Global Journal of Mathematics and Statistics
ISSN: 3078-8927
DOI: https://doi.org/10.61424/gjms.v3i1.930
Journal Homepage: https://bluemarkpublishers.com/index.php/GJME

GJMS
BLUEMARK PUBLISHERS

| RESEARCH ARTICLE

Calibration Estimation in Stratified Random Sampling Using Three Constraints

Namita Srivastava[1] and Ruchi Shakya[2] ✉
[12]*St. John's College, Agra (UP)*
Corresponding Author: Ruchi Shakya, **E-mail**: ruchi5293@gmail.com

| ABSTRACT
Calibration estimation improves the precision of population parameter estimates by incorporating specified auxiliary information. In this article, we have proposed a new calibration estimator for estimating the population mean in stratified sampling, based on three new calibration constraints derived from the coefficient of variation and correlation. A simulation study is conducted to evaluate the performance of the proposed estimator.

| KEYWORDS

Calibration estimation, auxiliary information, stratified sampling, coefficient of variation

| ARTICLE INFORMATION

ACCEPTED: 11 May 2026 **PUBLISHED:** 08 July 2026 **DOI:** https://doi.org/10.61424/gjms.v3i1.930

1. Introduction

The calibration-based estimation method helps improve survey estimates by using auxiliary information to adjust the initial design weights. A calibration estimator uses modified weights, which are known as calibrated weights. These calibrated weights are determined by minimizing a given distance function to the initial design weights, respecting a set of constraints associated with auxiliary information. In survey sampling many authors, such as Deville and Sarndal (1992), Estevao and Sarndal (2000), Arnab and Singh (2005), Farrell and Singh (2005), Kim and Park (2010), etc., defined some calibration estimators using different constraints. In stratified random sampling, a calibration approach is used to get optimum strata weights. Tracy et al. (2003), Kim et al. (2007) and Koyuncu (2012) define some calibration estimators in stratified random sampling. Koyuncu and Kadilar (2016), Clement and Enang (2017), Nidhi et al. (2017), Neha Garg & Meenakshi Pachori (2020) etc., have contributed in the development of calibrated estimators for different population parameters using different calibration constraints under various sampling schemes. Koyuncu and Kadilar (2013) define calibration estimator using different distance measures in stratified sampling. Koyunchu et al. (2015) introduced the calibration estimator by using three calibration constraints in stratified random sampling. Ozgul (2018) also introduce a calibration estimator by using three different calibration constraints in stratified random sampling.

In this article, we suggested a new calibration estimators for estimating population mean with the set of three new calibration constraints by using coefficient of variation and correlation. Section 2 gives some existing estimator in stratified sampling using three constraints defined in calibration estimation. Section 3 contains the proposed estimator in stratified random sampling. In Section 4, a simulation study of the proposed estimator is conducted on an artificial population as well as a real data set to evaluate the performance of the proposed estimator by using R programming software. Sections 5 and 6 give the results and conclusions of the paper.

2. Some Existing Estimators in Stratified Sampling Using Three Constraints

Different calibration estimators have been proposed in the literature for estimating the population mean of a stratified sampling using the auxiliary variable. Some existing calibration estimators using three constraints in stratified sampling are described as follows:

Nursel Koyuncu et al. (2015), introduced a new calibration estimator in stratified random sampling, which is given as

$$\bar{y}_{st(Nk)} = \sum_{h=1}^{L} \Psi_h \bar{y}_h \tag{1}$$

Where Ψ_h are new calibration weights minimizing chi-square distance function

$$\sum_{h=1}^{L} \frac{(\Psi_h - W_h)^2}{Q_h W_h} \tag{2}$$

And the calibration constraints are given in equation (3), (4) and (5)

$$\sum_{h=1}^{L} \Psi_h \bar{x}_h = \sum_{h=1}^{L} W_h \bar{X}_h \tag{3}$$

$$\sum_{h=1}^{L} \Psi_h s_{hx}^2 = \sum_{h=1}^{L} W_h S_{hx}^2 \tag{4}$$

$$\sum_{h=1}^{L} \Psi_h = \sum_{h=1}^{L} W_h \tag{5}$$

The proposed calibrated estimator by Nursel Koyuncu (2015) is expressed as

$$\bar{y}_{st(Nk)} = \sum_{h=1}^{L} W_h \bar{y}_h + \hat{\beta}_{1(Nk)}[\sum_{h=1}^{L} W_h(\bar{X}_h - \bar{x}_h)] + \hat{\beta}_{2(Nk)}[\sum_{h=1}^{L} W_h(S_{hx}^2 - s_{hx}^2)] \tag{6}$$

