Parte B

Introdução à Teoria da Probabilidade

Nota: a matéria dos *slides* 48 a 66 corresponde a revisões do ensino secundário. Não será lecionada nas aulas teóricas. Constitui estudo autónomo e será revista na 5ª aula prática.

Teoria da Probabilidade

Noções Preliminares

Definição | Fenómenos aleatórios

Fenómenos aleatórios são fenómenos sujeitos à influência do acaso. São caraterizados pela sua imprevisibilidade e regularidade estatística.

Definição | Experiência aleatória

Experiência aleatória é todo o procedimento que verifica as seguintes propriedades:

– pode repetir-se um grande número de vezes nas mesmas condições ou pelo menos em condições semelhantes;
– a sua realização dá um resultado de entre um conjunto de resultados possíveis;
– cada um dos resultados da experiência é imprevisível mas é possível considerar "estabilidade na frequência da sua ocorrência".

São exemplos de experiências aleatórias:

1. lançamento de dois dados equilibrados com as faces numeradas de 1 a 6 e registo do número de pontos da face que fica voltada para cima em cada um;
2. lançamento de uma moeda e observação da face que fica voltada para cima;
3. contagem do número mensal de acidentes de automóvel numa autoestrada;
4. registo do tempo de vida de uma pessoa, em anos;
5. registo do tempo de funcionamento de uma máquina até à primeira avaria.

Definição

Espaço de resultados ou espaço amostra é o conjunto de todos os resultados possíveis associados a uma experiência aleatória – representa-se por Ω.

Para os exemplos anteriores tem-se

1. $\Omega = \{(1,1),(1,2),(1,3),...,(6,5),(6,6)\}$;
2. $\Omega = \{$*'face valor', 'face país'*$\} = \{$'FV','FP'$\} = \{1,0\}$;
3. $\Omega = \mathbb{N}_0$;
4. $\Omega = \mathbb{N}$;
5. $\Omega = \mathbb{R}^+$.

Teoria da Probabilidade | Acontecimento

Definição

Acontecimento aleatório é qualquer subconjunto do espaço de resultados.

Seja Ω o espaço de resultados associado a uma experiência aleatória.

- Diz-se que $A \subset \Omega$ se realizou se o resultado, ω, da experiência é um elemento de A, i.e., $\omega \in A$.
- $A \subset B$, diz-se A subacontecimento de B, se e só se a realização de A implica a realização de B;
- A^c ou $\overline{A}$ diz-se acontecimento complementar ou contrário a A, é o conjunto de todos os elementos de Ω que não estão em A;

- $A \cup B$, diz-se união de A com B, é o acontecimento que consiste na realização de pelo menos um dos acontecimentos.
- AB ou $A \cap B$, diz-se produto ou intersecção, é o acontecimento que se realiza apenas quando ambos os acontecimentos se realizam.
- Os acontecimentos A e B dizem-se mutuamente exclusivos ou incompatíveis se e só se a realização de um implica a não realização do outro, i.e., se e só se $A \cap B = \emptyset$.
- $A - B = A \cap \overline{B}$ diz-se diferença dos acontecimentos A e B. É o acontecimento que se realiza se e só se A se realiza sem que B se realize.
- $\emptyset$ diz-se acontecimento impossível.
- Ω diz-se acontecimento certo.

Algumas propriedades das operações sobre acontecimentos:

Associativa	$(A \cap B) \cap C = A \cap (B \cap C)$ $(A \cup B) \cup C = A \cup (B \cup C)$
Comutativa	$A \cap B = B \cap A$ $A \cup B = B \cup A$
Distributiva	$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
Leis de De Morgan	$\overline{A \cap B} = \overline{A} \cup \overline{B}$ $\overline{A \cup B} = \overline{A} \cap \overline{B}$
Elemento neutro	$A \cap \Omega = A$ $A \cup \emptyset = A$
Elemento absorvente	$A \cap \emptyset = \emptyset$ $A \cup \Omega = \Omega$

Definição clássica de Laplace (séc. XIX)

Sob a hipótese de que **todos os casos são igualmente prováveis ou possíveis** (**princípio da simetria**).

Probabilidade de realização de um acontecimento A

$$P(A) = \frac{\text{número de casos favoráveis a } A}{\text{número total de casos possíveis}}$$

Na experiência aleatória 1, cada par (i, j), $i, j = 1, \cdots, 6$ tem igual possibilidade de ocorrer, assim a probabilidade de sair o mesmo número nos dois dados será

$$P(A) = \frac{\#\{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}}{\#\Omega} = \frac{6}{36} = \frac{1}{6}$$

Teoria da Probabilidade | Definições

Definição frequencista

Considere-se n repetições de uma experiência aleatória; n_A o nº de vezes que A se realiza. A frequência relativa do acontecimento A é n_A/n. De acordo com a teoria frequencista, a probabilidade do acontecimento A é a frequência relativa estimada para esse acontecimento, quando o número de vezes que se realiza a experiência (n) é muito elevado. Para n "muito elevado"

$$P(A) \approx \frac{n_A}{n}$$

Na experiência aleatória 2, suspeita-se que a moeda está viciada. A moeda foi lançada 1000 vezes e observou-se que a *'face valor'* saiu 647 vezes. Então a probabilidade do acontecimento *'face valor'* é aproximadamente 0.647, enquanto que a probabilidade do acontecimento *'face país'* é aproximadamente 0.353.

Teoria da Probabilidade | Definições

$\Omega-$ espaço de resultados associado a uma experiência aleatória.

Definição axiomática de Kolmogorov

Probabilidade, P, é uma aplicação que a cada acontecimento de Ω associa um número real satisfazendo o seguinte conjunto de axiomas:

A1) $P(A) \geq 0 \qquad \forall A \subset \Omega;$

A2) $P(\Omega) = 1;$

A3) $P(A \cup B) = P(A) + P(B)$ se $A \cap B = \emptyset$. (Axioma das probabilidades totais).

Se Ω é infinito,

A3*) $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ se $A_i \cap A_j = \emptyset, \quad i \neq j$ (Axioma completo das probabilidades totais).

Propriedades da Probabilidade

1. $P(\overline{A}) = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. $A \subset B \Rightarrow P(A) \leq P(B)$.
4. $P(A) \leq 1$.
5. $P(A - B) = P(A \cap \overline{B}) = P(A) - P(A \cap B)$.
6. Se $B \subset A \Rightarrow P(A - B) = P(A) - P(B)$.
7. Sejam $A_1, ..., A_n$ acontecimentos mutuamente exclusivos então $P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$
8. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Propriedades da Probabilidade

9. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

10. Generalização: Sejam $A_1, A_2, ..., A_n$ acontecimentos quaisquer
$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - ... - P(A_{n-1} \cap A_n)$
$+ P(A_1 \cap A_2 \cap A_3) + ... + P(A_{n-2} \cap A_{n-1} \cap A_n)$
$+ ... + (-1)^{n-1} P(A_1 \cap A_2 \cap ... \cap A_n)$.

Exemplo 1

Sejam A, B e C acontecimentos definidos num espaço de resultados Ω tais que $P(A) = P(B) = P(C) = \frac{1}{4}$; $P(A \cap B) = P(B \cap C) = 0$ e $P(A \cap C) = \frac{1}{8}$.

A probabilidade de se verificar pelo menos um dos acontecimentos A, B ou C é $P(A \cup B \cup C) =$

$$= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$
$$= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} - 0 - \frac{1}{8} - 0 + 0 = \frac{1}{4}.$$

Como $(A \cap B \cap C) \subset (A \subset B)$, $P(A \cap B \cap C) \leq P(A \cap B) = 0$.

Probabilidade condicional

Definição

Dados os acontecimentos A e B definidos em Ω, a probabilidade de A se realizar sabendo que B se realizou, ou seja, a probabilidade condicional de A dado B ou probabilidade de A se B representa-se por $P(A|B)$, com $P(B) > 0$ e define-se como

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Desta definição resulta o seguinte teorema:

Teorema da probabilidade composta

Se $P(A) > 0$ e $P(B) > 0$,

$$P(A \cap B) = P(A)\ P(B|A) = P(B)\ P(A|B)$$

Independência de acontecimentos

Definição

Dois acontecimentos A e B dizem-se mutuamente independentes se e só se

$$P(A \cap B) = P(A)\ P(B).$$

Da definição 9 conclui-se que se A e B são independentes então $P(A|B) = P(A)$ se $P(B) > 0$ e $P(B|A) = P(B)$ se $P(A) > 0$. Ou seja, dois acontecimentos são independentes quando a ocorrência de um não influencia a ocorrência do outro.

Teorema

Se A e B são independentes então
A e $\overline{B}$, $\overline{A}$ e B e $\overline{A}$ e $\overline{B}$, também são independentes.

Independência e exclusividade mútua

A independência de acontecimentos é um conceito "quase oposto" a exclusividade mútua. Enquanto que a independência corresponde à *não interferência* nas ocorrências, a exclusividade mútua corresponde ao *impedimento* da ocorrência conjunta.

Se $P(A) > 0$ e $P(B) > 0$,

A e B independentes $\Rightarrow$ A e B são não mutuamente exclusivos.

A e B mutuamente exclusivos $\Rightarrow$ A e B não são independentes.

(Demonstrar como exercício.)

Independência de 3 acontecimentos

Definição | Independência de três acontecimentos

Os acontecimentos A, B e C dizem-se mutuamente independentes ou apenas independentes se e só se

$P(ABC) = P(A)\ P(B)\ P(C)$ **e** $P(AB) = P(A)P(B)$ **e** $P(AC) = P(A)P(C)$ **e** $P(BC) = P(B)P(C)$.

Nota: A independência par a par não assegura independência de um conjunto de acontecimentos.

Probabilidade de acontecimentos | Exemplo 2

Uma empresa produz concentrado de tomate recorrendo a três processos de fabrico. Sabe-se que 20% da produção de concentrado provém do processo A, 30% do processo B e 50% do processo C. Nalgumas embalagens daquele concentrado tem-se verificado a ocorrência de defeitos. Sabe-se 1% das embalagens provenientes do processo A, 2% das provenientes do processo B e 8% das provenientes do processo C, respetivamente, têm defeito.

