STEP-DOWN PROCEDURE IN MULTIVARIATE ANALYSIS[1]

By J. Roy

University of North Carolina

1. Introduction and summary. Test criteria for (i) multivariate analysis of variance, (ii) comparison of variance-covariance matrices, and (iii) multiple independence of groups of variates when the parent population is multivariate normal are usually derived either from the likelihood-ratio principle [6] or from the "union-intersection" principle [2]. An alternative procedure, called the "step-down" procedure, has been recently used by Roy and Bargmann [5] in devising a test for problem (iii). In this paper the step-down procedure is applied to problems (i) and (ii) in deriving new tests of significance and simultaneous confidence-bounds on a number of "deviation-parameters."

The essential point of the step-down procedure in multivariate analysis is that the variates are supposed to be arranged in descending order of importance. The hypothesis concerning the multivariate distribution is then decomposed into a number of hypotheses—the first hypothesis concerning the marginal univariate distribution of the first variate, the second hypothesis concerning the conditional univariate distribution of the second variate given the first variate, the third hypothesis concerning the conditional univariate distribution of the third variate given the first two variates, and so on. For each of these component hypotheses concerning univariate distributions, well known test procedures with good properties are usually available, and these are made use of in testing the compound hypothesis on the multivariate distribution. The compound hypothesis is accepted if and only if each of the univariate hypotheses are accepted. It so turns out that the component univariate tests are independent, if the compound hypothesis is true. It is therefore possible to determine the level of significance of the compound test in terms of the levels of significance of the component univariate tests and to derive simultaneous confidence-bounds on certain meaningful parametric functions on the lines of [3] and [4].

The step-down procedure obviously is not invariant under a permutation of the variates and should be used only when the variates can be arranged on a priori grounds. Some advantages of the step-down procedure are (i) the procedure uses widely known statistics like the variance-ratio, (ii) the test is carried out in successive stages and if significance is established at a certain stage, one can stop at that stage and no further computations are needed, and (iii) it leads to simultaneous confidence-bounds on certain meaningful parametric functions.

1.1 *Notations.* The operator $\mathcal{E}$ applied to a matrix of random variables is used

Received January 17, 1958; revised May 5, 1958.

[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. Reproduction in whole or in part is permitted for any purpose of the United States Government.

to generate the matrix of expected values of the corresponding random variables. The form of a matrix is denoted by a subscript; thus $A_{n \times m}$ indicates that the matrix A has n rows and m columns. The maximum latent root of a square matrix B is denoted by $\lambda_{\max}(B)$. Given a vector $a = (a_1, a_2, \cdots, a_t)'$ and a subset T of the natural numbers 1, 2, $\cdots$, t, say $T = (j_1, j_2, \cdots, j_u)$ where $j_1 < j_2 < \cdots j_u$, the notation $T[a]$ will be used to denote the positive quantity:

$$T[a] = + \{a_{j_1}^2 + a_{j_2}^2 + \cdots + a_{j_t}^2\}^{1/2}.$$

$T[a]$ will be called the T-norm of a. Similarly, given a matrix $B_{t \times t}$, we shall write $B_{(T)}$ for the $u \times u$ submatrix formed by taking the j_1th, j_2th, $\cdots$, j_uth rows and columns of B. We shall call $B_{(T)}$ the T-submatrix of B.

2. Step-down procedure in multivariate analysis of variance.

2.1 *General linear hypothesis in univariate analysis.* Let the elements of $y_{n \times 1}$ be one-dimensional random variables distributed independently and normally with the same variance σ^2 and expectations given by

$$\mathcal{E}y = A\theta + X\beta \tag{1}$$

where elements of $\theta_{m \times 1}$ and $\beta_{q \times 1}$ are unknown parameters; $A_{n \times m}$ and $X_{n \times q}$ are matrices of known constants with rank $(A) = r$ and rank $(A:X) = r + q$, with $n > (r + q)$.

