CONFIDENCE BOUNDS ON THE "RATIO OF MEANS" AND THE "RATIO OF VARIANCE" FOR CORRELATED VARIATES

by

S.N.Roy and R.F. Potthoff
Department of Statistics, University of North Carolina

This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under contract No. AF 18(600)-83. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Institute of Statistics
Mimeograph Series No. 170
May, 1957

CONFIDENCE BOUNDS ON THE "RATIO OF MEANS" AND THE "RATIO OF VARIANCES" FOR CORRELATED VARIATES

by

S.N.Roy and R.F. Potthoff
Department of Statistics, University of North Carolina

1. Summary. In this paper confidence bounds are obtained (i) on the ratio of variances of a (possibly) correlated bivariate normal population, and then, by generalization, (ii) on a set of parametric functions of a (possibly) correlated p + p variate normal population, which plays the same role for a 2p-variate population as the ratio of variances does for the bivariate case, (iii) on the ratio of means of the population indicated in (i) and, by generalization, (iv) on a set of parametric functions of the population indicated in (ii), which plays the same role for this problem as the ratio of means does for the bivariate case. For (i) and (iii) the confidence coefficient is any pre-assigned $1-\alpha$ and the distribution involved is the central t-distribution, while for (ii) and (iv), the confidence statement in each case is a simultaneous one with a joint confidence coefficient greater than or equal to a preassigned $1-\alpha$. For (ii) the distribution involved is that of the central largest canonical correlation coefficient (squared), and for (iv) the distribution involved is that of the central Hotelling's T^2. As far as the authors are aware the results on (ii) and (iv) are new and so perhaps that on (i). But the result on (iii) has been in the field for a long time in various superficially different forms. An important point to keep in mind on these problems is that, for such confidence bounds and the associated tests of hypotheses to be physically meaningful, the two variates for the bivariate distribution should be comparable. For example, they might refer to the same characteristic of a set of individuals before and after a feed. Likewise, for a (p + p)-variate distribution, the p variates of the first set should be comparable to p variates of the second set. For example, they might refer to several

characteristics of a set of individuals before and after a treatment.

2. Confidence bounds for the case (i). Suppose we have a random sample of size $n(>2)$ from a population:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : N\left[\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}\right].$$

Let us denote the sample means by $\bar{x}_1$, $\bar{x}_2$, and the sample dispersion matrix by

$$\begin{bmatrix} s_1^2 & s_1 s_2 r \\ s_1 s_2 r & s_2^2 \end{bmatrix}.$$

Then for any constant λ, it is easy to check that covariance $(x_1-\lambda x_2, x_1+\lambda x_2)$ is $\text{var}(x_1)-\lambda^2\text{var}(x_2)=\sigma_1^2-\lambda^2\sigma_2^2$.

This will be zero if $\lambda^2 = \sigma_1^2/\sigma_2^2$. Thus, with a $\lambda^2=\sigma_1^2/\sigma_2^2$, the variates $x_1-\lambda x_2$ and $x_1+\lambda x_2$ will be uncorrelated and hence, denoting by r^* the sample correlation coefficient between these two variates, we have that r^* has the (central) r-distribution, i.e., $\sqrt{n-2}\, r^*/(1-r^{*2})^{1/2}$ has the (central) t-distribution with d.f. $(n-2)$. But it is easy to check that

$$(2.1)\quad r^*=(s_1^2-\lambda^2 s_2^2)/[(s_1^2+\lambda^2 s_2^2+2\lambda s_1 s_2 r)(s_1^2+\lambda^2 s_2^2-2\lambda s_1 s_2 r)]^{1/2}$$

$$= (s_1^2-\lambda^2 s_2^2)/[s_1^4+\lambda^4 s_2^4+2\lambda^2 s_1^2 s_2^2(1-2r^2)]^{1/2}.$$

Now, starting from the statement (with a probability $1-\alpha$)

$$(2.2)\quad \sqrt{n-2}\, r^*/(1-r^{*2})^{1/2} \leq t_{\alpha/2}(n-2), \text{ or } \leq t_{\alpha/2} \text{(more simply)},$$

where $t_{\alpha/2}(n-2)$ is the upper $\alpha/2$-point of the (central) t-distribution with d.f.

