# Introdução à Teoria da Medida e Probabilidade

Jorge Salazar

Janeiro 2016

# Introdução

A teoria da medida já estava 'no ar' no início do século vinte, a tese de L. Bachelier, "Théorie de la Spéculation," publicada em 1900 é prova. Nela, Bachelier propõe usar o movimento Browniano para modelar os mercados financeiros, mas nessa altura, o movimento Browniano ainda não era um objeto matemático bem definido. É geralmente aceite que a teoria da medida e da integração foi colocada em base firme a partir da tese de H. Lebesgue "Intégrale, longueur, aire," publicada em 1902. Em 1909, E. Borel fez a observação de que os dígitos da expansão binária dos números entre 0 e 1 (hoje conhecidas como funções de Rademacher), eram "independentes." Esta era a primeira vez que a noção probabilística de independência era aplicada a objetos matemáticos bem definidos. A monografia de Borel "Sur les probabilités dénombrables et leurs applications arithmétiques" [2] marca o início da teoria moderna das probabilidades. Pouco mais tarde, Carathéodory apresenta o seu teorema de existência e unicidade da extensão de uma (pre-)medida, ampliando enormemente o campo de aplicabilidade da teoria de Lebesgue. Muitos desenvolvimentos seguem-se durante a primeira metade do século vinte, o livro de Halmos [8] é uno dos primeiros tratados completos sobre a teoria da medida e uma referência histórica.

O livro que o leitor tem nas mãos foi desenvolvido a partir das notas de aulas da disciplina do mesmo nome, dadas ao curso de licenciatura em matemática aplicada, na Universidade de Évora, durante vários anos e na Escola Politécnica do Litoral, em Guayaquil durante a visita sabática em 2014. Nele tentamos combinar, a um nível introdutório, a teoria da medida desde os pontos de vista tanto da análise matemática como da teoria das probabilidades.

Damos agora uma descrição comentada do conteúdo deste livro. No Capítulo 1 definimos a estrutura básica dos espaços mensuráveis e estudamos o processo de geração dos mesmos a partir de estruturas mais simples. Aqui introduzimos os dois exemplos que encontraremos de forma sistemática ao longo de todo o livro, e que são: a) As sucessões de Bernoulli (como paradigma das probabilidades) e b) A reta real (como paradigma da análise matemática). No Capítulo 2 definimos as medidas sobre os espaços mensuráveis e ao mesmo tempo introduzimos a terminologia das probabilidades. Dão-se conceitos e propriedades essenciais para o estudo posterior da integração, no Capítulo 4, como são medida imagem (Secção 2.2) e os teoremas de convergência para conjuntos (Secção 2.3). O problema da existência/construção de medidas (Teorema de Carathéodory) é uma das partes mais técnicas do livro e por esse motivo, mostramos primeiro a sua aplicação aos dois exemplos fundamentais que nos acompanham ao longo da aprendizagem, a medida de Lebesgue em $\mathbb{R}$ e a medida produto sobre as sucessões de Bernoulli. No Capítulo 3, fazemos um compêndio de tudo o que vamos precisar sobre as funções reais. Introduzimos ademais a medida de Lebesgue-Stieltjes e damos vários exemplos de distribuições de probabilidade, encontradas habitualmente em Estatística. O Capítulo 4 é dedicado ao de integral de Lebesgue, suas propriedades fundamentais e os teoremas de convergência. Debruçamo-nos sobre algúns aspetos da convergência de variáveis aleatórias, como a convergência em probabilidades e as leies dos grandes números. Finalizamos este capítulo com um tratamento elementar dos espaços $L^p$. O Capítulo 5 trata da integração com respeito á medida produto, o teorema de Fubini-Tonelli e a sua relação com a independência estatística.

Esperamos que este livro guie o aluno através da aprendizagem desta matéria de fundamental importância tanto na matemática pura como nas suas aplicações. Para aprofundar nestes temas, o leitor pode consultar as obras que compilamos na lista bibliográfica no final do livro. A teoria da integração sobre espaços topológicos gerais tem duas correntes, a partir de medidas Borelianas abstratas, como é feito no livro de W. Rudin [10] e é a linha seguida neste livro. A outra alternativa é considerar medidas mais restritivas como as medidas de Radón, com a vantagem delas interatuar melhor com a estrutura topológica do espaço subjacente, este ponto de vista é apresentado nos volumes III e IV do tratado de Análise [12] e [13] de L. Schwartz, ou no volume II dos “Éléments d'Analyse” [5] de J. Dieudonné. Se queremos uma teoria da integração impecável sobre espaços topológicos gerais, então podemos estudar o livro de Schwartz [11] “Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures”, ou “Intégration” [3] (Capítulo 9) de Bourbaki. Para aplicações em Física e Engenharia, temos outro livro de L. Schwartz [14] “Méthodes mathématiques pour les sciences physiques”. Para tópicos mais especializados de análise, veja-se o livro de Evans e Gariepy [6]. Para os leitores mais inclinados às probabilidades, os livros de W. Feller [7], P. Billingsley [1] e/ou K. L. Chung [4] são uma grande fonte de consulta e inspiração. A monografia de M. Kac [9] é um excelente tratado sobre a noção de independência estatística.

# Notações

$(X, \mathcal{F})$ Espaço mensurável, onde $X$ é um conjunto e $\mathcal{F}$ uma $\sigma-$álgebra sobre $X$.

$(\Omega, \mathcal{F}, \mathrm{P})$ Espaço de probabilidade genérico.

$\mathcal{B}$ Álgebra$-\sigma$ de Borel ou Borelianos de $\mathbb{R}$.

$\mathcal{B}_X$ Álgebra$-\sigma$ de Borel de um conjunto $X$.

$\mathcal{F} \vee \mathcal{G}$ Menor maiorante das álgebras$-\sigma$ $\mathcal{F}$ e $\mathcal{G}$. Coincide com $\sigma[\mathcal{F} \cup \mathcal{G}]$.

$\mathcal{F} \wedge \mathcal{G}$ Maior minorante das álgebras$-\sigma$ $\mathcal{F}$ e $\mathcal{G}$. Coincide com $\mathcal{F} \cap \mathcal{G}$.

$m[\mathcal{S}]$ Família monótona gerada por $\mathcal{S}$.

$X$, $Y$, $Z$ Conjuntos genéricos.

$\mathbb{N}$ Conjunto dos números naturais, i.e. $\mathbb{N} = \{1, 2, ...\}$.

$\mathbb{Q}$ Conjunto dos números racionais, i.e. $\mathbb{Q} = \left\{\frac{p}{q};\ p \in \mathbb{Z},\ q \in \mathbb{N}\right\}$.

$\mathbb{R}$ Conjunto dos números reais.

$\mathbb{Z}$ Conjunto dos números inteiros, i.e. $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$.

$\chi_A$ Função característica de um subconjunto $A \subseteq X$.

$\delta_a$ Medida de Dirac em $a \in X$.

$(X, \mathcal{F}, \mu)$ Espaço de medida sobre o espaço mensurável $(X, \mathcal{F})$.

$\{0, 1\}^{\mathbb{N}}$ Sucessões de Bernoulli.

$\{0, 1\}^n$ Cartesiano de $n$ cópias idénticas do conjunto $\{0, 1\}$.

$\mathcal{B}^n$ Álgebra$-\sigma$ de Borel de $\mathbb{R}^n$.

$\mathcal{F} \otimes \mathcal{G}$ Álgebra$-\sigma$ produto.

$\mathcal{L}(X)$ Funções de $X$ em $\overline{\mathbb{R}}$, $\mathcal{F}/\mathcal{B}-$mensuráveis.

$\mathcal{L}^p(X, \mu)$ Funções cuja potência $p$ é $\mu-$integrável.

$\mathcal{L}^+(X)$ Funções mensuráveis, não negativas.

$\mathcal{L}^1(X,\mu)$ Funções $\mu-$integráveis.

$\mathcal{L}^\infty(X\mu)$ Funções limitadas em $\mu-$quase todo ponto.

$\mathcal{L}^{s,+}(X)$ Funções mensuráveis, simples não negativas.

$\mathcal{L}^s(X)$ Funções mensuráveis simples.

$\mathcal{O}_X$ Família de abertos de um espaço topológico $X$.

$\mathcal{P}(X)$ Conjunto das partes do conjunto $X$, i.e. $\mathcal{P}(X) = \{A; A \subseteq X\}$.

$\mathrm{Card}\,(X)$ Número de elementos do conjunto finito (finito) $X$.

E(X) Média ou esperança matemática da variável aleatória X.

$\mathrm{F_X}$ Função de acumulação de uma variável aleatória $\mathrm{X} : \Omega \to \overline{\mathbb{R}}$.

H Função de Heaviside.

Var(X) Variância da variável aleatória X.

X, Y, Z Variáveis aleatórias.

$\mu^*(A)$ Medida exterior $(= \inf\{\sum_{i=1}^\infty \mu_0(A_i)\,;\ A \subseteq \bigcup_{n=1}^\infty A_i\ \wedge\ \forall\, i \in \mathbb{N},\, A_i \in \mathcal{F}_0\})$.

$\nu_{\mathrm{X}}$ Distribuição de probabilidad ou lei de X $(= \mathrm{X}_\#\mathrm{P})$.

$\overline{\mathbb{R}}$ Compactificação de $\mathbb{R}$ por dois pontos.

$\sigma[f]$ Álgebra$-\sigma$ gerada por $f$.

$\sigma[\mathcal{S}]$ Álgebra$-\sigma$ gerada por $\mathcal{S}$.

$f_\#\mu$ Medida imagem de $\mu$ por $f$.

$L^p(X,\mu)$ Classes de funções cuja potência $p$ é $\mu-$integrável.

$L^1(X,\mu)$ Classes de equivalência das funções $\mu-$integráveis.

$L^\infty(X,\mu)$ Classes de funções limitadas em $\mu-$quase todo ponto.

# Conteúdo


**Introdução** **iii**

**Notações** **v**

**1 Estruturas básicas de famílias de subconjuntos** **1**

1.0.1 Conjunto das partes . . . . . . . . . . . . . . . . . . . . . . . 1

1.1 Álgebras $-\sigma$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.2 Álgebra$-\sigma$ Envolvente . . . . . . . . . . . . . . . . . . . . . . . 7

1.2.1 álgebra$-\sigma$ gerada por uma família de subconjuntos . . . . . 8

1.2.2 Álgebra$-\sigma$ de Borel . . . . . . . . . . . . . . . . . . . . . . . 8

1.2.3 Estrutura reticulada sobre as álgebras$-\sigma$ . . . . . . . . . . . 9

1.3 Sucessões de Bernoulli . . . . . . . . . . . . . . . . . . . . . . . 10

1.3.1 Álgebra$-\sigma$ sobre as sucessões de Bernoulli . . . . . . . . . . 11

1.3.2 Filtração de Bernoulli . . . . . . . . . . . . . . . . . . . . . . 13

1.4 Funções mensuráveis . . . . . . . . . . . . . . . . . . . . . . . . 14

1.4.1 Álgebra$-\sigma$ gerada por uma função . . . . . . . . . . . . . . . 15

**2 Medidas** **21**

2.1 Medida . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21

2.2 Transporte de medidas . . . . . . . . . . . . . . . . . . . . . . . 25

2.2.1 Medida imagem . . . . . . . . . . . . . . . . . . . . . . . . . 25

2.2.2 Medida invariante . . . . . . . . . . . . . . . . . . . . . . . . 27

2.2.3 Medida preimagem . . . . . . . . . . . . . . . . . . . . . . . 27

2.3 Teoremas de convergência para conjuntos . . . . . . . . . . . . . 28

2.4 Medidas a partir de premedidas . . . . . . . . . . . . . . . . . . 31

2.4.1 Premedidas . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31

2.4.2 Medida de Comprimento sobre $(0, 1]$ . . . . . . . . . . . . . 33

2.4.3 Probabilidade produto sobre $\{0, 1\}^{\mathbb{N}}$ . . . . . . . . . . . . . . 37

2.5 Medida exterior . . . . . . . . . . . . . . . . . . . . . . . . . . . 42

2.6 Construção de medidas . . . . . . . . . . . . . . . . . . . . . . . 46

2.7 Famílias monótonas . . . . . . . . . . . . . . . . . . . . . . . . . 50

2.8 Medidas completas . . . . . . . . . . . . . . . . . . . . . . . . . . . 54
2.8.1 Conjuntos de medida nula . . . . . . . . . . . . . . . . . . . 54
2.8.2 A medida de Lebesgue . . . . . . . . . . . . . . . . . . . . . 57

**3 Os Borelianos e as medidas de Borel em $(\mathbb{R}, \mathcal{B})$ 63**
3.1 Reta real estendida . . . . . . . . . . . . . . . . . . . . . . . . . . . 63
3.1.1 Álgebra$-\sigma$ de Borel . . . . . . . . . . . . . . . . . . . . . . 65
3.2 Mais sobre as funções mensuráveis . . . . . . . . . . . . . . . . . . . 68
3.2.1 Mensurabilidade das funções contínuas . . . . . . . . . . . . 68
3.2.2 Caracterização das funções mensuráveis reais . . . . . . . . . 69
3.2.3 Álgebra$-\sigma$ produto . . . . . . . . . . . . . . . . . . . . . . 71
3.2.4 Borelianos de $\mathbb{R}^n$ . . . . . . . . . . . . . . . . . . . . . . . 73
3.3 Medidas de Lebesgue-Stieltjes . . . . . . . . . . . . . . . . . . . . . 74
3.3.1 Variáveis aleatórias . . . . . . . . . . . . . . . . . . . . . . . 74
3.3.2 Construção das medidas de Lebesgue-Stieltjes . . . . . . . . 81
3.3.3 Classificação de medidas Borelianas . . . . . . . . . . . . . . 84
3.3.4 Variáveis aleatórias independentes . . . . . . . . . . . . . . . 88
3.3.5 Exemplos de distribuições de probabilidade . . . . . . . . . . 89

**4 Integral de Lebesgue 101**
4.1 Integração de funções . . . . . . . . . . . . . . . . . . . . . . . . . . 101
4.1.1 Integral de funções não negativas . . . . . . . . . . . . . . . 103
4.1.2 Continuidade absoluta . . . . . . . . . . . . . . . . . . . . . 107
4.1.3 Convergência monótona . . . . . . . . . . . . . . . . . . . . 110
4.1.4 Funções Integráveis . . . . . . . . . . . . . . . . . . . . . . . 113
4.1.5 Convergência dominada . . . . . . . . . . . . . . . . . . . . 120
4.2 Variáveis aleatórias integráveis . . . . . . . . . . . . . . . . . . . . . 122
4.2.1 Média e Variância . . . . . . . . . . . . . . . . . . . . . . . . 123
4.2.2 Lei(es) dos grandes números . . . . . . . . . . . . . . . . . . 125
4.3 Espaços $L^p(X, \mu)$ . . . . . . . . . . . . . . . . . . . . . . . . . . . 131
4.3.1 Estrutura de espaço vetorial normado sobre $L^p(X, \mu)$ . . . . 132

**5 Integração no espaço produto 137**
5.1 Medida produto . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 138
5.1.1 Teorema de Fubini-Tonelli para conjuntos . . . . . . . . . . 140
5.2 Teorema de Fubini-Tonelli . . . . . . . . . . . . . . . . . . . . . . . 145
5.3 Independência e medida produto . . . . . . . . . . . . . . . . . . . 148

# Capítulo 1

# Estruturas básicas de famílias de subconjuntos

Vamos começar por estudar as $\sigma-$álgebras, que são famílias de subconjuntos de um conjunto dado, mas que possuem uma certa estrutura. Lembre-se o que acontece com os conjuntos abertos do plano (ou dum espaço topológico em geral). Os conjuntos abertos têm certa estabilidade como por exemplo a união arbitrária de abertos é um aberto, como também é a interseção finita de conjuntos abertos. Os conjuntos abertos servem para definir o que é uma função contínua. De forma semelhante, uma $\sigma-$álgebra serve para definir a noção de função mensurável que são os "integrandos" da teoria da integração. As $\sigma-$álgebras são famílias de subconjuntos de um espaço ou conjunto de base, sobre as quais vamos definir as medidas, que são os "integradores" desta teoria. As $\sigma-$álgebras representam também a informação disponível, através dos eventos que fazem parte dela, especialmente nas filtrações de $\sigma-$álgebras associadas com processos estocásticos. Mas antes de nada, damos a seguinte definição.

### 1.0.1 Conjunto das partes

**Definição 1.1.** Seja $X$ um conjunto qualquer. ***O conjunto das partes de*** $X$ é a família de todos os subconjuntos de $X$ e denota-se $\mathcal{P}(X)$. Ou seja,

$$\mathcal{P}(X) = \{A; A \subseteq X\}.$$

### Exemplos

(a) Se $X = \emptyset$, então $\mathcal{P}(X) = \{\emptyset\}$. Este exemplo é simples demais e não é muito didático. Vamos ao segundo exemplo.

(b) Se $X$ contiver um único elemento, digamos $X = \{\emptyset\}$, então $\mathcal{P}(X) = \{\emptyset, X\}$.

(c) Sendo $X = \{0, 1\}$, o conjunto das partes é $\mathcal{P}(X) = \{\emptyset, \{0\}, \{1\}, X\}$.

### Conjuntos finitos

**Notação.** Sendo $X$ um conjunto finito, o **cardinal** de $X$ é o número de elementos de $X$ e denota-se $\mathrm{Card}\,(X)$.

**Proposição 1.1.** Se $X$ é finito, então $\mathcal{P}(X)$ também é finito e $\mathrm{Card}\,(\mathcal{P}(X)) = 2^n$ (onde $n$ é o número de elementos de $X$).

***Demonstração.*** Basta observar que se escrevermos o conjunto $X$ por extenso, digamos que $X = \{x_1, ..., x_n\}$, e mantermos a ordem fixa, podemos associar com cada $A \subset X$, uma sequência ordenada de 0's e 1's, de comprimento $n$. Sendo que 0 no lugar $i$ significa que o elemento $x_i$ não pertence a $A$. Igualmente, 1 no lugar $i$ significa $x_i \in A$. É facil ver que o número de subconjuntos de $X$ é igual ao número de sequências ordenada de 0's e 1's, de comprimento $n$, ou seja $2^n$. :-)

***Observação.*** Noutras palabras, a Proposição 1.1 diz que existe uma relação biunívoca entre o conjunto das partes $\mathcal{P}(X)$ e o conjunto $\{0, 1\}^n$ (sendo $n$ o número de elementos de $X$). Por exemplo,

$$\mathcal{P}(\{0, 1\}) \sim \{(0, 0), (1, 0), (0, 1), (1, 1)\}.$$

Explicitamente,

$$\begin{array}{ccc} \emptyset & \longleftrightarrow & (0,0) \\ \{0\} & \longleftrightarrow & (1,0) \\ \{1\} & \longleftrightarrow & (0,1) \\ \{0,1\} & \longleftrightarrow & (1,1) \end{array}$$

### Conjuntos infinitos

No caso em que $X$ é finito, o conjunto $\mathcal{P}(X)$ é muito maior (mais numeroso). Mas o que é que podemos dizer se $X$ for infinito, para além de $\mathcal{P}(X)$ também ser infinito?

Quando comparamos dois conjuntos infinitos, acontece muitas vezes que uma parte é *equipotente* ao todo, quer dizer, não é inusual existir uma bijeção entre um conjunto e um subconjunto estrito.

Por exemplo, com respeito a $\mathbb{N}$, podemos dizer que um de cada dois números naturais é par e portanto, em certo sentido, a "mitade" dos números naturais são pares. Mas por outro lado, a aplicação $n \to 2n$ é uma bijeção entre os números naturais e os números pares. Desde este ponto de vista, há tantos números pares como números naturais.

Independentemente de $X$ ser finito ou não, $\mathcal{P}(X)$ é "muito maior" do que $X$ num sentido muito preciso, como vamos ver de seguida.

**Proposição 1.2.** Sendo $X$ um conjunto qualquer, nenhuma função de $X$ em $\mathcal{P}(X)$ é sobrejetiva.

***Demonstração.*** Fixamos uma aplicação qualquer $\Psi : X \to \mathcal{P}(X)$ e construimos, a partir dela, um subconjunto de $X$ que não é "atingido" por $\Psi$. Quer dizer, procuramos um subconjun to $A \subset X$, tal $A \neq \Psi(x)$, para todo $x \in X$. Desta forma teremos mostrado que esta função $\Psi$, em particular, não é sobrejetiva no conjunto de chegada $\mathcal{P}(X)$. Mas sendo $\Psi$ arbitrária, vemos que não existe nenhuma aplicação $\Psi : X \to \mathcal{P}(X)$ que seja sobrejetiva. Agora só falta-nos construir o conjunto $A$, apropriado para este $\Psi$ particular. Seja

$$A = \{x \in X; x \notin \Psi(x)\}.$$

Vamos mostrar que este conjunto não pertence à imagem de $\Psi$. De facto, observe que para todo $x \in X$, necessariamente $x \in \Psi(x)$ ou $x \in A$. Por outro lado, é claro que $x \notin \Psi(x) \cap A$. Portanto, para todo $x \in X$, $A \neq \Psi(x)$. :-)

**Definição 1.2.** Dizemos que um conjunto $X$ é ***numerável*** se existe uma aplicação bijetiva $\psi : \mathbb{N} \to X$.

Note que os conjuntos finitos não são numeráveis, segundo esta definição.

**Nota 1.1.** Pela Proposição 1.2, $\mathcal{P}(\mathbb{N})$ não é numerável.

**Função característica**

**Definição 1.3.** Seja $X$ um conjunto e $A \subseteq X$. A função

$$\chi_A(x) = \begin{cases} 1, & \text{se } x \in A \\ 0, & \text{se } x \notin A, \end{cases} \tag{1.1}$$

definida sobre $X$, chama-se ***função característica*** de $A$.

Outra forma de representar os subconjuntos de $X$ (independentemente de $X$ ser finito ou não) é através do conjunto de todas as funções de $X$ em $\{0, 1\}$, que denotaremos $F(X; \{0, 1\})$ ou $\{0, 1\}^X$. Como o afirma a seguinte proposição.

**Proposição 1.3.** Qualquer que seja o conjunto $X$, temos

$$\mathcal{P}(X) \sim \{0, 1\}^X .$$

***Demonstração.*** De facto, se $f \in F(X; \{0, 1\})$ $(= \{0, 1\}^X)$, então $f^{-1}(1) \subset X$. Inversamente, se $A \subset X$, então $\chi_A \in F(X; \{0, 1\})$.

## 1.1 Álgebras $-\sigma$

Nesta secção definiremos as álgebras$-\sigma$, que são as famílias dos conjuntos *mensuráveis*, também chamados *eventos* em probabilidades.

**Definição 1.4.** Uma família $\mathcal{F} \subseteq \mathcal{P}(X)$ é uma ***álgebra*** sobre $X$, se se verificarem as condições seguintes.

1. $X \in \mathcal{F}$.
2. Se $A \in \mathcal{F}$, então $X \setminus A \in \mathcal{F}$.
3. Se $A, B \in \mathcal{F}$, então $A \cup B \in \mathcal{F}$.

**Proposição 1.4.** Se $\mathcal{F}$ é uma álgebra e se $A, B \in \mathcal{F}$, então

$$A \cap B \in \mathcal{F} \quad \text{e} \quad A \setminus B \in \mathcal{F}.$$

***Demonstração.*** cf. Exercício 6.

**Exemplo 1.1.** Sendo $X = \{0, 1, 2, 3\}$, ambas as famílias

$$\mathcal{F} = \{\emptyset, \{0, 2\}, \{1, 3\}, X\}] \quad \text{e} \quad \mathcal{G} = \{\emptyset, \{0, 1\}, \{2, 3\}, X\}$$

são álgebras sobre $X$.

**Definição 1.5.** Uma família $\mathcal{F} \subseteq \mathcal{P}(X)$ é uma ***álgebra***$-\sigma$ sobre $X$, se for uma álgebra (cf. Definição 1.4) e para além disso, verificar

4. Para toda família $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$, *crescente* (i.e. $A_1 \subseteq ... \subseteq A_i \subseteq ...$), temos

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F} \tag{1.2}$$

**Observações**

1. Vê-se imediatamente que $\mathcal{P}(X)$ é, ela própria, uma álgebra$-\sigma$. Por outro lado, comprova-se facilmente que a família $\{\emptyset, X\}$ é também uma álgebra$-\sigma$.
2. Toda álgebra finita é uma álgebra$-\sigma$. (cf. Exercício 3.)

**Nota 1.2.** (cf. Exercício 2.) Para toda álgebra $\mathcal{F}$ sobre $X$ (em particular para toda álgebra$-\sigma$ sobre $X$), temos

$$\{\emptyset, X\} \subseteq \mathcal{F} \subseteq \mathcal{P}(X).$$

Deduz-se que $\{\emptyset, X\}$ é uma álgebra$-\sigma$ (como já dissemos) e é a mais pequena de entre todas as álgebras de $X$. Ao mesmo tempo, $\mathcal{P}(X)$ é uma álgebra$-\sigma$ e é a maior de entre todas as álgebras de $X$.

**Lema 1.1.** *$\mathcal{F}$ é uma álgebra$-\sigma$ se e só se*

*1.* $X \in \mathcal{F}$.

*2. Se* $A \in \mathcal{F}$, *então* $X \setminus A \in \mathcal{F}$.

*5. Para toda família* $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$, *temos*

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}.$$

***Demonstração.*** Comparando com a alinha 4 da Definição 1.5, a hipótese 5 deste lema diz que não é necessário que a sucessão $\{A_i\}_{i\in\mathbb{N}}$ seja crescente, para concluir que $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$. Evidentemente, o que devemos mostrar é que a hipótese 5 deste lema é equivalente as propriedades 3 e 4 juntas (cf. Definições 1.4 e 1.5).

Suponhamos que $\mathcal{F}$ é uma álgebra$-\sigma$ e que $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$.

Para todo $n \in \mathbb{N}$, define-se

$$B_n = \bigcup_{i=1}^{n} A_i.$$

É claro (por indução) que para todo $n \in \mathbb{N}$, $B_n \in \mathcal{F}$. Efetivamente, se $B_n \in \mathcal{F}$, a alinha 3 e o facto que $A_{n+1} \in \mathcal{F}$, implicam que

$$B_{n+1} = B_n \cup A_{n+1} \in \mathcal{F}.$$

Mas como $B_1 = A_1 \in \mathcal{F}$, o raciocínio por indução está completo.

Como $\{B_n\}_{n\in\mathbb{N}}$ é uma família **crescente** de elementos de $\mathcal{F}$, então, por 4,

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{n=1}^{\infty} B_n \in \mathcal{F}.$$

A demonstração inversa (5 implica 3 e 4 ) é muito mais simples e fica como exercício (cf. Exercício 5). :-)

**Proposição 1.5.** Se $\mathcal{F}$ é uma álgebra$-\sigma$ e se $\{A_i\}_{i\in\mathbb{N}}$ é uma família numerável de elementos de $\mathcal{F}$, então

$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}. \tag{1.3}$$

***Demonstração.*** cf. Exercício 7.

**Nomenclatura.** Um par ordenado $(X, \mathcal{F})$, onde $X$ é um conjunto e $\mathcal{F}$ uma álgebra$-\sigma$ sobre $X$, chama-se ***espaço mensurável***. (Não se aplica quando $\mathcal{F}$ é simplesmente uma álgebra.)

Os elementos da álgebra$-\sigma$ $\mathcal{F}$ chamam-se ***conjuntos*** $\mathcal{F}-$***mensuráveis***. Quando a álgebra$-\sigma$ $\mathcal{F}$ é fixa ou está subentendida no contexto, podemos eliminar o prefixo e chamar aos elementos de $\mathcal{F}$ simplesmente ***conjuntos mensuráveis***.

## 1.2 Álgebra$-\sigma$ Envolvente

**Lema 1.2.** *A interseção de uma família não vazia de álgebras$-\sigma$ (resp. álgebras) sobre um conjunto $X$ é uma álgebra$-\sigma$ (resp. álgebra).*

*Explícitamente, sendo $\Phi$ uma família não vazia de álgebras$-\sigma$ (resp. álgebras) sobre $X$,*

$$\mathcal{C} = \bigcap_{\mathcal{F}\in\Phi} \mathcal{F}$$

*é uma álgebra$-\sigma$ (resp. álgebra) sobre $X$.*

***Demonstração.*** Para mostrar que $\mathcal{C}$ é uma álgebra$-\sigma$, vamos verificar uma a uma as propriedades 1, 2 e 5, do Lema 1.1.

1. Como todo $\mathcal{F} \in \Phi$ é uma álgebra$-\sigma$ (resp. álgebra), então para todo $\mathcal{F} \in \Phi$, $X \in \mathcal{F}$. Logo,
$$X \in \mathcal{C}.$$

2. Se $A \in \mathcal{C}$, então $A \in \mathcal{F}$ para todo $\mathcal{F} \in \Phi$. Como todo $\mathcal{F} \in \Phi$ é uma álgebra$-\sigma$ (resp. álgebra), $X \setminus A \in \mathcal{F}$, qualquer que seja $\mathcal{F} \in \Phi$. Então
$$X \setminus A \in \mathcal{C}.$$

5. Seja $\{A_i\}_{i\in\mathbb{N}}$ uma família numerável de elementos de $\mathcal{C}$. Então, para todo $\mathcal{F} \in \Phi$, $\{A_i\}_{i\in\mathbb{N}}$ é uma família numerável de elementos de $\mathcal{F}$, mas como $\mathcal{F}$ é uma álgebra$-\sigma$, então $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$, para todo $\mathcal{F} \in \Phi$. Por tanto,
$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{C}.$$

No caso de álgebras, só é necessário verificar que $\mathcal{C}$ é fechado para a operação $\cup$ (união de dois elementos). :-)

***Pergunta.*** Onde entra a hipótese $\Phi \neq \emptyset$?

A primeira consequência deste importante lema é a existência da álgebra$-\sigma$ gerada por uma família qualquer de subconjuntos e é o que vamos mostrar na próxima subsecção.

### 1.2.1 álgebra$-\sigma$ gerada por uma família de subconjuntos

**Lema 1.3.** *Seja $\mathcal{S} \subseteq \mathcal{P}(X)$, uma família qualquer de subconjuntos de $X$. Existe uma álgebra$-\sigma$, denotada $\sigma[\mathcal{S}]$, que é a álgebra$-\sigma$ mais pequena que contém $\mathcal{S}$. i.e. $\sigma[\mathcal{S}]$ verifica cumulativamente as condições seguintes*

1. *$\sigma[\mathcal{S}]$ é uma álgebra$-\sigma$.*
2. *$\mathcal{S} \subseteq \sigma[\mathcal{S}]$.*
3. *Se $\mathcal{F}$ é uma álgebra$-\sigma$ tal que $\mathcal{S} \subseteq \mathcal{F}$, então $\sigma[\mathcal{S}] \subseteq \mathcal{F}$.*

***Demonstração.*** Define-se

$$\Phi = \{\mathcal{F} \subseteq \mathcal{P}(X) \,; \mathcal{F} \text{ é uma álgebra} - \sigma \text{ e } \mathcal{S} \subseteq \mathcal{F}\}.$$

Pelo lema anterior (note que $\Phi \neq \emptyset$, já que $\mathcal{P}(X) \in \Phi$),

$$\sigma[\mathcal{S}] = \bigcap_{\mathcal{F} \in \Phi} \mathcal{F} \tag{1.4}$$

é uma álgebra$-\sigma$ sobre $X$. As outras duas propriedades do enunciado são consequências imediatas da definição de interseção. :-)

**Definição 1.6.** Seja $\mathcal{S} \subseteq \mathcal{P}(X)$, uma família qualquer de subconjuntos de $X$. A álgebra$-\sigma$ mais pequena que contém $\mathcal{S}$, denotada $\sigma[\mathcal{S}]$, cuja existência ficou estabelecida no lema anterior, chama-se ***álgebra$-\sigma$ gerada por $\mathcal{S}$*** ou ***álgebra$-\sigma$ envolvente de $\mathcal{S}$***.

***Observação.*** O mesmo procedimento permite-nos definir a álgebra gerada/envolvente por $\mathcal{S}$. Mas, a álgebra envolvente não é utilizada e por isso, não usamos nenhuma notação em especial para a designar.

### 1.2.2 Álgebra$-\sigma$ de Borel

Uma consequência extremamente importante do Lema 1.3 é a existência da $\sigma-$álgebra de Borel (cuja definição damos a seguir), a qual está relacionada com a topologia do espaço $X$ (assumindo que $X$ é um espaço topológico) e permite estabelecer uma relação entre funções contínuas e funções mensuráveis (definidas mais adelante).

Mas os únicos espaços topológicos que vamos considerar, agora no início, são: $\mathbb{R}$, com a topologia habitual, algúns dos seus subconjuntos (com a topologia inducida) e a compactificação $\overline{\mathbb{R}}$ de $\mathbb{R}$ da qual falaremos no capítulo 3. (Mais a frente falaremos também de $\mathbb{R}^n$, $n = 2, 3, ....$)

**Definição 1.7.** Seja $X$ um espaço topológico e $\mathcal{O}_X$ a família dos subconjuntos abertos de $X$. A ***álgebra−$\sigma$ de Borel*** de $X$ é a álgebra−$\sigma$ gerada pelos abertos de $X$ e denota-se $\mathcal{B}_X$. i.e.

$$\mathcal{B}_X = \sigma\left[\mathcal{O}_X\right].$$

Os elementos de $\mathcal{B}_X$ chamam-se ***conjuntos de Borel*** ou ***Borelianos*** de $X$.

Por enquanto, o único o espaço topológico que vamos considerar é a reta real. Em particular, álgebra−$\sigma$ de Borel de $\mathbb{R}$ denota-se $\mathcal{B}_{\mathbb{R}}$.

### 1.2.3 Estrutura reticulada sobre as álgebras−$\sigma$

**Relação de ordem**

**Definição 1.8.** Uma ***relação de ordem*** $\preceq$, sobre um conjunto $\mathcal{X}$, é uma relação que verifica as seguintes propriedades.

1. ***Reflexiva*:** Para todo $x \in \mathcal{X}$, $x \preceq x$.
2. ***Transitiva*:** Se $x \preceq y$ e $y \preceq z$, então $x \preceq z$.
3. ***Antisimétrica*:** Se $x \preceq y$ e $y \preceq x$, então $x = y$.

Um conjunto $\mathcal{X}$, junto com uma relação de ordem $\preceq$ é um ***conjunto ordenado***. Por exemplo, o conjunto $\mathcal{P}(X)$, junto com a relação de inclusão $\subseteq$ é um conjunto ordenado.

Um conjunto ordenado $\mathcal{X}$ é ***totalmente ordenado*** se

$$\forall\, x,\, y \in \mathcal{X},\;\; x \preceq y \;\text{ ou }\; y \preceq x.$$

Por exemplo, os conjuntos de números $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

O ***maior minorante*** de $x$ e $y \in \mathcal{X}$, denotado $\min(x, y)$ ou $x \wedge y$, é um elemento de $\mathcal{X}$ que verifica as seguintes condições.

($a$) $x \wedge y \preceq x$ e $x \wedge y \preceq y$.

($b$) Para todo $z \in \mathcal{X}$, se $z \preceq x$ e $z \preceq y$, então $z \preceq x \wedge y$.

Note que $x \wedge y$ pode não existir, mas se existir, é único.

O ***menor maiorante*** de $x$ e $y \in \mathcal{X}$, denotado $\max(x, y)$ ou $x \vee y$, é um elemento de $\mathcal{X}$ que verifica as seguintes condições.

$(a)$ $x \preceq x \vee y$ e $y \preceq x \vee y$.

$(b)$ Para todo $z \in \mathcal{X}$, se $x \preceq z$ e $y \preceq z$, então $x \vee y \preceq z$.

Da mesma forma, $x \vee y$ pode não existir, mas se existir, é único.

Um conjunto ordenado $\mathcal{X}$ é ***reticulado*** se

$$\forall\, x,\, y \in \mathcal{X},\;\; x \wedge y \;\; \text{e} \;\; y \vee x \;\; \text{existem.}$$

Outra consequência importante do Lema 1.3 é a seguinte proposição.

**Proposição 1.6.** O conjunto de todas as álgebras$-\sigma$ de um conjunto $X$ tem uma estrutura reticulada com respeito à relação de inclusão. Quer dizer que dadas duas álgebras$-\sigma$ $\mathcal{F}$ e $\mathcal{G}$ (ambas sobre o mesmo conjunto $X$), existem tanto o ***maior minorante*** como o ***menor maiorante*** de $\mathcal{F}$ e $\mathcal{G}$.

***Demonstração.*** Visto que a intersecção de álgebras$-\sigma$ é uma álgebra$-\sigma$, então o maior minorante das álgebras$-\sigma$ $\mathcal{F}$ e $\mathcal{G}$ é igual ao maior minorante dos *conjuntos* $\mathcal{F}$ e $\mathcal{G}$. i.e.,

$$\mathcal{F} \wedge \mathcal{G} = \mathcal{F} \cap \mathcal{G}.$$

Ao contrário do que acontece com a interseção, a união de álgebras$-\sigma$ não é, em geral, uma álgebra$-\sigma$. Mas graças ao Lema 1.3, a álgebra$-\sigma$ gerada por $\mathcal{F} \cup \mathcal{G}$ é o menor maiorante. i.e.

$$\mathcal{F} \vee \mathcal{G} = \sigma\left[\mathcal{F} \cup \mathcal{G}\right]. \quad :-)$$

***Observação.*** Pela Nota 1.2, existe um único elemento minimal que é $\{\emptyset, X\}$ e um único elemento maximal que é $\mathcal{P}(X)$.

**Exemplo 1.2.** (cf. Exemplo 1.1.) Sejam

$$\mathcal{F} = \{\emptyset, \{0,2\}, \{1,3\}, X\} \quad \text{e} \quad \mathcal{G} = \{\emptyset, \{0,1\}, \{2,3\}, X\}.$$

Então,

$$\mathcal{F} \wedge \mathcal{G} = \{\emptyset, X\} \quad \text{e} \quad \mathcal{F} \vee \mathcal{G} = \mathcal{P}\left(X\right).$$

## 1.3 Sucessões de Bernoulli

Um dos modelos probabilísticos mais utilizados para descrever a ocorrência ou não dum fenómeno associado a uma série de experiências são as sucessões de Bernoulli, que definimos a continuação.

**Definição 1.9.** Chamamos ***sucessões de Bernoulli*** às sucessões (infinitas) de 0's e 1's. i.e.

$$\{0,1\}^{\mathbb{N}} = \{(\omega_1, \omega_2, ...) ; \omega_i \in \{0,1\}, i \in \mathbb{N}\}. \tag{1.5}$$

***Observação.*** O conjunto $\{0,1\}^{\mathbb{N}}$ é igual ao conjunto de todas as aplicações de $\mathbb{N}$ em $\{0,1\}$. Pela Proposição 1.3, $\{0,1\}^{\mathbb{N}} \sim \mathcal{P}(\mathbb{N})$. Em particular $\{0,1\}^{\mathbb{N}}$ não é numerável. (cf. Nota 1.1.)

Utilizam-se as sucessões de Bernoulli como espaço básico para descrever os resultados (positivo ou negativo) duma sequência de "experiências aleatórias." Na realidade, uma sequência de experiências é sempre finita, mas é importante ter um espaço abstrato onde possamos dar sentido matemático concreto às noções eurísticas que queremos estudar, e este tem de ser infinito, por razões que só serão aparentes mais tarde.

Naturalmente, as sequências finitas fazem parte importante do modelo e é através delas que vamos definir a estrutura do espaço mensurável de Bernoulli e é delas que tratamos na subsecção seguinte.

**Nota 1.3.** Devemos ter atenção para não confundir "experiência aleatória" com "variável aleatória." As variáveis aleatórias são certo tipo de funções dentro da Teoria das Probabilidades, que serão definidas mais a frente. i.e. As variáveis aleatórias são entes puramente matemáticos e, na realidade, nada têm de aleatório.

### 1.3.1 Álgebra$-\sigma$ sobre as sucessões de Bernoulli

#### Sequências finitas

Sendo $n \in \mathbb{N}$ qualquer, denotamos $\{0,1\}^n$ o produto Cartesiano de $n$ cópias idénticas do conjunto $\{0,1\}$. i.e.

$$\{0,1\}^n = \{(\xi_1, \xi_2, ..., \xi_n) ; \xi_i \in \{0,1\}, i = 1, 2, ..., n\}$$

é o conjunto das sequências de 0's e 1's, de comprimento fixo $n$. e.g.

$$\{0,1\}^2 = \{(0,0) ; (0,1) ; (1,0) ; (1,1)\}.$$

***Observação.*** O conjunto $\{0,1\}^n$ tem $2^n$ elementos. $\{0,1\}^n$ serve, entre outras coisas, para representar o conjunto das partes de qualquer conjunto $X$ con $n$ elementos. (cf. Proposição 1.1.)

**Nomenclatura.** Para facilitar a leitura, estabelecemos a seguinte convenção: Quando falamos de sequências, elas são sempre finitas, enquanto que as sucessões são sempre infinitas.

Para além disso, note-se a utilização diferenciada das letras gregas $\omega$ e $\xi$. Exceto nos exercícios 17 e 18, onde $\xi = (\xi_1, \xi_2, ...) \in \{0,1\}^{\mathbb{N}}$, normalmente $\xi_i$ representa um valor dentro duma sequência finita, enquanto que $\omega_i$ denota um valor dentro de uma sucessão infinita. Da mesma forma, $\omega$ representa uma sucessão de Bernoulli ($\omega \in \{0,1\}^{\mathbb{N}}$), enquanto que $\xi$ representa uma sequência, independentemente do seu comprimento, i.e.

$$\xi \in \bigcup_{n=1}^{\infty} \{0,1\}^n .$$

### Função de truncadura das sucessões de Bernoulli

**Definição 1.10.** Para todo $n \in \mathbb{N}$ fixo, a função ***truncadura*** de comprimento $n$ é definida da seguinte forma.

$$\forall\, \omega = (\omega_1, ..., \omega_n, \omega_{n+1}, ...) \in \{0,1\}^{\mathbb{N}}, \ \mathrm{T}_n(\omega) = (\omega_1, ..., \omega_n) \in \{0,1\}^n . \tag{1.6}$$

### Feixe de sucessões de Bernoulli

Por outro lado, todas as sucessões de Bernoulli cujos primeiros $n$ termos são iguais a uma sequência fixa formam um "feixe" de sucessões de Bernoulli, todas tendo a mesma imagem através da função $\mathrm{T}_n$.

**Definição 1.11.** Sendo $\xi = (\xi_1, ..., \xi_n) \in \{0,1\}^n$ um elemento qualquer fixo, o ***feixe de parte inicial*** $\xi$, denotado $A_\xi$, é o conjunto das sucessões de Bernoulli que têm a parte inicial igual a $\xi$. i.e.

$$A_\xi = \mathrm{T}_n^{-1}(\{\xi\}) = \left\{(\omega_1, \omega_2, ...) \in \{0,1\}^{\mathbb{N}} ; \, \omega_i = \xi_i \, , \, i = 1, ..., n\right\}. \tag{1.7}$$

**Nota 1.4.** Sendo $n \in \mathbb{N}$ fixo, o conjunto de todos os feixes

$$\mathcal{I}_n = \{A_\xi ; \, \xi \in \{0,1\}^n\}$$

é uma partição de $\{0,1\}^{\mathbb{N}}$. Além disso,

$$\mathcal{I}_n \sim \{0,1\}^n .$$

**Proposição 1.7.** A álgebra$-\sigma$ gerada pela partição $\mathcal{I}_n$, denotada $\mathcal{F}_n$, é finita e é simplesmente igual às uniões de elementos de $\mathcal{I}_n$. i.e.

$$\mathcal{F}_n = \sigma\left[\mathcal{I}_n\right] = \left\{ \bigcup_{\xi \in J} A_\xi \,;\;\; J \subseteq \{0,1\}^n \right\}. \tag{1.8}$$

Noutras palabras,

$$\mathcal{F}_n \sim \mathcal{P}\left(\{0,1\}^n\right).$$

Em particualr,

$$\mathrm{Card}\left(\,\mathcal{F}_n\right) = \mathrm{Card}\left(\mathcal{P}\left(\{0,1\}^n\right)\right) = 2^{2^n}.$$

Para além disso, $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ é uma sucessão crescente de álgebras$-\sigma$ sobre $\{0,1\}^{\mathbb{N}}$. (cf. Exercício 14.) i.e.

$$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq ... \subseteq \mathcal{F}_n \subseteq ....$$

***Observação.*** Apesar do resultado do Exercício 11, o qual implica que

$$\bigcup_{n=1}^{\infty} \mathcal{F}_n$$

é uma álgebra, ela não é uma álgebra$-\sigma$. (cf. Exercício 15.)

**Definição 1.12.** A álgebra$-\sigma$ naturalmente associada com $\{0,1\}^{\mathbb{N}}$ é a álgebra$-\sigma$ $\mathcal{F}_\infty$, gerada pela união de todas as álgebras$-\sigma$ $\mathcal{F}_n$. Ou seja

$$\mathcal{F}_\infty := \bigvee_{n=1}^{\infty} \mathcal{F}_n := \sigma\left[\bigcup_{n=1}^{\infty} \mathcal{F}_n\right], \tag{1.9}$$

à qual nos referiremos como Álgebra$-\sigma$ de Bernoulli.

## 1.3.2 Filtração de Bernoulli

Todas as álgebras$-\sigma$ $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ são importantes e, no seu conjunto, formam o que denominamos de ***filtração*** do espaço (mensurável) $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty\right)$.

**Definição 1.13.** Em geral, uma ***filtração*** de um espaço mensurável $(X, \mathcal{G})$ é uma família de álgebras$-\sigma$, digamos $\{\mathcal{G}_t\}_{t\in T}$, indexada por $T = \mathbb{N}$ (no caso discreto) ou $T = \mathbb{R}^+$ (no caso contínuo), crescente, todas elas contidas em $\mathcal{G}$. i.e.

$$\forall\, t_1, t_2 \in T,\; t_1 \leq t_2,\; \mathcal{G}_{t_1} \subseteq \mathcal{G}_{t_2} \subseteq \mathcal{G}.$$

***Observação.*** Este conceito é ocupa um lugar central no estudo dos processos estocásticos e encontra-se normalmente associado à noção de informação estatística. Intuitivamente, as filtrações restringem às funções ou toma de decisões à informação disponível até o tempo $n$, quer dizer, elas refletem matemáticamente o facto de não podermos adivinhar o futuro.

## 1.4 Funções mensuráveis

**Definição 1.14.** Sejam $(X, \mathcal{F})$ e $(Y, \mathcal{G})$ dois espaços mensuráveis. Uma função $f : X \to Y$ é $\mathcal{F}/\mathcal{G}-$***mensurável*** se

$$\forall A \in \mathcal{G},\ f^{-1}(A) \in \mathcal{F}.$$

Quando as álgebras$-\sigma$ $\mathcal{F}$ e $\mathcal{G}$ estão subentendidas no contexto, podemos eliminar a referência a elas e falar simplesmente de funções mensuráveis em lugar de $\mathcal{F}/\mathcal{G}-$mensuráveis.

**Proposição 1.8.** Seja $(X, \mathcal{F})$ um espaço mensurável. Para todo $A \subseteq X$,

$$A \text{ é } \mathcal{F}-\text{mensurável (i.e. } A \in \mathcal{F}) \text{ se e só se } \chi_A \text{ é } \mathcal{F}/\mathcal{P}(\{0,1\})-\text{mensurável.}$$

***Demonstração.*** (cf. Exercício 30.)

**Exemplos**

1. Sejam $X = \{0,1,2,3\}$, $\mathcal{F} = \{\emptyset, \{0,2\}, \{1,3\}, X\}$ e $\mathcal{G} = \{\emptyset, \{0,1\}, \{2,3\}, X\}$. (cf. Exemplo 1.1.) A função definida sobre $X$ por

$$f(0) = f(2) = 0 \quad \text{e} \quad f(1) = f(3) = 1$$

é $\mathcal{F}/\mathcal{P}(\{0,1\})-$mensurável, mas não é $\mathcal{G}/\mathcal{P}(\{0,1\})-$mensurável.

2. $\mathrm{T}_n$ (cf. Definição 1.10) é $\mathcal{F}_m/\mathcal{P}(\{0,1\}^n)-$mensurável, para todo $m \geq n$. Mas não é mensurável si $m < n$. (cf. (1.8).)

3. Imaginemos um jogo no qual ganha-se ou perde-se 1€ depois de cada jogada. Se jogarmos $n$ partidas, o capital que ganhamos ou perdemos, como função de $\{0,1\}^{\mathbb{N}}$, é uma função $\mathcal{F}_n-$mensurável. (cf. Exercício 22.)

   O espaço de chegada é, naturalmente, $\{-n, ..., -1, 0, 1, ..., n\}$ com a álgebra$-\sigma$ das suas partes.

4. Imaginemos que entramos no jogo do exemplo 3 com um capital de 1€ e o jogo estende-se até ficarmos sem capital. (Supondo que a casa nunca fecha e que nós temos a eternidade para jogar). O número de partidas que jogamos até ficar sem capital é uma função $\mathcal{F}_\infty-$mensurável. (cf. Exercício 23.)

   O espaço de chegada neste exercício e no próximo é $\mathbb{N}\cup\{\infty\}$ com a álgebra$-\sigma$ das suas partes.

5. Imaginemos que saimos do jogo do exemplo 3 depois da primeira vez que perdemos. A variável que regista o número de vezes que ganhamos antes de sair é $\mathcal{F}_\infty-$mensurável. (cf. Exercício 24.)

***Observação.*** A demonstração de todas as afirmações nestes exemplos torna-se mais simples se aplicarmos o resultado da alinha 2 do Lema 1.4, a seguir.

### 1.4.1 Álgebra$-\sigma$ gerada por uma função

O seguinte lema é muito importante e será utilizado em vários sítios mais a frente. A demonstração é direta a partir das definições e fica como exercício para o leitor (cf. Exercício 31).

**Lema 1.4.** Sejam $X$ e $Y$ dois conjuntos e seja $f : X \to Y$ uma função qualquer. Então,

(*i*) se $\mathcal{G}$ é uma álgebra$-\sigma$ sobre $Y$, a família

$$\left\{f^{-1}(A)\,;\; A \in \mathcal{G}\right\}$$

é uma álgebra$-\sigma$ sobre $X$.

(*ii*) para todo $\mathcal{S} \subseteq \mathcal{P}(Y)$,

$$\left\{f^{-1}(A)\,;\; A \in \sigma[\mathcal{S}]\right\} = \sigma\left[\left\{f^{-1}(A)\,;\; A \in \mathcal{S}\right\}\right].$$

(*iii*) se $\mathcal{F}$ é uma álgebra$-\sigma$ sobre $X$ e $\mathcal{G}$ é uma álgebra$-\sigma$ sobre $Y$, então a função $f : X \to Y$ é $\mathcal{F}/\mathcal{G}-$mensurável se e só se

$$\left\{f^{-1}(A)\,;\; A \in \mathcal{G}\right\} \subset \mathcal{F}.$$

**Definição 1.15.** Seja $X$ um conjunto, $(Y, \mathcal{G})$ um espaço mensurável e $f : X \to Y$ uma função qualquer. A ***álgebra$-\sigma$ gerada por*** $f$, denotada $\sigma[f]$, é a mais pequena álgebra$-\sigma$ sobre $X$, tal que a função $f$ é mensurável.

Pelo Lema 1.4,

$$\sigma\left[f\right]=\left\{f^{-1}\left(A\right);\,A\in\mathcal{G}\right\}.$$

***Observação.*** A álgebra$-\sigma$ $\sigma\left[f\right]$, para além de depender de $f$, depende também de $\mathcal{G}$, mas geralmente $\mathcal{G}$ é fixo no contexto.

