Estatística I

Licenciatura em Gestão do Desporto (LGD)

2.º Ano/1.º Semestre

2023/2024

Aulas Teórico-Práticas N.ºs 9 e 10 (Semana 5)

Docente: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/

Conteúdos Programáticos

Aulas TP (Semanas 1 e 3)	Aulas TP (Semanas 3 a 6)	Aulas TP (Semanas 7 a 9)	Aulas TP (Semanas 10 a 12)
• **Capítulo 1**: Análise Descritiva • **Capítulo 2**: Probabilidades	• **Capítulo 3**: Variáveis Aleatórias Unidimensionais	• **Capítulo 4**: Variáveis Aleatórias Multidimensionais	• **Capítulo 5**: Variáveis Aleatórias Especiais

Material didático: Exercícios do Livro Murteira et al (2015), Formulário e Tabelas Estatísticas

Bibliografia: B. Murteira, C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta e F. Pimenta; *Introdução à Estatística*, 2ª ed., Escolar Editora, 2015.

3. Variáveis aleatórias unidimensionais

3.1. Variável aleatória

3.2. Função de distribuição

3.3. Classificação de variáveis aleatórias.

3.4. Variável aleatória discreta

3.5. Variável aleatória contínua

3.6. Função distribuição de uma função de uma variável aleatória

3.7. Valores esperados de variáveis aleatórias

3.8. Valores esperados de funções de variáveis aleatórias

3.9. Propriedades dos valores esperados

3.10. Momentos em relação à origem

3.11. Momentos em relação à média

3.12. Variância de uma variável aleatória

Variáveis Aleatórias Discretas: Exercícios

Murteira et al (2015)

1

Exercício do Livro Murteira et al (2015)

3. Seja

$$F(x)=\begin{cases}0 & (x<1)\\ x/3 & (1\le x<2)\\ 1 & (x\ge 2).\end{cases}$$

a) Verifique que se trata de uma função de distribuição.
b) Classifique a variável aleatória em causa.

Exercício 3 a): V.a. Discreta e Função de Distribuição

P2) F é não decrescente

- para $(x < 1)$ e $(x \geq 2)$, F é não decrescente pois é sempre constante e igual a 0 e 1 respectivamente. ✓

- para $(1 \leq x < 2)$, F é não decrescente, pois $\left(\frac{x}{3}\right)' = \frac{1}{3} > 0$ ✓

P3) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$ ✓

$F(+\infty) = \lim_{x \to +\infty} F(x) = 1 = 1$ ✓

Exercício 3 a): V.a. Discreta e Função de Distribuição

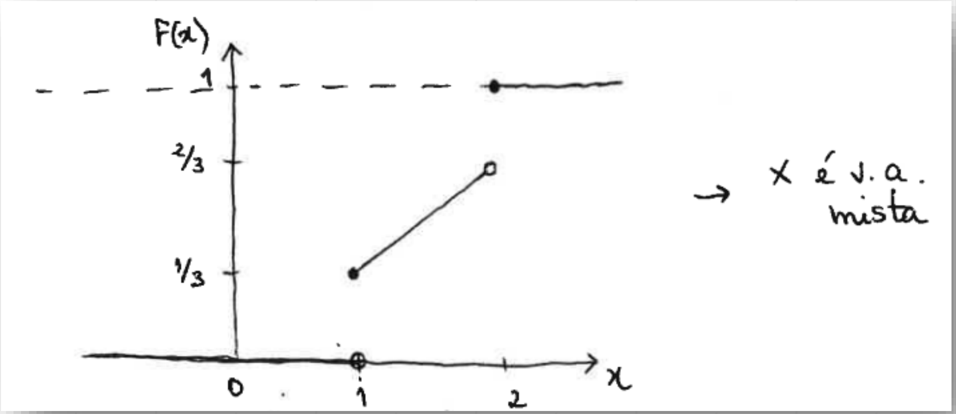
P5) F é contínua à direita : $F(a+0) = F(a)$

Pontos de transição de ramo : $x = 1$ e $x = 2$.

- $x = 1 \Rightarrow F(1+0) = \lim\limits_{x \to 1+0} F(x) = \lim\limits_{x \to 1+0} 1/3 = 1/3 = F(1)$ ✓

- $x = 2 \Rightarrow F(2+0) = \lim\limits_{x \to 2+0} F(x) = \lim\limits_{x \to 2+0} 1 = 1 = F(2)$ ✓

Exercício 3 b): V.a. Discreta e Função de Distribuição



Exercício do Livro Murteira et al (2015)

6. Seja X uma variável aleatória discreta com função de distribuição dada por,

$$F(x) = \begin{cases} 0 & (x < -1) \\ 0.2 & (-1 \le x < 0) \\ 0.7 & (0 \le x < 1) \\ 1 & (x \ge 1). \end{cases}$$

a) Determine a função probabilidade de X.

b) Calcule $P(X \ge 1)$.

c) Calcule $P(X < 0.5 \mid X \ge 0)$.

