

# Multivariate Statistical Analysis

## Lecture 11

Fudan University

luoluo@fudan.edu.cn

# Outline

- 1 The Likelihood Ratio Criterion
- 2 The Asymptotic Distribution of Sample Correlation
- 3 The Wishart Distribution

# Outline

- 1 The Likelihood Ratio Criterion
- 2 The Asymptotic Distribution of Sample Correlation
- 3 The Wishart Distribution

# The Likelihood Ratio Criterion

The likelihood ratio criterion:

- ① Let  $L(\mathbf{x}, \theta)$  be the likelihood function of the observation  $\mathbf{x}$  and the parameter vector  $\theta \in \Omega$ .
- ② Let a null hypothesis be defined by a proper subset  $\omega$  of  $\Omega$ . The likelihood ratio criterion is

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \omega} L(\mathbf{x}, \theta)}{\sup_{\theta \in \Omega} L(\mathbf{x}, \theta)}.$$

- ③ The likelihood ratio test is the procedure of rejecting the null hypothesis when  $\lambda(\mathbf{x})$  is less than a predetermined constant.

# Test $\rho = \rho_0$ by the Likelihood Ratio Criterion

We consider the likelihood ratio test of the hypothesis that  $\rho = \rho_0$  based on a sample  $\mathbf{x}_1, \dots, \mathbf{x}_N$  from

$$\mathcal{N}_2 \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho \\ \sigma_1 \sigma_2 \rho & \sigma_2^2 \end{bmatrix} \right).$$

Define the set

$$\Omega = \{(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) : \boldsymbol{\mu} \in \mathbb{R}^2, \sigma_1 > 0, \sigma_2 > 0, \boldsymbol{\Sigma} \succ \mathbf{0}\}$$

and its subset

$$\omega = \{(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) : \boldsymbol{\mu} \in \mathbb{R}^2, \sigma_1 > 0, \sigma_2 > 0, \boldsymbol{\Sigma} \succ \mathbf{0}, \rho = \rho_0\}.$$

We also follow the notation

$$r = \frac{a_{12}}{\sqrt{a_{11}} \sqrt{a_{22}}}, \quad \mathbf{A} = \sum_{\alpha=1}^N (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top \quad \text{and} \quad \bar{\mathbf{x}} = \frac{1}{N} \sum_{\alpha=1}^N \mathbf{x}_\alpha.$$

# Test $\rho = \rho_0$ by the Likelihood Ratio Criterion

The likelihood ratio criterion is

$$\frac{\sup_{\theta \in \omega} L(\mathbf{x}, \theta)}{\sup_{\theta \in \Omega} L(\mathbf{x}, \theta)} = \left( \frac{(1 - \rho_0^2)(1 - r^2)}{(1 - \rho_0 r)^2} \right)^{\frac{N}{2}}.$$

The likelihood ratio test is

$$\frac{(1 - \rho_0^2)(1 - r^2)}{(1 - \rho_0 r)^2} \leq c$$

where  $c$  is chosen by the prescribed significance level.

# The Maximum Likelihood Estimators

Let  $\phi : \mathcal{S} \rightarrow \mathcal{S}^*$  (may be not one-to-one) and

$$\phi^{-1}(\theta^*) = \{\theta : \theta^* = \phi(\theta)\}.$$

and define (the induced likelihood function)

$$g(\theta^*) = \sup\{f(\theta) : \theta^* = \phi(\theta)\}.$$

If  $\theta = \hat{\theta}$  maximize  $f(\theta)$ , then  $\theta^* = \phi(\hat{\theta})$  also maximize  $g(\theta^*)$ .

## Test $\rho = \rho_0$ by the Likelihood Ratio Criterion

The critical region can be written equivalently as

$$(\rho_0^2 c - \rho_0^2 + 1)r^2 - 2\rho_0 c r + c - 1 + \rho_0^2 \geq 0,$$

that is,

$$r > \frac{\rho_0 c + (1 - \rho_0^2)\sqrt{1 - c}}{\rho_0^2 c - \rho_0^2 + 1} \quad \text{and} \quad r < \frac{\rho_0 c - (1 - \rho_0^2)\sqrt{1 - c}}{\rho_0^2 c - \rho_0^2 + 1}.$$

Thus the likelihood ratio test of  $H : \rho = \rho_0$  against alternatives  $\rho \neq \rho_0$  has a rejection region of the form  $r > r_1$  and  $r < r_2$ .

