4 Hypothesis Testing

4.1 Introduction

We will exploit the distributional results discussed in the previous chapter in order to construct and perform tests for structural relationships between components of the (population) mean vector, and then extend this to tests (on the parameters) between two independent populations characterized by the same covariance matrix.

4.2 Testing for Structural relationships amongst components of the mean

Here, we test whether the data are consistent with a particular hypothesis regarding the relationships between the components of $\mu$.

As usual, it will be assumed that the data arise from the $N_p(\mu, \Sigma)$ distribution, where both $\mu$ and $\Sigma$ are unknown.

Example 4.1 (Equality of the components of the mean vector)

Determine whether or not the components of the (population) mean vector are equal to each other.

Solution:

Formally, we seek a test for

$H_0 : \mu_j = \mu_1, \ j=2, \ldots, p, \ vs. \ H_1 : H_0 \ false$

i.e.

$H_0 : \mu_1 - \mu_j = 0, \ j=2, \ldots, p, \ vs. \ H_1 : H_0 \ false$

i.e.

$H_0 : \begin{bmatrix} \mu_1 - \mu_2 \\ \mu_1 - \mu_3 \\ \vdots \\ \mu_1 - \mu_p \end{bmatrix} = 0$

which is a $(p - 1) \times 1$ column vector.
Again, this can be rewritten as

$$
\begin{bmatrix}
1 & -1 & 0 & 0 & \ldots & 0 \\
1 & 0 & -1 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \ddots \\
1 & 0 & 0 & 0 & \ldots & -1 \\
\end{bmatrix}
\begin{bmatrix}
\mu_1 \\
\mu_2 \\
\vdots \\
\mu_p \\
\end{bmatrix}
= 0
$$

or

$$
C'\mu = 0
$$

where

$$
C =
\begin{bmatrix}
1 & 1 & 1 & \ldots & \ldots & 1 & 1 \\
-1 & 0 & 0 & \ldots & \ldots & 0 & 0 \\
0 & -1 & 0 & \ldots & \ldots & 0 & 0 \\
0 & 0 & -1 & \ldots & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ldots & \ldots & 0 & -1 \\
\end{bmatrix}
$$

which is a \( p \times (p - 1) \) matrix of constants.

The hypotheses to test are

\[ H_0 : C'\mu = 0 \text{ vs. } H_1 : C'\mu \neq 0. \]

Now, for

\[ X_i \sim N_p(\mu, \Sigma), \ i = 1, \ldots, n, \]

and independent, then

\[ \bar{X} \sim N_p(\mu, \frac{1}{n}\Sigma) \]

and so

\[ C'\bar{X} \sim N_{p-1}(C'\mu, \frac{1}{n}C'SC). \]

Also

\[ (n - 1)C'SC \sim W_{p-1}(n - 1, C'SC) \]

independently (using the fact that \((n - 1)S \sim W_p(n - 1, \Sigma)\) and 3.11 (iv)).

Hence

\[
T^2 = n(C'\bar{X} - C'\mu)'(C'SC)^{-1}(C'\bar{X} - C'\mu) \\
= n\bar{X}'C(C'SC)^{-1}C'X \sim T^2_{p-1}(n - 1)
\]

\( n > p - 1 \), under \( H_0 \).

Hence, we reject \( H_0 \) at the 100\( \alpha \)% level of significance if

\[
\frac{n - p + 1}{(n - 1)(p - 1)}T^2_{obs} > F_{p-1,n-p+1,\alpha}
\]

and accept \( H_0 \) o.w.
Example 4.2 (Symmetric electrical activity)
Sensors are placed in a symmetrical fashion across patients’ heads in order to measure the amounts of electrical activity taking place.

Construct a test to determine whether or not the amount of activity across the head is asymmetric or not.

Solution:
In essence we ask whether the mean amount of activity, in the population under consideration, is symmetrical or not.

Model the readings of the $i$-th member of a random sample of size $n$ as $X_i \sim N_p(\mu, \Sigma)$, $i = 1, \ldots, n$, where $p$ is odd, and $\mu, \Sigma$ are both unknown.

