TESTS OF MULTIPLE INDEPENDENCE AND THE ASSOCIATED CONFIDENCE BOUNDS[1]

By S. N. Roy and R. E. Bargmann

University of North Carolina

1. Summary. In this paper a test based on the union-intersection principle is proposed for overall independence between p variates distributed according to the multivariate normal law, and this is extended to the hypothesis of independence between several groups of variates which have a joint multivariate normal distribution. Methods used in earlier papers [3, 4] have been applied in order to invert these tests for each situation, and to obtain, with a joint confidence coefficient greater than or equal to a preassigned level, simultaneous confidence bounds on certain parametric functions. These parametric functions are, in case I, the moduli of the regression vectors: (a) of the variate p on the variates $(p - 1)$, $(p - 2)$, $\cdots$, 2, 1, or on any subset of the latter; (b) of the variate $(p - 1)$ on the variates $(p - 2)$, $(p - 3)$, $\cdots$, 2, 1, or any subset of the latter, etc.; and finally, (c) of the variate 2 on the variate 1. For case II, parallel to each case considered above, there is an analogous statement in which the regression vector is replaced by a regression matrix, β, say, and the "modulus" of the regression vector is replaced by the (positive) square-root of the largest characteristic root of $(\beta\beta')$. Simultaneous confidence bounds on these sets of parameters are given. As far as the proposed tests of hypotheses of multiple independence are concerned they are offered as an alternative to another class of tests based on the likelihood-ratio criterion [5, 6] which has been known for a long time. So far as the confidence bounds are concerned it is believed, however, that no other easily obtainable confidence bounds are available in this area. One of the objects of these confidence bounds is the detection of the "culprit variates" in the case of rejection of the hypothesis of multiple independence, for the "complex" hypothesis is, in this case, the intersection of several more "elementary" hypotheses of two-by-two independence.

2. Introduction, notation, and preliminaries. Case I, which deals with the question of independence among p normally distributed variates, represents a well known situation which has occurred repeatedly in applications. For case II, which deals with the question of independence between k sets of normally distributed variates (where each set contains one or more variates), a number of potential applications has been described by *Wilks* [6]. In addition to the situations mentioned by *Wilks*, an interesting application concerns the problem of "unreliable measurement". If we consider the p_i variates x_i $(i = 1, 2, \cdots, k)$ as

Received June 13, 1957.

[1] 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.

different measurements of a physically identical quantity which, because of the inaccuracy of the measuring instrument, are not in perfect correlation, the procedures outlined in sections 5 and 6 will study the independence or dependence of the k underlying "true" physical quantities. The "correction-for-attenuation" technique which is widely used in this situation has two serious drawbacks which the present method tries to overcome: (1) It assumes equal error variance for each fallible measurement, and (2) it makes use of a statistic, the "correlation corrected for attenuation", which is not a correlation coefficient, for it may attain values greater than unity. The present method is free from these shortcomings. The confidence bounds discussed in section 6 will then give an indication of the maximum attainable degree of prediction if infallible measurements could be made.

Suppose we have a random sample of size $n + 1$ from an $N[\zeta(p \times 1), \Sigma(p \times p)]$, with $p \leqq n$. Then, denoting by $S(p \times p)$ the sample dispersion matrix, we know that S is symmetric and everywhere at least p.s.d., (and also p.d., a.e.). It is also well known that, a.e., there exists a one-to-one transformation from $S(p \times p)$ to $\tilde{T}(p \times p)$ given by $nS = \tilde{T}\tilde{T}'$, where $\tilde{T}$ is a lower triangular matrix with positive diagonal elements. Let t_{ij} $(i \geqq j = 1, 2, \cdots, p)$ denote the elements of $\tilde{T}$, s_{ij} and σ_{ij} $(s_{ij} = s_{ji}, \sigma_{ij} = \sigma_{ji}, i, j = 1, 2, \cdots, p)$ denote the elements of S and Σ, and let s^{ij} and σ^{ij} denote the elements of S^{-1} and Σ^{-1}. Furthermore, let $r_{p \cdot 1,2,\cdots,(p-1)}$, $r_{(p-1)\cdot 1,2,\cdots,(p-2)}$, $\cdots$, $r_{3\cdot 1,2}$, and $r_{2\cdot 1}$ denote, respectively, the multiple correlation coefficient of (p) with $(1, 2, \cdots, p-1)$, of $(p-1)$ with $(1, 2, \cdots, p-2)$, and so on, and finally the simple correlation coefficient of (2) with (1). It may be noted that all except the last are non-negative and a.e. positive. These multiple correlation coefficients will be called the *step-down correlations*. Likewise, let

$$(2.1)\quad \boldsymbol{\beta}'_{i\cdot 1,2,\cdots,i-1}(1 \times \overline{i-1}) = [\sigma_{1i}\sigma_{2i}\cdots\sigma_{i-1,i}]\begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1,i-1} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ \sigma_{1,i-1} & \sigma_{2,i-1} & \cdots & \sigma_{i-1,i-1} \end{bmatrix}^{-1}$$

(for $i = p, p-1, \cdots, 2$) denote the population regression vector of (i) on $(1, 2, \cdots, i-1)$ and

$$(2.2)\quad \mathbf{b}'_{i\cdot 1,2\cdots i-1}(1 \times \overline{i-1}) = [s_{1i}s_{2i}\cdots s_{i-1,i}]\begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1,i-1} \\ s_{12} & s_{22} & \cdots & s_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ s_{1,i-1} & s_{2,i-1} & \cdots & s_{i-1,i-1} \end{bmatrix}^{-1}$$

(for $i = p, p-1, \cdots, 2$) denote the corresponding sample regression vector. These regression vectors will be called the *step-down regression vectors*.

