UMA INTRODUÇÃO AOS PROCESSOS ESTOCÁSTICOS COM APLICAÇÕES

Adrian Hinojosa e Aniura Milanés

Departamento de Estatística
ICEx. UFMG.

Sumário

CAPíTULO 1

Introdução

1. Definições e exemplos

Os processos estocásticos representam sistemas nos quais o estado muda ao longo do tempo. Estas mudanças não são totalmente previssíveis, mas elas estão associadas a distribuições de probabilidade. Diversos fenômenos reais admitem a modelagem através dos processos estocásticos. Vejamos alguns exemplos.

Exemplo 1.1 (Cadeia de montagem). *Os produtos finais de uma cadeia de montagem, após uma supervisão à que são submetidos, podem ser considerados defeituosos ou não. Se o n-ésimo produto não tiver defeito, fazemos* $X_n = 1$, *caso contrário* $X_n = 0$. *Suponha que um produto é defeituoso independentemente dos outros produtos e que a probabilidade de que isto aconteça é* p, *então* $X_1, X_2, \ldots, X_n, \ldots$ *é uma sequência de variáveis aleatórias independentes Bernoulli com parâmetro* p *de sucesso.*

Exemplo 1.2 (Estoque). *Uma pequena loja de equipamentos eletrodomésticos vende um 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 antesd e começar o expediente. Vamos chamar de* X_n *à 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). *Consideremos a história de várias gerações de uma família que ao longo do tempo tem somente um filho. Neste modelo simples, a observação da classe social (alta, média ou baixa) da família para cada geração permitiria descrever sua evolução social ao longo do tempo.*

Se tivermos uma sociedade composta por famílias deste tipo, podemos escolher ao acaso uma família e para cada geração n *chamar de* X_n *à uma quantidade que valerá* 1 *se a família for de classe alta,* 2 *se ela for de classe média e* 3 *se for de classe baixa. Desta forma, cada* X_n *será uma variável aleatória e a sua evolução ao longo do tempo, permitirá tirar conclusões sobre as mudanças na estrutura da sociedade.*

Exemplo 1.4 (Lucro de uma companhia seguradora). *Suponha que uma seguradora recebe* c *unidades monetárias (u.m.) pelo total dos prêmios que ela cobra dos segurados* *dentro de uma determinada carteira por período de tempo (mês, semestre, por exemplo). Assuma também que a seguradora coleta os prêmios regularmente e que as indenizações são pagas quando os sinistros ocorrem. Além disto, não vamos considerar aqui eventuais despesas administrativas, ganhos ou perdas por investimentos, etcétera. Desta forma, a reserva desta seguradora será afetada somente pela cobrança dos prêmios ou por pagamentos de indenizações na ocorrência de sinistros. Em particular, o lucro da companhia no n-ésimo período será $c - Z_n$ u.m., sendo Z_n o valor total de indenizações pago pela seguradora nesse período. Se chamarmos de L_n ao lucro da seguradora desde que essa carteira começa a operar até o final do n-ésimo período, teremos que*

$$L_n = cn - \sum_{j=1}^{n} Z_j.$$

O comportamento desta quantidade ao longo do tempo influenciará significativamente a saúde financeira da seguradora.

Em todas as situações acima, as magnitudes de interesse são famílias de variáveis aleatórias.

Chamaremos de **processo estocástico** a qualquer família de variáveis aleatórias X_t, com $t \in \mathbb{T}$ e sendo $\mathbb{T}$ algum espaço de parâmetros [1]. .

Na maioria das situações reais, o espaço de parâmetros representa o tempo, mas isto nem sempre é assim (veja o exemplo 1.1). Observe que em todos nos exemplos acima, vale $\mathbb{T} = \mathbb{N}$. No entanto, muitos processos estocásticos importantes têm como espaço de parâmetros $\mathbb{T} = \mathbb{R}$ ou $\mathbb{T} = [0, +\infty)$. Ainda neste capítulo veremos alguns exemplos. Quando $\mathbb{T}$ é enumerável diremos que o processo estocástico correspondente é **a tempo discreto**. Se $\mathbb{T}$ for um intervalo, o processo estocástico será chamado **a tempo contínuo**. Nos exemplos de 1.1 até 1.4, todos os processos são a tempo discreto.

Os valores que tomam as variáveis do processo serão chamados de **estados** e o conjunto E destes valores será o **espaço de estados**. Observe que os estados não precisam ser quantidades numéricas.

Os processos estocásticos podem ter espaço de **estados discreto** ou **espaço de estados contínuo** em correspondência com a natureza do conjunto E. No exemplo 1.1 temos que que $E = \{0, 1\}$, no exemplo 1.2, $E = \{0, 1, 2, 3, 4, 5\}$, no exemplo 1.3, $E = \{1, 2, 3\}$ e finalmente no exemplo 1.4, $E = \mathbb{R}$. Os espaços de estados nos exemplos 1.1-1.3 são todos discretos enquanto o espaço de estados do exemplo 1.4 é contínuo.

No caso de um processo estocástico, uma observação será uma coleção de valores no espaço de estados da forma $\{x_t : t \in \mathbb{T}\}$ que é chamada de **trajetória** ou **realização** deste processo. Para o processo do exemplo 1.1, uma possível trajetória seria $\{1, 0, 0, 0, 0, 0, 0, \dots\}$. Ela corresponde ao caso em que o primeiro produto produzido não teve defeito e todos os demais foram defeituosos. No exemplo 1.3 a trajetória

[1] Uma definição rigorosa dos processos estocásticos pode ser consultada na seção 3

$\{1,3,1,3,1,3,1,3,\dots\}$ corresponde à escolha de uma família que alterna entre as classes alta e baixa ao longo das gerações.

Vejamos agora um exemplo de um processo estocástico a tempo contínuo.

EXEMPLO 1.5 (O processo de Poisson). *Suponha que num laboratório é possível contar a quantidade de partículas emitidas por uma certa substância radioativa a partir de um instante de tempo determinado. Suponha também que os intervalos de tempo decorrido entre duas emissões sucessivas de partículas formem uma sequência de variáveis aleatórias independentes e com distribuição exponencial com parâmetro* λ *(*$\lambda > 0$*). Se chamarmos de* N_t *à quantidade de partículas emitidas até o instante* t*, então o processo a tempo contínuo* $N = \{N_t\}_{t\geq 0}$ *será chamado de processo de Poisson. Este processo tem espaço de estados* $E = \mathbb{N}$ *e cada variável* N_t *tem distribuição* $Poisson[\lambda(t)]$*. As trajetórias deste processo são como na figura (1). Elas têm saltos de tamanho um.*

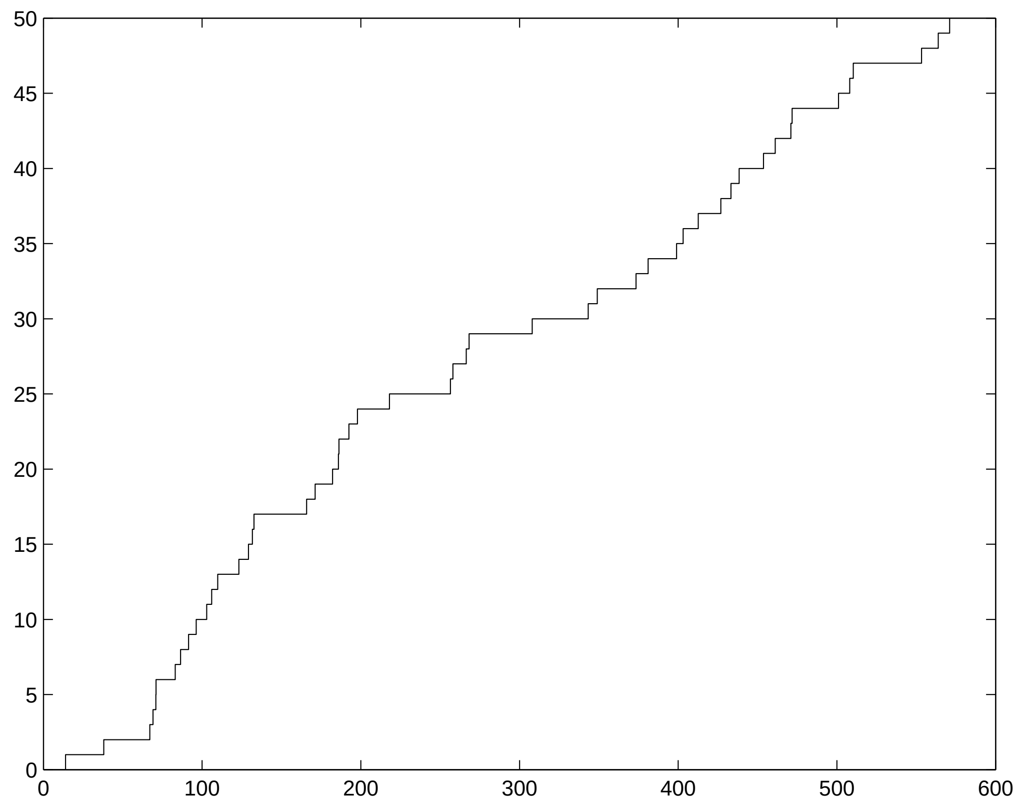
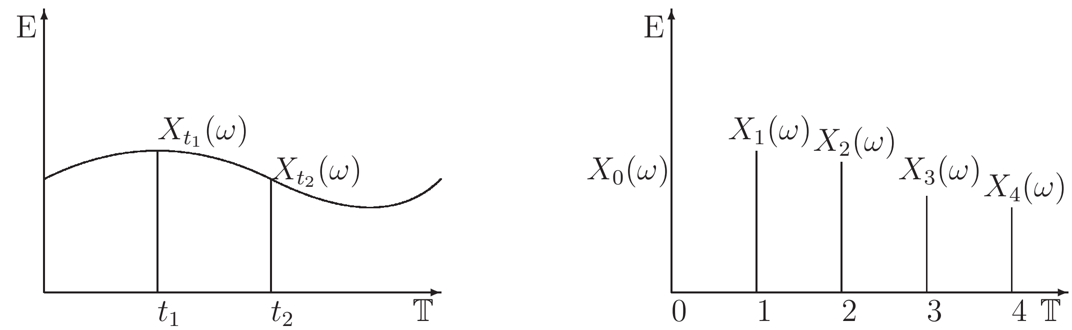


FIGURA 1. Trajetória do processo de Poisson

Este processo será estudado em detalhe no capítulo 4.

Dependendo do espaço de parâmetros, um processo estocástico pode ter vários tipos de trajetórias que são ilustradas graficamente nas figuras (2(a)) e (2(b)).



(a) Trajetória contínua (b) Trajetória Discreta

FIGURA 2.

Resumindo, temos as seguintes possibilidades para os processos estocásticos.

PROCESSOS ESTOCÁSTICOS

	E enumerável	E não enumerável
$\mathbb{T}$ enumerável	Processo a tempo discreto com espaço de estados discreto (Exemplos 1.1-1.3)	Processo a tempo discreto com espaço de estados contínuo (Exemplo 1.4)
$\mathbb{T}$ intervalo	Processo a tempo contínuo com espaço de estados discreto (Exemplo 1.5)	Processo a tempo contínuo com espaço de estados contínuo (Exemplo 1.10)

TABELA 1. Classificação dos processos estocásticos

Observe que podemos enxergar os processos estocásticos como generalizações de variáveis e vetores aleatórios. De fato, se a cardinalidade do espaço de parâmetros $\mathbb{T}$ for finita, o processo correspondente será um vetor aleatório. Por exemplo, se $\mathbb{T} = \{1, 2\}$, o processo correspondente pode ser representado como (X_1, X_2), que é um vetor aleatório bivariado. Em particular, se $\mathbb{T}$ contiver somente um elemento, o processo será uma variável aleatória.

Sabemos que 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, pois as distribuições marginais das coordenadas não determinam a distribuição do vetor no sentido que existem vetores aleatórios com distribuições diferentes e com as mesmas marginais como ilustramos a seguir.

Exemplo 1.6. *Consideremos os vetores aleatórios discretos* (X_1, X_2) *e* (Y_1, Y_2) *com funções de probabilidade conjunta*

$X_1\backslash X_2$	0	1
0	1/4	1/4
1	1/4	1/4

e

$Y_1\backslash Y_2$	0	1
0	0	1/2
1	1/2	0

respectivamente. Estas funções de probabilidade são diferentes, no entanto, as marginais coincidem pois

$$P(X_1 = 0) = P(Y_1 = 0) = \frac{1}{2} = P(X_2 = 0) = P(Y_2 = 0).$$

Parece bastante natural pensar que para processos estocásticos, as distribuições de cada uma das variáveis aleatórias que o formam (marginais do processo) não devem ser suficientes para determinar o seu comportamento probabilístico. Com o seguinte exemplo pretendemos esclarecer melhor esta idéia.

Exemplo 1.7. *No exemplo 1.1 tratamos com um processo formado por uma sequência de variáveis aleatórias independentes, todas elas com distribuição de Bernoulli com parâmetro* p*. Tais processos são chamados de processos de Bernoulli de parâmetro* p *(veja a seção 2).*

Seja Z *um processo de Bernoulli de parâmetro* $\sqrt{p}$*. De acordo com a definição isto quer dizer que as variáveis* $Z_1, Z_2, \ldots$ *são independentes e todas elas seguem a distribuição de Benoulli(*$\sqrt{p}$*). Vamos usar agora o processo* Z *para definir um outro processo* Y *da forma seguinte. Faremos* $Y_1 = Z_1Z_2, Y_2 = Z_2Z_3, Y_3 = Z_1Z_3, Y_4 = Z_4Z_5, Y_5 = Z_5Z_6, Y_6 = Z_4Z_6, \ldots,$ *e em geral*

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

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

Determinemos a distribuição de cada variável Y_j*. De fato, pela sua própria definição cada* Y_j *toma somente os valores* 0 *e* 1 *e podemos calcular, por exemplo*

$$\begin{aligned} P(Y_1 = 1) &= P(Z_1 = 1, Z_2 = 1), \\ &= P(Z_1 = 1)P(Z_2 = 1)(\text{ pela independência}), \\ &= (\sqrt{p})^2 = p. \end{aligned}$$

Por meio de cálculos análogos pode-se provar que $P(Y_j = 1) = p$ *para* $j \in \mathbb{N}$*, por tanto, cada variável* Y_j *tem distribuição Bernoulli(*p*). Será isto suficiente para garantir que* Y *é um processo de Bernoulli? Intuitivamente devemos esperar que não, pois pela* *maneira que foi definido este processo, os valores das variáveis Y_{3n-1} e Y_{3n}, por exemplo, sempre estarão relacionados. Em particular,*

$$\begin{aligned} P(Y_{3n-1}=1, Y_{3n}=0) &= P(Z_{3n-2}=0, Z_{3n-1}=1, Z_{3n}=1) \\ &= p(1-\sqrt{p})) \end{aligned}$$

e portanto (se $p<1$), $p(1-p) = P(Y_{3n-1}=1)P(Y_{3n}=0) \neq P(Y_{3n-1}=1, Y_{3n}=0)$ e Y_{3n-1} e Y_{3n} não são independentes.

Isto quer dizer que Y ***não é um processo de Bernoulli pois as variáveis aleatórias que formam o processo não são independentes.***

A maneira de conclusão, podemos dizer que:

O comportamento probabilístico de um processo estocástico está caracterizado não somente pelas distribuições marginais das variáveis coordenadas, mas também pelas relações de dependência entre elas.

Entre os estados de um processo X podemos encontrar diferentes tipos de relações de dependência. Neste contexto é interessante lembrar da interpretação do espaço de parâmetros $\mathbb{T}$ como tempo.

a) *Estados Independentes*
A relação de dependência entre variáveis aleatórias mais simples que podemos pensar seria a ausência total dela. Chamaremos de **processo de estados independentes** a aquele processo estocástico 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*
Consideremos os instantes $t_1, t_2, \ldots, t_n, t \in \mathbb{T}$, com $t_1 < t_2 < \cdots < t_n < t$. Um processo X é chamado de **processo de Markov** quando para todos $a, b, a_1, \ldots, a_n \in E$, vale

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

ou seja o estado X_t do processo depende da sua história anterior nos instantes $t_1, t_2, \ldots, t_n$ somente através do **presente** X_{t_n} e não do **passado** $X_{t_1}, X_{t_2}, \ldots, X_{t_{n-1}}$. Os processos de estados independentes são exemplos muito simples de processos de Markov. Todos os processos que estudaremos a partir do próximo capítulo serão processos de Markov.

d) *Martingais*
X será chamado de **martingal** quando para todos os instantes $t_1, t_2, \ldots, t_n, t$ com $t_1 < t_2 < \cdots < t_n < t_{n+1}$ e $a_1, a_2, \ldots, a_n \in E$ tivermos

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

Em outras palavras, poderiamos 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 1.8. *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 subtraimos 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

$$Z_j = \begin{cases} A, & \textit{se sair cara no j-ésimo lançamento,} \\ -A, & \textit{se sair coroa no j-ésimo lançamento,} \end{cases} \tag{1.1}$$

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

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

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

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

A seguinte definição será útil no que segue.

Definição 1.9. Se $\{X_t\}_{t \in \mathbb{T}}$ é um processo estocástico, chamaremos de **incremento** correspondente ao intervalo $(s, t]$ à variável aleatória $X_t - X_s$, $t > s$, $s, t \in \mathbb{T}$.

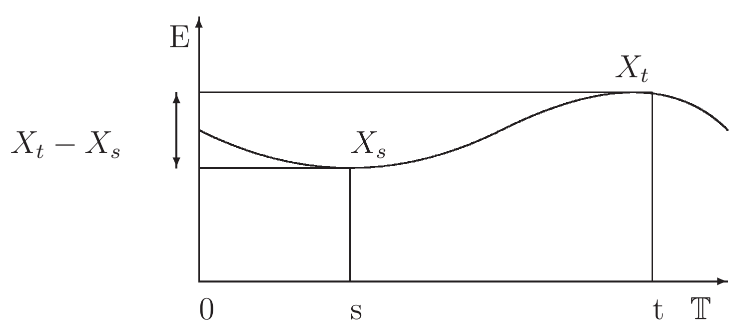


Figura 3. Incremento do processo $\{X_t\}_{t \in \mathbb{T}}$.

Diremos que o processo tem **incrementos estacionários** quando a distribuição de $X_t - X_s$ depende dos instantes s e t somente através da sua diferença $t-s$. Se para todos $t_0 < t_1 < \cdots < t_n$, 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 o processo tem incrementos independentes. Em outras palavras, um processo terá incrementos independentes quando os incrementos correspondentes a intervalos disjuntos sejam independentes.

No capítulo 4 veremos que o processo de Poisson, apresentado no exemplo 1.5, tem incrementos estacionários e independentes.

Terminaremos esta seção com um exemplo de um processo estocástico muito importante.

Exemplo 1.10. 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:*

(i) $X_0 = 0$;
(ii) X *tem incrementos independentes;*
(iii)
$$P(X_t - X_s \leq x) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{x} e^{-\frac{u^2}{2(t-s)}} du,$$
i.e. $X_t - X_s \sim N(0, t-s)$;
(iv) X *possui trajetórias contínuas.*

X *é conhecido como movimento Browniano ou processo de Wiener.*

A figura 4 sugere um comportamento bastante irregular das trajetórias do processo de Wiener. Observe que este processo tem incrementos estacionários e independentes.

