DUAL TO RATIO AND PRODUCT TYPE EXPONENTIAL ESTIMATORS OF FINITE POPULATION MEAN IN DOUBLE SAMPLING FOR STRATIFICATION

Dr.Arpita Lakhre Assistant Professor, Department of Statistics, Govt.Adarsh Girls College, Sheopur (MP),
Dr.Rajesh Tailor Associate Professor, School of Studies in Statistics, Vikram University, Ujjain (MP)
Dr.Asha Mishra Assistant Professor, Pioneer Institute of Professional Studies, Indore (MP)

Abstract
For estimating the finite population mean, two exponential type estimators in double sampling for stratification have been suggested with their properties. The biases and mean squared errors of the suggested estimators are obtained upto the first degree of approximation. The suggested estimators have been compared with usual unbiased estimator of population mean, ratio and product estimators, dual to ratio and product type estimators and ratio and product type exponential estimators in double sampling for stratification. To judge the merits of the suggested estimators, an empirical study has been carried out.

Keywords : double sampling, two exponential type estimators in double sampling, dual to ratio and product type estimators, Bias, mean squared error.

1. Introduction

In Stratified random sampling, it is assumed that the strata weights as well as sampling frame are available in advance. But in many practical situations, strata weights may be available but sampling frame within strata may not be available. For example in households survey in a city, number of households in different colonies might be available but list of households may not available. In this type of situation, post stratification technique is applied. There might be a situation when strata weights are not be available or if available, strata weights are outdated and can't be used. This type of situation occurs during the households survey, when investigator does not have information about newly added households in different colonies. In the field of agriculture there might be a situation when researcher may not have information about newly aided field yielding the crop under study. This situation leads investigator to use double sampling for stratification. Theory of double sampling was developed by Neyman (1938).

Let us consider a finite population $U = \{U_1, U_2, U_3, ..., U_N\}$ of size N in which strata weight $\frac{N_h}{N}, \{h = 1, 2, 3, ..., L\}$ are unknown. In such situation, we use double sampling for stratification which is described below

(a) A first phase of sample S of size n' using simple random sampling without replacement is drawn and only auxiliary variate x is observed.

(b) The sample is stratified into L strata on the basis of observed variable x. Let n'_h denotes the number of units in h^{th} stratum $(h = 1, 2, 3, ..., L)$ such that $n' = \sum_{h=1}^{L} n'_h$.

(c) From each n'_h unit, a sample of size $n_h = v_h n'_h$ is drawn where $0 < v_h < 1$, where $\{h = 1, 2, 3, ..., L\}$, is the predetermined probability of selecting a sample of size n_h from each stratum of size n'_h and it constitutes a sample S' of size $n = \sum_{h=1}^{L} n_h$. In S' both study variate y and the auxiliary variate x are observed.

Let y and x be the study variate and the auxiliary variate respectively. Then we define

$n = \sum_{h=1}^{L} n_h$, $(n = n_1, n_2, n_3, ..., n_L)$: Size of the sample S',

$w_h = \frac{n'_h}{n'} : h^{th}$ Stratum weight in the second phase sample,

$\bar{Y}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{hi} : h^{th}$ Stratum mean for the study variate y,

$$\bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} x_{hi} : h^{th} \text{ Stratum mean for the auxiliary variate } x,$$

$$\bar{Y} = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{N_h} y_{hi} \ : \text{ Population mean of the study variate } y,$$

$$\bar{X} = \frac{1}{N}\sum_{h=1}^{L}\sum_{i=1}^{N_h} x_{hi} \ : \text{Population mean of the auxiliary variate } x,$$

$$S_y^2 = \frac{1}{N-1}\sum_{h=1}^{L}\sum_{i=1}^{N_h}\left(y_{hi} - \bar{Y}_h\right)^2 : \text{Population mean square of the study variate } y,$$

$$S_x^2 = \frac{1}{N-1}\sum_{h=1}^{L}\sum_{i=1}^{N_h}\left(x_{hi} - \bar{X}_h\right)^2 : \text{Population mean square of the auxiliary variate } x,$$

$$S_{yh}^2 = \frac{1}{N_h-1}\sum_{i=1}^{N_h}\left(y_{hi} - \bar{Y}_h\right)^2 : \ h^{th} \text{ Stratum population mean square of the study variate}$$

y,

$$S_{xh}^2 = \frac{1}{N_h-1}\sum_{i=1}^{N_h}\left(y_{hi} - \bar{X}_h\right)^2 : h^{th} \text{ Stratum population mean square of the auxiliary}$$

