CALIBRATION APPROACH RATIO ESTIMATORS OF POPULATION MEDIAN IN STRATIFIED RANDOM SAMPLING DESIGN

Abstract

Median as a measure of location gives a more robust estimate than the mean when dealing with heavy tailed or skewed distributions. It can also be used in cases of qualitative variables and open end intervals. Calibration, an approach that adjusts the original design weight by incorporating auxiliary information is employed using the chi square distance measure on a ratio median estimator under stratified random sampling to propose some estimators of population median. These proposed estimators are: the regression and ratio-type calibrated estimators with one constraint and the regression and ratio-type calibrated estimators with two constraints. The estimators of variance of these proposed estimators are also obtained. Empirical investigations on the performance of these estimators are carried out using R software simulated data set under underlying distributional assumptions of Cauchy and Lognormal, for sample sizes of 10%, 20% and 25%. The results showed that the proposed regression and ratio-type calibrated estimators with one constraint and the regression-type calibrated estimator with two constraints were more efficient than the existing ratio estimator and the proposed ratio-type calibrated estimator under two constraints for both Cauchy and the lognormal distributions.

Keywords: Median, Calibration, Ratio-type estimator, Regression-type estimator, stratified random sampling, Auxiliary variables etc.

1 Introduction

The median is a commonly used measure of properties of a data set in statistics and is most times regarded as a more appropriate measure of location than the mean because it is not skewed so much by small proportions of extremely large or small values and so may give a better idea of a typical value. Auxiliary variables provide more efficient estimators through the relationship that exist with the study variable. Whenever there is auxiliary information available, the investigator may want to use it in the method of estimation through either the ratio, product or regression estimators to obtain more efficient results. Calibration was first introduced by Deville and Sarndal [5] as a technique for minimizing a distance measure between an initial weight and a calibrated weight subject to a single calibration constraint. Also, Singh and Arnab [20]obtained a calibrated mean estimator using two constraints. The second constraint confirmed that the sum of calibrated weight was equal to the sum of design weight.

In stratified random sampling, calibration approach is used to adjust the strata weight for improving the precision of survey estimates of population parameters. In existing literatures, Gross [9]considered median estimation in stratified random sampling without replacement. Kuk and Mak [14] proposed a ratio estimator for population median in simple random sampling which was a simple modification of the ratio estimator for mean alongside with two other estimators derived from different approaches. Singh, Singh and Upadhyaya [20] proposed a chain-ratio and regression-type median estimators. Aladag and Cingi [2] motivated by Kuk and

Mak, suggested ratio median estimators in stratified random sampling. They also dealt with the estimation of population median in simple and stratified random sampling using auxiliary information such as mode, range and correlation coefficient. Though, these proposed estimators are improvement over existing median estimators, no effort has yet been made to developing a median estimator using calibration approach.

In this study, new median estimators are developed using the approach of calibration under single and double constraints.

2. Existing Median Estimators in Stratified Random Sampling

Let $U = (U_1, U_2, ...U_N)$ be a finite population of size N, Y be the study variable and X the auxiliary variable associated with each unit $U_i, (i = 1,2,3...N)$ of the population.

Suppose the population consist of H strata with $N_h, (h = 1,2,3...H)$ units in the h^{th} stratum such that $\sum_{h=1}^{H} N_h = N$. A simple random sample without replacement of size n_h is drawn from the h^{th} population stratum N_h, such that $\sum_{h=1}^{H} n_h = n$. Let y_{hi} be the characteristic of interest for the i^{th} element in the h^{th} stratum and x_{hi} be the auxiliary information for the i^{th} element in the h^{th} stratum.

Assuming that M_{hx} and M_{hy} are population medians in the h^{th} stratum for both the auxiliary and the study variable, m_{hx} and m_{hy} are respective sample medians, $\rho_{M_{hX} M_{hY}}$ is the correlation coefficient in the h^{th} stratum between sampling distributions of M_{hx} and M_{hy} which is defined as $\rho_{M_{hX} M_{hY}} = 4 p_{h_{11}} - 1$ and $p_{h_{11}}$ is the proportion of units in the h^{th} stratum of the population with $Y_h < M_{hy}$ and $X_h < M_{hx}$. Let $M_x = \sum_{h=1}^{H} W_h M_{hx}$ and $M_y = \sum_{h=1}^{H} W_h M_{hy}$ be the population weighted medians, $m_{xst} = \sum_{h=1}^{H} W_h m_{hx}$ and $m_{yst} = \sum_{h=1}^{H} W_h m_{hy}$ be the sample weighted medians for X and Y respectively.

