Testing a Hypothesis about the Mean

- Consider as usual a random sample $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- Recall: if $\boldsymbol{\Sigma}$ is known, we can test the null hypothesis

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$$

with the size-α critical region

$$N\left(\bar{\mathbf{x}} - \boldsymbol{\mu}_0\right)' \boldsymbol{\Sigma}^{-1}\left(\bar{\mathbf{x}} - \boldsymbol{\mu}_0\right) > \chi_p^2(\alpha)$$

- If $\boldsymbol{\Sigma}$ is unknown, the natural statistic to use is

$$T^2 = N\left(\bar{\mathbf{x}} - \boldsymbol{\mu}_0\right)' \mathbf{S}^{-1}\left(\bar{\mathbf{x}} - \boldsymbol{\mu}_0\right).$$

- We need to justify its use, and find its null distribution.

Likelihood Ratio Test

- The likelihood function is

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

- The generalized likelihood ratio statistic is

$$\lambda = \frac{\max_{\boldsymbol{\Sigma}} L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu},\boldsymbol{\Sigma}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})}$$

- Write $\hat{\boldsymbol{\Sigma}}_\omega$ for the mle of $\boldsymbol{\Sigma}$ under H_0, when $\boldsymbol{\mu} = \boldsymbol{\mu}_0$, and $\hat{\boldsymbol{\Sigma}}_\Omega$ for the mle of $\boldsymbol{\Sigma}$ under the alternative hypothesis, when $\boldsymbol{\mu}$ is unconstrained.
- Then

$$\hat{\boldsymbol{\Sigma}}_\omega = \frac{1}{N}\sum_{\alpha=1}^{N}\left(\mathbf{x}_\alpha - \boldsymbol{\mu}_0\right)\left(\mathbf{x}_\alpha - \boldsymbol{\mu}_0\right)'$$

  and

$$L\left(\boldsymbol{\mu}_0, \hat{\boldsymbol{\Sigma}}_\omega\right) = \max_{\boldsymbol{\Sigma}} L\left(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}\right) = \frac{1}{\left(\sqrt{\det 2\pi\hat{\boldsymbol{\Sigma}}_\omega}\right)^N} \times e^{-\frac{1}{2}pN}.$$

- Similarly

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

  and

$$L\left(\bar{\mathbf{x}}, \hat{\boldsymbol{\Sigma}}_\Omega\right) = \max_{\boldsymbol{\mu},\boldsymbol{\Sigma}} L\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right) = \frac{1}{\left(\sqrt{\det 2\pi\hat{\boldsymbol{\Sigma}}_\Omega}\right)^N} \times e^{-\frac{1}{2}pN}.$$

- So the generalized likelihood ratio statistic is

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

$$= \left\{ \frac{\det \left[\sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}}) (\mathbf{x}_\alpha - \bar{\mathbf{x}})' \right]}{\det \left[\sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \boldsymbol{\mu}_0) (\mathbf{x}_\alpha - \boldsymbol{\mu}_0)' \right]} \right\}^{\frac{1}{2}N}$$

- Now

$$\sum_{\alpha=1}^{N} \left(\mathbf{x}_\alpha - \boldsymbol{\mu}_0\right)\left(\mathbf{x}_\alpha - \boldsymbol{\mu}_0\right)'$$

$$= \sum_{\alpha=1}^{N} \left(\mathbf{x}_\alpha - \bar{\mathbf{x}}\right)\left(\mathbf{x}_\alpha - \bar{\mathbf{x}}\right)' + N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'$$

$$= (N-1)\mathbf{S} + N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'$$

whence

$$\det\left[\sum_{\alpha=1}^{N} \left(\mathbf{x}_\alpha - \boldsymbol{\mu}_0\right)\left(\mathbf{x}_\alpha - \boldsymbol{\mu}_0\right)'\right]$$

$$= \det[(N-1)\mathbf{S}]\left\{1 + N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'[(N-1)\mathbf{S}]^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)\right\}$$