(n-2), and remembering that $\lambda=\sigma_1/\sigma_2$ and substituting from (2.1) for r^* in terms of s_1, s_2 and r, we have, for σ_1^2/σ_2^2, the following confidence equation (2.3) and confidence bounds (2.4)(with a confidence coefficient $1-\alpha$)

$$(2.3) \qquad \lambda^4 - \left[2 + \frac{4}{n-2}\, t^2_{\alpha/2}(1-r^2) \right] \frac{s_1^2}{s_2^2}\, \lambda^2 + \frac{s_1^4}{s_2^4} \leq 0 \quad ,$$

and

$$(2.4) \qquad \frac{s_1^2}{s_2^2} \left[\left(1 + \frac{2}{n-2}\, t^2_{\alpha/2}\overline{1-r^2}\right) - \left\{ \left(1 + \frac{2}{n-2}\, t^2_{\alpha/2}\,\overline{1-r^2}\right)^2 - 1 \right\}^{1/2} \right] \leq \frac{\sigma_1^2}{\sigma_2^2}$$

$$\leq \frac{s_1^2}{s_2^2} \left[\left(1 + \frac{2}{n-2}\, t^2_{\alpha/2}\,\overline{1-r^2}\right) + \left\{ \left(1 + \frac{2}{n-2}\, t^2_{\alpha/2}\,\overline{1-r^2}\right)^2 - 1 \right\}^{1/2} \right] .$$

We notice that $\lambda = \sigma_1/\sigma_2 = 1$ if and only if $\sigma_1=\sigma_2$, so that we accept $H_0: \sigma_1=\sigma_2$ if the confidence interval (2.3) includes 1.

3. <u>Confidence bounds for the case (ii).</u> Suppose we have

$$\underline{x}(2p \times 1) = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \begin{matrix} p \\ p \end{matrix} : N \left[\begin{bmatrix} \underline{\xi}_1 \\ \underline{\xi}_1 \end{bmatrix} \begin{matrix} p \\ p \end{matrix} , \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix} \begin{matrix} p \\ p \end{matrix} \right]$$
$$\qquad\qquad 1 \qquad\qquad\quad 1 \qquad\quad p \quad\; p$$

$$= N \left[\underline{\xi}(2p \times 1),\ \Sigma(2p \times 2p) \right] \text{(say)} ,$$

and a random sample of size $n(>2p)$ from this population, with a sample dispersion matrix denoted by

$$(3.1) \qquad \begin{bmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{bmatrix} \begin{matrix} p \\ p \end{matrix} = S(2p \times 2p) \text{ (say)} \quad .$$
$$p \qquad p$$

It is well known [3] that we can choose (non-singular) matrices $\mu(p \times p)$ and

$\nu(p \times p)$ such that

(3.2) $$\Sigma_{11}=\mu\mu', \ \Sigma_{22}=\nu\nu' \text{ and } \Sigma_{12}=\mu D_{\sqrt{\gamma}}\nu' \quad ,$$

where γ's, i.e., $\gamma_1, \gamma_2, \ldots, \gamma_p$ are the characteristic roots of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'$ and $D_{\sqrt{\gamma}}$ is a diagonal matrix whose diagonal elements are $\sqrt{\gamma_1}, \ldots, \sqrt{\gamma_p}$. It is also well known [3] that these roots are all non-negative, that the number of positive roots is the same as the rank of Σ_{12} and that all the roots are zero if, and only if, $\Sigma_{12}=0$.

Now introduce a new variate $x^*(2p \times 1)$ defined by

(3.3) $$\underline{x}^*(2p \times 1) = \begin{matrix} p \\ p \end{matrix}\begin{bmatrix} \underline{x}_1^* \\ \underline{x}_2^* \end{bmatrix} \text{ (say)} = A(2p \times 2p)\, \underline{x}\, (2p \times 1) \quad ,$$
$$\qquad\qquad 1$$

where

(3.4) $$A(2p \times 2p) = \begin{bmatrix} I & -\mu\nu^{-1} \\ I & \mu\nu^{-1} \end{bmatrix}\begin{matrix} p \\ p \end{matrix} = \begin{bmatrix} I & -\lambda \\ I & \lambda \end{bmatrix} \text{ (say)} \quad .$$
$$\qquad p \qquad\qquad p$$

Then this $\underline{x}^*$ is $N(\underline{\xi}^*, \Sigma^*)$, where $\underline{\xi}^* = A\underline{\xi}$ and

(3.5) $$\Sigma^* = \begin{bmatrix} \Sigma_{11}^* & \Sigma_{12}^* \\ \Sigma_{12}^{*\prime} & \Sigma_{22}^* \end{bmatrix} \text{ (say)} = A\,\Sigma\,A' \quad ,$$

whence we have that

(3.6) $$\Sigma_{11}^* = 2(\Sigma_{11}-\mu D_{\sqrt{\gamma}}\mu'), \ \Sigma_{22}^* = 2(\Sigma_{11}+\mu D_{\sqrt{\gamma}}\mu') \quad \text{and}$$
$$\Sigma_{12}^* = \Sigma_{11} - \mu\nu^{-1}\Sigma_{12}' + \Sigma_{12}\nu'^{-1}\mu' - \mu\nu^{-1}\Sigma_{22}\nu'^{-1}\mu'$$
$$= \Sigma_{11} - \mu D_{\sqrt{\gamma}}\mu' + \mu D_{\sqrt{\gamma}}\mu' - \Sigma_{11} = 0 \quad .$$

