International Journal of Applied Mathematics
& Statistical Sciences (IJAMSS)
ISSN (P): 2319-3972; ISSN (E): 2319-3980
Vol. 11, Issue 2, Jul–Dec 2022; 13–32

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

GENERALIZED RATIO ESTIMATORS IN TWO-PHASE SAMPLING WITH TWO AUXILIARY VARIABLES

R.K. Sahoo[1], Ajit Kumar Sabat[2] & L.N. Sahoo[3]

[1,2]School of Statistics, Gangadhar Meher University, Sambalpur 768004, India

[3]Institute of Mathematics & Applications, Andharua, Bhubaneswar 751003, India

ABSTRACT

This paper considers estimation of a finite population mean under two-phase sampling procedure involving two auxiliary variables with the assumption that population mean of the first (main) auxiliary variable is unknown whereas population mean of the second (additional) auxiliary variable is known accurately. This issue has been addressed by bringing out two generalized ratio-type estimators constituting two separate families/classes of estimators of course not necessarily disjoint. Some optimum properties of the proposed generalized estimators have been investigated and sufficient conditions for their superiority over the classical two-phase ratio estimator have been reported. After identifying some ratio/ratio-type estimators as specific cases of the said generalized estimators, both analytical and empirical comparisons among various estimators have been undertaken to show the effectiveness of the proposed estimation technique.

KEYWORDS: *Auxiliary Variable, Ratio Estimator, Two-Phase Sampling*

Article History

Received: 06Aug 2022 | Revised: 09Aug 2022 | Accepted: 19 Aug 2022

INTRODUCTION

Let the study variable y and an auxiliary variable x be defined on a finite population U of N units with $(y_i, x_i), i = 1, 2, \ldots., N$ as their observed values on the ith unit. When the correlation coefficient between the two variables has a high positive value and no prior information is available on the population mean $\overline{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$, then one of the most advantageous estimation strategy for the population mean $\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ is the classical ratio estimator in conjunction with two-phase or double sampling. Here, for our purpose, let the two-phase sampling methodology be described in the following manner:

- In the first phase, a large initial sample (called first phase sample) $s_1 (s_1 \subset U)$ of $n_1$units is taken from the population by simple random sampling without replacement (SRSWOR) to obtain an acceptable estimate of $\overline{X}$ by measuring the values of x for all the $n_1$sampled units.
- In the second phase, a sub-sample (called second phase sample) $s_2 (s_2 \subset s_1)$ of $n_2$units is again selected froms_1 by SRSWOR to measure the main characteristic under study y for each of these $n_2$units.

Let $\overline{x}_1 = \frac{1}{n_1}\sum_{i\in s_1} x_i$be the sample mean of x based on the first phase sample of n_1 units; $\overline{y}_2 = \frac{1}{n_1}\sum_{i\in s_1} y_i$ and $\overline{x}_2 = \frac{1}{n_2}\sum_{i\in s_2} x_i$be the sample means of y and x respectively based on the second phase sample of n_2 units. Then the two-phase sampling classical ratio estimator for $\overline{Y}$is defined by $t_R = \overline{y}_2 \frac{\overline{x}_1}{\overline{x}_2}$.

Although t_R is biased, for large sample sizes the bias is usually negligible and the approximate expression for the mean square error (MSE) is given by

$$M(t_R) = \overline{Y}^2\left[\theta_2 C_y^2 + (\theta_2 - \theta_1)\left(C_x^2 - 2C_{yx}\right)\right], \qquad (1)$$

where $\theta_1 = \frac{1}{n_1} - \frac{1}{N}, \theta_2 = \frac{1}{n_2} - \frac{1}{N}, C_y^2 = S_y^2/\overline{Y}^2, C_x^2 = S_x^2/\overline{X}^2$ and $C_{yx}/\overline{Y}\,\overline{X}$ such that

$S_y^2 = \sum_{i=1}^N \left(y_i - \overline{Y}\right)^2 / (N-1), S_x^2 = \sum_{i=1}^N \left(x_i - \overline{X}\right)^2/(N-1)$ as the population variances of y and x,

$S_{yx} = \sum_{i=1}^N \left(y_i - \overline{Y}\right)\left(x_i - \overline{X}\right)/(N-1)$as the population covariance between y and x.

Improvements over t_R is also attainable either by redesigning the sampling scheme or by reshaping the estimator to bring considerable variance reduction compared to t_R. But, another course of action for achieving this is the involvement of one or more additional auxiliary variables. In this work, we apply some modification techniques to t_R with the aid of an additional auxiliary variable z to build up some new ratio-type estimators.

2. ASSOCIATION OF A SECOND AUXILIARY VARIABLE

Following Chand (1975) and Kiregyera (1980, 1984), let us consider a real life situation where information on $\overline{X}$ is lacking before the start of a survey operation whereas the values of a secondary auxiliary variables z are known for the entire finite population and the population mean $\overline{Z}$ is known accurately. Here it is also expected that, like$x$; z is highly correlated with y. For instance, let us refer to a survey conducted for the estimation of cattle population of a backward district with villages as the sampling units, andy_i, x_i and $z_i$are respectively as the cattle population, total area under grass lands and geographical area of the ith village. In this case, we may not get the value of x_i but the value of z_i may be known from the district records and accordingly the exact value of $\overline{Z}$ can be calculated.

The two-phase sampling mechanism in the present context is such that the preliminary sample $s_1$is used to collect measurements on (x, z) whereas the second phase sample s_2 is used to collect measurements on y only. The key idea behind this is to make reasonable estimates for $\overline{X}$ based on the measured values of$(x_i, z_i), i\epsilon s_1$. Of course, the precision of such an estimate is influenced by the correlation strength between x and z.

Estimation of $\overline{Y}$ under the above framework was fingered for the first time by Chand (1975), Sukhatme and Chand (1977) and subsequently studied by Kiregyera (1980, 1984) in greater detail. In due course of time, several authors inspired by Chand (1975) and Kiregyera (1980, 1984) and composed large varieties of estimators (ratio- or product- or regression-type). But, in our study we emphasize on the creation of estimators considering the two-phase ratio estimator $t_R$as the base.

By convention, if z has a high positive correlation with x, the ratio estimator $\overline{x}_1\,\overline{Z}/\overline{z}_1$ will estimate $\overline{X}$ more accurately than $\overline{x}_1$. Thus, replacing $\overline{x}_1$ by $\overline{x}_1\,\overline{Z}/\overline{z}_1$ in t_R, Chand (1975) suggested a ratio-in-ratio estimator defined by

$t_{RR} = \overline{y}_2 \frac{\overline{x}_1}{\overline{x}_2} \frac{\overline{Z}}{\overline{z}_1}$,

where $\overline{z}_1 = \frac{1}{n_1} \sum_{i \in s_1} z_i$. On the contrary, if z has a high negative correlation with x, the product estimator $\overline{x}_1\, \overline{z}_1 / \overline{Z}$ will estimate $\overline{X}$ more accurately than $\overline{x}_1$. Accordingly, we may consider a product-in-ratio estimator of the form

$t_{PR} = \overline{y}_2 \frac{\overline{x}_1}{\overline{x}_2} \frac{\overline{z}_1}{\overline{Z}}$.

