International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Some Useful Estimators for Estimation of Population Mean Under PPS Sampling

Sangeeta Malik[1], Poonam Devi[2]

[1]Professor, Dept. of Mathematics, Baba Mastnath University, Rohtak, Haryana, India
Corresponding Author: Email: sangeetastat[at]gmail.com
[2]Research Scholar, Dept. of Mathematics, Baba Mastnath University, Rohtak, Haryana, India
Author: Email: poonamsharawat9[at]gmail.com

Abstract: ***To estimate population parameter(s) using auxiliary information, many authors have given a variety of estimation techniques. In this research paper we have proposed estimators using trigonometric type estimator to estimate the population's parameters under PPS sampling, which yield beneficial results. First-degree scale estimators have been shown using data from an auxiliary variable. For the suggested measurement scales, an evaluation of bias and mean squared error is carried out. The recommended estimator types outperform comparable estimates for each unit when compared to other affective estimates. We'll compare the outcomes with existing estimators. We have done an empirical illustration and graphical representation is also included to justify the utility of the proposed estimators.***

Keywords: PPS Sampling, Trigonometry Type estimator, Mean Square Errors

1.Introduction

It is commonly known that using several auxiliary variables increases the estimators' efficiency in surveys with large populations. The following method may be applied throughout the estimator design procedure or when choosing a sample from the population. The auxiliary variables were successfully employed in this work. It is commonly known that using pps estimators of the population mean instead of the traditional estimator simple mean for equal probability sampling results in a significant efficiency gain when the research variable and selection probabilities are substantially connected. The pps estimate is unacceptable since it depends on multiplicity but not on order. Then an outcome, the final estimator is not as practical as the initial pps estimator. This research proposes different estimators for population mean under pps sampling. As a result, the final estimator is not as practical as the initial pps estimator. This research proposes different estimators for population mean under pps sampling. Anita and Shashi Bahl [2] proposed an alternative estimator for the population mean under the probability proportional to size sampling. John Graunt used the ratio estimator for the first time to determine the ratio $\frac{y}{x}$,where x was the estimated total number of births that were recorded in the identical locations during the previous year and y was the overall population. If the correlation coefficient (ρ) is positive then we use the ratio estimator. If the correlation coefficient (ρ) is negative then we cannot use the ratio estimator. In such cases Goodman has proposed another type estimator say product estimator. S. Bhal and Tuteja [13] have recommended a ratio and product type estimator. C. Kadilar and H. Cingi, Ratio [6] have proposed an estimator for the population variance in simple and stratified random sampling. Mishra Madhulika, B.P. Singh, Rajesh Singh [8] have recommended an estimation of population mean using two auxiliary variables in stratified random sampling. Nikita and Sangeeta malik [10] have proposed a generalized logarithmic ratio and product type estimators in simple random sampling. P.A. Patel and Shraddha Bhatt [11] have suggested some estimation of finite population total under probability proportional to size sampling in presence of extra auxiliary information. Sangeeta malik and Kusum [12] have proposed a new log type estimator in simple random sampling.

Let U represent a size M finite population. For every unit i, let (y_i,x_i) represent a pair of values corresponding to the study variable y and an auxiliary variable x.

$\bar{y}_{pps}=\frac{1}{m}\sum_{i=1}^{m}\frac{y_i}{p_i}$ and $\bar{x}_{pps}=\frac{1}{m}\sum_{i=1}^{m}\frac{x_i}{p_i}$

$MSE(\bar{y}_{pps}) = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{y_i}{p_i}-Y\right)^2 p_i$ and

$MSE(\bar{x}_{pps}) = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{x_i}{p_i}-X\right)^2 p_i$

2.Notations

We take a finite population of M units for the current study. The study variable y and auxiliary variable x the obtaining mean estimates $\bar{y}$ and $\bar{x}$ of the population mean $\bar{Y}$ and $\bar{X}$. To understand the bias and MSE. Let define

$$q_i = \frac{y_i}{Mk_i}, p_i = \frac{x_i}{Mk_i}$$

$$\bar{q}_m = \frac{1}{m}\sum_{i=1}^{m} q_i = \bar{y}_{pps}, \quad \bar{p}_m = \frac{1}{m}\sum_{i=1}^{m} p_i = \bar{x}_{pps}$$