This means that the transformed p-set $\underline{x}_1^*$ is uncorrelated with transformed p-set $\underline{x}_2^*$. We shall put simultaneous confidence bounds on the largest and smallest characteristic roots of $\lambda\lambda'$, i.e., of $\mu\nu^{-1}\nu'^{-1}\mu'$ and then show at the end of this section how these roots are, in a sense, a generalization of σ_1^2/σ_2^2 for case (i). We may note here, incidentally, that for $p = 1$, λ does, in fact, reduce to σ_1/σ_2. Next, denoting by S^* the sample dispersion matrix of $\underline{x}^*$, we have

$$(3.7) \qquad S^*(2p \times 2p) = \begin{bmatrix} S_{11}^* & S_{12}^* \\ S_{12}^{*'} & S_{22}^* \end{bmatrix} \begin{matrix} p \\ p \end{matrix} \text{ (say)} = A\, S\, A'$$

$$\qquad\qquad\qquad\qquad\qquad\qquad p \qquad\quad p$$

$$= \begin{bmatrix} I & -\lambda \\ I & \lambda \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{bmatrix} \begin{bmatrix} I & I \\ -\lambda' & \lambda' \end{bmatrix} ,$$

whence we have

$$(3.8) \qquad \begin{aligned} S_{11}^* &= S_{11} - \lambda S_{12}' - S_{12}\lambda' + \lambda S_{22}\lambda' \\ S_{12}^* &= S_{11} - \lambda S_{12}' + S_{12}\lambda' - \lambda S_{22}\lambda' \\ S_{22}^* &= S_{11} + \lambda S_{12}' + S_{12}\lambda' + \lambda S_{22}\lambda' \end{aligned} ,$$

Now we go back to (3.6). Note that, since $\Sigma_{12}^* = 0$, the transformed $\underline{x}_1^*$-set is uncorrelated with the transformed $\underline{x}_2^*$-set, and also that, in this case, the joint distribution of the canonical correlation coefficients and also, in particular, of the largest canonical correlation coefficient is known. Thus we can find a $c_\alpha(p, p, n-1) = c_\alpha$(say) such that

$$(3.9) \qquad P\left[c_{max}(S_{11}^{*-1} S_{12}^* S_{22}^{*-1} S_{12}^{*'}) \le c_\alpha \,\middle|\, \Sigma_{12}^* = 0 \right] = 1-\alpha .$$

The problem now is to start from (3.9), use (3.8) and try to obtain confidence bounds on functions connected with $\lambda(=\mu\nu^{-1})$. For this we proceed as follows. Let c be a characteristic root of the matrix in (3.9). Then

$$(3.10) \qquad \left| cS^*_{11} - S^*_{12} S^{*-1}_{22} S^{*'}_{12} \right| = 0 \quad .$$

Putting $c = 1-4d$, this reduces to

$$(3.11) \qquad \left| dS^*_{11} - \frac{1}{4} S^*_{11} + \frac{1}{4} S^*_{12} S^{*-1}_{22} S^{*'}_{12} \right| = 0 \quad .$$

Now, using (3.8), we have

$$(3.12) \qquad -\frac{1}{4} S^*_{11} = -S_{11} + \frac{1}{4}(S^*_{12} + S^{*'}_{12} + S^*_{22})$$

$$= -S_{11} + \frac{1}{4}(S^*_{12} + S^*_{22}) S^{*-1}_{22} (S^{*'}_{12} + S^*_{22}) - \frac{1}{4} S^*_{12} S^{*-1}_{22} S^{*'}_{12} \quad .$$

Hence

$$(3.13) \qquad \left| dS^*_{11} - S_{11} + \left(\frac{S^*_{12} + S^*_{22}}{2} \right) S^{*-1}_{22} \left(\frac{S^{*'}_{12} + S^*_{22}}{2} \right) \right| = 0$$

or

$$\left| dS^*_{11} - S_{11} + (S_{11} + S_{12}\lambda') S^{*-1}_{22} (S_{11} + \lambda S'_{12}) \right| = 0$$