A set of t linearly independent linear functions $\phi_{t \times 1} = B_{t \times m}\theta$, where B is a given matrix of rank t, is said to be estimable if for each element of ϕ there exists an unbiassed estimate linear in y, for all values of θ and β. If ϕ is estimable, there exists an estimator $\hat{\phi}_{t \times 1}$ of ϕ, the elements of which are linear in y and minimum variance unbiassed estimators of the corresponding elements in ϕ. Denote the variance-covariance matrix of $\hat{\phi}$ by $C \cdot \sigma^2$, where $C_{t \times t}$ is a positive-definite matrix. Let $s^2/(n - q - r)$ denote the usual error mean square with $(n - q - r)$ degrees of freedom giving an unbiassed estimator of σ^2. Then it is well known that the statistics $u = (\hat{\phi} - \phi)'C^{-1}(\hat{\phi} - \phi)/\sigma^2$ and $v = s^2/\sigma^2$ are distributed independently as chi-squares with t and $(n - q - r)$ degrees of freedom respectively, so that

$$F \equiv \frac{(\hat{\phi} - \phi)'C^{-1}(\hat{\phi} - \phi)/t}{s^2/(n - q - r)} \tag{2}$$

is distributed as a variance-ratio with t and $(n - q - r)$ degrees of freedom.

Let α be a preassigned constant, $0 < \alpha < 1$, and f the upper 100α per cent point of the variance-ratio distribution with t and $(n - q - r)$ degrees of freedom. Setting $\mathcal{L}^2 = tf/(n - q - r)$ we then have

$$(\hat{\phi} - \phi)'C^{-1}(\hat{\phi} - \phi) \leqq \ell^2 s^2 \tag{3}$$

with probability $(1 - \alpha)$.

Now, the left-hand side of (3) is a positive definite quadratic form in $(\hat{\phi} - \phi)$ and consequently, we have

$$(\hat{\phi} - \phi)'C^{-1}(\hat{\phi} - \phi) \geqq (\hat{\phi} - \phi)'(\hat{\phi} - \phi)/\lambda_{\max}(C). \tag{4}$$

We thus have

$$(\hat{\phi} - \phi)'(\hat{\phi} - \phi) \leqq \ell^2 s^2 \lambda_{\max}(C) \tag{5}$$

with probability not less than $(1 - \alpha)$.

Now, let T be any subset of the natural numbers $1, 2, \cdots, t$ and consider the T-norms $T[\phi]$ of ϕ and $T[\hat{\phi}]$ of $\hat{\phi}$. Then (3) implies that

$$T[\hat{\phi}] - \ell s \lambda_{\max}^{1/2}(C_{(T)}) \leqq T[\phi] \leqq T[\hat{\phi}] + \ell s \lambda_{\max}^{1/2}(C_{(T)}) \tag{6}$$

for all subsets T of $(1, 2, \cdots, t)$, where C_T is the T-submatrix of C. The statement (6) thus provides simultaneous confidence-bounds on the parameters $T[\phi]$ for all T with probability not less than $(1 - \alpha)$. We note that there are in all $(2^t - 1)$ parameters of the type $T[\phi]$ and these in a sense measure the deviations from the hypothesis $\mathcal{H}_0$ that $\phi = 0$. The analysis of variance test for $\mathcal{H}_0$ at level of significance α, of course, is given by the rule

$$\text{accept } \mathcal{H}_0 \text{ if } \frac{\hat{\phi}'C^{-1}\hat{\phi}/t}{s^2/(n - q - r)} \leqq f; \tag{7}$$

otherwise reject $\mathcal{H}_0$.

However, simultaneous confidence-bounds of the type (6) are more interesting than the test (7) itself, because the direction of departure from the null hypothesis is indicated.

2.2 *Customary tests in multivariate analysis of variance.* We have a matrix $Y_{n \times p}$ of random variables, such that the rows are distributed independently, each row having a p-variate normal distribution with the same variance-covariance matrix $\Sigma_{p \times p}$ which is positive-definite. The expected values are given by

$$\mathcal{E}Y = A\Theta, \tag{8}$$

where $A_{n \times m}$ is a matrix of known constants of rank r, $r \leqq (n - p)$, and $\Theta_{m \times p}$ is a matrix of unknown parameters. As before, a set of linear parametric functions $\Phi_{t \times p} = B_{t \times m}\Theta$ is said to be estimable if, for all Θ, there exist unbiassed estimates of Φ linear in Y. If Φ is estimable, customary tests for the hypothesis