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

$$A_4 = (\sum_{h=1}^{L} Q_h W_h \bar{x}_h \bar{y}_h)[(\sum_{h=1}^{L} Q_h W_h)(\sum_{h=1}^{L} Q_h W_h s_{hx}^4) - (\sum_{h=1}^{L} Q_h W_h s_{hx}^2)^2] - (\sum_{h=1}^{L} Q_h W_h s_{hx}^2 \bar{y}_h)[(\sum_{h=1}^{L} Q_h W_h s_{hx}^2 \bar{x}_h)(\sum_{h=1}^{L} Q_h W_h) - (\sum_{h=1}^{L} Q_h W_h \bar{x}_h)(\sum_{h=1}^{L} Q_h W_h s_{hx}^2)] + (\sum_{h=1}^{L} Q_h W_h \bar{y}_h)[(\sum_{h=1}^{L} Q_h W_h s_{hx}^2)(\sum_{h=1}^{L} Q_h W_h s_{hx}^2 \bar{x}_h) - (\sum_{h=1}^{L} Q_h W_h \bar{x}_h)(\sum_{h=1}^{L} Q_h W_h s_{hx}^4)]$$

$$A_5 = (\sum_{h=1}^{L} Q_h W_h \bar{x}_h \bar{y}_h)[(\sum_{h=1}^{L} Q_h W_h \bar{x}_h)(\sum_{h=1}^{L} Q_h W_h s_{hx}^2) - (\sum_{h=1}^{L} Q_h W_h)(\sum_{h=1}^{L} Q_h W_h s_{hx}^2 \bar{x}_h)] + (\sum_{h=1}^{L} Q_h W_h s_{hx}^2 \bar{y}_h)[(\sum_{h=1}^{L} Q_h W_h)(\sum_{h=1}^{L} Q_h W_h \bar{x}_h^2) - (\sum_{h=1}^{L} Q_h W_h \bar{x}_h)^2] + (\sum_{h=1}^{L} Q_h W_h \bar{y}_h)[(\sum_{h=1}^{L} Q_h W_h \bar{x}_h)(\sum_{h=1}^{L} Q_h W_h s_{hx}^2 \bar{x}_h) - (\sum_{h=1}^{L} Q_h W_h \bar{x}_h^2)(\sum_{h=1}^{L} Q_h W_h s_{hx}^2)]$$

$$B = (\sum_{h=1}^{L} Q_h W_h)(\sum_{h=1}^{L} Q_h W_h s_{hx}^4)(\sum_{h=1}^{L} Q_h W_h \bar{x}_h^2) - (\sum_{h=1}^{L} Q_h W_h \bar{x}_h)^2(\sum_{h=1}^{L} Q_h W_h s_{hx}^4) - (\sum_{h=1}^{L} Q_h W_h)(\sum_{h=1}^{L} Q_h W_h s_{hx}^2 \bar{x}_h)^2 - (\sum_{h=1}^{L} Q_h W_h s_{hx}^2)^2(\sum_{h=1}^{L} Q_h W_h \bar{x}_h^2) + 2(\sum_{h=1}^{L} Q_h W_h \bar{x}_h)(\sum_{h=1}^{L} Q_h W_h s_{hx}^2)(\sum_{h=1}^{L} Q_h W_h s_{hx}^2 \bar{x}_h)$$

Ozqul (2018), developed a calibrated estimator is given as

$$\bar{y}_{st(O)} = \sum_{h=1}^{L} \Omega_h \bar{y}_h \tag{7}$$

Where Ω_h are the new weights obtained by minimizing the chi square distance function

$$\sum_{h=1}^{L} \frac{(\Omega_h - W_h)^2}{Q_h W_h} \tag{8}$$

And the calibration constraints are given in equation (9), (10) and (11)

$$\sum_{h=1}^{L} W_h = 1 \tag{9}$$

$$\sum_{h=1}^{L} \Omega_h \bar{x}_h = \sum_{h=1}^{L} W_h \bar{X}_h \tag{10}$$

$$\sum_{h=1}^{L} \Omega_h c_{hx} = \sum_{h=1}^{L} W_h C_{hx} \tag{11}$$

The proposed estimator by Ozqul (2018) is expressed as:

$$\bar{y}_{st(O)} = \sum_{h=1}^{L}\left[\bar{y}_h + \hat{\beta}_{1(O)}(\bar{X}_h - \bar{x}_h) + \hat{\beta}_{2(O)}(C_{hx} - c_{hx})\right] \quad (12)$$

Where,

$\hat{\beta}_{1(O)} = \frac{\gamma_1}{\eta}$ and $\hat{\beta}_{2(O)} = \frac{\gamma_2}{\eta}$

$$\gamma_1 = (\sum_{h=1}^{L} W_h Q_h \bar{y}_h)[(\sum_{h=1}^{L} W_h Q_h c_{hx})(\sum_{h=1}^{L} W_h Q_h \bar{x}_h c_{hx}) - (\sum_{h=1}^{L} W_h Q_h \bar{x}_h)(\sum_{h=1}^{L} W_h Q_h c_{hx}^2)] + (\sum_{h=1}^{L} W_h Q_h \bar{x}_h \bar{y}_h)[(\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h c_{hx}^2) - (\sum_{h=1}^{L} W_h Q_h c_{hx})^2] + (\sum_{h=1}^{L} W_h Q_h c_{hx} \bar{y}_h)[(\sum_{h=1}^{L} W_h Q_h \bar{x}_h)(\sum_{h=1}^{L} W_h Q_h c_{hx}) - (\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h \bar{x}_h c_{hx})]$$

