PROPORTION BASED EXPONENTIAL RATIO ESTIMATORS OF POPULATION MEAN IN CASE OF SINGLE AND DOUBLE SAMPLING PLANS

Sajad Hussain*†, Manish Sharma*, Hukum Chandra** and Vilayat Ali***

*Division of Statistics and Computer Science, FBSc, SKUAST-Jammu, Chatha-180009, India.

**ICAR-Indian Agricultural Statistics Research Institute (IASRI), New Delhi-110012, India.

***Department of Statistics, Pondicherry University, R V Nagar, Kalapet, Puducherry-605014, India.

ABSTRACT

This study proposes exponential ratio type estimators of population mean in case of simple random sampling without replacement ($SRSWOR$) when the auxiliary variable is an attribute. The double sampling version of these estimators is also proposed. Expressions for Bias and mean square error (MSE) have been evaluated. The proposed estimators were found to be more precise and also work as an alternative to different types of estimators as already existing. In order to support the theoretical findings, efficiency of the proposed estimators has also been analyzed numerically by using different population datasets.

KEYWORDS: Point bi-serial correlation, Bias, Auxiliary attribute, Exponential estimator, Mean square error.

MSC: 62D05

RESUMEN

Este trabajo propone un estimador tipo razón exponencial para la media de una población en el caso de muestreo simple aleatorio con remplazo en el que la variable auxiliar es un atributo. También se propone la versión de muestreo doble. Se evalúan las expresiones obtenidas para el sesgo y el error cuadrático medio. Se encontró que el estimador propuesto es más preciso. Además es una alternativa a otros estimadores existentes. Su eficiencia se analiza numéricamente usando diferentes bases de datos lo que soporta los resultados teóricos obtenidos.

PALABRAS CLAVE: Atributo auxiliar, Correlación bi-serial puntual, Error cuadrático medio, Estimador exponencial, Sesgo.

1. INTRODUCTION

In sampling theory the precision of an estimate can be increased at the estimation stage by using information of an instrumental variable called auxiliary variable which is correlated (highly) with the

†sajad.stat321@gmail.com

study variable. The method of utilizing auxiliary information depends upon the form in which it is available. When the auxiliary information is quantitative in nature, the estimators such as ratio estimator of Cochran [4], product type estimator of Robson [11], exponential type estimators of Bahl and Tuteja [3], the estimators of population mean proposed by Al-Omari and Bouza [2], Shalabh and Tsai [14], Singh and Vishwakarma [17], Mehta et.al [6] etc. can be used. There are many practical situations where the auxiliary information is qualitative in nature, in other words the auxiliary variable correlated with the study variable is an attribute. The authors Jhajj et al. [5] and Shabir & Gupta [12] have elaborated through the examples (a) the height of a person may depend on its sex (b) amount of milk produced by a cow may depend on its breed (c) the yield of a crop may depend on its variety. In all these situations the estimators where the auxiliary information is quantitative in nature cannot be used as there exists a point bi-serial correlation between the study and the auxiliary variable . Therefore Naik and Gupta [8], Jhajj et al. [5], Singh et.al [18], Shabir & Gupta ([12], [13]) and Abd-Elfattah et.al [1], have made some attempts in this direction and proposed the estimators of population mean by using prior knowledge of the parameters of auxiliary attribute. So taking a note of the above discussion the present study is carried with the following objectives:

- To propose proportion based exponential ratio type estimators under single and double sampling plans.
- To examine the efficiency of the proposed estimators empirically.

Use SRSWOR procedure for drawing a sample of size n from the population containing a total of N units. Let Y_i and ϕ_i denote the i^{th} unit $(i = 1, 2, \ldots, N)$ associated with the study variable and the auxiliary attribute respectively. Suppose there is complete dichotomy in the population with respect to the presence or absence of an attribute ϕ (say) and it is assumed that this attribute takes two values 0 and 1, as

$$\phi_i = \begin{cases} 1, & \text{if the } i^{th} \text{ unit of the population possesses the given attribute } \phi. \\ 0, & \text{otherwise.} \end{cases}$$

Consider $A = \sum_{i=1}^{N} \phi_i$ and $a = \sum_{i=1}^{n} \phi_i$, the number of units possessing the attribute ϕ in the population and sample respectively. Therefore $P = \frac{A}{N}$ and $p = \frac{a}{n}$ is the proportion of units possessing the given attribute ϕ in the population and sample respectively. Some formulas that have been used to compute various measures of the study are given below as

Population Estimates

$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} X_i$, is the mean of study variable.

$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$, is the mean square of study variable.

$S_\phi^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\phi_i - P)^2$, is the mean square of auxiliary variable.

Sample Estimates

: $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} x_i$, is the mean of study variable.

: $s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$, is the mean square of study variable.

: $s_\phi^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\phi_i - p)^2$, is the mean square of auxiliary variable.

$S_{y\phi} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i\phi_i - NP\bar{Y})^2$, is the covariance between study variable and auxiliary attribute.

: $s_{y\phi} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i\phi_i - np\bar{y})^2$, is the covariance between study variable and auxiliary attribute.

Further,
$C_y = \frac{S_y}{\bar{Y}}$ and $C_p = \frac{S_\phi}{P}$, is the coefficient of variation of Y and ϕ respectively.
$\rho_{pb} = \frac{S_{y\phi}}{S_y S_\phi}$, is the correlation between Y and ϕ .
$\hat{\beta} = \frac{s_{y\phi}}{s_\phi^2}$, is the sample regression coefficient.
$\gamma = \frac{1-f}{n}, \gamma_1 = \left(\frac{1}{\eta} - \frac{1}{N}\right), \gamma_2 = \left(\frac{1}{n} - \frac{1}{\eta}\right), \gamma_3 = \gamma + \gamma_1$, where $f = \frac{n}{N}$ is the sampling fraction.

