Lisbon School of Economics & Management
Universidade de Lisboa

Estatística I

Licenciatura em Gestão
2.º Ano/1.º Semestre
2023/2024

Aulas Teóricas 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 Teóricas (Semanas 1 a 3)	Aulas Teóricas (Semanas 3 a 5)	Aulas Teóricas (Semanas 5 a 7)	Aulas Teóricas (Semanas 8 a 13)
• **Capítulo 2**: Probabilidades	• **Capítulo 3**: Variáveis Aleatórias Unidimensionais	• **Capítulo 4**: Variáveis Aleatórias Multidimensionais	• **Capítulo 5**: Distribuições Teóricas • **Capítulo 6**: Amostragem. Distribuições por Amostragem.

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.

https://cas.iseg.ulisboa.pt

Aula 4	**Início do Capítulo 2**: variável aleatória, definição e exemplos. Função de distribuição. Propriedades.
Aula 5	Classificação de variáveis aleatórias. V.a. discreta: Função probabilidade: propriedades e exemplos. V.a. contínua: Função densidade de probabilidade: propriedades e exemplos.
Aula 6	V.a. mistas, exemplo. Funções de uma v.a.: método para v.a. discretas e método geral. Exemplos.
Aula 7	Funções de uma v.a. Valor esperado de uma v.a. discreta e valor esperado de uma v.a. contínua. Exemplos. Valor esperado de uma função de uma v.a.: caso discreto.

Aula 9	Coeficiente de assimetria e coeficiente de kurtosis. Exemplo. Quantil de ordem alfa para distribuições contínuas. Mediana como medida de localização, amplitude inter-quartis como medida de dispersão. Exemplo.
Aula 10	**Início do capítulo 3:** Variáveis aleatórias bidimensionais. Função de distribuição conjunta, propriedades. Funções de distribuição marginais. Independência de variáveis aleatórias. Variáveis aleatórias bidimensionais discretas. Função probabilidade conjunta.

1

Variáveis Aleatórias Discretas: Exercícios

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

Exercício Suplementar que não consta do livro Murteira et al (2015)

3.2 Numa fábrica existem três máquinas iguais de uma mesma marca, que trabalham independentemente. A probabilidade de cada máquina avariar num dado espaço de tempo é 0.1. Seja X a variável aleatória que representa o número de máquinas que findo esse período de tempo estão a trabalhar. Determine:

(a) A função de probabilidade de X.

(b) A função de distribuição de X.

(c) O valor esperado, moda, mediana e variância de X.

Exercício 3.2 (a): Variável Aleatória

- **Experiência aleatória**

  Classificação de 3 máquinas (que funcionam de forma independente) quanto a estarem avariadas (A) ou não ($\overline{A}$).

- **Eventos chave**

  A = máquina avariada

  $\overline{A}$ = máquina a trabalhar

- **Probabilidades**

  $P(A) = 0.1$

  $P(\overline{A}) = 1 - P(A) = 1 - 0.1 = 0.9$

- **Espaço de resultados**

  $\Omega = \{A\,A\,A, \quad \overline{A}\,A\,A, A\,\overline{A}\,A, A\,A\,\overline{A}, \quad \overline{A}\,\overline{A}\,A, \overline{A}\,A\,\overline{A}, A\,\overline{A}\,\overline{A}, \quad \overline{A}\,\overline{A}, \overline{A}\}.$

  $A\,A\,A \equiv A_1 \cap A_2 \cap A_3$ (1a., 2a. e 3a. máquinas avariadas)

  etc.

  $\#\Omega = 2^3 = 8$ (eventos elementares)

- **V.a. de interesse**

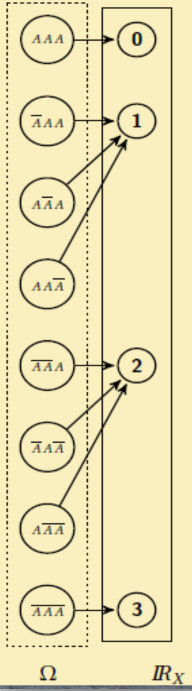
  X = número de máquinas que findo o período estão a trabalhar

  $X : \Omega \to \mathbb{R}_X$

Exercício 3.2 (a): Contradomínio da V.a.

