PROCESSOS ESTOCÁSTICOS E APLICAÇÕES

JOSÉ PEDRO GAIVÃO

Conteúdo

Date: 01-06-2017.

1. Noções Gerais

Os processos estocásticos são famílias de variáveis aleatórias. São geralmente usados para modelar fenómenos aleatórios que variam ao longo do tempo e onde se podem realizar observações em momentos no tempo espaçados entre si. Como exemplo, considere-se o lançamento de uma moeda um número arbitrário de vezes ou a evolução temporal da cotação de um índice em bolsa.

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

1.1. **Relembrar de teoria de probabilidades.** 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 e

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

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].

Uma **variável aleatória** é uma função mensurável $X : \Omega \to \mathbb{R}$. Ser **mensurável** significa que a pré-imagem (ou imagem inversa) de qualquer conjunto aberto é um evento, ou equivalente, $X^{-1}(]a, b[) \in \mathcal{F}$ 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 dois tipos especiais de variáveis aleatórias: discretas e contínuas. Uma variável aleatória diz-se **discreta** se tomar um conjunto contável (finito ou infinito) de valores, isto é, $X(\Omega) = \{x_1, x_2, \ldots\}$. Uma variável aleatória diz-se **contínua** se $\mathbb{P}(X = x) = F_X(x) - F_X(x^-) = 0$, ou seja,

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

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

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

TABLE 1. Exemplos de variáveis aleatórias discretas

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 2. Exemplos de variáveis aleatórias contínuas

F_X é uma função contínua. 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. Ver Tabelas 1 e 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

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

Finalmente, dado um conjunto de variáveis aleatórias $X_1,\dots X_n$ designa-se por $F_{X_1,\dots,X_n}$ a sua distribuição de probabilidade conjunta,

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

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

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

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

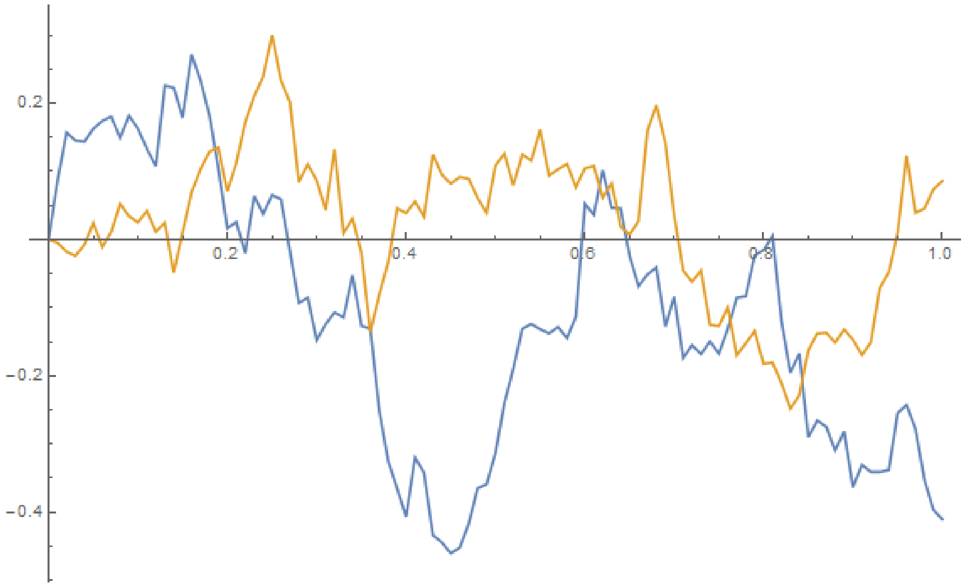

FIGURE 1. Duas realizações de um processo estocástico a tempo contínuo

Definição 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 $T = \{0, 1, 2, \dots\}$ ou $T = \mathbb{Z}$, então o processo estocástico é de **tempo discreto**. Neste contexto, uma sucessão de variáveis aleatórias $X_1, X_2, \dots$ 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 1 e 2.

