PSYCHOMETRIKA—VOL. 24, NO. 2
JUNE, 1959

ON METHODS IN THE ANALYSIS OF PROFILE DATA

SAMUEL W. GREENHOUSE AND SEYMOUR GEISSER*

NATIONAL INSTITUTE OF MENTAL HEALTH

This paper is concerned with methods for analyzing quantitative, non-categorical profile data, e.g., a battery of tests given to individuals in one or more groups. It is assumed that the variables have a multinormal distribution with an arbitrary variance-covariance matrix. Approximate procedures based on classical analysis of variance are presented, including an adjustment to the degrees of freedom resulting in conservative F tests. These can be applied to the case where the variance-covariance matrices differ from group to group. In addition, exact generalized multivariate analysis methods are discussed. Examples are given illustrating both techniques.

Much research in the social sciences is of the multivariate type; multiple observations are made on individuals who have been sampled from one or more populations. In particular, when the observations are in the form of a battery of tests or a set of items, there is the problem of profile analysis, wherein it is customary to test for differences in the levels and in the shapes of the group profiles. If the variables being observed are assigned to columns and the individuals to rows, the resulting matrix of observations is very suggestive of the data usually analyzed by analysis of variance. Furthermore, since the rows are random and the columns can be considered in almost all instances as fixed, the appropriate model is the mixed model.

As is well known, in order that the usually computed ratios of mean squares in this model [7, 14, 16] be exactly distributed as the F distribution, it is necessary that columns (variables), in addition to being normally distributed, have equal variances and be mutually independent or, at most, have equal correlations. But these assumptions seem much too restrictive. In most investigations, it is unrealistic to assume that three or more tests, items, or treatment schedules have the same pairwise correlations or that they have the same variances. It seemed obvious, therefore, that this problem of multiple observations should be considered in its greatest generality, namely, that an individual vector $x_1, x_2, \cdots, x_p$ is sampled from a p-variate normal distribution with an arbitrary variance-covariance matrix.

Exact procedures for analyzing data of this type have been known for some time and are usually referred to as the generalized multivariate analysis of variance [1, 10, 12, 13, 17]. These, however, require considerably more computations than that demanded by the arithmetic of the analysis of

*We are indebted to Mrs. Norma French for performing all the calculations appearing in this paper.

variance. Furthermore, an analysis of variance approach permits the analysis of a set of data which cannot be handled by multivariate procedures, namely, the case where n, the number of random vectors, is less than p, the number of variables. Although these multivariate methods are discussed subsequently and an example is given for the case of two groups, our main purpose is to utilize the simpler, and more familiar, conventional univariate analysis of variance techniques under the more general assumptions. Our results concerning the approximate distributions of the F statistics are based upon the work of Box [5, 6] with regard to one group and its extension, by Geisser and Greenhouse [8], to several groups. In addition, the latter have found certain adjustments to the approximate tests leading to conservative tests which can be used, when the group sample sizes are the same, in the case of unequal variance-covariance matrices among the groups.

It is of interest that Block, Levine, and McNemar [2] were also primarily concerned with the application of the analysis of variance to the profile problem. They presented F tests for testing the homogeneity of variable (columns) means, the homogeneity of over-all group means (profile levels) and the equality of profile shapes. However, they assumed equal variance, among the variables and, since they imply that the F tests are exact, it can only be inferred that they also assumed the variables to be independent or equally correlated.

The Problem

Our notation is almost identical to that used by Block, Levine, and McNemar. Let p tests, $x_1 , x_2 , \cdots , x_p$, be given to each of n_k individuals ($k = 1, 2, \cdots , g$) in each of g established groups. Assume both the p tests and the g groups to be fixed, i.e., they are not random elements sampled from larger populations. This model, which fixes interest in the tests and groups under study, conforms to many experimental situations met with in practice. The totality of Np observed scores ($N = \sum_{k=1}^{g} n_k$) can be classified according to the scheme at the top of the next page.