Next, we recall that for a non-singular $M_4(q \times q)$ we have

$$(3.14) \qquad \begin{vmatrix} M_1 & M_2 \\ M_3 & M_4 \end{vmatrix} \begin{matrix} p \\ q \end{matrix} = |M_4| \, |M_1 - M_2 M_4^{-1} M_3|$$
$$\begin{matrix} p & q \end{matrix}$$

and, using this, we observe that (3.13) $\Longleftrightarrow$

$$(3.15) \qquad \begin{vmatrix} S_{11} - dS^*_{11} & S_{11} + S_{12}\lambda' \\ S_{11} + \lambda S'_{12} & S_{11} + \lambda S'_{12} + S_{12}\lambda' + \lambda S_{22}\lambda' \end{vmatrix} = 0$$

$$\Longleftrightarrow \begin{vmatrix} S_{11} - dS^*_{11} & S_{12}\lambda' + dS^*_{11} \\ S_{11} + \lambda S'_{12} & S_{12}\lambda' + \lambda S_{22}\lambda' \end{vmatrix} = 0 \Longleftrightarrow \begin{vmatrix} S_{11} - dS^*_{11} & S_{12}\lambda' + dS^*_{11} \\ \lambda S'_{12} - dS^*_{11} & \lambda S_{22}\lambda' - dS^*_{11} \end{vmatrix} = 0$$

$$\Longleftrightarrow \left| \begin{bmatrix} S_{11} & S_{12}\lambda' \\ \lambda S'_{12} & \lambda S_{22}\lambda' \end{bmatrix} \begin{matrix} p \\ p \end{matrix} - d \begin{bmatrix} S^*_{11} & -S^*_{11} \\ -S^*_{11} & S^*_{11} \end{bmatrix} \right| = 0 \quad .$$

$$\begin{matrix} p & & p \end{matrix}$$

But we have

$$(3.16) \qquad \begin{bmatrix} S_{11} & S_{12}\lambda' \\ \lambda S'_{12} & \lambda S_{22}\lambda' \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S'_{12} & S_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \lambda' \end{bmatrix}$$

and

$$\begin{bmatrix} S^*_{11} & -S^*_{11} \\ -S^*_{11} & S^*_{11} \end{bmatrix} = \begin{bmatrix} I \\ -I \end{bmatrix} S^*_{11} \begin{bmatrix} I & -I \end{bmatrix}$$

$$= \begin{bmatrix} I \\ -I \end{bmatrix} \begin{bmatrix} I & -\lambda \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S'_{12} & S_{22} \end{bmatrix} \begin{bmatrix} I \\ -\lambda' \end{bmatrix} \begin{bmatrix} I & -I \end{bmatrix} .$$

Hence (3.15) reduces to

$$\left| \begin{bmatrix} I & 0 \\ 0 & \lambda \end{bmatrix} S \begin{bmatrix} I & 0 \\ 0 & \lambda' \end{bmatrix} - d \begin{bmatrix} I \\ -I \end{bmatrix} \begin{bmatrix} I & -\lambda \end{bmatrix} S \begin{bmatrix} I \\ -\lambda' \end{bmatrix} \begin{bmatrix} I & -I \end{bmatrix} \right| = 0 \quad ,$$

$$\Longleftrightarrow$$

$$(3.17) \qquad \left| eS - \begin{bmatrix} I & 0 \\ 0 & \lambda^{-1} \end{bmatrix} \begin{bmatrix} I \\ -I \end{bmatrix} \begin{bmatrix} I & -\lambda \end{bmatrix} S \begin{bmatrix} I \\ -\lambda' \end{bmatrix} \begin{bmatrix} I & -I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \lambda'^{-1} \end{bmatrix} \right| = 0 \ ,$$

where $e = 1/d$, which again reduces to

$$(3.18) \qquad \left| e\, I(2p \times 2p) - S^{-1} \beta S \beta' \right| = 0 \quad ,$$

where

(3.19) $$\beta(2p \times 2p) = \begin{bmatrix} I & 0 \\ 0 & \lambda^{-1} \end{bmatrix} \begin{bmatrix} I \\ -I \end{bmatrix} [I \quad -\lambda] = \begin{bmatrix} I & -\lambda \\ -\lambda^{-1} & I \end{bmatrix} .$$

Now we go back to (3.9), recall that $e = 1/d = 4/(1-c)$, put $e_\alpha = 4/(1-c_\alpha)$, observe that $c_{max} \le c_\alpha \Longleftrightarrow e_{max} \le e_\alpha$, and hence that (3.9) $\Longleftrightarrow$

$$P[c_{max}[S^{-1}\beta S\beta'] \le e_\alpha \quad \Sigma_{12}^* = 0] = 1-\alpha ,$$

or

(3.20) $$P\left[\frac{(\underline{a}'\beta\underline{b})^2}{(\underline{a}'\underline{a})(\underline{b}'\underline{b})} \le e_\alpha \frac{\underline{a}'S\underline{a}}{\underline{a}'\underline{a}} \cdot \frac{\underline{b}'S^{-1}\underline{b}}{\underline{b}'\underline{b}} \,\middle|\, \Sigma_{12}^* = 0, \text{ for all non null } \underline{a}(2p \times 1) \text{ and } \underline{b}(2p \times 1)\right] = 1-\alpha .$$