Next, we will present the expression for the multiple correlation coefficient by treating it as a special case of a canonical correlation (which will be convenient for later purposes). We have

$$(2.3)\qquad r^2_{i\cdot 1,2,\cdots,i-1} = s_{ii}^{-1}[s_{1i}\cdots s_{i-1,i}]\begin{bmatrix} s_{11} & \cdots & s_{1,i-1}\\ \cdot & \cdots & \cdot \\ s_{1,i-1} & \cdots & s_{i-1,i-1}\end{bmatrix}^{-1}\begin{bmatrix} s_{1i}\\ \cdot \\ s_{i-1,i}\end{bmatrix}$$

for $i = p, p-1, \cdots, 2$. Next we have, by using $nS = \tilde{T}\tilde{T}'$,

$$ns_{ii} = [t_{i1}\cdots t_{ii}][t_{i1}\cdots t_{ii}]' = \sum_{j=1}^{i} t_{ij}^2,$$

$$(2.3.1)\qquad n[s_{1i}\cdots s_{i-1,i}] = [t_{i1}\cdots t_{i,i-1}]\begin{bmatrix} t_{11} & 0 & \cdots & 0\\ t_{21} & t_{22} & \cdots & 0\\ \cdot & \cdot & \cdots & \cdot\\ t_{i-1,1} & \cdot & \cdots & t_{i-1,i-1}\end{bmatrix}'$$

and

$$n\begin{bmatrix} s_{11} & \cdots & s_{1,i-1}\\ \cdot & \cdots & \cdot \\ s_{1,i-1} & \cdots & s_{i-1,i-1}\end{bmatrix} = \begin{bmatrix} t_{11} & 0 & \cdots & 0\\ \cdot & \cdot & \cdots & \cdot\\ t_{i-1,1} & \cdot & \cdots & t_{i-1,i-1}\end{bmatrix}\begin{bmatrix} t_{11} & 0 & \cdots & 0\\ \cdot & \cdot & \cdots & \cdot\\ t_{i-1,1} & \cdot & \cdots & t_{i-1,i-1}\end{bmatrix}'$$

It is then easy to check, by substituting the expressions (2.3.1) into (2.3), that

$$(2.4)\qquad r_{i\cdot 1,2,\cdots,i-1} = +\left[\sum_{j=1}^{i-1} t_{ij}^2 \Big/ \sum_{j=1}^{i} t_{ij}^2\right]^{1/2},$$

for $i = p, p-1, \cdots, 2$.

Now, let us turn to the case of a $(p_1 + p_2 + \cdots + p_k)$-variate $(= p$-variate, say) normal distribution and partition the population dispersion matrix, $\sum$, into

$$(2.5)\qquad \sum(p\times p) = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1k}\\ \Sigma_{12}' & \Sigma_{22} & \cdots & \Sigma_{2k}\\ \cdot & \cdot & \cdots & \cdot\\ \Sigma_{1k}' & \Sigma_{2k}' & \cdots & \Sigma_{kk}\end{bmatrix}\begin{matrix}(p_1)\\(p_2)\\ \\(p_k)\end{matrix}$$
$$\qquad\qquad (p_1)\quad (p_2)\qquad\quad (p_k)$$

and the sample dispersion matrix, S, into

$$(2.6)\qquad S(p\times p) = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1k}\\ S_{12}' & S_{22} & \cdots & S_{2k}\\ \cdot & \cdot & \cdots & \cdot\\ S_{1k}' & S_{2k}' & \cdots & S_{kk}\end{bmatrix}\begin{matrix}(p_1)\\(p_2)\\ \\(p_k).\end{matrix}$$
$$\qquad\qquad (p_1)\quad (p_2)\qquad\quad (p_k)$$

Regarding each submatrix as an element let us say that there are k "pseudo-rows" and k "pseudo-columns" in the matrices on the right sides of (2.5) and (2.6).

Let $\beta_{i\cdot 1,2,\cdots,i-1}$ and $B_{i\cdot 1,2,\cdots,i-1}$ (for $i = k, k-1, \cdots, 2$) denote the population and the sample regression matrix of the (p_i)-set on the $(p_{i-1} + p_{i-2} + \cdots$

$+ p_1)$-set. These matrices are given by the expressions:

$$\beta_{i\cdot 1,2,\cdots,i-1}(p_i \times \overline{p_{i-1} + p_{i-2} + \cdots + p_1})$$

(2.7)
$$= [\Sigma'_{1i}\Sigma'_{2i} \cdots \Sigma'_{i-1,i}] \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1,i-1} \\ \Sigma'_{12} & \Sigma_{22} & \cdots & \Sigma_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ \Sigma'_{1,i-1} & \Sigma'_{2,i-1} & \cdots & \Sigma_{i-1,i-1} \end{bmatrix}^{-1}$$

and

$$B_{i\cdot 1,2,\cdots i-1}(p_i \times \overline{p_{i-1} + p_{i-2} + \cdots + p_1})$$

(2.8)
$$= [S'_{1i}S'_{2i} \cdots S'_{i-1,i}] \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1,i-1} \\ S'_{12} & S_{22} & \cdots & S_{2,i-1} \\ \cdot & \cdot & \cdots & \cdot \\ S'_{i,i-1} & S'_{2,i-1} & \cdots & S_{i-1,i-1} \end{bmatrix}^{-1},$$

for $i = k, k - 1, \cdots, 2$. The β's and B's will be called the *step-down regression matrices.* Also let us denote by

