

# 21<sup>o</sup> COLÓQUIO BRASILEIRO DE MATEMÁTICA

ACOPLAMENTO  
E PROCESSOS ESTOCÁSTICOS

P.A.FERRARI

A.GALVES

IMPA 21 - 25 JULHO, 1997

PABLO FERRARI (IME/USP)  
ANTONIO GALVES (IME/USP)

COPYRIGHT  by **Pablo Ferrari e Antonio Galves**  
CAPA by **Sara Müller**

ISBN 85-244-0132-X

**Conselho Nacional de Desenvolvimento Científico e Tecnológico**

**INSTITUTO DE MATEMÁTICA PURA E APLICADA**

**Estrada Dona Castorina, 110 - Jardim Botânico**

**22460-320 - Rio de Janeiro, RJ, Brasil**

# ACOPLAMENTO EM PROCESSOS ESTOCÁSTICOS

P. A. Ferrari e A. Galves

## SUMÁRIO

1. Construção de cadeias de Markov. Exemplos: processo de mistura, Ehrenfest. ....	7
2. Medidas invariantes. Irredutibilidade. Lema de Kac. ....	19
3. Perda de memória e convergência ao equilíbrio. Coeficiente de ergodicidade de Dobrushin. ....	37
4. Teoria da renovação a tempo discreto e cadeias de Markov. Teorema chave.....	53
5. Processos de Poisson. Método da projeção.....	69
6. Processos Markovianos de salto. Construção. Explosões. ....	93
7. Reversibilidade e o teorema de Burke.....	113
8. Sistemas de partículas: Modelo do votante. Processo de exclusão simples. Percolação Orientada. ....	129



*Dedicado a Ted Harris  
que nos ensinou a recortar e colar trajetórias  
e a Charlotte e Karin  
por mais coisas do que é possível dizer.*

## INTRODUÇÃO

Estas notas de aula têm a intenção de apresentar de maneira elementar um conjunto de noções, resultados e modelos básicos necessários a quem queira trabalhar com Processos Estocásticos. Apesar de seu caráter elementar e introdutório, este texto utiliza ferramentas e adota um ponto de vista que são representativos de uma parcela importante da pesquisa atual na área. Mais modestamente, essas ferramentas e ponto de vista são aqueles que os autores destas linhas adotam em sua prática cotidiana de pesquisa.

Este curso é elementar, não só porque evita utilizar a Teoria da Medida, mas também porque adota o ponto de vista contrutivista inerente à noção de *acoplamento*.

Acoplar dois processos estocásticos significa construí-los simultaneamente, através do mesmo dispositivo aleatório. O primeiro a explorar essa idéia foi Doebelin, no final da década de 30. Para demonstrar a convergência ao equilíbrio de um Processo de Markov, Doebelin fez evoluir simultaneamente duas trajetórias independentes do processo, uma começando com uma distribuição qualquer e a outra já começando em equilíbrio, e demonstrou que as duas trajetórias se encontram num tempo finito.

Talvez a morte prematura e trágica de Doebelin e a extrema originalidade de suas idéias, fizeram com que a noção de acoplamento só retornasse à literatura no final dos anos 60, em artigos de Athreya e Ney, Harris, Spitzer e Toom, entre outros. Estes últimos três estão na origem de uma área nova em Processos Estocásticos, os chamados *Sistemas Markovianos com um Grande Número de Componentes*, em seguida desenvolvida extensivamente por Holley e Liggett. No último capítulo destas notas esses sistemas são apresentados brevemente, através do *Modelo do Votante*, do *Processo de Exclusão simples simétrico* e do *Processo de Percolação orientada*.

A arte do acoplamento consiste em encontrar a melhor maneira de construir simultaneamente dois processos ou, mais geralmente, duas medidas de probabilidade. Assim, por exemplo, para estudar a velocidade de convergência ao equilíbrio de um processo de Markov, construímos simultaneamente duas trajetórias do processo e estimamos o tempo que elas levam para se encontrar. Esse tempo depende da lei conjunta das duas trajetórias. Trata-se então de encontrar a construção conjunta de duas trajetórias que minimiza esse tempo. A construção original de Doeblin, com duas trajetórias que evoluem independentemente até se encontrar, é aquela que *a priori* utiliza o maior tempo possível para realizar esse objetivo e, portanto, fornece a pior majoração para a velocidade de convergência do processo de Markov ao equilíbrio. A partir dessa idéia inicial, muitas outras construções são possíveis, algumas das quais são apresentadas nestas notas. Uma discussão da velocidade de convergência e do chamado *coeficiente de ergodicidade de Dobruschin* é apresentada no capítulo 3 destas notas.

A idéia central atrás de um acoplamento pode ser apresentada através de um exemplo muito simples. Suponhamos que temos duas moedas, uma delas caindo em cara com probabilidade  $p$  e a outra caindo em cara com probabilidade  $q$ , com  $0 < p < q < 1$ . Queremos construir um dispositivo estocástico equivalente ao lançamento simultâneo das duas moedas, e de tal maneira que sempre que a primeira moeda (aquela associada a  $p$ ) caia em cara, a segunda (associada a  $q$ ) também caia em cara. Vamos chamar de  $X$  e  $Y$  os resultados dos lançamentos da primeira e da segunda moeda, respectivamente. Vamos convencionar que 1 indicará a saída de cara e 0, a saída de coroa. Queremos então construir as variáveis  $X$  e  $Y$ , de tal maneira que

$$\mathbb{P}(X = 1) = p = 1 - \mathbb{P}(X = 0)$$

$$\mathbb{P}(Y = 1) = q = 1 - \mathbb{P}(Y = 0)$$

$$X \leq Y.$$

As duas primeiras condições dizem que  $X$  e  $Y$  expressam realmente os resultados dos lançamentos de duas moedas que caem em cara com probabilidades  $p$  e  $q$ , respectivamente. A terceira condição é a

propriedade desejada de acoplamento. Ela diz que o resultado

$$X = 1, \quad Y = 0,$$

correspondendo à situação em que a primeira moeda cai em cara, mas não a segunda, está excluído.

Para construirmos  $X$  e  $Y$  satisfazendo simultaneamente eses tres vínculos, utilizamos uma variável auxiliar  $U$ , com distribuição uniforme no intervalo  $[0, 1]$ . Definimos

$$X = 1\{U \leq p\} \quad \text{e} \quad Y = 1\{U \leq q\}.$$

É imediato que as variáveis  $X$  e  $Y$  assim definidas satisfazem as tres condições acima. Este acoplamento é o protótipo daquele conduzindo ao *coeficiente de ergodicidade de Dobruschin*.

O estilo destas notas está inspirado no *método gráfico* de Ted Harris. Ele ensinou como construir Sistemas de Partículas através de gráficos aleatórios, recortando e colando pedaços, de forma a por em evidência da maneira a mais elementar possível as propriedades do processo. Em particular, nossa construção no capítulo 6 do Processo Markoviano de Saltos, com ajuda de um Processo Pontual de Poisson no espaço-tempo, foi antecipada pela construção feita por Harris de um Sistema Markoviano com Mudança de Spins, no final dos anos 60. Por toda essa influência e inspiração estas notas lhe são dedicadas.

Este pequeno livro foi constituído com material selecionado nas notas de aulas que nós escrevemos nos últimos tres anos para nossos cursos de introdução aos Processos Estocásticos, tanto para o bacharelado como para a pós-graduação, do Departamento de Estatística do Instituto de Matemática e Estatística da Universidade de São Paulo. Diversas versões preliminares destas notas foram lidas por Hervé Guiol, Lane Pereira Alencar e Nancy Lopes Garcia que nos apontaram inúmeras gralhas e fizeram preciosas sugestões para tornar o texto mais claro. Antes deles, estudantes das diversas turmas a quem lecionamos a matéria já tinham nos ajudado ao longo dos cursos a melhorar o texto. A todos eles os autores manifestam seus agradecimentos.

São Paulo, Junho de 1997





## 1. CONSTRUÇÃO DE CADEIAS DE MARKOV.

Um processo estocástico é uma família  $(X_t)_{t \in \mathbb{T}}$  de variáveis aleatórias, onde o índice  $t$  representa o tempo. Informalmente falando, um processo estocástico descreve uma história que se desenvolve de forma aleatória ao longo de um período de tempo representado por  $\mathbb{T}$ . Nos exemplos que estudaremos  $\mathbb{T}$  será igual ao conjunto dos números naturais  $\mathbb{N} = \{0, 1, \dots\}$ , ao conjunto dos números inteiros  $\mathbb{Z}$ , ao conjunto dos números reais positivos  $[0, +\infty)$  ou ao conjunto dos números reais  $\mathbb{R}$ . Nos dois primeiros casos o processo será dito *em tempo discreto* e nos dois últimos casos ele será dito *em tempo contínuo*. Nos casos de processos em tempo discreto, há uma preferência pela letra  $n$  para indicar o instante de tempo, deixando a letra  $t$  para os casos de processos a tempo contínuo. Suponhamos que as variáveis  $X_n$  ou  $X_t$  assumam valores no conjunto  $E$ . Esse conjunto doravante será chamado de *espaço de estados* ou ainda de *espaço de fase*. Neste curso nós só consideraremos espaços de estados finitos ou enumeráveis.

Nesta seção vamos definir uma classe muito importante de processos a tempo discreto, as chamadas *cadeias de Markov*. Começemos por uma primeira definição que embora seja ligeiramente mais restritiva do que o necessário, tem o grande mérito de ser operacional.

**Definição 1.1.** Diremos que o processo  $(X_n)_{n \in \mathbb{N}}$  tendo  $E$  como espaço de estados é uma cadeia de Markov se existir uma função  $F : E \times [0, 1] \rightarrow E$  tal que para todo  $n \geq 1$

$$X_n = F(X_{n-1}, U_n) ,$$

onde  $U_1, U_2, \dots$  é uma sequência de variáveis aleatórias independentes entre si e distribuídas uniformemente no intervalo  $[0, 1]$ .

**Exemplo 1.2.** Tomamos  $E = \{0, 1\}$  e

$$F(x, u) = 1\{u > h(x)\} ,$$

onde  $h(0)$  e  $h(1)$  são dois números quaisquer fixados no intervalo  $[0, 1]$ . Informalmente isso significa que em cada instante  $n$  o processo

atualiza o seu valor, assumindo o valor 0 ou 1, segundo  $U_n$  seja menor ou maior do que  $h(X_{n-1})$ .

**Exemplo 1.3.** Tomamos  $E = \{0, 1\}$  e

$$F(x, u) = 1 - x, \text{ se } u > g(x) \text{ e}$$

$$F(x, u) = x, \text{ em caso contrário,}$$

onde  $g(0)$  e  $g(1)$  são dois números quaisquer fixados no intervalo  $[0, 1]$ . Informalmente, isso significa que em cada instante  $n$  o processo decide mudar o valor que ele tinha no instante anterior ou mantê-lo, conforme a variável uniforme  $U_n$  seja maior ou menor do que  $g(X_{n-1})$ .

Estes exemplos embora simples nos servirão de laboratório para apresentar e testar alguns dos resultados centrais da teoria. Para isso, vamos apresentar algumas de suas propriedades básicas cujas demonstrações serão apresentadas como exercícios no final do capítulo. Definimos

$$Q(x, y) = \mathbb{P}\{X_n = y \mid X_{n-1} = x\},$$

onde  $n \geq 1$  e  $x$  e  $y$  assumem os valores 0 e 1.

Designamos por  $Q_h$  e por  $Q_g$  as probabilidades condicionais correspondentes respectivamente ao primeiro e ao segundo exemplo. É fácil ver que se  $h(0) = h(1)$ , então

$$Q_h(0, y) = Q_h(1, y).$$

Ou seja, neste caso, a cadeia de Markov correspondente à função  $h$  é, na verdade, uma sequência de variáveis aleatórias independentes e identicamente distribuídas.

Por outro lado, se  $h(0) = g(0)$  e  $1 - h(1) = g(1)$ , então

$$Q_g(x, y) = Q_h(x, y)$$

para todo  $x$  e todo  $y$ .

Essa última igualdade significa que neste caso, se nós escolhermos o mesmo ponto inicial para as cadeias definidas nos exemplos 1 e 2,

então as duas terão a mesma lei conjunta. Desse ponto de vista, as duas construções são equivalentes.

No entanto, a primeira será aquela que nós em geral utilizaremos. A razão dessa preferência está ligada à noção de acoplamento que é definida a seguir para esta cadeia.

*Acoplamento.* Sejam  $(X_n^0)_{n \in \mathbb{N}}$  e  $(X_n^1)_{n \in \mathbb{N}}$  duas cadeias construídas como no exemplo 1.3 com a ajuda da mesma sequência  $U_1, U_2, \dots$  de variáveis uniformes, e tendo 0 e 1 respectivamente como pontos iniciais, isto é  $X_0^0 = 0$  e  $X_0^1 = 1$ . Denotemos por  $K$  o primeiro instante em que as duas cadeias assumem o mesmo valor, isto é

$$K = \inf\{n \geq 1 : X_n^0 = X_n^1\}$$

Supondo que  $0 < h(0) < h(1) < 1$ , veremos nos exercícios que

$$\mathbb{P}\{K > n\} = \rho^n,$$

onde  $\rho = h(1) - h(0)$ . Isso significa que ao cabo de um tempo aleatório com distribuição geométrica as cadeias  $(X_n^0)_{n \in \mathbb{N}}$  e  $(X_n^1)_{n \in \mathbb{N}}$  são indistinguíveis. Este resultado é o protótipo do teorema de convergência exponencialmente rápida para o equilíbrio que será demonstrado mais para a frente. Esse tipo de construção simultânea de duas cadeias se chama *acoplamento* e é uma importante ferramenta no estudo dos Processos Estocásticos.

Na discussão sobre acoplamento nós calculamos a probabilidade condicional do valor do processo no instante  $n$ , dado o valor do processo no instante anterior. Obtivemos assim uma função  $Q : E \times E \rightarrow [0, 1]$ . Esse tipo de função fornece uma maneira cômoda de caracterizar cadeias de Markov.

**Definição 1.4.** Seja  $E$  um conjunto finito ou enúmerável. Uma função  $Q : E \times E \rightarrow [0, 1]$  é dita uma *probabilidade de transição* se para todo elemento  $x$  de  $E$  valer  $\sum_{y \in E} Q(x, y) = 1$ .

Nos Exemplos 1.2 e 1.3, a propriedade  $\sum_{y \in E} Q(x, y) = 1$  vinha do fato de que para todo  $x$  fixado  $Q(x, \cdot)$  era uma probabilidade condicional.

É cômodo representar probabilidades de transição como matrizes. Assim, por exemplo, se  $E = \{1, 2\}$  e se  $Q(1, 1) = p$ ,  $Q(1, 2) = 1 - p$ ,  $Q(2, 1) = 1 - q$  e  $Q(2, 2) = q$ , onde  $p$  e  $q$  são números entre 0 e 1, então é natural representar  $Q$  como a matriz

$$Q = \begin{pmatrix} p & 1 - p \\ 1 - q & q \end{pmatrix}.$$

Com essa representação a propriedade  $\sum_{y \in E} Q(x, y) = 1$ , se traduz como a soma dos elementos de uma linha é igual a 1.

Daqui para a frente nós nos referiremos indiferentemente tanto a probabilidades de transição como a matrizes de transição.

**Proposição 1.5.** *Seja  $E$  um conjunto finito ou enumerável. Dados uma matriz de transição  $Q$  definida em  $E$  e um elemento  $a$  de  $E$  quaisquer é sempre possível construir uma cadeia de Markov  $(X_n^a)_{n \in \mathbb{N}}$ , tendo  $Q$  como probabilidade de transição e  $a$  como ponto inicial.*

**Prova.** Tudo o que temos a fazer é exibir a função  $F : E \times [0, 1] \rightarrow E$  da Definição 1.1 com a propriedade

$$\mathbb{P}(F(x, U) = y) = Q(x, y)$$

Já vimos que essa função pode ser obtida de várias maneiras diferentes. Vamos aqui descrever uma construção genérica.

Começamos construindo, para cada elemento  $x$  de  $E$ , uma partição disjunta  $(I_y^x)_{y \in E}$  do intervalo  $[0, 1]$ . Isto é, vamos construir uma família de conjuntos tais que, para cada  $x$  em  $E$ , os conjuntos  $(I_y^x)_{y \in E}$  são disjuntos entre si e  $\bigcup_{y \in E} I_y^x = [0, 1]$ . Vamos escolher essa partição de tal maneira que cada um dos conjuntos seja uma união de intervalos, de tal forma que nós possamos medir seus comprimentos. Vamos impor a condição que  $|I_y^x| = Q(x, y)$ , onde  $|I|$  designa o comprimento do conjunto  $I$ .

Por exemplo, em 1.2 a partição era formada pelos intervalos  $I_0^x = [0, h(x))$  e  $I_1^x = [h(x), 1]$ .

Há várias maneiras de definir uma família desse tipo. Por exemplo, podemos começar ordenando os elementos de  $E = \{y_1, y_2, \dots\}$ . Em seguida, para cada elemento  $x$  fixado de  $E$  definimos

$$I_{y_i}^x = [l_{i-1}^x, l_i^x]$$

onde

$$l_i^x = \begin{cases} 0 & \text{se } i = 0; \\ l_{i-1}^x + Q(x, y_i), & \text{se } i \geq 1. \end{cases}$$

Em seguida, definimos a função  $F$  da seguinte maneira:

$$F(x, u) = \sum_{y \in E} y 1\{u \in I_y^x\}.$$

Em outras palavras,  $F(x, u) = y$ , se e somente se  $u \in I_y^x$ .

Com essa função, basta construir a cadeia  $(X_n^a)_{n \in \mathbb{N}}$ , definindo  $X_0^a = a$ . Para ver que essa cadeia tem probabilidades de transição  $Q$ , calculemos

$$\mathbb{P}(X_n = y \mid X_{n-1} = x) = \mathbb{P}(F(x, U_n) = y)$$

Esta última expressão, pela definição construtiva da cadeia, é igual a

$$\mathbb{P}(U_n \in I_y^x) = |I_y^x| = Q(x, y). \quad \square$$

**Observação.** Quando demonstrarmos o teorema de convergência para o equilíbrio faremos uma construção mais refinada da partição  $(I_y^x)_{y \in E}$ .

A Proposição 1.5 fornece uma porta de passagem a uma definição mais abstrata de cadeia de Markov.

**Definição 1.6.** Um processo estocástico  $(Y_n)_{n \in \mathbb{N}}$ , tendo como espaço de estados o conjunto enumerável  $E$ , é dito uma cadeia de Markov com matriz de transição  $Q$ , se para todo  $n \geq 1$  e toda sequência  $x_0, x_1, \dots, x_n$  de elementos de  $E$ , tais que  $\mathbb{P}\{Y_0 = x_0, \dots, Y_{n-1} = x_{n-1}, Y_n = x_n\} > 0$ , valer a seguinte igualdade

$$(1.7) \quad \begin{aligned} \mathbb{P}\{Y_n = x_n \mid Y_0 = x_0, \dots, Y_{n-1} = x_{n-1}\} \\ = \mathbb{P}\{Y_n = x_n \mid Y_{n-1} = x_{n-1}\} = Q(x_{n-1}, x_n). \end{aligned}$$

Essa definição diz que para uma cadeia de Markov a previsão do próximo passo, conhecendo toda a história passada do processo desde o seu início é tão boa quanto a previsão feita conhecendo-se apenas o valor do processo no presente.

A Proposição 1.5 diz que dado um processo  $(Y_n)_{n \in \mathbb{N}}$  satisfazendo (1.7) é sempre possível construir um outro processo  $(\bar{Y}_n)_{n \in \mathbb{N}}$  de mesma lei usando o algoritmo descrito na Definição Operacional 1.1.

A partir de uma variável aleatória com distribuição uniforme é possível construir qualquer outra lei em  $\mathbb{R}$ . Assim sendo, a exigência na Definição 1.1 de Cadeia de Markov de que as variáveis  $U_1, U_2, \dots$  tenham uma distribuição uniforme não é realmente uma restrição. Daqui para a frente suporemos simplesmente que as variáveis uniformes  $U_1, U_1, \dots$  sejam independentes e identicamente distribuídas. Isso não é nada mais do que uma reescritura da definição que tem o mérito de simplificar a apresentação dos exemplos.

Vamos terminar esta seção com mais alguns exemplos importantes.

**Exemplo 1.8.** *Passeio aleatório no hipercubo.* Seja  $E = \{0, 1\}^N$ , onde  $N$  é um número inteiro estritamente positivo. Quando  $N = 2$  podemos pensar em  $E$  como sendo o conjunto de vértices de um quadrado. Quando  $N = 3$  podemos pensar em  $E$  como sendo o conjunto de vértices de um cubo. Quando  $N \geq 4$  podemos pensar em  $E$  como sendo um hipercubo. Nos casos em que  $N = 2$  ou  $N = 3$  um desenho ajuda a entender as relações geométricas de vizinhança entre pontos de  $E$ . Seja  $\xi = (\xi(1), \dots, \xi(N))$  um elemento de  $E$ . Seus vizinhos são todos os elementos que tem todas as coordenadas, menos uma, iguais às coordenadas de  $\xi$ . Se  $j$  é um inteiro entre 1 e  $N$  vamos representar por  $\xi^j$  o elemento de  $E$  que tem todas as coordenadas exceto a  $j$ -ésima, iguais às coordenadas de  $\xi$ , isto é

$$\xi^j(i) = \begin{cases} \xi(i) & \text{se } i \neq j; \\ 1 - \xi(j) & \text{se } i = j. \end{cases}$$

Os vizinhos de  $\xi$  são assim os pontos  $\xi^1, \dots, \xi^N$ .

Isso nos leva a definir uma distância  $d$  em  $E$  da seguinte maneira:

$$d(\xi, \zeta) = \sum_{i=1}^N |\xi(i) - \zeta(i)|$$

onde  $\xi$  e  $\zeta$  são dois elementos quaisquer de  $E$ . A distância entre  $\xi$  e  $\zeta$  é o número de coordenadas onde  $\xi$  e  $\zeta$  diferem. Essa distância é conhecida na literatura como *distância de Hamming*. Dois pontos de  $E$  são vizinhos, quando a distância entre eles é igual a 1.

Nós queremos construir uma cadeia de Markov, tendo o hipercubo  $E$  como espaço de estados e se comportando da seguinte maneira. Em cada instante, a cadeia decide mudar ou não de posição, lançando uma moeda honesta. Nos casos em que decide mudar, ela salta para uma posição vizinha escolhida ao acaso, com distribuição uniforme. Esse processo pode ser construído da seguinte maneira.

Seja  $I_1, I_2, \dots$  uma sequência de variáveis aleatórias independentes e uniformemente distribuídas no conjunto  $\{1, 2, \dots, N\}$ . Seja também  $V_1, V_2, \dots$  uma outra sequência de variáveis aleatórias, independentes entre si e independentes também da primeira sequência. Vamos supor também que as variáveis  $V_n$  sejam identicamente distribuídas com distribuição uniforme no conjunto  $\{0, 1\}$ .

Fixado um elemento qualquer  $\zeta$  de  $E$  como ponto inicial, definimos a cadeia  $(\xi_n^\zeta)_{n \in \mathbb{N}}$  da seguinte maneira:

$$\xi_n^\zeta = \begin{cases} \zeta & \text{se } n = 0 \\ F(\xi_{n-1}^\zeta, I_n, V_n) & \text{se } n \geq 1 \end{cases}$$

onde  $F : E \times \{1, \dots, N\} \times \{0, 1\} \rightarrow E$  é a função definida da seguinte maneira. Suponhamos que  $F(\xi, i, v) = \eta$ . Então

$$\eta(j) = \begin{cases} \xi(j) & \text{se } j \neq i; \\ v & \text{se } j = i. \end{cases}$$

Ou seja, o processo escolhe um vértice com probabilidade  $1/N$  e em seguida, usando uma moeda honesta, escolhe o valor em  $\{0, 1\}$  que esse vértice vai adotar no próximo instante.

Para verificar que o processo acima construído corresponde à descrição que dele fizemos anteriormente, basta calcular a sua matriz de transição o que é feito nos exercícios.

**Exemplo 1.9.** *Modelo de Ehrenfest.* O passeio aleatório no hipercubo pode ser visto como um modelo grosseiro da evolução de um gás entre dois compartimentos fechados. Inicialmente todo o gás está confinado num único compartimento. A experiência é iniciada, abrindo-se uma válvula que comunica esse compartimento com um segundo compartimento, ficando as moléculas do gás livres de se deslocarem de um para o outro compartimento. Se nós indicarmos o primeiro compartimento com o algarismo 0 e o outro com o algarismo 1, e se suposermos que  $N$  é o número de moléculas do gás (que é algo da ordem de  $10^{23}$ ), então o passeio pode ser visto como uma descrição detalhada da evolução das posições das  $N$  moléculas, bastando para isso que se interprete o número  $\xi_n^\zeta(i)$  como o número do compartimento no qual se encontra a molécula  $i$  no instante  $n$ , sendo que  $\zeta$  descreve a posição inicial das moléculas, isto é  $\zeta(i) = 0$  para todo  $i = 1, \dots, N$ .

Muitas vezes esse modelo é apresentado, dando-se ênfase ao número total de moléculas que se encontram em cada instante em um determinado compartimento. Esse novo processo pode ser definido diretamente a partir do primeiro. Definimos

$$Z_n = \sum_{i=1}^N \xi_n^\zeta(i).$$

Note que esse processo tem como espaço de estados o conjunto  $\{0, 1, \dots, N\}$ . Pode-se dizer que o Modelo de Ehrenfest é uma descrição *macroscópica* da evolução do sistema *microscópico* correspondendo ao passeio aleatório no hipercubo.

**Exemplo 1.10.** *Urna de Polya.* Seja  $\{U_n\}$  uma sequência de variáveis aleatórias uniformes em  $[0, 1]$  independentes. Dados dois números inteiros estritamente positivos  $B_0$  e  $V_0$ , definimos para todo  $n \geq 0$

$$B_{n+1} = \mathbf{1}\{U_{n+1} < s_n\} + B_n; \quad V_{n+1} = \mathbf{1}\{U_{n+1} > s_n\} + V_n$$



onde

$$s_n = \frac{B_n}{B_n + V_n}$$

O processo assim definido é chamado de urna de Polya. Na urna de Polya inicialmente temos  $V_0$  bolas vermelhas e  $B_0$  brancas. Em cada instante, uma bola é extraída da urna e devolvida junto com outra da mesma cor. O processo  $X_n \in \mathbb{N}^2$  definido por

$$X_n = (V_n, B_n)$$

é uma cadeia de Markov. No entanto o processo  $s_n$  não é de Markov. Isso será visto nos exercícios.

### Exercícios.

**1.11.** Considere uma cadeia de Markov com espaço de estados  $\{1, 2, 3\}$  e matriz de transição

$$Q = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

(a) Construa uma função  $F$  definindo uma cadeia com essa matriz de transição. Dica: Exiba uma família de intervalos  $I_y^x$  para  $x, y \in \{1, 2, 3\}$  com  $|I_y^x| = Q(x, y)$ .

(b) Calcule  $\mathbb{P}(X_2 = 1 \mid X_1 = 2)$

(c) Calcule  $\mathbb{P}(X_3 = 2 \mid X_1 = 1)$

(d) Usando a função do item (a) e uma tabela de números aleatórios, simule 10 instantes dessa cadeia, com estado inicial  $X_0 = 1$ .

**1.12.** Considere duas cadeias  $X_n$  e  $Y_n$  em  $E = \{0, 1\}$ . A cadeia  $X_n$  está definida pela função  $F(x, u) = 1\{u > h(x)\}$  (para valores fixos  $h(0)$  e  $h(1)$ ) e a cadeia  $Y_n$  está definida pela função  $F(x, u) = 1 - x$ , se  $u > g(x)$  e  $F(x, u) = x$  em caso contrário (para valores fixos  $g(0)$  e  $g(1)$ ).

(a) Calcule

$$Q_h(x, y) = \mathbb{P}\{X_n = y \mid X_{n-1} = x\},$$

$$Q_g(x, y) = \mathbb{P}\{Y_n = y \mid Y_{n-1} = x\}$$

onde  $n \geq 1$  e  $x$  e  $y$  assumem os valores 0 e 1.

(b) Mostre que se  $h(0) = h(1)$ , então o valor de  $Q_h(x, y)$  não depende de  $x$ , só de  $y$ . Ou seja, nesse caso, a cadeia de Markov correspondente à função  $h$  é, na verdade, uma sequência de variáveis aleatórias independentes e identicamente distribuídas.

(c) Simule 10 instantes da cadeia  $X_n$  com  $h(0) = 1/3$ ,  $h(1) = 1/2$ .

(d) Simule 10 instantes da cadeia  $Y_n$  com  $g(0) = 1/3$ ,  $g(1) = 1/2$ .

**1.13. Acoplamento.** Sejam  $(X_n^0)_{n \in \mathbb{N}}$  e  $(X_n^1)_{n \in \mathbb{N}}$  duas cadeias construídas como a cadeia  $X_n$  do exercício anterior com a ajuda da mesma sequência  $U_1, U_2, \dots$  de variáveis uniformes, e tendo 0 e 1 respectivamente como pontos iniciais, isto é  $X_0^0 = 0$  e  $X_0^1 = 1$ . Denotemos por  $K$  o primeiro instante em que as duas cadeias assumem o mesmo valor, isto é

$$K = \inf\{n \geq 1 : X_n^0 = X_n^1\}$$

Supondo que  $0 < h(0) < h(1) < 1$ , mostre que

$$\mathbb{P}\{K > n\} = \rho^n,$$

onde  $\rho = h(1) - h(0)$ .

**1.14.** Calcule a matriz de transição da cadeia  $(Z_n)_{n \in \mathbb{N}}$  do Modelo de Ehrenfest.

**1.15.** Prove que a partir de uma variável aleatória com distribuição uniforme é possível construir qualquer outra lei em  $\mathbb{R}$ .

**1.16.** Mostre que o processo  $(X_n) = (B_n, V_n)$  correspondente ao modelo da urna de Polya é uma cadeia de Markov e calcule sua matriz de transição. Mostre também que o processo  $(B_n/(B_n + V_n))$  não é de Markov.

**Comentários.** Cadeias de Markov foram introduzidas por A. A. Markov no começo do século XX como modelo linguístico. A construção apresentada na Proposição 1.5 se inspira na construção gráfica de Harris para sistemas de partículas. A noção de acoplamento que é esboçada aqui foi introduzida por Doebelin nos anos 30. O modelo de Ehrenfest foi introduzido por P. e T. Ehrenfest em um célebre artigo em defesa das idéias de Boltzmann em 1905. Em particular, este modelo ilustra como é possível ter ao mesmo tempo irreversibilidade (a nível macroscópico) e reversibilidade (ao nível microscópico). O exemplo da urna de Polya ilustra o fato que, em geral, a imagem de uma cadeia de Markov por uma função não é uma cadeia de Markov.



## 2. IRREDUTIBILIDADE E MEDIDAS INVARIANTES.

Nesta seção apresentaremos as noções de medida invariante e reversibilidade cadeias de Markov. Em particular veremos o Lema de Kâc que relaciona a medida invariante com os valores médios dos tempos de retorno aos estados.

Começemos com um pouco de álgebra linear. Vimos na seção anterior que podíamos pensar na probabilidade de transição como uma matriz. Nossa próxima proposição mostra que isso é mais do que um simples recurso gráfico de representação.

**Proposição 2.1.** *Seja  $(X_n^a)_{n \in \mathbb{N}}$  uma cadeia de Markov tendo como espaço de estados o conjunto enúmerável  $E$ ,  $Q$  como probabilidade de transição e  $a$  como ponto inicial. Para qualquer instante  $n \geq 1$  e para qualquer  $b \in E$  vale a igualdade*

$$\mathbb{P}\{X_n^a = b\} = Q^n(a, b).$$

Na expressão acima  $Q^n(a, b)$  designa o elemento da linha  $a$  e da coluna  $b$  da matriz produto  $Q^n$ .

**Prova.** Por indução. A igualdade é verdadeira para  $n = 1$  pela própria definição da matriz  $Q$ .

Vamos agora supor que ela seja verdadeira para um certo  $n \geq 1$ , e demonstrar que isso implica que a igualdade é verdadeira também para  $n + 1$ . Com efeito

$$(2.2) \quad \mathbb{P}\{X_{n+1}^a = b\} = \sum_{z \in E} \mathbb{P}\{X_n^a = z, X_{n+1}^a = b\},$$

pois os conjuntos  $(\{X_n^a = z\})_{z \in E}$  formam uma partição disjunta de  $\Omega$ . Podemos reescrever a probabilidade conjunta, condicionando na posição da cadeia no instante  $n$ , obtendo:

$$\sum_{z \in E} \mathbb{P}\{X_n^a = z, X_{n+1}^a = b\} = \sum_{z \in E} \mathbb{P}\{X_n^a = z\} \mathbb{P}\{X_{n+1}^a = b \mid X_n^a = z\}.$$

Pela hipótese de indução

$$(2.3) \quad \mathbb{P}\{X_n^a = z\} = Q^n(a, z).$$

Por outro lado, por definição

$$\begin{aligned}\mathbb{P}\{X_{n+1}^a = b \mid X_n^a = z\} &= \mathbb{P}\{F(X_n^a, U_{n+1}) = b \mid X_n^a = z\} \\ &= \mathbb{P}\{F(z, U_{n+1}) = b\}.\end{aligned}$$

Como as variáveis aleatórias  $U_1, U_2, \dots$  tem todas a mesma distribuição uniforme, vale a igualdade

$$\mathbb{P}\{F(z, U_{n+1}) = b\} = \mathbb{P}\{F(z, U_1) = b\}.$$

Por definição

$$(2.4) \quad \mathbb{P}\{F(z, U_1) = b\} = \mathbb{P}\{X_1^z = b\} = Q(z, b).$$

Usando (2.3) e (2.4) podemos reescrever (2.2) como

$$\mathbb{P}\{X_{n+1}^a = b\} = \sum_{z \in E} Q^n(a, z)Q(z, b)$$

e esta é precisamente a definição de  $Q^{n+1}(a, b)$ :

$$Q^{n+1}(a, b) = \sum_{z \in E} Q^n(a, z)Q(z, b),$$

o que termina a demonstração.  $\square$

**Exemplo 2.5.** Seja  $(X_n)_{n \in \mathbb{N}}$  a cadeia de Markov tendo como espaço de estados o conjunto  $E = \{1, 2\}$  e tendo como matriz de transição

$$Q = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}.$$

Se quisermos calcular  $\mathbb{P}\{X_2^1 = 1\}$  temos que somar as probabilidades de dois caminhos  $(X_1^1 = 1, X_2^1 = 1)$  e  $(X_1^1 = 2, X_2^1 = 1)$ . Ou seja,

$$\mathbb{P}\{X_2^1 = 1\} = Q(1, 1)Q(1, 1) + Q(1, 2)Q(2, 1) = p^2 + (1-p)(1-q),$$

que é exatamente o elemento que está na primeira linha e na primeira coluna da matriz

$$Q^2 = \begin{pmatrix} p^2 + (1-p)(1-q) & p(1-p) + (1-p)q \\ (1-q)p + q(1-q) & q^2 + (1-q)(1-p) \end{pmatrix}.$$

Infelizmente, para valores maiores de  $n$  o cálculo direto se torna complicado, como veremos a seguir. Para simplificar a escritura, vamos usar a notação  $P(n) = \mathbb{P}\{X_n^1 = 1\}$ . Então

$$(2.6) \quad \begin{aligned} P(n) &= P(n-1)Q(1,1) + (1-P(n-1))Q(2,1) \\ &= P(n-1)p + (1-P(n-1))(1-q). \end{aligned}$$

Esse sistema de equações a diferenças finitas tem solução simples. Com efeito, (2.6) se reescreve como

$$(2.7) \quad P(n) = P(n-1)(p+q-1) + (1-q).$$

Substituindo  $P(n-1)$  pela expressão equivalente a (2.7) para  $n-1$  obtemos

$$(2.8) \quad \begin{aligned} P(n) &= [P(n-2)(p+q-1) + (1-q)](p+q-1) + (1-q) \\ &= P(n-2)(p+q-1)^2 + (1-q)(p+q-1) + (1-q). \end{aligned}$$

Iterando esse procedimento  $n$  vezes, obtemos finalmente

$$(2.9) \quad P(n) = P(0)(p+q-1)^n + (1-q) \sum_{k=0}^{n-1} (p+q-1)^k.$$

Lembramos que, por definição,  $P(0) = 1$ . Lembramos também a fórmula da soma dos termos de uma progressão geométrica

$$\sum_{k=0}^{n-1} \theta^k = \frac{1-\theta^n}{1-\theta}.$$

Com isso, podemos reescrever (2.8) como

$$(2.10) \quad \begin{aligned} P(n) &= (p+q-1)^n + \frac{(1-q)[1-(p+q-1)^n]}{(1-p)+(1-q)} \\ &= \frac{1-q}{(1-p)+(1-q)} + (p+q-1)^n \frac{1-p}{(1-p)+(1-q)}. \end{aligned}$$

Isso encerra a conta. Observe que  $P(n)$  converge exponencialmente rápido ao valor

$$\frac{1-q}{(1-p)+(1-q)}.$$

Cálculos análogos mostram que  $Q^n$  converge (exponencialmente rápido em cada entrada) à matriz

$$\left( \begin{array}{cc} \frac{1-p}{(1-p)+(1-q)} & \frac{1-q}{(1-p)+(1-q)} \\ \frac{1-p}{(1-p)+(1-q)} & \frac{1-q}{(1-p)+(1-q)} \end{array} \right).$$

A situação na verdade é pior do que parece. O cálculo acima embora complicado é dos poucos exemplos de probabilidade de transição em  $n$  etapas que conseguimos realizar explicitamente. O que nos salva é que nos bons casos a complicada matriz  $Q^n$  converge, quando  $n \rightarrow \infty$ , para um objeto que pode ser calculado facilmente. A demonstração do resultado de convergência será feita na próxima seção. A próxima definição introduz o objeto constitutivo desse limite.