1. Qual a percentagem de embalagens, produzidas naquela empresa, que apresentam defeitos?
2. Verifica-se que uma embalagem escolhida ao acaso apresenta defeito. Qual a probabilidade de ter sido fabricada pelo processo A?

Teorema da probabilidade total

A resolução da Pergunta 1. baseia-se no seguinte teorema

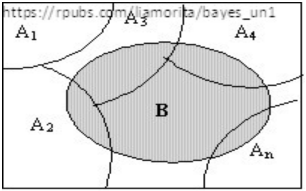
Teorema da probabilidade total

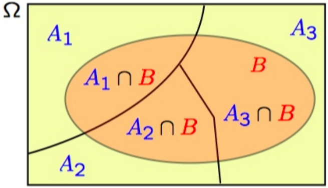
Sejam $A_1, A_2, ..., A_n$ acontecimentos definindo uma **partição de Ω**, i.e.,

$$A_1 \cup A_2 \cup .... \cup A_n = \Omega \qquad \text{e} \qquad A_i \cap A_j = \emptyset, \quad \forall i, j, \; i \neq j.$$

Se $P(A_i) > 0$, então para qualquer acontecimento $B \subset \Omega$ tem-se

$$P(B) = \sum_{i=1}^{n} P(A_i)\; P(B|A_i).$$





Exemplo 2 (continuação)

Resolução da Pergunta 1.

A experiência aleatória consiste em escolher ao acaso uma embalagem. A embalagem pode ter sido produzida pelos processo A, B ou C. A embalagem pode ter ou não defeito. Considerem-se os acontecimentos e respetivas probabilidades:
A: embalagem produzida pelo processo A $\rightarrow P(A) = 0.2$
B: embalagem produzida pelo processo B $\rightarrow P(B) = 0.3$
C: embalagem produzida pelo processo C $\rightarrow P(C) = 0.5$
D: embalagem defeituosa $\rightarrow P(D) =?$

Sabe-se que $P(D|A) = 0.01$, $P(D|B) = 0.02$ e $P(D|C) = 0.08$.

Note-se que A, B e C constituem uma partição de Ω (conjunto de todas as embalagens).

Pelo teorema da probabilidade total,

$$P(D) = P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C) = 0.048$$

.

Teorema de Bayes

Relativamente à Pergunta 2., pretende-se *atualizar* a probabilidade de um acontecimento *a priori*, à custa da informação *a posteriori*.
O seguinte teorema formaliza a resposta à questão:

Teorema de Bayes

Sejam $A_1, A_2, ..., A_n$ acontecimentos formando uma partição de Ω , onde $P(A_i) > 0$. Seja B um outro acontecimento de Ω, tal que $P(B) > 0$. Então para $k = 1, ..., n$ tem-se

$$P(A_k|B) = \frac{P(A_k).P(B|A_k)}{P(B)} = \frac{P(A_k).P(B|A_k)}{\sum_{i=1}^{n} P(A_i).P(B|A_i)}$$

Resolução da Pergunta 2.

$$P(A|D) = \frac{P(A).P(D|A)}{P(D)} = \frac{0.2 \times 0.01}{0.048} = 0.041(6).$$

Variável aleatória

Uma variável aleatória permite associar valores numéricos aos resultados de uma experiência aleatória.

Por exemplo, na experiência aleatória 1, pode-se associar a cada par de resultados a soma das faces dos dois dados lançados ao acaso.

Definição

Designa-se variável aleatória (v.a.) e costuma representar-se por X, uma função com domínio Ω e contradomínio em $\mathbb{R}$, cujo valor é determinado pelo resultado de uma experiência aleatória, i.e,

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

$$X(\omega) = x$$

Variável aleatória

Uma variável aleatória diz-se discreta se assume um conjunto finito ou infinito numerável de valores.

Exemplos:

- soma das faces viradas para cima no lançamento de dois dados;
- número de pessoas na fila para a cantina do ISA às 12h30.

Uma variável aleatória diz-se contínua se pode tomar qualquer valor real num dado intervalo, que pode ser a reta real

Exemplos:

- o peso de um frango com 2 meses de idade;
- a espessura da cortiça recém extraída ao nível de 1,30m.

Variável aleatória | probabilidade

A probabilidade de uma variável aleatória X tomar um conjunto de valores é a probabilidade do acontecimento de Ω cuja transformação por X originou o conjunto de valores.

Por exemplo, na experiência aleatória que consiste em escolher ao acaso um indivíduo, o espaço de resultados Ω é o conjunto de indivíduos. Considerando a v.a. X que representa o peso de um indivíduo, a probabilidade de o peso de um indivíduo estar entre 60Kg e 80Kg, $P[60 < X < 80]$, é a probabilidade de escolher ao acaso um indívíduo ω cujo peso $X(\omega)$ está entre 60 e 80Kg.

V.a. discreta | função massa de probabilidade

Uma forma de definir uma v.a. discreta consiste em identificar os valores possíveis para essa variável assim como a probabilidade de a v.a. tomar cada um desses valores.

Seja X uma v.a. tomando k valores, $x_1, \ldots, x_k$, cada um deles com probabilidade $p_1, \ldots, p_k$, respetivamente, i.e.,
$p_i = P[X = x_i]$, $i = 1, \ldots, k$.

Definição

Chama-se função massa de probabilidade da v.a. X à aplicação que a cada valor x_i faz corresponder um valor p_i, tal que

$$p_i = P[X = x_i]$$

A função massa de probabilidade satisfaz:
$p_i \geq 0$, $i = 1, \ldots, k$ **e** $\sum_{i=1}^{k} p_i = 1$.

V.a. discreta | distribuição de probabilidade

Designa-se distribuição de probabilidade da v.a. X ao conjunto de pares (x_i, p_i), $i = 1, \ldots, k$

Habitualmente a distribuição de probabilidade (lei) da v.a. X dispõe-se na forma:

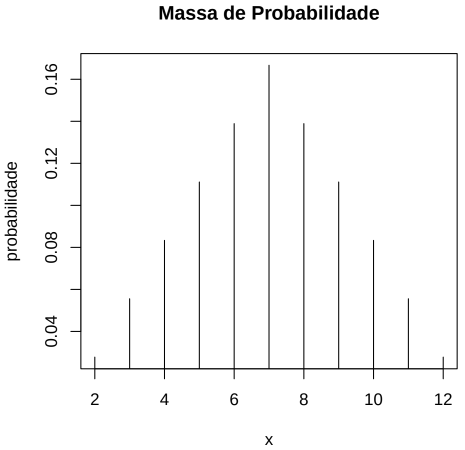
$$X = \begin{cases} x_1 & x_2 & \ldots & x_k \\ p_1 & p_2 & \ldots & p_k \end{cases} \quad \text{ou} \quad \begin{array}{c|cccc} x_i & x_1 & x_2 & \ldots & x_k \\ \hline P[X = x_i] & p_1 & p_2 & \ldots & p_k \end{array}$$

Por exemplo, a v.a. X que representa a soma das faces viradas para cima no lançamento de dois dados equilibrados tem a seguinte distribuição de probabilidade (verificar)

$$X = \begin{cases} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \end{cases}$$

V.a. discreta | distribuição de probabilidade

A distribuição de probabilidade de uma v.a. discreta representa-se graficamente através de um diagrama de barras. No exemplo acima tem-se



Exemplo 3

O número de aparelhos de micro-ondas vendidos diariamente num estabelecimento é uma variável aleatória X com a seguinte distribuição de probabilidade

$$X = \begin{cases} 0 & 1 & 2 & 3 & 4 \\ 0.3 & 0.3 & 0.2 & 0.1 & 0.1 \end{cases}$$

a) Verifique que se trata de uma distribuição de probabilidade.

b) Determine $P[1 \leq X \leq 3]$. Interprete esta probabilidade.

c) Num dia em que é vendido pelo menos 1 micro-ondas, qual é a probabilidade de serem vendidos mais de 2?

Exemplo 3 (continuação)

Resolução:

X é uma v.a. discreta que toma apenas cinco valores ($k = 5$).

a) Basta verificar que cada $p_i \geq 0$, $i = 1, \cdots, 5$ e $\sum_{i=1}^{5} p_i = 1$.

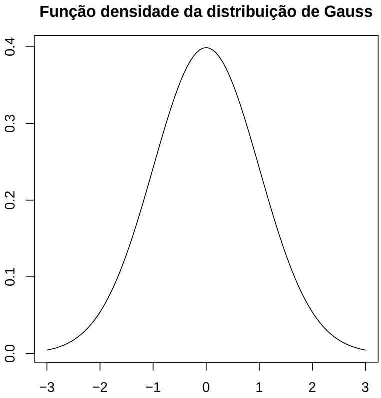
b) $P[1 \leq X \leq 3] = P[X = 1] + P[X = 2] + P[X = 3] = 0.3 + 0.2 + 0.1 = 0.6$.
Em 60% dos dias, são vendidos 1, 2 ou 3 aparelhos de micro-ondas.

c) $P[X > 2|X \geq 1] = \dfrac{P[X > 2 \wedge X \geq 1]}{P[X \geq 1]} = \dfrac{P[X > 2]}{1 - P[X < 1]} = \dfrac{0.1 + 0.1}{1 - 0.3} = \dfrac{2}{7}$

V.a. contínua | função densidade

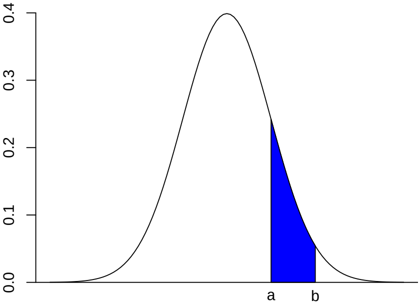
Contrariamente às variáveis discretas, as variáveis contínuas podem assumir qualquer valor numérico real dentro de um determinado intervalo. Entre quaisquer dois valores de uma variável aleatória contínua X, podem ocorrer infinitos outros valores.

Uma v.a. contínua pode definir-se através de uma função densidade, representada graficamente por uma curva em que a ordenada de cada ponto é a densidade de probabilidade desse ponto. Note-se que a altura correspondente ao valor 2.4 **não é** a probabilidade de X tomar o valor 2.4.



