

Dinâmica Estocástica: três exemplos

S. Friedli

Universidade Federal de Minas Gerais

1^o Colóquio da Região Sudeste

Abril de 2011

Sumário

1	O gás de Ehrenfest	3
1.1	Introdução	3
1.2	O modelo	5
1.3	Tempos longos: distribuição invariante	8
1.3.1	Tempo de permanência	11
1.3.2	Tempo de primeiro retorno	12
1.3.3	Sobre a existência do limite $\lim_{n \rightarrow \infty} \mu^{(n)}$	13
1.4	Exercícios	14
1.5	Referências bibliográficas	14
2	O passeio aleatório	15
2.1	Introdução	15
2.2	O modelo	17
2.3	Recorrência/Transiência	19
2.4	Exercícios	23
2.5	Referências bibliográficas	23
3	O processo de ramificação	25
3.1	Introdução	25
3.2	O modelo de Galton-Watson	26
3.3	Prova do Teorema	28
3.4	Exercícios	32
3.5	Referências bibliográficas	33
A	Lembrete de Probabilidade	35
B	Soluções dos Exercícios	37
	Referências Bibliográficas	43

Introdução

Este minicurso tem como objetivo estudar três fenômenos bem conhecidos:

1. a difusão de um gás, de um container cheio para um container vazio;
2. o movimento errático de uma partícula de pólen suspensa em um líquido;
3. a dinâmica do número de indivíduos vivos numa população.

Cada um desses fenômenos apresenta, por um lado, um comportamento *microscópico* complexo e, por outro lado, um comportamento *macroscópico* simples, determinístico. A complexidade sendo grande demais para ser estudada exatamente com ferramentas clássicas da mecânica (i.e. o cálculo diferencial), as leis clássicas da dinâmica microscópica serão trocadas por dinâmicas *estocásticas*.

Assim encontraremos, por exemplo, desde a primeira aula, o conceito chave de *cadeia de Markov*, central na Teoria da Probabilidade. Os modelos introduzidos neste minicurso foram de grande importância no desenvolvimento da mecânica estatística e da teoria da probabilidade no século 20 e continuam a ser usados por boa parte dos pesquisadores contemporâneos.

As ferramentas matemáticas necessárias serão apresentadas ao longo do curso. Somente será suposto que o leitor tenha noções (de cálculo e) de combinatória e probabilidade elementar. O fim do texto contém um apêndice com alguns lembretes, que o leitor poderá usar, se for necessário. O fim de cada capítulo contém algumas referências bibliográficas e sugestões de leitura.

O símbolo “:=” será usado para definir um objeto. Por exemplo,

$$c := \min\{n > 4 : n \text{ é primo}\}$$

significa que c é definido como o menor inteiro primo maior do que 4.

Para comparar sequências, usaremos as seguintes notações:

$$a_n \sim b_n \Leftrightarrow \frac{a_n}{b_n} \rightarrow 1 \text{ quando } n \rightarrow \infty.$$

$$a_n \simeq b_n \Leftrightarrow \text{existem } c_-, c_+ > 0 \text{ tais que } c_- \leq a_n/b_n \leq c_+$$

Escreveremos também $A \simeq B$ informalmente, para indicar que A e B são da mesma ordem de grandeza, por exemplo $2^{-8} \simeq 0.004$, $998 \simeq 1000$.

Capítulo 1

O gás de Ehrenfest

Ao abrir um frasco, um gás vaza e se espalha pelo quarto. Qual é a probabilidade das moléculas que compõem o gás voltarem, de repente, para o frasco?

1.1 Introdução

Considere um sistema formado por dois containers, A e B . Estudaremos um gás, formado por partículas monoatômicas idênticas que podem passar de um container para o outro por meio de uma pequena abertura.

Suporemos que a quantidade de gás é fixa (nem aumenta, nem diminui). Suporemos também que não há interações com o exterior e que, inicialmente, o gás está contido em A , com B vazio. Em termos de pressão: $p_A > 0$, $p_B \simeq 0$.

Deixemos o sistema evoluir, regido pelas leis da mecânica e da termodinâmica. Sabemos, pela nossa experiência, que o gás vaza para o container B e que ele atinge, depois de um certo tempo, um estado de *equilíbrio macroscópico*, em que as duas metades tem pressões iguais. Na Figura 1.1 se encontram três imagens de simulações¹ deste sistema.

Essa experiência levanta algumas perguntas:

- O que é exatamente um estado de *equilíbrio*? Por um lado, do ponto de vista *macroscópico*, o equilíbrio é atingido quando as pressões dos dois containers são iguais: $p_A \simeq p_B$.

Por outro lado, *microscopicamente*, o número total de partículas contido nos containers, denotado por N , é fixo, grande, da ordem do *número de Avogadro*:

$$N \simeq 10^{23}.$$

O estado inicial descrito acima é tal que $N_A \simeq N$ e $N_B \simeq 0$ e o equilíbrio corresponde a $N_A \simeq N/2$, $N_B \simeq N/2$. Mesmo se o número de partículas em cada container se aproxima dos valores $N_A \simeq N/2$, $N_B \simeq N/2$, a diferença $N_A - N_B$ *não converge para zero*, mas *flutua* em torno de zero.

¹Todas as simulações descritas neste curso (algumas reproduzidas durante as aulas) foram feitas usando o programa **Algodoo**. De uso muito simples, ele não requer o conhecimento de nenhuma linguagem de programação. Ele se encontra na internet, também na sua versão anterior chamada **Phun**.

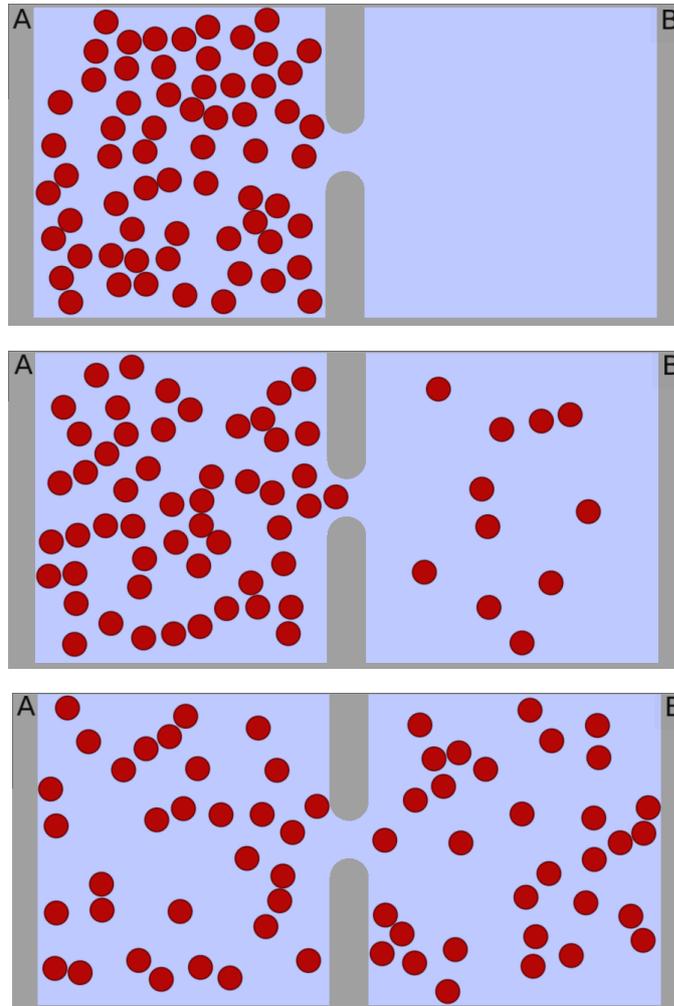


Figura 1.1: Simulação de um gás com $N = 64$ partículas. (1) No tempo inicial, com $N_A = 64$, $N_B = 0$. (2) Alguns instantes depois do tempo inicial. (3) Depois de um tempo longo, em equilíbrio termodinâmico, com $p_A \simeq p_B$.

Logo, vemos que a noção de equilíbrio depende da *escala* na qual o sistema é considerado. Na escala macroscópica, as quantidades que descrevem o gás (pressão, temperatura, etc.) não mudam. Na escala microscópica, há flutuações das grandezas (por exemplo $N_A - N_B$), de natureza aparentemente aleatória.

- Suponhamos que o sistema esteja em equilíbrio. Já que o número de partículas é finito, é possível, mas muito improvável, que em algum instante muito remoto no tempo, cada partícula tenha retornado ao seu container inicial. Qual é a probabilidade disso acontecer? E se acontecer, acontecerá depois de quanto tempo?

Essa questão foi um dos principais dilemas dos físicos do início do século 20, pela seguinte razão. Por um lado, quando o gás atingiu o seu estado de equilíbrio, a Segunda Lei da termodinâmica diz que ele não pode voltar ao seu estado inicial. Por outro lado, as leis da mecânica clássica são *reversíveis com respeito ao tempo*: se todas as velocidades das partículas do gás são invertidas, $v \mapsto -v$, então o gás deve voltar ao seu estado inicial, o que seria uma contradição com a Segunda Lei.

Os problemas levantados nessa discussão são clássicos na área chamada *mecânica estatística*, cujo objetivo é *explicar o comportamento macroscópico a partir das leis microscópicas*. Sem querer apresentar os fundamentos dessa teoria, introduziremos um modelo simples de *evolução estocástica* que permitirá responder a algumas das perguntas acima.

Antes de começar, vejamos porque esse tipo de problema não deve ser estudado com métodos clássicos, isto é, usando as leis da mecânica. Suponha que existam exatamente 10^{23} partículas. Uma descrição clássica requer um conhecimento *exato* da posição e da velocidade de cada partícula. A i -ésima partícula é determinada pela sua posição $\vec{x}^i \in \mathbb{R}^3$ e a sua velocidade $\vec{v}^i \in \mathbb{R}^3$. Se \vec{a}^i denotar a sua aceleração e \vec{F}^i a soma das forças da interação entre i e as outras $10^{23} - 1$ partículas, a equação de Newton para a partícula i (de massa m) se escreve $\vec{F}^i = m\vec{a}^i$. Isto é, para uma condição inicial dada, a evolução das posições $\vec{x}^i = \vec{x}^i(t)$ é solução do sistema

$$\begin{aligned}\vec{F}^1 &= m\vec{a}^1 \\ \vec{F}^2 &= m\vec{a}^2 \\ &\dots \\ \vec{F}^{10^{23}} &= m\vec{a}^{10^{23}}\end{aligned}$$

Assim, seria necessário resolver um sistema não linear de $3 \cdot 10^{23}$ equações, com $3 \cdot 10^{23}$ incógnitas...

1.2 O modelo

Definimos agora um modelo simples, mas complexo o suficiente para apresentar um comportamento interessante (em particular, fornecer uma explicação para o problema da irreversibilidade). Para começar suporemos que:

- o sistema contém um número N , fixo, de partículas *indistinguíveis*;

- nem as posições, nem as velocidades exatas de cada partícula dentro de cada container importam. O que interessa é o *número* de partículas em cada container;
- olharemos o sistema somente nos instantes *discretos* $t = \{0, 1, 2, 3, \dots\}$. Como esses tempos são inteiros, usaremos a letra n em vez de t .

Como o número total de partículas é fixo, o número de partículas numa caixa se deduz imediatamente a partir do outro. Logo, *o estado do sistema será determinado pela variável*

$X_n :=$ número de partículas no container B no instante n .

Como o container B pode conter zero, uma, duas, etc. até no máximo N partículas, temos $X_n \in \mathcal{S}_N$, em que

$$\mathcal{S}_N := \{0, 1, 2, \dots, N\}$$

é o **espaço dos estados**. Observe que o índice “ n ” em X_n é um parâmetro que descreve a variável “tempo”, enquanto N é o número de partículas no sistema, que é fixo.

Suporemos que no tempo $n = 0$, a condição inicial à dada por

$$X_0 = 0,$$

o que significa que todas as partículas estão no container A e que o container B está vazio. Em seguida, é preciso introduzir uma *dinâmica* para a evolução do sistema, isto é dar regras para determinar a evolução do número de partículas no container B em função do tempo,

$$X_0, X_1, X_2, X_3, \dots$$

Suporemos que *a cada instante, uma partícula (e uma só) passa de um container para o outro*. Em termos da variável X_n , isso significa que $X_{n+1} = X_n \pm 1$. Assim, o número X_n pode ser interpretado como a posição de uma *partícula virtual*, que evolui no conjunto \mathcal{S}_N , pulando de um ponto para um dos seus vizinhos (ver Figura (1.2)).

