Multivariate Statistics

Lecture 11

Fudan University

Outline

Outline

Univariate Least Squares

Consider scalar variables $x_1, \ldots, x_N$ drawn with expected values $\boldsymbol{\beta}^\top \mathbf{z}_1, \ldots, \boldsymbol{\beta}^\top \mathbf{z}_N$ respectively, where each $\mathbf{z}_\alpha \in \mathbb{R}^q$ is known and we shall estimate $\boldsymbol{\beta}$.

1. The least squares estimator of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = \left(\frac{1}{N}\sum_{i=1}^{N} \mathbf{z}_\alpha \mathbf{z}_\alpha^\top\right)^{-1} \left(\frac{1}{N}\sum_{i=1}^{N} x_\alpha \mathbf{z}_\alpha\right).$$

2. If the populations are normal, the vector $\hat{\boldsymbol{\beta}}$ is the maximum likelihood estimator of $\boldsymbol{\beta}$.

3. The unbiased estimator of the common variance σ^2 is

$$s^2 = \frac{1}{N-q}\sum_{\alpha=1}^{N}(x_\alpha - \boldsymbol{\beta}^\top \mathbf{z}_\alpha)^2.$$

4. Under the normality assumption, the maximum likelihood estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{(N-q)s^2}{N}.$$

The Estimation in Multivariate Linear Regression

Theorem 1

Suppose $\mathbf{x}_\alpha$ is an observation from $\mathcal{N}_q(\mathbf{B}\mathbf{z}_\alpha, \mathbf{\Sigma})$ for $\alpha = 1, \dots, N$, where $[\mathbf{z}_1, \dots, \mathbf{z}_N] \in \mathbb{R}^{N \times q}$ of rank q is given and $N \geq p + q$, the maximum likelihood estimator of $\mathbf{B}$ is given by

$$\hat{\mathbf{B}} = \mathbf{C}\mathbf{A}^{-1},$$

where

$$\mathbf{C} = \sum_{\alpha=1}^{N} \mathbf{x}_\alpha \mathbf{z}_\alpha^\top \qquad \text{and} \qquad \mathbf{A} = \sum_{\alpha=1}^{N} \mathbf{z}_\alpha \mathbf{z}_\alpha^\top;$$

the maximum likelihood estimator of $\mathbf{\Sigma}$ is give by

$$\hat{\mathbf{\Sigma}} = \frac{1}{N} \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \hat{\mathbf{B}}\mathbf{z}_\alpha)(\mathbf{x}_\alpha - \hat{\mathbf{B}}\mathbf{z}_\alpha)^\top.$$

Properties of the Estimators

The likelihood function is

$$L(\mathbf{B}, \mathbf{\Sigma}) = \frac{1}{(2\pi)^{\frac{Np}{2}} (\det(\mathbf{\Sigma}))^{\frac{N}{2}}} \exp\left(-\frac{1}{2} \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \mathbf{B}\mathbf{z}_\alpha)^\top \mathbf{\Sigma}^{-1} (\mathbf{x}_\alpha - \mathbf{B}\mathbf{z}_\alpha) \right).$$

We shall find $\hat{\mathbf{H}}$ such that

$$\begin{aligned}
&\sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \mathbf{B}\mathbf{z}_\alpha)(\mathbf{x}_\alpha - \mathbf{B}\mathbf{z}_\alpha)^\top \\
=&\sum_{\alpha=1}^{N} \left((\mathbf{x}_\alpha - \hat{\mathbf{H}}\mathbf{z}_\alpha)(\mathbf{x}_\alpha - \hat{\mathbf{H}}\mathbf{z}_\alpha)^\top + (\hat{\mathbf{H}}\mathbf{z}_\alpha - \mathbf{B}\mathbf{z}_\alpha)(\hat{\mathbf{H}}\mathbf{z}_\alpha - \mathbf{B}\mathbf{z}_\alpha)^\top \right).
\end{aligned}$$

Lemma 1

If $\mathbf{A} \in \mathbb{R}^{p\times p}$ and $\mathbf{G} \in \mathbb{R}^{p\times p}$ are positive definite, then $\operatorname{tr}\left(\mathbf{F}\mathbf{A}\mathbf{F}^\top \mathbf{G}\right) > 0$ for non-zero $\mathbf{F} \in \mathbb{R}^{p\times p}$.