$$\gamma_2 = (\sum_{h=1}^{L} W_h Q_h \bar{y}_h)[(\sum_{h=1}^{L} W_h Q_h \bar{x}_h)(\sum_{h=1}^{L} W_h Q_h \bar{x}_h c_{hx}) - (\sum_{h=1}^{L} W_h Q_h c_{hx})(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2)] + (\sum_{h=1}^{L} W_h Q_h \bar{x}_h \bar{y}_h)[(\sum_{h=1}^{L} W_h Q_h \bar{x}_h)(\sum_{h=1}^{L} W_h Q_h c_{hx}) - (\sum_{h=1}^{L} W_h Q_h \bar{x}_h c_{hx})(\sum_{h=1}^{L} W_h Q_h)] + (\sum_{h=1}^{L} W_h Q_h c_{hx} \bar{y}_h)[(\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2) - (\sum_{h=1}^{L} W_h Q_h \bar{x}_h)^2]$$

$$\eta = (\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2)(\sum_{h=1}^{L} W_h Q_h c_{hx}^2) - (\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h \bar{x}_h c_{hx}) - (\sum_{h=1}^{L} W_h Q_h \bar{x}_h^2)(\sum_{h=1}^{L} W_h Q_h c_{hx})^2 - (\sum_{h=1}^{L} W_h Q_h c_{hx}^2)(\sum_{h=1}^{L} W_h Q_h \bar{x}_h)^2 + 2(\sum_{h=1}^{L} W_h Q_h \bar{x}_h)(\sum_{h=1}^{L} W_h Q_h c_{hx})(\sum_{h=1}^{L} W_h Q_h \bar{x}_h c_{hx})$$

3. Proposed Calibration Estimator

In stratified sampling scheme the classical unbiased estimator of the population mean is given by

$$\bar{y}_{st} = \sum_{h=1}^{L} W_h \bar{y}_h \quad (13)$$

Where, $W_h = \frac{N_h}{N}$ is the stratum weights. The calibration estimator for stratified random sampling is defined by Tracy et al. (2003) as

$$\bar{y}_{st}(Tr) = \sum_{h=1}^{L} \Omega_h \bar{y}_h \quad (14)$$

Now following Tracy et al. (2003) and motivated by Nursel Koyuncu et al. (2015) and Ozqul (2018), we introduced a new calibration given as

$$\bar{y}_{st}(N) = \sum_{h=1}^{L} \Omega_h^* \bar{y}_h \quad (15)$$

Where Ω_h^* are the new calibrated weights obtained by minimizing the chi-square distance type measure $\sum_{h=1}^{L} \frac{(\Omega_h^* - W_h)^2}{Q_h W_h}$, subject to two calibration constraints given as

$$\sum_{h=1}^{L} \Omega_h^* = \sum_{h=1}^{L} \mathrm{W}_h \quad (16)$$

$$\sum_{h=1}^{L} \Omega_h^* \rho_{xh} = \sum_{h=1}^{L} W_h \rho_{Xh} \quad (17)$$

$$\sum_{h=1}^{L} \Omega_h^* c_{xh} = \sum_{h=1}^{L} W_h C_{Xh} \quad (18)$$

Here $c_{xh} = s_{hx}/\bar{x}_h$ and $C_{Xh} = S_{hx}/\bar{X}_h$ are the h^{th} stratum sample and population coefficient of variation of the auxiliary variable, respectively

The Lagrange function is given as

$$L = \sum_{h=1}^{L} \frac{(\Omega_h^* - W_h)^2}{Q_h W_h} - 2\lambda_1(\sum_{h=1}^{L} \Omega_h^* - \sum_{h=1}^{L} \mathrm{W}_h) - 2\lambda_2(\sum_{h=1}^{L} \Omega_h^* \rho_{xh} - \sum_{h=1}^{L} W_h \rho_{Xh}) - 2\lambda_3(\sum_{h=1}^{L} \Omega_h^* c_{xh} - \sum_{h=1}^{L} W_h C_{Xh}) \quad (19)$$

Where λ_1, λ_2 and λ_3 are the Lagrange multipliers to find the optimum value of Ω_h^*.

Now, differentiate the eq. (19) with respect to Ω_h^* and equate to zero.

$$\frac{\partial \Delta}{\partial \Omega_h^*} = \frac{2(\Omega_h^* - W_h)}{Q_h W_h} - 2\lambda_1 - 2\lambda_2 \rho_{xh} - 2\lambda_3 c_{xh}$$

$$\frac{(\Omega_h^*-W_h)}{Q_h W_h}-\lambda_1-\lambda_2\rho_{xh}-\lambda_3 c_{xh}=0$$

$$\Omega_h^*=W_h+Q_h W_h(\lambda_1+\lambda_2\rho_{xh}+\lambda_3 c_{xh}) \tag{20}$$

Substituting the weights Ω_h^* in (16), (17) and (18) we get

$$\lambda_1\sum_{h=1}^{L}W_hQ_h+\lambda_2\sum_{h=1}^{L}W_hQ_h\rho_{xh}+\lambda_3\sum_{h=1}^{L}W_hQ_hc_{xh}=0 \tag{21}$$

$$\lambda_1\sum_{h=1}^{L}W_hQ_h\rho_{xh}+\lambda_2\sum_{h=1}^{L}W_hQ_h\rho_{xh}^2+\lambda_3\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh}=\sum_{h=1}^{L}W_h(\rho_{Xh}-\rho_{xh}) \tag{22}$$

$$\lambda_1\sum_{h=1}^{L}W_hQ_hc_{xh}+\lambda_2\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh}+\lambda_3\sum_{h=1}^{L}W_hQ_hc_{xh}^2=\sum_{h=1}^{L}W_h(C_{Xh}-c_{xh}) \tag{23}$$