# Outline

- 1 The Likelihood Ratio Criterion
- 2 The Asymptotic Distribution of Sample Correlation
- 3 The Wishart Distribution

# The Asymptotic Distribution of Sample Correlation

For a sample  $\mathbf{x}_1, \dots, \mathbf{x}_N$  from a normal distribution  $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , we are interested in the asymptotic behavior of sample correlation coefficient

$$r(n) = \frac{a_{ij}(n)}{\sqrt{a_{ii}(n)}\sqrt{a_{jj}(n)}}$$

where  $n = N - 1$ ,

$$a_{ij}(n) = \sum_{\alpha=1}^N (x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j) = \sum_{\alpha=1}^n \begin{bmatrix} z_{i\alpha} \\ z_{j\alpha} \end{bmatrix} \begin{bmatrix} z_{i\alpha} & z_{j\alpha} \end{bmatrix}$$

with

$$\begin{bmatrix} z_{i\alpha} \\ z_{j\alpha} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{ii} & \sigma_{ij} \\ \sigma_{ji} & \sigma_{jj} \end{bmatrix} \right) \quad \text{and} \quad \bar{x}_i = \frac{1}{N} \sum_{\alpha=1}^N x_{i\alpha}.$$

# The Asymptotic Distribution of Sample Correlation

## Theorem

Let

$$\mathbf{A}(n) = \sum_{\alpha=1}^N (\mathbf{x}_\alpha - \bar{\mathbf{x}}_N)(\mathbf{x}_\alpha - \bar{\mathbf{x}}_N)^\top,$$

where  $\mathbf{x}_1, \dots, \mathbf{x}_N$  are independently distributed according to  $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  and  $n = N - 1$ . Then the limiting distribution of

$$\mathbf{B}(n) = \frac{1}{\sqrt{n}}(\mathbf{A}(n) - n\boldsymbol{\Sigma})$$

is normal with mean  $\mathbf{0}$  and covariance of the entries of  $\mathbf{B}(n)$  is

$$\mathbb{E}[b_{ij}(n)b_{kl}(n)] = \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}.$$

# The Asymptotic Distribution of Sample Correlation

We achieve

$$\lim_{n \rightarrow \infty} \frac{\sqrt{n}(r(n) - \rho)}{1 - \rho^2} \sim \mathcal{N}(0, 1).$$

by applying the following theorem.

## Theorem (Serfling (1980), Section 3.3)

Let  $\{\mathbf{u}(n)\}$  be a sequence of  $m$ -component random vectors and  $\mathbf{b}$  a fixed vector such that

$$\lim_{n \rightarrow \infty} \sqrt{n}(\mathbf{u}(n) - \mathbf{b}) \sim \mathcal{N}(\mathbf{0}, \mathbf{T}).$$

Let  $\mathbf{f}(\mathbf{u})$  be a vector-valued function of  $\mathbf{u}$  such that each component  $f_j(\mathbf{u})$  has a nonzero differential at  $\mathbf{u} = \mathbf{b}$ , and define  $\Phi_{\mathbf{b}}$  with its  $(i, j)$ -th component being

$$\left. \frac{\partial f_j(\mathbf{u})}{\partial u_i} \right|_{\mathbf{u}=\mathbf{b}}.$$

Then  $\sqrt{n}(\mathbf{f}(\mathbf{u}(n)) - \mathbf{f}(\mathbf{b}))$  has the limiting distribution  $\mathcal{N}(\mathbf{0}, \Phi_{\mathbf{b}}^{\top} \mathbf{T} \Phi_{\mathbf{b}})$ .

# Outline

- 1 The Likelihood Ratio Criterion
- 2 The Asymptotic Distribution of Sample Correlation
- 3 The Wishart Distribution

# John Wishart



John Wishart (November 28th, 1898 – July 14th, 1956)

# The Wishart Distribution

We consider the distribution of

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

where  $\mathbf{x}_1, \dots, \mathbf{x}_N$  are independent, each with the distribution  $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  and  $N > p$ .

The distribution of  $\mathbf{A}$  is often called Wishart distribution with  $n$  degrees of freedom and scale parameter  $\boldsymbol{\Sigma}$ , written as

$$\mathbf{A} \sim \mathcal{W}_p(\boldsymbol{\Sigma}, n) \quad \text{where } n = N - 1 \geq p.$$

# The Wishart Distribution

We can write

$$\mathbf{A} = \sum_{\alpha=1}^n \mathbf{z}_\alpha^\top \mathbf{z}_\alpha,$$

where  $\mathbf{z}_1, \dots, \mathbf{z}_n$  are independent, each with the distribution  $\mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$  and  $N = n - 1$ .

