Capítulo II: Variáveis Aleatórias Reais

Probabilidades e Aplicações

Licenciatura em Matemática
Licenciatura em Ciências da Computação
Universidade do Minho
Ano Letivo 2025/2026

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

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

em que Ω é o espaço amostral da experiência aleatória.

Exemplo:

Suponhamos que estamos interessados em estudar o número de faces par obtidas em 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 k-dimensional

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

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** que são 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 $X^{-1}(\{1,2\}) = X^{-1}(\,[1,+\infty[\,)$.

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 $\mathcal{F}$, a σ-álgebra sobre o espaço amostral Ω, de modo a que a sua probabilidade esteja bem definida.

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

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

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 $E \in \mathcal{B}(\mathbb{R})$.

1. Definição e Exemplos. Considerações gerais.

Definição [Variável aleatória real (v.a.r.)]

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

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

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

Teorema

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

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

[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

$$X(\omega) = \begin{cases} 1 & se \quad \omega = cara \\ 0 & se \quad \omega = coroa \end{cases}$$

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 $(\Omega, \mathcal{F}, P)$ o espaço de probabilidade associado a uma experiência aleatória e B um acontecimento (i.e., $B \in \mathcal{F}$). Vamos provar que a função indicatriz do conjunto B, i.e., a função $\mathbf{1}_B : \Omega \to \mathbb{R}$ definida por

$$\mathbf{1}_B(\omega) = \begin{cases} 1, & \text{se} \quad \omega \in B \\ 0 & \text{se} \quad \omega \in \overline{B} \end{cases} ,$$

é uma v.a.r..

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

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

Sendo $\mathcal{F}$ uma σ-álgebra sobre Ω, tem-se que $\emptyset \in \mathcal{F}$, $\overline{B} \in \mathcal{F}$ (porque $B \in \mathcal{F}$) e $\Omega \in \mathcal{F}$, concluindo-se assim que, qualquer que seja $c \in \mathbb{R}$, $\mathbf{1}_B^{-1}(\,]-\infty, c]\,) \in \mathcal{F}$, e ficando assim provado que $\mathbf{1}_B$ é uma v.a.r..

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

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

Definição [σ- álgebra gerada por uma v.a.r.]

Sejam $(\Omega, \mathcal{F}, P)$ 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

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

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

$$\begin{array}{ccccc} & & \Omega & \xrightarrow{X} & \mathbb{R} \\ & & & & \\ [0,1] & \xLeftarrow{P} & X^{-1}(\mathcal{B}(\mathbb{R})) & \xLeftarrow{X^{-1}} & \mathcal{B}(\mathbb{R}) \end{array}$$

Definição [Lei de probabilidade de uma v.a.r.]

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 exercício da Folha Prática 4 para a demonstração].

2. Lei de probabilidade e função de distribuição.

Definição [Função de distribuição de uma v.a.r.]

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. Já sabemos que o espaço de probabilidade é $(\Omega, \mathcal{P}(\Omega), P)$, com $\Omega = \{cara,\ coroa\}$ e P a medida de probabilidade 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

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

2. Lei de probabilidade e função de distribuição.

Propriedades de uma função de distribuição:

Sejam $(\Omega, \mathcal{F}, P)$ 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)] Exercícios da Folha Prática 4 .

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

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 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (ver Propriedade VII) de uma medida de probablidade).

Pela unicidade do limite, concluimos então que

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

c.q.d

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

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

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 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (ver Prop. VI) de medida de probablidade). Concluimos então que

$$P_X(\{x_0\}) = 0 \Leftrightarrow \lim_{c\to x_0^-} F(c) = F(x_0).$$

c.q.d.

No que se segue, $(\Omega, \mathcal{F}, P)$ 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 [v.a.r. discreta; lei de probabilidade discreta; contradomínio]

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.
Se X é uma v.a.r. 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, $(\Omega, \mathcal{F}, P)$ 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 [v.a.r. discreta; lei de probabilidade discreta; contradomínio]

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.
Se X é uma v.a.r. 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 X é uma v.a.r. discreta (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

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/P_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.

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

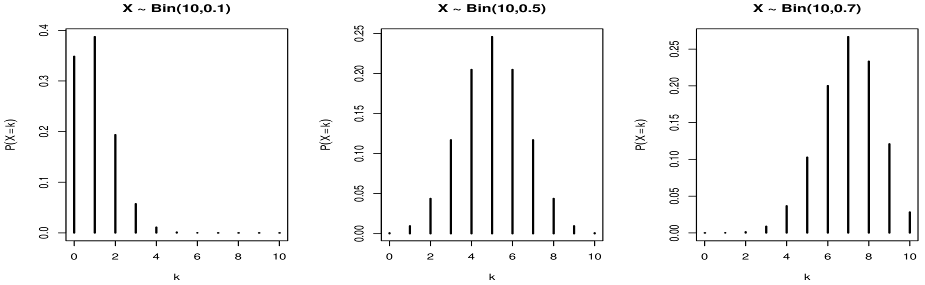
$$f(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & se \quad k \in \{0, 1, \ldots, n\} \\ 0 & se \quad c.c. \end{cases} .$$

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”.