Exercício 6 a): V.a. Discreta e Função de Probabilidade

$f(x) = F(x) - F(x^-)$

- $x = -1 \rightarrow f(-1) = F(-1) - F(-1^-) = 0.2 - 0 = 0.2$
- $x = 0 \rightarrow f(0) = F(0) - F(0^-) = 0.7 - 0.2 = 0.5$
- $x = 1 \rightarrow f(1) = F(1) - F(1^-) = 1 - 0.7 = 0.3$

Logo,

$$f(x) = \begin{cases} 0.2 & (x = -1) \\ 0.5 & (x = 0) \\ 0.3 & (x = 1) \end{cases}$$

Exercício 6 b): V.a. Discreta e Probabilidades

$$P(X \geq 1) = 1 - P(X < 1) = 1 - F(1^-) = 1 - 0.7 = 0.3$$

Exercício 6 c): V.a. Discreta e Probabilidades

$$P(X<0.5 \mid X\geq 0) = \frac{P(X<0.5 \wedge X\geq 0)}{P(X\geq 0)} = \frac{P(0\leq X<0.5)}{1-P(X<0)} = \frac{F(0.5^-)-F(0^-)}{1-F(0^-)}$$

$$= \frac{0.7-0.2}{1-0.2} = \frac{5}{8} = 0.625$$

Exercício do Livro Murteira et al (2015)

7. O número de automóveis encomendados mensalmente num *stand*, é uma variável aleatória X com a seguinte função probabilidade:

x	0	1	2	3	4
$f(x)$	0.3	0.3	0.2	0.1	0.1

a) Calcule a função de distribuição de X.

b) Determine o número mínimo de automóveis que o *stand* deve ter num mês para que a probabilidade de satisfazer todas as encomendas não seja inferior a 0.75.

c) Num mês em que haja apenas dois automóveis em *stock* no *stand*, calcule a probabilidade de serem todos vendidos, e especifique a distribuição da variável aleatória que representa as vendas nesse mês.

Exercício 7 a): V.a. Discreta e Função de Distribuição

$$F(x) = P(X \le x) = \sum_{x_i \le x} f(x_i)$$

- $x < 0 \longrightarrow F(x) = 0$
- $0 \le x < 1 \longrightarrow F(x) = f(0) = 0.3$
- $1 \le x < 2 \longrightarrow F(x) = f(1) + f(0) = 0.3 + 0.3 = 0.6$
- $2 \le x < 3 \longrightarrow F(x) = f(2) + f(1) + f(0) = 0.2 + 0.3 + 0.3 = 0.8$
- $3 \le x < 4 \longrightarrow F(x) = f(3) + f(2) + f(1) + f(0) = 0.1 + 0.2 + 0.3 + 0.3 = 0.9$
- $x \ge 4 \longrightarrow F(x) = f(4) + f(3) + f(2) + f(1) + f(0) = 0.1 + 0.1 + 0.2 + 0.3 + 0.3 = 1$

Logo,

$$F(x) = \begin{cases} 0 & (x < 0) \\ 0.3 & (0 \le x < 1) \\ 0.6 & (1 \le x < 2) \\ 0.8 & (2 \le x < 3) \\ 0.9 & (3 \le x < 4) \\ 1 & (x \ge 4) \end{cases}$$

Exercício 7 b): V.a. Discreta e Probabilidades

- Stand não tem automóveis $\Rightarrow P(E) = P(X=0) = f(0) = 0.3 < 0.75$
- Stand tem 1 automóvel $\Rightarrow P(E) = P(X \leq 1) = F(1) = 0.6 < 0.75$
- Stand tem 2 automóveis $\Rightarrow P(E) = P(X \leq 2) = F(2) = 0.8 > 0.75$

Logo, para que a probabilidade de satisfazer todas as encomendas não seja inferior a 0.75, o stand deve ter, no mínimo, 2 automóveis num mês.

Exercício 7 c): V.a. Discreta e Probabilidades

$P(V) = P(X \geq 2) = 1 - P(X < 2) = 1 - F(2^-) = 1 - 0.6 = 0.4$

seja,

Y - nº automóveis vendidos num mês com 2 em stock

$$Y = \begin{cases} 0 & (X = 0) \\ 1 & (X = 1) \\ 2 & (X \geq 2) \end{cases} \longrightarrow Y \text{ é v.a. discreta}$$

Exercício 7 c): V.a. Discreta e Probabilidades

Função probabilidade de Y

- $y = 0 \rightarrow f_Y(0) = P(Y=0) = P(X=0) = f_X(0) = 0.3$
- $y = 1 \rightarrow f_Y(1) = P(Y=1) = P(X=1) = f_X(1) = 0.3$
- $y = 2 \rightarrow f_Y(2) = P(Y=2) = P(X \geq 2) = 1 - F_X(2^-) = 1 - 0.6 = 0.4$

Logo,

$$f_Y(y) = \begin{cases} 0.3 & (y=0) \\ 0.3 & (y=1) \\ 0.4 & (y=2) \end{cases}$$

Função distribuição de Y

$$F_Y(y) = \begin{cases} 0 & (y<0) \\ 0.3 & (0 \leq y < 1) \\ 0.6 & (1 \leq y < 2) \\ 1 & (y \geq 2) \end{cases}$$

Exercício do Livro Murteira et al (2015)

9. Calcule o valor de k de modo que as funções dadas nas alíneas seguintes definam funções probabilidade da variável aleatória X.

a) $f(x) = kx \quad (x = 1, 2, \ldots, 10)$;

b) $f(x) = k\left(\frac{1}{5}\right)^x \quad (x = 1, 2, 3, \ldots)$.

