DUAL OF RATIO ESTIMATORS OF FINITE POPULATION MEAN OBTAINED ON USING LINEAR TRANSFORMATION TO AUXILIARY VARIABLE

H. S. Jhajj*, M. K. Sharma* and Lovleen Kumar Grover**

The efficiencies of the ratio-type estimators have been increased by using linear transformation on auxiliary variable in the literature. But such type of estimators requires the additional knowledge of unknown population parameters, which restricts their applicability. Keeping in view such restrictions, we have proposed two unbiased estimators of population mean of study variable on applying linear transformation to auxiliary variable by using its extreme values in the population that are generally available in practice. The comparison of the proposed estimators with the existing ones have been done with respect to their variances. It has also been shown that the proposed estimators have greater applicability and are more efficient than the mean per unit estimator even when the existing estimators are less efficient. We have also shown that under some known conditions the choice of most efficient estimators among the considered ones can be made for a given population. The theoretical results obtained are shown diagrammatically and have been verified numerically by taking some empirical populations.

Keywords and phrases: Bias, efficiency, most efficient estimator, preference region, simple random sampling without replacement, unbiased estimator, variance.

1. Introduction

For estimating the population mean $\bar{Y}$ of the variable under study $y$, the ratio estimator $\bar{y}_R$ has been widely used when there is a high positive correlation between study variable $y$ and auxiliary variable $x$. In literature, it has been shown by various authors viz Mohanty and Das (1971), Reddy (1974), Reddy and Rao (1977), Srivenkataramana (1978), Chaudhuri and Adhikari (1979) that the bias and the mean square error of the ratio estimator $\bar{y}_R$ can be reduced with the application of transformation on the auxiliary variable $x$. By using such transformation on auxiliary variable, the construction of the estimator of population mean $\bar{Y}$ requires the knowledge of unknown parameters, which restrict the applicability of these estimators. To overcome such type of restrictions, Mohanty and Sahoo (1995) have defined two ratio estimators by making the transformation on auxiliary variable $x$, using its minimum value $X_m$ and the maximum value $X_M$ in the population, when the values $X_m$ and $X_M$ in addition to its population mean $\bar{X}$ are available in advance. The information of extreme values $X_m$ and $X_M$ is generally available in practice, otherwise it can be obtained approximately from


*Department of Statistics, Punjabi University, Patiala 147 002, India.

**Department of Mathematics, G.N.D. University, Amritsar 143 005, India. Email: lovleen_2@yahoo.co.in
either the past experience or pilot sample survey, inexpensively. On assuming auxiliary variable $x$ as positive variable, Mohanty and Sahoo (1995) made the following transformations on $x$ as

\[
\begin{align*}
\nu &= \frac{x + X_m}{X_M + X_m} \\
\omega &= \frac{x + X_M}{X_M + X_m}
\end{align*}
\]

(1.1)

From (1.1), we see that $\nu \in (0, 1]$ and $\omega \in [1, 2)$. It should be noted that the correlation coefficients for the bivariate $(y, \nu)$ and $(y, \omega)$ remain same as that of $(y, x)$.

Assuming that the variables $y$ and $x$ are positively correlated and prior information on population mean $\bar{X}_x$ of the auxiliary variable $x$ is available then ratio estimator $\bar{y}_R$ is defined as

\[
\bar{y}_R = \frac{\bar{y}}{\bar{x}}
\]

(1.2)

where $\bar{y}$ and $\bar{x}$ are sample means of variables $y$ and $x$ respectively.

Under the transformations as given in (1.1), Mohanty and Sahoo (1995) defined the following ratio estimators of $\bar{Y}$ by using the known values of population means $\bar{V}_v$ and $\bar{\Omega}_\omega$ of variables $v$ and $\omega$ respectively,

\[
t_{MS1} = \frac{\bar{y}}{\bar{v}} \bar{V}
\]

(1.3)

and

\[
t_{MS2} = \frac{\bar{y}}{\bar{\omega}} \bar{\Omega}
\]

(1.4)

where $\bar{v}$ and $\bar{\omega}$ denote the sample means of variables $v$ and $\omega$ respectively.

Srivenkataramana (1980) defined a dual to the conventional ratio estimator $\bar{y}_R$ as

\[
\bar{y}_{RD} = \frac{\bar{y}(N\bar{X} - n\bar{x})}{(N - n)\bar{X}}
\]

(1.5)

where $n$ and $N$ denote the sample size and the population size respectively. He has shown that the exact expression for the bias of the estimator $\bar{y}_{RD}$ can be obtained even for a positive correlation between variables $y$ and $x$ whereas the exact expression for the bias of the estimator $\bar{y}_R$ is not available.