3.2 Leis de Probabilidade Discretas Mais Conhecidas

A título de exemplo, seguem-se três gráficos com a representação da função de probabilidade de algumas leis binomiais, todas com $n = 10$ e p igual a 0.1, 0.5 e 0.7, respetivamente. É nítido o efeito que a variação de p provoca na assimetria da distribuição.



Nota: Atente também no efeito que a variação de p tem na localização e na variabilidade. Em particular, note que, para o mesmo valor de n, a variabilidade máxima é obtida quando $p = 0.5$.

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

Sejam $(\Omega, \mathcal{F}, P)$ 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 $0 \leq M \leq R$ possuem um certo atributo, A, e os restantes $R-M$ elementos não têm este atributo A.

Considere agora 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

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

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

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

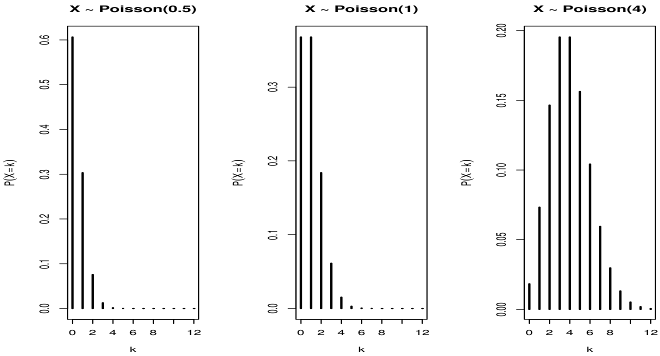
$$f(k)=\begin{cases}\frac{\lambda^k}{k!}e^{-\lambda} & \text{se} \quad k\in\mathbb{N}_0\\ 0 & \text{se} \quad \text{c.c.}\end{cases}.$$

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 $n \times p$.

A título de exemplo, seguem-se três gráficos com a representação da função de probabilidade de algumas leis de de Poisson, com λ igual a 0.5, 1 e 4, respetivamente. É nítido o efeito que a variação de λ tem na localização e na variabilidade.



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., as repetições são independentes) e seja T a v.a.r. 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 $C_T = \{1, 2, \ldots\} = \mathbb{N}$.

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:

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

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 <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:

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

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

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

3.2 Leis de Probabilidade Discretas Mais Conhecidas

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:

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

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

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

$\vdots$

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

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.,

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

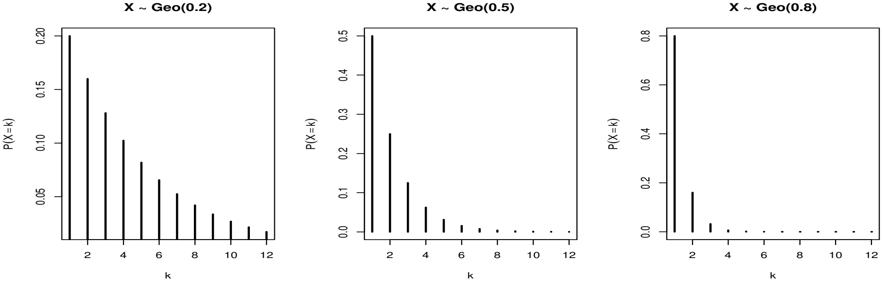
para todo o $k, n \in \mathbb{N}$.

Exemplo/Exercício(TPC): 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.r. relevante para o problema e que tenha distribuição Geométrica.

A título de exemplo, seguem-se três gráficos com a representação da função de probabilidade de algumas leis Geométricas com p igual a 0.2, 0.5 e 0.8, respetivamente. É nítido o efeito que a variação de p tem na localização e na variabilidade.



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 [v.a.r./lei de probabilidade difusa]

Sejam $(\Omega, \mathcal{A}, P)$ 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 lei de probabilidade difusa sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, i.e., se P_X for tal que

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

Notas:

1) 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.

2) Se X é uma v.a.r. discreta então X não é difusa.

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*.

Definição [Função densidade de probabilidade sobre $\mathbb{R}$]

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?

Definição [Função densidade de probabilidade sobre $\mathbb{R}$]

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?

- 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 [v.a.r./lei de probabilidade absolutamente contínua]

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

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

À função f chamamos *função densidade de probabilidade da v.a.r.* X ou *função densidade de probabilidade da lei* P_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

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

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 respetiva 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.,

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

Observação: Deste último teorema, resulta trivialmente que, se X é uma v.a.r. absolutamente contínua, então a respetiva 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

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

4.2 Leis absolutamente contínuas

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.