Exercício 9 a): V.a. Discreta e Função de Probabilidade

Sabe-se que : $\sum_{x \in D} f(x) = 1$

$f(x) = Kx \quad (x = 1, 2, \dots, 10)$

$\sum_{x \in D} f(x) = 1 \Leftrightarrow \sum_{x=1}^{10} f(x) = 1 \Leftrightarrow \sum_{x=1}^{10} Kx = 1 \Leftrightarrow K \sum_{x=1}^{10} x = 1 \Leftrightarrow$

$\Leftrightarrow K(1+2+\dots+10) = 1 \Leftrightarrow 55K = 1 \Leftrightarrow K = \frac{1}{55} \approx 0.018\dot{2}$

Exercício 9 b): V.a. Discreta e Função de Probabilidade

$$f(x) = K\left(\frac{1}{5}\right)^x \qquad (x = 1, 2, 3, \ldots)$$

$$\sum_{x \in D} f(x) = 1 \Leftrightarrow \sum_{x=1}^{+\infty} K\left(\frac{1}{5}\right)^x = 1 \Leftrightarrow K \underbrace{\sum_{x=1}^{+\infty}\left(\frac{1}{5}\right)^x}_{\text{série geométrica}} = 1 \Leftrightarrow$$

$$\Leftrightarrow K\left(\frac{\frac{1}{5}}{1-\frac{1}{5}}\right) = 1 \Leftrightarrow K\left(\frac{1/5}{4/5}\right) = 1 \Leftrightarrow K\left(\frac{1}{4}\right) = 1 \Leftrightarrow K = 4$$

Variáveis Aleatórias Contínuas

2

Variável Aleatória, Função de Distribuição, Função Densidade de Probabilidade, Valor Esperado, Moda, Variância e Quantis

Variável Aleatória Contínua

Definição 1.1: (Variável aleatória contínua) Uma variável aleatória X é contínua se assume valores num intervalo da reta real, ou seja, o número de valores que X pode assumir é não enumerável.

EXEMPLOS:

- Tempo até a cura de uma doença;
- Peso das peças em uma linha de produção;
- Salário dos estatísticos em João Pessoa.

Probabilidade II (ufpb.br)

Variável Aleatória Contínua

- Uma vez que os valores possíveis de X não são enumeráveis, não podemos falar do i-ésimo valor de X, e, por isso, $p(x_i)$ se torna sem sentido.

- Em vez de atribuir, como no caso discreto, probabilidades aos valores da variável, pode-se atribuir probabilidades a intervalos de valores da variável contínua por meio de uma função.

- A ideia então é substituir a função p definida somente para $x_1, x_2, \ldots$ por uma função f definida para todos os valores de x.

Probabilidade II (ufpb.br)

V.a. Contínua: Função Densidade de Probabilidade (fdp)

Para variáveis aleatórias contínuas introduzimos a **função densidade de probabilidade (fdp),** tal que,

$$(a) \quad f(x) \geq 0$$

$$(b) \quad \int_{-\infty}^{+\infty} f(x)\, dx = 1$$

A condição (a) implica que a densidade é uma função não negativa e condição (b) corresponde ao fato de que a soma (integral no caso de variáveis contínuas) das probabilidades é igual a um. A integração de ±oo significa que devemos integrar sobre todos os valores de x em que f(x) é definida.

Qualquer função f(x) satisfazendo (a) e (b) é uma fdp.

Note que para variáveis contínuas as somas são substituídas por integrais.

Tópico_08_.pdf (usp.br)

Variável Aleatória Contínua: fdp

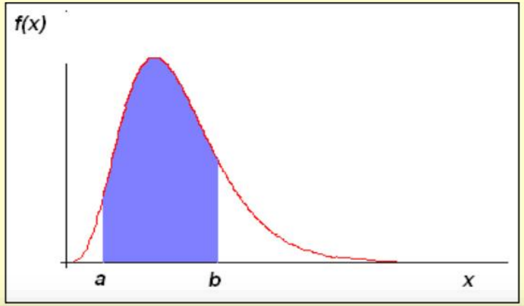
Uma v.a. X contínua é caracterizada por sua ***função densidade de probabilidade $f(x)$***, com as propriedades:

(i) A área sob a curva de densidade é 1;

(ii) $P(a \leq X \leq b)$ = área sob a curva da densidade $f(x)$ e acima do eixo x, entre os pontos a e b;

(iii) $f(x) \geq 0$, para todo x;

(iv) $P(X = x_0) = 0$, para x_0 fixo.



Assim

Note que, como a probabilidade da variável x assumir valor num dado ponto é nula, temos que: $P(a < x < b) = P(a \leq x \leq b) = P(a < x \leq b) = P(a \leq x < b)$

Variável Aleatória Contínua: Probabilidades

Para variáveis contínuas não podemos obter a probabilidade de x ter o valor num ponto $(x = a)$, ou seja, temos que considerar a probabilidade de x assumir valores num intervalo $a < x < b$.