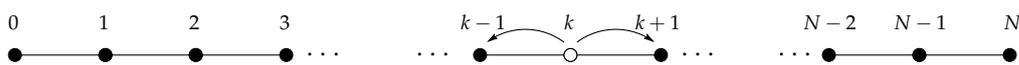


Figura 1.2: A evolução do gás de N partículas se reduz ao estudo de uma única partícula virtual, cujo espaço de estados $\mathcal{S}_N = \{0, 1, 2, \dots, N\}$. Se a imagem representa o estado do sistema no tempo n , isso significa que $X_n = k$: tem k partículas no container B e $N - k$ no A . No tempo $n + 1$, a partícula virtual estará ou em $k - 1$ (uma partícula passou de B para A), ou em $k + 1$ (uma partícula passou de A para B).

Para não ter que olhar de muito perto os detalhes microscópicos das partículas perto da abertura, adotaremos um ponto de vista *probabilístico*, determinando X_{n+1} a partir de X_n , com uma certa probabilidade. Suporemos então que *todas as partículas de um mesmo container tem a mesma probabilidade de passar pela abertura*.

Começemos com o primeiro estado imediatamente após a condição inicial $X_0 = 0$, isto é X_1 . Pelas regras que foram impostas acima, a única coisa que pode acontecer

entre $n = 0$ e $n = 1$ é de uma das N partículas do container A passar para o container B . Isto é: $X_1 = 1$. Esse primeiro passo na evolução será representado como

$$P(0 \rightarrow 1) = 1,$$

o que significa que quando $N_B = 0$, então uma partícula passa de A para B , com probabilidade 1.

Suponha que a partícula virtual se encontre em k no tempo n : $X_n = k$. No tempo $n + 1$ ela se encontrará ou em $k - 1$ ou em $k + 1$. Agora definiremos as probabilidades $P(k \rightarrow k - 1)$, $P(k \rightarrow k + 1)$. Pela nossa hipótese, as partículas do gás são indistinguíveis e equivalentes com respeito à mudança de container. Assim, a partícula que muda de container pode simplesmente ser escolhida aleatoriamente entre as N partículas do sistema. Como a probabilidade de escolher uma partícula que está em B é k/N ,

$$P(k \rightarrow k - 1) = \frac{k}{N}$$

e como a probabilidade de escolher uma partícula que está em A é $(N - k)/N$,

$$P(k \rightarrow k + 1) = \frac{N - k}{N}.$$

Lembramos também os dois casos particulares $k = 0$ e $k = N$, em que

$$P(0 \rightarrow 1) = 1, \quad P(N \rightarrow N - 1) = 1.$$

Assim vemos que o container mais cheio tende a mandar as suas partículas para o outro, o que reflete o fato do sistema querer igualar as pressões dos dois containers.

Os números $P(k \rightarrow k \pm 1)$ são chamados de **probabilidades de transição**. Estas podem ser escritas como as entradas de uma matriz $(N + 1) \times (N + 1)$, denotada Q , cujos elementos $Q_{i,j}$ ($i, j \in \mathcal{S}_N$) são

$$Q_{i,j} := P(i \rightarrow j).$$

A matriz Q é chamada de **matriz de transição**.

$$Q = \begin{pmatrix} 0 & 1 & 0 & \cdots & & \cdots & 0 \\ \frac{1}{N} & 0 & \frac{N-1}{N} & 0 & & & \vdots \\ 0 & \frac{2}{N} & 0 & \frac{N-2}{N} & 0 & & \\ \vdots & 0 & \frac{3}{N} & 0 & \frac{N-3}{N} & 0 & \\ & & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & & 0 & \frac{N-2}{N} & 0 & \frac{2}{N} & 0 \\ \vdots & & & & 0 & \frac{N-1}{N} & 0 & \frac{1}{N} \\ 0 & \cdots & & & \cdots & 0 & 1 & 0 \end{pmatrix} \quad (1)$$

Tirando as duas diagonais vizinhas da diagonal principal, todos os elementos de Q são nulos. Observe que

1. $Q_{i,j} \geq 0$ para todo i, j ;

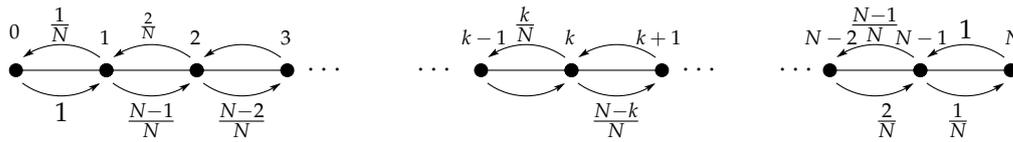


Figura 1.3: A Cadeia de Markov do Modelo de Ehrenfest.

2. $\sum_j Q_{i,j} = 1$ para todo i .

Toda matriz que satisfaz a essas duas propriedades é chamada de **matriz estocástica**.

Dada uma condição inicial X_0 , a sequência X_1, X_2, \dots aleatória é chamada de **cadeia de Markov**² com espaço de estados \mathcal{S}_N e matriz de transição Q . Essa descrição do modelo dos dois containers foi introduzida pelo físico P. Ehrenfest³.

Tentaremos agora entender o comportamento assintótico da sequência aleatória X_n , iniciada em $X_0 = 0$.

1.3 Tempos longos: distribuição invariante

Como já sabemos (e como as simulações mostram), a evolução do sistema com a condição inicial $N_A = N, N_B = 0$, tende a se aproximar de um estado de equilíbrio em que o número de partículas em cada container é $N_A \simeq N_B \simeq N/2$. Uma simulação da cadeia com $N = 500$, $X_0 = 0$, é ilustrada na Figura 1.4. Como provar esse comportamento a partir da cadeia de Markov X_n e da sua matriz de transição Q ?

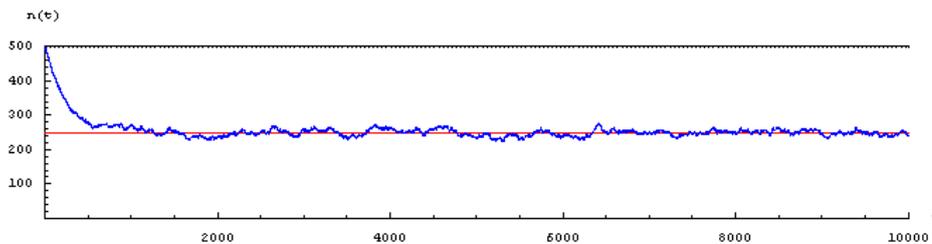


Figura 1.4: Simulação do número de partículas no container B em função do tempo, X_n , com a condição inicial $X_0 = 500$. Observe a convergência rápida para o estado de equilíbrio $N_A \simeq N_B \simeq 250$. Observe também que no equilíbrio, há *flutuações* em torno desse valor.

Podemos representar a condição inicial $X_0 = 0$ por um vetor linha:

$$\mu^{(0)} = (1, 0, \dots, 0, 0),$$

²Andrei Andreievitch Markov, Riazan (Rússia) 1856 - São Petersburgo (Rússia) 1922.

³Paul Ehrenfest: Viena (Austria) 1880 - Amsterdam, 1933.

que significa que no tempo $n = 0$,

$$\mu_k^{(0)} := P(X_0 = k) = \begin{cases} 1, & \text{se } k = 1, \\ 0, & \text{caso contrário.} \end{cases}$$

Ao longo da evolução, a probabilidade contida no vetor $\mu^{(0)}$, concentrada em 0 no tempo 0 vai se espalhar pelo sistema. No tempo $n = 1$,

$$\mu^{(1)} = (0, 1, \dots, 0, 0),$$

mas no tempo $n = 2$,

$$\mu^{(2)} = \left(\frac{1}{N}, 0, \frac{N-1}{N}, 0, \dots, 0, 0 \right),$$

De modo geral, denotamos por

$$\mu^{(n)} = (\mu_0^{(n)}, \dots, \mu_N^{(n)})$$

o vetor chamado **distribuição de probabilidade no tempo n** . As componentes satisfazem $\mu_k^{(n)} \geq 0$ e, por ser uma probabilidade,

$$\sum_{k \in \mathcal{S}_N} \mu_k^{(n)} = 1.$$

A interpretação é a seguinte:

$$\mu_k^{(n)} = P(X_n = k) = P(\text{a partícula virtual está em } k \text{ no tempo } n).$$

Estudaremos agora a sequência $\mu^{(0)}, \mu^{(1)}, \mu^{(2)}, \dots$. Se a partícula virtual estiver em n no tempo k , então no tempo $n - 1$ ela estava ou em $k - 1$ ou em $k + 1$ e depois pulou para k (ver Figura 1.5).

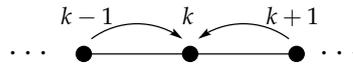


Figura 1.5: Antes de estar em k , a partícula virtual esteve ou em $k - 1$, ou em $k + 1$.

Logo, temos a seguinte identidade:

$$\mu_k^{(n)} = \mu_{k-1}^{(n-1)} P(k-1 \rightarrow k) + \mu_{k+1}^{(n-1)} P(k+1 \rightarrow k),$$

o que equivale a

$$\mu_k^{(n)} = \sum_{j \in \mathcal{S}_N} \mu_j^{(n-1)} Q_{j,k}.$$

Em notação vetorial,

$$\mu^{(n)} = \mu^{(n-1)} Q. \quad (2)$$

Isto é, a distribuição no tempo n se calcula à partir da distribuição no tempo $n - 1$ pelo produto matricial (à direita) da distribuição no tempo $n - 1$ com a matriz de transição. Iterando (2),

$$\mu^{(n)} = \mu^{(n-1)}Q = (\mu^{(n-2)}Q)Q = \mu^{(n-2)}Q^2 = \dots = \mu^{(0)}Q^n,$$

em que Q^n é a n -ésima potência de Q . Logo, o comportamento assintótico da distribuição de probabilidade,

$$\lim_{n \rightarrow \infty} \mu^{(n)}, \quad (3)$$

pode ser deduzido a partir do comportamento de Q^n no limite $n \rightarrow \infty$. Por exemplo, suponha que Q seja diagonalizável: $Q = BDB^{-1}$, em que B é uma matriz ortogonal formada a partir dos autovetores de Q e D a matriz diagonal dos autovalores. Logo,

$$Q^n = (BDB^{-1})(BDB^{-1}) \dots (BDB^{-1})(BDB^{-1}) = BD^nB^{-1},$$

e dependendo dos autovalores de Q , D^n pode possuir um limite ou não. Esse método será implementado em detalhes no Exercício (1.2).

Por enquanto, suponhamos que (3) exista ⁴. Isto é, para cada $k \in \mathcal{S}_N$,

$$\pi_k := \lim_{n \rightarrow \infty} \mu_k^{(n)} \equiv \lim_{n \rightarrow \infty} P(X_n = k).$$

É claro que o vetor π , com componentes π_k , é uma distribuição de probabilidade em \mathcal{S}_N . π deve ser interpretado da seguinte maneira: *no equilíbrio*, a posição da partícula virtual é aleatória, espalhada em \mathcal{S}_N , dada por

$$\pi_k = P(\text{em equilíbrio, a partícula virtual estar em } k). \quad (4)$$

Tentaremos agora calcular π . Para ver como obter π sem passar pelo cálculo do limite quando $n \rightarrow \infty$, observemos que

$$(\pi Q)_k = \left(\lim_{n \rightarrow \infty} \mu^{(n)} Q \right)_k = \left(\lim_{n \rightarrow \infty} \mu^{(n+1)} \right)_k = (\pi)_k = \pi_k.$$

Logo,

$$\pi Q = \pi. \quad (5)$$

Uma distribuição que satisfaz (5) é chamada **distribuição invariante (com respeito a Q)**. Com Q dada em (1), pode ser verificado (Exercício 1.1) que Q possui uma distribuição invariante, dada pela **distribuição de Bernoulli**,

$$\pi_k := \frac{1}{2^N} \binom{N}{k}, \quad k \in \{0, 1, 2, \dots, N\}.$$

Na Figura 1.6 representamos os valores de π_k , $k \in \mathcal{S}_N$, no caso $N = 8$.

É fácil verificar que π_k é simétrica em torno de $N/2$ (se N for par) e que $\pi_k < \pi_{k'}$ para todo $0 \leq k < k' \leq N/2$. Isso já mostra que no equilíbrio, o ponto onde a partícula virtual tem maior probabilidade de estar é perto de $N/2$, o que corresponde a ter o mesmo número de partículas em cada container.

Veremos agora que, com π na mão, dois resultados fundamentais da Teoria das cadeias de Markov permitem obter informações interessantes sobre o gás de Ehrenfest.

⁴Observe que a existência do limite (3) não contradiz o fato da sequência X_n não possuir limite. De fato, a partícula virtual não para de pular para um dos seus vizinhos, então a sua posição não tende a lugar nenhum. Porém, faz sentido dizer que a *probabilidade de ela estar num ponto k* tende a um valor fixo quando $n \rightarrow \infty$.

