SOME FURTHER RESULTS IN SIMULTANEOUS CONFIDENCE INTERVAL ESTIMATION

by

S. N. Roy
Institute of Statistics
University of North Carolina

Institute of Statistics
Mimeograph Series No. 86
December, 1953

SOME FURTHER RESULTS IN SIMULTANEOUS CONFIDENCE INTERVAL ESTIMATION

by

S. N. Roy
Institute of Statistics, University of North Carolina

1. Summary. In this paper confidence bounds on the characteristic roots of Σ and of $\Sigma_1\Sigma_2^{-1}$ are given, where Σ stands for the dispersion matrix of one p-variate, and Σ_1 and Σ_2 for the dispersion matrices of two p-variate normal populations, and where the confidence coefficient is to be greater than or equal to a preassigned level. To this end certain confidence statements from a previous paper [1] serve as the starting point and then certain results in matrix algebra, taken over from another paper [2], and certain further results stated and proved here, are used to obtain the bounds of this paper.

2. Introduction.

2.1. Statement of the Problems. We take over from [1] the two confidence statements (5.1.5) and (5.2.4) and renumber them as

$$(2.1.1)\quad \underline{a}'\underline{a}\,\theta_{1\alpha}(p,n) \leq \underline{a}'(D_{1/\sqrt{\Theta}}\, n\Gamma' S \Gamma D_{1/\sqrt{\Theta}})\,\underline{a} \leq \underline{a}'\underline{a}\,\theta_{2\alpha}(p,n)$$

and

$$(2.1.2)\quad \frac{n_2}{n_1}\,\theta_{1\alpha}(p,n_1,n_2)\,\underline{b}'S_2\underline{b} \leq \underline{b}'(\mu D_{1/\sqrt{\Theta}}\,\mu^{-1} S_1 \mu'^{-1} D_{1/\sqrt{\Theta}}\,\mu')\,\underline{b}$$

$$\leq \frac{n_2}{n_1}\,\theta_{2\alpha}(p,n_1,n_2)\,\underline{b}'S_2\underline{b}\ .$$

The statements (2.1.1) and (2.1.2) are supposed to hold respectively for all non-null $\underline{a}$ (p x 1) and $\underline{b}$ (p x 1), and each with a confidence coefficient $1 - \alpha$. In (2.1.1) S stands for the sample dispersion matrix, n + 1 for the sample size, Θ's for the characteristic roots of Σ, Γ is an orthogonal matrix given by $\Sigma = \Gamma D_{\Theta} \Gamma'$ and $\theta_{1\alpha}(p,n)$ and $\theta_{2\alpha}(p,n)$ are subject to the only restriction

$$(2.1.3) \qquad P(\theta_{1\alpha} \leq \theta_1 \leq \theta_p \leq \theta_{2\alpha} \mid \Sigma) = 1 - \alpha ,$$

where θ_1 and θ_p are the smallest and the largest characteristic roots of nS. $\theta_{1\alpha}$ and $\theta_{2\alpha}$ are otherwise, for the moment, left flexible, unlike what was done in the previous paper [1].

In (2.1.2) S_1 and S_2 stand for the two sample dispersion matrices, n_1+1 and n_2+1 for the two sample sizes, Θ's for the characteristic roots of $\Sigma_1\Sigma_2^{-1}$ μ is a non-singular matrix given by $\Sigma_1 = \mu D_{\Theta} \mu'$ and $\Sigma_2 = \mu\mu'$ and $\theta_{1\alpha}(p,n_1,n_2)$ and $\theta_{2\alpha}(p, n_1,n_2)$ are subject to the only restriction

$$(2.1.4) \qquad P(\theta_{1\alpha} \leq \theta_1 \leq \theta_p \leq \theta_{2\alpha} \mid \Sigma_1 = \Sigma_2) = 1 - \alpha ,$$

where θ_1 and θ_p are the smallest and the largest characteristic roots of $(n_1/n_2)\, S_1S_2^{-1}$. $\theta_{1\alpha}$ and $\theta_{2\alpha}$ are otherwise, for the moment, left free, unlike the development of the previous paper [1].