V.a. contínua | função densidade

Como uma v.a. contínua pode tomar infinitos valores num intervalo, **a probabilidade de tomar um único valor é zero**.



A probabilidade de uma v.a. contínua tomar um qualquer valor entre *a* e *b* é a área abaixo da curva da sua função densidade entre *a* e *b*, ou seja é o integral da função densidade entre *a* e *b*.

Como a probabilidade de a v.a. tomar um qualquer valor real é 1, a área abaixo da curva terá que ser 1.

V.a. contínua | função densidade

Definição

Uma função densidade de probabilidade ou apenas função densidade é uma função real de variável real, habitualmente designada por f, que verifica as seguintes condições:

$$f(x) \geq 0, \forall x \in \mathbb{R} \qquad \text{e} \qquad \int_{-\infty}^{+\infty} f(x)\, dx = 1$$

Sendo X uma v.a. **contínua** com função densidade f,

$$P[a < X < b] = P[a \leq X \leq b] = \int_a^b f(x)\, dx$$

Exemplo 4

O tempo de vida (em anos) de um dado equipamento pode ser modelado por uma variável aleatória X com função densidade

$$f(x) = \begin{cases} \dfrac{1}{5}\, e^{-x/5} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

a) Mostre que f é de facto uma função densidade.

b) Qual a probabilidade de esse equipamento durar entre 1 e 3 anos?

Exemplo 4

Resolução:

a) Basta verificar que $f(x) \geq 0$, $\forall x \in \mathbb{R}$, e que $\int_{-\infty}^{+\infty} f(x)\, dx = 1$

$$\begin{aligned}
\int_{-\infty}^{+\infty} f(x)\, dx &= \int_{-\infty}^{0} f(x)\, dx + \int_{0}^{+\infty} f(x) dx \\
&= \lim_{y \to -\infty} \int_{y}^{0} f(t)\, dt + \lim_{x \to +\infty} \int_{0}^{x} f(t)\, dt \\
&= \lim_{y \to -\infty} \int_{y}^{0} 0\, dt + \lim_{x \to +\infty} \int_{0}^{x} \frac{1}{5}\, e^{-t/5}\, dt \\
&= 0 + \lim_{x \to +\infty} \left[-e^{-t/5} \right]_{0}^{x} = \lim_{x \to +\infty} -e^{-x/5} + e^{0} = 1
\end{aligned}$$

b)

$$\begin{aligned}
P[1 < X < 3] &= \int_{1}^{3} f(x)\, dx = \int_{1}^{3} \frac{1}{5}\, e^{-x/5}\, dx \\
&= \left[-e^{-x/5} \right]_{1}^{3} = -e^{-3/5} - (-e^{-1/5}) \approx 0.2699
\end{aligned}$$

Variável aleatória | função distribuição cumulativa

Uma forma alternativa de definir uma variável aleatória X consiste na utilização de uma função real de variável real, cujo contradomínio é o intervalo $[0, 1]$, que associa a todo o $x \in \mathbb{R}$ a probabilidade de X tomar qualquer valor inferior ou igual a x.

Definição

A função distribuição cumulativa associada à variável aleatória X representa-se por F ou F_X e define-se como

$$F : \mathbb{R} \rightarrow [0, 1] \qquad \text{tal que} \qquad F(x) = P[X \leq x].$$

Variável aleatória | f.d.c. e probabilidade

O conhecimento da f.d.c. $F(.)$ permite calcular qualquer probabilidade:

- $P(X \leq b) = F(b)$ (por definição);
- $P(X < b) = \lim_{x \to b^-} F(x) = F(b^-)$;
- $P(X \geq a) = 1 - P(X < a) = 1 - F(a^-)$;
- $P(X > a) = 1 - P(X \leq a) = 1 - F(a)$;
- $P(X = a) = P(X \leq a) - P(X < a) = F(a) - F(a^-)$;
- $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$;
- $P(a < X < b) = P(X < b) - P(X \leq a) = F(b^-) - F(a)$;
- $P(a \leq X \leq b) = P(X \leq b) - P(X < a) = F(b) - F(a^-)$;
- $P(a \leq X < b) = P(X < b) - P(X < a) = F(b^-) - F(a^-)$.

Variável aleatória | f.d.c. e massa de probabilidade

Se X é uma v.a. **discreta** com distribuição de probabilidade (x_i, p_i), $F(x)$ é a **soma** das probabilidades de X tomar qualquer valor inferior ou igual a x, para todo o número real x.

X variável aleatória **discreta**

$$F(x) = P[X \leq x] = \sum_{x_i \leq x} P[X = x_i] = \sum_{x_i \leq x} p_i, \ \forall x \in \mathbb{R}.$$

e

$$p_i = P[X = x_i] = F(x_i) - F(x_i^-), \ \forall i = 1, 2, \cdots$$

Exemplo 3 (continuação)

$$X = \begin{cases} 0 & 1 & 2 & 3 & 4 \\ 0.3 & 0.3 & 0.2 & 0.1 & 0.1 \end{cases}$$

d) Determine a função de distribuição cumulativa de X e represente-a graficamente.

e) Determine $P[1 \leq X \leq 3]$ utilizando a função da alínea anterior.

Resolução:

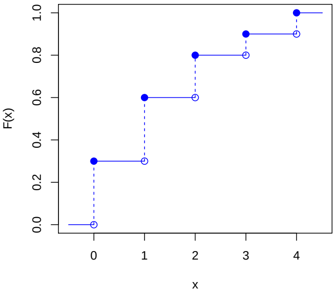
X é uma v.a. discreta que toma apenas cinco valores.

d) Para determinar a f.d.c. é necessário calcular $P[X \leq x] = \sum_{x_i \leq x} P[X = x_i], \forall x \in \mathbb{R}$.

Exemplo 3 (continuação)

- se $x < 0$, $F(x) = 0$
- se $0 \leq x < 1$, $F(x) = 0.3$
- se $1 \leq x < 2$, $F(x) = 0.3 + 0.3 = 0.6$
- se $2 \leq x < 3$, $F(x) = 0.3 + 0.3 + 0.2 = 0.8$
- se $3 \leq x < 4$, $F(x) = 0.3 + 0.3 + 0.2 + 0.1 = 0.9$
- se $x \geq 4$, $F(x) = 0.3 + 0.3 + 0.2 + 0.1 + 0.1 = 1$

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



Exemplo 3 (continuação)

No caso de uma variável aleatória discreta a função distribuição cumulativa é uma função em escada, onde os pontos de salto são os valores que a variável assume.

e) $P[1 \leq X \leq 3] = F(3) - F(1^-) = F(3) - F(0) = 0.9 - 0.3 = 0.6.$

Variável aleatória | f.d.c. e função densidade

Se X é uma v.a. **contínua** com função densidade f, $F(x)$ é o **integral** indefinido (e impróprio) da função densidade entre $-\infty$ e x, para todo o número real x.

X variável aleatória **contínua**

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} f(t)\,dt,\ \forall x \in \mathbb{R}.$$

e

$$f(x) = F'(x)$$

quando existe derivada; quando não existe pode-se arbitrar $f(x) = 0$.

Nota:

$$\begin{aligned}\int_a^b f(x)\,dx &= \int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx = F(b) - F(a)\\ &= P(a < X \leq b) = P(a \leq X \leq b)\\ &= P(a \leq X < b) = P(a < X < b),\ \forall a < b.\end{aligned}$$

$$f(x) = \begin{cases} \frac{1}{5}\, e^{-x/5} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

c) Determine a função de distribuição cumulativa de X e represente-a graficamente.

d) Qual a probabilidade de esse equipamento durar entre 1 e 3 anos? Utilize o resultado da alínea anterior.

Resolução:

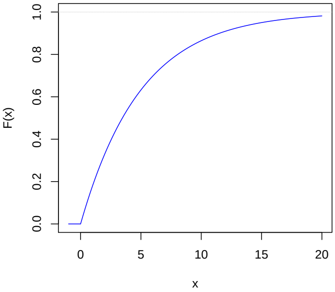
X é uma v.a. contínua.

Exemplo 4 (continuação)

c) A f.d.c. de X é $F(x) = \int_{-\infty}^{x} f(t)\,dt$, $x \in \mathbb{R}$:

- se $x \leq 0$, $F(x) = \int_{-\infty}^{x} 0\,dt = 0$
- se $x > 0$, $F(x) = \int_{-\infty}^{0} 0\,dt + \int_{0}^{x} \frac{1}{5}\, e^{-t/5}\,dt = 0 + \left[-e^{-t/5}\right]_0^x$
  $= -e^{-x/5} - (-e^0)$

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-x/5}, & x > 0 \end{cases}$$



Exemplo 4 (continuação)

No caso de uma variável aleatória contínua a função distribuição cumulativa é uma função contínua, mas não necessariamente derivável $\forall x \in \mathbb{R}$.
(Recorde as propriedades do integral indefinido estudado em Análise Matemática.)

d) $$\begin{aligned} P[1 < X < 3] &= F(3) - F(1) = \left(1 - e^{-3/5}\right) - \left(1 - e^{-1/5}\right) \\ &= e^{-1/5} - e^{-3/5} \approx 0.2699. \end{aligned}$$

Propriedades da função distribuição cumulativa

As seguintes propriedades da f.d.c. foram ilustradas nos Exemplos 3 e 4:

1. $0 \leq F(x) \leq 1$
2. $lim_{x \to -\infty} F(x) = 0$
   $lim_{x \to +\infty} F(x) = 1.$
3. F é uma função monótona não decrescente, i.e., dados dois números reais x_1 e x_2 tais que $x_1 < x_2$, tem-se $F(x_1) \leq F(x_2)$
4. $F(x)$ é contínua à direita, i.e., $lim_{x \to x_0^+} F(x) = F(x_0)$.
5. $P(X = x_0) = F(x_0) - F(x_0^-)$ onde $F(x_0^-) = lim_{x \to x_0^-} F(x)$

Se X é uma v.a. contínua, então F é contínua (também à esquerda) e portanto $F(x_0^-) = F(x_0)$ e $P(X = x_0) = 0$, para todo o número real x_0.