In the present paper, we have proposed two unbiased estimators of population mean $\bar{Y}$, which are respectively dual to the ratio estimators $t_{MS1}$ and $t_{MS2}$. The expressions for their variances have been obtained. The comparisons of the proposed estimators with the existing estimators have been made with respect to their variances. The preference regions of various estimators have also been obtained. The results have also been illustrated diagrammatically as well as numerically.
2. Proposed estimators and their variances

Suppose a simple random sample of size \( n \) is drawn from a finite population of size \( N \) without replacement and observation on variable \( y \) and \( x \) are taken. When the values of \( X_m, X_M \) and \( \bar{X} \) are known in advance then under the transformation \( v = \frac{x+X_m}{X_m+X_M} \) and \( \omega = \frac{x+X_M}{X_M+X_m} \) (same as given in (1.1)), the values of \( \bar{V} \) and \( \bar{\Omega} \) will also be known. Since the variables \( y \) and \( x \) are positively correlated therefore by using the known values of \( \bar{V} \) and \( \bar{\Omega} \), we consider the following estimators of \( \bar{Y} \) which are dual to the ratio estimators \( t_{MS1} \) and \( t_{MS2} \) respectively

\[
\hat{y}_{D1} = \frac{\bar{y}(N\bar{V} - n\bar{v})}{(N-n)\bar{V}}
\]

and

\[
\hat{y}_{D2} = \frac{\bar{y}(N\bar{\Omega} - n\bar{\omega})}{(N-n)\bar{\Omega}}.
\]

The biases of estimator \( \hat{y}_{D1} \) and \( \hat{y}_{D2} \) are obtained as

\[
B(\hat{y}_{D1}) = -\frac{1}{N} \bar{Y} \rho_{yx} C_y \bar{X} C_x \frac{X_m + \bar{X}}{X_m + \bar{X}}
\]

(2.3)

and

\[
B(\hat{y}_{D2}) = -\frac{1}{N} \bar{Y} \rho_{yx} C_y \bar{X} C_x \frac{X_M + \bar{X}}{X_M + \bar{X}}.
\]

(2.4)

where \( C_y \) and \( C_x \) are coefficients of variations of variables \( y \) and \( x \) respectively; \( \rho_{yx} \) is the correlation coefficient between \( y \) and \( x \). The biases of estimators \( \hat{y}_{D1} \) and \( \hat{y}_{D2} \) given in (2.3) and (2.4) are constant so can be estimated on the basis of same sample. Hence the corresponding proposed unbiased estimators in place of \( \hat{y}_{D1} \) and \( \hat{y}_{D2} \) are

\[
\hat{y}_{D1}^* = \frac{\bar{y}(N\bar{V} - n\bar{v})}{(N-n)\bar{V}} + \frac{s_{xy}}{N(X_m + \bar{X})}
\]

(2.5)

\[
\hat{y}_{D2}^* = \frac{\bar{y}(N\bar{\Omega} - n\bar{\omega})}{(N-n)\bar{\Omega}} + \frac{s_{xy}}{N(X_M + \bar{X})}.
\]

(2.6)

Srivokkamana (1980) has also defined the unbiased estimator in place of \( \bar{y}_{RD} \) as

\[
\hat{y}_{RD}^* = \frac{\bar{y}(N\bar{X} - n\bar{x})}{(N-n)\bar{X}} + \frac{s_{xy}}{NX}.
\]

(2.7)

The results obtained are given in the following theorem.

**Theorem 2.1.** For the simple random sampling without replacement (SRSWOR), the variances of the proposed estimators \( \hat{y}_{D1}^* \) and \( \hat{y}_{D2}^* \), up to the terms of order \( n^{-1} \), are

\[
V(\hat{y}_{D1}^*) = \left( \frac{1}{n} - \frac{1}{N} \right) \bar{Y}^2 \left[ C_y^2 + \left( \frac{n}{N-n} \right)^2 \frac{X^2 C_x^2}{(X_m + \bar{X})^2} \right]
\]

\[
- 2 \left( \frac{n}{N-n} \right) \rho_{yx} C_y \frac{\bar{X} C_x}{X_m + \bar{X}}
\]

and

\[
V(\hat{y}_{D2}^*) = \left( \frac{1}{n} - \frac{1}{N} \right) \bar{Y}^2 \left[ C_x^2 + \left( \frac{n}{N-n} \right)^2 \frac{X^2 C_y^2}{(X_M + \bar{X})^2} \right]
\]

\[
- 2 \left( \frac{n}{N-n} \right) \rho_{xy} C_x \frac{\bar{X} C_y}{X_M + \bar{X}}.
\]
\[(2.9) \quad V(\hat{y}^*_D) = \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y}^2 \left[ C_y^2 + \left(\frac{n}{N-n}\right)^2 \frac{X^2 C_x^2}{(X_M + X)^2} \right. \]
\[- 2 \left(\frac{n}{N-n}\right) \rho_{yx} C_y \frac{\bar{X} C_x}{X_M + X} \].

