

# A Distribuição Normal Multivariada

Prof. Caio Azevedo

# Brevíssima revisão de cálculo de probabilidades

- Como usual, denotaremos por uma letra maiúscula, e.g.  $Y$ , uma variável aleatória (va) e por uma letra minúscula,  $y$ , um valor observado (realização de um experimento aleatório) desta va.
- Um vetor aleatório (vea)  $\mathbf{Y} = (Y_1, \dots, Y_p)$  é uma coleção (arranjo) de variáveis aleatórias.
- As va's que compõem um vea podem apresentar alguma estrutura de dependência e/ou serem de diferentes tipos (discretas, contínuas ou mistas).

- Função densidade de probabilidade ou função de probabilidade:

$$f_{\mathbf{Y}}(\mathbf{y})$$

- Função de distribuição acumulada  $F_{\mathbf{Y}}(\mathbf{y}) = P(Y_1 \leq y_1, \dots, Y_p \leq y_p)$ .

- Vetor de médias:  $\boldsymbol{\mu} = \mathcal{E}(\mathbf{Y}) = \begin{bmatrix} \mathcal{E}(Y_1) \\ \mathcal{E}(Y_2) \\ \vdots \\ \mathcal{E}(Y_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$

- Matriz de covariâncias:  $\boldsymbol{\Sigma} = \text{Cov}(\mathbf{Y}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \dots & \sigma_p^2 \end{bmatrix}$

- Função geradora de momentos:

$$M_{\mathbf{Y}}(\mathbf{t}) = \int \left( \dots \left( \int \left( \int e^{\mathbf{y}'\mathbf{t}} f_{\mathbf{Y}}(\mathbf{y}) dy_1 \right) dy_2 \right) \dots \right) dy_p$$

- Função característica:

$$\Phi_{\mathbf{Y}}(\mathbf{t}) = \int \left( \dots \left( \int \left( \int e^{i\mathbf{y}'\mathbf{t}} f_{\mathbf{Y}}(\mathbf{y}) dy_1 \right) dy_2 \right) \dots \right) dy_p$$

- Sejam  $\mathbf{A}$  e  $\mathbf{B}$  matrizes não aleatórias, então

- $\mathcal{E}(\mathbf{AY}) = \mathbf{A}\mathcal{E}(\mathbf{Y})$ .
- $\text{Cov}(\mathbf{AY}, \mathbf{BX}) = \mathbf{A}\text{Cov}(\mathbf{Y}, \mathbf{X})\mathbf{B}'$ .
- $\text{Cov}(\mathbf{AY}) = \mathbf{A}\text{Cov}(\mathbf{Y})\mathbf{A}'$ .

# Distribuição Normal multivariada

- Dizemos que  $\mathbf{Y} = (Y_1, \dots, Y_p) \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  se sua fdp é dada por

$$f_{\mathbf{Y}}(\mathbf{y}) = |\boldsymbol{\Sigma}|^{-1/2} (2\pi)^{-p/2} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu}) \right\} \mathbb{1}_{\mathcal{R}^p}(\mathbf{y})$$

- $\boldsymbol{\mu}$  é o vetor de médias e  $\boldsymbol{\Sigma}$  é a matriz de covariâncias.
- $M_{\mathbf{Y}}(\mathbf{t}) = \exp \left\{ \boldsymbol{\mu}' \mathbf{t} + \frac{1}{2} \mathbf{t}' \boldsymbol{\Sigma} \mathbf{t} \right\}$

# Propriedades

- Fechada sob marginalização:  $Y_i \sim N(\mu_i, \sigma_i^2)$ .
- $Y_i \perp Y_j, \forall i \neq j \Leftrightarrow \sigma_{ij} = 0$ .
- Se  $\mathbf{A}_{(q \times p)}$  for uma matriz não aleatória, então  $\mathbf{V} = \mathbf{A}\mathbf{Y} \sim N_q(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$ .
- Se  $\mathbf{A}_{(p \times p)}$  for uma matriz não aleatória, simétrica e idempotente de rank = p,  $\boldsymbol{\mu} = \mathbf{0}$  e  $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_{(p \times p)}$ , então  $V = \frac{1}{\sigma^2} \mathbf{Y}'\mathbf{A}\mathbf{Y} \sim \chi_r^2, r = \text{tr}(\mathbf{A})$ . Em particular, se  $\mathbf{A} = \mathbf{I}$ , então  $\frac{1}{\sigma^2} \mathbf{Y}'\mathbf{Y} \sim \chi_p^2$ .
- Se  $\mathbf{A}_{(p \times p)}$  for uma matriz não aleatória, então  $\mathcal{E}(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = \text{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$ .