Now from equation (21), (22) and (23)

$$\begin{bmatrix}\sum_{h=1}^{L}W_hQ_h & \sum_{h=1}^{L}W_hQ_h\rho_{xh} & \sum_{h=1}^{L}W_hQ_hc_{xh}\\ \sum_{h=1}^{L}W_hQ_h\rho_{xh} & \sum_{h=1}^{L}W_hQ_h\rho_{xh}^2 & \sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh}\\ \sum_{h=1}^{L}W_hQ_h\rho_{xh} & \sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh} & \sum_{h=1}^{L}W_hQ_hc_{xh}^2\end{bmatrix}\begin{bmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{bmatrix}=\begin{bmatrix}0\\ \sum_{h=1}^{L}W_h(\rho_{Xh}-\rho_{xh})\\ \sum_{h=1}^{L}W_h(C_{Xh}-c_{xh})\end{bmatrix} \tag{24}$$

$D=(\sum_{h=1}^{L}W_hQ_h)[(\sum_{h=1}^{L}W_hQ_h\rho_{xh}^2)(\sum_{h=1}^{L}W_hQ_hc_{xh}^2)-(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})^2]-$
$(\sum_{h=1}^{L}W_hQ_h\rho_{xh})[(\sum_{h=1}^{L}W_hQ_h\rho_{xh})(\sum_{h=1}^{L}W_hQ_hc_{xh}^2)-(\sum_{h=1}^{L}W_hQ_h\rho_{xh})(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})]+$
$(\sum_{h=1}^{L}W_hQ_hc_{xh})[(\sum_{h=1}^{L}W_hQ_h\rho_{xh})(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})-(\sum_{h=1}^{L}W_hQ_h\rho_{xh})(\sum_{h=1}^{L}W_hQ_h\rho_{xh}^2)]$

$N_1=\sum_{h=1}^{L}W_h(\rho_{Xh}-\rho_{xh})\,[(\sum_{h=1}^{L}W_hQ_hc_{xh})(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})-(\sum_{h=1}^{L}W_hQ_h\rho_{xh}^2)(\sum_{h=1}^{L}W_hQ_hc_{xh}^2)]-$
$\sum_{h=1}^{L}W_h(C_{Xh}-c_{xh})\,[(\sum_{h=1}^{L}W_hQ_h\rho_{xh}^2)(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})+(\sum_{h=1}^{L}W_hQ_hc_{xh})(\sum_{h=1}^{L}W_hQ_h\rho_{xh}^2)]$

$N_2=\sum_{h=1}^{L}W_h(\rho_{Xh}-\rho_{xh})\,[(\sum_{h=1}^{L}W_hQ_h)(\sum_{h=1}^{L}W_hQ_hc_{xh}^2)-(\sum_{h=1}^{L}W_hQ_hc_{xh})^2]+\sum_{h=1}^{L}W_h(C_{Xh}-$
$c_{xh})\,[(\sum_{h=1}^{L}W_hQ_hc_{xh})(\sum_{h=1}^{L}W_hQ_h\rho_{xh})-(\sum_{h=1}^{L}W_hQ_h)(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})]$

$N_3=\sum_{h=1}^{L}W_h(\rho_{Xh}-\rho_{xh})\,[(\sum_{h=1}^{L}W_hQ_h\rho_{xh})(\sum_{h=1}^{L}W_hQ_hc_{xh})-(\sum_{h=1}^{L}W_hQ_h)(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})]+$
$\sum_{h=1}^{L}W_h(C_{Xh}-c_{xh})\,[(\sum_{h=1}^{L}W_hQ_h)(\sum_{h=1}^{L}W_hQ_h\rho_{xh}^2)-(\sum_{h=1}^{L}W_hQ_h\rho_{xh})^2]$

Then $\lambda_1=\frac{N_1}{D}$, $\lambda_2=\frac{N_2}{D}$, $\lambda_3=\frac{N_3}{D}$

Substituting the value of $\lambda_1,\lambda_2\ and\ \lambda_3$ in equation (20)

$$\Omega_h^*=W_h+Q_hW_h\left(\frac{N_1}{D}+\frac{N_2}{D}\rho_{xh}+\frac{N_3}{D}c_{xh}\right)$$

Now the proposed calibrated estimator is written as

$$\bar{y}_{st}(N)=\bar{y}_{st}+\hat{\beta}_{1(N)}[\sum_{h=1}^{L}W_h(\rho_{Xh}-\rho_{xh})]+\hat{\beta}_{2(N)}[\sum_{h=1}^{L}W_h(C_{Xh}-c_{xh})]$$

Where $\hat{\beta}_{1(N)}=\frac{N_4}{D}\quad and\quad \hat{\beta}_{2(N)}=\frac{N_5}{D}$

