ON ESTIMATING THE MEAN OF SYMMETRICAL POPULATIONS

BY

ASHOK SAHAI AND S. K. RAY
Lucknow University, Lucknow.

(Received : April, 1978)

1. INTRODUCTION

For estimating the mean $\bar{Y}$ of a population having the variance σ^2, Searles [1] proposed the estimator

$$t=\bar{y}\ (1+c/n)^{-1}$$

where $\bar{y}$ is the mean of o random sample of size m and c $(=\sigma^2/\bar{Y}^2)$, the square of the coefficient of variation, is considered to be known.

In practical situations, however, c is very rarely known. If c is estimated by its sample counterpart $c_e=s^2/\bar{y}^2$ where s^2 is the sample variance, the above estimator becomes

$$t_e=\bar{y}^3/(\bar{y}^2+s^2/n)$$

Srivastava (1974) suggested another estimator

$$t'_e=\bar{y}-(s^2/n\bar{y})$$

The bias and mean square error of t_e and t'_e to the order $O(n^{-2})$, are given by

$$B\ (t_e)=-(c/n)\ (1-\sqrt{\beta_1\ c/n})\ \bar{Y}$$

$$B\left(t'_e\right)=-(c/n)\ \ (1+(c-\sqrt{\beta_1 c})/n)\ \ \bar{Y}$$

and

$$M\ (t_e)=M\left(t'_e\right)=(\sigma^2/n)\ \ (1+(3c-2\sqrt{\beta_1 c})/n)$$

where β_1 $\left(=\mu_3^2/\sigma^6\right)$ is the coefficient of skewness of the population To simplify expressions of the bias and mean square error of these and other such estimators, the population is assumed to be symmetrical (*e.g.* Normal) for which $\beta_1=0$. For such estimators concerning symmetrical populations, to the order $O(n^{-2})$,

$$M(t_e)=M\left(t_e''\right)=(\sigma^2/n)(1+3c/n)>(\sigma^2/n)=\text{Var. }(\bar{y})$$

For symmetrical populatians, Upadhyaya and Srivastva [3] proposed the estimator

$$t_e''=\bar{y}+s^2/n\bar{y} \qquad \cdots(1)$$

for which the bias and mean square error, to the order $O(n^{-2})$ are given by

$$B\left(t_e''\right)=(c\bar{Y}/n)(1+c/n) \qquad \cdots(2)$$

$$M\left(t_e''\right)=(\sigma^2/n)(1-c/n) \qquad \cdots(3)$$

It is seen that

$$B\left(t_e''\right)=B\left(t_e'\right)$$

$$M\left(t_e''\right)<M\left(t_e'\right)$$

so that t_e'' is a better estimator than t_e'.

2. Alternative Estimators

We notice that the estimator t_e can be writen as
$t_e=\bar{y}-(s^2\bar{y})/(n\bar{y}^2+s^2)$

The relationship between t_e' and t_e'' leads us to propose an alternative estimator

$$t_e^*=\bar{y}+(s^2\bar{y})/(n\bar{y}^2+s^2) \qquad \cdots(4)$$

which is shown to be better than t_e''. As in t_e'', we consider the population to be symmetrical. To evaluate the bias and the mean square error of the estimator to order $O(n^{-2})$, let

$\bar{y}=\bar{Y}+u$, $s^2=\sigma^2+v$, where u and v are of order $O(n^{-\frac{1}{2}})$ with $E(u)=E(v)=O$. Thus, upto the order $O(n^{-2})$ we have

$$t_e^*=((\bar{Y}+u)+(s^2/n\bar{y})(1-s^2/n\bar{y}^2))$$

$$=\bar{Y}((1+u/\bar{Y})+(c+v/\bar{Y}^2)(1-u/\bar{Y}+u^2/\bar{Y}^2)/n-c^2/n^2)$$

so that,

$$\left(t_e^* - \bar{Y} \right) = \bar{Y} \left(u/\bar{Y} + c/n + (v/\bar{Y}^2 - cu/\bar{Y})/n + (cu^2/\bar{Y}^2 - c^2/n - uv/\bar{Y}^3)/n \right)$$

Hence,

$$B\left(t_e^* \right) = c\bar{Y}/n \qquad \cdots(5)$$

and

$$M\left(t_e^* \right) = (\sigma^2/n)\,(1-c/n) \qquad \cdots(6)$$

From (2), (3), (5) and (6), we have

$$B\left(t_e^* \right) < B\left(t_e'' \right) \text{ and } M\left(t_e^* \right) = M\left(t_e'' \right).$$

Therefore, our estimator (4) is superior to (1).

Further, as noted by Srivastava [2], even though c may not be exactly known, we may have figure fairly close to the true value of the coefficient of variation from our long association with experimentat material or from other empirical investigations or from some extraneous source. The coefficient of variation may, often, exhibit a stability in repeated experiments and a reasonably close guess may, thus, be available. Therefore, we may know or guess that c has an upper bound, say,

$$c \leqslant c_0$$

In that case we propose the estimator

$$t_e^{(k+1)} = \bar{y} + (s^2\bar{y})/(n\bar{y}^2 + (k+1)\, s^2) \qquad \cdots(7)$$

where $k = n/c_0$. We find the bias and the mean square error of this estimator, to the order $O\,(n^{-2})$ are as below.

$$B\left(t_e^{(k+1)} \right) = (c\bar{Y}/n)\,(1 - c/c_0) \qquad \cdots(8)$$

$$M\left(t_e^{(k+1)} \right) = (\sigma^2/n)\,(1 - c/n) \qquad \cdots(9)$$

Thus, the estimator (7) has the same mean square error as the estimators (1) and (4) but is much less biased than the latter ones. In fact, this estimator is almost unbiased to the extent of the closeness of the known or guessed upper bound to the true value of c and is unbiased if $c = c_0$.

REFERENCES

[1] Searles, D.T. (1964) : 'The utilization of a known coefficient of variation in estimation procedures' Jour. Amer. Stat. Assn., 59, 1225-6.

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

[3] Upadhyaya, L.N. & Srivastava, S.R. (1976) : 'A note on the use of coefficient of variation in estimating mean,' Jour. Ind. Soc. Agri. Stat., 28 (2), 97-99.