. Here we take a population of size 2650. **Table 1** shows the distribution of generated population to perform the study. **Table 2** represents the transformations which is used to generate correlated random variables.

**Table 1. Population parameters for the  $h^{th}$  stratum**

Strata no.	Stratum Size ( $N_h$ )	Sample size ( $n_h$ )	Distribution of study variable ( $y_{hi}$ )	Distribution of auxiliary variable ( $x_{hi}$ )
I	600	120	$N(22, 2)$	$N(20, 2)$
II	700	140	$N(33, 3)$	$N(25, 3)$

III	550	110	$N(44, 4)$	$N(30, 4)$
IV	800	160	$N(55, 5)$	$N(40, 5)$

**Table 2. Number of strata, correlation coefficient and the transformation**

Strata no.	Correlation between $(y_{hi})$ and $(x_{hi})$	Transformed auxiliary variable $(x_{hi})$
I	$r_{yx1} = 0.85$	$y_1 = r_{yx1}x_{21} + \sqrt{1 - r_{yx1}^2} y_{21}$
II	$r_{yx2} = 0.88$	$y_2 = r_{yx2}x_{22} + \sqrt{1 - r_{yx2}^2} y_{22}$
III	$r_{yx3} = 0.80$	$y_3 = r_{yx3}x_{23} + \sqrt{1 - r_{yx3}^2} y_{23}$
IV	$r_{yx4} = 0.75$	$y_4 = r_{yx4}x_{24} + \sqrt{1 - r_{yx4}^2} y_{24}$

We finally calculate the Average MSE (AMSE) or variance of the estimators using the following formulas

$$AMSE(\bar{y}_{st}(n)) = \frac{1}{5000} \sum_{l=1}^{5000} (\bar{y}_{st}(n) - \bar{Y})^2; l = 1, 2, \dots, 5000$$

$$AMSE(\bar{y}_{sr}) = \frac{1}{5000} \sum_{l=1}^{5000} (\bar{y}_{sr} - \bar{Y})^2; l = 1, 2, \dots, 5000$$

$$AMSE(\bar{y}_{sr}(c)) = \frac{1}{5000} \sum_{l=1}^{5000} (\bar{y}_{sr}(c) - \bar{Y})^2; l = 1, 2, \dots, 5000$$

**Table 3. Mean Square Error (MSE) and Percent Relative Efficiency (PRE) of Estimator in case of stratified sampling**

Estimator	MSE	PRE
$\bar{y}_{st}(n)$	0.03259475	100
$\bar{y}_{sr}$	0.02264014	203.6183

$\bar{y}_{sr}(c)$	<b>0.01111891</b>	<b>287.6743</b>
-------------------	-------------------	-----------------

## 5. Real Data Application

To examine the performance of the proposed calibration estimator, we performed a simulation study. empirical study on real data set. An empirical study also on real data set also carried out to support the theoretical result. The population data was obtained from the Kaggle website. <https://www.kaggle.com/datasets/darshansavaliya/gujratcropdataanalysis> is the link provided. The average production of Gujarat in 2011 as it changes over the several agricultural seasons in Gujarat is what we are trying to determine. In this dataset, we take into account X as an area and Y as a production in (tonnes) (hectares). Our population is divided into four groups based on the harvest season: Kharif, Rabi, Summer, and Whole Year. **Table 4** provides more information about the population's statistical parameters.

**Table 4. Values of the parameter for the  $h^{th}$  stratum**

Strata No.	Ni	$\bar{Y}_h$	$\bar{X}_h$	$\rho_{y_xh}$
1	292	61016.1	27080.82	0.91
2	163	43096.93	13474.23	0.78
3	112	19500.89	8243.75	0.92
<b>4</b>	<b>100</b>	<b>169716</b>	<b>5560</b>	<b>0.80</b>

**Table 5. Mean Square Error (MSE) and Percent Relative Efficiency (PRE) of the Estimator in case of stratified sampling**

Estimator	MSE	PRE
$\bar{y}_{st}(n)$	1151285003	100
$\bar{y}_{sr}$	715864702	163.381
$\bar{y}_{sr}(c)$	<b>438156732</b>	<b>287.6743</b>

## 5. CONCLUSIONS

A new calibration estimator was proposed to estimate the population mean in a stratified sample design using the information about auxiliary variables. In this paper, we obtained some new optimal stratum weights for which the estimator would achieve a minimum MSE or variance. The performance of the proposed calibration estimator was compared with respect to the stratified sampling calibration estimator

using the information auxiliary variables. The theoretical facts are supported by the empirical study. The empirical study was carried out using both an artificially generated data set and a real data set. **Table 3** and **Table 5** shows that the proposed calibration estimator gives better estimates as compared to the existing calibration estimator.

## REFERENCES

1. Deville JC, Särndal CE. Calibration estimators in survey sampling. *Journal of the American statistical Association*. 1992 Jun 1;87(418):376-82.
2. Kim JM, Sungur EA, Heo TY. Calibration approach estimators in stratified sampling. *Statistics & probability letters*. 2007 Jan 1;77(1):99-103.
3. Koyuncu N, Kadilar C. Calibration estimator using different distance measures in stratified random sampling. *International Journal of Modern Engineering Research*. 2013;3(1):415-9.
4. Nidhi, Sisodia BV, Singh S, Singh SK. Calibration approach estimation of the mean in stratified sampling and stratified double sampling. *Communications in Statistics-Theory and Methods*. 2017 May 19;46(10):4932-42.
5. Tracy DS, Singh SA, Arnab RA. Note on calibration in stratified and double sampling. *Survey Methodology*. 2003;29(1):99-104.
6. Wu C, Sitter RR. A model-calibration approach to using complete auxiliary information from survey data. *Journal of the American Statistical Association*. 2001 Mar 1;96(453):185-93.
7. Farrell PJ, Singh S. MODEL-ASSISTED HIGHER-ORDER CALIBRATION OF ESTIMATORS OF VARIANCE. *Australian & New Zealand Journal of Statistics*. 2005 Sep;47(3):375-83.
8. Arnab R, Singh S. A note on variance estimation for the generalized regression predictor. *Australian & New Zealand Journal of Statistics*. 2005 Jun;47(2):231-4.
9. Estevao VM, Särndal CE. Survey estimates by calibration on complex auxiliary information. *International Statistical Review*. 2006 Aug;74(2):127-47.
10. Reddy MK, Rao KR, Boiroju NK. Comparison of ratio estimators using Monte Carlo simulation. *International Journal of Agriculture and Statistical Sciences*. 2010 Dec 1;6(2):517-27.