Capítulo II: Variáveis Aleatórias Reais

Probabilidades e Aplicações
Licenciatura em Ciências da Computação
Universidade do Minho
Ano Letivo 2024/2025

1. Definição e Exemplos. Considerações gerais.

É frequente estarmos interessados em associar aos resultados de uma experiência aleatória uma, ou mais, características numéricas.

Por exemplo, para a experiência aleatória que consiste em efetuar dois lançamentos consecutivos de um dado, podemos estar interessados em estudar:

- o número de faces par obtidas;
- a soma das faces obtidas;
- a diferença, em valor absoluto, entre as faces obtidas;
- etc...

1. Definição e Exemplos. Considerações gerais.

Se estivermos interessados em apenas uma característica numérica, matematicamente tal é formalizado através de uma função que a cada elemento do espaço amostral, $\omega \in \Omega$, faz corresponder um número real, i.e., uma função

$$\begin{array}{rccc} X: & \Omega & \rightarrow & \mathbb{R} \\ & \omega & \rightarrow & X(\omega) \end{array}$$

em que Ω é o espaço amostral da experiência aleatória.

Exemplo:

Suponhamos que estamos interessados em estudar o número de faces par obtidas nos dois lançamentos consecutivos do dado. Neste caso,

$$\Omega = \{(x_1, x_2) : x_i \in \{1, 2, 3, 4, 5, 6\}, i \in \{1, 2\}\}$$

e a função X é tal que:

$$X((1,1)) = X((1,3)) = X((1,5)) = X((3,1)) = X((3,3)) = X((3,5)) =$$
$$= X((5,1)) = X((5,3)) = X((5,5)) = 0,$$

$$X((2,2)) = X((2,4)) = X((2,6)) = X((4,2)) = X((4,4)) = X((4,6)) =$$
$$= X((6,2)) = X((6,4)) = X((6,6)) = 2$$

e $X(\omega) = 1$ para os restantes elementos de Ω.

Se quisermos estudar, em simultâneo, k características numéricas, com $k \in \mathbb{N}$, somos conduzidos a uma função vectorial

$$\begin{array}{cccc} \mathbf{X}: & \Omega & \rightarrow & \mathbb{R}^k \\ & \omega & \rightarrow & (X_1(\omega), X_2(\omega), \ldots, X_k(\omega)), \end{array}$$

com $X_i : \Omega \rightarrow \mathbb{R}$ a i-ésima característica de interesse, $i \in \{1, \ldots, k\}$.

Neste capítulo, iremos estudar apenas o caso unidimensional ($k = 1$). O próximo capítulo será dedicado ao caso $k \geq 2$.

1. Definição e Exemplos. Considerações gerais.

Os **acontecimentos**, cujas probabilidades nos interessa calcular, são agora **expressos através de subconjuntos reais** elementos de $\mathcal{B}(\mathbb{R})$

No exemplo em que a função X representa o número de faces par nos 2 lançamentos, o acontecimento "não saiu qualquer face par" é dado por:

$$\begin{aligned} X^{-1}(\{0\}) &= \{\omega \in \Omega : X(\omega) = 0\} \\ &= \{(1,1),(1,3),(1,5),(3,1),(3,3),(3,5),(5,1),(5,3),(5,5)\}. \end{aligned}$$

E o acontecimento "saiu pelo menos uma face par" corresponde a

$$\begin{aligned} X^{-1}(\{1,2\}) &= \{\omega \in \Omega : X(\omega) \in \{1,2\}\} \\ &= \{(1,2),(1,4),(1,6),(2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ &\quad (3,2),(3,4),(3,6),(4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ &\quad (5,2),(5,4),(5,6),(6,1),(6,1),(6,3),(6,4),(6,5),(6,6)\}. \end{aligned}$$

Note que, para **esta** função X particular, tem-se $X^{-1}(\{1,2\}) = X^{-1}(\,[1,+\infty[\,)$.

1. Definição e Exemplos. Considerações gerais.

Para simplificar a notação, os acontecimentos anteriores podem ser abreviados por “$X = 0$” e “$X \in \{1, 2\}$”, respetivamente, i.e.,

$$X^{-1}(\{0\}) \equiv (X = 0)$$

e

$$X^{-1}(\{1, 2\}) \equiv (X \in \{1, 2\})$$

Observe que na descrição destes dois acontecimentos foram usados subconjuntos reais que são elementos de $\mathcal{B}(\mathbb{R})$, nomeadamente:

- $\{0\}$ no primeiro caso;
- $\{1, 2\}$ no segundo caso.

No segundo caso, poderia ser usado também $[1, +\infty[$.

1. Definição e Exemplos. Considerações gerais.

De um modo geral, estaremos interessados em calcular probabilidades de acontecimentos da forma

$$X^{-1}(E) = \{\omega \in \Omega : X(\omega) \in E\} \equiv (X \in E),$$

em que E é um subconjunto real elemento de $\mathcal{B}(\mathbb{R})$.

Observe que é preciso garantir que estes acontecimentos pertencem a $\mathcal{A}$, a σ-álgebra sobre o espaço amostral Ω, de modo a que a sua probabilidade esteja bem definida.

Assim, sendo $(\Omega, \mathcal{A}, P)$ o espaço de probabilidade da experiência, a função X deverá ser tal que

$$\forall E \in \mathcal{B}(\mathbb{R}),\ X^{-1}(E) \in \mathcal{A},$$

de modo a que

$$P(X \in E) \equiv P(X^{-1}(E)) = P(\{\omega \in \Omega : X(\omega) \in E\})$$

esteja definida para todo o $E \in \mathcal{B}(\mathbb{R})$.

Definição

Sejam $(\Omega, \mathcal{A}, P)$ um espaço de probabilidade e $X : \Omega \to \mathbb{R}$ uma função. X diz-se uma variável aleatória real (v.a.r.) se

$$\forall E \in \mathcal{B}(\mathbb{R}),\ X^{-1}(E) \in \mathcal{A}.$$

É difícil provar, usando a definição, que uma certa função é uma v.a.r.. Na prática, usaremos resultados mais simples para verificar se uma certa função é ou não uma v.a.r.. Em particular, o teorema seguinte será muito útil.

Teorema

Sejam $(\Omega, \mathcal{A}, P)$ um espaço de probabilidade e $X : \Omega \to \mathbb{R}$ uma função. X é v.a.r. sse

$$\forall c \in \mathbb{R},\ X^{-1}(\,]-\infty, c]\,) \in \mathcal{A}.$$

[Demonstração] [Ver Lopes e Gonçalves, 2000]

Exemplos: **1)** Considere a experiência aleatória que consiste em lançar uma moeda equilibrada. O espaço de probabilidade é $(\Omega, \mathcal{P}(\Omega), P)$, em que $\Omega = \{cara\,,\, coroa\}$ e P é a medida de Laplace. Usando o teorema anterior, vamos provar que a função $X : \Omega \to \mathbb{R}$ definida por $X(cara) = 1$ e $X(coroa) = 0$ é uma v.a.r..
Seja $c \in \mathbb{R}$ qualquer. Tem-se,

$$X^{-1}(\,]-\infty, c]\,) = \begin{cases} \emptyset & se & c < 0 \\ \{coroa\} & se & 0 \leq c < 1 \\ \Omega & se & c \geq 1 \end{cases} .$$

Uma vez que $\emptyset \in \mathcal{P}(\Omega)$, $\{coroa\} \in \mathcal{P}(\Omega)$ e $\Omega \in \mathcal{P}(\Omega)$, podemos afirmar que, qualquer que seja $c \in \mathbb{R}$, $X^{-1}(\,]-\infty, c]\,) \in \mathcal{P}(\Omega)$, ficando assim provado que X é uma v.a.r..
Observe que o acontecimento "saiu uma cara" corresponde a $X^{-1}(\{1\})$ (ou simplesmente "$X = 1$") e a sua probabilidade é

$$P(X = 1) \equiv P\left(X^{-1}(\{1\})\right) = P(\{cara\}) = \frac{1}{2}.$$

Exemplos: **2)** Sejam $(\Omega, \mathcal{A}, P)$ o espaço de probabilidade associado a uma certa experiência aleatória e B um acontecimento (i.e., $B \in \mathcal{A}$). Vamos provar que a função indicatriz do conjunto B, i.e., a função $\mathbf{1}_B : \Omega \to \mathbb{R}$ definida por $\mathbf{1}_B(\omega) = 1,$ se $\omega \in B$, e $\mathbf{1}_B(\omega) = 0,$ se $\omega \in \overline{B}$, é uma v.a.r..

Seja $c \in \mathbb{R}$ qualquer. Tem-se,

$$\mathbf{1}_B^{-1}(\,]-\infty, c]\,) = \begin{cases} \emptyset & se \quad c < 0 \\ \overline{B} & se \quad 0 \leq c < 1 \\ \Omega & se \quad c \geq 1 \end{cases}.$$

Uma vez que $\emptyset \in \mathcal{A}$, $\overline{B} \in \mathcal{A}$ (porque $B \in \mathcal{A}$) e $\Omega \in \mathcal{A}$, podemos concluir que, qualquer que seja $c \in \mathbb{R}$, $\mathbf{1}_B^{-1}(\,]-\infty, c]\,) \in \mathcal{A}$, e fica assim provado que $\mathbf{1}_B$ é uma v.a.r..

