Likelihood Ratio Test for Order Restrictions against all Alternatives in Multivariate Normal Distribution

Abouzar Bazyari, Rahim Chinipardaz and Ali Akbar Rasekhi

Department of Statistics, Shahid Chamran University, Ahvaz, Iran

Abstract

The problem of testing the isotonic of several $p$-variate normal mean vectors against all alternatives is considered. This is a multivariate extension of Robertson and Wegrman (1978). In the present paper, two cases are considered. First, it is assumed that the covariance matrices are known and second that they have an unknown scale factor. For both cases, we propose the test statistic, critical values and estimate the power of tests. The $p$-values are obtained by simulation study.

Keywords: Isotonic regression, Likelihood ratio test, Multivariate normal distribution, Simulation.

Introduction

Suppose that $X_1, X_2, \ldots, X_m$ are random variables from a $p$-variate normal distribution with mean vector $\mu_v$ and nonsingular covariance matrix $\Sigma_v$, $N_p(\mu_v, \Sigma_v)$, $v = 1, 2, \ldots, k$. Throughout this paper except in section 5, $\Sigma_v$ is assumed to be known.

Consider the likelihood ratio test for testing

$$H_1: \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k,$$

against all alternatives, where $\mu_v \leq \mu_j$ means that all the elements of $\mu_j - \mu_v$ are non-negative.

Let $H_2$ be the hypothesis placing no restriction on $\mu$'s. Then it is desired to test the hypothesis $H_1$ against the alternative hypothesis $H_2 - H_1$.

Such tests may be used in some fields. Their applications can be found in clinical trails design to test superiority of a combination therapy (Laska and Meisner, 1989
and Sarka et al., 1995). More examples may be found in the case of ordered treatment means or the testing in which a treatment is better than control when the responses are ordinal.

The extensive literature concerning this problem has been appeared by Barlow et al. (1972), Robertson et al. (1988) and Silvapulle and Sen (2005). The last reference includes some applicable examples as well as some new methods considered by authors in this area.

Such tests started by Bartholomew (1959a, b) who derived a likelihood ratio test for homogeneity of \( k \) univariate normal means against ordered alternatives. Chacko (1963), Shorack (1967) and Barlow et al. (1972) followed the problems with ordered parameters. Kudo (1963) considered a \( p \)-dimensional normal distribution with unknown mean \( \mu = (\mu_1, \mu_2, \ldots, \mu_p) \) and known covariance matrix \( \Sigma \). The problem of testing was \( H_0 : \mu = 0 \) against the restricted alternative \( H_1 : \mu_i \geq 0 (i = 1, \ldots, p) \), where the inequality is strict for at least one value of \( i \). He obtained a statistic based on the likelihood ratio criterion and discussed its existence and geometric nature and also gave a scheme for its computation.

Perlman (1969) studied this problem assuming that \( \Sigma \) is completely unknown. Kudo and Choi (1975) generalized this work where the alternative hypothesis is that the mean vector lies in a convex polyhedral cone. Sasabuchi (1980) considered the problem that the mean vector lies on the boundary of a convex polyhedral cone against the mean vector corresponds in the interior. It was a complete generalization of Inada (1978) who studied it for the bivariate case (see also Sasabuchi, 1988a, b).

The problem of testing the homogeneity of several multivariate normal means against the unrestricted alternative hypothesis was given by Anderson (1984). Sasabuchi et al. (1983) extended Bartholomew's (1959a) problem to multivariate normal mean vectors with known covariance matrices. They obtained the likelihood ratio test statistic and proposed an iterative algorithm for computing the bivariate isotonic regression. Kulatunga and Sasabuchi (1984) studied its null distribution only in some special situations. Kulatunga et al. (1990) proposed some test procedures when the covariance matrices are not diagonal, and studied them by simulation. Sasabuchi et al. (1992) generalized the iterative algorithm to multivariate isotonic regression.

Sasabuchi et al. (2003) considered this problem of testing in case that the covariance matrices are common but unknown. He proposed a test statistic, studied its upper tail probability under the null hypothesis and estimated its critical values. Sasabuchi (2007) then gave some more powerful tests than Sasabuchi et al. (2003). In all cases, the authors examined the hypothesis

\[
H_0 : \mu_1 = \mu_2 = \cdots = \mu_k
\]  

which seems to be easier than given in (1).

In fact, the hypotheses given in this paper are the multivariate versions of those given by Robertson and Wegman (1978) who obtained the likelihood ratio test statistic for testing the isotonic of several univariate normal means against all
Likelihood Ratio Test for Order Restrictions

alternative hypotheses. We are interested in a multivariate extension of Robertson and Wegman's (1978) problem in both cases when the covariance matrices are known and when they have an unknown scale factor. We obtain the test statistic based on the likelihood ratio method, the critical values and power of test in both cases. Our results include the results given by Robertson and Wegman (1978) as a special case.

