International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

An Efficient Estimator of Population Mean using Auxiliary Information on Both Occasions

Yeesha Verma[1], Vyas Dubey[2]

[1]*Statistician cum Assistant Professor, Department of Community Medicine, Shri Rawatpura Sarkar Institute of Medical Sciences and Research, Nava Raipur, India*

[2]*Professor, School of Studies in Statistics, Pt. Ravishankar Shukla University, Raipur, India*

***Abstract*-- In this paper, we consider an estimator of the population mean based on successive sampling. In this study, we use auxiliary information on both occasions to improve the efficiency of the proposed estimators. The expressions of bias and mean square error of the proposed estimator are derived, and it is seen that they are more efficient than other existing estimators for the optimum choice of constants. Finally, numerical illustrations were made to support the numerical results.**

***Keyword*-- Auxiliary information, Successive sampling, Bias, Mean Square Error, Relative efficiency.**

I. INTRODUCTION

In sample surveys, it is well known that the auxiliary information is used to improve the efficiency of estimators. In the case of successive sampling, it has been seen that the information collected on previous occasions is used to improve the current estimates. Jessen (1942) was the first who use the information collected on previous occasions to estimate the population mean due to a matched portion of the sample over current occasions. Sukhatme et al. (1984) considered a regression-type estimator where the information on previous occasion was used as an auxiliary variable.

Let $U = U_1, U_2 \ldots\ldots\ldots U_N$ be a finite population consisting of N identifiable units and let study variable be denoted by y_1 and y_2 on the first and the second occasions respectively, assume two auxiliary variables x and z, positively related to y. Let a simple random sample of size n be selected on the first occasion from population N and assume that a random sub-sample of $m = np$ units are retained (matched) on second occasion and remaining (unmatched) $u = n - m = qn$, $(q = 1 - p)$ units are drawn afresh on the second occasion from $N - n$ units.

Sukhatme et al. (1984) suggested fallowing estimator of $\bar{Y}_2$ based on common units

$$\bar{y}_{l2(m)} = \bar{y}_{2m} + \beta_{21}(\bar{x}_{1n} - \bar{x}_{1m}) \quad (1.1)$$

which is unbiased and has variance

$$V(\bar{y}_{l2(m)}) = \left(\frac{1}{m} - \frac{1}{n}\right)(1 - \rho_{yx}^2)S_{2y}^2 + \left(\frac{1}{n} - \frac{1}{N}\right)S_{2y}^2 \quad (1.2)$$

From combining $\bar{y}_{2u}$, with unmatched portion of samples on second occasion, further he proposed weighed estimator of $\bar{Y}_2$.

$$\hat{T}_1 = \phi_1 \bar{y}_{2u} + (1 - \phi_1)\bar{y}_{l2(m)} \quad (1.3)$$

Where,

$\bar{Y}_i \ (i = 1,2)$ be the population mean of y and $\bar{X}_1$ be the population mean of x.

$\bar{y}_{1n}$, $\bar{x}_{1n}$ - Sample mean of y and x based on n units observed on the first occasion.

$\bar{y}_{2m}$ - Sample mean of y based on m units observed on the second occasion, which is common to both the occasions.

$\bar{x}_{1m}$, $\bar{z}_{1m}$ - Sample mean of x based on m units observed on first occasion, which is common to both the occasions.

$\bar{y}_{2u}$ - Sample mean of y based newly draw u units observed on second occasion, which are unmatched units.

β_{21} - Regression coefficient of second occasion values on first occasions values.

$$V(\bar{y}_{2m}) = \left(\frac{1}{m} - \frac{1}{N}\right)S_{2y}^2, \quad V(\bar{y}_{2u}) = \left(\frac{1}{u} - \frac{1}{N}\right)S_{2y}^2$$

$$V(\bar{x}_{1m}) = \left(\frac{1}{m} - \frac{1}{N}\right)S_x^2, \quad V(\bar{x}_{1n}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_x^2$$

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

In many survey situations, information on auxiliary variables may be readily available on the first as well as on the second occasions. Assume that the auxiliary variable Z is available at the second occasion, for estimating $\bar{Y}_2$ Singh and Singh (2001), Proposed two independent estimators first based on sample of size u and second based on the sample of size m. Biradar and Singh (2001) suggested estimator using auxiliary information on both occasions.

