Advances in Research

21(6): 16-20, 2020; Article no.AIR.57430
ISSN: 2348-0394, NLM ID: 101666096

Unbiased Estimation of Population Mean Using Lahiri-Midzuno-Sen Type Sampling Scheme

Chandni Kumari[1*] and Ratan Kumar Thakur[1]

[1]Department of Statistics, Babasaheb Bhimrao Ambedkar, University (A Central University), Lucknow-226025, India.

Authors' contributions

This work was carried out in collaboration between both the authors. Author CK designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Authors RKT managed the literature searches and final editing of the manuscript. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AIR/2020/v21i630209
Editor(s):
(1) Dr. Martin Kröger, Swiss Federal Institute of Technology (ETH Zürich), Switzerland.
Reviewers:
(1) José Francisco de Oliveira-Júnior, Universidade Federal de Alagoas – UFAL, Brazil.
(2) Paul Inuwa Dalatu, Adamawa State University, Nigeria
Complete Peer review History: http://www.sdiarticle4.com/review-history/57430

Original Research Article

Received 20 March 2020
Accepted 27 May 2020
Published 09 June 2020

ABSTRACT

This paper considers the problem of estimating the population mean under double sampling. We have suggested the generalized class of estimators under Lahiri (1951) to Midzuno (1952) and Sen (1952) type sampling scheme and its properties are studied up to the first order of approximation. Further, we compare the proposed sampling strategy with some conventional estimators under the simple random sampling without replacement. On the basis of suitable range information, we give some concluding remarks related to propose sampling strategy. An empirical study is given in support of the present study.

Keywords: Simple random sampling; ratio type estimator; Lahiri-Midzuno-Sen type sampling scheme.

1. INTRODUCTION

The use of probability in sampling theory came to be recognized as a reliable tool in drawing inferences about the populations, whether finite or not. The main objective of this work is to present the theory and techniques of sample surveys with their applications in different types

**Corresponding author: E-mail: chandnithakur897@gmail.com;*

of problems under double sampling. This type of approach has been considered by [1,2,3,4,5,6] in which they improve the existing sampling strategies by using some auxiliary information under double sampling set up. The purpose of this paper is to give the class of unbiased estimators under the proposed sampling strategy, but they are theoretically biased in simple random sampling without replacement.

Several authors [7,8,11,13,14] had proposed various biased ratio and product type estimators where auxiliary information about the population parameters is not known. In this paper, we make all these estimators unbiased by using [9,10,12] type sampling scheme.

Let the population consists of *N* units. Y_i and X_i denote the i^{th} characteristics of the population. The population mean of the study variable is denoted by $\bar{Y}$ and population mean of the auxiliary variable which is not known is denoted by $\bar{x}'$ which is given as $\bar{x}' = 1/n' \sum x_i$. The population variance of the study variable and the auxiliary variable is denoted by S_y^2 and S_x^2. Let $\mu_{pq} = N^{-1}\sum_{i=1}^{N}(X_i - \bar{X})^p(Y_i - \bar{Y})^q$ be the population product moment between *x* and *y*. C_x and C_y be the coefficient of variation of auxiliary and study variable respectively. Thus, ρ be the correlation coefficient between the variable under study and auxiliary variable which measures the degree of linear relationship between two variables and it is given as ρ = *Cov(x,y)/σxσy*. Now, let us consider the simple random sample of size n' drawn without replacement. From this sample, we estimate the unknown population mean of auxiliary variable. After that, we take a random sample of size *n* from this large sample of size n' drawn without replacement to estimate the sample mean of auxiliary variable. Let y_i be the i^{th} characteristics of the study variable of the sample and x_i be the i^{th} characteristics of the auxiliary variable of the sample. The sample mean of study variable for estimating the population mean is denoted by $\bar{y} = n^{-1}\sum_{i=1}^{n} y_i$. The sample mean of auxiliary variable is denoted by $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$.

