A STRATIFIED EXPONENTIAL RATIO TYPE ESTIMATOR FOR ESTIMATING FINITE POPULATION MEAN IN DOUBLE SAMPLING

Namita Srivastava, Anupama Bhadauriya[1]
Department of Statistics, St. John's College Agra (UP India)

ABSTRACT
In this article, we proposed a stratified exponential ratio type estimator for estimating the finite population mean in double sampling. The properties of the proposed estimator are obtained and comparison is made some of the existing estimators. The proposed estimator is found to perform better than usual unbiased estimator, Ige Tripathi (1987), Tailor et al. (2014) estimators in double sampling for stratification. To judge the performance of the proposed estimator an empirical study has been carried out.

KEYWORDS: Auxiliary Variable, Stratified Sampling, Double Sampling, Bias, MSE.

MSC: 62D05

RESUMEN:
En este artículo, propusimos un estimador de tipo razón exponencial estratificada para estimar la media de la población finita en muestreo doble. Se obtienen las propiedades del estimador propuesto y se comparan algunos de los estimadores existentes. Se encontró que el estimador propuesto funciona mejor que los estimadores insesgados habituales, Ige Tripathi (1987), Tailor et al. (2014) en doble muestreo para estratificación. Para juzgar el desempeño del estimador propuesto se realizó un estudio empírico se ha llevado a cabo.

PALABRAS CLAVE: Variable auxiliar, muestreo estratificado, doble muestreo, sesgo, MSE.

1. INTRODUCTION

Use of auxiliary information in the estimation of population mean provides efficient estimators. Out of many, ratio, product and regression methods of estimation are good examples in this context. Cochran (1940) initiated the use of auxiliary information at estimation stage and proposed ratio estimator for population mean. It is well established fact that ratio type estimator provides better efficiency in comparison to simple mean estimator if the study variate and auxiliary variate are positively correlated.

A large amount of work has been carried out in estimating the population mean using simple random sampling (SRS) with or without replacement (WOR) scheme, for instance, see Singh (1986), Singh (2003) among other, Bahl and Tuteja (1991) pioneered ratio and type exponential estimators using an exponential function in simple random sampling. Usually, heterogeneous populations are encountered in practice. In such a situation, stratification is extensively used procedure in sample surveys to provide samples that are representatives of major sub-groups of a population. When the sampling frame within strata is known, stratified sampling is used, but there are many situations of practical importance where strata weights are known but a frame within the strata is not available. For example, in household survey in a city, number of households in different colonies may be available, but list of households may not be available. In such a situation post-stratification is used. However, in other situations with the passage of time, the stratum weights may not be available exactly as they become out-of-date. Further, the information on the stratification variable may not be readily available but could be made available by diverting a part of the survey budget to its collection.

This type of situation occurs during the household surveys, when the investigator does not have information about newly added households in different colonies. This situation leads investigator to use double sampling for stratification which was developed by Neyman (1938). For more studies on this topic the reader is referred to the papers by Rao (1973), Ige and Tripathi (1987, 1991), Singh and Vishwakarma (2007), Vishwakarma and Singh (2012), Tailor et al. (2014), Tailor and Lone (2014), Vishwakarma and Zeeshan (2018) and Singh and Nigam (2020 a, b),Clement, E. P. (2021).Recently, Singh, H.P., Nigam, P. (2022) worked for the estimation of finite population mean in Double Sampling for Stratification.

[1]Email Id: bhadauriyaanu@gmail.com

Following Tailor et al. (2014) and using transformation proposed by Bedi (1996), we proposed a stratified exponential ratio type estimator for estimating the finite population mean in double sampling.