O acontecimento "$\mathbf{1}_B = 0$" corresponde a "não se obteve um elemento do conjunto B na realização da experiência aleatória" e tem-se

$$P(\mathbf{1}_B = 0) \equiv P\left(\mathbf{1}_B^{-1}(\{0\})\right) = P(\overline{B}\,) = 1 - P(B).$$

1. Definição e Exemplos. Considerações gerais.

Definição

Sejam $(\Omega, \mathcal{A}, P)$ um espaço de probabilidade e $X : \Omega \to \mathbb{R}$ uma v.a.r.. Chamamos σ-álgebra gerada por X à seguinte família de subconjuntos de Ω

$$X^{-1}(\mathcal{B}(\mathbb{R})) = \{X^{-1}(B), B \in \mathcal{B}(\mathbb{R})\}.$$

Note que, pela definição de v.a.r., tem-se obviamente

$$X^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{A}.$$

1. Definição e Exemplos. Considerações gerais.

Na prática, muitas vezes temos que lidar com uma função de uma v.a.r.. A questão que se coloca é em que condições é que a composição de uma função, real de variável real, com uma v.a.r. resulta ainda numa v.a.r..

Teorema

Sejam $(\Omega, \mathcal{A}, P)$ um espaço de probabilidade, $X : \Omega \to \mathbb{R}$ uma v.a.r. e $\phi : D \subseteq \mathbb{R} \to \mathbb{R}$. Se ϕ é uma função contínua então $\phi(X)$ também é uma v.a.r..

[Demonstração] [ver Lopes e Gonçalves, 2000]

Exemplos: Se X é uma v.a.r. então

- X^2 é uma v.a.r. (de uma forma geral, X^k, com $k \in \mathbb{N}$, é uma v.a.r.);
- e^X é uma v.a.r.;
- $|X|$ é uma v.a.r.;
- se $X > 0$, $\log(X)$ é uma v.a.r..

2. Lei de probabilidade e função de distribuição.

Sejam $(\Omega, \mathcal{A}, P)$ um espaço de probabilidade e $X : \Omega \to \mathbb{R}$ uma v.a.r..
Tem-se o seguinte diagrama:

$$\begin{array}{ccccc} & \Omega & \xrightarrow{X} & \mathbb{R} \\ \overset{P}{\Longleftarrow} & X^{-1}(\mathcal{B}(\mathbb{R})) & \overset{X^{-1}}{\Longleftarrow} & \mathcal{B}(\mathbb{R}) \end{array}$$

Definição

A função $P_X : \mathcal{B}(\mathbb{R}) \to [0,1]$ definida por $P_X = P \circ X^{-1}$, ie,

$$P_X(B) = P\left(X^{-1}(B)\right) \equiv P(X \in B),\ B \in \mathcal{B}(\mathbb{R}),$$

é designada de *lei de probabilidade* da v.a.r. X.

Observação: P_X é uma medida de probabilidade sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ [ver Ex. 1 da Folha Prática 4 para a demonstração].

2. Lei de probabilidade e função de distribuição.

Definição

A função $F_X : \mathbb{R} \to [0,1]$ definida por: para $c \in \mathbb{R}$,

$$F_X(c) = P_X(\,]-\infty, c]\,) = P(\,X^{-1}(\,]-\infty, c]\,)\,) \equiv P(X \in]-\infty, c]) \equiv P(X \leq c),$$

é designada de *função de distribuição* da v.a.r. X ou *função de distribuição da lei de probabilidade* P_X.

Observação: **[V. IMP.]**
Uma vez que $\pi(\mathbb{R}) = \{\,]-\infty, c], c \in \mathbb{R}\}$ é um π-sistema sobre $\mathbb{R}$ e é tal que $\sigma(\pi(\mathbb{R})) = \mathcal{B}(\mathbb{R})$, a lei de probabilidade P_X fica caracterizada pela respectiva função de distribuição F_X (recorde o Lema enunciado no final do Capítulo I). Assim, se uma outra v.a.r. Y tiver a mesma função de distribuição que X (i.e., se $F_X = F_Y$), então a lei de probabilidade de Y coincide com a lei de probabilidade de X, ou seja, tem-se

$$\forall\, B \in \mathcal{B}(\mathbb{R}),\; P_X(B) = P_Y(B).$$

2. Lei de probabilidade e função de distribuição.

Exemplo: Voltemos à experiência que consiste em lançar uma moeda equilibrada. O espaço de probabilidade é $(\Omega, \mathcal{P}(\Omega), P)$, em que $\Omega = \{cara,\ coroa\}$ e P é a medida de Laplace.
A função de distribuição da v.a.r. $X : \Omega \to \mathbb{R}$, definida por $X(cara) = 1$ e $X(coroa) = 0$, é dada por

$$\begin{aligned} c \in \mathbb{R},\ F_X(c) &= P(X \leq c) \equiv P\left(X^{-1}(\,]-\infty, c]\,)\right) \\ &= \begin{cases} P(\emptyset) & se \quad c < 0 \\ P(\{coroa\}) & se \quad 0 \leq c < 1 \\ P(\Omega) & se \quad c \geq 1 \end{cases} \\ &= \begin{cases} 0 & se \quad c < 0 \\ \frac{1}{2} & se \quad 0 \leq c < 1 \\ 1 & se \quad c \geq 1 \end{cases} . \end{aligned}$$

2. Lei de probabilidade e função de distribuição.

Propriedades de uma função de distribuição:

Sejam $(\Omega, \mathcal{A}, P)$ um espaço de probabilidade, $X : \Omega \to \mathbb{R}$ uma v.a.r. e F a função de distribuição de X. F tem as seguintes propriedades:

i) F é monótona não-decrescente;

ii) $\lim_{x \to -\infty} F(x) = 0$ e $\lim_{x \to +\infty} F(x) = 1$;

iii) para todo o a, $b \in \mathbb{R}$, com $a < b$, tem-se

$$P_X(\,]a, b]\,) \equiv P(a < X \leq b) = F(b) - F(a)$$

em que P_X é a lei de probabilidade da v.a.r. X;

iv) F é contínua à direita;

v) F é contínua em $x_0 \in \mathbb{R}$ sse $P_X(\{x_0\}) \equiv P(X = x_0) = 0$;

vi) F tem, quando muito, uma infinidade numerável de pontos de descontinuidade.

[Demonstração de i)-iii)] Ver Folhas Práticas .

[Demonstração de vi)] Ver livro Lopes & Gonçalves, 2000.

2. Lei de probabilidade e função de distribuição.

Demonstração de iv): Pretende-se mostrar que, para todo o $a \in \mathbb{R}$,

$$\lim_{c \to a^+} F(c) = F(a).$$

Como F é monótona e limitada, existe $\lim_{c \to a^+} F(c)$.

Vamos considerar uma sucessão $(c_n)_{n\in\mathbb{N}} \searrow a$, i.e., uma sucessão de números reais decrescente e convergente para a.

Observe que

$$\lim_{n\to\infty} F(c_n) = \lim_{n\to\infty} P_X(\,]-\infty, c_n]\,) = P_X\left(\bigcap_{n\in\mathbb{N}}]-\infty, c_n]\right) = P_X(\,]-\infty, a]\,) = F(a)$$

A segunda igualdade deve-se ao facto de $(]-\infty, c_n])_{n\in\mathbb{N}}$ ser uma sucessão decrescente de elementos de $\mathcal{B}(\mathbb{R})$ e de P_X ser uma medida de probabilidade sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Pela unicidade do limite, concluimos então que

$$\lim_{c \to a^+} F(c) = F(a).$$

c.q.d

Demonstração de v): Já vimos que F é contínua à direita. Falta apenas provar que F é contínua à esquerda de x_0 sse $P_X(\{x_0\}) = 0$. Observe que

$$\begin{aligned}
P_X(\{x_0\}) &= P_X(\,]-\infty, x_0]\,) - P_X(\,]-\infty,\ x_0[\,) \\
&= F(x_0) - P_X\left(\bigcup_{n\in\mathbb{N}} \left]-\infty,\ x_0 - \frac{1}{n}\right] \right) \\
&= F(x_0) - \lim_{n\to\infty} P_X\left(\left]-\infty,\ x_0 - \frac{1}{n}\right] \right) \\
&= F(x_0) - \lim_{n\to\infty} F\left(x_0 - \frac{1}{n} \right) \\
&= F(x_0) - \lim_{c\to x_0^-} F(c)
\end{aligned}$$

A terceira igualdade deve-se ao facto de $\left(\left]-\infty,\ x_0 - \frac{1}{n}\right]\right)_{n\in\mathbb{N}}$ ser uma sucessão crescente de elementos de $\mathcal{B}(\mathbb{R})$ e P_X ser uma medida de probabilidade sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Concluimos então que $P_X(\{x_0\}) = 0 \Leftrightarrow \lim_{c\to x_0^-} F(c) = F(x_0)$. c.q.d.

3. Variáveis Aleatórias Reais Discretas
3.1 Definição. Função de Probabilidade.

No que se segue, $(\Omega, \mathcal{A}, P)$ é um espaço de probabilidade, $X : \Omega \to \mathbb{R}$ é uma v.a.r. e P_X representa a lei de probabilidade de X.

Definição

Se existe um subconjunto real $D \in \mathcal{B}(\mathbb{R})$ finito ou infinito numerável tal que $P_X(D) \equiv P(X \in D) = 1$, diz-se que X é uma v.a.r. discreta e que P_X é uma lei de probabilidade discreta.

Ao menor subconjunto real $D \in \mathcal{B}(\mathbb{R})$ que verifica a condição $P_X(D) = 1$, chama-se *contradomínio* ou *suporte* da v.a.r. X e denota-se por C_X.

3. Variáveis Aleatórias Reais Discretas
3.1 Definição. Função de Probabilidade.

No que se segue, $(\Omega, \mathcal{A}, P)$ é um espaço de probabilidade, $X : \Omega \to \mathbb{R}$ é uma v.a.r. e P_X representa a lei de probabilidade de X.

