Processos Estocásticos e Aplicações

José Pedro Gaivão

May 3, 2019

Contents

Chapter 1

Noções gerais de teoria de probabilidades

Antes de iniciarmos o nosso estudo dos processos estocásticos convém relembrar alguns conceitos elementares de teoria de probabilidades.

1.1 Espaços de probabilidade e variáveis aleatórias

Seja Ω um conjunto. Designamos por Ω o **espaço dos resultados**. Uma colecção $\mathcal{F}$ de subconjuntos de Ω é uma **σ-álgebra** se as seguintes 3 condições se verificarem:

1. $\emptyset \in \mathcal{F}$,
2. $A \in \mathcal{F}$ sse $A^c \in \mathcal{F}$,
3. $\bigcup_n A_n \in \mathcal{F}$ para quaisquer $A_1, A_2, \ldots \in \mathcal{F}$.

Os elementos de uma σ-álgebra $\mathcal{F}$ designam-se por **eventos**. Uma **medida de probabilidade** é uma função $\mathbb{P}\colon \mathcal{F} \to [0,1]$ tal que

$$\mathbb{P}(\emptyset) = 0 \quad \text{e} \quad \mathbb{P}\left(\bigcup_n A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n),$$

para quaisquer eventos $A_1, A_2, \ldots \in \mathcal{F}$ disjuntos entre si. Ao triplo $(\Omega, \mathcal{F}, \mathbb{P})$ designamos por **espaço de probabilidade**. Os axiomas que definem um espaço de probabilidade foram introduzidos por Andrey Kolmogorov[1].

[1] Andrey Kolmogorov (1903-1987) foi um matemático russo que fez inúmeras contribuições fundamentais em diversas áreas da matemática

Variável discreta	$X(\Omega)$	Distribuição de probabilidade
Bernoulli	$\{0,1\}$	$\mathbb{P}(X=1)=p, \quad p\in[0,1]$
Binomial	$\{0,\dots,n\}$	$\mathbb{P}(X=k)=C^n_k p^k(1-p)^{n-k}$
Poisson	$\{0,1,2,\dots\}$	$\mathbb{P}(X=k)=e^{-\lambda}\frac{\lambda^n}{n!}, \quad \lambda>0$

Table 1.1: Exemplos de variáveis aleatórias discretas

Uma **variável aleatória** é uma função $\mathcal{F}$-mensurável $X:\Omega\to\mathbb{R}$. Ser $\mathcal{F}$-**mensurável** significa que a pré-imagem (ou imagem inversa) de qualquer conjunto aberto é um evento de $\mathcal{F}$, i.e., $X^{-1}(]a,b[)\in\mathcal{F}$ para quaisquer reais $a<b$.

A **função de distribuição de probabilidade**[2] de X é a função $F_X:\mathbb{R}\to\mathbb{R}$ definida por

$$F_X(x)=\mathbb{P}(X\leq x).$$

Há duas classes particulares de variáveis aleatórias: discretas e contínuas. Uma variável aleatória diz-se **discreta** se tomar um conjunto contável (finito ou infinito) de valores, isto é, $X(\Omega)=\{x_1,x_2,\dots\}$. Uma variável aleatória diz-se **contínua** se $\mathbb{P}(X=x)=0$, ou seja, F_X é uma função contínua[3]. Se F_X for continuamente diferenciável, então

$$F_X(x)=\int_{-\infty}^{x} f_X(u)\,du,$$

onde $f_X=F'_X$ é a **densidade de probabilidade** de X. Nesta disciplina vamos apenas considerar variáveis aleatórias discretas ou contínuas com densidade de probabilidade, também conhecidas por **absolutamente contínuas**. Ver Tabelas 1.1 e 1.2.

Se X é integrável relativamente a $\mathbb{P}$ então definimos o **valor esperado** de X como

$$E(X)=\int_\Omega X\,d\mathbb{P}.$$

No caso de X ser discreta ou contínua temos que

$$E(X)=\sum_n x_n\mathbb{P}(X=x_n) \quad \text{ou} \quad E(X)=\int_{-\infty}^{\infty} xf_X(x)\,dx.$$

[2]Relembrar que F_X é uma função monótona crescente, contínua à direita, $\lim_{x\to-\infty}F_X(x)=0$ e $\lim_{x\to\infty}F_X(x)=1$

[3]Uma vez que $\mathbb{P}(X=x)=F_X(x)-F_X(x^-)$.

Variável contínua	$X(\Omega)$	Distribuição de probabilidade
Uniforme	$[a,b]$	$f_X(x) = \begin{cases} 1/(b-a) & a \leq x \leq b \\ 0 & \text{caso contrário} \end{cases}$
Gaussiana	$\mathbb{R}$	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad \sigma > 0, \mu \in \mathbb{R}$
Exponencial	$[0,+\infty[$	$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{caso contrário} \end{cases}, \quad \lambda > 0$

Table 1.2: Exemplos de variáveis aleatórias contínuas

Em geral, definimos o **momento de ordem** n de X como $E(X^n)$ e o **momento central de ordem** n de X como $E((X-\mu)^n)$ onde $\mu = E(X)$. Ao momento central de segunda ordem chamamos **variância** de X,

$$V(X) = E((X-\mu)^2).$$

Da linearidade do valor esperado segue que

$$V(X) = E(X^2) - (E(X))^2.$$

Dado um vector de variáveis aleatórias $(X_1, \ldots X_n) : \Omega \to \mathbb{R}^n$ designa-se por $F_{X_1,\ldots,X_n}$ a sua **distribuição de probabilidade conjunta**,

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n).$$

Recordar, que duas variáveis aleatórias X e Y são **independentes** sse

$$F_{X,Y}(x,y) = F_X(x)F_Y(y).$$

Definimos a covariância de variáveis aleatórias X e Y como

$$\operatorname{cov}(X,Y) = E\left[(X - E(X))(Y - E(Y))\right].$$

Se X e Y são independentes, então $\operatorname{cov}(X,Y) = 0$.

Exercício 1. Seja X uma variável aleatória contínua em $[0,+\infty[$. Mostre que

$$E(X) = \int_0^\infty \mathbb{P}(X > x)\, dx$$

Exercício 2. Sejam X e Y duas variáveis aleatórias independentes com densidade de probabilidade f_X e f_Y, respectivamente. Mostre que a densidade de probabilidade de $Z = X + Y$ é

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-\eta) f_Y(\eta)\, d\eta$$

1.2 Esperança Condicional

A esperança condicional é um conceito central em teoria de probabilidades, em particular no estudo de processos estocásticos. Considere uma variável aleatória $X : \Omega \to \mathbb{R}$ definida num espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. Suponha que desejamos obter a melhor "aproximação" de X dado que conhecemos alguma informação de $\mathcal{F}$ (alguns dos seus eventos). A melhor aproximação de X é num certo sentido dada pela esperança condicional.

Definição 1.2.1. Seja $\mathcal{G}$ uma sub-σ-álgebra de $\mathcal{F}$, i.e., $\mathcal{G} \subset \mathcal{F}$. A esperança condicional $E(X|\mathcal{G}) : \Omega \to \mathbb{R}$ é uma variável aleatória que satisfaz:

1. $E(X|\mathcal{G})$ é $\mathcal{G}$-mensurável,
2. Para todo o evento $A \in \mathcal{G}$,

$$\int_A E(X|\mathcal{G})\, d\mathbb{P} = \int_A X\, d\mathbb{P}.$$

Observação 1.2.1. A σ-álgebra $\mathcal{G}$ representa alguma informação que é conhecida da experiência aleatória. Na definição está implícito que X é integrável, ou seja, $E(X) < \infty$. A existência da esperança condicional é consequência do teorema de Radon-Nikodym.

Usando a definição geral de esperança condicional podemos condicionar X numa outra variável aleatória $Y : \Omega \to \mathbb{R}$. Chamamos à seguinte coleção de eventos a informação de Y,

$$\sigma(Y) = \left\{Y^{-1}(B) \colon B \text{ é Boreliano}\right\}.$$

É fácil verificar que $\sigma(Y)$ é uma σ-álgebra de Ω. De facto, é a menor σ-álgebra que torna Y mensurável. Assim, define-se a **esperança condicional de X dado Y** como

$$E(X|Y) = E(X|\sigma(Y)).$$

Note-se que $E(X|Y)$ é $\sigma(Y)$-mensurável. Portanto, $E(X|Y)$ é constante nos eventos $\{Y = y\}$. Logo, existe uma função $E(X|Y = y)$ de variável independente y tal que

$$E(X|Y = y) = E(X|Y)(\omega), \quad \forall\, \omega \in \{Y = y\}.$$

Vejamos como calcular $E(X|Y = y)$ quando X e Y são discretas ou absolutamente contínuas.

1.2.1 Esperança condicional: caso discreto

Seja $Y : \Omega \to \mathbb{R}$ uma variável aleatória discreta. Suponha que $Y(\Omega) = \{y_1, y_2, \ldots\}$ e que $P(Y = y_n) > 0$ para todo o $n \geq 1$. A esperança condicional de X dado Y é calculada da seguinte forma,

$$E(X|Y = y) = \begin{cases} E(X|Y = y_1), & y = y_1 \\ E(X|Y = y_2), & y = y_2 \\ \vdots \end{cases}$$

onde

$$E(X|Y = y_n) = \frac{1}{\mathbb{P}(Y = y_n)} \int_{\{Y = y_n\}} X \, d\mathbb{P}\,.$$

Exemplo 1.2.2. Considere duas moedas de 20 e 50 cêntimos. Lançam-se as moedas ao ar e somam-se os montantes das moedas que ficaram com a face voltada para cima. Esse montante é o valor da variável aleatória X. Seja Y a variável aleatória que retorna o montante da primeira moeda, de 20 cêntimos, caso esta se encontre de face voltada para cima. Caso contrário tem o valor zero. Note-se que

$$\{Y = 0\} = \{EF, EE\} \quad \text{e} \quad \{Y = 20\} = \{FF, FE\},$$

onde E simboliza escudo e F face. Então

$$E(X|Y = 0) = 25 \quad \text{e} \quad E(X|Y = 20) = 45\,.$$

Logo,

$$E(X|Y = y) = \begin{cases} 25 & \text{se } y = 0 \\ 45 & \text{se } y = 20 \end{cases}$$

Suponhamos, tal como no exemplo anterior, que X é uma variável aleatória discreta (tal como Y). Então neste caso definimos a **função de massa de probabilidade condicionada** de X dado Y,

$$p_{X|Y}(x|y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}, \quad \text{se} \quad \mathbb{P}(Y = y) > 0.$$

É usual também escrever-se

$$\mathbb{P}(X = x|Y = y) = p_{X|Y}(x|y).$$

Com esta definição expressa-se $E(X|Y = y)$ através da fórmula

$$E(X|Y = y) = \sum_n x_n \, p_{X|Y}(x_n|y).$$

Usando a **lei de probabilidade total**,

$$\mathbb{P}(X = x) = \sum_n p_{X|Y}(x|y_n)\mathbb{P}(Y = y_n)$$

obtemos a seguinte proposição.

Proposição 1.2.3 (Lei de probabilidade total para a esperança condicional)**.** *Sejam X e Y duas variáveis aleatórias discretas. Então*

$$E(X) = \sum_n E(X|Y = y_n)\mathbb{P}(Y = y_n)$$

Demonstração. Exercício. □

1.2.2 Esperança condicional: caso contínuo

Suponhamos agora que X e Y são duas variável aleatória absolutamente contínuas com função densidade de probabilidade f_X e f_Y, respectivamente. A **função de densidade de probabilidade condicional de X dado Y** é definida por

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

A esperança condicional $E(X|Y = y)$ é calculada da seguinte maneira,

$$E(X|Y = y) = \int_{-\infty}^{+\infty} x f_{X|Y}(x|y)\, dx.$$

Usando a lei de probabilidade total

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X|Y}(x|y) f_Y(y)\, dy$$

obtém-se a seguinte proposição.

Proposição 1.2.4. *Sejam X e Y duas variável aleatória contínuas com função densidade de probabilidade f_X e f_Y. Então*

$$E(X) = \int_{-\infty}^{+\infty} E(X|Y=y) f_Y(y)\, dy$$

Exemplo 1.2.5. Suponha que X e Y têm a função de distribuição de densidade conjunta

$$f_{X,Y}(x,y) = 1/y e^{-\frac{x}{y}-y}. \quad x, y > 0.$$

Então

$$f_Y(y) = \int_0^{\infty} f_{X,Y}(x,y)\, dx = e^{-y}.$$

Logo

$$f_{X|Y}(x|y) = 1/y e^{-x/y}.$$

Portanto, a variável aleatória X condicionada a $Y = y$ tem distribuição exponencial com parâmetro $1/y$. Logo $E(X|Y = y) = y$.

1.2.3 Esperança condicional: caso vectorial

Dado um vector aleatório $(Y_1, \ldots, Y_n)$ define-se de forma análoga a esperança condicional no caso de discreto,

$$E(X|Y_1 = y_1, \ldots, Y_n = y_n) = \sum_k x_k\, p_{X|Y_1,\ldots,Y_n}(x_k|y_1, \ldots, y_n),$$

onde

$$p_{X|Y_1,\ldots,Y_n}(x|y_1, \ldots, y_n) = \frac{\mathbb{P}(X = x, Y_1 = y_1, \ldots Y_n = y_n)}{\mathbb{P}(Y_1 = y_1, \ldots Y_n = y_n)},$$

e no caso contínuo

$$E(X|Y_1 = y_1, \ldots, Y_n = y_n) = \int_{-\infty}^{+\infty} x f_{X|Y_1,\ldots,Y_n}(x|y_1, \ldots, y_n) dx$$

onde

$$f_{X|Y_1,\ldots,Y_n}(x|y_1, \ldots, y_n) = \frac{f_{X,Y_1,\ldots,Y_n}(x, y_1, \ldots, y_n)}{f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n)}$$

1.2.4 Propriedades

A seguinte proposição reúne a principais propriedades da esperança condicional.

Proposição 1.2.6. *Sejam X, X_1, X_2, Y variáveis aleatórias com valor esperado finito e $g, h : \mathbb{R} \to \mathbb{R}$ duas funções mensuráveis. A esperança condicional satisfaz as seguintes propriedades:*

1. $E(\alpha_1 X_1 + \alpha_2 X_2 | \mathcal{G}) = \alpha_1 E(X_1|\mathcal{G}) + \alpha_2 E(X_2|\mathcal{G})$,
2. $E(X|\mathcal{G}) \geq 0$ *se* $X \geq 0$
3. $E(g(Y)X|Y = y) = g(y)E(X|Y = y)$
4. $E(h(X)|Y = y) = E(h(X))$ *se* X *e* Y *forem independentes*
5. *Se* $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$, *então* $E(X|\mathcal{H}) = E(E(X|\mathcal{G})|\mathcal{H})$.

Demonstração. Procurar na bibliografia. □

Note-se que da propriedade (3) obtém-se

$$E(XY|Y = y) = yE(X|Y = y).$$

No caso de X e Y serem independentes, então (4) implica

$$E(X|Y = y) = E(X).$$

Ou seja, a esperança condicional é uma constante e não depende de y.

Exercício 3. Uma empresa produz diariamente N componentes electrónicas, onde N é uma variável aleatória com distribuição de Poisson e parâmetro $\lambda > 0$. Cada componente pode ter um defeito, independentemente das restantes, com probabilidade p. Supomos também que o defeito de cada componente é independente do número N de componentes electrónicas. Seja D o número diário de componentes electrónicas com defeito. Determine:

1. $E(D|N = n)$
2. $E(D)$
3. $E(N|D = d)$

Chapter 2

Noções gerais de processos estocásticos

2.1 Noções gerais

Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade.

Definição 2.1.1. Um **processo estocástico** X é uma colecção $X = \{X_t : t \in T\}$ de variáveis aleatórias $X_t : \Omega \to \mathbb{R}$ indexadas por um parâmetro $t \in T \subset \mathbb{R}$.

Quando o **conjunto dos parâmetros** T é um conjunto contável, tipicamente $T = \{0, 1, 2, \ldots\}$ ou $T = \mathbb{Z}$, então o processo estocástico é de **tempo discreto**. Neste contexto, uma sucessão de variáveis aleatórias $X_1, X_2, \ldots$ é um processo estocástico de tempo discreto.

Quando T é um intervalo, tipicamente $T = [0, \infty[$ ou $T = \mathbb{R}$, então o processo estocástico é de **tempo contínuo**. Neste contexto, é usual escrever-se $X(t)$ quando o processo é de tempo contínuo.

Ao conjunto dos valores que as variáveis aleatórias X_t podem tomar designa-se por E, o **conjunto dos estados do processo estocástico**. Quando o conjunto dos estados E é contável (finito ou infinito) então o processo estocástico diz-se **discreto**, caso contrário diz-se **contínuo**.

Para cada $\omega \in \Omega$ a função

$$T \ni t \mapsto X_t(\omega)$$

é designada por **trajectória** ou **realização** do processo. Uma trajectória de um processo referente a um período limitado de tempo é designada por **série temporal**. Ver Figuras 2.1 e 2.2.

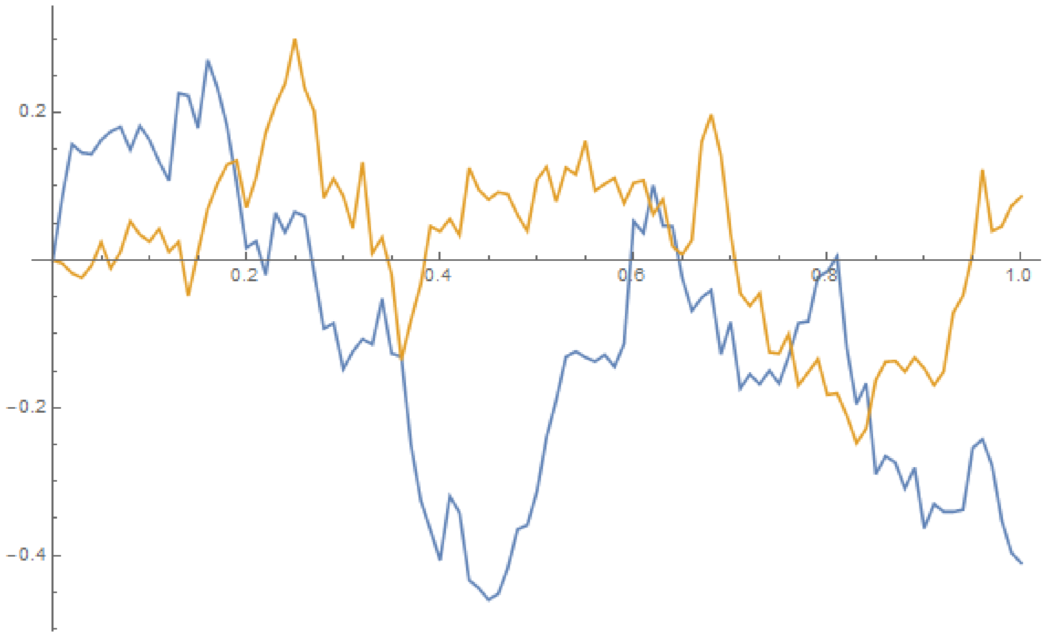

Figure 2.1: Duas realizações de um processo estocástico a tempo contínuo

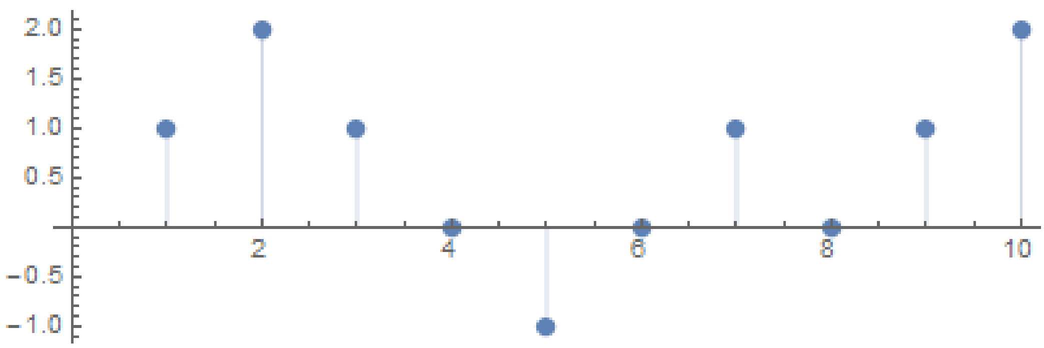

Figure 2.2: Realização de um passeio aleatório

A **lei de probabilidade** do processo estocástico é dada por todas as distribuições de probabilidade conjuntas de um número finito de variáveis aleatórias $(X_{t_1}, \ldots, X_{t_n})$. Portanto, a lei de probabilidade do processo consiste na família de funções de distribuição de probabilidade conjuntas,

$$F_{X_{t_1},\ldots,X_{t_n}}(x_1, \ldots, x_n) = P(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n)\,,$$

onde $t_1, \ldots, t_n$ são quaisquer conjunto finito de índices pertencentes a T e $n \geq 1$.

2.2 Processos iid

Definição 2.2.1. Um processo estocástico $X = \{X_t : t \in T\}$ é **iid (independente e identicamente distribuído** se

1. Para todo o conjunto finito de índices $t_1, \ldots, t_n \in T$ com $n \in \mathbb{N}$, as variáveis $X_{t_1}, \ldots, X_{t_n}$ são independentes, i.e.,

$$F_{X_{t_1},\ldots,X_{t_n}} = F_{X_{t_1}} \cdots F_{X_{t_n}}$$

2. As variáveis aleatórias são identicamente distribuídas, i.e.,

$$F_{X_t} = F_{X_s}, \quad \forall\, s, t \in T.$$

Exemplo 2.2.1 (Passeio aleatório simétrico)**.** Considere a sucessão $(X_n)_{n\geq 1}$ de variáveis aleatórias **i**ndependentes e **i**denticamente **d**istribuídas (iid) com distribuição

$$\mathbb{P}(X_n = -1) = \mathbb{P}(X_n = 1) = \frac{1}{2}.$$

Seja

$$S_n = X_1 + \cdots + X_n.$$

O processo estocástico de tempo discreto $S = \{S_n : n \in \mathbb{N}\}$ é designado por **passeio aleatório simétrico**. Ver Figura 2.2. A variável aleatória S_n por ser interpretada como o deslocamento até ao instante n, podendo ser escrita

$$S_n = S_{n-1} + X_n\,.$$

Definição 2.2.2. Dois processos estocásticos $X = \{X_t : t \in T\}$ e $Y = \{Y_t : t \in T\}$ dizem-se **identicamente distribuídos** sse tiverem a mesma família de funções de distribuição de probabilidade conjuntas, ou seja,

$$F_{X_{t_1},\ldots,X_{t_n}} = F_{Y_{t_1},\ldots,Y_{t_n}}$$

para todo o conjunto finito de índices $t_1, \ldots, t_n \in T$.

Definição 2.2.3. Dois processos estocásticos $X = \{X_t : t \in T\}$ e $Y = \{Y_t : t \in T\}$ dizem-se **independentes** sse os vectores aleatórios

$$(X_{t_1}, \ldots, X_{t_n}) \quad \text{e} \quad (Y_{t_1}, \ldots, Y_{t_n})$$

são independentes para quaisquer $t_1, \ldots, t_n \in T$ e $n \in \mathbb{N}$.

2.3 Estacionariedade

Em termos gerais, um processo estacionário é um processo cujas características de aleatoriedade não se alteram ao longo do tempo. Existem diversas definições de estacionariedade.

Definição 2.3.1. Um processo X_t diz-se **estacionário em média** sse

$$E(X_t) = \mu \,, \quad \forall\, t \in T \,.$$

Num processo estacionário em média o seu valor esperado não evolui ao longo do tempo.

Definição 2.3.2. Um processo X_t diz-se de **covariâncias estacionárias** sse

1. Os momentos de segunda ordem são finitos, isto é
$$E(X_t^2) < \infty \,, \quad \forall t \in T \,,$$
2. Existe uma função $\gamma : T \to \mathbb{R}$ tal que
$$\operatorname{Cov}(X_s, X_t) = \gamma(t-s) \,, \quad \forall s \leq t \,.$$

Observação 2.3.1. Num processo X_t com covariâncias estacionárias tem-se

$$\operatorname{Var}(X_t) = \gamma(0) \,.$$

Ou seja, a variância do processo não se altera ao longo do tempo.

Definição 2.3.3. Um processo X_t diz-se **estacionário até à segunda ordem** sse for estacionário em média e tiver covariâncias estacionárias.

Uma definição de estacionariedade mais forte que as anteriores é a seguinte.

Definição 2.3.4. Um processo X_t diz-se **fortemente ou estritamente estacionário de ordem** $k \in \mathbb{N}$ sse para quaisquer instantes $t_1, \ldots, t_k \in T$ e qualquer h tal que $t_i + h \in T$, $i = 1, \ldots, k$, os vectores aleatórios $(X_{t_1}, \ldots, X_{t_k})$ e $(X_{t_1+h}, \ldots, X_{t_k+h})$ são identicamente distribuídos.