**Definição 2.11.** Seja  $E$  um conjunto finito ou enumerável e  $Q$  uma matriz de transição definida em  $E$ . Uma medida de probabilidade  $\mu$  definida em  $E$  é dita **invariante** com respeito a  $Q$  se para todo elemento  $x$  de  $E$  valer a igualdade:

$$(2.12) \quad \mu(x) = \sum_{a \in E} \mu(a)Q(a, x).$$

Esta é uma definição puramente algébrica. Ela diz simplesmente que uma probabilidade invariante é um autovetor de  $Q$  tendo 1 como



autovalor. Ela tem, no entanto, um profundo sentido estatístico que começa a ser revelado na próxima proposição que diz informalmente o seguinte. Suponhamos que o estado inicial de uma cadeia de Markov seja escolhido aleatoriamente, usando para isso uma probabilidade invariante em relação à matriz de transição da cadeia. Nessas condições, a lei da cadeia em todos os instantes será igual à sua lei no instante inicial, isto é, a probabilidade de encontrar a cadeia no instante  $n$  numa posição  $a$  será precisamente  $\mu(a)$ , quaisquer que sejam  $n$  e  $a$ , onde  $\mu$  designa a probabilidade invariante com respeito a  $Q$  que foi usada para escolher a posição inicial.

Antes de formular a próxima proposição, vamos introduzir uma notação para descrever situações nas quais o estado inicial da cadeia é escolhido segundo uma certa distribuição de probabilidade. Sejam  $(X_n^a)_{n \in \mathbb{N}}$  uma cadeia de Markov tendo como espaço de estados o conjunto enumerável  $E$ . Se seu ponto  $a$  for escolhido com lei  $\mu$  usaremos a notação

$$\mathbb{P}\{X_n^\mu = b\} = \sum_{a \in E} \mu(a) \mathbb{P}\{X_n^a = b\}.$$

**Proposição 2.13.** *Seja  $(X_n^\mu)_{n \in \mathbb{N}}$  uma cadeia de Markov tendo como espaço de estados o conjunto enumerável  $E$ ,  $Q$  como probabilidade de transição e cujo estado inicial é escolhido aleatoriamente com lei  $\mu$ . Se  $\mu$  for invariante com respeito a  $Q$ , então*

$$\mathbb{P}\{X_n^\mu = b\} = \mu(b),$$

*quaisquer que sejam  $n$  e  $b$ .*

**Prova.** Por indução. A igualdade é verdadeira para  $n = 1$ , por hipótese.

Vamos agora supor que ela seja verdadeira para um certo  $n \geq 1$ , e demonstrar que isso implica sua validade também para  $n + 1$ . Com efeito

(2.14)

$$\mathbb{P}\{X_{n+1}^\mu = b\} = \sum_{z \in E} \sum_{a \in E} \mu(a) \mathbb{P}\{X_n^a = z\} \mathbb{P}\{X_{n+1}^a = b \mid X_n^a = z\}.$$

A igualdade (2.4) na demonstração da proposição anterior diz precisamente que

$$(2.15) \quad \mathbb{P}\{X_{n+1}^a = b \mid X_n^a = z\} = Q(z, b).$$

A hipótese de indução supõe que para todo  $z$  valha a igualdade

$$(2.16) \quad \sum_{a \in E} \mu(a) \mathbb{P}\{X_n^a = z\} = \mu(z).$$

Usando (2.15) e (2.16) podemos reescrever (2.14) como

$$\mathbb{P}\{X_{n+1}^\mu = b\} = \sum_{z \in E} \mu(z) Q(z, b).$$

Como, por hipótese, a medida de probabilidade  $\mu$  é invariante com respeito a  $Q$  vale a igualdade

$$\sum_{z \in E} \mu(z) Q(z, b) = \mu(b),$$

o que encerra a demonstração.  $\square$

**Definição 2.17.** Seja  $E$  um conjunto finito ou enumerável e  $Q$  uma matriz de transição definida em  $E$ . Uma medida de probabilidade  $\mu$  definida em  $E$  é dita **reversível** com respeito a  $Q$  se para todo par de elementos  $x$  e  $y$  de  $E$  valer a igualdade:

$$(2.18) \quad \mu(x)Q(x, y) = \mu(y)Q(y, x).$$

**Proposição 2.19.** Se  $\mu$  é reversível com respeito a  $Q$ , então  $\mu$  é invariante com respeito a  $Q$ .

**Prova.** Deixada como exercício (fácil) ao leitor.  $\square$

Uma grande vantagem da reversibilidade é que o sistema de equações (2.18) é mais simples que o sistema de equações (2.12). Outra

vantagem é que ele fornece um esquema simples para construir cadeias de Markov tendo como medida invariante uma probabilidade  $\mu$  dada, e isso é a base do chamado método de Monte Carlo para simulação de medidas de probabilidade.

Consideremos agora uma cadeia de Markov com probabilidades de transição  $Q(x, y)$  que aceita uma medida invariante  $\mu$ . Defina

$$(2.20) \quad Q^*(x, y) = \frac{\mu(x)}{\mu(y)} Q(y, x).$$

É fácil ver que  $Q^*$  define uma cadeia de Markov. Essa cadeia é chamada de cadeia reversa (em relação a  $\mu$ ). Uma definição alternativa de reversibilidade é dizer que  $\mu$  é reversível para  $Q$  se  $Q = Q^*$ , onde  $Q^*$  é definida em (2.20). A demonstração do seguinte lema segue da definição (2.20).

**Lema 2.21.** *Se  $\mu$  é invariante para  $Q$  e  $Q^*$  é a matriz de transição da cadeia reversa em relação a  $\mu$ , então*

$$\mu(x_0)Q(x_0, x_1) \dots Q(x_{n-1}, x_n) = \mu(x_n)Q^*(x_n, x_{n-1}) \dots Q^*(x_1, x_0).$$

Se  $x_n = x_0$

$$(2.22) \quad Q(x_0, x_1) \dots Q(x_{n-1}, x_n) = Q^*(x_n, x_{n-1}) \dots Q^*(x_1, x_0).$$

O Lema 2.21 diz que sob a distribuição de equilíbrio o processo visto para trás no tempo se comporta como uma cadeia de Markov com matriz de transição  $Q^*$ . A segunda parte diz que a probabilidade de percorrer um ciclo fechado de estados na cadeia direta é a mesma que a probabilidade de percorrê-lo no sentido contrário para o processo reverso.

Agora vamos estudar duas questões. Quando é que uma cadeia admite uma probabilidade invariante? Quando é que uma cadeia admite uma única probabilidade invariante? O exemplo seguinte nos mostra uma situação na qual tipicamente há mais de uma probabilidade invariante.

**Exemplo 2.23.** Seja  $E = \{1, 2\}$  e seja  $Q$  a matriz assim definida:

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

Essa é a matriz de transição da cadeia constante. É fácil ver que qualquer medida de probabilidade é invariante com respeito a essa cadeia. Vê-se que o problema aqui reside no fato de que os estados 1 e 2 não se comunicam através dessa cadeia. Vamos considerar um outro exemplo.

**Exemplo 2.24.** Seja  $E = \{1, 2, 3, 4\}$  e  $Q$  a matriz assim definida:

$$Q = \begin{pmatrix} p & 1-p & 0 & 0 \\ 1-p & p & 0 & 0 \\ 0 & 0 & q & 1-q \\ 0 & 0 & 1-q & q \end{pmatrix},$$

onde  $p$  e  $q$  são dois números estritamente maiores do que 0 e estritamente menores do que 1. Neste caso, os estados 1 e 2 se comunicam entre si, mas não se comunicam com os estados 3 e 4 que, por sua vez, só se comunicam entre si. Temos na verdade duas cadeias separadas, uma com espaço de estados  $E_1 = \{1, 2\}$  e matriz de transição

$$Q_1 = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix}.$$

e outra com espaço de estados  $E_2 = \{3, 4\}$  e matriz de transição

$$Q_2 = \begin{pmatrix} q & 1-q \\ 1-q & q \end{pmatrix}.$$

Cada uma delas tem como probabilidade invariante a distribuição uniforme no seu espaço de estados. É fácil ver que qualquer combinação linear convexa dessas duas probabilidades invariantes, consideradas agora como medidas sobre  $E = \{1, 2, 3, 4\}$ , produz uma probabilidade invariante para a cadeia de matriz  $Q$ . Isso sugere a seguinte definição

**Definição 2.25.** Uma matriz de transição  $Q$  definida sobre um conjunto finito ou enumerável  $E$  é dita **irredutível**, se, para todo par  $(x, y)$  de elementos de  $E$ , existir um número  $n = n(x, y)$ , tal que

$$Q^n(x, y) > 0.$$

A questão da existência de uma probabilidade invariante é bem mais delicada. É fácil ver que se  $E$  for finito, então o sistema de equações dado por (2.12) é sempre satisfeito por ao menos uma medida de probabilidade (o número de incógnitas é menor ou igual ao número de equações). No entanto, se  $E$  for enumerável, os exercícios abaixo fornecem exemplos de matrizes de transição que não admitem nenhuma probabilidade invariante. A próxima proposição fornece uma caracterização probabilística para esse fato algébrico. Trata-se de um caso particular de um resultado mais geral demonstrado por Mark Kác, no início da década de 50.

**Teorema 2.26.** *Lema de Kác.* *Seja  $(X_n^a)_{n \in \mathbb{N}}$  uma cadeia de Markov tendo como espaço de estados o conjunto enumerável  $E$ , tendo  $a$  como estado inicial e  $Q$  como matriz de transição. Suponhamos que  $Q$  seja irredutível e que  $\mu$  seja uma probabilidade invariante com respeito a  $Q$ . Vamos também supor que  $\mu(a) > 0$ . Definimos o instante do primeiro retorno da cadeia ao estado  $a$  por*

$$T^{a \rightarrow a} = \inf\{n \geq 1 : X_n^a = a\}.$$

Então

$$\mathbb{E}(T^{a \rightarrow a}) = \frac{1}{\mu(a)}.$$

O Lema de Kác diz que o tempo médio de retorno a um estado é inversamente proporcional à probabilidade invariante desse estado. A prova do Teorema 2.26 baseia-se em três lemas. Para enunciá-los de forma concisa, vamos introduzir a seguinte notação

$$\mathbb{P}\{T^{\mu \rightarrow a} = n\} = \begin{cases} \sum_{b \in E \setminus \{a\}} \mu(b) \mathbb{P}\{T^{b \rightarrow a} = n\}, & \text{se } n > 0; \\ \mu(a), & \text{se } n = 0; \end{cases}$$

onde

$$T^{b \rightarrow a} = \inf\{n \geq 0 : X_n^b = a\}.$$

**Lema 2.27.** *Seja  $a \in E$  tal que  $\mu(a) > 0$ . Então vale a seguinte identidade*

$$\mathbb{P}\{T^{\mu \rightarrow a} = n\} = \mu(a)\mathbb{P}\{T^{a \rightarrow a} > n\}.$$

**Prova.** Para  $n = 0$  a identidade é imediata pois

$$\mathbb{P}(T^{a \rightarrow a} > 0) = 1.$$

Para  $n > 0$  escrevamos

$$\begin{aligned} (2.28) \quad \mathbb{P}\{T^{\mu \rightarrow a} = n\} &= \sum_{x_0 \in E \setminus \{a\}} \mu(x_0) \mathbb{P}\{X_0^{x_0} \neq a, \dots, X_{n-1}^{x_0} \neq a, X_n^{x_0} = a\} \\ &= \sum_{x_0 \neq a} \sum_{x_1 \neq a} \cdots \sum_{x_{n-1} \neq a} \mu(x_0) Q(x_0, x_1) \cdots Q(x_{n-1}, a). \end{aligned}$$

Como, por hipótese,  $\mu$  é invariante com respeito a  $Q$ , podemos usar o Lema 2.21, para reescrever a expressão (2.28) como

$$(2.29) \quad \sum_{x_0 \neq a} \sum_{x_1 \neq a} \cdots \sum_{x_{n-1} \neq a} \mu(a) Q^*(a, x_{n-1}) \cdots Q^*(x_1, x_0),$$

onde  $Q^*$  é a matriz de transição do processo reverso dada por (2.20). A expressão (2.29) é precisamente igual a

$$\mu(a) \mathbb{P}\{(T^*)^{a \rightarrow a} > n\},$$

onde  $(T^*)^{a \rightarrow a}$  é o tempo de retorno a  $a$  para a cadeia reversa. Pela identidade (2.22), a distribuição de  $(T^*)^{a \rightarrow a}$  e de  $T^{a \rightarrow a}$  coincidem. Isto encerra a demonstração do lema.  $\square$

**Lema 2.30.** *Seja  $a \in E$  tal que  $\mu(a) > 0$ . Então a probabilidade  $\mathbb{P}\{T^{a \rightarrow a} > n\}$  tende a 0, quando  $n \rightarrow \infty$ . Consequentemente  $\mathbb{P}(T^{a \rightarrow a} < \infty) = 1$*

**Prova.** É claro que

$$(2.31) \quad 1 \geq \sum_{n=0}^{\infty} \mathbb{P}\{T^{\mu \rightarrow a} = n\}.$$

Pelo Lema (2.28)

$$(2.32) \quad \sum_{n=0}^{\infty} \mathbb{P}\{T^{\mu \rightarrow a} = n\} = \sum_{n=0}^{\infty} \mu(a) \mathbb{P}\{T^{a \rightarrow a} > n\}.$$

Como, por hipótese,  $\mu(a) > 0$ , (2.31) e (2.32) implicam que

$$\sum_{n=0}^{\infty} \mu(a) \mathbb{P}\{T^{a \rightarrow a} > n\} < \infty.$$

Consequentemente

$$(2.33) \quad \lim_{n \rightarrow \infty} \mathbb{P}\{T^{a \rightarrow a} > n\} = 0,$$

o que prova a primeira parte do lema. Agora note que

$$\mathbb{P}(T^{a \rightarrow a} = \infty) \leq \mathbb{P}\{T^{a \rightarrow a} > n\} \rightarrow 0 \quad (\text{as } n \rightarrow \infty)$$

como queríamos demonstrar.  $\square$

**Lema 2.34.** *A probabilidade  $\mathbb{P}\{T^{\mu \rightarrow a} > n\}$  tende a 0, quando  $n \rightarrow \infty$ .*

**Prova.** Inicialmente vamos verificar que a medida  $\mu$  atribui massa estritamente positiva a qualquer elemento de  $E$ . Com efeito, como a matriz  $Q$  é irredutível, qualquer que seja  $b \in E$ , existe uma sequência finita de estados  $x_0, x_1, \dots, x_k$ , tais que  $x_0 = a, x_k = b$  e

$$(2.35) \quad Q(x_0, x_1)Q(x_1, x_2) \dots Q(x_{k-1}, x_k) > 0.$$

Como  $\mu$  é invariante, pelo Lemma 2.21 vale a igualdade

$$(2.36) \quad \begin{aligned} \mu(a)Q(a, x_1)Q(x_1, x_2) \dots Q(x_{k-1}, x_k) \\ = \mu(b) \cdot Q^*(b, x_{k-1})Q^*(x_{k-1}, x_{k-2}) \dots Q^*(x_1, a). \end{aligned}$$

Como, por hipótese,  $\mu(a) > 0$ , (2.35) garante que o termo à esquerda da igualdade em (2.36) é estritamente positivo. Consequentemente

$Q^*(b, x_{k-1}) > 0$ ,  $Q^*(x_{k-1}, x_{k-2}) > 0$ ,  $\dots$ ,  $Q^*(x_1, a > 0)$  e  $\mu(b) > 0$ . Por definição,  $Q^*(x, y) > 0$  se e somente se  $Q(x, y) > 0$ . Portanto

$$(2.37) \quad Q(b, x_{k-1}) > 0, Q(x_{k-1}, x_{k-2}) > 0, \dots, Q(x_1, a > 0).$$

Como  $\mu(b) > 0$ , os resultados dos Lemas 2.27 e 2.30 também se aplicam a  $b$ . Consequentemente, pelo Lema 2.30

$$(2.38) \quad \mathbb{P}\{T^{b \rightarrow b} < \infty\} = 1.$$

Isso significa que a cadeia, partindo de  $b$ , volta a  $b$  num número finito de etapas, com probabilidade 1. Consideremos as trajetórias do processo entre duas visitas consecutivas a  $b$ . Cada uma dessas excursões pode passar, ou não, pelo ponto  $a$ . Seja  $a = x_0, x_1, \dots, x_k = b$  a sequência que nós consideramos na dedução de (2.36). De cada vez que o processo volta a  $b$ , a probabilidade de passar por  $a$  é no mínimo igual a

$$\begin{aligned} \delta &= Q(b, x_{k-1})Q(x_{k-1}, x_{k-2}) \dots Q(x_1, a) \\ &\quad \times Q(a, x_1)Q(x_1, x_2) \dots Q(x_{k-1}, x_k) > 0, \end{aligned}$$

por (2.37).

Isso nos fornece uma majoração para a probabilidade da cadeia partindo de  $b$  não visitar  $a$  em  $n \geq k$  etapas:

$$(2.39) \quad \mathbb{P}\{T^{b \rightarrow a} > n\} \leq (1 - \delta)^{[\frac{n}{k}]},$$

onde  $[\frac{n}{k}]$  designa a parte inteira de  $\frac{n}{k}$ .

Como  $\delta > 0$ , a desigualdade (2.39) implica que

$$(2.40) \quad \lim_{n \rightarrow \infty} \mathbb{P}\{T^{b \rightarrow a} > n\} = 0.$$

Para concluir a demonstração, escrevemos

$$(2.41) \quad \lim_{n \rightarrow \infty} \mathbb{P}\{T^{\mu \rightarrow a} > n\} = \lim_{n \rightarrow \infty} \sum_{b \in E} \mu(b) \mathbb{P}\{T^{b \rightarrow a} > n\}.$$



Passando o limite para o interior da soma, obtemos finalmente

$$\lim_{n \rightarrow \infty} \mathbb{P}\{T^{\mu \rightarrow a} > n\} = \sum_{b \in E} \mu(b) \lim_{n \rightarrow \infty} \mathbb{P}\{T^{b \rightarrow a} > n\} = 0,$$

como queríamos demonstrar.

Restá-nos justificar a passagem do limite para dentro da somatória em (2.41). Se o espaço de estados  $E$  for finito, isso é autorizado pela continuidade da função soma. Se o espaço de estados for enumerável temos que invocar um teorema um pouco mais avançado que em Teoria da Integração é geralmente chamado de *Teorema da Convergência Monótona*.  $\square$

**Prova do Teorema 2.26.** Se faz agora em tres linhas. O Lema 2.34 diz que

$$\mathbb{P}\{T^{\mu \rightarrow a} < \infty\} = 1.$$

Consequentemente, pelo Lema 2.27,

$$1 = \sum_{n=0}^{\infty} \mathbb{P}\{T^{\mu \rightarrow a} = n\} = \mu(a) \sum_{n=0}^{\infty} \mathbb{P}\{T^{a \rightarrow a} > n\}.$$

Basta agora observar que

$$\sum_{n=0}^{\infty} \mathbb{P}\{T^{a \rightarrow a} > n\} = \mathbb{E}(T^{a \rightarrow a}),$$

para concluir a demonstração.  $\square$

## Exercícios

**2.42.** Mostre que se  $\mu$  é reversível com respeito a  $Q$ , então  $\mu$  é invariante com respeito a  $Q$ .

**2.43.** Mostre que se  $\mu$  é invariante para  $Q$ , então  $Q^*$  definida por

$$(2.44) \quad Q^*(x, y) = \frac{\mu(x)}{\mu(y)} Q(y, x).$$

é a matriz de transição de uma cadeia de Markov.

2.45. Prove que a identidade (2.22) implica que  $T^{a \rightarrow a}$  e  $(T^*)^{a \rightarrow a}$  tem a mesma distribuição.

2.46. Calcule as medidas de probabilidades invariantes com respeito às cadeias de Markov apresentadas nos exemplos 1.2, 1.3 e no exercício 1.11.

2.47. De um exemplo de uma cadeia que tem uma probabilidade invariante que não satisfaz a condição de reversibilidade.

2.48. *Passeio aleatório em  $\mathbb{Z}$ .* Seja  $E = \mathbb{Z}$  e  $U_1, U_2, \dots$  uma sequência de variáveis aleatórias i.i.d. assumindo valores no conjunto  $\{-1, +1\}$  com

$$\mathbb{P}\{U_n = +1\} = p = 1 - \mathbb{P}\{U_n = -1\},$$

onde  $p \in [0, 1]$ . Seja  $a$  um ponto qualquer fixado de  $\mathbb{Z}$ . Definimos o passeio aleatório  $(S_n^a)_{n \in \mathbb{N}}$ , tendo  $a$  como ponto inicial, da seguinte maneira:

$$S_0^a = a$$

e

$$S_n^a = S_{n-1}^a + U_{n-1}, \text{ se } n \geq 1.$$

Essa cadeia é chamada passeio aleatório e é um modelo simplificado unidimensional do movimento Browniano.

i) Mostre que

$$\mathbb{P}\{S_n^0 = x\} = \begin{cases} \binom{n}{\frac{x+n}{2}} p^k (1-p)^{n-k}, & \text{se } x+n \text{ for par e } |x| \leq n; \\ 0, & \text{senão.} \end{cases}$$

ii) Vamos agora fazer  $p = \frac{1}{2}$ . Neste caso calcule  $\lim_{n \rightarrow \infty} \mathbb{P}\{S_n^0 = 0\}$ . Sugestão: use a fórmula de Stirling.

iii) Suponha que  $p > \frac{1}{2}$ . Use a Lei dos Grandes Números e o Lema de Borel-Cantelli, para mostrar que

$$\mathbb{P}\{S_n^0 = 0, \text{ para uma infinidade de valores de } n\} = 0.$$

iv) Tente encontrar alguma probabilidade invariante com respeito ao passeio aleatório. Tente estabelecer uma conexão entre os resultados deste item e dos dois anteriores.

**2.49.** O processo de *nascimento e morte* é uma cadeia de Markov tendo como espaço de estados o conjunto  $\mathbb{N}$  dos números naturais e tendo matriz de transição da forma seguinte:

$$Q(x, x + 1) = q(x) = 1 - Q(x, x - 1), \text{ se } x \geq 1,$$

$$Q(0, 1) = q(0) = 1 - Q(0, 0),$$

onde  $(q(x))_{x \in \mathbb{N}}$  é uma sequência de números reais *estritamente* maiores do que 0 e *estritamente* menores do que 1. Que condição deve ser satisfeita pela sequência  $(q(x))_{x \in \mathbb{N}}$  para que o processo de nascimento e morte tenha uma probabilidade invariante? Traduza esse resultado para o caso particular em que  $q(x) = p$  para todo  $x \in \mathbb{N}$ , interpretando-o.

**2.50. Método de Monte Carlo.** O método de Monte Carlo para simular uma medida de probabilidade consiste em simular uma cadeia de Markov tendo essa medida como probabilidade invariante. Para obter uma amostra da medida a partir da trajetória da cadeia deixa-se o processo evoluir até que atinja uma situação de equilíbrio. Na próxima seção nós trataremos da questão do tempo necessário para que uma cadeia de Markov atinja uma situação de equilíbrio. Aqui vamos exibir uma cadeia que pode ser utilizada para a simulação. Para isso vamos usar a noção de reversibilidade.

i) Seja  $E$  um conjunto finito e  $\mu$  uma medida de probabilidade definida em  $E$ . Considere as seguintes probabilidades de transição definidas em  $E$ : para  $y \neq x$ ,

$$Q_1(x, y) = \frac{1}{N} \frac{\mu(y)}{\mu(x) + \mu(y)}$$

$$Q_1(x, x) = 1 - \sum_{y \neq x} Q_1(x, y)$$

$$Q_2(x, y) = \frac{1}{N} \left[ \mathbf{1}\{\mu(x) \leq \mu(y)\} + \frac{\mu(y)}{\mu(x)} \mathbf{1}\{\mu(x) > \mu(y)\} \right]$$

$$Q_2(x, x) = 1 - \sum_{y \neq x} Q_2(x, y).$$

Verifique que  $\mu$  é reversível com respeito a  $Q_1$  e  $Q_2$ .

ii) Sejam  $E = \{0, 1\}^N$  e  $\mu$  a distribuição uniforme em  $E$  (isto é,  $\mu(\zeta) = 2^{-N}$ , qualquer que seja  $\zeta \in E$ ). Defina uma matriz de transição  $Q$  sobre  $E$  que satisfaça as seguintes condições:

a)  $Q(\xi, \zeta) = 0$ , se  $d(\xi, \zeta) > 2$ , onde  $d$  é a distância de Hamming, introduzida na seção anterior;

b)  $\mu$  é reversível com respeito a  $Q$ .

iii) Faça uma simulação concreta da cadeia de Markov correspondendo à matriz  $Q$  que for escolhida. Para isso use o algoritmo introduzido na definição 1.1. Tente determinar empiricamente o tempo necessário para que a trajetória da simulação atinja uma situação de equilíbrio. Estime empiricamente a densidade de valores iguais a 1 numa configuração escolhida com a medida  $\mu$ . Compare o resultado empírico com o resultado teórico dado pela Lei dos Grandes Números.

iv) Faça uma simulação da distribuição binomial de parâmetros  $\frac{1}{2}$  e  $N$ . Sugestão: utilize o modelo de Ehrenfest como base de uma simulação de Monte Carlo.

**2.51.** i) Verifique diretamente que o limite da probabilidade  $P(n)$ , dada pela fórmula (2.10) existe e calcule-o explicitamente. Calcule também  $\lim_{n \rightarrow \infty} \mathbb{P}\{X_n^2 = 1\}$ . Identifique esses limites.

ii) Para essa mesma cadeia calcule  $\mathbb{E}(T^{1 \rightarrow 1})$ , onde

$$T^{1 \rightarrow 1} = \inf\{n \geq 1 : X_n^1 = 1\}.$$

Relacione os resultados destes dois itens.

**2.52.** Sejam  $E = \mathbb{N}$  e  $Q$  a matriz de transição definida em  $E$  da seguinte maneira. Para todo  $x \in \mathbb{N}$

$$Q(0, x) = p(x) \text{ e}$$

$$Q(x, x-1) = 1 \text{ se } x \geq 1,$$

onde  $p$  é uma medida de probabilidade sobre  $\mathbb{N}$ . Seja  $(X_n^0)_{n \in \mathbb{N}}$  uma cadeia de Markov, partindo do ponto 0 e tendo  $Q$  como probabilidade de transição.

- i) Que condição deve ser satisfeita pela probabilidade  $p$ , para que a cadeia admita uma probabilidade invariante?
- ii) Calcule  $\mathbb{E}(T^{1 \rightarrow 1})$ .

Relacione os resultados destes dois ítems.

**2.53. Processo de Mistura.** O processo de mistura é uma cadeia de Markov tendo como espaço de estados o hipercubo  $E = \{0, 1\}^N$  e definido através do algoritmo seguinte. Seja  $\mathcal{P}_N$  o conjunto de todas as permutações possíveis dos elementos da sequência  $(1, 2, \dots, N)$ , isto é, o conjunto das bijeções do conjunto  $\{1, 2, \dots, N\}$  em si mesmo. Seja  $\pi$  um elemento de  $\mathcal{P}_N$ . Com a ajuda de  $\pi$  nós podemos definir uma transformação  $F_\pi: E \rightarrow E$  da seguinte maneira. Para todo elemento  $\xi \in E$

$$F_\pi(\xi)(i) = \xi(\pi(i)).$$

Em outras palavras,  $F_\pi$  permuta os valores de cada configuração  $\xi$  colocando na coordenada  $i$  o valor da coordenada  $\pi(i)$ .

Seja  $(\Pi_1, \Pi_2, \dots)$  uma sequência de variáveis aleatórias independentes entre si, identicamente distribuídas e assumindo valores no conjunto  $\mathcal{P}_N$ . O processo de mistura  $(\eta_n^\zeta)_{n \in \mathbb{N}}$  tendo  $\zeta$  como estado inicial é definido da seguinte maneira:

$$\eta_n^\zeta = \begin{cases} \zeta, & \text{se } n = 0; \\ F_{\Pi_n}(\eta_{n-1}^\zeta), & \text{se } n \geq 1. \end{cases}$$

- i) Mostre que o Processo de Mistura não é irredutível.
- ii) Suponha que as variáveis  $\Pi_n$  tenham todas distribuição uniforme. Calcule todas as probabilidades invariantes do processo de Mistura neste caso.
- iii) Seja  $\mathcal{V}_N$  o conjunto das permutações que só alteram as posições respectivas de dois pontos vizinhos da sequência  $(1, 2, \dots, N)$ . Um

elemento típico de  $\mathcal{V}_N$  é a permutação  $\pi^k$ , onde  $k \in \{1, 2, \dots, N\}$ , assim definida:

$$\pi^k(i) = \begin{cases} i, & \text{se } i \neq k, i \neq k+1, \\ k+1, & \text{se } i = k, \\ k, & \text{se } i = k+1. \end{cases}$$

Na representação acima, a adição é feita módulo  $N$ , isto é,  $N+1 = 1$ . Suponha que as variáveis  $\Pi_n$  tenham distribuição uniforme no conjunto  $\mathcal{V}_N$ . Calcule todas as probabilidades invariantes do processo de Mistura neste caso.

Compare os resultados dos itens ii) e iii).

### 3. PERDA DE MEMÓRIA E CONVERGÊNCIA PARA O EQUILÍBRIO.

Nesta seção aparece novamente a noção de *acoplamento* que é a principal ferramenta utilizada nestas notas. O objetivo é o de provar um teorema que garante a convergência em distribuição de uma cadeia de Markov em direção à sua medida de probabilidade invariante. Apresentaremos duas demonstrações desse fato. Em ambos os casos a ferramenta básica é o acoplamento. *Acoplar* duas medidas de probabilidade significa construí-las simultaneamente. A idéia central da demonstração é a seguinte. Inicialmente faremos uma construção simultânea de duas trajetórias da cadeia, cada trajetória tendo um ponto de partida diferente. É essa construção simultânea que nós chamamos de acoplamento. Essa construção é feita de tal maneira que as duas trajetórias se encontrem o mais rapidamente possível e a partir desse momento permaneçam juntas para sempre. Uma primeira versão simplificada desse fato é o Teorema 3.3. Os Lemas 3.5 e 3.13 mostrarão sucessivamente como utilizar um resultado de perda de memória da condição inicial, como 3.3, para deduzir um resultado de convergência em distribuição, assim como um resultado de unicidade da medida de probabilidade invariante. O Teorema 3.19 retoma de forma mais refinada a demonstração do teorema 3.3, introduzindo o chamado *coeficiente de ergodicidade de Dobrushin*. Isso nos leva finalmente à apresentação final do Teorema de Convergência, feita em 3.21. No meio do caminho, nós introduziremos a noção de cadeia *aperiódica*, necessária para formular com a devida generalidade as hipóteses do Teorema de Convergência.

Vamos nos colocar no quadro seguinte. Temos um espaço de estados  $E$  munido de uma matriz de transição  $Q$ . Através de um algoritmo do tipo daquele que foi introduzido na Proposição 1.5 construímos duas cadeias de Markov,  $(X_n^a)_{n \in \mathbb{N}}$  e  $(X_n^b)_{n \in \mathbb{N}}$ , tendo  $Q$  como matriz de transição e partindo dos elementos  $a$  e  $b$  de  $E$ , respectivamente.

**Definição 3.1.** O instante  $\tau^{a,b}$  em que as cadeias assumem pela

primeira vez o mesmo valor é definido da seguinte maneira

$$\tau^{a,b} = \begin{cases} +\infty, & \text{se } X_n^a \neq X_n^b, \text{ para todo } n \geq 0; \\ \min\{n \geq 1 : X_n^a = X_n^b\}, & \text{em caso contrário.} \end{cases}$$

Na sua primeira versão simplificada, a velocidade de perda de memória da cadeia será controlada pelo **coeficiente de ergodicidade**  $\beta$  que definiremos a seguir.

**Definição 3.2.** O coeficiente de ergodicidade  $\beta(Q)$  de uma matriz de transição  $Q$  é definido como

$$\beta(Q) = \sum_{x \in E} \min_{a \in E} Q(a, x).$$

**Teorema 3.3.** *Se for  $E$  finito e se a matriz de transição  $Q$  tiver todos os elementos estritamente positivos, então existe um algoritmo de construção simultânea das cadeias de Markov  $(X_n^a)_{n \in \mathbb{N}}$  e  $(X_n^b)_{n \in \mathbb{N}}$ , tal que*

$$\mathbb{P}\{\tau^{a,b} > n\} \leq (1 - \beta(Q))^n.$$

**Prova.** Retomamos a demonstração da Proposição 1.5. Lá explicávamos que para construir a função  $F : E \times [0, 1] \rightarrow E$  prevista na Definição 1.1 tudo o que tínhamos a fazer era definir para cada elemento  $x$  de  $E$  uma partição disjunta  $(I_y^x)_{y \in E}$  do intervalo  $[0, 1]$ , de tal forma que  $|I_y^x| = Q(x, y)$ , onde  $|A|$  designa o comprimento do conjunto  $A$ . Nesse texto  $A$  será sempre uma união finita de intervalos.

Nesta demonstração vamos definir essas partições de tal forma que as partes comuns  $(I_y^a \cap I_y^b)_{y \in E}$  sejam suficientemente grandes.

Para facilitar a redação, vamos começar, ordenando os elementos de  $E = \{y_1, y_2, \dots, y_N\}$ , onde  $N$  é o cardinal de  $E$ .

Em seguida, para cada  $i = 1, \dots, N$ , definimos

$$J_{y_i} = [l_{i-1}, l_i)$$



onde

$$l_i = \begin{cases} 0, & \text{se } i = 0; \\ l_{i-1} + \min_{a \in E} Q(a, y_i), & \text{se } i \geq 1. \end{cases}$$

Isso define uma partição do intervalo  $[0, l_N]$ . Vamos agora definir uma família de partições do intervalo complementar  $(l_N, 1]$ , indexada superiormente pelos elementos de  $E$ . Para cada elemento  $a$  de  $E$ , definimos

$$\tilde{J}_{y_i}^a = [\tilde{l}_{i-1}^a, \tilde{l}_i^a)$$

onde

$$\tilde{l}_i^a = \begin{cases} l_N, & \text{se } i = 0; \\ \tilde{l}_{i-1}^a + Q(a, y_i) - \min_{b \in E} Q(b, y_i) & \text{se } i \geq 1. \end{cases}$$

Finalmente definimos

$$I_{y_i}^a = J_{y_i} \cup \tilde{J}_{y_i}^a.$$

É fácil ver que

$$|I_{y_i}^a| = Q(a, y_i).$$

Em seguida, definimos a função  $F : E \times [0, 1] \rightarrow E$  da seguinte maneira:

$$F(a, u) = \sum_{i=1}^N y_i \mathbf{1}\{u \in I_{y_i}^a\}.$$

Em outras palavras,  $F(a, u) = y_i$ , se e somente se  $u \in I_{y_i}^a$ .

Note que

$$\bigcap_{a \in E} I_{y_i}^a = J_{y_i}.$$

Em consequência, se  $u < l_N$ , então

$$(3.4) \quad F(x, u) = F(z, u),$$

quaisquer que sejam  $x$  e  $z$ . Esta é a propriedade da função  $F$  que garantirá o resultado.

