DOI: https://doi.org/10.36094/sc.v89.2023.Comparison_of_Different_Existing_Estimators.Sunani_Bhoi.284

Sci. and Cult. 89 (7–8) : 284-287 (2023)

Comparison of Different Existing Estimators for Estimating Finite Population Mean: A Simulation Study

***Abstract*: This article explains on some comparison of different existing estimators suggested by Cochran[1], Robson[2] and Murthy[3], Srivenkataramana[4] and Bandopadhyay[5], Bahl and Tuteja[6], Swain[7]. The bias and mean square error (MSE) has been calculated to see the behaviour of the existing estimators. The numerical comparison was made using the Python program.**

Keywords: Population mean, auxiliary variable, study variable, bias, mean square error (MSE), percent relative efficiency (PRE)

In sample survey literature, estimation of finite population parameters with a greater accuracy has become a matter of concern among the researchers for a long decade. For this the auxiliary information has been used at several stages like pre-selection stage or selection stage or even post selection stage or at several stages simultaneously. Consider a finite population having N distinct and identifiable units and y_i is the value of the study variable for the ith unit of the population (i = 1, 2, ..., N) with unknown population mean $\bar{Y}$. In order to estimate, we select a simple random sample of size n from this population using without replacement scheme. In complete absence of any auxiliary information, we use traditional mean per unit estimator

$$t_0 = \bar{y} \tag{1}$$

for estimating population mean $\bar{Y}$ of the study variable **y**. But when the auxiliary information is available or can be made available by sacrificing a little cost from the total cost of survey, then it can be used to design more efficient estimators. Let x be the auxiliary variable highly correlated with y is available and x_i is the value of for the i-th unit of the population. Using auxiliary information at estimation stage was considered by (Cochran[1]) by suggesting the ratio estimator

$$t_1 = \bar{y}\frac{\bar{X}}{\bar{x}} \tag{2}$$

Robson[2] and Murthy[3] suggested product estimator of the population mean as

$$t_2 = \bar{y}\frac{\bar{x}}{\bar{X}} \tag{3}$$

Both of these estimators assumes that the population mean of the auxiliary variable $\bar{X}$ is known in advance. Cochran[8] advocated in favour of the ratio estimator t_1 that it is preferable to use when the population correlation coefficient ρ between the study variable (y) and the auxiliary variable (x) is highly positive, i.e., $\rho > \frac{1}{2}$ in case when the population ratio $R = \frac{\bar{Y}}{\bar{X}}$ is positive. On contrary to this, we prefer the product estimator t_2, when the population correlation coefficient is highly negative, i.e., $\rho < -\frac{1}{2}$.

By considering $\frac{\bar{X} - f\bar{x}}{1-f}$ as an unbiased estimator of $\bar{X}$, Srivenkataramana[4] used the linear transformation from x to x' such that $x_i' = (1+g)\bar{X} - gx_i$, $g = \frac{f}{(1-f)}$, $f = \frac{n}{N}$ is called as the sampling fraction , and suggested the dual to ratio estimator as

$$t_3 = \bar{y}\frac{\bar{x}}{\bar{X}} \tag{4}$$

This estimator performs better when y and x are negatively correlated with each other. Bandopadhyay[5] suggested dual to product estimator for estimating $\bar{Y}$ as

$$t_4 = \bar{y}\frac{\bar{X}}{\bar{x}} \tag{5}$$

The journal is in the category 'Group A' of UGC-CARE list and falls under the broad category of Multidisciplinary Sciences covering the areas Arts and Humanities, Science and Social Sciences.

The usual exponential-ratio estimators and product estimator proposed by Bahl and Tuteja[6], when using single auxiliary variables is given by

$$t_5 = y \exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) \quad (6)$$

$$t_6 = \bar{y} \exp\left(\frac{\bar{x} - \bar{X}}{\bar{x} + \bar{X}}\right) \quad (7)$$

The square root estimator proposed by Swain (2014) and it is defined as

$$t_6 = \bar{y}\left(\frac{\bar{X}}{\bar{X}}\right)^{1/2} \quad (8)$$

Properties of the Existing Estimators: To study the large sample properties of the class of estimators, we consider

