International Journal of Advance Research in Science and Engineering
Volume No. 11, Issue No. 08, August 2022
www.ijarse.com
IJARSE
ISSN 2319 - 8354

Estimator for estimating the finite population mean using auxiliary variable in Double Sampling Scheme

Kusum[1,a], Sangeeta Malik[1,b]

Department of Mathematics, Baba Mastnath University, Asthal Bohar,Rohtak,Haryana,India

Authors Emails

Corresponding author:sangeetastat@gmail.com

Author:kusumsansanwal9@gmail.com

ABSTRACT:

In this paper we have suggested a double sampling estimator for evaluation the population mean of the study variable. The presentation of suggested estimator corresponding to the mean per unit, ratio type estimator based on double sampling. The estimation of double sampling is generally affected under the excite that one of the samples is presented in the other. The first phase sample remaining secondary information that is virtually in wide to ascend since the second phase sample contain the adaptable of awareness. A simulated study has been made with usual estimators existing in the prose. An empirical study has also been carried out to determine the theoretical results.

Keywords: Double sampling, Bias, Mean square error, Auxiliary variable, study variable.

1. INTRODUCTION:

The prose on survey sampling designates a great variation of methods for using auxiliary information to attain more efficient estimators. The estimation of double sampling is generally performed under the impression that one of the sample is inserted within the other. The first phase sample present auxiliary information that is almost in extensive to derive since the second phase sample include the variable of interest. A comparative study has been made with usual estimators available in the literature. First, a random sample of size n' is drawn from a population of size N and again a random sample of size n is drawn from the first sample of size n'. So, the sample mean in this sampling is a function of the two-phase sampling is to furnish a good estimate of $\bar{X}$. This method is appropriate when the information about x_i is a file cards that have not been tabulated. Two stage sampling scheme is a particular kind of multi stage sampling where the units are chosen in two different stages. The auxiliary information may be utilize to enhance the precision of the estimates in two stage sampling also. There are several use of auxiliary information in increasing the precision of the estimators under two stage sampling scheme. Two-phase sampling is a powerful and cost-effective technique with a long history. Nyman (1938) was first to purpose it. S. Bhushan and C Kamari (2018) work on estimation of variance of finite population using double sampling scheme. C Kamari, S Bhushan, RK Thakur has studied on modified ratio estimators using two auxiliary information for estimating population variance in two-phase sampling.

If the population mean $\bar{X}$ is not known, than the double sampling technique is applied. Take a large initial sample of size n' by SRSWOR to estimate the population mean $\bar{X}$ as

$\bar{X} = \bar{X}' = \frac{1}{n'} \sum_{i=1}^{n} x_i$

Let us consider a finite population U= $(U_1, U_2, U_3, \dots \dots \dots \dots, U_N)$ $of\ size\ N\ and\ let\ Y_i$ be the observation on the study variable y for the i^{th} unit of the population (i=1, 2, 3, ... , N). Also

X_i and $Y_i$be the observation on the auxiliary variable.

2. PROPOSED ESTIMATOR:

$$\bar{y}_{kd1} = \bar{y} - \alpha\left(\frac{\bar{x}}{\bar{x'}}\right)^2 + \alpha \qquad (2.1)$$

$\bar{y} = \bar{Y}\,(1+e_0)$, $\quad \bar{x} = \bar{X}\,(1+e_1)$, $\quad \bar{x'} = \bar{X}\,(1+e_2)$

$=\bar{Y}\,(1+e_0) - \alpha\left(\frac{\bar{X}(1+e_1)}{\bar{X}\,(1+e_2)}\right)^2 + \alpha$