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

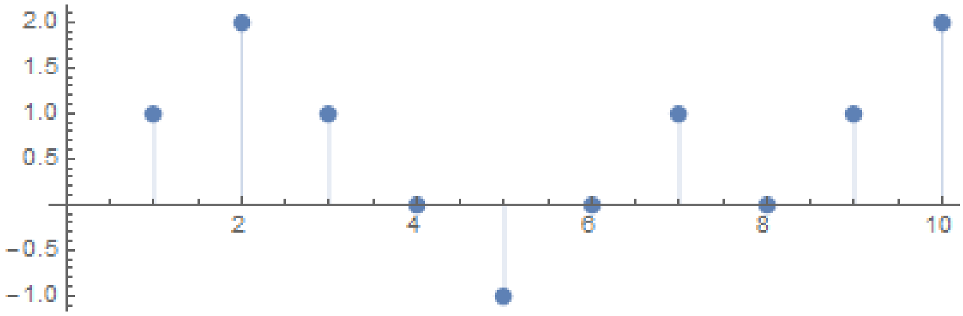

FIGURE 2. Realização de um passeio aleatório

conjuntas,

$$F_{X_{t_1},\dots,X_{t_n}}(x_1,\dots,x_n) = P(X_{t_1} \le x_1,\dots,X_{t_n} \le x_n)\,,$$

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

Definição 1.2. 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,

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

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

Exemplo 1.1 (Passeio aleatório simétrico)**.** Considere a sucessão $(X_n)_{n\ge1}$ 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 $S = \{S_n \colon n \in \mathbb{N}\}$ é designado por **passeio aleatório simétrico**. Ver Figura 2. A variável aleatória S_n por ser interpretada como o deslocamento até ao instante n, podendo ser escrita

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

2. Esperança Condicional

A esperança condicional é um conceito central em teoria da probabilidade, 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 é num certo sentido dada pela esperança condicional.

2.1. **Esperança condicional dado um evento.**

Definição 2.1. Dada uma variável aleatória X e um evento $A \in \mathcal{F}$ tal que $\mathbb{P}(A) > 0$, a **esperança condicional de** X **dado** A é definida por

$$E(X|A) = \frac{1}{P(A)} \int_A X \, d\mathbb{P} \, .$$

Exemplo 2.1. 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 A o evento de uma (e uma só) moeda ficar com a face voltada para cima. Vamos calcular $E(X|A)$. Note-se que $A = \{EF, FE\}$ onde E simboliza escudo e F face. Então

$$X(EF) = 50 \quad \text{e} \quad X(FE) = 20 \, .$$

Logo,

$$E(X|A) = \frac{1}{\frac{1}{2}} \left(50 \times \frac{1}{4} + 20 \times \frac{1}{4} \right) = 35 \, .$$

2.2. **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 $P(Y = y_n) > 0$. Desejamos condicionar uma variável aleatória X dado a variável aleatória Y. Como não sabemos *a priori* qual dos eventos $A_n = \{Y = y_n\}$ pode ocorrer, é necessário considerar todas as esperanças condicionais,

$$E(X|Y = y_1), E(X|Y = y_2), \ldots \, .$$

Definição 2.2. Seja X uma variável aleatória e Y uma variável aleatória discreta. A **esperança condicional de** X **dado** Y é a função $E(X|Y) : \Omega \to \mathbb{R}$ definida por

$$E(X|Y)(\omega) = E(X|A_n) \quad \text{se } \omega \in \{Y = y_n\} \, .$$

Uma vez que $E(X|Y)$ é constante nos eventos $\{Y = y_n\}$, podemos definir de forma análoga a esperança condicional de X dado Y como uma função $E(X|Y = y)$ de $\mathbb{R}$ em $\mathbb{R}$,

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

Exemplo 2.2. Continuando com o exemplo anterior, seja Y a variável aleatória que retorna o montante da primeira moeda, de 20 cêntimos, caso esta se encontre de face voltada para cima. Note-se

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

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 calcula-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,

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

obtemos a seguinte proposição.