Vamos construir as cadeias  $(X_n^a)_{n \in E}$  e  $(X_n^b)_{n \in E}$ , usando a função  $F$ .

Seja

$$K = \inf\{n \geq 1 : U_n < l_N\},$$

onde, seguindo a notação introduzida na definição 1.1,  $(U_1, U_2, \dots)$  é a sequência de variáveis aleatórias independentes e uniformes utilizadas na construção simultânea das cadeias de Markov  $(X_n^a)_{n \in E}$  e  $(X_n^b)_{n \in E}$ .

A propriedade (3.4) implica que

$$\tau^{a,b} \leq K,$$

para todo  $a$  e  $b$ .

Note-se que a variável aleatória  $K$  tem distribuição geométrica

$$\mathbb{P}\{K > n\} = \mathbb{P}\{U_1 > l_N, \dots, U_n > l_N\} =$$

$$\prod_{i=1}^n \mathbb{P}\{U_i > l_N\} = (1 - l_N)^n.$$

Para concluir a demonstração basta agora observar que

$$l_N = \beta(Q). \quad \square$$

Vamos agora mostrar como teoremas de convergência podem ser deduzidos de um resultado de perda de memória do tipo do Teorema 3.3. A idéia da demonstração é a seguinte. Escolhemos aleatoriamente o ponto de partida de uma das trajetórias, usando para isso a própria probabilidade invariante da cadeia. Assim sendo, vista isoladamente, esta componente da cadeia dupla que construímos estará sempre em equilíbrio. Como consequência do acoplamento podemos achar uma cota superior para a distância entre a medida invariante e a medida do processo no instante  $t$ .

**Teorema 3.5.** *Suponhamos que exista um algoritmo de construção simultânea (acoplamento) de uma família de cadeias de Markov  $(X_n^a)$ ,  $a \in E$ , tendo como espaço de estados o conjunto finito ou enumerável  $E$  e tendo  $Q$  como matriz de transição, tal que quaisquer que sejam  $a$  e  $b$ ,  $X_n^a = X_n^b$ , para todo  $n \geq \tau^{a,b}$ . Então*

$$(3.6) \quad \sup_{(a,y)} |\mathbb{P}\{X_n^a = y\} - \mu(y)| \leq \sup_{a,b} \mathbb{P}\{\tau^{a,b} > n\}.$$

**Prova.** Por hipótese, a probabilidade  $\mu$  é invariante com respeito à cadeia. Isso quer dizer que

$$\sum_{b \in E} \mu(b) \mathbb{P}\{X_n^b = y\} = \mu(y).$$

Isso nos permite reescrever a diferença de probabilidades no lado esquerdo da expressão (3.6) como

$$(3.7) \quad |\mathbb{P}\{X_n^a = y\} - \mu(y)| = \left| \mathbb{P}\{X_n^a = y\} - \sum_{b \in E} \mu(b) \mathbb{P}\{X_n^b = y\} \right|.$$

A expressão do lado direito da igualdade (3.7) é majorada por

$$(3.8) \quad \sum_{b \in E} \mu(b) |\mathbb{P}\{X_n^a = y\} - \mathbb{P}\{X_n^b = y\}|.$$

Como as cadeias  $(X_n^a)_{n \geq 1}$  e  $(X_n^b)_{n \geq 1}$  foram construídas simultaneamente, através das mesmas variáveis auxiliares  $U_n$ , nós podemos reescrever a diferença de probabilidades em (3.8) como a esperança da diferença de duas funções indicadoras

$$(3.9) \quad |\mathbb{P}\{X_n^a = y\} - \mathbb{P}\{X_n^b = y\}| = |\mathbb{E}[\mathbf{1}\{X_n^a = y\} - \mathbf{1}\{X_n^b = y\}]| \\ \leq \mathbb{E} |\mathbf{1}\{X_n^a = y\} - \mathbf{1}\{X_n^b = y\}|.$$

Finalmente, observamos que

$$(3.10) \quad |\mathbf{1}\{X_n^a = y\} - \mathbf{1}\{X_n^b = y\}| \leq \mathbf{1}\{X_n^a \neq X_n^b\}.$$

A hipótese (i) do lema diz que

$$(3.11) \quad \mathbf{1}\{X_n^a \neq X_n^b\} = \mathbf{1}\{\tau^{a,b} > n\}.$$

Utilizando (3.7), (3.8), (3.9), (3.10) e (3.11) sucessivamente em (3.6) obtemos a majoração

$$(3.12) \quad \begin{aligned} |\mathbb{P}\{X_n^a = y\} - \mu(y)| &\leq \sum_{b \in E} \mu(b) \mathbb{P}\{\tau^{a,b} > n\} \\ &\leq \sup_{a,b} \mathbb{P}(\tau^{a,b} > n). \end{aligned}$$

O que conclui a demonstração.  $\square$

**Corolário 3.13.** *Se*

$$\sup_{a,b} \mathbb{P}\{\tau^{a,b} > n\} \leq (1 - \gamma)^n,$$

onde  $\gamma$  é uma constante positiva estritamente menor do que 1 e se  $\mu$  for uma probabilidade invariante com respeito a  $Q$ , então

$$(3.14) \quad \sup_{(a,y)} |\mathbb{P}\{X_n^a = y\} - \mu(y)| \leq (1 - \gamma)^n.$$

**Lema 3.15.** *Nas condições do Corolário 3.13, se a matriz de transição admitir uma medida de probabilidade invariante, ela será única.*

**Prova.** Suponhamos que  $\mu$  e  $\nu$  sejam duas probabilidades invariantes com respeito à matriz  $Q$ . Como no Teorema 3.5 construímos duas cadeias cujos estados iniciais são escolhidos respectivamente segundo as distribuições  $\mu$  e  $\nu$ . Em seguida demonstramos exatamente como no Teorema 3.5 e no Corolário 3.13 que a distância entre as distribuições das duas cadeias no instante  $n$  converge a 0, quando  $n \rightarrow \infty$ . Mas como, por hipótese, as medidas  $\mu$  e  $\nu$  são invariantes, essas leis coincidem respectivamente com  $\mu$  e  $\nu$ , qualquer que seja  $n \geq 0$ . Segue que  $\mu = \nu$ . Os detalhes, idênticos aos do Teorema 3.5, são deixados como exercício ao leitor.  $\square$

Supor, como no teorema 3.3, que  $Q(x, y) > 0$  é certamente muito restritivo. Evidentemente bastaria supor que existe um inteiro  $k \geq 1$  tal que  $Q^k(x, y) > 0$ , o que nos daria um resultado semelhante simplesmente com o fator  $(1 - \beta(Q^k))^{\frac{1}{k}}$  no lugar de  $(1 - \beta(Q))$ . A questão que se coloca então naturalmente é: quais são as matrizes de transição sobre um espaço de estados finito que tem todos os valores positivos a partir de uma certa potência. Os exemplos de matrizes irredutíveis apresentados na seção anterior mostram que a condição de *irredutibilidade* é necessária. No entanto ela não é suficiente como nos mostra o próximo exemplo.

**Exemplo 3.16.** Consideremos em  $E = \{1, 2\}$  a matriz de transição  $Q$  assim definida

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

Trata-se uma matriz claramente irredutível; no entanto qualquer uma de suas potências terá sempre elementos nulos. Com efeito, qualquer que seja  $k \geq 0$ , temos as igualdades

$$Q^{2k} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

e

$$Q^{2k+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

O problema dessa matriz é que as transições de 1 a 2 ou de 2 a 1 só são possíveis num número ímpar de etapas, enquanto que as transições de 1 a 1, ou de 2 a 2, só são possíveis num número par de etapas. Esse tipo de situação nos conduz à noção de periodicidade, introduzida na próxima definição. Antes, no entanto, observemos que a matriz apresentada neste exemplo admite uma medida invariante, a saber a distribuição uniforme no conjunto  $E = \{1, 2\}$ . No entanto, a distribuição da cadeia no instante  $n$  não converge para coisa alguma, quando  $n \rightarrow \infty$ . Isso indica que a hipótese de existência de uma potência a partir da qual a cadeia tem todos os valores estritamente positivos é mais do que uma simples condição suficiente para o teorema de convergência.

**Definição 3.17.** Suponhamos que  $Q$  seja uma matriz de transição definida num espaço de estados finito ou enumerável  $E$ . Um elemento  $x$  de  $E$  é dito *periódico* de período  $d$ , se

$$\text{mdc} \{n \geq 1 : Q^n(x, x) > 0\} = d.$$

O elemento será dito **aperiódico** se  $d = 1$ .

Por exemplo, a matriz do exemplo 3.16 torna os estados de  $E = \{1, 2\}$  periódicos de período 2.

As demonstrações das duas proposições seguintes, de caráter puramente algébrico, serão omitidas nestas notas. Elas são elementares e podem ser encontradas nos livros introdutórios de cadeias de Markov.

**Proposição 3.18.** *Seja  $E$  um conjunto finito ou enumerável munido de uma matriz de transição  $Q$ . Se  $Q$  for irredutível, então todos os elementos de  $E$  tem o mesmo período.*

Essa proposição, quando a cadeia for irredutível, nos autoriza a falar de *cadeia de período  $d$*  e de *cadeia aperiódica*.

**Proposição 3.19.** *Seja  $E$  um conjunto finito munido de uma matriz de transição  $Q$ . Se  $Q$  for irredutível e aperiódica, então existe um inteiro  $k$  tal que  $Q^j$  tem todos os elementos positivos, qualquer que seja  $j \geq k$ .*

Vamos agora apresentar uma construção mais cuidadosa do par de cadeias acopladas. Isso permitirá um controle mais fino da velocidade de perda de memória da cadeia através do **coeficiente de ergodicidade** de Dobrushin que definiremos a seguir.

**Definição 3.20.** O **coeficiente de ergodicidade** de Dobrushin da matriz de transição  $Q$  é definido como

$$\alpha(Q) = \min_{a,b} \sum_{x \in E} \min\{Q(a, x), Q(b, x)\}.$$

**Teorema 3.21.** *Se for  $E$  finito e se  $Q$  for uma matriz de transição definida em  $E$ , tendo todos os elementos estritamente positivos, então existe um algoritmo de construção simultânea das cadeias de Markov  $(X_n^a)_{n \in \mathbb{N}}$  e  $(X_n^b)_{n \in \mathbb{N}}$ , tal que*

$$\mathbb{P}\{\tau^{a,b} > n\} \geq (1 - \alpha(Q))^n.$$

**Prova.** Retomamos a demonstração do teorema 3.3. Agora a família de partições será construída de tal forma que as partes comuns de cada par de partições indexadas por elementos distintos de  $E$  (isto é,  $(I_y^a \cap I_y^b)_{y \in E}$ ) sejam o maior possível. Isso nos obriga a utilizar um algoritmo algo mais detalhado.

Para facilitar a redação, vamos começar, ordenando os elementos de  $E = \{y_1, y_2, \dots, y_N\}$ , onde  $N$  é o cardinal de  $E$ .

Para cada par  $a$  e  $b$  de elementos fixados de  $E$  definimos

$$J_{y_i}^{a,b} = [l_{i-1}^{a,b}, l_i^{a,b}]$$

onde

$$l_i^{a,b} = \begin{cases} 0, & \text{se } i = 0; \\ l_{i-1}^{a,b} + \min\{Q(a, y_i), Q(b, y_i)\} & \text{se } i \geq 1. \end{cases}$$

Isso define uma partição do intervalo  $[0, l_N^{a,b}]$ . Restá-nos fazer uma partição do intervalo complementar  $(l_N^{a,b}, 1]$ . Como as partes comuns foram já totalmente utilizadas restá-nos alinhar os restos, de forma a respeitar a condição de que o comprimento total ser igual à probabilidade da transição. Vamos aqui simplesmente dar um exemplo de como isso pode ser feito. Definimos

$$\tilde{J}_{y_i}^{a|b} = [\tilde{l}_{i-1}^{a|b}, \tilde{l}_i^{a|b}]$$

onde

$$\tilde{l}_i^{a|b} = \begin{cases} l_N^{a,b}, & \text{se } i = 0; \\ \tilde{l}_{i-1}^{a|b} + \max\{0, (Q(a, y_i) - Q(b, y_i))\}, & \text{se } i \geq 1. \end{cases}$$

Observe que, contrariamente ao que acontecia na definição da primeira partição, nesta definição, tem importância a ordem em que  $a$  e  $b$  aparecem na expressão  $a|b$  ( que se lê:  $a$  dado  $b$  ) .

Finalmente definimos

$$I_{y_i}^{a|b} = J_{y_i}^{a,b} \cup \tilde{J}_{y_i}^{a|b}.$$

É fácil ver que, independentemente de  $b$ , vale a igualdade

$$|I_{y_i}^{a|b}| = Q(a, y_i).$$

Definimos a função  $\bar{F} : E \times E \times [0, 1] \rightarrow E \times E$  da seguinte maneira:

$$\bar{F}(a, b, u) = \sum_{i=1}^N \sum_{j=1}^N (y_i, y_j) 1_{\{u \in I_{y_i}^{a|b} \cap I_{y_j}^{b|a}\}}.$$

Em outras palavras,  $\bar{F}(a, b, u) = (y_i, y_j)$ , se e somente se  $u \in I_{y_i}^{a|b} \cap I_{y_j}^{b|a}$ . Note que

$$I_{y_i}^{a|b} \cap I_{y_j}^{b|a} = \begin{cases} J_{y_i}^{a,b}, & \text{se } i = j \\ \emptyset, & \text{em caso contrário.} \end{cases}$$

Em consequência, quaisquer que sejam  $a$  e  $b$ , se  $u < l_N^{a,b}$ , então

$$(3.22) \quad \bar{F}(a, b, u) = (x_i, x_i),$$

para algum  $i \in \{1, \dots, N\}$ .

Vamos construir o par de cadeias  $(X_n^a, X_n^b)_{n \geq 0}$  da seguinte maneira:

$$(X_n^a, X_n^b) = \begin{cases} (a, b), & \text{se } n = 0; \\ \bar{F}((X_{n-1}^a, X_{n-1}^b), U_n), & \text{se } n \geq 1, \end{cases}$$

onde  $(U_1, U_2, \dots)$  é uma seqüência de variáveis aleatórias independentes e uniformemente distribuídas no intervalo  $[0, 1]$ .



O restante da demonstração é idêntico ao do teorema 3.3, com a única diferença de que agora o intervalo de coincidência muda de um instante para outro, em função dos valores assumidos pelo par de cadeias.

Seja

$$K^{a,b} = \inf\{n \geq 1 : U_n < l_N^{X_n^a, X_n^b}\}.$$

A propriedade (3.22) implica que

$$\tau^{a,b} \leq K^{a,b}.$$

Note-se que a distribuição da variável aleatória  $K^{a,b}$  é majorada por uma distribuição geométrica

$$\begin{aligned} \mathbb{P}\{K^{a,b} > n\} &= \mathbb{P}\{U_1 > l_N^{X_0^a, X_0^b}, \dots, U_n > l_N^{X_{n-1}^a, X_{n-1}^b}\} \leq \\ &\leq \mathbb{P}\{U_1 > \min_{x,y} l_N^{x,y}, \dots, U_n > \min_{x,y} l_N^{x,y}\} = \\ &= \prod_{i=1}^n \mathbb{P}\{U_i > \min_{x,y} l_N^{x,y}\} = (1 - \min_{x,y} l_N^{x,y})^n. \end{aligned}$$

Para concluir a demonstração basta agora observar que

$$\min_{x,y} l_N^{x,y} = \alpha(Q). \quad \square$$

Finalmente podemos enunciar o Teorema de Convergência que conclui esta seção.

**Teorema 3.23.** *Se for  $E$  finito e se  $Q$  for uma matriz de transição definida em  $E$  irredutível e aperiódica, então*

$$\sup_{(a,b)} |\mathbb{P}\{X_n^a = b\} - \mu(b)| \leq (1 - \alpha(Q^k))^{\frac{n}{k}},$$

onde  $\mu$  é a única probabilidade invariante da cadeia e  $k$  é menor número inteiro positivo, tal que  $Q^k$  tem todos os elementos positivos.

**Prova.** Segue imediatamente do Teorema 3.5, do Corolário 3.13, da proposição 3.19 e do teorema 3.21.  $\square$

Vamos terminar essa seção mostrando outro tipo de perda de memória. O resultado que segue diz essencialmente que se a probabilidade de passar de qualquer estado a qualquer outro é limitada inferiormente por uma constante positiva, então existe um tempo aleatório que não depende do futuro tal que nesse tempo a cadeia se encontra na distribuição invariante.

**Teorema 3.24.** *Seja  $X_n$  uma cadeia de Markov em um espaço finito  $E$  tal que  $Q(x, y) \geq c > 0$  para todo  $x, y \in E$ . Seja  $\nu$  a única medida invariante para  $X_n$ . Então existe um tempo aleatório  $T$  tal que a distribuição de  $X_T$  é  $\nu$ . O tempo aleatório  $T$  tem distribuição geométrica de parâmetro  $c$  (isto é,  $\mathbb{P}(T > n) = (1 - c)^n$ ) e tem a propriedade de que o evento  $\{T = t\}$  é independente de  $\{X_s : s \geq t\}$ .*

**Prova.** A prova será feita através de uma construção particular de  $X_n$  e  $T$ . Sejam

$$U_1, U_2, \dots, V_1, V_2, \dots, W_1, W_2, \dots$$

três famílias de variáveis independentes identicamente distribuídas com distribuição uniforme no intervalo  $[0, 1]$ . Escrevemos

$$Q(x, y) = c\nu(y) + (1 - c)\tilde{Q}(x, y),$$

onde

$$\tilde{Q}(x, y) = \frac{Q(x, y) - c\nu(y)}{1 - c}.$$

Na fórmula acima suposemos que não estamos no caso trivial em que  $E$  é um conjunto unitário, e, portanto,  $c < 1$ .

Vamos definir duas partições,  $\{J_y : y \in E\}$  e  $\{\tilde{I}_y^x : x, y \in E\}$ . do intervalo  $[0, 1]$  tais que

$$|J_y| = \nu(y)$$

e

$$|\tilde{I}_y^x| = \tilde{Q}(x, y).$$

Com a ajuda dessas partições e das variáveis uniformes, para qualquer  $X_0 \in E$  fixado, vamos definir o processo  $(X_n)$  da seguinte maneira

(3.25)

$$X_{n+1} = \sum_y y \left( \mathbf{1}\{U_n \leq c, W_n \in J_y\} + \mathbf{1}\{U_n > c, V_n \in \tilde{I}_y^{X_n}\} \right).$$

O processo  $X_n$  assim definido é uma cadeia de Markov, com matriz de transição  $Q$ . Isso é deixado como exercício para o leitor.

Agora definimos

$$T = \inf\{n \geq 1 : U_n < c\} + 1.$$

Vamos calcular

$$\begin{aligned} \mathbb{P}(X_T = y) &= \sum_n \mathbb{P}(X_n = y, T = n) \\ (3.26) \quad &= \sum_n \mathbb{P}(X_n = y, U_1 > c, \dots, U_{n-1} > c, U_n \leq c). \end{aligned}$$

Segue da definição de  $X_n$  que a seguinte identidade entre eventos é verdadeira:

$$\{X_{n+1} = y, U_n < c\} = \{W_n \in J_y, U_n < c\}.$$

Portanto (3.26) é igual a

$$\begin{aligned} &\sum_{n \geq 1} \mathbb{P}(W_n \in J_y, U_1 > c, \dots, U_{n-1} > c, U_n \leq c) \\ &= \sum_{n \geq 1} \mathbb{P}(W_n \in J_y) \mathbb{P}(U_1 > c, \dots, U_{n-1} > c, U_n \leq c) \\ &= \nu(y) \sum_{n \geq 1} \mathbb{P}(T = n) = \nu(y), \end{aligned}$$

onde a primeira identidade segue da independência de  $W_n$  e  $(U_n)$ . Isso prova que o processo no instante aleatório  $T$  tem distribuição invariante.

Para ver que  $T$  tem distribuição geométrica basta observar que para  $n \geq 1$ ,

$$\{T = n\} = \{U_1 > c, \dots, U_{n-1} > c, U_n \leq c\},$$

e que, portanto,

$$\mathbb{P}(T \geq n + 1) = (1 - c)^n.$$

□

**Corolário 3.27.** *Nas condições do teorema 3.24 vale a desigualdade*

$$\sup_x |\mathbb{P}\{X_n = x\} - \nu(x)| \leq (1 - c)^n.$$

**Prova.** Segue imediatamente do teorema 3.24  $\square$

### Exercícios.

**3.28.** Verifique quais das cadeias apresentadas nos Exemplos 1.9, 1.10, 2.5, 2.24, e nos exercícios 1.11, 2.48 e 2.53 são periódicas e determine os seus períodos. Para aquelas que, além de aperiódicas são irredutíveis, determine a menor potência  $k$  a partir da qual todas as entradas da matriz de transição  $Q^k$  são estritamente positivas.

**3.29.** Seja  $E = \{1, 2\}$  e  $Q$  a matriz de transição definida da seguinte maneira

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

(a) Mostre que existe  $\bar{n}$ , tal que, para todo  $n \geq \bar{n}$  valem as desigualdades

$$0,45 \leq Q^n(1, 2) \leq 0,55 \text{ e}$$

$$0,45 \leq Q^n(2, 2) \leq 0,55.$$

Faça uma estimativa de  $\bar{n}$ .

(b) Obtenha os resultados equivalentes para  $Q^n(1, 1)$  e  $Q^n(2, 1)$ .

**3.30.** Prove que o processo definido recursivamente pela expressão (3.25) é de Markov e tem matriz de transição  $Q$ . Dica: calcule  $\mathbb{P}(X_{n+1} = y \mid X_n = x)$  e prove que é igual a  $Q(x, y)$ .

**3.31.** Determine  $c$ ,  $\beta$  e  $\alpha$  de todas as cadeias aperiódicas e irredutíveis listadas no exercício 3.28. Nos casos em que os cálculos forem muito complicados, tente determinar minorantes razoáveis para  $c$ ,  $\alpha$  e  $\beta$ . Comente os casos em que  $\alpha$  fornece uma estimativa da velocidade de convergência mais fina do que  $\beta$ . Compare com a estimativa dada por  $c$ .

**3.32.** *Acoplamento de Doeblin.* Dada uma matriz de transição  $Q$  sobre o conjunto finito ou enumerável  $E$ , definimos uma nova matriz  $\bar{Q}$  em  $E \times E$  da seguinte maneira

$$\bar{Q}((a, b), (x, y)) = Q(a, x)Q(b, y), \text{ se } a \neq b;$$

$$\bar{Q}((a, a), (x, y)) = Q(a, x), \text{ se } x = y;$$

$$\bar{Q}((a, a), (x, y)) = 0, \text{ se } x \neq y.$$

Verifique que  $\bar{Q}$  é efetivamente uma matriz de transição em  $E \times E$ , isto é, verifique que para todo  $(a, b) \in E \times E$  vale a igualdade

$$\sum_{(x, y) \in E \times E} \bar{Q}((a, b), (x, y)) = 1.$$

Observe também que a cadeia correspondente a  $\bar{Q}$  descreve a evolução de duas cadeias de Markov de matriz  $Q$  que evoluem independentemente uma da outra até se encontrarem pela primeira vez. Se isso acontecer, as duas cadeias ficam grudadas e evoluem juntas a partir de então.

**3.33.** Nas condições do Teorema 3.5, se a matriz de transição admitir uma medida de probabilidade invariante, ela será única. Sugestão: proceda como na demonstração do Teorema 3.5.

**Comentários e referências.** A técnica de acoplamento introduzida por Doeblin na década de 30. Doeblin construiu duas cadeias que evoluem independentemente até se encontrarem pela primeira vez, como no exercício 3.32. O coeficiente de Dobrushin é uma adaptação para cadeias de Markov das idéias de Dobrushin (1968), usadas para provar a existência de um único estado de Gibbs. O Teorema 3.24 que usa o fato que há regenerações em cadeias de Markov é devido a Athreya e Ney (1978).



#### 4. PROCESSOS DE RENOVAÇÃO.

Um processo de renovação é uma seqüência estritamente crescente

$$0 \leq T_1 < T_2 < \cdots < T_k < \cdots$$

de variáveis aleatórias, assumindo valores em  $\mathbb{N}$  ou em  $\mathbb{R}$  e satisfazendo as seguintes condições:

- (a) As variáveis  $T_1, T_2 - T_1, \dots, T_{k+1} - T_k, k \geq 1$  são independentes entre si.
- (b) As variáveis  $T_{k+1} - T_k, k \geq 1$  são identicamente distribuídas.

As variáveis  $(T_k)_{k \geq 1}$  modelam os instantes sucessivos de ocorrência de um determinado fenômeno que se repete, após cada ocorrência, independentemente da história passada. Por exemplo, as variáveis  $T_k$  poderiam ser os instantes sucessivos nos quais um determinado telefone recebe chamadas. A condição (a) da definição expressa de alguma forma o fato que as pessoas que efetuam as chamadas, fazem-no de forma independente. Em consequência os intervalos entre as chamadas sucessivas são variáveis independentes. A condição (b) será verificada se nós nos limitarmos a um período no qual as condições de ocorrência do fenômeno se mantem inalteradas. Isso num telefone residencial significa modelar separadamente os instantes de chamada recebidas nos períodos diurnos e noturnos e dentre estes estabelecer ainda subperíodos, de vez que é mais frequente receber telefonemas até às 10 da noite, por exemplo, do que entre meia-noite e seis da manhã. O fato de excluirmos  $T_1$  da condição (b) expressa o fato de que o tempo que decorre até a primeira ocorrência pode depender de outros fatores ligados à observação. Por exemplo, se nós estivermos modelando os instantes de passagem do metrô, num sistema perfeitamente regular, no qual os intervalos entre dois trens é constante, por exemplo 3 minutos, então  $T_{k+1} - T_k = 3$ , para todo  $k \geq 1$ . No entanto,  $T_1$  dependerá do momento que tomarmos como ponto de partida das observações. Vamos a seguir apresentar alguns exemplos.

**Exemplo 4.1.** Seja  $U_1, U_2, \dots$  uma seqüência de variáveis aleatórias

i.i.d. assumindo valores no conjunto  $\{-1, +1\}$  com

$$\mathbb{P}\{U_n = +1\} = p = 1 - \mathbb{P}\{U_n = -1\},$$

onde  $p \in [0, 1]$ . Definimos

$$T_1 = \inf\{n \geq 1 : U_n = -1\} \text{ e}$$

$$T_k = \inf\{n \geq T_{k-1} + 1 : U_n = -1\}, \text{ para todo } k \geq 2.$$

Neste caso, a independência das variáveis aleatórias  $U_n$  acarreta imediatamente a independência exigida na condição (a) e também a identidade das leis dos incrementos  $T_{k+1} - T_k$ , para todo  $k \geq 1$ . Neste caso é fácil ver que os incrementos, assim como  $T_1$ , tem distribuições geométricas.

**Exemplo 4.2.** Seja  $(X_n^a)_{n \in \mathbb{N}}$  uma cadeia de Markov irredutível, assumindo valores no espaço de estados enumerável  $E$  e tendo  $a$  como ponto de partida. Seja  $b$  um elemento qualquer fixado de  $E$ . Definimos os instantes sucessivos de passagem da cadeia pelo ponto  $b$ :

$$T_1^b = \inf\{n \geq 1 : X_n^a = b\} \text{ e}$$

$$T_k^b = \inf\{n \geq T_{k-1}^b : X_n^a = b\}, \text{ para todo } k \geq 2.$$

A propriedade de Markov implica que a sequência crescente  $(T_n^b)$  satisfaz as condições (a) e (b). Note que, neste caso, se  $a = b$ , então a lei de  $T_1^b$  será igual às leis de  $T_{k+1}^b - T_k^b$ .

**Definição 4.3.** Dado um processo de renovação  $(T_n)_{n \geq 1}$ , para todo par de instantes  $s \geq t$ , nós definimos a função de contagem  $N[s, t]$  da seguinte maneira:

$$N[s, t] = \sum_{k \geq 1} \mathbf{1}\{s \leq T_k \leq t\}.$$

Como o seu nome indica, a função de contagem  $N[s, t]$  conta o número de ocorrências do processo de renovação entre os instantes  $s$  e  $t$ .



Quando o processo de renovação assumir valores em  $\mathbb{N}$ , nós também usaremos a notação

$$N\{t\} = N[t, t] = \sum_{k \geq 1} \mathbf{1}\{T_k = t\}.$$

**Lema 4.4.** *Seja  $(T_n)_{n \geq 1}$  um processo de renovação assumindo valores em  $\mathbb{N}$  e seja  $t$  um ponto qualquer de  $\mathbb{N}$ . Então,*

$$\sum_{k \geq 1} \mathbf{1}\{T_k = t\} = \mathbf{1}\{t \in \{T_k : k \geq 1\}\},$$

e, portanto,

$$(4.5) \quad \mathbb{P}\{N\{t\} = 1\} = \mathbb{P}\{t \in \{T_k : k \geq 1\}\}.$$

**Prova.** Segue diretamente das definições e é deixada como exercício simples ao leitor.  $\square$

As questões básicas sobre os processos de renovação são as seguintes.

**Questão 1:** (*Lei dos Grandes Números.*) Determinar o limite (em média, ou em probabilidade, ou quase certamente) de

$$\lim_{t \rightarrow +\infty} \frac{N[0, t]}{t}.$$

**Questão 2:** Determinar a lei de  $T_1$  que torna a lei da função de contagem  $N[s, s + h]$  independente de  $s$ .

**Questão 3:** Determinar o limite

$$\lim_{s \rightarrow +\infty} \mathbb{P}\{N[s, s + h] = n\}.$$

Note que nas questões 1 e 3 a própria existência do limite não está a priori assegurada.

Nesta seção nós vamos estudar essas questões, limitando-nos ao caso de processo de renovação assumindo valores em  $\mathbb{N}$ . Nós mostraremos que todas as questões referentes a um processo de renovação em  $\mathbb{N}$  se traduzem em questões referentes a uma cadeia de Markov associada, tendo como espaço de estados o conjunto  $\mathbb{N}$ . Obteremos assim respostas às tres questões básicas, simplesmente utilizando o que já aprendemos sobre as cadeias de Markov.

**Lema 4.6.** (Lema Básico de Tradução.) *Seja  $(T_n)_{n \geq 1}$  um processo de renovação assumindo valores em  $\mathbb{N}$  e seja  $\nu$  a lei comum dos incrementos  $T_{k+1} - T_k$ , isto é*

$$\mathbb{P}\{T_{k+1} - T_k = n\} = \nu(n),$$

para todo  $n \in \mathbb{N}$ . Então

$$N[m, n] = \sum_{m \leq t \leq n} \mathbf{1}\{X_t^a = 0\} \quad e$$

$$\mathbb{P}\{t \in \{T_k : k \geq 1\}\} = \mathbb{P}\{X_t^a = 0\},$$

onde  $a = T_1$  e  $(X_n^a)_{n \in \mathbb{N}}$  é a cadeia de Markov tendo o conjunto  $\mathbb{N}$  dos números naturais como espaço de estados, tendo  $a$  como ponto de partida e tendo como matriz de transição  $Q_\nu$ , assim definida:

$$(4.7) \quad Q_\nu(0, x) = \nu(x + 1), \quad \text{para todo } x \in \mathbb{N};$$

e

$$(4.8) \quad Q_\nu(x, x - 1) = 1, \quad \text{para todo } x \geq 1.$$

**Prova.** Faz-se, desenhando um gráfico. No instante 0, marca-se o ponto de ordenada  $a = T_1$ . Em seguida, traça-se uma reta que desce uma unidade no eixo das ordenadas, para cada unidade avançada no eixo das abcissas. Essa reta parte do ponto  $(0, a)$ , ortogonalmente à diagonal, indo encontrar o eixo das abcissas no ponto  $(a, 0)$ . Em seguida, marca-se o ponto  $(a + 1, D_1)$ , onde  $D_k = T_{k+1} - T_k - 1$ , qualquer que seja  $k \geq 1$ . Recomeça-se o procedimento, traçando-se um segmento de reta perpendicular à diagonal, unindo os pontos  $(a + 1, D_1)$  e  $(a + 1 + D_1, 0)$ . Em seguida marcamos o ponto  $(a + 2 + D_1, D_2)$  e recomeçamos o procedimento, descendo uma reta até  $(a + 2 + D_1 + D_2, 0)$ . Em geral, marcamos o ponto  $(a + k + \sum_{j=1}^{k-1} D_j, D_k)$  e descemos um segmento de reta perpendicular à diagonal até o ponto  $(a + k + \sum_{j=1}^k D_j, 0)$ .

Para cada  $t \in \mathbb{N}$ , definimos  $X_t^a$  como sendo a ordenada do ponto de abscissa  $t$  no gráfico construído acima. É fácil ver, e é deixado

como exercício ao leitor, que o processo  $(X_t^a)_{t \in \mathbb{N}}$ , assim definido, é efetivamente uma cadeia de Markov, tendo  $Q_\nu$  como matriz de transição.

Segue também dessa construção que os pontos  $T_k$ , do processo de renovação, são precisamente os instantes nos quais a cadeia  $(X_t^a)_{t \in \mathbb{N}}$  visita o ponto 0. Isso, juntamente com o lema 4.4, encerra a demonstração.  $\square$

Daqui para a frente vamos utilizar sistematicamente a letra  $\nu$  para designar a lei comum dos incrementos  $T_{k+1} - T_k$ .

**Definição 4.9.** Seja  $\nu$  uma distribuição de probabilidades no conjunto  $\{1, 2, \dots\}$  de média  $\theta$  finita, isto é:

$$\theta = \sum_{n \geq 1} n\nu(n) < +\infty.$$

Definimos a medida de probabilidades  $G^\nu$  em  $\mathbb{N}$  da seguinte maneira. Para todo  $x \in \mathbb{N}$

$$(4.10) \quad G^\nu(x) = \frac{\nu((x, +\infty))}{\theta},$$

onde

$$\nu((x, +\infty)) = \sum_{y > x} \nu(y).$$

**Observação:** É deixado como exercício fácil para o leitor a verificação da igualdade

$$(4.11) \quad \theta = \sum_{n \geq 1} n\nu(n) = \sum_{x \geq 0} \nu((x, +\infty)).$$

A igualdade (4.11) garante que  $G^\nu$  é efetivamente uma medida de probabilidade sempre que  $\theta < +\infty$ .

Vamos agora utilizar o Lema Básico de Tradução para resolver a Questão 2.

**Proposição 4.12.** *Seja  $(T_n)_{n \geq 1}$  um processo de renovação assumindo valores em  $\mathbb{N}$  e seja  $\nu$  a lei comum dos incrementos  $T_{k+1} - T_k$ . Suponhamos que a média  $\theta$  de  $\nu$  seja finita. Então  $\mathbb{P}\{N\{t\} = 1\}$  é constante em  $t$  se e somente se  $T_1$  tem  $G^\nu$  como lei.*

**Prova.** Pelo Lema Básico de Tradução 4.6 sabemos que

$$\mathbb{P}\{N\{t\} = 1\} = \sum_{k \geq 0} \mathbb{P}(T_1 = k) \mathbb{P}\{X_t^k = 0\},$$

onde o índice superior  $k$  indica o ponto de partida da cadeia. Assim sendo, a constância em  $t$  de  $\mathbb{P}\{N\{t\} = 1\}$  é equivalente à invariância da distribuição  $\{\mathbb{P}(T_1 = k), k \geq 0\}$  para a cadeia  $(X_t)_{t \in \mathbb{N}}$ .

Isso significa que o ponto de partida da cadeia, isto é  $T_1$  é escolhido seguindo a distribuição invariante com respeito à matriz  $Q_\nu$ , definida no Lema Básico de Tradução.

Note-se que essa matriz é irredutível no conjunto

$$E_\nu = \{0, \sup\{x \geq 0 : \nu(x+1) > 0\}\}.$$

Note também que a condição  $\theta < +\infty$  é equivalente a dizer que a cadeia de matriz  $Q_\nu$  é recorrente positiva, o que assegura a existência de uma medida de probabilidade em  $\mathbb{N}$ , invariante com respeito a  $Q_\nu$ . Resta-nos verificar que a distribuição  $G^\nu$  é invariante com respeito à matriz de transição  $Q_\nu$ . Essa é uma conta simples e direta que é deixada como exercício ao leitor.  $\square$

**Corolário 4.13.** *Nas condições da Proposição 4.12, supondo que  $\theta < +\infty$  e que  $T_1$  tem  $G^\nu$  como lei, para todo  $t \in \mathbb{N}$  vale a igualdade:*

$$\mathbb{P}\{N\{t\} = 1\} = \frac{1}{\theta}.$$

**Definição 4.14.** Diremos o processo de renovação é **estacionário**, quando a lei de  $N\{t\}$  for independente de  $t$ .