- So

$$\lambda^{\frac{2}{N}} = \frac{1}{1 + N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'[(N-1)\mathbf{S}]^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)}$$
$$= \frac{1}{1 + \dfrac{T^2}{N-1}}$$

- Since λ 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 $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathbf{X}^* = \mathbf{C}\mathbf{X}$ for a nonsingular $\mathbf{C}$, then $\mathbf{X}^* \sim N_p(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*)$, where

$$\boldsymbol{\mu}^* = \mathbf{C}\boldsymbol{\mu},\ \boldsymbol{\Sigma}^* = \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}'$$

- The null hypothesis $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ is equivalent to the null hypothesis $H_0^* : \boldsymbol{\mu}^* = \mathbf{C}\boldsymbol{\mu}_0$.
- If T^2 is the statistic for testing H_0, based on $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$, and T^{*2} is the statistic for testing H_0^*, based on $\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_N^*$, then $T^{*2} = T^2$.
- That is, the statistic T^2, and the test based on it, are invariant to the mapping $\mathbf{X} \mapsto \mathbf{X}^*$, $\boldsymbol{\mu} \mapsto \boldsymbol{\mu}^*$.
- 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 $\mathbf{Y}'\mathbf{S}^{-1}\mathbf{Y}$, where $\mathbf{Y} \sim N_p(\boldsymbol{\nu}, \boldsymbol{\Sigma})$, and $\mathbf{S}$ is of the form
$$n\mathbf{S} = \sum_{\alpha=1}^{n} \mathbf{Z}_\alpha \mathbf{Z}_\alpha'$$
where $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_n$ are iid $N_p(\mathbf{0}, \boldsymbol{\Sigma})$, independent of $\mathbf{Y}$.
- Without loss of generality, we can take $\boldsymbol{\Sigma} = \mathbf{I}_p$ (see above, "Invariance").
- Construct a $p \times p$ orthogonal matrix $\mathbf{Q}$ with first row $\mathbf{Y}'/\sqrt{\mathbf{Y}'\mathbf{Y}}$, and the remaining rows arbitrary.
- Then
$$\mathbf{QY} = \begin{pmatrix} \sqrt{\mathbf{Y}'\mathbf{Y}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

- Hence

$$\begin{aligned} T^2 &= \mathbf{Y}'\mathbf{S}^{-1}\mathbf{Y} \\ &= (\mathbf{QY})'(\mathbf{QSQ}')^{-1}(\mathbf{QY}) \\ &= (\mathbf{Y}'\mathbf{Y})(\mathbf{QSQ}')^{1,1} \end{aligned}$$

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

$$n\mathbf{QSQ}' = \mathbf{B} = \begin{pmatrix} b_{1,1} & \mathbf{b}'_{(1)} \\ \mathbf{b}_{(1)} & \mathbf{B}_{2,2} \end{pmatrix}$$

  and note that

$$n\mathbf{QSQ}' = \mathbf{B} = \sum_{\alpha=1}^{n} (\mathbf{QZ}_\alpha)(\mathbf{QZ}_\alpha)'$$

- $\mathbf{Q}$ is random, but conditionally on $\mathbf{Q}$, $\mathbf{Q}\mathbf{Z}_\alpha$ has the same distribution as $\mathbf{Z}_\alpha$, so $\mathbf{B}$ has the same distribution as $n\mathbf{S}$.
- Since this *conditional* distribution does not depend on $\mathbf{Q}$, it is also the *unconditional* distribution, and $\mathbf{B}$ is independent of $\mathbf{Q}$, and hence of $\mathbf{Y}$.
- Also

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

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

- So finally,

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

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

- Consequently,

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

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

- Under the null hypothesis $\boldsymbol{\mu} = \boldsymbol{\mu}_0$, $\boldsymbol{\nu} = \mathbf{0}$, and the distribution is central.