2. SOME RATIO TYPE ESTIMATORS OF POPULATION MEAN

In order to have a rough idea about the population mean when there is no auxiliary information available, one can use the sample mean estimator given as

$$t_1 = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

The estimator t_1 is unbiased and its variance is as

$$V(t_1) = \gamma\bar{Y}^2{C_y}^2. \tag{2.1}$$

Naik and Gupta [8] was the pioneer to propose ratio estimator of population mean if the auxiliary variable is an attribute as

$$t_2 = \bar{y}\left[\frac{P}{p}\right].$$

The MSE of the estimator t_2 is as

$$MSE(t_2) = \gamma\bar{Y}^2({C_y}^2 + {C_p}^2 - 2C_{yp}). \tag{2.2}$$

Later Singh et al. [18] proposed exponential ratio type estimator of population mean if the auxiliary information available is an attribute as

$$t_3 = \bar{y}exp\left[\frac{P-p}{P+p}\right].$$

The MSE of estimator t_3 is as

$$MSE(t_3) = \gamma\bar{Y}^2\left({C_y}^2 + \frac{1}{4}{C_p}^2 - C_{yp}\right). \tag{2.3}$$

Abd-Elfattah et al. [1] used the information of auxiliary attribute (ϕ), such as point biserial correlation (ρ_{pb}), coefficient of variation (C_p), coefficient of kurtosis ($\beta_2(\phi)$) etc. and proposed following ratio type estimators of population mean as

$$t_4 = \bar{y}\left[\frac{P+\beta_2(\phi)}{p+\beta_2(\phi)}\right].$$

$$t_5 = \bar{y}\left[\frac{P+C_p}{p+C_p}\right].$$

$$t_6 = \bar{y}\left[\frac{P\beta_2(\phi)+C_p}{p\beta_2(\phi)+C_p}\right].$$

$$t_7 = \bar{y}\left[\frac{PC_p+\beta_2(\phi)}{pC_p+\beta_2(\phi)}\right].$$

$$t_8 = \bar{y}\left[\frac{P+\rho_{pb}}{p+\rho_{pb}}\right].$$

The MSE of estimators t_4 to t_8 are as

$$MSE(t_i) = \gamma\bar{Y}^2({C_y}^2 + {C_p}^2\delta_i^2 - 2\delta_i C_{yp}). \quad ; i = 4,5,6,7,8 \tag{2.4}$$

Where,

$$\delta_4 = \frac{P}{P+\beta_2(\phi)}, \delta_5 = \frac{P}{P+C_p}, \delta_6 = \frac{P\beta_2(\phi)}{P\beta_2(\phi)+C_p}, \delta_7 = \frac{PC_p}{\beta_2(\phi)+PC_p}, \delta_8 = \frac{P}{P+\rho_{pb}}.$$

3. PROPOSED ESTIMATORS

The proposed proportion based exponential ratio type estimators of population mean $\bar{Y}$ are as

$$t_{pr1} = \bar{y}exp\left[\frac{P-p}{SP}\right].$$

$$t_{pr2} = \bar{y}exp\left[\frac{P-p}{Rp}\right].$$

Where, S and R are suitable constants $(S \neq 0, R \neq 0)$.
For obtaining the theoretical expressions of Bias and MSE for t_{pr1} and t_{pr2}, let us define $e_y = \bar{Y}^{-1}(\bar{y}-\bar{Y})$ and $e_\phi = P^{-1}(p-P)$.
Under SRSWOR, the expected values of different quantities are obtained as $E(e_y) = E(e_\phi) = 0, E(e_y^2) = \gamma C_y^2, E(e_\phi^2) = \gamma C_p^2, E(e_y e_\phi) = \gamma C_{yp}$.
On rewriting the estimators t_{pr1} and t_{pr2} in terms of e_y and e_ϕ, the following expressions are obtained as

$$t_{pr1} = \bar{Y}(1+e_y)\exp\left(-S^{-1}e_\phi\right). \tag{3.1}$$

$$t_{pr2} = \bar{Y}(1+e_y)\exp\left(-R^{-1}e_\phi(1+e_\phi)^{-1}\right). \tag{3.2}$$

Solving the equations (5) & (6) and keeping the terms up to second degree of e'^{s}, the following expressions are obtained as

$$t_{pr1} = \bar{Y}\left(1 + e_y - \frac{e_\phi}{S} - \frac{e_y e_\phi}{S} + \frac{e_\phi^2}{2S^2}\right). \tag{3.3}$$

$$t_{pr2} = \bar{Y}\left(1 + e_y - \frac{e_\phi}{R} - \frac{e_y e_\phi}{R} + \frac{e_\phi^2}{R} + \frac{e_\phi^2}{2R^2}\right). \tag{3.4}$$

The Bias and MSE of the estimators t_{pr1} and t_{pr2} is obtained from equations (7) & (8) respectively as

$$Bias(t_{pr1}) = \gamma\bar{Y}\frac{1}{S}\left(\frac{1}{2S}C_p^2 - C_{yp}\right). \tag{3.5}$$

$$Bias(t_{pr2}) = \gamma\bar{Y}\frac{1}{R}\left(C_p^2 + \frac{1}{2R}C_p^2 - C_{yp}\right). \tag{3.6}$$

$$MSE(t_{pr1}) = \gamma\bar{Y}^2\left({C_y}^2 + \frac{1}{S^2}{C_p}^2 - \frac{2}{S}C_{yp}\right). \tag{3.7}$$

$$MSE(t_{pr2}) = \gamma\bar{Y}^2\left({C_y}^2 + \frac{1}{R^2}{C_p}^2 - \frac{2}{R}C_{yp}\right). \tag{3.8}$$

Special Cases: Some special cases have been derived from the estimators t_{pr1} and t_{pr2} and are as
Case1. The estimators t_{pr1} and t_{pr2} have the same MSE as that of the ratio type estimator proposed by Naik and Gupta [8] for S = R = 1.
Case2. The estimators t_{pr1} and t_{pr2} have the same MSE as that of the exponential ratio type estimator proposed by Singh et al. [18] for S = R = 2.
Case3. The estimators t_{pr1} and t_{pr2} have the same MSE as that of the product type estimator proposed by Naik and Gupta [8] for S = R = −1.
Case4. The estimators t_{pr1} and t_{pr2} have the same MSE as that of the exponential product type estimator proposed by Singh et al. [18] for S = R = −2.

