

AD-A121 470

TESTING HYPOTHESES WHICH ARE UNIONS OF LINEAR SUBSPACES

1/1

(U) FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS

R L BERGER ET AL. SEP 82 FSU-STATISTICS-M-639

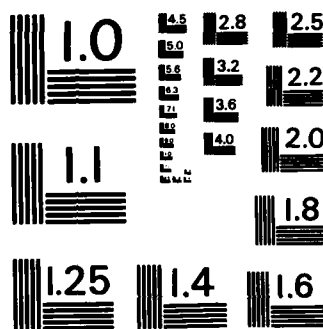
UNCLASSIFIED

ARO-16713. 22-MA

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A121470

TESTING HYPOTHESES WHICH ARE
UNIONS OF LINEAR SUBSPACES¹

by

Roger L. Berger² and Dennis F. Sinclair³

FSU Statistics Report No. M-639
USARO Report No. D-55

September, 1982

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

DTIC
ELECTE
NOV 9 1982
H

¹This research was supported in part by U. S. Army Research Office Contract DAAG-29-79-C-0158.

²Roger L. Berger is currently a Visiting Associate Professor at the Department of Statistics, North Carolina State University, Raleigh, NC 27650.

³Dennis F. Sinclair is currently at CSIRO Private Mailbag, P.O. Aitkenvale, Queensland, Australia 4814.

DTIC FILE COPY

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

TESTING HYPOTHESES WHICH ARE
UNIONS OF LINEAR SUBSPACES

by

Roger L. Berger and Dennis F. Sinlair

ABSTRACT

→ The likelihood ratio test (LRT) for hypotheses which are unions of linear subspaces is derived for the normal theory linear model. A more powerful variant of the LRT is proposed for the case in which the subspaces are not all of the same dimension. A theorem is proved which may be used to identify hypotheses which are unions of linear subspaces. Some hypotheses, of particular relevance in ecology, concerning the spacings between normal means are shown to be unions of linear subspaces and are therefore testable using the LRT. Finally, the computation of the LRT statistic is discussed. ←



Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Avail and/or	
Dist	Special
A	

1. Introduction.

Character displacement is an ecological process by which coexisting species diverge in size to reduce competition (see Grant, 1972, and Sinclair, Mosimann, and Meeter, 1982, for details). The size variable is chosen to reflect competition between species, e.g., bill length for birds. A typical situation is as follows. An island has been colonized by J species within some family from the nearby mainland. If character displacement is occurring, the species should be more dissimilar on the island, where the variety of resources is limited, than on the mainland. Recently, there has been much controversy in the ecological literature concerning the existence of character displacement (Strong, Szyska, and Simberloff, 1979; Grant and Abbott, 1980; Hendrickson, 1981; and Strong and Simberloff, 1981). We list below four hypotheses arising in character displacement studies. These hypotheses are all in a class for which we derive the LRT. Tests of these hypotheses may help resolve the controversy surrounding character displacement.

Let the mean size measurements be denoted μ_{ij} , $i = 1$ (island), 2 (mainland), $j = 1, \dots, J$ (species). Let the ordered species means on the mainland and the island be denoted

$$\mu_{i(1)} \leq \mu_{i(2)} \leq \dots \leq \mu_{i(J)}, \quad i = 1, 2.$$

The four character displacement hypotheses relate to these ordered means. They are the following.

$$H_0^1: \mu_{1(j+1)} - \mu_{1(j)} = \mu_{1(J+1-j)} - \mu_{1(J-j)}, \quad j = 1, \dots, [(J-1)/2],$$

where $[s]$ denotes the greatest integer less than or equal to s .

$$H_0^2: \mu_1(j+1) - \mu_1(j) = c_j(\mu_1(2) - \mu_1(1)), \quad j = 2, \dots, J-1,$$

where $c_2, \dots, c_{J-1}$ are specified positive constants with $c_j = c_{J-j}$, $j = 2, \dots, J-2$, and $c_{J-1} = 1$.

$$H_0^3: \mu_1(j+1) - \mu_1(j) = \mu_2(j+1) - \mu_2(j), \quad j = 1, \dots, J-1.$$

$$H_0^4: \begin{aligned} \mu_1(j+1) - \mu_1(j) &= \mu_2(j+1) - \mu_2(j), & j &= 1, \dots, J-1 \\ \mu_1(j+1) - \mu_1(j) &= c_j(\mu_1(2) - \mu_1(1)), & j &= 2, \dots, J-1 \end{aligned}$$

where the c_j are as in H_0^2 .

Hypothesis H_0^1 states that the island means are arranged in a symmetric fashion. Hypothesis H_0^2 specifies the relative sizes of the spacings between the means in the symmetric pattern. A commonly claimed manifestation of character displacement is equally spaced species sizes (Strong, Szyska, and Simberloff, 1979) corresponding to $c_j = 1$, $j = 2, \dots, J-1$, in H_0^2 . Hypothesis H_0^3 states that the corresponding spacings on the island and the mainland are equal while the relative sizes of the spacings are specified in H_0^4 .

In Section 2 we derive the LRT for hypotheses which are unions of subspaces. (Throughout this paper, "subspace" refers to a linear subspace.) In Section 3 we prove a theorem which may be used to identify hypotheses with this property. We use the theorem to show that the above four hypotheses are unions of subspaces and thus may be tested with the LRT of Section 2. Finally, in Section 4 we discuss how the test statistic may be computed.

2. Likelihood Ratio Test.

Let $X_1, \dots, X_K$ denote independent normal random variables with means $\xi_1, \dots, \xi_K$ and common variance σ^2 . We assume that $\xi = (\xi_1, \dots, \xi_K)'$ lies in ω , a subspace of R^K of dimension $J < K$. For example, $X_1, \dots, X_K$ may be comprised of independent samples from J populations ($J < K$). We will discuss testing hypotheses about ξ which are unions of subspaces of ω . In this section we derive the LRT of

$$(2.1) \quad H_0: \xi \in \omega_0 \quad \text{versus} \quad H_1: \xi \in \omega - \omega_0$$

where $\omega_0 = \bigcup_{i=1}^m \omega_i$ and each ω_i is a q_i dimensional subspace of ω .

