

An Effective Exponential Ratio and Product Type Estimator in Post-Stratification

Bharath Srinivasaiah¹, Shobhit Agrawal²

¹PhD Scholar, Department of Mathematics B.M.U Rohtak (Haryana), India.
bharath1995@gmail.com

²Professor, Department of Mathematics B.M.U Rohtak (Haryana), India.
shobhit@gmail.com

Corresponding Author

Abstract: *In this paper we present a unique estimator for the population parameter under post-stratification. By segmenting the population into homogeneous subgroups, post-stratification is a widely used strategy in survey sampling that increases the accuracy of estimators. In the post-stratification scenario, this paper addressed the issue of estimating the mean of a finite population. For post-stratification, better separate ratio and product exponential type estimators are proposed. This study's primary goal is to evaluate our suggested estimators' performance against that of current estimators. We perform a thorough simulation research to assess the precision and effectiveness of our estimator. The mean squared errors and biases of the proposed estimators are obtained to the first degree of approximation. Theoretical and practical researches have demonstrated that the proposed estimators are more efficient than other estimators that were taken into consideration.*

Keywords: Finite populations mean, post-stratification, bias, mean square error, percent relative efficiency.

1. Introduction

The problem of post-stratification was first discussed by Hansen et al. (1953). Ige and Tripathi (1989) studied the properties of classical ratio and product estimators of population mean in the case of post-stratification. Chouhan (2012) studied the Bahl and Tuteja (1991) estimators in the case of post-stratification. Many researchers including Kish (1965), Fuller (1966), Raj (1972), Holt and Smith (1979), Agrawal and Pandey (1993), Lone and Tailor (2014), Jatwa (2014), Lone and Tailor (2015), Tailor et al. (2015) contributed significantly to this area of research.

Bahl and Tuteja (1991) envisaged a ratio and a product type exponential estimator of population mean in simple random sampling. Following Srivenkataramana (1980) and Bondyopadhyay (1980), Lone and Tailor (2014, 2015) proposed dual to separate ratio and product type exponential estimators in the case of post-stratification.

Obtaining an estimator of a population parameter that is capable of taking into account the essential characteristics of the population should be one of the primary goals of any challenge involving estimating. If the population that is being studied is homogenous with regard to the feature that is being examined, then the method of simple random sampling will produce a homogeneous sample, and as a result, the mean of the sample will be a good estimate of the mean of the population. Therefore, if the population is consistent with regard to the quality that is being investigated, then the sample that is generated using a method of simple random sampling should offer a sample that is representative of the community. In addition, the variation of the sample mean is not only dependent on the

sample size and the sampling percentage, but it is also dependent on the variance of the population. We need to make use of a method of sampling that has the potential to lessen the degree to which the population is diverse so that we can improve the accuracy of our estimates. The use of stratified sampling as one of these sampling procedures is appropriate when the population being studied is diverse with regard to the trait that is being investigated. The fundamental concept that underpins the stratified sampling method is to.

- Divide the whole heterogeneous population into smaller groups or subpopulations, such that the sampling units are homogeneous with respect to the characteristic under study within the subpopulation and
- Heterogeneous with respect to the characteristic under study between/among the subpopulations. Such subpopulations are termed as strata.
- Treat each subpopulation as a separate population and draw a sample by SRS from each stratum.

Let us consider a finite population. $P = (P_1, P_2, \dots, P_N)$ of size N , is divided into M strata of size N_j ($j = 1, 2, \dots, M$). Let y be the study variate and x be auxiliary variates taking values y_{jk} , x_{jk} and z_{jk} ($j = 1, 2, \dots, M; k = 1, 2, \dots, N_j$), respectively, on k^{th} unit of the j^{th} stratum. Here x is positively correlated with study variate y while z is negatively correlated to study variates y . A sample of size n_j is drawn from each stratum which constitutes a sample of size $\sum_{j=1}^M n_j$.

