AN EXTENSION OF WILKS' TEST FOR THE EQUALITY OF MEANS

BY

I. OLKIN and S. S. SHRIRKANDE

TECHNICAL REPORT NO. 33
FEBRUARY 29, 1969

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT 6681

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
AN EXTENSION OF WILKS' TEST FOR THE EQUALITY OF MEANS

BY

I. OLKIN and S. S. SHRIKHANDE

TECHNICAL REPORT NO. 33
FEBRUARY 29, 1969

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT 6681

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
AN EXTENSION OF WILKS' TEST FOR THE EQUALITY OF MEANS

I. Olkin and S. S. Shrikhande
Stanford University and University of Bombay

1. Introduction. When \( x = (x_1, \ldots, x_p) \) has a multivariate normal distribution with mean \( \theta \) and covariance matrix \( \Sigma \), then Hotelling's \( T^2 \) test may be used for testing \( \theta = 0 \) versus \( \theta \neq 0 \). However, when part of the mean vector is known, i.e., \( \hat{\theta} = (\theta_1, \ldots, \theta_k) = 0 \), and we wish to test that the remaining part is zero, i.e., \( \tilde{\theta} = (\theta_{k+1}, \ldots, \theta_p) = 0 \) against \( \tilde{\theta} \neq 0 \), then the \( T^2 \)-statistic should not be used. The likelihood ratio test is given by \( (1 + T^2_k)/(1 + T^2_p) \), where \( T^2_m \) is the usual \( T^2 \) statistic in which the first \( m \) variates is used. This test, its properties, and alternative tests have been studied by a number of authors (all the references may be obtained from Olkin and Shrikhande (1954) and Kabe (1965)).

When the variables are interchangeable with respect to variances and covariances—the intraclass correlation model—the test of the hypothesis \( \theta_1 = \cdots = \theta_p \) versus \( -\infty < \theta_j < \infty, j = 1, \ldots, p \), was obtained by Wilks (1946). If we know that \( \theta_1 = \cdots = \theta_k \), and wish to test for the equality of all the means, the test statistic should be a variant of that obtained by Wilks. The present paper deals with this problem.

Define \( \Sigma_1 = \sigma^2[(I-r)e' + pe'e] \), where \( e = (1, \ldots, 1) \), and regions

\(^{1}\)Supported in part by the National Science Foundation.
\[ \omega_1 = \{ \theta, \Sigma_1 : \theta_1 = \cdots = \theta_p, \Sigma_1 > 0 \}, \]
\[ \omega_2 = \{ \theta, \Sigma_1 : \theta_1 = \cdots = \theta_k, -\infty < \theta_j < \infty, j = k+1, \ldots, p, \Sigma_1 > 0 \}, \]
\[ \omega_3 = \{ \theta, \Sigma_1 : -\infty < \theta_j < \infty, j = 1, \ldots, p, \Sigma_1 > 0 \}, \]

where \( \Sigma_1 > 0 \) means that \( \Sigma_1 \) is positive definite. The hypothesis \((\theta, \Sigma_1) \in \omega_1 \) versus \( \omega_3 \) is the Wilks hypothesis, we now consider \( \omega_1 \) versus \( \omega_2 \), and \( \omega_2 \) versus \( \omega_3 \). For each problem, the likelihood ratio statistic (LRS), its central and non-central distributions are obtained.

2. **Derivation of Tests.** Given a sample of size \( N \), we have the sufficient statistic \((\bar{x}, S)\), where \( \bar{x} \) is the sample mean and \( S \) is the matrix of sample cross-products. Then \( \bar{x} \) and \( S \) are independently distributed, \( L(\bar{x}) = N(\theta, \Sigma/N) \), \( L(S) = W(\Sigma, p, n) \), where \( n = N-1 \), i.e., \( \bar{x} \) and \( S \) have the joint density function

\[
(2.1) \quad p(\bar{x}, S) = c |\Sigma|^{-N/2} |S|^{(n-p-1)/2} \exp - \frac{1}{2} [N(\bar{x}-\theta) \Sigma^{-1} (\bar{x}-\theta)' + \text{tr} \Sigma^{-1} S].
\]

The maximum likelihood estimator (MLE) of the parameters for \((\theta, \Sigma_1) \in \omega_1 \) or \( \omega_3 \) are known, so that we need only consider the case \( \omega_2 \). Actually, a slight modification of the case \( \omega_2 \) yields the results for \( \omega_1 \) and \( \omega_3 \).
To obtain the MLE for \( \omega_2 \), let

\[
y = \bar{x} \bar{\Gamma}, \quad S = \bar{\Gamma} S \bar{\Gamma}^\prime, \text{ where } \bar{\Gamma} = \begin{pmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{pmatrix},
\]

and \( \Gamma_1 : k \times k, \Gamma_2 : \ell \times \ell \) (\( k + \ell = p \)) are orthogonal matrices with first row \( e / \sqrt{k} \) and \( e / \sqrt{\ell} \), respectively. Then

\[
\mathcal{L}(y) = N((\hat{\Theta}_1, \hat{\Theta}_2), \bar{\Gamma} \Sigma \bar{\Gamma}^\prime / N), \quad \mathcal{L}(\bar{s}) = W(\bar{\Gamma} \Sigma \bar{\Gamma}^\prime, p, n).
\]

