DIGITALCOMMONS
@WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 17 | Issue 1 Article 20

7-19-2018

Estimation of finite population mean by using minimum and maximum values in stratified random sampling

Umer Daraz
Quaid-i-Azam University, umerdaraz09@yahoo.com

Javid Shabbir
Quaid-i-Azam University, javidshabbir@gmail.com

Hina Khan
Government College University, hinakhan@gcu.edu.pk

Recommended Citation

Daraz, Umer; Shabbir, Javid; and Khan, Hina (2018) "Estimation of finite population mean by using minimum and maximum values in stratified random sampling," *Journal of Modern Applied Statistical Methods*: Vol. 17 : Iss. 1 , Article 20.
DOI: 10.22237/jmasm/1532007537
Available at: https://digitalcommons.wayne.edu/jmasm/vol17/iss1/20

Estimation of finite population mean by using minimum and maximum values in stratified random sampling

Cover Page Footnote

Dear Editor, We have revised the manuscript as referees have suggested.

Journal of Modern Applied Statistical Methods
May 2018, Vol. 17, No. 1, eP2548
doi: 10.22237/jmasm/1532007537

ISSN 1538 − 9472

Estimation of Finite Population Mean by Using Minimum and Maximum Values in Stratified Random Sampling

Ummer Draz
Quaid-i-Azam University
Islamabad, Pakistan

Javid Shabbir
Quaid-i-Azam University
Islamabad, Pakistan

Hina Khan
Government College University
Lahore, Pakistan

In this paper we have suggested an improved class of ratio type estimators in estimating the finite population mean when information on minimum and maximum values of the auxiliary variable is known. The properties of the suggested class of estimators in terms of bias and mean square error are obtained up to first order of approximation. Two data sets are used for efficiency comparisons.

Keywords: Minimum and maximum values, stratified random sampling, mean squared error, efficiency

Introduction

In survey sampling, auxiliary information is frequently used in various forms to increase the precision of estimators by taking advantage of the correlation between the study and auxiliary variables. In the proposed class of estimators, the minimum and maximum values of the auxiliary variable are used. Mohanty and Sahoo (1995), Khan and Shabbir (2013), Khan (2015), Walia, Kaur, and Sharma (2015), and Cekim and Cingi (2016) proposed some estimators for the estimation of finite population mean under minimum and maximum values using known information of the auxiliary variable.

Consider a finite population $\mathbf{U} = (U_1, U_2, U_3, \ldots, U_N)$ of size N which is divided into L strata, each of size N_h ($h = 1, 2, \ldots, L$), such that $\sum_{h=1}^{L} N_h = N$. Let Y_{hi} and X_{hi} be the values of the study variable and the auxiliary variable, respectively, in the h^{th} stratum of the i^{th} unit, $i = 1, 2, \ldots, N_h$. We select a sample of size n_h from the h^{th} stratum by using simple random sampling without replacement. Let

doi: 10.22237/jmasm/1532007537 | Accepted: May 24, 2017; Published: July 19, 2018.
Correspondence: Javid Shabbir, javidshabbir@gmail.com

$$\bar{Y}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} Y_{hi} \quad \text{and} \quad \bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} X_{hi}$$

be the population means of the study variable Y and the auxiliary variable X in the h^{th} stratum corresponding to the population means

$$\bar{Y} = \frac{1}{N_h}\sum_{h=1}^{L} W_h \bar{Y}_h \quad \text{and} \quad \bar{X} = \frac{1}{N_h}\sum_{h=1}^{L} W_h \bar{X}_h$$

respectively, where $W_h = N_h/N$ is the known stratum weight. Let

$$S_{yh}^2 = \frac{\sum_{i=1}^{N_h}\left(Y_{hi} - \bar{Y}_h\right)^2}{N_h - 1} \quad \text{and} \quad S_{xh}^2 = \frac{\sum_{i=1}^{N_h}\left(X_{hi} - \bar{X}_h\right)^2}{N_h - 1}$$

be the population variances in the h^{th} stratum. Let $C_{yh} = S_{yh}/\bar{Y}_h$ and $C_{xh} = S_{xh}/\bar{X}_h$ be the population coefficients of variations in the h^{th} stratum. Let $C_{yxh} = \rho_{yxh}C_{yh}C_{xh}$ be the population co-variance between the study variable (Y) and the auxiliary variable (X) in the h^{th} stratum. Let