- **Contradomínio e representação esquemática de X**

  Atendendo ao número de máquinas, os valores possíveis de X são $0, 1, 2, 3$, i.e., $\mathbb{R}_X = \{0, 1, 2, 3\}$.



Exercício 3.2 (a): Função Massa de Probabilidade

$$
\begin{aligned}
P(X=0) &= P(\text{nenhuma máquina a trabalhar}) \\
&= P(\text{três máquinas avariadas}) \\
&= P\left(A_1 \cap A_2 \cap A_3\right) \\
&= P(A_1) \times P(A_2) \times P(A_3) && \text{[eventos completamente independentes]} \\
&= 0.1^3 && \text{[eventos equiprováveis]} \\
&= 0.001
\end{aligned}
$$

$$
\begin{aligned}
P(X=1) &= P(\text{1 máquina a trabalhar}) \\
&= P(\text{duas máquinas avariadas}) \\
&= P\left(A_1 \cap A_2 \cap \overline{A_3}\right) + P\left(A_1 \cap \overline{A_2} \cap A_3\right) + P\left(\overline{A_1} \cap A_2 \cap A_3\right) \\
&= P(A_1) \times P(A_2) \times P(\overline{A_3}) + P(A_1) \times P(\overline{A_2}) \times P(A_3) + P(\overline{A_1}) \times P(A_2) \times P(A_3) \\
& && \text{[eventos completamente independentes]} \\
&= 3 \times (1-0.1) \times 0.1^2 && \text{[eventos equiprováveis]} \\
&= 0.027
\end{aligned}
$$

Exercício 3.2 (a): Função Massa de Probabilidade

$$
\begin{aligned}
P(X=2) &= P(2 \text{ máquinas a trabalhar}) \\
&= P(1 \text{ máquina avariada}) \\
&= P\left(A_1 \cap \overline{A_2} \cap \overline{A_3}\right) + P\left(\overline{A_1} \cap A_2 \cap \overline{A_3}\right) + P\left(\overline{A_1} \cap \overline{A_2} \cap A_3\right) \\
&= P(A_1) \times P(\overline{A_2}) \times P(\overline{A_3}) + P(\overline{A_1}) \times P(A_2) \times P(\overline{A_3}) + P(\overline{A_1}) \times P(\overline{A_2}) \times P(A_3) \\
& \hspace{6cm} [\text{eventos completamente independentes}] \\
&= 3 \times 0.1 \times (1-0.1)^2 \times \qquad [\text{eventos equiprováveis}] \\
&= 0.243
\end{aligned}
$$

$$
\begin{aligned}
P(X=3) &= P\left(\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}\right) \\
&= P(\overline{A_1}) \times P(\overline{A_2}) \times P(\overline{A_3}) \qquad [\text{eventos completamente independentes}] \\
&= (1-0.1)^3 \qquad [\text{eventos equiprováveis}] \\
&= 0.729.
\end{aligned}
$$

Exercício 3.2 (a): Função Massa de Probabilidade

Seja X uma variável aleatória discreta assumindo valores em R_X = {x1, x2, ...}

A função (massa de probabilidade) de X é dada por

$$f_X(x) = P(X = x) = \begin{cases} P(X = x_i), & x = xi \in R_X \\ 0, & caso\ contrário \end{cases}$$

- **F.p. de** X (cont.)

$$P(X = x) = \begin{cases} 0.001, & x = 0 \\ 0.027, & x = 1 \\ 0.243, & x = 2 \\ 0.729, & x = 3 \\ 0, & \text{outros valores de x} \end{cases}$$

- **Obs.** — Alternativamente, poderíamos tirar partido do facto de $X \sim \text{Binomial}(n = 3, p = 0.9)$, logo

$$P(X = x) \overset{form.}{=} \binom{3}{x} 0.9^x (1 - 0.9)^{3-x},\ x = 0.1, 2, 3.$$

Escusado será dizer que obteríamos o mesmo resultado.