The paper is organized as follows. In section 2, we review the results given by Barlow et al. (1972) and Sasabuchi et al. (1983). We describe the problem, compute the likelihood ratio test when the covariance matrices are known and study its null distribution.

In section 3, for bivariate and trivariate normal distribution, the critical values are given. The power of test for different values of significance levels and its exact p-value are estimated by simulation in section 4. In section 5, we obtain the likelihood ratio test when the covariance matrices have an unknown scale factor.

Review the results and likelihood ratio test
Suppose that \( X_{v1}, X_{v2}, \ldots, X_{vk} \) are random variables from a \( p \)-variate normal distribution with mean vector \( \mu_v \) and nonsingular covariance matrix \( \Sigma_v, N_p(\mu_v, \Sigma_v) \), \( v = 1,2,\ldots,k \). Throughout this paper, except in section 5, \( \Sigma_v \) is assumed to be known.

**Definition (Sasabuchi et al., 1983).** Given \( p \)-variate real vectors \( X_1, X_2, \ldots, X_k \) and \( p \times p \) positive definite matrices \( \Sigma_1, \Sigma_2, \ldots, \Sigma_k \), a \( p \times k \) real matrix \( (\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k) \) is said to be the multivariate isotonic regression (MIR) of \( X_1, X_2, \ldots, X_k \) with weights \( \Sigma_1^{-1}, \Sigma_2^{-1}, \ldots, \Sigma_k^{-1} \) if \( (\hat{\mu}_1, \leq \hat{\mu}_2, \leq \cdots, \leq \hat{\mu}_k) \) and \( (\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k) \) satisfies
\[
\min_{\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k} \sum_{v=1}^{k} (X_v - \mu_v)’ \Sigma_v^{-1} (X_v - \mu_v) = \sum_{v=1}^{k} (X_v - \hat{\mu}_v)’ \Sigma_v^{-1} (X_v - \hat{\mu}_v),
\]
where \( \hat{\mu}_v \)'s can be computed with iterative algorithm proposed by Sasabuchi et al. (1983).

In fact, this definition includes the definition given by Barlow et al. (1972) for univariate variables.

Now, consider the problem of testing
\[
H_1: \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k \text{ versus } H_2 - H_1,
\]
where \( \mu_v \leq \mu_j \) means that all the elements of \( \mu_j - \mu_v \) are non-negative.

The likelihood ratio test for testing \( H_1 \) against \( H_2 - H_1 \) is based on
A likelihood ratio test rejects $H_1$ for small values of $\lambda$.

**Theorem 1.** Suppose that $\Sigma_v$ is known. The likelihood ratio test for $H_1$ against $H_2 - H_1$ rejects $H_1$ for the large values of

$$T = -2 \ln \lambda = \sum_{v=1}^{k} n_v (\hat{\mu}_v - \overline{X}_v)\Sigma_v^{-1}(\hat{\mu}_v - \overline{X}_v),$$

where $\overline{X}_v$ is the maximum likelihood estimate of $\mu_v$ under the alternative hypothesis and $(\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k)$ is the multivariate isotonic regression (MIR) of $\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_k$ with weights $n_1\Sigma_1^{-1}, n_2\Sigma_2^{-1}, \ldots, n_k\Sigma_k^{-1}$ and $\overline{X}_v = \frac{1}{n_v} \sum_{j=1}^{n_v} X_{vj}$, $v = 1, 2, \ldots, k$.

**Proof.** With some easy manipulations we have

$$\lambda = \exp\{-1/2 \sum_{v=1}^{k} n_v (X_{vj} - \hat{\mu}_v)'\Sigma_v^{-1}(X_{vj} - \hat{\mu}_v) - \sum_{v=1}^{k} \sum_{j=1}^{n_v} (X_{vj} - \overline{X}_v)'\Sigma_v^{-1}(X_{vj} - \overline{X}_v)\},$$

where $\overline{X}_v$ and $\mu_v$ are as before.

$$-2 \ln \lambda = \left[\sum_{v=1}^{k} n_v (\hat{\mu}_v - \overline{X}_v)'\Sigma_v^{-1}(\hat{\mu}_v - \overline{X}_v) - \sum_{v=1}^{k} \sum_{j=1}^{n_v} (X_{vj} - \overline{X}_v)'\Sigma_v^{-1}(X_{vj} - \overline{X}_v)\right]$$

$$= \sum_{v=1}^{k} \left[ n_v (\hat{\mu}_v - \overline{X}_v)'\Sigma_v^{-1}(\hat{\mu}_v - \overline{X}_v) - 2 \sum_{j=1}^{n_v} (X_{vj} - \hat{\mu}_v)'\Sigma_v^{-1}(\hat{\mu}_v - \overline{X}_v)\right].$$