$$\bar{y}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} Y_{hi} \quad \text{and} \quad \bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} X_{hi}$$

be the sample means of the study variable (Y) and the auxiliary variable (X) in the h^{th} stratum. The usual unbiased estimator for the population mean $\bar{Y}$ is given by

$$\bar{y}_{\text{st}} = \sum_{h=1}^{L} W_h \bar{y}_h \tag{1}$$

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

$$\mathrm{V}\left(\bar{y}_{\text{st}}\right) = \sum_{h=1}^{L} W_h^2 \theta_h \bar{Y}_h^2 C_{yh}^2 \tag{2}$$

where

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

The usual ratio estimator for $\bar{Y}$ in stratified random sampling is given by

$$\bar{y}_{\mathrm{Rs}} = \sum_{h=1}^{L} W_h \bar{y}_h \left(\frac{\bar{X}_h}{\bar{x}_h} \right) \tag{3}$$

The bias and mean square error of $\bar{y}_{\mathrm{Rs}}$ in stratified random sampling are given by

$$\mathrm{Bias}\left(\bar{y}_{\mathrm{Rs}}\right) \cong \sum_{h=1}^{L} W_h \theta_h \bar{Y}_h \left(C_{xh}^2 - C_{ysh} \right) \tag{4}$$

and

$$\mathrm{MSE}\left(\bar{y}_{\mathrm{Rs}}\right) \cong \sum_{h=1}^{L} W_h^2 \theta_h \bar{Y}_h^2 \left(C_{yh}^2 + C_{xh}^2 - 2C_{ysh} \right) \tag{5}$$

The separate linear regression estimator $\bar{y}_{\mathrm{lrs}}$ of the population mean $\bar{Y}$ is given by

$$\bar{y}_{\mathrm{lrs}} = \sum_{h=1}^{L} W_h \left[\bar{y}_h + \hat{\beta}_{yxh} \left(\bar{X}_h - \bar{x}_h \right) \right] \tag{6}$$

where

$$\hat{\beta}_{yxh} = \frac{s_{yxh}}{s_{xh}^2}$$

is the sample regression coefficient whose population regression coefficient is β_{yxh}.

Expressions for bias and mean square error of $\bar{y}_{\mathrm{lrs}}$ are given by

$$\mathrm{Bias}\left(\bar{y}_{\mathrm{lrs}}\right) \cong \sum_{h=1}^{L} W_h \theta_h \beta_{yxh} \bar{X}_h C_{xh} \left[\lambda_{03h} - \frac{\lambda_{12h}}{\rho_{yxh}} \right] \tag{7}$$

and

$$\mathrm{MSE}\left(\bar{y}_{\mathrm{lrs}}\right) \cong \sum_{h=1}^{L} W_h^2 \theta_h \bar{Y}_h^2 C_{yh}^2 \left(1-\rho_{yxh}^2\right) \tag{8}$$

where

$$\lambda_{rsh} = \frac{\mu_{rsh}}{\mu_{02h}^{r/2} \mu_{20h}^{s/2}}$$

with

$$\mu_{rsh} = \frac{\sum_{i=1}^{N_h} \left(Y_{hi} - \bar{Y}_h\right)^r \left(X_{hi} - \bar{X}_h\right)^2}{N_h - 1} \quad \text{for } h = 1, 2, 3, \ldots, L$$

Mohanty and Sahoo (1995) proposed two estimators by using known information of minimum and maximum values of the auxiliary variable. The transformations are given by

$$v_h = \frac{X_{hi} + X_{\mathrm{m}h}}{X_{\mathrm{M}h} + X_{\mathrm{m}h}} \tag{9}$$

and

$$z_h = \frac{X_{hi} + x_{\mathrm{M}h}}{X_{\mathrm{M}h} + X_{\mathrm{m}h}} \tag{10}$$

where $X_{\mathrm{m}h}$ and $X_{\mathrm{M}h}$ are the minimum and maximum values of the auxiliary variable in the h^{th} stratum.

