Estimation of Average Paddy Production of Pira Nagar Village at Barabanki District in India

S. K. Yadav*, Shanya Baghel, Sugandha Saxena and Avinash Kumar Singh

Department of Statistics, Babasaheb Bhimrao Ambedkar University, Lucknow, India
E Mail: drskystats@gmail.com; baghel.shanya@gmail.com; saxenasugandha0911@gmail.com; avinashsingh543@gmail.com
Corresponding Author

Received 09 August 2019; Accepted 10 June 2020;
Publication 16 October 2020

Abstract

The idea of the present paper is the use of the known information on study variable for enhanced estimation of average paddy production of Pira Nagar village at Barabanki District in India under the Simple Random Sampling Scheme. This known information is utilzed in the form of median of primary variable as it is readily available and does not require every unit of the population to be inquired. The Bias and MSE of the suggested estimator are derived up to approximation of degree one.The minimum value of the MSE of suggested estimator is also obtained by optimizing the characterizing scalar. The MSE has also been compared with the considered competing estimators both theoretically and empirically. The theoretical efficiency conditions of the suggested estimator to be better than the considered estimators are verified using natural population on primary data collected from Pira Nagar Village at Barabanki District of Uttar Pradesh state in India.

Keywords: Study variable, median of study variable, simple random sampling, bias, MSE, PRE.

Journal of Reliability and Statistical Studies, Vol. 13, Issue 1 (2020), 127–148.
doi: 10.13052/jrss2229-5666.1317
© 2020 River Publishers

1 Introduction

Cochran (1940) was the first who used the auxiliary information for enhancing the efficiency of the estimator of population mean. Since then a number of researchers have modified the usual ratio estimator, utilizing the auxiliary information in the form of its various parameters. Taking data on auxiliary variable requires extra cost. Sometimes it is not feasible to bear the extra cost and sometimes it becomes impossible to get information on auxiliary variable. In such situations we need to find an alternative to the auxiliary information. One of the solutions for such situations is to obtain some easily accessible chracteristics on study variable itself as supplementary information. For example Let the study variable is monthly salary of the workers working at a place. Most of the workers are unwilling to reveal their exact salary or in case of free lancers they don't have a fixed salary. But they can tell whether their salary lies between 10000–15000, 15000–20000 and so on. Hence it is easier for them to tell the range within which their earning lies. In such kind of situations median can be obtained for the whole data and can be used for elevated estimation of average salary of the workers. In the present investigation, we make use of the known median of the main variable to obtain the estimate of average production of paddy crop at Pira Nagar Village of Barabanki District of Uttar Prdesh state in India.

Let $U = U_1, U_2, \ldots, U_N$ be population containing N units which are distinct and may be identified. The problem is to estimate the population mean $\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ of the main variable Y with higher efficiency. The most suitable estimator for $\overline{Y}$ is the sample mean $\overline{y}$. When it costs high to get auxiliary information, we may consider additional information on Y and may suggest a modified ratio type estimator for enhanced estimation of $\overline{Y}$. If we go for biased estimator, we can obtain much lesser MSE than the variance of $\overline{y}$ and MSE/Variance of other existing biased and unbiased estimators of $\overline{Y}$.

1.1 Notation

N:	Population Size
n:	Sample Size
f: $\frac{n}{N}$**:**	Sampling Fraction
$^{N}C_n$**:**	All possible samples of size n
Y:	Study variable
M:	Median of the Y
X:	Auxiliary variable

$\overline{Y}, \overline{X}$:	Population means
$\overline{y}, \overline{x}$:	Sample means
ρ:	Correlation coefficient between X and Y
β:	Regression coefficient of Y on X
β_1:	Coefficient of Skewness of X
β_2:	Coefficient of Kurtosis of X
$\overline{M}$:	Average of sample medians of Y
m:	Sample median of Y
Q_r:	Interquartile range
B(·):	Bias of the estimator
V(·):	Variance of the estimator
MSE(·):	Mean squared error of the estimator
Q_1, Q_2, Q_3:	Quartiles of X
$C_y, C_x,\ C_m$:	Coefficient of variation of x, y and m respectively
C_{yx}, C_{ym}:	Relative Covariances
	$\lambda : \frac{1-f}{n}$
PRE(e,p):	Percentage relative efficiency of the proposed
$\frac{MSE(e)}{MSE(p)} \times 100$:	estimator(p) with respect to the existing estimator (e)

1.2 Formulae

Variance of Study Variable:

$$\boldsymbol{V(\overline{y}) = \frac{1}{\binom{N}{n}} \sum_{i=1}^{\binom{N}{n}} (y_i - \overline{Y})^2 = \frac{1-f}{n} S_y^2}$$

Variance of Auxiliary variable:

$$\boldsymbol{V(\overline{x}) = \frac{1}{\binom{N}{n}} \sum_{i=1}^{\binom{N}{n}} (x_i - \overline{X})^2 = \frac{1-f}{n} S_x^2}$$

where

$$S_y^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \overline{Y})^2 \quad \text{and} \quad S_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{X})^2$$

