Wishart Distribution

Konstantin Yakovlev

MIPT, 2022

September 13, 2022

Definition

Definition

Let $\{\mathbf{x}_i\}_{i=1}^n, \mathbf{x}_i \in \mathbb{R}^p$ be i.i.d. samples from $\mathcal{N}(\mathbf{x}|\mathbf{0}, \mathbf{\Sigma})$. Then

$$\mathbf{M} = \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top \sim \mathcal{W}_p(\mathbf{\Sigma}, n),$$

where n is a number of degrees of freedom, $\mathbf{\Sigma}$ is a scale matrix.

Note that when $p = 1$, it is easy to see that $\mathbf{M} \sim \chi_n^2$, i.e. $\mathcal{W}_1(1, n) =^d \chi_n^2$.

Properties

Property 1

Let $\mathbf{M} \sim \mathcal{W}_p(\mathbf{M}|\mathbf{\Sigma}, n)$. Let $\mathbf{A} \in \mathbb{R}^{q \times p}$ be a a constant matrix. Then

$$\mathbf{AMA}^\top \sim \mathbf{W}_q(\mathbf{A\Sigma A}^\top, n).$$

Proof.

Rewrite the definition of $\mathbf{AMA}^\top$:

$$\mathbf{AMA}^\top = \sum_{i=1}^{n} \mathbf{A}\mathbf{x}_i\mathbf{x}_i^\top\mathbf{A}^\top.$$

We know that $\mathbf{y} = \mathbf{Ax} \sim \mathcal{N}(\mathbf{y}|\mathbf{0}, \mathbf{A\Sigma A}^\top)$. Then

$$\sum_{i=1}^{n} \mathbf{y}_i\mathbf{y}_i^\top \sim \mathcal{W}_q(.|\mathbf{A\Sigma A}^\top, n).$$

□

Properties

Cochran's theorem

Let $\mathbf{P} \in \mathbb{R}^{n \times n}$ be a projection matrix ($\mathbf{P} = \mathbf{P}^2$) of rank r. Let $\mathbf{X} = [\mathbf{x}_1^\top, \dots, \mathbf{x}_n^\top]^\top \in \mathbb{R}^{n \times p}$. Rows are i.i.d.. Each row $\mathbf{x}_i \sim \mathcal{N}(\mathbf{x}_i | \mathbf{0}, \mathbf{\Sigma})$. Let $\mathrm{rank}\mathbf{\Sigma} = p$. Then

$$\mathbf{X}^\top \mathbf{P} \mathbf{X} \sim \mathcal{W}(.|\mathbf{\Sigma}, r).$$

$$\mathbf{X}^\top (\mathbf{I} - \mathbf{P}) \mathbf{X} \sim \mathcal{W}(.|\mathbf{\Sigma}, n - r).$$

Moreover, $\mathbf{X}^\top \mathbf{P} \mathbf{X}$ and $\mathbf{X}^\top (\mathbf{I} - \mathbf{P}) \mathbf{X}$ are independant.

Property 2

Let $\{\mathbf{x}_i\}_{i=1}^n$ be i.i.d. samples from $\mathcal{N}(\mathbf{x} | \boldsymbol{\mu}, \mathbf{\Sigma})$. Then

$$n\mathbf{S} = \sum_{i=1}^{n} (\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^\top \sim \mathcal{W}_p(.|\mathbf{\Sigma}, n - 1).$$

Proof of property 2

Let $\mathbf{P} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top \in \mathbb{R}^{n\times n}, \mathrm{rank}\mathbf{P} = 1.$ Consider
$\mathbf{P}^2 = \mathbf{I} + \frac{1}{n^2}\mathbf{1}\mathbf{1}^\top\mathbf{1}\mathbf{1}^\top - \frac{2}{n}\mathbf{1}\mathbf{1}^\top = \mathbf{P}$.
Now consider $\mathbf{X}^\top\mathbf{P}\mathbf{X} = \sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^\top - \frac{1}{n}\mathbf{X}^\top\mathbf{1}\mathbf{1}^\top\mathbf{X} = n\mathbf{S}$. Here we used the fact that $\overline{\mathbf{x}} = \frac{1}{n}\mathbf{X}^\top\mathbf{1}$.
Hence, using Cochran's theorem,