Passemos agora à Questão 1.

**Proposição 4.15.** *Lei dos Grande Números para as esperanças. Nas condições da Proposição 4.12, supondo que  $\sum_{x \geq 1} xG^\nu(x) < +\infty$  e que  $\mathbb{E}(T_1) < +\infty$ , então vale o limite*

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}(N[0, t])}{t + 1} = \frac{1}{\theta}.$$

**Prova.** A demonstração será feita acoplando-se de uma maneira adequada o processo de renovação  $(T_k)_{k \geq 1}$  com um outro processo de renovação  $(S_k)_{k \geq 1}$ , cujos incrementos  $(S_{k+1} - S_k)_{k \geq 1}$  têm a mesma lei  $\nu$  que os incrementos  $(T_{k+1} - T_k)_{k \geq 1}$  do processo original, mas que além disso é estacionário, isto é,  $S_1$  tem lei  $G^\nu$  —nesta última afirmação estamos utilizando a Proposição 4.12. Este acoplamento será feito de tal forma que os dois processos de renovação, com exceção do primeiro intervalo de tempo, começando na origem, tenham em seguida exatamente os mesmos incrementos.

Começamos escolhendo  $S_1$ , com lei  $G^\nu$ , independentemente do processo  $(T_k)_{k \geq 1}$ . Em seguida, para todo  $k \geq 2$ , definimos

$$S_k = S_{k-1} + T_k - T_{k-1}.$$

Sejam  $N_T[0, t]$  e  $N_S[0, t]$ , respectivamente, as funções de contagem dos processos  $(T_k)_{k \geq 1}$  e  $(S_k)_{k \geq 1}$ , isto é

$$N_T[0, t] = \sum_{k \geq 1} \mathbf{1}\{T_k \leq t\},$$

$$N_S[0, t] = \sum_{k \geq 1} \mathbf{1}\{S_k \leq t\}.$$

Por construção, temos a desigualdade

$$(4.16) \quad |N_T[0, t] - N_S[0, t]| \leq S_1 + T_1.$$

A desigualdade (4.16) implica a desigualdade

$$(4.17) \quad \frac{1}{t+1} \mathbb{E}|N_T[0, t] - N_S[0, t]| \leq \frac{1}{t+1} \mathbb{E}(S_1 + T_1).$$

Por hipótese

$$\mathbb{E}(T_1) < +\infty \quad \text{e} \quad \mathbb{E}(S_1) = \sum_{x \geq 1} x G^\nu(x) < +\infty.$$

Nessas condições, (4.16) implica que

$$\lim_{t \rightarrow \infty} \frac{|\mathbb{E}(N_T[0, t]) - \mathbb{E}(N_S[0, t])|}{t + 1} = 0.$$

Basta agora observarmos que processo de renovação  $(S_k)_{k \geq 1}$  sendo estacionário vale a igualdade

$$\mathbb{E}N_S[0, t] = \sum_{s=0}^t \mathbb{E}(N_S\{s\}) = (t + 1) \frac{1}{\theta},$$

o que encerra a demonstração.  $\square$

**Proposição 4.18.** *Lei dos Grande Números quase certa. Se  $E_\nu$  é finito, então vale o limite*

$$(4.19) \quad \mathbb{P} \left( \lim_{t \rightarrow \infty} \frac{N[0, t]}{t} = \frac{1}{\theta} \right) = 1.$$

Na verdade a Proposição 4.18 vale sob a hipótese mais fraca que  $\mathbb{E}T_1 < \infty$  e  $\mathbb{E}(T_{k-1} - T_k) < \infty$ . Preferimos nos manter no caso finito porque a idéia da prova é a mesma mas muito mais elementar. Antes de demonstrarmos a proposição, vamos dar a seguinte definição.

**Definição.** Se (4.19) vale diremos que  $\frac{N[0, t]}{t}$  converge quase certamente a  $1/\theta$ .

**Prova.** Nessa prova vamos explorar a seguinte lei dos grandes números para variáveis aleatórias independentes: se  $\{X_i : i \geq 0\}$  é uma sequência de variáveis aleatórias independentes identicamente distribuídas, então

$$\mathbb{P} \left( \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n X_i}{n} = \mathbb{E}X_1 \right) = 1.$$

Escrevendo  $t = T_{N[0,t]+1} - T_1 + (T_{N[0,t]+1} - t) + T_1$ , obtemos

$$(4.20) \quad \frac{t}{N[0,t]} = \frac{T_{N[0,t]+1} - T_1}{N[0,t]} - \frac{T_{N[0,t]+1} - t}{N[0,t]} + \frac{T_1}{N[0,t]}.$$

Vamos tratar cada um desses três termos separadamente. O primeiro pode ser escrito como

$$\frac{\sum_{i=2}^{N[0,t]+1} (T_i - T_{i-1})}{N[0,t]}.$$

Mas  $N[0,t]$  vai para infinito quando  $t$  vai para infinito pois  $N[0,t] \geq t$  (isso porque assumimos  $\nu(0) = 0$ , mas quando  $\nu(0) \in (0, 1)$  é também possível provar que  $N[0,t]$  vai para infinito). Portanto podemos aplicar a lei dos grandes números para variáveis independentes para o primeiro termo e provar que esse termo vai para  $\theta = \mathbb{E}(T_i - T_{i-1})$ .

Como a distribuição do tempo entre renovações se concentra em um conjunto finito, o numerador do segundo termo em (4.20) é limitado. Como  $N[0,t]$  vai para infinito, esse termo converge a zero quase certamente.

O terceiro termo também vai para zero porque se trata do quociente entre uma variável aleatória fixa e outra que está indo para infinito.  $\square$

Vamos agora finalmente resolver a Questão 3. Na literatura, este resultado é muitas vezes chamado de *Teorema Chave* ou de *Teorema da Renovação*. Para não sairmos do quadro finito da seção 3, inicialmente nós o enunciaremos de forma mais restritiva do que o necessário supondo o conjunto  $E_\nu$  finito.

**Teorema 4.21.** *Teorema Chave - versão finita. Nas condições da Proposição 4.12, supondo que  $E_\nu$  é finito e que a matriz  $Q_\nu$  é aperiódica, qualquer que seja a distribuição de  $T_1$  em  $E_\nu$ , existe o limite*

$$\lim_{t \rightarrow +\infty} \mathbb{P}\{N\{t\} = 1\} = \frac{1}{\theta}.$$

Alem disso, se  $a_1, \dots, a_n \in \{0, 1\}$  são arbitrários, então para qualquer  $k$

$$\lim_{t \rightarrow +\infty} \mathbb{P}\{N\{t+i-k\} = a_i, i = 1, \dots, n\} = H(a_1, \dots, a_n),$$

onde

$$H(a_1, \dots, a_n) = \mathbb{P}\{N_S\{i\} = a_i, i = 1, \dots, n\}.$$

onde  $N_S$  é o processo de renovação estacionário introduzido na Proposição 4.15.

**Prova.** Note que  $E_\nu$  é finito implica  $\theta < +\infty$ . O resultado segue diretamente do Lema Básico de Tradução 4.6, da Proposição 4.12 e do Teorema de Convergência 3.23.  $\square$

Dois conceitos importantes em teoria de renovação são a idade e o tempo residual de um processo. Definimos a idade  $A(t)$  e o tempo residual  $R(t)$  do processo no instante  $t$  da seguinte forma

$$A(t) = t - T_{N[0,t]}; \quad R(t) = T_{N[0,t]+1} - t.$$

Intuitivamente, se pensamos que há uma renovação cada vez que uma lâmpada é trocada,  $A(t)$  representa a idade da lâmpada que está funcionando no instante  $t$  e  $R(t)$  representa o tempo de vida que ainda resta a essa lâmpada. Uma consequência do teorema chave é que podemos calcular a distribuição assintótica dessas variáveis aleatórias quando  $t \rightarrow \infty$ .

**Corolário 4.22.** *Sob as hipóteses do teorema chave, quando  $t \rightarrow \infty$ , a distribuição de  $A(t)$  e a de  $R(t)$  convergem para  $G^\nu$ .*

**Prova.** Note que

$$\mathbb{P}(A(t) = k) = \mathbb{P}(N\{t-k\} = 1, N\{t-k+1, t\} = 0),$$

o que, pelo teorema chave, converge quando  $t \rightarrow \infty$  a

$$\begin{aligned} \mathbb{P}(N_S\{0\} = 1, N_S[1, k] = 0) &= \mathbb{P}(S_1 = 0)\mathbb{P}(S_2 - S_1 > k) \\ &= \frac{1}{\theta} \nu(k, \infty) = G^\nu(k). \end{aligned}$$



Por outro lado,

$$\mathbb{P}(R(t) = k) = \mathbb{P}(N[t, t+k-1] = 0, N\{t+k\} = 1),$$

o que, pelo teorema chave, converge quando  $t \rightarrow \infty$  a

$$\mathbb{P}(N_S[t, t+k-1] = 0, N_S\{t+k\} = 1) = \mathbb{P}(S_1 = k) = G''(k). \quad \square$$

Uma prova alternativa elementar do Teorema Chave que vale para o caso  $E_\nu$  infinito, pode ser feita usando diretamente a construção do processo de renovação como fizemos a construção da cadeia de Markov no Capítulo 1. Seja  $\nu$  a lei de  $T_{i+1} - T_i$  para  $i \geq 1$  e  $\nu'$  a lei de  $T_1$ . Defina

$$\rho_k = \frac{\nu(k)}{\sum_{i \geq k} \nu(i)}; \quad \rho'_k = \frac{\nu'(k)}{\sum_{i \geq k} \nu'(i)}.$$

Sejam  $U_i$ ,  $i \geq 1$ , variáveis aleatórias independentes uniformes em  $[0, 1]$ . Defina

$$T_1 = \min\{n : U_n \leq \rho'(n)\}$$

e

$$T_{k+1} = \min\{n > T_k : U_n \leq \rho(n - T_k)\}.$$

**Lema 4.23.** *As variáveis  $T_1$  e  $T_{k+1} - T_k$  tem distribuição  $\nu'$  e  $\nu$  respectivamente para  $k \geq 1$ . Além disso, são independentes. Em outras palavras, o processo acima construído é uma versão do processo de renovação com distribuições  $\nu'$  e  $\nu$  para  $T_1$  e  $T_{k+1} - T_k$  respectivamente.*

**Prova.** Vamos provar que  $\mathbb{P}(T_1 = k) = \nu'(k)$ .

$$\mathbb{P}(T_1 = k) = \mathbb{P}(U_1 > \rho'_1, \dots, U_{k-1} > \rho'_{k-1}, U_k \leq \rho'_k).$$

Como as variáveis  $U_i$  são independentes, a expressão acima é igual a

$$\mathbb{P}(U_1 > \rho'_1) \dots \mathbb{P}(U_{k-1} > \rho'_{k-1}) \mathbb{P}(U_k \leq \rho'_k) = (1 - \rho'_1) \dots (1 - \rho'_{k-1}) \rho'_k.$$

Mas  $1 - \rho'_i = \nu'[i+1, \infty) / \nu'[i, \infty)$ . Portanto ficamos com

$$\frac{\nu'[2, \infty)}{\nu'[1, \infty)} \dots \frac{\nu'[k, \infty)}{\nu'[k-1, \infty)} \frac{\nu'(k)}{\nu'[k, \infty)} = \nu'(k).$$

(Usamos que  $\nu'[1, \infty) = 1$ .) O restante da prova é deixada como exercício fácil para o leitor.  $\square$

**Teorema 4.24.** *Teorema Chave com velocidade de convergência. Se existe uma constante  $\gamma \in (0, 1)$  tal que para todo  $n \geq 0$*

(4.25)

$$\mathbb{P}(T_1 = n \mid T_1 \geq n) \geq \gamma; \quad \mathbb{P}(T_{k+1} - T_k = n \mid T_{k+1} - T_k \geq n) \geq \gamma.$$

Então,

$$|\mathbb{P}(N\{t\} = 1) - (1/\theta)| \leq \gamma^t.$$

**Prova.** Em primeiro lugar note que sob as hipóteses do teorema,

$$\rho'_n > \gamma, \quad \text{e} \quad \rho_n > \gamma,$$

para todo  $n \geq 1$ . Por outro lado, para a variável  $S_1$  que tem distribuição  $G^\nu$ , valem também as desigualdades

$$\mathbb{P}(S_1 = n \mid S_1 \geq n) \geq \gamma.$$

Para provar essas desigualdades, note primeiro que elas são equivalentes a

$$(4.26) \quad \frac{\mathbb{P}(S_1 > n + 1)}{\mathbb{P}(S_1 > n)} \leq 1 - \gamma.$$

Para provar (4.26), inicialmente notemos que

$$\mathbb{P}(S_1 > n) = \sum_{i=n+1}^{\infty} \mathbb{P}(S_1 = i) = \frac{1}{\theta} \sum_{i=n+1}^{\infty} \nu[i, \infty).$$

Por definição,

$$\nu[i, \infty) = \mathbb{P}(T_{k+1} - T_k > i - 1) \geq \frac{1}{1 - \gamma} \mathbb{P}(T_{k+1} - T_k > i),$$

a desigualdade seguindo da hipótese do teorema. Por outro lado, aplicando de novo a definição, a última expressão é igual a

$$\frac{1}{1 - \gamma} \nu[i + 1, \infty).$$

Portanto

$$\mathbb{P}(S_1 > n) \geq \frac{1}{1-\gamma} \frac{1}{\theta} \sum_{i=n+2}^{\infty} \nu[i, \infty) = \frac{1}{1-\gamma} \mathbb{P}(S_1 > n+1),$$

onde as desigualdades seguem termo a termo da hipótese da proposição. Agora, definindo

$$\rho_S(n) = \frac{\mathbb{P}(S_1 = n)}{\mathbb{P}(S_1 > n)},$$

das desigualdades (4.26) segue que  $\rho_S(n) > \gamma$ .

Vamos acoplar o processo  $N[0, t]$  com distribuição inicial  $\nu'$  com o processo  $N_S[0, t]$  com salto inicial  $S_1$  com distribuição  $G^\nu$ . Sabemos que esse segundo processo é estacionário. Isso implica que  $\mathbb{P}(N_S\{t\} = 1) = 1/\theta$  para todo  $t$ . Definimos

$$\tau = \min\{n : N\{n\} = 1, N_S\{n\} = 1\}.$$

Temos então que

$$\begin{aligned} & |\mathbb{P}(N\{t\} = 1) - (1/\theta)| \\ &= |\mathbb{P}(N\{t\} = 1) - \mathbb{P}(N_S\{t\} = 1)| \\ &= |\mathbb{E}(\mathbf{1}\{N\{t\} = 1\} - \mathbf{1}\{N_S\{t\} = 1\})| \\ &= 0 \times \mathbb{P}(\tau \leq t) + |\mathbb{E}(\mathbf{1}\{N\{t\} = 1\} - \mathbf{1}\{N_S\{t\} = 1\} | \tau > t)| \mathbb{P}(\tau > t) \\ &\leq \mathbb{P}(\tau > t). \end{aligned}$$

A última desigualdade segue porque a diferença das funções indicadoras é no máximo 1. A esperança da diferença das funções indicadoras indica que os dois processos estão sendo realizados com as mesmas variáveis aleatórias  $U_n$ . Como para todo  $n$ ,  $\rho(n) > \gamma$ ,  $\rho'(n) > \gamma$  e  $\rho_S(n) > \gamma$ , sabemos que  $\{U_n < \gamma\} \subset \{N\{n\} = 1, N_S\{n\} = 1\}$ . Portanto  $\tau$  está dominada por uma variável geométrica  $\tau^0$  de parâmetro  $\gamma$  definida por

$$\tau^0 = \min\{n : U_n < \gamma\}. \quad \square$$

### Exercícios

4.27. Seja  $U_1, U_2, \dots$  uma sequência de variáveis aleatórias i.i.d. assumindo valores no conjunto  $\{-1, +1\}$  com

$$\mathbb{P}\{U_n = +1\} = p = 1 - \mathbb{P}\{U_n = -1\},$$

onde  $p \in [0, 1]$ . Definimos

$$T_1 = \inf\{n \geq 1 : U_n = -1\} \text{ e}$$

$$T_k = \inf\{n \geq T_{k-1} : U_n = -1\} \text{ para todo } k \geq 2.$$

i) Mostre que as variáveis aleatórias  $T_1, T_2 - T_1, \dots, T_{k+1} - T_k, \dots$  são independentes entre si e identicamente distribuídas com

$$\mathbb{P}\{T_1 = n\} = \mathbb{P}\{T_{k+1} - T_k = n\} = p^{n-1}(1-p), \text{ para todo } n \geq 1.$$

ii) Sejam  $m < n$  dois elementos quaisquer de  $\mathbb{N}$ . Calcule

$$\mathbb{P}\{N[m, n] = k\}, \text{ para todo } k \in \mathbb{N}.$$

Calcule

$$\mathbb{E}(N[m, n]).$$

4.28. Seja  $(X_n^a)_{n \in \mathbb{N}}$  uma cadeia de Markov irredutível, assumindo valores no espaço de estados enumerável  $E$  e tendo  $a$  como ponto de partida. Seja  $b$  um elemento qualquer fixado de  $E$ . Definimos os instantes sucessivos de passagem da cadeia pelo ponto  $b$

$$T_1^b = \inf\{n \geq 1 : X_n^a = b\} \text{ e}$$

$$T_k^b = \inf\{n \geq T_{k-1}^b : X_n^a = b\} \text{ para todo } k \geq 2.$$

i) Mostre que a sequência crescente  $(T_n^b)_{n \geq 1}$  é um processo de renovação.

ii) Suponha que a cadeia seja recorrente positiva e que seu ponto inicial seja escolhido com a distribuição invariante  $\mu$ . Sejam  $m < n$  dois elementos quaisquer de  $\mathbb{N}$ . Mostre que

$$\mathbb{E}(N[m, n]) = (n - m + 1)\mu(b).$$

**4.29.** Seja  $\nu$  uma distribuição de probabilidades no conjunto dos números naturais  $\{1, 2, \dots\}$ . Mostre que

$$\theta = \sum_{n \geq 1} n\nu(n) = \sum_{x \geq 0} \nu((x, +\infty)).$$

o que demonstra a igualdade (4.11).

**4.30.** Calcule a forma exata da distribuição  $G^\nu$ , definida pela equação (4.10), nos seguintes casos.

i) A distribuição  $\nu$  é degenerada e dá massa 1 ao ponto  $c$ , isto é

$$\nu(x) = \begin{cases} 1 & \text{se } x = c; \\ 0 & \text{se } x \neq c. \end{cases}$$

ii) Para todo  $p \in (0, 1)$ ,  $\nu$  é a distribuição geométrica de média  $\frac{1}{1-p}$  em  $\{1, 2, \dots\}$ , isto é

$$\nu(n) = p^{n-1}(1-p), \text{ para todo } n \geq 1.$$

**4.31.** Seja  $\nu$  uma distribuição de probabilidades no conjunto dos números naturais  $\{1, 2, \dots\}$  de média  $\theta$  finita, isto é:

$$\theta = \sum_{n \geq 1} n\nu(n) < +\infty.$$

Sejam  $G^\nu$  e  $Q_\nu$  a medida de probabilidade e matriz de transição no conjunto dos números naturais  $\mathbb{N}$ , definidas respectivamente pelas fórmulas (4.10), (4.7) e (4.8).

i) Mostre que a matriz  $Q_\nu$  é irredutível no conjunto

$$E_\nu = \{0, \sup\{x \geq 0 : \nu(x+1) > 0\}\}.$$

ii) Dê exemplos de condições suficientes garantindo a aperiodicidade de  $Q_\nu$ .

iii) Mostre que  $G^\nu$  é invariante com respeito a  $Q_\nu$ .

**4.32.** Nas condições do corolário 4.13 determine a distribuição da função de contagem  $N[t, t+1]$ , onde  $t$  é um número natural qualquer.

**4.33.** (Paradoxo da inspeção) Usando o Corolário 4.22 ao Teorema chave, calcule a distribuição do comprimento do intervalo que contém o ponto  $t$ , quando  $t \rightarrow \infty$ . Prove que a esperança desse comprimento é igual a  $2\theta$ .

**Comentários e referências.** O teorema chave é um célebre resultado de Blackwell. Sua demonstração via cadeias de Markov utilizando o lema de tradução faz parte do folklore probabilístico mas não conhecemos nenhum lugar onde esteja escrito. O resultado do decaimento exponencial assumindo taxa de falha limitada por baixo é uma espécie de resultado de “alta temperatura” que logicamente não é ótimo, mas de demonstração muito simples. Lindvall (1992) apresenta demonstrações alternativas para o teorema chave na versão mais geral usando acoplamento. Em particular demonstra que se  $\nu$  tem decaimento exponencial, então o teorema chave vale com convergência exponencialmente rápida (como no Teorema 4.24).

## 5. PROCESSOS DE POISSON

**Construção do processo de Poisson unidimensional.**

Um processo pontual em  $\mathbb{R}$  é uma sequência estritamente crescente de variáveis aleatórias  $\dots S_{-1} \leq S_0 \leq 0 \leq S_1, \dots$  em  $\mathbb{R}$ . As variáveis  $S_n$  podem ser interpretadas como os instantes sucessivos de ocorrência de um evento.

Vamos construir um processo pontual  $S_n \in \mathbb{R}$ ,  $n \in \mathbb{Z}$ , que será chamado de processo de Poisson.

Começamos fazendo uma partição de  $\mathbb{R}$  em intervalos  $A_i$ ,  $i \in \mathbb{Z}$ ,  $A_i = [a_i, a_{i+1})$ , com  $a_i < a_{i+1}$ . Chamamos  $|A_i|$  o comprimento  $a_{i+1} - a_i$  do intervalo  $A_i$ . Assim  $\dot{\cup}_{i \in \mathbb{Z}} A_i = \mathbb{R}$ , onde  $\dot{\cup}$  significa união disjunta.

O segundo passo consiste em considerar para cada  $i$  uma variável aleatória  $Y_i$  com distribuição de Poisson de média  $\lambda|A_i|$ . Isto é

$$\mathbb{P}(Y_i = k) = \frac{e^{-\lambda|A_i|} (\lambda|A_i|)^k}{k!}.$$

Assumimos que as variáveis  $Y_i$  são independentes.

Para cada  $i \in \mathbb{Z}$  consideramos uma sequência de variáveis independentes  $\{U_{i,j} : j = 1, 2, \dots\}$ , uniformes em  $A_i$ :

$$\mathbb{P}(U_{i,j} \in A \cap A_i) = \frac{|A \cap A_i|}{|A_i|}.$$

Seja  $S$  o conjunto aleatório

$$S = \bigcup_{i \in \mathbb{Z}} S_i,$$

onde

$$S_i = \begin{cases} \{U_{i,j} : 1 \leq j \leq Y_i\}, & \text{se } Y_i \geq 1; \\ \emptyset, & \text{se } Y_i = 0. \end{cases}$$

Ou seja, em cada intervalo  $A_i$  colocamos  $Y_i$  pontos distribuídos uniformemente no intervalo.

O processo pontual é obtido ordenando os pontos de  $\mathbf{S}$ . Seja  $\{S_n\}$  a sequência ordenada de pontos de  $\mathbf{S}$ , sendo  $S_1$  o primeiro ponto à direita da origem. Mais formalmente

$$S_1 = \min\{s > 0 : s \in \mathbf{S}\}$$

e

$$(5.1) \quad S_n = \begin{cases} \min\{s > S_{n-1} : s \in \mathbf{S}\}, & \text{se } n \geq 2; \\ \max\{s < S_{n+1} : s \in \mathbf{S}\}, & \text{se } n \leq 0. \end{cases}$$

Para  $A \subset \mathbb{R}$ , definimos  $N_{\mathbf{S}}(A)$  = número de pontos do conjunto  $\mathbf{S} \cap A$ . É claro que

$$N_{\mathbf{S}}(A) = \sum_n \mathbf{1}\{S_n \in A\}.$$

Quando não houver margens à dúvida escreveremos apenas  $N(A)$  em lugar de  $N_{\mathbf{S}}(A)$ .

O processo de Poisson que acabamos de construir é um caso particular de processo pontual. Já vimos no caso discreto outro exemplo: o processo de renovação. A definição formal de Processo pontual em  $\mathbb{R}$  é a seguinte. Consideremos o conjunto

$$\mathcal{M} = \{\mathbf{S} \subset \mathbb{R} : N_{\mathbf{S}}(A) < \infty\}$$

das possíveis realizações de um processo pontual que não tem pontos de acumulação.

Definir um processo pontual consiste em definir uma medida de probabilidade em  $\mathcal{M}$ . Como  $\mathcal{M}$  não é enumerável, isso passa por uma construção sofisticada da noção de evento em  $\mathcal{M}$  usando teoria da medida, que ultrapassa o escopo deste texto. Vamos nos limitar a apresentar os eventos essenciais que são da forma

$$\mathcal{A} = \{\mathbf{S} : N_{\mathbf{S}}(B_i) = b_i, i = 1, \dots, \ell\},$$

onde  $B_i$  são retângulos de dimensão  $d$  de área finita e  $b_i$  são números naturais finitos arbitrários.

Na sequência usaremos o seguinte teorema que será apresentado sem demonstração.



**Teorema 5.2.** *Uma medida de probabilidade em  $\mathcal{M}$  é inteiramente definida pelas probabilidades dos eventos pertencentes a  $\mathcal{A}$ .*

### Propriedades.

**Lema 5.3.** *Para cada intervalo  $A$ , a variável aleatória  $N(A)$  tem distribuição de Poisson com média  $\lambda|A|$ .*

**Prova.** Note primeiro que como  $A \cap A_i$  são conjuntos disjuntos, as variáveis  $N(A \cap A_i)$  são independentes. Calculemos a distribuição de  $N(A \cap A_i)$ . Por construção,

$$\begin{aligned} \mathbb{P}(N(A \cap A_i) = k) &= \sum_{h \geq k} \mathbb{P}(Y_i = h, N(A \cap A_i) = k) \\ &= \sum_{h \geq k} \mathbb{P} \left( Y_i = h, \sum_{j=0}^h \mathbf{1}\{U_{i,j} \in A \cap A_i\} = k \right). \end{aligned}$$

Como as variáveis  $U_{i,j}$  são independentes de  $Y_i$ , podemos fatorar a última probabilidade, obtendo

$$= \sum_{h \geq k} \mathbb{P}(Y_i = h) \mathbb{P} \left( \sum_{j=0}^h \mathbf{1}\{U_{i,j} \in A \cap A_i\} = k \right).$$

Mas  $U_{i,j}$  são variáveis independentes e uniformes em  $A_i$ . Portanto  $\mathbf{1}\{U_{i,j} \in A \cap A_i\}$  são variáveis independentes de Bernoulli com parâmetro  $\frac{|A \cap A_i|}{|A_i|}$ . A soma dessas variáveis tem distribuição binomial com parâmetros  $n$  e  $\frac{|A \cap A_i|}{|A_i|}$ . Assim a última expressão é igual a

$$\begin{aligned} \sum_{h \geq k} \frac{e^{-\lambda|A_i|} (\lambda|A_i|)^h}{h!} \binom{h}{k} \left( \frac{|A \cap A_i|}{|A_i|} \right)^k \left( 1 - \frac{|A \cap A_i|}{|A_i|} \right)^{h-k} \\ = \frac{e^{-\lambda|A \cap A_i|} (\lambda|A \cap A_i|)^k}{k!}. \end{aligned}$$

Para terminar a prova observe que  $N(A) = \sum N(A \cap A_i)$  é uma soma de variáveis aleatórias independentes com distribuição de Poisson de médias  $\lambda|A \cap A_i|$ . Como os  $A_i$  são disjuntos,  $\sum |A \cap A_i| = |A|$ . Assim,  $N(A)$  é Poisson com média  $|A|$ .  $\square$

**Lema 5.4.** *Para cada família finita de intervalos disjuntos  $B_l$ ,  $l = 1, \dots, L$ , as variáveis aleatórias  $N(B_l)$  são independentes e tem distribuição de Poisson com média  $\lambda|B_l|$ , respectivamente.*

**Prova.** A demonstração segue o mesmo padrão do lema anterior. É fácil verificar que para  $i$  fixo, dado  $N(A_i) = h_i$ , as variáveis

$$\{N(B_l \cap A_i) : l = 1, \dots, L\}$$

tem distribuição multinomial, isto é, para valores arbitrários inteiros  $k_{l,i} \geq 0$ , tais que  $\sum_{l=1}^L k_{l,i} = h_i$ ,

$$(5.5) \quad \mathbb{P}(N(B_l \cap A_i) = k_{l,i} \mid N(A_i) = h_i) = \frac{h_i!}{\prod_{l=1}^{L+1} k_{l,i}!} \prod_{l=1}^{L+1} (b_{l,i})^{k_{l,i}},$$

onde

$$k_{L+1,i} = h_i - \sum_{l=1}^L k_{l,i}, \quad b_{l,i} = \frac{|A_i - \cup_{l=1}^L B_l \cap A_i|}{|A_i|}$$

para  $1 \leq l \leq L$  e  $b_{L+1} = 1 - \sum_{l=1}^L b_l$ .

Como as variáveis  $\{N(A_i) : i \in \mathbb{Z}\}$  são independentes, segue de (5.5) que  $\{N(B_l \cap A_i) : l = 1, \dots, L, i \in \mathbb{Z}\}$  é uma família de variáveis aleatórias independentes com distribuição de Poisson com parâmetros  $\lambda|N(B_l \cap A_i)|$ , respectivamente. Basta somar em  $i$  e utilizar o fato que soma de variáveis aleatórias independentes com distribuição de Poisson é Poisson para concluir a prova.  $\square$

**Lema 5.6.** *Para qualquer intervalo  $A$ , a distribuição condicionada dos pontos de  $\mathbf{S} \cap A$  dado que  $N(A) = n$  é a mesma que a de  $n$  variáveis aleatórias independentes distribuídas uniformemente em  $A$ .*

**Prova.** Vamos utilizar o Teorema 5.2. Seja  $B_1, \dots, B_L$  uma partição de  $A$  e  $n_1, \dots, n_L$  inteiros não negativos tais que  $n_1 + \dots + n_L = n$ .

$$\begin{aligned} \mathbb{P}(N(B_l) = n_l, l = 1, \dots, L \mid N(A) = n) \\ &= \frac{\mathbb{P}(N(B_l) = n_l, l = 1, \dots, L)}{\mathbb{P}(N(A) = n)} \\ &= \frac{n!}{n_1! \dots n_L!} \prod_{l=1}^L \left( \frac{|B_l|}{|A|} \right)^{n_l} \end{aligned}$$

Tendo em vista o Teorema 5.2, basta agora provar que se  $U_1, \dots, U_L$  são variáveis aleatórias independentes com distribuição uniforme em  $A$ , e  $M(B) = \sum_i \mathbf{1}\{U_i \in B\}$ , então

$$\mathbb{P}(M(B_l) = n_l, l = 1, \dots, L) = \frac{n!}{n_1! \dots n_L!} \prod_{l=1}^L \left( \frac{|B_l|}{|A|} \right)^{n_l},$$

o que é deixado como exercício para o leitor.  $\square$

**Corolário 5.7.** *A distribuição condicionada de  $(S_1, S_2, \dots, S_n)$  dado  $N([0, t]) = n$  é a mesma que a de  $(Y_1, Y_2, \dots, Y_n)$ , onde  $Y_i$  são as estatísticas de ordem de  $(U_1, U_2, \dots, U_n)$ , variáveis independentes distribuídas uniformemente no intervalo  $[0, t]$ :*

$$Y_1 = \min\{U_1, U_2, \dots, U_n\};$$

$$Y_i = \min(\{U_1, U_2, \dots, U_n\} \setminus \{Y_1, \dots, Y_{i-1}\}), i = 2, \dots, n.$$

**Observação 5.8.** Os Lemas 5.4 e 5.6 provam que a escolha dos conjuntos  $A_i$  é inessencial para a construção do processo.

### Propriedade Markoviana do processo de Poisson.

Vamos apresentar uma construção alternativa do processo de Poisson unidimensional. Seja  $T_1$  uma variável aleatória exponencial de média  $1/\lambda$  e  $\tilde{N}(\cdot)$  um processo de Poisson de taxa  $\lambda$ , independente de  $T_1$ .

Seja  $N(\cdot)$  o processo definido por

$$(5.9) \quad N(B) = \mathbf{1}\{T_1 \in B\} + \tilde{N}(B - T_1),$$

onde

$$B - t = \{x \in \mathbb{R} : x - t \in B\}.$$

Em outras palavras, o processo  $N(\cdot)$  é obtido fixando o primeiro evento com a variável  $T_1$  e colando a continuação do instante desse evento um processo de Poisson independente.

**Teorema 5.10.** *O processo pontual definido por (5.9) é um processo de Poisson de taxa  $\lambda$ .*

**Prova.** Em vista do Teorema 5.2, basta provar que para o processo  $N(\cdot)$  acima construído vale o Lema 5.4. Queremos então calcular

$$\mathbb{P}(N(B_l) = k_l, l = 1, \dots, L),$$

para intervalos  $B_l$ ,  $k_l \in \mathbb{N}$  e  $L \geq 1$  arbitrários. Para simplificar a apresentação da prova vamos apenas considerar  $L = 1$  e  $B_1 = [a, c]$ . A extensão para todo  $L$  é deixada como exercício. Vamos condicionar ao valor assumido por  $T_1$ .

$$\begin{aligned} \mathbb{P}(N([a, c]) = k) &= \mathbb{P}(N([a, c]) = k, T_1 < a) \\ &\quad + \mathbb{P}(N([a, c]) = k, T_1 \in [a, c]) + \mathbb{P}(N([a, c]) = k, T_1 > c). \end{aligned}$$

Vamos primeiro assumir  $k = 0$ . Nesse caso o termo do meio é zero e condicionando no valor de  $T$  obtemos que o primeiro termo é igual a

$$\begin{aligned} \mathbb{P}(N([a, c]) = 0, T_1 < a) &= \int_0^a \lambda e^{-\lambda t} \mathbb{P}(\tilde{N}([a-t, c-t]) = 0) dt \\ &= \int_0^a \lambda e^{-\lambda t} e^{-\lambda[c-t-(a-t)]} dt \\ &= e^{-\lambda[c-a]}(1 - e^{-\lambda a}). \end{aligned}$$

Por outro lado, o terceiro termo é igual a

$$\mathbb{P}(T_1 > c) = e^{-\lambda c}.$$

Assim, somando os dois termos obtemos

$$\mathbb{P}(N([a, c]) = 0) = e^{-\lambda[c-a]},$$

que é o que queríamos demonstrar para  $k = 0$ .

Para  $k > 0$  o terceiro termo é nulo e o primeiro termo é

$$\begin{aligned} \mathbb{P}(N[a, c] = k, T_1 < a) &= \int_0^a \lambda e^{-\lambda t} \mathbb{P}(\tilde{N}([a-t, c-t]) = k) dt \\ &= \int_0^a \lambda e^{-\lambda t} \frac{e^{-\lambda[c-t-(a-t)]} \lambda^k [c-t-(a-t)]^k}{k!} dt \\ &= \frac{e^{-\lambda[c-a]} \lambda^k [c-a]^k}{k!} (1 - e^{-\lambda a}) \end{aligned}$$

e o segundo termo é igual a

$$\begin{aligned} \mathbb{P}(N[a, c] = k, T_1 \in [a, c]) &= \int_a^c \lambda e^{-\lambda t} \mathbb{P}(\tilde{N}([0, c-t]) = k-1) dt \\ &= \int_a^c \lambda e^{-\lambda t} \frac{e^{-\lambda[c-t]} \lambda^{k-1} [c-t]^{k-1}}{(k-1)!} dt \\ &= \frac{e^{-\lambda c} \lambda^k [c-a]^k}{k!}. \end{aligned}$$

A soma desses dois termos fornece o desejado

$$\mathbb{P}(N[a, c] = k) = \frac{e^{-\lambda[c-a]} \lambda^k [c-a]^k}{k!},$$

o que termina a prova do teorema.  $\square$

**Corolário 5.11.** *As variáveis  $X_n = S_n - S_{n-1}$ ,  $n \geq 2$ , são independentes, identicamente distribuídas com distribuição exponencial de taxa  $\lambda$ .*

**Prova.** Vimos no Teorema 5.10 que um processo de Poisson  $N_0(\cdot)$  pode ser construído colocando o instante do primeiro evento de acordo a uma variável exponencial  $T_1$  e a continuação um processo de Poisson  $N_1(\cdot)$  independente de  $T_1$ . O processo  $N_1(\cdot)$ , por seu lado, pode ser construído com uma variável exponencial que chamaremos  $S_2$  e um processo de Poisson independente chamado  $N_2(\cdot)$ . Continuando por indução temos que os instantes entre eventos sucessivos

formam uma sequência de variáveis independentes  $S_1, S_2, \dots$  com distribuição exponencial.  $\square$

### Definições alternativas de processos de Poisson.

No começo desse capítulo vimos uma definição construtiva do processo de Poisson. Nos livros aparecem usualmente outras definições. Nós consideraremos esses processos como definidos em toda a reta real  $\mathbb{R}$ . Usaremos a notação  $N(t)$  para  $N([0, t])$ . Diremos que um processo unidimensional  $N(\cdot)$  tem *incrementos estacionários* se a distribuição de  $N[s, t]$  é a mesma que a de  $N[s + r, t + r]$ . Diremos que um processo  $N(\cdot)$  tem *incrementos independentes* se para  $A$  e  $B$  disjuntos  $N(A)$  é independente de  $N(B)$ .

