International Journal of Engineering Sciences & Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164

IJESRT

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

ENHANCING EFFICIENCY OF RATIO ESTIMATOR OF POPULATION MEDIAN BY CALIBRATION TECHNIQUES

Iseh, Matthew Joshua, Ph.D.
Department of Statistics, Akwa Ibom State University, Mkpat Enin, Nigeria

DOI: https://doi.org/10.29121/ijesrt.v9.i8.2020.2

ABSTRACT

The use of calibration estimation techniques in survey sampling have been found to improve the precision of estimators. This paper adopts the calibration approach with the assumption that the population median of the auxiliary variable is known to obtain a more efficient ratio-type estimator in estimating population median in stratified sampling. Conditions necessary for efficiency comparison have been obtained which show that the proposed estimator will always perform better than the existing asymptotically unbiased separate estimators in stratified random sampling. Numerical evaluations have been carried out through simulation and real-life data to compliment the theoretical claims. Results from the simulation study carried out under three distributional assumptions, namely the chi square, lognormal and Cauchy distributions with different sample settings showed that the new estimator provided better estimate of the median with greater gain in efficiency. In addition, result from the real-life data further supports the superiority of the proposed estimator over the existing ones considered in this study.

KEYWORDS: Calibration estimation; Mean square error; Median estimation; Ratio-type estimator; Stratified random sampling.

1. INTRODUCTION

Most times in survey sampling, some researchers do not take into consideration the tool that will be most appropriate in the measure of location. As a result, users of Statistics often go for the mean or total which has been widely discussed in the finite population sampling literature unlike the median which is more complicated to deal with since it has to do with ordered data. However, it has been established that the median unlike the mean performs better when the considered variables are from a highly skewed distribution. In surveys involving the estimation of income, expenditure, scores, etc., it is very reasonable to assume that the population median unlike the population mean is known, hence the possibility of incorporating auxiliary information through formulation of calibration constraints.

Deville and Sarndal (1992) first proposed the calibration estimation technique in survey sampling in order to incorporate the auxiliary variable into an existing estimator. Several authors have made useful contributions to improve the precision of survey estimates of population parameters using the calibration approach. Notable among them are Sarndal (2007), Arnab and Singh (2014), Clement and Enang (2017), Clement (2015, 2017), Koyuncu and Kadilar (2013). Although authors like Gross (1980), Kuk and Mak (1989), H. P .Singh, S. Singh and Puertas (2003), Singh and Solanki (2013), Aladag and Cingi (2015), have made useful contributions in estimation of population median, it should also be noted that not much has been done using calibration technique in estimating population median. However, Garg and Pachori (2019) made use of the median in calibration estimation of the finite population mean in stratified sampling.

Aladag and Cingi (2015) made useful suggestions on improving estimations of population medians in simple random sampling and stratified sampling using auxiliary information. As evident in both the theoretical derivations and numerical examples, the authors proposed some asymptotically unbiased estimators of population median in stratified sampling that outperformed other existing and suggested estimators in terms of gain in efficiency. Based on this development as a benchmark, a new estimator of the population median is sought after for higher precision. This study seeks to extend the Vishwakarma and Singh (2011) separate ratio-product

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

estimator in stratified random sampling to population median and a further improvement of the estimator is made using the calibration technique.

2. NOTATIONS AND REVIEW OF EXISTING LITERATURE

Notations

Consider a finite population $U = \{u_1, u_2, \ldots, u_N\}$ with size N. Let Y and X be the study and auxiliary variables respectively. Let y_{hi} be the characteristic of interest and x_{hi} be the auxiliary variable known for every unit in the population for the i^{th} element in the h^{th} stratum respectively and are non-negative. Suppose the population size N is stratified into H strata with h^{th} stratum containing N_h units, where $h = 1,2, \ldots, H$ such that $\sum_{h=1}^{H} N_h = N$ and stratum weight given as $W_h = \frac{N_h}{N}$. A simple random sample of size n_h is drawn without replacement from the h^{th} stratum such that $\sum_{h=1}^{H} n_h = n$. Let M_{Yh} and M_{Xh} be the population median for the study and auxiliary variables respectively in the h^{th} stratum, m_{yh} and m_{xh} are the respective sample medians. Suppose $y_{h(1)}, y_{h(2)}, \ldots, y_{h(n)}$, are the y_{hi} values of the sample units in ascending order. Also, suppose r be the integer satisfying $Y_{hr} \leq M_{hY} \leq M_{h(r+1)}$ and $P_h = \frac{r}{n}$ be the proportion of y_{hi} values in the sample that are less than or equal to the median value M_{hY} which denotes the unknown population parameter. If $\varphi_y(r)$ denote the $r - quantile$ of Y_h then, $m_{yh} = \varphi_y(0.5)$. Kuk and Mak (1989) defined a matrix of proportion P_{ij} as shown in Table 1.

Also, $\rho_{M_{Yh}M_{Xh}}$ the correlation coefficient in the h^{th} stratum between the sampling distributions of M_{Yh} and M_{Xh} is defined as $\rho_{M_{Yh}M_{Xh}} = 4(P_{11h} - 0.25)$, where P_{11h} is the proportion of units in the population in the h^{th} stratum with $Y_h \leq M_{Yh}$ and $X_h \leq M_{Xh}$.

