TESTING FOR EQUALITY OF MEANS, EQUALITY OF VARIANCES
AND EQUALITY OF COVARIANCES UNDER RESTRICTIONS UPON
THE PARAMETER SPACE (PRELIMINARY REPORT)

BY

LEON JAY GLESER and INGRAM OLKIN

TECHNICAL REPORT NO. 25
July 30, 1967

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT 6681

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
TESTING FOR EQUALITY OF MEANS, EQUALITY OF VARIANCES
AND EQUALITY OF COVARIANCES UNDER RESTRICTIONS UPON
THE PARAMETER SPACE (PRELIMINARY REPORT)

By

Leon Jay Gleser *
The Johns Hopkins University

and

Ingram Olkin
Stanford University

TECHNICAL REPORT NO. 25
July 30, 1967

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT 6681

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

*Supported in part by the Office of Naval Research, Contract No. 4010(09).
1. Introduction

Suppose that the $p$-dimensional random row vector $x = (x_1, x_2, \ldots, x_p)$ has a multivariate normal distribution with mean vector $\mu = (\mu_1, \mu_2, \ldots, \mu_p)$ and covariance matrix $\Sigma = (\sigma_{ij})$. We are interested in testing the hypothesis that the distribution of $x$ is symmetric with respect to the components $x_1, x_2, \ldots, x_p$. Because of the assumption that $x$ has a normal distribution, this hypothesis is equivalent to the hypothesis $H_{\text{MVC}}$ that the components of the mean vector $\mu$ are equal ($\mu = (\eta, \ldots, \eta)$, $\eta$ a constant), that the variances are equal ($\sigma_{ii}^2 = \sigma^2$, $i = 1, 2, \ldots, p$), and that the covariances are equal ($\sigma_{ij}^2 = \sigma^2 \rho$, $i \neq j$). The less restricted hypothesis $H_{\text{VC}}$ that the variances are equal and that the covariances are equal (i.e., that $\Sigma$ has the \textit{intra-class correlational structure} $\Sigma = \sigma^2 (I - \rho I + \rho ee^T)$), where $e = (1, \ldots, 1)$, may also be of concern.

Likelihood ratio tests for testing $H_{\text{MVC}}$ against the general alternatives $\mu$ unrestricted, $\Sigma > 0$, for testing $H_{\text{VC}}$ against general alternatives, and for testing $H_{\text{MVC}}$ against $H_{\text{VC}}$ have been derived by Wilks (1946). The moments of these likelihood ratio test statistics under

\footnote{1\textsuperscript{st} Supported in part by the Office of Naval Research, Contract No. 4010(09).}
\footnote{2\textsuperscript{nd} Supported in part by the National Science Foundation, Grant No. 6681.}
their respective null distributions and exact expressions for the null distributions in certain special cases are also obtained by Wilks. Box (1949) has suggested a general method for obtaining approximations to the null distributions of these (and other) test statistics.

Wilks' (1946) interest in the above statistical problem arose from a problem in psychological testing theory in which we wish to test the hypothesis that $k$ examinations are "parallel forms" of the same examination. To test this hypothesis, the $k$ examinations are given to each of $N$ subjects. If the hypothesis is true, and if test scores for each individual have a $k$-dimensional multivariate normal distribution, then the means, variances, and covariances of the test scores should obey $H^m_{mvc}$.

However, further consideration of the above psychological problem suggests that if the hypothesis of "parallel forms" holds, then a more restricted hypothesis concerning the means, variances, and covariances is appropriate. Since "parallel forms" of an examination presumably measure a common aptitude (or aptitudes) in each subject, one expects that the $k$ examination scores for each subject are equally and positively correlated with one another. Thus, although under the hypothesis $H^m_{mvc}$ (or the hypothesis $H^c_{vc}$) we only require that the common correlation $\rho$ be restricted to the range from $-1/(p-1)$ to 1 (so as to guarantee that $\Sigma$ is positive definite), we actually are interested in the restricted hypothesis $H^c_{mvc}$ (or the restricted hypothesis $H^c_{vc}$) that requires $\rho$ to fall in the range from $\rho_0$ to 1, $\rho_0$ some nonnegative number.

An hypothesis of the form $H^c_{mvc}$ is also of concern when we are testing that the observations from a balanced one way analysis of variance design have a joint distribution obeying the assumptions of the Model II Analysis of Variance. If $x$ is one complete replication of the design
(one observation in each cell), then the Model II assumptions state that
\( x \) can be written as
\[
x = (v, v, \ldots, v) + \varepsilon = v\varepsilon + \varepsilon,
\]
where \( v \) is a scalar random variable, \( v \sim N(\theta, \tau_0^2) \), \( \varepsilon \sim N(0, \tau_1^2 I_p) \), and \( \varepsilon \) and \( \varepsilon \) are independently distributed. Consequently, if the model II assumptions hold, \( x \sim N(\theta, \tau_1^2 I_p (1-\rho) I + \rho \varepsilon \varepsilon') \), where \( 0 \leq \rho = \frac{\tau_0^2}{\tau_0^2 + \tau_1^2} \leq 1 \). To test the fit of the model II assumptions, we should test \( H_{\text{mvc}}^F \) (with \( 0 = \rho_0 \leq \rho < 1 \)) against general alternatives. Box (1950) considers the problem of testing the model II assumptions and obtains an approximate test by ignoring the restriction \( \rho \geq 0 \). Herbach (1959) has extensively studied some tests related to the model \( H_{\text{mvc}}^F \) with \( \rho_0 = 0 \), and obtains some optimal properties of these tests.