In stratified sampling, the ratio estimators are given by

$$\bar{y}_{h1\mathrm{s}} = \sum_{h=1}^{L} W_h \bar{y}_h \left(\frac{\bar{V}_h}{\bar{v}_h}\right) \tag{11}$$

and

$$\bar{y}_{h2s} = \sum_{h=1}^{L} W_h \bar{y}_h \left(\frac{\bar{Z}_h}{\bar{z}_h} \right) \tag{12}$$

Expressions for biases and mean squared errors of these estimators are given by

$$\text{Bias}\left(\bar{y}_{h1s}\right) \cong \sum_{h=1}^{L} W_h \theta_h \bar{Y}_h \left(t_{1h}^2 C_{xh}^2 - t_{1h} C_{yxh} \right) \tag{13}$$

$$\text{Bias}\left(\bar{y}_{h2s}\right) \cong \sum_{h=1}^{L} W_h \theta_h \bar{Y}_h \left(t_{2h}^2 C_{xh}^2 - t_{2h} C_{yxh} \right) \tag{14}$$

and

$$\text{MSE}\left(\bar{y}_{h1s}\right) \cong \sum_{h=1}^{L} W_h^2 \theta_h \bar{Y}_h^2 \left(C_{yh}^2 - 2t_{1h} C_{yxh} + t_{1h}^2 C_{xh}^2 \right) \tag{15}$$

$$\text{MSE}\left(\bar{y}_{h2s}\right) \cong \sum_{h=1}^{L} W_h^2 \theta_h \bar{Y}_h^2 \left(C_{yh}^2 - 2t_{2h} C_{yxh} + t_{2h}^2 C_{xh}^2 \right) \tag{16}$$

where

$$t_{1h} = \frac{\bar{X}_h}{\bar{X}_h + X_{mh}} \quad \text{and} \quad t_{2h} = \frac{\bar{X}_h}{\bar{X}_h + X_{Mh}}$$

Walia et al. (2015) proposed some estimators by using known information on minimum and maximum values of the auxiliary variable. The transformation is given by

$$z_{1h} = X_{hi} + \frac{X_{Mh}}{X_{mh}} \tag{17}$$

Walia et al. (2015) defined two estimators:

$$\bar{y}_{m1s} = \sum_{h=1}^{L} W_h \bar{y}_h \left(\frac{\bar{Z}_{1h}}{\bar{z}_{1h}} \right) \tag{18}$$

and

$$\bar{y}_{\mathrm{m2s}} = \sum_{h=1}^{L} W_h \bar{y}_h \left(\frac{\bar{Z}_{1h} + C_{z1h}}{\bar{z}_{1h} + C_{z1h}} \right) \tag{19}$$

where

$$C_{z1h} = \frac{S_{z1h}}{\bar{Z}_{1h}} = \frac{S_{xh}}{\bar{X}_h + \dfrac{X_{\mathrm{M}h}}{X_{\mathrm{m}h}}}$$

Expressions for biases and mean square errors of these estimators, up to first order approximation, are given by

$$\mathrm{Bias}\left(\bar{y}_{\mathrm{m1s}}\right) \cong \sum_{h=1}^{L} W_h \theta_h \bar{Y}_h \left(C_{1h}^2 C_{xh}^2 - C_{1h} \rho_{yxh} \right) \tag{20}$$

$$\mathrm{Bias}\left(\bar{y}_{\mathrm{m2s}}\right) \cong \sum_{h=1}^{L} W_h \theta_h \bar{Y}_h \left(C_{2h}^2 C_{xh}^2 - C_{2h} \rho_{yxh} \right) \tag{21}$$

and

$$\mathrm{MSE}\left(\bar{y}_{m1s}\right) \cong \sum_{h=1}^{L} W_h^2 \theta_h \bar{Y}_h^2 \left(C_{yh}^2 - 2C_{1h} \rho_{yxh} + C_{1h}^2 C_{xh}^2 \right) \tag{22}$$

$$\mathrm{MSE}\left(\bar{y}_{m2s}\right) \cong \sum_{h=1}^{L} W_h^2 \theta_h \bar{Y}_h^2 \left(C_{yh}^2 - 2C_{2h} \rho_{yxh} + C_{2h}^2 C_{xh}^2 \right) \tag{23}$$

where

$$C_{1h} = \frac{\bar{X}_h}{\bar{X}_h + \dfrac{X_{\mathrm{M}h}}{X_{\mathrm{m}h}}} \quad \text{and} \quad C_{2h} = \frac{\bar{X}_h \left(\bar{X}_h + \dfrac{X_{\mathrm{M}h}}{X_{\mathrm{m}h}} \right)}{\left(\bar{X}_h + \dfrac{X_{\mathrm{M}h}}{X_{\mathrm{m}h}} \right)^2 + S_{xh}}$$