$$\bar{z}_j = \frac{1}{N_j} \sum_{k=1}^{N_j} z_{jk} : j^{\text{th}} \text{ stratum mean of the study variate } z$$

$$\begin{aligned} \bar{Y}_j &= \frac{1}{N_j} \sum_{k=1}^{N_j} y_{jk}: j^{\text{th}} \text{ stratum mean of the study variate } y \\ \bar{X}_j &= \frac{1}{N_j} \sum_{k=1}^{N_j} x_{jk}: j^{\text{th}} \text{ stratum mean of the auxiliary variate } x \\ \bar{Y} &= \frac{1}{N} \sum_{j=1}^L \sum_{k=1}^{N_j} y_{jk} = \sum_{j=1}^L W_j \bar{Y}_j: \text{Population mean study variate } y \\ \bar{X} &= \frac{1}{N} \sum_{j=1}^L \sum_{k=1}^{N_j} x_{jk} = \sum_{j=1}^L W_j \bar{X}_j: \text{Population mean auxiliary variate } x \end{aligned}$$

$$\begin{aligned} \bar{Z} &= \frac{1}{N} \sum_{j=1}^L \sum_{k=1}^{N_j} z_{jk} \\ &= \sum_{j=1}^L W_j \bar{Z}_j: \text{Population mean of the auxiliary variate } z. \end{aligned}$$

In the case of post-stratification, the usual unbiased estimator of population mean \bar{Y} is defined as

$$\bar{Y}_{PS} = \sum_{j=1}^M W_j \bar{y}_j \quad (1.1)$$

Where

$W_j = \frac{N_j}{N}$ is the weight of the j^{th} stratum and $\bar{Y}_j = \frac{1}{n_j} \sum_{k=1}^{n_j} y_{jk}$ is the sample mean of n_j sample units that fall in the j^{th} stratum.

Using the results from Stephen (1945), the variance of \bar{Y}_{PS} to the first degree of approximation is obtained as

$$\text{Var}(\bar{Y}_{PS}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{j=1}^M W_j^2 S_{y_j}^2 \quad (1.2)$$

Where, $S_{y_j}^2 = \frac{1}{N_j - 1} \sum_{k=1}^{N_j} (y_{jk} - \bar{Y}_j)^2$.

Separate ratio and product type estimators of population mean \bar{Y} in the case of post-stratification are defined as

$$\hat{Y}_{PSR} = \sum_{j=1}^M W_j \bar{y}_j \left(\frac{\bar{X}_j}{\bar{x}_j}\right) \quad (1.3)$$

And

$$\hat{Y}_{PSP} = \sum_{j=1}^M W_j \bar{Y}_j \left(\frac{\bar{z}_j}{\bar{z}_j}\right) \quad (1.4)$$

Up to the first degree of approximation, biases and mean squared errors of the estimators \hat{Y}_{PSR} and \hat{Y}_{PSP} are obtained as

$$B(\hat{Y}_{PSR}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{j=1}^M \bar{Y}_j (C_{x_j}^2 - \rho_{yx_j} C_{x_j} C_{y_j}) \quad (1.5)$$

$$\begin{aligned} \text{MSE}(\hat{Y}_{PSR}) &= \left(\frac{1}{n} - \frac{1}{N}\right) \\ &\left[\sum_{j=1}^M W_j S_{y_h}^2 + \sum_{j=1}^M W_j R_{1j}^2 S_{x_j}^2 - 2 \sum_{j=1}^M W_j R_{1j} S_{y_{x_j}} \right] \quad (1.6) \end{aligned}$$

$$B(\hat{Y}_{PSP}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{j=1}^M \bar{Y}_j C_{y_j} C_{z_j} \rho_{yz_j} \quad (1.7)$$

And

$$\text{MSE}(\hat{Y}_{PC}) = \bar{Y}^2 \left[\sum_{j=1}^M W_j^2 \gamma_j (C_{y_j}^2 + C_{x_j}^2 + 2\rho_{yx_j} C_{x_j} C_{y_j}) \right] \quad (1.8)$$

2. Suggested combined ratio exponential type estimator

We suggest the improved combined ratio exponential type estimator for population mean \bar{Y} in the case of post-stratification as

$$\begin{aligned} \hat{Y}_{PTR}^{(\alpha)} &= \sum_{j=1}^M W_j \bar{y}_j \left(\frac{\bar{X}_j}{\bar{x}_j}\right)^\alpha \\ &\exp\left(\frac{\bar{X}_j - \bar{x}_j}{(\alpha + 1)\bar{X}_j + (1 - \alpha)\bar{x}_j}\right) \quad (2.1) \end{aligned}$$

Where, α is a real constant.