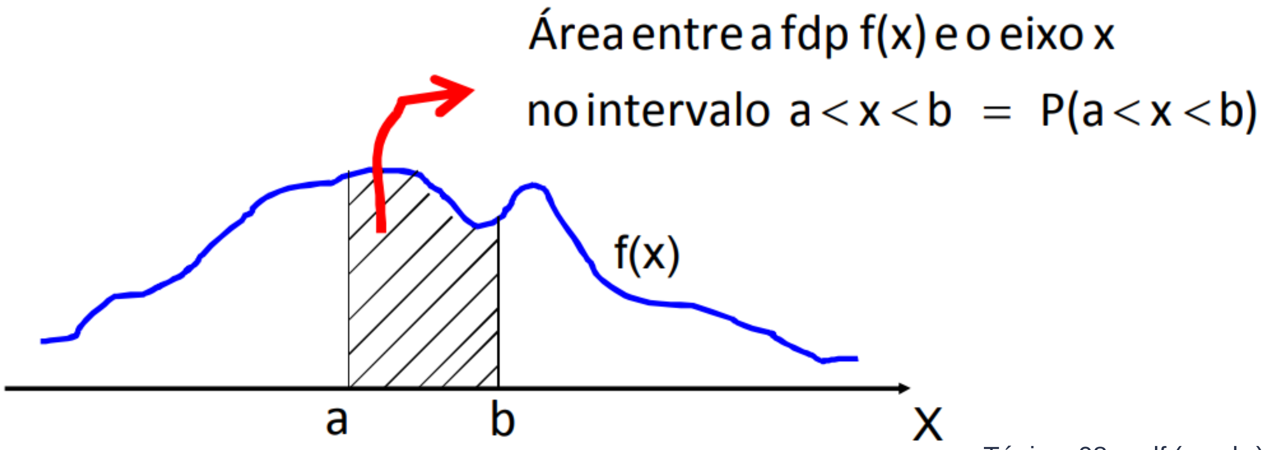
$$P(a < x < b) = \int_a^b f(x)\,dx$$

representa a probabilidade de x assumir valores no intervalo $a < x < b$.

Tópico_08_.pdf (usp.br)

Variável Aleatória Contínua: Probabilidades

Geometricamente



Tópico_08_.pdf (usp.br)

Cálculo de Probabilidades para uma Distribuição Contínua: Exemplo

Exemplo: Dada a função

$$f(x) = \begin{cases} \frac{1}{3}, & 1 \leq x \leq 4 \\ 0, & \text{outros valores} \end{cases}$$

(a) verifique se f(x) é uma fdp.

(b) Calcule a probabilidade de X assumir valores no intervalo $2 < x < 3$

Tópico_08_.pdf (usp.br)

Cálculo de Probabilidades para uma Distribuição Contínua: Exemplo

(a) Temos que $f(x) \geq 0$ e $\int_{-\infty}^{+\infty} f(x)\,dx = \int_1^4 \frac{1}{3}\,dx = \frac{1}{3}\,x\Big|_1^4 = \frac{1}{3}(4\text{-}1) = 1$

portanto f(x) é uma fdp.

(b) Temos: $P(2 < x < 3) = \int_2^3 \frac{1}{3}\,dx = \frac{1}{3}\,x\Big|_2^3 = \frac{1}{3}(3\text{-}2) = \frac{1}{3}$

Tabela 1.1: Tabela de Primitivas Elementares

f	$Pf=F$
$c, c \in IR$	$c\,x$
x^α $(\alpha \neq -1)$	$\frac{x^{\alpha+1}}{\alpha+1}$
$\frac{1}{x}$	$\log\lvert x\rvert$
e^x	e^x
$\cos x$	$\sin x$
$\sin x$	$-\cos x$
$\sec^2 x$	$\text{tg}x$
$\text{cosec}^2 x$	$-\text{cotg}x$
$\frac{1}{1+x^2}$	$\text{arctg}\,x$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$
$\cosh x$	$\sinh x$
$\sinh x$	$\cosh x$

Tópico_08_.pdf (usp.br)

Média, Variância e Desvio-Padrão para V.a. Contínua

Do mesmo modo que para variáveis discretas, podemos definir para variáveis contínuas a média, a variância e o desvio padrão. Lembrando que introduzimos a fdp associada a variável aleatória considerada e substituímos as somas por integrais.

Assim, para a variável aleatória X, temos:

Média (μ): $\mu = \int_{-\infty}^{+\infty} x\, f(x)\, dx$

Variância (σ^2): $\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2\, f(x)\, dx$

Desvio-padrão: $\sigma = \sqrt{\sigma^2}$

$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X)$

[Tópico_08_.pdf (usp.br)]

Média, Variância e Desvio-Padrão para V.a. Contínua: Exemplo

Exemplo: Considere a fdp considerada anteriormente

$$f(x) = \begin{cases} \frac{1}{3}, & 1 \leq x \leq 4 \\ 0, & \text{outros valores} \end{cases}$$

(a) Calcule a média da variável aleatória X.

(b) Calcule a variância e o desvio padrão de X.

