Apresentação da Disciplina e Revisão de Algumas Distribuições de Probabilidade

ESTAT0078 – Inferência I

Prof. Dr. Sadraque E. F. Lucena

sadraquelucena@academico.ufs.br

http://sadraquelucena.github.io/inferencia1

UNIVERSIDADE FEDERAL DE SERGIPE

Canais de Comunicação e Materiais da Disciplina

- Site: http://sadraquelucena.github.io/inferencia1
- Grupo no WhatsApp: http://tiny.cc/inf1wpp

Informações da disciplina

- **Componente curricular:** ESTAT0078 – Inferência I
- **Período:** 7º semestre
- **Carga horária:** 60 horas (4 créditos)
- **Horário:**
  - Terças - 19h00 às 20h30
  - Quintas - 20h45 às 22h15
- **Docente:** Prof. Dr. Sadraque E. F. Lucena

Objetivo da Disciplina

- Introduzir fundamentos da inferência estatística;
- Desenvolver habilidades de análise e interpretação de estimadores;
- Compreender a aplicabilidade da inferência na análise de dados.

Ementa

- Amostras e distribuições amostrais.
- Estimação pontual e por intervalo.
- Estudo de estimadores mais comumente usados: método dos momentos, máxima verossimilhança, estimador de Bayes.
- Intervalos de confiança; métodos para construção de intervalos de confiança.

Conteúdo Programático

Parte 1. Princípios da Estimação Pontual

1.1. Conceitos Iniciais: Parâmetros, Estimador e Estatística

1.2. Distribuição amostral de algumas estatísticas

1.3. Propriedades dos estimadores:

- 1.3.1. Estimador não viesado
- 1.3.2. Viés e Erro quadrático Médio
- 1.3.3. Eficiência de um Estimador

1.4. Estatística Suficiente

Conteúdo Programático

Parte 2. Métodos de Estimação

2.1. Família Exponencial

2.2. O Método da Máxima Verossimilhança

2.3. Propriedades dos Estimadores de Máxima Verossimilhança

2.4. Família Exponencial e o Método da Máxima Verossimilhança

2.5. O Método dos Momentos

2.6. Introdução ao estimador de Bayes

Conteúdo Programático

Parte 3. Estimação por intervalo

3.1. Resultados de amostras de população normal

3.2. Método da Quantidade Pivotal

3.3. Intervalos de confiança para populações normais para:

    3.3.1. Uma amostra

    3.3.2. Duas amostras

3.4. Intervalos de confiança aproximados

Avaliação

- Serão realizadas **3 avaliações**.
- A aprovação requer uma média maior ou igual a 5,0.
- Uma **avaliação substitutiva** será realizada ao final do período. Esta prova abrange todo o conteúdo, substitui apenas uma nota e é permitida em dois casos:
  1. Para repor falta em alguma das avaliações.
  2. Para substituir a menor nota (apenas para quem tiver média inferior a 5,0).

Datas Importantes

Avaliações (previsão):

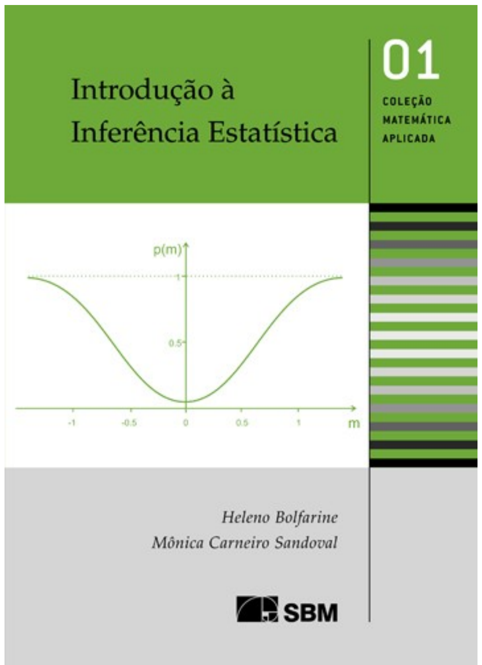
- **13/Nov/25 (quinta):** Avaliação 1
- **18/Dez/25 (quinta):** Avaliação 2
- **10/Fev/26 (terça):** Avaliação 3
- **12/Fev/26 (quinta):** Avaliação Substitutiva

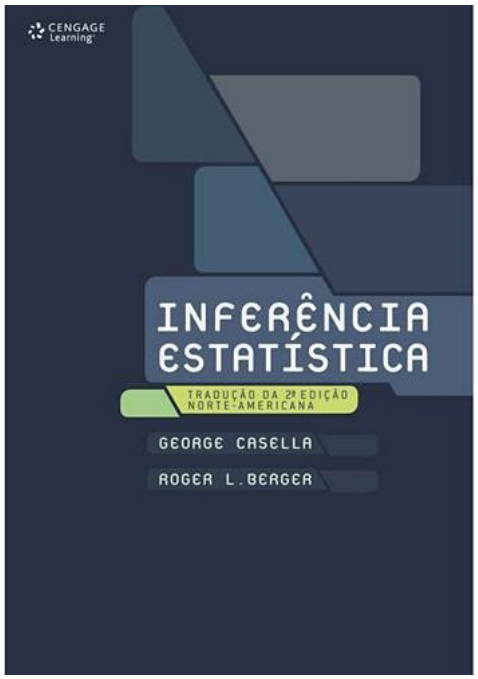
Não haverá aula:

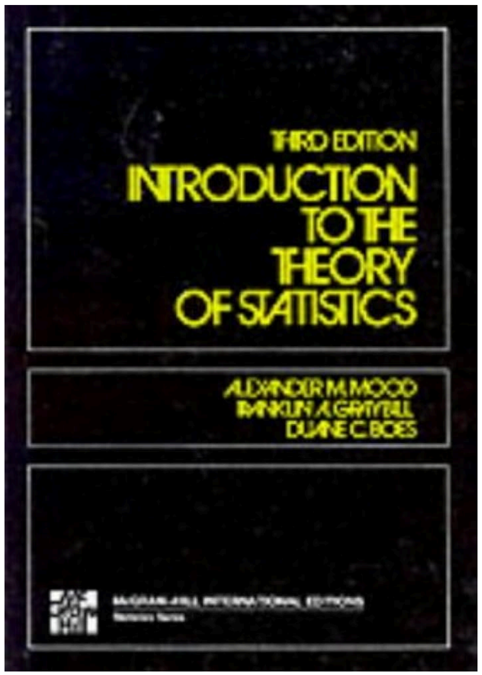
- **20/Nov/25:** Dia Nacional de Zumbi e da Consciência Negra (feriado nacional)
- **25 e 27/Nov/25:** XI SEMAC
- **22 a 31/12/2025:** Recesso de final de ano
- **01/01/2026:** Confraternização Universal (feriado nacional) e Aniversário de São Cristóvão (feriado municipal)
- **02 a 10/01/2026:** Férias coletivas para docentes

