



Processos Estocásticos e Aplicações

Licenciatura em Matemática Aplicada à Economia e Gestão

Nuno M. Brites

Setembro de 2025

Conteúdo

	5
1 Introdução aos processos estocásticos	7
1.1 Conceitos fundamentais	7
1.2 Tipos clássicos de processos estocásticos	12
1.2.1 Processos de incrementos independentes e estacionários	12
1.2.2 Processo estocástico real de 2ª ordem	13
1.2.3 Processos estacionários	14
1.2.4 Martingalas	17
1.2.5 Processos de Markov	18
2 Cadeias de Markov em tempo discreto	19
2.1 Introdução	19
2.1.1 Conceitos básicos	19
2.2 Classificação de estados de uma C.M.	23
2.2.1 Decomposição do espaço de estados	30
2.3 Probabilidades de absorção em estados recorrentes	33
2.4 Teoremas limite	35
2.4.1 Distribuição estacionária e distribuição limite	35
2.4.2 Comportamento limite de P_{ij}^n quando $n \rightarrow +\infty$	37
3 Cadeias de Markov em tempo contínuo	43
3.1 Processo de Poisson homogêneo	43
3.2 Processo de nascimento puro	49
3.3 Processo de nascimento e morte	52
3.3.1 Definição e equações de Chapman-Kolmogorov	52
3.3.2 Tempo de espera	54
3.3.3 Equações diferenciais de processos de nascimento e morte	55
4 Complementos de processos estocásticos	59
4.1 Processo de Wiener	59
4.2 O integral de Itô	63
5 Bibliografia	77



Todas as informações relacionadas com esta UC encontram-se no Fénix.

Agradeço ao Professor Alfredo Egídio dos Reis (ISEG) a cedência de alguns exercícios presentes neste texto. Agradeço igualmente ao Professor Pedro Soares (ISEG) pela profunda e competente revisão do texto.

Todos os erros e omissões são da minha inteira responsabilidade. Caso detete algum erro ou gralha, muito agradeço que me informe. Sugestões e comentários também serão muito bem-vindos.

Obrigado,

Nuno M. Brites

nbrites@iseg.ulisboa.pt

ISEG, Setembro de 2025

Todos os direitos reservados. É expressamente proibida a reprodução, cópia, distribuição, comunicação pública, transformação ou qualquer outra forma de utilização, total ou parcial, dos conteúdos deste sítio, incluindo textos, código e imagens, sem autorização prévia e por escrito do autor. Qualquer utilização não autorizada constitui violação dos direitos de autor e poderá dar lugar à responsabilidade civil e criminal nos termos da lei em vigor.

2025 | Nuno M. Brites | nbrites@iseg.ulisboa.pt

Capítulo 1

Introdução aos processos estocásticos

1.1 Conceitos fundamentais

Nesta secção procede-se a uma revisão sumária de noções basilares de probabilidade e de variáveis aleatórias. Seguidamente, introduz-se o conceito de processo estocástico, entendido como uma família de variáveis aleatórias definidas sobre um espaço de probabilidade e indexadas por um conjunto de parâmetros, usualmente interpretados como o tempo. Finalmente, analisam-se algumas classes fundamentais de processos estocásticos, em particular os processos com incrementos independentes e estacionários, bem como os processos estacionários em sentido forte e em sentido fraco.

Designa-se, como habitualmente, por **espaço amostral** o conjunto de todos os resultados possíveis de uma experiência aleatória, representado por Ω . No que se segue, consideramos que Ω é um conjunto não vazio.

Definição 1.1 (Sigma-álgebra). Uma σ -álgebra é uma família \mathcal{F} de subconjuntos de Ω que satisfaz as seguintes propriedades:

- i) $\emptyset \in \mathcal{F}$ e $\Omega \in \mathcal{F}$;
- ii) Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$, onde A^c denota o complementar de A relativamente a Ω ;
- iii) Se $A_n \subseteq \mathcal{F}$, $n \in \mathbb{N}$, então $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$.

Os elementos de \mathcal{F} designam-se por conjuntos mensuráveis (ou \mathcal{F} -mensuráveis, para explicitar a σ -álgebra a que pertencem).

Definição 1.2 (Medida de probabilidade). Uma medida de probabilidade P na σ -álgebra \mathcal{F} é uma função $P : \mathcal{F} \rightarrow [0, 1]$ que satisfaz as seguintes propriedades:

- i) $P(\emptyset) = 0$;
- ii) $P(\Omega) = 1$;
- iii) Se $(A_n)_{n \in \mathbb{N}}$ é uma família de conjuntos dois a dois disjuntos em \mathcal{F} , então

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n).$$

Definição 1.3 (Espaço de probabilidade). Um espaço de probabilidade é um terno (Ω, \mathcal{F}, P) , onde Ω é um conjunto, \mathcal{F} é uma σ -álgebra em Ω e P é uma medida de probabilidade em \mathcal{F} .

Os elementos de \mathcal{F} chamam-se acontecimentos; $P(A)$, $A \in \mathcal{F}$, representa a probabilidade do acontecimento A .

Definição 1.4 (Sigma-álgebra de Borel). Uma σ -álgebra de Borel, \mathcal{B} , definida num conjunto E satisfaz as seguintes propriedades:

- $\emptyset \in \mathcal{B}$ e $E \in \mathcal{B}$;
- \mathcal{B} é fechada relativamente ao complementar, isto é, $\forall A \in \mathcal{B} : A^c \in \mathcal{B}$;
- \mathcal{B} é fechada relativamente à reunião numerável, isto é, se $A_i \in \mathcal{B}$ para todo $i \in \mathbb{N}$, então $\bigcup_{i=1}^n A_i \in \mathcal{B}$.

Uma σ -álgebra de Borel é um caso particular de uma σ -álgebra e aplica-se aos conjuntos abertos de E . A σ -álgebra de Borel mais comum é a σ -álgebra de Borel em \mathbb{R} , que se denota por $\mathcal{B}_{\mathbb{R}}$, ou simplesmente \mathcal{B} caso não existam ambiguidades.

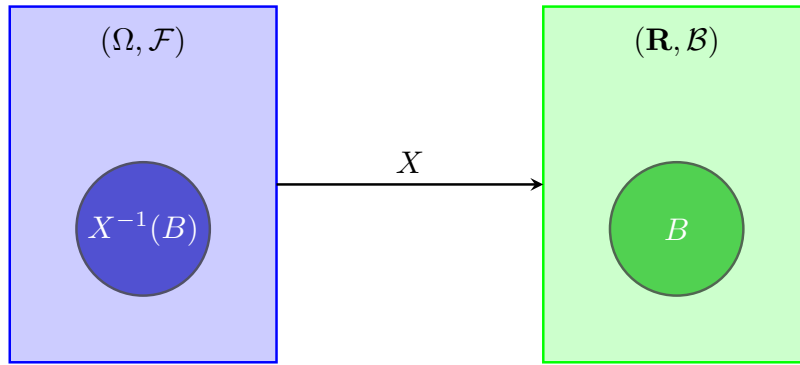
Definição 1.5 (Variável aleatória). Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Diz-se que uma função $X : \Omega \rightarrow \mathbb{R}$ é uma variável aleatória (v.a.) se

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

onde \mathcal{B} denota a σ -álgebra de Borel em \mathbb{R} .

Adicionalmente, diz-se que X é \mathcal{F} -mensurável, ou simplesmente mensurável quando a σ -álgebra associada estiver subentendida.

Em termos gráficos,



Teorema 1.1. Seja $X : \Omega \rightarrow \mathbb{R}$ uma variável aleatória. Defina-se

$$\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}\}.$$

Então, $\sigma(X)$ é a menor σ -álgebra sobre Ω para a qual X é mensurável. Esta σ -álgebra, que está contida em \mathcal{F} , designa-se por **σ -álgebra gerada por X** .

Definição 1.6 (Média e variância). Sejam (Ω, \mathcal{F}, P) um espaço de probabilidade e $X : \Omega \rightarrow \mathbb{R}$ uma variável aleatória. Define-se o **valor esperado** (ou média) e a **variância** de X da seguinte forma:

1. Caso geral (medida de probabilidade P):

$$E(X) = \int_{\Omega} X dP, \quad \text{Var}(X) = \int_{\Omega} (X - E(X))^2 dP,$$

desde que estes integrais existam e sejam finitos.

2. Caso discreto:

Se X assume valores em um conjunto discreto $\{x_1, x_2, \dots\}$ com probabilidades $p_i = P(X = x_i)$, então

$$E(X) = \sum_i x_i p_i, \quad \text{Var}(X) = \sum_i (x_i - E(X))^2 p_i.$$

3. Caso contínuo:

Se X possui densidade $f_X(x)$ relativamente à medida de Lebesgue, então

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx, \quad \text{Var}(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f_X(x) dx.$$

Definição 1.7. Sejam (Ω, \mathcal{F}, P) um espaço de probabilidade e X uma variável aleatória definida nesse espaço.

i) Diz-se que X é uma variável aleatória de quadrado integrável quando

$$E(X^2) < +\infty;$$

ii) O espaço L^2 é o conjunto das variáveis aleatórias de quadrado integrável definidas em (Ω, \mathcal{F}, P) ;

iii) A norma L^2 é a norma definida por

$$\forall X \in L^2 : \|X\|_{L^2} = (E(X^2))^{1/2}.$$

Nota . Relativamente à definição de espaço L^2 , na realidade deveríamos dizer: “espaço constituído pelas classes de equivalência de variáveis aleatórias...”, isto é, para duas variáveis aleatórias X e Y definidas em (Ω, \mathcal{F}, P) , considere-se a relação de equivalência

$$X \sim Y \iff P(X \neq Y) = 0$$

e constrói-se o espaço L^2 a partir da classe de equivalência $[X] = \{Y : X \sim Y\}$.

Definição 1.8. Seja $(X_n : n \in \mathbb{N})$ uma sucessão de variáveis aleatórias em L^2 . Diz-se que $(X_n : n \in \mathbb{N})$ converge para X em L^2 se

$$\|X_n - X\|_{L^2} \rightarrow 0 \quad \text{quando} \quad n \rightarrow +\infty,$$

ou, de modo equivalente,

$$E((X_n - X)^2) \rightarrow 0 \quad \text{quando} \quad n \rightarrow +\infty.$$

A este tipo de convergência chama-se convergência em média quadrática e representa-se por

$$X_n \xrightarrow{m.q.} X \quad \text{quando} \quad n \rightarrow +\infty$$

ou

$$l.i.m._{n \rightarrow +\infty} X_n = X.$$

Definição 1.9. Sejam X uma variável aleatória e $(X_n : n \in \mathbb{N})$ uma sucessão de variáveis aleatórias definidas no espaço de probabilidade (Ω, \mathcal{F}, P) .

i) Diz-se que X_n converge quase certamente (q.c.), ou que converge com probabilidade 1 para X , e denota-se por

$$X_n \xrightarrow{q.c.} X \quad \text{ou} \quad \lim_{n \rightarrow +\infty} X_n = X \quad q.c.,$$

se $X_n(\omega) \rightarrow X(\omega)$ para todo $\omega \in \Omega \setminus N$, onde $N \in \mathcal{F}$ é um conjunto de medida nula, isto é, $P(N) = 0$.

ii) Diz-se que X_n converge em probabilidade (ou converge estocasticamente) para X , e denota-se por

$$X_n \xrightarrow{P} X \quad \text{ou} \quad P - \lim_{n \rightarrow +\infty} X_n = X,$$

se, para todo $\delta > 0$,

$$P(|X_n - X| > \delta) \rightarrow 0 \quad \text{quando} \quad n \rightarrow +\infty.$$

Quando se pretende estudar fenómenos que não têm qualquer evolução, usam-se **amostras aleatórias** (repetições de observações i.i.d.'s). Mas, e se estivermos perante **variáveis aleatórias** que já se observaram (ou podiam observar) no passado e que poderemos observar no futuro? Tal ocorre quando pretendemos estudar, por exemplo:

- cotação diária de uma ação na bolsa de valores;
- evolução da taxa de desemprego num dado período;
- número de pessoas que chegam a uma certa fila para serem atendidas;
- evolução da temperatura num local;
- ...

Nos casos acima descritos dispomos apenas de uma única observação (chamada **trajetória**) a partir da qual se pretende extrair conclusões. Nesta trajetória não existe independência entre observações. Tipicamente pretendemos fazer:

- previsão de observações futuras;
- identificação do tipo de evolução;
- filtragem (previsão com a ajuda de observações parciais).

Definição 1.10 (Processo estocástico). Um **processo estocástico** (PE) é uma família de v.a $\{X_t, t \in T\}$, definida sobre o mesmo espaço de probabilidade (Ω, \mathcal{F}, P) e assumindo valores num mesmo espaço mensurável (E, \mathcal{B}) , onde:

- T : espaço dos parâmetros (ou do tempo);
- Ω : espaço de resultados possíveis;
- \mathcal{F} : σ -álgebra definida em Ω ;
- P : medida de probabilidade;
- E : conjunto de espaço de estados (a definir posteriormente);
- \mathcal{B} : σ -álgebra de Borel definida em E .

Nota .

- Dado um espaço de probabilidade (Ω, \mathcal{F}, P) e um conjunto arbitrário T , um PE é uma função $X(t, \omega)$ definida em $T \times \Omega$, tal que, para cada $t \in T$, $X_t(\omega)$ é uma v.a..
- O conceito de PE generaliza o de v.a. fazendo-a depender de um parâmetros t com domínio em T . Assim, podemos interpretar um PE como uma família ordenada de v.a.'s.

- Para cada ω_0 fixo, $\omega_0 \in \Omega$, $X(\omega_0, t)$ é uma função não aleatória de t . Deste modo, um PE pode identificar-se com um sistema que a cada ponto $\omega \in \Omega$, faz corresponder uma função de parâmetro t . Cada uma dessas funções diz-se uma trajetória ou realização do processo X .

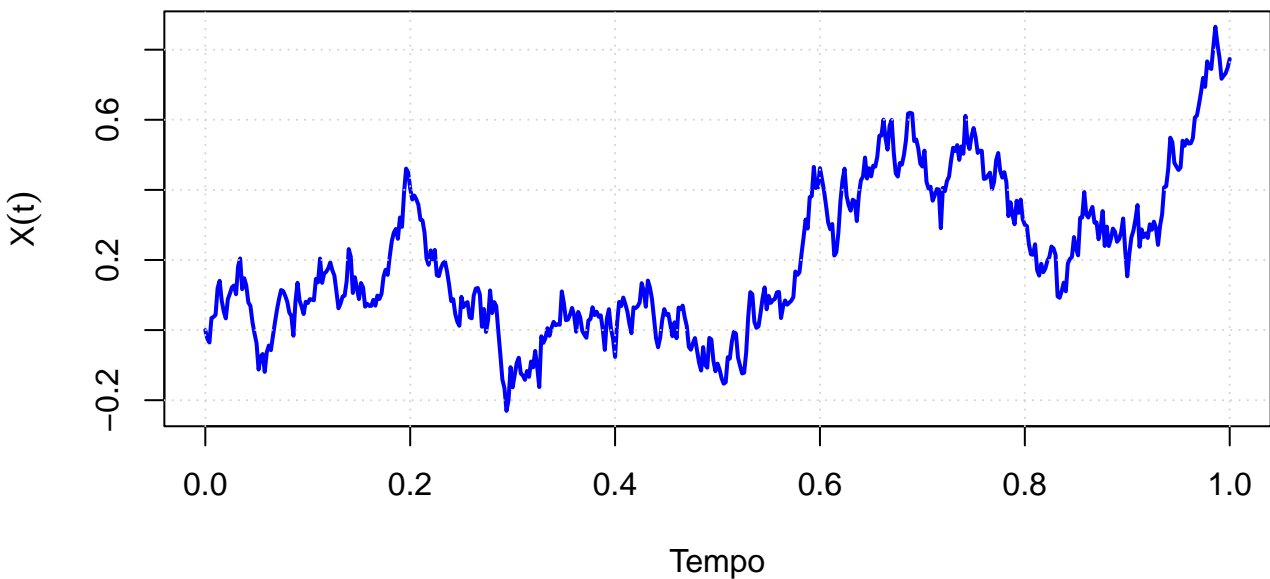
Definição 1.11 (Trajetória de um processo estocástico). Chama-se **trajetória** ou **realização** de um processo estocástico X à coleção $\{X_t(\omega), t \in T\}$, $\forall \omega \in \Omega$.

Nota . Em geral $(E, \mathcal{B}) = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$, onde:

- \mathbb{R}^n : conjunto dos possíveis valores do processo X_t ;
- $\mathcal{B}_{\mathbb{R}^n}$: σ -álgebra dos borelianos de \mathbb{R}^n ;
- Se $n = 1$ o PE chama-se processo estocástico univariado;
- Se $n > 1$ o PE chama-se processo estocástico multivariado;
- t : instante onde é feita a observação ou o período relativo a essa observação;
- Se E for finito ou infinito numerável então X é um PE de espaço de estados discreto;
- Se $E = \mathbb{R}$ então X é um PE de valores reais;
- Se T for finito ou infinito numerável então X é um PE de tempo discreto (tipicamente $T = \mathbb{N}_0$ ou $T = \mathbb{Z}$);
- Se T for infinito não numerável então X é um PE de tempo contínuo (tipicamente $T = \mathbb{R}_0^+$ ou $T = \mathbb{R}$).

Segue-se um exemplo de uma trajetória de um PE:

Exemplo de uma trajetória de um processo estocástico



Exercício 1.1. Para cada um dos seguintes processos estocásticos indique o espaço parâmetro e o espaço de estados:

- (a) Sejam X_i a quantidade de cerveja (em litros) pedida pelo i -ésimo cliente que entrou num bar e $N(t)$ o

número de clientes que chegaram ao bar até ao instante t . O processo estocástico é

$$Z_t = \sum_{i=1}^{N(t)} X_i, \quad t \geq 0,$$

onde Z_t representa a quantidade de cerveja pedida até ao instante t .

- (b) Trinta e seis pontos são escolhidos aleatoriamente no Alaska de acordo com alguma distribuição de probabilidade. Centrado em cada um desses pontos é desenhado um círculo de raio aleatório originando assim uma região Δ do Alaska. Seja $X(A)$ o preço do petróleo extraído no solo da região $A \cap \Delta$. O processo é

$$(X(B) : B \subset \text{Alaska}).$$

- (c) Um bebé dorme numa de três posições: (i) de barriga para cima com feição radiante; (ii) enrolada na posição fetal; (iii) na posição fetal, chupando o dedo polegar. Seja X_t a posição de dormir do bebé no instante t . O processo é $(X_t : t \geq 0)$.
- (d) Seja X_n o estado (ligado ou desligado) de uma fotocopiadora de um escritório ao meio-dia do n -ésimo dia. O processo é $(X_n : n = 1, 2, \dots)$.

Exercício 1.2.

Seja $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ com $P(\omega_i) = 1/4$, para $i = 1, 2, 3, 4$. Considere-se o processo estocástico $\{X(t, \omega) : t \geq 0\}$ tal que

$$X(t, \omega_i) = t \times i, \quad i = 1, 2, 3, 4.$$

- (a) Classifique o processo em causa;
- (b) Determine a função distribuição de X para $t = 1$;
- (c) Indique as trajectórias do processo;
- (d) Determine a função distribuição conjunta de $(X(1), X(2), X(3))$.

Exercício 1.3.

Considere uma sucessão infinita de provas de Bernoulli. Seja X_t o número de provas até obter um sucesso pela t -ésima vez, $t = 1, 2, \dots$

- (a) Defina o exposto como um processo estocástico, indicando o espaço dos parâmetros e dos estados.
- (b) Determine, para cada t , a função de probabilidade de X_t .
- (c) Represente graficamente uma trajectória do processo.
- (d) Determine a distribuição conjunta de (X_2, X_3, X_4) .
- (e) Calcule $P(X_4 = x \mid X_3 = x_3, X_2 = x_2)$ e $P(X_4 = x \mid X_3 = x_3)$. Comente o resultado.
- (f) Determine a distribuição da v.a. “tempo ou número de provas entre dois sucessos de Bernoulli”.
- (g) Determine a distribuição da v.a. “número de provas necessárias até à ocorrência de dois sucessos consecutivos de Bernoulli”.

1.2 Tipos clássicos de processos estocásticos

1.2.1 Processos de incrementos independentes e estacionários

Definição 1.12 (Processo com incrementos independentes). $\{X_t, t \in T\}$ é um PE com **incrementos independentes** sse

$$\forall n \in \mathbb{N}, \forall t_1, \dots, t_n \in T : t_1 < t_2 < \dots < t_n \implies X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_n} - X_{t_{n-1}}$$

são v.a.'s mutuamente independentes.

Definição 1.13 (Processo com incrementos estacionários). $\{X_t, t \in T\}$ tem **incrementos estacionários** sse $\forall s, t \in T, s < t$, a distribuição de $X_t - X_s$ depende apenas da amplitude $t - s$.

Nota . Num PE com incrementos estacionários, a distribuição de $X_{t_1+h} - X_{t_1}$ é a mesma de $X_{t_2+h} - X_{t_2}$, $\forall t_1, t_2 \in T$ e $\forall h \in \mathbb{R}_0^+$ tais que $t_1 + h, t_2 + h \in T$.

Do ponto de vista da modelação, a propriedade de independência de incrementos pode ser postulada para o modelo quando os resultados obtidos em intervalo de tempo disjuntos forem independentes. Adicionalmente, a propriedade de estacionariedade de incrementos pode ser postulada para o modelo quando for plausível que a distribuição de resultados em qualquer intervalo de tempo depende apenas da amplitude desse intervalo.

Definição 1.14 (Processo de incrementos independentes e estacionários). Dado um PE $X := \{X_t, t \in T\}$, onde T está munido de uma relação de ordem, X é um PE de **incrementos independentes e estacionários** sse tiver incrementos independentes e incrementos estacionários.

1.2.2 Processo estocástico real de 2ª ordem

Definição 1.15 (Processo Gaussiano). Diz-se que $\{X_t, t \in T\}$ é um **Processo Gaussiano** se

$$\forall n \in \mathbb{N}, \forall t_1, \dots, t_n \in T, (X_{t_1}, X_{t_2}, \dots, X_{t_n}) \sim \mathcal{N}_n(\mu, \Sigma),$$

isto é, qualquer vetor finito de variáveis aleatórias do processo tem distribuição normal multivariada.

Definição 1.16 (Processo estocástico real de 2ª ordem). Diz-se que $\{X_t, t \in T\}$ é um **processo estocástico real de 2ª ordem** se, e só se,

$$\forall t \in T : E(X_t^2) < +\infty.$$

Nestes casos, a descrição do processo faz-se habitualmente em termos dos seus dois primeiros momentos:

- **função média:** $m(t) = E(X_t), \quad \forall t \in T$;
- **função de covariância:** $\Gamma(s, t) = \text{Cov}(X_s, X_t), \quad \forall s, t \in T$.

Em geral, a informação fornecida por $m(t)$ e $\Gamma(s, t)$ não determina completamente a distribuição do processo. Contudo, no caso particular de um processo Gaussiano, a especificação destes dois primeiros momentos é suficiente para caracterizar completamente o processo.

Exemplo 1.1 (Ruído Branco Gaussiano). Chama-se **Ruído Branco Gaussiano** a um PE $\{\varepsilon_t, t \in T\}$ que satisfaz:

- $\forall t \in T, E(\varepsilon_t) = 0$;
- $\forall t \in T, \text{Var}(\varepsilon_t) = \sigma^2$;
- $\forall s, t \in T, s \neq t, \text{Cov}(\varepsilon_s, \varepsilon_t) = 0$;
- $\forall n \in \mathbb{N}, \forall t_1, t_2, \dots, t_n \in T : (\varepsilon_{t_1}, \varepsilon_{t_2}, \dots, \varepsilon_{t_n})$ é um vetor aleatório Gaussiano.

1.2.3 Processos estacionários

Definição 1.17 (Processo estacionário em sentido forte). Diz-se que um PE $\{X_t, t \in T\}$ é **estacionário em sentido forte** (ou fortemente estacionário) se:

$$\forall n \in \mathbb{N}, \forall t_1, \dots, t_n \in T, \forall h \in \mathbb{R} \text{ tal que } t_1 + h, \dots, t_n + h \in T,$$

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{d}{=} (X_{t_1+h}, \dots, X_{t_n+h}),$$

ou seja, a distribuição conjunta de qualquer vetor finito de variáveis do processo é invariante por translação do tempo.

Como consequência da estacionariedade forte, temos o seguinte Teorema:

Teorema 1.2. Se $\{X_t, t \in T\}$ é um PE de 2ª ordem e se é fortemente estacionário, então:

- $E(X_t) = m$, isto é, a média do processo é independente de t ;
- $\forall h \in T, \Gamma(t, t+h) = \text{Cov}(X_t, X_{t+h}) = \text{Cov}(X_0, X_h) = \gamma(h)$, independente de t .

