

Uma Introdução aos Processos estocásticos

Adrian Hinojosa e Aniura Milanés

Departamento de Estatística
ICE_x. UFMG.

Conteúdo

1. Prefácio.	5
Capítulo 1. Introdução.	7
1. Definições e exemplos.	7
2. O Processo de Bernoulli e outros processos estocásticos associados	13
Capítulo 2. Cadeias de Markov a Tempo Discreto	21
1. Definições e exemplos	21
2. Matrizes de transição de ordem superior.	29
3. Comportamento assintótico de cadeias com dois estados.	35
4. Classificação dos estados. Parte I.	38
5. Tempos de parada	43
6. Visitas a um Estado	46
7. Classificação dos Estados. Parte II	51
8. Probabilidades de Absorção	55
9. Exemplos	57
Capítulo 3. O Processo de Poisson	73
1. Introdução.	73
2. O processo de Poisson.	74
3. Tempos de Chegada	81
4. Superposição de Processos de Poisson	85
5. Decomposição de Processos de Poisson	86
6. Processo de Poisson Composto.	87
7. Processo de Poisson não homogêneo.	89
Capítulo 4. Cadeias de Markov a Tempo Contínuo	93
1. Definição e exemplos.	93
2. Estrutura de uma cadeia de Markov a tempo contínuo.	97
3. O gerador infinitesimal.	99
4. As equações diferenciais de Kolmogorov.	100
5. Distribuição estacionária e comportamento assintótico.	103
6. Aplicações.	106

1. Prefácio.

O presente texto foi elaborado a partir de notas de aula utilizadas no curso de graduação de Processos Estocásticos do departamento de Estatística do ICEX/UFMG. Ele é baseado na literatura colocada nas referências. O nosso objetivo não era apresentar resultados novos dos processos estocásticos e sim redigir um texto básico em português com os resultados fundamentais desta teoria.

O material foi organizado da forma seguinte. No primeiro capítulo falamos sobre a teoria geral dos processos estocásticos e colocamos como exemplos ilustrativos o processo de Bernoulli e outros processos associados a ele de forma natural. No segundo capítulo tratamos a teoria básica das cadeias de Markov a tempo discreto. Damos aqui as definições e resultados mais importantes. Em ordem de dificuldade seguiriam as cadeias de Markov a tempo contínuo, mas elas são tratadas somente no capítulo 4. Antes, no capítulo 3 tratamos sobre o processo de Poisson, que é um processo importante por si mesmo além de ser um exemplo de cadeia de Markov a tempo contínuo. Apresentamos as suas principais propriedades e generalizações.

O texto contém um conjunto de exemplos e figuras que visam fazer mais amena e ilustrativa a leitura. Ele deve ser aprimorado mais para frente com a adição de mais exemplos e com a confecção de exercícios. Uma outra ausência que será preenchida no futuro é a apresentação de aplicações das cadeias de Markov a métodos de simulação e estimação não paramétrica

Por último lembramos que esta é uma primeira versão que possivelmente terá muitos erros. Ficaremos muito gratos de qualquer comunicação que possa nos ajudar a corrigí-los.

Os autores. Belo Horizonte, 2 de dezembro de 2005.

CAPÍTULO 1

Introdução.

1. Definições e exemplos.

Consideremos os seguintes exemplos.

EXEMPLO 1.1. *Cadeia de montagem.* Os produtos finais de uma cadeia de montagem, após uma supervisão a que são submetidos, são considerados defeituosos, $X_n = 1$, ou sem defeito, $X_n = 0$. Suponha que um produto é defeituoso independentemente dos outros produtos e que a probabilidade de que isto aconteça é p , então $\{X_n\}_{n \geq 1}$ é uma sequência de variáveis aleatórias Bernoulli independentes.

EXEMPLO 1.2. *Estoque.* Uma pequena loja de equipos electrodomésticos vende certo tipo de máquina de lavar roupa. No entanto, ela somente pode ter em estoque no máximo cinco unidades. Então se no final do dia a loja tem no estoque somente uma unidade ou nenhuma, o gerente manda buscar tantas unidades quantas sejam necessárias para ter cinco na loja no dia seguinte antes de começar o expediente. Vamos chamar de X_n a quantidade de unidades na loja no final do n -ésimo dia. Elas podem ser consideradas variáveis aleatórias, pois é razoável supor que não temos como prever a quantidade de máquinas de lavar que serão compradas cada dia.

EXEMPLO 1.3. *Mobilidade social.* Para cada $n \in \mathbb{N}$ denotaremos como X_n a classe social de certa família na geração n , assumindo os valores 1=alta, 2=média e 3=baixa.

Nas três situações acima, as magnitudes de interesse podem ser modeladas usando processos estocásticos.

DEFINIÇÃO 1.4. *Seja \mathbb{T} um conjunto de índices e $E \subset \mathbb{R}$. Um processo estocástico indexado por \mathbb{T} com espaço de estados E é uma família de variáveis aleatórias $X = \{X_t : t \in \mathbb{T}\}$ definidas num espaço amostral Ω e tomando valores no conjunto E .*

O conjunto de índices costuma ser interpretado como tempo. Assim $t \in \mathbb{T}$ seria o instante de tempo no qual é observada a variável X_t , que é chamada de estado do processo nesse instante.

- Se \mathbb{T} for um conjunto enumerável (por exemplo, $\mathbb{T} = \{0, 1, 2, \dots\}$) então dizemos que o processo estocástico é **a tempo discreto**.
- Em caso contrário diremos que o processo estocástico é **a tempo contínuo**.

Nos exemplos 1.1-1.3 encontramos processos estocásticos a tempo discreto.

- Se $E \subseteq \mathbb{Z}$ então o processo tem espaço de estados discreto.
- Se $E \subseteq \mathbb{R}$ então o processo tem espaço de estados contínuo, ou é um processo estocástico real.

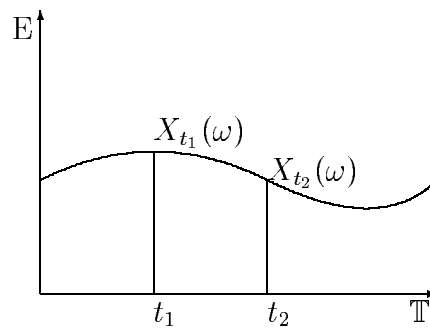
Os espaços de estados nos exemplos 1.1-1.3 são todos discretos. De fato temos para o exemplo 1.1 que $E = \{0, 1\}$, no exemplo 1.2, $E = \{0, 1, 2, 3, 4, 5\}$ e no exemplo 1.3, $E = \{1, 2, 3\}$.

Podemos pensar um processo estocástico X como uma função:

$$\begin{aligned} X : \mathbb{T} \times \Omega &\longmapsto E \\ (t, \omega) &\longmapsto X(t, \omega) \end{aligned}$$

Fixando um evento $\omega \in \Omega$, obtemos uma coleção de valores $\{X_t(\omega) : t \in \mathbb{T}\}$ que é chamada de **trajetória** ou **realização** de este processo. Podemos ter vários tipos de trajetórias que são ilustradas graficamente a seguir.

Trajетória contínua:



Trajетória Discreta:

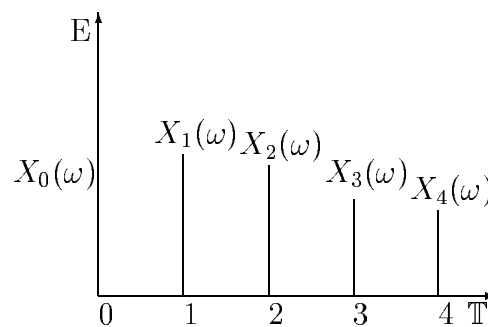


Figura 1

Vemos então que os processos estocásticos são generalizações de variáveis e vetores aleatórios. De fato, se a cardinalidade do conjunto \mathbb{T} for finita, o processo correspondente será um vetor aleatório. Em particular, se \mathbb{T} contiver somente um elemento, o processo será uma variável aleatória.

O comportamento probabilístico de variáveis aleatórias é descrito através da função de distribuição. No caso de vetores aleatórios precisamos usar a função de distribuição conjunta e sabemos que as marginais das coordenadas não determinam a distribuição do vetor no sentido que existem vetores aleatórios com distribuições diferentes e com as mesmas marginais. No caso de processos estocásticos acontece algo semelhante.

EXERCÍCIO 1.5. *O processo descrito no exemplo 1.1 é chamado de processo de Bernoulli de parâmetro p . Seja X um processo de Bernoulli de parâmetro \sqrt{p} e defina o processo Y como segue*

$$\begin{aligned} Y_{3n-2} &= X_{3n-2}X_{3n-1}, \\ Y_{3n-1} &= X_{3n-1}X_{3n}, \\ Y_{3n} &= X_{3n-2}X_{3n}, \end{aligned}$$

para cada $n \in \mathbb{N}$.

Prove que cada Y_n tem distribuição de Bernoulli com parâmetro p , mas que os estados do processo Y não são independentes e portanto ele não é um processo de Bernoulli.

Para caracterizar o comportamento probabilístico de um processo estocástico, devemos considerar a família das funções de distribuição de todos os vetores aleatórios formados com estados do processo. Mais precisamente, teremos a seguinte definição.

DEFINIÇÃO 1.6. *Seja $X = \{X_t, t \in \mathbb{T}\}$ um processo estocástico. Consideremos para cada conjunto $t_1 < t_2 < \dots < t_n$, $t_j \in \mathbb{T}$, $n \in \mathbb{N}$, a função de distribuição conjunta do vetor aleatório $(X_{t_1}, \dots, X_{t_n})$ que denotaremos por F_{t_1, t_2, \dots, t_n} .*

*A família $\{F_{t_1, t_2, \dots, t_n}\}$ das **funções de distribuição finito-dimensionais** de X é chamada de **lei do processo**.*

É claro que estas funções de distribuição conjunta estão definidas de maneira única e satisfazem a seguinte propriedade de consistência,

$$\lim_{x_k \uparrow \infty} F_{t_1, t_2, \dots, t_n}(x_1, \dots, x_n) = F_{t_1, \dots, t_{k-1}, t_{k+1}, \dots, t_n}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n). \quad (1.1)$$

A todo processo estocástico corresponde uma família de funções de distribuição satisfazendo (1.1).

NOTA 1.7. *Pode-se provar que dada uma família consistente de funções de distribuição, podem se encontrar um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ e um processo estocástico X tais que esta família constitua a lei deste processo. Este resultado é fundamental na hora de provar a existência de processos estocásticos.*

Supondo que as seguintes expressões existem, as esperanças e variâncias

$$\mu(t) = \mathbb{E}(X(t)), \quad \sigma^2(t) = \text{Var}(X(t)),$$

respectivamente, a cada instante $t \in \mathbb{T}$ e as covariâncias

$$C(s, t) = \mathbb{E}((X(s) - \mu(s))(X(t) - \mu(t)))$$

em distintos momentos $s, t \in \mathbb{T}$, dão alguma informação sobre a variabilidade no tempo do processo correspondente.

Os **processos estacionários (no sentido estrito)**, são aqueles cujas distribuições finito-dimensionais são invariantes no tempo, i.e.,

$$F_{t_1+h, t_2+h, \dots, t_n+h} = F_{t_1, t_2, \dots, t_n}$$

para todo $t_i, t_{i+h} \in \mathbb{T}$, $i = 1, 2, \dots, n$, $n \in \mathbb{N}$ e $h > 0$. Por outro lado, se existir uma constante μ e uma função $c : \mathbb{R} \rightarrow \mathbb{R}$ tais que

$$\mu(t) = \mu, \quad \sigma^2(t) = c(0) \text{ e } C(s, t) = c(t - s),$$

para todos $s, t \in \mathbb{T}$, então diremos que o processo é **estacionário no sentido amplo**.

O processo de Bernoulli (e em geral qualquer processo estocástico com estados identicamente distribuídos) é um exemplo de processo estacionário no sentido estrito. É claro que todo processo estacionário no sentido estrito tem que sê-lo também no sentido amplo. O contrário não necessariamente tem que ocorrer.

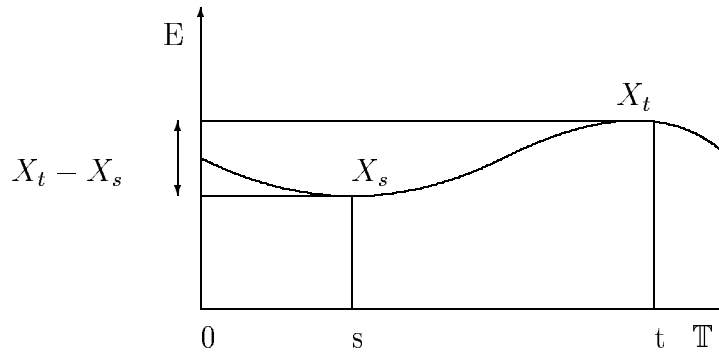


Figura 2

A seguinte definição será útil no que segue.

DEFINIÇÃO 1.8. $X_t - X_s$, $t > s$, $s, t \in \mathbb{T}$ é dito **incremento** correspondente ao intervalo $(s, t]$. Diremos que o **processo X tem incrementos estacionários** quando a distribuição de $X_t - X_s$ depende dos instantes s e t somente a través da sua diferença $t - s$. Se para todos $t_0 < t_1 < \dots < t_n$, com $t_1, \dots, t_n \in \mathbb{T}$ e para todo $n \in \mathbb{N}$ vale:

$$X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}} \quad \text{são independentes,}$$

então diremos que X é um processo de incrementos independentes.

Vejamos agora dois exemplos clássicos de processos estocásticos a tempo contínuo.

EXEMPLO 1.9. O processo de Poisson

Dada uma constante positiva λ , considere o processo a tempo contínuo $N = \{N_t\}_{t \geq 0}$, com espaço de estados $E = \mathbb{N}$ e que tem as seguintes características:

- a) $N_0 = 0$;
- b) N tem incrementos independentes;
- c) para todos $0 \leq s < t$, $N_t - N_s \sim \text{Poisson}[\lambda(t - s)]$.

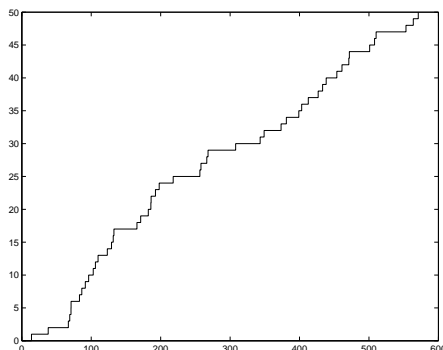


Figura 3. Trajetória do processo de Poisson

Este é chamado processo de Poisson de taxa λ . Observe que os seus incrementos além de serem independentes são também estacionários. As trajetórias deste processo são como na Figura 1.9. Elas têm saltos de tamanho um.

NOTA 1.10. Este processo é usado para construir um tipo de processo chamado cadeias de Markov a tempo contínuo.

EXEMPLO 1.11. O movimento Browniano

Em 1827, o botânico escocês Robert Brown observou e descreveu o movimento irregular executado por pequenos grãos de pólen suspensos em água. Esta observação aparentemente sem muita importância, tornou-se especialmente relevante alguns anos depois. Embora L. Bachelier em 1900 e A. Einstein em 1905 tenham sido os primeiros a abordar quantitativamente o estudo deste fenômeno, foi o matemático norteamericano Norbert Wiener quem em 1923 estudou e formalizou rigorosamente o modelo matemático motivado no fenômeno físico do movimento browniano. É por isso que ele é chamado de processo de Wiener ou movimento browniano, sendo que este último nome dá mais ênfase ao processo físico.

Considere o processo a tempo contínuo $X = \{X_t\}_{t \geq 0}$, com espaço de estados $E = \mathbb{R}$, que tem as seguintes características:

- a) $X_0 = 0$;
- b) X tem incrementos independentes;
- c)

$$P(X_t - X_s \leq x) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^x e^{-\frac{u^2}{2(t-s)}} du,$$

i.e. $X_t - X_s \sim N(0, t-s)$;

- d) X possui trajetórias contínuas.

X é conhecido como movimento Browniano ou processo de Wiener.

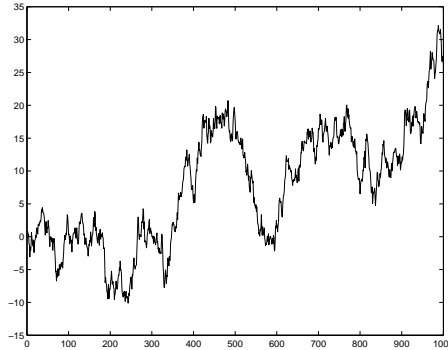


Figura 4. Trajetória do movimento Browniano

A Figura 1.11 sugere um comportamento bastante irregular das trajetórias do processo de Wiener.

EXERCÍCIO 1.12. Para o processo de Wiener, calcule

- (1) a função de densidade conjunta de dois estados.
- (2) a função de densidade conjunta de n estados.

Sugestão: Para instantes $0 < s < t$ você pode calcular a densidade conjunta de $Y = X_s - X_0 = X_s$ e $Z = X_t - X_s$. A densidade de X_s e X_t pode ser obtida usando a mudança de variáveis. O caso geral é análogo.

Entre os estados de um processo X podemos encontrar diferentes tipos de relações de dependência.

- a) *Estados Independentes*

Um **processo de estados independentes** é simplesmente aquele tal que todos os seus estados constituem uma família de variáveis aleatórias independentes. Um exemplo é o processo de Bernoulli de parâmetro p .

b) *Processos de Markov*

Um processo X é chamado de **processo de Markov** quando para todos $a, b, a_1, \dots, a_n \in E$, $t_1 < t_2 < \dots < t_n < t$ com $t_1, t_2, \dots, t_n, t \in \mathbb{T}$ vale

$$P[a \leq X_t \leq b | X_{t_1} = a_1, X_{t_2} = a_2, \dots, X_{t_n} = a_n] = P[a \leq X_t \leq b | X_{t_n} = a_n]$$

i.e. o processo só depende do *presente* t_n e não do *passado* $t_1 < t_2 < \dots < t_{n-1}$.

NOTA 1.13.

$$P(x, s; t, A) := P(X_t \in A | X_s = x)$$

é chamada *probabilidade de transição do processo* $\{X_t\}_{t \geq 0}$.

d) *Martingais*

X será chamado de **martingal** quando para todos $t_1 < t_2 < \dots < t_n < t_{n+1}$ com $t_1, t_2, \dots, t_n, t \in \mathbb{T}$ e $a_1, a_2, \dots, a_n \in E$ temos

$$\mathbb{E}[X_{t_{n+1}} | X_{t_1} = a_1, X_{t_2} = a_2, \dots, X_{t_n} = a_n] = X_{t_n} = a_n.$$

Em outras palavras, poderíamos dizer que para martingais vale que o que pode ser previsto sobre o estado do processo num instante futuro (t_{n+1}) sendo que são conhecidos n estados anteriores é exatamente o estado no instante presente (t_n).

EXEMPLO 1.14. *Um exemplo de martingal aparece em jogos simples de azar como o seguinte. Suponhamos que no n -ésimo lançamento de uma moeda honesta acrescentamos um valor A ao capital do jogador se sair cara subtraímos a mesma quantidade se sair coroa. O jogador começa o jogo com capital K e é admitido ter capital negativo. Vamor supor também que os lançamentos são independentes.*

Fazendo

$$Z_j = \begin{cases} A, & \text{se sair cara no } j\text{-ésimo lançamento} \\ -A & \text{se sair coroa no } j\text{-ésimo lançamento,} \end{cases} \quad (1.2)$$

teremos que o capital do jogador no instante do n -ésimo lançamento será

$$X_n = K + Z_1 + Z_2 + \dots + Z_n.$$

Observando que Z_1, Z_2, \dots são variáveis aleatórias independentes com $\mathbb{E}Z_i = 0$ vemos que é fácil verificar que X_n é um martingal, pois:

$$\begin{aligned} & \mathbb{E}[X_{n+1} | X_1 = a_1, X_2 = a_2, \dots, X_n = a_n] \\ &= \mathbb{E}[(X_n + Z_{n+1}) | X_1 = a_1, X_2 = a_2, \dots, X_n = a_n] \\ &= \mathbb{E}[(a_n + Z_{n+1}) | X_1 = a_1, X_2 = a_2, \dots, X_n = a_n] \\ &= a_n + \mathbb{E}[Z_{n+1} | X_1 = a_1, X_2 = a_2, \dots, X_n = a_n] \\ &= a_n + \mathbb{E}[Z_{n+1}] = a_n = X_n \end{aligned}$$

2. O Processo de Bernoulli e outros processos estocásticos associados

No exemplo 1.1 foi introduzido o processo de Bernoulli que trataremos agora com mais detalhe.

2.1. O Processo de Bernoulli. O processo a tempo discreto $\{X_n : n = 1, 2, \dots\}$ é chamado processo de Bernoulli, com probabilidade de sucesso p se:

- (1) X_1, X_2, \dots são v.a. independentes
- (2) $\forall n \geq 1 \ P(X_n = 1) = p, \ P(X_n = 0) = 1 - p = q.$

O processo apresentado no exemplo 1.1 é um processo de Bernoulli.

PROPRIEDADES 2.1. *Seja $\{X_n\}_{n \geq 1}$ um processo de Bernoulli, então:*

- (1) $\mathbb{E}X_n = p$
- (2) $\text{Var}X_n = pq$
- (3) $\mathbb{E}\alpha^{X_n} = q + \alpha p$

EXERCÍCIO 2.2. *Prove as propriedades anteriores.*

2.2. O Número de Sucessos no Processo de Bernoulli. Considere o processo de Bernoulli $\{X_n\}_{n \geq 1}$, com probabilidade p de sucesso. Defina:

$$N_n = \begin{cases} 0, & n = 0 \\ X_1 + \dots + X_n, & n = 1, 2, \dots \end{cases}$$

N_n é o número de sucessos nos n primeiros 'ensaios' do processo de Bernoulli. $\{N_n\}_{n \geq 0}$ é um processo estocástico.

Observe que $N_n \sim b(n, p)$ ($b(n, p)$ é a distribuição Binomial de parâmetros n e p), pois $N_n = X_1 + \dots + X_n$. Além disso, as 'probabilidades de transição' do processo N_n são:

$$P(N_{n+1} = k | N_n = j) = P(N_{n+1} = k - j) = \begin{cases} p, & \text{se } k - j = 1 \\ q, & \text{se } k = j \\ 0, & \text{caso contrário.} \end{cases}$$

Como $N_{n+m} - N_n$ é o número de sucessos nos *ensaios* independentes $n + 1, n + 2, \dots, n + m$, então $N_{n+m} - N_n \sim b(m, p)$ e $P(N_{n+m} - N_n = k) = \binom{m}{k} p^k (1 - p)^{m-k}$.

PROPRIEDADES 2.3.

- (1) *Para todo n :*
 $P(N_{n+m} - N_n = k | N_0, \dots, N_m) = P(N_{n+m} - N_n = k) = \binom{m}{k} p^k (1 - p)^{m-k}.$
- (2) *Para todos $n_0 \leq n_1 \leq n_2 \leq \dots \leq n_{m-1} \leq n_m$ vale que os incrementos:*
 $N_{n_1} - N_{n_0}, N_{n_2} - N_{n_1}, \dots, N_{n_m} - N_{n_{m-1}}$ *são independentes.*

EXEMPLO 2.4. *Calcule $P(N_5 = 4, N_7 = 5, N_{13} = 8)$.*

2. O PROCESSO DE BERNOULLI E OUTROS PROCESSOS ESTOCÁSTICOS ASSOCIADOS

$$\begin{aligned}
 P(N_5 = 4, N_7 = 5, N_{13} = 8) &= P(N_5 = 4, N_7 - N_5 = 1, N_{13} - N_7 = 3) \\
 &= P(N_5 = 4)P(N_7 - N_5 = 1)P(N_{13} - N_7 = 3) \\
 &= \binom{5}{4}p^4q \binom{2}{1}pq \binom{6}{3}p^3q^3
 \end{aligned}$$

Obs. $0 = N_0 \leq N_1 \leq N_2 \leq N_3 \leq \dots$

EXEMPLO 2.5. Calcule $\mathbb{E}(N_5N_8)$.
 Observe que $N_8 = N_5 + (N_8 - N_5)$

$$\begin{aligned}
 \mathbb{E}(N_5N_8) &= \mathbb{E}[N_5(N_5 + (N_8 - N_5))] \\
 &= \mathbb{E}N_5^2 + \mathbb{E}[N_5(N_8 - N_5)] \\
 &= \mathbb{E}N_5^2 + \mathbb{E}N_5\mathbb{E}[N_8 - N_5] \\
 &= (25p^2 + 5pq) + 5p3p
 \end{aligned}$$

Como vimos acima, o processo $\{N_n\}_{n \geq 0}$ tem incrementos independentes e estacionários. podemos provar também que:

TEOREMA 2.6.

- (1) $\mathbb{E}[N_{n+1}|N_0, \dots, N_n] = \mathbb{E}[N_{n+1}|N_n]$.
- (2) $P(N_{n+1} = k|N_0 = i_1, \dots, N_n = i_n) = (N_{n+1} = k|N_n = i_n)$,
i.e. o processo é Markoviano.

DEMONSTRAÇÃO. Exercício □

EXEMPLO 2.7.

$$\begin{aligned}
 \mathbb{E}(N_{11}|N_5) &= \mathbb{E}[N_5 + (N_{11} - N_5)|N_5] \\
 &= \mathbb{E}(N_5|N_5) + \mathbb{E}[N_{11} - N_5|N_5] \\
 &= N_5 + \mathbb{E}[N_{11} - N_5] \\
 &= N_5 + 6p
 \end{aligned}$$

EXEMPLO 2.8.

$$\begin{aligned}
 \mathbb{E}(N_5N_8) &= \mathbb{E}(\mathbb{E}[N_5N_8|N_5]) \\
 &= \mathbb{E}(N_5\mathbb{E}[N_8|N_5]) \\
 &= \mathbb{E}(N_5^2 + 3pN_5) \\
 &= 25p^2 + 5pq + 3p\mathbb{E}N_5 = 25p^2 + 5pq + 3p(5p)
 \end{aligned}$$

EXEMPLO 2.9.

$$\begin{aligned}
\mathbb{E}[N_{11}N_5|N_2, N_3] &= \mathbb{E}[\mathbb{E}(N_{11}N_5|N_0, N_1, \dots, N_5)|N_2, N_3] \\
&= \mathbb{E}[\mathbb{E}(N_{11}N_5|N_5)|N_2, N_3] = \mathbb{E}[N_5\mathbb{E}(N_{11}|N_5)|N_2, N_3] \\
&= \mathbb{E}[N_5(N_5 + 6p)|N_2, N_3] = \mathbb{E}[N_5^2 + 6pN_5|N_2, N_3] \\
&= \mathbb{E}[N_5^2 + (N_5)6p|N_3] = \mathbb{E}[(N_3 + (N_5 - N_3))^2 + 6pN_5|N_3] \\
&= \mathbb{E}[N_3^2 + 2N_3(N_5 - N_3) + (N_5 - N_3)^2 + 6pN_5|N_3] \\
&= N_3^2 + 2N_3\mathbb{E}[N_5 - N_3] + \mathbb{E}[(N_5 - N_3)^2] + 6p\mathbb{E}[N_5|N_3] \\
&= N_3^2 + 2N_3\mathbb{E}[N_5 - N_3|N_3] + \mathbb{E}[(N_5 - N_3)^2|N_3] \\
&+ 6p(\mathbb{E}[N_5 - N_3|N_3] + \mathbb{E}[N_3|N_3]) \\
&= N_3^2 + 2N_3(2p) + (4p^2 + 2pq) + 6p(N_3 - 2p) \\
&= N_3^2 + 10pN_3 + 16p^2 + 2pq
\end{aligned}$$

2.3. O Tempo dos Sucessos. Considere o processo de Bernoulli $\{X_n\}_{n \geq 0}$, com probabilidade p de sucesso. Sejam T_1, T_2, \dots os tempos nos quais acontecem os sucessos (i.e. $X_n = 1$).

EXEMPLO 2.10. Considere a realização $X_1 = 0, X_2 = 1, X_3 = 0, X_4 = 1, X_5 = 1$ do processo de Bernoulli, a trajetória do processo ate o instante $n = 5$ seria:

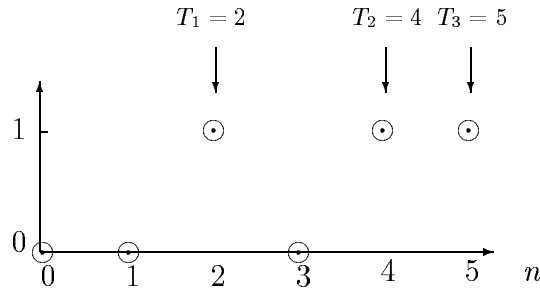


Figura 4.

logo $T_1 = 2, T_2 = 4$ e $T_3 = 5$.

A seqüência de v.a. $\{T_n\}_{n \geq 1}$ é chamada, processo dos tempos dos sucessos ou processo das chegadas.

EXEMPLO 2.11. Cada segundo do tempo observamos se nesse instante passa um carro num sitio fixo do caminho, suponha que os carros passam independentemente uns de outros. Este experimento corresponde a um processo de Bernoulli $\{X_n\}_{n \geq 1}$, com $X_n = 1$ se

2. O PROCESSO DE BERNOULLI E OUTROS PROCESSOS ESTOCÁSTICOS ASSOCIADOS

observamos um carro no instante n . Logo N_n seria o número de carros que passaram até o instante n e T_n o instante em que passou o n -ésimo carro.

Observe que para todo k $X_{T_k} = 1$, vejamos isso com um desenho:

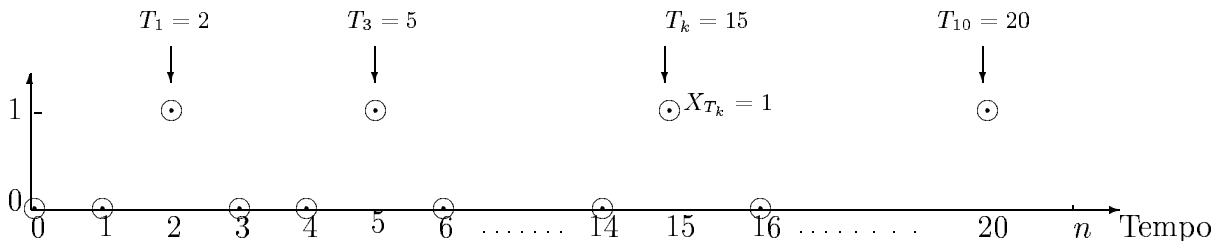


Figura 5

A partir da figura podemos concluir que:

LEMA 2.12.

- (1) $T_k \leq n$ se e somente se $N_n \geq k$.
- (2) $T_k = n$ se e somente se $N_{n-1} = k - 1, X_n = 1$.

Usando este lema vamos a provar este teorema:

TEOREMA 2.13. Para todo $k \in \mathbb{N}$:

- 1) $P(T_k \leq n) = \sum_{j=k}^n \binom{n}{j} p^j (1-p)^{n-j}$.
- 2) $P(T_k = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}$.

DEMONSTRAÇÃO.

