CONFIDENCE BOUNDS ON VECTOR ANALOGUES OF THE "RATIO OF MEANS" AND THE "RATIO OF VARIANCES" FOR TWO CORRELATED NORMAL VARIATES AND SOME ASSOCIATED TESTS

S. N. Roy and R. F. Potthoff

University of North Carolina

1. Summary and Introduction. 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 preassigned $1 - \alpha$ 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. In each case the confidence bounds are obtained by inverting the test of a certain hypothesis, which is indicated at its proper place. Thus, for the $(p + p)$-variate problem, we assume that there are p pairs of comparable variates and it is the pairwise comparison for these p pairs that seems, in this situation, to be physically more meaningful than anything else. Any general bounds that will be obtained in this paper are to be regarded, in a large measure, as a means to this end, although there could conceivably be physical questions, some of which will be illustrated in a later applied paper to be published elsewhere, to which these more general bounds would be pertinent.

Received May 14, 1957; revised February 18, 1958.

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

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

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., $\sqrt{n-2}\,r^* / (1 - r^{*2})^{\frac{1}{2}}$ has the (central) t-distribution with d.f. $(n - 2)$. But it is easy to check that

$$\text{(2.1)}\qquad \begin{aligned} r^* &= \frac{(s_1^2 - \lambda^2 s_2^2)}{[(s_1^2 + \lambda^2 s_2^2 + 2\lambda s_1 s_2 r)(s_1^2 + \lambda^2 s_2^2 - 2\lambda s_1 s_2 r)]^{\frac{1}{2}}} \\ &= \frac{(s_1^2 - \lambda^2 s_2^2)}{[s_1^4 + \lambda^4 s_2^4 + 2\lambda^2 s_1^2 s_2^2(1 - 2r^2)]^{\frac{1}{2}}}. \end{aligned}$$

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

$$\text{(2.2)}\qquad \sqrt{n-2}\,| r^*/(1 - r^{*2})^{\frac{1}{2}} | \leqq t_{\alpha/2}(n - 2), \text{ or } \leqq t_{\alpha/2} \text{ (more simply)},$$

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 $1 - \alpha$)

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

and

$$\text{(2.4)}\qquad \begin{aligned} &\frac{s_1^2}{s_2^2}\left[\left(1 + \frac{2}{n-2} t_{\alpha/2}^2(1 - r^2)\right) - \left\{\left(1 + \frac{2}{n-2} t_{\alpha/2}^2(1 - r^2)\right)^2 - 1\right\}^{\frac{1}{2}}\right] \leqq \frac{\sigma_1^2}{\sigma_2^2} \\ &\leqq \frac{s_1^2}{s_2^2}\left[\left(1 + \frac{2}{n-2} t_{\alpha/2}^2(1 - r^2)\right) + \left\{\left(1 + \frac{2}{n-2} t_{\alpha/2}^2(1 - r^2)\right)^2 - 1\right\}^{\frac{1}{2}}\right]. \end{aligned}$$

We notice that $\lambda = \sigma_1/\sigma_2 = 1$ if and only if $\sigma_1 = \sigma_2$.

Notice that (2.2) or (2.3) can be used as an acceptance region for the hypothesis $\sigma_1/\sigma_2 = \lambda$ (any specific value) against the alternative $\sigma_1/\sigma_2 \neq \lambda$. Since the paper was written it has been brought to the notice of the authors that this region, for the case of $\sigma_1/\sigma_2 = 1$, i.e., for $\sigma_1 = \sigma_2$, has been explicitly given by Walker and Lev [5].

3. Confidence bounds for the case (ii). Suppose we have

$$\mathbf{x}\,(2p \times 1) = \underset{1}{\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}}\begin{matrix} p \\ p \end{matrix} : N\left[\underset{1}{\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}}\begin{matrix} p \\ p \end{matrix},\ \begin{bmatrix} \Sigma_{11} & \Sigma_{11} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix}\begin{matrix} p \\ p \end{matrix}\right]$$
$$\qquad\qquad\qquad\quad p \quad p$$
$$= N[\xi(2p \times 1),\ \Sigma(2p \times 2p)] \quad \text{(say)},$$