To obtain the large sample properties for the suggested separate estimators, the followings are obtainable as the relative errors up to first order of approximation and their expectations;

$m_{yh} = M_{Yh}(1 + e_{0h}), \quad m_{xh} = M_{Xh}(1 + e_{1h}) \quad , \quad R_h = \frac{m_{xh}}{M_{Xh}}$ and $R_h^{-1} = \frac{M_{Xh}}{m_{xh}}$

$e_{0h} = \frac{m_{yh} - M_{Yh}}{M_{Yh}}, \quad e_{1h} = \frac{m_{xh} - M_{Xh}}{M_{Xh}}, \quad E(e_{0h}) = E(e_{1h}) = 0$

$E(e_{0h}^2) = \gamma_h C_{M_{Yh}}^2, \quad E(e_{1h}^2) = \gamma_h C_{M_{Xh}}^2, \quad E(e_{0h}e_{1h}) = \gamma_h C_{M_{Yh}} C_{M_{Xh}} \rho_{M_{Yh}M_{Xh}}$

$\gamma_h = \frac{1 - f_h}{4n_h}, \quad C_{M_{Yh}} = \{M_{Yh} f_Y(M_{Yh})\}^{-1}, \; C_{M_{Xh}} = \{M_{Xh} f_X(M_{Xh})\}^{-1}$

where, it is also assumed that the distribution function $f_Y(M_{Yh})$ and $f_X(M_{Xh})$ are nonnegative.

Table 1. Matrix of proportion for stratified sampling

	$X_h \leq M_{Xh}$	$X_h > M_{Xh}$	Total
$Y_h \leq M_{Yh}$	P_{11}	P_{12}	P_{1}
$Y_h > M_{Yh}$	P_{21}	P_{22}	P_{2}
Total	$P_{.1}$	$P_{.2}$	1

Some Existing Estimators

This section considers some ratio-type estimators in stratified sampling in estimating population median as suggested by Aladag and Cingi (2015) as follows;

Motivated by Kuk and Mak (1989), Aladag and Cingi (2015) suggested a separate ratio estimator for population median in stratified sampling as

$$\widehat{M}_{YRS} = \sum_{h=1}^{H} W_h m_{yh} \frac{M_{Xh}}{m_{xh}} \quad \text{(i)}$$

with bias and mean square error (MSE) as follows:

$$Bias(\widehat{M}_{YRS}) \cong \sum_{h=1}^{H} W_h M_{Yh} \gamma_h \left[C_{M_{Xh}}^2 - C_{M_{Yh}} C_{M_{Xh}} \rho_{M_{Yh}M_{Xh}}\right] \quad \text{(ii)}$$

$$MSE(\widehat{M}_{YRS}) = \sum_{h=1}^{H} W_h^2 M_{Yh}^2 \gamma_h \left[C_{M_{Yh}}^2 + C_{M_{Xh}}^2 - 2C_{M_{Yh}} C_{M_{Xh}} \rho_{M_{Yh}M_{Xh}}\right] \quad \text{(iii)}$$

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Also, following Singh, et al. (2003), Aladag and Cingi (2015) gave the median version of the difference estimator in stratified sampling as

$$\widehat{M}_{YDS} = \sum_{h=1}^{H} W_h\left[m_{yh} + d_{MSh}(M_{Xh} - m_{xh})\right] \qquad \text{(iv)}$$

The bias and MSE was obtained as

$$Bias(\widehat{M}_{YDS}) \cong 0 \qquad \text{(v)}$$

$$MSE(\widehat{M}_{YDS}) = \sum_{h=1}^{H} W_h^2\,\gamma_h\left\{[f_Y(M_{Yh})]^{-2} + d_{MSh}^2[f_X(M_{Xh})]^{-2} - 2d_{MSh}[f_Y(M_{Yh})f_X(M_{Xh})]^{-1}\rho_{M_{Yh}M_{Xh}}\right\} \qquad \text{(vi)}$$

The minimum variance of the unbiased estimator was further obtained by minimizing for d_{MSh} as

$$V_{Min}(\widehat{M}_{YDS}) = \sum_{h=1}^{H} W_h M_{Yh}^2\,\gamma_h C_{M_{Yh}}^2(1 - \rho_{M_{Yh}M_{Xh}}{}^2) \qquad \text{(vii)}$$

where

$$d_{MSh} = \frac{f_X(M_{Xh})}{f_Y(M_{Yh})}\rho_{M_{Yh}M_{Xh}}$$

Aladag and Cingi (2015) also proposed another version of the separate estimator as

$$\widehat{M}_{iS}^{(\alpha)} = \sum_{h=1}^{H} W_h m_{yh}\left(\frac{\psi_{ih}M_{Xh}+\zeta_{ih}}{\psi_{ih}m_{xh}+\zeta_{ih}}\right)^{\alpha} \qquad \text{(viii)}$$

For $i = 1,2,\ldots,9,$ where ψ_{ih} and ζ_{ih} are constants that take values $1, \rho_{M_YM_X}, M_0$ *for mode, and* R_X *for range*, depending on researcher's choice. Different combinations of these constants are detailed in Table 2 of Aladag and Cingi (2015).