In \( \omega_2 \), \( \hat{\Theta} = \Theta_0 e \), so that \( \hat{\Theta}_1 = \Theta_0 \sqrt{k} (1, 0, \ldots, 0), \hat{\Theta}_2 = \eta \), and

\[
\bar{\Gamma} \Sigma \bar{\Gamma}^\prime = \begin{pmatrix} D_1 & M \\ M^\prime & D_2 \end{pmatrix},
\]

where

\[
D_1 = \text{diag } \sigma^2 (1 - \rho + k \rho, 1 - \rho, \ldots, 1 - \rho),
\]

\[
D_2 = \text{diag } \sigma^2 (1 - \rho + \ell \rho, 1 - \rho, \ldots, 1 - \rho),
\]

\[
M = (m_{ij}), \quad m_{ij} = 0 \text{ for } (i,j) \neq (1,1) \text{ and } m_{11} = \sigma^2 p \sqrt{k \ell}.
\]

A simplification can be achieved by the reparameterization

\[
\tau = \sigma^2 (1 - \rho), \quad \varphi = \sigma^2 \rho / \tau,
\]

so that \( D_1 = \text{diag } \tau (1 + k \rho, 1, \ldots, 1), D_2 = \text{diag } \tau (1 + \ell \rho, 1, \ldots, 1), \)

\[
m_{11} = \sqrt{k \ell} \tau \varphi. \text{ Now relabel the variables } (y, \bar{s}):\]
\[ u = (u_1, u_2) = (y_1, y_{k+1}), \quad \dot{v} = (y_2, \ldots, y_k), \]

\[ \ddot{v} = (y_{k+2}, \ldots, y_p), \quad v = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}, \]

\[ V_{11} = \begin{pmatrix} \bar{V}_{11} & \bar{V}_{12} \\ \bar{V}_{21} & \bar{V}_{22} \end{pmatrix} = \begin{pmatrix} \bar{s}_{11} & \bar{s}_{1,k+1} \\ \bar{s}_{1,k+1} & \bar{s}_{k+1,k+1} \end{pmatrix}, \]

where \( V \) is a permutation transformation of \( \bar{S} \). Then \( u, \dot{v}, \ddot{v}, V \) are independently distributed.

\[ \mathcal{L}(u) = N(\mu, \tau\psi/S), \quad \mathcal{L}(\dot{v}) = N(0, \tau I/N), \]

\[ \mathcal{L}(\ddot{v}) = N(\ddot{v}, \tau I/N), \quad \mathcal{L}(v) = \tilde{w}(\tau\psi, \psi, p, n), \]

where \( \mu = Eu, \ddot{v} = Ev, \) and

\[ \psi = \begin{pmatrix} 1 + k\varphi & \sqrt{kk'} \varphi \\ \sqrt{kk'} \varphi & 1 + k\varphi \end{pmatrix} = I + \varphi a' a, \quad a = (\sqrt{k}, \sqrt{k'}). \]

The joint density function of \( u, v = (\dot{v}, \ddot{v}), \) and \( V \) is

\[ p(u,v,V) = \frac{(n-p-1)/2}{\tau^{N/2} |\psi|^{N/2}} \exp - \frac{N}{2\tau} [(u-\mu) \psi^{-1}(u-\mu)' + (\ddot{v}-\ddot{v})(\ddot{v}-\ddot{v})'] \]

\[ \times \exp - \frac{1}{2\tau} [N \psi \psi' + \text{tr } \psi^{-1} V_{11} + \text{tr } V_{22}]. \]

It is immediate that
\( \hat{\mu} = u, \quad \hat{\psi} = \psi, \quad pN^2 = N \ddot{\psi} + \text{tr} \psi^{-1} \psi_{11} + \text{tr} \psi_{22} \),

so that

\[
\begin{align*}
\max_{\mu, \psi, \tau} p(u, \psi, \psi) &= \frac{c|V|^{(n-p-1)/2} e^{-\frac{1}{2} \frac{pN}{(pN)^{N/2}}}}{|\psi|^{N/2} (N \ddot{\psi} + \text{tr} \psi^{-1} \psi_{11} + \text{tr} \psi_{22})^{(p-l)N/2}}.
\end{align*}
\]

