Jour. Ind. Soc. Ag. Statistics
Vol. XLII, No. 1 (1990), pp. 98-107

ESTIMATION OF PARAMETERS IN NORMAL PARENT

HOUSILA PRASAD SINGH
J.N.K. Vishwa Vidyalaya, Jabalpur, M.P.
(Received : June, 1986)

SUMMARY

For estimating the square of the population mean μ^2 and variance σ^2, some estimators are defined and their properties are analyzed in the context of normal population under large sample approximation.

Keywords : Coefficient of variation, Bias, Mean Squared Error, Normal Parent.

Introduction

In many situations of practical importance, the coefficient of variation exhibits stability and its value may be fairly accurately known. In such cases the estimation of variance reduces to the estimation of the square of the population mean μ^2. There are other instances when μ^2 may be the parameter of interest, see, e.g. Govindarazulu and Sahai [2] and Upadhyaya and Singh [8].

Suppose a random sample of size n is drawn from a normal parent $N(\mu, C^2\mu^2)$ where $C^2 = \sigma^2/\mu^2$ is the square of the coefficient of variation. If C^2 is exactly known, an unbiased estimator of μ^2 is

$$t = \frac{\bar{x}^2}{1 + C^2/n} = \frac{n}{n + C^2}\,\bar{x}^2 = \left[1 - \frac{C^2}{n + C^2}\right]\bar{x}^2 \quad (1.1)$$

and its relative variance is

$$RV(t) = E\left[\frac{t - \mu^2}{\mu^2}\right]^2 = \frac{2C^2\,(2n + C^2)}{(n + C^2)^2} \quad (1.2)$$

where $\bar{x} = n^{-1} \sum_{i=1}^{n} x_i$ is the sample mean based on n observations.

In practice C^2 may not be exactly known. Then the estimator t is of little utility. However, if C^2 has shown stability in repeated experiments, one can use a guessed value C_g^2 instead of C^2. But if nothing is known about C^2, then only alternative is to estimate it from the sample data at hand. We may employ the following two consistent estimators of C^2 (see, for example, Srivastava [4 : 34-35]) :

$$\left.\begin{aligned} \hat{C}_1^2 &= \frac{s^2}{\bar{x}^2}\left(1 - \frac{s^2}{n\bar{x}^2}\right)^{-1} \\ \hat{C}_2^2 &= \frac{s^2}{\bar{x}^2} \end{aligned}\right\} \quad (1.3)$$

where $s^2 = (n-1)^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ is an unbiased estimator of σ^2 based on n observations.

The substitution of $\hat{C}_1^2$ and $\hat{C}_2^2$ in (1,1) led Das [1] and Pandey [3] to formulate some estimators for μ^2. Some known estimators of μ^2 when C^2 is unknown are compiled in Table 1.

TABLE 1—ESTIMATORS OF μ^2 WHEN C^2 IS UNKNOWN

Reference	*Estimator*
An estimator	$T_0 = \bar{x}^2$
Minimum variance unbiased estimator (MVUE)	$T_1 = \bar{x}^2 - \frac{s^2}{n}$
Srivastava, Dwivedi and Bhatnagar [5]	$T_2 = \bar{x}^2 - \left(\frac{n-1}{n+1}\right)\frac{s^2}{n}$
Das [1]	$T_3 = \bar{x}^2\left(1 + \frac{s^2}{n\bar{x}^2}\right)^{-1}$
Pandey [3]	$T_4 = \bar{x}^2\left[1 + \frac{s^2}{n\bar{x}^2}\left(1 + \frac{s^2}{n\bar{x}^2}\right)^{-1}\right]$

Das [1] indicated that the probability of T_1 being negative may be significant for small samples and defined the following modified estimator :

$$T_1^1 = \begin{cases} T_1 & \text{if } \bar{x}^2 \geqslant \frac{s^2}{n} \\ 0 & \text{if } \bar{x}^2 < \frac{s^2}{n} \end{cases}$$

Owing to lack of exact expressions for the biases and mean squared errors of T_1^1 and T_3 considered by Das [1] and T_4 reported by Pandey [3], resort was made to Monte Carlo experiments for comparative study of estimators. Such experiments are not very conclusive because of their well known limitations. The exact expressions for biases and mean squared errors T_1^1 and T_3 are derived by Srivastava, Dwivedi and Bhatnagar [5]. But they too fail to draw clear cut inference as the results obtained by them are quite intricate. It is to be noted that the expressions for bias and mean squared error of Pandey [3] estimator T_4 can also be obtained following the procedure cited in [5]. However, it may be observed that T_1^1 will be very close to T_1 for large n and thus will possess the same large sample properties as T_1. For T_3 and T_4, the large sample approximations for the relative biases and relative mean squared errors are derived by Srivastava, Dwivedi and Bhatnagar [5] and Pandey [3], respectively.