$$\sigma_q^2 = \sum_1^M k_i(q_i-\bar{Y})^2, \frac{\sigma_q^2}{\bar{y}^2} = c_q^2$$

$$\sigma_p^2 = \sum_1^M k_i(p_i-\bar{X})^2, \frac{\sigma_p^2}{\bar{x}^2} = c_p^2$$

$$\rho_{pq} = \frac{\sum_{i=1}^{M} k_i\{q_i-\bar{Y}(p_i-\bar{X}_2)\}}{\sigma_q\sigma_p}$$

$$\bar{Y}_R = \frac{\bar{q}_m}{\bar{p}_m}\bar{X}, \ \bar{Y}_P = \frac{\bar{p}_m\bar{q}_m}{\bar{X}}$$

3.Proposed Estimators

$$\bar{y}_{pd1,r}=\bar{y}_{pps}[1+\lambda_1 \sin(\bar{X}-\bar{x}_{pps})] \quad (3.1)$$

$$\bar{y}_{pd1,p}=\bar{y}_{pps}[1+\lambda_2 \sin(\bar{x}_{pps}-\bar{X})] \quad (3.2)$$

$$\bar{Y}_{pd2,r}=\bar{y}_{pps}\left[1+\lambda_3 \sin\left(\frac{\bar{X}-\bar{x}_{pps}}{\bar{X}+\bar{x}_{pps}}\right)\right] \quad (3.3)$$

$$\bar{Y}_{pd2,p}=\bar{y}_{pps}\left[1+\lambda_4 \sin\left(\frac{\bar{x}_{pps}-\bar{X}}{\bar{x}_{pps}+\bar{X}}\right)\right] \quad (3.4)$$

Where λ_i are optimum constant.

Let $\bar{y} = \bar{Y}(1 + e_0)$	$E(e_0^2) = c_q^2$
$\bar{x} = \bar{X}(1 + e_1)$	$E(e_1^2) = c_p^2$
$E(e_0) = E(e_1) = 0$	$E(e_0 e_1) = \rho_{pq} c_p c_q$

Table 3.1: Bias and Mean Square Errors of the Proposed Estimators

Bias	MSE
$B(\bar{Y}_{pd1,r}) = 0$	$V(\bar{Y}_{pd1,r}) = \bar{Y}^2 (c_q^2 + \lambda_1^2 \bar{X}^2 c_p^2 - 2\lambda_1 \bar{X} c_{pq})$
$B(\bar{Y}_{pd1,p}) = \lambda_2 \bar{Y} \bar{X} c_{pq}$	$M(\bar{Y}_{pd1,p}) = \bar{Y}^2 (c_q^2 + \lambda_2^2 \bar{X}^2 c_p^2 + 2\lambda_2 \bar{X} c_{pq})$
$B(\bar{Y}_{pd2,r}) = \lambda_3 \bar{Y} \left(\frac{c_p^2}{4} - \frac{c_{pq}}{2}\right)$	$M(\bar{Y}_{pd2,r}) = \bar{Y}^2 \left(\frac{\lambda_3^2}{4} c_p^2 + c_q^2 - \lambda_3 c_{pq}\right)$
$B(\bar{Y}_{pd2,p}) = \lambda_4 \bar{Y} \left(\frac{c_{pq}}{2} - \frac{c_p^2}{4}\right)$	$M(\bar{Y}_{pd2,r}) = \bar{Y}^2 \left(\frac{\lambda_4^2}{4} c_p^2 + c_q^2 + \lambda_4 c_{pq}\right)$

4.Optimum Variance

The optimum values of λ_i are respectively

$$\lambda_{1opt} = \frac{c_{pq}}{\bar{X}\, c_p^2}\ , \lambda_{2opt} = -\frac{c_{pq}}{\bar{X}\, c_p^2}\ , \lambda_{3opt} = \frac{2c_{pq}}{c_p^2}$$

and $\lambda_{4opt} = -\frac{2c_{pq}}{c_p^2}$

The minimized optimum values of mean square errors for these λ_i are

$$\text{MSE}\ (\bar{Y}_{pd1,r})_{min} = \bar{Y}^2 \left(c_q^2 - \frac{c_{pq}^2}{c_p^2}\right)$$

$$\text{MSE}\ (\bar{Y}_{pd1,p})_{min} = \bar{Y}^2 \left(c_q^2 - \frac{c_{pq}^2}{c_p^2}\right)$$

$$\text{MSE}\ (\bar{Y}_{pd2,r})_{min} = \bar{Y}^2 \left(c_q^2 - \frac{c_{pq}^2}{c_p^2}\right)$$

