Testing a Hypothesis about the Mean

- Consider as usual a random sample $x_1, x_2, \ldots, x_N$ from $N_p(\mu, \Sigma)$.
- Recall: if $\Sigma$ is known, we can test the null hypothesis
  \[ H_0 : \mu = \mu_0 \]
  with the size-$\alpha$ critical region
  \[ N (\bar{x} - \mu_0)' \Sigma^{-1} (\bar{x} - \mu_0) > \chi^2_p(\alpha) \]
- If $\Sigma$ is unknown, the natural statistic to use is
  \[ T^2 = N (\bar{x} - \mu_0)' S^{-1} (\bar{x} - \mu_0). \]
- We need to justify its use, and find its null distribution.
Likelihood Ratio Test

- The likelihood function is

\[
L(\mu, \Sigma) = \frac{1}{\left(\sqrt{\det 2\pi \Sigma}\right)^N} \exp \left[ -\frac{1}{2} \sum_{\alpha=1}^{N} (x_\alpha - \mu)' \Sigma^{-1} (x_\alpha - \mu) \right].
\]

- The generalized likelihood ratio statistic is

\[
\lambda = \frac{\max_{\Sigma} L(\mu_0, \Sigma)}{\max_{\mu,\Sigma} L(\mu, \Sigma)}
\]
Write $\hat{\Sigma}_\omega$ for the mle of $\Sigma$ under $H_0$, when $\mu = \mu_0$, and $\hat{\Sigma}_\Omega$ for the mle of $\Sigma$ under the alternative hypothesis, when $\mu$ is unconstrained.

Then

$$\hat{\Sigma}_\omega = \frac{1}{N} \sum_{\alpha=1}^{N} (x_\alpha - \mu_0)(x_\alpha - \mu_0)'$$

and

$$L(\mu_0, \hat{\Sigma}_\omega) = \max_{\Sigma} L(\mu_0, \Sigma) = \frac{1}{\left(\sqrt{\det 2\pi \hat{\Sigma}_\omega}\right)^{N}} \times e^{-\frac{1}{2}pN}.$$ 

Similarly

$$\hat{\Sigma}_\Omega = \frac{1}{N} \sum_{\alpha=1}^{N} (x_\alpha - \bar{x})(x_\alpha - \bar{x})'$$

and

$$L(\bar{x}, \hat{\Sigma}_\Omega) = \max_{\mu,\Sigma} L(\mu, \Sigma) = \frac{1}{\left(\sqrt{\det 2\pi \hat{\Sigma}_\Omega}\right)^{N}} \times e^{-\frac{1}{2}pN}.$$
So the generalized likelihood ratio statistic is

\[\lambda = \left(\frac{\sqrt{\det 2\pi \hat{\Sigma}_\Omega}}{\sqrt{\det 2\pi \hat{\Sigma}_\omega}}\right)^N\]

\[= \left\{ \frac{\det \left[ \sum_{\alpha=1}^{N} (x_\alpha - \bar{x})(x_\alpha - \bar{x})' \right]}{\frac{1}{2}N} \right\} \left[ \sum_{\alpha=1}^{N} (x_\alpha - \mu_0)(x_\alpha - \mu_0)' \right] \]
Now

\[
\sum_{\alpha=1}^{N} (x_{\alpha} - \mu_0)(x_{\alpha} - \mu_0)'
\]

\[
= \sum_{\alpha=1}^{N} (x_{\alpha} - \bar{x})(x_{\alpha} - \bar{x})' + N(\bar{x} - \mu_0)(\bar{x} - \mu_0)'
\]

\[
= (N - 1)S + N(\bar{x} - \mu_0)(\bar{x} - \mu_0)'
\]

whence

\[
\det \left[ \sum_{\alpha=1}^{N} (x_{\alpha} - \mu_0)(x_{\alpha} - \mu_0)' \right]
\]

\[
= \det[(N - 1)S] \left\{ 1 + N(\bar{x} - \mu_0)'[(N - 1)S]^{-1}(\bar{x} - \mu_0) \right\}
\]
So

\[
\lambda^{\frac{2}{N}} = \frac{1}{1 + N(\bar{x} - \mu_0)'[(N - 1)S]^{-1}(\bar{x} - \mu_0)}
\]

\[
= \frac{1}{1 + \frac{T^2}{N - 1}}
\]

Since \( \lambda \) is a monotone function of \( T^2 \), they define equivalent tests.

