*Sankhyā : The Indian Journal of Statistics*
1983, Volume 45, Series A, Pt. 3 pp. 372–376.

# USE OF PRIOR INFORMATION ON SOME PARAMETERS IN ESTIMATING POPULATION MEAN

*By* T. P. TRIPATHI, and PULAKESH MAITI
*Indian Statistical Institute*
and
S. D. SHARMA
*Rajendra Agricultural University, Bihar, India*

*SUMMARY.* We consider the problem of estimating population mean $\bar{Y}$, of a character $y$, using information on some other parameters of $\bar{y}$. A class of estimators, which are linear function of $\bar{y}$ and a suitably chosen statistic $t$, is presented; general properties of the class are studied and the optimum weights and the resulting optimum mean square error are found. A general technique of generating estimators better than sample mean $\bar{y}$ and Searls' estimator (1964) is given and a member of such biased estimators are identified, for some choices of $t$, under very moderate conditions depending on the prior knowledge of the quantities which are smaller or greater than the actual values of some population parameters.

## 1. Introduction

Let $y$ be a variate (real) with population mean $\bar{Y}$, variance $\sigma_y^2$ and coefficient of variation $C_y(=\sigma_y/\bar{Y})$. Searls (1964) considered an estimator

$$T_1 = \lambda_1 \bar{y} \quad \text{... (1.1)}$$

for $\bar{Y}$ where $\bar{y}$ is the sample mean based on a simple random sample of size $n$ and $\lambda_1$ is a suitably chosen constant. In case $C_y$ is known exactly, the so called Searls' estimator

$$T_2 = [n/(n+C_y^2)]\bar{y}$$

with $\quad$ ... (1.2)

$$M(T_2) = \bar{Y}^2 C_y^2/(n+C_y^2)$$

is the best (in the sense of having smallest MSE) in the class of estimators $T_1$.

Hirano (1972) considered an estimator

$$T_3 = [n/(n+C_o^2)]\bar{y} \quad \text{... (1.3)}$$

in case a good guessed value of $C_y^2$, say $C_o^2$, is known.

In this paper, we have considered a class of estimators, for $\bar{Y}$, defined by

$$C_{\lambda v} = \{d : d = \boldsymbol{\lambda}'\boldsymbol{v}\} \quad \ldots \quad (1.4)$$

where

$$\boldsymbol{v}' = (\bar{y}, t) \text{ and } \boldsymbol{\lambda}' = (\lambda_1, \lambda_2)$$

t, being a suitably chosen statistics such that its variance $\sigma_t^2$ exist; and $\lambda_1$ and $\lambda_2$ being suitably chosen constants. Our main object is to present estimators better (in the sense of having smaller MSE) than those in $T_1$ (and hence better than $\bar{y}$, $T_2$ and $T_3$ also).

Let $\mu_r$ denote the $r$-th central moment of the character $y$ and

$$\beta_1 = \mu_3^2/\mu_2^3, \quad \beta_2 = \frac{\mu_4}{\mu_2^2}, \; \mu_2 = \sigma_y^2. \quad \ldots \quad (1.5)$$

Depending upon the situation, we shall assume the knowledge of $C_{(2)}$, $\beta_2^{(2)}$ which are such that

$$0 < C \leqslant C_2$$

and $\ldots \quad (1.6)$

$$\beta_2 \leqslant \beta_2^{(2)}.$$

## 2. Properties of the proposed class of estimators

The estimators in the proposed class $C_{\lambda v}$ are, in general, biased and their biases and mean square errors are given by

$$B(d) = \boldsymbol{\lambda}'\boldsymbol{\Psi} - \bar{Y} \quad \ldots \quad (2.1)$$

and

$$M(d) = \boldsymbol{\lambda}'G\boldsymbol{\lambda} - 2\bar{Y}\boldsymbol{\lambda}'\boldsymbol{\Psi}\bar{Y}^2 \quad \ldots \quad (2.2)$$

respectively, where

$$G = \begin{pmatrix} E\bar{y}^2 & E(\bar{y}t) \\ E(\bar{y}t) & Et^2 \end{pmatrix}, \quad \boldsymbol{\Psi} = \begin{pmatrix} \bar{Y} \\ Et \end{pmatrix}$$