Properties of the Estimators

The density then can be written as

$$\frac{1}{(2\pi)^{\frac{Np}{2}}(\det(\mathbf{\Sigma}))^{\frac{N}{2}}}\exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\left(N\hat{\mathbf{\Sigma}}+(\hat{\mathbf{B}}-\mathbf{B})\mathbf{A}(\hat{\mathbf{B}}-\mathbf{B})^{\top}\right)\right)\right).$$

Then $\hat{\mathbf{B}}$ and $\hat{\mathbf{\Sigma}}$ form a sufficient set statistics for $\mathbf{B}$ and $\mathbf{\Sigma}$.

Distribution of the Estimators

Let β_{ig} (or $\hat{\beta}_{ig}$) be the (i, g)-th element of $\mathbf{B}$ (or $\hat{\mathbf{B}}$).

1. The joint distribution of $\hat{\beta}_{ig}$ is normal since the $\hat{\beta}_{ig}$ are linear combinations of the $x_{i\alpha}$.
2. We have $\mathbb{E}[\hat{\mathbf{B}}] = \mathbf{B}$, which means $\hat{\mathbf{B}}$ is an unbiased estimator of $\mathbf{B}$.
3. The covariance between $\hat{\boldsymbol{\beta}}_i^\top$ and $\hat{\boldsymbol{\beta}}_j^\top$ (two rows of $\hat{\mathbf{B}}$) is $\sigma_{ij}\mathbf{A}^{-1}$.

Distribution of the Estimators

It follows that

$$N\hat{\boldsymbol{\Sigma}} = \sum_{\alpha=1}^{N} \mathbf{x}_\alpha \mathbf{x}_\alpha^\top - \hat{\mathbf{B}}\mathbf{A}\hat{\mathbf{B}}^\top$$

is distributed according to $\mathcal{W}(\boldsymbol{\Sigma}, N-q)$.

Theorem 2

Suppose $\mathbf{y}_1, \ldots, \mathbf{y}_m$ are independent with $\mathbf{y}_\alpha$ distributed according to $\mathcal{N}(\boldsymbol{\Gamma}\mathbf{w}_\alpha, \boldsymbol{\Phi})$, where $\mathbf{w}_\alpha$ is an r-component vector. Let $\mathbf{H} = \sum_{\alpha=1}^{m} \mathbf{w}_\alpha \mathbf{w}_\alpha^\top$ assumed non-singular, $\mathbf{G} = \sum_{\alpha=1}^{m} \mathbf{y}_\alpha \mathbf{w}_\alpha^\top \mathbf{H}^{-1}$ and

$$\mathbf{C} = \sum_{\alpha=1}^{m} (\mathbf{y}_\alpha - \mathbf{G}\mathbf{w}_\alpha)(\mathbf{y}_\alpha - \mathbf{G}\mathbf{w}_\alpha)^\top = \sum_{\alpha=1}^{m} \mathbf{y}_\alpha \mathbf{y}_\alpha^\top - \mathbf{G}\mathbf{H}\mathbf{G}^\top.$$

Then $\mathbf{C}$ is distributed as $\sum_{\alpha=1}^{m-r} \mathbf{u}_\alpha \mathbf{u}_\alpha^\top$ where $\mathbf{u}_1, \ldots, \mathbf{u}_{m-r}$ are independently distributed according to $\mathcal{N}(\mathbf{0}, \boldsymbol{\Phi})$ independently of $\mathbf{G}$.

The Best Linear Unbiased Estimator

Let β_{ig} be the (i, g)-th entry of $\mathbf{B}$.

An estimator F is a linear estimator of β_{ig} if

$$F = \sum_{\alpha=1}^{N} \mathbf{f}_\alpha^\top \mathbf{x}_\alpha.$$

It is a linear unbiased estimator of β_{ig} if

$$\beta_{ig} = \mathbb{E}[F] = \mathbb{E}\left[\sum_{\alpha=1}^{N} \mathbf{f}_\alpha^\top \mathbf{x}_\alpha\right] = \sum_{\alpha=1}^{N} \mathbf{f}_\alpha^\top \mathbf{B}\mathbf{z}_\alpha = \sum_{\alpha=1}^{N}\sum_{j=1}^{p}\sum_{h=1}^{q} f_{j\alpha}\beta_{jh}z_{h\alpha},$$

is an identity in $\mathbf{B}$, that is, if

$$\sum_{\alpha=1}^{N} f_{j\alpha}z_{h\alpha} = \begin{cases} 1, & j = i, h = g, \\ 0, & \text{otherwise.} \end{cases}$$

The Best Linear Unbiased Estimator

A linear unbiased estimator F is best if it has minimum variance over all linear unbiased estimators; that is, if $\mathbb{E}[(F - \beta_{ig})^2] \leq \mathbb{E}[(G - \beta_{ig})^2]$ for $G = \sum_{\alpha=1}^{N} \mathbf{g}_\alpha^\top \mathbf{x}_\alpha$ and $\mathbb{E}[G] = \beta_{ig}$.

