A MODIFIED RATIO ESTIMATOR USING COEFFICIENT OF VARIATION OF AUXILIARY VARIABLE

By

B.V.S. Sisodia and V.K. Dwivedi

Indian Veterinary Research Institute Izatnagar-243 122.

(Received : October, 1979)

Summary

A modified ratio estimator using coefficient of variation of auxiliary variable X is proposed in this paper. It is found that the absolute bias of the modified ratio estimator is always less than that of the ratio estimator when α is positive and greater than βc_x. The proposed estimator is found to be more efficient than both ratio and simple mean estimator when ρ lies between a certain range.

1. Introduction

The utilization of a prior value of coefficient of variation (c.v.) in the estimation stage appears to have started sprouting up with the work of Searles (1964). He developed an estimate which involved a prior value of c.v. of the character under study and the estimator so obtained, though biased was found to be more precise than simple mean estimator. Following Searles, some others have also used the c.v. of the character under study to increase the efficiency of estimators. The knowledge of the c.v. of the character under study is seldom available. However, the c.v. of an auxiliary variable can easily be obtained and it can be used to increase the efficiency of estimators. To estimate the population mean $\bar{Y}_N=\frac{I}{N}\sum_{i=1}^{N} yi$ of a variable Y of a finite population of size N, the ratio estimator based on a simple random sample of size n, is defined by

$$\hat{\bar{Y}}_R=\frac{\bar{Y}_n}{\bar{x}_n}\bar{X}_N \qquad \ldots(1.1)$$

where $\bar{Y}_n$ is the sample mean of the variable y and $\bar{x}_n$ and $\bar{X}_N$ are respectively the sample mean and the population mean of the auxiliary variable X. The bias (B) and mean-square error (MSE, denoted by M) of ratio estimator up to first order of approximation (Sukhatme and Sukhatme, 1970) are given by

$$B=\bar{Y}_N\theta \;(c_x^2 - \rho c_x c_y)=\theta\, \alpha c_x^2 \qquad ...(1.2)$$

and

$$M=\bar{Y}_N^2\; \theta \;(c_y^2 + c_x^2 - 2\rho\, c_x\, c_y) \qquad ...(1.3)$$

where $\theta=\left(\frac{1}{n}-\frac{1}{N}\right)$, c_x and c_y are c.v. of X and Y respectively, and ρ is correlation coefficient between X and Y. Here, α is intercept of the regression line of Y on X in the population.

Utilising the value of c_x, a modified ratio estimator is proposed and its properties are discussed in the following section. It is compared with ratio estimator and simple mean estimator in section 3. The paper is concluded with an empirical comparison on the basis of some real data.

2. Modified Ratio Estimator

We transform the auxiliary variable X_i and write $X_i'=X_i+c_x$. Then, the population mean of the transformed variable would be $\bar{X}_N'=\bar{X}_N+c_x$ and the sample mean $\bar{x}_n'=\bar{x}_n+c_x$. We propose a modified ratio estimator of $\bar{Y}_N$,

$$\hat{\bar{Y}}_{MR}=\bar{y}_n\left(\frac{\bar{X}_N+c_x}{\bar{x}_n+c_x}\right) \qquad .. (2.1)$$

The bias and MSE of the proposed estimator are derived up to first order of approximation on the same lines as in case of usual ratio estimator (Sukhatme and Sukhatme, 1970). The bias (B^*) of $\hat{\bar{Y}}_{MR}$ is

$$B^*=\bar{Y}_N\; \theta \left(c_x'^2 - \rho c_x'\, c_y \right) \qquad .. (2.2)$$

where $c_x'=S_x/\bar{X}_N'$, and S_x^2 is the population variance of the auxiliary variable. Alternatively, B^* can also be written as under

$$B^*=\theta\alpha'\, c_x'^2 \qquad ...(2.3)$$

Where $\alpha'=\alpha-\beta c_x$ is the intercept of the new regression line of Y on transformed variable X'. B^* vanishes when $\alpha'=0$. i.e., $\alpha/\beta=c_x$ which also implies that