Simplifying the second term
Likelihood Ratio Test for Order Restrictions

\[
2 \sum_{j=1}^{n} (X_{ij} - \hat{\mu}_v) \Sigma_v^{-1} (\hat{\mu}_v - \bar{X}_v) = 2 \sum_{j=1}^{n} X_{ij}' \Sigma_v^{-1} \hat{\mu}_v - \sum_{j=1}^{n} X_{ij}' \Sigma_v^{-1} \bar{X}_v - \sum_{j=1}^{n} \hat{\mu}_v \Sigma_v^{-1} \hat{\mu}_v + \sum_{j=1}^{n} \hat{\mu}_v \Sigma_v^{-1} \bar{X}_v
\]

\[
= 2 n_v (\hat{\mu}_v - \bar{X}_v)' \Sigma_v^{-1} (\hat{\mu}_v - \bar{X}_v).
\]

and hence we have

\[
-2 \ln \lambda = \sum_{v=1}^{k} n_v (\hat{\mu}_v - \bar{X}_v)' \Sigma_v^{-1} (\hat{\mu}_v - \bar{X}_v).
\]

The distribution of the \( T \) statistic

Without the loss of generality, for the simplicity, we suppose that \( n_v = 1 \) for \( v = 1, 2, \ldots, k \), then the likelihood ratio test is

\[
T = -2 \ln \lambda = \sum_{v=1}^{k} (\hat{\mu}_v - X_v)' \Sigma_v^{-1} (\hat{\mu}_v - X_v),
\]

where \( X_v \) is a column vector for \( v \)th population.

Consider another hypothesis:

\[
H_0: \mu_1 = \mu_2 = \cdots = \mu_k.
\]

Then it is clear that \( H_0 \subset H_1 \subset H_2 \). In fact, \( H_0 \) is the least favorable among hypotheses satisfying \( H_1 \) with the largest type I error probability.

Suppose that \( \Sigma_v \) is diagonal with diagonal elements \( \sigma_{v1}, \sigma_{v2}, \ldots, \sigma_{vk} \), \( v = 1, 2, \ldots, k \). Then the likelihood ratio test statistic leads to

\[
T = -2 \ln \lambda = \sum_{i=1}^{p} \sum_{v=1}^{k} \frac{1}{\sigma_{vi}} (\hat{\mu}_v - X_{vi})^2,
\]

where \( \hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k \) is the univariate isotonic regression of \( X_{i1}, X_{i2}, \ldots, X_{ik} \) with weights \( \sigma_{i1}^{-1}, \sigma_{i2}^{-1}, \ldots, \sigma_{ik}^{-1} \).

The distribution of \( T \) when \( \Sigma_v \) is diagonal

The \( T \) statistic can be written as

\[
T = -2 \ln \lambda = \sum_{i=1}^{p} (T_{i1})^2 - \sum_{i=1}^{p} \sum_{v=1}^{k} \frac{1}{\sigma_{vi}} (\hat{\mu}_v - X_{vi})^2.
\]
The null distribution of $T$ with this form is given by the following theorem.

**Theorem 2.** If $H_0$ be true, then for $t > 0$, the null distribution of $T$ is

$$P(T \geq t) = \sum_{l_p, l_{p-1}, \ldots, l_1} P(L_1 = l_1, \ldots, L_p = l_p) \cdot P(\chi^2_{(k-1)} \geq t)$$

where $l = \sum_{i=1}^{p} l_i$.

**Proof.** Robertson and Wegman (1978) showed that when $H_0$ is true the distribution of $(T_{12})_i$ for $i = 1, 2, \ldots, p$ is given as

$$P((T_{12})_i \geq t) = \sum_{i=1}^{k-1} P(l_i, k) P(\chi^2_{(k-l_i)} \geq t)$$

$$P((T_{12})_i = 0) = P(k, k),$$

where $P(l_i, k)$ is the probability that the isotonic regression function $\hat{\mu}_{vi}$ takes exactly $l_i$ distinct values and we note that

$$P(1, k) = \frac{1}{k}, \quad P(k, k) = \frac{1}{k!}$$

$$P(l_i, k) = \frac{1}{k} P(l_i - 1, k - 1) + \frac{k - 1}{k} P(l_i, k - 1).$$

We can write

$$P((T_{12})_i \geq t \mid l_i) = P(\chi^2_{(k-l_i)} \geq t).$$

Let $T \neq 0$.