Therefore the proposed exponential ratio estimators t_{pr1} and t_{pr2} can be used as an alternative to ratio and product type estimators of Naik & Gupta [8] and the exponential ratio and product type estimators of Singh et al. [18]. The optimum value of S and R is obtained from equations (11) and (12) by partial differentiation and was found to be $\frac{C_p}{\rho_{pb}C_y} = \lambda(say)$. Thus the Bias and MSE of t_{pr1} and t_{pr2} at this optimum value is as

$$Bias(t_{pr1}) = -\frac{1}{2}\gamma\bar{Y}\rho_{pb}^2 C_y^2. \tag{3.9}$$

$$Bias(t_{pr2}) = \gamma\bar{Y}\left(C_{yp} - \frac{1}{2}\rho_{pb}^2 C_y^2\right). \tag{3.10}$$

$$MSE_{min}(t_{prj}) = \gamma\bar{Y}^2{C_y}^2(1 - \rho_{pb}^2). \quad ; j = 1, 2. \tag{3.11}$$

That estimator t_{pr2} is unbiased, if

$$C_p = \frac{1}{2}\rho_{pb}C_y. \tag{3.12}$$

The expression (15) is same as that of the variance of $\bar{y}_{lr} = \bar{y} + \hat{\beta}(\bar{X} - \bar{x})$, that is the linear regression estimator. The value of λ is assumed to be known in advance, its value may either be obtained from the past experience or from the pilot survey; for instance see Murthy [7], Reddy ([9], [10]), Singh & Kumar [16] and Singh & Karpe [15]. If the value of λ is not known, it is advisable to replace λ by its estimate, $\hat{\lambda} = \hat{\beta}(\frac{P}{\bar{y}})$ (See Upadhyaya et al. [20]). Therefore the estimators t_{pr1} and t_{pr2} will take the form

$$\hat{t}_{pr1} = \bar{y}exp\left[\frac{\bar{y}(P-p)}{\hat{\beta}P^2}\right].$$

$$\hat{t}_{pr2} = \bar{y}exp\left[\frac{\bar{y}(P-p)}{\hat{\beta}Pp}\right].$$

The MSE of $\hat{t}_{pr1}$ and $\hat{t}_{pr2}$ to the first degree of approximation is found same as that of the MSE of t_{pr1} and t_{pr2} as given in equation (15).

4. EFFICIENCY COMPARISON

The procedure for the efficiency comparison of the proposed estimators t_{pr1} and t_{pr2}, is as follows

4.1. Efficiency comparison of t_{pr1} and t_{pr2}, when optimal value of S and R is taken

From equations (1), (2), (3), (4) and (15), we get the following conditions under which t_{pr1} and t_{pr2} are more efficient than the estimators t_1, t_2, t_3 and $t_i (i = 4, 5, 6, 7, 8)$ as

$$\begin{gathered} MSE_{min}(t_{prj}) < V(t_1) \\ \Rightarrow \gamma\bar{Y}^2{C_y}^2(1-\rho_{pb}^2) < \gamma\bar{Y}^2{C_y}^2, if \\ \rho_{pb}^2\bar{Y}^2 > 0. \end{gathered} \tag{4.1}$$

$$\begin{gathered} MSE_{min}(t_{prj}) < MSE(t_2) \\ \Rightarrow \gamma\bar{Y}^2{C_y}^2(1-\rho_{pb}^2) < \gamma\bar{Y}^2({C_y}^2 + {C_p}^2 - 2C_{yp}), if \\ (C_p - \rho_{pb}C_y)^2 > 0. \end{gathered} \tag{4.2}$$

$$\begin{gathered} MSE_{min}(t_{prj}) < MSE(t_3) \\ \Rightarrow \gamma\bar{Y}^2{C_y}^2(1-\rho_{pb}^2) < \gamma\bar{Y}^2\left({C_y}^2 + \frac{1}{4}{C_p}^2 - C_{yp}\right), if \\ (C_p - 2\rho_{pb}C_y)^2 > 0. \end{gathered} \tag{4.3}$$

$$\begin{gathered} MSE_{min}(t_{prj}) < MSE(t_i) \\ \Rightarrow \gamma\bar{Y}^2{C_y}^2(1-\rho_{pb}^2) < \gamma\bar{Y}^2({C_y}^2 + {C_p}^2\delta_i^2 - 2\delta_i C_{yp}), if \\ (\delta_i C_p - \rho_{pb}C_y)^2 > 0. \end{gathered} \tag{4.4}$$

Which is true in all the cases, therefore the proposed estimators are theoretically efficient.

4.2. Efficiency comparison of t_{pr1} and t_{pr2}, when different values of S and R are taken

From equations (1), (2), (3), (4) and (11), (12) the following conditions under which t_{pr1} and t_{pr2} are more efficient than the estimators t_1, t_2, t_3 and $t_i (i = 4, 5, 6, 7, 8)$ are obtained as

$$MSE_{min}(t_{pr1}) < V(t_1), if S > \left\{ \frac{C_p^2}{2C_{yp}} \right\}. \tag{4.5}$$

$$MSE_{min}(t_{pr2}) < V(t_1), if R > \left\{ \frac{C_p^2}{2C_{yp}} \right\}. \tag{4.6}$$

$$MSE_{min}(t_{pr1}) < MSE(t_2), if$$

$$min\left\{1, \frac{C_p^2}{2C_{yp} - C_p^2}\right\} < S < max\left\{1, \frac{C_p^2}{2C_{yp} - C_p^2}\right\}, \frac{C_{yp}}{C_p^2} > \frac{1}{2}. \tag{4.7}$$

$$Or\ S > 1,\ 0 \le \frac{C_{yp}}{C_p^2} \le \frac{1}{2}.$$

$$MSE_{min}(t_{pr2}) < MSE(t_2), if$$

$$min\left\{1, \frac{C_p^2}{2C_{yp} - C_p^2}\right\} < R < max\left\{1, \frac{C_p^2}{2C_{yp} - C_p^2}\right\}, \frac{C_{yp}}{C_p^2} > \frac{1}{2}. \tag{4.8}$$

