The Annals of Mathematical Statistics
1970, Vol. 41, No. 2, 683–687

AN EXTENSION OF WILKS' TEST FOR THE EQUALITY OF MEANS

BY I. OLKIN[1] AND S. S. SHRIKHANDE

Stanford University and University of Bombay

1. Introduction. When $x = (x_1, \cdots, x_p)$ has a multivariate normal distribution with mean vector θ and covariance matrix Σ, 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., $\dot{\theta} = (\theta_1, \cdots, \theta_k) = 0$, and we wish to test that the remaining part is zero, i.e., $\ddot{\theta} = (\theta_{k+1}, \cdots, \theta_p) = 0$ against $\ddot{\theta} \neq 0$, then the T^2-statistic should not be used. The likelihood ratio test is given by $(1+T_k^{\,2})/(1+T_p^{\,2})$, where $T_m^{\,2}$ 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 $-\infty < \theta_j < \infty$, $j = 1, \cdots, p$, 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 $\Sigma_I = \sigma^2[(1-\rho)I + \rho e'e]$, where $e = (1, \cdots, 1)$, and regions

$$\begin{aligned}
&\omega_1 = \{\theta, \Sigma_I : \theta_1 = \cdots = \theta_p, \Sigma_I > 0\},\\
(1.1)\quad &\omega_2 = \{\theta, \Sigma_I : \theta_1 = \cdots = \theta_k, \theta_{k+1} = \cdots = \theta_p, \Sigma_I > 0\},\\
&\omega_3 = \{\theta, \Sigma_I : \theta_1 = \cdots = \theta_k, -\infty < \theta_j < \infty, j = k+1, \cdots, p, \Sigma_I > 0\},\\
&\omega_4 = \{\theta, \Sigma_I : -\infty < \theta_j < \infty, j = 1, \cdots, p, \Sigma_I > 0\},
\end{aligned}$$

where $\Sigma_I > 0$ means that Σ_I is positive definite. The hypothesis $(\theta, \Sigma_I) \in \omega_1$ versus ω_4 is the Wilks hypothesis; we now consider testing ω_1 versus ω_2, ω_1 versus ω_3, and ω_2 versus ω_3. For each problem the likelihood ratio statistic (LRS) and its non-central distribution are obtained. In order to avoid degeneracies, we assume $p-1 \geqq k \geqq 2$. However, note that the test of ω_3 versus ω_4, for example, reduces to Wilks hypothesis when $p = k$.

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, $\mathscr{L}(\bar{x}) = N(\theta, \Sigma/N)$, $\mathscr{L}(S) = W(\Sigma, p, n)$, where $n = N-1$, i.e., $\bar{x}$ and S have the joint density function

$$p(\bar{x}, S) = c\,|\Sigma|^{-N/2}\,|S|^{(n-p-1)/2} \exp\{-\tfrac{1}{2}\operatorname{tr}\Sigma^{-1}[S + N(\bar{x}-\theta)'(\bar{x}-\theta)]\},$$

where $c = N^{p/2}[2^{pN/2}\pi^{p(p+1)/4}\prod_1^p \Gamma(\tfrac{1}{2}(n-i+1))]^{-1}$.

Received December 19, 1958; revised September 5, 1969.
[1] Supported in part by the National Science Foundation.

The maximum likelihood estimators (MLE) of the parameters for $(\theta, \Sigma_I) \in \omega_1$ or ω_4 are known, so that we need only consider the cases ω_2 and ω_3. Actually, a slight modification of the case ω_2 yields the result for all ω's.

To obtain the MLE's, we first reduce the problem by letting

$$(2.1) \qquad y = \bar{x}\bar{\Gamma}, \qquad \bar{S} = \bar{\Gamma}' S \bar{\Gamma}, \qquad \text{where } \bar{\Gamma} = \begin{pmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{pmatrix},$$

and $\Gamma_1 : k \times k$, $\Gamma_2 : \ell \times \ell$ $(k + \ell = p)$ are orthogonal matrices with first column $e'/k^{\frac{1}{2}}$ and $e'/\ell^{\frac{1}{2}}$, respectively. Define $\dot{\theta} = (\theta_1, \cdots, \theta_k)$, $\ddot{\theta} = (\theta_{k+1}, \cdots, \theta_p)$, then