Next, consider, for all non null $\underline{a}$ and $\underline{b}$, the statement

(3.21) $$\frac{(\underline{a}'\beta\underline{b})^2}{(\underline{a}'\underline{a})(\underline{b}'\underline{b})} \le e_\alpha \frac{\underline{a}'S\underline{a}}{\underline{a}'\underline{a}} \cdot \frac{\underline{b}'S^{-1}\underline{b}}{\underline{b}'\underline{b}} .$$

Now specialize $\underline{a}'(2p \times 1)$ and $\underline{b}'(2p \times 1)$ into $[\underset{p}{\underline{a}_1'} \quad \underset{p}{\underline{0}}]1$ and $[\underset{p}{\underline{b}_1'} \quad \underset{p}{\underline{0}}]1$, and also into $[\underset{p}{0} \quad \underset{p}{\underline{a}_2'}]1$ and $[\underset{p}{0} \quad \underset{p}{\underline{b}_2'}]1$.

We next set

(3.22) $$S^{-1}(2p \times 2p) = \begin{bmatrix} S^{11} & S^{12} \\ S^{12'} & S^{22} \end{bmatrix} \begin{matrix} p \\ p \end{matrix} ,$$
$$\qquad\qquad\quad p \qquad\; p$$

whence we have

(3.23) $$S^{11} = (S_{11} - S_{12}S_{22}^{-1}S_{12}')^{-1}, \quad S^{22} = (S_{22} - S_{12}'S_{11}^{-1}S_{12})^{-1} \text{ and}$$

$$S^{12} = -S^{11}S_{12}S_{22}^{-1} = -S_{11}^{-1}S_{12}S^{22} .$$

Back in (3.21) we now observe that (3.21) $\Longrightarrow$

$$(3.24) \qquad \frac{(\underline{a}_1'\lambda\underline{b}_2)^2}{(\underline{a}_1'\underline{a}_1)(\underline{b}_2'\underline{b}_2)} \leq e_\alpha \frac{\underline{a}_1'S_{11}\underline{a}_1}{\underline{a}_1'\underline{a}_1} \quad \frac{\underline{b}_2'S^{22}\underline{b}_2}{\underline{b}_2'\underline{b}_2}$$

for all non null $\underline{a}_1$ and $\underline{b}_2$, and that (3.21) also $\Longrightarrow$

$$(3.25) \qquad \frac{(\underline{a}_2\lambda^{-1}\underline{b}_1)^2}{(\underline{a}_2'\underline{a}_2)(\underline{b}_1'\underline{b}_1)} \leq e_\alpha \frac{\underline{a}_2'S_{22}\underline{a}_2}{\underline{a}_2'\underline{a}_2} \cdot \frac{\underline{b}_1'S^{11}\underline{b}_1}{\underline{b}_1'\underline{b}_1} \quad ,$$

for all non null $\underline{a}_2$ and $\underline{b}_1$. If now we consider the left side of (3.24) then it follows from Cauchy's inequality that for all non null $\underline{b}_2$,

$$(\underline{a}_1'\lambda\underline{b}_2)^2/(\underline{a}_1'\underline{a}_1)(\underline{b}_2'\underline{b}_2) \leq (\underline{a}_1'\lambda\lambda'\underline{a}_1)/(\underline{a}_1'\underline{a}_1)$$

and it is also well known that for all non null $\underline{a}_1$,

$$c_{min}(\lambda\lambda') \leq (\underline{a}_1'\lambda\lambda'\underline{a}_1)/(\underline{a}_1'\underline{a}_1) \leq c_{max}(\lambda\lambda') \ .$$

We have also exactly similar results by interchanging $\underline{a}_1$ and $\underline{b}_2$, and similar results on the left side of (3.25), in terms of λ^{-1} and $\underline{a}_2$ and $\underline{b}_1$ and then by the interchange of $\underline{a}_2$ and $\underline{b}_1$.

Next, maximizing the left side of (3.24) w.r.t. $\underline{a}_1$ and $\underline{b}_2$, we observe [2,3,4] that (3.24) and hence (3.21) $\Longrightarrow$

$$c_{max}(\lambda\lambda') \leq e_\alpha c_{max}(S_{11})c_{max}(S^{22}),$$

or, using (3.23),

$$(3.26) \qquad c_{max}(\lambda\lambda') \leq e_\alpha c_{max}(S_{11})/c_{min}(S_{22}-S_{12}'S_{11}^{-1}S_{12}) \ .$$

Likewise, maximizing the left side of (3.25) w.r.t. $\underline{a}_2$ and $\underline{b}_1$, we observe [4] that (3.25) and hence (3.21) $\Longrightarrow$

