JIRSS (2024)
Vol. 23, No. 01, pp 117-130
DOI: 10.22034/jirss.2024.1999153.1017

A Dual Problem of Calibration Ratio-Type Estimator under Stratified Systematic Sampling Scheme

Alka[1], **Piyush Kant Rai**[2], **Muhammad Qasim**[3]

[1] Department of Mathematics and Statistics, Banasthali University, Rajasthan, India.
[2] Department of Statistics, Banaras Hindu University, Varanasi, India.
[3] Department of Economics, Finance & Statistics, Jonkoping University, Sweden.

Received: 28/03/2023, Accepted: 02/08/2024, Published online: 20/11/2024

Abstract. This article introduces a dual problem of widely used calibration ratio-type estimators for estimating population mean of the study variable considering auxiliary information under dual constraints using stratified systematic sampling design. Under large sample approximations, the expression for bias and variance of the proposed estimator are derived. In addition, the optimality condition for the proposed estimator and hence optimum variance expression is also obtained for the same. Moreover, a study based on real-life data is carried out to judge the performance of the proposed calibration estimator in terms of minimum relative bias and relative root mean squared error criterion. The study reveals that the calibration ratio-type estimator under dual constraints may be preferred in practice as it provides consistent and more precise parameter estimates.

Keywords. Auxiliary Information, Calibration Estimation, Ratio-type Estimator, Systematic Sampling.

MSC: 62D05.

1 Introduction

Deville and Särndal (1992) first studied calibration estimators in survey sampling and is widely used by many survey statisticians as Estevao and Särndal (2006); Särndal (2007);

Corresponding Author: Alka (singhalka2889@gmail.com)
Piyush Kant Rai (raipiyush5@gmail.com)
Muhammad Qasim (muhammad.qasim@ju.se).

Kim and Park (2010); Clement and Enang (2017); Ozgul (2019); Alka (2019, 2021); Zaman and Bulut (2023); Rai et al. (2024) among many others. The basic purpose is to develop calibration estimator parallel to Horvitz Thompson estimator by considering auxiliary information to get calibrated weights close to design weights. Koyuncu and Kadilar (2013, 2016) defined some calibration estimators in stratified random sampling for population characteristics. Clement et al. (2014) defined calibration estimators for domain totals in stratified random sampling. Singh et al. (2009, 2011, 2012), Choudhury and Singh (2012a, b) and Onyeka (2012, 2013) proposed various improved classes of ratio estimators for estimating population mean under simple random sampling without replacement (SRSWOR) and stratified random sampling schemes. Some extensions to two-phase and post-stratification sampling are also considered and analysed by these authors.

In the recent decade, ample work has been done to develop such estimators in practice utilizing several aspects such as suitability of distance function, different calibration constraints, multi-auxiliary characteristics, and concept of multi-frame, along with varying sampling schemes. Singh (2013) discussed the dual problem of calibration of design weights. In connection with the above literature, an alternative estimator for estimating the population mean utilizing auxiliary information in the best possible way calibration based approach is welcomed which provides relatively precise estimation with a smaller mean squared error.

Therefore, in the present article an effort is made to get aligned ratio estimator under well suitable and demanded stratified systematic sampling design with better precision as compared to single constraint calibrated estimator. Here, we introduce the theory of calibration estimator to ratio estimation under stratified systematic sampling scheme and following Solanki et al. (2012), we propose a class of ratio estimators for population mean $\bar{Y}$ of the study variable (y) using auxiliary variable (x) under dual problem of calibration-based approach. We also study the properties for e.g. bias and variance of proposed class of estimators and the conditions under which the estimator will provide an optimal performance. Relative bias and root mean squared error of proposed estimator are computed and compared under single and dual constraints.

2 Proposed Calibration Estimator

Consider a finite population $\Omega = \{1, 2, ..., N\}$ which is divided into H strata with the h^{th} stratum containing N_h ($h = 1, 2, ..., H$) units, such that $\sum_{h=1}^{H} N_h = N$. A simple random sample of size n_h is drawn using without replacement scheme from the h^{th} stratum such that $\sum_{h=1}^{H} n_h = n$. The study and auxiliary variables are denoted as y and x, respectively. Furthermore, y_{h_i} and x_{h_i} denote the values associated with the i^{th} ($i = 1, 2, ..., N_h$) unit from h^{th} ($h = 1, 2, ..., H$) stratum in the population by y and x. Let the population mean of the auxiliary variable $\bar{X} = \dfrac{1}{N_h} \sum_{i=1}^{N_h} x_{hi}$ be known. The primary aim is to estimate the population mean $\bar{Y} = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{hi}$.