Proposição 2.3 (Lei de probabilidade total para a esperança condicional)**.** *Sejam X e Y duas variáveis aleatórias discretas. Então*

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

2.3. **Esperança condicional: caso contínuo.** Suponhamos agora que X e Y são duas variável aleatória 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

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

Definição 2.3. Dadas variáveis aleatórias X e Y contínuas com densidade de probabilidade, a esperança condicional $E(X|Y = y)$ é

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

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

2.4. **Propriedades.**

Proposição 2.6. *Sejam X, X_1, X_2, Y variáveis aleatórias com valor esperado finito e $g : \mathbb{R}^2 \to \mathbb{R}$, $h : \mathbb{R} \to \mathbb{R}$ duas funções integráveis. A esperança condicional satisfaz as seguintes propriedades:*

(1) $E(\alpha_1 X_1 + \alpha_2 X_2 | Y = y) = \alpha_1 E(X_1|Y = y) + \alpha_2 E(X_2|Y = y)$,
(2) $E(X|Y = y) \geq 0$ *se* $X \geq 0$
(3) $E(g(X, Y)|Y = y) = E(g(X, y)|Y = y)$
(4) $E(h(X)|Y = y) = E(h(X))$ *se X e Y forem independentes*

Demonstração. Exercício. □

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.

2.5. **Esperança condicional: caso geral.** 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

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

onde

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

2.6. **Exercícios.**

Exercício 1. 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 $\{X_t : t \geq 0\}$ 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 2. Seja $(\xi_n)_{n\geq 1}$ uma sucessão de variáveis aleatórias de Bernoulli independentes e

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

(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 $\{X_n : n \geq 1\}$ indicando 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 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) $E(N|D = d)$

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 $S = \{0, 1, 2, \dots\}$.

Definição 3.1. 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.

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 3.2. 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. *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 3.2.

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

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

Exemplo 3.3. 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 3.4 (Passeio Aleatório)**.** Considere a sucessão $(\xi_n)_{n \geq 0}$ de variáveis aleatórias **i**ndependentes e **i**denticamente **d**istribuídas (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,

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

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 3.5 (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 3.6 (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 3.7. *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 3.8. Usando a matriz de transição do Exemplo 3.3 podemos calcular, usando a Proposição 3.7, 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 3.9. 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

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

Exemplo 3.10 (Processo de Bernoulli)**.** Considere a cadeia de Markov $(X_n)_{n\geq 0}$ onde $S = \mathbb{N}$, $0 < p < 1$ e

$$\mathbb{P}(X_{n+1} = i + 1|X_n = i) = p, \quad \mathbb{P}(X_{n+1} = i|X_n = i) = 1 - p,$$

para quaisquer $n \geq 0$, $i \in S$. A matriz de transição P associada é

$$P = \begin{pmatrix}1-p & p & 0 & & \cdots\\ 0 & 1-p & p & 0 & \cdots\\ \vdots & & \ddots & \ddots & \end{pmatrix}$$

Se $\mathbb{P}(X_0 = 0) = 1$, então a variável aleatória X_n conta o número de sucessos (com probabilidade p) em n experiência aleatórias independentes.

De facto, é fácil verificar que

$$P^2 = \begin{pmatrix}(1-p)^2 & 2p(1-p) & p^2 & & \cdots\\ 0 & (1-p)^2 & 2p(1-p) & p^2 & \cdots\\ \vdots & & \ddots & \ddots & \end{pmatrix}.$$

Usando indução, podemos mostrar que

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

Exemplo 3.11 (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 3.12 (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,

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

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,

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

3.1. **Exercícios.**

Exercício 4. 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 5. Considere um modelo de aprovisionamento (descrito no Exemplo 3.11) 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 3.3. 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 3.13. *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 3.14 (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,

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

$\square$

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

Teorema 3.15. *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[3].

Usando os mesmos argumentos da prova do Teorema 3.15 obtemos a seguinte condição necessária para um estado ser transiente (ou suficiente para ser recorrente).

Corolário 3.16. *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 $\sum_n P_{i,j}^{(n)} < +\infty$. □

Exemplo 3.17. Considere uma cadeia de Markov com dois estados $\{0, 1\}$ e matriz de transição

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

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.

[3] 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.

Exemplo 3.18. 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 3.19. *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 3.16,

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

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. □

Exemplo 3.20. 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 3.21 (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 3.4. 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 3.22. *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 3.19, quando o espaço de estados é finito, todos os estados recorrentes são positivos.

Proposição 3.23. *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[4] $\lim_{n\to\infty} P_{i,j}^{(n)} = 0$. Mas,

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

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

Exemplo 3.24. Os estados recorrentes dos exemplos 3.17 e 3.18 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 3.25 (Simétrico)**.** Considere o passeio aleatório simétrico. As probabilidades de transição são

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

[4]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

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[5],

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

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.