Aladag and Cingi (2015) proposed a separate ratio median estimator under stratified random sampling as:

$$\hat{M}_{YR1} = \sum_{h=1}^{H} W_h \frac{m_{hy}}{m_{hx}} M_{hx} \qquad (1)$$

with:

$$Bias(\hat{M}_{YR1}) = \sum_{h=1}^{H} W_h \lambda_h M_{hy} \left(C_{M_{hx}}^2 - \rho_{M_{hx} M_{hy}} C_{M_{hy}} C_{M_{hx}} \right) \qquad (2)$$

and variance

$$V(\hat{M}_{YR1}) = \sum_{h=1}^{H} W_h^2 \lambda_h M_{hy}^2 (C_{M_{hy}}^2 + C_{M_{hx}}^2 - 2\rho_{M_{hx}M_{hy}} C_{M_{hy}} C_{M_{hx}}) \quad (3)$$

where

$$\lambda_h = \frac{(1-f_h)}{4n_h}$$

$$C_{M_{hy}}^2 = \left[M_{hy} F_{hy}(M_{hy})\right]^{-2}$$

$$C_{M_{hx}}^2 = \left[M_{hx} F_{hx}(M_{hx})\right]^{-2}$$

$F_{hx}(M_{hx})$ and $F_{hy}(M_{hy})$ are the distribution functions of the auxiliary and the study variables X and Y respectively.

3.1 The Proposed Calibrated Ratio Estimators Under One Constraint

Motivated by Kuk and Mak (1989) and Aladag and Cingi (2015), a new ratio median estimator under one constraint is given as:

$$\hat{M}_{YRC1} = \sum_{h=1}^{H} W_{h1}^* \frac{m_{hy}}{m_{hx}} M_{hx} \quad (4)$$

With W_{h1}^* (calibrated weight) is obtained such that a chi square distance measure of the form:

$$\sum_{h=1}^{H} \frac{(W_{h1}^* - W_h)^2}{2W_h q_h} \quad (5)$$

is minimized subject to a single constraint:

$$\sum_{h=1}^{H} W_{h1}^* m_{hx} = M_x \quad (6)$$

Where W_{h1}^* is the calibrated weight, q_h are positive weights uncorrelated with W_{h1}^*, called the tuning parameter and W_h is the initial design weight.

Minimizing (5) subjected to (6) gives the calibrated weight as:

$$W_{h1}^* = W_h + \frac{W_h q_h m_{hx}}{\sum_{h=1}^{H} W_h q_h m_{hx}^2} (M_x - \sum_{h=1}^{H} W_h m_{hx}) \quad (7)$$

Substituting (7) into (4) gives

$$\hat{M}_{YRC1} = \sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}} + \frac{\sum_{h=1}^{H} W_h q_h m_{hy} M_{hx}}{\sum_{h=1}^{H} W_h q_h m_{hx}^2} (M_x - \sum_{h=1}^{H} W_h m_{hx}) \quad (8)$$

Equation (8) can be written in form of a generalized regression (GREG) estimator as:

$$\hat{M}_{YRC1} = \hat{M}_{YR} + \hat{\beta}_{st.1}(M_x - m_{xst}) \quad (9)$$

Where $\hat{M}_{YR} = \sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}}$ is the estimator in (1), $m_{xst} = \sum_{h=1}^{H} W_h m_{hx}$ is the estimator of the auxiliary variable M_x and $\hat{\beta}_{st.1} = \frac{\sum_{h=1}^{H} W_h q_h m_{hy} M_{hx}}{\sum_{h=1}^{H} W_h q_h m_{hx}^2}$

By letting $q_h = 1$, equation (8) gives

$$\hat{M}_{YRC11} = \sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}} + \frac{\sum_{h=1}^{H} W_h m_{hy} M_{hx}}{\sum_{h=1}^{H} W_h m_{hx}^2}(M_x - \sum_{h=1}^{H} W_h m_{hx}) \quad (10)$$

which is the proposed regression-type calibrated estimator of the population median M_y in stratified random sampling under single constraint.

Also, letting $q_h = \frac{1}{m_{hx}}$, equation (8) becomes the proposed ratio-type calibrated estimator of the population median M_y in stratified random sampling under single constraint i.e

$$\hat{M}_{YRC12} = \sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}} + \frac{\sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}}}{\sum_{h=1}^{H} W_h m_{hx}}(M_x - \sum_{h=1}^{H} W_h m_{hx}) \quad (11)$$

3.2 Bias and Variance Estimators of the Proposed Calibrated Estimators Under One Constraint

Let $e_{hx} = \frac{m_{hx} - M_{hx}}{M_{hx}}$, and $e_{hy} = \frac{m_{hy} - M_{hy}}{M_{hy}}$ (12)

Where $E(e_{hx}) = E(e_{hy}) = 0$

$E(e_{hx}^2) = \lambda_h C_{M_{hx}}^2$

$E(e_{hy}^2) = \lambda_h C_{M_{hy}}^2$

$E(e_{hx} e_{hy}) = \lambda_h C_{M_{hx}} C_{M_{hy}} \rho_{M_{hx} M_{hy}}$

Such that $m_{hx} = M_{hx}(1 - e_{hx})$ and $m_{hy} = M_{hy}(1 - e_{hy})$ (13)

$\lambda_h = \frac{(1 - f_h)}{4n_h}$ and $\rho_{M_{hX} M_{hY}} = 4p_{h_{11}} - 1$,

$p_{h_{11}}$ is the proportion of units in the h^{th} stratum of the population with $Y_h < M_{hy}$ and $X_h < M_{hx}$.