Mean of medians of possible samples

$$\overline{M} = \frac{1}{\binom{N}{n}} \sum_{i=1}^{\binom{N}{n}} m_i$$

Variance of Sample Median of Y

$$V(m) = \frac{1}{\binom{N}{n}} \sum_{i=1}^{\binom{N}{n}} (m_i - M)^2$$

Covariance of $\overline{y}$ and $\overline{x}$

$$Cov(\overline{y}, \overline{x}) = \frac{1}{\binom{N}{n}} \sum_{i=1}^{\binom{N}{n}} (y_i - \overline{Y})(x_i - \overline{X}) = \frac{1-f}{n} S_{yx}$$

where

$$S_{yx} = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \overline{Y})(x_i - \overline{X})$$

Covariance of $\overline{y}$ and m

$$\mathrm{Cov}\,(\overline{\mathrm{y}}, \mathrm{m}) = \frac{1}{\binom{N}{n}} \sum_{\mathrm{i}=1}^{\binom{N}{n}} \left(\mathrm{y_i} - \overline{\mathrm{Y}}\right)(\mathrm{m_i} - \mathrm{M})$$

Coefficient of Variations

$$C_{xx} = \frac{V(\overline{x})}{\overline{X}^2} = {C_x}^2$$

$$C_{yy} = \frac{V(\overline{y})}{\overline{Y}^2} = {C_y}^2$$

$$C_{mm} = \frac{V(m)}{M^2} = {C_m}^2$$

$$C_{ym} = \frac{Cov(\overline{y}, m)}{M\overline{Y}}$$

$$C_{yx} = \frac{Cov(\overline{y}, \overline{x})}{\overline{Y}\overline{X}}$$

2 Literature Review of Existing Estimators

Under this section, various estimators of $\overline{\mathrm{Y}}$ along with their biases and MSEs are presented in Table 1. It is well known that in simple random sampling technique ($\overline{\mathrm{y}}$, $\overline{\mathrm{x}}$) are unbiased estimators for ($\overline{\mathrm{Y}}$, $\overline{\mathrm{X}}$) respectively.

Table 1 Various existing estimators of $\overline{\boldsymbol{Y}}$ along with their biases & MSEs

S.No.	Estimators	Bias	MSE/Variance
1	$t_0 = \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ Sample Mean	–	$\frac{1-f}{n}\overline{Y}^2 C_y^2$
2	$\widehat{\overline{Y}}_{lr} = \overline{y} + \beta(\overline{X} - \overline{x})$ Watson(1937)	–	$\frac{1-f}{n}\overline{Y}^2 C_y^2(1-\rho^2)$
3	$t_1 = \overline{y}\left(\frac{\overline{X}}{\overline{x}}\right)$ Cochron(1940)	$\frac{1-f}{n}\overline{Y}\left(C_x^2 - \rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2\left(C_y^2 + C_x^2 - 2\rho C_y C_x\right)$
4	$t_2 = \overline{X}\overline{r}$ $t_2' = \overline{X}\overline{r} + \frac{n(N-1)}{N(n-1)}(\overline{y} - \overline{r}\overline{x})$ Goodman(1958)	–	$\frac{1-f}{n}\overline{Y}^2\left(C_y^2 + C_x^2 - 2\rho C_y C_x\right)$
5	$t_3 = (1-\alpha)\overline{y} + \alpha\overline{y}\frac{\overline{X}}{\overline{x}}$ Chakrabarty (1979)	$\frac{1-f}{n}\overline{Y}\left(\frac{\alpha}{2}C_x^2 - \alpha\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2\left(C_y^2 + \alpha^2 C_x^2 - 2\alpha\rho C_y C_x\right)$
6	$t_4 = \overline{y}\left\{2 - \left(\frac{\overline{x}}{\overline{X}}\right)^{w}\right\}$ Sahai (1980)	$\frac{1-f}{n}\overline{Y}(-\frac{w(1-w)}{2}C_x^2 - w\rho C_y C_x)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + w^2 C_x^2 - 2w\rho C_y C_x)$
7	$t_5 = \overline{y}\left(\frac{\overline{X} + C_x}{\overline{x} + C_x}\right)$ Sisodia (1981)	$\frac{1-f}{n}\overline{Y}(R_5^2 C_x^2 - R_5\rho C_y C_x)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_5^2 C_x^2 - 2R_5\rho C_y C_x)$
8	$t_6 = \overline{y}\exp\left(\frac{\overline{X} - \overline{x}}{\overline{X} + \overline{x}}\right)$ Tuteja(1991)	$\frac{1-f}{8n}\overline{Y}\left(3C_x^2 - 4\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2\left(C_y^2 + \frac{C_x^2}{4} - \rho C_y C_x\right)$
10	$t_8 = \overline{y}\exp\left(\frac{\overline{X}\beta_2 - C_x}{\overline{x}\beta_2 + C_x}\right)$ Upadhyaya(1999)	$\frac{1-f}{n}\overline{Y}\left(R_8^2 C_x^2 - R_8\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2\left(C_y^2 + R_8^2 C_x^2 - 2R_8\rho C_y C_x\right)$

(*Continued*)