The least squares estimator $\hat{\mathbf{B}}$ is the best linear unbiased estimator of $\mathbf{B}$.

1. Let $\tilde{\beta}_{ig} = \sum_{\alpha=1}^{N} \sum_{j=1}^{p} f_{j\alpha} x_{j\alpha}$ be arbitrary unbiased estimator of β_{ig}.
2. Then we have

$$
\begin{aligned}
&\mathbb{E}\left[(\tilde{\beta}_{ig} - \beta_{ig})^2\right] \\
=&\mathbb{E}\left[(\hat{\beta}_{ig} - \beta_{ig})^2\right] + 2\mathbb{E}\left[(\hat{\beta}_{ig} - \beta_{ig})(\tilde{\beta}_{ig} - \hat{\beta}_{ig})\right] + \mathbb{E}\left[(\tilde{\beta}_{ig} - \hat{\beta}_{ig})^2\right] \\
=&\mathbb{E}\left[(\hat{\beta}_{ig} - \beta_{ig})^2\right] + \mathbb{E}\left[(\tilde{\beta}_{ig} - \hat{\beta}_{ig})^2\right] \\
\geq&\mathbb{E}\left[(\hat{\beta}_{ig} - \beta_{ig})^2\right].
\end{aligned}
$$

Outline

Likelihood Ratio Criteria

We partition

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \end{bmatrix}$$

so that $\mathbf{B}_1$ has q_1 columns and $\mathbf{B}_2$ has q_2 columns.

We shall derive the likelihood ratio criterion for testing the hypothesis

$$H : \mathbf{B}_1 = \mathbf{B}_1^*,$$

where $\mathbf{B}_1^*$ is a given matrix.

Likelihood Ratio Criteria

The maximum of the likelihood function L for the sample $\mathbf{x}_1, \ldots, \mathbf{x}_N$ is

$$\max_{\mathbf{B}\in\mathbb{R}^{p\times q},\, \boldsymbol{\Sigma}\in\mathbb{S}_p^{++}} L(\mathbf{B}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{pN}{2}} \det\left(\hat{\boldsymbol{\Sigma}}_\Omega\right)^{-\frac{N}{2}} \exp\left(-\frac{pN}{2}\right),$$

where

$$\hat{\boldsymbol{\Sigma}}_\Omega = \frac{1}{N}\sum_{\alpha=1}^{N}(\mathbf{x}_\alpha - \hat{\mathbf{B}}\mathbf{z}_\alpha)(\mathbf{x}_\alpha - \hat{\mathbf{B}}\mathbf{z}_\alpha)^\top.$$

Likelihood Ratio Criteria

To find the maximum of the likelihood function with restricted to $\mathbf{B}_1 = \mathbf{B}_1^*$, we partition

$$\mathbf{z}_\alpha = \begin{bmatrix} \mathbf{z}_\alpha^{(1)} \\ \mathbf{z}_\alpha^{(2)} \end{bmatrix}.$$

Let $\mathbf{y}_\alpha = \mathbf{x}_\alpha - \mathbf{B}_1^* \mathbf{z}_\alpha^{(1)}$, then $\mathbf{y}_\alpha \sim \mathcal{N}\big(\mathbf{B}_2 \mathbf{z}_\alpha^{(2)}, \mathbf{\Sigma}\big)$.

