

# Some useful Estimators for Estimation of Population Mean under PPS Sampling

Sangeeta Malik<sup>1</sup>, Poonam Devi<sup>2</sup>

<sup>1</sup>Professor, Department of Mathematics, Baba Mastnath University, Rohtak (Haryana)  
Corresponding Author Email: [sangeetastat\[at\]gmail.com](mailto:sangeetastat[at]gmail.com)

<sup>2</sup>Research Scholar, Department of Mathematics, Baba Mastnath University, Rohtak (Haryana)  
Email: [poonamsharawat9\[at\]gmail.com](mailto: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 ( $\rho$ ) is positive then we use the ratio estimator. If the correlation coefficient ( $\rho$ ) 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<sup>[6]</sup> 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.

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

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

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

## 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

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

$$\bar{q}_m = \frac{1}{m} \sum_{i=1}^m q_i = \bar{y}_{pps}$$

$$\bar{p}_m = \frac{1}{m} \sum_{i=1}^m p_i = \bar{x}_{pps}$$

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

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

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

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

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

## 3. Proposed Estimators

$$\bar{Y}_{pd1} = \bar{y}_{pps} + \lambda_1 \sin(\bar{X} - \bar{x}_{pps}) \quad \text{and} \quad \text{-----(1)}$$

$$\bar{Y}_{pd2} = \bar{y}_{pps} \left[ 1 + \lambda_2 \sin \left( \frac{\bar{X} - \bar{x}_{pps}}{\bar{X} + \bar{x}_{pps}} \right) \right] \quad \text{-----(2)}$$

Where  $\lambda_1$  and  $\lambda_2$  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 1:** Bias and Mean Square Errors of the Proposed Estimators

Bias	MSE
$B(\bar{Y}_{pd1}) = 0$	$M(\bar{Y}_{pd1}) = [\bar{Y}^2 c_q^2 + \lambda_1^2 \bar{X}^2 c_p^2 - 2\lambda_1 \bar{Y} \bar{X} \rho_{pq} c_p c_q]$
$B(\bar{Y}_{pd2}) = \frac{\lambda_2 \bar{Y} c_p^2}{4} - \frac{\lambda_2 \bar{Y} \rho_{pq} c_p c_q}{2}$	$M(\bar{Y}_{pd2}) = \bar{Y}^2 \left[ \frac{\lambda_2^2 c_p^2}{4} + c_q^2 - \lambda_2 \rho_{pq} c_p c_q \right]$

**4. Optimum Variance**

The optimum values of  $\lambda_1$  and  $\lambda_2$  are respectively

$$\lambda_{1opt} = \frac{\rho_{pq} \bar{Y} c_q}{\bar{X} c_p} \text{ and } \lambda_{2opt} = \frac{2c_q}{c_p^2} (\rho_{pq} c_p - c_q)$$

The minimized optimum values of mean square errors for these  $\lambda_1$  and  $\lambda_2$  are

$$MSE(\bar{Y}_{pd1})_{min} = (1 - \rho_{pq}) \bar{Y}^2 c_q^2$$

$$\text{And } MSE(\bar{Y}_{pd2})_{min} = \bar{Y}^2 \left[ \frac{2c_q^3}{c_p} (1 - \rho_{pq}) - \rho_{pq}^2 c_q^2 \right]$$

**Efficiency Comparison**

**4.11 Comparison of  $\bar{Y}_{pd1}$  with mean per unit estimator**

$$MSE(\bar{Y}_{pd1})_{min} < MSE(\bar{Y})$$

$$(1 - \rho_{pq}) \bar{Y}^2 c_q^2 < c_q^2 \bar{Y}^2$$

$$\rho_{pq} > 0$$

**4.12 Comparison of  $\bar{Y}_{pd1}$  with ratio estimator:**

$$MSE(\bar{Y}_{pd1})_{min} < MSE(\bar{Y}_R)$$

$$(1 - \rho_{pq}) \bar{Y}^2 c_q^2 < \bar{Y}^2 [c_q^2 - 2\rho_{pq} c_p c_q + c_p^2]$$

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

$$\rho_{pq} (2c_p c_q - c_q^2) < c_p^2$$

$$\rho_{pq} < \frac{c_p^2}{(2c_p c_q - c_q^2)}$$

**4.13 Comparison of  $\bar{Y}_{pd1}$  with product estimator:**

$$MSE(\bar{Y}_{pd1})_{min} < MSE(\bar{Y}_p)$$

$$(1 - \rho_{pq}) \bar{Y}^2 c_q^2 < \bar{Y}^2 [c_q^2 + 2\rho_{pq} c_p c_q + c_p^2]$$