$$P(T \geq t) = P(\sum_{i=1}^{p} (T_{12})_i \geq t)$$

$$= \sum_{l_1, \ldots, l_p} P(L_1 = l_1, \ldots, L_p = l_p) \cdot P(\sum_{i=1}^{p} (T_{12})_i \geq t \mid L_1 = l_1, \ldots, L_p = l_p)$$

$$= \sum_{l_p} \sum_{l_{p-1}, \ldots, l_1} P(L_1 = l_1, \ldots, L_p = l_p) \cdot P(\sum_{i=1}^{p} (T_{12})_i \geq t \mid L_1 = l_1, \ldots, L_p = l_p).$$
Likelihood Ratio Test for Order Restrictions

Since \((T_{12})_i\), given \(l_i\), has \(\chi^2\) distribution with \(k - l_i\) degrees of freedom, so given \(L_1 = l_1, \ldots, L_p = l_p\), \(\sum_{i=1}^{p} (T_{12})_i\) is distributed as \(\chi^2\) with \((kp - \sum_{i=1}^{p} l_i)\) degrees of freedom. For \(T = 0\),

\[
P(T = 0) = P((T_{12})_1 = \cdots = (T_{12})_p = 0) = P(L_1 = \cdots = L_p = k) = \prod_{i=1}^{p} P(L_i = k) = \left(\frac{1}{k!}\right)^p.
\]

Application for two populations

To illustrate this, we consider first the case \(k = 2\). Then we have

\[
(T_{12})_i = \begin{cases} 
\sum_{v=l}^{2} \sigma_{vi}^{-1} (\overline{X}_i - X_{vi})^2 & \text{if } X_{1i} > X_{2i} \\
0 & \text{if } X_{1i} \leq X_{2i}.
\end{cases}
\]

Now, for tests of significance the probability is

\[
P((T_{12})_i \geq t) = P(X_{1i} > X_{2i}, \sum_{v=l}^{2} \sigma_{vi}^{-1} (\overline{X}_i - X_{vi})^2 \geq t) = P(X_{1i} > X_{2i}, \frac{1}{\sigma_{vi}} (\overline{X}_i - X_{vi})^2 \geq t).
\]

Under the hypothesis \(H_0 : \mu_1 = \mu_2\), \(P(X_{1i} > X_{2i}) = \frac{1}{2}\). The distribution of \(\frac{1}{\sigma_{vi}} (\overline{X}_i - X_{vi})^2\) is \(\chi^2\) with one degree of freedom. So, if \(H_0\) be true, we have finally

\[
P((T_{12})_i \geq t) = \frac{1}{2} P(\chi^2_1 \geq t).
\]

For larger values of \(k\) the method is similar.

\[
P((T_{12})_i \geq t) = \sum_{l_i = 2}^{k} P(l_i, k) P((T_{12})_i \geq t | l_i)
\]

\[
P((T_{12})_i = 0) = P(k, k) = \frac{1}{k!}.
\]
Therefore, we have

\[ P((T_{l_2}) \geq t \mid l_{i}) = P(\chi^2_{(k-l_{i})} \geq t). \]

If the probability \( P(L_{i} = l_{i}) \) is found, then the distribution of \( T \) will be determined. To compute these probabilities see Barlow et al. (1972).

**Theorem 3.** Suppose that \( \Sigma_{s} \) are common and diagonal. If \( H_0 \) be true, then

\[
P(T \geq t) = \sum_{l=p}^{kp-1} Q(l, k, p) P(\chi^2_{kp-l} \geq t)
\]

\[ P(T = 0) = \left( \frac{1}{k!} \right)^p , \]

where \( Q(l, k, p) \) is the convolutions of the probabilities \( P(l, k) \) used in the univariate order restricted inference and defined as \( Q(l, k, p) = \sum_{l_1+\ldots+l_p = l} P(L_1 = l_1, \ldots, L_p = l_p) \) by Kulatunga (1984).

**Critical values of the test statistic**

In this section we compute the critical values of the test statistic, \( T \), for some of the bivariate and trivariate normal distributions. These values are given in tables 3.1 and 3.2.