The estimator has minimum bias and minimum MSE as follows:

$$Bias_{Min}(\widehat{M}_{iS}^{(\alpha)}) \cong \sum_{h=1}^{H} W_h M_{Yh}\gamma_h \frac{C_{M_{Yh}}\rho_{M_{Yh}M_{Xh}}}{2}\left[\delta_{ih}^{*}C_{M_{Xh}} - C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}}\right] \qquad \text{(ix)}$$

$$MSE_{Min}(\widehat{M}_{iS}^{(\alpha)}) \cong \sum_{h=1}^{H} W_h M_{Yh}^2\,\gamma_h C_{M_{Yh}}^2(1 - \rho_{M_{Yh}M_{Xh}}{}^2) \qquad \text{(x)}$$

Where $\alpha = \frac{C_{M_{Yh}}\rho_{M_{Yh}M_{Xh}}}{C_{M_{Xh}}\delta_{ih}^{*}}$ and $\delta_{ih}^{*} = \frac{\psi_{ih}M_{Xh}}{\psi_{ih}M_{Xh}+\zeta_{ih}}$

Following Kadilar and Cingi (2004), the authors further suggested the median estimator in stratified sampling of the form

$$\widehat{M}_{Pis} = \sum_{h=1}^{H} W_h\left[m_{yh} - \alpha(\psi_{ih}M_{Xh} - \psi_{ih}m_{xh})\right]\left(\frac{\psi_{ih}M_{Xh}+\zeta_{ih}}{\psi_{ih}m_{xh}+\zeta_{ih}}\right) \qquad \text{(xi)}$$

with minimum bias and minimum MSE as

$$Bias(\widehat{M}_{Pis}) \cong 0 \qquad \text{(xii)}$$

$$MSE_{Min}(\widehat{M}_{Pis}) \cong \sum_{h=1}^{H} W_h M_{Yh}^2\,\gamma_h C_{M_{Yh}}^2(1 - \rho_{M_{Yh}M_{Xh}}{}^2) \qquad \text{(xiii)}$$

where $\alpha = \frac{M_{Yh}}{M_{Xh}}\left(\frac{C_{M_{Yh}}\rho_{M_{Yh}M_{Xh}}}{C_{M_{Xh}}\psi_{ih}} - \frac{\delta_{ih}^{*}}{\psi_{ih}}\right)$

3. THE PROPOSED ESTIMATOR

Vishwakarma and Singh (2011) proposed separate ratio-product estimator for population mean in stratified sampling as

$$\bar{y}_{st}(\alpha) = \sum_{h=1}^{H} W_h\bar{y}_h\left[\alpha\frac{\bar{X}_h}{\bar{x}_h} + (1-\alpha)\frac{\bar{x}_h}{\bar{X}_h}\right] \qquad \text{(xiv)}$$

Consequently, an extension of equation (xiv) to the estimation of population median in stratified random sampling given that the population median of the auxiliary variable X is known can be obtained as:

$$\hat{M}_{st}(\alpha_h) = \sum_{h=1}^{H} W_h m_{yh}[\alpha_h R_h^{-1} + (1-\alpha_h)R_h] \quad \text{(xv)}$$

$$\hat{M}_{st}(\alpha_h) = \sum_{h=1}^{H} W_h M_{Yh}(1+e_{0h})[\alpha_h R_h^{-1} + (1-\alpha_h)R_h]$$

with the assumption that $|e_{1h}| < 1$, $(1+e_{1h})^{-1}$ is considered as a second order Taylor series expansion and neglecting the higher order terms, gives

$$\hat{M}_{st}(\alpha_h) = \sum_{h=1}^{H} W_h M_{Yh}[1 + e_{0h} + e_{1h} + e_{1h}e_{0h} + \alpha_h(e_{1h}^2 - 2e_{1h} - 2e_{1h}e_{0h})]$$

$$\hat{M}_{st}(\alpha_h) - M = \sum_{h=1}^{H} W_h M_{Yh}[(e_{0h} + e_{1h} + e_{1h}e_{0h}) + \alpha_h(e_{1h}^2 - 2e_{1h} - 2e_{1h}e_{0h})] \quad \text{(xvi)}$$

Taking expectation of both sides of (xvi) and using the results in section 2.1, we have

$$E[\hat{M}_{st}(\alpha_h) - M] = E[\sum_{h=1}^{H} W_h M_{Yh}[(e_{0h} + e_{1h} + e_{1h}e_{0h}) + \alpha_h(e_{1h}^2 - 2e_{1h} - 2e_{1h}e_{0h})]]$$

$$Bias\left(\hat{M}_{st}(\alpha_h)\right) = E[\hat{M}_{st}(\alpha_h) - M]$$

$$Bias\left(\hat{M}_{st}(\alpha_h)\right) = \sum_{h=1}^{H} W_h M_{Yh}\gamma_h\left[(\alpha_h C_{MXh}^2 + (1-2\alpha_h)C_{M_{Yh}}C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}}\right] \quad \text{(xvii)}$$

Squaring both sides of (xvi) and retaining terms to the second degree, we have

$$[\hat{M}_{st}(\alpha_h) - M]^2 = [\sum_{h=1}^{H} W_h M_{Yh}[(e_{0h} + e_{1h} + e_{1h}e_{0h}) + \alpha_h(e_{1h}^2 - 2e_{1h} - 2e_{1h}e_{0h})]]^2$$

$$[\hat{M}_{st}(\alpha_h) - M]^2 = \sum_{h=1}^{H} W_h^2 M_{Yh}^2[e_{0h}^2 + e_{1h}^2 + 4\alpha_h^2 e_{1h}^2 - 4\alpha_h e_{1h}^2 + 2e_{1h}e_{0h} - 4\alpha_h e_{1h}e_{0h} + \cdots] + \sum_{h=1}^{H}\sum_{h^1\neq 1}^{H^1} W_h W_{h^1} M_{Yh} M_{Yh^1}[e_{0h}e_{0h^1} + e_{1h}e_{1h^1} + 4\alpha_h^2 e_{1h}e_{1h^1} - 4\alpha_h e_{1h}e_{1h^1} + \cdots] \quad \text{(xviii)}$$