Further, Dubey and Shukla (2020) proposed modified estimator on two occasions. They consider the following estimator due to matched portion of sample

$$\hat{\bar{Y}}^* = W_1 \bar{y}_{2m} + W_2 \bar{x}_{1m} + (1 - W_1 - W_2)\bar{x}_{1n} \qquad (1.4)$$

Where W_1 and W_2 are suitable constants.

Combining estimator $\bar{y}_{2u}$ with the above estimator, he proposed combine estimator as

$$T_D^* = \phi_3 \bar{y}_{2u} + (1 - \phi_3)\hat{\bar{Y}}^* \qquad (1.5)$$

II. PROPOSED ESTIMATOR OF $\bar{Y}_2$

For estimating $\bar{Y}_2$ here we proposed two estimators. Assuming that the auxiliary variable z is available at the second occasion we suggest an estimator for unmatched portion based on sample of size u

$$\hat{\bar{Y}}_{lu} = \bar{y}_{2u} + \beta_{yz}(\bar{Z} - \bar{z}_{2u}) \qquad (2.1)$$

To estimate the population mean $\bar{Y}_2$ based on matched m units, we proposed estimator a modified regression type estimator using two auxiliary variables

$$T_m = \delta_1 \bar{y}_{2m} + \delta_2(\bar{x}_{1m} - \bar{x}_{1n}) + \delta_3(\bar{z}_{1m} - \bar{z}_{1n}) + (1 - \delta_1)\bar{x}_{1n} \qquad (2.2)$$

where δ_1 , δ_2 and δ_3 are suitable constant.

$\bar{Z}$ - Population mean of variable z

$\bar{z}_{1u}$ - Sample means of z based on u units observed on second occasion,

$\bar{z}_{2m}$ - Sample means of z based on m units observed on second occasion, which is common to both the occasions.

β_{yz} - Regression coefficient of y and z at the second occasion.

III. PROPERTIES OF PROPOSED ESTIMATOR

Proposed estimator $\hat{\bar{Y}}_{lu}$ is an unbiased estimator and variance is given by

$$V\left(\hat{\bar{Y}}_{lu}\right) = \left(\frac{1}{u} - \frac{1}{N}\right)\left(1 - \rho_{yz}^2\right)S_{2y}^2 \qquad (3.1)$$

Bias and MSE of the proposed estimator are given as

$$Bias(T_m) = \bar{Y}_2(\delta_1 - 1)(1 - P_0) \qquad (3.2)$$

and

$$\begin{aligned} MSE(T_m) = &+\delta_1^2 V(\bar{y}_{2m}) + \delta_2^2 V(\bar{x}_{1m} - \bar{x}_{1n}) + \delta_3^2 V(\bar{z}_{1m} - \bar{z}_{1n}) \\ &+ (1-\delta_1)^2 V(\bar{x}_{1n}) + 2\delta_1\delta_2 Cov(\bar{y}_{2m}, \bar{x}_{1m} - \bar{x}_{1n}) \\ &+ 2\delta_1\delta_3 Cov(\bar{y}_{2m}, \bar{z}_{1m} - \bar{z}_{1n}) + 2\delta_1(1-\delta_1)Cov(\bar{y}_{2m}, \bar{x}_{1n}) \\ &+ 2\delta_2\delta_3 Cov((\bar{x}_{1m} - \bar{x}_{1n}), (\bar{z}_{1m} - \bar{z}_{1n})) \\ &+ 2\delta_2(1-\delta_1)Cov(\bar{x}_{1n}, (\bar{x}_{1m} - \bar{x}_{1n})) \\ &+ 2\delta_3(1-\delta_1)Cov(\bar{x}_{1n}, (\bar{z}_{1m} - \bar{z}_{1n})) + \bar{Y}_2^2(\delta_1 - 1)^2(1 - P_0)^2 \end{aligned} \qquad (3.3)$$