In the present paper, we are interested in testing the hypotheses \( H_{\text{mvc}}^F \) and \( H_{\text{vc}}^F \) against general alternatives. Once it has been established that one of these hypotheses is an appropriate model for the data, we are also interested in the estimation of the parameters \( \rho \) and \( \sigma^2 \). (For the case \( \rho_0 = 0 \), Herbach (1959) obtains the MLE and discusses completeness.)

In Section 2, maximum likelihood estimators for \( \rho \) and \( \sigma^2 \) under \( H_{\text{mvc}}^F \) and \( H_{\text{vc}}^F \) are obtained. The exact and asymptotic distributions of these estimators are also discussed in Section 2. Section 3 is concerned with deriving the likelihood ratio tests for \( H_{\text{mvc}}^F \) and \( H_{\text{vc}}^F \) against general alternatives, and with finding the asymptotic null distributions of these test statistics.
2. Maximum Likelihood Estimators for the Parameters of the Covariance Matrix Under $H_{mvc}^\pi$ and $H_{vc}^\pi$

2.1 Preliminaries. Assume that we have $N$ independent observations $x(1), x(2), \ldots, x(N)$ upon the random p-dimensional row vector $x$. Recall that $x$ has a multivariate normal distribution with mean vector $\mu$ and covariance matrix $\Sigma$. A sufficient statistic for $(\mu, \Sigma)$ is $(\overline{x}, S)$, where \[ \overline{x} = \left( \overline{x}_1, \overline{x}_2, \ldots, \overline{x}_p \right) = \frac{1}{N} \sum_{j=1}^{N} x(j) \] is the sample mean vector and \[ S = \sum_{j=1}^{N} (x(j) - \overline{x})(x(j) - \overline{x})' \] is the sample cross-product matrix (where $S$ is the sample covariance matrix). It is a well-known result that $\overline{x}$ and $S$ are independently distributed, that $\sqrt{N} \overline{x} \sim N(\mu, N^{-1} \Sigma)$, and that $S$ has the Wishart distribution $\mathcal{W}(\Sigma; p, n)$. The joint density of $\overline{x}$ and $S$ is thus:

\[
(2.1) \quad p(\overline{x}, S) = \left(2 \pi \right)^{-p} \left( \frac{N}{2} \right)^{p/2} \left| \frac{N-p-1}{2} \right| \exp \left\{ \frac{1}{2} \left( \text{tr} \, S^{-1} S + N \text{tr} \, \Sigma^{-1} (\overline{x} - \mu)' (\overline{x} - \mu) \right) \right\},
\]

where $n = N - 1$, $c(p, n) = \left( \frac{1}{2} \pi \right)^{p/2} \Gamma(p-1)/2 \prod_{i=1}^{p-1} \Gamma\left( \frac{n-i+1}{2} \right)$, and the range of definition for $p(\overline{x}, S)$ is $S$ positive definite ($S > 0$), $\overline{x}$ unrestricted.

We make use of the following lemma in Section 2.2.

**Lemma 2.1.** For any matrix $A$ having the form $A = aI + b e e'$:

(i) $|A - \lambda I| = (a-\lambda)^{p-1} [a + b(p-1) - \lambda]$, 

(ii) $A$ can be represented in the form $A = \Gamma D \Gamma'$, where $\Gamma$ is any $p \times p$ orthogonal matrix having first row $p^{1/2}(1,1, \ldots, 1)$, and $D = \text{diag}(a + b(p-1), a, a, \ldots, a)$.

Thus, under $H_{vc}^\pi$, the covariance matrix $\Sigma$ can be reduced to diagonal form by an orthogonal matrix whose elements are independent of the parameters. This fact allows us to considerably simplify our calculations.

Before deriving the maximum likelihood estimators of $\rho$ and $\sigma^2$ under $H_{mvc}^\pi$ and $H_{vc}^\pi$, we indicate the precise form of the parameter space for each of these two hypotheses. The parameter space $\Omega_{mvc}^\pi$ for $H_{mvc}^\pi$
takes the form:

\[(2.2) \quad \Omega_{\text{mve}}^x = \{(\mu, \Sigma) : \mu = \eta e, \eta \text{ a constant}, \Sigma = \sigma^2[(1-\rho)I + \rho ee']\}, \]
\[\sigma^2 > 0, \rho_0 \leq \rho < 1\]

the parameter space \(\Omega_{\text{vc}}^x\) for \(K_{\text{vc}}^x\) has the form:

\[(2.3) \quad \Omega_{\text{vc}}^x = \{(\mu, \Sigma) : \mu \text{ unrestricted}, \Sigma = \sigma^2[(1-\rho)I + \rho ee'], \sigma^2 > 0, \rho_0 \leq \rho < 1\}, \]

where \(\rho_0\) is a given constant, \(-1/(p-1) < \rho_0 < 1\). Note that we have allowed \(\rho_0\) to be negative. For applications to model II analysis of variance, \(\rho_0\) is zero; for applications to the psychological testing problem mentioned by Wilks (1946), \(\rho_0\) is of moderate size and positive.

