SOME FURTHER RESULTS IN SIMULTANEOUS CONFIDENCE INTERVAL ESTIMATION

By S. N. Roy

University of North Carolina

1. Summary. This paper is a follow-up of a previous paper [1], the full implications of some of the results there being brought out here in terms that are physically more meaningful. Two cases of simultaneous confidence bounds, I and II, are given, in each case with a confidence coefficient which is to be greater than or equal to a preassigned level.

Case I relates to the characteristic roots of Σ and $\Sigma_1\Sigma_2^{-1}$, 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. Case II relates to a $(p + q)$-variate normal population $(p \leqq q)$, for which the matrix of regression of the p-set on the q-set is defined in a natural manner. This matrix is denoted by $\beta(p \times q)$ and simultaneous confidence bounds are given on all bilinear compounds of this matrix (with arbitrary coefficient vectors of unit modulus).

Confidence bounds on the characteristic roots of Σ and $\Sigma_1\Sigma_2^{-1}$ are given respectively by (3.1.3) and (3.2.8). Confidence bounds on the bilinear compounds of the regression matrix β are given by (4.7).

2. Introduction. Let us denote by A' the transpose of a matrix A, and shorten positive definite into p.d. and positive semidefinite 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. A $p \times p$ diagonal matrix whose diagonal elements are, say, c_1, $c_2, \cdots, c_p$ will be denoted by $D_c(p \times p)$ or simply by D_c. A $p \times p$ unit matrix will be denoted by $I(p)$.

2.1. *Statement and reduction of the problem for the case of Σ and $\Sigma_1\Sigma_2^{-1}$.* We take over from the previous paper [1] the two confidence statements (5.1.5) and (5.2.4) and renumber them as

$$\underset{\sim}{a}'\underset{\sim}{a}\theta_{1\alpha}(p, n) \leqq \underset{\sim}{a}'(D_{1/\sqrt{\theta}}n\Gamma'S\Gamma D_{1/\sqrt{\theta}})\underset{\sim}{a} \leqq \underset{\sim}{a}'\underset{\sim}{a}\theta_{2\alpha}(p, n), \tag{2.1.1}$$

$$\begin{aligned}(n_2/n_1)\theta_{1\alpha}(p, n_1, n_2)\underset{\sim}{b}'S_2\underset{\sim}{b} &\leqq \underset{\sim}{b}'(\mu D_{1/\sqrt{\theta}}\mu^{-1}S_1\mu'^{-1}D_{1/\sqrt{\theta}}\mu')\underset{\sim}{b}\\ &\leqq (n_2/n_1)\theta_{2\alpha}(p, n_1, n_2)\underset{\sim}{b}'S_2\underset{\sim}{b}.\end{aligned} \tag{2.1.2}$$

These statements are supposed to hold respectively for all nonnull $\underset{\sim}{a}(p \times 1)$ and $\underset{\sim}{b}(p \times 1)$, 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, and the Θ's for the characteristic roots of Σ. Here Γ is an orthogonal matrix

Received 1/8/54, revised 6/29/54.

given by $\Sigma = \Gamma D_\Theta \Gamma'$, and $\theta_{1\alpha}(p, n)$ and $\theta_{2\alpha}(p, n)$ are subject only to the restriction

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

where θ_1 and θ_p are the smallest and the largest characteristic roots of nS. Otherwise $\theta_{1\alpha}$ and $\theta_{2\alpha}$ are, 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, and the Θ's for the characteristic roots of $\Sigma_1\Sigma_2^{-1}$. Here μ is a nonsingular matrix given by $\Sigma_1 = \mu D_\Theta \mu'$ and $\Sigma_2 = \mu\mu'$, while $\theta_{1\alpha}(p, n_1, n_2)$ and $\theta_{2\alpha}(p, n_1, n_2)$ are subject only to the restriction

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

where θ_1 and θ_p are the smallest and the largest characteristic roots of $(n_1/n_2)S_1S_2^{-1}$. Otherwise $\theta_{1\alpha}$ and $\theta_{2\alpha}$ are, 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

