Jour. Ind. Soc. Ag. Statistics
54(1), 2001 : 13–18

An Estimation of Population Mean in the Presence of Measurement Errors

Manisha and R. Karan Singh
Lucknow University, Lucknow, U.P.
(Received : February, 2000)

SUMMARY

The effect of measurement errors on a new estimator obtained as combination of ratio and mean per unit estimators, are examined. A comparative study is made among the proposed estimator, the ratio estimator and the mean per unit estimator in the presence of measurement errors.

Key words : Measurement errors, Ratio estimator, Mean square error, Efficiency.

1. Introduction

For a simple random sample of size n, let (x_i, y_i) be the pair of values instead of the true values (X_i, Y_i) on the two characteristics (X, Y) respectively for the i^{th} $(i = 1, 2 \ldots n)$ unit in the sample. Let the observational or measurement errors are

$$u_i = y_i - Y_i \tag{1.1}$$

$$v_i = x_i - X_i \tag{1.2}$$

which are stochastic in nature and are uncorrelated with mean zero and variances σ_u^2 and σ_v^2 respectively. Further, let the population means of (X, Y) be (μ_X, μ_Y), population variances of (X, Y) be σ_X^2, σ_Y^2 respectively and ρ be the population correlation coefficient between X and Y.

Assuming μ_X to be known for the estimation of population mean μ_Y, the ratio estimator (see Cochran, [1] and Sukhatme *et al.* [3]) is given by

$$t_R = \frac{\bar{y}}{\bar{x}} \cdot \mu_X \tag{1.3}$$

where $(\bar{y}, \bar{x})$ are the means of the sample observations on (Y, X) respectively.

Combining the ratio estimator t_R and the mean per unit estimator $\bar{y}$ for the estimation of the population mean μ_Y of Y, proposed estimator $\bar{y}_\theta$ is defined as

$$\bar{y}_\theta = \theta \, . \, t_R + (1-\theta)\,\bar{y} \tag{1.4}$$

where θ is the characterising scalar to be chosen suitably.

2. *Bias and Mean Square Error (MSE) of $\bar{y}_\theta$*

We first introduce the following notations

$$W_U = \frac{1}{n^{1/2}} \sum_i u_i, \; W_Y = \frac{1}{n^{1/2}} \sum_i (Y_i - \mu_Y)$$

$$C_Y = \frac{\sigma_Y}{\mu_Y}$$

Similarly, we can define W_V, W_X and C_X for X.

We have

$$\bar{y} - \mu_Y = \frac{1}{n} \sum_i \{(Y_i - \mu_Y) + u_i\}$$

$$= \frac{1}{n^{1/2}} (W_Y + W_U)$$

so that, variance of $\bar{y}$ is

$$V(\bar{y}) = \frac{\sigma_Y^2}{n}\left(1 + \frac{\sigma_U^2}{\sigma_Y^2}\right) \tag{2.1}$$

Further, we have

$$\bar{y}_\theta = \theta \, t_R + (1-\theta)\,\bar{y}$$

$$= \theta \, . \, \frac{\bar{y}}{\bar{x}} \mu_X + (1-\theta)\,\bar{y}$$

$$= \bar{y}\left\{1 + \theta\left(\frac{\mu_X}{\bar{x}} - 1\right)\right\}$$

$$= \left\{\mu_Y + \frac{1}{n^{1/2}}(W_Y + W_U)\right\}\left\{1 + \theta\left(\frac{\mu_X}{\mu_X + \frac{1}{n^{1/2}}(W_X + W_V)} - 1\right)\right\}$$

$$= \left\{\mu_Y + \frac{1}{n^{1/2}}(W_Y + W_U)\right\} \left\{1 + \theta\left[\left\{1 + \left(\frac{W_X + W_V}{n^{1/2}\,\mu_X}\right)\right\}^{-1} - 1\right]\right\}$$

$$= \left[\mu_Y + \frac{1}{n^{1/2}}(W_Y + W_U)\right] \left[1 + \theta\left\{-\frac{(W_X + W_V)}{n^{1/2}\,\mu_X} + \frac{W_X^2 + W_V^2 + 2W_V W_X}{n\,\mu_X^2} \cdots\right\}\right]$$

or

$$\bar{y}_\theta - \mu_Y = -\frac{\theta\mu_Y(W_X + W_V)}{n^{1/2}\,\mu_X} + \frac{\theta\,\mu_Y\,(W_X^2 + W_V^2 + 2W_X\,W_V)}{n\,\mu_X^2} + \frac{1}{n^{1/2}}(W_Y + W_U) - \frac{\theta}{n\,\mu_X}(W_Y + W_U)\,(W_X + W_V) + 0\left(\frac{1}{n^{3/2}}\right) \tag{2.2}$$