$0 \leq f_X(x) \leq 1, \forall\, x \in \mathcal{R}$

$$\sum_{x \in \mathcal{R}} f_X(x) = 1$$

A distribuição Binomial será abordada mais à frente!!!

Exercício 3.2 (b): Função de Distribuição

- **F.d. de X**

Comecemos por preencher a tabela abaixo com alguns valores da f.d. de X.

x	$F_X(x) = P(X \le x) = \sum_{x_i \le x} P(X = x_i)$
-0.5	$F_X(-0.5) = P(X \le -0.5) = 0$
0	$F_X(0) = P(X \le 0) = P(X = 0) = 0.001$
0.3	$F_X(0.3) = P(X \le 0.3) = P(X = 0) = 0.001$
1	$F_X(1) = P(X \le 1) = P(X = 0) + P(X = 1) = 0.001 + 0.027 = 0.028$
1.4	$F_X(1.4) = P(X \le 1.4) = P(X \le 1) = 0.028$
2	$F_X(2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.001 + 0.027 + 0.243 = 0.271$
2.8	$F_X(2.8) = P(X \le 2) = 0.271$
3	$F_X(3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$ $= 0.001 + 0.027 + 0.243 + 0.729 = 1$
10.5	$F_X(10.5) = P(X \le 3) = 1$

Podemos então concluir que:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ 0.001, & 0 \le x < 1 \\ 0.028, & 1 \le x < 2 \\ 0.271, & 2 \le x < 3 \\ 1, & x \ge 3. \end{cases}$$

Exercício 3.2 (c): Valor Esperado e Moda

- **Valor esperado de** X

$$\begin{aligned} E(X) &= \sum_{x=0}^{3} x \times P(X=x) \\ &= 0 \times 0.001 + 1 \times 0.027 + 2 \times 0.243 + 3 \times 0.729 \\ &= 2.7 \end{aligned}$$

$$P(X=x) = \begin{cases} 0.001, & x=0 \\ 0.027, & x=1 \\ 0.243, & x=2 \\ 0.729, & x=3 \\ 0, & \text{outros valores de x} \end{cases}$$

- **Moda de** X

  Represente-se a moda de X por $mo(X)$. Então

$$mo(X): \quad P[X = mo(X)] = \max_{x \in \{0,1,2,3\}} P(X=x).$$

  Atendendo a que $P(X=3) = 0.729$ é superior a qualquer dos restantes valores da f.p. de X, temos que o valor mais frequente de X é efectivamente $mo(X) = 3$.

Exercício 3.2 (c): Mediana

- **Mediana de X**

Represente-se a mediana de X por $me(X)$. Então

$$me(X) \quad : \quad \frac{1}{2} \leq F_X[me(X)] \leq \frac{1}{2} + P[X = me(X)] \qquad (1)$$

$$F_X[me(X)^-] \leq \frac{1}{2} \leq F_X[me(X)]. \qquad (2)$$

Ora, tirando partido da definição de mediana em (1) e do facto de

$$\frac{1}{2} \leq F_X(3) \overset{(a)}{=} 1 \quad \leq \quad \frac{1}{2} + P(X=3) = \frac{1}{2} + 0.729 = 1.229,$$

concluímos que $me(X) = 3$.

Em alternativa, notemos que

$$F_X(3) = P(X \leq 3) \overset{(a)}{=} 1 \geq \frac{1}{2};$$

mais, de uma consulta da f.d. de X tem-se

Podemos então concluir que:

$$F_X(x) \quad = \quad \begin{cases} 0, & x < 0 \\ 0.001, & 0 \leq x < 1 \\ 0.028, & 1 \leq x < 2 \\ 0.271, & 2 \leq x < 3 \\ 1, & x \geq 3. \end{cases}$$

Exercício 3.2 (c): Mediana

$$F_X(3^-) = F_X(2) = 0.271 \leq \frac{1}{2}.$$

Logo

$$F_X(3^-) \leq \frac{1}{2} \leq F_X(3),$$

pelo que o resultado (2) leva-nos a concluir que $me(X) = 3$.