Let us denote by c(M) any characteristic root of the matrix M. Then it is well known that the statements (2.1.1) and (2.1.2) are respectively equivalent to

$$(2.1.5) \quad \frac{1}{n}\,\theta_{1\alpha}(p,n) \leq \text{all } c(D_{1/\sqrt{\Theta}}\,\Gamma' S \Gamma D_{1/\sqrt{\Theta}}) \leq \frac{1}{n}\,\theta_{2\alpha}(p,n) \text{ and}$$

$$(2.1.6) \quad \frac{n_2}{n_1}\,\theta_{1\alpha}(p,n_1,n_2) \leq \text{all } c(\mu D_{1/\sqrt{\Theta}}\,\mu^{-1} S_1 \mu'^{-1} D_{1/\sqrt{\Theta}}\,\mu' S_2^{-1})$$

$$\leq \frac{n_2}{n_1}\,\theta_{2\alpha}(p,n_1,n_2) \quad .$$

Notice that $\Theta_i = c_i(\Sigma)$ in (2.1.5) and $= c_i(\Sigma_1\Sigma_2^{-1})$ in (2.1.6) $(i=1,\ldots,p)$. It is now our purpose to try to obtain confidence bounds on Θ_i's (or their functions) in terms of $c_i(S)$'s (or their functions) in the case of (2.1.5) and of $c_i(S_1), c_i(S_2)$ (or their functions) in the case of (2.1.6). For $c_i(\Sigma)$'s the confidence bounds are given by (3.1.3) and (3.1.6) and for $c_i(\Sigma_1\Sigma_2^{-1})$ by (3.2.8). To derive these we need the following results in matrix algebra.

2.2. <u>Some auxiliary matrix results.</u> Let us denote by A' the transpose of A, and shorten positive definite into p.d. and positive semi-definite into p.s.d. Also let $c_{min}(M)$ and $c_{max}(M)$ denote the smallest and the largest characteristic root of a p.d. matrix M and, if any matrix B is p x p, let $tr_s(B)$ $(s=1, \ldots, p)$ stand for the sum of all s-th order principal minors of B. It is well known that

$$(2.2.1) \qquad tr_s(B) \quad \sum_{i_1 \neq i_2 \neq \ldots \neq i_s = 1}^{p} c_{i_1}(B) c_{i_2}(B) \ldots c_{i_s}(B) \quad ,$$

and, in particular, that

$$tr_1(B) = \sum_{i=1}^{p} c_i(B) = \sum_{i=1}^{p} b_{ii} \text{ and } tr_p(B) = \prod_{i=1}^{p} c_i(B) = |B| ,$$

$$(2.2.2) \qquad c\,[A(p \times p)B(p \times p)] = c\,[B(p \times p)A(p \times p)] \text{ and}$$

(2.2.3) the product of two p.d. matrices is p.d. and if $A(p \times q)$ [rank $r \leq \min(p,q)$] is a matrix with real elements, then AA' is p.s.d. of rank r.

We take over from [2] the following:

$$(2.2.4) \qquad c_{min}(A)c_{min}(B) \leq \text{all } c(AB) \leq c_{max}(A)c_{max}(B) \quad ,$$

where A and B are two symmetric matrices of which one is p.d. and the other at least p.s.d. The generalization to the product of a finite number of matrices is obvious and is also given in [1].

We also take over from [2], the following result:

$$(2.2.4.1) \qquad c_{min}(MM') \leq c^2(M) \leq c_{max}(MM') \quad ,$$

where M is a square matrix with real characteristic roots. From (2.2.4) it is easy to see, by replacing A by AB^{-1} (if B is non-singular), that

$$(2.2.5) \qquad c_{min}(AB^{-1})c_{min}(B) \leq \text{all } c(A) \leq c_{max}(AB^{-1})c_{max}(B) \quad .$$

Next, we establish that

(2.2.6) if (a) $d_1 \leq \text{all } c(AB^{-1}) \leq d_2$, then

$$\text{(b)}(d_1)^t tr_t(A) \leq tr_t(A) \leq (d_2)^t tr_t(B) \quad (t = 1, \ldots, p) \quad ,$$

where A and B are two p x p p.d. matrices and d_1 and d_2 any two positive numbers such that $d_1 \leq d_2$. Notice that (b) is a necessary (though not a sufficient) condition for (a).