Variáveis aleatórias | valor médio

As variáveis aleatórias têm associados “indicadores” numéricos designados parâmetros.
Parâmetros de uma distribuição são números reais que caraterizam essa distribuição.

Definição | Valor Médio

O valor médio, valor esperado ou esperança matemática de uma variável aleatória X representa-se por $E[X]$, μ_X ou μ e define-se como:

$$E[X] = \sum_{i=1}^{n} x_i \, p_i \quad \text{se } X \text{ é v.a. discreta com distr. de prob. } (x_i, p_i)$$

$$E[X] = \int_{-\infty}^{+\infty} x \, f(x) \, dx \quad \text{se } X \text{ é v.a. contínua com f. densidade } f(\cdot)$$

Variáveis aleatórias | valor médio

Se X for v.a. discreta com uma infinidade numerável de valores tem-se $E[X] = \sum_{i=1}^{\infty} x_i \; p_i$. Neste caso só existe valor médio se aquela "soma infinita" existir.

Analogamente, no caso contínuo só existe valor médio, $E[X] = \int_{-\infty}^{\infty} x \; f(x) \; dx$, se o integral impróprio $\int_{-\infty}^{\infty} |x| \; f(x) \; dx$ for convergente.

V.a. | valor médio de uma função de X

Se X é uma v.a. e $\varphi : \mathbb{R} \to \mathbb{R}$ é uma função real de variável real, define-se valor médio de $\varphi(X)$ como

$$E[\varphi(X)] = \sum_i \varphi(x_i)\, p_i \quad \text{se } X \text{ é v.a. discreta com dist. probab. } (x_i, p_i)$$

$$E[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi(x)\, f(x)\, dx \quad \text{se } X \text{ é v.a. contínua com f. densidade } f$$

Para que exista valor médio exige-se que exista aquela “soma infinita” (no caso de se tratar de uma v.a. discreta com uma infinidade de valores) ou que o integral de $|\varphi(x)|\ f(x)$ seja convergente.

Propriedades do Valor Médio de uma v.a.

1. Linearidade
   - $E[a] = a$.
   - $E[a + bX] = a + b\,E[X]$.
   - $E[\varphi(X) + \psi(X)] = E[\varphi(X)] + E[\psi(X)]$

2. Positividade

   Se $X \geq 0$, i.e. a variável toma apenas valores ≥ 0, tem-se $E[X] \geq 0$.

Variáveis aleatórias | Variância e Desvio Padrão

Definição

A variância de uma variável aleatória X representa-se por $Var[X]$, σ_X^2 ou apenas σ^2 e define-se como

$$\sigma_X^2 = E\left[(X-\mu)^2\right]$$

$\sigma_X = \sqrt{Var[X]}$ designa-se desvio padrão de X.

Nota:

$$Var[X] = E[X^2] - \mu^2$$

(Demonstrar como exercício.)

Variáveis aleatórias | Variância e Desvio Padrão

Propriedades

1. $Var[X] \geq 0$
2. $Var[a + b\ X] = b^2\ Var[X]$.

Para o desvio padrão tem-se $\sigma_{(a+b\ X)} = |b|\ \sigma_X$

Exemplo 3 (continuação)

O número de aparelhos de micro-ondas vendidos diariamente num estabelecimento é uma variável aleatória X com a seguinte distribuição de probabilidade

$$X = \left\{ \begin{array}{lllll} 0 & 1 & 2 & 3 & 4 \\ 0.3 & 0.3 & 0.2 & 0.1 & 0.1 \end{array} \right.$$

f) Qual o valor esperado do número de aparelhos de micro-ondas vendidos por dia?
g) Se cada micro-ondas é vendido por 85 Euros, qual é a distribuição de probabilidade da receita bruta da venda de micro-ondas por dia?
h) Calcule a receita bruta esperada da venda de micro-ondas por dia.

Exemplo 3 (continuação)

Resolução:

X é uma v.a. discreta que toma cinco valores.

f) $E[X] = \sum_{i=1}^{5} x_i p_i = 0 \times 0.3 + 1 \times 0.3 + 2 \times 0.2 + 3 \times 0.1 + 4 \times 0.1 = 1.4.$
Em média são vendidos 1.4 aparelhos de micro-ondas por dia.

g) Sendo R a v.a. que carateriza a receita bruta da venda por dia, $R = 85X$, logo

$$R = \begin{cases} 0 & 85 & 170 & 255 & 340 \\ 0.3 & 0.3 & 0.2 & 0.1 & 0.1 \end{cases}$$

h) $E[R] = E[85X] = 85E[X] = 119€$. A receita bruta é, em média, 119€ por dia.

Exemplo 4 (continuação)

O tempo de vida (em anos) de um dado equipamento pode ser modelado por uma variável aleatória X com função densidade

$$f(x) = \begin{cases} \dfrac{1}{5}\, e^{-x/5} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

e) Qual é o tempo médio de vida deste equipamento?

f) Qual é a variância do tempo de vida deste equipamento?

g) Se o valor de retoma (V em €) do equipamento depender do seu tempo de vida através da expressão $V = 20 - 2X$, qual é o valor médio e a variância da retoma deste equipamento?

Exemplo 4 (continuação)

Resolução:

e) X é uma v.a. contínua.

$$E[X] = \int_{-\infty}^{+\infty} x.f(x)\,dx = \int_{-\infty}^{0} x.f(x)\,dx + \int_{0}^{+\infty} x.f(x)dx$$

$$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{+\infty} x.\frac{1}{5}\,e^{-x/5}dx = 0 + \lim_{x\to+\infty}\int_{0}^{x} t.\frac{1}{5}\,e^{-t/5}\,dt$$

$$\underbrace{=}_{(*)} \lim_{x\to+\infty}\left[-t.e^{-t/5} - 5.e^{-t/5}\right]_0^x$$

$$= \lim_{x\to+\infty}\left(\underbrace{-x.e^{-x/5}}_{\to 0} - 5.\underbrace{e^{-x/5}}_{\to 0} + 0 + 5.e^0\right) = 5$$

$(*)$ utilizou-se o método de primitivação por partes, $P(fg) = Fg - P(Fg')$ com $f = \frac{1}{5}\,e^{-t/5}$ e $g = t$ e portanto $F = Pf = -e^{-t/5}$ e $g' = 1$.

Em média, este equipamento dura 5 anos.

Exemplo 4 (continuação)

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

$$E[X^2] = \int_{-\infty}^{+\infty} x^2.f(x)\ dx = \int_{-\infty}^{0} x^2.f(x)\ dx + \int_{0}^{+\infty} x^2 f(x) dx$$

$$= \int_{-\infty}^{0} 0\ dx + \int_{0}^{+\infty} x^2.\frac{1}{5}\ e^{-x/5} dx = 0 + \lim_{x\to+\infty} \int_0^x t^2.\frac{1}{5}\ e^{-t/5}\ dt$$

$$\underbrace{=}_{(*)} \lim_{x\to+\infty} \left[-t^2.e^{-t/5} - 6t.e^{-t/5} - 30.e^{-t/5}\right]_0^x$$

$$= \lim_{x\to+\infty} \left(\underbrace{-x^2.e^{-x/5}}_{\to 0} \underbrace{-6x.e^{-x/5}}_{\to 0} -30.\underbrace{e^{-x/5}}_{\to 0} +0 + 30.e^0\right) = 30$$

$(*)$ utilizou-se o método de primitivação por partes, $P(fg) = Fg - P(Fg')$ com $f = t.\frac{1}{5}\ e^{-t/5}$ e $g = t$ e portanto $F = Pf = -te^{-t/5} - 5e^{-t/5}$ e $g' = 1$.
Logo, $Var[X] = E[X^2] - (E[X])^2 = 30 - 5^2 = 5(\text{ano})^2$.

g) $E[V] = E[20 - 2X] = 20 - 2E[X] = 10€$ e
$Var[V] = Var[20 - 2X] = Var[-2X] = (-2)^2 Var[X] = 4 \times 5 = 20€^2$.

Definição

O quantil de probabilidade p $(0 < p < 1)$ de uma v.a. X representa-se por χ_p e define-se como o menor valor da variável aleatória X tal que $F_X(\chi_p) \geq p$.

Se $p = 0.5$, $\chi_{0.5}$ designa-se mediana de X e é o menor valor da variável tal que $F_X(\chi_{0.5}) \geq 0.5$.

Nota: Se X é uma v.a. contínua em que F_X não tem patamares (usual),

- o quantil de probabilidade p é o valor χ_p tal que $F_X(\chi_p) = p$, ou seja é a solução da equação $F_X(x) = p$.
- a mediana $\chi_{0.5}$, é a solução de $F_X(x) = 0.5 \iff \int_{-\infty}^{x} f(t)dt = 0.5$.

Vetores aleatórios

Quando ao resultado de uma experiência aleatória se associam 2 ou mais atributos numéricos, obtêm-se um vetor aleatório.

No caso de serem 2 atributos obtém-se um par aleatório representado frequentemente por (X_1, X_2) ou (X, Y).

Exemplos:

- quantidade de precipitado P e volume V de gás numa reação química, (P, V);
- altura A e diâmetro à altura do peito D do tronco de uma árvore seleccionada ao acaso, (A, D).

Pares aleatórios

Definição

Um par aleatório (X, Y) é uma aplicação

$$\begin{aligned}(X, Y) : \Omega &\to \mathbb{R}^2 \\ \omega &\to (x, y)\end{aligned}$$

Vão considerar-se os tipos de pares aleatórios:

- par aleatório discreto $\Rightarrow$ componentes são ambas variáveis aleatórias discretas;
- par aleatório contínuo $\Rightarrow$ componentes são ambas variáveis aleatórias contínuas.

Pares aleatórios discretos

(X, Y) é um par aleatório discreto se toma os valores (x_i, y_j) com probabilidades $p_{ij} = P[X = x_i, Y = y_j]$.

