ELEVATED FAMILY OF ESTIMATORS FOR ESTIMATING THE POPULATION MEAN UTILIZING THE KNOWN POPULATION MEDIAN OF THE MAIN VARIABLE

Dinesh K. Sharma, University of Maryland Eastern Shore, USA
Subhash Kumar Yadav, Babasaheb Bhimrao Ambedkar University, India
Kate Brown, University of Maryland Eastern Shore, USA

ABSTRACT

In this paper, we present an improved family of estimators of the population mean of the primary variable utilizing known information on the population median of the study variable under a simple random sampling without replacement scheme. The known population median of study variable is often available without increasing the cost of the sample. The bias and MSE of the suggested family have been derived up to the approximation of degree one. We compare the proposed family of estimators with other competing estimators and also find the conditions under which it performs better than the competing estimators. These conditions are verified with an empirical study and show that the proposed family performs best among the competing estimators of the population mean of the study variable.

Key words: Study variable, auxiliary variable, Estimator, Bias, Mean Squared Error.

INTRODUCTION

Sampling is of paramount importance for large populations as it saves time and money. The most common estimator for the population mean is the corresponding sample mean. Although the sample mean is unbiased, its sampling variance can be large, which means it is not the most efficient estimator of the population mean. Improved efficiency can be found with the use of an auxiliary variable which has a high degree of both positive and negative correction with the study variable (Cochran, 1940; Robson, 1957; Murthy, 1964). Information about the auxiliary variable often increases the cost of the sample. Therefore, we look for known parameters of the study variable which are readily available to improve the estimation of the population mean. One such parameters is the population median (Subramani, 2016).

In this paper, we suggest an improved family of estimators of the population mean using the known population median of the study variable which is often easily available.

REVIEW OF SOME EXISTING ESTIMATORS

A review of existing estimators of the population mean making use of auxiliary information along with the sample mean has been presented in Table 1.

Where,

$$C_y = \frac{S_y}{\overline{Y}}\,,\ S_y^2 = \frac{1}{N-1}\sum_{i-1}^{N}(Y_i - \overline{Y})^2 = \frac{1}{{}^N C_n}\sum_{i-1}^{{}^N C_n}(\overline{y}_i - \overline{Y})^2\,,\ C_{yx} = \rho_{yx}\, C_y C_x\,,\ f = \frac{n}{N}\,,\ C_x = \frac{S_x}{\overline{X}}\,,$$

$$S_x^2 = \frac{1}{N-1}\sum_{i-1}^{N}(X_i - \overline{X})^2 = \frac{1}{{}^N C_n}\sum_{i-1}^{{}^N C_n}(\overline{x}_i - \overline{X})^2\,,\ Cov(x,y) = \frac{1}{N-1}\sum_{i-1}^{N}(Y_i - \overline{Y})(X_i - \overline{X})\,,$$

$\rho_{yx} = \dfrac{Cov(x,y)}{S_x S_y}$, $R_5 = \dfrac{\bar{Y}}{M}$, $C_m = \dfrac{S_m}{M}$, $S_{ym} = \dfrac{1}{{}^N C_n}\sum_{i-1}^{{}^N C_n}(\bar{y}_i - \bar{Y})(m_i - M)$, $C_{ym} = \dfrac{S_{ym}}{\bar{Y}M}$ and

$S_m^2 = \dfrac{1}{{}^N C_n}\sum_{i-1}^{{}^N C_n}(m_i - M)^2$.

Table 1: Various estimators with their mean squared errors

S. No.	Estimator	Mean Squared Error
1.	$t_o = \bar{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$	$V(t_0) = \dfrac{1-f}{n} S_y^2 = \dfrac{1-f}{n}\bar{Y}^2 C_y^2$
2.	$t_1 = \bar{y} + \beta(\bar{X} - \bar{x})$ Watson (1937)	$V(t_1) = \dfrac{1-f}{n}\bar{Y}^2 C_y^2 (1-\rho_{yx}^2)$
3.	$t_2 = \bar{y}\dfrac{\bar{X}}{\bar{x}}$ Cochran (1940)	$MSE(t_2) = \dfrac{1-f}{n}\bar{Y}^2[C_y^2 + C_x^2 - 2C_{yx}]$
4.	$t_3 = \bar{y}\exp\left[\dfrac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right]$ Bahl and Tuteja (1991)	$MSE(t_3) = \dfrac{1-f}{n}\bar{Y}^2[C_y^2 + \dfrac{C_x^2}{4} - C_{yx}]$
7.	$t_4 = \bar{y}\left[\dfrac{\bar{X}}{\bar{X} + \alpha(\bar{x} - \bar{X})}\right]$ Reddy (1974)	$MSE_{\min}(t_4) = \dfrac{1-f}{n}\bar{Y}^2 C_y^2 (1-\rho_{yx}^2)$ for $\alpha_{opt} = C_{yx} \big/ C_x^2$
13.	$t_5 = \bar{y}\dfrac{M}{m}$ Subramani (2016)	$MSE(t_5) = \dfrac{1-f}{n}\bar{Y}^2[C_y^2 + R_5^2 C_m^2 - 2R_5 C_{ym}]$