Proof. It is easy to check that [$d_1 < \text{all } c(AB^{-1})$]

$$\Longrightarrow A-d_1 B \text{ is p.d.} \Longrightarrow A_t - d_1 B_t (t=1, \ldots, p) \text{ is p.d.}$$

(where $A_t - d_1 B_t$ is a submatrix formed by the intersection of any t rows of $A-d_1B$ with t columns bearing the same numbers) $\Longrightarrow$ $d_1 < \text{all } c(A_t B_t^{-1})$. Now, if all $c(A_t B_t^{-1}) > d_1$, one consequence is that

$$(2.2.7) \qquad \prod_{i=1}^{t} c_i(A_t B_t^{-1}) > (d_1)^t, \text{ i.e., } |A_t| / |B_t| > (d_1)^t \quad ,$$

$$\text{i. e., } |A_t| > (d_1)^t \, |B_t| \quad .$$

For a given t, summing over different possible submatrices we have

(2.2.8) $$\mathrm{tr}_t A > (d_1)^t \, \mathrm{tr}_t B \quad .$$

Using the same kind of argument for the other half of the inequality and remembering that $t = 1, 2, \ldots, p$, and combining, we have the result that

(2.2.9) if $d_1 <$ all $c(AB^{-1}) < d_2$, then

$$(d_1)^t \mathrm{tr}_t(B) < \mathrm{tr}_t(A) < (d_2)^t \mathrm{tr}_t(B) \ (t = 1, \ldots, p).$$

By a slight rephrasing (which is obviously permissible here) we have from (2.2.9) the result (2.2.6).

3. Confidence bounds on $c(\Sigma)$'s and $c(\Sigma_1\Sigma_2^{-1})$'s.

It is well known that "If E_1, then E_2"

$\Longrightarrow$ "E_2 is a necessary condition for E_1" $\Longrightarrow$ "$E_1 \subset E_2$",

$\longrightarrow$ $P(E_1) \leq P(E_2)$, the last one being a necessary (though not a sufficient) condition for the other statements. This will be used in the derivation of the confidence bounds.

3.1. Bounds on $c(\Sigma)$'s. Starting from (2.1.5) and noting that

(3.1.1) $$c(D_{1/\sqrt{\Theta}}\Gamma' S\Gamma D_{/\sqrt{\Theta}}) = c(S\Gamma D_{1/\Theta}\Gamma') = c(S\Sigma^{-1}) \quad ,$$

we have, with a confidence coefficient $1-\alpha$, the confidence bounds:

(3.1.2) $$\frac{1}{n}\,\theta_{1\alpha}(p,n) \leq \text{all } c\ (S\Sigma^{-1}) \leq \frac{1}{n}\,\theta_{2\alpha}(p,n), \quad \text{or}$$

$$n\ \theta_{1\alpha}^{-1}\ (p,n) \geq \text{all } c\ (\Sigma S^{-1}) \geq n\ \theta_{2\alpha}^{-1}\ (p,n).$$

From (2.2.5) we observe that (3.1.2) ——> the following:

(3.1.3) $$n\theta_{1\alpha}^{-1}(p,n)\ c_{max}(S) \geq \text{all } c(\Sigma) \geq n\theta_{2\alpha}^{-1}(p,n)\ c_{min}(S),$$

which is thus also another set of simultaneous confidence bounds with a confidence coefficient $\geq 1-\alpha$.