$=\bar{Y}\,(1+e_0) - \alpha\,[(1+2e_1)\,(1-2e_2)] + \alpha$

$E\,(\bar{y}_{kd1}) = \bar{Y}\,(1+e_0) - \alpha\,[2e_1 - 2e_2 - 4\,e_1\,e_2]$

$E\,(\bar{y}_{kd1}) = \bar{Y} + 4\alpha\, E(\,e_1\,e_2)$

$\text{Bias}\,(\bar{y}_{kd1}) = E\,(\bar{y}_{kd1}) - \bar{Y}$

$=4\alpha\left(\frac{1}{n'} - \frac{1}{N}\right)\frac{s_x^2}{\bar{X}^2}$

$\text{Var.}\,(\,\bar{y}_{kd1}\,) = E[\,\bar{y}_{k1d} - E\,(\,\bar{y}_{kd1})]^2$

$=E\left[\bar{y} - \alpha\left(\frac{\bar{x}}{\bar{x'}}\right)^2 + \alpha - \bar{Y} + 4\alpha e_1 e_2\,\right]^2$

$=E[\bar{Y}\,(1+e_0) - \alpha\,(1 - 2e_2 + 2e_1 - 4e_1e_2) + \alpha - \bar{Y} + 4\alpha e_1 e_2)]^2$

$=E\,[\bar{Y}^2 {e_0}^2 + 4\alpha^2 e_2^2 + 4\alpha^2 e_1^2 + 4\alpha e_0 e_1 \bar{Y} - 4\alpha e_0 e_2 \bar{Y} - 8\alpha^2 e_2 e_1]$

$$=\left(\frac{1}{n'} - \frac{1}{N}\right)s_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(s_y^2 + 4\alpha^2\,\frac{s_x^2}{\bar{X}^2} + 4\alpha\,\frac{s_{xy}}{\bar{X}}\right) \qquad (2.2)$$

3. EFFICIENCY COMPARISON:

3(a).Comparison of $\bar{y}_{kd1}$ with Mean per Unit Estimator:

$\text{Var.}\,(\bar{y}_{kd1}) - \text{var.}\,(\bar{y}) < 0$

This implies

$\left(\frac{1}{n'} - \frac{1}{N}\right)s_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(s_y^2 + 4\alpha^2\,\frac{s_x^2}{X^2} + 4\alpha\,\frac{s_{xy}}{\bar{X}}\right) < \left(\frac{1}{n'} - \frac{1}{N}\right)s_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)s_y^2$

i.e. $\rho < \alpha\frac{s_x}{s_y}\bar{X}$

3(b).comparison of $\bar{y}_{kd1}$ with Ratio Estimator:

$\text{Var.}\,(\bar{y}_{kd1}) - \text{Var.}\,(\bar{y}_r) < 0$

This implies

$\left(\frac{1}{n'} - \frac{1}{N}\right)s_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(s_y^2 + 4\alpha^2\,\frac{s_x^2}{X^2} + 4\alpha\,\frac{s_{xy}}{\bar{X}}\right) < \left(\frac{1}{n'} - \frac{1}{N}\right)s_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)(s_y^2 + R^2\,s^2 - 2R\rho s_x s_y)$

i.e. $\rho < \frac{1}{2}(R - 2\alpha)\frac{s_x}{s_y}\bar{X}$

3(c).Comparison of $\bar{y}_{kd1}$ with Product Estimator:

$\text{Var.}\,(\bar{y}_{kd1}) - \text{Var.}\,(\bar{y}_p) < 0$

This implies

$\left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)(s_y^2+4\alpha^2\,\frac{s_x^2}{\bar{X}^2}+4\alpha\,\frac{s_{xy}}{\bar{X}}) < \left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)(s_y^2+R^2\,s^2+\ \ 2R\rho s_x s_y)$

i.e. $\rho<\frac{3}{2}(R-2\alpha)\frac{s_x}{s_y}\bar{X}$

4. OPTIMUM VALUE OF PROPOSED ESTIMATOR:

Optimum value of α at which MSE in minimum is given by

$\alpha=\rho\frac{1}{2}\frac{s_y}{s_x}\bar{X}$

The minimum MSE of the estimator $\bar{y}_{kd1}$ is given by:

$\left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)(s_y^2(1-\rho^2)$

5. GENERALIZE FORM OF PROPOSED ESTIMATOR:

$\bar{y}_{kd2}=\bar{y}-\alpha\left(\frac{\bar{x}}{\bar{x}'}\right)^{\beta}+\alpha$ (5.1)