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

$$\rho_{pq} (-2c_p c_q - c_q^2) < c_p^2$$

$$\rho_{pq} > \frac{-c_p^2}{(2c_p c_q - c_q^2)}$$

**4.1.4 Comparison of  $\bar{Y}_{pd2}$  with mean per unit estimator:**

$$MSE(\bar{Y}_{pd2})_{min} < MSE(\bar{Y})$$

$$\bar{Y}^2 \left[ \frac{2c_q^3}{c_p} (1 - \rho_{pq}) - \rho_{pq}^2 c_q^2 \right] < c_q^2 \bar{Y}^2$$

$$\rho_{pq}^2 + \frac{2c_q \rho_{pq}}{c_p} - \frac{2c_q}{c_p} + 1 > 0$$

$$\rho_{pq}^2 + \frac{2c_q \rho_{pq}}{c_p} + \left(1 - \frac{2c_q}{c_p}\right) > 0$$

$$\rho_{pq} > \frac{-c_p}{c_q} \pm \sqrt{\left(\frac{c_p}{c_q} + 1\right)^2 - 2}$$

**4.15 Comparison of  $\bar{Y}_{pd2}$  with ratio estimator:**

$$MSE(\bar{Y}_{pd2})_{min} < MSE(\bar{Y}_R)$$

$$\bar{Y}^2 \left[ \frac{2c_q^3}{c_p} (1 - \rho_{pq}) - \rho_{pq}^2 c_q^2 \right] < \bar{Y}^2 [c_q^2 - 2\rho_{pq} c_p c_q + c_p^2]$$

$$\frac{2c_q^3}{c_p} - \frac{2c_q^3 \rho_{pq}}{c_p} - \rho_{pq}^2 c_q^2 - c_q^2 + 2\rho_{pq} c_p c_q - c_p^2 < 0$$

$$\rho_{pq}^2 c_q^2 + 2\rho_{pq} c_q \left(\frac{c_q}{c_p} - c_p\right) + \left(c_p^2 + c_q^2 - \frac{c_q^2}{c_p}\right) > 0$$

$$\rho_{pq} > \frac{1}{c_q} \left[ \left(\frac{c_q}{c_p} - c_p\right) \pm \sqrt{\left(\frac{c_q}{c_p} - c_p\right)^2 - \left(c_p^2 + c_q^2 - \frac{c_q^2}{c_p}\right)} \right]$$

**4.1.6 Comparison of  $\bar{Y}_{pd2}$  with product estimator:**

$$MSE(\bar{Y}_{pd2})_{min} < MSE(\bar{Y}_p)$$

$$\bar{Y}^2 \left[ \frac{2c_q^3}{c_p} (1 - \rho_{pq}) - \rho_{pq}^2 c_q^2 \right] < \bar{Y}^2 [c_q^2 + 2\rho_{pq} c_p c_q + c_p^2]$$

$$\frac{2c_q^3}{c_p} - \frac{2c_q^3 \rho_{pq}}{c_p} - \rho_{pq}^2 c_q^2 - c_q^2 - 2\rho_{pq} c_p c_q - c_p^2 < 0$$

$$\rho_{pq}^2 c_q^2 + 2\rho_{pq} c_q \left(\frac{c_q}{c_p} + c_p\right) + \left(c_p^2 + c_q^2 - \frac{c_q^2}{c_p}\right) > 0$$

$$\rho_{pq} > \frac{1}{c_q} \left[ \left(\frac{c_q}{c_p} - c_p\right) \pm \sqrt{\left(\frac{c_q}{c_p} + c_p\right)^2 - \left(c_p^2 + c_q^2 - \frac{c_q^2}{c_p}\right)} \right]$$

**5. Empirical Study**

We have calculated the percent relative efficiency of the estimators with respect to the estimators  $\bar{Y}$ ,  $\bar{Y}_p$  applying two natural populations in order to determine the effectiveness of the suggested estimators  $\bar{Y}_{pdi}$  in comparison to the remaining estimators. The formula as follows is used to determine the estimators  $\bar{Y}_{pdi}$ 's (i=1,2) PRE.

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

Where  $\phi$  is some estimator of population mean  $\bar{Y}$

We used data on percentage relative efficiency to evaluate the effectiveness of the suggested estimators.