- (1) Obvio.
- (2) Usando o Lema 2.12 (2) temos que:

$$\begin{aligned}
 P(T_k = n) &= P(N_{n-1} = k - 1, X_n = 1) \\
 &= P(N_{n-1} = k - 1)P(X_n = 1) \\
 &= \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} p \\
 &= \binom{n-1}{k-1} p^k (1-p)^{n-k}
 \end{aligned}$$

□

Vamos a mostrar que o processo das chegadas $\{T_k\}_{k \geq 1}$ é Markoviano. Vejamos porque com um exemplo. Suponha que para certa realização ω do processo de Bernoulli é tal que:

$$T_1(\omega) = 3, T_2(\omega) = 4, T_3(\omega) = 5, T_4(\omega) = 12.$$

Considere o evento $T_5(\omega) = 17$, sabendo que $T_4(\omega) = 12$, este evento vai a acontecer se e somente se:

$$X_{13}(\omega) = 0, X_{14}(\omega) = 0, X_{15}(\omega) = 0, X_{16}(\omega) = 0, X_{17}(\omega) = 1.$$

i.e. os tempos T_1, T_2 e T_3 não intervêm na determinação do evento $T_5(\omega) = 17$, somente precisamos dos valores de T_4 (e dos valores de X_{12}, \dots, X_{16} que são eventos que acontecem depois do tempo T_4), ver a figura 6:

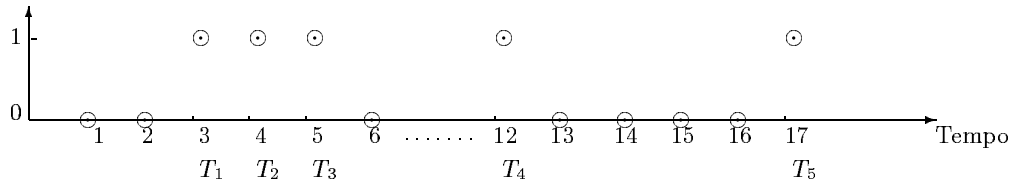


Figura 6

Provemos, então que:

TEOREMA 2.14. *Para todo $k \in \mathbb{N}$, e para todo $n \geq k$ vale:*

$$P(T_{k+1} = n | T_0, \dots, T_k) = P(T_{k+1} = n | T_k)$$

DEMONSTRAÇÃO. Sejam $0 < t_1 < \dots < t_k = t$ e $n > t$, então:

$$\begin{aligned} & P(T_{k+1} = n | T_0 = 0, T_1 = t_1, \dots, T_k = t) \\ &= P(X_{t+1} = 0, X_{t+2} = 0, \dots, X_n = 1) \\ &= p(1-p)^{n-(1+t)} = p(1-p)^{n-(1+T_k)} \\ &= P(T_{k+1} = n | T_k = t) \end{aligned}$$

□

Decorre da prova anterior que:

$$P(T_{k+1} = T_k + m | T_0, T_1, \dots, T_k) = p(1-p)^{(T_k+m)-(1+T_k)} = p(1-p)^{m-1}$$

logo:

TEOREMA 2.15. *Para todo $k \in \mathbb{N}$, e para todo $m \geq 0$ vale:*

$$P(T_{k+1} - T_k = m) = P(T_{k+1} - T_k = m | T_0, T_1, \dots, T_k) = p(1-p)^{m-1}$$

2. O PROCESSO DE BERNOULLI E OUTROS PROCESSOS ESTOCÁSTICOS ASSOCIADOS

Mais isto significa que $T_{k+1} - T_k \sim \text{Geométrica}(p)$ e que o processo $\{T_k\}_{k \geq 1}$ tem incrementos independentes e estacionários, pois a distribuição de $T_{k+1} - T_k$ não depende de k .

De outra parte:

$$T_k = (T_k - T_{k-1}) + (T_{k-1} - T_{k-2}) + \cdots + (T_2 - T_1) + T_1$$

Logo T_k é a soma de k variáveis aleatórias i.i.d com distribuição Geométrica(p). Isto é $T_k \sim \text{Binomial Negativa}(k, p)$, logo:

$$\begin{aligned}\mathbb{E}[T_k - T_{k-1}] &= \frac{1}{p}, \\ \text{Var}[T_k - T_{k-1}] &= \frac{1-p}{p^2}, \\ \mathbb{E}[T_k] &= \frac{k}{p} \quad \text{e} \quad \text{Var}[T_k] = k \frac{1-p}{p^2}\end{aligned}$$

EXEMPLO 2.16. Calcule $P(T_1 = 3, T_5 = 9, T_7 = 17)$.

$$\begin{aligned}P(T_1 = 3, T_5 = 9, T_7 = 17) &= P(T_1 = 3, T_5 - T_1 = 6, T_7 - T_5 = 8) \\ &= P(T_1 = 3)P(T_5 - T_1 = 6)P(T_7 - T_5 = 8) \\ &= P(T_1 = 3)P(T_4 - T_0 = 6)P(T_2 - T_0 = 8) \\ &= P(T_1 = 3)P(T_4 = 6)P(T_2 = 8) \\ &= \binom{3-1}{1-1} p^1 (1-p)^2 \binom{6-1}{4-1} p^4 (1-p)^2 \binom{8-1}{2-1} p^2 (1-p)^6 \\ &= (1p^1(1-p)^2) \left(\frac{5 \cdot 4 \cdot 3}{1 \cdot 2 \cdot 3} p^4 (1-p)^2 \right) \left(\frac{7}{1} p^2 (1-p)^6 \right) = 70p^7(1-p)^{10}\end{aligned}$$

EXEMPLO 2.17. Os componentes de um certo dispositivo tem tempo de vida (o tempo até falhar) aleatório. Suponha que os componentes são reemplazados imediatamente após falhar e que o tempo de vida, $U_k, k \geq 1$, de cada um dos componentes não depende dos outros e tem distribuição Geométrica(p).

Se chamamos T_k os tempos nos quais acontecem as falhas, então: $U_k = T_k - T_{k-1}$ e $P(U_k) = p(1-p)^{m-1}, m \geq 1$. Estes tempos $T_k, k \geq 1$ seriam os tempos dos sucessos de um processo de Bernoulli.

Suponha que foram observados $T_1 = 3, T_2 = 12, T_3 = 14$, queremos estimar T_5 . Para isso vamos a calcular

$$\mathbb{E}[T_5 | T_1 = 3, T_2 = 12, T_3 = 14]$$

Pela propriedade de Markov:

$$\mathbb{E}[T_5 | T_1 = 3, T_2 = 12, T_3 = 14] = \mathbb{E}[T_5 | T_3 = 14]$$

Observe que:

$$\begin{aligned}\mathbb{E}[T_5|T_3] &= \mathbb{E}[T_5 - T_3 + T_3|T_3] \\ &= \mathbb{E}[T_5 - T_3|T_3] + \mathbb{E}[T_3|T_3] \\ &= \mathbb{E}[T_5 - T_3] + T_3 = \mathbb{E}[T_2 - T_0] + T_3 = \frac{2}{p} + T_3\end{aligned}$$

logo:

$$\mathbb{E}[T_5|T_1 = 3, T_2 = 12, T_3 = 14] = \frac{2}{p} + 14$$

CAPÍTULO 2

Cadeias de Markov a Tempo Discreto

1. Definições e exemplos

Para introduzir a noção de cadeia de Markov, vamos considerar alguns exemplos.

EXEMPLO 1.1. *Ruina do jogador*

Considere um jogo no qual em cada aposta você perde um real com probabilidade 0,6 ou o ganha com probabilidade 0,4. Suponha também que você decide parar de jogar se a sua fortuna atingir N reais e se ela atingir 0 reais o casino não deixa você jogar mais.

Seja X_n a quantidade de dinheiro que você tem depois de n apostas. Observe que para prever o próximo estado X_{n+1} , é suficiente conhecer a sua fortuna no "presente" (X_n) e não no "passado" ($X_{n-1}, X_{n-2}, \dots, X_0$). De fato, se $X_n = i$, com $0 < i < N$, então independentemente dos valores i_0, \dots, i_{n-1} , teremos que

$$P(X_{n+1} = i + 1 | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = 0,4$$

pois isso significa que você ganha a aposta $n + 1$, a sua fortuna vai ser acrescentada em um e portanto é suficiente conhecer o valor da sua fortuna no presente.

DEFINIÇÃO 1.2. *Uma cadeia de Markov é um processo estocástico $\{X_n\}_{n \in \mathbb{T}}$, com o tempo discreto, $T = \{0, 1, 2, \dots\}$, o espaço de estados E finito ou enumerável e que tem a propriedade de Markov:*

$$P(X_{n+1} = j | X_0 = i_0, \dots, X_n = i_n) = P(X_{n+1} = j | X_n = i_n), \quad (1.3)$$

para todos $i_0, \dots, i_n, j \in E$ e todo $n \in \mathbb{T}$. Se $X_n = i$ dizemos que o processo no instante n esta no estado i .

A equação (1.3) diz que o estado *futuro* do processo, $X_{n+1} = j$, não depende do *passado*, $X_0 = i_0, \dots, X_{n-1} = i_{n-1}$, e depende só do *presente*, $X_n = i_n$.

A probabilidade condicional (1.3) é chamada **probabilidade de transição**.

Vamos restringir o nosso estudo às cadeias de Markov **homogêneas**, isto é aquelas cadeias nas quais (1.3) não depende do tempo n :

$$P(X_{n+1} = j | X_n = i) = \dots = P(X_1 = j | X_0 = i) = P_{i,j}, \quad i, j \in E,$$

logo $P_{i,j}$ é a probabilidade de passar, *em qualquer instante*, do estado i ao estado j .

É comum arranjar as probabilidades de transição $P_{i,j}$ numa matriz P .

(1) Se E é finito, por exemplo $E = \{0, 1, \dots, N\}$, então:

$$P = \begin{bmatrix} P_{0,0} & P_{0,1} & \dots & P_{0,N} \\ P_{1,0} & P_{1,1} & \dots & P_{1,N} \\ \dots & \dots & \dots & \dots \\ P_{N,0} & P_{N,1} & \dots & P_{N,N} \end{bmatrix} \leftarrow \text{transições do estado } 0 \text{ aos estados } 0, 1, \dots, N$$

No caso do exemplo do jogador, as probabilidades de transição não nulas valem

$$P_{i,i+1} = 0.4, \quad P_{i,i-1} = 0.6, \quad \text{se } 0 < i < N,$$

$$P_{0,0} = 1 = P_{N,N}.$$

Para $N = 5$, a matriz de transição é

$$\begin{array}{c} \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{pmatrix} 1.0 & 0 & 0 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

(2) Se E é infinito, por exemplo $E = \{0, 1, 2, \dots\}$, então:

$$P = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \dots \\ P_{1,0} & P_{1,1} & P_{1,2} & \dots \\ P_{2,0} & P_{2,1} & P_{2,2} & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

Também podemos descrever as transições como grafos:

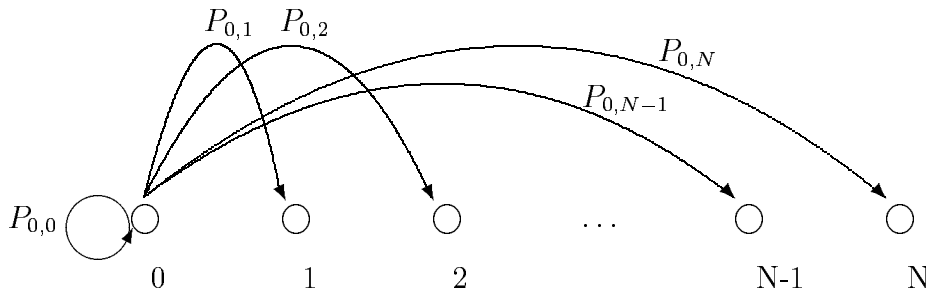


Figura 7

As setas entre os estados correspondem as transições, e o grafo é chamado **topologia** da cadeia.

EXEMPLO 1.3. **Cadeia de Ehrenfest** Suponha que o total de bolas contidas em duas urnas é N . A cada instante de tempo n , pegamos uma bola da primeira urna e a colocamos na segunda ou viceversa, pegamos uma da segunda e a colocamos na primeira. Definamos X_n como a quantidade de bolas na primeira urna. Então X_n é uma cadeia de Markov com espaço de estados $E = \{0, 1, \dots, N\}$. Calculemos as probabilidades de transição.

Observe que se em algum instante não tivermos bolas na primeira urna então necessariamente no instante seguinte teremos que passar uma bola da segunda urna para a primeira. Portanto $P_{0,1} = 1$. Analogamente teremos que $P_{N,N-1} = 1$. Se $1 < i < N$, então $P_{i,i-1} = i/N$ e $P_{i,i+1} = (N-i)/N$.

Para $N = 3$ a matriz de transição é

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

EXERCÍCIO 1.4. No exemplo 1.2, assuma que o número de clientes que quer comprar uma máquina de lavar cada dia é 0, 1, 2 ou 3 com probabilidade 0.3, 0.4, 0.2 e 0.1 respectivamente. Prove que $\{X_n\}$ é uma cadeia de Markov e determine a sua matriz de transição.

Vimos que a cada cadeia de Markov corresponde uma matriz de transição. Que tipo de matrizes dão lugar a cadeias de Markov?

DEFINIÇÃO 1.5. A matriz $P = (P_{i,j})_{i,j \in E}$ é uma **matriz estocástica** se:

- (1) $P_{i,j} \geq 0 \quad i, j \in E$.
- (2) $\sum_{j \in E} P_{i,j} = 1 \quad i \in E$.

Em outras palavras, todas as entradas de uma matriz estocástica são não negativas e qualquer linha tem soma um. Observe que toda matriz de transição é uma matriz estocástica. De fato, a condição (1) corresponde a que as entradas são valores de probabilidades e a (2) a que se o processo está no estado i no instante n , então no próximo instante ele terá que estar em algum dos estados $j \in E$. Por outro lado, vale também que toda matriz estocástica determina uma cadeia de Markov X_n da seguinte forma. Estando no estado i , sorteamos um dos valores $j \in E$ de acordo com a distribuição dada por $P_{ij}, j \in E$.

Um exemplo muito importante de cadeia de Markov com espaço de estados infinito é o seguinte.

EXEMPLO 1.6. Passeio aleatório simples.

No passeio aleatório simples o espaço de estados são os inteiros não negativos, i.e. $E = \mathbb{Z}$. As transições só ocorrem entre estados vizinhos, $P_{i,i+1} = p = 1 - P_{i,i-1}$, com $0 \leq p \leq 1$. Se $p = 0$ as transições são somente para a esquerda e se $p = 1$ elas são só para a direita.

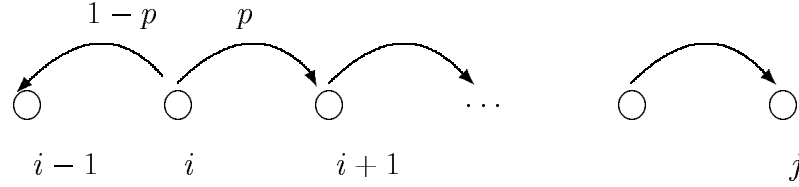


Figura 8

Quando $p = 1/2$ as transições satisfazem,

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

e a cadeia vai de um estado para o da esquerda ou para o da direita com a mesma probabilidade. Por esta razão neste caso, o processo chama-se de passeio aleatório simples simétrico.

EXEMPLO 1.7. Considere uma seqüência de variáveis aleatórias ξ_1, ξ_2, \dots i.i.d. discretas e com distribuição $P(\xi_i = k) = p_k$, $k \in \mathbb{N} \cup \{0\}$ com $\sum_{k=0}^{\infty} p_k = 1$.

- (1) Seja X_0 uma variável aleatória independente da seqüência $\{\xi_j\}_{j \in \mathbb{N}}$, e defina $X_n = \xi_n$ para $n \in \mathbb{N}$. As probabilidades de transição desta cadeia são:

$$P_{i,j} = P(X_{n+1} = j | X_n = i) = P(\xi_{n+1} = j | \xi_n = i) = P(\xi_{n+1} = j) = p_j,$$

e a matriz de transição é:

$$P = \begin{bmatrix} p_0 & p_1 & p_2 & \dots \\ p_0 & p_1 & p_2 & \dots \\ p_0 & p_1 & p_2 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

- (2) Seja $X_0 = \xi_0 = 0$ e $X_n = \xi_1 + \dots + \xi_n$, observe que $X_{n+1} = X_n + \xi_{n+1}$. Então:

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

Logo:

$$P_{i,j} = \begin{cases} p_{j-i}, & j \geq i \\ 0, & \text{c.c.} \end{cases}$$

é a matriz de transição é:

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

EXEMPLO 1.8. (*Número de Sucessos*)

Considere o número de sucessos, $N_n, n \geq 0$ num processo de Bernoulli com probabilidade p de sucesso. O espaço de estados é $E = \{0, 1, 2, \dots\}$, e devido ao fato que $N_0 = 0$, a distribuição inicial é $\pi(0) = 1$ e $\pi(j) = 0, j \geq 1$. Pelo exemplo anterior, N_n é uma cadeia de Markov com transições:

$$P_{i,j} = P(N_{n+1} = j | N_n = i) = P(X_{n+1} = j - i) = \begin{cases} p, & j - i = 1 \\ 1 - p & j = i \\ 0 & \text{c.c.} \end{cases}$$

Logo a matriz de transição é:

$$P = \begin{bmatrix} 1 - p & p & 0 & 0 & \dots \\ 0 & 1 - p & p & 0 & \dots \\ 0 & 0 & 1 - p & p & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

EXERCÍCIO 1.9. (1) Represente a topologia do processo N_n .

(2) Reobtenha a distribuição das N_n usando 2.5.

EXERCÍCIO 1.10. (*Tempo dos Sucessos*)

Observe que o processo dos tempos dos sucessos num processo de Bernoulli também é um caso particular do exemplo 1.7. Determine a matriz de transição e a topologia deste processo.

Sugestão: Faça $\xi_j = T_j - T_{j-1}$.

No que segue descreveremos outras aplicações das cadeias de Markov.

1.1. Fila a Tempo Discreto. Os usuários de certo serviço fazem uma *fila* para ser atendidos, eles são atendidos na ordem de chegada. O tempo que demora o atendimento é fixo, digamos um minuto. Se não tem gente na fila não é feito nenhum serviço. Durante o serviço chegam novos clientes aleatoriamente. O número ξ_n de clientes que chegam no instante n não depende dos que chegaram antes e tem a mesma distribuição, i.e. $P(\xi_n = k) = p_k, k = 0, 1, \dots$. Se no início do serviço na fila tem i pessoas, depois de um período de serviço o número de pessoas na fila será:

(1) $i - 1 + \xi_1$ se $i \geq 1$.

(2) ξ_1 se $i = 0$.

Considere X_n , o número de pessoas na fila no instante n . Então X_n é uma c.m. Para calcular as probabilidades de transição observe que:

$$X_{n+1} = (X_n - 1)^+ + \xi_{n+1}, \quad \text{onde } a^+ = \max\{a, 0\}.$$

Usando isto, obtemos:

$$P_{0,j} = P(X_{n+1} = j | X_n = 0) = P(\xi_{n+1} = j) = p_j$$

$$\begin{aligned} P_{i,j} &= P(X_{n+1} = j | X_n = i) = P((i-1)^+ + \xi_{n+1} = j) \\ &= P((i-1) + \xi_{n+1} = j) = p_{j-i+1}, j \geq i \geq 1. \end{aligned}$$

Logo,

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

1.2. Inventário. Um produto é armazenado para satisfazer certa demanda. Suponha que a demanda, Z_n , no n -ésimo dia é aleatória e que a seqüência de demandas diárias $\{Z_n\}_{n \geq 1}$ são variáveis aleatórias i.i.d e independentes do estoque inicial do produto.

O estoque do produto é completado no fim do dia de acordo a seguinte estratégia, que chamaremos de (s, S) , com $0 < s < S \in \mathbb{N}$. Se, após satisfazer a demanda Z_n , o estoque atinge o nível (inferior) s então é feita a reposição do estoque até o nível (superior) S . Se o estoque não atinge o nível inferior s então não é feita a reposição. Seja X_n a quantidade no estoque depois de satisfazer a demanda e antes de utilizar a estratégia (s, S) para completar novamente o estoque.

Vejamos uma realização do processo X_n . Suponha que $s = 2$, $S = 5$ e que inicialmente o estoque tem 4 unidades (i.e. $X_0 = 4$).

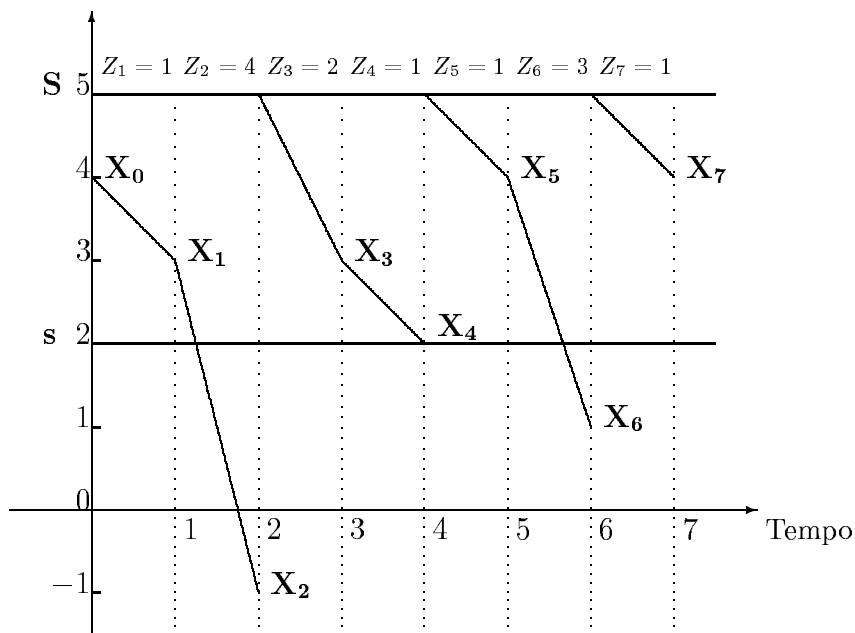


Figura 9

Logo:

$$X_{n+1} = \begin{cases} X_n - Z_{n+1}, & s < X_n \leq S. \\ S - Z_{n+1}, & X_n \leq s. \end{cases}$$

EXERCÍCIO 1.11. *Suponha que $P(Z_n = 1) = p = 1 - P(Z_n = 2)$. Comprove que X_n definido como acima é uma cadeia de Markov e ache a sua matriz de transição P .*

EXERCÍCIO 1.12. *Comprove que o exercício 1.4 representa uma cadeia de Markov do tipo descrito acima com $(s, S) = (1, 5)$, $P(Z_1 = 0) = 0.3$, $P(Z_1 = 1) = 0.4$, $P(Z_1 = 2) = 0.2$ e $P(Z_1 = 3) = 0.1$.*

1.3. Corrida de Sucessos. Considere uma seqüência de v.a. independentes $\{T_n\}_{n \geq 0}$ que toman so dois valores s (sucesso) ou f (falha) com probabilidade $P(T_n = s) = p$ e $P(T_n = f) = q$ com $p + q = 1$. Seja X_n o número de sucessos consecutivos (ou corrida de sucessos) no instante n , suponha que X_n é zero se tiver uma falha no instante n .

Uma realização seria:

n	1	2	3	4	5	6	7	8	9	10	...
T_n	s	s	f	f	s	f	f	s	s	s	...
X_n	1	2	0	0	1	0	0	1	2	3	...

X_n é uma cadeia de Markov com topologia:

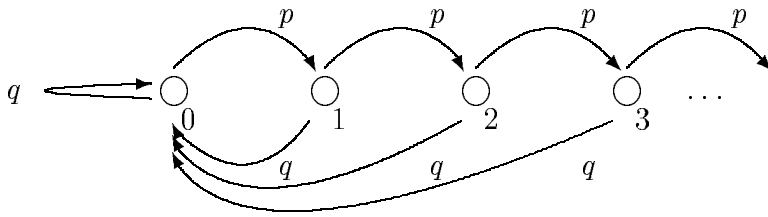


Figura 10

e matriz de transição

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

1.4. Processo de Ramificação (Processo de Galton-Watson). Considere a seguinte seqüência de variáveis aleatórias independentes com valores inteiros não negativos:

- 1) $\{Z_1^{(j)}\}_{j \geq 1}$ i.i.d.
- 2) $\{Z_2^{(j)}\}_{j \geq 1}$ i.i.d.
- ...
- n) $\{Z_n^{(j)}\}_{j \geq 1}$ i.i.d.
- ...

Defina o seguinte processo:

$$X_0 = 1, \quad X_{n+1} = \sum_{k=1}^{X_n} Z_n^{(k)}, \quad \text{e} \quad X_{n+1} = 0 \quad \text{se} \quad X_n = 0$$

X_n representa o número de indivíduos na geração n . Observe que no instante $n + 1$ somamos os descendentes dos X_n indivíduos da geração anterior, i.e. $Z_n^{(k)}$ são os descendentes do k -ésimo indivíduo ($1 \leq k \leq X_n$) da geração n . O processo X_n é chamado processo de **ramificação**. Para entender o porquê deste nome, observe a seguinte representação de uma realização do processo.

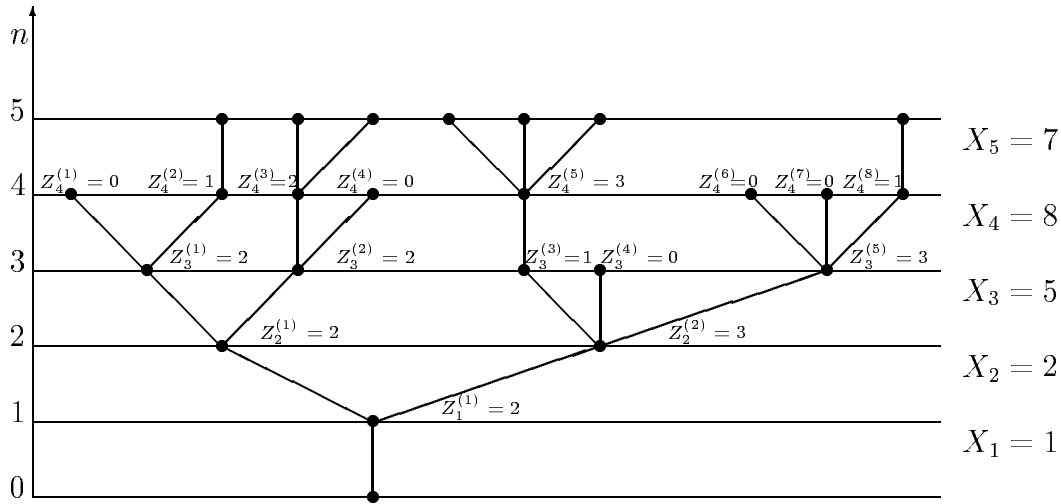


Figura 11

Vale que

$$P_{i,j} = P(X_{n+1} = j | X_n = i) = P(Z_n^{(1)} + Z_n^{(2)} + \dots + Z_n^{(i)} = j)$$

1.5. Processo de vida e morte. Uma cadeia de Markov com espaço de estados $E = \{0, 1, \dots, d\}$ (ou espaço de estados $E = \{0, 1, \dots, \infty\}$, no caso infinito) e com probabilidades de transição:

$$P_{i,j} = \begin{cases} q_i, & j = i - 1, \\ r_i, & j = i, \\ p_i, & j = i + 1, \end{cases}$$

onde $p_i + r_i + q_i = 1$ e $q_0 = 0$ e ($p_d = 0$ se $d < \infty$), é chamada *Processo de Nascimento e Morte*.

2. Matrizes de transição de ordem superior.

EXEMPLO 2.1. No exemplo 1.3 o processo X é uma cadeia de Markov se supormos que as mudanças de status estão dadas pela seguinte matriz

$$\begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix} \end{matrix}$$

Ela nos diz, por exemplo, que a probabilidade de que os filhos de uma família de classe baixa permaneçam nesta classe é 0.7.

EXEMPLO 2.2. Para exemplo acima, suponha que a família começa na classe média (estado 2) na geração 0. Qual a probabilidade que a geração 1 ascenda à classe alta (estado 3) e a geração 2 desca para a baixa (estado 1)?

Devemos calcular a probabilidade $P(X_1 = 3, X_2 = 1 | X_0 = 2)$. Observe que esta é a probabilidade de ir em um passo do estado 2 para o 1 e depois do 1 para o 3. Intuitivamente, pela propriedade de Markov, esta probabilidade deve ser $P_{2,1}P_{1,3}$. Vejamos que isto é assim.

$$\begin{aligned} P(X_1 = 3, X_2 = 1 | X_0 = 2) &= \frac{P(X_1 = 3, X_2 = 1, X_0 = 2)}{P(X_2 = 1, X_0 = 2)} \frac{P(X_2 = 1, X_0 = 2)}{P(X_0 = 2)} \\ &= P(X_1 = 3 | X_2 = 1, X_0 = 2) P(X_2 = 1 | X_0 = 2) \\ &= P(X_1 = 3 | X_2 = 1) P(X_2 = 1 | X_0 = 2) \end{aligned}$$

Generalizando o exemplo:

TEOREMA 2.3. Sejam $i_0, i_1, i_2, \dots, i_m \in E$, vale que:

$$\begin{aligned} &P(X_{n+1} = i_1, \dots, X_{n+m} = i_m | X_n = i_0) \\ &= P(X_{n+m} = i_m | X_{n+m-1} = i_{m-1}) \dots P(X_{n+1} = i_1 | X_n = i_0) \\ &= P_{i_{m-1}, i_m} P_{i_{m-2}, i_{m-1}} \dots P_{i_0, i_1} \\ &= P_{i_0, i_1} P_{i_1, i_2} \dots P_{i_{m-1}, i_m} \end{aligned}$$

DEMONSTRAÇÃO. Exercício. □

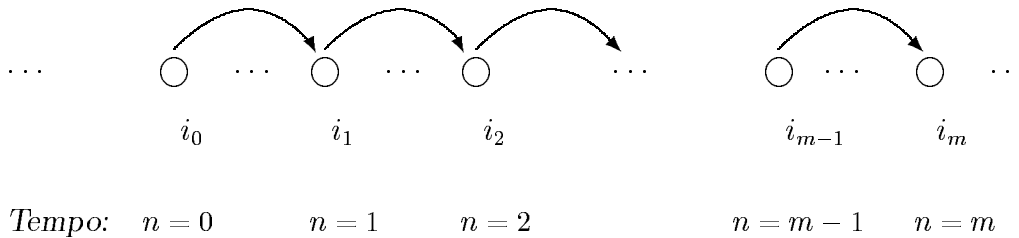


Figura 12

Voltando ao exemplo 2.2,

EXEMPLO 2.4. Suponha de novo que a família começa na classe média (estado 2) na geração 0. Qual a probabilidade que a geração 2 desca para a classe baixa (estado 1)?

Para resolver este problema, devemos considerar os três estados possíveis para a geração 1 e usar o teorema 2.3.

$$\begin{aligned}
 P(X_2 = 1|X_0 = 2) &= \sum_{k=1}^3 P(X_1 = k, X_2 = 1|X_0 = 2) \\
 &= \sum_{k=1}^3 P_{2,k}P_{k,1} \\
 &= 0.3 \cdot 0.7 + 0.5 \cdot 0.3 + 0.2 \cdot 0.2 \\
 &= 0.4
 \end{aligned}$$

De forma similar é possível provar que para $i, j \in \{1, 2, 3\}$ vale

$$P(X_2 = j|X_0 = i) = \sum_{k=1}^3 P_{i,k}P_{k,j}.$$

Observe que o termo da direita na igualdade anterior é o coeficiente (i, j) da matriz $P^2 = P \cdot P$. O termo da esquerda é a probabilidade de passar do estado i ao j em dois passos.

De forma geral vale que a probabilidade de transição em m passos de uma cadeia de Markov X

$$P_{i,j}^m = P(X_{m+n} = j|X_n = i)$$

é a entrada (i, j) da m -ésima potência da matriz de transição P , isto é, da matriz $P^m = P \cdot P \cdots P$ onde há m termos no produto. Provaremos isto a seguir.

Para as transições de **ordem dois**, isto é entre os tempos n e $n + 2$ vale:

$$\begin{aligned}
 P_{i,j}^2 &= P(X_{n+2} = j|X_n = i) \\
 &= \sum_{k \in E} P(X_{n+2} = j, X_{n+1} = k|X_n = i) \\
 &= \sum_{k \in E} P(X_{n+2} = j|X_{n+1} = k)P(X_{n+1} = k|X_n = i) \\
 &= \sum_{k \in E} P_{i,k}P_{k,j}.
 \end{aligned}$$

O termo direito da última expressão é o elemento (i, j) da matriz $P^2 = P \cdot P$. Analogamente podemos encontrar as transições de ordem três (i.e. entre os tempos n e $n + 3$):

$$\begin{aligned} P_{i,j}^3 &= P(X_{n+3} = j | X_n = i) \\ &= \sum_{k \in E} P_{i,k} P_{k,j}^2 \\ &= \sum_{k \in E} P_{i,k} \left[\sum_{l \in E} P_{k,l} P_{l,j} \right]. \end{aligned}$$

Na última expressão aparece agora o elemento (i, j) da matriz $P^3 = PPP = PP^2$. Em geral vale que $P_{i,j}^m$ é o elemento (i, j) da matriz P^m . Isto decorre do seguinte resultado.

TEOREMA 2.5. (*Equações de Chapman Kolmogorov*)

$$P_{i,j}^{m+n} = \sum_{k \in E} P_{i,k}^m P_{k,j}^n,$$

Observe que se chamarmos de $P^{(m)} = (P_{i,j}^m)$ à matriz de transição de ordem m , teremos que o teorema acima afirma que $P^{(m+n)} = P^{(m)} \cdot P^{(n)}$. Como $P^{(1)} = P$, temos então que $P^{(n+1)} = P \cdot P^{(n)}$ e usando um argumento indutivo obtemos que $P^{(n)} = P^n$.