$N_4=\sum_{h=1}^{L}W_hQ_h\,\bar{y}_h[(\sum_{h=1}^{L}W_hQ_hc_{xh})(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})-(\sum_{h=1}^{L}W_hQ_h\rho_{xh})(\sum_{h=1}^{L}W_hQ_hc_{xh}^2)]+$
$\sum_{h=1}^{L}W_hQ_h\,\bar{y}_h\rho_{xh}[(\sum_{h=1}^{L}W_hQ_h)(\sum_{h=1}^{L}W_hQ_hc_{xh}^2)-(\sum_{h=1}^{L}W_hQ_hc_{xh})^2]+$
$\sum_{h=1}^{L}W_hQ_h\,\bar{y}_hc_{xh}[(\sum_{h=1}^{L}W_hQ_h\rho_{xh})(\sum_{h=1}^{L}W_hQ_hc_{xh})-(\sum_{h=1}^{L}W_hQ_h)(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})]$

$N_5=\sum_{h=1}^{L}W_hQ_h\,\bar{y}_h[(\sum_{h=1}^{L}W_hQ_h\rho_{xh})(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})+(\sum_{h=1}^{L}W_hQ_hc_{xh})(\sum_{h=1}^{L}W_hQ_h\rho_{xh}^2)]+$
$\sum_{h=1}^{L}W_hQ_h\,\bar{y}_h\rho_{xh}[(\sum_{h=1}^{L}W_hQ_hc_{xh})(\sum_{h=1}^{L}W_hQ_h\rho_{xh})-(\sum_{h=1}^{L}W_hQ_h)(\sum_{h=1}^{L}W_hQ_hc_{xh}\rho_{xh})]+$
$\sum_{h=1}^{L}W_hQ_h\,\bar{y}_hc_{xh}[(\sum_{h=1}^{L}W_hQ_h)(\sum_{h=1}^{L}W_hQ_h\rho_{xh}^2)-(\sum_{h=1}^{L}W_hQ_h\rho_{xh})^2]$

Here we describe the different forms of the first proposed estimator for varying values of Q_h as follows:

Case 1: when $\boldsymbol{Q_h=1}$

The proposed estimator is written as

$$\bar{y}_{st}(N)=\bar{y}_{st}+\hat{\beta}_{1(N)}\left[\sum_{h=1}^{L}W_h(\rho_{Xh}-\rho_{xh})\right]+\hat{\beta}_{2(N)}\left[\sum_{h=1}^{L}W_h(C_{Xh}-c_{xh})\right]$$

Where $\hat{\beta}_{1(N)} = \frac{N_4}{D}$ $\quad and \quad$ $\hat{\beta}_{2(N)} = \frac{N_5}{D}$

$N_4 = \sum_{h=1}^{L} W_h \bar{y}_h [(\sum_{h=1}^{L} W_h c_{xh})(\sum_{h=1}^{L} W_h c_{xh}\rho_{xh}) - (\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}^2)] + \sum_{h=1}^{L} W_h \bar{y}_h \rho_{xh} [(\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h c_{xh}^2) - (\sum_{h=1}^{L} W_h c_{xh})^2] + \sum_{h=1}^{L} W_h \bar{y}_h c_{xh} [(\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}) - (\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h c_{xh}\rho_{xh})]$

$N_5 = \sum_{h=1}^{L} W_h \bar{y}_h [(\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}\rho_{xh}) + (\sum_{h=1}^{L} W_h c_{xh})(\sum_{h=1}^{L} W_h \rho_{xh}^2)] + \sum_{h=1}^{L} W_h \bar{y}_h \rho_{xh} [(\sum_{h=1}^{L} W_h c_{xh})(\sum_{h=1}^{L} W_h \rho_{xh}) - (\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h c_{xh}\rho_{xh})] + \sum_{h=1}^{L} W_h \bar{y}_h c_{xh} [(\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h \rho_{xh}^2) - (\sum_{h=1}^{L} W_h \rho_{xh})^2]$

$D = (\sum_{h=1}^{L} W_h)[(\sum_{h=1}^{L} W_h \rho_{xh}^2)(\sum_{h=1}^{L} W_h c_{xh}^2) - (\sum_{h=1}^{L} W_h c_{xh}\rho_{xh})^2] - (\sum_{h=1}^{L} W_h \rho_{xh})[(\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}^2) - (\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}\rho_{xh})] + (\sum_{h=1}^{L} W_h c_{xh})[(\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}\rho_{xh}) - (\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h \rho_{xh}^2)]$

Case 2: when $\boldsymbol{Q_h = 1/c_{xh}}$

The proposed estimator is written as

$$\bar{y}_{st}(N) = \bar{y}_{st} + \hat{\beta}_{1(N)} \left[\sum_{h=1}^{L} W_h(\rho_{Xh} - \rho_{xh})\right] + \hat{\beta}_{2(N)} \left[\sum_{h=1}^{L} W_h(C_{Xh} - c_{xh})\right]$$

Where $\hat{\beta}_{1(N)} = \frac{N_4}{D}$ $\quad and \quad$ $\hat{\beta}_{2(N)} = \frac{N_5}{D}$