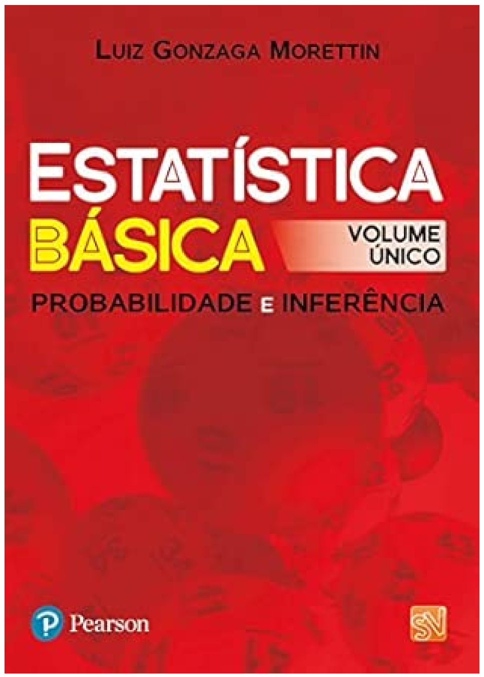
Bibliografia Recomendada

Básica



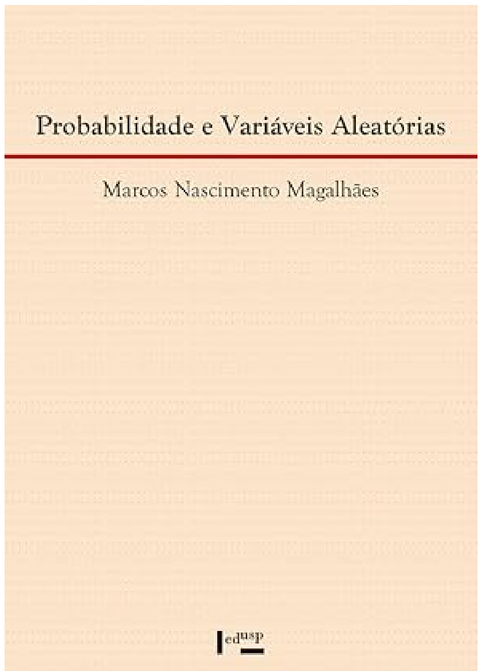


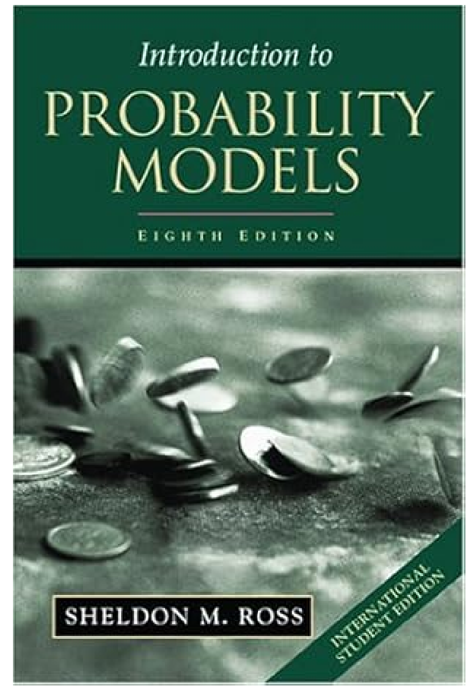




Bibliografia Recomendada

Complementar





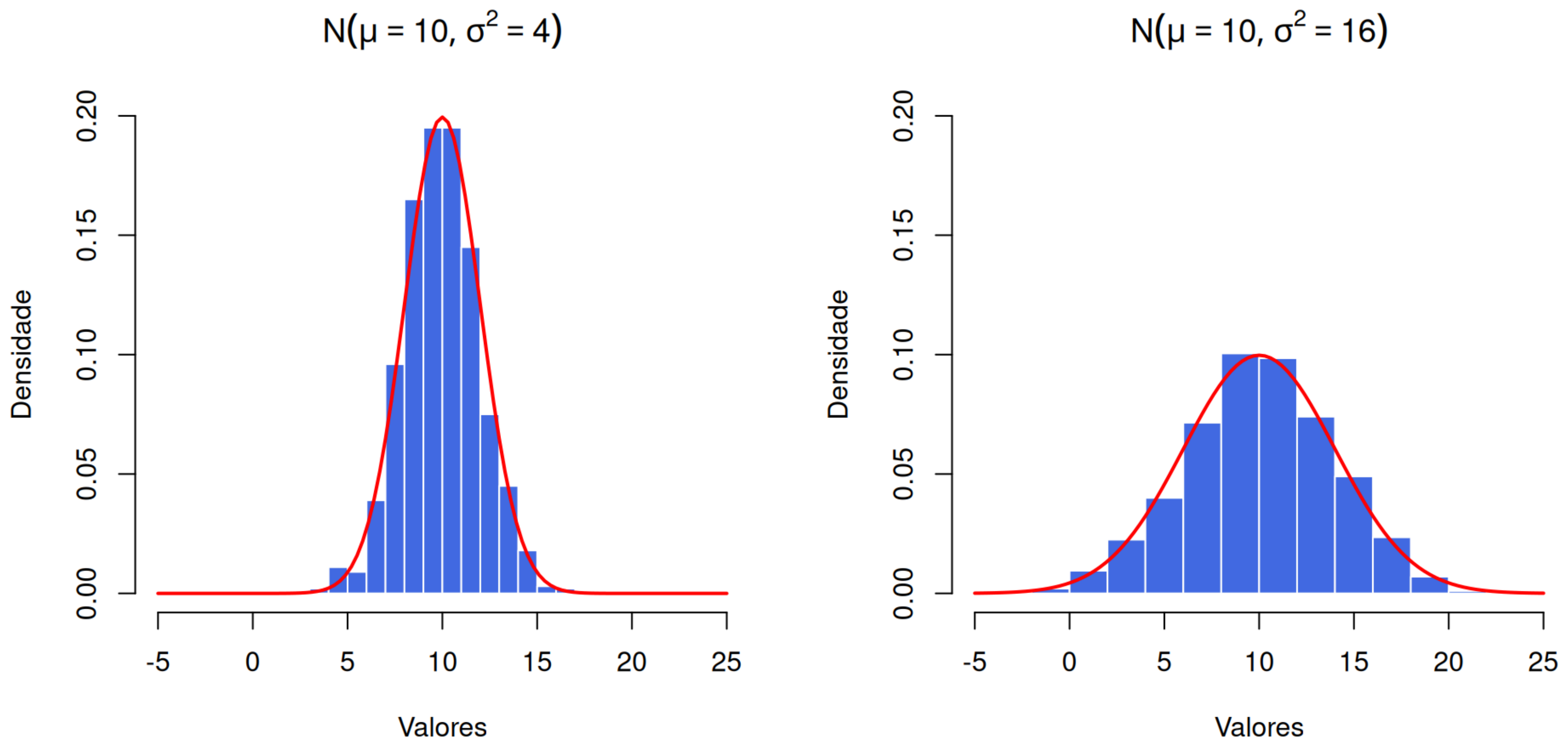
Introdução

Inferência: transformando dados em conhecimento

- As distribuições de probabilidade possuem parâmetros.
- **Exemplos:**
  - Normal(μ, σ^2)
  - Poisson(λ)
  - Exponencial(α)
- Essas distribuições são usadas para representar dados (e assim entender sua variabilidade e quantificar probabilidades).