**Exemplo 1.3.** (cf. Exercício 21.) A álgebra$-\sigma$ gerada por $\mathrm{T}_n$ (Definição 1.10) não é outra coisa que a álgebra$-\sigma$ $\mathcal{F}_n$ definida por (1.8) (cf. Proposição 1.7). i.e.

$$\mathcal{F}_n=\sigma\left[\mathrm{T}_n\right].$$

### Restrição de uma álgebra$-\sigma$ a um subconjunto

Como aplicação da alinha 1 do Lema 1.4, temos a seguinte definição.

**Definição 1.16.** Sendo $(X,\mathcal{F})$ um espaço mensurável e $E\subseteq X$ um subconjunto qualquer (mensurável ou não). A restrição de $\mathcal{F}$ a $E$, denotada $\left.\mathcal{F}\right|_E$, é a álgebra$-\sigma$ gerada pela função $\mathrm{id}_E : E\to X$, definida por $\mathrm{id}_E(x)=x$, $x\in E$. i.e.

$$\left.\mathcal{F}\right|_E=\sigma\left[\mathrm{id}_E\right]=\left\{A\cap E;\,A\in\mathcal{F}\right\}.$$

***Observação.*** Desde logo, $\left.\mathcal{F}\right|_E$ é uma álgebra$-\sigma$ sobre $E$ e $(E,\left.\mathcal{F}\right|_E)$ é um espaço mensurável. Para além disso, se $E\in\mathcal{F}$, então $\left.\mathcal{F}\right|_E\subseteq\mathcal{F}$. No entanto, em caso nenhum $\left.\mathcal{F}\right|_E$ é uma sub-álgebra$-\sigma$ de $\mathcal{F}$ (exceto se $E=X$).

# Exercícios

1. Sendo $X$ um conjunto qualquer, utilize a mesma ideia explorada na demonstração da Proposição 1.2, para demonstrar que não existe nenhuma aplicação **sobrejetiva** de $X$ em $F\left(X;\{0,1\}\right)$, sendo $F\left(X;\{0,1\}\right)$ o conjunto de todas as funções de $X$ em $\{0,1\}$.

2. Sendo $X$ um conjunto qualquer, mostre que para toda álgebra $\mathcal{F}$ sobre $X$ (em particular para toda álgebra$-\sigma$ sobre $X$), temos

$$\{\emptyset,X\}\subseteq\mathcal{F}\subseteq\mathcal{P}(X).$$

3. Mostre que toda álgebra finita é uma álgebra$-\sigma$.

4. Mostre que toda álgebra$-\sigma$ ou é finita ou é não numerável.

5. Complete a demonstarção do Lema 1.1.

6. Seja $\mathcal{F}$ uma álgebra e sejam $A, B \in \mathcal{F}$, mostre que
$$A \cap B \in \mathcal{F} \quad \text{e} \quad A \setminus B \in \mathcal{F}.$$

7. Sendo $\mathcal{F}$ uma álgebra$-\sigma$ e $\{A_i\}_{i\in\mathbb{N}}$ uma família numerável de elementos de $\mathcal{F}$, mostre que
$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}.$$

8. Utilize o exercício 7 para mostrar que $\{x\} \in \mathcal{B}_{\mathbb{R}}$, qualquer que seja $x \in \mathbb{R}$. ($\mathcal{B}_{\mathbb{R}}$ denota a álgebra$-\sigma$ de Borel de $\mathbb{R}$, cf. Definição 1.7.)

9. Seja $\mathcal{F}_0$ a família dos subconjuntos de $\mathbb{N}$ que são finitos ou que o seu complementar é finito. Quer dizer,
$$A \in \mathcal{F}_0 \Leftrightarrow \text{Card}\,(A) < \infty \text{ ou } \text{Card}\,(\mathbb{N} \setminus A) < \infty,$$
(sendo $\text{Card}\,(A)$ o número de elementos do conjunto $A$). Mostre que $\mathcal{F}_0$ é uma álgebra, mas não é uma álgebra$-\sigma$.

10. Seja $\mathcal{G}_0$ a família composta pelo conjunto vazio e por uniões finitas de elementos do conjunto
$$\mathcal{R} = \{(a, b] \subseteq (0, 1]\,;\, 0 \leq a < b \leq 1\}\,.$$
Mostre que $\mathcal{G}_0$ é uma álgebra sobre o intervalo $(0, 1]$, mas não é uma álgebra$-\sigma$.

11. Seja $\Phi$ uma família de álgebras. Suponhamos que $\Phi$ é uma cadeia no sentido da teoria de conjuntos, ou seja $\Phi$ é totalmente ordenada com respeito à inclusão. i.e.
$$\forall\, \mathcal{F}, \mathcal{G} \in \Phi,\ \mathcal{F} \subseteq \mathcal{G} \text{ ou } \mathcal{F} \subseteq \mathcal{G}.$$
Mostre que
$$\bigcup_{\mathcal{F}\in\Phi} \mathcal{F}$$
é uma álgebra.

12. Seja $\mathcal{U}$ a família dos subconjuntos unitários de $\mathbb{N}$. Quer dizer,
$$\mathcal{U} = \{\{n\}\,;\, n \in \mathbb{N}\}\,.$$
Descreva tanto a álgebra como a álgebra$-\sigma$ geradas por $\mathcal{U}$. Comparar com a álgebra do exercício 9.

13. (cf. Exercício 10).) Sendo

$$\mathcal{D} = \left\{ \left(0, \frac{k}{2^n}\right] \subseteq (0,1]\,;\, n \in \mathbb{N} \wedge k = 1, ..., 2^n \right\},$$

mostre que $\mathcal{R} \subseteq \sigma[\mathcal{D}]$ e conclua que $\sigma[\mathcal{D}] = \sigma[\mathcal{R}]$.

Nos exercícios 14 ao 23, o conjunto de base é $\{0,1\}^{\mathbb{N}}$ (cf. Definição 1.9). $\mathcal{F}_n$ é a álgebra$-\sigma$ definida por (1.8) (cf. Proposição 1.7) e $\mathcal{F}_\infty$ e a definida por (1.9) (cf. Definição 1.12). A função $\mathrm{T}_n$ foi definida em (1.10) (cf. Definição 1.10).

14. Demonstre as afirmações da Proposição 1.7.

15. Seja $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ a filtração natural do espaço mensurável $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty\right)$. Mostre que $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ não é uma álgebra$-\sigma$. Quer dizer,

$$\mathcal{F}_\infty \neq \bigcup_{n=1}^{\infty} \mathcal{F}_n. \tag{1.10}$$

16. Mostre que os subconjuntos unitários de $\{0,1\}^{\mathbb{N}}$ pertencem a $\mathcal{F}_\infty$. i.e.

$$\forall\, \omega \in \{0,1\}^{\mathbb{N}},\ \{\omega\} \in \mathcal{F}_\infty.$$

17. Sendo $\xi \in \{0,1\}^{\mathbb{N}}$ e $I \subset \mathbb{N}$ fixos, mostre que

$$A_{\xi,I} = \left\{\omega \in \{0,1\}^{\mathbb{N}}\,;\, \omega_i = \xi_i,\ \forall\, i \in I\right\} \in \mathcal{F}_\infty.$$

Este exercício generaliza o exercício 16, já que $A_{\xi,\mathbb{N}} = \{\xi\}$.

18. Mostre que a família

$$\left\{A_{\xi,I} \subseteq \{0,1\}^{\mathbb{N}}\,;\, \xi \in \{0,1\}^{\mathbb{N}},\ I \subset \mathbb{N}\right\},$$

onde $A_{\xi,I}$ são os subconjuntos definidos no exercício 17, não é uma álgebra (inclusive aumentando o conjunto vazio). Mas, demonstre que

$$\mathcal{F}_\infty = \sigma\left[\left\{A_{\xi,I} \subseteq \{0,1\}^{\mathbb{N}}\,;\, \xi \in \{0,1\}^{\mathbb{N}},\ I \subset \mathbb{N}\right\}\right].$$

19. Consideremos a família de subconjuntos unitários,

$$\mathcal{S}_u = \left\{\{\omega\} \subseteq \{0,1\}^{\mathbb{N}}\,;\, \omega \in \{0,1\}^{\mathbb{N}}\right\}. \tag{1.11}$$

Qual é a álgebra$-\sigma$ gerada por $\mathcal{S}_u$?

20. Sendo $n \in \mathbb{N}$ fixo, mostre que $\mathrm{T}_n$ é $\mathcal{F}_m-$mensuráveis, para todo $m \geq n$, mas não o é para $m < n$.

21. (cf. Exemplo 1.3.) Mostre que para todo $n \in \mathbb{N}$,

$$\mathcal{F}_n = \sigma\left[\mathrm{T}_n\right].$$

22. Mostre que num jogo no qual ganha-se ou perde-se 1€ em cada, o capital que ganhamos ou perdemos depois de $n$ partidas é uma função $\mathcal{F}_n-$mensurável.

    O conjunto de chegada é $\{-n, ..., -1, 0, 1, ..., n\}$ (sem a função ser sobrejetiva), com a álgebra$-\sigma$ das partes. Encontre a imagem propriamente dita.

23. Imaginemos entrar no jogo do exercício 22 com um capital de1€ e jogar até o capital esgotar (supondo que a casa nunca fecha e que nós temos a eternidade para jogar). Mostre que o número de partidas que jogamos até ficar sem capital é uma função $\mathcal{F}_\infty-$mensurável.

    O conjunto de chegada é $\mathbb{N}\cup\{\infty\}$, com a álgebra$-\sigma$ das suas partes. Encontre a imagem propriamente dita.

24. Imaginemos que saimos do jogo do exercício 22 depois da primeira vez que perdemos. Mostre que a variável que regista o número de vezes que ganhamos antes de sair é $\mathcal{F}_\infty-$mensurável.

25. Sendo $X = \{1, 2, 3\}$ e $\mathcal{F} = \{\emptyset, \{1\}, \{2, 3\}, \{1, 2, 3\}\}$, mostre que a função

$$f(1) = f(3) = 1 \quad \text{y} \quad f(2) = 0$$

    não é mensurável.

26. Sendo $X = \{0, 1, 2, 3\}$ e $\mathcal{F} = \{\emptyset, \{0, 2\}, \{1, 3\}, X\}$, encontrar todas as funções mensuráveis de $X$ em $\{0, 1\}$.

27. Seja $\mathcal{F}$ como no exercício 26 e $\mathcal{G} = \{\emptyset, \{0, 1\}, \{2, 3\}, X\}$. Encontrar todas as aplicações de $X$ em si mesmo, $\mathcal{F}/\mathcal{G}-$mensuráveis.

28. Sejam $\mathcal{F}$ e $\mathcal{G}$ duas álgebras$-\sigma$ sobre um conjunto $X$ qualquer. Mostre que a função

$$\mathrm{id}_X(x) = x, \ \forall\, x \in X$$

    é $\mathcal{F}/\mathcal{G}-$mensurável se e só se $\mathcal{G} \subseteq \mathcal{F}$.

29. Sejam $(X, \mathcal{F})$ e $(Y, \mathcal{G})$ dois espaços mensuráveis. Demonstre as afirmações seguintes.

    (a) Se $\mathcal{F} = \mathcal{P}(X)$, toda função $f : X \to Y$ é $\mathcal{P}(X)/\mathcal{G}-$mensurável.

(b) Se $\mathcal{G} = \{\emptyset, Y\}$, toda função $f : X \to Y$ é $\mathcal{F}/\{\emptyset, Y\}-$mensurável.

(c) Para toda álgebra$-\sigma$ $\mathcal{F}' \supseteq \mathcal{F}$ e toda álgebra$-\sigma$ $\mathcal{G}' \subseteq \mathcal{G}$, se

$$f : X \to Y$$

é $\mathcal{F}/\mathcal{G}-$mensurável, então, $f$ é também $\mathcal{F}'/\mathcal{G}'-$mensurável.

(d) Toda função $f : X \to Y$, constante (i.e. $f(x) = c$, para todo $x \in X$, sendo $c \in Y$ fixo), é $\mathcal{F}/\mathcal{G}-$mensurável, quaisquer que sejam as álgebras$-\sigma$ $\mathcal{F}$ e $\mathcal{G}$ de $X$ e $Y$, respetivamente.

30. Seja $(X, \mathcal{F})$ um espaço mensurável qualquer. Mostre que para todo $A \subseteq X$, $A$ é mensurável ($A \in \mathcal{F}$) se e só se a função característica $\chi_A$ (cf. Definição 1.3) é mensurável.

31. Demonstre o Lema 1.4.

# Capítulo 2

# Medidas

## 2.1 Medida

**Definição 2.1.** Dado um espaço mensurável $(X, \mathcal{F})$, uma ***medida*** sobre $\mathcal{F}$ (ou mais explicitamente, uma medida sobre $(X, \mathcal{F})$) é uma aplicação

$$\mu : \mathcal{F} \to \overline{\mathbb{R}}^+ \qquad \left(\overline{\mathbb{R}}^+ = [0, \infty]\right)$$

tal que

1. ***Nulidade do vazio***: $\mu(\emptyset) = 0$,

2. ***Aditividade numerável para conjuntos disjuntos***:

   Se $\{A_i\}_{i \in \mathbb{N}}$ é uma família numerável de elementos de $\mathcal{F}$, disjuntos dois a dois (i.e. $\forall i \neq j, A_i \cap A_j = \emptyset$), então

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i). \tag{2.1}$$

***Observação.*** Esta definição é extremamente geral e não é muito satisfatória em certos contextos. Normalmente, as medidas sobre conjuntos con estruturas matemáticas ricas (como por exemplo as medidas de Radon) verificam outras propriedades para além das enunciadas acima, de forma a fazer o casamento com a estrutura do conjunto de base. Neste livro vamos considerar unicamente medidas gerais, mas que verificam a propriedade da Definição 2.2 mais a frente.

### Propriedades imediatas

**Proposição 2.1.** Toda medida $\mu$, sobre um espaço $(X, \mathcal{F})$, verifica as seguintes propriedades.

1. ***Aditividade para conjuntos disjuntos***: $\forall\, A, B \in \mathcal{F}$ tais que $A \cap B = \emptyset$,
$$\mu\,(A \cup B) = \mu\,(A) + \mu\,(B)\,. \tag{2.2}$$

2. ***Monotonia***: $\forall\, A, B \in \mathcal{F}$. Se $A \subseteq B$, então
$$\mu\,(A) \leq \mu\,(B)\,.$$

3. ***Regra geral de aditividade***: $\forall\, A, B \in \mathcal{F}$,
$$\mu\,(A \cup B) + \mu\,(A \cap B) = \mu\,(A) + \mu\,(B)\,. \tag{2.3}$$

4. ***Subaditividade finita***: $\forall\, A, B \in \mathcal{F}$,
$$\mu\,(A \cup B) \leq \mu\,(A) + \mu\,(B)\,. \tag{2.4}$$

***Demonstração.*** Vamos mostrar unicamente a alinha 3, ficando as outras três a cargo do leitor. (cf. Exercício 1.)

Note que $A \setminus B$, $B \setminus A$ e $A \cap B$ são disjuntos dois a dois e, além disso,
$$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)\,.$$

Logo,
$$\mu\,(A \cup B) = \mu\,(A \setminus B) + \mu\,(B \setminus A) + \mu\,(A \cap B)\,. \tag{2.5}$$

Da mesma forma,
$$A = (A \setminus B) \cup (A \cap B) \quad \text{e} \quad B = (B \setminus A) \cup (A \cap B)\,.$$

Então,
$$\mu\,(A) = \mu\,(A \setminus B) + \mu\,(A \cap B) \quad \text{e} \quad \mu\,(B) = \mu\,(B \setminus A) + \mu\,(A \cap B)\,.$$

Somando estas duas igualdades termo a termo, obtemos
$$\mu\,(A) + \mu\,(B) = [\,\mu\,(A \setminus B) + \mu\,(B \setminus A) + \mu\,(A \cap B)\,] + \mu\,(A \cap B)\,. \tag{2.6}$$

substituindo (2.5) em (2.6), obtemos (2.3). :-)

**Medidas $\sigma-$finitas**

**Definição 2.2.** Uma medida $\mu$, sobre um espaço $(X, \mathcal{F})$, chama-se $\sigma-$***finita*** se existir uma família numerável $\{X_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$ (que podemos assumir crescente), tal que

$$\forall i \in \mathbb{N},\ \mu(X_i) < \infty \qquad \text{e} \qquad \bigcup_{i=1}^{\infty} X_i = X.$$

***Observação.*** Se restringirmos o domínio de uma medida $\sigma-$finita $\mu$ a uma sub-álgebra$-\sigma$ de $\mathcal{F}$, a medida resultante pode deixar de ser $\sigma-$finita (Exercício 2). Apesar desta patologia, esta propriedade é fundamental para demonstrar a unicidade da medida obtida pelo processo de extensão de Carathéodory. (cf. Teorema 2.1 e Proposição 2.6, mais a frente.)

**Definição 2.3.** O tripleto $(X, \mathcal{F}, \mu)$, formado por um conjunto $X$, uma álgebra$-\sigma$ $\mathcal{F}$ sobre $X$ e uma medida sobre o espaço mensurável $(X, \mathcal{F})$, chama-se ***espaço de medida***.

Se para além disso $\mu(X) = 1$, dizemos que $\mu$ é uma ***probabilidade*** ou ***medida de probabilidade***. Nesse caso, $(X, \mathcal{F}, \mu)$ chama-se ***espaço de probabilidade*** ou ***espaço probabilístico*** $(X, \mathcal{F}, \mu)$.

Tradicionalmente, um espaço de probabilidade genérico, denota-se por $(\Omega, \mathcal{F}, \mathrm{P})$, em vez de $(X, \mathcal{F}, \mu)$. Os subconjuntos de $\Omega$, pertencentes a $\mathcal{F}$, chamam-se ***eventos***.

**Exercício.** Consideremos a partição do conjunto $\mathbb{N}$ pela relação de equivalência *módulo* 3. i.e. para todo $a, b \in \mathbb{N}$, $a \sim b$ se e só se 3 for um divisor de $|a - b|$. Encontre a medida de probabilidade P, sobre a álgebra$-\sigma$ gerada pela partição $\{\overline{0}, \overline{1}, \overline{2}\}$, que toma os valores

$$\mathrm{P}\left(\overline{1}\right) = \frac{1}{3} \quad \text{e} \quad \mathrm{P}\left(\overline{2}\right) = \frac{1}{6}.$$

***Solução.*** Sendo $\{\overline{0}, \overline{1}, \overline{2}\}$ uma partição em três partes, a álgebra$-\sigma$ gerada por esta partição, que vamos denotar $\mathcal{P}_3$, tem só oito elementos e podemos a escrever por extenso.

$$\mathcal{P}_3 = \sigma[\{\overline{0}, \overline{1}, \overline{2}\}] = \{\emptyset,\ \overline{0},\ \overline{1},\ \overline{2},\ \overline{0} \cup \overline{1},\ \overline{0} \cup \overline{2},\ \overline{1} \cup \overline{2},\ \mathbb{N}\}.$$

Para definir uma medida sobre $\mathcal{P}_3$, basta estipular o seu valor nos "átomos" (i.e. nos elementos da partição). Nos outros elementos de $\mathcal{P}_3$, a medida será simplesmente a soma dos valores dos átomos que conformam o conjunto. e.g.

$$\mathrm{P}\left(\overline{0}\cup\overline{1}\right)=\mathrm{P}\left(\overline{0}\right)+\mathrm{P}\left(\overline{1}\right).$$

Foram dados os valores em duas das três partes. Sendo P uma medida de probabilidade, que dizer que $\mathrm{P}(\mathbb{N})=1$, e como

$$\mathrm{P}(\mathbb{N})=\mathrm{P}\left(\overline{0}\right)+\mathrm{P}\left(\overline{1}\right)+\mathrm{P}\left(\overline{2}\right),$$

então

$$\mathrm{P}\left(\overline{0}\right)=1-\mathrm{P}\left(\overline{1}\right)-\mathrm{P}\left(\overline{2}\right)=1-\frac{1}{3}-\frac{1}{6}=\frac{1}{2}. :-)$$

**Exemplos**

(a) A medida de contagem dos subconjuntos de $\mathbb{N}$ é uma medida $\sigma-$finita sobre o espaço mensurável $(\mathbb{N},\mathcal{P}(\mathbb{N}))$.

(b) Estudaremos mais a frente (cf. Teorema 2.2) a medida de "comprimento" sobre os números reais. (O Teorema 2.2 trata só do intervalo $(0,1]$, mas o raciocínio para a reta toda é o mesmo). A medida de comprimento é uma medida definida sobre $\mathcal{B}_{\mathbb{R}}$ (os borelianos de $\mathbb{R}$), sendo a medida de qualquer intervalo, o comprimento do mesmo.

Note que o facto de decidir (arbitrariamente) o valor da medida sobre os intervalos, não garante que de facto existe uma medida com essa característica (a medida dos intervalos é o seu comprimento). Para além disso, mesmo existindo uma tal medida, ainda temos de garantir que não há mais do que uma, pois estamos a definir a medida unicamente através do seu valor nos intervalos. Noutras palabras, será que conhecendo o valor da medida sobre os intervalos é suficiente para extender a definição de maneira única a todos os borelianos. A resposta é afirmativa, como mostraremos no Teorema 2.2, que por sua vez baseia-se no teorema de Carathéodory (cf. Teorema 2.1).

(c) A medida "produto" $\mathrm{P}_{\frac{1}{2}}$ sobre as sucessões de Bernoulli é a medida definida sobre $\mathcal{F}_\infty$ (cf. Definição 1.12). Nela, a medida de um feixe $A_\xi$, com $\xi\in\{0,1\}^n$, é $2^{-n}$. De forma semelhante ao que acontece no exemplo anterior, temos de provar que uma tal medida existe. (cf. Teorema 2.3.)

## 2.2 Transporte de medidas

### 2.2.1 Medida imagem

**Lema 2.1.** *Sendo $(X, \mathcal{F}, \mu)$ um espaço de medida, $(Y, \mathcal{G})$ um espaço mensurável e $f : X \to Y$ uma função $\mathcal{F}/\mathcal{G}-$mensurável, a função*

$$A \in \mathcal{G} \to \mu\left(f^{-1}(A)\right). \tag{2.7}$$

*é uma medida sobre o espaço mensurável $(Y, \mathcal{G})$.*

***Demonstração.*** Como $f^{-1}(\emptyset) = \emptyset$, então $\nu(\emptyset) = \mu(f^{-1}(\emptyset)) = 0$. Por outro lado, seja $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{G}$, disjuntos dois a dois (i.e. $\forall i \neq j, A_i \cap A_j = \emptyset$), então

$$f^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} f^{-1}(A_i).$$

Para além disso, os conjuntos $f^{-1}(A_i)$ são disjuntos dois a dois. Por (2.1),

$$\begin{aligned}
\nu\left(\bigcup_{i=1}^{\infty} A_i\right) &= \mu\left(f^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right)\right) \\
&= \mu\left(\bigcup_{i=1}^{\infty} f^{-1}(A_i)\right) \\
&= \sum_{i=1}^{\infty} \mu\left(f^{-1}(A_i)\right) \\
&= \sum_{i=1}^{\infty} \nu(A_i). \quad :-)
\end{aligned}$$

**Definição 2.4.** A medida $\nu$, definida pela fórmula (2.7) do Lema 2.1, chama-se ***medida imagem*** de $\mu$ por $f$ e denota-se $f_{\#}\mu$. i.e. Para todo $A \in \mathcal{G}$,

$$f_{\#}\mu(A) := \mu\left(f^{-1}(A)\right).$$

#### Exemplos

(a) Seja $X$ uma urna com 16 bolas, com a medida de contagem. Se pintarmos quatro bolas de azul, três de vermelho, duas de verde e sete de amarelo, então no conjunto das cuatro cores, podemos atribuir a cada cor a medida igual ao número de bolas com esa cor. i.e.

$$\nu(\{\text{azul}\}) = 4,\ \nu(\{\text{vermelho}\}) = 3,\ \ \nu(\{\text{verde}\}) = 2,\ \ \nu(\{\text{amarelo}\}) = 7.$$

(b) Seja $X$ o "universo" de todas as pessoas deste planeta, dotado da medida de contagem. Se fizermos corresponder a cada pessoa o animal do zodíaco chinese correspondente ao ano do seu nascimento. O conjunto imagem é o conjunto dos 12 animais do zodíaco e a medida imagem é a medida que associa a cada animal o número de pessoas com ese signo.

Para sermos rigorosos, devemos dizer que a medida imagem associa a cada **conjunto unitário** de um animal, o número de pessoas com ese signo. Mas tratando-se de conjuntos finitos, como é o caso destes dois exemplos, é usual determinar a medida através da função, sobre o conjunto de base, que a cada elemento associa a medida do conjunto unitário contendo esse elemento.

(c) (cf. Teorema 2.4.) A aplicação, definida em (2.36),

$$\psi : (\omega_1, \omega_2, ...) \in \{0,1\}^{\mathbb{N}} \longrightarrow \sum_{n=1}^{\infty} \frac{\omega_n}{2^n} \in [0,1]$$

transforma a medida de probabilidade $\mathrm{P}_{\frac{1}{2}}$, sobre $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty\right)$, na medida de comprimento sobre $(0,1]$. i.e.

$$\lambda = \psi_{\#}\mathrm{P}_{\frac{1}{2}}.$$

Inversamente, a aplicação, definida em (2.37),

$$\varphi : x = \sum_{n=1}^{\infty} \frac{\omega_n}{2^n} \in [0,1] \longrightarrow (\omega_1, \omega_2, ...) \in \{0,1\}^{\mathbb{N}}$$

transforma a medida $\lambda$, sobre $([0,1], \mathcal{B})$ na medida $\mathrm{P}_{\frac{1}{2}}$, sobre $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty\right)$, i.e.

$$\varphi_{\#}\lambda = \mathrm{P}_{\frac{1}{2}}.$$

(Para os números diádicos, a representação em série não é única, cf. Nota 2.7.)

**Nota 2.1.** Esta "quase" bijeção $(\psi/\phi)$ apareceu no célebre artigo de Émile Borel [2], "Sur les probabilités dénombrables et leurs applications arithmétiques," em 1909, pois está relacionada com o teorema de Borel dos números normais. Os números normais são aqueles que no seu desenvolvimento binário, têm, asimptóticamente em média, tantos 1's como 0's. (cf. Definição 4.15.)

Segundo o teorema de Borel (cf. Teorema 4.8), com respeito à medida de comprimento, quase todos os números reais entre 0 e 1 são normais. Este resultado é uma consequência da lei forte dos grandes números (cf. Teorema 4.7) e o facto que $\lambda = \psi_{\#}\mathrm{P}_{\frac{1}{2}}$.

Este artigo passou à história como sendo o primeiro a aplicar a noção de independência a objetos matemáticos bem definidos.

***Observação.*** O conceito de medida imagem está sempre presente na teoria da medida e integração. Por exemplo na mudança de variável de integração. (cf. Teorema 4.3.)

### 2.2.2 Medida invariante

**Definição 2.5.** Seja $f : X \to X$ uma função mensurável dum espaço de medida $(X, \mathcal{F}, \mu)$ em si mesmo. Dizemos que $\mu$ é ***invariante*** por $f$ ou que $f$ deixa $\mu$ invariante, se para todo $A \in \mathcal{F}$

$$\mu(A) = \mu\left(f^{-1}(A)\right).$$

i.e. $\mu$ é ***invariante*** por $f$, se $f_{\#}\mu = \mu$.

**Exemplos**

(a) Seja $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathrm{Card}\,)$ o espaço dos númeors naturais, com a medida de contagem. Qualquer função **bijetiva** de $\mathbb{N}$ em $\mathbb{N}$ deixa invariante a medida de contagem.

(b) Sendo $a \in \mathbb{R}$ fixo, a função $x \to a+x$ deixa invariante a medida de comprimento sobre $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. (cf. Exercício 10.)

(c) O "shift" é a aplicação

$$(\omega_1, \omega_2, \omega_3, ...) \in \{0,1\}^{\mathbb{N}} \to (\omega_2, \omega_3, ...) \in \{0,1\}^{\mathbb{N}}.$$

Esta função deixa invariante a medida de probabilidade $\mathrm{P}_p$, sobre $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty\right)$, qualquer que seja $p \in [0,1]$. (cf. Teorema 2.3.)

***Observação.*** O conceito de medida invariante é central na teoria ergódica. Mas neste livro, não exploramos esta vertente.

### 2.2.3 Medida preimagem

**Definição 2.6.** Seja $(Y, \mathcal{G}, \nu)$ é um espaço de medida e $f : X \to Y$ uma função definida sobre um conjunto cualquer $X$, a valores em $Y$. A ***medida preimagem***

de $\nu$ é uma medida sobre o espaço mensurável $(X, \sigma[f])$, denotemos-lhe $\mu$, tal que
$$\nu = f_{\#}\mu.$$

**Nota 2.2.** A medida preimagem, se existir é única. Mas, ao contrário da medida imagem, a medida preimagem pode não existir. A condição necessária e suficiente para ela existir é dada pela proposição seguinte.

**Proposição 2.2.** A fórmula
$$f^{-1}(A) \to \nu(A)$$
define a medida preimagem sobre $(X, \sigma[f])$, se e só se
$$\forall\, A, B \in \mathcal{G},\ f^{-1}(A) = f^{-1}(B) \Rightarrow \nu(A) = \nu(B)\,, \tag{2.8}$$

**Nota 2.3.** A condição (2.8) está garantida, por exemplo, se $f$ for sobrejetiva. Efetivamente, como
$$\forall\, A \subseteq Y,\ f\left(f^{-1}(A)\right) = A,$$
então
$$\forall\, A, B \subseteq Y,\ \ f^{-1}(A) = f^{-1}(B) \Rightarrow A = B.$$

Portanto, no mínimo para as funções sobrejetivas, podemos transportar a medida do espaço medido de chegada $(Y, \mathcal{G}, \nu)$, para o espaço mensurável $(X, \sigma[f])$ de saída. Note que a álgebra$-\sigma$ do espaço de saída também é determinada por $f$.

## 2.3 Teoremas de convergência para conjuntos

Em toda esta secção, $(X, \mathcal{F}, \mu)$ é um espaço de medida fixo.

### Uniões crescentes e limites

**Lema 2.2.** *Seja $\{A_i\}_{i\in\mathbb{N}}$ uma família crescente de conjuntos $\mathcal{F}-$mensuráveis. Então,*
$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \mu(A_i)\,. \tag{2.9}$$

***Demonstração.*** Pondo $A_0 = \emptyset$, consideramos a sucessão auxiliar definida por

$$A'_i = A_i \setminus A_{i-1}, \ i \in \mathbb{N}.$$

Note que para todo $i \neq j$, $A'_i \cap A'_j = \emptyset$. Além disso,

$$\forall\, n \in \mathbb{N}, \ \bigcup_{i=1}^{n} A'_i = A_n \quad \text{e} \quad \bigcup_{i=n+1}^{\infty} A_i = \bigcup_{i=n+1}^{\infty} A'_i. \tag{2.10}$$

Por (2.1),

$$\mu\left(\bigcup_{i=1}^{\infty} A'_i\right) = \sum_{i=1}^{\infty} \mu\left(A'_i\right) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu\left(A'_i\right) = \lim_{n\to\infty} \mu\left(\bigcup_{i=1}^{n} A'_i\right). \tag{2.11}$$

Substituindo (2.10) em (2.11), obtemos (2.9). :-)

***Observação.*** O Lema 2.2 é uma versão rudimentar do teorema da convergência monótona. (cf. Teorema 4.2.)

### Subaditividade numerável

**Lema 2.3.** *Para toda família $\{A_i\}_{i\in\mathbb{N}}$, de conjuntos $\mathcal{F}-$mensuráveis, temos*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu\left(A_i\right). \tag{2.12}$$

***Demonstração.*** Como $\bigcup_{i=1}^{n} A_i$ é crescente em $n$, por (2.9) e (2.4), temos

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu\left(\bigcup_{i=1}^{n} A_i\right) \leq \lim_{n\to\infty} \sum_{i=1}^{n} \mu\left(A_i\right) = \sum_{i=1}^{\infty} \mu\left(A_i\right). \ :-)$$

### Interseção decrescente de conjuntos

**Lema 2.4.** *Seja $\{A_i\}_{i\in\mathbb{N}}$ uma família decrescente de conjuntos $\mathcal{F}-$mensuráveis. Suponhamos que existe $i_0 \in \mathbb{N}$, tal que $\mu\left(A_{i_0}\right) < \infty$. Então,*

$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \mu\left(A_i\right). \tag{2.13}$$

***Demonstração.*** Como $A_i \subseteq A_{i_0}$, para todo $i \geq i_0$,

$$\mu(A_{i_0}) = \mu(A_{i_0} \setminus A_i) + \mu(A_i).$$

Pela mesma razão,

$$\begin{aligned} \mu(A_{i_0}) &= \mu\left(A_{i_0} \setminus \bigcap_{i=i_0}^{\infty} A_i\right) + \mu\left(\bigcap_{i=i_0}^{\infty} A_i\right) \\ &= \mu\left(\bigcup_{i=i_0}^{\infty} (A_{i_0} \setminus A_i)\right) + \mu\left(\bigcap_{i=i_0}^{\infty} A_i\right). \end{aligned}$$

Igualando estas duas expressões e tomando o limite quando $i \to \infty$,

$$\mu\left(\bigcup_{i=i_0}^{\infty} (A_{i_0} \setminus A_i)\right) + \mu\left(\bigcap_{i=i_0}^{\infty} A_i\right) = \lim_{i\to\infty} \mu(A_{i_0} \setminus A_i) + \lim_{i\to\infty} \mu(A_i). \tag{2.14}$$

Por outro lado, observe que $\{A_{i_0} \setminus A_i\}_{i \geq i_0}$ é uma família crescente, de conjuntos $\mathcal{F}-$mensuráveis. Por (2.9),

$$\mu\left(\bigcup_{i=i_0}^{\infty} (A_{i_0} \setminus A_i)\right) = \lim_{i\to\infty} \mu(A_{i_0} \setminus A_i).$$

Como este valor é finito (limitado por $\mu(A_{i_0})$), pode-se substrair de ambos os lados da equação (2.14), obtendo-se precisamente (2.13).

**Nota 2.4.** O Lema 2.4 não é completamente simétrico com respeito ao Lema 2.2, devido à condição $\mu(A_{i_0}) < \infty$. Podemos considerar este teorema como uma versão rudimentar do teorema de convergência dominada (cf. Teorema 4.4) que é uno dos resultados fundamentais da teoria da integração.

A importância e a natureza profunda da hipótese $\mu(A_{i_0}) < \infty$, pode-se apreciar no seguinte exemplo.

**Exemplo 2.1.** Sendo $i \in \mathbb{N}$, pomos $A_i = \{j \in \mathbb{N};\ j \geq i\}$. Todos estes conjuntos têm um número infinito de elementos. i.e. Para todo $i \in \mathbb{N}$, $\mathrm{Card}(A_i) = \infty$.

Por outro lado $\bigcap_{i=1}^{\infty} A_i = \emptyset$ e portanto, $\mathrm{Card}\left(\bigcap_{i=1}^{\infty} A_i\right) = 0$.

### Definição equivalente de medida

Aquí vamos demonstrar que a aditividade finita e a subaditividade numerável caracterizam as medidas. Este resultado será invocado várias vezes nas secções seguintes.

**Lema 2.5.** *Dado um espaço mensurável* $(X, \mathcal{F})$*, uma aplicação* $\mu : \mathcal{F} \to \overline{\mathbb{R}}^+$ *é uma medida se e só se se verificarem as três condições seguintes.*

1. **Nulidade do vazio**: $\mu(\emptyset) = 0$,

2. **Superaditividade finita**: *Para todo* $A, B \in \mathcal{F}$*, tal que* $A \cap B = \emptyset$*, temos*

$$\mu(A \cup B) \geq \mu(A) + \mu(B). \tag{2.15}$$

3. **Subaditividade numerável**: *Para todo* $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i). \tag{2.16}$$

***Demonstração.*** Pela Definição 2.1, a Proposição 2.1 e o Lema 2.3, uma medida verifica estas propriedades. Inversamente, para demonstrar que estas propriedades caracterizam as medidas, devemos mostrar a aditividade numerável, i.e. a validade de (2.1) (cf. Definição 2.1).

Graças a (2.15), $\mu$ é monótona. i.e. Para todo $A, B \in \mathcal{F}$, tal que $A \subseteq B$,

$$\mu(A) \leq \mu(B). \tag{2.17}$$

Agora, fixemos uma família numerável $\{A_i\}_{i \in \mathbb{N}}$ de conjuntos $\mathcal{F}-$mensuráveis, disjuntos dois a dois.

Por (2.15) (aplicada reiteradamente) e por (2.17), para todo $n \in \mathbb{N}$,

$$\sum_{i=1}^{n} \mu(A_i) \leq \mu\left(\bigcup_{i=1}^{n} A_i\right) \leq \mu\left(\bigcup_{i=1}^{\infty} A_i\right).$$

Fazendo tender $n \to \infty$, obtemos

$$\sum_{i=1}^{\infty} \mu(A_i) \leq \mu\left(\bigcup_{i=1}^{\infty} A_i\right).$$

A outra desigualdade é precisamente a hipótese (2.16). :-)

## 2.4 Medidas a partir de premedidas

### 2.4.1 Premedidas

**Definição 2.7.** Seja $\mathcal{F}_0$ uma álgebra sobre um conjunto $X$. Uma aplicação $\mu_0 : \mathcal{F}_0 \to \overline{\mathbb{R}}^+$ é uma ***premedida*** se se cumplirem as seguintes propriedades.

1. ***Nulidade do vazio***:
$$\mu_0(\emptyset) = 0.$$

2. ***Aditividade numerável***:

   Seja $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}_0$, disjuntos dois a dois (i.e. $\forall\, i \neq j,\ A_i \cap A_j = \emptyset$).
$$\text{Se } \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_0, \text{ então } \mu_0\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu_0(A_i). \tag{2.18}$$

**Nota 2.5.** A única diferença entre uma medida e uma premedida é a estrutura do domínio. Por esta razão, a equação (2.18) verifica-se sob a condição $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_0$. De facto sem essa condição $\mu_0\left(\bigcup_{i=1}^{\infty} A_i\right)$ poderia não ter sentido.

Portanto, as premedidas verificam as mesmas propriedades das medidas, sempre e quando as expressões involucradas tenham sentido. Em particular e para uma referência expedita vamos reformular o Lema 2.5 e dar a versão adaptada ao contexto das premedidas. A demonstração é igual à demonstração do Lema 2.5 e é omissa.

**Lema 2.6.** *Sejam $X$ um conjunto e $\mathcal{F}_0$ uma álgebra sobre $X$. $\mu_0 : \mathcal{F}_0 \to \overline{\mathbb{R}}^+$ é uma premedida se e só se se verificarem as três condições seguintes.*

*1.* **Nulidade do vazio***:*
$$\mu_0(\emptyset) = 0. \tag{2.19}$$

*2.* **Superditividade finita***: Para todo $A, B \in \mathcal{F}_0$, tal que $A \cap B = \emptyset$,*
$$\mu_0(A \cup B) \geq \mu_0(A) + \mu_0(B). \tag{2.20}$$

*3.* **Subaditividade numerável***: Para todo $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}_0$, para todo $A \in \mathcal{F}_0$, tal que $A \subseteq \bigcup_{i=1}^{\infty} A_i$,*
$$\mu_0(A) \leq \sum_{i=1}^{\infty} \mu_0(A_i). \tag{2.21}$$

**Definição 2.8.** Uma premedida $\mu_0$ sobre uma álgebra $\mathcal{F}_0$ chama-se $\sigma-$finita, se existir uma familia numerável $\{X_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}_0$ tal que
$$\forall i \in \mathbb{N},\ \mu_0(X_i) < \infty \qquad \text{e} \qquad \bigcup_{i=1}^{\infty} X_i = X.$$

***Observação.*** A importância das premedidas radica no facto que, usualmente, as medidas são descritas primeiro e de forma explícita sobre álgebras e só depois, graças ao teorema de Carathéodory (cf. Teorema 2.1), são estendidas de maneira única (quando a premedida é $\sigma-$finita) à álgebra$-\sigma$ gerada pela álgebra onde a premedida está definida originalmente. Os dois grandes exemplos que acompanham a exposição da matéria, a medida de Lebesgue (comprimento) sobre $\mathbb{R}$ (cf. Teorema 2.2) e as medidas produto $\mathrm{P}_p$, $p \in (0,1)$, sobre $\{0,1\}^{\mathbb{N}}$ (cf. Teorema 2.3) são prova disso.

**Teorema 2.1.** (Carathéodory) *Seja $\mathcal{F}_0$ uma álgebra sobre um conjunto $X$ e $\mu_0 : \mathcal{F}_0 \to \overline{\mathbb{R}}^+$ uma premedida $\sigma-$finita. Existe uma única medida sobre o espaço mensurável $(X, \sigma[\mathcal{F}_0])$, que denotaremos $\mu$, tal que*

$$\forall A \in \mathcal{F}_0,\ \mu(A) = \mu_0(A). \tag{2.22}$$

**Nota 2.6.** A demonstração deste teorema fundamental é algo técnica e necessita de novos conceitos, que serão apresentados nas secções 2.5 e 2.6. Por enquanto, vamos nos concentrar na igualmente importante tarefa de aplicar este teorema para demonstrar a existência das duas medidas que acompanham o nosso discurso: A medida de Lebesgue (comprimento) sobre $(0,1]$ e as medidas produto sobre as sucessões de Bernoulli. Fica a demonstração do Teorema 2.1 para ser desenvolvida em duas partes, a existência na Secção 2.5 (cf. Proposição 2.4) e a unicidade na Secção 2.6 (cf. Proposição 2.6).

### 2.4.2 Medida de Comprimento sobre $(0,1]$

**Teorema 2.2.** (Lebesgue) *Existe uma única medida sobre $\big((0,1], \mathcal{B}_{(0,1]}\big)$, que denotaremos $\lambda$, tal que para todo intervalo $(a,b]$, $0 \leq a < b \leq 1$,*

$$\lambda((a,b]) = b - a. \tag{2.23}$$

***Demonstração.*** Pelo Teorema 2.1, basta mostrar a existência de uma premedida que verifique (2.23), definida sobre uma álgebra cuja álgebra$-\sigma$ envelope sejam os Borelianos de $(0,1]$. A demonstração será feita em vários passos.

***Passo 1.*** Vamos primeiro determinar a álgebra $\mathcal{F}_0$. Para isto, consideramos a família dos conjuntos obtidos como uniões dum número ***finito*** de intervalos semi-abertos $(a,b]$, $0 \leq a < b \leq 1$, disjuntos e ***separados***, incluindo também o conjunto vazio. i.e. $\mathcal{F}_0$ é a família composta pelos conjuntos da forma

$$(a_1, b_1] \cup (a_2, b_2] \cup ... \cup (a_n, b_n]$$

sendo $n \in \mathbb{N}$ qualquer e $0 \leq a_1 < b_1 < a_2 < b_2 < ... < a_n < b_n \leq 1$, e o vazio. (No Exercício 4, pede-se para demonstrar que $\mathcal{F}_0$ é de facto uma álgebra.)

Para além disso, observe que para todo $A \in \mathcal{F}_0$, $A \neq \emptyset$, a representação de $A$ como união de intervalos disjuntos e separados é **única** (cf. Exercício 5). Quer dizer, para todo $A \in \mathcal{F}_0$, $A \neq \emptyset$, existe um único $n \in \mathbb{N}$ e uma sequência de números $0 \leq a_1 < b_1 < a_2 < b_2 < ... < a_n < b_n \leq 1$, estrictamente crescente, de comprimento $2n$, única, tal que

$$A = (a_1, b_1] \cup (a_2, b_2] \cup ... \cup (a_n, b_n]\,. \tag{2.24}$$

Os intervalos (disjuntos e separados), $(a_i, b_i]$, $i = 1, 2, ..., n$, que intervêm em (2.24), chamam-se ***componentes*** de $A$.

***Passo 2.*** Agora vamos definir a premedida sobre $\mathcal{F}_0$. Tendo em conta (2.23), a escolha é evidente. Para todo $A = (a_1, b_1] \cup (a_2, b_2] \cup ... \cup (a_n, b_n] \in \mathcal{F}_0$,

$$\lambda_0 (A) = b_1 - a_1 + b_2 - a_2 + ... + b_n - a_n, \tag{2.25}$$

Se $A = \emptyset$, pomos $\lambda_0 (\emptyset) = 0$.

Mostremos agora que a função $\lambda_0$, definida por (2.25) é de facto uma premedida sobre $\mathcal{F}_0$. Segundo o Lema 2.6, para além da nulidade do vazio, temos de provar:

2. *Superaditividade finita*: Para todo $A, B \in \mathcal{F}_0$ tal que $A \cap B = \emptyset$,

$$\lambda_0 (A \cup B) \geq \lambda_0 (A) + \lambda_0 (B)\,. \tag{2.26}$$

*Demonstração.* Sejam $A, B \in \mathcal{F}_0$, tais que $A \cap B = \emptyset$. Em função das suas componentes respetivas, escrevemos

$$\begin{aligned} A &= (a_1, b_1] \cup (a_2, b_2] \cup ... \cup (a_n, b_n]\,, \\ B &= (c_1, d_1] \cup (c_2, d_2] \cup ... \cup (c_m, d_m] \\ A \cup B &= (e_1, f_1] \cup (e_2, f_2] \cup ... \cup (e_r, f_r]\,, \end{aligned}$$

onde $0 \leq a_1 < b_1 < a_2 < b_2 < ... < a_n < b_n \leq 1$, $0 \leq c_1 < d_1 < c_2 < d_2 < ... < c_m < d_m \leq 1$ e $0 \leq e_1 < f_1 < e_2 < f_2 < ... < e_r < f_r \leq 1$, sendo $r \leq n + m$.

Cada componente de $A \cup B$, i.e. cada intervalo $(e_i, f_i]$, $i = 1, 2, ..., r$, é a união de um ou vários intervalos disjuntos, mas ***adjacentes***, pertencentes tanto às componentes de $A$ como às de $B$.

Portanto, para mostrar (2.26), basta ver que dados três números quaisquer $0 \leq a < b < c \leq 1$,

$$\lambda_0 ((a, b] \cup (b, c]) = \lambda_0 ((a, b]) + \lambda_0 ((b, c]),$$

o qual é evidente.

3. *Subaditividade numerável*: Seja $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}_0$ e seja $A \in \mathcal{F}_0$, tal que $A \subseteq \bigcup_{i=1}^{\infty} A_i$. Devemos mostrar que

$$\mu_0 (A) \leq \sum_{i=1}^{\infty} \mu_0 (A_i).$$

*Demonstração.* Podemos assumir, sem pérda de geralidade, que os conjuntos $A_i$ são (eles próprios) intervalos semiabertos e que $A$ é um intervalo semiaberto, digamos que

$$A_i = (a_i, b_i] \quad \text{e} \quad (a, b] \subseteq \bigcup_{i=1}^{\infty} (a_i, b_i],$$

onde $a, a_i \in [0, 1)$, e $b, b_i \in (0, 1]$, $i = 1, 2, ....$ Agora, o nosso problema reduz-se a mostrar

$$b - a \leq \sum_{i=1}^{\infty} (b_i - a_i). \tag{2.27}$$

Utilizaremos a compacidade dos intervalos fechados e limitados. Para isto, fixamos $0 < \epsilon < b - a$ e consideramos os seguintes conjuntos e suas inclusões:

$$[a + \epsilon, b] \subseteq (a, b] = \bigcup_{i=1}^{\infty} (a_i, b_i] \subseteq \bigcup_{i=1}^{\infty} \left(a_i, b_i + \frac{\epsilon}{2^i}\right). \tag{2.28}$$

Note que $\left(a_i, b_i + \frac{\epsilon}{2^i}\right)$, $i = 1, 2, ...$ é um recobrimento aberto de $[a + \epsilon, b]$, que é compacto, por ser um intervalo fechado e limitado. Graças à compacidade, existe $n_0 \in \mathbb{N}$ tal que

$$[a + \epsilon, b] \subseteq \bigcup_{i=1}^{n_0} \left(a_i, b_i + \frac{\epsilon}{2^i}\right). \tag{2.29}$$

Não podemos aplicar $\lambda_0$ em ambos os lados desta inclusão como gostariamos, pois estes intervalos não pertencem a $\mathcal{F}_0$. Por esta razão, consideramos as seguintes inclusões.

$$(a+\epsilon, b] \subseteq [a+\epsilon, b] \subseteq \bigcup_{i=1}^{n_0} \left(a_i, b_i + \frac{\epsilon}{2^i}\right) \cap (0,1] \subseteq \bigcup_{i=1}^{n_0} \left(a_i, b_i + \frac{\epsilon}{2^i}\right] \cap (0,1] \, .$$

Aplicando $\lambda_0$ em (2.29), pela monotonía de $\lambda_0$ (cf. Exercício 6),

$$\lambda_0\left((a+\epsilon, b]\right) \leq \lambda_0 \left( \bigcup_{i=1}^{n_0} \left(a_i, b_i + \frac{\epsilon}{2^i}\right] \cap (0,1] \right).$$

Por (2.25) e a subaditivade finita (cf. Exercício 6),

$$b - a - \epsilon \leq \sum_{i=1}^{n_0} b_i + \frac{\epsilon}{2^i} - a_i \leq \epsilon + \sum_{i=1}^{\infty} b_i - a_i.$$

Como $\epsilon$ é arbitrário, a desigualdade (2.27) fica estabelecida.

Sendo $\lambda_0$ uma premedida finita, pelo Teorema 2.1, existe uma medida única, definida sobre $((0,1], \sigma[\mathcal{F}_0])$, que denotaremos $\lambda$. Por construção, $\lambda$ verifica (2.23).

***Passo 3.*** Só falta ver que

$$\sigma[\mathcal{F}_0] = \mathcal{B}_{(0,1]}.$$

Com efeito, para todo $0 \leq a < b \leq 1$,

$$(a, b] = \bigcap_{n=1}^{\infty} \left(a, b + \frac{1}{n}\right) \in \mathcal{B}_{(0,1]}$$

Por outro lado,

$$(a, b) = \bigcup_{n=1}^{\infty} \left(a, b - \frac{1}{n}\right] \in \sigma[\mathcal{F}_0] \, .$$

Como $\mathbb{Q}$ é numerável e $\mathbb{Q} \cap (0,1]$ é denso em $(0,1]$, todo aberto de $(0,1]$ é união numerável de intervalos abertos. Portanto, os intervalos abertos geram a álgebra$-\sigma$ de Borel. (cf. Lema 2.7.)

Desta forma, a demonstração do Teorema 2.2 fica completa. :-)

***Observação.*** Veremos mais a frente (cf. Definição 2.16), que a medida $\lambda$ pode ser estendida a uma álgebra$-\sigma$ maior, contendo estritamente os Borelianos. A medida $\lambda$ (estendida) conhece-se como ***medida de Lebesgue***, em honor a Henri Lebesgue quem pôs a teoria da integração sobre $\mathbb{R}$ em base firme em 1902, com a publicação da sua tese "Intégrale, longueur, aire" no Annali di Matematica, anterior ao teorema de extensão de Carathéodory.

**Lema 2.7.** *Para todo $A \subseteq (0,1]$, aberto, existe uma família $\{A_i\}_{i\in J}$, $J \subseteq \mathbb{N}$ (numerável ou finita), de intervalos abertos em $(0,1]$, tal que para todo $i \neq j$,*

$$A_i \cap A_j = \emptyset \quad \text{e} \quad A = \bigcup_{i\in J} A_i.$$

***Demonstração.*** Sendo $A \subseteq (0,1]$, um conjunto aberto, definimos uma relação de equivalência em $A$ (cf. Exercício 7), da forma seguinte

$$\forall\, a, b \in A,\ a \sim b \text{ se } [\min\{a,b\}, \max\{a,b\}] \subseteq A. \tag{2.30}$$

Note que as classes de equivalência são intervalos disjuntos, abertos em $(0,1]$. (cf. Exercício 8.)