Definição

Se existe um subconjunto real $D \in \mathcal{B}(\mathbb{R})$ finito ou infinito numerável tal que $P_X(D) \equiv P(X \in D) = 1$, diz-se que X é uma v.a.r. discreta e que P_X é uma lei de probabilidade discreta.
Ao menor subconjunto real $D \in \mathcal{B}(\mathbb{R})$ que verifica a condição $P_X(D) = 1$, chama-se *contradomínio* ou *suporte* da v.a.r. X e denota-se por C_X.

Teorema

Se P_X é uma lei de probabilidade discreta então o contradomínio de X é o conjunto de pontos de descontinuidade da respectiva função de distribuição.

[Demonstração] Ver livro de Lopes & Gonçalves

3.1 Definição. Função de Probabilidade.

Teorema

Seja X uma v.a.r. discreta de contradomínio C_X. A lei de probabilidade P_X é caracterizada pela função $f : \mathbb{R} \to [0,1]$ definida por

$$f(a) = \begin{cases} P_X(\{a\}) \equiv P(X = a) & se \quad a \in C_X \\ 0 & se \quad c.c. \end{cases} .$$

f é designada de *função de probabilidade da v.a.r.* X ou *função de probabilidade da lei* P_X. Também é usual chamar *função massa de probabilidade de X*.

[Demonstração] É evidente que dada P_X a função f fica completamente determinada. Suponhamos agora que f é conhecida e provemos que P_X também fica completamente definida. Ora, tem-se que, $\forall E \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} P_X(E) &= P_X\left(E \cap (C_X \cup \overline{C_X})\right) = P_X(\underbrace{E \cap C_X}_{\text{numerável}}) + \underbrace{P_X\left(E \cap \overline{C_X}\right)}_{=0} \\ &= \sum_{x \in (E \cap C_X)} P_X(\{x\}) = \sum_{x \in (E \cap C_X)} f(x) \end{aligned}$$

c.q.d.

3.1 Definição. Função de Probabilidade.

Observação:

Uma vez conhecida a função de probabilidade, f, da v.a.r. X de contra-domínio $C_X = \{x_1, x_2, \ldots, x_n, \ldots\}$, com $x_1 < x_2 < \ldots < x_n < \ldots$, a função de distribuição de X obtém-se do seguinte modo:

$$\begin{aligned} F_X(c) &= P_X(\,]-\infty, c]\,) \equiv P(X \leq c) = \sum_{x_i \in C_X \,:\, x_i \leq c} f(x_i) \\ &= \begin{cases} 0 & se \quad c < x_1 \\ f(x_1) & se \quad x_1 \leq c < x_2 \\ f(x_1) + f(x_2) & se \quad x_2 \leq c < x_3 \\ \vdots & \\ f(x_1) + f(x_2) + \ldots + f(x_n) & se \quad x_n \leq c < x_{n+1} \\ \vdots & \end{cases} \end{aligned}$$

3.1 Definição. Função de Probabilidade.

Observe ainda que, se X é uma v.a.r. discreta, de contradomínio C_X e função de probabilidade f, então:

1. qualquer que seja $E \in \mathcal{B}(\mathbb{R})$, tem-se

$$P_X(E) \equiv P(X \in E) = \sum_{a \in (C_X \cap E)} f(a),$$

   como já foi visto na demonstração do último teorema e usado na construção da função de distribuição.

2. da definição de C_X, resulta obviamente que

$$\sum_{a \in C_X} f(a) = 1.$$

3.2 Leis de Probabilidade Discretas Mais Conhecidas

Nesta secção vamos estudar algumas leis de probabilidade discretas que são frequentemente utilizadas na prática.

3.2 Leis de Probabilidade Discretas Mais Conhecidas

Nesta secção vamos estudar algumas leis de probabilidade discretas que são frequentemente utilizadas na prática.

I) Lei Binomial com parâmetros n e p, com $n \in \mathbb{N}$, $p \in]0,1[$

Seja $(\Omega, \mathcal{F}, P)$ o espaço de probabilidade de uma experiência aleatória, ξ, e seja S um acontecimento que, numa realização de ξ, ocorre com probabilidade $0 < p < 1$, i.e., $S \in \mathcal{F}$ tal que $P(S) = p$.
Considere agora a v.a.r. X que representa o número de vezes que o acontecimento S ocorre em n repetições independentes de ξ.
Tem-se que X é uma v.a.r. discreta, de contradomínio $C_X = \{0, 1, \ldots, n\}$ e com função de probabilidade $f : \mathbb{R} \to [0,1]$ dada por

$$f(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & se \quad k \in \{0, 1, \ldots, n\} \\ 0 & se \quad c.c. \end{cases} .$$

Nestas condições, diz-se que a v.a.r. *X segue a lei Binomial com parâmetros n e p*, e abrevia-se por $X \sim Bin(n, p)$.
Nota: O acontecimento S é usualmente designado de "sucesso".

II) Lei de Bernoulli com parâmetro p, com $p \in]0,1[$

Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade e $A \in \mathcal{F}$ tal que $P(A) = p$, com $0 < p < 1$. Já vimos que a função $X : \Omega \to \mathbb{R}$ definida por

$$X(w) = \begin{cases} 0 & se \quad w \notin A \\ 1 & se \quad w \in A \end{cases}$$

é uma v.a.r. e a sua lei de probabilidade, P_X, é tal que $P_X(\{1\}) = p$ e $P_X(\{0\}) = 1 - p$. Como $p \in]0,1[$, temos que $C_X = \{0,1\}$, pelo que X é uma v.a.r. discreta. A função de probabilidade de X, $f : \mathbb{R} \to [0,1]$, é dada por

$$f(k) = \begin{cases} p & se \quad k = 1 \\ 1-p & se \quad k = 0 \\ 0 & se \quad c.c. \end{cases} .$$

Nestas condições, diz-se que a v.a.r. X segue a lei de Bernoulli com parâmetro p, e abrevia-se por $X \sim Bernoulli(p)$.

Observação: A lei $Bernoulli(p)$ coincide com a lei $Bin(1,p)$.

III) Lei Hipergeométrica com parâmetros R, M e n

Suponha que numa caixa estão R elementos, dos quais M, com $0 \leq M \leq R$, possuem um certo atributo A e os restantes $R - M$ elementos não têm este atributo.
Considere a experiência aleatória que consiste em recolher uma amostra, sem reposição, de n elementos retirados da caixa e seja X a v.a.r. que representa o número de elementos da amostra que possuem o atributo A. Tem-se que X é uma v.a.r. discreta, de contradomínio

$$C_X = \{\max(0, n-(R-M)), \ldots, \min(n, M)\}$$

e função de probabilidade

$$f(k) = \begin{cases} \frac{\binom{M}{k}\binom{R-M}{n-k}}{\binom{R}{n}} & \text{se} \quad k \in C_X \\ 0 & \text{se} \quad c.c. \end{cases} .$$

Nestas condições, diz-se que a v.a.r. *X segue a lei Hipergeométrica com parâmetros R, M e n*, e abrevia-se por $X \sim HG(R, M, n)$.
Nota: Se a amostra for feita com reposição, então $X \sim \text{Bin}\left(n, \frac{M}{R}\right)$.

Exemplo: Considere um lote de 10 peças, em que 4 são defeituosas e as restantes 6 não são defeituosas.

- Se escolhermos, ao acaso e sem reposição, 8 peças deste lote e considerarmos X a v.a.r. que representa o número de peças defeituosas entre as 8 escolhidas, temos que

$$X \sim HG(10, 4, 8)$$

  e que $C_X = \{2, 3, 4\}$.

- Se escolhermos, ao acaso e com reposição, 8 peças deste lote e considerarmos Y a v.a.r. que representa o número de peças defeituosas entre as 8 escolhidas, temos que

$$Y \sim Bin\left(8, \frac{4}{10}\right)$$

  e que $C_Y = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$.

[Fim da matéria para o 1.º teste]

IV) Lei de Poisson com parâmetro λ, com $\lambda \in \mathbb{R}^+$:

Seja $(X_n)_{n\in\mathbb{N}}$ uma sucessão de v.a.r.'s, todas definidas sobre o mesmo espaço de probabilidade $(\Omega, \mathcal{A}, P)$, e tais que $X_n \sim Bin(n, p_n)$ com os parâmetros n e p_n a satisfazer

$$\lim_{n\to+\infty} p_n = 0 \text{ e } \lim_{n\to+\infty} np_n = \lambda,$$

com $\lambda \in \mathbb{R}^+$.
Nestas condições, tem-se que

$$\frac{P(X_n = k)}{P(X_n = k-1)} = \frac{\binom{n}{k} p_n^k (1-p_n)^{n-k}}{\binom{n}{k-1} p_n^{k-1} (1-p_n)^{n-k+1}} = \frac{n-k+1}{k} \frac{p_n}{1-p_n} \underset{n\to+\infty}{\longrightarrow} \frac{\lambda}{k}.$$

Isto permite-nos concluir que, para n suficientemente grande, a função de probabilidade da v.a.r. X_n comporta-se como a de uma v.a.r. discreta, Y, de contradomínio $C_Y = \mathbb{N}_0$, e tal que

$$P(Y = k) = \frac{\lambda}{k} P(Y = k-1),\; k \in \mathbb{N}.$$

Trabalhando esta última igualdade, concluimos que

$$P(Y=k)=\frac{\lambda}{k}P(Y=k-1)=\frac{\lambda^2}{k(k-1)}P(Y=k-2)=\ldots=\frac{\lambda^k}{k!}P(Y=0).$$

Como $C_Y=\mathbb{N}_0$, temos

$$1=\sum_{k=0}^{+\infty}P(Y=k)\Leftrightarrow 1=\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!}P(Y=0)\Leftrightarrow 1=P(Y=0)e^{\lambda}\Leftrightarrow P(Y=0)=e^{-\lambda}$$

Concluimos assim que, para todo o $k\in\mathbb{N}_0$, $P(Y=k)=\frac{\lambda^k}{k!}e^{-\lambda}$ e que a função de probabilidade de Y é dada por

$$f(k)=\begin{cases}\frac{\lambda^k}{k!}e^{-\lambda} & \text{se} \quad k\in\mathbb{N}_0\\ 0 & \text{se} \quad \text{c.c.}\end{cases}.$$

Nestas condições, diz-se que Y segue a lei de Poisson com parâmetro λ, e abrevia-se por $Y\sim Poisson(\lambda)$.