Table 1 Continued

S.No.	Estimators	Bias	MSE/Variance
11	$t_9 = \overline{y}\left(\frac{\overline{X}^2}{\overline{x}^2}\right)$ Kadilar(2003)	$\frac{1-f}{n}\overline{Y}\left(3C_x^2 - 2\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2\left(C_y^2 + 4C_x^2 - 4\rho C_y C_x\right)$
12	$t_{10} = \overline{y}\left(\frac{\overline{X}\beta_1 + S_x}{\overline{x}\beta_1 + S_x}\right)$ Singh(2003)	$\frac{1-f}{n}\overline{Y}\left(R_{(10)}^2 C_x^2 - R_{(10)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(10)}^2 C_x^2 - 2R_{(10)}\rho C_y C_x)$
13	$t_{11} = \overline{y}\left(\frac{\overline{X} + \rho}{\overline{x} + \rho}\right)$ Singh (2003)	$\frac{1-f}{n}\overline{Y}\left(R_{(11)}^2 C_x^2 - R_{(11)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(10)}^2 C_x^2 - 2R_{(10)}\rho C_y C_x)$
14	$t_{12} = \overline{y}\left(\frac{\overline{X} + \beta_2}{\overline{x} + \beta_2}\right)$ Singh et al.	$\frac{1-f}{n}\overline{Y}\left(R_{(11)}^2 C_x^2 - R_{(11)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(12)}^2 C_x^2 - 2R_{(12)}\rho C_y C_x)$
15	$t_{13} = \overline{y}\left(\frac{\overline{X} + \beta_1}{\overline{x} + \beta_1}\right)$ Yan and Tian (2010)	$\frac{1-f}{n}\overline{Y}\left(R_{(13)}^2 C_x^2 - R_{(13)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(13)}^2 C_x^2 - 2R_{(13)}\rho C_y C_x)$
16	$t_{14} = \overline{y}\left(\frac{\overline{X}\beta_1 + \beta_2}{\overline{x}\beta_1 + \beta_2}\right)$ Yan and Tian (2010)	$\frac{1-f}{n}\overline{Y}\left(R_{(14)}^2 C_x^2 - R_{(14)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(14)}^2 C_x^2 - 2R_{(14)}\rho C_y C_x)$
17	$t_{15} = \overline{y}\left(\frac{\overline{X}C_x + \beta_1}{\overline{x}C_x + \beta_1}\right)$ Yan and Tian (2010)	$\frac{1-f}{n}\overline{Y}\left(R_{(15)}^2 C_x^2 - R_{(15)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(15)}^2 C_x^2 - 2R_{(15)}\rho C_y C_x)$
18	$t_{16} = \overline{y}\left(\frac{\overline{X}\beta_2 + \beta_1}{\overline{x}\beta_2 + \beta_1}\right)$ Yan and Tian (2010)	$\frac{1-f}{n}\overline{Y}\left(R_{(16)}^2 C_x^2 - R_{(16)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(16)}^2 C_x^2 - 2R_{(16)}\rho C_y C_x)$

19	$t_{17} = \overline{y}\frac{\overline{X}}{\overline{x}}\left(1 - \frac{k\overline{x}s^2}{n\overline{x}^3}\right)^{-1}$ Pandey (2011)	$\frac{1-f}{n}\overline{Y}\left(kC_x^2 - \rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + k^2C_x^2 - 2k\rho C_y C_x)$
20	$t_{18} = \overline{y}\left(\frac{\overline{X} + Q_3}{\overline{x} + Q_3}\right)$ Al-Omari (2012)	$\frac{1-f}{n}\overline{Y}\left(R_{(18)}^2 C_x^2 - R_{(18)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(18)}^2 C_x^2 - 2R_{(18)}\rho C_y C_x)$
21	$t_{19} = \overline{y}\left(\frac{\overline{X} + Q_r}{\overline{x} + Q_r}\right)$ Al-Omari (2012)	$\frac{1-f}{n}\overline{Y}\left(R_{(19)}^2 C_x^2 - R_{(19)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(19)}^2 C_x^2 - 2R_{(19)}\rho C_y C_x)$
22	$t_{20} = \overline{y}\left(\frac{\overline{X} + M_d}{\overline{x} + M_d}\right)$ Subramani and Kumarpandiyan (2012)	$\frac{1-f}{n}\overline{Y}\left(R_{(20)}^2 C_x^2 - R_{(20)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(20)}^2 C_x^2 - 2R_{(20)}\rho C_y C_x)$
23	$t_{21} = \overline{y}\left(\frac{\overline{X}C_x + M_d}{\overline{x}C_x + M_d}\right)$ Subramani and Kumarpandiyan (2012)	$\frac{1-f}{n}\overline{Y}\left(R_{(21)}^2 C_x^2 - R_{(21)}\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2 + R_{(21)}^2 C_x^2 - 2R_{(21)}\rho C_y C_x)$
24	$t_{22} = \overline{y}\left(\frac{\overline{X}}{\overline{x}}\right)^{1/2}$ Swain (2014)	$\frac{1-f}{8n}\overline{Y}\left(3C_x^2 - 4\rho C_y C_x\right)$	$\frac{1-f}{n}\overline{Y}^2\left(C_y^2 + \frac{C_x^2}{4} - \rho C_y C_x\right)$
25	$t_{23} = \alpha' + t_{Re} + (1-\alpha')t_{Pe}$ Yadav and Mishra (2015)	$\overline{Y}[\frac{1}{8}(4\alpha - 1 - 7f + 4f^2)C_x^2 - \frac{(2\alpha - 1)}{2(1-f)}\rho C_y C_x]$	$\frac{\overline{Y}^2(1-f)}{n}\left(C_y^2 + \alpha^2\frac{C_x^2}{4} - \alpha\rho C_y C_x\right)$

(*Continued*)