Chegamos assim à definição de estacionariedade mais forte.

Definição 2.3.5. Um processo é **fortemente ou estritamente estacionário** sse for fortemente estacionário de ordem k para qualquer $k \in \mathbb{N}$.

Esta definição de estacionariedade implica todas as outras. Adicionalmente, num processo fortemente estacionário todas as variáveis aleatórias X_t tem a mesma distribuição. Analogamente, todos os pares (X_t, X_s) são identicamente distribuídos desde os instantes estejam igualmente espaçados no tempo.

Exemplo 2.3.2 (Processo de Bernoulli)**.** Sejam $(X_n)_{n\geq 1}$ variáveis aleatórias iid com distribuição

$$\mathbb{P}(X_n = 1) = p, \quad \mathbb{P}(X_n = 0) = 1 - p$$

onde $p \in [0, 1]$. O processo de Bernoulli é fortemente estacionário.

Exemplo 2.3.3. Qualquer sucessão de variáveis aleatórias iid é um processo fortemente estacionário.

Exemplo 2.3.4. Considere um processo estocástico $\{X_n : n = 1, 2, \ldots\}$ tal que $E(X_n) = 0$ e $\mathrm{Var}(X_n) = \sigma^2$ para todo $n \geq 1$ e $\mathrm{Cov}(X_n, X_m) = 0$ para $n \neq m$. Este processo é estacionário até à segunda ordem, sendo usualmente designado por **ruído branco**.

Exercício 4. Seja $\Omega = \{0, 1, 2\}$ e considere a família de variáveis aleatórias $X_t : \Omega \to \mathbb{R}$,

$$X_t(\omega) = \omega t, \quad t \geq 0.$$

1. Classifique o processo estocástico $\{X_t \colon t \geq 0\}$ indicando o conjunto dos parâmetros T e o conjuntos dos estados E.

2. Determine todas as realizações do processo.

3. Calcule $\mathbb{P}(X_1 = 1, X_2 = 1)$ e $\mathbb{P}(X_2 = 2 | X_1 = 1)$.

Exercício 5. Seja $(\xi_n)_{n\geq 1}$ uma sucessão de variáveis aleatórias de Bernoulli independentes e

$$X_n = \min\{k \geq 1 \colon \xi_1 + \cdots + \xi_k = n\}, \quad n \geq 1.$$

1. Supondo que ξ_n representa o resultado (sucesso ou insucesso) de uma experiência aleatória, descreva o significado da variável aleatória X_n. Que valores pode tomar?

2. Classifique o processo estocástico $\{X_n : n \geq 1\}$ indicative o conjunto dos parâmetros T e o conjunto dos estados E.

3. Determine a distribuição de probabilidade de X_n.

4. Calcule $\mathbb{P}(X_3 = x_3 | X_2 = x_2)$.

Exercício 6. Seja $(X_n)_{n\geq 1}$ uma sucessão de variáveis aleatórias iid, $\theta \in \mathbb{R}$ e $Y_n = X_n - \theta X_{n-1}$. Mostre que o processo estocástico $Y = \{Y_n : n \in \mathbb{N}\}$, usualmente designado por processo de médias móveis de primeira ordem MA(1), é estacionário até à segunda ordem.

Chapter 3

Cadeias de Markov

Seja S um conjunto contável (finito ou infinito) e $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade. É frequente tomar $S = \{0, 1, 2, \ldots\}$.

Definição 3.0.1. Uma sucessão de variáveis aleatórias $(X_n)_{n\geq 0}$ é uma **cadeia de Markov com valores em** S se

1. $X_n(\Omega) \subseteq S$ para todo $n \geq 0$.

2. Para quaisquer $j, i_0, \ldots, i_n \in S$ e $n \geq 0$,

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n).$$

Note-se que as variáveis aleatórias X_n são discretas e as probabilidades condicionadas são calculadas de acordo com as expressões da secção anterior.

A equação da definição anterior é conhecida como a **propriedade de Markov**. Uma cadeia de Markov não tem memória do passado. A probabilidade de um estado futuro depende exclusivamente do estado presente.

3.1 Probabilidades de transição

A probabilidade de X_{n+1} estar no estado j dado que X_n está no estado i é conhecida como a probabilidade de transição a 1 passo,

$$P^n_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i), \quad i, j \in S.$$

Definição 3.1.1. Uma cadeia de Markov diz-se homogénea se $P^n_{i,j}$ não depende de n.

Vamos apenas considerar cadeias de Markov homogéneas. Neste caso escrevemos $P_{i,j}$ para denotar as probabilidades de transição e definimos a **matriz de probabilidades de transição**:

$$P = (P_{i,j})_{i,j \in S}.$$

Proposição 3.1.1. *A matriz de probabilidades de transição P satisfaz as seguintes propriedades:*

1. *$P_{i,j} \geq 0$ para quaisquer $i, j \in S$,*
2. *A soma das linhas de P é igual a 1, ou seja,*

$$\sum_{j \in S} P_{i,j} = 1.$$

Demonstração. A propriedade (1) é óbvia. Para demonstrar (2) note-se que

$$\begin{aligned}
\sum_{j \in S} P_{i,j} &= \sum_{j \in S} \mathbb{P}(X_{n+1} = j | X_n = i) \\
&= \frac{1}{\mathbb{P}(X_n = i)} \sum_{j \in S} \mathbb{P}(X_{n+1} = j, X_n = i) \\
&= 1
\end{aligned}$$

□

Uma matriz que satisfaz (1) e (2) da proposição anterior diz-se **estocástica**.

Suponhamos que $p_{i_0} = \mathbb{P}(X_0 = i_0)$. Dados estados $i_1, \ldots, i_n \in S$, como calcular as distribuições conjuntas

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) \quad ?$$

Usando a definição de probabilidade condicionada,

$$\begin{aligned}
\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_n = i_n | X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) \\
\times \mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}).
\end{aligned}$$

Pela propriedade de Markov,

$$\mathbb{P}(X_n = i_n | X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) = \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}) = P_{i_{n-1}, i_n}.$$

Logo,

$$\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) = P_{i_{n-1}, i_n} \mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}).$$

Iterando esta equação obtemos a seguinte proposição.

Proposição 3.1.2.

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = p_{i_0} P_{i_1,i_2} \cdots P_{i_{n-1},i_n}.$$

Esta proposição demonstra que todas as distribuição finitas da cadeia de Markov são determinadas pelas probabilidades de transição da matriz P.

Exemplo 3.1.3. Considere uma cadeia de Markov com três estados $S = \{0, 1, 2\}$ e com a seguinte matriz de transição

$$P = \begin{pmatrix} 1/4 & 1/4 & 1/2 \\ 9/10 & 1/10 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}.$$

Considere a distribuição inicial $\pi_0 = (p_0, p_1, p_2)$ onde

$$p_0 = \mathbb{P}(X_0 = 0) = 3/10, \quad p_1 = \mathbb{P}(X_0 = 1) = 2/5 \quad \text{e} \quad p_2 = \mathbb{P}(X_0 = 2) = 3/10.$$

Logo,

$$\mathbb{P}(X_0 = 0, X_1 = 1, X_2 = 2) = p_0 P_{0,1} P_{1,2} = 0,$$

e

$$\mathbb{P}(X_1 = 1, X_2 = 1 | X_0 = 2) = P_{2,1} P_{1,1} = 1/30.$$

Exemplo 3.1.4 (Passeio Aleatório)**.** Considere a sucessão $(\xi_n)_{n\geq 0}$ de variáveis aleatórias iid com distribuição

$$\mathbb{P}(\xi_n = 1) = p \quad \text{e} \quad \mathbb{P}(\xi_n = -1) = 1 - p,$$

onde $0 < p < 1$. Seja

$$X_n = \xi_0 + \cdots + \xi_n.$$

O processo estocástico de tempo discreto $X = \{X_n \colon n \in \mathbb{N}\}$ é designado por **passeio aleatório**. Vamos mostrar que $(X_n)_{n\geq 0}$ é uma cadeia de Markov homogénea. Em primeiro lugar, o espaço dos estados é $S = \mathbb{Z}$. Verifiquemos agora a propriedade de Markov:

$$\begin{aligned}
\mathbb{P}(X_{n+1} = j | X_0 = i_0, \ldots, X_n = i_n) &= \mathbb{P}(X_n + \xi_{n+1} = j | X_0 = i_0, \ldots, X_n = i_n) \\
&= \mathbb{P}(\xi_{n+1} = j - i_n | X_0 = i_0, \ldots, X_n = i_n) \\
&= \mathbb{P}(\xi_{n+1} = j - i_n) \,.
\end{aligned}$$

Por outro lado,

$$\begin{aligned}
\mathbb{P}(X_{n+1} = j | X_n = i_n) &= \mathbb{P}(X_n + \xi_{n+1} = j | X_n = i_n) \\
&= \mathbb{P}(\xi_{n+1} = j - i_n | X_n = i_n) \\
&= \mathbb{P}(\xi_{n+1} = j - i_n) \,.
\end{aligned}$$

Logo,

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n).$$

Assim, temos as seguintes probabilidades de transição:

$$P_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} p & \text{se } j = i + 1 \\ 1 - p & \text{se } j = i - 1 \\ 0 & \text{caso contrário} \end{cases}$$

Exemplo 3.1.5 (Incrementos iid)**.** Generalizando o exemplo anterior, é possível demonstrar que qualquer sucessão de variáveis aleatórias discretas $(X_n)_{n \geq 0}$ com incrementos iid, isto é:

1. As variáveis aleatórias $X_{n+1} - X_n$, $n \geq 0$ são independentes;
2. As variáveis aleatórias $X_{n+1} - X_n$ têm todas a mesma distribuição de probabilidade;

então $(X_n)_{n \geq 0}$ é uma cadeia de Markov homogénea.

Importante no estudo das cadeias de Markov são as probabilidades de transição a n passos,

$$P_{i,j}^{(n)} = \mathbb{P}(X_n = j | X_0 = i).$$

Proposição 3.1.6 (Equação de Chapman-Kolmogorov)**.**

$$P_{i,j}^{(n)} = \sum_{k \in S} P_{i,k} P_{k,j}^{(n-1)},$$

onde

$$P_{i,j}^{(0)} = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

Demonstração.

$$
\begin{aligned}
P_{i,j}^{(n)} &= \mathbb{P}(X_n = j | X_0 = i) \\
&= \sum_{k\in S} \mathbb{P}(X_n = j, X_1 = k | X_0 = i) \\
&= \sum_{k\in S} \mathbb{P}(X_n = j | X_1 = k, X_0 = i)\mathbb{P}(X_1 = k | X_0 = i) \\
&= \sum_{k\in S} \mathbb{P}(X_n = j | X_1 = k)\mathbb{P}(X_1 = k | X_0 = i) \\
&= \sum_{k\in S} \mathbb{P}(X_{n-1} = j | X_0 = k)\mathbb{P}(X_1 = k | X_0 = i) \\
&= \sum_{k\in S} P_{i,k} P_{k,j}^{(n-1)}
\end{aligned}
$$

□

Definindo a **matriz de transição a n passos**

$$P^{(n)} = (P_{i,j}^{(n)})_{i,j\in S}$$

as equação de Chapman-Kolmogorov pode ser escrita de forma matricial,

$$P^{(n)} = P \cdots P$$

como o produto de n cópias da matriz P.

Proposição 3.1.7. *Suponha que X_0 tem distribuição $\pi_0 = (\pi_{0,i})_{j\in S}$ onde $\pi_{0,i} = \mathbb{P}(X_0 = i)$. Então a distribuição π_n de X_n é*

$$\pi_n = \pi_0 P^{(n)}$$

Demonstração.

$$\mathbb{P}(X_n = j) = \sum_i \mathbb{P}(X_n = j | X_0 = i)\mathbb{P}(X_0 = i) = \sum_i \pi_{0,i} P_{i,j}^{(n)}$$

Logo,

$$\pi_n = \pi_0 P^{(n)}$$

onde $\pi_n = (\pi_{n,j})_{j\in S}$ e $\pi_{n,j} = \mathbb{P}(X_n = j)$. □

Exemplo 3.1.8. Usando a matriz de transição do Exemplo 3.1.3 podemos calcular, usando a Proposição 3.1.7, a distribuição de X_1 sabendo a distribuição de X_0. Suponhamos que a distribuição de X_0 é

$$\pi_0 = \begin{pmatrix} 4/9 & 5/9 & 0 \end{pmatrix}.$$

Então X_1 tem distribuição,

$$\begin{aligned}\pi_1 = \pi_0 P &= \begin{pmatrix} 4/9 & 5/9 & 0 \end{pmatrix} \begin{pmatrix} 1/4 & 1/4 & 1/2 \\ 9/10 & 1/10 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \\ &= \begin{pmatrix} 11/18 & 1/6 & 2/9 \end{pmatrix}\end{aligned}$$

Exemplo 3.1.9. Considere a seguinte matriz estocástica

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix},$$

associada a uma cadeia de Markov com dois estados $S = \{0, 1\}$. Então

$$P^2 = \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix} \quad \text{e} \quad P^3 = \begin{pmatrix} 5/8 & 3/8 \\ 3/4 & 1/4 \end{pmatrix}.$$

Note-se que $\mathbb{P}(X_1 = 1|X_0 = 1) = 0$, mas $\mathbb{P}(X_n = 1|X_0 = 1) > 0$ para $n = 2, 3$. Pode-se calcular, por exemplo diagonalizando a matriz P, a potência

$$P^n = \begin{pmatrix} \frac{2}{3} - \frac{1}{3}(-1)^{n+1}2^{-n} & \frac{1}{3} + \frac{1}{3}(-1)^{n+1}2^{-n} \\ \frac{2}{3} - \frac{1}{3}(-1)^{n}2^{1-n} & \frac{1}{3} + \frac{1}{3}(-1)^{n}2^{1-n} \end{pmatrix}.$$

Concluímos que

$$\lim_{n\to+\infty} \mathbb{P}(X_n = 0|X_0 = 0) = \frac{2}{3} \quad \text{e} \quad \lim_{n\to+\infty} \mathbb{P}(X_n = 1|X_0 = 0) = \frac{1}{3}.$$

Exercício 7. Seja $0 < p < 1$. Considere $X_0 = 0$ e

$$X_n = X_{n-1} + \xi_n, \quad n = 1, 2, \ldots$$

onde $(\xi_n)_{n\geq 1}$ é uma sucessão iid de variáveis aleatórias com distribuição

$$\mathbb{P}(\xi_n = 1) = p, \quad \mathbb{P}(\xi_n = 0) = 1 - p.$$

1. Mostre que $(X_n)_{n\geq 0}$ é uma cadeia de Markov.

2. Determine a matriz de transição da cadeia.

3. Desenhe o diagrama da cadeia.

4. Calcule $\mathbb{P}(X_2 = 1)$.

5. Mostre usando indução que

$$\mathbb{P}(X_n = j|X_0 = 0) = C_j^n p^j (1-p)^{n-j}.$$

Exemplo 3.1.10 (Aplicação: Modelo de aprovisionamento)**.** Considere um determinado artigo que é aprovisionado de modo a satisfazer determinada procura. Assumimos que a reposição do artigo é realizado no final de cada período $n = 0, 1, 2, \ldots$, e que a procura total do produto em cada período é uma variável aleatória ξ_n com distribuição de probabilidade,

$$\mathbb{P}(\xi_n = k) = a_k, \quad k = 0, 1, 2, \ldots.$$

Claro que $a_k \geq 0$ e $\sum_k a_k = 1$. Assumimos também que as procuras nos diversos períodos são independentes, ou seja, as variáveis ξ_n são independentes.

O nível de stock do artigo é examinado no final de cada período. Suponhamos que o nível de stock máximo é $m > 0$. A reposição do artigo é feita de acordo com um valor crítico não-negativo $s < m$. Seja X_n a quantidade de artigos no final do período n. Então a reposição segue a regra,

$$X_{n+1} = \begin{cases} X_n - \xi_{n+1} & \text{se } s < X_n \\ m - \xi_{n+1} & \text{se } X_n \leq s \end{cases}.$$

Note-se que os estados do processo são

$$S = \{\ldots, -3, -2, -1, 0, 1, \ldots, m-1, m\}.$$

É fácil verificar que $(X_n)_{n\geq 0}$ satisfaz a propriedade de Markov. Portanto, $(X_n)_{n\geq 0}$ é uma cadeia de Markov em S. As probabilidades de transição podem ser calculas explicitamente,

$$\begin{aligned} P_{i,j} = \mathbb{P}(X_{n+1} = j|X_n = i) &= \begin{cases} \mathbb{P}(X_n - \xi_{n+1} = j|X_n = i) & \text{se } X_n > s \\ \mathbb{P}(m - \xi_{n+1} = j|X_n = i) & \text{se } X_n \leq s \end{cases} \\ &= \begin{cases} \mathbb{P}(\xi_{n+1} = i - j) & \text{se } i > s \\ \mathbb{P}(\xi_{n+1} = m - j) & \text{se } i \leq s \end{cases} \\ &= \begin{cases} a_{i-j} & \text{se } i > s \\ a_{m-j} & \text{se } i \leq s \end{cases} \end{aligned}$$

Quantidades relevantes num modelo de aprovisionamento são a probabilidade de a procura exceder a oferta no longo prazo,

$$\lim_{n\to\infty}\sum_{j<0}\mathbb{P}(X_n=j),$$

e o nível médio do stock no longo prazo,

$$\lim_{n\to\infty}\sum_{j>0}j\mathbb{P}(X_n=j).$$

Estas questões mostram que é importante estudar o comportamento assimptótico da distribuição de X_n.

Exemplo 3.1.11 (Aplicação: Filas de espera)**.** Considere um serviço onde os clientes chegam e tomam o seu lugar numa fila de espera. Durante cada período $n = 0, 1, 2, \ldots$, um único cliente é servido, desde que haja clientes na fila de espera para serem servidos. Durante cada período em que é realizado um serviço novos clientes chegam e tomam o seu lugar na fila de espera. Suponhamos que o número de clientes que chega no período n é uma variável aleatória com distribuição,

$$\mathbb{P}(\xi_n = k) = a_k, \quad k = 0, 1, 2, \ldots.$$

Claro que $a_k \geq 0$ e $\sum_k a_k = 1$. Supomos também que as variáveis ξ_n são independentes. Se X_n é o número de clientes na fila de espera no início do período n, então

$$X_{n+1} = \begin{cases} X_n - 1 + \xi_n & \text{se } X_n \geq 1 \\ \xi_n & \text{se } X_n = 0 \end{cases}.$$

É fácil verificar que $(X_n)_{n\geq 0}$ é uma cadeia de Markov no espaço de estados $S = \{0, 1, 2, \ldots\}$. As probabilidade de transição podem ser calcular explicitamente,

$$\begin{aligned} P_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i) &= \begin{cases} \mathbb{P}(X_n - 1 + \xi_n = j | X_n = i) & \text{se } X_n \geq 1 \\ \mathbb{P}(\xi_n = j | X_n = i) & \text{se } X_n = 0 \end{cases} \\ &= \begin{cases} \mathbb{P}(\xi_n = j + 1 - i) & \text{se } i \geq 1 \\ \mathbb{P}(\xi_n = j) & \text{se } i = 0 \end{cases} \\ &= \begin{cases} a_{j+1-i} & \text{se } i \geq 1 \\ a_j & \text{se } i = 0 \end{cases} \end{aligned}$$

Colocando as probabilidades de transição numa matriz estocástica obtemos,

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & a_0 & a_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

É intuitivamente claro que se o valor esperado do número de novos clientes que chegam à fila de espera for maior que 1, isto é,

$$E(\xi_n) = \sum_{k=0}^{\infty} k a_k > 1,$$

então com o passar do tempo o número de clientes na fila de espera cresce sem limite. No entanto, se $E(\xi_n) < 1$, então as probabilidades de transição a n passos convergem para uma distribuição limite,

$$\lim_{n\to\infty} \mathbb{P}(X_n = k | X_0 = j) = \pi_k > 0, \quad k = 0, 1, 2, \dots$$

independente de j. Quantidades relevantes no estudo de filas de espera são a probabilidade de a longo prazo não haver clientes na fila de espera, π_0, e o tempo médio, a longo prazo, que um cliente espera na fila,

$$\sum_{k=0}^{\infty} (1+k)\pi_k.$$

Exercício 8. Uma cadeia de Markov $(X_n)_{n\geq 0}$ com estados $S = \{0, 1, 2\}$ tem a seguinte matriz de transição

$$P = \begin{pmatrix} 0.1 & 0.1 & 0.8 \\ 0.2 & 0.2 & 0.6 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}.$$

Calcule:

1. $\mathbb{P}(X_3 = 1 | X_1 = 0)$
2. $\mathbb{P}(X_2 = 1 | X_0 = 0)$
3. $\mathbb{P}(X_3 = 1, X_2 = 1, X_1 = 0)$

Exercício 9. Considere um modelo de aprovisionamento (descrito no Exemplo 3.1.10) onde apenas 0, 1, ou 2 artigos são procurados em cada período ($a_k = 0$ para $k \geq 3$) com probabilidade

$$\mathbb{P}(\xi_n = 0) = \frac{1}{2}, \quad \mathbb{P}(\xi_n = 1) = \frac{2}{5} \quad \text{e} \quad \mathbb{P}(\xi_n = 2) = \frac{1}{10}.$$

Suponha que $s = 0$ e $m = 2$.

1. Determine a matriz de transição de probabilidades P para a cadeia de Markov $(X_n)_{n\geq 0}$ onde X_n é o número de artigos em stock no fim do período n. (Dica: $S = \{2, 1, 0, -1\}$.)

2. Calcule $\mathbb{P}(X_2 = 1, X_1 = 2|X_0 = 0)$.

3. Calcule $\mathbb{P}(X_3 = 0|X_1 = 1)$.

3.2 Recorrência

Seja $(X_n)_{n\geq 0}$ uma cadeia de Markov homogénea com espaço de estados S.

Definição 3.2.1. Um estado $i \in S$ diz-se **recorrente** se a cadeia X_n eventualmente regressa a i partindo de i, ou seja,

$$r_i := \mathbb{P}(X_n = i \text{ para algum } n \geq 1|X_0 = i) = 1\,.$$

Um estado não recorrente diz-se **transiente**. Note-se que r_i é a probabilidade de a cadeia **reentrar** no estado i algures no futuro.

Considere a seguinte variável aleatória

$$T_i = \min\{n \geq 1 \colon X_n = i\}.$$

A variável aleatória $T_i : \Omega \to \mathbb{N} \cup \{+\infty\}$ devolve o tempo que a cadeia demora até visitar o estado i independentemente do estado de partida. A variável T_i é conhecida por **tempo de paragem** ou **primeiro tempo de retorno a** i. Podemos calcular a sua distribuição de probabilidade:

$$\mathbb{P}(T_i = n) = \mathbb{P}(X_1 \neq i, \ldots, X_{n-1} \neq i, X_n = i)$$

e

$$\mathbb{P}(T_i = +\infty) = \mathbb{P}(X_1 \neq i, X_2 \neq i, X_3 \neq i, \ldots).$$

Da igualdade dos eventos,

$$\{X_n = i \text{ para algum } n \geq 1\} = \bigcup_{n\geq 1} \{X_1 \neq i, \ldots, X_{n-1} \neq i, X_n = i\},$$

concluímos que a probabilidade de a cadeia reentrar no estado i é dada por

$$r_i = \sum_{n=1}^{\infty} \mathbb{P}(T_i = n | X_0 = i).$$

Assim, obtemos o seguinte critério para a recorrência.

Proposição 3.2.1. *Um estado $i \in S$ é recorrente sse*

$$\mathbb{P}(T_i < +\infty | X_0 = i) = 1.$$

Portanto, um estado $i \in S$ é recorrente se a probabilidade de a cadeia demorar tempo infinito para regressar a i é igual a zero, ou seja,

$$\mathbb{P}(T_i = +\infty | X_0 = i) = 0.$$

Na prática há um critério mais simples para determinar se um estado é recorrente ou transiente. Esse critério, baseado nas probabilidades de transição $P_{i,i}^{(n)}$ a n passos, é obtido da seguinte forma.

Seja N_i a variável aleatória que conta o número de visitas ao estado i, ou seja,

$$N_i = \sum_{n=0}^{\infty} \chi_{\{X_n = i\}} = \chi_{\{X_0 = i\}} + \chi_{\{X_1 = i\}} + \cdots$$

onde $\chi_{\{X_n = i\}}$ é uma variável aleatória (função indicadora) que toma valor 1 se $X_n = i$ ou valor 0 caso contrário.