Denotando $\mathcal{C}$ o conjunto das classes de equivalência e $\overline{x}$ a classe de equivalência que contem o ponto $x \in A$, como $\overline{x}$ é aberto e $\mathbb{Q} \cap (0,1]$ é denso em $(0,1]$, então $\overline{x} \cap \mathbb{Q} \neq \emptyset$.

Portanto, a aplicação

$$x \in A \cap \mathbb{Q} \to \overline{x} \in \mathcal{C}$$

é **sobrejetiva**. Sendo $A \cap \mathbb{Q}$ numerável, a imagem $\mathcal{C}$ é necessariamente numerável ou finita. Como $\mathcal{C}$ é uma partição de $A$, $A$ é a união disjunta, finita ou numerável, dos elementos de $\mathcal{C}$. :-)

### 2.4.3 Probabilidade produto sobre $\{0,1\}^{\mathbb{N}}$

Sendo $p \in (0,1)$, define-se

$$\nu_p(\{1\}) = p \qquad \text{e} \qquad \nu(\{0\}) = 1 - p. \tag{2.31}$$

Claramente, (2.31) define uma probabilidade sobre $(\{0,1\}, \mathcal{P}(\{0,1\}))$. Para simplificar a notação, escrevemos $\nu_p(1)$ e $\nu_p(0)$, em vez de $\nu_p(\{1\})$ e $\nu_p(\{0\})$, respetivamente.

**Teorema 2.3.** *Dado $p \in (0,1)$, existe uma única medida de probabilidad $\mathrm{P}_p$ sobre $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty\right)$, tal que para todo $n \in \mathbb{N}$, para todo $\xi = (\xi_1, \xi_2, ..., \xi_n) \in \{0,1\}^n$,*

$$\mathrm{P}_p(A_\xi) = \nu_p(\xi_1) \cdot \nu_p(\xi_2) \cdot ... \cdot \nu_p(\xi_n), \tag{2.32}$$

*sendo $A_\xi = \left\{(\omega_1, \omega_2, ...) \in \{0,1\}^{\mathbb{N}}; \omega_i = \xi_i, i = 1, 2, ..., n\right\}$ o feixe das sucessões de Bernoulli que têm a parte inicial igual à $(\xi_1, \xi_2, ..., \xi_n)$.* (cf. Definição 1.11.)

***Demonstração.*** Pelo Teorema 2.1, basta exibir uma premedida com estas características, sobre uma álgebra geradora de $\mathcal{F}_\infty$. Vamos conseguir isto em vários passos.

***Passo 1.*** Sendo $n$ fixo, para todo $J \subseteq \{0,1\}^n$, $J \neq \emptyset$, define-se

$$\nu_{p,n}(J) = \sum_{(\xi_1,\xi_2,\dots,\xi_n)\in J} \nu_p(\xi_1)\cdot\nu_p(\xi_2)\cdot\ldots\cdot\nu_p(\xi_n), \tag{2.33}$$

pomos ademais $\nu_{p,n}(\emptyset) = 0$. Observe que $\nu_{p,n}$ define uma probabilidade sobre $(\{0,1\}^n, \mathcal{P}(\{0,1\}^n))$. (cf. Exercício 11.)

A probabilidade $\nu_{p,n}$ é a probabilidade produto sobre o espaço $(\{0,1\}^n, \mathcal{P}(\{0,1\}^n))$, induzida por $\nu_p$.

Pela Proposição 1.7, para todo $A \in \mathcal{F}_n$, existe um único $J \subseteq \{0,1\}^n$, tal que

$$A = \bigcup_{\xi\in J} A_\xi. \tag{2.34}$$

(Por convenção, a união sobre o conjunto vazio é igual ao conjunto vazio.)

Para todo $A \in \mathcal{F}_n$, põe-se

$$\mathrm{P}_{p,n}(A) = \nu_{p,n}(J),$$

onde $J$ é o único conjunto verificando (2.34).

Note que $\mathrm{P}_{p,n}$ é a medida preimagem de $\nu_{p,n}$ pela função sobrejetiva $\mathrm{T}_n$ (cf. Definição 1.10). i.e.

$$\nu_{p,n} = (\mathrm{T}_n)_\#\mathrm{P}_{p,n}.$$

(cf. Proposição 2.2 e Nota 2.3.)

Note também que $\mathrm{P}_{p,n}$ verifica (2.32), para todo $\xi = (\xi_1,\xi_2,\dots,\xi_n) \in \{0,1\}^n$. Mas à diferença do Teorema 2.3, aqui $n$ é fixo.

***Passo 2.*** De seguida, vamos estabelecer a compatibilidade de $\{\mathrm{P}_{p,n}\}_{n\in\mathbb{N}}$. i.e. vamos ver que para todo $n \in \mathbb{N}$, para todo $A \in \mathcal{F}_n$,

$$\mathrm{P}_{p,n+1}(A) = \mathrm{P}_{p,n}(A). \tag{2.35}$$

De facto, pela Proposição 1.7, para todo $A \in \mathcal{F}_n$, existe um único $J \subseteq \{0,1\}^n$, tal que

$$A = \bigcup_{\xi\in J} A_\xi.$$

Para cada $\xi = (\xi_1,\dots,\xi_n) \in J$, põe-se

$$\xi^0 = (\xi_1,\dots,\xi_n,0) \quad \text{e} \quad \xi^1 = (\xi_1,\dots,\xi_n,1).$$

Como $\xi^0$ e $\xi^1 \in \{0,1\}^{n+1}$, $A_{\xi^0}$ e $A_{\xi^1} \in \mathcal{F}_{n+1}$. Portanto,

$$\bigcup_{\xi \in J} A_{\xi^0} \quad \text{e} \quad \bigcup_{\xi \in J} A_{\xi^1} \in \mathcal{F}_{n+1}.$$

Para além disso, $A_\xi = A_{\xi^0} \cup A_{\xi^1}$. Então,

$$A = \bigcup_{\xi \in J} (A_{\xi^0} \cup A_{\xi^1}) = \left( \bigcup_{\xi \in J} A_{\xi^0} \right) \bigcup \left( \bigcup_{\xi \in J} A_{\xi^1} \right) \in \mathcal{F}_{n+1}.$$

Logo,

$$\begin{aligned}
\mathrm{P}_{p,n+1}(A) &= \mathrm{P}_{p,n+1}\left( \bigcup_{\xi \in J} A_{\xi^0} \right) + \mathrm{P}_{p,n+1}\left( \bigcup_{\xi \in J} A_{\xi^1} \right) \\
&= \sum_{\xi \in J} \left( \nu_p(\xi_1) \cdot \ldots \cdot \nu_p(\xi_n) \cdot \nu_p(0) + \nu_p(\xi_1) \cdot \ldots \cdot \nu_p(\xi_n) \cdot \nu_p(1) \right) \\
&= \sum_{\xi \in J} \nu_p(\xi_1) \cdot \ldots \cdot \nu_p(\xi_n) \cdot (\nu_p(0) + \nu_p(1)) \\
&= \mathrm{P}_{p,n}(A),
\end{aligned}$$

já que $\nu_p(0) + \nu_p(1) = 1$. (cf. (2.31).)

***Passo 3.*** Agora vamos definir a premedida sobre $\bigcup_{n=1}^{\infty} \mathcal{F}_n$, que é uma álgebra (cf. Exercício 11 do Capítulo 1) e além disso, gera $\mathcal{F}_\infty$ (cf. Definição 1.12).

Por (2.35), a aplicação

$$A \in \bigcup_{n=1}^{\infty} \mathcal{F}_n \to \overline{\mathrm{P}}_p(A) = \mathrm{P}_{p,k}(A), \text{ se } A \in \mathcal{F}_k$$

está bem definida.

Como $\mathrm{P}_{p,n}$ é uma probabilidade, qualquer que seja $n \in \mathbb{N}$, $\overline{\mathrm{P}}_p$ verifica (2.15). Pelo Lema 2.6, só fica por demonstrar (2.21).

Em Topologia, o clássico teorema de Tychonoff estabelece que o produto Cartesiano de espaços topológicos compactos é compacto com respeito à topologia produto. Sendo $\{0,1\}^{\mathbb{N}}$ o produto Cartesiano de uma quantidade numerável de cópias do espaço compacto $\{0,1\}$, $\{0,1\}^{\mathbb{N}}$ é compacto com respeito à topologia produto.

Observe que todos os elmentos de $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ são abertos e ao mesmo tempo fechados e por tanto compactos na topologia produto. ($\{0,1\}^{\mathbb{N}}$, com a topologia produto, é completamente disconexo.)

Seja $\{A_i\}_{i\in\mathbb{N}} \subseteq \bigcup_{n=1}^{\infty} \mathcal{F}_n$ e $A \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$, tal que

$$A \subseteq \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_0.$$

Sendo $A$ compacto e $\{A_i\,;\ i \in \mathbb{N}\}$ um recobrimento de $A$ com abertos, existe $r \in \mathbb{N}$, tal que

$$A \subseteq \bigcup_{i=1}^{r} A_i.$$

Para cada $i = 1, ...r$, $A_i \in \mathcal{F}_{n_i}$, para algúm $n_i \in \mathbb{N}$, os conjuntos $A, A_1, A_2, ..., , A_r \in \mathcal{F}_N$, sendo $N = \max\{n_1, n_2, ..., n_r\}$.

Como $\mathrm{P}_{p,N}$ é uma probabilidade e portanto verifica (2.4),

$$\overline{\mathrm{P}}_p(A) = \mathrm{P}_{p,N}(A) \leq \sum_{i=1}^{r} \mathrm{P}_{p,N}(A_i) = \sum_{i=1}^{r} \overline{\mathrm{P}}_p(A_i) \leq \sum_{i=1}^{\infty} \overline{\mathrm{P}}_p(A_i)\,. \ \ : -)$$

**Isomorphismo entre** $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty, \mathrm{P}_{\frac{1}{2}}\right)$ **e** $([0,1], \mathcal{B}, \lambda)$

Quando as probabilidades de 0 e de 1 são iguais (i.e. $p = 1/2$), os espaços $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty, \mathrm{P}_{\frac{1}{2}}\right)$ e $([0,1], \mathcal{B}, \lambda)$ são isomorphos no sentido da teoria da medida, quer dizer que existem duas funções entre $\{0,1\}^{\mathbb{N}}$ e $[0,1]$, uma em cada sentido, mensuráveis, tal que transformam a medida do um na medida do outro. Estas duas funções são na realidade, e como é de esperar, inversas uma da outra, mas a condição de restringir o domínio de uma delas a um subconjunto de medida 1. Explicitamente, temos o seguinte teorema.

**Teorema 2.4.** *A aplicação*

$$\psi : (\omega_1, \omega_2, ...) \in \{0,1\}^{\mathbb{N}} \longrightarrow \sum_{n=1}^{\infty} \frac{\omega_n}{2^n} \in [0,1] \tag{2.36}$$

*transforma a medida* $\mathrm{P}_{\frac{1}{2}}$*, sobre* $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty\right)$*, na medida* $\lambda$*, sobre* $([0,1], \mathcal{B})$*. i.e.*

$$\psi_{\#}\mathrm{P}_{\frac{1}{2}} = \lambda$$

*Inversamente, a aplicação*

$$\varphi : x = \sum_{n=1}^{\infty} \frac{\omega_n}{2^n} \in [0,1] \longrightarrow (\omega_1, \omega_2, ...) \in \{0,1\}^{\mathbb{N}} \tag{2.37}$$

*transforma a medida $\lambda$, sobre $([0,1],\mathcal{B})$ na medida $\mathrm{P}_{\frac{1}{2}}$, sobre $\left(\{0,1\}^{\mathbb{N}},\mathcal{F}_\infty\right)$, i.e.*

$$\varphi_{\#}\lambda = \mathrm{P}_{\frac{1}{2}}$$

**Nota 2.7.** Para que $\varphi$, em (2.37), esteja bem definida, a representação de $x \in [0,1]$ em série de potências de $1/2$ deveria ser única. Mas a representação não é única para os números diádicos diferentes de 0 e de 1, quer dizer os números da forma

$$\frac{k}{2^n},\ k \in \{1,2,\,...\,,2^n-1\}\,,\ n \in \mathbb{N},$$

já que para representar estes números existem exatamente duas séries. Por exemplo,

$$\frac{1}{2} = \sum_{n=2}^{\infty}\frac{1}{2^n}.$$

Para que a função $\varphi$ fique bem definida, quando temos um número diádico, diferente de 0, escolhemos a série que termina com 1's, quer dizer, a série que é sempre 1 a partir de certa posição. (A escolha não tem relevância neste contexto, podiamos ter escolhido a série que termina em 0's.)

***Demonstração.*** Basta observar que para qualquer intervalo diádico

$$\left[\frac{k}{2^n}\,,\,\frac{k+1}{2^n}\right],\ k \in \{0,1,2,\,...\,,2^n-1\}\,,\ n \in \mathbb{N}$$

$$\psi{-1}\left(\left[\frac{k}{2^n}\,,\,\frac{k+1}{2^n}\right]\right) = A_\xi,$$

sendo $\xi = (\xi_1,\,...\,,\xi_n)$ o desenvolvimento de $k$ em base 2. i.e.

$$k = \xi_1\cdot 2^{n-1} + \cdots + \xi_{n-1}\cdot 2 + \xi_n.$$

Por outro lado, dado qualquer feixe $A_\xi$ (as sucessões com parte inicial fixa, igual a $\xi = (\xi_1,...,\xi_n) \in \{0,1\}^n$, cf.Definição 1.11), então,

$$\varphi^{-1}(A_\xi) = \left(\sum_{i=1}^{n}\frac{\xi_i}{2^i}\,,\,\frac{1}{2^n}+\sum_{i=1}^{n}\frac{\xi_i}{2^i}\right]$$

ou

$$\varphi^{-1}(A_\xi) = \left[0\,,\,\frac{1}{2^n}\right]$$

se $\xi = (0,...,0) \in \{0,1\}^n$.

Como

$$\mathrm{P}_{\frac{1}{2}}\left(\psi{-1}\left(\left[\frac{k}{2^n},\frac{k+1}{2^n}\right]\right)\right)=\mathrm{P}_{\frac{1}{2}}\left(A_\xi\right)=\frac{1}{2^n}=\lambda\left(\left[\frac{k}{2^n},\frac{k+1}{2^n}\right]\right),$$

$$\lambda\left(\varphi^{-1}\left(A_\xi\right)\right)=\lambda\left(\left[\sum_{i=1}^{n}\frac{\xi_i}{2^i},\frac{1}{2^n}+\sum_{i=1}^{n}\frac{\xi_i}{2^i}\right]\right)=\frac{1}{2^n}=\mathrm{P}_{\frac{1}{2}}\left(A_\xi\right)$$

e como os intervalos diádicos geram os borelianos (cf. Exercício **??**), o teorema está demonstrado. :-)

## 2.5 Medida exterior

Nesta secção vamos introduzir o conceito de "medida exterior," o qual tem muitas applicações na análise fina de funções. Mas o intuito principal neste contexto é de demonstrar a primeira parte do Teorema 2.1, ou seja a existência de uma medida que estende uma premedida dada. (cf. Proposição 2.4.)

**Definição 2.9.** Uma ***medida exterior***, sobre os subconjuntos de $X$, é uma função

$$\Psi:\mathcal{P}\left(X\right)\to\overline{\mathrm{R}}^{+}$$

que verifica as seguintes propriedades.

1. ***Nulidade do vazio***: $\Psi\left(\emptyset\right)=0$,
2. ***Monotonia***: Se $A\subseteq B\subseteq X$, então
$$\Psi\left(A\right)\leq\Psi\left(B\right).\tag{2.38}$$
3. ***Subaditividade numerável***: Para todo $\left\{A_i\right\}_{i\in\mathbb{N}}\subseteq\mathcal{P}\left(X\right)$,
$$\Psi\left(\bigcup_{i=1}^{\infty}A_i\right)\leq\sum_{i=1}^{\infty}\Psi\left(A_i\right).\tag{2.39}$$

**Exemplos**

(a) Seja $X$ um conjunto qualquer, a função $\Psi_0\left(A\right)\equiv 0$, para todo $A\subseteq X$ é uma medida exterior sobre $\mathcal{P}\left(X\right)$.

(b) A função $\Psi_1\left(A\right)=1$, para todo $A\neq\emptyset$, é uma medida exterior sobre $\mathcal{P}\left(X\right)$ (com $\Psi_1\left(\emptyset\right)=0$). Também podemos escrever como $\Psi_1\left(A\right)=1-\chi_{\{\emptyset\}}\left(A\right)$.

(c) Sendo $A \subseteq \mathbb{N} \times \mathbb{N}$, a função

$$\Psi(A) = \operatorname{Card}(\{n \in \mathbb{N};\ \exists\, m \in \mathbb{N},\ (n,m) \in A\})$$

é uma medida exterior sobre $\mathcal{P}(\mathbb{N} \times \mathbb{N})$.

***Observação.*** As medidas verificam estas três propriedades (cf. Definição 2.1, Proposição 2.1, Lema 2.3) mas o domínio das medidas é, em geral, mais restrito. Para apreciar a independência desta noção com respeito às medidas, nenhum dos exemplos anteriores é uma medida (exceto o primeiro!).

**Definição 2.10.** Seja $\Psi$ uma medida exterior e $A \subseteq X$. Diz-se que o conjunto $A$ é $\Psi-$***mensurável*** se

$$\forall\, E \subseteq X,\ \Psi(E \cap A) + \Psi(E \cap A^{\mathrm{c}}) \leq \Psi(E). \tag{2.40}$$

A família dos conjuntos $\Psi-$mensuráveis é denotada $\mathcal{F}_\Psi$. i.e.

$$\mathcal{F}_\Psi = \{A \subseteq X;\ \forall E \subseteq X,\ \Psi(E \cap A) + \Psi(E \cap A^{\mathrm{c}}) \leq \Psi(E)\} \tag{2.41}$$

**Nota 2.8.** $A \in \mathcal{F}_\Psi$ se e só se $A^{\mathrm{c}} \in \mathcal{F}_\Psi$. Para além disso, $\emptyset \in \mathcal{F}_\Psi$ e $X \in \mathcal{F}_\Psi$. De forma mais geral, temos o seguite resultado, cuja demonstração fica como exercício. (cf. Exercício 13.)

**Proposição 2.3.** Seja $\Psi$ uma medida exterior e $A \subseteq X$.

$$\text{Se } \Psi(A) = 0, \text{ então } A \in \mathcal{F}_\Psi.$$

**Exemplos**

(a) Seja $X$ um conjunto qualquer e $\Psi_0(A) \equiv 0$. Então $\mathcal{F}_{\Psi_0} = \mathcal{P}(X)$.

(b) Sendo $\Psi_1(A) = 1 - \chi_{\{\emptyset\}}(A)$, então $\mathcal{F}_{\Psi_1} = \{\emptyset, X\}$.

(c) Seja $\Psi(A) = \operatorname{Card}(\{n \in \mathbb{N};\ \exists\, m \in \mathbb{N},\ (n,m) \in A\})$, sendo $A \subseteq \mathbb{N} \times \mathbb{N}$. Então

$$\mathcal{F}_\Psi = \{J \times \mathbb{N};\ J \subset \mathbb{N}\}.$$

O resultado principal desta secção é o seguinte teorema.

**Teorema 2.5.** *Seja $\Psi$ uma medida exterior. Então, $\mathcal{F}_\Psi$ é uma álgebra$-\sigma$ sobre $X$ e $\Psi|_{\mathcal{F}_\Psi}$ é uma medida sobre $(X, \mathcal{F}_\Psi)$.* ($\Psi|_{\mathcal{F}_\Psi}$ é a restrição de $\Psi$ a $\mathcal{F}_\Psi$.)

***Demonstração.*** Por ser necessário na demonstração da primeira parte ($\mathcal{F}_\Psi$ é uma álgebra$-\sigma$), vamos provar primeiro a segunda afirmação, na versão seguinte:

***Passo 1.*** Se $\mathcal{F} \subset \mathcal{F}_\Psi$ é uma álgebra$-\sigma$, então $\Psi|_\mathcal{F}$ é uma medida sobre $(X, \mathcal{F})$.

Para isto, temos de verificar as três propriedades do Lema 2.5. A primeira e a terceira resultam diretamente da Definição 2.9. A segunda, i.e. a igualdade (2.15), é consequência do seguinte facto, um pouco mais geral.

Para todo $A \in \mathcal{F}_\Psi$, para todo $B \subseteq A^c$ e todo $E \subseteq X$,

$$\Psi\left(E \cap (A \cup B)\right) \geq \Psi\left(E \cap A\right) + \Psi\left(E \cap B\right). \tag{2.42}$$

Com efeito, como $A \in \mathcal{F}_\Psi$, tomando $E \cap (A \cup B)$ como conjunto "teste" em (2.40), temos

$$\Psi\left(E \cap (A \cup B)\right) \geq \Psi\left(E \cap (A \cup B) \cap A\right) + \Psi\left(E \cap (A \cup B) \cap A^c\right). \tag{2.43}$$

Claramente, $(A \cup B) \cap A = A$ e como $B \subseteq A^c$, então $(A \cup B) \cap A^c = B$. Substituindo estas igualdades em (2.43), obtemos (2.42).

***Observação.*** Unicamente assumimos que $A \in \mathcal{F}_\Psi$, mas não é necessário supor que $B \in \mathcal{F}_\Psi$.

Para demonstrar a primeira parte do Teorema 2.5, i.e. $\mathcal{F}_\Psi$ é uma álgebra$-\sigma$, vamos primeiro provar que $\mathcal{F}_\Psi$ é uma álgebra. Só no *Passo 3* é que vamos estabelecer a estabilidade de $\mathcal{F}_\Psi$ por união numerável crescente. Uma vez que hajamos mostrado que $\mathcal{F}_\Psi$ é uma álgebra$-\sigma$, a demonstração do Teorema 2.5 estará completa, pois, pelo *Passo 1*, a medida exterior $\Psi$, restringida a $\mathcal{F}_\Psi$ é automáticamente uma medida.

***Passo 2.*** $\mathcal{F}_\Psi$ é uma álgebra.

Claramente, os conjuntos $\emptyset$ e $X$ verificam (2.40), logo $\emptyset$ e $X$ pertencem a $\mathcal{F}_\Psi$. Ademais, como a Definição 2.9 é simétrica com respeito à complementação, temos que $A \in \mathcal{F}_\Psi$ se e só se $A^c \in \mathcal{F}_\Psi$.

Para mostrar que $\mathcal{F}_\Psi$ é estável por união finita, fixemos dois conjuntos $A, B \in \mathcal{F}_\Psi$. Segundo (2.40) (cf. Definição 2.10), temos de mostrar que

$$\Psi\left(E \cap (A \cup B)\right) + \Psi\left(E \cap (A \cup B)^c\right) \leq \Psi\left(E\right). \tag{2.44}$$

qualquer que seja $E \subseteq X$.

Como $B \in \mathcal{F}_\Psi$, aplicando (2.40) a $B$, com cada um dos subconjuntos $E \cap A$ e $E \cap A^c$ como conjuntos "teste", obtemos

$$\Psi\left(E \cap A \cap B\right) + \Psi\left(E \cap A \cap B^c\right) \leq \Psi\left(E \cap A\right) \tag{2.45}$$

e

$$\Psi\left(E\cap A^{\mathrm{c}}\cap B\right)+\Psi\left(E\cap A^{\mathrm{c}}\cap B^{\mathrm{c}}\right)\leq\Psi\left(E\cap A^{\mathrm{c}}\right). \tag{2.46}$$

Somando os lados correspondentes de (2.45) e (2.46) e aplicando (2.40) à soma do lado direito (lembre-se que $A\in\ \mathcal{F}_\Psi$), obtemos

$$\Psi\left(E\cap A\cap B\right)+\Psi\left(E\cap A\cap B^{\mathrm{c}}\right)+\Psi\left(E\cap A^{\mathrm{c}}\cap B\right)+\Psi\left(E\cap A^{\mathrm{c}}\cap B^{\mathrm{c}}\right)\leq\Psi\left(E\right). \tag{2.47}$$

Por (2.39), tendo em conta que $(A\cap B)\cup(A\cap B^{\mathrm{c}})\cup(A^{\mathrm{c}}\cap B)=A\cup B$, temos

$$\Psi\left(E\cap(A\cup B)\right)\leq\Psi\left(E\cap A\cap B\right)+\Psi\left(E\cap A\cap B^{\mathrm{c}}\right)+\Psi\left(E\cap A^{\mathrm{c}}\cap B\right).$$

Somando am ambos os lados $\Psi\left(E\cap A^{\mathrm{c}}\cap B^{\mathrm{c}}\right)$ e utilizando (2.47), obtemos (2.44).

***Passo 3.*** Para provar que $\mathcal{F}_\Psi$ é uma álgebra$-\sigma$, só falta estabelecer a alinha 4 da Definição 1.5. Para isto, fixamos uma família crescente $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{F}_\Psi$ e definimos

$$B_i=A_i\setminus A_{i-1},\ i\in\mathbb{N},\ A_0=\emptyset.$$

Como $\mathcal{F}_\Psi$ é uma álgebra (cf. *Passo 2*), então para todo $A,B\in\mathcal{F}_\Psi$, $A\cap B\in\mathcal{F}_\Psi$ e $A\setminus B\in\mathcal{F}_\Psi$. (cf. Exercício 6 do Capítulo 1.)

Portanto, para todo $i\in\mathbb{N}$, $B_i\in\mathcal{F}_\Psi$. Para além disso, os conjuntos $B_i$ são disjuntos dois a dois e

$$A_n=\bigcup_{i=1}^{n}B_i,\ \left(\forall\,n\in\mathbb{N}\right).$$

Aplicando a fórmula (2.42) do *Passo 1* inductivamente, qualquer que seja $E\subseteq X$, para todo $n\in\mathbb{N}$, temos

$$\Psi\left(E\cap A_n\right)=\Psi\left(E\cap\bigcup_{i=1}^{n}B_i\right)\geq\sum_{i=1}^{n}\Psi\left(E\cap B_i\right).$$

Mas como $A_n\in\mathcal{F}_\Psi$, por (2.40), temos

$$\Psi\left(E\right)\geq\Psi\left(E\cap A_n\right)+\Psi\left(E\cap A_n^{\mathrm{c}}\right)\geq\sum_{i=1}^{n}\Psi\left(E\cap B_i\right)+\Psi\left(E\cap A_n^{\mathrm{c}}\right).$$

Fazendo tender $n\to\infty$, como $A_n^{\mathrm{c}}\searrow\left(\bigcup_{n=1}^{\infty}A_n\right)^{\mathrm{c}}$, por (2.38),

$$\Psi\left(E\right)\geq\sum_{i=1}^{\infty}\Psi\left(E\cap B_i\right)+\Psi\left(E\cap\left(\bigcup_{n=1}^{\infty}A_n\right)^{\mathrm{c}}\right). \tag{2.48}$$

Como
$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{i=1}^{\infty} B_i,$$
por (2.39)
$$\Psi\left(E \cap \bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{i=1}^{\infty} \Psi\left(E \cap B_i\right).$$
Somando $\Psi\left(E \cap \left(\bigcup_{n=1}^{\infty} A_n\right)^{\mathrm{c}}\right)$ em ambos os lados e aplicando (2.48), obtemos
$$\Psi\left(E \cap \bigcup_{n=1}^{\infty} A_n\right) + \Psi\left(E \cap \left(\bigcup_{n=1}^{\infty} A_n\right)^{\mathrm{c}}\right) \leq \Psi\left(E\right).$$
Portanto, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_{\Psi}$. Aqui termina a demonstração do Teorema 2.5. :-)

***Observação.*** A notação "$\mathcal{F}_{\Psi}$" (semelhante às álgebras$-\sigma$), utilizada para o conjunto definido em (2.41), está justificada pelo Teorema 2.5. Mas note a inversão de papeis (de certo modo). Em quanto que as medidas são definidas sobre álgebras$-\sigma$, $\mathcal{F}_{\Psi}$ é definida em função de uma medida exterior e so após estabelecer o Teorema 2.5, temos as coisas direitas, uma medida sobre uma álgebra$-\sigma$.

## 2.6 Construção de medidas

A finalidade desta secção é provar a existência da extensão de que fala o Teorema 2.1. Para isto, vamos construir uma medida exterior, denotada $\mu^*$, que estende a premedida dada, digamos $\mu_0$, a todos os subconjuntos de $X$. De seguida restringiremos esta medida exterior à álgebra$-\sigma$ gerada por $\mathcal{F}_0$, não sem antes mostrar que os elementos de $\mathcal{F}_0$ são $\mu^*-$mensuráveis e portanto $\sigma\left[\mathcal{F}_0\right] \subseteq \mathcal{F}_{\mu^*}$. A conclusão final é consequência do Teorema 2.5.

**Notação.** Ao longo desta secção manteremos fixos um conjunto $X$, uma álgebra $\mathcal{F}_0$ sobre $X$ e uma premedida $\mu_0 : \mathcal{F}_0 \to \overline{\mathbb{R}}^+$, e referir-nos-emos a eles sem mais apresentações.

**Definição 2.11.** Para todo $A \subseteq X$ definimos
$$\mu^*\left(A\right) = \inf\left\{\sum_{i=1}^{\infty} \mu_0\left(A_i\right);\ A \subseteq \bigcup_{n=1}^{\infty} A_i \ \wedge\ \forall\, i \in \mathbb{N},\ A_i \in \mathcal{F}_0\right\}. \tag{2.49}$$

Para que este objeto seja a resposta ao nosso problema. i.e. Encontrar uma medida sobre $(X, \sigma\left[\mathcal{F}_0\right])$, que coincida com $\mu_0$ na álgebra $\mathcal{F}_0$, há vários pontos que devemos verificar: Primeiro, a função $\mu^*$, definida em (2.49), é de facto uma medida exterior. Segundo, os elementos de $\mathcal{F}_0$ são $\mu^*-$mensuráveis (cf. propriedade (2.40), Definição 2.10). Terceiro, $\mu^*$ e $\mu_0$ coincidem em $\mathcal{F}_0$. Vamos a desenvolver este programa em três lemas respetivamente.

**Lema 2.8.** $\mu^*$ é uma medida exterior.

***Demonstração.***

1. *Nulidade do vazio*: Como $\emptyset \in \mathcal{F}_0$ e $\mu_0(\emptyset) = 0$, por (2.49), $\mu^*(\emptyset) = 0$.
2. *Monotonia*: Se $A \subseteq B \subseteq X$, como
$$\left\{\sum_{i=1}^{\infty} \mu_0(A_i);\ A_i \in \mathcal{F}_0,\ B \subseteq \bigcup_{n=1}^{\infty} A_i\right\} \subseteq \left\{\sum_{i=1}^{\infty} \mu_0(A_i);\ A_i \in \mathcal{F}_0,\ A \subseteq \bigcup_{n=1}^{\infty} A_i\right\},$$
tomando o ínfimo temos $\mu^*(A) \leq \mu^*(B)$.
3. *Subaditividade numerável*: Seja $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{P}(X)$, temos de provar que
$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu^*(A_i). \tag{2.50}$$
Se existir $i \in \mathbb{N}$ tal que $\mu^*(A_i) = \infty$, a desigualdade anterior verifica-se trivialmente e não há mais nada que demonstrar.

   Caso contrário, fixamos $\epsilon > 0$, e aproximamos o valor de $\mu^*(A_i)$ (para cada $i \in \mathbb{N}$), da seguinte forma.

   Por (2.49), existe uma família $\left\{A_i^j\right\}_{j\in\mathbb{N}} \subseteq \mathcal{F}_0$ tal que
$$A_i \subseteq \bigcup_{j=1}^{\infty} A_i^j \quad \text{e} \quad \sum_{j=1}^{\infty} \mu_0\left(A_i^j\right) \leq \mu^*(A_i) + \frac{\epsilon}{2^i}. \tag{2.51}$$
Observe que $\left\{A_i^j\right\}_{i,j\in\mathbb{N}}$ é uma família numerável de elementos de $\mathcal{F}_0$ e
$$\bigcup_{i=1}^{\infty} A_i \subseteq \bigcup_{i,j=1}^{\infty} A_i^j.$$
Por (2.49) e (2.51),
$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i,j=1}^{\infty} \mu_0\left(A_i^j\right) \leq \sum_{i=1}^{\infty}\left(\mu^*(A_i) + \frac{\epsilon}{2^i}\right) = \epsilon + \sum_{i=1}^{\infty} \mu^*(A_i).$$
Sendo $\epsilon$ arbitrário, a desigualdade (2.50) fica estabelecida. :-)

***Observação.*** A hipótese (implícita) do Lema 2.8 é que $\mu_0$ é uma premedida. Mas a demonstração do lema emprega só a propriedade (2.19) (cf. Lema 2.6), na alinha 1. (nulidade do vazio).

**Lema 2.9.** $\mathcal{F}_0 \subseteq \mathcal{F}_{\mu^*}$

***Demonstração.*** Dado $A \in \mathcal{F}_0$, para provar que $A$ é $\mu^*-$mensurável (cf. (2.40), Definição 2.10), temos de provar que para todo $E \subseteq X$,

$$\mu^*(E \cap A) + \mu^*(E \cap A^c) \leq \mu^*(E) \tag{2.52}$$

Dado $E \subseteq X$, se $\mu^*(E) = \infty$, a desigualdade anterior verifica-se trivialmente e não há mais nada que demonstrar.

Se $\mu^*(E) < \infty$, por (2.49), para todo $\epsilon > 0$ fixo, existe uma família $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}_0$ tal que

$$E \subseteq \bigcup_{i=1}^{\infty} A_i \tag{2.53}$$

e

$$\sum_{i=1}^{\infty} \mu_0(A_i) \leq \mu^*(E) + \epsilon. \tag{2.54}$$

Tomando a interseção com $A$ em ambos os lados de (2.53), idem com $A^c$, obtemos

$$E \cap A \subseteq \bigcup_{i=1}^{\infty} A_i \cap A \quad \text{e} \quad E \cap A^c \subseteq \bigcup_{i=1}^{\infty} A_i \cap A^c. \tag{2.55}$$

Note que para todo $i \in \mathbb{N}$, $A_i \cap A$ e $A_i \cap A^c \in \mathcal{F}_0$. Então, por (2.49) e (2.55),

$$\mu^*(E \cap A) \leq \sum_{i=1}^{\infty} \mu_0(A_i \cap A) \quad \text{e} \quad \mu^*(E \cap A^c) \leq \sum_{i=1}^{\infty} \mu_0(A_i \cap A^c).$$

Somando os lados correspondentes destas duas desigualdades, temos

$$\mu^*(E \cap A) + \mu^*(E \cap A^c) \leq \sum_{i=1}^{\infty} \left( \mu_0(A_i \cap A) + \mu_0(A_i \cap A^c) \right). \tag{2.56}$$

Como $\mu_0$ é uma premedida, por (2.20),

$$\mu_0(A_i \cap A) + \mu_0(A_i \cap A^c) \leq \mu_0(A_i). \tag{2.57}$$

Substituindo em (2.56) e tendo em conta (2.54),

$$\mu^*(E \cap A) + \mu^*(E \cap A^c) \leq \sum_{i=1}^{\infty} \mu_0(A_i) \leq \mu^*(E) + \epsilon.$$

Sendo $\epsilon$ arbitrário, a desigualdade (2.52) fica demonstrada. :-)

***Observação.*** A demonstração do Lema 2.9 emprega unicamente a propriedade (2.20) das premedidas (cf. Lema 2.6), precisamente na equação (2.57).

**Lema 2.10.** $\mu^*|_{\mathcal{F}_0} = \mu_0$

***Demonstração.*** Se $A \in \mathcal{F}_0$, tomando o recubrimento $\{A, \emptyset, \emptyset, ...\}$, por (2.49) temos

$$\mu^*|_{\mathcal{F}_0}(A) = \mu^*(A) \leq \mu_0(A).$$

Por outro lado, se $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}_0$ e $A \subseteq \bigcup_{i=1}^{\infty} A_i$, então por (2.21),

$$\mu_0(A) \leq \sum_{i=1}^{\infty} \mu_0(A_i).$$

Tomando o ínfimo sobre todos os recubrimientos $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}_0$ de $A$,

$$\mu_0(A) \leq \mu^*(A) = \mu^*|_{\mathcal{F}_0}(A). :-)$$

***Observação.*** Por sua vez, a demonstração do Lema 2.10 emprega só a propriedade (2.21) (Lema 2.6).

Agora estamos prontos para demonstrar o Teorema 2.1, no que se refere à existência de uma medida que estende uma premedida dada. Recolhemos este resultado na proposição seguinte.

**Proposição 2.4.** (Teorema de Carathéodory: Existência) *Seja $\mathcal{F}_0$ uma álgebra sobre um conjunto $X$ e $\mu_0 : \mathcal{F}_0 \to \overline{\mathbb{R}}^+$ uma premedida sobre $X$. Existe uma medida sobre o espaço mensurável $(X, \sigma[\mathcal{F}_0])$, que denotaremos $\mu$, tal que*

$$\forall A \in \mathcal{F}_0,\ \mu(A) = \mu_0(A).$$

***Demonstração.*** Pelo Lema 2.8, $\mu^*$ é uma medida exterior. Pelo Teorema 2.5, $\mu^*$ é uma medida sobre $(X, \mathcal{F}_{\mu^*})$. Pelo Lema 2.9, $\mathcal{F}_0 \subseteq \mathcal{F}_{\mu^*}$. Logo, $\mu^*|_{\sigma[\mathcal{F}_0]}$ é uma medida sobre $(X, \sigma[\mathcal{F}_0])$. Finalmente, pelo Lema 2.10, $\mu^*|_{\mathcal{F}_0} = \mu_0$.

Portanto, basta tomar $\mu = \mu^*|_{\sigma[\mathcal{F}_0]}$. :-)

***Observação.*** Note que até o momento, ainda não foi necessária a hipótese do Teorema 2.1 da premedida ser $\sigma-$finita.

## 2.7 Famílias monótonas

Esta secção pode parecer isolada, pois vamos estudar uma estrutura de subconjuntos completamente nova, as "famílias monótonas." Vamos ver como esta estrutura se relaciona com as álgebras$-\sigma$. Utilizaremos este resultado para mostrar que a extensão obtida na Proposição 2.4 é única, sob a condição da premedida ser $\sigma-$finita.

**Definição 2.12.** Uma família $\mathcal{M} \subseteq \mathcal{P}(X)$, $\mathcal{M} \neq \emptyset$, é uma ***família monótona*** se ela verificar simultaneamente as seguintes condições:

1. ***União numerável crescente***: Se $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{M}$ é uma família crescente (i.e. $\forall\, i \in \mathbb{N}$, $A_i \subseteq A_{i+1}$), então
$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}. \tag{2.58}$$

2. ***Interseção numerável decrescente***: Se $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{M}$ é uma família decrescente (i.e. $\forall\, i \in \mathbb{N}$, $A_i \supseteq A_{i+1}$), então
$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{M}. \tag{2.59}$$

**Exemplo.** Seja $A \subseteq X$, $\{A\}$ é uma família monótona "minimal."

**Lema 2.11.** *Uma álgebra de conjuntos é uma álgebra$-\sigma$ se e só se ela é uma família monótona.*

***Demonstração.*** Por (1.2) (cf. Definição 1.5) e (1.3) (cf. Proposição 1.5), toda álgebra$-\sigma$ é uma família monótona. Inversamente, uma álgebra que verifica (2.58), é uma álgebra$-\sigma$ (cf. Definição 1.5).

**Proposição 2.5.** A interseção não vazia de famílias monótonas é uma família monótona. i.e. Sendo $\Phi$ um conjunto não vazio de famílias monótonas,
$$\mathcal{C} = \bigcap_{\mathcal{M}\in\Phi} \mathcal{M}$$
é uma família monótona.

***Demonstração.*** A demonstração desta proposição segue os mesmos passos da demonstração do Lema 1.2 e fica a cargo do leitor. (cf. Exercício 17.)

Uma consequência imediata do Proposição 2.5 é o seguinte corolário.

**Corolário 2.1.** Seja $\mathcal{S} \subseteq \mathcal{P}(X)$, uma família qualquer de subconjuntos de $X$. Existe a família monótona mais pequena que contém $\mathcal{S}$. i.e. Existe uma família monótona, que denotaremos $m[\mathcal{S}]$, tal que:

1. $\mathcal{S} \subseteq m[\mathcal{S}]$.
2. Para toda família monótona $\mathcal{M}$, se $\mathcal{S} \subseteq \mathcal{M}$, então $m[\mathcal{S}] \subseteq \mathcal{M}$.

**Definição 2.13.** Sendo $\mathcal{S} \subseteq \mathcal{P}(X)$, uma família qualquer de subconjuntos de $X$, a família monótona mais pequena que contém $\mathcal{S}$, cuja existência é garantida pelo Corolário 2.1, chama-se ***família monótona gerada por*** $\mathcal{S}$ e denota-se, como já dissemos, $m[\mathcal{S}]$.

O teorema seguinte é o resultado principal desta secção.

**Teorema 2.6.** *Seja $X$ um conjunto e $\mathcal{F}_0$ uma álgebra sobre $X$. Então*

$$m[\mathcal{F}_0] = \sigma[\mathcal{F}_0].$$

***Demonstração.*** Pela primeira parte do Lema 2.11, toda álgebra$-\sigma$ é uma família monótona. Então $\sigma[\mathcal{F}_0]$ é uma família monótona. Pela minimalidade de $m[\mathcal{F}_0]$,

$$m[\mathcal{F}_0] \subseteq \sigma[\mathcal{F}_0].$$

Para demonstrar a inclusão inversa, pela minimalidade de $\sigma[\mathcal{F}_0]$, basta provar que $m[\mathcal{F}_0]$ é uma álgebra$-\sigma$. Sendo $m[\mathcal{F}_0]$ uma família monótona, pela segunda parte do Lema 2.11, é suficiente mostrar que $m[\mathcal{F}_0]$ é uma álgebra.

Para isto, empregaremos um truque clássico e muito utilizado nas probabilidades e teoria da medida. Em lugar de tentar descrever este truque, o vamos aplicar sem mais comentários, neste exemplo particular.

Começaremos definindo três famílias de conjuntos e concluiremos a demonstração do Teorema 2.6, depois de provar que todas estas famílias são iguais.

Sendo

$$\mathcal{M}_0 = \{A \in m[\mathcal{F}_0]\,;\ A^{\mathrm{c}} \in m[\mathcal{F}_0]\}, \tag{2.60}$$

$$\mathcal{M}_1 = \{A \in m[\mathcal{F}_0]\,;\ \forall\, B \in \mathcal{F}_0,\ A \cup B \in m[\mathcal{F}_0]\} \text{ e} \tag{2.61}$$

$$\mathcal{M}_2 = \{A \in m[\mathcal{F}_0]\,;\ \forall\, B \in m[\mathcal{F}_0],\ A \cup B \in m[\mathcal{F}_0]\}, \tag{2.62}$$

queremos provar que

$$\mathcal{M}_0 = \mathcal{M}_1 = \mathcal{M}_2 = m[\mathcal{F}_0] \tag{2.63}$$

Primeiro observe-se que os conjuntos $\mathcal{M}_0$, $\mathcal{M}_1$ e $\mathcal{M}_2$ são todos famílias monótonas (cf. Exercício 18) e estão todos contidos em $m[\mathcal{F}_0]$.

Por outro lado, como $\mathcal{F}_0$ é uma álgebra, os elementos de $\mathcal{F}_0$ verificam as condições que definem $\mathcal{M}_0$ e $\mathcal{M}_1$. Logo, $\mathcal{F}_0 \subseteq \mathcal{M}_0$ e $\mathcal{F}_0 \subseteq \mathcal{M}_1$. Pela minimalidade de $m[\mathcal{F}_0]$, $\mathcal{M}_0 = \mathcal{M}_1 = m[\mathcal{F}_0]$. Em particular, visto que $\mathcal{M}_1 = m[\mathcal{F}_0]$, temos

$$\forall A \in \mathcal{F}_0, \forall B \in m[\mathcal{F}_0], \; A \cup B \in m[\mathcal{F}_0].$$

Logo, $\mathcal{F}_0 \subseteq \mathcal{M}_2$ e, pela minimalidade de $m[\mathcal{F}_0]$, $\mathcal{M}_2 = m[\mathcal{F}_0]$.

Agora, observe que $\mathcal{M}_0 \cap \mathcal{M}_2$ é uma álgebra. (cf. Exercício 15.) Como $\mathcal{M}_0 \cap \mathcal{M}_2 = m[\mathcal{F}_0]$, então $m[\mathcal{F}_0]$ é uma álgebra e para além disso verifica (2.58) por ser uma família monótoca. Pela Definição 1.5, $m[\mathcal{F}_0]$ é necessariamente uma álgebra$-\sigma$ (cf. Lema 2.11). Pela minimalidade de $\sigma[\mathcal{F}_0]$,

$$\sigma[\mathcal{F}_0] \subseteq m[\mathcal{F}_0]. :-)$$

Finalmente, estamos em posição de demonstrar a segunda parte do Teorema 2.1 ou seja a unicidade da extensão duma premedida $\sigma-$finita. (cf. Corolário 2.2.) Estabeleceremos este resultado numa versão mais conveviente. (cf. Proposição 2.6.)

**Proposição 2.6.** *Sejam $\mu_1$ e $\mu_2$ duas medidas sobre $(X, \sigma[\mathcal{F}_0])$, tal que*

$$\forall A \in \mathcal{F}_0, \; \mu_1(A) = \mu_2(A).$$

*Para além disso, supomos que existe uma sucessão crescente $\{X_j\}_{j\in\mathbb{N}} \subseteq \mathcal{F}_0$, tal que*

$$\forall j \in \mathbb{N}, \; \mu_1(X_j) = \mu_2(X_j) < \infty \quad \text{e} \quad \bigcup_{j=1}^{\infty} X_j = X. \tag{2.64}$$

*Então,*

$$\forall A \in \sigma[\mathcal{F}_0], \; \mu_1(A) = \mu_2(A). \tag{2.65}$$

***Demonstração.*** Para demonstrar (2.65), vamos empregar novamente o truque do Teorema 2.6, estudando a família dos "bem comportados."

Explicitamente, define-se a família

$$\mathcal{M} = \{A \in \sigma[\mathcal{F}_0]; \forall j \in \mathbb{N}, \mu_1(X_j \cap A) = \mu_2(X_j \cap A)\}. \tag{2.66}$$

Queremos mostrar que $\mathcal{M} = \sigma[\mathcal{F}_0]$.

Evidentemente, $\mathcal{F}_0 \subseteq \mathcal{M}$. Em particular, $\mathcal{M} \neq \emptyset$. Para além disso, $\mathcal{M}$ é estável por união crescente e interseção decrescente, como vamos ver a seguir.

*1. União numerável crescente*: Seja $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{M}$ uma família crescente. Por (2.66), para todo $i$, $j \in \mathbb{N}$,

$$\mu_1 (X_j \cap A_i) = \mu_2 (X_j \cap A_i) .$$

Tomando o limite quando $i \to \infty$, por (2.9) (cf. Lema 2.2), obtemos

$$\mu_1 \left( X_j \cap \bigcup_{i=1}^{\infty} A_i \right) = \lim_{i \to \infty} \mu_1 (X_j \cap A_i) = \lim_{i \to \infty} \mu_2 (X_j \cap A_i) = \mu_2 \left( X_j \cap \bigcup_{i=1}^{\infty} A_i \right) .$$

Como $j$ é arbitrário,

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}.$$

*2. União numerável decrescente*: Seja $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{M}$ uma família decrescente.

Por (2.66) e (2.64), para todo $i$, $j \in \mathbb{N}$,

$$\mu_1 (X_j \cap A_i) = \mu_2 (X_j \cap A_i) \leq \mu_1 (X_j) = \mu_2 (X_j) < \infty.$$

Tomando o limite quando $i \to \infty$, por (2.13) (cf. Lema 2.4),

$$\mu_1 \left( X_j \cap \bigcap_{i=1}^{\infty} A_i \right) = \lim_{i \to \infty} \mu_1 (X_j \cap A_i) = \lim_{i \to \infty} \mu_2 (X_j \cap A_i) = \mu_2 \left( X_j \cap \bigcap_{i=1}^{\infty} A_i \right) . \tag{2.67}$$

Como $j$ é arbitrário,

$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{M}.$$

Pelo Teorema 2.6, $\sigma [\mathcal{F}_0] = \mathcal{M}$. Conclui-se que para todo $A \in \sigma [\mathcal{F}_0]$, para todo $j \in \mathbb{N}$,

$$\mu_1 (X_j \cap A) = \mu_2 (X_j \cap A) .$$

Como $\{X_j\}_{j \in \mathbb{N}}$ é crescente, tomando o limite quando $j \to \infty$ e utilizando (2.9) (cf. Lema 2.2), obtemos

$$\mu_1 (A) = \mu_2 (A) . \ \ : -)$$

**Corolário 2.2.** (Teorema de Carathéodory: Unicidade) *Seja $\mathcal{F}_0$ uma álgebra sobre um conjunto $X$ e $\mu_0$ uma premedida $\sigma-$finita sobre $\mathcal{F}_0$. Então existe uma única medida sobre o espaço mensurável $(X, \sigma [\mathcal{F}_0])$, que denotaremos $\mu$, tal que*

$$\forall A \in \mathcal{F}_0, \ \mu (A) = \mu_0 (A) .$$

***Observação.*** No Teorema 2.1 (teorema de Carathéodory), supõe-se que a premedida é $\sigma-$finita. Mas o único sítio onde esta hipótese é utilizada é na obtenção de (2.67), pois estamos a aplicar o Lema 2.4. Para ver que esta hipótese é mesmo necessária, vamos dar um exemplo onde existem várias extensões duma mesma premedida.

**Exemplo 2.2.** Consideramos o conjunto $X = \mathbb{N} \cup \{\infty\}$ e denotamos Card a medida de contagem sobre o espaço $(X, \mathcal{P}(X))$. Claramente,

$$\mathrm{Card}(\{\infty\}) = 1.$$

Agora, definamos

$$A \subseteq X \to \mathrm{Kard}(A) = \mathrm{Card}(A \cap \mathbb{N}).$$

Kard é também uma medida sobre $(X, \mathcal{P}(X))$, mas

$$\mathrm{Kard}(\{\infty\}) = 0.$$

Ambas as medidas coinciden sobre a álgebra $\mathcal{F}_0$, gerada pela família

$$\mathcal{S} = \{A \subseteq \mathbb{N};\ \mathrm{Card}(A) < \infty\}$$

e, para além disso, $\sigma[\mathcal{S}] = \mathcal{P}(X)$. No entanto, isto não constitui uma violação da Proposição 2.6. (cf. Exercício 19.)

## 2.8 Medidas completas

### 2.8.1 Conjuntos de medida nula

**Definição 2.14.** Seja $(X, \mathcal{F}, \mu)$ um espaço de medida. $A \subseteq X$ é um ***conjunto de medida nula*** (ou ***conjunto de medida***$-\mu$ ***nula***), se existir um conjunto $B \in \mathcal{F}$, tal que

$$A \subseteq B \ \text{ e } \ \mu(B) = 0.$$

Um conjunto $A \subseteq X$ de medida$-\mu$ nula chama-se também $\mu-$***desprezável***.

***Observação.*** Os conjuntos de medida$-\mu$ nula não são necessáriamente mensuráveis.

**Exemplos**

(a) O conjunto $\mathbb{R} \setminus \{0\}$ é desprezável com respeito à medida de Dirac $\delta_0$.

(b) O intervalo aberto $(0, 1)$ é desprezável com respeito à medida de contagem de $\mathbb{Z}$ (como medida em $\mathbb{R}$). i.e., a medida $A \subseteq \mathbb{R} \to \mathrm{Card}\,(A \cap \mathbb{Z})$.

(c) Pelo Exercício 9, para todo $x \in \mathbb{R}$, $\{x\}$ é desprezável com respeito á medida de comprimento de $\mathbb{R}$.