2. PROCEDURE, NOTATIONS AND DEFINITIONS:

Consider a finite population $U = \{ U_1, U_2, \dots \dots U_N\}$ having$N$ distinct and identifiable unit partioned into L Strata. Let y and x be the study and auxiliary variate, respectively. Let $\bar{Y}$ be the population mean of the study variate$y$. suppose we want to estimate the population mean $\bar{Y}$ of y and consider it desirable to stratify the population on the basis of the values of an auxiliary variate x but the frequency distribution of xis unknown. Let the population of size N be stratified in to L strata of size N_h with strata weight $W_h = \frac{N_h}{N}$, $(h = 1,2, \dots . L)$ when strata weights are unknown, double sampling for stratification is used. It consists the following steps. [see Rao (1973) and Ige and Tripathi (1987)].
(a) A first phase sample S of size $n'$is drawn using simple random sampling without replacement and only auxiliary variate x is observed.
(b) The sample S is stratified in to L strata on the basis of auxiliary variable x. Let n'_hbe the number of units in h^{th} stratum $(h = 1,2..L)$ such that $\sum_{h=1}^{L} n'_h = n'$.
(c) From eachn'_h units, a sample of size $n_h = v_h n'_h$ is drawn where 0<v_h<1is the predetermined probability of selecting a sample of size n_h from strata of size n'_h and it constitutes a sample S of size $n = \sum_{h=1}^{L} n_h$. In sample S' both study variate y and auxiliary variate x are observed.
Let y and x be the study variate and auxiliary variate respectively. Then we define
$\sum_{h=1}^{L} n_h = n.$ Size of the sample S'
$W_h = \frac{N_h}{N}$is strata weight.
$\bar{Y}_h$=$\frac{1}{N_h}\sum_{i=1}^{N_h} y_{hi}$ 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} stratum mean for the auxiliary variate x.
$S^2{}_y = \frac{1}{N-1}\sum_{h=1}^{L}\sum_{i=1}^{N_h}(y_{hi} - \bar{Y}_h)^2$; Population mean square of the study variate y.
$S^2{}_x = \frac{1}{N-1}\sum_{h=1}^{L}\sum_{i=1}^{N_h}(x_{hi} - \bar{X}_h)^2$: Population mean square of the auxiliary variate$x$.
$S^2{}_{yh} = \frac{1}{N_h-1}\sum_{i=1}^{N_h}(y_{hi} - \bar{Y}_h)^2$: h^{th} Stratum mean square of the study variate y.
$S^2{}_{xh} = \frac{1}{N_h-1}\sum_{i=1}^{N_h}(x_{hi} - \bar{X}_h)^2$:$h^{th}$ Stratum mean square of the auxiliary variate x.
$S_{yxh} = \frac{1}{N_h-1}\sum_{i=1}^{N_h}(x_{hi} - \bar{X}_h)(y_{hi} - \bar{Y}_h)$ Covariance between y *and* x.
$C^2{}_{xh} = \frac{S^2{}_{xh}}{\bar{X}^2}$, $C^2{}_{yh} = \frac{S^2{}_{yh}}{\bar{Y}^2}$ is the coefficient of variation of h^{th} stratum.
$C_{hxy} = \rho_h C_{hy} C_{hx}, \rho_h$is the correlation coefficient between y *and* x
$\bar{y}_{ds} = \sum_{h=1}^{L} W_h \bar{y}_h$: Unbiased estimator of population mean $\bar{Y}$ in second phase or double sampling mean of the study variate y.
$\bar{x}_{ds} = \sum_{h=1}^{L} W_h \bar{x}_h$: Unbiased estimator of population mean $\bar{X}$ in second phase or double sampling mean of the auxiliary variate x.
$\bar{x}' = \sum_{h=1}^{L} W_h \bar{x}_h{}'$: Unbiased estimator of population mean $\bar{X}$ in first phase sampling mean of the auxiliary variate x.
$\bar{y}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} \bar{y}_{hi}$: Mean of the second phase sample taken from h^{th} stratum for the study variate$y$.
$\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} \bar{x}_{hi}$: Mean of the second phase sample taken from h^{th} stratum for the study variate x.
$\bar{x}_h{}' = \frac{1}{n_{h'}}\sum_{i=1}^{n_h{}'} \bar{x}_{hi}$: Mean of the first phase sample taken from h^{th} stratum x.
$f = \frac{n'}{N}$: First phase sampling fraction.
To obtain the bias and MSE of the suggested estimator, we define

$$\bar{y}_{ds} = \bar{Y}(1 + e_0), \ \bar{x}_{ds} = \bar{X}(1 + e_1) \text{ And } \bar{x}' = \bar{X}(1 + e_1{}')$$

$$E(e_0) = E(e_1) = E(e'_1) = 0$$

$$E(e_0^2) = \left[C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\left(\frac{1}{v_h} - 1\right)C^2{}_{yh}\right],$$

$$E(e_1^2) = \left[C_x^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\left(\frac{1}{v_h} - 1\right)C^2{}_{xh}\right],$$

$$E(e'^2_1) = \left[C_x^2\left(\frac{1-f}{n'}\right)\right],$$
$$E(e_0e_1) = \left[C_{xy}\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\left(\frac{1}{v_h} - 1\right)C_{xyh}\right],$$
$$E(e_0e'_1) = \left[C_{xy}\left(\frac{1-f}{n'}\right)\right],$$
$$E(e_1e'_1) = \left[C_x^2\left(\frac{1-f}{n'}\right)\right],$$

We have reviewed only those estimators from which our proposed estimator would be better.