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

Hotelling's T^2 dustribution

Definition

Let $\mathbf{x} \sim \mathcal{N}(x|\mathbf{0}, \mathbf{I}), \mathbf{x} \in \mathbb{R}^p$. Let $\mathbf{M} \sim \mathcal{W}_p(\mathbf{M}|\mathbf{I}, n), \mathbf{M} \in \mathbb{R}^{p \times p}$. $\mathbf{M}$ and $\mathbf{x}$ are independent. Then

$$n\mathbf{x}^\top \mathbf{M}^{-1}\mathbf{x} \sim T^2(p, n).$$

Property 3

Let $\mathbf{x} \sim \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\mathbf{M} \sim \mathcal{W}_p(\mathbf{M}|\boldsymbol{\Sigma}, n)$ are independent. Let also $\det \boldsymbol{\Sigma} \neq 0$. Then

$$n(\mathbf{x} - \boldsymbol{\mu})^\top \mathbf{M}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \sim T^2(p, n).$$

Proof.

Let $\mathbf{y} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{x} - \mu)$ is a standard normal random vector. Let $\mathbf{Z} = \boldsymbol{\Sigma}^{-1/2}\mathbf{M}\boldsymbol{\Sigma}^{-1/2} \sim \mathcal{W}_p(\mathbf{Z}|\mathbf{I}, n)$. Then $n\mathbf{y}^\top \mathbf{Z}\mathbf{y} \sim T^2(p, n)$. $\square$

Properties

Theorem

$$\frac{n-p}{(n-1)p} n\mathbf{x}^\top \mathbf{M}^{-1}\mathbf{x} \sim F_{p,n-p},$$

where $F_{\cdot,\cdot}$ is a Fisher distribution.

Corollary

From property 2, property 3 and Cochran's theorem follows that

$$\frac{n-p}{p}(\overline{\mathbf{x}} - \boldsymbol{\mu})\mathbf{S}^{-1}(\overline{\mathbf{x}} - \mu) \sim F_{p,n-p}.$$

Statistical tests

Known $\boldsymbol{\Sigma}$

Given $\{\mathbf{x}_i\}_{i=1}^n$ from $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}), \mathbf{x} \in \mathbb{R}^p$.
$H_0: \quad \boldsymbol{\mu} = \boldsymbol{\mu}_0$
$H_1: \quad \boldsymbol{\mu} <\neq> \boldsymbol{\mu}_0$
$T(\mathbf{X}) = n(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^\top \boldsymbol{\Sigma}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$
Null distribution: χ^2_p.

Unknown $\boldsymbol{\Sigma}$

Given $\{\mathbf{x}_i\}_{i=1}^n$ from $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}), \mathbf{x} \in \mathbb{R}^p$.
$H_0: \quad \boldsymbol{\mu} = \boldsymbol{\mu}_0$
$H_1: \quad \boldsymbol{\mu} <\neq> \boldsymbol{\mu}_0$
$T(\mathbf{X}) = \frac{n-p}{p}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^\top \mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$.
Null distribution: $F_{p,n-p}$

Fast Sampling

Odell and Feiveson, 1966

Let $\{V_i\}_{i=1}^{p}$ be independent random variables: $V_i \sim \chi^2_{p-i+1}$. Let $\{N_{ij}\}_{ij=1}^{n}$ be indepentent standard normal random variables idendependent of V_i. Define $b_{ij} = bji$ for $i, j \in \overline{1, p}$:

$$b_{ii} = V_i + \sum_{r=1}^{i-1} N_{ri}^2,$$

$$b_{ij} = N_{ij}\sqrt{V_i} + \sum_{r=1}^{i-1} N_{ri}N_{rj}.$$

Then $\mathbf{B} = \{b_{ij}\} \sim \mathcal{W}_p(\mathbf{I}, n)$.

Note that here we have to sample $O(p^2)$ stardard normal random variables instead of $O(np)$ is naive sampling.

Literature

1. **Definition, properties and tests**
   https://rich-d-wilkinson.github.io/MATH3030/7-multinormal.html
2. **Fast sampling**
   https://www.math.wustl.edu/ sawyer/hmhandouts/Wishart.pdf