Table 1 Continued

S.No.	Estimators	Bias	MSE/Variance
26	$t_{24}=\overline{y}\left(\frac{\overline{X}+n}{\overline{x}+n}\right)$ Jerjuddin and Kishun (2016)	$\frac{1-f}{n}\overline{Y}\left(R_{(24)}^2C_x^2-R_{(24)}\rho C_yC_x\right)$	$\frac{1-f}{n}\overline{Y}^2(C_y^2+R_{(24)}^2C_x^2-2R_{(24)}\rho C_yC_x)$
27	$t_{25}=\overline{y}\left(\frac{\overline{X}+C_x}{\overline{x}+C_x}\right)^{b_1}$ Soponviwatkul and Lawson (2017)	$\frac{1-f}{n}\overline{Y}\left[\frac{b_1(b_1+1)}{2}R_5^2C_x^2-b_1R_5\rho C_yC_x\right]$	$\frac{1-f}{n}\overline{Y}^2C_y^2(1-\rho^2)$
28	$t_{26}=\overline{y}\left(\frac{\overline{X}+\rho}{\overline{x}+\rho}\right)^{b_2}$ Soponviwatkul and Lawson (2017)	$\frac{1-f}{n}\overline{Y}\left[\frac{b_2\,(b_2+1)}{2}R_{11}^2C_x^2-b_2R_{11}\rho C_yC_x\right]$	$\frac{1-f}{n}\overline{Y}^2C_y^2(1-\rho^2)$
29	$t_{27}=\omega_1\overline{y}+(1-\omega_1)(\overline{y}\frac{\overline{X}}{\overline{x}})$ Ijaz and Ali (2018)	$\frac{1-f}{n}\overline{Y}\rho C_y(C_x-\rho)$	$\frac{1-f}{n}\overline{Y}^2C_y^2(1-\rho^2)$
30	$t_{28}=$ $\omega_2\overline{y}+(1-\omega_2)(\overline{y}\mathrm{exp}\frac{\overline{X}-\overline{x}}{\overline{x}-\overline{X}})$ Ijaz and Ali (2018)	$\frac{1-f}{n}\overline{Y}\rho C_y\left(\frac{1}{4}C_x-\rho C_y\right)$	$\frac{1-f}{n}\overline{Y}^2C_y^2(1-\rho^2)$
31	$t_{29}=\overline{y}\left(\frac{ab\overline{X}+cd}{ab\overline{x}+cd}\right)$ Yadav et al. (2019)	–	$\frac{1-f}{n}\overline{Y}^2\left(C_y^2-\frac{C_{yx}^2}{C_x^2}\right)$

Note: We will denote our measures in terms of Coefficient of Variation because it is a relative measure and most suitable to compare two series.
where,

$$R_5 = \frac{\overline{X}}{\overline{X} + C_x}, R_7 = \frac{\overline{X}C_x}{\overline{X}C_x + \beta_2}, R_8 = \frac{\overline{X}\beta_2}{\overline{X}\beta_2 + C_x}, R_{10} = \frac{\overline{X}\beta_1}{\overline{X}\beta_1 + S_x},$$

$$R_{11} = \frac{\overline{X}}{\overline{X} + \rho}, R_{12} = \frac{\overline{X}}{\overline{X} + \beta_2}, R_{13} = \frac{\overline{X}}{\overline{X} + \beta_1}, R_{14} = \frac{\overline{X}\beta_1}{\overline{X}\beta_1 + \beta_2},$$

$$R_{15} = \frac{\overline{X}C_x}{\overline{X}C_x + \beta_1}, R_{16} = \frac{\overline{X}\beta_2}{\overline{X}\beta_2 + \beta_1}, k = \rho\frac{C_y}{C_x}, R_{18} = \frac{\overline{X}}{\overline{X} + Q_3},$$

$$R_{19} = \frac{\overline{X}}{\overline{X} + Q_r}, R_{20} = \frac{\overline{X}}{\overline{X} + M_d}, R_{21} = \frac{\overline{X}C_x}{\overline{X}C_x + M_d}, \alpha' = \frac{\rho}{4}\frac{C_y}{C_x},$$

$$\omega_1 = w/R_5, \omega_2 = w/R_{11}, w = \rho\frac{C_y}{C_x}$$

a,b,c,d = Constants or Parametric Value

3 Proposed Estimator Based on Median

If the median **M** of **Y** is known so by utilizing it, we may suggest an elevated estimator of $\overline{Y}$ as,

$$\boldsymbol{t} = \overline{\boldsymbol{y}} + \boldsymbol{\alpha}\log\frac{\boldsymbol{m}}{\boldsymbol{M}} \tag{1}$$

Where α is chosen such that MSE(t) is minimum. The bias and MSE of t up to approximation of order one is given by,

$$B(t) = \frac{\alpha B(m)}{M} - \frac{\alpha}{2}\lambda C_m^2 \tag{2}$$

$$MSE(t) = \lambda\overline{Y}^2 C_y^2 + \lambda\alpha^2 C_m^2 + 2a\overline{Y}\lambda C_{y_m} \tag{3}$$

where $\lambda = \frac{1-f}{n}$

Hence

$$\alpha_{min} = \frac{-\overline{Y}C_{y_m}}{C_m^2} \tag{4}$$

$$\text{and } MSE_{min}(t) = \lambda\overline{Y}^2\left(C_y^2 - \frac{C_{Y_m}^2}{C_m^2}\right) \tag{5}$$