Assuming that the regression line of x on z is linear without touching the origin, Kiregyera (1980) recommended the use of a regression estimator$\overline{x}_1 - b_{xz(1)}\left(\overline{z}_1 - \overline{Z}\right)$ in place of $\overline{x}_1$ and proposed a regression-in-ratio estimator given by

$$t_{RGR} = \overline{y}_2 \frac{\left[\overline{x}_1 - b_{xz(1)}\left(\overline{z}_1 - \overline{Z}\right)\right]}{\overline{x}_2}$$

where $b_{xz(1)} = \frac{\sum_{i \in s_1}(x_i - \overline{x}_1)(z_i - \overline{z}_1)}{\sum_{i \in s_1}(z_i - \overline{z}_1)^2}$ is the sample regression coefficient of x on z for s_1.

Asymptotic expressions for the MSEs of t_{RR}, t_{PR} and t_{RGR} are as follows:

$$M(t_{RR}) = M(t_R) + \overline{Y}^2 \theta_1 \left(C_z^2 - 2C_{yz}\right) \tag{2}$$

$$M(t_{PR}) = M(t_R) + \overline{Y}^2 \theta_1 \left(C_z^2 + 2C_{yz}\right) \tag{3}$$

$$M(t_{RGR}) = M(t_R) + \overline{Y}^2 \theta_1 \left(\rho_{xz}^2 C_x^2 - 2\rho_{yz}\rho_{xz} C_y C_x\right) \tag{4}$$

where $C_z^2 = S_z^2 / \overline{Z}^2$, $C_{yz} = S_{yz} / \overline{Y}\, \overline{Z}$, ρ_{yz} and ρ_{xz} are respectively correlation coefficients between y and z, x and z, $S_z^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(z_i - \overline{Z}\right)^2$, $S_{yz} = \frac{1}{N-1} \sum_{i=1}^{N} \left(y_i - \overline{Y}\right)\left(z_i - \overline{Z}\right)$.

Comparing MSEs, we see that t_{RR}, t_{PR} and t_{RGR} are likely to be more efficient than t_R if

$$\rho_{yz} \frac{C_y}{C_z} > \frac{1}{2}, \rho_{yz} \frac{C_y}{C_z} < -\frac{1}{2} \text{ and } \rho_{yz} > \frac{1}{2} \rho_{xz} \frac{C_x}{C_y} \tag{5}$$

respectively. These conditions therefore indicate that the strength and magnitude of the relationship between y and z also play an essential role in searching of different alternative estimators for the unknown mean $\overline{X}$ using z as an auxiliary variable.

Instead of considering$\overline{x}_1$, $\overline{x}_1\, \overline{Z} / \overline{z}_1$, $\overline{x}_1\, \overline{z}_1 / \overline{Z}$ and $\overline{x}_1 - b_{xz(1)}\left(\overline{z}_1 - \overline{Z}\right)$ as estimators of $\overline{X}$, we may also consider more generally the difference estimator $\overline{x}_1 - d\left(\overline{z}_1 - \overline{Z}\right)$, and accordingly interprete a generalized estimator for $\overline{Y}$ by

$t = \overline{y}_2 \frac{\left[\overline{x}_1 - d\left(\overline{z}_1 - \overline{Z}\right)\right]}{\overline{x}_2}$.

The estimators t_R, t_{RR}, t_{PR} and t_{RGR} run out as its special cases when $d = 0$, $\overline{x}_1 / \overline{z}_1$, $-\overline{x}_1 / \overline{Z}$ and $b_{xz(1)}$ respectively.

3. TWO GENERALIZED ESTIMATORS UNDER A MODIFIED APPROACH

If we analyze formulation techniques of different estimators included in the preceding review as well as some others available in the literature, we would like to remark that the estimators are simply recommended by the concerned authors. But no explanations have been given on the technique adopted for their construction [see for example Mukerjee *et al.* (1987), Srivenkataramana and Tracy (1989), Srivastava *et al.* (1988, 1990)]. However, in this work it is desired to develop a general framework to address questions of how to effectively use the complete available auxiliary information on z at the estimation stage. The purpose is to gain better improvements over $t_R$in respect of efficiency taking into account certain modifications over the approach of Chand (1975) and Kiregyera (1980), henceforth may be called as Chand-Kiregyera approach.

An inspection of the compositions of t_{RR}, t_{PR} and t_{RGR} explained in the previous section shows that selections of $\bar{x}_1\,\bar{Z}/\bar{z}_1$, $\bar{x}_1\,\bar{z}_1/\bar{Z}$ and $\bar{x}_1 - b_{xz(1)}(\bar{z}_1 - \bar{Z})$ over $\bar{x}_1$ in the standard two-phase ratio estimator t_R were just due to the fact that the former estimators are more efficient than the later one for estimating $\bar{X}$ under certain conditions. However, when the question of efficiency comes, we also believe that $\bar{x}_2$ provides a less efficient estimate of $\bar{X}$ than $\bar{x}_1$. Hence, this school of thought also encourages for the selection of alternative estimators for $\bar{x}_2$ in term of the second covariate z. But, to generalize our estimation methodology we prefer to use difference estimators for which we have two options: use of either $\bar{x}_2 - \eta(\bar{z}_2 - \bar{z}_1)$ or $\bar{x}_2 - \omega(\bar{z}_2 - \bar{Z})$ in place of $\bar{x}_2$. At the same time, we would also like to consider $\bar{x}_1 - d_1(\bar{z}_1 - \bar{Z})$ or $\bar{x}_1 - d_2(\bar{z}_1 - \bar{Z})$ as alternatives to$\bar{x}_1$. These arrangements give rise to the following generalized ratio-type estimators:

$$t_1^{(G)} = \bar{y}_2 \frac{\bar{x}_1 - d_1(\bar{z}_1 - \bar{Z})}{\bar{x}_2 - \eta(\bar{z}_2 - \bar{z}_1)},$$

$$t_2^{(G)} = \bar{y}_2 \frac{\bar{x}_1 - d_2(\bar{z}_1 - \bar{Z})}{\bar{x}_2 - \omega(\bar{z}_2 - \bar{Z})}.$$

The coefficients η, d_1, ω and $d_2$appearing in the generalized estimators are either suitable picked out constants or random variables converging to some finite values. But in actual practice, the said coefficients are determined so as to control the mean square errors of the estimators.