Using stratified random sampling scheme, the classical unbiased estimator of population mean is given by

$$\bar{Y}_{st} = \sum_{h=1}^{H} w_h \bar{y}_h, \tag{2.1}$$

where $w_h = N_h/N$ is the stratum weight and $f_h = n_h/N_h$ be the sample fraction for h^{th} stratum in the population.

Based on n_h observations, let the h^{th} stratum means of the study variable $\bar{y}_{hi} = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$ and auxiliary variable $\bar{x}_{hi} = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{hi}$ be the unbiased estimator of the population mean $\bar{Y}_{hi} = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{hi}$ and $\bar{X}_{hi} = \frac{1}{N_h} \sum_{i=1}^{N_h} x_{hi}$ respectively. Cochran (1977) proposed the estimate of the population mean $\bar{Y}$ in stratified systematic sampling scheme:

$$\bar{y}_{stsy} = \sum_{h=1}^{H} w_h \bar{y}_{syh}, \tag{2.2}$$

where $\bar{y}_{syh}$ is the mean of systematic sample in the h^{th} stratum.

A ratio-type estimator under Simple Random Sampling Without Replacement (SRSWOR) proposed by Solanki et al. (2012) is as follows:

$$t(\alpha, \delta) = \bar{y}\left[2 - \left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha} exp\left(\frac{\delta(\bar{x} - \bar{X})}{(\bar{x} + \bar{X})}\right)\right]. \tag{2.3}$$

where α and δ are suitably chosen scalars.

Etuk et al. (2016) proposed an alternative ratio estimator of population mean under systematic sampling following Solanki et al. (2012) estimator which is given as

$$\bar{y}_{R,stsy} = \sum_{h=1}^{H} w_h \bar{y}_{syh}\left[\lambda_h - r_h \left(\frac{\bar{x}_{syh}}{\bar{X}_h}\right)^{\alpha_h} exp\left(\frac{\delta_h(\bar{x}_{syh} - \bar{X}_h)}{(\bar{x}_{syh} + \bar{X}_h)}\right)\right], \tag{2.4}$$

where λ_h, r_h, α_h and δ_h are suitably chosen scalars, under the condition $\lambda_h = 1 + r_h$ is satisfied by λ_h and r_h.

Here, we propose the family of estimators under calibration approach-based method of estimation as below

$$\bar{y}^{c}_{R,stsy} = \sum_{h=1}^{H} \Omega_h \bar{y}_{syh}\left[\lambda_h - r_h \left(\frac{\bar{x}_{syh}}{\bar{X}_h}\right)^{\alpha_h} exp\left(\frac{\delta_h(\bar{x}_{syh} - \bar{X}_h)}{(\bar{x}_{syh} + \bar{X}_h)}\right)\right], \tag{2.5}$$

where the calibrated weights are denoted by Ω_h. These weights can be obtained by optimizing the estimator of popuation mean of the auxiliary variable

$$\sum_{h=1}^{H} \Omega_h \bar{x}_{syh} = \hat{\hat{X}}, \tag{2.6}$$

under the constraint

$$\sum_{h=1}^{H} w_h = \sum_{h=1}^{H} \Omega_h, \tag{2.7}$$

and a new dual constraint γ defined as

$$\gamma = \frac{1}{2} \sum_{h=1}^{H} \frac{(\Omega_h - w_h)^2}{w_h q_h}, \tag{2.8}$$

where q_h are suitably chosen weights called tuning weights. This constraint is in the form of chi-square distance function. The form of the estimator in Eq.(2.5) depends upon the choice of q_h. The lagrangian function will be given as

$$L = \sum_{h=1}^{H} \Omega_h \bar{x}_{syh} - \lambda_1 \left(\sum_{h=1}^{H} \Omega_h - \sum_{h=1}^{H} w_h \right) - \lambda_2 \left(\frac{1}{2} \sum_{h=1}^{H} \frac{(\Omega_h - w_h)^2}{w_h q_h} - \gamma \right), \tag{2.9}$$

where λ_1 and λ_2 are lagrange's multiplier. On differentiating the lagrangian function with respect to Ω_h and equating to 0, we get

$$w_h q_h \bar{x}_{syh} - \lambda_1 w_h q_h - \lambda_2 (\Omega_h - w_h) = 0. \tag{2.10}$$

Now, on taking summation over all the possible sampled values, we have

$$\sum_{h=1}^{H} w_h q_h \bar{x}_{syh} - \lambda_1 \sum_{h=1}^{H} w_h q_h - \lambda_2 \left(\sum_{h=1}^{H} \Omega_h - \sum_{h=1}^{H} w_h \right) = 0.$$

