

# Distribuições multivariadas

Tiago M. Magalhães

Departamento de Estatística - ICE-UFJF

Juiz de Fora, 15 de outubro de 2025



# Roteiro

1 Distribuição multinomial

2 Distribuição normal multivariada

3 Distribuição Wishart

4 Distribuição Dirichlet

5 Referências bibliográficas



# Roteiro

1 Distribuição multinomial

2 Distribuição normal multivariada

3 Distribuição Wishart

4 Distribuição Dirichlet

5 Referências bibliográficas



# Experimento multinomial

Um experimento multinomial tem **3 características**:



# Experimento multinomial

Um experimento multinomial tem **3 características**:

- ① uma realização com  $n$  **repetições independentes**;



# Experimento multinomial

Um experimento multinomial tem **3 características**:

- ① uma realização com  $n$  **repetições independentes**;
- ② cada repetição admite  $k$  resultados:  $A_1, A_2, \dots, A_k$ ,



# Experimento multinomial

Um experimento multinomial tem **3 características**:

- ① uma realização com  $n$  **repetições independentes**;
- ② cada repetição admite  $k$  resultados:  $A_1, A_2, \dots, A_k$ , mutuamente exclusivos;



# Experimento multinomial

Um experimento multinomial tem **3 características**:

- ① uma realização com  $n$  **repetições independentes**;
- ② cada repetição admite  $k$  resultados:  $A_1, A_2, \dots, A_k$ , mutuamente exclusivos;
- ③ a probabilidade,  $p_i = \mathbb{P}(A_i)$ , de ocorrência de cada possível resultado  $A_i$ ,  
 $i = 1, 2, \dots, k$ ,



# Experimento multinomial

Um experimento multinomial tem **3 características**:

- ① uma realização com  $n$  **repetições independentes**;
- ② cada repetição admite  $k$  resultados:  $A_1, A_2, \dots, A_k$ , mutuamente exclusivos;
- ③ a probabilidade,  $p_i = \mathbb{P}(A_i)$ , de ocorrência de cada possível resultado  $A_i$ ,  $i = 1, 2, \dots, k$ , **não se altera** em cada repetição.



# Experimento multinomial

Um experimento multinomial tem **3 características**:

- ① uma realização com  $n$  **repetições independentes**;
- ② cada repetição admite  $k$  resultados:  $A_1, A_2, \dots, A_k$ , mutuamente exclusivos;
- ③ a probabilidade,  $p_i = \mathbb{P}(A_i)$ , de ocorrência de cada possível resultado  $A_i$ ,  $i = 1, 2, \dots, k$ , **não se altera** em cada repetição. Note que  $\sum_{i=1}^k p_i = 1$ .



# Experimento multinomial

Um experimento multinomial tem **3 características**:

- ① uma realização com  $n$  **repetições independentes**;
- ② cada repetição admite  $k$  resultados:  $A_1, A_2, \dots, A_k$ , mutuamente exclusivos;
- ③ a probabilidade,  $p_i = \mathbb{P}(A_i)$ , de ocorrência de cada possível resultado  $A_i$ ,  $i = 1, 2, \dots, k$ , **não se altera** em cada repetição. Note que  $\sum_{i=1}^k p_i = 1$ .



# Distribuição multinomial

Seja a variável aleatória discreta (VAD) multidimensional

$$\mathbf{X} = (X_1, X_2, \dots, X_k)^\top,$$

em que  $X_i$ : é o número de vezes que  $A_i$  ocorre



# Distribuição multinomial

Seja a variável aleatória discreta (VAD) multidimensional

$$\mathbf{X} = (X_1, X_2, \dots, X_k)^\top,$$

em que  $X_i$ : é o número de vezes que  $A_i$  ocorre nas  $n$  repetições de um experimento multinomial,



# Distribuição multinomial

Seja a variável aleatória discreta (VAD) multidimensional

$$\mathbf{X} = (X_1, X_2, \dots, X_k)^\top,$$

em que  $X_i$ : é o número de vezes que  $A_i$  ocorre nas  $n$  repetições de um experimento multinomial, com probabilidade de  $A_i$  ocorrer igual a  $p_i$ ,  $i = 1, 2, \dots, k$ .



# Distribuição multinomial

Seja a variável aleatória discreta (VAD) multidimensional

$$\mathbf{X} = (X_1, X_2, \dots, X_k)^\top,$$

em que  $X_i$ : é o número de vezes que  $A_i$  ocorre nas  $n$  repetições de um experimento multinomial, com probabilidade de  $A_i$  ocorrer igual a  $p_i$ ,  $i = 1, 2, \dots, k$ .



# Distribuição multinomial

## Observação

Os  $X_i$  não são VAD independentes,



# Distribuição multinomial

## Observação

Os  $X_i$  **não são** VAD independentes, pois  $\sum_{i=1}^k X_i = n$ .



# Distribuição multinomial

## Observação

Os  $X_i$  **não são** VAD independentes, pois  $\sum_{i=1}^k X_i = n$ . Em consequência, logo que os valores quaisquer  $(k-1)$  dessas VAD sejam conhecidos, o valor da  $k$ -ésima ficará determinado.



# Distribuição multinomial

## Observação

Os  $X_i$  **não são** VAD independentes, pois  $\sum_{i=1}^k X_i = n$ . Em consequência, logo que os valores quaisquer  $(k - 1)$  dessas VAD sejam conhecidos, o valor da  $k$ -ésima ficará determinado.



# Distribuição multinomial

Então, a VAD  $\mathbf{X}$  tem distribuição multinomial, com parâmetros  $n$  e  $\mathbf{p} = (p_1, p_2, \dots, p_k)^\top$ ,



# Distribuição multinomial

Então, a VAD  $\mathbf{X}$  tem distribuição multinomial, com parâmetros  $n$  e  $\mathbf{p} = (p_1, p_2, \dots, p_k)^\top$ , de forma alternativa,

$\mathbf{X} \sim \text{multinomial}(n, \mathbf{p})$  ou  $\mathbf{X} \sim \text{multinom}(n, \mathbf{p})$ .



# Distribuição multinomial

Então, a VAD  $\mathbf{X}$  tem distribuição multinomial, com parâmetros  $n$  e  $\mathbf{p} = (p_1, p_2, \dots, p_k)^\top$ , de forma alternativa,

$\mathbf{X} \sim \text{multinomial}(n, \mathbf{p})$  ou  $\mathbf{X} \sim \text{multinom}(n, \mathbf{p})$ .



# Distribuição multinomial

A função de probabilidades de uma variável aleatória discreta multidimensional  $\mathbf{X} \sim \text{multinomial}(n, \mathbf{p})$  é dada por:

$$p(x_1, x_2, \dots, x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} \mathbb{I}_{\{n\}}(x_1 + x_2 + \dots + x_k),$$



# Distribuição multinomial

A função de probabilidades de uma variável aleatória discreta multidimensional  $\mathbf{X} \sim \text{multinomial}(n, \mathbf{p})$  é dada por:

$$p(x_1, x_2, \dots, x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} \mathbb{I}_{\{n\}}(x_1 + x_2 + \dots + x_k),$$

em que  $p_i \in (0, 1)$ ,  $i = 1, 2, \dots, k$ .