DEMONSTRAÇÃO. Temos que,

$$P(X_{n+m} = j | X_0 = i) = \sum_{k \in E} P(X_{n+m} = j, X_m = k | X_0 = i). \quad (2.4)$$

Usando a definição da probabilidade condicional, cada um dos somandos pode ser escrito da forma,

$$\begin{aligned} P(X_{n+m} = j, X_m = k | X_0 = i) &= \frac{P(X_{n+m} = j, X_m = k, X_0 = i)}{P(X_0 = i)} \\ &= \frac{P(X_{n+m} = j, X_m = k, X_0 = i)}{P(X_m = k, X_0 = i)} \frac{P(X_m = k, X_0 = i)}{P(X_0 = i)} \\ &= P(X_{n+m} = j | X_m = k, X_0 = i) P(X_m = k | X_0 = i) \\ &= P(X_{n+m} = j | X_m = k) P(X_m = k | X_0 = i), \end{aligned}$$

onde, na última linha usamos a propriedade de Markov. Substituindo na igualdade 2.4, obtemos o resultado desejado. \square

Veremos no que segue como as distribuições conjuntas de estados do processo estão determinadas pela matriz de transição e a distribuição de probabilidade do estado inicial.

DEFINIÇÃO 2.6. (*Distribuição Inicial*)
Seja π uma distribuição de probabilidades no conjunto E :

$$\pi(i) \geq 0, i \in E, \quad \sum_{i \in E} \pi(i) = 1,$$

dizemos que π é a distribuição inicial da cadeia se para todo $i \in E$ vale $P(X_0 = i) = \pi(i)$.

Observe que a distribuição inicial de uma cadeia é a função de probabilidade do seu estado inicial X_0 .

O teorema da probabilidade total nos permite obter a distribuição de qualquer um dos estados em função da matriz de transição e da distribuição inicial. De fato, para todo $n \in \mathbb{N}$,

$$P(X_n = k) = \sum_{i \in E} P(X_n = k | X_0 = i) P(X_0 = i) = \sum_{i \in E} P_{i,k}^n \pi(i) = \pi^t P^n. \quad (2.5)$$

Aqui π representa o vetor coluna dos valores da distribuição inicial da cadeia.

Usando o teorema 2.3, podemos obter o seguinte resultado.

PROPOSIÇÃO 2.7. *Seja π a distribuição inicial da cadeia $\{X_n\}_{n \geq 0}$ que tem matriz de transição $P = (P_{i,j})_{i,j \in E}$. Sejam $i_0, i_1, i_2, \dots, i_m \in E$ então vale:*

$$P(X_0 = i_0, X_1 = i_1, \dots, X_m = i_m) = \pi(i_0) P_{i_0, i_1} \dots P_{i_{m-1}, i_m}$$

DEMONSTRAÇÃO.

$$\begin{aligned} & P(X_0 = i_0, X_1 = i_1, \dots, X_m = i_m) \\ &= P(X_1 = i_1, \dots, X_m = i_m | X_0 = i_0) P(X_0 = i_0) \\ &= P_{i_0, i_1} \dots P_{i_{m-1}, i_m} \pi(i_0) \end{aligned}$$

□

Mais geralmente é possível provar,

PROPOSIÇÃO 2.8. (*Distribuições Finito-Dimensionais*)
Seja π a distribuição inicial da cadeia $\{X_n\}_{n \geq 0}$ que tem matriz de transição $P = (P_{i,j})_{i,j \in E}$. Sejam $i_0, i_1, i_2, \dots, i_m \in E$ e $n_1 < n_2 < \dots < n_m$ então vale:

$$P(X_0 = i_0, X_{n_1} = i_1, \dots, X_{n_m} = i_m) = \pi(i_0) P_{i_0, i_1}^{n_1} P_{i_1, i_2}^{n_2 - n_1} \dots P_{i_{m-1}, i_m}^{n_m - n_{m-1}}$$

DEMONSTRAÇÃO. Exercício.

□

EXEMPLO 2.9. Considere uma cadeia de Markov X_n com espaço de estados $E = \{a, b, c\}$ e matriz de transição:

$$P = \begin{array}{ccc|c} a & b & c & \\ \hline \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & a \\ \frac{2}{3} & 0 & \frac{1}{3} & b \\ \frac{3}{5} & \frac{2}{5} & 0 & c \end{array}$$

- (1) A partir desta matriz podemos construir o grafo das transições ou topologia da cadeia.

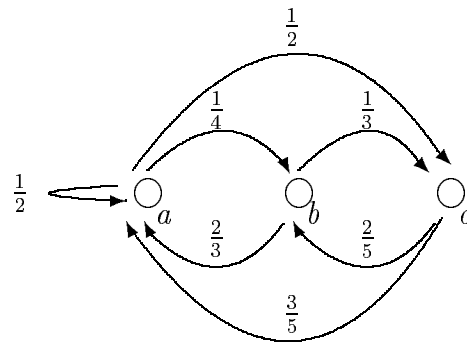


Figura 13

Reciprocamente podemos achar a matriz de transição P a partir do grafo.

- (2) Calcule $P(X_1 = b, X_2 = c, X_3 = a | X_0 = a)$.

$$\begin{aligned} P(X_1 = b, X_2 = c, X_3 = a | X_0 = a) &= P(X_1 = b | X_0 = a)P(X_2 = c | X_1 = b)P(X_3 = a | X_2 = c) \\ &= P_{a,b}P_{b,c}P_{c,a} = \frac{1}{4} \frac{1}{3} \frac{3}{5} = \frac{1}{20} \end{aligned}$$

- (3) Calcule $P(X_6 = c | X_4 = b)$.
Achemos primeiro P^2 .

$$P^2 = \begin{bmatrix} \frac{17}{30} & \frac{9}{40} & \frac{5}{24} \\ \frac{8}{15} & \frac{3}{10} & \frac{1}{6} \\ \frac{17}{30} & \frac{3}{20} & \frac{17}{60} \end{bmatrix}$$

Agora observe que $P(X_6 = c|X_4 = b) = P_{b,c}^2 = \frac{1}{6}$

(4) Calcule $P(X_6 = c, X_4 = a|X_3 = a)$.

$$\begin{aligned} P(X_6 = c, X_4 = a|X_3 = a) &= P(X_4 = a|X_3 = a)P(X_6 = c|X_4 = a) \\ &= P_{a,a}P_{a,c}^2 = \frac{1}{2} \frac{5}{24} = \frac{5}{48}. \end{aligned}$$

(5) Se a distribuição inicial da cadeia é $\pi(a) = 0.3$, $\pi(b) = 0.3$ e $\pi(c) = 0.4$, determine a função de probabilidade de X_2 .

Chamando de π ao vetor coluna

$$\pi = \begin{bmatrix} \pi(a) \\ \pi(b) \\ \pi(c) \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.3 \\ 0.4 \end{bmatrix},$$

sabemos que o vetor da função de probabilidades de X_2 está dado por

$$\pi^t P^2 = [0.3 \ 0.3 \ 0.4] \begin{bmatrix} \frac{17}{30} & \frac{9}{40} & \frac{5}{24} \\ \frac{8}{15} & \frac{3}{10} & \frac{1}{6} \\ \frac{17}{30} & \frac{3}{20} & \frac{17}{60} \end{bmatrix} = \left[\frac{167}{300} \ \frac{87}{400} \ \frac{271}{1200} \right].$$

Portanto,

$$P(X_2 = a) = \frac{167}{300}, \quad P(X_2 = b) = \frac{87}{400} \quad e \quad P(X_2 = c) = \frac{271}{1200}.$$

3. Comportamento assintótico de cadeias com dois estados.

Vimos que quando o espaço de estados é finito podemos calcular as probabilidades de transição de ordem superior multiplicando a matriz de transição P por ela mesma. Fazendo isto podemos observar que em alguns casos, depois de certa ordem as filas vão se aproximando entre si. Este fenômeno é chamado de **estacionariedade**. Vejamos isto com um exemplo.

EXEMPLO 3.1. Considere a cadeia de Markov X_n com espaço de estados $E = \{0, 1\}$ e matriz de transição,

$$P = \begin{bmatrix} 0,5 & 0,5 \\ 0,3 & 0,7 \end{bmatrix}$$

Então:

$$P^2 = \begin{bmatrix} 0,40 & 0,60 \\ 0,36 & 0,64 \end{bmatrix} \quad P^3 = \begin{bmatrix} 0,38 & 0,62 \\ 0,372 & 0,628 \end{bmatrix} \quad P^4 = \begin{bmatrix} 0,3760 & 0,6240 \\ 0,3744 & 0,6256 \end{bmatrix}.$$

Vamos examinar o exemplo com mais cuidado. Como antes, o espaço de estados será $E = \{0, 1\}$. A matriz de transição necessariamente toma a forma

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

com a topologia correspondente,

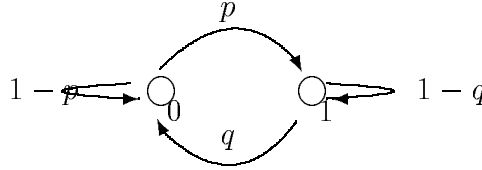


Figura 14

Isto é

$$\begin{aligned} P(X_{n+1} = 0 | X_n = 0) &= 1-p & P(X_{n+1} = 1 | X_n = 0) &= p \\ P(X_{n+1} = 0 | X_n = 1) &= q & P(X_{n+1} = 1 | X_n = 1) &= 1-q \end{aligned}$$

Observe que se $p + q > 0$ podemos escrever a matriz P como

$$P = \frac{1}{p+q} \begin{bmatrix} q & p \\ q & p \end{bmatrix} + \frac{1-p-q}{p+q} \begin{bmatrix} p & -p \\ -q & q \end{bmatrix}.$$

Usando as relações

$$\begin{bmatrix} q & p \\ q & p \end{bmatrix}^2 = (p+q) \begin{bmatrix} q & p \\ q & p \end{bmatrix}, \quad \begin{bmatrix} q & p \\ q & p \end{bmatrix} \begin{bmatrix} p & -p \\ -q & q \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

e

$$\begin{bmatrix} p & -p \\ -q & q \end{bmatrix}^2 = \frac{1-p-q}{p+q} \begin{bmatrix} p & -p \\ -q & q \end{bmatrix},$$

é possível provar por um argumento indutivo que

$$P^n = \frac{1}{p+q} \begin{bmatrix} q & p \\ q & p \end{bmatrix} + \frac{(1-p-q)^n}{p+q} \begin{bmatrix} p & -p \\ -q & q \end{bmatrix}. \quad (3.6)$$

Estudemos o comportamento de P^n quando $n \rightarrow \infty$. Para isto devemos considerar três casos.

- $p + q = 0$, i.e. $p = 0$ e $q = 0$

Neste caso P é a matriz identidade de ordem dois, vale que $P^n = P$ para todo $n \in \mathbb{N}$ e portanto as linhas não se aproximam entre si. Ambos os estados 0 e 1 são absorventes, portanto a cadeia vai visitar em todos os instantes o estado do qual ela começou.

- $p + q = 2$, i.e. $p = 1$ e $q = 1$

Agora

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

Se n for par, teremos que P^n é a matriz identidade de ordem dois e para n ímpar, $P^n = P$. Como consequência disto temos que o limite de P^n quando $n \rightarrow \infty$ não existe pois a matriz oscila entre duas matrizes fixas.

- $0 < p + q < 2$

Neste caso vale que $|1 - p - q| < 1$ e portanto $(1 - p - q)^n \rightarrow 0$ quando $n \rightarrow \infty$. Usando (3.6), obtemos

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{bmatrix}.$$

No último caso considerado, as linhas da matriz convergem para uma distribuição que chamaremos de π_{est} , isto é,

$$\lim_{n \rightarrow \infty} P_{0,0}^n = \lim_{n \rightarrow \infty} P_{1,0}^n = \pi_{est}(0),$$

$$\lim_{n \rightarrow \infty} P_{0,1}^n = \lim_{n \rightarrow \infty} P_{1,1}^n = \pi_{est}(1),$$

com $\pi_{est}(0) = \frac{q}{p+q}$ e $\pi_{est}(1) = \frac{p}{p+q}$. Estas quantidades podem ser interpretadas como as probabilidades da cadeia estar a longo prazo no estado 0 ou no 1, por isto π_{est} é chamada de **distribuição estacionária**. A relação (3.6) permite obter também uma estimativa para a taxa de convergência das probabilidades de transição em n passos para a distribuição estacionária.

PROPOSIÇÃO 3.2. *Para uma cadeia de Markov $\{X_n\}_{n \geq 1}$ com dois estados, $E = \{0, 1\}$ e tal que $0 < p + q < 2$, vale*

$$|P_{i,0}^n - \pi_{est}| = \left| P_{i,0}^n - \frac{q}{p+q} \right| \leq |1 - p - q|^n,$$

com $i = 0$ ou $i = 1$.

Estas probabilidades de transição se aproximam do equilíbrio com velocidade exponencial, ou seja, muito rápido. Por isso de forma geral quando $p + q < 2$, observamos a proximidade entre as linhas de P^n e a distribuição estacionária para valores de n pequenos.

EXEMPLO 3.3. No exemplo (3.1) temos que $p = 0,5$ e $q = 0,3$, portanto $p + q < 2$ e as linhas da matriz devem convergir para $(\pi_{est}(0) \ \pi_{est}(1)) = (3/8 \ 5/8) = (0,375 \ 0,625)$. A diferença entre elas vai para zero mais rápido que $0,2^n$.

A estacionariedade significa que para instantes de tempo grandes, a cadeia "esquece" o estado onde ela começou. É natural pensar que se no lugar de sair no instante inicial de um estado fixo, considerarmos para X_0 uma distribuição de probabilidade qualquer, ela também será "esquecida": a distribuição de X_n para valores grandes do n se aproxima de π_{est} . Provemos que isto é assim.

Suponha que a distribuição inicial é π com $\pi(0) = \pi_0$, $\pi(1) = \pi_1$ e $\pi_0 + \pi_1 = 1$. Em outras palavras, temos $P(X_0 = 0) = \pi_0$ e $P(X_0 = 1) = \pi_1$. Sabemos que $(P(X_n = 0) \ P(X_n = 1)) = (\pi_0 \ \pi_1) \cdot P^n$. Usando a expressão que temos para P^n e o fato que π é uma distribuição, obtemos

$$\begin{aligned} P(X_n = 0) &= \frac{q}{p+q} + (1-p-q)^n \left[\pi_0 - \frac{q}{p+q} \right], \\ P(X_n = 1) &= \frac{p}{p+q} + (1-p-q)^n \left[\frac{q}{p+q} - \pi_0 \right]. \end{aligned}$$

Como no caso que estamos considerando vale $(1-p-q)^n \rightarrow 0$ quando $n \rightarrow \infty$, concluímos que

$$\lim_{n \rightarrow \infty} P(X_n = 0) = \frac{q}{p+q} = \pi_{est}(0), \quad \lim_{n \rightarrow \infty} P(X_n = 1) = \frac{p}{p+q} = \pi_{est}(1),$$

que era o que queríamos provar.

Concluímos que quando o espaço de estados possui somente dois elementos é possível determinar completamente o comportamento assintótico da cadeia. Não há distribuição limite em dois casos. Um é quando a matriz de transição é a identidade de ordem dois, ou seja, a cadeia restrita a cada um dos estados é uma cadeia de Markov (constante). No outro caso, a cadeia oscila entre os dois estados. Veremos mais adiante que no caso finito, estas são essencialmente as duas possibilidades nas quais não há distribuição limite. A estrutura do espaço de estados é de fundamental importância no comportamento assintótico das cadeias de Markov.

4. Classificação dos estados. Parte I.

DEFINIÇÃO 4.1. Um estado $i \in E$ é dito **não essencial** se é possível sair dele numa quantidade finita de passos e não voltar nunca mais. Isto é, se existe um instante $n \in \mathbb{N}$ tal que $P_{i,j}^n > 0$ e vale $P_{j,i}^m = 0$ para todo $m \in \mathbb{N}$.

Todos os demais estados do espaço de estados são chamados estados **essenciais**. Se a cadeia atingir um estado essencial, nunca mais volta para um não essencial. Estes são os estados mais interessantes. Observe que para um estado i ser essencial tem que ocorrer uma das duas alternativas seguintes.

- (1) Não é possível sair de i .

Neste caso $P_{i,j}^n = 0$ para todo $j \in E, j \neq i$ e para todo $n \in \mathbb{N}$. Em particular $P_{i,j} = 0$ para todo $j \neq i$ e portanto $P_{i,i} = 1$. Quando isto ocorre dizemos que o estado i é **absorvente**, pois uma vez que a cadeia atinge um estado deste tipo, nunca mais sai.

- (2) Sempre que $P_{i,j}^n > 0$ teremos que existe $m \in \mathbb{N}$ tal que $P_{j,i}^m > 0$.

Ou sempre que a cadeia sair de i numa quantidade finita de passos, pode voltar a este estado também numa quantidade finita de passos.

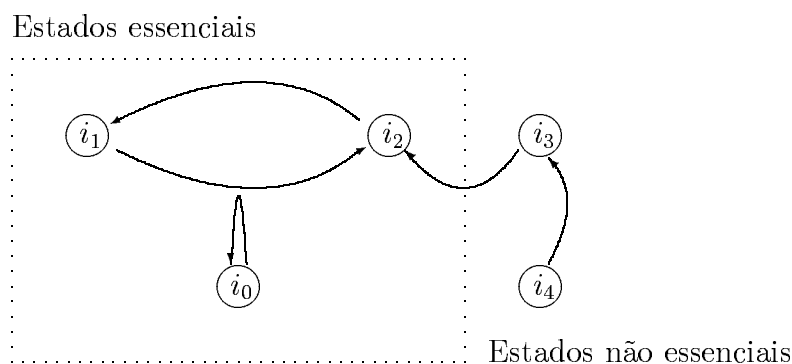
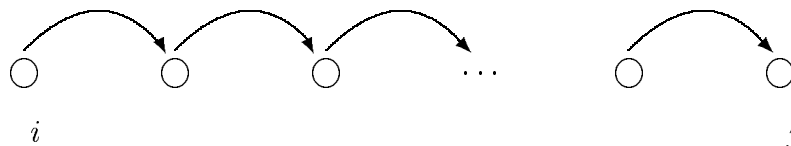


Figura 15

DEFINIÇÃO 4.2. Dizemos que o estado j é **acessível desde o estado i** , $i \rightarrow j$, se existe um instante $n > 0$ tal que $P_{i,j}^n > 0$. Se $i \rightarrow j$ e $j \rightarrow i$ diremos que **i e j estão comunicados** e o denotaremos por $i \leftrightarrow j$.



j acessível desde i

Figura 16.

EXEMPLO 4.3. No passeio aleatório simples com $0 < p < 1$, $i \rightarrow j$ para todo $i, j \in E$. Portanto, todos os estados são essenciais e estão comunicados. Se $p = 0$ ou $p = 1$, todos os estados são não essenciais pois sempre será possível sair deles e não voltar mais.

Repare que para todo estado essencial i vale que se $i \rightarrow j$ então $i \leftrightarrow j$. A relação de comunicação \leftrightarrow é reflexiva, simétrica e transitiva, isto é,

PROPOSIÇÃO 4.4. (1) $i \leftrightarrow i$. (consequência de $P_{i,i}^0 = 1$)
 (2) $i \leftrightarrow j \Rightarrow j \leftrightarrow i$. (consequência da definição)
 (3) $i \leftrightarrow j$ e $j \leftrightarrow k \Rightarrow i \leftrightarrow k$ (consequência de Chapman-Kolmogorov).

Para cada estado essencial, podemos considerar a classe de todos os estados comunicados com ele. Desta maneira obtemos uma decomposição do conjunto dos estados essenciais em classes disjuntas de estados comunicados com a propriedade de que é impossível para a cadeia passar de uma classe para a outra.

DEFINIÇÃO 4.5. Classes fechadas de estados

- (1) Um conjunto \mathcal{C} de estados é fechado se nenhum estado $j \notin \mathcal{C}$ é acessível desde algum estado $i \in \mathcal{C}$.
- (2) $\mathcal{C} = \{i\}$ é fechado se e somente se i é absorvente.

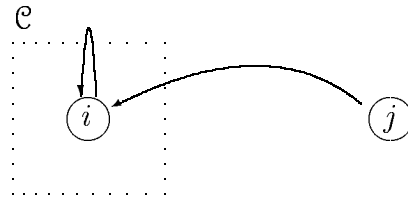


Figura 17

- (3) Um conjunto fechado é dito conjunto ou classe irredutível se ele não possui nenhum subconjunto próprio fechado.
- (4) Uma cadeia de Markov é dita irredutível se o conjunto de estados E for uma classe irredutível.

As classes irredutíveis coincidem com as classes disjuntas de estados comunicados. Logo, uma cadeia é irredutível se e somente se todos os seus estados se comunicam.

EXEMPLO 4.6. O passeio aleatório simples é uma cadeia irredutível quando $0 < p < 1$.

EXEMPLO 4.7. O passeio aleatório com barreiras absorventes tem espaço de estados $E = \{1, 2, \dots, n\}$ e topologia:

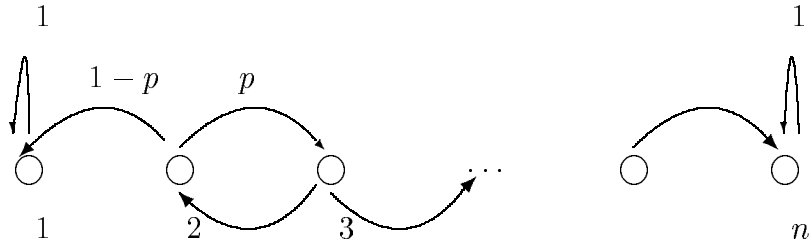


Figura 18

Os estados 1 e n são absorventes e são as únicas classes irredutíveis da cadeia. Os demais estados são não essenciais.

PROPOSIÇÃO 4.8. Suponha que a classe $\mathcal{C} = \{c_1, c_2, \dots, c_n\} \subset E$ é fechada e defina as seguintes probabilidades de transição na classe \mathcal{C} :

$$\tilde{P}_{i,j} = P_{c_i, c_j} \quad c_i, c_j \in \mathcal{C}$$

Então $\tilde{P} = (\tilde{P}_{i,j})$ é uma matriz de transição.

DEMONSTRAÇÃO. Observe que $\tilde{P}_{i,j} = P_{c_i, c_j} \geq 0$ e que:

$$\sum_{j=1}^n \tilde{P}_{i,j} = \sum_{j=1}^n P_{c_i, c_j} = \sum_{c_j \in \mathcal{C}} P_{c_i, c_j} + \sum_{k \notin \mathcal{C}} P_{c_i, k} = \sum_{l \in E} P_{c_i, l} = 1$$

pois $P_{c_i, k} = 0$ se $k \notin \mathcal{C}$. □

Pelo resultado anterior temos que uma cadeia de Markov restrita a um conjunto fechado é uma cadeia de Markov com matriz \tilde{P} .

EXEMPLO 4.9. Considere uma cadeia de Markov com espaço de estados $E = \{a, b, c, d, e\}$ e matriz de transição:

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

A topologia da cadeia é:

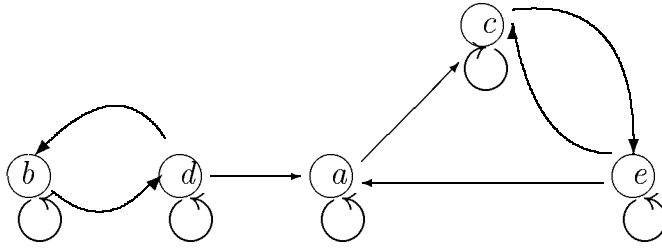


Figura 19

Os estados não essenciais são $\{b, d\}$ e a classe $\mathcal{C} = \{a, c, e\}$ é irreduzível, logo a cadeia não é irreduzível e a matriz

$$\tilde{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

é a matriz de transição da cadeia restrita à classe \mathcal{C} .

Uma cadeia restrita a uma das suas classes irreduzíveis será irreduzível. Neste sentido, poderemos nos limitar ao estudo de cadeias irreduzíveis.

EXEMPLO 4.10. A cadeia de Ehrenfest (exemplo 1.3) é irreduzível. Consideremos de novo a cadeia de Ehrenfest com 3 bolas. A matriz de transição é

$$P = \begin{array}{ccccc} & \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{bmatrix} & \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} \end{array}$$

e para a segunda potência de P , os zeros se deslocam

$$P^2 = \begin{array}{cccc|c} & \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \\ \mathbf{0} & 1/3 & 0 & 2/3 & 0 & \mathbf{0} \\ \mathbf{1} & 0 & 7/9 & 0 & 2/9 & \mathbf{1} \\ \mathbf{2} & 2/9 & 0 & 7/9 & 0 & \mathbf{2} \\ \mathbf{3} & 0 & 2/3 & 0 & 1/3 & \mathbf{3} \end{array}$$

Esta oscilação dos zeros vai continuar para matrizes de ordens superiores, pois observe que se partirmos de um estado par, necessariamente depois de um número ímpar de passos vamos passar a um estado ímpar e depois de um número par de passos passaremos a algum estado par. Devido a esta oscilação, não existe distribuição limite neste caso.

Se dividirmos o espaço de estados nos conjuntos $D_0 = \{0, 2\}$ e $D_1 = \{1, 3\}$, vemos que toda vez que a cadeia sair de D_0 tem que entrar em D_1 e viceversa. Isto sugere a decomposição das classes irredutíveis em subclasses cíclicas.

DEFINIÇÃO 4.11. Diremos que um estado i tem período $d = d(i)$ se as seguintes condições são satisfeitas:

- (1) $P_{i,i}^n > 0$ só para valores de n da forma $n = dm$, isto é, para n múltiplo de d ;
- (2) d é o maior valor satisfazendo (1).

Em outras palavras, d é o maior divisor comum dos números n para os quais vale $P_{i,i}^n > 0$.

5. Tempos de parada

Precisaremos nos referir aqui a instantes de tempo aleatórios.

DEFINIÇÃO 5.1. Uma variável aleatória $\tau : \Omega \rightarrow \mathbb{Z}^+ \cup \{\infty\}$ é chamada de **tempo de parada** em relação ao processo $\{X_n\}_{n \geq 0}$ quando para todo $n \in \mathbb{Z}^+$ a ocorrência do evento da forma $\{\tau = n\}$ depende somente dos valores de X_0, X_1, \dots, X_n .

EXEMPLO 5.2. Tempo de chegada. Considere o conjunto $A \subseteq E$, onde E é o espaço de estados da cadeia de Markov $\{X_n\}_{n \geq 0}$. O tempo de chegada ao conjunto A , τ_A e o menor instante depois do instante inicial no qual o processo $\{X_n\}_{n \geq 0}$ assume um valor no conjunto A :

$$\tau_A = \begin{cases} \min \{n > 0 : X_n \in A\}, \\ \infty, & \text{se } \forall n \quad X_n \notin A. \end{cases}$$

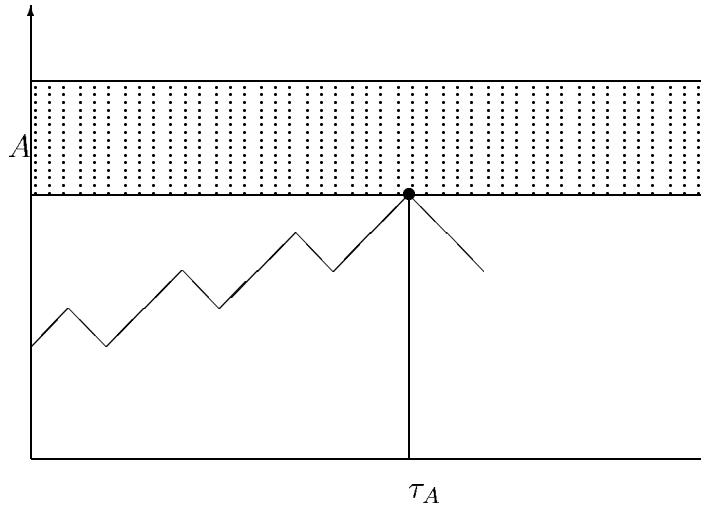


Figura 20

Observe que $[\tau_A = n] = [X_n \in A] \cap [X_j \notin A, \text{ para todo } j \text{ com } 1 \leq j \leq n-1]$, portanto τ_A é um tempo de parada.

Se o conjunto A tiver um só elemento, $A = \{a\}$, escrevemos $\tau_A = \tau_a$.

EXEMPLO 5.3. *Tempos determinísticos.* Um tempo constante ($\tau = c$, com $c \in \mathbb{Z}^+$) é um tempo de parada.

Os tempos de parada são não antecipativos no sentido que pode-se decidir se $[\tau = m]$ olhando somente os estados do processo até o instante m , ou seja, X_0, X_1, \dots, X_m .

EXEMPLO 5.4. *Contraexemplo: tempos de saída.* O tempo de saída do conjunto A , U_A e o último instante que o conjunto é visitado pela cadeia $\{X_n\}_{n \geq 0}$:

$$U_A = \begin{cases} \max \{n > 0 : X_n \in A\}, \\ 0, & \text{se } \forall n \ X_n \notin A, \\ +\infty & \text{se } \forall n \ X_n \in A. \end{cases}$$

Agora $[U_A = n] = [X_n \in A] \cap [X_j \notin A, \text{ para todo } j > n]$ e U_A não é um tempo de parada, pois ele fornece informação sobre o processo em instantes posteriores a ele próprio.

Não é difícil verificar se τ_1 e τ_2 são dois tempos de parada então que a sua soma $\tau = \tau_1 + \tau_2$ e o mínimo entre eles $\tau = \min\{\tau_1, \tau_2\} = \tau_1 \wedge \tau_2$ também o são.

Observe que com a notação introduzida, $X_{\tau \wedge 0}, \dots, X_{\tau \wedge n}$ são estados visitados pela cadeia em instantes consecutivos e anteriores a τ . Vamos definir a seguir a **propriedade de Markov forte**.

TEOREMA 5.5. *Seja $\{X_n\}_{n \geq 0}$ uma cadeia de Markov com espaço de estados E e matriz de transição P e seja τ um tempo de parada em relação a esta cadeia. Então para cada estado $i \in E$, dado que $X_\tau = i$ vale o seguinte:*

- (1) *O processo depois de τ e o processo antes são independentes,*
- (2) *O processo depois de τ é uma cadeia de Markov com matriz de transição P .*

DEMONSTRAÇÃO. (1) Devemos provar que para todos os tempos $k \geq 1$ e $n \geq 0$ e todos os estados $i_0, i_1, \dots, i_n, i, j_1, \dots, j_k$,

$$\begin{aligned} P(X_{\tau+1} = j_1, \dots, X_{\tau+k} = j_k | X_\tau = i, X_{\tau \wedge 0} = i_0, \dots, X_{\tau \wedge n} = i_n) \\ = P(X_{\tau+1} = j_1, \dots, X_{\tau+k} = j_k | X_\tau = i). \end{aligned}$$

Provaremos na verdade uma igualdade mais simples:

$$P(X_{\tau+k} = j | X_\tau = i, X_{\tau \wedge n} = i_n) = P(X_{\tau+k} = j | X_\tau = i), \quad (5.7)$$

pois o caso geral pode ser obtido por argumentos similares. O termo da esquerda é igual a

$$\frac{P(X_{\tau+k} = j, X_\tau = i, X_{\tau \wedge n} = i_n)}{P(X_\tau = i, X_{\tau \wedge n} = i_n)}$$

e vale

$$P(X_{\tau+k} = j, X_\tau = i, X_{\tau \wedge n} = i_n) = \sum_{r \geq 0} P(\tau = r, X_{r+k} = j, X_r = i, X_{r \wedge n} = i_n). \quad (5.8)$$

Mas

$$\begin{aligned} P(\tau = r, X_{r+k} = j, X_r = i, X_{r \wedge n} = i_n) \\ = P(X_{r+k} = j | X_r = i, X_{r \wedge n} = i_n, \tau = r) P(X_r = i, X_{r \wedge n} = i_n, \tau = r) \end{aligned}$$

e como $r \wedge n \leq r$ e $[\tau = r]$ depende dos valores de X_0, \dots, X_r ¹, o evento $[X_{r \wedge n} = i_n, \tau = r]$ também depende destes valores ou em outras palavras, depende da história da cadeia até o instante r . Pela propriedade de Markov $P(X_{r+k} = j | X_r = i, X_{r \wedge n} = i_n, \tau = r) = P(X_{r+k} = j | X_r = i) = P_{i,j}^k$ e (5.8) se reduz a

$$\sum_{r \geq 0} P_{i,j}^k P(\tau = r, X_{r \wedge n} = i_n, X_r = i) = P_{i,j}^k P(X_{\tau \wedge n} = i_n, X_\tau = i)$$

e portanto o termo esquerdo de (5.7) coincide com $P_{i,j}^k$. Cálculos similares mostram que o termo da direita também o faz, pelo qual ficaria provado o resultado desejado.