Solving for λ_1 by using Eq.(2.7), we have

$$\lambda_1 = \frac{\sum_{h=1}^{H} w_h q_h \bar{x}_{syh}}{\sum_{h=1}^{H} w_h q_h}. \tag{2.11}$$

On putting the value of λ_1 from Eq.(2.11) to Eq.(2.10)

$$(\Omega_h - w_h) = \frac{w_h q_h}{\lambda_2} \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h q_h \bar{x}_{syh}}{\sum_{h=1}^{H} w_h q_h} \right). \tag{2.12}$$

Substituting the value of $(\Omega_h - w_h)$ in Eq.(2.8) and solving

$$\lambda_2 = \pm \frac{\sqrt{\sum_{h=1}^{H} w_h q_h \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h q_h \bar{x}_{syh}}{\sum_{h=1}^{H} w_h q_h} \right)^2}}{\sqrt{\gamma}}, \tag{2.13}$$

where the sign of λ_2 is to be determined by the choice of the sign for γ. From Eqs.(2.12) and (2.13), the new calibrated weights are given by

$$\Omega_h = w_h + \frac{w_h q_h \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h q_h \bar{x}_{syh}}{\sum_{h=1}^{H} w_h q_h} \right)}{\sqrt{\sum_{h=1}^{H} w_h q_h \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h q_h \bar{x}_{syh}}{\sum_{h=1}^{H} w_h q_h} \right)^2}} \sqrt{\gamma}. \tag{2.14}$$

The optimal calibrated weights so obtained in Eq.(2.14) can be used for estimating the population mean of the study variable by means of the estimator defined in Eq.(2.5) as

$$\bar{y}^{c}_{R,stsy} = \sum_{h=1}^{H} w_h \bar{y}_{syh} \left[\lambda_h - r_h \left(\frac{\bar{x}_{syh}}{\bar{X}_h} \right)^{\alpha_h} exp\left(\frac{\delta_h (\bar{x}_{syh} - \bar{X}_h)}{(\bar{x}_{syh} + \bar{X}_h)} \right) \right] + \hat{\delta}\sqrt{\gamma}, \tag{2.15}$$

where,

$$\hat{\delta} = \frac{\sum_{h=1}^{H} w_h q_h \bar{y}_{syh} \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h q_h \bar{x}_{syh}}{\sum_{h=1}^{H} w_h q_h} \right)}{\sqrt{\sum_{h=1}^{H} w_h q_h \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h q_h \bar{x}_{syh}}{\sum_{h=1}^{H} w_h q_h} \right)^2}} \left[\lambda_h - r_h \left(\frac{\bar{x}_{syh}}{\bar{X}_h} \right)^{\alpha_h} exp\left(\frac{\delta_h (\bar{x}_{syh} - \bar{X}_h)}{(\bar{x}_{syh} + \bar{X}_h)} \right) \right].$$

Writing Eq.(2.14) under $q_h = \bar{x}^{-1}_{syh}$, we have

$$\Omega_h = w_h + \frac{w_h \bar{x}^{-1}_{syh} \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h \bar{x}^{-1}_{syh}} \right)}{\sqrt{\sum_{h=1}^{H} w_h \bar{x}^{-1}_{syh} \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h \bar{x}^{-1}_{syh}} \right)^2}} \sqrt{\gamma}. \tag{2.16}$$

We may also develop another set of family of ratio estimators by plugging Ω_h obtained above in Eq.(2.16) into Eq.(2.5). The best choice of $\sqrt{\gamma}$ can be

$$\frac{(\bar{X} - \sum_{h=1}^{H} w_h \bar{x}_{syh})}{\sqrt{\sum_{h=1}^{H} w_h q_h \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h q_h \bar{x}_{syh}}{\sum_{h=1}^{H} w_h q_h} \right)^2}} \sim N(0,1).$$

Any other unit free value of $\sqrt{\gamma}$ can be used like

$$\frac{(S_{hx}-s_{hx})}{\sqrt{\sum_{h=1}^{H} w_h q_h \left(\bar{x}_{syh}-\frac{\sum_{h=1}^{H} w_h q_h \bar{x}_{syh}}{\sum_{h=1}^{H} w_h q_h}\right)^2}}, \tag{2.17}$$

where S_{hx} & s_{hx} are the population and sample standard deviation for h^{th} stratum.