# Distribuição multinomial

A função de probabilidades de uma variável aleatória discreta multidimensional  $\mathbf{X} \sim \text{multinomial}(n, \mathbf{p})$  é dada por:

$$p(x_1, x_2, \dots, x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} \mathbb{I}_{\{n\}}(x_1 + x_2 + \dots + x_k),$$

em que  $p_i \in (0, 1)$ ,  $i = 1, 2, \dots, k$ .



# Distribuição multinomial

A esperança e variância de uma VAD multidimensional  $\mathbf{X}$  multinomial são dadas, respectivamente, por  $\mathbb{E}(\mathbf{X}) = [\mathbb{E}(X_1), \mathbb{E}(X_2), \dots, \mathbb{E}(X_k)]^\top$  e

$$\text{Var}(\mathbf{X}) = \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_k) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \cdots & \text{Cov}(X_2, X_k) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_k, X_1) & \text{Cov}(X_k, X_2) & \cdots & \text{Var}(X_k) \end{bmatrix},$$



# Distribuição multinomial

A esperança e variância de uma VAD multidimensional  $\mathbf{X}$  multinomial são dadas, respectivamente, por  $\mathbb{E}(\mathbf{X}) = [\mathbb{E}(X_1), \mathbb{E}(X_2), \dots, \mathbb{E}(X_k)]^\top$  e

$$\text{Var}(\mathbf{X}) = \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_k) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \cdots & \text{Cov}(X_2, X_k) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_k, X_1) & \text{Cov}(X_k, X_2) & \cdots & \text{Var}(X_k) \end{bmatrix},$$



# Distribuição multinomial

em que

- $\mathbb{E}(X_i) = np_i,$



# Distribuição multinomial

em que

- $\mathbb{E}(X_i) = np_i,$
- $\text{Var}(X_i) = np_i(1 - p_i),$



# Distribuição multinomial

em que

- $\mathbb{E}(X_i) = np_i,$
- $\text{Var}(X_i) = np_i(1 - p_i),$
- $\text{Cov}(X_i, X_j) = -np_i p_j, \text{ para } i \neq j,$



# Distribuição multinomial

em que

- $\mathbb{E}(X_i) = np_i,$
- $\text{Var}(X_i) = np_i(1 - p_i),$
- $\text{Cov}(X_i, X_j) = -np_i p_j, \text{ para } i \neq j,$

com  $i, j = 1, 2, \dots, k.$



# Distribuição multinomial

em que

- $\mathbb{E}(X_i) = np_i,$
- $\text{Var}(X_i) = np_i(1 - p_i),$
- $\text{Cov}(X_i, X_j) = -np_i p_j, \text{ para } i \neq j,$

com  $i, j = 1, 2, \dots, k.$  De maneira geral, isto é, na forma matricial, nós podemos escrever



# Distribuição multinomial

em que

- $\mathbb{E}(X_i) = np_i,$
- $\text{Var}(X_i) = np_i(1 - p_i),$
- $\text{Cov}(X_i, X_j) = -np_i p_j, \text{ para } i \neq j,$

com  $i, j = 1, 2, \dots, k$ . De maneira geral, isto é, na forma matricial, nós podemos escrever  $\mathbb{E}(\mathbf{X}) = \mathbf{np}$  e  $\text{Var}(\mathbf{X}) = n \left[ \text{diag}(\mathbf{p}) - \mathbf{pp}^\top \right],$



# Distribuição multinomial

em que

- $\mathbb{E}(X_i) = np_i,$
- $\text{Var}(X_i) = np_i(1 - p_i),$
- $\text{Cov}(X_i, X_j) = -np_i p_j, \text{ para } i \neq j,$

com  $i, j = 1, 2, \dots, k$ . De maneira geral, isto é, na forma matricial, nós podemos escrever  $\mathbb{E}(\mathbf{X}) = n\mathbf{p}$  e  $\text{Var}(\mathbf{X}) = n \left[ \text{diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}^T \right]$ , sendo  $\text{diag}(\mathbf{p})$  uma matriz diagonal, com a diagonal principal igual a  $\mathbf{p}$ .



# Distribuição multinomial

em que

- $\mathbb{E}(X_i) = np_i,$
- $\text{Var}(X_i) = np_i(1 - p_i),$
- $\text{Cov}(X_i, X_j) = -np_i p_j, \text{ para } i \neq j,$

com  $i, j = 1, 2, \dots, k$ . De maneira geral, isto é, na forma matricial, nós podemos escrever  $\mathbb{E}(\mathbf{X}) = n\mathbf{p}$  e  $\text{Var}(\mathbf{X}) = n \left[ \text{diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}^\top \right]$ , sendo  $\text{diag}(\mathbf{p})$  uma matriz diagonal, com a diagonal principal igual a  $\mathbf{p}$ .



# Distribuição multinomial

Seja  $\mathbf{t} = (t_1, t_2, \dots, t_k)^\top \in \mathbb{R}^k$ , a função geradoras de momentos de  $\mathbf{X}$  é dada por



# Distribuição multinomial

Seja  $\mathbf{t} = (t_1, t_2, \dots, t_k)^\top \in \mathbb{R}^k$ , a função geradoras de momentos de  $\mathbf{X}$  é dada por

$$M(\mathbf{t}) = \mathbb{E}(e^{\mathbf{t}^\top \mathbf{X}}) = \mathbb{E}\left(e^{t_1 X_1 + t_2 X_2 + \dots + t_k X_k}\right) = \left(\sum_{i=1}^k p_i e^{t_i}\right)^n.$$



# Distribuição multinomial

Seja  $\mathbf{t} = (t_1, t_2, \dots, t_k)^\top \in \mathbb{R}^k$ , a função geradoras de momentos de  $\mathbf{X}$  é dada por

$$M(\mathbf{t}) = \mathbb{E}(e^{\mathbf{t}^\top \mathbf{X}}) = \mathbb{E}\left(e^{t_1 X_1 + t_2 X_2 + \dots + t_k X_k}\right) = \left(\sum_{i=1}^k p_i e^{t_i}\right)^n.$$



# Distribuição multinomial

## Exemplos:

- Em um lançamento de um dado 10 vezes,



# Distribuição multinomial

## Exemplos:

- Em um lançamento de um dado 10 vezes, o interesse é contar o número de ocorrências de cada uma das 6 faces;



# Distribuição multinomial

Exemplos:

- Em um lançamento de um dado 10 vezes, o interesse é contar o número de ocorrências de cada uma das 6 faces;
- Em um banco de dados com 1.500 clientes,



# Distribuição multinomial

## Exemplos:

- Em um lançamento de um dado 10 vezes, o interesse é contar o número de ocorrências de cada uma das 6 faces;
- Em um banco de dados com 1.500 clientes, o interesse é contar quantos destes teriam um pedido de empréstimo



# Distribuição multinomial

Exemplos:

- Em um lançamento de um dado 10 vezes, o interesse é contar o número de ocorrências de cada uma das 6 faces;
- Em um banco de dados com 1.500 clientes, o interesse é contar quantos destes teriam um pedido de empréstimo aprovado na totalidade,



# Distribuição multinomial

Exemplos:

- Em um lançamento de um dado 10 vezes, o interesse é contar o número de ocorrências de cada uma das 6 faces;
- Em um banco de dados com 1.500 clientes, o interesse é contar quantos destes teriam um pedido de empréstimo aprovado na totalidade, 50% do valor pedido



# Distribuição multinomial

Exemplos:

- Em um lançamento de um dado 10 vezes, o interesse é contar o número de ocorrências de cada uma das 6 faces;
- Em um banco de dados com 1.500 clientes, o interesse é contar quantos destes teriam um pedido de empréstimo aprovado na totalidade, 50% do valor pedido e recusado;



# Distribuição multinomial

Exemplos:

- Em um lançamento de um dado 10 vezes, o interesse é contar o número de ocorrências de cada uma das 6 faces;
- Em um banco de dados com 1.500 clientes, o interesse é contar quantos destes teriam um pedido de empréstimo aprovado na totalidade, 50% do valor pedido e recusado;
- Equilíbrio de Hardy-Weinberg (Hardy, 1908; Weinberg, 1908).