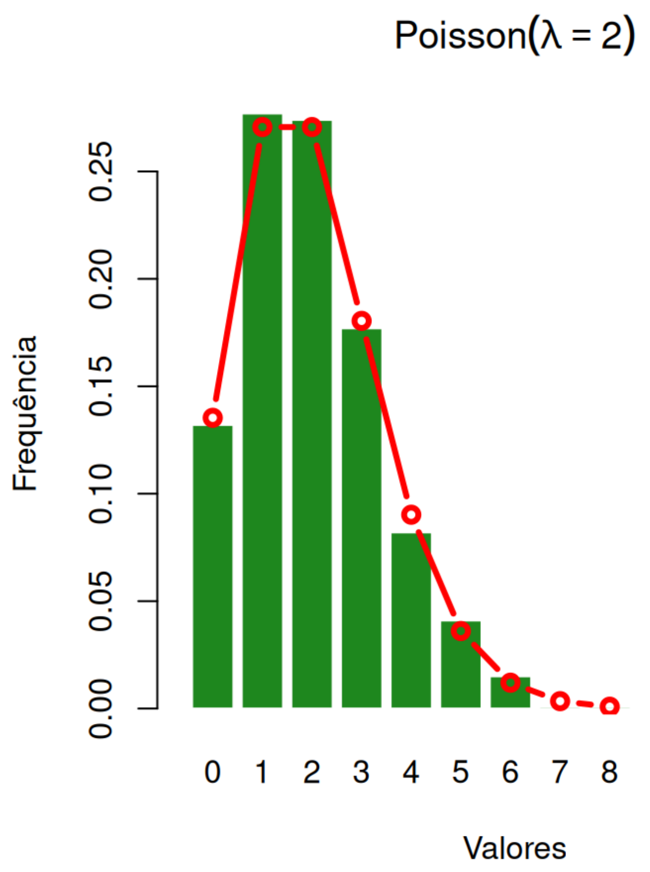
Inferência: transformando dados em conhecimento

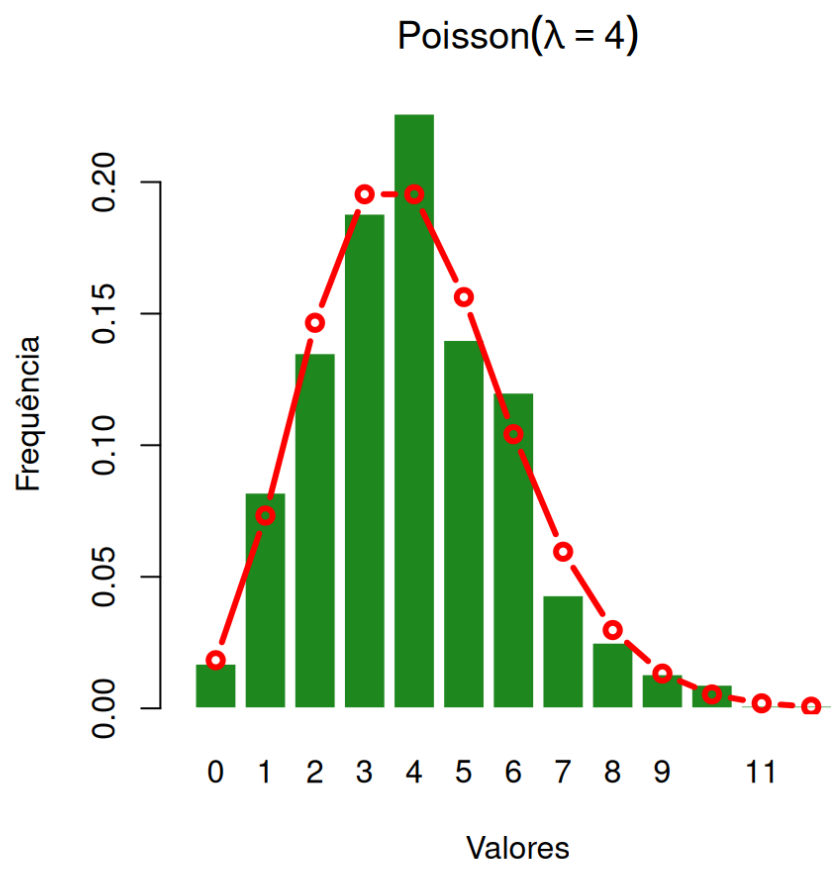
- Normal(μ, σ^2)



Inferência: transformando dados em conhecimento

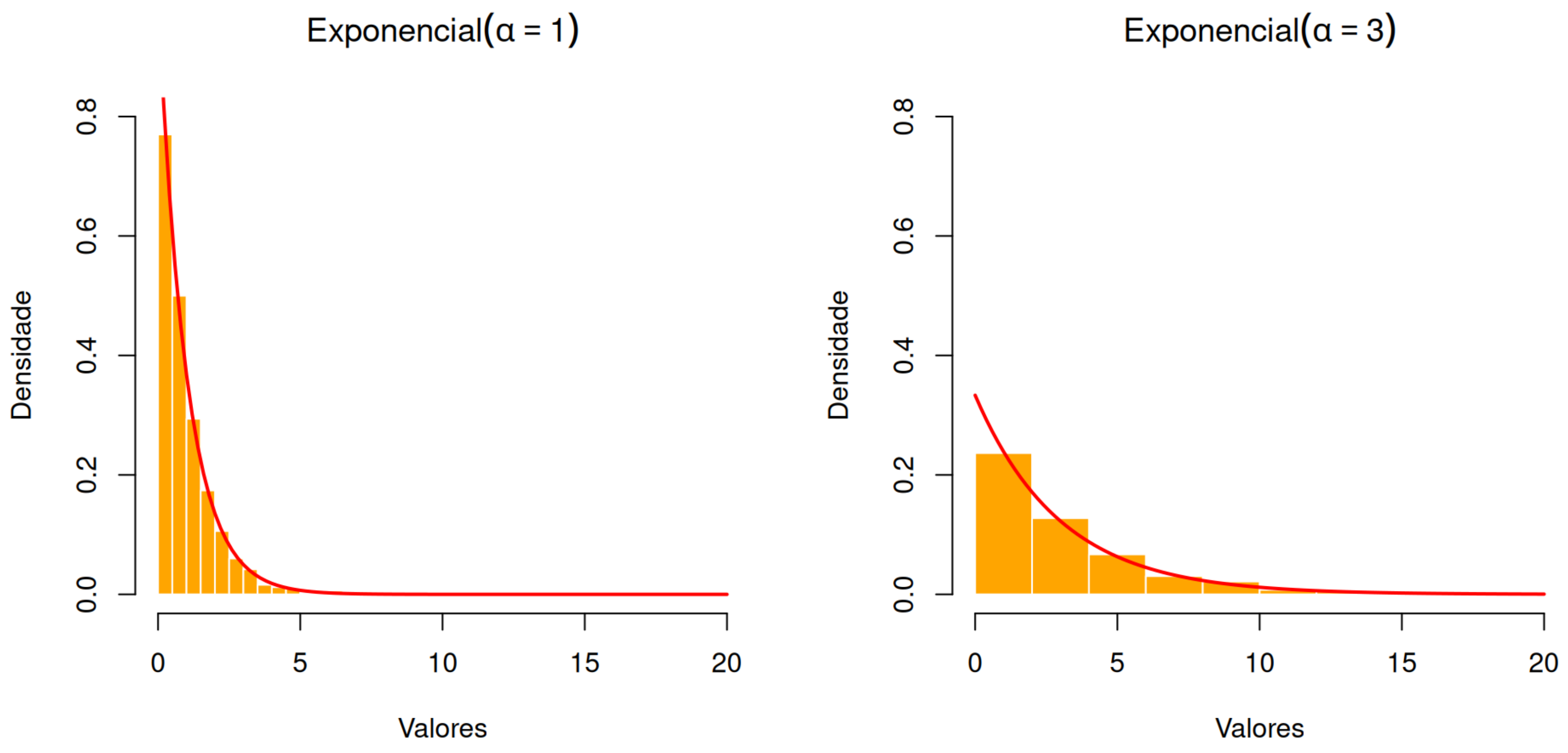
- Poisson(λ)