3.3. **Exercícios.**

Exercício 6. 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

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

[5] ver wikipedia

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

Exercício 7. 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 3.5. 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 3.6. Um estado $i \in S$ que é recorrente positivo e aperiódico diz-se que é **ergódico**.

Exemplo 3.26. Considere a cadeia de Markov do Exemplo 3.17. 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 3.27. Considere o passeio aleatório simétrico do Exemplo 3.25. 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 3.28. Considere a cadeia de Markov do Exercício 6. 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 3.29. 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 3.7. 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

$$i \to j$$

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

$$i \leftrightarrow j$$

Lemma 3.30. *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 $P_{i,j}^{(n)} > 0$ $(i \to j)$ e $P_{j,k}^{(m)} > 0$ $(j \to k)$ 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, $i \to k$. De forma análoga mostramos que $k \to i$.

□

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

Definição 3.8. 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 3.31. *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 3.9. 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 3.32. *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 3.33 (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,

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

□

Observação 3.34. 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 3.23, todos os estados recorrentes são positivos.

Observação 3.35. Se uma classe de comunicação $C(i)$ não é fechada, então segue da Proposição 3.32 que todos os estados da classe $C(i)$ são transientes.

Exemplo 3.36. 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 3.23.

3.6. **Exercícios.**

Exercício 8. 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 6 e escolha p e q adequadamente)

Exercício 9. Mostre que se $P_{i,i} > 0$, então o estado i é aperiódico.

Exercício 10. Mostre que se $i \leftrightarrow j$, então $Per(i) = Per(j)$.

Exercício 11. 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 12. 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 13. 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 3.7 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 3.10. Uma distribuição de probabilidade[6] π em S diz-se **estacionária** se

$$\pi = \pi P.$$

Observação 3.37. 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[7] 1, então a cadeia não tem distribuição estacionária.

Observação 3.38. 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

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

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,

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

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

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

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

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

Exemplo 3.39 (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 3.40 (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 3.41 (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 3.20.

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

Proposição 3.42. *Seja $\pi = (\pi_j \colon j \in S)$ 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 3.16). Então

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

□

Exemplo 3.43. Considere a cadeia de Markov do Exemplo 3.18. 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 3.44. *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 $\pi = (\pi_j \colon j \in S)$ 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. □

Observação 3.45. Se o espaço de estados é finito, então sabemos das Proposições 3.19 e 3.23 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 3.46. Considere a cadeia de Markov do Exemplo 3.17. 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 3.47. 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 3.48. *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[8]. 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**.

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

Observação 3.49. 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 3.48. Estabelece uma condição suficiente para a convergência das probabilidades de transição, definindo no limite uma única distribuição estacionária.

Teorema 3.50 (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 3.51. *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 $\pi = (\pi_j : j \in S)$ dada por*

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

Demonstração. Segue da Proposição 3.48 e do Teorema 3.50. □

Observação 3.52. A razão $\pi_j = 1/m_j$ representa a fracção de tempo passado no estado j.

Observação 3.53. 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 3.54. Considere a cadeia do Exercício 6. É irredutível e aperiódica. Tem uma única distribuição estacionária dada por,

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

Exemplo 3.55. Considere a cadeia de Markov do Exemplo 3.36. 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 3.56 (Aplicação: Modelo de Aprovisionamento)**.** Considere o modelo de aprovisionamento descrito no Exemplo 3.11, 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 3.51, 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 3.57 (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 3.51, 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 14. 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 15. 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 16. 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 3.11. 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 3.58.

(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 3.59. 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[9]. 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[10]

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

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

Proposição 3.60. *O vector de probabilidades de absorção em A, $h = (h_i \colon i \in S)$, é solução do seguinte sistema de equações lineares,*

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

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 3.61. 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

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

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

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

Compare com a equação que define uma distribuição estacionária.

Exemplo 3.62. Considere a cadeia de Markov do Exemplo 3.59. 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 3.63. *Se o vector de tempos médios de absorção em A, $t = (t_i \colon i \in S)$ for finito*[11]*, então t é solução do seguinte sistema de equações lineares,*

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

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

Exemplo 3.64. 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 17. 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.

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

Exercício 18. 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 19. 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.