(3.27) $$c_{max}(\lambda'^{-1}\lambda^{-1}) \leq e_\alpha c_{max}(S_{22})c_{max}(S^{11}) .$$

Now recall that [3], since all non zero roots of $\lambda^{-1}\lambda'^{-1}$ are also roots of $\lambda'^{-1}\lambda^{-1}$, i.e., of $(\lambda\lambda')^{-1}$ and λ is non singular, therefore, $c_{min}(\lambda^{-1}\lambda'^{-1})=c_{min}(\lambda\lambda')^{-1}$ $= 1/c_{max}(\lambda\lambda')$ and also similarly that $c_{min}(\lambda^{-1}\lambda'^{-1})=1/c_{max}(\lambda\lambda')$. At this point, using (3.23) we observe that (3.27) and hence (3.25) and hence (3.21) $\Longrightarrow$

(3.28) $$c_{min}(\lambda\lambda') \geq \frac{1}{e_\alpha} c_{min}(S_{11}-S_{12}S_{22}^{-1}S_{12}')/c_{max}(S_{22}) .$$

Also, giving back to (3.24) and first maximizing the left side of it w.r.t. $\underline{b}_2$ and then minimizing the right side w.r.t. $\underline{a}_1$, we observe [4] that (3.24) and hence (3.21) $\Longrightarrow$

(3.29) $$c_{min}(\lambda\lambda') \leq e_\alpha c_{min}(S_{11})/c_{min}(S_{22}-S_{12}'S_{11}^{-1}S_{12}) ,$$

and, furthermore, first maximizing the left side w.r.t. $\underline{a}_1$ and then minimizing the right side w.r.t. $\underline{b}_2$, we observe [4] that (3.24) and hence (3.21) also $\Longrightarrow$

(3.30) $$c_{min}(\lambda\lambda') \leq e_\alpha c_{max}(S_{11})/c_{max}(S_{22}-S_{12}'S_{11}^{-1}S_{12}) .$$

Likewise, back in (3.25), first maximizing the left side w.r.t. $\underline{b}_1$ and then minimizing the right side w.r.t. $\underline{a}_2$, we observe [4] that (3.25) and hence (3.21) $\Longrightarrow$

(3.31) $$c_{max}(\lambda\lambda') \geq \frac{1}{e_\alpha} c_{min}(S_{11}-S_{12}S_{22}^{-1}S_{12}')/c_{min}(S_{22}) ,$$

and first maximizing the left side w.r.t. $\underline{a}_2$ and then minimizing the right side w.r.t. $\underline{b}_1$, we observe [4] that (3.25) and hence (3.21) also $\Longrightarrow$

(3.32) $$c_{max}(\lambda\lambda') \geq \frac{1}{e_\alpha} c_{max}(S_{11}-S_{12}S_{22}^{-1}S_{12}')/c_{max}(S_{22}) .$$

Now combining (3.26), (3.28), (3.29)-(3.32), we observe that (3.21) $\Longrightarrow$ all these statements, and hence, going back to (3.20), we have with a joint probability

$\geq 1-\alpha$, the bounds

(3.33) $$\frac{1}{e_\alpha} c_{min}(S_{11}-S_{12}S_{22}^{-1}S_{12}')/c_{max}(S_{22}) \leq c_{min}(\lambda\lambda')$$

$$\leq e_\alpha \min [c_{min}(S_{11})/c_{min}(S_{22}-S_{12}'S_{11}^{-1}S_{12}), c_{max}(S_{11})/c_{max}(S_{22}-S_{12}'S_{22}^{-1}S_{12})]$$

and

(3.34) $$\frac{1}{e_\alpha}\max[c_{min}(S_{11}-S_{12}S_{22}^{-1}S_{12}')/c_{min}(S_{22}), c_{max}(S_{11}-S_{12}S_{22}^{-1}S_{12}')/c_{max}(S_{22})]$$

$$\leq c_{max}(\lambda\lambda') \leq e_\alpha c_{max}(S_{11})/c_{min}(S_{22}-S_{12}'S_{11}^{-1}S_{12}) .$$

It is interesting to use [3] and check that the lower bound of (3.33) is $\leq$ the upper bound of (3.34), but that the upper bound of (3.33) might be $\geq$ or $<$ the lower bound of (3.34). However, it is to be always remembered that $c_{min}(\lambda\lambda') \leq c_{max}(\lambda\lambda')$, which should imply an obvious restriction on combined bounds on $c_{max}(\lambda\lambda')$ and $c_{min}(\lambda\lambda')$.