$$Or\ R > 1,\ 0 \le \frac{C_{yp}}{C_p^2} \le \frac{1}{2}.$$

$$MSE_{min}(t_{pr1}) < MSE(t_3), if$$

$$min\left\{2, \frac{2C_p^2}{4C_{yp} - C_p^2}\right\} < S < max\left\{2, \frac{2C_p^2}{4C_{yp} - C_p^2}\right\}, \frac{C_{yp}}{C_p^2} > \frac{1}{4}. \tag{4.9}$$

$$Or\ S > 2,\ 0 \le \frac{C_{yp}}{C_p^2} \le \frac{1}{4}.$$

$$MSE_{min}(t_{pr2}) < MSE(t_3), if$$

$$min\left\{2, \frac{2C_p^2}{4C_{yp} - C_p^2}\right\} < R < max\left\{2, \frac{2C_p^2}{4C_{yp} - C_p^2}\right\}, \frac{C_{yp}}{C_p^2} > \frac{1}{4}. \tag{4.10}$$

$$Or\ R > 2,\ 0 \le \frac{C_{yp}}{C_p^2} \le \frac{1}{4}.$$

$$MSE_{min}(t_{pr1}) < MSE(t_i), if$$

$$min\left\{\frac{1}{\delta_i}, \frac{C_p^2}{2C_{yp} - \delta_i C_p^2}\right\} < S < max\left\{\frac{1}{\delta_i}, \frac{C_p^2}{2C_{yp} - \delta_i C_p^2}\right\}, \frac{C_{yp}}{C_p^2} > \frac{\delta_i}{2}. \tag{4.11}$$

$$Or\ S > \frac{1}{\delta_i},\ 0 \le \frac{C_{yp}}{C_p^2} \le \frac{\delta_i}{2}.$$

$$MSE_{min}(t_{pr2}) < MSE(t_i), if$$

$$min\left\{\frac{1}{\delta_i}, \frac{C_p^2}{2C_{yp} - \delta_i C_p^2}\right\} < R < max\left\{\frac{1}{\delta_i}, \frac{C_p^2}{2C_{yp} - \delta_i C_p^2}\right\}, \frac{C_{yp}}{C_p^2} > \frac{\delta_i}{2}. \tag{4.12}$$

$$Or\ R > \frac{1}{\delta_i},\ 0 \le \frac{C_{yp}}{C_p^2} \le \frac{\delta_i}{2}.$$

5. NUMERICAL STUDY

In order to get a rough idea about the gain in efficiency by t_{pr1} and t_{pr2}, an empirical study has been carried out. For that purpose data of two populations P1 and P2 has been taken from Sukhatme and Sukhatme [19] where the auxiliary information is an attribute. In the population P1, the study variable (Y) is the number of villages in the circles and the auxiliary information (ϕ) is a circle consisting of more than five villages. In the population P2, the study variable (Y) is the area under wheat crop in the circles (in acres) and the auxiliary attribute (ϕ) is a circle consisting of more than five villages. The requisite constants of the two populations are given below

Table 1: **Characteristics of population data sets.**

Population	N	n	$\bar{Y}$	P	ρ_{pb}	C_y	C_p	$\beta_2(\phi)$	C_{yp}
P1	89	23	3.36	0.124	0.766	0.60	2.678	6.162	1.230
P2	89	23	1102	0.124	0.624	0.65	2.678	6.162	1.086

The above data shows that from a population of size 89 units, a sample of size 23 is drawn. The population mean of study variable (Y) in P2 is very high as compared to P1. The correlation coefficient of P1 is greater than P2 and the population proportion of both the populations is known. Further the coefficient of variation of the study variable is higher in population P1 than in P2 and for the auxiliary attribute the cofficient of variation is same in both the populations .

Table 2: **Values of S & R for t_{pr1} & t_{pr2} to be more efficient than the estimator's t_1 to t_8**

Estimator	**Range of S & R for P1 and P2**			
	P1		**P2**	
t_1	S > 2.913	R > 2.913	S > 3.302	R > 3.302
t_2	S > 1.000	R > 1.000	S > 1.000	R > 1.000
t_3	S > 2.000	R > 2.000	S > 2.000	R > 2.000
t_4	S$\in$(3.093, 50.694)	R$\in$(3.093, 50.694)	S$\in$(3.532, 50.694)	R$\in$(3.532, 50.694)
t_5	S$\in$(3.347, 22.597)	R$\in$(3.347, 22.597)	S$\in$(3.867, 22.596)	R$\in$(3.867, 22.596)
t_6	S$\in$(4.505, 8.262)	R$\in$(4.505, 8.262)	S$\in$(12.365, 4.505)	R$\in$(12.365, 4.505)
t_7	S$\in$(3.426, 19.556)	R$\in$(3.426, 19.556)	S$\in$(3.973, 19.556)	R$\in$(3.973, 19.556)
t_8	S$\in$(4.909, 7.177)	R$\in$(4.909, 7.177)	S$\in$(6.032, 7.295)	R$\in$(6.032, 7.295)
Optimum value (λ)	5.827		6.603	

Table 3: **MSE and PRE of various estimators.**

Estimator	Population			
	P1		P2	
	MSE	PRE	MSE	PRE
t_1	2.611	100.000	280808.610	100.000
t_2	1.846	141.441	212291.210	132.275
t_3	0.336	777.083	44214.980	635.098
t_4	0.114	2290.351	14974.400	1875.258
t_5	0.097	2691.735	13328.730	2106.792
t_6	0.061	4280.328	11498.370	2442.160
t_7	0.092	2838.043	12927.770	2172.135
t_8	0.057	4580.702	10159.170	2764.090
t_{prj}	0.054	4835.185	10101.590	2779.846

The MSE and percent relative efficiency (PRE) of the estimators t_1 to t_8 and t_{pr1} & t_{pr2} have been obtained and given in the Table-3. It is found that the proposed estimators t_{pr1} & t_{pr2} have the lowest MSE and thus the highest PRE among all other estimators considered. The results of PRE are also shown graphically as

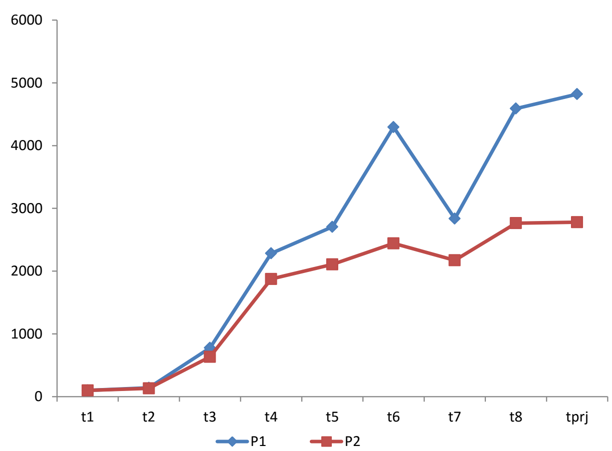

Figure 1: PRE of various estimators.

