Multivariate Statistical Analysis

Lecture 12

Fudan University

luoluo@fudan.edu.cn

Outline

1. The Density of Wishart Distribution
2. T^2-Statistic and F-Distribution
3. The Inverted Wishart Distribution

Outline

1. The Density of Wishart Distribution
2. T^2-Statistic and F-Distribution
3. The Inverted Wishart Distribution

The Density of Wishart Distribution

Theorem

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

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

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

Sketch of the proof:

1. Observe that $\mathbf{B} = \mathbf{\Sigma}^{-1/2}\mathbf{A}\mathbf{\Sigma}^{-1/2} \sim \mathcal{W}_p(\mathbf{I}_p, n)$.
2. Find the density of $\mathbf{B} \sim \mathcal{W}_p(\mathbf{I}_p, n)$ by induction.
3. Recall that the Jacobian of transform from $\mathbf{A}$ to $\mathbf{B}$ has determinant

$$(\det(\mathbf{\Sigma}^{-1/2}))^{p+1} = (\det(\mathbf{\Sigma}))^{-\frac{p+1}{2}}.$$

4. Achieve the density of $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}, n)$.

The multivariate gamma function is defined as

$$\Gamma_p(t) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p} \Gamma\left(t - \frac{1}{2}(i-1)\right).$$

We also write the density function of Wishart distribution as

$$w_p(\mathbf{A} \mid \mathbf{\Sigma}, n) = \frac{(\det(\mathbf{A}))^{\frac{n-p-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{A}\right)\right)}{2^{\frac{np}{2}}\Gamma_p\left(\frac{n}{2}\right)(\det(\mathbf{\Sigma}))^{\frac{n}{2}}}.$$

Properties of Wishart Distribution

Let $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}, n)$ and partition $\mathbf{A}$ and $\mathbf{\Sigma}$ into q and $p-q$ rows and columns as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \qquad \text{and} \qquad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix},$$

then we have

(a) $\mathbf{A}_{11} \sim \mathcal{W}_q(\mathbf{\Sigma}_{11}, n)$ and $\mathbf{A}_{22} \sim \mathcal{W}_{p-q}(\mathbf{\Sigma}_{22}, n)$;

(b) if $q = 1$, then

$$\mathbf{a}_{21} \mid \mathbf{A}_{22} \sim \mathcal{N}_{p-q}(\mathbf{A}_{22}\mathbf{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_{21}, \sigma_{11.2}^2\mathbf{A}_{22})$$

where $\sigma_{11.2}^2 = \sigma_{11} - \boldsymbol{\sigma}_{21}^\top\mathbf{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_{21}$;

(c) if $n > p - q$, then

$$\mathbf{A}_{11.2} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} \sim \mathcal{W}_q(\mathbf{\Sigma}_{11.2}, n - p + q)$$

is independent on $\mathbf{A}_{22}$ and $\mathbf{A}_{12}$, where $\mathbf{\Sigma}_{11.2} = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$.

Outline

1 The Density of Wishart Distribution

2 T^2-Statistic and F-Distribution

3 The Inverted Wishart Distribution

The Distribution of Sample Covariance

Recall that we define

$$\mathbf{A} = \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top \quad \text{and} \quad \mathbf{S} = \frac{1}{n} \sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top$$

where $\mathbf{x}_1, \dots, \mathbf{x}_N$ are independent, each with the distribution $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and $n = N - 1 \geq p$.

We have $\mathbf{A} \sim \mathcal{W}_p(\boldsymbol{\Sigma}, n)$, then

$$\mathbf{S} \sim \mathcal{W}_p\left(\frac{1}{n}\boldsymbol{\Sigma}, n\right)$$

T^2-Statistic

In hypothesis testing for the mean with unknown variance, we consider the t-student variable

$$t = \frac{\bar{x} - \mu}{s/\sqrt{N}},$$

where $\bar{x} = \frac{1}{N}\sum_{\alpha=1}^{N} x_\alpha$ and $s^2 = \frac{1}{N-1}\sum_{\alpha=1}^{N}(x_\alpha - \bar{x})^2$.

We have $t^2 = \frac{N(\bar{x}-\mu)^2}{s^2}$ and its multivariate analog is

$$T^2 = N(\bar{\mathbf{x}} - \boldsymbol{\mu})^\top \mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}),$$

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

F-Distribution

The F-distribution with d_1 and d_2 degrees of freedom is the distribution of

$$z = \frac{y_1/d_1}{y_2/d_2} = \frac{d_2 y_1}{d_1 y_2},$$

where $y_1 \sim \chi^2_{d_1}$ and $y_2 \sim \chi^2_{d_2}$ are independent, written as

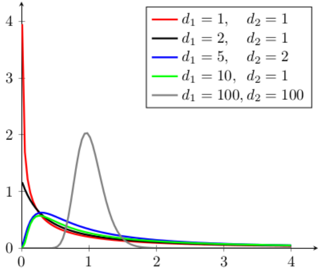
$$z \sim F_{d_1,d_2}.$$

F-Distribution

The density function of F-distribution is

$$f(z; d_1, d_2) = \frac{1}{B\left(\frac{d_1}{2}, \frac{d_2}{2}\right)} \left(\frac{d_1}{d_2}\right)^{\frac{d_1}{2}} z^{\frac{d_1}{2}-1} \left(1 + \frac{d_1 z}{d_2}\right)^{-\frac{d_1+d_2}{2}}$$

for $z > 0$, where $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,\mathrm{d}t$.