**Table 3.1.** Critical values of the test statistic, \( T \), for \( p = 2 \).

<table>
<thead>
<tr>
<th></th>
<th>( k )</th>
<th>( .1 )</th>
<th>( .05 )</th>
<th>( .025 )</th>
<th>( .01 )</th>
<th>( .005 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.952</td>
<td>4.231</td>
<td>5.337</td>
<td>7.289</td>
<td>8.628</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5.479</td>
<td>7.059</td>
<td>8.614</td>
<td>10.640</td>
<td>12.157</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7.858</td>
<td>9.669</td>
<td>11.418</td>
<td>13.663</td>
<td>15.323</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10.188</td>
<td>12.196</td>
<td>14.111</td>
<td>16.543</td>
<td>18.329</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14.783</td>
<td>17.121</td>
<td>19.317</td>
<td>22.071</td>
<td>24.071</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>21.583</td>
<td>24.327</td>
<td>26.871</td>
<td>30.024</td>
<td>32.292</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>23.833</td>
<td>26.697</td>
<td>29.344</td>
<td>32.615</td>
<td>34.963</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>26.076</td>
<td>29.055</td>
<td>31.799</td>
<td>35.183</td>
<td>37.607</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>28.314</td>
<td>31.401</td>
<td>34.239</td>
<td>37.731</td>
<td>40.228</td>
<td></td>
</tr>
</tbody>
</table>
### Likelihood Ratio Test for Order Restrictions

Table 3.2: Critical values of the test statistic, $T$, for $p = 3$.  

<table>
<thead>
<tr>
<th>$k$</th>
<th>.1</th>
<th>.05</th>
<th>.025</th>
<th>.01</th>
<th>.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4.010</td>
<td>5.434</td>
<td>6.861</td>
<td>8.746</td>
<td>10.171</td>
</tr>
<tr>
<td>4</td>
<td>10.689</td>
<td>12.752</td>
<td>14.715</td>
<td>17.205</td>
<td>19.286</td>
</tr>
<tr>
<td>6</td>
<td>17.237</td>
<td>19.751</td>
<td>22.098</td>
<td>25.026</td>
<td>27.144</td>
</tr>
<tr>
<td>8</td>
<td>23.768</td>
<td>26.650</td>
<td>29.314</td>
<td>32.605</td>
<td>34.967</td>
</tr>
<tr>
<td>9</td>
<td>27.028</td>
<td>30.075</td>
<td>32.881</td>
<td>36.335</td>
<td>38.807</td>
</tr>
<tr>
<td>10</td>
<td>30.284</td>
<td>33.487</td>
<td>36.426</td>
<td>40.033</td>
<td>42.609</td>
</tr>
<tr>
<td>11</td>
<td>33.538</td>
<td>36.887</td>
<td>39.951</td>
<td>43.705</td>
<td>46.379</td>
</tr>
<tr>
<td>12</td>
<td>36.787</td>
<td>40.276</td>
<td>43.461</td>
<td>47.353</td>
<td>50.120</td>
</tr>
<tr>
<td>13</td>
<td>40.033</td>
<td>43.655</td>
<td>46.955</td>
<td>50.979</td>
<td>53.837</td>
</tr>
<tr>
<td>14</td>
<td>43.276</td>
<td>47.025</td>
<td>50.436</td>
<td>54.588</td>
<td>57.531</td>
</tr>
<tr>
<td>15</td>
<td>46.515</td>
<td>50.388</td>
<td>53.904</td>
<td>58.179</td>
<td>61.205</td>
</tr>
<tr>
<td>16</td>
<td>49.751</td>
<td>53.742</td>
<td>57.361</td>
<td>61.754</td>
<td>64.861</td>
</tr>
<tr>
<td>17</td>
<td>52.983</td>
<td>57.089</td>
<td>60.807</td>
<td>65.315</td>
<td>68.500</td>
</tr>
<tr>
<td>18</td>
<td>56.213</td>
<td>60.430</td>
<td>64.244</td>
<td>68.863</td>
<td>72.123</td>
</tr>
<tr>
<td>19</td>
<td>59.439</td>
<td>63.765</td>
<td>67.672</td>
<td>72.399</td>
<td>75.732</td>
</tr>
<tr>
<td>20</td>
<td>62.663</td>
<td>67.093</td>
<td>71.091</td>
<td>75.923</td>
<td>79.328</td>
</tr>
</tbody>
</table>
Simulation study to compute the power and \( p \)-value

This section is devoted to a simulation study to estimate the power of test in bivariate normal distribution. The results may be generalized to more dimensional normal distributions. Consider the simple order \( H_1: \mu_1 \leq \mu_2 \leq \ldots \leq \mu_k \). Use the critical values from the table 3.1 for \( k = 2, 3 \) and two significance levels \( \alpha = .05 \) and \( \alpha = .01 \). We suppose that the covariance matrices are diagonal. Based on 1000 replications of multivariate normal the power is calculated. In both cases the mean vectors are \( \mu_v = (\beta \nu, \beta \nu + 1) \), \( \nu = 1, 2, \ldots, k \), \( k = 2, 3 \) and \( \beta = 3, 2, \frac{3}{2}, 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{10}, \frac{1}{20}, \ldots, \frac{1}{70} \).