¹e é aqui que usamos o fato de τ ser um tempo de parada

(2) Devemos provar que para todos os estados $i, j, k, i_{n-1}, \dots, i_1$

$$\begin{aligned} P(X_{\tau+n+1} = k | X_{\tau+n} = j, X_{\tau+n-1} = i_{n-1}, \dots, X_{\tau} = i) \\ = P(X_{\tau+n+1} = k | X_{\tau+n} = j) = P_{j,k}. \end{aligned}$$

A primeira igualdade é conseqüência de (1) para o tempo de parada $\tau' = \tau + n$, os processos antes e depois de τ' são independentes dado $X_{\tau'} = j$. A segunda igualdade pode ser obtida usando cálculos similares aos apresentados acima. \square

6. Visitas a um Estado

Fixemos um estado $j \in E$ e estudemos as sucessivas visitas que o processo $\{X_n\}_{n \geq 0}$ faz a esse estado. Vamos a definir os tempos das visitas:

1) O primeiro tempo τ_1 seria:

$$\tau_1 = \min \{n : X_n = j\}$$

com $\tau_1 = \infty$ se o estado j não é visitado (i.e. se $\forall n : X_n \neq j$).

2) O segundo tempo τ_2 depois de τ_1 :

$$\tau_2 = \min \{n : n \geq \tau_1, X_n = j\}$$

com $\tau_2 = \infty$ se $\forall n \geq \tau_1 : X_n \neq j$.

.....

k) O k -éssimo tempo τ_k depois de τ_{k-1} :

$$\tau_k = \min \{n : n \geq \tau_{k-1}, X_n = j\}$$

com $\tau_k = \infty$ se $\forall n \geq \tau_{k-1} : X_n \neq j$.

Se para algum k temos que $\tau_k = \infty$, então definiremos $\infty = \tau_{k+1} - \tau_k = \tau_{k+2} - \tau_{k+1} = \dots$

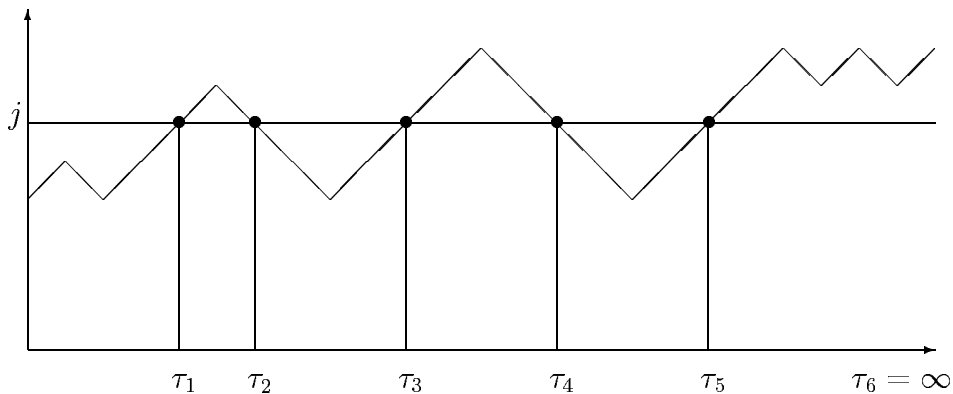


Figura 21

Observe que $\tau_1 \leq \tau_2 \leq \dots \leq \tau_k \leq \dots$.

Lembrando que usamos a notação $P_i(A)$ para representar a probabilidade $P_i(A) = P(A|X_0 = i)$, com $A \subseteq E$. Pela propriedade forte de Markov vale que:

PROPOSIÇÃO 6.1. *Seja $k \in E$, então:*

$$P_i(\tau_{m+1} - \tau_m = k | \tau_1, \tau_2, \dots, \tau_m) = \begin{cases} 0, & \{\tau_m = \infty\}. \\ P_j(\tau_1 = k), & \{\tau_m < \infty\}. \end{cases}$$

O resultado acima indica que dado que a cadeia saindo de i visita m vezes o estado j , ela o visitará de novo depois de k instantes de tempo com a mesma probabilidade que se "apagarmos" a história anterior e olhássemos a probabilidade dela voltar em j pela primeira vez no instante k .

Vamos agora estudar a probabilidade de começando no estado i , a primeira visita ao estado j acontecer no instante k , em símbolos: $F_k(i, j) = P_i(\tau_1 = k)$. Para achar $F_k(i, j)$ usaremos recursão no tempo k . Para $k = 1$, observe que:

$$\begin{aligned} F_1(i, j) &= P_i(\tau_1 = 1) = P(\tau_1 = 1 | X_0 = i) \\ &= P(X_1 = j | X_0 = i) \\ &= P_{i,j} \end{aligned}$$

Para $k \geq 2$:

$$\begin{aligned} F_k(i, j) &= P_i(\tau_1 = k) \\ &= P_i(X_1 \neq j, X_2 \neq j, \dots, X_{k-1} \neq j, X_k = j) \\ &= \sum_{b \neq j} P_i(X_1 = b, X_2 \neq j, \dots, X_{k-1} \neq j, X_k = j) \\ &= \sum_{b \neq j} P_i(X_2 \neq j, \dots, X_{k-1} \neq j, X_k = j | X_1 = b) P_i(X_1 = b) \\ &= \sum_{b \neq j} P_b(\tau_1 = k - 1) P_{i,b}, \end{aligned} \tag{6.9}$$

logo

$$F_k(i, j) = \begin{cases} P_{i,b}, & k = 1. \\ \sum_{b \neq j} F_{k-1}(b, j) P_{i,b}, & k \geq 2. \end{cases}$$

Observe que, pela propriedade de Markov, $F_k(i, j)$ é a probabilidade de ir do estado i ao estado j em k passos, sem visitar nesse tempo o estado j .

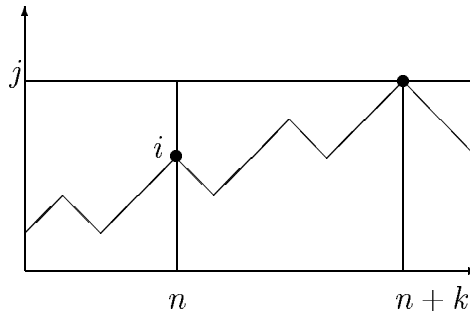


Figura 21

EXEMPLO 6.2. Seja $\{X_n\}_{n \geq 0}$ uma cadeia de Markov com espaço de estados $E = \{1, 2, 3\}$ e matriz de transição:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{3} & \frac{3}{5} & \frac{1}{15} \end{bmatrix}$$

Fixemos o estado $j = 3$. Vamos a achar as probabilidades de ir ao estado j em k transições: $f_k(i) = F_k(i, 3)$ com $i \in E$. Observe que no caso $k = 1$ é fácil, pois $f_1(i) = P_{i,3}$. Logo:

$$f_1 = \begin{bmatrix} 0 \\ \frac{1}{3} \\ \frac{1}{15} \end{bmatrix}$$

Para $k \geq 2$ considere as transições a $j \neq 3$:

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

e usando \tilde{P} recursivamente obtemos $f_k = \tilde{P}f_{k-1}$, assim:

$$f_2 = \begin{bmatrix} 0 \\ \frac{1}{18} \\ \frac{1}{15} \end{bmatrix}, \quad f_3 = \begin{bmatrix} 0 \\ \frac{1}{108} \\ \frac{1}{30} \end{bmatrix}, \quad f_4 = \begin{bmatrix} 0 \\ \frac{1}{648} \\ \frac{1}{180} \end{bmatrix} \cdots$$

Podemos escrever os resultados assim:

$$\begin{aligned} F_k(1, 3) &= 0, \quad k = 1, 2, \dots \\ F_k(2, 3) &= \frac{1}{3} \left(\frac{1}{6}\right)^{k-1}, \quad k = 1, 2, \dots \\ F_k(3, 3) &= \begin{cases} \frac{1}{15}, & k = 1 \\ \frac{3}{5} \left(\frac{1}{6}\right)^{k-2} \left(\frac{1}{3}\right), & k = 2, 3, \dots \end{cases} \end{aligned}$$

Usando estas relações podemos encontrar:

- $P_1(\tau_1 = \infty) = 1$.
- $P_2(\tau_1 = \infty) = 1 - P_2(\tau_1 < \infty) = 1 - \sum_{k=1}^{\infty} P_2(\tau_1 = k) = 1 - \sum_{k=1}^{\infty} \frac{1}{3} \left(\frac{1}{6}\right)^{k-1} = \frac{3}{5}$.
- $P_3(\tau_1 = \infty) = 1 - \sum_{k=1}^{\infty} P_3(\tau_1 = k) = \frac{52}{75}$.

Chamaremos de $F(i, j)$ à probabilidade de começando no estado i eventualmente visitar o estado j ,

$$F(i, j) = P_i(\tau_1 < \infty), \quad \text{com } \tau_1 = \min \{n : X_n = j\}. \quad (6.10)$$

A equação (6.10) pode expressar-se recursivamente como:

$$F(i, j) = P_i(\tau_1 < \infty) = \sum_{k=1}^{\infty} F_k(i, j) \quad (6.11)$$

$$= P_{i,j} + \sum_{b \neq j} P_{i,b} F(b, j) \quad (6.12)$$

O número de visitas ao estado j , N_j será definido como

$$N_j = \sum_{n=0}^{\infty} \mathbf{1}_{X_n=j}.$$

PROPOSIÇÃO 6.3. *Seja $j \in E$, então vale que:*

- (1) $P_j(N_j = m) = F(j, j)^{m-1}(1 - F(j, j))$, $m = 1, 2, 3, \dots$
- (2) Se $i \neq j$, $P_i(N_j = m) = \begin{cases} 1 - F(i, j), & m = 0 \\ F(i, j)F(j, j)^{m-1}(1 - F(j, j)), & m = 1, 2, 3, \dots \end{cases}$

DEMONSTRAÇÃO. Para $m = 0$ o resultado decorre da definição de $F(i, j)$.

Se $m > 0$ e $i \neq j$, o evento $\{N_j = m\}$ é equivalente ao evento $\{\tau_1 < \infty, \dots, \tau_m < \infty, \tau_{m+1} = \infty\}$ que por sua vez é equivalente a que ocorram os seguintes eventos independentes:

$$\{\tau_1 < \infty\}, \{\tau_2 - \tau_1 < \infty\}, \dots, \{\tau_m - \tau_{m-1} < \infty\}, \{\tau_{m+1} - \tau_m = \infty\}$$

que tem probabilidade:

$$F(i, j) \underbrace{F(j, j) \cdots F(j, j)}_{m-1} [1 - F(j, j)].$$

Se $i = j$ podemos usar o mesmo argumento trocando m por $m - 1$. □

Segue desta proposição que:

- (1) Se $F(j, j) = 1$ então $\forall m P_j(N_j = m) = 0$, logo $P_j(N_j < \infty) = 0$, i.e. $P_j(N_j = \infty) = 1$ e $\mathbb{E}_j[N_j] = \infty$.
- (2) Se $F(j, j) < 1$ então

$$\begin{aligned} P_j(N_j < \infty) &= \sum_{m=1}^{\infty} F(j, j)^{m-1} (1 - F(j, j)) \\ &= \frac{1}{1 - F(j, j)} (1 - F(j, j)) = 1 \end{aligned}$$

Observe também que $N_j | X_0 = j \sim \text{Geométrica}(1 - F(j, j))$, logo

$$\mathbb{E}_j[N_j] = \frac{1}{1 - F(j, j)}.$$

DEFINIÇÃO 6.4. (*Matriz Potencial*) A **matriz potencial** da cadeia de Markov $\{X_n\}_{n \geq 1}$ é a matriz $R = (R(i, j))_{i, j \in E}$ tal que $R(i, j) = \mathbb{E}_i[N_j]$.

Observe que é possível que as entradas desta matriz tomem o valor ∞ .

PROPOSIÇÃO 6.5. *Sejam $i, j \in E$, vale:*

- (1) $R(j, j) = \frac{1}{1 - F(j, j)}$.
- (2) $R(i, j) = F(i, j)R(j, j)$, $i \neq j$.

Este resultado é consequência da proposição 6.3.

Observe que $N_j = \sum_{n=0}^{\infty} \mathbf{1}_j(X_n)$, logo

$$R(i, j) = \mathbb{E}_i \sum_{n=0}^{\infty} \mathbf{1}_j(X_n) = \sum_{n=0}^{\infty} P_i(X_n = j) = \sum_{n=0}^{\infty} P_{i, j}^n.$$

Em particular se o espaço de estados E é finito vale,

$$\begin{aligned} R &= I + P + P^2 + \dots, \\ RP &= PR = P + P^2 + P^3 + \dots = R - I, \\ R - RP &= I, \quad R(I - P) = I = (I - P)R. \end{aligned}$$

7. Classificação dos Estados. Parte II

Chamemos de $T_j = \min \{n > 0 : X_n = j\}$ ao instante da primeira visita ao estado j . Então $F(i, j) = P_i(T_j < \infty)$ é a probabilidade da cadeia, que inicia no estado i , visitar eventualmente, i.e. num tempo finito, o estado j . Vamos classificar os estados da cadeia como segue.

DEFINIÇÃO 7.1. *Seja $j \in E$.*

- (1) *O estado j é **recorrente** se $P_j(T_j < \infty) = F(j, j) = 1$.*
- (2) *O estado j é **transitório** se $P_j(T_j < \infty) = F(j, j) < 1$.*

Usando os resultados da seção anterior, podemos caracterizar os estados recorrentes e transitórios da cadeia.

TEOREMA 7.2. (1) *O estado j é transitório se e somente se*

$$P_j(N_j < \infty) = 1 \Leftrightarrow R(j, j) < \infty.$$

Além disto,

$$P_i(N_j < \infty) = 1, \forall i \in E \text{ e } R(i, j) = \frac{F(i, j)}{1 - F(j, j)} < \infty, \forall i \in E.$$

(2) *O estado j é recorrente se e somente se*

$$P_j(N_j = \infty) = 1 \Leftrightarrow R(j, j) = \infty.$$

Além disto, $F(i, j) = P_i(T_j < \infty) = P_i(N_j = \infty)$, $\forall i \in E$. Se $F(i, j) = 0$ então $R(i, j) = 0$ e se $F(i, j) > 0$ então $R(i, j) = \infty$

DEMONSTRAÇÃO. Pelos resultados da seção anterior temos

$$P_i(N_j = \infty) = \lim_{m \rightarrow \infty} P_i(N_j \geq m) = \lim_{m \rightarrow \infty} F(i, j)F(j, j)^{m-1},$$

$$\begin{aligned} R(i, j) &= \mathbb{E}_i(N_j) \\ &= \sum_{m=1}^{\infty} m F(i, j) F(j, j)^{m-1} (1 - F(j, j)). \end{aligned}$$

(7.13)

(1) Se o estado j é transitório, então:

$$\begin{aligned}
 P_i(N_j = \infty) &= 0, \quad \text{pois } F(j, j) < 1. \\
 R(i, j) &= F(i, j) \sum_{m=1}^{\infty} m F(j, j)^{m-1} (1 - F(j, j)) \\
 &= \frac{F(i, j)}{1 - F(j, j)} \tag{7.14}
 \end{aligned}$$

(2) Se o estado j é recorrente, então observe que $F(j, j) = 1$, logo:

$$P_i(N_j = \infty) = F(i, j).$$

Além disto, se $F(i, j) = 0$ teremos $R(i, j) = 0$ e caso contrário $R(i, j) = \infty$, pois $P_i(N_j = \infty) > 0$.

Como já sabemos que $P_j(N_j = \infty) = 1 \Leftrightarrow P_j(T_j < \infty) = 1$, o resultado segue. \square

Em outras palavras, os estados recorrentes são os estados que a cadeia visita infinitas vezes enquanto para os estados transitórios sempre existirá um instante de tempo a partir do qual a cadeia nunca mais volta neles. Como consequência disto, vale o seguinte resultado.

PROPOSIÇÃO 7.3. *Uma cadeia de Markov com espaço de estados finito tem que possuir pelo menos um estado recorrente.*

DEMONSTRAÇÃO. A prova que apresentaremos aqui é um pouco informal. Posteriormente apresentaremos outra prova rigorosa.

Vamos raciocinar pelo absurdo. Suponha que $E = \{1, 2, \dots, M\}$ e que todos estes estados são transitórios. Então existe um instante N_1 tal que depois dele a cadeia nunca mais visita o estado 1. Analogamente existirá N_2 tal que depois deste instante a cadeia nunca mais visita o estado 2 e assim podemos definir os valores N_j até chegar no N_M que será o último instante que a cadeia visita o estado M .

Seja então $N = \max\{N_1, N_2, \dots, N_M\}$. Pela construção feita chegamos a que X_{N+1} não pode ser nenhum dos estados $\{1, 2, \dots, M\}$, só que não tem mais estados no conjunto E . Esta contradição implica que a nossa suposição era falsa e tem que existir pelo menos um estado recorrente. \square

EXERCÍCIO 7.4. *Qual é a falha da prova "informal" apresentada acima?*

Existem cadeias com espaço de estados infinito sem estados recorrentes (veja o exemplo 7.11). Algumas consequências do teorema 7.2 são as seguintes.

PROPOSIÇÃO 7.5. *Um estado j do espaço de estados de uma cadeia de Markov é transitório se e somente se $\sum_{n=0}^{\infty} P_{j,j}^n < \infty$.*

DEMONSTRAÇÃO. Isto é somente uma reescritura da relação

$$\mathbb{E}_j(N_j) < \infty \Leftrightarrow j,$$

que vale para j transitório. □

PROPOSIÇÃO 7.6. *Se o estado j do espaço de estados de uma cadeia de Markov é transitório então para todo $i \in E$ valem*

- (1) $\sum_{n=0}^{\infty} P_{i,j}^n < \infty$;
- (2) $\lim_{n \rightarrow \infty} P_{i,j}^n = 0$.

DEMONSTRAÇÃO. O resultado decorre de $\mathbb{E}_i(N_j) < \infty$ e de que o termo geral de uma série convergente tem que convergir para zero. □

PROPOSIÇÃO 7.7. *Se o estado j é recorrente e $j \rightarrow k$ então*

- (1) $k \rightarrow j$ e $F(k, j) = 1$;
- (2) k é recorrente.

DEMONSTRAÇÃO. (1) Se $j \rightarrow k$ então $F(j, k) > 0$. Mais $F(j, j) = 1$ logo a probabilidade de não voltar ao estado j é $1 - F(j, j) = 0$. De outra parte, uma maneira de não voltar a j seria ir de j a k sem passar em j e depois nunca mais voltar em j , portanto

$$1 - F(j, j) \geq F(j, k)(1 - F(k, j)) \geq 0$$

logo $1 - F(k, j) = 0$, $F(k, j) = 1$ e $k \rightarrow j$.

(2) Mostremos agora que $F(k, k) = 1$, i.e. que k é recorrente. Sejam n_1 e n_2 tais que $P_{k,j}^{n_1} > 0$ e $P_{j,k}^{n_2} > 0$. Por Chapman-Kolmogorov temos $P^{n_1+n_2}(k, k) \geq P_{k,j}^{n_1} P_{j,j}^{n_2} P_{j,k}^{n_2}$. Então como a série $\sum_{n=0}^{\infty} P_{j,j}^n$ é divergente, a série $\sum_{m=0}^{\infty} P_{k,k}^m$ também o será e portanto k é recorrente. □

Observe que $F(i, j) > 0$ se e somente se existe $n > 0$ tal que $P_{i,j}^n > 0$. Os seguintes resultados decorrem facilmente do exposto acima.

TEOREMA 7.8. *Se a classe \mathcal{C} é irredutível, com estados recorrentes, então $F(i, j) = 1$, $P_i(N_j = \infty) = 1$ e $R(i, j) = \infty$ para todo $i, j \in \mathcal{C}$.*

TEOREMA 7.9. *Se a classe \mathcal{C} é irredutível e finita todos os estados são recorrentes.*

A proposição 7.7 afirma que se um estado $j \in E$ é acessível desde um estado recorrente i , então necessariamente $j \rightarrow i$, ou seja, i é um estado essencial, em outras palavras,

COROLÁRIO 7.10. *Todos os estados recorrentes são essenciais.*

Valerá o recíproco do resultado acima? Isto é, serão necessariamente recorrentes os estados essenciais? Para responder esta questão consideremos primeiro uma cadeia com espaço de estados finito. Suponha que um estado $i \in E$ é essencial. A classe $\mathcal{C}(i)$ irredutível de todos os estados (essenciais) comunicados com o estado i tem que ser recorrente porque é finita. Obtemos então que quando E é finito, os estados essenciais e os recorrentes coincidem. Veremos no seguinte exemplo que isto não vale em geral.

EXEMPLO 7.11. *Considere o passeio aleatório simples com $0 < p < 1$.*

Como esta cadeia é irredutível, se provarmos que 0 é transitório, estaremos provando que todos os seus estados o são. Pela proposição 7.5, basta provar que $\sum_{n=0}^{\infty} P_{0,0}^n < \infty$. Mas como esta cadeia só pode voltar a um estado depois de uma quantidade par de passos, na verdade provaremos

$$\sum_{n=0}^{\infty} P_{0,0}^{2n} < \infty.$$

$P_{0,0}^{2n}$ é a probabilidade de saindo do zero, voltar nele em $2n$ transições. Isto só é possível dando n passos para a esquerda e n para a direita, portanto $P_{0,0}^{2n}$ é a probabilidade de uma variável aleatória binomial com parâmetros $2n$ e p ser igual a n , ou seja,

$$P_{0,0}^{2n} = \frac{(2n)!}{(n!)^2} p^n (1-p)^n.$$

A fórmula de Stirling

$$n! \sim \sqrt{2\pi n} n^n e^{-n},$$

permite concluir

$$P_{0,0}^{2n} \sim \frac{(4p(1-p))^n}{\sqrt{\pi n}}.$$

Ou seja, $\sum_{n=0}^{\infty} P_{0,0}^n < \infty \Leftrightarrow p \neq 1/2$. A cadeia é recorrente quando $p = 1/2$ (passeio aleatório simples simétrico). Em todos os demais casos ela é transitória.

EXEMPLO 7.12. *Considere a cadeia de Markov do exemplo 4.9, a classe $\mathcal{C} = \{a, c, e\}$ é finita e irredutível, logo todos os seus estados são recorrentes. Observe que os estados que estão fora de \mathcal{C} são não essenciais e portanto, transitórios.*

TEOREMA 7.13. *Seja \mathcal{C}_R o conjunto de estados recorrentes da classe \mathcal{C} , então podemos decompor \mathcal{C}_R como uma união disjunta de classes:*

$$\mathcal{C}_R = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \dots$$

Onde cada classe \mathcal{C}_i é irredutível.

EXEMPLO 7.14. A cadeia de Markov $X_n \in E = \{1, 2, \dots, 10\}$ definida pela seguinte matriz de transição:

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

Tem a seguinte topologia:

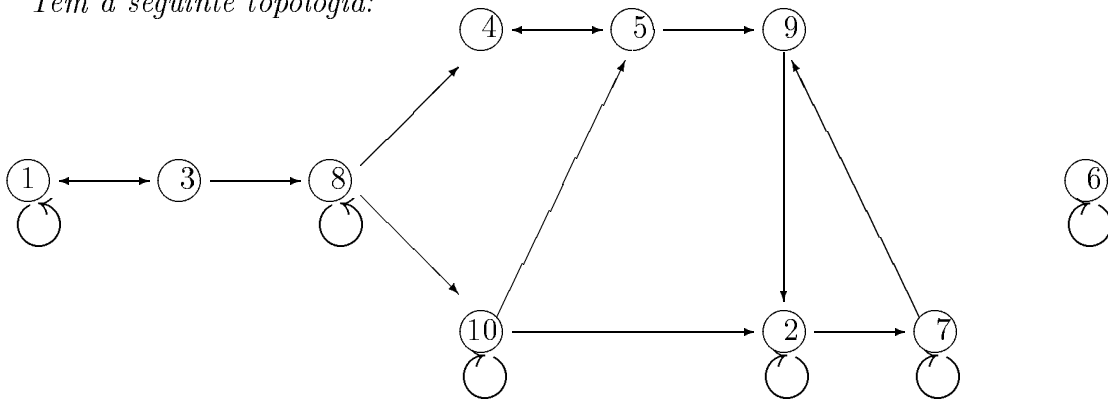


Figura 22

Temos então tres classes recorrentes: $\mathcal{C}_1 = \{1, 3\}$, $\mathcal{C}_2 = \{2, 7, 9\}$ e $\mathcal{C}_3 = \{6\}$. Logo $\mathcal{C}_R = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3$ e $\mathcal{C}_T = \{4, 5, 8, 10\}$ seriam os estados transitórios.

8. Probabilidades de Absorção

Considere a classe \mathcal{C} irredutível, vamos a considerar as transições de estados i à classe \mathcal{C} .

DEFINIÇÃO 8.1. Seja $i \in E$ um estado, definimos a probabilidade de absorção do estado i na classe \mathcal{C} como:

$$\rho_{\mathcal{C}}(i) = P_i(T_{\mathcal{C}} < \infty)$$

i.e. é a probabilidade de chegar na classe \mathcal{C} num tempo finito.

Observe que se $\mathcal{C} = \{j\}$, então o estado j é absorvente e portanto $\rho_{\mathcal{C}}(i) = P_{i,j}$; em particular se $i = j$ então $\rho_{\mathcal{C}}(j) = P_{j,j} = 1$. Também observe que para todo $i \in \mathcal{C}$ vale que

$\rho_{\mathcal{C}}(i) = 1$ e se $i \in \mathcal{C}'$ com $\mathcal{C} \neq \mathcal{C}'$ então $\rho_{\mathcal{C}}(i) = 0$. Logo basta ver o que acontece com os estados i transitórios.

Não é difícil provar que se j e k são dois estados da mesma classe recorrente \mathcal{C} e i é transitório, então $F(i, j) = F(i, k) = \rho_{\mathcal{C}}(i)$. Usando isto e (6.12) obtem-se que se \mathcal{C}_T é o conjunto dos estados transitórios e se \mathcal{C} é uma classe recorrente, então vale:

$$\rho_{\mathcal{C}}(i) = \sum_{j \in \mathcal{C}} P_{i,j} + \sum_{k \in \mathcal{C}_T} P_{i,k} \rho_{\mathcal{C}}(k)$$

No caso que a classe \mathcal{C}_T seja finita é fácil achar a solução desta equação.

EXEMPLO 8.2. *A cadeia de Markov $\{X_n\}$ tem espaço de estados $E = \{0, 1, 2, 3, 4, 5\}$ e a seguinte matriz de transição:*

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 & 0 \\ 0 & 1/5 & 2/5 & 1/5 & 0 & 1/5 \\ 0 & 0 & 0 & 1/6 & 1/3 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1/4 & 0 & 3/4 \end{bmatrix}$$

As classes recorrentes são $\mathcal{C}_1 = \{0\}$ e $\mathcal{C}_2 = \{3, 4, 5\}$. Os estados transitórios são $\mathcal{C}_T = \{1, 2\}$. Após reordenar as filas e columnas da matriz P podemos escrever P como:

$$P = \begin{array}{c} \begin{array}{cccccc} 0 & 3 & 4 & 5 & 1 & 2 \end{array} \\ \left[\begin{array}{c|ccc|cc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/6 & 1/3 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/4 & 0 & 3/4 & 0 & 0 \\ \hline 1/4 & 0 & 0 & 0 & 1/2 & 1/4 \\ 0 & 1/5 & 0 & 1/5 & 1/5 & 2/5 \end{array} \right] \begin{array}{l} 0 \\ 3 \\ 4 \\ 5 \\ 1 \\ 2 \end{array} \end{array}$$

É fácil reconhecer nesta matriz as matrizes de transição correspondentes às classes \mathcal{C}_1 e \mathcal{C}_2 . Seja $\rho_{1,0} = \rho_0(1)$ e $\rho_{2,0} = \rho_0(2)$, então temos de resolver:

$$\begin{aligned} \rho_{1,0} &= P_{1,0} + P_{1,2}\rho_{2,0} + P_{1,0}\rho_{1,0} \\ \rho_{2,0} &= P_{2,1}\rho_{1,0} + P_{2,2}\rho_{2,0} \end{aligned}$$

i.e.

$$\begin{aligned} \rho_{1,0} &= \frac{1}{4} + \frac{1}{4}\rho_{2,0} + \frac{1}{2}\rho_{1,0} \\ \rho_{2,0} &= \frac{1}{5}\rho_{1,0} + \frac{2}{5}\rho_{2,0} \end{aligned}$$

cuja solução é $\rho_{1,0} = \frac{3}{5}$ e $\rho_{2,0} = \rho_0(1)$. Observe que

$$\sum_j \rho_{\mathcal{C}_j}(i) = 1.$$

Logo $\rho_0(1) + \rho_{3,4,5}(1) = 1$, i.e. $\rho_{3,4,5}(1) = \frac{2}{5}$. Também, como $\rho_0(2) = \frac{1}{5}$ temos que $\rho_{3,4,5}(2) = \frac{4}{5}$. Mas, $\rho_{i,j} = \rho_{\mathcal{C}}(i)$ com $i \in \mathcal{C}_T$ e $j \in \mathcal{C}$, então:

$$\begin{aligned} \rho_3(1) &= \rho_4(1) = \rho_5(1) = \frac{2}{5}, \\ \rho_3(2) &= \rho_4(2) = \rho_5(2) = \frac{4}{5}. \end{aligned}$$

Quando existe somente uma classe recorrente \mathcal{C} todo estado transitório é absorvido por esta classe recorrente e vale $\rho_{\mathcal{C}}(i) = 1$ para todo estado i transitório.

9. Exemplos

9.1. O processo de vida e morte. Para a cadeia do Exemplo 1.5, considere três estados $i < j < k$ e defina a probabilidade de começando no estado j , visitar o estado i antes de visitar o estado k : $u(j) = P_j(T_i < T_k)$, onde $u(i) = 1$ e $u(k) = 1$.

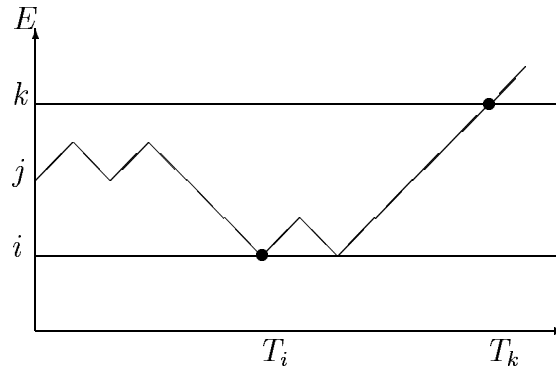


Figura 23

Então vale a seguinte relação:

$$u(j) = q_j u(j-1) + r_j u(j) + p_j u(j+1),$$

lembrando que $r_j = 1 - p_j - q_j$ obtemos

$$u(j) = q_j u(j-1) + (1 - p_j - q_j) u(j) + p_j u(j+1)$$

logo

$$u(j) + q_j u(j) - q_j u(j-1) = u(j) + p_j (u(j+1) - u(j))$$

de onde:

$$u(j+1) - u(j) = \frac{q_j}{p_j} (u(j) - u(j-1)) = \dots = \frac{q_j q_{j-1} \dots q_{i+1}}{p_j p_{j-1} \dots p_{i+1}} (u(i+1) - u(i))$$

Seja $\gamma_j = \frac{q_j q_{j-1} \dots q_1}{p_j p_{j-1} \dots p_1}$, então:

$$u(j+1) - u(j) = \frac{\gamma_j}{\gamma_i} (u(i+1) - u(i))$$

isto é:

$$u(j) - u(j+1) = \frac{\gamma_j}{\gamma_i} (u(i) - u(i+1))$$

e somando $j = i, \dots, k-1$ obtemos:

$$\sum_{j=i}^{k-1} [u(j) - u(j+1)] = \left(\sum_{j=i}^{k-1} \frac{\gamma_j}{\gamma_i} \right) (u(i) - u(i+1)),$$

logo

$$u(i) - u(k) = \frac{\sum_{j=i}^{k-1} \gamma_j}{\gamma_i} (u(i) - u(i+1)).$$

Mais $u(i) = 1$ e $u(k) = 0$ implica:

$$\frac{u(i) - u(i+1)}{\gamma_i} = \left(\sum_{j=i}^{k-1} \gamma_j \right)^{-1}$$

logo

$$u(j) - u(j+1) = \frac{\gamma_j}{\sum_{j=i}^{k-1} \gamma_j}$$

Somando $i = l, \dots, k-1$, com $i < l < k$:

$$\sum_{j=l}^{k-1} (u(j) - u(j+1)) = \frac{\sum_{j=l}^{k-1} \gamma_j}{\sum_{j=i}^{k-1} \gamma_j}$$

de onde, finalmente:

$$u(l) = \frac{\sum_{j=l}^{k-1} \gamma_j}{\sum_{j=i}^{k-1} \gamma_j}, \quad i \leq l < k$$

i.e.

$$P_l(T_i < T_k) = \frac{\sum_{j=l}^{k-1} \gamma_j}{\sum_{j=i}^{k-1} \gamma_j}, \quad P_l(T_k < T_i) = \frac{\sum_{j=i}^{l-1} \gamma_j}{\sum_{j=i}^{k-1} \gamma_j}$$

vejamos num exemplo a utilidade destas relações.