variate x,

$$\rho_{yxh} = \frac{S_{yxh}}{S_{yh}S_{xh}} : \text{Correlation coefficient between } y \text{ and } x \text{ in the } h^{th} \text{ stratum,}$$

$$\bar{y}_{ds} = \sum_{h=1}^{L} w_h\bar{y}_h : \text{Unbiased estimator of population mean } \bar{Y} \text{ at second phase or double}$$

sampling mean of the study variate y ,

$\bar{x}_{ds} = \sum\limits_{h=1}^{L} w_h \bar{x}_h$: Unbiased estimator of population mean $\overline{X}$, at second $\quad\quad$ phase

or double sampling mean of the auxiliary variate x ,

$\bar{y}_h = \dfrac{1}{n_h} \sum\limits_{i=1}^{n_h} \bar{y}_{hi}$:Mean of the second phase sample taken from h^{th} stratum for the study

variate y ,

$\bar{x}_h = \dfrac{1}{n_h} \sum\limits_{i=1}^{n_h} \bar{x}_{hi}$:Mean of the second phase sample taken from h^{th} stratum for the study

auxiliary x ,

$\bar{x}'_h = \dfrac{1}{n'_h} \sum\limits_{i=1}^{n_h} x_{hi}$:First phase sample mean of the hth stratum for the auxiliary variate x

,

$f = \dfrac{n'}{N}$: First phase sampling fraction.

Ige and Tripathi (1987) defined ratio estimator in double sampling for stratification as

$$d_R = \bar{y}_{ds} \left(\dfrac{\bar{x}'}{\bar{x}_{ds}} \right). \quad\quad\quad\quad\quad\quad (1.1)$$

Similarly, the product type estimator in double sampling for stratification can be defined as

$$d_P = \bar{y}_{ds} \left(\dfrac{\bar{z}_{ds}}{\bar{z}'} \right), \quad\quad\quad\quad\quad\quad (1.2)$$

where z is an auxiliary variate which is negatively correlated with the study variate y and notations $\bar{z}_{ds}$ and $\bar{z}'$ have their usual meanings.

Biases and mean squared errors of the estimators d_R and d_p upto the first degree of approximation are expressed as

$$B(d_R) = \frac{1}{\overline{X}}\left[\sum_{h=1}^{L}\frac{W_h}{n'}\left(\frac{1}{v_h}-1\right)\left\{R_1 S_{yh}^2 - S_{yxh}\right\}\right], \tag{1.3}$$

$$B(d_P) = \frac{1}{\overline{Z}}\left[\sum_{h=1}^{L}\frac{W_h}{n'}\left(\frac{1}{v_h}-1\right)S_{yzh}\right], \tag{1.4}$$

$$MSE(d_R) = S_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\left[S_{yh}^2 + R_1^2 S_{xh}^2 - 2R_1 S_{yxh}\right], \tag{1.5}$$

$$MSE(d_P) = S_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\left[S_{yh}^2 + R_2^2 S_{zh}^2 + 2R_2 S_{yxh}\right]. \tag{1.6}$$

$$f = \frac{n'}{N}.$$

Srivankataramana (1980) and Bondyopadhyay (1980) defined dual to classical ratio and product estimators in simple random sample without replacement by using transformation $x_i^* = \dfrac{NX - n\,x_i}{N-n}$ on auxiliary variate x as

$$\overline{y}_R^d = \overline{y}\left(\frac{\overline{x}^*}{\overline{X}}\right), \tag{1.7}$$

and

$$\overline{y}_P^d = \overline{y}\left(\frac{\overline{Z}}{\overline{z}^*}\right), \tag{1.8}$$

where $\bar{x}^* = \dfrac{N\bar{X} - n\,\bar{x}}{N-n}$ and $\bar{z}^* = \dfrac{N\bar{Z} - n\,\bar{z}}{N-n}$ are unbiased estimators of population

mean $\bar{X}$ and $\bar{Z}$ respectively. Here z is the auxiliary variate which is assumed to be negatively correlated with the study variate y.