Our generalized estimators are very much flexible in the sense of being reduced to a large number of estimators using either one or two supplementary variables for suitable selections of the coefficients. Hence, they generate classes/families of estimators for $\bar{Y}$. For the simplest case when $\eta = \omega = 0$ and$d_1 = d_2 = d$, this refers to the case of using one auxiliary variable x and we have$t_1^{(G)} = t_2^{(G)} = t$, the generalized estimator defined in section 2. On the other hand, when $\eta = \omega = 0$ and $d_1 = d_2 = 0$,$t_1^{(G)} = t_2^{(G)} = t_R$, the two-phase classical ratio estimator that is our base estimator. These results imply that the classes of estimators defined by $t_1^{(G)}$ and $t_2^{(G)}$ are not necessarily disjoint. We also briefly present some specific cases of $t_1^{(G)}$ and $t_2^{(G)}$ for suitable selections of their coefficients and define the new estimators so produced as follows:

Estimators Arising Out from Both $t_1^{(G)}$and $t_2^{(G)}$

$$t_{11} = \bar{y}_2 \frac{\bar{x}_1 \bar{z}_2}{\bar{x}_2 \bar{z}_1}, \; t_{12} = \bar{y}_2 \frac{\bar{x}_1 \bar{z}_1}{\bar{x}_2 \bar{z}_2}$$

Estimators Arising Out from $t_1^{(G)}$

$$t_{13}=\bar{y}_2\frac{\bar{x}_1}{\bar{x}_2-b_{xz(2)}(\bar{z}_2-\bar{z}_1)},\quad t_{14}=\bar{y}_2\frac{\bar{x}_1\bar{z}_2\bar{Z}}{\bar{x}_2(\bar{z}_1)^2},\quad t_{15}=\bar{y}_2\frac{\bar{x}_1(\bar{z}_1)^2}{\bar{x}_2\bar{z}_2\bar{Z}},\quad t_{16}=\bar{y}_2\frac{\bar{x}_1-b_{xz(1)}(\bar{z}_1-\bar{Z})}{\bar{x}_2-b_{xz(2)}(\bar{z}_2-\bar{z}_1)}$$

where $b_{xz(2)}=\frac{\sum_{i\in s_2}(x_i-\bar{x}_2)(z_i-\bar{z}_2)}{\sum_{i\in s_2}(z_i-\bar{z}_2)^2}$.

Estimators Arising Out from $t_2^{(G)}$

$$t_{21}=\bar{y}_2\frac{\bar{x}_1\bar{z}_2}{\bar{x}_2\bar{Z}},\quad t_{22}=\bar{y}_2\frac{\bar{x}_1\bar{Z}}{\bar{x}_2\bar{z}_2},\quad t_{23}=\bar{y}_2\frac{\bar{x}_1}{\bar{x}_2-b_{xz(2)}(\bar{z}_2-\bar{Z})},\quad t_{24}=\bar{y}_2\frac{\bar{x}_1-b_{xz(1)}(\bar{z}_1-\bar{Z})}{\bar{x}_2-b_{xz(2)}(\bar{z}_2-\bar{Z})}$$

4. COMPARISON OF SOME SPECIFIC ESTIMATORS

In order to establish goodness of our modified approach over Chand-Kiregyera approach, we shall now make MSE comparisons between some specific estimators derived in section 3 and considered in section 2. But to make our comparative study manageable, we restrict ourselves to the situations in which three variables y, x and z are positively correlated. For this reason, we exclude t_{PR}, t_{12}, t_{15} and t_{22} from the comparison, and considering structural resemblance we compare t_{11}, t_{14}, t_{21} with t_R, t_{RR} and t_{13}, t_{16}, t_{23}, t_{24} with t_R, t_{RGR}.

Asymptotic expressions for the MSEs of the comparable estimators are presented below:

$$M(t_{11})=M(t_R)+\bar{Y}^2(\theta_2-\theta_1)\left(C_z^2+2C_{yz}-2C_{xz}\right) \tag{6}$$

$$M(t_{13})=M(t_{24})=M(t_R)-\bar{Y}^2(\theta_2-\theta_1)\left(\rho_{xz}^2C_x^2-2\rho_{yz}\rho_{xz}C_yC_x\right) \tag{7}$$

$$M(t_{14})=M(t_R)+\bar{Y}^2\left[\theta_2\left(C_z^2+2C_{yz}-2C_{xz}\right)-2\theta_1\left(2C_{yz}-C_{xz}\right)\right] \tag{8}$$

$$M(t_{16})=M(t_R)-\bar{Y}^2(\theta_2-2\theta_1)\left(\rho_{xz}^2C_x^2-2\rho_{yz}\rho_{xz}C_yC_x\right) \tag{9}$$

$$M(t_{21})=M(t_R)+\bar{Y}^2\left[\theta_2\left(C_z^2+2C_{yz}-2C_{xz}\right)+2\theta_1C_{xz}\right] \tag{10}$$

$$M(t_{23})=M(t_R)-\bar{Y}^2\left[\theta_2\left(\rho_{xz}^2C_x^2-2\rho_{yz}\rho_{xz}C_yC_x\right)-2\theta_1\rho_{xz}^2C_x^2\right] \tag{11}$$

In subsections 4.1 and 4.2, we expose certain sufficient conditions to show how situations do arise where the estimators coming out under modified approach perform better than their respective counterparts under Chand-Kiregyera approach. However, it may be specified here that extraction of necessary conditions are difficult.

4.1. Comparison of t_{11}, t_{14} and t_{21} with t_R and t_{RR}

From (1), (2) and (6), $M(t_{11})<M(t_R)$ if

$$\rho_{xz}\frac{C_x}{C_z}>\frac{1}{2}+\rho_{yz}\frac{C_y}{C_z}, \qquad 12)$$

and $M(t_{11})<M(t_{RR})$ if

$$\rho_{xz}\frac{C_x}{C_z}>k_1\left(\frac{1}{2}+\rho_{yz}\frac{C_y}{C_z}\right), \tag{13}$$

where $k_1 = \frac{\theta_2 - 2\theta_1}{\theta_2 - \theta_1}$.

Note that $0 < k_1 < 1$ if $n_2 < \frac{n_1}{2}$ and $M(t_{RR}) < M(t_R)$ if $\rho_{yz}\frac{C_y}{C_z} > \frac{1}{2}$. Hence, we may conclude that if t_{RR} is superior to t_R, then t_{11} is superior to both t_R and t_{RR} if

$$\rho_{xz}\frac{C_x}{C_z} > 1, \tag{14}$$

provided $n_2 < \frac{n_1}{2}$.

The condition $n_2 < \frac{n_1}{2}$ is a very mild restriction satisfied in many survey situations, and decided by the sampler at the planning stage without any appreciable increase in cost.

Precisely in a similar way and omitting details of the derivations, we also deduce that when t_{RR} is more precise than t_R, then t_{14} is more precise than both t_R and t_{RR} if

$$\rho_{xz}\frac{C_x}{C_z} > \max(k_2, 1),$$

$$\Rightarrow \quad \rho_{xz}\frac{C_x}{C_z} > k_2, \tag{15}$$

and t_{21} is more precise than both t_R and t_{RR} if

$$\rho_{xz}\frac{C_x}{C_z} > \max(k_3, k_4),$$

$$\Rightarrow \quad \rho_{xz}\frac{C_x}{C_z} > k_4, \tag{16}$$

where $k_2 = \frac{\theta_2 - \theta_1}{\theta_2 - 2\theta_1}\left(> 0 \text{ for } n_2 < \frac{n_1}{2}\right)$,$k_3 = \frac{\theta_2}{\theta_2 - \theta_1}(> 0)$ and $k_4 = \frac{1}{\theta_2 - \theta_1}(> 0)$.