Taking expectation on both sides of (2.2), we have

$$E(\bar{y}_\theta - \mu_Y) = \frac{\theta\,.\,\mu_Y\,(\sigma_X^2 + \sigma_V^2)}{n\,\mu_X^2} - \frac{\theta\,\rho\,\sigma_X\,\sigma_Y}{n\,\mu_X} + 0\left(\frac{1}{n^{3/2}}\right)$$

so that the bias of $\bar{y}_\theta$ up to the terms of order $0\left(\frac{1}{n}\right)$ is

$$\text{Bias}\,(\bar{y}_\theta) = \theta\left\{\frac{\mu_Y}{n\,\mu_X^2}(\sigma_X^2 + \sigma_V^2) - \frac{1}{n\,\mu_X}\rho\,\sigma_X\,\sigma_Y\right\}$$

$$= \theta.\ \text{Bias}\,(t_R) \qquad \text{(see Shalabh, [2])} \tag{2.3}$$

Further, from (2.2) we have

$$(\bar{y}_\theta - \mu_Y)^2$$

$$= \theta^2\,\mu_Y^2\left[-\left(\frac{W_X + W_V}{n^{1/2}\,\mu_X}\right)\right]^2 + \frac{1}{n}(W_Y + W_U)^2 + \frac{2\theta}{n^{1/2}}\mu_Y\left[-\left(\frac{W_X + W_V}{n^{1/2}\,\mu_X}\right)\right](W_Y + W_U) - \frac{(W_Y + W_U)\,(W_X + W_V)}{n^{1/2}\,\mu_X} + 0\left(\frac{1}{n^{3/2}}\right).$$

Taking expectation on both sides, we obtain mean squared error of $\bar{y}_\theta$ up to terms of order $0\left(\frac{1}{n}\right)$ to be

$$\text{MSE}(\bar{y}_\theta) = E\left[\frac{\theta^2 \mu_Y^2}{n\,\mu_X^2}(W_X^2 + W_V^2 + 2W_X W_V) + \frac{1}{n}(W_Y^2 + W_U^2 + 2W_Y W_U) - \frac{2\theta\,\mu_Y}{n\,\mu_X}(W_X W_Y + W_X W_U + W_V W_Y + W_V W_U)\right]$$

$$= \frac{\theta^2 \mu_Y^2}{n\,\mu_X^2}(\sigma_X^2 + \sigma_V^2) + \frac{1}{n}(\sigma_Y^2 + \sigma_U^2) - \frac{2\,\theta\,\mu_Y}{n\,\mu_X}\rho\,\sigma_X\,\sigma_Y$$

$$= \frac{\sigma_Y^2}{n}\left[1 + \theta^2 \frac{\mu_Y^2}{\mu_X^2}\frac{\sigma_X^2}{\sigma_Y^2} - 2\theta\,\rho\,\frac{\mu_Y}{\mu_X}\frac{\sigma_X}{\sigma_Y}\right] + \frac{1}{n}\left[\frac{\theta^2 \mu_Y^2}{\mu_X^2}\sigma_V^2 + \sigma_U^2\right]$$

$$= \frac{\sigma_Y^2}{n}\left[1 - \theta\frac{C_X}{C_Y}\left(2\rho - \theta\frac{C_X}{C_Y}\right)\right] + \frac{1}{n}\left[\theta^2 \frac{\mu_Y^2}{\mu_X^2}\sigma_V^2 + \sigma_U^2\right] \tag{2.4}$$

3. *Concluding Remarks*

(a) All the results obtained by Shalabh [2] are the special cases of this study for $\theta = 0$ and $\theta = 1$ when the observations are subject to measurement errors.

(b) From Shalabh [2], the bias of t_R is

$$\text{Bias}\,(t_R) = \frac{\mu_Y}{n\,\mu_X^2}(\sigma_X^2 + \sigma_V^2) - \frac{1}{n\,\mu_X}\rho\,\sigma_X\,\sigma_Y$$

and from (2.3), the bias of $\bar{y}_\theta$ is

$$\text{Bias}\,(\bar{y}_\theta) = \theta.\ \text{Bias}\,(t_R)$$

so that the absolute bias of $\bar{y}_\theta$ is less than that of the ratio estimator t_R as long as the value of $|\theta| < 1$.