To obtain the bias and mean squared error of the suggested estimator $\hat{Y}_{PTR}^{(\alpha)}$, we write

$$\bar{y}_j = \bar{Y}_j (1 + e_{0j}), \bar{x}_j = \bar{X}_j (1 + e_{1j}) \text{ such that}$$

$$E(e_{0j}) = E(e_{1j}) = 0$$

$$E(e_{0j}^2) = \left(\frac{1}{nW_j} - \frac{1}{N_j}\right) C_{y_j}^2$$

$$E(e_{1j}^2) = \left(\frac{1}{nW_j} - \frac{1}{N_j}\right) C_{x_j}^2$$

$$E(e_{0j}e_{1j}) = \left(\frac{1}{nW_j} - \frac{1}{N_j}\right) \rho_{yx_j} C_{y_j} C_{x_j}$$

Expressing (2.1) in terms of e_{hj} ($h = 0, 1, 2$) and expanding the exponential function on the right-hand side, we get

$$\begin{aligned} \hat{Y}_{PTR}^{(\alpha)} &= \sum_{j=1}^M W_j \bar{Y}_j (1 + e_{0j}) \\ &(1 + e_{1j})^{-\alpha} \exp\left[-\frac{e_{1j}}{2} \left\{1 + \left(\frac{1-\alpha}{2}\right) e_{1j}\right\}^{-1}\right] \text{ And} \\ \hat{Y}_{PTR}^{(\alpha)} &= \sum_{j=1}^M W_j \bar{Y}_j \left[1 + e_{0j} - \left(\alpha + \frac{1}{2}\right) e_{1j} \right. \\ &\quad \left. + \frac{(4\alpha^2 + 6\alpha + 3)e_{1j}^2}{8} - \left(\alpha + \frac{1}{2}\right) e_{0j}e_{1j}\right] \\ (\hat{Y}_{PTR}^{(\alpha)} - \bar{Y}) &= \sum_{j=1}^M W_j \bar{Y}_j \left[e_{0j} - \left(\alpha + \frac{1}{2}\right) e_{1j} + \right. \\ &\quad \left. \frac{(4\alpha^2 + 6\alpha + 3)e_{1j}^2}{8} - \left(\alpha + \frac{1}{2}\right) e_{0j}e_{1j}\right] \quad (2.2) \end{aligned}$$

Now, taking expectation of both sides of (2.2), the bias of the suggested estimator $\hat{Y}_{PTR}^{(\alpha)}$ to the first degree of approximation is obtained as

$$B(\hat{Y}_{PTR}^{(\alpha)}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{j=1}^M \frac{1}{\bar{X}_j} \left[\frac{(4\alpha^2 + 6\alpha + 3)R_{1j}S_{xj}^2}{8} - \left(\alpha + \frac{1}{2}\right)S_{yxj} \right] \quad (2.3)$$

Squaring both sides of (2.2) and then taking expectation, we get the mean squared error of the suggested estimator $\hat{Y}_{PTR}^{(\alpha)}$ up to the first degree of approximation as

$$MSE(\hat{Y}_{PTR}^{(\alpha)}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{j=1}^M W_j \left(S_{yj}^2 + \left(\alpha + \frac{1}{2}\right)^2 R_{1j}^2 S_{xj}^2 - 2\left(\alpha + \frac{1}{2}\right)R_{1j}S_{yxj} \right) \quad (2.4)$$

Which is minimized for

$$\alpha = \left(\frac{S_{yxj}}{R_{1j}S_{xj}^2} - \frac{1}{2} \right) \quad (2.5)$$

Where $R_{1j} = \frac{\bar{Y}_j}{\bar{X}_j}$

Putting (2.5) in (2.4), we get the minimum mean squared error of the estimator $\hat{Y}_{PTR}^{(\alpha)}$ up to the first degree of approximation as

$$\min. MSE(\hat{Y}_{PTR}^{(\alpha)}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{j=1}^M W_j S_{yj}^2 (1 - \rho_j^2) \quad (2.6)$$

where $\rho_j = \frac{S_{yxj}}{S_{yj}S_{xj}}$.