Definição 1.18 (Processo estacionário em sentido fraco). Um PE $\{X_t, t \in T\}$ é **estacionário em sentido fraco** (ou estacionário de 2ª ordem), sse:

- $\forall t \in T, E(X_t^2) < +\infty$;
- $\forall t \in T, E(X_t) = m$, independente de t ;
- $\forall t \in T, \forall h \in T, \text{Cov}(X_t, X_{t+h}) = \gamma(h)$, isto é, a covariância apenas depende de h .

Nota . A função $\gamma(h)$, $\forall h \in T$, chama-se **função de autocovariância**. Se $h = 0$, então $\text{Cov}(X_t, X_{t+h}) = \text{Var}(X_t) = \gamma(0)$, $\forall t \in T$. A esta propriedade chama-se propriedade da homocedasticidade.

Vejamos agora que o Ruído Branco, $\{\varepsilon_t, t \in T\}$, é um exemplo de um PE estacionário de 2ª ordem:

Exemplo 1.2.

- $E(\varepsilon_t) = 0$;
- $\text{Var}(\varepsilon_t) = \sigma^2 \implies E(\varepsilon_t^2) < +\infty$;
- $t \neq s, \text{Cov}(\varepsilon_s, \varepsilon_t) = 0, \implies$ independência de t e de s .

Assim,

$$\gamma(h) = \begin{cases} \sigma^2, & h = 0, \\ 0, & h \neq 0. \end{cases}$$

Logo, estão satisfeitas as condições de estacionariedade fraca.

Nota (Observação importante).

$$\text{Estacionariedade forte} + E(X_t^2) < +\infty \implies \text{Estacionariedade fraca.}$$

$$\text{Estacionariedade fraca} \not\Rightarrow \text{Estacionariedade forte.}$$

Exemplo 1.3. Considere o PE $(X_t, t \in \mathbb{N})$ onde X_t tem distribuição de Cauchy, isto é, com f.d.p. $f(x) = \frac{1}{\pi(1+x^2)}$. Uma vez que não existe $E(X_t)$, então $E(X_t^2)$ não está definido. Assim, o processo é fortemente estacionário mas não é fracamente estacionário.

Propriedade 1.1 (Propriedades da função de autocovariância em processos estacionários). *A função de autocovariância $\gamma(h)$ goza das seguintes propriedades:*

- $\gamma(h) = \gamma(-h), \forall h \in \mathbb{Z}$, isto é, a função de autocovariância é par;
- $\forall n \in \mathbb{N}, \forall a_j \in \mathbb{R}, \forall t_j \in \mathbb{Z}, j = 1, \dots, n :$

$$\forall n \in \mathbb{N}, \forall a_1, \dots, a_n \in \mathbb{R}, \forall t_1, \dots, t_n \in \mathbb{Z}, \sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma(t_j - t_k) \geq 0,$$

isto é, a função de autocovariância define uma forma quadrática não-negativa.

Definição 1.19 (Função de autocorrelação em processos estacionários). Seja $\{X_t, t \in T\}$ um PE estacionário. Chama-se **função de autocorrelação** à função ρ definida por:

$$\rho(h) = \text{Corr}(X_t, X_{t+h}) = \frac{\text{Cov}(X_t, X_{t+h})}{\sqrt{V(X_t)}\sqrt{V(X_{t+h})}} = \frac{\gamma(h)}{\gamma(0)}.$$

Propriedade 1.2 (Propriedades da função de autocorrelação em processos estacionários). *A função de autocorrelação $\rho(h)$ goza das seguintes propriedades:*

- $\rho(h) = \rho(-h), \forall h \in \mathbb{Z}$, isto é, a função de autocovariância é par;
- $\forall n \in \mathbb{N}, \forall a_j \in \mathbb{R}, \forall t_j \in \mathbb{Z}, j = 1, \dots, n :$

$$\sum_{j=1}^n \sum_{k=1}^n a_j a_k \rho(t_j - t_k) \geq 0,$$

isto é, trata-se de uma função semi-definida positiva.

Exercício 1.4.

Sejam X e Y duas variáveis aleatórias com média nula, não correlacionadas e com a mesma variância $\sigma^2 > 0$. Considere-se o PE $(Z_t : t \in \mathbb{Z})$ definido por:

$$Z_t = f(t) \cdot X + g(t) \cdot Y, \quad t \in \mathbb{Z},$$

onde f e g são funções determinísticas.

- (a) Encontre expressões para f e g de modo a garantir que o processo $(Z_t : t \in \mathbb{Z})$ admita variância constante mas não seja necessariamente estacionário em sentido fraco.
- (b) Concretize f e g de modo a que $(Z_t : t \in \mathbb{Z})$ seja fracamente estacionário.

Exercício 1.5. Seja $\varepsilon = (\varepsilon_t : t \in \mathbb{Z})$ um ruído branco de variância $\sigma^2 > 0$. Considere os processos estocásticos $X = (X_t : t \in \mathbb{Z})$ e $Y = (Y_t : t \in \mathbb{Z})$ definidos do seguinte modo:

$$X_t = \varepsilon_t \quad \text{e} \quad Y_t = (-1)^t \varepsilon_t, \quad \forall t \in \mathbb{Z}.$$

- (a) Prove que X e Y são fracamente estacionários.

(b) Mostre que o processo $(Z_t = X_t + Y_t : t \in \mathbb{Z})$ é um processo não estacionário.

Exercício 1.6. Considere um processo estocástico $Y = (Y_t : t \in \mathbb{Z})$ tal que $Y_t = \varepsilon_t - \theta \varepsilon_{t-1}$, $\theta \in [-1, 1]$, onde $(\varepsilon_t : t \in \mathbb{Z})$ é um ruído branco gaussiano de variância $\sigma^2 > 0$.

- (a) Mostre que Y é gaussiano.
- (b) Determine a distribuição da variável aleatória Y_t , $\forall t \in \mathbb{Z}$.
- (c) Determine a função de autocorrelação de Y .
- (d) O que pode concluir quanto à estacionariedade forte e fraca de Y ?

Exercício 1.7. Seja $X = (X_t : t \geq 0)$ um processo estocástico, definido sobre o espaço de probabilidade (Ω, \mathcal{F}, P) , tal que, para todo $t \geq 0$, $X_t \sim \mathcal{N}(0, t)$, e $P(X_0 = 0) = 1$.

- (a) Diga em que condições será X um processo de incrementos independentes e estacionários.
- (b) Supondo que X é um processo de incrementos independentes e estacionários, mostre que: (i) $\forall t, s \in [0, +\infty[$, com $t > s$, tem-se que $X_t - X_s \sim \mathcal{N}(0, |t - s|)$; (ii) X é um processo gaussiano centrado.
- (c) Considere o processo estocástico $Y = (Y_t : t \geq 0)$ tal que:

$$Y(t) = \begin{cases} t, & X_t \geq 0 \\ -t, & X_t < 0. \end{cases}$$

Mostre que Y é um processo estocástico de segunda ordem centrado. Será Y estacionário em algum sentido? Justifique.

Exercício 1.8. Sejam $X = (X_t : t \in \mathbb{Z})$ e $(\varepsilon_t : t \in \mathbb{Z})$ dois processos estocásticos definidos sobre o espaço de probabilidade (Ω, \mathcal{F}, P) , tais que:

$$\forall t \in \mathbb{Z}, \quad X_t = \sum_{j=0}^{+\infty} \left(\frac{4}{5}\right)^j \varepsilon_{t-j}.$$

- (a) Explique em que condições será ε um ruído branco.
- (b) Suponha que ε é um ruído branco tal que $E[\varepsilon_t^2] = 9/50$. (i) Prove que X é fracamente estacionário e indique as respectivas função média e função de autocovariância; (ii) Suponha agora que X é um processo gaussiano. Indique a distribuição do vector aleatório (X_t, X_s) , $\forall t, s \in \mathbb{Z}$.
- (c) Considere o processo estocástico $Y = (Y_t : t \in \mathbb{Z})$ tal que:

$$Y_t = \begin{cases} 1/2, & X_t \geq 0 \\ -1, & X_t < 0, \end{cases}$$

admitindo que X está nas condições da alínea b) ii). Calcule a função média de Y e mostre que Y é fracamente estacionário.

Exercício 1.9. Seja $(\varepsilon_t : t \in \mathbb{Z})$ um ruído branco gaussiano de variância $\sigma^2 > 0$. Considere um outro processo estocástico $(Y_t : t \in \mathbb{Z})$ definido por:

$$Y_t = \varepsilon_t - \theta \varepsilon_{t-1} - \frac{\theta}{2} \varepsilon_{t-2}, \quad \theta \in [-1, 1].$$

- (a) Defina processo gaussiano e mostre que Y é gaussiano.
- (b) Determine a função de autocorrelação do processo Y .

1.2.4 Martingalas

Do ponto de vista da modelação, as martingalas são apropriadas para modelar fenómenos aleatórios, tais como jogos de azar.

Definição 1.20 (Martingala). Um PE $\{X_t, t \in T\}$ é uma **Martingala** sse:

- $E(|X_t|) < +\infty$;
- $\forall n \in \mathbb{N}, \forall t_1 < \dots < t_{n+1} \in T : E(X_{t_{n+1}} | X_{t_1}, \dots, X_{t_n}) = X_{t_n}$.

Exemplo 1.4. Considere-se E discreto e $T = \mathbb{N}$. Se interpretarmos X_n como a fortuna de um jogador após a realização do n —ésimo jogo, então a 2ª condição da definição anterior estabelece que a fortuna **esperada** após a $(n+1)$ —ésima partida do jogo é igual à fortuna depois do n —ésimo jogo, independentemente do que ocorreu anteriormente.

Nota . Na definição de Martingala, podemos ainda considerar,

- Submartingalas, quando $\forall n \in \mathbb{N}, \forall t_1 < \dots < t_{n+1} \in T : E(X_{t_{n+1}} | X_{t_1}, \dots, X_{t_n}) \leq X_{t_n}$.
- Supermartingalas, quando $\forall n \in \mathbb{N}, \forall t_1 < \dots < t_{n+1} \in T : E(X_{t_{n+1}} | X_{t_1}, \dots, X_{t_n}) \geq X_{t_n}$.

Exercício 1.10.

Sejam X_0, X_1, \dots v.a.'s independentes com média finita e nula e $S_n = \sum_{i=0}^n X_i$. Mostre que o PE $\{S_n : n \in \mathbb{N}_0\}$ é uma Martingala.

Exercício 1.11.

Considere um jogo no qual, em cada jogada, o jogador pode ganhar ou perder um euro, com igual probabilidade. Após n jogadas o ganho desse jogador é dado por $S_n = \sum_{i=1}^n X_i$, onde X_1, X_2, \dots são v.a.'s independentes. Mostre que o PE $\{S_n : n \in \mathbb{N}\}$ é uma Martingala.

Exercício 1.12.

Sejam X_1, X_2, \dots são v.a.'s independentes com média unitária. Mostre que o PE $\{Z_n : n \in \mathbb{N}\}$, definido por

$$Z_n = \prod_{i=1}^n X_i$$

é uma Martingala.

Exercício 1.13. Seja $(X_n, n = 0, 1, 2, \dots)$ um PE com espaço de estados \mathbb{N}_0 , com média unitária para $n \geq 1$, com incrementos independentes e tal que $P(X_0 = 0) = 1$.

- (a) O que significa dizer que o processo X tem incrementos independentes?
- (b) Prove que o processo $(X_n, n = 0, 1, 2, \dots)$ é uma Martingala.
- (c) Sabendo que $Var(X_n) = 1$, o que pode afirmar quanto à estacionariedade fraca do processo $(X_n, n = 0, 1, 2, \dots)$?

1.2.5 Processos de Markov

Os processos de Markov são apropriados na modelação de fenómenos aleatórios cujo comportamento futuro não é alterado pelo conhecimento do seu passado, apenas interessa conhecer o estado presente, ou seja, a probabilidade de que o sistema físico esteja num determinado estado num dado instante t pode deduzir-se a partir do conhecimento desse estado num instante qualquer anterior e essa probabilidade não depende da “história” do sistema antes de t .

Definição 1.21 (Processo de Markov). Um PE $\{X_t, t \in T\}$ com espaço de estados E diz-se um **processo de Markov** (ou **Markoviano**) sse $\forall n \in \mathbb{N}, \forall t_1 < \dots < t_{n+1} \in T, \forall x_1, \dots, x_{n+1} \in E, \forall B \in \mathcal{B}$:

$$P(X_{t_{n+1}} \in B \mid X_{t_1} = x_1, \dots, X_{t_n} = x_n) = P(X_{t_{n+1}} \in B \mid X_{t_n} = x_n).$$

Teorema 1.3. Se E for discreto e $T = \mathbb{N}$, a propriedade de Markov da definição anterior é equivalente à seguinte:

$$\begin{aligned} \forall n \in \mathbb{N}, \forall x_0, \dots, x_{n+1} \in E : P(X_0 = x_0, \dots, X_n = x_n) > 0, \text{ tem-se que} \\ P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} \mid X_n = x_n). \end{aligned}$$

Nota . Os processos de Markov, como quaisquer processos, são classificados de acordo com a natureza do espaço de estados E e do espaço dos parâmetros T . Uma classe especial de processos de Markov são as **Cadeias de Markov** (C.M.): processos de Markov com espaço de estados E **discreto**.

Assim, uma cadeia de Markov pode interpretar-se com um PE cujo desenvolvimento se pode considerar como uma série de transições entre valores determinados que têm a propriedade de que a distribuição de probabilidade do estado futuro do processo, sabendo-se que ele está num dado estado, depende apenas deste estado e não do modo de como o processo lá chegou. As C.M. são classificadas em **discretas** ou **contínuas**. Nesta UC iremos abordar ambos os casos.

Capítulo 2

Cadeias de Markov em tempo discreto

2.1 Introdução

Uma cadeia de Markov em tempo discreto, $\{X_t, t \in T\}$, é um PE de Markov cujo espaço de estados é **finito** ou **infinito numerável**.

2.1.1 Conceitos básicos

Definição 2.1 (Cadeia de Markov em tempo discreto). Um PE em tempo discreto $(X_n, n \in \mathbb{N}_0)$ com espaço de estados E discreto é uma **C.M. em tempo discreto** sse satisfaz a propriedade de Markov

$\forall n \in \mathbb{N}, \forall x_0, \dots, x_{n+1} \in E^{n+1} : P(X_0 = x_0, \dots, X_n = x_n) > 0$, tem-se que

$$P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} \mid X_n = x_n).$$

Exercício 2.1. Seja $(X_n, n = 1, 2, \dots)$ uma sucessão de variáveis aleatórias i.d.d. com distribuição de probabilidade definida por:

$$P(X_n = x) = p(1-p)^x, \quad x = 0, 1, 2, \dots; \quad p \in (0, 1).$$

Considere ainda o PE

$$Y = (Y_n = \sum_{i=1}^n X_i : n = 1, 2, \dots).$$

- (a) Identifique o espaço de estados de Y .
- (b) Mostre que $P(Y_{n+1} = y \mid Y_n = x)$ é independente de n .
- (c) Verifique que Y é uma Cadeia de Markov.
- (d) Calcule $P(Y_1 = y_1, Y_3 = y_3)$.

Exercício 2.2. Seja $(X_n, n = 1, 2, \dots)$ uma sucessão de variáveis aleatórias i.d.d. com distribuição de probabilidade definida por:

$$P(X_n = 1) = p = 1 - P(X_n = -1), \quad n = 1, 2, \dots; \quad p \in (0, 1).$$

Considere o PE $S = (S_n : n \geq 0)$, conhecido como passeio aleatório simples, definido por:

$$S_0 = 0 \quad \text{e} \quad S_n = \sum_{i=1}^n X_i, \quad n = 1, 2, \dots$$

- (a) Identifique o espaço de estados do processo S .

- (b) Prove que S é uma Cadeia de Markov para qualquer valor de p .
- (c) Determine para que valores de p , o passeio aleatório S é uma Martingala.
- (d) Calcule a função de autocovariância do processo S e verifique se o processo é estacionário em sentido estrito.

Associada a uma C.M. em tempo discreto tem-se a **função de probabilidade de transição a um passo**:

$$P(X_{n+1} = j \mid X_n = i) := P_{ij}(n, n+1),$$

que representa a probabilidade de X_{n+1} estar no estado j no instante $n+1$ sabendo que no instante n a cadeia estava no estado i .

Se as probabilidades $P_{ij}(n, n+1)$ não dependerem de n , então a C.M. em tempo discreto diz-se **homogénea**. Assim, numa C.M. homogénea observa-se

$$P(X_{n+1} = j \mid X_n = i) := P_{ij}, \forall i, j \in E, \forall n \in \mathbb{N}_0.$$

Nota . A expressão $P(X_{n+1} = j \mid X_n = i) := P_{ij}$:

- representa a probabilidade de cadeia ir do estado i para o estado j **num só passo**;
- é independente de n , isto é, é homogénea no tempo.

Nesta UC apenas iremos estudar C.M. homogéneas. As probabilidades de transição, P_{ij} são fundamentais para o estudo da estrutura probabilística das C.M.

Definição 2.2 (Matriz de transição). Define-se **Matriz de transição** ou **Matriz de probabilidade de transição** de uma C.M. homogénea à matriz

$$\mathbb{P} = [P_{ij}]_{i,j \in E} = \begin{bmatrix} P_{00} & P_{01} & \dots & P_{0j} & \dots \\ P_{10} & P_{11} & \dots & P_{1j} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{i0} & P_{i1} & \dots & P_{ij} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}, \quad E = \mathbb{N}_0$$

definida pelas probabilidades de transição P_{ij} do processo.

Nota . Na matriz de transição \mathbb{P} , observa-se:

- $0 \leq P_{ij} \leq 1$, $i, j = 0, 1, \dots$, uma vez que P_{ij} representa uma probabilidade;
- $\sum_{j=0}^{+\infty} P_{ij} = 1$, uma vez que $\sum_{j \in E} P_{ij} = \sum_{j \in E} P(X_{n+1} = j \mid X_n = i)$.

Exercício 2.3. Quatro bolas, duas brancas e duas pretas, são distribuídas em duas caixas A e B , de tal forma que, em cada caixa, ficam duas bolas. Tira-se uma bola de cada caixa e coloca-se cada uma na caixa oposta. Seja X_0 o número de bolas brancas que existiam inicialmente na caixa A . Para $n \geq 1$, seja X_n o número de bolas brancas que existirão na caixa A depois de se terem efetuado n trocas de bolas.

- (a) Identifique o espaço de estados.
- (b) Determine a matriz das probabilidades de transição.

Teorema 2.1. *Um PE $(X_t, t \in \mathbb{N}_0)$ é uma C.M. homogénea sse existe uma distribuição de X_0 e uma matriz estocástica \mathbb{P} tal que:*

$$\forall (i_0, i_1, \dots, i_n) \in \mathbb{N}_0^{n+1} : P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = P(X_0 = i_0)P_{i_0 i_1} \dots P_{i_{n-1} i_n}.$$

Este Teorema mostra que a distribuição do vetor aleatório (X_0, X_1, \dots, X_n) fica completamente determinada se conhecermos a distribuição inicial de X_0 e matriz de transição \mathbb{P} .

Exemplo 2.1 (Passeio aleatório). Seja $(Y_n)_{n \in \mathbb{N}_0}$ uma sequência de variáveis aleatórias independentes que tomam valores em \mathbb{Z} . A primeira variável, Y_0 , pode ter uma distribuição diferente, mas as outras Y_n , para $n \geq 1$, têm todas a mesma distribuição $p(x)$, para $x \in \mathbb{Z}$.

O processo definido por

$$X_n = \sum_{m=0}^n Y_m, \quad n \in \mathbb{N}_0,$$

chama-se **passeio aleatório** (ou deslocação de um ponto no plano), tendo estado inicial X_0 e distribuição p .

Exercício 2.4.

Mostre que o processo aleatório definido no exemplo anterior é uma C.M. homogénea com matriz de transição com probabilidades

$$P_{xy} = p(y - x), \quad \forall (x, y) \in \mathbb{Z}.$$

Exemplo 2.2 (Passeio aleatório unidimensional). Um passeio aleatório a uma dimensão é uma C.M. cujo espaço de estados é um conjunto finito ou infinito numerável de inteiros, no qual a partícula estando no estado i , pode deslocar-se, numa transição, para o estado $i + 1$, $i - 1$, ou permanecer no estado i , isto é,

$$P(X_{n+1} = j \mid X_n = i) = \begin{cases} q_i, & j = i - 1 \\ r_i, & j = i, \\ p_i, & j = i + 1. \end{cases}$$

A matriz de transição é definida por

$$\mathbb{P} = \begin{bmatrix} r_0 & p_0 & 0 & 0 & \dots \\ q_1 & r_1 & p_1 & 0 & \dots \\ 0 & q_2 & r_2 & p_2 & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix},$$

onde $q_i > 0$, $r_i > 0$, $p_i > 0$; $q_i + r_i + p_i = 1$, $i = 1, 2, \dots$; $p_0 \geq 0$, $r_0 \geq 0$: $r_0 + p_0 = 1$.

Em particular, quando no instante n o processo está no estado i , isto é, quando $X_n = i$, observa-se:

- $P(X_{n+1} = i + 1 \mid X_n = i) = p_i$;
- $P(X_{n+1} = i - 1 \mid X_n = i) = q_i$;
- $P(X_{n+1} = i \mid X_n = i) = r_i$.

Exemplo 2.3 (Versão discreta do movimento Browniano). Trata-se de um caso particular do passeio aleatório unidimensional com:

$$q_i = q, \quad r_i = r, \quad p_i = p, \quad \forall i.$$

Interpretação: uma partícula, no instante t , pode efetuar 3 movimentos:

- deslocar-se para a direita com probabilidade p ;
- deslocar-se para a esquerda com probabilidade q ;
- manter-se na mesma posição com probabilidade r ;

Se X_n representar a posição da partícula ao fim de n movimentos, tem-se

$$X_n = \sum_{i=0}^n Y_i,$$

onde

$$Y_i = \begin{cases} 1, & \text{partícula deslocou-se para a direita,} \\ -1, & \text{partícula deslocou-se para a esquerda,} \\ 0, & \text{partícula não se deslocou.} \end{cases}$$

O que representam os processos X e Y ?

- $(Y_n, n \in \mathbb{N}_0)$: sucessão de v.a.'s independentes (são independentes da posição X_n da partícula);
- $(X_n, n \in \mathbb{N}_0)$: passeio aleatório, isto é, CM homogénea, com espaço de estados \mathbb{Z} e matriz de transição \mathbb{P} onde:

$$P_{ij} = P(X_n = j \mid X_{n-1} = i) = \begin{cases} p, & j = i + 1, \\ q, & j = i - 1, \\ r, & j = i. \end{cases}$$

Assim, $P_{ij} = p(j - i)$ representa a lei de probabilidade (distribuição de probabilidade) de Y_n . Tipicamente considera-se $p = q = 0.5$ e $r = 0$.

Nota . Notação:

$$P_{ij}^m = P(X_{n+m} = j \mid X_n = i)$$

representa a probabilidade de, em m passos, a cadeia de Markov homogénea passar do estado i para o estado j .

Teorema 2.2. *Seja $\mathbb{P} = [P_{ij}]$ a matriz de transição a um passo de uma C.M. $(X_n, n \in \mathbb{N}_0)$. Então,*

$$P_{ij}^m = \sum_{k=0}^{+\infty} P_{ik}^r P_{kj}^s,$$

onde $(r, s) \in \mathbb{N}_0^2$ tal que $r + s = m$ e $P_{ij}^0 = 1$ se $i = j$, e $P_{ij}^0 = 0$ se $i \neq j$.

Exercício 2.5.

Considere $m = 2$ e prove o Teorema anterior.

Teorema 2.3 (Equações de Chapman-Kolmogorov). *O Teorema anterior pode ser re-escrito como*

$$P_{ij}^{m+n} = \sum_{k \in E} P_{ik}^m P_{kj}^n, \quad \forall i, j \in E.$$

Assim, $P_{i,j}^m$ representa o elemento (i, j) da matriz potência de ordem m de \mathbb{P} .

2.2 Classificação de estados de uma C.M.

Torna-se importante o estudo limite de $P_{i,j}^n$ quando $n \rightarrow +\infty$. Espera-se que a influência do estado inicial i diminua com o tempo, e que o limite de $P_{i,j}^n$ quando $n \rightarrow +\infty$ seja independente de i .

Para se poder analisar o comportamento assintótico do processo, vamos introduzir alguns critérios de classificação de estados de uma C.M.. Consideremos, no que se segue, uma C. M. $(X_n, n \in \mathbb{N}_0)$ com matriz de transição $\mathbb{P} = [P_{i,j}]$, $i, j \in E$.