# Derivadas matriciais úteis

- Sejam  $\mathbf{A}_{(m \times n)}$  e  $\mathbf{x}_{(n \times 1)}$ , então

$$\begin{aligned}\frac{\partial \mathbf{Ax}}{\partial \mathbf{x}} &= \mathbf{A}' \\ \frac{\partial \mathbf{x}'\mathbf{A}}{\partial \mathbf{x}} &= \mathbf{A} \\ \frac{\partial \mathbf{x}'\mathbf{x}}{\partial \mathbf{x}} &= 2\mathbf{x} \\ \frac{\partial \mathbf{x}'\mathbf{Ax}}{\partial \mathbf{x}} &= (\mathbf{A} + \mathbf{A}')\mathbf{x}\end{aligned}\tag{1}$$

## Dem. de que $f_{\mathbf{Y}}(\cdot)$ é uma fdp

- Queremos demonstrar que

$$I = \int_{\mathcal{R}^p} f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{R}} \int_{\mathcal{R}} \dots \int_{\mathcal{R}} f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} = 1$$

Note que, se  $\mathbf{\Sigma} = \mathbf{\Psi}\mathbf{\Psi}'$  (decomposição de Cholesky), temos que:

$$\begin{aligned} I &= \int_{\mathcal{R}^p} |\mathbf{\Sigma}|^{-1/2} (2\pi)^{-p/2} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \boldsymbol{\mu})' (\mathbf{\Psi}\mathbf{\Psi}')^{-1} (\mathbf{y} - \boldsymbol{\mu}) \right\} d\mathbf{y} \\ &= \int_{\mathcal{R}^p} |\mathbf{\Sigma}|^{-1/2} (2\pi)^{-p/2} \exp \left\{ -\frac{1}{2} [\mathbf{\Psi}^{-1} (\mathbf{y} - \boldsymbol{\mu})]' (\mathbf{\Psi})^{-1} (\mathbf{y} - \boldsymbol{\mu}) \right\} d\mathbf{y} \end{aligned}$$

considere a transformação  $\mathbf{z} = \mathbf{\Psi}^{-1} (\mathbf{y} - \boldsymbol{\mu}) \Leftrightarrow \mathbf{y} = \mathbf{\Psi}\mathbf{z} + \boldsymbol{\mu}$ . Assim, temos que  $d\mathbf{y} = |\mathbf{\Psi}'| d\mathbf{z}$  e  $|J| = |\mathbf{\Psi}'|$ .



- Além disso,

$$|\Psi'| = |\Psi'|^{1/2} |\Psi'|^{1/2} = |\Psi|^{1/2} |\Psi'|^{1/2} = |\Psi\Psi'|^{1/2} = |\Sigma|^{1/2}$$

- Assim,

$$\begin{aligned} I &= \int_{\mathcal{R}^p} |\Sigma|^{-1/2} |\Sigma|^{1/2} (2\pi)^{-p/2} \exp\left\{-\frac{1}{2}\mathbf{z}'\mathbf{z}\right\} dz \\ &= \prod_{i=1}^p \underbrace{\int_{\mathcal{R}} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z_i^2\right\} dz_i}_1 = 1 \end{aligned}$$

# Obtenção das Marginais

- Note que, para um dado  $j$ ,  $M_{Y_j}(t_j) = M_{\mathbf{Y}}(\mathbf{t}^*)$ , em que

$$\mathbf{t}^* = [0 \ 0 \dots \underbrace{1}_{\text{posição } j} \dots 0 \ 0].$$

- Logo, temos que  $M_{Y_j}(t_j) = \exp \left\{ \mu_j t_j + \frac{\sigma_j^2 t_j^2}{2} \right\}$ .
- A fgm acima corresponde à fgm de uma va com distribuição  $N(\mu_j, \sigma_j^2)$ .

# Distribuições condicionais

- Seja  $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2)$ ,  $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$  e

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

em que  $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}'_{12}$

- Então  $\mathbf{Y}_1 | \mathbf{Y}_2 = \mathbf{y}_2 \sim N(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}})$ , em que

$$\bar{\boldsymbol{\mu}} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{y}_2 - \boldsymbol{\mu}_2); \bar{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21}$$

- Estimadores de máxima verossimilhança (dada uma amostra aleatória)  $\hat{\boldsymbol{\mu}} = \bar{\mathbf{Y}} = \frac{1}{n} \sum_{i=1}^n \mathbf{Y}_i$  e  $\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^n (\mathbf{Y}_i - \bar{\mathbf{Y}})' (\mathbf{Y}_i - \bar{\mathbf{Y}})$ .