Jour. Ind. Soc. Ag. Statistics
Vol. XXXVII, No. 3 (1985), pp. 226-230

A NOTE ON ESTIMATION OF AMOUNT OF INFORMATION IN NORMAL SAMPLES

AMARENDRA MISHRA
Patna University, Patna-800 005

(Received : January, 1980)

SUMMARY

An unbiased estimator and a minimum mean square error estimator of the amount of information provided by an observed value x regarding the unknown parameter μ of the normal population, when the population variance σ^2 is unknown, have been suggested. The relative efficiencies of these estimators have also been obtained.

Keywords : Unbiased estimator, efficiency.

Introduction

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from a normal population with mean μ and variance σ^2. When the parameter under consideration is μ, the amount of information provided by each x_i $(i = 1, 2, \ldots, n)$ is $\theta = 1/\sigma^2$. In case σ^2 is unknown, Fisher [1] using the student's t distribution obtained the estimator $T_1 = n/(n + 2) \cdot (1/s^2)$ which is biased for θ. Here s^2 is an unbiased estimator of σ^2. In this paper an unbiased estimator T_2 of θ has been obtained which is better than T_1. An estimator $T_3 = M/s^2$ where M is a scalar and is to be determined such that the mean square error of T_3 is minimum, has also been considered. All the three estimator have been compared with each other.

2. Unbiased Estimator T_2

It can be shown that T_1 is not an unbiased estimator of θ, for, evaluating $E\,(1/s^2)$ from the sampling distribution of s^2 for normal samples

$$E\,(1/s^2) = \frac{n-1}{n-3}\,\theta \tag{1}$$

and thus,

$$E(T_1) = \frac{n(n-1)}{n(n-3)(n+2)}\,\theta \neq \theta \tag{2}$$

However, (1) suggests that the unbiased estimator of θ is given by

$$T_2 = \frac{n-3}{n-1} \cdot \frac{1}{s^2} \tag{3}$$

defined for $n > 3$.

Take the difference between T_1 and T_2,

$$T_1 - T_2 = \frac{\sigma}{(n-1)(n+2)} \cdot \frac{1}{s^2} \tag{4}$$

which is always a positive quantity. Therefore, T_1 is greater than T_2 for all n.

Evaluating $E(1/s^2)^2$, we get

$$E(1/s^2)^2 = \frac{(n-1)^2}{(n-3)(n-5)}\,\theta^2 \tag{5}$$

From (1) and (5) we get

$$\text{Var}(1/s^2) = \frac{2(n-1)^2}{(n-3)^2(n-5)}\,\theta^2 \tag{6}$$

and thus

$$\text{Var}\,(T_2) = \frac{(n-3)^2}{(n-1)^2}\,\text{Var}\,(1/s^2) = \frac{2}{n-5}\,\theta \tag{7}$$

The MSE of T_1 using (1) and (5) is obtained as

$$\text{MSE}(T_1) = \frac{2(n+3)(n^2-2n+10)}{(n+2)^2(n-3)(n-5)}\,\theta^2 \tag{8}$$

Therefore, the relative efficiency of T_2 with respect to T_1 is

$$\begin{aligned} \text{REF}\,(T_2, T_1) = \frac{\text{MSE}(T_1)}{\text{MSE}(T_2)} &= \frac{(n+3)(n^2-2n+10)}{(n-3)(n+2)^2} \\ &= 1 + \frac{12n+42}{(n-3)(n+2)^2} \end{aligned} \tag{9}$$

which shows that T_2 is more efficient than T_1 for all $n > 3$. T_2 is defined only for $n > 3$ and so it is meaningful to talk of relative efficiency for $n \geqslant 4$ only.

Table 1 gives the relative efficiency of T_2 with respect to T_1 at different values of $n \geqslant 4$.

TABLE 1—THE RELATIVE EFFICIENCY (%) OF T_2 WITH RESPECT TO T_1

Sample Size	4	5	6	7	8
Efficiency	350.00	204.08	157.37	138.89	127.60
Sample Size	9	10	15	20	30
Efficiency	120.66	116.07	106.40	103.43	101.45
Sample Size	50	100	200		
Efficiency	100.51	100.12	100.12		

It can be seen that in the beginning there is sharp fall in the relative efficiency, as n increases, after $n = 15$ the fall is very slow and after $n = 100$ the relative efficiency is constant upto the fourth place of decimal.

3. Minimum MSE Estimator T_3

Define a new estimator T_3 as

$$T_3 = M/s^2 \tag{10}$$

where M is an unknown scalar. Determine M to minimize the MSE of T_3. We have

$$\text{MSE}\,(T_3) = \text{Var}\,(T_3) + [\text{bias}\,(T_3)]^2 \tag{11}$$

Now

$$\text{Var}\,(T_3) = M^2\,\text{Var}\,(1/s^2) = \frac{2M^2(n-1)^2}{(n-3)^2(n-5)}\,\theta^2 \tag{12}$$

and
$$[\text{bias}\,(T_3)]^2 = \left(\frac{M(n-1)}{(n-3)} - 1\right)^2 \theta^2 \tag{13}$$

Therefore, from (12) and (13), we have

$$\text{MSE }(T_3) = \frac{2M^2(n-1)^2}{(n-3)^2(n-5)}\,\theta^2 + \left(\frac{M(n-1)}{(n-3)} - 1\right)^2\theta^2 \tag{14}$$

To minimize (14) differentiating it with respect to M and equating to zero, we get

$$\frac{d\,\text{MSE }(T_3)}{dM} = \frac{4M(n-1)^2\theta^2}{(n-3)^2(n-5)} + \frac{2(n-1)}{n-3}\left(\frac{M(n-1)}{(n-3)} - 1\right)\theta^2 = 0 \tag{15}$$

which gives $M = \dfrac{n-5}{n-1}$

Therefore, the estimator T_3 is obtained as

$$T_3 = \frac{n-5}{n-1}\cdot\frac{1}{s^2} \tag{16}$$

defined for $n > 5$. From (14) we have

$$\text{MSE }(T_3) = \frac{2}{n-3}\,\theta^2 \tag{17}$$

The relative efficiency of T_3 with respect to T_2 is

$$\text{REF }(T_3, T_2) = \frac{\text{MSE }(T_2)}{\text{MSE }(T_3)} = \frac{n-3}{n-5} = 1 + \frac{2}{n-5} \tag{18}$$

which shows that T_3 is more efficient than T_2 for all $n > 5$.

As T_3 is defined only for $n > 5$ the relative efficiency of T_3 with respect to T_2 is meaningless for $n \leqslant 5$. Table 2 on p. 230 gives the relative efficiency of T_3 with respect to T_2 at different values of $n \gtr 5$. Further,

$$\text{REF }(T_3, T_1) = \frac{\text{MSE}(T_1)}{\text{MSE}(T_3)} = \frac{(n+3)(n^2-2n+10)}{(n+2)^2(n-5)} = 1 + \frac{2(n+5)^2}{(n+2)^2(n-5)} \tag{19}$$

This shows that T_3 is more efficient than T_1 and T_2 for all $n \geqslant 6$.

TABLE 2—THE RELATIVE EFFICIENCY (%) OF T_3 WITH RESPECT TO T_2

Sample Size	6	7	8	9	10	15
Efficiency	300.00	200.00	166.67	150.00	140.00	120.00
Sample Size	20	25	30	50	100	200
Efficiency	113.33	110.00	108.00	104.44	102.11	101.03
						500
						100.40

REFERENCE

[1] Fisher, R. A. (1935) : *The Design of Experiments*, 9th Ed., Hafner Press, New York.