Definição 2.3 (Estado acessível). Diz-se que o estado $j \in E$ é **acessível** a partir do estado $i \in E$, se para algum $n \in \mathbb{N}_0$ se observa $P_{ij}^n > 0$. Representação:

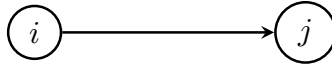


Figure 2.1: Estado acessível

Definição 2.4 (Estados em comunicação). Se dois estados $i, j \in E$ são acessíveis um relativamente ao outro, diz-se que intercomunicam ou que estão **em comunicação**. Representação:

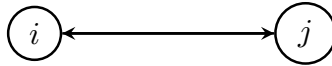


Figure 2.2: Estados em comunicação

Nota . Se dois estados $i, j \in E$ não comunicam, então:

$$\forall n \in \mathbb{N}_0, P_{ij}^n = 0 \vee P_{ji}^n = 0.$$

Teorema 2.4. A intercomunicação dos estados define uma relação de equivalência.

Exercício 2.6.

Mostre o Teorema anterior. Sugestão: mostre que a intercomunicação entre estados é reflexiva, simétrica e transitiva.

Nota . A relação de equivalência do último Teorema induz uma partição do conjunto de todos os estados em classe de equivalência. Dentro de cada classe todos os estados comunicam entre si.

Exemplo 2.4. Considere a seguinte matriz de transição, com $E = \{0, 1, 2, \dots, r\}$, de um passeio aleatório:

$$\mathbb{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ q & 0 & p & 0 & \dots & 0 & 0 \\ 0 & q & 0 & p & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$$

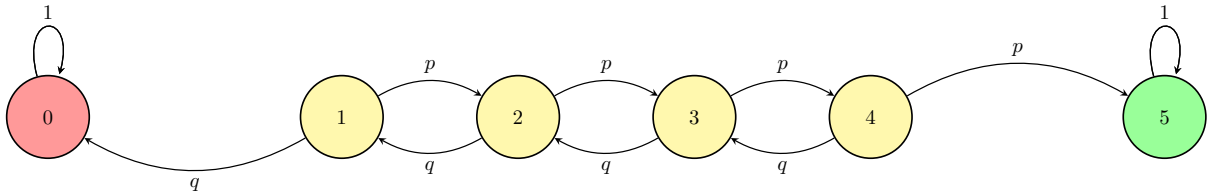
A relação de equivalência de comunicação induz as classes:

$$\{0\}, \{1, 2, \dots, r-1\}, \{r\},$$

donde se conclui que: do estado 0 só se pode ir para o estado 0; todos os estados $1, 2, \dots, r-1$ comunicam entre si; do estado r só se pode ir para o estado r . Para $r = 5$, isto é, 6 estados, a matriz de transição é:

$$\mathbb{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & 0 \\ 0 & 0 & q & 0 & p & 0 \\ 0 & 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

e a representação gráfica é:



As classes de equivalência estão representadas por cores.

Definição 2.5 (Período). Chama-se **período** do estado i , e representa-se por $d(i)$, ao máximo divisor comum de todos os inteiros $n \geq 1$ tais que $P_{ii}^n > 0$, isto é,

$$d(i) = \text{m.d.c.}(n \geq 1 : P_{ii}^n > 0).$$

Convenção: se $\forall n \geq 1, P_{ii}^n = 0$, então $d(i) = 0$.

Nota . Dois estados em comunicação têm o mesmo período, isto é,

$$i \leftrightarrow j \iff d(i) = d(j).$$

Exercício 2.7. Determine o período dos estados do exemplo anterior.

Exercício 2.8. Considere a C.M. de espaço de estados finito com matriz de transição

$$\mathbb{P} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad E = \{0, 1, \dots, n-1\}.$$

Mostre que $d(i) = n, \forall i \in E$.

Definição 2.6 (Estado aperiódico. Cadeia aperiódica.). Um estado diz-se **aperiódico** se tem período um. A cadeia é **aperiódica** se todos os estados acessíveis entre si (isto é, na mesma classe de comunicação) são aperiódicos.

Definição 2.7 (Estado recorrente). Um estado i diz-se **recorrente** se, partindo de i , a probabilidade de regressar a i em algum instante futuro é igual a 1.

Seja

$$T_i = \inf\{n \geq 1 : X_n = i \mid X_0 = i\},$$

o **tempo de retorno ao estado** i , isto é, o primeiro instante (após a partida) em que a cadeia volta a atingir o estado i .

Então, o estado i é recorrente se

$$P(T_i < +\infty) = 1.$$

De forma equivalente, um estado i é recorrente se

$$\sum_{n=1}^{\infty} P_{ii}^n = +\infty,$$

onde $P_{ii}^n = P(X_n = i \mid X_0 = i)$ representa a probabilidade de estar em i ao fim de n passos, partindo de i .

Definição 2.8 (Estado transitório). Um estado i diz-se **transitório** se, partindo do estado i , a probabilidade de retorno a i é menor que 1, isto é,

$$P(T_i < +\infty) < 1,$$

equivalentemente,

$$\sum_{n=1}^{\infty} P_{ii}^n < +\infty.$$

Definição 2.9 (Tempo mínimo de passagem). Considere-se a v.a. T_{ij} , que representa o **tempo da primeira passagem** do estado i ao estado j , definida por

$$T_{ij} = \min\{n \geq 1 : X_n = j\}.$$

A função de probabilidade de T_{ij} , denotada por f_{ij}^n , é dada por:

$$f_{ij}^n = P_i(X_n = j, X_1 \neq j, \dots, X_{n-1} \neq j), \quad n \geq 1.$$

Isto é, f_{ij}^n representa a probabilidade de, partindo do estado i , chegar ao estado j pela primeira vez ao fim de n passos.

Em particular:

$$f_{ij}^1 = P_{ij}, \quad \text{e} \quad f_{ij}^0 = 0, \quad \forall i, j \in E, \quad i \neq j.$$

Uma outra forma de interpretar T_{ij} é como o tempo de espera até à primeira chegada ao estado j , tendo começado em i .

Para $n \geq 2$ e $i \neq j$, a função de primeira passagem satisfaz a seguinte relação de recorrência:

$$f_{ij}^n = \sum_{k \in E \setminus \{j\}} P_{ik} f_{kj}^{n-1}.$$

Esta fórmula diz-nos que, para chegar pela primeira vez a j no passo n , o processo deve ir de i para um estado $k \neq j$ no primeiro passo (com probabilidade P_{ik}), e depois ir de k até j pela primeira vez ao fim de $n - 1$ passos (com probabilidade f_{kj}^{n-1}).

Condições iniciais:

$$f_{ij}^1 = P_{ij}, \quad \text{e} \quad f_{ij}^0 = 0, \quad \forall i, j \in E, \quad i \neq j.$$

Teorema 2.5.

$$\forall n \in \mathbb{N} : P_{ij}^n = \sum_{k=1}^n f_{ij}^k P_{jj}^{n-k}.$$

Exercício 2.9. Faça a demonstração do Teorema anterior.

Teorema 2.6. Considere-se a probabilidade:

$$f_{ij} = P(\text{partindo de } i, \text{ eventualmente chegar a } j, \text{ num número finito de passos}).$$

Então,

$$f_{ij} = \sum_{n=1}^{+\infty} f_{ij}^n,$$

isto é, f_{ij} representa a soma da probabilidade de sair de i , chegar a j , pela 1ª vez, em k passos, e

$$f_{ij} > 0 \iff j \text{ é acessível a partir de } i.$$

Definição 2.10 (Estado recorrente). O estado i diz-se **recorrente** (ou persistente) sse

$$\sum_{n=1}^{+\infty} f_{ii}^n = 1,$$

isto é, dizer que o estado i é **recorrente** significa que, partindo de i , volta-se com probabilidade 1 a i , num número finito de transições. Por outras palavras, há sempre a possibilidade de sair de i e chegar a i em n passos.

Em termos da v.a.

$$T_{ij} = \text{tempo mínimo para transitar do estado } i \text{ para o estado } j,$$

uma vez que

$$f_{ij}^n = P(T_{ij} = n),$$

então a definição de recorrência pode ser reescrita como:

$$i \text{ é recorrente} \iff P(T_{ii} < +\infty) = 1.$$

Definição 2.11 (Estado transitório). O estado i diz-se **transitório** se a probabilidade de voltar a i , partindo de i , for estritamente menor que 1, ou seja:

$$\sum_{n=1}^{+\infty} f_{ii}^n < 1.$$

Isto significa que, partindo de i , existe uma probabilidade positiva de **não regressar** a i . Em termos da variável aleatória T_{ii} , isto é equivalente a:

$$i \text{ é transitório} \iff P(T_{ii} < +\infty) < 1.$$

Muito usado para a verificação se um estado é recorrente é o seguinte Teorema:

Teorema 2.7 (Teorema de Abel). Um estado i é recorrente sse

$$\sum_{n=1}^{+\infty} P_{ii}^n = +\infty.$$

Teorema 2.8 (Critério de recorrência e transitoriedade).

$$\sum_{n=1}^{+\infty} P_{ii}^n = +\infty \iff i \text{ é recorrente.}$$

$$\sum_{n=1}^{+\infty} P_{ii}^n < +\infty \iff i \text{ é transitório.}$$

$$i \longleftrightarrow j \text{ e } i \text{ é recorrente} \implies j \text{ é recorrente.}$$

Nota . O último critério apresentado no Teorema anterior mostra que a recorrência é uma propriedade da classe, isto é, numa classe de equivalência todos os estados são recorrentes ou não são recorrentes. Consequentemente:

- de um estado recorrente só se podem alcançar estados recorrentes;
- o conjunto dos estados recorrentes forma num conjunto fechado.

Teorema 2.9. Se i é um estado recorrente e se j é acessível a partir de i , então $f_{ij} = 1$.

Definição 2.12 (Conjunto fechado). Seja um C.M. com espaço de estados E e matriz de transição $\mathbb{P} = [P_{ij}]_{i,j \in E}$. Diz-se que um conjunto $C \subset E$ é **fechado** sse

$$\forall i \in C, \forall j \notin C, P_{ij} = 0,$$

ou seja, sse nenhum estado **não pertencente** a C é acessível a partir de qualquer estado do conjunto C . Isto quer dizer que:

- não existem “setas” de dentro para fora de C ;
- uma classe de estados não é necessariamente um conjunto fechado.

Resumo: critérios para estabelecer a recorrência/transitoriedade

- Critério 1:

$$\sum_{n=1}^{+\infty} f_{ii}^n = 1 \implies \text{estado recorrente.}$$

$$\sum_{n=1}^{+\infty} f_{ii}^n < 1 \implies \text{estado transitório.}$$

- Critério 2:

$$P(T_{ii} < +\infty) = 1 \implies \text{estado recorrente.}$$

$$P(T_{ii} < +\infty) < 1 \implies \text{estado transitório.}$$

- Critério 3:

$$\sum_{n=0}^{+\infty} P_{ii}^n = +\infty \implies \text{estado recorrente.}$$

$$\sum_{n=0}^{+\infty} P_{ii}^n < +\infty \implies \text{estado transitório.}$$

Resumo: classificação de estados recorrentes

Defina-se $E[T_{ii}] = \mu_i = \sum_{n=1}^{+\infty} n f_{ii}^n$, isto é, μ_i representa o tempo de recorrência médio de um estado i (recorrente).

O estado i é:

- **recorrente positivo** se $\mu_i < +\infty$, isto é, o tempo médio de recorrência do estado i é finito;
- **recorrente nulo** se $\mu_i = +\infty$;
- **recorrente ergódico** se i for recorrente positivo e aperiódico.

Definição 2.13 (Cadeia de Markov irredutível). Uma cadeia de Markov com espaço de estados E diz-se **irredutível** se, para quaisquer estados $i, j \in E$, existe um número natural $n \geq 1$ tal que

$$P_{ij}^n = P(X_n = j \mid X_0 = i) > 0,$$

isto é, existe uma probabilidade positiva de se atingir o estado j a partir do estado i num número finito de passos.

Em termos informais, isto significa que todos os estados da cadeia **comunicam entre si**, ou seja, é possível ir de qualquer estado para qualquer outro (talvez em vários passos), com probabilidade positiva.

Exemplo 2.5. Considere uma cadeia de Markov com espaço de estados $\{0, 1, 2\}$ e matriz de transição

$$\mathbb{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Esta cadeia é irredutível porque, partindo de qualquer estado, é possível atingir qualquer outro estado num número finito de passos, com probabilidade positiva. Por exemplo:

$$0 \longrightarrow 1 \longrightarrow 2 \longrightarrow 0,$$

o que permite circular por todos os estados.

Por último, temos o seguinte Teorema que ajuda a classificar os estados de uma cadeia de Markov:

Teorema 2.10. Numa **cadeia de Markov irredutível** com espaço de estados finito ou infinito numerável, **ou** todos os estados são transitórios, **ou** todos os estados são recorrentes nulos, **ou** todos os estados são recorrentes positivos.

Exercício 2.10. Seja X_0, X_1, X_2, \dots uma sucessão de variáveis aleatórias i.i.d.'s tal que:

$$P(X_i = -1) = P(X_i = 1) = 0.5, \quad i = 0, 1, 2, \dots$$

(a) Prove que $(X_n : n \in \mathbb{N}_0)$ é uma C.M. homogênea com matriz de transição

$$\mathbb{P} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

(b) Determine a probabilidade de que o processo $(X_n : n \in \mathbb{N}_0)$, partindo do estado 1, volte a atingir este estado, pela primeira vez, num número par de passos.

(c) Considere o PE $(Y_n : n \in \mathbb{N})$ definido por:

$$Y_n = \begin{cases} X_{n-1} \cdot X_{n+1}, & n \text{ par} \\ X_n, & n \text{ ímpar} \end{cases}.$$

Verifique se Y é um ruído branco e, em caso afirmativo, identifique a sua variância. Prove que Y não é uma C.M.

Exercício 2.11. Considere uma estação de táxis, onde de s em s segundos chega um táxi. Se não existem clientes a ser servidos, o táxi parte de imediato, chegando outro s segundos depois. Caso contrário, os clientes vão sendo atendidos por ordem de chegada, havendo sempre um período de tempo constante de s segundos entre cada serviço. Durante esse período de tempo podem, no entanto, chegar novos clientes. Suponhamos que o número de chegadas ao n —ésimo período de tempo é uma variável aleatória Z_n cuja distribuição é independente do período em que ocorrem as chegadas, e é dada por

$$P(k \text{ clientes chegarem num intervalo de tempo entre 2 chegadas consecutivas de táxi}) = a_k,$$

onde $a_k \geq 0$, $\forall k \in \mathbb{N}_0$, e $\sum_{k=0}^{+\infty} a_k = 1$.

O estado do sistema no início do $(n+1)$ —ésimo período de tempo entre duas chegadas consecutivas de táxi é definido pelo número de clientes que esperam para serem atendidos, sendo esse número representado por X_n , $n \in \mathbb{N}_0$.

- Prove que $\{X_n, n \in \mathbb{N}_0\}$ é uma C.M. homogênea. Indique os respectivos espaço de estados e matriz de transição.
- Sabendo que no início de um determinado intervalo a fila tem zero clientes, qual a probabilidade de o sistema voltar a atingir este estado (pela primeira vez) ao fim de três chegadas de táxis?

Exercício 2.12. Seja $\{Y_i^{(n)} : i, n \in \mathbb{N}\}$ uma família de variáveis aleatórias independentes e identicamente distribuídas com distribuição binomial $B(2; 0.5)$.

Considere um processo estocástico $\{X_n\}_{n \in \mathbb{N}_0}$ definido por:

$$X_0 = a, \text{ com } a \in \mathbb{N} \text{ fixo, e para todo } n \geq 1 : \quad X_n = \begin{cases} \sum_{i=1}^{X_{n-1}} Y_i^{(n)}, & \text{se } X_{n-1} \geq a. \\ a, & \text{se } X_{n-1} < a \end{cases}$$

- Identifique o espaço de estados do processo X .
- Prove que X é uma C.M. homogênea e identifique a matriz de transição.

Exercício 2.13. Determinado ser vivo produz, durante a sua vida, um número de descendentes Y de acordo com uma distribuição dada por:

$$\forall k \in \mathbb{N}_0, P(Y = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \lambda > 0 \text{ constante.}$$

A cada indivíduo i da população associamos uma v.a. Y_i = número de descendentes do indivíduo i . Para cada $n \in \mathbb{N}_0$, define-se a v.a. X_n que representa o tamanho da população na geração de ordem n . Considere o PE $(X_n : n \in \mathbb{N}_0)$.

- Prove que $(X_n : n \in \mathbb{N}_0)$ é uma C.M. homogênea.
- Calcule a probabilidade de existirem $k \in \mathbb{N}_0$ indivíduos na geração de ordem $n+2$, sabendo que na geração de ordem n existiam $j \in \mathbb{N}_0$.

Exercício 2.14. Considere lançamentos repetidos de um dado honesto. Seja X_n o máximo dos números que ocorreram nos n primeiros lançamentos.

- Indique o espaço de estados da C.M. $(X_n : n \in \mathbb{N})$ e a respetiva matriz de transição.
- Determine $P(X_i = i)$, $i = 1, 2, \dots, 6$ e $P(X_2 = 3)$.
- Desenhe o grafo da cadeia, classifique os seus estados, e analise o tipo de recorrência.
- Verifique que a cadeia é aperiódica.

Exercício 2.15. Seja $(X_n : n \in \mathbb{N})$ uma C.M. com espaço de estados $E = \mathbb{N}_0$ e probabilidades de transição P_{ij} tais que:

$$P_{k0} = \frac{1}{k+2} \quad \text{e} \quad P_{k,k+1} = \frac{k+1}{k+2}.$$

Mostre que a cadeia é irredutível.

Exercício 2.16. Seja $(Z_n : n \in \mathbb{N}_0)$ uma C.M. homogénea com espaço de estados $E = \mathbb{N}_0$ e probabilidades de transição:

$$\mathbb{P} = \begin{bmatrix} 1-a_0 & a_0 & 0 & 0 & \dots \\ 1-a_1 & 0 & a_1 & 0 & \dots \\ 1-a_2 & 0 & 0 & a_2 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}, \quad 0 < a_i < 1, \quad i = 0, 1, 2, \dots$$

- A cadeia dada é irredutível e aperiódica? Justifique.
- Determine a probabilidade f_{00}^n de que a cadeia, partindo do estado 0, volte novamente a esse estado, pela primeira vez, em n passos. De seguida, mostre que:

$$\sum_{n=1}^{M+1} f_{00}^n = 1 - \prod_{i=0}^M a_i.$$

- Atendendo aos resultados das alíneas anteriores, enuncie, em termos dos a_i 's, uma condição necessária e suficiente para que todos os estados sejam recorrentes. Justifique a sua resposta.

Exercício 2.17. Uma dada empresa identificou seis estados associados ao comportamento diário dos seus colaboradores: 0, 1, 2, 3, 4, 5. As transições de estado para estado podem ser modeladas por uma C.M. com matriz de transição:

$$\mathbb{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 & 0 \\ 0.1 & 0 & 0.5 & 0.3 & 0 & 0.1 \\ 0 & 0 & 0 & 0.7 & 0.1 & 0.2 \\ 0.3 & 0 & 0 & 0.3 & 0.4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

- Desenhe o grafo de \mathbb{P} .
- Identifique os estados transitórios e os estados recorrentes.

2.2.1 Decomposição do espaço de estados

Agora pretendemos decompor o espaço de estados de uma cadeia de Markov finita em subclasses. O objetivo é estudar propriedades da cadeia pela análise das propriedades de cada classe separadamente.

Definição 2.14 (Classe fechada). Uma **classe fechada** é uma classe de comunicação C tal que, se $i \in C$ e $P_{ij} > 0$, então $j \in C$.

Note-se que, a cadeia nunca pode sair de C depois de entrar.

Coloca-se agora a questão: como encontrar todos os estados que pertencem a uma mesma classe de comunicação (e, em particular, a uma classe fechada)? Seguem-se os passos:

- **Passo 1:** inclui-se em C todos os estados j para os quais $P_{ij} > 0$, isto é, $i \rightarrow j$;
- **Passo 2:** inclui-se em C todos os estados k para os quais $P_{jk} > 0$, $\forall j \in C$, isto é, $i \rightarrow k$ (propriedade da transitividade);
- **Passo 3:** repetir o Passo 2 até não se poder incluir mais estados em C ;
- **Passo 4:** verificar se algum estado de C tem transições para fora.
 - Se não tiver, C é uma **classe fechada**.
 - Caso contrário, C é apenas uma **classe de comunicação** (não fechada).

Nota: se começarmos num estado **transitório** e aplicarmos o algoritmo, o conjunto construído acabará por conduzir a uma **classe fechada**. Assim, as classes fechadas são precisamente os **menores subconjuntos** do espaço de estados obtidos pela aplicação do algoritmo a todos os estados.

O resultado seguinte permite uma decomposição do espaço de estados de uma cadeia de Markov, chamada **decomposição canónica**.

Teorema 2.11 (Decomposição canónica). *Seja \mathbb{P} a matriz de transição de uma C.M. com espaço de estados E . Então, E pode ser decomposto numa união (finita ou infinita enumerável) da forma:*

$$E = T \cup C_1 \cup C_2 \cup \dots,$$

onde:

- T é o conjunto dos estados transitórios;
- C_1, C_2, \dots são classes de comunicação disjuntas, fechadas e constituídas apenas por estados recorrentes.

Para cada $j \in C_a$, a probabilidade de atingir um estado $k \in E$ num número finito de passos é dada por:

$$\sum_{n=1}^{+\infty} f_{jk}^n = \begin{cases} 1, & \text{se } k \in C_a, \\ 0, & \text{se } k \notin C_a, \end{cases}$$

onde $f_{jk}^n = P(X_n = k, X_1 \neq k, \dots, X_{n-1} \neq k \mid X_0 = j)$ é a probabilidade de atingir k pela primeira vez no instante n , começando em j .

Adicionalmente, reordenando os estados de forma conveniente, a matriz de transição \mathbb{P} pode ser escrita na forma em blocos:

$$\mathbb{P} = \begin{bmatrix} R & S \\ 0 & Q \end{bmatrix},$$

onde:

- R descreve as transições entre estados em T ;
- S descreve as transições de T para as classes C_1, C_2, \dots ;
- Q é uma matriz bloco-diagonal com submatrizes P_1, P_2, \dots , onde cada $P_a = [P_{ij}]_{i,j \in C_a}$ representa as transições internas em cada classe recorrente.

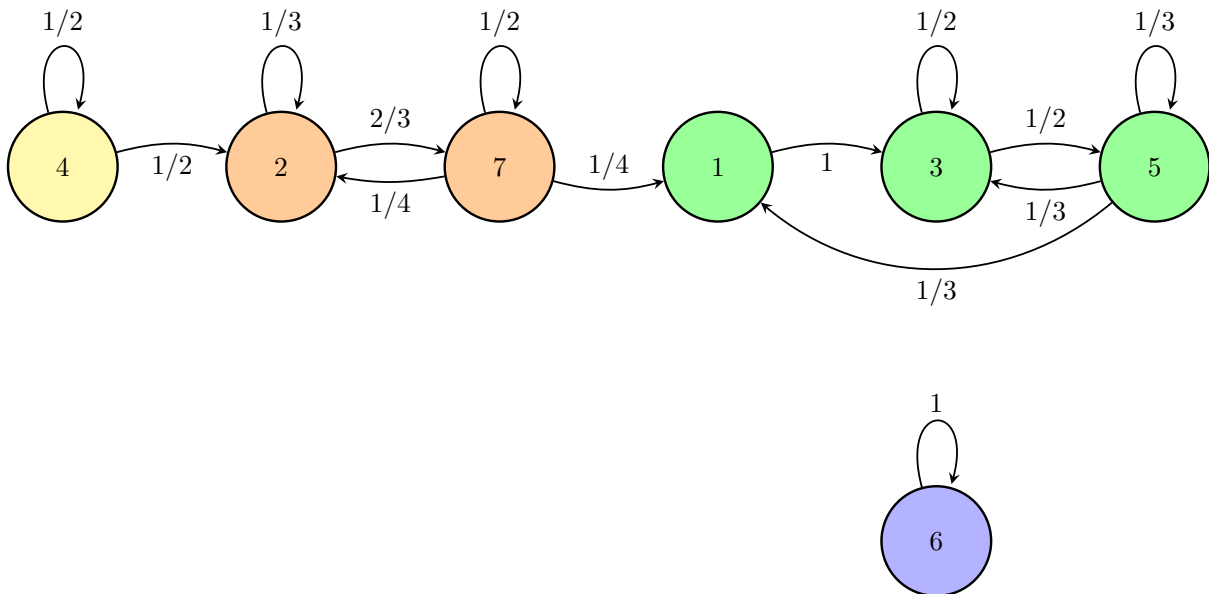
Nota . A decomposição canónica do espaço de estados de uma C.M. é muito útil para analisar as propriedades da cadeia, como a recorrência e a transitoriedade dos estados, bem como a classificação dos estados em classes de comunicação. Em particular, quando o espaço de estados E de uma C.M. é finito, a classificação dos estados pode fazer-se da seguinte forma:

- Passo 1: decompor E em classes de equivalência (classes de comunicação);
- Passo 2: identificar quais dessas classes são fechadas;
- Passo 3:
 - em cada classe fechada, todos os estados são recorrentes positivos (porque a classe é finita e fechada);
 - nas classes não fechadas, todos os estados são transitórios.