Podemos então concluir que:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ 0.001, & 0 \leq x < 1 \\ 0.028, & 1 \leq x < 2 \\ 0.271, & 2 \leq x < 3 \\ 1, & x \geq 3. \end{cases}$$

Exercício 3.2 (c): Variância

- **Variância de X**

  Como o valor esperado de X é igual a $E(X) = 2.97$ e o 2o. momento de X é dado por

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{3} x^2 \times P(X=x) \\
&= 0^2 \times 0.01 + 1^2 \times 0.027 + 2^2 \times 0.243 + 3^2 \times 0.729 \\
&= 7.56,
\end{aligned}
$$

  segue-se que

$$
\begin{aligned}
V(X) &= E(X^2) - E^2(X) \\
&= 7.56 - 2.7^2 \\
&= 0.27.
\end{aligned}
$$

$$
\boxed{P(X=x) = \begin{cases} 0.001, & x=0 \\ 0.027, & x=1 \\ 0.243, & x=2 \\ 0.729, & x=3 \\ 0, & \text{outros valores de x} \end{cases}}
$$

- **Obs.** — Alternativamente, poderíamos ter tirado partido do facto de $X \sim \text{Binomial}(n=3, p=0.9)$ e concluir que

$$
\begin{aligned}
E(X) &\stackrel{form.}{=} n\,p = 3 \times 0.9 = 2.7 \\
V(X) &\stackrel{form.}{=} n\,p\,(1-p) = 3 \times 0.9 \times (1-0.9) = 0.27.
\end{aligned}
$$

2

Variáveis Aleatórias Contínuas

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

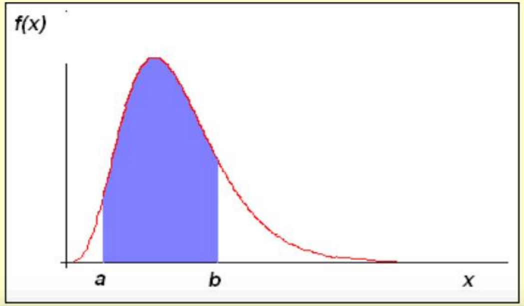
Uma v.a. X contínua é caracterizada por sua ***função densidade de probabilidade f(x)***, com as propriedades:

(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

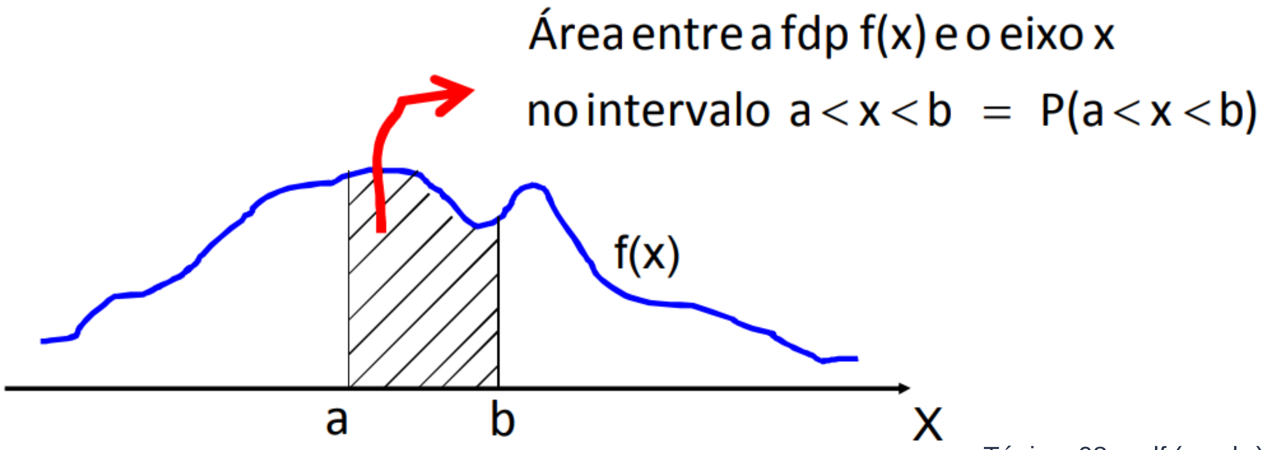
Para variáveis contínuas não podemos obter a probabilidade de x ter o valor num ponto ($x = a$), ou seja, temos que considerar a probabilidade de x assumir valores num intervalo $a < x < b$.

$$P(a < x < b) = \int_a^b f(x)\,dx$$ 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 $\int_{-\infty}^{+\infty} f(x)\,dx = \int_1^4 \frac{1}{3}\,dx = \frac{1}{3}\,x\Big|_1^4 = \frac{1}{3}(4-1) = 1$

portanto f(x) é uma fdp.