PROPOSED FAMILY OF ESTIMATOR

Motivated by various authors in the literature along with Subramani (2016), we propose the following family of estimators of the population mean of the study variable using the known population median of study variable as,

$$t_p = [\varphi_1 \bar{y} + \varphi_2 (M - m)]\left(\frac{M}{m}\right)^{\delta} \exp\left[\frac{\eta(M - m)}{M + m}\right] \quad (1)$$

Where φ_1, φ_2, δ and η are constants to be determined such that $\varphi_1 > 0$ and $-\infty < \varphi_2 < \infty$.

By putting various values of the above four constants, we can find various known and unknown members of this proposed family of estimators. Some of those estimators are presented in Table 2.

The bias and the mean squared error (MSE) of the proposed family of estimators (1) up to the approximation of order one are respectively given by,

$$B(t_p) = \bar{Y}[(\varphi_1 - 1) - B_1 \frac{B(m)}{M} - B_2 \lambda C_{ym} + B_3 \lambda C_m^2] \quad (2)$$

Table 2: Some known and unknown members of the proposed family

S. No.	Estimator	φ_1	φ_2	δ	η
1.	$\bar{y}$ [Mean per unit estimator]	1	0	0	0
2.	$\varphi_1\bar{y}$ [Searls (1964) estimator]	φ_1	0	0	0
3.	$\bar{y}+\varphi_2(M-m)$ [Yadav and Pandey (2017) difference type estimator]	1	φ_2	0	0
4.	$\bar{y}+b_{ym}(M-m)$ [Yadav *et al.* (2017) regression type estimator]	1	b_{ym}	0	0
5.	$\bar{y}\left(\frac{M}{m}\right)$ [Subramani (2016) ratio estimator]	1	0	1	0
6.	$t_{p1}^{*}=[\varphi_1\bar{y}+\varphi_2(M-m)]\left(\frac{M}{m}\right)\exp\left[\frac{(M-m)}{M+m}\right]$	φ_1	φ_2	1	1
7.	$t_{p2}^{*}=[\varphi_1\bar{y}+\varphi_2(M-m)]\left(\frac{M}{m}\right)^2\exp\left[\frac{(M-m)}{M+m}\right]$	φ_1	φ_2	2	1
8.	$t_{p3}^{*}=[\varphi_1\bar{y}+\varphi_2(M-m)]\left(\frac{M}{m}\right)$	φ_1	φ_2	1	0
9.	$t_{p4}^{*}=[\varphi_1\bar{y}+\varphi_2(M-m)]\left(\frac{M}{m}\right)^2$	φ_1	φ_2	2	1
10.	$t_{p5}^{*}=[\varphi_1\bar{y}+\varphi_2(M-m)]\exp\left[\frac{(M-m)}{M+m}\right]$	φ_1	φ_2	0	1

$$MSE(t_p)=\bar{Y}^2[1+D_1\varphi_1^2+D_2\varphi_2^2+2D_3\varphi_1\varphi_2-2D_4\varphi_1-2D_5\varphi_2] \quad (3)$$

where,

$$D_1=[1+\lambda C_y^2+(A_1+2A_2)\lambda C_m^2-2A_1\lambda C_{ym}-2A_1\frac{B(m)}{M}],\ D_2=\lambda C_m^2R_m^2,$$

$$D_3=R_m[2A_1\lambda C_m^2-\lambda C_{ym}-\frac{B(m)}{M}],\ D_4=[1-A_1\lambda C_{ym}+A_2\lambda C_m^2-A_1\frac{B(m)}{M}],$$

$$D_5=R_m[\frac{B(m)}{M}-A_1\lambda C_m^2]$$

The $MSE(t_p)$ is minimized for the optimum φ_1 and φ_2 when

$$\begin{cases}\varphi_1=\dfrac{D_2D_4-D_3D_5}{D_1D_2-D_3^2}=\varphi_1^{*}(say)\\[2ex] \varphi_2=\dfrac{D_1D_5-D_3D_4}{D_1D_2-D_3^2}=\varphi_2^{*}(say)\end{cases} \quad (4)$$