$$\mathscr{L}(y) = N((\dot{\theta}\Gamma_1, \ddot{\theta}\Gamma_2), \bar{\Sigma}_I/N), \qquad \mathscr{L}(\bar{S}) = W(\bar{\Sigma}_I, p, n),$$

$$\bar{\Sigma}_I \equiv \bar{\Gamma}' \Sigma_I \bar{\Gamma} = \begin{pmatrix} D_1 & M \\ M' & D_2 \end{pmatrix},$$

where

$$D_1 = \operatorname{diag} \sigma^2(1 - \rho + k\rho, 1 - \rho, \cdots, 1 - \rho),$$

$$D_2 = \operatorname{diag} \sigma^2(1 - \rho + \ell\rho, 1 - \rho, \cdots, 1 - \rho),$$

$$M = (m_{ij}), \quad m_{11} = \sigma^2 \rho (k\ell)^{\frac{1}{2}}, \quad \text{and} \quad m_{ij} = 0 \quad \text{for} \quad (i, j) \neq (1, 1).$$

A further simplification can be achieved by the reparameterization

$$\tau = \sigma^2(1 - \rho) > 0, \qquad \varphi = \sigma^2 \rho / \tau > -1/p,$$

so that $D_1 = \operatorname{diag} \tau(1 + k\varphi, 1, \cdots, 1)$, $D_2 = \operatorname{diag} \tau(1 + \ell\varphi, 1, \cdots, 1)$, $m_{11} = (k\ell)^{\frac{1}{2}}\tau\varphi$. Now relabel the variables $(y, \bar{S})$:

$$(2.2) \qquad u = (u_1, u_2) \equiv (y_1, y_{k+1}), \qquad \dot{v} = (y_2, \cdots, y_k),$$

$$\ddot{v} = (y_{k+2}, \cdots, y_p), \qquad v = (\dot{v}, \ddot{v}), \qquad V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix},$$

$$v_{11} = \begin{pmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{pmatrix} = \begin{pmatrix} \bar{s}_{11} & \bar{s}_{1,k+1} \\ \bar{s}_{1,k+1} & \bar{s}_{k+1,k+1} \end{pmatrix},$$

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

$$(2.3) \quad \mathscr{L}(u) = N(\mu, \tau\psi/N), \qquad \mathscr{L}(v) = N(\nu, \tau I/N), \qquad \mathscr{L}(V) = W\left(\tau \begin{pmatrix} \psi & 0 \\ 0 & I \end{pmatrix}, p, N\right),$$

where $\mu = (\dot{\theta}e'/k^{\frac{1}{2}}, \ddot{\theta}e'/\ell^{\frac{1}{2}})$, $\dot{\nu} = \dot{\theta}\Gamma_{12}$, $\ddot{\nu} = \ddot{\theta}\Gamma_{22}$, $\Gamma_1 = (e'/k^{\frac{1}{2}}, \Gamma_{12})$, $\Gamma_2 = (e'/\ell^{\frac{1}{2}}, \Gamma_{21})$,

$$\psi = \begin{pmatrix} 1 + k\varphi & (k\ell)^{\frac{1}{2}}\varphi \\ (k\ell)^{\frac{1}{2}}\varphi & 1 + \ell\varphi \end{pmatrix} = I + \varphi a' a, \qquad a = (k^{\frac{1}{2}}, \ell^{\frac{1}{2}}).$$

The joint density function of u, v, and V is

$$(2.4)\quad p(u,v,V) = c(V)\exp\{-\tfrac{1}{2}\tau^{-1}[N(u-\mu)\psi^{-1}(u-\mu)'+(v-\nu)(v-\nu)' + \operatorname{tr}\psi^{-1}V_{11}+\operatorname{tr}V_{22}]\},$$

where $c(V) = c\,|V|^{(n-p-1)/2}$.