4.2. Comparison of t_{13}, t_{16} and t_{23} with t_R and t_{RGR}

From (1), (4), (7) and (9) we directly see that both t_{13} and t_{16} would be more efficient than both t_R and t_{RGR} if

$$\rho_{yz} < \frac{1}{2}\rho_{xz}\frac{C_x}{C_y}. \tag{17}$$

But, under this condition t_{RGR} is less efficient thant_R. Hence, both t_{13} and t_{16} are superior to t_R and t_{RGR} just when t_{RGR} is inferior tot_R. In this sense the estimators t_{13} and t_{16} may be considered as complementary to t_{RGR}.

Comparing $M(t_{23})$ with $M(t_R)$ and $M(t_{RGR})$, we have $M(t_{23}) < M(t_R)$ if

$$\rho_{yz} < \frac{1}{2}k_5\rho_{xz}\frac{C_x}{C_y}, \tag{18}$$

and $M(t_{23}) < M(t_{RGR})$ if

$$\rho_{yz} < \frac{1}{2}k_6\rho_{xz}\frac{C_x}{C_y}, \tag{19}$$

where $k_5 = \frac{\theta_2 - 2\theta_1}{\theta_2}\left(> 0 \text{ for} n_2 < \frac{n_1}{2}\right)$ and $k_6 = \frac{\theta_2 - \theta_1}{\theta_2 + 2\theta_1}(> 0)$.

Combining preceding results, it should be clear that t_{23} would be more productive than t_R and t_{RGR} when

$$\rho_{yz} < \frac{1}{2}\min(k_5, k_6)\rho_{xz}\frac{C_x}{C_y}, \quad (20)$$

provided $n_2 < \frac{n_1}{2}$.

To have an idea on a specific real life situation where the condition (20) is realized, consider for example

$N = 200, n_1 = 30$ and $n_2 = 12$ which implies that $k_5 = 0.282$ and $k_6 = 0.373 > k_5$. Then, $\rho_{yz} < 0.141\, \rho_{xz}\frac{C_x}{C_y}$ would be sufficient for t_{23} to be more efficient than t_R and t_{RGR}.

In passing, we would like to remark that from the point of achievability and practicability, the conditions derived in this subsection are not as smooth as those derived in the preceding subsection.

5. SOME DESIGN-BASED PROPERTIES OF $t_1^{(G)}$ AND $t_2^{(G)}$

We now move to study qualities of the generalized ratio estimators on the grounds of their design-based bias and MSE. Since the exact expressions for these measures under a finite population set-up are not easily derivable, we rely only on the approximate expressions derived using Taylor linearization method. We provide these results in the following sub-sections and mention our observations and remarks.

5.1. Biases of $t_1^{(G)}$ and$t_2^{(G)}$

Derived asymptotic expressions for the biases of $t_1^{(G)}$ and$t_2^{(G)}$ are

$$B\left(t_1^{(G)}\right) = B(t_R) + \overline{Y}D\left[(\theta_2 - \theta_1)\eta\left(\eta DC_z^2 + C_{yz} - 2C_{xz}\right) - \theta_1 d_1\left(C_{yz} - C_{xz}\right)\right], \quad (21)$$

$$B\left(t_2^{(G)}\right) = B(t_R) + \overline{Y}D\left[(\theta_2 - \theta_1)\omega\left(\omega DC_z^2 + C_{yz} - 2C_{xz}\right) + \theta_1(\omega - d_2)\left(\omega DC_z^2 + C_{yz} - C_{xz}\right)\right], \quad (22)$$

where $D = \overline{Z}/\overline{X}$ and $B(t_R)$ is the asymptotic expression for the bias of t_R given by

$$B(t_R) = \overline{Y}(\theta_2 - \theta_1)\left(C_x^2 - C_{yx}\right). \quad (23)$$

As is already known, $B(t_R) = 0$ *i.e.*, t_R is approximately unbiased when

$$C_x^2 - C_{yx} = 0$$

$$\Rightarrow \quad \beta_{yx} = R, \quad (24)$$

which means that the regression line of y on x is linear passing through the origin, where $\beta_{yx} = S_{yx}/S_x^2$. Since the expressions (21) and (22) are not simple, it is not so easy to draw a similar conclusion on the biases of $t_1^{(G)}$ and$t_2^{(G)}$. However, we see that their biases are small for larger samples. In the following, we shall just derive some sufficient conditions for which $B\left(t_1^{(G)}\right) = B\left(t_2^{(G)}\right) = 0$ assuming that $B(t_R) = 0$*i.e.*, under the fulfillment of the restriction (24). Because, here our objective is to achieve improvements over t_R in certain sense.

From (21), $B\left(t_1^{(G)}\right) = 0$ when $B(t_R) = 0$, either $\eta = 0$ or $\eta DC_z^2 + C_{yz} - 2C_{xz} = 0$, and either $d_1 = 0$ or $C_{yz} - C_{xz} = 0$. But, we cannot consider both η and d_1 as zero for which $t_1^{(G)}$ is not defined. Further,

$$\eta DC_z^2 + C_{yz} - 2C_{xz} = 0 \text{ and } C_{yz} - C_{xz} = 0$$

$$\Rightarrow \qquad \eta = \frac{2C_{xz}-C_{yz}}{DC_z^2} \text{ subject to } C_{yz} - C_{xz} = 0$$

$$\Rightarrow \qquad \eta = \frac{C_{xz}}{DC_z^2} \text{ i.e.}, \eta = \beta_{xz}. \tag{25}$$

Hence, $B(t_1^{(G)}) = 0$ when $\eta = \beta_{xz}$ subject to the conditions that $d_1 \neq 0$ and $\beta_{yx} = R$.

To derive some impressive results from (22), let us rewrite the equation in the following alternative form:

$$B(t_2^{(G)}) = B(t_R) + \bar{Y}D[\theta_2\omega(\omega DC_z^2 + C_{yz} - 2C_{xz}) - \theta_1\{\omega d_2 DC_z^2 + d_2C_{yz} - (\omega + d_2)C_{xz}\}]. \tag{26}$$

Assuming that $\omega \neq 0$ and equating second term in the right side of (26) to zero, we have

$$\omega = \frac{2C_{xz}-C_{yz}}{DC_z^2}. \tag{27}$$

On the other hand, equating third term in the right side of (26) to zero subject to (27), we also find after considerable simplification that

$$d_2 = \frac{2C_{xz}-C_{yz}}{DC_z^2}. \tag{28}$$

Therefore, if (25) is satisfied $B(t_2^{(G)}) = 0$ when

$$d_2 = \omega = \frac{2C_{xz}-C_{yz}}{DC_z^2} = 2\beta_{xz} - \frac{\beta_{yz}}{R}, \tag{29}$$

where $\beta_{yz} = S_{yz}/S_z^2$ and $\beta_{xz} = S_{xz}/S_z^2$. But, for $\beta_{yx} = R$, we have

$$d_2 = \omega = 2\beta_{xz} - \frac{\beta_{yz}}{\beta_{yx}}. \tag{30}$$

In view of the preceding results, we have the following conclusions:

Under the assumption that t_R is asymptotically unbiased i.e., $\beta_{yx} = R$, $t_1^{(G)}$ is asymptotically unbiased if $\eta = \beta_{xz}$ and $d_1 \neq 0$, and$t_2^{(G)}$ is asymptotically unbiased if $d_2 = \omega = 2\beta_{xz} - \frac{\beta_{yz}}{\beta_{yx}}$.