2.1 Bias and Variance of the Proposed Estimator

To obtain the expression of Bias and Variance of the proposed estimator under dual constraint, let us assume that

$$\bar{y}_{syh}=\bar{Y}_h(1+e_0) \qquad \text{and} \qquad \bar{x}_{syh}=\bar{X}_h(1+e_1), \tag{2.18}$$

such that

$$E(e_0)=E(e_1)=0, \tag{2.19}$$

$$\begin{aligned} E(e_1^2) &= \frac{E(\bar{x}_{syh}-\bar{X}_h)^2}{\bar{X}_h^2} = \frac{1}{\bar{X}_h^2}\frac{N_h-1}{N_h}\frac{S_{hx}^2}{n_h}\left[1+(n_h-1)\rho_{hx}\right] \\ &= \left(\frac{N_h-1}{N_h n_h}\right)\frac{S_{hx}^2}{\bar{X}_h^2}\left[1+(n_h-1)\rho_{hx}\right] \\ &= \theta_h\left[1+(n_h-1)\rho_{hx}\right]C_{hx}^2. \end{aligned} \tag{2.20}$$

Similarly

$$E(e_0^2)=\theta_h(1+(n_h-1)\rho_{hy})C_{hy}^2, \tag{2.21}$$

$$E(e_0e_1)=\theta_h(1+(n_h-1)\rho_{hx})^{1/2}(1+(n_h-1)\rho_{hy})^{1/2}\rho_{hxy}C_{hx}C_{hy}, \tag{2.22}$$

where

$$\theta_h=\left(\frac{N_h-1}{N_h n_h}\right); \qquad C_{hx}^2=\frac{S_{hx}^2}{\bar{X}_h^2}; \qquad C_{hy}^2=\frac{S_{hy}^2}{\bar{Y}_h^2}; \qquad \rho_{hxy}=\frac{S_{hxy}}{S_{hx}S_{hy}};$$

$$S_{hx}^2=\frac{1}{N_h-1}\sum_{i=1}^{N_h}(x_{hi}-\bar{X}_h)^2; \qquad S_{hy}^2=\frac{1}{N_h-1}\sum_{i=1}^{N_h}(y_{hi}-\bar{Y}_h)^2;$$

$$S_{hxy}^2=\frac{1}{N_h-1}\sum_{i=1}^{N_h}(x_{hi}-\bar{X}_h)(y_{hi}-\bar{Y}_h).$$

Rewriting Eq.(2.5) using above notations, we get

$$\begin{aligned}\bar{y}^c_{R,stsy} &= \sum_{h=1}^{H} \Omega_h \bar{Y}_h (1+e_0)\left[\lambda_h - r_h \left(\frac{\bar{X}_h(1+e_1)}{\bar{X}_h}\right)^{\alpha_h} exp\left(\frac{\delta_h(\bar{X}_h(1+e_1)-\bar{X}_h)}{(\bar{X}_h(1+e_1)+\bar{X}_h)}\right)\right] \\ &= \sum_{h=1}^{H} \Omega_h \bar{Y}_h \left[\lambda_h - r_h(1+e_1)^{\alpha_h} exp\left(\frac{\delta_h e_1}{2}\left(1+\frac{1}{2}e_1\right)^{-1}\right)\right] \\ &\quad + \sum_{h=1}^{H} \Omega_h \bar{Y}_h e_0 \left[\lambda_h - r_h(1+e_1)^{\alpha_h} exp\left(\frac{\delta_h e_1}{2}\left(1+\frac{1}{2}e_1\right)^{-1}\right)\right].\end{aligned} \tag{2.23}$$

Expanding the terms $(1+e_1)^{\alpha_h}$, $\left(1+\frac{1}{2}e_1\right)^{-1}$ and $exp\left[\frac{\delta_h e_1}{2}\left(1+\frac{1}{2}e_1\right)^{-1}\right]$ in the first order of e_1 and neglecting higher order terms, we obtain

$$\bar{y}^c_{R,stsy} = \sum_{h=1}^{H} \Omega_h \bar{Y}_h + \sum_{h=1}^{H} \Omega_h \bar{Y}_h \left[\left\{\frac{-r_h(2\alpha_h+\delta_h)}{2}\left(e_1 + \frac{2\alpha_h+\delta_h-2}{4}e_1^2 + e_0 e_1\right)\right\} + e_0\right],$$

$$\bar{y}^c_{R,stsy} - \bar{Y} = \sum_{h=1}^{H} \Omega_h \bar{Y}_h \left[\left\{\frac{-r_h(2\alpha_h+\delta_h)}{2}\left(e_1 + \frac{2\alpha_h+\delta_h-2}{4}e_1^2 + e_0 e_1\right)\right\} + e_0\right]. \tag{2.24}$$