Nota:
A lei de Poisson é adequada para modelar o número de ocorrências de um fenómeno raro (i.e. um fenómeno que tem baixa probabilidade de ocorrer) quando não limitamos o número de repetições da experiência aleatória.

Em particular, a função de probabilidade da lei de Poisson é usada para obter um valor aproximado da função de probabilidade de uma v.a.r. $Z \sim Bin(n, p)$ quando n é grande e p é pequeno. O parâmetro λ a utilizar na aproximação será igual a $n \times p$.

V) Lei Geométrica, com parâmetro p, com $p \in]0,1[$:

Considere uma experiência aleatória na qual um certo acontecimento, que designamos por “sucesso”, ocorre com probabilidade $0 < p < 1$ (e ocorre “insucesso” com probabilidade $1-p$). Suponhamos agora que se repete esta experiência, sempre nas mesmas condições (i.e., repetições independentes) e seja T a v.a. que representa o número de vezes que se efetua a experiência até ocorrer “sucesso” pela primeira vez.

Naturalmente, tem-se que T é discreta e que $C_T = \{1, 2, \ldots\} = \mathbb{N}$.

Para determinar a função de probabilidade de T, considerem-se os seguintes acontecimentos:

A_i:"ocorreu insucesso na i-ésima vez que se efetuou a experiência", $i = 1, 2, \ldots$.
Note-se que $P(A_i) = 1 - p$ e $P(\overline{A_i}) = p$, para todo o i.

Usando estes acontecimentos (e observe que $A_1, A_2, \ldots, A_k$ formam uma família de k acontecimentos independentes) podemos facilmente obter alguns dos valores da função de probabilidade de T:
$P(T = 1) = P(\overline{A_1}) = p$;

3.2 Leis de Probabilidade Discretas Mais Conhecidas

Para determinar a função de probabilidade de T, considerem-se os seguintes acontecimentos:

A_i:"ocorreu <u>insucesso</u> na i-ésima vez que se efetuou a experiência", $i = 1, 2, \ldots$.

Note-se que $P(A_i) = 1 - p$ e $P(\overline{A_i}\,) = p$, para todo o i.

Usando estes acontecimentos (e observe que $A_1, A_2, \ldots, A_k$ formam uma família de k acontecimentos independentes) podemos facilmente obter alguns dos valores da função de probabilidade de T:

$P(T = 1) = P(\overline{A_1}\,) = p;$

$P(T = 2) = P(A_1 \cap \overline{A_2}\,) = P(A_1) \times P(\overline{A_2}) = (1 - p)p;$

Para determinar a função de probabilidade de T, considerem-se os seguintes acontecimentos:

A_i:"ocorreu insucesso na i-ésima vez que se efetuou a experiência", $i = 1, 2, \ldots$.
Note-se que $P(A_i) = 1 - p$ e $P(\overline{A_i}) = p$, para todo o i.

Usando estes acontecimentos (e observe que $A_1, A_2, \ldots, A_k$ formam uma família de k acontecimentos independentes) podemos facilmente obter alguns dos valores da função de probabilidade de T:

$P(T = 1) = P(\overline{A_1}) = p$;

$P(T = 2) = P(A_1 \cap \overline{A_2}) = P(A_1) \times P(\overline{A_2}) = (1 - p)p$;

$P(T = 3) = P(A_1 \cap A_2 \cap \overline{A_3}) = P(A_1) \times P(A_2) \times P(\overline{A_3}) = (1 - p)^2 p$;

Para determinar a função de probabilidade de T, considerem-se os seguintes acontecimentos:

A_i:"ocorreu insucesso na i-ésima vez que se efetuou a experiência", $i = 1, 2, \ldots$.

Note-se que $P(A_i) = 1 - p$ e $P(\overline{A_i}) = p$, para todo o i.

Usando estes acontecimentos (e observe que $A_1, A_2, \ldots, A_k$ formam uma família de k acontecimentos independentes) podemos facilmente obter alguns dos valores da função de probabilidade de T:

$P(T = 1) = P(\overline{A_1}) = p;$

$P(T = 2) = P(A_1 \cap \overline{A_2}) = P(A_1) \times P(\overline{A_2}) = (1 - p)p;$

$P(T = 3) = P(A_1 \cap A_2 \cap \overline{A_3}) = P(A_1) \times P(A_2) \times P(\overline{A_3}) = (1 - p)^2 p;$

$\vdots$

$P(T = k) = P(A_1 \cap A_2 \cap \ldots \cap A_{k-1} \cap \overline{A_k}) = p(1 - p)^{k-1}, \quad k \in \mathbb{N}$

3.2 Leis de Probabilidade Discretas Mais Conhecidas

Tem-se assim a seguinte definição:

Definição

Sejam T uma v.a. discreta e $p \in]0,1[$.
Diz-se que *T segue a distribuição Geométrica com parâmetro p*, e abrevia-se por $T \sim Geo(p)$, se o seu contradomínio é $\mathbb{N}$ e a sua função de probabibilidade é dada por

$$f(k) = \begin{cases} p(1-p)^{k-1} & se \quad k \in \mathbb{N} \\ 0 & se \quad c.c. \end{cases}$$

Observações: Se $T \sim Geo(p)$,
1) Facilmente se verifica que, para todo o $k \in \mathbb{N}$,

$$P(T > k) = (1-p)^k.$$

2) T tem a conhecida propriedade de falta de memória, i.e.,

$$P(T = k+n | T > k) = P(T = n),$$

para todo o $k, n \in \mathbb{N}$.

Exemplo/Exercício: Imagine um bêbado que tem n chaves na sua carteira e que, ao chegar a casa, não consegue identificar a única chave que abre a porta. Como está tão bêbado, de cada vez que ele tenta uma chave que não é a certa, não consegue colocá-la de lado pelo que na tentativa seguinte volta a ter n chaves disponíveis para a escolha. Calcule a probabilidade de ele:

i) acertar à primeira;
ii) acertar pela primeira vez na terceira tentativa;
iii) errar as primeiras 5 tentativas;
iv) acertar pela primeira vez na oitava tentativa, sabendo que errou nas primeiras 5.

Sugestão: Identificar uma v.a. relevante para o problema e que tenha distribuição Geométrica.

VI) Lei Uniforme num conjunto finito U:

Seja $U = \{u_1, u_2, \ldots, u_n\}$ um subconjunto real finito, com n elementos. Diz-se que uma v.a.r. X segue a lei Uniforme no conjunto U, abrevia-se por $X \sim Uniforme(U)$, se a função de probabilidade é dada por

$$f(a) = \begin{cases} \frac{1}{n} & \text{se} \quad a \in U \\ \\ 0 & \text{se} \quad \text{c.c.} \end{cases} .$$

Na prática, esta lei é utilizada sempre que se escolhe, ao acaso, um elemento do conjunto U e os diferentes elementos de U têm igual probabilidade de serem escolhidos.

4. Variáveis aleatórias reais (absolutamente) contínuas
4.1 Leis Difusas

Definição

Seja $(\Omega, \mathcal{A}, P)$ um espaço de probabilidade e $X : \Omega \to \mathbb{R}$ uma v.a.r..
X diz-se difusa se a sua lei de probabilidade, P_X, for uma medida de probabilidade difusa sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, i.e., se P_X for tal que

$$P_X(\{a\}) \equiv P(X = a) = 0, \forall a \in \mathbb{R}.$$

Nota: Se X é uma v.a.r. difusa então a função de distribuição de X, F_X, é uma função contínua. Ver propriedades da função de distribuição.

Entre as leis difusas, existe um subconjunto particularmente importante e que vamos estudar: o das leis de probabilidade absolutamente contínuas. Tais leis caracterizam-se à custa de uma função, real de variável real, chamada de *função densidade de probabilidade*.

4.2 Leis absolutamente contínuas

Definição

Uma função $f : \mathbb{R} \to \mathbb{R}$ diz-se uma função densidade de probabilidade sobre $\mathbb{R}$ se:

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

Exemplos: Das funções indicadas no Exercício 5, da Folha Prática 1, quais as que são funções densidade de probabilidade?

4.2 Leis absolutamente contínuas

Definição

Uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ diz-se uma função densidade de probabilidade sobre $\mathbb{R}$ se:

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

Exemplos: Das funções indicadas no Exercício 5, da Folha Prática 1, quais as que são funções densidade de probabilidade?

- i), iv), v) e vi) são funções densidade de probabilidade;
- iii) não é uma função densidade de probabilidade porque não é integrável.
- ii), vii) e viii) não são funções densidade de probabilidade porque, apesar de integráveis, não satisfazem a condição $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Nota: Recorde que todas estas funções são não-negativas.

4.2 Leis absolutamente contínuas

Definição

Uma v.a.r. X diz-se absolutamente contínua se a sua lei de probabilidade, P_X, é uma lei absolutamente contínua sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, i.e., se existe uma função densidade de probabilidade sobre $\mathbb{R}$, f, tal que

$$\underset{a,b\in\mathbb{R},\ a<b}{\forall}, \ P_X(\,]a,b[\,) \equiv P(\,X \in]a,b[\,) = \int_a^b f(x)dx.$$

À função f chamamos função densidade de probabilidade da v.a.r. X (ou da lei de probabilidade P_X). Ao menor elemento de $\mathcal{B}(\mathbb{R})$ onde a função f é estritamente positiva chamamos suporte ou contradomínio de X.