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 4.1. 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 4.1. *Sejam $T_1, \ldots, T_n$, $n \geq 1$ 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[12] $W_{n+1} = W_n + T_{n+1}$ 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,

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

□

[12]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,

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

4.1. Definição do processo de Poisson.

Definição 4.2. Uma variável aleatória X tem **distribuição de Poisson** com parâmetro $\lambda > 0$ se

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

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 4.2. *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 considerando 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 4.3 (Processo de Poisson)**.** $\{N(t), t \geq 0\}$ é um **processo de Poisson** com parâmetro $\lambda > 0$ se

(1) $N(0) = 0$,

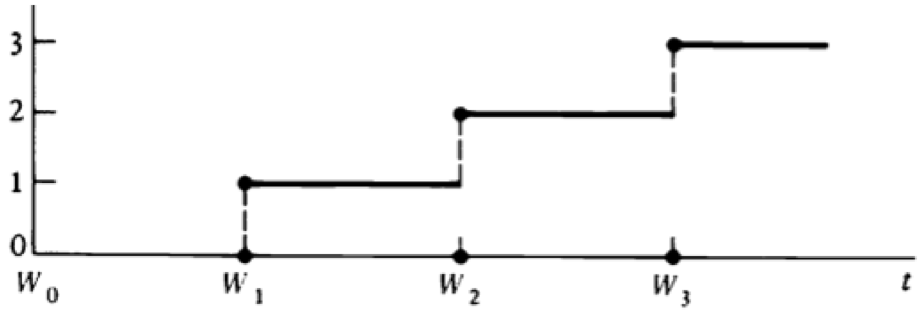

FIGURE 3. Realização de $N(t)$

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

(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 4.4. 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 4.3. *$N(t)$ 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 4.1 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 4.4. *$N(t+s)-N(s)$ 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 3). 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 4.5. *$N(t)$ 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 4.6. *$N(t)$ 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 4.7. *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 4.8. *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. 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 4.9. *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. □

Exemplo 4.10. 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 20. 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 21. 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 22. 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 23. 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 $\lambda_1 + \lambda_2$.

5. Processos de Markov em tempo contínuo

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

Definição 5.1. Dizemos que $\{X(t) \colon t \geq 0\}$ é um **processo (ou cadeia) de Markov a tempo contínuo** sse

$$\begin{aligned}\mathbb{P}(X(t+s) = j | X(s) = i, X(s_n) = i_n \ldots, &X(s_0) = i_0) \\ &= \mathbb{P}(X(t+s) = j | X(s) = i)\end{aligned}$$

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

Exemplo 5.1. 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

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

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 $c(s,t) = \operatorname{Cov}(X(s), X(t))$. Então[13] existem $\mu \in \mathbb{R}$ e $\sigma > 0$ tal que

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

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

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

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

Proposição 5.2 (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**. A matriz formada pelas intensidades dos saltos é conhecida por **matriz de intensidades** do processo,

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

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.

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

Exemplo 5.3. O processo de Poisson tem incrementos estacionários e independentes, logo é um processo de Markov. Por outro lado, das equações que caracterizam o processo

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

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 5.4. 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 5.5. 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

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

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

Exemplo 5.6. 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.

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

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 $0 < \lambda_i < \infty$. Quando $\lambda_i = 0$ então o processo é constante no estado i. Define-se

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

que corresponde à "probabilidade" do processo entrar em j uma vez que saiu de i. 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

$$T_n = \frac{\tau_{n-1}}{\lambda_{Y_0}} \quad \text{e} \quad W_n = T_1 + \cdots + T_n, \quad n = 1, 2, \ldots$$

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[14], 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

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

$[0,t] \cup [t,t+h]$ 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 5.7 (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

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

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 5.2. 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 5.8. *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. 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 5.9. *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 5.10. 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.

5.4. **Exercícios.**

Exercício 24. 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 25. 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 26. Considere um processo de nascimento e morte com parâmetros $\lambda_n = \lambda$ e $\mu_n = \mu n$ para $n = 0, 1, 2, \ldots$. 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 27. 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 28. Seja $X(t)$ um processo de nascimento e morte com estados $\{0, 1, \ldots, N\}$ e parâmetros $\lambda_n = \alpha(N - n)$ e $\mu_n = \beta n$ onde $\alpha, \beta > 0$. Determine a distribuição estacionária.