It may be shown that the optimum choice $\boldsymbol{\lambda}_0$ of $\boldsymbol{\lambda}$, which minimizes $M(d)$, is a solution of

$$G\boldsymbol{\lambda} = \bar{Y}\boldsymbol{\Psi} \quad \ldots \quad (2.3)$$

which is a consistent equation, i.e., always yields a solution for $\boldsymbol{\lambda}$, and hence

$$\boldsymbol{\lambda}_0 = \bar{Y}\bar{G}\boldsymbol{\Psi} \quad \ldots \quad (2.4)$$

where $\bar{G}$ is a $g$-inverse of the matrix $G$.

The resulting (optimum) MSE of $d$ would be given by

$$M_0(d) = \bar{Y}^2[1-\Psi'(\bar{G})\Psi] \quad \dots \quad (2.5)$$

and the resulting bias would be

$$B_0(d) = -M_0(d)/\bar{Y}. \quad \dots \quad (2.6)$$

It may be noted that

$$M_0(d) = \bar{Y}^2[1-\lambda_0'\Psi/\bar{Y}]. \quad \dots \quad (2.7)$$

It may be noted that the matrix $G$, in general, is non-negative and would be non singular, if we exclude the trivial cases $\bar{y} = 0$ a.s. and $t = b\bar{y}$, where $b$ is a constant. In case $G$ is a positive definite matrix, $\lambda_0 = (\lambda_{01}, \lambda_{02})'$ and $M_0(d)$ would be given by

$$\lambda_{01} = \bar{Y}[(E\bar{y})(Et^2)-(Et)(Et\bar{y})]/D(\bar{y}, t) \quad \dots \quad (2.8)$$

$$\lambda_{02} = \bar{Y}[(Et)(E\bar{y}^2)-(E\bar{y}(E\bar{y}t)]/D(\bar{y}, t)$$

and

$$M_0(d) = \bar{Y}^2[1-N(\bar{y}, t)/D(\bar{y}, t)] \quad \dots \quad (2.9)$$

where

$$D(\bar{y},t) = (E\bar{y}^2)(Et^2)-(E\bar{y}t)^2$$
$$= \bar{Y}^2(Et)^2[(1-\rho^2)C_{\bar{y}}^2C_t^2+C_{\bar{y}}^2+C_t^2-2\rho C_{\bar{y}}C_t]$$

$$N(\bar{y}, t) = \bar{Y}^2(Et)^2[C_{\bar{y}}^2-2\rho C_{\bar{y}}C_t+C_t^2]$$

$C_{\bar{y}}$ and $C_t$ being the coefficients of variation of $\bar{y}$ and $t$ respectively and $\rho$ being the correlation coefficient between $\bar{y}$ and $t$.

It may be noted, from (2.9), that for any fixed $\rho$ the $M_0(d)$ is an increasing function of $C_t$. Hence if $t'$ and $t''$ are two choices of $t$, both having the same correlation with $\bar{y}$, then the use of $t'$ in $d$ would be preferrable over that $t''$ iff $C_{t'} < C_{t''}$.

From (2.9) and (1.2) it is found that if $\lambda_0 = (\lambda_{01}, \lambda_{02})'$ is known exactly, the optimum estimator $d_0 = \lambda_{01}\bar{y}+\lambda_{02}t$ would always be better than the Searls' estimator $T_2$ and hence than $\bar{y}$ and $T_1 = \lambda_1\bar{y}$ also. However, if $t$ is such that $\rho = C_{\bar{y}}/C_t$, $T_2$ and $d_0$ would be equally efficient.

In practice $\lambda_0$ would not be known, as it depends upon a number of parameters. The following technique would help, in that case, to generate estimators from $d$ better than $\bar{y}$, $T_1$ and $T_2$.