Exemplo 2.6. Considere-se uma C.M com espaço de estados $E = \{1, 2, \dots, 7\}$ e com matriz de transição

$$\mathbb{P} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 0 & 0 & 2/3 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1/4 & 1/4 & 0 & 0 & 0 & 0 & 1/2 \end{bmatrix}.$$

- O grafo associado à matriz \mathbb{P} é:



- Classes de equivalência (partição de E com base na relação de comunicação $i \leftrightarrow j$):

$$\{1, 3, 5\}, \{2, 7\}, \{4\}, \{6\}.$$

- Classes fechadas (classes de equivalência sem saída):

$$\{1, 3, 5\}, \{6\}.$$

- Conjuntos fechados (qualquer subconjunto sem transições para fora):

$$\{1, 3, 5\}, \{6\}.$$

- Estados recorrentes positivos (estados em classes fechadas finitas):

$$1, 3, 5, 6.$$

- Estados transitórios (estados que podem sair da sua classe, ou que não são atingidos novamente):

$$2, 4, 7.$$

- Decomposição (separação entre transitórios e classes recorrentes):

$$E = T \cup C_1 \cup C_2 = \{2, 4, 7\} \cup \{1, 3, 5\} \cup \{6\}.$$

- A matriz de transição pode ser reescrita como:

$$\mathbb{P} = \begin{bmatrix} \color{red}{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \color{blue}{0} & \color{blue}{1} & \color{blue}{0} & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & \color{blue}{1/3} & \color{blue}{1/3} & \color{blue}{1/3} & 0 & 0 & 0 \\ \color{green}{0} & 0 & 0 & 0 & \color{red}{1/3} & 0 & \color{red}{2/3} \\ \color{green}{0} & 0 & 0 & 0 & \color{red}{1/2} & \color{red}{1/2} & 0 \\ \color{green}{0} & \color{red}{1/4} & 0 & 0 & \color{red}{1/4} & 0 & \color{red}{3/4} \end{bmatrix} = \begin{bmatrix} \color{red}{P_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \color{blue}{P_2} & \mathbf{0} \\ \color{green}{Q_1} & \color{red}{Q_2} & \color{red}{Q_3} \end{bmatrix},$$

onde P_i representa a matriz de probabilidades de transição da classe C_i e Q_i é a matriz associada a estados de transição.

Nota . Nem todos os conjuntos fechados são classes. Atente-se ao exemplo anterior:

- $\{1, 3, 5, 6\}$ é fechado, mas não é uma classe de equivalência, pois 6 não comunica com os outros.
- todas as classes fechadas são conjuntos fechados, mas o inverso não é verdadeiro.
- classes de equivalência disjuntas partilham zero comunicação entre si (por definição).

2.3 Probabilidades de absorção em estados recorrentes

Um dos cálculos de interesse na teoria das cadeias de Markov está relacionado com o tempo (ou número de transições) necessário, para que, a cadeia partindo de algum estado inicial, **atinja algum estado terminal de interesse**.

Este assunto está muitas vezes associado ao problema da determinação de probabilidades de absorção, com a seguinte formulação.

- Seja $E = T \cup C_1 \cup C_2 \cup \dots$ a decomposição canónica do espaço de estados E , onde T é definido pelos estados transitórios da cadeia, e C_i são classes fechadas e recorrentes.
- Se a cadeia parte de um estado recorrente em C_a , nunca mais deixará C_a (C_a é fechada).
- Se a cadeia parte do estado transitório de T , a cadeia poderá ser absorvida por uma das classes C_a , $a = 1, 2, \dots$.

Nestas circunstâncias estamos interessados nas probabilidades de absorção.

Definição 2.15. Seja $C := \bigcup_a C_a$ a união das classes fechadas de estados recorrentes, e seja T o conjunto de estados transitórios. Seja S a variável aleatória definida por

$$S := \min\{n \geq 1 : X_n \in C\},$$

isto é, o instante da primeira entrada numa classe recorrente, partindo de um estado transitório.

Dado $X_0 = i$, com $i \in T$, o valor

$$a_{ij} := P(X_S = j \mid X_0 = i), \quad j \in C,$$

representa a probabilidade de a cadeia, partindo de i , ser absorvida no estado recorrente j .

Esta quantidade é chamada **probabilidade de absorção** do estado transitório i no estado recorrente j .

Teorema 2.12 (Probabilidades de absorção). *Seja E um espaço de estados finito, $T \subset E$ o conjunto de estados transitórios, e $C = \bigcup_a C_a$ a união das classes fechadas de estados recorrentes.*

*Para cada par $(i, j) \in T \times C$, a **probabilidade de absorção** no estado j , partindo de i , é dada por*

$$a_{ij} = P_{ij} + \sum_{k \in T} P_{ik} a_{kj},$$

onde P_{ik} são as probabilidades de transição de 1 passo da matriz \mathbb{P} .

Interpretação: a cadeia pode ser absorvida no estado $j \in C$ de forma imediata (via P_{ij}) ou, caso isso não ocorra, pode passar para um estado intermédio $k \in T$, a partir do qual será absorvida em j .

Exemplo 2.7. Num estudo no Reino Unido, após a Segunda Guerra Mundial, sobre a mobilidade social entre gerações foram identificados 3 níveis: 1 - superior, 2 - médio e 3 - inferior. Foram estimadas as probabilidades condicionais de um filho pertencer a uma classe social (nível superior, médio, ou inferior) mediante o nível social dos pais ser superior, médio ou inferior. Os resultados são apresentados na tabela seguinte:

Pai	Filho		
	Superior	Médio	Inferior
Superior	0.45	0.48	0.07
Médio	0.05	0.70	0.25
Inferior	0.01	0.50	0.49

Admitamos que as transições entre classes de gerações sucessivas é uma família que pode ser considerada como transições de uma cadeia de Markov.

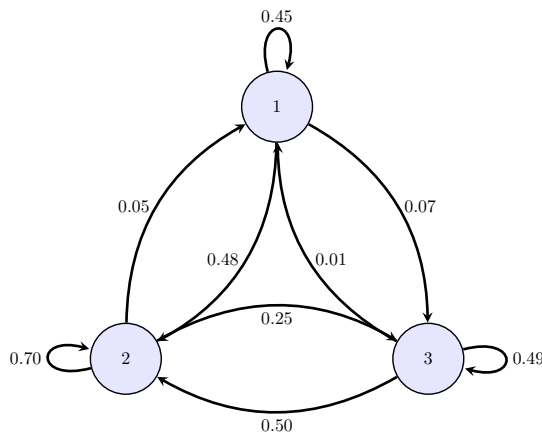
1. Qual a probabilidade de um neto de uma família com nível médio seja o primeiro descendente a ser considerado com um nível social superior, isto é, qual o valor de f_{21}^2 ?
2. Qual a probabilidade de que, em alguma geração, de uma família com nível social inferior, seja atingida pela primeira vez o nível superior?

Solução

A matriz de transição é dada por

$$\mathbb{P} = \begin{bmatrix} 0.45 & 0.48 & 0.07 \\ 0.05 & 0.70 & 0.25 \\ 0.01 & 0.50 & 0.49 \end{bmatrix}, \quad E = \{1, 2, 3\} = \{2, 3\} \cup \{1\} = T \cup C.$$

O grafo é



1. Pretende-se determinar a probabilidade de, começando no estado 2, o primeiro momento em que a cadeia entra no estado 1 seja no instante 2. Para tal, iremos usar a probabilidade de primeira passagem do estado 2 para o estado 1 em 2 passos, isto é,

$$\begin{aligned}
 f_{21}^2 &= P(X_2 = 1, X_1 \neq 1 \mid X_0 = 2) \\
 &= P(X_2 = 1, X_1 = 2 \mid X_0 = 2) + P(X_2 = 1, X_1 = 3 \mid X_0 = 2) \\
 &= \sum_{k \neq 1} P_{2k} P_{k1} \\
 &= P_{22} P_{21} + P_{23} P_{31} \\
 &= 0.70 \times 0.05 + 0.25 \times 0.01 \\
 &= 0.0375.
 \end{aligned}$$

2. Nesta questão, estamos interessados no cálculo do tempo necessário para que a cadeia, partindo de um estado inicial ("nível social inferior"), atinja um estado recorrente ("nível superior"). Podemos, portanto, utilizar as probabilidades de absorção e aplicar o Teorema 2.12. Assim,

$$a_{31} = P_{31} + \sum_{k \in T} P_{3k} a_{k1} = P_{31} + P_{32} a_{21} + P_{33} a_{31}.$$

Uma vez que não sabemos o valor de a_{21} , podemos resolver o sistema:

$$\begin{cases} a_{31} = P_{31} + P_{32} a_{21} + P_{33} a_{31} \\ a_{21} = P_{21} + P_{22} a_{21} + P_{23} a_{31} \end{cases},$$

donde se obtém $a_{31} = 1$ (e $a_{21} = 1$), ou seja, com probabilidade 1, a cadeia partindo do estado 3 (nível social inferior) atingirá o estado 1 (nível superior) em alguma geração futura.

Exercício 2.18. Os negócios do José flutuam em anos sucessivos entre 3 estados: 0 (bancarrota), 1 (perto da bancarrota) e 2 (solvência). A matriz de transição que indica a probabilidade de passagem de um estado para outro é:

$$\mathbb{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.25 \end{bmatrix}.$$

- (a) Qual a probabilidade dos negócios do José conduzirem a uma bancarrota sabendo que ele começou no estado de solvência?
- (b) A mãe do José considera ser mau para o nome da família permitir que os negócios do seu filho vão à bancarrota. Assim, quando o estado 0 é atingido, a mãe do José dá-lhe dinheiro efetivo de modo a que os negócios do José passem ao estado de solvência com probabilidade 1. A matriz de transição desta nova cadeia de Markov é dada por:

$$\mathbb{P} = \begin{bmatrix} 0 & 0 & 1 \\ 0.5 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.25 \end{bmatrix}.$$

A nova cadeia de Markov é irredutível e aperiódica? Sabendo que os negócios do José estão a correr bem (estado 2), qual a probabilidade da mãe do José ter necessidade de dar novamente dinheiro ao filho apenas daqui a 3 anos?

2.4 Teoremas limite

2.4.1 Distribuição estacionária e distribuição limite

Seja $(X_n : n \in \mathbb{N}_0)$ uma C. M. definida num espaço de estados E , com matriz de transição \mathbb{P} e distribuição inicial $P(X_0 = i)$, $i \in E$. Existem duas questões pertinentes:

- Qual o comportamento de X_n após um “longo” número de transições?
- Poderá a cadeia atingir um “comportamento estável” após um “longo” número de transições?

Em geral, a sucessão de v.a.'s $(X_n : n \in \mathbb{N}_0)$ não converge para um estado específico, já que a cadeia goza da propriedade da flutuação aleatória inerente ao próprio processo definida pela matriz de transições. No entanto, pode acontecer que a distribuição de X_n estabilize de algum modo após um elevado número de transições.

Definição 2.16 (Distribuição limite). Se existir

$$\lim_{n \rightarrow +\infty} P(X_n = j), \quad j \in E,$$

e esse limite, usualmente denotado por π_j , definir uma distribuição de probabilidade, então o vetor

$$\tilde{\pi} = (\pi_j, \quad j \in E)$$

designa-se por **distribuição limite** da cadeia.

O estudo da existência de uma distribuição limite da cadeia, nomeadamente de uma distribuição de probabilidade $(\pi_j, \quad j \in E)$, tal que

$$\forall j \in E, \quad \pi_j = \lim_{n \rightarrow +\infty} P(X_n = j)$$

está relacionado com a existência da chamada distribuição estacionária.

Definição 2.17 (Distribuição estacionária). Diz-se que a C.M. possui **distribuição estacionária** se existir $\tilde{\pi} = (\pi_j, \quad j \in E)$ tal que:

(i) $\tilde{\pi}$ define uma distribuição de probabilidade $(\pi_j > 0, \quad \forall j \in E \text{ e } \sum_{j \in E} \pi_j = 1)$;

(ii) $\tilde{\pi} = \tilde{\pi} \mathbb{P}$, isto é,

$$\forall j \in E : \quad \pi_j = \sum_{i \in E} \pi_i P_{ij}.$$

Nota .

1. $\tilde{\pi} = [\pi_0 \quad \pi_1 \quad \pi_2 \quad \dots]$ se $E = \mathbb{N}_0$.
2. A razão pela qual uma distribuição de probabilidade discreta $\tilde{\pi} = (\pi_j, \quad j \in E)$ que satisfaz (ii) é chamada de distribuição estacionária, tem a ver com o facto de

$$\begin{aligned} \tilde{\pi} = \tilde{\pi} \mathbb{P} &\implies \tilde{\pi} = (\tilde{\pi} \mathbb{P}) \mathbb{P} = \tilde{\pi} \mathbb{P}^2 \\ &= (\tilde{\pi} \mathbb{P}) \mathbb{P}^2 = \tilde{\pi} \mathbb{P}^3 \\ &= \dots \\ &= \tilde{\pi} \mathbb{P}^n, \quad \forall n \in \mathbb{N}. \end{aligned}$$

3. Por outro lado, pelas propriedades da distribuição das C. M., sabemos que a distribuição de X_n é dada por:

$$P(X_n = j) = \sum_{i \in E} P(X_0 = i) P_{ij}^n, \quad \forall n \in \mathbb{N}_0.$$

Assim, comparando as relações (2) e (3), se a distribuição inicial for a distribuição estacionária da cadeia, isto é, se

$$P(X_0 = i) = \pi_i, \quad i \in E,$$

então a distribuição de probabilidade de X_n também é a distribuição estacionária da cadeia, ou seja,

$$P(X_n = j) = \pi_j, \quad j \in E, \quad n \in \mathbb{N}_0.$$

Consequentemente, nestas condições, teríamos:

- a distribuição de X_n é estacionária com a passagem do tempo (isto é, não se altera com o número de transições n);
- os vetores aleatórios

$$(X_{n_1}, \dots, X_{n_k}) \text{ e } (X_{n_1+m}, \dots, X_{n_k+m})$$

têm a mesma distribuição, isto é, a cadeia é um processo estritamente estacionário.

Do exposto, podemos observar o seguinte Teorema:

Teorema 2.13. *Se a C.M. tem distribuição estacionária $\tilde{\pi} = (\pi_j, j \in E)$ e a distribuição inicial da cadeia for também $\tilde{\pi}$, isto é,*

$$P(X_0 = i) = \pi_i, \quad \forall i \in E,$$

então existe distribuição limite da cadeia e é dada por

$$\lim_{n \rightarrow +\infty} P(X_n = j) = \pi_j, \quad j \in E.$$

Nota . Para n suficientemente elevado, se os X_n tiverem a mesma distribuição de probabilidade, e se tomarmos essa distribuição como a distribuição inicial, então a C.M. terá um comportamento estável após um elevado número de transições dos estados da cadeia.

Em que condições a C.M. tem distribuição estacionária? Atente-se ao seguinte Teorema:

Teorema 2.14. *Uma C.M. irredutível tem uma distribuição estacionária*

$$\tilde{\pi} = (\pi_j, j \in E),$$

se os estados são todos recorrentes positivos. Neste caso, $\tilde{\pi}$ é a única distribuição estacionária e é dada por

$$\pi_j = \frac{1}{\mu_j}, \quad j \in E,$$

onde μ_j é o tempo médio de recorrência do estado j .

Nota . Este Teorema fornece um método para testar se uma C.M. irredutível é recorrente positiva. Para tal basta procurar a distribuição estacionária. Como se obtém esta distribuição?

$$\tilde{\pi} \text{ satisfaz: } \begin{cases} \tilde{\pi} = \tilde{\pi} \mathbb{P} \\ \sum_{j \in E} \pi_j = 1. \end{cases}$$

Se existir $\tilde{\pi}_j$ então a C.M. é recorrente positiva e $\pi_j = 1/\mu_j$.

Corolário 2.1. *Uma C.M. irredutível, recorrente positiva, e com distribuição inicial estacionária $\tilde{\pi} = (\pi_j, j \in E)$, tem distribuição limite dada por*

$$\pi_j = \lim_{n \rightarrow +\infty} P(X_n = j).$$

2.4.2 Comportamento limite de P_{ij}^n quando $n \rightarrow +\infty$

Qual a relação entre a existência de uma distribuição estacionária e o comportamento limite das probabilidades de transição a n passos, quando $n \rightarrow +\infty$?

Teorema 2.15. Se existir uma distribuição de probabilidade $\tilde{\pi} = (\pi_j, j \in E)$, tal que

$$\forall i, j \in E : \pi_j = \lim_{n \rightarrow +\infty} P_{ij}^n,$$

então,

- (i) $\lim_{n \rightarrow +\infty} P(X_n = j) = \pi_j, \forall j \in E$, ou seja, π é a distribuição limite da cadeia, independente da distribuição inicial;
- (ii) $\tilde{\pi}$ é estacionária: $\tilde{\pi} = \tilde{\pi}P$.

Nota . Pode acontecer que uma C.M. tenha uma única distribuição estacionária $\tilde{\pi}$, mas não ser a distribuição limite da cadeia, isto é, a cadeia pode não ser **estável**. No entanto, se esta distribuição estacionária única for tal que

$$\pi_j = \lim_{n \rightarrow +\infty} P_{ij}^n, \quad j \in E$$

então, necessariamente $\tilde{\pi}$ é a distribuição limite da C.M. Assim, podemos concluir que se existir o $\lim_{n \rightarrow +\infty} P_{ij}^n$, e se este definir uma distribuição de probabilidade, independentemente do estado inicial i , então essa distribuição definirá a distribuição limite da cadeia. Neste sentido, o estudo do comportamento limite da cadeia **não dependerá** da distribuição inicial $P(X_0 = i), i \in E$, mas apenas do comportamento limite das probabilidades de transição a n passos, quando $n \rightarrow +\infty$, ou seja, de

$$\lim_{n \rightarrow +\infty} P_{ij}^n, \quad \forall i, j \in E.$$

Mas existirá sempre este limite? Não! E, quando existe, como se calcula? Em geral, os cálculos para a obtenção de P^n são morosos e difíceis, principalmente para valores de n elevados.

Para contornar esta dificuldade, analisa-se o comportamento das probabilidades P_{ij}^n quando $n \rightarrow +\infty$. Caso estas sejam fáceis de obter, poderão ser utilizadas como aproximações das probabilidades de transição a um número finito de n passos, para n suficientemente elevado. Atente-se ao seguinte Teorema:

Teorema 2.16. Se uma C.M. é irreduzível, aperiódica, e recorrente, então

$$\exists \lim_{n \rightarrow +\infty} P_{ij}^n \quad \text{e} \quad \lim_{n \rightarrow +\infty} P_{ij}^n = \lim_{n \rightarrow +\infty} P_{jj}^n = \frac{1}{\mu_j}, \quad \forall i, j \in E.$$

Teorema 2.17. Em cadeias de Markov irreduzíveis e aperiódicas, todos os seus estados são recorrentes positivos, isto é, a cadeia é **ergódica**:

- (i) sse o sistema de equações lineares

$$\begin{cases} \pi_j = \sum_{i \in E} \pi_i P_{ij}, & j \in E, \\ \sum_{j \in E} \pi_j = 1 \end{cases},$$

tiver solução

$$\pi = [\pi_0 \quad \pi_1 \quad \dots], \quad \text{com } E = \mathbb{N}_0.$$

- (ii) se tiver solução, será única, estritamente positiva e,

$$\pi_j = \lim_{n \rightarrow +\infty} P_{ij}^n, \quad \forall i, j \in E.$$

Nota .

1. O Teorema anterior apresenta uma condição necessária e suficiente para que cadeia irreduzível e aperiódica seja ergódica;

2. A solução do sistema, quando existir, é simultaneamente distribuição estacionária e distribuição limite.
3. Uma vez que

$$\lim_{n \rightarrow +\infty} P_{ij}^n = \pi_j,$$

pode-se aproximar a matriz \mathbb{P}^n à matriz $[\pi_j]_{j \in E}$ para valores elevados de n .

Nota . Numa C.M. irredutível com um número finito de estados não pode haver nem estados transitórios, nem recorrentes nulos, pois em qualquer dos casos

$$\lim_{n \rightarrow +\infty} P_{ij}^n = 0.$$

Assim, numa C.M. irredutível finita, como todos os estados intercomunicam, eles serão todos recorrentes positivos. Atente-se ao Corolário seguinte.

Corolário 2.2. *Numa C.M. irredutível, aperiódica e finita, o sistema de equações lineares*

$$\begin{cases} \pi_i = \sum_{j \in E} \pi_j P_{ji}, \\ \sum_{j \in E} \pi_j = 1 \end{cases},$$

tem solução única, estritamente positiva e $\pi_j = \lim_{n \rightarrow +\infty} P_{ij}^n$.

Exercício 2.19. Num dado centro comercial existem 4 restaurantes A, B, C e D . Desde que a Ana trabalha numa das lojas do centro, ela almoça regularmente num dos 4 restaurantes. Sabe-se ainda que a escolha diária do restaurante está de acordo uma C.M. homogénea com matriz de transição:

$$\mathbb{P} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0 & 0.1 & 0.6 \\ 0.4 & 0 & 0.3 & 0.3 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Sabe-se que no primeiro dia de trabalho qualquer um dos 4 restaurantes tinha igual probabilidade de ser selecionado pela Ana.

- (a) Defina distribuição inicial de uma C.M. e indique a distribuição inicial da cadeia dada.
- (b) Calcule a probabilidade de no segundo dia de trabalho a Ana selecione o restaurante B para almoçar.
- (c) Classifique quanto à recorrência todos os estados da cadeia.
- (d) Sabendo que a Ana almoçou hoje no restaurante B , qual a probabilidade da Ana nos próximos dias optar pelo restaurante A antes de selecionar o restaurante D ?
- (e) Qual o número médio de dias entre dois almoços no restaurante B ?

Exercício 2.20. Considere uma C.M. sobre o espaço um espaço de estados $E = \{1, 2, 3, 4, 5, 6\}$ e com matriz de transição:

$$\mathbb{P} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0.25 & 0.75 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0 & 0 \\ 0.25 & 0 & 0.25 & 0.25 & 0 & 0.25 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix}.$$

- (a) Classifique os estados da cadeia identificando as classes fechadas e a periodicidade de todos os estados da cadeia.

- (b) Determine a probabilidade de que a cadeia, partindo do estado 1, regresse novamente a 1, pela primeira vez, em n passos.
- (c) Determine o número médio de transições necessárias para que a cadeia, partindo de 1 volte a 1, isto é, o tempo médio de recorrência do estado 1.
- (d) Encontre a distribuição estacionária relativa à C.M. restrita ao sub-espaço $\{1, 2\} \subset E$. A partir desta, determine o tempo médio de recorrência para o estado 1.

Exercício 2.21. Relativamente ao funcionamento de uma máquina analisa-se a durabilidade, em número de dias completos, de um certo tipo de peça. Para tal, considere-se que sempre que a peça falha a máquina pára, procedendo-se à substituição da peça por outra idêntica, de modo que no dia seguinte a máquina retoma o seu funcionamento com a nova peça.

Seja Z_{n+1} o tempo de vida (contado em dias completos) da peça instalada no n -ésimo dia, e denote por p_k a probabilidade de que uma peça nova dure k dias completos, com $k = 0, 1, 2, \dots$

Represente por X_n o tempo de vida (contado em dias completos) que resta à peça que está em uso no n -ésimo dia de observação do processo.

- (a) Prove que o processo $(X_n : n \in \mathbb{N}_0)$ é uma C.M. homogénea sobre o espaço \mathbb{N}_0 e com matriz de transição:

$$\mathbb{P} = \begin{bmatrix} p_0 & p_1 & p_2 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}.$$

- (b) A cadeia $(X_n : n \in \mathbb{N}_0)$ é irredutível e aperiódica? Justifique.
- (c) Defina distribuição estacionária de uma C.M. e mostre que a cadeia dada possui distribuição estacionária sse o tempo médio de vida das peças novas é finito, isto é,

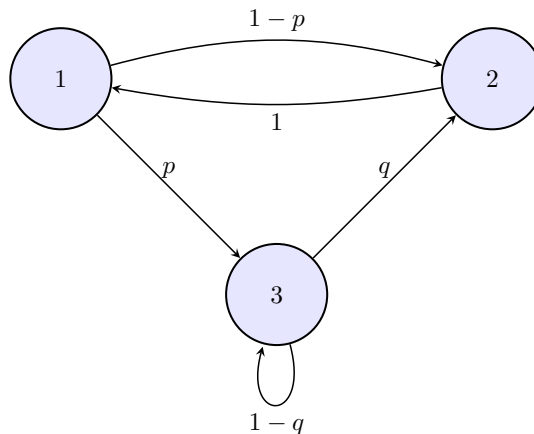
$$\sum_{k=0}^{+\infty} kp_k < +\infty.$$

Note que $p_0 + p_1 + p_2 + \dots = 1$.