3. REVIEW OF SOME EXISTING ESTIMATORS:

The usual unbiased estimator for population mean $\bar{Y}$ is defined by

$$\bar{y}_{ds} = \sum_{h=1}^{L} w_h\bar{y}_h \tag{3.1}$$

The variance /MSE of $\bar{y}_{ds}$ is given by

$$V(\bar{y}_{ds}) = MSE(\bar{y}_{ds}) = \left[C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\left(\frac{1}{v_h} - 1\right)C^2{}_{yh}\right] \tag{3.2}$$

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

$$d_R = \bar{y}_{ds}\frac{\bar{x}'}{\bar{x}_{ds}} \tag{3.3}$$

$$B(d_R) = \frac{1}{\bar{X}}\left[\sum_{h=1}^{L}\frac{W_h}{n'}\left(\frac{1}{v_h} - 1\right)\left\{R_1C_{yh}^2 - C_{yxh}\right\}\right] \tag{3.4}$$

$$MSE(d_R) = C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\left(\frac{1}{v_h} - 1\right)\left[C^2{}_{yh} + R_1^2C_{xh}^2 - 2R_1C_{yxh}\right] \tag{3.5}$$

Tailor et al. (2014) Estimator

$$d_{Re} = \bar{y}_{ds}exp\left[\frac{\bar{x}' - \bar{x}_{ds}}{\bar{x}' + \bar{x}_{ds}}\right] \tag{3.6}$$

$$B(d_{Re}) = \frac{1}{8\bar{X}}\left[3C_{yx}\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\left(\frac{1}{v_h} - 1\right)\left(3C_{xh}^2 - C_{yxh}\right)\right] \tag{3.7}$$

$$MSE(d_{Re}) = C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\left(\frac{1}{v_h} - 1\right)\left(C^2{}_{yh} + \frac{R_1^2}{4}C_{xh}^2 - \frac{1}{4}R_1C_{yxh}\right) \tag{3.8}$$

We have reviewed only those estimators from which our proposed estimator would be better.

4. PROPOSED ESTIMATOR:

Following Tailor et al. (2014) and using transformation proposed by Bedi (1996), We proposed a stratified exponential ratio type estimator for estimating the finite population mean in double sampling.

$$d_p = [k_1\bar{y}_{ds} + k_2(\bar{x}' - \bar{x}_{ds})]exp\left[\frac{\bar{x}' - \bar{x}_{ds}}{\bar{x}' + \bar{x}_{ds}}\right] \tag{4.1}$$

Where $k_1, k_2$are constant whose value are to be determined later. And expressing the proposed estimator d_p in terms of e's we have

$$d_p = [k_1\bar{y}_{ds}]exp\left[\frac{\bar{x}' - \bar{x}_{ds}}{\bar{x}' + \bar{x}_{ds}}\right] + [k_2((\bar{x}' - \bar{x}_{ds})]exp\left[\frac{\bar{x}' - \bar{x}_{ds}}{\bar{x}' + \bar{x}_{ds}}\right]$$

$$d_p = k_1\bar{Y}(1+e_0)exp\left[\frac{\bar{X}(1+e_1')-\bar{X}(1+e_1)}{\bar{X}(1+e_1')+\bar{X}(1+e_1)}\right] + k_2\left[\left((\bar{X}(1+e_1')-\bar{X}(1+e_1)\right)\right]exp\left[\frac{\bar{X}(1+e_1')-\bar{X}(1+e_1)}{\bar{X}(1+e_1')+\bar{X}(1+e_1)}\right]$$

$$d_p = k_1\bar{Y}(1+e_0)exp[(e_1'-e_1)(2+e_1'+e_1)^{-1}] + k_2\bar{X}(e_1'-e_1)exp[(e_1'-e_1)((2+e_1'+e_1)^{-1})]$$

$$d_p = [k_1\bar{Y}(1+e_0)+k_2\bar{X}(e_1'-e_1)]exp[(e_1'-e_1)(2+e_1'+e_1)^{-1}]$$