An individual i in group k has the profile

$$(x_{i1k} , \cdots , x_{ijk} , \cdots , x_{ipk}) \quad \begin{cases} i = 1, \cdots , n_k & \text{individuals in group } k \\ j = 1, \cdots , p & \text{variables} \\ k = 1, \cdots , g & \text{groups.} \end{cases}$$

And the group profile for group k, say, is represented by

$$(\bar{x}_{.1k} , \bar{x}_{.2k} , \cdots , \bar{x}_{.pk}).$$

Group	Ind.	Tests x_1	...	x_j	...	x_p	
1	1	x_{111}		x_{1j1}		x_{1p1}	
	n_1	$x_{n_1 11}$		$x_{n_1 j1}$		$x_{n_1 p1}$	
Means: Gr.	1	$\bar{x}_{\cdot 11}$		$\bar{x}_{\cdot j1}$		$\bar{x}_{\cdot p1}$	$\bar{x}_{\cdot\cdot 1}$
k	1	x_{11k}		x_{1jk}		x_{1pk}	
	i	x_{i1k}		x_{ijk}		x_{ipk}	
	n_k	$x_{n_k 1k}$		$x_{n_k jk}$		$x_{n_k pk}$	
Means: Gr.	k	$\bar{x}_{\cdot 1k}$		$\bar{x}_{\cdot jk}$		$\bar{x}_{\cdot pk}$	$\bar{x}_{\cdot\cdot k}$
g	1	x_{11g}		x_{1jg}		x_{1pg}	
	n_g	$x_{n_g 1g}$		$x_{n_g jg}$		$x_{n_g pg}$	
Means: Gr.	g	$\bar{x}_{\cdot 1g}$		$\bar{x}_{\cdot jg}$		$\bar{x}_{\cdot pg}$	$\bar{x}_{\cdot\cdot g}$
Means: All Groups		$\bar{x}_{\cdot 1\cdot}$		$\bar{x}_{\cdot j\cdot}$		$\bar{x}_{\cdot p\cdot}$	$\bar{x}_{\cdot\cdot\cdot}$

Assume that each individual profile is a random vector sampled from a p-variate normal distribution with an arbitrary variance-covariance matrix,

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1p}\sigma_1\sigma_p \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2p}\sigma_2\sigma_p \\ \vdots & \vdots & & \vdots \\ \rho_{1p}\sigma_1\sigma_p & \rho_{2p}\sigma_2\sigma_p & \cdots & \sigma_p^2 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix}.$$

Also assume that the p variables have the same metric. This is necessary to give meaning to the question of whether the group profiles have the same

shape, and not because of any statistical considerations. This restriction results in no loss of generality if there already exists a large body of data on these p tests so that the standard deviations can be assumed known. For in this instance, equal metrics can be obtained by standardization of the p test scores.

The questions that are most often asked in profile analysis are:

(i) Are the groups on the same level, i.e., do the groups arise from populations having the same group means, namely, $E(\bar{x}_{..1}) = E(\bar{x}_{..2}) = \cdots = E(\bar{x}_{..g})$, where E denotes the expectation?

(ii) Do the groups have the same shape, i.e., do the groups arise from populations having parallel group profiles?

Another question that may be asked of these data, although not too frequently in profile analysis, is whether the p tests have the same means.

With regard to the question on shape, it becomes necessary to define a statistic which reflects the concept of equally shaped group profiles. In a larger sense, profiles having the same shape can be considered to be parallel curves. As Box [4] and Block, Levine, and McNemar [2] point out, parallelism can be measured by the group-test interaction mean square. That is, if the curves are parallel, the group-test interaction should be zero and the mean square should not differ significantly from an appropriate error mean square. If, on the other hand, the curves have different shapes, the interaction mean square should be significantly greater than the error mean square.

This is made clear by reference to two group profiles:

$$\bar{x}_{.11} , \bar{x}_{.21} , \cdots , \bar{x}_{.p1}$$

and

$$\bar{x}_{.12} , \bar{x}_{.22} , \cdots , \bar{x}_{.p2} .$$

Denote the corresponding differences between group means for each test by

$$d_1 , d_2 , \cdots , d_p .$$

If the two profiles are parallel it is clear that $d_1 = d_2 = d_3 = \cdots = d_p$. On the other hand, if $d_1 = d_2 = d_3 = \cdots = d_p$, then the two profiles must be parallel. Hence, a necessary and sufficient condition that the two group profiles possess the same shape is that

$$d_1 = d_2 = d_3 = \cdots = d_p .$$

But the equality of these differences is exactly what is meant by no interaction between groups and tests, and the extent to which these differences are unequal corresponds to the existence of the group-test interaction. Therefore a test of the group-test interaction is also a test of whether group profiles have the same shape.

Tests of Significance In the Mixed Model

If the p test scores have equal variances and are independent (or, at

most, are equally correlated in pairs), so that

$$\Sigma_0 = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}$$

or

$$\Sigma_1 = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix},$$

then the given scheme constitutes the classical mixed model for g samples, with proportionate numbers of observations among the samples. The appropriate analysis of variance breakdown is shown in Table 1. The analysis under either of the above assumptions on the covariance matrix follows along classical lines. The F_1 , F_2 , and F_3 statistics used to test hypotheses of homogeneity of test (variable) means, of group means (level) and the nonexistence of a group-test interaction (equal shapes of group profiles), respectively, are exact.