$$\mathcal{H}_0: \Phi = 0$$

are based on two $p \times p$ matrices of random variables

$$S_e = Y'EY \quad \text{and} \quad S_h = Y'HY, \tag{9}$$

called respectively the sum of products matrix due to error and the sum of products matrix due to hypothesis. Here E and H are $n \times n$ symmetric idempotent matrices with non-stochastic elements, E of rank $(n - r)$ and H of rank t, E being a function of A, and H of both A and B. The likelihood-ratio test [6] is

$$\text{accept } \mathcal{H}_0 \text{ if } L \equiv \frac{|S_e|}{|S_e + S_h|} > c, \tag{10}$$

otherwise reject $\mathcal{H}_0$,

where c is a preassigned constant depending on the level of significance. The test based on the largest latent root [3] is

$$\text{(11)} \qquad \begin{array}{l} \text{accept } \mathcal{H}_0 \text{ if } \lambda_{\max}(S_h S_e^{-1}) < d, \\ \text{otherwise reject } \mathcal{H}_0 , \end{array}$$

where d is a constant depending on the level of significance. Simultaneous confidence-bounds on certain meaningful parametric functions have been derived by the largest (or the largest-smallest roots) procedure, [3] [4], whereas no such bounds are available as of now from the likelihood-ratio procedure.

2.3 *The step-down procedure.* We shall denote the ith columns of the matrices Y and Θ in section 2.2 by y_i and θ_i respectively and write $Y_i = [y_1 \, y_2 \, \cdots \, y_i]$ and $\Theta_i = [\theta_1 \, \theta_2 \, \cdots \, \theta_i]$. Further, we shall denote the top left-hand $i \times i$ submatrix of $\Sigma \equiv ((\sigma_{ij}))$ by Σ_i.

Then, under the condition that Y_i is fixed, the n elements of the vector y_{i+1} are distributed independently and normally each with the same variance σ_{i+1}^2 and expectations given by

$$\text{(12)} \qquad \mathcal{E} y_{i+1} = A \eta_{i+1} + Y_i \beta_i ,$$

where β_i is a vector of the form $i \times 1$ given by

$$\text{(13)} \qquad \beta_i = \Sigma_i^{-1} \begin{bmatrix} \sigma_{1,i+1} \\ \sigma_{2,i+1} \\ \cdots\cdot \\ \sigma_{i,i+1} \end{bmatrix}, \qquad \beta_0 = 0,$$

amd η_{i+1} is a vector of the form $m \times 1$ given by

$$\text{(14)} \qquad \eta_{i+1} = \theta_{i+1} - \Theta_i \beta_i$$

and

$$\text{(15)} \qquad \sigma_{i+1}^2 = \frac{|\Sigma_{i+1}|}{|\Sigma_i|},$$

with the understanding that $|\Sigma_0| = 1$ so that $\sigma_1^2 = \sigma_{11}$, $i = 0, 1, 2, \cdots, (p-1)$. The elements of the vectors β_i, η_{i+1} may then be regarded as unknown parameters. We shall call β_i the ith order step-down regression coefficient and σ_{i+1}^2 the ith order step-down residual variance.

Let us now consider linear functions

$$\text{(16)} \qquad \phi_i = B \eta_i \qquad (i = 1, 2, \cdots, p).$$

If Y_i is fixed, (12) is of the same form as (1). Let us now, with an easily understood notation similar to that used in Section 2.1, construct the statistics

$$\text{(17)} \qquad F_i \equiv \frac{(\hat{\phi}_i - \phi_i)' C_i^{-1} (\hat{\phi}_i - \phi_i)'/t}{s_i^2/(n - r - i + 1)} \qquad (i = 1, 2, \cdots, p).$$

Obviously, when Y_{i-1} is fixed, the statistic F_i is distributed as a variance ratio with t and $(n - r - i + 1)$ degrees of freedom $(i = 2, 3, \cdots, p)$. Finally, we note that in its functional form F_i involves only Y_i $(i = 1, 2, \cdots, p$ and that the conditional distribution of F_i, given Y_{i-1} does not involve Y_{i-1} $(i = 2, 3, \cdots, p)$ and hence $F_{i-1}, \cdots, F_1$. Also, F_1 is marginally distributed as a variance-ratio with t and $(n - r)$ degrees of freedom. Therefore the statistics $F_1, F_2, \cdots, F_p$ are independent. This can be verified in a straight-forward manner by using the transformation to rectangular coordinates as in [5] or any other set of step-down variates, or even otherwise.