Um fato muito interessante é que as condições (i)-(iv) acima são suficientes para caracterizar de forma única a lei do processo. Em particular, usando estas propriedades, podemos calcular as chamadas ***distribuições finito-dimensionais*** *do processo, que são simplesmente as distribuições conjuntas de todas as famílias finitas de estados.*

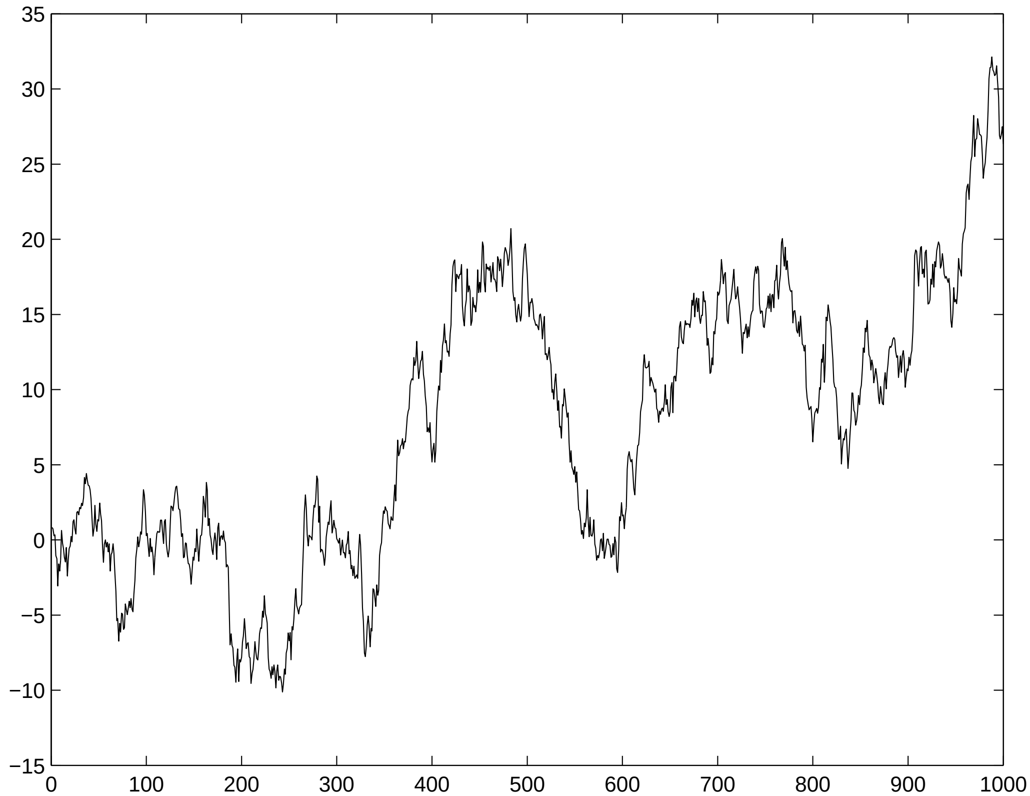


FIGURA 4. Trajetória do movimento Browniano

Vejamos por exemplo, como calcular a conjunta de X_s e X_t para dois instantes fixados s e t tais que $0 \leq s < t$. Se fizermos

$$\begin{cases} Y = X_s - X_0 = X_s, \\ Z = X_t - X_s, \end{cases}$$

usando a independência e a estacionariedade dos incrementos, teremos que

$$\begin{aligned} f_{Y,Z}(y,z) &= f_Y(y) f_Z(z), & (1.2) \\ &= \left(\frac{1}{\sqrt{2\pi s}} \exp\left\{ -\frac{y^2}{2s} \right\} \right) \left(\frac{1}{\sqrt{2\pi(t-s)}} \exp\left\{ -\frac{z^2}{2(t-s)} \right\} \right), & (1.3) \\ &= \frac{1}{2\pi\sqrt{s(t-s)}} \exp\left\{ -\frac{1}{2} \left(\frac{y^2}{s} + \frac{z^2}{t-s} \right) \right\}. & (1.4) \end{aligned}$$

A relação entre (X_s, X_t) e (Y, Z) é muito simples, pois

$$\begin{cases} X_s = Y, \\ X_t = Y + Z. \end{cases}$$

Podemos usar o método do jacobiano para calcular a densidade desejada. Observando que o jacobiano da transformação acima vale 1*, teremos*

$$f_{X_s,X_t}(v,w) = f_{Y,Z}(v, w-v).$$

Colocando (1.4) na expressão acima e fazendo alguns cálculos obteremos sem muita dificuldade

$$f_{X_s,X_t}(v,w) = \frac{1}{2\pi\sqrt{s(t-s)}} \exp\left\{-\frac{1}{2}\left(\frac{tv^2 - 2svw + sw^2}{s(t-s)}\right)\right\},$$

ou seja, o vetor (X_s, X_t) *segue a distribuição normal bivariada com vetor de médias nulo e* $Cov(X_s, X_t) = s$.

2. O Processo de Bernoulli e outros processos estocásticos associados

Se tivermos que "desenhar"um processo estocástico "interessante"e o mais elementar possível, podemos pensar num processo de estados independentes, a tempo discreto e com espaço de estados finito. Assumindo que todos os estados tem a mesma distribuição e escolhendo a distribuição de Bernoulli, obtemos o chamado processo de Bernoulli (que foi introduzido já no exemplo 1.1). Nesta seção tentaremos ilustrar algumas questões de interesse sobre os processos estocásticos atravês do estudo do processo de Bernoulli e de outros processos estocásticos a ele relacionados.

Definição 2.1. O processo a tempo discreto $\{X_n\}_{n\geq 1}$ é chamado processo de Bernoulli, com probabilidade de sucesso p se

(i) $X_1, X_2, \ldots$ são variáveis aleatórias independentes e
(ii) $\forall n \geq 1$ $P(X_n = 1) = p, \quad P(X_n = 0) = 1 - p = q.$

Uma trajetória do processo de Bernoulli será uma sequência de sucessos e fracassos obtidos em ensaios de Bernoulli consecutivos e independentes e com probabilidade p de sucesso.

Proposição 2.2. *Seja* $\{X_n\}_{n\geq 1}$ *um processo de Bernoulli, então:*

(i) $\mathbb{E}X_n = p$
(ii) $VarX_n = pq$
(iii) $\mathbb{E}\alpha^{X_n} = q + \alpha p$, $\alpha \neq 0$.

Demonstração. Exercício. □

Vários outros processos estocásticos aparecem associados ao processo de Bernoulli, de forma bastante natural. Consideremos mais uma vez o exemplo 1.1. Poderiamos contar entre os n primeiros produtos supervisionados, quantos deles não apresentaram defeito. Se chamarmos de N_n a esta nova quantidade, a coleção de variáveis aleatórias $\{N_n\}_{n\geq 0}$ representa um novo processo estocástico.

DEFINIÇÃO 2.3. Considere um processo de Bernoulli $\{X_n\}_{n\geq 1}$ com probabilidade p de sucesso e defina

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

N_n é o número de sucessos nos n primeiros ensaios do processo de Bernoulli. $\{N_n\}_{n\geq 0}$ é um processo estocástico que será chamado de **processo do número de sucessos no processo de Bernoulli**.

Observe que para $n > 0$ temos $N_n \sim b(n,p)$, pois $N_n = X_1 + \cdots + X_n$. Alem disso, o processo do número de sucessos no processo de Bernoulli é crescente, ou seja toda trajetória satisfaz $0 = N_0 \leq N_1 \leq N_2 \leq N_3 \leq \ldots$ e para $k \geq 1$ $N_k - N_{k-1} \leq 1$.

Teremos então que

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

Repare que podemos interpretar a propriedade acima como "probabilidades de transição"do processo $\{N_n\}_{n\geq 0}$ do estado j para o estado k.

Como $N_{n+m} - N_n$ é o número de sucessos nos *ensaios* independentes $n+1$, $n+2$, $\ldots, n+m$, então $N_{n+m} - N_n \sim b(n,p)$ e $P(N_{n+m} - N_n = k) = \binom{m}{k} p^k (1-p)^{m-k}$. Esta propriedade junto com o seguinte resultado, permitem concluir que o processo $\{N_n\}_{n\geq 0}$ tem incrementos estacionários e independentes.

PROPOSIÇÃO 2.4.

(i) Para todo n,
$P(N_{n+m} - N_n = k | N_0, \ldots, N_m) = P(N_{n+m} - N_n = k) = \binom{m}{k} p^k (1-p)^{m-k}$.

(ii) Para todos $n_0 < n_1 < n_2 < \cdots < n_{m-1} < n_m$ vale que os incrementos
$N_{n_1} - N_{n_0}, N_{n_2} - N_{n_1}, \ldots, N_{n_m} - N_{n_{m-1}}$ *são independentes.*

Vejamos agora como utilizar nos cálculos as propriedades enunciadas.

EXEMPLO 2.5. *Calcule $P(N_5 = 4, N_7 = 5, N_{13} = 8)$.*

Solução:

A princípio, para calcular a probabilidade pedida, precisariamos da função de probabilidade conjunta das variáveis N_5, N_7 e N_{13}, que não conhecemos, no entanto, conhecemos a das variáveis $N_5, N_7 - N_5$ e $N_{13} - N_7$ pela proposição 2.4. Podemos utilizar o fato dos eventos $A = [N_5 = 4, N_7 = 5, N_{13} = 8]$ e $B = [N_5 = 4, N_7 - N_5 = 1, N_{13} - N_7 = 3]$ serem *iguais, obtendo*

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

Este tipo de artifício sera usado repetidas vezes neste texto.

Exemplo 2.6. *Calcule* $\mathbb{E}(N_5N_8)$.
Solução:

Mais uma vez, podemos introduzir a variável incremento $N_8 - N_5$. *Escrevendo* $N_8 = N_5 + (N_8 - N_5)$ *e usando que a esperança do produto de variáveis independentes fatora no produto das esperanças, obtemos*

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

Exemplo 2.7. *Calcule* $\mathbb{E}(N_{11}|N_5)$.
Solução: *Escrevendo* $N_{11} = N_5 + (N_{11} - N_5)$ *e utilizando a proposição 2.4, obtemos*

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

Como vimos acima, o processo $\{N_n\}_{n\geq 0}$ tem incrementos independentes e estacionários. Podemos provar ainda que ele satisfaz a propriedade de Markov.

Teorema 2.8 (Propriedade de Markov para o número de sucessos).

(i) $\mathbb{E}\left[N_{n+1}|N_0, \ldots, N_n\right] = \mathbb{E}\left[N_{n+1}|N_n\right]$.
(ii) $P(N_{n+1} = k|N_0 = i_1, \ldots, N_n = i_n) = (N_{n+1} = k|N_n = i_n)$.

Demonstração. Exercício. □

Exemplo 2.9. *Calcule* $\mathbb{E}(N_5N_8)$.
Solução: *Raciocinando como nos exemplos acima e utilizando (1) no teorema anterior,*

podemos escrever

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

EXEMPLO 2.10. *Calcule* $\mathbb{E}\left[N_{11}N_5|N_2,N_3\right]$.
Solução: *Usando propriedades da esperança condicional, podemos escrever*

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

Voltando agora ao nosso exemplo 1.1, suponha que registramos o número de todos os produtos supervisionados e que não foram defeituosos. Por exemplo, se nos primeiros cinco produtos observados tivemos que somente o primeiro e o terceiro foram defeituosos, teriamos que $X_1=0, X_2=1, X_3=0, X_4=1, X_5=1$. Então, os produtos não defeituosos foram aqueles numerados com os índices $T_1=2, T_2=4$ e $T_3=5$. Graficamente, podemos representar esta trajetória da forma seguinte.

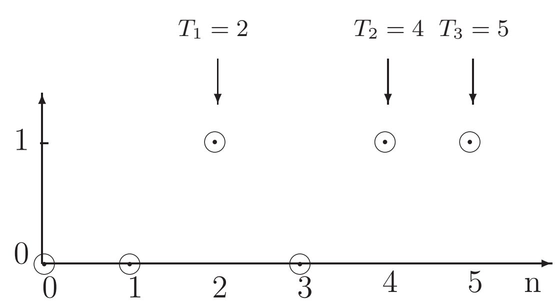


FIGURA 5. 2, 4 e 5 são os produtos sem defeito observados.

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

Considere um processo de Bernoulli $\{X_n\}_{n\geq 0}$, com probabilidade p de sucesso. Chamaremos de $T_1, T_2, \ldots$ aos instantes nos quais acontecem os sucessos consecutivos. Fica definida desta forma uma sequência de variáveis aleatórias formando o processo estocástico $\{T_n\}_{n\geq 1}$ chamado de **processo dos instantes dos sucessos ou processo das chegadas**.

Exemplo 2.11. *Cada segundo do tempo observamos se nesse instante passa um carro num lugar fixo de um caminho. Suponha que os carros passan independentemente uns dos outros. Este experimento corresponde a um processo de Bernoulli $\{X_n\}_{n\geq 1}$, com $X_n = 1$ se 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, $X_{T_k} = 1$, pois T_k é o instante de ocorrência do k-ésimo sucesso. Isto é ilustrado no gráfico da figura (6).

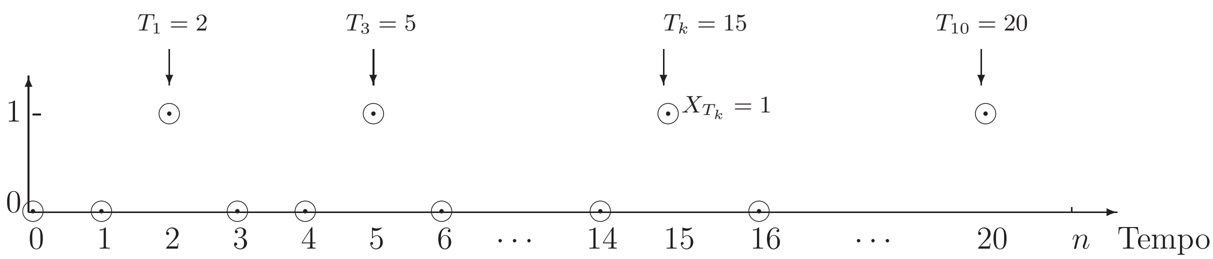


Figura 6. Instantes dos sucessos no processo de Bernoulli.

A partir da figura vemos também que

Proposição 2.12.
(i) $T_k \leq n$ *se e somente se* $N_n \geq k$.
(ii) $T_k = n$ *se e somente se* $N_{n-1} = k-1, X_n = 1$.

Demonstração.
(i) De fato, $T_k \leq n$ significa que o k-ésimo sucesso ocorre no máximo no instante n. Ou seja, o número de sucessos até o instante n (igual a N_n) deve ser no mínimo igual a k.

(ii) Teremos qu n é o instante do k-ésimo sucesso se
  (a) há um sucesso no instante n, ou seja, $X_n = 1$ e;
  (b) há $k-1$ sucessos até o instante $n-1$, ou seja $N_{n-1} = k-1$.

□

Usando esta proposição podemos provar o seguinte teorema.

Teorema 2.13. *Para todo $k \in \mathbb{N}$,* *(i)* $P(T_k \leq n) = \sum_{j=k}^{n} \binom{n}{k} p^k (1-p)^{n-k}$,

(ii) $P(T_k = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}$.

Demonstração.

(i) Usando (i) na proposição 2.12 e o fato de $N_n \sim bin(n,p)$.

(ii) Usando (ii) na proposição 2.12 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 mostrar que o processo das chegadas $\{T_k\}_{k\geq 1}$ é Markoviano. Vejamos por que 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 (veja a figura 6)

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

Ou seja, 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 $X_{12}, \ldots, X_{16}$ que são eventos que acontecem depois do tempo T_4).

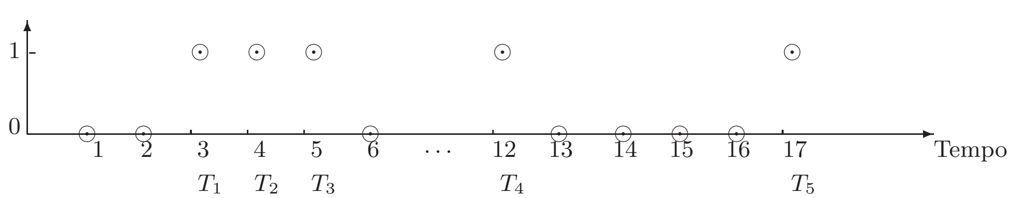


Figura 7.

Provemos, então que

Teorema 2.14 (ropriedade de Markov para os instantes dos sucessos). *Para todo* $k \in \mathbb{N}$, *e para todo* $n \geq k$ *vale*

$$P\left(T_{k+1} = n | T_0, \ldots, T_k\right) = P\left(T_{k+1} = n | T_k\right).$$

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

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

□

Decorre da prova anterior que

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

logo,

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

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

Mas isto significa que $T_{k+1} - T_k \sim$ 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.s com distribuição Geométrica(p). Isto é $T_k \sim$ Binomial Negativa(k, p), logo:

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

$$\mathbb{E}\left[T_k\right] = \frac{k}{p} \quad \text{e} \quad \operatorname{Var}\left[T_k\right] = k\frac{1-p}{p^2}$$

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

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

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 chamarmos de T_k aos instantes nos quais acontecem as falhas, então os tempos entre falhas consecutivas serão $U_k = T_k - T_{k-1}$ e $P(U_k = m) = p(1-p)^{m-1}, m \geq 1$. Estes instantes $T_k, k \geq 1$ serão os instantes dos sucessos de um processo de Bernoulli.

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

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

Pela propiedade de Markov,

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

Observe que

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

Portanto, a nossa estimativa seria

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

3. Alguma observações mais rigorosas sobre os processos estocásticos*

DEFINIÇÃO 3.1. 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.

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.

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

A família $\{F_{t_1,t_2,\ldots,t_n}\}$ 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,

$$\lim_{x_k \uparrow \infty} F_{t_1,t_2,\ldots,t_n}(x_1, \ldots, x_n) = F_{t_1,\ldots,t_{k-1},t_{k+1},\ldots,t_n}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n). \tag{3.5}$$

A todo processo estocástico corresponde uma família de funções de distribuição satisfazendo (3.5).

NOTA 3.3. 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

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

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.,

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

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

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

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 distribuidos) 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 necessáriamente tem que ocorrer.

4. Exercícios

1. Num cruzamento em T, aproximadamente 60% dos carros viram à esquerda. Defina X_n como sendo 1 ou 0 em dependendo se o n-ésimo carro virou à esquerda ou à direita. Suponha que os motoristas decidem para onde virar independentemente um do outro. Então $X = \{X_n, n \in \mathbb{N}\}$ é um processo de Bernouli com probabilidade 0,6 de sucesso. Num certo dia um pedestre observou o que faziam 10 carros que passaram consecutivamente e fez a anotação $(E, D, E, E, E, D, E, D, E, D)$, onde E=esquerda e D=direita. Quais seriam os valores correspondentes de

(a) $X_1, X_2, \ldots, X_{10}$
(b) $N_1, N_2, \ldots, N_{10}$
(c) $T_1, X_2, \ldots, T_6$?

2. Para um processo de Bernoulli com $p = 0,7$ interprete e calcule as seguintes quantidades

(a) $P(N_1 = 0, N_2 = 0, N_3 = 1, N_4 = 1)$;
(b) $P(T_1 = 2, T_2 = 3, T_3 = 5)$;
(c) $P(T_4 - T_3 = 12)$;

3. Para o processo de Wiener, calcule a função de densidade conjunta de n estados. Sugestão: Generalize o resultado utilizado no exemplo 1.10.

CAPíTULO 2

Cadeias de Markov a Tempo Discreto

1. Introdução

EXEMPLO 1.1. ***Jogo de azar***.

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 cassino(??) não deixa você jogar mais.

Seja X_n a quantidade de dinheiro que você tem depois de n apostas. Observe que são esta quantidade e o resultado do próximo sorteio que vão determinar a sua fortuna depois da aposta seguinte. Qualquer que tenha sido a evolução da sua fortuna no "passado"(ou seja, os valores de $X_{n-1}, X_{n-2}, \ldots, X_0$), para prever o próximo estado X_{n+1}, é suficiente conhecer a sua fortuna no "presente"(X_n). De fato, se $X_n = i$, com $0 < i < N$, então independentemente dos valores $i_0, \ldots, i_{n-1}$, teremos que

$$P(X_{n+1} = i+1 | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = 0,4, \tag{1.6}$$

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

Em cada aposta a sua fortuna somente poderá aumentar ou diminuir em um real com uma chance que não depende do número de apostas que você fez. Em outras palavras, a probabilidade condicional em (1.6) não depende de n.

Fixemos, por exemplo, o valor $N = 5$. Então os valores que pode tomar a sua fortuna são $\{0,\ 1,\ 2,\ 3,\ 4,\ 5\}$. Suponha, por exemplo, que depois de certa quantidade de apostas, você tem 2 reais. Podemos considerar as probabilidades de sua fortuna na próxima aposta, tomar algum destes valores possíveis. Como já foi observado, depois de uma aposta, você terá ou 1 real ou 3 reais, dependendo da sua sorte. As probabilidades mencionadas podem ser arranjadas em un vetor linha da forma $(0,\ 0.6,\ 0,\ 0.4,\ 0)$. Repare que a soma destas probabilidades todas é um, pois estamos considerando todos os valores possíveis da sua fortuna, ou seja, este vetor corresponde a uma distribuição de probabilidade. Fazendo isto para cada valor possível da sua fortuna podemos arranjar os vetores de probabilidades como linhas de uma matriz que ficaria da forma seguinte,

$$\begin{array}{c} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \end{array} \begin{array}{c} \begin{array}{cccccc} \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

O exemplo acima corresponde a um tipo de processo estocástico muito importante, as chamadas cadeias de Markov.

Definição 1.2 (Cadeia de Markov). Uma cadeia de Markov é um processo estocástico $\{X_n\}_{n\in\mathbb{T}}$, com o tempo discreto, $\mathbb{T} = \{0, 1, 2, \ldots\}$, o espaço de estados E finito ou enumerável e que possui a propriedade de Markov,

$$P(X_{n+1} = j|X_0 = i_0, \ldots, X_n = i_n) = P(X_{n+1} = j|X_n = i_n), \tag{1.7}$$

para todos os estados $i_0, \ldots, i_n, j$ e todo instante n. Se $X_n = i$ dizemos que o processo no instante n está no estado i.

A equação (1.7) diz que o estado *futuro* do processo, $X_{n+1} = j$, não depende do *passado*, $X_0 = i_0, \ldots, X_{n-1} = i_{n-1}$, e só depende do *presente*, $X_n = i_n$.

A probabilidade condicional (1.7) é chamada **probabilidade de transição**.

Vamos restringir o nosso estudo às cadeias de Markov **homogêneas**, isto é, aquelas cadeias nas quais (1.7) não depende do tempo n,

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

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, como foi feito no exemplo 1.1, que é chamada de **matriz de transição**.

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

$$P = \begin{pmatrix} P_{0,0} & P_{0,1} & \ldots & P_{0,N} \\ P_{1,0} & P_{1,1} & \ldots & P_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ P_{N,0} & P_{N,1} & \ldots & P_{N,N} \end{pmatrix}. \quad \begin{array}{l} \longleftarrow \textit{transições do estado 0 aos estados 0,1,}\ldots\textit{,N} \\ \\ \\ \\ \end{array}$$

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

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

Se E é infinito, por exemplo $E = \{0, 1, 2, \ldots\}$, então a matriz P será infinita:

$$P = \begin{pmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \dots \\ P_{1,0} & P_{1,1} & P_{1,2} & \dots \\ P_{2,0} & P_{2,1} & P_{2,2} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Também podemos descrever as transições como grafos como ilustrado na figura 1.