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

$$(2.1.6) \qquad (n_2/n_1)\theta_{1\alpha}(p, n_1, n_2) \leqq \text{all } c(\mu D_{1/\sqrt{\Theta}}\mu^{-1}S_1\mu'^{-1}D_{1/\sqrt{\Theta}}\mu' S_2^{-1}) \leqq (n_2/n_1)\theta_{2\alpha}(p, n_1, n_2).$$

We notice that $\Theta_i = c_i(\Sigma)$ in (2.1.5) and $= c_i(\Sigma_1\Sigma_2^{-1})$ in (2.1.6), with $i = 1, \cdots, p$. It is now our purpose 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 in terms of $c_i(S_1)$ and $c_1(S_2)$ (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 (3.1.4), and for $c_i(\Sigma_1\Sigma_2^{-1})$ by (3.2.8).

2.2. *Statement and reduction of the problem for the case of the regression matrix* β. We recall the confidence statement ([1], (6.1.4)), with a confidence coefficient $1 - \alpha$:

$$(2.2.1) \qquad b - \frac{t_\alpha(n-2)}{\sqrt{n-2}}\sqrt{1-r^2}\,\frac{s_1}{s_2} \leqq \beta \leqq b + \frac{t_\alpha(n-2)}{\sqrt{n-2}}\sqrt{1-r^2}\,\frac{s_1}{s_2},$$

where β (which is now a scalar) stands for the population regression of x_1 on x_2 (where x_1 and x_2 have a bivariate normal distribution), b for the sample regression (in a random sample of size $n \geqq 3$), r for the sample correlation, s_1 and s_2 for the two sample standard deviations, and t_α for the upper $\frac{1}{2}\alpha$-point of the t-distribution with D.F. $(n - 2)$.

We also note that

$$(2.2.2) \qquad b = rs_1/s_2 = rs_1s_2/s_2^2, \qquad \beta = \rho\sigma_1\sigma_2/\sigma_2^2,$$

where ρ, σ_1, and σ_2 stand respectively for the population correlation coefficient and the two standard deviations.

We now start ([1], Sec. 6.2) with a random sample of size n, with $n > p + q$ and $p \leqq q$, from a $(p + q)$-variate normal population. Next we reduce for the means and set

$$(n - 1)\begin{pmatrix} S_{11} & S_{12} \\ S'_{12} & S_{22} \end{pmatrix} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} (Y'_1 \quad Y'_2),$$

where S_{11}, S_{22}, and S_{12} stand respectively for the sample dispersion submatrix of the p-set, that of the q-set, and that between the p-set and the q-set. Here Y_1 and Y_2 have p.d.f. proportional to

$$\exp\left[-\frac{1}{2}\operatorname{tr}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma'_{12} & \Sigma_{22} \end{pmatrix}^{-1}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}(Y'_1 \; Y'_2)\right]. \tag{2.2.3}$$

We next recall ([1], Sec. 6.2) that there exist nonsingular $\mu_1(p \times p)$ and $\mu_2(q \times q)$ such that

$$\begin{aligned} \Sigma_{11}(p \times p) &= \mu_1(p \times p)\mu'_1(p \times p), \qquad \Sigma_{22}(q \times q) = \mu_2(q \times q)\mu'_2(q \times q), \\ \Sigma_{12}(p \times q) &= \mu_1(p \times p)[D_{\sqrt{\theta}} \quad 0]\mu'_2(q \times q), \end{aligned} \tag{2.2.4}$$

where $D_{\sqrt{\theta}}$ is a $p \times p$ matrix and the θ's are the characteristic roots (all non-negative) of the matrix $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{22}$ (i.e., the squares of the population canonical correlations between the p-set and the q-set). As in ([1], Sec. 6.2), we have