Lemma 3.2.2 (Valor médio do número de reentradas em i)**.**

$$E(N_i | X_0 = i) = \frac{1}{1 - r_i}.$$

Demonstração. A probabilidade de a cadeia reentrar exactamente n vezes no estado i é

$$\mathbb{P}(N_i = n | X_0 = i) = r_i^{n-1}(1 - r_i).$$

Logo,

$$\begin{aligned}
E(N_i|X_0 = i) &= \sum_{n=1}^{\infty} n\mathbb{P}(N_i = n|X_0 = i) \\
&= \sum_{n=1}^{\infty} n r_i^{n-1}(1 - r_i) \\
&= (1 - r_i) \sum_{n=1}^{\infty} n r_i^{n-1} \\
&= (1 - r_i) \frac{1}{(1 - r_i)^2} \\
&= \frac{1}{1 - r_i}.
\end{aligned}$$

A soma da série anterior pode ser obtida derivando em ordem a x ambos os lados da igualdade,

$$\frac{1}{1 - x} = \sum_{n=0}^{\infty} x^n, \quad |x| < 1.$$

□

Usando o lemma anterior obtemos o seguinte critério para determinar se um estado é recorrente ou transiente.

Teorema 3.2.3. *Um estado $i \in S$ é recorrente sse*

$$\sum_{n=1}^{\infty} P_{i,i}^{(n)} = +\infty.$$

Demonstração. Note-se que $E(\chi_{\{X_n=i\}}) = \mathbb{P}(X_n = i)$. Logo, segue da definição da variável N_i que

$$\begin{aligned}
E(N_i|X_0 = i) &= \sum_{n=0}^{\infty} E(\chi_{\{X_n=i\}}|X_0 = i) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(X_n = i|X_0 = i) \\
&= \sum_{n=0}^{\infty} P_{i,i}^{(n)}.
\end{aligned}$$

Portanto, do lemma anterior concluímos que

$$\sum_{n=0}^{\infty} P_{i,i}^{(n)} = \frac{1}{1 - r_i}.$$

Assim, $r_1 = 1$ (estado i recorrente) se e só se $\sum_n P_{i,i}^{(n)} = +\infty$. $\square$

Note-se que $i \in S$ é transiente se e só se

$$E(N_i|X_0 = i) = \sum_{n=1}^{\infty} P_{i,i}^{(n)} < +\infty.$$

Portanto, se $i \in S$ é transiente, então $\lim_{n\to+\infty} P_{i,i}^{(n)} = 0$. Por outro lado, o valor médio do número de reentradas em i é finito, portanto

$$\mathbb{P}(X_n = i \text{ para uma infinidade de } n \geq 1|X_0 = i) = 0.$$

Concluímos que uma cadeia de Markov entra apenas um número finito de vezes em qualquer estado transiente[1].

Usando os mesmos argumentos da prova do Teorema 3.2.3 obtemos a seguinte condição necessária para um estado ser transiente (ou suficiente para ser recorrente).

Corolário 3.2.4. *Se $j \in S$ é transiente, então*

$$\sum_{n=1}^{\infty} P_{i,j}^{(n)} < +\infty, \quad \forall\, i \in S.$$

Demonstração. Por um lado temos,

$$E(N_j|X_0 = i) = \sum_{n=1}^{\infty} P_{i,j}^{(n)}.$$

Por outro lado,

$$E(N_j|X_0 = i) \leq \frac{1}{1 - r_j}.$$

Logo, se $r_j < 1$ (j transiente), então $\sum_n P_{i,j}^{(n)} < +\infty$. $\square$

Exemplo 3.2.5. Considere uma cadeia de Markov com dois estados $\{0, 1\}$ e matriz de transição

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

[1] É possível mostrar que se i é recorrente então $\mathbb{P}(X_n = i \text{ para uma infinidade de } n \geq 1|X_0 = i) = 1$. No entanto, a demonstração é bastante complicada.

Então

$$P^n = \begin{cases} P, & \text{se } n \text{ é impar} \\ I, & \text{se } n \text{ é par} \end{cases}$$

onde I é a matriz identidade. Logo,

$$\sum_{n=1}^{\infty} P_{0,0}^{(n)} = +\infty \quad \text{e} \quad \sum_{n=1}^{\infty} P_{1,1}^{(n)} = +\infty.$$

Portanto, os dois estados são recorrentes.

Exemplo 3.2.6. Considere uma cadeia de Markov com dois estados $\{0, 1\}$ e matriz de transição

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}.$$

Então

$$P^n = \begin{pmatrix} \frac{1}{2^n} & 1 - \frac{1}{2^n} \\ 0 & 1 \end{pmatrix}$$

Assim,

$$\sum_{n=1}^{\infty} P_{0,0}^{(n)} = 1 \quad \text{e} \quad \sum_{n=1}^{\infty} P_{1,1}^{(n)} = +\infty.$$

Portanto, 0 é transiente e 1 é recorrente.

É fácil ver que se o espaço de estados é finito, ou seja, a cadeia de Markov contém apenas um número finito de estados, então existe sempre um estado recorrente.

Proposição 3.2.7. *Se o espaço de estados S é finito, então existe um estado $i \in S$ que é recorrente.*

Demonstração. Suponhamos por absurdo que todos os estados $j \in S$ da cadeia de Markov são transientes, ou seja, usando o Corolário 3.2.4,

$$\sum_{n=1}^{\infty} P_{i,j}^{(n)} < +\infty, \quad \forall\, i \in S.$$

Logo,

$$\sum_{j \in S} \sum_{n=1}^{\infty} P_{i,j}^{(n)} < +\infty.$$

No entando,

$$\sum_{j \in S} \sum_{n=1}^{\infty} P_{i,j}^{(n)} = \sum_{n=1}^{\infty} \sum_{j \in S} P_{i,j}^{(n)} = \sum_{n=1}^{\infty} 1 = +\infty,$$

o que é absurdo. $\square$

Exemplo 3.2.8. Considere uma cadeia de Markov com espaço de estados contável (infinito) $S = \{0, 1, 2, \ldots\}$ e matriz de transição

$$P = \begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} & a_{0,4} & \cdots \\ 0 & a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & \cdots \\ 0 & 0 & a_{2,2} & a_{2,3} & a_{2,4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

onde $0 < a_{i,j} < 1$. Uma vez que P é triangular superior, é fácil calcular as probabilidades de transição $P_{i,i}^{(n)}$ a n passos,

$$P_{i,i}^{(n)} = a_{i,i}^n, \quad i = 0, 1, 2, \ldots.$$

Assim,

$$\sum_{n=1}^{\infty} P_{i,i}^{(n)} = \sum_{n=1}^{\infty} a_{i,i}^n = \frac{a_{i,i}}{1 - a_{i,i}} < +\infty.$$

Logo, todos os estados da cadeia de Markov são transientes, ou seja, não há estados recorrentes!

Exemplo 3.2.9 (Assimétrico)**.** Considere o passeio aleatório assimétrico. Trata-se de uma cadeia de Markov com espaço de estados $S = \mathbb{Z}$ e probabilidades de transição

$$P_{i,j} = \begin{cases} p, & \text{se } j = i + 1 \\ 1 - p, & \text{se } j = i - 1 \\ 0, & \text{caso contrário} \end{cases}$$

onde $p \neq \frac{1}{2}$. É fácil calcular a probabilidade de transição $P_{i,i}^{(n)}$. No caso em que n é par, $n = 2k$, obtemos,

$$\begin{aligned} P_{i,i}^{(2k)} = \mathbb{P}(X_{2k} = i | X_0 = i) &= C_k^{2k} p^k (1-p)^k \\ &= \frac{(2k)!}{(k!)^2} p^k (1-p)^k \end{aligned}$$

e no caso em que n é ímpar obtemos $P_{i,i}^{(2k+1)} = 0$. Logo, como $p \neq \frac{1}{2}$, a série converge,

$$\sum_{n=1}^{\infty} P_{i,i}^{(n)} = \sum_{k=1}^{\infty} \frac{(2k)!}{(k!)^2} p^k (1-p)^k < +\infty$$

pelo critério d'Alembert (ou critério da razão) para séries,

$$\frac{P_{i,i}^{(2k+2)}}{P_{i,i}^{(2k)}} = \frac{(2k+2)(2k+1)}{(k+1)^2} p(1-p) \underset{k\to\infty}{\longrightarrow} 4p(1-p) < 1$$

porque $p \neq \frac{1}{2}$. Assim, concluímos que todos os estados do passeio aleatório assimétrico são transientes. O caso simétrico $p = \frac{1}{2}$ tem de ser analisado separadamente.

Os estados recorrentes dividem-se em dois tipos: os estados recorrentes com tempo finito de retorno e os estados recorrentes com tempo infinito de retorno. Os primeiros designam-se por **recorrentes positivos**.

Definição 3.2.2. Um estado recorrente $i \in S$ é **recorrente positivo** se o **tempo médio de reentrada** é finito,

$$m_i := E(T_i|X_0 = i) < +\infty.$$

Caso contrário dizemos que i é **recorrente nulo**.

Note-se que

$$m_i = \left(\sum_{n=1}^{\infty} n\mathbb{P}(T_i = n|X_0 = i)\right) + \infty \cdot \mathbb{P}(T_i = +\infty|X_0 = i).$$

Logo, se i é transiente, ou seja, $\mathbb{P}(T_i = +\infty|X_0 = i) > 0$, então $m_i = +\infty$. Portanto, se o tempo médio de reentrada é finito então o estado é recorrente.

O seguinte resultado estabelece um critério para decidir se um estado é recorrente nulo.

Teorema 3.2.10. *Um estado $i \in S$ recorrente nulo se e só se*

$$\lim_{n\to+\infty} P_{i,i}^{(n)} = 0.$$

Demonstração. Procurar na bibliografia da disciplina. □

Tal como na Proposição 3.2.7, quando o espaço de estados é finito, todos os estados recorrentes são positivos.

Proposição 3.2.11. *Se o espaço de estados é finito, então todos os estados recorrentes são positivos.*

Demonstração. Suponhamos por absurdo que existe um estado $i \in S$ tal que i é recorrente nulo. Ou seja, $\lim_{n\to\infty} P_{i,i}^{(n)} = 0$. Então para todo $j \in S$ temos que[2] $\lim_{n\to\infty} P_{i,j}^{(n)} = 0$. Mas,

$$\sum_{j\in S} P_{i,j}^{(n)} = 1,$$

[2]Este facto não é trivial, mas pode ser demonstrado com recurso à equação de Chapman-Kolmogorov e usando a noção de classes de comunicação que será introduzida posteriormente. Pode também encontrar uma demonstração na bibliografia da disciplina

contradizendo o facto de $\lim_n P_{i,j}^{(n)} = 0$ (porque S é finito). $\square$

Exemplo 3.2.12. Os estados recorrentes dos exemplos 3.2.5 e 3.2.6 são todos recorrentes positivos. No primeiro exemplo a sucessão $P_{i,i}^{(n)}$ tem dois pontos de acumulação, portanto não é convergente para zero. No segundo exemplo $P_{1,1}^{(n)} = 1$ para todo $n \geq 1$.

Exemplo 3.2.13 (Simétrico)**.** Considere o passeio aleatório simétrico. As probabilidades de transição são

$$P_{i,j} = \begin{cases} \frac{1}{2}, & \text{se } j = i+1 \\ \frac{1}{2}, & \text{se } j = i-1 \\ 0, & \text{caso contrário} \end{cases}$$

Tal como no passeio aleatório assimétrico temos,

$$P_{i,i}^{(n)} = \mathbb{P}(X_n = i | X_0 = i) = \begin{cases} \frac{n!}{((n/2)!)^2} \frac{1}{2^n} & \text{se n é par} \\ 0 & \text{se n é ímpar} \end{cases}$$

Seja $a_k = \frac{(2k)!}{(k!)^2} \frac{1}{2^{2k}}$. Queremos determinar se a série $\sum_k a_k$ converge ou diverge. O critério d'Alembert é inconclusivo porque o limite da razão $\frac{a_{k+1}}{a_k}$ é igual a 1 (ver exemplo do passeio aleatório assimétrico). Portanto, vamos usar a conhecida aproximação de Stirling[3],

$$k! \sim \sqrt{2\pi k} \left(\frac{k}{e}\right)^k$$

para estimar a_k. Assim,

$$a_k \sim \frac{\sqrt{4\pi k} \left(\frac{2k}{e}\right)^{2k}}{2\pi k \left(\frac{k}{e}\right)^{2k}} \frac{1}{2^{2k}} = \frac{1}{\sqrt{\pi k}}.$$

Logo,

$$\sum_{n=1}^{\infty} P_{i,i}^{(n)} = \sum_{k=1}^{\infty} a_k \sim \sum_{k=1}^{\infty} \frac{1}{\sqrt{\pi k}} = +\infty$$

Portanto, todos os estados do passeio aleatório simétrico são recorrentes! Mais, como

$$\lim_{n \to +\infty} P_{i,i}^{(n)} = \lim_{k \to +\infty} \frac{1}{\sqrt{\pi k}} = 0,$$

concluímos que todos os estados são recorrentes nulos. Ou seja, apesar dos estados serem recorrentes, o tempo médio de reentrada é infinito.

[3]ver wikipedia

3.3 Exercícios

Exercício 10. Considere uma cadeia de Markov homogénea com estados $S = \{0, 1\}$ e matriz de transição:

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \quad 0 < p < 1, \quad 0 < q < 1.$$

1. Mostre por indução que

$$P^n = \frac{1}{p+q}\left(\begin{pmatrix} q & p \\ q & p \end{pmatrix} + (1-p-q)^n \begin{pmatrix} p & -p \\ -q & q \end{pmatrix}\right), \quad n \geq 1.$$

2. Determine se os estados são recorrentes ou transientes.

3. Calcule
$$\pi_i = \lim_{n\to\infty} \mathbb{P}(X_n = i | X_0 = 0), \quad i \in \{0, 1\}.$$

4. Mostre que
$$\pi = \pi P$$
onde π é o vector linha formado por π_0 e π_1 calculados na alínea anterior.

Exercício 11. Uma urna contém inicialmente duas bolas, uma branca e outra preta. Uma bola é retirada ao acaso e substituída por uma bola de cor oposta. O processo é repetido um número infinito de vezes. Seja X_n o número de bolas brancas na urna no instante $n \geq 0$.

1. Mostre que X_n é uma cadeia de Markov homogénea e determine a matriz de transição P.

2. Calcule P^n.

3. Suponhamos que inicialmente a urna contém duas bolas brancas. No passo seguinte, uma bola branca é substituída por uma preta e por aí adiante. Determine a probabilidade de reencontrar ao longo do processo duas bolas brancas na urna.

3.4 Periodicidade

Definição 3.4.1. O período de um estado $i \in S$ é o máximo divisor comum dos inteiros positivos $n \in \mathbb{N}$ tais que $P_{i,i}^{(n)} > 0$, isto é,

$$Per(i) = \text{mdc}\{n \geq 1 \colon P_{i,i}^{(n)} > 0\}.$$

Dizemos que o estado i é **periódico** se $Per(i) \geq 2$. Caso contrário, dizemos que i é **aperiódico**.

Algumas observações a reter:

1. se $P_{i,i}^{(n)} > 0$ para todo $n \geq 1$, então o estado i é aperiódico;

2. se n não é um múltiplo de $Per(i)$, então $P_{i,i}^{(n)} = 0$.

Definição 3.4.2. Um estado $i \in S$ que é recorrente positivo e aperiódico diz-se que é **ergódico**.

Exemplo 3.4.1. Considere a cadeia de Markov do Exemplo 3.2.5. Temos para qualquer $i \in \{0, 1\}$ que

$$P_{i,i}^{(n)} = \begin{cases} 0 & \text{se } n \text{ é ímpar} \\ 1 & \text{se } n \text{ é par} \end{cases}$$

Logo,

$$Per(i) = \text{mdc}\{n \geq 1 \colon P_{i,i}^{(n)} > 0\} = \text{mdc}\{2, 4, 6, 8, \ldots\} = 2.$$

Portanto, todos os estados da cadeia são periódicos com período 2.

Exemplo 3.4.2. Considere o passeio aleatório simétrico do Exemplo 3.2.13. Tal como no exemplo anterior, $P_{i,i}^{(n)} > 0$ se n é par 2 e $P_{i,i}^{(n)} = 0$ se n é ímpar. Assim, todos os estados da cadeia são periódicos com período 2.

Exemplo 3.4.3. Considere a cadeia de Markov do Exercício 10. Como para todo $n \geq 1$, $P_{i,i}^{(n)} > 0$ temos que todos os estados i da cadeia são aperiódicos. Como também são recorrentes positivos então os estados da cadeia são ergódicos.

Exemplo 3.4.4. Considere a seguinte cadeia de Markov homogénea com estados $S = \{0, 1, 2, 3\}$ e matriz de transição

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

Fazendo o diagrama da cadeia de Markov é fácil verificar que $P_{0,0}^{(n)} = 1$ para todo $n \geq 1$ e para os restantes estados $i \in \{1, 2, 3\}$,

$$P_{i,i}^{(n)} = \begin{cases} 1 & \text{se } n \text{ é múltiplo de 3} \\ 0 & \text{caso contrário} \end{cases}$$

Portanto, todos os estados são recorrentes positivos. O estado 0 é aperiódico e os restantes têm período 3. Logo, o estado 0 é ergódico.

3.5 Classes de equivalência

Definição 3.5.1. Dizemos que um estado i **comunica** com um estado j se existir uma probabilidade de transitar de i para j, isto é, se existir $n \geq 0$ tal que

$$P_{i,j}^{(n)} = \mathbb{P}(X_n = j | X_0 = i) > 0.$$

Se i comunicar com j escrevemos

$$i \rightarrow j$$

Se $i \rightarrow j$ e $j \rightarrow i$ então i e j **comunicam entre si** e escrevemos

$$i \leftrightarrow j$$

Lemma 3.5.1. *A relação de intercomunicação $\leftrightarrow$ é uma relação de equivalência, isto é, satisfaz as seguintes propriedades:*

1. $i \leftrightarrow i$
2. *se* $i \leftrightarrow j$ *então* $j \leftrightarrow i$
3. *se* $i \leftrightarrow j$ *e* $j \leftrightarrow k$ *então* $i \leftrightarrow k$.

Demonstração.

1. $i \leftrightarrow i$ vem do facto que $P_{i,i}^{0} = 1 > 0$.
2. $i \leftrightarrow j$ é equivalente a $P_{i,j}^{(n)} > 0$ e $P_{j,i}^{(m)} > 0$ para algum $n, m \geq 0$. Logo $j \leftrightarrow i$.
3. Supondo que $P_{i,j}^{(n)} > 0$ $(i \rightarrow j)$ e $P_{j,k}^{(m)} > 0$ $(j \rightarrow k)$ usando a equação de Chapman-Kolmogorov,

$$P_{i,k}^{n+m} = \sum_{j \in S} P_{i,j}^{(n)} P_{j,k}^{(m)} \geq P_{i,j}^{(n)} P_{j,k}^{(m)} > 0.$$

   Logo, $i \rightarrow k$. De forma análoga mostramos que $k \rightarrow i$.

$\square$

Vamos de seguida introduzir algumas definições para particionar o espaço dos estados em subconjuntos que partilham propriedades comuns.

Definição 3.5.2. Ao conjunto $C(i)$ formado por todos os estados da cadeia que intercomunicam com o estado i designamos por **classe de comunicação de** i,

$$C(i) = \{j \in S \colon i \leftrightarrow j\}$$

Um subconjunto de estados diz-se **irredutível** se todos os estados nesse subconjunto comunicam entre si. Neste sentido, as classes de comunicação são irredutíveis. De facto, $C(i)$ é o maior conjunto irredutível contendo o estado i. Se todos os estados da cadeia comunicam entre si (existe apenas uma única classe de comunicação) então dizemos que a cadeia é **irredutível**.

Proposição 3.5.2. *Se $j \in C(i)$, então*

1. *i e j têm o mesmo período.*
2. *i é recorrente sse j é recorrente.*
3. *i é recorrente positivo sse j é recorrente positivo.*
4. *i é ergódico sse j é ergódico.*

Demonstração.

1. Procurar a prova na bibliografia da disciplina.
2. Como $i \leftrightarrow j$, temos que $P_{i,j}^{(n)} > 0$ e $P_{j,i}^{(m)} > 0$ para algum $n, m \geq 0$. Seja $\epsilon = P_{i,j}^{(n)} P_{j,i}^{(m)} > 0$. Usando sucessivamente a equação de Chapman-Kolmogorov,

$$\begin{aligned} P_{i,i}^{n+r+m} &= \sum_{k \in S} P_{i,k}^{(n)} P_{k,i}^{(r+m)} \\ &\geq P_{i,j}^{(n)} P_{j,i}^{(r+m)} \\ &= P_{i,j}^{(n)} \sum_{s \in S} P_{j,s}^{(r)} P_{s,i}^{(m)} \\ &\geq P_{i,j}^{(n)} P_{j,j}^{(r)} P_{j,i}^{(m)}. \end{aligned}$$

Assim,

$$P_{i,i}^{n+r+m} \geq \epsilon \, P_{j,j}^{(r)}.$$

Portanto, $\sum_{r=1}^{\infty} P_{j,j}^{(r)} = +\infty$ implica que $\sum_{r=1}^{\infty} P_{i,i}^{(r)} = +\infty$, ou seja, se j é recorrente, então i também é recorrente. Trocando os papeis de i e j obtemos que a propriedade de recorrência é preservada nas classe de comunicação.

3. O mesmo se verifica com a recorrência positiva usando a desigualdade anterior.

4. Segue do (1) e do (3).

□

Portanto, recorrência positiva (ou nula) e periodicidade são propriedades partilhadas por todos os estados da mesma classe de comunicação. Assim, dizemos que a cadeia é aperiódica quando todos os estados são aperiódicos, recorrente quando todos os estados são recorrentes, etc.

Definição 3.5.3. Dizemos que um subconjunto $C \subset S$ de estados é **fechado** se para todo $i \in C$ e $j \notin C$ tem-se $P_{i,j} = 0$. Ou seja, um subconjunto C de estados é fechado se a probabilidade de sair de C é zero.

Proposição 3.5.3. *Se i é recorrente, então $C(i)$ é fechado.*

Demonstração. Suponhamos por absurdo que $C(i)$ não é fechado. Então existe um estado $s \in C(i)$ recorrente e $j \notin C(i)$ tal que $P_{s,j} > 0$. Ou seja, $s \to j$ mas $j \nrightarrow s$, caso contrário j estaria em $C(i)$. Assim,

$$\begin{aligned}\mathbb{P}(T_s = +\infty | X_0 = s) &= \mathbb{P}(X_1 \neq s, X_2 \neq s, \ldots | X_0 = s) \\ &\geq \mathbb{P}(X_1 = j, X_2 \neq s, \ldots | X_0 = s) \\ &= \mathbb{P}(X_1 = j | X_0 = s) = P_{s,j} > 0,\end{aligned}$$

uma vez que $j \nrightarrow s$. Portanto, da desigualdade anterior concluímos que s é transiente. Absurdo. □

Chegamos assim a um resultado central que estabelece uma decomposição dos estados de uma cadeia de Markov homogénea.

Teorema 3.5.4 (Decomposição dos estados). *Qualquer espaço de estados de uma cadeia de Markov homogénea pode ser decomposto numa união (disjunta e contável) de estados transientes T e subconjuntos fechados e irredutíveis $C_1, C_2, \ldots$ de estados recorrentes, isto é,*

$$S = T \cup C_1 \cup C_2 \cup \cdots$$

Demonstração. Seja $R = S \setminus T$ o conjunto dos estados recorrentes da cadeia de Markov. Para cada $i \in R$, a respectiva classe de comunicação $C(i)$ é irredutível e contém apenas estados recorrentes. Segue da proposição anterior que é fechada. Assim, existem $i_1, i_2, \ldots \in R$, tal que o conjunto dos estados recorrentes R pode ser escrito como uma união disjunta de classes de comunicação $C(i_1), C(i_2), \ldots$ fechadas e irredutíveis. Portanto,

$$S = T \cup C(i_1) \cup C(i_2) \cup \cdots$$

$\square$

Observação 3.5.5. Se o espaço de estados é finito, então existe pelo menos um conjunto fechado irredutível e não-vazio C_i de estados recorrentes na decomposição do teorema anterior. Adicionalmente, pela Proposição 3.2.11, todos os estados recorrentes são positivos.

Observação 3.5.6. Se uma classe de comunicação $C(i)$ não é fechada, então segue da Proposição 3.5.3 que todos os estados da classe $C(i)$ são transientes.