3.1 Proposed Estimator Based on Median: Detailed Study

Let

$$e_0 = \frac{\overline{y} - \overline{Y}}{\overline{Y}} \quad \text{and} \quad e_1 = \frac{m - M}{M}$$
$$\text{then } \overline{y} = \overline{Y}(1 + e_0) \quad \text{and} \quad m = M(1 + e_1)$$

So,

$$E\left(e_0\right) = E\left(\frac{\overline{y} - \overline{Y}}{\overline{Y}}\right) = \frac{E\left(\overline{y}\right) - \overline{Y}}{\overline{Y}} = 0 \quad \text{or} \quad E(e_0) = 0 \tag{6}$$

$$E\left(e_1\right) = E\left(\frac{m - M}{M}\right) = \frac{E\left(m\right) - M}{M} \quad \text{or}$$

$$E(e_1) = \frac{\overline{M} - M}{M} = \frac{B(m)}{M} \tag{7}$$

$$E(e_0^2) = E\left(\frac{\overline{y} - \overline{Y}}{\overline{Y}}\right)^2 = \frac{E\left(\overline{y} - \overline{Y}\right)^2}{\overline{Y}^2} = \frac{V(\overline{y})}{\overline{Y}^2} = \frac{1-f}{n} C_y^2$$

$$E(e_0^2) = \lambda C_y^2 \tag{8}$$

$$E(e_1^2) = E\left(\frac{m - M}{M}\right)^2 = \frac{E(m - M)^2}{M^2} = \frac{V(m)}{M^2} = \frac{1-f}{n} C_m^2$$

$$E(e_1^2) = \lambda C_m^2 \tag{9}$$

$$E(e_0 e_1) = E[(\frac{(\overline{y} - \overline{Y})}{\overline{Y}})(\frac{m - M}{M})] = \frac{Cov(\overline{y}, m)}{\overline{Y} M} = \lambda C_{y_m}$$

$$E(e_0 e_1) = C_{y_m} \tag{10}$$

Hence the Estimator can be rewritten as,

$$t = \overline{Y} + \overline{Y} e_0 + \alpha = \left(e_1 - \frac{{e_1}^2}{2}\right) \tag{11}$$

For Biasedness,

$$t = \overline{Y} + \overline{Y} e_0 + a(e_1 - \frac{{e_1}^2}{2})$$
$$(t - \overline{Y}) = \overline{Y} e_0 + a e_1 - \frac{a {e_1}^2}{2}$$
$$E(t - \overline{Y}) = \overline{Y} E(e_0) + a E(e_1) - \frac{a}{2} E({e_1}^2)$$
$$B(t) = 0 + a \frac{B(m)}{M} - \frac{a}{2} \lambda C_m^2$$

(from Equation (6), (7) and (8))

Hence Biasedness is :

$$\boldsymbol{B(t) = \alpha\frac{B(m)}{M} - \frac{\alpha}{2}\lambda C_m^2} \tag{12}$$

For MSE,

$$\begin{aligned}
MSE(t) &= E(t-\overline{Y})^2\\
&= E\big(\overline{Y}e_0 + \alpha e_1\big)^2\\
&\text{(second and higher order terms are ignored)}\\
&= E\left(\overline{Y}^2 e_0^2 + \alpha^2 e_1^2 + 2\overline{Y}\alpha e_0 e_1\right)\\
&= \overline{Y}^2 E(e_0^2) + \alpha E(e_1^2) + 2\alpha E(e_0 e_1)
\end{aligned}$$

Hence,

$$\boldsymbol{MSE(t) = \lambda\overline{Y}^2 C_y^2 + \lambda\alpha^2 C_m^2 + 2\alpha\overline{Y}\lambda C_{y_m}} \tag{13}$$

(from equation (8), (9) and (10))

Minimum Value of a

For Minimum value of α, we should have

$$\frac{\partial MSE(t)}{\partial a} = 0 \quad \text{and} \quad \frac{\partial^2 MSE(t)}{\partial \alpha^2} > 0$$

So,

$$\frac{\partial MSE(t)}{\partial \alpha} = 02\alpha\lambda C_m^2 + 2\overline{Y}\lambda C_{y_m} = 0$$

$$\boldsymbol{\alpha_{min} = \frac{-\overline{Y}C_{y_m}}{C_m^2}} \tag{14}$$

For Minima,

$$\frac{\partial^2 MSE(t)}{\partial \alpha^2} = 2\lambda C_m^2 > 0$$

Hence a has a minimum value.

Minimum Value of MSE

Hence Minimum value of MSE(t) is obtained by putting the Value of α in Equation (3.3)

$$MSE_{min}(t) = \lambda\overline{Y}^2 C_y^2 + \lambda\overline{Y}^2\frac{C_{y_m}^2 C_m^2}{C_m^4} - 2\frac{\overline{Y}C_{y_m}}{C_m^2}\overline{Y}\lambda C_{y_m}$$

$$= \lambda \overline{Y}^2 C_y^2 + \lambda \overline{Y}^2 \frac{C_{y_m}^2}{C_m^2} - 2\lambda \frac{\overline{Y}^2 C_{y_m}^2}{C_m^2}$$

$$\boldsymbol{MSE_{min}(t) = \lambda \overline{Y}^2 \left(C_y^2 - \frac{C_{y_m}^2}{C_m^2} \right)} \tag{15}$$

4 Efficiency Comparison

In this section the suggested estimator is being compared with the competing estimators of $\overline{Y}$ and the efficiency conditions are presented in Table 2.