And $\text{MSE}\ (\bar{Y}_{pd2,p})_{min} = \bar{Y}^2 \left(c_q^2 - \frac{c_{pq}^2}{c_p^2}\right)$

4. Efficiency Comparison:

4.1 Comparison of $\bar{y}_{pd1,r}$ with mean per unit estimator

$\text{MSE}(\bar{y}_{pd1,r}) < \text{MSE}(\bar{Y})$

$\bar{Y}^2 (c_q^2 + \lambda_1^2 \bar{X}^2 c_p^2 - 2\lambda_1 \bar{X} \rho_{pq} c_p c_q) < c_q^2 \bar{Y}^2$

$\lambda_1^2 \bar{X}^2 c_p^2 - 2\lambda_1 \bar{X} \rho_{pq} c_p c_q < 0$

$$\rho_{pq} > \frac{\lambda_1 \bar{X} c_p}{2c_q} \qquad (4.1)$$

4.2 Comparison of $\bar{y}_{pd1,r}$ with ratio estimator

$\text{MSE}(\bar{y}_{pd1,r})_{min} < \text{MSE}(\bar{Y_R})$

$\bar{Y}^2 (c_q^2 + \lambda_1^2 \bar{X}^2 c_p^2 - 2\lambda_1 \bar{X} \rho_{pq} c_p c_q) < \bar{Y}^2 [c_q^2 - 2\rho_{pq} c_p c_q + c_p^2]$

$\lambda_1^2 \bar{X}^2 c_p^2 - 2\lambda_1 \bar{X} \rho_{pq} c_p c_q + 2\rho_{pq} c_p c_q - c_p^2 < 0$

$(\lambda_1^2 \bar{X}^2 - 1) c_p^2 < 2 (\lambda_1 \bar{X} - 1) \rho_{pq} c_p c_q$

$$\rho_{pq} > \frac{(\lambda_1 \bar{X} + 1) c_p}{2\ c_q} \qquad (4.2)$$

4.3 Comparison of $\bar{y}_{pd1,r}$ with product estimator

$\text{MSE}(\bar{y}_{pd1,r})_{min} < \text{MSE}(\bar{Y_P})$

$\bar{Y}^2 (c_q^2 + \lambda_1^2 \bar{X}^2 c_p^2 - 2\lambda_1 \bar{X} \rho_{pq} c_p c_q) < \bar{Y}^2 [c_q^2 + 2\rho_{pq} c_p c_q + c_p^2]$

$\lambda_1^2 \bar{X}^2 c_p^2 - 2\lambda_1 \bar{X} \rho_{pq} c_p c_q - 2\rho_{pq} c_p c_q - c_p^2 < 0$

$(\lambda_1^2 \bar{X}^2 - 1) c_p^2 < 2 (\lambda_1 \bar{X} + 1) \rho_{pq} c_p c_q$

$$\rho_{pq} > \frac{(\lambda_1 \bar{X} - 1) c_p}{2\ c_q} \qquad (4.3)$$

4.4Comparison of $\bar{y}_{pd1,p}$ with mean per unit estimator

$\text{MSE}(\bar{y}_{pd1,p})_{min} < \text{MSE}(\bar{Y})$

$\bar{Y}^2 (c_q^2 + \lambda_2^2 \bar{X}^2 c_p^2 + 2\lambda_2 \bar{X} \rho_{pq} c_p c_q) < c_q^2 \bar{Y}^2$

$\lambda_2^2 \bar{X}^2 c_p^2 + 2\lambda_2 \bar{X} \rho_{pq} c_p c_q < 0$

$$\rho_{pq} < -\frac{\lambda_2 \bar{X} c_p}{2c_q} \qquad (4.4)$$

4.5 Comparison of $\bar{y}_{pd1,p}$ with ratio estimator

$\text{MSE}(\bar{y}_{pd1,p})_{min} < \text{MSE}(\bar{Y_R})$

$\bar{Y}^2 (c_q^2 + \lambda_2^2 \bar{X}^2 c_p^2 + 2\lambda_2 \bar{X} \rho_{pq} c_p c_q) < \bar{Y}^2 [c_q^2 - 2\rho_{pq} c_p c_q + c_p^2]$