# Distribuição multinomial

Exemplos:

- Em um lançamento de um dado 10 vezes, o interesse é contar o número de ocorrências de cada uma das 6 faces;
- Em um banco de dados com 1.500 clientes, o interesse é contar quantos destes teriam um pedido de empréstimo aprovado na totalidade, 50% do valor pedido e recusado;
- Equilíbrio de Hardy-Weinberg (Hardy, 1908; Weinberg, 1908).



# Roteiro

1 Distribuição multinomial

2 Distribuição normal multivariada

3 Distribuição Wishart

4 Distribuição Dirichlet

5 Referências bibliográficas



# Distribuição normal bivariada

Seja  $(X_1, X_2)^\top$  uma variável aleatória contínua (VAC) bidimensional, tomando todos os valores no plano euclidiano. Nós diremos que  $(X_1, X_2)^\top$  tem distribuição normal bidimensional se sua função densidade de probabilidade (FDP) conjunta for dada pela seguinte expressão



# Distribuição normal bivariada

Seja  $(X_1, X_2)^\top$  uma variável aleatória contínua (VAC) bidimensional, tomando todos os valores no plano euclidiano. Nós diremos que  $(X_1, X_2)^\top$  tem distribuição normal bidimensional se sua função densidade de probabilidade (FDP) conjunta for dada pela seguinte expressão

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} & (1) \\ &\times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\} \\ &\times \mathbb{I}_{\mathbb{R}}(x_1)\mathbb{I}_{\mathbb{R}}(x_2), \end{aligned}$$



# Distribuição normal bivariada

Seja  $(X_1, X_2)^\top$  uma variável aleatória contínua (VAC) bidimensional, tomando todos os valores no plano euclidiano. Nós diremos que  $(X_1, X_2)^\top$  tem distribuição normal bidimensional se sua função densidade de probabilidade (FDP) conjunta for dada pela seguinte expressão

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \left( \frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \left( \frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right] \right\} \times \mathbb{I}_{\mathbb{R}}(x_1) \mathbb{I}_{\mathbb{R}}(x_2), \quad (1)$$

em que  $\mu_1 \in \mathbb{R}$ ,  $\mu_2 \in \mathbb{R}$ ,  $\sigma_1^2 \in \mathbb{R}_+$ ,  $\sigma_2^2 \in \mathbb{R}_+$  e  $\rho \in (-1, 1)$ .



# Distribuição normal bivariada

Seja  $(X_1, X_2)^\top$  uma variável aleatória contínua (VAC) bidimensional, tomando todos os valores no plano euclidiano. Nós diremos que  $(X_1, X_2)^\top$  tem distribuição normal bidimensional se sua função densidade de probabilidade (FDP) conjunta for dada pela seguinte expressão

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \left( \frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \left( \frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right] \right\} \times \mathbb{I}_{\mathbb{R}}(x_1) \mathbb{I}_{\mathbb{R}}(x_2), \quad (1)$$

em que  $\mu_1 \in \mathbb{R}$ ,  $\mu_2 \in \mathbb{R}$ ,  $\sigma_1^2 \in \mathbb{R}_+$ ,  $\sigma_2^2 \in \mathbb{R}_+$  e  $\rho \in (-1, 1)$ .



# Distribuição normal bivariada

## Propriedades.

- O parâmetro  $\rho$  será o coeficiente de correlação entre  $X_1$  e  $X_2$ ;



# Distribuição normal bivariada

## Propriedades.

- O parâmetro  $\rho$  será o coeficiente de correlação entre  $X_1$  e  $X_2$ ;
- As distribuições marginais são, respectivamente,



# Distribuição normal bivariada

## Propriedades.

- O parâmetro  $\rho$  será o coeficiente de correlação entre  $X_1$  e  $X_2$ ;
- As distribuições marginais são, respectivamente,

$$X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2);$$



# Distribuição normal bivariada

## Propriedades.

- O parâmetro  $\rho$  será o coeficiente de correlação entre  $X_1$  e  $X_2$ ;
- As distribuições marginais são, respectivamente,

$$X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2);$$

- As distribuições condicionais são, respectivamente,



# Distribuição normal bivariada

## Propriedades.

- O parâmetro  $\rho$  será o coeficiente de correlação entre  $X_1$  e  $X_2$ ;
- As distribuições marginais são, respectivamente,

$$X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2);$$

- As distribuições condicionais são, respectivamente,

$$X_1 | X_2 = x_2 \sim N \left[ \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (x_2 - \mu_2), \sigma_1^2 (1 - \rho^2) \right],$$

$$X_2 | X_1 = x_1 \sim N \left[ \mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x_1 - \mu_1), \sigma_2^2 (1 - \rho^2) \right].$$



# Distribuição normal bivariada

## Propriedades.

- O parâmetro  $\rho$  será o coeficiente de correlação entre  $X_1$  e  $X_2$ ;
- As distribuições marginais são, respectivamente,

$$X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2);$$

- As distribuições condicionais são, respectivamente,

$$X_1 | X_2 = x_2 \sim N \left[ \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (x_2 - \mu_2), \sigma_1^2 (1 - \rho^2) \right],$$

$$X_2 | X_1 = x_1 \sim N \left[ \mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x_1 - \mu_1), \sigma_2^2 (1 - \rho^2) \right].$$



# Distribuição normal bivariada

## Observações.

- É possível ter-se uma FDP conjunta que **não seja** normal bidimensional,



# Distribuição normal bivariada

## Observações.

- É possível ter-se uma FDP conjunta que **não seja** normal bidimensional, e no entanto, as FDP marginais de  $X_1$  e  $X_2$  serem normais unidimensionais;



# Distribuição normal bivariada

## Observações.

- É possível ter-se uma FDP conjunta que **não seja** normal bidimensional, e no entanto, as FDP marginais de  $X_1$  e  $X_2$  serem normais unidimensionais;
- Se  $\rho = 0$ ,



# Distribuição normal bivariada

## Observações.

- É possível ter-se uma FDP conjunta que **não seja** normal bidimensional, e no entanto, as FDP marginais de  $X_1$  e  $X_2$  serem normais unidimensionais;
- Se  $\rho = 0$ , a FDP conjunta de  $(X_1, X_2)^\top$  poderá ser fatorada e, consequentemente,  $X_1$  e  $X_2$  serão independentes.