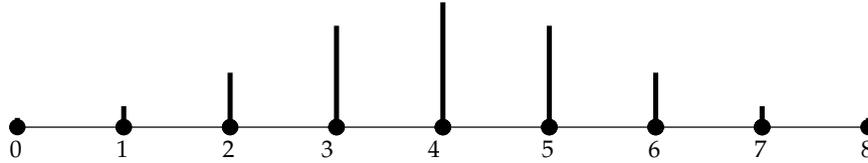


Figura 1.6: A distribuição invariante para o modelo de Ehrenfest com $N = 8$ partículas. Observe o máximo e a simetria em torno de $N/2 = 4$.

1.3.1 Tempo de permanência

Para um estado $k \in \mathcal{S}_N$, considere a **fração de tempo de permanência em k** , definida pelo limite

$$\tau_k := \lim_{m \rightarrow \infty} \frac{\#\{0 < n \leq m : X_n = k\}}{m}.$$

Se τ_k existir, significa que a partícula virtual visitou o ponto k aproximadamente $\tau_k n$ vezes até o tempo n . O seguinte resultado é uma consequência do *Teorema Ergódico para cadeias de Markov*:

Teorema 1.1 *Para qualquer condição inicial, τ_k existe quase certamente para todo $k \in \mathcal{S}_N$ e vale*

$$\tau_k = \pi_k.$$

O teorema mostra que a distribuição invariante representa também, assintoticamente, a *fração de tempo* passado num ponto pela partícula virtual ao longo da sua trajetória.

Observe, por exemplo, que no caso $N = 8$ (ver Figura 1.6),

$$\pi_4 = \frac{1}{2^8} \binom{8}{4} = \frac{1}{2^8} \frac{8!}{4!^2} \simeq 0.273.$$

Pelo Teorema Ergódico, isso significa que, ao longo da evolução, o sistema passa em torno de 25% do seu tempo no estado em que as partículas estão igualmente distribuídas nas caixas ($N_A = N_B = 4$). Por outro lado,

$$\pi_0 = \frac{1}{2^8} \binom{8}{0} = \frac{1}{2^8} \simeq 0.004,$$

o que implica que, no equilíbrio, a fração de tempo que o sistema passa no estado em que todas as partículas estão no container A é menor do que 1%.

Vejamos agora o que acontece quando o número N é grande, usando a **fórmula de Stirling**:

$$N! \sim \sqrt{2\pi N} e^{-N} N^N.$$

Por um lado, a probabilidade de existir exatamente o mesmo número de partículas em cada caixa é dada por (ver Exercício 1.1)

$$\pi_{N/2} = \frac{1}{2^N} \frac{N!}{(N/2)!^2} \sim \sqrt{\frac{2}{\pi N}}, \quad (6)$$

que converge para zero quando $N \rightarrow \infty$: devido às flutuações, ter *exatamente* o mesmo número de partículas em cada caixa é um evento raro. Por outro lado, a probabilidade de todas as partículas estarem na caixa A é

$$\pi_0 = \frac{1}{2^N}.$$

Logo, pelo Teorema Ergódico, a fração de tempo passado no estado 0 é negligível comparada com a fração de tempo passada no estado $N/2$ (quando N é grande).

1.3.2 Tempo de primeiro retorno

Vamos agora voltar ao problema posto no início do capítulo:

Quanto tempo demora para o sistema voltar ao seu estado inicial?

Suponhamos que a cadeia é iniciada em $k \in \mathcal{S}_N$, $X_0 = k$ e considere o **tempo de primeiro retorno para k** :

$$T_k := \inf\{n > 0 : X_n = k\}.$$

Observe que, como o sistema é finito, esse tempo é sempre finito⁵, mas o interessante é conhecer o *valor esperado* de T_k . Um resultado clássico da teoria das cadeias de Markov permite calcular, quando a partícula virtual é iniciada em k , o valor esperado de T_k , denotado $E_k[T_k]$.

Teorema 1.2 Para todo $k \in \mathcal{S}_N$,

$$E_k[T_k] = \frac{1}{\pi_k}. \quad (7)$$

Para apreciar esse resultado, suponhamos que o tempo n seja medido em segundos⁶ e consideremos por exemplo um gás com $N = 1000$ partículas. Iniciando o gás com 500 partículas em cada caixa temos, por (7) e (6),

$$E_{500}[T_{500}] = \frac{1}{\pi_{500}} = \frac{2^{500}}{\binom{1000}{500}} \simeq \sqrt{500\pi} \simeq 40 \text{ segundos.}$$

Mas, iniciando o gás com todas as partículas na caixa A ,

$$E_0[T_0] = \frac{1}{\pi_0} = 2^{1000} \text{ segundos} \simeq 10^{993} \text{ anos, ...}$$

Para dar uma idéia da ordem de grandeza desse número, lembramos que o tempo que se passou desde o Big Bang, está estimado em torno de 1.37×10^{10} anos, e que o número de partículas no universo é da ordem de 10^{80} . A conclusão é que mesmo com somente 1000 partículas, um físico provavelmente nunca viverá tempo o suficiente

⁵De fato, a cada N unidades de tempo, considere o evento em que a partícula virtual anda N passos em direção a sua posição inicial. Como esses eventos são independentes e de probabilidade positiva, um deles acaba acontecendo.

⁶Essa escolha é arbitrária; poderíamos também supor que a unidade de tempo é uma fração de segundo. De qualquer forma, a informação qualitativa fornecida abaixo não depende desta escolha.

para observar o gás retornar ao seu estado inicial. A situação piora se ele considerar um número mais realista de partículas, da ordem de $N \simeq 10^{23}$. Neste caso,

$$E_0[T_0] \simeq 2^{10^{23}} \simeq 10^{100.000.000.000.000.000} \text{ anos.}$$

Esses números dão uma explicação razoável do dilema da irreversibilidade mencionado no início do capítulo: o tempo necessário para o gás retornar ao seu estado inicial é *finito*, mas em escalas de tempo *humanas* a evolução do gás pode ser considerada como *irreversível*.

1.3.3 Sobre a existência do limite $\lim_{n \rightarrow \infty} \mu^{(n)}$

Vamos terminar a nossa discussão sobre a existência do limite $\lim_{n \rightarrow \infty} Q^n$.

Teorema 1.3 (Perron-Frobenius) *Seja Q uma matriz estocástica com a seguinte propriedade: existe um inteiro $m \geq 1$ tal que $(Q^m)_{i,j} > 0$ para todo i, j . Então:*

1. Q possui o autovalor $\lambda^0 = 1$ e qualquer outro autovalor λ' de Q satisfaz $|\lambda'| < \lambda^0$. Além disso, λ_0 é não-degenerado: existe um único autovetor π^0 associado, $\pi^0 Q = \pi^0$, que pode ser escolhido tal que $\sum_i \pi_i^0 = 1$;
2. para qualquer distribuição inicial μ , $\mu Q^n \rightarrow \pi^0$.

Infelizmente, a matriz Q definida em (1) não satisfaz à condição do teorema! Na verdade, será verificado no Exercício 1.2 que o limite $\lim_{n \rightarrow \infty} Q^n$ nem mesmo existe. Mas isso acontece por razões acidentais. De fato, considere a seguinte modificação da matriz Q : para $0 < \epsilon < 1/N$,

$$Q_\epsilon = \begin{pmatrix} 2\epsilon & 1-2\epsilon & 0 & \dots & \dots & 0 \\ \frac{1}{N} - \epsilon & 2\epsilon & \frac{N-1}{N} - \epsilon & 0 & & \vdots \\ 0 & \frac{2}{N} - \epsilon & 2\epsilon & \frac{N-2}{N} - \epsilon & 0 & \\ \vdots & 0 & \frac{3}{N} - \epsilon & 2\epsilon & \frac{N-3}{N} - \epsilon & 0 \\ & & \ddots & \ddots & \ddots & \vdots \\ & & & 0 & \frac{N-2}{N} - \epsilon & 2\epsilon & \frac{2}{N} - \epsilon & 0 \\ \vdots & & & & 0 & \frac{N-1}{N} - \epsilon & 2\epsilon & \frac{1}{N} - \epsilon \\ 0 & \dots & & & \dots & 0 & 1 - \epsilon & 2\epsilon \end{pmatrix}$$

Q_ϵ é uma pequena perturbação de Q , pois $\lim_{\epsilon \rightarrow 0} Q_\epsilon = Q$. A diferença é que se a dinâmica for regida pela matriz Q_ϵ , a partícula virtual tem uma pequena probabilidade de ficar onde ela está antes de pular para um dos seus vizinhos. Em particular, para um par de pontos quaisquer i, j em \mathcal{S}_N ,

$$\begin{aligned} (Q_\epsilon^N)_{i,j} &\geq P(\text{a partícula virtual pula } N - |j - i| \text{ vezes em } i \text{ para depois} \\ &\quad \text{dar } |j - i| \text{ pulos em direção a } j) \\ &\geq (2\epsilon)^{N-|j-i|} (1/N - \epsilon)^{|j-i|} > 0. \end{aligned}$$

Logo, Q_ϵ satisfaz à condição do teorema, com $m = N$. Portanto, existe uma distribuição invariante π_ϵ^0 e $\lim_{n \rightarrow \infty} \mu Q_\epsilon^n = \pi_\epsilon^0$ para qualquer distribuição inicial μ . O Exercício 1.2 propõe provar o Teorema de Perron-Frobenius nesse caso particular.

1.4 Exercícios

Exercício 1.1 Considere a distribuição de Bernoulli π no conjunto \mathcal{S}_N .

1. Mostre que π é uma distribuição de probabilidade, invariante para a matriz Q definida em (1).
2. No equilíbrio, calcule a probabilidade do container B conter a) 0 partículas, b) $N/2$ partículas. Compare a) e b) quando N é grande, usando a Fórmula de Stirling.

Exercício 1.2 Considere Q no caso $N = 2$ e calcule Q^n . O limite $\lim_{n \rightarrow \infty} Q^n$ existe? Em seguida, fixe $0 < \epsilon < 1/2$ e considere a matriz dada por

$$Q_\epsilon := \begin{pmatrix} 0 & 1 & 0 \\ 1/2 - \epsilon & 2\epsilon & 1/2 - \epsilon \\ 0 & 1 & 0 \end{pmatrix}.$$

1. Verifique que Q_ϵ é uma matriz de transição e interprete a cadeia de Markov associada.
2. Ache a distribuição invariante para Q_ϵ e mostre que esta coincide, no limite $\epsilon \rightarrow 0$, com a distribuição π do Exercício 1.1.
3. Para uma distribuição inicial $\mu^{(0)}$ qualquer, mostre sem usar o Teorema 1.3, que o vetor $\mu^{(n)} := \mu^{(0)} Q_\epsilon^n$ converge para π_ϵ no limite $n \rightarrow \infty$. Dica: siga o método sugerido na aula, diagonalizando Q_ϵ .

1.5 Referências bibliográficas

Usamos sem provas (em particular nos Teoremas 1.1, 1.2 e 1.3) vários resultados da Teoria das Cadeias de Markov, que se acham em vários lugares na literatura. Os três primeiros capítulos do pequeno livro do Lawler [8] dão uma excelente introdução aos resultados fundamentais, em particular sobre a convergência da sequência $\mu^{(0)} Q^n$. Os livros de Chung [2] e de Kemeny, Snell e Knapp [7] são mais completos e ainda introdutórios. Um pouco mais avançado, o livro de Grimmet e Stirzaker [4] contém um capítulo sobre cadeias de Markov, com muitas aplicações. Mais difícil e mais abstrato, o livro de Durrett [3]. Uma boa referência para a prova do Teorema de Perron-Frobenius é o livro de Seneta [11].

O modelo de Ehrenfest é, em geral, apresentado como uma aplicação da teoria. Sugiro dar uma olhada na internet para encontrar outras informações ou simulações. Por exemplo, não perca a *Ehrenfest Experiment*, em

<http://www.math.uah.edu/stat/markov/Ehrenfest.shtml>.

Para ler mais sobre grandes números: <http://en.wikipedia.org/wiki/Googol>.

Capítulo 2

O passeio aleatório

Depois de muitas horas passadas num bar, um bêbado resolve voltar para casa. Só que ele não se lembra onde mora e começa a andar aleatoriamente pela cidade: a cada esquina, ele escolhe uma das quatro possíveis direções, sem preferência, e anda até o próximo quarteirão. No caminho, ele tenta calcular a probabilidade de passar em frente à sua casa... (Descrição do passeio aleatório bidimensional, devida a Einstein.)

2.1 Introdução

Consideremos um corpo C imerso em um líquido e suponhamos que a densidade de C seja igual à densidade do líquido, o que faz com que ele não afunde.