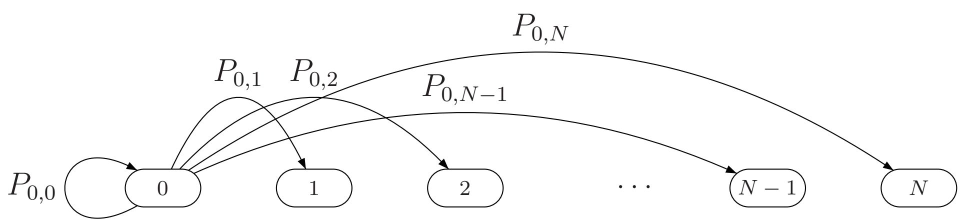


FIGURA 1. As setas entre os estados correspondem as transições, e o grafo é chamado de **topologia** da cadeia.

Vejamos agora outros exemplos.

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 vice-versa. 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

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

Vimos que a cada cadeia de Markov corresponde uma matriz de transição. Que propriedades caracterizam estas matrizes?

DEFINIÇÃO 1.4. A matriz $P = (P_{i,j})_{i,j \in E}$ é uma **matriz estocástica** se

(i) $P_{i,j} \geq 0$, para todos $i, j \in E$ e;

(ii) para todo $i \in E$, $\sum_{j \in E} P_{i,j} = 1$.

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 (i) corresponde a que as entradas são valores de probabilidades e a (ii) a que se o processo está no estado i no instante n então no próximo instante ele terá que estar en algum dos estados $j \in E$.

Não é difícil ver que uma matriz estocástica determina uma cadeia de Markov $\{X_n\}_{n\geq 0}$. De fato, o primeiro estado pode ser sorteado de uma distribuição discreta qualquer em E e estando no estado i, para determinar o estado da cadeia no próximo instante, 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 1.5. ***Passeio aleatório simples***

No passeio aleatório simples o espaço de estados são os números inteiros, 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.*

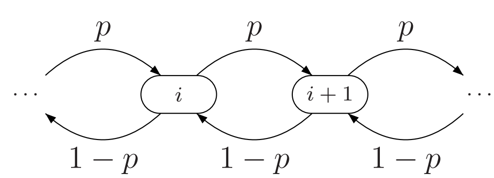


Figura 2. Topologia da cadeia para o passeio aleatório simples com infinitos estados.

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.

A partir de seqüências de variáveis aleatórias i.i.d.s podem se construir cadeias de Markov de diferentes formas, como ilustramos a seguir.

Exemplo 1.6. *Considere uma seqüência de variáveis aleatórias* $\xi_1, \xi_2, \ldots$ *i.i.d. discretas e com distribuição* $P(\xi_i = k) = p_k$, *com* $\sum_{k=0}^{\infty} p_k = 1$.

(a) *Seja X_0 uma variável aleatória com valores inteiros não negativos e independente da seqüência $\{\xi_j\}_{j\in\mathbb{N}}$ e defina $X_n = \xi_n$ para $n \in \mathbb{N}$. Podemos ver que*

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

Portanto, a família $\{X_n\}_{n\geq 1}$ é uma cadeia de Markov com matriz de transição

$$P = \begin{pmatrix} p_0 & p_1 & p_2 & \cdots \\ p_0 & p_1 & p_2 & \cdots \\ p_0 & p_1 & p_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Não é difícil perceber que o raciocínio acima funciona para qualquer outro espaço de estados, em particular para um espaço de estados finito, com a única diferença nesse caso que a matriz P seria finita. Em outras palavras, acabamos de ver que toda sequência de variáveis aleatórias discretas i.i.d.s pode ser considerada como uma cadeia de Markov cuja matriz de transição tem todas as linhas iguais. No exercício 9 pedimos para provar que a afirmação recíproca também vale.

(b) *Seja agora $X_0 = \xi_0 = 0$ e $X_n = \xi_1 + \cdots + \xi_n$, 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,

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

e a matriz de transição é $P = \begin{pmatrix} p_0 & p_1 & p_2 & \cdots \\ 0 & p_0 & p_1 & \cdots \\ 0 & 0 & p_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$

Observe que as linhas da matriz obtida neste caso são "deslocamentos" à direita das linhas da matriz do caso i.i.d. Por esta razão não é imediato como adaptar esta estrutura ao caso finito. Você teria alguma proposta? (fazer exercicio)

A construção neste exemplo é um caso particular de uma outra que é apresentada no exercício ??.

O seguinte exemplo é um caso particular do exemplo 1.6 (b).

Exemplo 1.7. ***Número de Sucessos de um processo de Bernoulli***

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 como $N_0 = 0$, a distribuição *inicial é* $\pi_0(0) = 1$ *e* $\pi_0(j) = 0$, $j > 0$. *Pelo exemplo anterior (caso (b)),* N_n *é uma cadeia de Markov com transições*

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

Logo a matriz de transição é $P = \begin{pmatrix} 1-p & p & 0 & 0 & \dots \\ 0 & 1-p & p & 0 & \dots \\ 0 & 0 & 1-p & p & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$

Na próxima seção descreveremos outras aplicações das cadeias de Markov.

2. Outros exemplos de cadeias de Markov

Exemplo 2.1 (Fila a Tempo Discreto). *Os usuários de certo serviço fazem uma fila para serem 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 aleatóriamente. 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 tem* i *pessoas na fila, 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 cadeia de Markov cujas probabilidades de transição podem ser calculadas como segue.*

Observe que

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

Usando isto, obtemos

$$\begin{aligned} P_{0,j} &= P(X_{n+1} = j | X_n = 0) = P(\xi_{n+1} = j) = p_j, \\ \\ P_{i,j} &= P(X_{n+1} = j | X_n = i) = P((i-1)^+ + \xi_{n+1} = j), \\ &= P((i-1) + \xi_{n+1} = j) = p_{j-i+1}, j \geq i \geq 1. \\ \\ P_{i,i-1} &= P(X_{n+1} = i-1 | X_n = i) = P((i-1)^+ + \xi_{n+1} = i-1) \\ &= P((i-1) + \xi_{n+1} = i-1) = p_{i-1-i+1} = p_0 \end{aligned}$$

Logo,

$$P = \begin{pmatrix} p_0 & p_1 & p_2 & \dots \\ p_0 & p_1 & p_2 & \dots \\ 0 & p_0 & p_1 & \dots \\ 0 & 0 & p_0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

EXEMPLO 2.2 (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 $0 < s < S$. Se após satisfazer a demanda Z_n, o estoque atingir o nível (inferior) s então é feita a reposição do estoque até o nível (superior) S. Se o estoque não atingir o nível inferior s então não é feita a reposição. Chamaremos de X_n à quantidade de produtos 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 na figura (3). Suponha que $s = 2$, $S = 5$ e que inicialmente o estoque tem 4 unidades (i.e. $X_0 = 4$). Logo:

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

EXEMPLO 2.3. *Sequência de sucessos Considere uma seqüência de v.a. independentes $\{T_n\}_{n\geq 0}$ que tomam só 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 seqüência 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

e matriz de transição:

$$P = \begin{pmatrix} q & p & 0 & 0 & \dots \\ q & 0 & p & 0 & \dots \\ q & 0 & 0 & p & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

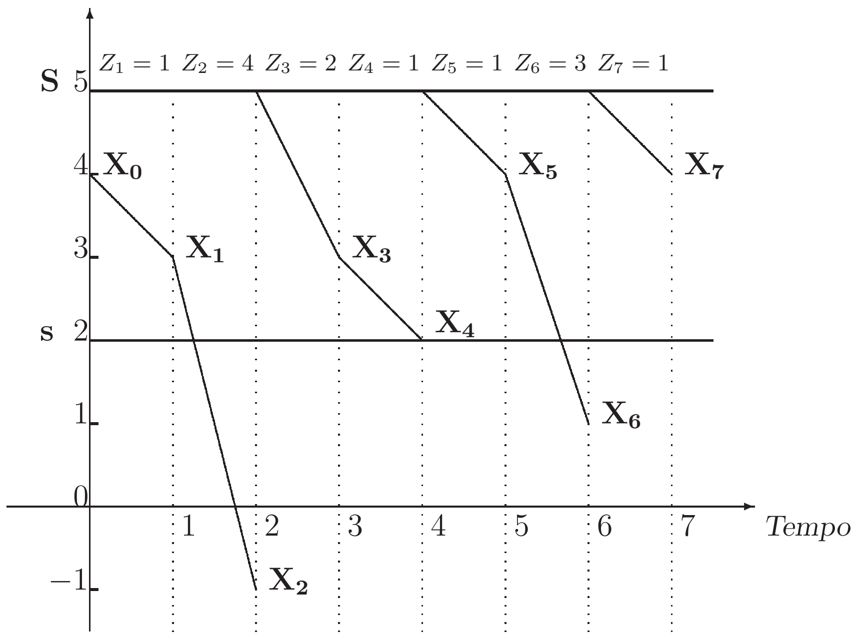


FIGURA 3. Inventário: uma realização do processo $\{X_n\}_{n\geq 0}$.

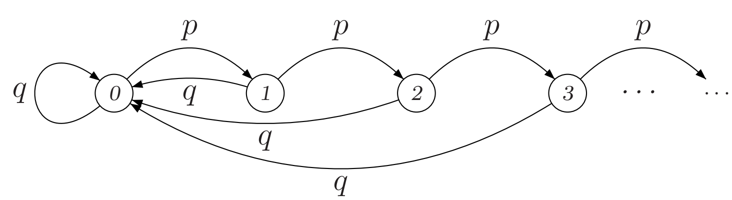


FIGURA 4. Topologia da cadeia de corrida de sucessos.

EXEMPLO 2.4 (Processo de Ramificação (Processo de Galton-Watson)). *Considere a seguinte seqüência de variaveis aleatórias independentes com valores inteiros não negativos:*

1) $\left\{Z_1^{(j)}\right\}_{j\geq 1}$ *i.i.d.*

2) $\left\{Z_2^{(j)}\right\}_{j\geq 1}$ *i.i.d.*

$\vdots \qquad \vdots$

n) $\left\{Z_n^{(j)}\right\}_{j\geq 1}$ *i.i.d.*

$\vdots \qquad \vdots$

Defina o seguinte processo:

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

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.*

Vale que

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

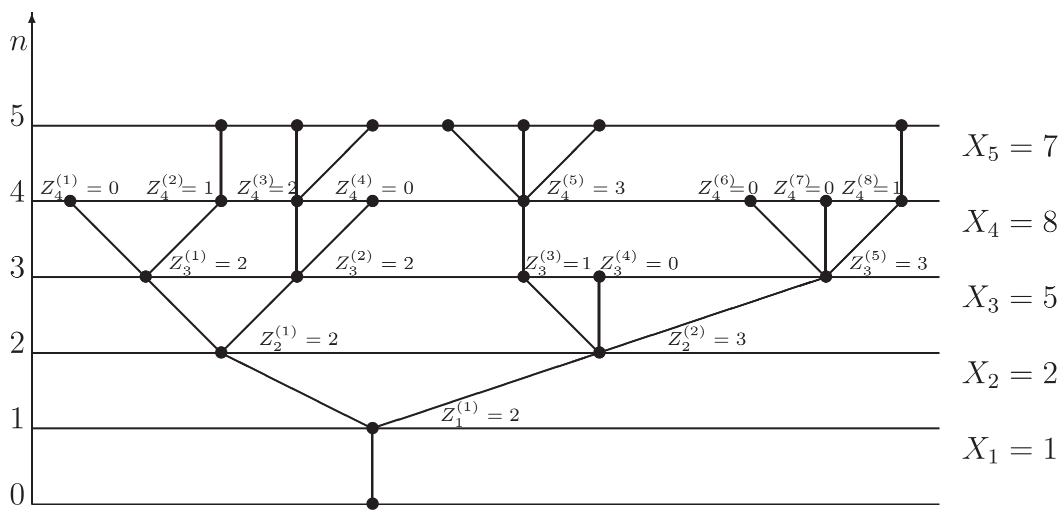


FIGURA 5. Processo de Ramificação.

EXEMPLO 2.5 (Processo de nascimento e morte). *Uma cadeia de Markov com espaço de estados $E = \{0, 1, \ldots, d\}$ (ou espaço de estados $E = \{0, 1, \ldots, \infty\}$, 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.*

3. Matrizes de transição de ordem superior

EXEMPLO 3.1. *No exemplo (1.3), o processo* X_n *é uma cadeia de Markov se supormos que as mudanças de classe social estão dadas pela seguinte matriz* [1] *discutiremos a aplicabilidade destes modelos na teoria social.*

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

Ela nos diz, por exemplo, que a probabilidade de que os filhos de uma família de classe baixa (estado **1***) permaneçam nesta classe é 0.7.*

EXEMPLO 3.2. *Para o 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 desça 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.*

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

Generalizando o exemplo, podemos obter o seguinte resultado.

TEOREMA 3.3. *Sejam* $i_0, i_1, i_2, \ldots, i_m \in E$, *vale que:*

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

[1] Na seção **??**

Demonstração. Exercício. □

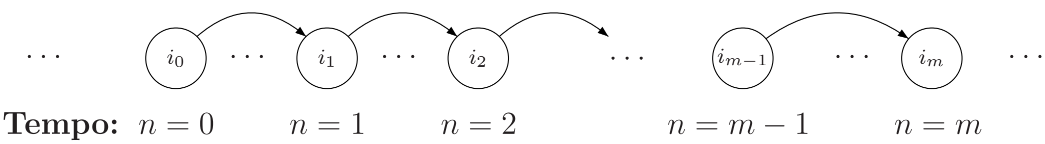


Figura 6. Transições correspondentes ao teorema 3.3.

Voltando ao exemplo (3.2),

Exemplo 3.4. *Suponha de novo que a família começa na classe média (estado* **2***) na geração 0. Qual a probabilidade que a geração 2 desça para a classe baixa (estado* **1***)?*

Para resolver este problema, devemos considerar os três estados possíveis para a geração 1 e usar o teorema 3.3.

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

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

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

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:

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

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 $n+3$):

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

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

Teorema 3.5. *(Equações de Chapman-Kolmogorov)*

$$P^{(m+n)}_{i,j} = \sum_{k\in E} P^{(m)}_{i,k}P^{(n)}_{k,j}.$$

Observe que se chamarmos de $P^{(m)} = (P^{(m)}_{i,j})$ 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,

$$P(X_{n+m}=j|X_0=i) = \sum_{k\in E} P(X_{n+m}=j, X_m=k|X_0=i). \tag{3.8}$$

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 (3.8), obtemos o resultado desejado. □

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 3.6. (Distribuição Inicial)
Seja π_0 uma distribuição de probabilidades no conjunto E,

$$\pi_0(i) \geq 0, i \in E, \quad \sum_{i \in E} \pi_0(i) = 1,$$

dizemos que π_0 é a distribuição inicial da cadeia se para todo $i \in E$ vale $P(X_0 = i) = \pi_0(i)$.

Em outras palavras, a distribuição inicial de uma cadeia é simplesmente 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 $n \in \mathbb{N}$,

$$\begin{aligned} P(X_n=k) &= \sum_{i\in E} P(X_n=k|X_0=i)P(X_0=i), \\ &= \sum_{i \in E} P_{i,k}^{(n)} \pi_0(i), \\ &= \pi_0^t P^n. \end{aligned}$$

Aqui π_0 representa o vetor coluna dos valores da distribuição inicial da cadeia.

Usando o teorema 3.3, podemos obter o seguinte resultado.

Proposição 3.7. *Seja π_0 a distribuição inicial da cadeia $\{X_n\}_{n\geq 0}$ que tem matriz de transição $P = (P_{i,j})_{i,j\in E}$. Sejam $i_0, i_1, i_2, \ldots, i_m \in E$ então vale*

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

Demonstração.

$$\begin{aligned} & P(X_0 = i_0, X_1 = i_1, \ldots, X_m = i_m) \\ &= P(X_1 = i_1, \ldots, X_m = i_m | X_0 = i_0) P(X_0 = i_0), \\ &= P_{i_{m-1},i_m} \cdots P_{i_0,i_1} \pi_0(i_0). \end{aligned}$$

□

Mais geralmente é possível provar,

Proposição 3.8. *(Distribuições Finito-Dimensionais)*
Seja π_0 a distribuição inicial da cadeia $\{X_n\}_{n\geq 0}$ que tem matriz de transição $P = (P_{i,j})_{i,j\in E}$. Sejam $i_0, i_1, i_2, \ldots, i_m \in E$ e $n_1 < n_2 < \cdots < n_m$ então vale

$$P(X_0 = i_0, X_{n_1} = i_1, \ldots, X_{n_m} = i_m) = \pi_0(i_0) P^{n_1}_{i_0,i_1} P^{n_2-n_1}_{i_1,i_2} \cdots P^{n_m-n_{m-1}}_{i_{m-1},i_m}.$$

Demonstração. Exercício. □

Exemplo 3.9. *Considere uma cadeia de Markov X_n com espaço de estados $E = \{a, b, c\}$ e matriz de transição:*

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

(a) A partir desta matriz podemos construir o grafo das transições ou topologia da cadeia. Reciprocamente podemos achar a matriz de transição P a partir do grafo.

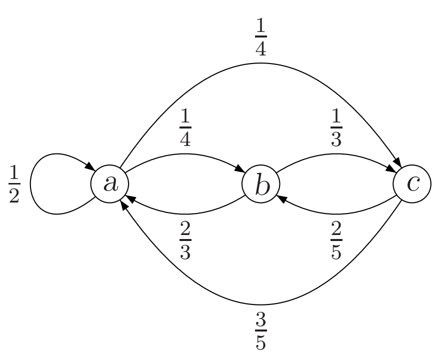


Figura 7.

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

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

(c) Calcule $P(X_6 = c|X_4 = b)$.

Achemos primeiro P^2.

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

$$P = \begin{array}{c c} & \begin{array}{ccc} \mathbf{a} & \mathbf{b} & \mathbf{c} \end{array} \\ \begin{array}{c} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{array} & \begin{pmatrix} \frac{17}{30} & \frac{9}{40} & \frac{5}{24} \\ \frac{8}{15} & \frac{3}{10} & \frac{1}{6} \\ \frac{17}{30} & \frac{3}{20} & \frac{17}{60} \end{pmatrix} \end{array}.$$

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

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

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

(e) Se a distribuição inicial da cadeia esta dada por $\pi_0(a) = 0.3$, $\pi_0(b) = 0.3$ *e* $\pi_0(c) = 0.4$, *determine a função de probabilidade de* X_2.

Chamaremos de π_0 *ao vetor coluna*

$$\pi_0 = \begin{pmatrix} \pi_0(a) \\ \pi_0(b) \\ \pi_0(c) \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.3 \\ 0.4 \end{pmatrix}.$$

Sabemos que o vetor da função de probabilidades de X_2 *está dado por*

$$\pi_0^t P^2 = [0.3\ 0.3\ 0.4] \begin{pmatrix} \frac{17}{30} & \frac{9}{40} & \frac{5}{24} \\ \frac{8}{15} & \frac{3}{10} & \frac{1}{6} \\ \frac{17}{30} & \frac{3}{20} & \frac{17}{60} \end{pmatrix} = \left(\frac{167}{300}\ \frac{87}{400}\ \frac{271}{1200} \right).$$

Portanto,

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

4. 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. Nesta seção analisaremos este fenômeno no caso de cadeias com dois estados. Começaremos com um exemplo.

Exemplo 4.1. *Considere a cadeia de Markov X_n com espaço de estados $E = \{0,1\}$ e matriz de transição,*

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

Então:

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

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

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \tag{4.9}$$

onde $0 \leq p \leq 1, 0 \leq q \leq 1$, com a topologia correspondente,

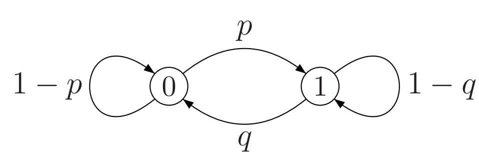


Figura 8.

Isto é,

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

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

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

Usando as relações

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

e

$$\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}^2 = (p+q)\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}.$$

É possível provar por um argumento indutivo que

$$P^n = \frac{1}{p+q}\begin{pmatrix} q & p \\ q & p \end{pmatrix} + \frac{(1-p-q)^n}{p+q}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}. \tag{4.10}$$

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

- $p+q=0$, i.e., $p=0$ e $q=0$
  Neste caso P é a matriz identidade de dimensão dois, vale que $P^n = P$ para todo $n \in \mathbb{N}$ e portanto as linhas não se aproximam entre si. 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
  $$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$
  Se n for par, teremos que P^n é a matriz identidade de ordem dois e para n ímpar, $P^n = P$. Como conseqüência disto temos que o limite de P^n quando $n \to \infty$ não existe pois a matriz oscila entre duas matrizes fixas.
- $0 < p+q < 2$
  Neste caso vale que $|1-p-q| < 1$ e portanto $(1-p-q)^n \to 0$ quando $n \to \infty$. Usando (4.10), obtemos
  $$\lim_{n\to\infty} P^n = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix}.$$

No último caso considerado, as linhas da matriz convergem para o vetor de probabilidades de uma distribuição que denotaremos como π_∞, isto é,

$$\lim_{n\to\infty} P^n_{0,0} = \lim_{n\to\infty} P^n_{1,0} = \pi_\infty(0),$$

$$\lim_{n\to\infty} P^n_{0,1} = \lim_{n\to\infty} P^n_{1,1} = \pi_\infty(1),$$

com $\pi_\infty(0) = \dfrac{q}{p+q}$ e $\pi_\infty(1) = \dfrac{p}{p+q}$.

