AN IMPROVED REGRESSION TYPE ESTIMATOR OF FINITE POPULATION MEAN USING AUXILIARY INFORMATION

Dr. Raghavendra Singh

Department of Statistics, Lucknow Christian Degree College, Lucknow-226018(India)

ABSTRACT

For estimating population mean, an estimator using prior auxiliary information is proposed, its bias and mean square error are found, and its comparative study with the usual ratio, product and linear regression estimators is made. An empirical illustration is also included to justify the practical utility of the proposed estimator.

Key words: *Auxiliary information, Coefficient of variation, Bias, Mean square error and efficiency.*

1. INTRODUCTION

Using auxiliary information in the form of known population mean $\overline{X}$ of auxiliary variable x,estimators like ratio, product and regression are available in the literature. In many practical situations, the coefficient of variation $C_x$of auxiliary variable$x$may be available, hence we can utilize this information along with $\overline{X}$ for the efficient estimation of population mean $\overline{Y}$ of study variable y.

Let the variable of interest y be related with an auxiliary variable$x$, information on which is available .Suppose n pairs $(x_i, y_i), i = 1, 2, 3..., n$of observations are taken on n units sampled under simple random sampling without replacement (SRSWOR) from a population of size N.

The ratio, product and linear regression estimators for $\overline{Y}$ are given by

$$\overline{y}_r = \overline{y}(\frac{\overline{X}}{\overline{x}}) \quad \textbf{(1.1)}$$

$$\overline{y}_p = \overline{y}(\frac{\overline{x}}{\overline{X}}) \quad \textbf{(1.2)}$$

And $\overline{y}_{lr} = \overline{y} + b(\overline{X} - \overline{x})$ **(1.3)**

Where $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$,$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2, s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})^2$$

$$s_{yx} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x}),$$

and $b = \frac{s_{yx}}{s_x^2}$.

Further, for $Y_1, Y_2, ..., Y_N$ being population values on study variable y, and $X_1, X_2, ..., X_N$ being the population values on the auxiliary variable x,let

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i,\ \overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i,$$

$$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \overline{Y})^2, \qquad S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})^2,$$

$$S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y}), \mu_{rs} = \frac{1}{N}\sum_{i=1}^{N}(X_i - \overline{X})^r (Y_i - \overline{Y})^s,$$

$$C_y = \frac{\sqrt{\mu_{02}}}{\overline{Y}},\ C_x = \frac{\sqrt{\mu_{20}}}{\overline{X}}, \beta = \frac{S_{yx}}{S_x^2} = \rho\frac{S_y}{S_x}, \rho = \frac{\mu_{11}}{\sqrt{\mu_{20}\mu_{02}}} = \frac{S_{yx}}{S_x S_y},$$

$$\gamma_1(x) = \frac{\mu_{30}}{\mu_{20}^{3/2}} = \frac{\mu_{30}}{\overline{X}^3 C_x^3} \text{and} \beta_2(x) = \frac{\mu_{40}}{\mu_{20}^2} = \frac{\mu_{40}}{\overline{X}^4 C_x^4}.$$

For the known population coefficient of variation C_x and population mean $\overline{X}$ of auxiliary variable x, the proposed estimator for population mean $\overline{Y}$ is

$$\overline{y}_k = \overline{y} + b(\overline{X} - \overline{x}) + k(\overline{x} - \frac{s_x}{C_x}) \quad \textbf{(1.4)}$$

where k is the characterizing scalar to be chosen suitably.

2. BIAS AND MEAN SQUARE ERROR OF $\overline{y}_k$

For simplicity, it assumed that population size N is large enough as compared to the sample size n so that finite population correction terms may be ignored. However, if N is not large as compared to n,the expressions concerning bias and mean square error will be multiplied by the finite population correction factor $\frac{N-n}{N}$ which tends to unity if N is large enough as compared to n.

Further, let

$$\overline{y} = \overline{Y} + e_0,\ \overline{x} = \overline{X} + e_1$$

$$s_{yx} = S_{yx} + e_2,\, s_x^2 = S_x^2 + e_3$$

so that

$$E(e_0) = E(e_1) = E(e_2) = E(e_3) = 0$$

and $E(e_0^2) = \frac{S_y^2}{n}$, $E(e_1^2) = \frac{S_x^2}{n}$, $E(e_3^2) = \{\frac{\beta_2(x)-1}{n}\}.S_x^4$,

$E(e_0e_1) = \frac{S_{yx}}{n} = \frac{\beta S_x^2}{n}$, $E(e_0e_3) = \frac{\mu_{21}}{n}$, $E(e_1e_2) = \frac{\mu_{21}}{n}$, $E(e_1e_3) = \frac{\mu_{30}}{n}$.