T^2-Statistic and F-Distribution

Theorem

Let $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}, n)$ *and* $\mathbf{y} \sim \mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma})$ *be independent with* $n \geq p$*, then*

$$\frac{n-p+1}{p} \cdot \mathbf{y}^\top \mathbf{A}^{-1} \mathbf{y} \sim F_{p,n-p+1}.$$

For T^2-statistic

$$T^2 = N(\bar{\mathbf{x}} - \boldsymbol{\mu})^\top \mathbf{S}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}),$$

we have

$$\frac{N-p}{(N-1)p} \cdot T^2 \sim F_{p,n-p+1}.$$

Outline

1. The Density of Wishart Distribution

2. T^2-Statistic and F-Distribution

3. The Inverted Wishart Distribution

The Inverted Wishart Distribution

If $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}, m)$, then $\mathbf{B} = \mathbf{A}^{-1}$ has the inverted Wishart distribution with m degrees of freedom and scale parameter $\mathbf{\Psi} = \mathbf{\Sigma}^{-1}$, written as

$$\mathbf{B} \sim \mathcal{W}_p^{-1}(\mathbf{\Psi}, m).$$

The density function of $\mathbf{B}$ is

$$w^{-1}(\mathbf{B} \mid \mathbf{\Psi}, m) = \frac{(\det(\mathbf{\Psi}))^{\frac{m}{2}} (\det(\mathbf{B}))^{-\frac{m+p+1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{\Psi}\mathbf{B}^{-1}\right)\right)}{2^{\frac{mp}{2}}\Gamma_p\left(\frac{m}{2}\right)},$$

where

$$\Gamma_p(t) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p} \Gamma\Big(t - \frac{1}{2}(i-1)\Big).$$

Quiz

Define $\bar{\mathbb{S}}^p \to \mathbb{R}^{p\times p}$ as

$$\mathbf{F}(\mathbf{X}) = \mathbf{X}^{-1},$$

where $\bar{\mathbb{S}}^p = \{\mathbf{X} \in \mathbb{R}^{p\times p} : \mathbf{X} = \mathbf{X}^\top \text{ and } \mathbf{X} \text{ is non-singular}\}$.

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

Quiz

Let $\mathbf{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. We define $\boldsymbol{\Psi} = \boldsymbol{\Sigma}^{-1}$ as the precision matrix.

1. It is well-known that

$$\sigma_{ij} = 0 \text{ if and only if } x_i \text{ and } x_j \text{ are independent.}$$

2. What is the meaning of $\psi_{ij} = 0$?

The Conjugate Prior for the Covariance

Theorem

If $\mathbf{A} \sim \mathcal{W}_p(\mathbf{\Sigma}, n)$ and $\mathbf{\Sigma}$ has a prior distribution $\mathcal{W}^{-1}(\mathbf{\Psi}, m)$, then the conditional distribution of $\mathbf{\Sigma}$ given $\mathbf{A}$ is the inverted Wishart distribution

$$\mathcal{W}^{-1}(\mathbf{A} + \mathbf{\Psi}, n + m).$$

Let each of $\mathbf{x}_1, \dots, \mathbf{x}_N$ has distribution $\mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma})$ independently and $n = N - 1$, then the sample covariance

$$\mathbf{S} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top \sim \mathcal{W}_p(\mathbf{\Sigma}, n).$$

If $\mathbf{\Sigma} \sim \mathcal{W}_p^{-1}(\mathbf{\Psi}, m)$, then we have

$$\mathbf{\Sigma} \,|\, \mathbf{S} \sim \mathcal{W}^{-1}(n\mathbf{S} + \mathbf{\Psi}, n + m).$$

The Inverted Wishart Distribution

Theorem

Let $\mathbf{x}_1, \dots, \mathbf{x}_N$ *be observations from* $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *Suppose* $\boldsymbol{\mu}$ *and* $\boldsymbol{\Sigma}$ *have prior densities*

$$n\left(\boldsymbol{\mu} \;\middle|\; \boldsymbol{\nu}, \frac{\boldsymbol{\Sigma}}{K}\right) \qquad \text{and} \qquad w^{-1}(\boldsymbol{\Sigma} \mid \boldsymbol{\Psi}, m)$$

respectively, where $n = N - 1$. *Then the posterior density of* $\boldsymbol{\mu}$ *and* $\boldsymbol{\Sigma}$ *given*

$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{\alpha=1}^{N} \mathbf{x}_\alpha \quad \text{and} \quad \mathbf{S} = \frac{1}{N-1}\sum_{\alpha=1}^{N} (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})^\top$$

is

$$n\left(\boldsymbol{\mu} \;\middle|\; \frac{N\bar{\mathbf{x}} + K\boldsymbol{\nu}}{N+K}, \frac{\boldsymbol{\Sigma}}{N+K}\right) \cdot w^{-1}\left(\boldsymbol{\Sigma} \mid \boldsymbol{\Psi} + n\mathbf{S} + \frac{NK(\bar{\mathbf{x}} - \boldsymbol{\nu})(\bar{\mathbf{x}} - \boldsymbol{\nu})^\top}{N+K}, N + m\right).$$