**Remark.** It is noted that under the given transformation i.e. \(x\) to \(v\) or \(x\) to \(\omega\), as defined in (1.1), there is no change in the form of usual linear regression estimator \(\bar{y}_{lr} = \bar{y} + b_{yx} (\bar{X} - \bar{x})\).

3. **Comparison of proposed estimators with existing ones**

For comparing the proposed estimators with existing ones, we first write the expressions for the variances of the existing estimators, namely \(\bar{y}, \bar{y}_R, \bar{y}_RD, t_{MS1}\) and \(t_{MS2}\) (under the same sampling scheme) up to the terms of order \(n^{-1}\) (the biases of these estimators if exist is of order \(n^{-1}\), so their contributions to the mean square errors will be of order \(n^{-2}\)) as

\[(3.1) \quad V(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y}^2 C_y^2 \]
\[(3.2) \quad V(\bar{y}_R) = \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y}^2 (C_y^2 + C_x^2 - 2 \rho_{yx} C_y C_x) \]
\[(3.3) \quad V(\bar{y}_RD) = \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y}^2 \left[ C_y^2 + \left(\frac{n}{N-n}\right)^2 C_x^2 - 2 \left(\frac{n}{N-n}\right) \rho_{yx} C_y C_x \right] \]
\[(3.4) \quad V(t_{MS1}) = \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y}^2 \left[ C_y^2 + \frac{\bar{X}^2 C_x^2}{(X_m + X)^2} - 2 \rho_{yx} C_y \frac{\bar{X} C_x}{X_M + X} \right] \]
\[(3.5) \quad V(t_{MS2}) = \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y}^2 \left[ C_y^2 + \frac{\bar{X}^2 C_x^2}{(X_m + X)^2} - 2 \rho_{yx} C_y \frac{\bar{X} C_x}{X_M + X} \right]. \]

Reddy (1978) has shown that the values of parameter \(K = \rho_{yx} \frac{C_y}{C_x}\) remain stable in any repetitive survey. So we find the conditions on the values of \(K\) under which the proposed estimators are superior to the existing ones. For the present situation, we note that \(0 < K < \frac{C_y}{C_x}\). In the survey sampling situations, usually \(\frac{n}{N-n} < 0.1\) but we assume that \(\frac{n}{N-n} < 1\) which hold good in almost all the situations of survey sampling. For the sake of comparison in the compact form, we take

\[(3.6) \quad m_1 = 1 + \frac{X_m}{X} \]
\[(3.7) \quad m_2 = 1 + \frac{X_M}{X} \]
\[(3.8) \quad \ell = \frac{n}{N-n}. \]

Noting that \(m_1 \in (1, 2]\) and \(m_2\) is a finite number so that \(m_2 \geq 2\) and by assumption we note that \(0 < \ell < 1\).
Using the expressions (2.8), (2.9) and (3.1) to (3.5), the results obtained are given in the following theorems.

**Theorem 3.1.** Up to the terms of order $n^{-1}$, we have

\[ V(\hat{y}^*_{D1}) < V(\bar{y}) \text{ for } K > \frac{\ell}{2m_1} \]
\[ V(t_{MS1}) < V(\bar{y}) \text{ for } K > \frac{1}{2m_1} \]
\[ V(\hat{y}^*_{D2}) < V(\bar{y}) \text{ for } K > \frac{\ell}{2m_2} \]
\[ V(t_{MS2}) < V(\bar{y}) \text{ for } K > \frac{1}{2m_2} \]

**Remark.** From the results of above theorem, we see that the proposed estimators $\hat{y}^*_{D1}$ and $\hat{y}^*_{D2}$ which are dual to $t_{MS1}$ and $t_{MS2}$ respectively are superior to the mean per unit estimator $\bar{y}$ even for smaller values of $K$ when $t_{MS1}$ and $t_{MS2}$ are inferior than $\bar{y}$.

**Theorem 3.2.** Up to the terms of order $n^{-1}$, the variance of the proposed estimator $\hat{y}^*_{D1}$ is less than the variances of all the other estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_R$, $\bar{y}^*_{MS1}$, $\bar{y}^*_{MS2}$ and $\hat{y}^*_{D2}$ for $K \in I_1$, where

\[
I_1 = \begin{cases} 
  \left( \frac{\ell}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right), \frac{\ell}{2} \left( \frac{1}{m_1} + 1 \right) \right) & \text{when } 0 < \ell < \frac{1}{m_2} \\
  \left( \frac{\ell}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right), \frac{1}{2} \left( \frac{\ell}{m_1} + \frac{1}{m_2} \right) \right) & \text{when } \frac{1}{m_2} < \ell < \frac{m_1}{m_2} \\
  \left( \frac{1}{2} \left( \frac{\ell}{m_1} + \frac{1}{m_2} \right), \frac{\ell}{2} \left( \frac{1}{m_1} + 1 \right) \right) & \text{when } \frac{m_1}{m_2} < \ell < \frac{1}{m_1} \text{ with } m_1^2 < m_2 \\
  \left( \frac{1}{2} \left( \frac{\ell}{m_1} + \frac{1}{m_2} \right), \frac{1}{2m_1}(\ell + 1) \right) & \text{when } \frac{1}{m_1} < \ell < 1 \text{ with } m_1^2 < m_2 \\
  \text{or } \frac{m_1}{m_2} < \ell < 1 \text{ with } m_1^2 > m_2.
\end{cases}
\]