6. DOUBLE SAMPLING

For estimating the population mean using the estimators t_{pr1} and t_{pr2}, the value of population proportion (P) must be known in advance. Sometimes a researcher many face a situation where the

value of P is not known in advance, under such situations the method of double sampling is used in which a part of the budget is used to collect information of the auxiliary variable. Collect a large first phase preliminary sample of size η, suppose p_1 is the proportion of units possessing the attribute ϕ in this phase. Draw a second phase (*small*) sample of size n which is nested within the first phase sample ($n < \eta$). Let p is the proportion of units possessing the attribute ϕ and $\bar{y}$ be the mean of study variable (Y) in this second phase sample.
Naik and Gupta [8] proposed the double sampling ratio type estimator of population mean as

$$v_1 = \bar{y}\left[\frac{p_1}{p}\right].$$

The MSE of the estimator v_1 is as

$$MSE(v_1) = \bar{Y}^2(\gamma {C_y}^2 + \gamma_2 {C_p}^2 - 2\gamma_2 C_{yp}). \tag{6.1}$$

The double sampling exponential ratio type estimator of population mean proposed by Singh et al. [18] is as

$$v_2 = \bar{y}exp\left[\frac{p_1 - p}{p_1 + p}\right].$$

With the MSE as

$$MSE(v_2) = \bar{Y}^2\left(\gamma {C_y}^2 + \frac{1}{4}\gamma_2 {C_p}^2 - \gamma_2 C_{yp}\right). \tag{6.2}$$

7. PROPOSED ESTIMATORS IN DOUBLE SAMPLING

The proposed proportion based double sampling exponential ratio estimators are as

$$v_{pr1} = \bar{y}exp\left[\frac{p_1 - p}{Mp_1}\right].$$

$$v_{pr2} = \bar{y}exp\left[\frac{p_1 - p}{Np}\right].$$

Where, M and N are suitable constants ($M \neq 0, N \neq 0$).
Let
$e_0 = \bar{Y}^{-1}(\bar{y} - \bar{Y}), e_\phi = P^{-1}(p - P), e_\phi^{'} = P^{-1}(p_1 - P)$.
Therefore the following expected values are obtained as
$E(e_0) = E(e_\phi) = E(e_\phi^{'}) = 0, E(e_0^2) = \gamma C_y^2, E(e_\phi^2) = \gamma C_p^2$.
$E(e_\phi^{'2}) = \gamma_1 C_p^2, E(e_0 e_\phi) = \gamma C_{yp}, E(e_0 e_\phi^{'}) = \gamma_1 C_{yp}, E(e_\phi e_\phi^{'}) = \gamma_1 C_p^2$.
Rewriting the proposed estimators v_{pr1} and v_{pr1} in terms of e'^{s}, the following expressions are obtained as

$$v_{pr1} = \bar{Y}(1 + e_0)\exp\left(\frac{e_\phi^{'} - e_\phi}{M(1 + e_\phi^{'})}\right). \tag{7.1}$$

$$v_{pr2} = \bar{Y}(1+e_0)\exp\left(\frac{e_\phi^{'} - e_\phi}{N(1+e_\phi)}\right). \tag{7.2}$$

After solving the expressions (31) & (32) and retaining the terms up to second order only, the expressions obtained are as

$$v_{pr1} = \bar{Y}\left(1 + e_0 + \frac{e_\phi^{'} - e_\phi - e_\phi^{'2} + e_\phi e_\phi{}' + e_0 e_\phi^{'} - e_\phi e_0}{M} + \frac{(e_\phi^{'} - e_\phi)^2}{2M^2}\right). \tag{7.3}$$

$$v_{pr2} = \bar{Y}\left(1 + e_0 + \frac{e_\phi^{'} - e_\phi + e_\phi^{2} - e_\phi e_\phi{}' + e_0 e_\phi^{'} - e_\phi e_0}{N} + \frac{(e_\phi^{'} - e_\phi)^2}{2N^2}\right). \tag{7.4}$$

Thus the theoretical expressions of Bias and MSE for the estimators v_{pr1} and v_{pr2} obtained from equations (33) and (34) respectively are as

$$Bias(v_{pr1}) = \gamma_2 \bar{Y}\frac{1}{M}\left(\frac{1}{2M}C_p^2 - C_{yp}\right). \tag{7.5}$$

$$Bias(v_{pr2}) = \gamma_2 \bar{Y}\frac{1}{N}\left(C_p^2 + \frac{1}{2N}C_p^2 - C_{yp}\right). \tag{7.6}$$

$$MSE(v_{pr1}) = \bar{Y}^2\left(\gamma {C_y}^2 + \frac{1}{M^2}\gamma_2 {C_p}^2 - \frac{2}{M}\gamma_2 C_{yp}\right). \tag{7.7}$$

$$MSE(v_{pr2}) = \bar{Y}^2\left(\gamma {C_y}^2 + \frac{1}{N^2}\gamma_2 {C_p}^2 - \frac{2}{N}\gamma_2 C_{yp}\right). \tag{7.8}$$

Special Cases: Some special cases that have been derived from v_{pr1} and v_{pr2} are discussed as
Case1. The estimators v_{pr1} and v_{pr2} have the same MSE as that of the double sampling ratio type estimator proposed by Naik and Gupta [8] for M = N = 1.
Case2. The estimators v_{pr1} and v_{pr2} have the same MSE as that of the double sampling exponential ratio type estimator proposed by Singh et al. [18] for M = N = 2.
Case3. The estimators v_{pr1} and v_{pr2} have the same MSE as that of the double sampling product type estimator proposed by Naik and Gupta [8] for M = N = −1.
Case4. The estimators v_{pr1} and v_{pr2} have the same MSE as that of the double sampling exponential product type estimator proposed by Singh et al. [18] for M = N = −2.