$$d_p = [k_1\bar{Y}(1+e_0)+k_2\bar{X}(e_1'-e_1)]\ exp\left[(e_1'-e_1)\frac{1}{2}\left(1+\frac{e_1'+e_1}{2}\right)^{-1}\right]$$

$$d_p = [k_1\bar{Y}(1+e_0)+k_2\bar{X}(e_1'-e_1)]\ exp\left[(e_1'-e_1)\frac{1}{2}\left(1-\frac{e_1'+e_1}{2}+\frac{(e_1'+e_1)^2}{4}+\cdots\right)\right]$$

$$d_p = [k_1\bar{Y}(1+e_0)+k_2\bar{X}(e_1'-e_1)]\ exp\left[\frac{(e_1'-e_1)}{2}-\frac{e_1'^2-e_1^2}{4}+\frac{(e_1'+e_1)^2(e_1'-e_1)}{8}+\cdots\right]$$

$$d_p = [k_1\bar{Y}(1+e_0)+k_2\bar{X}(e_1'-e_1)]\ exp\frac{(e_1'-e_1)}{2}\ exp\frac{(e_1'^2-e_1^2)}{4}\ exp\frac{(e_1'+e_1)^2(e_1'-e_1)}{8}$$

$$d_p = [k_1\bar{Y}(1+e_0)+k_2\bar{X}(e_1'-e_1)]\left[1+\frac{(e_1'-e_1)}{2}+\frac{(e_1'-e_1)^2}{8}+\cdots\right]\left[1+\frac{(e_1'^2-e_1^2)}{4}+\cdots\right]$$

$$d_p = [k_1\bar{Y}(1+e_0)+k_2\bar{X}(e_1'-e_1)]\left[1+\frac{(e_1'^2-e_1^2)}{4}+\frac{(e_1'-e_1)}{2}+\frac{(e_1'-e_1)^2}{8}+\cdots\right]$$

Keeping terms only up to order two in e's, we get

$$d_p-\bar{Y} = \bar{Y}\left[(k_1-1)+k_1\left(\frac{3e_1^2-e_1'^2-2e_1'e_1+4e_1'e_0-4e_1e_0}{8}+e_0+\frac{(e_1'-e_1)}{2}\right)+\cdots\right] \qquad (4.2)$$

$$+k_2\bar{X}\left[e_1'-e_1+\frac{e_1'^2+e_1^2-2e_1'e_1}{2}+\cdots\right]$$

Taking expectation both side of equation (4.2) and neglecting the higher order term

$$E(d_p-\bar{Y}) = \bar{Y}\ E\left[(k_1-1)+k_1\left(\frac{3e_1^2-e_1'^2-2e_1'e_1+4e_1'e_0-4e_1e_0}{8}+e_0+\frac{(e_1'-e_1)}{2}\right)+\cdots\right]$$

$$+k_2\bar{X}\left[e_1'-e_1+\frac{e_1'^2+e_1^2-2e_1'e_1}{2}+\cdots\right]$$

$$B(d_p) = \bar{Y}\left[(k_1-1)+k_1\left(\frac{3}{8n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2\right)-\frac{1}{2n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hyx}\right]+ \qquad (4.3)$$

$$+k_2\bar{X}\left[\frac{1}{2n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2\right]$$

On squaring both side of equation (4.2) and neglecting the higher order term. We take terms up to the first order of approximation because MSE can be obtained by taking the terms up to the first order of approximation.

$$(d_p-\bar{Y})^2 = \bar{Y}^2\left[\left[(k_1-1)+k_1\left(\frac{3e_1^2-e_1'^2-2e_1'e_1+4e_1'e_0-4e_1e_0}{8}+e_0+\frac{(e_1'-e_1)}{2}\right)+\cdots\right]+k_2\bar{X}\left[e_1'-e_1+\frac{e_1'^2+e_1^2-2e_1'e_1}{2}+\cdots\right]\right]^2$$

$$(d_p-\bar{Y})^2 = \left[\bar{Y}^2(k_1-1)^2+k_1^2\bar{Y}^2\left(e_0+\frac{(e_1'-e_1)}{2}\right)^2+k_2{}^2\bar{X}^2(e_1'^2+e_1{}^2-2e_1'e_1)\right.$$