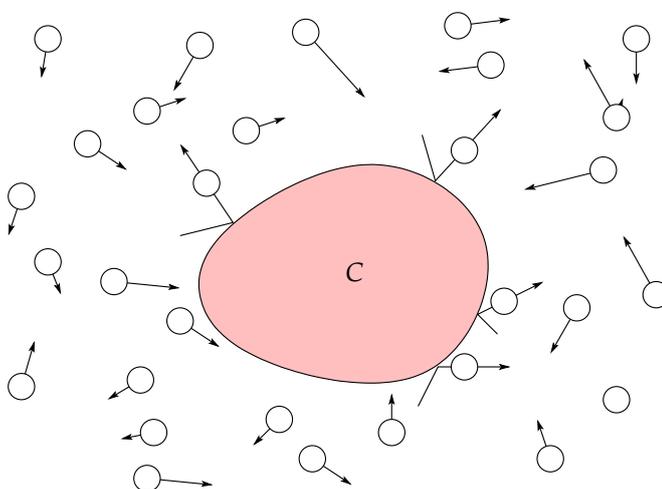


Figura 2.1: Um corpo C imerso, sujeito aos choques das moléculas que compõem o líquido.

Do ponto de vista *macroscópico*, o líquido está em equilíbrio termodinâmico, em repouso, com pressão e temperatura fixa. Do ponto de vista *microscópico*, ele é composto de moléculas, cujas posições e velocidades mudam constantemente com o tempo. Portanto, o objeto está sob influência constante dos choques das moléculas contra a sua superfície, o que pode lhe comunicar alguma energia cinética.

Por um lado, se C for de tamanho muito maior do que o tamanho das moléculas, então os choques são uniformes na superfície do objeto: na média, eles se compensam e o corpo não ganha energia cinética.

Por outro lado, se C for microscópico, porém mais pesado do que as moléculas, então, durante cada intervalo de tempo, a média dos choques não se compensa uniformemente na superfície e ele ganha um impulso. Já que as mudanças na escala molecular são muito mais rápidas do que o movimento do objeto, pode-se supor que o impulso ganho pelo objeto a cada instante é aleatório em direção e intensidade.

Esse movimento aleatório foi observado pela primeira vez em 1827 pelo botanista R. Brown ¹. Brown observou partículas de pólen suspensas em um líquido. As partículas eram de massa bem maior do que as moléculas do líquido, mas leves o suficiente para sentir os choques. As conclusões que ele tirou dessas observações não eram modestas: as moléculas existem!

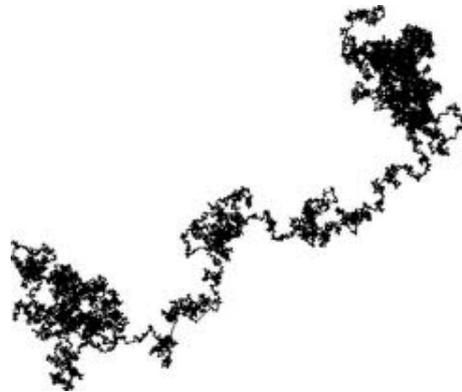


Figura 2.2: O corpo C , visto de longe, segue uma trajetória irregular chamada movimento Browniano.

Einstein ² foi um dos primeiros físicos a elaborar uma teoria matemática do movimento Browniano.

Para simplificar, consideremos a evolução aleatória da partícula de pólen (o corpo C) na reta. Seja $p(t, x)$ a densidade de probabilidade descrevendo a trajetória da partícula: $p(t, x)\delta x$ é a probabilidade da partícula estar num intervalo de tamanho δx , centrado em x , no tempo t . Einstein mostrou que $p(t, x)$ obedece à seguinte equação diferencial, chamada *equação do calor*:

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2}. \quad (1)$$

$D > 0$ é o coeficiente de difusão, que depende das características do líquido e do grão de pólen. Com uma condição inicial em que a partícula se encontra na origem, a solução de (1) é dada por

$$p(t, x) = \frac{1}{\sqrt{2\pi Dt}} e^{-\frac{x^2}{Dt}}. \quad (2)$$

¹Robert Brown: Montrose (Escócia), 1773 - Soho (Escócia), 1858.

²Albert Einstein: Ulm (Alemanha), 1879- Princeton (Estados Unidos), 1955.

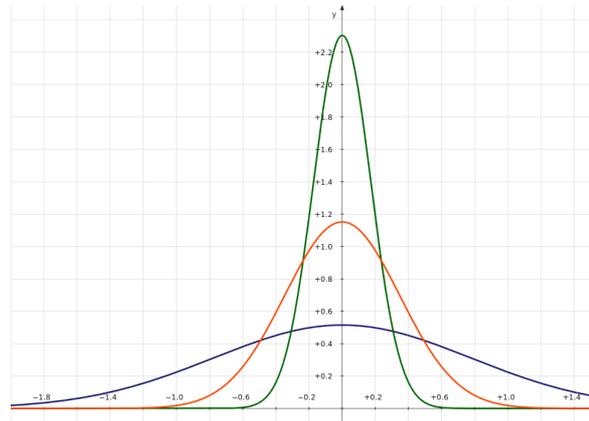


Figura 2.3: A solução (2) da equação do calor, com coeficiente $D = 3$, para os tempos $t = 0.004$, $t = 0.01$ e $t = 0.2$. O fato da distribuição se espalhar ao longo da evolução explica porque esse processo é chamado de *difusivo*.

Em 1923, N. Wiener ³ apresentou a primeira construção rigorosa do movimento Browniano como um processo estocástico. Na construção, a partícula de pólen é modelada como um ponto evoluindo em \mathbb{R}^d , cuja trajetória é parametrizada por uma função

$$t \mapsto X_t, \quad t \in [0, \infty), X_t \in \mathbb{R}^d.$$

Os choques devidos às moléculas do líquido ambiente são introduzidos associando uma probabilidade a cada trajetória $t \mapsto X_t$. Por ser um objeto *contínuo, em tempo contínuo*, a descrição probabilística de uma trajetória aleatória no espaço \mathbb{R}^d requer ferramentas de probabilidade que estão bem além deste curso.

Para simplificar, estudaremos uma versão *discreta* do movimento Browniano, o *passo aleatório*, cujas características principais são parecidas, mas cujo estudo pode ser feito usando ferramentas elementares de combinatória.

2.2 O modelo

Em vez da partícula evoluir em \mathbb{R}^d em tempo contínuo, suponhamos que as suas coordenadas só podem tomar valores inteiros e que o tempo é discreto. Isto é, as trajetórias serão modeladas por

$$n \mapsto X_n, \quad n \in \{0, 1, 2, \dots\}, X_n \in \mathcal{S},$$

em que \mathcal{S} é a **rede**

$$\mathcal{S} \equiv \mathbb{Z}^d := \{x = (x_1, \dots, x_d) : x_i \in \mathbb{Z}\}.$$

Como no capítulo anterior, aleatoriedade é introduzida na dinâmica da partícula por meio das probabilidades de transição de um ponto x para um outro ponto y . Por razões de simetria, suporemos que nenhuma direção (da rede) é privilegiada.

³Norbert Wiener, Columbia (Estados Unidos) 1894 - Stockholm (Suécia), 1964.

Por exemplo, no caso unidimensional, $d = 1$, suporemos que a cada instante, a partícula pode pular para um dos seus dois vizinhos com probabilidade $1/2$ (ver Figura 2.4):

$$P(x \rightarrow x \pm 1) = 1/2.$$

(Observe que nesse caso, a evolução é a mesma que a da partícula virtual do gás de Ehrenfest perto de $N/2$, quando N é grande.)

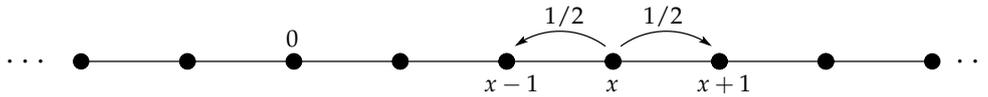


Figura 2.4: O passeio aleatório simétrico em dimensão $d = 1$.

Em dimensões superiores, a definição é análoga: a cada passo, a partícula pode pular para qualquer um dos seus $2d$ vizinhos, com probabilidade $1/2d$. Isto é, se $\mathbf{e}_1, \dots, \mathbf{e}_d$ denotam os vetores da base canônica de \mathbb{R}^d ,

$$P(x \rightarrow x \pm \mathbf{e}_j) = \frac{1}{2d}, \quad \forall j = 1, \dots, d.$$

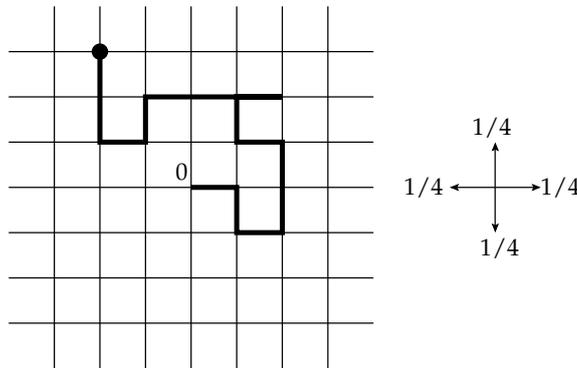


Figura 2.5: O passeio aleatório simétrico em dimensão $d = 2$. Vista de longe (para tempos longos), a trajetória é parecida com uma trajetória do movimento Browniano (Figura 2.2).

Suporemos sempre que a partícula começa na origem no tempo $n = 0$, $X_0 = 0$ e que a cada instante, ela pula para um dos seus próximos vizinhos com as probabilidades de transição dadas acima. O processo estocástico assim definido é uma cadeia de Markov no conjunto S , chamada de **passeio aleatório simples simétrico**.

Por ter um espaço de estados de cardinalidade infinita ($|\mathcal{S}| = \infty$), o passeio aleatório levanta perguntas novas. A principal pergunta que vai nos interessar é:

Qual é a probabilidade do passeio voltar para o seu ponto de partida?

O passeio aleatório à uma cadeia de Markov, mas por sua matriz estocástica ser infinita, o Teorema de Perron-Frobenius não se aplica; outras técnicas devem ser desenvolvidas ⁴.

2.3 Recorrência/Transiência

Se o passeio volta para a origem, então o tempo de seu primeiro retorno após $n = 0$,

$$T_0 := \inf\{n > 0 : X_n = 0\},$$

é finito. Mas uma trajetória pode também nunca voltar para a origem, nesse caso definimos

$$T_0 := \infty.$$

Estamos interessados nas probabilidades dos eventos $\{T_0 < \infty\}$ e $\{T_0 = \infty\}$. Por ser uma probabilidade, $P(T_0 < \infty) \leq 1$.

Se $P(T_0 < \infty) = 1$, o passeio é dito **recorrente**; se $P(T_0 < \infty) < 1$, ele é **transiente**.

Se o passeio for recorrente, ele volta para a origem pelo menos uma vez, quase certamente. Se ele for transiente, ele pode sair do seu ponto de partida para nunca voltar; a probabilidade $P(T_0 = \infty)$ chama-se **probabilidade de escapar**. No restante do capítulo daremos uma prova do seguinte resultado:

Teorema 2.1 *O passeio é recorrente em dimensões $d = 1, 2$ e transiente em dimensões $d \geq 3$.*

Para estudar $P(T_0 < \infty)$ usaremos um método clássico de combinatória. Começaremos com a seguinte observação: o evento $\{T_0 < \infty\}$ pode ser decomposto com respeito aos valores possíveis de T_0 :

$$P(T_0 < \infty) = \sum_{k \geq 0} P(T_0 = k). \quad (3)$$

Considere a sequência de coeficientes

$$a_k := P(T_0 = k), k \geq 0,$$

(observe que $a_0 = 0$) e a sua **função geradora**, definida pela série

$$\mathbb{A}(s) := \sum_{k \geq 0} a_k s^k.$$

Como $0 \leq a_k \leq 1$, a série converge para todo $-1 < s < 1$. Logo, o seu raio de convergência é maior do que ou igual a 1. Podemos então expressar (3) como

$$\begin{aligned} P(T_0 < \infty) &= \sum_{k \geq 0} a_k = \sum_{k \geq 0} a_k 1^k = \sum_{k \geq 0} \lim_{s \nearrow 1} a_k s^k \\ &= \lim_{s \nearrow 1} \sum_{k \geq 0} a_k s^k \equiv \lim_{s \nearrow 1} \mathbb{A}(s). \end{aligned} \quad (4)$$

⁴O leitor interessado pode procurar a referência [10] para ter uma idéia do que pode ser feito sobre esse modelo.