From **(1.4)**, we have

$$\overline{y}_k = (\overline{Y} + e_0) + \left(\frac{S_{yx}+e_2}{S_x^2+e_3}\right)(-e_1) + k\left[\overline{X} + e_1 - \overline{X}\frac{(S_x^2+e_3)^{1/2}}{S_x}\right]$$

$$\overline{y}_k - \overline{Y} = e_0 - \beta e_1\left(1 + \frac{e_2}{S_{yx}}\right)\left(1 + \frac{e_3}{S_x^2}\right)^{-1} + k\left[\overline{X} + e_1 - \overline{X}\left(1 + \frac{e_3}{S_x^2}\right)^{1/2}\right] \tag{2.1}$$

Expanding $\left(1 + \frac{e_3}{S_x^2}\right)^{-1}$ and $\left(1 + \frac{e_3}{S_x^2}\right)^{1/2}$, multiplying out and retaining the terms of $e_i^{'}s$ up to the second degree, we obtain

$$\overline{y}_k - \overline{Y} = (e_0 - \beta e_1) + k\overline{X}\left(\frac{e_1}{\overline{X}} - \frac{e_3}{2S_x^2}\right) + \beta\left(\frac{e_1e_3}{S_x^2} - \frac{e_1e_2}{S_{yx}}\right) + \frac{k\overline{X}e_3^2}{8S_x^4} + \dots \tag{2.2}$$

Taking expectation on both sides of **(2.2)**, we have the bias of $\overline{y}_k$ up to terms of order $O(\frac{1}{n})$ to be

$$Bias(\overline{y}_k) = \frac{\beta}{n}\left(\frac{\mu_{30}}{S_x^2} - \frac{\mu_{21}}{S_{yx}}\right) + \frac{k\overline{X}}{8n}\left(\beta_2(x) - 1\right)$$

$$= \frac{1}{n}\left[\frac{k\overline{X}}{8}\left(\beta_2(x) - 1\right) - \beta\left(\frac{\mu_{21}}{S_{yx}} - \frac{\mu_{30}}{S_x^2}\right)\right] \tag{2.3}$$

Squaring both sides of **(2.2)** and taking expectation, the mean square error of $\overline{y}_k$ given by $E(\overline{y}_k - \overline{Y})^2$ up to terms of order $O(\frac{1}{n})$ is

$$MSE(\overline{y}_k) = E(e_0 - \beta e_1)^2 + k^2\overline{X}^2\left[\frac{E(e_1^2)}{\overline{X}^2} + \frac{E(e_3^2)}{4S_x^4} - \frac{E(e_1e_3)}{\overline{X}S_x^2}\right]$$

$$+2k\overline{X}\left[\frac{E(e_0e_1)}{\overline{X}}-\frac{E(e_0e_3)}{2S_x^2}-\frac{\beta E(e_1^2)}{\overline{X}}+\frac{\beta E(e_1e_3)}{2S_x^2}\right]$$

$$=\frac{\overline{Y}^2C_y^2(1-\rho^2)}{n}+\frac{k^2\overline{X}^2}{4n}\left[4C_x^2+\beta_2(x)-1-4\gamma_1C_x\right]$$
$$+\frac{k\overline{X}}{n}\left[\frac{2S_{yx}}{\overline{X}}-\frac{\mu_{21}}{\overline{X}^2C_x^2}-\frac{2\beta S_x^2}{\overline{X}}+\beta\gamma_1(x)\overline{X}C_x\right]$$

$$=\frac{\overline{Y}^2C_y^2(1-\rho^2)}{n}+\frac{k^2\overline{X}^2[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]}{4n}+\frac{k[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]}{n\overline{X}C_x^2} \tag{2.4}$$