EXEMPLO 9.1. Num jogo as apostas que faz um jogador valem um real, ele ganha um real ou perde a aposta com probabilidade $p = 9/19$, $q = 10/19$ respectivamente. Ele entra no jogo com um capital de 10 reais e se retira do jogo se fica sem dinheiro ou se ganha 25 reais (i.e. 35 reais em total). Ache a probabilidade de que saia do jogo ganhando.

Considere o processo de vida e morte X_n com espaço de estados $E = \{0, 1, \dots, 35\}$, X_n é o capital do jogador no n -ésimo jogo. Se $X_0 = 10$ e $p_i = 9/19$ e $q_i = 10/19$ para $0 < i < 35$, e os estados 0 e 35 são absorventes, queremos achar $P_{10}(T_{35} < T_0)$. Observe que:

$$\gamma_i = \left(\frac{10/19}{9/19}\right)^i = \left(\frac{10}{9}\right)^i, \quad 0 \leq i \leq 34.$$

logo

$$P_{10}(T_{35} < T_0) = \frac{\sum_{i=0}^{34} \left(\frac{10}{9}\right)^i}{\sum_{i=0}^{34} \left(\frac{10}{9}\right)^i} = \frac{\left(\frac{10}{9}\right)^{35} - 1}{\left(\frac{10}{9}\right)^{35} - 1} = 0.47$$

Consideremos o caso infinito $d = \infty$, suponha que, para todo i , $p_i > 0$ e $q_i > 0$ logo a cadeia é irredutível (não finita). Queremos encontrar os estados que são recorrentes e transitórios. Para isso basta ver se o estado 0 é recorrente ou transitório, uma vez que todos os estados se comunicam, isso implica que todos os demais estados serem recorrentes ou transitórios. Primeiro observe que:

$$P_1(T_0 < T_n) = 1 - \frac{1}{\sum_{j=0}^{n-1} \gamma_j}, \quad n > 1$$

e que $\{T_0 < T_n\} \subseteq \{T_0 < T_{n+1}\}$, pois se $X_0 = 1$ então $T_2 < T_3 < \dots$ e $T_n \rightarrow \infty$, logo:

$$\lim_{n \rightarrow \infty} P(T_0 < T_n) = P_1(T_0 < \infty)$$

e isto implica que

$$F(1, 0) = 1 - \frac{1}{\sum_{j=0}^{\infty} \gamma_j}$$

Agora suponha que a cadeia X_n é recorrente, então $F(1, 0) = 1$ e isto implica que $\sum_{j=0}^{\infty} \gamma_j = \infty$. Reciprocamente suponha que $\sum_{j=0}^{\infty} \gamma_j = \infty$, então $F(1, 0) = 1$ e observe que:

$$P_0(T_0 < \infty) = P_{0,0} + P_{0,1}P_1(T_0 < \infty)$$

logo

$$P_0(T_0 < \infty) = r_0 + p_0 = 1$$

i.e. 0 é recorrente e a cadeia é recorrente.

Em resumo o processo de vida e morte é recorrente se e somente se

$$\sum_{j=0}^{\infty} \gamma_j = \sum_{j=0}^{\infty} \frac{q_j q_{j-1} \cdots q_1}{p_j p_{j-1} \cdots p_1} = \infty. \quad (9.15)$$

EXEMPLO 9.2. *Considere o processo de vida e morte dado pela probabilidades:*

$$p_i = \frac{i+2}{2(i+1)}, \quad q_i = \frac{i}{2(i+1)}, \quad i \geq 0.$$

Observe que $\frac{q_i}{p_i} = \frac{i}{i+2}$, logo:

$$\gamma_i = \frac{1 \cdot 2 \cdots i}{3 \cdot 4 \cdots (i+2)} = \frac{2}{(i+1)(i+2)} = 2 \left(\frac{1}{i+1} - \frac{1}{i+2} \right)$$

logo,

$$\sum_{j=0}^{\infty} \gamma_j = 2 \sum_{j=0}^{\infty} \left(\frac{1}{i+1} - \frac{1}{i+2} \right) = 2 \left(\frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \frac{1}{4} - \dots \right) = 2 \left(\frac{1}{2} \right) = 1 < \infty$$

logo X_n é transitório.

9.2. O processo de ramificação. O processo $\{X_n\}$ do exemplo 1.4, representa o tamanho da população, com espaço de estados $E = \{0, 1, 2, \dots\}$. O problema que é considerado neste modelo é a probabilidade de extinção (i.e. $X_n = 0$). Observe que 0 é um estado absorvente do processo. Seja ρ a probabilidade de eventualmente a população ficar extinta se começarmos o processo com um indivíduo:

$$\rho = P_1(T_0 < \infty) = F(1, 0),$$

se a população começa com i pais, então a probabilidade da população ficar extinta seria: $F(i, 0) = P_i(T_0 < \infty)$. Usando o fato que a descendência de cada pai é independente concluímos que

$$F(i, 0) = P_i(T_0 < \infty) = [P_1(T_0 < \infty)]^i = F(1, 0)^i.$$

Assim, sem perda de generalidade, podemos considerar a população inicial composta de um só indivíduo. Lembre que $X_{n+1} = \sum_{k=1}^{X_n} Z_n^{(k)}$ onde as v.a. $Z_n^{(k)}$ representavam o número de descendentes do k -ésimo pai na n -ésima geração. Suponha que $Z_n^{(k)}$ são i.i.d. todas elas com a distribuição de ξ , onde $P(\xi = i) = p_i$. Observe que se $p_1 = 1$ então a população é estática, $X_n = 1$ e $\rho = 0$. Portanto, vamos a considerar o caso $p_1 < 1$, onde o estado 0 é absorvente e os demais estados são transitórios. Temos duas possibilidades, $X_n \rightarrow 0$ ou $X_n \rightarrow \infty$, que correspondem respectivamente a $\rho = 1$ ou $\rho < 1$. Queremos achar condições necessárias para cada caso. As duas possibilidades vão depender do comportamento de ξ , i.e. do número de descendentes; por isso vamos a usar a função geradora de probabilidades de ξ , $M_\xi(t) = \sum_{k=0}^{\infty} t^k p_k = p_0 + \sum_{k=1}^{\infty} t^k p_k$.

Mostremos que ρ é um ponto fixo da função M_ξ , i.e. $M_\xi(\rho) = \rho$

$$\begin{aligned} \rho = F(1, 0) &= P_{1,0} + \sum_{k=1}^{\infty} P_{1,k} F(k, 0), \\ &= P_{1,0} + \sum_{k=1}^{\infty} P_{1,k} \rho^k, \\ &= p_0 + \sum_{k=1}^{\infty} p_k \rho^k, \\ &= M_\xi(\rho). \end{aligned}$$

- (1) Se $\mu = \mathbb{E}\xi \leq 1$ então $M_\xi(t) \neq t$ para todo $t \in [0, 1)$ como se conclui da seguinte figura

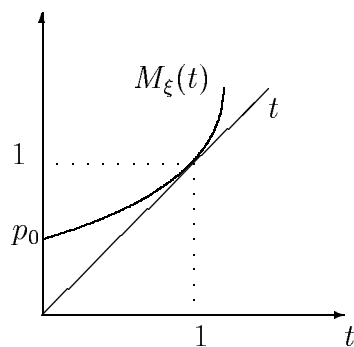


Figura 24

logo $M_\xi(t)$ só tem $\rho = 1$ como ponto fixo, e a probabilidade de extinção é 1.

- (2) Se $\mu = \mathbb{E}\xi > 1$ então M_ξ tem um ponto fixo $\rho \in [0, 1)$, pois nesse caso temos a seguinte figura:

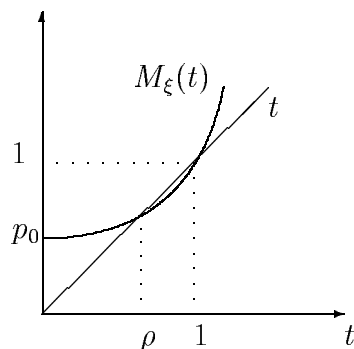


Figura 25

logo $M_\xi(\rho) = \rho$ e $\rho < 1$, e $P(X_n \rightarrow \infty) > 0$.

EXEMPLO 9.3. *Suponha que numa população cada homem de uma família tem 3 filhos, e que a probabilidade de que cada um deles seja homem é $1/2$ e $1/2$ de ser mulher, independentemente dos outros filhos. Logo, se ξ é o número de filhos homens de cada pai, então ξ tem distribuição binomial, $\xi \sim b(3, 1/2)$. Se $P(\xi = k) = p_k$ então*

$$p_0 = \frac{1}{8}, p_1 = \frac{3}{8}, p_2 = \frac{3}{8}, p_3 = \frac{1}{8}$$

e $\mu = \frac{3}{2} > 1$. Se X_n o número de homens na n -ésima geração então, usando os resultados anteriores, a probabilidade de que a linha paterna sobreviva é positiva. Para mostrar isto, procuremos os pontos fixos de M_ξ . Observe que $M_\xi(t) = (\frac{1}{2} + \frac{1}{2}t)^3$ e

$$\left(\frac{1}{2} + \frac{1}{2}t\right)^3 = t, \implies (t-1)(t^2 + 4t - 1) = 0$$

as soluções da equação anterior são: $t = 1$, $t = -\sqrt{5}-2$ e $t = \sqrt{5}-2$. Logo $\rho = \sqrt{5}-2 > 0$ e $\rho < 1$.

Observe que se cada homem tem somente 2 filhos, então $\xi \sim b(2, 1/2)$ e $\mu = 1$, logo $\rho = 1$ e a linha paterna termina certamente.

9.3. Filas. Estudaremos a cadeia X_n que representa o tamanho da fila de clientes que esperam ser atendidos, após o n -ésimo serviço (Exemplo ??). Se as v.a. ξ_n , $n \geq 0$ representam o número de clientes que chegam no instante n (vamos supor que são independentes e identicamente distribuídas, com distribuição $P(\xi = k) = p_k$), então $X_{n+1} = (X_n - 1)^+ + \xi_{n+1}$. Observe que $P_{0,i} = P(\xi = i) = p_i$ e $P_{i,j} = P(\xi = j - i + 1) = p_{j-i+1}$. Seja $\mu = \mathbb{E}\xi$.

Seja $\rho = F(0, 0)$, mostremos que $M_\xi(\rho) = \rho$.

Observe que para todo $j \in E = \{0, 1, 2, \dots\}$ vale que $P_{0,j} = P_{1,j}$. Logo $F(0, 0) = F(1, 0)$. Mais também temos que

$$F(0, 0) = P_{0,0} + \sum_{j=1}^{\infty} P_{0,j} F(j, 0) \implies \rho = p_0 + \sum_{j=1}^{\infty} p_j F(j, 0) \quad (9.16)$$

De outra parte o evento $\{T_{j-1} = n\}$ é equivalente ao evento:

$$\{n = \min \{k > 0 : j + (\xi_1 - 1) + (\xi_2 - 1) + \dots + (\xi_k - 1) = n - 1\}\}$$

isto é

$$\{n = \min \{k > 0 : \xi_1 + \xi_2 + \dots + \xi_k = k - 1\}\}$$

que não depende de j . Logo $P_j(T_{j-1} = n)$ e $F(j, j-1) = P_j(T_{j-1} < \infty)$ também não dependem de j . Portanto,

$$F(j, j-1) = F(j-1, j-2) = \dots = F(1, 0) = \rho$$

e aplicando a propriedade de Markov:

$$F(j, 0) = F(j, j-1)F(j-1, j-2) \cdots F(1, 0) = \rho^j$$

usando a última equação e (9.16) obtemos $\rho = p_0 + \sum_{j=1}^{\infty} p_j \rho^j = M_{\xi}(\rho)$. A partir de $\rho = M_{\xi}(\rho)$ é possível deduzir, como no caso do processo de ramificação, quais estados são recorrentes e quais são transitórios:

- Se $\mu \leq 1$ e a fila é irredutível então o 0 é recorrente, e portanto todos os estados são recorrentes.
- Se $\mu > 1$ e a fila é irredutível então todos os estados são transitórios.

9.4. Comportamento assintótico e distribuição estacionária. Caso geral. Precisaremos do seguinte resultado,

PROPOSIÇÃO 9.4.

$$F(i, j) = \begin{cases} 1, & j \text{ recorrente e } i \leftrightarrow j, \\ 0, & i, j \text{ recorrentes não comunicados,} \\ 0, & i \text{ recorrente e } j \text{ transitório,} \\ \frac{\mathbb{E}_i(N_i)}{\mathbb{E}_j(N_j)}, & i \text{ e } j \text{ transitórios,} \\ 1 - \frac{1}{\mathbb{E}_i(N_i)}, & i = j \text{ transitório,} \\ \rho_{e(j)}(i), & j \text{ recorrente e } i \text{ transitório.} \end{cases}$$

Aqui $\rho_{e(j)}(i)$ denota a probabilidade do estado i ser absorvido pela classe do j .

Um estado recorrente $i \in E$ de uma cadeia de Markov satisfaz $P_i(T_i < +\infty) = 1$. Quando além disto vale $\mathbb{E}_i(T_i) < +\infty$, i é chamado **recorrente positivo**. Se $\mathbb{E}_i(T_i) = +\infty$, então ele é chamado **recorrente nulo**. Os estados recorrentes nulos são visitados infinitas vezes pela cadeia, só que o tempo entre duas visitas consecutivas é muito grande. Nesse sentido eles são quase transitórios. De fato, estados recorrentes nulos e transitórios têm muitas propriedades em comum. Isto é consequência do seguinte resultado que caracteriza o comportamento assintótico de uma cadeia de Markov.

LEMA 9.5. *Seja j um estado recorrente aperiódico.*

(1) *Se j é positivo,*

$$P_{i,j}^n \rightarrow \frac{F(i, j)}{\mathbb{E}_j(T_j)}, \quad n \rightarrow \infty.$$

(2) *Se j é nulo,*

$$P_{i,j}^n \rightarrow 0, \quad n \rightarrow \infty.$$

Este resultado não será provado aqui. Ele tem várias consequências importantes.

PROPOSIÇÃO 9.6. *Se i é recorrente nulo e $i \leftrightarrow j$ então j também é recorrente nulo.*

DEMONSTRAÇÃO. □

PROPOSIÇÃO 9.7. *Uma classe irredutível finita não possui estados recorrentes nulos.*

DEMONSTRAÇÃO. Seja \mathcal{C} uma classe recorrente irreduzível e finita de uma cadeia de Markov. Pela proposição anterior, todos os seus estados são ou recorrentes nulos ou recorrentes positivos. Seja $P_{i,j}^\infty = \lim_{n \rightarrow \infty} P_{i,j}^n$. Como $1 = \sum_{j \in \mathcal{C}} P_{i,j}^n = \sum_{j \in E} P_{i,j}^\infty$, os termos $P_{i,j}^\infty$, $j \in \mathcal{C}$ não podem ser todos nulos. Pelo lema 9.5, os estados de \mathcal{C} devm ser todos recorrentes positivos. \square

Voltemos agora ao problema da existência de distribuição limite das matrizes de transição de ordem n .

PROPOSIÇÃO 9.8. *Uma cadeia de Markov aperiódica com uma única classe irreduzível recorrente positiva possui distribuição limite.*

DEMONSTRAÇÃO. Pelo lema 9.5 obtemos que

- (1) para j transitório, $P_{i,j}^\infty = 0$,
- (2) para j recorrente positivo, $P_{i,j}^\infty = \frac{1}{\mathbb{E}_j(T_j)}$. As filas são todas iguais e existe uma distribuição limite \square

PROPOSIÇÃO 9.9. *Uma cadeia de Markov aperiódica finita com uma única classe irreduzível possui distribuição limite.*

DEMONSTRAÇÃO. Numa cadeia de Markov aperiódica finita as classes irreduzíveis são recorrentes positivas. Aplicando o resultado anterior, o nosso aqui segue. \square

PROPOSIÇÃO 9.10. *Uma cadeia de Markov finita e tal que existe algum $n_0 \in \mathbb{N}$ para o qual*

$$\min_{i,j \in E} P_{i,j}^{n_0} > 0,$$

possui distribuição limite.

DEMONSTRAÇÃO. A cadeia é irreduzível porque todos os estados estão comunicados. Ela também é periódica pois $P_{i,i}^{n_0} > 0$ e $P_{i,i}^{n_0+1} > P_{i,j} P_{j,i}^{n_0} > 0$, para j tal que $P_{i,j} > 0$. Então $d(i) \leq m.d.c.\{n_0, n_0 + 1\} = 1$, portanto $d(i) = 1$ e a cadeia é aperiódica. \square

Veremos agora que quando a distribuição limite existe, ela deve satisfazer certa equação.

PROPOSIÇÃO 9.11. *Seja X_n uma cadeia de Markov com matriz de transição P e tal que os limites*

$$\lim_{n \rightarrow \infty} P_{i,j}^n = \pi(j) \tag{9.17}$$

existem para qualquer par de estados i e j e não dependem de i . Então para todo $j \in E$,

$$\sum_{i \in E} \pi(i) P_{i,j} = \pi(j). \tag{9.18}$$

Além disto, ou $\pi(i) = 0$, para todo $i \in E$ ou $\sum_{i \in E} \pi(i) = 1$. Neste último caso π é a única distribuição que satisfaz (9.18).

DEFINIÇÃO 9.12. Toda distribuição π que satisfaça (9.18) será chamada de **distribuição estacionária** da cadeia X_n .

NOTA 9.13. Repare que no caso que E é finito, a equação (9.18) pode ser reescrita da forma

$$\pi^t \cdot P = \pi^t \quad (9.19)$$

EXEMPLO 9.14. Mostramos que quando temos somente dois elementos no espaço de estados, para cadeias tais que $p+q < 2$ vale (9.18). Portanto existe uma única distribuição estacionária dada por

$$\pi_{est} = \left[\begin{array}{c} \frac{q}{p+q} \\ \frac{p}{p+q} \end{array} \right]. \quad (9.20)$$

Se $p = q = 1$, não existe distribuição limite quando $n \rightarrow \infty$. No entanto, observe que neste caso (9.19) fica da forma

$$[\pi(0) \ \pi(1)] = [\pi(0) \ \pi(1)] \cdot \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right] = [\pi(1) \ \pi(0)],$$

que tem como única solução $\pi(0) = \pi(1) = 1/2$ e vale também neste caso que a única distribuição estacionária é da forma (??).

Se $p = q = 0$, vemos que $P = Id(2)$ e para cada distribuição de probabilidade π vale $\pi^t \cdot P = \pi^t \cdot Id(2) = \pi^t$. Ou seja, qualquer distribuição é estacionária para esta cadeia.

Já temos exemplos de cadeias com somente uma e com infinitas distribuições estacionárias. A seguir apresentamos um exemplo de uma cadeia sem distribuições estacionárias.

EXEMPLO 9.15. *Passeio aleatório simples simétrico.* É possível provar que os estados do passeio aleatório simples simétrico são recorrentes nulos. Pela proposição 9.11, obtemos que esta cadeia não possui distribuição limite. Provaremos a seguir que ela também não possui distribuição estacionária. Para isso podemos raciocinar pelo absurdo.

Suponha que π é uma distribuição de probabilidades em \mathbb{Z} tal que (9.18) vale, onde P é a matriz de transição do passeio aleatório simples simétrico. Logo, para cada $i \in \mathbb{Z}$ temos

$$\pi(i) = \frac{\pi(i-1) + \pi(i+1)}{2}$$

ou

$$\pi(i) - \pi(i-1) = \pi(i+1) - \pi(i).$$

Ou seja, a quantidade $\pi(i+1) - \pi(i)$ é independente do estado i e podemos fazer $m = \pi(i+1) - \pi(i)$. É claro que $M > 0$ porque se todos os valores $\pi(i)$ fossem iguais, a soma deles não poderia ser um. Pegue agora um estado $i > 0$. Teremos

$$\begin{aligned} \pi(i) &= (\pi(i) - \pi(i-1)) + (\pi(i-1) - \pi(i-2)) + \cdots + (\pi(1) - \pi(0)) + \pi(0) \\ &= i \cdot m + \pi(0). \end{aligned}$$

No entanto, por ser π uma distribuição, deve valer que $\pi(i) \leq 1$ para todo $i \in \mathbb{Z}$ mas isto é impossível porque a expressão $i \cdot m + \pi(0)$ fica ilimitada para valores grandes do i .

Isto contradiz a nossa suposição de que π é uma distribuição estacionária da cadeia e portanto tais distribuições não existem.

Como veremos a seguir, uma distribuição π , estacionária para a cadeia X_n , é invariante para as transições da cadeia; por isto é chamada também de distribuição invariante da cadeia.

PROPOSIÇÃO 9.16. *Se a distribuição π é estacionária para a cadeia X_n então valem:*

- (1) *Para todo $j \in E$, $\sum_{i \in E} \pi(i)P_{i,j}^2 = \pi(j)$. Em geral, para todo n , vale que $\sum_{i \in E} \pi(i)P_{i,j}^n = \pi(j)$.*
- (2) *Se X_n tem distribuição inicial π (i.e. $P(X_0 = i) = \pi(i)$) então, para todo n , vale $P(X_n = i) = \pi(i)$.*
- (3) *Se $\lim_{n \rightarrow \infty} P_{i,j}^n = \pi(j)$ então $\lim_{n \rightarrow \infty} P(X_n = j) = \pi(j)$.*

DEMONSTRAÇÃO. (1) Considere o caso $n = 2$:

$$\begin{aligned} \sum_{i \in E} \pi(i)P_{i,j}^2 &= \sum_{i \in E} \pi(i) \left[\sum_{k \in E} P_{i,k}P_{k,j} \right] \\ &= \sum_{k \in E} P_{k,j} \left[\sum_{i \in E} \pi(i)P_{i,k} \right] \\ &= \sum_{k \in E} P_{k,j}\pi_k = \pi(j). \end{aligned}$$

Para obter o resultado para qualquer n podemos proceder por indução.

- (2) Se X_n tem distribuição inicial π ,

$$P(X_n = j) = \sum_{i \in E} \pi(i)P_{i,j}^n = \pi(j)$$

- (3) Se $P_{i,j}^n \rightarrow \pi(j)$ quando $n \rightarrow \infty$ então:

$$P(X_n = j) = \sum_{i \in E} P(X_0 = i)P_{i,j}^n \rightarrow \sum_{i \in E} P(X_0 = i)\pi(j) = \pi(j) \sum_{i \in E} P(X_0 = i) = \pi(j)$$

□

NOTA 9.17. *O segundo resultado acima fala que se uma cadeia de Markov começa com uma distribuição estacionária então todos os estados X_n terão esta distribuição como distribuição de probabilidade para qualquer instante n . Além disto é possível provar que tal cadeia é um processo estocástico estacionário no sentido estrito. Isto é consequência do resultado citado e da proposição 2.8.*

Para determinar a distribuição estacionária de uma cadeia de Markov com espaço de estados finito, basta encontrar um vetor π com entradas não negativas e tal que $\sum_{i \in E} \pi(i) = 1$ e $\pi^t \cdot P = \pi^t$. Vejamos isso no seguinte exemplo.

EXEMPLO 9.18. *Considere uma cadeia de Markov com espaço de estados $E = \{0, 1, 2\}$ e matriz de transição:*

$$P = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/6 & 1/3 & 1/2 \end{bmatrix}$$

A equação $\pi^t P = \pi^t$ é equivalente ao sistema:

$$\begin{cases} 1/3(\pi_0) + 1/4(\pi(1)) + 1/6(\pi(2)) = \pi_0 \\ 1/3(\pi_0) + 1/2(\pi(1)) + 1/3(\pi(2)) = \pi(1) \\ 1/3(\pi_0) + 1/4(\pi(1)) + 1/2(\pi(2)) = \pi(2) \end{cases}$$

mas este sistema é indeterminado pois a soma das três equações resulta na igualdade trivial

$$\pi(1) + \pi(2) + \pi(3) = \pi(1) + \pi(2) + \pi(3).$$

Tirando a primeira equação ficamos com o sistema

$$\begin{cases} 1/3(\pi_0) - 1/2(\pi(1)) + 1/3(\pi(2)) = 0 \\ 1/3(\pi_0) + 1/4(\pi(1)) - 1/2(\pi(2)) = 0. \end{cases}$$

Para poder encontrar a solução basta acrescentar ao sistema anterior a equação:

$$\pi_0 + \pi(1) + \pi(2) = 1.$$

Fazendo isto encontramos que a solução é

$$\pi_0 = 6/5, \quad \pi(1) = 2/5, \quad \pi(2) = 9/25.$$

EXEMPLO 9.19. *Voltemos ao exemplo da mobilidade social. A matriz de transição era*

$$\begin{array}{c} \begin{matrix} & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix} \end{array}$$

Para encontrar a sua distribuição estacionária devemos resolver o sistema

$$\begin{cases} 0,7\pi(1) + 0,3\pi(2) + 0,2\pi(3) = \pi(1) \\ 0,2\pi(1) + 0,5\pi(2) + 0,4\pi(3) = \pi(2) \\ \pi(1) + \pi(2) + \pi(3) = 1, \end{cases}$$

cujas soluções são $(11/23 \ 9/23 \ 3/23)$.

EXEMPLO 9.20. Vamos a achar a distribuição estacionária π do processo de vida e morte (exemplo 1.5).

$$\text{Se } \sum_{i \in E} \pi(i)P_{i,j} = \pi(j), \quad \text{então } \begin{cases} \pi(0)r_0 + \pi(1)q_1 = \pi(0), \\ \pi(i-1)p_{i-1} + \pi(i)r_i + \pi(i+1)q_{i+1} = \pi(i), \quad i \geq 1 \end{cases}$$

mas $p_i + r_i + q_i = 1$, logo:

$$\begin{cases} \pi(1)q_1 - \pi(0)p_0 = 0, \\ \pi(i+1)q_{i+1} - \pi(i)p_i = \pi(i)q_i - \pi(i-1)p_{i-1}, \quad i \geq 1, \end{cases}$$

então $\pi(i+1)q_{i+1} - \pi(i)p_i = 0$, $i \geq 1$ e portanto $\pi(i+1) = \frac{p_i}{q_{i+1}}\pi(i)$. Iterando este resultado obtemos:

$$\pi(i) = \frac{p_0 p_1 \cdots p_{i-1}}{q_1 q_2 \cdots q_i} \pi(0) \quad i \geq 1.$$

Observe que a distribuição estacionária depende de $\pi(0)$ e de

$$\Gamma_i = \begin{cases} 1, & i = 0, \\ \frac{p_0 p_1 \cdots p_{i-1}}{q_1 q_2 \cdots q_i}, & i \geq 1 \end{cases}$$

de modo que $\pi(i) = \Gamma_i \pi(0)$. Logo, somando a equação anterior, $1 = \sum_{i \in E} \pi(i) = \pi(0) \sum_{i \in E} \Gamma_i$ obtemos o valor de $\pi(0)$, desde que $\sum_{i \in E} \Gamma_i < \infty$. Em resumo:

- Se $\sum_{i=0}^{\infty} \Gamma_i < \infty$ então existe a distribuição estacionária π , que pode expressar-se como:

$$\pi(i) = \frac{\Gamma_i}{\sum_{k=0}^{\infty} \Gamma_k}.$$

- Se $\sum_{i=0}^{\infty} \Gamma_i = \infty$ então não existe a distribuição estacionária.

No caso finito, $d < \infty$, a distribuição estacionária sempre existe e vem dada por:

$$\pi(i) = \frac{\Gamma_i}{\sum_{k=0}^d \Gamma_k}.$$

EXEMPLO 9.21. **Cadeia de Ehrenfest** A cadeia de Ehrenfest corresponde a um processo de vida e morte, logo

$$\Gamma_0 = 1, \quad \Gamma_1 = \frac{1}{\frac{1}{3}} = 3, \quad \Gamma_2 = \frac{1 \frac{2}{3}}{\frac{1}{\frac{2}{3}}} = 3, \quad \Gamma_3 = \frac{1 \frac{2}{3} \frac{1}{3}}{\frac{1}{\frac{2}{3} \frac{1}{3}}} = 1$$

e portanto,

$$\pi(0) = \frac{1}{8}, \quad \pi(1) = \frac{3}{8}, \quad \pi(2) = \frac{3}{8}, \quad \pi(3) = \frac{1}{8}.$$

9.5. Teorema Ergodico. Considere o passeio aleatorio finito, com espaço de estados $E = \{1, 2, 3, \dots, d\}$. Isto é

$$P_{i,i+1} = p, P_{i,i-1} = q, p + q = 1, P_{0,1} = 1, P_{d,d-1} = 1$$

O passeio é uma cadeia de Markov periodica, de periodo 2. Isto significa que não é possível voltar ao mesmo estado num numero impar de transições, em particular $P_{i,i}^n = 0$ se n é impar e $P_{i,i}^n \neq 0$ se n é par. Logo não existe $\lim_{n \rightarrow \infty} P_{i,i}^n$.

Para resolver este problema observe que vale o seguinte resultado:

$$\lim_{n \rightarrow \infty} a_n = L \Rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n a_n = L$$

para toda sequencia de numeros a_n . Assim, se $\lim_{n \rightarrow \infty} P_{i,j}^n = \pi(j)$, então $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P_{i,j}^k = \pi(j)$.

Mais, estamos interessados no reciproco do anterior resultado, que é conhecido como teorema ergodico:

TEOREMA 9.22. *Cosidere uma cadeia de Markov irredutivel X_n , finita; e seja $f : E \mapsto R$ uma função limitada, então vale que:*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}f(X_k) = \sum_{i \in E} f(i)\pi(i)$$

onde π é uma distribuição no conjunto E . Em particular, se $f = \delta_j$ (i.e. $f(k) = 1$ se $k = j$ e $f(k) = 0$ se $k \neq j$), então:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P_{i,j}^k = \pi(j)$$

o que implica que existe a distribuição estacionária e que é igual a π .

Um outra maneira de expressar o teorema ergodico é decir que a media temporal da cadeia é igual a media espacial. Esta ideia é a base dos algoritmos usados nos metodos de MCMC (Monte Carlo Markov Chains) que se usam, por exemplo, na estatística Bayesiana. Onde, para obter amostras de uma certa distribuição construímos uma cadeia

de Markov que tenha como distribuição estacionária a distribuição que estamos procurando.

Usando o teorema ergodico é possível caracterizar a distribuição estacionária. Para isso observe que:

$$P_{i,j}^n = P_i(X_n = j) = \mathbb{E}_i(\delta_j(X_n))$$

De outra parte, seja $N_n(j) = \sum_{k=1}^n \delta_j(X_k)$, então

$$\mathbb{E}N_n(j) = \sum_{k=1}^n P_{i,j}^k$$

e se chamamos $R_n(i, j) = \sum_{k=1}^n P_{i,j}^k$ então

$$R(i, j) = \lim_{n \rightarrow \infty} \sum_{k=1}^n R_n(i, j) = \lim_{n \rightarrow \infty} \sum_{k=1}^n P_{i,j}^k$$

onde $R(i, j)$ é o termo (i, j) da matriz potencial R . Lembre que usamos ela para definir os estados transitórios e recorrentes. Vamos a ver a relação entre os estados transitórios e recorrentes e a distribuição estacionária.