(c) From Shalabh [2], mean squared error of the ratio estimator t_R under measurement errors to the terms of order $0\left(\frac{1}{n}\right)$ is

$$\text{MSE}(t_R) = \frac{\sigma_Y^2}{n}\left[1 - \frac{C_X}{C_Y}\left(2\rho - \frac{C_X}{C_Y}\right)\right] + \frac{1}{n}\left[\frac{\mu_Y^2}{\mu_X^2}\sigma_V^2 + \sigma_U^2\right] \tag{3.1}$$

The mean squared error of $\bar{y}_\theta$ to the terms of order $0\left(\frac{1}{n}\right)$ is given by (2.4). Further, the mean squared error of $\bar{y}$ is same as $V(\bar{y})$ given by (2.1). Hence, from (2.4) and (2.1), we see that

MSE $(\bar{y}_\theta)$ < MSE $(\bar{y})$, that is, (2.4) < (2.1)

if $$-2\rho\left(\frac{\sigma_Y}{\sigma_X}\right)\cdot\left(\frac{\mu_Y}{\mu_X}\right)\theta+\theta^2\left(1+\frac{\sigma_V^2}{\sigma_X^2}\right)\left(\frac{\mu_Y}{\mu_X}\right)^2 < 0$$

or if

$$\rho > \frac{\theta}{2}\left(1+\frac{\sigma_V^2}{\sigma_X^2}\right)\cdot\frac{C_X}{C_Y}$$

when, μ_X and μ_Y have the same signs (noting that $\theta > 0$) (3.2)

$$\rho < -\frac{\theta}{2}\left(1+\frac{\sigma_V^2}{\sigma_X^2}\right)\cdot\frac{C_X}{C_Y}$$

when, μ_X and μ_Y have opposite signs (3.3)

In particular when C_X and C_Y are identical in magnitudes, the efficiency conditions (3.2) and (3.3) respectively reduce to

$$\rho > \frac{\theta}{2}\left(1+\frac{\sigma_V^2}{\sigma_X^2}\right)$$ when, μ_X and μ_Y have the same signs (3.4)

$$\rho < -\frac{\theta}{2}\left(1+\frac{\sigma_V^2}{\sigma_X^2}\right)$$ when, μ_X and μ_Y have the opposite signs (3.5)

It may be mentioned here that, for $\theta = 1$ the efficiency conditions (3.2) and (3.3) respectively reduce to

$$\rho > \frac{1}{2}\left(1+\frac{\sigma_V^2}{\sigma_X^2}\right)\cdot\frac{C_X}{C_Y}$$ for μ_X and μ_Y having same signs (3.6)

$$\rho < -\frac{1}{2}\left(1+\frac{\sigma_V^2}{\sigma_X^2}\right)\cdot\frac{C_X}{C_Y}$$ for μ_X and μ_Y having opposite signs (3.7)

which are the same efficiency conditions as obtained by Shalabh [2] for the superiority of the ratio estimator t_R over the mean per unit estimator $\bar{y}$ under measurement errors.

For $0 < \theta < 1$, comparing (3.2) with (3.6), it is clear that the efficiency condition (3.2) for the superiority of the estimator $\bar{y}_\theta$ over $\bar{y}$ is wider than the efficiency conditions (3.6) for the superiority of the ratio estimator t_R over $\bar{y}$, hence in the extended range of the efficiency condition (3.2) over that of the efficiency condition (3.6), the estimator $\bar{y}_\theta$ is superior to both the estimators t_R and $\bar{y}$. Similar remarks hold when for $0 < \theta < 1$, we compare the superiority condition (3.3) of $\bar{y}_\theta$ over $\bar{y}$ with the superiority condition (3.7) of t_R over $\bar{y}$.

ACKNOWLEDGEMENT

The authors wish to thank the referee for the valuable suggestions.

REFERENCES

[1] Cochran, W.G., (1963). *Sampling Techniques.* Second Edition, Wiley Eastern Private Limited, New Delhi.

[2] Shalabh, (1997). Ratio method of estimation in the pressence of measurement errors. *J. Ind. Soc. Agril. Stat.*, **52,** 150-155.

[3] Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C., (1984). *Sampling Theory of Surveys with Applications.* Iowa State University Press, Ames.