ME731 - Métodos em Análise Multivariada
– Distribuição Normal Multivariada II –

Prof. Carlos Trucíos
ctrucios@unicamp.br
ctruciosm.github.io

Instituto de Matemática, Estatística e Computação Científica,
Universidade Estadual de Campinas

Aula 05

UNICAMP

Agenda I

1. Introdução
2. Distribuição Normal
3. Distribuição Wishart
4. Distribution T^2 de Hotelling
5. Teorema Central do Limite

Introdução

Introdução

- A distribuição Normal multivariada é uma generalização da distribuição Normal univariada.
- A distribuição Normal multivariada é completamente definida pelos seus dois primeiros momentos, um total de $\frac{1}{2}p(p+1)+p=\frac{1}{2}p(p+3)$ parâmetros.
- Correlação zero implica independência (isto não acontece necessariamente em outras distribuições).
- Funções lineares de um vetor aleatorio com distribuição Normal multivariada tem distribuição normal univariada.
- Mesmo quando os dados não são normalmente distribuidos, sob algumas suposições, podemos invocar o TCL.

Distribuição Normal

Distribuição Normal

Estimadores de Máxima Verossimilhança

Sejam $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim N_p(\mu, \Sigma)$. Então,

$$\hat{\mu} = \bar{\mathbf{X}} \quad e \quad \hat{\Sigma} = \mathbf{S} = \frac{1}{n}\sum_{j=1}^{n}(\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})'$$

são os estimadores de máxima verossimilhança de μ e Σ, respectivamente.

Distribuição Normal

Estimadores de Máxima Verossimilhança

Sejam $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim N_p(\mu, \Sigma)$. Então,

$$\hat{\mu} = \bar{\mathbf{X}} \quad e \quad \hat{\Sigma} = \mathbf{S} = \frac{1}{n}\sum_{j=1}^{n}(\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})'$$

são os estimadores de máxima verossimlhança de μ e Σ, respectivamente.

Resultado útil para a demostração: Seja $\mathbf{B}_{p\times p} > 0$ uma matriz simétrica e seja o escalar $b > 0$, pode-se provar que

$$\frac{1}{|\Sigma|^b} e^{-Tr(\Sigma^{-1}\mathbf{B})/2} \leq \frac{1}{|\mathbf{B}|^b}(2b)^{pb} e^{-bp},$$

para qualquer matriz definida positiva Σ. A Igualdade acontece se $\Sigma = (1/2b)\mathbf{B}$. [Demonstração em Johnson e Wichern (2007)].

Distribuição Normal

Demostração:

Distribuição Normal

Demostração:

Obs 1: continua valendo a propriedade de invariância, ou seja, o EMV de $h(\theta)$ é $h(\hat{\theta})$.

Distribuição Normal

Demostração:

Obs 1: continua valendo a propriedade de invariância, ou seja, o EMV de $h(\theta)$ é $h(\hat{\theta})$.

Obs 2: $\bar{\mathbf{X}}$ e $\mathbf{S}$ são estatísticas suficientes.

Distribuição Normal

Teorema

Se $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim^{iid} f(\mathbf{x}, \theta)$ com $\mathbb{E}(\mathbf{X}_i) = \mu$ e $\mathbb{V}(\mathbf{X}_i) = \Sigma$. Então,

- $\mathbb{E}(\bar{\mathbf{X}}) = \mu$
- $\mathbb{V}(\bar{\mathbf{X}}) = \frac{1}{n}\Sigma$
- $\mathbb{E}(\mathbf{S}) = \frac{n-1}{n}\Sigma$

Distribuição Normal

Teorema

Se $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim^{iid} f(\mathbf{x}, \theta)$ com $\mathbb{E}(\mathbf{X}_i) = \mu$ e $\mathbb{V}(\mathbf{X}_i) = \Sigma$. Então,

- $\mathbb{E}(\bar{\mathbf{X}}) = \mu$
- $\mathbb{V}(\bar{\mathbf{X}}) = \frac{1}{n}\Sigma$
- $\mathbb{E}(\mathbf{S}) = \frac{n-1}{n}\Sigma$

Demostração:

Distribuição Normal

Teorema

Se $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim N_p(\mu, \Sigma)$. Então

$$\bar{\mathbf{X}} \sim N_p(\mu, n^{-1}\Sigma)$$

Distribuição Normal

Teorema

Se $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim N_p(\mu, \Sigma)$. Então

$$\bar{\mathbf{X}} \sim N_p(\mu, n^{-1}\Sigma)$$

Demostração:

Distribuição Normal

Teorema

Se $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim N_p(\mu, \Sigma)$. Então

$$\bar{\mathbf{X}} \sim N_p(\mu, n^{-1}\Sigma)$$

Demostração:

$$\varphi_{\bar{\mathbf{X}}}(\mathbf{t}) = \varphi_{\frac{1}{n}\mathbf{X}_1 + \cdots + \frac{1}{n}\mathbf{X}_n}(\mathbf{t}) = \prod_{j=1}^{n} \varphi_{\frac{1}{n}\mathbf{X}_j}(\mathbf{t}) = \prod_{j=1}^{n} \varphi_{\mathbf{X}_j}(\frac{1}{n}\mathbf{t})$$

Distribuição Normal

Teorema

Se $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim N_p(\mu, \Sigma)$. Então

$$\bar{\mathbf{X}} \sim N_p(\mu, n^{-1}\Sigma)$$

Demostração:

$$\varphi_{\bar{\mathbf{X}}}(\mathbf{t}) = \varphi_{\frac{1}{n}\mathbf{X}_1 + \cdots + \frac{1}{n}\mathbf{X}_n}(\mathbf{t}) = \prod_{j=1}^{n} \varphi_{\frac{1}{n}\mathbf{X}_j}(\mathbf{t}) = \prod_{j=1}^{n} \varphi_{\mathbf{X}_j}(\frac{1}{n}\mathbf{t})$$

$$= \prod_{j=1}^{n} e^{i\frac{1}{n}\mathbf{t}'\mu - \frac{1}{n^2}\mathbf{t}'\Sigma\mathbf{t}} = e^{i\mathbf{t}'\mu - \frac{1}{n}\mathbf{t}'\Sigma\mathbf{t}}$$

Distribuição Normal

Teorema

Se $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim N_p(\mu, \Sigma)$. Então

$$\bar{\mathbf{X}} \sim N_p(\mu, n^{-1}\Sigma)$$

Demostração:

$$\varphi_{\bar{\mathbf{X}}}(\mathbf{t}) = \varphi_{\frac{1}{n}\mathbf{X}_1+\cdots+\frac{1}{n}\mathbf{X}_n}(\mathbf{t}) = \prod_{j=1}^{n} \varphi_{\frac{1}{n}\mathbf{X}_j}(\mathbf{t}) = \prod_{j=1}^{n} \varphi_{\mathbf{X}_j}(\frac{1}{n}\mathbf{t})$$

$$= \prod_{j=1}^{n} e^{i\frac{1}{n}\mathbf{t}'\mu - \frac{1}{n^2}\mathbf{t}'\Sigma\mathbf{t}} = e^{i\mathbf{t}'\mu - \frac{1}{n}\mathbf{t}'\Sigma\mathbf{t}}$$

Que é a função característica de uma $N_p(\mu, n^{-1}\Sigma)$

Distribuição Normal

Notação:

Sejam $\mathbf{X}_1, \cdots, \mathbf{X}_n \sim N_p(\mu, \Sigma)$. Então, a matriz

$$\mathbf{X}_{n\times p} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)' = \begin{pmatrix} \mathbf{X}_1' \\ \vdots \\ \mathbf{X}_n' \end{pmatrix} = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ \vdots & \ldots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{pmatrix}$$

é dita $N_p(\mu, \Sigma)$

Distribuição Normal

Teorema

Se $\mathbf{X}_{n\times p} \sim N_p(\mu, \Sigma)$ e se $\mathbf{Y}_{m\times q} = A\mathbf{X}_{n\times p}B$, $\mathbf{Y}$ tem distribuição Normal sss

- $A\mathbf{1} = \alpha\mathbf{1}$ para algum α ou $B'\mu = 0$ e
- $AA' = \beta\mathbf{I}$ para algum β ou $B'\Sigma B = 0$.

Se ambas as condições são satisfetias,

$$\mathbf{Y} \sim N_q(\alpha B'\mu, \beta B'\Sigma B)$$

Demostração: fora do escopo desta matéria.

Distribuição Wishart

Distribuição Wishart

Definição

Se $\mathbf{W}$ pode ser escrito como $\mathbf{W} = \mathbf{X}'_{n\times p}\mathbf{X}_{n\times p}$ tal que $\mathbf{X}_{n\times p} \sim N_p(0, \Sigma)$, então

$$\mathbf{W} \sim W_p(\Sigma, n),$$

em que $W_p(\Sigma, n)$ denota uma distribuição Wishart (centrada) com parâmetro de escala Σ e graus de liberdade n. Quando $\Sigma = I$, a distribuição é dita estar na sua forma padrão.