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

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

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

$$\text{(3.2)} \qquad \Sigma_{11} = \mu\mu', \qquad \Sigma_{22} = \nu\nu' \quad \text{and} \quad \Sigma_{12} = \mu D_{\gamma^{1/2}}\nu',$$

where γ's, i.e., $\gamma_1, \gamma_2, \cdots, \gamma_p$ are the characteristic roots of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'$ and $D_{\gamma^{1/2}}$ is a diagonal matrix whose diagonal elements are $\gamma_1^{\frac{1}{2}}, \cdots, \gamma_p^{\frac{1}{2}}$. 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 $\mathbf{x}^*$ $(2p \times 1)$ defined by

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

where

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

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

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

whence we have that

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

This means that the transformed p-set $\mathbf{x}_1^*$ is uncorrelated with transformed p-set $\mathbf{x}_2^*$. 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 $\mathbf{x}^*$, we have

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

whence we have

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

Now we go back to (3.6). Note that, since $\Sigma_{12}^* = 0$, the transformed $\mathbf{x}_1^*$-set is uncorrelated with the transformed $\mathbf{x}_2^*$-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

$$P[c_{\max}(S_{11}^{*-1}S_{12}^*\, S_{22}^{*-1}S_{12}^{*\prime}) \leqq c_\alpha \mid \Sigma_{12}^* = 0] = 1 - \alpha. \tag{3.9}$$

The set over which the probability statement (3.9) is made, namely,

$$c_{\max}(S_{11}^{*-1}S_{12}^*\, S_{22}^{*-1}S_{12}^{*\prime}) \leqq c_\alpha,$$

can be used as an acceptance region for the hypothesis that $\mu\nu^{-1}$ has a particular (matrix) value, and, in particular, that $\mu\nu^{-1} = I(p)$, or in other words, $\Sigma_{11} = \Sigma_{22}$. 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

$$| cS_{11}^* - S_{12}^*\, S_{22}^{*-1}S_{12}^{*\prime} | = 0. \tag{3.10}$$

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

$$| dS_{11}^* - \tfrac{1}{4}S_{11}^* + \tfrac{1}{4}S_{12}^*\, S_{22}^{*-1}S_{12}^{*\prime} | = 0. \tag{3.11}$$

Now, using (3.8), we have

$$\begin{aligned} -\tfrac{1}{4}S_{11}^* &= -S_{11} + \tfrac{1}{4}(S_{12}^* + S_{12}^{*\prime} + S_{22}^*) \\ &= -S_{11} + \tfrac{1}{4}(S_{12}^* + S_{22}^*)S_{22}^{*-1}(S_{12}^{*\prime} + S_{22}^*) - \tfrac{1}{4}S_{12}^*\, S_{22}^{*-1}S_{12}^{*\prime}. \end{aligned} \tag{3.12}$$

Hence

$$\left| dS_{11}^{*} - S_{11} + \left(\frac{S_{12}^{*} + S_{22}^{*}}{2}\right) S_{22}^{*-1} \left(\frac{S_{12}^{*\prime} + S_{22}^{*}}{2}\right) \right| = 0 \tag{3.13}$$

or

$$| \, dS_{11}^{*} - S_{11} + (S_{11} + S_{12}\lambda')S_{22}^{*-1}(S_{11} + \lambda S_{12}') \, | = 0.$$

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

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

and, using this, we observe that (3.13) is equivalent to

$$\begin{vmatrix} S_{11} - dS_{11}^{*} & S_{11} + S_{12}\lambda' \\ S_{11} + \lambda S_{12}' & S_{11} + \lambda S_{12}' + S_{12}\lambda' + \lambda S_{22}\lambda' \end{vmatrix} = 0, \tag{3.15}$$

that is,

$$\begin{vmatrix} S_{11} - dS_{11}^{*} & S_{12}\lambda' + dS_{11}^{*} \\ S_{11} + \lambda S_{12}' & S_{12}\lambda' + \lambda S_{22}\lambda' \end{vmatrix} = 0,$$

that is,

$$\begin{vmatrix} S_{11} - dS_{11}^{*} & S_{12}\lambda' + dS_{11}^{*} \\ \lambda S_{12}' + dS_{11}^{*} & \lambda S_{22}\lambda' - dS_{11}^{*} \end{vmatrix} = 0,$$