(A justificativa da troca do limite $\lim_{s \nearrow 1^-}$ com o somatório será feita no Exercício 2.3). Portanto, é preciso saber se $\lim_{s \nearrow 1^-} \mathbb{A}(s)$ é igual a 1 ou menor do que 1. Para estudar \mathbb{A} na vizinhança do ponto $s = 1$, consideremos a sequência

$$b_n := P(X_n = 0), n \geq 0.$$

Observe que, pela condição inicial, $b_0 = 1$. A função geradora da sequência b_n é definida por

$$\mathbb{B}(s) := \sum_{n \geq 0} b_n s^n = 1 + \sum_{n \geq 1} b_n s^n.$$

Observe que o raio de convergência da série também é maior do que ou igual a 1. Além disso, as sequências a_n e b_n são ligadas por uma **equação de renovação**:

Lema 2.2 Para todo $n \geq 1$,

$$b_n = \sum_{m=0}^n b_m a_{n-m}. \quad (5)$$

Demonstração: De fato, o evento $\{X_n = 0\}$ que define b_n pode ser decomposto com respeito ao tempo do último retorno na origem, T_* , cujos valores possíveis são $1, 2, \dots, n-1$ (ver Figura 2.6). Suponhamos que esse tempo seja m . Observe que, no

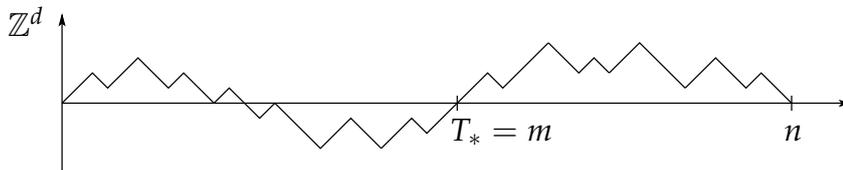


Figura 2.6: A propriedade de renovação.

tempo m , a partícula está na origem e que a sua evolução recomeça como se fosse a partir do tempo inicial (ela esquece tudo que aconteceu até m). Mas por m ser o tempo da sua *última* visita na origem antes do tempo n , n pode ser visto como o tempo de primeira visita em 0 depois do tempo m . As seguintes linhas resumem essa discussão:

$$\begin{aligned} P(X_n = 0) &= \sum_{m=1}^{n-1} P(X_n = 0, T_* = m) \\ &= \sum_{m=1}^{n-1} P(X_{n-m} = 0) P(T_* = m) \\ &= \sum_{m=1}^{n-1} b_m a_{n-m}. \end{aligned}$$

□

Multiplicando (5) por s^n (para $0 < s < 1$) e somando no índice n ,

$$\begin{aligned}\mathbb{B}(s) &= 1 + \sum_{n \geq 1} \left(\sum_{m=0}^n b_m a_{n-m} \right) s^n \\ &= 1 + \sum_{m \geq 0} b_m \left(\sum_{n \geq m} a_{n-m} s^n \right)\end{aligned}\quad (6)$$

$$\begin{aligned}&= 1 + \sum_{m \geq 0} b_m \left(s^m \sum_{n' \geq 0} a_{n'} s^{n'} \right) \\ &\equiv 1 + \mathbb{B}(s) \mathbb{A}(s).\end{aligned}\quad (7)$$

Em (6) trocamos os somatórios, pois todos os termos somados são positivos. Em (7) fizemos a mudança de índice $n' := n - m$.

Assim foi provada uma identidade simples entre \mathbb{A} e \mathbb{B} :

$$\mathbb{B}(s) = \frac{1}{1 - \mathbb{A}(s)}.$$

Logo, o limite $\lim_{s \nearrow 1^-} \mathbb{A}(s)$ em (4) pode ser estudado por meio do limite $\lim_{s \nearrow 1^-} \mathbb{B}(s)$. Mas, pela definição de $\mathbb{B}(s)$,

$$\lim_{s \nearrow 1^-} \mathbb{B}(s) = \sum_{n \geq 0} P(X_n = 0).$$

Provamos até agora o seguinte critério:

Proposição 2.3

$$\begin{aligned}O \text{ passeio é recorrente} &\iff \sum_{n \geq 0} P(X_n = 0) = \infty. \\ O \text{ passeio é transiente} &\iff \sum_{n \geq 0} P(X_n = 0) < \infty.\end{aligned}$$

Logo, o que sobra é estudar a convergência da série $\sum_n P(X_n = 0)$. Para isto veremos a vantagem de trabalhar com \mathbb{B} (ao invés de \mathbb{A}): o comportamento assintótico da sequência $P(X_n = 0)$ pode ser determinado usando métodos combinatórios elementares.

Primeiro, é claro que $P(X_n = 0) > 0$ se, e somente se, n é par. Logo estudaremos $P(X_{2j} = 0)$ para $j \in \mathbb{N}$. Até o tempo $2j$, o passeio seguiu uma certa trajetória, e o evento $\{X_{2j} = 0\}$ pode ser visto como o conjunto das trajetórias (ou *caminhos*) de tamanho $2j$ que começam na origem no tempo 0 e terminam na origem no tempo $2j$ (uma trajetória com essa propriedade está representada na Figura 2.6). Pela definição da probabilidade,

$$P(X_{2j} = 0) = \frac{\#\{\text{caminhos de tam. } 2j \text{ que terminam em } 0 \text{ no tempo } 2j\}}{\#\{\text{caminhos de tam. } 2j\}}. \quad (8)$$

Consideremos inicialmente o caso da dimensão $d = 1$, mais fácil. Calcularemos a fração em (8) pensando da seguinte maneira: um caminho unidimensional pode ser identificado com a sequência dos seus passos, isto é uma sequência de $2j$ símbolos:

$$\leftarrow, \rightarrow, \rightarrow, \leftarrow, \rightarrow, \leftarrow, \leftarrow, \dots, \leftarrow, \rightarrow, \leftarrow, \leftarrow,$$

em que \leftarrow significa que o passo é para a esquerda e \rightarrow que o passo é para a direita. Assim, o denominador é igual ao número total de tais sequências, que é igual a 2^{2j} (cada símbolo pode ser ou \leftarrow , ou \rightarrow):

$$\#\{\text{caminhos de tam. } 2j\} = 2^{2j}.$$

Por outro lado, para um caminho estar de volta em 0 no tempo $2j$, uma única restrição deve ser satisfeita: o número de passos feitos para a esquerda, $n(\leftarrow)$, deve ser igual ao número de passos feitos para a direita, $n(\rightarrow)$. Como $n(\leftarrow) + n(\rightarrow) = 2j$ obtemos $n(\leftarrow) = n(\rightarrow) = j$. Logo, o numerador em (8) se calcula contando o número de jeitos de colocar j símbolos \leftarrow dentro de $2j$ caixas, que é $\binom{2j}{j}$. Logo,

$$\#\{\text{caminhos de tam. } 2j \text{ que terminam em } 0 \text{ no tempo } 2j\} = \binom{2j}{j}$$

Portanto,

$$P(X_{2j} = 0) = \frac{1}{2^{2j}} \binom{2j}{j} = \frac{1}{2^{2j}} \frac{(2j)!}{j!^2}$$

e para determinar se o passeio unidimensional é recorrente ou transiente, precisamos estudar a série

$$\sum_j \frac{1}{2^{2j}} \frac{(2j)!}{j!^2}.$$

Usando a Fórmula de Stirling,

$$\frac{1}{2^{2j}} \frac{(2j)!}{j!^2} \sim \frac{1}{2^{2j}} \frac{\sqrt{2\pi 2j} (2j)^{2j} e^{-2j}}{(\sqrt{2\pi j} j e^{-j})^2} = \sqrt{\frac{2}{\pi j}}. \quad (9)$$

Logo,

$$\sum_n P(X_n = 0) = \sum_{j \geq 1} P(S_{2j} = 0) \simeq \sqrt{\frac{2}{\pi}} \sum_{j \geq 1} \frac{1}{\sqrt{j}} = +\infty.$$

Usando a Proposição 2.3, está provada a primeira afirmação do Teorema 2.1: em dimensão 1, o passeio aleatório simples simétrico é recorrente. Observe que se a probabilidade do passeio voltar para a origem é 1, então pode ser mostrado (e é bastante fácil de acreditar) que a probabilidade dele voltar um número infinito de vezes é 1 também. Na verdade, pode ser mostrado que o passeio unidimensional visita qualquer ponto de \mathbb{Z} um número infinito de vezes.

Esboçemos o argumento no caso da dimensão $d = 2$, em que a combinatória é parecida. Os detalhes e uma prova alternativa serão vistos no Exercício 2.1.

Em dimensão 2, para a partícula estar na origem no tempo $2j$, é necessário ela ter feito um número de passos para cima igual ao número de passos para baixo, $n(\uparrow) = n(\downarrow) = k$, e um número de passos para esquerda igual ao número de passos para direita, $n(\leftarrow) = n(\rightarrow) = i$, com $2k + 2i = 2j$. Somando sobre os valores possíveis de $n(\uparrow)$,

$$P(X_{2j} = 0) = \frac{1}{4^{2j}} \sum_{k=0}^j \binom{2j}{k} \binom{2j-k}{k} \binom{2j-2k}{j-k} \simeq \frac{2}{\pi j}.$$

Portanto, $\sum_j P(X_{2j} = 0) = \infty$: o passeio é recorrente também em dimensão 2. Como no caso unidimensional, pode ser mostrado que a trajetória visita qualquer ponto de \mathbb{Z}^2 um número infinito de vezes (com probabilidade 1).

Em dimensão $d = 3$, o método é o mesmo mas a conclusão é outra. Por uma conta de combinatória parecida, pode ser mostrado que, para j grande,

$$P(X_{2j} = 0) \simeq \frac{1}{j^{3/2}}.$$

Logo, $\sum_j P(X_{2j} = 0)$ converge e o passeio é transiente: há uma probabilidade positiva dele não voltar para a origem (de *escapar*).

Em dimensões maiores, $d \geq 4$, a combinatória se complica e, em geral, são métodos analíticos que permitem mostrar que

$$P(X_{2j} = 0) \simeq \frac{1}{j^{d/2}}. \quad (10)$$

Concluiremos com a seguinte citação do matemático japonês Kakutani ⁵:

A drunk man will find his way home, but a drunk bird may get lost forever.

2.4 Exercícios

Exercício 2.1 Mostre que o passeio aleatório simétrico em dimensão 2 é recorrente, completando o esboço dado acima.

Exercício 2.2 Considere o passeio aleatório em \mathbb{Z} , com probabilidades de transição $p := P(x \rightarrow x + 1)$ e $q := P(x \rightarrow x - 1) = 1 - p$. Mostre que se esse passeio é assimétrico, isto é se $p \neq q$, então ele é transiente.

Exercício 2.3 Mostre que se $a_n(s) \geq 0$ e é não-decrescente em s para todo $n \geq 0$, e se $\lim_{s \nearrow s_0} a_n(s) = a_n(s_0)$ para todo $n \geq 0$, então

$$\lim_{s \nearrow s_0} \sum_{n \geq 1} a_n(s) = \sum_{n \geq 0} a_n(s_0).$$

2.5 Referências bibliográficas

Existem inúmeros textos, resultados e técnicas sobre o passeio aleatório e as suas generalizações. Um dos mais acessível é o texto de Lawler [8]. Uma referência padrão é o livro antigo do Spitzer [12]. O estudo de $P(X_{2j} = 0)$ em dimensão 3 está bem feito em [3]. Mais recente e muito detalhado (em particular para (10) em dimensão qualquer): o livro de Lawler e Limic [9]. Difícil, mas fascinante, é o livro de Révész [10].

⁵Shizuo Kakutani, Osaka (Japão), 1911– New Haven (EUA), 2004

Capítulo 3

O processo de ramificação

*O Senhor Tal*¹ estava preocupado com a sobrevivência de seu sobrenome. A curiosidade o levou a se fazer a seguinte pergunta: qual é a probabilidade do meu sobrenome sobreviver em gerações futuras?

3.1 Introdução

Suponhamos que o Senhor Tal seja a única pessoa sobrando no mundo com esse sobrenome e que ele queira garantir que muitas (se for possível: *infinitas*) gerações futuras tenham o privilégio de herdar seu sobrenome. Para garantir isso, ele precisa ter pelo menos um filho, se possível muitos, e que estes também tenham filhos quando adultos, etc., para garantir que tenha sempre pelo menos um descendente com o sobrenome “Tal”. O perigo é que a vontade de ter filhos não é necessariamente a mesma para cada indivíduo. Assim, corre o risco de haver, talvez num futuro muito remoto, uma geração composta essencialmente de descendentes que não têm interesse em criar filhos. Com uma geração dessa, a sobrevivência do sobrenome “Tal” estaria em perigo.

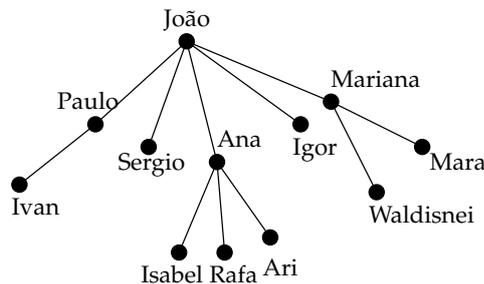


Figura 3.1: Os descendentes vivos do Senhor Tal. Na segunda geração, a sobrevivência do sobrenome depende inteiramente de Ivan, Isabel, Rafa, Ari, Waldisnei e Mara. Se pelo menos um desses netos tiver um filho, a sobrevivência do sobrenome é garantida para mais uma geração. Caso contrário, haverá extinção.

¹Mikhail Nekhemievich Tal, Riga (Letonia) 1936- Moscou (Russia) 1992.