For a preassigned constant α_i, $0 < \alpha_i < 1$, let f_i denote the upper $100\alpha_i$ per cent point of the variance-ratio distribution with t and $(n - r - i + 1)$ degrees of freedom. Then the probability P that simultaneously

$$F_i \leqq f_i, \qquad i = 1, 2, \cdots, p, \tag{18}$$

is given by

$$P = \prod_{i=1}^{p} (1 - \alpha_i). \tag{19}$$

Therefore, for any subset T of the natural numbers $1, 2, \cdots, t$ writing as in (6), $T[\phi_i]$ and $T[\hat{\phi}_i]$ for the T-norms of ϕ_i and $\hat{\phi}_i$ respectively, and setting

$$\ell_i^2 = tf_i/(n - r - i + 1) \tag{20}$$

and writing $C_{i(T)}$ for the T-submatrix of C_i, we have the simultaneous confidence bounds

$$T[\hat{\phi}_i] - \ell_i s_i \lambda_{\max}^{1/2}(C_{i(T)}) \leqq T[\phi_i] \leqq T[\hat{\phi}_i] + \ell_i s_i \lambda_{\max}^{1/2}(C_{i(T)}) \tag{21}$$

for all subsets T of $(1, 2, \cdots, t)$ and $i = 1, 2, \cdots, p$ with probability greater than P.

To derive a test of the hypothesis $\mathcal{H}_0$ that $\Phi = 0$, we note that $\mathcal{H}_0$ is true if and only if the hypothesis $\mathcal{H}_i$ that $\phi_i = 0$ holds for all $i = 1, 2, \cdots, p$. Using the result (17), we set up the following procedure for testing $\mathcal{H}_0$:

$$\text{accept } \mathcal{H}_0 \text{ if } u_i \equiv \frac{\hat{\phi}_i' C_i^{-1} \hat{\phi}_i / t}{s_i^2/(n - r - i + 1)} \leqq f_i \qquad \text{for all } i = 1, 2, \cdots, p; \tag{22}$$

otherwise reject $\mathcal{H}_0$.

Obviously, the level of significance for this test is $1 - P$ where P is given by (19). The arbitrariness in determining the f_i's when the level of significance is preassigned may be removed by stipulating that $\alpha_1 = \alpha_2 = \cdots = \alpha_p$. From the fact that the variance-ratio test (7) is uniformly unbiassed, it can be seen after a little consideration, that the test procedure (22) is also uniformly unbiassed.

To carry out the test one should first compute u_1. If $u_1 > f_1$, $\mathcal{H}_0$ is rejected and no further computations are needed. If $u_1 \leqq f_1$, the next step is to compute u_2. If $u_2 > f_2$, $\mathcal{H}_0$ is rejected and no further computations are needed. If $u_2 \leqq f_2$, one proceeds to compute u_3 and so on. This way one need compute u_i if and only if $u_j \leqq f_j$ for $j = 1, 2, \cdots, i - 1$. Much computational labor is saved thereby.

It is well known that the likelihood-ratio statistic L given by (10) can be expressed as

$$L = \prod_{i=1}^{p} \frac{(n - r - i + 1)}{t + (n - r - i + 1)u_i} \tag{23}$$

and this has been utilized [1] to obtain the moments of L when $\mathcal{H}_0$ is true. However, the step-down procedure based on the individual u_i's rather than on a single function L, is advantageous from the point of view of (i) setting up simultaneous confidence bounds and (ii) saving computational labor, specially in the situation indicated in the introduction.

3. Step-down procedure for variance-covariance matrices. Let $S_{p \times p} \equiv ((s_{ij}))$ be a symmetric matrix of random variables, distributed in Wishart's form with n degrees of freedom, $n > p$, so that S/n provides an unbiassed estimate for the variance-covariance matrix Σ of a p-variate normal population. In the same way as in Section 2.3, we shall write S_i for the $i \times i$ top left-hand submatrix of S and let

$$b_i = S_i^{-1} \begin{bmatrix} s_{1,i+1} \\ s_{2,i+1} \\ \cdots \\ s_{i,i+1} \end{bmatrix}, \qquad b_0 = 0, \tag{24}$$