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma'_{12} & \Sigma_{22} \end{pmatrix}^{-1} = \left[\begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}\begin{pmatrix} I(p) & \vdots & (D_{\sqrt{\theta}}\; 0) \\ \cdots & & \cdots \\ \begin{pmatrix} D_{\sqrt{\theta}} \\ 0 \end{pmatrix} & \vdots & I(q) \end{pmatrix}\begin{pmatrix} \mu'_1 & 0 \\ 0 & \mu'_2 \end{pmatrix}\right]^{-1}$$

$$= \left[\begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}\begin{pmatrix} D_{\sqrt{1-\theta}} & \vdots & (D_{\sqrt{\theta}}\; 0) \\ \cdots & & \cdots \\ 0 & \vdots & I(q) \end{pmatrix}\begin{pmatrix} D_{\sqrt{1-\theta}} & \vdots & 0 \\ \cdots & & \cdots \\ \begin{pmatrix} D_{\sqrt{\theta}} \\ 0 \end{pmatrix} & \vdots & I(q) \end{pmatrix}\begin{pmatrix} \mu'_1 & 0 \\ 0 & \mu'_2 \end{pmatrix}\right]^{-1} \tag{2.2.5}$$

$$= \begin{pmatrix} \mu_1'^{-1} & 0 \\ 0 & \mu_2'^{-1} \end{pmatrix}\begin{pmatrix} D_{\sqrt{1/(1-\theta)}} & \vdots & 0 \\ \cdots & & \cdots \\ -\begin{pmatrix} D_{\sqrt{\theta/(1-\theta)}} \\ 0 \end{pmatrix} & \vdots & I(q) \end{pmatrix}\begin{pmatrix} D_{\sqrt{1/(1-\theta)}} & \vdots & -(D_{\sqrt{\theta/(1-\theta)}}\; 0) \\ \cdots & & \cdots \\ 0 & \vdots & I(q) \end{pmatrix}$$

$$\times\begin{pmatrix} \mu_1^{-1} & 0 \\ 0 & \mu_2^{-1} \end{pmatrix}.$$

Going back to (2.2.3) and using the result that tr (AB) = tr (BA), we have

$$
\begin{aligned}
&\operatorname{tr}\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{12}' & \Sigma_{22}\end{pmatrix}^{-1}\begin{pmatrix}Y_1\\ Y_2\end{pmatrix}(Y_1'\ Y_2')\\
&= \operatorname{tr}\begin{pmatrix}D_{\sqrt{1/(1-\theta)}} & \vdots & -(D_{\sqrt{\theta/(1-\theta)}}\ \ 0)\\ \cdots & \vdots & \cdots\\ 0 & \vdots & I(q)\end{pmatrix}\begin{pmatrix}\mu_1^{-1} & 0\\ 0 & \mu_2^{-1}\end{pmatrix}\begin{pmatrix}Y_1\\ Y_2\end{pmatrix}(Y_1'\ Y_2') \qquad (2.2.6)\\
&\qquad \times\begin{pmatrix}\mu_1'^{-1} & 0\\ 0 & \mu_2'^{-1}\end{pmatrix}\begin{pmatrix}D_{\sqrt{1/(1-\theta)}} & \vdots & 0\\ \cdots & \vdots & \cdots\\ -\begin{pmatrix}D_{\sqrt{\theta/(1-\theta)}}\\ 0\end{pmatrix} & \vdots & I(q)\end{pmatrix} = \operatorname{tr}\begin{pmatrix}Z_1\\ Z_2\end{pmatrix}(Z_1'\ Z_2'),
\end{aligned}
$$

where

$$(2.2.7)\qquad Z_1 = \sqrt{1/(1-\theta)}\,\mu_1^{-1}Y_1 - (D_{\sqrt{\theta/(1-\theta)}}\ \ 0)\mu_2^{-1}Y_2, \qquad Z_2 = \mu_2^{-1}Y_2.$$

Thus it is easy to check from (2.2.3), (2.2.6), and (2.2.7) that (Z_1, Z_2) have a p.d.f. proportional to

$$(2.2.8)\qquad \exp\left\{-\tfrac{1}{2}\operatorname{tr}\begin{pmatrix}Z_1\\ Z_2\end{pmatrix}(Z_1'\ Z_2')\right\}.$$