If, on the other hand, the validity of these two models is suspect, on the basis either of prior evidence or of a statistical test, the given F ratios are not distributed like the tabulated F distribution. In this situation where the covariance matrix is assumed to be arbitrary and given by Σ, Roy [13], Rao [12], and others have approached the problem through the multivariate analysis of variance. However, it is of interest, and possibly of considerable practical importance, to investigate the distribution of the computed F statistics.

Tests of Significance for Arbitrary Covariance Matrix

Geisser and Greenhouse [8], in extending to several groups Box's work [5, 6] relative to one group, have shown that Q_1 and Q_4 are each independent of Q_5 , and Q_2 is independent of Q_3 . They have also shown that, under the null hypothesis,

$$E(Q_1) = A, \quad \text{say}, \quad E(Q_4) = (g - 1)A, \qquad E(Q_5) = (N - g)A,$$

and

$$E(Q_2) = (g - 1)B, \quad \text{say}, \quad E(Q_3) = (N - g)B.$$

Table 2 gives the mean square (M.S.) and the expectations of the mean

TABLE 1

Analysis of Variance

Source	d.f.	Sum of Squares	F
Tests	p-1	$Q_1 = N\sum_{j=1}^{p}(\bar{x}_{.j.} - \bar{x}_{...})^2$	$F_1 = \frac{(N-g)\ Q_1}{Q_5}$
Groups	g-1	$Q_2 = p\sum_{k=1}^{g} n_k(\bar{x}_{..k} - \bar{x}_{...})^2$	$F_2 = \frac{(N-g)\ Q_2}{(g-1)\ Q_3}$
Individuals (within Groups)	N-g	$Q_3 = p\sum_{k=1}^{g}\sum_{i=1}^{n_k}(\bar{x}_{i.k} - \bar{x}_{..k})^2$	
Group x Tests	(p-1)(g-1)	$Q_4 = \sum_{k=1}^{g}\sum_{j=1}^{p} n_k(\bar{x}_{.jk} - \bar{x}_{.j.} - \bar{x}_{..k} + \bar{x}_{...})^2$	$F_3 = \frac{(N-g)\ Q_4}{(g-1)\ Q_5}$
Indiv. x Tests (within Groups)	(p-1)(N-g)	$Q_5 = \sum_{k=1}^{g}\sum_{i=1}^{n_k}\sum_{j=1}^{p}(x_{ijk} - \bar{x}_{.jk} - \bar{x}_{i.k} - \bar{x}_{..k})^2$	
Total	Np-1	$Q_6 = \sum_{k}\sum_{i}\sum_{j}(x_{ijk} - \bar{x}_{...})^2$	

square (E.M.S.) for each of the five sources of variation in the analysis of variance.

From the results presented in Tables 1 and 2, it follows that each of the three F ratios, F_1 , F_2 , and F_3 , is a ratio of two independent mean squares

TABLE 2

Analysis of Variance

Source	M.S.	E.M.S.
Tests	$(p-1)^{-1}Q_1$	$(p-1)^{-1}A$
Groups	$(g-1)^{-1}Q_2$	B
Individuals within Groups	$(N-g)^{-1}Q_3$	B
Groups x Tests	$(p-1)^{-1}(g-1)^{-1}Q_4$	$(p-1)^{-1}A$
Individuals x Tests within Groups	$(p-1)^{-1}(N-g)^{-1}Q_5$	$(p-1)^{-1}A$

with the same expectations under the null hypothesis. Making use of the fact that each of the quadratic forms involved in the three F statistics is exactly distributed like a linear sum of independent χ^2 variables with the same degrees of freedom (theorem 6.1, Box [5, 6]), F_1 is approximately distributed like $F[(p-1)\epsilon, (p-1)(N-g)\epsilon]$, F_3 is approximately distributed like $F[(p-1)(g-1)\epsilon, (p-1)(N-g)\epsilon]$, and F_2 is exactly distributed like $F(g-1, N-g)$, where

$$\epsilon = p^2(\bar{\sigma}_{tt} - \bar{\sigma}_{..})^2/(p-1)(\Sigma\Sigma\sigma_{ts}^2 - 2p\Sigma\bar{\sigma}_{t.}^2 + p^2\bar{\sigma}_{..}^2);$$

σ_{ts} are the elements of the matrix Σ, $\bar{\sigma}_{tt}$ is the mean of the diagonal terms, $\bar{\sigma}_{t.}$ is the mean of the tth row (or tth column), and $\bar{\sigma}_{..}$ is the grand mean. Thus, the effect of the arbitrary variance-covariance matrix, which must be the same from group to group, is to assess the significance of the F_1 and F_3 statistics in the ordinary tabulated F distribution but with reduced degrees of freedom. The F_2 test on group means, it will be noted, remains unchanged from the standard F test since it results from a one-way analysis of variance with all observations having the same variance.