For  $p = 1$ , we have

$$a \sim \mathcal{W}_1(\sigma^2, n) \quad \text{and} \quad \frac{a}{\sigma^2} \sim \chi^2(n).$$

# Properties of Wishart Distribution

## Theorem

Let  $\mathbf{A} \sim \mathcal{W}_p(\boldsymbol{\Sigma}, n)$  and  $\mathbf{C} \in \mathbb{R}^{q \times p}$ , then

$$\mathbf{CAC}^\top \sim \mathcal{W}_p(\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^\top, n).$$

For any  $\mathbf{t} \in \mathbb{R}^p$ , we have

$$\mathbf{t}^\top \mathbf{A} \mathbf{t} \sim \mathcal{W}_1(\mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}, n) \quad \text{and} \quad \frac{\mathbf{t}^\top \mathbf{A} \mathbf{t}}{\mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}} \sim \chi^2(n).$$

# Properties of Wishart Distribution

## Theorem

If  $\mathbf{A}_1, \dots, \mathbf{A}_k$  are independently distributed with  $\mathbf{A}_i \sim \mathcal{W}(\boldsymbol{\Sigma}, n_i)$  for  $i = 1, \dots, k$ , then

$$\mathbf{A} = \sum_{i=1}^k \mathbf{A}_i \sim \mathcal{W}\left(\boldsymbol{\Sigma}, \sum_{i=1}^k n_i\right).$$

# Properties of Wishart Distribution

## Theorem

Let

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1^\top \\ \vdots \\ \mathbf{z}_n^\top \end{bmatrix} \in \mathbb{R}^{n \times p},$$

where  $\mathbf{z}_1, \dots, \mathbf{z}_n$  are independent, each with the distribution

$$\mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma}).$$

For projection matrix  $\mathbf{Q} \in \mathbb{R}^{n \times n}$  with rank- $r$ , we have

$$\mathbf{Z}^\top \mathbf{Q} \mathbf{Z} \sim \mathcal{W}_p(\boldsymbol{\Sigma}, r).$$

# The Density of Wishart Distribution

The density of  $\mathbf{A} \sim \mathcal{W}_p(\boldsymbol{\Sigma}, n)$  is

$$\frac{(\det(\mathbf{A}))^{\frac{n-p-1}{2}} \exp\left(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{A})\right)}{2^{\frac{np}{2}} \pi^{\frac{p(p-1)}{4}} (\det(\boldsymbol{\Sigma}))^{\frac{n}{2}} \prod_{i=1}^p \Gamma\left(\frac{1}{2}(n+1-i)\right)}.$$

for positive definite  $\mathbf{A}$  and 0 elsewhere.

# Quiz 1

Let  $\mathbf{A} \in \mathbb{R}^{p \times p}$  and define  $\mathbf{F} : \mathbb{R}^{p \times q} \rightarrow \mathbb{R}^{p \times q}$  as

$$\mathbf{F}(\mathbf{X}) = \mathbf{AX}.$$

What is the determinant of Jacobian of  $\mathbf{F}(\mathbf{X})$ ?

## Quiz 1.5

Let  $\mathbf{A} \in \mathbb{R}^{p \times p}$ ,  $\mathbf{B} \in \mathbb{R}^{q \times q}$  and define  $\mathbf{F} : \mathbb{R}^{p \times q} \rightarrow \mathbb{R}^{p \times q}$  as

$$\mathbf{F}(\mathbf{X}) = \mathbf{AXB}.$$

What is the determinant of Jacobian of  $\mathbf{F}(\mathbf{X})$ ?

## Quiz 2

Let  $\mathbf{A} \in \mathbb{R}^{p \times p}$  be non-singular and define  $\mathbf{F} : \mathbb{S}^p \rightarrow \mathbb{R}^{p \times p}$  as

$$\mathbf{F}(\mathbf{X}) = \mathbf{AXA}^\top,$$

where  $\mathbb{S}^p = \{\mathbf{X} \in \mathbb{R}^{p \times p} : \mathbf{X} = \mathbf{X}^\top\}.$

What is the determinant of Jacobian of  $\mathbf{F}(\mathbf{X})$ ?