A relação (4.10) permite obter também uma estimativa para a taxa de convergência das probabilidades de transição em n passos para a distribuição π_∞.

Proposição 4.2. *Para uma cadeia de Markov $\{X_n\}_{n\geq 1}$ com dois estados, $E = \{0,1\}$ e tal que $0 < p+q < 2$, vale*

$$|P^n_{i,0} - \pi_\infty(0)| = \left| P^n_{i,0} - \frac{q}{p+q} \right| \leq |1-p-q|^n,$$

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

Estas probabilidades de transição se aproximam de π_∞ com velocidade exponencial, ou seja, muito rápido. Por isso, de forma geral quando $0 < p+q < 2$, observamos a proximidade entre as linhas de P^n mesmo para valores de n pequenos.

Exemplo 4.3. *No exemplo 4.1 temos que $p = 0.5$ e $q = 0.3$, portanto $0 < p+q < 2$ e as linhas da matriz devem convergir para $(\pi_\infty(0)\ \ \pi_\infty(1)) = (3/8\ \ 5/8) = (0.375\ \ 0.625)$. A diferença entre elas vai para zero mais rápido que $(0.2)^n$. Observe que para $n = 4$ obtivemos uma precisão de duas casas decimais.*

Quando existe π_∞, temos que para instantes de tempo grandes, a matriz P^n se aproxima da matriz de transição de ordem n de uma sequência de variáveis aleatórias i.i.d.s. (veja o exercício 22). Isto significa que para instantes de tempo grandes, os estados irão se tornando "menos dependentes" e é natural pensar que a cadeia "esquece" o estado onde ela começou, ou seja, a sua distribuição inicial. Provemos que isto é assim.

Suponha que a distribuição inicial da cadeia é π_0, ou seja, temos $P(X_0 = 0) = \pi_0(0)$ e $P(X_0 = 1) = \pi_0(1)$. Sabemos que $(P(X_n = 0)\ \ P(X_n = 1)) = (\pi_0(0),\ \pi_0(1)) \cdot P^n$. Usando a expressão que temos para P^n e o fato que π_0 é uma distribuição, obtemos

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

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

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

que era o que queriamos provar.

As quantidades $\pi_\infty(0) = \dfrac{q}{p+q}$ e $\pi_\infty(1) = \dfrac{p}{p+q}$ podem ser interpretadas como as probabilidades da cadeia estar a longo prazo no estado 0 ou no 1, respectivamente, por isto π_∞ é chamada de **distribuição assintótica** da cadeia.

Exemplo 4.4. *Um pai que está ensinando ao seu filho de cinco anos a ler, observou que se que o menino faz um erro numa palavra, ele fará um erro na seguinte no texto também em 25% dos casos e se ele ler uma palavra bem, a próxima é lida corretamente em 90%*

das vezes. Se a criança ler um texto de 100 palavras, dê uma aproximação para o número delas que ele lerá corretamente.

Solução:
Dado um texto que o menino deve ler, podemos considerar uma cadeia de Markov $\{X_n\}_{n\geq 1}$ na qual cada variável X_n toma os valores **M** ou **B** quando a n-ésima palavra tenha sido lida com ou sem erros, respectivamente.

O enunciado está dizendo que esta cadeia tem como matriz de transição

$$P = \begin{array}{c c} & \begin{array}{cc} \mathbf{M} & \mathbf{B} \end{array} \\ \begin{array}{c} \mathbf{M} \\ \mathbf{B} \end{array} & \begin{pmatrix} 0.25 & 0.75 \\ 0.10 & 0.90 \end{pmatrix} \end{array}.$$

Comparando com a matriz 4.9, vemos que neste caso teremos $p = 0.75$ e $q = 0.10$, portanto $\pi_\infty = \left(\dfrac{2}{17}, \dfrac{15}{17}\right)$. Pela proposição 4.2, $\dfrac{15}{17}$ é uma aproximação para a probabilidade de ler corretamente uma palavra. Obtemos então que o menino lerá corretamente aproximadamente 88 palavras do texto.

Para saber o quanto seria boa esta aproximação, observe que neste caso, $1-p-q = 0.15$, como $(0.15)^4 = 5 \cdot 10^{-4}$, para $n \geq 4$ já devemos ter uma precisão de, no mínimo duas casas decimais.

Concluimos que quando o espaço de estados possui somente dois elementos é possível determinar completamente o comportamento da cadeia no limite. Não há distribuição assintótica em dois casos. Um é quando a matriz de transição é a identidade, 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 existe distribuição assintótica.

5. Distribuição invariante.

Apresentaremos a seguir a noção de distribuição invariante.

PROPOSIÇÃO 5.1. *Para uma cadeia $\{X_n\}_{n\geq 0}$, suponha que π é uma distribuição satisfazendo as equações*

$$\sum_{i\in E} \pi(i) P_{i,j} = \pi(j), \tag{5.11}$$

para todo $j \in E$. Então valem as seguintes afirmações.

(i) Para todo $j \in E$,

$$\sum_{i\in E} \pi(i) P^2_{i,j} = \pi(j).$$

Em geral, para todo $n \geq 1$, vale que

$$\sum_{i\in E} \pi(i) P^n_{i,j} = \pi(j).$$

(ii) Se a cadeia tem ditribuição inicial π, então para todo $n \geq 1$ vale

$$P(X_n = i) = \pi(i). \tag{5.12}$$

Demonstração. (i) Considere o caso $n = 2$,

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

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

(ii) Se X_n tem distribuição inicial π, então

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

□

A igualdade (5.12) afirma que se a cadeia começar com uma distribuição π satisfazendo (5.11), então a distribuição em todos os instantes será a mesma e igual a π. Em outras palavras, a distribuição marginal permanece invariante ante as transições da cadeia, isto motiva a seguinte definição.

Definição 5.2. Toda distribuição π que satisfaça (5.11) será chamada de **distribuição invariante** da cadeia $\{X_n\}_{n \geq 0}$.

É possível provar que uma cadeia começando com uma distribuição invariante é um processo estocástico estacionário no sentido estrito. Isto é consequência de (5.12) e da proposição 3.8.

Nota 5.3. Repare que no caso que E é finito, a equação (5.11) pode ser reescrita da forma

$$\pi^t \cdot P = \pi^t. \tag{5.13}$$

Para determinar a distribuição invariante de uma cadeia de Markov com espaço de estados finito, basta encontrar um vetor π satisfazendo (5.13), com entradas não negativas e tal que $\sum_{i \in E} \pi(i) = 1$. Vejamos isso no seguinte exemplo.

Exemplo 5.4. *Considere uma cadeia de Markov com espaço de estados* $E = \{0, 1, 2\}$ *e matriz de transição:*

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

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

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

Para encontrar a sua distribuição invariante devemos resolver o sistema

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

cuja solução é $(11/23 \;\; 9/23 \;\; 3/23)$.

Mais adiante veremos que quando o espaço de estados é finito, toda cadeia de Markov possui pelo menos uma distribuição invariante (veja também o exercício 24). No caso de E infinito isto não vale e um exemplo disto é o passeio aleatório simples simétrico.

Exemplo 5.6 (Passeio aleatório simples simétrico).

Para provar que esta caseia não possui distribuição invariante vamos raciocinar pelo absurdo.

Suponha que π é uma distribuição de probabilidades em $\mathbb{Z}$ tal que (5.11) 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 invariante da cadeia e portanto tais distribuições não existem.

É possível provar também que nenhum passeio aleatório simples nos inteiros possui distribuição invariante. No exercício 15 lhe pediremos para encontrar a (única) distribuição invariante de certos passeios aleatórios nos inteiros.

Exemplo 5.7. *Vamos a achar a distribuição invariante π do processo de nascimento e morte (exemplo 2.5).*

Se $\displaystyle\sum_{i\in E} \pi(i)P_{i,j} = \pi(j),$ *então* $\left\{ \begin{array}{l} \pi(0)r_0 + \pi(1)q_1 = \pi(0), \\ \pi(i-1)p_{i-1} + \pi(i)r_i + \pi(i+1)q_{i+1} = \pi(i), \quad i \geq 1 \end{array} \right.$

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

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

então $\pi(i+1)q_{i+1} - \pi(i)p_i = 0, \quad i \geq 1$ 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 invariante 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, sumando 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 $\sum_{i \in E} \Gamma_i < \infty$. Resumindo,

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

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

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

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

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

Exemplo 5.8. ***Cadeia de Ehrenfest***

A cadeia de Ehrenfest corresponde a um processo de nascimento e morte, logo

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

e portanto,

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

Você consegue identificar a distribuição obtida? (Veja o exercício **??***).*

Passaremos a estudar agora o papel da distribuição invariante no comportamento assintótico das cadeias de Markov. Para isto começaremos com a seguinte definição.

Definição 5.9. Toda distribuição π_∞ que satisfaça

$$\lim_{n \to \infty} P^n_{i,j} = \pi_\infty(j), \tag{5.14}$$

para todo $j \in E$, será chamada de **distribuição assintótica** da cadeia $\{X_n\}_{n \geq 0}$.

Observe que pela própria definição temos que a distribuição assintótica é única.

Em (5.14) estamos pedindo que as linhas da matriz de transição de ordem n da cadeia se aproximem de certa distribuição no limite quando n cresce. O resultado a seguir justifica por que chamamos a esta distribuição de distribuição assintótica.

Proposição 5.10. *A distribuição π_∞ será a distribuição assintótica da cadeia $\{X_n\}_{n\geq 0}$ se e somente se para todo $j \in E$,*

$$\lim_{n\to\infty} P(X_n = j) = \pi_\infty(j).$$

Para provar este resultado é interessante esclarecer a relação entre a distribuição assintótica de uma cadeia e as sua possíveis distribuições invariantes. Veremos agora que quando a distribuição assintótica existe, ela é a única distribuição invariante.

Proposição 5.11. *Seja $\{X_n\}_{n\geq 0}$ uma cadeia de Markov com matriz de transição P e com distribuição assintótica π_∞. Então π_∞ é a* ***única*** *distribuição invariante da cadeia.*

Demonstração. Faremos a prova para um espaço de estados finito.

Observe que para todo $j \in E$, $P(X_n = j) \to \pi_\infty(j)$ quando $n \to \infty$ pois

$$P(X_n = j) = \sum_{i\in E} P(X_0 = i)P^n_{i,j} \to \sum_{i\in E} P(X_0 = i)\pi_\infty(j) = \pi_\infty(j)\sum_{i\in E} P(X_0 = i) = \pi_\infty(j).$$

Como $P(X_{n+1} = j) = \sum_{i\in E} P(X_n = i)P_{i,j}$, passando ao limite em n obteremos que π_∞ satisfaz a igualdade (5.11), sendo portanto uma distribuição invariante.

Para provar que ela é a única podemos raciocinar pelo absurdo e supor que existe uma outra distribuição invariante π_*. Se a escolhermos como distribuição inicial teremos, por um lado que $P(X_n = j) \to \pi_\infty(j)$, como visto acima, e por outro que para todo $n \geq 0$, $P(X_n = j) = \pi_*$ pelo fato de π_* ser invariante. Isto contradiz a nossa suposição de π_* e π_∞ serem diferentes e portanto elas devem coincidir. □

Com este resultado, a prova da proposição 5.10 resulta bastante simples e fica como exercício para o leitor.

Exemplo 5.12. *Na seção 4 mostramos que quando temos somente dois elementos no espaço de estados, para cadeias tais que $0 < p + q < 2$ existe a distribuição assintótica. Portanto nesse caso existe uma única distribuição invariante dada por*

$$\pi^t_\infty = \left(\frac{q}{p+q}, \frac{p}{p+q}\right). \tag{5.15}$$

Se $p = q = 1$, não existe distribuição assintótica quando $n \to \infty$. No entanto, observe que neste caso (5.13) toma a forma

$$(\pi(0)\ \ \pi(1)) = (\pi(0)\ \ \pi(1)) \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (\pi(1)\ \ \pi(0)),$$

que tem como única solução $\pi(0) = \pi(1) = 1/2$ e vale também neste caso que a única distribuição invariante satisfaz (5.15).

Se $p = q = 0$, vimos 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 invariante para esta cadeia.

Que condições garantiriam a existência da distribuição assintótica? Vimos acima que quando uma cadeia tem distribuição assintótica ela é a única distribuição invariante. Portanto, cadeias sem distribuição invariante ou com mais de uma, não poderão ter distribuição assintótica. Por outro lado, acabamos de ver um exemplo de uma cadeia com uma única distribuição invariante e sem distribuição assintótica, então além da existência da distribuição invariante precisamos pedir alguma coisa a mais. A seguir apresentamos um resultado básico de convergência das cadeias de Markov que é uma espécie de recíproca da proposição 5.11.

Teorema 5.13 (Teorema básico de convergência).
Se uma cadeia de Markov $\{X_n\}_{n\geq 0}$ com distribuição invariante π_{inv} for irredutível e aperiódica então π_{inv} será a sua distribuição assintótica.

Ainda não sabemos o que são cadeias irredutíveis e aperiódicas. Estas noções estão relacionadas com a estrutura do espaço de estados e serão estudadas na próxima seção, mas observe que como consequência do teorema, teriamos que tais cadeias podem ter no máximo uma distribuição invariante, pois toda distribuição invariante seria assintótica e a distribuição assintótica é única.

6. Classificação dos estados: parte I

Definição 6.1. Dizemos que o **estado j é acessível desde o estado i**, $i \to j$, se existe um instante $n > 0$ tal que $P^n_{i,j} > 0$. Se $i \to j$ e $j \to i$ diremos que **i e j estão comunicados** e o denotaremos por $i \leftrightarrow j$.

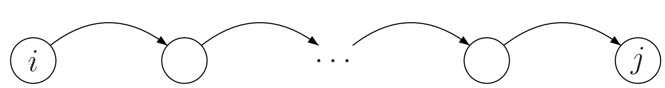


Figura 9. j acessível desde i

A relação de comunicação $\leftrightarrow$ é reflexiva, simétrica e transitiva, isto é,

Proposição 6.2.
(i) $i \leftrightarrow i$. *(pois* $P^0_{i,i} = 1$*)*
(ii) $i \leftrightarrow j \Rightarrow j \leftrightarrow i$. *(pela definição)*
(iii) $i \leftrightarrow j$ *e* $j \leftrightarrow k \Rightarrow i \leftrightarrow k$ *(conseqüência de Chapman-Kolmogorov).*

Definição 6.3. Um estado $i \in E$ é dito **não essencial** se for possível sair dele numa quantidade finita de passos e não voltar nunca mais. Isto é, se existirem um instante $n \in \mathbb{N}$ e $j \in E$ tais que $P^n_{i,j} > 0$ e $P^m_{j,i} = 0$ para todo $m \in \mathbb{N}$.

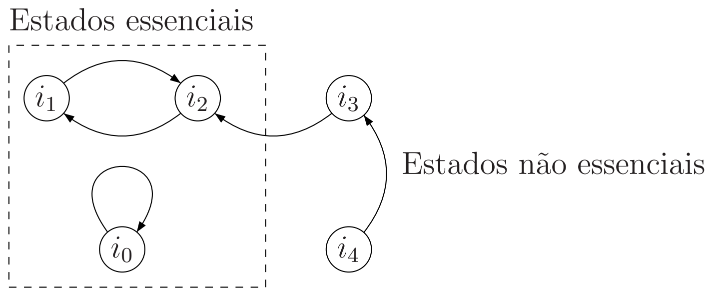


FIGURA 10.

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.

Repare que para todo estado essencial i vale que se $i \to j$ então $i \leftrightarrow j$. Além disto, para um estado i ser essencial tem que ocorrer uma das duas alternativas seguintes.

- Não é possível sair de i.
  Neste caso $P^n_{i,j} = 0$ 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. Tais estados estão comunicados só com eles mesmos.
- É possível sair de i.
  Sempre que $P^n_{i,j} > 0$ teremos que existe $m \in \mathbb{N}$ tal que $P^m_{j,i} > 0$.
  Ou seja, existe $j \neq i$ tal que $P^n_{i,j} > 0$ para algum $n \geq 1$ ($i \to j$) e nesse caso temos algum $m \in \mathbb{N}$ tal que $P^m_{j,i} > 0$ $j \to i$.

EXEMPLO 6.4 (Jogo de azar (continuação)).
No exemplo 1.1 com $N = 5$ a matriz de transição tinha a forma,

$$
\begin{array}{c}
\\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5}
\end{array}
\begin{array}{c}
\begin{array}{cccccc} \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0.6 & 0 & 0.4 & 0 & 0 & 0 \\
0 & 0.6 & 0 & 0.4 & 0 & 0 \\
0 & 0 & 0.6 & 0 & 0.4 & 0 \\
0 & 0 & 0 & 0.6 & 0 & 0.4 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{array}.
$$

A topologia correspondente seria

Observe que qualquer valor $1 < i < 5$ *representa um estado não essencial. Os estados* $i = 0$ *(ruína) e* $i = 5$ *(lucro máximo) são estados absorventes da cadeia.*

EXEMPLO 6.5. *No passeio aleatório simples com* $0 < p < 1$, $i \to j$ *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.*

Esta cadeia não possui estados absorventes.

Para cada estado essencial, podemos considerar a classe de todos os estados comunicados com ele. Em toda classe haverá no mínimo um elemento (ele próprio), sendo que haverá exatamente um se e somente se o estado for absorvente. Desta forma, 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. Estas classes serão chamadas de **classes comunicantes de estados essenciais ou classes irredutíveis**.

EXEMPLO 6.6. *Considere uma cadeia de Markov com a seguinte matriz de transição,*

$$\begin{array}{c} \\ \mathbf{0}\\ \mathbf{1}\\ \mathbf{2}\\ \mathbf{3}\\ \mathbf{4}\\ \mathbf{5}\\ \mathbf{6}\\ \mathbf{7}\\ \mathbf{8} \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} & \mathbf{7} & \mathbf{8} \end{array}\\
\begin{pmatrix}
0.2 & 0.8 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0.1 & 0.3 & 0.5 & 0 & 0 & 0 & 0.1 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & 0.3 & 0 & 0 & 0 & 0 & 0.7\\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0.4 & 0 & 0 & 0.6 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0.3 & 0 & 0 & 0 & 0.3 & 0 & 0.4 & 0\\
0 & 0 & 0 & 0.2 & 0 & 0 & 0 & 0 & 0.8
\end{pmatrix}
\end{array}$$

e topologia representada na figura 11.

Os estados essenciais são $\mathbf{3}, \mathbf{4}, \mathbf{5}, \mathbf{7}$ *e* $\mathbf{9}$. *Os não essenciais seriam* $\mathbf{1}, \mathbf{2}, \mathbf{6}$ *e* $\mathbf{8}$.

As classes irredutíveis são $\mathcal{C}_1 = \{\mathbf{5}\}$, $\mathcal{C}_2 = \{\mathbf{4}, \mathbf{9}\}$ *e* $\mathcal{C}_3 = \{\mathbf{3}, \mathbf{7}\}$.

DEFINIÇÃO 6.7. Uma cadeia de Markov é dita irredutível quando o conjunto de estados E é uma classe irredutível, ou seja, quando todos os estados estão comunicados. Caso contrário, a cadeia será chamada de redutível.

EXEMPLO 6.8. *Um estado* i *é absorvente se e somente se* $\mathcal{C} = \{i\}$ *é uma classe irredutível.*

EXEMPLO 6.9. *A cadeia do exemplo 6.6 é redutível, pois possui três classes irredutíveis.*

EXEMPLO 6.10. *O passeio aleatório simples é uma cadeia irredutível quando* $0 < p < 1$.

EXEMPLO 6.11. *O passeio aleatório com barreiras absorventes tem espaço de estados* $E = \{1, 2, \ldots, n\}$ *e topologia:*

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

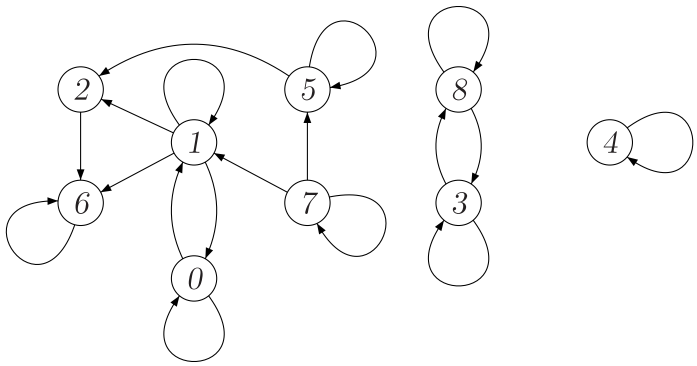


Figura 11.

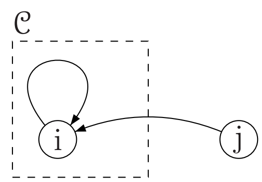


Figura 12.