2.2 Derivation of the Maximum Likelihood Estimators of \(\rho\) and \(\Sigma\)

Under \(\Omega_{\text{mve}}^x\) and \(\Omega_{\text{mc}}^x\). Since \((\bar{x}, \bar{y})'\) is a sufficient statistic for \((\mu, \Sigma)\), we begin our derivation of the maximum likelihood estimators (MLE) of \(\rho\) and \(\sigma^2\) by considering the density \((2.1)\). Motivated by Lemma 2.1 (and the remark following that lemma), we let \(y = \sqrt{n} \bar{x} \Gamma'\) and \(V = \Gamma \Sigma \Gamma'\), where \(\Gamma\) is any \(p \times p\) orthogonal matrix with first row \(\rho^{-1/2} e\), \(e = (1,1, \ldots, 1)\). Then, under \(\Omega_{\text{vc}}^x\) (and under \(\Omega_{\text{mve}}^x \subset \Omega_{\text{vc}}^x\)), we have that \(y\) and \(V\) are independently distributed, that \(y \sim N(\xi, D)\), and that \(V \sim \chi^2(D; p, n)\), where \(\xi = \sqrt{n} \mu \Gamma'\), \(D = \Gamma \Sigma \Gamma'\) = diag\(\{\sigma_1^2, \sigma_2^2, \sigma_3^2, \ldots, \sigma_p^2\}\), and

\[(2.4) \quad \sigma_1^2 = \sigma^2(1 + (p-1)\rho), \quad \sigma_2^2 = \sigma^2(1-\rho). \]

Note that the joint distribution of \(y\) and \(V\) becomes

\[(2.5) \quad p(y, V) = (2\pi)^{-p/2} \frac{n}{\sigma_1^2 \sigma_2^2 N(p-1)} |V|^{-n/2} \exp \left\{-\frac{1}{2} \frac{1}{\sigma_1^2} (v_{11} + (y_1 - \xi_1)^2) \right\}
\[+ \frac{\sigma_2^2}{\sum_{i=2}^p v_{ii} + \sum_{i=2}^p (y_i - \xi_i)^2}\].\]
From (2.5) it is easily verified that \((y, v_{11}, \frac{p}{i=2} v_{i1})\) is a sufficient statistic for \((\xi, \sigma_1^2, \sigma_2^2)\). It may be directly verified that \(y, v_{11}, \text{ and } \sum_{i=2}^{p} v_{i1}\) are independently distributed, that \(\sigma_1^{-2} v_{11} \sim \chi^2_n\), and that \(\sigma_2^{-2} \sum_{i=2}^{p} v_{i1} \sim \chi^2_{n(p-1)}\).

To find the MLE of \(\mu\) and \(\sigma^2\), first maximize with respect to \(\hat{\xi}\) in (2.5).

**Case I.** For \(\hat{\xi} = \sqrt{N} \mu \Gamma\) subject to the restrictions of the parameter space \(\Omega_{\nu}^r\) (i.e., \(\xi\) unrestricted), the MLE of \(\hat{\xi}\) is \(\hat{\xi} = y\) so that

\[
\text{(2.6a)} \quad \max_{\hat{\xi}} p(y, V) = \sigma_1^{-N} \sigma_2^{-N(p-1)} \exp \left[ -\frac{1}{2} (\sigma_1^{-2} v_{11} + \sigma_2^{-2} \sum_{i=2}^{p} v_{i1}) \right].
\]

**Case II.** For \(\hat{\xi}\) subject to the restrictions of \(\Omega_{\nu}^r\), we have that \(\hat{\xi} = (\xi_1, 0, 0, \ldots, 0)\), \(-\infty < \xi_1 < \infty\). Thus \(\hat{\xi}_1 = y_1\), and

\[
\text{(2.6b)} \quad \max_{\xi=(\xi_1, 0, 0, \ldots, 0)} \quad \text{where } -\infty < \xi_1 < \infty.
\]

**Case III.** We can also consider the case \(\Omega_{\nu}^r\) when \(\mu\) is assumed to be zero. Then

\[
\text{(2.6c)} \quad \max_{\hat{\xi}} p(y, V) = \sigma_1^{-N} \sigma_2^{-N(p-1)} \exp \left[ -\frac{1}{2} (\sigma_1^{-2} v_{11} + \sigma_2^{-2} \sum_{i=2}^{p} (v_{i1} + y_1^2)) \right].
\]

Expressions (2.6a), (2.6b), and (2.6c) have a common form, namely:

\[
\text{(2.7)} \quad p(u, w) = \max_{\hat{\xi}} p(y, V) = c \sigma_1^{-N} \sigma_2^{-N(p-1)} \exp \left[ -\frac{1}{2} (\sigma_1^{-2} u + \sigma_2^{-2} w) \right],
\]

where \(c\) is a normalizing constant and where \(u\) and \(w\) are given by:

**Case I.** \(u = v_{11} \sim \chi^2_{1} \), \(v_{11} \sim \chi^2_{n(p-1)}\),

**Case II.** \(u = v_{11} \sim \chi^2_{1} \), \(v_{11} \sim \chi^2_{n(p-1)}\),

**Case III.** \(u = v_{11} + y_1^2 \sim \chi^2_{1} \), \(v_{11} \sim \chi^2_{n(p-1)}\).
In all three cases, $u$ and $w$ are independently distributed.

Setting $\phi = \sigma_1^{-2}$, $\theta = \sigma_2^{-2}$, we have:

\[
\phi^2 = \frac{\sigma_1^2 + (p-1)\sigma_2^2}{p} = \frac{\Phi + (p-1)\theta}{p\Phi},
\]

\[
\phi = \frac{\sigma_1 - \sigma_2}{\sigma_1^2 + (p-1)\sigma_2^2} = \frac{\Phi - \theta}{\Phi + (p-1)\theta}.
\]

The condition $\rho \geq \rho_0$ then becomes

\[
8 = \phi - \theta\gamma_0 \geq 0, \quad \gamma_0 = (1-\rho_0)^{-1}(1 + (p-1)\rho_0).
\]

Reparametrizing (2.7) in terms of $\theta$ and $\delta$ and taking logarithms, we obtain

\[
2 \log p(u, w) = N \log \theta + N(p-1) \log(\delta + \theta\gamma_0) - \theta(u + w\gamma_0) - w\delta.
\]