The reduction in the degrees of freedom for this approximate test is a function of the elements of the population variance-covariance matrix. This is almost never known, and therefore ϵ will have to be estimated from the sample variances and covariances. However, the effect of using an estimated ϵ on the approximate F distributions involved is unknown. Hence, unless the variance-covariance matrix is estimated with a large number of degrees of freedom, use of the conservative test given below is suggested.

A Conservative Test

The preceding approximate procedure requires some computations on the elements of a known variance-covariance matrix. In many profile problems, the number of tests may be as high as 50 if not more. This results in a 50×50 matrix, necessitating some laborious arithmetic. Furthermore, in almost all problems variances and covariances are unknown and the extent to which ϵ is changed by using sample estimates has not been investigated. As a result it is useful to obtain a lower bound on ϵ; it can be shown that

$$\epsilon > \frac{1}{p-1}.$$

This minimum value of ϵ is independent of the elements of the variance-covariance matrix.

With this new correction to the degrees of freedom, the F_1 and F_3 statistics are now judged for significance by entering the tabulated F distribution with 1 and $N - g$ degrees of freedom and with $g - 1$ and $N - g$ degrees of freedom respectively. These tests are called conservative since the minimum value of ϵ gives the maximum reduction in degress of freedom.

An Example

Five groups of mothers, classified into their groups according to some external criteria, were given a maternal attitude questionnaire containing 23 scales. For purposes of this illustration, six of these scales have been selected. Thus $p = 6$, $g = 5$, and $N = 128$. The group profiles and group means are given in Table 3.

The five variance-covariance matrices were first tested for homogeneity. The likelihood ratio test, the multivariate analogue of Bartlett's test for

TABLE 3

Mean Profiles for Five Groups of Mothers on Selected Scales of a Maternal Attitude Questionnaire*

Groups	No. of Mothers	Scale 1	3	6	9	13	14	Group Mean
A	59	17.02	10.97	13.24	11.47	9.80	15.44	12.99
B	13	17.92	13.85	17.23	14.00	12.23	17.38	15.44
C	15	18.87	11.60	14.13	8.93	8.27	17.73	13.26
D	32	16.75	14.47	15.41	11.78	9.91	15.94	14.04
E	9	18.33	10.78	13.89	14.44	12.11	18.78	14.72
All Groups	128	17.35	12.20	14.34	11.72	10.05	16.27	13.65

*We are indebted to Dr. Richard Q. Bell of the Laboratory of Psychology, National Institute of Mental Health, for permitting us to use part of his data for this example.

homogeneity of variances, can be found in Box [3, 4]. (Kullback [11] derives an equivalent test through information theory.) The test statistic is

$$M = N \log_e | S | - \sum_{i=1}^{5} n_i \log_e | S_i | = 112.6565.$$

In the above $| S |$ is the determinant of the pooled variance-covariance matrix, and $| S_i |$ is the determinant of the sample variance-covariance matrix in the ith group. Now compute

$$A_1 = \frac{2p^2 + 3p - 1}{6(p + 1)(g - 1)} \left(\sum_{i=1}^{g} \frac{1}{n_i} - \frac{1}{N} \right) = .17012,$$

and

$$f_1 = \tfrac{1}{2}p(p + 1)(g - 1) = 84,$$

and enter $(1 - A_1)M = 93.4$ in the χ^2 distribution with 84 degrees of freedom. Since the probability of getting this value of χ^2 or larger is fairly high, the null hypothesis of equal variance-covariance matrices is not rejected.