Symmetry can be posed as

$$
\mu_1 = \mu_p, \mu_2 = \mu_{p-1}, \ldots, \mu_{\frac{p-1}{2}} = \mu_{\frac{p+1}{2}}
$$

i.e.

$$
\begin{align*}
\mu_1 - \mu_p &= 0 \\
\mu_2 - \mu_{p-1} &= 0 \\
&\vdots \\
\mu_{\frac{p-1}{2}} - \mu_{\frac{p+1}{2}} &= 0 
\end{align*}
$$

or $C'\mu = 0$,

where

$$
C = \begin{bmatrix}
1 & 0 & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & 0 & \ldots & 0 \\
0 & 0 & 1 & 0 & \ldots & 0 \\
& & & & & \\
& & & & & \\
& & & & & \\
0 & 0 & -1 & 0 & \ldots & 0 \\
0 & -1 & 0 & 0 & \ldots & 0 \\
-1 & 0 & 0 & 0 & \ldots & 0 
\end{bmatrix}
$$

which is a $p \times \frac{p-1}{2}$ matrix.

Since

$$
C'\bar{X} \sim N(\frac{p-1}{2}; C'\mu, \frac{1}{n}C'\Sigma C)
$$
and 
\[(n - 1)C'SC \sim W_{(p-1)/2}(n - 1, C'\Sigma C)\]

independently, then, under \(H_0\),
\[
T^2 = n(C'\bar{X} - 0)'(C'SC)^{-1}(C'\bar{X} - 0)
\]
\[
= n\bar{X}'C(C'SC)^{-1}C'\bar{X} \sim T^2_{m/2}(n - 1)
\]

Thus, we reject the hypothesis of symmetry at the 100\(\alpha\)% level of significance if
\[
\frac{n - (p - 1)/2}{(p - 1)(n - 1)/2}T^2_{obs} > F_{(p-1)/2,n-(p-1)/2,\alpha}
\]

and accept \(H_0\) o.w.

Generalization

Assume that \(X_i \sim N_p(\mu, \Sigma), i = 1, \ldots, n\), mutually independent, where both \(\mu\) and \(\Sigma\) are unknown.

Seek a test for \(H_0 : C'\mu = \phi\) vs. \(H_1 : C'\mu \neq \phi\)

where \(C\) is a \(p \times m\) matrix of constants, \(\text{rank}(C) = m < n\), and \(\phi\) is a \(m \times 1\) vector of constants.

Use
\[
T^2 = n(C'\bar{X} - \phi)'(C'SC)^{-1}(C'\bar{X} - \phi).
\]

Since \(C'\bar{X}\) is unbiased for \(C'\mu\), then if \(H_0\) were true, \(T^2\) should be 'small'; on the other hand if \(H_0\) were false, then \(T^2\) should be 'large'.

However, it can be shown that, under \(H_0\), \(T^2 \sim T^2_m(n - 1)\) (setting \(f = n - 1\) and \(q = m\)).

Thus, we reject \(H_0\) if
\[
\frac{n - m}{(n - 1)m}T^2_{obs} > F_{m,n-m,\alpha}
\]

\(n > m\), and accept \(H_0\) o.w.

4.3 Two Sample Tests

Suppose that we draw two independent random samples of sizes \(n_1\) and \(n_2\) from populations 1 and 2 respectively.

Observations from population \(i\) are assumed to be mutually independent with distribution \(N_p(\mu_i, \Sigma)\), where both \(\mu_i\), and \(\Sigma\), are unknown.
We wish to devise a test to determine whether the two populations have the same mean or not, i.e.

\[ H_0 : \mu_1 = \mu_2 \text{ vs. } H_1 : \mu_1 \neq \mu_2. \]

Notice that we have assumed that the two populations share the same covariance matrix.

Let \( \overline{X}_1 \) and \( S_1 \) be the sample mean vector and sample covariance matrix for the random sample drawn from population 1.

Similarly, let \( \overline{X}_2 \) and \( S_2 \) be the sample mean vector and the sample covariance matrix for the random sample from population 2, drawn independently of the 1st.

Since

\[ \overline{X}_i \sim N_p(\mu_i, \frac{1}{n_i} \Sigma) \]

\( i = 1, 2 \), independently, then it can be shown that

\[ \overline{X}_1 - \overline{X}_2 \sim N_p(\mu_1 - \mu_2, \left( \frac{1}{n_1} + \frac{1}{n_2} \right) \Sigma) \] (1)

Also, define

\[ S = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2} \]

which is the pooled 'within-groups' estimator of \( \Sigma \).

Indeed \( E[S] = \Sigma \).