**Theorem 3.3.** Up to the terms of order $n^{-1}$, the variance of the proposed estimator $\hat{y}^*_{D2}$ is less than the variances of all the other estimators $\bar{y}$, $\bar{y}_R$, $\bar{y}^*_R$, $\bar{y}^*_{MS1}$, $\bar{y}^*_{MS2}$ and $\hat{y}^*_{D1}$ for $K \in I_2$, where

\[
I_2 = \begin{cases} 
  \left( \frac{\ell}{2m_2}, \frac{\ell}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right) \right) & \text{when } 0 < \ell < \frac{m_1}{m_2} \\
  \left( \frac{\ell}{2m_2}, \frac{1}{2m_2}(\ell + 1) \right) & \text{when } \frac{m_1}{m_2} < \ell < 1.
\end{cases}
\]
Theorem 3.4. Up to the terms of order $n^{-1}$, the estimator $\bar{y}$ will become most efficient among the estimators $\bar{y}, \bar{y}_R, \bar{y}_{RD}, t_{MS1}, t_{MS2}, \bar{y}_{D1}^*$ and $\bar{y}_{D2}^*$ for

$$K \in \left(0, \frac{\ell}{2m_2}\right) = I_3.$$  \hspace{1cm} (3.10)

Theorem 3.5. Up to the terms of order $n^{-1}$, the estimator $\bar{y}_R$ will become most efficient among the estimators $\bar{y}, \bar{y}_R, \bar{y}_{RD}, t_{MS1}, t_{MS2}, \bar{y}_{D1}^*$ and $\bar{y}_{D2}^*$ for $K \in I_4$, where

$$I_4 = \begin{cases} \left(\frac{1}{2} \left(1 + \frac{1}{m_1}\right), \frac{C_y}{C_x}\right) & \text{when } \ell < \frac{1}{m_1} \\ \left(\frac{1}{2} (1 + \ell), \frac{C_y}{C_x}\right) & \text{when } \ell > \frac{1}{m_1} \end{cases}.$$  \hspace{1cm} (3.11)

Theorem 3.6. Up to the terms of order $n^{-1}$, the estimator $\bar{y}_{RD}^*$ will become most efficient among the estimators $\bar{y}, \bar{y}_R, \bar{y}_{RD}, t_{MS1}, t_{MS2}, \bar{y}_{D1}^*$ and $\bar{y}_{D2}^*$ for $K \in I_5$, where

$$I_5 = \begin{cases} \left(\frac{1}{2} \left(1 + \frac{1}{m_1}\right), \frac{1}{2} \left(\ell + \frac{1}{m_2}\right)\right) & \text{when } 0 < \ell < \frac{1}{m_2} \\ \left(\frac{1}{2} \left(\ell + \frac{1}{m_2}\right), \frac{1}{2} \left(\ell + \frac{1}{m_1}\right)\right) & \text{when either } \frac{1}{m_2} < \ell < \frac{m_1}{m_2} \text{ with } m_1^2 < m_2 \\ \text{or } \frac{1}{m_2} < \ell < \frac{1}{m_1} \text{ with } m_1^2 > m_2 \\ \left(\frac{1}{2} \left(\ell + \frac{1}{m_1}\right), \frac{1}{2} (\ell + 1)\right) & \text{when } \frac{1}{m_1} < \ell < \frac{m_1}{m_2} \text{ with } m_1^2 > m_2 \\ \left(\frac{1}{2} \left(\ell + \frac{1}{m_1}\right), \frac{1}{2} (\ell + 1)\right) & \text{when } \frac{m_1}{m_2} < \ell < \frac{1}{m_1} \text{ with } m_1^2 < m_2 \\ \left(\frac{1}{2} \left(\ell + \frac{1}{m_1}\right), \frac{1}{2} (\ell + 1)\right) & \text{when either } \frac{1}{m_1} < \ell < 1 \text{ with } m_1^2 < m_2 \\ \text{or } \frac{m_1}{m_2} < \ell < 1 \text{ with } m_1^2 > m_2 \end{cases}.$$  \hspace{1cm} (3.12)
Theorem 3.7. Up to the terms of order $n^{-1}$, the estimator $t_{MS1}$ will become most efficient among the estimators $\bar{y}, \bar{y}_R, \bar{y}_{RD}^*, t_{MS1}, t_{MS2}, \hat{y}_D^*$ and $\hat{y}_D^*$ for $K \in I_6$, where