This function is strictly concave in $\theta$ and $\delta$, so that its maximum is achieved either in the interior of the positive orthant $\theta > 0$, $\delta > 0$, or on the boundary $\theta > 0$, $\delta = 0$. Differentiating (2.11) with respect to $\theta$ and $\delta$, we find that $p(u, w)$ is maximized for

\[
\hat{\theta} = \frac{N}{u}, \quad \hat{\delta} = \frac{N}{u\gamma_0} [(p-1)u - \gamma_0 w],
\]

provided that $(p-1)u - \gamma_0 w > 0$. Otherwise, the maximum is achieved when $\delta = 0$, in which case

\[
\hat{\theta} = \frac{pN}{u + w\gamma_0}, \quad \hat{\delta} = 0.
\]

Summarizing, we have
\[(2.12) \quad \sigma_1^2 = \begin{cases} \frac{u}{N}, & \text{if } \frac{u}{w} > \frac{\gamma_0}{p-1} , \\ \frac{u + w \gamma_0}{pN}, & \text{if } \frac{u}{w} \leq \frac{\gamma_0}{p-1} , \end{cases} \]

\[(2.13) \quad \sigma_2^2 = \begin{cases} \frac{w}{(p-1)N}, & \text{if } \frac{u}{w} > \frac{\gamma_0}{p-1} , \\ \frac{u + w \gamma_0}{pN}, & \text{if } \frac{u}{w} \leq \frac{\gamma_0}{p-1} , \end{cases} \]

or in terms of \( \tilde{\sigma}^2 \) and \( \hat{\rho} \),

\[(2.14) \quad \tilde{\sigma}^2 = \begin{cases} \frac{u}{N}, & \text{if } \frac{u}{w} > \frac{\gamma_0}{p-1} , \\ \frac{(u+w \gamma_0)(\gamma_0+\rho-1)}{p^2 p_0}, & \text{if } \frac{u}{w} \leq \frac{\gamma_0}{p-1} . \end{cases} \]

\[(2.15) \quad \hat{\rho} = \begin{cases} \frac{(p-1)u - w}{(p-1)(uw)}, & \text{if } \frac{u}{w} > \frac{\gamma_0}{p-1} , \\ \rho_0, & \text{if } \frac{u}{w} \leq \frac{\gamma_0}{p-1} . \end{cases} \]

The expressions for \( \tilde{\sigma}^2 \) and \( \hat{\rho} \) in terms of \( \bar{x} \) and \( S \) can be easily worked out for each case in (2.7) from (2.14), (2.15), and the relations:

\[ y_1^2 = Np^{-1}((x_1 \bar{x})^2), \quad \sum_{i=2}^p y_i^2 = N(\bar{x}^2 \bar{x}) - p^{-1}(x_1 \bar{x})^2, \]

\[ v_{11} = p^{-1}eSe^t, \quad \sum_{i=2}^p v_{1i} = \text{tr } S - p^{-1}eSe^t. \]

2.3 A Distributional Representation for the MLE. We now make use of the fact that if two random variables \( s_1 \) and \( s_2 \) are independently distributed, each having a chi-square distribution, then \( s_1 + s_2 \) is independent of \( s_1/(s_1+s_2) \). Thus if we let

\[(2.16) \quad s = \frac{u}{\sigma_1^2} + \frac{w}{\sigma_2^2}, \quad r = \frac{\bar{x}}{\sigma_2^2} / \left( \frac{u}{\sigma_1^2} + \frac{w}{\sigma_2^2} \right), \]

...
it follows that \( r \) and \( s \) are independent with \( s \sim \chi^2_{\nu_1 + \nu_2} \) and \( r \sim \text{B} \left( \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right) \), where \((\nu_1, \nu_2)\) is equal to \((n, (p-1)n)\) in Case I, \((n, (p-1)N)\) in Case II, and \((N, (p-1)N)\) in Case III.

The above transformation permits us to represent \( \hat{\rho} \) as a function of \( r \) alone, and to represent \( \hat{\sigma}^2 \) as a function of \( r \) multiplied by a function of \( s \), so that we can take advantage of independence. Thus, we have

\[
(2.17) \quad \hat{\rho} = \text{max} \left\{ \frac{(p-1)\gamma(1-r) - r}{(p-1)[r + \gamma(1-r)]}, \frac{\rho}{\rho_0} \right\},
\]

where \( \gamma = \frac{1 + (p-1)\rho}{1 - \rho} = \sigma^2 / \sigma_1^2 \). Also,

\[
(2.18) \quad \hat{\sigma}^2 = \sigma^2 \frac{s}{Np} h(r),
\]

where

\[
(2.19) \quad h(r) = \left\{ \begin{array}{ll}
1 + \rho(1-r) - 1, & \text{if } 0 \leq r \leq \lambda, \\
\left[ \frac{(1 + (p-1)\rho)(1-r) + (1-\rho)\rho_0^2}{\rho_0^2} \right] (\gamma_0 + p-1), & \text{if } \lambda < r \leq 1,
\end{array} \right.
\]

and \( \lambda = \gamma(p-1)/[\gamma(p-1) + \gamma_0] \). Note that although \( r \) and \( s \) are functions of \( \sigma^2 \) and \( \rho \), the RHS of (2.17) and (2.18) are independent of these parameters, so that \( \hat{\rho} \) and \( \hat{\sigma}^2 \) are valid estimators.