5.2. Mean Square Errors of $t_1^{(G)}$ and $t_2^{(G)}$

Asymptotic expressions for the MSEs of the generalized ratio estimators are obtained as

$$M(t_1^{(G)}) = M(t_R) + \bar{Y}^2 D[(\theta_2 - \theta_1)\eta(\eta DC_z^2 + 2C_{yz} - 2C_{xz}) + \theta_1 d_1(d_1 DC_z^2 - 2C_{yz})], \tag{31}$$

$$M(t_2^{(G)}) = M(t_R) + \bar{Y}^2 D[(\theta_2 - \theta_1)\omega(\omega DC_z^2 + 2C_{yz} - 2C_{xz}) + \theta_1\{(\omega - d_2)^2 DC_z^2 + 2(\omega - d_2)C_{yz}\}]. \tag{32}$$

These expressions tentatively decide possible ranges or intervals for the coefficients so that the generalized estimators would be better thant_R.

From (31) we note that $M(t_1^{(G)}) < M(t_R)$ *i.e.*, $t_1^{(G)}$ would be more efficient than t_R when

$$Q_1 = \eta(\eta DC_z^2 + 2C_{yz} - 2C_{xz}) < 0 \tag{33}$$

and

$$Q_2 = d_1\left(d_1 D C_z^2 - 2C_{yz}\right) < 0. \tag{34}$$

These conditions hold iff the roots of the quadratic equations $Q_1 = 0$ in η and $Q_2 = 0$ in d_1 are real and distinct, and η and d_1 lie between them. This leads to the restrictions

$$0 < \eta \leq 2\left(\beta_{xz} - \frac{\beta_{yz}}{R}\right) \text{ and } 0 < d_1 \leq 2\frac{\beta_{yz}}{R} \tag{35}$$

or

$$2\left(\beta_{xz} - \frac{\beta_{yz}}{R}\right) \leq \eta < 0 \text{ and } 2\frac{\beta_{yz}}{R} \leq d_1 < 0, \tag{36}$$

according as $0 < \frac{\beta_{yz}}{R} < \beta_{xz}$ or $\beta_{xz} < \frac{\beta_{yz}}{R} < 0$.

Combining ranges for η and d_1 given in (35) and (36), we further have

$$0 < \eta + d_1 < 2\beta_{xz} \tag{37}$$

and

$$2\beta_{xz} < \eta + d_1 < 0 \tag{38}$$

if $0 < \frac{\beta_{yz}}{R} < \beta_{xz}$ and $\beta_{xz} < \frac{\beta_{yz}}{R} < 0$ respectively.

The derived ranges for η and d_1 in (35), (36), (37) and (38) provide certain guidelines to select their values in order to improve accuracy of $t_1^{(G)}$ compared to t_R.

From (32), $t_2^{(G)}$ would be superior to t_R if

$$Q_3 = \omega\left(\omega D C_z^2 + 2C_{yz} - 2C_{xz}\right) < 0 \tag{39}$$

and

$$Q_4 = (\omega - d_2)^2 D C_z^2 + 2(\omega - d_2)C_{yz} < 0. \tag{40}$$

The equation (39) directly implies that

$$0 < \omega \leq 2\left(\beta_{xz} - \frac{\beta_{yz}}{R}\right) \tag{41}$$

and

$$2\left(\beta_{xz} - \frac{\beta_{yz}}{R}\right) \leq \omega < 0 \tag{42}$$

for $0 < \frac{\beta_{yz}}{R} < \beta_{xz}$ and $\beta_{xz} < \frac{\beta_{yz}}{R} < 0$ respectively.

Further, from (40) we have

$$-2\frac{\beta_{yz}}{R} \leq (\omega - d_2) < 0 \tag{43}$$

for $\beta_{yz} > 0$, and

$$0 < (\omega - d_2) \leq -2\frac{\beta_{yz}}{R} \tag{44}$$

for $\beta_{yz} < 0$.

Now we see that a plausible value of ω can be determined from (41) or (42) directly whereas such a value for d_2 cannot be determined independently from (43) or (44). But, after deciding ω in light of (41) or (42),d_2 can be decided using (43) or (44).

The upper and lower limits of the ranges for η, ω, d_1 and d_2 calculated here exclusively depend on β_{yz}, β_{xz} and R. These ranges would be competent enough to provide suitable values of the coefficients so as to make the proposed generalized ratio estimators more efficient than the classical ratio estimator. Sometimes this of course may not be feasible in the absence of known values of the said parameters. However, prior knowledge from the past data or surveys or experience or even guessed values having close approximations to the true values may be very much helpful for this purpose.

6. DETERMINATION OF OPTIMAL COEFFICIENTS FOR $t_1^{(G)}$ AND $t_2^{(G)}$

As said earlier, proper selections of the coefficients make the proposed generalized estimators more effective and practicable. Towards a solution of this problem, our discussions in the preceding section is helpful to some extent by constructing appropriate ranges in terms of certain population parameters. But, here we would like to obtain the best values *i.e.*, optimal values of the coefficients $d_1,\ \eta,\ d_2$ and ω which minimize variances of the concerned estimators.

For minimizing $M\left(t_1^{(G)}\right)$, we differentiate the equation (31) w.r.t. η and d_1, and equate the resulting equations to zero to obtain the following normal equations:

$$\frac{\partial M\left(t_1^{(G)}\right)}{\partial \eta} = 0$$

$$\Rightarrow \qquad \eta D C_z^2 + C_{yz} - C_{xz} = 0, \tag{45}$$

$$\frac{\partial M\left(t_1^{(G)}\right)}{\partial d_1} = 0$$

$$\Rightarrow \qquad d_1 D C_z^2 - C_{yz} = 0, \tag{46}$$

Letting $\hat{\eta}$ and $\hat{d}_1$ as optimum values of η and d_1 respectively, from (45) and (46) we directly get

$$\hat{\eta} = \beta_{xz} - \frac{\beta_{yz}}{R} \tag{47}$$

and

$$\hat{d}_1 = \frac{\beta_{yz}}{R}. \tag{48}$$

After making use of these optimal values of the coefficients, we find the minimum value of $M\left(t_1^{(G)}\right)$ as

$$M_{min}\left(t_1^{(G)}\right) = M(t_R) - \overline{Y}^2 C_y^2 \left[(\theta_2 - \theta_1)\left(\rho_{yz} - \frac{C_x}{C_y}\rho_{xz}\right)^2 + \theta_1 \rho_{yz}^2\right]. \tag{49}$$