3. Efficiency comparisons of the suggested improved ratio exponential type estimator $\hat{Y}_{PTR}^{(\alpha)}$ with \hat{Y}_{PS} and \hat{Y}_{RSP} .

From (1.2), (1.6) and (2.4), it is observed that the suggested estimator $\hat{Y}_{PTR}^{(\alpha)}$ would be more efficient than

(i) The usual unbiased estimator \hat{Y}_{PS} if

$$\sum_{j=1}^M R_{1j}W_j \left(R_{1j} \left(\alpha + \frac{1}{2}\right)^2 S_{xj}^2 - 2\left(\alpha + \frac{1}{2}\right)S_{yxj} \right) < 0 \quad (3.1)$$

(ii) The usual separate ratio estimator \hat{Y}_{RSP} if

$$\sum_{j=1}^M W_j \left(R_{1j}^2 S_{xj}^2 \left\{ \alpha_j^2 + \alpha_j - \frac{3}{4} \right\} - 2S_{yxh}R_{1h} \left\{ \alpha_j - \frac{1}{2} \right\} \right) < 0 \quad (3.2)$$

4. Improved combined product exponential type estimator

Improved combined product exponential type estimator for population mean \bar{Y} in the case of post-stratification is being suggested as

$$\hat{Y}_{PTP}^{(\beta)} = \sum_{j=1}^M W_j \bar{y}_j \left(\frac{\bar{X}_j}{\bar{x}_j} \right)^\beta \exp \left(\frac{\bar{z}_j - \bar{Z}_j}{(\beta + 1)\bar{X}_j + (1 - \beta)\bar{x}_j} \right) \quad (4.1)$$

Where, β is a real constant.

The estimator $\hat{Y}_{PTP}^{(\beta)}$ in (5.1) expressing in terms of e_{hj} ($h = 0,1,2$) and expanding the exponential function on the right-hand side, we get

$$\hat{Y}_{PTP}^{(\beta)} = \sum_{j=1}^M W_j \bar{y}_j \left[1 + e_{0j} + e_{1j}^\beta \exp \left[\frac{e_{1j}}{2} \left\{ 1 + \left(\frac{1 - \beta}{2}\right)e_{1j} \right\}^{-1} \right] \right] \quad (4.2)$$

And

$$\hat{Y}_{PTP}^{(\beta)} = \sum_{j=1}^M W_j \bar{y}_j \left[1 + e_{0j} - \left(\beta - \frac{1}{2}\right)e_{1j} + \frac{(4\beta^2 + 2\beta - 1)e_{1j}^2}{8} - \left(\beta - \frac{1}{2}\right)e_{0j}e_{1j} \right]$$

$$(\hat{Y}_{PTP}^{(\beta)} - \bar{Y}) = \sum_{j=1}^M W_j \bar{y}_j \left[e_{0j} - \left(\beta - \frac{1}{2}\right)e_{1j} + \frac{(4\beta^2 + 2\beta - 1)e_{1j}^2}{8} - \left(\beta - \frac{1}{2}\right)e_{0j}e_{1j} \right]$$

Using the standard procedure, the bias and mean squared error of the suggested estimator $\hat{Y}_{PTP}^{(\beta)}$ are obtained as

$$B(\hat{Y}_{PTP}^{(\beta)}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{j=1}^M \frac{1}{\bar{X}_j} \left[\frac{(4\beta^2 + 2\beta - 1)R_{2j}S_{xj}^2}{8} - \left(\beta - \frac{1}{2}\right)S_{yxj} \right] \quad (4.3)$$

And

$$MSE(\hat{Y}_{PTP}^{(\beta)}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{j=1}^M W_j \left(S_{yj}^2 + \left(\beta - \frac{1}{2}\right)^2 R_{2j}^2 C_{xj}^2 - 2\left(\beta - \frac{1}{2}\right)R_{2j}S_{yxj} \right) \quad (4.4)$$

