Probabilidade e inferência estatística com aplicações em R

Aula 3 - Distribuições contínuas

Felipe Barletta

Departamento de Estatística

22 setembro, 2020

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

Sumário

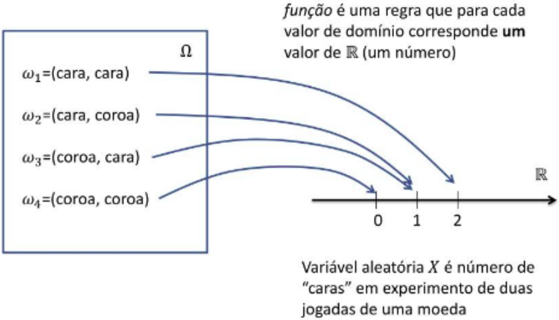
Variável aleatória

- **Variável aleatória** - Descrição numérica do resultado de um fenômeno aleatório.
- Em probabilidade, uma função X que associa a cada evento do espaço amostral um número real $X(\omega) \in \mathbb{R}$, é denominada uma variável aleatória (V.A.).
- Dessa forma, podemos atribuir probabilidades aos eventos desse espaço amostral, dando origem ao conceito de variável aleatória

Notação

- X denota a variável aleatória.
- x denota os valores realizados da V.A.
- Probabilidade de X assumir o valor x é denotada $P[X = x]$.

Variável aleatória



Distribuições de probabilidade

- Há diversos modelos probabilísticos
- Distribuições de probabilidade de V.A. (discretas ou contínuas)
- Descrição das probabilidades associadas com os possíveis valores de X
  - **Variáveis discretas** $\Rightarrow$ suporte em um conjunto de valores enumeráveis (finitos ou infinitos)
  - **Variáveis contínuas** $\Rightarrow$ suporte em um conjunto não enumerável de valores

Distribuições de probabilidade

Definição: função densidade de probabilidade

A função densidade de probabilidade $f(x)$ de uma V.A. contínua, é uma função definida em um intervalo de valores na reta dos reias, ou seja,

$$P[a \leq X \leq b] = \int_b^a f_x(x).$$

i. A área total da abaixo da curva de $f(x)$ é 1.

$$\int_{-\infty}^{\infty} f_x(x) = 1$$

ii. A função $f(x)$ é positiva ou zero.

$$f_x(x) \geq 0$$

Exemplo

Em um laboratório, a corrente em um circuito é medida usando um amperímetro. Devido a vários fatores aleatórios, a medida X varia. Registros anteriores indicam que essa corrente varia ao longo do intervalo [2, 6], mas não uniformemente. Valores mais altos de X têm maiores probabilidades de ocorrência. Foi descoberto que um bom modelo para o os dados são representados pelo seguinte f.d.p.:

$$\begin{cases} f(x) = 0.025x + 0.15, & 2 < x < 6 \\ 0, & \text{caso contrário} \end{cases}$$

Exemplo - Verificar se $f(x)$ é uma f.d.p.

$$\begin{aligned}
\int_2^6 (0.25x + 0.15)dx &= 0.025\frac{x^2}{2} + 0.15x\Big|_2^6 \\
&= (0.025 * \frac{36}{2} + 0.15 * 6) - (0.025 * \frac{4}{2} + 0.15 * 2) \\
&= 1.35 - 0.35 \\
&= 1
\end{aligned}$$

UTFPR UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

Exemplo - $P[2 < X < 3]$

$$\begin{aligned}
\int_2^3 (0.25x + 0.15)dx &= 0.025\frac{x^2}{2} + 0.15x\Big|_2^3 \\
&= (0.025 * \frac{9}{2} + 0.15 * 3) - (0.025 * \frac{4}{2} + 0.15 * 2) \\
&= 0.5625 - 0.35 \\
&= 0.2125
\end{aligned}$$

Função de distribuição (ou Função de distribuição acumulada)

- Em muitas situações, é útil calcularmos a probabilidade **acumulada** até um certo valor.
- Definimos a **função de distribuição** ou **função acumulada de probablidade** de uma V.A. X pela expressão:

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} f(t)dt$$

para qualquer número real x.