We will show that the critical value for the LRT is a multiple of a percentile from an F distribution. We will also describe a modification of the LRT which has a higher power than the LRT if at least two of the ω_i have different dimensions. Throughout we assume that (2.1) has been expressed in such a way that the ω_i are all distinct, that is, there do not exist i and j , $1 \leq i, j \leq m$, $i \neq j$, such that $\omega_i \subset \omega_j$.

Let the density of $\underline{X} = (X_1, \dots, X_K)'$ be denoted by

$$(2.2) \quad p(\underline{x}; \underline{\xi}, \sigma) = (2\pi\sigma^2)^{-K/2} \exp\left(-\sum_{i=1}^K (x_i - \xi_i)^2 / 2\sigma^2\right).$$

Let $\theta_i = \{(\underline{\xi}, \sigma): \underline{\xi} \in \omega_i, \sigma > 0\}$, $\theta_0 = \{(\underline{\xi}, \sigma): \underline{\xi} \in \omega_0, \sigma > 0\} = \bigcup_{i=1}^m \theta_i$

and $\theta = \{(\underline{\xi}, \sigma): \underline{\xi} \in \omega, \sigma > 0\}$. The LRT statistic for testing H_0 is defined as

$$(2.3) \quad \lambda(\underline{x}) = \frac{\sup_{\theta_0} p(\underline{x}; \underline{\xi}, \sigma)}{\sup_{\theta} p(\underline{x}; \underline{\xi}, \sigma)}.$$

If we let $\hat{\xi}$ denote the projection of $\underline{x}$ on ω and $\hat{\xi}_i$ denote the projection of $\underline{x}$ on ω_i , then $\lambda(\underline{x})$ is

$$\begin{aligned}
 \lambda(\underline{x}) &= \frac{\max_{1 \leq i \leq m} \sup_{\theta_i} p(\underline{x}; \xi, \sigma)}{\sup_{\theta} p(\underline{x}; \xi, \sigma)} \\
 &= \max_{1 \leq i \leq m} \frac{\sup_{\theta_i} p(\underline{x}; \xi, \sigma)}{\sup_{\theta} p(\underline{x}; \xi, \sigma)} \\
 (2.4) \quad &= \max_{1 \leq i \leq m} \left(\frac{\|\underline{x} - \hat{\xi}\|^2}{\|\underline{x} - \hat{\xi}_i\|^2} \right)^{K/2} \\
 &= \left(\frac{\|\underline{x} - \hat{\xi}\|^2}{\min_{1 \leq i \leq m} \|\underline{x} - \hat{\xi}_i\|^2} \right)^{K/2}
 \end{aligned}$$

where $\|\underline{y}\|^2 = \underline{y}'\underline{y}$. The third equality in (2.4) is a standard result from linear models theory. The last expression in (2.4) reflects the fact that the maximum likelihood estimate of ξ under H_0 is the projection of $\underline{x}$ on the nearest subspace ω_i . Since $\|\underline{x} - \hat{\xi}_i\|^2 = \|\underline{x} - \hat{\xi}\|^2 + \|\hat{\xi} - \hat{\xi}_i\|^2$, rejecting H_0 if $\lambda(\underline{x}) < c$ is equivalent to rejecting H_0 if $\lambda^*(\underline{x}) > c^*$ where

$$(2.5) \quad \lambda^*(\underline{x}) = \frac{\min_{1 \leq i \leq m} \|\hat{\xi} - \hat{\xi}_i\|^2}{\|\underline{x} - \hat{\xi}\|^2}.$$

The value of c^* which produces a size α test is given in Theorem 1.

Theorem 1. Let $F_{\alpha,a,b}$ denote the upper 100 α -percentile of an F distribution with a and b degrees of freedom. Let $c_{\alpha}^* = (\max_{1 \leq i \leq m} (J-q_i) F_{\alpha, J-q_i, K-J}) / (K-J)$. The test which rejects H_0 if $\lambda^*(\underline{X}) > c_{\alpha}^*$ is a size α test, that is, the test satisfies

$$(2.6) \quad \sup_{\theta_0} P_{\xi, \sigma}(\lambda^*(\underline{X}) > c_{\alpha}^*) = \alpha.$$

Proof. For any $\xi \in \omega_j$ and $\sigma > 0$,

$$\begin{aligned} P_{\xi, \sigma}(\lambda^*(\underline{X}) > c_{\alpha}^*) &\leq P_{\xi, \sigma}(\|\hat{\xi} - \hat{\xi}_j\|^2 / \|\underline{x} - \hat{\xi}\|^2 > c_{\alpha}^*) \\ &\leq P_{\xi, \sigma}(\|\hat{\xi} - \hat{\xi}_j\|^2 / \|\underline{x} - \hat{\xi}\|^2 > (J-q_j) F_{\alpha, J-q_j, K-J} / (K-J)) \\ &= \alpha. \end{aligned}$$

The last equality is true since a standard result from linear models theory states that for any $\xi \in \omega_j$ and $\sigma > 0$, $[(K-J)/(J-q_j)] \|\hat{\xi} - \hat{\xi}_j\|^2 / \|\underline{x} - \hat{\xi}\|^2$ has an F distribution with $J-q_j$ and $K-J$ degrees of freedom. Thus

$$(2.7) \quad \sup_{\theta_0} P_{\xi, \sigma}(\lambda^*(\underline{X}) > c_{\alpha}^*) \leq \alpha.$$

To prove the reverse inequality, let j be such that

$$(J-q_j) F_{\alpha, J-q_j, K-J} / (K-J) = \max_{1 \leq i \leq m} (J-q_i) F_{\alpha, J-q_i, K-J} / (K-J) = c_{\alpha}^*. \text{ We have}$$

assumed that ω_j is not a subset of ω_i for any $i \neq j$. So for every

$i = 1, \dots, m, i \neq j$, $\omega_i \cap \omega_j$ is a subspace of dimension at most $q_j - 1$.