Similar to the derivation of $\hat{\mathbf{B}}$, the estimator of $\mathbf{B}_2$ is

$$\hat{\mathbf{B}}_{2\omega} = \sum_{\alpha=1}^{N} \mathbf{y}_\alpha \mathbf{z}_\alpha^{(2)} \mathbf{A}_{22}^{-1} = \sum_{\alpha=1}^{N} \big(\mathbf{x}_\alpha - \mathbf{B}_1^* \mathbf{z}_\alpha^{(1)}\big) \mathbf{z}_\alpha^{(2)} \mathbf{A}_{22}^{-1} = \big(\mathbf{C}_2 - \mathbf{B}_1^* \mathbf{A}_{12}\big) \mathbf{A}_{22}^{-1},$$

with

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}.$$

Likelihood Ratio Criteria

The estimator of $\boldsymbol{\Sigma}$ is given by

$$
\begin{aligned}
N\hat{\boldsymbol{\Sigma}}_\omega &= \sum_{\alpha=1}^{N} \left(\mathbf{y}_\alpha - \hat{\mathbf{B}}_{2\omega}\mathbf{z}_\alpha^{(2)}\right)\left(\mathbf{y}_\alpha - \hat{\mathbf{B}}_{2\omega}\mathbf{z}_\alpha^{(2)}\right)^\top \\
&= \sum_{\alpha=1}^{N} \mathbf{y}_\alpha\mathbf{y}_\alpha^\top - \hat{\mathbf{B}}_{2\omega}\mathbf{A}_{22}^{-1}\hat{\mathbf{B}}_{2\omega}^\top \\
&= \sum_{\alpha=1}^{N} \left(\mathbf{x}_\alpha - \mathbf{B}_1^*\mathbf{z}_\alpha^{(1)}\right)\left(\mathbf{x}_\alpha - \mathbf{B}_1^*\mathbf{z}_\alpha^{(1)}\right)^\top - \hat{\mathbf{B}}_{2\omega}\mathbf{A}_{22}^{-1}\hat{\mathbf{B}}_{2\omega}^\top.
\end{aligned}
$$

Thus the maximum of the likelihood function over ω is

$$
(2\pi)^{-\frac{pN}{2}} \det\left(\hat{\boldsymbol{\Sigma}}_\omega\right)^{-\frac{N}{2}} \exp\left(-\frac{pN}{2}\right).
$$

The Likelihood Ratio Criterion for Testing

The likelihood ratio criterion for testing H is

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

In testing H, one rejects the hypothesis if $\lambda < \lambda_0$ where λ_0 is a suitably chosen number.

The likelihood ratio criterion for testing the null hypothesis $\mathbf{B}_1 = \mathbf{0}$ is invariant with respect to transformations $\mathbf{x}_\alpha^* = \mathbf{D}\mathbf{x}_\alpha$ for $\alpha = 1, \ldots, N$ and non-singular $\mathbf{D}$.

Outline

Testing Equality of Means with Common Covariance

Let $\mathbf{x}_\alpha^{(g)}$ be an observation from the g-th population $\mathcal{N}(\boldsymbol{\mu}^{(g)}, \boldsymbol{\Sigma})$ for $\alpha = 1, \ldots, N_g$, $g = 1, \ldots, q$.

We wish to test the hypothesis

$$H_0 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_g.$$

The likelihood function is

$$L = \prod_{g=1}^{q} \frac{1}{(2\pi)^{\frac{pN_g}{2}} (\det(\boldsymbol{\Sigma}))^{\frac{N_g}{2}}} \exp\left(-\frac{1}{2} \sum_{\alpha=1}^{N_g} \left(\mathbf{x}_\alpha^{(g)} - \boldsymbol{\mu}^{(g)}\right)^\top \boldsymbol{\Sigma}^{-1} \left(\mathbf{x}_\alpha^{(g)} - \boldsymbol{\mu}^{(g)}\right) \right).$$

1. The space Ω is the parameter space in which $\boldsymbol{\Sigma}$ is positive definite and each $\boldsymbol{\mu}^{(g)}$ is any vector.
2. The space ω is the parameter space in which $\boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_g$ (positive definite) and $\boldsymbol{\Sigma}$ is any positive definite matrix.

Testing Equality of Means with Common Covariance

Let $\mathbf{x}_\alpha^{(g)}$ be an observation from the g-th population $\mathcal{N}(\boldsymbol{\mu}^{(g)}, \boldsymbol{\Sigma}_g)$ for $\alpha = 1, \dots, N_g$, $g = 1, \dots, q$.

We wish to test the hypothesis $H_1 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_g$.

