ON CLASS OF DOUBLE SAMPLING EXPONENTIAL RATIO TYPE ESTIMATOR USING AUXILIARY INFORMATION ON AN ATTRIBUTE AND AN AUXILIARY VARIABLE

SHASHI BHUSHAN, SACHIN KUMAR YADAV AND S. K. PANDEY

Abstract: In this paper, we have proposed a class of double sampling exponential ratio type estimator using the auxiliary information in both the form attribute and variable. The proposed class of estimator utilizes the auxiliary information on means of the auxiliary variable and attribute available in the first phase sample. The bias and mean square error (MSE) of the proposed class of estimator is obtained. In the conclusion, it has been shown that the proposed class of double sampling estimator is better than the most commonly used double sampling estimators discussed in the literature. An empirical study is included for illustration.

Keywords: Auxiliary information, Double Sampling Exponential type Estimator, Bias, Mean Square Error

Introduction: It is well known to all that the auxiliary information is used to improve the precision of the estimator, for instance see for instance see Shabbir, Gupta and Ahmad (2008), Bhushan (2007), Bahl and Tuteja (1991), Bandyopadhyay (1980), Sukhatme (1962) to name a few. So in the case when auxiliary information is not known, double sampling strategy helps in improving the precision of the estimator. Many biased double sampling ratio type, double sampling ratio type and the biased double sampling estimator obtained through parametric combination of ratio type and the usual unbiased estimators are available for estimating the population mean. The use of an auxiliary variable and an attribute to improve the efficiency of the population mean has been discussed recently by Yadav and Bhushan (2013), Nirmala Sawan (2010), Naik and Gupta (1996), Shabbir and Gupta(2007) among others.

Consider the following notations

Y = Study variable

X = Auxiliary Variable

Φ = Auxiliary Attribute

N= Size of the population

n = Size of the sample

$\bar{Y}=\frac{1}{N}\sum_{i=1}^{N}Y_i$ = Population mean of study variable

$\bar{X}=\frac{1}{N}\sum_{i=1}^{N}X_i$ = Population mean of auxiliary variable

$P=\frac{1}{N}\sum_{i=1}^{N}\phi_i$ = Population mean of auxiliary attribute

$S_Y^2=\frac{1}{N}\sum_{i=1}^{N}(Y_i-\bar{Y})^2$ = Population variance of study variable

$S_X^2=\frac{1}{N}\sum_{i=1}^{N}(X_i-\bar{X})^2$ = Population variance of auxiliary variable

$S_\phi^2=\frac{1}{N}\sum_{i=1}^{N}(\phi_i-P)^2$ = Population variance of auxiliary attribute (1.1)

If the information about the auxiliary variable and attribute is not known then in double sampling scheme these auxiliary characteristics are replaced by the corresponding sample values comes from the large preliminary simple random sample of size n' drawn without replacement from a population of size N in the first phase. Also the characteristic of interest Y and the auxiliary characteristic X and ϕ are observed on the second phase sample of size n drawn from the first phase sample by simple random sample without replacement.

Let $\bar{x}'=\frac{1}{n'}\sum_{i=1}^{n'}X_i$ = First phase sample mean of auxiliary variable

$p'=\frac{1}{n'}\sum_{i=1}^{n'}\phi_i$ = First phase sample mean of auxiliary attribute

$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}Y_i$ = Second phase sub-sample mean value of study variable

$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}X_i$ = Second phase sub-sample mean value of auxiliary variable

$p=\frac{1}{n}\sum_{i=1}^{n}\phi_i$ = Second phase sub-sample value of auxiliary attribute (1.2)

Consider the following usual unbiased and biased double sampling ratio and exponential ratio type estimator of population mean developed in the past

(i) General estimator of population mean in case of SRSWOR

$\hat{\bar{Y}}_1=\bar{y}$ = Sample mean

with $MSE(\hat{\bar{Y}}_1) = f_n \bar{Y}^2 C_Y^2$ (1.3)

(ii) Double sampling ratio estimator using auxiliary variable

$$\hat{\bar{Y}}_2 = \bar{y}\frac{\bar{x}'}{\bar{x}}$$

with

$$MSE(\hat{\bar{Y}}_2) = \bar{Y}^2\left[f_n C_Y^2 + f_{nn'}(C_X^2 - 2\rho_{YX}C_y C_X)\right] \quad (1.4)$$