Expressing (4) in terms of the 'e' terms in (13) above, we obtain

$$\hat{M}_{YRC1} = \sum_{h=1}^{H} W_{h1}^{*} M_{hy} (1 + e_{hy})(1 + e_{hx})^{-1} \quad (14)$$

Assuming $|e_{hx}| \leq 1$, the Taylor series expansion on $(1+e_{hx})^{-1}$ is obtained. By substituting, multiplying out and ignoring terms with $e_x e_y$ of higher order than two, we obtain:

$$\hat{M}_{YRC1} = \sum_{h=1}^{H} W_{h1}^{*} M_{hy} (1+e_{hy} - e_{hx} - e_{hx} e_{hy} + e_{hx}^{2}) \tag{15}$$

$$(\hat{M}_{YRC1} - M_Y) = \sum_{h=1}^{H} W_{h1}^{*} M_{hy} (e_{hy} - e_{hx} - e_{hx} e_{hy} + e_{hx}^{2}) \tag{16}$$

Taking the expectation of both sides of (16) and using the results for the expectation of the 'e' terms above gives the bias to the first order of approximation as:

$$Bias(\hat{M}_{YRC1}) = E(\hat{M}_{YRC1} - M_y) = \sum_{h=1}^{H} W_{h1}^{*} M_{hy} E(e_{hy} - e_{hx} - e_{hx} e_{hy} + e_{hx}^{2})$$

$$Bias(\hat{M}_{YRC1}) = \sum_{h=1}^{H} W_{h1}^{*} M_{hy} \lambda_h (C_{M_{hx}}^{2} - C_{M_{hy}} C_{M_{hx}} \rho_{M_{hx} M_{hy}}) \tag{17}$$

Squaring both sides of (16) , taking its expectation, substituting the results of the e terms above and retaining terms to the second degree, gives the variance to the first order of approximation as:

$$V_T(\hat{M}_{YRC1}) = E(\hat{M}_{YRC1} - M_Y)^2 = \sum_{h=1}^{H} (W_{h1}^{*})^2 M_{hy}^{2} \lambda_h (C_{M_{hy}}^{2} + C_{M_{hx}}^{2} - 2C_{M_{hx}} C_{M_{hy}} \rho_{M_{hx} M_{hy}}) \tag{18}$$

And variance estimator:

$$\hat{V}_T(\hat{M}_{YRC1}) = \sum_{h=1}^{H} (W_{h1}^{*})^2 m_{hy}^{2} \lambda_h (\hat{C}_{M_{hy}}^{2} + \hat{C}_{M_{hx}}^{2} - 2\hat{C}_{M_{hx}} \hat{C}_{M_{hy}} \hat{\rho}_{M_{hx} M_{hy}})$$

or

$$\hat{V}_T(\hat{M}_{YRC1}) = \sum_{h=1}^{H} (W_{h1}^{*})^2 \lambda_h \left((\hat{F}_{hy}(M_{hy}))^{-2} + \hat{R}_m^{2} (F(M_{hx}))^{-2} - 2\hat{R}_m (\hat{F}_{hx}(M_{hx}) \hat{F}_{hy}(M_{hy}))^{-1} \hat{\rho}_{M_{hx} M_{hy}} \right) \tag{19}$$

Where

$$\hat{R}_m = \frac{m_{hy}}{m_{hx}} \tag{20}$$

$\hat{\rho}_{M_{hx} M_{hy}}$ = the correlation coefficient in the h[th] stratum sample between the sampling distributions of m_{hx} and m_{hy}

$$\hat{C}_{M_{hy}}^{2} = \left(m_{hy} \hat{F}(M_{hy}) \right)^{-2} \tag{21}$$

$$\hat{C}_{M_{hx}}^{2} = \left(m_{hx} \hat{F}(M_{hx}) \right)^{-2} \tag{22}$$

$$\hat{F}_{hy}(M_{hy}) = \frac{\sum_{i=1}^{n_h} \pi_{hi}^{-1} \Delta(M_{hy} - Y_{hi})}{\sum_{i=1}^{n_h} \pi_{hi}^{-1}} \qquad (23)$$

$$\hat{F}_{hx}(M_{hx}) = \frac{\sum_{i=1}^{n_h} \pi_{hi}^{-1} \Delta(M_{hx} - X_{hi})}{\sum_{i=1}^{n_h} \pi_{hi}^{-1}} \qquad (24)$$

$\pi_{hi} = \frac{n_h}{N_h}$ is the inclusion probability for the i^{th} element in the h^{th} stratum and

$$\Delta_{(x)} = \begin{cases} 1, x \geq 0 \\ 0, elsewhere \end{cases}$$

Now substituting for the calibrated weight from (7) we obtain:

$$(W_{h1}^*)^2 = \left[W_h + \frac{W_h q_h m_{hx}}{\sum_{h=1}^{H} W_h q_h m_{hx}^2} (M_x - \sum_{h=1}^{H} W_h m_{hx}) \right]^2 \qquad (25)$$