2.4 Remarks on Expected Values. Using (2.18), we have

\[
E \hat{\sigma}^2 = \sigma^2 \frac{s}{Np} E[h(r)]
\]

\[
= \sigma^2 \frac{\nu_1 + \nu_2}{Np} E[h(r)],
\]

\[
E \hat{\sigma}^4 = \sigma^4 \frac{\nu_1 + \nu_2 + 2}{N^2 p^2} E[h(r)]^2.
\]

The main complication arises from the term containing \( h(r) \). By a direct computation we have
\[ E[h(r)] = \frac{(p-1)\rho_0 (1+(p-1)\rho)}{1+(p-1)\rho_0} \left( I_{\lambda, \frac{v_2}{2}, \frac{v_1}{2}} - \frac{v_2}{(v_1+v_2)} \frac{p(1+\rho_0 - (p\rho - 1))}{1-\rho_0} I_{\lambda, \frac{v_2+2}{2}, \frac{v_1}{2}} \right) \\
+ (p-1+\gamma_0) \left( \frac{1+(p-1)\rho}{p\rho_0} + \frac{v_2}{v_1+v_2} [\gamma_0 - 1 - (p-1+\gamma_0)\rho] \right), \]

for \( I_{\lambda(a,b)} = \left[ B(a,b) \right]^{-1} \int_0^\lambda y^{a-1} (1-y)^{b-1} dy. \) The computation of \( E[h(r)]^2 \)
is straightforward, though tedious, and results in a linear combination of
\( I_{\lambda, \frac{v_2}{2}, \frac{v_1}{2}}, I_{\lambda, \frac{v_2+2}{2}, \frac{v_1}{2}}, \) and \( I_{\lambda, \frac{v_2+2}{2}, \frac{v_1}{2}}. \)

A direct computation yields \( \lambda \geq (p-1)/\rho, \) with equality if and
only if \( \rho = \rho_0. \) Furthermore, as \( N \to \infty, \) \( r \) converges to \((p-1)/\rho\) with
probability one. Consequently, \( I_{\lambda, \frac{v_2}{2}, \frac{v_1}{2}} = P(0 \leq r \leq \lambda) \) and \( I_{\lambda, \frac{v_2+2}{2}, \frac{v_1}{2}} \)
converge to one when \( \rho > \rho_0. \) A more refined result (see Appendix) can
actually be obtained when \( \rho > \rho_0, \) namely,

\[ 1 - I_{\lambda, \frac{v_2}{2}, \frac{v_1}{2}} \leq c_1 \sqrt{N} e^{-\tau N}(1 + O(N^{-1})), \]
\[ 1 - I_{\lambda, \frac{v_2+2}{2}, \frac{v_1}{2}} \leq c_2 \sqrt{N} e^{-\tau N}(1 + O(N^{-1})), \]

where \( \tau > 0. \) It follows that

\[ E[h(r)] = 1 + (p-1)\rho - p\rho v_2 (v_1 + v_2)^{-1} + o(e^{-\tau N}), \]
or

\[ E[h(r)] = 1 + o(e^{-\tau N}). \]

Hence

\[ E[\hat{\sigma}^2] = \sigma^2 + O(N^{-1}). \]

From (2.17) and (2.20),

\[ E[\tilde{\rho}] \leq E[\hat{\rho}] \leq E[\tilde{\rho}] + (\rho_0 + \frac{1}{p-1}) O(e^{-\tau N}), \]

where \( \tilde{\rho} \) is the unrestricted MLE of \( \rho, \) i.e.,
(2.23) $\hat{\rho} = \frac{(p-1)y - r[(p-1)y + 1]}{(y-1)[y + (1-7)r]}$.

The result $E \tilde{\rho} = \rho + O(N^{-1})$ can be directly computed from the representation (2.23). Therefore,

(2.24) $E \hat{\rho} = \rho + O(N^{-1})$.

2.3 Asymptotic Distribution of the MLE. For $\rho > \rho_0$, we have seen that $\lim (\hat{\rho} - \rho) = 0$. Indeed, the rate of convergence is very rapid as can be seen from (2.20). The limiting distribution of $\hat{\rho}$ is known to be (viz., Olkin and Pratt (1958)),

$$\lim_{N \to \infty} z(\hat{\rho}) = \mathcal{N}(0, \omega)$$

where

$$\omega = \frac{2(1-\rho)^2 (1+(p-1)\rho)^2}{p(p-1)}$$

Hence, for $\rho > \rho_0$,

(2.25) $\lim_{N \to \infty} z(\hat{\rho}) = \mathcal{N}(0, \omega)$

For $\rho = \rho_0$, we still obtain asymptotic normality, but now the distribution is censored, i.e.,

(2.26) $\lim_{N \to \infty} z(\hat{\rho}) = \mathcal{N}^+(0, \omega)$

where $\mathcal{N}^+(0, \omega)$ represents the normal distribution censored at 0.

When $\rho > \rho_0$, $P(\lambda \leq r \leq 1)$ rapidly converges to 0 as $N \to \infty$.

Thus

$$\lim_{N \to \infty} P(\hat{\rho}^2 < \frac{\tilde{\rho}^2 - \lambda}{\tilde{\rho}^2}) = \lim_{N \to \infty} P(\lambda \leq r \leq 1) = 0$$

It follows that

(2.27) $\lim_{N \to \infty} z(\hat{\rho}^2) = \lim_{N \to \infty} z(\hat{\rho}^2) = \mathcal{N}(0, \frac{2\tilde{\rho}}{p}[1 + (p-1)\tilde{\rho}^2])$.