Neste capítulo descreveremos um modelo, introduzido por Galton e Watson ², em que esse problema pode ser estudado de maneira qualitativa.

3.2 O modelo de Galton-Watson

A primeira hipótese principal é que a população considerada é composta de seres assexuados (!), chamados de indivíduos. De fato, o importante para o nosso problema é saber *quantos* filhos cada indivíduo produz e não se estes são homens ou mulheres. Essa hipótese também permite evitar o problema das mulheres que utilizam o nome do marido. Suporemos também que cada indivíduo é filho de exatamente um ancestral.

A segunda hipótese, é que o número de filhos que um indivíduo pode ter ao longo da sua vida é uma quantidade *aleatória* e que *todos os indivíduos (não importa de qual geração) tem um número de filhos igual a k com a mesma probabilidade*

$$p_k := P(\text{um indivíduo ter } k \text{ filhos}) \in [0, 1].$$

Por ser uma distribuição de probabilidade,

$$\sum_{k \geq 0} p_k = 1.$$

Suporemos também que os indivíduos tem filhos de modo *independente*. Isto é: o número de filhos de um não influe nos números de filhos dos outros.

A escolha da sequência p_k depende do tipo de população considerado e será mantida *fixa*. Uma escolha possível para p_k é por exemplo

p_0	p_1	p_2	p_3	p_4	$p_k, k \geq 5$
0.5	0.2	0.2	0.05	0.05	0

Uma escolha mais realista, em que um indivíduo pode ter um número arbitrário de filhos, é, por exemplo $p_k := 2^{-k-1}$. Isto, é um indivíduo tem zero filhos com probabilidade $p_0 = 0.5$, 1 filho com probabilidade $p_1 = 0.25$, etc. (ver o Exercício 3.3).

Seja Z_n o número de indivíduos na n -ésima geração. Suporemos que

$$Z_0 = 1,$$

o que significa que a população começa com um único indivíduo “original”: o Senhor Tal. O número Z_n é aleatório com valores $\{0, 1, 2, \dots\}$. Por exemplo, na Figura 3.1 temos $Z_0 = 1$, $Z_1 = 5$ e $Z_2 = 6$.

Denotaremos por $X_j^{(n)}$ o número de filhos do j -ésimo indivíduo da n -ésima geração. (Em particular, $Z_1 = X_1^{(0)}$ é o número de filhos do único indivíduo original.) Como dito acima, estamos supondo que

$$P(X_j^{(n)} = k) = p_k, \quad \forall k \in \{0, 1, 2, \dots\}$$

²Sir Francis Galton, Birmingham (Inglaterra) 1822 - Haslemere (Inglaterra) 1911. Henri William Watson, Londres, 1827 - Berkswell (Inglaterra), 1903. Originalmente, Galton e Watson estavam interessados em estudar a sobrevivência dos sobrenomes da aristocracia Inglesa do século XIX. O trabalho deles levou a uma publicação em 1874, intitulada *On the probability of extinction of families*.

Observe que o número de indivíduos da $(n + 1)$ -ésima geração é obtido somando o número de filhos da n -ésima geração:

$$Z_{n+1} = X_1^{(n)} + X_2^{(n)} + \dots + X_{Z_n}^{(n)}. \tag{1}$$

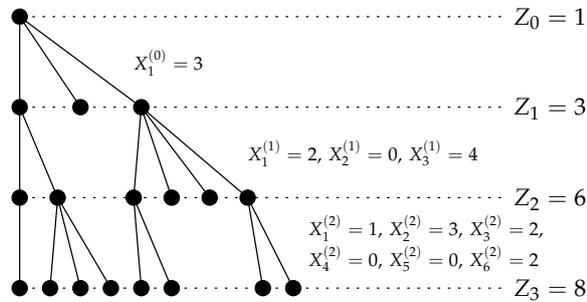


Figura 3.2: As três primeiras gerações de um processo de ramificação.

Observação 3.1 Por razões óbvias, a sequência aleatória Z_n é também chamada de **processo de ramificação**³. O processo de ramificação tem um papel central na Teoria da Probabilidade. Por exemplo, ele é usado para modelar grafos aleatórios, reações em cadeia, processos de crescimento, filas de espera, etc.

É claro que se $Z_n = 0$ para algum n , então $Z_{n'} = 0$ para todo $n' > n$. Nesse caso diremos que há **extinção**. Estamos interessados em calcular a **probabilidade de extinção**, definida por

$$P(\text{extinção}) := P(\text{existe } n \text{ tal que } Z_n = 0).$$

O evento complementar da extinção é o evento em que $Z_n > 0$ para todo n (todas as gerações futuras têm pelo menos um filho) e será chamado de **sobrevivência**. A probabilidade de haver sobrevivência será calculada da seguinte maneira:

$$P(\text{sobrevivência}) = 1 - P(\text{extinção}).$$

Para tornar o problema interessante, suporemos que

$$0 < p_0 < 1, \tag{2}$$

o que implica que há sempre a possibilidade da população morrer já na primeira geração. Observe que se $p_0 = 0$, então qualquer indivíduo tem pelo menos um filho e o problema da sobrevivência é trivial. Se (2) for satisfeita, então a probabilidade de ter extinção é sempre positiva, pois

$$P(\text{extinção}) \geq P(\text{o Senhor Tal não ter filho}) = p_0 > 0.$$

³Em inglês: *branching process*.

Portanto, se trata de saber se $P(\text{extinção})$ é igual a 1 ou estritamente menor do que 1. Temos as seguintes relações:

$$P(\text{extinção}) \begin{cases} = 1 \\ < 1 \end{cases} \iff P(\text{sobrevivência}) \begin{cases} = 0 \\ > 0 \end{cases}. \quad (3)$$

Intuitivamente, uma população tem chance de sobreviver se, *em média*, o número de filhos por indivíduo é suficientemente grande. Logo, um parâmetro natural de se considerar é o **valor esperado do número de filhos por indivíduo**,

$$m := E[Z_1] = \sum_{k \geq 0} k p_k.$$

Se m for pequeno (como em alguns países da Europa), as chances da população sobreviver são pequenas; se m é grande (como em boa parte dos países da África ou da América do Sul), então as chances de sobreviver são grandes. Parece também razoável admitir que, para uma população sobreviver, cada indivíduo precisa ter, na média, pelo menos um filho. Logo, o valor $m = 1$ parece ter um papel importante. O seguinte teorema justifica essa discussão:

Teorema 3.2 *Suponha (2) satisfeita e seja $m := E[Z_1]$. Então*

$$\begin{aligned} m > 1 &\Rightarrow P(\text{extinção}) < 1, \\ m \leq 1 &\Rightarrow P(\text{extinção}) = 1. \end{aligned}$$

Como consequência do teorema e de (3) temos

$$\begin{aligned} m > 1 &\Rightarrow P(\text{sobrevivência}) > 0, \\ m \leq 1 &\Rightarrow P(\text{sobrevivência}) = 0, \end{aligned}$$

o que resolve o problema considerado na introdução.

O modelo é chamado de **subcrítico** (resp. **crítico**, **supercrítico**) se $m < 1$ (resp. $m = 1$, $m > 1$). Observe que o teorema acima diz também que um modelo crítico $m = 1$ sofre extinção com probabilidade 1. Na Figura 3.3 está representada uma simulação⁴ do processo de ramificação **supercrítico** com $m = 1.2 > 1$, em um caso em que cada indivíduo pode ter no máximo dois filhos, $k = 0, 1, 2$.

3.3 Prova do Teorema

Pode ser mostrado que a sequência Z_n é uma cadeia de Markov no conjunto $\mathcal{S}_+ := \{0, 1, 2, \dots\}$, mas com uma matriz de transição infinita difícil de estudar. Em vez disso, a prova do Teorema 3.2 será baseada no uso da **função geradora da variável** Z_n , definida por

$$G_n(s) := E[s^{Z_n}] = \sum_{k \geq 0} P(Z_n = k) s^k.$$

⁴Agradeço ao Prof. Niels Berglund (Université d'Orléans) por me deixar usar essa imagem.

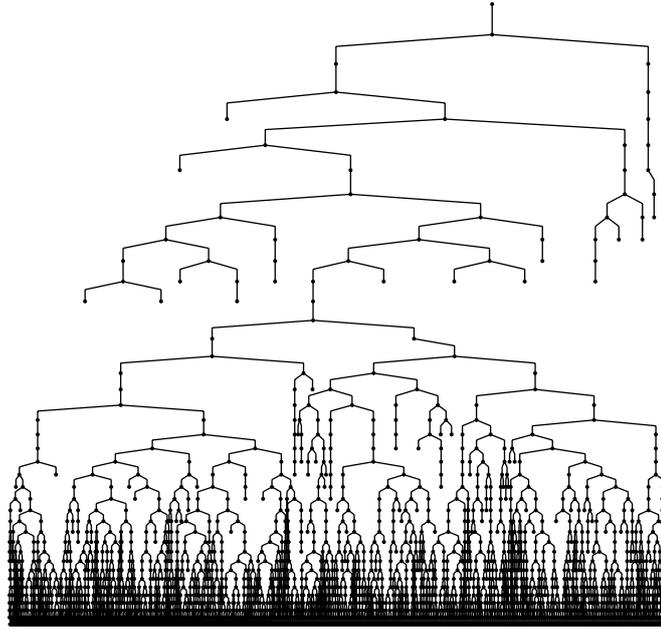


Figura 3.3: Um processo de ramificação supercrítico.

Observe que a série acima converge absolutamente para todo $|s| \geq 1$, pois

$$\sum_{k \geq 0} P(Z_n = k) |s|^k \leq \sum_{k \geq 0} P(Z_n = k) = P(Z_n \in \{0, 1, 2, \dots\}) = 1 < \infty. \quad (4)$$

Logo, $G_n(s)$ está bem definida para todo $s \in [-1, 1]$. Sabemos, dos cursos de cálculo, que $G(s)$ também define uma função analítica da variável $s \in (-1, 1)$ e que a sua derivada neste intervalo é dada por

$$G'_n(s) = \sum_{k \geq 1} k s^{k-1} P(Z_n = k). \quad (5)$$

Em particular,

$$\lim_{s \nearrow 1} G'_n(s) = \sum_{k \geq 1} \lim_{s \nearrow 1} k P(Z_n = k) s^{k-1} = \sum_{k \geq 1} P(Z_n = k) k \equiv E[Z_n]. \quad (6)$$

Assim, propriedades analíticas de G_n fornecem informações probabilísticas sobre Z_n . A segunda derivada se calcula também termo a termo: para $s \in (0, 1)$,

$$G''_n(s) = \sum_{k \geq 2} k(k-1) s^{k-2} P(Z_n = k) \geq 0. \quad (7)$$

Como consequência, G_n é convexa em $(0, 1)$.

Para $n = 1$ escreveremos simplesmente $G_1 \equiv G$. Como Z_1 é o número de filhos do indivíduo original, $P(Z_1 = k) \equiv p_k$, temos

$$G(s) = \sum_{k \geq 0} p_k s^k.$$

Em particular,

$$G_n(0) = P(Z_n = 0). \quad (8)$$

Mesmo não sendo possível, em geral, calcular a função G_n explicitamente, essa satisfaz uma propriedade de recorrência interessante:

Lema 3.3 Para todo $n \geq 1$ e todo $0 \leq s < 1$,

$$\begin{aligned} G_{n+1}(s) &= G_n(G(s)) \\ &= G(G_n(s)). \end{aligned} \quad (9)$$

Demonstração: Observe primeiro que por (4), $0 \leq s \leq 1$ implica $0 \leq G_n(s) \leq 1$ para todo n . Logo, o lado direito das identidades (9) e (10) está bem definido. Por definição,

$$G_{n+1}(s) = E[s^{Z_{n+1}}] = \sum_{k \geq 0} P(Z_{n+1} = k) s^k. \quad (11)$$

Lembramos de (1) que Z_{n+1} é obtida a partir de Z_n e dos filhos da n -ésima geração:

$$Z_{n+1} = X_1^{(n)} + X_2^{(n)} + \cdots + X_{Z_n}^{(n)}.$$

O que torna o estudo de Z_{n+1} difícil é que essa é uma soma de variáveis aleatórias, cujo número de termos é aleatório. Logo, é natural estudar Z_{n+1} separadamente para os valores possíveis de Z_n . Façamos então a seguinte decomposição (ver (1) no Apêndice A):

$$\begin{aligned} P(Z_{n+1} = k) &= \sum_{j \geq 0} P(Z_{n+1} = k, Z_n = j) \\ &= \sum_{j \geq 0} P(X_1^{(n)} + \cdots + X_j^{(n)} = k, Z_n = j) \\ &= \sum_{j \geq 0} P(X_1^{(n)} + \cdots + X_j^{(n)} = k) P(Z_n = j). \end{aligned} \quad (12)$$