that is,

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

But we have

$$\begin{bmatrix} S_{11} & S_{12}\lambda' \\ \lambda S_{12}' & \lambda S_{22}\lambda' \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \lambda' \end{bmatrix} \tag{3.16}$$

and

$$\begin{bmatrix} S_{11}^{*} & -S_{11}^{*} \\ -S_{11}^{*} & S_{11}^{*} \end{bmatrix} = \begin{bmatrix} I \\ -I \end{bmatrix} S_{11}^{*} \, [I \quad -I]$$
$$= \begin{bmatrix} I \\ -I \end{bmatrix} [I \quad -\lambda] \begin{bmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{bmatrix} \begin{bmatrix} I \\ -\lambda' \end{bmatrix} [I \quad -I].$$

Hence (3.15) reduces to

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

which is equivalent to

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

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

$$(3.18)\qquad | eI(2p \times 2p) - S^{-1} \beta S \beta' | = 0,$$

where

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

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

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

or to

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

Next, consider, for all non null **a** and **b**, the statement

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

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

We next set

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

whence we have

$$(3.23)\qquad \begin{aligned} S^{11} &= (S_{11} - S_{12} S_{22}^{-1} S_{12}')^{-1}, \qquad S^{22} = (S_{22} - S_{12}' S_{11}^{-1} S_{12})^{-1}, \\ S^{12} &= -S^{11} S_{12} S_{22}^{-1} = -S_{11}^{-1} S_{12} S^{22}. \end{aligned}$$

Back in (3.21) we now observe that (3.21) implies

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

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

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

for all non null $\mathbf{a}_2$ and $\mathbf{b}_1$. If now we consider the left side of (3.24), then it follows from Cauchy's inequality that for all non null $\mathbf{b}_2$, $(\mathbf{a}_1'\lambda\mathbf{b}_2)^2/(\mathbf{a}_1'\mathbf{a}_1)(\mathbf{b}_2'\mathbf{b}_2) \leqq (\mathbf{a}_1'\lambda\lambda'\mathbf{a}_1)/(\mathbf{a}_1'\mathbf{a}_1)$, and it is also well known that for all non null $\mathbf{a}_1$, $c_{\min}(\lambda\lambda') \leqq (\mathbf{a}_1'\lambda\lambda'\mathbf{a}_1)/(\mathbf{a}_1'\mathbf{a}_1) \leqq c_{\max}(\lambda\lambda')$. We have also exactly similar results by interchanging $\mathbf{a}_1$ and $\mathbf{b}_2$, and similar results on the left side of (3.25), in terms of λ^{-1} and $\mathbf{a}_2$ and $\mathbf{b}_1$ and then again by the interchange of $\mathbf{a}_2$ and $\mathbf{b}_1$.

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

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

or, after substitution from (3.23),

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

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

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

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 nonsingular, 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) imply

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

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

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

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

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

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

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

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

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

Now combining (3.26), (3.28), (3.29)–(3.32), we observe that (3.21) implies all these statements, and hence, going back to (3.20), we have with a joint probability $\geqq 1 - \alpha$, the bounds

$$(3.33) \qquad \begin{aligned} &\frac{1}{e_\alpha} c_{\min}(S_{11} - S_{12} S_{22}^{-1} S_{12}') \Big/ c_{\max}(S_{22}) \leqq c_{\min}(\lambda\lambda') \\ &\leqq e_\alpha \min\,[c_{\min}(S_{11})/c_{\min}(S_{22} - S_{12}' S_{11}^{-1} S_{12}),\, c_{\max}(S_{11})/c_{\max}(S_{22} - S_{12}' S_{22}^{-1} S_{12})] \end{aligned}$$

and

$$(3.34) \qquad \begin{aligned} &\frac{1}{e_\alpha} \max\,[c_{\min}(S_{11} - S_{12} S_{22}^{-1} S_{12}')/c_{\min}(S_{22}),\, c_{\max}(S_{11} - S_{12} S_{22}^{-1} S_{12}')/c_{\max}(S_{22})] \\ &\leqq c_{\max}(\lambda\lambda') \leqq e_\alpha c_{\max}(S_{11})/c_{\min}(S_{22} - S_{12}' S_{11}^{-1} S_{12}). \end{aligned}$$