That is, the test based on \( T^2 \), known as Hotelling’s \( T^2 \), is equivalent to the generalized ratio test.
Invariance

- If $X \sim N_p(\mu, \Sigma)$ and $X^* = CX$ for a nonsingular $C$, then $X^* \sim N_p(\mu^*, \Sigma^*)$, where
  \[
  \mu^* = C\mu, \quad \Sigma^* = C\Sigma C'
  \]

- The null hypothesis $H_0: \mu = \mu_0$ is equivalent to the null hypothesis $H_0^*: \mu^* = C\mu_0$.

- If $T^2$ is the statistic for testing $H_0$, based on $x_1, x_2, \ldots, x_N$, and $T^{*2}$ is the statistic for testing $H_0^*$, based on $x_1^*, x_2^*, \ldots, x_N^*$, then $T^{*2} = T^2$.

- That is, the statistic $T^2$, and the test based on it, are invariant to the mapping $X \mapsto X^*$, $\mu \mapsto \mu^*$.

- The test based on Hotelling’s $T^2$ is the uniformly most powerful test with this invariance property.
Distribution of $T^2$

- $T^2$ may be written as $Y' S^{-1} Y$, where $Y \sim N_p(\nu, \Sigma)$, and $S$ is of the form

$$nS = \sum_{\alpha=1}^{n} Z_\alpha Z'_\alpha$$

where $Z_1, Z_2, \ldots, Z_n$ are iid $N_p(0, \Sigma)$, independent of $Y$.

- Without loss of generality, we can take $\Sigma = I_p$ (see above, “Invariance”).

- Construct a $p \times p$ orthogonal matrix $Q$ with first row $Y' / \sqrt{Y'Y}$, and the remaining rows arbitrary.

- Then

$$QY = \begin{pmatrix} \sqrt{Y'Y} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$
Hence

\[ T^2 = Y'S^{-1}Y \]
\[ = (QY)'(QSQ')^{-1}(QY) \]
\[ = (Y'Y)(QSQ')^{1,1} \]

where \( A^{i,j} \) denotes the \((i,j)\) entry of \( A^{-1} \).

Write

\[ nQSQ' = B = \begin{pmatrix} b_{1,1} & b_{(1)}' \\ b_{(1)} & B_{2,2} \end{pmatrix} \]

and note that

\[ nQSQ' = B = \sum_{\alpha=1}^{n} (QZ_\alpha)(QZ_\alpha)' \]
Q is random, but conditionally on Q, $QZ_\alpha$ has the same distribution as $Z_\alpha$, so $B$ has the same distribution as $nS$.

Since this conditional distribution does not depend on Q, it is also the unconditional distribution, and $B$ is independent of $Q$, and hence of $Y$.

Also

$$\frac{1}{b_{1,1}} = b_{1,1} - b'_{(1)} B^{-1}_{2,2} b_{(1)} = b_{1,1 \cdot 2, \ldots, p}$$

a sample partial (or residual) variance, which is therefore distributed as $\chi^2$ with $n - (p - 1)$ degrees of freedom.
So finally,

\[
\frac{T^2}{n} = \frac{\mathbf{Y}'\mathbf{Y}}{b_{1,1,2,\ldots,p}}
\]

is the ratio of a (possibly non-central) \(\chi^2\) random variable with \(p\) degrees of freedom to an independent central \(\chi^2\) random variable with \(n - (p - 1)\) degrees of freedom.

Consequently,

\[
\frac{n - (p - 1)}{p} \times \frac{T^2}{n} = \frac{\mathbf{Y}'\mathbf{Y}/p}{b_{1,1,2,\ldots,p}/[n - (p - 1)]}
\]

has the \(F\) distribution with \(p\) and \(n - (p - 1)\) degrees of freedom.

Under the null hypothesis \(\mu = \mu_0, \nu = 0\), and the distribution is central.