Let, $e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}} \Rightarrow \bar{y} = \bar{Y}(1 + e_0)$ and

$$e_1 = \frac{\bar{x} - \bar{X}}{\bar{X}} \Rightarrow \bar{x} = \bar{X}(1 + e_1)$$

and their expectations are given as

$$E(e_0) = 0, E(e_1) = 0,$$

$$E\left(e_0^2\right) = \frac{1}{\bar{Y}^2} E\left(\bar{y} - \bar{Y}\right)^2 = \frac{1}{\bar{Y}^2}\frac{N-n}{Nn} S_y^2 = \frac{1-f}{n}\frac{S_y^2}{\bar{Y}^2}$$

$$= \frac{1-f}{n} C_y^2,$$

$$E\left(e_1^2\right) = \frac{1-f}{n} c_x^2$$

$$E(e_0 e_1) = \frac{1}{\bar{Y}\,\bar{X}} E\left[\left(\bar{x} - \bar{X}\right)\left(\bar{y} - \bar{Y}\right)\right]$$

$$= \frac{1}{\bar{Y}\,\bar{X}}\frac{N-n}{Nn}\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)\left(y_i - \bar{Y}\right)$$

$$= \frac{1}{\bar{Y}\,\bar{X}}\frac{1-f}{n} S_{y_x} = \frac{1-f}{n}\rho c_y c_x = \frac{1-f}{n} k c_x^2 \left(k = \frac{\rho c_y}{c_x}\right)$$

Bias of Different Estimators: The bias of different estimators are given below

Table: 1. Bias of different estimators

$B(t_1) = \bar{Y}\frac{1-f}{n}\left(c_x^2 - \rho c_x c_y\right)$	(9)
$B(t_2) = \bar{Y}\frac{1-f}{n}\rho c_x c_y$	(10)
$B(t_3) = -\bar{Y}\frac{1-f}{n} g\rho c_y c_x$	(11)
$B(t_4) = \bar{Y}\frac{1-f}{n} g\left(g c_x^2 + \rho c_y c_x\right)$	(12)
$B(t_5) = \bar{Y}\frac{1-f}{n}\left(\frac{3}{8}c_x^2 - \frac{1}{2}\rho c_x c_y\right)$	(13)
$B(t_6) = \bar{Y}\frac{1-f}{n}\left(\frac{\rho c_x c_y}{2} - \frac{c_x^2}{8}\right)$	(14)
$B(t_7) = \bar{Y}\frac{1-f}{n}\left(\frac{3}{8}c_x^2 - \frac{1}{2}\rho c_x c_y\right)$	(15)

Mean Square Error of the Different Existing Estimators: The variance of different estimators are given below

Table: 2.Mean square error (MSE) of different estimators

$V(t_0) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2 C_y^2$	(16)
$MSE\ (t_1) = \bar{Y}^2\frac{1-f}{n}\left(c_x^2 + c_y^2 - 2\rho c_x c_y\right)$	(17)
$MSE\ (t_2) = Y^2\frac{1-f}{n}\left(c_x^2 + c_y^2 + 2\rho c_x c_y\right)$	(18)
$MSE\ (t_3) = \bar{Y}^2\frac{1-f}{n}\left(c_y^2 + g^2 c_x^2 - g\rho c_x c_y\right)$	(19)
$MSE\ (t_4) = \bar{Y}^2\frac{1-f}{n}\left(c_y^2 + g^2 c_x^2 + g\rho c_x c_y\right)$	(20)
$MSE\ (t_5) = \bar{Y}^2\frac{1-f}{n}\left(c_y^2 + \frac{1}{4}c_x^2 - \rho c_x c_y\right)$	(21)
$MSE\ (t_6) = \bar{Y}^2\frac{1-f}{n}\left(c_y^2 + \frac{1}{4}c_x^2 + \rho c_x c_y\right)$	(22)
$MSE\ (t_7) = \bar{Y}^2\frac{1-f}{n}\left(c_y^2 + \frac{1}{4}c_x^2 - \rho c_x c_y\right)$	(23)

Empirical Study: To study the behaviour of the existing estimators, we have considered three natural populations, which are available in different text books. The bias and mean square error (MSE) of the different existing estimators are tabulated in table no. (1 & 2) as well as the percent relative efficiency (PRE). We have used Python (Jupyter Notebook) software for numerical illustration. Table.3. represents the sources and description of the populations, table.4 represents the characteristics of the different population parameters, the bias and mean square error of the different estimators are listed in table no. (5 & 6) and the percent relative efficiency is listed in table no 7.