$$s_{i+1}^2 = \frac{|S_{i+1}|}{|S_i|}, \qquad s_1^2 = s_{11}, \tag{25}$$

for $i = 1, 2, \cdots, p - 1$. Let β_{i-1} and σ_i^2 be defined by (13) and (15) for $i = 1, 2, \cdots, p$. Then it is well known that when S_i is fixed, the distribution of b_i is independent of the distribution of s_{i+1}^2; the distribution of b_i is i-variate normal with expectation β_i and variance-covariance matrix $\sigma_{i+1}^2 S_i^{-1}$, and s_{i+1}^2/σ_{i+1}^2 has the chi-square distribution with $(n - i)$ degrees of freedom, $i = 1, 2, \cdots, (p - 1)$. Finally s_1^2/σ_1^2 has the chi-square distribution with n degrees of freedom.

When more than one variance-covariance matrix is involved, we shall distinguish them by a superscript under parentheses. Thus with a number of population variance-covariance matrices $\Sigma^{(j)}$ and the corresponding Wishart matrices $S^{(j)}$, the quantities $\beta_i^{(j)}$, $\sigma_i^{(j)}$, $b_i^{(j)}$, $s_i^{(j)}$, etc., will be defined in the same way as in (13), (15), (24), and (25) for $j = 1, 2, \cdots$, etc.

3.1 *One variance-covariance matrix.* On the basis of a matrix S distributed in Wishart's form with n degrees of freedom, with S/n providing an unbiassed estimate for Σ, it is possible to set up simultaneous confidence-bounds on parameters which are functions of the elements of Σ by the step-down procedure as follows.

When S_i is fixed, the statistics $u = (b_i - \beta_i)'S_i(b_i - \beta_i)/\sigma_{i+1}^2$ and $v =$

s_{i+1}^2/σ_{i+1}^2 are distributed independently as chi-squares, u with i degrees of freedom and v with $n - i$ degrees of freedom. Therefore, given pre-assigned positive constants a_i, c_{i+1}, and d_{i+1}, where $c_{i+1} < d_{i+1}$, the probability P_{i+1} that

$$(26) \qquad \begin{aligned} (b_i - \beta_i)' S_i (b_i - \beta_i)/s_{i+1}^2 &\leqq a_i^2, \\ c_{i+1} \leqq s_{i+1}^2/\sigma_{i+1}^2 &\leqq d_{i+1} \end{aligned}$$

holds for fixed S_i, is a constant depending only on n, i, a_i, c_{i+1}, and d_{i+1}. As a matter of fact,

$$(27) \qquad P_{i+1} = \int_{c_{i+1}}^{d_{i+1}} G_i(a_i^2 x) g_{n-i}(x)\, dx \qquad (i = 1, 2, \cdots, p - 1),$$

where

$$(28) \qquad G_\nu(x) = \int_0^x g_\nu(\xi)\, d\xi$$

and

$$(29) \qquad g_\nu(x) = \frac{e^{-x} x^{\frac{1}{2}\nu - 1}}{2^{\frac{1}{2}\nu}\Gamma(\frac{1}{2}\nu)}.$$

Also, given preassigned positive constants b_1, $c_1(b_1, < c_1)$, the marginal probability P_1 that

$$(30) \qquad c_1 \leqq s_1^2/\sigma_1^2 \leqq d_1$$

is given by

$$(31) \qquad P_1 = \int_{c_1}^{d_1} g_n(x)\, dx.$$

By an argument similar to that which follows (17) in section 2.3, we obtain the probability P that simultaneously

$$(32) \qquad \begin{aligned} c_i \leqq s_i^2/\sigma_i^2 &\leqq d_i & (i = 1, 2, \cdots, p), \\ (b_i - \beta_i)' S_i (b_i - \beta_i)/s_{i+1}^2 &\leqq a_i^2 & (i = 1, 2, \cdots, p - 1) \end{aligned}$$

as

$$P = \prod_{i=1}^{p} P_i.$$

Now, as in Section 2.3, for a given subset T_i of the integers $1, 2, \cdots, i$, writing $T_i[\beta_i]$ and $T_i[b_i]$ for the T_i-norms of β_i and b_i respectively, and writing $U_{i(T_i)}$ for the T_i-submatrix of S_i^{-1},