The relative biases (RB) and relative mean squared errors (RMSE) of the estimators T_i, $i = 0$ to 4 cited in Table 1 are compiled in Table 2.

TABLE 2—THE RELATIVE BIASES AND RELATIVE MEAN SQUARED ERRORS OF T_i, $i = 0$ to 4.

Estimator	*Relative bias to Terms of Order* 0 (n^{-3})	*Relative MSE to Terms of Order 0* (n^{-3})
T_0	$\text{RB}(T_0) = C^2/n$	$\text{RMSE}(T_0) = \frac{C^2}{n}\left(4 + \frac{3\,C^2}{n}\right)$
T_1	$\text{RB}(T_1) = 0$	$\text{RMSE}(T_1) = \frac{2\,C^2}{n}\left(2 + \frac{C^2}{n}\right)$
T_2	$\text{RB}(T_2) = \frac{2\,C^2}{n\,(n+1)}$	$\text{RMSE}(T_2) = \frac{2\,C^2}{n}\left[2 + \left(\frac{n+2}{n+1}\right)\frac{C^2}{n}\right]$
T_3	$\text{RB}(T_3) = \frac{C^4}{n^2}\left[1 + \frac{2\,(1+C^2)}{n}\right]$	$\text{RMSE}(T_3) = \frac{2\,C^2}{n}\left[2 + \left(\frac{n+1}{n^2}\right)C^2\right]$
T_4	$\text{RB}(T_4) = \frac{C^6}{n_3}$	$\text{RMSE}(T_4) = \frac{2\,C^2}{n}\left[2 + \left(\frac{n+1}{n^2}\right)C^2\right]$

In Sections 2 and 3 some estimators are suggested for μ^2 and σ^2, respectively and their properties studied under large sample approximation.

2. Estimators For μ^2

Consider the following estimators for μ^2 :

$$t_1 = \bar{x}^2\left[1 + \frac{s^2}{n\,\bar{x}^2}\left(1 - \frac{s^2}{n\,\bar{x}^2}\right)^{-1}\right]^{-1} \tag{2.1}$$

$$t_2 = \bar{x}^2\left[1 + \frac{s^2}{n\bar{x}^2}\left\{1 + \frac{s^2}{n\bar{x}^2}\left(1 - \frac{s^2}{n\,\bar{x}^2}\right)^{-1}\right\}\right]^{-1} \tag{2.2}$$

$$t_3 = \bar{x}^2\left[1 + \frac{s^2}{n\,\bar{x}^2}\left\{1 + \frac{s^2}{n\,\bar{x}^2}\left(1 + \frac{s^2}{n\,\bar{x}^2}\right)\right\}\right]^{-1} \tag{2.3}$$

$$t_4 = \bar{x}^2\left[1 - \frac{s^2}{n\,\bar{x}^2}\left(1 - \frac{s^2}{n\,\bar{x}^2}\right)^{-1}\left\{1 + \frac{s^2}{n\,\bar{x}^2}\left(1 - \frac{s^2}{n\,\bar{x}^2}\right)^{-1}\right\}^{-1}\right] \tag{2.4}$$

In order to derive the expressions for bias and mean squared error of t_i, $i = 1$ to 4, under large sample approximation, we write

$\bar{x} = \mu\,(1 + \varepsilon_0)$ and $s^2 = \sigma^2\,(1 + \varepsilon_1)$

such that $E\,(\varepsilon_i) = 0$, $i = 0, 1$ and $|\,\varepsilon_i\,| < 1$, $i = 0, 1$.