Following Sukhatme et al. (1984), we have

$$V(\bar{x}_{1m} - \bar{x}_{1n}) = \left(\frac{1}{m} - \frac{1}{n}\right)S_x^2$$

$$V(\bar{z}_{1m} - \bar{z}_{1n}) = \left(\frac{1}{m} - \frac{1}{n}\right)S_z^2$$

$$Cov(\bar{y}_{2m}, \bar{x}_{1n}) = \left(\frac{1}{n} - \frac{1}{N}\right)S_{yx}$$

$$Cov(\bar{y}_{2m}, (\bar{x}_{1m} - \bar{x}_{1n})) = \left(\frac{1}{m} - \frac{1}{n}\right)S_{yx}$$

$$Cov(\bar{y}_{2m}, (\bar{z}_{1m} - \bar{z}_{1n})) = \left(\frac{1}{m} - \frac{1}{n}\right)S_{yz}$$

$$Cov((\bar{x}_{1m} - \bar{x}_{1n}), (\bar{z}_{1m} - \bar{z}_{1n})) = \left(\frac{1}{m} - \frac{1}{n}\right)S_{xz}$$

$$Cov(\bar{x}_{1n}, (\bar{x}_{1m} - \bar{x}_{1n})) = 0$$

$$Cov(\bar{x}_{1n}, (\bar{z}_{1m} - \bar{z}_{1n})) = 0$$

$$\theta = \left(\frac{1}{n} - \frac{1}{N}\right), \quad \theta_1' = \left(\frac{1}{m} - \frac{1}{N}\right), \quad \theta'' = \left(\frac{1}{m} - \frac{1}{n}\right)$$

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

So

$$MSE(T_m) = \delta_1^2\left(\frac{1}{m}-\frac{1}{N}\right)S_y^2 + \theta''\left[\delta_2^2 S_x^2 + \delta_3^2 S_z^2 + 2\delta_1\lambda_2 S_{yx} + 2\delta_1\lambda_3 S_{yz} + 2\delta_2\delta_3 S_{xz}\right]$$
$$+\theta\left[(1-\delta_1)^2 S_x^2 - \left(\delta_1-\delta_1^2\right)S_{yx}\right] + \bar{Y}_2^2(\delta_1-1)^2(1-P_0)^2 \qquad (3.4)$$

The optimum values of δ_1, δ_2 and δ_3 are

$$\delta_{01} = \frac{(1-p_0)^2 + \theta\left(p_0^2 C_x^2 - p_0 C_{yx}\right)}{(1-p_0)^2 + \theta\left(p_0^2 C_x^2 - 2p_0 C_{yx}\right) + C_{\bar{y}l2(m)}^2 - \theta'' C_{2y}^2} \qquad (3.5)$$

where

$$C_{\bar{y}l2(m)}^2 = \frac{V\left(\bar{y}_{l2(m)}\right)}{\bar{Y}^2} = \left(\frac{1}{m}-\frac{1}{n}\right)\left(1-\rho_{yx}^2\right)C_{2y}^2 + \left(\frac{1}{n}-\frac{1}{N}\right)C_{2y}^2$$

$$P_0 = \frac{\bar{X}_1}{\bar{Y}_2}, \qquad b = \frac{\left(\rho_{yx}-\rho_{yz}\rho_{xz}\right)^2}{\left(1-\rho_{xz}^2\right)}$$

$$\delta_{02} = \delta_{01}\frac{\beta_{yx}}{\rho_{yx}}\left[\frac{\rho_{yz}\rho_{xz}-\rho_{yx}}{\left(1-\rho_{xz}^2\right)}\right] \qquad (3.6)$$

$$\delta_{03} = -\delta_{01}\frac{\beta_{yz}}{\rho_{yz}}\left[\rho_{yz} + \frac{\rho_{yz}\rho_{xz}^2-\rho_{yx}\rho_{xz}}{\left(1-\rho_{xz}^2\right)}\right] \qquad (3.7)$$

From equation (3.4), (3.5), (3.6) and (3.7), we have minimum MSE

$$MSE(T_m)_{\min} = \frac{\left[(1-p_0)^2 + \theta p_0^2 C_x^2\right]\left[V\left(\bar{y}_{l2(m)}\right) - \theta S_y^2 b\right] - \theta^2 C_y^2 S_x^2 \rho_{yx}^2}{(1-p_0)^2 + \theta\left(p_0^2 C_x^2 - 2p_0 C_{yx}\right) + C_{\bar{y}l2(m)}^2 - \theta'' C_{2y}^2 b} \qquad (3.8)$$

3.1 EFFICIENCY COMPRESSION

From comparing efficiency, we have

$$MSE(\hat{\bar{Y}}^*) = \frac{V\left(\bar{y}_{l2(m)}\right)\left\{(1-P_0)^2 + \theta P_0^2 C_x^2\right\} - \theta^2 C_y^2 s_x^2 \rho_{yx}^2}{V\left(C_{\bar{y}l2(m)}^2\right) + (1-P_0)^2 + \theta\left(P_0^2 C_x^2 - 2P_0 C_{yx}\right)} \qquad (3.9)$$