Let $N = \sum_{g=1}^q N_g$, $\mathbf{A} = \sum_{g=1}^q \mathbf{A}_g$,

$$\mathbf{A}_g = \sum_{\alpha=1}^{N_g} \left(\mathbf{x}_\alpha^{(g)} - \bar{\mathbf{x}}^{(g)}\right)\left(\mathbf{x}_\alpha^{(g)} - \bar{\mathbf{x}}^{(g)}\right)^\top \quad \text{and} \quad \mathbf{B} = \sum_{g=1}^{q} \sum_{\alpha=1}^{N_g} \left(\mathbf{x}_\alpha^{(g)} - \bar{\mathbf{x}}\right)\left(\mathbf{x}_\alpha^{(g)} - \bar{\mathbf{x}}\right)^\top.$$

The maximum likelihood estimators of $\boldsymbol{\mu}^{(g)}$ and $\boldsymbol{\Sigma}$ in Ω are given by

$$\hat{\boldsymbol{\mu}}_\Omega^{(g)} = \bar{\mathbf{x}}^{(g)} \quad \text{and} \quad \hat{\boldsymbol{\Sigma}}_\Omega = \frac{1}{N}\mathbf{A}.$$

The maximum likelihood estimators of $\boldsymbol{\mu}^{(g)}$ and $\boldsymbol{\Sigma}$ in ω are given by

$$\hat{\boldsymbol{\mu}}_\omega^{(g)} = \bar{\mathbf{x}} \quad \text{and} \quad \hat{\boldsymbol{\Sigma}}_\omega = \frac{1}{N}\mathbf{B}.$$

Lemma 2

If $\mathbf{D} \in \mathbb{R}^{p \times p}$ is positive definite, the maximum of

$$f(\mathbf{G}) = -N \ln \det(\mathbf{G}) - \mathrm{tr}(\mathbf{G}^{-1}\mathbf{D})$$

with respect to positive definite matrices $\mathbf{G}$ exists, occurs at $\mathbf{G} = \frac{1}{N}\mathbf{D}$.

Testing Equality of Means with Common Covariance

The likelihood ratio criterion for testing H_0 is

$$\lambda_0 = \frac{\left(\det\left(\hat{\mathbf{\Sigma}}_\Omega\right)\right)^{\frac{N}{2}}}{\left(\det\left(\hat{\mathbf{\Sigma}}_\omega\right)\right)^{\frac{N}{2}}} = \frac{(\det(\mathbf{A}))^{\frac{N}{2}}}{(\det(\mathbf{B}))^{\frac{N}{2}}}.$$

The critical region is

$$\lambda_0 \leq \lambda_0(\epsilon),$$

where $\lambda_0(\epsilon)$ is defined so that above inequality holds with probability ϵ when H_0 is true.

Outline

Testing Equality of Several Covariance Matrices

Let $\mathbf{x}_\alpha^{(g)}$ be an observation from the g-th population $\mathcal{N}(\boldsymbol{\mu}^{(g)}, \boldsymbol{\Sigma}_g)$ for $\alpha = 1, \ldots, N_g$, $g = 1, \ldots, q$.

We wish to test the hypothesis

$$H_1 : \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_g.$$

The likelihood function is

$$L = \prod_{g=1}^{q} \frac{1}{(2\pi)^{\frac{pN_g}{2}} (\det(\boldsymbol{\Sigma}_g))^{\frac{N_g}{2}}} \exp\left(-\frac{1}{2} \sum_{\alpha=1}^{N_g} \left(\mathbf{x}_\alpha^{(g)} - \boldsymbol{\mu}^{(g)}\right)^\top \boldsymbol{\Sigma}_g^{-1} \left(\mathbf{x}_\alpha^{(g)} - \boldsymbol{\mu}^{(g)}\right) \right).$$

1. The space Ω is the parameter space in which each $\boldsymbol{\Sigma}_g$ is positive definite and $\boldsymbol{\mu}^{(g)}$ are any vector.
2. The space ω is the parameter space in which $\boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_g$ (positive definite) and $\boldsymbol{\mu}^{(g)}$ are any vector.

Testing Equality of Several Covariance Matrices

Let

$$N = \sum_{g=1}^{q} N_g, \quad \mathbf{A}_g = \sum_{\alpha=1}^{N_g} \left(\mathbf{x}_\alpha^{(g)} - \bar{\mathbf{x}}^{(g)}\right)\left(\mathbf{x}_\alpha^{(g)} - \bar{\mathbf{x}}^{(g)}\right)^\top \text{ and } \mathbf{A} = \sum_{g=1}^{q} \mathbf{A}_g.$$

The maximum likelihood estimators of $\boldsymbol{\mu}^{(g)}$ and $\boldsymbol{\Sigma}_g$ in Ω are given by

$$\hat{\boldsymbol{\mu}}_\Omega^{(g)} = \bar{\mathbf{x}}^{(g)} \quad \text{and} \quad \hat{\boldsymbol{\Sigma}}_{g\Omega} = \frac{1}{N_g}\mathbf{A}_g.$$