The least MSE of the proposed estimator (t_p) is thus obtained as,

$$MSE_{\min}(t_p)=\bar{Y}^2(1-D) \quad (5)$$

where, $D = \frac{(D_2 D_4^2 - 2D_3 D_4 D_5 + D_1 D_5^2)}{(D_1 D_2 - D_3^2)}$

EMPIRICAL STUDY

For a numerical comparison, we have considered the population given in Subramani (2016). The parameters for this population are given below.

$N = 34$, $n = 5$, ${}^N C_n = 278256$, $\bar{Y} = 856.4118$, $\bar{M} = 736.9811$, $M = 767.50$, $\bar{X} = 208.8824$, $R_{12} = 1.1158$, $C_y^2 = 0.125014$, $C_x^2 = 0.088563$, $C_m^2 = 0.100833$, $C_{ym} = 0.07314$, $C_{yx} = 0.047257$, $\rho_{yx} = 0.4491$, $\rho_{ym} = 0.72538$

The MSE and percentage relative efficiency (PRE) with respect to the mean per unit estimator is given in Table 3.

Table 3. MSE and PRE of the existing estimators w.r.t sample mean estimator

Estimator	MSE	PRE
t_0	15585.63	100.00
t_1	12486.75	124.82
t_2	14895.27	104.64
t_3	12498.01	124.71
t_6	12486.75	124.82
t_{12}	10926.53	142.64
t_{p1}	8807.83	176.95
t_{p1}^*	8504.56	183.26
t_{p2}^*	8665.67	179.86
t_{p3}^*	8694.34	179.26
t_{p4}^*	8637.75	180.44
t_{p5}^*	8677.34	179.61

RESULTS AND CONCLUSION

In this paper, we proposed a family of estimators of the population mean using the known population median of the study variable and studied its properties up to the approximation of order one. The MSE of various competing and proposed estimators are calculated for comparison purposes and are presented in Table 3. It can be seen from Table 3 that the proposed estimator has smaller MSE and the MSE of the first unknown member of the proposed family has the smallest MSE demonstrating the value of the new estimators. Thus the proposed family should be used for improved estimation of population mean of study variable when its population median is known.

REFERENCES

Abid, M., Abbas, N. Sherwani, R.A.K. and Nazir, H. Z. (2016). Improved Ratio Estimators for the population mean using non-conventional measure of dispersion. *Pakistan Journal of Statistics and Operations Research*, XII (2), 353-367.

Bahl, S. and Tuteja, R.K. (1991). Ratio and product type exponential estimator, *Information and Optimization Sciences*, XII (I), 159-163.

Cochran, W. G. (1940). The Estimation of the Yields of the Cereal Experiments by Sampling for the Ratio of Grain to Total Produce. *The Journal of Agric. Science*, 30, 262-275.

Reddy, V.N. (1974). On a transformed ratio method of estimation. *Sankhya*, C, 36(1), 59-70.

Searls, D. T. (1964). The Utilization of a Known Coefficient of Variation in the Estimation Procedure. Journal of the American Statistical Association. 59: 1225-1226.

Subramani, J. (2016). A new median based ratio estimator for estimation of the finite population mean, *Statistics in Transition New Series*, 17, 4, 1-14.

Watson, D.J. (1937). The estimation of leaf area in field crops, *The Journal of Agricultural Science*, 27, 3,

Yadav, S.K. and Pandey, H. (2017). A new difference type median based estimator of the Finite population mean, *International journal of Agricultural and Statistical Sciences*, 13, 1, 289-295.

Yadav, S. K., Singh, L., Mishra, S. S., Mishra, P. P. and Kumar, S. (2017). A median based regression type estimator of the finite population mean, *International journal of Agricultural and Statistical Sciences*, 13, 1, 265-271.

(A complete list of references is available upon request from Dinesh K. Sharma at dksharma@umes.edu)