(2.8.1)
$$0 \leqq c^{(1)}_{i\cdot 1,2,\cdots,i-1} \leqq c^{(2)}_{i\cdot 1,2,\cdots,i-1} \leqq \cdots \leqq c^{(p_i)}_{i\cdot 1,2,\cdots,i-1} \leqq 1$$

the p_i characteristic roots of the matrix

(2.8.2)
$$S^{-1}_{ii}[S'_{1i}S'_{2i} \cdots S'_{i-1,i}] \begin{bmatrix} S_{11} & \cdots & S_{1,i-1} \\ \cdot & \cdots & \cdot \\ S'_{1,i-1} & \cdots & S_{i-1,i-1} \end{bmatrix}^{-1} \begin{bmatrix} S_{1i} \\ \cdot \\ S_{i-1,i} \end{bmatrix},$$

for $i = k, k - 1, \cdots, 2$. It will be noticed that these c's are the squares of the canonical correlation coefficients of the (p_i)-set with the $(p_1 + p_2 + \cdots + p_{i-1})$-set and that, a.e., the inequalities in (2.8.1) will be strict. For any (p_i)-set, all the p_i canonical correlation coefficients (or rather, as will be seen later, the largest of them) will play the same role as $r_{i\cdot 1,2,\cdots,i-1}$ in the previous case. These will be called the *step-down sets of canonical correlations.* We are assuming here for simplicity of discussion, but without loss of generality, that the sets are so numbered as to make $\sum_{j=1}^{i-1} p_j \geqq p_i$, (for $i = 2, 3, \cdots, k$). The matrix corresponding to $\tilde{T}$ in the previous case will be introduced in a later section.

Sections 3 and 4 will be concerned with the first case, i.e., the case of a p-variate normal distribution; section 3 will describe a test of the hypothesis $H_0: \sigma_{ij} = 0$ $(i \neq j = 1, 2, \cdots, p)$, and section 4 will present simultaneous confidence bounds, for $i = p, p - 1, \cdots, 2$, on $(\beta'_{i\cdot 1,2,\cdots,i-1}\beta_{i\cdot 1,2\cdots,i-1})^{1/2}$ (and on truncations obtained by deleting any $1, 2, \cdots, (i - 2)$ variates of the $(i - 1)$-set).

Sections 5 and 6 will be concerned with the second case, i.e., that of a $(p_1 + p_2 + \cdots + p_k)$-variate normal distribution; section 5 will describe a test of the hypothesis $H_0: \sum_{ij} = 0$ $(i \neq j = 1, 2, \cdots, k)$, and section 6 will present simultaneous confidence bounds, for $i = k, k - 1, \cdots, 2$, on the largest characteristic root of $(\beta_{i\cdot 1,2,\cdots,i-1}\beta'_{i\cdot 1,2,\cdots,i-1})$ (and on truncations obtained by deleting any

1, 2, $\cdots$, $(i-2)$ of the sets (p_1), (p_2), $\cdots$, (p_{i-1})). It will be noted that, in case I, the variate (i) is independent of the variates 1, 2, $\cdots$, $i-1$ if and only if $\beta'_{i\cdot 1,2,\cdots,i-1}\beta_{i\cdot 1,2,\cdots,i-1} = 0$, and, in case II, independence of the (p_i)-set on the $[(p_1), (p_2), \cdots, (p_{i-1})]$-set implies, and is implied by, the vanishing of the largest characteristic root of $(\beta_{i\cdot 1,2,\cdots,i-1}\beta'_{i\cdot 1,2,\cdots,i-1})$.

3. Independence in the p-variate problem.

3.1. Independence (in distribution) of the step-down correlations, under the null hypothesis. The joint distribution of the t_{ij}'s, for general $\sum$, is well known and given by

$$\text{const} \cdot \exp\left[-\frac{1}{2}\, tr \textstyle\sum^{-1} \tilde{T}\tilde{T}'\right] \prod_{i=1}^{p} t_{ii}^{n-i} \prod_{i \geq j=1}^{p} dt_{ij}. \tag{3.1.1}$$

Among various proofs, a recent one is given in [1, 2]. Under the null hypothesis we have $\sum = D_{\sigma_{ii}}$, where the right side denotes a diagonal matrix with elements σ_{11}, σ_{22}, $\cdots$, σ_{pp}. In this situation, (3.1.1) reduces to

$$\text{const} \cdot \exp\left[-\frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{i} t_{ij}^2 \Big/ \sigma_{ii}\right] \prod_{i=1}^{p} t_{ii}^{n-i} \prod_{i=1}^{p}\prod_{j=1}^{i} dt_{ij}. \tag{3.1.2}$$

It should be noticed that the t_{ii}'s vary from 0 to ∞, and the t_{ij}'s from $-\infty$ to $+\infty$. Now a comparison with the expression (2.4) shows immediately that the $r_{i\cdot 1,2,\cdots,i-1}$'s $(i = p, \cdots, 2)$ are independently distributed, and their joint distribution is given by

$$\text{const} \cdot \prod_{i=p}^{2} (r^2_{i\cdot 1,2,\cdots,i-1})^{(i-3)/2}(1 - r^2_{i\cdot 1,2,\cdots,i-1})^{(n-i-1)/2} \times d(r^2_{i\cdot 1,2,\cdots,i-1}). \tag{3.1.3}$$