Setting $q_h = 1$, we obtain

$$(W_{h11}^*)^2 = \left[W_h + \frac{W_h m_{hx}}{\sum_{h=1}^{H} W_h m_{hx}^2} (M_x - \sum_{h=1}^{H} W_h m_{hx}) \right]^2 \qquad (26)$$

Also, setting $q_h = \frac{1}{m_{hx}}$ we obtain

$$(W_{h12}^*)^2 = \left(\frac{W_h M_x}{\sum_{h=1}^{H} W_h m_{hx}} \right)^2 \qquad (27)$$

Substituting (26) into (19) we obtain the variance estimator for Regression-type Ratio calibrated estimator under one constraint as:

$$\hat{V}_T(\hat{M}_{YRC11}) = \sum_{h=1}^{H} \left(W_h + \frac{W_h m_{hx}}{\sum_{h=1}^{H} W_h m_{hx}^2} (M_x - \sum_{h=1}^{H} W_h m_{hx}) \right)^2 \lambda_h \left((\hat{F}_{hy}(M_{hy}))^{-2} + \hat{R}_m^2 (\hat{F}_{hx}(M_{hx}))^{-2} \right.$$

$$\left. - 2 \hat{R}_m (\hat{F}_{hx}(M_{hx}) \hat{F}_{hy}(M_{hy}))^{-1} \hat{\rho}_{MhxMhy} \right) \qquad (28)$$

Also, putting (27) into (19) we obtain the variance estimator for the ratio-type calibrated estimator under one constraint as:

$$\hat{V}_T(\hat{M}_{YRC12}) = E(\hat{M}_{YRC1} - M_Y)^2 = \sum_{h=1}^{H}\left(\frac{W_h M_x}{\sum_{h=1}^{H} W_h m_{hx}}\right)^2 \lambda_h\Bigg((\hat{F}_{hy}(M_{hy}))^{-2} + \hat{R}_m^2 (\hat{F}_{hx}(M_{hx}))^{-2} \text{ (107)}$$

$$- 2\hat{R}_m (\hat{F}_{hx}(M_{hx}) \hat{F}_{hy}(M_{hy}))^{-1} \hat{\rho}_{MhxMhy} \Bigg) \tag{29}$$

4.1 The Proposed Calibrated Ratio Estimators Under Two Constraints

Motivated by Singh & Arnab (2014), a new median estimator is also obtained using calibration under two constraints as:

$$\hat{M}_{YRC2} = \sum_{h=1}^{H} \frac{W_{h2}^{*} m_{hy} M_{hx}}{m_{hx}} \tag{30}$$

where W_{h2}^{*} are calibrated weights chosen to minimize the chi-square distance measure:

$$\sum_{h=1}^{H} \frac{(W_{h2}^{*} - W_h)^2}{2W_h q_h} \tag{31}$$

subject to two constraints:

$$\sum_{h=1}^{H} W_{h2}^{*} = \sum_{h=1}^{H} W_h \tag{32}$$

$$\sum_{h=1}^{H} W_{h2}^{*} m_{hx} = M_x \tag{33}$$

The calibrated weights under two constraints are obtained by minimizing (31) subject to (32) and (33) as:

$$W_{h2}^{*} = W_h + \frac{W_h q_h m_{hx} (\sum_{h=1}^{H} W_h q_h) - W_h q_h (\sum_{h=1}^{H} W_h q_h m_{hx})}{(\sum_{h=1}^{H} W_h q_h m_{hx}^2)(\sum_{h=1}^{H} W_h q_h) - (\sum_{h=1}^{H} W_h q_h m_{hx})^2} (M_x - \sum_{h=1}^{H} W_h m_{hx}) \tag{34}$$

Substituting (34) into (30) gives:

$$\hat{M}_{YRC2} = \sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}} + \frac{(\sum_{h=1}^{H} \frac{W_h q_h m_{hx} m_{hy} M_{hx}}{m_{hx}})(\sum_{h=1}^{H} W_h q_h) - (\sum_{h=1}^{H} \frac{W_h q_h m_{hy} M_{hx}}{m_{hx}})(\sum_{h=1}^{H} W_h q_h m_{hx})}{(\sum_{h=1}^{H} W_h q_h m_{hx}^2)(\sum_{h=1}^{H} W_h q_h) - (\sum_{h=1}^{H} W_h q_h m_{hx})^2} \tag{35}$$

$$\left(M_x - \sum_{h=1}^{H} W_h m_{hx} \right)$$

Equation (35) can be written in the form of the generalized regression (GREG) estimator as:

$$\hat{M}_{YRC2} = \hat{M}_{YR} + \hat{\beta}_{st.2} (M_x - \sum_{h=1}^{H} W_h m_{hx}) . \tag{36}$$

Where

$$\hat{\beta}_{st.2} = \frac{(\sum_{h=1}^{H} \frac{W_h q_h m_{hx} m_{hy} M_{hx}}{m_{hx}})(\sum_{h=1}^{H} W_h q_h) - (\sum_{h=1}^{H} \frac{W_h q_h m_{hy} M_{hx}}{m_{hx}})(\sum_{h=1}^{H} W_h q_h m_{hx})}{(\sum_{h=1}^{H} W_h q_h m_{hx}^2)(\sum_{h=1}^{H} W_h q_h) - (\sum_{h=1}^{H} W_h q_h m_{hx})^2}$$

and $\hat{M}_{YR} = \sum_{h=1}^{H} \frac{W_h m_{hy} m_{hx}}{m_{hx}}$.