Definição | massa de probabilidade conjunta

Designa-se distribuição de probabilidades conjunta do par (X, Y) aos valores (x_i, y_j) e respectivas probabilidades p_{ij}

p_{ij} é designada função massa de probabilidade conjunta e deve verificar as seguintes condições:

$$p_{ij} \geq 0 \qquad \forall (i, j) \qquad \text{e} \qquad \sum_i \sum_j p_{ij} = 1.$$

Pares aleatórios discretos

Um modo de representar a distribuição de probabilidades conjuntas de um par aleatório discreto (X, Y) é na forma de um quadro

$X \backslash Y$	y_1	y_2	...	y_n	
x_1	p_{11}	p_{12}	...	p_{1n}	$p_{1\cdot}$
x_2	p_{21}	p_{22}	...	p_{2n}	$p_{2\cdot}$
.	.	.	...	.	.
.	.	.	...	.	.
.	.	.	...	.	.
x_m	p_{m1}	p_{m2}	...	p_{mn}	$p_{m\cdot}$
	$p_{\cdot 1}$	$p_{\cdot 2}$	...	$p_{\cdot n}$	1

$p_{i\cdot} = P[X = x_i] = \sum_{j=1}^{n} p_{ij}$ e $p_{\cdot j} = P[Y = y_j] = \sum_{i=1}^{m} p_{ij}$

são as probabilidades marginais de X e Y respetivamente.

Pares aleatórios discretos | Exemplo 5

(Adaptado do 1º Teste 2021/22)
4. Um estudante está a estudar intensamente para os seus exames. Para descontrair nos intervalos da sua preparação costuma realizar o seguinte jogo: faz séries de 2, 3 ou 4 lançamentos a um cesto de basquetebol. Seja (X, Y) o par aleatório discreto em que: X é o número de lançamentos realizados numa série e Y é o número de lançamentos acertados numa série. A função massa probabilidade conjunta de (X, Y) é dada por:

X \ Y	0	1	2	3	4	$p_{i\cdot}$
2	0.01	0.02	0.04	0	0	0.07
3	0.02	0.03	0.04	0.04	0	0.13
4	0	0.1	0.2	0.2	0.3	0.8
$p_{\cdot j}$	0.03	0.15	0.28	0.24	0.3	1

- o estudante faz 3 lançamentos e acerta todos em 4% das séries;
- o estudante faz 2 lançamentos em 7% das séries;
- o estudante acerta 4 lançamentos em 30% das séries;
- o estudante acerta todos os lançamentos em 38% das séries.

Pares aleatórios discretos

Definição | probabilidade condicional de X dado Y

A probabilidade condicional de X dado $Y = y_j$ (fixo) com $P[Y = y_j] > 0$ é definida como

$$P[X = x_i | Y = y_j] = \frac{P[X = x_i, Y = y_j]}{P[Y = y_j]} = \frac{p_{ij}}{p_{\cdot j}}$$

Definição | probabilidade condicional de Y dado X

Do mesmo modo a probabilidade condicional de Y dado $X = x_i$ (fixo) com $P[X = x_i] > 0$ é definida como

$$P[Y = y_j | X = x_i] = \frac{P[X = x_i, Y = y_j]}{P[X = x_i]} = \frac{p_{ij}}{p_{i \cdot}}$$

Pares aleatórios discretos | Exemplo 5 (cont.)

- nas séries de 2 lançamentos, o estudante acerta os 2 lançamentos em $4/7$ das vezes, acerta 1 lançamento em $2/7$ das vezes e falha os 2 lançamentos em $1/7$ das vezes;
- nas séries em que falha todos os lançamentos, o estudante faz 2 lançamentos em $1/3$ das séries e 3 lançamentos em $2/3$ das vezes;
- o estudante acerta 2 lançamentos em $4/7$ das vezes em que faz 2 lançamentos; em $4/13$ das vezes em que faz 3 lançamentos e em $1/4$ das vezes em que faz 4 lançamentos.

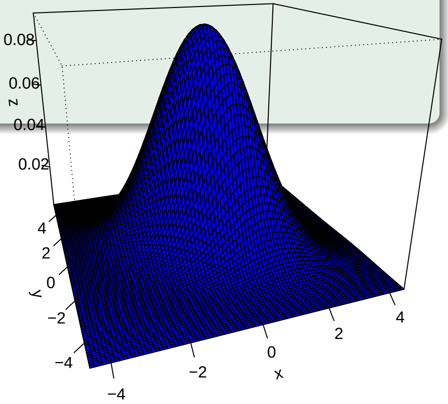
(Verificar os valores a azul como exercício)

Pares aleatórios contínuos

Definição | função densidade conjunta

Um par aleatório (X, Y) diz-se contínuo se existir uma função $f(x, y)$, designada função densidade (de probabilidade) conjunta, que verifica as seguintes condições:

- $f(x, y) \geq 0$
- $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dxdy = 1.$



Pares aleatórios contínuos

Dado $A \subset \mathbb{R}^2$ tem-se $\quad P[(X, Y) \in A] = \iint_A f(x, y) dx dy.$

É o volume da região limitada inferiormente por A no plano $x0y$ e superiormente pelo gráfico de f.

Definição | densidade marginal

densidade marginal de X: $\quad f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$

densidade marginal de Y: $\quad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$

Independência de variáveis aleatórias

Definição

Dado o par aleatório (X, Y) diz-se que as variáveis X e Y são independentes se e só se

- $p_{ij} = p_{i.} \times p_{.j} \quad \forall i, j$, no caso de (X, Y) ser um par aleatório discreto
- $f(x, y) = f_X(x) \times f_Y(y) \quad \forall (x, y) \in \mathbb{R}^2$ no caso de (X, Y) ser um par aleatório contínuo.

Valor Médio de uma função de um par aleatório

Definição

Dado o par aleatório (X, Y), e $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, define-se

$$E[g(X, Y)] = \sum_i \sum_j g(x_i, y_j)\; p_{ij}\;, \qquad \text{no caso discreto}$$

$$E[g(X, Y)] = \iint_{R^2} g(x, y)\; f(x, y)\; dxdy\;, \qquad \text{no caso contínuo.}$$

Por exemplo, se (X, Y) é um par aleatório discreto:

- $E[X] = \sum_i \sum_j x_i\; p_{ij} = \sum_i x_i \sum_j\; p_{ij} = \sum_i x_i p_{i.}$
- $E[XY] = \sum_i \sum_j x_i y_j\; p_{ij}$

Valor Médio de uma função de um par aleatório

Propriedades do Valor Médio:

1. Aditividade $\qquad E[X \pm Y] = E[X] \pm E[Y]$

2. Se X e Y variáveis aleatórias independentes

$$\Downarrow$$

$$E[XY] = E[X]E[Y]$$

Valor Médio de uma função de um par aleatório

Nota:

O recíproco da Propriedade 2 não é verdadeiro.

Se X e Y são v. a.'s com a seguinte distribuição de probabilidades

X \ Y	-1	0	1
0	0	1/2	0
1	1/4	0	1/4

tem-se $E[XY] = E[X] \times E[Y]$ e no entanto X e Y não são independentes.
(Verificar como exercício.)

Covariância de um par aleatório

Definição

Dado o par aleatório (X, Y) define-se covariância de X e Y como

$$Cov[X, Y] \equiv \sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$$

Nota: $Cov[X, Y] = E[XY] - E[X]E[Y]$ (Demonstrar como exercício.)

Propriedades

1. $Var[X \pm Y] = Var[X] + Var[Y] \pm 2Cov[X, Y]$
2. Se X e Y são v.a.'s <u>independentes</u> $\Longrightarrow Cov[X, Y] = 0$.

   Nota: O recíproco não é verdadeiro.
3. $Cov[a + bX, c + dY] = bd\ Cov[X, Y]$, com $a, b, c, d \in \mathbb{R}$.
4. $|Cov[X, Y]| \leq \sigma_X\ \sigma_Y$, onde σ_X designa o desvio padrão da v.a. X.

Coeficiente de correlação de um par aleatório

Definição

Dado o par aleatório (X, Y) define-se coeficiente de correlação de X e Y como

$$\rho \equiv \rho_{X,Y} = \frac{Cov[X, Y]}{\sigma_X \, \sigma_Y}$$

$(\sigma_X > 0$ e $\sigma_Y > 0)$.

Propriedades

1. $-1 \leq \rho_{X,Y} \leq 1$
2. Se X e Y são v. a. independentes $\Longrightarrow \quad \rho_{X,Y} = 0.$
3. $\rho_{a+bX,c+dY} = \begin{cases} \rho_{X,Y} & \text{se } bd > 0 \\ -\rho_{X,Y} & \text{se } bd < 0 \end{cases}$

Modelos (Distribuições) Discretos

Há fenómenos que podem ser modelados por distribuições conhecidas. Algumas das mais importantes distribuições discretas são:

- Distribuição uniforme discreta
- Distribuição de Bernoulli e binomial
- Distribuição de Poisson

Distribuição uniforme discreta

Definição

Seja X uma v.a. que toma os valores $x_1, \ x_2 \ , \ \ldots \ , x_k$ com igual probabilidade, isto é

$$X = \begin{cases} x_1 & x_2 & \cdots & x_k \\ 1/k & 1/k & \cdots & 1/k \end{cases}$$

então diz-se que X tem distribuição uniforme discreta,
$X \frown UD(x_1, \ x_2 \ , \ \ldots \ , x_k)$.

Caso particular

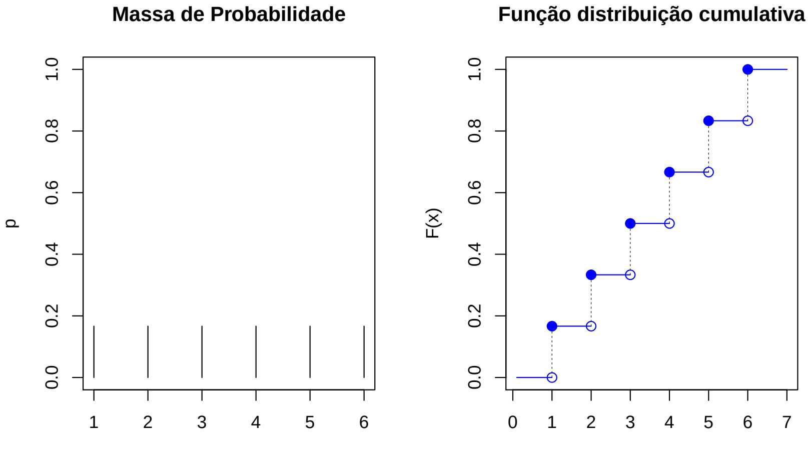
$$X = \begin{cases} 1 & 2 & \cdots & n \\ 1/n & 1/n & \cdots & 1/n \end{cases}$$

$X \frown UD(1, \ 2 \ , \ \ldots \ , n)$.