Table 4.1: Simulation of power for the test statistic, \( T \). significance level, \( \alpha = .05, .01 \) and alternatives \( \mu_v = (\beta \nu, \beta \nu + 1) \).

<table>
<thead>
<tr>
<th>( k )</th>
<th>( \beta )</th>
<th>( \alpha = .05 )</th>
<th>( \alpha = .05 )</th>
<th>( \alpha = .01 )</th>
<th>( \alpha = .01 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>0.782</td>
<td>1.000</td>
<td>0.637</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.735</td>
<td>1.000</td>
<td>0.663</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{3}{2} )</td>
<td>0.550</td>
<td>1.000</td>
<td>0.512</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{2} )</td>
<td>0.647</td>
<td>0.847</td>
<td>0.524</td>
<td>0.822</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{3} )</td>
<td>0.612</td>
<td>0.681</td>
<td>0.482</td>
<td>0.728</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{4} )</td>
<td>0.513</td>
<td>0.814</td>
<td>0.407</td>
<td>0.761</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{5} )</td>
<td>0.575</td>
<td>0.794</td>
<td>0.511</td>
<td>0.757</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{6} )</td>
<td>0.442</td>
<td>0.643</td>
<td>0.393</td>
<td>0.664</td>
</tr>
</tbody>
</table>
Likelihood Ratio Test for Order Restrictions

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>0.624</th>
<th>0.743</th>
<th>0.518</th>
<th>0.531</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma'$</td>
<td>0.421</td>
<td>0.828</td>
<td>0.374</td>
<td>0.443</td>
</tr>
<tr>
<td>$\gamma''$</td>
<td>0.359</td>
<td>0.721</td>
<td>0.306</td>
<td>0.624</td>
</tr>
<tr>
<td>$\gamma'''$</td>
<td>0.316</td>
<td>0.624</td>
<td>0.036</td>
<td>0.544</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>0.420</td>
<td>0.473</td>
<td>0.331</td>
<td>0.584</td>
</tr>
<tr>
<td>$\gamma_{20}$</td>
<td>0.186</td>
<td>0.426</td>
<td>0.125</td>
<td>0.642</td>
</tr>
<tr>
<td>$\gamma_{30}$</td>
<td>0.241</td>
<td>0.393</td>
<td>0.180</td>
<td>0.427</td>
</tr>
<tr>
<td>$\gamma_{40}$</td>
<td>0.117</td>
<td>0.648</td>
<td>0.102</td>
<td>0.566</td>
</tr>
<tr>
<td>$\gamma_{50}$</td>
<td>0.124</td>
<td>0.627</td>
<td>0.084</td>
<td>0.344</td>
</tr>
<tr>
<td>$\gamma_{60}$</td>
<td>0.094</td>
<td>0.178</td>
<td>0.045</td>
<td>0.342</td>
</tr>
<tr>
<td>$\gamma_{70}$</td>
<td>0.036</td>
<td>0.273</td>
<td>0.026</td>
<td>0.297</td>
</tr>
</tbody>
</table>

Following the simulation method for computing the exact $p$-value of the test statistic, $T$, for testing $H_1$ against $H_2 - H_1$:

1) Generate independent observations from bivariate normal $N_2(\mu_0, \Sigma_0)$, where $(\mu_0, \Sigma_0)$ can have any arbitrary values and the covariance matrices are diagonal.
2) The $T$ statistic is estimated.
3) Repeat the previous two steps $N = 10000$ times, and estimate the $p$-value by $M/N$ where $M$ is the number of times the simulated value of $T$ statistic in the second step is not less than its sample value.

In this simulation study, to estimate the $p$-value of the test, we consider the mean vectors $\mu_v = (\beta_v, \beta_v + 1)$, $\beta = 3, 2, \frac{3}{2}, 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{10}, \frac{1}{20}, \ldots, \frac{1}{70}$, $v = 1, 2, \ldots, k$ and $k = 2, 3$. The results are given in table 4.2.

| Table 4.2. Simulation of p-value for the test statistic $T$. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| $\beta$ $\ y/ k$ | 3   | 2   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  |
| 2   | .235 | .048 | .145 | .080 | .325 | .313 | .280 | .024 | .203 | .315 | .152 | .048 | .123 | .162 | .332 | .074 |

Likelihood ratio and its null distribution when the covariance matrices have an unknown scale factor
Suppose that the covariance matrices have the form $\Sigma_v = \sigma^2 \Lambda_v$, $v = 1, 2, \ldots, k$, where $\Lambda_v$'s are known nonsingular matrices and $\sigma^2$ is unknown.
Let $\hat{\sigma}^2$ is an estimator for $\sigma^2$ which is independent of $X_v$'s and the estimator $\frac{q\hat{\sigma}^2}{\sigma^2}$ is distributed with $\chi^2$ distribution with $q$ degrees of freedom.