Dual to Ige and Tripathi (1987) ratio and product type estimators are defined as

$$d_R^* = \bar{y}_{ds}\left(\frac{\bar{x}_{ds}^*}{\bar{x}'}\right),\qquad\qquad (1.9)$$

and

$$d_P^* = \bar{y}_{ds}\left(\frac{\bar{z}'}{\bar{z}_{ds}^*}\right).\qquad\qquad (1.10)$$

Bahl and Tuteja (1991) suggested ratio and product type exponential estimators of population mean $\bar{Y}$ in simple random sampling respectively as

$$t_{Re} = \bar{y}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right),\qquad\qquad (1.11)$$

and

$$t_{Pe} = \bar{y}\exp\left(\frac{\bar{z}-\bar{Z}}{\bar{z}+\bar{Z}}\right),\qquad\qquad (1.12)$$

where $\bar{X}$ and $\bar{Z}$ are known population mean of auxiliary variates x and z respectively.

Tailor et al. (2014) studied Bhal and Tuteja (1991) ratio and product type exponential estimators in double sampling for stratification as,

$$d_{\text{Re}} = \bar{y}_{ds}\,\exp\left(\frac{\bar{x}' - \bar{x}_{ds}}{\bar{x}' + \bar{x}_{ds}}\right), \tag{1.13}$$

and

$$d_{Pe} = \bar{y}_{ds}\,\exp\left(\frac{\bar{z}_{ds} - \bar{z}'}{\bar{z}_{ds} + \bar{z}'}\right), \tag{1.14}$$

where $\bar{y}_{ds}$ and $\bar{x}_{ds}$ are unbiased estimators of population mean $\bar{Y}$ of the study variate y and auxiliary variate x respectively.

To the first degree of approximation, mean squared error of d_{Re}, d_{Pe}, d_R^* and d_P^* are obtained as

$$MSE(d_R^*) = S_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h} - 1\right)\left[S_{yh}^2 + g^2 R_1^2 S_{yh}^2 - 2\,gR_1 S_{yxh}\right], \tag{1.15}$$

$$MSE(d_P^*) = S_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h} - 1\right)\left[S_{yh}^2 + g^2 R_2^2 S_{zh}^2 + 2gR_2 S_{yzh}\right], \tag{1.16}$$

$$MSE(d_{\text{Re}}) = S_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h} - 1\right)\left[S_{yh}^2 + \frac{1}{4}R_1^2 S_{xh}^2 - R_1 S_{yxh}\right], \tag{1.17}$$

and

$$MSE(d_{Pe}) = S_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h} - 1\right)\left[S_{yh}^2 + \frac{1}{4}R_2^2 S_{zh}^2 + R_2 S_{yzh}\right], \tag{1.18}$$

2. Suggested Dual Ratio Type Exponential Estimator

Srivenkataramana (1980) and Bondyopadhyay (1980) motivates to suggest dual to Tailor et al. (2014) ratio type exponential estimator in double sampling for stratification as,

$$d_{\text{Re}}^{*} = \bar{y}_{ds} \exp\left(\frac{\bar{x}_{ds}^{*} - \bar{x}'}{\bar{x}_{ds}^{*} + \bar{x}'}\right), \tag{2.1}$$

where $\bar{x}_{ds}^{*} = \dfrac{N\bar{x}' - n\,\bar{x}_{ds}}{N-n}$.

To obtain the bias and mean squared error of the suggested estimator d_{Re}^{*}, we write

$$\bar{y}_{ds} = \bar{Y}\left(1 + e_o\right) \;,\; \bar{x}_{ds} = \bar{X}\left(1 + e_1\right) \qquad \text{and} \quad \bar{x}' = \bar{X}(1 + e_1'),$$

such that $E(e_o) = E(e_1) = 0 = E(e_1') = 0$ and

$$E(e_0^2) = \frac{1}{\bar{Y}^2} V(\bar{y}_{ds}) = \frac{1}{\bar{Y}^2}\left[S_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\, S_{yh}^2\left(\frac{1}{v_h} - 1\right) \right],$$

$$E(e_1^2) = \frac{1}{\bar{X}^2} V(\bar{x}_{ds}) = \frac{1}{\bar{X}^2}\left[S_x^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\, S_{xh}^2\left(\frac{1}{v_h} - 1\right) \right],$$

$$E(e_1'^2) = \frac{1}{\bar{X}^2} V(\bar{x}_{ds}) = \frac{1}{\bar{X}^2} S_x^2\left(\frac{1-f}{n'}\right),$$

$$E(e_0 e_1) = \frac{1}{\bar{Y}\,\bar{X}} Cov(\bar{y}_{ds}, \bar{x}_{ds})$$

$$= \frac{1}{\bar{Y}\,\bar{X}}\left[\left(\frac{1-f}{n'}\right)S_{yxh} + \frac{1}{n'}\sum_{h=1}^{L} W_h\, S_{yxh}\left(\frac{1}{v_h} - 1\right) \right],$$

$$E(e_0 e_1') = \frac{1}{\bar{Y}\,\bar{X}}\left(\frac{1-f}{n'}\right)S_{yxh}, \quad \text{and}$$

$$E\left(e_1 e_1'\right) = \frac{1}{\overline{X}^2}\left(\frac{1-f}{n'}\right)S_x^2.$$