$N_4 = \sum_{h=1}^{L} W_h \bar{y}_h / c_{xh} \left[(\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h \rho_{xh}) - \left(\sum_{h=1}^{L} \frac{W_h \rho_{xh}}{c_{xh}}\right)(\sum_{h=1}^{L} W_h c_{xh})\right] + \sum_{h=1}^{L} W_h \bar{y}_h \rho_{xh} / c_{xh} [(\sum_{h=1}^{L} W_h / c_{xh})(\sum_{h=1}^{L} W_h c_{xh}) - (\sum_{h=1}^{L} W_h)^2] + \sum_{h=1}^{L} W_h \bar{y}_h [(\sum_{h=1}^{L} W_h \rho_{xh}/c_{xh})(\sum_{h=1}^{L} W_h) - (\sum_{h=1}^{L} W_h / c_{xh})(\sum_{h=1}^{L} W_h \rho_{xh})]$

$N_5 = \sum_{h=1}^{L} W_h \bar{y}_h / c_{xh} \left[\left(\sum_{h=1}^{L} \frac{W_h \rho_{xh}}{c_{xh}}\right)(\sum_{h=1}^{L} W_h \rho_{xh}) + (\sum_{h=1}^{L} W_h)\left(\frac{\sum_{h=1}^{L} W_h \rho_{xh}^2}{c_{xh}}\right)\right] + \sum_{h=1}^{L} W_h \bar{y}_h \rho_{xh} / c_{xh} [(\sum_{h=1}^{L} W_h Q_h c_{xh})(\sum_{h=1}^{L} W_h Q_h \rho_{xh}) - (\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h c_{xh}\rho_{xh})] + \sum_{h=1}^{L} W_h Q_h \bar{y}_h c_{xh} [(\sum_{h=1}^{L} W_h Q_h)(\sum_{h=1}^{L} W_h Q_h \rho_{xh}^2) - (\sum_{h=1}^{L} W_h Q_h \rho_{xh})^2]$

$D = (\sum_{h=1}^{L} W_h)[(\sum_{h=1}^{L} W_h \rho_{xh}^2)(\sum_{h=1}^{L} W_h c_{xh}^2) - (\sum_{h=1}^{L} W_h c_{xh}\rho_{xh})^2] - (\sum_{h=1}^{L} W_h \rho_{xh})[(\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}^2) - (\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}\rho_{xh})] + (\sum_{h=1}^{L} W_h c_{xh})[(\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}\rho_{xh}) - (\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h \rho_{xh}^2)]$

Case 3: when $\boldsymbol{Q_h = 1/\rho_{xh}}$

$$\bar{y}_{st}(N) = \bar{y}_{st} + \hat{\beta}_{1(N)} \left[\sum_{h=1}^{L} W_h(\rho_{Xh} - \rho_{xh})\right] + \hat{\beta}_{2(N)} \left[\sum_{h=1}^{L} W_h(C_{Xh} - c_{xh})\right]$$

Where $\hat{\beta}_{1(N)} = \frac{N_4}{D}$ $\quad and \quad$ $\hat{\beta}_{2(N)} = \frac{N_5}{D}$

$N_4 = \sum_{h=1}^{L} W_h \bar{y}_h / \rho_{xh} [(\sum_{h=1}^{L} W_h c_{xh}/\rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}) - (\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h c_{xh}^2 / \rho_{xh})] + \sum_{h=1}^{L} W_h \bar{y}_h [(\sum_{h=1}^{L} W_h / \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}^2 / \rho_{xh}) - (\sum_{h=1}^{L} W_h c_{xh}/\rho_{xh})^2] + \sum_{h=1}^{L} W_h \bar{y}_h c_{xh} / \rho_{xh} [(\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h c_{xh}/\rho_{xh}) - (\sum_{h=1}^{L} W_h / \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh})]$

$N_5 = \sum_{h=1}^{L} W_h \bar{y}_h / \rho_{xh} [(\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h c_{xh}) + (\sum_{h=1}^{L} W_h c_{xh})(\sum_{h=1}^{L} W_h)] + \sum_{h=1}^{L} W_h \bar{y}_h [(\sum_{h=1}^{L} W_h c_{xh} / \rho_{xh})(\sum_{h=1}^{L} W_h) - (\sum_{h=1}^{L} W_h / \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh})] + \sum_{h=1}^{L} W_h \bar{y}_h c_{xh} / \rho_{xh} [(\sum_{h=1}^{L} W_h / \rho_{xh})(\sum_{h=1}^{L} W_h \rho_{xh}) - (\sum_{h=1}^{L} W_h)^2]$

$D = (\sum_{h=1}^{L} W_h / \rho_{xh})[(\sum_{h=1}^{L} W_h \rho_{xh})(\sum_{h=1}^{L} W_h c_{xh}^2 / \rho_{xh}) - (\sum_{h=1}^{L} W_h c_{xh})^2] - (\sum_{h=1}^{L} W_h)[(\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h c_{xh}^2 / \rho_{xh}) - (\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h c_{xh})] + (\sum_{h=1}^{L} W_h c_{xh} / \rho_{xh})[(\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h c_{xh}) - (\sum_{h=1}^{L} W_h)(\sum_{h=1}^{L} W_h \rho_{xh})]$

4. Simulation Study

To study the performance of proposed calibration estimators in a stratified random sampling design with respect to the other estimators, we have conducted a simulation study on two data sets one is an artificial dataset generated using R-software and other is a real data taken from Sarndal et al. (2003).

