INFERENCE CONCERNING THE MEAN VECTOR WHEN THE COVARIANCE MATRIX IS TOTALLY REDUCIBLE

BY

DENNIS L. YOUNG

TECHNICAL REPORT NO. 97
APRIL 1, 1975

PREPARED UNDER THE AUSPICES OF
NATIONAL SCIENCE FOUNDATION GRANT GP-32326X

Ingram Olkin, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
INFERENCES CONCERNING THE MEAN VECTOR WHEN THE
COVARIANCE MATRIX IS TOTALLY REDUCIBLE

BY

DENNIS L. YOUNG

TECHNICAL REPORT NO. 97
APRIL 1, 1975

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT GP-32326X

Ingram Olkin, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
INFERENCES CONCERNING THE MEAN VECTOR WHEN THE
COVARIANCE MATRIX IS TOTALLY REDUCIBLE

Dennis L. Young
New Mexico State University and
Stanford University

ABSTRACT

Given a random sample from a p-variate normal distribution \( N_p(\mu, \Sigma) \),
the likelihood ratio criterion and an information theory criterion, which
is an analogue of Hotelling's \( T^2 \), are given for testing \( H_0: \mu = \mu_0 \)
(known) under the condition that \( \Sigma \) is totally reducible, that is, diag-
onalizable by an orthogonal matrix which depends upon the pattern but not
on the unknown elements of \( \Sigma \). Exact, approximate and asymptotic distri-
butions of the test criteria are considered. Various confidence regions
concerning \( \mu \) are provided. Application to repeated measures experiments
is discussed.

1. INTRODUCTION

An important model for the parameters of a multivariate normal dis-
tribution is the one which specifies the pattern or structure of the co-
variance matrix \( \Sigma \). Among the covariance patterns which have been
suggested and studied are: intraclass (Wilks (1946)), compound symmetric
(Votaw (1948)), circular symmetric (Olkin and Press (1969)), circular
symmetric in blocks (Olkin (1972)), intraclass of order \( k \) (Rogers and
Young (1974)), intraclass in blocks (Arnold (1973) and Fleiss (1966)),
simplex (Mukherjee (1966) and Jöreskog (1970)), linear (Anderson (1970,
1973)) and Kronecker (Rogers and Young (1975)). A number of these
patterns (intraclass, circular symmetric, intraclass of order $k$ for example) are special cases of covariance patterns which are totally reducible, that is, there is an orthogonal matrix which diagonalizes the covariance matrix independent of its elements. Totally reducible covariance patterns have been studied by Srivastava (1966), Mukherjee (1970) and Rogers and Young (1975).

Tests of hypotheses concerning the mean vector when the covariance matrix is patterned are also of considerable importance and find use in certain analysis of variance problems. Inference concerning the mean vector when the covariance pattern is intraclass are found in Wilks (1946), Geisser (1963), Olkin and Skrikhande (1970) and Marrison (1972). When the covariance pattern is compound symmetric, Votaw (1948), Geisser (1963) and Morrison (1972) discuss problems concerning the mean vector. Fleiss (1966), Olkin and Press (1969), Olkin (1972) and Arnold (1973) also develop tests for the mean vector under various covariance patterns.

The problem to be considered here is that of making inferences about the mean vector when the covariance matrix has a pattern which is totally reducible. The results here generalize a number of studies mentioned earlier for specific reducible patterns. The likelihood ratio test and a test using an information theory approach, which is based on an analogue of Hotelling's $T^2$, are developed for the hypothesis $H_0: \mu = \mu_0$ (known). Exact, approximate and asymptotic distributions of the test criteria are discussed as well as construction of simultaneous confidence regions for linear functions on the components of the mean vector. Attention is also given to the problem of testing equality of the components of the mean vector. These results have application in repeated
measures experiments. The extension to the multipopulation case will be undertaken in a future study.

2. A CANONICAL FORM

First the following definitions and results about patterned covariance matrices are stated (see Rogers and Young (1975)).

Definition 1.