Since \( |\psi| = 1 + pp, \psi^{-1} = I - \varphi a' a/(1+pp) \), we need to minimize

\[
(1 + pp) [N \ddot{\psi} + \text{tr} \psi - \varphi a \psi_{11} a'/(1+pp)]^P.
\]

This is easily found to be \( \hat{\varphi} = aV_{11} a' - \text{tr} \psi - N \ddot{\psi} \), and hence

\[
\begin{align*}
\sup_{\omega_2} p(u, \psi, \psi) &= \frac{c|V|^{(n-p-1)/2} e^{-\frac{1}{2} \frac{pN}{(pN)^{N/2}}}}{(aV_{11} a')^{N/2} (N \ddot{\psi} + \text{tr} \psi - aV_{11} a' / p)^{(p-l)N/2}}.
\end{align*}
\]

In a similar way, when \( \omega_3 \) holds, (2.4) is replaced by

\[
\begin{align*}
\mathcal{L}(u) &= N(\mu, \tau \psi / N), \quad \mathcal{L}(\hat{\psi}) = N(\hat{\psi}, \tau \bar{\psi} / N),
\end{align*}
\]

\[
\begin{align*}
\mathcal{L}(\psi) &= N(\psi, \tau \bar{\psi} / N), \quad \mathcal{L}(\hat{\psi}) = W(\tau(\frac{\psi}{0} \frac{1}{1}), p, n),
\end{align*}
\]

so that

\[
\begin{align*}
\sup_{\omega_2} p(u, \psi, \psi) &= \frac{c|V|^{(n-p-1)/2} e^{-\frac{1}{2} \frac{pN}{(pN)^{N/2}}}}{(aV_{11} a')^{N/2} (\text{tr} \psi - aV_{11} a' / p)^{N/2}}.
\end{align*}
\]
When $\omega_1$ holds, it is simpler to let $x^* = x^*_T$, $S^* = \Gamma^* S \Gamma^*$, where $\Gamma^* : p \times p$ is an orthogonal matrix with first row $e / \sqrt{p}$. By a development similar to that leading to (2.6), we obtain

\begin{equation}
\sup_{\omega_1} p(x,S) = \frac{c|S|^{(n-p-1)/2} e^{-\frac{1}{2} \frac{pN}{2} (pN)(p-1)N/2 (p-1)(p-1)N/2}}{(eS'e')^{N/2} \left( \text{tr } S - eS'e'/p + \bar{x}x - (xe'/\sqrt{p})^2 \right)^{(p-1)N/2}}.
\end{equation}

Further, note that $|V| = |S|$, $\text{tr } V = \text{tr } S$, $k v_{11} = eS_{11} e'$, $k v_{22} = eS_{22} e'$, $\sqrt{k} \bar{V} v_{12} = e S_{12} e'$, where $S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$, $S_{11} : k \times k$. Consequently, $aV_{12}a' = eS'e'$. Let $W_m = I_m - e'e/\sqrt{m}$, and $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_k)$, so that

\begin{align*}
\ddot{x} \ W_k \dot{x} &= \ddot{x} \ x' - (\ddot{x} e'/\sqrt{k})^2 = yy' - y^2_1 = \dot{v} \dot{v}', \\
\bar{x} \ W_p \bar{x}' &= yy' - (y \bar{\Gamma}' e' e \bar{\Gamma} y')/p \\
&= uu' + \dot{v} \dot{v}' + \psi \psi' - (u_1 \sqrt{k} + u_2 \sqrt{k})^2 / p \\
&= vv' + (u_1 \sqrt{k} - u_2 \sqrt{k})^2 / p.
\end{align*}

From (2.6), (2.8), and (2.9), the LRS for testing $\omega_1$ versus $\omega_2$ (labeled $\lambda_{12}$), and for testing $\omega_1$ versus $\omega_3$ (labeled $\lambda_{23}$), are given by
\( (2.10) \quad \lambda_{12}^{2/(p-1)} = \frac{\text{tr} S - eS'e'/p + N \bar{x}_k \bar{x}'}{\text{tr} S - eS'e'/p + N \bar{x}_p \bar{x}'} \)

\[ = \frac{\text{tr} V - aV_{11} a'/p + N \bar{v}v'}{\text{tr} V - aV_{11} a'/p + N \bar{v}v'} + N(u_1 \sqrt{\ell} - u_2 \sqrt{k})^2/p' \]

\[ (2.11) \quad \lambda_{23}^{2/(p-1)} = \frac{\text{tr} S - eS'e'/p}{\text{tr} S - eS'e'/p + N x_k x} \]

\[ = \frac{\text{tr} V - aV_{11} a'/p}{\text{tr} V - aV_{11} a'/p + \bar{v}v} \]