Taking expectations of both sides of (xviii), and applying the results earlier given in section 2.1, we obtain the MSE of $\hat{M}_{st}(\alpha_h)$ to the first order of approximation as:

$$MSE\left(\hat{M}_{st}(\alpha_h)\right) = \sum_{h=1}^{H} W_h^2 M_{Yh}^2 \gamma_h[C_{M_{Yh}}^2 + C_{M_{Xh}}^2(1-2\alpha_h)^2 + 2C_{M_{Yh}}C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}}(1-2\alpha_h)] \quad \text{(xix)}$$

Minimizing the MSE of $\hat{M}_{st}(\alpha_h)$ with respect to α_h, we have

$$\alpha_h = \frac{1+K_h}{2} \quad \text{(xx)}$$

where

$$k_h = \frac{C_{M_{Yh}}\rho_{M_{Yh}M_{Xh}}}{C_{M_{Xh}}}$$

Substituting (xx) into (xix) gives

$$MSE\left(\hat{M}_{st}(\alpha_h)\right) = \sum_{h=1}^{H} W_h^2 M_{Yh}^2 \gamma_h[C_{M_{Yh}}^2(1-\rho_{M_{Yh}M_{Xh}}{}^2)] \quad \text{(xxi)}$$

which can also be written as

$$MSE\left(\hat{M}_{st}(\alpha_h)\right) = \sum_{h=1}^{H} W_h^2 \gamma_h\{f_Y(M_{Yh})\}^{-2}(1-\rho_{M_{Yh}M_{Xh}}{}^2) \quad \text{(xxii)}$$

An Improved Version of the Suggested Median Estimator

Equation (xv) can be written as

$$\hat{M}_{st}^*(\alpha_h) = \sum_{h=1}^{H} \Omega_h m_{yh}\,\beta \quad \text{(xxiii)}$$

where the coefficient β is given as

$$\beta = [\alpha_h R_h^{-1} + (1-\alpha_h)R_h]$$

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

And Ω_h is the calibration weights chosen such that the chi square distance measure

$$\sum_{h=1}^{H}\left(\frac{\Omega_h - W_h}{q_h W_h}\right)^2 \quad \text{(xxiv)}$$

is minimized subject to the constraints

$$\sum_{h=1}^{H}\Omega_h\{f_X(M_{Xh})\}^{-2} = V(m_{xs}) \quad \text{(xxv)}$$

Minimizing the distance measure in (xxiv) subject to the constraint in (xxv) gives the calibration weights as

$$\Omega_h = W_h + \left[V(m_{xs}) - \sum_{h=1}^{H} W_h\{f_X(M_{Xh})\}^{-2}\right]\frac{q_h W_h\{f_X(M_{Xh})\}^{-2}}{\sum_{h=1}^{H} q_h W_h\{f_X(M_{Xh})\}^{-4}} \quad \text{(xxvi)}$$

Squaring both sides of (xxvi) and assuming the tuning parameter $q_h = \{f_X(M_{Xh})\}^2$, then

$$\Omega_h^2 = W_h^2\left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2 \quad \text{(xxvii)}$$

where $V(m_{xs}) = \sum_{h=1}^{H} W_h^2\gamma_h\{f_X(M_{Xh})\}^{-2}$ and $\hat{V}(m_{xs}) = \sum_{h=1}^{H} W_h\{f_X(M_{Xh})\}^{-2}$
Substituting equation (xxvi) into (xxiii) gives the improved separate ratio-product estimator for population median as

$$\hat{M}_{st}^{*}(\alpha_h) = \sum_{h=1}^{H}\left[W_h + \left[V(m_{xs}) - \sum_{h=1}^{H} W_h\{f_X(M_{Xh})\}^{-2}\right]\frac{q_h W_h\{f_X(M_{Xh})\}^{-2}}{\sum_{h=1}^{H} q_h W_h\{f_X(M_{Xh})\}^{-4}}\right] m_{yh}\,\beta \quad \text{(xxviii)}$$

Variance Estimator
Suppose

$$\hat{M}_{st}^{*}(\alpha_h) - M_y = \sum_{h=1}^{H}\Omega_h m_{yh}\,\beta - M_y$$