4.1 Simulation study with artificial data set

To study the performance of the proposed estimator a limited simulation study has been carried out for with three different artificially generated populations. The population given in Tracy et.al. (2003) and Nursel Koyuncu (2015). Where X_{hi} and y_{hi} values are form different distributions. The populations given in Table. 1. The correlation coefficients between study and auxiliary variables for each stratum taken as $\rho_{xy1} = 0.5$, $\rho_{xy2} = 0.7, \rho_{xy3} = 0.9$ respectively. The quantities $S_{1x} = 4.5$,$S_{2x} = 6.2$,$S_{3x} = 8.4$ and S_{1y} ,S_{2y} ,$S_{3y} = 4.8$ were fixed in each stratum. All four population consist three strata having 1500 units. This process have been repeated 50000 time independently that is 500 samples of size 200, 300, 400 units have been drawn from each stratum from given population. The formula gives the MSE of Tracy's estimator and the proposed estimator

$$MSE\big(\bar{y}_{st}(\alpha)\big) = \frac{\sum_{k=1}^{(50000)}[\bar{y}_{st}(\alpha) - \bar{Y}]}{(50000)}, \alpha = \bar{y}_{st(Nk)}, \bar{y}_{st(O)}, \bar{y}_{st(N)}$$

Table 1: Parameters and distributions of study and auxiliary variables

Populations	Parameters and distribution of the study Variable	Parameters and distributions of the auxiliary variable
Population $h = 1,2,3$	$f(y_{hi}^*) = \frac{1}{\Gamma(1.5)} y_{hi}^{*1,5-1} e^{-y_{hi}^*}$	$f(x_{hi}^*) = \frac{1}{\Gamma(0.3)} x_{hi}^{*0,3-1} e^{-x_{hi}^*}$
2. Population $h = 1,2,3$	$f(y_{hi}^*) = \frac{1}{\Gamma(0.3)} y_{hi}^{*0.3-1} e^{-y_{hi}^*}$	$f(x_{hi}^*) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_{hi}^{*2}}{2}}$
4. Population $h = 1,2,3$	$f(y_{hi}^*) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y_{hi}^{*2}}{2}}$	$f(x_{hi}^*) = \frac{1}{\Gamma(0.3)} x_{hi}^{*0,3-1} e^{-x_{hi}^*}$

Table 2: Properties of strata

Strata	Study variable	Auxiliary variable
Stratum	$y_{1i} = 50 + y_{1i}^*$	$x_{1i} = 15 + \sqrt{(1-\rho_{xy1}^2)x_{1i}^*} + \rho_{xy1}\left(\frac{S_{1x}}{S_{1y}}\right) y_{1i}^*$
Stratum	$y_{2i} = 150 + y_{2i}^*$	$x_{2i} = 100 + \sqrt{(1-\rho_{xy2}^2)x_{2i}^*} + \rho_{xy2}\left(\frac{S_{2x}}{S_{2y}}\right) y_{2i}^*$
Stratum	$y_{3i} = 100 + y_{3i}^*$	$x_{3i} = 200 + \sqrt{(1-\rho_{xy3}^2)x_{3i}^*} + \rho_{xy3}\left(\frac{S_{3x}}{S_{3y}}\right) y_{3i}^*$

Table 3: The Mean square error (MSE) of different calibration estimators in Stratified Sampling design for artificial data.

Populations	**Estimators**	$\boldsymbol{Q=1}$	$\boldsymbol{Q=1/c_{xh}}$	$\boldsymbol{Q=1/\rho_{xh}}$
I	$\bar{y}_{st(Nk)}$	410262.8	702012.8	303107.5
	$\bar{y}_{st(O)}$	7379.371	47134.05	4739.451
	$\bar{y}_{st(N)}$	**2772.786**	**70.63716**	**292.2711**
II	$\bar{y}_{st(Nk)}$	372400.1	602499	305422.8
	$\bar{y}_{st(O)}$	1513.792	27242.59	1167.385
	$\bar{y}_{st(N)}$	**30.45064**	**64.27899**	**258.4545**
III	$\bar{y}_{st(Nk)}$	308027.1	510899.9	229310.8
	$\bar{y}_{st(O)}$	8388.492	49980.68	5282.875
	$\bar{y}_{st(N)}$	**1725.127**	**62.34677**	**304.0844**

4.2 Simulation study with real data set

The simulation study with real data was carried out to study the performance of the proposed calibration estimators. The population MU284 is taken from Appendix B of Sarndal et al. (2003, pp). We consider 246 units from that population after deleting extreme units. Stratification is done according to the REG variable and hence we have eight strata. The study variable y was the population of the year 1985 (P85) and the auxiliary variable x was the population of the year 1982 (CS82). A simulated study generating 50,000 samples is done using R-software. A simulated study generating 50,000 samples is done using R-software is worked out for stratified sampling.

$$\boldsymbol{MSE\left(\bar{y}_{st}(\alpha)\right) = \frac{\sum_{k=1}^{(50000)}\left(\bar{y}_{st}(\alpha) - \bar{Y}\right)^2}{50000}}, \quad \boldsymbol{\alpha = \bar{y}_{st(Nk)}, \bar{y}_{st(O)}, \bar{y}_{st(N)}}$$

Table 4: The Mean square error (MSE) of different calibration estimators in Stratified Sampling design for real data.