(d) Pelo Lema 2.12 a seguir e a alinha anterior, qualquer subconjunto numerável de $\mathbb{R}$ é desprezável com respeito á medida de comprimento. Em particular, o conjunto $\mathbb{Q}$ (os números racionais). Além disso $\mathbb{Q} \in \mathcal{B}_{\mathbb{R}}$ (pelo Exercício 8 do Capítulo 1 e (1.2), Definição 1.5), portanto $\mathbb{Q}$ *tem* comprimento nulo. i.e. $\lambda\,(\mathbb{Q}) = 0$.

**Espaços de medida completos**

**Lema 2.12.** *A família*

$$\mathcal{F}_{\mu-\mathrm{nulo}} = \{A \subseteq X;\, A \text{ ou } A^{\mathrm{c}} \text{ é de medida } \mu - \text{nula}\}$$

*é uma álgebra$-\sigma$ sobre $X$. Em particular, a união numerável de conjuntos de medida nula é um conjunto de medida nula.*

***Demonstração.*** cf. Exercício 14.

**Definição 2.15.** Dizemos que $(X, \mathcal{F}, \mu)$ é um ***espaço de medida completo***, se $\mathcal{F}$ contem todos os conjuntos de medida$-\mu$ nula. i.e. $(X, \mathcal{F}, \mu)$ é completo, se

$$\mathcal{F}_{\mu-\mathrm{nulo}} \subseteq \mathcal{F}.$$

Também podemos dizer que $\mu$ é uma ***medida completa***, ou que a álgebra$-\sigma$ $\mathcal{F}$ é $\mu-$***completa***.

A completude das medidas é uma propriedade necessária, por exemplo, para garantir a mensurabilidade de certas variáveis aleatórias, como *tempos de paragem*, definidas em base a processos estocásticos. Mas dado um espaço de medida, $(X, \mathcal{F}, \mu)$, é sempre possível estender a medida $\mu$ a uma álgebra$-\sigma$ completa. Este é o conteúdo do próximo lema.

**Lema 2.13.** *Dado um espaço de medida $(X, \mathcal{F}, \mu)$, existe uma medida $\tilde{\mu}$, definida sobre uma álgebra$-\sigma$ $\tilde{\mathcal{F}} \supseteq \mathcal{F}$, tal que*

$$\forall\, A \in \mathcal{F},\ \tilde{\mu}\,(A) = \mu\,(A)$$

*e, além disso, $\tilde{\mu}$ é completa.*

***Demonstração.*** Para todo $A \subseteq X$ definimos

$$\mu^*(A) = \inf_{B \in \mathcal{F},\, A \subseteq B} \mu(B). \tag{2.68}$$

(Compare com a Definição 2.11.)

Pelo Lema 2.8, $\mu^*$ é uma medida exterior. Pelo Teorema 2.5, a família $\mathcal{F}_{\mu^*}$, dos conjuntos $\mu^*-$mensuráveis (cf. Definição 2.10), é uma álgebra$-\sigma$ sobre $X$ e $\mu^*|_{\mathcal{F}_{\mu^*}}$ é uma medida sobre $(X, \mathcal{F}_{\mu^*})$. Pela Proposição 2.3, o espaço $\left(X, \mathcal{F}_{\mu^*}, \mu^*|_{\mathcal{F}_{\mu^*}}\right)$ é completo.

Como $\mathcal{F} \subseteq \mathcal{F}_{\mu^*}$ (cf. Lema 2.9) e $\mu^*|_{\mathcal{F}} = \mu$ (cf. Lema 2.10), pondo $\tilde{\mathcal{F}} = \mathcal{F}_{\mu^*}$ e $\tilde{\mu} = \mu^*|_{\mathcal{F}_{\mu^*}}$, o lema fica demonstrado. :-)

***Observação.*** Pela Proposição 2.3 e o Teorema 2.5, a álgebra$-\sigma$ $\mathcal{F}_\Psi$ (cf. Definição 2.10) é $\Psi|_{\mathcal{F}_\Psi}-$completa.

**Proposição 2.7.** *Seja $\mu$ uma medida $\sigma-$finita sobre $(X, \mathcal{F})$, $\mu^*$ a medida exterior definida por* (2.68) *e $\mathcal{F}_{\mu^*}$ a álgebra$-\sigma$ dos conjuntos $\mu^*-$mensuráveis (cf. Definição 2.10). Então, $A \in \mathcal{F}_{\mu^*}$ se e só se existem $B$ e $C \in \mathcal{F}$ tais que $C \subseteq A \subseteq B$ e $\mu(B \setminus C) = 0$.*

***Demonstração.*** Se $\mu^*(A) < \infty$, existe uma sucessão $\{B_i\}_{i \in \mathbb{N}} \subset \mathcal{F}$ tal que para todo $i \in \mathbb{N}$,

$$A \subseteq B_i \quad \text{e} \quad \lim_{i \to \infty} \mu(B_i) = \mu^*(A).$$

Como $\mathcal{F}$ é uma álgebra$-\sigma$,

$$B = \bigcap_{i \in \mathbb{N}} B_i \in \mathcal{F}.$$

Claramente, $A \subseteq B$ e $\mu^*(A) = \mu(B)$.

Se $A \in \mathcal{F}_{\mu^*}$, então $B \setminus A \in \mathcal{F}_{\mu^*}$ e $\mu^*(B \setminus A) = 0$. Pelo raciocínio anterior, existe $D \in \mathcal{F}$, tal que $B \setminus A \subset D$ e $\mu(D) = 0$. Pondo $C = B \setminus D$, fica demonstrada a condição necessária, quando $\mu^*(A) < \infty$. No caso em que $\mu^*(A) = \infty$, consideramos uma sucessão crescente para $X$, de conjuntos de medida finita e aplicamos o resultado anterior. Os detalhes ficam a cargo do leitor (cf. Exercício 20).

Agora suponhamos que existem $B$ e $C \in \mathcal{F}$ tais que $C \subseteq A \subseteq B$ e $\mu(B \setminus C) = 0$. Seja $E \subseteq X$, pela monotonia e a subaditividade da medida exterior, e o facto de $B$ ser $\mu^*-$mensurável (já que $B \in \mathcal{F}$), temos

$$\mu^*(E \cap A) + \mu^*(E \setminus A) \leq \mu^*(E \cap B) + \mu^*(E \setminus C)$$

$$\leq \mu^*(E \cap B) + \mu^*(E \setminus B) + \mu^*(B \setminus C)$$
$$\leq \mu^*(E),$$

pois $\mu^*(B \setminus C) = \mu(B \setminus C) = 0$.

A demonstração da proposição está concluída. Note que para provar a condição suficiente, não é necessário $\mu$ ser $\sigma-$finita. :-)

**Corolário 2.3.** Com as notações da Proposição 2.7,

$$\mathcal{F}_{\mu^*} = \sigma\left[\mathcal{F} \cup \mathcal{F}_{\mu-\text{nulo}}\right].$$

Noutras palabras, o espaço $(X, \mathcal{F}, \mu)$ é completo se e só se $\mathcal{F}_{\mu^*} = \mathcal{F}$.

## 2.8.2 A medida de Lebesgue

Como exemplo dum espaço de medida completo, neste parágrafo vamos considerar o completado da medida de comprimento, $\lambda$, sobre $\left((0,1], \mathcal{B}_{(0,1]}\right)$ (cf. Teorema 2.2), ou sobre $(\mathbb{R}, \mathcal{B})$, para todos os efeitos.

**Definição 2.16.** Sendo $\lambda^*$ a medida exterior definida por (2.68) (com $\lambda$ no lugar de $\mu$), os conjuntos $\lambda^*-$mensuráveis (cf. Definição 2.10) chamam-se conjuntos mensuráveis à Lebesgue e $\mathcal{F}_{\lambda^*}$ chama-se ***álgebra*$-\sigma$ *de Lebesgue***.

É usual omitir o asterisco quando se trata da medida de Lebesgue e escrever $(\mathbb{R}, \mathcal{F}_\lambda, \lambda)$, já que a medida de Lebesgue é, na realidade, a medida de comprimento sobre a álgebra$-\sigma$ de Lebesgue e não sobre a álgebra$-\sigma$ de Borel. O espaço $(\mathbb{R}, \mathcal{F}_\lambda, \lambda)$ é um espaço de medida completo.

### Exemplo de um conjunto não mensurável

Utilizando o axioma de escolha, é fácil construir um conjunto que não seja mensurável à Lebesgue. Efetivamente, consideremos a relação de equivalência, definida em $[0,1] \times [0,1]$ por $a \sim b$ se $a - b \in \mathbb{Q}$. Seja $A$ um conjunto contendo exatamente um elemento de cada classe de equivalência. Os conjuntos $A_q := q + A \mod (1)$, $q \in \mathbb{Q} \cap [0,1)$, que são as translacões racionais de $A$ em $[0,1]$, i.e.

$$A_q := \{a + q;\, a \in A,\, a + q \in [0,1]\} \cup \{a + q - 1;\, a \in A,\, a + q \in (1,2]\},$$

são disjuntos dois a dois e a união de todos eles é igual a $R$. (cf. Exercício 21.) Como, para além disso, $Q \cap [0,1)$ é numerável, temos

$$\lambda([0,1]) = \sum_{q \in Q \cap [0,1)} \lambda(A_q),$$

isto é, se os $A_q$ forem mensuráveis à Lebesgue. Mas, $\lambda([0,1]) = 1$ e $\lambda$ é invariante por translação, i.e. $\lambda(A_q) = \lambda(A)$ (cf. Exercício 10). Portanto, se $A$ for mensurável, teriamos

$$1 = \sum_{q \in Q \cap [0,1)} \lambda(A) = \infty \cdot \lambda(A).$$

O qual é impossível.

Um problema aberto durante muito tempo foi o de construir subconjuntos da reta que não sejam mensuráveis, mas sem aplicacar o axioma de escolha. Em 1970, Robert M. Solovay construiu um modelo no qual verificam-se todos os axiomas da teoria de conjuntos de Zermelo–Fraenkel, exceto o axioma de escolha. Mas nesse modelo, todos os subconjuntos são mensuráveis à Lebesgue. Desta forma, Solovay mostrou que o axioma de escolha é necessário para provar a existência de um subconjunto que não seja mensurável à Lebesgue.

Para terminar esta discussão, note que por um argumento de cardinalidade, existem conjuntos mensuráveis que não são borelianos, já que o conjunto de Cantor tem medida nula e portanto todos os seus subconjuntos são mensuráveis. Mas por outro lado, o conjunto de Cantor tem a mesma cardinalidade que $\mathbb{R}$ e portanto o conjunto dos seus subconjuntos é de uma cardinalidade estritamente superior, o que não acontece com os Borelianos.

## Propriedades quase verdadeiras/quase falsas

Sendo $(X, \mathcal{F}, \mu)$ um espaço de medida, diz-se que uma propriedade é verificada em $\mu-$***quase nenhuma parte*** (ou ***quase nenhuma parte***) se a propriedade é verdadeira unicamente num conjunto de medida nula.

Diz-se que uma propriedade é verificada em $\mu-$***quase todo ponto*** (ou ***quase todo ponto***) se o conjunto dos pontos onde a propriedade é falsa é um conjunto de medida nula, ou seja, se a propriedade é falsa em quase nenhuma parte.

Num espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$, o complementar dum conjunto de medida nula é um ***conjunto de medida total***. i.e. $A \subseteq \Omega$ é um conjunto de medida total se e só se existe um evento $B \in \mathcal{F}$, tal que $B \subseteq A$ e $\mu(B) = 1$.

Num espaço de probabilidade, uma propriedade é ***quase segura*** se ela se verificar num conjunto de medida total. Dizemos também que a propriedade verifica-se ***quase seguramente***.

# Exercícios

1. Complete a demonstração da Proposição 2.1. Sendo $\mu$ uma medida sobre um espaço $(X, \mathcal{F})$, demonstre as seguintes propriedades.

   (a) *Aditividade para conjuntos disjuntos*: Sejam $A, B \in \mathcal{F}$, disjuntos, então
   $$\mu(A \cup B) = \mu(A) + \mu(B).$$

   (b) *Monotonía*: Sendo $A, B \in \mathcal{F}$. Se $A \subseteq B$, então
   $$\mu(A) \leq \mu(B).$$

   (c) *Subaditividade finita*: Para todo $A, B \in \mathcal{F}$, temos
   $$\mu(A \cup B) \leq \mu(A) + \mu(B).$$

2. Dê um exemplo de uma medida $\sigma-$finita tal que se restringirmos o domínio da medida a uma certa sub-álgebra$-\sigma$, a medida resultante deixe de ser $\sigma-$finita.

3. Mostre que a soma (numerável ou finita) de medidas é uma medida.

4. Prove que a família $\mathcal{F}_0$, dos conjuntos obtidos como uniões dum número finito de intervalos semiabertos $(a, b]$, $0 \leq a < b \leq 1$, disjuntos e separados, incluindo também o conjunto vazio, é uma álgebra sobre $(0, 1]$. i.e.
   $$\mathcal{F}_0 = \{(a_1, b_1] \cup ... \cup (a_n, b_n]\,; n \in \mathbb{N},\, 0 \leq a_1 < b_1 < ... < a_n < b_n \leq 1\} \cup \{\emptyset\}$$
   é uma álgebra sobre $(0, 1]$. (Ver o *Passo 1* da demonstração do Teorema 2.2.)

5. Sendo $\mathcal{F}_0$ a álgebra do exercício 4, mostre que para todo $A \in \mathcal{F}_0$, $A \neq \emptyset$, a representação de $A$ como união de intervalos disjuntos e separados é única.

6. Sendo $\mathcal{F}_0$ a álgebra do exercício 4 e $\lambda_0$ a função definida em (2.25), prove as duas afirmações seguintes.

   (a) *Monotonia*: Para todo $A, B \in \mathcal{F}_0$, se $A \subseteq B$, então
   $$\lambda_0(A) \leq \lambda_0(B).$$

   (b) *Subaditividade finita*: Para todo $A, B \in \mathcal{F}_0$,
   $$\lambda_0(A \cup B) \leq \lambda_0(A) + \lambda_0(B).$$

7. Sendo $A \subseteq (0,1]$, um conjunto aberto, mostre que

$$\forall\, a, b \in A,\ \ a \sim b \ \text{ se } \ [\min\{a,b\}\,, \max\{a,b\}] \subseteq A$$

define uma relação de equivalência em $A$ (Ver a definição(2.30), na demonstração do Lema 2.7.)

8. Mostre que as classes de equivalência, definidas no exercício 7, são intervalos abertos e disjuntos de $(0,1]$.

9. Mostre que sendo $\lambda$ a medida de Lebesgue sobre $(0,1]$, $\lambda(\{x\}) = 0$. Como consequência disto, prove que para todo $0 < a < b \leq 1$,

$$\lambda([a,b]) = \lambda((a,b)) = b - a.$$

(Pelo Exercício 8 do Capítulo 1, para todo $x \in (0,1]$, $\{x\} \in \mathcal{B}_{(0,1]}$.)

10. Sendo $a \in \mathbb{R}$ fixo, mostre que a função $x \to a + x$ deixa invariante a medida de Lebesgue $\lambda$ sobre $(\mathbb{R}, \mathcal{B})$.

11. Demonstre que a aplicação $\nu_{p,n}$, definida em (2.33), é uma probabilidade sobre $(\{0,1\}^n\,, \mathcal{P}(\{0,1\}^n))$.

12. Seja $\Psi$ uma medida exterior e $A \subseteq X$. Prove que $A$ é $\Psi-$mensurável se e só se

$$\forall E \subseteq X,\ \Psi(E \cap A) + \Psi(E \cap A^{\mathrm{c}}) = \Psi(E)\,.$$

13. Seja $\Psi$ uma medida exterior e $A \subseteq X$. Mostre que se $\Psi(A) = 0$, então $A \in \mathcal{F}_{\Psi}$.

14. Mostre que a família

$$\mathcal{F}_{\mu-\text{nulo}} = \{A \subseteq X;\ A \text{ ou } A^{\mathrm{c}} \text{ é de medida nula}\}$$

é uma álgebra$-\sigma$ sobre $X$.

15. Sendo $\mathcal{M}_0$ e $\mathcal{M}_2$ as famílias definidas por (2.60) e (2.62) respetivamente, prove que $\mathcal{M}_0 \cap \mathcal{M}_2$ é uma álgebra.

16. Utilizando as famílias monótonas, dê outra demonstração do *Passo 3* da demonstração do Teorema 2.5.

17. Demonstre que a interseção não vazia de famílias monótonas é uma família monótona. (cf. Proposição 2.5.)

18. Mostre que os conjuntos

$$\mathcal{M}_0 = \{A \in m[\mathcal{F}_0]\,;\ A^{\mathrm{c}} \in m[\mathcal{F}_0]\}$$

$$\mathcal{M}_1 = \{A \in m\,[\mathcal{F}_0]\,;\ \forall\, B \in \mathcal{F}_0,\ A \cup B \in m\,[\mathcal{F}_0]\}$$

$$\mathcal{M}_2 = \{A \in m\,[\mathcal{F}_0]\,;\ \forall\, B \in m\,[\mathcal{F}_0]\,,\ A \cup B \in m\,[\mathcal{F}_0]\}$$

definidos em (2.60), (2.61) e (2.62), respetivamente (demonstração do Teorema 2.6), são famílias monótonas.

19. Sendo $X = \mathbb{N} \cup \{\infty\}$, descreva a álgebra $\mathcal{F}_0$ gerada pela família

$$\mathcal{S} = \{A \subseteq \mathbb{N};\ \mathrm{Card}\,(A) < \infty\}\,.$$

Encontre a medida que se obtem pelo processo de Carathéodory (cf. Proposição 2.4), a partir da medida de contagem sobre $\mathcal{F}_0$ e compare com as duas medidas consideradas no Exemplo 2.2. Explique por qué a extensão não é única e compare com o que acontece no Corolário 2.2.

20. (cf. Proposição 2.7.) Sendo $\mu$ uma medida $\sigma-$finita sobre um espaço mensurável $(X, \mathcal{F})$ e $\mu^*$ a medida exterior definida por (2.68), mostre que se $A \in \mathcal{F}_{\mu^*}$ (cf. Definição 2.10), então, existem $B$ e $C \in \mathcal{F}$ tais que $C \subseteq A \subseteq B$ e $\mu(B \setminus C) = 0$. (Considerar o caso $\mu^*(A) = \infty$.)

21. Mostre que a relação $a \sim b$, se $a - b \in \mathbb{Q}$, é uma relação de equivalência em $\mathbb{R} \times \mathbb{R}$.

# Capítulo 3

# Os Borelianos e as medidas de Borel em $(\mathbb{R}, \mathcal{B})$

Neste capítulo, vamos profundizar um pouco o estudo das funções mensuráveis em geral e das funções mensuráveis, reais em particular.

## 3.1 Reta real estendida

Designamos por $\overline{\mathbb{R}}$ a compactificação de $\mathbb{R}$ por dois pontos, a qual consiste, intuitivamente falando, em aumentar um ponto no "extremo-direito", denotado $+\infty$, e outro no "extremo-esquerdo", denotado $-\infty$. Nesta secção vamos descrever a relação de ordem, a topologia, as operações aritméticas e a álgebra$-\sigma$ de Borel deste espaço topológico, como preparação para o estudo das funções mensuráveis que tomam valores nele.

#### Relação de ordem na reta real estendida

Lembre-se que uma relação de ordem é uma relação reflexiva, transitiva e antisimétrica (cf. 1.2.3). A ordem sobre o conjunto $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ é a ordem usual de $\mathbb{R}$, junto com as relações:

$$\forall\, a \in \overline{\mathbb{R}},\ -\infty \leq a \leq +\infty. \tag{3.1}$$

Por este motivo, $[-\infty, +\infty]$ é outra forma de denotar $\overline{\mathbb{R}}$.

### Topologia da reta real estendida

Dotamos $\overline{\mathbb{R}}$ da topologia usual de $\mathbb{R}$, aumentando as vizinhanças de $+\infty$ e $-\infty$, que são os subconjuntos que contêm uma semireta. Recolhemos isto na seguinte definição.

**Definição 3.1.**

(a) Para todo $x \in \mathbb{R}$ e $E \subset \overline{\mathbb{R}}$, $E$ é uma vizinhança de $x$ se existe um intervalo aberto $(a, b) \subset \mathbb{R}$ tal que $x \in (a, b) \subset E$.

(b) $E \subset \overline{\mathbb{R}}$ é uma vizinhança de $+\infty$ (resp. $-\infty$) se $+\infty \in E$ (resp. $-\infty \in E$) e, para algúm $a \in \mathbb{R}$, $(a, +\infty) \subset E$ (resp. $(-\infty, a) \subset E$).

Lembre-se que um subconjunto dum espaço topológico é aberto se (por definição) é vizinhança de todos os seus pontos. Note que se $A \subseteq \mathbb{R}$ é aberto em $\mathbb{R}$, então $A$ é aberto em $\overline{\mathbb{R}}$. Em particular, $\mathbb{R}$ é aberto em $\overline{\mathbb{R}}$. Inversamente, se $A \subseteq \overline{\mathbb{R}}$ é aberto em $\overline{\mathbb{R}}$, então $A \cap \mathbb{R}$ é aberto em $\mathbb{R}$. (cf. Exercício 1.)

### Compacidade

**Lema 3.1.** *$\overline{\mathbb{R}}$ é uma compactificação de $\mathbb{R}$. Quer dizer, $\overline{\mathbb{R}}$ é compacto e $\mathbb{R}$ é denso em $\overline{\mathbb{R}}$ (i.e. a clausura de $\mathbb{R}$ em $\overline{\mathbb{R}}$ é $\overline{\mathbb{R}}$).*

***Demonstração.*** A demonstração deste lema é um exercício clássico em topologia e deixamos ao leitor. (cf. Exercício 2.)

### Operações aritméticas

**Definição 3.2.** As operações aritméticas entre elementos de $\mathbb{R}$ são as mesmas. As regras quando intervém $+\infty$ e $-\infty$ são as seguintes:

1. ***Nulidade***: $\forall\, a \in \overline{\mathbb{R}},\ a \cdot 0 = 0$. (Em particular, $\pm\infty \cdot 0 = 0$.)
2. ***Produto***: $\forall\, a \in \overline{\mathbb{R}}$, $a \neq 0,\ a \cdot (\pm\infty) = \begin{cases} \pm\infty & \text{se } a > 0 \\ \mp\infty & \text{se } a < 0 \end{cases}$.
3. ***Soma***: $\forall\, a \in \mathbb{R},\ a + (\pm\infty) = \pm\infty$.
4. ***Soma de infinitos do mesmo sinal***: $(\pm\infty) + (\pm\infty) = \pm\infty$.
5. ***Indefinido***: A soma de infinitos de sinal oposto, $(\pm\infty) + (\mp\infty)$, não está definida.

### 3.1.1 Álgebra$-\sigma$ de Borel

**Nota 3.1.** Para descrever $\mathcal{B}_{\overline{\mathbb{R}}}$ (a álgebra$-\sigma$ de Borel de $\overline{\mathbb{R}}$), primeiro observe que os conjuntos unitários $\{+\infty\}$ e $\{-\infty\}$ são borelianos. (A demonstração é semelhante à do exercício 8 do Capítulo 1.) Portanto, qualquer que seja $A \subseteq \overline{\mathbb{R}}$, $A \setminus \mathbb{R} \in \mathcal{B}_{\overline{\mathbb{R}}}$, já que tanto $\{+\infty\}$, como $\{-\infty\}$ e $\{+\infty, -\infty\} \in \mathcal{B}_{\overline{\mathbb{R}}}$.

Isto sugere que podemos expressar $\mathcal{B}_{\overline{\mathbb{R}}}$ em termos de $\mathcal{B}_{\mathbb{R}}$ (a álgebra$-\sigma$ de Borel de $\mathbb{R}$) e vice-versa, como vemos no seguinte resultado.

**Lema 3.2.**
$$\forall\, A \subseteq \overline{\mathbb{R}},\ A \in \mathcal{B}_{\overline{\mathbb{R}}} \Leftrightarrow A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}.$$

***Demonstração.*** A família

$$\left\{A \subseteq \overline{\mathbb{R}};\ A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\right\}$$

é uma álgebra$-\sigma$ de $\overline{\mathbb{R}}$ e contém todos os abertos de $\overline{\mathbb{R}}$. (cf. Exercício 3.)

Então
$$\mathcal{B}_{\overline{\mathbb{R}}} \subseteq \left\{A \subseteq \overline{\mathbb{R}};\ A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\right\}.$$

Isto mostra a implicação direta: $A \in \mathcal{B}_{\overline{\mathbb{R}}} \Rightarrow A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}$.

Por outro lado, a restrição de $\mathcal{B}_{\overline{\mathbb{R}}}$ a $\mathbb{R}$, i.e.

$$\mathcal{B}_{\overline{\mathbb{R}}}|_{\mathbb{R}} = \{A \cap \mathbb{R};\ A \in \mathcal{B}_{\overline{\mathbb{R}}}\}$$

é uma álgebra$-\sigma$ de $\mathbb{R}$, que contém todos os abertos de $\mathbb{R}$. (Ver a Definição 1.16 e a *Observação* a seguir.)

Então, pela minimalidade de $\mathcal{B}_{\mathbb{R}}$ e como $\mathbb{R} \in \mathcal{B}_{\overline{\mathbb{R}}}$,

$$\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{B}_{\overline{\mathbb{R}}}|_{\mathbb{R}} \subseteq \mathcal{B}_{\overline{\mathbb{R}}}. \tag{3.2}$$

Para concluir, suponhamos que $A \subseteq \overline{\mathbb{R}}$ e que $A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}$.

Como, por (3.2), $A \cap \mathbb{R} \in \mathcal{B}_{\overline{\mathbb{R}}}$ e, pela Nota 3.1, $A \setminus \mathbb{R} \in \mathcal{B}_{\overline{\mathbb{R}}}$,

$$A = (A \cap \mathbb{R}) \cup (A \setminus \mathbb{R}) \in \mathcal{B}_{\overline{\mathbb{R}}}.$$

Isto mostra a implicação inversa: $A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}} \Rightarrow A \in \mathcal{B}_{\overline{\mathbb{R}}}$. :-)

### Famílias geradoras dos Borelianos de $\mathbb{R}$

**Lema 3.3.** *As álgebras–$\sigma$ geradas por*

$$\begin{aligned} \mathcal{S}_1 &= \{(-\infty, a]\,;\, a \in \mathbb{R}\} & \mathcal{S}_2 &= \{(a, \infty)\,;\, a \in \mathbb{R}\} \\ \mathcal{S}_3 &= \{(-\infty, a)\,;\, a \in \mathbb{R}\} & \mathcal{S}_4 &= \{[a, \infty)\,;\, a \in \mathbb{R}\} \\ \mathcal{S}_5 &= \{(a, b)\,;\, a, b \in \mathbb{R}\} & \mathcal{S}_6 &= \{[a, b]\,;\, a, b \in \mathbb{R}\}\,, \end{aligned} \tag{3.3}$$

*são todas iguais a $\mathcal{B}_{\mathbb{R}}$. i.e.*

$$\sigma[\mathcal{S}_1] = \sigma[\mathcal{S}_2] = \sigma[\mathcal{S}_3] = \sigma[\mathcal{S}_4] = \sigma[\mathcal{S}_5] = \sigma[\mathcal{S}_6] = \mathcal{B}_{\mathbb{R}}. \tag{3.4}$$

### Famílias geradoras dos Borelianos de $\overline{\mathbb{R}}$

**Lema 3.4.** *As álgebras–$\sigma$ geradas por*

$$\begin{aligned} \overline{\mathcal{S}}_1 &= \{[-\infty, a]\,;\, a \in \mathbb{R}\} & \overline{\mathcal{S}}_2 &= \{(a, \infty]\,;\, a \in \mathbb{R}\} \\ \overline{\mathcal{S}}_3 &= \{[-\infty, a)\,;\, a \in \mathbb{R}\} & \overline{\mathcal{S}}_4 &= \{[a, \infty]\,;\, a \in \mathbb{R}\} \\ \overline{\mathcal{S}}_5 &= \{(a, b)\,;\, a, b \in \mathbb{R}\} \cup \{\{\infty\}\} & \overline{\mathcal{S}}_6 &= \left\{[a, b]\,;\, a, b \in \overline{\mathbb{R}}\right\} \end{aligned} \tag{3.5}$$

*são todas iguais a $\mathcal{B}_{\overline{\mathbb{R}}}$. i.e.*

$$\sigma\left[\,\overline{\mathcal{S}}_1\,\right] = \sigma\left[\,\overline{\mathcal{S}}_2\,\right] = \sigma\left[\,\overline{\mathcal{S}}_3\,\right] = \sigma\left[\,\overline{\mathcal{S}}_4\,\right] = \sigma\left[\,\overline{\mathcal{S}}_5\,\right] = \sigma\left[\,\overline{\mathcal{S}}_6\,\right] = \mathcal{B}_{\overline{\mathbb{R}}}.$$

***Demonstração.*** Demonstraremos unicamente este último lema, sendo a demonstração do anterior praticamente idéntica. (cf. Exercício 4.)

Por definição, $I \in \overline{\mathcal{S}}_1 \Leftrightarrow I^{\mathrm{c}} \in \overline{\mathcal{S}}_2$, então $\sigma\left[\,\overline{\mathcal{S}}_1\,\right] = \sigma\left[\,\overline{\mathcal{S}}_2\,\right]$. Pela mesma razão, $\sigma\left[\,\overline{\mathcal{S}}_3\,\right] = \sigma\left[\,\overline{\mathcal{S}}_4\,\right]$.

Ademais observe que para todo $a \in \mathbb{R}$,

$$[-\infty, a) = \bigcup_{n=1}^{\infty} \left[-\infty, a - \frac{1}{n}\right] \in \sigma\left[\,\overline{\mathcal{S}}_1\,\right].$$

Portanto, $\overline{\mathcal{S}}_3 \subseteq \sigma\left[\,\overline{\mathcal{S}}_1\,\right]$. Por outro lado, $\overline{\mathcal{S}}_1 \subseteq \sigma\left[\,\overline{\mathcal{S}}_3\,\right]$, já que

$$[-\infty, a] = \bigcap_{n=1}^{\infty}\left[-\infty, a + \frac{1}{n}\right) \in \sigma\left[\,\overline{\mathcal{S}}_3\,\right].$$

Isto mostra a igualdade entre as quatro primeiras álgebras$-\sigma$.

Continuando, como $\overline{\mathcal{S}}_1 \subseteq \overline{\mathcal{S}}_6$ e

$$[a,b] = [a,+\infty] \cap [-\infty, b] \in \sigma\left[\,\overline{\mathcal{S}}_1 \cup \overline{\mathcal{S}}_4\,\right] = \sigma\left[\,\overline{\mathcal{S}}_1\,\right],$$

então $\sigma\left[\,\overline{\mathcal{S}}_1\,\right] = \sigma\left[\,\overline{\mathcal{S}}_6\,\right]$.

Finalmente, observe que

$$(a,+\infty] = \bigcup_{n=1}^{\infty}(a,n) \cup \{+\infty\} \in \sigma\left[\,\overline{\mathcal{S}}_5\right],$$

e

$$(a,b) = (a,+\infty] \cap [-\infty, b) \in \sigma\left[\,\overline{\mathcal{S}}_2 \cup \overline{\mathcal{S}}_3\,\right] = \sigma\left[\,\overline{\mathcal{S}}_2\,\right]$$

e ainda

$$\{+\infty\} = \bigcap_{n=1}^{\infty}(n,+\infty) \in \sigma\left[\,\overline{\mathcal{S}}_2\right].$$

Então, $\sigma\left[\,\overline{\mathcal{S}}_3\,\right] = \sigma\left[\,\overline{\mathcal{S}}_5\,\right]$. Com isto terminamos a demonstração da igualdade das seis álgebras$-\sigma$.

Falta-nos demonstrar que todas elas são iguais a $\mathcal{B}_{\overline{\mathbb{R}}}$. Para isto, denotamos $\overline{\mathcal{S}}_7$ o conjunto de todos os intervalos abertos de $\overline{\mathbb{R}}$, i.e.

$$\overline{\mathcal{S}}_7 = \mathcal{S}_2 \cup \overline{\mathcal{S}}_2 \cup \mathcal{S}_3 \cup \overline{\mathcal{S}}_3 \cup \mathcal{S}_5$$

e denotamos $\overline{\mathcal{S}}_8$ o conjunto de todos os abertos de $\overline{\mathbb{R}}$.

Como $\mathbb{Q}$ é numerável e denso em $\overline{\mathbb{R}}$, todo aberto de $\overline{\mathbb{R}}$ é união numerável de intervalos abertos disjuntos. (A demonstração é a mesma que a do Lema 2.7.) Logo,

$$\overline{\mathcal{S}}_8 \subseteq \sigma\left[\,\overline{\mathcal{S}}_7\,\right]. \quad :-)$$

**Notação.** Para simplificar a notação e devido à semelhança entre $\mathcal{B}_{\mathbb{R}}$ e $\mathcal{B}_{\overline{\mathbb{R}}}$, estas duas álgebras$-\sigma$ denotam-se ambas por $\mathcal{B}$, sendo geralmente claro no contexto a qual delas referimo-nos. As medidas sobre $(\mathbb{R}, \mathcal{B})$ chamam-se ***medidas de Borel*** ou ***medidas Borelianas***.

## 3.2 Mais sobre as funções mensuráveis

**Composição de funções mensuráveis**

**Lema 3.5.** *Sejam $(X, \mathcal{F})$, $(Y, \mathcal{G})$ e $(Z, \mathcal{H})$ três espaços mensuráveis. Consideramos duas funções $f : X \to Y$ e $g : Y \to Z$, a primeira $\mathcal{F}/\mathcal{G}-$mensurável e a segunda $\mathcal{G}/\mathcal{H}-$mensurável. Então*

$$g \circ f : X \to Z \text{ é } \mathcal{F}/\mathcal{H}-\text{mensurável.}$$

***Demonstração.*** Seja $A \in \mathcal{H}$. Pela Definição 1.14, $g^{-1}(A) \in \mathcal{G}$, por $g : Y \to Z$ ser $\mathcal{G}/\mathcal{H}-$mensurável. Como $f : X \to Y$ é $\mathcal{F}/\mathcal{G}-$mensurável, então

$$(g \circ f)^{-1}(A) = f^{-1}\left(g^{-1}(A)\right) \in \mathcal{F}. \ :-)$$

**Caracterização das funções mensuráveis**

**Lema 3.6.** *Sejam $(X, \mathcal{F})$ e $(Y, \mathcal{G})$ dois espaços mensuráveis e seja $\mathcal{S} \subseteq \mathcal{G}$ uma família geradora de $\mathcal{G}$, i.e. $\sigma[\mathcal{S}] = \mathcal{G}$.*

*Então, $f : X \to Y$ é $\mathcal{F}/\mathcal{G}-$mesurável se e só se*

$$\forall A \in \mathcal{S},\ f^{-1}(A) \in \mathcal{F}.$$

***Demonstração.*** O Lema 3.6 é uma reformulação, simplificada, do Lema 1.4, para facilitar a sua aplicabilidade. (cf. Exercício 5.)

### 3.2.1 Mensurabilidade das funções contínuas

Sejam $X$ e $Y$ dois espaços topológicos. Denotamos por $\mathcal{O}_X$ e $\mathcal{O}_Y$ as famílias de abertos de $X$ e $Y$, respetivamente.

**Lema 3.7.** *Toda função contínua $f : X \to Y$ é $\mathcal{B}_X/\mathcal{B}_Y-$mensurável.* (Sendo $\mathcal{B}_X$ e $\mathcal{B}_Y$ as álgebras$-\sigma$ de Borel de $X$ e $Y$, respetivamente.)

***Demonstração.*** Pela caracterização das funções contínuas, $f : X \to Y$ é contínua se e só se

$$\forall A \in \mathcal{O}_Y,\ f^{-1}(A) \in \mathcal{O}_X.$$

Como $\mathcal{O}_Y$ gera $\mathcal{B}_Y$ e $\mathcal{O}_X \subseteq \mathcal{B}_X$, pelo Lema 3.6, $f$ é $\mathcal{B}_X/\mathcal{B}_Y-$mensurável. :-)

***Observação.*** Em particular, toda função contínua $f : D \subseteq \mathbb{R} \to \mathbb{R}$ é mensurável.

### 3.2.2 Caracterização das funções mensuráveis reais

**Lema 3.8.** *Seja $(X, \mathcal{F})$ um espaço mensurável. Uma função $f : X \to \overline{\mathbb{R}}$ é mensurável se e só se se verificar uma das seguintes afirmações (equivalentes).*

(a) $\forall\, a \in \mathbb{R},\ \{x \in X;\ f(x) \leq a\} \in \mathcal{F}$.

(b) $\forall\, a \in \mathbb{R},\ \{x \in X;\ f(x) < a\} \in \mathcal{F}$.

(c) $\forall\, a \in \mathbb{R},\ \{x \in X;\ f(x) \geq a\} \in \mathcal{F}$.

(d) $\forall\, a \in \mathbb{R},\ \{x \in X;\ f(x) > a\} \in \mathcal{F}$.

***Demonstração.*** Este lema é uma consequência imediata do Lema 3.4 e o Lema 3.6, já que (por exemplo)

$$\{x \in X;\ f(x) \leq a\} = f^{-1}([-\infty, a]).$$

#### Máximo e mínimo de funções mensuráveis

**Lema 3.9.** *Seja $(X, \mathcal{F})$ um espaço mensurável e sejam $f$ e $g$ duas funções de $X$ em $\overline{\mathbb{R}}$, $\mathcal{F}/\mathcal{B}-$mensuráveis, definidas. Então, as seguintes funções são mensuráveis.*

$$(f \wedge g)(x) = \min\{f(x), g(x)\} \quad e \quad (f \vee g)(x) = \max\{f(x), g(x)\}.$$

***Demonstração.*** Por pelo Lema 3.8, basta observar que

$$\{x \in X;\ (f \wedge g)(x) \leq a\} = \{x \in X;\ f(x) \leq a\} \cup \{x \in X;\ g(x) \leq a\} \in \mathcal{F}$$

e

$$\{x \in X;\ (f \vee g)(x) \leq a\} = \{x \in X;\ f(x) \leq a\} \cap \{x \in X;\ g(x) \leq a\} \in \mathcal{F}.\ :-)$$

**Corolário 3.1.** Se $f$ é mensurável, as seguintes funções são mensuráveis.

$$f^+ = f \vee 0 \ \text{ e } \ f^- = (-f) \vee 0. \tag{3.6}$$

**Nomenclatura.** Dizemos que uma função $f : X \to \overline{\mathbb{R}}$ é positiva se $f(x) \geq 0$, para todo $x \in X$. Se para todo $x \in X$, $f(x) > 0$, diremos que $f$ é estrítamente positiva. (Na literatura em Inglés estas funções chamam-se "nonnegative" e "positive", respetivamente.)

**Nota 3.2.** Observe que toda função mensurável é diferença de duas funções mensuráveis **positivas**. De facto,

$$f = f^+ - f^-.$$

### Limites de funções mensuráveis

**Teorema 3.1.** *Seja* $(X, \mathcal{F})$ *um espaço mensurável e* $f_i : X \to \overline{\mathbb{R}}$, $i \in \mathbb{N}$, *uma sucessão de funções mensuráveis. Então, as seguintes funções são mensuráveis.*

$$x \in X \to \left(\sup_{i\in\mathbb{N}} f_i\right)(x) = \sup\left\{f_i(x);\ i \in \mathbb{N}\right\}$$

*e*

$$x \in X \to \left(\inf_{i\in\mathbb{N}} f_i\right)(x) = \inf\left\{f_i(x);\ i \in \mathbb{N}\right\}.$$

***Demonstração.*** Este teorema demonstra-se de forma semelhante ao Lema 3.9. (cf. Exercício 6.)

**Corolário 3.2.** Seja $\{f_i\}_{i\in\mathbb{N}}$ uma sucessão de funções mensuráveis. Então, as seguintes funções são mensuráveis.

$$\limsup_{i\to\infty} f_i = \lim_{i\to\infty}\left(\sup_{j\geq i} f_j\right) = \inf_{i\in\mathbb{N}}\left(\sup_{j\geq i} f_j\right)$$

e

$$\liminf_{i\to\infty} f_i = \lim_{i\to\infty}\left(\inf_{j\geq i} f_j\right) = \sup_{i\in\mathbb{N}}\left(\inf_{j\geq i} f_j\right).$$

Em particular, se para todo $x \in X$, $\lim_{i\to\infty} f_i\,(x)$ existe, então a função

$$x \in X \to \lim_{i\to\infty} f_i\,(x) \tag{3.7}$$

é mensurável.

### 3.2.3 Álgebra$-\sigma$ produto

**Definição 3.3.** Dados dois espaços mensuráveis, $(X, \mathcal{F})$ e $(Y, \mathcal{G})$, define-se a ***álgebra***$-\sigma$ ***produto***, denotada $\mathcal{F} \otimes \mathcal{G}$, como sendo a álgebra$-\sigma$ gerada por

$$\mathcal{R}_{X \times Y} = \{A \times B \subseteq X \times Y;\, A \in \mathcal{F},\, B \in \mathcal{G}\}\,. \tag{3.8}$$

Ou seja,

$$\mathcal{F} \otimes \mathcal{G} = \sigma\left[\mathcal{R}_{X \times Y}\right]. \tag{3.9}$$

***Observação.*** Utilizamos a notação do produto tensorial, $\otimes$, devido a que $\mathcal{F} \otimes \mathcal{G}$ é de facto um produto tensorial na categoria das álgebras$-\sigma$ com os functores das preimagens das funções mensuráveis. Por simplicidade da exposição, aqui não exploraremos este "ângulo". Mas a notação é útil na mesma, para não confundir com o produto Cartesiano.

**Definição 3.4.** O ***espaço (mensurável) produto***,

$$(X \times Y, \mathcal{F} \otimes \mathcal{G})\,,$$

é o produto Cartesiano $X \times Y$, dotado da ***álgebra***$-\sigma$ ***produto*** $\mathcal{F} \otimes \mathcal{G}$.

#### Mensurabilidade das projeções

**Proposição 3.1.** As projeções $p_X : (x, y) \to x$ e $p_Y : (x, y) \to y$ são mensuráveis. Aliás, $\mathcal{F} \otimes \mathcal{G}$ é a mais pequena álgebra$-\sigma$ tal que as duas projeções são mensuráveis. i.e.

$$\mathcal{F} \otimes \mathcal{G} = \sigma\left[p_X, p_Y\right]$$

onde $p_X(x, y) = x$, $p_Y(x, y) = y$ e $\sigma\left[p_X, p_Y\right] = \sigma\left[\sigma\left[p_X\right] \cup \sigma\left[p_Y\right]\right]$.

***Demonstração.*** Note que

$$\sigma\left[p_X\right] = \{A \times Y;\, A \in \mathcal{F}\} \subset \mathcal{R}_{X \times Y} \subset \mathcal{F} \otimes \mathcal{G}.$$

Idem, $\sigma\left[p_Y\right] = \{X \times B;\, B \in \mathcal{G}\} \subset \mathcal{R}_{X \times Y} \subset \mathcal{F} \otimes \mathcal{G}$. Então

$$\sigma\left[p_X\right] \cup \sigma\left[p_Y\right] \subset \mathcal{F} \otimes \mathcal{G}.$$

Pela minimalidade de $\sigma\left[p_X, p_Y\right]$,

$$\sigma\left[p_X, p_Y\right] = \sigma\left[\sigma\left[p_X\right] \cup \sigma\left[p_Y\right]\right] \subseteq \mathcal{F} \otimes \mathcal{G}.$$

Inversamente, para todo $A \in \mathcal{F}$ e todo $B \in \mathcal{G}$,

$$A \times B = A \times Y \cap X \times B \in \sigma\left[\sigma\left[p_X\right] \cup \sigma\left[p_Y\right]\right] = \sigma\left[p_X, p_Y\right].$$

Pela minimalidade de $\mathcal{F} \otimes \mathcal{G}$,

$$\mathcal{F} \otimes \mathcal{G} \subseteq \sigma\left[p_X, p_Y\right]. \;\; :-)$$

**Associatividade do produto de álgebras**$-\sigma$

**Definição 3.5.** Sejam $(X, \mathcal{F})$, $(Y, \mathcal{G})$ e $(Z, \mathcal{H})$ três espaços mensuráveis. Define-se

$$\mathcal{F} \otimes \mathcal{G} \otimes \mathcal{H} = \sigma\left[\mathcal{R}_{X \times Y \times Z}\right],$$

onde

$$\mathcal{R}_{X \times Y \times Z} = \{A \times B \times C;\, A \in \mathcal{F},\, B \in \mathcal{G},\, C \in \mathcal{H}\}.$$

**Lema 3.10.** *Sejam $(X, \mathcal{F})$, $(Y, \mathcal{G})$ e $(Z, \mathcal{H})$ três espaços mensuráveis. Então*

$$(\mathcal{F} \otimes \mathcal{G}) \otimes \mathcal{H} = \mathcal{F} \otimes (\mathcal{G} \otimes \mathcal{H}) = \mathcal{F} \otimes \mathcal{G} \otimes \mathcal{H}.$$

***Demonstração.*** Para todo $A \in \mathcal{F}$, $B \in \mathcal{G}$ e $C \in \mathcal{H}$,

$$A \times B \times C \in (\mathcal{F} \otimes \mathcal{G}) \otimes \mathcal{H} \;\text{ e }\; A \times B \times C \in \mathcal{F} \otimes (\mathcal{G} \otimes \mathcal{H}).$$

Portanto, pela minimalidade de $\mathcal{F} \otimes \mathcal{G} \otimes \mathcal{H}$,

$$\mathcal{F} \otimes \mathcal{G} \otimes \mathcal{H} \subseteq (\mathcal{F} \otimes \mathcal{G}) \otimes \mathcal{H} \;\cap\; \mathcal{F} \otimes (\mathcal{G} \otimes \mathcal{H}).$$

Por outro lado, pelo exercício 7,

$$\sigma\left[p_{X \times Y}\right] = \sigma\left[\{A \times B \times Z;\, A \in \mathcal{F},\, B \in \mathcal{G}\}\right] \subseteq \mathcal{F} \otimes \mathcal{G} \otimes \mathcal{H}.$$

Claramente,

$$\sigma\left[p_Z\right] = \{X \times Y \times C;\, C \in \mathcal{H}\} \subseteq \mathcal{F} \otimes \mathcal{G} \otimes \mathcal{H}.$$

Como $(\mathcal{F} \otimes \mathcal{G}) \otimes \mathcal{H}$ é a mais pequena álgebra$-\sigma$ tal que as duas projeções, $p_{X \times Y}$ e $p_Z$, são mensuráveis, então

$$(\mathcal{F} \otimes \mathcal{G}) \otimes \mathcal{H} \subseteq \mathcal{F} \otimes \mathcal{G} \otimes \mathcal{H}.$$

A demonstração de $\mathcal{F} \otimes (\mathcal{G} \otimes \mathcal{H}) \subseteq \mathcal{F} \otimes \mathcal{G} \otimes \mathcal{H}$ é semelhante. :-)

### 3.2.4 Borelianos de $\mathbb{R}^n$

Para mostrar, por exemplo, que a soma de duas funções reais, mensuráveis é mensurável, convém trabalhar em duas dimensões. Por isso, vamos introduzir e caraterizar as funções mensuráveis a valores vetoriais (tomando valores em $\mathbb{R}^n$, $n \geq 2$).

**Teorema 3.2.** *A álgebra$-\sigma$ de Borel de $\mathbb{R}^n$, denotada $\mathcal{B}^n$, é igual à álgebra$-\sigma$ produto de $n$ cópias de $\mathcal{B}$, a álgebra$-\sigma$ de Borel de $\mathbb{R}$. i.e.*

$$\mathcal{B}^n = \mathcal{B} \otimes \cdots \otimes \mathcal{B} := \sigma\left[\{A_1 \times \cdots \times A_n;\ \forall\, i = 1, \cdots, n,\ A_i \in \mathcal{B},\ \}\right].$$

(Lembre-se que $\mathcal{B}^n$ é a álgebra$-\sigma$ gerada pelos abertos de $\mathbb{R}^n$.)

***Demonstração.*** Como as projeções $\Pi_i : (x_1, \cdots, x_n) \in \mathbb{R}^n \to x_i \in \mathbb{R}$ são contínuas, então elas são $\mathcal{B}^n/\mathcal{B}-$mensuráveis. Mais $\mathcal{B} \otimes \cdots \otimes \mathcal{B}$ é a mais pequena álgebra$-\sigma$ tal que as $n$ projeções são contínuas, então

$$\mathcal{B} \otimes \cdots \otimes \mathcal{B} \subseteq \mathcal{B}^n.$$

Para mostrar a otra inclusão, observe que:

*i.* Para todo $a, b \in \mathbb{Q}$, $a < b$, $(a, b) \times \cdots \times (a, b) \in \mathcal{B} \otimes \cdots \otimes \mathcal{B}$.

*ii.* O conjunto $Q = \{(a, b) \times \cdots \times (a, b)\,;\ a, b \in \mathbb{Q},\ a < b\}$ é numerável.

*iii.* Todo conjunto aberto $A \subseteq \mathbb{R}^n$ é igual à união de todos aqueles elementos de $Q$ que estão contidos em $A$.

Como $Q$ é numerável e $Q \subseteq \mathcal{B} \otimes \cdots \otimes \mathcal{B}$, então, para todo $A \subseteq \mathbb{R}^n$, aberto, $A \in \mathcal{B} \otimes \cdots \otimes \mathcal{B}$. Pela minimalidade de $\mathcal{B}^n$,

$$\mathcal{B}^n \subseteq \mathcal{B} \otimes \cdots \otimes \mathcal{B}.\ \ :-)$$

O corolário seguinte é muito útil e será invocado um sem-número de vezes. A demonstração fica como exercício. (cf. Exercício 8.)

**Corolário 3.3.** Uma função $f = (f_1, \cdots, f_n) : X \to \mathbb{R}^n$ é $\mathcal{F}/\mathcal{B}^n-$mensurável se e só se todas as componentes $f_i : X \to \mathbb{R}$, $i = 1, \cdots, n$ , são $\mathcal{F}/\mathcal{B}-$mensuráveis.

#### Operações com funções mensuráveis

**Lema 3.11.** *Sejam $(X, \mathcal{F})$ um espaço mensurável, $f : X \to \mathbb{R}$ e $g : X \to \mathbb{R}$ duas funções $\mathcal{F}/\mathcal{B}-$mensuráveis e $\alpha \in \mathbb{R}$. Então, as seguintes funções são mensuráveis:*

$$
\begin{array}{rrll}
\textit{(i)} & \alpha f: & x \in X & \longrightarrow \alpha \cdot f(x) \in \mathbb{R}. \\
\textit{(ii)} & f+g: & x \in X & \longrightarrow f(x)+g(x) \in \mathbb{R}. \\
\textit{(iii)} & f \cdot g: & x \in X & \longrightarrow f(x) \cdot g(x) \in \mathbb{R}. \\
\textit{(iv)} & \frac{1}{f}: & x \in X \setminus \{f=0\} & \longrightarrow \frac{1}{f(x)} \in \mathbb{R}.
\end{array}
$$

***Demonstração.*** Pelo Corolário 3.3, a aplicação

$$\Theta : x \in X \to (f(x), g(x)) \in \mathbb{R}^2$$

é mensurável.

Por outro lado, as aplicações

$$
\begin{array}{rll}
h_1: & (u,v) \in \mathbb{R}^2 & \longrightarrow \alpha \cdot u + \beta \cdot v \in \mathbb{R}. \\
h_2: & (u,v) \in \mathbb{R}^2 & \longrightarrow u \cdot v \in \mathbb{R}. \\
h_3: & u \in \mathbb{R} \setminus \{0\} & \longrightarrow \frac{1}{u} \in \mathbb{R}.
\end{array}
$$

($\alpha$ e $\beta$ fixos) são contínuas e portanto mensuráveis (cf. Lema 3.7).