Squaring both sides and taking expectation yields

$$E\left[\hat{M}_{st}^{*}(\alpha_h) - M_y\right]^2 = E\left[\sum_{h=1}^{H}\Omega_h m_{yh}\,\beta - M_y\right]^2$$
$$= E\left[\sum_{h=1}^{H}\Omega_h m_{yh}\,\beta\right]^2 - 2\,\beta M_y E\left(\sum_{h=1}^{H}\Omega_h m_{yh}\right) + M_y^2$$
$$= \left[Var\left(\sum_{h=1}^{H}\Omega_h m_{yh}\right)\beta\right] + \left[E\left(\sum_{h=1}^{H}\Omega_h m_{yh}\beta\right)\right]^2 - 2\,\beta M_y E\left(\sum_{h=1}^{H}\Omega_h m_{yh}\right) + M_y^2$$
$$= \beta^2 Var\left(\sum_{h=1}^{H}\Omega_h m_{yh}\right) + \beta^2\left(\sum_{h=1}^{H}\Omega_h M_{Yh}\right)^2 - 2\,\beta M_y\left(\sum_{h=1}^{H}\Omega_h M_{Yh}\right) + M_y^2$$
$$= \beta^2\sum_{h=1}^{H}\Omega_h^2\,Var(m_{yh}) + M_y^2(\beta - 1)^2$$
$$MSE\left(\hat{M}_{st}^{*}(\alpha_h)\right) = \beta^2\sum_{h=1}^{H}\Omega_h^2\,\gamma_h C_{M_{Yh}}^2 + M_y^2(\beta - 1)^2 \quad \text{(xxix)}$$

Substituting (xxvii) into (xxix), gives

$$MSE\left(\hat{M}_{st}^{*}(\alpha_h)\right) = \beta^2\left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2\sum_{h=1}^{H} W_h^2\,\gamma_h C_{M_{Yh}}^2 + M_y^2(\beta - 1)^2 \quad \text{(xxx)}$$

Minimizing equation (xxx) with respect to β gives

$$\hat{\beta} = \frac{M_y^2}{\left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2\sum_{h=1}^{H} W_h^2\gamma_h C_{M_{Yh}}^2 + M_y^2} \quad \text{(xxxi)}$$

Substituting (xxxi) into (xxx), yields

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

$$MSE\left(\hat{M}_{st}^{*}(\alpha_h)\right)=\frac{M_y^2\left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2\sum_{h=1}^{H}W_h^2\gamma_h C_{M_{Yh}}^2}{\left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2\sum_{h=1}^{H}W_h^2\gamma_h C_{M_{Yh}}^2+M_y^2} \quad \text{(xxxii)}$$

4. EFFICIENCY COMPARISON

Here, the efficiency comparison of the suggested estimator in equation (xv) is obtained by comparing its MSE at its minimum in equation (xxii) with that of the classical median estimator in stratified random sampling $V(m_{ys})=\sum_{h=1}^{H}W_h^2\gamma_h\{f_Y(M_{Yh})\}^{-2}$, that of the separate ratio estimator in equation (iii), the separate difference estimator in equation (vii) and that of the separate product estimator in equation (xiii). Furthermore, the efficiency conditions of the proposed calibration estimator in (xxviii) will be done by comparing the MSE in equation (xxxii) with that of equation (xxii).

Condition 1: Efficiency comparison between (xxii) and the classical median estimator

$MSE\left(\hat{M}_{st}(\alpha_h)\right) < V(m_{ys})$ if

$$\sum_{h=1}^{H}W_h^2\gamma_h\{f_Y(M_{Yh})\}^{-2}(1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h=1}^{H}W_h^2\gamma_h\{f_Y(M_{Yh})\}^{-2}$$

$\Rightarrow (1-\rho_{M_{Yh}M_{Xh}}{}^2) < 1$

$\Rightarrow \rho_{M_{Yh}M_{Xh}}{}^2 > 0$ is always satisfied.

Hence, $MSE\left(\hat{M}_{st}(\alpha_h)\right)$ is more efficient than $V(m_{ys})$

Condition 2: Efficiency comparison between (xxii) and (iii)

$MSE\left(\hat{M}_{st}(\alpha_h)\right) < MSE(\hat{M}_{YRS})$ if

$$\sum_{h=1}^{H}W_h^2\gamma_h M_{Yh}^2 C_{M_{Yh}}^2\left(1-\rho_{M_{Yh}M_{Xh}}{}^2\right) < \sum_{h=1}^{H}W_h^2 M_{Yh}^2\gamma_h\left[C_{M_{Yh}}^2+C_{M_{Xh}}^2-2C_{M_{Yh}}C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}}\right]$$

$\Rightarrow \rho_{M_{Yh}M_{Xh}}{}^2 C_{M_{Yh}}^2 - 2C_{M_{Yh}}C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}} + C_{M_{Xh}}^2 > 0$

$\Rightarrow \left(\rho_{M_{Yh}M_{Xh}}C_{M_{Yh}} - C_{M_{Xh}}\right)^2 > 0$ or $\left(\rho_{M_{Yh}M_{Xh}} - \frac{C_{M_{Xh}}}{C_{M_{Yh}}}\right)^2 > 0$ will always be satisfied.

From the above condition, it is evident that the $MSE\left(\hat{M}_{st}(\alpha_h)\right)$ will be more efficient than $MSE(\hat{M}_{YRS})$.