Taking expectations on both sides, we have

$$E(\bar{y}^c_{R,stsy} - \bar{Y}) = \sum_{h=1}^{H} \Omega_h \bar{Y}_h \gamma_h \left[\frac{-r_h(2\alpha_h+\delta_h)}{2} C^2_{hx}\left(\frac{2\alpha_h+\delta_h-2}{4} + K_h\right)\right],$$

$$\text{where } \gamma_h = \theta_h(1+(n_h-1)\rho_{hx}),\; K_h = \left(\frac{1+(n_h-1)\rho_{hy}}{1+(n_h-1)\rho_{hx}}\right)^{1/2} \rho_{hxy}\frac{C_{hy}}{C_{hx}}. \tag{2.25}$$

The Bias will be zero if $\alpha_h = -\frac{1}{2}\delta_h$.

Now, to find the variance of the proposed estimator, squaring Eq.(2.24) and neglecting the third and higher order terms of e_0 and e_1.

$$(\bar{y}^c_{R,stsy} - \bar{Y})^2 = \sum_{h=1}^{H} \Omega_h^2 \bar{Y}_h^2 \left(\frac{r_h^2(2\alpha_h+\delta_h)^2 e_1^2}{4} + e_0^2 - 2\frac{r_h(2\alpha_h+\delta_h)}{2}e_0 e_1\right).$$

On taking expectations, we get

$$\begin{aligned}E(\bar{y}^c_{R,stsy} - \bar{Y})^2 &= \sum_{h=1}^{H} \Omega_h^2 \bar{Y}_h^2 \left[\frac{r_h^2(2\alpha_h+\delta_h)^2}{4}\theta_h(1+(n_h-1)\rho_{hx})C^2_{hx} + \theta_h(1+(n_h-1)\rho_{hy})C^2_{hy}\right. \\ &\quad \left. -r_h(2\alpha_+\delta_h)\theta_h(1+(n_h-1)\rho_{hx})^{1/2}(1+(n_h-1)\rho_{hy})^{1/2}\rho_{hxy}C_{hx}C_{hy}\right],\end{aligned}$$

$$V(\bar{y}^c_{R,stsy}) = \sum_{h=1}^{H} \Omega_h^2 \bar{Y}_h^2 \theta_h \psi_h \left[\rho^{*2} C^2_{hy} + \frac{r_h(2\alpha_h+\delta_h)}{4} C^2_{hx}(r_h(2\alpha_h+\delta_h) - 4\rho^* K_h)\right], \tag{2.26}$$

where $\rho^* = \left(\frac{1+(n_h-1)\rho_{hy}}{1+(n_h-1)\rho_{hx}}\right)^{1/2}$, $\psi_h = 1+(n_h-1)\rho_{hx}$. On setting $q_h = \bar{x}_{syh}^{-1}$ and substituting the value of Ω_h in Eq.(2.26), we get

$$V(\bar{y}^c_{R,stsy}) = \sum_{h=1}^{H}\left(w_h + \frac{w_h\bar{x}_{syh}^{-1}\left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h\bar{x}_{syh}^{-1}}\right)}{\sqrt{\sum_{h=1}^{H} w_h\bar{x}_{syh}^{-1}\left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h\bar{x}_{syh}^{-1}}\right)^2}}\sqrt{\gamma}\right)^2 \bar{Y}_h^2\theta_h\psi_h\Big[\rho^{*2}C^2_{hy} + \frac{r_h(2\alpha_h+\delta_h)}{4}C^2_{hx}(r_h(2\alpha_h+\delta_h) - 4\rho^* K_h)\Big]. \tag{2.27}$$

Rewriting above equation by using $\eta = r_h(2\alpha_h+\delta_h)$, we get

$$V(\bar{y}^c_{R,stsy}) = \sum_{h=1}^{H}\left(w_h + \frac{w_h\bar{x}_{syh}^{-1}\left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h\bar{x}_{syh}^{-1}}\right)}{\sqrt{\sum_{h=1}^{H} w_h\bar{x}_{syh}^{-1}\left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h\bar{x}_{syh}^{-1}}\right)^2}}\sqrt{\gamma}\right)^2 \bar{Y}_h^2\theta_h\psi_h\Big[\rho^{*2}C^2_{hy} + \frac{1}{4}C^2_{hx}(\eta^2 - 4\eta\rho^* K_h)\Big]. \tag{2.28}$$