An estimate of the matrix Σ is given by the pooled variance-covariance matrix

$$S = \begin{bmatrix} 3.100 & .101 & -.279 & -.083 & -.009 & 1.557 \\ .101 & 5.780 & 1.013 & -.114 & -1.014 & .039 \\ -.279 & 1.013 & 5.560 & 1.039 & 1.366 & -.169 \\ -.083 & -.114 & 1.039 & 5.600 & 3.080 & .258 \\ -.009 & -1.014 & 1.366 & 3.080 & 6.820 & .222 \\ 1.557 & .039 & -.169 & .258 & .222 & 5.170 \end{bmatrix}.$$

Consider now whether the hypothesis of equal variances and equal covariances is consistent with S. The best estimate of the uniform variance-covariance matrix under this hypothesis is given by

$$S_1 = \begin{bmatrix} 5.3888 & .467 & \cdots & .467 \\ .467 & 5.338 & \cdots & .467 \\ \vdots & \vdots & & \vdots \\ .467 & .467 & \cdots & 5.338 \end{bmatrix},$$

where the diagonal element is an average of the 6 variances in S and the covariance is an average of the $15[\frac{1}{2}p(p - 1)]$ covariances in S. The reason for testing this hypothesis is that if S_1 is consistent with the data then classical analysis of variance procedures are applicable. The test used is again a likelihood ratio test, also given by Box [3, 4]. The test statistic is

$$M = -(N - g) \log_e \frac{|S|}{|S_1|} = -123 \log \frac{10{,}806.42}{21{,}040.49} = 81.956,$$

where $(N - g) = 123$ is the degrees of freedom entering into the computation of any element in S or S_1. Now compute

$$A_1 = \frac{p(p+1)^2(2p-3)}{6(N-g)(p-1)(p^2+p-4)} = .01887,$$

and

$$f_1 = (p^2 + p - 4)/2 = 19,$$

and enter $(1 - A_1)M = 80.4$ in the χ^2 tables with $f_1 = 19$ degrees of freedom. The probability of this result is well below .001; the hypothesis of equal variances and equal covariances must be rejected.

The analysis of variance yields the numerical results of Table 4.

TABLE 4

Analysis of Variance

Source	d.f.	SS	M.S.	F
Tests	5	5092.56		
Groups	4	509.12	127.28	$F_2 = 16.51$
Individuals within Groups	123	948.41	7.71	
Groups x Tests	20	644.74	32.24	$F_3 = 6.63$
Individuals x Tests within Groups	615	2991.04	4.86	

Of primary interest is the test of the homogeneity of group profiles, which is a test for the existence of the group-test interaction. For this purpose enter the F_3 value in the F table with $(g - 1)(p - 1)\epsilon$ and $(N - g)(p - 1)\epsilon$, or with 20ϵ and 615ϵ, degrees of freedom. From the previous formula, and the elements in the S matrix, ϵ is estimated to be .8194. Therefore the effective degrees of freedom are 16 and 503. The observed $F_3 = 6.63$ is greater than the .001 point for F with 15 and 120 degrees of freedom. One therefore rejects the hypothesis of no interaction and concludes that the mean profiles differ in shape from group to group.

The conservative test, which of course does not require the computation of ϵ, would enter $F_3 = 6.63$ in the F tables with $g - 1 = 4$ and $N - g = 123$ degrees of freedom. The .001 point for F with these degrees of freedom is 4.95. In this case, therefore, the conservative test yields the same conclusion

as the approximate test, namely, the probability that the group profiles differ in shape due to chance is less than .001.

The groups clearly differ with regard to levels as can be seen from the very large F_2 value.

Other Procedures

The foregoing procedures present approximate and conservative tests of significance resulting from the analysis of variance utilizing readily available tables of the F distribution. As mentioned earlier there are available exact procedures in the multivariate analysis of variance. These procedures lead to exact tests of the general hypothesis in multivariate analysis of the equality of vector means among g populations and of the existence of the group-item or group-test interaction of interest in profile analysis. However, all of these procedures require laborious computations involving the inversion of $(p \times p)$ matrices (p equal to the number of tests or items) and the computation of latent roots or the evaluation of determinants. A further complication is the lack of tabled probability values for the appropriate test statistics. Recently, however, distribution tables have appeared relating to the approach of multivariate analysis initially taken by Roy [13]. Under this view, the distribution of the test statistic is dependent upon the distribution of the maximum characteristic root of certain matrices. The most comprehensive tables or charts thus far available are those given by Heck [9]. Heck, incidentally, specifically considers the problem of profile analysis.

The case for two groups will be developed in some detail to illustrate the principles involved and then the extension to g groups as given by Heck will be summarized briefly. The former situation leads to Hotelling's generalized T^2 statistic and is implied in the literature on multivariate analysis.