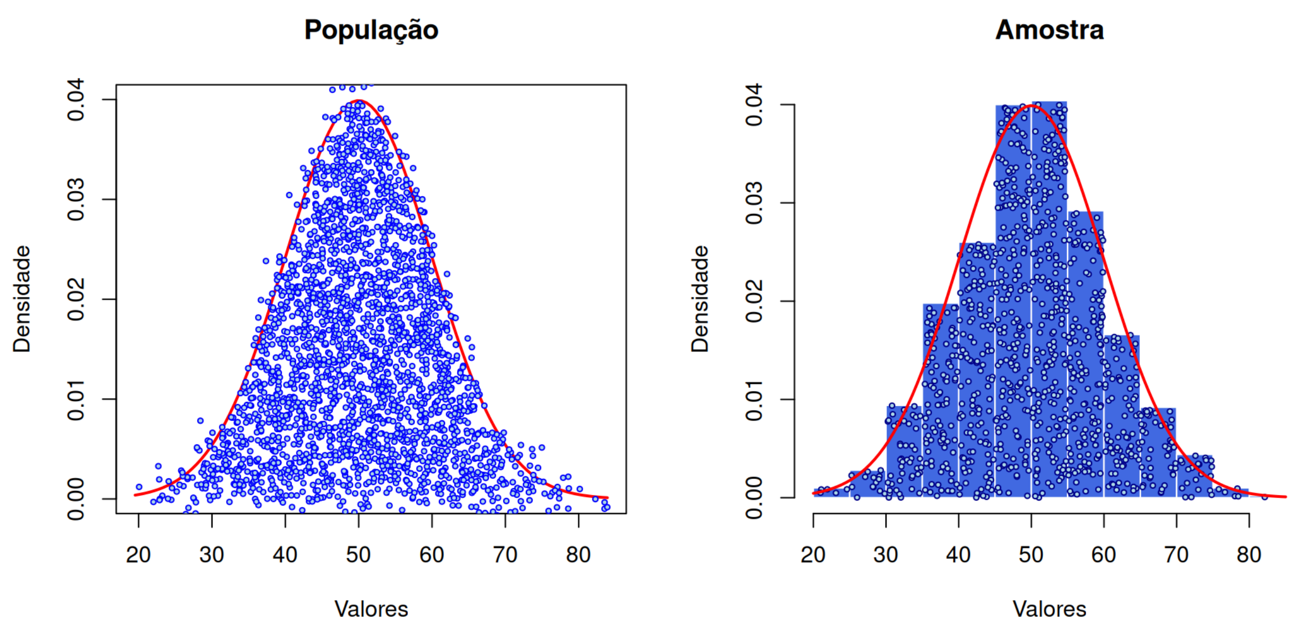
Inferência: transformando dados em conhecimento

- Exponencial(α)

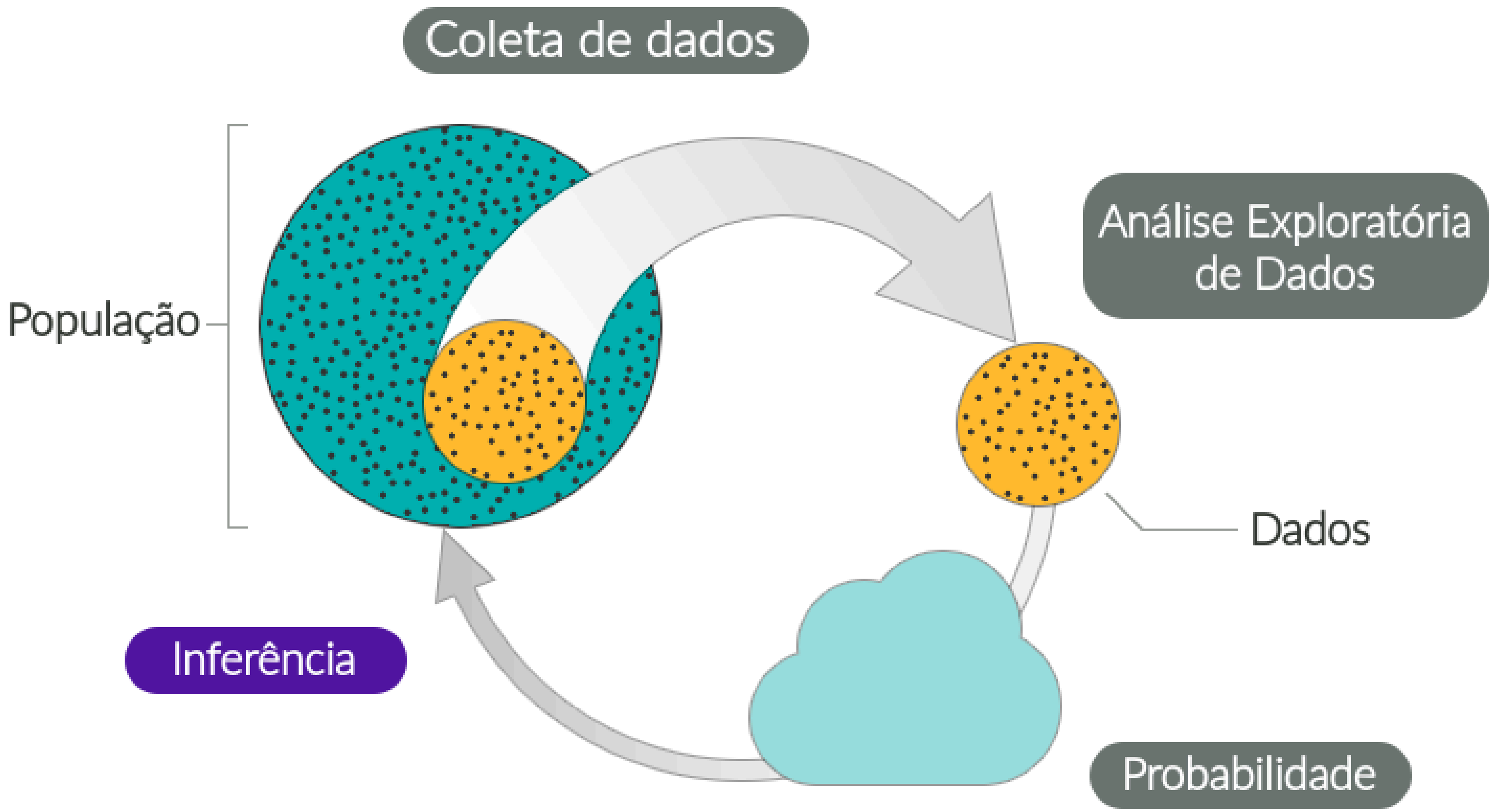


Inferência: transformando dados em conhecimento

- O objetivo da inferência estatística é estimar os parâmetros das distribuições com base nos dados observados.
  - Isso permite fazer afirmações sobre a população e quantificar a probabilidade de estarmos certos (ou errados).