$$ I_6 = \begin{cases} 
\left( \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right), \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_1} \right) \right) & \text{when } 0 < \ell < \frac{1}{m_2} \\
\left( \frac{1}{2} \left( \ell + \frac{1}{m_1} \right), \frac{1}{2} \left( \frac{1}{m_1} + \ell \right) \right) & \text{when either } \frac{1}{m_2} < \ell < \frac{m_1}{m_2} \text{ with } m_1^2 < m_2 \\
or \frac{1}{m_2} < \ell < \frac{1}{m_1} \text{ with } m_1^2 > m_2 \\
\left( \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right), \frac{1}{2} \left( \ell + \frac{1}{m_1} \right) \right) & \text{when either } \frac{1}{m_1} < \ell < 1 \text{ with } m_1^2 > m_2 \\
or \frac{m_1}{m_2} < \ell < 1 \text{ with } m_1^2 < m_2 
\end{cases} $$

(3.13) $I_6 = \ldots$

Theorem 3.8. Up to the terms of order $n^{-1}$, the estimator $t_{MS2}$ will become most efficient among the estimators $\bar{y}, \bar{y}_R, \bar{y}_{RD}^*, t_{MS1}, t_{MS2}, \hat{y}_D^*$ and $\hat{y}_D^*$ for $K \in I_7$, where

$$ I_7 = \begin{cases} 
\left( \frac{1}{2} \left( \frac{1}{m_1} + \ell \right), \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right) \right) & \text{when } 0 < \ell < \frac{1}{m_2} \\
\left( \frac{1}{2} \left( \ell + \frac{1}{m_1} \right), \frac{1}{2} \left( \frac{1}{m_1} + \ell \right) \right) & \text{when either } \frac{1}{m_2} < \ell < \frac{m_1}{m_2} \text{ with } m_1^2 < m_2 \\
or \frac{1}{m_2} < \ell < \frac{1}{m_1} \text{ with } m_1^2 > m_2 \\
\left( \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right), \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right) \right) & \text{when either } \frac{1}{m_1} < \ell < 1 \text{ with } m_1^2 > m_2 \\
or \frac{m_1}{m_2} < \ell < 1 \text{ with } m_1^2 < m_2 \\
\left( \frac{1}{2} \left( \ell + \frac{1}{m_1} \right), \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right) \right) & \text{when either } \frac{1}{m_1} < \ell < 1 \text{ with } m_1^2 < m_2 \\
or \frac{m_1}{m_2} < \ell < 1 \text{ with } m_1^2 > m_2 
\end{cases} $$

(3.14) $I_7 = \ldots$
4. **Diagrammatic representation of preference regions of various estimators**

The preference region for the estimator is the interval of \( K \) on which the estimator is more efficient than the other estimators. To have a bird’s eye view of the preference regions for the different estimators obtained earlier in the Section 3, we have made an effort by representing them diagrammatically. For the sake of convenience to mark the limits of the preference regions in the diagrams, we take

\[
P_1 = \frac{l}{2m_2}, \quad P_2 = \frac{l}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right), \quad P_3 = \frac{l}{2} \left( 1 + \frac{1}{m_1} \right),
\]

\[
P_4 = \frac{1}{2} \left( \frac{1}{m_2} + l \right), \quad P_5 = \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right), \quad P_6 = \frac{1}{2} \left( \frac{1}{m_1} + 1 \right),
\]

\[
P_7 = \frac{1}{2} \left[ \frac{l}{m_1} + \frac{1}{m_2} \right], \quad P_8 = \frac{1}{2} \left( \frac{1}{m_1} + l \right), \quad P_9 = \frac{1}{2} (1 + l),
\]

\[
P_{10} = \frac{1}{2m_2} (1 + l), \quad P_{11} = \frac{1}{2m_1} (1 + l).
\]

The order of the various preference regions is shown in the following diagrams by considering the whole range of \( K \) under various situations.

**Diagram 4.1.** When \( 0 < \ell < \frac{1}{m_2} \) then we have

\[
0 \quad P_1 \quad P_2 \quad P_3 \quad P_4 \quad P_5 \quad P_6 \quad \frac{C_y}{C_5} \quad K
\]

**Diagram 4.2.** When either \( \frac{1}{m_2} < \ell < \frac{m_1}{m_2} \) with \( m_1^2 < m_2 \) or \( \frac{1}{m_2} < \ell < \frac{1}{m_1} \) with \( m_1^2 > m_2 \) then we have

\[
0 \quad P_1 \quad P_2 \quad P_7 \quad P_4 \quad P_8 \quad P_6 \quad \frac{C_y}{C_5} \quad K
\]