Proposed Estimator

Motivated by Koyuncu (2012), an improved class of estimators is proposed for estimating the finite population mean using known information on minimum and maximum values of the auxiliary variable under stratified random sampling. The proposed estimator is given by

$$\bar{y}_{\text{DS}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{y}_h \left(\frac{\bar{X}_h}{\bar{x}_h} \right)^{\alpha_1} + k_{2h} \left(\frac{\bar{X}_h}{\bar{x}_h} \right)^{\alpha_2} \right] \exp\left[\frac{\left(\bar{X}_h - \bar{x}_h\right)}{\left(\bar{X}_h + \bar{x}_h\right) + 2 j_h} \right] \qquad (24)$$

where k_{ih}, $i = 1, 2$, are unknown constants whose values are to be determined such that the bias and mean square error are minimum, $j_h = X_{\text{M}h} - X_{\text{m}h}$, and α_1, α_2 are scalar quantities which contain the values (0, –1, 1). From (24), we obtain the different classes of estimators which are given in Table 1.

Table 1. Some classes of estimators in stratified random sampling

Subsets of $\bar{y}_{\text{DS}}$	α_1	α_2
$\bar{y}_{\text{DS1}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{y}_h + k_{2h} \left(\frac{\bar{x}_h}{\bar{X}_h} \right) \right] L_h$	0	–1
$\bar{y}_{\text{DS2}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{y}_h \left(\frac{\bar{x}_h}{\bar{X}_h} \right) + k_{2h} \right] L_h$	–1	0
$\bar{y}_{\text{DS3}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{y}_h + k_{2h} \left(\frac{\bar{X}_h}{\bar{x}_h} \right) \right] L_h$	0	1
$\bar{y}_{\text{DS4}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{y}_h \left(\frac{\bar{X}_h}{\bar{x}_h} \right) + k_{2h} \right] L_h$	1	0
$\bar{y}_{\text{DS5}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{y}_h \left(\frac{\bar{X}_h}{\bar{x}_h} \right) + k_{2h} \left(\frac{\bar{x}_h}{\bar{X}_h} \right) \right] L_h$	1	–1
$\bar{y}_{\text{DS6}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{y}_h \left(\frac{\bar{x}_h}{\bar{X}_h} \right) + k_{2h} \left(\frac{\bar{X}_h}{\bar{x}_h} \right) \right] L_h$	–1	1
$\bar{y}_{\text{DS7}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{y}_h \left(\frac{\bar{X}_h}{\bar{x}_h} \right) + k_{2h} \left(\frac{\bar{X}_h}{\bar{x}_h} \right) \right] L_h$	1	1
$\bar{y}_{\text{DS8}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{y}_h \left(\frac{\bar{x}_h}{\bar{X}_h} \right) + k_{2h} \left(\frac{\bar{x}_h}{\bar{X}_h} \right) \right] L_h$	–1	–1

where

$$L_h = \exp\left[\frac{\left(\bar{X}_h - \bar{x}_h\right)}{\left(\bar{X}_h + \bar{x}_h\right) + 2\,j_h}\right]$$

To obtain the properties of the proposed estimator, we define the following error terms: Let

$$e_{0h} = \left(\frac{\bar{y}_h - \bar{Y}_h}{\bar{Y}_h}\right) \quad \text{and} \quad e_{1h} = \left(\frac{\bar{x}_h - \bar{X}_h}{\bar{X}_h}\right)$$

such that $\mathrm{E}(e_{ih}) = 0$ for $i = 0,\ 1$, $\mathrm{E}\left(e_{0h}^2\right) = \theta_h C_{yh}^2$, $\mathrm{E}\left(e_{1h}^2\right) = \theta_h C_{xh}^2$, and $\mathrm{E}\left(e_{0h}e_{1h}\right) = \theta_h C_{yxh}$.