3 Optimal Conditions for the Proposed Calibration Ratio Estimator

The optimal condition for the proposed calibration ratio-type estimator under dual constraint can be given as:

$$\frac{\partial V(\bar{y}^c_{R,stsy})}{\partial \eta} = 0 \text{ which implies that } 2\eta - 4\rho^* K_h = 0,$$

or,

$$\eta = 2\rho^* K_h \text{ that is } r_h(2\alpha_h+\delta_h) = 2\rho^* K_h,$$

Therefore,

$$\alpha_{h,opt} = \frac{2\rho^* K_h - r_h\delta_h}{2r_h}. \tag{3.1}$$

The calibration estimator for population mean in Eq.(2.5) using stratified sampling under dual constraint will be asymptotically optimum on substituting the value of

$\alpha_{h,opt}$ from Eq.(3.1) and is given as

$$\bar{y}^{c}_{R,stsy} = \sum_{h=1}^{H} \Omega_h \bar{y}_{syh} \left(\lambda_h - r_h \left(\frac{\bar{x}_{syh}}{\bar{X}_h} \right)^{\left(\frac{2\rho^* K_h - r_h \delta_h}{2 r_h} \right)} exp \left[\frac{\delta_h (\bar{x}_{syh} - \bar{X}_h)}{(\bar{x}_{syh} + \bar{X}_h)} \right] \right). \tag{3.2}$$

Also, the variance of optimum calibration estimator Eq.(3.2) can be obtained on putting the value of $\alpha_{h,opt}$ in Eq.(2.27) as

$$V(\bar{y}^{c}_{R,stsy}) = \sum_{h=1}^{H} \left(w_h + \frac{w_h \bar{x}^{-1}_{syh} \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h \bar{x}^{-1}_{syh}} \right)}{\sqrt{\sum_{h=1}^{H} w_h \bar{x}^{-1}_{syh} \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h \bar{x}^{-1}_{syh}} \right)^2}} \sqrt{\gamma} \right)^2 \bar{Y}^2_h \theta_h \psi_h \left[\rho^{*2} C^2_{hy} + \frac{2\rho^* K_h}{4} C^2_{hx} (2\rho^* K_h - 4\rho^* K_h) \right]. \tag{3.3}$$

Rewriting the Eq.(3.3) by using Eq.(3.1), we get

$$V(\bar{y}^{c}_{R,stsy}) = \sum_{h=1}^{H} \left(w_h + \frac{w_h \bar{x}^{-1}_{syh} \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h \bar{x}^{-1}_{syh}} \right)}{\sqrt{\sum_{h=1}^{H} w_h \bar{x}^{-1}_{syh} \left(\bar{x}_{syh} - \frac{\sum_{h=1}^{H} w_h}{\sum_{h=1}^{H} w_h \bar{x}^{-1}_{syh}} \right)^2}} \sqrt{\gamma} \right)^2 \bar{Y}^2_h \theta_h \psi_h \rho^{*2} \left[C^2_{hy} - K^2_h C^2_{hx} \right]. \tag{3.4}$$

Further, the performance of the proposed estimator with the help of a real-life dataset is evaluated in the next section.

4 An Application

Consider the Forced Expiratory Volume (FEV) data set which was used by Singh (2013). This real data set of FEV is an index of pulmonary function that measures the volume of air expelled after one second of constant effort and can be downloaded from http://www.amstat.org/publications/jse/datasets/fev.dat.txt. The FEV data set was studied in East Boston, Massachusetts, 1980, on 654 children aged from 3 to 19 years who were seen in the childhood respiratory disease.

The study variable Y is defined as FEV, and the auxiliary variable X as Children age, from 3-19 years of age. For this data set, $\bar{Y} = 2.6367$, $\bar{X} = 9.5841$, and correlation $\rho(X, Y) = 0.75646$. Also, we set $r = -1, \lambda = 0, \alpha = -1, \delta = 3/5, k = 1.70$ as in Etuk et al. (2016) satisfying the condition $\lambda = 1 + r$ and $(2\alpha + \delta) = \frac{2k}{r}$. The author suggested to choose the other set of values for $r, \lambda, \alpha, \delta, k$. Our primary aim is to estimate $\bar{Y} = 2.6367$ (assumed unknown), when $\bar{X} = 9.5841$ is assumed to be known. A pictorial representation of the real life datasets is shown in the Figure 1.

The population ($N = 654$) is divided into three non-overlapping, mutually exclusive strata $h = 1, 2, 3$ containing $N_1 = 128, N_2 = 283, N_3 = 243$ units respectively and within each stratum, samples of proportion 5%, 10%, 15%, 20%, 25%, 30%, 35% and 40% of total population are drawn under systematic sampling design. The values of the estimators are simulated M=5000 using R. Then, the calibration ratio-type estimator for population mean using systematic sampling design under single and dual constraints are analysed in terms of percent Relative Bias (RB%) and percent Relative Root Mean Squared Error (RRMSE%).