**Definição 5.12.** Diremos que um processo  $N(\cdot)$  é de Poisson com taxa  $\lambda$  se

- i. O processo tem incrementos independentes e estacionários.
- ii. A variável  $N[s, t]$  tem distribuição de Poisson com taxa  $\lambda(t - s)$ .

**Notação.** Diremos que uma função  $f$  é um  $o(h)$  se

$$\lim_{h \rightarrow 0} \frac{f(h)}{h} = 0.$$

**Definição 5.13.** Diremos que um processo  $N(\cdot)$  é de Poisson com taxa  $\lambda$  se

- i. O processo tem incrementos independentes e estacionários;
- ii.  $\mathbb{P}(N[t, t + h] = 1) = \lambda h + o(h)$ ;
- iii.  $\mathbb{P}(N[t, t + h] \geq 2) = o(h)$ .

**Observação.** Cada uma das definições acima definem univocamente um processo como consequência do Teorema 5.2.

**Proposição 5.14.** *As definições 5.12 e 5.13 são equivalentes.*

**Prova.** Não é muito difícil provar que a Definição 5.12 implica na Definição 5.13. Para isso usando que  $N[t, t + h]$  tem distribuição de

Poisson de parâmetro  $\lambda h$ , podemos escrever

$$\begin{aligned}\mathbb{P}(N[t, t+h] = 1) &= \lambda h e^{-\lambda h} = \lambda h \left( 1 - \lambda h + \frac{(\lambda h)^2}{2} + \dots \right) \\ &= \lambda h + \lambda h \left( -\lambda h + \frac{(\lambda h)^2}{2} + \dots \right),\end{aligned}$$

usando o desenvolvimento em série de  $e^{-\lambda h}$ . Mas o segundo termo do último membro é justamente um  $o(h)$ . Isso prova 5.13.ii. Por outro lado,

$$\mathbb{P}(N[t, t+h] \geq 2) = e^{-\lambda h} \left( \frac{(h\lambda)^2}{2!} + \frac{(h\lambda)^3}{3!} + \dots \right) = o(h).$$

Isso prova 5.13.iii.

A prova que Definição 5.13 implica na Definição 5.12 é mais complicada e envolve a solução de equações diferenciais. Essa prova será omitida nesse texto, mas pode ser encontrada por exemplo em Ross (1983).  $\square$

**Proposição 5.15.** *O processo  $S_n$  definido em (5.1) satisfaz as condições das Definições 5.12 e 5.13.*

**Prova.** Por construção o processo tem incrementos independentes e estacionários e pelo Lema 5.3 a distribuição de pontos em  $[s, t]$  tem distribuição de Poisson de parâmetro  $\lambda(t-s)$ . Isso implica que satisfaz a Definição 5.12. Pela proposição precedente, também satisfaz a Definição 5.13.  $\square$

### O paradoxo da inspeção.

Suponha que os ônibus de uma cidade saem com intervalos de uma hora e de duas horas alternativamente. Os horários são 0, 1, 3, 4, 6, 7, 9, 10, 12, etc. A metade dos intervalos entre chegadas tem comprimento 1 e a outra metade tem comprimento 2. Se escolhermos um instante de tempo  $T$  ao acaso entre as 0 e as 12 horas quanto teremos que esperar em média para a saída do próximo ônibus?

Um raciocínio seria assumir que os intervalos estão guardados em uma urna. Portanto o tempo médio seria calculado escolhendo primeiro ao acaso um intervalo e depois calculando o tempo médio de espera para o intervalo escolhido. Nesse caso escolher um intervalo de tamanho 1 tem probabilidade  $1/2$  e para esse tamanho o tempo médio de espera é  $1/2$ . A probabilidade de escolha de um intervalo de tamanho 2 é  $1/2$ , e para esse intervalo o tempo médio de espera é 1 hora. O resultado seria

$$ET = \frac{1}{2} \frac{1}{2} + \frac{1}{2} 1 = \frac{3}{4}.$$

Isso corresponde a calcular o comprimento médio de meio intervalo. Esse raciocínio está **errado**.

O problema é que, ao se escolher um um ponto uniformemente no intervalo  $[0, 12]$ , a probabilidade dele pertencer a um intervalo de comprimento 2 é  $8/12$  enquanto que a probabilidade de pertencer a um intervalo de comprimento 1 é  $4/12$ . Isso dá

$$ET = \frac{4}{12} \frac{1}{2} + \frac{8}{12} 1 = \frac{5}{6} > \frac{3}{4}.$$

Isto é, o tempo médio de espera é maior que o comprimento médio de meio intervalo. Isso porque a probabilidade de escolher um intervalo comprido é mais alta que a de escolher um intervalo curto. O mesmo acontece quando temos um processo pontual estacionário na reta: o comprimento do intervalo que contém a origem é em geral maior que o comprimento do intervalo típico. Vejamos por exemplo o que acontece com o processo de Poisson.

Seja  $S_1$  o tempo da primeira ocorrência do processo à direita da origem e  $S_0$  o instante da primeira ocorrência à esquerda. Já vimos que  $S_1$  tem distribuição exponencial de taxa  $\lambda$ . O mesmo argumento serve para provar que  $S_0$  tem também a mesma distribuição. Como o processo tem incrementos independentes, a distância do primeiro ponto à direita da origem é independente da distância do primeiro ponto à esquerda e portanto a distribuição do tamanho do intervalo que contém a origem  $S_1 - S_0$  é a mesma que a da soma de duas exponenciais independentes. Isto é uma gama de parâmetros 2 e  $\lambda$ :

$$\mathbb{P}(S_1 - S_0 > t) = \int_t^\infty \lambda x e^{-\lambda x} dx.$$



e o valor médio deste intervalo é

$$\mathbb{E}(S_1 - S_0) = \int_0^{\infty} \lambda x^2 e^{-\lambda x} dx = \frac{2}{\lambda}.$$

que é o dobro do valor médio do intervalo típico, que como vimos é uma exponencial de taxa  $\lambda$ .

Notemos agora que o intervalo que contem um instante qualquer  $t$  também tem comprimento médio 2. Para  $t \geq 0$ , esse intervalo é  $S_{N(t)+1} - S_{N(t)}$ . Pelo mesmo raciocínio anterior,  $S_{N(t)+1} - t$  e  $t - S_{N(t)}$  são variáveis independentes com distribuição exponencial. Portanto  $S_{N(t)+1} - S_{N(t)}$  tem distribuição gama de parâmetro 2.

O único processo no qual o comprimento do intervalo que contem um ponto é o mesmo que o do intervalo típico é o processo para o qual os tempos entre chegadas são constantes.

### Processos de Poisson em duas ou mais dimensões.

Nessa seção vamos construir um processo de Poisson em duas dimensões. O processo obtido será utilizado nas seções seguintes para construir processos unidimensionais não homogêneos e superposição de processos de Poisson.

Vamos construir um conjunto aleatório de pontos em  $\mathbb{R}^2$ . A mesma construção serve em  $\mathbb{R}^d$  para  $d \geq 2$  mas por simplicidade ficaremos em dimensão 2. A construção é muito parecida com a de dimensão 1.

Começamos com uma partição de  $\mathbb{R}^2$  em retângulos finitos  $A_i$ . Por exemplo os  $A_i$  podem ser os quadrados determinados pelo reticulado  $\mathbb{Z}^2$ . Denotamos  $|A_i|$  a área de  $A_i$ . Temos

$$\dot{\cup}_{i \in \mathbb{Z}^2} A_i = \mathbb{R}^2,$$

onde  $\dot{\cup}$  significa união disjunta.

Para cada  $i$ , seja  $Y_i$  uma variável aleatória com distribuição de Poisson de média  $\lambda|A_i|$ . Isto é

$$\mathbb{P}(Y_i = k) = \frac{e^{-\lambda|A_i|} (\lambda|A_i|)^k}{k!}.$$

Assumimos que as variáveis  $Y_i$  são independentes.

Finalmente, para cada  $i$  consideramos uma sequência de variáveis independentes  $(U_{i,j})_{j \geq 1}$  uniformes em  $A_i$ :

$$\mathbb{P}(U_{i,j} \in A \cap A_i) = \frac{|A \cap A_i|}{|A_i|}.$$

O processo pontual será o seguinte conjunto (aleatório):

$$\mathbf{S} = \cup_{i \in \mathbb{Z}} \cup_{j=1}^{Y_i} \{U_{i,j}\},$$

onde usamos a convenção  $\cup_{j=1}^0 \{U_{i,j}\} = \emptyset$ . Ou seja, em cada retângulo  $A_i$  colocamos  $Y_i$  pontos distribuídos uniformemente no retângulo.

Até aqui repetimos o procedimento de  $d = 1$ . A diferença agora é que não há uma maneira satisfatória de ordenar os pontos do conjunto aleatório  $\mathbf{S}$ . Mas isso não é importante.

Para  $A \subset \mathbb{R}$ , conjunto mensurável de  $\mathbb{R}^2$ , definimos  $M(A) =$  número de pontos do conjunto  $\mathbf{S} \cap A$ . Usaremos a letra  $M$  para duas dimensões e a letra  $N$  para uma dimensão.

Valem as seguintes propriedades do processo. As provas são idênticas às dos lemas correspondentes da sessão anterior e por isso omitidas.

**Lema 5.16.** *Para cada conjunto  $A$  finito a variável aleatória  $M(A)$  tem distribuição de Poisson com média  $\lambda|A|$ .*

**Lema 5.17.** *Para cada família finita de conjuntos disjuntos mensuráveis  $B_i$ , as variáveis aleatórias  $M(B_i)$  tem distribuição de Poisson com média  $\lambda|B_i|$  e são independentes.*

**Lema 5.18.** *Para qualquer conjunto mensurável  $A$ , a distribuição dos pontos de  $\mathbf{S} \cap A$  dado que  $M(A) = n$  é a mesma que a de  $n$  variáveis aleatórias independentes distribuídas uniformemente em  $A$ .*

**Exemplo 5.19.** Dada essa construção podemos calcular a distribuição da variável aleatória que mede a distância do ponto que

está mais perto da origem. Seja  $V = \inf\{|x| : x \in \mathbf{S}\}$ .

$$\mathbb{P}(V > b) = \mathbb{P}(N(B(0, b)) = 0) = e^{-\lambda\pi b^2},$$

onde  $B(0, b)$  é o círculo centrado na origem de raio  $b$ .

### Construção de processos unidimensionais como projeção de processos bidimensionais.

Suponhamos que a função aleatória  $M(\cdot)$  descreve o processo de Poisson bi-dimensional de parâmetro 1. Queremos construir um processo de dimensão 1 que será representado pela função aleatória  $N(\cdot)$ . Para isso definimos para um intervalo  $I \subset \mathbb{R}$

$$(5.20), \quad N(I) = M(I \times [0, \lambda])$$

ou seja, o número de pontos no intervalo  $I$  será o mesmo que o número de pontos no retângulo  $I \times [0, \lambda]$ . Isso é equivalente a projetar os pontos na faixa  $\mathbb{R} \times [0, \lambda]$  em  $\mathbb{R}$ .

**Lema 5.21.** *O processo determinado por  $N(\cdot)$  é um processo de Poisson unidimensional de taxa  $\lambda$ .*

**Prova.** Basta provar que para  $I_1, \dots, I_n$  disjuntos,

$$\mathbb{P}(N(I_i) = k_i) = \prod_{i=1}^n \frac{e^{-\lambda|I_i|} (\lambda|I_i|)^{k_i}}{k_i!},$$

o que segue imediatamente da definição (5.20).

A estas alturas o leitor deve estar se perguntando se dois pontos do processo bi-dimensional podem ser projetados para um único ponto do unidimensional. A resposta é não e a prova está dada no seguinte lema.

**Lema 5.22.** *Seja  $I$  um intervalo finito. O evento “dois pontos do processo de Poisson bidimensional são projetados em um no intervalo  $I$ ” tem probabilidade zero.*

**Prova.** Vamos assumir que  $I = [0, 1]$  sem perda de generalidade para facilitar a notação. Particionamos  $I$  em intervalinhos de comprimento  $\delta$ :  $I_n^\delta = (n\delta, (n+1)\delta]$ . A probabilidade de dois pontos

serem projetados em 1 é menor ou igual à probabilidade de dois pontos pertencer ao mesmo intervalinho, que é dada por

$$\mathbb{P}(\cup_{n=1}^{\lfloor I/\delta \rfloor} \{M(I_n^\delta \times [0, \lambda]) \geq 2\}) \leq \sum_{n=1}^{\lfloor I/\delta \rfloor} \mathbb{P}(M(I_n^\delta \times [0, \lambda]) \geq 2) \leq \frac{\lfloor I \rfloor}{\delta} o(\delta),$$

que tende a zero quando  $\delta$  vai para zero.  $\square$

A grande vantagem de construir o processo unidimensional a partir do bi-dimensional é que podemos superpor processos unidimensionais com taxas diferentes de tal maneira que um tenha sempre mais eventos que o outro.

Lembremos da definição de acoplamento. O vetor aleatório  $(X, Y)$  é um acoplamento de  $X$  e  $Y$  se as marginais do vetor  $(X, Y)$  tem a mesma distribuição de  $X$  e  $Y$ , respectivamente.

Por exemplo, duas versões independentes de  $X$  e  $Y$  formam um acoplamento de  $X$  e  $Y$ . Mas em geral esse acoplamento não é muito interessante.

Vamos construir um acoplamento de dois processos de Poisson  $N_1(\cdot)$  e  $N_2(\cdot)$  com parâmetros respectivamente  $\lambda_1$  e  $\lambda_2$ . Para  $i = 1, 2$  sejam

$$(5.23) \quad N_i(I) = M(I \times [0, \lambda_i]).$$

**Lema 5.24.** *Assuma  $\lambda_1 \geq \lambda_2$ . Então para o processo acoplado definido em (5.23) vale*

$$N_1(I) \geq N_2(I),$$

para todo intervalo  $I \subset \mathbb{R}$ .

**Prova.** Segue da definição que  $N_2$  projeta menos pontos que  $N_1$ .  $\square$

### Superposição de processos de Poisson.

Suponha que em um banco chegam clientes homens de acordo com um processo de Poisson de taxa  $p\lambda$  e clientes mulheres de acordo com um processo de Poisson independente do anterior com taxa  $(1-p)\lambda$ .

Uma maneira de modelar o processo das chegadas incluindo o atributo de cada chegada (homem ou mulher) é construir um processo bidimensional  $M(\cdot)$  como na subseção precedente e definir  $N_1(\cdot)$  e  $N_2(\cdot)$  da maneira seguinte:

$$N_1(I) = M(I \times [0, p\lambda]); \quad N_2(I) = M(I \times [p\lambda, \lambda]),$$

ou seja, os pontos que estão na faixa  $[0, \lambda] \times \mathbb{R}$  são projetados e indicam os instantes de chegada de clientes independentemente do sexo. Os pontos que estão na faixa  $[0, p\lambda] \times \mathbb{R}$  são marcados com 1 (que faremos corresponder a homem) e os pontos que estão na faixa determinada pelo intervalo vertical  $[p\lambda, \lambda] \times \mathbb{R}$  são marcados com 0 (que faremos corresponder mulher).

Vimos na seção anterior que o processo  $N(\cdot)$  definido por  $N(I) = N_1(I) + N_2(I)$  é um processo de Poisson de taxa  $\lambda$  porque podemos exprimir  $N(I) = M([0, \lambda] \times I)$ .

Sejam  $S_n$  os instantes de chegada de clientes independentemente do sexo, ou seja os instantes de chegada do processo  $N(\cdot)$ . Cada um desses pontos tem uma marca 1 ou 0 de acordo com a faixa donde foi projetado. Seja

$$G(S_i) = \begin{cases} 1 & \text{se } S_i \text{ está marcado 1} \\ 0 & \text{se } S_i \text{ está marcado 0.} \end{cases}$$

**Proposição 5.25.** *As variáveis  $G(S_i)$  são independentes, identicamente distribuídas com marginais*

$$\mathbb{P}(G(S_i) = 1) = p.$$

**Prova.** Sejam  $V_1, V_2, \dots$  e  $W_1, W_2, \dots$  duas sequências de variáveis aleatórias independentes em  $[0, 1]$  e independentes entre si.

Para cada  $n$  seja  $\pi_n$  a seguinte permutação aleatória do conjunto  $\{1, \dots, n\}$  (isto é, uma bijeção desse conjunto em si mesmo.) definida por: para  $i = 1, \dots, n-1$ ,

$$V_{\pi_n(i)} \leq V_{\pi_n(i+1)}.$$

Ou seja,  $\pi_n(i)$  é o índice da  $i$ -ésima variável quando os primeiros  $n$   $V_j$  são ordenados de menor a maior.

Para cada  $n$  fixo construímos uma sequência de variáveis aleatórias independentes no retângulo  $[0, t] \times [0, \lambda]$  da seguinte maneira:

$$(5.26) \quad U_i = (tV_{\pi_n(i)}, \lambda W_i).$$

Deixamos para o leitor provar que  $(U_i)$  é um conjunto de  $n$  variáveis independentes uniformemente distribuídas em  $[0, t] \times [0, \lambda]$ . A vantagem dessa construção é que  $W_i$  é a ordenada do  $i$ -ésimo dos  $n$   $U_j$ , quando esses vetores são ordenados de acordo com a abscissa, de menor a maior.

Seja  $L$  fixo, e para  $1 \leq k \leq L$  sejam  $a_k \in \{0, 1\}$  arbitrários. Defina

$$A_L = \{G(S_1) = a_1, \dots, G(S_L) = a_L\}.$$

Por uma versão discreta do Teorema 5.2, para provar que  $(G(S_i))$  são Bernoulli independentes de parâmetro  $p$  basta provar que

$$(5.27) \quad \mathbb{P}(A_L) = p^{\sum a_k} (1-p)^{\sum (1-a_k)}.$$

Para cada  $t$  fixo

$$\mathbb{P}(A_L) = \sum_{n=L}^{\infty} \mathbb{P}(A_L, N(0, t) = n) + \mathbb{P}(A_L, N(0, t) < n).$$

Pela lei dos grandes números para processos de Poisson unidimensionais, o segundo termo do membro da direita vai a zero quando  $t$  vai para infinito. Portanto se formos capazes de provar que o primeiro termo é igual a (5.27) para qualquer  $t$ , a prova estará terminada. Agora

$$(5.28) \quad \mathbb{P}(A_L, N(0, t) = n) = \mathbb{P}(A_L | N(0, t) = n) \mathbb{P}(N(0, t) = n).$$

Mas  $N(0, t) = M([0, t] \times [0, \lambda])$  e dado que  $M([0, t] \times [0, \lambda]) = n$ , a distribuição dos pontos nesse retângulo é a mesma que a de  $n$  uniformes independentes. Sejam

$$I_k = \begin{cases} [0, \lambda p] & \text{se } a_k = 1 \\ [\lambda p, \lambda] & \text{se } a_k = 0. \end{cases}$$

Podemos então usar a construção (5.26) para obter que

$$\begin{aligned} \mathbb{P}(A_L \mid N(0, t) = n) &= \mathbb{P}(A_L \mid M([0, t] \times [0, \lambda] = n) \\ &= \mathbb{P}(U_{\pi_n(k)} \in [0, t] \times I_k, 1 \leq k \leq n) \\ &= \mathbb{P}(\lambda W_k \in I_k, 1 \leq k \leq n) \\ &= p^{\sum a_k} (1 - p)^{\sum (1 - a_k)}. \end{aligned}$$

Como essa identidade não depende de  $n$ , somando (5.28) obtemos o resultado desejado.  $\square$

A prova da proposição acima nos leva à seguinte construção alternativa para um processo de Poisson bidimensional  $M(\cdot)$  de taxa 1 na faixa  $[0, \infty] \times [0, \lambda]$ . Seja  $T_1$  uma variável exponencial de parâmetro  $\lambda$ ,  $W_1$  uma variável uniforme em  $[0, \lambda]$  e  $M_1(\cdot)$  um processo de Poisson bidimensional de taxa 1 na faixa  $[0, \infty] \times [0, \lambda]$ . Assuma que  $T_1$ ,  $W_1$  e  $M_1(\cdot)$  são independentes.

Defina  $M(\cdot)$  como o processo

$$M(A) = \mathbf{1}\{(T_1, W_1) \in A\} + M_1(A - T_1),$$

onde

$$A - t = \{(x, y) \in \mathbb{R}^2 : (x + t, y) \in A\}.$$

Argumentos muito parecidos com os do Teorema 5.10, do corolário 5.11 e da Proposição 5.25 mostram o seguinte teorema.

**Teorema 5.29.** *Seja  $M(\cdot)$  um processo de Poisson bidimensional de taxa 1. Sejam  $S_1, S_2, \dots$  os instantes ordenados de ocorrência de eventos na faixa  $[0, \lambda]$ . Sejam  $W_1, W_2, \dots$  as segundas coordenadas desses eventos. Então,  $(S_{i+1} - S_i)_{i \geq 1}$  são variáveis independentes com distribuição exponencial de taxa  $\lambda$  e  $(W_i)_{i \geq 1}$  são variáveis independentes com distribuição uniforme em  $[0, \lambda]$ . Além disso  $(S_{i+1} - S_i)_{i \geq 1}$  e  $(W_i)_{i \geq 1}$  são independentes.*

### Processos não homogêneos.

Seja  $\lambda(t)$  uma função contínua por partes e não negativa. Assumiremos que em qualquer intervalo finito  $I \subset \mathbb{R}$ , o número de discontinuidades de  $\lambda(t)$  é finito.

Queremos construir um processo pontual com incrementos independentes que tenha taxa instantanea  $\lambda(t)$ . Ou seja, pretendemos achar um processo estocástico  $N(\cdot)$  tal que, para os pontos de continuidade de  $\lambda(t)$ ,

$$(5.30) \quad \mathbb{P}(N([t, t+h]) = 1) = h\lambda(t) + o(h)$$

$$(5.31) \quad \mathbb{P}(N([t, t+h]) \geq 2) = o(h).$$

(Veja a Definição 5.13.)

Para isso consideramos o processo bidimensional  $M(\cdot)$  e definimos

$$(5.32) \quad N(I) = M(F(I)),$$

onde  $F(I) = \{(x, y) \in \mathbb{R}^2 : x \in I \text{ e } y \leq \lambda(x)\}$ . Ou seja  $N(\cdot)$  é o processo obtido ao projetar os pontos de  $M(\cdot)$  que estão abaixo da função  $\lambda(t)$ .

**Lema 5.33.** *O processo construído em (5.32) satisfaz as condições (5.30) e (5.31).*

**Prova.** Pela definição (5.32),

$$\mathbb{P}(N([t, t+h]) = 1) = \mathbb{P}(M(F[t, t+h]) = 1).$$

Sejam  $y_0 = y_0(t, h)$  e  $y_1 = y_1(t, h)$  o ínfimo e o supremo, respectivamente, de  $\lambda(t)$  no intervalo  $[t, t+h]$ . Assim obtemos as seguintes cotas

$$M([t, t+h] \times [0, y_0]) \leq M(F[t, t+h]) \leq M([t, t+h] \times [0, y_1]).$$

Como a função  $\lambda(t)$  é continua em  $t$ ,

$$\lim_{h \rightarrow 0} y_0(t, h) = \lim_{h \rightarrow 0} y_1(t, h) = \lambda(t)$$

e  $y_1(t, h) - y_0(t, h) = O(h)$ , por continuidade, onde  $O(h)$  é uma notação para indicar uma função de  $h$  que quando dividida por  $h$  e mandando  $h$  a zero, permanece limitada por cima e por baixo. Portanto,

$$\mathbb{P}(M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) \geq 1) = o(h).$$



Assim,

$$\begin{aligned} \mathbb{P}(M(F[t, t+h]) = 1) \\ &= \mathbb{P}(M(F[t, t+h]) = 1, \\ &\quad M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) = 0) \\ &\quad + \mathbb{P}(M(F[t, t+h]) = 1, M([t, t+h] \times [0, y_1]) \\ &\quad \quad - M([t, t+h] \times [0, y_0]) \geq 1). \end{aligned}$$

O primeiro termo é igual a

$$\begin{aligned} \mathbb{P}(M([t, t+h] \times [0, y_0]) = 1, \\ M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) = 0), \end{aligned}$$

o que, usando independência do processo  $M(\cdot)$ , nos dá

$$\begin{aligned} \mathbb{P}(M([t, t+h] \times [0, y_0]) = 1) \\ \quad \times \mathbb{P}(M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) = 0) \\ = h\lambda(t) + o(h). \end{aligned}$$

O segundo termo é limitado por

$$\mathbb{P}(M([t, t+h] \times [0, y_1]) - M([t, t+h] \times [0, y_0]) \geq 1) = o(h). \quad \square$$

O processo de Poisson não homogêneo construído em (5.32) tem a propriedade que o número de pontos em um intervalo qualquer é uma variável de Poisson cujo parâmetro é a área sob a função  $\lambda(t)$  no intervalo. Isto é

$$\mathbb{P}(N([0, t]) = n) = \frac{e^{-\mu(t)} (\mu(t))^n}{n!},$$

onde  $\mu(t) = \int_0^t \lambda(s) ds$ .

A definição (5.32) de processo não homogêneo tem a vantagem de permitir a construção conjunta de dois ou mais processos com taxas diferentes com a seguinte propriedade. Sejam  $\lambda_1(t)$  e  $\lambda_2(t)$  duas funções contínuas por partes que satisfazem

$$\lambda_1(t) \leq \lambda_2(t) \text{ para todo } t$$

Então é possível acoplar processos de Poisson não homogêneos  $N_1(\cdot)$  e  $N_2(\cdot)$  de taxas  $\lambda_1(t)$  e  $\lambda_2(t)$  respectivamente em tal forma que para todo intervalo  $I$ ,

$$N_1(I) \leq N_2(I).$$

A prova dessa afirmação é imediata a partir da definição.

### Exercícios.

**5.34.** Sejam  $X$  e  $Y$  são variáveis aleatórias independentes com distribuição exponencial de parâmetros  $\lambda_1$  e  $\lambda_2$  respectivamente,

(a) Calcule a distribuição de  $Z = \min(X, Y)$ .

(b) Qual a distribuição condicional de  $Z$  dado que  $X = x$ ?

**5.35.** Prove que a soma de  $n$  variáveis independentes de Poisson com médias  $\lambda_1, \dots, \lambda_n$ , respectivamente, tem distribuição de Poisson com média  $\lambda_1 + \dots + \lambda_n$ .

**5.36.** Em uma urna são colocadas  $N_1$  bolas tipo 1,  $N_2$  bolas tipo 2 e  $N_3$  bolas tipo 3, onde  $N_i$  são variáveis de Poisson independentes com média  $\lambda_i$ .

(a) Uma bola é escolhida ao acaso da urna. Prove que, dado que  $N_1 + N_2 + N_3 \geq 1$ , a probabilidade da bola escolhida ser de tipo 1 é  $\lambda_1/(\lambda_1 + \lambda_2 + \lambda_3)$ .

(b) Prove que o resultado de (a) é o mesmo se condicionarmos a  $N_1 + N_2 + N_3 = n$  para algum  $n \geq 1$ .

(c) Prove que dado  $N_1 + N_2 + N_3 = n \geq 1$ , a distribuição dos tipos das bolas é uma trinomial de parâmetros  $n$  e  $p_i = \lambda_i/(\lambda_1 + \lambda_2 + \lambda_3)$ .

(d) Generalize o item (c) para  $k \geq 1$  tipos diversos de bolas:

$$\mathbb{P}(N_i = n_i \mid N = n) = \frac{n!}{n_1! \dots n_k!} \lambda_1^{n_1} \dots \lambda_k^{n_k},$$

para  $n_1 + \dots + n_k = n \geq 1$ .

**5.37.** Seja  $N(t)$  um processo de Poisson de parâmetro  $\lambda$ . Para

$s < r < t$  calcule

$$P(N(s) = k, N(r) - N(s) = j | N(t) = n).$$

**5.38.** Prove que as variáveis  $U_k$  definidas em (5.26) são uniformes em  $[0, t] \times [0, \lambda]$  e independentes.

**5.39.** Prove o Teorema 5.29.

**5.40.** (i) Seja  $N$  uma variável aleatória com distribuição de Poisson de parâmetro  $\lambda$ . Seja  $B = X_1 + \dots + X_N$ , onde  $X_i$  é Bernoulli de parâmetro  $p$  independente de  $N$ . Prove que  $B$  é Poisson de parâmetro  $\lambda p$ . Sugestão: construa um processo  $M(\cdot)$  em  $[0, 1] \times [0, \lambda]$  e marque os pontos de acordo com o fato de estarem por cima ou por baixo da linha  $y = \lambda p$ .

(ii) Prove que se  $T_n$  são os instantes sucessivos de um processo de Poisson de parâmetro  $\lambda$ , então o processo definido pelos tempos

$$S_n = \inf\{T_i > S_{n-1} : X_i = 1\},$$

com  $S_1 > 0$ , é um processo de Poisson de parâmetro  $\lambda p$ . Sugestão: construa simultaneamente  $(S_n)$  e  $(T_n)$  usando a Definição (5.20).

**5.41.** Suponha que uma variável aleatória satisfaz a propriedade seguinte. Para todo  $t \geq 0$

$$\mathbb{P}(X \in (t, t+h) | X > t) = \lambda h + o(h).$$

Prove que  $X$  tem distribuição exponencial de taxa  $\lambda$ .

**5.42.** Para um processo de Poisson de parâmetro  $\lambda$ ,

(a) Calcule a distribuição de  $T_2$ , o instante da segunda chegada.

(b) Calcule a distribuição de  $T_2$  dado que  $N(0, t) = 4$ .

**5.43.** Seja  $N(0, t)$  um processo de Poisson de parâmetro  $\lambda$ . Calcule

(a)  $\mathbb{E}N(0, 2)$ ,  $\mathbb{E}N(4, 7)$ ,

(b)  $\mathbb{P}(N(4, 7) > 2 | N(2, 4) \geq 1)$ ,

- (c)  $\mathbb{P}(N(0, 1) = 2 \mid N(0, 3) = 6)$ ,  
 (d)  $\mathbb{P}(N(0, t) = \text{ímpar})$ ,  
 (e)  $\mathbb{E}(N(0, t)N(0, t + s))$ .

**5.44.** Para um processo de Poisson  $N(0, t)$  calcule a distribuição conjunta de  $S_1, S_2, S_3$ , os instantes de chegadas dos três primeiros eventos.

**5.45.** Usando as mesmas idéias que na Proposição 4.18 e o fato de conhecermos as distribuições da idade  $A(t)$  e do resíduo  $R(t)$  para processos de Poisson estacionários, prove que  $(N(t)/t) \rightarrow 1/\lambda$ , quando  $t \rightarrow \infty$ .

**5.46.** Em um banco chegam mulheres de acordo com um processo de Poisson de parâmetro 3 por minuto e homens de acordo com um processo de Poisson independente do primeiro de taxa 4 por minuto. Calcule as seguintes probabilidades:

- (a) Chegue um homem antes de uma mulher.  
 (b) Chegam 3 homens antes de quinta mulher.  
 (c) Cheguem 3 clientes nos primeiros 3 minutos.  
 (d) Cheguem 2 homens e nenhuma mulher nos primeiros 3 minutos.  
 (e) Cheguem 3 mulheres dado que chegaram 7 clientes nos primeiros 10 minutos.

**5.47.** Suponha que no caixa de um supermercado chegam apenas dois tipos de clientes: aqueles que pagam com cartão, fazem isso de acordo com um processo de Poisson  $\{N_1(t), t \geq 0\}$  com parâmetro 6 por hora e aqueles que pagam com cheque que chegam de acordo com um processo de Poisson  $\{N_2(t), t \geq 0\}$  com taxa 8 por hora.

- (a) Prove que o processo  $N(t) = \{N_1(t) + N_2(t), t \geq 0\}$  de chegada de clientes no caixa é um processo de Poisson com taxa 14.  
 (b) Qual a probabilidade de o primeiro cliente a chegar no banco pagar com cartão?  
 (c) Se nos primeiros dez minutos não chegou nenhum cliente, qual a

probabilidade de os três próximos clientes pagarem com cheque?

**5.48.** Em um posto de pedágio chegam veículos de acordo com um processo de Poisson de parâmetro 3 por minuto (180 por hora). A probabilidade de cada veículo ser um carro (e não um caminhão) depende do instante de chegada do veículo e está dada pela função  $p(s)$ :

$$p(s) = \begin{cases} s/12 & \text{se } 0 < s < 6 \\ 1/2 & \text{se } 6 \leq s < 12 \\ 1/4 & \text{se } 12 \leq s < 18 \\ (1/4) - s/24 & \text{se } 18 \leq s < 24. \end{cases}$$

(a) Qual a probabilidade que chegue pelo menos um caminhão entre 5 da manhã e 10 da manhã?

(b) Dado que chegaram 30 veículos entre 0 e 1 hora da manhã, qual a distribuição do tempo de chegada do primeiro veículo do dia?

**5.49.** Construa um processo de Poisson não homogêneo com taxa  $\lambda(t) = e^{-t}$ . Calcule o número médio de eventos no intervalo  $[0, \infty]$ .

**5.50.** Prove que para o processo não homogêneo definido em (5.32), a distribuição de  $\{S_1, \dots, S_n\}$  em  $[0, t]$  dado  $N(0, t) = n$  é a mesma que a de  $n$  variáveis independentes com densidade

$$\frac{\lambda(r)}{\int_0^t \lambda(s) ds}; \quad r \in [0, t].$$

**Comentários e referências.** A construção de processo de Poisson apresentada neste capítulo é um caso particular da construção geral apresentada por Neveu (1977). A construção de processos de dimensão menor usando processos de Poisson de dimensão maior foi introduzida por Kurtz (1989) e utilizada por Garcia (1995).



## 6. PROCESSOS DE MARKOV A TEMPO CONTINUO

### Processos Markovianos de Salto.

A diferença essencial entre as cadeias de Markov vistas no capítulo 2 e os processos de salto que veremos neste capítulo é a seguinte. Uma cadeia de Markov muda de estado nos instantes inteiros de tempo:  $1, 2, \dots$ . Os processos de salto mudam em instantes reais aleatórios  $\tau_1, \tau_2, \dots$ .

Vamos definir um processo  $X_t$  em um conjunto enumerável de estados que por simplicidade tomaremos  $\mathbb{Z}$ . Queremos que a taxa de salto de um estado  $x$  a um estado  $y$  seja fixa e dada por uma função  $q(x, y) \geq 0$ . Isto é, queremos que nosso processo satisfaça

$$(6.1) \quad \mathbb{P}(X_{t+h} = y \mid X_t = x) = hq(x, y) + o(h).$$

Assumimos que as taxas de saída estejam uniformemente limitadas, isto é

$$(6.2) \quad \lambda = \sup_x \sum_y q(x, y) = \sup_x \mu(x) < \infty.$$

Observe que se o espaço de estados for finito, a condição (6.2) estará satisfeita automaticamente. Seja

$$\mu(x) = \sum_y q(x, y),$$

a taxa de saída do estado  $x$ .

Vamos construir o processo da seguinte forma. Seja  $M(\cdot)$  um processo de Poisson bi-dimensional. Seja  $I^x = [0, \mu(x)]$ . Para cada estado  $x$  fazemos uma partição de  $I^x$  em intervalos disjuntos  $I_y^x$  de comprimento  $q(x, y)$ .

Suponhamos que  $X_0 = x_0$ . Seja  $\tau_1$  o primeiro instante em que aparece uma marca do processo  $M(\cdot)$  no intervalo  $I^{x_0}$ :

$$\tau_1 = \inf\{t > 0 : M(I^{x_0} \times [0, t]) > 0\}.$$

Defina  $x_1 = y$ , se

$$\begin{aligned} & \inf\{t > 0 : M(I^{x_0} \times [0, t]) > 0\} \\ & = \inf\{t > 0 : M(I_y^{x_0} \times [0, t]) > 0\}, \end{aligned}$$

ou seja,  $x_1$  está determinado pelo intervalo  $I_{x_1}^{x_0}$  que realiza o ínfimo.