The set, $\bigcup_{\substack{i=1 \\ i \neq j}}^m (\omega_i \cap \omega_j)$, cannot contain the q_j dimensional set ω_j since each

set in the union is at most $(q_j - 1)$ dimensional. Thus there exists ξ^* , such that $\xi^* \in \omega_j$ and $\xi^* \notin \omega_i$ for any $i = 1, \dots, m, i \neq j$.

Fix $\sigma > 0$. We will consider the sequence of parameter points $(k\xi^*, \sigma)$, $k = 1, 2, \dots$. Note that $k\xi^* \in \omega_j$ for every k since ω_j is a subspace. Let ξ_i^* denote the projection of ξ^* on ω_i . Then the projection of $k\xi^*$ on ω_i is $k\xi_i^*$ and $\|k\xi^* - k\xi_i^*\|^2 = k^2 \|\xi^* - \xi_i^*\|^2 \rightarrow \infty$ as $k \rightarrow \infty$ since $\|\xi^* - \xi_i^*\|^2 > 0$ for $i = 1, \dots, m, i \neq j$.

For $i = 1, \dots, m$, let $R_i = \{x: \|\hat{\xi} - \hat{\xi}_i\|^2 / \|\hat{x} - \hat{\xi}\|^2 > c_\alpha^*\}$. At $(k\xi^*, \sigma)$ $[(K-J)/(J-q_i)] \|\hat{\xi} - \hat{\xi}_i\|^2 / \|\hat{x} - \hat{\xi}\|^2$ has a noncentral F distribution with $J-q_i$ and $K-J$ degrees of freedom and noncentrality parameter $\delta_{k,i} = k^2 \|\xi^* - \xi_i^*\|^2 / \sigma^2$. For $i = 1, \dots, m, i \neq j$, $\delta_{k,i} \rightarrow \infty$ as $k \rightarrow \infty$. So for $i = 1, \dots, m, i \neq j$,

$$(2.8) \quad P_{k\xi^*, \sigma}(R_i) \rightarrow 1 \text{ as } k \rightarrow \infty.$$

On the other hand $[(K-J)/(J-q_j)] \|\hat{\xi} - \hat{\xi}_j\|^2 / \|\hat{x} - \hat{\xi}\|^2$ has a central F distribution with $J-q_j$ and $K-J$ degrees of freedom. Furthermore,

$c_\alpha^* = (J-q_j)F_{\alpha, J-q_j, K-J}/(K-J)$. Thus

$$(2.9) \quad P_{k\xi^*, \sigma}(R_j) = \alpha \quad \text{for every } k = 1, 2, \dots$$

Using (2.8) and (2.9) we obtain

$$\begin{aligned} \lim_{k \rightarrow \infty} P_{k\xi^*, \sigma}(\lambda^*(X) > c_\alpha^*) &= \lim_{k \rightarrow \infty} P_{k\xi^*, \sigma}\left(\bigcap_{i=1}^m R_i\right) \\ &= \lim_{k \rightarrow \infty} \left[1 - P_{k\xi^*, \sigma}\left(\bigcup_{i=1}^m R_i^c\right)\right] \\ &\geq 1 - \lim_{k \rightarrow \infty} \sum_{i=1}^m P_{k\xi^*, \sigma}(R_i^c) \end{aligned}$$

$$= 1 - (1-\alpha) - \lim_{k \rightarrow \infty} \sum_{\substack{i=1 \\ i \neq j}}^m P_{k\xi^*, \sigma}(R_i^c)$$

$$= \alpha - 0 = \alpha.$$

Since $(k\xi^*, \sigma) \in \theta_j \subset \theta_0$ for every k ,

$$(2.10) \quad \sup_{\theta_0} P_{\xi, \sigma}(\lambda^*(X) > c_\alpha^*) \geq \lim_{k \rightarrow \infty} P_{k\xi^*, \sigma}(\lambda^*(X) > c_\alpha^*) \geq \alpha.$$

Combining (2.7) and (2.10) yields (2.6). ||

We believe that in most applications all of the ω_i , $i = 1, \dots, m$, will have the same dimension q , in which case $c_\alpha = (J-q)F_{\alpha, J-q, K-J}/(K-J)$. This is the case for all the examples we discuss in Section 3. But if the dimensions of some of the ω_i differ, there is a modification of the LRT which is also a size α test and has higher power than the LRT. This test is described in Theorem 2; it is the LRT if all the ω_i have the same dimension.

Theorem 2. Let $F_i = [(K-J)/(J-q_i)] \|\hat{\xi} - \hat{\xi}_i\|^2 / \|\hat{x} - \hat{\xi}\|^2$, $i = 1, \dots, m$. The test λ^{**} which rejects H_0 if and only if $F_i > F_{\alpha, J-q_i, K-J}$ for every $i = 1, \dots, m$ is a size α test. The test λ^{**} has a power which is greater than or equal to the power of the LRT for every $(\xi, \sigma) \in \theta$.

Proof. The proof that λ^{**} is a size α test is almost identical to the proof given in Theorem 1 that λ^* is a size α test. In this case any ω_i , $i = 1, \dots, m$, can play the special role played by ω_j in the second half of Theorem 1.

The $\{x: \lambda^*(x) > c_\alpha^*\} \subset \{x: F_i > F_{\alpha, J-q_i, K-J}, i = 1, \dots, m\}$. That is, the rejection region for the LRT is a subset of the rejection region for λ^{**} . Thus the power of λ^{**} is greater than or equal to the power of the LRT. ||

Unless all of the quantities $(J-q_i)F_{\alpha, J-q_i, K-J}/(K-J)$, $i = 1, \dots, m$, are equal, the rejection region for the LRT is a proper subset of the rejection region for λ^{**} and the power for the LRT is strictly smaller than the power of λ^{**} for every parameter in H_1 . This provides an example, like that of C. Stein (see Bickel and Doksum, 1977, p. 239), of a LRT whose power is everywhere dominated by the power of another test.