Tópico_08_.pdf (usp.br)

Média, Variância e Desvio-Padrão para V.a. Contínua: Exemplo

(a) Temos: $\mu = \int_{-\infty}^{+\infty} x\, f(x)\, dx = \int_{1}^{4} x\left(\frac{1}{3}\right) dx = \frac{1}{3}\int_{1}^{4} x\, dx = \frac{1}{3}\left(\frac{x^2}{2}\right)\Big|_{1}^{4} = \frac{1}{6}(4^2 - 1^2) = \frac{15}{6} = \frac{5}{2}$

(b) Temos: $\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2\, f(x)\, dx = \int_{1}^{4} (x - \frac{5}{2})^2\left(\frac{1}{3}\right) dx = \frac{1}{3}\int_{1}^{4} (x - \frac{5}{2})^2\, dx = \frac{1}{3}\frac{\left(x - \frac{5}{2}\right)^3}{3}\Bigg|_{1}^{4} = \frac{3}{4}$

e assim: $\sigma = \sqrt{\sigma^2} = \frac{\sqrt{3}}{2}$

Tópico_08_.pdf (usp.br)

Variável Aleatória Contínua: Função de Distribuição (fd)

Definição 1.2: (Variável aleatória absolutamente contínua) Uma variável aleatória X é absolutamente contínua se existir uma função não negativa f tal que para todo $x \in \mathbb{R}$,

$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

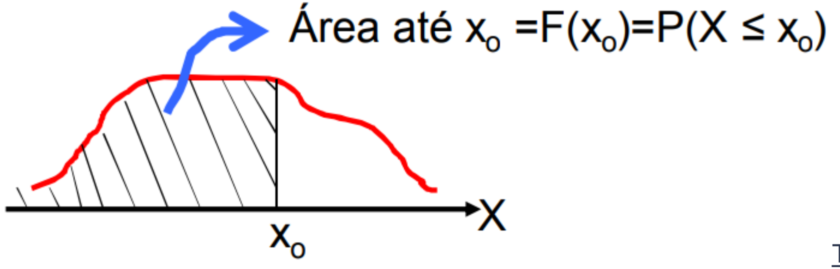
em que F é a função de distribuição acumulada e f é a função densidade da variável aleatória X.

Variável Aleatória Contínua: fd

A função distribuição acumulada (fda) F(x) é a probabilidade de que $X \leq x_o$, ou seja,

$$F(x_o) = P(X \leq x_o) = \int_{-\infty}^{x_o} f(x)\,dx$$

F(x) é análoga a distribuição de frequências relativas acumuladas (ou % acumuladas) estudadas no início do curso



[Tópico_08_.pdf (usp.br)]

Variável Aleatória Contínua: Probabilidades

IMPORTANTE: A partir da definição de função de distribuição, tem-se que

$$P(a \leq X \leq b) = \int_a^b f(x)dx = F(b) - F(a).$$

Note que a integral acima não se altera com a inclusão ou não dos extremos *a* e *b*. Dessa forma, para as variáveis contínuas, a probabilidade da variável ser igual a um particular valor é zero.

Assim,

$$P(X(\omega) \in [a, b]) = P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b).$$

$P(a \leq X \leq b)$ representa a área sob a curva da função densidade entre *a* e *b*.

Variável Aleatória Contínua: fdp vs fd

IMPORTANTE:A função de densidade serve para a caracterização da variável contínua. Dada a função de densidade, a função de distribuição é obtida por integração. Por outro lado, derivando a função de distribuição, obtemos a densidade.

Ou seja, utilizando o teorema fundamental de Cálculo, se $f(x)$ for contínua, temos a seguinte relação

$$\frac{d}{dx}F(x) = f(x)$$

fdp vs fd: Exemplo 1

Encontre a função de distribuição acumulada da seguinte função de densidade de probabilidade

$$f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \textit{para quaisquer outros valores} \end{cases}$$

fdp vs fd: Exemplo 1

$$f(x)=\begin{cases}2x, & 0<x<1,\\ 0, & \textit{para quaisquer outros valores}\end{cases}$$

$$F(x)=\begin{cases}0, & se\ x<0,\\ \int_0^x (2s)ds = x^2, & se\ 0\leq x<1,\\ 1, & se\ x\geq 1\end{cases}$$

fdp vs fd: Exemplo 2

Arqueólogos estudaram uma certa região e mediram o *comprimento de fósseis* encontrados (em cm). Chamamos de **C** a v.a. contínua comprimento de fósseis. Suponha que C possui a função densidade de probabilidade:

$$f(c) = \begin{cases} \frac{1}{40}\left(\frac{c}{10}+1\right) & \text{se} \quad 0 \le c \le 20; \\ 0 & \text{caso contrário.} \end{cases}$$

Qual a probabilidade de um fóssil, escolhido ao acaso nessa região, apresentar comprimento inferior a 8 cm?

fdp vs fd: Exemplo 2

$$f(c)=\begin{cases}\dfrac{1}{40}\left(\dfrac{c}{10}+1\right) & \text{se}\quad 0\le c\le 20;\\ 0 & \text{caso contrário.}\end{cases}$$

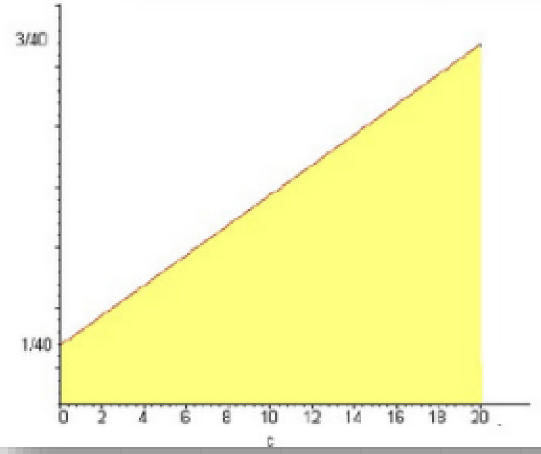
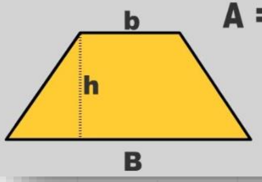


Gráfico da função densidade de probabilidade

ÁREA DO TRAPÉZIO



$$A=\frac{(B+b).h}{2}$$

A **área** do **trapézio** é a soma das bases vezes a altura dividido por dois.