We defined RB and RRMSE in percentage as

$$RB(\bar{Y}^c_{R,stsy}) = \frac{\frac{1}{M}\sum_{i=1}^{M} \bar{Y}^c_{R,stsy} - \bar{Y}}{\bar{Y}} \times 100\% \tag{4.1}$$

$$RRMSE(\bar{Y}^c_{R,stsy}) = \sqrt{\frac{1}{M}\sum_{i=1}^{M} (\bar{Y}^c_{R,stsy} - \bar{Y})^2 \times 100\%} \tag{4.2}$$

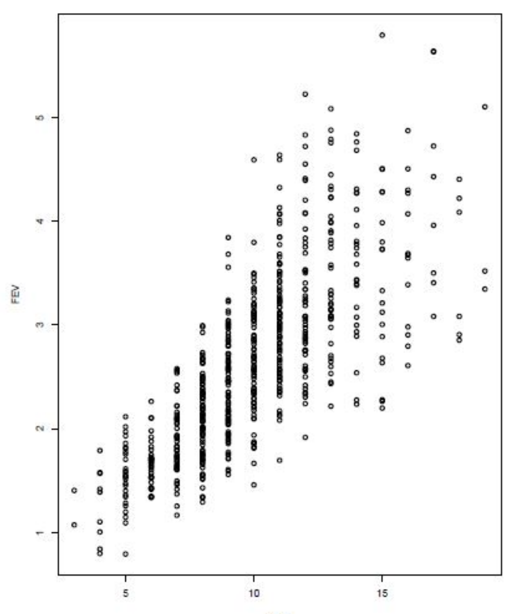

Figure 1: Scatterplot of the two variables (FEV and Age) considered in the study.

Table 1: RB(%) values of the calibration estimator $\bar{Y}^{c}_{R,stsy}$

Sample Proportions	Under Single Constraint	Under Dual Constraints
0.05	4.2041	0.5816
0.10	3.8930	0.2871
0.15	3.7410	0.1197
0.20	3.6736	0.0640
0.25	3.6701	0.0743
0.30	3.6675	0.0695
0.35	3.6697	0.0727
0.40	3.6545	0.0466

Table 2: The RRMSE (%) values of the estimator $\bar{Y}^{c}_{R,stsy}$

Sample Proportions	Under Single Constraint	Under Dual Constraints
0.05	9.5439	7.6607
0.10	7.1389	5.3444
0.15	5.9716	4.1616
0.20	5.3603	3.4828
0.25	4.9426	2.9422
0.30	4.7517	2.6877
0.35	4.5246	2.3749
0.40	4.3846	2.1172

On analyzing the calibration ratio-type estimator under single and dual constraints, the RB(%) values under single constraint varies from 4.2041 to 3.6545 in the samples of proportions 5% to 40% while it varies from 0.5816 to 0.0466 under dual constraints showing significant reduction throughout the samples. From Table 1, we see that RB(%) values under dual constraints are compartively less under single constraint. Furthermore, it can be observed from RRMSE(%) values (from Table 2) which varies from 9.5439 to 4.3846 and 7.6607 to 2.1172 under dual constraints. Also, RRMSE (%) values for dual constraint based calibration estimator are less than that for single constraint based estimator. This supports the fact that, the use of supplementary information in survey sampling increases the precision of the estimates of the population parameter(s) of interest (Cochran, 1977). So, the calibration ratio-type estimator under dual constraints outperforms that of the estimator under single constraint.

5 Conclusion

This article introduces the theory of calibration ratio-type estimator for population mean under dual constraints using stratified systematic sampling design. Expression for bias and variance have been derived upto second order of approximations. Furthermore, the optimal conditions for the proposed estimator and variance expression is given. Here, the conclusion can be made that the calibration ratio-type estimator under dual cosntraints should be preferred in practice as it provides consistent and more precise parameter estimates. From Tables 1 and 2, it is seen that proposed estimator is more efficient in terms of RB(%) and RRMSE(%).

Therefore, utilizing auxiliary information through calibration approach provides considerable precision of estimates under dual constraints as compared to single constraint, so it is strongly recommended to consider the estimator for application in real scenario as supported by present study.

References

Alka, Rai, P.K., and M. Qasim. (2019). Two-step calibration of design weights under two auxiliary variables in sample survey. *Journal of Statistical Computation and Simulation*, 89(12),2316–2327.