From (2.2.6) we also observe that (3.1.2) ——> the following:

(3.1.4) $$[n\theta_{1\alpha}^{-1}(p,n)]^t\ tr_t(S) \geq tr_t(\Sigma) \geq [n\theta_{2\alpha}^{-1}(p,n)]^t\ tr_t(S),\ (t=1,2,\ldots,p),$$

which is thus a set of simultaneous confidence bounds with a confidence coefficient $\geq 1-\alpha$. Notice that, using (2.2.1), $tr_t(S)$ and $tr_t(\Sigma)$ are easily calculated in terms of θ_i's and Θ_i's.

3.2. Bounds on $c(\Sigma_1\Sigma_2^{-1})$'s. Starting from (2.16) we have, with a confidence coefficient $1-\alpha$, the confidence bounds:

$$(3.2.1)\quad \frac{n_1}{n_2}\,\theta_{1\alpha}^{-1}(p,n_1,n_2) \geq \text{all } c(S_2(\mu')^{-1}D_{\sqrt{\Theta}}\mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\mu^{-1}) \geq \frac{n_1}{n_2}\theta_{2\alpha}^{-1}(p,n_1,n_2).$$

Using (2.2.2) and (2.2.5) we have

$$(3.2.2)\quad c_{max}(S_2(\mu')^{-1}D_{\sqrt{\Theta}}\mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\mu^{-1})c_{max}(S_2^{-1}) \geq$$

$$\text{all } c((\mu')^{-1}D_{\sqrt{\Theta}}\ \mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\ \mu^{-1}),\ \text{i.e., all } c(S_1^{-1}\ \Delta) \geq$$

$$c_{min}(S_2(\mu')^{-1}\ D_{\sqrt{\Theta}}\ \mu' S_1^{-1}\ \mu D_{\sqrt{\Theta}}\ \mu^{-1})\ c_{min}(S_1^{-1}),$$

where

$$(3.2.3)\quad \Delta = (\mu D_{\sqrt{\Theta}}\mu^{-1})((\mu')^{-1}D_{\sqrt{\Theta}}\mu') = (\mu D_{\sqrt{\Theta}}\mu^{-1})(\mu D_{\sqrt{\Theta}}\mu^{-1})'.$$

In the same way we have

$$(3.2.4)\quad c_{max}(S_1^{-1}\Delta)\ c_{max}(S_1) \geq \text{all } c(\Delta) \geq c_{min}(S_1^{-1}\Delta)\ c_{min}(S_1).$$

Furthermore noting that

$$(3.2.5)\quad c(\mu D_{\sqrt{\Theta}}\,\mu^{-1}) = c(D_{\sqrt{\Theta}}) = \sqrt{\Theta} = c((\mu')^{-1}\, D_{\sqrt{\Theta}}\,\mu'),$$

and using (2.2.4), we have

$$(3.2.6)\quad c_{max}(\Delta) \geq \text{all } c^2(\mu D_{\sqrt{\Theta}}\,\mu^{-1}),\ \text{i.e., all } c^2(D_{\sqrt{\Theta}}),\ \text{i.e., all } \Theta_i\text{'s} \geq c_{min}(\Delta)$$

Combining (3.2.2), (3.2.4) and (3.2.6) we have

$$(3.2.7)\quad c_{max}(S_2(\mu')^{-1}D_{\sqrt{\Theta}}\,\mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\,\mu^{-1})\ c_{max}(S_2^{-1})\ c_{max}(S_1) \geq$$

$$\text{all } \Theta_i\text{'s} \geq c_{min}(S_2(\mu')^{-1}D_{\sqrt{\Theta}}\,\mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\,\mu^{-1})\ c_{min}(S_2^{-1})\ c_{min}(S_1).$$

From this it is easy to check that (3.2.1) $\longrightarrow$ the following:

$$(3.2.8)\quad \frac{n_1}{n_2}\,\theta_{1\alpha}^{-1}(p,n_1,n_2)\ c_{max}(S_2^{-1})\ c_{max}(S_1) \geq \text{all } c(\Sigma_1\Sigma_2^{-1}) \geq$$

$$\frac{n_1}{n_2}\,\theta_{2\alpha}^{-1}(p,n_1,n_2)\ c_{min}(S_2^{-1})\ c_{min}(S_1),$$

which is thus a set of simultaneous confidence bounds with a confidence coefficient $\geq 1-\alpha$. Notice that

$$c_{max}(S_2^{-1}) = 1/c_{min}(S_2) \text{ and } c_{min}(S_2^{-1}) = 1/c_{max}(S_2).$$

Confidence bounds in terms of tr_t could also be given as in (3.1.4), but in this case the bounds would be more complicated and do not appear to be so worthwhile as in the previous case.