For $\rho = \rho_0$, the representation (2.18) of $\hat{\rho}^2$ in terms of $s^2$ and $r$ becomes
\[ \hat{\sigma}^2 = \frac{\sigma^2}{Np} \max[1, 1 + (p-1)r_0^0 - p^0 r] = R(s) g(r). \]

Now \[ \lim_{N \to \infty} \mathcal{N}(\max \{s, \max \{0, s\}\}) = \mathcal{N}(0, 2p^{-1} \sigma^4) \]

while \[ \lim_{N \to \infty} \mathcal{N}(\max \{s, \max \{0, s\}\}) = \mathcal{N}(\max \{0, s\} \sigma^2) \]

where \( z \sim \mathcal{N}(0, p^{-1} (p-1) \sigma^2) \). Using the Delta method, we conclude that the asymptotic distribution of \( \sqrt{N} (\hat{\sigma} - \sigma) \) as \( N \to \infty \) is the same as the distribution of \( z^* + \sigma^2 \max[0, \max] \), where \( z^* \) and \( z \) are independently distributed, \( z \sim \mathcal{N}(0, 2p^{-1} \sigma^4) \), and \( z^* \sim \mathcal{N}(0, 2p^{-1} \sigma^4) \).
3. Tests of Hypotheses for $\mathcal{H}_n^{r}$ and $\mathcal{H}_v^{r}$

3.1 Derivation of the Likelihood Ratio Tests. We are interested in testing $\mathcal{H}_n^{r}: (\mu, \Sigma) \in \Omega_n^{r}$ and $\mathcal{H}_v^{r}: (\mu, \Sigma) \in \Omega_v^{r}$ against general alternatives $\mathcal{H}: (\mu, \Sigma) \in \Omega$, where $\Omega = \{ (\mu, \Sigma): \mu$ unrestricted, $\Sigma > 0 \}$. To obtain likelihood ratio tests for these hypotheses, we first obtain the supremum of the likelihood function (2.5) over each of the regions $\Omega_n^{r}$, $\Omega_v^{r}$, and $\Omega$. From (2.7), (2.12), (2.13), we find that:

\[
\sup_p(y, V) = \begin{cases} 
\exp\left(-\frac{1}{2} pN\right) \frac{1}{(p-1)N}, & \text{if } \frac{v_{11}}{p \Sigma} \sum_{i=2}^{p} (v_{ii} + y_i^2) \leq \frac{\gamma_0}{p - 1}, \\
\frac{1}{N} \left[ \frac{v_{11}}{p \Sigma} \sum_{i=2}^{p} (v_{ii} + y_i^2) \right]^{(p-1)N}, & \text{if } \frac{v_{11}}{p \Sigma} \sum_{i=2}^{p} (v_{ii} + y_i^2) > \gamma_0.
\end{cases}
\]

(3.1)

\[
\sup_p(y, V) = \begin{cases} 
\exp\left(-\frac{1}{2} pN\right) \frac{1}{(p-1)N}, & \text{if } \frac{v_{11}}{p \Sigma} \sum_{i=2}^{p} v_{ii} \leq \frac{\gamma_0}{p - 1}, \\
\frac{1}{pN} \left[ \frac{v_{11}}{p \Sigma} \sum_{i=2}^{p} v_{ii} \right]^{(p-1)N}, & \text{if } \frac{v_{11}}{p \Sigma} \sum_{i=2}^{p} v_{ii} > \gamma_0.
\end{cases}
\]

(3.2)
Finally, it is well known that

\[
\sup_{\mathbf{y}} \mathbf{p}(\mathbf{y}, \mathbf{v}) = \frac{\frac{p}{N} \exp\left\{ -\frac{1}{2} \mathbf{y}^T \mathbf{M} \mathbf{y} \right\}}{\left| \mathbf{v} \right|^2}.
\]

We find the likelihood ratio test statistic \( \lambda_1 \) for testing \( H_{mvc}^T \) against general alternatives by forming the ratio of (3.4) with (3.5). The result can be expressed in the following form:

\[
\lambda_1 = \lambda_{mvc} \lambda_0|_{mvc},
\]

where \( \lambda_{mvc} \) is Wilks' (1946) likelihood ratio test statistic for testing \( H_{mvc}^T \) against general alternatives, that is,

\[
\lambda_{mvc}^{-2/N} = \left\{ \begin{array}{l}
\frac{\mathbf{v}_1^T \mathbf{1}}{\mathbf{v}_1^T \mathbf{1}} \left[ \frac{p}{\Sigma} \left( \mathbf{v}_i^T \mathbf{y}_i^2 \right) \right]^{p-1}, \quad \text{if } \frac{\mathbf{v}_1^T \mathbf{1}}{\mathbf{v}_1^T \mathbf{1}} > \frac{\gamma_0}{p-1} \\
\end{array} \right.
\]

and where \( \lambda_0|_{mvc} \) is the likelihood ratio test of \( \rho > \rho_0 \) versus \( \rho = \rho_0 \) given that \( H_{mvc}^T \) is true, i.e.,

\[
\lambda_0|_{mvc}^{-2/N} = \left\{ \begin{array}{l}
\frac{\mathbf{v}_1^T \mathbf{1} + \gamma_0 \Sigma (\mathbf{v}_i^T \mathbf{y}_i^2)}{\mathbf{v}_1^T \mathbf{1}} \left[ \frac{p}{\Sigma} \left( \mathbf{v}_i^T \mathbf{y}_i^2 \right) \right]^{p-1}, \quad \text{if } \frac{\mathbf{v}_1^T \mathbf{1}}{\mathbf{v}_1^T \mathbf{1}} > \frac{\gamma_0}{p-1} \\
\end{array} \right.
\]
Similarly, we can obtain the likelihood ratio test statistic $\lambda_2$ for testing $H^T_{vc}$ against general alternatives by dividing (3.2) by (3.3). We find that

\begin{equation}
\lambda_2 = \lambda_{vc} \lambda_0|vc,
\end{equation}

where $\lambda_{vc}$ is Wilks' (1946) likelihood ratio test statistic for testing $H_{vc}$ against general alternatives, that is,

\begin{equation}
\lambda_{vc}^{-2/N} = \frac{|V|}{\left(\frac{1}{p-1} \Sigma_{i=2}^{p} v_{ii} \right)^{p-1} C_{11}},
\end{equation}

and where $\lambda_0|vc$ is the likelihood ratio test statistic for testing $\rho > \rho_0$ against $\rho = \rho_0$ given that $H_{vc}$ is true, namely,

\begin{equation}
\lambda_0|vc^{-2/N} = \begin{cases} 
1, & \text{if } \frac{v_{11}}{p \Sigma_{i=2}^{p} v_{ii}} > \frac{\gamma_0}{p-1}, \\
\left[\frac{v_{11}^\gamma_0 + \Sigma_{i=2}^{p} v_{ii}}{p-1} \left(\frac{\gamma_0}{p \Sigma_{i=2}^{p} v_{ii}}\right)^{p-1} \right], & \text{if } \frac{v_{11}}{p \Sigma_{i=2}^{p} v_{ii}} < \frac{\gamma_0}{p-1}.
\end{cases}
\end{equation}