**Diagram 4.3.** When \( \frac{1}{m_1} < \ell < \frac{m_1}{m_2} \) with \( m_1^2 > m_2 \) then we have

\[
0 \quad P_1 \quad P_2 \quad P_7 \quad P_5 \quad P_8 \quad P_9 \quad \frac{C_y}{C_5} \quad K
\]
Diagram 4.4. When \( \frac{m_1}{m_2} < \ell < \frac{1}{m_1} \) with \( m_1^2 < m_2 \) then we have

\[
\begin{array}{cccccccc}
0 & P_1 & P_{10} & P_7 & P_3 & P_8 & P_6 & \frac{C_y}{C_x} \\
\end{array}
\]

Diagram 4.5. When either \( \frac{1}{m_1} < \ell < 1 \) with \( m_1^2 < m_2 \) or \( \frac{m_1}{m_2} < \ell < 1 \) with \( m_1^2 > m_2 \) then we have

\[
\begin{array}{cccccccc}
0 & P_1 & P_{10} & P_7 & P_{11} & P_8 & P_9 & \frac{C_y}{C_x} \\
\end{array}
\]

The above diagrams 4.1 to 4.5 clearly indicate that

(i) There is no advantage of using known auxiliary information for improving the estimator of \( \bar{Y} \) when the value of \( K \) is very small i.e. close to zero.

(ii) The use of ordinary ratio estimator \( \bar{y}_R \) is very beneficial for estimating \( \bar{Y} \) when the value of \( K \) is very large.

(iii) The proposed estimators \( \hat{y}^*_{D1} \) and \( \hat{y}^*_{D2} \) which are dual to \( t_{MS1} \) and \( t_{MS2} \) respectively are superior to the mean per unit estimator \( \bar{y} \) even for smaller values of \( K \) when \( t_{MS1} \) and \( t_{MS2} \) are inferior to \( \bar{y} \).

(iv) The interval \( I_3 \) is always adjacent to \( I_4 \) in the diagrams which shows that the estimator \( \hat{y}^*_{D2} \) is better than \( \bar{y} \) even when \( \bar{y}_R, \bar{y}^*_{RD}, t_{MS1}, t_{MS2} \) and \( \hat{y}^*_{D1} \) are not better than \( \bar{y} \).

Hence we conclude that the use of the proposed estimators \( \hat{y}^*_{D1} \) and \( \hat{y}^*_{D2} \) for estimating \( \bar{Y} \) will be more beneficial for moderate values of \( K \). So the choice of sample size can be made accordingly for which the proposed estimators \( \hat{y}^*_{D1} \) and \( \hat{y}^*_{D2} \) will become most efficient with known values of \( K, m_1 \) and \( m_2 \). On the other hand, for a given sample size, the choice of most efficient estimator among the considered estimators namely \( \bar{y}, \bar{y}_R, \bar{y}^*_{RD}, t_{MS1}, t_{MS2}, \hat{y}^*_{D1} \) and \( \hat{y}^*_{D2} \) can also be made accordingly with known values of \( K, m_1 \) and \( m_2 \).

5. An empirical study

To get a rough idea about the efficiencies of the estimators, we have taken seven empirical populations from the literature. The description of the populations and the values of population constants are given in Tables 1 and 2 respectively. The expressions of biases of ratio estimator and ratio estimators defined by Mohanty and Sahoo (1995) are given in Appendix A, so the biases of these estimators are calculated which are given in Table 3. The preference
**Table 1. Description of populations.**

<table>
<thead>
<tr>
<th>Population number</th>
<th>Source of population</th>
<th>( N )</th>
<th>Variable ( y )</th>
<th>Variable ( x )</th>
<th>( n )</th>
<th>( \ell )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Singh and Chaudhary (1986) p-166</td>
<td>16</td>
<td>Area under wheat during 1979–80</td>
<td>Total cultivated area during 1978–79</td>
<td>5</td>
<td>0.4545</td>
</tr>
<tr>
<td>2</td>
<td>Cochran (1977) p-325</td>
<td>10</td>
<td>Number of persons in a block</td>
<td>Number of rooms in a block</td>
<td>4</td>
<td>0.6667</td>
</tr>
<tr>
<td>3</td>
<td>Singh and Chaudhary (1986) p-306</td>
<td>10</td>
<td>Number of inhabitants ('000) in 1981–82</td>
<td>Number of inhabitants ('000) in 1980–81</td>
<td>4</td>
<td>0.6667</td>
</tr>
<tr>
<td>4</td>
<td>Singh and Chaudhary (1986) p-155</td>
<td>17</td>
<td>Number of milch animals in survey (1977–78)</td>
<td>Number of milch animals in census (1976)</td>
<td>6</td>
<td>0.5454</td>
</tr>
<tr>
<td>5</td>
<td>Singh and Chaudhary (1986) p-177(1–15)</td>
<td>15</td>
<td>Area under wheat in 1973</td>
<td>Area under wheat in 1971</td>
<td>5</td>
<td>0.5000</td>
</tr>
<tr>
<td>6</td>
<td>Sampford (1962) p-61(1–9)</td>
<td>9</td>
<td>Acreage under oats in 1957</td>
<td>Acreage of crops and grass in 1947</td>
<td>3</td>
<td>0.5000</td>
</tr>
<tr>
<td>7</td>
<td>Panse and Sukhatme (1967) p-124 (1–10)</td>
<td>10</td>
<td>Parental plot mean (mm)</td>
<td>Parental plant value (mm)</td>
<td>4</td>
<td>0.6667</td>
</tr>
</tbody>
</table>