POPULATION I: [Source: Cochran, 4]

N=80, n=23,  $\bar{Y}$ =90.813,  $\bar{X}$ =104.515,  $c_q$ =0.392,  $c_p$ =0.379,  $\rho_{pq}$ =0.628

POPULATION II: [Source: Kalidar and Cingi (2006) ,6]

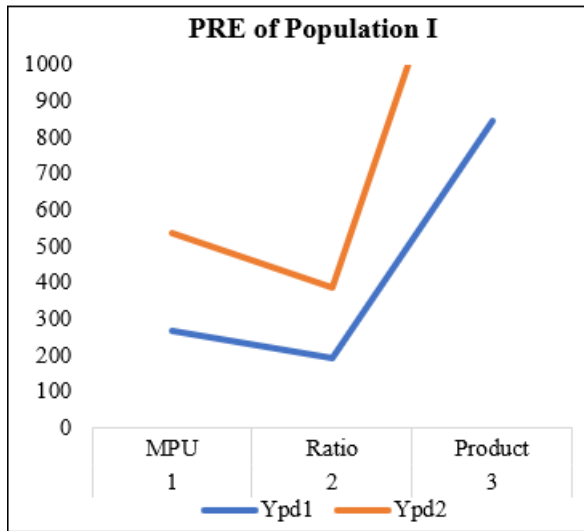
M=200, m=50,  $\bar{Y}$ =500,  $\bar{X}$ =25,  $c_q$ =15,  $c_p$ =2,  $\rho_{pq}$ =0.9

**Table 2:** PRE of minimized optimum value of proposed estimator with the corresponding existing estimator based on simple mean  $\bar{Y}$  on population I

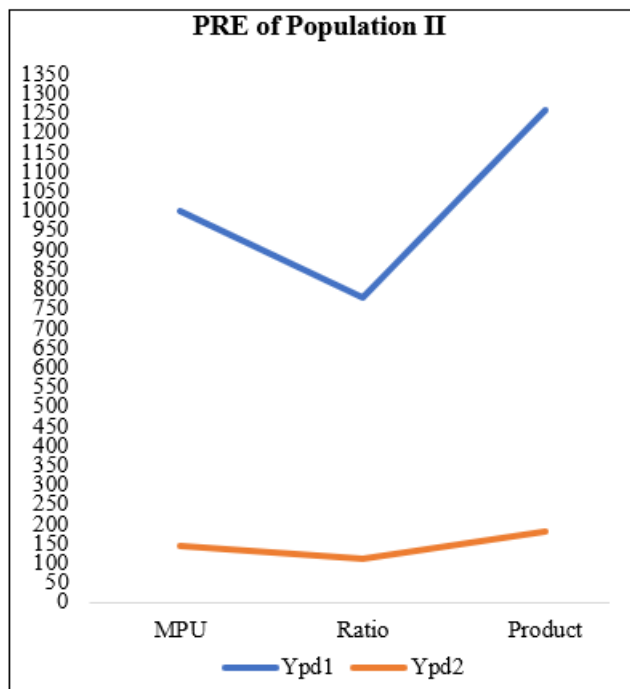
S. No.	Estimators	PRE (Mean Per Unit)	PRE (Ratio)	PRE (Product)
1	$\bar{Y}_{pd1}$	268.82	193.66	846.54
2	$\bar{Y}_{pd2}$	266.57	192.04	839.46

**Table 3:** PRE of minimized optimum value of proposed estimator with the corresponding existing estimator based on simple mean  $\bar{Y}$  population II

S. No.	Estimators	PRE (Mean Per Unit)	PRE (Ratio)	PRE (Product)
1	$\bar{Y}_{pd1}$	1000	777.78	1257.78
2	$\bar{Y}_{pd2}$	144.93	112.72	182.29



**Figure 1:** Relative efficiency of proposed estimator and existing estimators.



**Figure 2:** Relative efficiency of proposed estimator and existing estimators

### 6. Conclusion

Using the current mean square error and comparative effectiveness as presented in the table 1,2 and 3 suggested trigonometric estimators  $\bar{Y}_{pd1}$  and  $\bar{Y}_{pd2}$  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., McGraw-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 Reliability 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] Waikhon Warseen Chaun and B.K. Singh, Global Journal of Science Frontier Research: mathematics and Decision Science, 14(2), 68-81(2014).