Suggested estimator d_{Re}^* can be defined in terms of e_is as

$$d_{\mathrm{Re}}^* = \overline{Y}(1+e_0)\exp\left(\frac{\overline{X}(1+e_0)-\overline{X}(1+e_1')}{\overline{X}(1+e_0)+\overline{X}(1+e_1')}\right)$$

Adopting the usual procedure for finding the bias and mean squared error, the bias and mean squared error of the suggested estimators d_{Re}^* upto the first degree of approximation are obtained as

$$B\left(d_{\mathrm{Re}}^*\right) = \frac{1}{8\overline{X}}\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\left[g^2 RS_{xh}^2 + 4g S_{yxh}\right] \tag{2.3}$$

and

$$MSE\left(d_{\mathrm{Re}}^*\right) = S_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\left[S_{yh}^2 + \frac{1}{4}R_1^2 g^2 S_{xh}^2 - g R_1 S_{yxh}\right]. \tag{2.4}$$

3. Efficiency Comparisons for Dual to Ratio Type Exponential Estimator d_{Re}^*

Variance of usual unbiased estimator $\overline{y}_{ds}$ in double sampling for stratification is given as

$$V\left(\overline{y}_{ds}\right) = S_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h S_{yh}^2\left(\frac{1}{v_h}-1\right). \tag{3.1}$$

Efficiency comparison for ratio type exponential estimator d_{Re}

Comparisons of (1.5), (1.15), (1.17) and (2.4) show that the suggested estimator d^*_{Re} would be more efficient than

(i) usual unbiased estimator $\bar{y}_{ds}$ if

$$R_1 < \frac{4B}{Ag},$$
(3.2)

(ii) Ige and Tripathi (1987) ratio type estimator d_R if

$$R_1 < \frac{4B}{A(g+2)},$$
(3.3)

(iii) dual to Ige and Tripathi (1987) ratio type estimator d^*_R if

$$R_1 > \frac{4B}{3Ag},$$
(3.4)

(iv) Tailor et al. (2014) ratio type exponential estimator d_{Re} if

$$R_1 < \frac{4B}{A(g+1)}.$$
(3.5)

4 Suggested Dual to Product Type Exponential Estimator

In case of negative correlation between the study variate y and auxiliary variate z, suggested dual to product type exponential estimator in double sampling for stratification i.e. dual to Tailor et al. (2014) product type exponential estimator is defined as

$$d^*_{Pe} = \bar{y}_{ds} \exp\left(\frac{\bar{z}' - \bar{z}^*_{ds}}{\bar{z}' + \bar{z}^*_{ds}}\right),$$
(4.1)

where $\bar{z}^*_{ds} = \dfrac{N\bar{z}' - n\bar{z}_{ds}}{N-n}$ is an unbiased estimator of population mean $\bar{Z}$.

Using the standard procedure, the bias and mean squared error of the suggested dual to product type exponential estimator are obtained as

$$B\left(d^*_{Pe}\right) = \frac{1}{8\bar{Z}}\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\left[3g^2R^2S^2_{zh}+4gR_2S_{yzh}\right], \qquad (4.2)$$

$$MSE\left(d^*_{Pe}\right) = S^2_y\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\left[S^2_{yh}+\frac{1}{4}R_2^{\,2}g^2S^2_{zh}+gR_2S_{yzh}\right] \quad (4.3)$$

5 Efficiency Comparisons for Product Type Exponential Estimator d_{Pe}

From (4.1.16), (4.1.18), (4.3.1) and (4.4.3), it is observed that the suggested product type exponential estimator d^*_{pe} would be more efficient than