2. PROPERTIES OF PROPOSED SAMPLING STRATEGY

The ratio estimator become unbiased by using the proposed sampling strategy even under the double sampling set up for estimating the population mean. Here, Lahiri-Midzuno-Sen type sampling scheme is used for selecting a unit from sample. Thus, at phase I, we select a large sample of size n' drawn from population of size N for estimating the unknown auxiliary variable with probability proportional to x and at phase II, we select a subsample s of size n from this large sample for estimating the sample mean of study and auxiliary variable by simple random sampling without replacement (SRSWOR). Now, the probability of selecting a sample under Lahiri-Midzuno-Sen type sampling scheme is given as

$$P(s) = \sum_{i=1}^{n}(Prob. of\ selecting\ y_i\ at\ first\ draw) \text{ X } \{Prob\ of\ selecting\ (n-1) units\ out\ of\ (n'-1 units\ by\ SRSWOR \quad (1)$$

$$= \frac{\sum_{i=1}^{n}\left(\alpha\bar{X} + \beta + A\alpha(x_i - \bar{x}')\right)}{\sum_{i=1}^{n}(\alpha\bar{X} + \beta)} \frac{1}{\binom{n'-1}{n-1}}$$

$$= \frac{\left(\alpha\bar{X} + \beta + A\alpha(\bar{x} - \bar{x}')\right)}{(\alpha\bar{X} + \beta)\binom{n'}{n}} \quad (2)$$

When the random sample *s* is selected by simple random sampling without replacement, we propose the following new classes of log-type estimators for the population mean $\bar{Y}$ as :

$$\hat{\bar{y}}_A' = \frac{\bar{y}_s(\alpha\bar{X} + \beta)}{(\alpha\bar{X} + \beta) + A\alpha(\bar{x} - \bar{x}')} \quad (3)$$

Where A is the characterizing scalar to be chosen suitably; α and β represents the prior information in the form of the parameters based on auxiliary characters.

Bias and MSE of proposed estimators

In order to obtain the bias and mean square error (MSE), let us consider

$E(\epsilon_0) = E(\epsilon_1) = E(\epsilon_1') = 0$, $E(\epsilon_0)^2 = \frac{1}{n}\left(\mu_{04} - \mu_{02}^2\right)$, $E(\epsilon_1)^2 = \frac{1}{n}\mu_{20}$, $E(\epsilon_1')^2 = \frac{1}{n'}\mu_{20}$, $E(\epsilon_0\epsilon_1) = \frac{1}{n}\mu_{12}$, $E(\epsilon_0\epsilon_1') = \frac{1}{n'}\mu_{12}$, $E(\epsilon_1'\epsilon_1) = \frac{1}{n'}\mu_{20}$ where $\left(\beta_{02} = \frac{\mu_{04}}{\mu_{02}^2}\right)$

Theorem1 $\hat{\bar{y}}_A'$ is an unbiased estimator of population mean under the proposed sampling design P(s) and the bias is given by

$Bias(\hat{\bar{y}}_A^{'}) = \bar{Y}$

Proof. Consider the estimator

$$\hat{\bar{y}}_A^{'} = \frac{\bar{y}_s(\alpha\bar{X}+\beta)}{(\alpha\bar{X}+\beta)+A\alpha(\bar{x}-\bar{x}^{'})}$$

Taking expectation under the proposed sampling design P(s), we get

$$E(\hat{\bar{y}}_A^{'}) = E_s\left[\frac{\bar{y}_s(\alpha\bar{X}+\beta)}{(\alpha\bar{X}+\beta)+A\alpha(\bar{x}-\bar{x}^{'})}\right]$$

$$= \sum_{s=1}^{\binom{n^{'}}{n}}\left[\frac{\bar{y}_s(\alpha\bar{X}+\beta)}{(\alpha\bar{X}+\beta)+A\alpha(\bar{x}-\bar{x}^{'})}\right]P(S)$$

$$= \sum_{s=1}^{\binom{n^{'}}{n}}\left[\frac{\bar{y}_s(\alpha\bar{X}+\beta)}{(\alpha\bar{X}+\beta)+A\alpha(\bar{x}-\bar{x}^{'})}\right]\frac{\left(\alpha\bar{X}+\beta+A\alpha(\bar{x}-\bar{x}^{'})\right)}{(\alpha\bar{X}+\beta)\binom{n^{'}}{n}}$$

$$= \sum_{s=1}^{\binom{n^{'}}{n}}\bar{y}_s\frac{1}{\binom{n^{'}}{n}}$$

$= E(\bar{y}_s)$ where E(.) denotes the expectation under SRSWOR

$$= \bar{Y} \qquad (4)$$

Thus $\hat{\bar{y}}_A^{'}$ is an unbiased estimator of population mean under the proposed sampling strategy.