We know that

\[ (n_1 - 1)S_1 \sim W_p(n_1 - 1, \Sigma) \]
\[ (n_2 - 1)S_2 \sim W_p(n_2 - 1, \Sigma) \]

independently.

Thus

\[ (n_1 + n_2 - 2)S = (n_1 - 1)S_1 + (n_2 - 1)S_2 \sim W_p(n_1 + n_2 - 2, \Sigma) \] (2)

using 3.11(iii).

Setting \( f = n_1 + n_2 - 2 \) and \( q = p \), the condition \( f > q - 1 \) amounts to \( n_1 + n_2 - 2 > p - 1 \), i.e. \( n_1 + n_2 > p + 1 \).

Consider the statistic

\[ T^2 = \left( \frac{1}{n_1} + \frac{1}{n_2} \right)^{-1} (\overline{X}_1 - \overline{X}_2)'S^{-1}(\overline{X}_1 - \overline{X}_2). \]
Under $H_0$, 
\[
X_1 - X_2 \sim N_p \left( 0, \left( \frac{1}{n_1} + \frac{1}{n_2} \right) \Sigma \right) 
\sim N_p \left( 0, \frac{1}{\left( \frac{1}{n_1} + \frac{1}{n_2} \right)^{-1}} \Sigma \right).
\]

Thus, under $H_0$, $T^2 \sim T^2_p (n_1 + n_2 - 2)$ and 
\[
\left( \frac{f - q + 1}{fq} \right) T^2 = \left( \frac{n_1 + n_2 - p - 1}{p(n_1 + n_2 - 2)} \right) T^2 \sim F_{p,n_1+n_2-p-1}.
\]

A test at the $100\alpha\%$ level of significance rejects $H_0 : \mu_1 = \mu_2$ if 
\[
\left( \frac{n_1 + n_2 - p - 1}{p(n_1 + n_2 - 2)} \right) T^2_{obs} > F_{p,n_1+n_2-p-1,\alpha}.
\]

Example 4.3 (Whale Data) 
Recall Ex. 1.3: An ocean research laboratory conducted a survey on two species of whale: the Blue whale and the Bowhead whale. Three measurements were taken on each one of the whales in the survey: body length (in metres), weight (in tonnes), and right flipper length (in metres). Data on two independent random samples from the two species are presented below. 

Blue Whale(A): 
(sample size $n_1=4$) 

<table>
<thead>
<tr>
<th>Body length</th>
<th>Weight</th>
<th>Flipper length</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.30</td>
<td>109.74</td>
<td>2.46</td>
</tr>
<tr>
<td>24.96</td>
<td>108.95</td>
<td>1.95</td>
</tr>
<tr>
<td>25.36</td>
<td>109.12</td>
<td>1.75</td>
</tr>
<tr>
<td>25.74</td>
<td>109.44</td>
<td>2.35</td>
</tr>
</tbody>
</table>

Bowhead Whale(B): 
(sample size $n_2=5$) 

<table>
<thead>
<tr>
<th>Body length</th>
<th>Weight</th>
<th>Flipper length</th>
</tr>
</thead>
<tbody>
<tr>
<td>22.39</td>
<td>83.07</td>
<td>2.53</td>
</tr>
<tr>
<td>22.45</td>
<td>81.84</td>
<td>2.62</td>
</tr>
<tr>
<td>22.75</td>
<td>82.81</td>
<td>3.39</td>
</tr>
<tr>
<td>20.92</td>
<td>81.90</td>
<td>2.94</td>
</tr>
<tr>
<td>21.64</td>
<td>82.65</td>
<td>2.19</td>
</tr>
</tbody>
</table>