# Distribuição normal bivariada

## Observações.

- É possível ter-se uma FDP conjunta que **não seja** normal bidimensional, e no entanto, as FDP marginais de  $X_1$  e  $X_2$  serem normais unidimensionais;
- Se  $\rho = 0$ , a FDP conjunta de  $(X_1, X_2)^\top$  poderá ser fatorada e, consequentemente,  $X_1$  e  $X_2$  serão independentes. Assim, nós verificamos que no caso de uma distribuição normal bidimensional, correlação zero e independência são equivalentes.



# Distribuição normal bivariada

## Observações.

- É possível ter-se uma FDP conjunta que **não seja** normal bidimensional, e no entanto, as FDP marginais de  $X_1$  e  $X_2$  serem normais unidimensionais;
- Se  $\rho = 0$ , a FDP conjunta de  $(X_1, X_2)^\top$  poderá ser fatorada e, consequentemente,  $X_1$  e  $X_2$  serão independentes. Assim, nós verificamos que no caso de uma distribuição normal bidimensional, correlação zero e independência são equivalentes.



# Distribuição normal multivariada

Seja  $\mathbf{X} = (X_1, X_2, \dots, X_k)^\top$  um vetor aleatório contínuo. Nós diremos que  $\mathbf{X}$  tem distribuição normal multidimensional (multivariada) se sua FDP conjunta for dada pela seguinte expressão



# Distribuição normal multivariada

Seja  $\mathbf{X} = (X_1, X_2, \dots, X_k)^\top$  um vetor aleatório contínuo. Nós diremos que  $\mathbf{X}$  tem distribuição normal multidimensional (multivariada) se sua FDP conjunta for dada pela seguinte expressão

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \mathbb{I}_{\mathbb{R}^k}(\mathbf{x}), \quad (2)$$



# Distribuição normal multivariada

Seja  $\mathbf{X} = (X_1, X_2, \dots, X_k)^\top$  um vetor aleatório contínuo. Nós diremos que  $\mathbf{X}$  tem distribuição normal multidimensional (multivariada) se sua FDP conjunta for dada pela seguinte expressão

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \mathbb{I}_{\mathbb{R}^k}(\mathbf{x}), \quad (2)$$



# Distribuição normal multivariada

em que  $\mu = (\mu_1, \mu_2, \dots, \mu_k)^\top \in \mathbb{R}^k$  é o vetor de médias,  $\Sigma \in \mathbb{R}^{k \times k}$  é a matriz de covariâncias (uma matriz positiva semi-definida), de dimensão  $k \times k$



# Distribuição normal multivariada

em que  $\mu = (\mu_1, \mu_2, \dots, \mu_k)^\top \in \mathbb{R}^k$  é o vetor de médias,  $\Sigma \in \mathbb{R}^{k \times k}$  é a matriz de covariâncias (uma matriz positiva semi-definida), de dimensão  $k \times k$  e  $|\Sigma|$  é o determinante de  $\Sigma$ .



# Distribuição normal multivariada

em que  $\mu = (\mu_1, \mu_2, \dots, \mu_k)^\top \in \mathbb{R}^k$  é o vetor de médias,  $\Sigma \in \mathbb{R}^{k \times k}$  é a matriz de covariâncias (uma matriz positiva semi-definida), de dimensão  $k \times k$  e  $|\Sigma|$  é o determinante de  $\Sigma$ . De forma alternativa, nós denotamos

$$\mathbf{X} \sim N_k(\mu, \Sigma). \quad (3)$$



# Distribuição normal multivariada

em que  $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_k)^\top \in \mathbb{R}^k$  é o vetor de médias,  $\boldsymbol{\Sigma} \in \mathbb{R}^{k \times k}$  é a matriz de covariâncias (uma matriz positiva semi-definida), de dimensão  $k \times k$  e  $|\boldsymbol{\Sigma}|$  é o determinante de  $\boldsymbol{\Sigma}$ . De forma alternativa, nós denotamos

$$\mathbf{X} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \quad (3)$$



# Distribuição normal multivariada

Seja  $\mathbf{t} = (t_1, t_2, \dots, t_k)^\top \in \mathbb{R}^k$ , a função geradoras de momentos de  $\mathbf{X}$  é dada por



# Distribuição normal multivariada

Seja  $\mathbf{t} = (t_1, t_2, \dots, t_k)^\top \in \mathbb{R}^k$ , a função geradoras de momentos de  $\mathbf{X}$  é dada por

$$M(\mathbf{t}) = \mathbb{E}(e^{\mathbf{t}^\top \mathbf{X}}) = \exp \left\{ \boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2} \mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t} \right\}.$$



# Distribuição normal multivariada

Seja  $\mathbf{t} = (t_1, t_2, \dots, t_k)^\top \in \mathbb{R}^k$ , a função geradoras de momentos de  $\mathbf{X}$  é dada por

$$M(\mathbf{t}) = \mathbb{E}(e^{\mathbf{t}^\top \mathbf{X}}) = \exp \left\{ \boldsymbol{\mu}^\top \mathbf{t} + \frac{1}{2} \mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t} \right\}.$$



# Distribuição normal multivariada

## Observação

A quantidade  $\sqrt{(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})}$  é conhecida como a **distância de Mahalanobis** (Mahalanobis, 1936).



# Distribuição normal multivariada

## Observação

A quantidade  $\sqrt{(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})}$  é conhecida como a **distância de Mahalanobis** (Mahalanobis, 1936).



# Distribuição normal multivariada

A normal bivariada, dada em (1), pode ser escrita na forma apresentada em (2), basta nós definirmos

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$



# Distribuição normal multivariada

A normal bivariada, dada em (1), pode ser escrita na forma apresentada em (2), basta nós definirmos

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

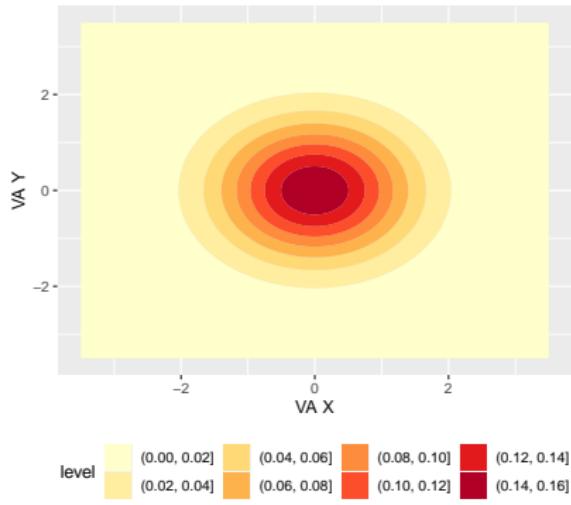


# Distribuição normal multivariada

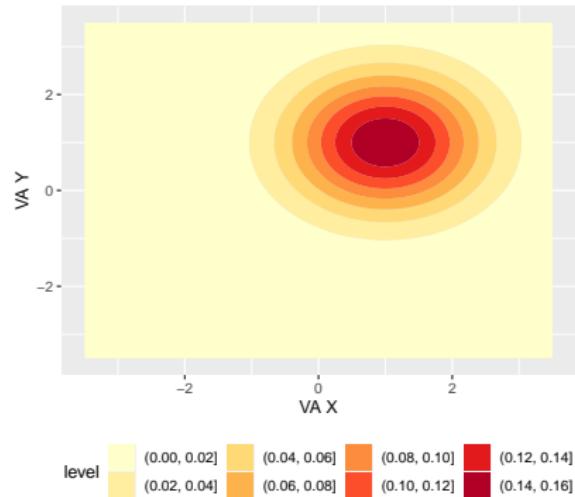
Exemplos:

- ①  $\mu_1 = 0, \mu_2 = 0, \sigma_1^2 = 1, \sigma_2^2 = 1$  e  $\rho = 0$ ;
- ②  $\mu_1 = 1, \mu_2 = 1, \sigma_1^2 = 1, \sigma_2^2 = 1$  e  $\rho = 0$ ;
- ③  $\mu_1 = 0, \mu_2 = 0, \sigma_1^2 = 2, \sigma_2^2 = 2$  e  $\rho = 0$ ;
- ④  $\mu_1 = 0, \mu_2 = 0, \sigma_1^2 = 2, \sigma_2^2 = 2$  e  $\rho = 0,45$ .



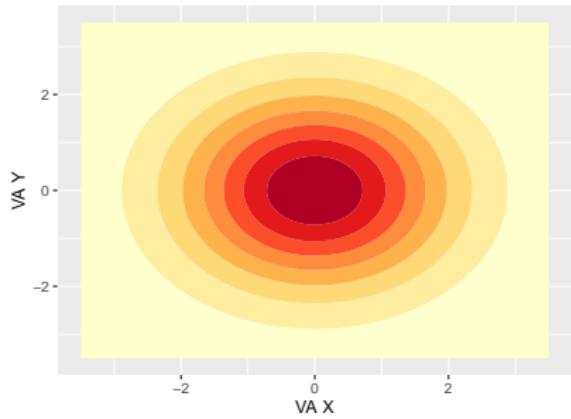


(a) Exemplo 1.

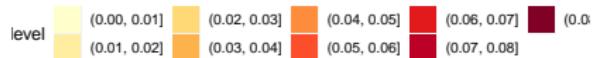
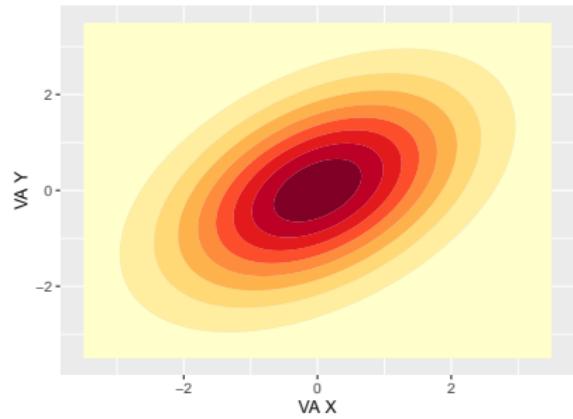


(b) Exemplo 2.

Figura 1: Gráfico de contornos de normais bidimensionais.



(a) Exemplo 3.



(b) Exemplo 4.

Figura 2: Gráfico de contornos de normais bidimensionais.

# Distribuição normal multivariada

Sejam  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)^\top$ , tal que  $Z_i \sim N(0, 1)$ ,  $i = 1, 2, \dots, n$ ,  $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_m)^\top \in \mathbb{R}^m$  e uma matriz  $\mathbf{A}$ , de dimensão  $m \times n$ . A condição para  $\mathbf{X} = (X_1, X_2, \dots, X_m)^\top$  ter distribuição normal multivariada é a seguinte:



# Distribuição normal multivariada

Sejam  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)^\top$ , tal que  $Z_i \sim N(0, 1)$ ,  $i = 1, 2, \dots, n$ ,  $\mu = (\mu_1, \mu_2, \dots, \mu_m)^\top \in \mathbb{R}^m$  e uma matriz  $\mathbf{A}$ , de dimensão  $m \times n$ . A condição para  $\mathbf{X} = (X_1, X_2, \dots, X_m)^\top$  ter distribuição normal multivariada é a seguinte:

$$\mathbf{X} \sim N_m(\mu, \Sigma) \Leftrightarrow \mathbf{A}\mathbf{Z} + \mu, \quad (4)$$

em que  $\Sigma = \mathbf{A}\mathbf{A}^\top$ .



# Distribuição normal multivariada

Sejam  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)^\top$ , tal que  $Z_i \sim N(0, 1)$ ,  $i = 1, 2, \dots, n$ ,  $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_m)^\top \in \mathbb{R}^m$  e uma matriz  $\mathbf{A}$ , de dimensão  $m \times n$ . A condição para  $\mathbf{X} = (X_1, X_2, \dots, X_m)^\top$  ter distribuição normal multivariada é a seguinte:

$$\mathbf{X} \sim N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Leftrightarrow \mathbf{A}\mathbf{Z} + \boldsymbol{\mu}, \quad (4)$$

em que  $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^\top$ . A equação (4) expressa em sistema lineares pode ser encontrada em Ross (2010, p. 432).



# Distribuição normal multivariada

Sejam  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)^\top$ , tal que  $Z_i \sim N(0, 1)$ ,  $i = 1, 2, \dots, n$ ,  $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_m)^\top \in \mathbb{R}^m$  e uma matriz  $\mathbf{A}$ , de dimensão  $m \times n$ . A condição para  $\mathbf{X} = (X_1, X_2, \dots, X_m)^\top$  ter distribuição normal multivariada é a seguinte:

$$\mathbf{X} \sim N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Leftrightarrow \mathbf{A}\mathbf{Z} + \boldsymbol{\mu}, \quad (4)$$

em que  $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^\top$ . A equação (4) expressa em sistema lineares pode ser encontrada em Ross (2010, p. 432).



# Distribuição normal multivariada

Sejam  $\mathbf{X}$  um vetor com distribuição normal multivariada, como dada em (3), e  $\mathbf{a} = (a_1, a_2, \dots, a_k)^\top \in \mathbb{R}^k$  um vetor de constantes. Agora, seja uma VA definida da seguinte forma:



# Distribuição normal multivariada

Sejam  $\mathbf{X}$  um vetor com distribuição normal multivariada, como dada em (3), e  $\mathbf{a} = (a_1, a_2, \dots, a_k)^\top \in \mathbb{R}^k$  um vetor de constantes. Agora, seja uma VA definida da seguinte forma:

$$Y = \mathbf{a}^\top \mathbf{X} = a_1 X_1 + a_2 X_2 + \dots + a_k X_k.$$



# Distribuição normal multivariada

Sejam  $\mathbf{X}$  um vetor com distribuição normal multivariada, como dada em (3), e  $\mathbf{a} = (a_1, a_2, \dots, a_k)^\top \in \mathbb{R}^k$  um vetor de constantes. Agora, seja uma VA definida da seguinte forma:

$$Y = \mathbf{a}^\top \mathbf{X} = a_1 X_1 + a_2 X_2 + \dots + a_k X_k.$$

Então,  $Y$  tem distribuição normal unidimensional com média  $\mu_Y = \mathbf{a}^\top \boldsymbol{\mu}$  e  $\sigma_Y^2 = \mathbf{a}^\top \boldsymbol{\Sigma} \mathbf{a}$ ,



# Distribuição normal multivariada

Sejam  $\mathbf{X}$  um vetor com distribuição normal multivariada, como dada em (3), e  $\mathbf{a} = (a_1, a_2, \dots, a_k)^\top \in \mathbb{R}^k$  um vetor de constantes. Agora, seja uma VA definida da seguinte forma:

$$Y = \mathbf{a}^\top \mathbf{X} = a_1 X_1 + a_2 X_2 + \dots + a_k X_k.$$

Então,  $Y$  tem distribuição normal unidimensional com média  $\mu_Y = \mathbf{a}^\top \boldsymbol{\mu}$  e  $\sigma_Y^2 = \mathbf{a}^\top \boldsymbol{\Sigma} \mathbf{a}$ , i.e., combinação linear de distribuições normais tem distribuição normal.