(iii) Naik and Gupta (1996) double sampling ratio estimator using auxiliary attribute

$$\hat{\bar{Y}}_3 = \bar{y}\frac{p'}{p}$$

with

$$MSE(\hat{\bar{Y}}_3) = \bar{Y}^2\left[f_n C_Y^2 + f_{nn'}(C_P^2 - 2\rho_{YP}C_y C_P)\right] \quad (1.5)$$

(iv) Bahl and Tuteja(1991) double sampling exponential ratio estimator using auxiliary variable

$$\hat{\bar{Y}}_4 = \bar{y}\exp\left(\frac{\bar{x}' - \bar{x}}{\bar{x}' + \bar{x}}\right)$$

with

$$MSE(\hat{\bar{Y}}_4) = \bar{Y}^2\left[f_n C_Y^2 + f_{nn'}(\frac{1}{4}C_X^2 - \rho_{YX}C_y C_X)\right] \quad (1.6)$$

(v) Following Bahl and Tuteja (1991), double sampling exponential estimator given by Nirmala Sawan(2010) using auxiliary attribute

$$\hat{\bar{Y}}_5 = \bar{y}\exp\left(\frac{p' - p}{p' + p}\right)$$

with

$$MSE(\hat{\bar{Y}}_5) = \bar{Y}^2\left[f_n C_Y^2 + f_{nn'}(\frac{1}{4}C_P^2 - \rho_{YP}C_y C_P)\right] \quad (1.7)$$

where $f_n = \left(\frac{1}{n} - \frac{1}{N}\right)$ and

$$f_{nn'} = \left(f_n - f_{n'}\right) = \left(\frac{1}{n} - \frac{1}{n'}\right).$$

Proposed Class of Double Sampling Estimators : Using the Bahl and Tuteja (1991) and Singh (1967) estimators, we propose a class of double sampling ratio type estimators assuming that the auxiliary population means and auxiliary population proportions are not known

$$\hat{\bar{Y}}_p = \bar{y}\left\{\exp\left(\frac{\bar{x}' - \bar{x}}{\bar{x}' + \bar{x}}\right)\right\}^{\alpha_1}\left\{\exp\left(\frac{p' - p}{p' + p}\right)\right\}^{\alpha_2} \quad (2.1)$$

It may be noted that the estimators given from (1.3), (1.6) and (1.7) are the special cases of the proposed study using no auxiliary variable, using one auxiliary variable and using one auxiliary attribute respectively.

In order to obtain the Bias and Mean square error (MSE), let us denote

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

with $E(e_0) = E(e_1) = E(e_2) = E(e_1') = E(e_2') = 0$ (2.2)

and the results given by Sukhatme and Sukhatme (1997)

$$E\left(e_0^2\right) = f_n C_Y^2$$
$$E\left(e_1^2\right) = f_n C_X^2,\ E\left(e_1'^2\right) = f_{n'} C_X^2$$
$$E\left(e_2^2\right) = f_n C_P^2,\ E\left(e_2'^2\right) = f_{n'} C_P^2$$
$$E\left(e_0 e_1\right) = f_n \rho_{YX} C_Y C_X,\ E\left(e_0 e_1'\right) = f_{n'} \rho_{YX} C_Y C_X$$
$$E\left(e_0 e_2\right) = f_n \rho_{YP} C_Y C_P,\ E\left(e_0 e_2'\right) = f_{n'} \rho_{YP} C_Y C_P$$
$$E\left(e_1 e_2\right) = f_n \rho_{XP} C_X C_P,$$
$$E\left(e_1' e_2\right) = E\left(e_1 e_2'\right) = E\left(e_1' e_2'\right) = f_{n'} \rho_{XP} C_X C_P \quad (2.3)$$

Substituting the values from (2.2) in (2.1) and on solving, we get

$$\hat{\bar{Y}}_p = \bar{Y}\begin{bmatrix} 1 + e_0 - \alpha_1\frac{(e_1 - e_1')}{2} + \alpha_1\frac{(e_1^2 - e_1'^2)}{4} \\ +\alpha_1^2\frac{(e_1^2 + e_1'^2 - 2e_1 e_1')}{8} - \alpha_2\frac{(e_2 - e_2')}{2} \\ +\alpha_2\frac{(e_2^2 - e_2'^2)}{4} + \alpha_2^2\frac{(e_2^2 + e_2'^2 - 2e_2' e_2)}{8} \\ -\alpha_1\frac{e_0(e_1 - e_1')}{2} - \alpha_2\frac{e_0(e_2 - e_2')}{2} \\ +\alpha_1\alpha_2\frac{(e_1 e_2 - e_1' e_2 - e_1 e_2' + e_1' e_2')}{4} \end{bmatrix} \quad (2.4)$$