The maximum likelihood estimators of $\boldsymbol{\mu}^{(g)}$ and $\boldsymbol{\Sigma}_g$ in ω are given by

$$\hat{\boldsymbol{\mu}}_\Omega^{(g)} = \bar{\mathbf{x}}^{(g)} \quad \text{and} \quad \hat{\boldsymbol{\Sigma}}_{g\Omega} = \frac{1}{N}\mathbf{A}.$$

Testing Equality of Several Covariance Matrices

The likelihood ratio criterion for testing H_1 is

$$\lambda_1 = \frac{\prod_{g=1}^{q} \left(\det \left(\hat{\boldsymbol{\Sigma}}_{g\Omega} \right) \right)^{\frac{N_g}{2}}}{\left(\det \left(\hat{\boldsymbol{\Sigma}}_{\omega} \right) \right)^{\frac{N}{2}}} = \frac{\prod_{g=1}^{q} (\det(\mathbf{A}_g))^{\frac{N_g}{2}}}{(\det(\mathbf{A}))^{\frac{N}{2}}} \cdot \frac{N^{\frac{pN}{2}}}{\prod_{g=1}^{q} N_g^{\frac{pN_g}{2}}}.$$

The critical region is

$$\lambda_1 \leq \lambda_1(\epsilon),$$

where $\lambda_1(\epsilon)$ is defined so that above inequality holds with probability ϵ when H_1 is true.

Testing Equality of Several Covariance Matrices

Bartlett (1937a) has suggested using the numbers of degrees of freedom. Except for constants, the statistic is

$$V_1 = \frac{\prod_{g=1}^{q} (\det(\mathbf{A}_g))^{\frac{n_g}{2}}}{(\det(\mathbf{A}))^{\frac{n}{2}}},$$

where $n_g = N_g - 1$ and $n = N - q$.

The statistic is invariant with respect to linear transformation

$$\mathbf{x}^{*(g)} = \mathbf{C}\mathbf{x}^{(g)} + \boldsymbol{\nu}^{(g)}.$$

Outline

Testing that Several Normal Distribution are Identical

Let $\mathbf{x}_\alpha^{(g)}$ be an observation from the g-th population $\mathcal{N}(\boldsymbol{\mu}^{(g)}, \boldsymbol{\Sigma}_g)$ for $\alpha = 1, \ldots, N_g$, $g = 1, \ldots, q$.

We wish to test

$$H_2 : \boldsymbol{\mu}^{(1)} = \cdots = \boldsymbol{\mu}^{(q)}, \quad \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_q. \tag{1}$$

1. Let Ω be the unrestricted parameter space of $\{\boldsymbol{\mu}^{(g)}, \boldsymbol{\Sigma}_g\}_{g=1}^q$, where $\boldsymbol{\Sigma}_g$ is positive definite; and ω^* consists of the space restricted by (1).
2. The likelihood function is

$$L = \prod_{g=1}^q \frac{1}{(2\pi)^{\frac{pN_g}{2}} (\det(\boldsymbol{\Sigma}_g))^{\frac{N_g}{2}}} \exp\left(-\frac{1}{2} \sum_{\alpha=1}^{N_g} \left(\mathbf{x}_\alpha^{(g)} - \boldsymbol{\mu}^{(g)}\right)^\top \boldsymbol{\Sigma}_g^{-1} \left(\mathbf{x}_\alpha^{(g)} - \boldsymbol{\mu}^{(g)}\right) \right).$$

Testing that Several Normal Distribution are Identical

Let $\mathbf{y}$ be an observation with density $f(\mathbf{y};\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a parameter vector in a space Ω.

1. Let H_a be the hypothesis $\boldsymbol{\theta} \in \Omega_a \subset \Omega$.
2. Let H_b be the hypothesis $\boldsymbol{\theta} \in \Omega_b \subset \Omega_a$ given $\boldsymbol{\theta} \in \Omega_a$.
3. Let H_{ab} be the hypothesis $\boldsymbol{\theta} \in \Omega_b$ given $\boldsymbol{\theta} \in \Omega$.