In terms of the new variables, the hypotheses (1.1) become

$$\begin{aligned}
&\omega_1 = \{\mu,\nu,\tau,\psi: \mu = \theta_0(k^{\frac{1}{2}}, l^{\frac{1}{2}}), \nu = 0, \tau > 0, \psi > 0\},\\
(2.5)\quad &\omega_2 = \{\mu,\nu,\tau,\psi: -\infty < \mu < \infty, \nu = 0, \tau > 0, \psi > 0\},\\
&\omega_3 = \{\mu,\nu,\tau,\psi: -\infty < \mu < \infty, \dot{\nu} = 0, -\infty < \ddot{\nu} < \infty, \tau > 0, \psi > 0\},\\
&\omega_4 = \{\mu,\nu,\tau,\psi: -\infty < \mu < \infty, -\infty < \nu < \infty, \tau > 0, \psi > 0\}.
\end{aligned}$$

It is immediate that for ω_2,

$$(2.6)\qquad \hat{\mu} = \mu, \qquad \hat{\nu} = 0, \qquad pN\hat{\tau} = Nvv' + \operatorname{tr}\psi^{-1}v_{11} + \operatorname{tr}V_{22},$$

so that

$$\max_{\mu,\nu,\tau} p(u,v,V) = \frac{c(V)\,e^{-pN/2}(pN)^{pN/2}}{|\psi|^{N/2}(Nvv' + \operatorname{tr}\psi^{-1}V_{11} + \operatorname{tr}V_{22})^{pN/2}}.$$

Since $|\psi| = 1 + p\varphi$, $\psi^{-1} = I - \varphi a'a/(1+p\varphi)$, we need to minimize $(1+p\varphi)[A - \varphi B/(1+p\varphi)]^p$, over $\varphi > -1/p$, where $A = Nvv' + \operatorname{tr}V$, $B = aV_{11}a'$. This is easily found to be $\hat{\varphi} = (B-A)/(pA-B)$, and hence

$$(2.7)\qquad \sup_{\omega_2} p(u,v,V) = \frac{c(V)\,e^{-pN/2}(pN)^{pN/2}(p-1)^{(p-1)N/2}}{(aV_{11}a')^{N/2}[pNvv' + p\operatorname{tr}V - aV_{11}a']^{(p-1)N/2}}.$$

Similarly, when ω_3 or ω_4 holds, (2.6) is replaced by

$$(2.8)\qquad \hat{\mu} = u, \qquad \hat{\dot{\nu}} = 0, \qquad \hat{\ddot{\nu}} = \ddot{v}, \qquad pN\hat{\tau} = N\dot{v}\dot{v}' + \operatorname{tr}\psi^{-1}V_{11} + \operatorname{tr}V_{22},$$

$$(2.9)\qquad \hat{\mu} = u, \qquad \hat{\nu} = v, \qquad pN\hat{\tau} = \operatorname{tr}\psi^{-1}V_{11} + \operatorname{tr}V_{22},$$

respectively, so that

$$(2.10)\qquad \sup_{\omega_3} p(u,v,V) = \frac{c(V)\,e^{-pN/2}(pN)^{pN/2}(p-1)^{(p-1)N/2}}{(aV_{11}a')^{N/2}(pN\dot{v}\dot{v}' + p\operatorname{tr}V - aV_{11}a')^{(p-1)N/2}},$$

$$(2.11)\qquad \sup_{\omega_4} p(u,v,V) = \frac{c(V)\,e^{-pN/2}(pN)^{pN/2}(p-1)^{(p-1)N/2}}{(aV_{11}a')^{N/2}(p\operatorname{tr}V - aV_{11}a')^{(p-1)N/2}}.$$

When ω_1 holds, we could make an alternative transformation to (2.1), namely, $x^* = \bar{x}\Gamma^*$, $S^* = \Gamma^{*\prime}S\Gamma^*$, where $\Gamma^*: p \times p$ is an orthogonal matrix with first column $e'/p^{\frac{1}{2}}$. However, the present formulation suffices with an additional minimization. Now $\hat{\nu} = 0$ and $\mu = \theta_0 a$. From (2.4) we easily find that the minimum of $(u-\theta_0 a)\psi^{-1}(u-\theta_0 a)'$ is achieved at

$$\hat{\theta}_0 = (a\psi^{-1}u')/(a\psi^{-1}a'),$$

from which, after some algebraic manipulation, we obtain

$$(u-\hat{\mu})\psi^{-1}(u-\hat{\mu})' = u[I-a'a/p]u' = (u_1 l^{\frac{1}{2}}-u_2 k^{\frac{1}{2}})^2/p$$

independent of φ. Hence, by a parallel argument

(2.12) $\sup_{\omega_1} p(u, v, V)$

$$= \frac{c(V)\, e^{-pN/2}\,(pN)^{pN/2}\,(p-1)^{(p-1)N/2}}{(aV_{11}a')^{N/2}[pNvv'+N(u_1 l^{\frac{1}{2}}-u_2 k^{\frac{1}{2}})^2+p\operatorname{tr} V-aV_{11}a']^{(p-1)N/2}}\,.$$