In the previous notation, x_{ijk} is an observation on item j for individual i in group k, and $\bar{x}_{.jk}$ is the mean of character j in group k. The range of subscripts here is $k = 1, 2; j = 1, 2, \cdots, p$; and $i = 1, 2, \cdots, n_k$. As before, assume that the random vector $x'_{i(k)} = (x_{i1k}, \cdots, x_{ipk})$ is $N(\mu_{(k)}, \Sigma)$, that is, the p variables have a multivariate normal distribution in population k with mean vector $\mu'_{(k)} = (\mu_{1k}, \cdots, \mu_{pk})$ and variance-covariance matrix Σ which is common to the g populations. The hypothesis to be tested for $g = 2$ is

$$\mu_{11} - \mu_{12} = \mu_{21} - \mu_{22} = \cdots = \mu_{p1} - \mu_{p2} .$$

Transform the p variates in x to $p - 1$ variates in y as follows (see [1], pp. 110–112 and [12], pp. 239–244):

$$y = \begin{bmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & & \vdots \\ c_{p-1,1} & \cdots & c_{p-1,p} \end{bmatrix} x$$

such that $\sum_{s=1}^{p} c_{rs} = 0$. The matrix C, subject to the restriction, can be perfectly arbitrary. For example,

$$C = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix}$$

subtracts $x_2 , \cdots , x_p$ from the first variate resulting in $y_1 = x_1 - x_2 , y_2 = x_1 - x_3 , \cdots , y_{p-1} = x_1 - x_p$. Or,

$$C = \frac{1}{-p} \begin{bmatrix} p-1 & -1 & -1 & \cdots & -1 & -1 \\ -1 & p-1 & -1 & \cdots & -1 & -1 \\ \vdots & & & & & \vdots \\ -1 & -1 & -1 & \cdots & p-1 & -1 \end{bmatrix},$$

which in effect subtracts from each of the p variates their mean $\bar{x} = (1/p) \sum_{i=1}^{p} x_i$ resulting in $y_1 = x_1 - \bar{x}, \cdots , y_{p-1} = x_{p-1} - \bar{x}$. Using the first transformation above, the vector $y'_{(k)} = (y_{1k} , \cdots , y_{(p-1)k})$ is multivariate normal with mean $\eta'_{(k)} = (\eta_{1k}, \cdots , \eta_{(p-1)k})$, $\eta_{jk} = \mu_{1k} - \mu_{(j+1)k}$, and variance-covariance matrix $C\Sigma C'$, where the prime denotes the transpose of a matrix.

After transforming the p x-variates into the $p - 1$ y-variates for each of the $n = n_1 + n_2$ individuals, the group means in the y's are

$$\bar{y}_{.11} , \bar{y}_{.21} , \cdots , \bar{y}_{.(p-1)1}$$

$$\bar{y}_{.12} , \bar{y}_{.22} , \cdots , \bar{y}_{.(p-1)2} ,$$

and the pooled sample variance-covariance matrix in the y's, $W = [w_{rs}]$, where

$$w_{rs} = \frac{1}{n_1 + n_2 - 2} \left\{ \sum_{i=1}^{n_1} (y_{ir1} - \bar{y}_{.r1})(y_{is1} - \bar{y}_{.s1}) + \sum_{i=1}^{n_2} (y_{ir2} - \bar{y}_{.r2})(y_{is2} - \bar{y}_{.s2}) \right\} ,$$

and $r, s = 1, 2, \cdots , p - 1$. It is easily seen that the null hypothesis in the x's is equivalent to the following hypothesis in the y's:

$$\eta_{j1} = \mu_{11} - \mu_{(j+1)1} = \eta_{j2} = \mu_{12} - \mu_{(j+1)2} , \qquad j = 1, 2, \cdots , (p - 1);$$

i.e., $\eta_{(1)} = \eta_{(2)}$. But this is the general hypothesis of multivariate analysis of the equality of mean vectors for two groups and it is well known that the appropriate statistic to test this hypothesis is T^2. Therefore

$$T^2 = \frac{n_1 n_2}{n_1 + n_2}(\bar{y}_{(1)} - \bar{y}_{(2)})'W^{-1}(\bar{y}_{(1)} - \bar{y}_{(2)})$$

$$= \frac{n_1 n_2}{n_1 + n_2} \sum_{r}^{p-1} \sum_{s}^{p-1} w^{rs}(\bar{y}_{.r1} - \bar{y}_{.r2})(\bar{y}_{.s1} - \bar{y}_{.s2}),$$

where w^{rs} is element rs in the inverse matrix W^{-1}. This statistic has the T^2 distribution with $n_1 + n_2 - 2$ degrees of freedom.