Exemplo - (Bussab $\&$ Moretin, pág. 167, exercício 1)

- Mostre que $f(x)$ é uma função densidade de probabilidade
- Calcule a probabilidade de que $X > 1$
- Calcule a probabilidade de que $0,2 < X < 0,8$

$$f(x) = \begin{cases} 2e^{-2x}, & \text{se} \quad x \geq 0 \\ 0, & \text{se} \quad x < 0 \end{cases}$$

Modelo Uniforme contínuo

Definição:

Dizemos que X segue o modelo **Uniforme contínuo** se atribui a mesma probabilidade ao longo de um intervalo finito $[a, b]$.
Então, sua função de probabilidade é dada por

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

Notação: $X \sim \mathrm{U}[a, b]$

Distribuições de probabilidade discreta

- Para resolução dos exercícios vamos explorar as seguintes funcionalidades do R para operações com distribuição de probabilidade:
  - *d:* calcula a densidade de probabilidade $p(x)$ no ponto
  - *p:* calcula a função de probabilidade acumulada $F(x)$ no ponto
  - *q:* calcula o quantil correspondente a uma dada probabilidade
  - *r:* gera uma amostra aleatória da distribuição

Modelo Exponencial

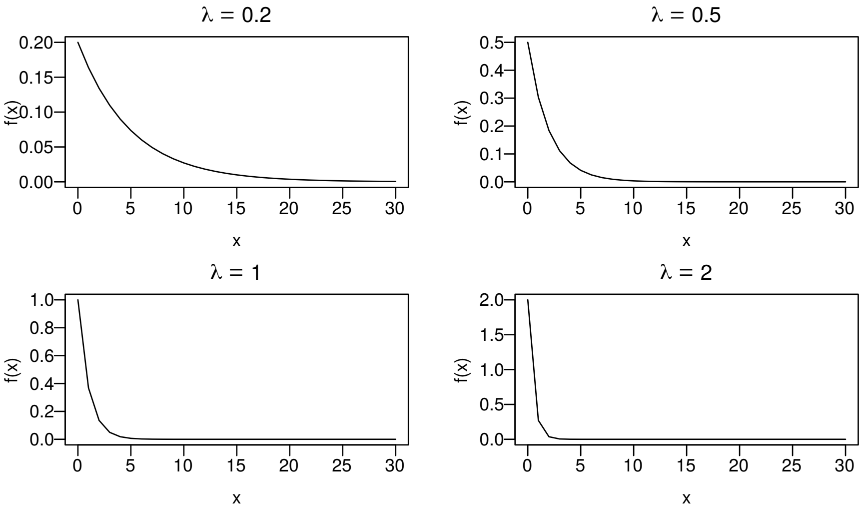
Definição:

uma V.A. contínua X assumindo valores não negativos, segue o modelo exponencial com parâmetro $\lambda > 0$ se sua densidade é dada por

$$f(x) = \lambda e^{-\lambda x} I_{(0,\infty)}(x)$$

Notação: $X \sim \text{Exp}(\lambda)$

Modelo Exponencial



Exemplo: Modelo Exponencial

- Seja uma variável aleatória $X \sim \exp(\lambda = 500)$.
  - Calcular a probabilidade $P[X \geq 400)$.
    1. Use a função básica do R
    2. Crie uma função no R e compare os resultados

Modelo Gaussiano(ou Normal)