Inferência: transformando dados em conhecimento



Fonte: Adaptado de

https://stats.libretexts.org/Courses/Lumen_Learning/Concepts_in_Statistics_(Lumen)/06%3A_Probability_and_Probability_Distributions/6.0_Probability_and_Probability_Distributions

Revisão de Algumas Distribuições de Probabilidade

Distribuição Binomial

- Descreve o número de sucessos em uma sequência fixa de tentativas independentes.
- Notação: $X \sim$ Binomial(n, p)
- Parâmetros da distribuição:
  - n: número de tentativas
  - p: probabilidade de sucesso em cada tentativa, $0 < p < 1$
- Função de probabilidade:

$$P(X = x) = \binom{n}{p} p^x (1 - p)^{n-x}, \qquad x = 0, 1, 2, \ldots, n$$

Distribuição Binomial

- Propriedades importantes:
  - Média: $E(X) = np$
  - Variância: $Var(X) = np(1 - p)$
- Modela pesquisas eleitorais (favorável ou não a um candidato), número de sinistros em uma carteira, etc.

Distribuição Binomial no R

Considere $X \sim \text{Binomial}(n, p)$. Para calcular:

- $P(X = x)$ use o comando `dbinom(x, n, p)`.
- $P(X \leq x)$ use o comando `pbinom(x, n, p)`.
- Para simular r ocorrências de uma binomial use `rbinom(r, n, p)`.

Distribuição Binomial

Exemplo 1.1

Seja $X \sim$ Binomial($n = 10; p = 0{,}3$).

a. $P(X = 4) = \binom{n}{p} p^x (1-p)^{n-x} = \binom{10}{4} 0{,}3^4 (1-0{,}3)^{10-4}$
$= 0{,}2001$

Cálculo no R:

```
dbinom(4, 10, 0.3)
```

```
[1] 0.2001209
```

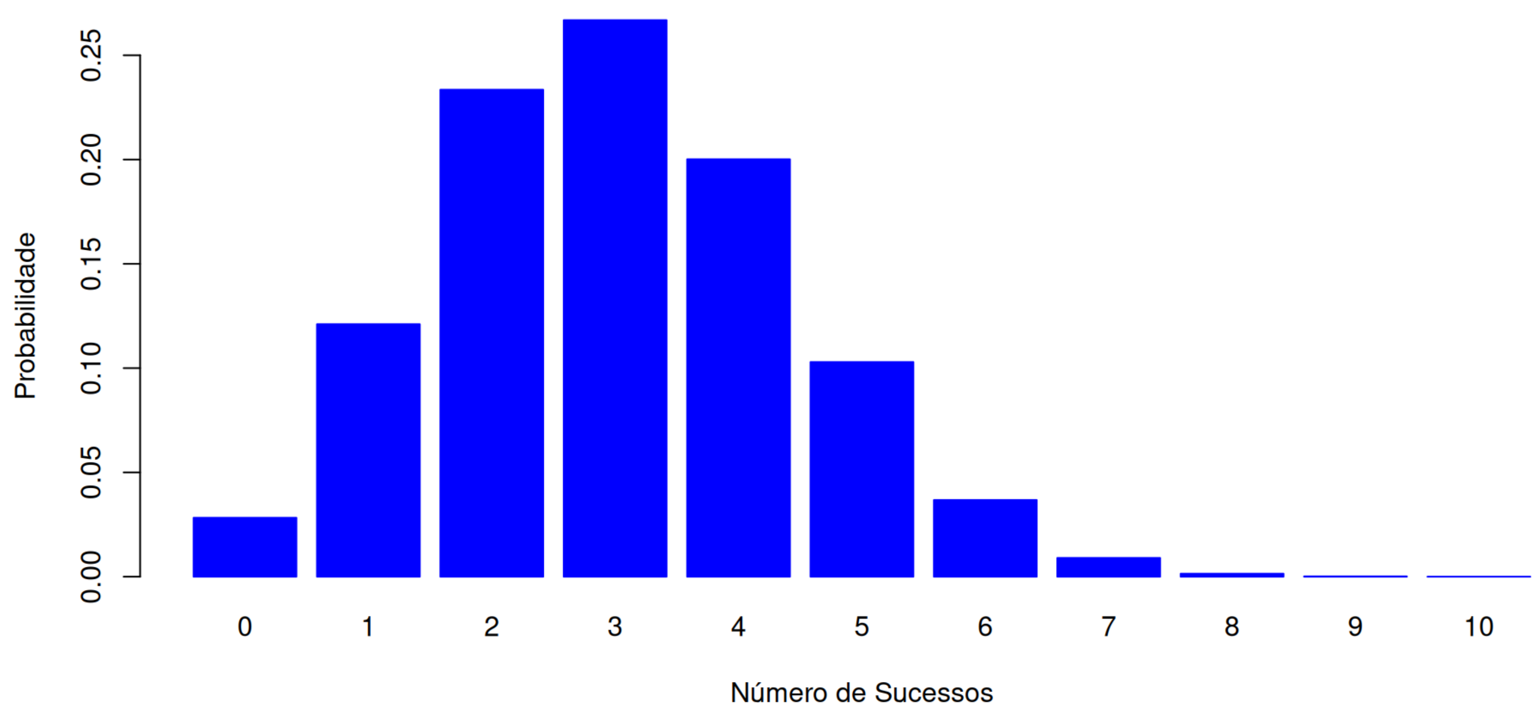
b. $E(X) = 10 \cdot 0{,}3 = 3$

c. $Var(X) = 10 \cdot 0{,}3(1 - 0{,}3) = 2{,}1$

Distribuição Binomial

Gráfico de $X \sim$ Binomial($n = 10; p = 0{,}3$):

```r
n <- 10; p <- 0.3
probabilidades <- dbinom(0:n, n, p)
barplot(probabilidades, names.arg = 0:n,
        xlab = "Número de Sucessos", ylab = "Probabilidade",
        col = "blue", border = "blue", ylim = c(0,.25))
```



Distribuição de Poisson

- Expressa a probabilidade de um número fixo de eventos ocorrer em um intervalo de tempo ou espaço, com uma taxa média conhecida e de forma independente.
- Notação: $X \sim \text{Poisson}(\lambda)$
- Parâmetro da distribuição:
  - λ: taxa média de eventos por unidade de tempo ou espaço
- Função de probabilidade:

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \qquad x = 0, 1, 2, \dots$$

Distribuição de Poisson

- Propriedade importante:
  - $E(X) = Var(X) = \lambda$ (média e variância iguais)

Distribuição de Poisson no R

Considere $X \sim \text{Poisson}(\lambda)$. Para calcular:

- $P(X = x)$ use o comando `dpois(x, lambda)`.
- $P(X \leq x)$ use o comando `pbinom(x, lambda)`.
- Para simular r ocorrências de uma Poisson use `rpois(r, lambda)`.