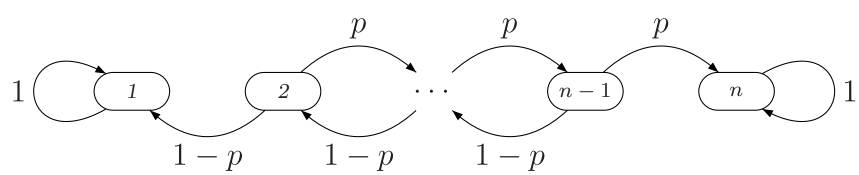


Figura 13. Topologia da cadeia com propriedades absorventes

Exemplo 6.12. *Considere uma cadeia de Markov com espaço de estados $E = \{a, b, c, d, e\}$ e matriz de transição*

$$P = \begin{array}{c} \\ \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \\ \mathbf{d} \\ \mathbf{e} \end{array} \begin{array}{c} \begin{array}{ccccc} \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \end{array} \\ \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \end{pmatrix} \end{array}$$

A topologia da cadeia é representada na figura 14.

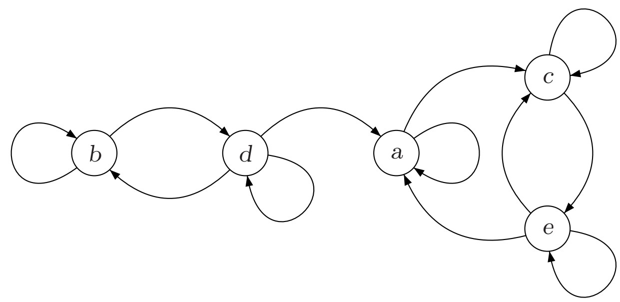


Figura 14.

Os estados não essenciais são $\{b, d\}$ e a classe $\mathcal{C} = \{a, c, e\}$ é irredutível, logo a cadeia não é irredutível. Reordenando os estados de forma que apareçam primeiro os estados essenciais, obtemos a matriz

$$\tilde{P} = \begin{array}{c} \\ \mathbf{a} \\ \mathbf{c} \\ \mathbf{e} \end{array} \begin{array}{c} \begin{array}{ccc} \mathbf{a} & \mathbf{c} & \mathbf{e} \end{array} \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \end{array}.$$

Observe que esta submatriz de P é a matriz de transição da cadeia restrita à classe $\mathcal{C}$.

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

Outra propriedade de interesse dos estados de uma cadeia é o período.

Definição 6.13. Seja $\{X_n\}_{n\geq 0}$ uma cadeia de Markov e $i \in E$ um estado. Chamaremos de período de i ao valor

$$d(i) = m.d.c.\{n \geq 1 : P_{ii}^{(n)} > 0\}, \tag{6.16}$$

ou seja, d é o maior inteiro que divide a todos os $n \geq 1$ tais que $P_{ii}^{(n)} > 0$. Se para todo $n \geq 1$, $P_{ii}^{(n)} = 0$ então $d(i) = 0$. Estados com período 1 serão chamados de **aperiódicos**.

Exemplo 6.14. *As entradas não nulas na diagonal de uma matriz estocástica correspondem a estados aperiódicos, pois se $P_{ii} > 0$, teremos $\{n \geq 1 : P_{ii}^{(n)} > 0\} = \{1, \dots\}$, cujo m.d.c. é 1. Por exemplo, na matriz*

$$P = \begin{array}{c} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \end{array} \begin{array}{c} \begin{array}{ccc} \mathbf{0} & \mathbf{1} & \mathbf{2} \end{array} \\ \begin{pmatrix} 0.2 & 0.1 & 0.7 \\ 0.5 & 0 & 0.5 \\ 0.8 & 0 & 0.2 \end{pmatrix} \end{array},$$

*os estados **0** e **2** são aperiódicos.*

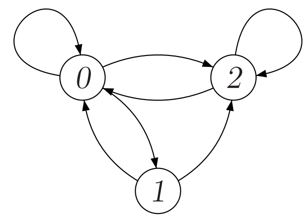


FIGURA 15. Topologia do exemplo 6.14

O estado $\mathbf{1}$ *também é aperiódico pois* $P^{(2)}_{\mathbf{11}} \geq P_{\mathbf{10}}P_{\mathbf{01}} = 0.05 > 0$ *e* $P^{(3)}_{\mathbf{11}} \geq P_{\mathbf{10}}P_{\mathbf{00}}P_{\mathbf{01}} = 0.010 > 0$. *Portanto* $\{n \geq 1 : P^{(n)}_{ii} > 0\} = \{2, 3, \ldots\}$. *como 2 e 3 são primos entre si, então* $d(\mathbf{1}) = \mathbf{1}$.

EXEMPLO 6.15. *Consideremos de novo a cadeia de Ehrenfest com 3 bolas. A matriz de transição é*

$$P = \begin{array}{c} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

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

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

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.

Calculemos por exemplo, $d(0)$. *Temos que para todo* $k \in \mathbb{N}$, $p^{(2k)}_{00} > 0$ *e* $p^{(2k+1)}_{00} = 0$, *portanto,* $\{n \geq 1 : p^{(n)}_{ii} > 0\} = \{2, 4, 6, \ldots\}$ *e* $d(0) = 2$.

Usando o mesmo raciocínio, vemos que se calcularmos o período de qualquer outro dos estados, ele terá o mesmo valor. Isto é um resultado que vale para classes irredutíveis de forma geral.

PROPOSIÇÃO 6.16. *Os períodos de todos os estados de uma classe irredutível coincidem.*

DEMONSTRAÇÃO. dd $\square$

Este resultado, permite definir o período de uma classe irredutível como o período de qualquer um dos seus estados. Analogamente, chamaremos de classe aperiódica a toda classe irredutível de período um.

Exemplo 6.17.

No exemplo 6.6, todas as classes irredutíveis são aperiódicas.

Exemplo 6.18.

A cadeia de Ehrenfest e o passeio aleatório simples com $0 < p < 1$, são exemplos de cadeias irredutíveis de período dois.

Exemplo 6.19 (Cadeia com período três).

Consideremos uma cadeia com a seguinte matriz de transição.

$$P = \begin{array}{c} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \end{array} \begin{array}{c} \begin{array}{cccccc} \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.6 & 0 & 0.2 \\ 0.5 & 0 & 0 & 0 & 0.5 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

A topologia da cadeia esta representada na figura 16.

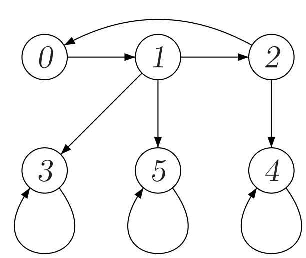


Figura 16. Topologia do exemplo 6.19.

Não é difícil comprovar que todos os estados estão comunicados portanto, a cadeia é irredutível. Calculemos o período do estado $\mathbf{1}$.

Observe que se sairmos de $\mathbf{1}$ *para* $\mathbf{2}$, *voltaremos no* $\mathbf{1}$ *no mínimo em três passos e o mesmo ocorre se sairmos pelo estado* $\mathbf{3}$ *ou* $\mathbf{5}$. *Ou seja,* $P^{(n)}_{\mathbf{11}} = 0$ *para* $1 \leq n \leq 2$ *e* $P^{(3)}_{\mathbf{11}} > 0$. *Além disto, numa trajetória aparecerão* $\mathbf{1}$*'s consecutivos somente depois de três transições, então* $P^{(n)}_{\mathbf{11}} > 0$ *só se* $n = 3k$, $k \in \mathbb{N}$. *Portanto* $d(\mathbf{1}) = \mathbf{3}$ *e a cadeia tem período três.*

7. Cadeias com espaço de estados finito

No caso em que o espaço de estados é finito, podemos reformular as noções e propriedades relativas às cadeias de Markov usando somente a álgebra linear. Veremos como fazer isto e obteremos uma prova do teorema básico de convergência a partir de resultados sobre a estrutura espectral de matrizes não negativas. Nesta seção suporemos que o espaço de estados é finito, particularmente $E = \{1, 2, \ldots, N\}$.

Observe que um vetor coluna $\mathbf{v}$ é um vetor de probabilidades correspondente a alguma distribuição se as suas entradas forem não negativas e somarem 1, ou seja, se

(i) $\mathbf{v} \geq \mathbf{0_N}$ e;
(ii) $\mathbf{v}^t \mathbf{1_N} = 1$.

Definição 7.1. Diremos que uma matriz A, de N linhas e M colunas é não negativa (respectivamente, positiva) se todas as suas entradas o forem. Em tal caso escreveremos $A \geq \mathbf{0_{N\times N}}$ (respectivamente $A > \mathbf{0_{N\times N}}$).

Uma matriz P quadrada e de dimensão N é estocástica se ela for não negativa e se a soma dos elementos das linhas forem todas iguais a 1, ou seja, se

(i) $P \geq \mathbf{0_{N\times N}}$, e;
(ii) $P \cdot \mathbf{1_N} = \mathbf{1_N}$.

Em outras palavras, P será uma matriz estocástica se todas as sua linhas forem vetores de probabilidades. Note também que a condição (ii) equivale a dizer que $\lambda = 1$ é autovalor de P com autovetor direito associado $\mathbf{1_N}$.

Se, para cada $n \geq 0$, representarmos a distribuição de probabilidade de X_n através do vetor π_n, teremos que

$$\begin{aligned} \pi_n^t &= \pi_{n-1}^t P, \\ &= \pi_0^t P^n, \end{aligned}$$

para qualquer $n \geq 1$.

Também sabemos que uma distribuição invariante pode ser representada por um vetor π_{inv} satisfazendo

$$\pi_{inv}^t = \pi_{inv}^t P.$$

Ou seja, distribuições invariantes correspondem a vetores de probabilidades que são autovetores esquerdos da matriz P associados ao autovalor $\lambda = 1$.

Exemplo 7.2. *No R para as matrizes dos exemplos 5.13.5.18*
calcular todos os autovalores
calcular autovetores esquerdo e direito associados a 1

Neste exemplo, constatamos que efetivamente, $\lambda = 1$ é um autovalor de P. Além disto, observe que nenhum autovalor tem módulo maior que um. Isto vale de forma geral para matrizes não negativas e é consequência do seguinte resultado sobre a estrutura do espectro deste tipo de matrizes.

Teorema 7.3 (Teorema de Perron-Frobenius para matrizes não negativas).
Seja $A = (a_{ij})$ uma matriz quadrada não negativa. Então as seguintes afirmações valem,

(i) Existe um autovalor real λ_{PF}, tal que qualquer outro autovalor λ satisfaz $|\lambda| \leq \lambda_{PF}$. λ_{PF} é chamado de autovalor de Perron-Frobenius.
(ii) Existe um autovetor esquerdo (respectivamente, direito) associado a λ_{PF} e com todas as entradas não negativas.
(iii) λ_{PF} satisfaz

$$\min_i \sum_j a_{ij} \leq \lambda_{PF} \leq \max_i \sum_j a_{ij}, \tag{7.17}$$

ou seja, λ_{PF} é maior que a menor soma de linhas e menor que a maior delas.

Aplicando este resultado a matrizes estocásticas, temos que (7.17) implica que para tais matrizes, $\lambda_{PF} = 1$. Além disto, se escolhermos um autovetor esquerdo não negativo $\mathbf{v}$, cuja existência é garantida por (ii), normalizado de forma que $\mathbf{v}^t \mathbf{1_N} = 1$, poderemos concluir que

Proposição 7.4. *Toda cadeia de Markov com espaço de estados finito possui pelo menos uma distribuição invariante.*

Vimos no exemplo 5.6 que o passeio aleatório simples simétrico nos inteiros não possui distribuição invariante, portanto, o resultado acima **não vale em dimensão infinita**.

Na seção 5 vimos que as noções de irredutibilidade e de periodicidade eram muito importantes na determinação do comportamento assintótico das cadeias de Markov. Por esta razão, noções análogas são utilizadas também para matrizes estocásticas.

Definição 7.5. Chamaremos de matriz estocástica irredutível a toda matriz P estocástica tal que para todo par de estados $i, j \in E$, com $i \neq j$, existe $n = n(i,j) \geq 1$ tal que $P_{ij}^{(n)} > 0$.

Em outras palavras, as matrizes estocásticas irredutíveis são as matrizes de transição das cadeias iredutíveis. Analogamente, pode ser definido o período das matrizes estocásticas irredutíveis e em particular, as matrizes estocásticas irredutíveis aperiódicas.

Definição 7.6. Chamaremos de matriz estocástica primitiva a toda matriz P estocástica tal que exista alguma potência $k \in \mathbb{N}$ tal que $P^n > \mathbf{0_{N\times N}}$, para todo $n \geq k$.

Observe que as duas definições são diferentes. Para uma matriz ser primitiva estamos pedindo que as suas potências sejam positivas **a partir de algum expoente** k. Por outro lado, uma matriz irredutível poderia ter por exemplo, $P_{ii}^{(n)} = 0$, para uma sequência infinita de valores de n, como por exemplo, a matriz

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

ou em geral como ocorre com toda matriz periódica. É possível provar que uma matriz estocástica é primitiva se e somente se ela for e irredutível e aperiódica (veja o exercício 31). Desta forma, o teorema básico sobre a convergência de cadeias de Markov (teorema 5.13), pode ser reescrito no caso de espaço de estados finito, da forma seguinte.

Teorema 7.7 (Teorema básico da convergência de cadeias de Markov com espaço de estados finito).

Toda cadeia de Markov com matriz de transição primitiva, possui distribuição assintótica.

Ainda nesta seção, veremos que este teorema é consequência da estrutura espectral das matrizes estocásticas primitivas.

Note que toda matriz estocástica positiva é primitiva. Isto nos dá um critério muito simples para concluir a existência da distribuição assintótica.

Corolário 7.8. *Toda cadeia de Markov com espaço de estados finito e matriz de transição positiva possui distribuição assintótica.*

Versões do teorema 7.3 para matrizes irredutíveis e para matrizes primitivas (veja [?], por exemplo) permitem formular o seguinte resultado.

Teorema 7.9. *Suponha que P é uma matriz estocástica irredutível. Então,*

(i) $\lambda_{PF} = 1$ *é raiz simples do polinômio característico;*
(ii) existe um autovetor esquerdo associado a $\lambda_{PF} = 1$*, com todas as entradas positivas;*
(iii) os autovetores esquerdo e direito associados a $\lambda_{PF} = 1$*, são únicos salvo constantes multiplicativas;*
(iv) qualquer outro autovalor $\lambda \neq 1$ *satisfaz que* $1 \geq |\lambda|$*, sendo* $1 > |\lambda|$ *se* P *for primitiva;*
(v) se P *for periódica com período* $d > 1$*, há exatamente* d *autovalores com módulo um, que são todas as raízes de* $\lambda^d = 1$*.*

Este teorema é extremamente poderoso e permite tirar várias conclusões importantes.

Proposição 7.10. *Toda cadeia de Markov irredutível e com espaço de estados finito possui uma única distribuição invariante. Tal distribuição aloca probabilidades positivas em todos os estados.*

No exercício **??**, propomos analisar melhor o problema da unicidade das distribuições invariantes. Indicamos ai como concluir que quando o espaço de estados é finito, a existência de uma única classe irredutível é necessária e suficiente para a unicidade da distribuição invariante.

O teorema 7.9 fornece uma outra maneira de calcularmos o período de uma cadeia irredutível com espaço de estados finito. De fato, observe que podemos simplesmente determinar quantas raízes diferentes e com módulo um tem o polinômio característico. Além disto, se $\lambda_{PF} = 1$ for a única raiz com módulo um deste polinômio, a matriz deve ser primitiva.

Exemplo 7.11. *O polinômio característico da matriz*

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

é $p(\lambda) = \lambda^2 - 1$, *confirmando o fato dela ser periódica de período* $d = 2$. *O único autovetor esquerdo* $\mathbf{v}$ *satisfazendo* $v_1 + v_2 = 1$ *foi calculado em (**??**), ou seja,* $\mathbf{v} = \pi_{inv} = \left(\frac{1}{2}\,\frac{1}{2}\right)^t$.

Exemplo 7.12. *Vimos que a cadeia do exemplo 6.19 era irredutível e aperiódica de período três. O polinômio característico da matriz de transição correspondente é* $p(\lambda) = \lambda^3(\lambda^3 - 1)$, *cujas raízes são* 0 *e as três raízes cúbicas de* 1, *confirmando o fato do período da cadeia ser três.*

A teoria espectral de matrizes não negativas, permite também estudar o problema da existência da distribuição assintótica. Em virtude dos resultados apresentados no final da seção 5 e do teorema 7.7, vemos que para uma matriz estocástica e primitiva P, com distribuição invariante π_{inv} temos que estudar a diferença entre as matrizes P^n e $\mathbf{1_N} \cdot \pi_{inv}^t$, para $n \to \infty$. Começaremos por um exemplo muito simples de uma matriz P primitiva e diagonalizável.

Exemplo 7.13.

Para matrizes primitivas não necessariamente diagonalizáveis, podemos usar um resultado correspondente também à teoria de Perron-Frobenius. Começaremos por fixar a notação.

Para uma matriz estocástica primitiva P, denotaremos como $1, \lambda_1, \dots, \lambda_k$ a todos os seus autovalores distintos, com multiplicidades $1, m_2, \dots, m_k$, respectivamente. Assumiremos também que $1 \geq |\lambda_1| \geq \cdots \geq |\lambda_k|$, e que se $|\lambda_i| = |\lambda_{i+1}|$ para algum $1 < i < k$, então $m_i \geq m_{i+1}$ para $1 < i < k$.

Exemplo 7.14. *matriz primitiva, autovalores, autovetores, multiplicidades*

Teorema 7.15. *Para uma matriz estocástica primitiva* P, *valem*

(i) *Se* $\lambda_2 = 0$, *então*
$$P^n = \mathbf{1_N} \cdot \pi_{inv}^t, \tag{7.18}$$
para $n \geq N - 1$.

(ii) *Se* $\lambda_2 \neq 0$, *então existem constantes* $c > 0$ *e* $0 < \alpha < 1$ *tais que quando* $n \to \infty$,
$$|P_{i,j}^n - \pi_{inv}(i)| \leq c\alpha^n, \tag{7.19}$$
para todos $i, j \in E$. *Além disto, a constante* α *pode ser escolhida como*
$$\alpha = \begin{cases} |\lambda_2| & , \quad se\ m_2 = 1, \\ |\lambda_2| + \epsilon, \ para\ qualquer\ \epsilon > 0 & , \quad se\ m_2 > 1. \end{cases}$$

Exemplo 7.16. $\lambda_2 = 0$

Exemplo 7.17. *exemplos, grafico de* $max(P_1^n - \mathbf{1_N} \cdot \pi_{inv}^t)$ *v.s.*$max(P_2^n - \mathbf{1_N} \cdot \pi_{inv}^t)$.

8. 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 8.1. 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 8.2. *(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

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

$$\begin{aligned} R(i,j) &= \mathbb{E}_i(N_j) \\ &= \sum_{m=1}^{\infty} mF(i,j)F(j,j)^{m-1}(1 - F(j,j)). \end{aligned} \tag{8.20}$$

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

$$P_i(N_j = \infty) = 0, \quad \text{pois } F(j,j) < 1.$$

$$\begin{aligned} R(i,j) &= F(i,j)\sum_{m=1}^{\infty} mF(j,j)^{m-1}(1-F(j,j)) \\ &= \frac{F(i,j)}{1-F(j,j)} \end{aligned} \tag{8.21}$$

(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.

□

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 consequeñcia disto, vale o seguinte resultado.

Proposição 8.3. *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 $E=\{1,2,\ldots,M\}$ 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 existitirá 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 $N=\max\{N_1,N_2,\ldots,N_M\}$. Pela construção feita chegamos a que X_{N+1} não pode ser nenhum dos estados $\{1,2,\ldots,M\}$, 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. □

4. Qual é a falha da prova ”informal”apresentada acima?

Existem cadeias com espaço de estados infinito sem estados recorrentes (veja o exemplo 8.10). Algumas conseqüências do teorema 8.2 são as seguintes.

Proposição 8.4. *Um estado j do espaço de estados de uma cadeia de Markov é transitório se e somente se $\sum_{n=0}^{\infty} P^n_{j,j}<\infty$.*

Demonstração. Isto é somente uma reescritura da relação

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

que vale para j transitório. □

Proposição 8.5. *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) $lim_{n\to\infty} P_{i,j}^n = 0$.

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 8.6. *Se o estado j é recorrente e $j \to k$ então*

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

Demonstração. (1) Se $j \to k$ 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 $k \to j$.