The test which rejects $H_{0i}: \xi \in \omega_i$ if $F_i > F_{\alpha, J-q_i, K-J}$ is a size α test of H_{0i} . The test of H_0 in Theorem 2 rejects H_0 if and only if for each i the test based on F_i rejects H_{0i} . Tests of this form have been called intersection-union tests by Gleser (1973); they have also been discussed by Berger (1982).

3. Hypotheses Consisting of Linear Subspaces.

In the remaining sections, we discuss some specific problems which fall into the general framework described in Section 2. These problems involve hypotheses about the spacings between normal means. The ecological problem which motivated our interest in these hypotheses was described in Section 1.

We will consider the following special case of the model presented in Section 2. Let X_{ijk} , $i = 1, \dots, I$, $j = 1, \dots, J_i$, $k = 1, \dots, K_{ij}$ denote $K = \sum_{i,j} K_{ij}$ independent normal observations. The mean of X_{ijk} is

μ_{ij} and all the X_{ijk} have a common variance of σ^2 . Let

$\underline{\mu} = (\mu_{11}, \dots, \mu_{1J_1}, \mu_{21}, \dots, \mu_{IJ_I})'$ and

$\underline{\xi} = (\mu_{11} \hat{L}_{K_{11}}', \dots, \mu_{1J_1} \hat{L}_{K_{1J_1}}', \mu_{21} \hat{L}_{K_{21}}', \dots, \mu_{IJ_I} \hat{L}_{K_{IJ_I}}')'$ where $\hat{L}_b$ is a

vector of b ones. In the formulation of Section 2 we would consider

$\underline{\xi} \in \omega \subset R^K$. Since here there is a one to one correspondence between $\underline{\mu}$ and $\underline{\xi}$, we may equivalently consider $\underline{\mu} \in R^J$ where $J = \sum_i J_i$. We now consider the subspaces ω_i , $i = 1, \dots, m$, as subspaces in R^J .

We will be concerned with permutations of $(\mu_{i1}, \dots, \mu_{iJ_i})$ for each $i = 1, \dots, I$. Let $\underline{\mu}_i = (\mu_{i1}, \dots, \mu_{iJ_i})'$ so that $\underline{\mu} = (\underline{\mu}_1', \dots, \underline{\mu}_I')'$. A map $\pi: R^J \rightarrow R^J$ is called a subpermutation if $\pi(\underline{\mu}) = (\pi_1(\underline{\mu}_1), \dots, \pi_I(\underline{\mu}_I))$ where $\pi_i(\underline{\mu}_i)$ is a permutation of $\underline{\mu}_i$, $i = 1, \dots, I$. There are $\prod_{i=1}^I J_i!$ subpermutations. A set $A \subset R^J$ is called subpermutation invariant if $\underline{\mu} \in A$ implies $\pi(\underline{\mu}) \in A$ for every subpermutation π . A set B is called a subpermutation of a set A if B is the image of A under some subpermutation π . It is easily verified that if A is a subspace then any subpermutation of A is also a subspace.

The following theorem may be used to identify hypotheses which are unions of subspaces. We shall use this theorem to show that the ordered mean hypotheses in Section 1 are unions of subspaces and hence are testable using the LRT of Section 2.

Theorem 3. Let ω_0 be a subpermutation invariant subset of R^J . Let H denote a $(J-q) \times J$ matrix of rank $J-q$. Let $N = \{\underline{\mu} \in R^J: H\underline{\mu} = \underline{Q}\}$. Let $O = \{\underline{\mu} \in R^J: \mu_{i1} \leq \mu_{i2} \leq \dots \leq \mu_{iJ_i}, i = 1, \dots, I\}$. If $N \cap O = \omega_0 \cap O$ and $N \subset \omega_0$ then $\omega_0 = \bigcup_{i=1}^m \omega_i$ where $m = \prod_{i=1}^I J_i!$, ω_1 is the q dimensional subspace N , and $\omega_2, \dots, \omega_m$ are the subpermutations of N . Thus ω_0 is the union of q dimensional subspaces.

Proof. Suppose $\underline{\mu} \in \bigcup_{i=1}^m \omega_i$. Then $\underline{\mu} = \pi(\underline{\mu}^*)$ for some $\underline{\mu}^* \in N$ and some sub-permutation π . Since $N \subset \omega_0$ and ω_0 is subpermutation invariant, $\underline{\mu} = \pi(\underline{\mu}^*) \in \omega_0$. Thus $\bigcup_{i=1}^m \omega_i \subset \omega_0$.

Now suppose $\underline{\mu} \in \omega_0$. Let $\underline{\mu}^*$ be a subpermutation of $\underline{\mu}$ such that $\mu_{i1}^* \leq \dots \leq \mu_{iJ_i}^*$, $i = 1, \dots, I$. By definition, $\underline{\mu}^* \in O$. Also, $\underline{\mu}^* \in \omega_0$ since ω_0 is subpermutation invariant. Thus, $\underline{\mu}^* \in \omega_0 \cap O = N \cap O \subset N$. Since $\underline{\mu}$ is a subpermutation of $\underline{\mu}^* \in N$, $\underline{\mu} \in \omega_i$ for some i . Thus $\omega_0 \subset \bigcup_{i=1}^m \omega_i$.

Finally, since $N = \omega_1$ is a q dimensional subspace of R^J , each of the subpermutations of N , $\omega_2, \dots, \omega_m$, is also a q dimensional subspace of R^J . ||

The subspaces $\omega_1, \dots, \omega_m$ defined in Theorem 3 will not be distinct. As will be seen in the following examples, the number of distinct subspaces will be much smaller than m (at most $m/2$, in fact). Recognizing this fact results in a saving of effort in the computation of the test statistic λ^* for which the minimum needs to be taken only over distinct subspaces. Taking the minimum over all m subspaces in Theorem 3 will, of course, give the same value of λ^* . It would just be inefficient since many of the subspaces are equal.

We now consider the four hypotheses about the spacings between normal means discussed in Section 1.