The PRE's of the estimators t is compared with the simple mean estimator t_0 is given by

$$PRE(t) = \frac{V(t_0)}{MSE(t)} \times 100$$

Table 3. Sources and Description of the populations

P.No.	Source	y	x
1	Cochran[8], p.182	Number of Placebo	Number of Paralytic Polio Case in the placebo
2	Singh and Choudhury[9], p.177	Area under wheat in 1974	Area under wheat in 1971
3	Singh and Choudhury[9], p.176	Number of Cow	Census Figure

Table 4. The Characteristics of the different Population Parameters

P.No.	N	n	g	$\bar{Y}$	$\bar{X}$	r	C_{yy}	C_{xx}	C_{yx}
1	34	3	0.096	4.923	2.588	0.732	1.047	1.52	0.924
2	34	6	0.214	199.44	777	0.832	0.567	0.347	0.369
3	20	3	0.176	816.45	641.05	0.893	0.499	0.684	0.522

Table 5. The Bias of different estimators

P.No.	t_1	t_2	t_3	t_4	t_5	t_6	t_7
1	0.892	1.383	0.133	0.112	0.161	0.407	0.161
2	0.607	10.114	2.167	1.730	1.492	3.869	1.492
3	37.616	120.828	21.322	16.388	0.997	40.608	0.997

Table: 6. Mean square error of different estimators

P. No.	t_0	t_1	t_2	t_3	t_4	t_5	t_6	t_7
1	7.715	5.293	32.548	17.330	3.702	8.479	7.160	3.702
2	3097.096	958.630	9027.923	5588.464	1553.817	3616.449	2751.881	1553.817
3	94271.008	26332.63	420935.0	225262.31	27961.115	115708.537	80890.67	27961.11

Table: 7.The Percent relative efficiency (PRE) of different estimators

P.No.	t_0	t_1	t_2	t_3	t_4	t_5	t_6	t_7
1	100	145.757	23.703	208.362	44.518	208.362	107.743	90.985
2	100	323.075	34.305	199.321	55.419	199.321	112.544	85.639
3	100	358.000	22.395	337.150	41.849	337.150	116.541	81.472

Conclusion

In this article, we have made a numerical comparison of different estimators for estimating population mean because these estimators are designed to be used in case of positive/negative correlation coefficient between the study variable and the auxiliary variable. A numerical comparison has been made between the mean per unit estimator, usual ratio estimator and dual to product estimator under SRSWOR scheme. On the support of the empirical evidences, we conclude that the estimators of t_3 due to Srivenkataramana[4] and t_5 due to Bahl and Tuteja[6] are to be preferred due to their maximum gain in efficiency. ❐

KALYANI SUNANI*
BISWAJIT BHOI

* Faculty, Department of Statistics,
Central University of Odisha,
Koraput-763004, Odisha, India
e-mail: sunanikalyani@gmail.com

1. W.G. Cochran, *Journal of Agricultural Science*.**30(2)**, 262-275 (1940).
2. D.S. Robson, *Journal of American Statistical Association*,**52**, p. 411-422 (1957).
3. M.N. Murthy, *Sankhya*, **21**, 381-392 (1964).
4. T. Srivenkataramana, *Biometrika*. **67**, 199-204 (1980).
5. B. Bandopadhyay, *Sankhya*, **242**, 45-49 (1980).
6. S. Bhal and R.K. Tuteja, *Journal of Information and Optimization Sciences*, 159-164 (1991).
7. A.K.P.C. Swain, *Revista Investigacion Operational*, **35**, 49-57 (2014).
8. W.G Cochran, Sampling techniques. Wiley Eastern Limited, New Delhi, III Edition (1977).
9. D. Singh and F.S. Chaudhary, Theory and Analysis of Sample Survey Designs. New Age International (P) Ltd., Chennai (1986).