The optimum value of k minimizing the mean square error of $\overline{y}_k$in (**2.4)** is given by

$$k_{opt}=-\frac{2[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]}{\overline{X}^3C_x^2[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]}=k_0(say) \tag{2.5}$$

and the minimum mean square error of $\overline{y}_k$ is given by

$$MSE(\overline{y}_{k_0})=\frac{\overline{Y}^2C_y^2(1-\rho^2)}{n}-\frac{[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]^2}{n\overline{X}^4C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]}. \tag{2.6}$$

3. ESTIMATOR BASED ON ESTIMATED OPTIMUM $\hat{k}$

In practice, the guessed value of k may not be known, hence the alternative is to replace it by its estimate based on sample values. Replacing S_{yx},μ_{21},$\gamma_1(x)$, $\beta_2(x)$, μ_{40}, and μ_{30}involved in optimum k by their unbiased or consistent estimators given by

$$\hat{S}_{yx}=s_{yx}=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{y})(x_i-\overline{x}),\ \hat{\mu}_{21}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2(y_i-\overline{y}),$$

$$\hat{\gamma}_1(x)=\frac{\hat{\mu}_{30}}{\overline{X}^3C_x^3},\ \hat{\beta}_2(x)=\frac{\hat{\mu}_{40}}{\overline{X}^4C_x^4},\ \hat{\mu}_{40}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^4$$ and

$$\hat{\mu}_{30}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^3$$,we get the estimated optimum $\hat{k}$ as

$$\hat{k}=-\frac{2[\overline{X}C_x\hat{\gamma}_1(x)s_{yx}-\hat{\mu}_{21}]}{\overline{X}^3C_x^2[4C_x^2+\hat{\beta}_2(x)-1-4\hat{\gamma}_1(x)C_x]}$$

$$=-\frac{2\left[\frac{\overline{X}C_x\hat{\mu}_{30}s_{yx}}{\overline{X}^3C_x^3}-\hat{\mu}_{21}\right]}{\overline{X}^3C_x^2\left[4C_x^2+\frac{\hat{\mu}_{40}}{\overline{X}^4C_x^4}-1-\frac{4\hat{\mu}_{30}C_x}{\overline{X}^3C_x^3}\right]} \tag{3.1}$$

and hence the resulting estimator of population mean $\overline{Y}$ based on estimated optimum $\hat{k}$ is

$$\overline{y}_{es} = \overline{y} + b(\overline{X} - \overline{x}) + \hat{k}(\overline{x} - \frac{s_x}{C_x}) \quad \textbf{(3.2)}$$

Let

$\hat{\mu}_{30} = \mu_{30} + e_4, \hat{\mu}_{21} = \mu_{21} + e_5$ and $\hat{\mu}_{40} = \mu_{40} + e_6$

so that from **(3.1)**, we get

$$\hat{k} = -\frac{2\left[\frac{\overline{X}C_x(\mu_{30}+e_4)(S_{yx}+e_2)}{\overline{X}^3 C_x^3} - (\mu_{21}+e_5)\right]}{\overline{X}^3 C_x^2\left[4C_x^2 + \frac{(\mu_{40}+e_6)}{\overline{X}^4 C_x^4} - 1 - \frac{4(\mu_{30}+e_4)C_x}{\overline{X}^3 C_x^3}\right]}$$

$$= -\frac{2\left[\frac{\overline{X}C_x\mu_{30}S_{yx}}{\overline{X}^3 C_x^3} + \frac{\mu_{30}e_2 + S_{yx}e_4 + e_2e_4}{\overline{X}^2 C_x^2} - \mu_{21} - e_5\right]}{\overline{X}^3 C_x^2\left[4C_x^2 + \frac{\mu_{40}}{\overline{X}^4 C_x^4} + \frac{e_6}{\overline{X}^4 C_x^4} - 1 - \frac{4\mu_{30}C_x}{\overline{X}^3 C_x^3} - \frac{4e_4}{\overline{X}^3 C_x^2}\right]}$$

$$= -\frac{2\left[\overline{X}C_x\gamma_1(x)S_{yx} - \mu_{21} + \frac{\mu_{30}e_2 + S_{yx}e_4 + e_2e_4 - \overline{X}^2 C_x^2 e_5}{\overline{X}^2 C_x^2}\right]}{\overline{X}^3 C_x^2\left[4C_x^2 + \beta_2(x) - 1 - 4\gamma_1(x)C_x + \frac{e_6 - 4\overline{X}C_x^2 e_4}{\overline{X}^4 C_x^4}\right]}$$

