EFFICIENT ESTIMATION OF NORMAL POPULATION PARAMETER USING ITS KNOWN COEFFICIENT OF VARIATION

BY

GOVIND PRASAD AND ASHOK SAHAI

University of Roorkee

SUMMARY

The minimum mean square error estimator from a bigger class of estimators linear in srmple means $\bar{x}$, sample standard deviation s and $(\bar{x}^3/s^2)$ has been obtained. The new estimator happens to be more efficient than that obtained by some other workers. The gain in efficiency is illustrated through an empirical study.

I. INTRODUCTION

Searles [3] considered the problem of estimating θ in $N(\theta, a\theta^2)$ from a random sample of size, say n; $x_1, x_2, \ldots x_n$. The coefficient of variation, say, C $(C^2=a)$ is supposed to be known. Let

$$\bar{x}=\frac{\Sigma x_i}{n} \text{ and } s^2=\Sigma (x_i-\bar{x})^2/(n-1)$$

be the sample mean and the sample varience, respectively. Searles [3] used the sample mean only, to generate his estimator exploiting the known value of $C(=+\sqrt{a})$; Whereas Khan [2] and Gleser and Healy [1] considered both '$\bar{x}$' as well as 's' $(=+\sqrt{s^2})$ to generate their estimators which are, respectively, the minimum variance unbiased (mvu) and the minimum mean square error (mmse) estimators in the class of estimators linear in '$\bar{x}$' and 's'.

We know that

$$\bar{x}\sim N(\theta, r\theta^2) \qquad \text{where} \qquad r=\frac{a}{n}$$

and

$$(n-1) \frac{s^2}{(a\theta^2)}\sim\chi^2_{n-1}.$$

Therefore $$E(s^l)=\left(\frac{2}{n-1}\right)^{l/2}\frac{\left|\frac{n+l-1}{2}\right.}{\left|\frac{l-1}{2}\right.}(a\theta^2)^{l/2};\ l=\pm 1,2,\pm 3$$

$$=K_{(l)}\,(a)^{l/2}\,\theta^l\ ;\ \text{say} \qquad ...(1.1a)$$

$$=K_{(l)}\,\theta^l,\quad \text{where}\quad K'_{(l)}=a^{l/2}\,K_{(l)} \qquad ...(1.1b)$$

Consequently $\bar{x}$ as well as $\left(\frac{s}{K'_{(l)}}\right)=s'$, say, are unbiased estimators of θ, with

$$\text{Var.}\,(\bar{x})=r\theta^2=v_1\theta^2,\ \text{say},$$

and $$\text{Var.}\,(s^t)=(1-K^2_{(1)})\theta^2/K^2_{(1)}=v_2\theta^2,\ \text{say}.$$

Then as found by Gleser and Healy [1]

$$T^*_{LMMS}=\alpha_1\,s+\alpha_2\,\bar{x},$$

where $$\alpha_i=\frac{v_i}{(v_1+v_2+v_1\,v_2)}\ ;\ i=1,2,$$

the mmse (or linear minimum mean square estimator) estimator and its mean square error, say M^*, is given by :

$$M^*=\frac{v_1\,v_2}{(v_1+v_2+v_1\,v_2)}.$$

Thus the Relative Efficiency (RE) of T^*_{LMMS} as compared to the usual sample mean estimator $\bar{x}$, say E_1, is given by

$$E_1=\frac{(v_1+v_2+v_1\,v_2)}{v_2}\times 100\%$$

2. The Alternative Estimator

We simply consider a bigger class of estimators linear in $\bar{x}$, s and $\bar{x}^3/s^2$ and obtain the mmse estimator in the class. We easily check, in this context, that

$$E(\bar{x})=\theta,\ E(\bar{x}^2)=b\theta^2, E(\bar{x}^3)=(3b-2)\,\theta^3,$$

$$E(\bar{x}^4)=(3b^2-2)\,\theta^4$$

and $$E(\bar{x}^6)=(15b^3-30b+16)\theta^6 \qquad ...(2.1)$$

where $$b=(1+r)$$

Let $$T'=A\bar{x}+Bs+C\,\frac{\bar{x}^3}{s^2}.$$

It follows from the stochastic independence of $\bar{x}$ and s, that

$$MSE(T')=M', \text{ say}$$
$$=A^2E(\bar{x}^2)+B^2E(s^2)+C^2E(\bar{x}^6).E(s^4)$$
$$+2\,ABE(\bar{x}).E(s)+2\,ACE(\bar{x}^4).E(s^{-2})$$
$$+2\,BCE(\bar{x}^3)E(s^{-1})-2\,A\theta E(\bar{x})-2\,B\theta E(s)$$
$$-2\,C\theta E(\bar{x}^3)\,E(s^{-2})+\theta^2.$$