(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 $P^{n_1+n+n_0}(k,k) \geq P_{k,j}^{n_1} P_{j,j}^{n} P_{j,k}^{n_0}$. 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 8.7. *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 8.8. *Se a classe $\mathcal{C}$ é irredutivel e finita todos os estados são recorrentes.*

A proposição 8.6 afirma que se um estado $j \in E$ é acessível desde um estado recorrente i, então necessariamente $j \to i$, ou seja, i é um estado essencial, em outras palavras,

Corolário 8.9. *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 8.10. *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 8.4, 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 8.11. *Considere a cadeia de Markov do exemplo 6.12, a classe* $\mathcal{C} = \{a, c, e\}$ *é finita e irredutível, logo todos os seus estados são recorrentes. Observe que os estados que estam fora de* $\mathcal{C}$ *são não essenciais e portanto, transitórios.*

Teorema 8.12. *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 8.13. *A cadeia de Markov $X_n \in E = \{1, 2, \ldots, 10\}$ 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:

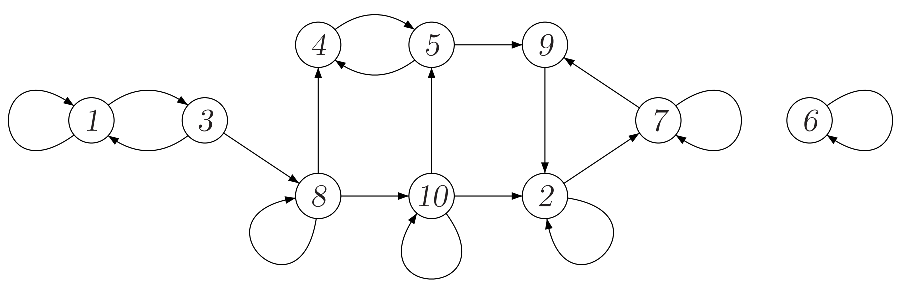


Figura 17.

Temos então três 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.

9. Probabilidades de Absorção

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

Definição 9.1. 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 (**??**) 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 9.2. *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:

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

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:*

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

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.

10. Exemplos

10.1. O processo de nascimento e morte. Para a cadeia do Exemplo (2.5), 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 $u(k) = 1$.

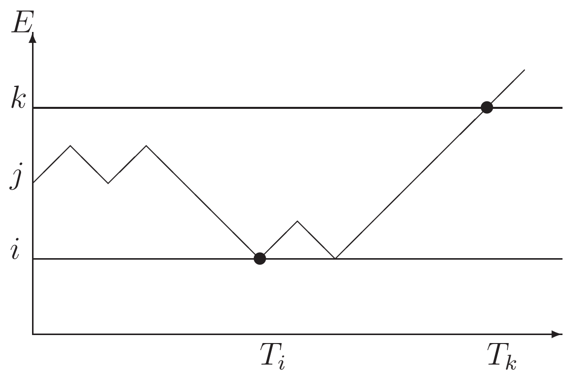


FIGURA 18. Processo de vida e morte.

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:

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

Seja $\gamma_j = \frac{q_j q_{j-1} \cdots q_1}{p_j p_{j-1} \cdots p_1}$, 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 10.1. *Num jogo as apostas que faz um jogador valem um real, ele ganha um real ou perde a aposta com probabilidade $p = 9/19$, $q = 10/19$ respetivamente. 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:*

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

logo

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

Consideremos o caso infinito $d = \infty$, suponha que, para todo i, $p_i > 0$ e $q_i > 0$ logo a cadeia é irredutivel (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 seram recorrentes ou transitórios. Primeiro observe que:

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

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

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

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.

Resumindo, o processo de vida e morte é recorrente se e somente se

$$\sum_{j=0}^{\infty} \gamma_j = \sum_{j=0}^{\infty} \frac{q_j q_{j-1} \cdots q_1}{p_j p_{j-1} \cdots p_1} = \infty. \tag{10.22}$$

Exemplo 10.2. *Considere o processo de vida e morte dado pela probabilidades:*

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

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,

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

logo X_n é transitório.

10.2. O processo de ramificação. O processo $\{X_n\}$ do exemplo 2.4, 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 comença com i pais, então a probabilidade da população ficar extinta sería: $F(i,0) = P_i(T_0 < \infty)$. Usando o fato que a descendência de cada pai é independente concluimos 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 individuo. Lembre que $X_{n+1} = \sum_{k=1}^{X_n} Z_n^{(k)}$ onde as v.a. $Z_n^{(k)}$ representabam 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, $X_n \to 0$ ou $X_n \to \infty$, que correspondem respectivamente a $\rho = 1$ ou $\rho < 1$. Queremos achar condições necessarias 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 na figura (19(a)). 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 figura (19(b)). Logo, $M_\xi(\rho) = \rho$, $\rho < 1$ e $P(X_n \to \infty) > 0$.

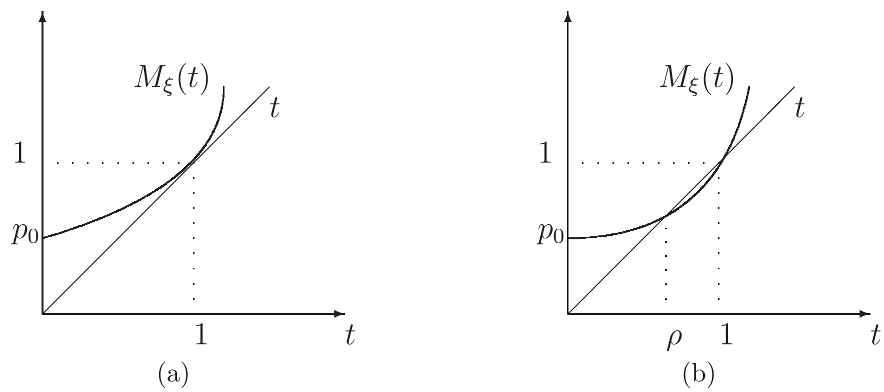


FIGURA 19.

EXEMPLO 10.3. *Suponha que numa população cada homen de uma familia tem 3 filhos, e que a probabilidade de que cada um deles seja homen 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

$$(\frac{1}{2} + \frac{1}{2}t)^3 = t, \quad \Longrightarrow (t-1)(t^2 + 4t - 1) = 0$$

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

10.3. Filas. Estudaremos a cadeia X_n que representa o tamanho da fila de clientes que esperan ser atendidos, após o n-ésimo serviço (Exemplo 2.1). 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 distribuidas, 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

$$F(0,0) = P_{0,0} + \sum_{j=1}^{\infty} P_{0,j} F(j,0) \Rightarrow \rho = p_0 + \sum_{j=1}^{\infty} p_j F(j,0) \tag{10.23}$$

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

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

isto é

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

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,

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

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 ultima equação e (10.23) obtemos $\rho = p_0 + \sum_{j=1}^{\infty} p_j \rho^j = M_\xi(\rho)$. A partir de $\rho = M_\xi(\rho)$ é possivel deducir, como no caso do processo de ramificação, quais estados são recorrentes e quais são transitórios:

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

11. Comportamento assintótico e distribuição invariante (caso geral)

Precisaremos do seguinte resultado,

Proposição 11.1.

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

Aqui $\rho_{\mathcal{C}(j)}(i)$ 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 conseqüência do seguinte resultado que caracteriza o comportamento assintótico de uma cadeia de Markov.

Lema 11.2. *Seja j um estado recorrente aperiódico.*

(1) Se j é positivo,

$$P^n_{i,j} \to \frac{F(i,j)}{\mathbb{E}_j(T_j)},\ n \to \infty.$$

(2) Se j é nulo,

$$P^n_{i,j} \to 0,\ n \to \infty.$$

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

Proposição 11.3. *Se i é recorrente nulo e $i \leftrightarrow j$ então j também é recorrente nulo.*

Demonstração. □

Proposição 11.4. *Uma classe irredutível finita não possui estados recorrentes nulos.*

Demonstração. Seja $\mathcal{C}$ uma classe recorrente irredutível e finita de uma cadeia de Markov. Pela proposição anterior, todos os seus estados são ou recorrentes nulos ou recorrentes positivos. Seja $P^\infty_{i,j} = \lim_{n\to\infty} P^n_{i,j}$. Como $1 = \sum_{j\in\mathcal{C}} P^n_{i,j} = \sum_{j\in E} P^\infty_{i,j}$, os termos $P^\infty_{i,j}$, $j \in \mathcal{C}$ não podem ser todos nulos. Pelo lema 11.2, os estados de $\mathcal{C}\mathcal{C}$ devm ser todos recorrentes positivos. □

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

Proposição 11.5. *Uma cadeia de Markov aperiódica com uma única classe irredutível recorrente positiva possui distribuição limite.*

Demonstração. Pelo lema 11.2 obtemos que

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

□

Proposição 11.6. *Uma cadeia de Markov aperiódica finita com uma única classe irredutível possui distribuição limite.*

Demonstração. Numa cadeia de Markov aperiódica finita as classes irredutíveis são recorrentes positivasAplicando o resultado anterior, o nosso aqui segue. □

Proposição 11.7. *Uma cadeia de Markov finita e tal que existe algum $n_0 \in \mathbb{N}$ para o qual*

$$\min P^{n_0}_{i,j\in E} > 0,$$

possui distribuição limite.

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

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

Proposição 11.8. *Seja X_n uma cadeia de Markov com matriz de transição P e tal que os limites*

$$\lim_{n\to\infty} P^n_{i,j} = \pi(j) \tag{11.24}$$

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

$$\sum_{i\in E} \pi(i)P_{i,j} = \pi(j). \tag{11.25}$$

*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 (5.11).*

11.1. 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 posivel voltar ao mesmo estado num numero impar de transições, em particular $P^n_{i,i} = 0$ se n é impar e $P^n_{i,i} \neq 0$ se n é par. Logo não existe $\lim\limits_{n\to\infty} P^n_{i,i}$.

Para resolver este problema observe que vale o seguinte resultado:

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

para toda sequencia de numeros a_n. Assim, se $\lim\limits_{n\to\infty} P^n_{i,j} = \pi(j)$, então $\lim\limits_{n\to\infty} \frac{1}{n}\sum\limits_{k=1}^{n} P^k_{i,j} = \pi(j)$.

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

TEOREMA 11.9. *Cosidere uma cadeia de Markov irredutivel X_n, finita; e seja $f : E \mapsto R$ uma função limitada, então vale que:*

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

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:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P^k_{i,j} = \pi(j)$$

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

Um outra manera de expresar o teorema ergodico é decir que a media temporal da cadeia é igual a media espacial. Esta idea é a base dos algoritmos usados nos metodos de **MCMC** (Monte Carlo Markov Chains) que se usam, por exemplo, na estatistica Bayesiana. Onde, para obter amostras de uma certa distribuição construimos uma cadeia de Markov que tenha como distribuição invariante a distribuição que estamos procurando.

Usando o teorema ergodico é possivel cracterizar a distribuição invariante. Para isso observe que:

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

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

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

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

$$R(i,j) = \lim_{n\to\infty}\sum_{k=1}^{n} R_n(i,j) = \lim_{n\to\infty}\sum_{k=1}^{n} P^k_{i,j}$$

onde $R(i,j)$ e 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 invariante.

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

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

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

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

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

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

Em particular,

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

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

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

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 11.10. *Uma cadeia de Markov finita e irredutivel tem todos seus estados recorrentes positivos.*

Exemplo 11.11. **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 irredutivel e tem matriz de transição:*

$$P = \begin{bmatrix} p_0 & p_1 & p_2 & p_3 & \cdots \\ p_0 & p_1 & p_2 & p_3 & \cdots \\ 0 & p_0 & p_1 & p_2 & \cdots \\ 0 & 0 & p_0 & p_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Observe que se $p_0 = 0$ então o estado 0 é transitório. Também, se $p_0 + p_1 = 1$ então $p_2 = p_3 = \cdots = 0$, 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 $\mu \le 1$ então X_n é recorrente.*

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

É possivel 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* $\pi(0) = 1 - \mu$.

12. Exercícios

5. Para quais valores de x, y e z as matrizes

$$\begin{pmatrix} 0,5 & 0,1 & x \\ y & 0,2 & 0,4 \\ 0,3 & z & 0,1 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 0,3 & x & x \\ 0,4 & 0,2 & 0,4 \\ 0,2 & 0,9 & z \end{pmatrix}$$

são matrizes estocásticas?

6. 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 sua matriz de transição.

7. Para o número dos sucessos no processo de Bernoulli,

(1) Represente a topologia do processo N_n.
(2) Reobtenha a distribuição das N_n usando (3.9).

8. Observe que o processo dos tempos dos sucessos num processo de Bernoulli também é um caso particular do exemplo (1.6). Determine a matriz de transição e a topologia deste processo. Sugestão: Faça $\xi_j = T_j - T_{j-1}$.

9. Prove que toda matriz estocástica com as linha iguais é a matriz de transição de uma sequência de variáveis aleatórias i.i.d.s.

10. No exemplo 2.2, suponha que $P(Z_n = 1) = p = 1 - P(Z_n = 2)$. Comprove que X_n é uma cadeia de Markov e ache a sua matriz de transição P.

11. Comprove que o processo do exercício 6 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 $P(Z_1 = 3) = 0.1$.

12. Os resultados de uma pesquisa revelam que 80% das filhas de mulheres trabalhadoras trabalham, enquanto só 30% das filhas de mulheres que não trabalham, o fazem. Esta situação pode ser modelada por uma cadeia de Markov $X = \{X_n\}_{n\in\mathbb{N}}$ onde X_n representa o estatus laboral de uma mulher na geração n, isto é, o espaço de estados é $E = \{T, NT\}$ (T=mulher que trabalha, NT=mulher que não trabalha).

(a) Escreva a matriz de transição para este modelo.
(b) Se numa época determinada a proporção de mulheres trabalhadoras é de 30%, qual será a proporção das netas destas mulheres que trabalham?
(c) Calcule a probabilidade de que uma mulher e a sua avó trabalhem.
(d) A longo prazo qual será a proporção de mulheres trabalhadoras?

13. Uma seguradora de carros classifica os seus clientes em três classes: não desejáveis, satisfatórios e preferenciais. A classificação de um cliente pode mudar de um ano para o outro. Não é possível passar de preferencial a não desejável nem viceversa. 40 %

dos segurados não desejáveis viram satisfatórios, 30% dos satisfatórios vira preferencial enquanto 10 % deles viram não desejáveis e 20% dos preferenciais viram satisfatórios.

Podemos representar esta situação usando uma cadeia de Markov $X = \{X_n\}_{n\in\mathbb{N}}$ onde X_n representa a classificação de um cliente no ano. Temos portanto, $E = \{$não desejável, satisfatório, preferencial$\}$.

(a) Escreva a matriz de transição do modelo.
(b) Represente a topologia da cadeia.
(c) Calcule a probabilidade de um cliente preferencial continuar a sê-lo no próximo ano e virar satisfatório no ano seguinte.
(d) Calcule a probabilidade de um cliente não desejável virar satisfatório depois de dois anos.
(e) A longo prazo quais serão as proporções de clientes em cada uma das classes?

14. Um professor tem duas lâmpadas na sua garagem. Quando as duas se queimam, elas são trocadas de forma tal que no começo do dia seguinte haverão duas lâmpadas funcionando. Suponha que quando as duas funcionam, exatamente uma delas para de funcionar com probabilidade $0,02$. No entanto, se há somente uma funcionando, ela se queimará com probabilidade $0,05$. A longo prazo qual é a fração do tempo que haverá exatamente uma lâmpada funcionando?

15. (Passeios aleatórios com distribuição invariante) No exemplo 5.6 mostramos um passeio aleatório nos inteiros sem distribuição invariante. Veremos agora variações deste exemplo para as quais existe a distribuição invariante. Em todos os casos a seguir estaremos supondo que X é uma cadeia de Markov definida nos inteiros.

(a) Prove que para $p \in (0,1)$, se $P_{i,i+1} = p$, quando $i \neq -1$ e $P_{i,0} = 1-p$ então a X tem uma única distribuição invariante. Calcule-a.
(b) Prove o mesmo assumindo que as probabilidades de transição são

$$\begin{aligned} P_{i,i+1} = 0.4 \text{ e } P_{i,i-1} = 0.6 \text{ para } i > 0,\\ P_{i,i+1} = 0.6 \text{ e } P_{i,i-1} = 0.4 \text{ para } i < 0,\\ P_{0,1} = 0.6 \text{ e } P_{0,-1} = 0.5. \end{aligned}$$

(c) Observe que ambos casos...Mais geralmente, prove que ...

16. Para a cadeia de Markov com espaço de estados $E = \{0, 1, 2\}$ e matriz de transição

$$P = \begin{array}{c} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \end{array} \begin{array}{c} \begin{array}{ccc} \mathbf{0} & \mathbf{1} & \mathbf{2} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1/3 & 2/3 \end{pmatrix} \end{array}$$

(a) Represente a topologia da cadeia.
(b) Prove que a cadeia é irredutível.
(c) Calcule o período
(d) Quantas distribuições invariantes existem? Por quê? Calcule-as.

17. Considere uma cadeia de Markov com matriz de transição

$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array}\begin{array}{c} \begin{array}{ccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 1/2 & 0 & 1/8 & 1/4 & 1/8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix} \end{array}$$

(a) Represente a topologia da cadeia.
(b) Classifique os estados em transitórios ou recorrentes positivos ou nulos.
(c) Determine todas as classes irredutíveis. A cadeia é irredutível? Por quê?
(d) Determine todos os conjuntos fechados.
(e) Esta cadeia possui algum estado periódico?

18. Considere as cadeias de Markov com as seguintes matrizes de transição:

$$P_1 = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \end{array}\begin{array}{c} \begin{array}{ccc} \mathbf{1} & \mathbf{2} & \mathbf{3} \end{array} \\ \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix} \end{array} \qquad P_2 = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \end{array}\begin{array}{c} \begin{array}{cccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

$$P_3 = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \end{array}\begin{array}{c} \begin{array}{ccccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} \end{array} \\ \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix} \end{array} \qquad P_4 = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \end{array}\begin{array}{c} \begin{array}{ccccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} \end{array} \\ \begin{pmatrix} 1/4 & 3/4 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1/3 & 2/3 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

(a) Represente as topologias correspondentes.
(b) Determine as classes irredutíveis e classifique os estados em transitórios ou recorrentes positivos ou nulos.
(c) Comprove que a cadeia com matriz de transição P_2 é irredutível e periódica com período 3.
(d) Para as cadeias com matrizes de transição P_1, P_3 e P_4 calcule o limite quando $n \to \infty$ das correspondentes matrizes de transição em n passos.
(e) Diga em cada caso se existe a distribuição estacionária e explique a sua resposta.

19. Para a cadeia de Markov com espaço de estados $E = \{0, 1\}$ e matriz de transição

$$P = \begin{array}{c} \\ \mathbf{0} \\ \mathbf{1} \end{array}\begin{array}{c} \begin{array}{cc} \mathbf{0} & \mathbf{1} \end{array} \\ \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \end{array}$$

calcule $P_0(T_0 = n)$ e $\mathbb{E}_0(T_0)$.

20. Considere a cadeia de Ehrenfest com 5 bolas. Suponha que inicialmente todas as bolas estão na primeira caixa.

(a) Calcule a esperança do número de transições necessárias para que o sistema volte a estar nesse estado.
(b) Dê uma aproximação da distribuição do número de bolas na primeira caixa depois de 1000 transições.
(c) Dê uma aproximação da distribuição do número de bolas na primeira caixa depois de 1001 transições.

21. Uma partícula se move seguindo uma cadeia de Markov com espaço de estados $E = \{1, 2, 3, 4\}$ da forma seguinte. Saindo de 1 ou 2 ela visita algum dos estados 3 ou 4 escolhido ao acaso e viceversa, saindo de 3 ou 4, algum dos estados 1 ou 2 será visitado com a mesma probabilidade.

(a) Escreva a matriz de transição desta cadeia e prove que ela é irredutível.
(b) Calcule a distribuição invariante da cadeia e interprete o resultado obtido.
(c) Esta cadeia pode ser generalizada como segue. Consideramos agora como estaço de estados $E = \{1, 2, \ldots, c+d\}$ on de c e d são dois números naturais. Saindo de algum dos primeiros c estados a cadeia vai para um outro escolhido ao acaso entre os d últimos e saindo de algum destes d últimos estados ela visita algum dos c primeiros escolhido ao acaso. Escreva a matriz de transição para esta nova cadeia, prove que é irredutível. Qual você esperaria que fosse a sua distribuição invariante? Tente calculá-la analíticamente para comprovar se a sua intuição estava certa.

22. Sabemos que uma matriz de transição P com todas as linhas iguais corresponde a uma sequência de variáveis i.i.d.s.

(a) Prove que $P^n = P$, para todo $n > 1$, ou seja, a matriz de transição de ordem n coincide com a matriz de transição da cadeia.
(b) Suponha que para a cadeia de Markov $\{X_n\}_{n\geq 0}$, existe um inteiro k tal que a matriz de transição de ordem k tem todas as linhas iguais. Prove que isto tem que valer também para todas as matrizes de transição de ordem superior a k, ou seja, $P^m = P^k$ para todo $m \geq k$.

Este último resultado tem a aplicação seguinte. Suponha que você está querendo obter aproximadamente a distribuição invariante de uma cadeia com muitos estados e para isso você decide calcular computacionalmente potências da matriz de transição. Se para a potência 5, por exemplo, você obtiver uma matriz com todas as linhas iguais, você pode parar o seu procedimento, e retornar o vetor linha dessa matriz como a distribuição invariante, pois não melhorará nada mais calculando potências superiores.

23. Uma matriz P é dita **duplamente estocástica** se ela for estocástica e a sua transposta também o for (isto é, se as somas dos elementos das colunas também são iguais a 1).

(a) Prove que todos os estados de uma cadeia de Markov com matriz de transição duplamente estocástica são essenciais.
(b) Suponha que uma cadeia de Markov X com espaço de estados $E = \{0, 1, ..., M\}$ tem matriz de transição duplamente estocástica primitiva. Prove que a distribuição invariante atribui igual probabilidade a cada estado.
(c) Você pode dar um exemplo (simples!!) de uma matriz de transição duplamente estocástica com mais de uma distribuição invariante?