Theorem 2 The optimum variance of $\hat{\bar{y}}_A^{'}$ under the proposed sampling design P(s) equal to the variance of linear regression estimator

$$Var(\hat{\bar{y}}_A^{'})_{opt} = \left\{\left(\frac{1}{n}-\frac{1}{n^{'}}\right)S_y^2(1-\rho^2)+\frac{1}{n^{'}}S_y^2\right\} \qquad (5)$$

under the optimum values of constant

$$A_{opt} = \frac{S_{xy}}{\eta\bar{Y}S_x^2} \qquad (6)$$

3. DOMINANCE CONDITION OF PROPOSED SAMPLING STRATEGY

3.1 Comparison with Mean per Unit Estimator

$Var(\hat{\bar{y}}_A^{'})_{opt} < MSE(\bar{y})$, if $A < \frac{2C}{\eta}$ (7)

3.2 Comparison with Ratio Estimator

$Var(\hat{\bar{y}}_A^{'})_{opt} < MSE(\bar{y}_r)$, if $A\eta + 1 < 2C$ (8)

3.3 Comparison with Product Estimator

$Var(\hat{\bar{y}}_A^{'})_{opt} < MSE(\bar{y}_p)$, if $A\eta - 1 < 2C$ (9)

3.4 Comparison with Linear Regression Estimator

$$Var(\hat{\bar{y}}_A^{'})_{opt} = MSE(\bar{y}_{lr}) \qquad (10)$$

Both are equally efficient but the proposed sampling strategy is an unbiased estimator for all values of A whereas $MSE(\bar{y}_{lr})$ is biased estimator for population mean.

4. EMPIRICAL STUDY

The numerical illustration is performed over four natural populations and the details of the population is given below.

Population 1. (Choudhary F. S. and Singh D., Year 2009, Pg. n°. 107). The data concerns the number of boats landing and the catch of fish at a particular centre on a particular day.

y : catch of fish (quintals)
x : number of boats.

Population 2. (Choudhary F. S. and Singh D., Year 2009, Pg. n°. 176). The data concerns the number of cows in milk enumerated from tehsil.

y : number of cows in milk enumerated
x : number of cows in milk in the previous year.

Population 3. (Sukhatme P. V., Year 2002, Pg. n°. 51). To study fertilizer practices for different crops was carried out by the I.C.A.R. in Raipur district of Madhya Pradesh in 1958-59. The data concerns the cultivated area and area under rice of villages of Baloda Bazar tehsil of the district.

y : area under rice (acres)
x : total cultivated area of the village.

Population 4. (Choudhary F. S. and Singh D., Year 2009, Pg. n°. 141). The data concerns number of lime trees and the area reported under lime, in each of the 22 villages growing lime in one of the tehsils of the Nellore district.

y : number of bearing lime trees
x : area under lime (in acres).

The details of the data is given in Table 1.

Table 2, showed that the propose sampling strategy has lesser mean squared error than other conventional estimators. So, this research work has proven to bring better class of estimators for estimation of population variance.

Table 3, indicates that proposed sampling strategy has maximum gain in efficiency than conventional estimators. Hence, the suggested work has greater significance when compared to existing estimators.

5. GRAPHICHAL REPRESENTATION OF PRE OF VARIOUS ESTIMATORS

The graphical representation of percent relative efficiency of various estimator which is evaluated in Table 3.

Table 1. Parameters of the data

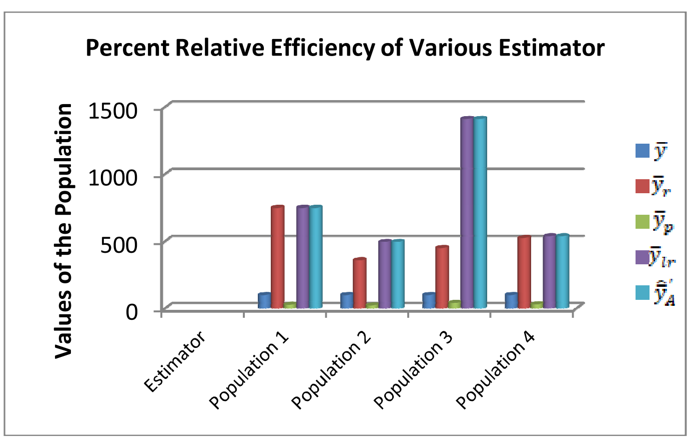
Parameters	Population 1	Population 2	Population 3	Population 4
n'	12	20	25	22
n	8	8	12	8
$\bar{X}$	27.33	641.05	927.36	22.62
$\bar{Y}$	550.33	816.45	735.8	1467.55
σ_x^2	429.59	562.21	785	2503.23
σ_y^2	19.92	517.11	568.97	32.29
ρ	0.93	0.89	0.96	0.90