Observação: É possível mostrar que toda a função densidade de probabilidade sobre $\mathbb{R}$ determina uma única medida de probabilidade Q sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ absolutamente contínua que verifica a condição

$$\underset{a,b\in\mathbb{R},\ a<b}{\forall}, \ Q(\,]a,b[\,) = \int_a^b f(x)dx.$$

Teorema

Se Q é uma lei de probabilidade sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ absolutamente contínua então Q é difusa.

[Demonstração]: Pretende-se provar que $Q(\{a\}) = 0$, para todo o $a \in \mathbb{R}$. Comece por observar que

$$\{a\} = \bigcap_{n \in \mathbb{N}} \left]a - \frac{1}{n}, a + \frac{1}{n}\right[.$$

Sendo Q absolutamente contínua então Q admite uma função densidade de probabilidade e seja f essa função. Tem-se, então, que

$$Q(\{a\}) = \lim_{n \to +\infty} Q\left(\left]a - \frac{1}{n}, a + \frac{1}{n}\right[\right) = \lim_{n \to +\infty} \int_{a-\frac{1}{n}}^{a+\frac{1}{n}} f(x)dx = 0.$$

A primeira igualdade deve-se ao facto de Q ser uma medida de probabilidade sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ e a segunda deve-se ao facto de Q ser uma lei absolutamente contínua com função densidade de probabilidade f.

c.q.d.

4.2 Leis absolutamente contínuas

Observação: Deste último teorema, resulta trivialmente que, se X é uma v.a.r. absolutamente contínua, então a respectiva lei, P_X, satisfaz as seguintes igualdades: para todo $a, b \in \mathbb{R},\ a < b$,

$$P_X(\,]a,b[\,) = P_X(\,[a,b[\,) = P_X(\,]a,b]\,) = P_X(\,[a,b]\,) = \int_a^b f(x)dx,$$

i.e.,

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

i.e.,

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

4.2 Leis absolutamente contínuas

Observação: Deste último teorema, resulta trivialmente que, se X é uma v.a.r. absolutamente contínua, então a respectiva lei, P_X, satisfaz as seguintes igualdades: para todo $a, b \in \mathbb{R},\ a < b$,

$$P_X(\,]a,b[\,) = P_X(\,[a,b[\,) = P_X(\,]a,b]\,) = P_X(\,[a,b]\,) = \int_a^b f(x)dx,$$

i.e.,

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

i.e.,

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

O teorema seguinte caracteriza as leis de probabilidade sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ que são absolutamente contínuas. Em particular, é estabelecida uma relação entre a função densidade de probabilidade e a função de distribuição de uma tal lei.

4.2 Leis absolutamente contínuas

Teorema

Uma condição necessária e suficiente para que uma v.a.r. X seja absolutamente contínua é que a sua função de distribuição, F_X, verifique a seguinte condição

$$\forall c \in \mathbb{R},\ F_X(c) = P_X(\,]-\infty, c]\,) \equiv P(X \leq c) = \int_{-\infty}^{c} f(x)dx,$$

para alguma função densidade de probabilidade f.

[Demonstração]
($\Rightarrow$) Suponhamos que X é uma v.a.r. absolutamente contínua. Então existe f, uma função densidade de probabilidade sobre $\mathbb{R}$, tal que

$$\forall\, a,\, b \in \mathbb{R}\ :\ a < b,\ \ P_X(\,]a, b[\,) = \int_a^b f(x)dx.$$

Então, para todo o $c \in \mathbb{R}$,

$$\begin{aligned} F_X(c) &= P_X(]-\infty, c]) = P_X(\,]-\infty, c[\,) = P_X\left(\bigcup_{n\in\mathbb{N}}]-n, c[\right) \\ &= \lim_{n\to+\infty} P_X(\,]-n, c[\,) \\ &= \lim_{n\to+\infty} \int_{-n}^{c} f(x)dx \\ &= \int_{-\infty}^{c} f(x)dx. \end{aligned}$$

Observe que a segunda igualdade deve-se ao facto de P_X ser uma lei difusa; a quarta igualdade deve-se a P_X ser medida de probabilidade sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (Propriedade VI)); a penúltima igualdade deve-se a X ser absolutamente contínua com função densidade de probabilidade f; a última igualdade deve-se ao facto de f ser integrável.

($\Leftarrow$) Suponhamos agora que X é uma v.a.r. cuja função de distribuição, F_X, é dada por

$$\forall\, c \in \mathbb{R}, \;\; F_X(c) = \int_{-\infty}^{c} f(x)dx$$

para alguma função densidade de probabilidade f. Se provarmos que

$$\forall\, a,\, b \in \mathbb{R} \,:\, a < b, \;\; P_X(\,]a,b[\,) = \int_a^b f(x)dx$$

fica demonstrado que X é uma v.a.r. absolutamente contínua. Para provar isto, basta mostrar as seguintes igualdades:

a) $P_X(\,]a,b]\,) = \int_a^b f(x)dx$,
b) $P_X(\{b\}) = 0$.

Ora

$$P_X(\,]a,b]\,) = F_X(b) - F_X(a) = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx = \int_{a}^{b} f(x)dx,$$

ficando assim provado a). Observe que a primeira igualdade se deve a uma das propriedades de uma função de distribuição (propriedade iii)). Adicionalmente,

$$\begin{aligned} P_X(\{b\}) &= P_X\left(\bigcap_{n\in\mathbb{N}} \left]b-\frac{1}{n}, b+\frac{1}{n}\right]\right) \\ &= \lim_{n\to+\infty} P_X\left(\left]b-\frac{1}{n}, b+\frac{1}{n}\right]\right) \\ &= \lim_{n\to+\infty} \int_{b-\frac{1}{n}}^{b+\frac{1}{n}} f(x)dx \\ &= 0, \end{aligned}$$

e fica provado b). Observe que a segunda igualdade se deve ao facto de P_X ser medida de probabilidade sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (Propriedade VII)) e a penúltima igualdade faz uso de a). . c.q.d.

4.2 Leis absolutamente contínuas

O último teorema caracteriza as v.a.r.'s absolutamente contínuas e estabelece como se determina a função de distribuição à custa da função densidade de probabilidade.

O último teorema caracteriza as v.a.r.'s absolutamente contínuas e estabelece como se determina a função de distribuição à custa da função densidade de probabilidade.
A questão que agora se coloca é: como obter a função densidade de probabilidade, f, à custa da função de distribuição, F?

4.2 Leis absolutamente contínuas

O último teorema caracteriza as v.a.r.'s absolutamente contínuas e estabelece como se determina a função de distribuição à custa da função densidade de probabilidade.

A questão que agora se coloca é: como obter a função densidade de probabilidade, f, à custa da função de distribuição, F?

É possível demonstrar [ver Lopes & Gonçalves] que a função de distribuição de uma v.a.r. absolutamente contínua é derivável excepto, quando muito, num subconjunto infinito numerável de $\mathbb{R}$. Mais, pode-se ainda provar que:

1. Se F' é contínua então $f(x) = F'(x),\ \forall x \in \mathbb{R}$.
2. Se F' é contínua excepto, quando muito, num subconjunto infinito numerável $D \subset \mathbb{R}$, e se F' é limitada então

$$f(x) = F'(x), \text{ para todo o } x \in \mathbb{R} \setminus D.$$

Neste último caso, para os pontos $x \in D$ convenciona-se $f(x) = 0$.

Na verdade, podemos atribuir um qualquer valor não negativo à função f no conjunto D.
Uma vez que D é, quando muito, infinito numerável, qualquer que seja o valor (não negativo) atribuído a $f(x)$ em pontos $x \in D$, tem-se sempre

$$P_X(D) = 0.$$

Deste modo, podem existir várias funções densidade de probabilidade para uma mesma v.a.r. X absolutamente contínua. No entanto, elas só podem diferir umas das outras num subconjunto finito ou infinito numerável de $\mathbb{R}$.

4.3 Leis de probabilidade contínuas mais conhecidas

Nesta secção vamos estudar algumas leis de probabilidade absolutamente contínuas que são frequentemente utilizadas na prática.

Nesta secção vamos estudar algumas leis de probabilidade absolutamente contínuas que são frequentemente utilizadas na prática.

I) Lei Uniforme no intervalo $[a,\, b]$, $a,\, b \in \mathbb{R}$, $a < b$:

Diz-se que uma v.a.r. absolutamente contínua, X, segue a lei Uniforme no intervalo $[a,\, b]$, abrevia-se por $X \sim U([a,\, b])$, se a função densidade de probabilidade de X é dada por

$$f(x) = \begin{cases} \frac{1}{b-a} & se \quad a \leq x \leq b \\ 0 & se \quad c.c. \end{cases} \quad .$$

Esta lei atribui a mesma probabilidade a intervalos de igual amplitude contidos em $[a,\, b]$. De facto, se $]c, d[\, \subseteq [a, b]$ tem-se que

$$P(X \in \,]c, d[\,) = \int_c^d f(x)dx = \left[\frac{x}{b-a}\right]_c^d = \frac{d-c}{b-a} = \frac{\text{amplitude de }]c,\, d[}{b-a}.$$

4.3 Leis de probabilidade contínuas mais conhecidas

Se $X \sim U([a, b])$ então a respectiva função de distribuição é dada por

$$F(c) = P_X(]-\infty, c]) = \int_{-\infty}^{c} f(x)dx = \begin{cases} 0 & se \quad c < a \\ \frac{c-a}{b-a} & se \quad a \leq c \leq b \\ 1 & se \quad c > b. \end{cases} .$$

Uma das utilizações mais importantes da lei uniforme é apresentada no teorema da transformação uniformizante, que diz o seguinte:

Teorema [Transformação Uniformizante]

Seja X uma v.a.r. absolutamente contínua e cuja função de distribuição, G, é estritamente crescente. Então a v.a.r. $Y = G(X) \sim U([0, 1])$.