$$+2\bar{Y}^2k_1(k_1-1)\left(\frac{3e_1^2-e_1'^2-2e_1'e_1+4e_1'e_0-4e_1e_0}{8}+e_0+\frac{(e_1'-e_1)}{2}\right)$$

$$\left.+2(k_1-1)k_2\bar{Y}\bar{X}\left(e_1'-e_1+\frac{e_1'^2+e_1{}^2-2e_1'e_1}{2}\right)+2k_1k_2\bar{Y}\bar{X}\left(e_0e_1'-e_0e_1+\frac{e_1'^2+e_1{}^2-2e_1'e_1}{2}\right)\right]$$

$$(d_p-\bar{Y})^2 = \left[(k_1-1)^2\bar{Y}^2+k_1{}^2\bar{Y}^2e_0{}^2+k_1{}^2\bar{Y}^2(e_1'^2+e_1{}^2-2e_1'e_1)+k_2{}^2\bar{X}^2(e_1'^2+e_1{}^2-2e_1'e_1)\right.$$

$$+2k_1{}^2\bar{Y}^2e_0e_1'-2k_1{}^2\bar{Y}^2e_0e_1+2k_1k_2\bar{X}\bar{Y}(e_1'^2+e_1{}^2-2e_1'e_1)+2k_1k_2\bar{X}\bar{Y}(e_0e_1'-e_1e_0)$$

$$\left.-k_1\bar{Y}^2\left(\frac{3}{4}e_1{}^2-\frac{e_1'^2}{4}-\frac{1}{2}e_1'e_1+e_0e_1'-e_0e_1+e_0+\frac{(e_1'-e_1)}{2}\right)\right]$$

$$E\left(d_p-\bar{Y}\right)^2 = E\left[(k_1-1)^2\bar{Y}^2 + k_1{}^2\bar{Y}^2 e_0{}^2 + k_1{}^2\bar{Y}^2\left(e_1'^2 + e_1{}^2 - 2e_1'e_1\right) + k_2{}^2\bar{X}^2\left(e_1'^2 + e_1{}^2 - 2e_1'e_1\right) + 2k_1{}^2\bar{Y}^2 e_0e_1' - 2k_1{}^2\bar{Y}^2 e_0e_1 + 2k_1k_2\bar{X}\bar{Y}\left(e_1'^2 + e_1{}^2 - 2e_1'e_1\right) + 2k_1k_2\bar{X}\bar{Y}\left(e_0e_1' - e_1e_0\right) - k_1\bar{Y}^2\left(\frac{3}{4}e_1{}^2 - \frac{e_1'^2}{4} - \frac{1}{2}e_1'e_1 + e_0e_1' - e_0e_1 + e_0 + \frac{(e_1' - e_1)}{2}\right)\right]$$

$$MSE\left(d_p\right) = \left[(k_1-1)^2\bar{Y}^2 + k_1{}^2\bar{Y}^2\left\{C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hy}^2\right\} + k_1{}^2\bar{Y}^2\left\{\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2\right\} + k_2{}^2\bar{X}^2\left\{\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2\right\} - 2k_1{}^2\bar{Y}^2\left\{\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hyx}\right\} + 2k_1k_2\bar{Y}\bar{X}\left\{\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2\right\} - 2k_1k_2\bar{Y}\bar{X}\left\{\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hyx}\right\}\right]$$

$$MSE\left(d_p\right) = k_1{}^2\bar{Y}^2\left[1 + C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hy}^2 + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2 - \frac{2}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hyx}\right] + \bar{Y}^2 - 2k_1\bar{Y}^2 + k_2{}^2\bar{X}^2\left[\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2\right] + 2k_1k_2\bar{Y}\bar{X}\left[\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2\right] - 2k_1k_2\bar{Y}\bar{X}\left[\frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hyx}\right] \quad (4.4)$$

Differentiate equation (4.4) w.r.to k_1, k_2 for obtaining the optimum value of k_1, k_2.

$$k_2 = k_1\left(\frac{\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hyx}}{\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2} - 1\right)\frac{\bar{Y}}{\bar{X}}$$

$$k_1 = \frac{1}{1 + C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hy}^2 - \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\frac{C_{hxy}^2}{C_{hx}^2}}$$

Substituting the value of k_1, k_2 in (4.4) we get minimum MSE of d_p is given by

$$MSE\left(d_p\right)_{min} = \bar{Y}^2\left[\frac{1}{1 + C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hy}^2 - \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\frac{C_{hxy}^2}{C_{hx}^2}} + 1 - \frac{2}{1 + C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hy}^2 - \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\frac{C_{hxy}^2}{C_{hx}^2}}\right]$$

Where $A = 1 + C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hy}^2 - \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)\frac{C_{hxy}^2}{C_{hx}^2}$

$$MSE\left(d_p\right)_{min} = \bar{Y}^2\left[1 - \frac{1}{A}\right] \quad (4.5)$$

5. THEORETICAL EFFICIENCY COMPARISON

In this section the proposed a stratified exponential ratio estimator in double sampling were compared theoretically with others existing estimator.