- (1) Considere o estado $j \in E$ e suponha que ele é transitório. Logo $P_j(N_j < \infty) = 1$; onde $N_j = \lim_{n \rightarrow \infty} N_n(j)$. Portanto:

$$\lim_{n \rightarrow \infty} \frac{N_n(j)}{n} = 0 \quad \text{e} \quad \lim_{n \rightarrow \infty} \frac{R_n(i, j)}{n} = 0$$

- (2) Considere o estado $j \in E$ e suponha que ele é recorrente. Seja $m_j = \mathbb{E}_j(T_j)$, onde T_j é o tempo do primeiro retorno ao estado j . Decorre do teorema ergodico que:

- $\lim_{n \rightarrow \infty} \frac{N_n(j)}{n} = \frac{1}{m_j}$, se $T_j < \infty$.
- $\lim_{n \rightarrow \infty} \frac{R_n(i, j)}{n} = \frac{F(i, j)}{m_j}$.

Observe que se a cadeia é irredutível ou se $i, j \in \mathcal{C}$, com \mathcal{C} uma classe fechada de estados, então:

$$\lim_{n \rightarrow \infty} \frac{R_n(i, j)}{n} = \frac{1}{m_j}$$

Em particular,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P_{i,j}^k = \frac{1}{m_j}$$

Suponha que $m_j = \mathbb{E}_j(T_j) < \infty$, então:

$$\frac{1}{m_j} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P_{i,j}^k = \lim_{n \rightarrow \infty} P_{i,j}^n = \pi(j)$$

Logo se j é recorrente e $m_j < \infty$ então $\pi(j) = \frac{1}{m_j}$. Como vimos antes, neste caso j é recorrente positivo. Se $m_j = \infty$ então j é recorrente negativo e $\pi(j) = 0$.

Concluimos também que:

TEOREMA 9.23. *Uma cadeia de Markov finita e irredutível tem todos seus estados recorrentes positivos.*

EXEMPLO 9.24. Fila Considere a fila $X_n \in E = \{0, 1, 2, \dots\}$, onde o número de clientes que chegam no instante n é ξ_n . As v.a. ξ_n são i.i.d. com distribuição $P(\xi_n = i) = p_i$. Um cliente é atendido em cada instante. Logo $X_{n+1} = (X_n - 1)^+ + \xi_{n+1}$. Suponha que $p_0 > 0$ e que $p_0 + p_1 < 1$, neste caso a cadeia é irredutível e tem matriz de transição:

$$P = \begin{bmatrix} p_0 & p_1 & p_2 & p_3 & \dots \\ p_0 & p_1 & p_2 & p_3 & \dots \\ 0 & p_0 & p_1 & p_2 & \dots \\ 0 & 0 & p_0 & p_1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

Observe que se $p_0 = 0$ então o estado 0 é transitório. Também, se $p_0 + p_1 = 1$ então $p_2 = p_3 = \dots = 0$, neste caso os estados 0, 1 são recorrentes e os demais estados são transitórios. Em geral, suponha que $p_0 + p_1 < 1$, neste caso a cadeia é irredutível. Seja $\mu = \mathbb{E}\xi_i$, vimos antes que:

- Se $\mu \leq 1$ então X_n é recorrente.
- Se $\mu > 1$ então X_n é transitório.

É possível provar que se $\mu \leq 1$ então $\mathbb{E}(T_0) = \frac{1}{1-\mu}$. Podemos concluir então que:

- (1) Se $\mu > 1$ então 0 é transitório logo X_n é transitório.
- (2) Se $\mu = 1$ então 0 é recorrente nulo.
- (3) Se $\mu < 1$ então 0 é recorrente positivo, e $\pi(0) = 1 - \mu$.

9.6. Teorema de Perron-Frobenius. Vimos antes que se uma cadeia de Markov irredutível (e finita) com matriz de transição $P = (P_{i,j})$ então:

$$\lim_{n \rightarrow \infty} P_{i,j}^n = \pi(j)$$

onde π é a distribuição estacionária. Logo se n for grande $P_{i,j}^n \sim \pi(j)$ Então, podemos fazermos a pergunta de qual é o erro nesta aproximação. O seguinte teorema (Perron-Frobenius) nos permite avaliar este erro.

TEOREMA 9.25. *Seja $P = (P_{i,j})_{i,j \in E}$ a matriz de transição de uma cadeia de Markov finita e irredutível, então vale:*

- (1) *Os autovalores $\lambda_1, \lambda_2, \dots$ da matriz P pertencem ao círculo unitário no conjunto \mathbb{C} , $|\lambda_i| \leq 1$.*
- (2) *Se ordenamos os autovalores: $|\lambda_1| \geq |\lambda_2| \geq \dots$, então $\lambda_1 = 1$ e o autovalor correspondente é igual a π .*
- (3) *Se λ_2 é o segundo maior autovalor, então:*

$$\max_{i \in E} \max_{j \in E} |P_{i,j}^n - \pi(j)| \leq c |\lambda_2|^n$$

Então, o erro na aproximação é dado pelo segundo autovalor.

EXERCÍCIO 9.26. *Considere as seguintes matrizes de transição:*

$$P = \begin{bmatrix} 0,4 & 0,6 \\ 0,6 & 0,4 \end{bmatrix} \quad \tilde{P} = \begin{bmatrix} 0,9 & 0,1 \\ 0,1 & 0,9 \end{bmatrix}$$

A matriz P vai mais rapidamente ao equilíbrio do que a matriz \tilde{P} , porque? Verificar isso:

- (1) *Ache os autovalores das matrizes e identifique o segundo autovalor em cada caso. Qual é maior?*
- (2) *Ache o autovetor (esquerdo) correspondente ao autovalor 1. Em outras palavras, procure π e $\tilde{\pi}$ tal que $\pi^T P = \pi$ (e $\tilde{\pi}^T \tilde{P} = \tilde{\pi}$)*
- (3) *Ache P^n , $n = 2, 3, 4, 5$. Avalie*

$$\max_{i \in E} \max_{j \in E} |P_{i,j}^n - \pi(j)|.$$

e compare com o valor de $|\lambda_2|^n$.

- (4) *repita o item anterior com \tilde{P} e $\tilde{\pi}$.*

CAPÍTULO 3

O Processo de Poisson

1. Introdução.

Apresentaremos aqui o processo de Poisson. Este processo estocástico deve o seu nome ao matemático francês Simon-Denis Poisson (1781 - 1840). Veremos que ele pode ser definido em termos das ocorrências de eventos, em outras palavras, se denotarmos este processo como $\{N_t\}_{t \geq 0}$ e fixarmos o tempo no instante t , teremos que N_t é um número inteiro que representa **o número de chegadas ou eventos até o instante t** . O processo de Poisson é portanto, um processo estocástico a tempo contínuo, isto é $T = [0, \infty)$, e com espaço de estados $E = \mathbb{N} = \{0, 1, 2, \dots\}$.

EXEMPLO 1.1. *Suponha que $N_t = 5$ e suponha que não chegam dois "eventos" no mesmo instante, uma realização do processo poderia ser*

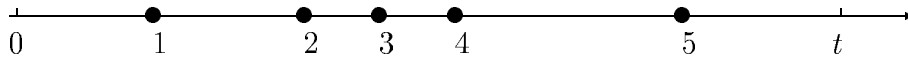


Figura 26

que pode ser também representada como uma trajetória

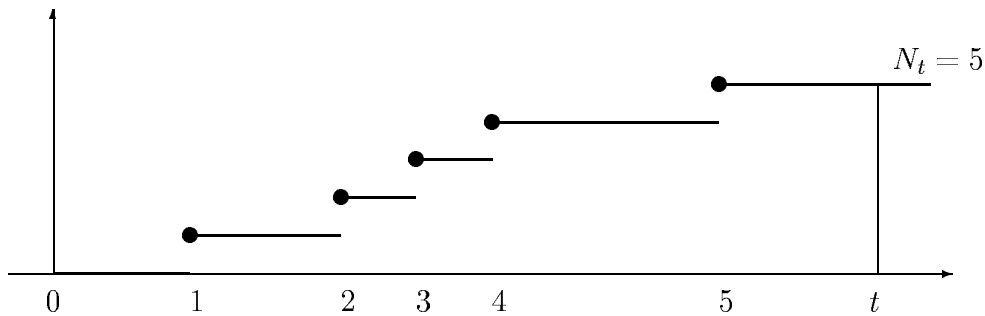


Figura 27

Como é ilustrado na figura acima, cada trajetória do processo é uma função escada. O número de eventos no intervalo $(t, t + s]$, $s \geq 0$ será $N_{t+s} - N_t$. Veremos que ele é

independente do número de eventos até o instante t , $\{N_u, u \leq t\}$. Em outras palavras, o processo tem incrementos independentes (ver ..).

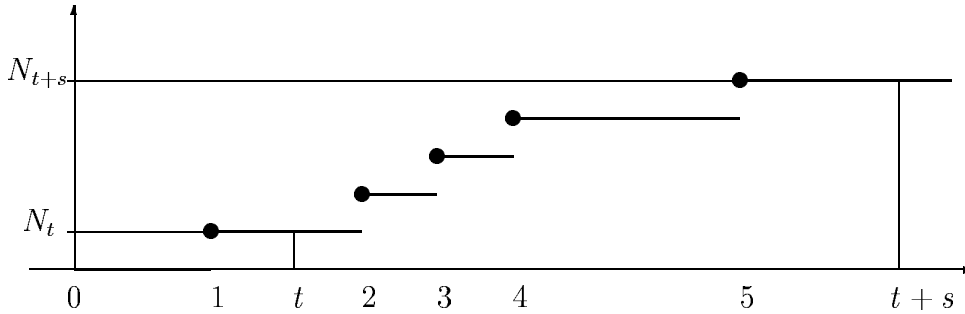


Figura 28

2. O processo de Poisson.

DEFINIÇÃO 2.1. Um processo a tempo contínuo $\{N_t\}_{t \geq 0}$ definido sobre um espaço amostral Ω , com espaço de estados $E = \mathbb{N}$ e tal que para todo evento elementar $\omega \in \Omega$, a trajetória correspondente, $t \rightarrow N_t(\omega)$, verifique

- (1) é não decrescente;
- (2) cresce somente com saltos. (i.e. é constante entre os saltos).
- (3) é contínua a direita e tem limite à esquerda.
- (4) $N_0(\omega) = 0$;

é chamado **processo de contagem** ou **processo das chegadas**.

Sejam T_1, T_2, \dots os tempos das chegadas (ou os tempos dos saltos ou dos instantes nos quais ocorrem os eventos). Estas variáveis definem um processo a tempo discreto. Uma trajetória típica deste processo é

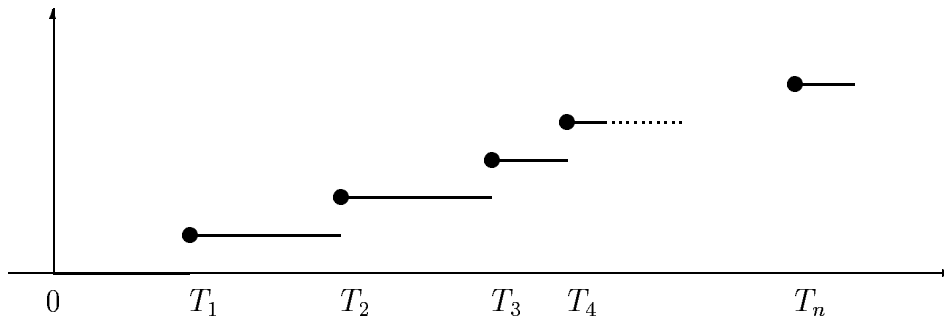


Figura 29

DEFINIÇÃO 2.2. O processo de contagem $\{N_t\}_{t \geq 0}$ é chamado de **processo de Poisson homogêneo** se

- (1) os saltos têm comprimento um;
- (2) $N_{t+s} - N_t$ é independente de $\{N_u, u \leq t\}$, para todo $t, s > 0$;
- (3) a distribuição de $N_{t+s} - N_t$ é independente de t .

Observe que se $N_{t+s} - N_t$ é independente de $\{N_u, u \leq t\}$, então $N_{t+s} - N_t$ será independente de $N_{t_1}, N_{t_2}, \dots, N_{t_n}$ para $t_1 < t_2 < \dots < t_n = t$ e portanto este incremento será independente do vetor dos incrementos $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$. Pelo mesmo argumento obtemos que $N_{t_n} - N_{t_{n-1}}$ deve ser independente do vetor $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_{n-1}} - N_{t_{n-2}}$ e procedendo recursivamente obteremos que incrementos correspondentes a intervalos disjuntos devem ser independentes, ou seja, o processo tem incrementos independentes. A propriedade (3) expressa que os incrementos do processo são estacionários.

O processo de Poisson fica definido então como um processo de contagem com saltos de valor um e incrementos estacionários e independentes.

Alguns exemplos de situações que podem ser modeladas usando o processo de Poisson aparecem a seguir.

- (1) O número de ligações que chegam numa central telefônica durante um intervalo de tempo determinado define um processo de Poisson, se supormos que o número de chamadas recebidas durante intervalos disjuntos são independentes, dependem somente do comprimento do intervalo e se pudermos assumir que há um número médio constante de chegadas por unidade de tempo. Em problemas mais realistas, este número médio depende do tempo, sendo mais apropriado o modelo de Poisson não homogêneo.
- (2) O número de fótons que chega num detector de fótons ao longo do tempo.
- (3) Os astrônomos podem considerar o número de estrelas num volume determinado no espaço como uma variável aleatória de Poisson e o número de estrelas em regiões disjuntas podem ser consideradas independentes. Com estas suposições, o número de estrelas observadas num volume V é um processo de Poisson tridimensional sobre o espaço definido pelo volume V (veja o exercício ..).

A definição que temos apresentado do processo de Poisson não coloca nenhuma restrição explícita sobre a distribuição das variáveis N_t . No entanto, estas distribuições estão determinadas pela definição dada. A seguir calcularemos a distribuição de N_t . Isto será consequência de uma série de proposições.

PROPOSIÇÃO 2.3. *Existe uma constante $\lambda \geq 0$ tal que para todo $t > 0$,*

$$P(N_t = 0) = e^{-\lambda t}.$$

DEMONSTRAÇÃO. Observe que o evento $\{N_{t+s} = 0\} = \{N_t = 0, N_{t+s} - N_t = 0\}$. Logo, usando a independência e a estacionariedade dos incrementos podemos obter

$$\begin{aligned} P(N_{t+s} = 0) &= P(N_t = 0)P(N_{t+s} - N_t = 0) \\ &= P(N_t = 0)P(N_s = 0). \end{aligned}$$

Considere a função $f(u) = P(N_u = 0)$. Usando equação anterior vemos que f satisfaz a equação de Cauchy

$$f(s+t) = f(s)f(t) \quad (2.21)$$

que tem como solução $f(t) = e^{-\lambda t}$ para algum $\lambda \in \mathbb{R}$ ou $f(t) = 0$, para todo $t \geq 0$.

Vejam porquê f não pode ser identicamente nula. Para isto podemos raciocinar pelo absurdo. Vamos supor então que para todo $t \geq 0$, $f(t) = 0$, ou seja, para um instante t qualquer teremos que $P(N_t \geq 1) = 1$. Podemos dividir o intervalo $[0, t]$ em n subintervalos de comprimento $1/n$. Seja $t_k = k \frac{t}{n}$, $k \in \{1, \dots, n\}$. Então

$$N_t = (N_t - N_{t_{n-1}}) + \dots + (N_{t_1}),$$

e sabemos que $P(N_t - N_{t_{n-1}} \geq 1) = \dots = P(N_{t_1} \geq 1) = 1$. Como tem n termos na soma acima, obtemos que $P(N_t \geq n) = 1$ e como isto foi feito para $n \in \mathbb{N}$ quaisquer, chegaríamos a que $P(N_t = +\infty) = 1$, mas isto contradiz o fato que $E = \mathbb{N}$.

Temos então que $f(t) = e^{-\lambda t}$ para algum $\lambda \in \mathbb{R}$. Para ver que $\lambda \geq 0$ basta observar que pelo fato do processo de Poisson ser não decrescente, vale que se $t, s \geq 0$, $\{N_{t+s} = 0\} \subset \{N_t = 0\}$. Ou seja, f é uma função decrescente e devemos ter $\lambda \geq 0$. O caso $\lambda = 0$ corresponde ao caso degenerado em que não existem chegadas, i.e, $N_t = 0$ para todo $t \geq 0$ com probabilidade 1. \square

PROPOSIÇÃO 2.4.

$$\lim_{t \rightarrow 0} \frac{1}{t} P(N_t \geq 2) = 0$$

Vamos a usar o fato que para todo $t \geq 0$, $\mathbb{E}N_t < \infty$ (não provaremos isto que decorre do fato do processo não ser **explosivo**, i.e. não acontecem dois o mas eventos no mesmo instante.)

DEMONSTRAÇÃO. Seja $h(t) = P(N_t \geq 2)$, e observe que $\{N_t \geq 2\} \subseteq \{N_{t+s} \geq 2\}$ implica $h(t) \leq h(t+s)$. Logo h é não decrescente. Defina

$$n_t = \max \left\{ n : n < \frac{1}{t} \right\}$$

então $t < \frac{1}{n_t}$ e $\frac{1}{t} < n_t + 1$. Usando que h é não decrescente temos que $h(t) \leq h(\frac{1}{n_t})$. Portanto:

$$0 \leq \frac{1}{t} h(t) \leq (n_t + 1) h\left(\frac{1}{n_t}\right) \leq \left(\frac{n_t + 1}{n_t}\right) \left(n_t h\left(\frac{1}{n_t}\right)\right)$$

Se $t \rightarrow 0$ então $n_t \rightarrow \infty$ e $\frac{n_t+1}{n_t} \rightarrow 1$. Logo basta mostrar que $nh\left(\frac{1}{n}\right) \rightarrow 0$. Para isso dividimos o intervalo $[0, 1]$ em subintervalos de comprimento $\frac{1}{n}$.

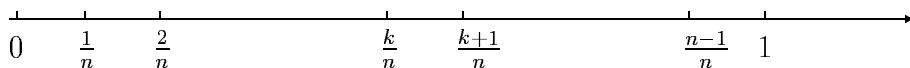


Figura 30

Seja S_n o número dos subintervalos onde temos mais de dois eventos, então S_n é o número de sucessos num processo de Bernoulli com probabilidade de sucesso $p = h\left(\frac{1}{n}\right)$. Logo

$$\mathbb{E}S_n = np = nh\left(\frac{1}{n}\right)$$

Considere uma realização ω do processo N_t , para n grande os subintervalos são suficientemente pequenos como para conter no máximo um evento. Logo $\lim_{n \rightarrow \infty} S_n(\omega) = 0$.

Agora usamos o fato de $\mathbb{E}N_t < \infty$, que implica $\mathbb{E}N_1 < \infty$, e o fato de $S_n \leq N_1$ para concluir que

$$\lim_{n \rightarrow \infty} nh\left(\frac{1}{n}\right) = \lim_{n \rightarrow \infty} \mathbb{E}S_n = 0.$$

□

PROPOSIÇÃO 2.5.

$$\lim_{t \rightarrow 0} \frac{1}{t} P(N_t = 1) = \lambda$$

DEMONSTRAÇÃO. Observe que $P(N_t = 1) = 1 - P(N_t = 0) - P(N_t \geq 2)$ implica:

$$\frac{1}{t} P(N_t = 1) = \frac{1 - e^{-\lambda t}}{t} - \frac{1}{t} P(N_t \geq 2)$$

e portanto $\lim_{t \rightarrow 0} \frac{1}{t} P(N_t = 1) = \lambda$.

□

Usando as proposições anteriores podemos mostrar que:

TEOREMA 2.6. *Seja $\{N_t\}_{t \geq 0}$ o processo de Poisson, então existe $\lambda \geq 0$ tal que*

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

*i.e. $N_t \sim \text{Poisson}(\lambda t)$. λ é chamada de **taxa** do processo.*

DEMONSTRAÇÃO. Seja $G(t) = \mathbb{E}[\alpha^{N_t}]$, $0 < \alpha < 1$, a função geradora de probabilidades de N_t . Então

$$G(t) = \sum_{k=0}^{\infty} \alpha^k P(N_t = k).$$

Basta mostrar que $G(t) = e^{-\lambda t(1-\alpha)}$, pois usando a expansão em série de Taylor da exponencial obteríamos

$$e^{-\lambda t(1-\alpha)} = e^{-\lambda t} \sum_{k=0}^{\infty} \alpha^k \frac{(\lambda t)^k}{k!}$$

e comparando os coeficientes correspondentes a potências iguais de α chegamos ao resultado desejado.

Observe que $N_{t+s} = N_t + (N_{t+s} - N_t)$ e a independência dos incrementos implicam que

$$\begin{aligned} G(t+s) &= \mathbb{E}[\alpha^{N_{t+s}}] = \mathbb{E}[\alpha^{N_t}] \mathbb{E}[\alpha^{N_{t+s}-N_t}] \\ &= G(t)G(s). \end{aligned} \tag{2.22}$$

Como $G(t) \geq P(N_t = 0) = e^{-\lambda t} > 0$, a equação de Cauchy (2.21) tem uma única solução da forma $G(t) = e^{tg(\alpha)}$, $t \geq 0$, com $g(\alpha)$ certa função que satisfaz $g(\alpha) = G'(0)$. Observe que $G(0) = 1$. Logo,

$$\begin{aligned} g(\alpha) &= \lim_{t \rightarrow 0} \frac{1}{t} [G(t) - 1] \\ &= \lim_{t \rightarrow 0} \frac{1}{t} [P(N_t = 0) - 1] + \lim_{t \rightarrow 0} \frac{1}{t} [\alpha P(N_t = 1)] + \lim_{t \rightarrow 0} \frac{1}{t} \left[\sum_{n=2}^{\infty} \alpha^n P(N_t = n) \right]. \end{aligned}$$

Usando as proposições anteriores obtemos,

- $\lim_{t \rightarrow 0} \frac{1}{t} [P(N_t = 0) - 1] = -\lambda$
- $\lim_{t \rightarrow 0} \frac{1}{t} [\alpha P(N_t = 1)] = \alpha\lambda$.

Como $0 < \alpha < 1$ verifica-se que $\lim_{t \rightarrow 0} \frac{1}{t} [\sum_{n=2}^{\infty} \alpha^n P(N_t = n)] = 0$, pois

$$\begin{aligned} 0 &\leq \lim_{t \rightarrow 0} \frac{1}{t} \left[\sum_{n=2}^{\infty} \alpha^n P(N_t = n) \right] \leq \lim_{t \rightarrow 0} \frac{1}{t} \left[\sum_{n=2}^{\infty} P(N_t = n) \right] \\ &= \lim_{t \rightarrow 0} \frac{1}{t} P(N_t \geq 2) = 0 \end{aligned}$$

Portanto $g(\alpha) = -\lambda + \lambda\alpha$, logo $G(t) = e^{-\lambda t + \lambda\alpha t}$ e o resultado fica provado. \square

Observe que $N_t \sim \text{Poisson}(\lambda t)$ implica que $\mathbb{E}(N_t) = \lambda t$ e $\text{Var}(N_t) = \lambda t$.

COROLÁRIO 2.7. *Seja $\{N_t\}_{t \geq 0}$ o processo de Poisson com taxa λ , então:*

$$P(N_{t+s} - N_t = k | N_u, u \leq t) = P(N_{t+s} - N_t = k) = \frac{(\lambda s)^k}{k!} e^{-\lambda s}, \quad k = 0, 1, \dots$$

EXEMPLO 2.8. *Seja $\{N_t\}_{t \geq 0}$ o processo de Poisson com taxa $\lambda = 8$. Achar $P(N_{2,5} = 17, N_{3,7} = 22, N_{4,3} = 36)$.*

Solução:

$$\begin{aligned}
& P(N_{2,5} = 17, N_{3,7} = 22, N_{4,3} = 36) \\
&= P(N_{2,5} = 17, N_{3,7} - N_{2,5} = 5, N_{4,3} - N_{3,7} = 14) \\
&= P(N_{2,5} = 17)P(N_{3,7} - N_{2,5} = 5)P(N_{4,3} - N_{3,7} = 14) \\
&= \frac{(8(2, 5))^{17}}{17!} e^{-8(2,5)} \frac{(8(3, 7 - 2, 5))^5}{5!} e^{-8(3,7-2,5)} \frac{(8(4, 3 - 3, 7))^{14}}{14!} e^{-8(4,3-3,7)}
\end{aligned}$$

O resultado apresentado a seguir dá uma outra caracterização dos processos de Poisson.

TEOREMA 2.9. $\{N_t\}_{t \geq 0}$ é o processo de Poisson com taxa λ se e somente se.

- (1) com probabilidade um os saltos das trajetórias $t \mapsto N_t(\omega)$ têm valor um;
- (2) Para todo $t, s > 0$ vale que $\mathbb{E}[N_{t+s} - N_t | N_u, u \leq t] = \lambda s$.

Considere o conjunto $(t, t + s] \subseteq \mathbb{R}^+$, $s > 0, t > 0$. O número de eventos que ocorrem no intervalo $(t, t + s]$ está definido como

$$N_{(t,t+s]} = N_{t+s} - N_t.$$

Pelo fato dos incrementos do processo de Poisson ser independentes e identicamente distribuídos,

$$N_{(t,t+s]} \sim N_s \sim \text{Poisson}(\lambda s),$$

mas $\lambda s = \lambda \text{comp}((t, t + s])$, onde $\text{comp}((t, t + s])$ é o comprimento do intervalo $(t, t + s]$. Isto é, se $B \subseteq \mathbb{R}^+$ é um intervalo então

$$\text{comp}(B) = \int_B dx.$$

Podemos estender esta definição a um conjunto $B \subseteq \mathbb{R}^+$ e considerar o número de eventos N_B que acontecem neste conjunto.

Qual será a distribuição de N_B ?. Quando $B = (t, t + s]$, sabemos que:

$$N_B \sim \text{Poisson}(\lambda \text{comp}(B)).$$

Vejamos que isto vale também quando B for uma união finita de intervalos. Suponha que A e B são dois intervalos disjuntos com comprimentos a e b respectivamente, por exemplo, $A = (t, t + a]$ e $B = (s, s + b]$ com $t + a \leq s$. Então N_A e N_B são variáveis aleatórias independentes com distribuição de Poisson com parâmetros λa e λb , respectivamente. Seja $C = (t + a, t + a + b]$. Pela estacionariedade dos incrementos, N_B e N_C seguem a mesma distribuição e N_A e N_C também são independentes. Portanto, $N_A + N_B \sim N_A + N_C$, mas esta última variável conta a quantidade de chegadas no intervalo $(t, t + a + b]$, logo $N_A + N_C \sim \text{Poisson}(\lambda(a + b))$. Obtemos então que $N_A + N_B \sim \text{Poisson}(\lambda(a + b))$. O mesmo argumento funciona para qualquer número de intervalos e prova a parte da necessidade no resultado a seguir.

TEOREMA 2.10. *O processo N_t é de Poisson, com taxa λ , se e somente se para todo $B \subseteq \mathbb{R}^+$ que seja união finita de intervalos disjuntos vale*

$$P(N_B = k) = \frac{(\lambda b)^k}{k!} e^{-\lambda b}, \quad \text{onde } b = \text{comp}(B).$$

Suponha agora que conhecemos que em instantes no conjunto $B = (t, t + b]$ ocorreram k eventos. Gostariamos de ter informação sobre os instantes em que eles ocorreram.

TEOREMA 2.11. *Sejam A_1, \dots, A_n intervalos disjuntos dois a dois e tais que $B = A_1 \cup A_2 \cdots \cup A_n$ e $\text{comp}(A_i) = a_i$. (i.e. $b = \text{comp}(B) = \sum_{i=1}^n a_i$). Sejam $k_1, k_2, \dots, k_n, k \in \mathbb{N}$ tais que $k_1 + k_2 + \cdots + k_n = k$. Então vale*

$$P(N_{A_1} = k_1, N_{A_2} = k_2, \dots, N_{A_n} = k_n) = \frac{k!}{k_1! k_2! \cdots k_n!} \left(\frac{a_1}{b}\right)^{k_1} \cdots \left(\frac{a_n}{b}\right)^{k_n}.$$

Logo $(N_{A_1}, N_{A_2}, \dots, N_{A_n}) | N_B = k \sim \text{Multinomial}(k, (\frac{a_1}{b}), \dots, (\frac{a_n}{b}))$

DEMONSTRAÇÃO. Observe que $\{N_{A_1} = k_1, N_{A_2} = k_2, \dots, N_{A_n} = k_n\} \subseteq \{N_B = k\}$, logo:

$$\begin{aligned} P(N_{A_1} = k_1, N_{A_2} = k_2, \dots, N_{A_n} = k_n) &= \frac{P(N_{A_1} = k_1, N_{A_2} = k_2, \dots, N_{A_n} = k_n)}{P(N_B = k)} \\ &= \frac{(\lambda a_1)^{k_1}}{k_1!} e^{-\lambda a_1} \cdots \frac{(\lambda a_n)^{k_n}}{k_n!} e^{-\lambda a_n} \frac{k!}{(\lambda b)^k} e^{\lambda b} \\ &= \frac{k!}{k_1! k_2! \cdots k_n!} \left(\frac{a_1}{b}\right)^{k_1} \cdots \left(\frac{a_n}{b}\right)^{k_n}. \end{aligned}$$

□

Suponha agora que sabemos que chegou um só evento no intervalo B , achemos a probabilidade de que o evento chegue no intervalo $A \subseteq B$. Suponha que $b = \text{comp}(B) > \text{comp}(A) = a$.

Logo:

$$\begin{aligned} P(N_A = 1 | N_B = 1) &= P(N_A = 1, N_{B \setminus A} = 0 | N_B = 1) \\ &= \frac{1!}{1!0!} \left(\frac{a}{b}\right)^1 \left(\frac{b-a}{a}\right)^0 = \frac{a}{b} \end{aligned}$$

Estudemos agora o comportamento assintótico do processo de Poisson $\{N_t\}_{t \geq 0}$. Para tanto observe que $\mathbb{E}N_t = \lambda t$. Portanto, se $t = n$, $\mathbb{E}N_n = \lambda n$ e podemos escrever

$$N_n = N_1 + (N_2 - N_1) + (N_3 - N_2) + \cdots + (N_n - N_{n-1}),$$

com $\{N_{i+1} - N_i\}_{i=1, n}$ i.i.d. e $\mathbb{E}(N_{i+1} - N_i) = \lambda$. Pela lei dos grandes números aplicada à sequência $\{N_{i+1} - N_i\}_{i=1, n}$ vale que quando $n \rightarrow \infty$:

$$\frac{N_n}{n} \longrightarrow \lambda.$$

Podemos generalizar este resultado:

TEOREMA 2.12. *Lei dos grandes Números do processo de Poisson.*

$$\frac{N_t}{t} \longrightarrow \lambda$$

se $t \rightarrow \infty$.

Logo, um estimador consistente da taxa do processo λ é:

$$\hat{\lambda} = \frac{\text{Número de eventos no intervalo } [0, T]}{T}$$

e um intervalo de confiança assintótico para este estimador pode ser obtido do seguinte resultado.

TEOREMA 2.13. *Teorema Central do Limite do processo de Poisson.*

$$\frac{N_t - \lambda t}{\sqrt{\lambda t}} \xrightarrow{d} N(0, 1)$$

se $t \rightarrow \infty$.

Observe que se t for grande N_t se comporta aproximadamente como uma normal de média λt e variância λt .

3. Tempos de Chegada

Vamos considerar agora os tempos de chegada do processo de Poisson. Eles são os tempos nos quais acontecem os eventos do processo. O n -ésimo tempo de chegada está definido por

$$T_n = \min \{t : N_t = n\}.$$

Observe que $N_{T_n} = n$. A distribuição de T_n pode ser obtida a partir da distribuição do processo N_t a partir da igualdade dos seguintes eventos

$$\{T_n \leq t\} = \{N_t \geq n\}.$$

Assim,

$$\begin{aligned} P(T_n \leq t) &= P(N_t \geq n) \\ &= \sum_{k=n}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \end{aligned}$$

A avaliação desta expressão é complicada. No lugar de fazer isto vamos mostrar por outra via que $T_n \sim \text{Gamma}(n, \lambda)$ e para isso vamos a estudar algumas propriedades do **processo dos tempos das chegadas**, $\{T_n\}_{n \geq 1}$.