(i) usual unbiased estimator $\bar{y}_{ds}$ if

$$R_2 < -\frac{4D}{gC}, \qquad (5.1)$$

(ii) Ige and Tripathi (1987) product type estimator d_p if

$$R_2 < -\frac{4D}{C(2+g)}, \qquad (5.2)$$

(iii) Ige and Tripathi (1987) dual to product type estimator d^*_P if

$$R_2 > -\frac{4D}{3Cg}, \qquad (5.3)$$

(iv) Tailor et al. (2014) product type exponential estimator d_{Pe} if

$$R_2 < -\frac{4D}{C(1+g)},$$
(5.4)

where

$$A = \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)S_{xh}^2 , \quad B = \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)S_{yxh} ,$$

$$C = \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)S_{zh}^2 \qquad D = \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)S_{yzh} ,$$

$$R_1 = \frac{\bar{Y}}{\bar{X}} \quad \text{and} \quad R_2 = \frac{\bar{Y}}{\bar{Z}}.$$

6. Empirical Study

To exhibit the performance of the suggested estimators in comparison to other estimators, a natural population data set is being considered. The description of the population is given below.

Population I [Source: Chouhan (2012)]

x: production in '000 Tons,

y: productivity (MT/Hectare) and

Z: area in '000 Hectare.

Constant	Stratum I	Stratum II
n_h	4	4
n_h'	7	7
N_h	10	10

$\bar{Y}_h$	1.70	3.67
$\bar{X}_h$	10.41	289.14
$\bar{Z}_h$	6.32	8.67
S_{yh}	0.50	1.41
S_{xh}	3.53	111.61
S_{zh}	1.19	10.82
S_{yxh}	1.60	144.87
S_{yzh}	-0.05	-7.04
S_{xzh}	1.38	92.02

Table .1 Percent relative Efficiencies of $\bar{y}_{ds}$, d_R, d_P, d_{Re}, d_{Pe}, d_R^*, d_P^*, d_{Re}^* **and** d_{Pe}^* **with respect to** $\bar{y}_{ds}$

Estimator	$\bar{y}_{ds}$	d_R	d_P	d_{Re}	d_{Pe}	d_R^*	d_P^*	d_{Re}^*	d_{Pe}^*
PRE	100.00	103.45	101.15	104.84	101.03	104.55	100.87	107.74	100.32

7. Conclusion

Section 3 provides the conditions under which suggested estimator d_{Re}^* has less mean squared error in comparison to usual unbiased estimator $\bar{y}_{ds}$, Ige and Tripathi (1987) ratio type estimator d_R , dual to Iga and Tripathi (1987) ratio type estimator d_R^* and Tailor et al. (2014) ratio type exponential estimator d_{Re}.

Section 5 provides the conditions under which suggested dual to product type exponential estimator d_{Pe}^* has less mean squared error in comparison to usual unbiased estimator $\bar{y}_{ds}$, Ige and Tripathi (1987) product type estimator d_P dual to Ige

and Tripathi (1987) product type estimator d_P^* and Tailor et al. (2014) product type exponential estimator d_{Pe} .

Table 1 exhibits that suggested estimator d_{Re}^* has maximum percent relative efficiency as compared to all other considered estimators. Hence it can be concluded that if conditions obtained in section 3 and 5 are satisfied suggested estimator may be used for the estimate of population mean. It is important to note that suggested estimator d_{Pe}^* has less percent relative efficiency as compared to Ige and Tripathi (1987) product type estimator d_P , dual to Ige and Tripathi (1987) product type estimator d_P^* as well as Tailor and Chouhan (2014) product type exponential estimator d_{Pe} . It may be due to not satisfying the conditions obtained in section 5.Thus there is need to explore another population data set for which conditions obtained in section 5 are satisfied.

References

Bahl, S. and Tuteja, R.K. (1991). Ratio and product type exponential estimators. Info. & Opti. Sci. XIII, 159-163.

Bandyopadhyay (1980). Improved ratio and product estimators. Sankhya, 42, C, 45-49.

Ige, A.F. and Tripathi. T.P. (1987). On double sampling for stratification and use of auxiliary information. J. Ind. Soc. Agri. Statist., 39, 191-201.

Neyman, J. (1938). Contribution to the theory of sampling human population. J. Amer. Statist. Assoc., 33, 101-116.

Srivenkataramana , T. (1980). A dual of ratio estimator in sample surveys. Biometrika, 67, 1, 199-204.

Tailor, R., Chouhan, S. and Kim, J.M. (2014). Ratio and product type exponential estimators of population mean in double sampling for stratification. Comm. Statist. Appl. & Meth. 21, 1, 1–9.