To test the hypothesis at level α, enter

$$\frac{T^2(n_1 + n_2 - p)}{(n_1 + n_2 - 2)(p - 1)}$$

in the F table with $p - 1$ and $n_1 + n_2 - p$ degrees of freedom. If

$$\frac{T^2(n_1 + n_2 - p)}{(n_1 + n_2 - 2)(p - 1)} > F_\alpha(p - 1, n_1 + n_2 - p)$$

reject the hypothesis; otherwise accept.

The general case for g populations, of which the above is a special case, is given by Heck [9]. The extension is obvious. From the g by $p - 1$ table of group means, one computes the between groups sums of squares and cross products to obtain the elements of the matrix B, say. Thus element rs of this matrix is

$$b_{rs} = \sum_{k=1}^{g} n_k(\bar{y}_{.rk} - \bar{y}_{.r.})(\bar{y}_{.sk} - \bar{y}_{.s.}) = \sum_{k=1}^{g} n_k \bar{y}_{.rk}\bar{y}_{.sk} - n\bar{y}_{.r.}\bar{y}_{.s.},$$

where $r, s = 1, 2, \cdots, (p - 1)$. For the error matrix W, compute similarly the sums of products, so that

$$w_{rs} = \sum_{k=1}^{g} \sum_{i=1}^{n_k} (y_{irk} - \bar{y}_{.rk})(y_{isk} - \bar{y}_{.sk}) = \sum_{k=1}^{g} \sum_{i}^{n_k} y_{irk}y_{isk} - \sum_{k=1}^{g} n_k \bar{y}_{.rk}\bar{y}_{.sk}.$$

In the above formula, $\bar{y}_{.rk} = n_k^{-1} \sum_{i=1}^{n_k} y_{irk}$. The various test statistics proposed are proportional to some function of the product matrix BW^{-1}.

In the literature on multivariate analysis, there have been three approaches to the distribution problem. Wilks [17], starting with the likelihood ratio criterion, derived the test statistic $| I + BW^{-1} |^{-1}$, which is obviously equal to the inverse of the product of the characteristic roots of $(I + BW^{-1})$, I being the identity matrix. Hotelling [10] has proposed the distribution of tr BW^{-1} or of the sum of the characteristic roots of BW^{-1}. Roy [13] has proposed the consideration of the distribution of the maximum characteristic root of BW^{-1}. For a further discussion of these three points of view consult Anderson ([1], pp. 221–224). There are no probability tables available for the first two test statistics although the exact cumulative distribution of the determinantal statistic is given by an infinite series of χ^2's, the first term of which, for any reasonable N, gives an excellent approximation to the whole

tive test provides a procedure which is more than "rough and ready" and yet saves considerable time since it does not require a matrix inversion nor even the computation of a covariance matrix. This is particularly true when p, the number of variables, is large and the number of samples is greater than two.

The question of electronic computers is another matter. Given the availability of a classical analysis of variance program and the availability of a combined program to carry out the multivariate analysis of variance involving the between samples variance-covariance matrix, the inverse of the error variance-covariance matrix, and the extraction of the maximum latent root of the product of the two matrices, it is very likely that the former would require less machine time. However, the difference is probably of no practical importance and the exact procedure should be used.

A more fundamental question relates to a comparison of the two exact tests involved. Are the multivariate analysis of variance procedures depending upon the distribution of BW^{-1} more powerful against all alternatives than the distribution of the ratio of linear sums of χ^2 variates? It is not clear that this is so, particularly with regard to the analysis of profile shapes where the former procedures must reduce the dimensionality of the random vector.

If one does decide to use the F tests in an analysis, the following series of steps are suggested. After finding the traditional analysis of variance table, first test the appropriate observed F value in the F distribution with full, i.e., unreduced, degrees of freedom. For F_3 , for example, this would be F with $(p - 1)(g - 1)$ and $(p - 1)(N - g)$ degrees of freedom. If F_3 is smaller than the α critical point, one can stop here, for the null hypothesis will not be rejected with further manipulation of degrees of freedom. If the observed F is significant, then one proceeds to the conservative test where the degrees of freedom are reduced by a factor equal to $1/(p - 1)$. For F_3 , the appropriate F distribution is $F(g - 1, N - g)$. If this test leads to significance at the α level, one can at this point reject the null hypothesis without further testing. However if the conservative test is not significant then it is suggested that the ϵ be estimated from the variance-covariance matrix and the approximate test be carried out.