$$\frac{R-\beta}{\beta}=\frac{c_x}{\bar{X}_N}$$

We shall now compare the absolute bias of ratio estimator and modified ratio estimator. The absolute bias of $\hat{\bar{Y}}_{MR}$ will be less than the absolute bias of $\hat{\bar{Y}}_R$ if

$$|\,B\,|-|\,B^*\,|>0$$

or

$$|\,\alpha c_x^2\,|-|\,(\alpha-\beta\, c_x)\, c_x'^2\,|>0$$

or

$$c_x^2\,/\,c_x'^2\,|\,\alpha\,|>|\,\alpha-\beta\, c_x\,|\,. \quad \text{...(2.4)}$$

When α is positive and greater than βc_x, the inequality (2.4) always holds. In other situations *viz.*, when (*i*) α is positive and less than βc_x and, (*ii*) α is negative, it is not possible to make a similar comparison. Some empirical comparisons are made in section 4. The MSE of $\hat{\bar{Y}}_{MR}$ denoted by M^*, is

$$M^*=\bar{Y}_N^2\,\theta\left(c_y^2+c_x'^2-2\,\rho c_x'\, c_y\right) \quad \text{...(2.5)}$$

It is easily verified that for $c_x=\alpha/\beta$, the MSE of $\bar{Y}_{MR}$ reduces to the MSE of usual linear regression estimator.

3. COMPARISON OF $\hat{\bar{Y}}_{MR}$ WITH $\hat{\bar{Y}}_R$ AND SIMPLE MEAN ESTIMATOR $\bar{y}_n$.

To examine the efficiency of $\hat{\bar{Y}}_{MR}$, we compare it with ratio estimator and simple mean estimator. The variance of simple mean estimator, $\bar{y}_n$, is given by

$$v\,(\bar{y}_n)=\bar{Y}_N^2\,\theta\, c_y^2 \quad \text{...(3.1)}$$

The estimator $\hat{\bar{Y}}_{MR}$ will be more efficient than $\bar{Y}_R$ if,

$$M>M^*$$

i.e.,

$$S_x\left[\frac{1}{\bar{X}_N^2}-\frac{1}{\bar{X}_N'^2}\right]>2\rho\, c_y\left[\frac{1}{\bar{X}_N}-\frac{1}{\bar{X}_N'}\right]$$

i.e.,

$$\rho<\tfrac{1}{2}\,\frac{c_x}{c_y}\left(\frac{2\bar{X}_N+c_x}{\bar{X}_N+c_x}\right) \quad \text{...(3.2)}$$

The estimator $\hat{\bar{Y}}_{MK}$ will be more efficient than simple mean estimator $\bar{y}_n$, if

$$v\,(\bar{y}_n) > M^*$$

i.e., $\rho > \frac{1}{2}\frac{c_x'}{c_y}$

i.e , $$\rho > \frac{1}{2}\,\frac{c_x}{c_y\left(1+\dfrac{c_x}{\bar{X}_N}\right)} \qquad \ldots(3.3)$$

Combining (3.2) and (3.3), we find that the modified ratio estimator is more efficient than the ratio as well as the simple mean estimator if,

$$\frac{1}{2}\;\frac{c_x}{c_y\,(1+c_x/\bar{X}_N)} < \rho < \frac{1}{2}\,\frac{c_x}{c_y}\left(\frac{2\bar{X}_N+c_x}{\bar{X}_N+c_x}\right) \qquad \ldots(3.4)$$

It is evident that lower limit of this range is slightly smaller than $\frac{1}{2}$ c_x/c_y.