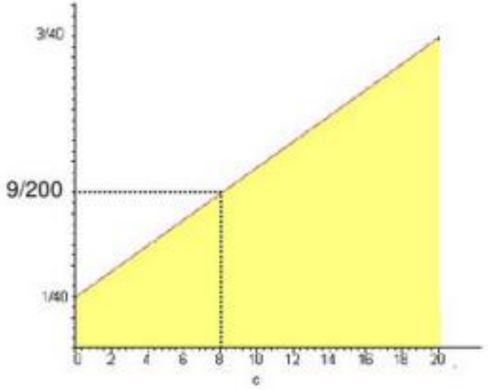
$$\text{Área sob } f(c)=\frac{(b+B)h}{2}=\frac{\left(\frac{1}{40}+\frac{3}{40}\right)20}{2}=1$$

$$\text{Área sob } f(c)=\int_0^{20}\frac{1}{40}\left(\frac{c}{10}+1\right)=1$$

- Como f(c) é positiva e a área é igual 1, podemos concluir que f(c) é efetivamente uma densidade.

fdp vs fd: Exemplo 2

- Qual a probabilidade de um fóssil, escolhido ao acaso nessa região, apresentar comprimento inferior a 8 cm?



$$f(8) = \frac{9}{200}$$

$$P(C < 8) = \frac{\left(\frac{1}{40} + \frac{9}{200}\right)8}{2} = \frac{7}{25}$$

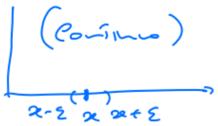
$$P(C < 8) = \int_0^8 f(c)dc = \frac{7}{25}$$

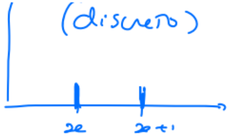
V.a. Discreta vs V.a. Contínua

Seja X uma v.a. contínua, que toma valores num intervalo de $\mathbb{R}$ (ou em $\mathbb{R}$)

i) $P(X=x) = 0, \quad \forall x$ [1ª grande diferença em relação às v.a. discretas]

Em vez de função de probabilidade, vamos definir:

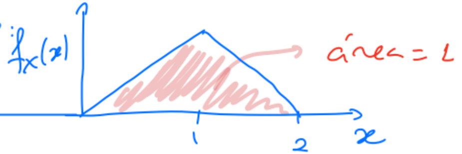




$$\boxed{f_X(x)} = \lim_{\varepsilon \to 0} \frac{P(x - \varepsilon/2 \le X \le x + \varepsilon/2)}{\varepsilon}$$

↳ função densidade de probabilidade

Exemplo:



Slides Professora Claúdia Nunes

V.a. Discreta vs V.a. Contínua

v.a. discreta

propriedades : $P(X=x) \in [0,1]$

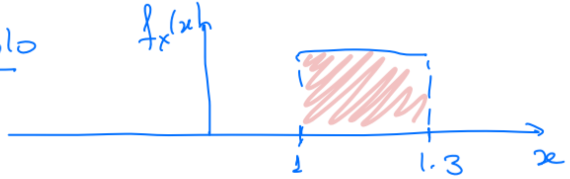
$$\sum_{x=-\infty}^{+\infty} P(X=x) = 1$$

v.a. contínua

propriedades : • $f_X(x) \geq 0, \forall x$

• $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$

Exemplo

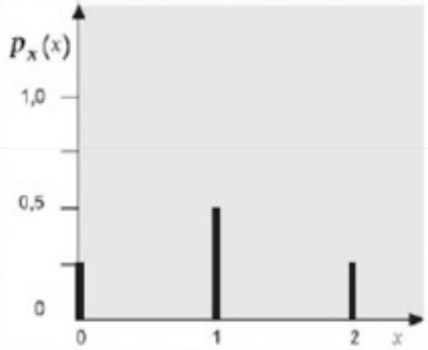


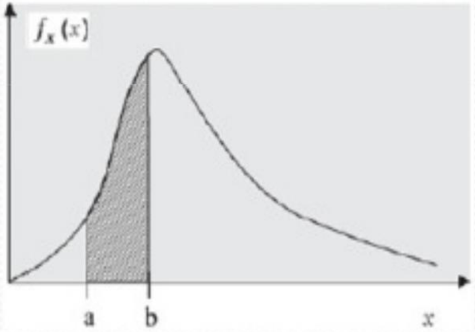
$$\text{área} = (1.3 - 1)\,k = 1 \Rightarrow 0.3k = 1 \Rightarrow$$

$$k = \frac{1}{0.3} = 3.33\ldots$$

V.a. Discreta vs V.a. Contínua

- **Função massa de probabilidade**
  - Indica com que probabilidade a variável aleatória x assume o valor x_o, *i.e.*, $P(x=x_o) = p_x(x_o)$.
  - A função massa de probabilidade se aplica a variáveis discretas.
- **Função densidade de probabilidade**
  - Equivale à função massa de probabilidade, sendo que se aplica a variáveis contínuas.