Therefore the proposed double sampling ratio estimators v_{pr1} and v_{pr2} can be used as an alternative to double sampling ratio and product type estimators of Naik & Gupta [8] and the double sampling exponential ratio and product type estimators of Singh et al. [18]. The optimal value of the constants M and N that will minimize the MSE of the proposed estimators v_{pr1} & v_{pr2} is obtained by partially differentiating the equations (37) and (38), the optimal value was found to be $\frac{C_p}{\rho_{pb}C_y} = \lambda_0(say)$. Therefore the Bias and MSE of v_{pr1} and v_{pr2} at this optimal value are as

$$Bias(v_{pr1}) = -\frac{1}{2}\gamma_2 \bar{Y}\rho_{pb}^2 C_y^2. \tag{7.9}$$

$$Bias(v_{pr2}) = \gamma_2 \bar{Y}\left(C_{yp} - \frac{1}{2}\rho_{pb}^2 C_y^2\right). \tag{7.10}$$

$$MSE_{min}(v_{prj}) = \gamma \bar{Y}^2 {C_y}^2(1-\rho_{pb}^2) + \gamma_1 \bar{Y}^2 \rho_{pb}^2 {C_y}^2. \quad ; j = 1, 2. \tag{7.11}$$

That estimator v_{pr2} is unbiased, if

$$C_p = \frac{1}{2}\rho_{pb}C_y. \tag{7.12}$$

The value of λ_0 is assumed to be known in advance or can be made available by different ways as explained in Section-3 of this paper. If the value of λ_0 is not known, it is advisable to replace λ_0 by its estimate, $\hat{\lambda}_0 = \hat{\beta}(\frac{p}{\bar{y}})$ (See Upadhyaya et al. [20]). Therefore the estimators v_{pr1} and v_{pr2} will take the form

$$\hat{v}_{pr1} = \bar{y}exp\left[\frac{\bar{y}(p_1-p)}{\hat{\beta}p_1^2}\right].$$

$$\hat{v}_{pr2} = \bar{y}exp\left[\frac{\bar{y}(p_1-p)}{\hat{\beta}p_1 p}\right].$$

The MSE of $\hat{v}_{pr1}$ and $\hat{v}_{pr2}$ to the first degree of approximation is found to be same as that of equation (41).

In case the second phase sample is not nested within the first phase sample, the Bias and MSE of the estimators v_{pr1} and v_{pr2} is as

$$Bias_0(v_{pr1}) = \bar{Y}\frac{1}{M}\left(\frac{1}{2M}\gamma_3 C_p^2 - \gamma_1 C_p^2 - \gamma C_{yp}\right). \tag{7.13}$$

$$Bias_0(v_{pr2}) = \bar{Y}\frac{1}{N}\left(\frac{1}{2N}\gamma_3 C_p^2 + \gamma C_p^2 - \gamma C_{yp}\right). \tag{7.14}$$

$$MSE_0(v_{pr1}) = \bar{Y}\left(\gamma C_y^2 + \frac{1}{M^2}\gamma_3 C_p^2 - \frac{2}{M}\gamma C_{yp}\right). \tag{7.15}$$

$$MSE_0(v_{pr2}) = \bar{Y}\left(\gamma C_y^2 + \frac{1}{N^2}\gamma_3 C_p^2 - \frac{2}{N}\gamma C_{yp}\right). \tag{7.16}$$

Similarly the optimal value of the constants M and N that will minimize $MSE_0(v_{pr1})$ and $MSE_0(v_{pr2})$ is $\frac{\gamma_3 C_p}{\gamma\rho_{pb}C_y} = \lambda_0^{'}(say)$. Therefore

$$Bias_0(v_{pr1}) = -\frac{\gamma}{\gamma_3}\bar{Y}\rho_{pb}C_y\left(\gamma_1 C_p + \frac{1}{2}\gamma\rho_{pb}C_y\right). \tag{7.17}$$

$$Bias_0(v_{pr2}) = \frac{\gamma^2}{\gamma_3}\bar{Y}\rho_{pb}C_y\left(C_p - \frac{1}{2}\rho_{pb}C_y\right). \tag{7.18}$$

$$MSE_{0(min)}(v_{prj}) = \gamma\bar{Y}^2{C_y}^2\left(1 - \frac{\gamma}{\gamma_3}\rho_{pb}^2\right).; j = 1,2. \tag{7.19}$$

The value of $\lambda_0^{'}$ is assumed to be known in advance or can be made available by different ways. If it is not known, the estimated value $\hat{\lambda}_0^{'} = \hat{\beta}\left(\frac{\gamma p_1}{\gamma_3\bar{y}}\right)$ is used. Therefore,

$$\hat{v}_{pr1} = \bar{y}exp\left[\frac{(p_1-p)\gamma_3\bar{y}}{\hat{\beta}\gamma p_1^2}\right].$$

$$\hat{v}_{pr2} = \bar{y}exp\left[\frac{(p_1-p)\gamma_3\bar{y}}{\hat{\beta}\gamma p_1 p}\right].$$

Up to O (n^{-1}), it can be proven that MSE of $\hat{v}_{pr1}$ and $\hat{v}_{pr2}$ is same as (49).