Na última linha, usamos a independência entre o número de filhos de uma geração e os seus ancestrais. Inserindo (12) em (11) e trocando a ordem do somatório, para $0 \leq s < 1$,

$$\begin{aligned} G_{n+1}(s) &= \sum_{j \geq 0} \left(\sum_{k \geq 0} s^k P(X_1^{(n)} + \cdots + X_j^{(n)} = k) \right) P(Z_n = j) \\ &\equiv \sum_{j \geq 0} E[s^{X_1^{(n)} + \cdots + X_j^{(n)}}] P(Z_n = j). \end{aligned} \quad (13)$$

Usando, em seguida, a independência das variáveis $X_i^{(n)}$ (ver (2) no Apêndice A),

$$E[s^{X_1^{(n)} + \cdots + X_j^{(n)}}] = E\left[\prod_{i=1}^j s^{X_i^{(n)}}\right] = \prod_{i=1}^j E[s^{X_i^{(n)}}] = \prod_{i=1}^j G(s) = G(s)^j.$$

Logo, (13) se escreve

$$G_{n+1}(s) = \sum_{j \geq 0} P(Z_n = j)G(s)^j \equiv G_n(G(s)),$$

o que prova (9). Para provar (10), é suficiente iterar (9):

$$G_{n+1}(s) = G_n(G(s)) = G_{n-1}(G(G(s))) = G_{n-1}(G_2(s)) = \cdots = G(G_n(s)).$$

□

Estudaremos agora a probabilidade $q := P(\text{extinção})$. Definamos

$$q_n := P(Z_n = 0).$$

Como $Z_n = 0$ implica $Z_{n+1} = 0$,

$$\{Z_1 = 0\} \subset \{Z_2 = 0\} \subset \cdots \subset \{Z_n = 0\} \subset \{Z_{n+1} = 0\} \subset \cdots \subset \{\text{extinção}\},$$

temos

$$q_1 \leq q_2 \leq \cdots \leq q_n \leq q_{n+1} \leq \cdots \leq q.$$

Como

$$\{\text{extinção}\} = \bigcup_{n \geq 0} \{Z_n = 0\},$$

e aplicando a continuidade da probabilidade obtemos

$$q_n \nearrow q.$$

Decorre de (8) e (10) que a sequência q_n obedece a uma relação de recorrência:

$$q_{n+1} = P(Z_{n+1} = 0) = G_{n+1}(0) = G(G_n(0)) = G(q_n).$$

Logo, temos uma sequência crescente q_n , com condição inicial $q_1 = P(Z_1 = 0) = p_0 > 0$, em que q_{n+1} é obtido a partir de q_n pela relação $q_{n+1} = G(q_n)$ e o objetivo é determinar se o limite $q = \lim_n q_n$ é igual a 1 ou menor do que 1. Como o comportamento assintótico de q_n depende inteiramente de G , olhemos mais de perto as propriedades da função $G(s)$, para $0 \leq s \leq 1$.

Lembramos que $G(0) = p_0 > 0$ e $G(1) = 1$. Sabemos, por (5) (com $n = 1$), que G é crescente e, de (7), que ela é convexa. Uma olhada nos gráficos qualitativos de $s \mapsto G(s)$ na Figura 3.4 mostra a importância do valor da derivada de G perto do ponto $s = 1$. Por (6),

$$\lim_{s \nearrow 1} G'(s) = E[Z_1] \equiv m.$$

Dependendo do valor de m , a dinâmica da sequência q_n muda (ver Figura 3.5). Assim temos dois casos para considerar:

1. caso $m \leq 1$: neste caso, como G é convexa, o seu gráfico fica acima da diagonal $y = s$, o que implica que $q_n \nearrow 1$. Logo, $q = 1$ e há extinção;

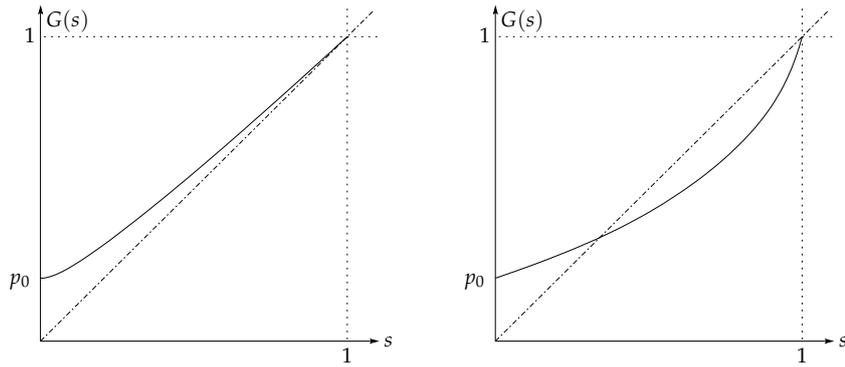


Figura 3.4: A função geradora do número de filhos, nos casos em que a) $m \leq 1$, b) $m > 1$.

2. caso $m > 1$: neste caso, o gráfico de G intersecta a diagonal em exatamente um ponto $q_* < 1$. Como $q_n \nearrow q_*$, isso mostra que $q \equiv q_* < 1$ e há sobrevivência.

Isso conclui a prova do Teorema 3.2. Na verdade, provamos a seguinte caracterização da probabilidade de extinção:

Teorema 3.4 A probabilidade de extinção $q = P(\text{extinção})$ é a menor solução positiva da equação

$$s = G(s). \tag{14}$$

Isto é, q é ponto fixo de G . Se $m \leq 1$, essa solução é $q = 1$; se $m > 1$, temos $q < 1$.

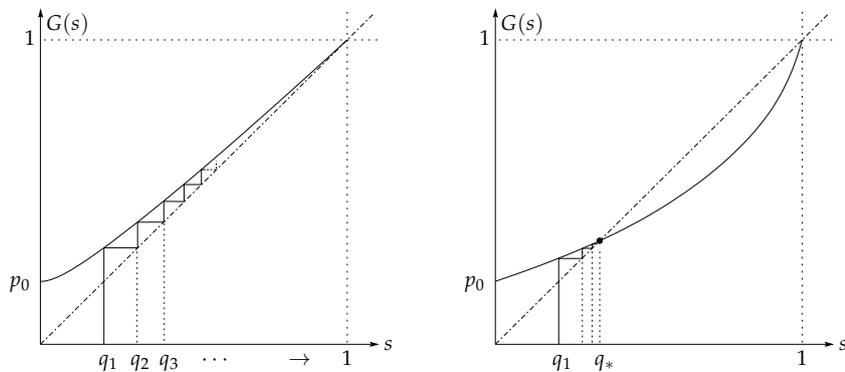


Figura 3.5: A evolução da sequência q_n , nos casos em que a) $m \leq 1$, b) $m > 1$.

3.4 Exercícios

Exercício 3.1 Mostre que para todo $n \geq 1$, $E[Z_n] = m^n$. (Dica: use a relação de recorrência.)

Exercício 3.2 Considere o modelo de Galton-Watson associado à sequência

$$p_k := bp^{k-1}, \quad \forall k \geq 1,$$

$p_0 := 1 - \sum_{k \geq 1} p_k$, em que $b > 0$ e $0 < p < 1$.

1. Mostre que se $0 < b < 1 - p$, a distribuição p_k é bem definida. Calcule p_0 .
2. Estude a sobrevivência dessa população em função dos parâmetros p e b . Em particular, calcule $P(\text{sobrevivência})$.

Exercício 3.3 Considere o modelo de Galton-Watson associado à sequência

$$p_k := \frac{1}{2^{k+1}}, \quad \forall k \geq 0.$$

1. Mostre que a população não sobrevive.
2. Calcule $G(s)$ e mostre, por indução, que para todo $n \geq 2$, G_n pode escrito na forma

$$G_n(s) = \frac{n - (n-1)s}{n + 1 - ns}. \quad (15)$$

Em particular, escreva a série de $G_n(s)$.

3. Mostre que $P(Z_n > 0) = \frac{1}{n+1} \rightarrow 0$ quando $n \rightarrow \infty$. (Dica: calcule primeiro $P(Z_n = 0)$ usando a definição da série de G_n .)

3.5 Referências bibliográficas

Em [4] e [3] se encontram algumas propriedades mais detalhadas do processo de ramificação. O Livro de Athreya e Ney [1] é uma referência completa sobre o assunto.

Apêndice A

Lembrete de Probabilidade

Lembremos aqui algumas definições elementares da Teoria da Probabilidade. Para mais detalhes, o leitor pode consultar [6], [5], [4], [3] ou qualquer outro livro sobre o assunto.

Um **espaço de probabilidade** (Ω, \mathcal{F}, P) é formado por

1. Um conjunto Ω chamado **espaço amostral**.
2. Uma família de eventos \mathcal{F} que satisfaz às seguintes condições: $\Omega \in \mathcal{F}$, se $A \in \mathcal{F}$ então $A^c := \Omega \setminus A \in \mathcal{F}$, e se $A_n \in \mathcal{F}$ então $\bigcup_n A_n \in \mathcal{F}$
3. Uma **probabilidade** $P : \mathcal{F} \rightarrow [0, 1]$: $P(\Omega) = 1$, e se $A_n \in \mathcal{F}$ é uma sequência de eventos dois a dois disjuntos, então $P(\bigcup_n A_n) = \sum_n P(A_n)$.

Em particular, a medida satisfaz à seguinte propriedade de **continuidade**: se A_n é uma sequência crescente, $A_n \subset A_{n+1}$, então $P(A_n) \nearrow P(A)$, em que $A := \bigcup_n A_n$.

Se os eventos $B_n \in \mathcal{F}$ formam uma **partição de Ω** , isto é, se eles são dois a dois disjuntos e $\bigcup_n B_n = \Omega$, então para todo $A \in \mathcal{F}$

$$P(A) = \sum_n P(A \cap B_n). \quad (1)$$

Uma função $X : \Omega \rightarrow \{x_0, x_1, \dots\}$ é chamada **variável aleatória** se $\{X = x_k\} \in \mathcal{F}$ para todo k . Suponha que X tome o valor x_k com probabilidade

$$p_k := P(X = x_k).$$

Observe que $\sum_{k \geq 0} p_k = 1$. Diz-se que X é **integrável** se $\sum_k |x_k| p_k < \infty$. A **esperança matemática** de uma variável integrável X é definida pela série

$$E[X] := \sum_{k \geq 0} x_k p_k.$$

Se $Y : \Omega \rightarrow \{y_0, y_1, \dots\}$ é uma outra variável aleatória, diz-se que X e Y são **independentes** se para todo j, k ,

$$P(X = x_j, Y = y_k) = P(X = x_j)P(Y = y_k).$$

Isto é, duas variáveis são independentes se o valor tomado por uma não influi a probabilidade da outra.

A esperança tem a seguinte propriedade: se X e Y são independentes, então

$$E[XY] = E[X]E[Y].$$

De fato, como o produto XY é uma variável aleatória que toma os valores $x_j y_k$,

$$\begin{aligned} E[XY] &= \sum_{j,k} x_j y_k P(X = x_j, Y = y_k) \\ &= \sum_{j,k} x_j y_k P(X = x_j) P(Y = y_k) \\ &= \left(\sum_j x_j P(X = x_j) \right) \left(\sum_k y_k P(Y = y_k) \right) \\ &= E[X]E[Y]. \end{aligned}$$

Da mesma forma podemos mostrar que se $f(X)$ e $g(Y)$ são independentes e integráveis, então

$$E[f(X)g(Y)] = E[f(X)]E[g(Y)]. \quad (2)$$

Apêndice B

Soluções dos Exercícios

Solução do Exercício 1.1. 1. Lembrando a fórmula do binômio de Newton: $\sum_{k=0}^N \binom{N}{k} a^k b^{N-k} = (a+b)^N$. Com $a = b = 1$, isso mostra que π é uma distribuição de probabilidade. Se $0 < k < N$,

$$\begin{aligned}(\pi Q)_k &= \sum_j \pi_j Q_{j,k} \\ &= \pi_{k-1} Q_{k-1,k} + \pi_{k+1} Q_{k+1,k} \\ &= \frac{1}{2^N} \binom{N}{k-1} \frac{N-(k-1)}{N} + \frac{1}{2^N} \binom{N}{k+1} \frac{k+1}{N} \\ &= \frac{1}{2^N} \frac{k}{N} \binom{N}{k} + \frac{1}{2^N} \frac{N-k}{N} \binom{N}{k} \\ &\equiv \pi_k.\end{aligned}$$

Se $k = 0$, $(\pi Q)_0 = \pi_0 Q_{0,0} + \pi_1 Q_{1,0} = 1/2^N = \pi_0$. Do mesmo jeito, se $k = N$, $(\pi Q)_N = \pi_N$. Isso mostra que π é invariante com respeito a Q .