Uma observação importante é que conhecer o processo até o instante T_n , i.e. $\{N_t : t \leq T_n\}$ é o mesmo que conhecer o processo dos tempos das chegadas até o instante n , i.e. $\{T_1, T_2, \dots, T_n\}$, isto é fácil de visualizar na seguinte figura.

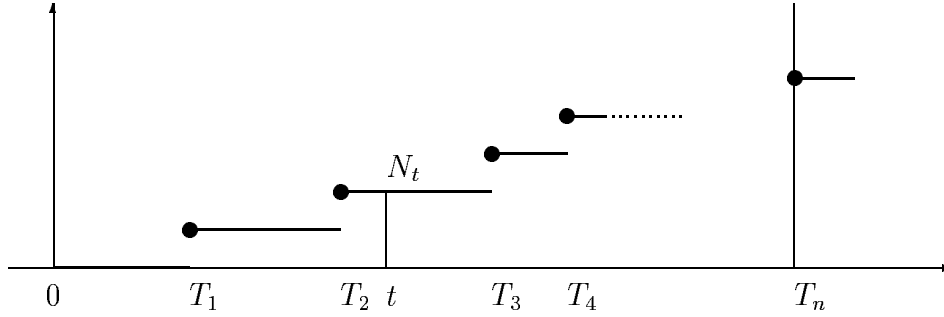


Figura 31

Sabemos que para $t, s \geq 0$ vale

$$P(N_{(t,t+s]} = 0 | N_u, u \leq t) = P(N_{(t,t+s]} = 0) = e^{-\lambda s}.$$

Esta propriedade vale também para os tempos T_n , ou seja,

$$P(N_{(T_n, T_n+s]} = 0 | N_u, u \leq T_n) = P(N_{(T_n, T_n+s]} = 0) = e^{-\lambda s}.$$

Então,

$$\begin{aligned} P(N_{T_n+s} - N_{T_n} = 0 | T_1, T_2, \dots, T_n) &= P(N_{T_n+s} - N_{T_n} = 0 | N_u, u \leq T_n) \\ &= P(N_{(T_n, T_n+s]} = 0 | N_u, u \leq T_n) \\ &= e^{-\lambda s} \end{aligned}$$

mais observe que

$$\{N_{T_n+s} - N_{T_n} = 0\} = \{T_{n+1} - T_n > s\}$$

e obtemos portanto,

TEOREMA 3.1. Para todo $t > 0$ e $n \geq 1$ vale que:

$$P(T_{n+1} - T_n \leq t | T_1, T_2, \dots, T_n) = 1 - e^{-\lambda t}$$

Assim o processo $\{T_n : n \geq 1\}$ é estacionário e tem incrementos independentes.

COROLÁRIO 3.2. As variáveis aleatórias $T_1, T_2 - T_1, T_3 - T_2, \dots$ são *i.i.d.* e $T_{n+1} - T_n \sim \exp(\lambda)$.

Logo os tempos entre dois eventos consecutivos no processo de Poisson tem distribuição exponencial. Lembremos que a distribuição exponencial tem a propriedade da perda da memória, i.e. se $X \sim \exp(\lambda)$ e $t, s > 0$, então

$$P(X > s + t | X > s) = P(X > t).$$

Assim, podemos concluir que o processo de Poisson *perde a memória*. Em geral vale o seguinte resultado.

TEOREMA 3.3. *Sejam T_1, T_2, T_3, \dots os tempos de chegada num processo de contagem N_t com $t \geq 0$.*

$$\{N_t\}_{t \geq 0} \text{ é um processo de Poisson}(\lambda) \Leftrightarrow \begin{cases} T_{n+1} - T_n, n \geq 0, \text{ i.i.d.}, \\ T_{n+1} - T_n \sim \text{exp}(\lambda). \end{cases}$$

Observe que $T_n = T_1 + (T_2 - T_1) + (T_3 - T_2) + \dots + (T_n - T_{n-1})$. Usando o fato que a soma de distribuições exponenciais i.i.d. tem distribuição Gamma podemos concluir que o tempo do n -ésimo evento $T_n \sim \text{Gamma}(n, \lambda)$. Logo,

$$\begin{aligned} \mathbb{E}T_n &= \frac{n}{\lambda}, \\ \text{Var}(T_n) &= \frac{n}{\lambda^2}, \\ \mathbb{E}e^{-\alpha T_n} &= \left[\frac{\lambda}{\lambda + \alpha} \right]^n. \end{aligned}$$

A distribuição $\text{Gamma}(n, \lambda)$ é chamada de distribuição de Erlang(n).

EXEMPLO 3.4. *Os tempos de falha de um chip de um computador tem distribuição exponencial com taxa λ . Cada vez que falha um chip ele é imediatamente substituído.*

Sejam X_1, X_2, \dots os tempos de duração de cada chip que foi trocado. Logo, $P(X_n < t) = 1 - e^{-\lambda t}$. Considere T_1, T_2, \dots os sucessivos instantes nos quais aconteceu uma falha no computador devido a uma falha do chip,

$$\begin{aligned} T_1 &= X_1 \\ T_2 &= X_1 + X_2 \\ &\dots \\ T_n &= X_1 + X_2 + \dots + X_n. \end{aligned}$$

Por exemplo $T_3 = X_1 + X_2 + X_3$ é o instante da falha do terceiro chip.

Suponha que $\lambda = 0,0002$ (em horas⁻¹), então a esperança de vida de um chip é

$$\mathbb{E}X_n = \frac{1}{\lambda} = 5000 \text{ horas}$$

e a variância é

$$\text{Var}X_n = \frac{1}{\lambda^2} = 25 \times 10^6.$$

Se N_t é o número de falhas até o instante $t > 0$ então N_t é um processo de Poisson com taxa λ .

Suponha que o custo de cada replazo é β reais e que a taxa de desconto é $\alpha > 0$ (α pode ser a taxa de juros), i.e. cada real gasto no instante t tem um valor no presente de

$e^{-\alpha t}$. Considere o custo presente da troca o n -ésimo chip, $\beta e^{-\alpha T_n}$. Somando todos os custos temos que o valor presente de todas as futuras trocas é

$$C = \sum_{n=1}^{\infty} \beta e^{-\alpha T_n}.$$

Logo $\mathbb{E}C = \sum_{n=1}^{\infty} \beta \mathbb{E}e^{-\alpha T_n}$.

$\mathbb{E}e^{-\alpha T_n} = \left[\frac{\lambda}{\alpha + \lambda}\right]^n$ e portanto,

$$\mathbb{E}C = \beta \sum_{n=1}^{\infty} \left(\frac{\lambda}{\alpha + \lambda}\right)^n = \beta \frac{\frac{\lambda}{\alpha + \lambda}}{1 - \frac{\lambda}{\alpha + \lambda}} = \frac{\beta \lambda}{\alpha}.$$

Em particular, se o tempo de vida médio é $\mathbb{E}T_1 = 5000$ horas e o custo de cada troca é $\beta = 800$ reais e a taxa de juros é 24% ao ano então,

$$\alpha = \frac{0.24}{365 \times 24} = \frac{0.01}{365} \text{ e } \mathbb{E}C = 800 \frac{36500}{5000} = 5840 \text{ reais}.$$

EXEMPLO 3.5. Fixemos um ponto numa estrada e chamemos de U_1, U_2, \dots os sucessivos instantes nos quais passam automóveis pelo ponto. Suponha que estes tempos são variáveis aleatórias independentes e identicamente distribuídas com distribuição,

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

com densidade:

$$f_{U_k} = \lambda e^{-\lambda t} - [\lambda e^{-\lambda t} + (\lambda^2 t) e^{-\lambda t}] = \lambda^2 t e^{-\lambda t}.$$

Observe que $U_k \sim \text{Gamma}(2, \lambda)$, logo U_k pode se interpretar como a soma de dois tempos entre chegadas consecutivas num processo de Poisson com taxa λ . Logo, $U_1 = T_2$, $U_1 + U_2 = T_4$, $U_1 + U_2 + U_3 = T_6$, ... onde $T_i, i \geq 1$ são os tempos das chegadas do processo de Poisson $\{N_t\}_{t \geq 0}$. Seja M_t o número de carros que passam até o instante t . Por exemplo

$$M_t = 6 \Leftrightarrow N_t = 12 \text{ ou } N_t = 13.$$

Logo,

$$\begin{aligned} P(M_t = k) &= P(N_t = 2k) + P(N_t = 2k + 1) \\ &= \frac{e^{-\lambda t} (\lambda t)^{2k}}{(2k)!} + \frac{e^{-\lambda t} (\lambda t)^{2k+1}}{(2k+1)!} \\ &= \frac{e^{-\lambda t} (\lambda t)^{2k}}{(2k)!} \left[1 + \frac{(\lambda t)}{2k+1} \right]. \end{aligned}$$

PROPOSIÇÃO 3.6. Seja $f \geq 0$, então vale:

$$\mathbb{E} \left[\sum_{n=1}^{\infty} f(T_n) \right] = \lambda \int_0^{\infty} f(t) dt.$$

DEMONSTRAÇÃO. Observe que

$$\mathbb{E}[f(T_n)] = \int_0^\infty f(t) \frac{\lambda e^{-\lambda t} (\lambda t)^{n-1}}{(n-1)!} dt.$$

Logo,

$$\begin{aligned} \mathbb{E} \left[\sum_{n=1}^{\infty} f(T_n) \right] &= \sum_{n=1}^{\infty} \int_0^\infty f(t) \frac{\lambda e^{-\lambda t} (\lambda t)^{n-1}}{(n-1)!} dt \\ &= \int_0^\infty \lambda f(t) e^{-\lambda t} \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} dt \\ &= \int_0^\infty \lambda f(t) e^{-\lambda t} e^{\lambda t} dt = \lambda \int_0^\infty f(t) dt. \end{aligned}$$

□

4. Superposição de Processos de Poisson

Sejam $L = \{L_t, t \geq 0\}$ e $M = \{M_t, t \geq 0\}$ dois processos de Poisson independentes, com taxas μ e λ respectivamente (logo $M_t \sim \text{Poisson}(\mu t)$ e $L_t \sim \text{Poisson}(\lambda t)$). O processo $N_t = L_t + M_t$ é chamado **superposição** dos processos L e M .

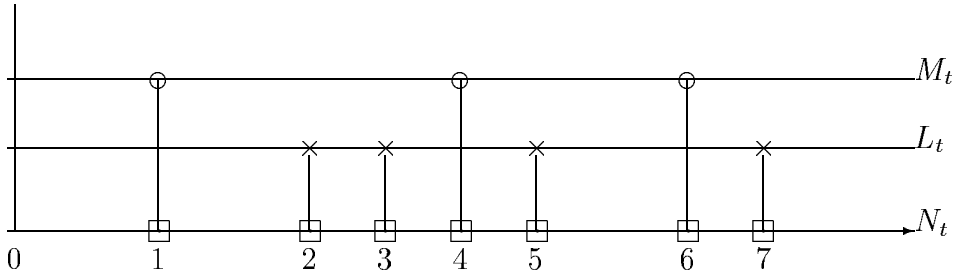


Figura 32

N_t também é chamado de processo de competência (entre os processos L_t e M_t).

EXEMPLO 4.1. Considere o exemplo anterior. Seja L_t o número de automóveis que chegam no ponto pela esquerda e R_t os que chegam no ponto pela direita. Então o número de automóveis que passam pelo ponto é $N_t = L_t + R_t$.

TEOREMA 4.2. A superposição de dois processo de Poisson independentes M e L com taxas μ e λ , respectivamente, é um processo de Poisson com taxa $\nu = \mu + \lambda$.

DEMONSTRAÇÃO. Basta tomar um intervalo $B \in \mathbb{R}^+$, observar que como L_B e M_B são independentes, então $N_B = L_B + M_B \sim \text{Poisson}((\lambda + \mu) \text{comp}(B))$ e usar o teorema 2.10. □

5. Decomposição de Processos de Poisson

Seja $\{X_n, n \geq 0\}$ um processo de Bernoulli com parâmetro p . X_n é o resultado do n -ésimo ensaio e $\{S_n, n \geq 0\}$ o número de sucessos nos n primeiros ensaios. Vamos supor que os tempos entre os ensaios são aleatórios. Para isto considere T_n , o instante do n -ésimo evento de um processo de Poisson $\{N_t, t \geq 0\}$ com taxa λ e suponha que os ensaios ocorrem nestes instantes. Assim, N_t seria o numero de ensaios até o instante t .

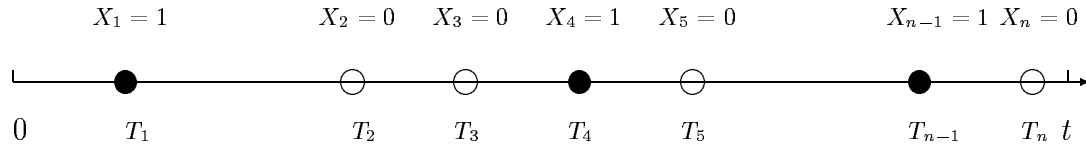


Figura 33

Finalmente, se M_t é o número de sucessos até o instante t então $M_t = S_{N_t}$. O número de fracassos até o instante t , L_t , seria $L_t = N_t - M_t$.

TEOREMA 5.1. *Os processos L_t e M_t são Poisson com taxas λp e $\lambda(1 - p)$ e são independentes*

EXEMPLO 5.2. *Suponha que os carros que chegam num determinado cruzamento viram à esquerda ou à direita independentemente com probabilidade $p = 0.6$ e $p = 0.4$, respectivamente.*

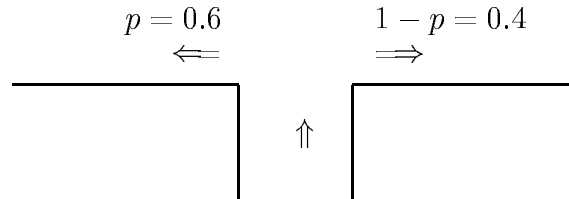


Figura 34

Seja N_t o número de carros que passaram no cruzamento até o instante t . Vamos supor que ele é um processo de Poisson com taxa $\lambda = 30$ carros por minuto.

Então, o número de carros que viraram à esquerda até o instante t , E_t , é um processo de Poisson com taxa $\lambda p = 30(0.6) = 18$ e o número de carros que viraram à direita até o instante t , D_t , é um processo de Poisson com taxa $\lambda(1 - p) = 30(0.4) = 12$. Os dois processos, D_t e E_t são independentes.

EXEMPLO 5.3. *Os carros que chegam num restaurante o fazem segundo um processo de Poisson com taxa $\lambda = 20$ por hora. Os vehiculos têm 1, 2, 3, 4 ou 5 ocupantes com*

probabilidades $p_i = 0.3, 0.3, 0.2, 0.1$ e 0.1 , respectivamente. Queremos achar o número esperado de pessoas que chegam no restaurante durante uma hora.

Sejam $N_t^{(1)}, N_t^{(2)}, N_t^{(3)}, N_t^{(4)}$ e $N_t^{(5)}$ o número de veículos que chegam com 1, 2, 3, 4 ou 5 ocupantes respectivamente até o instante t . Todos estes são processos de Poisson com taxas $\lambda p_i = 6, 6, 4, 2$ e 2 . Observe que $\mathbb{E}N_t^{(i)} = \lambda p_i$ e podemos assumir que eles são independentes. Logo, {

$$\begin{aligned} \text{Número esperado de pessoas numa hora} &= \mathbb{E}(1N_1^{(1)} + 2N_1^{(2)} + 3N_1^{(3)} + 4N_1^{(4)} + 5N_1^{(5)}) \\ &= (1 \times 6) + (2 \times 6) + (3 \times 4) + (4 \times 2) + (5 \times 2) \\ &= 48. \end{aligned}$$

}

6. Processo de Poisson Composto.

O processo do exemplo da soma dos sucessos em tempos aleatórios, $M_t = S_{N_t} = \sum_{n=0}^{N_t} X_n$ é um processo de Poisson composto. Vejamos como são definidos tais processos de forma geral.

DEFINIÇÃO 6.1. Chamaremos de **processo de Poisson composto** a um processo $\{Z_t\}_{t \geq 0}$ definido por

$$Z_t = \sum_{n=0}^{N_t} X_n,$$

onde $\{N_t\}_{t \geq 0}$ é um processo de Poisson e $\{X_n, n \geq 0\}$ é uma sequência de variáveis aleatórias i.i.d.

Tais processos tem as seguintes características,

- (1) as trajetórias têm um número finito de saltos sobre intervalos finitos;
- (2) $\forall t, s \geq 0, Z_{t+s} - Z_t$ é independente do passado $\{Z_u, u \leq t\}$;
- (3) $\forall t, s \geq 0$, a distribuição de $Z_{t+s} - Z_t$ é independente do t .

Observe que os saltos podem ser negativos. Uma possível trajetória é representada a seguir.

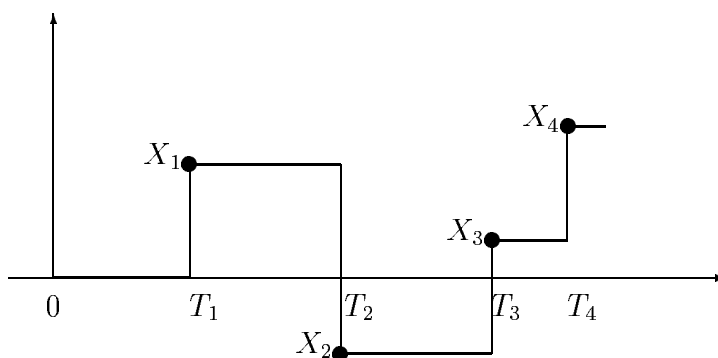


Figura 35

EXEMPLO 6.2. Suponha que o número de clientes que chegam num restaurante segue um processo de Poisson $\{N_t\}_{t \geq 0}$. O n -ésimo cliente gasta uma quantia X_n . Suponha que $\{X_n\}_{n \geq 1}$ é i.i.d.. Seja Y_n a quantia gasta por n clientes. Então

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

Logo, o total gasto pelos clientes até o instante t é um processo de Poisson composto,

$$Z_t = Y_{N_t} = \sum_{n=0}^{N_t} X_n.$$

EXEMPLO 6.3. O modelo clássico do risco.

O modelo clássico do risco na atividade seguradora é um processo estocástico

$$U(t) = u + ct - S(t),$$

onde $U(t)$ é o capital da seguradora no instante t (reserva de risco) e c é uma constante que representa o prêmio por unidade de tempo, de forma que ct será o prêmio que recebeu a seguradora até o instante t . u é a reserva inicial da seguradora e $S(t)$ representa o valor total das indenizações até o instante t ,

$$S(t) = \sum_{j=1}^{N_t} X_j$$

onde $\{X_n\}_{n \geq 1}$ é uma sequência de variáveis aleatórias não negativas que representam os valores das indenizações individuais que deve pagar a seguradora ante a ocorrência de sinistros e $\{N_t\}_{t \geq 0}$ é um processo de Poisson homogêneo das ocorrências das indenizações até o instante t .

Neste modelo, o total das indenizações $S(t)$ é um processo de Poisson composto.

Suponha que $X_n \in E = \{a, b, c, \dots\} \subseteq \mathbb{R}$. Considere somente o número de eventos, até o instante t , cuja magnitude é igual a a (i.e. $X_n = a$): $N_t^{(a)}$.

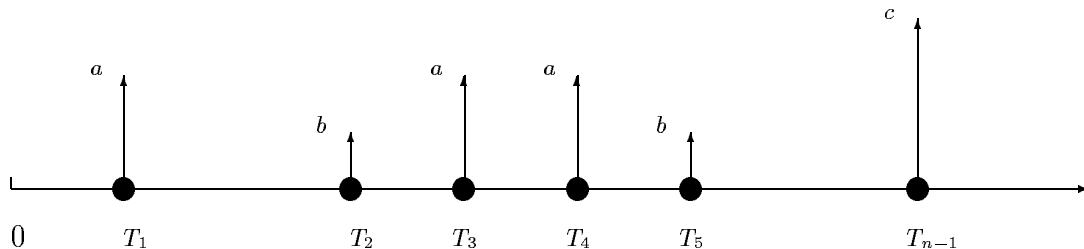


Figura 36

Então vale que:

$$\begin{aligned} N_t^{(a)} &\sim \text{Poisson com taxa } \lambda(a)t = \lambda P(X = a)t, \\ N_t^{(b)} &\sim \text{Poisson com taxa } \lambda(b)t = \lambda P(X = b)t, \\ &\dots \end{aligned}$$

Logo,

$$\begin{aligned} Z_t &= \sum_{n=0}^{N_t} X_n \\ &= \sum_{n: X_n=a}^{N_t^{(a)}} + \sum_{n: X_n=b}^{N_t^{(b)}} + \dots \\ &= aN_t^{(a)} + bN_t^{(b)} + \dots \end{aligned}$$

Observe que os somandos da última equação são independentes.

Uma outra propriedade do processo de Poisson composto é que $\mathbb{E}(Z_t) = \lambda t \mathbb{E}(X_n)$. Para provar isto, condicionamos em relação ao número de ocorrências do processo de Poisson,

$$\mathbb{E}(Z_t | N_t = n) = \mathbb{E} \left(\sum_{k=0}^n X_k \right) = \sum_{k=0}^n \mathbb{E}(X_k) = n \mathbb{E}(X_1),$$

portanto $\mathbb{E}(Z_t | N_t) = N_t \mathbb{E}(X_1)$. Logo,

$$\begin{aligned} \mathbb{E}(Z_t) &= \mathbb{E}(\mathbb{E}(Z_t | N_t)) = \mathbb{E}(N_t \mathbb{E}(X_1)) \\ &= \mathbb{E}(N_t) \mathbb{E}(X_1) = \lambda t \mathbb{E}(X_1). \end{aligned}$$

7. Processo de Poisson não homogêneo.

DEFINIÇÃO 7.1. O processo $\{N_t\}_{t \geq 0}$ é um processo de Poisson não homogêneo se

- (1) N_t tem saltos de tamanho um;
- (2) a distribuição de $N_{t+s} - N_t$ é independente de t .

Seja $a(t) = \mathbb{E}(N_t)$ e observe que $a(t)$ é crescente pois:

$$N_{t+s} \geq N_t \Rightarrow \mathbb{E}N_{t+s} \geq \mathbb{E}N_t \Rightarrow a(t+s) \geq a(t)$$

A função $a(t)$ é chamada de **taxa** do processo. Suponha que $a(t)$ é contínua e defina $\tau(u) = a^{-1}(u)$, a inversa de a .

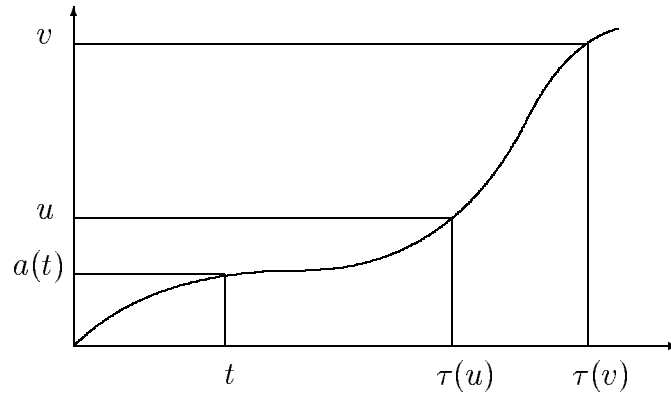


Figura 37

TEOREMA 7.2. *Seja $\{N_t\}_{t \geq 0}$ um processo de Poisson não homogêneo com taxa $a(t)$. Defina o processo $M_t = N_{\tau(t)}$. Então M_t é um processo de Poisson com taxa 1.*

Observe que $\mathbb{E}M_t = \mathbb{E}N_{\tau(t)} = a(\tau(t)) = t$.

TEOREMA 7.3. *Seja $\{N_t\}_{t \geq 0}$ um processo de Poisson não homogêneo com taxa $a(t)$ e seja $b(t, s) = a(t + s) - a(t)$, então*

(1) *os incrementos têm distribuição*

$$P(N_{t+s} - N_t = k) = \frac{e^{-b(t,s)} b(t,s)^k}{k!};$$

(2) *se T_n é o tempo do n -ésimo evento do processo, então*

$$P(T_{n+1} - T_n > t | T_1, T_2, \dots, T_n) = e^{-\lambda(a(T_n+t) - a(T_n))}.$$

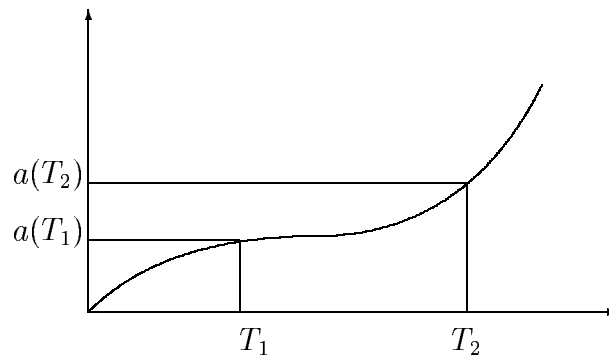


Figura 38

Em geral a taxa do processo é uma função descontínua.

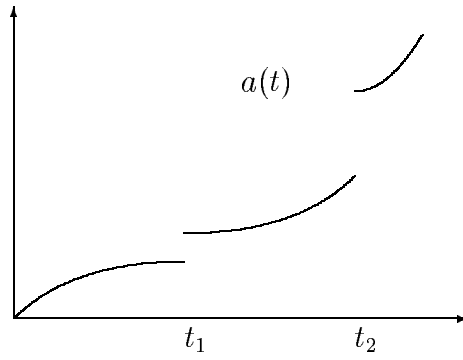


Figura 39

Nesse caso podemos decompor o processo como soma de dois processos N_t^c (um processo com taxa contínua) e N_t^d (Um processo que é soma de processos homogêneos definidos apenas em intervalos de tempo). As taxas dos dois processos podem ver-se na seguinte figura.

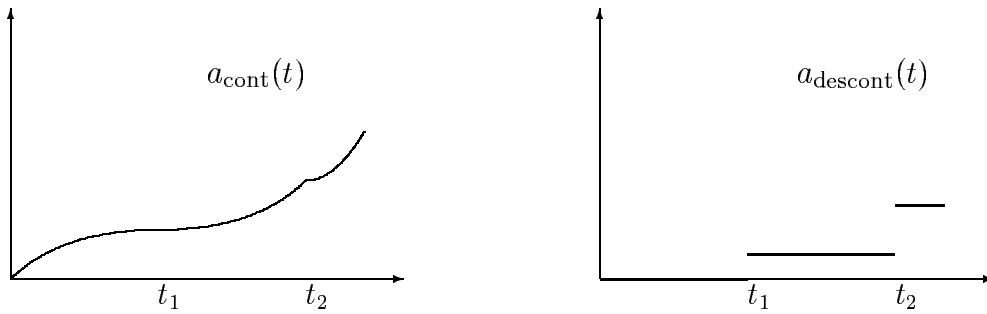


Figura 40

CAPÍTULO 4

Cadeias de Markov a Tempo Contínuo

Neste capítulo estudaremos as cadeias de Markov a tempo contínuo. Veremos que estes processos estão relacionados de maneira natural com as cadeias de Markov a tempo discreto e também que o processo de Poisson é um caso particular de cadeia de Markov a tempo contínuo.

1. Definição e exemplos.

DEFINIÇÃO 1.1. *Considere um processo estocástico a tempo contínuo $X = \{X_t\}_{t \geq 0}$ com espaço de estados E finito ou enumerável. Diremos que X é uma **cadeia de Markov a tempo contínuo** se e somente se para todo $t, s \geq 0$*

$$P(X_{t+s} = j | X_u, u \leq s) = P(X_{t+s} = j | X_s). \quad (1.23)$$

Se além disto, a probabilidade de transição entre dois estados depende somente do intervalo de tempo durante o qual ocorre a transição e não dos instantes de tempo nos que a cadeia ocupa esses estados, ou seja, quando

$$P(X_{t+s} = j | X_s = i) = P_{i,j}(t), \quad (1.24)$$

*a cadeia será chamada de **homogênea no tempo**.*

A equação (1.23) exprime a propriedade de Markov da cadeia que, como no caso discreto, significa que a predição que podemos fazer sobre o futuro depende da história anterior do processo somente através do instante presente.

Daqui para a frente consideraremos somente cadeias de Markov homogêneas no tempo. Para tais cadeias chamaremos à família de matrizes $P(t) = (P_{i,j}(t))_{i,j \in E}$ de **função de transição** da cadeia X . Ela satisfaz as seguintes propriedades.

- (1) $P(0) = I$.
- (2) Para todo $t \geq 0$, $P(t)$ é uma matriz estocástica.
- (3) Para $t, s \geq 0$, $P(t+s) = P(t)P(s)$.

As propriedades 1) e 2) decorrem da definição de P . A propriedade 3, pode ser provada de forma análoga às equações de Chapman-Kolmogorov das cadeias de Markov a tempo discreto, mas será apresentada aqui por motivos que ficarão esclarecidos mais para frente.

PROPOSIÇÃO 1.2. *Equações de Chapman-Kolmogorov para cadeias a tempo contínuo. Para todo $t, s \geq 0$, $i, j \in E$ vale*

$$P_{i,j}(t+s) = \sum_{k \in E} P_{i,k}(t)P_{k,j}(s). \quad (1.25)$$

DEMONSTRAÇÃO. Decompondo o espaço amostral da forma

$$\Omega = \bigcup_{k \in E} \{X_t = k\}$$

e usando a lei da probabilidade total e a propriedade de Markov podemos obter

$$\begin{aligned} P_{i,j}(t+s) &= P(X_{t+s} = j | X_0 = i) \\ &= \sum_{k \in E} P(X_{t+s} = j | X_t = k, X_0 = i) P(X_t = k | X_0 = i) \\ &= \sum_{k \in E} P(X_s = j | X_0 = k) P(X_t = k | X_0 = i) \\ &= \sum_{k \in E} P_{k,j}(s) P_{i,k}(t) \end{aligned}$$

e o resultado desejado fica provado. \square

Na prova acima condicionamos em relação a eventos que fixam o estado do processo no instante t . Para provar a igualdade $P(t+s) = P(s)P(t)$, que também vale pois os papéis de s e t podem ser trocados, teríamos que fixar o estado do processo no instante s .

EXEMPLO 1.3. *Seja $\{N_t\}_{t \geq 0}$ um processo de Poisson homogêneo com taxa $\lambda > 0$. Vejamos que ele é uma cadeia de Markov a tempo contínuo. Para isto consideremos $t, s \geq 0$.*

$$\begin{aligned} P(N_{t+s} = j | N_s = i, N_u = i(u), u < s) &= P(N_{t+s} - N_s = j - i | N_s = i, N_u = i(u), u < s) \\ &= P(N_{t+s} - N_s = j - i) \\ &= P(N_t = j - i). \end{aligned}$$

Obtemos então que

$$P_{i,j}(t) = P(N_t = j - i) = \begin{cases} \frac{e^{-\lambda t} (\lambda t)^{j-i}}{(j-i)!}, & j \geq i, \\ 0, & \text{caso contrário.} \end{cases}$$

EXEMPLO 1.4. *Consideremos mais uma vez um processo de Poisson $\{N_t\}_{t \geq 0}$ com taxa $\lambda > 0$ e defina $\{X_t\}_{t \geq 0}$, com $E = \{1, -1\}$ por*

$$X_t = X_0(-1)^{N_t},$$

sendo X_0 uma variável aleatória com valores em E e independente do processo $\{N_t\}_{t \geq 0}$.