Remark. The likelihood ratio test statistic $\lambda_2$ of $H^T_{mvc}$ versus $H^T_{vc}$ may be of interest. This statistic can be obtained as the ratio of $\lambda_1$ to $\lambda_2$, that is, $\lambda_2 = \lambda_1/\lambda_2$. From (3.4) and (3.7), we see that $\lambda_2 = \lambda_{m}^{0} \lambda_{mvc}^{0}|vc$, where $\lambda_{m}$ is Wilks' (1946) likelihood ratio test statistic for testing $H_{mvc}$ when $H_{vc}$ is known to be true. The ratio $\lambda_0|mvc|vc^{-1}$ is a complicated function of the observations, and the null
distribution of \( \lambda \) is quite involved even in the asymptotic case.

3.2 Asymptotic Null Distributions of the Likelihood Ratio Test
Statistics. Since under \( H_{\text{mvc}} \), the triple \((y_{11}^2 + y_{12}^2, y_{12}^2 + y_{13}^2)\) constitutes a complete and sufficient statistic for the parameter \((\theta_1, \theta_2)\) (Barbash (1959)), and since the distribution of \( \lambda_{\text{mvc}} \) is independent of \((\theta_1, \theta_2)\) while the distribution of \( \lambda_{0|\text{mvc}} \) depends upon these parameters, the statistics \( \lambda_{\text{mvc}} \) and \( \lambda_{0|\text{mvc}} \) are independent when \( H_{\text{mvc}} \) is true (Basu (1953)). The distribution of \( \lambda_{\text{mvc}} \) under \( H_{\text{mvc}} \) is known (viz., Tukey and Wilks (1946)) to be a product of independent Beta variables. Wilks (1946) has shown that (under \( H_{\text{mvc}} \)),

\[
(3.10) \quad \lim_{N \to \infty} z(-2 \log \lambda_{\text{mvc}}) = \frac{\chi^2}{(p+3)(p-2)}.
\]

The distribution of \( \lambda_{0|\text{mvc}} \) is more complicated. \( \lambda_{0|\text{mvc}} \) is equivalent to a one sided \( F_{n,N(p-1)} \) test based on

\[
\frac{\sum_{i=1}^{p} (y_{i1}^2 + y_{i2}^2)}{n \sum_{i=1}^{p} (y_{i1}^2 + y_{i2}^2)}.
\]

When \( \rho > \rho_0 \),

\[
\lim_{N \to \infty} \mathbb{P} \left( \frac{\sum_{i=1}^{p} (y_{i1}^2 + y_{i2}^2)}{n \sum_{i=1}^{p} (y_{i1}^2 + y_{i2}^2)} > \frac{\gamma_0}{p-1} \right) = 1,
\]

so that \( z(-2 \log \lambda_{0|\text{mvc}}) \) approaches a distribution degenerate at 0.

Thus for \( \rho > \rho_0 \),

\[
(3.11) \quad \lim_{N \to \infty} z(-2 \log \lambda_1) = \lim_{N \to \infty} z(-2 \log \lambda_{\text{mvc}}) = \frac{\chi^2}{(p+3)(p-2)}.
\]

When \( \rho = \rho_0 \),

\[
\lim_{N \to \infty} \mathbb{P} \left( \frac{\sum_{i=1}^{p} (y_{i1}^2 + y_{i2}^2)}{n \sum_{i=1}^{p} (y_{i1}^2 + y_{i2}^2)} > \frac{\gamma_0}{p-1} \right) = \frac{1}{2},
\]

and it can be shown that...
\[
\lim_{N \to \infty} x(-2 \log \lambda_0|_{\text{mvc}}) \Rightarrow \begin{cases} 
0, & \text{with probability } \frac{1}{2}, \\
\chi^2_1, & \text{with probability } \frac{1}{2}.
\end{cases}
\]

Therefore, when \( \rho = \rho_0 \),

\[(3.19) \quad \lim_{N \to \infty} x(-2 \log \lambda_1) \Rightarrow \begin{cases} 
\chi^2_{(p^2-1)/2}(p-2), & \text{with probability } \frac{1}{3}, \\
\chi^2_{(p+3)(p-2)} + 1, & \text{with probability } \frac{1}{2}.
\end{cases}
\]

Since the limiting law (3.12) is stochastically larger than the limiting law (3.11), it is recommended that the rejection region be chosen with reference to the limiting law (3.12). Indeed, under \( H_{\text{mvc}} \), the distribution of \( \lambda_1 \) has a monotone likelihood ratio in \( \rho \), so that even in finite samples the rejection region should be selected with reference to the distribution of \( \lambda_1 \) when \( H_{\text{mvc}} \) holds and \( \rho = \rho_0 \).