Distribuição uniforme discreta | Exemplo 6

Lançamento de um dado equilibrado e observação do número resultante:

$$X = \begin{cases} 1 & 2 & \cdots & 6 \\ 1/6 & 1/6 & \cdots & 1/6 \end{cases}$$



Distribuição uniforme discreta

Valor médio e variância

$$X = \begin{cases} 1 & 2 & \cdots & n \\ 1/n & 1/n & \cdots & 1/n \end{cases}$$

$$E[X] = \frac{n+1}{2} \qquad Var[X] = \frac{n^2-1}{12}$$

Demonstrar estes resultados como exercício, sabendo que a soma dos n primeiros números naturais é $\frac{n(n+1)}{2}$ e que a soma dos seus quadrados é $\frac{n(n+1)(2n+1)}{6}$.

Para o Exemplo 6, $E[X] = 3.5$ e $Var[X] = 2.9167$.

Distribuição de Bernoulli

Experiência (ou prova) de Bernoulli

Experiência aleatória com apenas dois resultados possíveis, $\Omega =${“**sucesso**”, “**insucesso**”}, em que a probabilidade de sucesso é p e portanto a probabilidade de insucesso é $1-p$.

Distribuição de Bernoulli com parâmetro p

A variável aleatória X que descreve o resultado de uma **única** experiência de Bernoulli, associando o valor **1** a **sucesso** e o valor **zero** a **insucesso**, tem massa de probabilidade

$$X = \begin{cases} 0 & 1 \\ 1-p & p \end{cases}$$

X tem distribuição de Bernoulli com parâmetro p, $X \frown Bernoulli(p)$.

Distribuição de Bernoulli

Exemplos:

- presença ou ausência de infeção numa planta
- artigo defeituoso ou não defeituoso numa linha de produção
- dia com chuva ou sem chuva
- animal vacinado ou não vacinado

Valor esperado e variância:

$X \frown Bernoulli(p)$

$$E[X] = p \qquad Var[X] = pq, \; q = 1 - p$$

Distribuição Binomial

Provas de Bernoulli são **independentes** quando o resultado de uma prova não interfere no resultado de outra.

Definição

A v.a. X que conta o número de sucessos em n provas de Bernoulli independentes, todas com a mesma probabilidade de sucesso p, tem distribuição binomial com parâmetros n e p, $X \frown \mathcal{B}(n, p)$.

Note-se que $X \frown \mathcal{B}(1, p) \Leftrightarrow X \frown \mathit{Bernoulli}(p)$.

Exemplo 7.

De uma embalagem de bolbos de túlipa, onde se indica que a taxa de germinação é 40%, colocaram-se 5 bolbos a germinar. A germinação ou não germinação de um bolbo é o resultado de uma experiência de Bernoulli com probabilidade de sucesso (*sucesso* = *germinação*) $p = 0.4$. O número de bolbos que germinam, em 5, pode ser considerado o número de sucessos em $n = 5$ provas de Bernoulli independentes com probabilidade de sucesso $p = 0.4$, $X \frown \mathcal{B}(5, 0.4)$.

- a probabilidade de germinarem todos os bolbos é $P[X = 5] = 0.4^5$
- a probabilidade de não germinar qualquer bolbo é $P[X = 0] = 0.6^5$
- a probabilidade de germinarem 3 bolbos é

$$P[X = 3] = \binom{5}{3} (0.4)^3 (0.6)^2$$

Note que a variável aleatória X apenas pode tomar os valores 0, 1, ..., 5.

Exercício: Calcular as probabilidades de X tomar cada um dos valores possíveis e confirmar que somam 1.

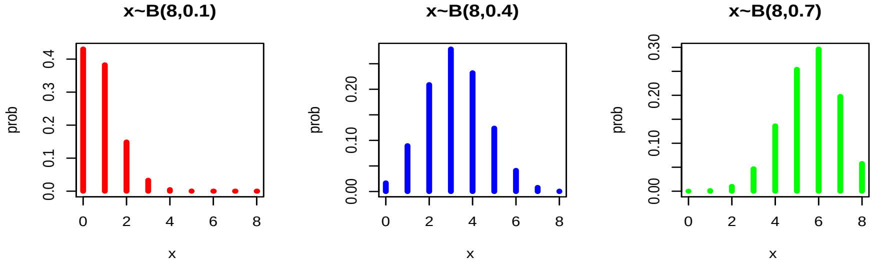
Distribuição Binomial

Caracterização da v.a. $X \frown \mathcal{B}(n, p)$:

$x = 0, 1, 2, ..., n$ $\longrightarrow$ $n^{\underline{o}}$ de "sucessos" nas n provas

$P[X = x] = \binom{n}{x}\, p^x\, (1-p)^{n-x}$ $\longrightarrow$ probabilidade de ocorrerem x "sucessos"

Para $n = 8$ e $p = 0.1,\ 0.4,\ 0.7$, a função massa de probabilidade é:



Distribuição Binomial

Podem-se obter valores da massa de probabilidade e/ou da função distribuição cumulativa da distribuição binomial recorrendo a um *software* estatístico ou a tabelas.

A distribuição binomial no R

A distribuição binomial no R é designada `binom`. Para cada distribuição existem 3 funções que se distinguem pelo prefixo `d, p, q`:

- `d` devolve valores da massa de probabilidade de uma distribuição discreta (ou da densidade de uma distribuição contínua)
- `p` devolve valores da função distribuição cumulativa
- `q` devolve quantis da distribuição

Para informações mais detalhadas, consultar a página `https://stat.ethz.ch/R-manual/R-devel/library/stats/html/Distributions.html`

Distribuição Binomial

No Exemplo 7, $X \frown \mathcal{B}(5, 0.4)$, $P[X = 3]$ obtém-se com o comando.

```
> # BINOMIAL NO R
> # dbinom(x, size, prob, log = FALSE)
> # pbinom(q, size, prob, lower.tail = TRUE, log.p = FALSE)
> # qbinom(p, size, prob, lower.tail = TRUE, log.p = FALSE)
> dbinom(3, 5, 0.4)
[1] 0.2304
> # que e' igual a
> pbinom(3,5,0.4)-pbinom(2,5,0.4)
[1] 0.2304
```

Tabelas da binomial, $X \frown \mathcal{B}(n, p)$

A tabela contém, para cada n e p, os valores da função distribuição cumulativa $F(x)$, para $x = 0, 1, \cdots, n$.

Para obter $P[X = a]$ será necessário utilizar
$P[X = a] = F(a) - F(a^-) = F(a) - F(a-1)$.

Distribuição Binomial

No Exemplo 7, $X \frown \mathcal{B}(5, 0.4)$,
$P[X = 3] = F(3) - F(2) = 0.9130 - 0.6826 = 0.2304.$

Função Distribuição Cumulativa da Binomial

Esta tabela foi criada com base no comando `pbinom` do *software R*, indicando os valores da Função Distribuição Cumulativa duma variável aleatória com distribuição Binomial, $X \cap B(n, p)$, para valores dos parâmetros n indicados na primeira coluna, valores do pârametro p indicados no topo de cada coluna, e valores da variável x indicados na segunda coluna. No corpo da tabela estão as probabilidades $P[X \leq x]$.

		p									
n	x	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500
	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400	0.7975	0.7500
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250
	1	0.9928	0.9720	0.9392	0.8960	0.8438	0.7840	0.7183	0.6480	0.5748	0.5000
	2	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360	0.9089	0.8750
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625
	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752	0.3910	0.3125
	2	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208	0.7585	0.6875
	3	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744	0.9590	0.9375
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0312
	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370	0.2562	0.1875
	2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826	0.5931	0.5000
	3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130	0.8688	0.8125
	4	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9947	0.9898	0.9815	0.9688

Distribuição Binomial

Valor médio e variância

$X \frown \mathcal{B}(n, p)$

$$E[X] = np \qquad Var[X] = npq, \; q = 1 - p$$

Relação entre as distribuições do número de sucessos e do número de insucessos

$$X \frown \mathcal{B}(n, p) \;\Rightarrow\; (n - X) \frown \mathcal{B}(n, 1 - p).$$

Distribuição Binomial

Teorema da estabilidade da soma

Se as v.a. $X_1, \cdots, X_k$ são independentes e $X_i \frown \mathcal{B}(n_i, p)$, $i = 1, \cdots, k$, então

$$\sum_{i=1}^{k} X_i \frown \mathcal{B}\left(n, p\right)$$

com $n = \sum_{i=1}^{k} n_i$.

Aqui `https://www.statology.org/binomial-distribution-real-life-examples/` encontra exemplos da vida real em que se utiliza a distribuição binomial.

Distribuição de Poisson

A distribuição de Poisson permite modelar o número de ocorrências num intervalo de tempo ou numa região do espaço.

Processo de Poisson refere-se ao número de ocorrências num intervalo de tempo ou numa região espacial que verifica as propriedades:

- os números de ocorrências em intervalos de tempo (regiões) disjuntos são independentes;
- a probabilidade de ocorrência num intervalo (região) muito pequeno é proporcional à amplitude (tamanho) do intervalo (região);
- a probabilidade de o número de ocorrências ser superior a um é nula em intervalos (regiões) muito pequenos.