Thus, in the range

$$\left\{\frac{1}{2}\;\frac{c_x}{c_y\,(1+c_x/\bar{X}_N)},\ \frac{1}{2}\,\frac{c_x}{c_y}\right\}$$

of ρ, the proposed estimator is better than simple mean estimator while the usual ratio estimator is inferior to it.However, in the range

$$\left\{\frac{1}{2}\,\frac{c_x}{c_y},\ \frac{1}{2}\,\frac{c_x}{c_y}\left(\frac{2\bar{X}_N+c_x}{\bar{X}_N+c_x}\right)\right\}$$

of ρ, the proposed estimator is superior to the usual ratio estimator. Since c_x is likely to be very small, it is expected that

$$\left(\frac{2\bar{X}_N+c_x}{\bar{X}_N+c_x}\right)$$

is quite close to 2. This indicates superiority of the proposed estimator over the usual ratio estimator for sufficiently wide range of values of ρ.

4. Empirical Comparison

For empirical comparison, we have considered six populations which are described in Table-1. The various parameters viz., $\bar{X}_N$, c_x, α, β, and ρ of these populations are computed and they are depicted in Table-1. Lower and upper bound of the inequality (3.4) are also shown in this table. Table-1 also contains the bias and MSE of different estimators to have an idea of reduction in the bias and MSE.

TABLE

Description of the populations, their parameters, bias and MSE of different estimate.

S. No.	*Source*	*X*	*Y*	*N*	X_N	c_x	α	β	ρ	*Lower bond of (3.4)*	*Upper bond of (3.4)*	$\frac{\lvert B \rvert}{\theta}$	$\frac{\lvert B^* \rvert}{\theta}$	$\frac{M}{\theta}$	$\frac{M^*}{\theta}$	$\frac{v(\bar{y}_n)}{\theta}$
1.	Sukhatme and Sukhatme (1970, p. 185)	Area under Wheat 1936	Area under Wheat 1937	34	218.41	0.77	17.95	0.84	0.93	0.51	0.99	9.79	6.53	3326.43	3241.23	2343.09×10
2.	Sukhatme and Sukhatme (1970, p. 256)	No. of villages	Area under Wheat	89	3.36	0.59	320.60	232.56	0.64	0.39	0.84	113.07	46.28	3396.69×10^2	3114.94×10^2	5130.85×10^2
3.	Cochran (1963, p. 32)	Family size	Weekly expenditure on food	33	7.73	0.41	16.79	2.87	0.43	0.50	0.60	2.82	1.99	131.87	114.9	103.46
4.	Cochran (1963, p. 32)	Weekly family income	-do-	33	72.54	0.15	157.48	2.55	0.25	0.20	0.40	0.22	0.18	99.45	98.69	103.46
5.	Cochran (1963, p. 156)	Number of inhabitant 1920	Number of in habitant, 1930	49	103.14	1.02	15.38	1.09	0.98	0.53	1.08	9.20	7.06	692.52	600.05	1505.23×10
6.	Cochran (1963, p. 175)	Number of peach trees in an orchard	Estimated production in bushels of peaches	256	44.45	1.40	6.25	1.13	0.89	0.48	0.98	10.76	7.39	1395.77	1306.11	6430.02

It is clear from Table-1 that the inequality (3.4) is satisfied for all the populations except the third. Thus, ratio and proposed estimato. are not applicable for third population while the proposed estimator is clearly more efficient than the ratio and the simple mean estimator for rest of the populations. It can also be seen from Table-1 that the proposed estimator is less biased than ratio estimator for all the populations.

ACKNOWLEDGEMENT

Authors are very grateful for referees comments which improved the earlier version of the paper.

REFERENCES

Cochran, W.G. (1963) : *Sampling Technique.* John Wiley & Sons Inc. London.

Mohanty, S, and Das, M.N. (1971) : Use of transformation in Sampling. *Jour. Ind. Soc. Agri. Stat.* Vol. 23, No. 2.

Searles; D.T. (1964) : Utilisation of known coefficient of variation in the estimation procedure. *Jour. Amer. Stat. Assoc.* Vol. 59, pp. 1225.

Sisodia B.V.S. and Singh. R. (1978) : A Note on use of coefficient of variation in successive sampling (Abstract). *Jour Indian Soc. Agri. Stat.* Vol. 30, No. 2.

Sukhatme, P.V. and Sukhatme, B.V. (1970) : Sampling theory of surveys with applications. Indian Society of Agricultural Statistics, New Delhi.