$\lambda_2^2 \bar{X}^2 c_p^2 + 2\lambda_2 \bar{X} \rho_{pq} c_p c_q + 2\rho_{pq} c_p c_q - c_p^2 < 0$

$2 (\lambda_2 \bar{X} + 1) \rho_{pq} c_p c_q < (1 - \lambda_2^2 \bar{X}^2) c_p^2$

$$\rho_{pq} < \frac{(1 - \lambda_2 \bar{X}) c_p}{2\ c_q} \qquad (4.5)$$

4.6 Comparison of $\bar{y}_{pd1,p}$ with product estimator

$\text{MSE}(\bar{y}_{pd1,p})_{min} < \text{MSE}(\bar{Y_P})$

$\bar{Y}^2 (c_q^2 + \lambda_2^2 \bar{X}^2 c_p^2 + 2\lambda_2 \bar{X} \rho_{pq} c_p c_q) < \bar{Y}^2 [c_q^2 + 2\rho_{pq} c_p c_q + c_p^2]$

$\lambda_2^2 \bar{X}^2 c_p^2 + 2\lambda_2 \bar{X} \rho_{pq} c_p c_q - 2\rho_{pq} c_p c_q - c_p^2 < 0$

$2 (\lambda_2 \bar{X} - 1) \rho_{pq} c_p c_q < (1 - \lambda_2^2 \bar{X}^2) c_p^2$

$$\rho_{pq} < \frac{(1 + \lambda_2 \bar{X}) c_p}{2\ c_q} \qquad (4.6)$$

4.7 Comparison of $\bar{y}_{pd2,r}$ with mean per unit estimator

$\text{MSE}(\bar{y}_{pd2,r})_{min} < \text{MSE}(\bar{Y})$

$\bar{Y}^2 \left(\frac{\lambda_3^2}{4} c_p^2 + c_q^2 - \lambda_3 \rho_{pq} c_p c_q\right) < c_q^2 \bar{Y}^2$

$\frac{\lambda_3^2}{4} c_p^2 - \lambda_3 \rho_{pq} c_p c_q < 0$

$$\rho_{pq} > \frac{\lambda_3 c_p}{4c_q} \qquad (4.7)$$

4.8 Comparison of $\bar{y}_{pd2,r}$ with ratio estimator

$\text{MSE}(\bar{y}_{pd2,r})_{min} < \text{MSE}(\bar{Y_R})$

$\bar{Y}^2 \left(\frac{\lambda_3^2}{4} c_p^2 + c_q^2 - \lambda_3 \rho_{pq} c_p c_q\right) < \bar{Y}^2 [c_q^2 - 2\rho_{pq} c_p c_q + c_p^2]$

$\frac{\lambda_3^2}{4} c_p^2 - \lambda_3 \rho_{pq} c_p c_q + 2\rho_{pq} c_p c_q - c_p^2 < 0$

$\left(\frac{\lambda_3^2}{4} - 1\right) c_p^2 - 2\rho_{pq} c_q c_p (\lambda_3 - 2) < 0$

$$\rho_{pq} > \frac{(2 + \lambda_3) c_p}{4\ c_q} \qquad (4.8)$$

4.9 Comparison of $\bar{y}_{pd2,r}$ with product estimator

$\text{MSE}(\bar{y}_{pd2,r})_{min} < \text{MSE}(\bar{Y_P})$

$\bar{Y}^2 \left(\frac{\lambda_3^2}{4} c_p^2 + c_q^2 - \lambda_3 \rho_{pq} c_p c_q\right) < \bar{Y}^2 [c_q^2 + 2\rho_{pq} c_p c_q + c_p^2]$

$$\frac{\lambda_3^2}{4}c_p^2 - \lambda_3\rho_{pq}c_pc_q - 2\rho_{pq}c_pc_q - c_p^2 < 0$$
$$\left(\frac{\lambda_3^2}{4} - 1\right)c_p^2 - 2\rho_{pq}c_qc_p(\lambda_3 + 2) < 0$$
$$\rho_{pq} > \frac{(\lambda_3 - 2)c_p}{4\ c_q} \quad (4.9)$$