Then the likelihood function is the joint density function of $(X, \hat{\sigma}^2)$, is

$$L = \prod_{v=1}^{k} c_0 \frac{1}{\sigma} e^{-\frac{1}{2\sigma^2}(X_v-\mu_v)^2} (\frac{q\hat{\sigma}^2}{\sigma^2})^{\frac{q-1}{2}} e^{-\frac{q\hat{\sigma}^2}{2\sigma^2} q}$$

$$= c_1 \frac{1}{\sigma^{k+q}} e^{-\frac{1}{2\sigma^2} \sum_{v=1}^{k} (X_v-\mu_v)^2} \frac{q\hat{\sigma}^2}{\sigma^2}$$

where $c_0$ and $c_1$ are positive constants which is independent of $\mu$ and $\sigma^2$.

The likelihood ratio test rejects $H_1$ for the small values of

$$\lambda = \frac{\sup_{H_1} L}{\sup_{H_0} L} = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_1^2}\right)^{(k+q)/2},$$

where $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are the maximum likelihood estimates of $\sigma^2$ under $H_1$ and $H_2 - H_1$ respectively.

$$\hat{\sigma}_1^2 = \frac{1}{k+q} \left[ \sum_{v=1}^{k} (X_v - \hat{\mu}_v)^2 (X_v - \hat{\mu}_v) + q \hat{\sigma}^2 \right]$$

and

$$\hat{\sigma}_2^2 = \frac{q \hat{\sigma}^2}{k+q}.$$

**Theorem 4.** The likelihood ratio test rejects $H_1$ for large values of

$$\bar{T} = 1 - \lambda^{(k+q)} = \frac{1}{\sum_{v=1}^{k} (X_v - \hat{\mu}_v)^2 (X_v - \hat{\mu}_v) + q \hat{\sigma}^2}$$

Now, suppose that the matrices $\Lambda_v$'s are diagonal with diagonal elements $d_{v1}, d_{v2}, \ldots, d_{vp}$, $v = 1, 2, \ldots, k$, then the $\bar{T}$ statistic can be written of the form
Likelihood Ratio Test for Order Restrictions

\[
T = \frac{\sum_{i=1}^{k} \sum_{v=1}^{k} \frac{1}{d_{vi}} (\hat{\mu}_{vi} - X_{vi})^2}{\sum_{i=1}^{k} \sum_{v=1}^{k} \frac{1}{d_{vi}} (\hat{\mu}_{vi} - X_{vi})^2 + q\hat{\sigma}^2},
\]

where \(\hat{\mu}_{i1}, \hat{\mu}_{i2}, \ldots, \hat{\mu}_{ik}\) is the univariate isotonic regression of \(X_{i1}, X_{i2}, \ldots, X_{ik}\) with weights \(\frac{1}{d_{i1}}, \frac{1}{d_{i2}}, \ldots, \frac{1}{d_{ik}}\).

**Theorem 5.** Suppose that \(A, \ldots, A\) are diagonal with diagonal elements \(d_{v1}, d_{v2}, \ldots, d_{vp}\), \(v = 1, 2, \ldots, k\). If \(H_0\) be true, then for \(t > 0\), the null distribution of \(T\) is

\[
P(T \geq t) = \frac{\sum_{l=1}^{k} \sum_{l_{1}, \ldots, l_{p}} P(L_{1} = l_{1}, \ldots, L_{p} = l_{p}) P(B_{\frac{1}{2}(l_{1}+\cdots+l_{p})-q} \geq t)}{\sum_{l=1}^{k} \sum_{l_{1}, \ldots, l_{p}} P(L_{1} = l_{1}, \ldots, L_{p} = l_{p}) P(T \geq t | L_{1} = l_{1}, \ldots, L_{p} = l_{p})}.
\]

where \(l = \sum_{i=1}^{k} l_{i}\) and \(B_{(\alpha, \beta)}\) is the Beta distribution with parameters \(\alpha\) and \(\beta\).