[Demonstração] A função de distribuição de Y, F_Y, é dada por

$$F_Y(c) = P(G(X) \leq c) = \begin{cases} 0 & se \quad c < 0 \\ P(X \leq G^{-1}(c)) = G(G^{-1}(c)) = c & se \quad 0 \leq c \leq 1 \\ 1 & se \quad c > 1 \end{cases} ,$$

que coincide com a função de distribuição da lei $U([0, 1])$. c.q.d.

II) Lei Exponencial de parâmetro λ, $\lambda \in \mathbb{R}^+$:

Diz-se que uma v.a.r. absolutamente contínua, X, segue a lei Exponencial de parâmetro λ, com $\lambda \in \mathbb{R}^+$, abrevia-se por $X \sim Exp(\lambda)$, se a função densidade de probabilidade de X é dada por

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

Se $X \sim Exp(\lambda)$ então a respectiva função de distribuição é dada por

$$F(c) = P_X(]-\infty, c]) = \int_{-\infty}^{c} f(x)dx = \begin{cases} 0 & se \quad c < 0 \\ 1 - e^{-\lambda c} & se \quad c \geq 0 \end{cases}.$$

A propriedade mais importante da lei Exponencial é a conhecida falta de memória: para todo o $x, t \in \mathbb{R}^+$ tem-se

$$P(T > t + x | T > t) = P(T > x).$$

Observação: Esta propriedade é partilhada com a lei Geométrica (que é uma lei discreta) [ver Secção 3.2].

III) Lei Normal (ou Gaussiana) de parâmetros μ e σ^2, $\mu \in \mathbb{R}$ e $\sigma \in \mathbb{R}^+$:
Diz-se que uma v.a.r. absolutamente contínua, X, segue a lei Normal de parâmetros μ e σ^2, abrevia-se por $X \sim N(\mu, \sigma^2)$, se a função densidade de probabilidade de X for dada por

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, \; x \in \mathbb{R}.$$

Nota: A lei $N(0,1)$ é designada de lei *normal standard* ou lei *normal centrada e reduzida* ou lei *normal padrão*.

A Normal tem várias propriedades interessantes, o que a torna um lei muito relevente em problemas de modelação e de inferência estatística.

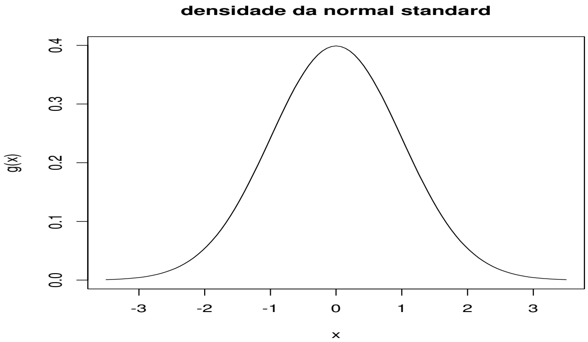
Vamos começar por provar que a função f acima é mesmo uma função densidade de probabilidade sobre $\mathbb{R}$, i.e., que $f(x) \geq 0$ (trivial) e que

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

4.3 Leis de probabilidade contínuas mais conhecidas

A título de exemplo, apresenta-se o gráfico da função densidade de probabilidade da lei $N(0,1)$, i.e., da função

$$g(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}, \; x \in \mathbb{R}.$$



Vamos começar pela função densidade de probabilidade da $N(0,1)$,

$$g(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}, \; x \in \mathbb{R}$$

e provar que $I \equiv \int_{-\infty}^{+\infty} g(x)dx = 1$. Como $g(x) > 0$, $\forall x \in \mathbb{R}$, se provarmos que $I^2 = 1$ teremos que a única possibilidade é $I = 1$. E, de facto:

$$\begin{aligned}
I^2 &= \int_{-\infty}^{+\infty} g(x)dx \times \int_{-\infty}^{+\infty} g(y)dy \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{2\pi} \exp\left\{-\frac{1}{2}(x^2+y^2)\right\} dx\, dy \\
&\overset{(*)}{=} \int_0^{+\infty}\int_0^{2\pi} \frac{1}{2\pi} \exp\left\{-\frac{1}{2}r^2\right\} r\, d\theta dr = \int_0^{+\infty} \frac{r}{2\pi} \exp\left\{-\frac{1}{2}r^2\right\} dr[\theta]_0^{2\pi} \\
&= \int_0^{+\infty} r \exp\left\{-\frac{1}{2}r^2\right\} dr = \left[-\exp\left\{-\frac{1}{2}r^2\right\}\right]_0^{+\infty} = 1.
\end{aligned}$$

A igualdade assinalada com $(*)$ deve-se à chamada substituição em coordenadas polares:

$$\left\{\begin{array}{lcl} x & = & r\cos(\theta) \\ y & = & r\sin(\theta) \end{array}\right. ,$$

com $\theta \in [0, 2\pi]$ e $r \geq 0$. Observe que o jacobiano associado a esta substituição é o determinante da seguinte matriz

$$J = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -r\sin(\theta) & r\cos(\theta) \end{bmatrix},$$

i.e.,

$$det(J) = r\cos^2(\theta) + r\sin^2(\theta) = r(\cos^2(\theta) + \sin^2(\theta)) = r.$$

Voltemos agora ao cálculo de $\int_{-\infty}^{+\infty} f(x)dx$.

$$\begin{aligned}
\int_{-\infty}^{+\infty} f(x)dx &= \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} dx \\
&\overset{(**)}{=} \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}y^2\right\} \sigma dy \\
&= I \\
&= 1.
\end{aligned}$$

$(**)$substituição: $y = \frac{x-\mu}{\sigma}$.

Algumas propriedades da lei Normal:

Sejam $\mu \in \mathbb{R}$ e $\sigma \in \mathbb{R}^+$.

1) Se $X \sim N(\mu, \sigma^2)$ então $\frac{X-\mu}{\sigma} \sim N(0,1)$.

2) Se $Z \sim N(0,1)$ então $\sigma Z + \mu \sim N(\mu, \sigma^2)$.

3) A lei $N(0, \sigma^2)$ é simétrica relativamente à origem, i.e., se $X \sim N(0, \sigma^2)$ então

$$P(X \in [a,b]) = P(X \in [-b,-a]), \forall a,\, b \in \mathbb{R},\, a < b.$$

4) A lei $N(\mu, \sigma^2)$ é simétrica relativamente a μ, i.e., se $X \sim N(\mu, \sigma^2)$ então

$$P(X \in [\mu+a, \mu+b]) = P(X \in [\mu-b, \mu-a]), \forall a,\, b \in \mathbb{R},\, a < b.$$

[Demonstração] (T.P.C.)

Observe que as propriedades 3) e 4) são, naturalmente, consequências da simetria, relativamente ao parâmetro μ, das respetivas funções densidade de probabilidade.

Observação: A função de distribuição da lei $N(\mu, \sigma^2)$, dada por

$$F(c) = \int_{-\infty}^{c} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} dx,$$

não tem uma expressão analítica.

No entanto, para cada $c \in \mathbb{R}$, é possível obter uma aproximação de $F(c)$ recorrendo a métodos numéricos. O resultado dessa aproximação está disponível na generalidade dos *software* de matemática/probabilidades e estatística.

Em particular, no R, para obter um valor aproximado de $F(c)$ temos que executar

$$\texttt{pnorm}(c,\ \mu,\ \sqrt{\sigma^2}).$$

5. Variáveis aleatórias reais com lei mista

Nesta secção, será feita uma muito breve caracterização das leis de probabilidade *mistas* ou *de mistura*: isto é, leis de probabilidade sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ que não são nem discretas nem difusas. Veremos também um exemplo prático de utilização de uma tal lei.

Teorema

Seja H uma lei de probabilidade sobre $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ nem discreta nem difusa. Então existem $\alpha \in (0, 1)$, uma lei de probabilidade discreta, H_1, e uma lei de probabilidade difusa, H_2, tais que

$$H = \alpha H_1 + (1 - \alpha) H_2.$$

Além disso, esta decomposição de H é única.

[Demonstração] Ver Lopes & Gonçalves

Exemplo: Seja X a v.a.r. que representa a quantidade de chuva (em metros cúbicos) que cai diariamente numa certa região do país. Sabe-se que, num qualquer dia escolhido ao acaso, a probabilidade de não chover é $1/3$, i.e., $P(X=0)=1/3$. Sabe-se também que, quando chove, a quantidade de chuva é uma v.a.r. absolutamente contínua que segue a lei $Exp(1)$, i.e.,

$$P(X \leq c | X > 0) = 1 - e^{-c},\ c > 0.$$

Observe que a função de distribuição de X é dada por:

$$F(c) = P(X \leq c) = \begin{cases} 0, & c < 0 \\ \frac{1}{3}, & c = 0 \\ \frac{1}{3} + \frac{2}{3}(1 - e^{-c}), & c > 0 \end{cases} = \begin{cases} 0, & c < 0 \\ \frac{1}{3} + \frac{2}{3}(1 - e^{-c}), & c \geq 0 \end{cases},$$

que pode ser escrita na forma

$$F(c) = \alpha F_1(c) + (1-\alpha) F_2(c),$$

com $\alpha = \frac{1}{3}$,

$$F_1(c) = \begin{cases} 0, & c < 0 \\ 1, & c \geq 0 \end{cases} \quad \text{e} \quad F_2(c) = \begin{cases} 0, & c < 0 \\ 1 - e^{-c}, & c \geq 0 \end{cases}.$$

Repare que F_1 é função de distribuição de uma lei de probabilidade discreta e F_2 é função de distribuição de uma lei de probabilidade absolutamente contínua (e, consequentemente, difusa). De facto, F_1 corresponde à função de distribuição de uma v.a.r. discreta, cujo contradomínio é o conjunto $\{0\}$ (é uma v.a.r. quase certa), e F_2 corresponde à função de distribuição de uma v.a.r. absolutamente contínua com lei $Exp(1)$.