This minimum MSE may be designated as the minimum MSE bound of $t_1^{(G)}$. An estimator whose MSE equals to $M_{min}\left(t_1^{(G)}\right)$ is termed as a minimum MSE bound estimator that can be obtained after substituting optimal values $\hat{\eta}$ and $\hat{d}_1$ in $t_1^{(G)}$. Here, this estimator turns out to be the following ratio-type estimator:

$$t_{1R}^{(G)} = \bar{y}_2 \frac{\bar{x}_1 - \frac{\beta_{yz}}{R}(\bar{z}_1 - \bar{Z})}{\bar{x}_2 - \left(\beta_{xz} - \frac{\beta_{yz}}{R}\right)(\bar{z}_2 - \bar{z}_1)}.$$

Differentiating $M\left(t_2^{(G)}\right)$ in (32) with respect to ω and d_2 partially, we also derive the following two normal equations:

$$\frac{\partial M\left(t_2^{(G)}\right)}{\partial \omega} = 0$$

$$\Rightarrow \quad \theta_2\left(\omega D C_z^2 + C_{yz} - C_{xz}\right) + \theta_1(-d_2 D C_z^2 + C_{xz}) = 0. \tag{50}$$

$$\frac{\partial M\left(t_2^{(G)}\right)}{\partial d_2} = 0$$

$$\Rightarrow \quad (\omega - d_2) D C_z^2 + C_{yz} = 0. \tag{51}$$

One can easily note that the normal equations (50) and (51) form a system of simultaneous equations. For this reason, the optimum values of ω and d_2 cannot be found out uniquely as they depend on each other in each of the equations. Denoting $\hat{\omega}$ and $\hat{d}_2$ as the optimum values of ω and d_2, from (51) we now have

$$\hat{d}_2 = \hat{\omega} + \frac{\beta_{yz}}{R}. \tag{52}$$

Using (52), from (50) we finally obtain

$$\hat{\omega} = \beta_{xz} - \frac{\beta_{yz}}{R}, \tag{53}$$

and

$$\hat{d}_2 = \beta_{xz}. \tag{54}$$

The foregoing analysis shows that the minimum MSE bound and the corresponding MSE bound estimator of $t_2^{(G)}$ are respectively

$$M_{min}\left(t_2^{(G)}\right) = M(t_R) - \bar{Y}^2 C_y^2 \left[(\theta_2 - \theta_1)\left(\rho_{yz} - \frac{C_x}{C_y}\rho_{xz}\right)^2 + \theta_1 \rho_{yz}^2\right], \tag{55}$$

and

$$t_{2R}^{(G)} = \bar{y}_2 \frac{\bar{x}_1 - \beta_{xz}(\bar{z}_1 - \bar{Z})}{\bar{x}_2 - \left(\beta_{xz} - \frac{\beta_{yz}}{R}\right)(\bar{z}_2 - \bar{Z})}.$$

Here the important point to note is that $M_{min}\left(t_1^{(G)}\right) = M_{min}\left(t_2^{(G)}\right)$, which implies that the minimum MSE bounds of $t_1^{(G)}$ and $t_2^{(G)}$ are the same.

As in the case of range determination, here the coefficients d_1, η, d_2 and ω also require known values of the parameters β_{yz}, β_{xz} and R otherwise the optimum estimators $t_{1R}^{(G)}$ and $t_{2R}^{(G)}$ cannot be computed from the survey data. But in most of the occasions the parameters remain as unknown quantities and the usual practice is to estimate those using data available from the second-phase sample s_2.

Let $b_{yz(2)} = \frac{\sum_{i \in s_2}(y_i - \bar{y}_2)(z_i - \bar{z}_2)}{\sum_{i \in s_2}(z_i - \bar{z}_2)^2}$, $b_{xz(2)} = \frac{\sum_{i \in s_2}(x_i - \bar{x}_2)(z_i - \bar{z}_2)}{\sum_{i \in s_2}(z_i - \bar{z}_2)^2}$ and $r_2 = \frac{\bar{y}_2}{\bar{x}_2}$ respectively be the consistent estimators of β_{yz}, β_{xz} and R based on s_2. Then for computational purposes the optimum estimators shall be defined in following manner:

$$\hat{t}_{1R}^{(G)} = \bar{y}_2 \frac{\bar{x}_1 - \frac{b_{yz(2)}}{r_2}(\bar{z}_1 - \bar{Z})}{\bar{x}_2 - \left(b_{xz(2)} - \frac{b_{yz(2)}}{r_2}\right)(\bar{z}_2 - \bar{z}_1)} \quad \text{and} \quad \hat{t}_{2R}^{(G)} = \bar{y}_2 \frac{\bar{x}_1 - b_{xz(2)}(\bar{z}_1 - \bar{Z})}{\bar{x}_2 - \left(b_{xz(2)} - \frac{b_{yz(2)}}{r_2}\right)(\bar{z}_2 - \bar{Z})}.$$

It is important to understand here that the use of sample estimates in places of the respective unknown parameters does not make any change in the asymptotic MSE expressions of the resulting estimators *i.e.*,

$$M_{min}\left(t_1^{(G)}\right) = M_{min}\left(t_2^{(G)}\right) = M\left(\hat{t}_{1R}^{(G)}\right) = M\left(\hat{t}_{2R}^{(G)}\right).$$

To sum up our preceding theoretical results, once again we would like to remark that the classes of estimators constructed by $t_1^{(G)}$ and $t_2^{(G)}$ although have the same minimum MSE bounds, their MSE bound estimators are different.

7. COMPARISON OF $t_1^{(G)}$ AND $t_2^{(G)}$ WITH t

As is said earlier, the estimator t defined in section 2 under the Chand-Kiregyera approach, produces a system of estimators covering t_R, t_{RR}, t_{PR} and t_{RGR} as its potential members. This system also remains as a subclass of estimators of the wider classes of estimators coming out of $t_1^{(G)}$ and $t_2^{(G)}$ for $\eta = \omega = 0$ and $d_1 = d_2 = d$. But to confirm that our formulated modified technique is better than the Chand-Kiregyera method, we need a comparison of $t_1^{(G)}$ and $t_2^{(G)}$ with t at least in respect of bias and MSE.

Considering $\eta = \omega = 0$ and substituting $d_1 = d_2 = d$, asymptotic expressions for the bias and MSE of t can be directly obtained from those expressions for $t_1^{(G)}$ or $t_2^{(G)}$. Hence we have

$$B(t) = B(t_R) - \bar{Y} d D \theta_1 \left(C_{yz} - C_{xz}\right) \tag{56}$$

and

$$M(t) = M(t_R) + \bar{Y}^2 d D \theta_1 \left(d D C_z^2 - 2C_{yz}\right). \tag{57}$$

See that unlike $t_1^{(G)}$ and $t_2^{(G)}$, $B(t)$ does not depend on the coefficient d and $B(t) = 0$, if

$$B(t_R) = 0 \text{ and } C_{yz} - C_{xz} = 0$$

$$\Rightarrow \quad \beta_{yx} = R \text{ and } \beta_{xz} = \frac{\beta_{yz}}{R} = \frac{\beta_{yz}}{\beta_{yx}}. \tag{58}$$

This mean that under the situations where t_R is asymptotically unbiased, t is equivalently unbiased if $\beta_{xz} = \frac{\beta_{yz}}{\beta_{yx}}$. Of course this restriction appears to be more severe than the similar restriction for $t_1^{(G)}$ but equally stringent to that of $t_2^{(G)}$.