Alka, Ekpenyong, E.J., Rai, P.K, and Tiwari, S. (2021). Calibration estimator of population mean under stratified systematic sampling design. *International Journal of Agricultural and Statistical Sciences*, 17, 2353–61.

Choudhury, S. and Singh, B. K. (2012a). An Efficient Class of Dual to Product-cum-Dual to Ratio Estimators of Finite Population Mean in Sample Surveys. *Global Journal of Science Frontier Research Mathematics & Decision Sciences*, 12(3): Version 1.0. ISSN: 2249–4626.

Choudhury, S. and Singh, B. K. (2012b). A Class of Chain ratio-cum-dual to Ratio Type Estimator with Two Auxiliary Characters under Double Sampling in Sample Surveys. *Statistics in Transition- New Series*, 13(1), 65–78.

Clement, E. P. and Enang, E. I. (2017). On the Efficiency of Ratio Estimator over the Regression Estimator. *Communications in Statistics-Theory and Methods*, 46 (11), 5357-5367.

Cochran, W. G. (1977). Sampling Techniques, 3rd edition, John Wiley, New York.

Clement, E. P., Udofia, G. A. and Enang, E. I. (2014). Sample design for domain calibration estimators. *International Journal of Probability and Statistics*, 3(1), 8-14.

Clement, E. P. and Enang, E. I. (2015). Calibration approach alternative ratio estimator for population mean in stratified sampling. *International Journal of Statistics and Economics*, 16(1), 83-93.

Deville, J. C. and Särndal, C. E. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 87, 376-382.

Estevao, V. M. and Särndal, C. E. (2006). Survey estimates by calibration on complex auxiliary information. *International Statistical Review*, 74, 127-147.

Etuk, S. I., Enang, E. I. and Ekpenyong, E. J. (2016). A Modified Class of Ratio Estimators for Population Mean. *Communications in Statistics -Theory and Methods*, 45(4), 837-849.

Kim, J. K. and Park, M. (2010). Calibration Estimation in Survey Sampling. *International Statistical Review*, 78(1), 21-29.

Koyuncu, N. and Kadilar, C. (2013). Calibration Estimators Using Different Measures in Stratified Random Sampling. *International Journal of Modern Engineering Research*, 3(1), 415-419.

Koyuncu, N. and Kadilar, C. (2016). Calibration weighting in stratified random sampling. *Communications in Statistics-Simulation and Computation*, 45(7), 2267-2275.

Onyeka, A.C. (2012). Estimation of Population Mean in Post-Stratified Sampling using Known Value of Some Population Parameter(s). *Statistics in Transition-New Series*, 13(1), 65–78.

Onyeka, A.C. (2013). Dual to Ratio Estimators of Population Mean in Post-Stratified Sampling using Known Value of Some Population Parameters. *Global Journal of Science Frontier Research Mathematics & Decision Sciences*, 13(2), Version 1.0. ISSN: 2249-4626.

Ozgul, N. (2019). New calibration estimator based on two auxiliary variables in stratified sampling. *Communications in Statistics-Theory and Methods*, 48(6), 1481-1492.

Rai, P. K., Tiwari, S., & Alka. (2024). Calibration approach-based estimator of population total under successive sampling design. *Communications in Statistics - Theory and Methods*, 1–15.

Särndal, C. E. (2007). The Calibration Approach in Survey Theory and Practice. *Survey Methodology*, 33, 99-119.

Singh, S. (2013) A Dual Problem of Calibration of Design Weights, *Statistics*, 47(3), 566-574.

Singh, H. P., Rathour, A. and Solanki, R. S. (2012). An Improved Dual to Chain Ratio Type Estimator for the Population Mean. *J. Stat*, 1(3), 1–6.

Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. (2009). Improvement in Estimating the Population Mean using Exponential Estimator in Simple Random Sampling. *Bulletin of Statistics & Economics*, 3(9), 13–18.

Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. (2011). Improved Exponential Estimator for Population Variance Using Two Auxiliary Variables. *Italian Journal of Pure and Applied Mathematics*, 28, 101–108.

Solanki, R. S., Singh, H. P. and Rathour, A. (2012): An Alternative Estimator for Estimating the Finite Population Mean Using Auxiliary Information in Sample Surveys. ISRN Probability and Statistics Article ID 657682, doi:10.5402/2012/657682.

Zaman, T., and Bulut, H. (2023). Robust calibration for estimating the population mean using stratified random sampling. *Scientia Iranica*, DOI: 0.24200/SCI.2023.59408.6224.