**Table 2. Constants of the populations.**

<table>
<thead>
<tr>
<th>Population number</th>
<th>( \rho_{yx} )</th>
<th>( C_x )</th>
<th>( C_y )</th>
<th>( m_1 )</th>
<th>( m_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9600</td>
<td>0.74</td>
<td>0.6900</td>
<td>1.21</td>
<td>4.06</td>
</tr>
<tr>
<td>2</td>
<td>0.6500</td>
<td>0.14</td>
<td>0.1500</td>
<td>1.82</td>
<td>2.22</td>
</tr>
<tr>
<td>3</td>
<td>0.8800</td>
<td>0.60</td>
<td>0.6400</td>
<td>1.54</td>
<td>3.43</td>
</tr>
<tr>
<td>4</td>
<td>0.4371</td>
<td>0.02</td>
<td>0.0165</td>
<td>1.98</td>
<td>2.04</td>
</tr>
<tr>
<td>5</td>
<td>0.1732</td>
<td>0.82</td>
<td>0.8903</td>
<td>1.08</td>
<td>4.59</td>
</tr>
<tr>
<td>6</td>
<td>0.0700</td>
<td>0.10</td>
<td>0.2900</td>
<td>1.86</td>
<td>2.12</td>
</tr>
<tr>
<td>7</td>
<td>0.0833</td>
<td>0.07</td>
<td>0.0647</td>
<td>1.92</td>
<td>2.13</td>
</tr>
</tbody>
</table>

**Table 3. Biases of estimators.**

<table>
<thead>
<tr>
<th>Population number</th>
<th>[ \frac{B_i}{\sqrt{\frac{3}{5}}} \times 100 ]</th>
<th>( \bar{y}_R )</th>
<th>( t_{MS1} )</th>
<th>( t_{MS2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.742400</td>
<td>3.108596</td>
<td>8.751211</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.595000</td>
<td>0.158284</td>
<td>0.217170</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.208000</td>
<td>6.763232</td>
<td>6.791946</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.025576</td>
<td>0.002918</td>
<td>0.002541</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>54.5956</td>
<td>45.93969</td>
<td>0.436784</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.797000</td>
<td>0.179911</td>
<td>0.126744</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.452273</td>
<td>0.113272</td>
<td>0.090291</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Preference regions of superiority and percentage relative efficiencies with respect to $\bar{y}$ of various estimators.