Substituting $q_h = 1$ in (35) gives the proposed regression-type calibrated estimator of the population median in stratified random sampling under two constraints as.

$$\hat{M}_{YRC21} = \sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}} + \frac{(\sum_{h=1}^{H} \frac{W_h m_{hx} m_{hy} M_{hx}}{m_{hx}})(\sum_{h=1}^{H} W_h) - \sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}} (\sum_{h=1}^{H} W_h m_{hx})}{(\sum_{h=1}^{H} W_h m_{hx}^2)(\sum_{h=1}^{H} W_h) - (\sum_{h=1}^{H} W_h m_{hx})^2} \left(M_x - \sum_{h=1}^{H} W_h m_{hx} \right) \tag{37}$$

Also by letting $q_h = \frac{1}{m_{hx}}$ in (35), we have the proposed ratio-type calibrated estimator of the population median in stratified random sampling under two constraints as:

$$\hat{M}_{YRC22} = \sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}} + \frac{(\sum_{h=1}^{H} \frac{W_h m_{hy} M_{hx}}{m_{hx}})(\sum_{h=1}^{H} \frac{W_h}{m_{hx}}) - (\sum_{h=1}^{H} W_h \frac{m_{hy}}{m_{hx}^2} M_{hx})(\sum_{h=1}^{H} W_h)}{(\sum_{h=1}^{H} W_h m_{hx})(\sum_{h=1}^{H} \frac{W_h}{m_{hx}}) - (\sum_{h=1}^{H} W_h)^2} \left(M_x - \sum_{h=1}^{H} W_h m_{hx} \right) \tag{38}$$

4.2 Bias and Variance Estimators of the proposed Calibrated Estimators Under Two Constraints

The bias and variance estimators of the proposed calibrated ratio median estimators under two constraints, are also obtained using same procedure of Taylor Linearization Technique as used under one constrain . Only the calibrated weight is changed We have:

$$Bias(\hat{M}_{YRC2}) = \sum_{h=1}^{H} W_{h2}^{*} M_{hy} \lambda_h (C_{M_{hx}}^2 - C_{M_{hy}} C_{M_{hx}} \rho_{M_{hx} M_{hy}}) \tag{37}$$

and

$$\hat{V}_T(\hat{M}_{YRC2}) = \sum_{h=1}^{H} (W_{h2}^{*})^2 \lambda_h \left((\hat{F}_{hy}(M_{hy}))^{-2} + \hat{R}_m^2 (F_{hx}(M_{hx}))^{-2} - 2\hat{R}_m (\hat{F}_{hx}(M_{hx}) \hat{F}_{hy}(M_{hy}))^{-1} \hat{\rho}_{M_{hx}M_{hy}} \right) \tag{38}$$

Where

$$W_{h2}^{*} = W_h + \frac{W_h q_h m_{hx} (\sum_{h=1}^{H} W_h q_h) - W_h q_h (\sum_{h=1}^{H} W_h q_h m_{hx})}{(\sum_{h=1}^{H} W_h q_h m_{hx}^2)(\sum_{h=1}^{H} W_h q_h) - (\sum_{h=1}^{H} W_h q_h m_{hx})^2} (M_x - \sum_{h=1}^{H} W_h m_{hx}) \tag{39}$$

is the calibrated weight for two constraints as stated in (34)

Setting $q_h = 1$, we obtain:

$$W_{h21}^{*} = W_h + \frac{W_h m_{hx} (\sum_{h=1}^{H} W_h) - W_h (\sum_{h=1}^{H} W_h m_{hx})}{(\sum_{h=1}^{H} W_h m_{hx}^2)(\sum_{h=1}^{H} W_h) - (\sum_{h=1}^{H} W_h m_{hx})^2} (M_x - \sum_{h=1}^{H} W_h m_{hx}) \tag{40}$$

Also setting $q_h = \frac{1}{m_{hx}}$, we obtain:

$$W_{h22}^{*} = W_h + \frac{(W_h)(\sum_{h=1}^{H} \frac{W_h}{m_{hx}}) - \frac{W_h}{m_{hx}}(\sum_{h=1}^{H} W_h)}{(\sum_{h=1}^{H} W_h m_{hx})(\sum_{h=1}^{H} \frac{W_h}{m_{hx}}) - (\sum_{h=1}^{H} W_h)^2}(M_x - \sum_{h=1}^{H} W_h m_{hx}) \tag{41}$$