24. Prove que se uma cadeia de Markov tem mais de uma distribuição invariante, então tem infinitas.
Sugestão: Para qualquer $\alpha \in (0,1)$ vale que se π_1 e π_2 são distribuições invariantes, $\pi_\alpha = \alpha\pi_1 + (1-\alpha)\pi_2$ também o é.

25. Considere uma cadeia de Markov irredutível com matriz de transição P. Prove que se $P > 0$, então todos os estados são aperiódicos. Mostre através de um exemplo que esta condição é necessária, mas não suficiente, ou seja, que existem cadeias de Markov irredutíveis e aperiódicas com alguma probabilidade de transição nula.

26. Seja π_{inv} uma distribuição invariante de uma cadeia de Markov X. Prove que se $\pi_{inv}(i) > 0$ e $i \to j$ então $\pi_{inv}(j) > 0$.

27. Prove que a recorrência positiva é uma propriedade de classe.

28. Considere uma cadeia de Markov irredutível em um espaço de N estados e suponha que X_0 segue a distribuição invariante da cadeia. Defina τ como o primeiro instante de retorno no estado inicial, ou seja, $\tau = \inf\{k > 0 : X_k = X_0\}$. Prove que $\mathbb{E}(\tau) = N$.

29. (Processos espaço-estado) Suponha que $X = \{X_n, n \in \mathbb{N}\}$ é uma cadeia de Markov com espaço de estados $E = \{0, 1\}$ e matriz de transição P e seja K uma variável aleatória com valores naturais e independente de X. Para cada $n \in \mathbb{N}$, fazemos $K_n = K + n$ e chamamos de $Y_n = (X_n, K_n)$.
(a) Prove que Y é uma cadeia de Markov com espaço de estados $F = E \times \mathbb{N}$.
(b) Calcule as probabilidades de transição desta cadeia.
(c) Prove os resultados anteriores supondo que E é finito ou infinito enumerável.
(d) Prove os resultados anteriores supondo que K é condicionalmente independente de $\{X_1, X_2, \dots\}$ dado X_0.

30. Seja $X = \{X_n, n \in \mathbb{N}\}$ um processo estocástico com espaço de estados enumerável E. Suponha que para todo $j \in E$ e $n \in \mathbb{N}, n \geq 2$, vale

$$\mathbb{P}(X_n | X_0, \dots, X_{n-1}) = \mathbb{P}(X_n | X_{n-2}, X_{n-1}),$$

ou sejan neste processo, o estado no futuro depende da história anterior somente através dos estados nos dois últimos instantes. Tal processo é chamado de cadeia de Markov de segunda ordem.

(a) Para cada $n \in \mathbb{N}$ defina $Y_n = (X_n, X_{n+1})$. Prove que o processo $Y = \{Y_n, n \in \mathbb{N}\}$ é uma cadeia de Markov com espaço de estados $F = E^2$. Qual seria a matriz de transição?
(b) Como você definiria cadeias de Markov de ordem k, $k \in \mathbb{N}$?
(c) Prove que o processo X do exemplo **??** não é uma cadeia de ordem k para nenhum valor de k.

31. (a) Prove que uma cadeia com espaço de estados finito é irredutível e aperiódica se e somente se a sua matriz de transição for primitiva.
(b) Use um exemplo para provar que a afirmação análoga não vale no caso de espaço de estados infinito.

32. Considere as seguintes matrices 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 equilibrio do que a matriz $\tilde{P}$, porque? Verificar isso:

(1) Ache os autovalores das matrices e identifique o segundo autovalor em cada caso. Qual é maior?
(2) Ache o autovetor (esquerdo) correspondente ao autovalor 1. En 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

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

e compare com o valor de $|\lambda_2|^n$.
(4) repita o item anterior com $\tilde{P}$ e $\tilde{\pi}$.

CAPíTULO 3

Tópicos adicionais sobre cadeias de Markov.

1. Modelagem através de cadeias de Markov

2. Algoritmos estocásticos

3. Inferência

4. Aplicações

4.1. Mobilidade social.

4.2. Sistema de Bonus-Malus.

A maioria das companhias seguradoras utilizam um sistema de tarifação dos seguros de automoveis no qual os segurados se dividem em l classes ordenadas (grupos tarifários), sendo que o prêmio depende da classe à que pertence o segurado. A cada ano, a classe de cada segurado é determinada com base na classe do ano anterior e na quantidade de indenizações reportadas durante esse ano. Se o segurado não teve indenizações durante um ano, ele vai ser reclassificado numa classe mais baixa, pagando possivelmente um prêmio menor. Caso contrário, o seguradovai para uma classe mais alta, pagando correspondentemente um prêmio possivelmente maior.

Este sistema é chamado de "sistema de Bonus-Malus"e teve a sua origem .... (ver Lemaire)

Os elementos que compõem um sistema de Bonus-Malus são os seguintes:

- as classes de tarifação, numeradas como $1, \dots, l$, sendo que a classe 1 é chamada de "suberbonus"e a l de "supermalus";
- uma escala de prêmios $b = (b_1, b_2, \dots, b_l)$, com $b_1 \le b_2 \le \cdots \le b_l$;
- regras de transição que especificam como passar de uma classe a outra dependendo do número de indenizações durante um ano,

$$k \text{ indenizações} \Rightarrow t_{i,j}(k) = \begin{cases} 1, & \text{se o segurado vai de } i \text{ a } j, \\ 0, & \text{caso contrário.} \end{cases}$$

- uma classe inicial i_0 na qual será alocado cada segurado novo entrando no sistema.

Assumindo que as indenizações reportadas durante um ano formam uma sequência de variáveis aleatórias independentes e identicamente distribuídas, teremos que se escolhermos um segurado ao acaso e chamarmos de X_n à sua classe de tarifação depois de n anos no seguro, a sequência $X = \{X_n, n \ge 0\}$ forma uma cadeia de Markov.

Alguns problemas interessantes neste contexto são os seguintes:

- Probabilidade de no n-ésimo ano estar na classe j.
- Média do prêmio pago ao longo de n anos.
- Estratégia ótima $x = (x_1, \ldots, x_l)$ para um segurado evitar o aumento do prêmio.

CAPíTULO 4

O Processo de Poisson

1. Introdução.

Apresentaremos aqui o processo de Poisson. Este processo estocástico deve o seu nome ao matemático francês Simï¿½on-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*

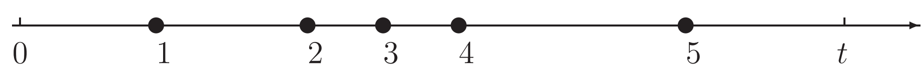


Figura 26

que pode ser também representada como uma trajetória

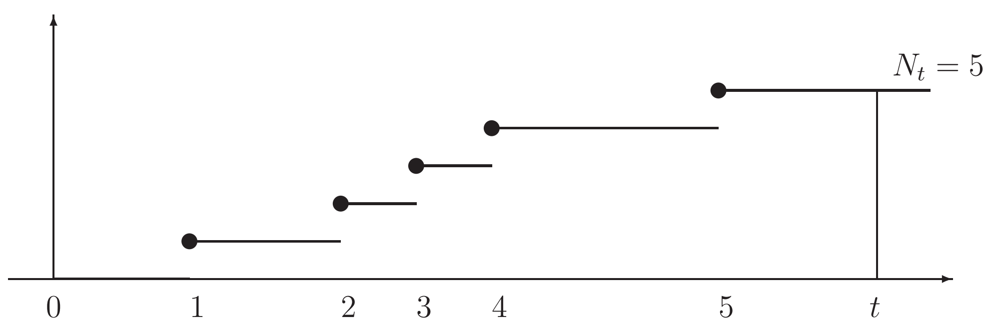


Figura 27

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 ..).

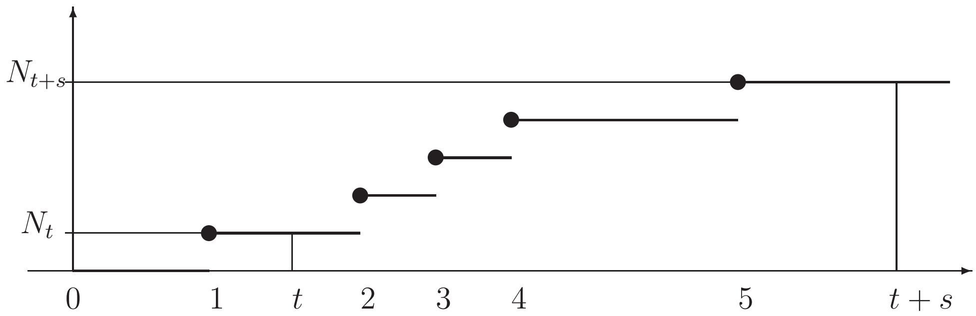


Figura 28

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, $t \to N_t(\omega)$, 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 $T_1, T_2, \ldots$ 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

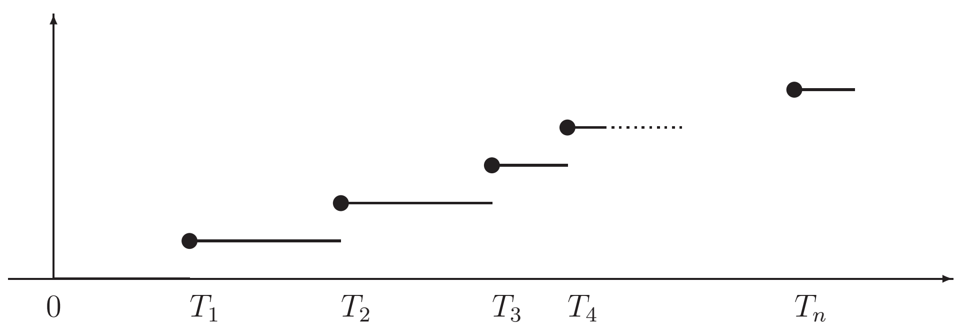


Figura 29

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 $N_{t_1}, N_{t_2}, \ldots, N_{t_n}$ para $t_1 < t_2 < \cdots < t_n = t$ e portanto este incremento será independente do vetor dos incrementos $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}}$. Pelo mesmo argumento obtemos que $N_{t_n} - N_{t_{n-1}}$ deve ser independente do vetor $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_{n-1}} - N_{t_{n-2}}$ 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 detetor 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

$$f(s+t) = f(s)f(t) \tag{2.26}$$

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

Vejamos 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 $t_k = k\frac{t}{n}$, $k \in \{1, \ldots, n\}$. Então

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

e sabemos que $P(N_t - N_{t_{n-1}} \geq 1) = \cdots = P(N_{t_1} \geq 1) = 1$. 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, chegariamos 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 descrescente, 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. □

Proposição 2.4.

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

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

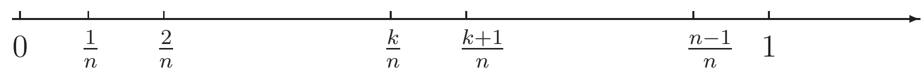


Figura 30

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 $\lim_{n\to\infty} S_n(\omega) = 0$.

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

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

□

Proposição 2.5.

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

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 $\lim_{t\to 0} \frac{1}{t} P(N_t = 1) = \lambda$. □

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*

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

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

Demonstração. Seja $G(t) = \mathbb{E}\left[\alpha^{N_t}\right]$, $0 < \alpha < 1$, 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 obteriamos

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

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

$$\begin{aligned} G(t+s) &= \mathbb{E}\left[\alpha^{N_{t+s}}\right] = \mathbb{E}\left[\alpha^{N_t}\right]\mathbb{E}\left[\alpha^{N_{t+s}-N_t}\right] \\ &= G(t)G(s). \end{aligned} \tag{2.27}$$

Como $G(t) \geq P(N_t = 0) = e^{-\lambda t} > 0$, a equação de Cauchy (2.26) 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,

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

Usando as proposições anteriores obtemos,

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

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

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

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

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:*

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

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:

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

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* $\mathbb{E}\left[N_{t+s} - N_t | N_u, u \leq t\right] = \lambda s$.

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*

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

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 comp$(A_i) = a_i$. (i.e. $b =$ 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*

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

Logo $(N_{A_1}, N_{A_2}, \dots, N_{A_n}) | N_B = k \sim$ 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:

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

□

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 $b =$ comp $(B) >$ 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 assintôtico 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 $n \to \infty$:

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

Podemos generalizar este resultado:

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

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

se $t \to \infty$.

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 assintôtico 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 $t \to \infty$.

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 $n-$éssimo tempo de chegada esta 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

$$\{T_n \le t\} = \{N_t \ge n\}.$$

Assim,

$$\begin{aligned} P(T_n \le t) &= P(N_t \ge n) \\ &= \sum_{k=n}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \end{aligned}$$

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**, $\{T_n\}_{n\ge1}$.

Uma observação importante é que conhecer o processo até o instante T_n, i.e $\{N_t : t \le T_n\}$ é o mesmo que conhecer o processo dos tempos das chegadas até o instante n, i.e. $\{T_1, T_2, \ldots, T_n\}$, isto é fácil de visualizar na seguinte figura.

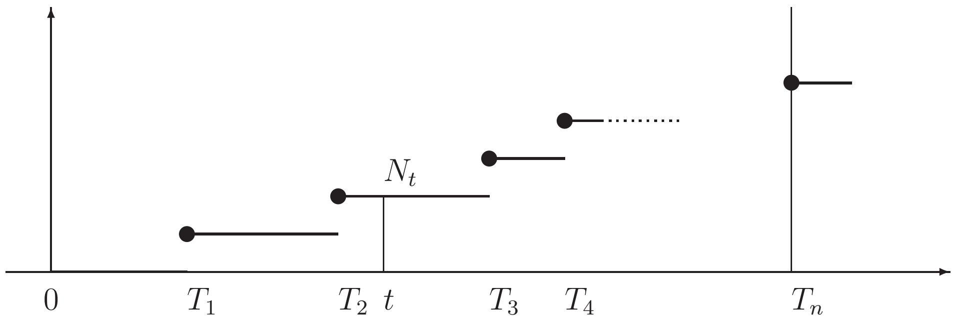


Figura 31

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,

$$\begin{aligned} P(N_{T_n+s} - N_{T_n} = 0 | T_1, T_2, \ldots, T_n) &= P(N_{T_n+s} - N_{T_n} = 0 | N_u, u \leq T_n) \\ &= P(N_{(T_n,T_n+s]} = 0 | N_u, u \leq T_n) \\ &= e^{-\lambda s} \end{aligned}$$

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:*

$$P(T_{n+1} - T_n \leq t | T_1, T_2, \ldots, T_n) = 1 - e^{-\lambda t}$$

Assim o processo $\{T_n : n \geq 1\}$ é estacionário e tem incrementos independentes.

Corolário 3.2. *As variáveis aleatórias $T_1, T_2 - T_1, T_3 - T_2, \ldots$ 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 memoria, 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 $T_1, T_2, T_3, \ldots$ os tempos de chegada num processo de contagem N_t com $t \geq 0$.*

$$\{N_t\}_{t\geq 0} \text{ é um processo de Poisson}(\lambda) \Leftrightarrow \begin{cases} T_{n+1} - T_n, n \geq 0, i.i.d., \\ T_{n+1} - T_n \sim exp(\lambda). \end{cases}$$

Observe que $T_n = T_1 + (T_2 - T_1) + (T_3 - T_2) + \cdots + (T_n - T_{n-1})$. Usando o fato que a soma de distribuicoes exponenciais i.i.d. tem distribuição Gamma podemos concluir que o tempo do $n-$éssimo evento $T_n \sim$ Gamma(n, λ). Logo,

$$\begin{aligned} \mathbb{E}T_n &= \frac{n}{\lambda}, \\ \mathrm{Var}(T_n) &= \frac{n}{\lambda^2}, \\ \mathbb{E}e^{-\alpha T_n} &= \left[\frac{\lambda}{\lambda + \alpha}\right]^n. \end{aligned}$$

A distribuição Gamma(n, λ) 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 $X_1, X_2, \ldots$ os tempos de duração de cada chip que foi trocado. Logo, $P(X_n < t) = 1 - e^{\lambda t}$. Considere $T_1, T_2, \ldots$ os sucessivos instantes nos quais aconteceu uma falha no computador devido a uma falha do chip,

$$\begin{aligned} T_1 &= X_1 \\ T_2 &= X_1 + X_2 \\ &\cdots \\ T_n &= X_1 + X_2 + \cdots + X_n. \end{aligned}$$

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

$$VarX_n = \frac{1}{\lambda^2} = 25 \times 10^6.$$

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 reemplazo 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* $n-$*éssimo 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 $\mathbb{E}C=\sum_{n=1}^{\infty}\beta\mathbb{E}e^{-\alpha T_n}$.

$\mathbb{E}e^{-\alpha T_n}=[\frac{\lambda}{\alpha+\lambda}]^n$ *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,*

$\alpha=\frac{0.24}{365\times 24}=\frac{0.01}{365}$ e $\mathbb{E}C=800\frac{36500}{5000}=5840$ *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 estos tempos são variáveis aleatórias independentes e identicamente distribuídas com distribuição,*

$$P(U_k\le t)=1-e^{-\lambda t}-\lambda te^{-\lambda t},\quad t\ge 0,$$

com densidade:

$$f_{U_k}=\lambda e^{-\lambda t}-[\lambda e^{-\lambda t}+(\lambda^2 t)e^{-\lambda t}]=\lambda^2te^{-\lambda t}.$$

Observe que $U_k\sim$ *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* $T_i, i\ge 1$ *são os tempos das chegadas do processo de Poisson* $\{N_t\}_{t\ge 0}$*. 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* $f\ge 0$*, então vale:*

$$\mathbb{E}\left[\sum_{n=1}^{\infty}f(T_n)\right]=\lambda\int_0^{\infty}f(t)dt.$$

DEMONSTRAÇÃO. Observe que

$$\mathbb{E}\left[f(T_n)\right] = \int_0^\infty f(t)\frac{\lambda e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!}dt.$$

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.

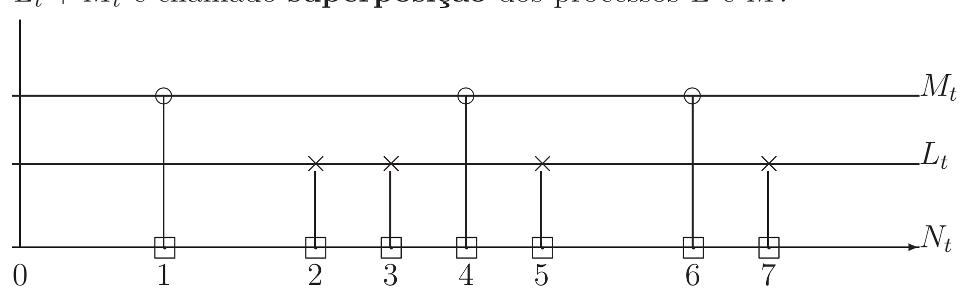


Figura 32

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 passan 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-éssimo 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-éssimo 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 sería o numero de ensaios até o instante t.

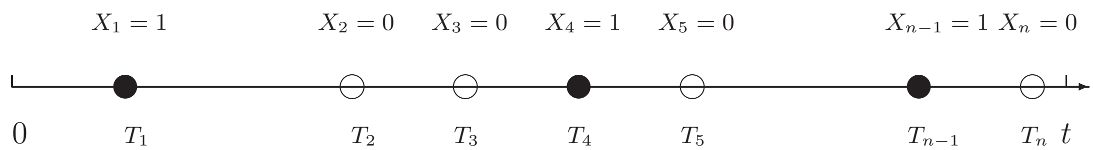


Figura 33

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, sería $L_t = N_t - M_t$.

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.*

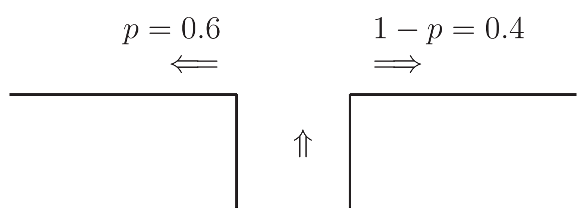


Figura 34

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 vehiculos 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,* {

$$\begin{aligned} \textit{Número esperado de pessoas numa hora} &= \mathbb{E}(1N_1^{(i)} + 2N_1^{(2)} + 3N_1^{(3)} + 4N_1^{(4)} + 5N_1^{(5)}) \\ &= (1 \times 6) + (2 \times 6) + (3 \times 4) + (4 \times 2) + (5 \times 2) \\ &= 48. \end{aligned}$$

}

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 un 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 trajet0́rias têm um numero finito de saltos sobre intervalos finitos;
(2) $\forall t, s \geq 0$, $Z_{t+s} - Z_t$ é independente do pasado $\{Z_u, u \leq t\}$;
(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.

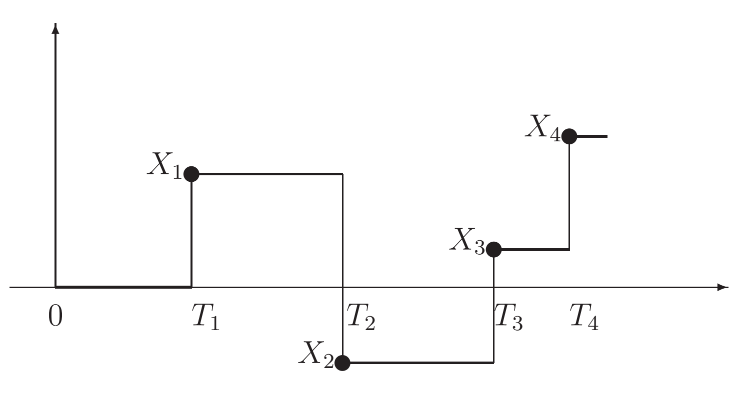


Figura 35

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*-éssimo 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. $X_n = a$): $N_t^{(a)}$.