Sendo composições de funções mensuráveis, as funções

$$\alpha f + \beta g = h_1 \circ \Theta, \quad f \cdot g = h_2 \circ \Theta \quad \text{e} \quad \frac{1}{f} = h_3 \circ f$$

são mensuráveis. (cf. Lema 3.5.) :-)

## 3.3 Medidas de Lebesgue-Stieltjes

Antes de discutir as medidas de Lebesgue-Stieltjes, vamos falar das funções de acumulação das distribuições de probabilidade, por ser mais conhecidas. Na Teoría das Probalidades, a função distribuição, também conhecida como função de acumulação, é utilizada para caracterizar as distribuições de probabilidade das "variáveis aleatórias" que definiremos mais a frente (cf. Definição 3.6).

Depois de estudar as medidas de Lebesgue-Stieltjes e as suas propriedades, voltaremos às probabilidades para falar da decomposição das leis de probabilidade em parte discreta, contínua e absolutamente contínua. Falaremos também da independência de variáveis aleatórias. No fim da secção, damos uma pequena lista comentada das probabilidades mais conhecidas em Estatística.

### 3.3.1 Variáveis aleatórias

Este parece-nos ser o lugar propício para discutir sobre variáveis aleatórias. Vimos no Capítulo 2 que um espaço de probabilidade é um tripleto $(\Omega, \mathcal{F}, \mathrm{P})$, onde $\Omega$ é um conjunto, $\mathcal{F}$ é uma álgebra$-\sigma$ sobre $\Omega$ e $\mathrm{P}$ é uma medida de probabilidade sobre $\mathcal{F}$, i.e. $\mathrm{P}(\Omega) = 1$. (cf. Definição 2.3.)

Note a mudança de notação de $(X, \mathcal{F}, \mu)$, no Capítulo 2, para $(\Omega, \mathcal{F}, \mathrm{P})$. A única razão é esta ser a notação tradicional na Teoria das Probabilidades e é a notação que utilizaremos especificamente para um espaço de probabilidade genérico.

Lembre-se que, tratando-se de um espaço de probabilidade, $(\Omega, \mathcal{F}, \mathrm{P})$, os elementos de $\mathcal{F}$ chamam-se ***eventos***.

**Definição 3.6.** Seja $(\Omega, \mathcal{F}, \mathrm{P})$ um espaço de probabilidade. Qualquer função

$$\mathrm{X} : \Omega \to \mathbb{R}^n,\ \mathcal{F}/\mathcal{B}^n\text{−mensurável},$$

chama-se ***variável aleatória*** $n-$dimensional.

***Observação.*** Tradicionalmente, em Estatística, as variáveis aleatórias denotam-se com as letras X, Y, Z, etc. Mais uma razão para mudar a notação de $(X, \mathcal{F}, \mu)$ para $(\Omega, \mathcal{F}, \mathrm{P})$.

Aqui, vamos concentrar-nos nas variáveis aleatórias reais (i.e. $n = 1$).

**Definição 3.7.** Sejam $(\Omega, \mathcal{F}, \mathrm{P})$ um espaço de probabilidade e $\mathrm{X} : \Omega \to \overline{\mathbb{R}}$ uma variável aleatória. A medida imagem, ou seja a medida

$$A \in \mathcal{B} \to \nu_{\mathrm{X}}(A) := \mathrm{P}(\mathrm{X} \in A) = \mathrm{P}(\{\omega \in \Omega;\ \mathrm{X}(\omega) \in A\}) \tag{3.10}$$

(i.e. $\nu_{\mathrm{X}} = \mathrm{X}_{\#}\mathrm{P}$, cf. Lema 2.1 e Definição 2.4) chama-se ***distribuição de probabilidad*** ou ***lei*** de X.

**Nota 3.3.** Num modelo estatístico, fala-se de um espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$, no qual as variáveis aleatórias estão definidas. (cf. Definição 3.6.) Mas, muitas das vezes, o espaço $(\Omega, \mathcal{F}, \mathrm{P})$ fica omisso. Quer dizer, não é explícito nem P, nem $\mathcal{F}$, nem sequer $\Omega$ e menos ainda X, simplesmente supõe-se que estes entes existem. Esta situação é habitual em Estatística, já que a existência de $(\Omega, \mathcal{F}, \mathrm{P})$ não coloca nenhum problema teórico, nem sequer no caso dos processos estocásticos (separáveis) em tempo contínuo. No entanto, esquecer $(\Omega, \mathcal{F}, \mathrm{P})$ e ficar unicamente com o espaço imagem $(\mathbb{R}, \mathcal{B}, \nu)$ reduz as Probabilidades à Teoria da Medida. De facto, o espaço $(\Omega, \mathcal{F}, \mathrm{P})$ possui informação sobre conceitos puramente probabilistas que de outra forma ficam obscurecidos.

Uma situação extrema e decididamente perigosa, que de facto nada tem a ver com a Estatística como ciência, mas com a percepção errada do público, é aquela na qual nem sequer a lei é dada e pensa-se nela como sendo algo intrínseco! Pergunta-se por exemplo: Qual é a probabilidade disto ou aquelo? sem nunca ter especificado

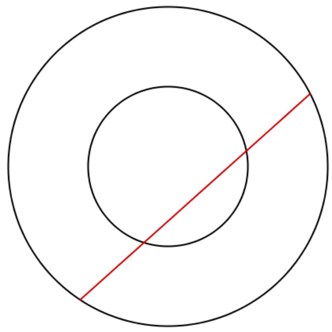


Figura 3.1: corda escolhida ao "azar"

distribuição nenhuma. Para ilustrar este facto, vamos relatar um exemplo clássico, dado por Joseph Louis François Bertrand (1822–1900) no seu livro "Calcul des probabilités," publicado em 1889.

**O paradoxo de Bertrand.** Numa circunferência de raio 2 cm, escolhemos uma corda ao "azar." (Uma corda é um segmento de reta que une dois pontos da circunferência. Ver a Figura 3.1.)

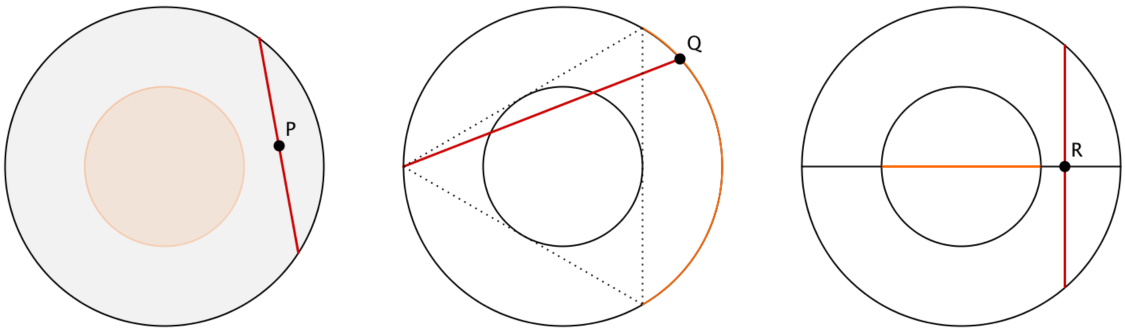
**Pergunta.** Qual é a probabilidade da corda, escolhida ao "azar", intersetar o círculo concêntrico de raio 1 cm?

A frase "*a corda interseta o círculo interior*" não é um "*evento*" no sentido das probabilidades, em quanto não se determinar o espaço mensurável. E mesmo estando especificado o espaço mensurável, não podemos falar de "*probabilidade*" de um evento, sem especificar a medida de probabilidade à qual estamo-nos a referir. Deste modo, podemos dar diferentes respostas, de acordo com diferentes interpretações destes termos, como se mostra a continuação.

***Solução 1.*** Todas as cordas (exceto os diámetros) estão determinadas pelo ponto meio, de forma unívoca. Toda corda interseta o disco de raio 1 se e só se o ponto meio pertence ao disco de raio 1. Logo, para escolher uma corda, basta escolher o seu ponto meio. (Ponto P no esquema da esquerda, Figura 3.2.)

Podemos convenir que o ponto é escolhido numa certa região, com uma probabilidade proporcional à área de essa região. Nesse caso, a resposta à pergunta: Qual é a probabilidade da corda, escolhida aleatoriamente, intersetar o círculo concéntrico de raio 1 cm? é o quociente entre a área do disco de raio 1 dividido pela área do disco de raio 2. Ou seja $1/4$.

***Solução 2.*** Por simetria, podemos considerar unicamente as cordas que nascem num ponto fixo da circunferência. Nesse caso, as cordas estão determinadas pelo




$$\frac{\textit{área disco interior}}{\textit{área disco exterior}} = 1/4 \qquad \frac{\textit{comprimento arco laranja}}{\textit{comprimento circunferência}} = 1/3 \qquad \frac{\textit{diámetro disco interior}}{\textit{diámetro disco exterior}} = 1/2$$

Figura 3.2: três interpretações diferentes

segundo ponto sobre a circunferência. As cordas que intersetam o disco de raio 1 são as que têm o segundo ponto no terço de circunferencia que fica frente ao ponto fixo. (Ponto Q no esquema central, Figura 3.2).

A medida natural associada com os subconjuntos da circunferência é proporcional ao comprimento de arco. Neste caso, a respuesta é o quociente entre o comprimento do arco "favorável," dividido pelo comprimento da circunferência. Quer dizer, $1/3$.

***Solução 3.*** Novamente por simetra, podemos considerar unicamente as cordas "verticais."Desta vez, as cordas estão determinadas pelo ponto de interseção com o diámetro horizontal. (Ponto R no esquema da direita, Figura 3.2).

A medida "uniforme"associada aos subconjuntos do diámetro é proporcional ao seu comprimento. Para a corda intersetar o disco interior, o ponto de interseção com o diámetro horizontal deve estar no segmento central, que corresponde ao diámetro do disco pequeno. O comprimento deste segmento é igual à mitade do comprimento do diámetro do disco exterior. Interpretando as coisas desta forma, a resposta é $1/2$.

O exemplo de Bertrand mostra que, em Probabilidades, "ao azar" ou "aleatoriamente" não são conceitos intrínsicos mas que devem ser definidos no contexto.

### Medidas de probabilidade em $(\mathbb{R}, \mathcal{B})$

**Definição 3.8.** *(Função de acumulação)* Sejam $(\Omega, \mathcal{F}, \mathrm{P})$ um espaço de probabilidade e $\mathrm{X} : \Omega \to \overline{\mathbb{R}}$ uma variável aleatória. A função

$$x \in \overline{\mathbb{R}} \to \mathrm{F_X}(x) = \mathrm{P}(\mathrm{X} \leq x) = \nu_{\mathrm{X}}([-\infty, x])$$

chama-se ***função distribuição*** ou ***função de acumulação*** de X.

**Nota 3.4.** Suporemos, como é razoável, que as distribuições de probabilidade assim obtidas (medidas imagem duma variável aleatória) verificam:

$$\nu_{\mathrm{X}}\left(\{+\infty,-\infty\}\right)=\mathrm{P}\left(\mathrm{X}=\pm\infty\right)=0.$$

No entanto, isto não significa que $\{\omega\in\Omega;\ \mathrm{X}(\omega)=\pm\infty\}=\emptyset$. Significa simplesmente que este conjunto tem medida nula.

Nesse caso,

$$\nu_{\mathrm{X}}\left(\mathbb{R}\right)=1$$

e é suficiente definir a função de acumulação em $\mathbb{R}$:

$$x\in\mathbb{R}\to\mathrm{F}_{\mathrm{X}}\left(x\right)=\mathrm{P}\left(\mathrm{X}\leq x\right)=\nu_{\mathrm{X}}\left((-\infty,x]\right). \tag{3.11}$$

(Excluimos $-\infty$ do intervalo do termo da direita, pois $\nu_{\mathrm{X}}(\{-\infty\})=0$.)

### Propriedades da função de acumulação

De seguida, vamos estudar as propriedades principais de uma função de acumulação qualquer, i.e. uma função de acumulação $\mathrm{F}_{\mathrm{X}}$, definida por (3.11), a partir duma variável aleatória $\mathrm{X}:\Omega\to\overline{\mathbb{R}}$ qualquer, sobre um espaço de probabilidade $(\Omega,\mathcal{F},\mathrm{P})$ igualmente arbitrário.

Como dissemos, suporemos sempre que $\mathrm{P}\left(\mathrm{X}=\pm\infty\right)=0$. (O qual náo é essencial, mas é prático e razoável.)

**Lema 3.12.** *Sejam* $(\Omega,\mathcal{F},\mathrm{P})$ *um espaço de probabilidade,* $\mathrm{X}:\Omega\to\overline{\mathbb{R}}$ *uma variável aleatória. Então, a função de acumulação* $\mathrm{F}_{\mathrm{X}}$ *é crescente e contínua à direita.*

*Além disso,*

$$\lim_{y\to x^-}\mathrm{F}_{\mathrm{X}}\left(y\right)=\nu_{\mathrm{X}}\left((-\infty,x)\right), \tag{3.12}$$

*onde* $\nu_{\mathrm{X}}$ *é a distribuição de probabilidad de* X*, definida por* (3.10)*.*

***Demonstração.*** Por (3.11) e pela monotonía de $\nu_{\mathrm{X}}$ (cf. alinha 2 da Proposição 2.1), $\mathrm{F}_{\mathrm{X}}$ é crescente. Como tal, para todo $x\in\mathbb{R}$, os limites

$$\lim_{y\to x^-}\mathrm{F}_{\mathrm{X}}\left(y\right)\quad\text{e}\quad\lim_{y\to x^+}\mathrm{F}_{\mathrm{X}}\left(y\right).$$

existem e podem ser avaliados tomando o limite duma sucesão monótona $y_n \to x$ qualquer (crescente ou decrescente, consoante o caso). Por exemplo, para calcular o segundo limite, podemos usar $y_n = x + \frac{1}{n}$:

$$\lim_{y\to x^+} \mathrm{F_X}(y) = \lim_{n\to\infty} \mathrm{F_X}\left(x + \frac{1}{n}\right) = \lim_{n\to\infty} \nu_{\mathrm{X}}\left(\left(-\infty, x + \frac{1}{n}\right]\right). \tag{3.13}$$

Por (2.13) (cf. Lema 2.4),

$$\lim_{n\to\infty} \nu_{\mathrm{X}}\left(\left(-\infty, x + \frac{1}{n}\right]\right) = \nu_{\mathrm{X}}\left(\bigcap_{n=1}^{\infty}\left(-\infty, x + \frac{1}{n}\right]\right) = \nu_{\mathrm{X}}\left((-\infty, x]\right) = \mathrm{F_X}(x). \tag{3.14}$$

Juntando (3.13) e (3.14), conclui-se que $\mathrm{F_X}$ é contínua pela direita.

Usando a sucessão $x - \frac{1}{n}$, obtemos

$$\lim_{y\to x^-} \mathrm{F_X}(y) = \lim_{n\to\infty} \mathrm{F_X}\left(x - \frac{1}{n}\right) = \lim_{n\to\infty} \nu_{\mathrm{X}}\left(\left(-\infty, x - \frac{1}{n}\right]\right). \tag{3.15}$$

Por (2.9) (cf. Lema 2.2),

$$\lim_{n\to\infty} \nu_{\mathrm{X}}\left(\left(-\infty, x - \frac{1}{n}\right]\right) = \nu_{\mathrm{X}}\left(\bigcup_{n=1}^{\infty}\left(-\infty, x - \frac{1}{n}\right]\right) = \nu_{\mathrm{X}}\left((-\infty, x)\right). \tag{3.16}$$

Juntando (3.15) e (3.16), obtemos (3.12). :-)

**Corolário 3.4.**

$$\lim_{x\to-\infty} \mathrm{F_X}(x) = 0 \quad \text{e} \quad \lim_{x\to+\infty} \mathrm{F_X}(x) = 1. \tag{3.17}$$

***Demonstração.*** (cf. Exercício 9.)

**Lema 3.13.** *Para todo par $a, b \in \mathbb{R}$, tal que $a < b$,*

$$\nu_{\mathrm{X}}((a, b]) = \mathrm{F_X}(b) - \mathrm{F_X}(a). \tag{3.18}$$

*Para além disso, $\nu_{\mathrm{X}}$ é a única medida sobre $(\mathbb{R}, \mathcal{B})$ que verifica* (3.18).

***Demonstração.*** Como $(-\infty, b] = (-\infty, a] \cup (a, b]$, por (2.2) (cf. Proposição 2.1, alinha 1.),

$$\mathrm{F_X}(b) = \nu_{\mathrm{X}}((-\infty, b]) = \nu_{\mathrm{X}}((-\infty, a]) + \nu_{\mathrm{X}}((a, b]) = \mathrm{F_X}(a) + \nu_{\mathrm{X}}((a, b]).$$

A unicidade de $\nu_{\mathrm{X}}$ deduz-se da Proposição 2.6. :-)

**Lema 3.14.**

1. *Nos pontos onde* $\mathrm{F_X}$ *não é contínua, caso existam,* $\mathrm{F_X}$ *dá um salto. (i.e. Os limites laterais existem, mas são distintos.)*

2. $\mathrm{F_X}$ *tem, no máximo, uma quantidade numerável de discontinuidades.*

3. *Para todo* $x \in \mathbb{R}$*,*
$$\nu_{\mathrm{X}}(\{x\}) = \mathrm{F_X}(x) - \lim_{y \to x^-} \mathrm{F_X}(y).$$
*(Lembre-se que* $\{x\} \in \mathcal{B}$ *e portanto* $\mu(\{x\})$ *tem sentido.)*

4. *O conjunto*
$$\{x \in \mathbb{R};\ \nu_{\mathrm{X}}(\{x\}) > 0\}$$
*é, no máximo, numerável.*

***Demonstração.*** Como $\mathrm{F_X}$ é crescente, o limite

$$\lim_{y \to x^-} \mathrm{F_X}(y)$$

existe. Portanto, as discontinuidades, se as houver, são discontinuidades de “salto”. Quer dizer, $\mathrm{F_X}$ é discontínua em $x \in \mathbb{R}$ se e só se

$$\mathrm{F_X}(x) - \lim_{y \to x^-} \mathrm{F_X}(y) > 0.$$

Desta forma fica provado o primeiro ponto .

Como $\mathrm{F_X}$ é crescente, por (3.17), a soma dos saltos não pode ser superior a 1, pois
$$\lim_{x \to +\infty} \mathrm{F_X}(x) - \lim_{x \to -\infty} \mathrm{F_X}(x) = \nu_{\mathrm{X}}(\mathbb{R}) = 1.$$

Portanto, qualquer que seja $n \in \mathbb{N}$, há no máximo $n$ saltos de tamanho igual ou superior a $\frac{1}{n}$. Isto demonstra o segundo ponto, já que

$$\left\{x \in \mathbb{R};\ \mathrm{F_X}(x) - \lim_{y \to x^-} \mathrm{F_X}(y) > 0\right\} = \bigcup_{n=1}^{\infty} \left\{x \in \mathbb{R};\ \mathrm{F_X}(x) - \lim_{y \to x^-} \mathrm{F_X}(y) \geq \frac{1}{n}\right\}.$$

O terceiro ponto resulta de (3.11) e (3.12). O ponto quarto é consequência dos pontos segundo e terceiro.

Os pontos $x \in \mathbb{R}$ nos quais $\nu_{\mathrm{X}}(\{x\}) > 0$ chamam-se ***atomos*** de $\nu_{\mathrm{X}}$.

### 3.3.2 Construção das medidas de Lebesgue-Stieltjes

Na subsecção 3.3.1, vimos as propriedades da função de acumulação $\mathrm{F_X}$, definida por (3.11). Nesta secção, partimos duma função $\mathrm{F} : \mathbb{R} \to \mathbb{R}$ crescente, contínua à direita, e construímos uma medida seguindo os mesmos passos da construção da medida de Lebesgue sobre $(0, 1]$. (cf. Teorema 2.2.)

**Teorema 3.3.** *Seja* $\mathrm{F} : \mathbb{R} \to \mathbb{R}$ *uma função crescente e contínua à direita. Existe uma única medida de Borel, chamemos-lhe* $\nu_{\mathrm{F}}$*, tal que para todo* $a, b \in \mathbb{R}$*,* $a < b$ *tenhamos*

$$\nu_{\mathrm{F}}\left((a, b]\right) = \mathrm{F}(b) - \mathrm{F}(a). \tag{3.19}$$

***Demonstração.*** Pelo Teorema 2.1, basta mostrar a existência de uma premedida que verifique (3.19), definida sobre uma álgebra que gira $\mathcal{B}$. Para isto, vamos seguir passo a passo a demonstração do Teorema 2.2.

Seja $\mathcal{F}_0$ a álgebra definida no *Passo 1* da demonstração do Teorema 2.2, mas desta vez os intervalos não estão limitados por 0 e 1. i.e. $\mathcal{F}_0$ é composto do conjunto vazio e dos conjuntos da forma

$$(a_1, b_1] \cup (a_2, b_2] \cup ... \cup (a_n, b_n],$$

onde $a_i, b_i \in \mathbb{R}$, $i = 1, ..., n$ e $a_1 < b_1 < a_2 < b_2 < ... < a_n < b_n$.

No *Passo 3* da demonstração do Teorema 2.2, mostramos que $\sigma[\mathcal{F}_0] = \mathcal{B}_{(0,1]}$, da mesma forma, podemos mostrar $\sigma[\mathcal{F}_0] = \mathcal{B}_{\mathbb{R}}\ (= \mathcal{B})$. (cf. Exercício 10).

Agora, define-se $\nu_0$ da seguinte maneira.

$$\nu_0\left((a_1, b_1] \cup ... \cup (a_n, b_n]\right) = \mathrm{F}(b_1) - \mathrm{F}(a_1) + ... + \mathrm{F}(b_n) - \mathrm{F}(a_n), \tag{3.20}$$

onde $A = (a_1, b_1] \cup ... \cup (a_n, b_n]$, sendo $(a_i, b_i]$, $i = 1, ..., n$, as componentes de $A$. Se $A = \emptyset$, pomos $\nu_0(\emptyset) = 0$.

Claramente, $\nu_0$ verifica a condição (3.19). Para demonstrar que $\nu_0$ é uma premedida $\sigma-$finita sobre $\mathcal{F}_0$, seguimos a linha de raciocínio do *Passo 2* da demonstração do Teorema 2.2. Mas veremos no desenvolvimento que se segue, onde e como intervêm as hipóteses de F ser crescente e contínua à direita.

Sendo $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}_0$ e $A \in \mathcal{F}_0$, tal que $A \subseteq \bigcup_{i=1}^{\infty} A_i$, devemos mostrar que

$$\nu_0(A) \leq \sum_{i=1}^{\infty} \nu_0(A_i).$$

Assumimos, sem pérda de geralidade, que os conjuntos $A_i$ são (eles próprios) intervalos semiabertos e que $A$ é um intervalo semiaberto. Ou seja, assumimos que

$$(a,b] \subseteq \bigcup_{i=1}^{\infty} (a_i, b_i]\,,$$

onde $a, b, a_i, b_i \in \mathbb{R}$, $a < b$, e $a_i < b_i$, $i = 1, 2, ...$ e queremos provar que

$$\mathrm{F}(b) - \mathrm{F}(a) \leq \sum_{i=1}^{\infty} (\,\mathrm{F}\,(b_i) - \mathrm{F}\,(a_i)\,)\,. \tag{3.21}$$

Note que a série do lado direito de (3.21) tem sentido em $\overline{\mathbb{R}}$, já que todos os termos são positivos (por F ser crescente).

Seguindo o *Passo 2* da demonstração do Teorema 2.2, fixamos $0 < \epsilon < b - a$ e, tornando os intervalos $(a_i, b_i]$ um cadinho maiores, conseguimos igualmente um recobrimento aberto de $[a + \epsilon, b]$. Mas neste ponto, temos de ser mais cuidadosos com a escolha dos intervalos, pois temos de ter em consideração a função F.

Como F é contínua pela direita, para todo $i \in \mathbb{N}$, existe $\delta_i > 0$ tal que

$$\mathrm{F}\,(b_i + \delta_i) \leq \mathrm{F}\,(b_i) + \frac{\epsilon}{2^i}. \tag{3.22}$$

Utilizamos a sucessão $\delta_i$ para construir o nosso recobrimento, e não $\frac{\epsilon}{2^i}$, como no Teorema 2.2.

$$[a + \epsilon, b] \subseteq (a, b] \subseteq \bigcup_{i=1}^{\infty} (a_i, b_i] \subseteq \bigcup_{i=1}^{\infty} (a_i, b_i + \delta_i)\,.$$

Graças à compacidade de $[a + \epsilon, b]$, existe $n_0 \in \mathbb{N}$ tal que

$$[a + \epsilon, b] \subseteq \bigcup_{i=1}^{n_0} (a_i, b_i + \delta_i)$$

e portanto

$$(a + \epsilon, b] \subseteq \bigcup_{i=1}^{n_0} (a_i, b_i + \delta_i]\,. \tag{3.23}$$

Aplicando $\nu_0$ em ambos os lados, pela monotonía de $\nu_0$ , temos

$$\nu_0\,((a + \epsilon, b]) \leq \nu_0 \left( \bigcup_{i=1}^{n_0} (a_i, b_i + \delta_i] \right).$$

Por (3.20), e a subaditivade finita de $\nu_0$,

$$\mathrm{F}(b) - \mathrm{F}(a + \epsilon) \leq \sum_{i=1}^{n_0} (\,\mathrm{F}\,(b_i + \delta_i) - \mathrm{F}\,(a_i)\,) \leq \sum_{i=1}^{\infty} (\,\mathrm{F}\,(b_i + \delta_i) - \mathrm{F}\,(a_i)\,)\,.$$

(cf. Exercício 11.)

Usando (3.22),

$$F(b) - F(a+\epsilon) \leq \sum_{i=1}^{\infty} \Big( F(b_i) + \frac{\epsilon}{2^i} - F(a_i) \Big) = \epsilon + \sum_{i=1}^{\infty} \left( F(b_i) - F(a_i) \right).$$

Fazendo tender $\epsilon \to 0$, pela continuidade à direita de F, temos $F(a+\epsilon) \to F(a)$. Sendo $\epsilon$ arbitrário, obtemos (3.21). :-)

**Definição 3.9.** Uma medida $\nu_F$ construída a partir duma função F crescente, contínua à direita, conhece-se como a ***medida de Lebesgue-Stieltjes*** de F.

### Propriedades das medidas de Lebesgue-Stieltjes

**Proposição 3.2.** Sendo $F : \mathbb{R} \to \mathbb{R}$ uma função crescente, contínua à direita e $\nu_F$ a medida de Lebesgue-Stieltjes de F, verificam-se as propriedades seguintes.

1. A medida $\nu_F$ é $\sigma-$finita. Mas $\nu_F$ é finita se e só se forem finitos ambos os limites
$$\lim_{x \to -\infty} F(x) \quad \text{e} \quad \lim_{x \to +\infty} F(x).$$
2. F tem, no máximo, um número numerável de discontinuidades (saltos).
3. Para todo $x \in \mathbb{R}$,
$$\nu_F(\{x\}) = F(x) - \lim_{y \to x^-} F(y).$$
4. O conjunto $\{x \in \mathbb{R};\ \nu_F(\{x\}) > 0\}$. é finito ou numerável.

***Demonstração.*** As demonstrações destas propriedades sào similares às do Lema 3.14. (cf. Exercício 12.)

### Função de Lebesgue-Stieltjes

A fórmula (3.11) (subsecção 3.3.1) define a função de acumulação, $F_X$, de medidas finitas. Se a medida $\nu$ não é finita, mas toma valores finitos nos intervalos limitados, ainda podemos construir uma função $F_\nu$ que verifique 3.19. Efetivamente, temos a seguinte proposição, cuja demonstração fica como exercício. (cf. Exercício 13.)

**Proposição 3.3.** Se $\nu$ é uma medida Boreliana, que toma valores finitos nos conjuntos compactos, então a função

$$\mathrm{F}_\nu(x) = \begin{cases} \nu([0,x]) & \text{se } x \geq 0, \\ -\nu((x,0)) & \text{se } x < 0 \end{cases} \tag{3.24}$$

é crescente, contínua à direita e para todo $a, b \in \mathbb{R}$, $a < b$, temos

$$\nu((a,b]) = \mathrm{F}_\nu(b) - \mathrm{F}_\nu(a) . \tag{3.25}$$

Ademais, se F é outra função que verifica (3.25), então $\mathrm{F} - \mathrm{F}_\nu$ é constante.

**Definição 3.10.** Qualquer função crescente, contínua à direita, que verifica (3.25), chama-se ***função de Lebesgue-Stieltjes*** de $\nu$.

***Observação.*** Se $\nu$ é uma medida de probabilidade, a função de acumulação (cf. Definição 3.8) é a única função de Lebesgue-Stieltjes que tende para 0 em $-\infty$.

**Nota 3.5.** Pelo Teorema 3.3, $\mathrm{F}_\nu$ caracteriza a medida $\nu$, já que $\nu$ é a única medida que verifica (3.25). i.e. $\nu = \nu_{\mathrm{F}_\nu}$ (cf. Definição 3.9).

### 3.3.3 Classificação de medidas Borelianas

Para falar da decomposição de medidas Borelianas, necessitamos introduzir os conceitos de medida estrangeira e medida absolutamente contínua. Estes conceitos serão tratados com maior profundidade no Capítulo 4. Aqui não podemos fazer mais do que nos limitar ao caso de medidas Borelianas, e unicamente em relação à medida de comprimento sobre $\mathbb{R}$ (cf. Teorema 2.2), também chamada, por abuso de linguagem, medida de Lebesgue em dimensão 1. (A diferença entre a medida de comprimento, definida sobre os Borelianos, e a medida de Lebesgue é explicada na Subsecção 2.8.2.) A pesar de que a discussão que se segue faz referência unicamente à álgebra$-\sigma$ de Borel, nos vamos chamar medida de Lebesgue à medida de comprimentos sobre os Borelianos, para nos habituar ao nome comummente utilizado.

**Definição 3.11.** Uma medida Boreliana, $\nu$, sobre $\mathbb{R}$ e a medida de Lebesgue são ***estrangeiras*** entre elas, se existe $A \in \mathcal{B}$ tal que

$$\nu(A^{\mathrm{c}}) = \lambda(A) = 0 \tag{3.26}$$

**Exemplo 3.1.** Uma clase importante de medidas estrangeiras (com respeito à medida de Lebesgue) são as medidas discretas, definidas em (3.29), já que as medidas discretas estão concentradas em conjuntos numeráveis e tais conjuntos têm comprimento nulo.

**Definição 3.12.** Uma medida Boreliana, $\nu$, que toma valores finitos nos intervalos de comprimento finito, é ***absolutamente contínua*** com respeito à medida de Lebesgue, se existe uma função $f : \mathbb{R} \to \mathbb{R}^+$, "localmente" integrável no sentido de Riemann, tal que a função de Lebesgue-Stieltjes de $\nu$, $\mathrm{F}_\nu$, definida pela fórmula (3.24) da Proposição 3.2, se escreve como

$$\mathrm{F}_\nu(x) = \int_0^x f_\nu(t)dt. \tag{3.27}$$

A função $f_\nu$ é a ***densidade*** de $\nu$ com respeito à medida de Lebesgue de $\mathbb{R}$.

***Observação.*** No Capítulo 4, daremos uma definição mais abrangente ao considerar funções integráveis no sentido de Lebesgue e não no sentido de Riemann, como nesta definição. Localmente integrável significa que a função, restringida a qualquer intervalo de comprimento finito, é integrável.

**Exemplo 3.2.** A função de acumulação de uma variável aleatória U, uniformemente distribuida em $[0, 1]$ é

$$\mathrm{F}_\mathrm{U}(x) = \begin{cases} 0 & \text{se } x < 0, \\ x & \text{se } 0 \leq x < 1, \\ 1 & \text{se } x \geq 1. \end{cases}$$

Também podemos expressar $\mathrm{F}_\mathrm{U}$ como

$$\mathrm{F}_\mathrm{U}(x) = \int_0^x \chi_{(0,1)}(t)dt.$$

Portanto, a densidade da distribuição uniforme é a função $\chi_{(0,1)}$, o que corresponde à derivada de $\mathrm{F}_\mathrm{U}$, nos pontos onde ela existe. De facto,

$$\mathrm{F}_\mathrm{U}{}'(x) = \begin{cases} 0 & \text{se } x < 0, \\ 1 & \text{se } 0 < x < 1, \\ 0 & \text{se } x > 1. \end{cases}$$

### Medidas discretas

**Definição 3.13.** A ***medida de Dirac*** em 0 é dada por

$$A \in \mathcal{B} \to \delta_0(A) = \begin{cases} 1 & \text{se } 0 \in A, \\ 0 & \text{se } 0 \notin A. \end{cases}$$

A medida de Dirac em 0 é uma medida de probabilidade. A sua função de acumulação é

$$\mathrm{H}\,(x) = \begin{cases} 0 & \text{se } x < 0, \\ 1 & \text{se } x \geq 0, \end{cases} \tag{3.28}$$

conhecida como ***função de Heaviside***.

A medida de Dirac em $a$ é a medida $\delta_a = T_{a\#}\delta_0$, onde $T_a : \mathbb{R} \to \mathbb{R}$ é a "translação" $x \to a + x$. i.e.

$$A \in \mathcal{B} \to \delta_a\,(A) = \begin{cases} 1 & \text{se } a \in A, \\ 0 & \text{se } a \notin A. \end{cases}$$

**Definição 3.14.** Uma ***medida discreta*** é uma combinação linear finita ou numerável de medidas de Dirac. i.e.

$$\sum_{i \in I} c_i \cdot \delta_{a_i}, \tag{3.29}$$

sendo $a_i \in \mathbb{R}$, $c_i > 0$, $i \in I \subseteq \mathbb{N}$ finito ou numerável. (cf. Exercício 14)

**Exemplo 3.3.** Note que uma medida discreta é necessáriamente $\sigma-$finita. Mas isto não significa que ela seja finita sobre os intevalos de comprimento finito, que é uma das propriedades das medidas de Lebesgue-Stieljes. Por exemplo, consideremos uma enumeração (ordenamento) qualquer do conjunto dos números racionais, digamos $q_1, q_2, \ldots$. A medida de contagem sobre $\mathbb{Q}$,

$$\nu = \sum_{i \in \mathbb{N}} \delta_{q_i},$$

é uma medida discreta e portanto $\sigma-$finita, mas existem uma infinidade de números racionais em qualquer intervalo de comprimento positivo. Como tal, não admite uma função de Lebesgue-Stieltjes.

**Exemplo 3.4.** Modificando este exemplo, podemos obter uma medida de probabilidade discreta, tal que qualquer intervalo de comprimento positivo tenha probabilidade estritamente positiva. Efetivamente, basta tomar

$$\nu_1 = \sum_{i \in \mathbb{N}} 2^{-i} \cdot \delta_{q_i}.$$

A função de Lebesgue-Stieltjes (função de acumulação), i.e.

$$\mathrm{F}_{\nu_1}(x) = \sum_{i=1}^{\infty} 2^{-i} \cdot \mathrm{H}\,(x - q_i)\,,$$

dá uma infinidade de saltos em qualquer intervalo de comprimento positivo.

**Nota 3.6.** Qualquer medida discreta (cf. Definição 3.14) é estrangeiras com respeito à medida de Lebesgue (cf. Definição 3.11)

### Medidas contínuas

**Definição 3.15.** Uma medida Boreliana, $\nu$, que toma valores finitos nos intervalos de comprimento finito, é ***contínua*** se a função de Lebesgue-Stieltjes de $\nu$, $\mathrm{F}_\nu$, é contínua.

***Observação.*** Ao contrário das medidas discretas, a Definição 3.15 aplica-se unicamente às medidas de Lebesgue-Stieltjes.

**Nota 3.7.** Se $\nu$ é absolutamente contínua com respeito à medida de Lebesgue (cf. Definição 3.12), a função de Lebesgue-Stieltjes de $\nu$, $\mathrm{F}_\nu$, é dada pelo integral (3.27). Portanto, $\mathrm{F}_\nu$ é contínua.

A inversa não é verdadeira. Quer dizer, $\mathrm{F}_\nu$ pode ser contínua, sem existir necessariamente uma densidade $f_\nu$ verificando (3.27). (cf. Corolário 4.9.)

### Decomposição de medidas Borelianas

**Teorema 3.4.** Seja $\nu$ uma medida Boreliana, que toma valores finitos nos intervalos finitos e seja $\mathrm{F}_\nu$ a função de Lebesgue-Stieltjes de $\nu$, definida por (3.24).

Então, $\nu$ decompõe-se na soma de uma medida discreta e outra contínua. i.e.

$$\nu = \nu_d + \nu_c$$

sendo $\nu_d$ discreta ($\nu_d$ tem a forma (3.29)) e $\nu_c$ contínua ($\mathrm{F}_{\nu_c}$ é contínua).

***Demonstração.*** Pela alinha 4 da Proposição 3.2, $D = \{x \in \mathbb{R};\ \nu(\{x\}) > 0\}$ é, no máximo, numerável. Logo, podemos escrever os elementos de $D$ (se não for vazio) numa certa ordem, digamos $a_1, a_2, ..., a_i, ...,\ i \in I$, sendo $I \subseteq \mathbb{N}$ finito ou numerável. Agora, para todo $i \in I$, define-se

$$c_i = \mathrm{F}(a_i) - \lim_{y \to {a_i}^-} \mathrm{F}(y)\,.$$

Se $D$ não for vazio, a medida

$$\nu_d = \sum_{i \in I} c_i \cdot \delta_{a_i}$$

é uma medida discreta e chama-se *parte discreta* de $\nu$. A função de Lebesgue-Stieltjes de $\nu_d$ é

$$\mathrm{F}_{\nu_d}(x) = \sum_{i\in I} c_i \cdot \mathrm{H}\left(x - a_i\right).$$

A medida $\nu_c = \nu - \nu_d$ (se $D$ é vazio, pomos $\nu_c = \nu$) é uma medida Boreliana, e chama-se *parte contínua* de $\nu$, já que função

$$\mathrm{F}_{\nu_c} = \mathrm{F}_{\nu} - \mathrm{F}_{\nu_d}, \tag{3.30}$$

que é a função de Lebesgue-Stieltjes de $\nu_c$, é contínua em $\mathbb{R}$. (cf. Exercício 15.)

### 3.3.4 Variáveis aleatórias independentes

**Definição 3.16.** Seja $(\Omega, \mathcal{F}, \mathrm{P})$ um espaço de probabilidade e sejam $\mathrm{X}_1, ..., \mathrm{X}_n$ uma família de $n$ variáveis aleatórias definidas sobre $\Omega$ tomando valores em $\overline{\mathbb{R}}$.

Dizemos que as variáveis aleatórias $\mathrm{X}_1, ..., \mathrm{X}_n$ são ***independentes***, se para todo $A_1, ..., A_n \in \mathcal{B}$,

$$\mathrm{P}\left(\mathrm{X}_1 \in A_1, \, \ldots, \mathrm{X}_n \in A_n\right) = \mathrm{P}\left(\mathrm{X}_1 \in A_1\right) \cdots \mathrm{P}\left(\mathrm{X}_n \in A_n\right). \tag{3.31}$$

Um conjunto qualquer, $\mathcal{V}$, de variáveis aleatórias definidas sobre $\Omega$ tomando valores em $\overline{\mathbb{R}}$ é uma família de variáveis aleatórias independentes se para toda subfamília finita $\mathrm{X}_1, ..., \mathrm{X}_n \in \mathcal{V}$, as variáveis aleatórias $\mathrm{X}_1, ..., \mathrm{X}_n$ são independentes.

**Exemplo 3.5.** Seja $(\{0,1,2,3\}, \mathcal{P}(\{0,1,2,3\}), \mathrm{P})$, onde

$$\mathrm{P}(\{0\}) = \frac{1}{9}, \quad \mathrm{P}(\{1\}) = \frac{2}{9}, \quad \mathrm{P}(\{2\}) = \frac{2}{9}, \quad \mathrm{P}(\{3\}) = \frac{4}{9}.$$

As variáveis

$$\mathrm{X}(0) = 0, \quad \mathrm{X}(1) = 0, \quad \mathrm{X}(2) = 1, \quad \mathrm{X}(3) = 1.$$

e

$$\mathrm{Y}(0) = 0, \quad \mathrm{Y}(1) = 1, \quad \mathrm{Y}(2) = 0, \quad \mathrm{Y}(3) = 1.$$

são independentes. (cf. Exercício 16.)

**Ensaios de Bernoulli**

**Proposição 3.4.** As projeções

$$\mathrm{X}_i : (\omega_1, \omega_2, \ldots) \in \mathbb{B} \longrightarrow \omega_i\,, \ i \in \mathbb{N} \tag{3.32}$$

são variáveis aleatórias independentes definidas sobre o espaço de probabilidade $(\mathbb{B}, \mathcal{F}_\infty, \mathrm{P}_p)$, sendo $p \in (0,1)$ qualquer. (cf. Exercício 18.)

**Definição 3.17.** Uma ***variável aleatória de Bernoulli*** é uma variável aleatória $\mathrm{X} : \Omega \to \mathbb{R}$, definida sobre um espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$, tomando unicamente dois valores, 0 e 1.

A lei de X é determinada por $p := \mathrm{P}\,(\mathrm{X} = 1)$, já que $\mathrm{P}\,(\mathrm{X} = 0) = 1 - p$.

**Exemplo.** As projeções definidas em (3.32) ou qualquer função característica, $\chi_A$, sendo $A \in \mathcal{F}$, são variaveis aleatórias de Bernoulli.

**Definição 3.18.** Uma ***sucessão ensaios de Bernoulli*** é qualquer sucessão de variáveis aleatórias de Bernoulli, $\mathrm{X}_i : \Omega \to \{0,1\}$, $i \in \mathbb{N}$, independentes, identicamente distribuidas, i.e. $\mathrm{P}\,(\mathrm{X}_i = 1) = p$, para todo $i \in \mathbb{N}$.

Retomaremos a noção de independência na subsecção 4.2, do próximo capítulo e mais tarde no Capítulo 5. Agora vamos dar exemplos de algumas das distribuições mais utilizadas em Estatística.

### 3.3.5 Exemplos de distribuições de probabilidade

As distribuições de probabilidade desta lista de exemplos ou são puramente discretas ou são absolutamente contínuas. Além disso, as densidades das variáveis aleatórias absolutamente contínuas são contínuas por partes. Portanto, neste contexto, a média e a variância, quando existem, obtêm-se através de séries ou de integrais que podem e serão calculados utilizando métodos do cálculo integral. Quando definamos o integral de Lebesgue, no Capítulo 4, daremos definições mais abrangentes de média e a variância de uma variável aleatória. A média e a variância de X serão denotadas por $m_{\mathrm{X}}$ e $V_{\mathrm{X}}$, respetivamente.

## Distribuições discretas

**Distribuições de Dirac.** A mais elementar das probabilidades é a distribuição de Dirac (cf. Definição 3.13). Nela não tem cabimento o "azar," pois toda a probabilidade concentra-se num ponto. No entanto, é a base para a construção de todas as outras distribuições discretas, que são combinações convexas de medidas de Dirac. (i.e. Combinações lineares com coeficientes positivos, cuja soma é 1.)

A média da distribuição de Dirac $\delta_a$ é $a$ e a variância é 0. Aliás

$$m_{\mathrm{X}} = a \cdot \delta_a(\{a\}) = a \;\text{ e }\; V_{\mathrm{X}} = (a - m_{\mathrm{X}})^2 \cdot \delta_a(\{a\}) = 0.$$

**Distribuições de Bernoulli.** Seja X uma variavel aleatória de Bernoulli. (cf. Definição 3.17.) A lei de X é

$$\nu_{\mathrm{X}}(\{1\}) = \mathrm{P}\,(\mathrm{X} = 1) = p \;\text{ e }\; \nu_{\mathrm{X}}(\{0\}) = \mathrm{P}\,(\mathrm{X} = 0) = 1 - \mathrm{P}\,(\mathrm{X} = 1) = 1 - p.$$

Quer dizer, para determinar a lei duma variavel aleatória de Bernoulli, basta dar um número $p \in (0,1)$, que representa a probabilidade do 1 e, ao mesmo tempo, a média. Efetivamente,

$$m_{\mathrm{X}} = 1 \cdot \mathrm{P}\,(\mathrm{X} = 1) + 0 \cdot \mathrm{P}\,(\mathrm{X} = 0) = p.$$

e, pondo $q = 1 - p$,

$$V_{\mathrm{X}} = (1 - m_{\mathrm{X}})^2 \cdot \mathrm{P}\,(\mathrm{X} = 1) + (0 - m_{\mathrm{X}})^2 \cdot \mathrm{P}\,(\mathrm{X} = 0) = (1-p)^2 \cdot p + p^2 \cdot q = pq.$$

A função de acumulação é

$$\mathrm{F}_{\mathrm{X}}(x) = q \cdot \mathrm{H}(x) + p \cdot \mathrm{H}(x - 1),$$

onde H é a função de Heaviside, definida em (3.28).

**Distribuição binomial.** Seja $\{\mathrm{X}_i\}_{i \in \mathbb{N}}$ uma sucessão de ensaios de Bernoulli de parâmetro $p \in (0,1)$ (cf. Defnição 3.18). A soma dos $n$ primeiros ensaios. $\mathrm{S}_n$, i.e.

$$\mathrm{S}_n := \mathrm{X}_1 + \cdots + \mathrm{X}_n,$$

representa o número de vezes que obtivemos um resultado positivo ao fim de $n$ experiências. A lei de $\mathrm{S}_n$ é a distribuição binomial $\mathfrak{B}_{p,n}$, dada por

$$\mathfrak{B}_{p,n}(k) = \mathrm{P}\,(\mathrm{S}_n = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, ..., n,$$

sendo $q = 1 - p$ e

$$\begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{k! \cdot (n-k)!}$$

o coeficiente binomial.

A média é

$$m_{\mathrm{S}_n} = \sum_{k=0}^{n} k \cdot \mathfrak{B}_{p,n}(k) = \sum_{k=0}^{n} k \cdot \frac{n!}{k! \cdot (n-k)!} p^k q^{n-k} = np.$$

A variância é

$$V_{\mathrm{S}_n} = \sum_{k=0}^{n} (k - np)^2 \cdot \mathfrak{B}_{p,n}(k) = \sum_{k=0}^{n} k^2 \cdot \frac{n!}{k! \cdot (n-k)!} p^k q^{n-k} \; - \; n^2 p^2 = npq.$$

O nome binomial deve-se ao binómio de Newton, já que

$$1 = (p+q)^n = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} p^k q^{n-k}.$$

**Distribuição geométrica.** Seja $\{\mathrm{X}_i\}_{i \in \mathbb{N}}$ uma sucessão de ensaios de Bernoulli de parámetro $p \in (0,1)$ (como no parágrafo anterior). Seja $\mathrm{T}_1$ a variável aleátoria que regista o número de 0's antes da ocorrência do primeiro 1 (número de fracassos até o primeiro sucesso), obtemos uma variável aleatória (cf. Exercício 24 do Capítulo 1) com distribuição geométrica, quer dizer

$$\mathfrak{b}_{p,1}(k) = \mathrm{P}\left(\mathrm{T}_1 = k\right) = p \cdot (1-p)^k, \quad k = 0, 1, 2, ...,$$

A média é

$$m_{\mathrm{T}_1} = \sum_{k=0}^{\infty} k \cdot \mathfrak{b}_{p,1}(k) = p \cdot \sum_{k=0}^{\infty} k \cdot (1-p)^k = \frac{1-p}{p}.$$

A variância é

$$V_{\mathrm{T}_1} = \sum_{k=0}^{\infty} \left(k - \frac{1-p}{p}\right)^2 \mathfrak{b}_{p,1}(k) = p \sum_{k=0}^{\infty} k^2 (1-p)^k \; - \; \left(\frac{1-p}{p}\right)^2 = \frac{1-p}{p^2}.$$

**Distribuição binomial negativa.** Generalizando o caso anterior, consideramos uma sucessão de ensaios de Bernoilli de parámetro $p \in (0,1)$, mas desta vez anotamos o lugar em que ocorre o $r-$ésimo 1 ($r-$ésimo sucesso). A variável aleatória obtida, denotada $\mathrm{T}_r$, segue a distribuição binomial negativa, cuja lei é

$$\mathfrak{b}_{p,r}(k) = \mathrm{P}\left(\mathrm{T}_r = k\right) = \frac{(k+r-1)!}{k! \cdot (r-1)!} \cdot p^r \cdot (1-p)^k, \quad k = 0, 1, ..., n.$$

A média é

$$\begin{aligned} m_{\mathrm{T}_r} &= \sum_{k=0}^{\infty} k \cdot \mathfrak{b}_{p,r}(k) \\ &= \sum_{k=0}^{\infty} k \cdot \frac{(k+r-1)!}{k! \cdot (r-1)!} \cdot p^r \cdot (1-p)^k \\ &= \frac{(1-p) \cdot p^r}{(r-1)!} \cdot \sum_{k=1}^{\infty} \frac{(k+r-1)!}{(k-1)!} \cdot (1-p)^{k-1} \\ &= \frac{r \cdot (1-p)}{p}. \end{aligned}$$

A variância é

$$\begin{aligned} V_{\mathrm{T}_r} &= \sum_{k=0}^{\infty} \left( k - \frac{(1-p)r}{p} \right)^2 \cdot \mathfrak{b}_{p,r}(k) \\ &= \sum_{k=0}^{\infty} k^2 \cdot \frac{(k+r-1)!}{k! \cdot (r-1)!} \cdot p^r \cdot (1-p)^k \; - \; \frac{(1-p)^2 r^2}{p^2} \\ &= \frac{(1-p)^2 p^r}{(r-1)!} \sum_{k=2}^{\infty} \frac{(k+r-1)!}{(k-2)!} (1-p)^{k-2} \; + \; \frac{(1-p)r}{p} - \frac{(1-p)^2 r^2}{p^2} \\ &= \frac{r \cdot (1-p)}{p^2}. \end{aligned}$$

**Distribuição de Poisson.** Utiliza-se para modelar o número de fenómenos independentes ocorridos num intervalo de tempo fixo. Por exemplo: O número de fotões emitidos por um material radioativo num segundo. Quantos clientes chegam, por hora, a uma fila de espera. Quantas chamadas recebem-se por minuto num call center. etc. (Compare com a lei exponencial.)

A distribuição de Poisson, de parámetro $\lambda > 0$, é dada por

$$\mathfrak{p}_\lambda(k) = \mathrm{e}^{-\lambda} \cdot \frac{\lambda^k}{k!}, \; k = 0, 1, \cdots . \tag{3.33}$$

$\lambda$ representa a média da "frequência" com que acontece o fenómeno que estamos a observar, num intervalo de tempo fixo. Efetivamente, a média de uma variável aleatória de Poisson de parámetro $\lambda$ é $\lambda$, já que

$$\sum_{k=0}^{\infty} k \cdot \mathfrak{p}_\lambda(k) = \mathrm{e}^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{k\lambda^k}{k!} = \lambda \cdot \mathrm{e}^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

A variância também é $\lambda$, como calculamos a seguir,

$$\sum_{k=0}^{\infty} (k-\lambda)^2 \cdot \mathfrak{p}_\lambda(k) = \mathrm{e}^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{k^2 \lambda^k}{k!} \; - \; \lambda^2 = \lambda.$$

**Distribuições absolutamente contínuas**

**Distribuição uniforme em** $[0,1]$**.** Não é outra coisa que a medida de Lebesgue, $\lambda$, discutida no Teorema 2.2. A função de acumulação de uma variável aleatória U, uniformemente distribuida em $[0,1]$ é

$$\mathrm{F_U}(x) = \begin{cases} 0 & \text{se } x < 0, \\ x & \text{se } 0 \leq x < 1, \\ 1 & \text{se } x \geq 1. \end{cases}$$

A "densidade" da distribuição uniforme é $\mathrm{F'_U}(x) = \chi_{(0,1)}$.