$$= -\frac{2[\overline{X}C_x\gamma_1(x)S_{yx} - \mu_{21}]}{\overline{X}^3 C_x^2[4C_x^2 + \beta_2(x) - 1 - 4\gamma_1(x)C_x]}\left[1 + \frac{\mu_{30}e_2 + S_{yx}e_4 + e_2e_4 - \overline{X}^2 C_x^2 e_5}{\overline{X}^2 C_x^2[\overline{X}C_x\gamma_1(x)S_{yx} - \mu_{21}]}\right]\left[1 + \frac{e_6 - 4\overline{X}C_x^2 e_4}{\overline{X}^4 C_x^4[4C_x^2 + \beta_2(x) - 1 - 4\gamma_1(x)C_x]}\right]^{-1}$$

$$= -\frac{2[\overline{X}C_x\gamma_1(x)S_{yx} - \mu_{21}]}{\overline{X}^3 C_x^2[4C_x^2 + \beta_2(x) - 1 - 4\gamma_1(x)C_x]}\left[1 + \frac{\mu_{30}e_2 + S_{yx}e_4 + e_2e_4 - \overline{X}^2 C_x^2 e_5}{\overline{X}^2 C_x^2[\overline{X}C_x\gamma_1(x)S_{yx} - \mu_{21}]} - \frac{e_6 - 4\overline{X}C_x^2 e_4}{\overline{X}^4 C_x^4[4C_x^2 + \beta_2(x) - 1 - 4\gamma_1(x)C_x]} + \dots\right]^{-1} \quad \textbf{(3.3)}$$

From **(3.2)**, we have

$$\overline{y}_{es} = (\overline{Y} + e_0) + \left(\frac{S_{yx}+e_2}{S_x^2+e_3}\right)(-e_1) + \hat{k}\left[\overline{X} + e_1 - \overline{X}\frac{(S_x^2+e_3)^{1/2}}{S_x}\right]$$

$$\overline{y}_{es} - \overline{Y} = e_0 - \beta e_1\left(1 + \frac{e_2}{S_{yx}}\right)\left(1 + \frac{e_3}{S_x^2}\right)^{-1} + \hat{k}\left[\overline{X} + e_1 - \overline{X}\left(1 + \frac{e_3}{S_x^2}\right)^{1/2}\right] \quad \textbf{(3.4)}$$

Expanding $\left(1 + \frac{e_3}{S_x^2}\right)^{-1}$ and $\left(1 + \frac{e_3}{S_x^2}\right)^{1/2}$, multiplying out and retaining the terms of $e_i^{,}s$ up to the second degree, we obtain

$$\overline{y}_{es} - \overline{Y} = (e_0 - \beta e_1) + \hat{k}\overline{X}\left(\frac{e_1}{\overline{X}} - \frac{e_3}{2S_x^2}\right) + \beta\left(\frac{e_1e_3}{S_x^2} - \frac{e_1e_2}{S_{yx}}\right) + \frac{\hat{k}\overline{X}e_3^2}{8S_x^4} + \dots \quad \textbf{(3.5)}$$

Substituting $\hat{k}$ from **(3.3)** in **(3.5)**, squaring both sides and taking expectation, the mean square error of $\overline{y}_{es}$ given by $E(\overline{y}_{es}-\overline{Y})^2$ to the terms of order $O(\frac{1}{n})$ is

$$MSE(\overline{y}_{es}) = E\left[(e_0-\beta e_1) - \frac{2[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]}{\overline{X}^2C_x^2[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]}\left(\frac{e_1}{\overline{X}}-\frac{e_3}{2S_x^2}\right)\right]^2$$

$$= E(e_0^2)+\beta^2E(e_1^2)-2\beta E(e_0e_1)+\frac{4[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]^2}{\overline{X}^4C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]^2}\left[\frac{E(e_1^2)}{\overline{X}^2}+\frac{E(e_3^2)}{4S_x^4}-\frac{E(e_1e_3)}{\overline{X}S_x^2}\right]$$

$$-\frac{4[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]}{\overline{X}^2C_x^2[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]}\left[\frac{E(e_0e_1)}{\overline{X}}-\frac{E(e_0e_3)}{2S_x^2}-\frac{\beta E(e_1^2)}{\overline{X}}+\frac{\beta E(e_1e_3)}{2S_x^2}\right]$$