Let$B = C_y^2\left(\frac{1-f}{n'}\right)$, $C = \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C^2{}_{yh}$, $D = \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hx}^2$, $E = \frac{1}{n'}\sum_{h=1}^{L}W_h\left(\frac{1}{v_h}-1\right)C_{hyx}$

(1) The MSE of proposed ratio estimator (d_p) is better than sample mean per unit estimator$(\bar{y}_{ds})$ if

$$MSE(\bar{y}_{ds}) - MSE(d_p)_{min} > 0$$

$$\left[C_y^2\left(\frac{1-f}{n'}\right) + \frac{1}{n'}\sum_{h=1}^{L} W_h\left(\frac{1}{v_h} - 1\right)C^2{}_{yh}\right] - \bar{Y}^2\left[1 - \frac{1}{A}\right] > 0$$

$$B + C + \frac{1}{A} > 1 \qquad (5.1)$$

(2) The MSE of proposed estimator (d_p) is better than The MSE of Ige and Tripathi estimator (d_R) if $MSE(\mathrm{d_R}) - MSE(d_p)_{min} > 0$

$$B + C + R_1^2 D + \frac{1}{A} > 1 + 2R_1 E \qquad (5.2)$$

(3) The MSE of proposed estimator (d_p) is better than The MSE of Tailor et. al (2014) Estimator (d_{Re}) if

$$MSE(\mathrm{d_{Re}}) - MSE(d_p)_{min} > 0$$

$$B + C + \frac{R_1^2}{4}D + \frac{1}{A} > 1 + \frac{R_1}{4}E \qquad (5.3)$$

Expression (5.1), (5.2), (5.3) shows the condition in which proposed estimator performs better than usual unbiased estimator, Ige and Tripathi, Tailor et al. estimators.

6. EMPIRICAL STUDY

To judge the performance of the proposed estimator in comparison to other existing estimators, two population data sets are being considered. The parametric value of population is given below
Source of Data: Tailor et al. (2014)

Parametric value of population 1

x: Production in '000 Tons and y:Productivity (MT/ Hectare)

Stratum	n_h	n'_h	N_h	$\bar{Y}_h$	$\bar{X}_h$	s_{yh}	S_{xh}	S_{yxh}	S_y^2
1	2	4	10	1.70	10.41	0.50	3.53	1.61	2.21
2	2	4	10	3.67	289.14	1.41	111.61	144.88	

Parametric value of population 2

x:Area in Hectare, y:Productivity (MT/ Hectare)

Stratum	n_h	n'_h	N_h	$\bar{Y}_h$	$\bar{X}_h$	s_{yh}	S_{xh}	S_{yxh}	S_y^2
1	4	7	10	142.80	1632	6.09	102.17	-239.30	528.43
2	4	7	10	102.60	2036	12.60	103.46	-655.30	

Estimator	Population 1st		Population 2nd	
	MSE	PRE	MSE	PRE
$\bar{y}_{ds}$	0.3056	100	11.7195	100
d_R	0.3049	100.23	11.7168	100.02
d_{Re}	0.2998	104.37	11.7164	100.03
d_p	**0.1826**	**167.36**	**10.47**	**111.93**

Table 6.1: MSE, PRE of $(\bar{y}_{ds})$, (d_R), (d_{Re}), (d_p) with respect to $(\bar{y}_{ds})$