Exemplo 3.5.7. Considere uma cadeia de Markov homogénea com estados $S = \{0, 1, 2, 3, 4, 5\}$ e matriz de transição

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

É fácil verificar que

$$0 \leftrightarrow 1, \quad 2 \leftrightarrow 3 \quad \text{e} \quad 4 \leftrightarrow 5.$$

Estas são as únicas relação de intercomunicação. Assim, temos três classes de comunicação

$$C(0) = \{0, 1\}, \quad C(2) = \{2, 3\} \quad e \quad C(4) = \{4, 5\}.$$

Para determinar os estados recorrentes, basta verificar se as classes de comunicação são fechadas ou não. Como as classes $C(0)$ e $C(4)$ são fechadas, contêm apenas estados recorrentes. A classe $C(2)$ não é fechada , logo é formada por estados transientes. Quanto à periodicidade, podemos concluir que todos os estados $i \in S$ são aperiódicos ($Per(i) = 1$) porque $P_{i,i} > 0$, logo $P_{i,i}^{(n)} > 0$ para todo $n \geq 1$. Relativamente à distribuição dos tempos de paragem T_i, podemos calcular sem dificuldade,

$$\mathbb{P}(T_0 = n | X_0 = 0) = \begin{cases} \frac{1}{2} & \text{se } n = 1 \\ P_{0,1} P_{1,1}^{n-2} P_{1,0} & \text{se } n \geq 2 \end{cases}$$

Expressões equivalentes podem ser obtidas para as distribuições dos restantes tempos de paragem. Logo,

$$
\begin{aligned}
m_0 = E(T_0|X_0 = 0) &= \sum_{n=1}^{\infty} n\mathbb{P}(T_0 = n|X_0 = 0) \\
&= \frac{1}{2} + \sum_{n=2}^{\infty} n\frac{1}{2}\left(\frac{3}{4}\right)^{n-2}\frac{1}{4} \\
&= \frac{1}{2} + \frac{1}{6}\left(\sum_{n=1} n\left(\frac{3}{4}\right)^{n-1} - 1\right) \\
&= \frac{1}{2} + \frac{1}{6}\left(\frac{1}{(1-\frac{3}{4})^2} - 1\right) \\
&= 3
\end{aligned}
$$

Portanto, o estado 0 é recorrente positivo. Um facto que já sabíamos pela Proposição 3.2.11.

3.6 Exercícios

Exercício 12. Construa uma cadeia de Markov homogénea onde $P_{i,i}^{(Per(i))} = 0$ para algum estado i da cadeia. (Dica: considere a cadeia de Markov do Exercício 10 e escolha p e q adequadamente)

Exercício 13. Mostre que se $P_{i,i} > 0$, então o estado i é aperiódico.

Exercício 14. Mostre que se $i \leftrightarrow j$, então $Per(i) = Per(j)$.

Exercício 15. É possível construir uma cadeia de Markov homogénea com espaço de estados finito e um estado recorrente nulo? e todos os estados transientes? Justifique.

Exercício 16. Determine a decomposição dos estados $S = \{0, 1, 2, 3, 4, 5\}$ da cadeia de Markov homogénea com a seguinte matriz de transição e calcule $m_i = E(T_i|X_0 = i)$ para todo $i \in S$,

$$
P = \begin{pmatrix}
\frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\
\frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
$$

Exercício 17. Determine e classifique todas as classes de comunicação e período de cada classe para a cadeia de Markov homogénea com estados $S = \{0, 1, 2, 3, 4, 5\}$ e matriz de transição:

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \end{pmatrix}$$

Calcule $m_5 = E(T_5|X_0 = 5)$.

3.7 Distribuição estacionária

Seja $(X_n)_{n\geq 0}$ uma cadeia de Markov estacionária com espaço de estados (finito ou infinito) S e matriz de transição P.

Dada uma distribuição de probabilidade inicial π_0 de X_0, usando a Proposição 3.1.7 a distribuição de X_n é dada por $\pi_n = \pi_0 P^n$. Um caso especial é quando $\pi_n = \pi_0$ para todo $n \geq 0$.

Definição 3.7.1. Uma distribuição de probabilidade[4] π em S diz-se **estacionária** se

$$\pi = \pi P.$$

Observação 3.7.1. A equação $\pi = \pi P$ determina os vectores próprios associados ao valor próprio 1 da matriz transposta de P. Portanto, se P não tiver valor próprio[5] 1, então a cadeia não tem distribuição estacionária.

Observação 3.7.2. Se a cadeia de Markov tem duas distribuições estacionárias distintas, então tem uma infinidade de distribuições estacionárias. De facto, sejam π e ν duas distribuições estacionárias. Então a combinação convexa

$$\mu = \lambda\pi + (1-\lambda)\nu, \quad \lambda \in [0, 1]$$

é também uma distribuição estacionária:

$$\mu P = \lambda\pi P + (1-\lambda)\nu P = \lambda\pi + (1-\lambda)\nu = \mu$$

e

$$\sum_{i\in S} \mu_i = \sum_{i\in S} \lambda\pi_i + (1-\lambda)\nu_i = \lambda \sum_{i\in S} \pi_i + (1-\lambda) \sum_{i\in S} \nu_i = 1.$$

[4]Uma distribuição de probabilidade é um vector linha $\pi = (\pi_i : i \in S)$ tal que $\sum_{i\in S} \pi_i = 1$ e $\pi_i \geq 0$ para todo $i \in S$.

[5]A matriz P e a sua transposta têm os mesmos valores próprios.

No caso em que a cadeia de Markov admite uma distribuição estacionária π, se X_0 tiver distribuição π então todas as variáveis X_n têm a mesma distribuição π. Segue que a cadeia de Markov é um **processo estacionário**, ou seja,

$$(X_{n_1}, \ldots, X_{n_k}) \stackrel{d}{=} (X_{n_1+m}, \ldots, X_{n_k+m})$$

para quaisquer $0 \leq n_1 < \cdots < n_k$ e $m \geq 0$. De facto,

$$\begin{aligned}\mathbb{P}(X_{n_1+m} = i_1, \ldots, X_{n_k+m} = i_k) &= \pi_{i_1} P_{i_1,i_2}^{(n_2-n_1)} \cdots P_{i_{k-1},i_k}^{(n_k-n_{k-1})} \\ &= \mathbb{P}(X_{n_1} = i_1, \ldots, X_{n_k} = i_k)\end{aligned}$$

Exemplo 3.7.3 (existência e unicidade)**.** Considere a cadeia de Markov com matriz de transição

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Seja $\pi = (\pi_0, \pi_1)$. Então $\pi = \pi P$ é equivalente a

$$\begin{aligned}\pi_0 &= \frac{1}{2}\pi_0 + \frac{1}{2}\pi_1 \\ \pi_1 &= \frac{1}{2}\pi_0 + \frac{1}{2}\pi_1\end{aligned}$$

Destas equações concluímos que $\pi_0 = \pi_1$. Como $\pi_0 + \pi_1 = 1$ concluímos que $\pi_0 = \pi_1 = \frac{1}{2}$ é a distribuição estacionária e é única.

Exemplo 3.7.4 (existência e não unicidade)**.** Considere a cadeia de Markov com matriz de transição

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Então $\pi = \pi P = \pi$ porque P é a matriz identidade. Ou seja, qualquer distribuição de probabilidade é estacionária.

Exemplo 3.7.5 (inexistência)**.** Considere a cadeia de Markov com estados $S = \{0, 1, 2, \ldots\}$ e matriz de transição

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} & \frac{1}{32} & \cdots \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} & \cdots \\ 0 & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

A equação $\pi = \pi P$ é equivalente ao sistema de equações (em número infinito)

$$\begin{aligned}
\pi_0 &= \frac{1}{2}\pi_0 \\
\pi_1 &= \frac{1}{4}\pi_0 + \frac{1}{2}\pi_1 \\
\pi_2 &= \frac{1}{8}\pi_0 + \frac{1}{4}\pi_1 + \frac{1}{2}\pi_2 \\
&\vdots
\end{aligned}$$

A solução do sistema é $\pi_i = 0$ para todo $i \in \{0, 1, 2, \ldots\}$. No entanto, o vector nulo $\pi = 0$ não é uma distribuição de probabilidade. Logo, a cadeia não tem distribuição estacionária. De facto, a matriz P é triangular superior com um unico valor proprio igual a $1/2$. Repare que todos os estados da cadeia são transientes! Compare com o Exemplo 3.2.8.

O próximo resultado mostra que os estados transientes não têm peso na distribuição estacionária.

Proposição 3.7.6. *Seja $\pi = (\pi_j : j \in S)$ uma distribuição estacionária da cadeia de Markov. Se $j \in S$ é um estado transiente, então $\pi_j = 0$.*

Demonstração. Como j é transiente, $\lim_{n\to\infty} P_{i,j}^{(n)} = 0$ para todo $i \in S$ (ver Corolário 3.2.4). Então

$$\pi_j = \sum_{i\in S} P_{i,j}^{(n)} \pi_i \underset{n\to\infty}{\longrightarrow} 0.$$

□

Exemplo 3.7.7. Considere a cadeia de Markov do Exemplo 3.2.6. Recorde-se que tem matriz de transição

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$$

Como 0 é um estado transiente, obtemos que $\pi_0 = 0$. De facto, a distribuição estacionária é $\pi = (0, 1)$.

O resultado seguinte relaciona os estados recorrentes com a distribuição estacionária e mostra que toda a cadeia de Markov com pelo menos um estado recorrente positivo admite pelo menos uma distribuição estacionária que é única no caso de existir uma única classe de comunicação entre estados recorrentes.

Teorema 3.7.8. *Toda a cadeia de Markov admite uma distribuição estacionária se e só se tiver pelo menos um estado recorrente positivo. Mais, a distribuição estacionária é única se e só se todos os estados recorrentes comunicam entre si. Nesse caso, a distribuição estacionária $\pi = (\pi_j : j \in S)$ é dada por*

$$\pi_j = \frac{1}{m_j},$$

onde m_j é o tempo médio de reentrada no estado j, ou seja, $m_j = E(T_j|X_0 = j)$.

Demonstração. Procurar na bibliografia. $\square$

Observação 3.7.9. Se o espaço de estados é finito, então sabemos das Proposições 3.2.7 e 3.2.11 que toda a cadeia de Markov tem pelo menos um estado recorrente positivo. Logo, admite distribuições estacionárias. A distribuição é única se tiver uma única classe de comunicação de estados recorrentes.

Exemplo 3.7.10. Considere a cadeia de Markov do Exemplo 3.2.5. A matriz de transição é

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

A cadeia de Markov é irredutível e todos os estados são recorrentes. É fácil calcular a distribuição estacionária $\pi = (\frac{1}{2}, \frac{1}{2})$ que é única. Concluímos também que os tempos médios de reentrada são $m_1 = m_2 = 2$.

Exemplo 3.7.11. O passeio aleatório (simétrico ou assimétrico) não tem estados recorrentes positivos, portanto não admite distribuições estacionárias.

De seguida iremos relacionar as distribuição limite das probabilidades de transição de uma cadeia de Markov com as distribuições estacionárias. O próximo resultado mostra que se o limite das probabilidades de transição existir e for independente do estado de partida, então definem uma distribuição estacionária que é única.

Proposição 3.7.12. *Se para quaisquer $i, j \in S$, o seguinte limite é convergente e independente de i,*

$$\pi_j := \lim_{n\to\infty} P_{i,j}^{(n)} < +\infty$$

e $\pi := (\pi_j : j \in S) \neq 0$ ou S é finito, então π é a única distribuição estacionária da cadeia de Markov.

Demonstração. Para simplificar a demonstração consideremos o caso de S finito[6]. Pela equação de Chapman-Kolmogorov,

$$\pi_j = \lim_{n\to\infty} P_{i,j}^{(n)} = \lim_{n\to\infty} \sum_{k\in S} P_{i,k}^{(n-1)} P_{k,j} = \sum_{k\in S} \lim_{n\to\infty} P_{i,k}^{(n-1)} P_{k,j} = \sum_{k\in S} \pi_k P_{k,j}$$

e

$$\sum_{j\in S} \pi_j = \sum_{j\in S} \lim_{n\to\infty} P_{i,j}^{(n)} = \lim_{n\to\infty} \sum_{j\in S} P_{i,j}^{(n)} = \lim_{n\to\infty} 1 = 1.$$

Logo, $\pi = \pi P$ e $\sum_j \pi_j = 1$, ou seja, π é uma distribuição estacionária. Para mostrar a sua unicidade, suponhamos que ν é outra distribuição estacionária. Como $\nu = \nu P^n$ temos que

$$\nu_j = \sum_{i\in S} P_{i,j}^{(n)} \nu_i \underset{n\to\infty}{\longrightarrow} \sum_{i\in S} \pi_j \nu_i = \pi_j.$$

□

A distribuição π obtida pelo limite das probabilidade de transição é conhecida por **distribuição limite**.

Observação 3.7.13. A distribuição limite de X_n não depende do estado inicial da cadeia,

$$\lim_{n\to\infty} \mathbb{P}(X_n = j) = \lim_{n\to\infty} \sum_{i\in S} P_{i,j}^{(n)} \mathbb{P}(X_0 = i) = \pi_j.$$

O próximo resultado complementa a Proposição 3.7.12. Estabelece uma condição suficiente para a convergência das probabilidades de transição, definindo no limite uma única distribuição estacionária.

Teorema 3.7.14 (Teorema Ergódico)**.** *Se a cadeia de Markov é irredutível e aperiódica, então*

$$\lim_{n\to\infty} P_{i,j}^{(n)} = \frac{1}{m_j}.$$

Demonstração. Procurar na bibliografia. □

Quando o conjunto dos estados S é finito obtemos o seguinte corolário.

Corolário 3.7.15. *Considere uma cadeia de Markov com espaço de estados S finito. Se a cadeia de Markov é irredutível e aperiódica, então admite uma única distribuição estacionária $\pi = (\pi_j \colon j \in S)$ dada por*

$$\pi_j := \lim_{n\to\infty} P_{i,j}^{(n)} = \frac{1}{m_j}.$$

[6] O caso S infinito é tecnicamente mais sofisticado.

Demonstração. Segue da Proposição 3.7.12 e do Teorema 3.7.14. □

Observação 3.7.16. A razão $\pi_j = 1/m_j$ representa a fracção de tempo passado no estado j.

Observação 3.7.17. Se $(X_n)_{n\geq 0}$ é uma cadeia de Markov irredutível com período d então $(X_{nd})_{n\geq 0}$ é uma cadeia aperiódica. Logo,

$$\lim_{n\to\infty} P_{i,j}^{(nd)} = \frac{d}{m_j}.$$

Exemplo 3.7.18. Considere a cadeia do Exercício 10. É irredutível e aperiódica. Tem uma única distribuição estacionária dada por,

$$\pi = \left(\frac{q}{p+q}, \frac{p}{p+q}\right).$$

Exemplo 3.7.19. Considere a cadeia de Markov do Exemplo 3.5.7. Recorde-se a matriz de transição:

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Existem duas classes de comunicação recorrentes $C(0)$ e $C(4)$. Cada classe de comunicação define uma cadeia de Markov a dois estados irredutível e aperiódica. Logo, todas as distribuições estacionárias da cadeia são uma combinação convexa das distribuições,

$$\nu = \left(\frac{1}{3}, \frac{2}{3}, 0, 0, 0, 0\right) \quad \text{e} \quad \mu = \left(0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}\right).$$

Exemplo 3.7.20 (Aplicação: Modelo de Aprovisionamento)**.** Considere o modelo de aprovisionamento descrito no Exemplo 3.1.10, onde apenas 0, 1, ou 2 artigos são procurados em cada período ($a_k = 0$ para $k \geq 3$) com probabilidade

$$\mathbb{P}(\xi_n = 0) = \frac{1}{2}, \quad \mathbb{P}(\xi_n = 1) = \frac{2}{5} \quad \text{e} \quad \mathbb{P}(\xi_n = 2) = \frac{1}{10}.$$

Suponha que $s = 0$ e $m = 2$. O espaço de estados da cadeia é $S = \{2, 1, 0, -1\}$ e a matriz de transição

$$P = \begin{pmatrix} 1/2 & 2/5 & 1/10 & 0 \\ 0 & 1/2 & 2/5 & 1/10 \\ 1/2 & 2/5 & 1/10 & 0 \\ 1/2 & 2/5 & 1/10 & 0 \end{pmatrix}.$$

É fácil verificar que todos os estados da cadeia comunicam entre si. Portanto, a cadeia é irredutível. Como $P_{2,2} > 0$, concluímos que o estado 2 é aperiódico, logo a cadeia também é. Pelo Corolário 3.7.15, a cadeia admite uma única distribuição estacionária. Podemos determinar a distribuição resolvendo o sistema $\pi = \pi P$ onde $\pi = (\pi_2, \pi_1, \pi_0, \pi_{-1})$,

$$\begin{aligned} \pi_2 &= \frac{1}{2}\pi_2 + \frac{1}{2}\pi_0 + \frac{1}{2}\pi_{-1} \\ \pi_1 &= \frac{2}{5}\pi_2 + \frac{1}{2}\pi_1 + \frac{2}{5}\pi_0 + \frac{2}{5}\pi_{-1} \\ \pi_0 &= \frac{1}{10}\pi_2 + \frac{2}{5}\pi_1 + \frac{1}{10}\pi_0 + \frac{1}{10}\pi_{-1} \\ \pi_{-1} &= \frac{1}{10}\pi_1 \end{aligned}$$

Resolvendo este sistema determinamos os coeficientes da distribuição estacionária

$$\pi = \left(\frac{5}{18}, \frac{4}{9}, \frac{7}{30}, \frac{2}{45}\right).$$

Como exemplo, concluímos que o tempo médio de reentrada da cadeia no estado 2 é 18/5. Mais, a probabilidade de a procura exceder a oferta a longo prazo é

$$\lim_{n\to\infty} \sum_{j<0} \mathbb{P}(X_n = j) = \frac{2}{45}$$

e o nível médio de stock a longo prazo é

$$\lim_{n\to\infty} \sum_{j>0} j\mathbb{P}(X_n = j) = 2 \times \frac{5}{18} + 1 \times \frac{4}{9} = 1.$$

Exemplo 3.7.21 (PageRank do Google)**.** O motor de busca na internet mais famoso do mundo lista os seus resultados de acordo com a probabilidade com que um utilizador entrando aleatoriamente em links chega a uma determinada pagina na internet.

Podemos representar a internet como um grafo $G = (V, E)$ onde os vértices correspondem às páginas e as arestas aos links ligando as diversas páginas. Designamos por $n = \#V$ o número de páginas que existem na internet e $L(i)$ o número distinto de links existentes na página $i \in V$. Um utilizador "aleatório" situado numa página $i \in V$ entra numa página $j \in V$ com probabilidade

$$\hat{P}_{i,j} = \begin{cases} \frac{1}{n} & \text{se } L(i) = 0 \\ \frac{1}{L(i)} & \text{se } L(i) > 0 \text{ e } (i,j) \in E \\ 0 & \text{se } L(i) > 0 \text{ e } (i,j) \notin E \end{cases}$$

A matriz $\hat{P} = (\hat{P}_{i,j})_{i,j \in V}$ é estocástica. Para garantir que a cadeia de Markov associada é irredutível e aperiódica fazemos uma pequena transformação,

$$P_{i,j} = (1 - \epsilon)\hat{P}_{i,j} + \frac{\epsilon}{n}$$

onde $\epsilon \in [0, 1[$ é um pequeno parâmetro. Assim, $P_{i,j} > 0$ para quaisquer $i, j \in V$ garantindo que a cadeia de Markov com matriz de transição $P = (P_{i,j})_{i,j \in V}$ é irredutível e aperiódica. Pelo Corolário 3.7.15, a cadeia admite uma única distribuição estacionária,

$$\pi = \pi P$$

Se $\pi_i > \pi_j$, então a página i é mais relevante que a página j. Um dos segredos do Google reside num método numérico eficiente e uma grande capacidade computacional para calcular as maiores entradas da distribuição estacionária π.

3.8 Exercícios

Exercício 18. Determine as distribuições estacionárias das cadeias de Markov com matrizes de transição:

$$\begin{pmatrix} \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \end{pmatrix}.$$

Exercício 19. Determine a distribuição limite da cadeia de Markov com matriz transição

$$P = \begin{pmatrix} q & p & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ q & 0 & 0 & p & 0 \\ q & 0 & 0 & 0 & p \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad 0 < p, q < 1, \quad p + q = 1.$$

Exercício 20. Considere a cadeia de Markov com estados $S = \{0, 1, 2, 3\}$ e matriz transição,

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 \end{pmatrix}$$

Determine:

1. A distribuição limite da cadeia de Markov.

2. Calcule $E(T_0 | X_0 = 1)$. (Dica: observe que a cadeia passa sempre do estado 0 para o estado 1)

3.9 Estados absorventes e probabilidade de absorção

Seja $(X_n)_{n \geq 0}$ uma cadeia de Markov homogénea com espaço de estados S (finito ou infinito).

Um tipo particular de estados recorrentes são os estados absorventes.

Definição 3.9.1. Um estado $i \in S$ é **absorvente** se e só se $P_{i,i} = 1$.

Ou seja, uma cadeia de Markov que esteja num estado absorvente i tem probabilidade zero de sair do mesmo, isto é, $P_{i,j} = 0$ para todo $j \neq i$.

Observação 3.9.1.

1. Se i é absorvente então $C(i) = \{i\}$, isto é, a classe de comunicação de i é o próprio i.

2. Se $C \subset S$ é um conjunto de estados fechado contendo apenas um estado i, então i é absorvente.

3. Estados absorventes são ergódicos. De facto, pela equação de Chapman-Kolmogorov,

$$P_{i,i}^{(n)} = \sum_{j\in S} P_{i,j}P_{j,i}^{(n-1)} = P_{i,i}P_{i,i}^{(n-1)} = P_{i,i}^{(n-1)}.$$

Logo, $P_{i,i}^{(n)} = 1$ para todo $n \geq 1$, e $Per(i) = 1$.

Exemplo 3.9.2. Considere a seguinte matriz estocástica associada a uma cadeia de Markov homogénea $(X_n)_{n\geq 0}$ com estados $S = \{0, 1, 2\}$,

$$P = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & \beta & \gamma \\ 0 & 0 & 1 \end{pmatrix}, \quad \alpha + \beta + \gamma = 1, \quad \alpha, \beta, \gamma > 0.$$

Os estados 0 e 2 são absorventes enquanto o estado 1 é transiente.

Seja $A \subset S$ um conjunto de estados fechado[7]. Em aplicações é importante calcular a **probabilidade de absorção em** A,

$$h_i := \mathbb{P}(T < \infty | X_0 = i), \quad i \in S$$

onde $T = \min\{n \geq 0 \colon X_n \in A\}$ é o primeiro tempo de retorno ao conjunto A. Designamos por **tempo médio de absorção** ao valor esperado

$$t_i := E(T | X_0 = i).$$

É fácil determinar h_i e t_i para certos estados $i \in S$. Note-se que se $i \in A$ então

$$h_i = 1 \quad e \quad t_i = 0.$$

Por outro lado, se i é um estado absorvente mas não pertence a A então[8]

$$h_i = 0 \quad \text{e} \quad t_i = \infty.$$

De seguida vamos deduzir umas equações que permitem calcular h_i e t_i para qualquer estado $i \in S$.

Proposição 3.9.3. *O vector de probabilidades de absorção em A, $h = (h_i \colon i \in S)$, é solução do seguinte sistema de equações lineares,*

$$\begin{cases} h_i = 1, & se\ i \in A \\ h_i = \sum_{j\in S} P_{i,j}h_j & se\ i \notin A \end{cases}$$

[7] O conjunto A pode ser formado por estados absorventes ou mais geral por classes de comunicação recorrentes.

[8] Mais geral, se $i \nrightarrow j$ para todo $j \in A$, então $h_i = 0$ e $t_i = \infty$.

Demonstração. Já vimos que se $i \in A$, então $h_i = 1$. Suponhamos que $i \notin A$. Então $T \geq 1$. Logo, usando a propriedade de Markov e a homogeneidade da cadeia,

$$\begin{aligned}
h_i &= \mathbb{P}(T < \infty | X_0 = i) \\
&= \sum_{j \in S} \mathbb{P}(T < \infty | X_1 = j, X_0 = i) \mathbb{P}(X_1 = j | X_0 = i) \\
&= \sum_{j \in S} \mathbb{P}(T < \infty | X_1 = j) P_{i,j} \\
&= \sum_{j \in S} h_j P_{i,j}
\end{aligned}$$

□

Observação 3.9.4. Seja $\hat{P}$ a matriz que se obtém removendo as linhas $i \in A$ de P. Analogamente, seja $\hat{h}$ o vector coluna que se obtém removendo as entradas $i \in A$ do vector h. Então

$$\hat{h} = \hat{P} h.$$

Compare com a equação que define uma distribuição estacionária.

Exemplo 3.9.5. Considere a cadeia de Markov do Exemplo 3.9.2. Seja $A = \{0\}$ o conjunto formado pelo estado absorvente 0. Aplicando a proposição anterior obtemos,

$$\begin{aligned}
h_0 &= 1 \\
h_1 &= \alpha + \beta h_1 + \gamma h_2 \\
h_2 &= 0.
\end{aligned}$$

Note-se que o estado 2 é absorvente. Portanto $h_2 = 0$. Assim, concluímos que

$$h_1 = \frac{\alpha}{1 - \beta} = \frac{\alpha}{\alpha + \gamma}.$$

Ou seja, a probabilidade de absorção no estado 0 partindo do estado 1 é $\frac{\alpha}{\alpha+\gamma}$.