Substituting (40) and 41) into (38), gives the regression and the ratio-type variance estimators under two constraints as:

$$\hat{V}_T(\hat{M}_{YRC21}) = \sum_{h=1}^{H}\left(W_h + \frac{(W_h m_{hx})(\sum_{h=1}^{H} W_h) - (W_h)(\sum_{h=1}^{H} W_h m_{hx})}{(\sum_{h=1}^{H} W_h m_{hx}^2)(\sum_{h=1}^{H} W_h) - (\sum_{h=1}^{H} W_h m_{hx})^2}(M_x - \sum_{h=1}^{H} W_h m_{hx})\right)^2 \lambda_h\Big((F_{hy}(M_{hy}))^{-2} + R_m^2(F_{hx}(M_{hx}))^{-2}$$

$$- 2\hat{R}_m(\hat{F}_{hx}(M_{hx})\hat{F}_{hy}(M_{hy}))^{-1}\hat{\rho}_{M_{hx}M_{hy}}\Big) \tag{42}$$

and

$$\hat{V}_T(\hat{M}_{YRC22}) = \sum_{h=1}^{H}\left(W_h + \frac{(W_h)(\sum_{h=1}^{H} \frac{W_h}{m_{hx}}) - (\frac{W_h}{m_{hx}})(\sum_{h=1}^{H} W_h)}{(\sum_{h=1}^{H} W_h m_{hx})(\sum_{h=1}^{H} \frac{W_h}{m_{hx}}) - (\sum_{h=1}^{H} W_h)^2}(M_x - \sum_{h=1}^{H} W_h m_{hx})\right)^2 \lambda_h\Big((F_{hy}(M_{hy}))^{-2} + R_m^2(F_{hx}(M_{hx}))^{-2}$$

$$- 2\hat{R}_m(\hat{F}_{hx}(M_{hx})\hat{F}_{hy}(M_{hy}))^{-1}\hat{\rho}_{M_{hx}M_{hy}}\Big) \tag{43}$$

5.1 Empirical Evaluation of Estimators

The performance of the proposed calibrated Ratio estimators shall be compared with the existing ratio estimator using two performance measures namely: Percentage Relative Efficiency and the Percentage Absolute Relative Bias.

By the percentage relative efficiency of two estimators, we mean the percentage ratio of their variances or mean square errors. It can be computed as:

$$\%RE(\hat{M}_{YRC}) = \left[\frac{MSE(\hat{M}_{YR})}{MSE(\hat{M}_{YRC})}\right] \times 100 \tag{44}$$

Also, given the calibration estimator $\hat{M}_{YRC}$, the percentage absolute relative bias with respect to the population median M_y is given as:

$$\%ARB(\hat{M}_{YRC}) = \left| \frac{1}{R}\sum_{r=1}^{R}\left(\frac{\hat{M}_{YRC}}{M_y} - 1\right) \right| \text{x } 100 \quad (45)$$

5.2 Simulation study

A simulation study using R software was done on the estimators for different sample sizes, with distributional assumptions of Cauchy and lognormal distributions.

For each sample of size $(10\%), (20\%)$ and (25%), selected in each simulation run r=1,2,3…R , (R=10,000), the estimates of $\hat{M}_{YRC}$ and $\hat{M}_{YR}$ was computed.

Suppose $\hat{M}^{r}_{YRC}$ and $\hat{M}^{r}_{YR}$ denotes the proposed and existing estimator respectively for the r^{th} run, r=1,2…R, then the mean square error for both estimators was also computed as

$$MSE(\hat{M}_{YRC}) = \frac{1}{R}\sum_{r=1}^{R}(\hat{M}^{r}_{YRC} - M_y)^2 \quad (46)$$

$$MSE(\hat{M}_{YR}) = \frac{1}{R}\sum_{r=1}^{R}(\hat{M}^{r}_{YR} - M_y)^2 \quad (47)$$

5.3 Discussion of Results

Tables 1 and 2 below, shows the percentage absolute relative bias (%ARB) and the percentage relative efficiency (%RE) with one and two constraints for Cauchy distribution, and lognormal distributions using sample sizes of 10%, 20% and 25% respectively.

From Table 1, under the Cauchy distribution, the ratio-type estimator with one constraint $\hat{M}_{YRC12}$ has the smallest %ARB of 95.6 to 102.0 as sample size increases from 10% to 25%, followed by the regression-type estimator with two constraints $\hat{M}_{YRC21}$ with a fluctuated %ARB of 5045.7 at 10% sample size with a decrease of 1263.8 to 1109.1 as sample size increased from 20% to 25% and the regression-type estimator with one constraint $\hat{M}_{YRC11}$ having a %ARB of 1169.5 to 2435.0 as sample size increased from 10% to 25% whereas, the existing estimator $\hat{M}_{YR}$ and the proposed ratio-type estimator with two constraints $\hat{M}_{YRC22}$ have very high %ARB of 1538.5 to 3277.7 and 5401.6 to 10706.9 respectively as sample size increased from 10% to 25%.