3.2. The proposed test and a reason for the allocation of component probabilities. The proposed test is as follows:

$$\begin{aligned} &\text{Accept } H_0 \text{ over } \bigcap_{i=2}^{p} [r^2_{i\cdot 1,2,\cdots,i-1} \leqq \mu], \quad \text{and} \\ &\text{reject } H_0 \text{ over } \bigcup_{i=2}^{p} [r^2_{i\cdot 1,2,\cdots,i-1} > \mu], \end{aligned} \tag{3.2.1}$$

where μ is given by

$$\prod_{i=2}^{p} P[r^2_{i\cdot 1,2,\cdots,i-1} \leqq \mu \mid \rho^2_{i\cdot 1,2,\cdots i-1} = 0] = 1 - \alpha. \tag{3.2.2}$$

To obtain μ, proceed as follows: Take a trial value, μ_1, say, between zero and one. Using this value for given n and for $i = 2, 3, \cdots, p$ obtain, from the *Tables of the Incomplete Beta Function*, the probabilities corresponding to the individual factors on the left side of (3.2.2); call these probabilities γ_2, γ_3, $\cdots$, γ_p, say; form the product of the γ_i's and denote it by γ. Proceed in the same manner for other trial values, μ_2, μ_3, etc., and plot the μ_i's against the resulting γ's. Then, on this plot, μ in (3.2.2) is that value of the μ_i's which corresponds to $\gamma = 1 - \alpha$, the preassigned confidence level.

It is obvious that, if the same μ goes with the different component regions $[r^2_{i\cdot 1,2,\cdots,i-1} \leqq \mu]$, the probability measures that go with these regions are all different. One reason why we make this kind of allocation of the different γ_i's is the following: Notice that the acceptance region for H_0 is the intersection (over $i = p, p-1, \cdots, 2$) of regions $[r^2_{i\cdot 1,2,\cdots,i-1} \leqq \mu]$; but, for a given i, $[r^2_{i\cdot 1,2,\cdots,i-1} \leqq \mu]$ is itself the intersection of regions of the type $[r^2_{i\cdot L(1,2,\cdots,i-1)} \leqq \mu]$, where $r_{i\cdot L(1,2,\cdots,i-1)}$ denotes the (simple) correlation of the ith variate with any linear combination of the variates $(1, 2, \cdots, i-1)$ which includes, as a special case, the variates $(1, 2, \cdots, i-1)$ *individually*. Thus, if we allocate the γ_i's in such a way as to make μ the same for each component region, we attach the same weight not only to the correlations between any pair of the observed variates but also to the correlations between each variate and linear combinations of some others. The reader will perceive that this allocation is not completely symmetric. While symmetry is preserved with respect to all correlations by pairs, the step-down procedure is asymmetric as regards the correlation of any variate with *any* linear combination of all the other variates. However, this is perhaps the best that could be done under this particular approach. It should be noted that, if the square of any simple correlation in the correlation matrix exceeds the value μ, we will have to reject the hypothesis of independence. If, however, the square of the largest correlation coefficient in the correlation matrix stays below μ, we will have to perform the step-down process in order to decide on acceptance or rejection of the hypothesis of independence.

3.3. Relation to the likelihood-ratio test. Since the determinant of R, the correlation matrix, equals $(1 - r^2_{p\cdot 1,2,\cdots,p-1}) \times (1 - r^2_{p-1\cdot 1,2,\cdots,p-2}) \cdots (1 - r^2_{2\cdot 1})$, a test based on the product of the complements of the squares of the step-down correlations is equivalent to the likelihood-ratio test. While the distribution of the determinant of R is fairly complicated, even under the null hypothesis of independence [7], its moments [6] are well known and easily obtained from the joint distribution of the correlation coefficients under the hypothesis of independence. It can be easily verified that they satisfy the recurrence relations:

$$\mu_1' = \prod_{i=1}^{p} \left(1 - \frac{i-1}{n}\right), \tag{3.3.1}$$

and

$$\mu_{\alpha+1}' = \mu_\alpha' \prod_{i=1}^{p} \left(1 - \frac{i-1}{n+2\alpha}\right). \tag{3.3.2}$$

From these relations it is quite simple to obtain the moments, hence the coefficients of skewness and kurtosis, and, from Table 42 of the Biometrika Tables for Statisticians [8], we can obtain, at least for moderately large n, very good approximations to the desired percentage points of the cdf. Thus, for testing the hypothesis of independence, the determinant test is quite useful and closely related to the step-down procedure presented in this paper. At the moment, however, we do not know of any method to use this determinant test for the construction of confidence bounds on parametric functions without running into complicated non-central distributions, whereas the step-down procedure, as described above, can be immediately inverted for the purpose of constructing simultaneous confidence bounds.

4. Confidence bounds associated with the test of independence for a p-variate problem. For shortness, let us now denote by just r_i the $r_{i\cdot 1,2,\cdots,i-1}$ defined by (2.3) and (2.4), by just $\boldsymbol{\beta}_i$ (with components $\beta_{i1}, \beta_{i2}, \cdots, \beta_{i,i-1}$, say) the $\boldsymbol{\beta}_{i\cdot 1,2,\cdots,i-1}$ defined by (2.1), and by just $\mathbf{b}_i$ (with components b_{i1}, $b_{i2}, \cdots, b_{i,i-1}$, say) the $\mathbf{b}_{i\cdot 1,2,\cdots,i-1}$ defined by (2.2). Assuming a general (symmetric, p.d.) Σ, let us now transform the original variates $x_1, x_2, \cdots, x_p$ to a new set $x_1^*, x_2^*, \cdots, x_p^*$ defined by

$$\text{(4.1)} \qquad \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \\ \cdot \\ x_*^p \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & & 0 \\ -\beta_{21} & 1 & 0 & \cdots & & 0 \\ -\beta_{31} & -\beta_{32} & 1 & \cdots & & 0 \\ \cdot & \cdot & \cdot & \cdots & & \cdot \\ -\beta_{p1} & -\beta_{p2} & -\beta_{p3} & \cdots & -\beta_{p,p-1} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ x_p \end{bmatrix}$$