Let \( \Sigma = \Sigma(\sigma) = \sum_{j=1}^{m} \sigma_j G_j \), where \( G_1, \ldots, G_m \) are known, real, symmetric, linearly independent \( p \times p \) matrices, \( \sigma_1, \ldots, \sigma_m \) are real numbers such that \( \Sigma \succeq 0 \) (positive semi-definite) and \( m \leq p(p + 1)/2 \). Then \( \Sigma \) is a patterned covariance matrix.

Definition 2.

A \( p \times p \) patterned covariance matrix \( \Sigma = \Sigma(\sigma) \) is totally reducible if and only if there is an orthogonal matrix \( P \) such that \( P \Sigma P^t \) is diagonal for all \( \sigma \in \Omega = \{ \sigma | \Sigma(\sigma) \succeq 0 \} \). (Note that \( P \) is independent of \( \sigma \in \Omega \).)

Rogers and Young (1975) give numerous necessary and sufficient conditions for a patterned covariance matrix to be totally reducible. One of these conditions states that \( \Sigma(\sigma) \) is totally reducible if and only if the characteristic roots of \( \Sigma(\sigma) \) are linear combinations of \( \sigma_1, \ldots, \sigma_m \). If \( q \) is the number of distinct characteristic roots of \( \Sigma(\sigma) \) and if \( \Sigma \) is patterned and totally reducible, then \( m \leq q \leq p \).
In the case that $m = q$ estimation of the characteristic roots of $\Sigma$ is equivalent to estimating the parameters $\sigma_1, \ldots, \sigma_m$. If $m < q$ estimates of the $\sigma$'s are not unique. The likelihood ratio test of the hypothesis that $\Sigma$ is totally reducible is given by Rogers and Young (1975).

Let $X_1, \ldots, X_n$ be a random sample of size $n$ from $N_p(\mu, \Sigma_X)$, the $p$-variate normal distribution. Let $\Sigma_X = \sum_{j=1}^m \sigma_j G_j$ be totally reducible by the orthogonal matrix $P$. Make the transformation $Y_i = PX_i$, $i = 1, \ldots, n$. If $L(X)$ denotes the probability law of $X$, then $L(Y_1) = N_p(\nu, \Sigma_Y)$ where $\nu = P\mu$, and $\Sigma_Y = P\Sigma_X P'$. Now $\Sigma_Y$ is diagonal and without loss of generality

$\Sigma_Y = \text{diag}(\tau_1, \ldots, \tau_1, \tau_2, \ldots, \tau_2, \ldots, \tau_q, \ldots, \tau_q)$ where $\tau_1, \ldots, \tau_q$ are the $q$ distinct characteristic roots of $\Sigma_X$, $\tau_i$ appears with multiplicity $p_i (\sum_{i=1}^q p_i = p)$ and $\tau_i = \sum_{j=1}^m \sigma_j d_{ij}$ with $d_{ij}$ the $i$-th diagonal element of $P G_j P'$, $i = 1, \ldots, p$, $j = 1, \ldots, m$. Note that the components of each $Y_i$ are independent.

Thus as our starting point we will assume the canonical form for our random sample, namely that $Y_1, \ldots, Y_n$ are a random sample from $N_p(\nu, \Sigma_Y)$ where $\Sigma_Y = \text{diag}(\tau_1, \ldots, \tau_1, \ldots, \tau_q, \ldots, \tau_q)$ with $\tau_i$ having multiplicity $p_i$, $i = 1, \ldots, q$. Then hypotheses about $\mu$ can be stated in terms of hypotheses about $\nu$ since $\nu = P\mu$.

3. TESTS OF $H_0: \mu = \mu_0$

In this section we consider testing $H_0: \mu = \mu_0$ (known) against $H_A: \mu \neq \mu_0$, or equivalently $H_0: \nu = \nu_0$ against $H_A: \nu \neq \nu_0$ where $\nu = P\mu$ and $\nu_0 = P\mu_0$. Two testing procedures are considered. One test
is the likelihood ratio test and the other is a test based on an information theory approach utilizing an analogue of Hotelling's $\mathbf{T}^2$ statistic. The canonical form for the random sample is assumed throughout.