When \( \rho < \rho_0 \),

\[
\lim_{N \to \infty} P\left( \frac{v_{11}}{\sum_{i=2}^{p} (v_{ii} + y_i^2)} > \frac{\gamma_0}{p-1} \right) = 0,
\]

so that

\[
\lim_{N \to \infty} \frac{-2 \log \lambda_0|_{\text{mvc}} + N \log \left[ \frac{v_{11}}{\sum_{i=2}^{p} (v_{ii} + y_i^2)} + \frac{\gamma_0}{p-1} \right] - \frac{v_{11}}{p-1} \gamma_0}{p} = 0.
\]

From this result,

\[
\lim_{N \to \infty} -2 \log \lambda_1 = \infty.
\]
Similar results hold for $\lambda_2$. Under $H_{vc}$, $-2 \log \lambda_2$ is the convolution of the independent test statistics $-2 \log \lambda_{vc}$ and $-2 \log \lambda_{0|vc}$. The distribution of $\lambda_{vc}$ under $H_{vc}$ can be expressed as the product of independent Beta variables, while $\lambda_{0|vc}$ is equivalent to the one-sided $F_{u,u(p-1)}$-test based on $\sum_{i=1}^{p} \frac{v_{1i}}{v_{ii}}$ and has a monotone likelihood ratio in $\rho$. Consequently, the rejection region of $\lambda_2$ should be derived from the distribution for $\lambda_2$ when $H_{vc}$ holds and $\rho = \rho_0$. The asymptotic distribution of $\lambda_2$ when $H_{vc}$ holds and $\rho > \rho_0$ is

$$\lim_{N \to \infty} z(-2 \log \lambda_2) = \frac{\chi^2_{p(p-1)}}{2} - 2$$

whereas the limiting distribution of $\lambda_2$ when $H_{vc}$ holds and $\rho = \rho_0$ is

$$\lim_{N \to \infty} z(-2 \log \lambda_2) = \begin{cases} \frac{\chi^2_{p(p-1)}}{2} - 2, & \text{with probability } \frac{1}{2}, \\ \frac{\chi^2_{p(p-1)}}{2} - 1, & \text{with probability } \frac{1}{2}. \end{cases}$$

When $H_{vc}$ holds and $\rho < \rho_0$,

$$\plim_{N \to \infty} -2 \log \lambda_2 = \infty.$$
Appendix. An Inequality for the Incomplete Beta Function

In this appendix we prove the following inequality.

Theorem. If $\alpha_1 = z$, $\alpha_2 = az + b$, $a > 0$, and if $\beta < 1/(a+1)$, then as $z \to \infty$,

$$I_{\lambda}(\alpha_1, \alpha_2) = \int_0^\beta \frac{y^{\alpha_1-1}}{\Gamma(\alpha_1) \Gamma(\alpha_2)} \frac{d^\alpha_2}{(1-y)^{\alpha_2-1}} \leq C \sqrt{e} e^{-\tau z(1 + o(z^{-1}))},$$

for some $\tau > 0$.

Proof. The maximum of $y^{\alpha_1-1} (1-y)^{\alpha_2-1}$ is achieved for $y = 1/(a+1)$. Further, this function is increasing for $y \leq 1/(a+1)$, so that for $0 \leq y \leq \beta < 1/(a+1)$,

$$y^{\alpha_1-1} (1-y)^{\alpha_2-1} \leq \beta^{\alpha_1-1} (1-\beta)^{\alpha_2-1}.\]

Consequently,

$$\int_0^\beta \frac{y^{\alpha_1-1}}{\Gamma(\alpha_1) \Gamma(\alpha_2)} \frac{d^\alpha_2}{(1-y)^{\alpha_2-1}} \leq \left[ \int_0^\beta (1-y)^{b+a-1} dy \right] \beta^{\alpha_1-1} (1-\beta)^{\alpha_2-1}.\]

Furthermore, from the well-known Sterling expansion of the Gamma function,

$$\Gamma(\alpha_1) \Gamma(\alpha_2) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1) \Gamma(\alpha_2)} = \frac{\sqrt{2\pi}}{e^{(\alpha_1+\alpha_2)}} \frac{1}{(\alpha_1+\alpha_2)^{\alpha_1+\alpha_2}} [1 + o(z^{-1})].$$

$$B(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1) \Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)} = \sqrt{2\pi} \frac{e^{-(\alpha_1+\alpha_2)}}{e^{(\alpha_1+\alpha_2)}} \frac{1}{(\alpha_1+\alpha_2)^{\alpha_1+\alpha_2}} [1 + o(z^{-1})]$$

$$= \sqrt{2\pi} (\alpha_1+\alpha_2)^{-\frac{1}{2}} \left( \frac{\alpha_1}{\alpha_1+\alpha_2} \right)^{\alpha_1+\alpha_2} \left( \frac{\alpha_2}{\alpha_1+\alpha_2} \right)^{\alpha_2+\alpha_2} [1 + o(z^{-1})].$$

Taking the ratio of (i) to (ii), we obtain
\[ I_{\lambda}(c_1, c_2) \leq \sqrt{2\pi} e^{-b} \left( \frac{a}{a+1} \right)^{-\frac{1}{2}} \left[ \frac{1 - (1-\beta)^{b+1}}{b+1} \right] \beta^{-1}(1-\beta)^{-a} \]

\[
\left\{ \left[ (a+1)\beta \right] \left[ a^{-1}(a+1)(1-\beta) \right]^a \right\}^2 \left[ 1 + O(z^{-\frac{1}{2}}) \right].
\]

But since \( \beta < 1/(a+1) \), the arithmetic-geometric mean inequality tells us that

\[
\frac{1}{[(a+1)\beta]^{a+1} \left[ a^{-1}(a+1)(1-\beta) \right]^{a+1}} < 1.
\]

Thus we can write

\[
\left[ (a+1)\beta \right] \left[ a^{-1}(a+1)(1-\beta) \right]^a = e^{-\tau},
\]

for some \( \tau > 0 \). This completes the proof. \( \|
\)
References