From (3.9) and (1.2)

$$MSE(\hat{\bar{Y}}^*)_{\min} < V(\bar{y}_{l2(m)}) \text{ if} \qquad (3.10)$$

$$\left\{\sqrt{\bar{y}_{l2(m)}} - \frac{\theta S_{yx}}{\sqrt{\bar{y}_{l2(m)}}}\right\}^2 > 0$$

From comparing (3.8) and (3.9)

$$MSE(T_m)_{\min} < MSE(\hat{\bar{Y}}^*)_{\min} \text{ if} \qquad (3.11)$$

$$\left\{(1-P_0)^2 + \theta P_0^2 C_x^2\right\} > 0$$

which condition holds always.

So it is conclude that proposed estimator is more efficient than others existing estimator.

IV. Combined Estimator

Now we proposed combine estimator for $\bar{Y}_2$

$$T^* = \phi\hat{\bar{Y}}_{lu} + (1-\phi)T_m \qquad (4.1)$$

The proposed estimator has bias

$$Bias(T^*) = \phi\bar{Y}_2(\lambda_1-1)(1-P_0) \qquad (4.2)$$

and MSE

$$MSE\left(T^*\right) = \phi^2 V\left(\hat{\bar{Y}}_{lu}\right) + (1-\phi)^2 M(T_m) + 2\phi(1-\phi)Cov\left(\hat{\bar{Y}}_{lu}, T_m\right) \qquad (4.3)$$

Differentiate equation with respect to ϕ, for obtaining optimum value of ϕ, for which the above MSE will be minimum

$$\phi = \frac{M(T_m) - Cov\left(\hat{\bar{Y}}_{lu}, T_m\right)}{V\left(\hat{\bar{Y}}_{lu}\right) + M(T_m) - 2Cov\left(\hat{\bar{Y}}_{lu}, T_m\right)}$$

Minimum MSE is given by

$$MSE\left(T^*\right)_{opt} = \frac{M(T_m)V\left(\hat{\bar{Y}}_{lu}\right) - \left\{Cov\left(\hat{\bar{Y}}_{lu}, T_m\right)\right\}^2}{M(T_m) + V\left(\hat{\bar{Y}}_{lu}\right) - 2Cov\left(\hat{\bar{Y}}_{lu}, T_m\right)}$$

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

Following Sukhatme et al. (1984), we have

$$Cov\left(\hat{\hat{Y}}_{lu}, T_m\right) = Cov\begin{pmatrix}\delta_1 \bar{y}_{2m} + \delta_2(\bar{x}_{1m} - \bar{x}_{1n}) + \delta_3(\bar{z}_{1m} - \bar{z}_{1n}) \\ + (1-\delta_1)\bar{x}_{1n}\end{pmatrix} \left(\bar{y}_{2u} + \beta_{yz}(\bar{Z} - \bar{z}_{2u})\right)$$

$$= Cov\{\delta_1 \bar{y}_{2m}, (\bar{y}_{2u} + \beta_{yz}(\bar{Z} - \bar{z}_{2u}))\}$$

$$= Cov\{\delta_1(\bar{y}_{2m}, \bar{y}_{2u}) + \delta_1 \beta_{yz} \bar{y}_{2m}(\bar{Z} - \bar{z}_{2u})\}$$

$$= \delta_1 Cov(\bar{y}_{2m}, \bar{y}_{2u}) + \delta_1 \beta_{yz} Cov(\bar{y}_{2m}, \bar{Z} - \bar{z}_{2u})$$

$$= -\delta_1 \frac{S_{2y}^2}{N} + \delta_1 \beta_{yz}\left(-\frac{S_{2z}^2}{N}\right)$$

$$= -\frac{\delta_1}{N}\left(S_{2y}^2 + \beta_{yz} S_{2z}^2\right)$$

$$MSE\left(T^*\right)_{opt} = \frac{M(T_m)V\left(\hat{\hat{Y}}_{lu}\right) - \left\{\frac{\delta_1}{N}\left(S_{2y}^2 + \beta_{yz} S_{2z}^2\right)\right\}^2}{M(T_m) + V\left(\hat{\hat{Y}}_{lu}\right) + 2\frac{\delta_1}{N}\left(S_{2y}^2 + \beta_{yz} S_{2z}^2\right)} \quad (4.4)$$