3.3. Determination of the constants $(\theta_{1\alpha}(p,n), \theta_{2\alpha}(p,n))$ and $(\theta_{\alpha}(p,n_1,n_2), \theta_{2\alpha}(p,n_1,n_2))$ occurring in the confidence bounds.

It has been stated in section 2 that the pair $\theta_{1\alpha}(p,n)$, $\theta_{2\alpha}(p,n)$ for the first problem and the pair $\theta_{1\alpha}(p,n_1,n_2)$, $\theta_{2\alpha}(p,n_1,n_2)$ for the second problem satisfy respectively the conditions (2.1.3) and (2.1.4), but are otherwise free. It is well known how the shortness (in the sense of probability) of a confidence interval (or intervals) ties in with the power of the associated test. Let us consider the associated tests, or rather, the acceptance regions of the respective hypotheses (i) $H(\Sigma = \Sigma_0)$, (ii) $H(\Sigma_1 = \Sigma_2)$. They are

(3.3.1) $H(\Sigma = \Sigma_0)$: $\theta_{1\alpha}(p,n) \leq \theta_1 \leq \theta_p \leq \theta_{2\alpha}(p,n)$ and

(3.3.2) $H(\Sigma_1 = \Sigma_2)$: $\theta_{1\alpha}(p,n_1,n_2) \leq \theta_1 \leq \theta_p \leq \theta_{2\alpha}(p,n_1,n_2)$.

In the first case it is possible to (so) choose $\theta_{1\alpha}$ and $\theta_{2\alpha}$ (and this choice will be unique) as to let the second kind of error (which, aside from p, n and α, depends only on the characteristic roots of $\Sigma \Sigma_0^{-1}$) have a (local) minimum, i.e., the power a local maximum at $\Sigma = \Sigma_0$($\Sigma \neq \Sigma_0$ is supposed to be the alternative). It so happens in this case that the resulting power function then monotonically increases as each $c_i(\Sigma \Sigma_0^{-1})$ tends away from unity, provided that all are ≥ 1 or ≤ 1, to begin with. We have an exactly similar situation in the second case, $H(\Sigma = \Sigma_0)$ being replaced by $H(\Sigma_1 = \Sigma_2)$ and $\Sigma \Sigma_0^{-1}$ being replaced by $\Sigma_1 \Sigma_2^{-1}$. The impact of this on the shortness, in the probability sense, of the resulting confidence bounds is obvious and need not be discussed in detail. The results just stated are proved in another paper to be shortly submitted to the <u>Annals of Mathematical Statistics</u>. It may be noticed, however, that for any pair ($\theta_{1\alpha}$, $\theta_{2\alpha}$) subject only to (2.1.3) or (2.1.4), we are going to get anyway the confidence bounds of subsections 3.1 and 3.2, with confidence coefficients $\geq 1-\alpha$, the only difference being that they would not have the property of "shortness" possessed by those that are based on ($\theta_{1\alpha}$, $\theta_{2\alpha}$) determined in the above way.

4. Concluding remarks. In a later paper this technique will be used to obtain the confidence bounds on "canonical regressions" discussed in section 6 of [1] and on certain other types of parameters in multivariate analysis.

REFERENCES

1. S. N. Roy and R. C. Bose, "Simultaneous confidence interval estimation", Annals of Mathematical Statistics, Vol. 24 (1953), pp.

2. S. N. Roy, "A useful theorem in matrix algebra", mimeographed paper.