**Proof.**

\[
P(T \geq t) = \sum_{l=1}^{k} \sum_{l_{1}, \ldots, l_{p}} P(L_{1} = l_{1}, \ldots, L_{p} = l_{p}) P(T \geq t | L_{1} = l_{1}, \ldots, L_{p} = l_{p})
\]

Now, since \(\sum_{i=1}^{k} \frac{1}{d_{vi}} (\hat{\mu}_{vi} - X_{vi})^2\) given \(L_{1} = l_{1}, \ldots, L_{p} = l_{p}\) is distributed as \(\sigma^2 \chi^2_{(kq-l)}\) and independent of \(q\hat{\sigma}^2\) which is distributed as \(\sigma^2 \chi^2_{q}\). Therefore, the \(T\) statistic is distributed as \(B_{\frac{1}{2}(kq-l-q)\frac{1}{2}}\). The proof of the second part is easily.

**Theorem 6.** Suppose that \(A, \ldots, A\) are common and diagonal. If \(H_0\) be true, then

\[
P(T \geq t) = \sum_{l=1}^{k} O(l, k, p) P(B_{\frac{1}{2}(l-1)\frac{1}{2}} \geq t)
\]

\[
P(T = 0) = \left(\frac{1}{k!}\right)^p.
\]

The critical values of the \(T\) statistic for trivariate normal distribution for when \(q = 2\) are computed in table 5.1.
Table 5.1. Critical values of the test statistic, $\bar{T}$, for $p = 3, q = 2$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>.1</th>
<th>.05</th>
<th>.025</th>
<th>.01</th>
<th>.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.633</td>
<td>0.741</td>
<td>0.817</td>
<td>0.884</td>
<td>0.918</td>
</tr>
<tr>
<td>3</td>
<td>0.787</td>
<td>0.853</td>
<td>0.898</td>
<td>0.936</td>
<td>0.955</td>
</tr>
<tr>
<td>4</td>
<td>0.854</td>
<td>0.900</td>
<td>0.931</td>
<td>0.957</td>
<td>0.970</td>
</tr>
<tr>
<td>5</td>
<td>0.890</td>
<td>0.925</td>
<td>0.948</td>
<td>0.968</td>
<td>0.978</td>
</tr>
<tr>
<td>6</td>
<td>0.913</td>
<td>0.941</td>
<td>0.959</td>
<td>0.975</td>
<td>0.982</td>
</tr>
<tr>
<td>7</td>
<td>0.928</td>
<td>0.951</td>
<td>0.967</td>
<td>0.979</td>
<td>0.986</td>
</tr>
<tr>
<td>8</td>
<td>0.939</td>
<td>0.959</td>
<td>0.972</td>
<td>0.983</td>
<td>0.988</td>
</tr>
<tr>
<td>9</td>
<td>0.947</td>
<td>0.964</td>
<td>0.975</td>
<td>0.985</td>
<td>0.989</td>
</tr>
<tr>
<td>10</td>
<td>0.953</td>
<td>0.969</td>
<td>0.978</td>
<td>0.987</td>
<td>0.991</td>
</tr>
<tr>
<td>11</td>
<td>0.958</td>
<td>0.972</td>
<td>0.981</td>
<td>0.988</td>
<td>0.992</td>
</tr>
<tr>
<td>12</td>
<td>0.962</td>
<td>0.975</td>
<td>0.983</td>
<td>0.989</td>
<td>0.993</td>
</tr>
<tr>
<td>13</td>
<td>0.966</td>
<td>0.977</td>
<td>0.984</td>
<td>0.990</td>
<td>0.993</td>
</tr>
<tr>
<td>14</td>
<td>0.968</td>
<td>0.979</td>
<td>0.985</td>
<td>0.991</td>
<td>0.994</td>
</tr>
<tr>
<td>15</td>
<td>0.971</td>
<td>0.980</td>
<td>0.987</td>
<td>0.992</td>
<td>0.994</td>
</tr>
<tr>
<td>16</td>
<td>0.973</td>
<td>0.982</td>
<td>0.987</td>
<td>0.992</td>
<td>0.998</td>
</tr>
<tr>
<td>17</td>
<td>0.974</td>
<td>0.983</td>
<td>0.988</td>
<td>0.993</td>
<td>0.995</td>
</tr>
<tr>
<td>18</td>
<td>0.976</td>
<td>0.984</td>
<td>0.989</td>
<td>0.993</td>
<td>0.995</td>
</tr>
<tr>
<td>19</td>
<td>0.978</td>
<td>0.985</td>
<td>0.989</td>
<td>0.994</td>
<td>0.996</td>
</tr>
<tr>
<td>20</td>
<td>0.979</td>
<td>0.986</td>
<td>0.990</td>
<td>0.994</td>
<td>0.996</td>
</tr>
</tbody>
</table>

References


Likelihood Ratio Test for Order Restrictions