A média é

$$m_{\mathrm{U}} = \int_0^1 x\,dx = \frac{1}{2}.$$

A variância é

$$V_{\mathrm{U}} = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \frac{1}{12}.$$

***Observação.*** Esta distribuição é muito importante, pois serve para modelar um sem-número de fenómenos. Por exemplo, se os ensaios de Bernoulli são simétricos, i.e. se $p = \frac{1}{2}$, então

$$\psi : (\omega_1, \omega_2, ...) \in \mathbb{B} \longrightarrow \sum_{n=1}^{\infty} \frac{\omega_n}{2^n} \in [0,1]$$

é uma variável aleatória uniforme. (cf. Teorema 2.4.)

Por outro lado, $\left([0,1], \mathcal{B}_{[0,1]}, \lambda\right)$ serve como espaço probabilístico universal, já que qualquer medida de probabilidade é a medida imagem da distribuição uniforme através de uma variável aleatória apropriada, como o afirma a seguinte proposição.

**Proposição 3.5.** Dada qualquer distribuição de probabilidade $\nu$ sobre $(\mathbb{R}, \mathcal{B})$, existe uma variável aleatória X sobre o espaço de probabilidade $\left([0,1], \mathcal{B}_{[0,1]}, \lambda\right)$, tal que a lei de X é $\nu$. i.e.

$$\forall\,\nu,\ \exists\,\mathrm{X} : [0,1] \to \mathbb{R} \text{ tal que } \mathrm{X}_{\#}\lambda = \nu.$$

***Demonstração.*** Efetivamente, basta tomar como variável aleatória a inversa da função de acumulação, se esta for contínua e (estritamente) crescente. Senão o for, é necessário maoir cuidado na sua definição, mas a ideia é a mesma. A demonstração pormenorizada, quando a função de acumulação não é contínua e/ou não é estritamente crescente, fica como exercício. (cf. Exercício 20.)

**Distribuição normal.** A distribuição normal é determinada pela média e a variância, tradicionalmente representadas por $\mu$ e $\sigma^2$, respetivamente. ($\sigma$ é o desvio padrão, i.e. a raiz quadrada da variância). Denotamos $\eta(\mu,\sigma)$ a distribuição normal de média $\mu$ e variância $\sigma^2$. A distribuição normal, centrada ($\mu$=0) e reduzida ($\sigma$=1), tem como densidade à função de Gauss "standard", que é dada por

$$f_{\eta(0,1)} = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{x^2}{2}}.$$

A expresão mais simples da função de acumulação é em forma integral:

$$\mathrm{F}_{\eta(0,1)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{y^2}{2}} \mathrm{d}y,$$

já que não há uma fórmula prática para a primitiva da função de Gauss.

Não entanto, vamos dar uma fórmula alternativa, que troca o integral por um somatório, o que pode resultar útil para cálculos numéricos.

$$\mathrm{F}_{\eta(0,1)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{y^2}{2}} \mathrm{d}y = \frac{1}{2} + \frac{x}{\sqrt{2\pi}} \cdot \sum_{i=0}^{\infty} \frac{(-1)^i \cdot x^{2i}}{(2i+1) \cdot i! \cdot 2^i}.$$

A média e a variância da distribuição normal são, respetivamente,

$$m_{\eta(0,1)} = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{x^2}{2}} dx = 0 \text{ e } V_{\eta(0,1)} = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{x^2}{2}} dx = 1.$$

A distribuição $\eta(\mu,\sigma)$ obtem-se como a imagem da distribuição "standard" $\eta(0,1)$ através da transformação afim $T_{\mu,\sigma}(x) = \sigma x + \mu$. (cf. Lema 2.1.) i.e.

$$\eta(\mu,\sigma) = T_{\mu,\sigma\#}\eta(0,1).$$

***Observação.*** A posição 'central' que esta distribuição ocupa em Probabilidade e Estatística deve-se ao teorema do limite central, segundo o qual, dada uma sucessão de variaveis aleatórias independentes, identicamente distribuídas, possuindo momentos de ordem 1 e 2 (i.e. posuindo media e variância), a distribuição da soma das $n$ primeiras funções (convenientemente centrada e normalizada) tende, 'em lei,' para a distribuição $\eta\left(0, \sqrt{pq}\right)$, quando $n \to \infty$. No caso dos ensaios de Bernoulli, a demonstração do teorema do limite central utiliza unicamente a fórmula de Stirling.

**Distribuição de Cauchy.** A densidade da distribuição de Cauchy é

$$\mathrm{F}'_{\mathfrak{c}}(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

A função de acumulação é

$$\mathrm{F}_{\mathfrak{c}}(x) = \frac{1}{\pi} \cdot \int_{-\infty}^{x} \frac{1}{1+t^2}\, dt = \frac{1}{2} + \frac{1}{\pi} \arctan(x),$$

A distribuição de Cauchy foi estudada desde o primeiro quarto do século 19 por Siméon Denis Poisson, quem salientou o facto de ela não ter média (portanto também não tem variância). Efetivamente, a função $x/1+x^2$ não é integrável, já que a primitiva, $1/2 \cdot \ln(1+x^2)$, é divergente quando $x \to \infty$. Historicamente, este resultado invalidou o uso do Teorema do Limite Central, que Laplace teria feito em algúns dos seus trabalhos. Augustin-Louis Cauchy só aparece ligado a esta distribuição a partir da segunda mitade do século 19. A distribuição de Cauchy também é conhecida como distribuição de Lorentz, especialmente em Física.

A distribuição de Cauchy é a distribuição de probabilidade do lugar em que um movimento Browniano, partindo do ponto $(0,1) \in \mathbb{R}^2$, interceta o eixo $x$ (i.e. a reta $\{y=0\}$). Portanto, a medida de Cauchy de um intervalo inmerso no eixo $x$ é o ângulo com que o intervalo é "visto" desde o ponto $(0,1)$, normalizado (i.e. dividido por $\pi$).

A distribuição de Cauchy é igual à distribuição$-t$ de Student con um grau de liberdade.

**Distribuição$-t$ de Student.** A função de acumulação da distribuição$-t$ de Student com dois graus de liberdade é

$$\mathrm{F}_{\mathrm{t},2}(x) = \frac{1}{2} + \frac{x}{2\sqrt{2+x^2}}.$$

A densidade é

$$\mathrm{F}'_{\mathrm{t},2}(x) = \left(2+x^2\right)^{-\frac{3}{2}}.$$

À diferença da distribuição de Cauchy, esta distribuição já tem média, que por simetría é 0. Mas a variância não existe na mesma, porque $x^2 \cdot (2+x^2)^{-\frac{3}{2}} \sim |x|^{-1}$, quando $|x| \to \infty$.

**Distribuição exponencial.** A distribuição exponencial emprega-se na modelização de tempos de espera. Ou seja, o tempo que um fenómeno tarda em ocorrer. Por exemplo: O tempo que demora um material radioativo em emitir um fotão.

Quando chega o próximo cliente a uma fila de espera. Quando chega a próxima chamada a um call center. etc. (Compare com a lei de Poisson.)

A função de acumulação da distribuição exponencial de razão $\lambda > 0$ é

$$\mathrm{F}_{\mathfrak{e}(\lambda)}(x) = \begin{cases} 0 & \text{se } x < 0, \\ 1 - \mathrm{e}^{-\lambda x} & \text{se } x \geq 0, \end{cases}$$

cuja densidade é $\mathrm{F}'_{\mathfrak{e}(\lambda)}(x) = \chi_{(0,\infty)} \cdot \lambda \mathrm{e}^{-\lambda x}$.

**Distribuição Gamma.** As distribuições Gamma são uma família de probabilidades "exponenciais" com dois parámetros positivos, $\alpha$ e $\beta$, classicamente denominados forma e razão, respetivamente. Esta mesma família é por vezes indexada pelos parámetros forma e escala (inverso da razão). A densidade é dada pela fórmula

$$f_{\mathfrak{g}(\alpha,\beta)}(x) = \frac{\beta^\alpha x^{\alpha-1} \mathrm{e}^{-\beta x}}{\Gamma(\alpha)}, \; x \geq 0,$$

sendo $\Gamma$ a função Gamma. i.e.

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} \mathrm{e}^{-t} \mathrm{d}t.$$

A função de acumulação é o integral

$$\mathrm{F}_{\mathfrak{g}(\alpha,\beta)}(x) = \frac{\int_0^x \beta^\alpha t^{\alpha-1} \mathrm{e}^{-\beta t} \mathrm{d}t}{\Gamma(\alpha)} = \frac{\int_0^{\beta x} t^{\alpha-1} \mathrm{e}^{-t} \mathrm{d}t}{\Gamma(\alpha)} = \frac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)}, \; x \geq 0,$$

onde $\gamma$ é a função Gamma incompleta inferior. i.e.

$$\gamma(\alpha, x) = \int_0^x t^{\alpha-1} \mathrm{e}^{-t} \mathrm{d}t.$$

Quando $\alpha$ é um número natural, podemos expresar $\mathrm{F}_{\mathfrak{g}(\alpha,\beta)}(x)$ como o somatório:

$$\mathrm{F}_{\mathfrak{g}(\alpha,\beta)}(x) = 1 - \mathrm{e}^{-\beta x} \cdot \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!} = \mathrm{e}^{-\beta x} \cdot \sum_{i=\alpha}^{\infty} \frac{(\beta x)^i}{i!}, \; x \geq 0. \tag{3.34}$$

Se $\alpha = 1$, obtemos a distribuição exponencial.

***Observação.*** A pesar de que $\alpha$ é estritamente positivo, podemos substituir $\alpha = 0$ na última expressão (extremo-direito de (3.34)), obtendo-se a função de Heaviside. No entanto, $\alpha = 0$ não faz sentido na primeira fórmula de (3.34) e ainda menos na densidade $f_{\mathfrak{g}(\alpha,\beta)}$!

**Distribuição $\chi-$Quadrado.** A distribuição $\chi-$Quadrado, com $k$ graus de liberdade, é a distribuição da soma dos quadrados de k variáveis aleatórias, independentes, normais standard (i.e. com distribuição $\eta(0,1)$). Utiliza-se em Estatística para teste de hipóteses ou na construção de intervalos de confiança. A distribuição $\chi-$Quadrado, com $k$ graus de liberdade, é igual à distribuição Gamma com $\alpha = \frac{k}{2}$ e $\beta = \frac{1}{2}$.

## Exercícios

1. Mostre que

   (a) Se $A \subseteq \mathbb{R}$ é aberto em $\mathbb{R}$, então $A$ é aberto em $\overline{\mathbb{R}}$.

   (b) Se $A \subseteq \overline{\mathbb{R}}$ é aberto em $\overline{\mathbb{R}}$, então $A \cap \mathbb{R}$ é aberto em $\mathbb{R}$.

   (Lembre-se que um subconjunto de um espaço topológico é aberto se (por definição) é vizinhança de todos os seus pontos.)

2. Prove que $\overline{\mathbb{R}}$ é uma compactificação de $\mathbb{R}$. i.e. $\overline{\mathbb{R}}$ é compacto e $\mathbb{R}$ é un subconjunto denso em $\overline{\mathbb{R}}$.

3. (cf. Lema 3.2.) Mostre que a família

$$\left\{A \subseteq \overline{\mathbb{R}};\ A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\right\} \tag{3.35}$$

   é uma álgebra$-\sigma$ de $\overline{\mathbb{R}}$ e contém todos os abertos de $\overline{\mathbb{R}}$.

4. Demonstre o Lema 3.3, tendo atenção às diferenças com o Lema 3.4.

5. (cf. Lema 3.6.) Sejam $(X, \mathcal{F})$ e $(Y, \mathcal{G})$ dois espaços mensuráveis. Seja $\mathcal{S} \subseteq \mathcal{G}$ uma família geradora de $\mathcal{G}$, i.e. $\sigma[\mathcal{S}] = \mathcal{G}$. Prove que $f : X \to Y$ é $\mathcal{F}/\mathcal{G}-$mesurável se e só se

$$\forall\, A \in \mathcal{S},\ f^{-1}(A) \in \mathcal{F}.$$

6. (cf. Lema 3.9 e Teorema 3.1.) Seja $\{f_i\}_{i\in\mathbb{N}}$ uma sucessão de funções mensuráveis. Prove que as funções

$$x \in X \to \left(\sup_{i\in\mathbb{N}} f_i\right)(x) = \sup\{f_i(x);\ i \in \mathbb{N}\}$$

   e

$$x \in X \to \left(\inf_{i\in\mathbb{N}} f_i\right)(x) = \inf\{f_i(x);\ i \in \mathbb{N}\}$$

   são mensuráveis.

7. (cf. Lema 3.10.) Mostre que

$$\sigma[p_{X\times Y}] = \sigma[\{A \times B \times Z;\ A \in \mathcal{F},\ B \in \mathcal{G}\}] \subseteq \mathcal{F} \times \mathcal{G} \times \mathcal{H}.$$

8. (cf. Corolário 3.3.) Prove que uma função $f = (f_1, \cdots, f_n) : X \to \mathbb{R}^n$ é $\mathcal{F}/\mathcal{B}^n-$mensurável se e só se todas as componentes

$$f_i : X \to \mathbb{R},\ i = 1, \cdots, n,\ \text{são } \mathcal{F}/\mathcal{B}-\text{mensuráveis}.$$

9. (cf. Corolário 3.4.) Mostre os limites seguintes

$$\lim_{x\to-\infty} \mathrm{F_X}(x) = 0 \quad \text{e} \quad \lim_{x\to+\infty} \mathrm{F_X}(x) = 1.$$

10. Seja $\mathcal{F}_0$ a álgebra composta do conjunto vazio e dos conjuntos da forma

$$(a_1, b_1] \cup (a_2, b_2] \cup ... \cup (a_n, b_n],$$

onde $a_i, b_i \in \mathbb{R}$, $i = 1, ..., n$ e $a_1 < b_1 < a_2 < b_2 < ... < a_n < b_n$. Demonstre que $\sigma[\mathcal{F}_0] = \mathcal{B}$.

11. (cf. Teorema 3.3.) Seja $\mathcal{F}_0$ a álgebra do exercício 10 e seja $\nu_0$ a premedida definida por (3.20). Prove as duas afirmações seguintes.

    (a) *Monotonia*: Para todo $A, B \in \mathcal{F}_0$, se $A \subseteq B$, então

    $$\nu_0(A) \leq \nu_0(B).$$

    (b) *Subaditividade finita*: Para todo $A, B \in \mathcal{F}_0$,

    $$\nu_0(A \cup B) \leq \nu_0(A) + \nu_0(B).$$

12. (cf. Proposição 3.2.) Seja $\mathrm{F} : \mathbb{R} \to \mathbb{R}$ uma função crescente, contínua à direita e $\nu_\mathrm{F}$ a medida de Lebesgue-Stieltjes de F. Demonstre as seguintes propriedades

    (a) A medida $\nu_\mathrm{F}$ é $\sigma-$finita. Mas $\nu_\mathrm{F}$ é finita se e só se forem finitos ambos os limites

    $$\lim_{x\to-\infty} \mathrm{F}(x) \quad \text{e} \quad \lim_{x\to+\infty} \mathrm{F}(x).$$

    (b) F tem, no máximo, um número numerável de discontinuidades (saltos).

    (c) Para todo $x \in \mathbb{R}$,

    $$\nu_\mathrm{F}(\{x\}) = \mathrm{F}(x) - \lim_{y\to x^-} \mathrm{F}(y).$$

(d) O conjunto $\{x \in \mathbb{R};\ \nu_{\mathrm{F}}(\{x\}) > 0\}$. é finito ou numerável.

13. (cf. Proposição 3.3.) Seja $\nu$ é uma medida Boreliana que toma valores finitos nos conjuntos compactos. Prove que a função

$$\mathrm{F}_\nu(x) = \begin{cases} \nu([0,x]) & \text{se } x \geq 0, \\ -\nu((x,0)) & \text{se } x < 0 \end{cases}$$

é crescente, contínua à direita e, para todo $a, b \in \mathbb{R}$, $a < b$,

$$\nu((a,b]) = \mathrm{F}_\nu(b) - \mathrm{F}_\nu(a).$$

Prove, para além disso, que se F é outra função que verifica (3.25), então $\mathrm{F} - \mathrm{F}_\nu$ é constante.

14. Prove que as medidas discretas (definidas por (3.29)) são efetivamente medidas Borelianas, $\sigma-$finitas.

    Mais geralmente, mostre que a soma de qualquer família numerável de medidas, definidas sobre o mesmo espaço mensurável, é uma medida.

15. Mostre que a função $\mathrm{F}_{\nu_c} = \mathrm{F}_\nu - \mathrm{F}_{\nu_d}$, definida em (3.30), é contínua em $\mathbb{R}$.

16. Demonstre que as variáveis

$$\mathrm{X}(0) = 0, \quad \mathrm{X}(1) = 0, \quad \mathrm{X}(2) = 1, \quad \mathrm{X}(3) = 1.$$

e

$$\mathrm{Y}(0) = 0, \quad \mathrm{Y}(1) = 1, \quad \mathrm{Y}(2) = 0, \quad \mathrm{Y}(3) = 1,$$

definidas em $\{0,1,2,3\}$, são independentes, com respeito à medida

$$\mathrm{P}(\{0\}) = \frac{1}{9}, \quad \mathrm{P}(\{1\}) = \frac{2}{9}, \quad \mathrm{P}(\{2\}) = \frac{2}{9}, \quad \mathrm{P}(\{3\}) = \frac{4}{9}.$$

17. Encontre um espaço $(\Omega, \mathcal{F}, \mathrm{P})$ e três variáveis aleatórias X, Y e Z, definidas sobre $\Omega$, independentes dois a dois, mas que em conjunto não sejam independentes.

18. Prove que as projecções $\mathrm{X}_i$, definidas em (3.32), são de variáveis aleatórias independentes, qualquer que seja a medida $\mathrm{P}_p$ associada com o espaço mensurável $(\mathbb{B}, \mathcal{F}_\infty)$. (i.e. Qualquer que seja $p \in (0,1)$.)

19. Escreva a função de acumulação da distribuição de Poisson como uma função sobre $\mathbb{N}$ e compare-a com a função de acumulação da distribuição Gamma (para $\alpha \in \mathbb{N}$) e em particular com a distribuição exponencial.

20. Demonstre a Proposição 3.5.

# Capítulo 4

# Integral de Lebesgue

. À diferença do que acontece com o integral de Riemann, que é um processo no qual o domínio é dividido em intervalos pequenos, no integral de Lebesgue divide-se a imagem. Sendo as funções mensuráveis, as preimagens dos intervalos pequenos são conjuntos mensuráveis. Aplicamos a medida a estes conjuntos e somamos os resultados multiplicados pelo valor aproximado da função, no conjunto em questão. Desta forma, a integração reduz-se à teoria da medida desenvolvida no Capítulo 2. Mesmo assim existem muitos pontos técnicos que vamos esclarecer neste capítulo.

## 4.1 Integração de funções

### Função simples

**Definição 4.1.** Sendo $X$ um conjunto, uma função $s : X \to \mathbb{R}$ chama-se ***simples*** se a imagem, $s\,(X)$, é um subconjunto finito de $\mathbb{R}$. i.e. $s$ é simples se existe um conjunto finito $\{c_1, \cdots, c_n\} \subset \mathbb{R}$, tal que

$$\forall\, x \in X,\ \ s\,(x) \in \{c_1, \cdots, c_n\}\,.$$

**Exemplo 4.1.** Depois das funções constantes, a mais elementar das funções simples é a função característica de um subconjunto $A \subseteq X$. Pois, para todo $x \in X$, $\chi_A(x) \in \{0, 1\}$.

**Lema 4.1.** *Se $s$ e $t : X \to \mathbb{R}$ são funções simples e $a, b \in \mathbb{R}$, então $as + bt$ é uma função simples. Isto mostra que o conjunto das funções simples é um espaço* *vetorial real. Para além disso, o conjunto das funções características*

$$\{\chi_A;\, A \subseteq X\}$$

*gera linearmente o espaço vetorial das funções simples de $X$ em $\mathbb{R}$.*

***Demonstração.*** De facto, para todo $x \in X$, $(as + bt)(x)$ tem a forma $ac_i + bd_j$, onde $c_i$ e $d_j$ pertencem às imagens de $s$ e $t$, respetivamente. Portanto, $(as + bt)(x)$ só toma um número finito de valores.

Por outro lado, se $A \subseteq X$, a função característica $\chi_A(x)$ é simples. (cf. Exemplo 4.1.) Inversamente, todas as funções simples são combinações lineares de funções características. Efetivamente, pondo

$$s\,(X) := \{c_1, \cdots, c_n\} \quad \text{e} \quad A_i := s^{-1}\left(\{c_i\}\right),\ i = 1, \cdots, n, \tag{4.1}$$

então

$$s\,(x) = \sum_{i=1}^{n} c_i \cdot \chi_{A_i}\,(x)\,. \tag{4.2}$$

***Observação.*** Diferentes combinações lineares de funções características podem dar o mesmo resultado. Por exemplo, si $A, B \subseteq X$,

$$\chi_A + \chi_B = \chi_{A \triangle B} + 2 \cdot \chi_{A \cap B},$$

onde $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

Por este motivo, damos a seguinte definição.

**Definição 4.2.** Sendo $s$ uma função simples, a representacão (4.2) de $s$, onde $n$, $c_i$ e $A_i$, $i = 1, \cdots, n$ são os elementos definidos por (4.1), chama-se ***representação canónica*** de $s$. A representação canónica é única. (cf. Exercício 1.)

**Proposição 4.1.** Seja $s$ uma função simples e sejam $A_i$, $i = 1, ..., n$, as preimagens dos valores de $s$, como em (4.1). Então, a família $A_i$, $i = 1, \cdots, n$, é uma partição finita de $X$. Para além disso,

$$\sigma\,[s] = \sigma\,[\{A_1, \cdots, A_n\}]\,.$$

***Demonstração.*** A primeira observação resulta do facto que $s$ é uma função. A segunda é igualmente evidente.

**Proposição 4.2.** Uma função $s : X \to \mathbb{R}$ é simples se e só se $\sigma[s]$ é finita.

***Demonstração.*** A condição é necessária, pela Proposição 4.1. Para mostrar que também é suficiente, note que se $\sigma[s]$ é finita, então ela é gerada por uma partição finita, digamos $B_1, \cdots, B_n$. Para além disso, $s$ é constante em cada $B_i$, $i = 1, \cdots, n$. (cf. Exercício 2.)

### 4.1.1 Integral de funções não negativas

No resto deste capítulo, $(X, \mathcal{F}, \mu)$ representa um espaço de medida qualquer, ao qual referir-nos-emos prescindindo da sua reapresentação em cada ocorrência. Da mesma forma, faremos referência aos seus elementos, $X$, $\mathcal{F}$ e $\mu$, sem voltar a dizer a cada vez o que eles representam.

**Definição 4.3.** O conjunto das funções de $X$ em $\overline{\mathbb{R}}$, que são $\mathcal{F}/\mathcal{B}-$mensuráveis, denota-se $\mathcal{L}(X)$. O espaço vetorial das funções mensuráveis, simples (e portanto finitas) denota-se $\mathcal{L}^s(X)$ e $\mathcal{L}^{s,+}(X)$ é o cone vetorial das funções mensuráveis, simples, não negativas. i.e.

$$\mathcal{L}^s(X) = \{s \in \mathcal{L}(X)\,;\, \mathrm{Card}\,(s(X)) < \infty,\ s(X) \subset \mathbb{R}\}.$$

e

$$\mathcal{L}^{s,+}(X) = \{s \in \mathcal{L}^s(X)\,;\ \ s(X) \subset [0, \infty)\}.$$

**Nota 4.1.** $\mathcal{L}(X)$ não possui estrutura vetorial, pois $(f+g)(x)$ não está definida num ponto onde $f(x) = -g(x) = +\infty$ ou viceversa. Este inconveniente será ultrapassado mais a frente, mas unicamente para as funções integráveis.

#### Preintegral de funções não negativas, simples

**Definição 4.4.** Seja $s \in \mathcal{L}^{s,+}(X)$ e seja

$$s = \sum_{i=1}^{n} c_i \cdot \chi_{A_i}, \tag{4.3}$$

a sua representação canónica (cf. Definição 4.2). O ***preintegral*** de $s$ com respeito a $\mu$ é definido pela fórmula

$$\int_X^* s(x)\,\mu(dx) := \sum_{i=1}^{n} c_i \cdot \mu(A_i) = \sum_{x \in s(X)} s(x) \cdot \mu\left(s^{-1}(\{s(x)\})\right). \tag{4.4}$$

***Observação.*** Mesmo que o somatório em (4.4) é finito, o seu valor pode ser $+\infty$, caso tivermos ao mesmo tempo para algúm $i \in \{1, \cdots, n\}$ $\mu(A_i) = +\infty$ e $c_i \neq 0$. Por outro lado, pelas regras aritméticas (cf. Defnição 3.2, alinha 1), o termo com coeficiente nulo (caso existir) é nulo, independentemente da medida do conjunto associado ser infinita ou não. Pois, $0 \cdot \mu\left(s^{-1}(\{0\})\right) = 0$.

**Proposição 4.3.** Sendo $s$ e $t \in \mathcal{L}^{s,+}(X)$, temos as propriedades seguintes.

($a$) *Monotonia.* Se $s \leq t$, então

$$\int_X^* s(x)\, d\mu(x) \leq \int_X^* t(x)\, d\mu(x). \tag{4.5}$$

($b$) *Linearidade.* Para todo $a, b \in \mathbb{R}^+$,

$$\int_X^* (as + bt)(x)\, d\mu(x) = a \cdot \int_X^* s(x)\, d\mu(x) + b \cdot \int_X^* t(x)\, d\mu(x). \tag{4.6}$$

***Demonstração.*** Estas duas propriedades são uma consequência imediata do Lema 4.2 a seguir. (cf. Exercício 3)

***Observação.*** A alinha 2 não é propriamente "linearidade," por aplicar-se unicamente a combinações lineares com coeficientes **não negativos**, de funções não negativas. Aliás, $\mathcal{L}^{s,+}(X)$ nem sequer é um espaço vetorial, é um ***cone*** em $\mathcal{L}^s(X)$. No entanto, por ser um nombre descritivo, fica assim.

**Lema 4.2.** Seja $s \in \mathcal{L}^{s,+}(X)$ e seja

$$s(x) = \sum_{j=1}^{m} d_j \cdot \chi_{B_j}(x)$$

uma forma qualquer de representar $s$ como combinação linear de funções características, mensuráveis, com coeficientes não negativos. Então

$$\int_X^* s(x)\, d\mu(x) = \sum_{j=1}^{m} d_j \cdot \mu(B_j). \tag{4.7}$$

***Demonstração.*** Seja

$$s = \sum_{i=1}^{n} c_i \cdot \chi_{A_i}$$

a sua representação canónica (cf. Definição 4.2). Temos de provar que

$$\sum_{i=1}^{n} c_i \cdot \mu(A_i) = \sum_{j=1}^{m} d_j \cdot \mu(B_j). \tag{4.8}$$

Como $A_1, ..., A_n$ é uma partição de $X$,

$$\mu(B_j) = \sum_{i=1}^{n} \mu(B_j \cap A_i).$$

Então

$$\sum_{j=1}^{m} d_j \cdot \mu(B_j) = \sum_{j=1}^{m} d_j \cdot \sum_{i=1}^{n} \mu(B_j \cap A_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} d_j \cdot \mu(B_j \cap A_i).$$

Portanto, para provar (4.8), basta provar que

$$c_i \cdot \mu(A_i) = \sum_{j=1}^{m} d_j \cdot \mu(B_j \cap A_i),$$

qualquer que seja $i = 1, ..., n$. Ou seja, podemos supor que $s = \chi_A$. Também podemos supor, sem perda de generalidade, que $d_j \neq 0$, para todo $j = 1, ..., m$. Nesse caso, $B_j \subseteq A$ (cf. Exercício 4) e o nosso problema reduz-se a mostrar

$$\mu(A) = \sum_{j=1}^{m} d_j \cdot \mu(B_j).$$

Designando por $J_x$ o conjunto de índices $j \in \{1, ..., m\}$ tal que $x \in B_j$, a família $\{J_x\}_{x \in A}$ é finita, já que $J_x \in \mathcal{P}(\{1, ..., m\})$ que é um conjunto finito. Para fixar a notação, digamos que $\{J_x\}_{x \in A}$ tem $r$ elementos que denotamos $I_1, ..., I_r$. i.e.

$$\{J_x\}_{x \in A} = \{I_1, ..., I_r\}.$$

Observe que para todo $k = 1, ..., r$,

$$\sum_{j \in I_k} d_j = 1. \tag{4.9}$$

(cf. Exercício 4.) Além disso, pondo

$$C_k = \bigcap_{j \in I_k} B_j, \ k = 1, ..., r,$$

então $C_1, ..., C_r$ é uma partição finita de $A$. (cf. Exercício 5.)

Como $B_j \subset A$, para todo $j = 1, ..., m$, temos

$$\mu(B_j) = \sum_{k=1}^{r} \mu(B_j \cap C_k).$$

Então,

$$\sum_{j=1}^{m} d_j \cdot \mu(B_j) = \sum_{j=1}^{m} d_j \cdot \sum_{k=1}^{r} \mu(B_j \cap C_k) = \sum_{k=1}^{r} \sum_{j=1}^{m} d_j \cdot \mu(B_j \cap C_k). \tag{4.10}$$

Mas, para todo $k = 1, ..., r$, para todo $j = 1, ..., m$, $C_k \subseteq B_j$ ou $C_k \cap B_j = \emptyset$. Para além disso, $C_k \subseteq B_j$ se e só se $j \in I_k$. Logo,

$$\begin{aligned} \mu(B_j \cap C_k) &= 0, & \forall\, j \notin I_k, \\ \mu(B_j \cap C_k) &= \mu(C_k), & \forall\, j \in I_k. \end{aligned} \tag{4.11}$$

Substituindo (4.11) em (4.10), obtemos (utilizando (4.9) na terceira igualdade),

$$\sum_{j=1}^{m} d_j \cdot \mu(B_j) = \sum_{k=1}^{r} \sum_{j \in I_k} d_j \cdot \mu(C_k) = \sum_{k=1}^{r} \mu(C_k) \cdot \sum_{j \in I_k} d_j = \sum_{k=1}^{r} \mu(C_k) = \mu(A),$$

já que $C_1, ..., C_r$ é uma partição de $A$. :-)

### Integral de funções não negativas

Vamos agora estender a noção de preintegral às funções mensuráveis, não negativas, definidas sobre $(X, \mathcal{F}, \mu)$, i.e. $\mathcal{L}^+(X)$.

**Definição 4.5.** Seja $f \in \mathcal{L}^+(X)$. O ***integral*** de $f$ com respeito a $\mu$ define-se da seguinte forma.

$$\int_X f(x)\, d\mu(x) := \sup\left\{ \int_X^* s(x)\, d\mu(x)\, ;\ s \in \mathcal{L}^{s,+}(X),\ s \leq f \right\}. \tag{4.12}$$

**Nota 4.2.** (cf. Exercício 6.) Para todo $s \in \mathcal{L}^{s,+}(X)$,

$$\int_X s(x)\, d\mu(x) = \int_X^* s(x)\, d\mu(x).$$

**Proposição 4.4.** Sendo $f$, $g \in \mathcal{L}^+(X)$, temos as propriedades seguintes.

$(a)$ *Monotonía.* Se $f \leq g$, então

$$\int_X f(x)\, d\mu(x) \leq \int_X g(x)\, d\mu(x). \tag{4.13}$$

$(b)$ *Linearidade.* Para todo $a, b \in \mathbb{R}^+$,

$$\int_X (a \cdot f + b \cdot g)(x)\, d\mu(x) = a \cdot \int_X f(x)\, d\mu(x) + b \cdot \int_X g(x)\, d\mu(x). \tag{4.14}$$

***Demonstração.*** A alinha 1 é imediata, pois o integral de $g$ é o supremo de um conjunto maior do que o integral de $f$.

Para demonstrar a alinha 2, fixamos duas sucessões $\{s_n\}_{n\in\mathbb{N}}$ e $\{t_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^{s,+}(X)$ crescentes, tais que $s_n \to f$ e $t_n \to g$, pontualmente. (cf. Lema 4.3, na próxima subsecção.)

Por (4.6) (cf. Propoição 4.3), para todo $n \in \mathbb{N}$,

$$\int_X^* (a \cdot s_n + b \cdot t_n)(x)\, d\mu(x) = a \cdot \int_X^* s_n(x)\, d\mu(x) + b \cdot \int_X^* t_n(x)\, d\mu(x).$$

Tomando o limite em ambos os lados e aplicando (4.26) (cf. Lema 4.4, na próxima subsecção), obtemos o resultado procurado. :-)

***Observação.*** Como acontece na Proposição 4.3, a alinha 2 da Proposição 4.4 não é propriamente "linearidade," por aplicar-se unicamente a combinações lineares com coeficientes não negativos, de funções não negativas.

## 4.1.2 Continuidade absoluta

### Medidas definidas através de uma densidade

**Definição 4.6.** Seja $f \in \mathcal{L}^+(X)$ e seja $A \in \mathcal{F}$. Define-se o ***integral de $f$ sobre $A$***, respeito a $\mu$, por

$$\int_A f(x)\, \mathrm{d}\mu(x) := \int_X (\chi_A \cdot f)(x)\, \mathrm{d}\mu(x). \tag{4.15}$$

Como $\chi_A$ e $f \in \mathcal{L}^+(X)$, pela alinha *(iii)* do Lema 3.11, $\chi_A \cdot f \in \mathcal{L}^+(X)$ e portanto o integral em (4.15) tem sentido.

**Teorema 4.1.** *Sendo $f \in \mathcal{L}^+(X)$ fixo, a aplicação $f \cdot \mu$, definida por*

$$A \in \mathcal{F} \to (f \cdot \mu)(A) := \int_A f(x)\, d\mu(x) \tag{4.16}$$

*é uma medida sobre $(X, \mathcal{F})$.*

***Demonstração.*** Temos de verificar as três propriedades exigidas no Lema 2.5. Estas serão o resultado das propriedades de $f \cdot \mu$ que vamos enunciar e demonstrar a seguir.

*1. Conjuntos de medida nula.* Para todo $A \in \mathcal{F}$, tal que $\mu(A) = 0$, tem-se

$$(f \cdot \mu)(A) = \int_A f(x)\, d\mu(x) \leq +\infty \cdot \mu(A) = 0. \tag{4.17}$$

Em particular, verifica-se a alinha 1 do Lema 2.5, i.e. $(f \cdot \mu)(\emptyset) = 0$.

*2. Superaditividade para conjuntos disjuntos.* Sejam $A, B \in \mathcal{F}$, tal que $A \cap B = \emptyset$ e sejam $s, t \in \mathcal{L}^{s,+}(X)$, tais que $s \leq \chi_A \cdot f$ e $t \leq \chi_B \cdot f$. Como $A$ e $B$ são disjuntos, então $\chi_A + \chi_B = \chi_{A \cup B}$ e

$$s + t \leq \chi_A \cdot f + \chi_B \cdot f = \chi_{A \cup B} \cdot f.$$

Logo,

$$\int_X^* (s+t)(x)\, d\mu(x) \leq \int_{A \cup B} f(x)\, d\mu(x).$$

Por (4.6),

$$\int_X^* s(x)\, d\mu(x) + \int_X^* t(x)\, d\mu(x) \leq \int_{A \cup B} f(x)\, d\mu(x).$$

Tomando separadamente o supremo sobre as funções $s, t \in \mathcal{L}^{s,+}(X)$, que verificam $s \leq \chi_A \cdot f$ e $t \leq \chi_B \cdot f$, vemos que

$$\int_A f(x)\, d\mu(x) + \int_B f(x)\, d\mu(x) \leq \int_{A \cup B} f(x)\, d\mu(x).$$

Ou seja,

$$(f \cdot \mu)(A) + (f \cdot \mu)(B) \leq (f \cdot \mu)(A \cup B).$$

Portanto, fica verificada a alinha 2 do Lema 2.5.

*3. Subaditividade numerável*: Sendo $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F}$ e $A = \bigcup_{i=1}^{\infty} A_i$, queremos provar

$$\int_A f(x)\, d\mu(x) \leq \sum_{i=1}^{\infty} \int_{A_i} f(x)\, d\mu(x). \tag{4.18}$$

Fixemos $s \in \mathcal{L}^{s,+}(X)$, tal que $s \leq \chi_A \cdot f$. Seja $s = c_1 \cdot \chi_{B_1} + \cdots + c_n \cdot \chi_{B_n}$ uma representação qualquer de $s$, com coeficientes $c_1, ..., c_n$, todos estritamente positivos. Por (4.7) (cf. Lema 4.2),

$$\int_X^* s(x)\, d\mu(x) = c_1 \cdot \mu(B_1) + \cdots + c_n \cdot \mu(B_n). \tag{4.19}$$

Como $s \leq \chi_A \cdot f$, então $B_j \subseteq A$, para todo $j = 1, ..., n$. Logo,

$$B_j = B_j \cap A = \bigcup_{i=1}^{\infty} (B_j \cap A_i).$$

Por (2.12) (subaditividade numerável de $\mu$, cf. Lema 2.3),

$$\mu(B_j) \leq \sum_{i=1}^{\infty} \mu(B_j \cap A_i).$$

Conclui-se que

$$\int_X^* s(x)\, d\mu(x) = \sum_{j=1}^{n} c_j \cdot \mu(B_j) \leq \sum_{j=1}^{n} c_j \cdot \sum_{i=1}^{\infty} \mu(B_j \cap A_i). \tag{4.20}$$

Trocado a ordem de somação (a soma finita com a série), temos

$$\sum_{j=1}^{n} c_j \cdot \sum_{i=1}^{\infty} \mu(B_j \cap A_i) = \sum_{i=1}^{\infty} \sum_{j=1}^{n} c_j \cdot \mu(B_j \cap A_i) = \sum_{i=1}^{\infty} \int_X^* (\chi_{A_i} \cdot s)(x)\, d\mu(x), \tag{4.21}$$

já que, para todo $i \in \mathbb{N}$, $\chi_{A_i} \cdot s = c_1 \cdot \chi_{B_1 \cap A_i} + \cdots + c_n \cdot \chi_{B_n \cap A_i}$ e portanto,

$$\int_X^* (\chi_{A_i} \cdot s)(x)\, d\mu(x) = \sum_{j=1}^{n} c_j \cdot \mu(B_j \cap A_i).$$

De (4.20) e (4.21) deduz-se

$$\int_X^* s(x)\, d\mu(x) \leq \sum_{i=1}^{\infty} \int_X^* (\chi_{A_i} \cdot s)(x)\, d\mu(x) \leq \sum_{i=1}^{\infty} \int_{A_i} f(x)\, d\mu(x) \tag{4.22}$$

já que $\chi_{A_i} \cdot s \leq \chi_{A_i} \cdot f$, para todo $i \in \mathbb{N}$, e portanto,

$$\int_X^* (\chi_{A_i} \cdot s)(x)\, d\mu(x) \leq \int_{A_i} f(x)\, d\mu(x).$$

Tomando o supremo no lado esquerdo de (4.22), sobre todas as funções $s \in \mathcal{L}^{s,+}(X)$, $s \leq \chi_A \cdot f$, obtemos (4.18). :-)

**Corolário 4.1.** Seja $f \in \mathcal{L}^+(X)$ e $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{F}$ uma família **crescente**. Então, pondo $A = \bigcup_{i=1}^{\infty} A_i$, temos

$$\int_A f(x)\, d\mu(x) = \lim_{i\to\infty} \int_{A_i} f(x)\, d\mu(x). \tag{4.23}$$

***Demonstração.*** Basta aplicar (2.9) (cf. Lema 2.2) à medida $f \cdot \mu$. :-)

**Proposição 4.5.** Se a medida $f \cdot \mu$, definida por (4.16) (cf. Teorema 4.1), é $\sigma-$finita, então $\mu(\{x \in X;\, f(x) = \infty\}) = 0$. Inversamente se $\mu$ é $\sigma-$finita e $\mu(\{x \in X;\, f(x) = \infty\}) = 0$ então $f \cdot \mu$ é $\sigma-$finita.

***Demonstração.*** (cf. Exercício 7.)

**Definição 4.7.** Seja $\nu$ outra medida sobre o espaço $(X, \mathcal{F})$ (para além de $\mu$). Dizemos que $\nu$ é ***absolutamente contínua*** com respeito à $\mu$ se existir uma função $f \in \mathcal{L}^+(X)$, tal que

$$\forall\, A \in \mathcal{F},\ \nu(A) = \int_A f(x)\, d\mu(x).$$

A função $f$ chama-se ***densidade*** de $\nu$ com respeito à $\mu$.

**Exemplos**

Todas as distribuições contínuas da subsecção 3.3.5 são absolutamente contínuas com respeito à medida de Lebesgue de $\mathbb{R}$. As medidas discretas, da mesma subsecção, são absolutamente contínuas com respeito à medida de contagem de $\mathbb{N}$.

### 4.1.3 Convergência monótona

#### Aproximação de funções positivas

Primeiro vamos ver que qualquer função positiva (mensurável ou não) é o limite crescente de funções simples.

**Lema 4.3.** *Seja $X$ um conjunto qualquer e $f : X \to \overline{\mathbb{R}^+}$ $(= [0, +\infty])$. Então, existe uma sucessão crescente de funções simples, $\{s_n\}_{n\in\mathbb{N}}$, tal que*

$$\forall\, x,\ s_n(x) \to f(x),$$

*quando $n \to \infty$. Além disso, a convergência é uniforme sobre todo conjunto da forma $\{f \leq M\}$, sendo $M \in \mathbb{R}^+$ um número positivo qualquer.*

***Demonstração.*** Fixemos $n \in \mathbb{N}$ e dividamos o intervalo $[0, n)$ em $n \cdot 2^n$ intervalos, todos eles de comprimento igual a $2^{-n}$. i.e. dividimos $[0, n)$ nos intervalos

$$\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right), \ i = 0, 1, \cdots, n \cdot 2^n - 1.$$

Para $i = 0, 1, \cdots, n \cdot 2^n - 1$, denotamos $A_i$ o seguinte conjunto.

$$A_i = f^{-1}\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right) = \left\{x \in X; \ \frac{i}{2^n} \leq f(x) < \frac{i+1}{2^n}\right\}. \tag{4.24}$$

Para $i = n \cdot 2^n$, pomos

$$A_{n \cdot 2^n} = f^{-1}([n, \infty]) = \{x \in X; \ f(x) \geq n\}.$$

A função

$$s_n = \sum_{i=0}^{n \cdot 2^n} \frac{i}{2^n} \cdot \chi_{A_i}, \tag{4.25}$$

verifica

$$0 \leq f(x) - s_n(x) < \frac{1}{2^n},$$

sobre o conjunto $\{x \in X; \ f(x) \leq n\}$.

Isto mostra que $s_n \to f$ uniformemente sobre $\{x \in X; \ f(x) \leq M\}$, qualquer que seja $M \in \mathbb{R}^+$, fixo.

Em particular, $s_n(x) \to f(x)$, sobre o conjunto $\{x \in X; \ f(x) < \infty\}$. No conjunto $\{x \in X; \ f(x) = \infty\}$, $s_n = n$, para todo $n \in \mathbb{N}$, obtendo-se a mesma conclusão.

Para ver que $s_n$ é crescente, observe que para todo $i = 0, 1, \cdots, n \cdot 2^n - 1$,

$$A_i = \left\{x; \ \frac{2i}{2^{n+1}} \leq f(x) < \frac{2i+1}{2^{n+1}}\right\} \bigcup \left\{x; \ \frac{2i+1}{2^{n+1}} \leq f(x) < \frac{2i+2}{2^{n+1}}\right\}.$$

Fixando $i = 0, 1, \cdots, n \cdot 2^n - 1$, se $\frac{2i}{2^{n+1}} \leq f(x) < \frac{2i+1}{2^{n+1}}$, então

$$s_n(x) = \frac{2i}{2^{n+1}} = s_{n+1}(x).$$

E se $\frac{2i+1}{2^{n+1}} \leq f(x) < \frac{2i+2}{2^{n+1}}$, então

$$s_n(x) = \frac{2i}{2^{n+1}} < \frac{2i+1}{2^{n+1}} = s_{n+1}(x).$$

Em todo caso, $s_n(x) \leq s_{n+1}(x)$, sobre $\{x \in X; \ f(x) < n\}$. Sobre o conjunto $\{x \in X; \ f(x) \geq n\}$, $s_n = n \leq s_{n+1}$. :-)

### Aproximação de funções mensuráveis, positivas

Voltando ao espaço de medida $(X, \mathcal{F}, \mu)$, note que se $f \in \mathcal{L}^+(X)$, então os conjuntos $A_i$, definidos em (4.24), pertencem a $\mathcal{F}$ e portanto as funções $s_n$, definidas em (4.25), são mensuráveis.

A questão que se coloca de maneira natural é de saber se a aproximação pontual, construída em (4.25), serve também para aproximar o integral de $f$, i.e. se o preintegral da função $s_n$, definida em (4.25), converge para $\int_X f(x)\, d\mu(x)$. A resposta a esta pergunta é afirmativa, como pode-se ver no lema seguinte.

**Lema 4.4.** *Seja $f \in \mathcal{L}^+(X)$ e seja $\{s_n\}_{n\in\mathbb{N}}$ a sucessão definida por* (4.25). *Então*

$$\int_X f(x)\, d\mu(x) = \lim_{n\to\infty} \int_X^* s_n(x)\, d\mu(x)\,. \tag{4.26}$$

***Demonstração.*** Este lema é uma consequência imediata do Teorema 4.2, a seguir (teorema da convergência monótona). :-)

### Teorema da convergência monótona

**Teorema 4.2.** (Convergência monótona.) *Seja $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^+(X)$ uma sucessão crescente. Então*

$$\int_X \lim_{n\to\infty} f_n(x)\, d\mu(x) = \lim_{n\to\infty} \int_X f_n(x)\, d\mu(x) \tag{4.27}$$

***Demonstração.*** Como $\{f_n\}_{n\in\mathbb{N}}$ é crescente, para todo $x \in X$, o limite

$$f(x) := \lim_{n\to\infty} f_n(x)$$

existe em $\overline{\mathbb{R}^+}$. Por (3.7), $f \in \mathcal{L}^+(X)$.

Pela monotonía do integral (desigualdade (4.13)), $\int_X f_n(x)\, d\mu(x)$ é crescente. Logo, o limite (denotado $\alpha$) existe em $\overline{\mathbb{R}^+}$ e verifica

$$\alpha := \lim_{n\to\infty} \int_X f_n(x)\, d\mu(x) \leq \int_X f(x)\, d\mu(x)\,.$$

Para provar a desigualdade inversa, fixemos $s \in \mathcal{L}^{s,+}(X)$, tal que $s \leq f$, fixemos também $c \in (0,1)$ e consideremos os conjuntos

$$E_n := \{x \in X;\ f_n(x) \geq c \cdot s(x)\}\,.$$

Note que $E_n = (f_n - c \cdot s)^{-1}\left(\overline{\mathbb{R}^+}\right) \in \mathcal{F}$. Como

$$f_n \geq \chi_{E_n} \cdot f_n \geq c \cdot \chi_{E_n} \cdot s,$$

por (4.13) (monotonia do integral) e (4.6) ("linearidade")

$$\alpha \geq \int_X f_n(x)\, d\mu(x) \geq \int_{E_n} f_n(x)\, d\mu(x) \geq \int_{E_n}^{*} c \cdot s(x)\, d\mu(x) = c \cdot \int_{E_n}^{*} s(x)\, d\mu(x).$$

Aplicando (4.23) (cf. Corolário 4.1), obtemos

$$\alpha \geq c \cdot \lim_{n\to\infty} \int_{E_n}^{*} s(x)\, d\mu(x) = c \cdot \int_X^{*} s(x)\, d\mu(x),$$

já que $X = \bigcup_{n=1}^{\infty} E_n$ (cf. Exercício 8).

Como $c \in (0,1)$ é arbitrário, conclui-se que

$$\alpha \geq \int_X^{*} s(x)\, d\mu(x).$$

Tomando o supremo sobre $s \in \mathcal{L}^{s,+}(X)$, $s \leq f$, por (4.12) (cf. Definição 4.5)

$$\alpha \geq \int_X f(x)\, d\mu(x). \quad :-)$$

**Corolário 4.2.** Seja $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^+(X)$. Então,

$$\int_X \sum_{n=1}^{\infty} f_n(x)\, d\mu(x) = \sum_{n=1}^{\infty} \int_X f_n(x)\, d\mu(x). \tag{4.28}$$

### 4.1.4 Funções Integráveis

**Definição 4.8.** Seja $f \in \mathcal{L}(X)$. O ***integral*** de $f$ com respeito a $\mu$ é

$$\int_X f(x)\, d\mu(x) = \int_X f^+(x)\, d\mu(x) - \int_X f^-(x)\, d\mu(x), \tag{4.29}$$

sob a condição de, no mínimo, um dos integrais da direita ser finito, pois de outra forma, (4.29) não teria sentido. As funções $f^+$ e $f^-$ denotam a parte positiva e a parte negativa de $f$, definidas em (3.6) (cf. Corolário 3.1).

**Definição 4.9.** Uma função $f \in \mathcal{L}(X)$ é $\mu-$***integrável*** (ou ***integrável***, quando $\mu$ está subentendida no contexto) se $\int_X |f(x)|\, d\mu(x) < \infty$. O conjunto das funções $\mu-$integráveis denota-se $\mathcal{L}^1(X, \mu)$. i.e.

$$\mathcal{L}^1(X, \mu) = \left\{ f \in \mathcal{L}(X);\ \int_X |f(x)|\, d\mu(x) < \infty \right\}. \tag{4.30}$$

**Nota 4.3.** Seja $f \in \mathcal{L}(X)$, então

$$f \in \mathcal{L}^1(X, \mu) \text{ se e só se } f^+ \text{ e } f^- \in \mathcal{L}^1(X, \mu).$$

Em particular, toda função integrável é igual à diferença de duas funções positivas, integráveis.

### O espaço vetorial normado $L^1(X, \mu)$

Vamos definir operações aritméticas entre funções integráveis. Como comentamos na Nota 4.1, $\mathcal{L}(X)$ não possui estrutura vetorial, pois $(f+g)(x)$ não está definida nos pontos onde $f(x) = -g(x) = +\infty$ ou viceversa. Por este motivo, provaremos previamente que as funções integráveis são finitas em quase todo ponto e portanto a soma de duas funções integráveis tem sentido em quase todo ponto.

**Proposição 4.6.** Para todo $f \in \mathcal{L}^1(X, \mu)$, o conjunto $\{x \in X;\ |f(x)| = \infty\}$ é desprezável (cf. Definição 2.14). i.e.

$$\mu(\{x \in X;\ |f(x)| = \infty\}) = 0. \tag{4.31}$$

***Demonstração.*** Como $\infty \cdot \chi_{\{x \in X;\, |f(x)| = \infty\}} \leq |f|$, aplicando a desigualdade (4.13), obtemos

$$\infty \cdot \mu(\{x \in X;\ |f(x)| = \infty\}) \leq \int_X |f(x)|\, d\mu(x) < \infty.$$

Pelas regras 1 e 2 da Definição 3.2, a desigualdade anterior só pode ser satisfeita se $\mu(\{x \in X;\ |f(x)| = \infty\}) = 0$. :-) (compare com o Exercício 7.)

**Corolário 4.3.** Sendo $f$ e $g \in \mathcal{L}^1(X, \mu)$, a função $f + g : x \in X \to f(x) + g(x)$ está definida em $\mu-$quase todo ponto.