6. Parâmetros de localização e dispersão

Os parâmetros de localização e dispersão não caracterizam as leis de probabilidade das respetivas v.a.r.'s, como acontece, por exemplo, com a função de distribuição ou com a função de probabilidade (no caso discreto) ou com a função densidade de probabilidade (no caso absolutamente contínuo).

No entanto, estes parâmetros fornecem informação importante sobre algumas características da variável como, por exemplo: a localização na recta real, a dispersão relativamente a algum valor central (em geral a média), a concentração em determinadas partes de $\mathbb{R}$, etc.

6.1 Esperança matemática

Definição

Seja X uma v.a.r. definida sobre o espaço de probabilidade $(\Omega, \mathcal{A}, P)$.

i) Se X é discreta de contradomínio $C_X = \{x_1, x_2, \ldots, x_n, \ldots\}$ e se

$$\sum_{x_i \in C_X} |x_i| P(X = x_i) < +\infty,$$

define-se a esperança matemática (ou valor médio) de X, usualmente denotada por $E[X]$, como sendo

$$E[X] = \sum_{x_i \in C_X} x_i P(X = x_i).$$

Se

$$\sum_{x_i \in C_X} |x_i| P(X = x_i) = +\infty$$

diz-se que X não tem esperança matemática (ou que $E[X]$ não existe).

6.1 Esperança matemática

Definição

ii) Se X é absolutamente contínua com função densidade de probabilidade f e se

$$\int_{-\infty}^{+\infty} |x| f(x) dx < +\infty,$$

define-se a esperança matemática (ou valor médio) de X, usualmente denotada por $E[X]$, como sendo

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

Se

$$\int_{-\infty}^{+\infty} |x| f(x) dx = +\infty$$

diz-se que X não tem esperança matemática (ou que $E[X]$ não existe).

Observações:

1) Se X é uma v.a.r. discreta de contradomínio finito então X tem esperança matemática.
Mas se X é discreta com contradomínio infinito numerável então $E[X]$ pode não existir.
Exemplo: Considere a v.a.r. X, com $C_X = \mathbb{Z} \setminus \{0\}$ e tal que

$$P(X = n) = P(X = -n) = \frac{1}{2n(n+1)}, n \in \mathbb{N}.$$

Note que a lei de probabilidade de X está bem definida uma vez que $P(X = n) > 0, \forall\, n \in \mathbb{Z} \setminus \{0\}$, e

$$\sum_{n \in \mathbb{Z} \setminus \{0\}} P(X = n) = 2\sum_{n=1}^{\infty} \frac{1}{2n(n+1)} = \sum_{n=1}^{\infty} \frac{1}{n} - \frac{1}{n+1} = 1.$$

X não tem esperança matemática pois

$$\sum_{n \in \mathbb{Z} \setminus \{0\}} |n| P(X = n) = 2\sum_{n=1}^{+\infty} n P(X = n) = 2\sum_{n=1}^{+\infty} \frac{n}{2n(n+1)} = \sum_{n=1}^{+\infty} \frac{1}{n+1} = +\infty.$$

2) Se X é uma v.a.r. absolutamente contínua, $E[X]$ também pode não existir.
Exemplo: Considere uma v.a.r. X, absolutamente contínua, com função densidade de probabilidade

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2},\ x \in \mathbb{R},$$

conhecida por densidade de Cauchy.
$E[X]$ não existe pois

$$\int_{-\infty}^{+\infty} |x| f(x) dx = 2\int_{0}^{+\infty} x\frac{1}{\pi}\frac{1}{1+x^2} dx \underset{(1)}{=} \frac{1}{\pi}\left[\log(1+x^2)\right]_0^{+\infty} = +\infty.$$

Observação: (1) a função $\log$ utilizada é a função logaritmo neperiano.

Exemplos/Exercícios: Mostre que:

i) se $X \sim Bin(n,p)$, com $n \in \mathbb{N}$ e $p \in]0,1[$, então $E[X]$ existe e
$$E[X] = np;$$

ii) se $X \sim Poisson(\lambda)$, com $\lambda \in \mathbb{R}^+$, então $E[X]$ existe e
$$E[X] = \lambda;$$

iii) se $X \sim N(0,1)$ então $E[X]$ existe e
$$E[X] = 0.$$

6.1 Esperança matemática

Vejamos agora a definição de esperança matemática de uma função de uma v.a.r. (sendo a definição anterior um caso particular desta).

Definição

Sejam X uma v.a.r. e $\phi : D \subseteq \mathbb{R} \to \mathbb{R}$ tal que $\phi(X)$ é uma v.a.r..

i) Se X é discreta, de contradomínio $C_X = \{x_1, x_2, \ldots, x_n, \ldots\}$, e se

$$\sum_{x_i \in C_X} |\phi(x_i)| P(X = x_i) < +\infty,$$

diz-se que $E[\phi(X)]$ existe e

$$E[\phi(X)] = \sum_{x_i \in C_X} \phi(x_i) P(X = x_i).$$

Se

$$\sum_{x_i \in C_X} |\phi(x_i)| P(X = x_i) = +\infty$$

diz-se que $E[\phi(X)]$ não existe.

Definição

ii) Se X é absolutamente contínua, com função densidade de probabilidade f, e se

$$\int_{-\infty}^{+\infty} |\phi(x)| f(x) dx < +\infty,$$

diz-se que $E[\phi(X)]$ existe e

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

Se

$$\int_{-\infty}^{+\infty} |\phi(x)| f(x) dx = +\infty$$

diz-se que $E[\phi(X)]$ não existe.

Observação: Podemos concluir que $E[X]$ existe sse $E[|X|]$ existe.

Exemplo/Exercício: Sejam $X \sim Exp(1)$ uma v.a.r. e $\phi : \mathbb{R} \to \mathbb{R}$ a função dada por

$$\phi(x) = \begin{cases} 1 - e^{-x} & se \quad x \geq 0 \\ 0 & se \quad x < 0 \end{cases}.$$

Mostre que a v.a.r. $Y = \phi(X)$ tem esperança matemática e determine então $E[Y]$.

Obs.: Recorde que se X é um v.a.r. e $\phi : D \subseteq \mathbb{R} \to \mathbb{R}$ é uma função contínua, então $\phi(X)$ também é uma v.a.r..

Propriedades da esperança matemática:

I) Sejam X uma v.a.r. e $\phi : D \subseteq \mathbb{R} \to \mathbb{R}$ uma função tal que $E[\phi(X)]$ existe. Então

$$E[a\phi(X) + b] = aE[\phi(X)] + b,$$

quaisquer que sejam $a,\ b \in \mathbb{R}$.
Observe que, como consequência deste resultado, tem-se que

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

quando $E[X]$ existe.

[Demonstração] Vamos fazer a demonstração apenas para o caso discreto. Caso em que X é v.a.r. absolutamente contínua - TPC.

6.1 Esperança matemática

Propriedades da esperança matemática (continuação):

Seja então X v.a.r. discreta de contradomínio $C_X = \{x_1, x_2, \ldots, x_n, \ldots\}$.
Então:

$$\begin{aligned}
E[|a\phi(X)+b|] &= \sum_{x_i \in C_X} |a\phi(x_i)+b| P(X=x_i) \\
&\leq \sum_{x_i \in C_X} |a\phi(x_i)| P(X=x_i) + \sum_{x_i \in C_X} |b| P(X=x_i) \\
&= |a| E[|\phi(X)|] + |b| < +\infty
\end{aligned}$$

Concluimos assim que $E[a\phi(X)+b]$ existe. Vamos agora calculá-la.

$$\begin{aligned}
E[a\phi(X)+b] &= \sum_{x_i \in C_X} [a\phi(x_i)+b] P(X=x_i) \\
&= \sum_{x_i \in C_X} a\phi(x_i) P(X=x_i) + \sum_{x_i \in C_X} bP(X=x_i) \\
&= aE[\phi(X)] + b
\end{aligned}$$

(c.q.d)

Propriedades da esperança matemática (continuação):

II) Sejam X uma v.a.r. e $\phi : D \subseteq \mathbb{R} \to \mathbb{R}$ uma função tal que $E[\phi(X)]$ existe. Então
$$|E[\phi(X)]| \leq E[|\phi(X)|].$$

[Demonstração] Será feita apenas para o caso discreto. Caso em que X é v.a.r. absolutamente contínua - TPC.

$$\begin{aligned}
|E[\phi(X)]| &= \Big| \sum_{x_i \in C_X} \phi(x_i) P(X = x_i) \Big| \\
&\leq \sum_{x_i \in C_X} |\phi(x_i)| P(X = x_i) \\
&= E[|\phi(X)|]
\end{aligned}$$

(c.q.d)

Propriedades da esperança matemática (continuação):

III) Sejam $X_1, X_2, \ldots X_n$ v.a.r.'s definidas sobre o mesmo espaço de probabilidade, todas admitindo esperança matemática. Então

$$E[a_1X_1 + a_2X_2 + \ldots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \ldots + a_nE[X_n]$$

quaisquer que sejam as constantes reais $a_1, a_2, \ldots, a_n$.

[Demonstração] - Ver Lopes $\&$ Gonçalves

Exemplo/Exercício: Mostre que se $X \sim N(\mu, \sigma^2)$ então $E[X] = \mu$.

6.2 Momentos, variância e desvio-padrão

Os momentos de uma v.a.r. são definidos à custa da esperança matemática de certas funções da variável em causa. Tais funções são contínuas e têm a particularidade de a sua composição com uma v.a.r. discreta (respectivamente, contínua) resultar ainda numa v.a.r. discreta (respectivamente, contínua).