- (d) Sob que condição a cadeia dada é ergódica? Justifique.

Exercício 2.22.

Considere uma cadeia de Markov homogénea definida pelo seguinte grafo:



- (a) Determine a matriz das probabilidades de transição.

- (b) Em que condições esta cadeia é irredutível e aperiódica?
- (c) Determine a distribuição estacionária $\pi = (\pi_1, \pi_2, \pi_3)$.
- (d) Calcule os valores de p e q tais que $\pi_1 = \pi_2 = \pi_3$.

Exercício 2.23.

Considere uma C.M. com espaço de estados $E = \{1, 2, 3, 4\}$ e matriz de transição

$$\mathbb{P} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1/3 & 2/3 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix}$$

Classifique os estados.

Exercício 2.24.

Considere uma C.M. homogênea $\{X_n, n \geq 0\}$, com espaço de estados $E = \{1, 2, 3\}$ e matriz de probabilidades de transição a um passo:

$$\mathbb{P} = \begin{bmatrix} p & 1-p & 0 \\ 0 & 0 & 1 \\ 0 & 1-q & q \end{bmatrix}, \quad 0 < p < 1 \quad \text{e} \quad 0 < q < 1.$$

- (a) Classifique, justificando, cada um dos estados da cadeia.
- (b) Calcule o limite

$$\lim_{n \rightarrow \infty} P(X_n = 3).$$

Exercício 2.25.

Uma urna contém 6 bolas, das quais 3 são encarnadas e 3 são verdes. São selecionadas ao acaso da urna 2 bolas simultaneamente. Se uma for verde e a outra for encarnada, então são postas de lado e são colocadas duas bolas azuis na urna. Se não for o caso, colocam-se de volta as bolas retiradas na urna. O processo repete-se até só haver bolas azuis na urna.

Seja X_n o número de bolas encarnadas na urna depois da tiragem n .

- (a) Justifique que $\{X_n\}$ é uma C.M homogênea. Defina o espaço de estados e construa a respectiva matriz das probabilidades de transição.
- (b) Classifique, justificando, os estados da cadeia.
- (c) Calcule a probabilidade de, a determinada altura, a urna apenas conter bolas azuis partindo de $X_0 = 3$.

Exercício 2.26.

Considere uma C.M com estados 0 e 1, e matriz de transição

$$\mathbb{P} = \begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix}, \quad 0 < a, b < 1.$$

- (a) Calcule a probabilidade do primeiro retorno ao estado 1 em n passos, f_{11}^n , $n = 1, 2, \dots$, e verifique que o estado 1 é recorrente positivo.
- (b) Calcule a distribuição limite (π_0, π_1) e discuta a sua existência. Relacione π_1 com a média do tempo de primeiro retorno ao estado 1.

Exercício 2.27.

Considere a C.M. $(X_n, n = 0, 1, \dots)$ com espaço de estados $E = \{0, 1\}$ e tal que,

$$0 < p_{00} < 1 \quad \text{e} \quad 0 < p_{11} < 1.$$

- (a) Prove que a cadeia é recorrente positiva.
- (b) Determine a distribuição estacionária da cadeia.

Exercício 2.28.

Considere uma C.M. com espaço de estados $E = \{1, 2, 3, 4\}$ e matriz de transição

$$\mathbb{P} = \begin{bmatrix} 1/3 & 2/3 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 0 & 1/4 & 1/2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- (a) Classifique, justificando, os estados da cadeia.
- (b) Calcule $f_{34}(n)$, a probabilidade de a primeira visita ao estado 4 ter lugar no n -ésimo passo, partindo de 3, e calcule a probabilidade de absorção no estado 4, partindo de 3.

Exercício 2.29.

Considere uma C.M. definida pela matriz das probabilidades de transição:

$$\mathbb{P} = \begin{bmatrix} 0 & 2/3 & 1/3 \\ 3/8 & 1/8 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

- (a) Verifique que a cadeia é irredutível e aperiódica.
- (b) Verifique a existência de distribuição limite e determine-a.

Exercício 2.30.

Considere uma C.M. $(X_n, n = 0, 1, \dots)$ com espaço de estados $E = \{1, 2\}$, em que $P_{12} = P_{21} = 1$ e $P(X_0 = 1) = P(X_0 = 2) = 1/2$.

- (a) O que pode concluir quanto à convergência de P_{ii}^n , $i = 1, 2$, quando $n \rightarrow +\infty$? Justifique.
- (b) Mostre que

$$P(X_n = 1) = P(X_n = 2) = 1/2, \quad \forall n = 1, 2, \dots,$$

isto é, a distribuição de X_n é estacionária. Comente e relacione justificadamente com a conclusão obtida na alínea anterior.

Capítulo 3

Cadeias de Markov em tempo contínuo

Neste capítulo iremos considerar $(X_t, t \in \mathbb{R}_0^+)$ uma C.M. com valores em \mathbb{N}_0 e espaço de parâmetro \mathbb{R}_0^+ .

Vamos admitir que $(X_t, t \in \mathbb{R}_0^+)$ é homogênea, isto é, tem probabilidade de transição estacionárias. Nestas condições, a função de probabilidade de transição

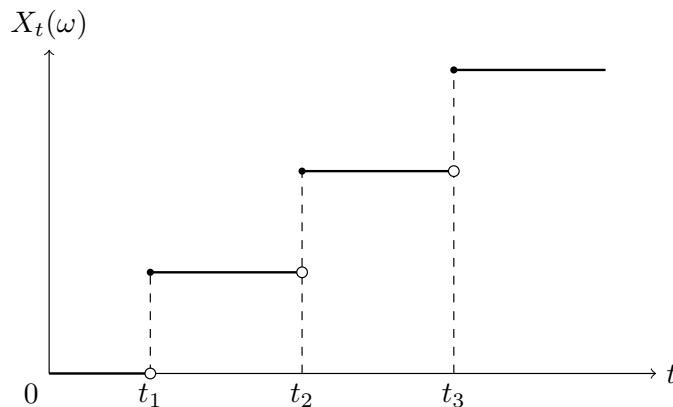
$$\forall t > 0, P_{ij}(t) = P(X_{t+n} = j \mid X_n = i), \quad i, j \in \mathbb{N}_0$$

é independente de $n \geq 0$.

3.1 Processo de Poisson homogêneo

O processo de Poisson homogêneo é um processo estocástico que modela a ocorrência de eventos aleatórios ao longo do tempo, onde os eventos ocorrem de forma independente e com uma taxa constante. É frequentemente utilizado para modelar fenômenos como chamadas telefônicas recebidas num call center, chegadas de clientes a um serviço, ou falhas em sistemas, entre outros.

Seja X_t uma função que conta o número de vezes que um determinado acontecimento ocorre durante o período de tempo de 0 a t . Assim, a aplicação $t \rightarrow X_t$ é uma função em escada, não decrescente, em que os saltos correspondem às ocorrências dos acontecimentos:



Hipótese 3.1 (Postulados do processo de Poisson).

- P1. O número de acontecimentos que ocorrem em intervalos de tempo disjuntos são v.a.'s independentes.
- P2. A v.a. $X_{t_0+t} - X_{t_0}$ (isto é, o acréscimo) depende apenas de t e não de t_0 ou de X_{t_0} .
- P3. A probabilidade de ocorrer pelo menos um acontecimento num intervalo de tempo pequeno de amplitude h é proporcional à amplitude desse intervalo. Assim,

$$P(h) = \lambda h + o(h), \quad h \rightarrow 0, \lambda > 0,$$

onde $g(t) = o(t)$, $t \rightarrow 0$, $\Leftrightarrow \lim_{t \rightarrow 0} \frac{g(t)}{t} = 0$.

- P4. A probabilidade de ocorrer mais do que um acontecimento num intervalo de tempo pequeno de amplitude h é negligível quando h é pequeno. Isto mostra que num intervalo pequeno, ou ocorre um acontecimento, ou não ocorre nenhum, excluindo a possibilidade de ocorrência simultânea de dois ou mais acontecimentos. Assim,

$$\sum_{i=2}^{+\infty} P_i(k) = o(h), \quad h \rightarrow 0,$$

onde $P_i(k) = P(X_k = i)$.

Representemos por $P_n(t)$ a probabilidade de ocorrerem n acontecimentos no intervalo de tempo $[0, t]$. Assim, num intervalo de amplitude $t + h$ temos:

- Para $n = 0$:

$$\begin{aligned} P_0(t+h) &= P_0(t) \cdot P_0(h) \\ &= P_0(t) \cdot (1 - P(\text{ocorrer mais do que 1 acontecimento})) \\ &= P_0(t) \cdot (1 - P(h)) \\ &= P_0(t) - P_0(t) \cdot P(h). \end{aligned}$$

Temos então que:

$$P_0(t+h) - P_0(t) = -P_0(t) \cdot P(h) \Leftrightarrow \frac{P_0(t+h) - P_0(t)}{h} = -P_0(t) \cdot \frac{P(h)}{h}.$$

Aplicando limites, e tendo em conta o Postulado P3, obtemos:

$$\lim_{h \rightarrow 0} \frac{P_0(t+h) - P_0(t)}{h} = -P_0(t) \cdot \lim_{h \rightarrow 0} \frac{P(h)}{h},$$

o que resulta em

$$\frac{d}{dt} P_0(t) = -P_0(t) \cdot \lim_{h \rightarrow 0} \frac{\lambda h + o(h)}{h},$$

ou seja,

$$\frac{d}{dt} P_0(t) = -\lambda P_0(t),$$

isto é, a probabilidade do acontecimento não se realizar no intervalo de tempo $[0, t]$, $P_0(t)$, satisfaz a equação diferencial

$$\boxed{P_0'(t) = -\lambda P_0(t).}$$

multiplicando pelo fator integrante $e^{\lambda t}$, a solução desta equação diferencial é,

$$P_0(t) = K \cdot e^{-\lambda t},$$

onde K é uma constante de integração. Como $P_0(0) = 1$, temos que $K = 1$. Assim, a solução da equação diferencial é

$$\boxed{P_0(t) = e^{-\lambda t}.}$$

- Para $n \geq 1$: se no intervalo $[0, t]$ ocorrerem n eventos, no intervalo $[t, t+h]$ ocorrem zero; se no intervalo $[0, t]$ ocorrerem $n-1$ eventos, no intervalo $[t, t+h]$ ocorre 1; se no intervalo $[0, t]$ ocorrerem $n-2$ eventos, no intervalo $[t, t+h]$ ocorrem 2; e assim sucessivamente. Logo,

$$\begin{aligned}
P_n(t+h) &= P_n(t) \cdot P_0(h) + P_{n-1}(t) \cdot P_1(h) + P_{n-2}(t) \cdot P_2(h) + \dots \\
&= P_n(t) \cdot P_0(h) + P_{n-1}(t) \cdot P_1(h) + \sum_{i \geq 2} P_{n-i}(t) \cdot P_i(h) \\
&= P_n(t) \cdot (1 - P(h)) + P_{n-1}(t) \cdot (P(h) + o(h)) + \sum_{i \geq 2} P_{n-i}(t) \cdot P_i(h),
\end{aligned}$$

donde se obtém:

$$P_n(t+h) - P_n(t) = -P_n(t) \cdot P(h) + P_{n-1}(t) \cdot P(h) + P_{n-1}(t) \cdot o(h) + o(h) + \sum_{i \geq 2} P_{n-i}(t) \cdot P_i(h).$$

Dividindo por h e aplicando o limite quando $h \rightarrow 0$, obtemos:

$$\begin{aligned}
\lim_{h \rightarrow 0} \frac{P_n(t+h) - P_n(t)}{h} &= -P_n(t) \cdot \lim_{h \rightarrow 0} \frac{P(h)}{h} + P_{n-1}(t) \cdot \lim_{h \rightarrow 0} \frac{P(h)}{h} + P_{n-1}(t) \cdot \lim_{h \rightarrow 0} \frac{o(h)}{h} \\
&\quad + \lim_{h \rightarrow 0} \frac{o(h)}{h} + \lim_{h \rightarrow 0} \frac{\sum_{i \geq 2} P_{n-i}(t) \cdot P_i(h)}{h},
\end{aligned}$$

ou seja,

$$\frac{d}{dt} P_n(t) = -P_n(t) \cdot \lambda + P_{n-1}(t) \cdot \lambda + 0 + 0,$$

o que equivale a escrever que a probabilidade do acontecimento se realizar pelo menos uma vez no intervalo $[0, t]$ satisfaz a equação diferencial

$$P'_n(t) = \lambda P_{n-1}(t) - \lambda P_n(t), \quad n \in \mathbb{N}.$$

A solução desta equação diferencial, tendo em conta que $P_n(0) = 0$, é dada por

$$P_n(t) = \frac{\lambda^n t^n}{n!} e^{-\lambda t}, \quad n \in \mathbb{N}.$$

Assim, podemos concluir que a probabilidade de ocorrerem n acontecimentos no intervalo de tempo $[0, t]$ segue uma distribuição de Poisson com parâmetro λt , ou seja,

$$X_t \sim Po(\lambda t),$$

donde

$$P(X_t = n) = P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad n \in \mathbb{N}_0.$$

Das propriedades da distribuição de Poisson, sabemos que $E(X_t) = \lambda t$, o que significa que o número esperado de acontecimentos num intervalo de amplitude t é proporcional à amplitude do intervalo.

No caso $t = 1$, temos que $E(X_1) = \lambda$, pelo que:

- λ representa o número médio de acontecimentos que ocorrem por unidade de tempo;
- λ designa a taxa de ocorrência, razão ou intensidade do processo de Poisson homogêneo.

Com base no exposto, podemos definir processo de Poisson homogêneo do seguinte modo:

Definição 3.1 (Processo de Poisson homogêneo). Um processo estocástico $(X_t, t \in \mathbb{R}_0^+)$ é um **processo de Poisson homogêneo** com taxa $\lambda > 0$ sse:

- (i) tem incrementos independentes e estacionários, e $X_0 = 0$ q.c. (quase certamente);
- (ii) X_t segue uma distribuição de Poisson com parâmetro λt , isto é, $X_t \sim Po(\lambda t)$, para todo $t \in \mathbb{R}_0^+$.

Nota (Observações sobre o processo de Poisson homogéneo).

1. Os postulados do processo de Poisson são suficientes para defini-lo e permitem também definir completamente a distribuição do vetor $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$. Com efeito, por um lado verifica-se

$$\forall s, t \in \mathbb{R}_0^+ : 0 \leq s \leq t, \quad X_t - X_s \sim Po(\lambda(t-s)),$$

isto é, o número de acontecimentos que ocorrem num certo intervalo só depende da amplitude desse intervalo:

$$\begin{aligned} X_t - X_s &\stackrel{d}{=} X_{t-s+s} - X_s \\ &\stackrel{d}{=} X_{t-s} - X_0 \\ &\stackrel{d}{=} X_{t-s} \\ &= Po(\lambda(t-s)). \end{aligned}$$

Por outro lado, como os incrementos são independentes, temos que, para $t_1 < t_2 < \dots < t_k$:

$$\begin{aligned} P(X_{t_1} = n_1, \dots, X_{t_k} = n_k) &= P(X_{t_1} = n_1, \dots, X_{t_k} - X_{t_{k-1}} = n_k - n_{k-1}) \\ &= P(X_{t_1} = n_1) \cdots P(X_{t_k} - X_{t_{k-1}} = n_k - n_{k-1}) \\ &= \frac{(\lambda t_1)^{n_1}}{n_1!} e^{-\lambda t_1} \cdots \frac{(\lambda(t_k - t_{k-1}))^{n_k - n_{k-1}}}{(n_k - n_{k-1})!} e^{-\lambda(t_k - t_{k-1})}. \end{aligned}$$

2. Ligada à v.a. X_t costuma definir-se ainda a v.a. L que representa o tempo de espera até ao primeiro acontecimento, isto é, $L = \inf\{t \in \mathbb{R}_0^+ : X_t > 0\}$. A v.a. L segue uma distribuição exponencial com parâmetro λ , ou seja, $L \sim Exp(\lambda)$. Vejamos:

- $F_L(t) = P(L \leq t) = 1 - P(L > t) = 1 - P(X_t = 0) = 1 - P_0(t) = 1 - e^{-\lambda t}$.
- $f_L(t) = \lambda e^{-\lambda t}, \quad t > 0$.
- $E(L) = 1/\lambda$.
- $Var(L) = 1/\lambda^2$.
- $M_L(t) = \frac{\lambda}{\lambda - t}, \quad t < \lambda$. (f.g.m.).

3. Considere-se a seguinte v.a.:

$$S_r = \text{tempo de espera até à ocorrência do } r\text{-ésimo acontecimento,}$$

isto é,

$$S_r = X_1 + X_2 + \dots + X_r, \quad X_i \text{ i.i.d.,}$$

onde X_i representa o tempo de espera desde o $(i-1)$ -ésimo acontecimento até ocorrer o acontecimento de ordem i . Assim, $X_i \sim Exp(\lambda)$, $\forall i = 1, \dots, r$. Assim, calculando a f.g.m. de S_r , e relacionando com a f.g.m. de X_i vem:

$$M_{S_r}(t) = \prod_{i=1}^r M_{X_i}(t) = (M_{X_i}(t))^r = \left(\frac{\lambda}{\lambda - t} \right)^r, \quad t < \lambda,$$

ou seja, trata-se da f.g.m. de uma v.a. com distribuição Gamma com parâmetros r e λ , ou seja, $S_r \sim \Gamma(r, \lambda)$, donde se obtém,

$$E(S_r) = r/\lambda \quad e \quad Var(S_r) = r/\lambda^2.$$

4. O processo de Poisson é uma cadeia de Markov em tempo contínuo homogénea, isto é, $\forall k_1, k_2, \dots, k_n, k \in \mathbb{N}_0, \forall t_1, t_2, \dots, t_n \in \mathbb{R}_0^+, t_1 \leq t_2 \leq \dots \leq t_n \leq t$:

$$P(X_t = k \mid X_{t_1} = k_1, \dots, X_{t_n} = k_n) = P(X_t = k \mid X_{t_n} = k_n) = P_{k_n k}(t - t_n).$$

Pelo exposto, podemos dar uma outra definição de processo de Poisson:

Definição 3.2 (Processo de Poisson homogéneo). Se $(X_t, t \in \mathbb{R}_0^+)$ é um **processo de Poisson homogéneo** com taxa $\lambda > 0$, então $(X_t, t \in \mathbb{R}_0^+)$ é uma cadeia de Markov com valores em \mathbb{N}_0 , tal que:

- (i) $P(X_{t+h} - X_t = 1 \mid X_t = x) = \lambda h + o(h)$, $x \in \mathbb{N}_0$, quando $h \rightarrow 0$.
- (ii) $P(X_{t+h} - X_t = 0 \mid X_t = x) = 1 - \lambda h + o(h)$, $x \in \mathbb{N}_0$, quando $h \rightarrow 0$.
- (iii) $X_0 = 0$ q.c.

Existem cadeias de Markov mais gerais e que nos permitem descrever fenómenos análogos aos descritos pelos processos de Poisson. É o que veremos nas secções seguintes.

Exercício 3.1.

Seja $X = (X_t : t \geq 0)$ um PE real tal que $X_0 = 0$ q.c. e $X_t \sim Po(\cdot)$.

- (a) Em que condições será X um processo de Poisson?
- (b) Supondo que X é um processo de Poisson, prove que $\forall t, s, h \in \mathbb{R}_0^+$, com $t > s > h$, e $\forall x, y \in \mathbb{N}_0$:

$$P(X_t - X_s = x, X_s - X_h = y) = \frac{e^{-\lambda(t-h)} \lambda^{x+y} (t-s)^x (s-h)^y}{x!y!}.$$

Exercício 3.2.

Considere uma estação de serviço de lavagem de automóveis na qual apenas um carro é atendido de cada vez e segundo a ordem de chegada. Um estudo realizado pela empresa permitiu concluir que as chegadas dos automóveis ocorrem segundo um processo de Poisson com intensidade média de 15 carros por hora. Designe por N_t , $t \geq 0$, o número de automóveis que chegam num intervalo de tempo de amplitude t minutos.

- (a) Identifique a distribuição de N_t , $\forall t \geq 0$. Justifique a sua resposta.
- (b) Mostre que

$$\lim_{h \rightarrow 0} \frac{P(N_h \geq 2)}{P(N_h = 1)} = 0.$$

- c) Prove que a condição expressa na alínea anterior é equivalente a

$$\lim_{h \rightarrow 0} P(N_h > 1 \mid N_h \geq 1) = 0.$$

O que pode concluir sobre o processo em causa?

- d) Qual o tempo médio de espera entre duas chegadas consecutivas?

Exercício 3.3.

Seja $(N_t, t \geq 0)$ um processo de Poisson de intensidade $\lambda > 0$.

- (a) Supondo que $s < t$, calcule:
 - i. $E(N_t - N_s)$.
 - ii. $Var(N_t - N_s)$.
 - iii. $Cov(N_t, N_s)$.
- (b) Os Clientes de um vendedor de jornais chegam segundo um processo de Poisson a uma velocidade média de 2 Clientes por minuto.

- i. Determine a probabilidade de não chegarem Clientes nos próximos três minutos sabendo que chegaram um ou mais Clientes nos últimos cinco minutos.
- ii. O vendedor costuma fazer a seguinte aposta: paga ao seu assistente um euro se o próximo Cliente não chegar dentro de um minuto, caso contrário o assistente paga-lhe um euro. Qual o valor que o vendedor espera ganhar?

Exercício 3.4.

O volume de vendas de um determinado produto constitui um processo de Poisson, com um volume médio de vendas de 4 unidades por dia.

- (a) Qual é a probabilidade de que, em dois dias, se vendam exatamente 6 unidades?
- (b) Qual é a probabilidade de que, em dois dias, se vendam mais de 6 unidades?
- (c) Determine o volume médio de vendas semanal.
- (d) Qual é a probabilidade de que um stock de 4 unidades dure menos de um dia?

Exercício 3.5.

Numa loja os clientes chegam de acordo com uma lei de Poisson à média de 30 por hora. Qual a probabilidade de que o intervalo de tempo entre chegadas sucessivas seja:

- (a) Superior a 2 minutos?
- (b) Inferior a 4 minutos?
- (c) Entre 1 e 3 minutos?

Exercício 3.6.

Uma v.a. T diz-se *sem memória* sse:

$$P(T > x + y \mid T > x) = P(T > y), \quad \forall x, y > 0.$$

Mostre que:

- (a) Se T for uma v.a. contínua, T é sem memória sse T for distribuída exponencialmente.
- (b) Se T tomar apenas valores inteiros e positivos, T é sem memória para x e y não negativos sse existe uma constante p tal que:

$$P(T = k) = p(1 - p)^{k-1}, \quad k = 1, 2, 3, \dots$$

Exercício 3.7.

A chegada de passageiros a uma paragem de autocarro segue um processo de Poisson com intensidade λ . Suponha que um autocarro partiu no instante $t = 0$, não tendo deixado nenhum passageiro em espera. Seja T o tempo de chegada do autocarro seguinte. Então, o número de pessoas na paragem aquando da sua chegada é $N(T)$. Suponha que o tempo de chegada T é independente do processo de Poisson e que T tem distribuição uniforme no intervalo $(1, 2)$.

Calcule a média e a variância de $N(T)$.

Exercício 3.8.

Sejam $(N_t, t \geq 0)$ um processo de Poisson com intensidade λ e $P_k(t) = P(N_t = k)$, $k = 0, 1, 2, \dots$

(a) Deduza as equações diferenciais:

$$\begin{aligned} P'_0(t) &= -\lambda P_0(t) \\ P'_k(t) &= -\lambda P_k(t) + \lambda P_{k-1}(t), \quad k = 1, 2, \dots \end{aligned}$$

(b) Encontre a partir das equações acima a função de probabilidade:

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \dots$$

3.2 Processo de nascimento puro

Considere-se uma sucessão de números positivos $\{\lambda_k, k \in \mathbb{N}_0\}$.