$$(33) \qquad \begin{aligned} s_i^2/d_i \leqq \sigma_i^2 &\leqq s_i^2/c_i \qquad \text{for } i = 1, 2, \cdots, p, \\ T_i[b_i] - a_i s_{i+1} \lambda_{\max}^{1/2}(U_{i(T_i)}) \leqq T_i[\beta_i] &\leqq T_i[b_i] + a_i s_{i+1} \lambda_{\max}^{1/2}(U_{i(T_i)}) \end{aligned}$$

for all subsets T_i of $(1, 2, \cdots, i)$ and $i = 1, 2, \cdots, p - 1$. The statement (33) thus provides simultaneous confidence-bounds on p parameters of the type σ_i^2 and $(2^p - p)$ parameters of the form $T_i[\beta_i]$ with probability not less than P.

It is to be noted that to set up simultaneous confidence bounds of the type (32), one has to evaluate the integral (27) which is not usually available in tabulated form. Another meaningful procedure, which, incidentally, avoids this difficulty, is to set up separate sets of simultaneous confidence bounds: one on $\sigma_1^2, \cdots, \sigma_p^2$, using the chi-square distribution for s_i^2/σ_i^2, with a preassigned probability and another set on the step-down regressions β_i, using the variance-ratio distribution for $(b_i - \beta_i)'S_i(b_i - \beta_i)/s_{i+1}^2$, and with a probability not less than a preassigned level.

We suggest a slightly different procedure for testing the hypothesis $\mathcal{H}_0$ that Σ has a specified value Σ_0. This hypothesis may be reformulated in terms of the step-down regression-coefficients and residual variances as follows: the hypothesis $\mathcal{H}_0$ is true if and only if each of the hypotheses

$$\mathcal{H}_{i1} : \sigma_i^2 = \sigma_{i0}^2, \qquad i = 1, 2, \cdots, p,$$

$$\mathcal{H}_{i2} : \beta_i = \beta_{i0}, \qquad i = 1, 2, \cdots, p - 1,$$

is true, where σ_{i0}^2, β_{i0} are derived from Σ_0 the same way as σ_i^2, β_i are derived from Σ. The test procedure suggested is:

accept $\mathcal{H}_0$ if

$$(34) \qquad \begin{aligned} c_i \leqq s_i^2/\sigma_{i0}^2 &\leqq d_i \qquad (i = 1, 2, \cdots, p), \\ (b_i - \beta_{i0})'S_i(b_i - \beta_{i0})/\sigma_{i+1,0}^2 &\leqq e_i^2 \quad (i = 1, 2, \cdots, p - 1); \end{aligned}$$

otherwise reject $\mathcal{H}_0$.

The level of significance α for this procedure is given by

$$(35) \qquad \alpha = 1 - \left\{\prod_{i=1}^{p} P_i'\right\}\left\{\prod_{i=1}^{p-1} P_i''\right\},$$

where

$$P_i' = \int_{c_i}^{d_i} g_{n-i+1}(x)\, dx,$$

$$P_i'' = G_i(e_i^2).$$

For a given α, the c_i, d_i, e_i's are not uniquely determined. The arbitrariness may be removed, for instance, by the further stipulation that

$$P_1' = P_2' = \cdots = P_p' = P_1'' = P_2'' = \cdots = P_{p-1}'' = \beta \text{ (say)}$$

and that (c_i, d_i) are the locally unbiassed partitioning of the 100 $(1 - \beta)$ per cent critical region based on the chi-square distribution with $n - i + 1$ degrees of freedom. With this choice of the constants c_i, d_i, e_i, the test procedure is locally unbiassed.

3.2 *Two variance-covariance matrices.* With two population variance-covariance matrices $\Sigma^{(1)}$, $\Sigma^{(2)}$ and two matrices of random variables $S^{(1)}$, $S^{(2)}$ distributed independently in Wishart's form with n_1 and n_2 degrees of freedom respectively, so that $S^{(j)}/n_j$ provides an unbiassed estimate for $\Sigma^{(j)}$, we can use the step-down procedure for testing the hypothesis $\mathcal{H}_0$ that the two variance-covariance matrices are identical or, in symbols,

$$\mathcal{H}_0 : \Sigma^{(1)} = \Sigma^{(2)},$$

and also set up simultaneous confidence bounds for parameters measuring deviations from $\mathcal{H}_0$.