It is simpler to consider monotone functions of \( \lambda_{12} \) and \( \lambda_{23} \), namely

\[ \lambda_{12}^* = \frac{N \bar{v}v' + N(u_1 \sqrt{\ell} - u_2 \sqrt{k})^2/p}{\text{tr} V_{11}(I - a'a/p) + \text{tr} V_{22} + N \bar{v}v'} \]

\[ \lambda_{23}^* = \frac{N \bar{v}v'}{\text{tr} V_{11}(I - a'a/p) + \text{tr} V_{22}} \]

respectively. We now obtain the central distribution of \( \lambda_{12}^* \) and \( \lambda_{23}^* \).

Whether the parameters are in \( \omega_1, \omega_2, \) or \( \omega_3 \), the statistics \( \lambda_{12}^* \) and \( \lambda_{23}^* \) are invariant under \((u,v,v) \to (cu, cv, cv')\), so that we can assume \( \tau = 1 \). Consequently, \( \zeta(\text{tr} V_{22}) = \chi^2_{(p-2)n} \). Since \( V_{11} \) and \( V_{22} \) are independent, we need the distribution of \( \text{tr} V_{11}(I-a'a/p) \).
Lemma. If \( L(Z) = W(I + \omega a'a, d, n), \alpha > -1/aa', \) then \( \text{tr} \ Z(I-a'a/aa') \) has a \( \chi^2_{(d-1)n} \) distribution independent of \( \alpha. \)

Proof. Let \( U = \Gamma Z \Gamma', \) where the first row of \( \Gamma \) is \( a/\sqrt{aa'}, \) then 
\[
L(U) = W(\Delta, d, n), \quad \Delta = \Gamma(I + \omega a'a) \Gamma' = \text{diag}(1 + \omega aa', 1, \ldots, 1).
\]
But \( \text{tr} \ Z(I - a'a/aa') = \text{tr} \ U \Gamma(I - a'a/aa') \Gamma' = \sum_2^d u_{ii}. \) The \( u_{ii} \) are independent with \( L(u_{ii}) = \chi^2_{n}, \) from which the result follows.

Under \( \omega_2, \) we have from (2.4) that 
\[
L(N \bar{v} v') = \chi^2_{k-1},
\]
\[
L(N \bar{v} v') = \chi^2_{k-1}(\bar{v} v') \quad \text{(i.e., the non-central \( \chi^2 \) distribution with parameter \( \bar{v} v' \)).}
\]
The statistic \( (u_1 \sqrt{\bar{k}} - u_2 \sqrt{\bar{k}})/\sqrt{p} \) is normally distributed with mean \( (\mu_1 \sqrt{\bar{k}} - \mu_2 \sqrt{\bar{k}})/\sqrt{p} = \mu^* \) and variance 
\[
[(1+\bar{k})\bar{v} + (1+\bar{k})\bar{k} - 2\bar{k}\bar{k}]/p = 1, \quad \text{and hence} \quad L((u_1 \sqrt{\bar{k}} - u_2 \sqrt{\bar{k}})^2/p) = \chi^2_{1}(\mu^*). \]

In conclusion, we have the result that the non-central distribution of \( \chi^2_{k-1}[(p-1)n + k-1]/\ell \) is that of a non-central \( F \)-distribution with parameter \( \delta = \bar{v} \bar{v}' + \mu^* \) and \((\ell, (p-1)n + k-1)\) degrees of freedom. In terms of the original parameters,
\[
\delta = \bar{v} \bar{v}' + k\bar{k}(\bar{\theta} - \bar{\theta})^2/p,
\]
where \( \bar{\theta} = \frac{1}{k} \sum_1^k \theta_1/k, \quad \bar{\theta} = \frac{1}{k+1} \sum_1^{k+1} \theta_1/k. \) Under \( \omega_1, \theta = \theta_0 e, \) so that \( \delta = 0, \) as is to be expected.

Under \( \omega_2, \) we have \( L(\bar{v}) = N(\bar{v}, \tau I/N), \) so that \( L(N \bar{v} \bar{v}') = \chi^2_{k-1}(\bar{v} \bar{v}') \). The denominator of \( \chi^2_{k-1} \) has a \( \chi^2_{(p-1)n} \) distribution, and hence the non-central distribution of \( \chi^2_{k-1}[(p-1)n/(k-1)] \) is that of
a non-central F-distribution with parameter \( \delta = \nu \nu' \) and

\((k-1, (p-1)n)\) degrees of freedom. In terms of the original parameters,

\[ \delta = \delta W_k \delta', \]

and under \( \omega_2 \), \( \delta = \delta_0 e \), so that \( \delta = 0 \).

REFERENCES