For comparison of different estimators, we calculate percent relative efficiency (PRE) of $(\bar{y}_{ds})$, $(d_R), (d_{Re}), (d_p)$ with respect to $(\bar{y}_{ds})$as

$$PRE(d_R, \bar{y}_{ds}) = \frac{V(\bar{y}_{ds})}{MSE\ (d_R)} * 100$$

$$PRE(d_{Re}, \bar{y}_{ds}) = \frac{V(\bar{y}_{ds})}{MSE\ (d_{Re})} * 100$$

$$PRE(d_P, \bar{y}_{ds}) = \frac{V(\bar{y}_{ds})}{MSE\ (d_P)} * 100$$

Table 6.1 gives the MSE, Percent Relative Efficiency of estimators $(\bar{y}_{ds})$, $(d_R), (d_{Re}), (d_p)$ with respect to usual unbiased estimator $(\bar{y}_{ds})$. it observe for Population 1st PRE of usual unbiased estimator $(\bar{y}_{ds})$ is 100 and PRE of Ige Tripathi estimator (d_R) is 100.23, PRE of Tailor et al. (d_{Re}),estimator is 104.37 and the PRE of proposed stratified exponential ratio type estimator in double sampling (d_p) is 167.36 which is highest. Also, in case of Population 2nd, the PRE of propose estimator is highest in comparison of other existing estimators.

7. SIMULATION STUDY

To see the performance of existing and proposed estimator in double sampling for stratification we have carried out a simulation study. We generate an artificial population of size $N = 1500$ and divided in to 5 strata. $N_h(h = 1,2 \ldots..5)$ as 100, 200, 300, 400, 500, we assume $X's$ distribution are chi square with 5 degree of freedom and Gamma distribution with shape parameter as 2.3 and scale parameter as 1. Five strata were formed according to increasing value of X.

The study variable Y generated using $Y_{hi} = \beta X_{hi} + \varepsilon_{hi,}$ $\ i = 1,2 \ldots. N_h$ where the random effects ε_{hi} generated from normal distribution with parameter $N(0,2)$, we choose $\beta = 2.5$. At first phase samples are selected by SRSWOR from each strata of sizes $\acute{n}_1 = 20, \acute{n}_2 = 40, \acute{n}_3 = 60, \acute{n}_4 = 80, \acute{n}_5 = 100$ and at second phase samples are selected in each strata from the samples drawn at first phase by SRSWOR of sizes $n_1 = 10, n_2 = 20, n_3 = 30, n_4 = 40, n_5 = 50$. The above experiment as repeated 500 times using R Software and the MSE and PRE of the estimators are computed through these iterations using the formula:

$$MSE(\hat{Y}_i) = \frac{\sum_{i=1}^{500}(\hat{Y}_i - \bar{Y})^2}{500},$$

where $\hat{Y}_i = \bar{y}_{ds},\ d_R, d_{Re}, d_P$

$$PRE(\hat{Y}_i) = \frac{MSE(\bar{y}_{ds})}{MSE(d_i)} \times 100,$$

where $d_i = d_R, d_{Re}, d_P$

Estimator	MSE	PRE
$\bar{y}_{ds}$	12.96	100
d_R	5.99	216.36
d_{Re}	9.25	140.10
d_P	5.23	247.80

Table 7.1. MSE, PRE of $(\bar{y}_{ds})$, $(d_R), (d_{Re}), (d_P)$ with respect to $(\bar{y}_{ds})$

Table (7.1) shows that PRE of proposed estimator is more efficient in comparison of other existing estimators.

8. CONCLUSION.

Section (5) provides the conditions under which the proposed estimator (d_p)performs better than usual unbiased estimator$(\bar{y}_{ds})$, Ige and Tripathi (d_R), Tailor et al. (d_{Re}) estimators.In Section 6, empirical study reveals that the proposed stratified exponential ratio type estimator(d_p) in double sampling has minimum MSE and maximum PRE in comparison the other existing estimators for population 1, 2.In Section 7, simulation study shows that the proposed ratio type exponential estimator (d_p) has minimum MSE and maximum PRE in comparison the other existing estimators for artificial population.

All comparison shows that the proposed estimator(d_p) is more efficient than usual unbiased estimator$(\bar{y}_{ds})$, Ige and Tripathi (d_R), Tailor et al.(d_{Re})estimators.

RECEIVED: APRIL, 2023.
REVISED: JUNE, 2024

REFERENCES

[1] BAHL, S. AND TUTEJA, R.K. (1991): Ratio and product type exponential estimator, **Journal of information and Optimization sciences,** 12, 159-163.

[2] BEDI, P.K. (1996): Efficient utilization of auxiliary information at estimating stage, **Biom. J.** 38(8), 973 -976.

[3] CHOUHAN, S. (2012): Improved Estimation of Parameters Using Auxiliary Information in Sample Surveys, **Unpublished Ph.D. Thesis,** Vikram University, Ujjain, M.P. India.