Table 2. MSE of various estimators

Estimator	Population 1	Population 2	Population 3	Population 4
$\bar{y}$	21530.53	23706.29	26703.08	498447.4
$\bar{y}_r$	2876.25	6621.88	5929.65	94906.49
$\bar{y}_p$	77726.36	105852.8	65139.19	160009
$\bar{y}_{lr}$	2876.07	4788.30	1892.50	92776.46
$\hat{\bar{y}}_A'$	2876.07	4788.30	1892.50	92776.46

Table 3. PRE of various estimators

Estimator	Population 1	Population 2	Population 3	Population 4
$\bar{y}$	100	100	100	100
$\bar{y}_r$	748.56	358.00	450.33	525.19
$\bar{y}_p$	27.70	22.4	40.99	31.15
$\bar{y}_{lr}$	748.61	495.09	1410.99	537.26
$\hat{\bar{y}}_A'$	748.61	495.09	1410.99	537.26



Graph 1. The above graph depicts the percent relative efficiency of some conventional estimators and our proposed strategy. We have clearly seen that proposed sampling strategy has maximum gain in efficiency

6. CONCLUSION

As mentioned in the previous sections, we conclude that the proposed generalized class of sampling strategies provides an unbiased estimation of the population mean for all values of the characterizing scalar A under Midzuno-Sen type sampling scheme. Furthermore, the optimum variance of the proposed sampling strategy is found which is same as that of mean squared error of the biased linear regression estimator under simple random sampling without replacement. Hence, the proposed sampling strategy provides a better estimation of population mean in terms of unbiasedness, efficiency and much more practical utility.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Kumari C, Thakur RK. A generalized class of ratio type estimator for population variance under Midzuno-Sen type sampling scheme. Int. J. Eng. Res. Appli. 2020;10(4):55-60.
2. Kumari C, Thakur RK. A family of logarithmic estimators for population variance under double sampling. Int. J. Maths. Tre. Tech. 2020;66(4):99-105.
3. Kumari C, Thakur RK. Improved sampling strategies for finite population using auxiliary information. Int. J. Curr. Adv. Res. 2020;9(4):21932-21935.
4. Kumari C, Thakur RK. Improved ratio type estimators using auxiliary ATTRIBUTE For population variance. Int. J. Sci. Res. 2020: 9(4):1491-1505 .
5. Kumari C, Bhushan S, Thakur RK. Optimal two parameter logarithmic estimators for estimating the population variance. Glo. J. Pure and App. Mathe. 2019;15(5):527-237.
6. Bhushan S, Kumari C. An unbiased family of sampling strategies for estimation of population variance. Int. J. Stat. Sys. 2018; 13(1):41-52.
7. Kumari C, Bhushan S, Thakur RK. Modified ratio estimators using two auxiliary information for estimating population variance in two-phase sampling. Int. J. Sci. Eng. Res. 2018;9(8): 1884-1892.
8. Bhushan S, Kumari C. A class of double sampling log-type estimators for population Variance using two auxiliary variable. Int. J. App. Eng Res. 2018;13(13):11151-11155.
9. Lahiri DB. A method of sample selection providing unbiased ratio estimates. Bull. Int. Stat. Inst. 1951;3:133-140.
10. Midzuno H. On the sampling system with probability proportional to the sum of the sizes, Ann. Ints. Stat. Math. 1952;3:99-107.
11. Murthy MN. Sampling theory and method. Statistical Publishing Society, Calcutta, India; 1967.
12. Sen AR. Present status of probability sampling and its use in the estimation of characteristics. Econometrika. 1952;20-103.
13. Walsh JE. Generalization of ratio estimator for population total, Sankhya, A. 1970;32: 99-106.
14. Das AK, Tripathi TP. Use of auxiliary information in estimating the finite population variance. Sankhya. 1978; 40(C):139-148.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/57430