Condition 3: Efficiency comparison between (xxii) and (vii)

$MSE\left(\hat{M}_{st}(\alpha_h)\right) < V_{Min}(\hat{M}_{YDS})$ if

$$\sum_{h=1}^{H}W_h^2\gamma_h M_{Yh}^2 C_{M_{Yh}}^2(1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h=1}^{H}W_h\gamma_h M_{Yh}^2 C_{M_{Yh}}^2(1-\rho_{M_{Yh}M_{Xh}}{}^2)$$

$MSE\left(\hat{M}_{st}(\alpha_h)\right)$ will be more efficient than $V_{Min}(\hat{M}_{YDS})$ if $W_h^2 < W_h$

Condition 4: Efficiency comparison between (xxii) and (xiii)

$MSE\left(\hat{M}_{st}(\alpha_h)\right) < MSE_{Min}(\hat{M}_{Pis})$ if

$$\sum_{h=1}^{H}W_h^2\gamma_h M_{Yh}^2 C_{M_{Yh}}^2(1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h=1}^{H}W_h\gamma_h M_{Yh}^2 C_{M_{Yh}}^2(1-\rho_{M_{Yh}M_{Xh}}{}^2)$$

$MSE\left(\hat{M}_{st}(\alpha_h)\right)$ will be more efficient than $V_{Min}(\hat{M}_{Pis})$ if $W_h^2 < W_h$

Condition 5: Efficiency comparison between (xxxii) and (xxii)

$MSE\left(\hat{M}_{st}^{*}(\alpha_h)\right) < MSE\left(\hat{M}_{st}(\alpha_h)\right)$ if

$$\frac{M_y^2\left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2\sum_{h=1}^{H}W_h^2\gamma_h C_{M_{Yh}}^2}{\left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2\sum_{h=1}^{H}W_h^2\gamma_h C_{M_{Yh}}^2+M_y^2} < \sum_{h=1}^{H}W_h^2\gamma_h M_{Yh}^2 C_{M_{Yh}}^2(1-\rho_{M_{Yh}M_{Xh}}{}^2)$$

$$\sum_{h=1}^{H} W_h^2 \gamma_h M_{Yh}^2 C_{M_{Yh}}^2 \left(1 - \rho_{M_{Yh}M_{Xh}}{}^2\right) > \frac{\left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2 \sum_{h=1}^{H} W_h^2 \gamma_h C_{M_{Yh}}^2}{\left[1 + \left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2 \sum_{h=1}^{H} W_h^2 \gamma_h C_{M_{Yh}}^2\right]}$$

$$\rho_{M_{Yh}M_{Xh}}{}^2 < 1 - \frac{\left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2 \sum_{h=1}^{H} W_h^2 \gamma_h C_{M_{Yh}}^2}{\left[1 + \left[\frac{V(m_{xs})}{\hat{V}(m_{xs})}\right]^2 \sum_{h=1}^{H} W_h^2 \gamma_h C_{M_{Yh}}^2\right] \sum_{h=1}^{H} W_h^2 \gamma_h M_{Yh}^2 C_{M_{Yh}}^2}$$ is satisfied, then

$MSE\left(\hat{M}_{st}^*(\alpha_h)\right)$ will be more efficient than $MSE\left(\hat{M}_{st}(\alpha_h)\right)$ and as such better than those considered in conditions 1-4.

5. APPLICATION

Simulation study with data generated based on model

To demonstrate the performance of the proposed estimators with respect to the existing ones,
a simulation study is considered by assuming that the interest and auxiliary variables follow certain kinds of probability distributions. Bivariate observations (x_i, y_i) are generated with a population size of 1200 using the model

$$Y_{hi} = \beta_0 + \beta_1 X_{hi} + e_{hi}, \qquad i = 1,2,\dots,N;\ h = 1,2,3$$

where X_{hi} are generated from different probability distributions that depict real life situations which include; chi square with 10 degrees of freedom, log normal with mean 0 and variance 5 and Cauchy with mean 0 and variance 1. The values of β_0 and β_1 are given as 5 and 1.5 respectively while e_{hi} are independently generated from normal distribution with mean 1 and variance 2. The population is further stratified into 3 strata with stratum population sizes 200, 400, 600 respectively. Using simple random sampling without replacement, different sample settings of 10%, 15%, and 20% are selected using proportional allocation from each stratum.

The MSE values of the existing (sample median, Aladag and Cingi (2015) separate ratio, separate difference, and separate product estimators) and proposed estimators are computed as shown in Table 2. Also, the percent relative efficiencies of the estimators are obtained as follows:

$$PRE = \frac{MSE(\hat{M}_{YS})}{MSE(.)} \times 100$$

where $MSE(\hat{M}_{YS})$ is the MSE of classical median estimator in stratified sampling which is the same as $V(m_{ys})$ and $MSE(.)$ denotes the MSE of estimators mentioned here. The results of analyses are shown in Tables 3.

Table 2. MSE of the Proposed and Existing Estimators

Distribution	Sample Size	$\hat{M}_{YS}$	$\hat{M}_{YRS}$	$\hat{M}_{YDS}$	$\hat{M}_{Pis}$	$\hat{M}_{st}(\alpha_h)$	$\hat{M}_{st}^*(\alpha_h)$
Chi square	10%	4.909	4713.652	0.814	0.814	0.791	8.573e-07
	15%	3.539	3281.177	0.536	0.536	0.533	2.927e-07
	20%	2.182	2094.957	0.362	0.362	0.352	7.527e-08
Log-norm	10%	4.909	98302.610	0.814	0.814	0.791	6.034e-06
	15%	3.539	68856.450	0.536	0.536	0.533	2.067e-06
	20%	2.182	43690.050	0.362	0.362	0.352	5.298e-07
Cauchy	10%	4.909	8107823.0	0.814	0.814	0.791	1.225e-05
	15%	3.539	6116420.0	0.536	0.536	0.533	4.168e-06
	20%	2.182	3603477.0	0.362	0.362	0.352	1.076e-06