3.1 Likelihood Ratio Test of $H_0: \nu = \nu_0$

Let $\omega = \{\nu, \tau \mid -\infty < \nu_i < \infty, i = 1, \ldots, p; \tau_i > 0, i = 1, \ldots, q\}$ and $\omega_0 = \{\nu, \tau \mid \nu = \nu_0; \tau_i > 0, i = 1, \ldots, q\}$ where $\tau = (\tau_1, \ldots, \tau_q)'$, $\nu = (\nu_1, \ldots, \nu_p)'$ and $\nu_0 = (\nu_{10}, \ldots, \nu_{p0})'$. Then the likelihood ratio criterion for testing $H_0: \nu = \nu_0$ against $H_a: \nu \neq \nu_0$ is

$$
\lambda = \frac{\sup_{\omega_0} L(\nu, \tau; Y_1^1, \ldots, Y_n^1)}{\sup_\omega L(\nu, \tau; Y_1, \ldots, Y_n^1)}
$$

where

$$L(\nu, \tau; Y_1, \ldots, Y_n^1) = (2\pi)^{-np/2} \left( \prod_{i=1}^q p_i \right)^{-n/2} \exp\left(-\frac{1}{2} \sum_{i=1}^q \sum_{j=1}^q (Y_{ij} - \nu_{ij})^2 / \tau_i \right),$$

$Y_\alpha = (Y_{\alpha 1}, \ldots, Y_{\alpha p})$ and $\sum_i$ denotes summation over the set of indices $I_i = \{ \sum_{j=1}^p p_j + 1, \ldots, \sum_{j=1}^q p_i \}$.

Now

$$\sup_{\omega_0} L(\nu, \tau; Y_1, \ldots, Y_n^1) = (2\pi)^{-np/2} \left( \prod_{i=1}^q \hat{\tau}_{i0} \right)^{-n/2} e^{-np/2},$$

where

$$\hat{\tau}_{i0} = \frac{n}{\sum_{\alpha=1}^q (Y_{\alpha j} - \nu_{j0})^2 / np_i},$$

and $p_i = \prod_{j=1}^q p_{ij}$.
and
\[
\sup_{\omega} L(\nu, \tau; \nu_1, \ldots, \nu_n) = (2\pi)^{-np/2} \left( \prod_{i=1}^{m} \tau_i \right)^{-n/2} e^{-np/2},
\]
where
\[
\hat{\tau}_i = \sum_{\alpha=1}^{n} \sum_i (y_{\alpha j} - \bar{y}_j)^2 / np_i.
\]

Note that \(\hat{\tau}_1, \ldots, \hat{\tau}_q\) are independent. Then
\[
\lambda^{2/n} = \prod_{i=1}^{q} \left( \frac{\hat{\tau}_i}{\hat{\tau}_{i0}} \right)^{p_i}
\]
and \(H_0\) is rejected if \(\lambda < c\).

To find the distribution of \(\lambda\), first note that \(\hat{\tau}_{i0} = \hat{\tau}_i + v_i\), where \(v_i = \sum_i (\bar{y}_j - \bar{y}_{j0})^2 / p_i\), \(\hat{\tau}_i\) and \(v_i\) are independent,
\[
L(np_i \hat{\tau}_i / \tau_i) = \chi^2(p_i(n - 1)) \quad \text{and} \quad L(np_i v_i / \tau_i) = \chi^2(p_i) \quad \text{if} \quad H_0 \quad \text{is true.}
\]

Thus
\[
\lambda^{2/n} = \prod_{i=1}^{q} \left( \frac{\hat{\tau}_i}{(\hat{\tau}_i + v_i)} \right)^{p_i}.
\]

Also \(L(\hat{\tau}_i / (\hat{\tau}_i + v_i)) = Be({\alpha_+ p_i(n - 1), \alpha_+ p_i})\), the beta distribution with parameters \(\alpha p_i(n - 1)\) and \(\alpha_+ p_i\). So \(L(\lambda) = L((\prod_{i=1}^{q} p_i n)^{n/2})\) where
\[
L(X_i) = Be({\alpha_+ p_i(n - 1), \alpha_+ p_i}), \quad i = 1, \ldots, q \quad \text{and the} \quad X_i's \quad \text{are independent.}
\]