Then it can be verified by induction or otherwise that the new variates are uncorrelated and hence that the step-down correlations of the new variates, $r_{i\cdot 1,2,\cdots,i-1}^*$, say, ($i = p, p-1, \cdots, 2$) are independently distributed. With a joint probability, $1 - \alpha$, say, let us now make the simultaneous statements:

$$\text{(4.2)} \qquad r_{i\cdot 1,2,\cdots,i-1}^{*2} \leqq \mu \qquad (\text{for } i = p, p-1, \cdots, 2).$$

We have already seen that, given μ, we can easily find α and also, given α, we can find μ. In order to invert the typical component statement (4.2) and thus make a confidence statement on $\boldsymbol{\beta}_{i\cdot 1,2,\cdots,i-1}$, i.e., on $[\beta_{i1}, \cdots, \beta_{i,i-1}]$, we observe that the multiple correlation coefficient between x_i^* and $[x_1^*, x_2^*, \cdots, x_{i-1}^*]$ is the same as that between x_i^* and $[x_1, x_2, \cdots, x_{i-1}]$, since the starred variates in the first square brackets are linear combinations of just the non-starred variates the first square brackets are linear combinations of just the non-starred variates in the second square brackets; this fact simplifies our calculation of the desired expressions in terms of the S matrix, i.e., the sample dispersion matrix of the original variates. We may now use the results obtained in reference [4] in connection with the confidence bounds on a (pseudo-) regression matrix of a p-set on a q-set ($p \leqq q$) for a $(p + q)$-variate normal distribution. Let us take expression (3.2) from reference [4] and renumber it as

$$\text{(4.3)} \qquad c_{\max}^{1/2}(BB') - \lambda c_{\max}^{1/2}(S_{1\cdot 2}) c_{\max}^{1/2}(S_{22}^{-1}) \leqq c_{\max}^{1/2}(\beta\beta') \leqq c_{\max}^{1/2}(BB') + \lambda c_{\max}^{1/2}(S_{1\cdot 2}) c_{\max}^{1/2}(S_{22}^{-1}),$$

where B and β are the sample and population regression matrices of the p-set on the q-set given, respectively, by $B = S_{12}S_{22}^{-1}$ and $\beta = \sum_{12}\sum_{22}^{-1}$; $S_{1\cdot 2}$ is the sample "residual" matrix of the p-set on the q-set given by $S_{1\cdot 2} =$

$S_{11} - S_{12}S_{22}^{-1}S_{12}'$; S_{22} is, of course, the sample dispersion matrix of the q-set. Given a preassigned confidence coefficient, $1 - \alpha$, the value of λ in (4.3) can be obtained from the central distribution of the square of the largest canonical correlation coefficient, for which recursion formulas are available [2].

For $p = 1$ and $q = i - 1$, (4.3) reduces to

$$(4.4)\quad (\mathbf{b}_i'\mathbf{b}_i)^{1/2} - \lambda(s_{ii} - \mathbf{s}_i'S_{i-1}^{-1}\mathbf{s}_i)^{1/2}\, c_{\max}^{1/2}(S_{i-1}^{-1}) \leqq (\boldsymbol{\beta}_i'\boldsymbol{\beta}_i)^{1/2} \leqq (\mathbf{b}_i'\mathbf{b}_i)^{1/2} + \lambda(s_{ii} - \mathbf{s}_i'S_{i-1}^{-1}\mathbf{s}_i)^{1/2} c_{\max}^{1/2}(S_{i-1}^{-1}),$$

where $\mathbf{b}_i$ and $\boldsymbol{\beta}_i$ are defined in the opening paragraph of section 4, and

$$\mathbf{s}_i'(1 \times \overline{i-1}) = [s_{1i}, s_{2i}, \cdots, s_{i-1,i}],$$

$$(4.4.1)\quad S_{i-1}\,(\overline{i-1} \times \overline{i-1}) = \begin{bmatrix} s_{11} & \cdots & s_{1,i-1} \\ \cdot & \cdots & \cdot \\ s_{1,i-1} & \cdots & s_{i-1,i-1} \end{bmatrix},$$

and, finally, $\lambda = +[\mu/(1 - \mu)]^{1/2}$, where μ is obtained by the procedure outlined in the sequel of (3.2.2). It then follows that the typical statement (4.2) $\Rightarrow$ (4.4), and therefore simultaneous statements (4.2) for $i = p, p - 1, \cdots, 2$ will imply, with a joint probability $\geqq 1 - \alpha$, simultaneous confidence bounds (4.4) on $\boldsymbol{\beta}_i'\boldsymbol{\beta}_i$ for $i = p, p - 1, \cdots, 2$.

In equation (3.1) of reference [4], we may put $p = 1$ and $q = i - 1$ and choose the vector $\mathbf{d}_2$ given there in such a way as to make any one, any two, etc., and finally any $(i - 2)$ components of $\mathbf{d}_2$ equal to zero; if then we make the corresponding transition from (3.1) to (3.2) given in reference [4], we will have, along with each typical statement (4.4) above, truncated statements where any one, two, etc., finally any $(i - 2)$ components of $\boldsymbol{\beta}_i$ and $\mathbf{b}_i$ have been deleted without, however, disturbing the expressions that occur with λ. Thus, statement (4.4) and the truncations mentioned above will result in $2^{i-1} - 1$ joint confidence statements for given i. Since i can take the values $p, p - 1, \cdots, 2$, we will have, altogether, $\sum_{i=2}^{p} (2^{i-1} - 1) = 2^p - p - 1$ confidence statements with a joint confidence coefficient $\geqq 1 - \alpha$.