Example 1. (Symmetric Spacings). For this example, $I = 1$, so we will denote the means by $\mu_1, \dots, \mu_J$ and the ordered means by $\mu_{(1)} \leq \dots \leq \mu_{(J)}$. By the "symmetric spacings hypothesis" we mean

$$(3.1) H_0^1: \mu_{(j+1)} - \mu_{(j)} = \mu_{(J+1-j)} - \mu_{(J-j)}, \quad j = 1, \dots, [(J-1)/2].$$

We shall use Theorem 3 to verify that H_0^1 is the union of subspaces. Let H be a $[(J-1)/2] \times J$ matrix such that $H\mu = 0$ is equivalent to the conditions $\mu_{j+1} - \mu_j = \mu_{J+1-j} - \mu_{J-j}$, $j = 1, \dots, [(J-1)/2]$. If $\mu \in 0$ then $\mu_j = \mu_{(j)}$, $j = 1, \dots, J$, whence $H\mu = 0$ is equivalent to $\mu_{(j+1)} - \mu_{(j)} = \mu_{(J+1-j)} - \mu_{(J-j)}$, $j = 1, \dots, [(J-1)/2]$. That is, $N \cap 0 = \omega_0 \cap 0$ as required by Theorem 3. Now, to verify that $N \subset \omega_0$. Let $\mu \in N$, that is, $H\mu = 0$. Let $\mu^* = (\mu_1 + \mu_J)/2$. For any $j = 1, \dots, [(J-1)/2]$, $\mu_j - \mu^* = \mu_1 - \mu^* + \sum_{i=1}^{j-1} (\mu_{i+1} - \mu_i) = \mu^* - \mu_J + \sum_{i=1}^{j-1} (\mu_{J+1-i} - \mu_{J-i}) = \mu^* - \mu_{J+1-j}$. So, each pair, μ_j and μ_{J+1-j} , is symmetrically placed about μ^* . (If J is odd, $[(J-1)/2] + 1 = J - [(J-1)/2]$ and $\mu_{[(J-1)/2]+1} = \mu_{J-[(J-1)/2]} = \mu^*$.) If $\mu_r = \mu_{(j)}$ and $\mu_s = \mu_{(j+1)}$ then $\mu_{J+1-r} = \mu_{(J+1-j)}$ and $\mu_{J+1-s} = \mu_{(J-j)}$. Thus $\mu_{(j+1)} - \mu_{(j)} = \mu_s - \mu^* - (\mu_r - \mu^*) = \mu^* - \mu_{J+1-s} - (\mu^* - \mu_{J+1-r}) = \mu_{(J+1-j)} - \mu_{(J-j)}$. Therefore, $\mu \in \omega_0$. Hence $N \subset \omega_0$ as required by Theorem 3. By Theorem 3, the symmetric spacings hypothesis is the union of subspaces of dimension $J - [(J-1)/2]$ and λ^* can be used to test H_0^1 .

This examples gives a good illustration of the fact that the m subspaces defined in Theorem 3 are not all distinct. Let $J = 4$. In this example, $m = 4! = 24$ but actually H_0^1 consists of only 3 distinct subspaces. The H defined in the previous paragraph can be written as $H = (1, -1, -1, 1)$. These eight permutations of $N = \{\mu: H\mu = 0\}$ all equal N itself: $(1, 2, 3, 4)$, $(1, 3, 2, 4)$, $(4, 2, 3, 1)$, $(4, 3, 2, 1)$, $(2, 1, 4, 3)$, $(2, 4, 1, 3)$, $(3, 1, 4, 2)$, $(3, 4, 1, 2)$. These eight permutations of N all yield the subspace defined by $(1, -1, 1, -1)\mu = 0$: $(1, 2, 4, 3)$, $(1, 3, 4, 2)$, $(4, 2, 1, 3)$, $(4, 3, 1, 2)$, $(2, 1, 3, 4)$, $(2, 4, 3, 1)$, $(3, 1, 2, 4)$,

(3, 4, 2, 1). These eight permutations of N all yield the subspace defined by $(1, 1, -1, -1)\underline{\mu} = 0$: (1, 4, 2, 3), (1, 4, 3, 2), (4, 1, 2, 3), (4, 1, 3, 2), (2, 3, 1, 4), (2, 3, 4, 1), (3, 2, 1, 4), (3, 2, 4, 1).

Example 2. (Symmetric Spacings with Specified Ratios). Again $I = 1$ so the notation of Example 1 is used. Now the hypothesis of interest is a subhypothesis of H_0^1 , namely

$$(3.2) \quad H_0^2: \mu_{(j+1)} - \mu_{(j)} = c_j(\mu_{(2)} - \mu_{(1)}), \quad j = 2, \dots, J-1$$

where $c_2, \dots, c_{J-1}$ are specified positive constants with $c_j = c_{J-j}$, $j = 2, \dots, J-2$, and $c_{J-1} = 1$. The restrictions on the c_j imply that the symmetric spacings are equal as in H_0^1 . H_0^2 can be used to test whether the means are spaced like the expected values of order statistics from some symmetric distribution. For example, if $c_j = 1$, $j = 2, \dots, J-1$, the distribution is the uniform. If $J = 5$, $c_2 = c_3 = .74111$ and $c_4 = 1$, the distribution is the normal. Theorem 3 can be used to verify that H_0^2 , which is subpermutation invariant, is the union of subspaces. Let H be a $(J-2) \times J$ matrix such that $H\underline{\mu} = \underline{0}$ is equivalent to the conditions $\mu_{j+1} - \mu_j = c_j(\mu_2 - \mu_1)$, $j = 2, \dots, J-1$. Arguing as in Example 1 it is easy to verify that $N \cap O = \omega_0 \cap O$. To verify that $N \subset \omega_0$, let $\underline{\mu} \in N$, that is, $H\underline{\mu} = \underline{0}$. If $\mu_1 \leq \mu_2$, then $\mu_1 \leq \mu_2 \leq \dots \leq \mu_J$ since $c_j > 0$, $j = 1, \dots, J-1$. Thus $\underline{\mu} \in N \cap O = \omega_0 \cap O \subset \omega_0$. If $\mu_1 \geq \mu_2$, then $\mu_1 \geq \mu_2 \geq \dots \geq \mu_J$. For any $j = 2, \dots, J-1$, $\mu_{(j+1)} - \mu_{(j)} = \mu_{J-j} - \mu_{J+1-j} = -c_{J-j}(\mu_2 - \mu_1) = c_{J-j}(\mu_{J-1} - \mu_J) = c_j(\mu_{J-1} - \mu_J) = c_j(\mu_{(2)} - \mu_{(1)})$. The second, third and fourth equalities are true since $H\underline{\mu} = \underline{0}$, $c_{J-1} = 1$, and $c_j = c_{J-j}$. Thus $\underline{\mu} \in \omega_0$. By Theorem 3, ω_0 is the union of 2-dimensional subspaces and λ^* can be used to test H_0^2 .