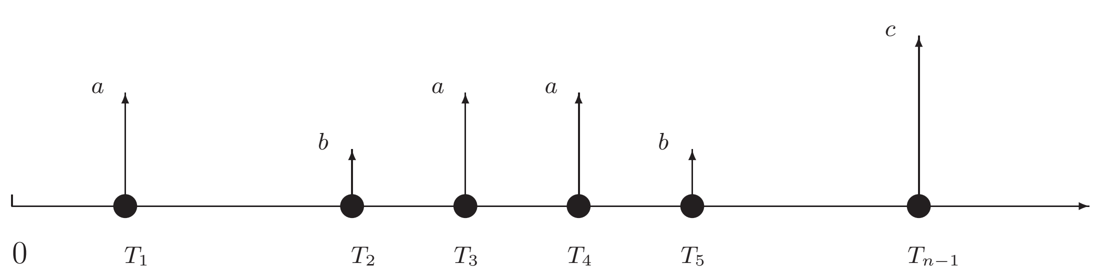


Figura 36

Então vale que:

$$\begin{array}{rcl} N_t^{(a)} & \sim & \text{Poisson com taxa } \lambda(a)t = \lambda P(X=a)t, \\ N_t^{(b)} & \sim & \text{Poisson com taxa } \lambda(b)t = \lambda P(X=b)t, \\ & \dots & \end{array}$$

Logo,

$$\begin{array}{rcl} Z_t & = & \displaystyle\sum_{n=0}^{N_t} X_n \\ & = & \displaystyle\sum_{n:X_n=a}^{N_t^{(a)}} + \sum_{n:X_n=b}^{N_t^{(b)}} + \dots \\ & = & aN_t^{(a)} + bN_t^{(b)} + \dots \end{array}$$

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,

$$\begin{array}{rcl} \mathbb{E}(Z_t) & = & \mathbb{E}(\mathbb{E}(Z_t|N_t)) = \mathbb{E}(N_t\mathbb{E}(X_1)) \\ & = & \mathbb{E}(N_t)\mathbb{E}(X_1) = \lambda t\mathbb{E}(X_1). \end{array}$$

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.

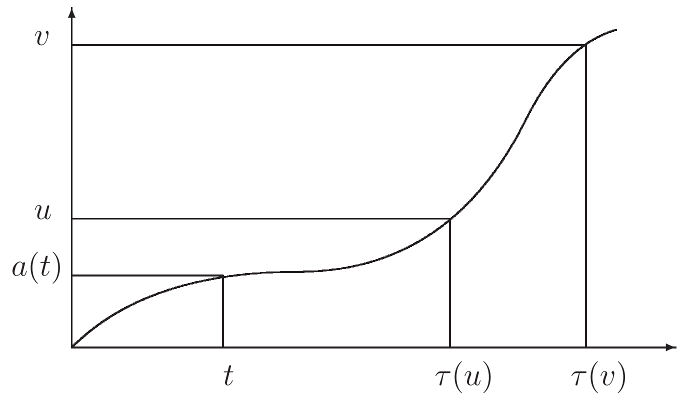


Figura 37

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

$$P\left(N_{t+s} - N_t = k\right) = \frac{e^{-b(t,s)} b(t,s)^k}{k!};$$

(2) se T_n é o tempo do n-éssimo evento do processo, então

$$P(T_{n+1} - T_n > t | T_1, T_2, \ldots, T_n) = e^{-\lambda(a(T_n+t) - a(T_n))}.$$

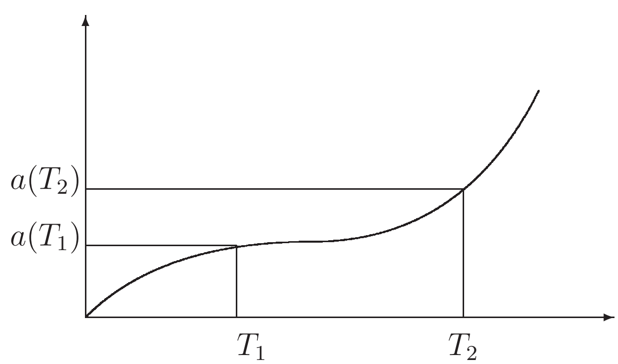


Figura 38

Em geral a taxa do processo é uma função descontínua.

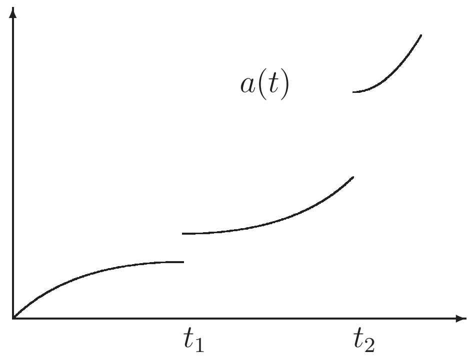


Figura 39

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.

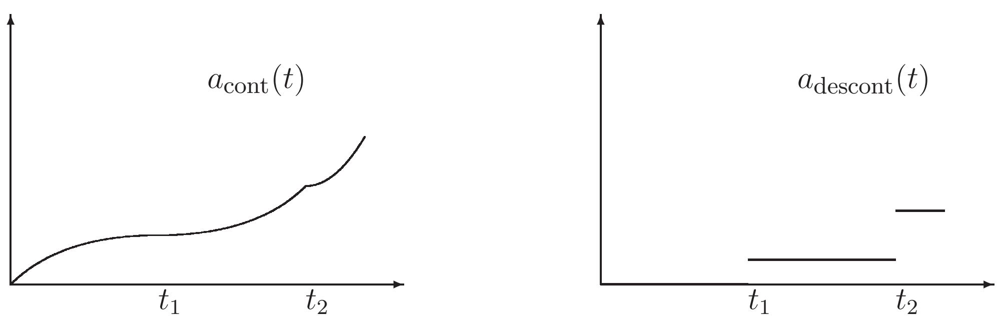


Figura 40

8. Exercícios

33. Suponha que $\{N_t\}_{t\geq 0}$ é um processo de Poisson de taxa 5 e que $\{T_n\}_{n\geq 1}$ é o correspondente processo dos tempos das chegadas.

(a) Interprete os seguintes eventos (p.ex, $N_{1,4} = 5$: houve 5 ocorrências até o instante $t = 1,4$).
   (i) $\{N_{1,4} = 0\}$;
   (ii) $\{N_{4,5} - N_{3,4} = 2\}$;
   (iii) $\{N_5 = 3, N_{10} = 5, N_{15} = 7\}$;
   (iv) $\{N_5 = 3, T_4 = 5,4\}$;
   (v) $\{T_1 = 3, T_4 = 5,4\}$;

(b) Calcule a probabilidade dos eventos que aparecem acima e dos seguintes
   (i) $\{N_5 = 3, N_4 = 4\}$;
   (ii) $\{N_5 = 3, T_4 = 4,5\}$;
   (iii) $\{T_2 = 3, T_4 = 2,5\}$;
   (iv) $\{T_2 = 3, T_4 = 4,5, T_5 = 5\}$;

(c) Escreva em termos de $\{N_t\}_{t\geq 0}$ os seguintes eventos (p.ex, $\{T_1 = 2\} = \{N_2 = 1, N_t = 0, 0 \leq t < 2\}$).
   (i) $\{T_2 = 3\}$;
   (ii) $\{T_2 = 3, T_4 = 4,5\}$;
   (iii) $\{T_2 = 3, T_4 = 4,5, T_5 = 5\}$;

(d) Calcule $P(N_5 = 3|N_2 = 2)$ e $P(N_2 = 2|N_5 = 3)$.

(e) Calcule $cov(N_{1,5}, N_3)$.

(f) Calcule $\mathbb{E}(N_t)$, $Var(N_t)$, $\mathbb{E}(N_t \cdot N_{t+s})$ e $\mathbb{E}(N_{t+s}|N_t)$ para $t, s > 0$.

(g) Calcule a densidade conjunta de T_1, T_2.

34. A cada cinco clientes que chegam numa loja, um ganha un presente. (I.e. os clientes nro. 5, 10, 15, etc ganham presentes). Se as chegadas dos clientes formam um processo de Poisson com taxa λ,

(a) Calcule a função de densidade dos tempos entre chegadas consecutivas de clientes que ganham presentes;

(b) Calcule $P(M_t = K)$ para o número de presentes M_t que foram dados pela loja até o instante t.

35. As chegadas de clientes numa loja formam um processo de Poisson com taxa $\lambda = 20$ por hora. Calcule a quantidade esperada de vendas feitas durante o expediente de oito horas durante un dia de trabalho se a probabilidade de um cliente comprar alguma coisa é 0.3.

36. Um shopping têm três andares. As chegadas a cada um deles formam um processo de Poisson com taxas $\lambda_1 = 110$, $\lambda_2 = 90$, $\lambda_3 = 160$ clientes por hora. 30% dos clientes são homens. A probabilidade de um cliente homem comprar alguma coisa é $0,8$ e a probabilidade de uma cliente comprar é $0,1$. As mercadorias custam em média 4,50 reais.

(a) Qual será a média do total de vendas num dia com expediente de 10 horas?
(b) Qual é a probabilidade de que a terceira cliente mulher que comprou alguma coisa chegue durante os primeiros 15 minutos? Qual é o valor esperado do momento da sua chegada?

37. Um dispositivo está submetido a batidas que ocorrem segundo um processo de Poisson N com taxa λ. O dispositivo pode falhar só devido a estas batidas e a probabilidade que uma batida provoque uma falha é independente da quantidade e dos instantes das falhas anteriores. Seja K o número total de btidas que o dispositivo leva antes de uma falha e seja $T = T_K$ o instante da falha.
(a) Calcule $\mathbb{E}(T)$ e $Var(T)$.
(b) Calcule $\mathbb{E}(T|K)$.
(c) Calcule $\mathbb{E}(T|K > 9)$.

38. Um processo de Poisson bidimensional de taxa λ é um processo estocástico que conta eventos que ocorrem aleatoriamente no plano (i.e. $\mathbb{T} = \mathbb{R}^2$) de forma que
(i) para cada região com área A, o número de eventos naquela região tem distribuição de Poisson com média λA;
(ii) o número de eventos numa região é independente do número de eventos em qualquer outra que não a intercepte (ou seja, disjunta com ela).

Para um tal processo considere um ponto arbitrário do plano e denote por X a sua distância até o evento mais próximo (distância euclidiana). Prove que
(a) $P(X > t) = e^{-\lambda\pi t^2}$;
(b) $\mathbb{E}(X) = \sqrt{\lambda}/2$

CAPíTULO 5

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 $t, s \geq 0$

$$P(X_{t+s} = j|X_u, u \leq s) = P(X_{t+s} = j|X_s). \tag{1.28}$$

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

$$P(X_{t+s} = j|X_s = i) = P_{i,j}(t), \tag{1.29}$$

a cadeia será chamada de **homogênea no tempo**.

A equação (1.28) 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*

$$P_{i,j}(t+s) = \sum_{k \in E} P_{i,k}(t) P_{k,j}(s). \tag{1.30}$$

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. □

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, teriamos 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 atravês 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 $0 \leq t_0 < t_1 < \cdots < t_n$ instantes de tempo e sejam $i_0, i_1, \ldots, i_n \in E$, então*

$$P(X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_n} = i_n | X_{t_0} = i_0) = P_{i_0,i_1}(t_1 - t_0) P_{i_1,i_2}(t_2 - t_1) \cdots P_{i_{n-1},i_n}(t_n - t_{n-1}). \quad (1.31)$$

Se além disto vale que π é a distribuição inicial da cadeia, ou seja, π é a distribuição de probabilidade de X_0, então

$$P(X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_n} = i_n) = \sum_{i_0 \in E} \pi(i_0) P_{i_0,i_1}(t_1) P_{i_1,i_2}(t_2 - t_1) \cdots P_{i_{n-1},i_n}(t_n - t_{n-1}). \quad (1.32)$$

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 1.31 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 $t \geq 0$,*

$$P(W_t > u | X_t = i) = e^{-q_i u},\ u \geq 0,$$

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

$$0 = T_0 < T_1 < \cdots < T_n < \cdots$$

e vale

$$X_t = \widehat{X}_n, \quad T_n \leq t < T_{n+1}.$$

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 $T_1 | \widehat{X}_0 = i \sim \exp(q_i)$. De forma análoga teremos que $T_{n+1} = T_n + W_{T_n}$ e $T_{n+1} - T_n | \widehat{X}_n = i \sim \exp(q_i)$.

Para poder fazer a construção acima para todo $t \in \mathbb{R}$ precisaremos que $\lim_{n\to\infty} T_n = +\infty$. 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 $m - 1$-éssima 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 dai 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*

$$\frac{d}{dt}\Big|_{t=0} P(t) = A.$$

Exemplo 3.3. *Para o exemplo 1.7 calcule o gerador infinitesimal, as taxas q_i e as probabilidades Q_{ij}.*

Solução:

$$A = \frac{d}{dt}\Big|_{t=0} \begin{bmatrix} 0,6 + 0,4e^{-5t} & 0,4 - 0,4e^{-5t} \\ 0,6 - 0,6e^{-5t} & 0,4 + 0,6e^{-5t} \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix}.$$

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* $n \geq 1$ $\left(e^{-\lambda t}(\lambda t)^n\right)' = \lambda^n e^{-\lambda t} t^{n-1}(n - \lambda t)$, *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* $\widehat{Y}$. *Se* i *não for absorvente para* $\widehat{Y}$ *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, pose 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 $P(t+h) = P(t)P(h)$, obteriamos 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 teriamos $P(t) = e^{At}$. Usando a série de Taylor da exponencial poderiamos 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 coficientes 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 assintôtico.

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 diferênça dos transitórios. Portanto, um estado será recorrente para a cadeia X se e somemente se ele o for para o seu esqueleto $\{\widehat{X}_n\}_{n\geq 0}$.

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 irredutível no caso discreto se todos os seus elementos estavam comunicados entre si. De forma análoga diremos agora que $\mathcal{C} \subseteq E$ é **irredutível** quando para todos $i, j \in E$ vale $P_i(\tau_j < \infty) > 0$, ou seja a cadeia vai de i para j com probabilidade positiva. Os conjuntos irredutíveis 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 $s \geq 0$,

$$\sum_{k\in E} \pi(k) P_{k,j}(s) = \pi(j). \tag{5.33}$$

No caso finito podemos escrever (5.33) da forma compacta

$$\pi^t \cdot P(s) = \pi^t. \tag{5.34}$$

Novamente, se iniciarmos a cadeia com a distribuição estacionária, teremos que todos os estados possuem a mesma distribuição pois pela igualdade (1.32) teriamos,

$$P(X_t = j) = \sum \pi(i_0) P_{i_0,j}(t) = \pi(j).$$

A diferênça do caso discreto, agora teremos uma quantidade náo 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 (5.33). 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) obteriamos

$$\sum_{k\in E} \pi(k) q_{k,j} = 0, \tag{5.35}$$

e no caso finito

$$\pi^t \cdot A = 0. \tag{5.36}$$

Podemos nos plantear também o problema no sentido inverso: será que uma distribuição que satisfaça (5.35) tem que ser uma distribuição estacionária? Para responder esta questão usaremos as equações de Kolmogorov retrospectivas.

Suponha que uma distribuição π satisfaz (5.35). 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 (5.35). 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 satisfazer (5.35).

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 satisfazer (5.35), 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 $\widehat{Y}$ e viceversa, toda distribuição estacionária de $\widehat{Y}$ será também distribuição estacionária para X.

Como no caso discreto, o comportamento assintôtico 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,*

$$\lim_{t \to \infty} P_i(X_t = j) = 0.$$

Teorema 5.6. *Seja X uma cadeia irredutível. Então existe*

$$\lim_{t \to \infty} P_i(X_t = j) = \pi(j)$$

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* $\{\widehat{X}_n\}_{n\geq 0}$ *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

$$A = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & 0 & \dots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & & \end{bmatrix}.$$

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 (10.22) 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

$$\pi(j) = \begin{cases} \frac{1}{c}, & j = 0 \\ \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{c \mu_1 \mu_2 \cdots \mu_k}, & j \geq 1 \end{cases},$$

onde

$$c = 1 + \sum_{k=1}^{+\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k}.$$

7. Exercícios

39. Considere uma cadeia de Markov a tempo contínuo com espaço de estados $E = \{0, 1, 2\}$ e gerador infinitesimal

$$A = \begin{bmatrix} -10 & 5 & 5 \\ 1 & -2 & 1 \\ 2 & 2 & -4 \end{bmatrix}$$

(a) Ache a equação prospectiva e a retrospectiva para $P'_{0,0}(t)$.
(b) Ache a matriz de transição Q do esqueleto da cadeia.
(c) Ache a distribuição estacionária da cadeia.

40. Considere um processo de nascimento e morte com espaço de estados $E = \{0, 1, 2, \dots\}$ e taxas de nascimento e morte: $\lambda_{i+1} = (i + 1)\lambda$ e $\mu_i = i\mu$, $i \geq 0$.
(a) Ache a equação retrospectiva para $P'_{i,j}(t)$.
(b) Ache a matriz de transição Q do esqueleto da cadeia.
(c) O processo é transitório ou recorrente? Porquê?

41. Considere uma cadeia de Markov a tempo contínuo com espaço de estados $E = \{0, 1, 2\}$ e matriz de transição do esqueleto

$$Q = \begin{bmatrix} 0 & 1 & 0 \\ 0,4 & 0 & 0,6 \\ 0 & 1 & 0 \end{bmatrix}.$$

Suponha que os tempos que a cadeia permanece em cada um dos estados são variáveis aleatórias exponenciais independentes entre si e com parâmetros 2,5 e 3, respectivamente.
(a) Calcule A.
(b) Ache a equação retrospectiva para $P'_{i,j}(t)$.
(c) Calcule a distribuição estacionária da cadeia. Ela é distribuição limite? Porquê?

42. Suponha que as chegadas a certo lugar seguem um processo de Poisson com taxa $\lambda > 0$ e suponha que cada chegada é do tipo a ou do tipo b com probabilidade p e q, respectivamente. Seja Y_t o tipo da última chegada antes que t.
(a) Prove que Y é um processo de Markov com espaço de estados $E = \{a, b\}$ e função de transição

$$P(t) = \begin{bmatrix} p + qe^{-\lambda t} & q - qe^{-\lambda t} \\ p - pe^{-\lambda t} & q + pe^{-\lambda t} \end{bmatrix}$$

(b) Ache a sua distribuição limite.

43. Considere uma fila do tipo M/M/1/∞ com a seguinte modificação: quando há dois clientes no sistema se um outro chegar ele vá embora e não volta nunca mais.

(a) Prove que nesse caso, X é uma cadeia de Markov com espaço de estados $E = \{0, 1, 2\}$ e gerador
$$A = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 \\ \mu & -(\mu+\lambda) & \mu \\ 0 & \mu & -\mu \end{bmatrix}.$$
(b) Escreva as equações retrospectivas.
(c) Calcule a distribuição estacionária da cadeia. Em que casos ela seria distribuição limite?

44. Seja X uma cadeia de Markov que visita os estados a, b e c na ordem cíclica $a, b, c, a, b, c, \ldots$ com tempos de permanência nos estados tendo esperanças 1, 3 e 5, respectivamente. Calcule a distribuição limite do processo.

45. Considere um processo de nascimento e morte com três estados, $E = \{0, 1, 2\}$ e taxas de nascimento e morte tais que $\lambda_0 = \mu_2$. Use as equações prospectivas para calcular $P_{0,k}(t)$, $k = 0, 1, 2$.

46. Prove que a cadeia de Markov do exemplo 1.4 da apostila é um caso especial de cadeia de Markov uniforme. Prove que o seu gerador infinitesimal é
$$A = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix}.$$

47. Para um processo puro de nascimentos com $\lambda_n > 0, \forall n \geq 0$, calcule $P_0(X_t = n)$, $\forall n \geq 0$, $t \geq 0$.

Referências

[1] P. Brémaud. *Markov Chains*. Springer, 2001.

[2] E. Çinlar. *Introduction to Stochastic Processes*. Prentice Hall, New Jersey, 1975.

[3] R. Durrett. *Elementary Probability for Applications, Cap. 5*. Springer; 2nd ed. edition-http://www.math.cornell.edu/ durrett/ep4a/ep4a.html, 1995.

[4] P. G. Hoel, S. C. S.C. Port, and C. J. Stone. *Introduction to Stochastic Processes*. Waveland Press, 1986.

[5] J.L. Snell J. G. Kemeny. *Finite Markov Chains: With a New Appendix Generalization of a Fundamental Matrix*. Springer, New York, 1983.

[6] S. Resnick. *adventures in Stochastic Processes*. Birkh 1992.

[7] S.M. Ross. *Stochastic Processes*. Wiley, 1995.

[8] S.M. Ross. *Introduction to Probability Models*. Academic Press, 1997.

[9] H.M. Taylor S. Karlin. *A First Course in Stochastic Processes*. Academic Press, 1975.

[10] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer, 2006.