# Distribuição normal multivariada

Sejam  $\mathbf{X}$  um vetor com distribuição normal multivariada, como dada em (3), e  $\mathbf{a} = (a_1, a_2, \dots, a_k)^\top \in \mathbb{R}^k$  um vetor de constantes. Agora, seja uma VA definida da seguinte forma:

$$Y = \mathbf{a}^\top \mathbf{X} = a_1 X_1 + a_2 X_2 + \cdots + a_k X_k.$$

Então,  $Y$  tem distribuição normal unidimensional com média  $\mu_Y = \mathbf{a}^\top \boldsymbol{\mu}$  e  $\sigma_Y^2 = \mathbf{a}^\top \boldsymbol{\Sigma} \mathbf{a}$ , i.e., combinação linear de distribuições normais tem distribuição normal.



# Roteiro

1 Distribuição multinomial

2 Distribuição normal multivariada

3 Distribuição Wishart

4 Distribuição Dirichlet

5 Referências bibliográficas



# Distribuição Wishart

Sejam  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  vetores aleatórios independentes,  $p \times 1$ , tal que

$$\mathbf{X}_i \sim \mathcal{N}_p(\mathbf{0}, \Sigma),$$

$i = 1, 2, \dots, n$ . Seja também,

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n),$$



# Distribuição Wishart

Sejam  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  vetores aleatórios independentes,  $p \times 1$ , tal que

$$\mathbf{X}_i \sim \mathcal{N}_p(\mathbf{0}, \Sigma),$$

$i = 1, 2, \dots, n$ . Seja também,

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n),$$

uma matriz aleatória, de dimensão  $p \times n$ .



# Distribuição Wishart

Sejam  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  vetores aleatórios independentes,  $p \times 1$ , tal que

$$\mathbf{X}_i \sim \mathcal{N}_p(\mathbf{0}, \Sigma),$$

$i = 1, 2, \dots, n$ . Seja também,

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n),$$

uma matriz aleatória, de dimensão  $p \times n$ .



# Distribuição Wishart

A matriz aleatória, de dimensão  $p \times p$ , construída da seguinte forma,

$$\mathbf{M} = \mathbf{X}\mathbf{X}^\top = \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^\top,$$



# Distribuição Wishart

A matriz aleatória, de dimensão  $p \times p$ , construída da seguinte forma,

$$\mathbf{M} = \mathbf{X}\mathbf{X}^\top = \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^\top,$$

é denominada de distribuição Wishart, com  $n$  graus de liberdade e matriz de covariâncias  $\Sigma$ .



# Distribuição Wishart

A matriz aleatória, de dimensão  $p \times p$ , construída da seguinte forma,

$$\mathbf{M} = \mathbf{X}\mathbf{X}^\top = \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^\top,$$

é denominada de distribuição Wishart, com  $n$  graus de liberdade e matriz de covariâncias  $\Sigma$ . Notação:  $\mathbf{M} \sim W_p(n, \Sigma)$ .



# Distribuição Wishart

A matriz aleatória, de dimensão  $p \times p$ , construída da seguinte forma,

$$\mathbf{M} = \mathbf{X}\mathbf{X}^\top = \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^\top,$$

é denominada de distribuição Wishart, com  $n$  graus de liberdade e matriz de covariâncias  $\Sigma$ . Notação:  $\mathbf{M} \sim W_p(n, \Sigma)$ .



# Distribuição Wishart

Se  $\mathbf{M} \sim W_p(n, \Sigma)$ , sua FDP conjunta é dada pela seguinte expressão

$$f(\mathbf{M}) = \frac{1}{2^{np/2} \Gamma_p\left(\frac{n}{2}\right) |\Sigma|^{n/2}} |\mathbf{M}|^{(n-p-1)/2} \exp\left\{-\frac{1}{2} \text{tr} [\Sigma^{-1} \mathbf{M}]\right\} \mathbb{I}_{\mathbb{R}_+^{p \times p}}(\mathbf{M}),$$



# Distribuição Wishart

Se  $\mathbf{M} \sim W_p(n, \Sigma)$ , sua FDP conjunta é dada pela seguinte expressão

$$f(\mathbf{M}) = \frac{1}{2^{np/2} \Gamma_p\left(\frac{n}{2}\right) |\Sigma|^{n/2}} |\mathbf{M}|^{(n-p-1)/2} \exp\left\{-\frac{1}{2} \text{tr} [\Sigma^{-1} \mathbf{M}]\right\} \mathbb{I}_{\mathbb{R}_+^{p \times p}}(\mathbf{M}),$$

em que  $\text{tr}[\cdot]$  é o operador traço e  $\Gamma_p(\cdot)$  é a função gama multivariada.



# Distribuição Wishart

Se  $\mathbf{M} \sim W_p(n, \Sigma)$ , sua FDP conjunta é dada pela seguinte expressão

$$f(\mathbf{M}) = \frac{1}{2^{np/2} \Gamma_p\left(\frac{n}{2}\right) |\Sigma|^{n/2}} |\mathbf{M}|^{(n-p-1)/2} \exp\left\{-\frac{1}{2} \text{tr} [\Sigma^{-1} \mathbf{M}]\right\} \mathbb{I}_{\mathbb{R}_+^{p \times p}}(\mathbf{M}),$$

em que  $\text{tr} [\cdot]$  é o operador traço e  $\Gamma_p(\cdot)$  é a função gama multivariada.



# Distribuição Wishart

Observações:

- ① se  $\Sigma = I_p$ ,



# Distribuição Wishart

Observações:

- ① se  $\Sigma = I_p$ , a distribuição de  $M$  é denominada de Wishart padrão;



# Distribuição Wishart

Observações:

- ① se  $\Sigma = \mathbf{I}_p$ , a distribuição de  $\mathbf{M}$  é denominada de Wishart padrão;
- ② a Wishart é a extensão multivariada da distribuição qui-quadrado;



# Distribuição Wishart

Observações:

- ① se  $\Sigma = \mathbf{I}_p$ , a distribuição de  $\mathbf{M}$  é denominada de Wishart padrão;
- ② a Wishart é a extensão multivariada da distribuição qui-quadrado;
- ③ se  $M \sim W_1(n, \sigma^2)$ ,



# Distribuição Wishart

Observações:

- ① se  $\Sigma = \mathbf{I}_p$ , a distribuição de  $\mathbf{M}$  é denominada de Wishart padrão;
- ② a Wishart é a extensão multivariada da distribuição qui-quadrado;
- ③ se  $\mathbf{M} \sim W_1(n, \sigma^2)$ , então  $\mathbf{M}/\sigma^2 \sim \chi_n^2$ .