At this point it may be useful to translate the variables (u, v, V) into the original variables, $(\bar{x}, S)$. Note that $|V| = |S|$, $\operatorname{tr} V = \operatorname{tr} S$, $kv_{11} = eS_{11}e'$, $lv_{22} = eS_{22}e'$, $(kl)^{\frac{1}{2}}v_{12} = eS_{12}e'$, where

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix},$$

$S_{11}: k \times k$. Consequently, $aV_{11}a' = eSe'$. Let $W_m = I_m - e'e\,m^{\frac{1}{2}}$, $\dot{x} = (\bar{x}_1, \cdots, \bar{x}_k)$, and $\dot{y} = (y_1, \cdots, y_k)$, so that

$$\begin{aligned}
\dot{x}W_k\dot{x}' &= \dot{x}\dot{x}' - (\dot{x}e'/k^{\frac{1}{2}})^2 = yy' - {y_1}^2 = \dot{v}\dot{v}',\\
\bar{x}W_p\bar{x}' &= yy' - (y\bar{\Gamma}'e'e\bar{\Gamma}y')/p\\
&= uu' + \dot{v}\dot{v}' + \ddot{v}\ddot{v}' - (u_1 k^{\frac{1}{2}} + u_2 l^{\frac{1}{2}})^2/p\\
&= vv' + (u_1 l^{\frac{1}{2}} - u_2 k^{\frac{1}{2}})^2/p.
\end{aligned}$$

From (2.7), (2.10), (2.11), and (2.12), the LRS for testing ω_i versus ω_j, denoted by λ_{ij}, may be determined. It is simpler to consider instead a monotone function L_{ij}, of λ_{ij}, obtained from $\lambda_{ij}^{2/(p-1)N} = 1/(1+L_{ij})$. Thus

$$\begin{aligned}
L_{12} &= \frac{N(u_1 l^{\frac{1}{2}}-u_2 k^{\frac{1}{2}})^2/p}{Nvv'+\operatorname{tr} V_{11}(I-a'a/p)+\operatorname{tr} V_{22}}\,,\\
L_{13} &= \frac{N\ddot{v}\ddot{v}'+N(u_1 l^{\frac{1}{2}}-u_2 k^{\frac{1}{2}})^2/p}{N\dot{v}\dot{v}'+\operatorname{tr} V_{11}(I-a'a/p)+\operatorname{tr} V_{22}}\,,\\
L_{23} &= \frac{N\ddot{v}\ddot{v}'}{N\dot{v}\dot{v}'+\operatorname{tr} V_{11}(I-a'a/p)+\operatorname{tr} V_{22}}\,,\\
L_{34} &= \frac{N\dot{v}\dot{v}'}{\operatorname{tr} V_{11}(I-a'a/p)+\operatorname{tr} V_{22}}\,.
\end{aligned}$$

We now obtain the distribution of the L_{ij}. Whether the parameters are in ω_1, ω_2, ω_3, or ω_4, the statistics L_{ij} are invariant under $(u, v, V) \to (\alpha u, \alpha v, \alpha^2 V)$, so that we can assume $\tau = 1$. Consequently, $\mathscr{L}(\operatorname{tr} V_{22}) = \chi^2_{(p-2)n}$. Since V_{11} and V_{22} are independent, we need the distribution of $\operatorname{tr} V_{11}(I-a'a/p)$.

LEMMA. *If* $\mathscr{L}(Z) = W(I+\alpha a'a, d, n)$, $\alpha > -1/aa'$, *then* $\operatorname{tr} Z(I-a'a/aa')$ *has a* $\chi^2_{(d-1)n}$ *distribution independent of* α.

