MNOHOROZMĚRNÉ NORMÁLNÍ ROZDĚLENÍ
3.10.2019

DEFINICE: Nechť $\mathbf{Z} = (Z_1, \dots, Z_r)^\top$, kde $Z_i \sim \mathsf{N}(0,1)$ jsou nezávislé náhodné veličiny. Nechť $\boldsymbol{\mu} \in \mathbb{R}^n$ a $\mathbf{A}_{n\times r} \in \mathbb{R}^{n\times r}$ je matice. Pak řekneme, že $\mathbf{X} = \boldsymbol{\mu} + \mathbf{AZ}$ má **n-rozměrné normální rozdělení** $\mathsf{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde $\boldsymbol{\Sigma} = \mathbf{AA}^T$.

Diskuze k definici:

– Je-li $\boldsymbol{\Sigma}_{n\times n}$ pozitivně semidefinitní matice s hodností $h(\boldsymbol{\Sigma}) = r \le n$, pak existuje matice $\mathbf{A}_{n\times r}$ taková, že $h(\mathbf{A}) = r$ a $\mathbf{AA}^T = \boldsymbol{\Sigma}$.

– Rozdělení $\boldsymbol{\mu} + \mathbf{AZ}$ z definice nezávisí na konkrétní volbě $\mathbf{A}$, závisí pouze na $\boldsymbol{\Sigma} = \mathbf{AA}^T$.

VLASTNOSTI: Ukažte, že platí:

1. Má-li $\mathbf{X} \sim \mathsf{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, pak $E\mathbf{X} = \boldsymbol{\mu}$, $\mathsf{Var}\,\mathbf{X} = \boldsymbol{\Sigma}$.

2. Je-li $\mathbf{B} \in \mathbb{R}^{k\times n}$, $\boldsymbol{\mu}_2 \in \mathbb{R}^k$, pak $\mathbf{BX} + \boldsymbol{\mu}_2$ má rozdělení $\mathsf{N}_k(\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\mu}_2, \mathbf{B\Sigma B}^T)$.

   Poznámka: Dokonce platí, že $\mathbf{X}$ má mnohorozměrné normální rozdělení $\Leftrightarrow$ $\mathbf{c}^T\mathbf{X}$ má normální rozdělení pro každé $\mathbf{c} \in \mathbb{R}^n$.

3. Existence hustoty:

   – Je-li $\boldsymbol{\Sigma}$ regulární, pak existuje hustota vzhledem k Lebesgueově míře a je tvaru

   $$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \boldsymbol{\Sigma}}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right\}, \quad \mathbf{x} \in \mathbb{R}^n.$$

   – Je-li $\boldsymbol{\Sigma}$ singulární, pak hustota vzhledem k Lebesgueově míře neexistuje.

4. Nechť $\mathbf{X} \sim \mathsf{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ a označme $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)^T$, kde $\mathbf{X}_1$ tvoří prvních k složek $\mathbf{X}$.

   (i) Marginální rozdělení $\mathbf{X}_1$ i $\mathbf{X}_2$ je normální, speciálně $\mathbf{X}_1 \sim \mathsf{N}_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$, kde $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)^T$ a $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^T & \boldsymbol{\Sigma}_{22} \end{pmatrix}$.

   (ii) Je-li $\boldsymbol{\Sigma}_{12} = \mathbf{0}$, pak jsou $\mathbf{X}_1$ a $\mathbf{X}_2$ nezávislé.

5. Nechť $X \sim \mathsf{N}(0,1)$ a nechť $1/2 = \mathsf{P}(Z = 1) = \mathsf{P}(Z = -1)$ jsou nezávislé náhodné veličiny. Definujme $Y = Z \cdot X$. Ukažte, že

   (a) X i Y mají rozdělení $\mathsf{N}(0,1)$,
   (b) veličiny X a Y jsou nekorelované, ale nejsou nezávislé,
   (c) sdružené rozdělení X a Y není normální.

6. Nechť $\mathbf{X} \sim \mathsf{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde $\boldsymbol{\Sigma}$ je regulární. Pak

   $$(\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2_n.$$

7. Nechť $\mathbf{X} \sim \mathsf{N}_n(\mathbf{0}, \boldsymbol{\Sigma})$. Pak náhodná veličina $||\mathbf{X}||^2 = \mathbf{X}^\top\mathbf{X} = \sum_{i=1}^n X_i^2$ má stejné rozdělení jako $\sum_{i=1}^n \lambda_i Y_i^2$, kde Y_i jsou iid veličiny s $\mathsf{N}(0,1)$ rozdělením a $\lambda_1, \dots, \lambda_n$ jsou vlastní čísla $\boldsymbol{\Sigma}$. Speciálně pak platí, že

- $\mathsf{E}||\mathbf{X}||^2 = \mathrm{tr}(\mathbf{\Sigma})$,
- pokud má $\mathbf{\Sigma}$ vlastní čísla pouze 0 nebo 1, pak má veličina $\|\mathbf{X}\|^2$ rozdělení χ^2 se stupni volnosti rovnými $\mathrm{h}(\mathbf{\Sigma}) = \mathrm{tr}(\mathbf{\Sigma})$.

8. Nalezněte asymptotickou aproximaci rozdělení χ^2_n pro velké n pomocí normálního rozdělení.

Opakování

Spektrální rozklad matice. Nechť $\mathbf{\Sigma}_{n\times n}$ je symetrická reálná matice. Pak existuje $\mathbf{Q}$ orthogonální a $\mathbf{\Lambda}$ diagonální matice takové, že

$$\mathbf{\Sigma} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T.$$

χ^2 rozdělení. Nechť $Z_1, \dots, Z_n$ jsou nezávislé náhodné veličiny s rozdělením $\mathsf{N}(0,1)$. Pak $Y = \sum_{i=1}^n Z_i^2$ má χ^2 rozdělení s n stupni volnosti.

Význam normálního rozdělení: Mnohorozměrná centrální limitní věta. Nechť $\{\mathbf{X}_n\}$ je náhodný výběr z rozdělení se střední hodnotou $\boldsymbol{\mu}$ a konečnou rozptylovou maticí $\mathbf{\Sigma}$. Pak platí

$$\sqrt{n}(\overline{\mathbf{X}}_n - \boldsymbol{\mu}) \xrightarrow{D} \mathsf{N}(0, \mathbf{\Sigma}), \quad n \to \infty.$$