Which is minimized for

$$\beta = \left(\frac{S_{yxj}}{R_{2j}S_{xj}^2} + \frac{1}{2} \right) \quad (4.5)$$

Putting (4.5) in (4.4), we get the minimum mean squared error of the estimator $\hat{Y}_{PTP}^{(\beta)}$ up to the first degree of approximation as

$$\min. MSE(\hat{Y}_{PTP}^{(\beta)}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{j=1}^M W_j S_{yj}^2 (1 - \rho_j^2) \quad (4.6)$$

where $\rho_j = \frac{S_{yxj}}{S_{yj}S_{xj}}$.

5. Efficiency comparisons of the suggested estimator $\hat{Y}_{STP}^{(\beta)}$ with \hat{Y}_{PS} and \hat{Y}_{PSP}

From (1.2), (1.8) and (4.4), it is concluded that the suggested estimator $\hat{Y}_{PTP}^{(\beta)}$ would be more efficient than

(i) The usual unbiased estimator \hat{Y}_{PS} if

$$\sum_{j=1}^M R_{1j}W_j \left(R_{1j} \left(\beta - \frac{1}{2}\right)^2 S_{xj}^2 - 2\left(\beta - \frac{1}{2}\right)S_{yxj} \right) < 0 \quad (5.1)$$

(ii) The usual separate product estimator \hat{Y}_{PSP} if

$$\sum_{j=1}^M W_j \left(R_{1j}^2 S_{xj}^2 \left\{ \beta^2 - \beta - \frac{3}{4} \right\} - 2S_{yxh} R_{1h} \left\{ \beta + \frac{1}{2} \right\} \right) < 0 \quad (5.2)$$

6. Empirical study

To judge the performance of the suggested estimators we are considering two natural population data sets, the descriptions of populations are given below:

Population I- [Source: Koyuncu & Kadilar (2009)]

Y: Numbers of Teachers.

X: Numbers of Students.

Table 6.1

Constant	Stratum	Constant	Stratum
N	923	S _{y1}	888.84
n	180	S _{y2}	644.922
N ₁	127	S _{y3}	1033.467
N ₂	117	S _{y4}	810.585
N ₃	103	S _{y5}	403.654
N ₄	170	S _{y6}	711.723
N ₅	205	S _{x1}	30486.75
N ₆	201	S _{x2}	15180.76
n ₁	31	S _{x3}	27549.7
n ₂	21	S _{x4}	18218.93
n ₃	29	S _{x5}	8497.776
n ₄	38	S _{x6}	23094.14
n ₅	22	S _{yx1}	25237154
n ₆	39	S _{yx2}	9747943
Y ₁	703.74	S _{yx3}	28294397
Y ₂	413	S _{yx4}	14523886
Y ₃	573.17	S _{yx5}	3393592
Y ₄	424.46	S _{yx6}	15864574
Y ₅	267.03	W ₁	0.1375948
Y ₆	393.84	W ₂	0.126761
X ₁	20804.59	W ₃	0.111593
X ₂	9211.79	W ₄	0.184182
X ₃	14309.30	W ₅	0.2221002
X ₄	9478.85	W ₆	0.217768
X ₅	5569.95	X̄	11440.49848
X ₆	12997.59	Ȳ	436.4330288
γ ₁	0.02438405	f ₁	0.2440945
γ ₂	0.03907204	f ₂	0.1794872
γ ₃	0.02477402	f ₃	0.2815534
γ ₄	0.02043344	f ₄	0.2235294
γ ₅	0.0405765	f ₅	0.1073171
γ ₆	0.0206659	f ₆	0.1940299

Table 6.2: Percent relative efficiencies of \bar{Y}_{PS} , \hat{Y}_{PSR} and $\hat{Y}_{PTR}^{(\infty)}$ with respect to \bar{Y}_{PS}