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

Tabela 1.1: Tabela de Primitivas Elementares

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

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.

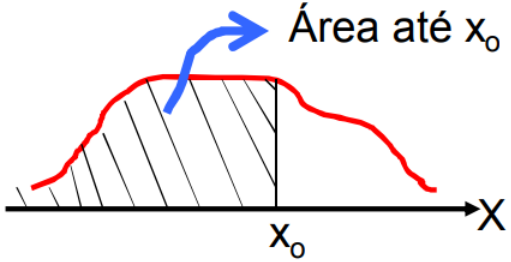
Variável Aleatória Contínua: fd

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

Área até x_o $= F(x_o) = P(X \leq x_o)$



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

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

Probabilidade II (ufpb.br)

fdp vs fd: Exemplo 1

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

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

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:

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

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

Probabilidade II (ufpb.br)

fdp vs fd: Exemplo 2

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

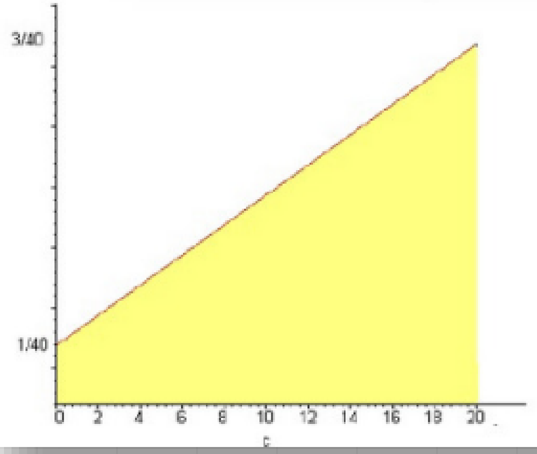
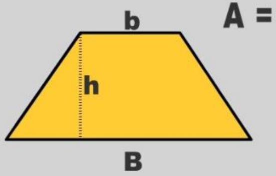


Gráfico da função densidade de probabilidade

ÁREA DO TRAPÉZIO



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

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

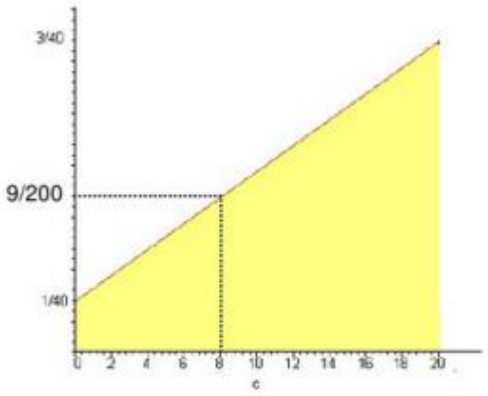
$$\text{Área sob } f(c)=\frac{(b+B)h}{2}=\frac{\left(\frac{1}{40}+\frac{3}{40}\right)20}{2}=1$$

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

- 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 $\mathbb{R}$),

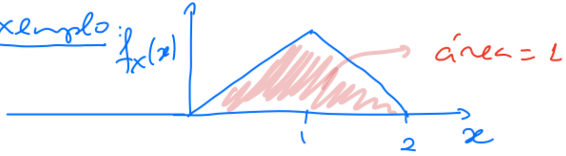
i) $P(X=x) = 0, \quad \forall\, x$ [1ª grande diferença em relação às v.a. discretas]

Em vez da 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:



V.a. Discreta vs V.a. Contínua

v.a. discreta

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

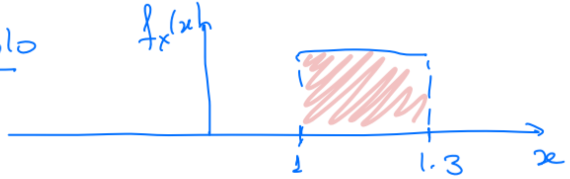
$$\sum_{x=-\infty}^{+\infty} P(X=x) = 1$$

v.a. contínua

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

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

Exemplo



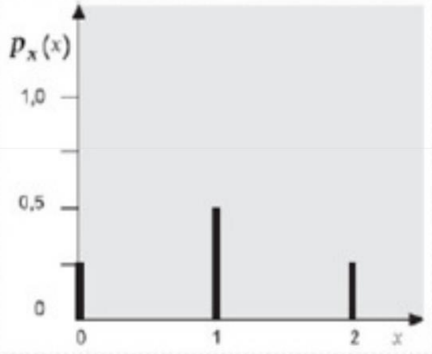
área $= (1.3 - 1)\,k = 1 \Rightarrow 0.3k = 1 \Rightarrow$

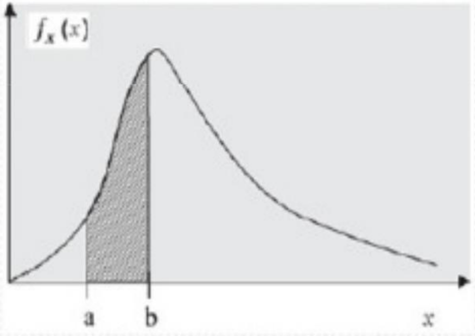
$k = \frac{1}{0.3} = 3.33\ldots$

Slides Professora Claúdia Nunes

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

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

V.a. Discreta vs V.a. Contínua

V.a. discretas	V.a. contínuas
• $E(X) = \sum_{x=-\infty}^{+\infty} x\, P(X=x)$	• $E(X) = \int_{-\infty}^{+\infty} x\, f_X(x)\, dx$

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

V.a. discretas	V.a. contínuas
• $E[g(X)] = \sum_{x=-\infty}^{+\infty} g(x)\, P(X=x)$	• $E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x)\, dx$

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

V.a. discretas	V.a. contínuas
• $var(X) = \sum_{x=-\infty}^{+\infty} x^2 P(X=x) - E^2(X)$	• $var(X) = \int_{-\infty}^{+\infty} x^2 f_X(x)\, dx - E^2(X)$
$= \sum_{x=-\infty}^{+\infty} x^2 P(X=x) - \left[\sum_{x=-\infty}^{+\infty} x\, P(X=x)\right]^2$	$= \int_{-\infty}^{+\infty} x^2 f_X(x)\, dx - \left[\int_{-\infty}^{+\infty} x f_X(x)\, dx\right]^2$

$$var(aX+b) = a^2 var(X)$$

V.a. discretas	V.a. contínuas
• mediana $= me$	• mediana $= me$
$0.5 \leq F_X(me) \leq 0.5 + P(X=me)$	$F_X(me) = 0.5$
	• 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