It is interesting to use [3] and check that the lower bound of (3.33) is $\leqq$ the upper bound of (3.34), but that the upper bound of (3.33) might be $\geqq$ or $<$ the lower bound of (3.34). However, it is to be always remembered that $c_{\min}(\lambda\lambda') \leqq c_{\max}(\lambda\lambda')$, 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 $\mathbf{a}_1$ and the *corresponding* elements of $\mathbf{b}_2$, $\mathbf{a}_2$, and $\mathbf{b}_1$ (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 $\mathbf{a}_1$ and the corresponding elements of $\mathbf{b}_2$, $\mathbf{a}_2$ and $\mathbf{b}_1$, 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 $1 + \binom{p}{1} + \binom{p}{2} + \cdots + \binom{p}{p-1}$, i.e., $2^p - 1$ pairs of statements like (and including) (3.33) and (3.34) with a joint probability $\geqq 1 - \alpha$. It should be noticed that on all these statements e_α, however, stays the same.

It follows from the above remarks that, with a joint confidence coefficient $\geqq 1 - \alpha$, (3.33) and (3.36) imply, among other things, the following set of confidence statements on the ratios $\sigma_{1i}^2/\sigma_{2i}^2$:

$$(3.34.1) \qquad \frac{1}{e_\alpha}\frac{s_{1i}^2}{s_{2i}^2}(1 - r_i^2) \leqq \frac{\sigma_{1i}^2}{\sigma_{2i}^2} \leqq e_\alpha \frac{s_{1i}^2}{s_{2i}^2(1 - r_i^2)} \qquad \text{for } i = 1, 2, \cdots, p,$$

where s_{1i}^2, s_{2i}^2, σ_{1i}^2, σ_{2i}^2 and r_i stand respectively for the sample variances of the ith variate for the two sets, the population variances of the ith variate for the two sets and the sample correlation coefficient between the ith variate for the first set and for the second set.

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

$$(3.35) \qquad \mu\nu^{-1} = A, \quad \text{i.e.,} \quad \mu = A\nu,$$

where A is any arbitrary orthogonal matrix. Going back to (3.2), we easily check that (3.35) implies

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

which, if we recall that A is orthogonal, 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 implies

$$(3.37) \qquad \Sigma_{12} = \mu D_{\gamma^{1/2}}\nu' = A\nu D_{\gamma^{1/2}}\nu' = A \times \text{a symmetric matrix},$$

where A is the same orthogonal matrix that occurs in (3.36). Thus (3.35) implies (3.36) and (3.37) and it is also easy to verify that (3.36) and (3.37) imply (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 hence (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 $\sigma_z^2 = \sigma_1^2 - 2q\rho\sigma_1\sigma_2 + q^2\sigma_2^2$. Thus

$$\sqrt{n}\,\bar{z}/s_z = \sqrt{n}(\bar{x}_1 - q\bar{x}_2)/(s_1^2 - 2qs_1s_2r + q^2s_2^2)^{\frac{1}{2}}$$

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

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

or

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

where $k = (1/n)t_{\alpha/2}^2$. We can use the statement within the parentheses in (4.1) as an acceptance region for the hypothesis that the population ratio of means has a specific value q. But such an acceptance is, of course, well known, at least in an implicit form.

Subject to the restriction that q is to have real values, the statement within the parentheses 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 actual confidence bounds on q are given by

$$(4.2)\quad \begin{aligned} &\frac{(\bar{x}_1\bar{x}_2 - ks_1s_2r) - [(\bar{x}_1\bar{x}_2 - ks_1s_2r)^2 - (\bar{x}_1^2 - ks_1^2)(\bar{x}_2^2 - ks_2^2)]^{\frac{1}{2}}}{(\bar{x}_2^2 - ks_2^2)} \leqq q \\ &\leqq \frac{(\bar{x}_1\bar{x}_2 - ks_1s_2r) + [(\bar{x}_1\bar{x}_2 - ks_1s_2r)^2 - (\bar{x}_1^2 - ks_1^2)(\bar{x}_2^2 - ks_2^2)]^{\frac{1}{2}}}{(\bar{x}_2^2 - ks_2^2)}. \end{aligned}$$