Table 3. PRE of the Proposed and Existing Estimators

Distribution	Sample Size	$\hat{M}_{YS}$	$\hat{M}_{YRS}$	$\hat{M}_{YDS}$	$\hat{M}_{Pis}$	$\hat{M}_{st}(\alpha_h)$	$\hat{M}_{st}^*(\alpha_h)$
Chi square	10%	100	0.104	603.170	603.170	620.592	5.726e+08
	15%	100	0.108	660.208	660.208	663.798	1.209e+09

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

	20%	100	0.104	603.170	603.170	620.593	2.899e+09
Log-norm	10%	100	4.993e-03	603.170	603.170	620.593	8.135e+07
	15%	100	5.141e-03	660.208	660.208	663.798	1.713e+08
	20%	100	4.993e-03	603.170	603.170	620.593	4.118e+08
Cauchy	10%	100	6.054e-05	603.170	603.170	620.593	4.007e+07
	15%	100	5.787e-05	660.208	660.208	663.798	8.492e+07
	20%	100	6.054e-05	603.170	603.170	620.593	2.027e+08

Numerical Example of Real-life Data

Data used by Aladag and Cingi (2015) will be considered in this work to compliment the theoretical findings. Computations on the development index of all district in Turkey about educational opportunities using the data gathered from schools by Ministry of National Education for 2006-2007 educational year. The authors obtained the development groups by clustering the districts with the same development level in the same group. Data for number of Teachers is considered as study variable in elementary schools for 923 districts in Turkey in 2007 and the number of students as auxiliary variable at six regions (1. Mediterranean, 2. Aegean, 3. East and Southeast Anatolia, 4. Central Anatolia, 5. Black Sea, 6. Marmara). Proportional allocation was used in determining the sample size of each stratum. The data are shown in Table 4.

Table 4. Data Statistics

$\boldsymbol{N_1 = 91}$	$\boldsymbol{\rho_{M_{Y1}M_{X1}} = 0.84}$	$\boldsymbol{f_X(M_{Y1}) = 0.003160}$
$\boldsymbol{n_1 = 18}$	$M_{Y1} = 81$	$f_X(M_{X1}) = 0.000190$
$\boldsymbol{P_{111} = 0.46}$	$M_{X1} = 1265$	$R_{X1} = 56862, M_{01} = 290$
$\boldsymbol{N_2 = 129}$	$\rho_{M_{Y2}M_{X2}} = 0.96$	$f_X(M_{Y2}) = 0.003180$
$\boldsymbol{n_2 = 26}$	$M_{Y2} = 93$	$f_X(M_{X2}) = 0.000240$
$\boldsymbol{P_{112} = 0.49}$	$M_{X2} = 1139$	$R_{X2} = 45559, M_{02} = 233$
$\boldsymbol{N_3 = 204}$	$\rho_{M_{Y3}M_{X3}} = 0.84$	$f_X(M_{Y3}) = 0.011510$
$\boldsymbol{n_3 = 41}$	$M_{Y3} = 24$	$f_X(M_{X3}) = 0.000486$
$\boldsymbol{P_{113} = 0.49}$	$M_{X3} = 614$	$R_{X3} = 42014, M_{03} = 468$
$\boldsymbol{N_4 = 145}$	$\rho_{M_{Y4}M_{X4}} = 0.88$	$f_X(M_{Y4}) = 0.000299$
$\boldsymbol{n_4 = 29}$	$M_{Y4} = 54$	$f_X(M_{X4}) = 0.004420$
$\boldsymbol{P_{114} = 0.47}$	$M_{X4} = 763$	$R_{X4} = 41652, M_{04} = 226$
$\boldsymbol{N_5 = 184}$	$\rho_{M_{Y5}M_{X5}} = 0.88$	$f_X(M_{Y5}) = 0.005120$
$\boldsymbol{n_5 = 29}$	$M_{Y5} = 44$	$f_X(M_{X5}) = 0.000523$
$\boldsymbol{P_{115} = 0.47}$	$M_{X5} = 533$	$R_{X5} = 26705, M_{05} = 140$
$\boldsymbol{N_6 = 170}$	$\rho_{M_{Y6}M_{X6}} = 0.96$	$f_X(M_{Y6}) = 0.000249$
$\boldsymbol{n_6 = 34}$	$M_{Y6} = 101$	$f_X(M_{X6}) = 0.000087$
$\boldsymbol{P_{116} = 0.49}$	$M_{X6} = 911$	$R_{X6} = 26823, M_{06} = 198$

Table 5. MSE and PRE of the estimators

Estimator	MSE	PRE
$\boldsymbol{\hat{M}_{YS}}$	6189.39679	100.00000
$\boldsymbol{\hat{M}_{YRS}}$	4080.03230	151.69970
$\boldsymbol{\hat{M}_{YDS}}$	867.99978	713.06433
$\boldsymbol{\hat{M}_{Pis}}$	867.99978	713.06433
$\boldsymbol{\hat{M}_{st}(\alpha_h)}$	689.19880	898.05681
$\boldsymbol{\hat{M}^*_{st}(\alpha_h)}$	0.0000007087	8.73×10^9

6. DISCUSSION OF RESULTS

Results from Tables 2 and 3 are indications that the numerical illustrations from the simulation study support the efficiency comparisons from the theoretical results obtained earlier. As observed, the proposed calibration estimator $\widehat{M}_{st}^{*}(\alpha_h)$ has a negligible MSE and an overwhelming performance in terms of higher gains in efficiency compared to $\widehat{M}_{st}(\alpha_h)$ (the median version of the Vishwakarma and Singh (2011) separate ratio-product estimator), the classical median estimator, the Aladag and Cingi (2015) separate ratio $\widehat{M}_{YRS}$, separate difference $\widehat{M}_{YDS}$, and separate product $\widehat{M}_{Pis}$ estimators. As a result, it is considered the best estimator with respect to separate estimators for estimating population median.