# Distribuição Wishart

Observações:

- ① se  $\Sigma = \mathbf{I}_p$ , a distribuição de  $\mathbf{M}$  é denominada de Wishart padrão;
- ② a Wishart é a extensão multivariada da distribuição qui-quadrado;
- ③ se  $\mathbf{M} \sim W_1(n, \sigma^2)$ , então  $\mathbf{M}/\sigma^2 \sim \chi_n^2$ .



# Roteiro

1 Distribuição multinomial

2 Distribuição normal multivariada

3 Distribuição Wishart

4 Distribuição Dirichlet

5 Referências bibliográficas



# Distribuição Dirichlet

Nós podemos assumir que o vetor aleatório  $\mathbf{X} = (X_1, X_2, \dots, X_D)^\top$  segue uma distribuição Dirichlet (Olkin e Rubin, 1964) com vetor de parâmetros  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_D)^\top$ ,  $\lambda_j > 0$ ,  $j = 1, 2, \dots, D$ , com  $D \geq 2$ ,



# Distribuição Dirichlet

Nós podemos assumir que o vetor aleatório  $\mathbf{X} = (X_1, X_2, \dots, X_D)^\top$  segue uma distribuição Dirichlet (Olkin e Rubin, 1964) com vetor de parâmetros  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_D)^\top$ ,  $\lambda_j > 0$ ,  $j = 1, 2, \dots, D$ , com  $D \geq 2$ , se sua FDP conjunta pode ser escrita como

$$f(\mathbf{x}, \boldsymbol{\lambda}) = \frac{\Gamma\left(\sum_{j=1}^D \lambda_j\right)}{\prod_{j=1}^D \Gamma(\lambda_j)} \prod_{j=1}^D x_j^{\lambda_j-1} \mathbb{I}_{(0,1)}(x_1 + x_2 + \dots + x_D), \quad (5)$$



# Distribuição Dirichlet

Nós podemos assumir que o vetor aleatório  $\mathbf{X} = (X_1, X_2, \dots, X_D)^\top$  segue uma distribuição Dirichlet (Olkin e Rubin, 1964) com vetor de parâmetros  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_D)^\top$ ,  $\lambda_j > 0$ ,  $j = 1, 2, \dots, D$ , com  $D \geq 2$ , se sua FDP conjunta pode ser escrita como

$$f(\mathbf{x}, \boldsymbol{\lambda}) = \frac{\Gamma\left(\sum_{j=1}^D \lambda_j\right)}{\prod_{j=1}^D \Gamma(\lambda_j)} \prod_{j=1}^D x_j^{\lambda_j-1} \mathbb{I}_{(0,1)}(x_1 + x_2 + \dots + x_D), \quad (5)$$

em que  $\Gamma(\cdot)$  é a função gama, isto é,  $\Gamma(\lambda) = \int_0^\infty v^{\lambda-1} e^{-v} dv$ .



# Distribuição Dirichlet

Nós podemos assumir que o vetor aleatório  $\mathbf{X} = (X_1, X_2, \dots, X_D)^\top$  segue uma distribuição Dirichlet (Olkin e Rubin, 1964) com vetor de parâmetros  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_D)^\top$ ,  $\lambda_j > 0$ ,  $j = 1, 2, \dots, D$ , com  $D \geq 2$ , se sua FDP conjunta pode ser escrita como

$$f(\mathbf{x}, \boldsymbol{\lambda}) = \frac{\Gamma\left(\sum_{j=1}^D \lambda_j\right)}{\prod_{j=1}^D \Gamma(\lambda_j)} \prod_{j=1}^D x_j^{\lambda_j-1} \mathbb{I}_{(0,1)}(x_1 + x_2 + \dots + x_D), \quad (5)$$

em que  $\Gamma(\cdot)$  é a função gama, isto é,  $\Gamma(\lambda) = \int_0^\infty v^{\lambda-1} e^{-v} dv$ .



# Distribuição Dirichlet

A distribuição Dirichlet é a generalização da distribuição beta. Além disso, se  $\mathbf{X} \sim \text{Dirichlet}(\boldsymbol{\lambda})$ ,



# Distribuição Dirichlet

A distribuição Dirichlet é a generalização da distribuição beta. Além disso, se  $\mathbf{X} \sim \text{Dirichlet}(\boldsymbol{\lambda})$ , nós temos que



# Distribuição Dirichlet

A distribuição Dirichlet é a generalização da distribuição beta. Além disso, se  $\mathbf{X} \sim \text{Dirichlet}(\boldsymbol{\lambda})$ , nós temos que

$$\mathbb{E}(X_j) = \frac{\lambda_j}{\sum_{j=1}^D \lambda_j}, \quad j = 1, 2, \dots, D.$$



# Distribuição Dirichlet

A distribuição Dirichlet é a generalização da distribuição beta. Além disso, se  $\mathbf{X} \sim \text{Dirichlet}(\boldsymbol{\lambda})$ , nós temos que

$$\mathbb{E}(X_j) = \frac{\lambda_j}{\sum_{j=1}^D \lambda_j}, \quad j = 1, 2, \dots, D.$$

Outras propriedades podem ser encontradas em Pereira e Stern (2008).



# Distribuição Dirichlet

A distribuição Dirichlet é a generalização da distribuição beta. Além disso, se  $\mathbf{X} \sim \text{Dirichlet}(\boldsymbol{\lambda})$ , nós temos que

$$\mathbb{E}(X_j) = \frac{\lambda_j}{\sum_{j=1}^D \lambda_j}, \quad j = 1, 2, \dots, D.$$

Outras propriedades podem ser encontradas em Pereira e Stern (2008).

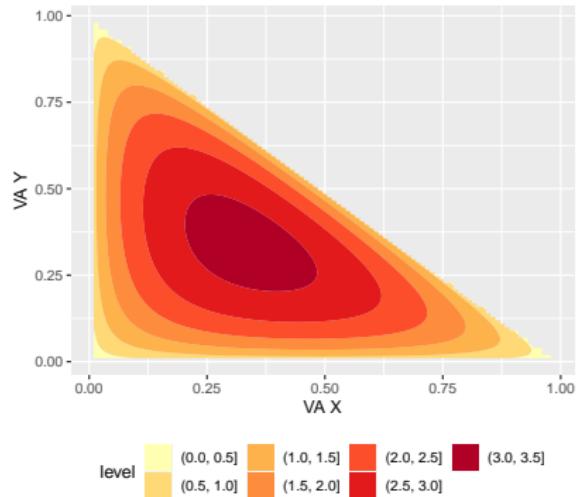


# Distribuição Dirichlet

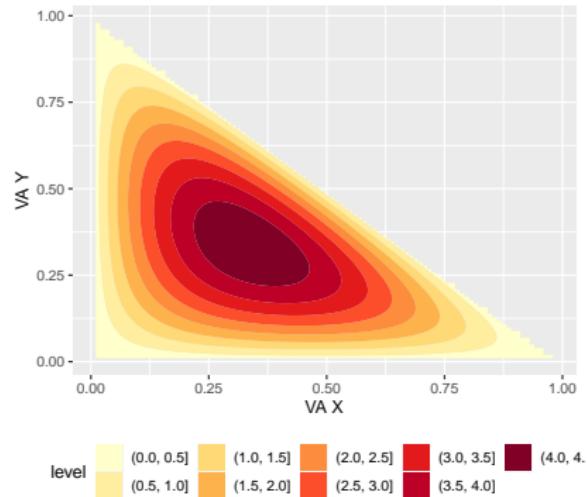
Exemplos:

- ①  $\lambda_1 = 1,5, \lambda_2 = 1,5$  e  $\lambda_3 = 1,5$ ;
- ②  $\lambda_1 = 2, \lambda_2 = 2$  e  $\lambda_3 = 2$ ;
- ③  $\lambda_1 = 2, \lambda_2 = 4$  e  $\lambda_3 = 8$ ;
- ④  $\lambda_1 = 8, \lambda_2 = 4$  e  $\lambda_3 = 2$ .