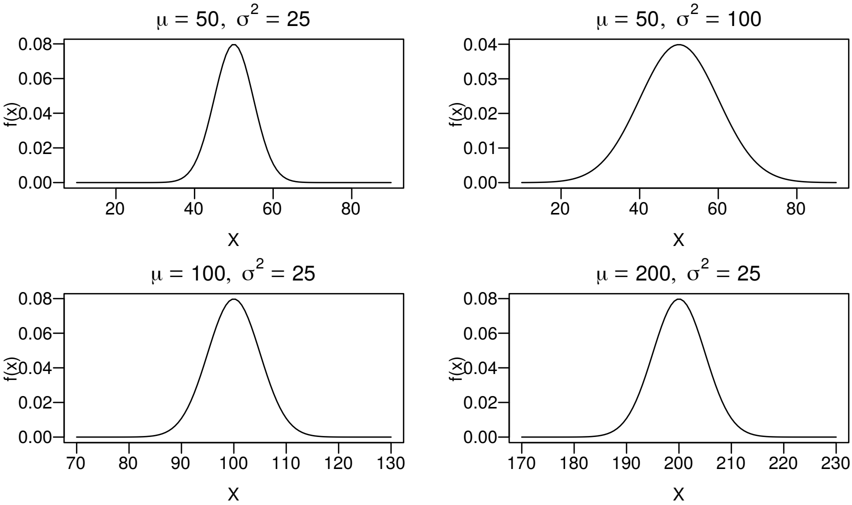
Definição:

Uma V.A. contínua é dita ter distribuição gaussiana com parâmetros μ e σ^2 se sua f.d.p. for:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty$$

Notação: $X \sim \mathsf{N}(\mu, \sigma^2)$

Modelo Gaussiano(ou Normal)



Modelo Gaussiano(ou Normal)

Característcas da curva normal:

- É **simétrica** em relação à μ
- O ponto máximo (moda) de $f(x)$ é o ponto $x = \mu$
- Os pontos de inflexão da função são $\mu - \sigma$ e $\mu + \sigma$
- A área total sob a curva é 1 ou 100%
- A curva é **assintótica** em relação ao eixo x

Normal padrão

Definição:

Uma V.A. contínua que segue uma distribuição normal com parâmetros $\mu = 0$ e $\sigma^2 = 1$, é conhecida com uma distribuição normal padrão.

Notação: $Z \sim \mathsf{N}(\mu = 0, \sigma^2 = 1)$

e sua f.d.p. é:

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(z)^2\right], \quad -\infty < z < \infty$$

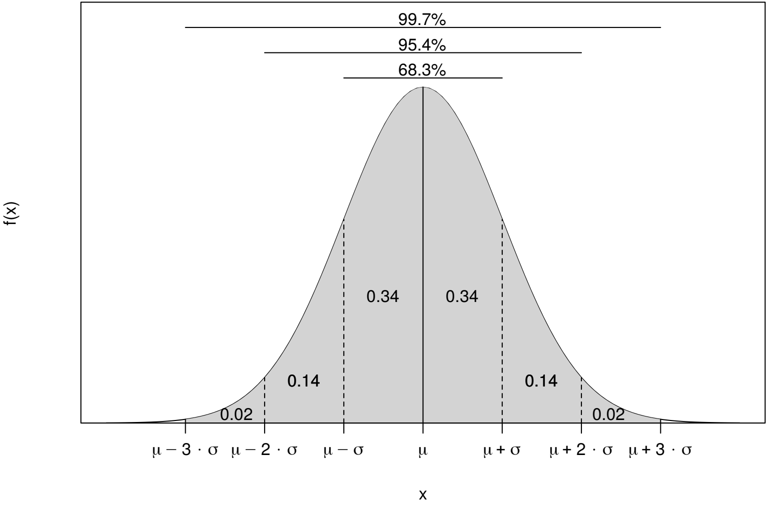
Gaussiano(ou Normal)

Para qualquer VA gaussiana X, valem as seguintes relações:

$$P[X > \mu] = P[X < \mu]$$
$$P[\mu - \sigma < X < \mu + \sigma] \approxeq 0,683$$
$$P[\mu - 2\sigma < X < \mu + 2\sigma] \approxeq 0,954$$
$$P[\mu - 3\sigma < X < \mu + 3\sigma] \approxeq 0,997$$

Portanto, 6σ é frequentemente referida como a **largura** de uma distribuição normal. Métodos mais avançados de integração podem ser utilizados para mostrar que a área sob a função densidade de probabilidade normal de $-\infty < x < \infty$ é igual a 1.

Regra empírica para uma distribuição gaussiana



Fim do módulo 1