Taking expectation and substituting the values from (2.3), we get

$$Bias\left(\hat{\bar{Y}}_p\right) = E\left(\hat{\bar{Y}}_p\right) - \bar{Y}$$

$$= (f_n - f_{n'})\bar{Y}\begin{bmatrix} \alpha_1\frac{C_X^2}{4} + \alpha_2\frac{C_P^2}{4} + \alpha_1^2\frac{C_X^2}{8} + \alpha_2^2\frac{C_P^2}{8} \\ -\alpha_1\frac{\rho_{XY}C_XC_Y}{2} - \alpha_2\frac{\rho_{PY}C_PC_Y}{2} \\ +\alpha_1\alpha_2\frac{\rho_{XP}C_XC_P}{4} \end{bmatrix}$$

$= f_{nn'}\bar{Y}A$ (Say) (2.5)

For mean square of error (MSE), taking (2.4) up to the first order of approximation, squaring and taking expectation both the sides, we get

$$MSE\left(\hat{\bar{Y}}_p\right) = E\left(\hat{\bar{Y}}_p - \bar{Y}\right)^2$$

$$= \bar{Y}^2\left[f_nC_Y^2 + f_{nn'}\begin{pmatrix} \alpha_1^2\frac{C_X^2}{4} + \alpha_2^2\frac{C_P^2}{4} + \alpha_1\alpha_2\frac{\rho_{XP}C_XC_P}{2} \\ -\alpha_1\rho_{YX}C_YC_X - \alpha_2\rho_{YP}C_YC_P \end{pmatrix}\right]$$

$= \bar{Y}^2\left[f_nC_Y^2 + f_{nn'}B\right]$ (Say) (2.6)

The minimum value of MSE is obtained if optimum values of α_1 and α_2 are

$$opt(\alpha_1) = \frac{2(\rho_{YX} - \rho_{YP}\rho_{PX})}{(1-\rho_{XP}^2)}\frac{C_Y}{C_X} \qquad (2.7)$$

$$opt(\alpha_2) = \frac{2(\rho_{YP} - \rho_{YX}\rho_{XP})}{(1-\rho_{XP}^2)}\frac{C_Y}{C_P} \qquad (2.8)$$

and the minimum mean square error under the optimizing values of the characteristic scalars is given by

$$\min MSE(\hat{\bar{Y}}_p) = \bar{Y}^2(f_n - f_{nn'}R_{Y.XP}^2)C_Y^2 = M(Say) \qquad (2.9)$$

Comparison with the prevalent estimators :

Consider the following estimator of the study variable

(i) Sample Mean in case of SRSWOR

$\hat{\bar{Y}}_1 = \bar{y}$ **Vs.** $\hat{\bar{Y}}_S^*$

From (1.3) and (2.9)

$$MSE(\hat{\bar{Y}}_1) - M = f_{nn'}R_{Y.XP}^2C_Y^2 \ge 0 \qquad (3.1)$$

(ii) Double sampling ratio estimator using auxiliary variable

$\hat{\bar{Y}}_2$ **Vs.** $\hat{\bar{Y}}_s^*$

From (1.4) and (2.9)

$$MSE(\hat{\bar{Y}}_2) - M = f_{nn'}\bar{Y}^2\left[(C_X - \rho_{YX}C_Y)^2 + (R_{Y.XP}^2 - \rho_{YX}^2)C_Y^2\right] \ge 0 \qquad (3.2)$$

as $R_{Y.XP}^2 \ge \rho_{YX}^2$

(iii) Double sampling ratio estimator using auxiliary attribute

$\hat{\bar{Y}}_3$ **Vs.** $\hat{\bar{Y}}_s^*$

From (1.4) and (2.9)

$$MSE(\hat{\bar{Y}}_3) - M = f_{nn'}\bar{Y}^2\left[(C_P - \rho_{YP}C_Y)^2 + (R_{Y.XP}^2 - \rho_{YP}^2)C_Y^2\right] \ge 0 \qquad (3.3)$$

since $R_{Y.XP}^2 \ge \rho_{YX}^2$

(iv) Bahl and Tuteja (1991) double sampling exponential ratio estimator using auxiliary attribute

$\hat{\bar{Y}}_4$ **Vs.** $\hat{\bar{Y}}_s^*$

From (1.5) and (2.9)

$$MSE(\hat{\bar{Y}}_4) - M = f_{nn'}\bar{Y}^2\left[(\frac{C_X}{2} - \rho_{YX}C_Y)^2 + (R_{Y.XP}^2 - \rho_{YX}^2)C_Y^2\right] \ge 0 \qquad (3.4)$$