8. EFFICIENCY COMPARISON

The efficiency comparison of the proposed estimators v_{pr1} and v_{pr2} with the other existing estimators is done below as

8.1. Efficiency comparison of v_{pr1} and v_{pr2}, when optimal value of M and N is taken

From equations (1), (29), (30) and (41), we get the following conditions under which v_{pr1} and v_{pr2} are more efficient than the estimators t_1, v_1 and v_2 as

$$\begin{aligned} & MSE_{min}(v_{prj}) < V(t_1) \\ & \Rightarrow \gamma\bar{Y}^2{C_y}^2(1-\rho_{pb}^2) + \gamma_1\bar{Y}^2\rho_{pb}^2{C_y}^2 < \gamma\bar{Y}^2{C_y}^2, if \\ & \gamma_2\rho_{pb}^2\bar{y}^2 > 0. \end{aligned} \tag{8.1}$$

$$\begin{aligned} & MSE_{min}(v_{prj}) < MSE(v_1) \\ & \Rightarrow \gamma\bar{Y}^2{C_y}^2(1-\rho_{pb}^2) + \gamma_1\bar{Y}^2\rho_{pb}^2{C_y}^2 < \bar{Y}^2(\gamma{C_y}^2 + \gamma_2{C_p}^2 - 2\gamma_2C_{yp}), if \\ & \gamma_2(C_p - \rho_{pb}C_y)^2 > 0. \end{aligned} \tag{8.2}$$

$$\begin{aligned} & MSE_{min}(v_{prj}) < MSE(v_2) \\ & \Rightarrow \gamma\bar{Y}^2{C_y}^2(1-\rho_{pb}^2) + \gamma_1\bar{Y}^2\rho_{pb}^2{C_y}^2 < \bar{Y}^2\left(\gamma{C_y}^2 + \frac{1}{4}\gamma_2{C_p}^2 - \gamma_2C_{yp}\right), if \\ & \gamma_2(C_p - 2\rho_{pb}C_y)^2 > 0. \end{aligned} \tag{8.3}$$

Which is true in all the cases, therefore the proposed estimators are theoretically efficient .

8.2. Efficiency comparison of v_{pr1} and v_{pr2}, when different values of M and N are taken

From equations (1), (29), (30) and (37), (38) the following conditions under which v_{pr1} and v_{pr2} are more efficient than the estimators t_1, v_1 and v_2 are obtained as

$$MSE_{min}(t_{pr1}) < V(t_1), if M > \left\{\frac{C_p^2}{2C_{yp}}\right\}. \tag{8.4}$$

$$MSE_{min}(t_{pr2}) < V(t_1), if N > \left\{\frac{C_p^2}{2C_{yp}}\right\}. \tag{8.5}$$

$$\begin{aligned} & MSE_{min}(t_{pr1}) < MSE(v_1), if \\ & min\left\{1, \frac{C_p^2}{2C_{yp} - C_p^2}\right\} < M < max\left\{1, \frac{C_p^2}{2C_{yp} - C_p^2}\right\},\ \frac{C_{yp}}{C_p^2} > \frac{1}{2}. \\ & Or\ M > 1,\ 0 \le \frac{C_{yp}}{C_p^2} \le \frac{1}{2}. \end{aligned} \tag{8.6}$$

$$MSE_{min}(t_{pr2}) < MSE(v_1), if$$

$$min\left\{1, \frac{C_p^2}{2C_{yp}-C_p^2}\right\} < N < max\left\{1, \frac{C_p^2}{2C_{yp}-C_p^2}\right\},\ \frac{C_{yp}}{C_p^2} > \frac{1}{2}. \tag{8.7}$$

$$Or\ N > 1,\ 0 \leq \frac{C_{yp}}{C_p^2} \leq \frac{1}{2}.$$

$$MSE_{min}(t_{pr1}) < MSE(v_2), if$$

$$min\left\{2, \frac{2C_p^2}{4C_{yp}-C_p^2}\right\} < M < max\left\{2, \frac{2C_p^2}{4C_{yp}-C_p^2}\right\}, \frac{C_{yp}}{C_p^2} > \frac{1}{4}. \tag{8.8}$$

$$Or\ M > 2,\ 0 \leq \frac{C_{yp}}{C_p^2} \leq \frac{1}{4}.$$

$$MSE_{min}(t_{pr2}) < MSE(v_2), if$$

$$min\left\{2, \frac{2C_p^2}{4C_{yp}-C_p^2}\right\} < N < max\left\{2, \frac{2C_p^2}{4C_{yp}-C_p^2}\right\}, \frac{C_{yp}}{C_p^2} > \frac{1}{4}. \tag{8.9}$$

$$Or\ N > 2,\ 0 \leq \frac{C_{yp}}{C_p^2} \leq \frac{1}{4}.$$

9. NUMERICAL STUDY

For comparing the efficiency of the proposed double sampling estimators v_{pr1} and v_{pr1} with the existing estimators, a first phase sample of size 45 is drawn from the populations used at Section-5 of this paper and a second phase sample which is nested within the first phase sample of size 23 is drawn. The description of the population is given below as

Table 4: **Characteristics of population data sets.**

Population	N	η	n	$\bar{Y}$	P	ρ_{pb}	C_y	C_p	$\beta_2(\phi)$	C_{yp}
P1	89	45	23	3.36	0.124	0.766	0.60	2.678	6.162	1.230
P2	89	45	23	1102	0.124	0.624	0.65	2.678	6.162	1.086

Table 5: **Values of M & N for v_{pr1} & v_{pr2} to be more efficient than the estimator's t_1, v_1 & v_2**

Estimator	Range of M & N for P1 and P2			
	P1		P2	
t_1	M > 2.913	N > 2.913	M > 3.302	N > 3.302
v_1	M > 1.000	N > 1.000	M > 1.000	N > 1.000
v_2	M > 2.000	N > 2.000	M > 2.000	N > 2.000
Optimum value (λ_0)	5.827		6.603	

Table 6: **MSE and PRE of various estimators.**

Estimator	Population			
	P1		P2	
	MSE	PRE	MSE	PRE
t_1	2.611	100.000	280808.610	100.000
v_1	1.261	207.058	145591.840	192.874
v_2	0.266	981.579	34786.030	807.245
v_{prj}	0.080	3263.750	12296.470	2283.653

The data of Table-6 clearly shows that the estimators v_{pr1} & v_{pr2} have the lowest MSE, so the highest percent relative efficiency as compared to the other estimators t_1, v_1 and v_2. The results of the table have also been shown graphically as shown in Figure 2.