The moments of \(\lambda\) and \(\lambda^{2/n}\) can be easily computed using the fact that \(\lambda\) has the same distribution as products of powers of independent beta variables.

\[
E(\lambda^h) = \prod_{i=1}^{q} \frac{\Gamma(\alpha_p n)}{\Gamma(\alpha_p (n - 1))} \prod_{i=1}^{q} \frac{\Gamma(\alpha_+ p_i n(1 + h) - \alpha_+ p_i)}{\Gamma(\alpha_+ p_i n(1 + h))} \quad (3.1)
\]
and

\[ E(\frac{\lambda^{2/n}}{h}) = \frac{q}{]\Pi_{i=1}^{q} \frac{\Gamma(\frac{1}{2} p_i n)}{\Gamma(\frac{1}{2} p_i (n-1))}} \frac{q}{\Pi_{i=1}^{q} \frac{\Gamma(\frac{1}{2} p_i (n-1) + h p_i)}{\Gamma(\frac{1}{2} p_i n + h p_i)}}. \] (3.2)

Using the Gauss-Legendre multiplication formula for the gamma function (3.2) can be rewritten as

\[ E(\frac{\lambda^{2/n}}{h}) = q \prod_{i=1}^{q} \frac{p_i^{-1}}{\Pi_{r=0}^{(n-1)/2 + r/p_i}} \frac{\Gamma(n/2 + r/p_i)}{\Gamma((n-1)/2 + r/p_i)} \frac{\Gamma((n-1)/2 + r/p_i + h)}{\Gamma(n/2 + r/p_i + h)}. \] (3.3)

Because \( 0 \leq \lambda \leq 1 \), the moments of \( \chi^2/n \) uniquely determine its distribution. Thus (3.3) shows that \( L(\chi^2/n) = L(\prod_{i=1}^{q} \prod_{j=1}^{p_i} X_{ij}) \) where \( L(X_{ij}) = \text{Be}((n-1)/2 + (j-1)/p_i, \frac{1}{2}) \), \( j=1, \ldots, p_i \), \( i=1, \ldots, q \). Since (3.3) gives the Mellin transform of \( \chi^2/n \), the results of Mathai and Saxena (1973) give the exact distribution of \( \chi^2/n \) in terms of Meijer's G function. Expansion of Meijer's G function representation of the distribution of \( \chi^2/n \) can be undertaken using the techniques of Mathai and Saxena (1973) for specific values of \( p_1, \ldots, p_q \).

A very useful approximation of the distribution of \( \lambda \) when \( H_0 \) is true can be obtained by the method developed by Box. See Gleser and Olkin (1972) for some simplified expressions for determining parameters in the Box approximation or see Anderson (1958, p203). Suppose that \( W = (n_{i=1}^{q} X_i^p)^r \) where \( L(X_i) = \text{Be}(sp_i, tp_i) \), \( r, s, t > 0 \), and the \( X \)'s are independent. Then

\[ E(W^h) = q \prod_{i=1}^{q} \frac{\Gamma(p_i (s+t))}{\Gamma(p_i s)} \frac{q}{\prod_{i=1}^{q} \frac{\Gamma(p_i r(1+h) + p_i (s-r))}{\Gamma(p_i r(1+h) + p_i (s+t-r))}} \]

and the asymptotic expansion of the distribution is given by
\[ P(-2 \rho \ln \lambda \leq x) = (1-\phi)P(X^2(f) \leq x) + \phi P(X^2(f+4) \leq x) + R(\gamma) \]

where \( f = 2pt, \ \rho = (2ps+pt-q)/(2pr), \ \gamma = pr \) and

\[
\phi = \frac{pt(t^2p^2 - 3q^2 + 2p\sum_{i=1}^{q} 1/p_i)}{6(2ps + pt - q)^2}.
\]

If \( s = O(\gamma) \), it can be shown that \( R(\gamma) = O(\gamma^{-3}) \) as \( \gamma \to \infty \). In the application here \( r = n/2 \) and so the remainder term is \( O(n^{-3}) \).