Suponha agora que  $\tau_{n-1}$  e  $x_{n-1}$  estejam determinados. Definimos indutivamente

$$\tau_n = \inf\{t > \tau_{n-1} : M(I^{x_{n-1}} \times (\tau_{n-1}, t]) > 0\}.$$

e

$$\begin{aligned} x_n = y \text{ se } & \inf\{t > \tau_{n-1} : M(I^{x_{n-1}} \times (\tau_{n-1}, t]) > 0\} \\ & = \inf\{t > 0 : M(I_y^{x_{n-1}} \times (\tau_{n-1}, t]) > 0\}. \end{aligned}$$

**Definição.** Para todo  $t \in [0, \infty)$  definimos  $\tau_\infty = \sup_n \tau_n$  e

$$(6.3) \quad X_t = x_n, \text{ se } t \in [\tau_n, \tau_{n+1}).$$

Dessa maneira, para cada realização (aleatória) do processo de Poisson bidimensional  $M(\cdot)$ , construímos (deterministicamente) uma realização do processo  $X_t$  para  $t \in [0, \tau_\infty)$ . Observe que  $\tau_n$  é o instante do  $n$ -ésimo salto do processo  $X_t$  e  $x_n$  é o  $n$ -ésimo estado assumido pelo processo.

**Proposição 6.4.** *O processo  $(X_t)_{t \in [0, \tau_\infty)}$  acima definido satisfaz (6.1).*

**Prova.** Por definição,

$$\begin{aligned} & \mathbb{P}(X_{t+h} = y \mid X_t = x) \\ & = \mathbb{P}\{M(I_y^x \times (t, t+h]) = 1\} + \mathbb{P}(\text{outras coisas}), \end{aligned}$$

onde o evento {outras coisas} está contido no evento

$$\{M([0, \mu(x)] \times (t, t+h]) \geq 2\}$$



(dois ou mais eventos acontecem no retângulo  $[0, \mu(x)] \times (t, t + h]$ ). É claro pela definição de  $M(\cdot)$  que

$$\begin{aligned}\mathbb{P}(M([0, \mu(x)] \times (t, t + h]) \geq 2) &= o(h) \quad \text{e} \\ \mathbb{P}(M(I_y^x \times (t, t + h]) = 1) &= hq(x, y) + o(h).\end{aligned}$$

Isto termina a prova da proposição.  $\square$

**Exemplo 6.5.** Vamos construir um processo que assume apenas 3 estados: 0, 1, 2. Um exemplo disso é um sistema formado por dois servidores que rejeitam clientes que chegam quando ambos já estão ocupados. Se as chegadas acontecem de acordo a um processo de Poisson de taxa  $\lambda$  e os serviços são exponenciais de taxa  $\mu$ , o processo terá taxas:

$$\begin{aligned}q(0, 1) &= q(1, 2) = \lambda \\ q(1, 0) &= \mu; \quad q(2, 1) = 2\mu \\ q(x, y) &= 0, \text{ nos outros casos.}\end{aligned}$$

A construção descrita na proposição pode ser feita com os seguintes intervalos:

$$\begin{aligned}I_1^0 &= I_2^1 = [0, \lambda] \\ I_0^1 &= [\lambda, \lambda + \mu]; \quad I_1^2 = [0, 2\mu].\end{aligned}$$

Todas as taxas são menores que  $\max\{\lambda + \mu, 2\mu\}$ .

**Exemplo 6.6.** *Processo de nascimento puro.* Vamos construir um processo  $X(t) \in \{0, 1, 2, \dots\}$  com as seguintes propriedades

$$(6.7) \quad \mathbb{P}(X(t+h) = x+1 \mid X(t) = x) = \lambda x h + o(h)$$

$$(6.8.) \quad \mathbb{P}(N((t, t+h]) \geq 2) = o(h)$$

Ou seja que a taxa de chegada no instante  $t$  é proporcional ao número de chegadas até esse instante. Dado o processo bidimensional  $M(\cdot)$  de taxa 1, construiremos uma curva aleatória  $\lambda(t)$  da seguinte forma. Seja

$$\tau_1 = \inf\{t > 0 : M([0, t] \times [0, \lambda]) = 1$$

e em geral

$$\tau_n = \inf\{t > \tau_{n-1} : M((\tau_{n-1}, t] \times [0, (n-1)\lambda]) = 1\}.$$

A partir disso podemos definir a curva  $\lambda(t)$  por

$$\lambda(t) = n\lambda \text{ se } t \in [\tau_n, \tau_{n+1}).$$

Definimos agora

$$(6.9) \quad N(I) = M(F(I))$$

onde  $F(I) = \{(x, y) \in \mathbb{R}^2 : x \in I \text{ e } y \leq \lambda(x)\}$ . Ou seja  $N(\cdot)$  é o processo obtido ao projetar os pontos de  $M(\cdot)$  que estão por baixo da função  $\lambda(t)$ . O processo construído em (6.9) satisfaz as condições (6.7) e (6.8). Veremos nos exercícios que é possível acoplar dois processos de nascimento puro com taxas  $\mu_1 \geq \mu_2$  de tal maneira que um fique sempre por cima do outro.

### Explosões.

É fácil ver que se o espaço de estados é finito, então  $\tau_\infty$  é infinito. Isto é, não pode ocorrer um número infinito de saltos em um tempo finito. Agora veremos o que pode acontecer quando o espaço de estados é infinito.

Vamos supor que o processo assuma valores em  $\mathbb{N}$  e tenha taxas

$$q(x, y) = \begin{cases} 2^x, & \text{se } y = x + 1 \\ 0, & \text{em caso contrário.} \end{cases}$$

Este processo pode ser construído a partir de um processo de Poisson bidimensional em  $[0, \infty] \times [0, \infty]$  da seguinte maneira. Defina  $\tau_n$  e  $X_t$  assumindo  $\tau_0 = 0$  e

$$\begin{aligned} \tau_n &= \inf\{t > \tau_{n-1} : M((\tau_{n-1}, t] \times [0, 2^n]) = 1\} \\ X_t &= n, \text{ para } t \in [\tau_n, \tau_{n+1}). \end{aligned}$$

Se  $x_0 = 0$ , o tempo transcorrido até o  $n$ -ésimo salto será

$$\tau_n = \sum_{i=0}^n T_i,$$

onde  $T_i$  é uma variável aleatória exponencial de média  $2^{-i}$ . Portanto,

$$\mathbb{E}\tau_n = \sum_{i=0}^n 2^{-i} \leq 2, \text{ para todo } n.$$

Definimos  $\tau_\infty = \sup_n \tau_n$ . Deixamos como exercício provar que a variável  $\tau_\infty$  é finita, isto é

$$\mathbb{P}(\tau_\infty < \infty) = 1.$$

Seja  $N(t)$  o número de saltos feitos até o instante  $t$ , isto é,

$$N(t) = \sup\{n : \tau_n < t\}.$$

Por definição,

$$N(t) > n \text{ se e somente se } \tau_n < t.$$

Portanto, para todo  $t > \tau_\infty$ , o processo faz infinitos saltos antes do instante  $t$ .

**Definição 6.10.** Diremos que um processo  $X_t$  explode se

$$\mathbb{P}(\lim_{n \rightarrow \infty} \tau_n < \infty) > 0.$$

Após o instante aleatório finito  $\tau_\infty$  o processo não está formalmente definido, mas podemos definir um processo explosivo adicionando um novo estado que será chamado  $\infty$  com taxas de transição  $q(\infty, x) = 0$  para todo  $x \in E$ .

Se o processo não explodir, isto é, se

$$\mathbb{P}(\lim_{n \rightarrow \infty} \tau_n < \infty) = 0,$$

então as taxas  $q(x, y)$  definem univocamente o processo que pode ser construído como na proposição acima.

Os processos que estudaremos na sequência não explodem.

### Propriedades.

O resultado que segue diz que o processo  $X_t$  que acabamos de construir é de Markov. Isto em palavras quer dizer que, dado o presente, o futuro e o passado são independentes.

**Proposição 6.11.** *O processo  $(X_t)$  definido em (6.3) satisfaz*

$$\mathbb{P}(X_{t+u} = y \mid X_t = x, X_s = x_s, 0 \leq s < t) = P(X_{t+u} = y \mid X_t = x).$$

**Prova.** De fato, na nossa construção, o processo nos instantes posteriores a  $t$  dependem apenas do processo de Poisson bi-dimensional  $M$  na região  $M((t, \infty) \times \mathbb{R}^+)$  e do valor de  $X_t$  no instante  $t$ . Dado  $X_t$ , isto é independente do que aconteceu antes de  $t$ .  $\square$

Dado que no instante  $\tau_n$  o processo  $X_t$  está no estado  $x$ , o tempo até o próximo salto é uma variável aleatória exponencial de parâmetro  $\mu(x)$  e uma vez que o processo decide saltar, escolhe o estado  $y$  com probabilidade

$$(6.12) \quad p(x, y) = \frac{q(x, y)}{\mu(x)}.$$

Esses resultados são demonstrados nos dois próximos teoremas. Vamos assumir que as taxas de saída são uniformemente limitadas (isto é,  $\sup_x \mu(x) \leq \lambda < \infty$ ).

**Teorema 6.13.** *Para um processo de Markov a tempo contínuo vale*

$$(6.14) \quad \mathbb{P}(\tau_{n+1} - \tau_n > t \mid X(\tau_n) = x) = e^{-t\mu(x)}.$$

**Prova.** Vamos fazer a prova para o caso em que as taxas  $\mu(x)$  estão uniformemente limitadas por  $\lambda \in (0, \infty)$ . O caso geral pode ser provado usando aproximações finitas. Usaremos a representação de  $M(\cdot)$  na faixa  $[0, \infty] \times [0, \lambda]$  do Teorema 5.29.

É claro que o conjunto  $(\tau_n)_n$  está contido no conjunto  $(S_n)_n$  definido no Teorema 5.29. De fato, dado  $x_0 \in E$ , podemos definir

$$(6.15) \quad \begin{aligned} \tau_n &= \min\{S_k > \tau_{n-1} : W_k < \mu(x_{n-1})\} \\ K_n &= \{k : S_k = \tau_n\} \\ x_n &= \{y : W_{K_n} \in I_y^{x_{n-1}}\}. \end{aligned}$$

Essa definição nada mais é que uma consequência da representação do processo de Poisson bidimensional do Teorema 5.29 e da construção do processo de Markov usando o Processo de Poisson bidimensional que culmina na Definição (6.3).

A distribuição de  $\tau_{n+1} - \tau_n$ , condicionada a  $X_{\tau_n} = x$  é dada por

$$\begin{aligned} \mathbb{P}(\tau_{n+1} - \tau_n > t \mid X(\tau_n) = x) \\ = \frac{\mathbb{P}(\tau_{n+1} - \tau_n > t, X(\tau_n) = x)}{\mathbb{P}(X(\tau_n) = x)}. \end{aligned}$$

Condicionando nos possíveis valores de  $K_n$ , o denominador pode ser escrito como

$$\begin{aligned} \sum_k \mathbb{P}(\tau_{n+1} - \tau_n > t, X(\tau_n) = x, K_n = k) \\ = \sum_k \mathbb{P}(\tau_{n+1} - S_k > t, X(S_k) = x, K_n = k) \\ = \sum_k \sum_\ell \mathbb{P}(S_{k+\ell} - S_k > t, W_{k+1} > \mu(x), \dots, W_{k+\ell-1} > \mu(x), \\ W_{k+\ell} < \mu(x)) \mathbb{P}(X(S_k) = x, K_n = k) \\ = e^{-\mu(x)} \sum_k \mathbb{P}(X(S_k) = x, K_n = k). \end{aligned}$$

Assim, pelo exercício 5.40.ii,

$$\frac{\mathbb{P}(\tau_{n+1} - \tau_n > t, X(\tau_n) = x)}{\mathbb{P}(X(\tau_n) = x)} = e^{-\mu(x)},$$

que é o que queríamos demonstrar.  $\square$

O processo discreto  $Y_n = X_{\tau_n}$  é chamado *esqueleto* do processo  $X_t$ .

**Teorema 6.16.** *O esqueleto  $Y_n$  é uma cadeia de Markov com probabilidades de transição  $\{p(x, y), x, y \in E\}$  dadas por (6.12).*

**Prova.** Queremos provar que

$$(6.17) \quad \mathbb{P}(X_{\tau_{n+1}} = y \mid X_{\tau_n} = x) = p(x, y).$$

Vamos usar de novo a construção dada em (6.15). Particionamos de acordo com os possíveis valores de  $K_n$ :

$$(6.18) \quad \begin{aligned} & \mathbb{P}(X_{\tau_{n+1}} = y, X_{\tau_n} = x) \\ &= \sum_k \mathbb{P}(X_{\tau_{n+1}} = y, X_{\tau_n} = x, K_n = k) \end{aligned}$$

Segue da construção que

$$\begin{aligned} & \{X_{\tau_{n+1}} = y, X_{\tau_n} = x, K_n = k\} \\ &= \bigcup_{l \geq 1} \{W_{k+1} > \mu(x), \dots, W_{k+l-1} > \mu(x), W_{k+l} \in I_y^x, \\ & \quad X_{S_k} = x, K_n = k\}, \end{aligned}$$

onde usamos a convenção que para  $l = 1$  o evento  $\{W_{k+1} > \mu(x), \dots, W_{k+l-1} > \mu(x), W_{k+l} \in I_y^x\}$  é apenas  $\{W_{k+1} \in I_y^x\}$ . Pela independência de  $(W_k)$  e  $S_k$ , a expressão (6.18) é igual a

$$\begin{aligned} &= \sum_k \sum_{l \geq 1} \mathbb{P}(W_{k+1} > \mu(x), \dots, W_{k+l-1} > \mu(x), \\ & \quad W_{k+l} \in I_y^x) \mathbb{P}(X_{S_k} = x, K_n = k) \end{aligned}$$

Mas,

$$(6.19) \quad \sum_{l \geq 1} \mathbb{P}(W_{k+1} > \mu(x), \dots, W_{k+l-1} > \mu(x), W_{k+l} \in I_y^x) = \frac{q(x, y)}{\mu(x)}.$$

Portanto (6.18) é igual a

$$\begin{aligned} & \frac{q(x, y)}{\mu(x)} \sum_k \mathbb{P}(X_{S_k} = x, K_n = k) \\ &= \frac{q(x, y)}{\mu(x)} \mathbb{P}(X_{\tau_n} = x), \end{aligned}$$

donde segue a prova do teorema por condicionamento.  $\square$

**Equações de Kolmogorov.**

Para processos de Markov a tempo contínuo é útil usar a seguinte notação matricial. Seja  $Q$  a matriz com entradas

$$q(x, y) \text{ se } x \neq y$$

$$q(x, x) = -\mu(x) = -\sum_{y \neq x} q(x, y).$$

Seja  $P_t$  a matriz que tem como entradas

$$p_t(x, y) = \mathbb{P}(X_t = y \mid X_0 = x).$$

**Proposição 6.20.** *Equações de Chapman-Kolmogorov. Valem as seguintes identidades*

$$(6.21) \quad P_{t+s} = P_t P_s.$$

**Prova.** *Vamos calcular*

$$\begin{aligned} p_{t+s}(x, y) &= \mathbb{P}(X_{t+s} = y \mid X_0 = x) \\ &= \sum_z \mathbb{P}(X_s = z \mid X_0 = x) \mathbb{P}(X_{t+s} = y \mid X_s = z) \\ &= \sum_z p_s(x, z) p_t(z, y). \quad \square \end{aligned}$$

**Proposição 6.22.** *Equações de Kolmogorov. Valem as seguintes identidades*

$$P_t' = Q P_t \quad \text{Equação para trás.}$$

$$P_t' = P_t Q \quad \text{Equação para frente.}$$

**Prova.** Vamos provar a equação para trás. Usando as equações de Chapman-Kolmogorov, temos que para qualquer  $h > 0$ :

$$\begin{aligned} p_{t+h}(x, y) - p_t(x, y) &= \sum_z p_h(x, z) p_t(z, y) - p_t(x, y) \\ &= (p_h(x, x) - 1) p_t(x, y) + \sum_{z \neq x} p_h(x, z) p_t(z, y). \end{aligned}$$

Dividindo por  $h$  e mandando  $h$  a zero, obtemos do lado esquerdo  $p'_t(x, y)$ . Pra calcular o lado direito observe que

$$p_h(x, x) = 1 - \mu(x)h + o(h).$$

Portanto

$$\lim_{h \rightarrow 0} \frac{p_h(x, x) - 1}{h} = -\mu(x) = q(x, x).$$

Analogamente,

$$p_h(x, y) = q(x, y)h + o(h)$$

e

$$\lim_{h \rightarrow 0} \frac{p_h(x, y) - 1}{h} = q(x, y).$$

Isso prova a equação para trás. A equação para frente prova-se em forma análoga e a deixamos como exercício. A diferença é que  $p_{t+h}(x, y)$  deve ser escrita assim:

$$p_{t+h}(x, y) = \sum_z p_t(x, z) + p_h(z, y). \quad \square$$

### Recorrência e transitividade.

**Definição 6.23.** Seja  $T^{x,y} = \inf\{t > \tau_1 : X_t^x = y\}$ , o primeiro instante de chegada em  $y$ . A exigência  $t > \tau_1$  é feita para evitar que  $T^{x,x} = 0$ .

**Definição 6.24.** Diremos que um estado  $x$  é

transitório, se  $\mathbb{P}(T^{x,x} = \infty) > 0$ ;

recorrente nulo, se  $\mathbb{P}(T^{x,x} = \infty) = 0$  e  $\mathbb{E}T^{x,x} = \infty$ ;

recorrente positivo, se  $\mathbb{E}T^{x,x} < \infty$ .

**Definição 6.25.** Diremos que um processo é irredutível se para todo par de estados  $x, y$ , a probabilidade de chegar de um estado a outro em um tempo finito é positiva:

$$P(T^{x,y} < \infty) > 0.$$



**Teorema 6.26.** *Uma processo  $X_t$  é irredutível, se e somente seu esqueleto  $Y_n$  for irredutível.*

*Um estado  $x$  é recorrente (respectivamente transitório) para o processo  $X_t$  se e somente se  $x$  é recorrente (respectivamente transitório) para o esqueleto  $Y_n$ .*

*Se as taxas de saída dos estados estiverem uniformemente limitadas por cima e por baixo, isto é, se*

$$0 < \inf_x \mu(x); \sup_x \mu(x) < \infty,$$

*então  $x$  é recorrente nulo para  $X_t$ , se e somente se  $x$  é recorrente nulo para o esqueleto  $Y_n$ .*

**Prova.** A prova é trabalhosa, porém direta.  $\square$

**Observação 6.27.** É claro que no caso finito as hipóteses do Teorema 6.26 são automaticamente satisfeitas.

A primeira parte do teorema diz que um processo é recorrente (respectivamente transitório) se e somente se o esqueleto dele é recorrente (respectivamente transitório). No entanto, é possível encontrar contra-exemplos no caso infinito (onde as condições do teorema são violadas) para os quais o esqueleto e o processo tem comportamentos qualitativamente diferentes. Já no salto da pulga podemos achar esses contra-exemplos.

**Exemplo 6.28.** Vamos apresentar um processo cujo esqueleto é recorrente positivo e o processo recorrente nulo. Considere as taxas  $q(x, 0) = 1/2^x$ ,  $q(x, x+1) = 1/2^x$ . Assim  $p(x, 0) = 1/2$ ,  $p(x, x+1) = 1/2$  e o esqueleto é recorrente positivo porque o tempo de retorno à origem é dado por uma geométrica de parâmetro  $1/2$ . Por outro lado, como o tempo medio de espera em cada estado  $x$  é  $2^x$ ,

$$\mathbb{E}T^{00} = \sum 2^x 1/2^{x+1} = \infty.$$

**Exemplo 6.29.** Um exemplo (simples porém explosivo) em que o esqueleto do salto da pulga é recorrente nulo e o processo contínuo é

recorrente positivo é o processo com as seguintes taxas:  $q(x, 0) = 1$ ,  $q(x, x+1) = x^2$ . As probabilidades de transição da cadeia de Markov definida pelo esqueleto são  $p(x, 0) = 1/(1+x^2)$  e  $p(x, x+1) = x^2/(1+x^2)$ . O tempo médio para a cadeia discreta voltar para a origem é

$$\sum_x x \left( \frac{1}{1+x^2} \right) \prod_{y=1}^{x-1} \left( \frac{y^2}{1+y^2} \right).$$

Deixamos para o leitor provar que essa soma é infinita. O tempo médio para o processo contínuo voltar à origem é

$$\sum_x \left( \sum_{y=1}^x \frac{1}{y^2} \right) \left( \frac{1}{1+x^2} \right) \left( \prod_{y=1}^{x-1} \left( \frac{y^2}{1+y^2} \right) \right).$$

Deixamos para o leitor provar que essa soma é finita.

### Medidas invariantes.

**Definição 6.30.** Diremos que  $\pi$  é uma medida invariante para  $(X_t)$  se

$$(6.31) \quad \sum_x \pi(x) p_t(x, y) = \pi(y)$$

$$(6.32) \quad \sum_x \pi(x) = 1$$

isto é, se a distribuição do estado inicial é dado por  $\pi$ , então a distribuição do estado do processo no instante  $t$  é também dado por  $\pi$  para qualquer  $t \geq 0$ .

**Teorema 6.33.** Uma medida  $\pi$  é invariante para um processo com taxas  $q(x, y)$  se e somente se

$$(6.34) \quad \sum_x \pi(x) q(x, y) = \pi(y) \sum_z q(y, z).$$

A condição (6.34) pode ser interpretada como que a taxa de entrada sob  $\pi$  no estado  $y$  é a mesma que a taxa de saída de  $y$ . Por isso as equações (6.34) são chamadas de equações de balanço.

**Prova.** Em notação matricial uma medida estacionária pode ser vista como um vetor fila que satisfaz

$$\pi P_t = \pi.$$

Podemos diferenciar esta equação para obter

$$\sum_x \pi(x) p'_t(x, y) = 0.$$

Aplicando as equações de Kolmogorov obtemos (6.34).

Reciprocamente, as equações (6.34) podem ser lidas como

$$\pi Q = 0.$$

Aplicando as equações de Kolmogorov para trás obtemos

$$(\pi P_t)' = \pi P_t' = \pi Q P_t = 0;$$

em outras palavras, começando com  $\pi$  a distribuição do processo é constante e igual a  $\pi$ . Isso porque  $P_0$  é a matriz identidade e  $\pi P_0 = \pi$ .

□

O seguinte é um resultado análogo ao Teorema 3.21. A novidade é que o resultado se aplica mesmo que o esqueleto seja periódico. A conclusão é que processos a tempo contínuo misturam mais que os de tempo discreto.

**Teorema 6.35.** *Um processo  $X_t$  irredutível em um espaço de estados finito tem uma única medida invariante  $\nu$ . Além disso,*

$$\sup_{x,y} |\nu(y) - P_t(x, y)| < e^{-\gamma t},$$

onde

(6.36)

$$\gamma = \min_{x,y} \left( \sum_{z \notin \{x,y\}} \min\{Q(x, z), Q(y, z)\} + Q(x, y) + Q(y, x) \right).$$

**Prova.** Vamos construir um acoplamento de duas cadeias com a mesma matriz das taxas de transição.  $X_t$  com estado inicial arbitrário e  $Y_t$  com estado inicial escolhido com a medida invariante  $\nu$ .

Assuma primeiro que os estados estão ordenados de alguma maneira. Para isso, para cada par de estados distintos  $x < y$ , vamos construir duas famílias de conjuntos disjuntos  $\{I_z^{y|x} : z \in E\}$  e  $\{I_z^{x|y} : z \in E\}$  tais que  $|I_z^{y|x}| = Q(y, z)$  e  $|I_z^{x|y}| = Q(x, z)$ .

Primeiro, começando da origem, construa uma família de intervalos sucessivos  $J_z^{x,y}$  de comprimento

$$|J_z^{x,y}| = \min\{Q(x, z), Q(y, z)\}; \quad z \neq x, y.$$

A continuação coloque um intervalo  $I_y^x$  de comprimento  $Q(x, y)$ , a continuação do qual coloque um intervalo  $I_x^y$  de comprimento  $Q(y, x)$ . A continuação coloque intervalos  $J_z^{x|y}$  de comprimento

$$|J_z^{x|y}| = Q(x, z) - \min\{Q(x, z), Q(y, z)\}; \quad z \neq x, y.$$

A continuação, coloque intervalos  $J_z^{y|x}$  de comprimento

$$|J_z^{y|x}| = Q(y, z) - \min\{Q(x, z), Q(y, z)\}; \quad z \neq x, y.$$

Agora chame, para  $z \neq x, y$

$$I_z^{y|x} = J_z^{y|x} \cup J_z^{x,y}; \quad I_z^{x|y} = J_z^{x|y} \cup J_z^{x,y}$$

Se  $x = y$ , apenas defina intervalos sucessivos  $I_x^x$  de comprimento  $Q(x, z)$ ; isto é,  $I_x^{x|x} = \emptyset$  para todo  $x \in E$ .

Seja

$$I^{x,y} = \left[ \bigcup_{z \neq x,y} (I_z^{y|x} \cup I_z^{x|y}) \right] \cup I_y^x \cup I_x^y.$$

Suponhamos que  $(X_0, Y_0) = (x_0, y_0)$ . Seja  $\tau_1$  o primeiro instante em que aparece uma marca do processo  $M(\cdot)$  no intervalo  $I^{x_0, y_0}$ :

$$\tau_1 = \inf\{t > 0 : M(I^{x_0, y_0} \times [0, t]) > 0\}.$$

Seja

$$x_1 = \begin{cases} z & \text{se } \inf\{t > 0 : M(I^{x_0, y_0} \times [0, t]) > 0\} \\ & = \inf\{t > 0 : M(I_z^{x_0|y_0} \times [0, t]) > 0\} \\ x_0 & \text{em caso contrário.} \end{cases}$$

ou seja  $x_1$  está determinado pelo intervalo  $I_z^{x_0|y_0}$  que realiza o ínfimo ou permanece igual a  $x_0$  se nenhum desses intervalos realiza o ínfimo. Analogamente,

$$y_1 = \begin{cases} z & \text{se } \inf\{t > 0 : M(I^{x_0, y_0} \times [0, t]) > 0\} \\ & = \inf\{t > 0 : M(I_z^{y_0|x_0} \times [0, t]) > 0\} \\ y_0 & \text{em caso contrário.} \end{cases}$$

$y_1$  está determinado pelo intervalo  $I_z^{y_0|x_0}$  que realiza o ínfimo, ou permanece inalterado se nenhum desses intervalos realiza o ínfimo. Seja  $\tau_n$  o primeiro instante em que aparece uma marca do processo  $M(\cdot)$  no intervalo  $I^{x_{n-1}, y_{n-1}}$ :

$$\tau_n = \inf\{t > \tau_{n-1} : M(I^{x_{n-1}, y_{n-1}} \times (\tau_{n-1}, t]) > 0\}.$$

Seja

$$x_n = \begin{cases} z & \text{se } \inf\{t > \tau_{n-1} : M(I^{x_{n-1}, y_{n-1}} \times (\tau_{n-1}, t]) > 0\} \\ & = \inf\{t > \tau_{n-1} : M(I_z^{x_{n-1}|y_{n-1}} \times (\tau_{n-1}, t]) > 0\} \\ x_{n-1} & \text{em caso contrário.} \end{cases}$$

ou seja  $x_n$  está determinado pelo intervalo  $I_z^{y_{n-1}|x_{n-1}}$  que realiza o ínfimo ou permanece igual a  $x_{n-1}$  se nenhum desses intervalos realiza o ínfimo. Analogamente,

$$y_n = \begin{cases} z & \text{se } \inf\{t > \tau_{n-1} : M(I^{x_{n-1}, y_{n-1}} \times (\tau_{n-1}, t]) > 0\} \\ & = \inf\{t > \tau_{n-1} : M(I_z^{y_{n-1}|x_{n-1}} \times (\tau_{n-1}, t]) > 0\} \\ y_{n-1} & \text{em caso contrário.} \end{cases}$$

ou seja  $y_n$  está determinado pelo intervalo  $I_z^{y_{n-1}|x_{n-1}}$  que realiza o ínfimo ou permanece igual a  $y_{n-1}$  se nenhum desses intervalos realiza o ínfimo.

Agora defina

$$(6.37) \quad (X_t, Y_t) = (x_n, y_n) \text{ se } t \in (\tau_n, \tau_{n+1}).$$

Deixamos como exercício para o leitor provar que o processo acima definido tem as marginais corretas.

O fato importante desta construção é que se o processo está no estado  $(x, y)$  e aparece uma marca do processo  $M$  dentro do intervalo

$$(U_z(J_z^{x,y}) \cup I_x^y \cup I_y^x),$$

então os dois processos seguem juntos para sempre. O comprimento desse intervalo é justamente

$$\sum_{z \notin \{x,y\}} \min\{Q(x,z), Q(y,z)\} + Q(x,y) + Q(y,x).$$

Como esses intervalos tem o extremo inferior na origem, os processos vão se juntar no primeiro instante em que uma marca do processo  $M(\cdot)$  aparecer na intersecção desses intervalos. Essa intersecção tem comprimento  $\gamma$  dado por (6.36).  $\square$

Um teorema mais geral, porém sem a velocidade de convergência é o seguinte

**Teorema 6.38.** *Se  $X_t$  é recorrente positiva e irredutível então o processo admite uma única medida invariante  $\pi$ . Além disso, para qualquer estado  $x$  inicial vale que*

$$\lim_{t \rightarrow \infty} p_t(x, y) = \pi(y).$$

**Prova.** omitida.  $\square$

**Relação entre a medida invariante do processo e seu esqueleto.** Assuma que o processo  $X_t$  e o esqueleto discreto  $Y_n = X_{\tau_n}$  são irredutíveis e recorrentes positivos. Vimos no capítulo 4 que uma medida invariante para o esqueleto discreto  $Y_n$  deve satisfazer o sistema de equações

$$\sum_x m(x)p(x, y) = m(y)$$

Por outro lado, a medida invariante  $\pi$  para o processo  $X_t$  deve satisfazer

$$\sum_x \pi(x)q(x, y) = \pi(y)\mu(y)$$

Isto implica que  $m$  é a medida invariante para  $Y_n$  se e somente se  $\pi$  definida por

$$(6.39) \quad \pi(x) = m(x)/\mu(x)$$

é invariante para  $X_t$ . Intuitivamente, os tempos de espera diferentes para distintos estados no processo contínuo requerem uma correção na medida invariante do processo discreto que leve em conta essas diferenças.

Como corolário temos que se os tempos de espera nos diferentes pontos não dependem do ponto, então uma medida é invariante para o processo se e somente se for invariante para o esqueleto. Isto é, se  $\mu(x) = \mu(y)$  para todo par  $x, y$ , então  $\pi(x) = m(x)$  para todo  $x$ .

### Processo de nascimento e morte.

O processo de nascimento e morte representa o crescimento (e decréscimo) de uma população. O valor  $X_t$  representa o número de indivíduos vivos no instante  $t$ . As taxas de crescimento e decréscimo dependem apenas do número de indivíduos vivos. Assim

$$(6.40) \quad q(x, x+1) = \lambda_x \quad \text{e} \quad q(x, x-1) = \mu_x,$$

onde  $\lambda_x, \mu_x$  são famílias de parâmetros não negativos. Vamos usar as equações de balanço (6.34) para ver sob quais condições o processo admite uma medida invariante. Estamos procurando um vetor  $\pi$  que satisfaça as equações de balanço

$$(6.41) \quad \begin{aligned} \pi(0)q(0, 1) &= \pi(1)q(1, 0) \\ \pi(x)q(x, x-1) + q(x, x+1) &= \pi(x-1)q(x-1, x) + \pi(x+1)q(x+1, x), \end{aligned}$$

para  $x \geq 1$ .

Dai não é difícil concluir que para todo  $x \geq 0$ ,

$$(6.42) \quad \pi(x+1)\mu_{x+1} - \pi(x)\lambda_x = 0,$$

donde

$$\pi(x+1) = \frac{\lambda_x}{\mu_{x+1}}\pi(x), \quad x \geq 0.$$

Portanto, para todo  $x \geq 1$

$$(6.43) \quad \pi(x) = \frac{\lambda_0 \dots \lambda_{x-1}}{\mu_1 \dots \mu_x} \pi(0).$$

É claro que  $\pi(x)$  assim construído satisfaz (6.34). Agora devemos fixar  $\pi(0)$  de tal forma que (6.32) seja satisfeita com  $\pi(0) > 0$  pois com  $\pi(0) = 0$  (6.32) não pode ser satisfeita. Para isso precisamos

$$(6.44) \quad \sum_{x \geq 1} \frac{\lambda_0 \dots \lambda_{x-1}}{\mu_1 \dots \mu_x} < \infty.$$

Se (6.44) é satisfeita, podemos definir

$$\pi(0) = \left( \frac{\lambda_0 \dots \lambda_{x-1}}{\mu_1 \dots \mu_x} \right),$$

e  $\pi(x)$  como em (6.43).  $\square$

### Exercícios.

**6.45.** Seja  $X_t$  um processo a tempo contínuo no espaço de estados  $\{0, 1\}$  e com taxas  $q(0, 1) = 1$ ,  $q(1, 0) = 2$ . Calcule as equações de Kolmogorov e ache a probabilidade  $p_t(1, 0)$ .

**6.46.** Demonstre que a variável  $\tau_\infty$  definida no Exemplo 6.5 é finita com probabilidade 1. Sugestão: observe que, pela desigualdade de Markov

$$\mathbb{P}(\tau_\infty > K) \leq \frac{\mathbb{E}\tau_\infty}{K}.$$



Em seguida use o fato que

$$\lim_n \mathbb{P}(\tau_n > K) = \mathbb{P}(\tau_\infty > K).$$

**6.47.** Prove que o processo de nascimento puro construído em (6.9) satisfaz as condições (6.7) e (6.8).

**6.48.** Acoplamento de processos de nascimento puro. Prove que é possível acoplar dois processos de nascimento puro  $(X_t^1)$  e  $(X_t^2)$  com taxas  $\mu_1 \geq \mu_2$ , respectivamente, de tal maneira que se  $X_0^1 \geq X_0^2$ , então  $X_t^1 \geq X_t^2$ .

**6.49.** Prove a identidade (6.19).

**6.50.** Prove que o processo apresentado no Exemplo 6.29 é recorrente positivo e seu esqueleto é recorrente nulo.

**6.51.** Prove que as marginais do processo conjunto definido em (6.37) tem a distribuição do processo. Isto é, prove que

$$\begin{aligned} \mathbb{P}(X_{t+h} = y \mid X_t = x) &= q(x, y)h + o(h) \\ \mathbb{P}(Y_{t+h} = y \mid Y_t = x) &= q(x, y)h + o(h). \end{aligned}$$

**6.52.** Seja  $X_t$  o processo em  $E = \{1, 2, 3\}$  com matriz  $Q$  de taxas de transição

$$Q = \begin{pmatrix} -4 & 1 & 3 \\ 2 & -3 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

(A diagonal é apenas a soma das taxas de saída do estado correspondente trocada de sinal.) Construa o acoplamento  $(X_t, Y_t)$  dado por (6.37). Quanto vale  $\gamma$ ?

**6.53.** Para o processo do Exercício 6.52 calcule as probabilidades de transição do esqueleto, a medida invariante para o processo contínuo e para o esqueleto e verifique (6.39).

**6.54.** Considere uma fila onde os clientes chegam em grupos de 2 a taxa  $\lambda$ . O serviço é feito de a um por vez a taxa  $\mu$ . Suponha que o sistema tem uma capacidade máxima de 4 clientes. Ou seja que se em algum momento tiver 3 clientes e chegar um grupo de 2, apenas um deles fica no sistema e o outro vai embora.

- (a) Estabeleça as equações de balanço e ache a medida invariante.
- (b) Qual a probabilidade de um grupo chegar no sistema e nenhum dos dois conseguir ficar nele?
- (c) Qual é o número medio de clientes no sistema?
- (d) Qual o tempo medio que um cliente fica no sistema?

**6.55.** Considere uma fila  $M/M/\infty$ , isto é, as chegadas acontecem de acordo a uma taxa  $\lambda$  e os serviços de acordo a uma taxa  $\mu$  mas agora o sistema tem infinitos servidores. Resolva os itens do exercício 6.54 nesse caso.