Truncation. Going back to (3.24) again we can proceed as in [4], equate to zero any element of $\underline{a}_1$ and the corresponding elements of $\underline{b}_2$, $\underline{a}_2$ and $\underline{b}_1$ (it has to be the corresponding elements, in order to make the process physically meaningful) and then apply the process of maximization, minimization etc., leading ultimately to the same kind of statements as (3.33) and (3.34) in terms, however, of truncated matrices everywhere, with one variate of the first p-set and the corresponding variate of the second p-set being cut out. Thus there will be $\binom{p}{1}$, i.e., p pairs of such statements. Likewise equating to zero any two elements of $\underline{a}_1$ and the corresponding elements of $\underline{b}_2$, $\underline{a}_2$ and $\underline{b}_1$, we are ultimately led to $\binom{p}{2}$ pairs of statements like (3.33) and (3.34) based on different possible sets of (p-2) variates, and so on. Ultimately we have $1 + \binom{p}{1} + \binom{p}{2} + \dots + \binom{p}{p-1}$, i.e., 2^p-1 pairs of statements like (and including)(3.33) and (3.34) with a joint probability $\geq 1-\alpha$. It should be noticed that on all these statements e_α, however, stays the same.

Interpretation of the role of the characteristic roots of $\lambda\lambda'$. The characteristic roots of $\lambda\lambda'$, i.e., of $\mu\nu^{-1}\nu'^{-1}\mu'$ are all equal to unity if and only if $\mu\nu^{-1}\nu'^{-1}\mu'$ is an identity matrix, i.e., if and only if

(3.35) $$\mu\nu^{-1} = A, \text{ that is, } \mu = A\nu \quad ,$$

where A is any arbitrary orthogonal matrix. Going back to (3.2) we easily check that (3.35) $\Longrightarrow$

(3.36) $$\Sigma_{11} = A\Sigma_{22}A' \quad ,$$

which, if we recall that A is $\perp$, and Σ_{11} and Σ_{22} are symmetric, is precisely the condition that Σ_{11} and Σ_{22} are to be similar matrices. Furthermore, using (3.2) again it is easy to see that (3.35) also $\Longrightarrow$

(3.37) $$\Sigma_{12} = \mu D_{\sqrt{\gamma}}\nu' = A\nu D_{\sqrt{\gamma}}\nu' = A \times \text{a symmetric matrix} \ ,$$

where A is the same $\perp$ matrix that occurs in (3.36). Thus (3.35) $\Longrightarrow$ (3.36) and (3.37) and it is also easy to verify that (3.36) and (3.37) $\Longrightarrow$ (3.35). Hence all the characteristic roots of $\lambda\lambda'$, i.e., of $\mu\nu^{-1}\nu'^{-1}\mu''$ being unity is a necessary and sufficient condition that the relations (3.36) and (3.37) should hold. The deviation of these characteristic roots from unity might be regarded as a (joint) measure of departure from the hypothesis given by (3.36) and (3.37) of which a very special case is the one that we get for the bivariate problem. Further statistical implications of (3.36) and (3.37) will be discussed in a later paper.

4. Confidence bounds for the case (iii). Starting from the bivariate normal distribution characterized in section 2, put $q = \xi_1/\xi_2$ and introduce a new variate $z = x_1 - qx_2$ (assume that $\xi_2 \neq 0$, i.e., $q \neq \pm\infty$). Then z is $N(0,\sigma_z^2)$, where $\sigma_z^2 = \sigma_1^2 - 2q \quad \sigma_1\sigma_2 + q^2\sigma_2^2$. Thus $\sqrt{n}\,\bar{z}/s_z = \sqrt{n}(\bar{x}_1 - q\bar{x}_2)/(s_1^2 - 2qs_1s_2r + q^2s_2^2)^{1/2}$ has the

(central) t-distribution with d.f.(n-1), so that we can find a $t_{\alpha/2}$ such that

$$P\left[n(\bar{x}_1 - q\bar{x}_2)^2/(s_1^2 - 2qs_1s_2r + q^2s_2^2) \le t_{\alpha/2}^2 \;\middle|\; q = \frac{\xi_1}{\xi_2} \right] = 1-\alpha$$

or

(4.1) $$P\left[(ks_2^2 - \bar{x}_2^2)q^2 - 2(ks_1s_2r - \bar{x}_1\bar{x}_2)q + (ks_1^2 - \bar{x}_1^2) \ge 0 \right] = 1 - \alpha$$

where $k = \frac{1}{n} t_{\alpha/2}^2$.

Subject to the restriction that q is to have real values, the statement within the parantheses in (4.1) gives the confidence bounds on $q = \xi_1/\xi_2$. There is also the further restriction that (4.1) is supposed to be a probability statement on $\bar{x}_1$, $\bar{x}_2$, s_1 and s_2 for all real values of $q = \xi_1/\xi_2$, except for $\xi_2 = 0$, i.e., for $q = \pm \infty$. Equating to zero the expression on the left side of the inequality statement under the probability sign in (4.1), we have an equation in q whose coefficients involve stochastic variates. The nature of the roots of this equation will determine the nature of the confidence region for q. Some of these regions will be meaningful and some not so. This pattern will also be repeated in the next problem, and the whole issue will be discussed in detail in a later paper.