Let us introduce the two sets of step-down regression-coefficients and residual variances: $\beta_i^{(j)}$, $\sigma_i^{(j)}$, $b_i^{(j)}$, and $s_i^{(j)}$. The hypothesis $\mathcal{H}_0$ may be reformulated in terms of the step-down parameters as follows: $\mathcal{H}_0$ is true if and only if the hypotheses

$$\begin{aligned} &\mathcal{H}_{i1} : \sigma_i^{(1)} = \sigma_i^{(2)}, && i = 1, 2, \cdots, p, \\ &\mathcal{H}_{i2} : \beta_i^{(1)} = \beta_i^{(2)}, && i = 1, 2, \cdots, p-1, \end{aligned} \tag{36}$$

are simultaneously true. We may take $\rho_i = \sigma_i^{(1)}/\sigma_i^{(2)}$ and $T_i[\delta_i]$ as measures of deviation from $\mathcal{H}_0$ where $\delta_i = \beta_i^{(1)} - \beta_i^{(2)}$, T_i is a subset of $(1, 2, \cdots, i)$ and $T_i[\delta_i]$ denotes the T_i-norm of δ_i. In this case, it has not been possible to set-up confidence bounds on all these parameters simultaneously. However, one may proceed as follows. Given pre-assigned positive constants c_i, d_i; $c_i < d_i$, and writing

$$r_i = \left(\frac{n_1 - i + 1}{n_2 - i + 1}\right)^{-1/2} s_i^{(1)}/s_i^{(2)}, \tag{37}$$

we find the probability that

$$r_i^2/d_i \leqq \rho_i^2 \leqq r_i^2/c_i, \qquad i = 1, 2, \cdots, p, \tag{38}$$

should hold simultaneously is given by

$$P = \prod_{i=1}^{p} P_i, \tag{39}$$

where

$$P_i = \int_{c_i}^{d_i} dF_{n_2-i+1}^{n_1-i+1}(x), \tag{40}$$

in which $F_n^m(x)$ stands for the distribution-function of the variance-ratio statistic with m degrees of freedom for the numerator and n degrees of freedom for the denominator. Therefore, (38) provides simultaneous confidence-bounds on ρ_i^2 $(i = 1, 2, \cdots, p)$ with probability P.

Let us now write $\hat{\delta}_i = b_i^{(1)} - b_i^{(2)}$ and note that if $S_i^{(1)}$ and $S_i^{(2)}$ are fixed, $\hat{\delta}_i$ is distributed in an i-variate normal form with expected value δ_i and variance-covariance matrix

$$\{\sigma_{i+1}^{(1)}\}^2\{S_i^{(1)}\}^{-1} + \{\sigma_{i+1}^{(2)}\}^2\{S_i^{(2)}\}^{-1}$$

distributed independently of $s_{i+1}^{(1)}$ and $s_{i+1}^{(2)}$. If $\mathcal{H}_{i+1,1}$ is true, we have $\sigma_{i+1}^{(1)} = \sigma_{i+1}^{(2)} = \sigma_{i+1}$, say. In that case, if $S_i^{(1)}$ and $S_i^{(2)}$ are fixed, $\hat{\delta}_i$ is distributed in an i-variate normal form with expected value δ_i and dispersion matrix $C_i . \sigma_{i+1}^2$ where

$$C_i = \{S_i^{(1)}\}^{-1} + \{S_i^{(2)}\}^{-1}. \tag{41}$$

Also, $\hat{\delta}_i$ is distributed independently of u_1 and u_2 where

$$u_j = (s_{i+1}^{(j)})^2/\sigma_{i+1}^2 \qquad (j = 1, 2) \tag{42}$$

and u_j is distributed as a chi-square with $(n_j - i)$ degrees of freedom. Consequently, writing

$$s_{i+1}^2 = (s_{i+1}^{(1)})^2 + (s_{i+1}^{(2)})^2 \tag{43}$$

we find that if $\mathcal{H}_{i+1,1}$ is true and $S_i^{(j)}$ are fixed $(j = 1, 2)$ the statistics

$$(\hat{\delta}_i - \delta_i)' C_i^{-1} (\hat{\delta}_i - \delta_i)/s_{i+1}^2 \tag{44}$$

and

$$\frac{n_2 - i}{n_1 - i} \left(\frac{s_{i+1}^{(1)}}{s_{i+1}^{(2)}} \right)^2 \tag{45}$$

are distributed independently as variance-ratios, (44) with i and $(n_1 + n_2 - 2i)$ degrees of freedom, and (45) with $(n_1 - i)$ and $(n_2 - i)$ degrees of freedom.