Distribuição de Poisson

O modelo de Poisson é então adequado quando queremos modelar o número de ocorrências num processo cujo comportamento médio é "estável"[2]: número de abelhas que regressam à colmeia durante períodos de 5 minutos, número de glóbulos vermelhos em cada cela de um hemacímetro, número de camarões que se recolhe num camaroeiro de determinado tamanho num tanque de um viveiro, número de ninhos de determinada espécie que existe em determinada área. É a tradução simples da nossa fé na regularidade, e consequente predictibilidade, dos fenómenos. Se pegarmos em punhados de arroz e os "semearmos" enquanto percorremos uma sala, a configuração que esperamos que resulte tem um padrão de aleatoriedade que corresponde a contagens de Poisson. O modelo de Poisson é também uma bitola no que respeita a dispersão, porque $\frac{\mathrm{var}(X)}{\mathbb{E}(X)} = 1$: Se

em: https://www.instituto-camoes.pt/images/stories/tecnicas_comunicacao_em_portugues/Matematica/Matematica-ModelosdeContagemePadroesdeAleatoriedade.pdf

Distribuição de Poisson

Seja $\lambda > 0$ o número médio de ocorrências num dado intervalo de tempo (ou numa região do espaço) num processo de Poisson. Então a v.a. que representa o número de ocorrências nesse intervalo de tempo (ou região), tem distribuição de Poisson com parâmetro λ, $X \frown \mathcal{P}(\lambda)$.

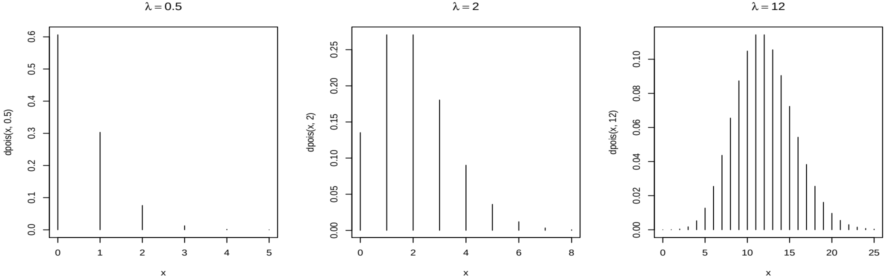
Função massa de probabilidade

$X \frown \mathcal{P}(\lambda), \quad \lambda > 0$

$$P[X = x] = \frac{e^{-\lambda}\,\lambda^x}{x!}, \qquad x = 0, 1, 2.....$$

Distribuição de Poisson

Para $\lambda = 0.5,\ 2,\ 12$, a função massa de probabilidade é:



Distribuição de Poisson

Podem-se obter valores da massa de probabilidade e/ou da função distribuição cumulativa da distribuição de Poisson recorrendo a um *software* estatístico ou a tabelas.

A distribuição de Poisson no R

A distribuição de Poisson no R é designada `pois`.
Se $X \frown \mathcal{P}(0.7)$, $P[X = 2]$ obtém-se com o comando:

```
> # POISSON NO R
> # dpois(x, lambda, log = FALSE)
> # ppois(q, lambda, lower.tail = TRUE, log.p = FALSE)
> # qpois(p, lambda, lower.tail = TRUE, log.p = FALSE)
> dpois(2, 0.7)
[1] 0.1216634
> # que e' igual a
> ppois(2, 0.7)-ppois(1, 0.7)
[1] 0.1216634
```

Tabelas da Poisson, $X \frown \mathcal{P}(\lambda)$

A tabela contem, para cada λ, os valores da função distribuição cumulativa $F(x)$, para $x = 0, 1, \cdots$.

Para obter $P[X = a]$ será necessário utilizar
$P[X = a] = F(a) - F(a^-) = F(a) - F(a-1)$.

Se $X \frown \mathcal{P}(0.7)$, $P[X = 2] = F(2) - F(1) = 0.966 - 0.844 = 0.122$.

Função Distribuição Cumulativa da Poisson

Esta tabela foi criada com base no comando `ppois` do *software R*, indicando os valores da Função Distribuição Cumulativa duma variável aleatória com distribuição Poisson, $X \frown P(\lambda)$, para valores do parâmetro λ indicados no topo de cada coluna, e valores da variável x indicados no início de cada linha. No corpo da tabela estão as probabilidades $P[X \leq x]$.

	$\lambda = E[X]$									
x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.905	0.819	0.741	0.670	0.607	0.549	0.497	0.449	0.407	0.368
1	0.995	0.982	0.963	0.938	0.910	0.878	0.844	0.809	0.772	0.736
2	1.000	0.999	0.996	0.992	0.986	0.977	0.966	0.953	0.937	0.920
3	1.000	1.000	1.000	0.999	0.998	0.997	0.994	0.991	0.987	0.981
4	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.996
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	$\lambda = E[X]$									
x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
0	0.333	0.301	0.273	0.247	0.223	0.202	0.183	0.165	0.150	0.135
1	0.699	0.663	0.627	0.592	0.558	0.525	0.493	0.463	0.434	0.406
2	0.900	0.879	0.857	0.833	0.809	0.783	0.757	0.731	0.704	0.677

Distribuição de Poisson

Valor médio e variância

$E[X] = \lambda \qquad Var[X] = \lambda.$

Teorema da estabilidade da soma

Se as v.a. $X_1, \cdots, X_k$ são independentes e $X_i \frown \mathcal{P}(\lambda_i)$, $i = 1, \cdots, k$, então

$$\sum_{i=1}^{k} X_i \frown \mathcal{P}\left(\sum_{i=1}^{k} \lambda_i\right).$$

Aproximação Binomial $\rightarrow$ Poisson

Se $X \frown \mathcal{B}(n, p)$ com $n \geq 20$ e $p \leq 0.05$ então $X \approx \mathcal{P}(\lambda)$, com $\lambda = np$.

Exemplo 8. (1º Teste de 2021/22)

Um sismo com magnitude superior a cinco é considerado um "grande sismo". Admita que o número de "grandes sismos" que ocorrem anualmente em Portugal Continental segue uma distribuição de Poisson. De acordo com o Catálogo Sísmico de Portugal Continental, o número médio anual desses "grandes sismos" é 0.7. Calcule a probabilidade de:

a) ocorrerem dois ou mais"grande sismos" num ano;

b) o número total de "grandes sismos" em 8 anos ser inferior a seis;

c) em 20 anos, existirem 10 anos sem "grandes sismos".

Resolução:

a) Seja X a v.a. que representa o número de "grandes sismos" que ocorrem em Portugal num ano, $X \frown \mathcal{P}(0.7)$.
$P[X \geq 2] = 1 - P[X < 2] = 1 - F_X(1) \underbrace{=}_{\text{tabela}} 1 - 0.844 = 0.156;$

Exemplo 8 (continuação da resolução):

b) seja X_i a v.a. que representa o número de "grandes sismos" que ocorrem em Portugal no ano i, $i = 1, \cdots, 8$, $X_i \frown \mathcal{P}(0.7)$. Admitindo que o número de "grandes sismos" é independente de ano para ano, $Y = \sum_{i=1}^{8} X_i \frown \mathcal{P}(8 \times 0.7)$. $P[Y < 6] = F_Y(5) = 0.512$ (tabela da Poisson com $\lambda = 5.6$);

c) seja W a v.a. que presenta o número de anos, em 20, em que não ocorrem "grandes sismos"; esta v.a. conta sucessos em $n = 20$ provas de Bernoulli independentes (admitindo que a ocorrência de "grandes sismos" é independente de ano para ano) em que a probabilidade de sucesso em cada prova é a probabilidade de não ocorrer um "grande sismo", $p = P[X = 0] = \dfrac{e^{-0.7}0.7^0}{0!} = 0.497$. Assim, $W \frown \mathcal{B}(20, 0.497)$. Pede-se $P[W = 10] = F_W(10) - F_W(9) \approx 0.5881 - 0.4119 = 0.1762$ (os valores de F_W foram lidos na tabela da binomial com $n = 20$ e $p = 0.5$).

Modelos (Distribuições) Contínuos

As principais distribuições contínuas incluem:

- Distribuição uniforme contínua
- Distribuição normal

Distribuição uniforme contínua

Definição

Uma v.a. contínua diz-se ter distribuição uniforme ou rectangular no intervalo (a, b) e representa-se por $X \frown \mathcal{U}(a, b)$ se a função densidade de probabilidade (f.d.p.) é da forma:

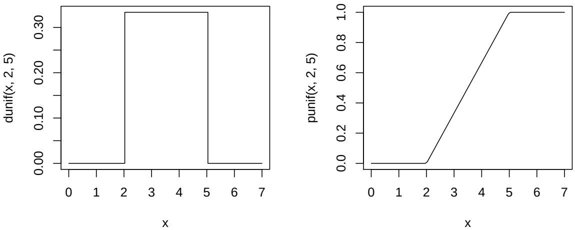
$$f(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & x \leq a \text{ ou } x \geq b. \end{cases}$$

A função distribuição cumulativa correspondente é (verificar como exercício):

$$F(x) = \begin{cases} 0 & x \leq a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \geq b. \end{cases}$$

Distribuição uniforme contínua

Gráficos da função densidade e da função distribuição cumulativa de uma v.a. $X \frown \mathcal{U}(2, 5)$:



Valor médio e variância

$X \frown \mathcal{U}(a, b)$

$$E[X] = \frac{a+b}{2} \qquad Var[X] = \frac{(b-a)^2}{12}$$

Distribuição normal ou de Gauss

É a distribuição contínua mais importante:

- é um modelo adequado para representar muitos fenómenos da vida real, nomeadamente características biométricas, erros de medição, etc;
- surge como a distribuição limite da soma (e da média) de muitas v.a. em certas condições (Teorema Limite Central, mais à frente);
- algumas distribuições (como a binomial e a Poisson) podem ser bem aproximadas pela normal.

Poderá ter interesse em ver este vídeo
`https://www.probabilitycourse.com/videos/chapter4/video4_9.php`

Distribuição normal ou de Gauss

Função densidade

A v.a. contínua X tem distribuição normal ou de Gauss com parâmetros μ e σ, $X \frown \mathcal{N}(\mu, \sigma)$ se a sua f.d.p. é da forma:

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

$x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma > 0.$

Os parâmetros da distribuição normal são:

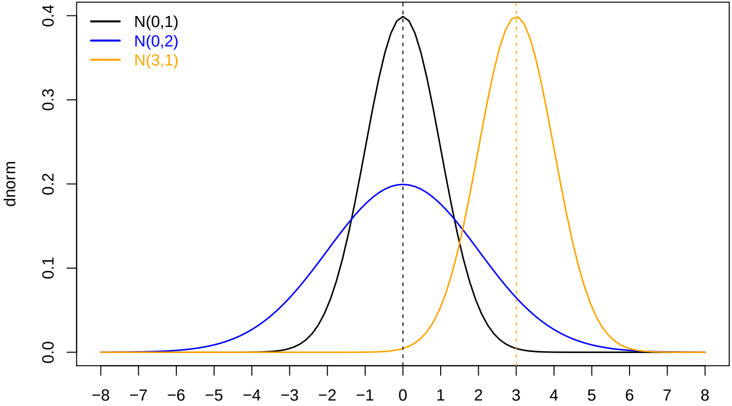
- $\mu \in \mathbb{R}$, o valor médio ou valor esperado de X
- $\sigma > 0$, o desvio padrão de X.