Table 2 Efficiency comparison

S.No.	$MSE(t) < MSE(\cdot)$	Condition
1	$MSE(t) \leq V(\overline{y})$	$\frac{C_{y_m}^2}{C_m^2} \geq 0$
2	$MSE(t) \leq MSE(t_{2'})$	$\frac{C_{y_m}^2}{C_m^2} \geq C_x(2\rho C_y - C_x)$
3	$MSE(t) \leq MSE(t_j), MSE(\overline{y}_{lr})$	$\frac{C_{y_m}}{C_m^2} \geq \rho^2 C_y^2; j = 3, 4, 25, \ldots, 29$
4	$MSE(t) \leq MSE(t_5)$	$\frac{C_{y_m}^2}{C_m^2} \geq 2R_5 C_x(2\rho C_y - \frac{R_5}{2} C_x)$
5	$MSE(t) \leq MSE(t_6)$	$\frac{C_{y_m}^2}{C_m^2} \geq C_x(\rho C_y - \frac{C_x}{4})$
6	$MSE(t) \leq MSE(t_j)$	$\frac{C_{y_m}^2}{C_m^2} \geq 2R_j C_x(2\rho C_y - \frac{R_j}{2} C_x);$ $j = 7, 8, 10 \ldots, 16, 18, \ldots, 21$
7	$MSE(t) \leq MSE(t_9)$	$\frac{C_{y_m}^2}{C_m^2} \geq 4C_x(\rho C_y - C_x)$
8	$MSE(t) \leq MSE(t_{17})$	$\frac{C_{y_m}}{C_m^2} \geq \rho^2 C_y^2$
9	$MSE(t) \leq MSE(t_{22})$	$\frac{C_{y_m}^2}{C_m^2} \geq C_x(\rho C_y - \frac{C_x}{4})$
10	$MSE(t) \leq MSE(t_{23})$	$\frac{C_{y_m}^2}{C_m^2} \geq \rho C_y C_x$
11	$MSE(t) \leq MSE(t_{24})$	$\frac{C_{y_m}^2}{C_m^2} \geq 2R_{24} C_x(\rho C_y - \frac{R_{24}}{2} C_x)$

5 Numerical Study

Under this section the efficiency conditions of the suggested estimator over competing estimators are verified using real data sets.

5.1 Data Collection

To verify the results we have obtained for the Paddy Production data from the Pira Nagar Village of Barabanki Distrcit.The details of the obtained data is as follows:

Village	Pira Nagar
District	Barabanki
Time	March 2018
Production	Paddy Production
Type of Data	Primary Data
Information Taken	Name of Resident
	Their Area of Cultivation (Unit in Hectares)
	Yield obtained for each area (Unit in Quintals):
	one Quintal=100 Kilogram

For our Numerical Justification we have taken:

Study Variable	Yield	denoted as Y
Auxiliary Variable	Area of Cultivation	denoted as X
Population Size	52	
Sample Size	3	

5.2 Population Parameters

Parameter	Value	Units in
N	52	–
n	3	–
${}^{N}C_n$	22100	–
$\overline{Y}$	14.721	Quintal
$\overline{X}$	0.46227	Hectare
M	10	Quintal
ρ	0.8046229	–
S_y^2	190.8668	Quintal
S_x^2	0.15675383	Hectare

Parameter	Value	Units in
S_{yx}^2	4.401155	–
C_y^2	0.8807379	–
C_x^2	0.7335474	–
C_{yx}	4.401155	–
β_1	8.1028894	–
β_2	14.146291	–
β	28.0769	–
Q_1	0.4040	Hectare
Q_3	0.5050	Hectare
Q_r	0.2525	Hectare
$\overline{M}$	12.0119	Quintal
V(m)	37.7429	Quintal
C_{ym}^2	0.2536	–
C_m^2	0.3774	–

5.3 Measurement on Population Parameters

On the basis of the data, The numerical values related to proposed estimators are obtained as follows:

t	14.022176
B(t)	1.5962096
a_{min}	9.8921161
$MSE_{min}(t)$	14.211079

5.4 Numerical Comparison

We have constructed a table for numerical comparison of the suggested estimator with the estimators in competition. The following table gives:

- The values of the suggested and competing estimators on the Basis of the data obtained and sample taken.
- The Biases of competing and suggested Estimators.
- Mean Square Error (MSE) of competing and Proposed Estimators.
- The percentage relative efficiencies (PRE) of the suggested over competing estimators of $\overline{Y}$.