Proposição 3.9.6. *Se o vector de tempos médios de absorção em A, $t = (t_i \colon i \in S)$ for finito*[9], *então t é solução do seguinte sistema de equações lineares,*

$$\begin{cases} t_i = 0, & \text{se } i \in A \\ t_i = 1 + \sum_{j \in S} P_{i,j} t_j & \text{se } i \notin A \end{cases}$$

[9] $t_i < \infty$ para todo $i \in S$

Demonstração. Análoga à prova da proposição anterior. Procurar na bibliografia. □

Exemplo 3.9.7. Continuando o exemplo anterior, vamos agora calcular os tempos médios de absorção em $A = \{0, 2\}$. Note-se que $t_0 = t_2 = 0$. Para calcular t_1 aplicamos a proposição anterior,

$$t_1 = 1 + \alpha t_0 + \beta t_1 + \gamma t_2$$

Portanto,

$$t_1 = \frac{1}{1-\beta} = \frac{1}{\alpha+\gamma}.$$

Ou seja, o tempo médio de absorção partindo do estado 1 é $\frac{1}{\alpha+\gamma}$.

3.10 Exercícios

Exercício 21. Considere uma cadeia de Markov com estados $S = \{0, 1, 2, 3\}$ e matriz de transição

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{2}{5} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Determine os tempos médios de absorção no estado 3.

Exercício 22. Considere uma cadeia de Markov com estados $S = \{0, 1, 2, 3\}$ e matriz de transição

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{2} & \frac{1}{8} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

1. Partindo do estado 1, determine a probabilidade de a cadeia ser absorvida pelo estado 0.

2. Determine o tempo médio de absorção em $A = \{0, 3\}$.

Exercício 23. Considere o seguinte jogo. Uma moeda perfeita é lançada sucessivamente ao ar até que apareçam duas caras sucessivas.

1. Modele o jogo usando uma cadeia de Markov. Determine a matriz de transição e o diagrama da cadeia. (Dica: $S = \{0, 1, 2\}$)

2. Determine a decomposição do espaço dos estados da cadeia.

3. Calcule o tempo médio de duração do jogo supondo que começa o jogo com duas coroas.

Chapter 4

Processo de Poisson

Processos de Poisson são usados para contar o número de ocorrências "aleatoriamente" espaçadas entre si ao longo do tempo. Por exemplo, o número de particular emitidas por um material radioativo ou o número de chamadas telefónicas que chegam a uma central. Antes de introduzir o Processo de Poisson vamos relembrar as propriedades da distribuição exponencial.

Definição 4.0.1. Uma variável aleatória T com valores em $[0, +\infty[$ tem **distribuição exponencial** com parâmetro $\lambda > 0$ se

$$\mathbb{P}(T \leq t) = 1 - e^{-\lambda t}, \quad t \geq 0.$$

A função densidade de probabilidade $f_T(t)$ é

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

Facilmente determinamos que $E(T) = 1/\lambda$ e $V(T) = 2/\lambda^2$.

A distribuição exponencial tem a propriedade de **não possuir memória**, ou seja,

$$\mathbb{P}(T > t + s | T > t) = \mathbb{P}(T > s).$$

De facto,

$$\begin{aligned} \mathbb{P}(T > t + s | T > t) &= \frac{\mathbb{P}(T > t + s|, T > t)}{\mathbb{P}(T > t)} \\ &= \frac{\mathbb{P}(T > t + s)}{\mathbb{P}(T > t)} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} \\ &= \mathbb{P}(T > s). \end{aligned}$$

Como demonstra o próximo resultado, a soma de variáveis aleatórias independentes com distribuição exponencial tem **distribuição gamma**.

Proposição 4.0.1. *Sejam $T_1, \dots, T_n$, $n \geq 1$ variáveis aleatórias independentes com distribuição exponencial com parâmetro λ e $W_n = T_1 + \cdots + T_n$. Então W_n tem distribuição gamma com parâmetros n e λ,*

$$f_{W_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, \quad t \geq 0.$$

Demonstração. A prova é por indução em n. Para $n = 1$ temos $W_1 = T_1$ que tem distribuição exponencial $f_{W_1}(t) = f_{T_1}(t) = \lambda e^{-\lambda t}$. Suponhamos que a proposição é verdadeira para n. Vamos mostrar o mesmo para $n+1$. Usando o facto[1] $W_{n+1} = W_n + T_{n+1}$ e a independência dos T_i's temos que

$$f_{W_{n+1}}(t) = \int_0^t f_{W_n}(s) f_{T_{n+1}}(t-s)\, ds.$$

Usando a hipótese de indução,

$$\begin{aligned} f_{W_{n+1}}(t) &= \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} \lambda e^{-\lambda(t-s)}\, ds \\ &= \lambda e^{-\lambda t} \lambda^n \int_0^t \frac{s^{n-1}}{(n-1)!}\, ds \\ &= \lambda e^{-\lambda t} \lambda^n \frac{t^n}{n!}. \end{aligned}$$

□

4.1 Definição do processo de Poisson

Definição 4.1.1. Uma variável aleatória X tem **distribuição de Poisson** com parâmetro $\lambda > 0$ se

$$\mathbb{P}(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}, \quad n = 0, 1, 2, \dots$$

[1]Dadas duas variáveis aleatórias X e Y independentes com densidade f_X e f_Y, então a densidade da soma $Z = X + Y$ é a convolução das densidades f_X e f_Y, ou seja,

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\, dx.$$

É fácil verificar que $E(X) = V(X) = \lambda$. Outra propriedade a reter é a soma de variáveis aleatórias independentes com distribuição de Poisson tem também distribuição de Poisson.

Proposição 4.1.1. *Sejam $X_1, \ldots, X_k$ variáveis aleatórias independentes com distribuição de Poisson de parâmetro $\lambda_1, \ldots, \lambda_k$, respectivamente. Então a soma $X_1 + \cdots + X_k$ tem distribuição de Poisson com parâmetro $\lambda_1 + \cdots + \lambda_k$.*

Demonstração. Exercício. □

A distribuição de Poisson surge considerendo o limite da distribuição Binomial quando a probabilidade do evento é λ/n e $n \to \infty$. Esta aproximação é conhecida por **lei dos eventos raros**. Suponhamos que λ representa o número médio de eventos que se observam num intervalo de tempo (ou espaço). Divida-se o intervalo em n subintervalos de comprimento igual tal que $n > \lambda$. Então λ/n representa a proporção de eventos que se observam em cada subintervalo. Assumimos que a ocorrência de um evento num subintervalo é uma variável aleatória com **distribuição de Bernoulli** com parâmetro λ/n . Assim, o número de eventos no intervalo inicial e uma **distribuição Binomial** com parâmetros n e λ/n, ou seja, a probabilidade de observar k eventos é

$$\frac{n!}{(n-k)!k!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}.$$

Reescrevendo obtemos

$$\frac{n(n-1)\cdots(n-k+1)}{n^k}\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

Tomando o limite $n \to \infty$, é fácil verificar que o 1º termo do produto tem limite 1, o 2º não depende de n, o 3º tem limite $e^{-\lambda}$ e o 4º tem limite 1. Portanto,

$$\lim_{n\to\infty}\frac{n!}{(n-k)!k!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Vamos agora definir o processo de Poisson.

Definição 4.1.2 (Processo de Poisson)**.** $\{N(t), t \geq 0\}$ é um **processo de Poisson** com parâmetro $\lambda > 0$ se

1. $N(0) = 0$,

2. os incrementos $N(t+s)-N(s)$ têm distribuição de Poisson com parâmetro λt,

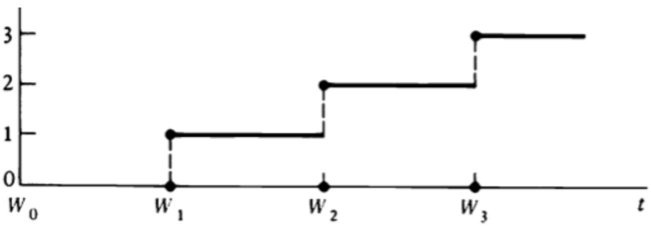

Figure 4.1: Realização de $N(t)$

3. $N(t)$ tem incrementos independentes, isto é, se $t_0 < t_1 < \cdots < t_n$ então

$$N(t_1) - N(t_0), \ldots, N(t_n) - N(t_{n-1}) \quad \text{são independentes.}$$

O Processo de Poisson $N(t)$ conta o número de eventos que ocorreram até ao instante t. É portanto um processo de contagem. Como têm incrementos estacionários e independentes, pertence a uma classe mais geral de processos que são as cadeias de Markov a tempo contínuo.

4.2 Construção do processo de Poisson

Definição 4.2.1. Sejam $T_1, \ldots, T_n$ variáveis aleatórias independentes com distribuição exponencial de parâmetro λ. Considere-se a soma $W_n = T_1 + \cdots + T_n$ com $W_0 = 0$ e

$$N(t) := \max\{n \geq 0 \colon W_n \leq t\}.$$

Mostremos agora que $N(t)$ é um processo de Poisson.

Lemma 4.2.1. *$N(t)$ tem distribuição de Poisson com parâmetro λt.*

Demonstração. Note-se que $N(t) = n$ se e só se $W_n \leq t < W_{n+1}$. Portanto,

$$\begin{aligned}
\mathbb{P}(N(t) = n) &= \mathbb{P}(W_n \leq t < W_{n+1}) \\
&= \int_0^t \mathbb{P}(W_n \leq t < W_{n+1} | W_n = s) f_{W_n}(s)\, ds \\
&= \int_0^t \mathbb{P}(T_{n+1} > t - W_n | W_n = s) f_{W_n}(s)\, ds \\
&= \int_0^t \mathbb{P}(T_{n+1} > t - s) f_{W_n}(s)\, ds.
\end{aligned}$$

Segue da Proposição 4.0.1 que

$$\mathbb{P}(N(t) = n) = \int_0^t e^{-\lambda(t-s)}\lambda e^{-\lambda s}\frac{(\lambda s)^{n-1}}{(n-1)!}\,ds = e^{-\lambda t}\frac{(\lambda t)^n}{n!}.$$

□

Lemma 4.2.2. *$N(t+s) - N(s)$ tem distribuição de Poisson com parâmetro λt e é independente de $N(r)$ para $0 \leq r \leq s$.*

Demonstração. Para fixar ideias, suponhamos que até ao instante s exactamente 3 eventos ocorreram nos instantes w_1, w_2 e w_3 (ver Figura 4.1). O tempo de espera até quarto evento é T_4 que por hipótese $s < w_3 + T_4$ (caso contrário 4 eventos teriam ocorrido). Como a distribuição exponencial não tem memória temos que,

$$\mathbb{P}(T_4 > s - w_3 + t | T_4 > s - w_3) = \mathbb{P}(T_4 > t).$$

Isto mostra que a distribuição do tempo de espera até ao primeiro evento depois de s é exponencial com parâmetro λ e independente de T_1, T_2 e T_3. Logo, $N(t+s) - N(s)$ tem a mesma distribuição de $N(t)$ e é independente de $N(r)$ com $0 \leq r \leq s$. □

Lemma 4.2.3. *$N(t)$ tem incrementos independentes.*

Demonstração. O lemma anterior mostra que $N(t_n) - N(t_{n-1})$ é independente de $N(r)$ com $r \leq t_{n-1}$. Logo é independente de $N(t_{n-1}) - N(t_{n-2})$. O resultado segue usando indução. □

Mostrámos assim o seguinte teorema.

Teorema 4.2.4. *$N(t)$ é um processo de Poisson.*

4.3 Caracterização

Descrevemos agora uma outra forma equivalente para caracterizar um processo de Poisson. Escrevemos $o(h)$ para designar uma função que satisfaz $\lim_{h\to 0} o(h)/h = 0$.

Teorema 4.3.1. *Seja $\{N(t) \colon t \geq 0\}$ um processo estocástico tal que $N(0) = 0$ e $N(t) \in \{0, 1, 2, \ldots\}$ para todo $t \geq 0$. O processo $N(t)$ é Poisson se e só se*

1. $N(t)$ tem incrementos independentes

2. $\mathbb{P}(N(t+h) - N(t) = 1) = \lambda h + o(h)$ quando $h \to 0$

3. $\mathbb{P}(N(t+h) - N(t) = 0) = 1 - \lambda h + o(h)$ quando $h \to 0$

Demonstração. Se $N(t)$ é um processo de Poisson, então $N(t+h) - N(t)$ tem distribuição Poisson com parâmetro λh. Portanto,

$$\mathbb{P}(N(t+h) - N(t) = 0) = e^{-\lambda h} \quad \text{e} \quad \mathbb{P}(N(t+h) - N(t) = 1) = \lambda h e^{-\lambda h}.$$

Logo,

$$\mathbb{P}(N(t+h) - N(t) = 0) = e^{-\lambda h} = 1 - \lambda h + o(h)$$

e

$$\mathbb{P}(N(t+h) - N(t) = 1) = \lambda h(1 - \lambda h + o(h)) = \lambda h + o(h)$$

Suponhamos agora que (1)-(3) se verificam e queremos mostrar que $N(t+s) - N(s)$ tem distribuição de Poisson com parâmetro λt. Consideremos o caso $s = 0$ uma vez que o caso $s > 0$ é análogo. Seja

$$p_n(t) = \mathbb{P}(N(t) = n).$$

Note-se que

$$\begin{aligned} p_n(t+h) = {} & \mathbb{P}(N(t) = n, N(t+h) - N(t) = 0) \\ & + \mathbb{P}(N(t) = n, N(t+h) - N(t) = 1) \\ & + \mathbb{P}(N(t) = n, N(t+h) - N(t) \geq 2) \end{aligned}$$

Segue de (1), (2) e (3) que

$$p_n(t+h) = p_n(t)p_0(h) + p_{n-1}(t)p_1(h) + o(h)$$

e

$$p_0(h) = 1 - \lambda h + o(h), \quad p_1(h) = \lambda h + o(h)$$

Logo,

$$p_n(t+h) - p_n(t) = -\lambda h p_n(t) + \lambda h p_{n-1}(t) + o(h)$$

Dividindo ambos os termos da equação por h e tomando o limite $h \to 0$ temos que

$$p_n'(t) = -\lambda p_n(t) + \lambda p_{n-1}(t).$$

Note-se que $p_{-1}(t) = 0$. Temos de resolver este sistema de equações diferenciais com condição inicial

$$p_n(0) = \mathbb{P}(N(0) = n) = \begin{cases} 1 & n = 0 \\ 0 & n \geq 1 \end{cases}$$

Quando $n = 0$ temos

$$p_0'(t) = -\lambda p_0(t) \quad \text{com} \quad p_0(t) = 1.$$

A solução é $p_0(t) = e^{-\lambda t}$. Usando indução é resolver as restantes equações e obter solução

$$p_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}.$$

□

4.4 Relação com a distribuição Binomial

Proposição 4.4.1. *Sejam $s < t$ e $m \leq n$. Então*

$$\mathbb{P}(N(s) = m | N(t) = n) = \frac{n!}{m!(n-m)!}\left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m}$$

Demonstração. Exercício. □

4.5 Processo de Poisson não homogéneo

Uma extensão do processo de Poisson considera intensidades que variam ao longo do tempo.

Definição 4.5.1. Um processo $\{N(t) \colon t \geq 0\}$ é um processo de Poisson com intensidade $\lambda(t)$ se

1. $N(0) = 0$
2. $N(t)$ tem incrementos independentes
3. $N(t) - N(s)$ tem distribuição de Poisson com parâmetro $\int_s^t \lambda(r)\, dr$

É possível caracterizar um processo de Poisson não homogéneo através das aproximações de primeira ordem dos incrementos.

Teorema 4.5.1. *Um processo $\{N(t) \colon t \geq 0\}$ é um processo de Poisson com intensidade $\lambda(t)$ sse*

1. *Tem incrementos independentes*
2. $\mathbb{P}(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$

3. $\mathbb{P}(N(t+h) - N(t) = 0) = 1 - \lambda(t)h + o(h)$

Demonstração. Procurar na bibliografia. $\square$

Exemplo 4.5.2. Suponha que a procura de um produto numa loja entre as 9h e as 13h segue um processo de Poisson com taxa

$$\lambda(t) = \begin{cases} 2t, & 0 \leq t < 1 \\ 2, & 1 \leq t < 2 \\ 4 - t, & 2 \leq t \leq 4 \end{cases}$$

Vamos determinar a probabilidade da procura nas primeiras 2h ser igual a 2 artigos. Para tal, calculamos a taxa média entre $t = 0$ e $t = 2$, portanto

$$\mu = \int_0^2 \lambda(t)\, dt = \int_0^1 2t\, dt + \int_1^2 2\, dt = 1 + 2 = 3.$$

Assim,

$$\mathbb{P}(N(2) = 2) = e^{-3}\frac{3^2}{2!}.$$

No entanto, se quisermos determinar a probabilidade da procura ser igual a 2 artigos entre as 11h e as 13h, então

$$\mu = \int_2^4 \lambda(t)\, dt = \int_2^4 4 - t\, dt = 2,$$

e

$$\mathbb{P}(N(4) - N(2) = 2) = e^{-2}\frac{2^2}{2!}.$$

4.6 Exercícios

Exercício 24. Um autocarro chega a uma determinada paragem de 10 em 10 minutos. Suponha que o número de autocarros que chegam à paragem segue um processo de Poisson.

1. Qual é a probabilidade de o intervalo entre chegadas sucessivas ser superior a 20 minutos?

2. Após perder um autocarro, quanto tempo tem um passageiro de esperar para apanhar o autocarro seguinte com probabilidade de 0.5?

3. Dado que na última hora não chegou nenhum autocarro qual é a probabilidade esperar mais uma hora?

Exercício 25. Seja T uma variável aleatória com distribuição de probabilidade contínua. Mostre que T tem distribuição exponencial se e só se T não tem memória, isto é,

$$\mathbb{P}(T > x + y | T > x) = \mathbb{P}(T > y).$$

Exercício 26. A chegada de passageiros a uma paragem de autocarro segue um processo de Poisson com intensidade λ_1. Seja T o tempo de chegada de um autocarro que é independente do processo de Poisson. Quando $t = 0$ não existem passageiros na paragem. Supondo que T segue uma distribuição exponencial com intensidade λ_2, calcule o número médio de pessoas na paragem no instante T.

Exercício 27. Suponha que N_1 e N_2 são processos de Poisson independentes com intensidades λ_1 e λ_2, respectivamente. Mostre que N_1+N_2 é um processo de Poisson com intensidade $\lambda_1 + \lambda_2$.

4.7 Processo de Poisson composto

Dado um processo de Poisson $N(t)$ com intensidade $\lambda > 0$ e uma sucessão $(X_n)_{n\geq 1}$ de variáveis aleatórias iid, e independentes do processo $N(t)$, designamos por $Z(t)$ o **processo de Poisson composto** definido por

$$Z(t) = \sum_{n=1}^{N(t)} X_n$$

Exemplo 4.7.1.

1. Em teoria do risco, $N(t)$ representa o número de sinistros observados até ao instante t, X_n a indemnização do n-ésimo sinistro, e $Z(t)$ a indemnização agregada até ao instante t.

2. Outro exemplo, em matemática financeira, usa o processo de Poisson composto para descrever a evolução do preço de um activo financeiro. Suponhamos que transações de um certo activo ocorrem segundo um processo de Poisson com intensidade $\lambda > 0$. A variável X_k representa a variação do preço do activo entre a $k - 1$ e a k-ésima transação. Supomos também que as variações são independentes entre si. Sob estas condições, $Z(t)$ representa a variação total do preço do activo até ao instante t.

A distribuição do processo de Poisson composto pode ser calculada explicitamente fazendo a convolução das distribuições dos incrementos do processo (ver Exercício 2). Suponhamos que as variáveis aleatórias X_n são contínuas e admitem densidade de probabilidade $f(x)$ (uma vez que são iid). Seja $f^{(n)}(x)$ a densidade de probabilidade de $X_1 + \cdots + X_n$, i.e.,

$$f^{(n)}(x) = f_{X_1+\cdots+X_n}(x).$$

Usando o Exercício 2, temos que

$$f^{(n)}(x) = \int_{-\infty}^{+\infty} f^{(n-1)}(x-y)f(y)\,dy, \quad n = 2, 3, \ldots$$

Assim,

$$\begin{aligned}
F^{(n)}(x) &= \mathbb{P}(X_1 + \cdots + X_n \leq x) \\
&= \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f^{(n-1)}(z-y)f(y)\,dy\,dz \\
&= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{x} f^{(n-1)}(z-y)\,dz \right) f(y)\,dy \\
&= \int_{-\infty}^{+\infty} F^{(n-1)}(x-y)f(y)\,dy
\end{aligned}$$

Então

$$\begin{aligned}
\mathbb{P}(Z(t) \leq x) &= \mathbb{P}\left(\sum_{k=1}^{N(t)} X_k \leq x \right) \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left(X_1 + \cdots + X_n \leq x | N(t) = n\right) \mathbb{P}(N(t) = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left(X_1 + \cdots + X_n \leq x\right) \mathbb{P}(N(t) = n) \\
&= \sum_{n=0}^{\infty} F^{(n)}(x) \frac{(\lambda t)^n}{n!} e^{-\lambda t}
\end{aligned}$$

Fórmulas para os momentos de primeira e segunda ordem podem ser obtidas analogamente.

Exercício 28. Mostre que

$$E(Z(t)) = \lambda \mu t \quad \text{e} \quad V(Z(t)) = \lambda(\nu^2 + \mu^2)t$$

onde $\mu = E(X_1)$ e $\nu^2 = V(X_1)$.

Expressões idênticas podem ser deduzidas quando as variáveis X_n são discretas. Neste caso, o integral é substituído por uma soma. Compare com o Exercício 3.

Exercício 29. Considere um sistema sujeito a impactos. Em média observa-se λ impactos por unidade de tempo. Seja $N(t)$ o número de impactos observados até instante t. Cada impacto provoca um estrago mensurável X_n no sistema. Supomos que as variáveis X_n são exponencialmente distribuídas com intensidade $\mu > 0$. O sistema permanece em funcionamento enquanto o estrago total não exceder um valor crítico $a > 0$. Usando o processo de Poisson composto para descrever o estrago total do sistema determine:

1. A distribuição de probabilidade do tempo de falha do sistema, i.e.,

$$T = \min\{t \geq 1 \colon Z(t) \geq a\}$$

2. O tempo médio de funcionamento $E(T)$.

Chapter 5

Processos de Markov em tempo contínuo

Seja $\{X(t)\colon t \geq 0\}$ um processo estocástico a tempo contínuo com espaço de estados $E = \{0, 1, 2, 3 \dots\}$.

Definição 5.0.1. Dizemos que $\{X(t)\colon t \geq 0\}$ é um **processo (ou cadeia) de Markov a tempo contínuo** sse

$$\mathbb{P}(X(t+s) = j | X(s) = i, X(s_n) = i_n \dots, X(s_0) = i_0) \\ = \mathbb{P}(X(t+s) = j | X(s) = i)$$

para todo $t > 0$, $s > s_n > \cdots > s_0 > 0$ e $j, i, i_n, \dots, i_0 \in E$.

Exemplo 5.0.1. Uma classe especial de processos de Markov são os processos com incrementos estacionários e independentes. Suponhamos que $X(0) = 0$. O processo $X(t)$ tem **incrementos estacionários e independentes** sse

1. $X(t) - X(s) \stackrel{d}{=} X(t-s)$ para $t > s$;

2. As variáveis aleatórias são independentes

$$X(t_n) - X(t_{n-1}), \dots, X(t_1) - X(t_0)$$

para quaisquer $t_0 < t_1 < \cdots < t_n$.

Um processo com incrementos estacionários e independentes tem uma expressão simples para a evolução da média e da covariância do processo. Sejam $m(t) = E(X(t))$ e $c(s,t) = \mathrm{Cov}(X(s), X(t))$. Então[1] existem $\mu \in \mathbb{R}$ e $\sigma > 0$ tal que

$$m(t) = \mu t \quad \text{e} \quad c(s,t) = \sigma^2 \min\{s, t\}.$$

[1] Prova-se usando o lemma de Cauchy. Ver `https://en.wikipedia.org/wiki/Cauchy%27s_functional_equation`

O processo de Markov é **homogéneo** se as probabilidades de transição não dependerem de s, ou seja,

$$\mathbb{P}(X(t+s)=j|X(s)=i)=\mathbb{P}(X(t)=j|X(0)=i).$$

Neste curso tratamos apenas dos processos homogéneos. Define-se

$$p_{i,j}(t)=\mathbb{P}(X(t)=j|X(0)=i).$$

É claro que

$$p_{i,j}(0)=\begin{cases}0, & i\neq j\\ 1, & i=j\end{cases}.$$

As equações de Chapman-Kolmogorov podem ser obtidas de forma equivalente ao caso discreto (cadeias de Markov).