[4] CLEMENT E. P. (2018B): A Note on Calibration Weightings in Stratified Double Sampling with Equal Probability, **Communication in Statistics: Theory and Methods,** 47, 2835-2847

[5] CLEMENT,E.P. (2021): Ratio-product estimator in stratified double sampling based on coefficient of skewness of the auxiliary variable, **International Journal of Statistics and Applied Mathematics** ,6, 24-28

[6] COCHRAN, W.G. (1940): The estimation of yield of cereal experiments by sampling for the ratio of gain to total produce, **Journal of Agricultural Science**, 30, 262 – 275.

[7] DIANA, G. GIORDAN, M. PERRI P.F. (2011): An improved class of estimators for the population means **Statistical Methods and Applications,** 20, 123-140.

[8] GUPTA S, SHABBIR J. (2008): On the improvement in estimating the population mean in simple random sampling, **Journal of Applied Statistics,** 35, 559-566.

[9] HANSEN M.H., HURWITZ W.N. AND GURNEY, M. (1946): Problems and methods of the sample survey of business, **journal of the American Statistical Association,** 41, 173-189.

[10] IGE. A.F. AND TRIPATHI, T.P. (1987): On double sampling for stratification and use of auxiliary information**, journal of the Indian Society of Agricultural Statistics.** 39. 191 – 201.

[11] KADILAR, C. AND CINGI, H. (2003): Ratio estimators in stratified random sampling, **Biometrical Journal,** 45, 218- 225.

[12] KOYUNCU, N. (2017). Calibration estimator of population mean under stratified rank set sampling design, **communication in statistics- Theory and Methods,10**, 1-9

[13] MURTHY, M. N. (1967): Sampling Theory and Methods, **Statistical Publishing Society,** Calcutta, India

[14] NEYMAN, J. (1938): Contribution to the theory of sampling human populations, **Journal of the American Statistical Association,** 37, 101-116.

[15] RAO, J. N. K. (1973), On double sampling for stratification and analytical surveys, **Biometrika,** 6, 125-133.

[16] SINGH H.P. AND VISHWAKARMA, G.K. (2007): A general procedure for estimating the mean using Double sampling for stratification, **Model Assisted statistics and Application,** 2, 225- 237.

[17] SINGH, H. P. (1986): A generalized class of estimators of ratio, product and mean using supplementary information on an auxiliary character in PPSWR sampling scheme, **Gujarat Statistical Review**, 13, 1-30.

[18] SINGH, H. P. AND NIGAM, P. (2020A): Product-product-type (PPT) exponential estimator in double sampling for stratification (DSS) of finite population mean. **International Journal of Mathematics and Statistics,** 21, 50-59.

[19] SINGH, H. P. AND NIGAM, P. (2020B): Ratio-ratio-type exponential estimator of finite population mean in double sampling for stratification, **International Journal of Agricultural and Statistical Sciences,** 16, 251-257.

[20] SINGH, H.P. AND NIGAM, P. (2022): A General Class of Product-cum-Ratio-Type Exponential Estimators in Double Sampling for Stratification of Finite Population Mean, **Statistics and Application,** 20,165-179

[21] SINGH, H.P. AND VISHWAKARMA, G.K. (2006): An efficient variant of the product and ratio estimators in stratified random sampling, **Statistics in Transition**, 7. 1311 – 1325.

[22] SINGH, S. (2003):**Advance sampling theory and applications-How Michael 'Selected' Amy,** vol. I, Springer Sciences and Business Media, Berlin.

[23] TAILOR, R. AND LONE, H. A., (2014): Dual to separate ratio type exponential estimator in post stratification, **Journal of Statistics Application and Probability**, 3, 425-432.

[24] TAILOR, R. CHOUHAN, S. AND KIM, J. M. (2014): Ratio and product type exponential estimators of population mean in double sampling for stratification, **Communications for Statistical Applications and Methods**, 21, 1-9.

[25] VISHWAKARMA, G. K. AND SINGH, H. P. (2012): a general procedure for estimating the mean using double sampling for stratification and multi-auxiliary information, **Journal of Statistical Planning and Inference,** 142, 1252-1261.

[26] VISHWAKARMA, G. K. AND ZEESHAN, S. M. (2018): a ratio-cum-product type exponential estimator using double sampling for stratification, **Romanian Statistical Review,** 2, 3-13.