The bounds will be physically meaningful only if the expression under the radical is non-negative, i.e., only if,

$$(4.3)\quad \frac{\bar{x}_1^2}{s_1^2} + \frac{\bar{x}_2^2}{s_2^2} \geqq 2\frac{\bar{x}_1}{s_1}\cdot\frac{\bar{x}_2}{s_2}r + k\cdot\frac{\bar{x}_1^2}{s_1^2}\cdot\frac{\bar{x}_2^2}{s_2^2}(1 - r^2).$$

Notice that $(\bar{x}_1^2/s_1^2) + (\bar{x}_2^2/s_2^2)$ is always greater than or equal to $2(\bar{x}_1/s_1)(\bar{x}_2/s_2)r$ but may not always be greater than or equal to the right side of (4.3). Thus, if in the sample, the inequality (4.3) breaks down we should not, in that situation, attempt to put any confidence bounds on ξ_1/ξ_2.

Going back to (4.1) and tying it up with (4.2) and (4.3) we now observe that α is the probability of choosing a sample such that either (4.2) is not a real interval or (4.2) is real but does not cover the true value.

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

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

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

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

we have sample dispersion matrix of $\mathbf{z}(p \times 1)$ given by

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

Also the sample mean vector of $\mathbf{z}(p \times 1)$ is given by

$$\bar{\mathbf{z}} = \bar{\mathbf{x}}_1 - D_q \bar{\mathbf{x}}_2 . \tag{5.4}$$

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

$$P\left[\bar{\mathbf{z}}' S_z^{-1} \bar{\mathbf{z}} \leqq \frac{1}{n} T_\alpha^2 \mid q\text{'s defined as above}\right] = 1 - \alpha. \tag{5.5}$$

The set over which the probability statement (5.5) is made, can be used as an acceptance region for the hypothesis that the population mean ratios have specific values q_i's. This, of course, is implicit in the possible applications of Hotelling's T^2. Now consider the statement within the parentheses in (5.5). It is well known that this statement is equivalent to the statement that all $c[\bar{\mathbf{z}}\bar{\mathbf{z}}' S_z^{-1}]' \leqq T_\alpha^2/n$, which again is equivalent to

$$\frac{\mathbf{a}'\bar{\mathbf{z}}\bar{\mathbf{z}}'\mathbf{a}}{\mathbf{a}'\mathbf{a}} \leqq \frac{T_\alpha^2}{n} \cdot \frac{\mathbf{a}' S_z \mathbf{a}}{\mathbf{a}'\mathbf{a}}, \tag{5.6}$$

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

$$\begin{aligned} c_{\max}(SA'A) &\leqq c_{\max}(S) c_{\max}(A'A), \text{ i.e., } \leqq c_{\max}(S) c_{\max}(AA'), \\ &\text{i.e., } \leqq c_{\max}(S) c_{\max}[I + D_{q^2}], \\ &\text{i.e., } \leqq c_{\max}(S) \max[1 + q_1^2, 1 + q_2^2, \cdots, 1 + q_p^2]. \end{aligned} \tag{5.7}$$

Now, if we go back to (5.6) and maximize the left side w.r.t. $\mathbf{a}$, it is easy to check that (5.6) implies

$$\frac{1}{n} T_\alpha^2 c_{\max}(S) \max[1 + q_1^2, 1 + q_2^2, \cdots, 1 + q_p^2] - \sum_{i=1}^{p} (\bar{x}_{1i} - q_i \bar{x}_{2i})^2 \geqq 0$$

or

$$\frac{1}{n} T_\alpha^2 c_{\max}(S) \max[1 + q_1^2, \cdots, 1 + q_p^2] - \bar{\mathbf{x}}_1'\bar{\mathbf{x}}_1 - \sum_{i=1}^{p} q_i^2 \bar{x}_{2i}^2 + 2 \sum_{i=1}^{p} q_i \bar{x}_{1i} \bar{x}_{2i} \geqq 0. \tag{5.8}$$