Consider now, for any two arbitrary nonnull vectors $\underline{a}_1(p \times 1)$ and $\underline{a}_2(q \times 1)$ and for a fixed positive θ_0, the statement

$$(2.2.9)\qquad \frac{(\underline{a}_1'Z_1Z_2'\underline{a}_2)^2}{(\underline{a}_1'Z_1Z_1'\underline{a}_1)(\underline{a}_2'Z_2Z_2'\underline{a}_2)} \leqq \theta_0.$$

This can be written in terms of Y_1 and Y_2 as

$$(2.2.10)\qquad \frac{[\underline{a}_1'\{D_{\sqrt{1/(1-\theta)}}\,\mu_1^{-1}Y_1Y_2'\mu_2'^{-1} - (D_{\sqrt{\theta/(1-\theta)}}\ \ 0)\mu_2^{-1}Y_2Y_2'\mu_2'^{-1}\}\underline{a}_2]^2}{(\underline{a}_2'\mu_2^{-1}Y_2Y_2'\mu_2'^{-1}\underline{a}_2)(\underline{a}_1'QQ'\underline{a}_1)} \leqq \theta_0,$$

where

$$(2.2.11)\qquad Q = D_{\sqrt{1/(1-\theta)}}\,\mu_1^{-1}Y_1 - (D_{\sqrt{\theta/(1-\theta)}}\ \ 0)\,\mu_2^{-1}Y_2.$$

Now putting

$$(2.2.12)\qquad \underline{b}_1'(1 \times p) = \underline{a}_1'D_{\sqrt{1/(1-\theta)}}\mu_1^{-1}, \qquad \underline{b}_2'(1 \times q) = \underline{a}_2'\mu_2^{-1}$$

and using (2.2.4), we check that (2.2.10) reduces to

$$\frac{[\underline{b}_1'(Y_1Y_2' - \beta Y_2Y_2')\underline{b}_2]^2}{[\underline{b}_2'Y_2Y_2'\underline{b}_2][\underline{b}_1'(Y_1 - \beta Y_2)(Y_1' - Y_2'\beta')\underline{b}_1]} \leqq \theta_0$$

or

$$(2.2.13)\qquad \frac{[\underline{b}_1'(S_{12} - \beta S_{22})\underline{b}_2]^2}{(\underline{b}_2'S_{22}\underline{b}_2)[\underline{b}_1'(S_{11} - S_{12}\beta' - \beta S_{12}' + \beta S_{22}\beta')\underline{b}_1]} \leqq \theta_0,$$

where

$$\beta(p \times q) = \mu_1(D_{\sqrt{\theta}} \quad 0)\mu_2^{-1} = \Sigma_{12}\Sigma_{22}^{-1}. \tag{2.2.14}$$

As defined by (2.2.14), β can be appropriately called the matrix of population regression of the p-set on the q-set. It is the only set of population parameters that occurs in statement (2.2.13).

For an $p \times p$ matrix B, let $\mathrm{tr}_s(B)$, for $s = 1, \cdots, p$, stand for the sum of all sth order principal minors of B. It is well known that

$$\mathrm{tr}_s(B) = \sum_{i_1 \neq i_2 \neq \cdots \neq i_s = 1}^{p} c_{i_1}(B)c_{i_2}(B)\cdots c_{i_s}(B), \tag{2.3.1}$$

and, in particular, that $\mathrm{tr}_1(B) = \sum c_i(B) = \sum b_{ii}$, and $\mathrm{tr}_p(B) = \prod c_i(B) = B$. Also well known is that

$$c[A(p \times p)B(p \times p)] = c[B(p \times p)A(p \times p)] \tag{2.3.2}$$

Furthermore we recall

LEMMA A. *The product of two* p.d. *matrices is* p.d. *If* $A(p \times q)$ [*rank* $r \leqq \min(p, q)$] *is a matrix with real elements, then* AA' *is* p.s.d. *of rank* r.