Estimators	Percent Relative Efficiency (PRE's)			
	α=0.37	α=0.38	α=0.4	α=0.45
\bar{Y}_{PS}	100.00	100.00	100.00	100.00
\hat{Y}_{PSR}	2368.02185	2368.02185	228.6985382	2368.02185
$\hat{Y}_{PTR}^{(\infty)}$	229.8513283	230.18506	230.6426	230.5483526

Table 6.3: Percent relative efficiencies of \bar{Y}_{PS} , \hat{Y}_{PSP} and $\hat{Y}_{PTP}^{(\beta)}$ with respect to \bar{Y}_{PS}

Estimators	Percent Relative Efficiency (PRE's)			
	β=0.1	β=0.5	β=1	β=1.5
\bar{Y}_{PS}	100.00	100.00	100.00	100.00
\hat{Y}_{PSP}	34.5401006	34.5401006	34.5401006	34.5401006
$\hat{Y}_{PTP}^{(\beta)}$	122.442413	179.819078	232.668191	159.336122

7. Conclusion

The circumstances under which the proposed estimators $\hat{Y}_{PTR}^{(\infty)}$ and $\hat{Y}_{PTP}^{(\beta)}$ tend to have lower mean squared errors than the standard unbiased estimator and separate ratio and product type estimators when compared to post-stratification are given in Sections 3 and 6. Table 6.1 presents the comparative efficiency of the suggested estimators compared to the distinct ratio and product type estimators, as well as the traditional unbiased estimator. Estimating the population mean when the conditions found in sections 3 and 6 are met is indicated to be done in practice using the recommended estimators.

References

- [1] Bahl, S., Tuteja, R. K. (1991). Ratio and product type exponential estimator. Inform. Optimiz. Sci. 12(1):159–163.
- [2] Chouhan, S., (2012). Improved estimation of parameters using auxiliary information in sample surveys. Ph. D. Thesis, Vikram University, Ujjain, M.P., India.
- [3] Cochran, W. G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce, Journal of Agricultural Science, 30, 262-275.
- [4] Cochran, William, G., Sampling Techniques, Third Edition, Wiley Eastern Limited, (1977).
- [5] Goldstion, H., Multilevel statistic model. 2nd edition. Edward Arnold, London, (1995).
- [6] HOLT, D., SMITH, T. M. F., (1979). Post-stratification. J. Roy. Statist. Soc., 142, A, 33-46.
- [7] Ige, A. F. and Tripathi, T. P. (1987). On double sampling for post-stratification and use of auxiliary information, Journal of the Indian Society of Agricultural Statistics, 39, 191-201.
- [8] Jatwa, N. K., (2014). Estimation of population parameters in presence of auxiliary information. Ph.D. Thesis, Vikram University, Ujjain, M.P. India.
- [9] Koyuncu, N. and Kadilar, C. (2009). Family of estimators of population mean using two auxiliary variable in stratified random sampling, Communication in Statistics - Theory and Methods, 38, 2398-2417.
- [10] Lone, H. A., Tailor, R., (2015). Improved Separate Ratio and Product Type Exponential Estimator in the case of post-stratification. STATISTICS IN TRANSITION new series, Spring 2015 16,1,53-64.
- [11] Malik, S. and K., modified unbiased estimator using auxiliary variables, 2597(1) (2022), <https://doi.org/10.1063/5.0115269>.
- [12] Murthy, M. N. (1967). Sampling Theory and Methods, Statistical Publishing Society, Calcutta.
- [13] RAJ, D., (1972). The design of sample surveys. McGraw Hill, New York.
- [14] Robson, D. S. (1957). Applications of multivariate polykays to the theory of unbiased radio-type estimation, Journal of the American Statistical Association, 52, 511–522.
- [15] Singh, M. P. (1967). Ratio cum product method of estimation. Metrika 12:34–42.

- [16] Stephan, F., (1945). The expected value and variance of the reciprocal and other negative powers of a positive Bernoullian variate. *Ann. Math. Statist.* 16, 50-61.
- [17] Tailor, R. and Lone, H. A. (2014). Radio-cum-product estimator of finite population mean in double sampling for post-stratification, *Journal of Reliability and Statistical Studies*, 7, 93-101.