V.a. Discreta vs V.a. Contínua

$$P(X=x) \leftrightarrow f_X(x)$$

$$\Sigma \leftrightarrow \int$$

V.a. Discreta

- $F_X(x) = P(X \le x) = \sum_{y=-\infty}^{x} P(X=y)$
  - $F_X(x) \in [0,1]$
  - $\lim_{x \to -\infty} F_X(x) = 0$
  - $\lim_{x \to +\infty} F_X(x) = 1$
  - $P(X=x) = F_X(x) - F_X(x^-)$
  - Não decrescente
  - $P(a < X \le b) = F_X(b) - F_X(a)$
  - $P(a \le X \le b) = F_X(b) - F_X(a^-)$
  - $P(a \le X < b) = F_X(b^-) - F_X(a^-)$
  - $P(a < X < b) = F_X(b^-) - F_X(a)$

$=]-\infty, b] -]-\infty, a[$

V.a. Contínua

- $F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(y)\,dy$
  - $F_X(x) \in [0,1]$
  - $\lim_{x \to -\infty} F_X(x) = 0$
  - $\lim_{x \to +\infty} F_X(x) = 1$
  - A função F_X é contínua
    $f_X(x) = F_X'(x)$
  - Não decrescente
  - $P(a < X \le b) = P(a \le X \le b)$
    $= P(a \le X < b) = P(a < X < b)$
    $= F_X(b) - F_X(a)$
    (porque $P(X=a) = P(X=b) = 0$)

V.a. Discreta vs V.a. Contínua

V.a. discretas

- $E[X] = \sum_{x=-\infty}^{+\infty} x\, P(X=x)$

$$E[aX+b] = aE[X]+b$$

- $E[g(X)] = \sum_{x=-\infty}^{+\infty} g(x)\, P(X=x)$

$$Var[X] = E[X^2] - E^2(X)$$

- $var(X) = \sum_{x=-\infty}^{+\infty} x^2 P(X=x) - E^2(X)$

$= \sum_{x=-\infty}^{+\infty} x^2 P(X=x) - \left[\sum_{x=-\infty}^{+\infty} x\, P(X=x)\right]^2$

$$Var[aX+b] = a^2 Var(X)$$

- mediana = me

$0.5 \le F_X(me) \le 0.5 + P(X=me)$

V.a. contínuas

- $E[X] = \int_{-\infty}^{+\infty} x\, f_X(x)\, dx$
- $E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x)\, dx$
- $var(X) = \int_{-\infty}^{+\infty} x^2 f_X(x)\, dx - E^2(X)$

$= \int_{-\infty}^{+\infty} x^2 f_X(x)\, dx - \left[\int_{-\infty}^{+\infty} x f_X(x)\, dx\right]^2$

- mediana = me

$F_X(me) = 0.5$

- percentil de prob. $p = x_p$

$F_X(x_p) = p \qquad p \in (0,1)$

Slides Professora Claúdia Nunes

Características da Distribuição

Localização/Tendência

- **Central**: Moda, Mediana, Média
- **Não central ou relativa**: Alguns Quantis (Quartis, Decis, Percentis, Mínimo, Máximo)

Dispersão

- Amplitude Total ou Amostral, Amplitude Interquartil, Variância, Desvio Padrão, Coeficiente de Variação