Number of Individuals Less than the Number of Variables

As indicated in the introduction, in the case of one group, if $(n - 1) < p$, or in the case of g groups, if $(N - g) < p$, it is not possible to apply multivariate procedures. The reason of course is that the error matrix, W, is singular. Such situations are not too uncommon, especially in research in clinical psychology and psychiatry. Clearly the approximate F tests presented are not applicable either since the reduction in degrees of freedom is dependent upon the elements of a singular matrix. However, the conservative test can be applied.

Unequal Variance-Covariance Matrices

Perhaps one of the most important uses of the conservative test is in the situation where one cannot assume the equality of the unknown variance-covariance matrices in the p-variate normal populations being sampled. For this case, there are no exact procedures available. It will be noted that this case, $p = 1$ and $g = 2$, reduces to the Fisher-Behrens problem.

Here, in order for the F statistics to be unbiased, it is necessary to work with equal sample sizes in the groups, i.e., $n_1 = \cdots = n_g = n$. Therefore, $N = \Sigma n_k = gn$. It can again be shown that the respective numerator and denominator quadratic forms entering into F_1, F_2, and F_3 are independent and have the same expectations. Now, however, when an F distribution is used to approximate these F statistics (see [5], theorem 6.1), it turns out that there are different factors reducing the numerator and denominator degrees of freedom, and these in turn differ for the three F statistics. Here again it can be shown that these ϵ's have lower limits which when applied to the appropriate degrees of freedom result in a conservative test for assessing the significance of F_1, F_2, and F_3 by entering these in the F distribution with 1 and $n - 1$ degrees of freedom. It is of interest that the F_2 test, when $p = 1$ and $g = 2$, is a conservative test for the various approximate solutions given to the Fisher-Behrens problem of testing the equality of two means with unequal variances (e.g., [15], p. 295).

REFERENCES

[1] Anderson, T. W. *Introduction to multivariate statistical analysis.* New York: Wiley, 1958.
[2] Block, J., Levine, L., and McNemar, Q. Testing for the existence of psychometric patterns. *J. abnorm. soc. Psychol.*, 1951, **46**, 356–359.
[3] Box, G. E. P. A general distribution theory for a class of likelihood criteria. *Biometrika*, 1949, **36**, 317–346.
[4] Box, G. E. P. Problems in the analysis of growth and wear curves. *Biometrics*, 1950, **6**, 362–389.
[5] Box, G. E. P. Some theorems on quadratic forms applied in the study of analysis of variance problems: I. Effect of inequality of variance in the one-way classification. *Ann. math. Statist.*, 1954, **25**, 290–302.
[6] Box, G. E. P. Some theorems on quadratic forms applied in the study of analysis of variance problems: II. Effects of inequality of variance and of correlation between errors in the two-way classification. *Ann. math. Statist.*, 1954, **25**, 484–498.
[7] Eisenhart, C. The assumptions underlying the analysis of variance. *Biometrics*, 1947, **3**, 1–21.
[8] Geisser, S. and Greenhouse, S. W. An extension of Box's results on the use of the F distribution in multivariate analysis. *Ann. math. Statist.*, 1958, **29**, 885–891.
[9] Heck, D. L. Some uses of the distribution of the largest root in multivariate analysis. Inst. Statist. Univ. North Carolina, Mimeo. Ser. No. 194, 1958.
[10] Hotelling, H. A generalized T test and measure of multivariate dispersion. *Proceedings of the second Berkeley symposium on mathematical statistics and probability.* Berkeley: Univ. Calif. Press, 1951, 23–42.

[11] Kullback, S. An application of information theory to multivariate analysis, II. *Ann. math. Statist.*, 1956, **27**, 122–146.
[12] Rao, C. R. *Advanced statistical methods in biometric research.* New York: Wiley, 1952.
[13] Roy, S. N. On a heuristic method of test construction and its use in multivariate analysis. *Ann. math. Statist.*, 1953, **24**, 220–238.
[14] Scheffé, H. A "mixed model" for the analysis of variance. *Ann. math. Statist.*, 1956, **27**, 23–36.
[15] Welch, B. L. Note on Mrs. Aspin's Tables and on certain approximations to the tabled functions. *Biometrika*, 1949, **36**, 293–296.
[16] Wilk, M. B. and Kempthorne, O. Fixed, mixed, and random models. *J. Amer. statist. Ass.*, 1955, **50**, 1144–1167.
[17] Wilks, S. S. Certain generalizations in the analysis of variance. *Biometrika*, 1932, **24**, 471–494.

Manuscript received 8/21/58

Revised manuscript received 12/1/58