- $\mathbb{E}(\mathbf{W}) = n\Sigma$
- $W_1(\sigma^2, n) = \sigma^2\chi^2_n$

Distribuição Wishart

Definição

Seja $\mathbf{W} \sim W_p(\Sigma, n)$ com $\Sigma > 0$ e $n \geq p$. Então a densidade de $\mathbf{W}$ é dada por

$$f_{\mathbf{W}}(\mathbf{w}) = \frac{1}{2^{np/2}\Gamma_p(\frac{n}{2})|\Sigma|^{\frac{n}{2}}}|\mathbf{w}|^{\frac{n-p-1}{2}} e^{-\frac{1}{2}Tr(\Sigma^{-1}\mathbf{w})}, \quad \mathbf{w} > 0.$$

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

- A função caracteristica de $\mathbf{W}$ é dada por

$$\varphi_{\mathbf{W}}(\mathbf{T}) = |\mathbf{I}_p - iM(\mathbf{T})\Sigma|^{-\frac{n}{2}},$$

em que $M(\mathbf{T}) = \sum_{i \leq j} t_{ij}(\mathbf{e}_i\mathbf{e}_j' + \mathbf{e}_j\mathbf{e}_i')$, $\mathbf{e}_i$ é a i-éssima coluna de $\mathbf{I}_p$.

Distribuição Wishart

Propriedades

- Se $\mathbf{W} \sim W_p(\Sigma, n)$ e $\mathbf{B}_{p\times q}$, então
$$\mathbf{B}'\mathbf{W}\mathbf{B} \sim W_q(\mathbf{B}'\Sigma\mathbf{B}, n).$$
- $\Sigma^{-1/2}\mathbf{W}\Sigma^{-1/2} \sim W_p(I, n)$
- Se $\mathbf{W} \sim W_p(I, n)$ e $\mathbf{B}_{p\times q}$ satisfaz $\mathbf{B}'\mathbf{B} = I$, então
$$\mathbf{B}'\mathbf{W}\mathbf{B} \sim W_q(I, n).$$
- Se $\mathbf{W} \sim W_p(\Sigma, n)$ e $\mathbf{a}_{p\times 1}$ tal que $\mathbf{a}'\Sigma\mathbf{a} \neq 0$, então
$$\mathbf{a}'\mathbf{W}\mathbf{a}/\mathbf{a}'\Sigma\mathbf{a} \sim \chi^2_n.$$
- Sejam $\mathbf{W}_1 \sim W_p\Sigma, n_1$ e $\mathbf{W}_2 \sim W_p\Sigma, n_2$ com $\mathbf{W}_1$ e $\mathbf{W}_2$ independentes, então
$$\mathbf{W}_1 + \mathbf{W}_2 \sim W_p\Sigma, n_1 + n_2$$

Distribuição Wishart

Teorema (Cochran, 1934)

Sejam $\mathbf{X}_{n\times p} \sim N_p(0,\Sigma)$ e $\mathbf{C}_{n\times n}$ simétrica, entao

1. $\mathbf{X}'\mathbf{C}\mathbf{X} \sim \sum_{i=1}^{n} \lambda_i W_p(\Sigma, 1)$, em que $\lambda_1, \cdots, \lambda_n$ são os autovalores de $\mathbf{C}$.
2. $\mathbf{X}'\mathbf{C}\mathbf{X}$ tem distribuição Wishart sss $\mathbf{C}^2 = \mathbf{C}$ em cujo caso
$$\mathbf{X}'\mathbf{C}\mathbf{X} \sim W_p(\Sigma, r),$$
com $r = Tr(\mathbf{C}) = rank(\mathbf{C})$.
3. Se $\mathbf{S}$ for a matriz de covariância amostral, então
$$n\mathbf{S} \sim W_p(\Sigma, n-1).$$

Distribution T^2 de Hotelling

Distribution T^2 de Hotelling

Definição

Sejam $\mathbf{X} \sim N_p(0, \mathbf{I})$ e $\mathbf{W} \sim W_p(\mathbf{I}, n)$ com $\mathbf{X}$ e $\mathbf{W}$ independentes, então

$$n\mathbf{X}'\mathbf{W}^{-1}\mathbf{X} \sim T^2_{p,n},$$

em que $T^2_{p,n}$ denota uma distribuição T^2 de Hotelling com parâmetros p e n.