We have also [2] that

$$c_{\min}(A)\, c_{\min}(B) \leqq \text{all } c(AB) \leqq c_{\max}(A)\, c_{\max}(B), \tag{2.3.3}$$

where A and B are two symmetric matrices of which one is p.d. and the other is at least p.s.d. The generalization to the product of a finite number of matrices is obvious [2]. We also take over from [2] the result that

$$c_{\min}(MM') \leqq c^2(M) \leqq c_{\max}(MM'), \tag{2.3.4}$$

where M is a square matrix with real characteristic roots. From (2.3.4) it is easy to see, by replacing A by AB^{-1} (if B is nonsingular), that

$$c_{\min}(AB^{-1})\, c_{\min}(B) \leqq \text{all } c(A) \leqq c_{\max}(AB^{-1})\, c_{\max}(B). \tag{2.3.5}$$

Next, we establish

LEMMA B. *If* $d_1 \leqq$ all $c(AB^{-1}) \leqq d_2$, *then*

$$(d_1)^t\, \mathrm{tr}_t(B) \leqq \mathrm{tr}_t(A) \leqq (d_2)^t\, \mathrm{tr}_t(B), \qquad t = 1, \cdots, p,$$

where A *and* B *are two* $p \times p$ *matrices and* d_1 *and* d_2 *any two positive numbers such that* $d_1 \leqq d_2$.

The conclusion is a necessary (though not a sufficient) condition for the hypothesis.

PROOF. It is easy to check that the statement $d_1 <$ all $c(AB^{-1})$ is equivalent to the statement "A-d_1B is p.d." which again is equivalent to the statement "$A_t - d_1B_t$, for $t = 1, 2, \cdots, p$, is p.d.", where $A_t - d_1B_t$ is a submatrix formed by the intersection of any t rows of $A - d_1B$ with t columns bearing the same numbers. The last statement again is equivalent to the statement $d_1 <$ all $c(A_tB_t^{-1})$.

Now, if all $c(A_tB_t^{-1}) > d_1$, one consequence is that

$$(2.3.6)\quad \prod_{i=1}^{t} c_i(A_tB_t^{-1}) > (d_1)^t, \quad \text{that is,} \quad \frac{|A_t|}{|B_t|} > (d_1)^t, \quad \text{that is,} \quad |A_t| > (d_1)^t\,|B_t|.$$

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

$$(2.3.7)\qquad \mathrm{tr}_t A > (d_1)^t\, \mathrm{tr}_t B.$$

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

$$(2.3.8)\quad \text{if} \quad d_1 < \text{all } c(AB^{-1}) < d_2, \quad \text{then} \quad (d_1)^t\, \mathrm{tr}_t(B) < \mathrm{tr}_t(A) < (d_2)^t\, \mathrm{tr}_t(B), \quad t = 1, \cdots, p.$$

By a slight rephrasing (which is obviously permissible here) we can obtain Lemma B) from (2.3.8). We recall the three following well known lemmas, repeatedly used in [2].

LEMMA C. *The statement "$g_1 \leqq$ all $c(M) \leqq g_2$ (for a $p \times p$ real matrix M with real roots)" is equivalent to the statement "$g_1 \leqq \underset{\sim}{d}'(1 \times p)M(p \times p)\underset{\sim}{d}(p \times 1) \leqq g_2$ (for all arbitrary vectors $\underset{\sim}{d}$ of unit modulus)".*

LEMMA D. *The statement "$g_1 \leqq$ all $c(M_1M_2^{-1}) \leqq g_2$ (for two $p \times p$ real matrices* M_1 and M_2 *with real roots, M_2 being nonsingular" is equivalent to the statement "$g_1 \leqq \underset{\sim}{d}'(1 \times p)M_1(p \times p)\underset{\sim}{d}(p \times 1) \,/\, \underset{\sim}{d}'(1 \times p)M_2(p \times p)\underset{\sim}{d}(p \times 1) \leqq g_2$ (for all arbitrary nonnull vectors $\underset{\sim}{d}$)".*

LEMMA E. *The statement "$\underset{\sim}{x}'(1 \times q)\underset{\sim}{x}(q \times 1) \leqq h\ (> 0)$" is equivalent to the statement "$|\underset{\sim}{x}'(1 \times q)\underset{\sim}{d}(q \times 1)| \leqq \sqrt{h}$ (for all arbitrary vectors $\underset{\sim}{d}$ of unit modulus)".*

2.4. *A result in set-theoretic logic.* It is well known that the statement "If E_1, then E_2" is equivalent to the statement "E_2 is a necessary condition for E_1" which again is equivalent to the statement "$E_1 \subset E_2$". All these statements imply that "$P(E_1) \leqq P(E_2)$", which is a necessary (though not a sufficient) condition for the other statements. This will be used in the derivation of the confidence bounds.