Example 3. (Equal Spacings in Two Sets of Means). In this example there are two sets of means of interest, $I = 2$ and $J_1 = J_2 = J' = J/2$. We are interested in testing the hypothesis that the spacings in the first set, $\mu_{11}, \dots, \mu_{1J'}$, are equal to the spacings in the second set, $\mu_{21}, \dots, \mu_{2J'}$. That is, we wish to test

$$(3.3) \quad H_0^3: \mu_{1(j+1)} - \mu_{1(j)} = \mu_{2(j+1)} - \mu_{2(j)}, \quad j = 1, \dots, J'-1.$$

To verify that H_0^3 , which is subpermutation invariant, is the union of subspaces, let H be a $(J'-1) \times J$ matrix such that $H\mu = 0$ is equivalent to the conditions $\mu_{1,j+1} - \mu_{1j} = \mu_{2,j+1} - \mu_{2j}$, $j = 1, \dots, J'-1$. As in Example 1, it is easily verified that $N \cap O = \omega_0 \cap O$. To verify that $N \subset \omega_0$, let $\mu \in N$. For any $j = 1, \dots, J'$, $\mu_{2j} = \sum_{r=1}^{j-1} (\mu_{2,r+1} - \mu_{2r}) + \mu_{21} = \sum_{r=1}^{j-1} (\mu_{1,r+1} - \mu_{1r}) + \mu_{21} = \mu_{1j} + (\mu_{21} - \mu_{11})$. Thus the set of means, $\mu_{21}, \dots, \mu_{2J'}$, is a translation of the set of means, $\mu_{11}, \dots, \mu_{1J'}$, the amount of translation being $(\mu_{21} - \mu_{11})$. Thus the spacings among $\mu_{21}, \dots, \mu_{2J'}$, are all equal to the spacings among $\mu_{11}, \dots, \mu_{1J'}$ and $\mu \in \omega_0$. By Theorem 3, ω_0 is the union of subspaces of dimension $J - J' + 1 = J/2 + 1$ and λ^* can be used to test H_0^3 .

This argument can be easily extended to the situation in which one wishes to test for equal spacings in I ($I > 2$) sets of means. In this case H is an $(I-1)(J'-1) \times J$ matrix and the subspaces are of dimension $J - (I-1)(J'-1) = I + J' - 1 = I + J/I - 1$.

Example 4. (Equal Spacings in Two Sets of Means with Specified Ratios). For this example, the notation is the same as in Example 3. We combine the ideas in Examples 2 and 3 to consider testing

$$(3.4) \quad H_0^4: \begin{aligned} \mu_{1(j+1)} - \mu_{1(j)} &= \mu_{2(j+1)} - \mu_{2(j)}, & j &= 1, \dots, J'-1 \\ \mu_{1(j+1)} - \mu_{1(j)} &= c_j(\mu_{1(2)} - \mu_{1(1)}), & j &= 2, \dots, J'-1 \end{aligned}$$

where the c_j satisfy the same conditions as in Example 2. This null hypothesis is also subpermutation invariant. If we let ω_r denote the subspaces in Example 2 (now considered as subspaces of R^J with $\mu_{21}, \dots, \mu_{2J}$ unrestricted) and let η_s denote the subspaces in Example 3, we see that for H_0^4 , $\omega_0 = (\cup_r \omega_r) \cap (\cup_s \eta_s) = \cup_{r,s} (\omega_r \cap \eta_s)$. But for every r and s , $\omega_r \cap \eta_s$ is a linear subspace of dimension 3. (The subspace ω_r involves $J-2$ restrictions and η_s involves an additional $J-1$ restrictions. This results in $\omega_r \cap \eta_s$ having dimension 3.) Thus λ^* can be used to test H_0^4 .

4. Computation of the Test Statistic.

As in standard linear models theory, the LRT statistic λ^* can be expressed as a product of matrices. This will simplify the computation of λ^* . Furthermore, for the hypotheses described in Section 3, λ^* can be computed without any minimization. These points will be discussed in this section.

Let the J dimensional subspace ω be defined as $\omega = \{\xi: \xi = W\beta, \beta \in R^J\}$ where W is a known $K \times J$ design matrix. Let the subspaces ω_i be defined by $\omega_i = \{\xi: \xi = W\beta, H_i\beta = 0\}$ where H_i is a known $(J-q_i) \times J$ matrix of rank $J-q_i$, $i = 1, \dots, m$. Let $\hat{\beta} = (W'W)^{-1}W'X$. Then analogous to standard linear models theory we can write

$$(4.1) \quad \lambda^*(X) = \frac{SSH_0}{SSR}$$

where

$$(4.2) \quad SSH_0 = \min_{1 \leq i \leq m} \hat{\beta}' H_i' (H_i (W'W)^{-1} H_i')^{-1} H_i \hat{\beta}$$

and

$$(4.3) \quad SSR = (\underline{X} - W\hat{\beta})'(\underline{X} - W\hat{\beta}).$$

Expression (4.2) is true since the numerator of λ^* in (2.5) is the minimum of the sums of squares associated with each of the hypotheses $H_{0i}: \xi \in \omega_i$. By standard linear models theory these sums of squares are the expressions given in (4.2).