4.1 Efficiency Comparison

For comparing efficiency, we have

$$V\left(\hat{\hat{Y}}_2\right)_{opt} = \frac{V(\bar{y}_{l2(m)})V(\bar{y}_{2u}) - \left(\frac{S_{2y}^2}{N}\right)^2}{V(\bar{y}_{l2(m)}) + V(\bar{y}_{2u}) + 2\frac{S_{2y}^2}{N}} \quad (4.5)$$

$$MSE\left(T_D^*\right)_{opt} = \frac{M(T_D)V(\bar{y}_{2u}) - \left\{W_1 \frac{S_{2y}^2}{N}\right\}^2}{M(T_D) + V(\bar{y}_{2u}) + 2W_1 \frac{S_{2y}^2}{N}} \quad (4.6)$$

It is difficult to compare the minimum MSE's of the estimator with proposed estimator. If that N is large, so that terms $\frac{S_{2y}^2}{N}$ and $\frac{S_{2z}^2}{N}$ can be ignore, in this case equation (4.4), (4.5) and (4.6) reduces as

$$V\left(\hat{\hat{Y}}_2\right)_{opt} = \frac{V(\bar{y}_{l2(m)})V(\bar{y}_{2u})}{V(\bar{y}_{l2(m)}) + V(\bar{y}_{2u})} \quad (4.7)$$

$$MSE\left(T_D^*\right)_{opt} = \frac{M\left(\hat{\hat{Y}}^*\right)V(\bar{y}_{2u})}{M\left(\hat{\hat{Y}}^*\right) + V(\bar{y}_{2u})} \quad (4.8)$$

$$MSE\left(T^*\right)_{opt} = \frac{M(T_m)V\left(\hat{\hat{Y}}_{lu}\right)}{M(T_m) + V\left(\hat{\hat{Y}}_{lu}\right)} \quad (4.9)$$

From comparing (4.7) and (4.8)

$$MSE(T_D^*)_{opt} < V(\hat{\hat{Y}}_2)_{opt} \text{ if} \quad (4.10)$$

$$MSE\left(\hat{\hat{Y}}^*\right) < V\left(\bar{y}_{l2(m)}\right)$$

From comparing (4.8) and (4.9)

$$MSE(T^*)_{opt} < MSE(T_D^*)_{opt} \text{ if} \quad (4.11)$$

$$MSE(T_m)_{\min} < MSE(\hat{\hat{Y}}^*)_{\min}$$

i.e. $\left\{(1-P_0)^2 + \theta P_0^2 C_x^2\right\} > 0$

V. NUMERICAL COMPRESSION

Data is considered from Sukhatme et al. (1984) which relates to the total cultivate area and also the area under wheat for a sample of 34 villages from a population of 170 villages. We have

$N = 34$, $n = 16$, $m = u = 8$

$\bar{Y}_2 = 201.412$, $\bar{X}_1 = 218.412$, $\bar{Z}_1 = 765.353$

$S_y^2 = 23154.856$, $S_x^2 = 28123.219$, $S_{yx} = 23730.128$

$\rho_{yx} = 0.9299$, $\rho_{yz} = 0.899$, $\rho_{xz} = 0.831$

The relative efficiency of estimators is given in Table5.1. Relative efficiency with respect to $\hat{\bar{Y}}_2$ is define by $\frac{V\left(\hat{\bar{Y}}_2\right)}{MSE(.)} \times 100$

Table 1

Estimator	Relative Efficiency
$\hat{\bar{Y}}_2$	100.00
T_D^*	103.91
T^*	241.677

From the above table 1, it is clear that the proposed estimator is more efficient than existing estimators.

REFERENCES

[1] Biradar, R.S. and Singh H.P. 2001. Successive Sampling Using Auxiliary Information on Both the Occasions, Calcutta Statistical Association Bulletin, 51, 203-204.

[2] Dubey, V. and Shukla A.K. 2020. Some modified estimators of population mean on two occasions, Advances in Mathematics. 9, 6457-6465.

[3] Jessen, R. J. 1942. Statistical investigation of a farm survey for obtaining farm facts. Iowa Agricultural Station Research Bulleting. 304, 54-59.

[4] Singh, G.N. and Singh V.K. 2001. On the use of auxiliary information in successive sampling, Journal of Indian Society of Agriculture and Statistics, 54(1), 1-12.

[5] Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Ashok, C. 1984. Sampling theory of survey applications, ISAS publication, New Delhi.