5. Independence in the $(p_1 + p_2 + \cdots + p_k)$-variate problem.

5.1. Independence (in distribution) of the step-down sets of canonical correlation coefficients, under the null hypothesis. Starting from (2.6) in section 2 we shall make a transformation from S to a partitioned triangular matrix, $\tilde{T}$, given by

$$nS(p \times p) = \begin{matrix} (p_1) \\ (p_2) \\ \\ (p_k) \end{matrix} \begin{bmatrix} \tilde{T}_{11} & 0 & \cdots & 0 \\ T_{21} & \tilde{T}_{22} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ T_{k1} & T_{k2} & \cdots & \tilde{T}_{kk} \end{bmatrix} \begin{bmatrix} \tilde{T}_{11} & 0 & \cdots & 0 \\ T_{21} & \tilde{T}_{22} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ T_{k1} & T_{k2} & \cdots & \tilde{T}_{kk} \end{bmatrix}'$$
$$\qquad\qquad (p_1) \quad (p_2) \qquad\quad (p_k)$$

The distribution of the elements of $\tilde{T}$, under a general $\sum$, will be given by

$$\text{(5.1.1)} \qquad \text{const} \cdot \exp\left[-\frac{1}{2} tr \textstyle\sum^{-1}\tilde{T}\tilde{T}'\right] \prod_{i=1}^{p} t_{ii}^{n-i} \prod_{i \geq j=1}^{k} dT_{ij},$$

where $p = p_1 + p_2 + \cdots + p_k$, and $T_{ii} = \tilde{T}_{ii}$ (i.e., triangular). Under H_0: $\sum_{ij} = 0$ $(i \neq j = 1, 2, \cdots, k)$, (5.1.1) reduces to

$$\text{(5.1.2)} \qquad \text{const} \cdot \exp\left[-\frac{1}{2}\sum_{i=1}^{k} tr \textstyle\sum_{ii}^{-1} (T_{i1}, \cdots, \tilde{T}_{ii})(T_{i1}, \cdots, \tilde{T}_{ii})'\right] \displaystyle\prod_{i=1}^{p} t_{ii}^{n-i} \prod_{i \geq j=1}^{k} dT_{ij}.$$

For shortness, let us write the matrix (2.8.2) in the form

$$\text{(5.1.3)} \qquad S_{ii}^{-1} S^{(i)\prime} S_{i-1}^{-1} S^{(i)},$$

where

$$\text{(5.1.4)} \qquad S^{(i)\prime}(p_i \times \overline{p_1 + p_2 + \cdots + p_{i-1}}) = \underset{(p_1)}{[S_{1i}'} \quad \underset{(p_2)}{S_{2i}'} \quad \cdots \quad \underset{(p_{i-1})}{S_{i-1,i}']} \; (p_i),$$

and

$$S_{i-1}(\overline{p_1 + \cdots + p_{i-1}} \times \overline{p_1 + \cdots + p_{i-1}}) = \begin{bmatrix} S_{11} & \cdots & S_{1,i-1} \\ \cdot & \cdots & \cdot \\ S_{1,i-1}' & \cdots & S_{i-1,i-1} \end{bmatrix}.$$

Also, let us denote the p_i characteristic roots of (5.1.3), ordered from the smallest to the largest, by

$$[c_i^{(1)}, c_i^{(2)}, \cdots, c_i^{(p_i)}].$$

It then follows directly that

$$\text{(5.1.5)} \qquad nS_{ii} = [T_{i1} \cdots \tilde{T}_{ii}][T_{i1} \cdots \tilde{T}_{ii}]',$$

$$nS^{(i)\prime} = [T_{i1} \cdots T_{i,i-1}] \begin{bmatrix} \tilde{T}_{11} & 0 & \cdots & 0 \\ T_{21} & \tilde{T}_{22} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ T_{i-1,1} & T_{i-1,2} & \cdots & \tilde{T}_{i-1,i-1} \end{bmatrix}',$$

and

$$nS_{i-1} = \begin{bmatrix} \tilde{T}_{11} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ T_{i-1,1} & \cdot & \cdots & \tilde{T}_{i-1,i-1} \end{bmatrix} \begin{bmatrix} \tilde{T}_{11} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ T_{i-1,1} & \cdot & \cdots & \tilde{T}_{i-1,i-1} \end{bmatrix}'.$$

Substituting the expressions (5.1.5) into (5.1.3) (or (2.8.2)), we see that (5.1.3) reduces to

$$\text{(5.1.6)} \qquad \left[\sum_{j=1}^{i} T_{ij} T_{ij}'\right]^{-1} \left[\sum_{j=1}^{i-1} T_{ij} T_{ij}'\right].$$

Now, since by (5.1.2) the sets $[T_{i1} \cdots \tilde{T}_{ii}]$ are independently distributed under the null hypothesis, for $i = k, k - 1, \cdots 1$, it follows that the sets of characteristic roots of (5.1.3), viz.,

$$[c_2^{(1)}, \cdots, c_2^{(p_2)}], \quad [c_3^{(1)}, \cdots, c_3^{(p_3)}], \cdots, [c_k^{(1)}, \cdots, c_k^{(p_k)}]$$

are independently distributed.