Under the lognormal distribution, the regression-type estimator with two constraints $\hat{M}_{YRC21}$ has the smallest %ARB of 1358.1 to 874.4 which decreased as sample size increased from 10% to 25% followed by the ratio and the regression-type estimators with one constraint $\hat{M}_{YRC12}$ and $\hat{M}_{YRC11}$ having %ARB of 1444.5 to 896.9 and 1713.4 to 1030.3 respectively which decreased as sample sizes increased from 10% to 25%. The existing Ratio estimator $\hat{M}_{YR}$ and the proposed ratio- type estimator with two constraints $\hat{M}_{YRC22}$ still having high %ARB of 1724.0 to 1149.6 and 12502.6 to 1457.2 respectively as sample size increased.

Table 2, under the Cauchy distribution, shows that the proposed regression and ratio-type estimators with one constraint $\hat{M}_{YRC11}$ and $\hat{M}_{YRC12}$, and the regression-type estimator with two constraints $\hat{M}_{YRC21}$ have high gains in efficiency of 173.0 to 181.2, 24605.9 to 103343.0 and 579.2 to 872.9 respectively as the sample size increased (with an initial loss of 9.3% at 10% sample size for the regression-type estimator with two constraints $\hat{M}_{YRC21}$) than the existing estimator $\hat{M}_{YR}$ whereas, the ratio -type estimator with two constraints $\hat{M}_{YRC22}$ has a loss ranging from 8.1 to 9.9 as sample size increased from 10% to 25%.

Under the lognormal distribution, the regression and the ratio-type estimators with one constraint $\hat{M}_{YRC11}$ and $\hat{M}_{YRC12}$ and the regression-type estimator with two constraints $\hat{M}_{YRC21}$ gained efficiencies ranging from 101.2 to 124.5, 142.4 to 164.3 and 161.2 to 174.3 respectively as sample increased from 10% to 25% than the existing estimator $\hat{M}_{YR}$, while the proposed ratio-type estimator with two constraints $\hat{M}_{YRC22}$ has a loss of 1.9 to 62.2 as sample size increased from 10% to 25%

TABLE 1

Percentage Absolute Relative Bias for Cauchy and Lognormal Distributions Under one and Two Constraints

		Existing Estimator	One Constraint		Two Constraints	
Sample Sizes	Distribution	$\hat{M}_{YR}$	$\hat{M}_{YRC11}$	$\hat{M}_{YRC12}$	$\hat{M}_{YRC21}$	$\hat{M}_{YRC22}$
10%	Cauchy	1538.5	1169.5	98.1	5045.7	5401.6
	Lognormal	1724.0	1713.4	1444.5	1358.1	12502.6

20%	Cauchy	3041.6	2351.1	95.6	1263.8	10706.9
	Lognormal	1222.3	1118.2	957.1	925.7	4603.8
25%	Cauchy	3277.7	2435.0	102.0	1109.1	10408.0
	Lognormal	1149.6	1030.3	896.9	874.4	1457.2

TABLE 2

Percentage Relative Efficiency for Cauchy, and Lognormal Distributions Under One and Two Constraints

		Existing Estimator	One Constraint		Two Constraints	
Sample Sizes	Distributions	$\hat{M}_{YR}$	$\hat{M}_{YRC11}$	$\hat{M}_{YRC12}$	$\hat{M}_{YRC21}$	$\hat{M}_{YRC22}$
10%	Cauchy	100	173.0	24605.9	9.3	8.1
	Lognormal	100	101.2	142.4	161.2	1.9
20%	Cauchy	100	167.4	101168.6	579.2	8.1
	Lognormal	100	119.5	163.0	174.3	7.0
25%	Cauchy	100	181.2	103343.0	872.9	9.9
	Lognormal	100	124.5	164.3	172.9	62.2

6. Conclusion

In this paper, four ratio median estimators (the regression-type estimator under one constraint $\hat{M}_{YRC11}$, the ratio-type estimator under one constraint $\hat{M}_{YRC12}$, the regression-type estimator under two constraints $\hat{M}_{YRC21}$ and the ratio-type estimator under two constraints $\hat{M}_{YRC22}$) were introduced using calibration for both single and double constraints under stratified random sampling. Also, their biases and variance estimators were obtained using Taylor Linearization.

Following the discussion of results above, we deduce that under the Cauchy and lognormal distributions which are strictly heavy tailed, the proposed regression and ratio-type estimators with one constraint $\hat{M}_{YRC11}$, $\hat{M}_{YRC12}$ and the regression-type estimator with two constraints $\hat{M}_{YRC21}$ are more efficient than the existing ratio estimator $\hat{M}_{YR}$ and the proposed ratio-type estimator under two constraints $\hat{M}_{YRC22}$. This work obviously shows that though one of the proposed estimators under two constraints (the ratio-type estimator under two constraints $\hat{M}_{YRC22}$) is bad, the other three proposed estimators work better for real life data that follow a skewed or heavy tailed distributions like the Cauchy and lognormal distributions. This agrees with the already known fact that robust statistics have good performance for data drawn from distributions that are skewed and generally supports the results of Deville and Sarndal (1992) that there is gain in efficiency for calibrations.