Proposição 5.0.2 (Equações de Chapman-Kolmogorov)**.**

$$p_{i,j}(t+s)=\sum_{k\in E}p_{i,k}(t)p_{k,j}(s).$$

Demonstração. Exercício. □

Suponhamos que as probabilidade de transição $p_{i,j}(t)$ são funções bem comportadas, ou seja, que podemos tomar o limite

$$q_{i,j}=\lim_{t\to 0^+}\frac{p_{i,j}(t)}{t},\quad i\neq j.$$

No caso em que o limite existe designamos por $q_{i,j}$ a **intensidade do salto de i para j**. Note-se que $q_{i,j}=p'_{i,j}(0)$. Da igualdade,

$$p_{i,i}(t)=1-\sum_{j\neq i}p_{i,j}(t),$$

segue que

$$q_{i,i}:=\lim_{t\to 0^+}\frac{p_{i,i}(t)-1}{t}=-\sum_{j\neq i}q_{i,j}.$$

Note-se que $q_{i,i}=p'_{i,i}(0)$. A matriz formada pelas intensidades dos saltos é conhecida por **matriz de intensidades** do processo,

$$Q=(q_{i,j})_{i,j\in E}.$$

Uma matriz de intensidades Q tem a soma das linhas igual a zero e as entradas da diagonal sempre não-positivas.

Veremos de seguida que é possível construir um processo de Markov a partir de uma matriz de intensidades. Mais, a matriz de intensidades especifica a lei de probabilidade do processo, ou seja, é possível determinar a distribuição de probabilidade do processo a partir da matriz de intensidades. Antes de prosseguir, vejamos alguns exemplos.

Exemplo 5.0.3. O processo de Poisson tem incrementos estacionários e independentes, logo é um processo de Markov. Por outro lado, das equações que caracterizam o processo

$$\begin{aligned}
\mathbb{P}(N(t+h)-N(t)=1) &= \lambda h + o(h)\\
\mathbb{P}(N(t+h)-N(t)=0) &= 1-\lambda h + o(h)
\end{aligned}$$

obtemos

$$p_{i,j}(t)=\begin{cases}1-\lambda t+o(t), & j=i\\ \lambda t+o(t), & j=i+1\\ 0, & \text{caso contrário}\end{cases}.$$

Assim,

$$q_{i,j}=\begin{cases}-\lambda, & j=i\\ \lambda, & j=i+1\\ 0, & \text{caso contrário}\end{cases}$$

com matriz de intensidades

$$Q=\begin{pmatrix}-\lambda & \lambda & 0 & 0 & \cdots\\ 0 & -\lambda & \lambda & 0 & \cdots\\ 0 & 0 & -\lambda & \lambda & \cdots\\ \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}$$

Exemplo 5.0.4. Os **Processos de Nascimento e Morte** são processos de Markov a tempo contínuo que generalizam os processos de Poisson, no sentido em que são permitidos decrementos do processo. O processo de nascimento e morte é caracterizado por,

1. $p_{i,i+1}(t)=\lambda_i t+o(t)$, $i=0,1,\ldots$
2. $p_{i,i-1}(t)=\mu_i t+o(t)$, $i=1,2,\ldots$
3. $p_{i,i}(t)=1-(\lambda_i+\mu_i)t+o(t)$, $i=0,1,\ldots$
4. $p_{i,j}(t)=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}$

5. $\mu_0 = 0$, $\lambda_0 > 0$ e $\lambda_i, \mu_i > 0$ para $i = 1, 2, \ldots$

É fácil verificar que o processo tem a seguinte matriz de intensidades

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Exemplo 5.0.5. O **Processo de Ramificação** modela a evoluçao de uma população onde cada indivíduo nasce a uma taxa $\lambda > 0$ e morre a uma taxa $\mu > 0$. Assim

$$q_{i,i+1} = \lambda i \quad \text{e} \quad q_{i,i-1} = \mu i.$$

Quando $\mu = 0$ obtém-se o **Processo de Yule**.

Exemplo 5.0.6. Numa fila de espera do tipo $(M/M/s)$ com s servidores temos

$$q_{i,i+1} = \lambda \quad \text{e} \quad q_{i,i-1} = \begin{cases} i\mu, & 0 \leq i \leq s \\ s\mu, & i > s \end{cases}$$

onde λ é a taxa de chegada de novos clientes à fila e μ a taxa de serviço de cada servidor. Filas de espera serão tratadas com maior detalhe na secção 5.5.

5.1 Construção do Processo de Markov

Dada uma matriz de intensidades Q vamos de seguida construir um processo de Markov $X(t)$ com intensidades Q. Seja

$$\lambda_i = \sum_{i \neq j} q_{i,j}$$

a intensidade com que o processo deixa o estado $i \in E$. Vamos supor que $0 \leq \lambda_i < \infty$. Define-se

$$r_{i,j} = \frac{q_{i,j}}{\lambda_i}, \quad i \neq j,$$

que corresponde à "probabilidade" do processo entrar em j uma vez que saiu de i. Quando $\lambda_i = 0$ então o processo é constante no estado i e definimos $r_{i,i} = 1$. Usando as probabilidades $r_{i,j}$ definimos uma cadeia de Markov $(Y_n)_{n \geq 0}$ com matriz de transição $R = (r_{i,j})$. As variáveis Y_n representarão a n-ésima ocorrência do processo de Markov. Sejam $\tau_0, \tau_1, \tau_2, \ldots$ variáveis aleatórias iid com distribuição exponencial de parâmetro 1 e

$$T_n = \frac{\tau_{n-1}}{\lambda_{Y_{n-1}}} \quad \text{e} \quad W_n = T_1 + \cdots + T_n, \quad n = 1, 2, \ldots$$

Convenientemente definimos $W_0 = 0$. Finalmente, definimos o processo de Markov como

$$X(t) = Y_n, \quad W_n \leq t < W_{n+1}.$$

Os tempos T_n são independentes e têm distribuição exponencial de parâmetro $\lambda_{Y_{n-1}}$. Correspondem aos tempos entre ocorrências com intensidade $\lambda_{Y_{n-1}}$ de transitar do estado Y_{n-1} para o estado Y_n. Os tempos W_n são os tempos de espera até ocorrer Y_n.

Esta construção do processo fornece um algoritmo para simular um processo de Markov a partir da matriz de intensidades Q.

5.2 Equações Regressivas/Progressivas de Kolmogorov

A partir da matriz Q podemos determinar a distribuição do processo de Markov. Iremos de seguida deduzir umas equações diferenciais que relacionam as probabilidades de transição $p_{i,j}(t)$. Usando a equação de Chapman-Kolmogorov,

$$\begin{aligned} p_{i,j}(t+h) - p_{i,j}(t) &= \sum_k p_{i,k}(h) p_{k,j}(t) - p_{i,j}(t) \\ &= \sum_{k \neq i} p_{i,k}(h) p_{k,j}(t) + (p_{i,i}(h) - 1) p_{i,j}(t). \end{aligned}$$

Dividindo todos os termos por h e tomando o limite obtemos as **equações regressivas de Kolmogorov**,

$$p'_{i,j}(t) = \sum_{k \neq i} q_{i,k} p_{k,j}(t) - \lambda_i p_{i,j}(t),$$

uma vez que

$$\lim_{h \to 0} \frac{p_{i,i}(h) - 1}{h} = - \lim_{h \to 0} \sum_{k \neq i} \frac{p_{i,k}(h)}{h} = - \sum_{k \neq i} q_{i,k} = -\lambda_i.$$

Podemos escrever as equações regressivas de Kolmogorov em forma matricial,

$$P'(t) = QP(t),$$

onde $P(t) = (p_{i,j}(t))_{i,j\in E}$ é a matriz das probabilidades de transição. A equação diferencial tem de ser resolvida usando a condicional inicial $P(0) = I$, ou seja,

$$p_{i,j}(0) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

A solução da equação é dada a partir da matriz exponencial

$$P(t) = e^{tQ} \stackrel{def}{=} \sum_{n=0}^{\infty} \frac{(tQ)^n}{n!}.$$

Note-se que $\frac{d}{dt}e^{tQ} = Qe^{tQ}$. Uma forma de calcular a matriz exponencial é diagonalizando a matriz Q. Se a matriz Q for diagonalizável[2], então existe uma matriz invertível U, formada pelos vectores próprios de Q, tal que

$$\Delta = U^{-1}QU$$

é uma matriz diagonal, ou seja, todas as entradas fora da diagonal são zero. Suponhamos que $a_1, a_2, \ldots$ são os elementos da diagonal de Δ. Então $e^{t\Delta}$ é a matriz diagonal com elementos $e^{a_1 t}, e^{a_2 t}, \ldots$. Assim,

$$e^{tQ} = Ue^{t\Delta}U^{-1}.$$

Na dedução das equações regressivas de Kolmogorov partimos o intervalo $[0, t+h]$ da forma $[0, h]\cup[h, t+h]$. Usando a partição $[0, t]\cup[t, t+h]$ e repetindo a dedução obtêm-se as **equações progressivas de Kolmogorov**,

$$p'_{i,j}(t) = \sum_{k\neq j} p_{i,k}(t)q_{k,j} - \lambda_j p_{i,j}(t),$$

que reescritas na formal matricial,

$$P'(t) = P(t)Q.$$

Note-se que

$$P(t)Q = QP(t).$$

Exemplo 5.2.1 (Processo de nascimento)**.** Suponhamos que o processo de nascimento começa em zero, i.e., $X(0) = 0$. Definimos

$$p_n(t) := \mathbb{P}(X(t) = n | X(0) = 0)$$

[2]Um condição suficiente para Q ser diagonalizável é ter os seus valores próprios todos distintos.

Note-se que

$$p_n(0) = \begin{cases} 1, & n = 0 \\ 0, & n > 0 \end{cases}$$

Escrevendo as equações progressivas para o processo de nascimento obtemos o seguinte sistema para calcular $p_n(t)$,

$$\begin{aligned} p_0'(t) &= -\lambda_0 p_0(t) \\ p_n'(t) &= -\lambda_n p_n(t) + \lambda_{n-1} p_{n-1}(t), \quad n \geq 1 \end{aligned}$$

Da primeira equação obtém-se

$$p_0(t) = e^{-\lambda_0 t}$$

As restantes equações podem ser resolvidas recursivamente, ou seja, sabendo $p_{n-1}(t)$, podemos calcular $p_n(t)$ integrando a equação diferencial usando, por exemplo, o método de variação dos parâmetros. Obtém-se assim,

$$p_n(t) = \lambda_{n-1} e^{-\lambda_n t} \int_0^t e^{\lambda_n s} p_{n-1}(s)\, ds, \quad n = 1, 2, \ldots$$

Para que a solução $(p_n(t))_{n \geq 0}$ obtida seja uma distribuição de probabilidade é necessário que

$$\sum_{n=0}^{\infty} p_n(t) = 1 \quad \text{e} \quad p_n(t) \geq 0, \forall\, n \geq 0.$$

A segunda condição é trivialmente satisfeita. No entanto, pode-se demonstrar que a primeira condição é satisfeita se e só se

$$\sum_{n=0}^{\infty} \frac{1}{\lambda_n} = \infty,$$

ou seja, é necessário um tempo médio infinito para que a população seja infinita.

Quando as intensidades λ_n são todas distintas, então podemos escrever expressões relativamente simples para as probabilidades $p_n(t)$,

$$p_n(t) = \lambda_0 \cdots \lambda_{n-1} \sum_{k=0}^{n} \left(\prod_{j=0, j \neq k}^{n} \frac{1}{\lambda_j - \lambda_k} \right) e^{-\lambda_k t}, \quad n = 1, 2, \ldots$$

Em particular,

$$p_1(t) = \lambda_0 \left(\frac{1}{\lambda_1 - \lambda_0} e^{-\lambda_0 t} + \frac{1}{\lambda_0 - \lambda_1} e^{-\lambda_1 t} \right)$$

e

$$\begin{aligned} p_2(t) = \lambda_0 \lambda_1 \Bigg(\frac{1}{(\lambda_1 - \lambda_0)(\lambda_2 - \lambda_0)} e^{-\lambda_0 t} &+ \frac{1}{(\lambda_0 - \lambda_1)(\lambda_2 - \lambda_1)} e^{-\lambda_1 t} \\ &+ \frac{1}{(\lambda_0 - \lambda_2)(\lambda_1 - \lambda_2)} e^{-\lambda_2 t} \Bigg) \end{aligned}$$

Exemplo 5.2.2 (Processo de Markov com 2 estados)**.** Considere um processo de Markov com estados $\{0, 1\}$ e matriz de intensidade

$$Q = \begin{pmatrix} -\lambda & \mu \\ \mu & -\lambda \end{pmatrix}$$

onde $\lambda, \mu > 0$. As equações regressivas de Kolmogorov são,

$$\begin{pmatrix} p'_{0,0}(t) & p'_{0,1}(t) \\ p'_{1,0}(t) & p'_{1,1}(t) \end{pmatrix} = \begin{pmatrix} -\lambda & \mu \\ \mu & -\lambda \end{pmatrix} \begin{pmatrix} p_{0,0}(t) & p_{0,1}(t) \\ p_{1,0}(t) & p_{1,1}(t) \end{pmatrix}$$

Como $p_{i,1} = 1 - p_{i,0}$ vamos determinar primeiro $p_{i,0}$. As equações para $p_{i,0}$ são

$$\begin{aligned} p'_{0,0}(t) &= \lambda p_{1,0}(t) - \lambda p_{0,0}(t) \\ p'_{1,0}(t) &= \mu p_{0,0}(t) - \mu p_{1,0}(t) \end{aligned}$$

Fazendo a diferença obtém-se a seguinte equação diferencial linear,

$$\frac{d}{dt} \left(p_{0,0}(t) - p_{1,0}(t) \right) = -(\lambda + \mu) \left(p_{0,0}(t) - p_{1,0}(t) \right),$$

com condição inicial $p_{0,0}(0) - p_{1,0}(0) = 1 - 0 = 1$. Logo,

$$p_{0,0}(t) - p_{1,0}(t) = e^{-(\lambda+\mu)t}.$$

Como $p'_{0,0}(t) = -\lambda e^{-(\lambda+\mu)t}$ segue que

$$p_{0,0}(t) = 1 - \int_0^t \lambda e^{-(\lambda+\mu)t} = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t}.$$

De forma análoga determinam-se as restantes probabilidades de transição. Assim

$$P(t) = \begin{pmatrix} \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} & \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} \\ \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t} & \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t} \end{pmatrix}$$

Note-se que a distribuição limite

$$\lim_{t\to\infty} p_{i,j}(t) = \begin{cases} \frac{\mu}{\lambda+\mu}, & j=0 \\ \frac{\lambda}{\lambda+\mu}, & j=1 \end{cases},$$

não depende do estado de partida i.

5.3 Distribuição estacionária

Dada uma distribuição inicial $\pi(0)$ para $X(0)$, a distribuição $\pi(t)$ de $X(t)$ é obtida da seguinte forma,

$$\pi(t) = \pi(0)P(t).$$

Um caso especial é quando $\pi(t)$ é constante e independente de t. Nesse caso $\pi(t) = \pi(0)$ para todo $t \geq 0$. Este facto motiva a seguinte definição (compare com as cadeias de Markov).

Definição 5.3.1. Uma distribuição π é estacionária sse $\pi P(t) = \pi$ para todo $t \geq 0$.

Se um processo de Markov admite uma distribuição estacionária π , então o processo é estacionário com distribuição inicial π.

Lemma 5.3.1. *π é uma distribuição estacionária sse $\pi Q = 0$.*

Demonstração. Se $\pi P(t) = \pi$, então

$$0 = \frac{d}{dt}\pi P(t) = \pi P'(t) = \pi P(t)Q = \pi Q.$$

Por outro lado, se $\pi Q = 0$, então

$$\frac{d}{dt}\pi P(t) = \pi P'(t) = \pi Q P(t) = 0.$$

Logo, $\pi P(t)$ é constante para todo $t \geq 0$. Em particular, $\pi P(t) = \pi P(0) = \pi$. □

De seguida vamos relacionar a distribuição estacionária com a distribuição limite do processo.

Definição 5.3.2. Um processo de Markov é **irredutível** se para todo $i, j \in E$ existe uma sequência de estados $i = k_0, k_1, \ldots, k_n = j$ tal que

$$q_{k_{m-1},k_m} > 0, \quad m = 1, \ldots, n.$$

Teorema 5.3.2. *Se o processo de Markov é irredutível e admite uma distribuição estacionária π, então*

$$\pi_j = \lim_{t\to\infty} p_{i,j}(t).$$

Demonstração. Ver na bibliografia. □

Exemplo 5.3.3. Considere o seguinte modelo simplificado para prever o tempo numa região. O processo é composto por três estados 0 = sol, 1 = nublado e 2 = chuva. Sabe-se que

1. Permanece sol durante um período de 3 dias até ficar nublado;
2. Permanece nublado durante um período de 4 dias até começar a chover;
3. Após 1 dia de chuva regressa o sol.

A matriz de intensidades do processo é

$$Q = \begin{pmatrix} -1/3 & 1/3 & 0 \\ 0 & -1/4 & 1/4 \\ 1 & 0 & -1 \end{pmatrix}.$$

Para determinar a distribuição estacionária vamos resolver o sistema $\pi Q = 0$, ou seja,

$$\begin{aligned} -\frac{1}{3}\pi_1 + \pi_3 &= 0 \\ \frac{1}{3}\pi_1 - \frac{1}{4}\pi_2 &= 0 \\ \frac{1}{4}\pi_2 - \pi_3 &= 0 \end{aligned}$$

Usando $\pi_1 + \pi_2 + \pi_3 = 1$ determinamos a solução

$$\pi = \left(\frac{3}{8}, \frac{1}{2}, \frac{1}{8}\right).$$

Como o processo é irredutível (todos os estados comunicam entre si), sabemos que

$$\lim_{t\to\infty} p_{j,0}(t) = \frac{3}{8}.$$

Ou seja, a longo prazo, 3 em cada 8 dias vão ser dias de sol.

Exemplo 5.3.4. Considere um processo de nascimento e morte com parâmetros $\lambda_n > 0$ e $\mu_n > 0$. Recorde que a matriz de intensidades,

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Vamos determinar as distribuições estacionárias do processo, i.e., $\pi Q = 0$. Temos de resolver o sistema

$$\begin{cases} -\lambda_0\pi_0 + \mu_1\pi_1 = 0 \\ \lambda_{n-1}\pi_{n-1} - (\lambda_n+\mu_n)\pi_n + \mu_{n+1}\pi_{n+1} = 0, \quad n = 1, 2, \dots \end{cases}$$

Definimos $\theta_0 = 1$ e

$$\theta_n = \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}, \quad n = 1, 2, \dots$$

Usando indução mostramos que

$$\pi_n = \theta_n \pi_0$$

De facto, o caso $n = 0$ verifica-se. Supondo a hipótese de indução temos que

$$\begin{aligned} \pi_{n+1} &= \frac{(\lambda_n+\mu_n)\pi_n - \lambda_{n-1}\pi_{n-1}}{\mu_{n+1}} \\ &= \frac{(\lambda_n+\mu_n)\theta_n - \lambda_{n-1}\theta_{n-1}}{\mu_{n+1}}\pi_0 \\ &= \frac{\lambda_n\theta_n}{\mu_{n+1}}\pi_0 + \frac{\mu_n\theta_n - \lambda_{n-1}\theta_{n-1}}{\mu_{n+1}}\pi_0 \\ &= \theta_{n+1}\pi_0 \end{aligned}$$

uma vez que $\mu_{n+1}\theta_{n+1} = \lambda_n\theta_n$. Assim,

$$1 = \sum_{n=0}^{\infty} \pi_n = \pi_0 \sum_{n=0}^{\infty} \theta_n.$$

Desta equação determinamos π_0 e consequentemente π_n. Logo, supondo que

$$\sum_{n=0}^{\infty} \theta_n < \infty \tag{5.3.1}$$

o processo de nascimento e morte tem uma única distribuição estacionária dada por,

$$\pi_n = \frac{\theta_n}{\sum_{n=0}^{\infty} \theta_n}, \quad n = 0, 1, 2, \ldots$$

Segue do Teorema 5.3.2, que sob a condição (5.3.1), a distribuição estacionária é a distribuição limite do processo.

Exemplo 5.3.5. O processo de nascimento e morte pode ser usado para descrever o crescimento linear de uma população $X(t)$ sujeita com imigração,

$$\lambda_n = n\lambda + a \quad \text{e} \quad \mu_n = n\mu$$

onde $a > 0$ é uma taxa de imigração constante. Suponhamos que $\lambda < \mu$, ou seja, que a intensidade de nascimento per capita é menor que a intensidade de morte. Nestas condições o processo tem uma distribuição estacionária. Calculamos,

$$\begin{aligned}
\theta_n &= \frac{a(a+\lambda)\cdots(a+(n-1)\lambda)}{\mu^n n!} \\
&= \left(\frac{\lambda}{\mu}\right)^n \frac{(a/\lambda)(a/\lambda+1)\cdots(a/\lambda+(n-1))}{n!} \\
&= C_n^{\alpha+n-1} \left(\frac{\lambda}{\mu}\right)^n
\end{aligned}$$

onde $\alpha = a/\lambda$. Assim[3],

$$\sum_{n=0}^{\infty} \theta_n = \sum_{n=0}^{\infty} C_n^{\alpha+n-1} \left(\frac{\lambda}{\mu}\right)^n = \left(1 - \frac{\lambda}{\mu}\right)^{-\alpha}$$

uma vez que $\lambda/\mu < 1$. Portanto,

$$\pi_n = C_n^{\alpha+n-1} \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)^{\alpha}$$

Usando a distribuição estacionária, podemos calcular o valor médio (no limite) da população,

$$\lim_{t\to\infty} E(X(t)) = \sum_{n=0}^{\infty} n\pi_n = \frac{a}{\mu - \lambda}.$$

[3]

$$(1-x)^{-\alpha} = \sum_{n=0}^{\infty} C_n^{\alpha+n-1} x^n, \quad |x| < 1$$

5.4 Exercícios

Exercício 30. Considere o crescimento linear com imigração do Exemplo 5.3.5. Mostre que o valor médio (no limite) da população é $a/(\mu-\lambda)$ quando $\lambda<\mu$.

Exercício 31. Seja $(Y_n)_{n\geq 0}$ uma cadeia de Markov com estados $\{0,1\}$ e matriz de transição

$$P=\begin{pmatrix} 0 & 1 \\ 1-\alpha & \alpha \end{pmatrix}, \quad \alpha>0.$$

Considere um processo de Poisson $N(t)$ com parâmetro $\lambda>0$. Mostre que

$$X(t)=Y_{N(t)}$$

é um processo de nascimento e morte com dois estados $\{0,1\}$. Determine os parâmetros do processo λ_0 e μ_1 em função de α e λ.

Exercício 32. Considere um processo $X(t)$ de nascimento (sem morte) com parâmetros $\lambda_0=1$, $\lambda_1=3$ e $\lambda_2=2$. Supondo que $X(0)=0$, determine $\mathbb{P}(X(t)=j)$ para $j=0,1,2$.

Exercício 33. Considere um processo de nascimento e morte com parâmetros $\lambda_n=\lambda$ e $\mu_n=\mu n$ para $n=0,1,2,\dots$. Verifique que as probabilidades

$$p_{0,j}(t)=\frac{(\lambda p)^j e^{-\lambda p}}{j!}, \quad \text{onde} \quad p=\frac{1-e^{-\mu t}}{\mu}$$

satisfazem as equações progressivas de Kolmogorov ($i=0$).

Exercício 34. Um camião viaja entre Lisboa, Castelo Branco e Porto com a seguinte matriz de intensidades (número de viagens por mês)

$$Q=\begin{pmatrix} -4 & 2 & 2 \\ 3 & -4 & 1 \\ 5 & 0 & -5 \end{pmatrix}$$

Determine:

1. A fracção de tempo (a longo prazo) que o camião permanece em cada cidade.

2. O número médio de viagens por ano de Castelo Branco para Lisboa.

Exercício 35. Seja $X(t)$ um processo de nascimento e morte com estados $\{0,1,\dots,N\}$ e parâmetros $\lambda_n=\alpha(N-n)$ e $\mu_n=\beta n$ onde $\alpha,\beta>0$. Determine a distribuição estacionária.

Exercício 36. Uma fábrica tem 2 máquinas e um técnico responsável pela sua manutenção. Suponha que cada máquina trabalha durante um tempo médio de 12 dias até necessitar de manutenção e o técnico demora em média 2 dias para realizar a respectiva manutenção.

1. Supondo que a intensidade de avaria é proporcional ao número de máquinas em funcionamento, use um processo de nascimento e morte para modelar o número de máquinas em activo. Determine a matriz de intensidades que caracteriza o processo.