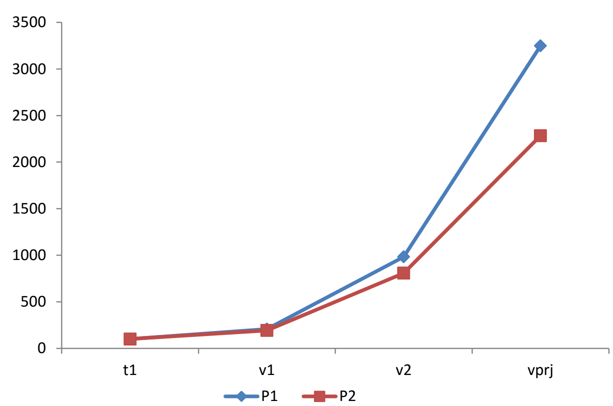

Figure 2: PRE of double sampling estimators.

10. CONCLUSION

- The proposed proportion based exponential ratio type estimators in single sampling plan are as

$$t_{pr1} = \bar{y}exp\left[\frac{\rho_{pb}C_y(P-p)}{PC_p}\right].$$

$$t_{pr2} = \bar{y}exp\left[\frac{\rho_{pb}C_y(P-p)}{pC_p}\right].$$

- The proposed proportion based exponential ratio type estimators in double sampling plan are as

$$v_{pr1} = \bar{y}exp\left[\frac{\rho_{pb}C_y(p_1-p)}{p_1C_p}\right].$$

$$v_{pr2} = \bar{y}exp\left[\frac{\rho_{pb}C_y(p_1-p)}{pC_p}\right].$$

- The estimators t_{pr2} & v_{pr2} are unbiased, if $C_p = \frac{1}{2}\rho_{pb}C_y$.

- The proposed proportion based exponential ratio estimators t_{pr1} and t_{pr2} are more efficient than the sample mean estimator, ratio estimators of Naik & Gupta [8] and Abd-Elfattah et al. [1] and the exponential ratio type estimator of Singh et al. [18].

- The proposed proportion based exponential ratio estimators v_{pr1} and v_{pr2} are more efficient than the sample mean estimator, double sampling ratio estimator of Naik & Gupta [8] and the double sampling exponential ratio estimator of Singh et al. [18].

RECEIVED: AUGUST, 2020.
REVISED: MAY, 2021.

Acknowledgement: The authors are gratefull to the editor in chief Prof. Carlos N. Bouza and the learned anonymous referees for their valuable comments/suggestions towards improvement of the manuscript. The authors are pleased for the swift and timely responses of the journal.

REFERENCES

[1] ABD-ELFATTAH, A. M., EL-SHERPIENY, E. A., MOHAMED, S. M., and ABDOUF, O. F. (2010): Improvement in estimating the population mean in simple random sampling using information on auxiliary attribute **Applied mathematics and computation**, 215(12):4198–4202.

[2] AL-OMARI, A. and BOUZA, C. (2015): Ratio estimators of the population mean with missing values using ranked set sampling **Environmetrics**, 26(2):67–76.

[3] BAHL, S. and TUTEJA, R. K. (1991): Ratio and product type exponential estimators **Journal of information and optimization sciences**, 12(1):159–164.

[4] COCHRAN, W. G. (1940): The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce **The journal of agricultural science**, 30(2):262–275.

[5] JHAJJ, H. S., SHARMA, M. K., and GROVER, L. K. (2006): A family of estimators of population mean using information on auxiliary attribute **Pakistan Journal of Statistics**, 22(1):43.

[6] MEHTA, V., SINGH, H. P., and PAL, S. K. (2020): A general procedure for estimating finite population mean using ranked set sampling **Investigación Operacional**, 41(1):80–92.

[7] MURTHY, M. N. (1967): Sampling theory and methods. **Sampling theory and methods.**

[8] NAIK, V. D. and GUPTA, P. C. (1996): A note on estimation of mean with known population proportion of an auxiliary character **Jour. Ind. Soc. Agr. Stat**, 48(2):151–158.

[9] REDDY, V. N. (1973): On ratio and product methods of estimation **Sankhyā: The Indian Journal of Statistics, Series B**, pages 307–316.

[10] REDDY, V. N. (1974): On a transformed ratio method of estimation **Sankhya C**, 36:59–70.

[11] ROBSON, D. S. (1957): Applications of multivariate polykays to the theory of unbiased ratio-type estimation **Journal of the American Statistical Association**, 52(280):511–522.

[12] SHABBIR, J. and GUPTA, S. (2007): On estimating the finite population mean with known population proportion of an auxiliary variable **Pakistan Journal of Statistics**, 23(1):1.

[13] SHABBIR, J. and GUPTA, S. (2010): Estimation of the finite population mean in two phase sampling when auxiliary variables are attributes **Hacettepe Journal of Mathematics and Statistics**, 39(1):121–129.

[14] SHALABAH and TSAI, J.-R. (2017): Ratio and product methods of estimation of population mean in the presence of correlated measurement errors **Communications in Statistics-Simulation and Computation**, 46(7):5566–5593.

[15] SINGH, H. P. and KAPRE, N. (2010): Estimation of mean, ratio and product using auxiliary information in the presence of measurement errors in sample surveys **Journal of Statistical Theory and Practice**, 4(1):111–136.

[16] SINGH, H. P. and KUMAR, S. (2008): A general family of estimators of finite population ratio, product and mean using two phase sampling scheme in the presence of non-response **Journal of Statistical Theory and Practice**, 2(4):677–692.

[17] SINGH, N. and VISHWAKARAMA, G. K. (2019): A generalised class of estimator of population mean with the combined effect of measurement errors and non-response in sample survey **Investigación Operacional**, 40(2):275–285.

[18] SINGH, R., CHAUHAN, P., SAWAN, N., and SMARANDACHE, F. (2007): Ratio-product type exponential estimator for estimating finite population mean using information on auxiliary attribute **Renaissance High press, USA**, 1:18–32.

[19] SUKHATME, P. V. and SUKHATME, B. V. (1970): Sampling theory of surveys with applications **Lowa State University Press, USA**.

[20] UPADHYAY, L. N., SINGH, H. P., CHATTERJEE, S., and YADAV, R. (2011): Improved ratio and product exponential type estimators **Journal of statistical theory and practice**, 5(2):285–302.