Definição 3.3 (Processo de nascimento puro). Um processo estocástico $(X_t, t \in \mathbb{R}_0^+)$, com valores em \mathbb{N}_0 , é um **processo de nascimento puro** com taxa (ou razão de nascimento) $\{\lambda_k, k \in \mathbb{N}_0\}$ se é uma cadeia de markov em tempo contínuo homogénea, satisfazendo os axiomas:

- (i) $P(X_{t+h} - X_t = 1 \mid X_t = k) = \lambda_k h + o_{1,k}(h) = P_{k,k+1}(h)$.
- (ii) $P(X_{t+h} - X_t = 0 \mid X_t = k) = 1 - \lambda_k h + o_{2,k}(h) = P_{k,k}(h)$.
- (iii) $P(X_{t+h} - X_t < 0 \mid X_t = k) = 0, k \in \mathbb{N}_0$, quando $h \rightarrow 0$.
- (iv) $X_0 = 0$ q.c.

Nota .

1. A condição (iv) é admita por conveniência.
2. X_t **não** representa o tamanho da população mas o número de nascimentos no intervalo $[0, t]$.
3. Uma vez que as probabilidades de transição dadas por (i) e (ii) são estacionárias, então $o_{1,k}(h)$ e $o_{2,k}(h)$ não dependem de t .
4. O processo de nascimento puro é uma generalização do processo de Poisson homogéneo, onde a probabilidade de um acontecimento ocorrer num certo instante depende do número de acontecimentos que já ocorreram. Assim, o processo de Poisson é um processo de nascimento puro de razão de nascimento constante e igual λ .

Considere-se agora

$$P_n(t) = P(X_t = n),$$

e atente-se ao Teorema seguinte:

Teorema 3.1. $P_n(t)$ satisfaz, para $t \geq 0$, o sistema de equações diferenciais

$$\begin{cases} P'_0(t) = -\lambda_0 P_0(t), \\ P'_n(t) = \lambda_{n-1} P_{n-1}(t) - \lambda_n P_n(t), \quad n \geq 1 \end{cases},$$

com as condições fronteira

$$\begin{cases} P_0(0) = P(X_0 = 0) = 1, \\ P_n(0) = P(X_0 = n) = 0, \quad n > 0. \end{cases}$$

Adicionalmente,

$$P_0(t) = P(X_t = 0) = e^{-\lambda_0 t}.$$

Consideremos agora a v.a.

$T_k =$ tempo compreendido entre o instante k e o $k + 1$ —ésimo nascimentos consecutivos,

isto é

$T_k =$ tempo de espera entre nascimentos consecutivos.

Tem-se que

$$P_n(t) = P(X_t = n) = P\left(\sum_{i=0}^{n-1} T_i \leq t \leq \sum_{i=0}^n T_i\right).$$

Considere-se

$$S_k = \sum_{i=0}^{k-1} T_i,$$

o tempo durante o qual ocorrem k nascimentos. Como se viu anteriormente,

$$\forall z > 0, P(T_0 \leq z) = 1 - P(X_z = 0) = 1 - e^{-\lambda_0 z},$$

isto é,

$$T_0 \sim \text{Exp}(\lambda_0).$$

É possível provar (ver Karlin & Taylor, por exemplo) que $(T_k, k \in \mathbb{N}_0)$ é uma sucessão de v.a.'s independentes, tais que, para cada $k \in \mathbb{N}_0$, $T_k \sim \text{Exp}(\lambda_k)$.

Adicionalmente, se $(X_t, t \geq 0)$ é um processo de Poisson, então S_n segue uma distribuição Gamma com parâmetros n e λ , ou seja, $S_n \sim \Gamma(n, \lambda)$, onde λ é a taxa do processo de Poisson.

Terminamos esta secção com o seguinte Teorema:

Teorema 3.2. $P_k(t)$ verifica a equação de recorrência

$$P_k(t) = \lambda_{k-1} e^{-\lambda_k t} \int_0^t e^{\lambda_k x} P_{k-1}(x) dx, \quad k \in \mathbb{N}.$$

Exercício 3.9.

Com vista ao bom funcionamento de determinado consultório médico, a direcção determinou que em cada instante, do período de funcionamento do mesmo, não poderia existir no consultório mais do que N doentes. Apenas um doente é atendido de cada vez e segundo a respetiva ordem de chegada. Os doentes chegam ao consultório segundo um processo de Poisson de intensidade $1/2$, ficando a aguardar a sua vez de atendimento apenas se nesse momento o número de utentes no consultório for inferior a N . As consultas são concluídas segundo um processo de Poisson de intensidade $1/3$.

(a) Designe por N_t , $t \geq 0$, o número de doentes que chegam num intervalo de amplitude t .

1. Prove que:

- i. $(N_t, t \geq 0)$ é uma C.M. de tempo contínuo homogénea e indica a respetiva probabilidade de transição.
- ii. $(N_t, t \geq 0)$ é um processo de nascimento puro.

2. Sendo T uma v.a. que representa o tempo de espera entre duas chegadas consecutivas, prove que

$$P(T > t) = e^{-t/2}, \quad t > 0.$$

3. Qual o tempo médio de espera entre chegadas?

(b) Seja agora X_t , $t \geq 0$, o número total de doentes no consultório no instante t . Supondo que:

$$\forall k \in \{0, 1, 2, \dots, N\} : P(X_t = k) = \left(\frac{3}{2}\right)^k P(X_t = 0),$$

determine:

1. a probabilidade de que existam k doentes à espera de serem atendidos, num qualquer instante t .
2. o número médio de doentes no consultório, num qualquer instante t .

Exercício 3.10.

Considere um quiosque no qual os Clientes chegam de acordo com um processo de Poisson à razão de 32 Clientes por dia, durante o horário diário de abertura do quiosque (o qual corresponde a 8 horas). Designe por N_t , $t \geq 0$, o número de Clientes que chegam ao quiosque num intervalo de tempo de amplitude t horas.

- a) Identifique, justificando, a distribuição de N_t .
- b) Sendo T_2 a v.a. que representa o instante de chegada (em horas) do segundo Cliente ao quiosque, em cada dia, mostre que:

$$P(T_2 > t) = e^{-4t}(1 + 4t), \quad t > 0.$$

Exercício 3.11.

Uma população de organismos evolui da forma seguinte: cada organismo existe independentemente dos outros, e vive durante determinado tempo, aleatório, com distribuição exponencial de parâmetro θ , dividindo-se então em dois novos organismos. Por sua vez, a sua existência é também independente dos outros organismos e têm um tempo de vida exponencialmente distribuído de parâmetro θ , e assim sucessivamente.

Seja $X(t)$ o número de organismos existentes no instante t . Suponha que $X(0) = 1$ e defina $P_n(t) = P(X(t) = n)$. Justifique que $X(t)$ é um processo de nascimento puro, ou seja, verifica

$$P'_n(t) = -\theta(n P_n(t) - (n-1) P_{n-1}(t)), \quad n = 1, 2, \dots$$

Exercício 3.12.

Considere uma população de dimensão $N(t)$ no instante t tal que $N(0) = 1$. Admita que qualquer dos membros desta população se divide em dois novos membros no intervalo $[t, t+h]$ com probabilidade $\lambda h + o(h)$ ou mantém-se inalterado neste intervalo com probabilidade $1 - \lambda h + o(h)$.

- (a) Prove que $(N(t), t \geq 0)$ é um processo de nascimento puro com taxa de natalidade $\lambda_n = n\lambda$, para todo o $n = 1, 2, \dots$.
- (b) Designe por $p_k(t) = P(N(t) = k)$, com $k = 1, 2, \dots$ e prove que:

$$p'_k(t) = (k-1)\lambda p_{k-1}(t) - k\lambda p_k(t), \quad k = 1, 2, \dots$$

- (c) Tendo em conta a equação diferencial anterior, conclua por indução que:

$$p_k(t) = e^{-k\lambda t} (e^{\lambda t} - 1)^{k-1}, \quad k = 1, 2, \dots$$

- (d) Seja $P(z, t) = \sum_{k=1}^{\infty} z^k p_k(t)$ a função geradora das probabilidades $p_k(t)$. Prove que:

$$P(z, t) = \frac{ze^{-\lambda t}}{1 - z + ze^{-\lambda t}}.$$

(e) Calcule $E(N(t))$.

Exercício 3.13.

Seja $(N(t), t \geq 0)$ um processo de nascimento puro com $N(0) = I$ e taxa de natalidade $\lambda_n = n\lambda$, sendo I um inteiro positivo. Designe por $P_n(t) = P(N(t) = n)$.

(a) Prove que:

$$\begin{aligned} P'_I(t) &= -I\lambda P_I(t) \\ P'_n(t) &= -n\lambda P_n(t) + (n-1)\lambda P_{n-1}(t), \quad n = I+1, I+2, \dots \end{aligned}$$

(b) Prove que:

$$P_k(t) = \binom{k-1}{I-1} e^{-I\lambda t} (1 - e^{-\lambda t})^{k-I}, \quad k \geq I.$$

3.3 Processo de nascimento e morte

3.3.1 Definição e equações de Chapman-Kolmogorov

Definição 3.4 (Processo de nascimento e morte). Um processo estocástico $(X_t, t \in \mathbb{R}_0^+)$, com valores em \mathbb{N}_0 , é um **processo de nascimento e morte** com taxas $\{\lambda_k, k \in \mathbb{N}_0\}$ e $\{\mu_k, k \in \mathbb{N}_0\}$ se é uma cadeia de markov em tempo contínuo homogénea, satisfazendo os axiomas:

1. $P_{i,i+1}(h) = P(X_{t+h} = i+1 \mid X_t = i) = \lambda_i h + o(h), \quad i \geq 0,$
2. $P_{i,i-1}(h) = P(X_{t+h} = i-1 \mid X_t = i) = \mu_i h + o(h), \quad i \geq 1,$
3. $P_{i,i}(h) = P(X_{t+h} = i \mid X_t = i) = 1 - (\lambda_i + \mu_i)h + o(h), \quad i \geq 0,$
4. $P_{i,j}(0) = P(X_t = j \mid X_t = i) = \delta_{ij},$
5. $\mu_0 = 0, \lambda_0 > 0, \mu_i, \lambda_i > 0, \quad i \in \mathbb{N},$

onde $o(h)$, em cada caso, pode depender de i e considera-se que $o(h) \rightarrow 0$ quando $h \rightarrow 0$.

Quando ocorre um nascimento o processo passa do estado E_k para o estado E_{k+1} e, quando ocorre uma morte, o processo passa do estado E_k para o estado E_{k-1} . Assim, o processo de nascimento e morte é um processo de nascimento puro com a adição de mortes.

Nota. Uma das generalizações óbvias dos processos de nascimento puro considerados consiste em permitir que X_t decresça, por exemplo através da morte dos seus membros. Assim, se no instante $t = 0$ o processo está no estado n ele poderá mudar-se para os estados vizinhos $n+1$ ou $n-1$ após um tempo de espera aleatório. Um processo de nascimento e morte pode assim ser interpretado como um passeio aleatório de tempo contínuo.

Num processo de nascimento e morte (com espaço de estados \mathbb{N}_0) observa-se, $\forall t \in \mathbb{R}_0^+$ e $\forall i, j \in \mathbb{N}_0$:

1. $P_{ij}(t) \geq 0$, ou seja, as probabilidades de transição são não negativas para quaisquer pares (i, j) e qualquer tempo t . Note-se que, em particular, $P_{ij}(0) = \delta_{ij}$, onde δ_{ij} é o símbolo de Kronecker, i.e., $\delta_{ij} = 1$ se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$;
2. $\sum_{j=0}^{\infty} P_{ij}(t) = 1$, ou seja, a soma das probabilidades de transição a partir do estado i para todos os estados possíveis é igual a 1, para qualquer tempo t ;

3. As probabilidades de transição são dadas por: $\forall s, t, \in \mathbb{R}_0^+$,

$$\begin{aligned}
 P_{ij}(t+s) &= P(X_{n+t+s} = j \mid X_n = i) \\
 &= \sum_{k=0}^{+\infty} P(X_{n+t+s} = j, X_{n+t} = k \mid X_n = i) \\
 &= \sum_{k=0}^{+\infty} P(X_{n+t+s} = j \mid X_{n+t} = k, X_n = i) \cdot P(X_{n+t} = k \mid X_n = i) \\
 &= \sum_{k=0}^{+\infty} P(X_{n+t+s} = j \mid X_{n+t} = k) \cdot P(X_{n+t} = k \mid X_n = i) \\
 &= \sum_{k=0}^{+\infty} P_{kj}(s) \cdot P_{ik}(t).
 \end{aligned}$$

Temos então as **Equações de Chapman-Kolmogorov**:

$$P_{ij}(t+s) = \sum_{k=0}^{+\infty} P_{kj}(s) P_{ik}(t). \quad (3.1)$$

Nesta dedução, em cada igualdade, utilizamos, respetivamente, as seguintes justificações:

1. Definição da probabilidade de transição do estado i para o estado j em $t+s$ passos, a partir do instante n .
2. Aplicação da regra da probabilidade total, ao considerar todos os possíveis estados k em que a cadeia pode estar no instante intermédio $n+t$.
3. Aplicação da regra do produto das probabilidades condicionadas.
4. Aplicação da propriedade de Markov, que garante que o futuro (a partir de $n+t$) depende apenas do estado atual $X_{n+t} = k$, e não do passado $X_n = i$.
5. Finalmente, identificamos as probabilidades de transição:

$$P_{kj}(s) = P(X_{n+t+s} = j \mid X_{n+t} = k) \text{ e } P_{ik}(t) = P(X_{n+t} = k \mid X_n = i).$$

As probabilidades de transição e as leis marginais caracterizam a distribuição do processo. As distribuições marginais dependem apenas da distribuição inicial e das probabilidades de transição. Assim, se

$$q_i := P(X_0 = i), \quad i \in \mathbb{N}_0,$$

tem-se, $\forall n \in \mathbb{N}_0$,

$$\begin{aligned}
 P(X_t = n) &= \sum_{i=0}^{+\infty} P(X_t = n, X_0 = i) \\
 &= \sum_{i=0}^{+\infty} P(X_0 = i) \cdot P(X_t = n \mid X_0 = i) \\
 &= \sum_{i=0}^{+\infty} P_{in}(t) \cdot q_i.
 \end{aligned}$$

Assim, as distribuições marginais do processo de nascimento e morte são dadas por

$$P(X_t = n) = \sum_{i=0}^{+\infty} P_{in}(t) q_i, \quad n \in \mathbb{N}_0.$$

3.3.2 Tempo de espera

Considere-se agora a v.a.

T_i = tempo de espera de X_t no estado i .

É possível mostrar (Karlin & Taylor, por exemplo) que, quando $h \rightarrow 0$,

$$P(T_i \geq t + h) = P(T_i \geq t) \cdot P(T_i \geq h).$$

Podemos ainda escrever a igualdade anterior do seguinte modo

$$\begin{aligned} P(T_i \geq t + h) &= P(T_i \geq t) \cdot P(T_i \geq h) \\ &= P(T_i \geq t) \cdot P(X_{n+h} = i, X_{s+n} = i \mid X_n = i), \quad \forall s \in [0, h] \\ &= P(T_i \geq t) \cdot (P_{ii}(h) + o(h)) \\ &= P(T_i \geq t) \cdot (1 - (\lambda_i + \mu_i)h + o(h)) \\ &= P(T_i \geq t) \cdot (1 - (\lambda_i + \mu_i)) + o(h). \end{aligned}$$

Designando por $G_i(t) = P(T_i \geq t)$, obtemos

$$G_i(t + h) = G_i(t) \cdot (1 - (\lambda_i + \mu_i)h) + o(h),$$

donde, dividindo por h e aplicando o limite quando $h \rightarrow 0$, obtemos

$$\frac{d}{dt}G_i(t) = -(\lambda_i + \mu_i)G_i(t),$$

pelo que,

$$G_i(t) = \exp\{-(\lambda_i + \mu_i)t\} = 1 - (1 - \exp\{-(\lambda_i + \mu_i)t\}),$$

isto é,

$$P(T_i \geq t) = 1 - \exp\{-(\lambda_i + \mu_i)t\}, \quad t \geq 0,$$

ou seja,

$$\boxed{T_i \sim \text{Exp}(\lambda_i + \mu_i),}$$

e o tempo médio de espera é dado por

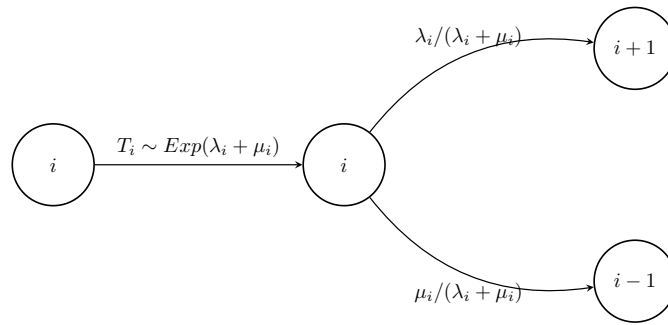
$$\frac{1}{\lambda_i + \mu_i}.$$

Nota . A descrição do movimento de X_t descreve-se a seguir. O processo permanece num certo estado i por um tempo aleatório T_i seguindo uma distribuição exponencial de parâmetro $\lambda_i + \mu_i$. Quando deixa o estado i , o processo ou entra no estado $i + 1$ ou $i - 1$:

- a probabilidade de que a transição seja para $i + 1$ é dada por $\lambda_i / (\lambda_i + \mu_i)$;

- a probabilidade de que a transição seja para $i - 1$ é dada por $\mu_i/(\lambda_i + \mu_i)$.

Esquemáticamente, temos:



Uma possível realização de X_t é:

$$i \longrightarrow i + 1 \longrightarrow i \longrightarrow i - 1 \longrightarrow \dots$$

O movimento é análogo ao do caminho aleatório, com a diferença de que o tempo de espera em cada estado é uma v.a. com distribuição exponencial.

3.3.3 Equações diferenciais de processos de nascimento e morte

As equações de Chapman-Kolmogorov descrevem a evolução das probabilidades de transição ao longo do tempo. A partir destas equações, podemos deduzir as equações diferenciais que descrevem a evolução das probabilidades de transição em função do tempo.

Pela relação (3.1), temos:

$$\begin{aligned} P_{ij}(t+h) &= \sum_{k=0}^{+\infty} P_{ik}(h) \cdot P_{kj}(t) \\ &= P_{ii}(h) \cdot P_{ij}(t) + P_{i,i+1}(h) \cdot P_{i+1,j}(t) + P_{i,i-1}(h) \cdot P_{i-1,j}(t) + \sum_{k \notin \{i-1, i, i+1\}} P_{ik}(h) \cdot P_{kj}(t) \\ &= (1 - (\lambda_i + \mu_i)h) P_{ij}(t) + \lambda_i h \cdot P_{i+1,j}(t) + \mu_i h \cdot P_{i-1,j}(t) + o(h), \end{aligned}$$

assumindo que o processo é de saltos de primeiro vizinho, isto é, que $P_{ik}(h) = o(h)$ para $|k - i| \geq 2$.

Dividindo por h e aplicando o limite quando $h \rightarrow 0$, obtemos:

$$\boxed{P'_{ij}(t) = \lambda_i P_{i+1,j}(t) + \mu_i P_{i-1,j}(t) - (\lambda_i + \mu_i) P_{ij}(t), \quad i \geq 1.}$$

Para $i = 0$, temos:

$$\begin{aligned} P_{0j}(t+h) &= P_{00}(h) \cdot P_{0j}(t) + P_{0,1}(h) \cdot P_{1j}(t) + \sum_{k \geq 2} P_{0k}(h) \cdot P_{kj}(t) \\ &= (1 - \lambda_0 h) P_{0j}(t) + \lambda_0 h \cdot P_{1j}(t) + o(h), \end{aligned}$$

donde se obtém:

$$\boxed{P'_{0j}(t) = \lambda_0 P_{1j}(t) - \lambda_0 P_{0j}(t).}$$

Estas equações designam-se por **Equações de Kolmogorov de atraso**, e descrevem como a probabilidade de transição evolui em função do **estado inicial** i .

As condições iniciais associadas são:

$$P_{ij}(0) = \delta_{ij}, \quad i, j \in \mathbb{N}_0,$$

onde δ_{ij} representa o delta de Kronecker.

Por outro lado, a equação de Chapman-Kolmogorov (3.1) também pode ser escrita na forma:

$$\begin{aligned} P_{ij}(t+h) &= \sum_{k=0}^{+\infty} P_{ik}(t) \cdot P_{kj}(h) \\ &= P_{ij}(t) \cdot P_{jj}(h) + P_{i,j+1}(t) \cdot P_{j+1,j}(h) + P_{i,j-1}(t) \cdot P_{j-1,j}(h) + \sum_{k \notin \{j-1,j,j+1\}} P_{ik}(t) \cdot P_{kj}(h) \\ &= (1 - (\lambda_j + \mu_j)h) P_{ij}(t) + \lambda_{j-1}h \cdot P_{i,j-1}(t) + \mu_{j+1}h \cdot P_{i,j+1}(t) + o(h), \end{aligned}$$

assumindo, tal como anteriormente, que os saltos ocorrem apenas entre estados vizinhos.

Dividindo por h e aplicando o limite $h \rightarrow 0$, obtemos:

$$P'_{ij}(t) = \lambda_{j-1}P_{i,j-1}(t) + \mu_{j+1}P_{i,j+1}(t) - (\lambda_j + \mu_j)P_{ij}(t), \quad j \geq 1.$$

Para $j = 0$, temos:

$$P'_{i0}(t) = \mu_1P_{i1}(t) - \lambda_0P_{i0}(t).$$

Estas equações designam-se por **Equações de Kolmogorov de avanço**, e descrevem como a probabilidade de transição evolui em função do **estado final** j .

Vejamos agora se o comportamento de $P_{ij}(t)$ estabiliza à medida que $t \rightarrow +\infty$.

É possível demonstrar (ver, por exemplo, Karlin & Taylor) que, sob condições adequadas (recorrência positiva e aperiocidade), as probabilidades de transição convergem para um valor limite, isto é:

$$\lim_{t \rightarrow +\infty} P_{ij}(t) = \pi_j,$$

onde π_j é a probabilidade de que o sistema se encontre no estado j em regime estacionário, independentemente do estado inicial i .

Por outras palavras, π_j representa a probabilidade limite de que o sistema se encontre no estado j , num instante arbitrário no futuro.

Neste caso, as equações de Kolmogorov de atraso (ou avanço) convergem para um sistema de equações algébricas conhecidas como **Equações de Kolmogorov estacionárias**, que descrevem o comportamento assintótico do processo de nascimento e morte:

$$\begin{cases} \lambda_0\pi_1 - \lambda_0\pi_0 = 0, & j = 0, \\ \lambda_{j-1}\pi_{j-1} + \mu_{j+1}\pi_{j+1} - (\lambda_j + \mu_j)\pi_j = 0, & j \geq 1. \end{cases}$$

Estas equações, a par da condição de normalização

$$\sum_{j=0}^{\infty} \pi_j = 1,$$

permitem determinar a distribuição de probabilidade estacionária $(\pi_j)_{j \in \mathbb{N}_0}$.

Exercício 3.14.

Considere um sistema de self-service em que a probabilidade de haver uma chegada em $(t, t+h)$ dado que existem j Clientes a servirem-se no instante t é igual a $ajh + o(h)$, $j \geq 0$, quando $h \rightarrow 0$, onde $a > 0$ é uma constante real positiva. Suponha que os Clientes acabam o seu serviço segundo um processo de Poisson de intensidade $2a$ e que estão reunidas todas as condições para poder modelar o sistema por um processo de nascimento e morte, N_t .

- Identifique um conjunto de axiomas que caracterize o processo $(M_t : t \geq 0)$, onde M_t representa o número de Clientes que acabam de se servirem num intervalo de tempo de amplitude t .
- Identifique, justificando, as taxas de nascimento e morte do processo N_t .
- Faça um diagrama de velocidade de probabilidades de transição para o processo N_t e escreva o correspondente sistema de equações de avanço de Kolmogorov.

Exercício 3.15.

Considere que os autocarros chegam a uma certa rua segundo um processo de Poisson de intensidade de 10 por hora, e que percorrem num intervalo de tempo que é sempre constante e igual a 10 minutos. Suponha que a rua não tem limitação para o número de veículos que nela podem transitar.

- Após associar ao problema um processo de nascimento e morte, determine a distribuição de equilíbrio e interprete o significado de π_0 .
- Determine o número médio de autocarros na rua depois de terem decorrido algumas horas desde o início da carreira.
- O número de autocarros tende a aumentar ou a diminuir com a passagem do tempo? Justifique.

Exercício 3.16.

Seja $(X(t), t \geq 0)$ um processo de nascimento e morte tal que:

$$\begin{aligned} \lambda_n &= \lambda q^n & 0 < q < 1, \lambda > 0 & \quad n = 0, 1, 2, \dots \\ \mu_n &= \mu & \mu > 0 & \quad n = 1, 2, \dots \\ \mu_0 &= 0 \end{aligned}$$

Designe por $P_n(t) = P(X(t) = n)$. Prove que

$$\begin{aligned} P'_0(t) &= -\lambda P_0(t) + \mu P_1(t) \\ P'_n(t) &= \lambda q^{n-1} P_{n-1}(t) - (\lambda q^n + \mu) P_n(t) + \mu P_{n+1}(t), \quad n \geq 1. \end{aligned}$$

Exercício 3.17.