Using (1.1a), (1.1b) and (2.1), it is easily checked that the normal equations for the minimum of M' are :

$$bA_o+K'_{(1)}B_o+(3b^2-2)\,K'_{(-2)}C_o=1 \qquad \ldots(2.2a)$$

$$K'_{(1)}\,A_o+K'_{(2)}\,B_o+(3b-2)K'_{(-1)}\,C_o=K'_{(1)} \qquad \ldots(2.2b)$$

$$(3b^2-2)K'_{(-2)}\,A_o+(3b-2)\,K'_{(-1)}\,B_o+(15b^3-30b+16)\,K'_{(-4)}C_o$$
$$=(3b-2)\,K'_{(-2)} \quad \ldots(2.2c)$$

Thus, we obtain A_o, B_o and C_o and hence our m.m.s.e. estimator, say T'_o with its m.s.e., say M_o, simplifying to :

$$M_o=r\theta^2\,[A_o+3b\,C'_o], \text{ in view of (2.2}a\text{)},$$

where $$C'_o=K'_{(-2)}\,C_o.$$

Therefore, the relative Efficiency (RE) of T'_o as compared to $\bar{x}$, say E_2, is given by

$$E_2=(A_o+3B\,C'_o)^{-1}\times 100\% \qquad \ldots(2.3)$$

Apparently T'_o is more efficient then T^*_{LMMS} as the former is the optimum estimator in a bigger class which includes the class relevant to the latter.

3. Empirical Study

In this section, we tabulate the Relative Efficiencies E_1 and E_2 for various sample sizes n=10, 20 and 50 corresponding to some example-values of 'a' (the square of the coefficient of variance 'c') =0.1, 0.2, 0.5, 1,0, 2.0, 5.0 and 10.0. The small empirical study is intended to bring out the possible gain in the relative efficiency.

R.Es. (%) of the proposed estimator (E_2) and that of Gleser and Healy (E_1).

$\frac{a}{C^2}$	E_1	118.541	119.260	119.702
$=0.1$	E_2	118.982	111.524	119.983
$\frac{a}{C^2}$	E_1	137.082	138.520	139.403
$=0.2$	E_2	138.012	133.181	140.171
$\frac{a}{C^2}$	E_1	192.706	196.299	198.508
$=0.5$	E_2	195.878	198.545	200.303
$\frac{a}{C^2}$	E_1	285.412	292.598	297.015
$=1.0$	E_2	293.371	298.185	301.997
$\frac{a}{C^2}$	E_1	470.825	485.197	494.031
$=2.0$	E_2	489.515	498.269	504.859
$\frac{a}{C^2}$	E_1	1,027.062	1.062.991	1.085.077
$=5.0$	E_2	1,076.794	1,098.659	1,119.729
$\frac{a}{C^2}$	E_1	1,954.124	2,025.983	2,070.154
$=10.0$	E_2	2,045.778	2,095.559	2.141.528

The above table shows that for relatively larger samples, we may not have practically significant gain in relative efficiency unless C is rather big. For example, for $a=10$; (the coefficient of variation $C=3.16278$), even for sample as large as $N=50$, the gain is worth going for.

References

[1] Gleser, L.J. and Healy, J.D. (1976) : Estimating the Mean of a Normal Distribution with known coefficient of variation. *Jr. Amer. Stat. Assn.* **71**, 977-981.

[2] Khan, R.A. (1968) : A Note on Estimating the Mean of a Normal Distribution with known Coefficient of Variation *Jr. Amer. Stat. Assn.*, **63**, 1039-41.

[3] Searles, D.T. (1964) : The Utilization of a Known Coefficient of Variation in Estimation Procedures, *Jr. Amer. Stat. Assn.* **59**, 1225-1226.