Would it be fair to say that the two species are dissimilar with respect to the three measurements taken on each whale?
> blue.l <- c(24.3, 24.96, 25.36, 25.74)
> blue.w <- c(109.74, 108.95, 109.12, 109.44)
> blue.f <- c(2.46, 1.95, 1.75, 2.35)
> bow.l <- c(22.39, 22.45, 22.75, 20.92, 21.64)
> bow.w <- c(83.07, 81.84, 82.81, 81.9, 82.65)
> bow.f <- c(2.53, 2.62, 3.39, 2.94, 2.19)
> blue.whale <- data.frame(blue.l, blue.w, blue.f)
> blue.whale
     blue.l blue.w blue.f
1  24.30  109.74  2.46
2  24.96  108.95  1.95
3  25.36  109.12  1.75
4  25.74  109.44  2.35
> bowhead.whale <- data.frame(bow.l, bow.w, bow.f)
> bowhead.whale
     bow.l bow.w bow.f
1  22.39  83.07  2.53
2  22.45  81.84  2.62
3  22.75  82.81  3.39
4  20.92  81.90  2.94
5  21.64  82.65  2.19
> sm.blue <- apply(blue.whale, 2, mean)
> sm.bow <- apply(bowhead.whale, 2, mean)
> cov.blue <- var(blue.whale)
> cov.bow <- var(bowhead.whale)
> sm.blue
     blue.l blue.w blue.f
1  25.09  109.3125  2.1275
> sm.bow
     bow.l bow.w bow.f
1  22.03  82.454  2.734
> cov.blue
     blue.l blue.w blue.f
blue.l 0.37880000 -0.08656667 -0.06563333
blue.w -0.08656667  0.12249167  0.10250833
blue.f -0.06563333  0.10250833  0.11135833
> cov.bow
     bow.l bow.w bow.f
bow.l 0.552150  0.189675  0.083625
bow.w 0.189675  0.307130 -0.010720
bow.f 0.083625 -0.010720  0.205830
> q <- 3
> m <- 4
> n <- 5
> s.pooled <- ((m - 1) * cov.blue + (n - 1) * cov.bow)/(m + n - 2)
> T2 <- ((m * n)/(m + n)) * t(sm.blue - sm.bow) %*% solve(s.pooled) %*% (sm.blue - sm.bow)
> f <- m + n - 2
> F.stat <- drop(((f - q + 1)/(f * q)) * T2)
> F.stat
[1] 1789.309
> p.value <- 1 - pf(F.stat, q, f - q + 1)
> p.value
[1] 5.380012e-008
The small \( p \)-value provides strong evidence that the two species are indeed dissimilar with respect to the measurements taken.

### 4.4 Further Two Sample Tests

The test procedure of the previous section can be generalized still further.

Suppose that independent random samples of sizes \( n_1 \) and \( n_2 \) are drawn from the \( N_p(\mu_1, \Sigma) \) and \( N_p(\mu_2, \Sigma) \) distributions, respectively.

Seek a test procedure for

\[
H_0 : C'(\mu_1 - \mu_2) = \phi
\]

vs.

\[
H_1 : C'(\mu_1 - \mu_2) \neq \phi
\]

where \( C \) is a \( p \times m \) matrix of constants, of rank \( m \), and \( \phi \) is a \( m \times 1 \) vector of constants.

From (1), it follows that

\[
C'(\bar{X}_1 - \bar{X}_2) \sim N_m \left( C'(\mu_1 - \mu_2), \left( \frac{1}{n_1} + \frac{1}{n_2} \right) C'\Sigma C \right)
\]

\[
\sim N_m \left( C'(\mu_1 - \mu_2), \frac{1}{\frac{n_1 + n_2}{n_1 + n_2}} C'\Sigma C \right).
\] (5)

Also, using (2),

\[
(n_1 + n_2 - 2)C'\Sigma C \sim W_m(n_1 + n_2 - 2, C'\Sigma C)
\]

independently, where \( S \) is the pooled sample covariance matrix.

Now, under \( H_0 \),

\[
C'(\bar{X}_1 - \bar{X}_2) \sim N_m \left( \phi, \frac{1}{\frac{n_1 + n_2}{n_1 + n_2}} C'\Sigma C \right).
\] (6)

Thus, from Prop. 3.13, it follows that, under \( H_0 \),

\[
T^2 = \frac{n_1 n_2}{n_1 + n_2} \left[ C'(\bar{X}_1 - \bar{X}_2) - \phi \right]' \left( C'\Sigma C \right)^{-1} \left[ C'(\bar{X}_1 - \bar{X}_2) - \phi \right] \sim T_m^2(n_1 + n_2 - 2)
\]

and

\[
\left[ f - q + 1 \right] T^2 = \left[ \frac{n_1 + n_2 - m - 1}{(n_1 + n_2 - 2)m} \right] T^2 \sim F_{m,n_1+n_2-m-1}.
\]

So we reject \( H_0 : C'(\mu_1 - \mu_2) = \phi \) at the 100\( \alpha \)% level of significance if

\[
\left[ \frac{n_1 + n_2 - m - 1}{(n_1 + n_2 - 2)m} \right] T_{\text{obs}}^2 > F_{m,n_1+n_2-m-1,\alpha}.
\]