Thus substituting the values of $\bar{x}$ and s^2 in terms of ε_i's in (2.1), (2.2), (2.3) and (2.4) we have

$$t_1 = \mu^2\,(1 + \varepsilon_0)^2\left[1 + \frac{C^2}{n}\,(1 + \varepsilon_1)\,(1 + \varepsilon_0)^{-2} \times \left\{1 - \frac{C^2}{n}\,(1 + \varepsilon_1)\,(1 + \varepsilon_0)^{-2}\right\}^{-1}\right]^{-1} \tag{2.5}$$

$$t_2 = \mu^2(1 + \varepsilon_0)^2\left[1 + \frac{C^2}{n}\,(1 + \varepsilon_1)\,(1 + \varepsilon_0)^{-2} \times \left\{1 + \frac{C^2}{n}\,(1 + \varepsilon_1)\,(1 + \varepsilon_0)^{-2}\left(1 - \frac{C^2}{n}\,(1 + \varepsilon_1)\,(1 + \varepsilon_0)^{-2}\right\}\right]^{-1} \tag{2.6}$$

$$t_3 = \mu^2 (1 + \varepsilon_0)^2 \left[1 + \frac{C^2}{n} (1 + \varepsilon_1)(1 + \varepsilon_0)^2 \right.$$

$$\left. \times \left\{ 1 + \frac{C^2}{n} (1 + \varepsilon_1)(1 + \varepsilon_0)^{-2} \left(1 + \frac{C^2}{n} (1 + \varepsilon_1)(1 + \varepsilon_0)^{-2} \right) \right\} \right]^{-1} \quad (2.7)$$

$$t_4 = \mu^2 (1 + \varepsilon_0) - \frac{(1 + \varepsilon_1)\mu^2 C^2}{n} \left\{ 1 - \frac{C^2}{n} (1 + \varepsilon_1)(1 + \varepsilon_0)^{-2} \right\}$$

$$\times \left\{ 1 + \frac{C^2}{n} (1 + \varepsilon_1)(1 + \varepsilon_0)^{-2} \left(1 - \frac{C^2}{n} (1 + \varepsilon_1)(1 + \varepsilon_0)^{-2} \right)^{-1} \right\}^{-1} \quad (2.8)$$

Now expanding the right hand sides of equations (2.5) to (2.8), retaining the terms to order $O(n^{-3})$ and simplifying, we get

$$E(t_i) = \mu_2 \Rightarrow \text{Bias}(t_i) = 0 \Rightarrow \text{RB}(t_i) = 0. \qquad i = 1 \text{ to } 4 \qquad (2.9)$$

The mean squared error of t_i, $i = 1$ to 4 can easily be obtained to terms of order $O(n^{-3})$ as

$$\text{MSE}(t_i) = \frac{2\mu^4 C^2}{n} \left[2 + \left(\frac{n+1}{n^2} \right) C^2 \right] \qquad i = 1 \text{ to } 4 \qquad (2.10)$$

$$\Rightarrow \text{RMSE}(t_i) = \frac{MSE(t_i)}{\mu^4} = \frac{2C^2}{n} \left[2 + \left(\frac{n+1}{n^2} \right) C^2 \right]$$

$$i = 1 \text{ to } 4 \qquad (2.11)$$

From Table 2, (2.9) and (2.11) we observe that

$$0 = RB(t_i) \leqslant RB(T_4) \leqslant RB(T_3) \qquad i = 1 \text{ to } 4 \qquad (2.12)$$

and

$$\text{RMSE}(t_i) = \text{RMSE}(T_3) = \text{RMSE}(T_4) \qquad i = 1 \text{ to } 4 \qquad (2.13)$$

Thus for larger sample sizes, the proposed estimator t_i, $i = 1$ to 4, which are unbiased, are preferable over Pandey [3] estimator.

3, Estimators for σ^2 When C^2 is Unknown

The population variance σ^2 is equal to $C^2 \mu^2$. Its estimators will be $\hat{\sigma}^2 = \hat{C}^2 \hat{\mu}^2$. On the basis of the estimators for μ^2 reported in Table 1, Pandey [3] forwarded the following estimators for σ^2 as

$$T_1^* = s^2 \tag{3.1}$$

$$T_2^* = s^2 - \frac{s^4}{n \bar{x}^2} \tag{3.2}$$

$$T_3^* = s^2 \left(1 + \frac{s^2}{n \bar{x}^2} \right)^{-1} \tag{3.3}$$

$$T_4^* = s^2 \left[1 + \frac{s^2}{n \bar{x}^2} \left(1 + \frac{s^2}{n \bar{x}^2} \right) \right]^{-1} \tag{3.4}$$

The maximum likelihood estimator of σ^2 is

$$T_5^* = \frac{(n-1)}{n} s^2 \tag{3.5}$$

The relative biases and relative mean squared errors of T_i^*, $i = 1$ to 5 to terms of order $O(n^{-2})$ are compiled in Table 3.