Table 4 Biases, MSE's and respective PRE's

Estimator	Bias	MSE	PRE
t_0	0	$V(t_0)$ =59.951751	421.866
$\overline{y}_{lr}$	0	$V(\overline{y}_{lr})$=21.137907	148.74245
t_1	0.40139256	21.837168	153.66298
$t_2^{'}$	0	21.837168	153.66298
t_3	−1.1203056	21.137907	148.74245
t_4	−2.3629292	21.137907	148.74245
t_5	−0.63149751	35.223464	247.8592
t_6	−0.22328973	28.411333	199.92382
t_7	−0.078904053	57.591622	405.25862
t_8	0.0076015701	21.138228	148.74471
t_9	4.1946732	83.587598	588.18614
t_{10}	0.06976912	21.163736	148.9242
t_{11}	−0.63958709	34.472784	242.5768
t_{12}	−0.091234028	57.215612	402.6127
t_{13}	−0.15151962	55.345217	389.4512
t_{14}	−0.47740469	43.707413	307.5587
t_{15}	−0.086059435	57.373669	403.7249
t_{16}	−0.65903753	30.58859	215.2447
t_{17}	0.27660967	21.493713	151.2462
t_{18}	−0.65448695	29.277583	206.0194
t_{19}	−0.51534035	23.893597	168.13359
t_{20}	−0.62993932	27.185932	191.30097
t_{21}	−0.64921611	28.605068	201.2871
t_{22}	−0.22328973	28.411333	199.9238
t_{23}	1.7460142	15.928206	112.0830
t_{24}	−0.33881397	49.086159	345.4077
t_{25}	−0.7940365	21.137907	148.74245
t_{26}	−0.8804347	21.137907	148.74245
t_{27}	0.1809299	21.137907	148.74245
t_{28}	−1.888758	21.137907	148.74245
t_{29}	–	21.137907	148.74245
(t)	**1.5962096**	**14.211079**	

6 Results and Discussion

From the Table 4 it is observed from all the estimators that,

6.1 Biasednesses are ranging from -2.3629 to 4.1947.
6.2 Mean Square Errors are ranging from 14.2111 to 83.5875.
6.3 The PREs of the suggested estimator over estimators in competition are ranging from 112.0830 to 588.1861.

6.4 The sample mean ($\boldsymbol{t_0}$), Usual Regression Estimator (Watson,1937) $\boldsymbol{\overline{y}_{lr}}$, Goodman and Hartely's revised estimator (1958), $\boldsymbol{t_{2'}}$; are unbiased for $\boldsymbol{\overline{Y}}$.
6.5 Among the existing estimators, the MSE of the estimator of Yadav and Mishra(2015), $\boldsymbol{t_{23}}$ is minimum i.e. 15.9282 and MSE of the estimator of Kadilar and Kingi (2003), $\boldsymbol{t_9}$ is maximum i.e. 83.5875.
6.6 The suggested estimator is 1.12% more efficient than $\boldsymbol{t_{23}}$ and 5.88% more efficient than $\boldsymbol{t_9}$.
6.7 Since Efficiency is stronger property than the unbiasedness. Hence here we prefer the biased estimator with minimum MSE instead of unbiased estimator with higher MSE.
6.8 It is observed from data that the given below inequalities hold good;
6.9 $\boldsymbol{MSE\,(t) \leq MSE\,(t_{23}) \leq V\,(\overline{y}_{lr}) \leq MSE\,(t_1) \leq V\,(\overline{y})}$
6.10 The proposed estimator has minimum MSE and comes out to be more efficient than the other estimators which was our aim of study.

7 Conclusion

7.1 We have applied our proposed estimator successfully for the estimation of Average Paddy production. We can also use it for other agricultural areas and productions with larger sample size and population size.
7.2 In case the exact values are not known but we may know the ranges within which our values may be supposed to lie, this Median based estimator will give a precise result.
7.3 The use of this estimator can be extended to other fields also and can be used as an alternative of SRSWOR sample mean when exact values of characteristic under study are not known.
7.4 Whenever taking the auxiliary information involves very high cost we can use the proposed estimator as an alternative of Ratio and Regression type estimators.
7.5 It has been shown theoretically as well as numerically that the suggested estimator is better than the competing estimators for the estimating $\overline{Y}$.
7.6 Mostly the minimum MSE of any ratio type estimator is equal to the variance of the regression estimator but the suggested ratio type estimator has its MSE less than the variance of the usual regression estimator.
7.7 Therefore, the suggested estimator is highly recommended for the practical applications.

Acknowledgment

The authors are very much thankful to the Editior-in-Chief of JRSS and the learned referees for critically examining the paper which improved the quality of the earlier draft.

References

Al-Omari, A. I. (2012). Ratio estimation of the population mean using auxiliary information in simple random sampling and median ranked set sampling, Statistics and Probability Letters, 82, p. 1883–18890.

Bahl, S. and Tuteja, R.K. (1991). Ratio and product type exponential estimator, Information and Optimization Sciences, XII (I), p. 159–163.

Chakrabarty, R.P. (1979). Some ratio estimators, Journal of the Indian Society of Agricultural Statistics, 31(1), p. 49–57.

Cochran, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce, The Journal of Agric. Science, 30, p. 262–275.

Goodman, L. A. and Hartley, H.O. (1958).The precision of unbiased ratio-type estimators, Journal of the American Statistical Association, 53(282), p. 491–508.

Hartley, H.O. and Ross, A. (1954). Unbiased ratio estimators, Nature, August, 7, Vol. 174, p. 270–271.

Ijaz, M., and Ali, H. (2018). Some improved ratio estimators for estimating mean of finite population, Research and Reviews: Journal of Statistics and Mathematical Sciences, 4(2), 18–23.

Jerajuddin, M. and Kishun, J. (2016). Modified ratio estimators for population mean using size of the sample selected from population, IJSRSET, 2(2), p. 10–16.

Kadilar, C., and Cingi, H. (2003). A study on the chain ratio type estimator, Hacettepe Journal of Mathematics and Statistics, Vol. 32, p. 105–108.

Pandey, H., Yadav, S.K. and Shukla, A.K. (2011). An Improved General Class of Estimators, Estimating Population Mean using Auxiliary Information, International Journal of Statistics and Systems, Vol. 6(1), p. 1–7.