The result from the real data as shown in Table 5 supports the simulation results that the proposed estimator has an outstanding higher percent relative efficiency than $\widehat{M}_{st}(\alpha_h)$, the classical median estimator and Aladag and Cingi (2015) unbiased estimators $\widehat{M}_{YDS}$ and $\widehat{M}_{Pis}$. Remarkably, in both simulation and real-life data, the results showed that, $\widehat{M}_{st}(\alpha_h)$ also performs exceedingly better than other existing estimators in this study which is supported in the efficiency comparison. However, unlike the simulated data, the separate ratio estimator $\widehat{M}_{YRS}$ performs better than $\widehat{M}_{YS}$ the classical median estimator.

This superiority in the gain in efficiency of the proposed estimator is as a result of the calibration constraint which is formulated based on the variance of the classical median estimator of the auxiliary variable. It becomes imperative to sought for an estimator with minimum MSE since most of the separate estimators of population median have been shown to be asymptotically unbiased as demonstrated by Aladag and Cingi (2015). The idea of calibration in this case has really paid off in improving the efficiency of the median estimator under stratified random sampling.

7. CONCLUSION

This study was based on formulating an improved estimator for population median by adopting the Vishwakarma and Singh (2011) separate ratio-product estimator of the population mean. The conditions necessary for this formulation assumed that the population median of the auxiliary variable is known. Calibration estimation technique was further applied to derive an improved version of the suggested estimator of population median as well as the mean square error.

Contrary to the results obtained by Aladag and Cingi (2015) that various transformations of the auxiliary variable do not affect the value of the minimum MSE of the separate estimators in stratified random sampling, proper formulation of calibration constraints has offered a more fruitful result. Results from both simulation and real data analyses justify the claims. Conclusively, it is evident to say that the proposed estimator is the optimum estimator in estimating population median when the population median of the auxiliary variable is known and positively correlated with the study variable. In addition, the proposed estimator will be suitable and highly recommended when the variable considered is from a distribution that is skewed.

REFERENCES

[1] Deville, J. C. and C. E. Sarndal, "Calibration estimators in survey sampling", Journal of the American Statistical Association, vol. 87, pp. 376-382, 1992.
[2] Sarndal, C. E, "The calibration approach in survey theory and practice", Survey Methodology, vol. 33 pp. 99-119, 2007.
[3] Singh, S. and R. Arnab, "On calibration of design weights", International Journal of Statistics, vol. 69, no. 2, pp. 185-205, 2014.
[4] Clement, E. P. and E. I. Enang. "On the efficiency of ratio estimator over the regression estimator", Communication in Statistics: Theory and Methods vol. 46, no. 11, pp. 5357-5367, 2017.
[5] Clement, E. P., "Calibration approach separate ratio estimator for population mean in stratified sampling", International Journal of Modern Mathematical Sciences, vol. 13, no. 4, pp. 377-384, 2015.
[6] Clement, E. P., "A new ratio estimator of mean in survey sampling by calibration estimation", Elixir Statistics, vol. 106, pp. 46461-46465, 2017.
[7] Koyuncu, N and C. Kadilar. 2013. Calibration estimators using different measures in stratified random sampling. International Journal of Modern Engineering research 3 (1): 414-419.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[8] Gross, T.S., "Median estimation in sample surveys", In American Statistical Association Proceedings of Survey Research methodology Section, pp. 181-184, 1980.
[9] Kuk, A. Y. C. and T. K. Mak, "Median estimation in the presence of auxiliary variable", Journal of Royal Statistical Society. Series B, vol. 51: pp. 261-269, 1989.
[10] Singh, H. P., S. Singh and S. M. Puertas. "Ratio-type estimators for the median of finite populations", Allegemeines Statistisches Archiv, vol. 87, pp. 369-382, 2003.
[11] Singh, H. P. and R. S. Solanki, "Some classes of estimators for the population median using auxiliary information", Communication in Statistics, vol. 42, pp.4222-4238, 2013.
[12] Aladag, S. and H. Cingi, "Improvement in estimating the population median in simple random sampling and stratified random sampling using auxiliary information", Communication in Statistics-Theory and Methods, vol. 45, no. 5, 1013-1032, 2015.
[13] Mursala, K. and Y. A. Abdullah, "A chain ratio-type estimator for population median under double sampling", Journal of Statistics Application and Probability, vol. 4, no.3, pp. 1-5, 2015.
[14] Garg, N. and M Pachori, "Use of median in calibration estimation of finite population mean in stratified sampling", International Journal of Computer Science and Engineering, vol. 7, no. 5, pp. 667-671, 2019.
[15] Vishwakarma, G. K. and H. P. Singh, "Separate ratio-product estimator for estimating population mean using auxiliary information", Journal of statistical Theory and Applications vol. 10, no. 4, pp. 653-664, 2011..