Distribuição de Poisson

Exemplo 1.2

Seja $X \sim \text{Poisson}(\lambda = 6)$.

a. $P(X = 4) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-6}6^4}{4!} = 0{,}1339$

Cálculo no R:

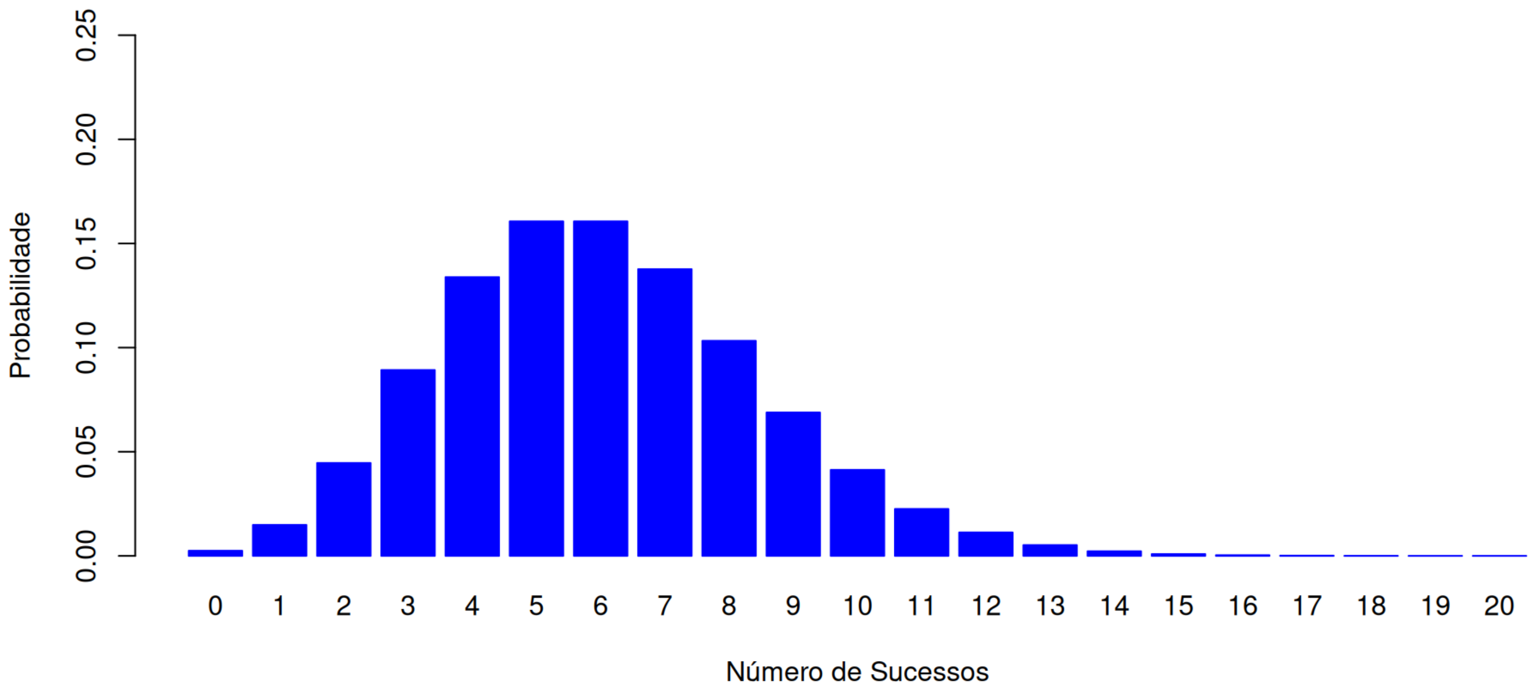
```r
dpois(4, 6)
```

```
[1] 0.1338526
```

Distribuição de Poisson

Gráfico da distibuição de $X \sim \text{Poisson}(\lambda = 6)$:

```
n <- 20; lambda <- 6
probabilidades <- dpois(0:n, lambda)
barplot(probabilidades, names.arg = 0:n,
        xlab = "Número de Sucessos", ylab = "Probabilidade",
        col = "blue", border = "blue", ylim = c(0,.25))
```



Distribuição Uniforme

- Descreve uma situação onde todos os resultados possíveis de uma variável aleatória têm a mesma probabilidade de ocorrer.
- Notação: $X \sim \text{Uniforme}(a, b)$
- Parâmetros da distribuição:
  - a: limite inferior (mínimo)
    - b: limite superior (máximo)
- Função de densidade de probabilidade:

$$f(x) = \frac{1}{b - a}, \quad a \leq x \leq b$$

Distribuição Uniforme

- Propriedades importantes:
  - Média: $E(X) = \frac{a+b}{2}$
    - Variância: $Var(X) = \frac{(b-a)^2}{12}$
    - A distribuição é simétrica em torno da média.
    - A variável aleatória pode assumir qualquer valor entre a e b.

Distribuição Uniforme no R

Considere $X \sim \text{Uniforme}(a, b)$. Para calcular:

- $P(X \leq x)$ use o comando `punif(x, min=a, max=b)`.
- Para simular r ocorrências de uma uniforme use `runif(r, min=a, max=b)`.

Distribuição Uniforme

Exemplo 1.3

Seja $X \sim$ Uniforme$(a = 1, b = 7)$.

a. $P(X \leq 6) = \int_a^x \frac{1}{b-a} dx = \int_1^6 \frac{1}{7-1} dx = \int_1^6 \frac{1}{6} dx$
$= \left[\frac{x}{6}\right]_1^6 = 1 - \frac{1}{6} = 0{,}8333$

Cálculo no R:

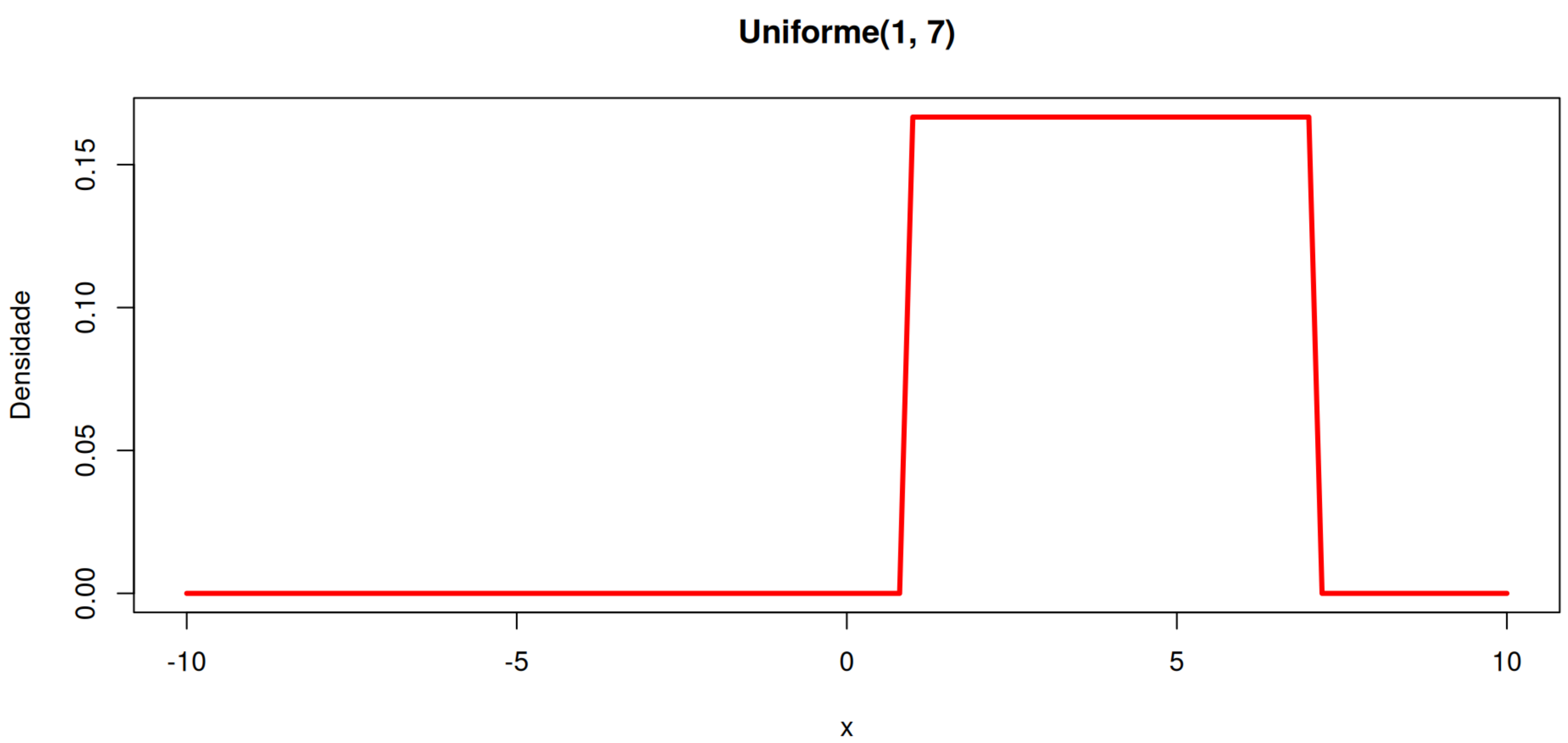
```
punif(6, min=1, max=7)
```

```
[1] 0.8333333
```

Distribuição Uniforme

Gráfico da distibuição de $X \sim$ Uniforme($a = 1, b = 7$):

```
a <- 1; b <- 7
curve(dunif(x, min=a, max=b), from = -10, to = 10,
      main = "Uniforme(1, 7)", xlab = "x", ylab = "Densidade",
      lwd = 3, col = "red")
```



Distribuição Normal

- Distribuição contínua, simétrica, e em forma de sino.
- Representa muitas variáveis naturais e mensuráveis, como altura, peso, e notas de testes.
- Notação: $X \sim \text{Normal}(\mu, \sigma^2)$
- Parâmetros da distribuição:
  - μ: média
    - σ^2: variância
- Função de densidade de probabilidade:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma^2 >$$

Distribuição Normal

- Propriedades importantes:
  - É perfeitamente simétrica em torno da média.
    - A maior parte dos dados está concentrada ao redor da média, com a probabilidade de eventos extremos diminuindo conforme a distância da média.
    - 68% dos dados estão dentro de $\mu \pm 1\sigma^2$.
    - 95% dos dados estão dentro de $\mu \pm 2\sigma^2$.
    - 99,7% dos dados estão dentro de $\mu \pm 3\sigma^2$.

Distribuição Normal no R

Considere $X \sim \text{Normal}(\mu, \sigma^2)$. Para calcular:

- $P(X \leq x)$ use o comando `pnorm(x, mean=`μ`, sd=`σ`)`.
- Para simular r ocorrências de uma normal use `rnorm(r, mean=`μ`, sd=`σ`)`.

Distribuição Normal

Exemplo 1.4

Seja $X \sim$ Normal$(\mu = 0, \sigma^2 = 4)$.

a. $P(X \leq 1) = \int_{-\infty}^{1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$

Cálculo no R:

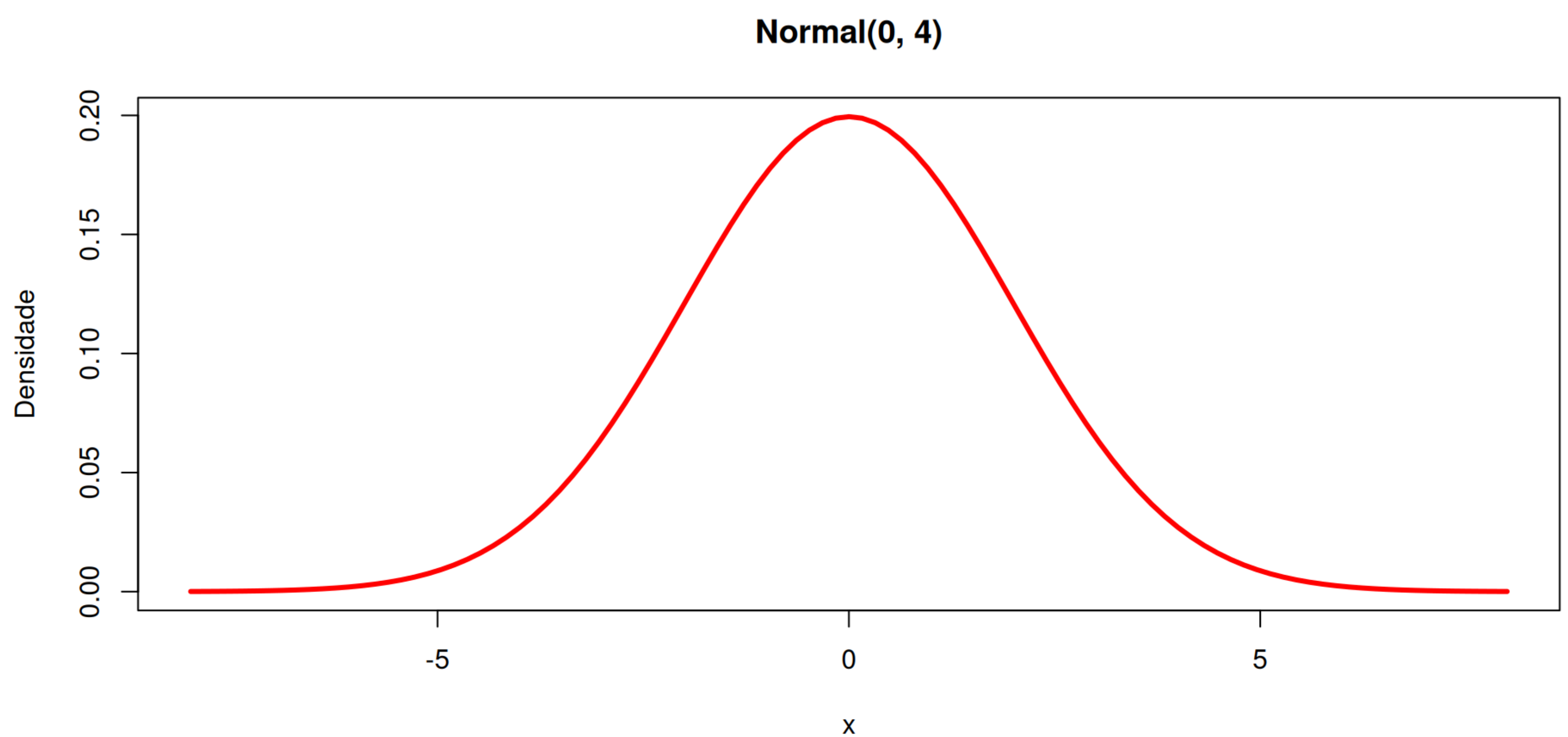
```
pnorm(1, mean=0, sd=sqrt(4))
```

```
[1] 0.6914625
```