4.10 Comparison of $\bar{y}_{pd2,\ p}$ with mean per unit estimator

MSE($\bar{y}_{pd2,\ r}$) $_{min}$ < MSE($\bar{Y}$)
$$\bar{Y}^2\left(\frac{\lambda_4^2}{4}c_p^2 + c_q^2 + \lambda_4\rho_{pq}c_pc_q\right) < c_q^2\bar{Y}^2$$
$$\frac{\lambda_4^2}{4}c_p^2 + \lambda_4\rho_{pq}c_pc_q < 0$$
$$\rho_{pq} > -\frac{\lambda_4 c_p}{4c_q} \quad (4.10)$$

4.11 Comparison of $\bar{y}_{pd2,\ p}$ with ratio estimator

MSE($\bar{y}_{pd2,\ r}$) $_{min}$ < MSE($\bar{Y}_R$)
$$\bar{Y}^2\left(\frac{\lambda_4^2}{4}c_p^2 + c_q^2 + \lambda_4\rho_{pq}c_pc_q\right) < \bar{Y}^2\left[c_q^2 - 2\rho_{pq}c_pc_q + c_p^2\right]$$
$$\frac{\lambda_4^2}{4}c_p^2 + \lambda_4\rho_{pq}c_pc_q + 2\rho_{pq}c_pc_q - c_p^2 < 0$$
$$\left(\frac{\lambda_4^2}{4} - 1\right)c_p^2 + 2\rho_{pq}c_qc_p(\lambda_4 + 2) < 0$$
$$\rho_{pq} < \frac{(2 - \lambda_4)c_p}{4\ c_q} \quad (4.11)$$

4.12 Comparison of $\bar{y}_{pd2,\ p}$ with product estimator

MSE($\bar{y}_{pd2,\ r}$) $_{min}$ < MSE($\bar{Y}_P$)
$$\bar{Y}^2\left(\frac{\lambda_4^2}{4}c_p^2 + c_q^2 + \lambda_4\rho_{pq}c_pc_q\right) < \bar{Y}^2\left[c_q^2 + 2\rho_{pq}c_pc_q + c_p^2\right]$$
$$\frac{\lambda_4^2}{4}c_p^2 + \lambda_4\rho_{pq}c_pc_q - 2\rho_{pq}c_pc_q - c_p^2 < 0$$
$$\left(\frac{\lambda_4^2}{4} - 1\right)c_p^2 + 2\rho_{pq}c_qc_p(\lambda_4 - 2) < 0$$
$$\rho_{pq} < -\frac{(2 + \lambda_4)c_p}{4\ c_q} \quad (4.12)$$

5.Empirical study

Using two natural populations, we have computed the percent relative efficiency of the estimators with respect to $\bar{Y}$, $\bar{Y}_P$ and $\bar{Y}_R$ in order to assess the efficacy of the proposed estimators $\bar{Y}_{pdi}$ in contrast to the other estimators. According to the following formula, the estimators $\bar{Y}_{pdi}$ (i=1,2) PRE are found.

$$PRE = \left[\frac{MSE(\phi)}{MSE(\bar{Y}_{pdi})}\right] \times 100$$

Where ϕ is some estimator of population mean $\bar{Y}$
We used data on percentage relative efficiency to evaluate the effectiveness of the suggested estimators.

POPULATION: [Source: Steel and Torrie (1960, p. 282)]

Y: Long of leaf burn in sec.
X: Potassium percentage.
M=30, m=6, $\bar{Y}$=0.6860, $\bar{X}$=4.6537, c_q=0.4803,
c_p=0.2295, ρ_{pq}= 0.1794

Table 5.1: Percentage relative efficiency of proposed estimator

Estimators	Variable Constant	PRE over $\bar{Y}$	PRE over $\bar{Y}_R$	PRE over $\bar{Y}_P$
$\bar{Y}_{pd1,r}$	0.5	103.31	109.18	144.60
	0.7	102.45	108.28	143.41
	0.9	100.40	106.11	140.54
$\bar{Y}_{pd1,p}$	0.5	90.38	95.52	126.51
	0.7	85.48	90.34	119.65
	0.9	80.31	84.87	112.41
$\bar{Y}_{pd2,r}$	0.5	102.94	108.80	144.10
	0.7	103.31	109.19	144.61
	0.9	103.19	109.06	144.44
$\bar{Y}_{pd2,p}$	0.5	99.91	105.60	139.86
	0.7	99.83	105.50	139.73
	0.9	99.74	105.41	139.61