**Lema 4.5.** (Desigualdade de Chebyshev.) *Para todo $f \in \mathcal{L}^1(X, \mu)$, para todo $\alpha > 0$,*

$$\mu(\{x \in X;\ |f(x)| \geq \alpha\}) \leq \frac{1}{\alpha} \int_X |f(x)|\, d\mu(x). \tag{4.32}$$

***Demonstração.*** Pondo $A_\alpha = \{x \in X;\ |f(x)| \geq \alpha\}$, temos

$$\alpha \cdot \chi_{A_\alpha} \leq \chi_{A_\alpha} \cdot |f| \leq |f|.$$

Por (4.13) (monotonia do integral, cf. Proposição 4.4),

$$\alpha \cdot \mu(A_\alpha) \leq \int_{A_\alpha} |f(x)|\, d\mu(x) \leq \int_X |f(x)|\, d\mu(x).$$

Dividindo por $\alpha$, obtemos (4.32). :-)

***Observação.*** A primeira aplicação da desigualdade de Chebyshev será mostrar que o integral das funções integráveis e mesmo a integrabilidade delas, não altera se alterarmos o valor da função num conjunto de medida nula.

**Lema 4.6.** *Para todo* $f \in \mathcal{L}(X, \mu)$,

$\int_X |f(x)|\, d\mu(x) = 0$ *se e só se* $f(x) = 0$, *em quase todo ponto* $x$.

***Demonstração.*** Pela desigualdade de Chebyshev (4.32), para todo $\alpha > 0$,

$$\mu(\{x \in X; |f(x)| \geq \alpha\}) \leq \frac{1}{\alpha} \int_X |f(x)|\, d\mu(x) = 0.$$

Como,

$$\{x \in X; |f(x)| > 0\} = \bigcup_{n=1}^{\infty} \left\{x \in X; |f(x)| \geq \frac{1}{n}\right\},$$

então, por (2.9),

$$\mu(\{x \in X; |f(x)| > 0\}) = \lim_{n\to\infty} \mu\left(\left\{x \in X; |f(x)| \geq \frac{1}{n}\right\}\right) = 0.$$

Inversamente, como

$$|f(x)| \leq \infty \cdot \chi_{\{x \in X; |f(x)| > 0\}},$$

pela monotonia do integral (cf. Proposição 4.4, desigualdade (4.13)),

$$\int_X |f(x)|\, d\mu(x) \leq \infty \cdot \mu(\{x \in X; |f(x)| > 0\}) = 0. \ :-)$$

**Corolário 4.4.** Sejam $f$ e $g \in \mathcal{L}(X)$, tais que $f(x) = g(x)$ em $\mu-$quase todo ponto $x \in X$. i.e.

$$\mu(\{x \in X;\, |f(x) - g(x)| > 0\}) = 0.$$

Então

$$\int_X |f(x)|\, d\mu(x) = \int_X |g(x)|\, d\mu(x).$$

(Podendo ambos os lados serem iguais a $+\infty$.)

**Proposição 4.7.** Sejam $f$ e $g \in \mathcal{L}(X)$, a relação $\sim$, definida por

$$f \sim g \ \text{ se } \ \mu(\{x \in X;\, |f(x) - g(x)| > 0\}) = 0 \tag{4.33}$$

($f \sim g$ se $f(x) = g(x)$ em $\mu-$quase todo ponto $x \in X$), é uma relação de equivalência sobre $\mathcal{L}(X)$.

***Demonstração.*** (cf. Exercícios 9.)

**Definição 4.10.** O conjunto das classes de equivalência das funções integráveis denota-se $L^1(X, \mu)$. i.e.

$$L^1(X, \mu) = \mathcal{L}^1(X, \mu) / \sim .$$

**Nota 4.4.** Como $f \sim g$ se e só se $f^+ \sim g^+$ e $f- \sim g^-$, pelo Corolário 4.4, o integral (se existir) tem o mesmo valor, qualquer que seja o elemento da clase de equivalência. i.e. O integral é uma aplicação definida em $L^1(X, \mu)$.

Além disso, sendo $f$ e $g \in \mathcal{L}^1(X, \mu)$, pelo Corolário 4.3, $f + g$ está definida em quase todo ponto. Isto é suficiente para definir a classe de equivalência correspondente à $f + g$, já que qualquer função mensurável $h \in \mathcal{L}(X \setminus E)$, definida parcialmente num conjunto $X \setminus E$, sendo $E \in \mathcal{F}$ um conjunto de medida nula, define claramente uma classe de equivalência, pois qualquer extensão (mensurável) de $h$ pertence à mesma classe. Ademais, a classe definida por $f + g$ não depende individualmente de $f$ e $g$, mas sim da classe de $f$ e da classe de $g$. Quer dizer que podemos definir a soma de elementos (classes) de $L^1(X, \mu)$.

Exploramos estes factos na próxima definição.

**Definição 4.11.** Sejam $f$ e $g \in \mathcal{L}^1(X, \mu)$ e $\lambda \in \mathbb{R}$. Denotando $\overline{f}$ e $\overline{g} \in L^1(X, \mu)$ as classes de equivalência respetivas, define-se

(a) $\overline{f} + \overline{g} := \overline{f + g}$.

(b) $\lambda \cdot \overline{f} := \overline{\lambda \cdot f}$.

(c) $\int_X \overline{f} d\mu := \int_X f(x)\, d\mu(x)$, sendo $f \in \overline{f}$ qualquer.

(d) $\left\|\overline{f}\right\|_1 := \int_X |f(x)|\, d\mu(x)$, sendo $f \in \overline{f}$ qualquer.

**Nomenclatura.** Em diante, denotaremos aos elementos de $L^1(X,\mu)$ da mesma forma que aos elementos de $\mathcal{L}^1(X,\mu)$, i.e. sem barras. Alias, às classes de equivalência chamaremos funções integráveis. Ou seja, a partir de agora, uma função integrável é um elemento $f \in L^1(X,\mu)$.

**Proposição 4.8.** $L^1(X,\mu)$, junto com as operações definidas nas alinhas (a) e (b) da Definição 4.11, é um espaço vetorial sobre $\mathbb{R}$. Para além disso, a aplicação $\|f\|_1$ é uma norma sobre $L^1(X,\mu)$.

***Demonstração.*** É fácil ver que a soma de classes de equivalência e o produto por escalar, definidas nas alinhas (a) e (b) da Definição 4.11, verificam as propriedades das operações de soma e produto por escalar de um espaço vetorial sobre $\mathbb{R}$.

Para provar que $\|f\|_1$ é uma norma sobre $L^1(X,\mu)$, vemos que

(i) Como $|f(x)| \geq 0$, pela monotonia do integral,

$$\|f\|_1 = \int_X |f(x)|\, d\mu(x) \geq 0.$$

Pelo Lema 4.6, $\|f\|_1 = 0$ se e só se $f = \overline{0}$. i.e. $f = 0$ em quase todo ponto.

(ii) Por (4.14) (cf. Proposição 4.4),

$$\|\alpha \cdot f\|_1 = \int_X |\alpha \cdot f(x)|\, d\mu(x) = |\alpha| \cdot \int_X |f(x)|\, d\mu(x) = |\alpha| \cdot \|f\|_1 .$$

(iii) Finalmente, como $|f+g| \leq |f| + |g|$, por (4.13) e (4.14) (monotonia e "linearidade" do integral, cf. Proposição 4.4),

$$\int_X |f(x) + g(x)|\, d\mu(x) \leq \int_X |f(x)|\, d\mu(x) + \int_X |g(x)|\, d\mu(x) .$$

Logo,

$$\|f+g\|_1 \leq \|f\|_1 + \|g\|_1 .$$

### Linearidade do integral sobre $L^1(X,\mu)$

A propriedade (4.14), chamada de "linearidade" (cf. Proposição 4.4), não é propriamente assim devido a que o conjunto $\mathcal{L}^+(X)$ nem sequer tem uma estrutura de espaço vetorial. Agora que hemos definido uma estrutura de espaço vetorial sobre $L^1(X,\mu)$, vamos mostrar a propriedade de linearidade do integral sobre este espaço.

**Lema 4.7.** *Sejam $f$ e $g \in L^1(X,\mu)$ e sejam $a$ e $b \in \mathbb{R}$. Então*

$$\int_X (a\cdot f + b\cdot g)(x)\, d\mu(x) = a\cdot \int_X f(x)\, d\mu(x) \; + \; b\cdot \int_X g(x)\, d\mu(x)\,. \tag{4.34}$$

***Demonstração.*** Se $a \geq 0$ e $b = 0$, então $(a\cdot f)^+ = a\cdot f^+$ e $(a\cdot f)^- = a\cdot f^-$. Por (4.29) (definição de integral) e por (4.14) (cf. Proposição 4.4), temos

$$\begin{aligned}
\textstyle\int_X a\cdot f(x)\, d\mu(x) \; &= \textstyle\int_X (a\cdot f)^+(x)\, d\mu(x) - \int_X (a\cdot f)^-(x)\, d\mu(x) \\
&= \textstyle\int_X a\cdot f^+(x)\, d\mu(x) - \int_X a\cdot f^-(x)\, d\mu(x) \\
&= \textstyle a\cdot \int_X f^+(x)\, d\mu(x) - a\cdot \int_X f^-(x)\, d\mu(x) \\
&= \textstyle a\cdot \int_X f(x)\, d\mu(x)\,.
\end{aligned}$$

Se $a < 0$ e $b = 0$, então $(a\cdot f)^+ = -a\cdot f^-$ e $(a\cdot f)^- = -a\cdot f^+$. Aplicando igualmente (4.29) e (4.14), a conclusão é a mesma.

Se $a = b = 1$, pomos $h := f + g = f^+ - f^- + g^+ - g^-$ e rearrangamos os termos no lado da equação em que ficam positivos.

$$h^+ + f^- + g^- = f^+ + g^+ + h^-.$$

Integrando ambos lados e aplicando (4.14), pois todas as funções são positivas,

$$\int_X h^+(x)\, d\mu(x) + \int_X f^-(x)\, d\mu(x) + \int_X g^-(x)\, d\mu(x)$$

$$= \int_X f^+(x)\, d\mu(x) + \int_X g^+(x)\, d\mu(x) + \int_X h^-(x)\, d\mu(x)\,.$$

Como todos os termos são finitos, podemos recombinar os integrais das partes positivas e negativas de cada função, obtendo-se, pela Definição 4.8, o resultado (4.34).

Para todos os outros valores de $a$ e $b$ podemos concluir a partir destes três casos. :-)

**Corolário 4.5.** *(Monotonia do integral sobre $L^1(X,\mu)$)* Se $f$ e $g \in L^1(X,\mu)$ e se $f \le g$ (i.e. $g - f \ge 0$ em quase todo ponto), então

$$\int_X f(x)\, d\mu(x) \le \int_X g(x)\, d\mu(x)\,. \tag{4.35}$$

***Demonstração.*** Por (4.13) (monotonia do integral para funções positivas, cf. Proposição 4.4) e pelo Lema 4.7,

$$0 \le \int_X (g-f)(x)\, d\mu(x) = \int_X g(x)\, d\mu(x) - \int_X f(x)\, d\mu(x).\ \ :-)$$

**Mudança de variável**

**Teorema 4.3.** *Seja $(X,\mathcal{F},\mu)$ um espaço de medida, $(Y,\mathcal{G})$ um espaço mensurável e $\psi : (X,\mathcal{F}) \to (Y,\mathcal{G})$ uma aplicação mensurável. Então, para toda função mensurável, $f : (Y,\mathcal{G}) \to (\mathbb{R},\mathcal{B})$,*

$$f \in L^1(Y,\psi_{\#}\mu) \text{ se e só se } f\circ\psi \in L^1(X,\mu).$$

*Para além disso, pondo $\nu = \psi_{\#}\mu$,*

$$\int_Y f(y)\, d\nu(y) = \int_X f\circ\psi(x)\, d\mu(x) \tag{4.36}$$

***Demonstração.*** Mostraremos primeiro o teorema para as funções simples, não negativas. Seja $s = c_1\cdot\chi_{B_1} + \cdots + c_r\cdot\chi_{B_r} \in \mathcal{L}^{s,+}(Y)$. Então,

$$\begin{aligned}
\int_Y s(y)\, d\nu(y) &= \int_Y (c_1\chi_{B_1} + \cdots + c_r\chi_{B_r})\, d\nu \\
&= c_1\cdot\psi_{\#}\mu(B_1) + \cdots + c_r\cdot\psi_{\#}\mu(B_r) \\
&= c_1\cdot\mu\left(\psi^{-1}(B_1)\right) + \cdots + c_r\cdot\mu\left(\psi^{-1}(B_r)\right) \\
&= \int_X \left(c_1\chi_{\psi^{-1}(B_1)} + \cdots + c_r\chi_{\psi^{-1}(B_r)}\right) d\mu \\
&= \int_X (c_1\chi_{B_1} + \cdots + c_r\chi_{B_r})\circ\psi(x)\ d\mu(x). \\
&= \int_X s\circ\psi(x)\ d\mu(x).
\end{aligned}$$

Agora assumimos que $f \in \mathcal{L}^+(X)$ e arranjamos uma sucessão $\{s_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^{s,+}(Y)$, tal que $s_n \nearrow f$. Note que $\{s_n\circ\psi\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^{s,+}(X)$ e $s_n\circ\psi \nearrow f\circ\psi$.

Por (4.26) (cf. Lema 4.4),

$$\int_Y f(y)\, d\nu(y) = \lim_{n\to\infty} \int_Y^* s_n(y)\, d\nu(y)$$

$$= \lim_{n\to\infty} \int_X^* s_n \circ \psi(x)\, d\mu(x)$$

$$= \int_X f \circ \psi(x)\, d\mu(x).$$

Para $f \in \mathcal{L}(X)$ qualquer, aplicamos (4.36) a $f^+$ e $f^-$ separadamente. :-)

### 4.1.5 Convergência dominada

Vamos terminar esta secção com um teorema que permite manipular o integral e os limites. Mas primeiro vamos dar um lema fundamental.

**Lema 4.8.** (Lema de Fatou) *Para toda sucessão* $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^+(X)$,

$$\int_X \liminf_{n\to\infty} f_n(x)\, d\mu(x) \leq \liminf_{n\to\infty} \int_X f_n(x)\, d\mu(x). \tag{4.37}$$

***Demonstração.*** Define-se

$$g_n := \inf_{i\geq n} f_i.$$

Como $g_n$ é crescente, por (4.27) (teorema da convergência monótona),

$$\int_X \lim_{n\to\infty} g_n(x)\, d\mu(x) = \lim_{n\to\infty} \int_X g_n(x)\, d\mu(x). \tag{4.38}$$

Como $g_n \leq f_i$, para todo $i \geq n$, por (4.13) (monotonía do integral),

$$\int_X g_n(x)\, d\mu(x) \leq \inf_{i\geq n} \int_X f_i(x)\, d\mu(x).$$

Tomando o limite quando $n \to \infty$,

$$\lim_{n\to\infty} \int_X g_n(x)\, d\mu(x) \leq \lim_{n\to\infty} \left( \inf_{i\geq n} \int_X f_i(x)\, d\mu(x) \right).$$

Por (4.38),

$$\int_X \lim_{n\to\infty} g_n(x)\, d\mu(x) \leq \lim_{n\to\infty} \left( \inf_{i\geq n} \int_X f_i(x)\, d\mu(x) \right).$$

Como

$$\lim_{n\to\infty} g_n = \lim_{n\to\infty}\left(\inf_{i\geq n} f_i\right) = \liminf_{n\to\infty} f_n$$

e de igual forma,

$$\lim_{n\to\infty}\left(\inf_{i\geq n}\int_X f_i(x)\,d\mu(x)\right) = \liminf_{n\to\infty}\int_X f_i(x)\,d\mu(x),$$

o lema de Fatou fica demonstrado. :-)

**Teorema 4.4.** (Teorema da convergência dominada) *Seja* $\{f_n\}_{n\in\mathbb{N}} \subseteq L^1(X,\mu)$*, uma sucessão convergente em quase todo ponto. Seja* $g \in L^1(X,\mu)$*, tal que*

$$\forall\, n\in\mathbb{N},\ |f_n(x)| \leq g(x),\ \textit{para } \mu\textit{-quase todo } x\in X.$$

*Então,*

$$\int_X \lim_{n\to\infty} f_n(x)\,d\mu(x) = \lim_{n\to\infty}\int_X f_n(x)\,d\mu(x). \tag{4.39}$$

***Demonstração.*** (Note que parte da conclusão é que o limite da direita existe.) A sucessão $\{g-f_n\}_{n\in\mathbb{N}}$ é uma sucessão de funções positivas. Aplicando (4.37) (lema de Fatou), obtemos

$$\int_X \liminf_{n\to\infty}(g-f_n)d\mu \leq \liminf_{n\to\infty}\int_X (g-f_n)d\mu.$$

Pela linearidade do integral,

$$\int_X g d\mu - \int_X \lim_{n\to\infty} f_n d\mu \leq \int_X g d\mu - \limsup_{n\to\infty}\int_X f_n d\mu$$

Como $\int_X g d\mu < \infty$, podemos restar $\int_X g d\mu$ em ambos os lados. Portanto,

$$\limsup_{n\to\infty}\int_X f_n d\mu \leq \int_X \lim_{n\to\infty} f_n d\mu.$$

Repetimos o processo anterior com $\{g+f_n\}_{n\in\mathbb{N}}$, que também é uma sucessão de funções positivas. Obtemos

$$\int_X \lim_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty}\int_X f_n d\mu.$$

Como $\liminf_{n\to\infty}\int_X f_n d\mu \leq \limsup_{n\to\infty}\int_X f_n d\mu$, a sucessão $\left\{\int_X f_n d\mu\right\}_{n\in\mathbb{N}}$ é convergente e o limite verifica (4.39). :-)

**Corolário 4.6.** Seja $\{f_n\}_{n\in\mathbb{N}} \subseteq L^1(X,\mu)$ tal que

$$\sum_{n=1}^{\infty}\int_X |f_n(x)|\, d\mu(x) < \infty.$$

Então, $\sum_{n=1}^{\infty} f_n \in L^1(X,\mu)$ e

$$\int_X \sum_{n=1}^{\infty} f_n(x)\, d\mu(x) = \sum_{n=1}^{\infty}\int_X f_n(x)\, d\mu(x). \tag{4.40}$$

***Demonstração.*** Por (4.28) (teorema da convergência monótona),

$$\int_X \sum_{n=1}^{\infty} |f_n(x)|\, d\mu(x) = \sum_{n=1}^{\infty}\int_X |f_n(x)|\, d\mu(x) < \infty.$$

Logo,

$$\sum_{n=1}^{\infty} |f_n| \in L^1(X,\mu).$$

Então, por (4.31), para quase todo $x \in X$,

$$\sum_{n=1}^{\infty} |f_n(x)| < \infty.$$

Ou seja, existe um conjunto $A \subset X$, $\mu(A) = 0$, tal que para todo $x \in X \setminus A$, $\sum_{n=1}^{\infty} f_n(x)$ é absolutamente convergente. i.e.

$$\sum_{n=1}^{\infty} f_n(x) = \lim_{N\to\infty}\sum_{n=1}^{N} f_n(x)$$

existe em quase todo $x \in X$. Para além disso,

$$\left|\sum_{n=1}^{N} f_n(x)\right| \le \sum_{n=1}^{\infty} |f_n(x)|.$$

Pela linearidade do integral,

$$\int_X \sum_{n=1}^{N} f_n(x)\, d\mu(x) = \sum_{n=1}^{N}\int_X f_n(x)\, d\mu(x).$$

Por (4.39) (cf. Teorema 4.4), obtemos (4.40). :-)

## 4.2 Variáveis aleatórias integráveis

Nesta secção, por se tratar de espaços de probabilidade, mudaremos a notação de $(X,\mathcal{F},\mu)$ para a notação clássica $(\Omega,\mathcal{F},\mathrm{P})$.

### 4.2.1 Média e Variância

**Definição 4.12.** Seja $X : \Omega \to \mathbb{R}$ uma variável aleatória integrável. i.e. $X \in L^1(\Omega, P)$. Define-se a ***média*** ou ***esperança matemática*** de X, denotada E(X), pela fórmula

$$E(X) := \int_\Omega X(\omega)\, dP(\omega). \tag{4.41}$$

Se para além disso, $X^2 \in L^1(\Omega, P)$, define-se também a ***variância*** de X,

$$\mathrm{Var}(X) := \int_\Omega (X(\omega) - E(X))^2\, dP(\omega) = \int_\Omega X^2(\omega)\, dP(\omega) - (E(X))^2 .$$

Ou, utilizando a notação que acabamos de introduzir em (4.41),

$$\mathrm{Var}(X) := E\left((X(\omega) - E(X))^2\right) = E\left(X^2\right) - (E(X))^2 .$$

**Nota 4.5.** Denotando $\nu$ a lei de X (i.e. $\nu = X_\# P$), pelo Teorema 4.3,

$$E(X) = \int_{\mathbb{R}} x\, d\nu(x) \quad \text{e} \quad \mathrm{Var}(X) = \int_{\mathbb{R}} x^2\, d\nu(x) - (E(X))^2 .$$

**Nota 4.6.** Se $X^2 \in L^1(\Omega, P)$, então $X \in L^1(\Omega, P)$. Efetivamente, num espaço de medida finita, as constantes são integráveis e $|X| \leq 1 + X^2$. O contrário não acontece necessariamente, como sucede com a distribuição$-t$ de Student com dois graus de liberdade, que possui média mas não possui variância. A lei de Cauchy não tem nem média nem variância. (cf. Subsecção 3.3.5.)

#### Produto de variáveis aleatórias integráveis, independentes

**Lema 4.9.** Sejam X e Y duas variáveis aleatórias integráveis, independentes (cf. Definição 3.16). Então, o produto $X \cdot Y$ é integrável e

$$E(X \cdot Y) = E(X) \cdot E(Y) . \tag{4.42}$$

Quer dizer,

$$\int_\Omega X(\omega) \cdot Y(\omega)\ dP(\omega) = \int_\Omega X(\omega)\ dP(\omega) \cdot \int_\Omega Y(\omega)\ dP(\omega) .$$

***Demonstração.*** Assumiremos que X e Y são não negativas. O caso geral segue-se facilmente, a partir deste.

Fixemos $n \in \mathbb{N}$ e definamos, como em (4.24) (cf. Lema 4.3),

$$A_i = \mathrm{X}^{-1}\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right) = \left\{\omega \in \Omega\,;\, \frac{i}{2^n} \leq \mathrm{X}\,(\omega) < \frac{i+1}{2^n}\right\}$$

e

$$B_i = \mathrm{Y}^{-1}\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right) = \left\{\omega \in \Omega\,;\, \frac{i}{2^n} \leq \mathrm{Y}\,(\omega) < \frac{i+1}{2^n}\right\}.$$

As variáveis aleatórias

$$s_n = \sum_{i=0}^{n\cdot 2^n} \frac{i}{2^n} \cdot \chi_{A_i} \quad \text{e} \quad t_n = \sum_{i=0}^{n\cdot 2^n} \frac{i}{2^n} \cdot \chi_{B_i}$$

são independentes e

$$\begin{aligned}
\mathrm{E}\,(s_n \cdot t_n) &= \mathrm{E}\left(\sum_{i,j=0}^{n\cdot 2^n} \frac{i}{2^n} \cdot \frac{j}{2^n} \cdot \chi_{A_i} \cdot \chi_{B_j}\right) \\
&= \sum_{i,j=0}^{n\cdot 2^n} \frac{i}{2^n} \cdot \frac{j}{2^n} \cdot \mathrm{P}\,(A_i \cap B_j) \\
&= \sum_{i,j=0}^{n\cdot 2^n} \frac{i}{2^n} \cdot \frac{j}{2^n} \cdot \mathrm{P}\,(A_i) \cdot \mathrm{P}\,(B_j) \\
&= \sum_{i=0}^{n\cdot 2^n} \frac{i}{2^n} \cdot \mathrm{P}\,(A_i) \cdot \sum_{j=0}^{n\cdot 2^n} \frac{j}{2^n} \cdot \mathrm{P}\,(B_j) \\
&= \mathrm{E}\,(s_n) \cdot \mathrm{E}\,(t_n)\,.
\end{aligned}$$

Como $s_n \nearrow \mathrm{X}$, $t_n \nearrow \mathrm{Y}$ e $s_n \cdot t_n \nearrow \mathrm{X} \cdot \mathrm{Y}$, pelo Teorema 4.2 (convergência monótona), obtemos (4.42). :-)

***Observação.*** A demonstração é a mesma para o produto de um número finito de variáveis aleatórias independentes.

Note também que o produto de variáveis aleatórias integráveis, nem sempre é integrável. Se o for, todas as variáveis aleatórias integráveis seriam de quadrado integrável. Mas é fácil encontrar exemplos de variáveis aleatórias integráveis cujo quadrado não é integrável.

### 4.2.2 Lei(es) dos grandes números

#### Convergência em probabilidade

**Definição 4.13.** Seja $\{X_n\}_{n\in\mathbb{N}}$ uma sucessão de variáveis aleatórias e seja X mais uma variável aleatória. Diz-se que $X_n$ ***converge*** para X ***em probabilidade***, quando $n \to \infty$, e denota-se $X_n \rightharpoonup^{P} X$, se para todo $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(|X_n - X| \geq \epsilon\right) = 0 \tag{4.43}$$

O lema seguinte é uma aplicação clássica da desigualdade de Chebyshev (Lema 4.5) e da esperança do produto de variáveis aleatórias independentes (Lema 4.9).

#### Lei fraca dos grandes números

**Lema 4.10.** (Lei da média.) *Seja $\{X_i\}_{i\in\mathbb{N}}$ uma sucessão de ensaios de Bernoulli de parâmetro $p \in (0,1)$. (i.e. Variáveis aleatórias de Bernoulli de parâmetro $p$, independentes.) Então,*

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \rightharpoonup^{P} p.$$

***Demonstração.*** Por (4.32) (desigualdade de Chebyshev, cf. Lema 4.5), para todo $\epsilon > 0$ fixo,

$$P\left(\left|\overline{X}_n - p\right| \geq \epsilon\right) \leq \frac{1}{\epsilon^2} E\left(\left|\overline{X}_n - p\right|^2\right) = \frac{1}{\epsilon^2 n^2} E\left(\left|\sum_{i=1}^{n} (X_i - p)\right|^2\right).$$

Mas

$$\begin{aligned} E\left(\left|\sum_{i=1}^{n} (X_i - p)\right|^2\right) &= \sum_{i,j=1}^{n} E\left((X_i - p)(X_j - p)\right) \\ &= \sum_{i=1}^{n} E\left((X_i - p)^2\right) \\ &= np(1-p), \end{aligned}$$

já que para $i \neq j$, por (4.42),

$$E\left((X_i - p)(X_j - p)\right) = E\left(X_i - p\right) E\left(X_j - p\right) = 0,$$

pois as variáveis $X_i - p$ e $X_j - p$ são independentes.

Logo,

$$\mathrm{P}\left(\left|\overline{\mathrm{X}}_n - p\right| \geq \epsilon\right) \leq \frac{p(1-p)}{\epsilon^2 n},$$

que tende para 0 quando $n \to \infty$. :-)

O próximo teorema generaliza o Lema 4.10 para sucessões que podem ter distribuições completamente diferentes, só exige-se que as variâncias sejam uniformemente limitadas. Este teorema é utilizado em Estatística para estimar a média de um fenómeno observado, do qual náo se conhece a lei de probabilidade. A sua demonstração é semelhante à do Lema 4.10. (cf. Exercício 10.)

**Teorema 4.5.** (Lei fraca dos grandes números.) *Seja* $\{\mathrm{X}_i\}_{i\in\mathbb{N}}$, *uma sucessão de variáveis aleatórias independentes (cf. Definição 3.16), de quadrado integrável (i.e. possuindo média e variância). Supomos para além disso, que a sucessão de variâncias é limitada. i.e. Existe* $K > 0$, *tal que* $\mathrm{Var}\left(\mathrm{X}_i\right) \leq K$, *para todo* $i \in \mathbb{N}$. *Então, pondo*

$$\overline{\mathrm{X}}_n := \frac{1}{n}\sum_{i=1}^{n} \mathrm{X}_i, \tag{4.44}$$

*temos*

$$\overline{\mathrm{X}}_n - \mathrm{E}\left(\overline{\mathrm{X}}_n\right) \rightharpoonup^{\mathrm{P}} 0.$$

### Lei forte dos grandes números

**Lema 4.11.** (Borel-Cantelli.) *Seja* $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ *uma sucessão de conjuntos mensuráveis, tal que*

$$\sum_{n=1}^{\infty} \mathrm{P}\left(A_n\right) < \infty.$$

*Então,*

$$\mathrm{P}\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} A_n\right) = 0.$$

***Demonstração.*** Para todo $k \in \mathbb{N}$,

$$\mathrm{P}\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} A_n\right) \leq \mathrm{P}\left(\bigcup_{n=k}^{\infty} A_n\right) \leq \sum_{n=k}^{\infty} \mathrm{P}\left(A_n\right) \to 0,$$

já que a série é convergente. :-)

**Nomenclatura.** Na literatura em inglês, se $x \in \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$, diz-se que $x \in A_n$ i.o. (infinitely often). Efetivamente, $x \in \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$ se e só se $x \in A_n$ para um número infinito de índices.

***Observação.*** O lema de Borel-Cantelli é enunciado no contexto das probabilidades, mas ele é válido em qualquer espaço de medida, pois não é necessário a medida ser finita. Historicamente, o lema de Borel-Cantelli foi formulado para demonstrar o teorema de Borel dos "números normais." (cf. Teorema 4.8.)

**Teorema 4.6.** (Lei forte dos grandes números.) *Seja* $\{\mathrm{X}_i\}_{i\in\mathbb{N}}$ *uma sucessão de variáveis aleatórias independentes, tal que para todo* $i \in \mathbb{N}$, $\mathrm{X}_i^4 \in L^1(\Omega, \mathrm{P})$. *Supomos que existe* $K > 0$, *tal que para todo* $i \in \mathbb{N}$, $\mathrm{E}\left((\mathrm{X}_i - \mathrm{E}(\mathrm{X}_i))^4\right) \leq K$ *(em particualr, a variância é limitada por* $1 + K$, *qualquer que seja* $i \in \mathbb{N}$*). Então, o conjunto*

$$\left\{\omega \in \Omega;\ \lim_{n\to\infty} \overline{\mathrm{X}}_n - \mathrm{E}\left(\overline{\mathrm{X}}_n\right) = 0\right\}$$

*tem medida total. i.e.*

$$\mathrm{P}\left(\left\{\omega \in \Omega;\ \lim_{n\to\infty} \overline{\mathrm{X}}_n - \mathrm{E}\left(\overline{\mathrm{X}}_n\right) = 0\right\}\right) = 1.$$

( $\overline{\mathrm{X}}_n$ é a média aritmética das primeiras $n$ variáveis, definida em (4.44) )

***Demonstração.*** A demonstração deste teorema é muito semelhante à demonstração do Teorema 4.7 a seguir, e fica como exercício. (cf. Exercício 12.)

O teorema seguinte é a versão da lei forte dos grandes números para os ensaios de Bernoulli. O enunciamos e demonstramos detalhadamente, pois é a base para o teorema de Borel dos "números normais." (cf. Teorema 4.8.)

**Teorema 4.7.** (Lei forte dos grandes números I.) *Seja* $\{\mathrm{X}_i\}_{i\in\mathbb{N}}$ *uma sucessão de ensaios de Bernoulli de parâmetro* $p \in (0, 1)$. *(i.e. Variáveis aleatórias de Bernoulli de parâmetro* $p$*, independentes.) Então o conjunto*

$$N_p := \left\{\omega \in \Omega;\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p\right\} \tag{4.45}$$

*tem medida total. i.e.*

$$\mathrm{P}(N_p) = 1.$$

( $\overline{\mathrm{X}}_n$ é a média aritmética das primeiras $n$ variáveis, definida em (4.44) )

***Demonstração.*** Note primeiro que o conjunto $N_p$ é mensurável, já que as variáveis aleatórias $\liminf_{n\to\infty} \overline{\mathrm{X}}_n$ e $\limsup_{n\to\infty} \overline{\mathrm{X}}_n$ são mensuráveis (cf. Corolário 3.2) e

$$N_p = \left\{\omega \in \Omega;\ \liminf_{n\to\infty} \overline{\mathrm{X}}_n(\omega) \geq p\right\} \cap \left\{\omega \in \Omega;\ \limsup_{n\to\infty} \overline{\mathrm{X}}_n(\omega) \leq p\right\}.$$

Agora, fixemos $\epsilon > 0$.

Pela desigualdade de Chebyshev (Lema 4.5),

$$\mathrm{P}\left(\left|\overline{\mathrm{X}}_n - p\right| \geq \epsilon\right) \leq \frac{1}{\epsilon^4}\mathrm{E}\left(\left|\overline{\mathrm{X}}_n - p\right|^4\right) = \frac{1}{\epsilon^4 n^4}\mathrm{E}\left(\left|\sum_{i=1}^{n}\left(\mathrm{X}_i - p\right)\right|^4\right).$$

Observe-se que

$$\mathrm{E}\left(\left|\sum_{i=1}^{n}\left(\mathrm{X}_i - p\right)\right|^4\right) = \sum_{i,j,k,l=1}^{n}\mathrm{E}\left(\left(\mathrm{X}_i - p\right)\left(\mathrm{X}_j - p\right)\left(\mathrm{X}_k - p\right)\left(\mathrm{X}_l - p\right)\right).$$

Pela independência das variáveis aleatórias,

$$\mathrm{E}\left(\left(\mathrm{X}_i - p\right)\left(\mathrm{X}_j - p\right)\left(\mathrm{X}_k - p\right)\left(\mathrm{X}_l - p\right)\right) = 0$$

sempre que um dos índices é diferente de todos os outros. Portanto, os termos não nulos são $\mathrm{E}\left(\left(\mathrm{X}_i - p\right)^4\right)$, $i = 1, ..., n$ e 3 termos iguais a $\mathrm{E}\left(\left(\mathrm{X}_i - p\right)^2\left(\mathrm{X}_j - p\right)^2\right)$, por cada par $i, j = 1, ..., n$, $i \neq j$. Como $\mathrm{X}_i^m = \mathrm{X}_i$, para todo $m > 0$, então

$$\mathrm{E}\left(\left(\mathrm{X}_i - p\right)^4\right) = \mathrm{E}\left(\mathrm{X}_i^4\right) - 4p\mathrm{E}\left(\mathrm{X}_i^3\right) + 6p^2\mathrm{E}\left(\mathrm{X}_i^2\right) - 4p^3\mathrm{E}\left(\mathrm{X}_i\right) + p^4 = p - 4p^2 + 6p^3 - 3p^4.$$

e

$$\mathrm{E}\left(\left(\mathrm{X}_i - p\right)^2\left(\mathrm{X}_j - p\right)^2\right) = \mathrm{E}\left(\left(\mathrm{X}_i - p\right)^2\right)\mathrm{E}\left(\left(\mathrm{X}_j - p\right)^2\right) = p^2\left(1 - p\right)^2.$$

Somando estes termos, tendo em conta que há $n$ termos iguais a $\mathrm{E}\left(\left(\mathrm{X}_i - p\right)^4\right)$ e $3n(n-1)$ termos iguais a $\mathrm{E}\left(\left(\mathrm{X}_i - p\right)^2\left(\mathrm{X}_j - p\right)^2\right)$, $i \neq j$, obtemos

$$np(1-p)(1-3p(1-p)) + 3n(n-1)p^2\left(1-p\right)^2 \leq n^2.$$

Logo,

$$\mathrm{P}\left(\left|\overline{\mathrm{X}}_n - p\right| \geq \epsilon\right) \leq \frac{1}{\epsilon^4 n^2}.$$

Tomando $\epsilon = n^{-\frac{1}{8}}$, obtemos

$$\mathrm{P}\left(\left|\overline{\mathrm{X}}_n - p\right| \geq n^{-\frac{1}{8}}\right) \leq n^{-\frac{3}{2}}.$$

Como

$$\sum_{n=1}^{\infty}\mathrm{P}\left(\left|\overline{\mathrm{X}}_n - p\right| \geq n^{-\frac{1}{8}}\right) \leq \sum_{n=1}^{\infty}\mathfrak{n}^{-\frac{3}{2}} \leq 3,$$

pelo lema de Borel-Cantelli (cf. Lema 4.11),

$$\mathrm{P}\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\left\{\omega \in \Omega;\ \left|\overline{\mathrm{X}}_n\left(\omega\right) - p\right| \geq n^{-\frac{1}{8}}\right\}\right) = 0.$$

Só falta verificar que

$$\Omega \setminus \left( \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \left\{ \omega \in \Omega;\ \left| \overline{\mathrm{X}}_n(\omega) - p \right| \geq n^{-\frac{1}{8}} \right\} \right) \subseteq N_p.$$

Efetivamente, se

$$\omega \notin \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \left\{ \omega \in \Omega;\ \left| \overline{\mathrm{X}}_n(\omega) - p \right| \geq n^{-\frac{1}{8}} \right\},$$

então existe $m \in \mathbb{N}$, tal que

$$\omega \notin \bigcup_{n=m}^{\infty} \left\{ \omega \in \Omega;\ \left| \overline{\mathrm{X}}_n(\omega) - p \right| \geq n^{-\frac{1}{8}} \right\}$$

Isto significa que para todo $n \geq m$, $\left| \overline{\mathrm{X}}_n(\omega) - p \right| < n^{-\frac{1}{8}}$. Portanto,

$$\lim_{n\to\infty} \left| \overline{\mathrm{X}}_n(\omega) - p \right| = 0. \ :-)$$

### Medidas estrangeiras

**Definição 4.14.** (cf. Definição 3.11.) Duas medidas, $\mu$ e $\nu$ sobre um espaço mensurável $(X, \mathcal{F})$ são ***estrangeiras***, se existe $A \in \mathcal{F}$ tal que

$$\mu(A^{\mathrm{c}}) = \nu(A) = 0.$$

**Corolário 4.7.** As probabilidades $\mathrm{P}_p$ sobre o espaço de Bernoulli $(\mathbb{B}, \mathcal{F}_\infty)$, dadas pelo Teorema 2.3, são estrangeiras entre elas. i.e. $\mathrm{P}_{p_1}$ e $\mathrm{P}_{p_2}$ são estrangeiras, para todo $p_1$, $p_2 \in (0,1)$, $p_1 \neq p_2$.

***Demonstração.*** Basta observar que as projeções coordenadas

$$\mathrm{X}_i : (\omega_1, \omega_2, \ldots) \in \mathbb{B} \longrightarrow \omega_i\,,\ i \in \mathbb{N},$$

definidas em (3.32), são ensaios de Bernoulli de parámetro $p$, definidos sobre o espaço de probabilidade $(\mathbb{B}, \mathcal{F}_\infty, \mathrm{P}_p)$, sendo $p \in (0,1)$ um número fixo qualquer. (cf. Proposição 3.4.)

Pelo Teorema 4.7,

$$\mathrm{P}_{p_1}\left( \left\{ \omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p_1 \right\} \right) = \mathrm{P}_{p_2}\left( \left\{ \omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p_2 \right\} \right) = 1.$$

Mas, para todo $p_1$, $p_2 \in (0,1)$, $p_1 \neq p_2$,

$$\left\{ \omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p_1 \right\} \bigcap \left\{ \omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p_2 \right\} = \emptyset. :-)$$

**Probabilidades contínuas, no entanto estrangeiras com respeito à medida de Lebesgue**

Seja $\psi(\omega_1, \omega_2, ...) = \sum_{i=1}^{\infty} \omega_i \cdot 2^{-i} \in [0,1]$ a aplicação definida em (2.36). Sendo $p \in (0,1)$ fixo, vamos estudar a medida imagem

$$\varrho_p := \psi_{\#} \mathrm{P}_p, \tag{4.46}$$

sobre $([0,1], \mathcal{B})$, da probabilidade produto sobre $(\mathbb{B}, \mathcal{F}_\infty)$, denotada $\mathrm{P}_p$ (obtida no Teorema 2.3).

Tirando as sucessões que terminam em zeros (ou as que terminam em uns), $\psi$ é injetiva. Ademais, as sucessões que terminam em zeros como as que terminam em uns estão excluídas dos conjuntos $N_p$ (cf. fórmula (4.45)), qualquer que seja $p \in (0,1)$. Portanto,

$$\psi^{-1}\left(\psi\left(\left\{\omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p\right\}\right)\right) = \left\{\omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p\right\}$$

e

$$\psi\left(\left\{\omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p_1\right\}\right) \bigcap \psi\left(\left\{\omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p_2\right\}\right) = \emptyset.$$

Para além disso, pela definição de medida imagem,

$$\varrho_p\left(\psi\left(\left\{\omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p\right\}\right)\right) = \mathrm{P}_p\left(\left\{\omega \in \mathbb{B};\ \lim_{n\to\infty} \overline{\mathrm{X}}_n(\omega) = p\right\}\right) = 1.$$

**Corolário 4.8.** As probabilidades $\varrho_p$, sobre $([0,1], \mathcal{B})$, $p \in (0,1)$, são estrangeiras entre elas (cf. Definição 3.11). i.e. Para todo $p_1$, $p_2 \in (0,1)$, $p_1 \neq p_2$, $\varrho_{p_1}$ e $\varrho_{p_2}$ são estrangeiras.

***Demonstração.*** É uma aplicação direta do Corolário 4.7. :-)

**Corolário 4.9.** As probabilidades $\varrho_p$, $p \in (0,1)$, $p \neq \frac{1}{2}$ são uma família de probabilidades contínuas sobre $([0,1], \mathcal{B})$ (Definição 3.15), a pesar de serem estrangeiras com respeito à medida de Lebesgue $\lambda$.

***Demonstração.*** Pelo Teorema 2.4, à medida de Lebesgue sobre $([0,1], \mathcal{B})$, denotada $\lambda$, coincide com $\varrho_{\frac{1}{2}}$ $(p = 1/2)$. Pelo Corolário 4.8, as probabilidades $\varrho_p$, $p \in (0,1)$, $p \neq \frac{1}{2}$ são estrangeiras com respeito a $\varrho_{\frac{1}{2}}$, que, como dizemos, é igual a $\lambda$. O facto de $\varrho_p$ ser contínua, deixa-se como exercício. (Exercício 11.) :-)

**Números normais**

**Definição 4.15.** Um número $t \in [0,1]$ é um ***número normal*** se

$$\lim_{n\to\infty} \frac{\omega_1 + \cdots + \omega_n}{n} = \frac{1}{2},$$

onde a sucessão $\omega = (\omega_1, \omega_2, \cdots)$ é o desenvolvimento em base dois de $t$. i.e. $t = \sum_{i=1}^{\infty} \omega_i \cdot 2^{-i}$.

***Observação.*** O conjunto dos números normais é $\psi\left(N_{\frac{1}{2}}\right)$, sendo $N_{\frac{1}{2}}$ o conjunto definido em (4.45) e $\psi$ a aplicação definida em (2.36).

**Teorema 4.8.** (Teorema de Borel dos números normais) *No espaço probabilístico* $([0,1], \mathcal{B}, \lambda)$ *(i.e. no intervalo* $[0,1]$*, dotado da probabilidade uniforme), quase todos os números são normais.*

***Demonstração.*** Este teorema é uma consequência imediata dos teoremas 4.7, 2.4 e da observação anterior. :-)

## 4.3 Espaços $L^p(X, \mu)$

**Definição 4.16.**

(a) Sendo $p \geq 1$ e $f$ uma função mensurável, dizemos que $f$ é $p-\mu-$***integrável*** (ou $p-$***integrável***) se $|f|^p \in L^1(X, \mu)$. i.e.

$$\int_X |f|^p(x)\, d\mu(x) < \infty.$$

O conjunto das funções $p-$integráveis denota-se

$$\mathcal{L}^p(X, \mu) := \left\{ f \in \mathcal{L}(X)\,;\ \int_X |f|^p(x)\, d\mu(x) < \infty \right\}$$

e

$$L^p(X, \mu) := \mathcal{L}^p(X, \mu) / \sim$$

denota o conjunto das classes de equivalência das funções $p-$integráveis.

(b) Seja $f$ uma função mensurável. Dizemos que $f$ é ***essencialmente limitada*** se existir $M > 0$ tal que

$$|f(x)| \leq M, \text{ para quase todo } x \in X. \text{ i.e.}$$

$$\mu\left(\{x \in X;\ |f(x)| > M\}\right) = 0.$$

O conjunto das (classes de equivalência das) funções essencialmente limitadas denota-se $L^{\infty}(X, \mu)$.

### 4.3.1 Estrutura de espaço vetorial normado sobre $L^p(X, \mu)$

**Nota 4.7.** Por (4.31) (cf. Proposição 4.6), qualquer que seja $p$ $(1 \leq p \leq \infty)$, para todo $f \in L^p(X, \mu)$,

$$\mu\left(\{|f| = +\infty\}\right) = \mu\left(\{|f|^p = +\infty\}\right) = 0.$$

Portanto, a soma e o produto por escalar dos elementos de $L^p(X, \mu)$ está bem definida (cf. Definição 4.11).

**Proposição 4.9.** $L^p(X, \mu)$ é um espaço vetorial sobre $\mathbb{R}$. $(1 \leq p \leq \infty.)$

***Demonstração.*** O único que falta por ver é que a soma de funções $p-$integráveis é uma função $p-$integrável. Para $L^{\infty}(X, \mu)$, é evidente, já que $|f + g| \leq |f| + |g|$. Para $p \geq 1$, como a função $t \to t^p$ é convexa,

$$\left(\frac{|f+g|}{2}\right)^p \leq \frac{|f|^p + |g|^p}{2}.$$

Integrando ambos os lados desta desigualdade, pela monotonia (4.13) e aditividade (4.14) (cf.Proposição 4.4),

$$\int_X |f+g|^p \, d\mu \leq 2^{p-1}\left(\int_X |f|^p \, d\mu + \int_X |g|^p \, d\mu\right). \quad :-)$$

**Norma em $L^p(X, \mu)$**

**Definição 4.17.** Sendo $p \geq 1$, para todo $f \in L^p(X, \mu)$, definimos

$$\|f\|_p := \left(\int_X |f|^p(x) \, d\mu(x)\right)^{\frac{1}{p}}. \tag{4.47}$$

Para $f \in L^{\infty}(X, \mu)$, define-se

$$\|f\|_{\infty} := \inf\left\{M \geq 0;\ \mu\left(\{x \in X;\ |f(x)| > M\}\right) = 0\right\}. \tag{4.48}$$

**Nota 4.8.** $\|f\|_\infty = \min\{M \geq 0;\ \mu(\{x \in X;\ |f(x)| > M\}) = 0\}$, já que

$$\mu(\{x \in X;\ |f(x)| > \|f\|_\infty\}) = 0.$$

(cf. Exercício 13.)

**Proposição 4.10.** $\|\cdot\|_p$ é uma norma sobre $L^p(X, \mu)$. $(1 \leq p \leq \infty.)$

***Demonstração.*** Na Proposição 4.8, provamos que $\|\cdot\|_1$ é uma norma sobre $L^1(X, \mu)$. As duas primeiras alinhas da demonstração da Proposição 4.8 adaptam-se facilmente a este caso.

Só falta provar a desigualdade triângular. No contexto dos espaços $L^p(X, \mu)$, a desigualdade triangular é conhecida como desigualdade de Minkowski, que vamos demonstrar na Proposição 4.12. Mas antes vamos mostrar a desigualdade de Hölder, por ser necessária para a demonstração da desigualdade de Minkowski. :-)

**Definição 4.18.** Sejam $p > 1$ e $q > 1$. Dizemos que $p$ e $q$ são ***exponentes conjugados*** ou simplesmente ***conjugados***, se

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{4.49}$$

i.e. $p$ e $q$ são conjugados se $p + q = pq$.

Para além disso, 1 e $\infty$ são conjugados.

### Desigualdade de Hölder

**Proposição 4.11.** Sejam $p$ e $q$ um par de exponentes conjugados. Então, para todo $f \in L^p(X, \mu)$ e todo $g \in L^q(X, \mu)$, $fg \in L^1$ e verifica-se a desigualdade seguinte

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q. \tag{4.50}$$

***Demonstração.*** Se $\|f\|_p = 0$ ou $\|g\|_q = 0$, $fg = 0$ em quase todo ponto e a desigualdade verifica-se trivialmente. Se, pelo contrário, $\|f\|_p \cdot \|g\|_q \neq 0$, dividimos $f$ e $g$ pelas suas normas respetivas. Pela desigualdade de Young (cf. Exercício **??**), para todo $x \in X$,

$$\frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p} \cdot \frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q} \cdot \frac{|g(x)|^q}{\|g\|_q^q}.$$

Integrando ambos os lados desta desigualdade, pela monotonia (4.13) e aditividade (4.14) (cf.Proposição 4.4), obtemos

$$\begin{aligned}\frac{1}{\|f\|_p \|g\|_q} \int_X |f(x)g(x)|\, d\mu(x) &\leq \frac{1}{p}\frac{1}{\|f\|_p^p} \int_X |f(x)|^p \, d\mu(x) + \frac{1}{q}\frac{1}{\|f\|_q^q} \int_X |g(x)|^q \, d\mu(x) \\ &= \frac{1}{p} + \frac{1}{q} = 1.\end{aligned}$$

multiplicando por $\|f\|_p \|g\|_q$ ambos os extremos desta desigualdade, obtemos (4.50).

***Observação.*** Quando $p = q = 2$, a desigualdade (4.50) é conhecida como desigualdade de Cauchy-Schwarz.

### Desigualdade de Minkowski

**Proposição 4.12.** Para todo $f$ e $g \in L^p(X, \mu)$, $1 \leq p \leq \infty$ qualquer,

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \,. \tag{4.51}$$

***Demonstração.*** Se $p = 1$ ou $p = \infty$, a desigualdade (4.51) é evidente, já que

$$|f + g| \leq |f| + |g| \,.$$

Se $p \in (1, \infty)$, multiplicamos ambos os lados da desigualdade anterior por $|f + g|^{p-1}$ e integramos. Pela monotonia (4.13) e aditividade (4.14) (cf.Proposição 4.4),

$$\int_X |f + g|^p \, d\mu \leq \int_X |f + g|^{p-1} |f| \, d\mu + \int_X |f + g|^{p-1} |g| \, d\mu. \tag{4.52}$$

Agora, aplicamos a desigualdade (4.50) (desigualdade de Holder) aos dois integrais da direita, separadamente. Note que $q := \frac{p}{p-1}$ é o conjugado de $p$.

$$\int_X |f + g|^{p-1} |f| \, d\mu \leq \left(\int_X |f + g|^p \, d\mu\right)^{\frac{p-1}{p}} \left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}}$$

e

$$\int_X |f + g|^{p-1} |g| \, d\mu \leq \left(\int_X |f + g|^p \, d\mu\right)^{\frac{p-1}{p}} \left(\int_X |g|^p \, d\mu\right)^{\frac{1}{p}}.$$

Somando estas duas desigualdades e aplicando (4.52), obtemos

$$\int_X |f + g|^p \, d\mu \leq \left(\int_X |f + g|^p \, d\mu\right)^{\frac{p-1}{p}} \left[\left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}} + \left(\int_X |g|^p \, d\mu\right)^{\frac{1}{p}}\right]. \tag{4.53}$$

Finalmente, se $\|f + g\|_p = 0$ a desigualdade (4.51) é trivial. Senão, dividimos (4.53) por $\|f + g\|_p^{p-1}$ para obter (4.51). :-)

## Exercícios

1. (cf. Definição 4.2.) Sendo $s$ uma função simples, definimos $n$, $c_i$ e $A_i$, $i = 1, \cdots, n$, por (4.1). Mostre que a representação (4.2) de $s$ é única.

2. Seja $\{X_j\}_{j\in J}$ uma partição de $X$ e $\mathcal{F}$ a $\sigma-$álgebra gerada pela partição $\{X_j\}_{j\in J}$, sendo $J$ um conjunto numerável de índices. Mostre que qualquer função $f : X \to \overline{\mathbb{R}}$ é $\mathcal{F}/\mathcal{B}-$mensurável se e só se para todo $j \in J$, $f$ é constante em $X_j$.

