EFFICIENT ESTIMATION OF COMMON PARAMETER OF TWO NORMAL POPULATIONS WITH KNOWN COEFFICIENTS OF VARIATION

BY

ASHOK SAHAI, GOVIND PRASAD AND SUNITA RANI

University of Roorkee, Roorkee-247672

(Received : August, 1980)

SUMMARY

Pandey and Singh [1] extended the minimum mean sqaure error estimation approach of Searles [2] to proposes analogous estimator of the common mean of two populations. The estimator has a lower mean square error (mse) through the use of known coefficient of variation, say, V_1 and V_2 of the two populations. This paper proposes a rather more efficient approach of utilising the knowledge, when the populations are normal, resulting in an estimator with mse lower than that of Pandey and Singh [1] estimator. The improvement has been illustrated by lisiting the relative efficiency of the proposed estimator with respect to the Pandey and Singh [1] estimator for some values of V_1, V_2 and n_1, n_3.

1. INTRODUCTION

Let us consider two normal populations with common mean, say, θ. Our investigations concern with the situations wherein the coefficients of variation are known for the two populations. Let $x_{11}, x_{12}, \ldots\ldots x_{1n_1}$ and $x_{21}, x_{22} \ldots\ldots x_{2n_2}$ be random samples of sizes n_1 and n_2 respectively, from the populations : $N(\theta, v_1^2\ \theta^2)$ and $N(\theta, v_2^2\ \theta^2)$. Further, let

$$\bar{X}_i = \sum_{j=1}^{n_j} (X_{ij}/n_i)$$

and

$$S_i^2 = (n_i - 1)^{-1} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2 , \; i = 1,\ 2$$

be the sample means and sample variances, respectively.

Pandey and Singh [1] like Searles [2] exploit the apriori information (in term of v_1 and v_2) to develope their estimator

$$Y^*=((n_1v_2^2)\,\bar{x}_1+(n_2v_1^2)\bar{x}_2)\,(n_1v_1^2+n_2v_1^2+v_1^2v_2^2)^{-1} \qquad ...(1.1)$$

of the common mean θ of the two populations. The mean square error (mse) of Y^* is found to be

$$MSE\ Y^*=(v_1^2v_2^2)(n_1v_2^2+n_2v_1^2+v_1^2v_2^2)^{-1}\,\theta^2=M^*,\ \text{say} \qquad ...(1.2)$$

It is worth noting that for the case under consideration two populations have only one parameter *i.e.* θ, the common mean. Thus the problem of estimating the standard deviations of the two populations ($V_1\theta$, and $V_2\theta$) is implicit in the estimation of θ. Essentially, V_1Y^* and V_2Y^* are the estimates of the standard deviation as per Pandey and Singh [1]'s approach. This simple observation motivates us to consider the sample standard deviations S_1 and S_2, alongwith the sample means, to evolve another estimator of the common parameter θ which apparently happens to have smaller mean square error (mse).

2. The New Estimator of The Common Parameter

Let us consider the class of estimators of the common parameter θ as the linear function :

$$Y_2=A\bar{X}_1+B\bar{X}_2+CS_1+DS_2$$

where A, B, C, and D are arbitrary scalars. We intend to determine A, B, C and D so as to minimise the mean square error (mse) of Y_2. The resultant estimator of θ in the class, say Y^{**}, is determined below. It will be the minimum mse (MMSE) estimator just as Pandey and Singh [1]'s Y^* is in the class of estimators Y_1. As the parent populations are normal, it is well known that

$$(n_i-1)\,S_i^2/(V_i^2\theta^2)\sim\chi^2\,(n_i-1);\ i=1,2$$

and $\bar{X}_i\sim N\,(\theta,\ r_i\,\theta^2)$, where $r_i=\dfrac{V_i^2}{n_i}$

$$b_i=1+r_i; \qquad i=1,2$$

Hence

$$E(s_i)=\left(\frac{2}{n_i-1}\right)\frac{\Gamma(n_i/2)}{\Gamma\left(\frac{n_i-1}{2}\right)}\,(V_i^2\ \theta^2)^{\frac{1}{2}}$$

$$=K_{(i)}^{(1)},\ \text{say}, \qquad ...(2.1)$$

and

$$E(s_i^2)=V_i^2\,\theta^2$$

$$=K_{(2)}^{(i)}, \text{ say,} \qquad i=1, 2 \qquad \ldots(2.2)$$

and $(V_i\theta)$ is the population standard deviation of the ith $(i=1, 2)$ population. Using (2.1) and (2.2) we get,

$$MSE\ (Y_2)=(A^2b_1+B^2+2AB+C^2\,K_{(2)}^{(1)}+D^2+K_{(2)}^{(2)}$$

$$+2CD\,K_{(1)}^{(1)}\,K_{(1)}^{(2)}+1-2C\,K_{(1)}^{(1)}-2D\,K_{(1)}^{(2)}$$

$$+2AC\,K_{(1)}^{(1)}+2BC\,K_{(1)}^{(1)}+2AD\,K_{(1)}^{(2)}$$

$$+2BD\,K_{(1)}^{(2)}-2A-2B)\,\theta^2 \qquad \ldots(2.3)$$

It is easy to verify from (2.3) that the values of A^*, B^*, C^* and D^* minimising $MES\ (Y_2)$ are obtained from the four normal equations as follows—