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

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

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

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

$$(3.1.2)\qquad \begin{aligned} (1/n)\theta_{1\alpha}(p, n) &\leqq \text{all } c(S\Sigma^{-1}) \leqq (1/n)\theta_{2\alpha}(p, n), \\ n\theta_{1\alpha}^{-1}(p, n) &\geqq \text{all } c(\Sigma S^{-1}) \geqq n\theta_{2\alpha}^{-1}(p, n). \end{aligned}$$

From (2.3.6) we observe that this implies

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

which is thus a set of simultaneous confidence bounds with a confidence coefficient $\geqq 1 - \alpha$. We note that, by using Lemma C, we can replace "all $c(\Sigma)$" occurring in the middle of (3.1.3) by "$\underset{\sim}{a}'\Sigma\underset{\sim}{a}$ (for all arbitrary vectors $\underset{\sim}{a}$ of unit modulus)."

From Lemma B we also observe that (3.1.2) implies

$$(3.1.4) \qquad [n\theta_{1\alpha}^{-1}(p, n)]^t \operatorname{tr}_t(S) \geqq \operatorname{tr}_t(\Sigma) \geqq [n\theta_{2\alpha}^{-1}(p, n)]^t \operatorname{tr}_t(S),$$

for $t = 1, 2, \cdots, p$, which is thus also another set of simultaneous confidence bounds with a confidence coefficient $\geqq 1 - \alpha$. Using (2.3.1), $\operatorname{tr}_t(S)$ and $\operatorname{tr}_t(\Sigma)$ 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 (2.1.6) we have, with a confidence coefficient $1 - \alpha$, the confidence bounds

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

Using (2.3.2) and (2.3.6) we have

$$(3.2.2) \qquad \begin{aligned} c_{\max}[S_2(\mu')^{-1}D_{\sqrt{\Theta}}\mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\mu^{-1}]c_{\max}(S_2^{-1}) \\ \geqq \text{all } c[(\mu')^{-1}D_{\sqrt{\Theta}}\mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\mu^{-1}] \equiv \text{all } c(S_1^{-1}\Delta) \\ \geqq c_{\min}[S_2(\mu')^{-1}D_{\sqrt{\Theta}}\mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\mu^{-1}]\, c_{\min}(S_2^{-1}), \end{aligned}$$

where

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

In the same way we have

$$(3.2.4) \qquad c_{\max}(S_1^{-1}\Delta)\, c_{\max}(S_1) \geqq \text{all } c(\Delta) \geqq c_{\min}(S_1^{-1}\Delta)\, c_{\min}(S_1).$$

Furthermore, noting that

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

and using (2.3.5), we have

$$(3.2.6) \quad c_{\max}(\Delta) \geqq \text{all } c^2(\mu D_{\sqrt{\Theta}}\mu^{-1}) \equiv \text{all } c^2(D_{\sqrt{\Theta}}) \equiv \text{all } \Theta_i \geqq c_{\min}(\Delta).$$

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

$$(3.2.7) \quad \begin{aligned} c_{\max}(S_2(\mu')^{-1}D_{\sqrt{\Theta}}\mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\mu^{-1})\, c_{\max}(S_2^{-1})\, c_{\max}(S_1) \\ \geqq \text{all } \Theta_i \geqq c_{\min}(S_2(\mu')^{-1}D_{\sqrt{\Theta}}\mu' S_1^{-1}\mu D_{\sqrt{\Theta}}\mu^{-1})\, c_{\min}(S_2^{-1})\, c_{\min}(S_1). \end{aligned}$$