<table>
<thead>
<tr>
<th>Population number</th>
<th>$K$</th>
<th>Preference region and efficiency, say $E$, of estimators</th>
<th>Most efficient estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8951</td>
<td>$\bar{y}$: (0, 0.0560) $\bar{y}<em>R$: (0.9132, 0.9324) $\bar{y}</em>{RD}$: (0.4151, 0.6405) $t_{MS1}$: (0.6405, 0.9132) $t_{MS2}$: (0.1791, 0.3110) $\bar{y}<em>{D1}$: (0.3110, 0.4151) $\bar{y}</em>{D2}$: (0.0560, 0.1791)</td>
<td>$t_{MS1}$ $\bar{y}<em>{RD}$ $\bar{y}<em>R$ $\bar{y}</em>{D1}$ $\bar{y}</em>{D2}$</td>
</tr>
<tr>
<td>2</td>
<td>0.6964</td>
<td>$\bar{y}$: (0, 0.1501) $\bar{y}<em>R$: (0.8333, 1.0714) $\bar{y}</em>{RD}$: (0.6080, 0.8333) $t_{MS1}$: (0.5000, 0.6080) $t_{MS2}$: (0.4084, 0.5000) $\bar{y}<em>{D1}$: (0.3333, 0.4084) $\bar{y}</em>{D2}$: (0.1501, 0.3333)</td>
<td>$\bar{y}<em>{RD}$ $\bar{y}<em>R$ $\bar{y}</em>{D1}$ $\bar{y}</em>{D2}$</td>
</tr>
<tr>
<td>3</td>
<td>0.9387</td>
<td>$\bar{y}$: (0, 0.0972) $\bar{y}<em>R$: (0.8333, 1.0667) $\bar{y}</em>{RD}$: (0.6580, 0.8333) $t_{MS1}$: (0.5411, 0.6580) $t_{MS2}$: (0.2429, 0.3622) $\bar{y}<em>{D1}$: (0.3622, 0.5411) $\bar{y}</em>{D2}$: (0.0972, 0.2429)</td>
<td>$\bar{y}<em>R$ $\bar{y}</em>{RD}$ $\bar{y}<em>{D1}$ $\bar{y}</em>{D2}$</td>
</tr>
<tr>
<td>4</td>
<td>0.3606</td>
<td>$\bar{y}$: (0, 0.1337) $\bar{y}<em>R$: (0.7727, 0.8250) $\bar{y}</em>{RD}$: (0.5252, 0.7727) $t_{MS1}$: (0.4976, 0.5252) $t_{MS2}$: (0.3828, 0.4976) $\bar{y}<em>{D1}$: (0.2714, 0.3828) $\bar{y}</em>{D2}$: (0.1337, 0.2714)</td>
<td>$\bar{y}<em>{RD}$ $\bar{y}</em>{D1}$ $\bar{y}_{D2}$</td>
</tr>
<tr>
<td>5</td>
<td>0.1880</td>
<td>$\bar{y}$: (0, 0.0545) $\bar{y}<em>R$: (0.9630, 1.0857) $\bar{y}</em>{RD}$: (0.4815, 0.7130) $t_{MS1}$: (0.7130, 0.9630) $t_{MS2}$: (0.1634, 0.3404) $\bar{y}<em>{D1}$: (0.3404, 0.4815) $\bar{y}</em>{D2}$: (0.0544, 0.1634)</td>
<td>$t_{MS1}$ $t_{MS2}$ $\bar{y}<em>{D1}$ $\bar{y}</em>{D2}$</td>
</tr>
<tr>
<td>6</td>
<td>0.2030</td>
<td>$\bar{y}$: (0, 0.1180) $\bar{y}<em>R$: (0.7688, 2.9000) $\bar{y}</em>{RD}$: (0.4858, 0.5188) $t_{MS1}$: (0.5188, 0.7688) $t_{MS2}$: (0.3702, 0.4858) $\bar{y}<em>{D1}$: (0.2523, 0.3702) $\bar{y}</em>{D2}$: (0.1180, 0.2523)</td>
<td>$\bar{y}<em>{D2}$ $\bar{y}</em>{D1}$ $\bar{y}_{D2}$</td>
</tr>
<tr>
<td>7</td>
<td>0.0770</td>
<td>$\bar{y}$: (0, 0.1565) $\bar{y}<em>R$: (0.8333, 0.9243) $\bar{y}</em>{RD}$: (0.5937, 0.8333) $t_{MS1}$: (0.4951, 0.5937) $t_{MS2}$: (0.4083, 0.4951) $\bar{y}<em>{D1}$: (0.3301, 0.4083) $\bar{y}</em>{D2}$: (0.1565, 0.3301)</td>
<td>$\bar{y}$</td>
</tr>
</tbody>
</table>

$E = 100$
regions of all the estimators in which one estimator is superior to all others and the efficiencies of all the estimators with respect to the estimator $\bar{y}$ are given in Table 4.

**Note.**

1. Bold figure in Table 3 indicates the minimum bias for the given population.
2. Bold figure in Table 4 indicates the maximum efficiency for the given population.
3. For Table 4, figure in bracket in each box indicates the preference region of the corresponding estimator for the given population.

From Table 4, we see that in all the populations the value of $K$ for the most efficient estimator really lies in the corresponding preference region of the same estimator. Hence it may be concluded that for the known value of $K$ in any population, we can choose the most efficient estimator among the estimators namely $\bar{y}$, $\bar{y}_R$, $\bar{y}_{RD}$, $t_{MS1}$, $t_{MS2}$, $\hat{y}_{D1}^*$ and $\hat{y}_{D2}^*$.

**Appendix A**

Up to the terms of order $n^{-1}$, we have

\begin{align*}
B(\bar{y}_R) &= \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y} \left(C_x^2 - \rho_{yx} C_y C_x\right) \\
B(t_{MS1}) &= \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y} \frac{1}{X_m + \bar{X}} \left(\frac{\bar{X}^2 C_x^2}{X_m + \bar{X}} - \rho_{yx} C_y x C_x\right) \\
B(t_{MS2}) &= \left(\frac{1}{n} - \frac{1}{N}\right) \bar{Y} \frac{1}{X_M + \bar{X}} \left(\frac{\bar{X}^2 C_x^2}{X_M + \bar{X}} - \rho_{yx} C_y x C_x\right). 
\end{align*}

**Acknowledgements**

Authors are very thankful to the referees for their valuable suggestions given for improving the original version of the manuscript.

**References**


Panse, V. G. and Sukhatme, P. V. (1967). *Statistical Methods for Agricultural Workers*, Indian Council of Agricultural Research, New Delhi, India.