$$= \frac{S_y^2}{n}+\frac{\beta^2S_x^2}{n}-\frac{2\beta S_{xy}}{n}+\frac{4[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]^2}{n\overline{X}^4C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]^2}\left[\frac{S_x^2}{\overline{X}^2}+\frac{\{\beta_2(x)-1\}S_x^4}{4S_x^4}-\frac{\mu_{30}}{\overline{X}S_x^2}\right]$$

$$-\frac{4[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]}{n\overline{X}^2C_x^2[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]}\left[\frac{S_{xy}}{\overline{X}}-\frac{\mu_{21}}{2S_x^2}-\frac{S_{xy}S_x^2}{\overline{X}S_x^2}+\frac{S_{xy}\mu_{30}}{2S_x^4}\right]$$

$$= \frac{1}{n}\left[S_y^2+\frac{\rho^2S_y^2S_x^2}{S_x^2}-\frac{2\rho^2S_y^2S_x}{S_x}\right]+\frac{[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]^2}{n\overline{X}^4C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]^2}[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]$$

$$-\frac{2[\overline{X}C_x\gamma_1(x)S_{xy}-\mu_{21}]}{n\overline{X}^4C_x^4[4C_x^4+\beta_2(x)-1-4\gamma_1(x)C_x]}[\overline{X}C_x\gamma_1(x)S_{xy}-\mu_{21}]$$

$$= \frac{\overline{Y}^2C_y^2(1-\rho^2)}{n}+\frac{[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]^2}{n\overline{X}^4C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]}-\frac{2[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]^2}{n\overline{X}^4C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]}$$

$$= \frac{\overline{Y}^2C_y^2(1-\rho^2)}{n}-\frac{[\overline{X}C_x\gamma_1(x)S_{yx}-\mu_{21}]^2}{n\overline{X}^4C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]} \quad \textbf{(3.6)}$$

which shows that the mean square error of the estimator $\overline{y}_{es}$ in **(3.2)** based on the estimated optimum, to the terms of order $O(\frac{1}{n})$ is the same as that of $\overline{y}_{k_0}$ given in **(2.6).**

4. EFFICIENCY COMPARISON

The mean square errors (MSE) of $\overline{y}$, $\overline{y}_r$, $\overline{y}_p$, $\overline{y}_{lr}$ and $\overline{y}_{es}$ to the first degree of approximation are respectively given by

$$MSE(\overline{y}) = \frac{\overline{Y}^2C_y^2}{n} \quad \textbf{(4.1)}$$

$$MSE(\overline{y}_r) = \frac{\overline{Y}^2}{n}\left[C_y^2+C_x^2-2\rho C_xC_y\right] \quad \textbf{(4.2)}$$

$$MSE(\overline{y}_p) = \frac{\overline{Y}^2}{n}\left[C_y^2+C_x^2+2\rho C_xC_y\right] \quad \textbf{(4.3)}$$

$$MSE(\overline{y}_{lr}) = \frac{\overline{Y}^2 C_y^2}{n}(1-\rho^2) \tag{4.4}$$

$$MSE(\overline{y}_{es}) = \frac{\overline{Y}^2 C_y^2(1-\rho^2)}{n} - \frac{[\overline{X} C_x \gamma_1(x) S_{yx} - \mu_{21}]^2}{n\overline{X}^4 C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]} \tag{4.5}$$

From **(4.5)** and **(4.1),** we get

$$MSE(\overline{y}) - MSE(\overline{y}_{es}) = \frac{\rho^2\overline{Y}^2 C_y^2}{n} + \frac{[\overline{X} C_x \gamma_1(x) S_{yx} - \mu_{21}]^2}{n\overline{X}^4 C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]} > 0,$$

from **(4.5)** and **(4.2)**, we get

$$MSE(\overline{y}_r) - MSE(\overline{y}_{es}) = \frac{\overline{Y}^2[\rho C_y - C_x]^2}{n} + \frac{[\overline{X} C_x \gamma_1(x) S_{yx} - \mu_{21}]^2}{n\overline{X}^4 C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]} > 0,$$