Using the moments of \( \lambda \) in (3.1) we see that \( r = n/2, s = (n-1)/2, t = 1/2 \). So the parameters in the Box approximation are \( f = p, \rho = (2pn - p - 2q)/(2pn) \) and

\[
\phi = \frac{p(p^2 - 12q^2 + 8p\sum_{i=1}^{q} 1/p_i)}{12(2pn - 2q - p)^2}.
\]

To obtain the asymptotic non-null distribution note that if \( \nu \neq \nu_0 \)
\[ L(np_i \nu_i / \tau_i) = X^2(p_i; n\delta_i) \] (the non-central \( X^2 \) distribution with \( p_i \) degrees of freedom) where \( \delta_i = \sum_i (\nu_j - \nu_{j0})^2 / \tau_i \), and that \( X^2/n \) is the product of \( q \) such independent random variables. Using the fact that if \( L(X) = X^2(\eta; \delta), L((X-(\eta+\delta))/[2(\eta+\delta)^{1/2}]) \to N(0, 1) \) as \( \eta \) or \( \delta \to \infty \) (Johnson and Kotz (1970)), and the delta method of Cramer (see Rao (1965), p321), we find that

\[
L((n/2)^{1/2}[\ln \lambda^{2/n} - \ln \prod_{i=1}^{q} (1 + \delta_i/p_i)^{-p_i}])
\to N(0, p - \sum_{i=1}^{q} p_i (1 + \delta_i/p_i)^{-2}).
\]
3.2 Information Theory Statistic for Testing $H_0: \nu = \nu_0$.

A second criterion for testing $H_0: \nu = \nu_0$ is the analogue of Hotelling's $T^2$ which is the trace of the product of two matrices. This statistic is derived from the information principle of Kullback (1959, Chapter 9) and is given by

$$U = ntr S_\nu^{-1}(\overline{\nu} - \nu_0)(\overline{\nu} - \nu_0)' = n(\overline{\nu} - \nu_0)'S_\nu^{-1}(\overline{\nu} - \nu_0),$$

where $S_\nu = \text{diag}(s_1^2, s_2^2, \ldots, s_q^2)$, $s_i^2$ has multiplicity $p_i$, $i = 1, \ldots, q$, and $s_i^2 = \frac{\sum_{i=1}^{n} \sum_{j=1}^{p_i} (y_{ij} - \overline{y}_j)^2}{(p_i(n-1))} = n\bar{\tau}_i/(n-1)$. Then

$$U = n \sum_{i=1}^{q} \frac{\sum_{j=1}^{p_i} (y_{ij} - \overline{y}_j)^2}{s_i^2} = \sum_{i=1}^{q} p_i F_i$$

where

$$F_i = n(n-1) \frac{\sum_{j=1}^{p_i} (y_{ij} - \overline{y}_j)^2}{\sum_{j=1}^{p_i} (y_{ij} - \overline{y}_j)^2} = (n-1) \frac{\nu_i}{\bar{\tau}_i}$$

and $L(F_i) = F(p_i, p_i(n-1))$ (the $F$ distribution) if $H_0$ is true. Thus $U$ is a linear combination of independent central $F$ variables. The null hypothesis $H_0: \nu = \nu_0$ is rejected if $U > u_0$.

The exact distribution of $U$ appears to be very complicated (see Morrison (1971) for a discussion of the exact distribution for $q = 2$). As an approximation to the distribution of $U$ Morrison (1971), (1972) suggests using $cF(a, b)$, where $a$, $b$ and $c$ are determined by equating the first three cumulants of $U$ to the first three cumulants of $cF$ where $L(F) = F(a, b)$. Now the $j^{th}$ cumulant of $U$ is
\[ \kappa_j(u) = \sum_{i=1}^{q} p_i^j \kappa_j(i) \]

where \( \kappa_j(i) \) is the \( j \)th cumulant for the \( F(p_i, p_i(n-1)) \) distribution. The first three cumulants of \( \text{cF} \) are

\[ \kappa_1(\text{cF}) = cb/(b-2) \]

\[ \kappa_2(\text{cF}) = 2c^2b^2(b+a-2)/[a(b-2)^2(b-4)] \]

and

\[ \kappa_3(\text{cF}) = 8c^3b^3(b+a-2)(2a+b-2)/[a^2(b-2)^3(b-4)(b-6)] . \]