Distribuição Normal

Gráfico da distibuição de $X \sim \text{Normal}(\mu = 0, \sigma^2 = 4)$:

```
mu <- 0; sigma <- sqrt(4)
curve(dnorm(x, mean=mu, sd=sigma), from = -8, to = 8,
      main = "Normal(0, 4)", xlab = "x", ylab = "Densidade",
      lwd = 3, col = "red")
```



Distribuição Exponencial

- Distribuição contínua que é frequentemente usada para modelar o tempo de vida de equipamentos.
- Notação: $X \sim$ Exponencial(α)
- Parâmetro da distribuição:
  - α: taxa média de ocorrências de um evento por unidade de tempo.
- Função de densidade de probabilidade:

$$f(x) = \alpha e^{-\alpha x}, \qquad x \geq 0, \alpha > 0$$

Distribuição Exponencial

- Propriedades importantes:
  - Média: $E(X) = \frac{1}{\alpha}$
  - Variância: $Var(X) = \frac{1}{\alpha^2}$
  - A distribuição exponencial tem a propriedade de falta de memória, ou seja, a probabilidade de um evento futuro não depende do tempo já decorrido.
  - O tempo entre eventos em um processo de Poisson é exponencialmente distribuído.

Distribuição Exponencial no R

Considere $X \sim$ Exponencial(α). Para calcular:

- $P(X \leq x)$ use o comando `pexp(x, rate=`α`)`.
- Para simular r ocorrências de uma Exponencial use `rexp(r, rate=`α`)`.

Distribuição Exponencial

Exemplo 1.4

Seja $X \sim$ Exponencial($\alpha = 2$).

a. $P(X \leq 1) = \int_0^x \alpha e^{-\alpha x} dx = \int_0^1 2e^{-2x} dx = -e^{-2x}\Big|_0^1$
$= 1 - e^{-2} = 0{,}8647$

Cálculo no R:

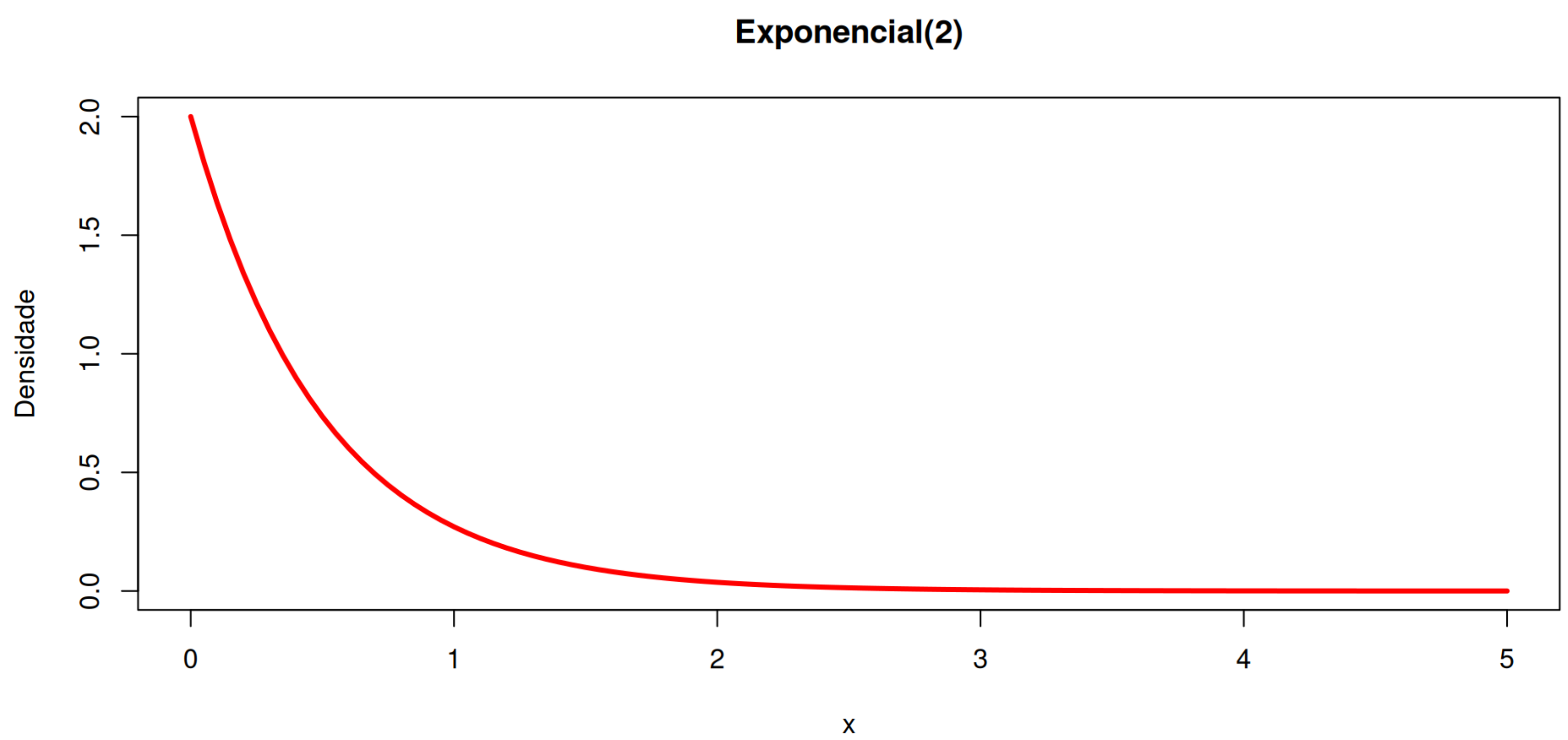
```
pexp(1, rate=2)
```

```
[1] 0.8646647
```

Distribuição Exponencial

Gráfico da distibuição de $X \sim \text{Exponencial}(\alpha = 2)$:

```
alpha <- 2
curve(dexp(x, rate=alpha), from = 0, to = 5,
      main = "Exponencial(2)", xlab = "x", ylab = "Densidade",
      lwd = 3, col = "red")
```



Fim