A distribuição T^2 de Hotelling é uma generalização da distribuição `T de Student` e desempenha um papel importante (como veremos ao longo da disciplina) em testes de hipóteses.

Distribution T^2 de Hotelling

Propriedades

- Se $\mathbf{X} \sim N_p(\mu, \Sigma)$ e $\mathbf{W} \sim W_p(\Sigma, n)$ com $\mathbf{X}$ e $\mathbf{W}$ independentes, então

$$n(\mathbf{X} - \mu)'\mathbf{W}^{-1}(\mathbf{X} - \mu) \sim T^2_{p,n}$$

Distribution T^2 de Hotelling

Propriedades

- Se $\mathbf{X} \sim N_p(\mu, \Sigma)$ e $\mathbf{W} \sim W_p(\Sigma, n)$ com $\mathbf{X}$ e $\mathbf{W}$ independentes, então

$$n(\mathbf{X} - \mu)'\mathbf{W}^{-1}(\mathbf{X} - \mu) \sim T^2_{p,n}$$

- Se $\bar{\mathbf{X}}$ e $\mathbf{S}$ são o vetor de médias e a matriz de covariância de uma amostra aleatória de tamanho n de uma $N_p(\mu, \Sigma)$, então

$$(n-1)(\bar{\mathbf{X}} - \mu)'S^{-1}(\bar{\mathbf{X}} - \mu) \sim T^2_{p,n-1}$$

Distribution T^2 de Hotelling

Propriedades

- Se $\mathbf{X} \sim N_p(\mu, \Sigma)$ e $\mathbf{W} \sim W_p(\Sigma, n)$ com $\mathbf{X}$ e $\mathbf{W}$ independentes, então

$$n(\mathbf{X} - \mu)'\mathbf{W}^{-1}(\mathbf{X} - \mu) \sim T^2_{p,n}$$

- Se $\bar{\mathbf{X}}$ e $\mathbf{S}$ são o vetor de médias e a matriz de covariância de uma amostra aleatória de tamanho n de uma $N_p(\mu, \Sigma)$, então

$$(n-1)(\bar{\mathbf{X}} - \mu)'S^{-1}(\bar{\mathbf{X}} - \mu) \sim T^2_{p,n-1}$$

- $T^2_{p,n} = \dfrac{np}{(n-p+1)} F_{p,n-p+1}.$

Distribution T^2 de Hotelling

Propriedades

- Seja $\bar{\mathbf{X}}$ o vetor de medias amostrais e seja $\mathbf{S}$ a matriz de covariância amostral de uma a.a. de tamanho $n \sim N_p(\mu, \Sigma)$. Então

$$\frac{n-p}{p}(\bar{\mathbf{X}} - \mu)'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \mu) \sim F_{p,n-p}$$

- Sejam $\mathbf{X} \sim N_p(0, I)$ e $\mathbf{W} \sim W_p(I, n)$ com $\mathbf{X}$ e $\mathbf{W}$ independentes, então

$$\mathbf{X}'\mathbf{X}\left(1 + \frac{1}{\mathbf{X}\mathbf{W}^{-1}\mathbf{X}}\right) \sim \chi^2_{n+1}$$

e é independente de $\mathbf{X}'\mathbf{W}^{-1}\mathbf{X}$.

Teorema Central do Limite

Teorema Central do Limite

TCL

Sejam $\mathbf{X}_1, \cdots, \mathbf{X}_n$ vetores atealtorios p-dimensionais *iid* com $\mathbb{E}(\mathbf{X}_1) = \mu$ e $\mathbb{V}(\mathbf{X}_1) = \Sigma$. Então,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\mathbf{X}_i - \mu) \equiv \sqrt{n}(\bar{\mathbf{X}} - \mu) \quad \underrightarrow{D} \quad N_p(0, \Sigma)$$

Referências

Referências

- Härdle, W. K., & Simar, L. (2019). Applied Multivariate Statistical Analysis. Fifth Editon. Springer Nature. Capítulo 5.
- Johnson, R. A., & Wichern, D. W. (2007). Applied multivariate statistical analysis. Sixth Edition. Pearson Prentice Hall. Capítulo 4.
- Mardia, K. V., Kent, J. T., & Bibby, J, M. (1979). Multivariate Analysis. Academic Press. Capítulo 3.
- Wilkinson, R. (2022). Multivariate Statistics. Capítulo 7.