$$Ab_1+B+C\,K_{(1)}^{(1)}+D\,K_{(1)}^{(2)}=1$$

$$A+Bb_2+C\,K_{(1)}^{(1)}+D\,K_{(1)}^{(2)}=1$$

$$AK_{(1)}^{(1)}+BK_{(1)}^{(1)}+CK_{(2)}^{(1)}+DK_{(1)}^{(1)}\,K_{(1)}^{(2)}=K_{(1)}^{(1)}$$

$$AK_{(1)}^{(2)}+BK_{(1)}^{(2)}+CK_{(1)}^{(1)}\,K_{(1)}^{(2)}+DK_{(2)}^{(2)}=K_{(1)}^{(2)}$$

Hence

$$MSE\ (Y^{**})=-A^*-B^*-C^*K_{(1)}^{(1)}-D^*\,K_{(1)}^{(2)}+1$$

$$=M^{**}, \text{ say.}$$

3. ILLUSTRATION

To illustrate the improvement achieved through the proposed estimator over that of Pandey and Singh [1] the Relative Efficiency (in percent) of the former has been tabulated with respect to the latter for some values of n_1, n_2 ; v_1, v_2 as below:

Thus it may be concluded that the gain in efficiency is quite significant for larger values of v_1 and v_2. However, the gain comes down with an increase in sample size (s)n_1 and/or n_2 Nevertheless the gain is rather substantial and worth going for.

TABLE 3.1

$n_1=5$, $n_2=10$; REF (y^{**}, y^*) (in %)

V_2 \ V_1	*.1*	*.25*	*.5*	*1.0*	*2.0*
.1	101.725	103.919	104.788	105,069	105.144
.25	102.395	110.748	121.407	128.464	131.021
.5	102.536	114.311	142.463	183,555	210.222
1.0	102.575	115.604	156.309	261.889	404.732
2.0	102.584	115.964	161.307	311.446	645.309

TABLE 3.2

$n_1=5$, $n_2=15$; REF (y^{**}, y^*) (in %)

V_2 \ V_1	*.1*	*.25*	*.5*	1.0	2.0
.1	101.645	104.443	105.868	106.380	106.522
.25	102.083	110.256	123.348	134.292	138.844
.50	102.165	112.613	140.642	191.445	233.011
1.00	102.187	113.382	149.879	256.762	437.642
2.00	102.192	113.589	152.884	290.841	648.669

TABLE 3.3

$n_1=5$, $n_2=25$; REF (y^{**}, y^*) (in%)

V \ V_1	*.1*	*.25*	*.5*	*1.0*	*2.0*
.1	101.642	105.471	108.202	109.371	109.717
.25	101.907	110.256	127.233	156.501	156.494
.50	101.952	111.706	140.736	207.153	280.883
1.0	101.965	112.141	146.501	259,002	502.373
2.0	101.968	112.254	148.206	280.883	679.890

TABLE 3.4

$n_1=10$, $n_2=5$; REF (y^{**}, y^*) (in%)

V_2 \ V_1	.1	.25	.5	1.0	0.2
.1	101.726	102.396	102.537	102.575	102.584
.25	103.919	110.748	114.311	115.604	115.961
.5	104.784	121.407	142.463	156.630	161.307
1.0	105.069	128.468	183.555	261.889	311.446
2.0	105.144	131.021	210.222	404.732	645.309

TABLE 3.5

$n_1=10$, $n_1=25$; REF (y^{**}, y^*) (in%)

V_2 \ V_1	.1	.25	0.5	1.0	0.2
.1	101.723	104.309	105.483	105.884	105.993
.25	102,267	110.761	123.129	132.456	136.095
.5	102.375	113.689	142.814	191.467	227.763
1.0	102.402	114.688	154.486	267.689	449.959
2.0	102.410	114.961	158.326	311.818	719.159

TABLE 3.6

$n_1=15$, $n_2=20$; REF (y^{**}, y^*) in%)

V_2 \ V_1	.1	.25	.5	1.0	2.0
.1	101.929	103.546	104.271	104.439	104.484
.25	103.012	112.334	121.029	125.865	127.441
.5	103.277	117.718	147.875	183.337	202.278
1.0	103.350	116.199	170.316	287.509	421.444
2.0	103.368	116.649	164.330	372.741	792.341

TABLE 3.7

$n_1=15$, $n_2=25$; REF (y, y*) (in %)**

V_2 \ V_1	*.1*	*.25*	*.5*	*1.0*	*2.0*
.1	101.683	103.549	104.209	104.417	104.472
.25	102.457	110.513	119.761	125.333	127.254
.5	102.630	114.617	141.854	178.356	200.202
1.0	102.677	116.199	158.091	264.356	402.858
2.0	102.689	116.649	164.330	326.507	712.600

Acknowledgement

The authors are thankful to the referee for suggesting some significant improvement in the original draft leading to the proposed estimator.

References

[1] Pandey and Singh (1978) : 'A note on the Use of known Coefficients of Variation in the Estimation of Common Mean of Two Populations', *Jr. Ind. Assoc.* **16**, 141-144.

[2] Searles, D.T. (1964) : 'The Utilization of a known Coefficient of Variations in Estimation Procedure, *Jr. Amer. Stat. Assoc.*, **59**, 1225-26.