Since $R_{Y.XP}^2 \ge \rho_{YX}^2$

(v) Following Bahl and Tuteja (1991), Proposed double sampling exponential ratio estimator using auxiliary attribute by Nirmal Sawan (2010)

$\hat{\bar{Y}}_5$ **Vs.** $\hat{\bar{Y}}_s^*$

From (1.6) and (2.9)

$$MSE(\hat{\bar{Y}}_5) - M = f_{nn'}\bar{Y}^2\left[(\frac{C_P}{2} - \rho_{YP}C_Y)^2 + (R_{Y.XP}^2 - \rho_{YP}^2)C_Y^2\right] \ge 0 \qquad (3.5)$$

Since $R_{Y.XP}^2 \ge \rho_{YP}^2$

Empirical Study : In order to study the performance of the proposed class of estimator we studied two populations and the results are reported below alongwith the details of the populations.

Population 1: [Source: William G. Cochran (1977), Page-34]

Y = Food Cost, X = Family Income, ϕ = Family of size more than 3

$\bar{Y}$ = 27.40, $\bar{X}$ = 72.55, P = 0.52, C_X = 0.146, C_Y = 0.369, C_P = 0.985, ρ_{YX} = 0.2521, ρ_{YP} = 0.388, ρ_{XP} = -

0.153, $R^2_{Y.XP}$ = 0.249879, n = 16, n' = 22, N = 33

Population II: [Source: Advance Data from Vital and Health Statistics, Number 347, October 7, 2004(CDC)]

Y = Height of the people, X = Weight of the people, ϕ = Sex of the people

$\bar{Y}$ = 140.18, $\bar{X}$ = 39.63, P = 0.50, C_Y = 0.191654, C_X = 0.482337, C_P = 1.014, ρ_{YX} = 0.973, ρ_{YP} = 0.07, ρ_{XP} = 0.073, $R^2_{Y.XP}$ = 0.94673, n = 12, n' = 22, N = 36.

Table 4.1: PRE of various estimators with respect to sample mean

Estimator	PRE	
	Population I	Population II
$\hat{\bar{Y}}_1$	100.00	100.00
$\hat{\bar{Y}}_2$	102.33	50.52
$\hat{\bar{Y}}_3$	27.21	5.11
$\hat{\bar{Y}}_4$	103.32	243.89
$\hat{\bar{Y}}_5$	71.70	18.12
$\hat{\bar{Y}}_s^*$	**115.25**	**282.09**

Conclusion : The comparative study of the proposed class of double sampling exponential ratio type estimators establishes its superiority in the sense of minimum mean square of error over the sample mean, double sampling ratio estimator, and double sampling exponential ratio estimator using auxiliary variable and attribute at one time under the estimated values of optimizing scalars

.

References

1. Bhushan.S (2007) "Some improved sampling strategies in finite population", unpublished PhD thesis submitted and accepted by Lucknow University.
2. Bahl, S. and Tuteja, R.K. (1991) : Ratio and Product type exponential estimator, Information and Optimization sciences, Vol.XII, I, 159-163.
3. Bandyopadhyay, S. (1980) : Improved ratio and product estimators. Sankhya, 42, 45-49.224
4. Cochran W. G. (1977): Sampling Techniques, John Wiley, New York
5. Naik, V.D. and Gupta, P.C. (1996): A note on estimation of mean with known population
6. Sukhatme, B. V. (1962):Some ratio-type estimators in two-phase sampling. Journal of the American Statistical Association, 57, 628-632.
7. Shabbir, J. and Gupta, S.(2007) : On estimating the finite population mean with known population proportion of an auxiliary variable. Pak. J. Statist., 23(1), 1-9.
8. Shabbir, J., Gupta, S. and Ahmad, S. Estimation of median in two phase sampling using two auxiliary variables, Communication in Statistics-Theory and Methods 37 (11), 1815–1822, 2008.
9. Sukhatme, P.V. and Sukhatme, B.V.(1997): Sampling theory of surveys with applications. Piyush Publications, India.
10. Bhushan S. and Yadav S. K. (2013): On Unbiased Class of Estimators for Population Mean Using Auxiliary Information on an Attribute and a Variable, *Proceedings of the VII International Symposium on Optimization and Statistics 2012*, 181 – 192, Department of Statistics and Operations Research, Aligarh Muslim University, India

* * *

[1]Department of Applied Statistics, Babasaheb Bhimrao Ambedkar University, Lucknow
shashi.py@gmail.com
[2]Department of Statistics, Lucknow University, Lucknow

.