**6.56.** Considere uma população de  $m$  indivíduos que no instante zero tem  $k$  indivíduos “infectados” e  $m - k$  sadios. Um indivíduo infectado fica sadio após um tempo exponencial de parâmetro  $\mu$ . Se tiver  $k$  indivíduos infectados, a taxa de cada um dos  $m - k$  indivíduos sadios restantes ficar infectado é  $\lambda(k + 1)$ .

- (a) Estabeleça as equações de balanço.
- (b) Para  $m = 5$ ,  $\lambda = 1$  e  $\mu = 2$  calcule a medida invariante.
- (c) Qual o número médio de indivíduos infectados?
- (d) Qual a probabilidade que em um instante dado todo mundo esteja infectado?

**Comentários e referências.** Neste capítulo susamos a construção do processo de Markov contínuo usando um processo de Poisson bidimensional homogêneo. Se trata de uma extensão natural do método da projeção introduzido no capítulo anterior. O acoplamento usando esse método é análogo ao acoplamento usado nas cadeias discretas.

## 7. REVERSIBILIDADE E O TEOREMA DE BURKE

Processos de Markov reversíveis.

Conceitos fundamentais em processos estocásticos são o de reversibilidade e o de reversibilidade dinâmica.

Lembremos de notação introduzida no capítulo anterior.  $X_t$  representa um processo de Markov a tempo contínuo com taxas dadas por

$$q(x, y) = \lim_{h \rightarrow 0} \frac{\mathbb{P}(X(t+h) = y \mid X_t = x)}{h}.$$

Lembremos da convenção

$$\mu(x) = \sum_{y \neq x} q(x, y).$$

Usamos também a notação

$$p_t(x, y) = \mathbb{P}(X_t = y \mid X_0 = x).$$

**Definição 7.1.** Diremos que uma função  $\pi : E \rightarrow [0, 1]$  é uma medida de probabilidade reversível para o processo  $X_t$  se para todo  $x, y \in E$ :

$$(7.2) \quad \pi(x)q(x, y) = \pi(y)q(y, x) \text{ e}$$

$$(7.3) \quad \sum_x \pi(x) = 1.$$

Em palavras, se o processo  $X_t$  tem como distribuição inicial uma medida  $\pi$  reversível, então a probabilidade de observar um salto na direção  $x \rightarrow y$  em um intervalo pequeno de tempo é a mesma que a de observar o salto  $y \rightarrow x$ . As equações (7.2) são chamadas de equações de balanço detalhado.

**Proposição 7.4.** Se  $\pi$  é reversível para  $X_t$  então  $\pi$  é invariante para  $X_t$ .

**Prova.** Basta somar sobre  $x$  em (7.2) para obter

$$\sum_x \pi(x)q(x, y) = \pi(y) \sum_z q(y, z),$$

que é a condição para uma medida ser invariante.  $\square$

Ou seja que medidas reversíveis são invariantes.

**Exemplo 7.5.** O caso mais simples de processo reversível é obtido com espaço de estados com apenas dois elementos: qualquer processo em um espaço com dois estados é reversível. Se os estados são chamados 0 e 1, a medida reversível é dada por

$$\pi(0) = \frac{q(1, 0)}{q(1, 0) + q(0, 1)}, \quad \pi(1) = \frac{q(0, 1)}{q(1, 0) + q(0, 1)}.$$

Se  $q(0, 1) = q(1, 0) = 0$ , então qualquer medida de probabilidade em  $\{0, 1\}$  é reversível.

**Exemplo 7.6.** Um processo em um espaço de estados finito que sempre aceita medidas reversíveis é aquele que tem taxas simétricas:

$$(7.7) \quad q(x, y) = q(y, x),$$

ou seja, aquele em que a taxa de passar de um estado para outro é a mesma que a de fazer o salto no outro sentido. A medida reversível está dada por

$$(7.8) \quad \pi(x) = \frac{\sum_{y: y \neq x} q(x, y)}{\sum_z \sum_{y: y \neq z} q(z, y)} = \frac{\mu(x)}{\sum_y \mu(y)}.$$

Um processo reversível tem a mesma distribuição que o processo quando olhado para trás no tempo. Este é o conteúdo do nosso próximo resultado.

**Proposição 7.9.** *Um processo de Markov  $(X_t)$  aceita  $\pi$  como medida reversível se e somente se  $\pi$  é uma probabilidade e para todo  $t \geq 0$*

$$(7.10) \quad \pi(x)p_t(x, y) = \pi(y)p_t(y, x).$$

**Prova.** É fácil ver que, calculando a derivada de (7.10) obtemos (7.2). Para provar a outra implicação, vamos assumir que as taxas

$q(x, y)$  são uniformemente limitadas por cima: existe um número real (positivo)  $\gamma$  tal que

$$\sup_x \mu(x) \leq \gamma < \infty.$$

Nesse caso podemos escrever

$$q(x, y) = \gamma P(x, y),$$

onde  $P$  é uma matriz de transição de probabilidade para uma cadeia de Markov dada por

$$P(x, y) = \begin{cases} \frac{q(x, y)}{\gamma} & \text{se } y \neq x \\ 1 - \frac{\mu(x)}{\gamma} & \text{se } y = x. \end{cases}$$

Com esta representação podemos realizar o processo da seguinte maneira: Em instantes de tempo indicados por um processo de Poisson de parâmetro  $\gamma$ , o processo salta do estado  $x$  ao estado  $y$ , de acordo com a matriz de transição  $P$ . Mais formalmente

$$X_t = Y_{N(t)},$$

onde  $Y_n$  é uma cadeia de Markov a tempo discreto com matriz de transição  $P$  e  $N(t)$  é um processo de Poisson de parâmetro  $\gamma$ . Além disso  $Y_n$  e  $N(t)$  são independentes. Os tempos de espera entre saltos de  $X_t$  dependem do valor de  $X_t$ . Em contraste, para  $Y_n$  os tempos de espera entre saltos são identicamente iguais a 1. Para compensar essa diferença,  $Y_n$  realiza saltos nulos, o que se reflete no fato de  $P(x, x)$  não ser necessariamente igual a zero.

Seja

$$m(x) = \pi(x)\mu(x).$$

É fácil ver que  $\pi$  é reversível para  $X_t$  se e somente se  $\pi$  é reversível para  $Y_n$ . Isto é, para  $x \neq y$ , as seguintes identidades são equivalentes

$$\begin{aligned} \pi(x)q(x, y) &= \pi(y)q(y, x) \\ \pi(x)P(x, y) &= \pi(y)P(y, x). \end{aligned}$$

Por outro lado é imediato que para qualquer sequencia finita de estados  $x_1, \dots, x_n$ :

$$\pi(x)P(x, x_1) \dots P(x_n, y) = \pi(y)P(y, x_n) \dots P(x_1, x).$$

Somando sobre todos os possíveis valores de  $x_1, \dots, x_n$ , obtemos

$$\pi(x)P^n(x, y) = \pi(y)P^n(y, x).$$

E como

$$(7.11) \quad p_t(x, y) = \sum_{n \geq 0} \mathbb{P}(N(t) = n)P^n(x, y),$$

segue que

$$\begin{aligned} \pi(x)p_t(x, y) &= \sum_{n \geq 0} \mathbb{P}(N(t) = n)\pi(x)P^n(x, y) \\ &= \sum_{n \geq 0} \mathbb{P}(N(t) = n)\pi(y)P^n(y, x) = \pi(y)p_t(y, x), \end{aligned}$$

o que termina a demonstração.  $\square$

Uma trajetória da cadeia  $X_t$  é caracterizada por dois processos independentes: os instantes de salto  $T_i$ , que são os instantes de ocorrência do processo de Poisson  $N(t)$ , e os saltos propriamente ditos, que estão caracterizados por  $Y_n$ . Sob uma medida reversível, uma trajetória tem a mesma lei que a trajetória vista de trás para frente. Uma maneira de formalizar esse resultado é o seguinte:

**Proposição 7.12.** *Sejam  $0 = t_1 < \dots < t_n < t$  e  $x_1, \dots, x_n$  estados possíveis da cadeia. Seja  $s_i = t - t_{n-i}$ .*

$$\mathbb{P}_\pi(X_{t_1} = x_1, \dots, X_{t_n} = x_n) = \mathbb{P}_\pi(X_{s_1} = x_n, \dots, X_{s_n} = x_1),$$

onde  $\mathbb{P}_\pi$  representa a distribuição do processo sob amedida estacionária.

**Prova.** Pelo lema anterior, para uma cadeia reversível vale

$$\pi(x)p_t(x, y) = \pi(y)p_t(y, x).$$

Por outro lado, pela propriedade markoviana:

$$\begin{aligned}
 & \mathbb{P}_\pi(X_{t_1} = x_1, \dots, X_{t_n} = x_n) \\
 &= \sum_x \pi(x) \sum_y p_{t_1}(x, x_1) \dots p_{t_n - t_{n-1}}(x_{n-1}, x_n) p_{t-t_n}(x_n, y) \\
 &= \sum_x \sum_y p_{s_n - s_{n-1}}(x_1, x) \dots p_{s_1 - s_0}(x_n, x_{n-1}) p_{s_0}(y, x_n) \pi(y) \\
 &= \mathbb{P}_\pi(X_{s_1} = x_n, \dots, X_{s_n} = x_1). \quad \square
 \end{aligned}$$

**Exemplo 7.13.** O exemplo mais simples de processo não reversível é um processo em um círculo que se movimenta somente na direção horária. Invertendo o tempo veremos que o processo se movimenta na direção anti-horária. Vejamos um exemplo com 3 estados. Seja  $q(0,1) = q(1,2) = q(2,0) = 1$ . Há uma única medida invariante porque o sistema é irredutível e ergódico em um espaço de estados finito. Portanto, da Proposição 7.4 segue que a única candidata a medida reversível é a medida invariante. A medida invariante pode facilmente ser calculada colocando as equações, mas nesse caso dado a simetria entre os tres valores podemos deduzir que  $\pi(0) = \pi(1) = \pi(2) = 1/3$  é invariante. A demonstração desse fato é deixada como exercício. Por outro lado,  $\pi$  não é reversível porque –por exemplo–

$$1/3 = \pi(0)q(0,1) \neq \pi(1)q(1,0) = 0.$$

Em geral, se a transição de um estado  $x$  a um estado  $y$  tem taxa positiva mas a transição de  $y$  a  $x$  tem taxa nula, então o processo não pode admitir uma medida reversível.

Vamos agora construir em forma conjunta o processo  $X_t$  e o processo obtido revertendo o tempo. Provaremos que ao fazer isso obtemos novamente um processo de Markov que será chamado de processo reverso  $X_t^*$ . Vamos assumir que o processo  $X_t$  está definido para todo  $t \in \mathbb{R}$  e que está em equilíbrio, isto é  $\mathbb{P}(X_t = x) = \pi(x)$  para todo  $t \in \mathbb{R}$  e que vem acontecendo desde o tempo  $-\infty$ . Definimos o processo reverso  $X_t^*$  por

$$(7.14) \quad X_t^* = \lim_{s \uparrow t} X_{-s}.$$

O processo  $X_t^*$  tem saltos nos mesmos instantes que o processo  $X_t$ . A definição usando o limite a esquerda é feita para garantir que as trajetórias sejam contínuas à direita.

**Proposição 7.15.** *Seja  $X_t$  um processo que tem  $\pi$  como medida invariante. Seja  $X_t^*$  o processo reverso em relação a  $\pi$  construído como indicado acima. Então,*

(1)  $\pi$  é invariante para  $X_t^*$ .

(2) O processo reverso  $X_t^*$  tem a mesma distribuição que o processo  $X_t$  em equilíbrio se e somente se  $\pi$  for reversível para  $(X_t)$ .

**Prova.** É deixada como exercício.  $\square$

### A Fila $(M|M|1)$ e o teorema de Burke.

A fila  $(M|M|1)$  é um processo de Markov  $(X_t)$  tendo  $\mathbb{N}$  como espaço de estados com taxas

$$q(x, y) = \begin{cases} \lambda & \text{se } y = x + 1 \\ \mu & \text{se } x \geq 1 \text{ e } y = x - 1 \\ 0 & \text{em caso contrário.} \end{cases}$$

O processo  $X_t$  indica o número de clientes que esperam numa fila para serem atendidos num guichê. O atendimento faz-se pela ordem de chegadas. Só se atende um único cliente por vez. Sempre que um cliente é atendido, o próximo cliente, se houver, tem seu atendimento iniciado imediatamente. As chegadas de clientes ocorrem de acordo a um processo de Poisson de parâmetro  $\lambda$  e o tempo de serviço é exponencial de taxa  $\mu$ . O primeiro  $M$  quer dizer chegadas markovianas, referindo-se ao fato que o processo de Poisson, como os processos de Markov não tem memória, e o segundo  $M$  quer dizer serviços markovianos, que no caso são representados pela distribuição exponencial. O "1" refere-se ao fato de ser apenas um guichê de atendimento.

A medida invariante para esse processo pode ser facilmente calculada, resolvendo-se as equações de balanço. Uma alternativa seria tentar a sorte e se perguntar se a cadeia aceita uma medida reversível.



Nesse caso as equações a resolver são mais simples. Procuramos números  $\pi(x)$  tal que para  $x \geq 0$

$$\pi(x)\lambda = \pi(x+1)\mu, \quad \sum_x \pi(x) = 1$$

É fácil ver que a solução é a distribuição geométrica:

$$(7.16) \quad \pi(x) = \left(\frac{\lambda}{\mu}\right)^x \left(1 - \frac{\lambda}{\mu}\right)$$

e que a segunda condição (que a soma total das probabilidades seja 1) é satisfeita se e somente se  $\lambda < \mu$  que será assumido daqui para frente. De fato, a fila  $M|M|1$  é um caso particular do processo de nascimento e morte introduzido em (6.40) com  $\lambda_i \equiv \lambda$  e  $\mu_i \equiv \mu$ . A condição (6.44) lê-se aqui como  $\lambda < \mu$ .

Como a medida  $\pi$  definida em (7.16) é reversível, se olharmos a fila em equilíbrio para trás no tempo, vemos de novo uma fila com chegadas markovianas de parâmetro  $\lambda$  e serviços exponenciais de parâmetro  $\mu$ . Vamos assumir que o processo  $X_t$  está definido para todo valor de  $t$  e que está em equilíbrio, isto é, para qualquer  $t$

$$\mathbb{P}(X_t = x) = \pi(x).$$

Vamos agora supor que não nos é permitido observar o processo de Poisson das chegadas, mas apenas os valores  $X_t$ ,  $t \in \mathbb{R}$ . Definimos então o processo de chegadas como uma função do processo  $\{X_t, t \in \mathbb{R}\}$  por

$$A_t - A_s = \sum_{u \in [s, t]} \mathbf{1}\{X_u - X_{u-} = 1\} = \sum_{n: T_n \in [s, t]} (Y_n - Y_{n-1})^+,$$

ou seja, o número de vezes que no intervalo  $[s, t]$  o processo  $X_u$  aumentou de valor. A cadeia  $Y_n$  é o esqueleto discreto introduzido no capítulo anterior. Note que a primeira soma, apesar de estar definida em princípio sobre um conjunto não enumerável, é na verdade uma

soma finita com probabilidade um. Da mesma maneira definimos o processo de partidas,  $D_t$  por

$$D_t - D_s = \sum_{u \in [s, t]} 1\{X_u - X_{u-} = -1\} = \sum_{n: T_n \in [s, t]} (Y_n - Y_{n-1})^-,$$

ou seja, o número de vezes que no intervalo  $[s, t]$  o processo diminuiu de valor.

Por exemplo, se no intervalo  $[s, t]$ ,  $X_u$  assumiu sucessivamente os valores

$$4, 3, 2, 3, 2, 3, 4, 3, 4, 5, 6, 5, 6, 5, 4, 3,$$

teremos que

$$A_t - A_s = 6,$$

porque 6 vezes (23, 23, 34, 34, 56, 56) a sequência de números cresceu.

Similarmente,

$$D_t - D_s = 8,$$

correspondente aos pares 43, 32, 32, 43, 65, 65, 54, 43.

**Teorema de Burke.** *Se a distribuição inicial de  $X_t$  é a distribuição estacionária dada por (7.16), então  $D_t$  é um processo de Poisson de parâmetro  $\lambda$ .*

**Prova.** Poderia se fazer uma conta explícita, mas vamos dar uma prova curta usando o processo reverso  $X_t^*$  definido acima. Defina o processo  $A_t^*$  de chegadas do reverso por

$$A_t^* - A_s^* = \sum_{u \in [s, t]} 1\{X_u^* - X_{u-}^* = 1\}.$$

Analogamente o processo de partidas  $D_t^*$  do reverso por

$$(7.17) \quad D_t^* - D_s^* = \sum_{u \in [s, t]} 1\{X_u^* - X_{u-}^* = -1\}.$$

Da definição do reverso (7.14) e de (7.17) segue que

$$A_t - A_s = D_{-s}^* - D_{-t}^* \quad \text{e} \quad D_t - D_s = A_{-s}^* - A_{-t}^*.$$

O teorema fica então demonstrado, porque o processo  $A_t^*$  é de Poisson.  $\square$

**Observação 7.18.** Note primeiro que o processo de chegadas  $A_t$  é um processo de Poisson de parâmetro  $\lambda$ , porque as chegadas não dependem do número de clientes na fila e ocorrem com taxa  $\lambda$ .

Por outro lado, as saídas acontecem com taxa  $\mu$ , quando tem alguém na fila, e com taxa zero quando não tem ninguém. O Teorema de Burke diz que se produz uma extraordinária compensação e que, se o sistema estiver em equilíbrio, e alguém observar apenas as pessoas que saem do sistema, verá que estas fazem isto de acordo com um processo de Poisson de taxa  $\lambda$ .

### Reversibilidade dinâmica.

Vamos supor que um processo de Markov  $X_t$  com taxas  $q(x, y)$  admita uma medida invariante  $\pi$ . Podemos então construir o processo para todo tempo  $t \in \mathbb{R}$  de maneira tal que

$$\mathbb{P}(X_t = x) = \pi(x).$$

Definimos o processo reverso como em (7.14). Mesmo se  $\pi$  não for reversível, o processo  $X_t^*$  assim definido será de Markov. Esse é o resultado do seguinte lema.

**Lema 7.19.** *Assuma que  $\pi$  é invariante para  $X_t$  e que  $X_t^*$  é definido por (7.14). Então  $X_t^*$  é um processo de Markov com taxas*

$$(7.20) \quad q^*(x, y) = \frac{q(y, x)\pi(y)}{\pi(x)}.$$

*Mais precisamente, usando a notação da Proposição 7.12:*

$$\mathbb{P}_\pi(X_{t_1} = x_1, \dots, X_{t_n} = x_n) = \mathbb{P}_\pi(X_{s_1}^* = x_n, \dots, X_{s_n}^* = x_1).$$

**Prova.** É similar à prova da Proposição 7.9. Seja

$$\mu^*(x) = \sum_{y \neq x} q^*(x, y).$$

Então, o leitor provará nos exercícios que

$$\mu(x) = \mu^*(x).$$

Defina a matriz de transição do esqueleto discreto dos processos direto e reverso da seguinte maneira

$$P(x, y) = \frac{q(x, y)}{\mu(x)}$$

$$P^*(x, y) = \frac{q^*(x, y)}{\mu^*(x)}.$$

Como

$$\sum_z q(x, z) = \sum_z q^*(x, z),$$

a definição (7.20) de  $q^*$  implica que

$$\pi(x)P(x, y) = \pi(y)P^*(y, x).$$

Portanto,

$$\pi(x)P^n(x, y) = \pi(y)(P^*)^n(y, x).$$

O restante da prova segue como no Lema 7.9.  $\square$

Um problema típico é o recíproco do lema acima: suponha que temos um processo de Markov  $X_t$  com taxas  $q(x, y)$  e queremos achar a medida invariante  $\pi$ . Classicamente devem ser resolvidas as equações de balanço. Mas isso pode ser difícil. Uma forma alternativa que muitas vezes funciona é adivinhar qual vai ser o processo reverso  $q^*(x, y)$  e a partir daí procurar a solução  $\pi$  das equações

$$(7.21) \quad \pi(x)q(x, y) = \pi(y)q^*(y, x).$$

O lema que segue diz que isso funcionará somente se para todo  $x$ ,

$$(7.22) \quad \mu(x) = \mu^*(x),$$

ou seja, se a taxa de saída do estado  $x$  para o processo direto é igual à taxa de saída de  $x$  para o processo reverso.

**Lema 7.23.** *Seja  $X_t$  um processo de Markov com taxas  $q(x, y)$ . Assuma que existem taxas  $q^*(x, y)$  e uma medida de probabilidade  $\pi$  tal que (7.21) e (7.22) são satisfeitas. Prove que  $\pi$  é invariante para  $X_t$  e que o processo com taxas  $q^*(x, y)$  é o reverso de  $X_t$  com relação a  $\pi$ .*

**Prova.** É deixada como exercício.  $\square$

**Exemplo 7.24.** Continuação do Exemplo 7.13 para um processo em um círculo com três estados. Como o processo para frente consiste em um indivíduo que depois de um tempo exponencial salta uma unidade no sentido horário, não é absurdo imaginar que o processo reverso no tempo seja um indivíduo que salta uma unidade em sentido antihorário. Assim definimos

$$q^*(1, 0) = q(2, 1) = q(0, 2) = 1; \quad q(x, y) = 0 \quad \text{em caso contrário.}$$

Já conhecemos a medida invariante  $\pi$  porque resolvemos as equações de balanço em um exercício. Mas vamos fazer de conta que não a conhecemos e que procuramos soluções  $\pi$  das equações

$$\pi(x)q(x, y) = \pi(y)q^*(y, x),$$

que no exemplo ficam:

$$\pi(0)q(0, 1) = \pi(1)q^*(1, 0), \quad \text{etc.}$$

mas como  $q(0, 1) = q^*(1, 0) = 1$  vemos que a solução deve satisfazer  $\pi(0) = \pi(1)$ . Analogamente obtemos  $\pi(2) = \pi(1)$ . Isso nos dá como solução a medida uniforme em  $\{0, 1, 2\}$ . Para provar que a medida é invariante e o processo assim definido é o processo reverso, é necessário verificar que para cada estado, a taxa de saída do processo reverso é a mesma que a do processo direto. Isto é, devemos provar que

$$\sum_y q^*(x, y) = \sum_y q(x, y).$$

Isso é imediato porque cada um dos lados dessa identidade é igual a 1 para cada um dos  $x \in \{0, 1, 2\}$ .

### Filas em série.

Vamos definir um sistema de duas filas em série. Com taxa  $\lambda$  os clientes chegam à fila 1, onde são atendidos com taxa  $\mu$ . Cada cliente que termina um serviço na fila 1, vai para a fila 2, onde de novo é atendido com taxa  $\mu$ . O espaço de estados é  $\{(x, y) : x, y \in \mathbb{N}\}$ , onde  $x$  é o número de clientes na primeira fila e  $y$  o número de clientes na segunda fila. As taxas são

$$(7.25) \quad q((x, y), (z, w)) = \begin{cases} \lambda & \text{se } z = x + 1 \text{ e } w = y \\ \mu & \text{se } x \geq 1, z = x - 1 \text{ e } w = y + 1 \\ \mu & \text{se } z = x, y \geq 1 \text{ e } w = y - 1 \\ 0 & \text{em caso contrário.} \end{cases}$$

Esse processo não pode ser reversível. Intuitivamente porque os clientes vão da fila 1 para a fila 2. Portanto, revertendo o tempo, os clientes irão da fila 2 para a fila 1. Formalmente porque –por exemplo–

$$q((2, 3), (1, 4)) = \mu \quad \text{e} \quad q((1, 4), (2, 3)) = 0.$$

Portanto não existe uma medida (não trivial) tal que

$$\pi(2, 3)q((2, 3), (1, 4)) = \pi(1, 4)q((1, 4), (2, 3)).$$

Para achar a medida invariante  $\pi$  vamos tentar adivinhar qual é o processo reverso  $X_t^*$ . O mais à mão é o processo que tem as seguintes propriedades:

- (i) chegadas de Poisson com taxa  $\lambda$  na fila 2;
- (ii) na fila 2 os clientes são atendidos a taxa  $\mu$ ;
- (iii) os clientes atendidos na fila 2 vão para a fila 1, onde são atendidos com taxa  $\mu$  após o qual saem do sistema.

As taxas para esse sistema de duas filas são

$$(7.26) \quad q^*((x, y), (z, w)) = \begin{cases} \lambda & \text{se } w = y + 1 \text{ e } z = x \\ \mu & \text{se } y \geq 1, w = y - 1 \text{ e } z = x + 1 \\ \mu & \text{se } w = y, x \geq 1 \text{ e } z = x - 1 \\ 0 & \text{em caso contrário.} \end{cases}$$

Não é difícil ver que para esse sistema

$$(7.27) \quad \mu(x, y) = \mu^*(x, y),$$

para todo estado  $(x, y)$ . Portanto esse é um candidato possível para ser o processo reverso.

Estamos procurando uma medida de probabilidade  $\pi(x, y)$ , satisfazendo as equações (7.21), que no nosso contexto se escrevem como:

$$(7.28) \quad \begin{aligned} \pi(x, y)\lambda &= \pi(x + 1, y)\mu \\ \pi(x + 1, y)\mu &= \pi(x, y + 1)\mu \\ \pi(x, y + 1)\mu &= \pi(x, y)\lambda. \end{aligned}$$

para todo  $x, y \geq 0$ . Não é difícil verificar que

$$(7.29) \quad \pi(x, y) = C \left( \frac{\lambda}{\mu} \right)^{x+y}$$

é solução do sistema (7.28), onde a constante de normalização  $C$  é dada por

$$C = \left( 1 - \frac{\lambda}{\mu} \right)^2.$$

Assim provamos que a medida invariante para o sistema com duas filas é o produto de duas geométricas de parâmetro  $\lambda/\mu$  e que o processo é dinamicamente reversível. O processo reverso é também um sistema de filas, mas os clientes vão na direção oposta aos do sistema direto. Sob a medida invariante  $\pi$ , o número de clientes na fila 1 é independente do número de clientes na fila 2. Isso já é bastante surpreendente. Agora vamos usar a reversibilidade dinâmica para provar o teorema de Burke para filas em série.

**Teorema de Burke 2.** *Em um sistema de filas em série com taxas de transição (7.25) e começando com a medida invariante (7.28), o processo  $D_t$  de saídas da fila 2 é um processo de Poisson de parâmetro  $\lambda$ .*

**Prova.** A ideia é a mesma que a da prova do teorema de Burke para uma fila. A única diferença é que no caso de uma fila o processo

reverso tinha a mesma distribuição que o processo direto. Aqui o processo reverso é diferente, mas a demonstração funciona da mesma maneira.

Seja  $A_t - A_s$  o número de clientes que entram na fila 1 no processo direto no intervalo  $[s, t]$ . Seja  $D_t - D_s$  o número de clientes que saem da fila 2 no mesmo intervalo. Seja  $A_t^* - A_s^*$  o número de clientes que entram na fila 2 no processo reverso e  $D_t^* - D_s^*$  o número de clientes que saem da fila 1 no mesmo intervalo de tempo.

Então  $A_t$  e  $A_t^*$  são processos de Poisson por definição dos processos e como

$$D_t - D_s = A_{-s}^* - A_{-t}^*,$$

o teorema fica provado.  $\square$

### Exercícios.

**7.30.** Prove que a medida dada por (7.29) é solução do sistema de equações (7.28).

**7.31.** Prove que a medida uniforme em  $\{1, 2, 3\}$  é invariante para o processo no círculo do exemplo 7.13.

**7.32.** Prove que se  $X_t$  é um processo em um espaço de estados unidimensional com saltos aos vizinhos mais próximos então as medidas invariantes para  $X_t$  são medidas reversíveis. Mais precisamente, Seja  $X_t \in \{0, 1, 2, \dots, N\}$  e  $q(x, y) = 0$  se  $|x - y| \neq 1$ . Prove que se  $\pi$  é invariante para  $X_t$ , então  $\pi$  é reversível.

**7.33.** Prove que a medida  $\pi$  definida na equação (7.8) é invariante para o processo com taxas simétricas do exemplo 7.6.

**7.34.** Prove a identidade (7.11).

**7.35.** Assuma que a medida  $\pi$  é reversível para o processo  $X_t$ .

(a) Prove que  $X_t^*$  tem a mesma distribuição que o processo  $X_t$  em equilíbrio. Isto é, prove que

$$\mathbb{P}_\pi(X_{t_1}^* = x_1, \dots, X_{t_n}^* = x_n) = \mathbb{P}_\pi(X_{t_1} = x_1, \dots, X_{t_n} = x_n).$$



(b) Prove que a probabilidade de o processo  $X_t$  assumir uma certa seqüência de valores  $\{a_1, a_2, \dots, a_n\}$  no intervalo  $[s, t]$  é a mesma que a do processo  $X_t^*$  assumir os valores  $\{a_n, \dots, a_1\}$  no intervalo  $[-t, -s]$ .

**7.36.** Prove que um processo é reversível se a probabilidade de percorrer um ciclo em um sentido é a mesma que a de percorre-lo no outro. Isto é, prove que (7.2) e (7.3) são satisfeitas se e somente se para seqüência de pontos e tempos definidos como no enunciado da Proposição 7.12 tais que  $x_1 = x_n$

$$\mathbb{P}(X_{t_1} = x_1, \dots, X_{t_n} = x_n) = \mathbb{P}(X_{s_1} = x_n, \dots, X_{s_n} = x_1).$$

**7.37.** Para o processo reverso definido em (7.20) vale a igualdade

$$\mu(x) = \mu^*(x),$$

para todo estado  $x$ .

**7.38.** Prove o Lema 7.23.

**7.39.** Prove que se  $P^*$  é a matriz de transição do processo reverso de  $P$  em relação a  $\pi$ , então para todo  $n$  vale que

$$\pi(x)P^n(x, y) = \pi(y)(P^*)^n(y, x).$$

**7.40.** Seja  $X_t$  um processo que tem  $\pi$  como medida invariante. Seja  $X_t^*$  o processo reverso em relação a  $\pi$ . Prove que

- (1)  $\pi$  é invariante para  $X_t^*$ .
- (2) O processo reverso  $X_t^*$  tem a mesma distribuição que o processo  $X_t$  em equilíbrio se e somente se  $\pi$  for reversível para  $(X_t)$  Isto é, prove que

$$\mathbb{P}_\pi(X_{t_1}^* = x_1, \dots, X_{t_n}^* = x_n) = \mathbb{P}_\pi(X_{t_1} = x_1, \dots, X_{t_n} = x_n).$$

- (3) Prove que a probabilidade de o processo  $X_t$  assumir uma certa seqüência de valores  $\{a_1, a_2, \dots, a_n\}$  no intervalo  $[s, t]$  é a mesma

que a do processo  $X_t^*$  assumir os valores  $\{a_n, \dots, a_1\}$  no intervalo  $[-t, -s]$ .

**7.41.** Prove que para o sistema definido em (7.26) vale a igualdade

$$\mu(x, y) = \mu^*(x, y),$$

para todo estado  $(x, y)$ .

**Comentários e referências.** As idéias deste capítulo estão essencialmente contidas em Kelly (1979) e suas referências.

## 8. SISTEMAS DE PARTÍCULAS

Nessa sessão apresentamos três exemplos de sistemas de partículas em tempo discreto para os quais provaremos alguns resultados elementares. Esses processos são o modelo do votante, o processo de exclusão simples simétrico e o processo do contacto ou percolação orientada.

Sistemas markovianos de partículas com interação são cadeias de Markov tendo como espaço de estados  $\{0, 1\}^{\mathbb{Z}^d}$ . Nessa sessão consideraremos apenas  $d = 1$ . Informalmente esses sistemas descrevem a evolução de partículas em  $\mathbb{Z}$  com diversos tipos de interações locais. O objetivo geral da teoria é derivar o comportamento global desses sistemas.

### MODELO DO VOTANTE

Vamos considerar um processo  $X_n \in \{0, 1\}^{\mathbb{Z}^d}$  que informalmente descreve a evolução da opinião de uma comunidade de indivíduos representados pelos pontos de  $\mathbb{Z}$ . Cada indivíduo pode ter uma de duas opiniões, representadas por 0 e 1;  $X_n(i) \in \{0, 1\}$  representa a opinião do indivíduo  $i$  no instante  $n$ . Ao longo do tempo cada indivíduo atualiza sua opinião levando em consideração a opinião de seus vizinhos mais próximos.

O processo pode ser construído utilizando a construção gráfica de Harris. O tempo é discreto. Os votantes localizados nos pontos pares de  $\mathbb{Z}$  atualizarão suas opiniões nos instantes pares e os votantes localizados nos pontos ímpares o farão nos instantes ímpares. Para construir o processo utilizaremos uma sequência de variáveis independentes  $\{U_n(i), i \in \mathbb{Z}, n \in \mathbb{N}\}$  uniformes em  $[0, 1]$ .

Dada uma configuração inicial  $X_0(i) = x_i$ , onde  $x_i$  é uma sequência duplamente infinita de zeros e uns, construímos o processo de

acordo à seguinte regra:

$$X_n(i) = \begin{cases} X_{n-1}(i-1)\mathbf{1}\{U_n(i) > p\} + X_{n-1}(i+1)\mathbf{1}\{U_n(i) \leq p\}, \\ \quad \text{se } i+n \text{ par} \\ X_n(i), \quad \text{se } i+n \text{ impar.} \end{cases}$$

Ou seja, em cada sítio de  $\mathbb{Z}$  há um votante que em um de cada dois instantes muda de opinião escolhendo ao acaso a opinião de um dos dois vizinhos mais próximos. Se  $p = 1/2$  terá a mesma chance de escolher qualquer um dos dois vizinhos. Se  $p > 1/2$  então haverá uma tendência a escolher o vizinho da direita.

O primeiro resultado diz que o modelo do votante mantém fixa a proporção de votantes com opinião 1 ao mesmo tempo que obriga à comunidade a chegar a uma posição unânime. Mais formalmente, vamos provar o seguinte teorema.

**Teorema 8.1.** *Assuma que  $\mathbb{P}(X_0(i) = 1) = \rho \in [0, 1]$  para todo  $i \in \mathbb{Z}$ . Então:*

(i)  $\mathbb{P}(X_n(i) = 1) = \rho$  para todo  $n$  e  $i$ .

(ii) Para quaisquer valores  $i, j$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n(i) = X_n(j)) = 1.$$

Antes de demonstrar o teorema vamos introduzir a noção de *dualidade* que é a ferramenta principal da prova.

Vamos fixar um tempo inteiro  $n > 0$ . Vamos definir uma família de processos estocásticos  $Y_k^{i,n}$  com espaço de estados em  $\mathbb{Z}$  indo para trás no tempo da seguinte maneira:

(8.2)

$$Y_0^{i,n} = i,$$

$$Y_1^{i,n} = i, \text{ se } i+n \text{ é impar,}$$

$$Y_k^{i,n} = (Y_k^{i,n} - 1)\mathbf{1}\{U_{n-k-1} > p\} + (Y_k^{i,n} + 1)\mathbf{1}\{U_{n-k-1} \leq p\}$$

para  $2 \leq k \leq n$ . Usaremos a notação  $Y_n^{i,n}$  para indicar o processo  $Y_n$  com ponto inicial  $i$ . Essa família de processos será chamada

de processo dual. Note que a distribuição marginal de  $Y_n^{i,n}$  para  $k = 1, \dots, n$  é a de um passeio aleatório com matriz de transição

$$p(x, y) = \begin{cases} p & \text{se } y = x + 1 \\ 1 - p & \text{se } y = x - 1 \\ 0 & \text{em caso contrário.} \end{cases}$$

A próxima proposição mostra a relação entre o modelo do votante e o processo dual.

**Proposição 8.3.** *Para toda realização das variáveis  $U_n(i)$ , a seguinte fórmula da dualidade é verdadeira:*

$$X_n(i) = X_0(Y_n^{i,n}).$$

**Prova.** Imediata.  $\square$

A Proposição diz que o votante no ponto  $i$  no instante  $n$  tem a mesma opinião que o votante no ponto aleatório  $Y_n^{i,n}$  no instante zero. Sublinhamos o fato que esta propriedade é verdadeira para cada realização das variáveis uniformes.

**Prova do Teorema.** Da fórmula da dualidade podemos concluir que:

$$\begin{aligned} \mathbb{P}(X_n(i) = 1) &= \mathbb{P}(X_0(Y_n^{i,n}) = 1) \\ &= \sum_j \mathbb{P}(X_0(j) = 1 | Y_n^{i,n} = j) P(Y_n = j). \end{aligned}$$

Como  $X_0$  e  $Y_n^{i,n}$  são independentes ( $Y_n^{i,n}$  dependem apenas das variáveis aleatórias  $U_n(i)$  que independem de tudo), temos que a expressão acima é igual a

$$= \sum_j \rho P(Y_n^{i,n} = j) = \rho.$$

Isto prova o ponto (i).