5. <u>Confidence bounds for the case (iv)</u>. Starting from the (p + p) variate normal distribution characterized in section 3, define a set of q's, $q_1, q_2, \ldots, q_p$ by $\underline{\xi}_1 = D_q \underline{\xi}_2$ where $D_q(p \times p)$ is a diagonal matrix whose diagonal elements are $q_1, \ldots, q_p$. Introduce a new variate $\underline{z}(p \times 1)$ defined by

(5.1) $$\underline{z}(p \times 1) = p\left[\underset{p}{I} \quad \underset{p}{-D_q} \right] \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \begin{matrix} p \\ p \end{matrix} = A(p \times 2p)\underline{x}(2p \times 1) \text{(say)} .$$

It is easy to check that $E(\underline{z}) = \underline{\xi}_1 - D_q\underline{\xi}_2 = 0$, whence $\underline{z}$ is $N(\underline{0}, \Sigma_z)$ where $\Sigma_z = A\Sigma A'$. Also given the sample dispersion matrix of $\underline{x}(2p \times 1)$

(5.2) $$S(2p \times 2p) = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{bmatrix} ,$$

we have that the sample dispersion matrix of $\underline{z}(p \times 1)$

(5.3) $$S_z = ASA' = S_{11} - D_q S_{12}' - S_{12} D_q + D_q S_{22} D_q .$$

Also the sample mean vector of $\underline{z}(p \times 1)$

(5.4) $$\bar{\underline{z}} = \bar{\underline{x}}_1 - D_q \bar{\underline{x}}_2 .$$

Thus, with the q's defined as above, $n\, \bar{\underline{z}}' S_z^{-1} \bar{\underline{z}}$ is distributed as (central) Hotelling's T^2, which means that we can find a T_α^2 such that

(5.5) $$P[\, \bar{\underline{z}}' S_z^{-1} \bar{\underline{z}} \leq \frac{1}{n} T_\alpha^2 \mid q\text{'s defined as above} \,] = 1 - \alpha .$$

Now consider the statement within the parentheses in (5.5). It is well known that this statement $\iff$ all $c[\, \bar{\underline{z}}\, \bar{\underline{z}}'\, S_z^{-1} \,] \leq T_\alpha^2/n$, which again $\iff$

(5.6) $$\frac{\underline{a}'\bar{\underline{z}}\,\bar{\underline{z}}'\underline{a}}{\underline{a}'\underline{a}} \leq \frac{T_\alpha^2}{n} \cdot \frac{\underline{a}' S_z \underline{a}}{\underline{a}'\underline{a}} ,$$

for all non null $\underline{a}(p \times 1)$'s. Considering the left side of (5.6), we use again Cauchy's inequality to obtain that for all non null $\underline{a}$'s $\underline{a}'\bar{\underline{z}}/(\underline{a}'\underline{a})^{1/2} \leq +(\bar{\underline{z}}'\bar{\underline{z}})^{1/2}$ whence we see that the left side of (5.6) $\leq \bar{\underline{z}}'\bar{\underline{z}}$, i.e., $\leq \sum_{i=1}^{p} (\bar{x}_{1i} - q_i \bar{x}_{2i})^2$, where $\bar{x}_{1i}$ and $\bar{x}_{2i}$ (for i = 1, 2, ..., p) stand for the i-th elements of the vectors $\bar{\underline{x}}_1$ and $\bar{\underline{x}}_2$. We also note that the largest value of the right side of (5.6) under variation of $\underline{a}$'s is, aside from the constant factor T_α^2/n, $c_{max}(S_z)$, i.e., $c_{max}(ASA')$, i.e., $c_{max}(SA'A)$. Now we use [1] to obtain that

Truncation. Here again, as in section 4, it is possible to go back again to (5.6), proceed in the same way as before and get statements like (5.8) or (5.10) on any (p-1) variate-pairs, or on any (p-2) variate-pairs, and so on, and finally any variate-pair, thus ultimately obtaining 2^p-1 confidence statements like (5.8) or (5.10), all of them with a joint confidence coefficient $\geq 1-\alpha$.

REFERENCES

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

[2] Roy, S.N., "Some further results in simultaneous confidence interval estimation," Ann. Math. Stat., Vol. 25(1954), pp. 752-761.

[3] Roy, S.N., "A report on some aspects of multivariate analysis," Institute of Statistics mimeograph series. No. 121.

[4] Roy, S.N. and R. Gnanadesikan, "Further contributions to multivariate confidence bounds," Institute of Statistics mimeograph series No. 155. (Also to be published in the Biometrika, June, 1957).