PROOF. Let $U = \Gamma Z \Gamma'$, where the first row of Γ is $a/(aa')^{\frac{1}{2}}$, then $\mathscr{L}(U) = W(\Delta, d, n)$, where $\Delta = \Gamma(I+\alpha a'a)\Gamma' = \text{diag}(1+\alpha aa', 1, \cdots, 1)$. 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 $\mathscr{L}(u_{ii}) = \chi_n{}^2$, from which the result follows. □

Hence, regardless of which hypothesis is true,

$$\mathscr{L}\{\text{tr}\, V_{11}(I - a'a/p) + \text{tr}\, V_{22}\} = \chi^2_{(p-1)n}.$$

From (2.3) and (2.5), we see that under ω_2, $\mathscr{L}(Nvv') = \chi^2_{p-2}$, $N^{\frac{1}{2}}(u_1 \ell^{\frac{1}{2}} - u_2 k^{\frac{1}{2}})/p^{\frac{1}{2}}$ is normally distributed with mean $N^{\frac{1}{2}}(\mu_1 \ell^{\frac{1}{2}} - \mu_2 k^{\frac{1}{2}})/p^{\frac{1}{2}} \equiv \mu^*$, and variance $[(1+k\varphi)\ell + (1+\ell\varphi)k - 2\ell k\varphi]/p = 1$; hence $\mathscr{L}(N(u_1 \ell^{\frac{1}{2}} - u_2 k^{\frac{1}{2}})^2/p) = \chi_1{}^2(\mu^{*2})$, i.e., the noncentral χ^2 distribution with parameter μ^{*2} and 1 df. Thus the noncentral distribution of $L_{12}[(p-1)n+p-2]$ is that of a noncentral F-distribution with parameter $\delta = \mu^{*2}$ and $(1, (p-1)n+p-2)$ degrees of freedom. In terms of the original parameters,

$$\delta = k\ell(\dot{\theta} - \ddot{\theta})^2/p,$$

where $\dot{\theta} = \sum_1^k \theta_i/k$, $\ddot{\theta} = \sum_{k+1}^p \theta_i/\ell$. Under ω_1, $\theta = \theta_0 e$, so that $\delta = 0$, as is to be expected.

Under ω_3

$$\mathscr{L}(N\dot{v}\dot{v}') = \chi^2_{k-1}, \quad \mathscr{L}(N\ddot{v}\ddot{v}') = \chi^2_{l-1}(N\ddot{v}\ddot{v}'), \quad \mathscr{L}(N(u_1 \ell^{\frac{1}{2}} - u_2 k^{\frac{1}{2}})^2/p = \chi_1{}^2(\mu^{*2}),$$

so that $L_{13}[(p-1)n+k-1]/\ell$ has a noncentral F-distribution with parameter

$$\delta = N\ddot{v}\ddot{v}' + \mu^{*2} = N\ddot{\theta}W_\ell\ddot{\theta}' + Nk\ell(\dot{\theta} - \ddot{\theta})^2/p,$$

and $(\ell, (p-1)n+k-1)$ degrees of freedom. When ω_1 holds, $\delta = 0$, and the F-distribution is central. Similarly, $L_{23}[(p-1)n+k-1]/(\ell-1)$ has a noncentral F-distribution with parameter

$$\delta = N\ddot{v}\ddot{v}' = N\ddot{\theta}W_\ell\ddot{\theta}',$$

and $(\ell-1, (p-1)n+k-1)$ degrees of freedom.

Under ω_4, $\mathscr{L}(N\dot{v}\dot{v}') = \chi^2_{k-1}(N\dot{v}\dot{v}')$, so that $L_{33}[(p-1)n/(k-1)]$ has a noncentral F-distribution with parameter

$$\delta = N\dot{\theta}W_k\dot{\theta}',$$

and $(k-1, (p-1)n)$ degrees of freedom.

REFERENCES

[1] KABE, D. G. (1965). On the noncentral distribution of Rao's U statistic. *Ann. Inst. Statist. Math.* **17** 75–80.

[2] OLKIN, I. and SHRIKHANDE, S. S. (1954). On a modified T^2 problem (Abstract). *Ann. Math. Statist.* **25** 808.

[3] WILKS, S. S. (1946). Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. *Ann. Math. Statist.* **17** 257–281.