TABLE 3—THE RELATIVE BIASES AND RELATIVE MEAN SQUARED ERRORS OF T_i^* ; $i = 1$ to 5

Estimator	*Relative Bias to Terms of Order 0 (n^{-2})*	*Relative Mean Squared Error to Terms of Order 0 (n^{-2})*
T_1^*	$RB(T_1^*) = 0$	$RMSE(T_1^*) = 2/(n-1)$
T_2^*	$RB(T) = -\frac{C}{n}\left[1 + \frac{(2 + C^2)}{n}\right]$,	$RMSE(T_2^*) = \frac{2}{n}\left[1 + \frac{C^2}{2n}(C^2 - 8)\right]$
T_3^*	$RB(T_3^*) = -\frac{C^2}{n}\left[1 + \frac{2(1 + C^2)}{n}\right]$,	$RMSE(T_3^*) = RMSE(T_2^*)$
T_4^*	$RB(T_4^*) = -\frac{C^2}{n}\left[1 + \frac{(2 + 3C^2)}{n}\right]$,	$RMSE(T_4^*) = RMSE(T_2^*)$
T_5^*	$RB(T_5^*) = -1/n$	$RMSE(T_5^*) = (2n-1)/n^2$

On the basis of similar arguments we propose four estimators for σ^2 as

$$t_1^* = s^2 \left[1 + \frac{s^2}{n\bar{x}^2} \left(1 - \frac{s^2}{n\bar{x}^2} \right)^{-1} \right]^{-1} \tag{3.6}$$

$$t_2^* = s^2 \left[1 + \frac{s^2}{n\bar{x}^2} \left\{ 1 + \frac{s^2}{n\bar{x}^2} \left(1 - \frac{s^2}{n\bar{x}^2} \right)^{-1} \right\}^{-1} \right]^{-1} \tag{3.7}$$

$$t_3^* = s^2 \left[1 + \frac{s^2}{n\bar{x}^2} \left\{ 1 + \frac{s^2}{n\bar{x}^2} \left(1 + \frac{s^2}{n\bar{x}^2} \right) \right\} \right]^{-1} \tag{3.8}$$

$$t_4^* = s^2 \left[1 - \frac{s^2}{n\bar{x}^2} \left(1 - \frac{s^2}{n\bar{x}^2} \right)^{-1} \left\{ 1 + \frac{s^2}{n\bar{x}^2} \left(1 - \frac{s^2}{n\bar{x}^2} \right)^{-1} \right\}^{-1} \right] \tag{3.9}$$

It is easy to verify that the relative biases and relative mean squared errors of t_i^*, $i = 1$ to 4, to terms of order $O(n^{-2})$, are

$$RB(t_i^*) = -\frac{C^2}{n} \left[1 + \frac{1}{n} (2 + 3C^2) \right], \quad i = 1 \text{ to } 4 \tag{3.10}$$

and

$$\text{RMSE}(t_i^*) = \frac{2}{n} \left[1 + \frac{C^2 (C^2 - 8)}{2n} \right], \quad i = 1 \text{ to } 4. \tag{3.11}$$

It follows from Table 3, (3.10) and (3.11) that the proposed estimators are as good as Pandey [3] T_4^* estimator. But these estimators are inferior to T_3^*, since T_3^* has smaller absolute relative bias and same RMSE. Further it may be concluded that these estimators are better than maximum likelihood estimator T_5^* if

(i) sample size is small but coefficient of variation is large, and
(ii) sample size is large but coefficient of variation is small.

Keeping in view the form of Srivastava and Banarasi [7] estimator of the population mean μ, we can propose a class of estimators for σ^2 as

$$t_5^* = s^2 \left[1 - \frac{s^2}{n\bar{x}^2} \left(1 + \frac{s^2}{n\bar{x}^2} \right)^{-\alpha} \right] \tag{3.12}$$

where α is nonstochastic.

To terms of order $O(n^2)$, the relative bias and relative mean squared error of T_5^* are, respectively, given by

$$RB(t_5^*) = -\frac{C^2}{n}\left[1 + \frac{1}{n}\left\{2 + (3 - \alpha) C^2\right\}\right], \tag{3.13}$$

and

$$\text{RMSE}(T_5^*) = \frac{2}{n}\left[1 + \frac{C^2(C^2 - 8)}{2n}\right] \tag{3.14}$$

From Table 3, (3.13) and (3.14) we have

$$\left.\begin{array}{l}\text{ARB}(t_5^*) = \text{RMSE}(T_3^*) \\ \text{RMSE}(t_5^*) = \text{RMSE}(T_3^*)\end{array}\right\} \text{ for } \alpha > 1 \tag{3.15}$$

and

$$\left.\begin{array}{l}\text{ARB}(t_5^*) < \text{ARB}(T_4^*) \\ \text{RMSE}(t_5^*) = \text{RMSE}(T_4^*)\end{array}\right\} \text{ for } \alpha > 0 \tag{3.16}$$

where ARB (.) = RB (.) stands for absolute relative bias.