Rewriting (24) in terms of errors, we have

$$\bar{y}_{\mathrm{DS}} = \sum_{h=1}^{L} W_h \left[k_{1h}\bar{Y}_h \left(1+e_{0h}\right)\left(1+e_{1h}\right)^{-\alpha_1} + k_{2h}\left(1+e_{1h}\right)^{-\alpha_2} \right] \exp\left[\frac{-g_{1h}e_{1h}}{2}\left(1+\frac{g_{1h}}{2}e_{1h}\right)^{-1}\right]$$

where

$$g_{1h} = \frac{\bar{X}_h}{\bar{X}_h + X_{\mathrm{M}h} - X_{\mathrm{m}h}}$$

Up to first order of approximation,

$$\begin{aligned} \bar{y}_{\mathrm{DS}} - \bar{Y} \cong \sum_{h=1}^{L} W_h \Bigg[-\bar{Y}_h + k_{1h}\bar{Y}_h \Bigg\{ & 1 + e_{0h} - e_{1h}\left(\alpha_1 + \frac{g_{1h}}{2}\right) \\ & + e_{1h}^2 \left(\frac{\alpha_1 g_{1h}}{2} + \frac{3g_{1h}^2}{8} + \frac{\alpha_1\left(\alpha_2+1\right)}{2}\right) - e_{0h}e_{1h}\left(\alpha_1 + \frac{g_{1h}}{2}\right)\Bigg\} \qquad (25) \\ & + k_{2h}\left\{1 - e_{1h}\left(\alpha_2 + \frac{g_{1h}}{2}\right) + e_{1h}^2\left(\frac{\alpha_2 g_{1h}}{2} + \frac{3g_{1h}^2}{8} + \frac{\alpha_2\left(\alpha_2+1\right)}{2}\right)\right\} \end{aligned}$$

Using (25), the bias of $\bar{y}_{\mathrm{DS}}$, up to first order of approximation, is given by

$$\text{Bias}\left(\bar{y}_{\text{DS}}\right) \cong \sum_{h=1}^{L} W_h \left(-\bar{Y}_h + k_{1h}\bar{Y}_h D_h + k_{2h} G_h\right) \tag{26}$$

where

$$D_h = \left[1 + \theta_h C_{xh}^2 \left(\frac{\alpha_1 g_{1h}}{2} + \frac{3g_{1h}^2}{8} + \frac{\alpha_1(\alpha_1 + 1)}{2}\right) - \theta_h C_{yxh}\left(\alpha_1 + \frac{g_{1h}}{2}\right)\right]$$

and

$$G_h = \left[1 + \theta_h C_{xh}^2 \left(\frac{\alpha_2 g_{1h}}{2} + \frac{3g_{1h}^2}{8} + \frac{\alpha_2(\alpha_2 + 1)}{2}\right)\right]$$

By squaring and taking expectations on both sides of (25), obtain the mean square error up to first order of approximation, which is given by

$$\begin{aligned}\text{MSE}\left(\bar{y}_{\text{DS}}\right) \cong \sum_{h=1}^{L} W_h^2 \Big(\bar{Y}_h^2 + \bar{Y}_h^2 k_{1h}^2 A_h + k_{2h}^2 B_h - 2\bar{Y}_h^2 k_{1h} D_h \\ -2\bar{Y}_h k_{2h} G_h + 2\bar{Y}_h k_{1h} k_{2h} F_h\Big)\end{aligned} \tag{27}$$

where

$$\begin{aligned}A_h = \Bigg[1 + \theta_h \Bigg\{C_{yh}^2 + C_{xh}^2 \Bigg\{\left(\alpha_1 + \frac{g_{1h}}{2}\right)^2 + \left(\alpha_1 g_{1h} + \frac{3g_{1h}^2}{4} + \alpha_1(\alpha_1 + 1)\right)\Bigg\} \\ -4C_{yxh}\left(a_1 + \frac{g_{1h}}{2}\right)\Bigg\}\Bigg]\end{aligned}$$