For the examples considered in Section 3, the minimum in (4.2) does not need to be computed. It is possible to determine which sum of squares will be the minimum by examining the order of the sample means, as is shown below. Thus only one sum of squares needs to be computed for these models.

For the remainder of this section, assume the model defined in Section 3. For this model $\hat{\beta} = \hat{\underline{\beta}} = (\hat{\underline{\beta}}_1', \dots, \hat{\underline{\beta}}_I')'$ where $\hat{\underline{\beta}}_i = (\hat{\beta}_{i1}, \dots, \hat{\beta}_{iJ_i})'$ and $\hat{\beta}_{ij} = \sum_{k=1}^{K_{ij}} x_{ijk}/K_{ij}$. Let H , the matrix from Theorem 3, be partitioned as $H = (C_1 : \dots : C_I)$ where C_i is a $(J-q) \times J_i$ matrix with columns $\xi_{i1}, \dots, \xi_{iJ_i}$, $i = 1, \dots, I$. For $i = 1, \dots, I$, suppose $\pi_i(1), \dots, \pi_i(J_i)$ is a permutation of $1, \dots, J_i$ such that $\bar{x}_{i\pi_i(1)} \leq \dots \leq \bar{x}_{i\pi_i(J_i)}$. Define a $(J-q) \times J$ matrix H^* by

$$(4.4) \quad H^* = (C_1^* : \dots : C_I^*)$$

where C_i^* is a $(J-q) \times J_i$ matrix with columns $\xi_{i1}^*, \dots, \xi_{iJ_i}^*$ and $\xi_{ij}^* = \xi_{i\pi_i^{-1}(j)}$. Then the numerator of λ^* , SSH_0 , is given in Theorem 4.

Theorem 4. If H , N , O , and ω_0 satisfy the conditions of Theorem 3 and H^* is defined by (4.4) then

$$(4.5) \quad SSH_0 = \hat{\underline{\beta}}' H^{*-'} (H^* (W'W)^{-1} H^{*-'})^{-1} H^* \hat{\underline{\beta}}.$$

Thus the matrix H^* can be constructed by permuting the columns of H in a way dictated by the order of the sample means. Then SSH_0 can be computed directly because H^* is the particular H_i which minimizes the sum of squares in (4.2). The proof of Theorem 4 uses the following two lemmas.

Lemma 1. Suppose H , N , O and ω_0 satisfy the conditions of Theorem 3 and H^* is defined by (4.4). Let $N^* = \{\underline{\mu} \in R^J: H^*\underline{\mu} = \underline{0}\}$. Let $O^* = \{\underline{\mu} \in R^J: \mu_{i\pi_i(1)} \leq \dots \leq \mu_{i\pi_i(J_i)}, i = 1, \dots, I\}$. Then $N^* \subset \omega_0$ and $N^* \cap O^* = \omega_0 \cap O^*$.

Proof. For any $\underline{\mu} \in R^J$, let $\pi(\underline{\mu}) = (\pi_1(\underline{\mu}_1), \dots, \pi_I(\underline{\mu}_I))'$ where $\pi_i(\underline{\mu}_i) = (\mu_{i\pi(1)}, \dots, \mu_{i\pi(J_i)})'$. Let $\underline{c}_{ij} = (c_{ij1}, \dots, c_{ij(J-q)})'$. Then for any $\underline{\mu} \in R^J$ and $1 \leq r \leq J-q$, the r^{th} coordinate of $H\pi(\underline{\mu})$ is $\sum_{i=1}^I \sum_{j=1}^{J_i} c_{ijr} \mu_{i\pi(j)}$ and the r^{th} coordinate of $H^*\underline{\mu}$ is $\sum_{i=1}^I \sum_{j=1}^{J_i} c_{ijr}^* \mu_{ij} = \sum_{i=1}^I \sum_{j=1}^{J_i} c_{i\pi^{-1}(j)r} \mu_{ij} = \sum_{i=1}^I \sum_{j=1}^{J_i} c_{ijr} \mu_{i\pi(j)}$. Thus $H\pi(\underline{\mu}) = H^*\underline{\mu}$ for every $\underline{\mu} \in R^J$.

If $\underline{\mu} \in N^*$, $\underline{0} = H^*\underline{\mu} = H\pi(\underline{\mu})$. Thus $\pi(\underline{\mu}) \in N \subset \omega_0$. Since ω_0 is sub-permutation invariant, $\underline{\mu} \in \omega_0$. Thus $N^* \subset \omega_0$, and hence $N^* \cap O^* \subset \omega_0 \cap O^*$.

If $\underline{\mu} \in \omega_0 \cap O^*$, then $\pi(\underline{\mu}) \in O$. Furthermore $\pi(\underline{\mu}) \in \omega_0$ since ω_0 is sub-permutation invariant. Thus $\pi(\underline{\mu}) \in N$. So $\underline{0} = H\pi(\underline{\mu}) = H^*\underline{\mu}$ and, hence, $\underline{\mu} \in N^*$. Thus $\underline{\mu} \in N^* \cap O^*$, and hence $\omega_0 \cap O^* \subset N^* \cap O^*$. ||

Lemma 2. Suppose $\bar{x}_{i\pi_i(1)} \leq \dots \leq \bar{x}_{i\pi_i(J_i)}$ for $i = 1, \dots, I$. Then for any $\underline{\mu} \in R^J$,

$$(4.6) \quad \sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij} (\bar{x}_{ij} - \mu_{ij})^2 \geq \sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij} (\bar{x}_{ij} - \mu_{ij}^*)^2$$

where μ^* is a subpermutation of $\underline{\mu}$ satisfying $\mu_{i\pi_i(1)}^* \leq \dots \leq \mu_{i\pi_i(J_i)}^*$ for $i = 1, \dots, I$.