Assimetria

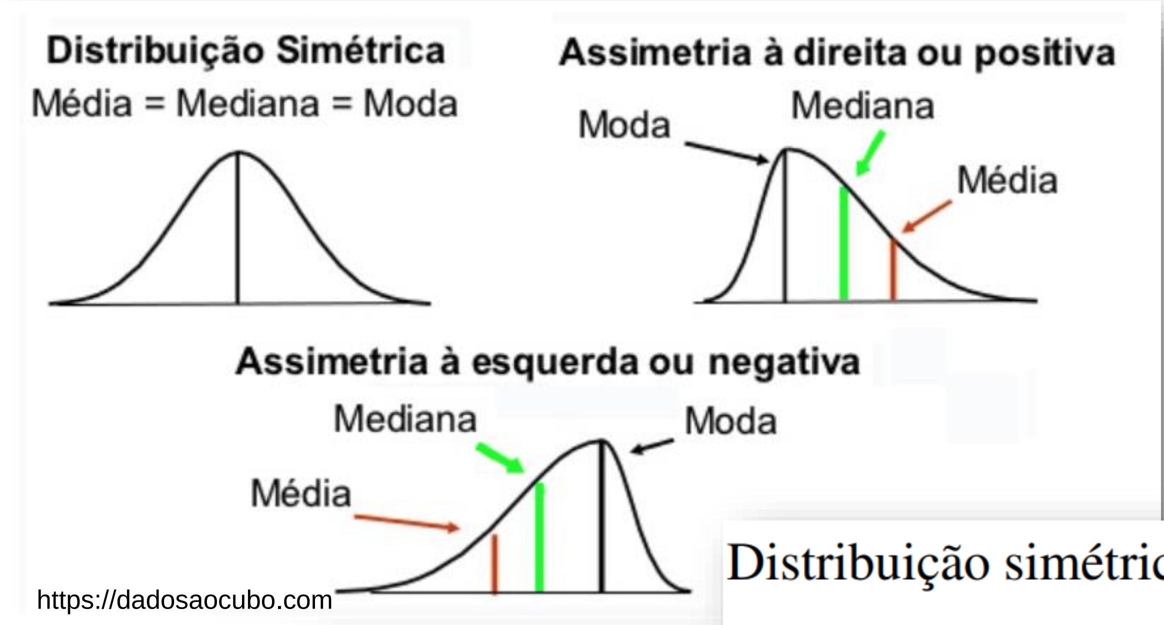
- Coeficientes de Assimetria

Achatamento/Curtose/Forma

- Coeficientes de Achatamento

Para avaliar o peso relativo da dispersão face à localização, utiliza-se o **coeficiente de variação**: $\text{CV} = \dfrac{\sigma}{\mu}$ (utiliza-se sobretudo se o suporte de $X \subset \Re^{+}$)

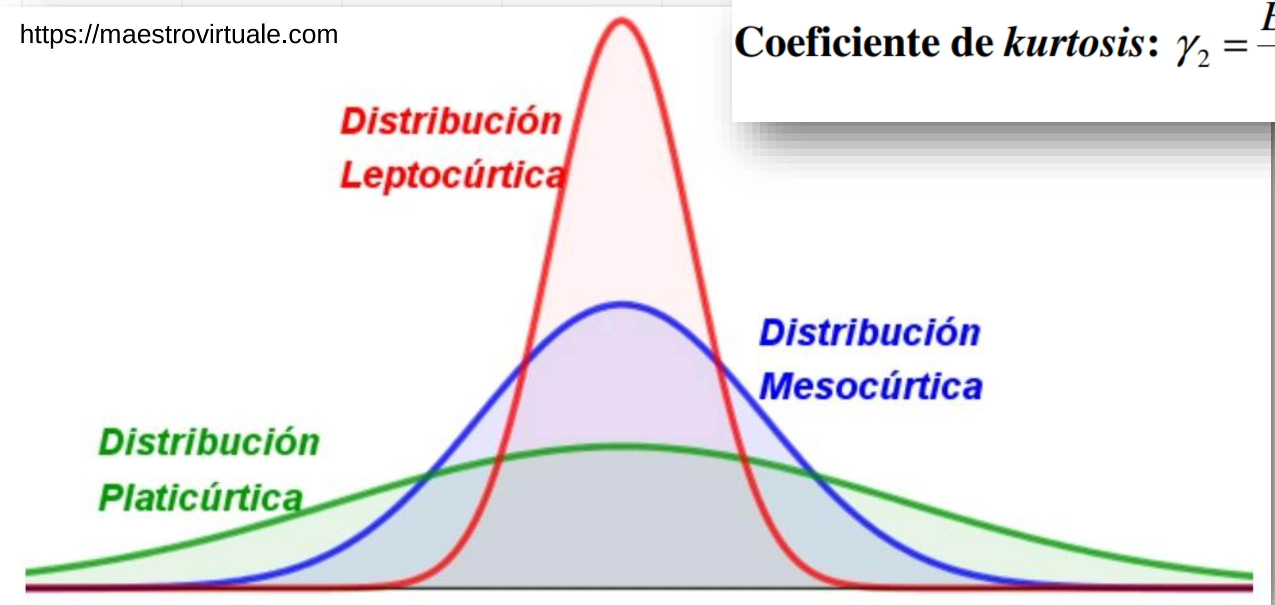
Assimetria/Enviesamento



$$\text{Distribuição simétrica} \rightarrow \gamma_1 = 0$$

$$\textbf{Coeficiente de assimetria: } \gamma_1 = \frac{E(X-\mu)^3}{\sigma^3} = \frac{\mu_3}{\sigma^3}$$

Achatamento/Curtose/Forma



Coeficiente de *kurtosis*: $\gamma_2 = \dfrac{E(X-\mu)^4}{\sigma^4} = \dfrac{\mu_4}{\sigma^4}$

o **grau de afastamento de uma certa distribuição em relação à distribuição normal**. Nesta distribuição tem-se $\gamma_1 = 0$, pois trata-se de uma distribuição simétrica, e $\gamma_2 = 3$.

Assimetria e Achatamento

Assimetria e curtose

A assimetria γ_1, a curtose β_2 e a curtose normalizada γ_2 são obtidas a partir das fórmulas dos momentos

$$\begin{cases} \gamma_1 = \dfrac{\mu_3}{\sigma^3} = 0 \\ \beta_2 = \dfrac{\mu_4}{\sigma^4} = 3 \\ \gamma_2 = \beta_2 - 3 = 0 \end{cases}.$$ [49]

A distribuição normal é um ponto de referência para comparação das espessuras de caudas longas. Se uma distribuição possui uma curtose normalizada $\gamma_2 > 0$, então a distribuição possui uma cauda longa mais grossa que a distribuição normal e é chamada leptocúrtica. Se $\gamma_2 < 0$, a distribuição possui uma cauda longa mais fina que a distribuição normal e é chamada platicúrtica. Se a distribuição possui uma curtose normalizada nula, então a distribuição possui uma cauda longa comparável à distribuição normal e é chamada mesocúrtica.[50][51]

Distribuição normal – Wikipédia, a enciclopédia livre (wikipedia.org)

Obrigada!

Questões?