Observe que

$$X_{t+s} = X_0(-1)^{N_{t+s}} = X_0(-1)^{N_s}(-1)^{N_{t+s}-N_s} = X_s(-1)^{N_{t+s}-N_s},$$

então os valores de X_{t+s} dependem da história do processo até o instante s através do estado em s , X_s e do incremento $N_{t+s}-N_s$, que é independente de $\{N_u, u \leq s\}$ e portanto de $\{X_u, u \leq s\}$. Logo,

$$\begin{aligned} P(X_{t+s} = j | X_s = i, X_u = i(u), u < s) &= P(X_s(-1)^{N_{t+s}-N_s} = j | X_s = i) \\ &= P(i(-1)^{N_{t+s}-N_s} = j) \\ &= P(i(-1)^{N_t} = j) \\ &= P_{i,j}(t). \end{aligned}$$

Por exemplo,

$$\begin{aligned} P_{-1,1}(t) &= P(N_t \text{ ímpar}) \\ &= \sum_{k=0}^{+\infty} \frac{e^{-\lambda t} (\lambda t)^{2k+1}}{(2k+1)!} \\ &= \frac{1 - e^{-2\lambda t}}{2} \end{aligned}$$

e procedendo de forma análoga nos outros casos obteremos

$$P(t) = \frac{1}{2} \begin{bmatrix} 1 + e^{-2\lambda t} & 1 - e^{-2\lambda t} \\ 1 - e^{-2\lambda t} & 1 + e^{-2\lambda t} \end{bmatrix}.$$

EXEMPLO 1.5. *Cadeia de Markov uniforme.*

Seja $\{\widehat{Y}_n\}_{n \geq 0}$ uma cadeia de Markov a tempo discreto com espaço de estados E e matriz de transição $K = (k_{ij})_{i,j \in E}$ e seja $\{\tau_n\}_n \geq 1$ o processo dos tempos das chegadas de um processo de Poisson $N = \{N_t\}_{t \geq 0}$ com taxa $\lambda > 0$. Suponha que $\{\widehat{Y}_n\}_{n \geq 0}$ e N são independentes.

O processo definido por

$$X_t = \widehat{Y}_{N_t}$$

chama-se *cadeia de Markov uniforme*, N é o *relógio* da cadeia e $\{\widehat{Y}_n\}_{n \geq 0}$ se chama de *cadeia subordinada*.

$$\begin{aligned} P_i(X_t = j) &= P_i(\widehat{Y}_{N_t} = j) \\ &= \sum_{n=0}^{+\infty} P_i(\widehat{Y}_n = j, N_t = n) \\ &= \sum_{n=0}^{+\infty} P_i(\widehat{Y}_n = j) P_i(N_t = n) \\ &= \sum_{n=0}^{+\infty} k_{i,j}^n \frac{e^{-\lambda t} (\lambda t)^n}{n!}, \end{aligned}$$

ou seja,

$$P(t) = \sum_{n=0}^{+\infty} \frac{e^{-\lambda t} (\lambda t)^n}{n!} K^n.$$

As propriedades a seguir afirmam que as distribuições finito-dimensionais do processo estão determinadas pela função de transição e pela distribuição inicial, o que também acontecia no caso discreto.

PROPOSIÇÃO 1.6. *Sejam $0 \leq t_0 < t_1 < \dots < t_n$ instantes de tempo e sejam $i_0, i_1, \dots, i_n \in E$, então*

$$P(X_{t_1} = i_1, X_{t_2} = i_2, \dots, X_{t_n} = i_n | X_{t_0} = i_0) = P_{i_0, i_1}(t_1 - t_0) P_{i_1, i_2}(t_2 - t_1) \cdots P_{i_{n-1}, i_n}(t_n - t_{n-1}). \quad (1.26)$$

Se além disto vale que π é a distribuição inicial da cadeia, ou seja, π é a distribuição de probabilidade de X_0 , então

$$P(X_{t_1} = i_1, X_{t_2} = i_2, \dots, X_{t_n} = i_n) = \sum_{i_0 \in E} \pi(i_0) P_{i_0, i_1}(t_1) P_{i_1, i_2}(t_2 - t_1) \cdots P_{i_{n-1}, i_n}(t_n - t_{n-1}). \quad (1.27)$$

EXEMPLO 1.7. *Considere a cadeia X com espaço de estados $E = \{a, b\}$ e função de transição*

$$P(t) = \begin{bmatrix} 0,6 + 0,4e^{-5t} & 0,4 - 0,4e^{-5t} \\ 0,6 - 0,6e^{-5t} & 0,4 + 0,6e^{-5t} \end{bmatrix}.$$

Calcule $P_a(X_{2,4} = b, X_{3,8} = a, X_{4,2} = a)$.

Usando a proposição 1.26 obtemos,

$$\begin{aligned} P_a(X_{2,4} = b, X_{3,8} = a, X_{4,2} = a) &= P(X_{2,4} = b, X_{3,8} = a, X_{4,2} = a | X_0 = a) \\ &= P_{a,b}(2,4) P_{b,a}(1,4) P_{a,a}(0,4) \\ &\approx 0,15. \end{aligned}$$

2. Estrutura de uma cadeia de Markov a tempo contínuo.

Podemos considerar o período de tempo que a cadeia permanece no estado que ela ocupa no instante t . Este será uma variável aleatória que chamaremos de W_t e pode ser definida da seguinte maneira,

$$W_t(w) = \inf\{s \geq 0 \text{ tal que } X_{t+s}(w) \neq X_t(w)\}.$$

Segundo o comportamento desta variável, os estados podem ser classificados como segue

- (1) i será chamado de **estado instantâneo** se $P(W_t = 0 | X_t = i) = 1$.

Observe que neste caso a cadeia fica no estado i somente no instante que ela chegou.

- (2) i será chamado de **estado absorvente** se $P(W_t < +\infty | X_t = i) = 0$.

Uma vez que a cadeia chega num estado absorvente ela fica nele para sempre.

- (3) i será chamado de **estado estável** se $P(0 < W_t < +\infty | X_t = i) = 1$.

Toda vez que a cadeia chega num estado estável, ela fica nele durante um período de tempo finito.

Cadeias com estados instantâneos são muito raras (mas existem!, veja [?]), pois é possível provar que

- E finito \Rightarrow não há estado instantâneos.
- Trajetórias contínuas direita com probabilidade um \Leftrightarrow não há estados instantâneos.

Para cadeias sem estados instantâneos podemos determinar a distribuição da variável W_t .

TEOREMA 2.1. *Seja $\{X_t\}_{t \geq 0}$ uma cadeia sem estados instantâneos. Para todo $i \in E$ e todo $t \geq 0$,*

$$P(W_t > u | X_t = i) = e^{-q_i u}, \quad u \geq 0,$$

para algum número $q_i \in [0, +\infty)$.

DEMONSTRAÇÃO. Fixemos um estado $i \in E$. Pela homogeneidade no tempo da cadeia teremos que $P(W_t > u | X_t = i)$ não depende do tempo. Chamemos esta função de $f(u)$.

Observe agora que $\{W_t > u + v\}$ se e somente se $\{W_t > u\}$ e $\{W_{t+u} > v\}$. Logo,

$$\begin{aligned} f(u + v) &= P(W_t > u + v | X_t = i) \\ &= P(W_t > u, W_{t+u} > v | X_t = i) \\ &= P(W_{t+u} > v | W_t > u, X_t = i) P(W_t > u | X_t = i) \end{aligned}$$

Examinemos o termo $P(W_{t+u} > v | W_t > u, X_t = i)$. Pela continuidade à direita das trajetórias teremos que

$$\{X_t = i, W_t > u\} = \{X_\tau = i, t \leq \tau \leq t + u\},$$

logo, pela propriedade de Markov e a homogeneidade no tempo,

$$\begin{aligned} P(W_{t+u} > v | W_t > u, X_t = i) &= P(W_{t+u} > v | X_\tau = i, t \leq \tau \leq t+u) \\ &= P(W_{t+u} > v | X_{t+u} = i) \\ &= P(W_t > v | X_t = i) \\ &= f(v). \end{aligned}$$

Portanto, a função f satisfaz a equação de Cauchy $f(u+v) = f(u)f(v)$ e então ou ela é identicamente nula ou existe uma constante $q_i \in \mathbb{R}$ tal que $f(u) = e^{-q_i u}$. Observe que f é identicamente nula se e somente se i é um estado instantâneo, que não está sendo considerado aqui. Por outro lado, como para $u > v$ vale $\{W_t > u\} \subseteq \{W_t > v\}$, temos que f é decrescente e portanto $q_i \geq 0$. \square

Repare que $q_i = 0 \Leftrightarrow i$ absorvente. Além disto, se admitirmos o valor $+\infty$ para q_i , entenderemos que $P(W_t > u | X_t = i) = 0$, para todo $u \geq 0$, ou seja, i seria um estado instantâneo. No que segue suporemos que os estados não são instantâneos. Cadeias sem estados instantâneos são chamadas de **processos de saltos**. As trajetórias dos processos de saltos são contínuas à direita com probabilidade um.

Suponha que a cadeia $\{X_t\}_{t \geq 0}$ é tal que todos os seus estados são estáveis. Para cada $n \in \mathbb{N}$, podemos considerar a variável aleatória T_n : instante de tempo no qual a cadeia muda de estado pela n -éssima vez. Definiremos $T_0 = 0$ por conveniência. Temos então que

$$0 = T_0 < T_1 < \dots < T_n < \dots$$

e vale

$$X_t = \hat{X}_n, \quad T_n \leq t < T_{n+1}.$$

Aqui estamos denotando por \hat{X}_n o estado visitado pela cadeia na n -éssima transição que ela faz quando $n \geq 1$ e \hat{X}_0 é o estado inicial.

Observe que $W_0 = T_1$ e portanto $T_1 | \hat{X}_0 = i \sim \exp(q_i)$. De forma análoga teremos que $T_{n+1} = T_n + W_{T_n}$ e $T_{n+1} - T_n | \hat{X}_n = i \sim \exp(q_i)$.

Para poder fazer a construção acima para todo $t \in \mathbb{R}$ precisaremos que $\lim_{n \rightarrow \infty} T_n = +\infty$. Não é difícil ver que isto é equivalente a pedir que as trajetórias de $\{X_t\}_{t \geq 0}$ tenham um número finito de saltos em cada intervalo finito de tempo. Quando isto ocorre fala-se que a cadeia é **regular**. No que segue trabalharemos somente com processos de saltos regulares.

Observe a sequência $\{\hat{X}_n\}_{n \geq 0}$ que acabamos de definir é uma cadeia de Markov a tempo discreto tal que as transições são feitas somente entre estados diferentes, ou seja, se chamarmos de $Q = (Q_{ij})_{i,j \in E}$ à matriz de transição desta cadeia, então ela deve satisfazer $Q_{ii} = 0$, para todo $i \in E$.

A cadeia $\{\hat{X}_n\}_{n \geq 0}$ determina a sequência dos estados visitados. Ela costuma ser chamada de **esqueleto** da cadeia $\{X_t\}_{t \geq 0}$. A sequência $\{T_n\}_{n \geq 0}$ determina os instantes em que são feitas as transições, portanto, ambas sequências determinam a cadeia $\{X_t\}_{t \geq 0}$.

Sabemos que a matriz de transição Q determina a cadeia a tempo discreto $\{\hat{X}_n\}_{n \geq 0}$. Por outro lado, é possível provar que dada $\{\hat{X}_n\}_{n \geq 0}$, a sequência $\{T_{n+1} - T_n\}_{n \geq 0}$ é independente, portanto $\{T_n\}_{n \geq 0}$ está determinada por Q e pelas constantes $\{q_i\}_{i \in E}$. Logo, $\{X_t\}_{t \geq 0}$ está determinada pelos elementos $\{q_i\}_{i \in E}$ e $\{Q_{ij}\}_{i,j \in E}$. Isto vale também para cadeias com estados absorventes, sendo que para i absorvente teremos $q_i = 0$, $Q_{ii} = 1$, $T_m = +\infty$ e $T_{n+1} - T_n = +\infty$, $n \geq m$ se i é visitado pela cadeia pela primeira vez na $m - 1$ -ésima transição.

3. O gerador infinitesimal.

Definamos para $i \neq j$, $q_{ij} = q_i Q_{ij}$. Esta é a taxa com que o processo faz uma transição desde i para j . Para $i = j$ façamos $q_{ii} = -q_i$.

DEFINIÇÃO 3.1. A matriz $A = (q_{ij})_{i,j \in E}$ é chamada de **gerador infinitesimal** da cadeia $\{X_t\}_{t \geq 0}$.

O gerador infinitesimal é obtido a partir das taxas $\{q_i\}_{i \in E}$ e das probabilidades de transição $\{Q_{ij}\}_{i,j \in E}$ e viceversa, as taxas e as probabilidades de transição podem ser obtidas a partir do gerador. De fato, temos que $q_i = -q_{ii}$ e que para $i \neq j$ vale $Q_{ij} = 0$ se $q_i = 0$ e $Q_{ij} = q_{ij}/q_i$ em caso contrário. Portanto, ele determina a cadeia e daí vem o seu nome. Ele também pode ser obtido a partir da função de transição como veremos a seguir.

PROPOSIÇÃO 3.2. Para cada $i, j \in E$, a função $P_{i,j}(t)$ é diferenciável e a sua derivada é contínua. Além disto, vale $P'_{i,j}(0) = q_{ij}$, ou seja

$$\left. \frac{d}{dt} \right|_{t=0} P(t) = A.$$

EXEMPLO 3.3. Para o exemplo 1.7 calcule o gerador infinitesimal, as taxas q_i e as probabilidades Q_{ij} .

Solução:

$$A = \left. \frac{d}{dt} \right|_{t=0} \begin{bmatrix} 0,6 + 0,4e^{-5t} & 0,4 - 0,4e^{-5t} \\ 0,6 - 0,6e^{-5t} & 0,4 + 0,6e^{-5t} \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix}.$$

As taxas são os opostos dos elementos na diagonal, portanto $q_a = 2$ e $q_b = 3$. Observe que ambos os estados são estáveis. Agora podemos calcular as probabilidades de transição. Sabemos que $2 = q_{ab} = q_a Q_{ab} = 2Q_{ab}$, portanto $Q_{ab} = 1$. Analogamente obtemos $Q_{ba} = 1$. Observe que como Q é uma matriz de transição de tamanho dois e os estados são estáveis, na verdade não precisávamos calcular as entradas pois se $Q_{aa} = Q_{bb} = 0$ necessariamente teremos $Q_{ab} = Q_{ba} = 1$, pois as somas das entradas por linhas é um.

EXEMPLO 3.4. Considere a cadeia de Markov com espaço de estados $E = \{a, b, c\}$, matriz de transição

$$Q = \begin{bmatrix} 0 & 1 & 0 \\ 0,4 & 0 & 0,6 \\ 0 & 1 & 0 \end{bmatrix}$$

e taxas $q_a = 2$, $q_b = 5$ e $q_c = 3$. Calcule o gerador infinitesimal deste processo.

Solução:

Como as taxas são todas não nulas, os estados são estáveis. Os elementos da diagonal de A são os opostos das taxas. As entradas restantes são obtidas multiplicando a taxas pelas probabilidades de transição. Temos então que $q_{ab} = q_a \cdot Q_{ab} = 2 \cdot 1 = 2$ e $q_{ac} = q_a \cdot Q_{ac} = 2 \cdot 0 = 0$. Procedendo de forma similar com as outras linhas, obtemos

$$A = \begin{bmatrix} -2 & 2 & 0 \\ 2 & -5 & 3 \\ 0 & 3 & -3 \end{bmatrix}.$$

EXEMPLO 3.5. Calculemos o gerador infinitesimal da cadeia de Markov uniforme (exemplo 1.5). Para isto, escreveremos a sua função de transição da forma seguinte.

$$P(t) = e^{-\lambda t} I + \sum_{n=1}^{+\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} K^n.$$

Usando que $(e^{-\lambda t})' = -\lambda e^{-\lambda t}$ e que para $n \geq 1$ $(e^{-\lambda t} (\lambda t)^n)' = \lambda^n e^{-\lambda t} t^{n-1} (n - \lambda t)$, obtemos que

$$A = P'(0) = \lambda(K - I).$$

Para cada estado $i \in E$, teremos $q_i = \lambda(1 - k_{ii})$. Logo $q_i = 0 \Leftrightarrow k_{ii} = 1$, em outras palavras, i será absorvente para X se e somente se ele o for para \hat{Y} . Se i não for absorvente para \hat{Y} ele será estável e para $j \neq i$ teremos $Q_{ij} = \frac{k_{ij}}{1 - k_{ii}}$.

4. As equações diferenciais de Kolmogorov.

Até agora sabemos como obter A a partir de $\{P(t)\}_{t \geq 0}$. Vejamos como podemos fazer o contrário. Para isto precisaremos das equações de Chapman-Kolmogorov. Obteremos equações diferenciais (as equações de Kolmogorov) para as transições $P(t)$ que envolvem as entradas da matriz A .

Calculemos a derivada de $P(t)$. Usando a definição de A e as equações de Chapman-Kolmogorov obtemos

$$\begin{aligned}
P'(t) &= \lim_{h \searrow 0} \frac{P(t+h) - P(t)}{h} \\
&= \lim_{h \searrow 0} \frac{P(h)P(t) - P(t)}{h} \\
&= \lim_{h \searrow 0} \frac{P(h) - I}{h} P(t) \\
&= AP(t).
\end{aligned}$$

Lembremos que para provar a relação $P(t+h) = P(h)P(t)$ tivemos que condicionar em relação a eventos que fixavam o estado da cadeia no instante h , que por estar convergindo para zero, pode ser considerado anterior ao instante t . Por esta razão as equações obtidas,

$$P'(t) = AP(t)$$

são chamadas de **equações de Kolmogorov retrospectivas** (backward Kolmogorov equations). Escritas componente a componente, elas ficariam da forma

$$P'_{ij}(t) = q_i \sum_{k \neq i} Q_{ik} P_{kj}(t) - q_i P_{ij}(t).$$

Se usarmos a igualdade $P(t+h) = P(t)P(h)$, obteríamos as chamadas **equações de Kolmogorov prospectivas** (forward Kolmogorov equations),

$$P'(t) = P(t)A$$

ou, componente a componente

$$P'_{ij}(t) = \sum_{k \neq j} q_k Q_{kj} P_{ik}(t) - q_j P_{ij}(t).$$

Observe que na dedução da equação de Chapman-Kolmogorov utilizada agora, precisamos condicionar em relação a eventos que fixam o estado da cadeia no instante t , posterior ao instante h . Dai o nome de equações prospectivas.

A derivação que fizemos das equações de Kolmogorov está correta no caso que E é finito, mas não no caso infinito, pois precisamos passar um limite para dentro de um somatório com infinitos termos. Isto pode ser justificado rigorosamente em particular quando vale a condição $\sup_{i \in E} q_i < \infty$, condição que será satisfeita pelos exemplos que consideraremos aqui.

EXEMPLO 4.1. *Cadeia com dois estados*

Calcule a função de transição de uma cadeia com dois estados que permanece no estado 0 durante um tempo exponencial com taxa $\lambda > 0$ antes de passar ao estado 1, onde estará durante um tempo exponencial com taxa $\mu > 0$ antes de voltar ao estado 0.

Solução:

Pelo enunciado da questão sabemos que $q_0 = \lambda$ e $q_1 = \mu$ e como os dois estados são estáveis, necessariamente a matriz Q toma a forma

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

A equação de Kolmogorov prospectiva com $i = j = 0$ será

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

onde a última equação foi obtida a partir de $P_{0,1}(t) = 1 - P_{0,0}(t)$. Portanto,

$$e^{(\lambda+\mu)t} [P'_{0,0} + (\lambda + \mu)P_{0,0}(t)] = \mu e^{(\lambda+\mu)t},$$

ou seja,

$$[e^{(\lambda+\mu)t} P_{0,0}(t)]' = \mu e^{(\lambda+\mu)t}.$$

Integrando entre 0 e s , rearranjando os termos e usando que $P_{0,0}(0) = 1$, podemos obter

$$P_{0,0}(s) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)s}.$$

Por simetria teremos também

$$P_{1,1}(s) = \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} e^{-(\lambda+\mu)s}$$

e usando o fato que $P(t)$ é uma matriz de transição,

$$P_{0,1}(s) = \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)s},$$

$$P_{1,0}(s) = \frac{\mu}{\lambda + \mu} - \frac{\mu}{\lambda + \mu} e^{-(\lambda+\mu)s}.$$

No exemplo anterior conseguimos obter $P(t)$ pois escrevemos um equação em termos de uma só função incôgnita. Isto não pode ser feito sempre, mas para matrizes finitas, existe uma forma de resolver estas equações.

EXEMPLO 4.2. *Cadeia com espaço de estados finito.*

Suponha que o espaço de estados de uma cadeia X tem N elementos. A matriz A será então uma matriz de $N \times N$. As equações de Kolmogorov serão

$$P'(t) = AP(t) = P(t)A$$

e sabemos que $P(0) = I$. Se $N = 1$, então simplesmente A seria um escalar e teríamos $P(t) = e^{At}$. Usando a série de Taylor da exponencial poderíamos escrever também

$$P(t) = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!}.$$

Esta última expressão faz sentido também no caso que A é uma matriz pois depende só das suas potências.

Para A matriz de $N \times N$ define-se

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n,$$

lembrando que $A^0 = I$. Prova-se que $P(t) = e^{tA}$ é neste caso a única solução das equações de Kolmogorov.

Para calcular e^{tA} não é preciso computar os coeficientes da série, pois existe uma relação natural entre a decomposição espectral da matriz A e aquela de e^{tA} . Isto pode ser consultado em livros de álgebra linear.

5. Distribuição estacionária e comportamento assintótico.

Suponha que no estado inicial a cadeia não se encontra no estado $i \in E$. Então

$$\tau_i = \inf \{t \geq T_1 : X_t = i\}$$

será o primeiro instante em que a cadeia visita o estado i . Por outro lado, se $X_0 = i$, τ_i representará o primeiro instante em que a cadeia voltou a este estado. τ_i pode ser interpretado então como o primeiro tempo de visita ou de retorno no estado i . Observe que τ_i pode tomar o valor $+\infty$ se a cadeia nunca visitar o estado i depois do instante T_1 ou se i for um estado absorvente.

Quando estudamos cadeias de Markov a tempo discreto, vimos que os estados eram classificados dependendo da frequência com que eles eram visitados. No caso contínuo esta frequência está representada por $P_i(\tau_i < \infty)$ e temos portanto, a seguinte classificação.

DEFINIÇÃO 5.1. O estado $i \in E$ será chamado de **transitório** se $P_i(\tau_i < \infty) < 1$. Caso contrário, ele será chamado de **recorrente**. Um estado $i \in E$ recorrente será chamado de **recorrente nulo** se $\mathbb{E}_i(\tau_i) = \infty$ e será de **positivo** em caso contrário. Os estados absorventes são considerados recorrentes positivos.

Observe que como no caso discreto, os estados recorrentes são visitados infinitas vezes a diferença dos transitórios. Portanto, um estado será recorrente para a cadeia X se e somente se ele o for para o seu esqueleto $\{\hat{X}_n\}_{n \geq 0}$.

Os estados recorrentes nulos são parecidos aos transitórios no sentido que os tempos entre visitas consecutivas são em média, muito grandes. A recorrência nula ou positiva da cadeia X e do seu esqueleto são propriedades diferentes. Um indício disto é que a esperança $\mathbb{E}_i(\tau_i)$ na definição acima depende da matriz Q , mas também das taxas $\{q_i\}$.

Um conjunto $\mathcal{C} \subseteq E$ era irreduzível no caso discreto se todos os seus elementos estavam comunicados entre si. De forma análoga diremos agora que $\mathcal{C} \subseteq E$ é **irreduzível** quando para todos $i, j \in \mathcal{C}$ vale $P_i(\tau_j < \infty) > 0$, ou seja a cadeia vai de i para j com probabilidade positiva. Os conjuntos irreduzíveis para X e para o seu esqueleto coincidem.

DEFINIÇÃO 5.2. Uma distribuição π sobre o espaço de estados E será chamada de **distribuição estacionária** da cadeia X se para todo $j \in E$ e todo $s \geq 0$,

$$\sum_{k \in E} \pi(k) P_{k,j}(s) = \pi(j). \quad (5.28)$$

No caso finito podemos escrever (5.28) da forma compacta

$$\pi^t \cdot P(s) = \pi^t. \quad (5.29)$$

Novamente, se iniciarmos a cadeia com a distribuição estacionária, teremos que todos os estados possuem a mesma distribuição pois pela igualdade (1.27) teríamos,

$$P(X_t = j) = \sum \pi(i_0) P_{i_0,j}(t) = \pi(j).$$

A diferença do caso discreto, agora teremos uma quantidade não enumerável de sistemas de equações que definem a distribuição estacionária. Seria desejável reduzir o cálculo a um só sistema e isto será possível graças ao gerador infinitesimal da cadeia.

Suponha que temos uma distribuição estacionária π . Então ela satisfaz (5.28). Podemos derivar ambos termos da equação em relação a t e avaliar no zero. Se pudéssemos passar a derivada para dentro do somatório (como suporemos aqui) obteríamos

$$\sum_{k \in E} \pi(k) q_{k,j} = 0, \quad (5.30)$$

e no caso finito

$$\pi^t \cdot A = 0. \quad (5.31)$$

Podemos nos plantear também o problema no sentido inverso: será que uma distribuição que satisfaça (5.30) tem que ser uma distribuição estacionária? Para responder esta questão usaremos as equações de Kolmogorov retrospectivas.

Suponha que uma distribuição π satisfaz (5.30). Calculemos $\frac{d}{ds} (\sum_{k \in E} \pi(k) P_{k,j}(s))$. Mais uma vez suporemos que podemos passar a derivada para dentro do somatório. Então,

$$\begin{aligned} \frac{d}{ds} \left(\sum_{k \in E} \pi(k) P_{k,j}(s) \right) &= \sum_{k \in E} \pi(k) \frac{d}{ds} P_{k,j}(s) \\ &= \sum_{k \in E} \pi(k) \sum_{l \in E} q_{kl} P_{l,j}(s) \\ &= \sum_{l \in E} P_{l,j}(s) \sum_{k \in E} \pi(k) q_{kl} \\ &= 0, \end{aligned}$$

pela equação (5.30). Ou seja, $\sum_{k \in E} \pi(k) P_{k,j}(s)$ não depende de s e portanto tem que ser igual a $\pi(j)$, que é o valor que toma esta expressão para $s = 0$.

Temos então que no caso finito ou no caso infinito sob suposições bastante razoáveis vale que uma distribuição é estacionária se e somente se ela satisfizer (5.30).

EXEMPLO 5.3. Seja X uma cadeia de Markov com gerador infinitesimal

$$A = \begin{bmatrix} -5 & 2 & 3 \\ 2 & -3 & 1 \\ 2 & 4 & -6 \end{bmatrix}.$$

Para encontrar a (ou as) distribuição estacionária, da cadeia basta encontrar as distribuições que sejam solução do sistema $\pi^t \cdot A = 0$. Observe que é suficiente encontrar uma solução (x, y, z) do sistema $[x \ y \ z] \cdot A = 0$ e depois dividir cada uma das componentes do vetor pela soma $x + y + z$ para obter uma distribuição. Temos que resolver então o sistema

$$\begin{aligned} -5x + 2y + 2z &= 0 \\ 2x - 3y + 4z &= 0 \\ 3x + y - 6z &= 0 \end{aligned}$$

Uma das equações é redundante, então podemos por exemplo, desconsiderar a terceira equação e resolver o sistema formado pelas duas primeiras. Este sistema possui infinitas soluções e para calcular uma delas basta fixar um valor (não nulo) para uma das incógnitas e resolver as duas equações resultantes. Se fixarmos $x = 42$, as equações que ficaram são

$$\begin{aligned} 2y + 2z &= 210 \\ -3y + 4z &= -84 \end{aligned}$$

cuja única solução é $y = 72$, $z = 33$ e obtemos

$$\begin{aligned} \pi(1) &= \frac{x}{x + y + z} = \frac{14}{49}, \\ \pi(2) &= \frac{y}{x + y + z} = \frac{24}{49}, \\ \pi(3) &= \frac{z}{x + y + z} = \frac{11}{49}, \end{aligned}$$

que é a única distribuição estacionária desta cadeia.

EXEMPLO 5.4. Vimos que o gerador infinitesimal da cadeia de Markov uniforme (exemplo 1.5) era

$$A = \lambda(K - I).$$

Uma distribuição π sobre E será estacionária para X se e somente se ela satisfizer (5.30), ou seja, se para todo $j \in E$ vale

$$\sum_{l \in E} \pi(l)k_{lj} = \pi(j).$$

Obtemos então que toda distribuição estacionária de X o será para \hat{Y} e viceversa, toda distribuição estacionária de \hat{Y} será também distribuição estacionária para X .

Como no caso discreto, o comportamento assintótico das probabilidades de transição de cadeias a tempo contínuo estará relacionado com a classificação dos estados.

PROPOSIÇÃO 5.5. *Seja j um estado transitório. Então para todo $i \in E$ tem-se,*

$$\lim_{t \rightarrow \infty} P_i(X_t = j) = 0.$$

TEOREMA 5.6. *Seja X uma cadeia irredutível. Então existe*

$$\lim_{t \rightarrow \infty} P_i(X_t = j) = \pi(j)$$

e não depende de i . As componentes de π são todas nulas se a cadeia for transitória ou recorrente nula. Caso contrário π é a única distribuição estacionária desta cadeia e vale

$$\pi(j) = \frac{1}{q_j \mathbb{E}_j(\tau_j)}.$$

Repare que agora na distribuição limite aparece também a média no tempo que a cadeia fica no estado j , que é o inverso da taxa q_j .

Existe uma relação entre a distribuição limite da cadeia e aquela do seu esqueleto.

PROPOSIÇÃO 5.7. *Suponha que o esqueleto $\{\hat{X}_n\}_{n \geq 0}$ da cadeia X é recorrente positivo e seja π_Q a sua distribuição limite. Então X será recorrente positivo se e somente se $\sum_{k \in E} \pi_Q(k)/q_k < \infty$ e sendo esse o caso, a sua distribuição limite π será*

$$\pi(j) = \frac{\pi_Q(j)/q_j}{\sum_{k \in E} \pi_Q(k)/q_k}.$$

6. Aplicações.

Processos de Nascimento e Morte.

Uma cadeia de Markov a tempo contínuo com espaço de estados $E = \{0, 1, \dots\}$ e gerador infinitesimal

$$A = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & 0 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 & 0 & \dots \\ 0 & 0 & \mu_3 & -(\lambda_3 + \mu_3) & \lambda_3 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}.$$

é chamada de **processo a tempo contínuo de nascimento e morte**.

O estado no processo no instante t pode ser interpretado como a quantidade de habitantes de uma população dada. Se $X_t = i$, a próxima transição será para $i+1$ (nascimento) com taxa λ_i ou para $i-1$ (morte) com taxa μ_i . Estas taxas são chamadas taxas de nascimento e de morte, respectivamente. Observe que um estado i será absorvente se e somente se ambas as taxas forem nulas.

As equações de Kolmogorov prospectivas tomam a forma,

$$\begin{aligned} P'_{i,0}(t) &= \mu_1 P_{i,1}(t) - \lambda_0 P_{i,0}(t) \\ P'_{i,j}(t) &= \lambda_{j-1} P_{i,j-1}(t) + \mu_{j+1} P_{i,j+1}(t) - (\lambda_j + \mu_j) P_{i,j}(t) \end{aligned}$$

Usando (9.15) na matriz de transição do esqueleto desta cadeia obteremos que o processo contínuo de nascimento e morte será recorrente se e somente se

$$\sum_{k=1}^{+\infty} \frac{\mu_1 \mu_2 \cdots \mu_k}{\lambda_1 \lambda_2 \cdots \lambda_k} = \infty.$$

Em tal caso, a distribuição estacionária existe se e somente se

$$\sum_{k=1}^{+\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k} = \infty,$$

ela será também distribuição limite e tem a forma

$$\pi(j) = \begin{cases} \frac{1}{c}, & j = 0 \\ \frac{\lambda_0 \lambda_1 \cdots \lambda_{j-1}}{c \mu_1 \mu_2 \cdots \mu_j}, & j \geq 1 \end{cases},$$

onde

$$c = 1 + \sum_{k=1}^{+\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k}.$$

Bibliografia

- [1] E. Çinlar, *Introduction to Stochastic Processes*, Prentice Hall, New Jersey (1975).
- [2] R. Durrett, *Elementary Probability for Applications*, Capítulo 5, <http://www.math.cornell.edu/~durrett/ep4a/ep4a.html>
- [3] S. Karlin, H.M. Taylor *A First Course in Stochastic Processes*, Academic Press, New York, (1975)
- [4] J. G. Kemeny, J.L. Snell, *Finite Markov Chains*, Springer, New York, (1975).
- [5] P. Brémaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulation and Queues*, Springer (2001).
- [6] S. Resnick, *adventures in Stochastic Processes*, Birkhäuser (1992).
- [7] S.M. Ross, *Stochastic Processes*, John Wiley & Sons (1983).
- [8] S.M. Ross, *Introduction to Probability Models*, Academic Press (1997).
- [9] P.G. Hoel, S.C. Port, C.J. Stone, *Introduction to Stochastic Processes*, Boston (1972).