2. Sabendo que $p_{0,0}(t) = \frac{6}{25}e^{-\frac{5t}{6}} + \frac{18}{25}e^{-\frac{5t}{12}} + \frac{1}{25}$ calcule as probabilidades de transição $p_{1,0}(t)$ e $p_{2,0}(t)$ do processo.

3. Determine a distribuição estacionária.

4. Qual é fracção de tempo (longo prazo) de funcionamento simultâneo das máquinas.

5.5 Filas de Espera

Uma fila de espera consiste num sistema em que clientes chegam aleatoriamente e esperam para ser servidos. Uma fila de espera é caracterizada por:

- Um processo de chegada
- Uma distribuição do tempo de serviço
- Disciplina de serviço: número de servidores e política de serviço.

Iremos considerar filas de espera onde o processo de chegada é o processo de Poisson, os tempos de serviço são distribuídos exponencialmente, e a política de serviço é *first in first out* (FIFO).

Eis algumas medidas (assimptóticas) que são importantes no estudo de uma fila de espera:

- a distribuição limite do número de clientes na fila
- taxa de utilização dos servidores
- comprimento médio do número de pessoas no sistema e em fila de espera
- tempo médio de espera até ser servido

Usaremos a notação de Kendall para designar um tipo de fila de espera

$$A/B/s$$

onde A representa o tipo de distribuição dos tempos entre chegadas, B o tempo de serviço de cada servidor, e s o número de servidores existentes no sistema. Usualmente, A e B são do tipo

- G, distribuição geral
- M, distribuição sem memória
- E_k, de Erlang, distribuição gamma de ordem k

De seguida iremos estudar as filas de espera

$$M/M/s$$

onde os tempos entre chegada/serviço são exponencialmente distribuídos. Seja $X(t)$ o número de clientes no sistema no instante t, i.e., o número de clientes em fila de espera mais o número de clientes a serem servidos. Não é difícil verificar que $\{X(t) \colon t \geq 0\}$ é um processo de nascimento e morte. Vejamos o caso $s = 1$.

Fila de Espera $M/M/1$

Numa fila de espera do tipo $M/M/1$ temos um processo de nascimento e morte com parâmetros,

$$\lambda_n = \lambda \quad \text{e} \quad \mu_n = \mu$$

onde λ é a intensidade de chegadas (o número médio de clientes que chegam à fila de espera por unidade de tempo) e μ a intensidade de serviço (o número médio de clientes que cada servidor consegue servir por unidade de tempo). Queremos calcular a distribuição limite

$$\pi_n = \lim_{t\to\infty} \mathbb{P}(X(t) = n), \quad n = 0, 1, 2, \ldots$$

Como o processo é irredutível, basta calcular a distribuição estacionária do processo (ver Exemplo 5.3.4). Temos que

$$\theta_n = \left(\frac{\lambda}{\mu}\right)^n \quad \text{e} \quad \sum_{n=0}^{\infty} \theta_n = \begin{cases} \infty, & \lambda \geq \mu \\ \frac{\mu}{\mu-\lambda}, & \lambda < \mu \end{cases}.$$

Seja

$$\rho = \frac{\lambda}{\mu}$$

a **intensidade de tráfego** no sistema. Portanto, se $\rho < 1$ então o sistema tem uma única distribuição estacionária dada por

$$\pi_n = \frac{\theta_n}{\sum \theta_n} = (1-\rho)\rho^n, \quad n \geq 0.$$

Note-se que π_n é a distribuição geométrica onde ρ é a probabilidade do insucesso. A probabilidade de uma pessoa ser servida imediatamente quando chega à fila de espera é,

$$\pi_0 = 1 - \rho,$$

ou seja, $1-\rho$ representa a taxa de desocupação do servidor, ou a proporção de tempo em que o servidor se encontra inactivo. Usando a distribuição estacionária podemos calcular o número médio de pessoas no sistema

$$L = \sum_{n=0}^{\infty} n\pi_n = \frac{\rho}{1-\rho}.$$

Note-se que L cresce de forma ilimitada à medida que ρ se aproxima se 1. De forma análoga, podemos calcular o número médio de pessoas em fila de espera (excepto aquela que se encontra a ser servida),

$$L_0 = \sum_{n=1}(n-1)\pi_n = \frac{\rho^2}{1-\rho}.$$

Para determinar o tempo médio de espera definimos a variável aleatória T que representa o tempo de espera de uma pessoa para sair do sistema e a variável aleatória T_n que é o tempo de espera T de uma pessoa dado que existem n pessoas no sistema à frente desta. Note-se que T_n é a soma de $n+1$ variáveis aleatórias independentes com distribuição exponencial de parâmetro μ. Portanto, T_n tem distribuição gama com parâmetros $n+1$ e μ. Assim,

$$\begin{aligned}
\mathbb{P}(T \leq t) &= \sum_{n=0}^{\infty} \mathbb{P}(T_n \leq t)\pi_n \\
&= \sum_{n=0}^{\infty} \left(\int_0^t \frac{\mu^{n+1} x^n e^{-\mu s}}{n!}\, ds \right) (1-\rho)\rho^n \\
&= \mu(1-\rho) \int_0^t e^{-\mu s} \sum_{n=0}^{\infty} \frac{(\lambda x)^n}{n!}\, ds \\
&= \mu(1-\rho) \int_0^t e^{-\mu(1-\rho)s}\, ds \\
&= 1 - e^{(\mu-\lambda)t}
\end{aligned}$$

Logo, T tem distribuição exponencial de parâmetro $\mu - \lambda$. Assim, o tempo médio de espera (desde que entra até que sai) de uma pessoa no sistema é

$$W = E(T) = \frac{1}{\mu - \lambda}.$$

Verificamos que

$$L = \lambda W$$

conhecida como a **equação fundamental** das filas de espera. Se excluirmos o tempo de serviço, então podemos calcular de forma análoga o tempo médio de espera W_0 de uma pessoa na fila de espera. Também se verifica a equação fundamental,

$$L_0 = \lambda W_0.$$

Fila de Espera $M/M/s$

No caso geral, uma fila de espera com s servidores pode ser estudado usando um processo de nascimento e morte com parâmetros,

$$\lambda_n = \lambda \quad \text{e} \quad \mu_n = \begin{cases} n\mu, & n = 0, 1, 2, \dots, s-1 \\ s\mu, & n \geq s \end{cases},$$

onde μ é a intensidade de serviço individual de cada servidor. Portanto, a intensidade de serviço μ_n é igual à intensidade agregada dos servidores em funcionamento. Como temos s servidores,

$$\mu_n = \min\{n, s\}\mu, \quad n \geq 0.$$

No caso geral, definimos

$$\rho = \frac{\lambda}{s\mu}$$

e verifica-se que

$$\theta_n = \begin{cases} \frac{s^n}{n!}\rho^n, & n = 0, 1, 2, \dots, s-1 \\ \frac{s^s}{s!}\rho^n, & n \geq s \end{cases}$$

Assim

$$\sum_{n=0}^{\infty} \theta_n = \begin{cases} \frac{s^s}{s!}\frac{\rho^s}{1-\rho} + \sum_{n=0}^{s-1} \frac{s^n}{n!}\rho^n, & \rho < 1 \\ \infty, & \rho \geq 1 \end{cases}$$

Portanto, supondo que

$$\rho < 1$$

concluímos que o sistema tem uma única distribuição estacionária,

$$\pi_n = \theta_n \pi_0, \quad n \geq 0,$$

onde

$$\pi_0 = \left(\frac{s^s}{s!} \frac{\rho^s}{1-\rho} + \sum_{n=0}^{s-1} \frac{s^n}{n!} \rho^n \right)^{-1}.$$

Usando a distribuição estacionária podemos calcular as seguintes medidas de desempenho do sistema:

$$\begin{aligned}
L &= \lambda W \\
W &= W_0 + \frac{1}{\mu} \\
L_0 &= \sum_{n=s}^{\infty} (n-s)\pi_n = \frac{s^s}{s!} \frac{\rho^{s+1}}{(1-\rho)^2} \pi_0 \\
W_0 &= \frac{L_0}{\lambda}
\end{aligned}$$

5.6 Exercícios

Exercício 37. Clientes chegam a uma mercearia com uma intensidade de 5 por hora. O empregado demora em média 10 minutos a atender cada cliente. Determine a probabilidade de encontrar, a longo prazo, 2 ou mais clientes na mercearia.

Exercício 38. Para uma fila de espera $M/M/1$ faça o gráfico a taxa de utilização do servidor $1 - \pi_0$ e do número médio L de pessoas no sistema como funções da intensidade de tráfego $\rho < 1$.

Exercício 39. Um supermercado dispõe de duas caixas de *checkout*. Uma caixa tem um operador e a outra é automática. A caixa com operador tem um tempo médio de serviço de 30 segundos, enquanto que a caixa automática tem um tempo médio de serviço de 50 segundos. Supondo que os clientes chegam às caixas com uma intensidade de 1 cliente por minuto, compare os comprimentos das filas de espera das duas caixas.

Exercício 40. Considere uma fila de espera $M/M/2$. Determine o tempo médio de espera W quando $\lambda = 2$ e $\mu = 1.2$. Compare com o tempo médio de espera numa fila $M/M/1$ com $\lambda = 1$ e $\mu = 1.2$. Explique a diferença dos tempos de espera.

Exercício 41. Suponha que clientes chegam a uma repartição das finanças seguindo um processo de Poisson com intensidade de 5 clientes por hora. A repartição tem um único funcionário que demora, em média, 10 minutos a atender cada cliente. Quando o número de clientes na repartição excede um certo número N o funcionário encerra a repartição. Determine o menor N tal que, a longo prazo, a probabilidade de o funcionário encerrar a repartição não exceda 0.01.

Chapter 6

Martingalas

O termo Marginala tem as suas origens nos jogos de apostas e descreve o conceito de jogo honesto. As martingalas encontram aplicações em diversas áreas, nomeadamente em matemática financeira, na modelação de activos financeiros.

Definição 6.0.1. Uma sucessão de variáveis aleatórias $(X_n)_{n\geq 1}$ é uma **martingala** se

1. X_n é integrável, ou seja, $E(|X_n|) < \infty \;\forall\, n \geq 1$.
2. Para todo $n \geq 1$,

$$E(X_{n+1}|X_1 = x_1, \ldots, X_n = x_k) = x_n.$$

Da definição de esperança condicional obtemos as seguintes propriedades.

Proposição 6.0.1.

1. *Uma martingala tem valor esperado constante, i.e.,*

$$E(X_n) = E(X_1), \quad \forall\, n \geq 1.$$

2. *Se $k \leq n$, então*

$$E(X_n|X_1 = x_1, \ldots, X_k = x_k) = x_k.$$

Demonstração. Faremos apenas a demonstração da primeira propriedade no caso particular em que a martingala X_n é contínua e qualquer vector de probabilidade do processo admite uma densidade de probabilidade conjunta. Não faremos a prova da segunda propriedade. Esta usa a propriedade da torre da esperança condicional que não foi abordada neste curso. Segue da lei de probabilidade total que

$$E(X_{n+1}) = \int_{\mathbb{R}^n} E(X_{n+1}|X_1 = x_1, \ldots, X_n = x_n) f_{X_1,\ldots,X_n}(\boldsymbol{x})\, d\boldsymbol{x}$$

onde $\boldsymbol{x} = (x_1, \ldots, x_n)$ e $d\boldsymbol{x} = dx_1 \cdots dx_n$. Pela definição de martingala obtemos,

$$E(X_{n+1}) = \int_{\mathbb{R}^n} x_n f_{X_1,\ldots,X_n}(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\mathbb{R}} x_n f_{X_n}(x_n)\, dx_n = E(X_n).$$

Portanto, $E(X_n) = E(X_1)$ para todo $n \geq 1$. $\square$

Exemplo 6.0.2 (Passeio aleatório)**.** Sejam $\xi_1, \xi_2, \ldots$ variáveis aleatórias independentes, $E(\xi_n|) < \infty$ e $E(\xi_n) = 0$ para todo $n \geq 1$. A soma

$$X_n = \xi_1 + \xi_2 + \cdots + \xi_n\,,$$

é uma martingala. De facto, a primeira condição da definição de martingala é fácil verificar usando a desigualdade triangular e a aditividade da esperança condicional. Relativamente à segunda condição temos que,

$$\begin{aligned} E(X_{n+1}|X_1 = x_1, \ldots, X_n = x_n) &= E(\xi_{n+1} + X_n|X_1 = x_1, \ldots, X_n = x_n) \\ &= E(\xi_{n+1}) + E(X_n|X_1 = x_1, \ldots, X_n = x_n) \\ &= x_n \end{aligned}$$

Exemplo 6.0.3. Sejam $\xi_1, \xi_2, \ldots$ variáveis aleatórias independentes, $E(\xi_n|) < \infty$ e $E(\xi_n) = 1$ para todo $n \geq 1$. O produto

$$X_n = \prod_{k=1}^{n} \xi_k\,,$$

é uma martingala.

No jogo da moeda, dois jogadores atiram uma moeda perfeita ao ar e observam o resultado. Se sair coroa, o primeiro jogador dá 1 euro ao segundo jogador. Se sair cara, o segundo jogador dá 1 euro ao primeiro jogador. Os jogadores continuam a jogar fazendo sucessivos lançamentos da moeda. Ao fim do n-ésimo lançamento, ganho total do primeiro jogador é,

$$X_n = \xi_1 + \cdots + \xi_n\,,$$

onde ξ_n são variáveis aleatórias iid que tomam valores ± 1 com igual probabilidade ($\xi_i = 1$ se sair cara, e $\xi_i = -1$ se sair coroa no i-ésimo lançamento). Concluímos do Exemplo 6.0.2 que X_n é uma martingala. O jogo da moeda perfeita é um exemplo de um **jogo justo**, ou seja, que perante a informação acumulada até ao n-ésimo lançamento, espera-se que no lançamento seguinte o ganho total não se altere.

Nem todos os jogos são justos, no sentido em que podem existir jogos mais favoráveis que outros. Um exemplo de um jogo favorável seria o do lançamento de dois dados em que o apostador ganharia se a soma dos dados fosse maior ou igual a 6.

Definição 6.0.2. Uma sucessão de variáveis aleatórias $(X_n)_{n\geq 1}$ é uma **submartingala** se

1. X_n é integrável, ou seja, $E(|X_n|) < \infty \; \forall\, n \geq 1$.

2. Para todo $n \geq 1$,

$$E(X_{n+1}|X_1 = x_1, \ldots, X_n = x_k) \geq x_n.$$

Portanto, uma submartingala representa um jogo favorável. Invertendo a desigualdade na definição anterior, temos a definição de **supermartingala** que representa um jogo desfavorável.

6.1 Estratégias

Considere um jogo de apostas em que em cada jogada é possível apostar uma certa quantidade que pode ou não perder com certa probabilidade. Ou seja, considere uma sucessão de variáveis aleatórias $(\xi_n)_{n\geq 1}$ tais que $E(|\xi_n|) < \infty$ representam o valor que pode ganhar ou perder em cada jogada. Se as suas apostas em cada jogada são igual à unidade então o ganho total no final de n jogadas é

$$X_n = \xi_1 + \cdots + \xi_n\,.$$

Suponhamos que decide em cada jogada alterar a sua aposta com base nas jogadas anteriores. Ou seja, na jogada n faz uma aposta α_n que é uma função dos ganhos anteriores $(X_1, \ldots, X_{n-1})$. É conveniente definir $\alpha_1 = 1$, uma vez que não é possível fazer uma aposta "informada" logo na primeira jogada. A sucessão das apostas $(\alpha_n)_{n\geq 1}$ é chamada uma **estratégia** relativa a $(X_n)_{n\geq 1}$. Então o ganho total até à n-ésima jogada usando uma estratégia é dado por,

$$Z_n = \alpha_1\xi_1 + \cdots + \alpha_n\xi_n\,.$$

Numa situação real, as apostas α_n terão de ser funções limitadas, uma vez que nenhum jogador pode dispor de fundos ilimitados. Portanto, consideraremos apenas **estratégias limitadas**, i.e., a função $\alpha_n(x_1, \ldots, x_{n-1})$ é limitada para todo $n \geq 1$.

Note-se que dada uma sucessão $(X_n)_{n\geq 1}$ de variáveis aleatórias, podemos sempre decompor X_n como uma soma de incrementos, $X_n = \xi_1 + \cdots + \xi_n$ onde $\xi_k = X_k - X_{k-1}$ para $k \geq 2$ e $\xi_1 = X_1$. Portanto, a discussão anterior aplica-se a qualquer sucessão de variáveis aleatórias integráveis.

A seguinte proposição demonstra que é impossível através de uma estratégia limitada alterar um jogo a nosso favor.

Proposição 6.1.1. *Se X_n é uma martingala e $(\alpha_n)_{n\geq 1}$ é uma estratégia limitada relativa a X_n, então $(Z_n)_{n\geq 1}$ é uma martingala.*

Demonstração. Seja $C_n = \max_{1\leq k\leq n} |\alpha_n|$. Note-se que $C_n < \infty$ por hipótese. Então,

$$E(|Z_n|) = E(|\alpha_1\xi_1 + \cdots + \alpha_n\xi_n|) \leq C_n \sum_{i=1}^{n} E(|\xi_i|) < \infty\,.$$

Verifiquemos agora a segunda condição na definição de martingala. Em primeiro lugar, note-se que

$$\begin{aligned} E(Z_{n+1}|Z_1 = z_1, \ldots, Z_n = z_n) &= E(Z_n + \alpha_{n+1}\xi_{n+1}|Z_1 = z_1, \ldots, Z_n = z_n) \\ &= z_n + E(\alpha_{n+1}\xi_{n+1}|Z_1 = z_1, \ldots, Z_n = z_n). \end{aligned}$$

Uma vez que α_{n+1} é função de $(Z_1, \ldots, Z_n)$ podemos retirar para fora da esperança condicional. Por outro lado,

$$E(\xi_{n+1}|Z_1 = z_1, \ldots, Z_n = z_n) = E(X_{n+1} - X_n|Z_1 = z_1, \ldots, Z_n = z_n) = 0$$

e portanto

$$E(\alpha_{n+1}\xi_{n+1}|Z_1 = z_1, \ldots, Z_n = z_n) = 0.$$

Logo, $E(Z_{n+1}|Z_1 = z_1, \ldots, Z_n = z_n) = z_n$. □

Exemplo 6.1.2 (A estratégia da martingala)**.** Considere novamente o jogo do lançamento de uma moeda perfeita. O ganho acumulado é dado por

$$X_n = \xi_1 + \cdots + \xi_n\,,$$

onde $\xi_1, \xi_2, \ldots$ são variáveis aleatórias iid tais que $\xi_n = \pm 1$ com igual probabilidade.

A **martingala** é um tipo de estratégia usada nos jogos de apostas que teve origem em França no século XVIII. A estratégia consiste em cada jogada duplicar a aposta anterior caso o resultado tenha sido desfavorável. Ou seja,

$$\alpha_n = \begin{cases} 2^{n-1} & \text{se } \xi_1 = \xi_2 = \cdots = \xi_{n-1} = -1 \\ 0 & \text{caso contrário} \end{cases}, \quad n \geq 2$$

com $\alpha_1 = 1$. Portanto, ao fim de n jogadas temos um ganho acumulado

$$Z_n = \alpha_1 \xi_1 + \cdots + \alpha_n \xi_n \,,$$

Segue da proposição anterior que Z_n é uma martingala. Das propriedades da martigala temos que

$$E(Z_n) = E(Z_1) = 0, \quad \forall\, n \geq 1.$$

Considere agora o menor n tal que o resultado da n-ésima jogada foi favorável, ou seja, a variável aleatória

$$\tau = \min\{n \in \mathbb{N} \colon \xi_n = 1\} \;.$$

Note-se que o acontecimento $\{\tau = n\} \in \mathcal{F}_n$. Dizemos que τ é um **tempo de paragem**. Usando o tempo de paragem podemos escrever o ganho acumulado,

$$Z_n = \begin{cases} -1 - 2 - \ldots - 2^{n-1} & \text{se } n < \tau \\ 1 & \text{se } n \geq \tau \end{cases}.$$

Um vez que $\mathbb{P}(\tau < \infty) = 1$ temos que $Z_\tau = 1$. Logo

$$E(Z_\tau) = 1 \,.$$

Por outro lado,

$$E(Z_{\tau-1}) = \sum_{n=2}^{\infty} (-1 - \ldots - 2^{n-2}) P(\tau = n) = \sum_{n=2}^{\infty} \frac{1 - 2^{n-1}}{2^n} = -\infty \,.$$

Ou seja, um jogador pode ir acumulando pequenos ganhos ao longo do tempo dando a ilusão de que a sua estratégia é ganhadora. No entanto, a possibilidade remota de ter um prejuízo catastrófico contrabalança esses ganhos e pode levar o jogador à falência rapidamente. Portanto, a estratégia da martingala é apenas bem sucedida se o jogador dispõe de recursos e tempo ilimitados.

6.2 Tempos de paragem

Motivados pelo exemplo anterior, fazemos a seguinte definição.

Definição 6.2.1. Um **tempo de paragem** relativo a uma sucessão $(X_n)_{n\geq 1}$ de variáveis aleatórias é uma variável aleatória $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ definida,

$$\tau = \min\{n \geq 1\colon X_n \in B\}$$

onde $B \subset \mathbb{R}$ é um Boreliano de $\mathbb{R}$ (por exemplo um intervalo).

É também usual chamar-se τ um **tempo de primeira entrada** de X_n em B. Dado um tempo de paragem τ podemos definimos

$$X_\tau = \sum_{n=1}^{\infty} X_n \cdot \chi_{\{\tau=n\}} \, .$$

que designamos por X_n **no momento exacto do tempo de paragem**. Definimos também uma nova sucessão,

$$X_n^\tau = \begin{cases} X_n & \text{se } n \leq \tau \\ X_\tau & \text{se } n > \tau \end{cases} ,$$

que representa a **sucessão** $(X_n)_{n\geq 1}$ **parada no instante** τ. Por exemplo, se $\tau = 3$ então

$$X_1^\tau = X_1, \quad X_2^\tau = X_2, \quad X_3^\tau = X_3, \quad X_4^\tau = X_3, \quad \text{etc.}$$

Vimos na secção anterior que não é possível através de uma estratégia limitada alterar um jogo justo a nosso favor. A seguinte proposição mostra um resultado semelhante no contexto dos tempos de paragem.

Proposição 6.2.1. *Seja X_n uma martingala e τ um tempo de paragem relativo a X_n. Então X_n^τ é uma martingala.*

Demonstração. Seja

$$\alpha_n = \begin{cases} 1 & \text{se } n \leq \tau \\ 0 & \text{se } n > \tau \end{cases} .$$

Note-se que $\alpha_n = 0$ se $X_k \in B$ para algum $k < n$, e $\alpha_n = 1$ caso contrário. Portanto, α_n é uma função do vector $(X_1, \ldots, X_{n-1})$. Claramente, α_n é limitada, portanto é uma estratégia limitada relativa a X_n. Assim,

$$Z_n = \alpha_1 \xi_1 + \cdots + \alpha_n \xi_n$$

onde $\xi_1 = X_1$ e $\xi_k = X_k - X_{k-1}$ para $k \geq 2$ é uma martingala pela Proposição 6.1.1. No entanto, é fácil verificar que $Z_n = X_n^\tau$. Logo X_n^τ é uma martingala. □

6.3 Teorema de paragem opcional

Vimos na secção anterior que se X_n é uma martingala e τ um tempo de paragem então X_n^τ é também uma martingala. Adicionalmente,

$$E(X_n^\tau) = E(X_n) = E(X_1)\,,$$

uma vez que martingalas têm valor esperado constante. No entanto, se considerarmos X_n no momento exacto do tempo de paragem, i.e., X_τ, então nem sempre $E(X_\tau) = E(X_1)$. Como exemplo veja-se o caso da estratégia de martingala (ver Exemplo 6.1.2). O teorema de paragem opcional estabelece condições suficientes para que a propriedade da martingala se estenda a tempos de paragem, i.e.,

$$E(X_\tau) = E(X_1)\,.$$

Antes de enunciar o teorema, necessitamos de uma definição técnica.

Definição 6.3.1. Uma sucessão de variáveis aleatórias $(Y_n)_{n\geq 1}$ diz-se **uniformemente integrável** sse existir uma função integrável $g : \Omega \to [0,\infty]$ tal que $|Y_n| \leq g$ para todo $n = 1, 2, \ldots$.

Estamos em condições de enunciar o teorema.

Teorema 6.3.1 (da paragem opcional de Doob)**.** *Seja $(X_n)_{n\geq 1}$ uma martingala e τ um tempo de paragem. Se $P(\tau < \infty) = 1$ e $(X_n^\tau)_{n\geq 1}$ é uniformemente integrável então*

$$E(X_\tau) = E(X_1).$$

Ou seja, não é possível alterar um jogo a nosso favor usando uma estratégia de paragem que satisfaça as condições do teorema. Antes de comentarmos acerca das condições do teorema vejamos a sua demonstração.