If the likelihood ratio criterion λ_a, λ_b and λ_{ab} for testing H_a, H_b and H_{ab} are uniquely defined for the observation vector $\mathbf{y}$, then we have

$$\lambda_a = \frac{\max_{\boldsymbol{\theta}\in\Omega_a} f(\mathbf{y};\boldsymbol{\theta})}{\max_{\boldsymbol{\theta}\in\Omega} f(\mathbf{y};\boldsymbol{\theta})}, \quad \lambda_b = \frac{\max_{\boldsymbol{\theta}\in\Omega_b} f(\mathbf{y};\boldsymbol{\theta})}{\max_{\boldsymbol{\theta}\in\Omega_a} f(\mathbf{y};\boldsymbol{\theta})} \quad \text{and} \quad \lambda_{ab} = \frac{\max_{\boldsymbol{\theta}\in\Omega_b} f(\mathbf{y};\boldsymbol{\theta})}{\max_{\boldsymbol{\theta}\in\Omega} f(\mathbf{y};\boldsymbol{\theta})}.$$

Hence, $\lambda_{ab} = \lambda_a\lambda_b$.

Testing that Several Normal Distribution are Identical

Recall that

1. $H_1 : \mathbf{\Sigma}_1 = \cdots = \mathbf{\Sigma}_g$;
2. $H_0 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_g$ (common covariance matrix);
3. $H_2 : \boldsymbol{\mu}^{(1)} = \cdots = \boldsymbol{\mu}^{(q)}, \quad \mathbf{\Sigma}_1 = \cdots = \mathbf{\Sigma}_q$.

Then we have

$$\begin{aligned}\lambda_2 = \lambda_1 \lambda_0 &= \frac{\prod_{g=1}^{q} (\det(\mathbf{A}_g))^{\frac{N_g}{2}}}{(\det(\mathbf{A}))^{\frac{N}{2}}} \cdot \frac{N^{\frac{pN}{2}}}{\prod_{g=1}^{q} N_g^{\frac{pN_g}{2}}} \cdot \frac{(\det(\mathbf{A}))^{\frac{N}{2}}}{(\det(\mathbf{B}))^{\frac{N}{2}}} \\ &= \left(\prod_{g=1}^{q} \frac{(\det(\mathbf{A}_g))^{\frac{N_g}{2}}}{N_g^{\frac{pN_g}{2}}} \right) \frac{N^{\frac{pN}{2}}}{(\det(\mathbf{B}))^{\frac{N}{2}}}.\end{aligned}$$

Outline

Testing that the Covariance is Proportional to $\mathbf{I}$

We use a sample of p-component vectors $\mathbf{x}_1, \ldots, \mathbf{x}_N$ from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to test the hypothesis

$$H : \boldsymbol{\Sigma} = \sigma^2 \mathbf{I},$$

where σ^2 is not specified.

The hypothesis H is a combination of the hypothesis:

1. H_1 : $\boldsymbol{\Sigma}$ is diagonal;
2. H_2 : The diagonal elements of $\boldsymbol{\Sigma}$ are equal given that $\boldsymbol{\Sigma}$ is diagonal.

The criterion for H_1 is

$$\lambda_1 = \frac{(\det(\mathbf{A}))^{\frac{N}{2}}}{\prod_{i=1}^{p} a_{ii}^{\frac{N}{2}}},$$

where $\mathbf{A} = \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top$ and a_{ij} is the (i, j)-th element of $\mathbf{A}$.

Testing that the Covariance is Proportional to $\mathbf{I}$

We can find λ_2 by considering test equality of several covariance matrices.

1. View the ith component of $\mathbf{x}_\alpha$ as the α-th observation from the i-th population.
2. p here is q in the section of testing equality of several covariance matrices; N here is N_g there; pN here is N there.
3. Thus, we have

$$\lambda_2 = \frac{\prod_{i=1}^{p}\left(\sum_{\alpha=1}^{N}(x_{i\alpha} - \bar{x}_i)^2\right)^{\frac{N}{2}}}{\left(\sum_{i=1}^{p}\sum_{\alpha=1}^{N}(x_{i\alpha} - \bar{x}_i)^2/p\right)^{\frac{pN}{2}}} = \frac{\prod_{i=1}^{p} a_{ii}^{\frac{N}{2}}}{(\mathrm{tr}(\mathbf{A})/p)^{\frac{pN}{2}}}.$$