$$B_h = \left[1 + \theta_h C_{xh}^2 \left\{\left(\alpha_2 + \frac{g_{1h}}{2}\right)^2 + \left(\alpha_2 g_{1h} + \frac{3g_{1h}^2}{4} + \alpha_2(\alpha_2 + 1)\right)\right\}\right]$$

and

$$F_h = \left[1+\theta_h\left\{C_{xh}^2\left\{\left(\frac{\alpha_1 g_{1h}}{2}+\frac{3g_{1h}^2}{8}+\frac{\alpha_1(\alpha_1+1)}{2}\right)+\left(\alpha_1+\frac{g_{1h}}{2}\right)\left(\alpha_2+\frac{g_{1h}}{2}\right)\right.\right.\right.$$
$$\left.\left.\left.+\left(\frac{\alpha_2 g_{1h}}{2}+\frac{3g_{1h}^2}{8}+\frac{\alpha_2(\alpha_2+1)}{2}\right)\right\}-C_{yxh}(\alpha_1+\alpha_2+g_{1h})\right\}\right]$$

The optimum values of k_{1h} and k_{2h} obtained by minimizing (27) are, respectively, given by

$$k_{1h(\text{opt})}=\frac{B_hD_h-F_hG_h}{A_hB_h-F_h^2}\quad\text{and}\quad k_{2h(\text{opt})}=\frac{\bar{Y}_h\left(A_hG_h-D_hF_h\right)}{A_hB_h-F_h^2}$$

Substituting the optimum values of k_{1h} and k_{2h} into (26) and (27) yields the minimum bias and MSE of $\bar{y}_{\text{DS}}$, given by

$$\text{Bias}\left(\bar{y}_{\text{DS}}\right)_{\min}\cong-\sum_{h=1}^{L}W_h\bar{Y}_h\left[1-\frac{\left(A_hG_h^2+B_hD_h^2-2D_hF_hG_h\right)}{A_hB_h-F_h^2}\right]\tag{28}$$

and

$$\text{MSE}\left(\bar{y}_{\text{DS}}\right)_{\min}\cong\sum_{h=1}^{L}W_h^2\bar{Y}_h^2\left[1-\frac{\left(A_hG_h^2+B_hD_h^2-2D_hF_hG_h\right)}{A_hB_h-F_h^2}\right]\tag{29}$$

Comparison of Estimators

Now, compare the proposed class of estimators $\bar{y}_{\text{DS}}$ with existing estimators $\bar{y}_{\text{st}}$, $\bar{y}_{\text{Rs}}$, $\bar{y}_{\text{lrs}}$, $\bar{y}_{h1\text{s}}$, $\bar{y}_{h2\text{s}}$, $\bar{y}_{\text{m1s}}$, and $\bar{y}_{\text{m2s}}$.

Condition (i) By (2) and (29), $V(\bar{y}_{\text{st}}) > \text{MSE}(\bar{y}_{\text{DS}})_{\min}$ if

$$\sum_{h=1}^{L}W_h^2\bar{Y}_h^2\left[\theta_hC_{yh}^2-1+\frac{\left(B_hD_h^2+A_hG_h^2-2D_hF_hG_h\right)}{A_hB_h-F_h^2}\right]>0$$

Condition (ii) By (5) and (29), MSE($\bar{y}_{Rs}$) > MSE($\bar{y}_{DS}$)$_{min}$ if

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[\theta_h \left(C_{yh}^2 + C_{xh}^2 - 2C_{yxh} \right) - 1 + \frac{\left(B_h D_h^2 + A_h G_h^2 - 2D_h F_h G_h \right)}{A_h B_h - F_h^2} \right] > 0$$

Condition (iii) By (8) and (29), MSE($\bar{y}_{lrs}$)$_{min}$ > MSE($\bar{y}_{DS}$)$_{min}$ if

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[\theta_h C_{yh}^2 \left(1 - \rho_{yxh}^2 \right) - 1 + \frac{\left(B_h D_h^2 + A_h G_h^2 - 2D_h F_h G_h \right)}{A_h B_h - F_h^2} \right] > 0$$

Condition (iv) By (15) and (29), MSE($\bar{y}_{h1s}$) > MSE($\bar{y}_{DS}$)$_{min}$ if