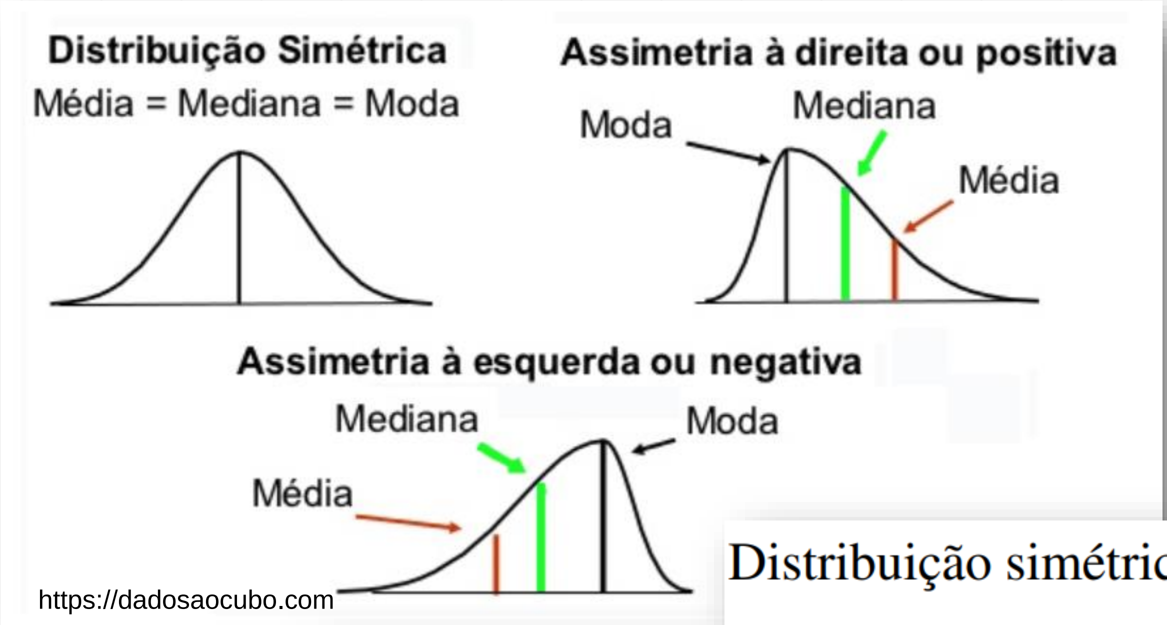
- Coeficientes de Assimetria

Achatamento/Curtose/Forma

- Coeficientes de Achatamento

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

Assimetria/Enviesamento

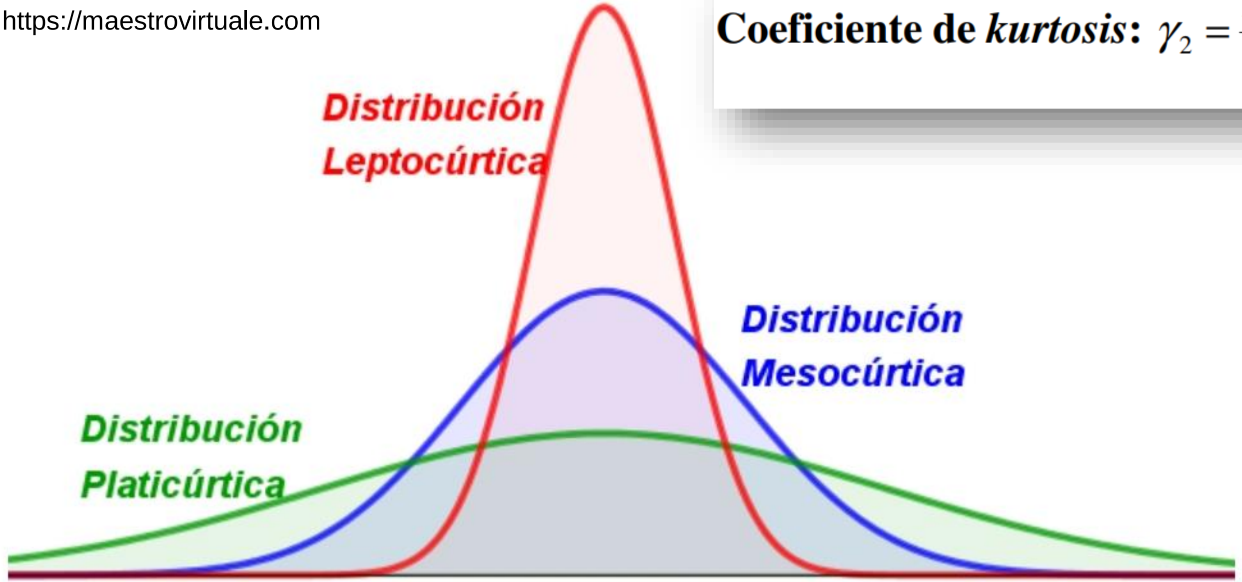


Distribuição simétrica $\rightarrow \gamma_1 = 0$

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

Achatamento/Curtose/Forma

https://maestrovirtuale.com



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?