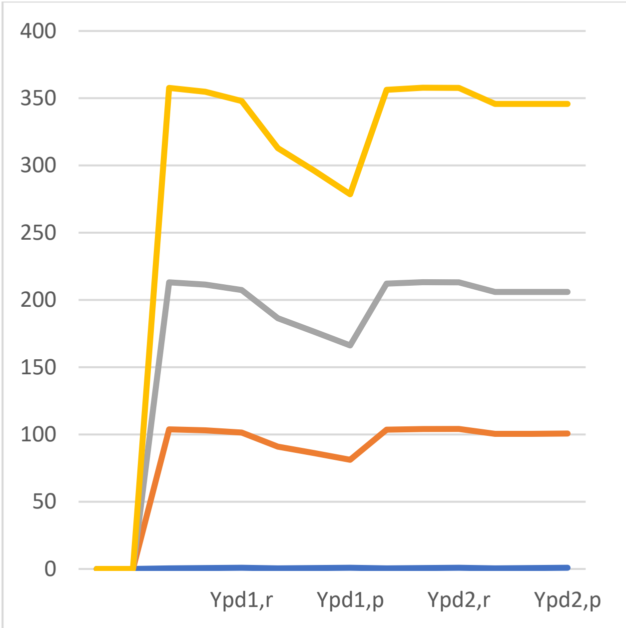

Figure 1: Relative efficiency of proposed estimator and existing estimators

6.Conclusion

Using the current mean square error and comparative effectiveness as presented in the table 5.1 suggested trigonometric estimators $\bar{Y}_{pdi}$ have been found to operate superior than the corresponding estimator found in the existing research. This conclusion is further verified by statistical analysis and the outcomes are both in theory and experiment satisfying. It follows that the proposed estimators operate more effectively relative to the estimators already in use, and their application in everyday situations is highly advised.

References

[1] A. C. Onekya, C.H. Izunobi and I.S. Iwueze, Open Journal of Statistics, 5, 27-34(2015).
[2] Anita and Shashi Bahl, "An Alternative Estimator for the Population Mean in PPS Sampling," International Journal of Science and Research, 3(4), pp-2316-2318(2015).
[3] Anita and Shashi Bahl, "Some New Estimators of Population Mean Using PPS Sampling," International

Journal of Science and Research, 4(8), pp-439-440(2015).

[4] Cochran, W.G., Journal of Estimators agriculture society, 30, 262-275(1940).

[5] C. Kadilar and H. Cingi," Ratio in Simple Random Sampling," Applied Mathematics and Computation, 151, 893-902(2004).

[6] C. Kadilar and H. Cingi," Ratio Estimator for the Population Variance in Simple and Stratified Random Sampling," Vol. 173,1047 – 1059 (2006).

[7] Johnston, J., McGrow-Hill Kongakusha, Ltd. (1982).

[8] Mishra Madhulika, B.P. Singh, Rajesh Singh, "Estimation of Population Mean Using Two Auxiliary Variables in Stratified Random Sampling," Journal of Reliablity and Statistical studies, 10(1), 59-68(2017).

[9] Murthy, M. N. Sampling Theory and Methods, Statistical Publishing Society Calcutta, India, (1967).

[10] Nikita and Sangeeta malik," Generalized Logarithmic Ratio and Product Type Estimators in Simple Random Sampling (SRS)," International Journal of Scientific Research in Mathematical and Statistical Sciences, 9(6), 56 – 60 (2022)

[11] P.A. Patel and Shraddha Bhatt, "Estimation of Finite Population Total Under PPS Sampling in Presence of Extra Auxiliary Information," International Journal of Statistics and Analysis, 1(6), pp–9-16(2016).

[12] Sangeeta malik and Kusum, "New Log Type Estimator in Simple Random Sampling," Mathematical Statistician and Engineering Applications ISSN: 2094-0343 2326-9865, 71(4), 992-998 (2022).

[13] S. Bhal and Tuteja, "Ratio and Product Type Estimator," information and Optimization Science, 8(1), 159-163(1991).

[14] Steel, R.G.D. and Torrie, J.H. (1960): Principles and Procedures of Statistics with Special Reference to the Biological Sciences. McGraw Hill, New York, 187-287.

[15] Waikhon Warseen Chaun and B.K. Singh, Global Journal of Science Frontier Research: mathematics and Decision Science, 14(2), 68-81(2014)

Paper ID: SR24425172020 DOI: https://dx.doi.org/10.21275/SR24425172020