Demonstração. Uma vez que $P(\tau < \infty) = 1$ segue que

$$\lim_{n\to\infty} X_n^\tau = X_\tau\,, \quad P\text{-q.c.}$$

Logo, segue do Teorema da convergência dominada que

$$E(X_\tau) = \lim_{n\to\infty} E(X_n^\tau) = \lim_{n\to\infty} E(X_1) = E(X_1)\,.$$

□

Das duas condições do teorema da paragem opcional a mais difícil de verificar é X_n^τ uniformemente integrável. Contudo existem alguns casos em que podemos verificar essa condição com facilidade.

Proposição 6.3.2. *A martingala $(X_n^\tau)_{n\geq 1}$ é uniformemente integrável se alguma das seguintes condições se verificar:*

1. *X_n toma um número finito de valores, i.e., o conjunto dos estados é finito.*

2. *A sucessão $(X_n^\tau)_{n\geq 1}$ é limitada, ou seja, existe um $M \geq 0$ tal que $|X_n^\tau| \leq M$ para todo $n \geq 1$.*

3. *O tempo de paragem é limitado, i.e., $P(\tau \leq k) = 1$ para algum $k \geq 1$.*

Demonstração.

1. Trivial, uma vez que $|X_n|$ é majorado pelo maior dos seus estados.

2. Segue directamente da Definição 6.3.1.

3. Se $\tau \leq k$ P-q.c. então

$$|X_n^\tau| \leq \max\{|X_1|, \ldots, |X_k|\} .$$

Logo, X_n^τ é uniformemente integrável.

□

Exemplo 6.3.3 (O jogo da ruína)**.** Retomemos o jogo do lançamento de uma moeda perfeita. Sejam $\xi_1, \xi_2, \ldots$ variáveis aleatórias iid tais que $\xi_n = \pm 1$ com igual probabilidade. Então o passeio aleatório,

$$X_n = \xi_1 + \ldots + \xi_n ,$$

é uma martingala. Considere-se o tempo de paragem,

$$\tau = \min\{n \in \mathbb{N} \colon X_n \in \{-a, b\}\} ,$$

onde $0 < a, b$. Ou seja, o jogo termina quando o jogador obtiver um prejuízo de a euros ou um lucro de b euros. Desejamos calcular a probabilidade $p_a = P(X_\tau = -a)$, isto é, a probabilidade de o jogador acabar o jogo com um prejuízo de a euros em vez de um lucro de b euros.

Suponhamos que as condições de aplicabilidade do teorema da paragem opcional se verificam,

- $P(\tau < \infty) = 1$

- X_n^τ é uniformemente integrável

Então, do teorema da paragem opcional segue que $E(X_\tau) = E(X_1) = 0$. Por outro lado,

$$0 = E(X_\tau) = -ap_a + b(1 - p_a)\,.$$

Logo,

$$p_a = \frac{b}{b+a}\,.$$

Portanto, se o jogador tiver a euros e estiver disposto jogar enquanto não os perder, então a probabilidade de ruína é,

$$\lim_{b\to\infty} p_a = \lim_{b\to\infty} \frac{b}{b+a} = 1\,.$$

Verifiquemos agora as condições do teorema.

- Uma vez que $X_n^\tau \in [-a, b]$ logo X_n^τ é uma sucessão limitada, portanto uniformemente integrável.
- Demonstremos $P(\tau < \infty) = 1$. O jogo da ruína pode ser visto como uma cadeia de Markov no espaço de estados $S = \{-a, \dots, b\}$ onde os estados $-a$ e b são absorventes, os restantes transientes e probabilidades de transição $P_{i,i+1} = P_{i,i-1} = 1/2$ com $i \in S \setminus \{-a, b\}$. Neste contexto, τ é o tempo de primeira entrada no conjunto fechado $\{-a, b\}$. Assim, $\mathbb{P}(\tau < \infty)$ é a probabilidade de absorção de partindo do estado 0 a cadeia ser absorvida em $\{-a, b\}$. Podemos calcular esta probabilidade resolvendo o sistema da Proposição 3.9.3. Deixa-se como exercício.

Proposição 6.3.4 (Equação de Wald)**.** *Seja X_n uma sucessão de variáveis aleatórias iid com $E(|X_n|) \leq M$ para algum $M \geq 0$ e τ um tempo de paragem relativo a X_n tal que $P(\tau < \infty) = 1$ e $E(\tau) < \infty$. Então*

$$E(\sum_{i=1}^{\tau} X_i) = \mu E(\tau)\,,$$

onde $\mu = E(X_n)$ para todo $n = 1, 2, \dots$.

Demonstração. Seja $Z_n = \sum_{i=1}^n X_i - n\mu$. É fácil ver que Z_n é uma martingala (Exercício 43). Para aplicarmos o teorema da paragem opcional temos de verificar que Z_n^τ é uniformemente integrável. Este facto não é trivial (a menos que X_n esteja nas condições da Proposição 6.3.2) e portanto omitimos a prova. Aplicando o teorema temos que

$$0 = E(Z_\tau) = E(\sum_{i=1}^{\tau} X_i - \tau\mu) = E(\sum_{i=1}^{\tau} X_i) - \mu E(\tau)\,.$$

□

Exemplo 6.3.5. Seja $X_1, X_2, \ldots$ variáveis aleatórias iid tais que $X_n \in \{0, 1\}$ e $\alpha = P(X_n = 1)$. Forme-se a soma,

$$S_n = X_1 + \cdots + X_n \,,$$

e considere-se o tempo de paragem,

$$\tau_k = \min \{n \in \mathbb{N} \colon S_n = k\} \ .$$

Uma vez que $P(\tau_k < \infty) = 1$ temos pela equação de Wald que

$$k = E(S_{\tau_k}) = \alpha E(\tau_k) \,.$$

Logo, $E(\tau_k) = \frac{k}{\alpha}$.

Exemplo 6.3.6. Retomando o jogo da ruína (Exemplo 6.3.3), desejamos calcular $E(\tau)$. A equação de Wald não ajuda uma vez que $E(X_\tau) = 0$ e $\mu = 0$. No entanto se considerarmos uma nova sucessão $Z_n = X_n^2 - n$ é possível calcular $E(\tau)$. Note-se que Z_n é uma martingala (Exercício 44), logo Z_n^τ é também uma martingala. Assim, pelo teorema da paragem opcional

$$0 = E(Z_1) = E(Z_\tau) = E(X_\tau^2 - \tau) \,.$$

Portanto,

$$E(X_\tau^2) = E(\tau) \,.$$

Mas como, $E(X_\tau^2) = a^2 \frac{b}{b+a} + b^2 (1 - \frac{b}{b+a}) = ab$, logo,

$$E(\tau) = ab \,.$$

O teorema da paragem opcional tem uma versão no caso das supermartingalas (submartingalas).

Teorema 6.3.7. *Seja $(X_n)_{n \geq 1}$ uma supermartingala (submartingala) e τ um tempo de paragem. Se $P(\tau < \infty) = 1$ e $(X_n^\tau)_{n \geq 1}$ é uniformemente integrável então $E(X_\tau) \leq E(X_1)$ ($E(X_\tau) \geq E(X_1)$).*

6.4 Convergência de martingalas

O resultado que se segue estabelece a convergência de uma martingala com probabilidade 1 desde que a sucessão $E(|X_n|)$ seja limitada. A sua demonstração pode ser consultada num dos livros da bibliografia recomendada.

Teorema 6.4.1. *Seja X_n uma martingala. Se existe um $M > 0$ tal que $E(|X_n|) \leq M$ para todo $n \geq 1$, então existe uma variável aleatória X_∞ integrável tal que*

$$X_\infty = \lim_{n\to\infty} X_n \,, \quad P - q.c.$$

Como corolário tem-se o seguinte resultado.

Corolário 6.4.2. *Se X_n é uma martingala não-negativa, i.e., $X_n \geq 0$, então existe uma variável aleatória X_∞ tal que*

$$X_\infty = \lim_{n\to\infty} X_n \,, \quad P - q.c.$$

Demonstração. Uma vez que X_n é não-negativa tem-se

$$E(|X_n|) = E(X_n) = E(X_1) \,, \quad \forall n \geq 1.$$

O resultado segue do teorema anterior. □

Exemplo 6.4.3. Considere novamente o passeio aleatório,

$$X_n = \xi_1 + \ldots + \xi_n \,,$$

onde ξ_n é uma sucessão de variáveis aleatórias iid tais que $\xi_n = \pm 1$ com igual probabilidade. Suponha que o jogador dispõe de crédito limitado, a euros, e só decide abandonar o jogo quando o seu prejuízo atingir esse limite,

$$\tau = \min \{n \geq 1 : X_n = -a\} \ .$$

Vejamos que $\mathbb{P}(\tau < \infty) = 1$ usando o teorema da convergência das martingalas. Em primeiro lugar, a martingala $X_n^\tau + a$ é não-negativa. Logo, segue do corolário anterior que

$$X_\tau = \lim_{n\to\infty} X_n^\tau = -a \,, \quad P - \text{q.c.}$$

Uma vez que

$$|X_{n+1} - X_n| = |\xi_{n+1}| = 1 \,, \quad \forall n \geq 1$$

então τ é finito P-q.c., caso contrário a sucessão X_n^τ não seria convergente P-q.c. para $-a$.

Conclui-se assim que com probabilidade 1, o jogador será obrigado a abandonar o jogo ao fim de um número finito de jogadas.

Exemplo 6.4.4. Sejam $X_1, X_2, \ldots$ variáveis aleatórias iid tais que $X_n \in \{0, 2\}$ com igual probabilidade. Uma vez que $E(X_n) = 1$ então

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

é uma martingala. Como $Z_n \geq 0$ então aplicando o corolário anterior tem-se

$$Z_\infty = \lim_{n\to\infty} Z_n\,, \quad P - \text{q.c.}$$

De facto, $Z_\infty = 0$, P-q.c.

6.5 Exercícios

Exercício 42. Mostre que a sucessão definida no Exemplo 6.0.3 é uma martingala.

Exercício 43. Seja $(X_n)_{n\geq 1}$ uma sucessão de variáveis aleatórias iid e $E(|X_n|) < \infty$ para todo $n \geq 1$. Mostre que

$$Z_n = \sum_{k=1}^{n} X_k - n\mu$$

é uma martingala, onde $\mu = E(X_1)$.

Exercício 44. Considere o passeio aleatório $X_n = \xi_1 + \cdots + \xi_n$ onde a sucessão $(\xi_n)_{n\geq 1}$ é iid e $\mathbb{P}(\xi_n = \pm 1) = \frac{1}{2}$. Mostre que

$$Z_n = X_n^2 - n$$

é uma martingala.

Exercício 45. Seja $X_n = \xi_1 + \ldots + \xi_n$ o passeio aleatório com $(\xi_n)_{n\geq 1}$ iid, $\xi_n = \pm 1$ com igual probabilidade. Mostre que

$$Z_n = (-1)^n \cos(\pi X_n)$$

é uma martingala.

Exercício 46. Seja $X_0 = 0$ e

$$X_n = X_{n-1} + \xi_n - 1\,, \quad n = 1, 2, \ldots$$

onde $(\xi_n)_{n\geq 1}$ é uma sucessão de variáveis aleatórias iid que têm uma distribuição de Poisson com valor esperado $\mu > 0$. Determine os valores de μ para os quais a sucessão $(X_n)_{n\geq 1}$ é uma martingala, submartingala ou supermartingala.

Exercício 47. Mostre que $\mathbb{P}(\tau < \infty) = 1$ no jogo da ruína.

Exercício 48. Sejam $(\xi_n)_{n\geq 1}$ variáveis aleatórias iid tais que $\xi_n = \pm 1$ com probabilidade $P(\xi_n = 1) = p$ e $P(\xi_n = -1) = q$ onde $p \neq q$. Considere o passeio aleatório,

$$X_n = \xi_1 + \ldots + \xi_n \,,$$

e o tempo de paragem

$$\tau = \min \{n \geq 1 : X_n \in \{-a, b\}\} \ ,$$

onde $a, b > 0$. Mostre que:

1. A sucessão $Z_n = (q/p)^{X_n}$ é uma martingala.

2. $$P(X_\tau = b) = \frac{1 - (q/p)^{-a}}{(q/p)^b - (q/p)^{-a}}$$

Exercício 49. Seja $X_n = \xi_1 + \ldots + \xi_n$ o passeio aleatório onde $(\xi_n)_{n\geq 1}$ são iid tais que $\xi_n = \pm 1$ com igual probabilidade. Supondo que $k \in \mathbb{N}$:

1. Mostre que $Z_n = (-1)^n \cos(\pi(X_n + k))$ é uma martingala.

2. Calcule $E((-1)^\tau)$ onde τ é o tempo de paragem

$$\tau = \min \{n \geq 1 : |X_n| = k\} \ .$$

Exercício 50. Sejam $X_1, X_2, \ldots$ variáveis aleatórias iid tais que $X_n \in \{-1, 1\}$ para todo $n = 1, 2, \ldots$. Considere o tempo de paragem

$$\tau = \min\{n \geq 1 \colon X_1 + \cdots + X_n = 1\} \, .$$

Determine $E(\tau)$.

Exercício 51 (difícil)**.** Seja $(X_n)_{n\geq 1}$ uma sucessão de variáveis aleatórias e τ um tempo de paragem relativo a X_n tal que para algum $k \in \mathbb{N}$ e algum $\epsilon > 0$,

$$\mathbb{P}(\tau \leq n + k \,|X_1 = x_1, \ldots, X_n = x_n) > \epsilon \,, \quad \forall\, n \geq 1.$$

Mostre que $\mathbb{P}(\tau < \infty) = 1$.

Chapter 7

Movimento Browniano

O movimento Browniano (ou processo de Wiener) é um processo estocástico usado para descrever o movimento descrito por uma partícula imersa num líquido quando esta colide com as moléculas do líquido. O movimento Browniano é amplamente utilizado em matemática financeira.

Definição 7.0.1. Um processo estocástico a tempo contínuo $\{B(t) : t \in [0, \infty[\}$ é um **movimento Browniano** se:

1. $B(0) = 0$,

2. tem incrementos independentes, i.e., para quaisquer $0 \leq t_0 < t_1 < \cdots < t_n$ as variáveis aleatórias

$$B(t_1) - B(t_0), \ldots, B(t_n) - B(t_{n-1})$$

   são independentes.

3. tem incrementos estacionários e normalmente distribuídos, i.e., para quaisquer $t, h > 0$ o incremento $B(t+h) - B(t)$ tem distribuição normal com média 0 e variância h.

4. as trajectórias $t \mapsto B(t)$ são contínuas $\mathbb{P}$-q.c.

Dado $x \in \mathbb{R}$, designamos por $B_x(t) = x + B(t)$ o movimento Browniano com início em x. Note-se que a lei de probabilidade de $B_x(t)$ é a mesma de $B(t)$.

Proposição 7.0.1.

1. $E(B(t)) = 0$

2. $V(B(t)) = t$

3. $Cov(B(t), B(s)) = \min\{t, s\}$

Demonstração. Exercício. $\square$

Definimos,

$$W(t) = x + \mu t + \sigma B(t).$$

Segue da proposição anterior que

$$E(W(t)) = x + \mu t \quad \text{e} \quad V(W(t)) = \sigma^2 t.$$

Os parâmetros μ e σ são denominados por *drift* e difusão, respectivamente. Ao processo estocástico $W(t)$ chamamos o **processo de Wiener** com parâmetros (x, μ, σ).

A proposição seguinte mostra como é possível determinar a lei de probabilidade do movimento Browniano.

Proposição 7.0.2. *Seja $f_{t_1,\ldots,t_n}$ a densidade de probabilidade conjunta de $(B(t_1), \ldots, B(t_n))$. Então*

$$f_{t_1,\ldots,t_n}(x_1, \ldots, x_n) = f_{t_1}(x_1) f_{t_2-t_1}(x_2 - x_1) \cdots f_{t_n-t_{n-1}}(x_n - x_{n-1}),$$

onde

$$f_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$$

é a densidade de probabilidade de $B(t)$.

Demonstração. Escrevemos

$$(B(t_1), B(t_2), \ldots, B(t_n)) = (U_1, U_1 + U_2, \ldots, U_1 + U_2 + \cdots + U_n)$$

onde U_i são incrementos, i.e., $U_i = B(t_i) - B(t_{i-1})$ com $i = 1, \ldots, n$ e $t_0 = 0$. Segue da definição do movimento Browniano que U_i tem distribuição normal de média 0 e variância $t_i - t_{i-1}$. Como os incrementos são independentes, a densidade de probabilidade de $(U_1, \ldots, U_n)$ é

$$f_{U_1,\ldots,U_n}(u_1, \ldots, u_n) = f_{t_1}(u_1) f_{t_2-t_1}(u_2) \cdots f_{t_n-t_{n-1}}(u_n),$$

onde $f_t(x)$ é a densidade da normal (média 0 e variância t). Seja $F(x_1, \ldots, x_n)$ a função de distribuição do vector $(B(t_1), B(t_2), \ldots, B(t_n))$. Então

$$\begin{aligned}
F(x_1, \ldots, x_n) &= \mathbb{P}(B(t_1) \leq x_1, B(t_2) \leq x_2, \ldots, B(t_n) \leq x_n) \\
&= \mathbb{P}(U_1 \leq x_1, U_1 + U_2 \leq x_2, \ldots, U_1 + \cdots + U_n \leq x_n) \\
&= \int_R f_{t_1}(u_1) f_{t_2-t_1}(u_2) \cdots f_{t_n-t_{n-1}}(u_n)\, du_1 du_2 \cdots du_n
\end{aligned}$$

onde a região de integração é

$$R = \{(u_1, \dots, u_n) \in \mathbb{R}^n \colon u_1 \le x_1, u_1 + u_2 \le x_2, \dots, u_1 + \dots + u_n \le x_n\}$$

Fazendo a mudança de variável $v_1 = u_1$, $v_2 = u_1 + u_2$, etc, $v_n = u_1 + \dots + u_n$ obtemos

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{t_1}(v_1) f_{t_2 - t_1}(v_2 - v_1) \cdots f_{t_n - t_{n-1}}(v_n - v_{n-1})\, d\boldsymbol{v}$$

onde $d\boldsymbol{v} = dv_1 \cdots dv_n$. □

7.1 Propriedades

O movimento Browniano tem a propriedade de invariância de escala.

Proposição 7.1.1. *Seja $a \neq 0$. O processo estocástico*

$$X(t) = \frac{1}{a} B(a^2 t)$$

é um movimento Browniano.

Demonstração. A continuidade das realização é óbvia e também $X(0) = 0$. A independência dos incrementos também segue directamente da definição. Resta calcular a distribuição do incremento. Seja $t > s$. Então

$$X(t) - X(s) = \frac{1}{a}(B(a^2 t) - B(a^2 s)).$$

Sabemos que $B(a^2 t) - B(a^2 s)$ tem distribuição normal com média 0 e variância $a^2(t - s)$. Portanto, $X(t) - X(s)$ tem distribuição normal com média 0 e variância $t - s$. □

O movimento Browniano também tem a propriedade de inversão do tempo.

Proposição 7.1.2. *O processo estocástico*

$$X(t) = \begin{cases} 0, & t = 0 \\ tB(1/t), & t > 0 \end{cases}$$

é um movimento Browniano.

Demonstração. É claro que $X(0) = 0$ e os incrementos são independentes. A continuidade das realizações é óbvia para $t > 0$. O caso $t = 0$ requer algum cuidado (procurar na bibliografia). Vamos calcular a distribuição do incremento. Dado $t > s > 0$ temos

$$\begin{aligned} X(s) - X(t) &= sB(1/s) - tB(1/t) \\ &= sB(1/s) - tB(1/t) - sB(1/t) + sB(1/t) \\ &= s(B(1/s) - B(1/t)) + (s-t)B(1/t) \end{aligned}$$

Uma vez que os incrementos $B(1/t)$ e $B(1/s) - B(1/t)$ são independentes e normalmente distribuídos, concluímos que $X(t) - X(s)$ também tem distribuição normal com média 0 e variância,

$$\begin{aligned} V(X(t) - X(s)) &= s^2 V(B(1/s) - B(1/t)) + (s-t)^2 V(B(1/t)) \\ &= s^2 \left(\frac{1}{s} - \frac{1}{t}\right) + (s-t)^2 \frac{1}{t} \\ &= t - s \end{aligned}$$

□

Como corolário da proposição anterior obtemos a lei dos grandes números para o movimento Browniano.

Proposição 7.1.3. *Se $B(t)$ é um movimento Browniano, então*

$$\lim_{t \to +\infty} \frac{B(t)}{t} = 0, \quad \mathbb{P} - q.c.$$

Demonstração. Considere o movimento Browniano $X(t)$ da proposição anterior. Então

$$\lim_{t \to +\infty} \frac{X(t)}{t} = \lim_{t \to +\infty} \frac{tB(1/t)}{t} = \lim_{t \to +\infty} B(1/t) = B(0) = 0.$$

□

Este resultado diz que a ordem do crescimento do movimento Browniano $B(t)$ é inferior ao crescimento linear. De facto, existe um resultado mais preciso que indica exactamente a ordem do crescimento. Esse resultado é conhecido pela **lei do logaritmo iterado** e diz que

$$\limsup_{t \to +\infty} \frac{B(t)}{\sqrt{2t \log \log t}} = 1$$

A proposição seguinte demonstra como as trajectórias do movimento Browniano podem ser muito irregulares (apesar de contínuas).

Proposição 7.1.4. *Quase certamente, o movimento Browniano não é monótono em qualquer intervalo* $[a, b]$.

Demonstração. Considere um intervalo $[a, b]$ e uma partição desse intervalo,

$$a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b$$

Se o movimento Browniano é monótono em $[a, b]$, então todos os incrementos $B(a_i) - B(a_{i-1})$ têm o mesmo sinal (positivo ou negativo). Como os incremento são independentes, o evento que consiste em todos os incrementos terem o mesmo sinal tem probabilidade $2 \times \frac{1}{2^n}$, uma vez que há duas escolhas possíveis para o sinal e n incrementos no total. Tomando $n \to \infty$, obtemos que a probabilidade do movimento Browniano ser monótono em $[a, b]$ é igual a zero. O mesmo é verdade para qualquer intervalo com extremos racionais. Assim, a probabilidade do movimento Browniano ser monótono em qualquer intervalo com extremidades racionais é zero (uma vez que a união numerável de eventos com probabilidade zero também tem probabilidade zero). Como qualquer intervalo contém um intervalo com extremidades racionais, temos demonstrada a proposição. □

De facto, apesar do movimento Browniano ter trajectórias contínuas, estas têm um comportamento bastante irregular, ao ponto de não serem sequer diferenciáveis.

Teorema 7.1.5. *Quase certamente, o movimento Browniano não é diferenciável em nenhum ponto. De facto,*

$$\liminf_{h \to 0} \frac{B(t+h) - B(t)}{h} = -\infty \quad \textit{ou} \quad \limsup_{h \to 0} \frac{B(t+h) - B(t)}{h} = +\infty$$

7.2 Exercícios

Exercício 52. Mostre que

$$Cov(W(t), W(s)) = \sigma^2 \min\{t, s\},$$

onde $W(t)$ é o processo de Wiener com parâmetros (x, μ, σ).

Exercício 53. Seja $B(t)$ um movimento Browniano. Mostre que

$$E(e^{B(t)}) = e^{t/2}$$

para todo $t \geq 0$.

Exercício 54. Seja $\{X(t)\colon t \geq 0\}$ um processo estocástico a tempo contínuo que tem as seguintes propriedades:

1. $X(0) = 0$
2. trajectórias contínuas $\mathbb{P}$-q.c.
3. para quaisquer $0 < t_1 < \cdots < t_n$ o vector aleatório $(X(t_1), \ldots, X(t_n))$ tem função de densidade de probabilidade conjunta,

   $$f_{t_1}(x_1) f_{t_2-t_1}(x_2 - x_1) \cdots f_{t_n - t_{n-1}}(x_n - x_{n-1})$$

   onde

   $$f_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}},$$

   é a função de densidade da distribuição normal.

Mostre que $X(t)$ é um movimento Browniano.

Exercício 55. Seja $B(t)$ um movimento Browniano. Mostre que dados $0 \leq s < t$ e A Boreliano de $\mathbb{R}$ então,

$$\mathbb{P}(B(t) \in A | B(s) = x) = \int_A f_{t-s}(y - x) dy,$$

onde $f_t(x) = \frac{1}{\sqrt{2\pi t}} e^{\frac{-x^2}{2t}}$.

Exercício 56. Seja $B(t)$ um movimento Browniano e considere instantes $0 < t_1 < t_2 < t_3$. Determine a distribuição de probabilidade das variáveis:

1. $B(t_1) + B(t_2)$
2. $B(t_1) + B(t_2) + B(t_3)$