Computing the first three cumulants of \( U \) yields

\[ \kappa_1(U) = (n-1) \sum_{i=1}^{q} p_i^2/[p_i(n-1) - 2] \]

\[ \kappa_2(U) = 2(n-1)^2 \sum_{i=1}^{q} p_i^3(p_i(n-2))/[(p_i(n-1) - 2)^2(p_i(n-1) - 4)] \]

and

\[ \kappa_3(U) = 8(n-1)^3 \sum_{i=1}^{q} \frac{p_i^4(p_i(n-2)(p_i(n+1) - 2)}{(p_i(n-1) - 2)^3(p_i(n-1) - 4)(p_i(n-1) - 6)} . \]

Finally equating the first three cumulants of \( U \) and \( \text{cF} \) gives

\[ b = \frac{6\kappa_1(U)\kappa_3(U) + 4[\kappa_1(U)]^2\kappa_2(U) - 8[\kappa_2(U)]^2}{\kappa_1(U)\kappa_3(U) - 2[\kappa_2(U)]^2} \]

\[ a = \frac{4(b-2)\kappa_1(U)\kappa_2(U)}{(b-6)\kappa_3(U) - 8\kappa_1(U)\kappa_2(U)} \]

and
\[ c = \frac{b - 2}{b} \kappa_1(U). \]

The non-null distribution of \( U \) is that of \( \sum_{i=1}^q p_i F'_i \) where \( L(F'_i) = F'(p_i, p_i(n-1); n\delta_i) \), the noncentral \( F \) distribution, with \( \delta_i = \sum_i (\nu_j - \nu_{j0})^2/\tau_i \) and \( F'_1, \ldots, F'_q \) are independent. The non-null distribution of \( U \) can again be approximated by equating the first three cumulants of \( U \) to the first three cumulants of \( cF(a, b) \). The details are omitted.

The asymptotic distribution of \( U \) if \( H_0 \) is true is \( \chi^2(p) \) since \( L(F'_i) \rightarrow L(\chi^2(p_i)/p_i) \) and \( n \rightarrow \infty \). If \( \nu \neq \nu_0 \), the delta method of Cramer can be applied as in Section 3.1 to yield the following asymptotic distribution

\[ L((n/2)^{1/2}(u/n - \sum_{i=1}^q \delta_i)) \rightarrow N(0, \sum_{i=1}^q p_i (\delta_i/p_i + 1)^2 - p). \]

3.3 Confidence Regions

To obtain simultaneous confidence intervals for the components of \( \nu = (\nu_i, \ldots, \nu_p)' \), we note that if

\[ T_j = \frac{\bar{y}_j - \nu_j}{[\sum_{\alpha=1}^n \sum_i (y_{\alpha k} - \bar{y}_k)^2/(p_i(n-1))]^{1/2}}, \]

then \( L(T_j) = t(p_i(n-1)), j \in I_i, i = 1, \ldots, q \). The joint distribution of \( \{T_j, j \in I_i\} \) is multivariate \( t(p_i > 1) \) since the \( T_j \)'s, \( j \in I_i \), are not independent (Dunnett and Sobel (1955)). Conservative confidence intervals can be obtained using the inequality

\[ P\left( \bigcap_{j \in I_i} \{T_j \in I_j\} \right) \geq \prod_{j \in I_i} P(T_j \in I_j). \]
where \( I_j \) is the interval \([a_j, b_j]\). Let \( P(T_j \in I_j) = 1 - \alpha, j \in I_i, i = 1, \ldots, q \). Since the sets \( \{T_j \}_{j \in I_1}, \ldots, \{T_j \}_{j \in I_q} \) are independent we have
\[
P[\cap_{i=1}^{q} \cap_{j \in I_i} (T_j \in I_j)] \geq \prod_{i=1}^{q} P(T_j \in I_j) = (1 - \alpha)^P.
\]

Conservative simultaneous confidence intervals for \( \nu_1, \ldots, \nu_p \) with confidence coefficient at least \( 1 - \gamma \) are given by \( \bar{y}_j \pm t_{\alpha}(p_i(n-1))s_j/\sqrt{n} \), \( j \in I_i, i = 1, \ldots, q \) and \( \alpha = 1 - (1-\gamma)^{1/p} \). Confidence intervals for \( \nu_1, \ldots, \nu_p \) can then be used to construct confidence intervals for \( \mu_1, \ldots, \mu_p \) since \( \mu = P'v \).