From this it is easy to check that (3.2.1) implies

$$(3.2.8) \quad \begin{aligned} (n_1/n_2)\Theta_{1\alpha}^{-1}(p, n_1, n_2)\, c_{\max}(S_2^{-1})\, c_{\max}(S_1) \geqq \text{all } c(\Sigma_1\Sigma_2^{-1}) \\ \geqq (n_1/n_2)\theta_{2\alpha}^{-1}(p, n_1, n_2)\, c_{\min}(S_2^{-1})\, c_{\min}(S_1), \end{aligned}$$

which is thus a set of simultaneous confidence bounds with a confidence coefficient $\geqq 1 - \alpha$. We observe that, by using Lemma D we can replace "all $c(\Sigma_1\Sigma_2^{-1})$" occurring in the middle of (3.2.8) by "$\underset{\sim}{a}'\Sigma_1\underset{\sim}{a}/\underset{\sim}{a}'\Sigma_2\underset{\sim}{a}$ (for all arbitrary nonnull vectors $\underset{\sim}{a}(p \times 1)$)." We notice that

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

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 would appear to be less worthwhile than 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) $H(\Sigma = \Sigma_0)$ and (ii) $H(\Sigma_1 = \Sigma_2)$. They are, respectively,

(3.3.1) $\quad H(\Sigma = \Sigma_0)\colon \theta_{1\alpha}(p, n) \leqq \theta_1 \leqq \theta_p \leqq \theta_{2\alpha}(p, n),$

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

In the first case it is possible to choose $\theta_{1\alpha}$ and $\theta_{2\alpha}$ (and this choice will be unique) so 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, that is to let the power have a local maximum at $\Sigma = \Sigma_0$, when $\Sigma \neq \Sigma_0$ is supposed to be the alternative. In this case it so happens that the resulting power function then monotonically increases as each $c_i(\Sigma\Sigma_0^{-1})$ tends away from unity, provided that all are $\geqq 1$ or $\leqq 1$, to begin with.

In the second case, we have an exactly similar situation, $H(\Sigma = \Sigma_0)$ being replaced by $H(\Sigma_1 = \Sigma_2)$ and $\Sigma\Sigma_0^{-1}$ being replaced by $\Sigma_1\Sigma_2^{-1}$. The effect 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 published shortly. 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 Sections 3.1 and 3.2, with confidence coefficients $\geqq 1 - \alpha$, the only difference being that they will 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. Confidence bounds on the regression matrix $\Sigma_{12}\Sigma_{22}^{-1}$ or β. It is well known [1] that the statement (2.2.13), for all arbitrary nonnull $\underset{\sim}{b}_1$ and $\underset{\sim}{b}_2$, is exactly equivalent to

(4.1) $\quad$ all $\theta_i \leqq \theta_0 \quad$ or $\quad \theta_p \leqq \theta_0,$

where the θ_i's, for $i = 1, \cdots, p$ and $0 \leqq \theta_1 \leqq \cdots \leqq \theta_p \leqq 1$, are the roots of the determinantal equation in θ

$$(4.2)\quad |\theta(S_{11} - S_{12}\beta' - \beta S_{12}' + \beta S_{22}\beta') - (S_{12} - \beta S_{22})S_{22}^{-1}(S_{12}' - S_{22}\beta')| = 0$$

Now put $\lambda \equiv \theta/(1 - \theta)$, so that we have, from (4.2), the determinantal equation in λ

$$(4.3)\quad |\lambda(S_{11} - S_{12}S_{22}^{-1}S_{12}') - (S_{12}S_{22}^{-1} - \beta)S_{22}(S_{22}^{-1}S_{12}' - \beta')| = 0.$$

Statement (4.1) can now be replaced by the statement that the largest root of (4.3) is not greater than $\lambda \equiv \theta_0/(1 - \theta_0)$, that is,

$$(4.4)\quad \text{all } c[(S_{11} - S_{12}S_{22}^{-1}S_{12}')^{-1}(B - \beta)S_{22}(B' - \beta')] \leqq \theta_0/(1 - \theta_0),$$

where $B(p \times q) = S_{12}S_{22}^{-1}$. This B may be called appropriately the matrix of sample regression of the p-set on the q-set.