Portanto $X \frown \mathcal{N}(\mu, \sigma) \Rightarrow E[X] = \mu; \quad Var[X] = \sigma^2$.

Distribuição normal ou de Gauss

A função densidade

- tem a forma de sino com máximo em $x = \mu$;
- é simétrica em relação à reta vertical $x = \mu$: $f(\mu - x) = f(\mu + x)$;
- tem pontos de inflexão em $\mu - \sigma$ e $\mu + \sigma$.



Distribuição normal ou de Gauss

Normal reduzida

O caso particular $\mu = 0$ e $\sigma = 1$ corresponde à distribuição normal reduzida, $Z \frown \mathcal{N}(0, 1)$, com f.d.p.

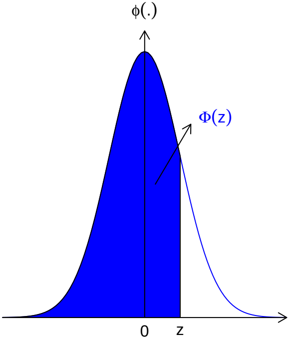
$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \, e^{-\frac{1}{2}z^2}$$

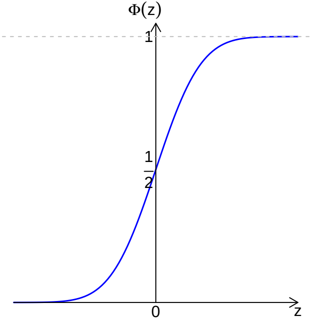
A função distribuição cumulativa é

$$\Phi(z) = P[Z \leq z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}u^2} \, du$$

$\Phi(z)$ é uma função não elementar; para o cálculo de probabilidades de uma v.a. normal é necessário recorrer a *software* ou tabelas.

Distribuição normal reduzida





Propriedades:

- $\Phi(0) = \frac{1}{2}$
- $\Phi(-z) = 1 - \Phi(z), \quad \forall z \in \mathbb{R}$

Distribuição normal ou de Gauss

Teorema

Seja $X \frown \mathcal{N}(\mu, \sigma)$ e $Y = a + bX$ então $Y \frown \mathcal{N}(a + b\mu,\ |b|\sigma)$.

O seguinte resultado permite usar as tabelas da distribuição $\mathcal{N}(0, 1)$ para calcular probabilidades de uma $\mathcal{N}(\mu, \sigma)$:

Corolário

Seja $X \frown \mathcal{N}(\mu, \sigma)$, então a v.a. $Z = \dfrac{X - \mu}{\sigma}$ tem distribuição normal reduzida, i.e., $Z = \dfrac{X - \mu}{\sigma} \frown \mathcal{N}(0, 1)$.

Distribuição normal ou de Gauss

Teorema da estabilidade da soma de normais

Sejam $X_1, ..., X_n$ v.a. normais independentes, tais que
$X_1 \frown \mathcal{N}(\mu_1, \sigma_1), \quad X_2 \frown \mathcal{N}(\mu_2, \sigma_2), \quad \cdots, \quad X_n \frown \mathcal{N}(\mu_n, \sigma_n).$
A v.a. $X_1 + X_2 + ... + X_n$ tem distribuição normal de parâmetros (μ, σ)
com $\mu = \mu_1 + \mu_2 + ... + \mu_n$ e $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2}$

Corolário

Sejam $X_1, ..., X_n$ v.a. normais independentes e semelhantes, i.e., tendo todas o mesmo valor médio μ e a mesma variância σ^2.

As variáveis aleatórias soma e média, definidas respetivamente como $S_n = \sum_{i=1}^n X_i$ e $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, têm distribuição normal assim definida

$$S_n \frown \mathcal{N}(n\mu, \sigma\sqrt{n}) \qquad \text{e} \qquad \overline{X}_n \frown \mathcal{N}(\mu, \sigma/\sqrt{n}).$$

Exemplo 9

Uma vacaria tem uma produção diária de leite que se admite seguir uma lei normal com $\mu = 950$ litro e $\sigma = 50$ litro.

a) Qual a probabilidade de se ter uma produção inferior a 1000 litros?

b) Qual a percentagem de dias em que a produção ultrapassa a produção média em mais de 100 litros?

c) Se na região existe outra vacaria, com uma produção diária que se admite normal com $\mu = 900$ *l* e $\sigma = 40$ *l*, funcionando independentemente da primeira, qual a probabilidade de num dado dia a produção total das duas vacarias ser superior a 1800 litros?

Exemplo 9 (resolução)

X: v.a. que representa a produção (litro) diária de leite,
$X \frown \mathcal{N}(950, 50)$

a) $P[X < 1000] = \Phi\left(\frac{1000 - 950}{50}\right) = \Phi(1) = 0.84134$ (Tabela)

b) $P[X > 950 + 100] = P\left[\frac{X - 950}{50} > \frac{100}{50}\right] = 1 - \Phi\left(\frac{100}{50}\right) =$
$1 - \Phi(2) \underbrace{=}_{\text{Tabela}} 1 - 0.97725 = 0.02275$

c) $Y \frown \mathcal{N}(900, 40)$
X, Y independentes $\Rightarrow \quad X + Y \frown \mathcal{N}(950 + 900, \sqrt{50^2 + 40^2})$
$P[X + Y > 1800] = 1 - \Phi\left(\frac{1800 - 1850}{\sqrt{50^2 + 40^2}}\right) \approx 1 - \Phi(-0.78) =$
$\Phi(0.78) = 0.7823$ (Tabela)

Exemplo 9 (Tabela da normal reduzida)

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169

Teorema Limite Central

O TLC estabelece que a **soma** ou **média** de um *elevado* número de v.a. independentes e identicamente distribuídas (com qualquer distribuição, *inclusive* discreta), têm distribuição **aproximadamente normal**:

Teorema Limite Central

Sejam $X_1, ..., X_n$ variáveis aleatórias independentes e identicamente distribuídas, com valor médio μ e variância σ^2 (finita).

Quando n é "grande", a soma $S_n = \sum_{i=1}^{n} X_i$ e a média $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ verificam:

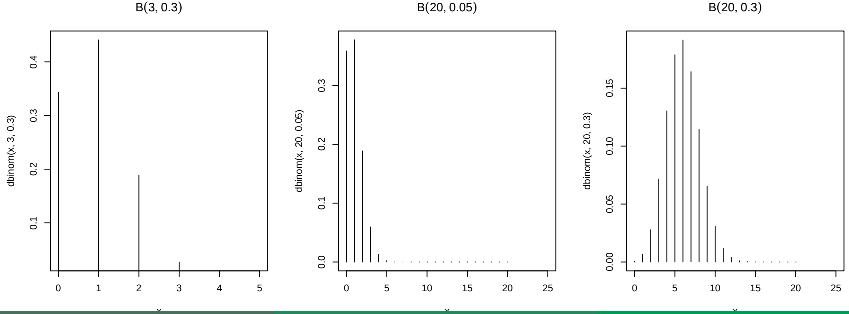
$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(0,1) \qquad \text{e} \qquad \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1).$$

Aplicações do Teorema Limite Central

Teorema de De Moivre ($\mathcal{B}inomial \rightarrow \mathcal{N}ormal$)

Seja X uma v.a. com distribuição $\mathcal{B}(n, p)$ com valor médio $\mu = np$ e variância $\sigma^2 = npq$. Então quando n é elevado e p afastado de 0 e de 1, em concreto $np > 5$ e $nq > 5$,

$$\frac{X - np}{\sqrt{npq}} \sim \mathcal{N}(0, 1)$$



Aplicações do Teorema Limite Central

Teorema: aproximação $\mathcal{P}oisson \rightarrow \mathcal{N}ormal$

Seja $X \frown \mathcal{P}(\lambda)$. Quando $\lambda \rightarrow \infty$ então $\dfrac{X-\lambda}{\sqrt{\lambda}} \sim \mathcal{N}(0,1)$.

Regra prática:
A aproximação é considerada boa para $\lambda \geq 12$.

Correção de continuidade

Quando se aproxima a distribuição binomial ou Poisson (discretas) pela distribuição normal (contínua) surge a seguinte desconformidade:

$X \frown \mathcal{B}$ ou $X \frown \mathcal{P} \Longrightarrow P[X \leq k] \neq P[X < k]$, k inteiro

mas
$X \frown \mathcal{N} \Longrightarrow P[X \leq k] = P[X < k]$

Para ultrapassar esta desconformidade, garantindo que $P[X \leq k] + P[X > k] = 1$, utiliza-se a correção de continuidade que consiste em representar cada número inteiro k pelo intervalo $\left(k - \frac{1}{2},\ k + \frac{1}{2}\right)$. Assim,

$$P[X < k] = P\left[X \leq k - \tfrac{1}{2}\right] \approx \Phi\left(\frac{k - \frac{1}{2} - \mu}{\sigma}\right)$$

$$P[X \leq k] = P\left[X \leq k + \tfrac{1}{2}\right] \approx \Phi\left(\frac{k + \frac{1}{2} - \mu}{\sigma}\right)$$

Correção de continuidade

Notas:

- as aproximações referidas são apenas necessárias quando se recorre a valores tabelados da função distribuição cumulativa da v.a. discreta (com limitações dos valores dos parâmetros);
- a correção de continuidade não garante uma melhor aproximação no sentido numérico;
- dispondo de *software* que tem implementada a função distribuição cumulativa das distribuições discretas para quaisquer valores dos seus parâmetros, não se utilizam as aproximações referidas.