$\bar{y}=\bar{Y}\,(1+e_0)$, $\bar{x}=\bar{X}\,(1+e_1)$, $\bar{x'}=\bar{X}\,(1+e_2)$

$=\bar{Y}\,(1+e_0)-\alpha\left(\frac{\bar{X}(1+e_1)}{\bar{X}\,(1+e_2)}\right)^{\beta}+\alpha$

$=\bar{Y}\,(1+e_0)-\alpha\,[(1+\beta e_1)\,(1-\beta e_2)]+\alpha$

$E\,(\bar{y}_{kd2})=\bar{Y}\,(1+e_0)-\alpha\,[\beta e_1-\ \beta e_2-\beta^2\,e_1\,e_2]+\alpha$

$E\,(\bar{y}_{kd2})=\bar{Y}+\alpha\,\beta^2E(\,e_1\,e_2)$

$Bias\,(\bar{y}_{kd2})=E\,(\bar{y}_{kd2})-\bar{Y}$

$=\alpha\beta^2\left(\frac{1}{n'}-\frac{1}{N}\right)\frac{s_x^2}{\bar{X}^2}$

$Var.\,(\ \bar{y}_{kd2}\,)=E[\ \bar{y}_{kd2}-E\ \ (\bar{y}_{kd2})]^2$

$=E\left[\{\,\bar{y}\ -\alpha\left(\frac{\bar{x}}{\bar{X}}\right)^{\beta}+\alpha\}-(\bar{Y}+\alpha\beta^2 e_1 e_2\,)\right]^2$

$=E[\bar{Y}\,(1+e_0)-\alpha\,(1-\beta e_2+\beta e_1-\beta^2 e_1 e_2)+\alpha-(\bar{Y}\ +\alpha\beta^2 e_1 e_2\,)\,]^2$

$=E\,[\bar{Y}^2{e_0}^2+\alpha^2\beta^2 e_2^2+\alpha^2\beta^2 e_1^2+2\alpha\beta e_0 e_1\bar{Y}-2\alpha\beta e_0 e_2\bar{Y}-2\alpha^2\beta^2 e_2 e_1]$

$=\left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)(s_y^2+\alpha^2\,\beta^2\,\frac{s_x^2}{\bar{X}^2}+2\alpha\beta\,\frac{s_{xy}}{\bar{X}})$ (5.2)

6. EFFICINCY COMPARISON:

6(a).Comparison of $\bar{y}_{kd2}$ with Mean per Unit Estimator:

$Var.\,(\bar{y}_{kd2})-var.\,(\bar{y})<0$

This implies

$\left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)(s_y^2+\alpha^2\,\beta^2\,\frac{s_x^2}{\bar{X}^2}+2\alpha\beta\,\frac{s_{xy}}{\bar{X}}) < \left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)s_y^2$

i.e. $\rho<\frac{1}{2}\alpha\beta\,\frac{s_x}{s_y}\bar{X}$

6(b).Comparison of $\bar{y}_{kd2}$ with Ratio Estimator:

$Var.\,(\bar{y}_{kd2})-Var.\,(\bar{y}_r)<0$

This implies

$\left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)(s_y^2+\beta^2\alpha^2\,\frac{s_x^2}{\bar{X}^2}\text{-}2\alpha\beta\,\frac{s_{xy}}{\bar{X}})<\left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)(s_y^2+R^2\,s_x^2-\ \ 2R\rho s_x s_y)$

i.e. $\rho>\frac{1}{2}(\beta\alpha+R)\frac{s_x}{s_y}\bar{X}$

6(c).Comparison of $\bar{y}_{kd2}$ with Product Estimator:

Var. ($\bar{y}_{kd2}$) - Var. ($\bar{y}_p$) <0

This implies

$\left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)(s_y^2+\alpha^2\,\beta^2\,\frac{s_x^2}{\bar{X}^2}+2\alpha\beta\,\frac{s_{xy}}{\bar{X}})<\left(\frac{1}{n'}-\frac{1}{N}\right)s_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)(s_y^2+R^2\,s_x^2+\ \ 2R\rho s_x s_y)$

i.e. $\rho>-\frac{1}{2}(\beta\alpha)\frac{s_x}{s_y}\bar{X}$

7. EMPIRICAL STUDY:

To examine the proposed estimator we have consider some population data sets. The data are given below:

Population I: Kadilar and Cingi (2006, pp. 1047-1059)

Y: Level of apple production (1 unit=100 tones)

X: Number of apple trees (1 unit=100 trees)

The population parameters are:

$$\bar{Y}=15.37,\quad s_y^2=4127.626,\quad s_{yx}=25940.47,$$

$$\bar{X}=243.76,\quad s_x^2=242453,\quad N=106,\quad n'=80,\quad n=20$$

Population II: D.N. (2004, pp.433)

Y: Average miles per gallon

X: Top speed, miles per hours

The population parameters are:

$$\bar{Y}=33.83457,\quad s_y^2=997797.2,\quad s_{yx}=-9689.10755,$$

$$\bar{X}=112.4568,\quad s_x^2=196.9697,\quad N=81,\quad n'=21,\quad n=9$$

Table I: PRE of Proposed Estimators with Usual Estimator using Population I:

Proposed estimator	Constant value α	variance	Relative efficiency (Mean per unit)	Relative efficiency (Ratio)	Relative efficiency (Product)
$\bar{y}_{kd1}$	α =-.1	529.5247	316.2448	654.4158	182.5222
$\bar{y}_{kd1}$	α =.13	779.9847	465.8254	963.9481	268.8533
$\bar{y}_{kd2}$	α =.6,β = .-3	463.4585	276.7884	572.7675	159.7498
$\bar{y}_{kd2}$	α =.2,β=-.8	401.4733	239.7694	496.1629	138.3841

Table 2: PRE of Proposed Estimators with Usual Estimator using Population II:

Proposed estimator	Constant value α	variance	Relative efficiency (Mean per unit)	Relative efficiency (Ratio)	Relative efficiency (Product)
$\bar{y}_{kd1}$	α =8	101574.4	103.0711	102.6842	103.4597
$\bar{y}_{kd1}$	α =9.5	152931.8	155.1853	154.6028	155.7704
$\bar{y}_{kd1}$	α =12	105488	107.0432	106.6414	107.4468
$\bar{y}_{kd2}$	α =1,β = 1	98549.44	100.0016	99.62	100.3786
$\bar{y}_{kd2}$	α =1,β=-1	98571.32	99.64834	100.0238	100.409
$\bar{y}_{kd2}$	α =3.5,β=5.5	102971.5	104.4888	104.0966	104.8828

8. CONCLUSION:

- In population I, the proposed and generalized ratio estimators in double sampling better than the mean per unit, ratio estimators and product estimators.
- In population II, the proposed and generalized estimators in double sampling are better than the mean per unit and product estimators but not with ratio estimators.
- When α=1, β=-1 then generalized estimator is equally efficient with ratio and product estimators but mean per unit is not better.
- At last we conclude that in many practical situations, the proposed estimators can be shown better than the usual estimators for different value of constant variables.

9. REFERENCES:

1. Neyman, J. Journal of American Statistical Association, 33, 101-116 (1938).
2. Kadilar, C. and Cingi, H. InterStat, 1-11 (2006).
3. Bhal, S. and Tuteja, R.K. Infrm. And Optim. Sci., XII, 1, 159-163 (1991).
4. Gujarati, D.N. and Sangeetha. Tata McGrow-Hill (2007).
5. olkin, I. Biometrika, 45, 154-165 (1958).
6. Murthy, M.N., Sankhya, A, 26, 69-74 (1964).
7. Srivastava, S.K., Calcutta Statistical Association. Bulletin, 16, 121-132 (1970).
8. Sukhatme, B. V., Journal of American Statistical Associtіon, 57, 628-632 (1962).
9. Srivastava, S.K., Canadian Journal of Statistical Association, 66, 404-407 (1971).