Garg, N. & Pachori, M. (2019). Use of Median in Calibration Estimation of the Finite Population Mean in Stratified Sampling. *International Journal of Computer Sciences and Engineering, 7(*5):2347-2693

Mouhamed, A.M., El-sheik, A. & Mouhamed, H.A. (2015). A New Calibration Estimator of Stratified Random Sampling. *Applied Mathematical Sciences,* 35(9):1735-1744

References

[1] Aladag, S.& Cingi, H. (2015). Improvement in Estimating the population Median in Simple Random Sampling and Stratified Random Sampling using Auxiliary Information. Communications in Statistics-Theory and Methods, 44(5): 1013-1032

[2] Aamir, M, Shabri, A. & Ishaq, M (2018). Improvement on Estimating Median for Finite Population using Auxiliary Variables in Double- Sampling. Journal Teknologi, 80(5),135-143

[3] Baig, A., Masood, S. & Tarray, T.A (2019). Improved Class of Difference-type Estimators for Population median in Survey Sampling. Communication in Statistics-Theory and Methods, 47(24), 1-16

[4] Clement, E.P. & Enang, I.E (2015). Calibration Approach Alternative Ratio Estimator for Population Mean in Stratified Random Sampling. International Journal of Statistics and Economics, 16(1), 0973-7022

[5] Clement, E.P & Enang, E.I (2017). On Efficiency of Ratio Estimators over Regression Estimators. Communication in Statistics-Theory and Methods, 46(1):5357_, 5367

[6] Deville, J.C Sarndal, C.E. (1992). Calibration Estimators in Survey Sampling. Journal of the American Statistical Association, 87(418), 376-382

[7] Enang, E., Etuk, S., Ekpenyong, E.T. & Akpan, V.M. (2016). An alternative Exponential Estimator of Population Median. International Journal of Statistics and Economics, 17(3), 0973-7022

[8] Garcia, M.R. & Cebrian, A.A. (2001). On Estimating the Median from Survey Data using Multiple Auxiliary information. Metrika, 54, 59-76

[9] Garg, N. & Pachori, M. (2019). Use of Median in Calibration Estimation of the Finite Population Mean in Stratified Sampling. International Journal of Computer Sciences and Engineering, 7(5), 2347-2693

[10] Gross, S.T. (1980). Median Estimation in Sample Surveys. Preceedings of the Survey Research Methods Section, American Statistical Association, 1814184, 181-184

[11] Heldal, J. (1992). A Method for Calibration of Weights in Survey Sampling. Central Bureau of Statistics Norway, 3, 1-18

[12] Jhaji, H.S. & Bhangu, H.K. (2013). Generalized Estimators of Population Median Using Auxiliary Information. International Journal of Science and Engineering Research, 4(11), 229-551

[13] Kayuncu, N.& Kadilar, C. (2013). Calibration Estimator using different Distance Measure in Stratified Random Sampling. International Journal of Modern Engineering Research, 3(1), 415-419

[14] Kayuncu, N.& Kadilar, C. (2017). Calibration Weighting in Stratified Random sampling. Communication in Statistics-Simulation and Computation. https://dx.doi.org/10.1080/03610918.2014.901354

[15] Kim, J.K & Park, M. (2009). Calibration Estimation in Survey Sampling. International Statistical Institute, 78 (1), 21-39

[16] Kuk, A.Y.C. & Mak, T.K. (1989). Median Estimation in the presence of Auxiliary Information. Journal of the Royal Statistical Society B, 51, 261-269

[17] Meeden, G. (1995). Median Estimation using Auxiliary Information. Survey Methodology, 21, 71-77

[18] Price, R.M, & Bonett, D., (2014). Estimating the Variance of the Sample Median. Journal of Statistical Computation and Simulation, 68(3), 295-305

[19] Sharma, P.& Singh, R. (2015). Generalized Class of Estimators for Population Median using Auxiliary Information. Hacettepe Journal of Mathematics and Statistics, 44(22), 443-453

[20] Singh, S., Joarder, A.H. & Tracy, D.S (2001). Median Estimation using Double Sampling. Australian & New Zealand Journal of Statistics, 43, 33-46

[21] Singh, S.S., Singh, H.P & Upadhyaya, L.N. (2006). Chain Ratio and Regression-type Estimators for Median Estimation Survey Sampling. Statistical Papers, 48, 23-46

[22] Singh, H.P, & Solanki, R.S (2013). Some Class of Estimators for Population Median using Auxiliary information. Communications in Statistics, 42(23), 4222-4238

[23] Singh, S. & Arnab, R. (2014). On Calibration of Designs Weight. Metron-International Journal of Statistics, 69(2),185-205

[24] Solanki, R.S. & Singh, H.P. (2015). Some Classes of Estimators for Median Estimation in Survey Sampling. Communications in Statistics-Theory & Methods, 44(7), 1450-1465