Considere o processo estocástico $N(t)$, que representa o número de linhas ocupadas numa central telefónica, a qual dispõe de um número elevado de linhas. Este processo é modelado por um sistema com chegadas espontâneas de chamadas e término aleatório de chamadas, com os seguintes pressupostos:

- As chamadas chegam à central a uma taxa constante λ , independentemente do número de linhas atualmente ocupadas.
- Cada chamada em curso termina a uma taxa μ , de forma independente das restantes. Assim, quando há k chamadas em curso (ou k linhas ocupadas), a taxa total de término é $k\mu$.

- (a) Mostre que as Equações de Kolmogorov de avanço associadas às probabilidades $P_i(t) = P(N(t) = i)$ são dadas por:

$$P'_i(t) = -(\lambda + i\mu) P_i(t) + \lambda P_{i-1}(t) + (i+1)\mu P_{i+1}(t), \quad i = 0, 1, 2, \dots$$

- (b) Suponha que, para cada $i \in \mathbb{N}_0$, a função $P_i(t)$ é dada por:

$$P_i(t) = \frac{(\lambda/\mu)^i}{i!} (1 - e^{-\mu t})^i \exp \left\{ -\frac{\lambda}{\mu} (1 - e^{-\mu t}) \right\}, \quad i = 0, 1, 2, \dots$$

Determine a probabilidade de todas as linhas estarem desocupadas no regime estacionário (isto é, quando $t \rightarrow +\infty$), e deduza a forma da distribuição estacionária do número de linhas ocupadas.

Exercício 3.18.

Seja Y_n , $n = 0, 1, \dots$, uma cadeia de Markov com espaço de estados $E = \{0, 1\}$ e matriz de transição \mathbb{P} dada por:

$$\mathbb{P} = \begin{bmatrix} 0 & 1 \\ 1 - \alpha & \alpha \end{bmatrix}.$$

Considere ainda um processo de Poisson com parâmetro λ , isto é, $(N(t), t \geq 0)$. Mostre que o processo definido por $X(t) = Y_{N(t)}$, para $t \geq 0$, é um processo de nascimento e morte com dois estados, e determine os parâmetros λ_0 e μ_1 em função de α e λ .

Capítulo 4

Complementos de processos estocásticos

4.1 Processo de Wiener

Definição 4.1 (Filtração). Seja $X = (X(t), t \in T)$ um processo estocástico definido no espaço de probabilidade (Ω, \mathcal{F}, P) , com conjunto de índices $T = [0, +\infty[$. Uma família de sub- σ -álgebras de \mathcal{F} , tal que para $s \leq t$ se tenha $\mathcal{F}_s \subset \mathcal{F}_t$, designa-se por **filtração**.

Denomina-se **filtração natural** do processo X a família

$$(\mathcal{F}_t = \sigma(X_s : 0 \leq s \leq t), t \in T),$$

formada pelas álgebras- σ geradas pelo processo X até ao instante t .

Um processo estocástico $X = (X(t), t \in T)$ está **adaptado** à filtração $(\mathcal{F}_t, t \in T)$ se, para todo $t \in T$, a variável aleatória $X(t)$ é \mathcal{F}_t -mensurável, isto é, as imagens inversas dos conjuntos $B \in \mathcal{B}$ estão contidas em \mathcal{F}_t .

Em 1828, o botânico inglês Robert Brown observou pequenas partículas de pólen imersas num líquido a movimentarem-se de forma completamente aleatória. Mais tarde, em 1905, Albert Einstein justificou este movimento com a constante colisão entre as partículas e as moléculas do líquido envolvente e caracterizou-o por um processo estocástico que viria a ser chamado processo de Wiener. Finalmente, em 1918, apareceu a primeira definição matemática do termo através do matemático Norbert Wiener.

Definição 4.2 (Processo de Wiener padrão (ou movimento Browniano)). Um **processo de Wiener padrão** (ou **movimento Browniano**) é um processo estocástico $W = (W_t)_{t \geq 0}$ definido num espaço de probabilidade (Ω, \mathcal{F}, P) , que satisfaz as seguintes propriedades:

1. **Condição inicial:** $W_0 = 0$ quase certamente, isto é,

$$P(W_0 = 0) = 1;$$

2. **Incrementos gaussianos:** Para quaisquer instantes $0 \leq s < t < \infty$, a variável aleatória $W_t - W_s$ é normalmente distribuída com média zero e variância $t - s$, ou seja,

$$W_t - W_s \sim \mathcal{N}(0, t - s);$$

3. **Incrementos independentes:** Para todo $n \in \mathbb{N}$ e qualquer sequência crescente de instantes $0 \leq t_0 < t_1 < \dots < t_n$, os incrementos

$$W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$$

são variáveis aleatórias independentes;

4. **Trajetoárias contínuas:** Com probabilidade 1, a aplicação $t \mapsto W_t(\omega)$ é contínua para todo $\omega \in \Omega$, ou seja,

$$P(W \in C([0, \infty[)) = 1,$$

onde $C([0, \infty[)$ denota o espaço das funções contínuas em $[0, \infty[$.

Nos processos estocásticos e, em particular, nas equações diferenciais estocásticas, o processo de Wiener representa o efeito acumulado das perturbações aleatórias na evolução de determinado fenómeno em estudo. Dada a importância deste processo, iremos apresentar algumas das suas propriedades.

Definição 4.3. Considere-se uma função $f : [0, t] \rightarrow \mathbb{R}$ e uma sequência de partições $\mathcal{P}_n = \{t_0^n, t_1^n, \dots, t_n^n\}$ do intervalo $[0, t]$, com $0 = t_0^n < t_1^n < \dots < t_n^n = t$ para cada $n \in \mathbb{N}$, tais que

$$\delta_n = \max_{0 \leq i \leq n-1} |t_{i+1}^n - t_i^n| \rightarrow 0 \quad \text{quando } n \rightarrow +\infty.$$

- A **variação** da função f no intervalo $[0, t]$ é definida por

$$V_f([0, t]) = V_f(t) := \lim_{n \rightarrow +\infty} \sum_{i=0}^{n-1} |f(t_{i+1}^n) - f(t_i^n)|.$$

- Diz-se que f tem **variação finita** no intervalo $[0, t]$ se $V_f(t) < +\infty$.
- Diz-se que f tem **variação limitada** no intervalo $[0, t]$ se

$$\sup_{u \in [0, t]} V_f(u) < k, \quad \text{para algum } k > 0.$$

- Diz-se que f tem **variação quadrática** no intervalo $[0, t]$ se existir e for finito o limite

$$V_f^2(t) := \lim_{n \rightarrow +\infty} \sum_{i=0}^{n-1} |f(t_{i+1}^n) - f(t_i^n)|^2.$$

Propriedade 4.1 (Propriedades do processo de Wiener). *O processo de Wiener, W_t , possui as seguintes propriedades:*

1. *Existe uma versão separável e contínua do processo, isto é, com trajectórias quase certamente contínuas;*
2. *Para todo $t \geq 0$, $W_t \sim \mathcal{N}(0, t)$;*
3. *A função de covariância é dada por $\text{Cov}[W_s, W_t] = E[W_s W_t] = s \wedge t$;*
4. *W_t é um processo de Markov homogéneo;*
5. *A distribuição condicional de $W_{s+\tau}$ dado $W_s = x$ é Normal com média x e variância τ ;*
6. *W_t é uma martingala em relação à sua filtração natural;*
7. *As trajectórias do processo de Wiener são, quase certamente, não diferenciáveis;*
8. *As trajectórias do processo de Wiener são, quase certamente, de variação ilimitada em qualquer intervalo;*
9. *Possui variação quadrática finita no intervalo $[a, b]$, igual a $b - a$.*

Nota (Ruído branco como derivada generalizada do processo de Wiener). Embora as trajectórias do processo de Wiener sejam, quase certamente, contínuas mas não diferenciáveis (propriedade 7), e tenham variação total infinita (propriedade 8), é possível interpretar a sua derivada no **sentido das distribuições generalizadas** (ou distribuições de Schwartz).

Neste contexto, define-se a **derivada generalizada** do processo de Wiener como

$$\frac{dW_t}{dt} = \xi_t,$$

onde ξ_t representa um **processo estocástico generalizado**, designado por **ruído branco** (ou *white noise*). Este processo não é uma função no sentido clássico, mas sim uma distribuição (ou funcional) que actua sobre funções teste suaves.

O ruído branco ξ_t caracteriza-se pelas seguintes propriedades formais:

- É um processo com média nula: $E(\xi_t) = 0$;
- A sua função de autocovariância é dada por:

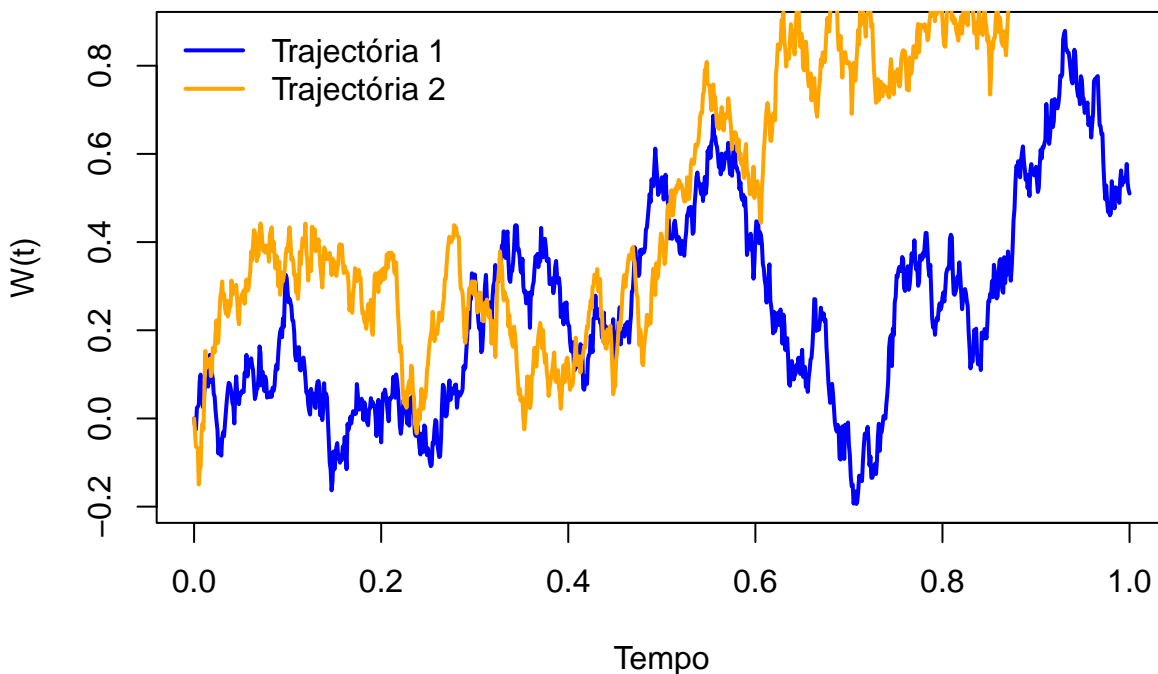
$$E(\xi_s \xi_t) = \delta(t - s),$$

onde δ é a função delta de Dirac, interpretada como distribuição.

Este formalismo é essencial na formulação de equações diferenciais estocásticas (EDEs), nas quais o ruído branco representa uma força aleatória infinitesimalmente perturbadora que actua continuamente ao longo do tempo.

Na imagem seguinte apresentam-se duas trajectórias de um processo de Wiener. As trajectórias foram obtidas por simulação numérica, considerando incrementos independentes e normalmente distribuídos com média zero e variância proporcional ao incremento temporal.

Exemplo de duas trajectórias do processo de Wiener



Exercício 4.1 (Braumann (2005)). Tirando partido das propriedades do processo de Wiener, calcule ou determine:

1. $P(W(2.7) > 1.5)$.
2. $P(-1.5 < W(2.7) < 1.5)$.

3. $P(W(2.7) < 1.5 \mid W(1.8) = 1)$.
4. $P(-1.5 < W(2.7) < 1.5 \mid W(1.8) = 1)$.
5. $E(W(t) \mid W(s), W(u))$ com $0 < u < s < t$.
6. $Var(W(t) \mid W(s), W(u))$ com $0 < u < s < t$.
7. $P(W(2.7) > 1.5 \mid W(1.8) = 1, W(0.5) = -2)$.
8. $E(W(2.7) \mid W(1.8) = 1, W(0.5) = -2)$.
9. $P(W(1.8) < 1 \mid W(2.7) = 1.5)$.
10. $P(W(1.8) = 1 \mid W(2.7) < 1.5)$.
11. $P(W(2.7) = 1.5, W(1.8) > 1)$.
12. $P(W(2.7) < 1.5, W(1.8) = 1)$.
13. $P(-1 < W(2.7) - W(1.8) < 1.4 \wedge 0.5 < W(1.6) - W(0.9) < 1.5)$.
14. $P(-1 < W(2.7) - W(1.8) < 1.4 \mid W(1.6) - W(0.9) = 1.5)$.

Exercício 4.2.

Considere um movimento Browniano standard $(B(t), t \geq 0)$ nos instantes $0 < u < u + v < u + v + w$, em que $u, v, w > 0$. Calcule

$$E(B(u)B(u+v)B(u+v+w)).$$

Exercício 4.3.

Seja $(B(t), t \geq 0)$ com $B(0) \equiv 3$, um movimento Browniano com variância σ^2 . Determine

$$Cov(B(t), B(s)), \quad t, s \geq 0.$$

Exercício 4.4.

Considere um movimento Browniano standard $(B(t), t \geq 0)$. Determine as funções de covariância para os processos estocásticos seguintes:

- (a) $U(t) = e^{-t}B(e^{2t}), \quad t \geq 0$.
- (b) $V(t) = (1-t)B\left(\frac{t}{1-t}\right)$, para $0 < t < 1$.
- (c) $W(t) = tB\left(\frac{1}{t}\right)$, com $W(0) = 0$.

Exercício 4.5.

Considere um movimento Browniano standard $(B(t), t \geq 0)$. Para t fixo e $M(t) = \max_{0 \leq u \leq t} B(u)$, mostre que:

- (a) $M(t)$ e $|B(t)|$ têm a mesma distribuição com f.d.p.

$$f_{M(t)}(x) = \frac{2}{\sqrt{t}}\phi(x/\sqrt{t}), \quad x > 0.$$

- (b) $E(M(t)) = \sqrt{2t/\pi}$.

Exercício 4.6.

Sejam $B_1(t)$ e $B_2(t)$ dois movimentos Brownianos independentes e $R(t) = \sqrt{B_1(t)^2 + B_2(t)^2}$, $t \geq 0$. Calcule $E(R(t))$.

Exercício 4.7.

As flutuações de preço das acções de determinada companhia são modeladas por um movimento Browniano $(A(t), t \geq 0)$. Suponha que a companhia entra em falência se o preço de mercado das acções atingir o nível zero.

Se o valor inicial das acções for $A(0) = 5$ u.m., determine a probabilidade de ...

- (a) ... a companhia entrar em falência no instante $t = 25$.
- (b) ... as acções estarem acima de 10 unidades monetárias no instante $t = 25$.

Exercício 4.8.

Considere um movimento Browniano com parâmetros $\mu = 0.1$ e $\sigma = 2$. Calcule a probabilidade do processo sair fora do intervalo $(a, b]$ no ponto b , partindo de $X(0) = 0$, para $b = 1, 10, 100$ e $a = -b$.

Exercício 4.9.

A flutuação do preço de determinado tipo de acções pode ser descrita por um movimento browniano geométrico com desvio-padrão $\alpha = 0$. Supondo que adquira estas acções, quais são as hipóteses de ver o seu capital investido duplicar?

4.2 O integral de Itô

Nota . No que se segue, adoptámos a seguinte notação para esperança matemática e probabilidade condicionadas:

$$E(\cdot | X_s = x) = E_{s,x}(\cdot)$$

e

$$P(\cdot | X_s = x) = P_{s,x}(\cdot).$$

Definição 4.4 (Processo de difusão). Seja (Ω, \mathcal{F}, P) um espaço de probabilidade e $(X_t, t \geq 0)$ um processo estocástico definido nesse espaço. Diz-se que X_t é um **processo de difusão** se satisfizer as seguintes propriedades:

- i) X_t é um processo de Markov;
- ii) As trajectórias de X_t são quase certamente contínuas;
- iii) $X_t \in L^2$, isto é, $E[X_t^2] < +\infty$;
- iv) Para todo $\varepsilon > 0$, tem-se

$$\lim_{\Delta \rightarrow 0^+} \frac{P_{s,x}(|X_{s+\Delta} - X_s| > \varepsilon)}{\Delta} = 0;$$

- v) Existe, e é finito, o limite

$$\lim_{\Delta \rightarrow 0^+} E_{s,x} \left[\frac{X_{s+\Delta} - X_s}{\Delta} \right] = a(s, x);$$

- vi) Existe, e é finito, o limite

$$\lim_{\Delta \rightarrow 0^+} E_{s,x} \left[\frac{(X_{s+\Delta} - X_s)^2}{\Delta} \right] = b(s, x).$$

Se as funções $a(s, x)$ e $b(s, x)$ forem independentes da variável temporal s , o processo diz-se **homogêneo**.

As funções $a(s, x)$ e $b(s, x)$ designam-se, respectivamente, por **coeficiente de tendência** (ou **momento infinitesimal de primeira ordem**) e **coeficiente de difusão** (ou **momento infinitesimal de segunda ordem**).

O coeficiente de tendência, $a(s, x)$, mede a velocidade da média do processo no instante s , enquanto que o coeficiente de difusão, $b(s, x)$, mede a intensidade das flutuações do processo, ou seja, mede a velocidade da variância do processo no instante s .

Nota: Existem na literatura definições alternativas para processo de difusão, algumas das quais assumem hipóteses adicionais ou diferentes.

Exercício 4.10 (Fonte: Braumann (2005)).

- (i) Mostre que o processo de Wiener W_t é um processo de difusão homogêneo com coeficiente de tendência nulo e coeficiente de difusão unitário.
- (ii) Mostre que $X_t = x_0 + \sigma W_t$, com x_0 e σ constantes, sendo um processo de Wiener (não-padrão), é um processo de difusão homogêneo com coeficiente de tendência nulo e coeficiente de difusão σ^2 .
- (iii) Mostre que $Z_t = x_0 + \mu t + \sigma W_t$, com x_0 , μ e σ constantes, conhecido como movimento browniano com tendência, é um processo de difusão homogêneo com coeficiente de tendência μ e coeficiente de difusão σ^2 .

Definição 4.5 (Função delta de Dirac). Chama-se **função delta de Dirac** à função generalizada $\delta(x)$ com as seguintes propriedades:

1. $\delta(x) = 0$, para todo o $x \neq 0$;
2. $\delta(0) = +\infty$;
3. $\int_{-\infty}^{+\infty} \delta(x) dx = 1$.

A caracterização, do ponto de vista probabilístico, de um processo de difusão recorre apenas aos seus momentos infinitesimais e às equações de Kolmogorov.

Teorema 4.1. *Seja X_t um processo de difusão, como definido anteriormente, com função densidade de transição $p(t, y | s, x)$, contínua em s , com derivadas parciais de primeira e segunda ordem em relação a x finitas e contínuas em s . Nessas condições, verificam-se:*

1. **Equação de Kolmogorov progressiva (ou equação de Fokker-Planck):**

$$\frac{\partial p}{\partial t} + \frac{\partial(a(s, x)p)}{\partial y} - \frac{1}{2} \frac{\partial^2(b(s, x)p)}{\partial y^2} = 0,$$

com condição inicial

$$\lim_{t \downarrow s} p(t, y | s, x) = \delta(x - y),$$

onde δ representa a função delta de Dirac, e (s, x) está fixo;

2. **Equação de Kolmogorov regressiva:**

$$\frac{\partial p}{\partial s} + a(s, x) \frac{\partial p}{\partial x} + \frac{1}{2} b(s, x) \frac{\partial^2 p}{\partial x^2} = 0,$$

com condição inicial

$$\lim_{t \uparrow s} p(t, y | s, x) = \delta(x - y),$$

onde δ representa a função delta de Dirac, e (t, y) está fixo.

Considere-se o ponto $X(0) = X_0 \in \mathbb{R}$ e o seguinte problema de Cauchy, induzido por uma equação diferencial ordinária:

$$\begin{cases} dX(t) = f(X(t)) dt, & \text{para } t > 0, \\ X(0) = X_0, \end{cases}$$

onde $f : \mathbb{R} \rightarrow \mathbb{R}$ é uma função diferenciável, e $X : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ é a solução do problema (4.2).

Se interpretarmos $X(t)$ como a trajectória de uma partícula, então $dX(t)/dt$ representa a sua velocidade. É natural admitir que essa velocidade apresente pequenas oscilações que não são explicadas pela função f , ou seja, o sistema descrito na equação (4.2) não incorpora o efeito aleatório que as flutuações ambientais induzem na trajectória de X . Assim, torna-se necessário adicionar um *ruído* ao problema (4.2), de modo a reflectir a influência dessas flutuações sobre a dinâmica do sistema:

$$\begin{cases} dX(t) = f(X(t)) dt + g(X(t)) \xi(t) dt, & \text{para } t > 0, \\ X(0) = X_0, \end{cases}$$

onde $g(\cdot)$, que mede a intensidade das flutuações ambientais, é uma função dependente de $X(t)$.

Considerando que $dW(t) = \xi(t) dt$, o sistema (4.2) pode reescrever-se da seguinte forma:

$$\begin{cases} dX(t) = f(X(t)) dt + g(X(t)) dW(t), \\ X(0) = X_0, \end{cases}$$

o qual representa uma **Equação Diferencial Estocástica (EDE)**. A solução deste sistema é dada por:

$$X(t) = X_0 + \int_0^t f(X(s)) ds + \int_0^t g(X(s)) dW(s), \quad t > 0,$$

em que o primeiro integral é um integral de Riemann-Stieltjes. Contudo, o segundo integral **não existe** neste sentido, dado que as trajectórias do processo de Wiener são, quase certamente, de variação ilimitada no intervalo $[0, t]$.

No entanto, como o processo de Wiener possui variação quadrática finita, é possível definir o segundo integral recorrendo à **definição de integral estocástico**.

Note-se que, como já referido, se omitiu a dependência explícita em ω na notação de $X(t)$.

Mostraremos de seguida como obter a solução (4.2), bem como a definição do integral estocástico

$$\int_0^t g(X(s)) dW(s).$$

Suponhamos que desejamos calcular o seguinte integral:

$$\int_0^t W(t) dW(t).$$

Se aplicarmos as regras de cálculo habituais, obtemos como solução:

$$\frac{1}{2} W^2(t).$$

Vamos verificar se esta solução está correta.

Seja $f : [0, t] \rightarrow \mathbb{R}^+$, com $f(u) = W(u)$, uma função, e sejam $\mathcal{P}_n = \{t_0^n, t_1^n, \dots, t_n^n\}$, $n = 1, 2, \dots$, partições do intervalo $[0, t]$ com

$$0 = t_0^n < t_1^n < \dots < t_n^n = t \geq 0,$$

tais que os diâmetros

$$\delta_n = \max_{0 \leq i \leq n-1} |t_{i+1}^n - t_i^n|$$

satisfazem $\delta_n \rightarrow 0$ quando $n \rightarrow +\infty$.

Consideremos as somas de Riemann-Stieltjes aproximadoras do integral $\int_0^t f(u) dW(u)$:

$$\sum_{i=0}^{n-1} W(\xi_i^n)(W(t_{i+1}^n) - W(t_i^n)),$$

com $\xi_i^n \in [t_i^n, t_{i+1}^n]$, e usemos limites em média quadrática quando $n \rightarrow +\infty$ como possível definição do integral.

Consideremos o caso particular $\xi_i^n = (1 - \lambda)t_i^n + \lambda t_{i+1}^n$, e definamos as somas de Riemann-Stieltjes:

$$S_\lambda(W(t)) = \sum_{i=0}^{n-1} W(\xi_i^n)(W(t_{i+1}^n) - W(t_i^n)).$$

Facilmente verificamos que, para λ fixo, o limite em média quadrática destas somas, quando $n \rightarrow +\infty$, é

$$\frac{W^2(t)}{2} + \left(\lambda - \frac{1}{2}\right)t.$$

Com efeito:

$$E \left[\left(S_\lambda(W(t)) - \frac{W^2(t)}{2} - \left(\lambda - \frac{1}{2}\right)t \right)^2 \right] \rightarrow 0.$$

Este limite depende da escolha do valor de λ e, consequentemente, do ponto intermédio $\xi_i \in [t_i, t_{i+1}]$. Assim, **não existe** o integral no sentido de Riemann-Stieltjes, pois falha a existência de um limite comum para todas as escolhas de pontos intermédios.