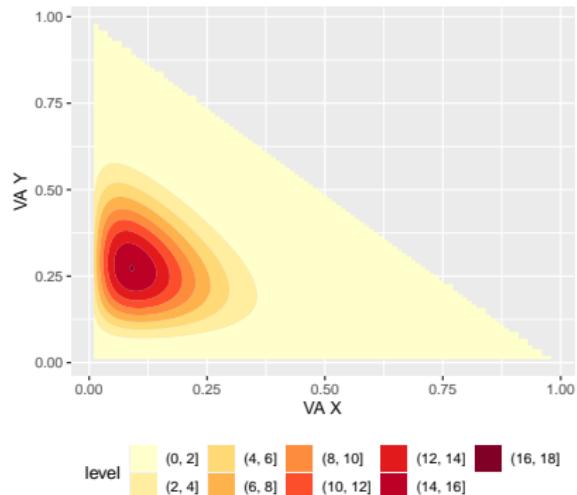


(a) Exemplo 1.

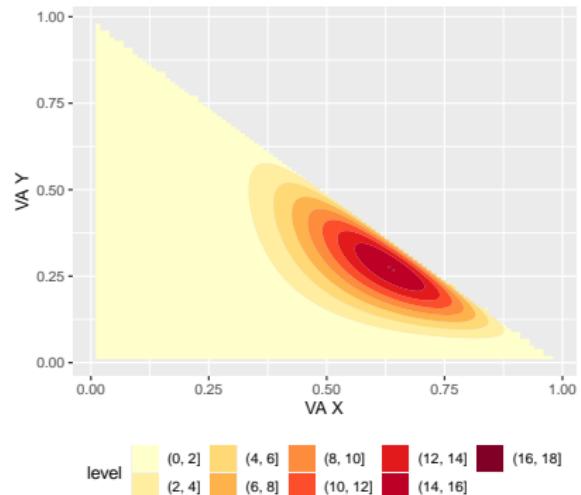


(b) Exemplo 2.

Figura 3: Gráfico de contornos de Dirichlets tridimensionais.



(a) Exemplo 3.



(b) Exemplo 4.

Figura 4: Gráfico de contornos de Dirichlets tridimensionais.

# Distribuição Dirichlet

Uma vantagem de assumir uma distribuição Dirichlet é que ela pode ser simulada a partir de uma distribuição gama, como demonstrado por Devroye (1986).



# Distribuição Dirichlet

Uma vantagem de assumir uma distribuição Dirichlet é que ela pode ser simulada a partir de uma distribuição gama, como demonstrado por Devroye (1986). O algoritmo de simulação é o seguinte:



# Distribuição Dirichlet

Uma vantagem de assumir uma distribuição Dirichlet é que ela pode ser simulada a partir de uma distribuição gama, como demonstrado por Devroye (1986). O algoritmo de simulação é o seguinte:

- ① Fixa-se  $\lambda = (\lambda_1, \dots, \lambda_D)^\top$ ;



# Distribuição Dirichlet

Uma vantagem de assumir uma distribuição Dirichlet é que ela pode ser simulada a partir de uma distribuição gama, como demostrado por Devroye (1986). O algoritmo de simulação é o seguinte:

- ① Fixa-se  $\lambda = (\lambda_1, \dots, \lambda_D)^\top$ ;
- ② Simulam-se  $D$  valores de  $y_j \sim \text{gama}(\lambda_j, 1)$ , com  $j = 1, \dots, D$ ;



# Distribuição Dirichlet

Uma vantagem de assumir uma distribuição Dirichlet é que ela pode ser simulada a partir de uma distribuição gama, como demostrado por Devroye (1986). O algoritmo de simulação é o seguinte:

- ① Fixa-se  $\lambda = (\lambda_1, \dots, \lambda_D)^\top$ ;
- ② Simulam-se  $D$  valores de  $y_j \sim \text{gama}(\lambda_j, 1)$ , com  $j = 1, \dots, D$ ;
- ③ Calculam-se  $x_j = \frac{y_j}{\sum_{j=1}^D y_j}$ , com  $j = 1, \dots, D$ .



# Distribuição Dirichlet

Uma vantagem de assumir uma distribuição Dirichlet é que ela pode ser simulada a partir de uma distribuição gama, como demostrado por Devroye (1986). O algoritmo de simulação é o seguinte:

- ① Fixa-se  $\lambda = (\lambda_1, \dots, \lambda_D)^\top$ ;
- ② Simulam-se  $D$  valores de  $y_j \sim \text{gama}(\lambda_j, 1)$ , com  $j = 1, \dots, D$ ;
- ③ Calculam-se  $x_j = \frac{y_j}{\sum_{j=1}^D y_j}$ , com  $j = 1, \dots, D$ .



# Roteiro

1 Distribuição multinomial

2 Distribuição normal multivariada

3 Distribuição Wishart

4 Distribuição Dirichlet

5 Referências bibliográficas



## Referências bibliográficas I

Devroye, L. (1986), *Non-Uniform Random Variate Generation*, Springer-Verlag, New York.

Hardy, G. H. (1908), 'Mendelian proportions in a mixed population', *Science* **28**(706), 49–50.

Mahalanobis, P. C. (1936), 'On the generalised distance in statistics', *Proceedings of the National Institute of Sciences of India* **2**(1), 49–55.

Olkin, I. e Rubin, H. (1964), 'Multivariate beta distributions and independence properties of the wishart distribution', *Annals of Mathematical Statistics* **35**(1), 261–269.



## Referências bibliográficas II

- Pereira, C. A. B. e Stern, J. M. (2008), 'Special characterizations of standard discrete models', *REVSTAT - Statistical Journal* **6**(3), 199–230.
- Ross, S. (2010), *Probabilidade: um curso moderno com aplicações*, 8 edn, Bookman, Porto Alegre.
- Weinberg, W. (1908), 'Über den nachweis der vererbung beim menschen', *Jahresheft des Vereins für vaterländische Naturkunde in Württemberg* **64**, 368–382.



# Obrigado!

✉ [tiago.magalhaes@ufjf.br](mailto:tiago.magalhaes@ufjf.br)

✉ [ufjf.br/tiago\\_magalhaes](http://ufjf.br/tiago_magalhaes)

🌐 Departamento de Estatística, Sala 319