From (2.2), we may write

$$M(d) = M(T_1)+\lambda_2^2 Et^2-2\lambda_2\{\bar{Y}(Et)-\lambda_1 E(\bar{y}t)\}. \qquad \ldots \quad (2.10)$$

For a specified $\lambda_1$, the estimator $d = T_1+\lambda_2 t$ would be better than $T_1 = \lambda_1\bar{y}$

$$\text{iff } \lambda_2 \text{ lies between 0 and } 2\lambda_{02}^* \qquad \ldots \quad (2.11)$$

where

$$\lambda_{02}^* = [(1-\lambda_1)\bar{Y}(Et)-\lambda_1 \text{ cov}(\bar{y}, t)]/E(t^2)$$

is the optimum choice of $\lambda_2$, for fixed $\lambda_1$. in $d$.

For a specific choice of the statistic $t$ and specified $\lambda_1$, we shall find, as in (2.11), that, the estimator $d$ would be better than $\bar{y}$ or $T_2$ iff $\lambda_2$ lies between 0 and $2\lambda_0^*$, say. Obviously, $\lambda_0^*$ would be a function of some unknown population parameters, say $\Phi$ and the vector $\Phi$ can be decomposed into component vectors say, $\Phi_1$, $\Phi_2$, $\Phi_3$ such that $|\lambda_0^*|$ is non-decreasing in each component of $\Phi_1$ and non-increasing in each component of $\Phi_2$. If $\Phi_1^*$, $\Phi_2^*$, $\Phi_3^*$ are known quantities such that $\Phi_1 \geqslant \Phi_1^*$, $\Phi_2 \leqslant \Phi_2^*$, $\Phi_3 = \Phi_3^*$ hold and moreover $\text{sgn}[\lambda_0^*(\phi)]$ is known then $\mu^* = \text{sgn}[\lambda_0^*(\phi)]\,|\lambda_0^*(\phi^*)|$ is a known quantity and then it is obvious that, for a given $\lambda_1$, we shall have $M(d) \leqslant V(\bar{y})$ or $M(d) < M(T_2)$ for all $\mu$ such that either $0 < \mu < 2\mu^*$ or $2\mu^* < \mu < 0$ holds.

Let $\lambda_1^* = n/[n+C_*^2]$, where $C_*^2 < C_y^2$ so that the estimator $T_1^* = \lambda_1^*\bar{y}$ is better than the sample mean $\bar{y}$.

Let

$$d^* = T_1^*+\lambda_2 t. \qquad \ldots \quad (2.12)$$

We shall make use of the notation (1.6) and

$$\Delta = \beta_2+(n^2-2n+3)/(n-1)$$

$$\Delta_{(2)} = \beta_2^{(2)}+(n^2-2n+3)/(n-1) \qquad \ldots \quad (2.13)$$

$$a = \begin{cases} +1 & \text{for positively skewed distributions} \\ 0 & \text{for symmetrical distributions} \\ -1 & \text{for negatively skewed distribution.} \end{cases}$$

Using (2.10), (2.11) it may be shown that

$$d_1^* = T_1^*+\lambda_2(s^2/\bar{y})$$

would be better than $T_1^*$ (and hence $\bar{y}$ too) iff $\lambda_2$ lies between

$$0 \text{ and } \frac{2[n(C_*^2+C_y^2)+C_y^2C_*^2-a\sqrt{\beta_1}C_y(C_*^2+n)]}{(n+C_*^2)C_y^2\Delta''} \quad \ldots \quad (2.14)$$

and a set of sufficient condition for $d_1^*$ to be better than $T_1^*$ and $\bar{y}$, in case of symmetrical distributions, would be

$$0 < \lambda_2 < \frac{2[n(C_*^2+C_{(2)}^2)+C_{(2)}^2C_*^2]}{(n+C_*^2)C_{(2)}^2(\Delta_{(2)}+3C_{(2)}^2)}\,. \quad \ldots \quad (2.15)$$

The sufficient condition in case of positively and negatively skewed distributions may be obtained likewise by using the appropriate bounds of involved parameters.

### Reference

Hirano, K. (1972): Using some approximately known coefficient of variation in estimating the mean, Research Memorandum No. 49, *Inst. Stat. Math.*, Tokyo, Japan.

Searls, D, T. (1964): The utilisation of a known coefficient of variation in the estimation procedure, *Jour. Amer. Statist. Assoc.*, 59, 1225–26.