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[\theta_h \left(C_{yh}^2 + t_{1h}^2 C_{xh}^2 - 2t_{1h} C_{yxh} \right) - 1 + \frac{\left(B_h D_h^2 + A_h G_h^2 - 2D_h F_h G_h \right)}{A_h B_h - F_h^2} \right] > 0$$

Condition (v) By (16) and (29), MSE($\bar{y}_{h2s}$) > MSE($\bar{y}_{DS}$)$_{min}$ if

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[\theta_h \left(C_{yh}^2 + t_{2h}^2 C_{xh}^2 - 2t_{2h} C_{yxh} \right) - 1 + \frac{\left(B_h D_h^2 + A_h G_h^2 - 2D_h F_h G_h \right)}{A_h B_h - F_h^2} \right] > 0$$

Condition (vi) By (22) and (29), MSE($\bar{y}_{m1s}$) > MSE($\bar{y}_{DS}$)$_{min}$ if

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[\theta_h \left(C_{yh}^2 + C_{1h}^2 C_{xh}^2 - 2C_{1h} C_{yxh} \right) - 1 + \frac{\left(B_h D_h^2 + A_h G_h^2 - 2D_h F_h G_h \right)}{A_h B_h - F_h^2} \right] > 0$$

Condition (vii) By (23) and (29), MSE($\bar{y}_{m2s}$) > MSE($\bar{y}_{DS}$)$_{min}$ if

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left[\theta_h \left(C_{yh}^2 + C_{2h}^2 C_{xh}^2 - 2C_{2h} C_{yxh} \right) - 1 + \frac{\left(B_h D_h^2 + A_h G_h^2 - 2D_h F_h G_h \right)}{A_h B_h - F_h^2} \right] > 0$$

Numerical Examples

Two data sets are used to obtain the percent relative efficiency (PRE) of the different estimators:

Data Set 1. (Bureau of Statistics, 2013, p. 226)

Y: Employment level in 2012 by divisions.
X: Number of registered factories in 2012 by divisions.

The data are divided into two groups:

Group 1: Gujranwala, Lahore, Rawalpindi, and Sargodha divisions.
Group 2: Bahawalpur, D.G Khan, Faisalabad, Multan, and Sahiwal divisions.

The summary statistics are given below:

$N_1 = 18$, $N_2 = 18$, $n_1 = 8$, $n_2 = 8$, $\bar{Y}_1 = 85572.11$, $\bar{Y}_2 = 19293.61$, $\bar{X}_1 = 414.5556$, $\bar{X}_2 = 257$, $X_{\text{M1}} = 2055$, $X_{\text{M2}} = 1674$, $X_{\text{m1}} = 24$, $X_{\text{m2}} = 52$, $S_{y1} = 248216$, $S_{y2} = 37979.33$, $S_{x1} = 521.6751$, $S_{x2} = 365.6955$, $\rho_{yx1} = 0.3473$, $\rho_{yx2} = 0.9796$, $C_{x1} = 1.2584$, $C_{x2} = 1.423$, $C_{y1} = 2.9007$, $C_{y2} = 1.9685$

Data Set 2. (Bureau of Statistics, 2013, p. 135)

Y: Total number of students enrolled in 2012 by divisions.
X: Total number of government primary and secondary schools for boys and girls in 2012 by divisions.

The data are divided into two groups:

Group 1: Gujranwala, Lahore, Rawalpindi, and Sargodha divisions.
Group 2: Bahawalpur, D.G Khan, Faisalabad, Multan, and Sahiwal divisions.

The summary statistics are given below:

$N_1 = 18$, $N_2 = 18$, $n_1 = 8$, $n_2 = 8$, $\bar{Y}_1 = 162979.3$, $\bar{Y}_2 = 134458$, $\bar{X}_1 = 962.0556$, $\bar{X}_2 = 1146.722$, $S_{y1} = 255887.7$, $S_{y2} = 50235.82$, $S_{x1} = 307.9531$, $S_{x2} = 469.9311$, $X_{\text{M1}} = 1530$, $X_{\text{M2}} = 2370$, $X_{\text{m1}} = 388$, $X_{\text{m2}} = 582$, $\rho_{yx1} = 0.1447$, $\rho_{yx2} = 0.787$, $C_{x1} = 0.3202$, $C_{x2} = 0.4098$, $C_{y1} = 1.5701$, $C_{y2} = 0.3736$