Quenouille, M. H. (1956). Notes on bias in estimation, Biometrika, 43, p. 353–60.

Sahai, A. and Ray, S.K. (1980). Efficient families of ratio and product type estimators, Biometrika, 67(1), p. 211–215.

Singh, G.N. (2003). On the improvement of product method of estimation in sample surveys. Jour. Ind. Soc. Agri. Statistics, 56(3), p. 267–275.

Singh, H. P., Tailor, R., Tailor, R. and Kakran, M. S. (2004): An improved estimator of population mean using power transformation, Journal of the Indian Society of Agricultural Statistics, 58(2), p. 223–230.

Singh, H.P. and Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean, Statistics in Transition, 6(4), p. 555–560.

Sisodia, B.V.S. and Dwivedi, V.K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable, Journal of the Indian Society of Agricultural Statistics, 33, p. 13–18.

Soponviwatkal, K., and Lawson, N. (2017). New ratio estimators for estimating population mean in simple random sampling using a coefficient of variation, correlation coefficient and a regression coefficient, Gazi University Journal of Science, 30(4), p. 610–621.

Subramani, J. and Kumarapandiyan, G. (2012a). Modified ratio estimator for population mean using median of the auxiliary variable, Proceedings of National Conference on Recent developments in the applications of Reliability Theory and Survival Analysis "Application of Reliability Theory and Survival Analysis", Bonfring Publisher, p. 224–231, ISBN:978-93-82338-25-3.

Subramani, J. and Kumarapandiyan, G. (2012b). Estimation of population mean using co-efficient of variation and median of an auxiliary variable. International Journal of Probability and Statistics, 1(4), p. 111–118.

Subramani, J. and Kumarapandiyan, G. (2012c): Modified ratio estimators for population mean using function of quartiles of auxiliary variable, Bonfring International Journal of Industrial Engineering and Management Science, 2(2), p. 19–23.

Sukhatme, P. V. and Sukhatme, B. V. (1970). Sample Theory of Surveys with Applications, Ames : Iowa State University Press.

Swain, A.K.P.C. (2014). On an improved ratio type estimator of finite population mean in sample surveys, Revista Investigacion Operacional, 35(1), p. 49–57.

Upadhyaya, L. N. and Singh, H. P. (1999). Use of transformed auxiliary variable in estimating the finite population mean, Biometrical Journal, 41(5), p. 627–636.

Watson, D.J. (1937). The estimation of leaf area in field crops, Jour. Agr. Sci., 27(3), p. 474–483.

Yadav, S.K. and Mishra, S.S. (2015). developing improved predictive estimator for finite population mean using auxiliary information, Statistika, 95(1), p. 76–85.

Yadav, S.K., Dixit, M. K., Dungana, H.N. and Mishra, S.S. (2019): Improved estimators for estimating average yield using auxiliary variable, International Journal of Mathematical, Engineering and Management Sciences, Vol. 4, No. 5, p. 1228–1238.

Yadav, S.K, Misra, S., Tiwari, V. and Shukla, A.K. (2010). Improved family of ratio-cum-dual to ratio estimators of finite population mean, IFRSA's Int. Journal of Computing, 1(1), 17–23.

Yadav, S.K, Misra, S., Tiwari, V. and Shukla, A.K. (2014). Use of coefficient of skewness and quartile deviation of auxiliary variable for improved estimation of population mean, International Journal of Engineering Sciences and Research Technology, 3(1), p. 404–410.

Yan, Z. and Tian, B. (2010). Ratio method to the mean estimation using coefficient of skewness of auxiliary variable, ICICA, Part II, CCIS, 106, p. 103–110.

Biographies

S. K. Yadav is a faculty in the Department of Statistics at Babasaheb Bhimrao Ambedkar University Lucknow, U.P., India. He earned his M.Sc. and Ph.D. degrees in Statistics from Lucknow University and qualified the National Eligibility Test. Dr. Yadav has published more than 40 research papers in national and international journals of repute indexed in Scopus/Web of science and two books from an international publisher. He is a referee for 20 reputed international journals. He has been awarded Young Scientist Award in

2016 for the contribution in the field survey sampling by Venus International Research Foundation, Chennai, India and best paper award in 2018 by MTMI, USA. He has presented papers in more than 20 national and international conferences and also delivered 18 invited talks in several conferences.

Shanya Baghel is pursuing her Ph.D in IIT(ISM) Dhanbad, Department of Mathematics and Computing as a JRF. She has completed her M.Phil. from Babasaheb Bhimrao Ambedkar University (BBAU), Lucknow, in 2020. She received her M.Sc. degree in Applied Statistics from BBAU, Lucknow in 2018. She attended University of Lucknow, India and received her B.Sc. in 2016. Her M.Phil work centers on Estimation of Population Mean using known Auxiliary Variable.

Sugandha Saxena is a postgraduate from Babasaheb Bhimrao Ambedkar University. She is an alumna of City Montessori School, Lucknow, India. She received a B.Sc. degree in Statistics and Mathematics from Mahila P.G. College, University of Lucknow. She completed her M.Sc. in Applied Statistics from Babasaheb Bhimrao Ambedkar University, Lucknow.

Avinash Kumar Singh received his M.Sc. degree in Applied Statistics form Babasaheb Bhimrao Ambedkar University, Lucknow, India. He attended the University of Lucknow, India and received his B.Sc. in Physics, Mathematics and Statistics.