Simultaneous confidence intervals for all linear functions of \( \nu \), say \( a'v \), can be found by employing Schiff's method to the ellipsoidal acceptance region obtained using the \( U \) test statistic. The \( 100(1 - \alpha)\% \) set of simultaneous confidence intervals for any linear compound \( a'v \) has general interval \( a'y \pm [u_\alpha a'S_y a/n]^{1/2} \), where \( u_\alpha \) is the upper \( \alpha \) percentage point of the \( U \) statistic. Since \( \mu = P'v \), simultaneous confidence intervals for all linear functions \( b'\mu \) can be found by letting \( a' = b'P \).

### 3.4 Comments

The two criteria for testing \( H_0: \nu = \nu_0 \) are functions of the statistics \( F_1, \ldots, F_q \) which are independent with \( L(F_i) = F(p_i, p_i(n-1)) \). Other functions of the \( F_i \) would lead to other test criteria with their various power functions. Different criteria could be useful against certain alternative hypotheses. One might also consider the test with the acceptance region the rectangular set \( \{F_i < c_i, i = 1, \ldots, q\} \) where
the $c_i$'s are selected so that the desired significance level is attained.

4. TESTS OF $H_0: \mu = \beta \epsilon$

Another hypothesis of importance is $H_0: \nu = \beta \epsilon$ where $\beta$ is an unknown scalar and $e = (e_1, \ldots, e_p)'$ is a known $p \times 1$ vector with $e_i \neq 0$, $i = 1, \ldots, p$. If $e = (1, \ldots, 1)'$ we have the so-called repeated measurements hypothesis $H_0: \mu_1 = \cdots = \mu_p$ (see Morrison (1972)). Now there exists a $(p - 1) \times p$ matrix $C$ so that $Ce = 0$ and $\text{rank}(C) = p - 1$. One such $C = (c_{ij})$ has elements $c_{ii} = e_{i+1}/e_i$, $c_{i, i+1} = -1$, $i = 1, \ldots, p - 1$ and all other entries zero. Then $H_0: \mu = \beta \epsilon$ is equivalent to $H_0: C \mu = 0$.

If $X_1, \ldots, X_n$ is a random sample from $N_p(\mu, \Sigma_X)$ where $\Sigma_X = L_{i=1}^m \sigma_i G_i$ is patterned, let $\omega_i = CX_i$, $i = 1, \ldots, n$. Then $L(\omega_i) = N_{p-1}(\psi, \Sigma_{\omega})$ with $\psi = C\mu$ and $\Sigma_{\omega} = C\Sigma_X C' = L_{i=1}^m \sigma_i H_i$ where $H_i = CG_i C'$. If $\Sigma_{\omega}$ is totally reducible by the orthogonal matrix $P$, let $V_i = P \omega_i$ where $L(V_i) = N_{p-1}(\nu, \Sigma_{\nu})$ and $\Sigma_{\nu}$ is diagonal. A necessary and sufficient condition for $\Sigma_{\omega}$ to be totally reducible is that $H_1, \ldots, H_m$ be commutative (i.e., $H_i H_j = H_j H_i$, $i, j = 1, \ldots, m$) (Rogers and Young (1975)). Note that if $\Sigma_{\omega}$ is totally reducible, the number of unknown parameters in $\Sigma_{\omega}$ must be less than or equal to $p - 1$, i.e., $m \leq q \leq p - 1$. Some covariance patterns which are reducible after the transformation by $C$ are the intraclass, circular symmetric and certain cases of the intraclass of order $k$.

If $\Sigma_{\omega}$ is totally reducible, the results of Section 3 can be used to test the hypothesis $H_0: \mu = \beta \epsilon$ which is equivalent to $H_0: \nu = PC \mu = 0$. Simultaneous confidence intervals for contrasts (or any linear compound) of the components of $\mu$ can be obtained using the treatment of Section 3.3.
BIBLIOGRAPHY