3. (cf. Proposição 4.3.) Sendo $s$ e $t \in \mathcal{L}^* (X)$, utilize o Lema 4.2, para verificar as seguintes propriedades.

   (a) *Linearidade*: Para todo $a, b \in \mathbb{R}^+$,
   $$\int_X^* (as + bt)(x) \, \mathrm{d}\mu(x) = a \cdot \int_X^* s(x) \, \mathrm{d}\mu(x) + b \cdot \int_X^* t(x) \, \mathrm{d}\mu(x).$$

   (b) *Monotonia*: Se $s \leq t$, então
   $$\int_X^* s(x) \, \mathrm{d}\mu(x) \leq \int_X^* t(x) \, \mathrm{d}\mu(x).$$

4. Mostre que se
   $$\chi_A = \sum_{j=1}^{m} d_j \cdot \chi_{B_j},$$
   onde $d_j > 0$, para todo $j = 1, ..., m$, então $B_j \subseteq A$.

   Sendo
   $$J_x = \{j \in \{1, ..., m\} \, ; \, x \in B_j\},$$
   mostre que qualquer que seja $x \in A$,
   $$\sum_{j\in I_k} d_j = 1.$$

5. Sendo $J_x$ o conjunto de índices definido no exercício 4, mostre que
   $$C_x = \bigcap_{j\in J_x} B_j \, , \; x \in A \, ,$$
   é uma partição *finita* de $A$.

6. (cf. Nota 4.2.) Prove que para todo $s \in \mathcal{L}^{s,+}(X)$,
   $$\int_X s(x) \, d\mu(x) = \int_X^* s(x) \, d\mu(x).$$

7. (cf. Propposição 4.5.) Mostre que se a medida $f \cdot \mu$, definida por (4.16) (cf. Teorema 4.1), é $\sigma-$finita, então $\mu\left(\{x \in X;\ f(x) = \infty\}\right) = 0$.

   Para além disso, prove que se $\mu$ é $\sigma-$finita e $\mu\left(\{x \in X;\ f(x) = \infty\}\right) = 0$ então $f \cdot \mu$ é $\sigma-$finita.

8. Sendo
$$E_n := \{x \in X;\ f_n(x) \geq c \cdot s(x)\},$$
   onde $f \in \mathcal{L}^+(X)$, $s \in \mathcal{L}^{s,+}(X)$, $s \leq f$ e $c \in (0,1)$, mostre que
$$X = \bigcup_{n=1}^{\infty} E_n,$$

9. Prove que a relação $\sim$, definida por (4.33) (cf, Proposição 4.7), é uma relação de equivalência sobre $\mathcal{L}(X, \mu)$.

10. (Lei fraca dos grandes números, cf. Teorema 4.5.) Seja $\{\mathrm{X}_i\}_{i\in\mathbb{N}}$, uma sucessão de variáveis aleatórias independentes, de quadrado integrável. Supondo que a sucessão de variâncias é limitada (i.e. Existe $K > 0$, tal que $\mathrm{Var}(\mathrm{X}_i) \leq K$, para todo $i \in \mathbb{N}$), mostre que
$$\overline{\mathrm{X}}_n - \mathrm{E}\left(\overline{\mathrm{X}}_n\right) \rightharpoonup^{\mathrm{P}} 0,$$
    onde $\overline{\mathrm{X}}_n$ é a média aritmética, definida em (4.44).

11. Mostre que as probabilidades $\varrho_p$, dadas por (4.46), são contínuas, qualquer que seja $p \in (0,1)$.

12. (cf. Teorema 4.6.) Seja $\{\mathrm{X}_i\}_{i\in\mathbb{N}}$ uma sucessão de variáveis aleatórias independentes, tal que $\mathrm{X}_i^4 \in L^1(\Omega, \mathrm{P})$. Supondo que existe $K > 0$, tal que para todo $i \in \mathbb{N}$, $\mathrm{E}\left((\mathrm{X}_i - \mathrm{E}(\mathrm{X}_i))^4\right) \leq K$, mostre que o conjunto
$$\left\{\omega \in \Omega;\ \lim_{n\to\infty} \overline{\mathrm{X}}_n - \mathrm{E}\left(\overline{\mathrm{X}}_n\right) = 0\right\}$$
    tem medida total. i.e.
$$\mathrm{P}\left(\left\{\omega \in \Omega;\ \lim_{n\to\infty} \overline{\mathrm{X}}_n - \mathrm{E}\left(\overline{\mathrm{X}}_n\right) = 0\right\}\right) = 1.$$
    $\overline{\mathrm{X}}_n$ denota a média aritmética das primeiras $n$ variáveis, ver (4.44).

13. (cf. Nota 4.8.)Mostre que
$$\mu\left(\{x \in X;\ |f(x)| > \|f\|_\infty\}\right) = 0.$$

# Capítulo 5

# Integração no espaço produto

Neste Capítulo, vamos construir a medida produto, definida sobre a $\sigma-$álgebra produto de dois espaços de medida, digamos $(X, \mathcal{F}, \mu)$ e $(Y, \mathcal{G}, \nu)$, com medidas $\mu$ e $\nu$, ambas $\sigma-$finitas. De seguida, trataremos o teorema de Fubini-Tonelli, que será uma consequência imediata de esta construção. Finalmente, veremos a relação estreita entre medida produto e variáveis aleatórias independentes. Na primeira e segunda parte pode-se apreciar alguma imflueência do livro de P. Billingsley [1], a última parte, sobre a independência, toma certos aspetos do livro de M. Kac [9].

Para ler esta secção, é conveniente que o leitor reveja, se necessário, a definição e construção da $\sigma-$álgebra produto (cf. Definição 3.3) e a Proposição 3.1, onde mostramos que as projeções são mensuráveis. Além disso, provamos que o produto de $\sigma-$álgebras é associativo (cf. Lema 3.10) e que a $\sigma-$álgebra de Borel de $\mathbb{R}^n$, gerada pelos abertos de $\mathbb{R}^n$ e denotada $\mathcal{B}^n$, é igual à $\sigma-$álgebra produto de $n$ cópias da $\sigma-$álgebra de Borel de $\mathbb{R}$ (qualquer que seja $n \in \mathbb{N}$). i.e. $\mathcal{B}^n = \mathcal{B} \otimes \cdots \otimes \mathcal{B}$ (cf. Teorema 3.2). Isto permite caraterizar as funções mensuráveis a valores em $\mathbb{R}^n$, como sendo exatamente aquelas cujas componentes são todas funções mensuráveis, reais (cf. Corolário 3.3).

O espaço probabilístico de Bernoulli, $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty, \mathrm{P}_p\right)$ (sendo $p \in (0,1)$ qualquer), construído na subsecção 2.4.3 (cf. Teorema 2.3), é um exemplo de um espaço produto, com a $\sigma-$álgebra produto e a medida produto de uma quantidade numerável de cópias do espaço $(\{0,1\}, \mathcal{P}(\{0,1\}), \nu_p)$, mesmo que, até o momento, não o tenhamos visto sob este ângulo.

## 5.1 Medida produto

**Lema 5.1.** *Sendo* $(X, \mathcal{F})$ *e* $(Y, \mathcal{G})$ *dois espaços mensuráveis, a álgebra* $\mathcal{A}_{X\times Y}$*, gerada pelos* **retângulos mensuráveis** *(i.e. os conjuntos da forma* $A \times B$*, sendo* $A \in \mathcal{F}$ *e* $B \in \mathcal{G}$*, cf.* (3.8)*) é igual à união finita de retângulos mensuráveis disjuntos dois a dois. i.e.*

$$\mathcal{A}_{X\times Y} = \left\{ \bigcup_{i=1}^{n} A_i \times B_i;\, n \in \mathbb{N},\, A_i \in \mathcal{F},\, B_i \in \mathcal{G},\, A_i \times B_i \cap A_j \times B_j = \emptyset, \forall\, i \neq j \right\}.$$

***Demonstração.*** cf. Exercício 1.

**Teorema 5.1.** *Sejam* $(X, \mathcal{F}, \mu)$ *e* $(Y, \mathcal{G}, \nu)$ *dois espaços de medida. Assumimos que as medidas* $\mu$ *e* $\nu$ *são ambas* $\sigma-$*finitas. Então, existe uma única medida, denotada* $\mu \otimes \nu$*, sobre* $(X \times Y, \mathcal{F} \otimes \mathcal{G})$ *(cf. Definição 3.3), tal que para todo* $A \in \mathcal{F}$ *e* $B \in \mathcal{G}$*,*

$$\mu \otimes \nu\,(A \times B) = \mu\,(A) \cdot \nu\,(B)\,. \tag{5.1}$$

***Demonstração.*** Pelo Teorema 2.1 (teorema de Carathéodory), basta definir uma premedida $\sigma-$finita sobre a álgebra $\mathcal{A}_{X\times Y}$ do Lema 5.1, verificando (5.1). Naturalmente, começamos por definir

$$\zeta\,(A \times B) = \mu\,(A) \cdot \nu\,(B)\,, \tag{5.2}$$

para todo $A \times B \in \mathcal{R}_{X\times Y}$ (retângulos mensuráveis, cf. (3.8)).

Pela propriedade aditiva das premedidas, estendemos $\zeta$ à $\mathcal{A}_{X\times Y}$ por adição,

$$\zeta\,(A_1 \times B_1 \cup \cdots \cup A_n \times B_n) = \mu\,(A_1) \cdot \nu\,(B_1) + \cdots + \mu\,(A_n) \cdot \nu\,(B_n)\,, \tag{5.3}$$

sendo $n$ o número de retângulos mensuráveis, $A_i \times B_i \in \mathcal{R}_{X\times Y}$, $i = 1, ..., n$ e $A_i \times B_i \cap A_j \times B_j = \emptyset$, para todo $i \neq j$.

Note que $\zeta$ está bem definida, já que o valor de $\zeta$ depende unicamente do conjunto $A_1 \times B_1 \cup \cdots \cup A_n \times B_n$ e não da sua representação em retângulos mensuráveis, disjuntos, que evidentemente não é única. (cf. Exercício 2.)

Claramente, esta função é $\sigma-$finita. Efetivamente: Por hipótese, existem duas famílias crescentes de conjuntos mensuráveis $X_i \nearrow X$ e $Y_i \nearrow Y$, de medida finita. Portanto $X_i \times Y_i \nearrow X \times Y$ e para todo $i \in \mathbb{N}$,

$$\zeta\,(X_i \times Y_i) = \mu(X_i) \cdot \nu(Y_i) < \infty.$$

Agora vamos provar que $\zeta$ é uma premedida. Por (5.3), as alinhas 1 e 2 do Lema 2.6 estão garantidas. Só falta demonstrar (2.21) (subaditividade numerável).

Pelo Exercício 6, basta provar que

$$\zeta\left(A\times B\right)\leq\sum_{i=1}^{\infty}\zeta\left(A_i\times B_i\right),$$

onde $\{A_i\times B_i\}_{i\in\mathbb{N}}\subseteq\mathcal{R}_{X\times Y}$ é uma sucessão de retângulos mensuráveis e $A\times B$ é mais um retângulo mensurável, tal que

$$A\times B\subseteq\bigcup_{i=1}^{\infty}A_i\times B_i.$$

Tendo em conta (5.2), devemos mostrar

$$\mu\left(A\right)\cdot\nu\left(B\right)\leq\sum_{i=1}^{\infty}\mu\left(A_i\right)\cdot\nu\left(B_i\right). \tag{5.4}$$

Pondo $I_x=\{i\in\mathbb{N};\ x\in A_i\}$, note que para todo $x\in A$,

$$B\subseteq\bigcup_{i\in I_x}B_i.$$

Pela monotonia e subaditividade numerável da medida $\nu$, (cf. Lema 2.3),

$$\nu\left(B\right)\leq\sum_{i\in I_x}\nu\left(B_i\right)=\sum_{i=1}^{\infty}\nu\left(B_i\right)\cdot\chi_{A_i}\left(x\right).$$

Logo,

$$\nu\left(B\right)\cdot\chi_A\leq\sum_{i=1}^{\infty}\nu\left(B_i\right)\cdot\chi_{A_i}. \tag{5.5}$$

Como $A$ e $A_i$, $i\in\mathbb{N}$, são mensuráveis, as funções $\nu\left(B\right)\cdot\chi_A$ e $\nu\left(B_i\right)\cdot\chi_{A_i}$, $i\in\mathbb{N}$, são mensuráveis, incluindo o somatório (cf. (3.7)). Integrando com respeito a $\mu$ ambos os lados da desigualdade (5.5), pela monotonia do integral (cf. (4.13), Proposição 4.4) e por (4.28) (cf. Corolário 4.2),

$$\nu\left(B\right)\cdot\mu\left(A\right)\leq\int_X\sum_{i=1}^{\infty}\nu\left(B_i\right)\cdot\chi_{A_i}(x)\,d\mu(x)=\sum_{i=1}^{\infty}\nu\left(B_i\right)\cdot\mu\left(A_i\right).\ \ :-)$$

### 5.1.1 Teorema de Fubini-Tonelli para conjuntos

De aquí para frente, $(X \times Y, \mathcal{F} \otimes \mathcal{G}, \mu \otimes \nu)$ denota o espaço produto de dois espaços de medida fixos, $(X, \mathcal{F}, \mu)$ e $(Y, \mathcal{G}, \nu)$, ambos dotados de medidas $\sigma-$finitas, $\mu$ e $\nu$, respetivamente.

Na demonstração do Teorema 5.1 fixamos $x \in A$ e consideramos o recobrimento de $B$ pelos conjuntos $B_i$, cujo índice $i$ verifica $x \in A_i$. Se fazemos ao contrário, ou seja fixar $y \in B$ e considerar o recobrimento de $A$ pelos conjuntos $A_i$, cujo índice $i$ verifica $y \in B_i$, obtemos naturalmente o mesmo resultado. Mas se em lugar de $A \times B$ considerarmos um conjunto $D \in \mathcal{F} \otimes \mathcal{G}$ geral e fizermos o tipo de cálculo que corresponde a esta operação, fixando $x$ e depois fixando $y$, vamos obter duas fórmulas diferentes para $\mu \otimes \nu\,(D)$, o qual constitui o teorema de Fubini-Tonelli para conjuntos. Este é o objetivo principal desta subsecção.

**Secções ao longo de $Y$**

**Definição 5.1.** Para todo $D \subseteq X \times Y$ e todo $x \in X$, a ***secção de $D$ em $x$*** é o conjunto

$$D_x = \{y \in Y;\ (x, y) \in D\}. \tag{5.6}$$

**Exemplo.** Sejam $A \subseteq X$ e $B \subseteq Y$. Então, para todo $x \in X$,

$$(A \times B)_x = \begin{cases} B, & \text{si } x \in A \\ \emptyset, & \text{si } x \notin A \end{cases} \tag{5.7}$$

**Proposição 5.1.** (a) Para todo $D$ e $E \subseteq X \times Y$ e todo $x \in X$,

$$(D \setminus E)_x = D_x \setminus E_x.$$

En particular

$$(E^{\mathrm{c}})_x = (E_x)^{\mathrm{c}} \tag{5.8}$$

(b) Seja $\{D_i\}_{i \in I}$ uma família de subconjuntos de $X \times Y$, indexada por um conjunto de índices $I$ (arbitrário). Então, para todo $x \in X$,

$$\left(\bigcup_{i \in I} D_i\right)_x = \bigcup_{i \in I} (D_i)_x \tag{5.9}$$

e

$$\left(\bigcap_{i \in I} D_i\right)_x = \bigcap_{i \in I} (D_i)_x. \tag{5.10}$$

***Demonstração.*** cf. Exercício 4

**Lema 5.2.** *Para todo* $D \in \mathcal{F} \otimes \mathcal{G}$ *e todo* $x \in X$*, temos*

$$D_x \in \mathcal{G}.$$

***Demonstração.*** Fixemos $x \in X$ e definamos a família

$$(\mathcal{F} \otimes \mathcal{G})_x = \{D \in \mathcal{F} \otimes \mathcal{G};\ D_x \in \mathcal{G}\}.$$

Como $\mathcal{G}$ é uma $\sigma-$álgebra, observe-se que:

1. $X \times Y \in (\mathcal{F} \otimes \mathcal{G})_x$ (por (5.7)).
2. Se $D \in (\mathcal{F} \otimes \mathcal{G})_x$, então $D^c \in (\mathcal{F} \otimes \mathcal{G})_x$ (por (5.8)).
3. Se $\{D_i\}_{i\in\mathbb{N}} \subseteq (\mathcal{F} \otimes \mathcal{G})_x$, então $\bigcup_{i=1}^{\infty} D_i \in (\mathcal{F} \otimes \mathcal{G})_x$ (por (5.9)).

Logo, $(\mathcal{F} \otimes \mathcal{G})_x$ é uma $\sigma-$álgebra. Para além disso, $(\mathcal{F} \otimes \mathcal{G})_x \supseteq \mathcal{R}_{X\times Y}$ (por (5.7)). Conclui-se que,

$$(\mathcal{F} \otimes \mathcal{G})_x \supseteq \sigma\left[\mathcal{R}_{X\times Y}\right] = \mathcal{F} \otimes \mathcal{G}.\ \ :-)$$

**Corolário 5.1.** Seja $(Z, \mathcal{H})$ um terceiro espaço mensurável e seja $f : X \times Y \to Z$ uma função $(\mathcal{F} \otimes \mathcal{G})/\mathcal{H}-$mensurável. Então, para todo $x \in X$, a função

$$f_x : y \in Y \to f(x, y) \in Z \ \text{ é } \ \mathcal{G}/\mathcal{H} - \text{mensurável.} \tag{5.11}$$

***Demonstração.*** Para todo $A \in \mathcal{H}$, temos

$$\begin{aligned} f_x^{-1}(A) &= \{y \in Y;\ f_x(y) = f(x, y) \in A\} \\ &= \{y \in Y;\ (x, y) \in f^{-1}(A)\} \\ &= \left(f^{-1}(A)\right)_x. \end{aligned}$$

Como $f$ é $(\mathcal{F} \otimes \mathcal{G})/\mathcal{H}-$mensurável, $f^{-1}(A) \in \mathcal{F} \otimes \mathcal{G}$. Pelo lema anterior,

$$f_x^{-1}(A) = \left(f^{-1}(A)\right)_x \in \mathcal{G}.\ \ :-)$$

Agora incorporamos a medida $\nu$ ao nosso estudo. Pelo Lema 5.2, para todo $D \in \mathcal{F} \otimes \mathcal{G}$, $D_x \in \mathcal{G}$. Portanto, faz todo o sentido considerar a função $x \to \nu(D_x)$.

**Lema 5.3.** *Para todo* $D \in \mathcal{F} \otimes \mathcal{G}$*, a função*

$$x \in X \to \nu(D_x).$$

*é mensurável. i.e.* $x \in X \to \nu(D_x) \in \mathcal{L}^+(X)$.

***Demonstração.*** Utilizaremos novamente o tipo de raciocínio que já estamos habituados. Consideramos a família dos conjuntos que verificam a propriedade desejada e vamos provar que esta família, devido a sua estrutura, é de facto toda a $\sigma-$álgebra produto. Efetivamente, definamos

$$\mathcal{D} = \left\{ D \in \mathcal{F} \times \mathcal{G};\ x \in X \to \nu\left(D_x\right) \in \mathcal{L}^+\left(X\right) \right\}.$$

Queremos mostrar que $\mathcal{D} = \mathcal{F} \otimes \mathcal{G}$.

Por (5.7), para todo $A \in \mathcal{F}$ e $B \in \mathcal{G}$,

$$\nu\left(\left(A \times B\right)_x\right) = \nu\left(B\right) \cdot \chi_A(x) \in \mathcal{L}^+\left(X\right). \tag{5.12}$$

Logo, $\mathcal{R}_{X\times Y} \subseteq \mathcal{D}$.

Agora consideremos uma união finita, digamos $(A_1 \times B_1) \cup \cdots \cup (A_n \times B_n)$, de retângulos mensuráveis, disjuntos dois a dois. i.e. $(A_i \times B_i) \cap (A_j \times B_j) = \emptyset$, para todo $i \neq j$. Como

$$\nu\left(\left(\left(A_1 \times B_1\right) \cup \cdots \cup \left(A_n \times B_n\right)\right)_x\right) = \nu\left(B_1\right) \cdot \chi_{A_1}(x) + \cdots + \nu\left(B_n\right) \cdot \chi_{A_n}(x),$$

a função $x \to \nu\left(\left(\left(A_1 \times B_1\right) \cup \cdots \cup \left(A_n \times B_n\right)\right)_x\right)$ pertence a $\mathcal{L}^+\left(X\right)$. Portanto, $\mathcal{A}_{X\times Y} \subset \mathcal{D}$.

Pelo Teorema 2.6 (teorema das famílias monótonas), $\sigma\left[\mathcal{A}_{X\times Y}\right] = m\left[\mathcal{A}_{X\times Y}\right]$. Logo

$$\mathcal{F} \otimes \mathcal{G} = \sigma\left[\mathcal{A}_{X\times Y}\right] = m\left[\mathcal{A}_{X\times Y}\right] \subseteq m[\mathcal{D}].$$

Para concluir, basta provar que $m[\mathcal{D}] = \mathcal{D}$, ou seja que $\mathcal{D}$ é uma família monótona. Tendo em conta a Definição 2.12, devemos verificar as duas propriedades seguintes.

1. *União numerável crescente*: Seja $\{D_i\}_{i\in\mathbb{N}} \subset \mathcal{D}$, uma sucessão crescente. Pela monotonia das medidas (cf. Proposição 2.1, alinha 2), a sucessão

   $$\left\{x \to \nu\left(\left(D_i\right)_x\right)\right\}_{i\in\mathbb{N}} \subset \mathcal{L}^+\left(X\right)$$

   é crescente.

   Por (3.7) (cf. Corolário 3.2), a função

   $$x \to \lim_{i\to\infty} \nu\left(\left(D_i\right)_x\right) \in \mathcal{L}^+\left(X\right).$$

   Por (2.9) (cf. Lema 2.2), temos

   $$\lim_{i\to\infty} \nu\left(\left(D_i\right)_x\right) = \nu\left(\bigcup_{i=1}^{\infty}\left(D_i\right)_x\right).$$

Além disso, por (5.9),

$$\bigcup_{i=1}^{\infty}(D_i)_x = \left(\bigcup_{i=1}^{\infty} D_i\right)_x .$$

Logo,

$$x \to \nu\left(\left(\bigcup_{i=1}^{\infty} D_i\right)_x\right) = \nu\left(\bigcup_{i=1}^{\infty}(D_i)_x\right) = \lim_{i\to\infty} \nu\left((D_i)_x\right) \in \mathcal{L}^+(X).$$

Isto mostra que

$$\bigcup_{i=1}^{\infty} D_i \in \mathcal{D}.$$

2. *União numerável decrescente*: Seja $\{D_i\}_{i\in\mathbb{N}} \subseteq \mathcal{D}$ uma família decrescente. Temos de mostrar que

$$x \to \nu\left(\left(\bigcap_{i=1}^{\infty} D_i\right)_x\right) \in \mathcal{L}^+(X).$$

Para maior claridade na exposição, assumimos que $\nu$ é uma medida finita. O caso em que a medida $\nu$ é $\sigma-$finita, segue-se desta demonstração, tomando uma sucessão $Y_n \nearrow Y$, tal que $\nu(Y_n) < \infty$ (cf. Exercício 5).

Por hipótese, para todo $i \in \mathbb{N}$, $x \to \nu((D_i)_x) \in \mathcal{L}^+(X)$. Para além disso, $\{x \to \nu((D_i)_x)\}_{i\in\mathbb{N}}$ é decrescente. Por (3.7) (cf. Corolário 3.2),

$$x \to \lim_{i\to\infty} \nu((D_i)_x) \in \mathcal{L}^+(X).$$

Logo, tendo em conta (2.13) (cf. Lema 2.4),

$$\nu\left(\bigcap_{i=1}^{\infty}(D_i)_x\right) = \lim_{i\to\infty} \nu((D_i)_x) \in \mathcal{L}^+(X). \quad :-)$$

**Secções ao longo de $X$**

Tudo o que fizemos até aqui, nesta subsecção, o podemos fazer com respeito ao espaço $Y$. Ou seja, para qualquer $D \subseteq X \times Y$ e $y \in Y$, define-se a secção

$$D^y = \{x \in X;\ (x,y) \in D\}.$$

ao longo de $y \in Y$. Por um reciocínio idéntico, demonstram-se as mesmas propriedades que acabamos de ver para as secções ao longo de $X$. Para facilitar referências posteriores, enunciamos estas propriedades de forma integrada, na próxima proposição.

**Proposição 5.2.**

1. Para todo $D \in \mathcal{F} \otimes \mathcal{G}$ e todo $(x, y) \in X \times Y$, $D_x \in \mathcal{G}$ e $D^y \in \mathcal{F}$.
2. Sendo $(Z, \mathcal{H})$ um terceiro espaço mensurável e $f : X \times Y \to Z$ uma função $(\mathcal{F} \otimes \mathcal{G}) / \mathcal{H}-$mensurável. Então, para todo $(x, y) \in X \times Y$, a função $f_x : y \in Y \to f(x, y) \in Z$ é $\mathcal{G}/\mathcal{H}-$mensurável e a função $f^y : x \in X \to f(x, y) \in Z$ é $\mathcal{F}/\mathcal{H}-$mensurável.
3. Para todo $D \in \mathcal{F} \otimes \mathcal{G}$ e todo $(x, y) \in X \times Y$, $x \in X \to \nu(D_x) \in \mathcal{L}^+(X)$ e $y \in Y \to \mu(D^y) \in \mathcal{L}^+(Y)$.

Agora já podemos estabelecer o resultado principal desta subsecção, o que podemos considerar como a primeira versão do teorema de Fubini-Tonelli, ou seja o teorema de Fubini-Tonelli para funções características.

**Teorema 5.2.** *Para todo $D \in \mathcal{F} \otimes \mathcal{G}$,*

$$\mu \otimes \nu(D) = \int_X \nu(D_x)\, d\mu(x) = \int_Y \mu(D^y)\, d\nu(y). \tag{5.13}$$

***Demonstração.*** Primeiro observe que os dois integrais de (5.13) têm sentido, já que $x \to \nu(D_x) \in \mathcal{L}^+(X)$ e $y \to \mu(D^y) \in \mathcal{L}^+(Y)$ (cf. Proposição 5.2).

Se $A \in \mathcal{F}$ e $B \in \mathcal{G}$, por (5.12),

$$\int_X \nu\left((A \times B)_x\right) d\mu(x) = \int_X \nu(B) \cdot \chi_A d\mu(x) = \mu(A) \cdot \nu(B). \tag{5.14}$$

Pelo Teorema 5.1, existe uma única medida verificando (5.14). Portanto, se mostrarmos que

$$D \in \mathcal{F} \otimes \mathcal{G} \to \int_X \nu(D_x)\, d\mu(x) \tag{5.15}$$

é uma medida sobre $\mathcal{F} \otimes \mathcal{G}$, a primeira igualdade de (5.13) fica demonstrada.

Começamos por observar que $\emptyset_x = \emptyset$ e $\int_{X \times Y} \nu(\emptyset_x)\, d\mu(x) = 0$.

Agora, se $\{D_i\}_{i \in \mathbb{N}} \subseteq \mathcal{F} \otimes \mathcal{G}$ é uma sucessão de conjuntos mensuráveis, disjuntos dois a dois. Por (5.9) e (2.1), temos

$$\nu\left(\left(\bigcup_{i=1}^{\infty} D_i\right)_x\right) = \nu\left(\bigcup_{i=1}^{\infty} (D_i)_x\right) = \sum_{i=1}^{\infty} \nu\left((D_i)_x\right),$$

já que para todo $x \in X$, $\{(D_i)_x\}_{i \in \mathbb{N}}$ é uma sucessão de conjuntos $\mathcal{G}-$mensuráveis, disjuntos dois a dois.

Integrando com respeito a $\mu$, por (4.28) (cf. Proposição 4.2),

$$\int_X \nu\left(\left(\bigcup_{i=1}^{\infty} D_i\right)_x\right) \mathrm{d}\mu(x) = \int_X \left(\sum_{i=1}^{\infty} \nu\left((D_i)_x\right)\right) \mathrm{d}\mu(x)$$
$$= \sum_{i=1}^{\infty} \int_X \nu\left((D_i)_x\right) \mathrm{d}\mu(x).$$

A segunda igualdade de (5.13) demonstra-se exatamente da mesma forma. :-)

## 5.2 Teorema de Fubini-Tonelli

A seguinte proposição é simplesmente uma forma mais simétrica e intuitiva de escrever o Teorema 5.2 e será útil para explicar e perceber melhor o resultado principal de esta secção.

**Proposição 5.3.** Para todo $D \in \mathcal{F} \otimes \mathcal{G}$,

$$\int_{X\times Y} \chi_D(x,y)\, d(\mu\otimes\nu)(x,y) = \int_X \left(\int_Y \chi_D(x,y)\, d\nu(y)\right) d\mu(x)$$
$$= \int_Y \left(\int_X \chi_D(x,y)\, d\mu(x)\right) d\nu(y). \tag{5.16}$$

***Demonstração.*** O lado esquerdo de (5.16) é simplesmente $\mu\otimes\nu(D)$. As duas integrais do lado direito precisam de ser interpretadas corretamente.

Note que $\chi_D \in \mathcal{L}^+(X\times Y)$. Portanto, estritamente falando, $\int_Y \chi_D(x,y)\, d\nu(y)$ e $\int_X \chi_D(x,y)\, d\mu(x)$ não têm muito sentido. Seria mais formal escrever estes integrais utilizando a notação usada em (5.11) (cf. Corolário 5.1). i.e.

$$\int_Y \chi_D(x,y)\, d\nu(y) = \int_Y (\chi_D)_x(y)\, d\nu(y) = \nu(D_x)$$

e

$$\int_X \chi_D(x,y)\, d\mu(x) = \int_X (\chi_D)^y(x)\, d\mu(x) = \mu(D^y),$$

já que $(\chi_D)_x = \chi_{D_x}$ e $(\chi_D)^y = \chi_{D^y}$.

Pelo Teorema 5.2, as duas igualdades em (5.16) ficam estabelecidas. :-)

**Teorema 5.3.** *Seja $f \in \mathcal{L}^+(X\times Y)$. Então,*

$$\int_{X\times Y} f(x,y)\, d(\mu\otimes\nu)(x,y) = \int_X \left(\int_Y f(x,y)\, d\nu(y)\right) d\mu(x)$$

$$= \int_Y \left( \int_X f(x,y)\, d\mu(x) \right) d\nu(y)\,. \qquad (5.17)$$

***Demonstração.*** Aquí, como na Proposição 5.3, cometemos o mesmo abuso de linguagem. Isto permite escrever o resultado com maior simetría e ao mesmo tempo simplificar a notação, deixando de lado as funções auxiliares $f_x$ e $f^y$ cujo papel principal é o de estabelecer a mensurabilidade das mesmas (cf. Proposição 5.2, alinha 2). Explicitamente, o sentido dos integrais em (5.17) é o seguinte.

$$\int_Y f(x,y)\, d\nu(y) = \int_Y f_x(y)\, d\nu(y)$$

e

$$\int_X f(x,y)\, \mathrm{d}\mu(x) = \int_X f^y(x)\, \mathrm{d}\mu(x)\,.$$

Começemos por mostrar (5.17) para as funções simples. Seja

$$s = \sum_{i=1}^{n} c_i \cdot \chi_{D_i} \in \mathcal{L}^{s,+}(X \times Y)\,,$$

uma combinação linear de funções características. Por (4.7) (cf. Lema 4.2) e por (5.16) (cf. Proposição 5.3),

$$\begin{aligned}
\int_{X\times Y} s(x,y)\, d(\mu\otimes\nu)(x,y) &= \sum_{i=1}^{n} c_i \cdot \int_{X\times Y} \chi_{D_i}(x,y)\, d(\mu\otimes\nu)(x,y) \\
&= \sum_{i=1}^{n} c_i \cdot \int_X \left( \int_Y \chi_{D_i}(x,y)\, d\nu(y) \right) d\mu(x) \\
&= \int_X \left( \int_Y \left( \sum_{i=1}^{n} c_i \cdot \chi_{D_i}(x,y) \right) d\nu(y) \right) d\mu(x) \\
&= \int_X \left( \int_Y s(x,y)\, d\nu(y) \right) d\mu(x)\,.
\end{aligned}$$

Da mesma forma obtemos

$$\int_{X\times Y} s(x,y)\, d(\mu\otimes\nu)(x,y) = \int_Y \left( \int_X s(x,y)\, d\mu(x) \right) d\nu(y)\,.$$

Agora, fixemos uma função $f \in \mathcal{L}^{+}(X \times Y)$ e uma sucessão crescente $\{s_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^{s,+}(X \times Y)$, convergente pontualmente para $f$ (por exemplo a aproximação construída no Lema 4.3).

Aplicando em cada paso o Teorema 4.27 (convergência monótona), obtemos

$$\begin{aligned}\int_{X\times Y} f(x,y)\, d(\mu\otimes\nu)(x,y) &= \lim_{n\to\infty}\int_{X\times Y} s_n(x,y)\, d(\mu\otimes\nu)(x,y)\\ &= \lim_{n\to\infty}\int_X\left(\int_Y s_n(x,y)\, d\nu(y)\right) d\mu(x)\\ &= \int_X\left(\lim_{n\to\infty}\int_Y s_n(x,y)\, d\nu(y)\right) d\mu(x)\\ &= \int_X\left(\int_Y f(x,y)\, d\nu(y)\right) d\mu(x).\end{aligned}$$

Da mesma forma obtemos

$$\int_{X\times Y} f(x,y)\, d(\mu\otimes\nu)(x,y) = \int_Y\left(\int_X f(x,y)\, d\mu(x)\right) d\nu(y). \quad :-)$$

**Teorema de Fubini-Tonelli en $L^1(X\times Y)$**

**Teorema 5.4.** *Seja $f\in\mathcal{L}(X\times Y)$ (uma função a valores na reta real estendida, mensurável). Então, as três afirmações seguintes são equivalentes:*

1. $f\in L^1(X\times Y,\mu\otimes\nu)$.
2. *A função* $x\to\int_Y |f|_x\, d\nu\in L^1(X,\mu)$.
3. *A função* $y\to\int_X |f|^y\, d\mu\in L^1(Y,\nu)$.

*Para além disso, se $f\in L^1(X\times Y,\mu\otimes\nu)$, então*

$$\int_{X\times Y} f\, d(\mu\otimes\nu) = \int_X\left(\int_Y f_x\, d\nu\right) d\mu(x) = \int_Y\left(\int_X f^y\, d\mu\right) d\nu(y). \tag{5.18}$$

***Demonstração.*** Por (5.17) (cf. Teorema 5.3),

$$\int_{X\times Y} |f|\, d(\mu\otimes\nu) = \int_X\left(\int_Y |f|_x\, d\nu\right) d\mu(x) = \int_Y\left(\int_X |f|^y\, d\mu\right) d\nu(y).$$

Então, as três afirmações são todas verdadeiras ou todas falsas, dependendo do valor comúm destes integrais ser finito ou infinito.

Agora, se $f\in L^1(X\times Y,\mu\otimes\nu)$, então $f^+$ e $f^-\in L^1(X\times Y,\mu\otimes\nu)$. Para além disso, graças a (5.17) (cf. Teorema 5.3),

$$\int_{X\times Y} f^+ d(\mu\otimes\nu) = \int_X\left(\int_Y \left(f^+\right)_x d\nu\right) d\mu = \int_Y\left(\int_X \left(f^+\right)^y d\mu\right) d\nu < \infty$$

e

$$\int_{X\times Y} f^- d(\mu\otimes\nu) = \int_X \left(\int_Y \left(f^-\right)_x d\nu\right) d\mu = \int_Y \left(\int_X \left(f^-\right)^y d\mu\right) d\nu < \infty.$$

Por (4.29) (cf. Definição 4.8),

$$\begin{aligned}\textstyle\int_{X\times Y} f\, d(\mu\otimes\nu) &= \textstyle\int_{X\times Y} f^+ d(\mu\otimes\nu) - \int_{X\times Y} f^- d(\mu\otimes\nu)\\ &= \textstyle\int_X \left(\int_Y f_x^+ d\nu\right) d\mu - \int_X \left(\int_Y f_x^- d\nu\right) d\mu \qquad (5.19)\\ &= \textstyle\int_X \left(\int_Y f_x d\nu\right) d\mu.\end{aligned}$$

Efetivamente, $\int_X \left(\int_Y (f_x)^+ \mathrm{d}\nu\right) \mathrm{d}\mu < \infty$ e, por (4.31) (cf. Proposição 4.6), $\int_Y |f_x|\, d\nu < \infty$ em $\mu-$quase todo $x \in X$. Portanto, nos pontos onde $f_x \in L^1(Y,\nu)$ ($\mu-$quase todos), temos

$$\int_Y f_x\, \mathrm{d}\nu = \int_Y f_x^+ \mathrm{d}\nu - \int_Y f_x^- \mathrm{d}\nu.$$

A igualdade com o outro integral iterado, prova-se da mesma forma. :-)

## 5.3 Independência e medida produto

Como aplicação do teorema de Fubini-Tonelli, vamos revisitar a noção de independência de variáveis aleatórias, e a sua relação com o espaço produto, em particular veremos uma generalização do Lema 4.9.

**Lema 5.4.** *Sejam* $\mathrm{X} : \omega \in \Omega \to \mathrm{X}(\omega) \in \overline{\mathbb{R}}$ *e* $\mathrm{Y} : \omega \in \Omega \to \mathrm{Y}(\omega) \in \overline{\mathbb{R}}$ *duas variáveis aleatórias, ambas definidas sobre um espaço de probabilidade* $(\Omega, \mathcal{F}, \mathrm{P})$*, com distribuições de probabilidade* $\nu_{\mathrm{X}}$ *e* $\nu_{\mathrm{Y}}$*, respetivamente. Então,* X *e* Y *são independentes se e só se a distribuição de probabilidade do vetor aleatório*

$$(\mathrm{X},\mathrm{Y}) : \omega \in \Omega \to (\mathrm{X}(\omega), \mathrm{Y}(\omega)) \in \overline{\mathbb{R}}^2$$

*é* $\nu_{\mathrm{X}} \otimes \nu_{\mathrm{Y}}$*. i.e.*

$$(\mathrm{X},\mathrm{Y})_{\#}\,\mathrm{P} = \nu_{\mathrm{X}} \otimes \nu_{\mathrm{Y}}.$$

***Demonstração.*** cf. Exercício 6

**Corolário 5.2.** Se $f : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ e $g : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ são duas funções mensuráveis, então a variável aleatória $f(\mathrm{X}) \cdot g(\mathrm{Y})$ é integrável se e só se $f$ é $\nu_{\mathrm{X}}$ integrável e $g$ é $\nu_{\mathrm{Y}}$ integrável. Além disso, no caso integrável temos a igualdade

$$\int_\Omega f(\mathrm{X}) \cdot g(\mathrm{Y})\, \mathrm{dP} = \int_\Omega f(\mathrm{X})\, \mathrm{dP} \cdot \int_\Omega g(\mathrm{Y})\, \mathrm{dP}. \tag{5.20}$$

***Demonstração.*** Por (4.36) (cf. Teorema 4.3, sobre a mudança de variável),

$$\int_\Omega f(\mathrm{X}) \cdot g(\mathrm{Y})\, \mathrm{dP} = \int_{\overline{\mathbb{R}}^2} f(x) \cdot g(y)\, d(\nu_{\mathrm{X}} \otimes \nu_{\mathrm{Y}})(x, y).$$

Pelo teorema de Fubini-Tonelli (cf. Teorema 5.4, equação (5.18)),

$$\begin{aligned}
\int_{\overline{\mathbb{R}}^2} f(x) \cdot g(y)\, d(\nu_{\mathrm{X}} \otimes \nu_{\mathrm{X}})(x, y) &= \int_{\overline{\mathbb{R}}} \left( \int_{\overline{\mathbb{R}}} f(x) \cdot g(y)\, d\nu_{\mathrm{X}}(x) \right) d\nu_{\mathrm{Y}}(y) \\
&= \int_{\overline{\mathbb{R}}} g(y) \cdot \left( \int_{\overline{\mathbb{R}}} f(x)\, d\nu_{\mathrm{X}}(x) \right) d\nu_{\mathrm{Y}}(y) \\
&= \left( \int_{\overline{\mathbb{R}}} f(x)\, d\nu_{\mathrm{X}}(x) \right) \cdot \int_{\overline{\mathbb{R}}} g(y)\, d\nu_{\mathrm{Y}}(y).
\end{aligned}$$

Aplicando novamente (4.36) para mudar a variável de integração de volta ao espaço $\Omega$, obtemos (5.20). :-)

***Observação.*** O Lema 4.9 é um caso particular deste corolário.

# Exercícios

1. (cf. Lema 5.1) Mostre que a álgebra $\mathcal{A}_{X \times Y}$, gerada pelos conjuntos da forma $A \times B$, sendo $A \in \mathcal{F}$ e $B \in \mathcal{G}$ (retângulos mensuráveis), é igual à união finita de retângulos mensuráveis, disjuntos dois a dois.

2. Prove que que a função $\zeta$, definida sobre $\mathcal{A}_{X \times Y}$ (a álgebra do Exercício 1), por (5.3), depende unicamente do conjunto $A_1 \times B_1 \cup \cdots \cup A_n \times B_n$ e não da sua representação em retângulos mensuráveis, disjuntos.

3. Mostre que para provar que

$$\zeta(C) \leq \sum_{i=1}^{\infty} \zeta(C_i),$$

para toda sucessão $\{C_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A}_{X\times Y}$ e $C \in \mathcal{A}_{X\times Y}$, tal que $C \subseteq \bigcup_{i=1}^{\infty} C_i$ (sendo $\mathcal{A}_{X\times Y}$ a álgebra do Exercício 1), basta demonstrar a mesma propriedade para os elementos de $\mathcal{R}_{X\times Y}$ (os retângulos mensuráveis, i.e. os conjuntos da forma $A \times B$, sendo $A \in \mathcal{F}$ e $B \in \mathcal{G}$, cf. (3.8)). Ou seja, é suficiente provar que

$$\zeta\left(A \times B\right) \leq \sum_{i=1}^{\infty} \zeta\left(A_i \times B_i\right),$$

para toda sucessão $\{A_i \times B_i\}_{i\in\mathbb{N}} \subseteq \mathcal{R}_{X\times Y}$ e $A \times B \in \mathcal{R}_{X\times Y}$, tal que

$$A \times B \subseteq \bigcup_{i=1}^{\infty} A_i \times B_i.$$

4. (cf. Proposição 5.1) Prove que para todo $D$ e $E \subseteq X \times Y$ e todo $x \in X$,

$$\left(D \setminus E\right)_x = D_x \setminus E_x \tag{5.21}$$

Para além disso, sendo $\{D_i\}_{i\in I}$ uma família de subconjuntos de $X \times Y$, indexada por um conjunto de índices $I$, mostre que para todo $x \in X$,

$$\left(\bigcup_{i\in I} D_i\right)_x = \bigcup_{i\in I} \left(D_i\right)_x \tag{5.22}$$

e

$$\left(\bigcap_{i\in I} D_i\right)_x = \bigcap_{i\in I} \left(D_i\right)_x. \tag{5.23}$$

5. Complete a demonstração do Lema 5.3, para o caso em que a medida $\nu$ é $\sigma-$finita. Nomeadamente, sendo $\{D_i\}_{i\in\mathbb{N}} \subseteq \mathcal{D}$ uma família decrescente, prove que

$$x \to \nu\left(\left(\bigcap_{i=1}^{\infty} D_i\right)_x\right) \in \mathcal{L}^+\left(X\right).$$

6. (cf. Lema 5.4) Mostre que duas variáveis aleatórias, X e Y, definidas sobre um mesmo espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$ são independentes se e só se a distribuição de probabilidade do vetor aleatório $\omega \in \Omega \to (\mathrm{X}(\omega), \mathrm{Y}(\omega)) \in \overline{\mathbb{R}}^2$ é $\nu_{\mathrm{X}} \otimes \nu_{\mathrm{Y}}$, sendo $\nu_{\mathrm{X}}$ e $\nu_{\mathrm{Y}}$ são as distribuições de probabilidade de X e Y, respetivamente.

# Índice Alfabético


Álgebra$-\sigma$ produto, 71
Álgebra $-\sigma$ de subconjuntos, 4
Álgebra de subconjuntos, 4
Álgebra$-\sigma$ de Bernoulli, 13
Álgebra$-\sigma$ de Borel, 8
Álgebra$-\sigma$ gerada por uma função, 15
Álgebra$-\sigma$ gerada/envolvente, 8

Função de acumulação, 78

Aproximação de funções mensuráveis, 112
Aproximação de funções positivas, 110
Associatividade do produto de álgebras$-\sigma$, 72

Boreliano/Conjunto de Borel, 9
Borelianos da reta estendida, 65
Borelianos de $\mathbb{R}^n$, 73

Caracterização das funções mensuráveis, 68
Caracterização das funções mensuráveis reais, 69
Cardinal, 2
Classificação de medidas Borelianas, 84
Compacidade, 64
Composição de funções mensuráveis, 68
Conjunto das partes, 1
Conjunto de medida nula, 54
Conjunto mensurável com respeito à uma medida exterior, 43
Conjunto reticulado, 10
Conjuntos finitos, 2
Conjuntos infinitos, 2
Conjuntos numeráveis, 3
Construção da medida de Lebesgue-Stieltjes, 81
Construção de medidas, 46
Continuidade absoluta, 107
Convergência dominada, 120
Convergência em probabilidade, 125
Convergência monótona, 110

Decomposição duma medida Boreliana, 87
Definição equivalente de medida, 30
Desigualdade de Hölder, 133
Desigualdade de Minkowski, 134
Distribuição $\chi-$Quadrado., 97
Distribuição binomial, 90
Distribuição binomial negativa, 91
Distribuição geométrica, 91
Distribuição de Bernoulli, 90

Distribuição de Cauchy, 95
Distribuição de Poisson, 92
Distribuição de probabilidade, 89
Distribuição exponencial, 95
Distribuição Gamma, 96
Distribuição normal, 94
Distribuição uniforme em $[0,1]$, 93
Distribuição$-t$ de Student, 95
Distribuições absolutamente contínuas, 93
Distribuições discretas, 90

Ensaios de Bernoulli, 89
Espaço completo, 55
Espaço probabilístico, 23
Espaço $L^p(X,\mu)$, 131
espaço (mensurável) produto, 71
Espaço de medida, 23
Espaço de probabilidade, 23
Espaço mensurável, 6
Espaço normado $L^1(X,\mu)$, 114
Espaço reticulado, 10
Espaço vetorial normado, 114
Estrutura reticulada, 10
Evento, 23
Exemplo de conjunto não mensurável, 57

Família monótona, 50
Família monótona envolvente/gerada, 51
Famílias geradoras dos Borelianos, 66
Feixe de sucessões de Bernoulli, 12
Filtração de Bernoulli, 13
Função de Lebesgue-Stieltjes, 83
Função característica, 4
função de Heaviside, 86
Função de truncadura, 12
Função Integrável, 113
Função mensurável, 14
Funções mensuráveis, 68
Funções simples, 101

Independência de variáveis aleatórias, 88
Integração, 101
Integral de Lebesgue, 101
Integral de uma função não negativa, 106
Interseção de álgebras, 7
Interseção de álgebras$-\sigma$, 7
Isomorphismo entre $\left(\{0,1\}^{\mathbb{N}}, \mathcal{F}_\infty, \mathrm{P}_{\frac{1}{2}}\right)$ e $([0,1], \mathcal{B}, \lambda)$, 40

Lei dos grandes números, 125
Lei forte dos grandes números, 126
Lei fraca dos grandes números, 125
Lema de Fatou, 120
Limites de funções mensuráveis, 70
Linearidade do integral, 117

Máximo e mínimo de funções mensuráveis, 69
Média, 123
Medida, 21
Medida $\sigma-$finita, 23
Medida absolutamente contínua, 85
Medida completa, 54
Medida contínua, 87
Medida de Comprimento sobre $(0,1]$, 33

medida de Dirac, 85
Medida de Lebesgue, 57
Medida de Lebesgue-Stieltjes, 74
Medida de probabilidade em $(\mathbb{R}, \mathcal{B})$., 77
Medida definida através de uma densidade, 107
Medida discreta, 86
medida duma interseção decrescente de conjuntos, 29
Medida exterior, 42
Medida imagem, 25
Medida invariante, 27
Medida preimagem, 27
Medidas de Lebesgue-Stieltjes, 81
Medidas estrangeiras, 84, 129
Mensurabilidade das funções contínuas, 68
Mensurabilidade das projeções, 71
Mudança de variável, 119

Números normais, 131
Norma sobre $L^p(X, \mu)$, 132

Operações com funções mensuráveis, 73
Operações aritméticas na reta estendida, 64

Paradoxo de Bertrand, 76
Preintegral, 103
Preintegral de uma função simples, 103
Premedida, 31
Premedida $\sigma-$finita, 32
Probabilidade produto sobre $\{0,1\}^{\mathbb{N}}$, 37
Probabilidades contínuas, no entanto estrangeiras com respeito à medida de Lebesgue, 130
Produto de variáveis aleatórias independentes, 123
Propriedade quase falsa, 58
Propriedade quase verdadeiras, 58

Relação de ordem, 9
Relação de ordem na reta real estendida, 63
Restrição de uma álgebra$-\sigma$ a um subconjunto, 16
Reta real estendida, 63

Sequências finitas, 11
Subaditividade da medida duma união numerável, 29
Sucessões de Bernoulli, 10

Teorema da convergência dominada, 121
Teorema da convergência monótona, 112
Teoremas de convergência para conjuntos, 28
Topologia da reta estendida, 64
Transporte de medidas, 25
Truncadura duma sucessão de Bernoulli, 12

Variáveis aleatórias integráveis, 122
Variável aleatória, 74
Variável aleatória independente, 88
Variável de Bernoulli, 89
Variância, 123
Vizinhança do $\infty$, 64

# Bibliografia


[1] Patrick Billingsley, *Probability and measure*, Anniversary Edition, John Willey & Sons, 2012.

[2] Émile Borel, *Sur les probabilités dénombrables et leurs applications arithmétiques*, Rendiconti del Circolo Matematico di Palermo **Vol. 27** (1909), no. 1, 247–271.

[3] Nicolas Bourbaki, *Intégration*, Oxford University Press, Oxford, 2006.

[4] Kai Lai Chung, *A course in probability theory*, Second edition, Academic Press, 1974.

[5] Jean Dieudonné, *Éléments d'analyse, tome 2, applications à la théorie de la mesure*, Jacques Gabay, Paris, 2003.

[6] Lawrence C. Evans and Ronald F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, Florida, 1991.

[7] William Feller, *An introduction to probability theory and its applications*, 3rd edition ed., vol. Vol. 1, Wiley, 1968.

[8] Paul Halmos, *Probability and measure*, Graduate Texts in Mathematics (Book 18), Springer; 1st ed. 1950. Corr. 2nd printing, 1978.

[9] Mark Kac, *Statistical independence in probability, analysis and number theory*, The Mathematical Association of America, New Jersey, 1959.

[10] Walter Rudin, *Real and complex analysis*, McGraw-Hill International Editions, Singapore, 1987.

[11] Laurent Schwartz, *Radon measures on arbitrary topological spaces and cylindrical measures*, Oxford University Press, Paris, 1974.

[12] ———, *Analyse, tome 3, calcul intégral*, Editions Hermann, Paris, 1997.

[13] ———, *Analyse, tome 4, applications à la théorie de la mesure*, Editions Hermann, Paris, 1997.

[14] ———, *Méthodes mathématiques pour les sciences physiques*, Editions Hermann, Paris, 1997.