From (57), $M(t) < M(t_R)$ if either

$d < 0$ and $dDC_z^2 - 2C_{yz} > 0$

or

$d > 0$ and $dDC_z^2 - 2C_{yz} < 0$.

These conditions further imply that t is more precise than t_R when either

$$2\frac{\beta_{yz}}{R} < d < 0 \tag{59}$$

or

$$0 < d < 2\frac{\beta_{yz}}{R} \tag{60}$$

according as $\beta_{yz} > 0$ or $\beta_{yz} < 0$. These conditions are equivalent to the second conditions of (35) and (36).

From (31) and (57)

$$M(t) - M\left(t_1^{(G)}\right) = -\overline{Y}^2 D\left[(\theta_2 - \theta_1)\eta\left(\eta DC_z^2 + 2C_{yz} - 2C_{xz}\right) + \theta_1(d_1 - d)\left\{(d_1 + d)DC_z^2 - 2C_{yz}\right\}\right]. \tag{61}$$

Hence, $t_1^{(G)}$ would be more efficient than t if

$$\eta > 2\left(\beta_{xz} - \frac{\beta_{yz}}{R}\right) \tag{62}$$

and

$$d_1 > 2\frac{\beta_{yz}}{R} - d\ . \tag{63}$$

But if $d_1 = d$, (62) is the only sufficient condition to make $t_1^{(G)}$ more effective than t.

From (32) and (57) we also have

$$M(t) - M\left(t_2^{(G)}\right) = -\overline{Y}^2 D\left[(\theta_2 - \theta_1)\omega\left(\omega DC_z^2 + 2C_{yz} - 2C_{xz}\right)\right.$$
$$\left.+\theta_1\left((\omega - d_2) - d\right)\left\{\left((\omega - d_2) + d\right)DC_z^2 - 2C_{yz}\right\}\right]. \tag{64}$$

Here we see that $M(t) < M\left(t_2^{(G)}\right)$, when

$$\omega > 2\left(\beta_{xz} - \frac{\beta_{yz}}{R}\right) \tag{65}$$

and

$$\omega - d_2 > 2\frac{\beta_{yz}}{R} - d. \tag{66}$$

After fixing the value of ω according to (65), a choice of d_2 is made according to (66). On the other hand, if $d_2 = d$ then

$$\omega > \max\left[2\left(\beta_{xz} - \frac{\beta_{yz}}{R}\right), 2\frac{\beta_{yz}}{R}\right] \tag{67}$$

is the sufficient condition so that $t_2^{(G)}$ would be more efficient than t.

It may be remarked here that the above derived sufficient conditions favoring $t_1^{(G)}$ and $t_2^{(G)}$ of course difficult to check in many occasions. But they clearly indicate that there is scope for improving upon the formulated estimation strategy over that considered in Chand (1975) and Kiregyera (1980). However, in the following discussion it has been shown that for optimum choices of the coefficients, the former strategy always yields higher efficiency gain over the later one.

7.1. Efficiency Comparison for Optimal Coefficients

From (58), the optimal value of d that minimizes $M(t)$ is

$$\hat{d} = \frac{\beta_{yz}}{R},$$

and the resulting minimum MSE bound and the minimum MSE bound estimators are respectively

$$M_{min}(t) = M(t_R) - \overline{Y}^2\theta_1 C_y^2 \rho_{yz}^2 \qquad (68)$$

and

$$t^{(G)} = \overline{y}_2 \frac{\left[\overline{x}_1 - \frac{\beta_{yz}}{R}(\overline{z}_1 - \overline{Z})\right]}{\overline{x}_2} \text{ or } \hat{t}^{(G)} = \overline{y}_2 \frac{\left[\overline{x}_1 - \frac{b_{yz(2)}}{r_2}(\overline{z}_1 - \overline{Z})\right]}{\overline{x}_2}$$

if β_{yz} and R are estimated.

Now we see that $M_{min}(t_1^{(G)}) = M_{min}(t_2^{(G)}) < M_{min}(t)$. This shows that t is less efficient than both $t_1^{(G)}$ and $t_2^{(G)}$ in respect on minimum MSE bound criterion.

8. SOME REMARKS ON THE EFFICIENCIES OF $\hat{t}^{(G)}$, $\hat{t}_{1R}^{(G)}$ AND $\hat{t}_{2R}^{(G)}$

It has already been shown earlier that the four estimators *viz.*, t_R, t_{RR}, t_{PR} and t_{RGR}, and the series of estimators t_{11}, t_{12}, t_{13}, t_{14}, t_{15}, t_{16}, t_{21}, t_{22}, t_{23} and t_{24} those considered in section 3 are some specific cases of either $t_1^{(G)}$ or $t_2^{(G)}$ or both. These estimators are therefore less efficient than their minimum MSE bound estimators $\hat{t}_{1R}^{(G)}$ and$\hat{t}_{2R}^{(G)}$. On the same ground,t_R, t_{RR}, t_{PR} and t_{RGR} being particular cases of t are always less efficient than $\hat{t}^{(G)}$.

Further from (49), (55) and (68) note that

$$M_{min}(t_1^{(G)}) = M_{min}(t_2^{(G)}) < M_{min}(t).$$

This leads to a conclusion that that $\hat{t}^{(G)}$ is definitely inferior to both $\hat{t}_{1R}^{(G)}$ and $\hat{t}_{2R}^{(G)}$.

9. EMPIRICAL STUDY

In the previous sections while studying precision of one estimator over others, we derived various sufficient conditions in terms of certain parametric functions. But some of the derived conditions are so complicated to be checked in a specific situation of practical interest. This makes the job of identifying a better estimator among others a difficult one. Hence, we need an analysis of the performance of different estimators quantitatively. For this here we do carry out an empirical study using data of 12 natural populations as described in table 1. This will not only help to evaluate the gain in efficiency of $\hat{t}_{1R}^{(G)}$

or $\hat{t}_{2R}^{(G)}$ over other estimators but also to identify a better estimator among others easier. But, as in section 4, to make the study controllable we deal with $\rho_{yx} > 0$, $\rho_{yz} > 0$ and $\rho_{xz} > 0$. Hence, the estimators under consideration are t_R, t_{RR}, t_{RGR}, t_{11}, t_{13}, t_{14}, t_{16}, t_{21}, t_{23} and t_{24}. But, keeping in mind $M(t_{24}) = M(t_{13})$ and $M(\hat{t}_{1R}^{(G)}) = M(\hat{t}_{2R}^{(G)})$, we quote results for t_{13} and $\hat{t}_{1R}^{(G)}$.

To examine relative performance of the selected estimators of $\bar{Y}$, we have computed their percentage relative efficiencies (PREs) compared to the conventional estimator $\bar{y}_2$ with $V(\bar{y}_2) = \bar{Y}^2 \theta_2 C_y^2$. Allowing SRSWOR scheme at each phase, computed values of the PREs of the comparable estimators for different values of n_1 and n_2 meeting the restriction $n_2 < \frac{n_1}{2}$ are displayed in table 2.

