Jour. Ind. Soc. Ag. Statistics
Vol. XXXVII, No. 2 (1985), pp. 154-157

IMPROVING THE ESTIMATION OF POPULATION MEAN: SOME REMARKS

V. K. SRIVASTAVA and T. D. DWIVEDI
Department of Statistics, Lucknow University, Lucknow-226007

(Received: January, 1982)

SUMMARY

The paper considers the problem of improving the conventional estimator of population mean and examines the relevance of large-sample approximations for analysing the properties of estimators.

Keywords: Improved estimator than sample mean; Large-sample approximations.

1. Introduction

Sample mean is known to be the minimum variance unbiased estimator of population mean. Improvements in the estimator can be made if we are prepared to sacrifice unbiasedness. One such estimator was proposed by Searles [1] assuming the coefficient of variation to be known. His estimator was essentially a minimum mean squared error estimator. When coefficient of variation is not known, Srivastava [2] considered two approximate minimum mean squared error estimators and analyzed their large-sample properties. In the context of normal population, one estimator did not possess any finite moments while the other estimator had a larger mean squared error, according to large-sample approximation, than the variance of sample mean. Later, Srivastava [3] worked out the exact expressions and observed that the estimator having finite moments had a smaller mean squared error than the variance of sample mean so long as the relative variance of sample mean was more than $\frac{1}{2}$. This poses a question. Can an improved estimator be suggested which will perform better than sample mean when the relative variance is less than $\frac{1}{2}$? This paper attempts to answer it. A simple estimator which merely calls for a change in sign is proposed and its properties are investigated. Finally, some remarks are placed indicating that large-sample approximations are not always a reliable guide. This point is, however, well recognised but our investigation furnishes a simple example to demonstrate it.

2. Estimators and their Properties

Suppose $\bar{y}$ is the sample mean based on a simple random sample of size n drawn from a population with mean $\bar{Y}$ and variance σ^2.

For estimating $\bar{Y}$, Srivastava [2] considered the following estimator as an approximation of the minimum mean squared error estimator proposed by Searles [1]:

$$t = \frac{\bar{y}^3}{\bar{y}^2 + s^2/n} \tag{2.1}$$

where s^2 is an unbiased estimator of σ^2.

The large-sample approximations, to order $O(n^{-2})$, of the relative bias and relative mean squared error are given by

$$RB(t) = E\left(\frac{t - \bar{Y}}{\bar{Y}}\right)$$

$$= \frac{c}{n}\left(1 - \frac{\sqrt{\beta_1 c}}{n}\right) \tag{2.2}$$

$$RMSE(t) = E\left(\frac{t - \bar{Y}}{\bar{Y}}\right)^2 \tag{2.3}$$

$$= \frac{c}{n}\left[1 + \frac{1}{n}(3c - 2\sqrt{\beta_1 c})\right]$$

where

$$c = \frac{\sigma^2}{\bar{Y}^2}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\sigma^3} \tag{2.4}$$

μ_3 being the third central moment of population. It is obvious that when

$$\sqrt{\frac{\beta_1}{c}} > \frac{3}{2} \tag{2.5}$$

the relative mean squared error of t is smaller than the relative variance (c/n) of $\bar{y}$. Thus the use of t may be recommended only for those populations which satisfy the inequality (2.5).

The above finding poses a natural question whether an improved estimator can be developed for those populations which do not satisfy (2.5). A simple estimator which is obtained merely through a change of plus sign in the denominator of t is

$$t^* = \frac{\bar{y}^3}{\bar{y}^2 - s^2/n} \tag{2.6}$$

The relative bias and relative mean squared error, both to order $O(n^{-2})$, can be obtained in the same way as indicated for t in Srivastava [2]:

$$RB(t^*) = E\left(\frac{t^* - \bar{Y}}{\bar{Y}}\right) \tag{2.7}$$

$$= \frac{c}{n}\left[1 + \frac{1}{n}\,(2c - \sqrt{\beta_1 c})\right]$$

$$RMSE(t^*) = E\left(\frac{t^* - \bar{Y}}{\bar{Y}}\right)^2 \tag{2.8}$$

$$= \frac{c}{n}\left[1 - \frac{1}{n}\,(c - 2\sqrt{\beta_1 c})\right]$$

Obviously, the relative mean squared error of t^* is smaller than the relative variance of $\bar{y}$ when

$$\sqrt{\frac{\beta_1}{c}} < \frac{1}{2} \tag{2.9}$$

It may be mentioned that an improved estimator over the range $(\frac{1}{2}, \frac{3}{2})$ of $\sqrt{\beta_1/c}$ can be designed but its intricate form renders it of little practical utility.

3. The Case of Normal Population

When the population is normal, we know that β_1 is O so that t^* appears to be better than $\bar{y}$ while t turns out to be worse according to our large-sample approximations.

Srivastava [3] worked out the exact expressions for the first two moments of t and therefrom compared the relative mean squared error of t and relative variance of $\bar{y}$ for few selected values of $(n - 1)$ and c/n. It was observed that t has smaller mean squared error than the relative variance of $\bar{y}$ so long as c/n exceeds $\frac{1}{2}$. Large-sample approximation for relative mean squared error fails to reflect this point simply because large n makes c/n small (given c) and the large-sample approximation is no more valid. Similarly, if we investigate the exact mean squared error of t^* (an expression for which can be derived exactly in the same manner as done by Srivastava [3] in case of t) and variance of $\bar{y}$, it is observed that the performance of t^* is poor in comparison to that of $\bar{y}$, according to mean squared error criterion, when c/n is large. The situation worsens to such an extent that large-sample approximation (2.8) in case of normal population provides a negative value of relative mean squared error for c/n greater than 1! This is, however, not true when exact mean squared error is considered.

ACKNOWLEDGEMENT

Partial financial support from NSERC Grant No. A 3989 is gratefully acknowledged.

REFERENCES

[1] Searles, D. T. (1964): The utilization of a known coefficient of variation in estimation procedure, *J. Amer. Statist. Ass.* **59**: 1225-26.

[2] Srivastava, V. K. (1974): On the use of coefficient of variation in estimating mean, *Jour. of Ind. Soc. of Agri. Statistics* **26**: 33-36.

[3] Srivastava, V. K. (1980): A note on the estimation of mean in normal population, *Metrika* **27**: 99-102.