Therefore, given pre-assigned positive quantities e_i^2 the probability P' that

$$(\hat{\delta}_i - \delta_i)' C_i^{-1} (\hat{\delta}_i - \delta_i)/s_{i+1}^2 \leqq e_i^2, \quad i = 1, 2, \cdots, p - 1, \tag{46}$$

should hold simultaneously is equal to

$$P' = \prod_{i=1}^{p-1} P_i', \tag{47}$$

where

$$P_i' = F_{n_1+n_2-2i}^{i}(e_i^2) \tag{48}$$

provided $\mathcal{H}_{i1}$ is true for $i = 2, 3, \cdots, p$. From (45), we get the following simultaneous confidence-bounds (49) on the T_i-norms of δ_i where T_i is a subset of $(1, 2, \cdots, i)$ (under the highly restrictive condition that $\mathcal{H}_{i1}$ is true) for $i = 2, 3, \cdots, p$:

$$T_i[\hat{\delta}_i] - e_i s_{i+1} \lambda_{\max}^{1/2}(C_{i(T_i)}) \leqq T_i[\delta_i] \leqq T_i[\hat{\delta}_i] + e_i s_{i+1} \lambda_{\max}^{1/2}(C_{i(T_i)}) \tag{49}$$

with probability not less than P', where $C_{i(T_i)}$ is the T_i-submatrix of C_i.

To test the hypothesis $\mathcal{H}_0$, the step-down procedure suggested is:

accept $\mathcal{H}_0$ if

$$\begin{aligned} &(\hat{\delta}_i - \delta_i)' C_i^{-1} (\hat{\delta}_i - \delta_i)/s_{i+1}^2 \leqq e_i^2, \quad i = 1, 2 \cdots, p - 1, \\ &c_i \leqq \frac{n_2 - i + 1}{n_1 - i + 1} \frac{s_i^{(1)}}{s_i^{(2)}} \leqq d_i, \qquad i = 1, 2, \cdots, p, \end{aligned} \tag{50}$$

and, otherwise, reject $\mathcal{H}_0$,

where e_i^2, c_i, $d_i(c_i < d_i)$ are pre-assigned positive constants. The level of significance α is given by

$$\alpha = 1 - \left\{\prod_{i=1}^{p} P_i\right\}\left\{\prod_{i=1}^{p-1} P_i'\right\}, \tag{51}$$

where P_i is given by (40) and P_i' by (48). For a pre-assigned value of α, the constants c_i, d_i, e_i^2 are uniquely determined if we stipulate that

$$P_1 = P_2 = \cdots = P_p = P_1' = P_2' = \cdots = P_{p-1}' = \beta, \quad \text{say,}$$

and that (c_i, d_i) gives an unbiassed partitioning of the $100(1 - \beta)$ per cent critical region of the variance-ratio distribution with i and $n_1 + n_2 - 2i$ degrees of freedom. With this choice the step-down test is locally unbiassed.

4. Acknowledgment. The author wishes to thank Professor S. N. Roy for kindly going through the manuscript and suggesting improvements.

REFERENCES

[1] C. R. Rao, "On transformations useful in the distribution problems of least squares," *Sankhyā*, Vol. 12 (1952-53), pp. 339-346.

[2] S. N. Roy, "On a heuristic method of test construction and its use in multivariate analysis," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 220-238.

[3] S. N. Roy, "A report on some aspects of multivariate analysis," North Carolina Institute of Statistics Mimeograph Series No. 121, 1954.

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

[5] S. N. Roy and R. E. Bargmann, "Tests of multiple independence and the associated confidence-bounds," North Carolina Institute of Statistics Mimeograph Series No. 175, 1957.

[6] S. S. Wilks, "Certain generalizations in the analysis of variance," *Biometrika*, Vol. 24 (1932), pp. 471-494.