Also notice that

$$\left|\sum_{i=1}^{p} q_1\, \bar{x}_{1i}\, \bar{x}_{2i}\right| \leqq \sum_{i=1}^{p} \left| q_i \right| \left| \bar{x}_{1i}\, \bar{x}_{2i} \right| \tag{5.9}$$
$$\leqq [\max(q_1^2, \cdots, q_p^2)]^{\frac{1}{2}} \sum_{i=1}^{p} \left| \bar{x}_{1i}\, \bar{x}_{2i} \right|,$$

and

$$-\sum_{i=1}^{p} q_i^2\, \bar{x}_{2i}^2 \leqq -\min(q_1^2, \cdots, q_p^2) \sum_{i=1}^{p} \bar{x}_{2i}^2.$$

Hence it is easy to check that (5.8) and hence (5.6) imply

$$\frac{1}{n} T_\alpha^2\, c_{\max}(S) \max[1 + q_1^2, \cdots, 1 + q_p^2]$$
$$+ 2 \sum_{i=1}^{p} \left| \bar{x}_{1i}\, \bar{x}_{2i} \right| \max[q_1^2, \cdots, q_p^2]^{\frac{1}{2}} \tag{5.10}$$
$$- \bar{\mathbf{x}}_1' \bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2' \bar{\mathbf{x}}_2 \min(q_1^2, \cdots, q_p^2) \geqq 0.$$

Going back to (5.5) we now observe that with a probability $\geqq 1 - \alpha$, we have the confidence statement (5.8) or (5.10).

Truncation. Here again, as in section 4, it is possible to go back 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 $> 1 - \alpha$.

If we are interested in pairwise comparisons we go back to (5.6), set $k = T_2^2/n$ and choose $\mathbf{a}$ to be the vector with 1 in the ith position and 0's elsewhere. The resulting inequality can be written as (4.2) (with $k = T_\alpha^2/n$). Thus (5.6) implies a set of inequalities like this for $i = 1, 2, \cdots, p$, and hence, with a confidence coefficient greater than or equal to a preassigned $1 - \alpha$, we have the set of confidence bounds on ξ_{1i}/ξ_{2i} given by

$$(e_{1i} - e_{2i}^{\frac{1}{2}})/e_{3i} \leqq q_i = \xi_{1i}/\xi_{2i} \leqq (e_{1i} + e_{2i}^{\frac{1}{2}})/e_{3i}, \tag{5.11}$$

where, for $i = 1, 2, \cdots, p$,

$$e_{1i} = \bar{x}_{1i}\bar{x}_{2i} - ks_{1i}s_{2i}r_{12i}, \qquad e_{3i} = \bar{x}_{2i}^2 - ks_{2i}^2, \tag{5.12}$$
$$e_{2i} = (\bar{x}_{1i}\bar{x}_{2i} - ks_{1i}s_{2i}r_{12i})^2 - (\bar{x}_{1i}^2 - ks_{1i}^2)(\bar{x}_{2i}^2 - ks_{2i}^2).$$

As in section 4, the bounds will be physically meaningful only if

$$\frac{\bar{x}_{1i}^2}{s_{1i}^2} + \frac{\bar{x}_{2i}^2}{s_{2i}^2} \geqq 2\, \frac{\bar{x}_{1i}\, \bar{x}_{2i}}{s_{1i}\, s_{2i}}\, r_{12i} + k\, \frac{\bar{x}_{1i}^2\, \bar{x}_{2i}^2}{s_{1i}^2\, s_{2i}^2}\, (1 - r_{12i}^2). \tag{5.13}$$

As in section 4 so also here, the remarks made after (4.3) will be pertinent again as an indication of how to use these bounds.

In conclusion it is a great pleasure to thank the referee and the associate editor for their valuable comments and suggestions. The result (5.11), in particular, is entirely due to the referee and provides shorter bounds than the ones obtained by the authors' originally, starting from (5.10) rather than directly from (5.6).

REFERENCES

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

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

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

[4] S. N. Roy and R. Gnanadesikan, "Further contributions to multivariate confidence bounds," Institute of Statistics mimeograph series no. 1551. (Also *Biometrika*, Vol. 44 (1957), pp. 399-410.

[5] Helen Walker and Joseph Lev, "Statistical Inference," Henry Holt and Company, New York, 1953.