Proof. Let $g(\bar{X}_i, \underline{\mu}_i) = - \sum_{j=1}^{J_i} K_{ij} (\bar{X}_{ij} - \mu_{ij})^2 = G(\bar{X}_i - \underline{\mu}_i)$. It is easily shown (Marshall and Olkin, 1979, p. 57) that G is Schur-concave. It follows by Lemma 2.2 of Hollander, Proschan, and Sethuraman (1977) that g is decreasing in transposition and, hence,

$$- \sum_{j=1}^{J_i} K_{ij} (\bar{X}_{ij} - \mu_{ij}^*)^2 \geq - \sum_{j=1}^{J_i} K_{ij} (\bar{X}_{ij} - \mu_{ij})^2.$$

Inequality (4.6) follows since the above inequality holds for each $i = 1, \dots, I$. ||

Proof of Theorem 4. Let N^* and O^π be as in Lemma 1 and let

$$g(\bar{X}, \underline{\mu}) = \sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij} (\bar{X}_{ij} - \mu_{ij})^2. \quad \text{The numerator of } \lambda^*(\bar{X}) \text{ is}$$

$$\begin{aligned} \text{SSH}_0 &= \inf_{\underline{\mu} \in \omega_0} g(\bar{X}, \underline{\mu}) \\ &\leq \inf_{\underline{\mu} \in N^*} g(\bar{X}, \underline{\mu}) \quad (\text{since } N^* \subset \omega_0 \text{ by Lemma 1}) \\ (4.7) \quad &\leq \inf_{\underline{\mu} \in \omega_0 \cap O^\pi} g(\bar{X}, \underline{\mu}) \quad (\text{since } \omega_0 \cap O^\pi \subset N^* \text{ by Lemma 1}) \\ &= \inf_{\underline{\mu} \in \omega_0} g(\bar{X}, \underline{\mu}), \end{aligned}$$

the last equality being true by Lemma 2 since ω_0 is subpermutation invariant. The second and last expression in (4.7) are equal so the inequalities are equalities. But by standard linear models theory, the expression involving N^* in (4.7) is the right hand side of (4.5). ||

We will now compute a small example using Theorem 4. Suppose we wish to test the symmetric spacings hypothesis of Example 1 with $J = 3$ populations. So $H = (1, -2, 1)$ can be used. Suppose $n = 4$ with two observations on population 2, so

$$W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Suppose the observed mean vector is $\bar{\underline{x}} = (1, 10, 4)'$. Then $H^* = (1, 1, -2)$. Equation (4.5) yields $SSH_0 = 18/11$. The maximum likelihood estimate of $\underline{\mu}$ obtained as the projection onto N^* is $\hat{\underline{\mu}} = (5/11, 107/11, 56/11)$. The sums of squares corresponding to the projections onto the other two subspaces in ω_0 , those defined by $H\underline{\mu} = 0$ and $(-2, 1, 1)\underline{\mu} = 0$, are $225/4$ and $288/11$. Clearly SSH_0 is the minimum of the three sums of squares.

REFERENCES

- Berger, R. L. (1982), "Multiparameter Hypothesis Testing and Acceptance Sampling," Technometrics, 24, in print.
- Bickel, P. J. and Doksum, K. A. (1977), Mathematical Statistics: Basic Ideas and Selected Topics, San Francisco: Holden Day, Inc.
- Gleser, L. J. (1973), "On a Theory of Intersection-Union Tests," Institute of Mathematical Statistics Bulletin, 2, 233. (Abstract)
- Grant, P. R. (1972), "Convergent and Divergent Character Displacement," Biological Journal of the Linnean Society, 4, 39-68.
- Grant, P. R. and Abbott, I. (1980), "Interspecific Competition, Island Biography and Null Hypotheses," Evolution, 34, 332-341.
- Hendrickson, J. A., Jr. (1981), Community-wide Character Displacement Reexamined, Evolution, 35, 794-809.
- Hollander, M., Proschan, F. and Sethuraman, J. (1977), "Functions Decreasing in Transposition and Their Applications in Ranking Problems," Annals of Statistics, 5, 722-733.
- Marshall, A. W. and Olkin, I. (1979), Inequalities: Theory of Majorization and Its Applications, New York: Academic Press.
- Sinclair, D. F., Mosimann, J. E. and Meeter, D. A. (1982), Tests for Character Displacement, Department of Statistics, The Florida State University, Tallahassee, Florida, Technical Report No. M-629.
- Strong, D. R. and Simberloff, D. S. (1981), "Straining at Gnats and Swallowing Ratios: Character Displacement," Evolution, 35, 810-812.
- Strong, D. R., Jr., Szyska, L. A., and Simberloff, D. S. (1979), "Tests of Community-wide Character Displacement Against Null Hypotheses," Evolution, 33, 897-913.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER FSU Statistics M639 USARO D-55	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and subtitle) Testing Hypotheses Which Are Unions of Linear Subspaces	5. TYPE OF REPORT & PERIOD COVERED Technical	
7. AUTHOR(s) Roger L. Berger and Dennis F. Sinclair	6. PERFORMING ORG. REPORT NUMBER FSU Statistics M639	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Florida State University Tallahassee, FL 32306	8. CONTRACT OR GRANT NUMBER(s) USARO Contract DAAG-29-79-C-0158	
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	12. REPORT DATE September, 1982	
	13. NUMBER OF PAGES 19	
	15. SECURITY CLASS. (of this report) Unclassified	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	

16. DISTRIBUTION STATEMENT (of this report)

Distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report)

18. SUPPLEMENTARY NOTES

THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT
ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSIDERED AS
AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DE-
CISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

19. KEY WORDS

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The likelihood ratio test (LRT) for hypotheses which are unions of linear subspaces is derived for the normal theory linear model. A more powerful variant of the LRT is proposed for the case in which the subspaces are not all of the same dimension. A theorem is proved which may be used to identify hypotheses which are unions of linear subspaces. Some hypotheses, of particular relevance in ecology, concerning the spacings between normal means are shown to be unions of linear subspaces and are therefore testable using the LRT. Finally, the computation of the LRT statistic is discussed.