Definição [Momento de ordem k de uma v.a.r.]

Sejam X uma v.a.r. e $k \in \mathbb{N}$.
Se $E[X^k]$ existe, chamamos momento de ordem k ao valor de $E[X^k]$.
Caso $E[X^k]$ não exista diz-se que X não tem momento de ordem k.

6.2 Momentos, variância e desvio-padrão

Propriedades dos momentos: Seja X uma qualquer v.a.r..

I) Se $E[X^k]$ existe então $E[X^n]$ existe, para todo o $n \leq k$, $n,\ k \in \mathbb{N}$.
[Demonstração] Ver Lopes $\&$ Gonçalves

II) Seja $k \in \mathbb{N}$. $E[X^k]$ existe sse, para todo o $a \in \mathbb{R}$, $E[(X-a)^k]$ existe.
[Demonstração]
$(\Leftarrow)$ Trivial. Basta fazer $a = 0$.
$(\Rightarrow)$ Usando o Binómio de Newton, tem-se que: para todo o $a \in \mathbb{R}$,

$$(X-a)^k = \sum_{n=0}^{k} \binom{k}{n} X^n a^{k-n} (-1)^{k-n}.$$

Logo,

$$|(X-a)^k| \leq \sum_{n=0}^{k} \binom{k}{n} |X^n|\, |a^{k-n}|.$$

Por hipótese, $E[X^k]$ existe. Então, pela propriedade anterior, $E[|X^n|]$ existe, para todo o $0 \leq n \leq k$. Logo $E[(X-a)^k]$ existe porque

$$E[|(X-a)^k|] \leq \sum_{n=0}^{k} \binom{k}{n} E[|X^n|]\, |a^{k-n}| < +\infty$$

(c.q.d.)

Definição [Momento centrado de ordem k]

Seja X uma v.a.r. tal que $E[X]$ existe e seja $k \in \mathbb{N}$. Chama-se momento centrado de ordem k, denota-se por μ_k, a

$$\mu_k = E[(X - E[X])^k]$$

quando μ_k existe.

Nota: $\mu_1 = E[X - E[X]] = E[X] - E[E[X]] = E[X] - E[X] = 0.$

6.2 Momentos, variância e desvio-padrão

Definição [Variância e desvio-padrão]

Seja X uma v.a.r.. Chama-se variância de X, denota-se por $Var[X]$, ao momento centrado de ordem 2 quando existe, i.e.,

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

Chama-se desvio-padrão de X, denota-se por σ_X, a

$$\sigma_X = \sqrt{Var[X]}.$$

$Var[X]$ e σ_X são parâmetros utilizados para medir a dispersão dos valores de X relativamente ao parâmetro de localização $E[X]$.

Observação: Qualquer que seja a natureza da v.a.r. X, tem-se

$$Var[X] \geq 0.$$

6.2 Momentos, variância e desvio-padrão

Propriedades da variância: Seja X uma v.a.r..

I) $E[X^2]$ existe sse $E[X]$ e $Var[X]$ existem.
[Demonstração] TPC [Ver Lopes & Gonçalves]

II) $Var[X] = 0$ sse X é uma v.a.r. quase certa, i.e., existe $a \in \mathbb{R}$ tal que

$$P(X = a) = 1.$$

[Demonstração] TPC [Ver Lopes & Gonçalves]

III) Se $E[X^2]$ existe então $Var[X] = E[X^2] - (E[X])^2$.
[Demonstração] É consequência directa da linearidade da esperança matemática.

$$\begin{aligned} Var[X] &= E[(X - E[X])^2] = E[X^2 - 2XE[X] + (E[X])^2] \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 \\ &= E[X^2] - (E[X])^2 \end{aligned}$$

(c.q.d.)

Propriedades da variância (continuação):

IV) Se X admite momento de ordem 2 então

$$Var[aX+b] = a^2 Var[X], \;\; \forall a,\, b \in \mathbb{R}.$$

[Demonstração] Comecemos por mostrar que $Var[aX+b]$ existe. Para isso, basta mostrar que $aX+b$ tem momento de 2ª ordem. Usando a linearidade da esperança matemática e o facto de $E[X^2]$ existir, tem-se que

$$E[|(aX+b)^2|] = E[|a^2X^2+2abX+b^2|] \leq a^2E[|X|^2] + 2|ab|E[|X|] + b^2 < +\infty,$$

e concluimos assim que $Var[aX+b]$ existe. Vamos agora calculá-la.

$$\begin{aligned} Var[aX+b] &= E[(aX+b-E[aX+b])^2] = E[(aX+b-aE[X]-b)^2] \\ &= E[(a(X-E[X]))^2) = E[a^2(X-E[X])^2] \\ &= a^2E[(X-E[X])^2] \\ &= a^2Var[X] \end{aligned}$$

(c.q.d.)

Exemplos/Exercícios: Mostre que

i) se $X \sim Poisson(\lambda)$, com $\lambda \in \mathbb{R}^+$, então $Var[X] = \lambda$;

ii) se $X \sim N(\mu, \sigma^2)$, com $\mu \in \mathbb{R}$ e $\sigma \in \mathbb{R}^+$, então $Var[X] = \sigma^2$.

Resolução de ii): Como é usual com a lei normal, vamos começar por provar que se $Y \sim N(0,1)$ então $Var[Y] = 1$ e, de seguida, estender o resultado à lei $N(\mu, \sigma^2)$, com quaisquer parâmetros μ e σ^2. Para este caso particular, faremos uso da propriedade IV) da variância (ver pág. anterior).

Vamos começar então por provar que $E[Y^2]$ existe, i.e., que

$$\int_{-\infty}^{+\infty} |x^2| f(x) dx < +\infty,$$

com f a função densidade de probabilidade da lei $N(0,1)$. Ora:

Exemplos/Exercícios: (continuação - resolução de ii))

$$\begin{aligned}
\int_{-\infty}^{+\infty} |x^2| f(x) dx &= 2\int_0^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx \\
&= \frac{2}{\sqrt{2\pi}} \left\{ \left[-x \exp\left\{-\frac{1}{2}x^2\right\} \right]_0^{+\infty} - \int_0^{+\infty} - \exp\left\{-\frac{1}{2}x^2\right\} dx \right\} \\
&= 2\int_0^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx = 2 \times \frac{1}{2} = 1 < +\infty.
\end{aligned}$$

Concluímos assim que $E[Y^2]$ existe e, portanto, que $Var[Y]$ também existe e, adicionalmente, que é dada por

$$Var[Y] = E[Y^2] - (E[Y])^2 = 1 - 0^2 = 1.$$

Recorde que já foi provado atrás que $E[Y] = 0$ e que $E[Y^2] = E[|Y^2|] = 1$. Resta agora conjugar as propriedades da lei Normal (enunciadas na Secção 4.3) e a propriedade IV) da variância para concluir que, se $X \sim N(\mu, \sigma^2)$, então $Var[X]$ existe e $Var[X] = \sigma^2$. (c.q.d.)

6.3 Quantis

Para terminar, vamos ainda estudar os *quantis* de uma v.a.r. que, sendo medidas de localização, podem também ser usados para quantificar a sua concentração em certas partes de $\mathbb{R}$.

Definição [Quantil de ordem p]

Sejam X uma v.a.r., com função de distribuição F, e $p \in]0,1[$. Define-se quantil de ordem p, denota-se por χ_p, a

$$\chi_p = \inf\{c \in \mathbb{R} : F(c) \geq p\}.$$

Observações:

1) Se X é uma v.a.r. absolutamente contínua e $F^{-1}(p)$ existe então

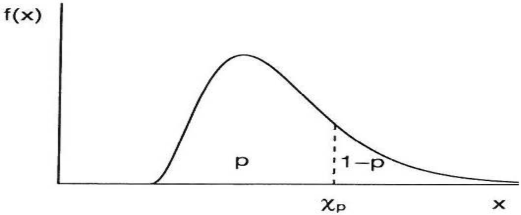
$$\chi_p = F^{-1}(p).$$

2) Em geral, o quantil de ordem p pode ser visto como o valor real que divide a lei da v.a.r. X em duas partes: uma de peso máximo igual a p (a parte inferior a χ_p) e outra de peso máximo igual $(1-p)$ (a parte superior a χ_p).

Na situação 1) (da observação anterior) a divisão é exacta, i.e.,

$$P(X < \chi_p) = p \;\text{ e }\; P(X > \chi_p) = 1 - p.$$

De facto, neste caso, χ_p é o valor que, no gráfico da função densidade de probabilidade, tem à sua esquerda área igual a p e à sua direita área igual a $(1-p)$ (conforme figura abaixo).



6.3 Quantis

3) Quando $p = 0.5$, o quantil é também designado de *mediana*. Grosso modo, a mediana divide a lei da v.a.r. em 2 partes de igual peso: 0.5 cada uma.

4) Os quantis $\chi_{1/4}$, $\chi_{2/4}$ e $\chi_{3/4}$ são designados de *quartis*, sendo $\chi_{1/4}$ o *1º quartil*, $\chi_{2/4}$ o *2º quartil* (que coincide com a mediana) e $\chi_{3/4}$ o *3º quartil*.
   Grosso modo, os quartis dividem a lei da v.a.r. em 4 partes de igual peso: 0.25 cada uma.

5) Fazendo $p = i/10$, com $i \in \{1, 2, \ldots, 9\}$, obtemos os *decis*. Grosso modo, os decis dividem a lei da v.a.r. em 10 partes de igual peso: 0.1 cada uma.

6) Fazendo $p = i/100$, com $i \in \{1, 2, \ldots, 99\}$, obtemos os *percentis*. Grosso modo, os percentis dividem a lei da v.a.r. em 100 partes de igual peso: 0.01 cada uma.