Ao fixarmos $\lambda = 0$, obtemos como ponto intermédio o ponto inicial do intervalo, isto é, $\xi_i = t_i$, e verificamos que

$$\int_0^t W(t) dW(t) = \frac{1}{2}W^2(t) - \frac{1}{2}t,$$

o que é um resultado **diferente** do indicado em (4.2). De facto, para diferentes valores de λ , obtemos diferentes integrais. Por exemplo, se considerarmos $\lambda = \frac{1}{2}$, o resultado do integral é:

$$\int_0^t W(t) dW(t) = \frac{1}{2}W^2(t).$$

O facto de diferentes valores de λ implicarem diferentes integrais levanta uma questão pertinente: **qual o valor de λ que devemos escolher?**

A escolha de $\xi_i = t_i$, ou seja, o ponto inicial, permite-nos definir integrais de funções mais gerais do que apenas o processo de Wiener. Isto conduz a integrais do tipo:

$$\int_0^t G(s) dW(s),$$

onde G pertence a uma vasta classe de funções com a propriedade de serem **não-antecipativas**. Veremos mais à frente como definir rigorosamente estas funções.

Como se referiu, a escolha de λ permite obter diferentes integrais. Assim:

- (i) Se $\lambda = 0$, escolhemos o ponto inicial do intervalo e obtemos o **integral de Itô**;
- (ii) Se $\lambda = \frac{1}{2}$, escolhemos o ponto intermédio do intervalo e obtemos o **integral de Stratonovich**.

Vamos agora dedicar-nos ao estudo do integral de Itô. Começamos com a introdução de algumas definições e resultados importantes.

Definição 4.6. Seja $W(t)$, $t \geq 0$, um processo de Wiener padrão definido num espaço de probabilidade (Ω, \mathcal{F}, P) .

1. Chama-se **filtração natural do processo de Wiener até ao instante $s > 0$** à σ -álgebra

$$\mathcal{M}_s = \sigma(W(u), 0 \leq u \leq s);$$

2. Chama-se **σ -álgebra dos incrementos futuros do processo de Wiener** à σ -álgebra

$$\mathcal{M}_s^+ = \sigma(W(u) - W(s), u \geq s);$$

3. Uma família $\{\mathcal{A}_s : 0 \leq s \leq t\}$ de σ -álgebras é chamada **filtração não-antecipativa**, relativamente a $W(s)$, se:

- $\mathcal{A}_s \supset \mathcal{M}_s$, $0 \leq s \leq t$;
- \mathcal{A}_s é independente de \mathcal{M}_s^+ , $\forall s \geq 0$.

Informalmente, podemos dizer que a filtração \mathcal{A}_s contém toda a informação disponível do processo até ao instante s .

A escolha da filtração não-antecipativa \mathcal{A}_s costuma coincidir com a própria filtração natural do processo de Wiener, \mathcal{M}_s , desde que não seja necessário incluir informação adicional sobre o processo. Caso contrário, considera-se uma filtração **maior** (por exemplo, de modo a incluir a condição inicial de um problema de Cauchy), desde que a mesma seja não-antecipativa.

Definição 4.7 (Processo não-antecipativo). Um processo estocástico $G(t)$ é chamado de **não-antecipativo**, relativamente à filtração \mathcal{A}_t , se $G(t)$ é \mathcal{A}_t -mensurável, para todo $t \geq 0$ (ou seja, $G(t)$ depende apenas da informação disponível até ao instante t).

Tendo em conta estas definições, podemos definir o integral de Itô para uma classe especial de funções não-antecipativas, as **funções em escada**. Nota: na realidade, para definir o integral de Itô, não basta que G seja não-antecipativa. É necessário que $G = G(t, \omega)$ seja **conjuntamente mensurável**.

Definição 4.8 (Espaço de Hilbert). Chama-se **espaço de Hilbert**, no intervalo $[0, t]$, e representa-se por $H^2[0, t]$, ao espaço das funções

$$G : [0, t] \times \Omega \rightarrow \mathbb{R}$$

que verificam as seguintes condições:

- G é conjuntamente mensurável relativamente à medida de Lebesgue l em $[0, t]$ e à medida de probabilidade P ;
- G é **não-antecipativa**;
- $\int_0^t E[G^2(u, \omega)] du < +\infty$.

Nota . Nesta última definição considerámos, de modo abusivo, uma qualquer função G com a propriedade de ser conjuntamente mensurável relativamente às medidas l e P .

Formalmente, deveríamos referir-nos ao conjunto das funções conjuntamente mensuráveis que sejam **quase iguais**, no seguinte sentido: duas funções G_1 e G_2 dizem-se quase iguais quando o conjunto de pontos (t, ω) onde diferem tem medida nula relativamente à medida produto $l \times P$.

Assim, a função G é, na realidade, um **representante da classe de equivalência** das funções conjuntamente mensuráveis relativamente à relação de equivalência de quase-igualdade.

Deste modo, para simplificar a linguagem, referimo-nos à função G como um representante da classe de equivalência, em vez de referir explicitamente a própria classe de equivalência.

Definição 4.9 (Função em escada). Uma função G , no espaço $H^2[0, t]$, é chamada de **função em escada** se existir uma partição $\{0 = t_0 < t_1 < \dots < t_n = t\}$ do intervalo $[0, t]$ tal que:

$$G(t) = G(t_i), \quad t_i \leq t \leq t_{i+1}, \quad i = 0, \dots, n-1.$$

Note-se que $G(t_i)$ é \mathcal{A}_{t_i} -mensurável, pois G é não-antecipativa.

Ao espaço de funções em escada de $H^2[0, t]$, chamamos $H_E^2[0, t]$.

Definição 4.10 (Integral de Itô para funções em escada). Seja G uma função em $H_E^2[0, t]$. O integral de Itô da função G no intervalo $[0, t]$ é dado por:

$$\int_0^t G(s) dW(s) = \sum_{i=0}^{n-1} G(t_i) (W(t_{i+1}) - W(t_i)).$$

Teorema 4.2 (Propriedades do integral de Itô). *Sejam F e G duas funções em $H_E^2[0, t]$, e $\alpha, \beta \in \mathbb{R}$ duas constantes. Verificam-se as seguintes propriedades:*

1. *Linearidade:*

$$\int_0^t (\alpha F(s) + \beta G(s)) dW(s) = \alpha \int_0^t F(s) dW(s) + \beta \int_0^t G(s) dW(s);$$

2. *Esperança nula:*

$$E \left[\int_0^t F(s) dW(s) \right] = 0;$$

3. *Isometria de Itô:*

$$E \left[\left(\int_0^t F(s) dW(s) \right)^2 \right] = E \left[\int_0^t (F(s))^2 ds \right] = \int_0^t E[(F(s))^2] ds.$$

Definimos, assim, o integral de Itô para funções em escada, ou seja, funções no espaço $H_E^2[0, t]$. Vamos agora generalizar este integral para funções genéricas em $H^2[0, t]$, através da existência de sucessões aproximadoras de funções em escada.

Teorema 4.3 (Aproximação em média quadrática). *Seja $G \in H^2[0, t]$ uma função. Então, existe uma sucessão de funções limitadas em escada, $G_n \in H_E^2[0, t]$, tal que:*

$$E \left[\int_0^t |G(s) - G_n(s)|^2 ds \right] \xrightarrow{m.q.} 0, \quad n \rightarrow +\infty.$$

Definição 4.11. Sejam G e G_n como no teorema anterior. O **integral de Itô** da função G no intervalo $[0, t]$ é definido como:

$$\int_0^t G(s) dW(s) = \lim_{n \rightarrow +\infty} \int_0^t G_n(s) dW(s),$$

onde o limite é tomado em **média quadrática**.

Teorema 4.4 (Propriedades do integral de Itô). *Sejam F e G duas funções em $H^2[0, t]$, e $\alpha, \beta \in \mathbb{R}$ duas constantes. Verificam-se as seguintes propriedades:*

1. **Linearidade:**

$$\int_0^t (\alpha F(s) + \beta G(s)) dW(s) = \alpha \int_0^t F(s) dW(s) + \beta \int_0^t G(s) dW(s).$$

2. **Esperança nula:**

$$E \left[\int_0^t F(s) dW(s) \right] = 0.$$

3. **Isometria de Itô:**

$$E \left[\left(\int_0^t F(s) dW(s) \right)^2 \right] = E \left[\int_0^t F(s)^2 ds \right] = \int_0^t E[F(s)^2] ds.$$

4. **Covariância:**

$$E \left[\int_0^t F(s) dW(s) \int_0^t G(s) dW(s) \right] = E \left[\int_0^t F(s) G(s) ds \right].$$

5. **Distribuição normal no caso determinístico (se $G(s)$ for determinística):**

$$\int_0^t G(s) dW(s) \sim \mathcal{N} \left(0, \int_0^t G^2(s) ds \right).$$

O integral de Itô para funções no espaço $H^2[0, t]$ pode ser estudado como função do seu limite superior, ou seja, como um **integral indefinido**.

A prova destas propriedades encontra-se fora do âmbito desta unidade curricular.

Definição 4.12 (Integral indefinido de Itô). Seja $G \in H^2[0, d]$ uma função, e $[0, d]$ um intervalo. O integral de Itô da função G , considerando t como limite superior de integração, é dado por:

$$Z(t) = \int_0^t G(s) dW(s) = \int_0^d G(s) I_{[0,t]}(s) dW(s).$$

Teorema 4.5. *Seja $Z(t)$ o processo estocástico definido acima. São válidas as seguintes propriedades:*

1. $Z(t)$ é uma martingala relativamente à filtração \mathcal{A}_t ;
2. $Z(t)$ possui uma versão contínua (com trajectórias quase certamente contínuas);
3. $Z(t)$ tem incrementos não correlacionados.

As classes de funções até aqui apresentadas são bastante simples. Na prática, interessa-nos estudar integrais de Itô em que a função G não pertence apenas ao espaço $H^2[0, t]$, mas sim a uma classe mais ampla: o espaço $M^2[0, t]$.

Definição 4.13. Dizemos que $G(s, \omega)$ é uma função no espaço $M^2[0, t]$ se:

1. É **conjuntamente mensurável**;
2. É **não-antecipativa** em relação à filtração \mathcal{A}_s ;
3. O integral

$$\int_0^t G^2(s) ds$$

existe e é **finito quase certamente**.

Note-se que a exigência

$$\int_0^t G^2(s) ds < +\infty$$

é mais fraca do que a condição exigida para o espaço H^2 . Assim, temos a inclusão:

$$H^2[0, t] \subset M^2[0, t]$$

A extensão do integral de Itô a funções do espaço $M^2[0, t]$ é feita de forma semelhante à aproximação por funções em escada em $H_E^2[0, t]$, com a diferença de que a **convergência requerida é mais fraca**.

Teorema 4.6. *Seja $G \in M^2[0, t]$. Então, existe uma sucessão de funções limitadas em escada $G_n \in H_E^2[0, t]$ tal que:*

$$\int_0^t (G(s) - G_n(s))^2 ds \rightarrow 0 \quad \text{quase certamente, } n \rightarrow +\infty$$

Definição 4.14. Sejam G e G_n como no teorema anterior. O **integral de Itô** da função G no intervalo $[0, t]$ é definido por:

$$\int_0^t G(s) dW(s) = P - \lim_{n \rightarrow +\infty} \int_0^t G_n(s) dW(s),$$

onde o limite é tomado **em probabilidade**.

Nota (Nota sobre propriedades do integral). Dada a natureza das funções no espaço $M^2[0, t]$, não existe garantia de que as propriedades clássicas do integral de Itô — tais como esperança nula, isometria, e covariância — se verifiquem, pois os respetivos momentos podem não existir.

Finda a apresentação do integral de Itô, é agora necessário introduzir as **regras de cálculo** destes integrais: o chamado **cálculo de Itô**.

O cálculo de Itô difere do cálculo clássico devido à introdução de uma nova regra de diferenciação — a **regra da cadeia de Itô**. Apresentamos de seguida a definição de **processo de Itô** e o respetivo **teorema de Itô**, base fundamental do cálculo de integrais estocásticos.

Definição 4.15 (Processo de Itô). Sejam:

- $(W(t), t \geq 0)$ o processo de Wiener;
- X_0 uma variável aleatória \mathcal{A}_0 -mensurável;
- F uma função conjuntamente mensurável, adaptada à filtração \mathcal{A}_s e tal que

$$\int_0^d |F(s)| ds < +\infty \quad \text{quase certamente;}$$

- $G \in M^2[0, d]$.

Define-se o **processo de Itô** no intervalo $t \in [0, d]$ como:

$$X(t) = X_0 + \int_0^t F(s) ds + \int_0^t G(s) dW(s).$$

Este processo pode também ser representado na forma diferencial:

$$dX(t) = F(t) dt + G(t) dW(t).$$

Teorema 4.7 (Teorema de Itô). *Seja $X(t, \omega)$ um processo de Itô como definido anteriormente, e seja $Y(t) = h(t, X(t))$, onde h , $h_t(t, x)$ e $h_{xx}(t, x)$ são funções contínuas. Então:*

- (i) $Y(t) = Y(t, \omega)$ é um processo de Itô com condição inicial $Y_0 = h(0, X_0)$;

(ii) a forma diferencial de $Y(t)$ é dada pela **regra da cadeia de Itô**:

$$dY_t = \left(\frac{\partial h(t, X_t)}{\partial t} + \frac{\partial h(t, X_t)}{\partial x} F(t) + \frac{1}{2} \frac{\partial^2 h(t, X_t)}{\partial x^2} G^2(t) \right) dt + \frac{\partial h(t, X_t)}{\partial x} G(t) dW_t;$$

(iii) a forma integral de $Y(t)$ é dada por:

$$Y_t = Y_0 + \int_0^t \left(\frac{\partial h(s, X_s)}{\partial s} + \frac{\partial h(s, X_s)}{\partial x} F(s) + \frac{1}{2} \frac{\partial^2 h(s, X_s)}{\partial x^2} G^2(s) \right) ds + \int_0^t \frac{\partial h(s, X_s)}{\partial x} G(s) dW_s.$$

Finda a apresentação de definições, propriedades e teoremas relativos ao cálculo de Itô, podemos agora abordar a resolução de equações diferenciais estocásticas, ou seja, o cálculo das suas soluções. Começamos pela definição de solução de uma equação diferencial estocástica de Itô.

No que se segue, consideramos:

- $W = (W_t, t \geq 0)$ é um processo de Wiener;
- X_0 é uma variável aleatória independente do processo de Wiener;
- $\mathcal{A}_t = \mathcal{F}(X_0, W_s), 0 \leq s \leq t$;
- F, G duas funções definidas em $[0, T]$, conjuntamente mensuráveis, com $T > 0$.

Definição 4.16 (Solução de uma EDE de Itô). Um processo estocástico X_t é solução da equação diferencial estocástica de Itô

$$\begin{cases} dX_t = F(X_t, t) dt + G(X_t, t) dW_t, & 0 \leq t \leq T \\ X(0) = X_0, \end{cases}$$

se satisfizer as seguintes condições:

- (i) X é \mathcal{F}_t -mensurável;
- (ii) F é não-antecipativa e

$$\int_0^T F(X_s, s) ds < +\infty;$$

- (iii) G é não-antecipativa e

$$\int_0^T G^2(X_s, s) ds < +\infty;$$

- (iv)

$$X_t = X_0 + \int_0^t F(X_s, s) ds + \int_0^t G(X_s, s) dW_s \quad q.c., \quad \forall t \in [0, T].$$

Teorema 4.8 (Teorema de existência e unicidade de soluções de EDE de Itô). *Sejam $F : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ e $G : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ duas funções contínuas que satisfazem as seguintes condições:*

- (i) $|F(x, t) - F(y, t)| \leq L|x - y|$ e $|G(x, t) - G(y, t)| \leq L|x - y|$, para todo o $t \in [0, T]$ e $x, y \in \mathbb{R}$;
- (ii) $|F(x, t)| \leq L(1 + |x|)$ e $|G(x, t)| \leq L(1 + |x|)$, para todo o $t \in [0, T]$ e $x \in \mathbb{R}$,

onde $L > 0$ é uma constante.

Seja X_0 uma variável aleatória, independente dos incrementos futuros do processo de Wiener, tal que

$$E(|X_0|^2) < +\infty.$$

Nestas condições, existe uma única solução X_t da equação diferencial estocástica de Itô:

$$\begin{cases} dX_t = F(X_t, t) dt + G(X_t, t) dW_t, & 0 \leq t \leq T \\ X(0) = X_0. \end{cases}$$

Esta solução é um processo de Markov e, se F e G forem contínuas em t , trata-se também de um processo de difusão.

A unicidade enunciada significa o seguinte: se X_t e Y_t forem soluções da equação (4.8), então

$$P(X_t = Y_t) = 1, \quad \forall t \in [0, T].$$

As condições impostas às funções F e G correspondem, respetivamente, a uma condição de Lipschitz (continuidade uniforme) e a uma restrição de crescimento linear.

A demonstração deste teorema recorre ao *Lema de Gronwall* e pode ser encontrada em qualquer bom livro sobre equações diferenciais estocásticas.

Exercício 4.11 (Fonte: Braumann (2005)). Determine $d(tW(t))$ e utilize o resultado para mostrar que

$$\int_0^t s dW(s) = tW(t) - \int_0^t W(s) ds.$$

Exercício 4.12 (Fonte: Braumann (2005)). Mostre que a equação $dY(t) = Y(t) dW(t)$, com $Y(0) = 1$, tem como solução

$$Y(t) = \exp\left(W(t) - \frac{t}{2}\right), \quad \text{para } t \geq 0.$$

Exercício 4.13. Considere a seguinte EDE:

$$dY(t) = \mu dt + \sigma dW(t), \quad Y(0) = y_0.$$

Mostre que a sua solução é dada por:

$$Y(t) = y_0 + \mu t + \sigma W(t).$$

Sugestão: esta EDE é linear com coeficientes constantes. Resolva-a diretamente por integração.

Exercício 4.14. Considere a seguinte EDE, conhecida como modelo de **Ornstein-Uhlenbeck**:

$$dX(t) = -\theta X(t) dt + \sigma dW(t), \quad X(0) = x_0.$$

Mostre que a solução deste modelo é dada por:

$$X(t) = x_0 e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-s)} dW(s).$$

Sugestão: aplique a mudança de variável $Z(t) = e^{\theta t} X(t)$ e resolva a EDE resultante.

Exercício 4.15. Considere a seguinte EDE, conhecida como modelo de **Vasicek**:

$$dY(t) = b(A - Y(t)) dt + \sigma dW(t), \quad Y(0) = y_0.$$

Mostre que a solução deste modelo é dada por:

$$Y(t) = A + (y_0 - A)e^{-bt} + \sigma \int_0^t e^{-b(t-s)} dW(s).$$

Sugestão: aplique a mudança de variável $Z(t) = Y(t) - A$ e resolva a EDE resultante.

Exercício 4.16. Considere a seguinte EDE, conhecida como modelo de **Gompertz** (ou de **Fox**):

$$dX(t) = rX(t)(\ln K - \ln X(t)) dt + \sigma X(t) dW(t), \quad X(0) = x_0.$$

Mostre que a solução deste modelo é dada por:

$$X(t) = \exp\left(\ln K + e^{-rt}(\ln x_0 - \ln K) - \frac{\sigma^2}{2r}(1 - e^{-rt}) + \sigma \int_0^t e^{-r(t-s)} dW_s\right).$$

Sugestão: aplique a mudança de variável $Z(t) = \ln X(t)$ e resolva a EDE resultante.

Exercício 4.17. Considere a seguinte EDE, conhecida como modelo de **Black-Scholes**:

$$dY(t) = rY(t) dt + \sigma Y(t) dW(t), \quad Y(0) = y_0.$$

Mostre que a solução deste modelo é dada por:

$$Y(t) = y_0 e^{\left(r - \frac{\sigma^2}{2}\right)t + \sigma W(t)}.$$

Sugestão: aplique a mudança de variável $Z(t) = \ln Y(t)$ e resolva a EDE resultante.

Exercício 4.18. Seja $X(t)$ o valor de uma ação no instante $t \geq 0$, assumindo que satisfaz o modelo de Black-Scholes, com $X(0) = \$52.800$, $r = 0.312/\text{trimestre}$ e $\sigma^2 = 0.087/\text{trimestre}$. Determine:

- (a) $P(X(2 \text{ trimestres}) > \$70.000 \mid X(1 \text{ trimestre}) = \$60.500)$
- (b) $E(X(1 \text{ trimestre}))$
- (c) $P(\$55.000 \leq X(1 \text{ trimestre}) \leq \$65.000)$
- (d) $Var(X(1 \text{ trimestre}))$
- (e) $E(X(2 \text{ trimestres}) \mid X(0.5 \text{ trimestre}) = \$54.200 \text{ e } X(1 \text{ trimestre}) = \$60.500)$
- (f) $P(X(2 \text{ trimestres}) > \$70.000 \mid X(0.5 \text{ trimestre}) = \$54.200 \text{ e } X(1 \text{ trimestre}) = \$60.500)$
- (g) $Var(X(2 \text{ trimestres}) \mid X(1 \text{ trimestre}) = \$60.500)$
- (h) $E(X(2 \text{ trimestres}) \mid X(1 \text{ trimestre}) = \$60.500)$

Sugestão: Use o facto de que $X(t) = X(0) \cdot e^{Z(t)}$, com $Z(t)$ gaussiano. Logo,

$$X^2(t) = X(0)^2 \cdot e^{2Z(t)}$$

e então,

$$E(X^2(t)) = X(0)^2 \cdot E(e^{2Z(t)}) = X(0)^2 \cdot \exp\left(E(2Z(t)) + \frac{1}{2}Var(2Z(t))\right).$$

Exercício 4.19. Considere a seguinte EDE, conhecida como modelo de **log-normal inverso**:

$$dY(t) = -\frac{\sigma^2}{2}Y(t) dt + \sigma Y(t) dW(t), \quad Y(0) = y_0.$$

Mostre que a solução deste modelo é dada por:

$$Y(t) = y_0 e^{\sigma W(t)}.$$

Sugestão: aplique a mudança de variável $Z(t) = \ln Y(t)$ e resolva a EDE resultante.

Exercício 4.20. Considere a seguinte EDE, conhecida como modelo de **Gompertz com parâmetro limite**:

$$dX(t) = (X(t) - \gamma)(\alpha - \beta \ln(X(t) - \gamma))dt + \sigma(X(t) - \gamma)dW(t), \quad X(0) = x_0$$

Mostre que a solução deste modelo é dada por:

$$X_t = \gamma + \exp \left\{ e^{-\beta t} \left(\ln(x_0 - \gamma) + \frac{1}{\beta} \left(\alpha - \frac{\sigma^2}{2} \right) (e^{\beta t} - 1) \right) + \sigma e^{-\beta t} \int_0^t e^{\beta s} dW_s \right\}.$$

Sugestão: aplique a mudança de variável $Y(t) = \ln(X(t) - \gamma)$ e resolva a EDE resultante.

Capítulo 5

Bibliografia

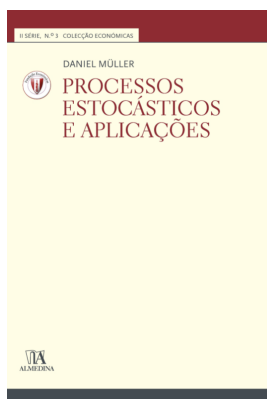
Principal

- Muller, D. (2007) Processos Estocásticos e Aplicações. II Série, nº3, Coleção Económicas. Almedina.

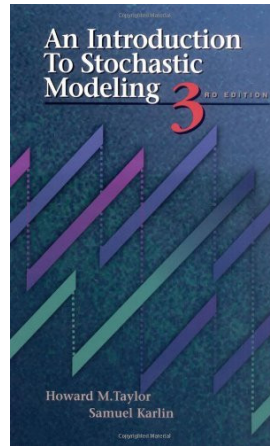


Secundária

- Muller, D. (2011) Probabilidades e Processos Estocásticos. II Série, nº17, Coleção Económicas. Almedina.



- Taylor, H. M., Karlin, S. (1998) An Introduction to Stochastic Modeling (3rd Edition), Academic Press, New York.



Todos os direitos reservados. É expressamente proibida a reprodução, cópia, distribuição, comunicação pública, transformação ou qualquer outra forma de utilização, total ou parcial, dos conteúdos deste sítio, incluindo textos, código e imagens, sem autorização prévia e por escrito do autor. Qualquer utilização não autorizada constitui violação dos direitos de autor e poderá dar lugar à responsabilidade civil e criminal nos termos da lei em vigor.