Estimators	$Q=1$	$Q=1/c_{xh}$	$Q=1/\rho_{xh}$
$\bar{y}_{st(Nk)}$	8479640696	1671466679	3101296249
$\bar{y}_{st(O)}$	20653915317	9803309512	518125763579
$\bar{y}_{st(N)}$	**190.488**	**13389.2**	**420720917**

5. Result

The mean square error of the proposed calibration estimators, which uses a new set of three constraints based on coefficient of variation and correlation of the auxiliary variable, is compared with other existing calibration estimators in stratified random sampling. The results are presented in Table 3 and Table 4 for both an artificially generated data set and a real data set.

From Table 3 and Table 4 it is evident that the proposed calibration estimator $\bar{y}_{st(N)}$ is better perform for the different situations of $Q_i's$. This result is consistent for both populations.

6. Conclusion

In this chapter we have considered the calibration estimation in stratified sampling. A new calibration estimators are proposed under stratified random sampling incorporating new set of three constraints.

To evaluate the performance of proposed estimators in comparison with other calibration estimators, a simulation study was conducted on two datasets. From this simulation study, we conclude that the proposed estimator is better perform than the other existing estimators for all conditions.

Moreover, the empirical studies highlighted the efficiency of proposed approach, showing significant improvements over the methods developed by Koyuncu et al. (2015) and Ozqul (2018).

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
ORCID
Namita Srivastava, ORCID ID: 0000-0001-8695-9148
Ruchi Shakya, ORCID ID: 0009-0008-1500-1045
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers
Artificial Intelligence (AI) Use Disclosure: The authors declare that no artificial intelligence tools were used in the preparation of this manuscript.

References

[1] Alam, S. and Shabbir J. (2020). Calibration estimation of mean by using double use of auxiliary information, Communications in Statistics- Simulations and Computation DOI: 10.1080/0361091802020.1749660.
[2] Babatunde, T. et al. (2022). Robust Calibration Estimation of population mean in stratified random sampling in presence of outliers, Research article in Research Square. DOI: 10.21203/rs.3,rs-2037825/v1.
[3] Clement EP. (2015). Calibration approach separate ratio estimator for population mean in stratified sampling. Int J Mod Math Sci., 377-384.
[4] Deville, J.C., and Sarndal, C.E. (1992). Calibration estimators in survey sampling, Journal of American Statistical Association, Vol.87, No. 418.
[5] Estevao, V.M., Sarndal, C.E., (2000). Afunctional form approach to calibration, Journal of Official Statistics, 16, 379-399.
[6] Farrell, P.J., Singh, S., (2005). Model-assisted higher order calibration of estimators of variance, Australian and New Zeland J. Statist., 47, 3,375-383.
[7] Kim, J.M., Sungur, E. A. Heo T.Y., (2007). Calibration Approach estimator in stratified sampling, Statistics and Probability Letters, 77, 1, 99-103.
[8] Kim, J.M., Park, M., (2010). Calibration estimation in Survey Sampling, International Statistical Review, 78, 1, 21-29.
[9] Garg, N. and Pachori, M. (2019). Use of coefficient of variation in calibration estimation of population mean in stratified sampling, Communication in Statistics Theory and Methods, Vol.49, No23 5842-5852.
[10] Koyuncu, N., and Kadilar, C. (2014). A new calibration estimator in stratified double sampling. Hacettepe Journal of Mathematics and Statistics, 43(2), 337-346.
[11] Koyuncu, N., and Kadilar, C. (2016). Calibration weighting in stratified random sampling. Communications in Statistics, Simulation and Computation 45(7), 2267-2275 doi:10.1080/03610918.2014.901354.
[12] Koyuncu, N. (2018). Calibration estimator of population mean under stratified ranked set sampling design. Communications in Statistics Theory and Methods, 47(23), 5845-5853.
[13] Nidhi, Sisodia B. V. S., Singh S. and Singh, S. K. (2017). Calibration approach estimation of the mean in stratified sampling and stratified double sampling, Communications in Statistics - Theory and Methods, 46:10, 4932-4942, DOI: 10.1080/03610926.2015.1091083.
[14] Ozgul, N. (2018). New calibration estimator in stratified sampling, Article in Journal of Statistical Computation (Hacttepe University) vol. 88(13).
[15] Ozgul, N. (2019). New Calibration Estimator Based on Tow Auxiliary Variables in stratified sampling, journal of communications in Statistics: Theory and methods.
[16] Singh, G., Rao, D., Khan, M. G. M. (2014). Calibration estimator of population mean in stratified random sampling, Asia-Pacific World Congress on Computer Science and Engineering. DOI:10.1109/APWCCSE.2014.7053875.
[17] Tracy, D.S., Singh, S. and Raghunath, A. (2003). Note on calibration in stratified and double sampling no. 12-001 Survey Methodology, vol 29 no.1 pp.99-104 .

[18] T. Babatunde et al. (2022). Robust Calibration Estimation of population mean in stratified random sampling in presence of outlier, research article in Research square.
[19] Wang, et al. (2016). Using ranked set sampling with cluster randomized designs for improved inference on treatment effects, J. Am. Stat. Assoc. 516, 1576-1590
[20] Shahzad, U (2023). Calibration Based Mean estimator under stratified Median Ranked set sampling, ∑ Mathematics (MDPI), 11, 1825.