Para provar o ponto (ii) note que, se para algum  $k$ ,

$$(8.4) \quad Y_k^{j,n} = Y_k^{i,n},$$

então para todo  $n \geq k$ ,

$$(8.5) \quad Y_n^{j,n} = Y_n^{i,n}.$$

Ou seja, podemos pensar que os  $Y_n^{i,n}$  são passeios aleatórios que evoluem independentemente até se encontrarem pela primeira vez. A partir do instante do encontro os passeios continuam juntos para sempre. Assim,

$$(8.6) \quad \begin{aligned} & \mathbb{P}(X_n(i) \neq X_n(j)) \\ &= \mathbb{P}(X_0(Y_n^{i,n}) \neq X_0(Y_n^{j,n})) \\ &= \mathbb{P}(X_0(Y_n^{i,n}) \neq X_0(Y_n^{j,n}) | Y_n^{j,n} = Y_n^{i,n}) \mathbb{P}(Y_n^{j,n} = Y_n^{i,n}) \\ &\quad + \mathbb{P}(X_0(Y_n^{i,n}) \neq X_0(Y_n^{j,n}) | Y_n^{j,n} \neq Y_n^{i,n}) \mathbb{P}(Y_n^{j,n} \neq Y_n^{i,n}) \\ &= \mathbb{P}(X_0(Y_n^{i,n}) \neq X_0(Y_n^{j,n}) | Y_n^{j,n} \neq Y_n^{i,n}) \mathbb{P}(Y_n^{j,n} \neq Y_n^{i,n}) \\ &\leq \mathbb{P}(Y_n^{j,n} \neq Y_n^{i,n}). \end{aligned}$$

Chamando  $Z_n^\ell = \frac{1}{2}(Y_n^{j,n} - Y_n^{i,n})$  para  $\ell = j - i$ , temos

$$(8.7) \quad \mathbb{P}(Y_n^{j,n} \neq Y_n^{i,n}) = \mathbb{P}(Z_n^\ell \neq 0),$$

onde  $Z_n$  é um passeio aleatório em  $\mathbb{Z}$  com as seguintes probabilidades de transição: para  $z \neq 0$

$$\begin{aligned} \mathbb{P}(Z_{n+1} = z + 1 | Z_n = z) &= p(1 - p) \\ \mathbb{P}(Z_{n+1} = z - 1 | Z_n = z) &= p(1 - p) \\ \mathbb{P}(Z_{n+1} = z | Z_n = z) &= p^2 + (1 - p)^2. \end{aligned}$$

Com efeito,

$$\begin{aligned} & \mathbb{P}(Z_{n+1} = z + 1 | Z_n = z) \\ &= \frac{\mathbb{P}(Z_{n+1} = z + 1 | Z_n = z)}{\mathbb{P}(Z_n = z)} \\ &= \sum_x \frac{\mathbb{P}(Y_{n+1}^{j,n} = x - 1, Y_{n+1}^{i,n} = x + z + 1, Y_n^{j,n} = x, Y_n^{i,n} = x + z)}{\mathbb{P}(Z_n = z)} \\ &= \sum_x \mathbb{P}(V_n(x) > p) \mathbb{P}(V_n(x + z) \leq p) \frac{\mathbb{P}(Y_n^{j,n} = x, Y_n^{i,n} = x + z)}{\mathbb{P}(Z_n = z)} \\ &= p(1 - p), \end{aligned}$$

onde  $V_n(i)$  é uma família de variáveis independentes uniformes em  $[0, 1]$ .

A demonstração é análoga para  $z - 1$ ; para  $z$  segue por ser esse o complementar dos outros dois casos. Observe-se que 0 é um ponto absorvente para esse passeio:

$$\mathbb{P}(Z_{n+1} = 0 | Z_n = 0) = 1.$$

Isso segue de (8.4) e (8.5), já que o passeio  $Z_n$  dá a distância entre os dois passeios duais e quando esses se encontram, continuam juntos para sempre.

Assim, concluímos que o passeio aleatório  $Z_n$  é simétrico e absorvente na origem.

A probabilidade de um passeio aleatório simétrico absorvente na origem ser igual a zero no instante  $t$  começando no ponto  $\ell$  é

$$(8.8) \quad \mathbb{P}(Z_n^\ell = 0) \sim 2 \left( 1 - \Phi \left( \frac{\ell}{\sqrt{n}} \right) \right),$$

onde  $\Phi$  é a distribuição acumulada da normal padrão:

$$\Phi(r) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^r e^{-u^2/2} du.$$

(Veja sessão de primeiras passagens do passeio aleatório no volume I do Feller (1968)). O lado direito de (8.8) converge a 1 quando  $n \rightarrow \infty$ . Portanto (8.7) converge a 0, o que implica que o lado esquerdo de (8.6) converge a 0. Isso conclui a demonstração de (ii).

□

Vamos usar essa mesma técnica para calcular as correlações assintóticas, quando  $\{X_0(i)\}_i$  é uma sequência de variáveis aleatórias independentes identicamente distribuídas de Bernoulli com parâmetro  $\rho$ , isto é, para qualquer conjunto finito  $A \subset \mathbb{Z}$ ,

$$(8.9) \quad \mathbb{P}(X_0(i) = 1, i \in A) = \rho^{|A|}.$$

**Proposição 8.10.** *Nas condições acima vale a seguinte identidade*

$$(8.11) \quad \text{cov}(X_n(i)X_n(j)) = \rho(1 - \rho)\mathbb{P}(Z_n^{i-j} = 0).$$

**Prova.** Para calcular a covariância entre  $X_n(i)$  e  $X_n(j)$  calculamos primeiro

$$\begin{aligned} \mathbb{E}(X_n(i)X_n(j)) &= \rho^2\mathbb{P}(Z_n \neq 0) + \rho\mathbb{P}(Z_n = 0) \\ &= \rho^2 + \rho(1 - \rho)\mathbb{P}(Z_n = 0). \end{aligned}$$

Portanto a covariância entre  $X_n(i)$  e  $X_n(j)$  é dada por

$$\text{cov}(X_n(i)X_n(j)) = \mathbb{E}(X_n(i)X_n(j)) - \rho^2 = \rho(1 - \rho)\mathbb{P}(Z_n^{i-j} = 0). \quad \square$$

**Observação 8.12.** Como o desvio padrão de  $X_n(i)$  é  $\sqrt{\rho(1 - \rho)}$ , o coeficiente de correlação é exatamente  $\mathbb{P}(Z_n^{i-j} = 0)$ , que é da ordem de  $1 - (1/\sqrt{n})$  para  $i - j$  fixo.

### PROCESSO DE EXCLUSÃO SIMPLES SIMÉTRICO.

O processo de exclusão simples descreve a evolução de um sistema de partículas que realizam passeios aleatórios independentes exceto pelo fato que as partículas são proibidas de pular sobre lugares já ocupados por outras partículas.

Vamos considerar apenas o caso unidimensional com saltos permitidos apenas para vizinhos mais próximos. A construção pelo método gráfico de Harris é feita da seguinte maneira. A variável  $X_n(i) \in \{0, 1\}$  para  $n \in \mathbb{Z}$  e  $i \in \mathbb{Z}$  indica a ocupação ou não do ponto  $i$  no instante  $n$ . A cada instante  $n$  associamos uma variável aleatória  $V_n$  com distribuição de Bernoulli de parâmetro  $1/2$ . Para cada ponto  $i \in \mathbb{Z}$  e  $n \geq 0$  consideramos as variáveis aleatórias uniformes  $U_n(i)$  independentes das  $V_n$  e de tudo mais.

Sejam

$$W_n(i) = \begin{cases} 1, & \text{se } i \text{ e } V_n \text{ tem a mesma paridade e } U_n(i) < p; \\ 0, & \text{em caso contrário.} \end{cases}$$



Fixada uma configuração inicial  $X_0(i)$ , definimos

$$X_n(i) = X_{n-1}(i+1)W_n(i) + X_{n-1}(i-1)W_n(i-1) \\ + X_{n-1}(i)(1 - W_n(i) - W_n(i-1)).$$

Ou seja, quando  $V_n = 0$  com probabilidade  $p$ , os conteúdos dos sítios pares e dos ímpares sucessivos são trocados independentemente para cada par com probabilidade  $p$ . Se  $V_n = 1$ , os conteúdos dos sítios ímpares e dos pares sucessivos são trocados independentemente com probabilidade  $p$ .

Esse processo pode ser chamado também de processo de mistura porque simplesmente a cada instante os conteúdos de sítios vizinhos são trocados. Note que o processo é conservativo: não há criação nem destruição de partículas ao longo da evolução. Se o processo for construído em uma caixa finita contida em  $\mathbb{Z}$ , o número total de sítios ocupados não muda com o tempo. Portanto é razoável pensar que se começarmos o processo com uma densidade  $\rho$  de sítios ocupados, a longo alcance também teremos densidade  $\rho$ .

Para provar esse e outros resultados referentes a esse processo vamos de novo utilizar a técnica da *dualidade*.

O processo dual é definido da seguinte maneira. Para cada  $i \in \mathbb{Z}$ , seja  $Y_n^{i,n}$  um processo que corre para trás no tempo, construído usando as mesmas variáveis aleatórias  $W_n(i)$  que usamos para construir o processo de exclusão simples. Para  $k \in \{0, \dots, n\}$  definimos

$$Y_{k+1}^{i,n} = Y_k^{i,n} + W_{n-k}(Y_k^{i,n}) - W_{n-k}(Y_k^{i,n} - 1).$$

Ou seja, dependendo de  $W_n(y)$  e  $W_n(y-1)$ , o processo  $Y_k^i$  pode dar um passo à direita, um passo para a esquerda ou ficar no mesmo lugar. Realiza um passo de comprimento 1 com probabilidade  $p$  ou fica no lugar com probabilidade  $1-p$ :

$$\mathbb{P}(Y_{k+1}^{i,n} = y+1 | Y_k^{i,n} = y) = \mathbb{P}(W_{n-k}(y) = 1) = \frac{1}{2}p$$

$$\mathbb{P}(Y_{k+1}^{i,n} = y-1 | Y_k^{i,n} = y) = \mathbb{P}(W_{n-k}(y-1) = 1) = \frac{1}{2}p$$

$$\mathbb{P}(Y_{k+1}^{i,n} = y | Y_k^{i,n} = y) = \mathbb{P}(W_{n-k}(y) = 0, W_{n-k}(y-1) = 0) = 1-p.$$

Note que se em algum momento dois passeios são vizinhos e um deles faz um passo na direção do outro, então o outro também simultaneamente faz um passo na direção do primeiro. Isto é, se  $Y_k^{i,n} = y$ ,  $Y_k^{j,n} = y + 1$  e  $W_{n-k}(y) = 1$ , então

$$Y_{k+1}^{i,n} = y + 1 \text{ e } Y_{k+1}^{j,n} = y.$$

A continuação estabelecemos a relação entre  $Y_n$  e  $X_n(\cdot)$ .

**Lema 8.13.** *Fórmula da Dualidade.* *A seguinte fórmula vale para todo  $i$  e para todo  $n$ :*

$$X_n(i) = X_0(Y_n^{i,n}).$$

**Prova.** A prova é imediata e pode ser feita indução.  $\square$

Uma consequência da fórmula da dualidade é que se começarmos com uma densidade  $\rho$  de lugares ocupados, então teremos a mesma densidade para todos os tempos sucessivos. Por outro lado, se a distribuição inicial de lugares ocupados obedece à chamada medida produto (quando se trata de variáveis independentes identicamente distribuídas), então em instantes sucessivos teremos a mesma distribuição.

Vamos definir uma distribuição sobre o conjunto de sequências de zeros e uns. Diremos que  $X \in \{0, 1\}^{\mathbb{Z}}$  tem distribuição produto com densidade  $\rho$  se para todo conjunto finito  $A$  de números inteiros vale que

$$\mathbb{P}(X(i) = 1, i \in A) = \rho^{|A|}.$$

**Teorema 8.14.** (i) *Se  $\mathbb{P}(X_0(i) = 1) = \rho$  para todo  $i$ , então para todo  $i$  e para todo  $n$  vale que  $\mathbb{P}(X_n(i) = 1) = \rho$ .*

(ii) *A medida produto é invariante para o processo de exclusão simples.*

**Prova.** O ponto (i) segue diretamente por dualidade. Com efeito

$$\begin{aligned} \mathbb{P}(X_n(i) = 1) &= \mathbb{P}(X_0(Y_n^{i,n}) = 1) \\ &= \sum_y \mathbb{P}(X_0(y) = 1 \mid Y_n^{i,n} = y) \mathbb{P}(Y_n^{i,n} = y) \\ &= \sum_y \mathbb{P}(X_0(y) = 1) \mathbb{P}(Y_n^{i,n} = y) \\ &= \rho \sum_y \mathbb{P}(Y_n^{i,n} = y) = \rho. \end{aligned}$$

Demonstrar (ii) significa provar que se  $\mathbb{P}(X_0(i) = 1, i \in A) = \rho^{|A|}$ , para todo  $A$  finito então  $\mathbb{P}(X_n(i) = 1, i \in J) = \rho^{|J|}$  para todo  $J$  finito, para todo  $n$ . Usando novamente dualidade,

$$\begin{aligned} \mathbb{P}(X_n(i) = 1, i \in J) &= \mathbb{P}(X_0(Y_n^{i,n}) = 1, i \in J) \\ &= \sum_{i \in J} \sum_{y_i \in \mathbb{Z}} \mathbb{P}(X_0(y_i) = 1, i \in J \mid Y_n^{i,n} = y_i, i \in J) \mathbb{P}(Y_n^{i,n} = y_i, i \in J) \\ &= \sum_{i \in J} \sum_{y_i \in \mathbb{Z}} \mathbb{P}(X_0(y_i) = 1, i \in J) \mathbb{P}(Y_n^{i,n} = y_i, i \in J) \\ &= \rho^{|J|} \sum_{i \in J} \sum_{y_i \in \mathbb{Z}} \mathbb{P}(Y_n^{i,n} = y_i, i \in J) = \rho^{|J|}. \quad \square \end{aligned}$$

### Limite hidrodinâmico.

Agora vamos considerar o que é chamado de limite hidrodinâmico. A motivação básica é derivar a equação do calor a partir do estudo do sistema microscópico de partículas modelado pelo processo de exclusão simples simétrico.

Começaremos com a seguinte proposição que descreve o comportamento assintótico de  $Y_n^{i,n}$ , isto é, de um passeio aleatório. Trata-se de um teorema limite central.

**Lema 8.15.** *Valem os limites seguintes:*

$$\lim_{n \rightarrow \infty} \mathbb{P}(Y_{tn}^{r\sqrt{pn}} \leq a\sqrt{pn}) = \Phi_t(a - r),$$

onde  $\Phi_t$  é a função acumulada da normal centrada com variância  $pt$ :

$$\Phi_t(r) = \frac{1}{\sqrt{2\pi tp}} \int_{-\infty}^r e^{x^2/2tp} dx.$$

**Prova.** A prova para  $r = 0$  segue do fato que  $Z_k^0 = Y_{k+1}^0 - Y_k^0$  são variáveis aleatórias independentes com a mesma distribuição de  $Y_1^0$ :

$$\mathbb{P}(Y_1^0 = 1) = \mathbb{P}(Y_1^0 = -1) = \frac{p}{2} \quad \text{e} \quad \mathbb{P}(Y_1^0 = 0) = 1 - p.$$

Ou seja, são variáveis aleatórias independentes e centradas com variância  $\text{Var } Y_1^0 = p$ . Portanto

$$\mathbb{P}(Y_{tn}^0 \leq a\sqrt{pn}) = \Phi_t(a).$$

Agora note que a distribuição de  $Y_n^{i,n}$  é a mesma que a de  $Y_{n-i}^0$ . Dai sai o restante do lema.  $\square$

Vamos introduzir o limite hidrodinâmico. Vamos supor que a configuração inicial é

$$X_0(i) = \mathbf{1}\{i \geq 0\}.$$

Isto é, não há partículas à esquerda da origem e todos os sítios estão ocupados à direita (incluindo a origem). Queremos ver como evolui esse processo. Para isso usamos novamente a fórmula da dualidade. Na origem, teremos

$$\begin{aligned} \mathbb{P}(X_n(0) = 1) &= \mathbb{P}(X_0(Y_n^{0,n}) = 1) \\ &= \sum_y \mathbb{P}(X_0(y) = 1 \mid Y_n^{0,n} = y) \mathbb{P}(Y_n^{0,n} = y) \\ &= \sum_y \mathbb{P}(X_0(y) = 1) \mathbb{P}(Y_n^{0,n} = y) \\ &= \sum_{y>0} \mathbb{P}(Y_n^{0,n} = y) = \mathbb{P}(Y_n^{0,n} > 0), \end{aligned}$$

porque  $X_0(y) = 1$  se e somente se  $y > 0$ . Mas o limite dessa última expressão é

$$\lim_{n \rightarrow \infty} \mathbb{P}(Y_n^{0,n} > 0) = 1 - \Phi_1(0) = 1/2,$$

pelo Lema 8.15.

Em geral teremos

$$\begin{aligned} \mathbb{P}(X_{nt}(r\sqrt{n}) = 1) &= \mathbb{P}(X_0(Y_{nt}^{r\sqrt{n}}) = 1) \\ &= \sum_y \mathbb{P}(X_0(y) = 1 \mid Y_{nt}^{r\sqrt{n}} = y) \mathbb{P}(Y_{nt}^{r\sqrt{n}} = y) \\ &= \sum_y \mathbb{P}(X_0(y) = 1) \mathbb{P}(Y_{nt}^{r\sqrt{n}} = y) \\ &= \sum_{y>0} \mathbb{P}(Y_{nt}^{r\sqrt{n}} = y) \\ &= \sum_{y>0} \mathbb{P}(Y_{nt}^0 = y - r\sqrt{n}) \\ &= \mathbb{P}(Y_{nt}^0 > -r\sqrt{n}) \\ &= \mathbb{P}(Y_{nt}^0 < r\sqrt{n}). \end{aligned}$$

Mas

$$\lim_{n \rightarrow \infty} \mathbb{P}(Y_{nt}^0 < r\sqrt{n}) = \Phi_t(r),$$

pelo teorema do limite central.  $\square$

Por outro lado, note que  $u(t, r) = \sqrt{pt}(1 - \Phi(r))$  é a solução da equação do calor:

$$\frac{\partial u}{\partial t} = p \frac{\partial^2 u}{\partial r^2},$$

com condições iniciais

$$u(0, r) = \mathbf{1}\{r \geq 0\}.$$

A função  $u(t, r)$  representa a temperatura no ponto  $r$  no instante  $t$  em uma barra infinita, que no instante 0 tinha temperatura 1 à direita da origem e temperatura 0 à esquerda.

## PROCESSO DE CONTACTO

O processo de contacto em tempo discreto, também conhecido como processo de percolação orientada é um modelo muito simples que apresenta transição de fase. Esse processo pode ser descrito informalmente como descrevendo a evolução de um sistema de partículas com nascimento e morte. Partículas morrem espontaneamente e criam novas partículas nos pontos vizinhos desocupados. Nessa sessão vamos construir uma família de processos de contacto indexados por um parâmetro  $p \in [0, 1]$ . Esse parâmetro é a probabilidade com a qual uma nova partícula é criada. O principal resultado desta sessão é que o modelo apresenta comportamentos qualitativamente diferentes para valores do parâmetro à direita e à esquerda de um valor crítico. Isto é chamado de transição de fase.

Para definir o processo de contacto usaremos duas sequências de variáveis aleatórias independentes  $E_n(i)$  e  $D_n(i)$  para  $i \in \mathbb{Z}$  e  $n \geq 0$  uniformemente distribuídas no intervalo  $[0, 1]$ . O parâmetro  $n$  joga o papel de tempo.

Vamos definir indutivamente o processo. Dado  $X_n \in \{0, 1\}^{\mathbb{Z}}$ , definimos

$$X_{n+1}(i) = \max\{X_n(i-1)\mathbf{1}\{D_n(i-1) \leq p\}, X_n(i+1)\mathbf{1}\{E_n(i-1) \leq p\}\}.$$

A sequência duplamente infinita de zeros e uns descrita por  $X_n$  é identificada com o subconjunto de  $\mathbb{Z}$  de pontos ocupados por partículas no instante  $n$ . Vemos que se no instante inicial apenas pontos pares estão ocupados, em instantes pares sucessivos somente pares poderão estar ocupados e em instantes ímpares sucessivos, somente ímpares poderão estar ocupados. Assim podemos pensar que o processo está definido em um sub-reticulado de  $\mathbb{Z}^2$  que chamaremos

$$\mathbf{Z} = \{(i, n) : i + n = \text{par}\}.$$

O reticulado  $\mathbf{Z}$  induz um grafo com laços orientados  $((i, n), (i \pm 1, n + 1))$ .

Podemos interpretar o processo de contacto como um processo de ramificação no qual cada indivíduo morre e dá lugar a no máximo

dois filhos, um em cada uma das posições vizinhas com probabilidade  $p$  independentemente. Somada a esse processo de ramificação há interação entre os indivíduos: no máximo pode haver um indivíduo por lugar.

Acabamos de construir uma família de processos de contacto indexada por  $p$ . Na definição acima o número inicial de 1's pode ser finito ou infinito.

**Definição.** Seja  $X_n^0$  o processo começando com apenas a origem ocupada no instante 0. Portanto  $X_0^0(0) = 1$  e  $X_0^0(i) = 0$  para todo  $i \neq 0$ . Seja  $|X_n|$  o número de pontos ocupados no instante  $n$ . Seja

$$\pi_0(p) = \mathbb{P}(|X_n^0| = 0 \text{ para algum } n).$$

Diremos que o processo *morre* se  $\pi_0 = 1$ . Em caso contrário, se  $\pi_0 < 1$ , diremos que o processo *sobrevive*.

Provaremos que este modelo apresenta transição de fase:

**Teorema 8.16.** *Existe um valor  $p_c \in (0, 1)$ , tal que se  $p > p_c$ , o processo sobrevive e se  $p < p_c$ , o processo morre.*

Evidentemente, se  $p = 0$  o processo morre e se  $p = 1$  o processo sobrevive. Por isso a parte não trivial desse teorema é provar que o valor crítico  $p_c$  está contido no intervalo  $(0, 1)$ .

**Prova:** Primeiro provaremos que  $p_c > 0$ . Isso é, se  $p$  é suficientemente pequeno então  $\pi_0(p) = 1$ . Vamos definir os seguintes passeios aleatórios, os dois começando da origem:  $Y_0 = Z_0 = 0$  e

$$Y_{n+1} = Y_n + 21\{D_n(Y_n) > p\} - 1$$

$$Z_{n+1} = Z_n - 21\{E_n(Y_n) > p\} + 1.$$

Assim o passeio  $Y_n$  faz um passo de uma unidade para a direita com probabilidade  $p$  e um passo de uma unidade para a esquerda com probabilidade  $1 - p$ . Já o passeio  $Z_n$  faz um passo de uma unidade para a esquerda com probabilidade  $p$  e um passo de uma unidade para a direita com probabilidade  $1 - p$ . Além disso  $Y_0 = Z_0 = 0$ .

Defina

$$R_n = \max\{i : X_n^0(i) = 1\}$$

$$L_n = \min\{i : X_n^0(i) = 1\},$$

se  $X_n^0(i) \neq 0$  para algum  $i$ . Se  $X_n^0(i) \equiv 0$ , definimos  $R_n = -\infty$  e  $L_n = +\infty$ .

O processo atinge a configuração vazia no instante  $T$  em que esses caminhos se cruzam, isto é

$$T = \min\{n : L_n > R_n\} = \min\{n : X_n^0 = \emptyset\}.$$

Portanto, para provar que o processo morre basta provar que

$$P(T < \infty) = 1.$$

Por outro lado, segue da construção do processo que

$$R_n \leq Y_n \text{ e } L_n \geq Z_n.$$

De fato, se para algum  $n$ ,  $R_n = Y_n$ , tanto  $R_n$  como  $Y_n$  darão um passo à direita, se  $D_n(Y_n) < p$ . Se  $D_n(Y_n) \geq p$ , então  $Y_{n+1}$  será igual a  $Y_n - 2$  e  $R_{n+a}$  vai se deslocar pelo menos duas unidades para a esquerda. Se  $R_n < Y_n$ , então a distância mínima entre as duas é de duas unidades e em um passo não poderá haver cruzamento. O mesmo raciocínio serve para  $L_n$  e  $Z_n$ .

A dominação acima descrita implica que

$$T \leq S = \min\{n : Y_n < Z_n\}$$

e

$$P(T < \infty) \geq P(S < \infty).$$

Ficamos assim com o problema de provar que, para  $p$  suficientemente pequeno, a probabilidade de  $Y_n$  ficar menor que  $Z_n$  para algum  $n$  é 1. Ou seja queremos provar que

$$P(Y_n - Z_n = -1 \text{ para algum } n) = 1.$$

Mas, chamando

$$W_n = \frac{Y_n - Z_n + 2}{2},$$



isso equivale a provar que

$$P(W_n = 0 \text{ para algum } n \mid W_0 = 1) = 1,$$

Qual é a lei de  $W_n$ ?

$$(8.17) \quad \begin{aligned} P(W_{n+1} = w + 1 \mid W_n = w) &= p^2 \\ P(W_{n+1} = w - 1 \mid W_n = w) &= (1 - p)^2 \\ P(W_{n+1} = w \mid W_n = w) &= 2p(1 - p). \end{aligned}$$

A probabilidade desse passeio aleatório ser absorvido na origem é 1 se  $p^2 < (1 - p)^2$  (esse resultado é pedido como exercício). Essa desigualdade é verdadeira se  $p < 1/2$ . Isso prova que  $p_c \geq 1/2$ .

Vamos agora provar que  $p_c < 1$ . Inicialmente vamos introduzir a noção de *aglomerado*.

**Definição 8.18.** O *aglomerado da origem* é o conjunto

$$\{(i, n) \in \mathbf{Z} : X_n^0(i) = 1\},$$

onde  $X_n^0$  é o processo obtido tomando como configuração inicial a configuração com uma única partícula na origem.

O aglomerado da origem descreve a história dos descendentes da partícula inicialmente localizada na origem. Dizer que o processo  $X_n^0$  sobrevive é equivalente a dizer que o aglomerado da origem é infinito. Assim para provar que  $p_c < 1$  basta mostrar que para  $p$  suficientemente grande mas ainda menor que 1, a probabilidade de existir um aglomerado infinito é positiva.

Para isso usaremos o famoso *argumento de Peierls*, que consiste essencialmente em contar as diferentes formas de acontecer o evento “o aglomerado da origem é finito”. Essa contagem é depois usada para achar uma majoração superior menor que 1 para a probabilidade desse evento quando  $p$  é suficientemente grande.

Começemos com uma sequência de definições.

**Definição 8.19.** (i) A rede *dual* é o conjunto  $\mathbf{D} = \{(i, n) : i + n \text{ ímpar}\}$ .

(ii) Dois pontos  $(i, n)$  e  $(j, m)$  de  $\mathbf{D}$  são ditos vizinhos se  $|i - j| = 1$  e  $|m - n| = 1$ .

(iii) Um laço é um par ordenado de pontos vizinhos.

(iv) Um contorno  $\gamma$  de comprimento  $k$  é definido por uma sequência  $k + 1$  pontos de  $\gamma = G_0, \dots, G_k \in \mathbf{D}$  satisfazendo as propriedades seguintes:

$$G_0 = (-1, 0); G_k = (1, 0).$$

Além disso

$$\text{se } G_{\ell-1} = (i, n), \text{ e } G_\ell = (i - 1, n + 1),$$

$$\text{então } G_{\ell+2} \in \{(i - 2, n + 2), (i, n + 2)\}$$

$$\text{se } G_{\ell-1} = (i, n), \text{ e } G_\ell = (i + 1, n + 1),$$

$$\text{então } G_{\ell+2} \in \{(i, n + 2), (i + 2, n + 2), (i + 2, n)\}$$

$$\text{se } G_{\ell-1} = (i, n), \text{ e } G_\ell = (i + 1, n - 1),$$

$$\text{então } G_{\ell+2} \in \{(i + 2, n), (i + 2, n - 2)\}.$$

Note que pontos sucessivos de um contorno são vizinhos. Assim sendo, podemos pensar em um contorno como uma sequência de laços.

Diremos que  $\gamma$  é um contorno fechado se as variáveis aleatórias  $D_n(i)$  e  $E_n(i)$  correspondentes aos laços que cortam laços do contorno crescentes na primeira das duas coordenadas estão fechados:

$$[(i, n), (i + 1, n + 1)] \in \gamma \text{ então } E_n(i + 1) > p$$

$$[(i, n), (i + 1, n - 1)] \in \gamma \text{ então } D_{n-1}(i) > p.$$

A cada aglomerado finito  $\xi = \cup_n (X_n, n)$  corresponde um contorno fechado sem loops (para o qual cada vértice é usado apenas uma vez)  $\gamma(\xi)$  cujo interior  $\dot{\gamma}$  contem os pontos de  $\xi$ . Esse contorno tem a propriedade que os pontos da rede  $\mathbf{Z}$  à esquerda de cada laço pertencem a  $\xi$  e os da direita não pertencem.

Dados  $D_n(i)$  e  $E_n(i)$  há uma maneira construtiva de definir  $\gamma(\xi)$ . Propomos ao leitor que formule um algoritmo para construir esse contorno supondo conhecidos os valores  $D_n(i)$  e  $E_n(i)$ .

Agora, o evento  $\{|\xi| < \infty\}$  é idêntico ao evento

$$\cup_k \cup_{\gamma \in \Gamma_k} \{\xi \subset \dot{\gamma}\},$$

ou seja, o aglomerado da origem é finito se e somente se estiver contido no interior de algum dos contornos finitos. Particionamos os contornos finitos de acordo ao comprimento.  $\Gamma_k$  é o conjunto de todos os contornos de comprimento exatamente igual a  $k$ . Note que o número de laços de cada contorno deve ser par (exercício). Note que a probabilidade de um contorno qualquer de comprimento  $k$  ser fechado é exatamente  $(1-p)^{(k/2)+1}$  (exercício). Assim

$$\mathbb{P}(\{|\xi| < \infty\}) = \mathbb{P}(\cup_k \cup_{\gamma \in \Gamma_k} \{\xi \subset \dot{\gamma}\}) \leq \sum_k N_k (1-p)^{k/2} (1-p),$$

onde  $N_k$  é o cardinal de  $\Gamma_k$  e  $(1-p)^k$  é a probabilidade de um contorno de comprimento  $k$  ser fechado.

Como de cada nô do contorno podem sair no máximo três ramas,

$$N_k \leq 3^k.$$

Assim,

$$\begin{aligned} \sum_k N_k (1-p)^{k/2} (1-p) &\leq \sum_{k \geq 1} \left(3\sqrt{1-p}\right)^k (1-p) \\ &= \frac{3\sqrt{1-p}(1-p)}{1-3\sqrt{1-p}} < 1, \end{aligned}$$

se  $3\sqrt{1-p} < 1$ . Mas isso equivale a  $p > 8/9$ . Portanto, uma condição suficiente para que o processo sobreviva é que  $p > 8/9$  o que implica que  $p_c < 8/9$ .

### Exercícios.

#### PARA O MODELO DO VOTANTE

**8.20.** Simule o processo de exclusão simples em uma caixa fechada de tamanho 5 por 6 instantes de tempo. Considere uma configuração inicial aleatória.

**8.21.** Prove a fórmula de dualidade para o modelo do votante.

**8.22.** Prove que o processo  $Y_n^{i,n}$  definido em (8.2) é uma cadeia de Markov e calcule as probabilidades de transição.

**8.23.** Na simulação do exercício 8.20 simule  $Y_4^{0,8}$  e  $Y_5^{0,8}$ .

**8.24.** Prove que para o passeio aleatório  $Z_n$  vale que

$$\mathbb{P}(Z_n^\ell = 0) \sim \frac{\text{constante}}{\sqrt{n}}.$$

Qual é a dependência em  $\ell$ ?

**8.25.** Forneça uma definição semelhante do modelo do votante em dimensões 2 e 3.

(i) Prove que em dimensão 2 há ainda unanimidade como em dimensão 1.

(ii) Prove que em dimensão 3 não há unanimidade. Isso implica em particular que em dimensão 3 há medidas invariantes para o processo que não são triviais. Sugestão: use dualidade e o fato que dois passeios aleatórios independentes em dimensão 3 tem uma probabilidade positiva de não se encontrarem nunca.

## PARA O PROCESSO DE EXCLUSÃO SIMPLES

**8.26.** Seja  $A_n = \sum_i i \mathbf{1}\{X_n^0(i) = 1\}$  a posição da única partícula no instante  $n$  no processo de exclusão simples quando o processo começou apenas com a origem ocupada. Prove que  $Q_n$  é um passeio aleatório e calcule a matriz de transição.

**8.27.** Prove a fórmula de dualidade do Lema 8.13.

**8.28.** Simule o processo de exclusão simples em uma caixa fechada de tamanho 5 por 6 instantes de tempo. Considere uma configuração inicial aleatória.

**8.29.** Prove que a distribuição de  $Y_n^{i,n}$  é a mesma que a de  $Y_{n-i}^{0,n}$ .

## PARA O PROCESSO DE CONTACTO

**8.30.** Simule o processo de percolação orientada para vários valores de  $p$  e faça uma estimativa grosseira do verdadeiro valor de  $p_c$ .

**8.31.** Simule simultaneamente o processo de contacto e os passeios aleatórios  $Y_n$  e  $Z_n$ . Verifique que o processo morre no momento em que  $Y_n$  e  $Z_n$  se cruzam.

**8.32.** Prove que se  $W_n = Y_n - Z_n$ , então a lei de  $W_n$  está dada pela expressão  $\text{equ}(un)$ .

**8.33.** Prove que se o processo de contacto sobrevive até o instante  $n$  para um determinado  $p$ , então também sobrevive para  $q > p$  para o mesmo  $n$ .

**8.34.** Usando o exercício 4, prove que há apenas um valor  $p_c$  tal que se  $p > p_c$  o processo sobrevive e se  $p < p_c$  o processo morre.

**8.35.** Seja  $W_n$  um passeio aleatório em  $\mathbb{N}$  com transições

$$\mathbb{P}(W_{n+1} = w + 1 \mid W_n = w) = p$$

$$\mathbb{P}(W_{n+1} = w - 1 \mid W_n = w) = q$$

$$\mathbb{P}(W_{n+1} = w \mid W_n = w) = r$$

e  $p + q + r = 1$ . Prove que a probabilidade de esse passeio começando do 1 chegar alguma vez à origem é 1 se  $p \leq q$  e menor que 1 se  $p > q$ .

**Comentários e referências.** Sistemas de Partículas é agora um tema clássico em processos estocásticos. As referências padrão são Liggett (1985) e Durrett (1988). Neste capítulo apresentamos alguns dos modelos clássicos mas para tempo discreto.



## REFERÊNCIAS

- Athreya, K. B.; Ney, P. (1978) A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* 245, 493-501.
- Dobrushin, R. L. (1968) Gibbsian random fields for lattice systems with pairwise interactions. *Funct. Anal. Appl.* 2, no. 4, 292-301.
- Dobrushin, R. L. (1968) The description of random fields by means of conditional probabilities and conditions of its regularity. *Theory Prob. Appl.* 13 no. 2, 197-224.
- Durrett, Richard (1988) Lecture notes on particle systems and percolation. The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA.
- Feller, W. (1968) An introduction of Probability Theory and its Applications. John Wiley & Sons, Ltd.
- Garcia, N.L. (1995) Birth and death processes as projections of higher dimensional Poisson processes. *Adv. Appl. Probab.* 27, 911-930.
- Kelly, Frank P. (1979) Reversibility and stochastic networks. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Ltd., Chichester.
- Kurtz, T.G. (1989) Stochastic processes as projections of Poisson random measures. Special invited paper at IMS meeting. Washington D.C.. Unpublished.
- Liggett, Thomas M. (1985) Interacting particle systems. Springer-Verlag, New York-Berlin.
- Lindvall, T. (1992) Lectures on the coupling method. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York.
- Neveu, J. (1977) Processus ponctuels. École d'Été de Probabilités

de Saint-Flour, VI—1976, pp. 249–445. Lecture Notes in Math., Vol. 598, Springer-Verlag, Berlin.

Ross, Sheldon M. (1983) Stochastic processes. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Lectures in Mathematics, 14. John Wiley & Sons, Inc., New York.



Impresso na Gráfica do



pele Sistema Xerox / 5390