Table 2. Percent relative efficiency of different estimators with respect to $\bar{y}_{st}$

Estimator	Data set 1	Data set 2
$\bar{y}_{\mathrm{st}}$	100.000	100.000
$\bar{y}_{\mathrm{Rs}}$	115.073	103.756
$\bar{y}_{\mathrm{lrs}}$	116.252	104.509
$\bar{y}_{h1\mathrm{s}}$	115.298	104.495
$\bar{y}_{h2\mathrm{s}}$	105.080	103.318
$\bar{y}_{\mathrm{m1s}}$	115.883	103.776
$\bar{y}_{\mathrm{m2s}}$	115.880	103.778
$\bar{y}_{\mathrm{DS1}}$	686.109	2142.170
$\bar{y}_{\mathrm{DS2}}$	131154.840	49578.150
$\bar{y}_{\mathrm{DS3}}$	1106.640	1724.010
$\bar{y}_{\mathrm{DS4}}$	75600.490	29063.170
$\bar{y}_{\mathrm{DS5}}$	893.565	3726.730
$\bar{y}_{\mathrm{DS6}}$	2145.380	2142.520
$\bar{y}_{\mathrm{DS7}}$	770.943	1235.560
$\bar{y}_{\mathrm{DS8}}$	782.772	2920.860

We use the following expression for efficiency comparisons:

$$PRE = \frac{\mathrm{V}\left(\bar{y}_{\mathrm{st}}\right)}{\mathrm{MSE}\left(\bar{y}_{k}\right)} \times 100$$

where k is one of Rs, lrs, h1s, h2s, m1s, m2s, or DSi (i = 1, 2,…, 8). The percent relative efficiencies are summarized for two data sets in Table 2.

Conclusion

An improved class of estimators was proposed in estimating the finite population mean using known minimum and maximum values of the auxiliary variable. Shown in Table 2 are the percent relative efficiencies of all estimators over the usual mean per unit estimator $\bar{y}_{\mathrm{st}}$ in stratified random sampling. It was observed that the performance of the suggested class of estimators $\bar{y}_{\mathrm{DS}i}$ (i = 1, 2,…, 8) is better as compared to all other considered estimators. Among the suggested class of estimators, $\bar{y}_{\mathrm{DS2}}$ is preferable because of least MSE.

References

Bureau of Statistics. (2013). *Punjab development statistics*. Lahore, Pakistan: Bureau of Statistics, Government of the Punjab. Retrieved from http://www.bos.gop.pk/system/files/Dev-2013.pdf

Cekim, H. O., & Cingi, H. (2016). Some estimator types for population mean using linear transformation with the help of the minimum and maximum values of the auxiliary variable. *Hacettepe Journal of Mathematics and Statistics, 46*(4). doi: 10.15672/hjms.201510114186

Khan, M. (2015). Improvement in estimating the finite population mean under maximum and minimum values in double sampling scheme. *Journal of Statistics Applications & Probability Letters*, *2*(2), 115-121. Retrieved from http://www.naturalspublishing.com/ContIss.asp?IssID=262

Khan, M., & Shabbir, J. (2013). Some improved ratio, product, and regression estimators of finite population mean when using minimum and maximum values. *The Scientific World Journal*, *2013*, 431868. doi: 10.1155/2013/431868

Koyuncu, N. (2012). Efficient estimators of population mean using auxiliary attributes. *Applied Mathematics and Computation, 218*(22), 10900-10905. doi: 10.1016/j.amc.2012.04.050

Mohanty, S., & Sahoo, J. (1995). A note on improving the ratio method of estimation through linear transformation using certain known population parameters. *Sankhyā: The Indian Journal of Statistics, Series B, 57*(1), 93-102. Available from https://www.jstor.org/stable/25052879

Walia, G. S., Kaur, H., & Sharma, M. (2015). Ratio type estimator of population mean through efficient linear transformation. *American Journal of Mathematics and Statistics, 5*(3), 144-149. Retrieved from http://article.sapub.org/10.5923.j.ajms.20150503.06.html