from **(4.5)** and **(4.3)**, we get

$$MSE(\overline{y}_p) - MSE(\overline{y}_{es}) = \frac{\overline{Y}^2[\rho C_y + C_x]^2}{n} + \frac{[\overline{X} C_x \gamma_1(x) S_{yx} - \mu_{21}]^2}{n\overline{X}^4 C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]} > 0,$$

and from **(4.5)** and **(4.4)**, we get

$$MSE(\overline{y}_{lr}) - MSE(\overline{y}_{es}) = \frac{[\overline{X} C_x \gamma_1(x) S_{yx} - \mu_{21}]^2}{n\overline{X}^4 C_x^4[4C_x^2+\beta_2(x)-1-4\gamma_1(x)C_x]} > 0$$

showing that the estimator $\overline{y}_{es}$ is superior to all the estimators, that is, mean per unit, ratio, product and linear regression estimators in the sense of having lesser mean square error.

5. EMPRIRICAL STUDY

In order to justify the practical utility of the proposed estimator, we have considered two natural populations. The description of the populations is given below.

POPULATION-I: Singh and Chaudhary (1986, page no.176)

y : Total no. of guava trees

x: Area under guava orchard (in acres)

$N = 13, \overline{X} = 5.66153846, \overline{Y} = 746.9231, S_x^2 = 13.5422808, S_y^2 = 202717.6$

$C_x = 0.64999714, C_y = 0.602795, \rho = 0.900596, \gamma_1(x) = 0.744442366, \beta_2(x) = 2.94538212,$

$\mu_{21} = 4835.446$

POPULATION-II: Mukhopadhyay (2008, page no.104)

y : Quantity of raw material

x: Number of laborers

$N = 20, \overline{X} = 441.95, \overline{Y} = 41.5, S_x^2 = 6432.044, S_y^2 = 88.25104$

$C_x = 0.181469, C_y = 0.226366, \rho = 0.675049, \gamma_1(x) = 1.670798, \beta_2(x) = 5.962528,$

$\mu_{21} = 87548.29$

The mean square errors of the estimators $\overline{y}$, $\overline{y}_r$, $\overline{y}_p$, $\overline{y}_{lr}$ and $\overline{y}_{es}$ are given in table **(5.1)** and the percentage relative efficiencies (PRE) of the estimators $\overline{y}_r$, $\overline{y}_p$, $\overline{y}_{lr}$ and $\overline{y}_{es}$ over the mean per unit estimator $\overline{y}$ are given in tables **(5.2)** .

	Mean Square Errors of $\overline{y}$, $\overline{y}_r$, $\overline{y}_p$, $\overline{y}_{lr}$ and $\overline{y}_{es}$				
	$MSE(\overline{y})$	$MSE(\overline{y}_r)$	$MSE(\overline{y}_p)$	$MSE(\overline{y}_{lr})$	$MSE(\overline{y}_{es})$
Population I	$\frac{202717.6}{n}$	$\frac{44700.61}{n}$	$\frac{832150.9}{n}$	$\frac{38298.71}{n}$	$\frac{36505.49}{n}$
Population II	$\frac{88.25104}{n}$	$\frac{49.45049}{n}$	$\frac{240.4819}{n}$	$\frac{48.03582}{n}$	$\frac{45.00982}{n}$

Table (5.1)

Percentage Relative Efficiencies (PRE) of $\overline{y}_r$, $\overline{y}_p$, $\overline{y}_{lr}$ and $\overline{y}_{es}$ over mean per unit estimator $\overline{y}$				
	$\overline{y}_r$	$\overline{y}_p$	$\overline{y}_{lr}$	$\overline{y}_{es}$
Population I	453.5006989	24.3606753	529.3065907	555.3071345
Population II	178.4634288	36.6975886	183.7192229	196.0706345

Table (5.2)

From table **(5.2)** or **(5.1),** it is clear that $\overline{y}_{es}$ is the most efficient estimator.

REFERENCES

[1]. Cochran, W.G. (1977). Sampling Techniques, New York: John Wiley and Sons.

[2]. Singh, Daroga and Chaudhary F.S. (1986). Theory and Analysis of Sample Survey Designs, New Age International (p) Limited Publishers, New Delhi.

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

[4]. Mukhopadhyay, P. (2008). Theory and Methods of Survey Sampling, 2nd Ed., PHI Learning Private Limited, New Delhi.