Table 1: Description of the Populations

Pop.No.	Source	*N*	*y*	*x*	*z*
1	Murthy(1977), p.339	34villages	areaunderwheatin 1964	areaunderwheatin 1963	cultivatedareain 1961
2	Perry(2007)	8011 households	householdnetdisposalincome	thehouseholdconsumption	thenumberof householdincome-earners
3	Sukhatme(1954), p.183	34villages	areaunderwheatin 1937	areaunderwheatin 1936	totalcultivatedarea in1931
4	Anderson(2003), p.109	25brothers	headlengthof secondson	headlengthoffirst son	headbreadthoffirst son
5	Sukhatmeand Chand(1977)	120trees	bushelsofapples harvestedin1964	appletreesof bearingagein1964	bushelsofapples harvestedin1959
6	Cochran(1977),p.181	34countries	number of placebo children	number of paralytic polio cases in the placebo group	number of paralytic polio cases in the not inoculated group
7	Shukla(1966)	50plants	fiberyield/plant	plantgreenweight	basediameter
8	Srivastava(1971)	50plants	yield/plant	heightoftheplant	basediameter
9	Tripathi(1980)	225households	personsinservice	educatedpersons	sizeofhouseholds
10	Murthy(1977), p.228	80factories	output	no.ofworkers	fixedcapital
11	Fisher(1936)	50Irisflowers (versicolor)	sepalwidth	sepallength	petallength
12	Fisher(1936)	50Irisflowers (virginica)	petalwidth	sepalwidth	petallength

After careful examination of the tabulated values on the PREs, we now summarize our numerical findings in the following manner:

- As the theory asserts, $\hat{t}_{1R}^{(G)}$ attains the maximum precision amongst all with appreciable efficiency gain for all populations taken into consideration.
- Both t_{RR} and t_{RGR} are more efficient than t_R. But, as is expected, performance of $\hat{t}^{(G)}$ is better than t_R, t_{RR} and t_{RGR} in all cases.
- The estimators t_{11}, t_{14} and t_{21} are more preferable to t_{RR} but t_{21} appears to be less preferable to both t_{11} and t_{14}.
- Amongst t_{13}, t_{16} and t_{23}, t_{23} comes out as the worst one. Although for all populations t_{13} and t_{16} are superior to t_{RGR}, t_{23} is the same for 8 populations (except first 4). This negative result for t_{23} is due to nonfulfillment of its favorable conditions.

Table 2: PREs of Different Estimators

Pop.No.	Samplesizes		Estimators										
	n_1	n_2	t_R	t_{RR}	t_{11}	t_{14}	t_{21}	t_{RGR}	t_{13}	t_{16}	t_{23}	$\hat{t}^{(G)}$	$\hat{t}_{1R}^{(G)}$
1	15	7	156.9	256.5	411.3	354.2	266.3	408.6	515.3	638.5	341.9	618.7	819.0
2	600	250	147.5	165.3	180.5	190.3	167.5	150.2	151.7	153.1	149.0	210.5	225.8
3	12	5	147.7	466.9	503.8	516.1	488.3	573.9	582.3	575.2	441.8	578.6	584.9
4	10	4	130.2	178.8	185.1	203.5	182.3	185.1	197.9	187.1	178.0	190.2	202.5
5	50	20	256.4	409.2	499.9	523.9	467.3	483.5	516.1	520.5	489.6	519.5	529.3
6	13	6	112.8	135.8	186.3	199.1	136.9	153.6	197.2	226.0	157.0	208.0	259.2
7	25	8	184.6	186.6	199.6	202.4	186.9	175.6	192.7	189.4	179.7	193.6	215.3
8	20	7	143.3	164.4	184.2	191.2	169.9	148.5	190.0	196.3	150.8	183.8	235.9
9	50	20	116.5	126.2	127.6	129.9	126.8	130.5	137.3	137.9	135.3	139.6	140.3
10	30	10	172.5	246.1	251.2	286.4	248.2	264.6	290.6	278.4	273.2	296.4	315.6
11	20	8	115.9	123.5	169.6	172.3	145.4	134.2	194.4	206.2	170.1	208.2	238.6
12	18	7	107.0	114.1	182.7	181.6	164.4	122.5	208.3	212.7	197.3	210.3	222.7

After further scrutinizing the empirical findings, it may be finally concluded that the proposed estimation method in terms of $t_1^{(G)}$ and $t_2^{(G)}$can be gainfully employed in many survey situations.

REFERENCES

1. *Anderson, T.W. (2003). An Introduction to Multivariate Statistical Analysis (3rd Edition). John Wiley & Sons, Inc.*
2. *Chand, L. (1975). Some ratio-type estimators based on two or more auxiliary variables. Unpublished Ph.D. Dissertation, Iowa State University, Ames, Iowa.*
3. *Cochran, W.G. (1977). Sampling Techniques (3rd Edition). Wiley Eastern Limited.*
4. *Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, 7, 179-188.*
5. *Kiregyera, B. (1980). A chain ratio-type estimator in finite population double sampling using two auxiliary variables. Metrika, 27, 217-223.*
6. *Kiregyera, B. (1984). Regression-type estimators using two auxiliary variables and the model of double sampling. Metrika, 31, 215-226.*
7. *Mukerjee, R., Rao, T.J. and Vijayan, K. (1987). Regression-type estimators using multiple auxiliary information. Australian Journal of Statistics, 29, 244-254.*
8. *Murthy, M.N. (1977). Sampling Theory and Methods. Statistical Publishing Society, Calcutta.*
9. *Perry, P. F. (2007). Improved ratio-cum-product type estimators. Statistics in Transition, 8, 51-69.*
10. *Shukla, G.K. (1966). An alternative multivariate ratio estimate for finite population. Calcutta Statistical Association Bulletin, 15, 127-134.*
11. *Srivastava (1971). Generalized estimator for the mean of a finite population using multi-auxiliary information. Journal of the American Statistical Association, 66, 404-407.*
12. *Srivastava, S. Rani, Khare, B.B. and Srivastava, S.R. (1988). On generalized chain estimators for ratio and product of two population means using auxiliary characters. Assam Statistical Review, 2, 21-29.*

13. *Srivastava, S. Rani, Khare, B.B. and Srivastava, S.R. (1990). A generalized chain ratio estimator for mean of finite population. Journal of the Indian Society of Agricultural Statistics, 42, 108-117.*

14. *Srivenkataramana, T. and Tracy, D.S. (1989). Two-phase sampling for selection with probability proportional to size in sample surveys. Biometrika, 76, 818-821.*

15. *Sukhatme, B.V. and Chand, L. (1977). Multivariate ratio type estimator. Proceedings, Social Statistics Section of American Statistical Association, 927-931.*

16. *Sukhatme, P.V. (1954). Sampling Theory of Surveys with Applications. Iowa State College Press, Ames, Iowa.*

17. *Tripathi, T.P. (1980). A general class of estimators for population ratio. Sankhya, C42, 63-75.*