We note that (4.4) is equivalent to (4.1) which again is equivalent to (2.2.9), so that θ_p is the largest characteristic root of the matrix $(Z_1Z_1')^{-1}(Z_1Z_2')(Z_2Z_2')^{-1} \times (Z_2Z_1')$, where (Z_1, Z_2) have the p.d.f. given by (2.2.8). The joint distribution of these central θ_i's, and also of the largest root θ_p are known; thus all that we have to do to make (4.4), that is (4.1), that is, (2.2.9), a simultaneous confidence statement with a joint confidence coefficient $1 - \alpha$ is to choose $\theta_0 = \theta_\alpha(p, q, n - 1) = \theta_\alpha$ (say), where θ_0 or θ_α is defined by $P(\text{central } \theta_p \geqq \theta_0) = \alpha$.

Now, as in Sections 3.1 and 3.2, using (2.3.5) and the result in Section 2.4, we have from (4.4), with a joint confidence coefficient $\geqq 1 - \alpha$, the simultaneous confidence statement that

$$(4.5)\quad \text{all } c[(B - \beta)(B' - \beta')] \leqq [\theta_\alpha/(1 - \theta_\alpha)]c_{\max}(S_{11} - S_{12}S_{22}^{-1}S_{12}')c_{\max}(S_{22}^{-1}).$$

We now note that

$$c_{\max}(S_{22}^{-1}) = 1/c_{\min}(S_{22}),$$
$$c_{\max}(S_{11} - S_{12}S_{22}^{-1}S_{12}') \leqq c_{\max}(S_{11})c_{\max}(I - S_{11}^{-1}S_{12}S_{22}^{-1}S_{12}'),$$
$$c_{\max}(I - S_{11}^{-1}S_{12}S_{22}^{-1}S_{12}') = 1 - c_{\min}(S_{11}^{-1}S_{12}S_{22}^{-1}S_{12}').$$

Using these, we check that (4.5) can be replaced (with a confidence coefficient $\geqq 1 - \alpha$) by

$$(4.6)\quad \text{all } c[(B - \beta)(B' - \beta')] \leqq \frac{\theta_\alpha}{1 - \theta_\alpha}[1 - c_{\min}(S_{11}^{-1}S_{12}S_{22}^{-1}S_{12}')]\frac{c_{\max}(S_{11})}{c_{\min}(S_{22})}.$$

Letting h denote the right side of (4.6), and applying the Lemmas C and E to (4.6) we have, with a joint confidence coefficient $\geqq 1 - \alpha$, the following equivalent simultaneous confidence statements for all arbitrary unit modulus vectors $\underline{d}_1(p \times 1)$ and $\underline{d}_2(q \times 1)$,

$$(4.7)\quad |\underline{d}_1'(B - \beta)\underline{d}_2| \leqq \sqrt{h}, \qquad \underline{d}_1'B\underline{d}_2 - \sqrt{h} \leqq \underline{d}_1'\beta\underline{d}_2 \leqq \underline{d}_1'B\underline{d}_2 + \sqrt{h}.$$

A set of simultaneous confidence bounds on just the elements β_{ij} of the β-matrix would be a subset of the bounds on the total set $\underline{d}_1'\beta\underline{d}_2$. It is worthwhile to check that, if $p = q = 1$, (4.7) reduces, as it should, to (2.2.1). Also, if $p = 1$, we should have another special case of (4.7) giving a set of simultaneous confidence bounds on all linear functions of the partial regressions of one variate on several others. Thus, in several ways, (4.7) seems to be an appropriate generalization of (2.2.1).

REFERENCES

[1] S. N. Roy and R. C. Rose, "Simultaneous confidence interval estimation," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 513–536.

[2] S. N. Roy, "A useful theorem in matrix theory," *Proc. Amer. Math. Soc.*, Vol. 5 (1954), pp. 635–638.