5.2. The proposed test of independence. We propose the following test:

$$\text{(5.2.1)} \qquad \text{Accept } H_0 \text{ over } \bigcap_{i=2}^{k} [c_i^{(p_i)} \leqq \lambda], \quad \text{and}$$

$$\text{reject } H_0 \text{ over } \bigcup_{i=2}^{k} [c_i^{(p_i)} > \lambda],$$

where λ is given by

$$\text{(5.2.2)} \qquad \prod_{i=2}^{k} P[c_i^{(p_i)} \leqq \lambda \mid \gamma_i^{(p_i)} = 0] = 1 - \alpha,$$

and γ denotes the largest characteristic root of $\sum_{ii}^{-1} \sum^{(i)'} \sum_{i-1}^{-1} \sum^{(i)}$, where the $\sum_{ij}$'s are obtained from $\sum$ in exactly the same way as the S_{ij}'s from S in equations (5.1.4). It will be noted that $\gamma_i^{(p_i)}$ is zero if and only if, for given i, $\sum_{ij} = 0$, for $j = 1, 2, \cdots, i - 1$. Analogous to section 3.2., we take the same value of λ for each factor on the left-hand side of (5.2.2); the reason is the same as that given in section 3.2. The procedure for obtaining λ is analogous to that given for μ in section (3.2.) except that the incomplete Beta function needs to be replaced by the central distribution function of the (square of) the largest canonical correlation coefficient. The distribution and recursion relations for particular values are discussed explicity in reference [2].

5.3. Relation to the likelihood-ratio test. Denoting the jth canonical correlation coefficient of the p_i-set on the $(p_1 + p_2 + \cdots + p_{i-1})$-set by $r_{i \cdot 1,2,\cdots,i-1}^{(j)}$, we see that

$$\text{(5.3.1)} \qquad 1 - r_{i\cdot 1,2,\cdots,i-1}^{(j)2} = c^{(j)}[S_{ii}^{-1}(S_{ii} - S^{(i)'}S_{i-1}^{-1}S^{(i)}]$$
$$= c^{(j)}[R_{ii}^{-1}(R_{ii} - R^{(i)'}R_{i-1}^{-1}R^{(i)}],$$

where $c^{(j)}$ denotes the jth characteristic root, and the R's are the sample correlation matrices corresponding to the covariance matrices, S.
Thus,

$$\text{(5.3.2)} \qquad \prod_{j=1}^{p_i} (1 - r_{i\cdot 1,2,\cdots,i-1}^{(j)2}) = \frac{|R_i|}{|R_{ii}|\,|R_{i-1}|},$$

and the product of the products of all step-down canonical correlation coefficients (or rather, of the complements of their squares) becomes

$$\text{(5.3.3)} \qquad \prod_{i=2}^{k} \prod_{j=1}^{p_i} (1 - r_{i\cdot 1,2,\cdots,i-1}^{(j)2}) = \frac{|R|}{\prod_{i=1}^{k} |R_{ii}|},$$

because $|R_1| = |R_{11}|$. Thus, a comparison with reference [6] shows that a test based on the product of products of all step-down correlations is closely related to the likelihood-ratio test. The distribution of this statistic, under H_0, is discussed in [7], and the moments are given in [6]. They can be readily obtained if, in the joint distribution of all r_{ij}'s (for a general matrix $(\zeta_{ij}) = \hat{R}$, say)

$$P(R \mid \hat{R}) = \frac{p\Gamma(np/2)}{\pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma(n - i + 1/2)} \cdot \frac{|R|^{n-p-1/2}}{|\hat{R}|^{n/2}} \tag{5.3.4}$$
$$\times \int_{\infty}^{\infty} \cdots \int_{\infty}^{\infty} \frac{d\beta_1 \cdots d\beta_{p-1}}{[tr\hat{R}^{-1} D_\beta R D_\beta]^{np/2}} \prod_{i>j=1}^{p} dr_{ij},$$

(where D_β is a diagonal matrix with elements $e^{-\beta_1/2}$,

$$e^{(\beta_1-\beta_2)/2}, \quad e^{(\beta_2-\beta_3)/2}, \cdots, e^{(\beta_{p-2}-\beta_{p-1})/2}, e^{\beta_{p-1}/2})$$

we set

$$\hat{R} = \begin{bmatrix} \hat{R}_{11} & 0 & \cdots & 0 \\ 0 & \hat{R}_{22} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \hat{R}_{kk} \end{bmatrix}.$$

The moments satisfy the convenient recurrence relation:

$$\mu_1' = \frac{\prod_{q=1}^{p} (n - q + 1)}{\prod_{j=1}^{k} \prod_{i=1}^{p_j} (n - i + 1)}, \tag{5.3.5}$$

and

$$\mu_{\alpha+1}' = \mu_\alpha' \frac{\prod_{q=1}^{p} (n + 2\alpha - q + 1)}{\prod_{j=1}^{k} \prod_{i=1}^{p_j} (n + 2\alpha - i + 1)}.$$

Thus, for testing the hypothesis of independence between k sets of normally distributed variates, the determinant test is quite useful and closely related to the step-down procedure. At the moment, however, we cannot easily construct simultaneous confidence bounds on parametric functions on the basis of the determinant test, whereas the step-down procedure can be inverted into a simultaneous confidence statement.