Thus the criterion for H is

$$\lambda_1\lambda_2 = \frac{(\det(\mathbf{A}))^{\frac{N}{2}}}{\prod_{i=1}^{p} a_{ii}^{\frac{N}{2}}} \cdot \frac{\prod_{i=1}^{p} a_{ii}^{\frac{N}{2}}}{(\mathrm{tr}(\mathbf{A})/p)^{\frac{pN}{2}}} = \frac{(\det(\mathbf{A}))^{\frac{N}{2}}}{(\mathrm{tr}(\mathbf{A})/p)^{\frac{pN}{2}}}.$$

Testing that the Covariance is Proportional to $\boldsymbol{\Psi}_0$

For the hypothesis

$$\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{\Psi}_0,$$

let $\mathbf{C}$ be matrix such that

$$\mathbf{C}\boldsymbol{\Psi}_0\mathbf{C}^\top = \mathbf{I},$$

$\mathbf{x}_\alpha^* = \mathbf{C}\mathbf{x}$, $\boldsymbol{\mu}^* = \mathbf{C}\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}^* = \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^\top$.

Then hypothesis is transformed into $\boldsymbol{\Sigma}^* = \sigma^2 \boldsymbol{\Psi}_0$ and the criterion is

$$\frac{(\det(\mathbf{A}\boldsymbol{\Psi}_0^{-1}))^{\frac{N}{2}}}{(\mathrm{tr}(\mathbf{A}\boldsymbol{\Psi}_0^{-1})/p)^{\frac{pN}{2}}}.$$

Outline

Testing that the Covariance is Equal to a Give Matrix

We use a sample of p-component vectors $\mathbf{x}_1, \ldots, \mathbf{x}_N$ from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to test the hypothesis

$$\boldsymbol{\Sigma} = \mathbf{I}.$$

The likelihood ratio criterion is

$$\lambda_1 = \frac{\max_{\boldsymbol{\mu} \in \mathbb{R}^p} L(\boldsymbol{\mu}, \mathbf{I})}{\max_{\boldsymbol{\mu} \in \mathbb{R}^p, \boldsymbol{\Sigma} \in \mathbb{S}_p^{++}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})},$$

where

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

Testing that the Covariance is Equal to a Give Matrix

Then we have

$$\lambda_1 = \frac{(2\pi)^{-\frac{pN}{2}} \exp\left(- \frac{1}{2} \sum_{\alpha=1}^{N} \left(\mathbf{x}_\alpha - \bar{\mathbf{x}}\right)^\top \left(\mathbf{x}_\alpha - \bar{\mathbf{x}}\right)\right)}{(2\pi)^{-\frac{pN}{2}} \det\left(\frac{1}{N}\mathbf{A}\right)^{-\frac{N}{2}} \exp\left(-\frac{pN}{2}\right)}$$
$$= \left(\frac{\mathrm{e}}{N}\right)^{\frac{pN}{2}} (\det(\mathbf{A}))^{\frac{N}{2}} \exp\left(-\frac{\mathrm{tr}(\mathbf{A})}{2}\right),$$

where $\mathbf{A} = \sum_{\alpha=1}^{N}(\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top$.

Testing that the Covariance is Equal to a Give Matrix

To test the hypothesis

$$H_1 : \mathbf{\Sigma} = \mathbf{\Sigma}_0.$$

The likelihood ratio criterion is

$$\lambda_1 = \left(\frac{\mathrm{e}}{N}\right)^{\frac{pN}{2}} (\det(\mathbf{A}\mathbf{\Sigma}_0^{-1}))^{\frac{N}{2}} \exp\left(-\frac{\mathrm{tr}(\mathbf{A}\mathbf{\Sigma}_0^{-1})}{2}\right).$$

Testing that the Mean and the Covariance Simultaneously

Theorem 3

Given the p-component observation vectors $\mathbf{x}_1, \ldots, \mathbf{x}_N$, from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the likelihood ratio criterion for testing the hypothesis

$$H : \boldsymbol{\mu} = \boldsymbol{\mu}_0, \quad \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$$

is

$$\lambda = \left(\frac{\mathrm{e}}{N}\right)^{\frac{pN}{2}} \left(\det\left(\mathbf{A}\boldsymbol{\Sigma}_0^{-1}\right)\right)^{\frac{N}{2}} \exp\left(-\frac{1}{2}\left(\operatorname{tr}\left(\mathbf{A}\boldsymbol{\Sigma}_0^{-1}\right) + N(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^\top \boldsymbol{\Sigma}_0^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)\right)\right),$$

where $\mathbf{A} = \sum_{\alpha=1}^{N}(\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top$.

We consider hypothesises

1. $H_1 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$;
2. $H_2 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ given $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$.