From (3.15) and (3.16) it follows that the estimator t_5^*, for all values of $\alpha > 1$, is superior to the proposed estimators t_i^*, $i = 1$ to 4 and the estimators T_2^*, T_3^* and T_4^* considered by Pandey [3].

Motivated by Srivastava and Bhatnagar [6], we suggest the two parameter family of estimators for σ^2 as

$$t_6^* = s^2\left[1 + \frac{Ks^2}{n\bar{x}^2}\left(1 + \frac{gs^2}{n\bar{x}^2}\right)^{-1}\right] \tag{3.17}$$

where K and g are characterizing scalars.

Assuming the characterizing scalars to be non-stochastic, it is easy to see that the relative bias and relative mean squared error of t_6^*, to terms of order $O(n^{-3})$ are

$$\text{RB}(t_6^*) = \frac{KC^2}{n}\left[1 + \frac{1}{n}\left\{2 + (3 - g) C^2\right\}\right] \tag{3.18}$$

and

$$\text{RMSE}\,(t_6^*) = \frac{2}{n}\left[1 + \frac{KC^2}{2n}\,(KC^2 + 8)\right] \tag{3.19}$$

It is to be noted that for $K = -1$ and $g = \alpha$ the expressions for relative bias and relative mean squared error of t_6^* turn out to be the same as t_5^*.

From Table 3 and (3.19) we see that the proposed estimator t_6^* is more efficient than Pandey [3] estimator T_3^*, if

$$\left.\begin{array}{ll} \text{either } -1 < K < \dfrac{(C - 8)}{C^2}; & C^2 > 8 \\ \text{or } \quad \dfrac{(C^2 - 8)}{C^2} < K < -1; & C^2 \ngtr 8 \end{array}\right\} \tag{3.20}$$

Minimization of RMSE (t_6^*) with respect to K yields

$$K_{min} = -4/C^2 \tag{3.21}$$

so that

$$\text{min. RMSE}\,(t_6^*) = \frac{2}{n}\left(1 - \frac{8}{n}\right) \tag{3.22}$$

Thus if C^2 is exactly known then the resulting estimator

$$t_6^{**} = s^2\left[1 - \frac{4}{C^2}\cdot\frac{s^2}{n\,\bar{x}^2}\left(1 + \frac{g\,s^2}{n\,\bar{x}^2}\right)^{-1}\right] \tag{3.23}$$

is more efficient than maximum likelihood estimator T_5^*.

ACKNOWLEDGEMENTS

The author is grateful to the referee for his valuable suggestions in improving the presentation of the paper. Author also thanks to Mr. A. S. Thakur for typing the paper.

REFERENCES

[1] Das, B. (1975) : Estimation of μ^2 in normal population, *Cal. Stat. Assn. Bull.* 24 : 135-40.

[2] Govindarajulu, Z. and Sahai, H. (1972) : Estimation of the parameters of a normal distribution with known coefficient of variation, *Rep. Stat. App. Res. JUSU*, 19 : 86-98.

[3] Pandey, B. N. (1981) : On estimation of variance in normal distribution, *Jour. Ind. Soc. Agri. Statist.* 33 (3) : 1-6.

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

[5] Srivastava, V. K., Dwivedi, T. D., and Bhatnagar, S. (1980) : Estimation of the square of mean in normal population. *Statistica*, XL (4) : 455-466.

[6] Srivastava, V. K. and Bhatnagar, S. (1983) : A revisit to the use of coefficient of variation in estimating mean, *Jour. Ind. Soc. Agri. Statist.* 35 (1) : 67-71.

[7] Srivastava, S. R. and Banarasi (1982) : A note on the estimation of mean of a symmetrical population, *Jour. Ind. Soc. Agri. Statist.* 34 (3) : 79-81.

[8] Upadhyaya, L. N. and Singh, H. P. (1984) : On the estimation of population mean with known coefficient of variation, *Biometrical Journal* 26 (6) : 915-922.