6. Confidence bounds associated with the test of independence for a $(p_1 + p_2 + \cdots + p_k)$-variate problem. Using (5.1.4) let us rewrite $\beta_{i\cdot 1,2,\cdots,i-1}$ in (2.7) and $B_{i\cdot 1,2,\cdots,i-1}$ in (2.8) as

$$\beta_i(p_i \times \overline{p_1 + p_2 + \cdots + p_{i-1}}) = \Sigma^{(i)\prime} \Sigma_{i-1}^{-1} \tag{6.1}$$

and

$$B_i(p_i \times (p_1 + p_2 + \cdots + p_{i-1})) = S^{(i)\prime} S_{i-1}^{-1}. \tag{6.2}$$

Next we partition β_i into

$$\begin{matrix}[\beta_{i1} & \beta_{i2} & \cdots & \beta_{i,i-1}] \\ (p_1) & (p_2) & & (p_{i-1})\end{matrix}$$

and B_i into

$$[B_{i1}\ B_{i2} \cdots B_{i,i-1}].$$

Assuming now a general (symmetric, p.d.) $\sum$, let us transform the original variates $\mathbf{x}_1(p_1 \times 1)$, $\mathbf{x}_2(p_2 \times 1)$, $\cdots$, $\mathbf{x}_k(p_k \times 1)$ into a new set of variates $\mathbf{x}_1^*$, $\mathbf{x}_2^*$, $\cdots$, $\mathbf{x}_k^*$ defined by

$$\begin{bmatrix}\mathbf{x}_1^* \\ \mathbf{x}_2^* \\ \cdot \\ \mathbf{x}_k^*\end{bmatrix}\begin{matrix}(p_1) \\ (p_2) \\ \\ (p_k)\end{matrix} = \begin{bmatrix} I(p_1) & 0 & 0 & \cdots & 0 & 0 \\ -\beta_{21} & I(p_2) & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ -\beta_{k1} & -\beta_{k2} & \cdot & \cdots & -\beta_{k,k-1} & I(p_k)\end{bmatrix}\begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2 \\ \cdot \\ \mathbf{x}_k\end{bmatrix} \tag{6.3}$$

Then it can be verified by induction or otherwise that the k sets of new (starred) variates are uncorrelated, and hence the step-down sets of (squares of) canonical correlations, $[c_2^{*(1)}, \cdots, c_2^{*(p_2)}]$, $[c_3^{*(1)}, \cdots, c_3^{*(p_3)}]$, $\cdots$, $[c_k^{*(1)}, \cdots, c_k^{*(p_k)}]$ are independently distributed. With a joint probability, $1 - \alpha$, say, we may thus make the simultaneous statement

$$c_i^{*(p_i)} \leqq \lambda \qquad (\text{for } i = k, k-1, \cdots, 2). \tag{6.4}$$

Analogous to section 4, with the modifications given in 5.2., we can find λ if α is preassigned. By the same argument as in section 4, and by using (4.3) we can obtain, with a joint confidence coefficient $\geqq 1 - \alpha$, the following sets of simultaneous confidence bounds, for $i = k, k - 1, \cdots, 2$:

$$\begin{aligned} c_{\max}^{1/2}(B_iB_i') - \lambda c_{\max}^{1/2}[S_{ii} - S^{(i)\prime}S_{i-1}^{-1}S^{(i)}]c_{\max}^{1/2}(S_{i-1}^{-1}) &\leqq c_{\max}^{1/2}(\beta_i\beta_i') \\ \leqq c_{\max}^{1/2}(B_iB_i') + \lambda c_{\max}^{1/2}[S_{ii} - S^{(i)\prime}S_{i-1}^{-1}S^{(i)}]c_{\max}^{1/2}(S_{i-1}^{-1}),& \end{aligned} \tag{6.5}$$

where β_1 and B_i are defined by (6.1) and (6.2), $S^{(i)}$ and S_{i-1} by (5.1.4), and $\sum^{(i)}$ and $\sum_{i-1}$ analogously. Following the argument presented in section 3 of reference [4] and in section 4 of this paper we see that, with a joint confidence coefficient $\geqq 1 - \alpha$, not only can we make the $(k - 1)$ statements (6.5) but, for each typical statement under (6.5), we can also make a number of truncated confidence statements by deleting any number of variates of the (p_i)-set and any number of variates of the $(p_1 + p_2 + \cdots + p_{i-1})$-set taking care only that the number of variates left in the $(p_1 + \cdots + p_{i-1})$-set is not less than that left in the p_i-set.

REFERENCES

[1] I. OLKIN AND S. N. ROY, "On multivariate distribution theory", *Ann. Math. Stat.*, vol. 25 (1954).

[2] S. N. ROY, "A report on some aspects of multivariate analysis", *Institute of Statistics, University of North Carolina*, mimeograph series No. 121.

[3] S. N. ROY, "Some further results in simultaneous confidence interval estimation", *Ann. Math. State.*, vol. 25 (1954).

[4] S. N. ROY AND R. GNANADESHIKAN, "Further contributions to multivariate confidence bounds", *Institute of Statistics, University of North Carolina*, mimeograph series No. 155. Also *Biomitrika*, vol. 44 (1957).

[5] S. S. WILKS, "Certain generalizations in the analysis of variance", *Biometrika*, vol. 24 (1932).

[6] S. S. WILKS, "On the independence of *k* sets of normally distributed statistical variables", *Econometrica*, vol. 3 (1935).

[7] A. WALD AND R. J. BROOKNER, "On the distribution of Wilks' statistic for testing the independence of several groups of variates", *Ann. Math. Stat.*, vol. 12 (1941).

[8] E. S. PEARSON AND H. O. HARTLEY, "Biometrika Tables for Statisticians", vol. I, 4th ed., *Cambridge University Press* (1954).