2. Usando a Fórmula de Stirling,

$$\begin{aligned}\pi_{N/2} &= \frac{1}{2^N} \binom{N}{N/2} = \frac{1}{2^N} \frac{N!}{(N/2)!^2} \\ &\sim \frac{1}{2^N} \frac{\sqrt{2\pi N} N^N e^{-N}}{(\sqrt{2\pi N/2} (N/2)^{N/2} e^{-N/2})^2} = \sqrt{\frac{2}{\pi N}}.\end{aligned}$$

Por outro lado, a probabilidade de ter zero partículas no container B é dada por

$$\pi_0 = \binom{N}{0} \frac{1}{2^N} = \frac{1}{2^N}.$$

Portanto,

$$\frac{\pi_0}{\pi_{N/2}} \simeq \frac{\sqrt{N}}{2^N} \rightarrow 0,$$

o que significa que, no equilíbrio, a probabilidade de existir exatamente a metade das partículas em cada container, apesar de ser pequena, é muito maior do que a de todas as partículas estarem no container A .

Solução do Exercício 1.2. É fácil ver que, para todo $k \geq 0$,

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

Logo, $\lim_{n \rightarrow \infty} Q^n$ não existe.

1. Todos os elementos de Q_ϵ são não negativos e a soma das entradas de cada linha é igual a 1. Logo, Q_ϵ é uma matriz de transição. A cadeia de Markov associada corresponde ao modelo de Ehrenfest com duas partículas em dois containers, quando existe exatamente uma partícula em cada caixa no tempo n . Então, no tempo $n + 1$: (1) as duas partículas estão na caixa A (com probabilidade $(1/2) - \epsilon$) (2) nenhuma partícula se mexeu (com probabilidade $2\epsilon > 0$), (3) as duas partículas estão na caixa B (com probabilidade $1 - 2\epsilon$). A vantagem desse modelo, em comparação com o modelo de Ehrenfest original, é que um pequeno $\epsilon > 0$ ajuda a *misturar* a probabilidade ao longo do tempo.

2. A distribuição invariante para Q_ϵ é definida por $\pi_\epsilon Q_\epsilon = \pi_\epsilon$. Escrevendo $\pi = (a, b, c)$ e resolvendo o sistema associado (lembrando que $a + b + c = 1$) obtemos

$$\pi_\epsilon = \left(\frac{1/2 - \epsilon}{2 - 2\epsilon}, \frac{1}{2 - 2\epsilon}, \frac{1/2 - \epsilon}{2 - 2\epsilon} \right).$$

Observe que $\lim_{\epsilon \rightarrow 0} \pi_\epsilon = (1/4, 1/2, 1/4)$, que coincide com a distribuição invariante do modelo de Ehrenfest.

3. Contas simples mostram que os autovalores de Q_ϵ são, em ordem decrescente: $\lambda_1 = 1$, $\lambda_2 = 0$, $\lambda_3 = 2\epsilon - 1$ e que os autovetores associados respectivos são

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 2\epsilon - 1 \\ 1 \end{pmatrix}.$$

Logo, a diagonalização de Q_ϵ é dada por $Q_\epsilon = B_\epsilon D_\epsilon B_\epsilon^{-1}$, em que

$$B_\epsilon := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2\epsilon - 1 \\ 1 & -1 & 1 \end{pmatrix}, \quad D_\epsilon := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2\epsilon - 1 \end{pmatrix}.$$

Observe que

$$\begin{aligned} Q_\epsilon^n &= (B_\epsilon D_\epsilon B_\epsilon^{-1})(B_\epsilon D_\epsilon B_\epsilon^{-1}) \cdots (B_\epsilon D_\epsilon B_\epsilon^{-1})(B_\epsilon D_\epsilon B_\epsilon^{-1}) \\ &= B_\epsilon D_\epsilon^n B_\epsilon^{-1}. \end{aligned}$$

Note que, como previsto pelo Teorema de Perron-Frobenius, 1 é o maior autovalor de Q_ϵ . Os outros são ambos menores do que 1 em valor absoluto, e desaparecem quanto tomamos o limite com $n \rightarrow \infty$: como $0 < \epsilon < 1/2$, temos

$$\lim_{n \rightarrow \infty} D_\epsilon^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Logo, para uma distribuição inicial $\mu^{(0)}$ qualquer,

$$\lim_{n \rightarrow \infty} \mu^{(0)} Q_\epsilon^n = \mu^{(0)} B_\epsilon \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} B_\epsilon^{-1} = \left(\frac{1/2 - \epsilon}{2 - 2\epsilon}, \frac{1}{2 - 2\epsilon}, \frac{1/2 - \epsilon}{2 - 2\epsilon} \right) \equiv \pi_\epsilon.$$

Solução do Exercício 2.1. Um caminho bidimensional de tamanho $2j$ pode ser escrito como uma sequência de $2j$ símbolos

$$\leftarrow, \downarrow, \rightarrow, \uparrow, \rightarrow, \leftarrow, \leftarrow, \dots, \downarrow, \rightarrow, \downarrow, \leftarrow,$$

com as seguintes condições: $n(\uparrow) = n(\downarrow) = k$, $n(\leftarrow) = n(\rightarrow) = i$, $i + k = j$. O número de sequências que satisfazem a essas duas condições é igual a

$$\binom{2j}{k} \times \binom{2j-k}{k} \times \binom{2j-2k}{i} \times 1.$$

O primeiro fator é o número de maneiras de escolher k símbolos \uparrow em $2j$ possibilidades, o segundo é o número de maneiras de escolher k símbolos \downarrow em $2j - k$ possibilidades, o terceiro é o número de maneiras de escolher i símbolos \leftarrow em $2j - 2k$ possibilidades. O "1" foi colocado para indicar que só há uma maneira de escolher i símbolos \rightarrow entre $2j - 2k - i = i$ possibilidades. Como $i = j - k$, somando sobre k dá o número total de caminhos:

$$\begin{aligned} \sum_{k=0}^j \binom{2j}{k} \binom{2j-k}{k} \binom{2j-2k}{j-k} &= \sum_{k=0}^j \frac{(2j)!}{k!k!(j-k)!(j-k)!} \\ &= \sum_{k=0}^j \frac{(2j)!}{j!j!} \frac{j!j!}{k!k!(j-k)!(j-k)!} \\ &= \binom{2j}{j} \sum_{k=0}^j \binom{j}{k}^2 \\ &= \binom{2j}{j}^2 \end{aligned}$$

Portanto, segue de (9) que

$$P(X_{2j} = 0) = \left(\frac{1}{2^{2j}} \binom{2j}{j} \right)^2 \simeq \frac{1}{j}.$$

Solução do Exercício 2.2. Qualquer trajetória que volta para a origem no tempo $2j$ fez exatamente j passos para a direita, j passos para a esquerda. Logo, como cada passo para direita (resp. esquerda) tem probabilidade p (resp. q), cada trajetória tem probabilidade $p^j q^j = (pq)^j$

$$P(X_{2j} = 0) = \binom{2j}{j} (pq)^j.$$

Já vimos, pela Fórmula de Stirling, que $\binom{2j}{j} \simeq \frac{1}{\sqrt{j}} 2^{2j}$. Logo,

$$P(X_{2j} = 0) \simeq \frac{1}{\sqrt{j}} (4pq)^j$$

Somando,

$$\sum_j P(X_{2j} = 0) \simeq \sum_j \frac{1}{\sqrt{j}} (4pq)^j.$$

Mas, se $p > q$, isto é, $p > 1/2$, então $\rho := 4pq = 4p(1-p) < 1$. Logo, a série $\sum_j \rho^j / \sqrt{j} < \infty$ converge e o passeio é transiente.

Solução do Exercício 2.3. Por um lado, $a_n(s) \leq a_n(s_0)$, o que implica

$$\limsup_{s \nearrow s_0} \sum_{n \geq 0} a_n(s) \leq \sum_{n \geq 0} a_n(s_0).$$

Por outro lado, para todo $N \geq 1$ temos $\sum_{n \geq 0} a_n(s) \geq \sum_{n=0}^N a_n(s)$, o que implica

$$\liminf_{s \nearrow s_0} \sum_{n \geq 0} a_n(s) \geq \sum_{n=0}^N a_n(s_0).$$

Tomando $N \rightarrow \infty$ obtemos a desigualdade desejada.

Solução do Exercício 3.1. Vimos que $E[Z_n] = G'_n(s)|_{s=1^-}$. Mas $G_n(s) = G(G_{n-1}(s))$. Logo, pela regra da cadeia,

$$E[Z_n] = G'(G_{n-1}(s))|_{s=1^-} G'_{n-1}(s)|_{s=1^-} = G'(t)|_{t=1^-} E[Z_{n-1}].$$

Como $G'(t)|_{t=1^-} = m$, isso implica $E[Z_n] = mE[Z_{n-1}]$. Iterando essa expressão obtemos $E[Z_n] = m^n E[Z_0] \equiv m^n$.

Solução do Exercício 3.2. 1. Como $\sum_{k \geq 1} p_k = b/(1-p)$, temos $p_0 = 1 - b/(1-p)$, que é maior do que 0 se, e somente se, $0 < b < 1-p$. 2. A função geradora se calcula facilmente (para $0 < s < 1$):

$$G(s) = p_0 + \sum_{k \geq 1} p_k s^k = p_0 + \frac{b}{p} \sum_{k \geq 1} (sp)^k = p_0 + \frac{bs}{1-sp}.$$

A sua derivada é dada por

$$G'(s) = \frac{b}{(1-sp)^2}.$$

Portanto, $m = \lim_{s \nearrow 1} G'(s) = b/(1-p)^2$. Logo, há sobrevivência se, e somente se, $m > 1$, isto é, se, e somente se, $b > (1-p)^2$. Observe que nessa região de parâmetros, $p_0 < p$ (a probabilidade de ter zero filhos é pequena). O diagrama de fase (isto é, a separação das regiões dos parâmetros (p, b) para os quais há sobrevivência ou extinção) está representado na Figura B.1. O Teorema 3.4 permite achar o valor de $P(\text{extinção})$ via uma equação de ponto fixo.

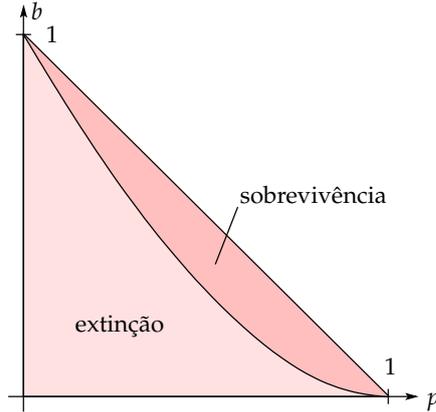


Figura B.1: O diagrama das fases. O modelo está bem definido no triângulo $T := \{(p, b) : 0 < p < 1, 0 < b < 1 - p\}$. A região dos parâmetros para os quais há sobrevivência é $S := \{(p, b) : 0 < p < 1, (1 - p)^2 < b < 1 - p\}$. A região dos parâmetros para os quais há extinção é $E := T \setminus S$.

Ora, $s = G(s)$ dá

$$ps^2 - (1 + p_0p - b)s + p_0 = ps^2 - (p + p_0)s + p_0 = 0,$$

cujas soluções são

$$s = \frac{p + p_0 \pm |p - p_0|}{2p} = \{1, p_0/p\}.$$

A solução $s = 1$ é a solução trivial; a outra dá $P(\text{extinção}) = p_0/p$. Isto é, $P(\text{sobrevivência}) = 1 - p_0/p = 1 - p^{-1} + b/(p(1 - p))$.

Solução do Exercício 3.3. 1. Temos

$$m = \sum_{k \geq 0} \frac{k}{2^{k+1}} = \frac{1}{4} \sum_{k \geq 1} (x^k)' \Big|_{x=1/2} = \frac{1}{4} \frac{d}{dx} \left(\sum_{k \geq 1} x^k \right) \Big|_{x=1/2} = \frac{1}{4} \frac{d}{dx} \left(\frac{x}{1-x} \right) \Big|_{x=1/2} = 1.$$

Pelo teorema provado na aula, $P(\text{extinção}) = 1$.

2. Usando de novo a série geométrica, é fácil calcular $G(s) = \sum_{k \geq 0} s^k / 2^{k+1} = (2 - s)^{-1}$, que é da forma (15) para $n = 1$. Suponha que G_n seja da forma (15) para n . Para $n + 1$,

$$G_{n+1}(s) = G_n(G(s)) = \frac{n - (n - 1)G(s)}{n + 1 - nG(s)} = \dots = \frac{n + 1 - ns}{n + 2 - (n + 1)s}.$$

o que prova que G_{n+1} também é da forma (15) para $n + 1$. Usando a série geométrica e juntando os termos obtemos

$$G_n(s) = \frac{n}{n + 1} + \frac{1}{n(n + 1)} \sum_{j \geq 1} \left(\frac{ns}{n + 1} \right)^j$$

3. Como $P(Z_n = 0)$ é o termo constante da série de G_n , temos $P(Z_n = 0) = \frac{k}{k+1}$. Também, $P(Z_n > 0) = 1 - P(Z_n = 0) = \frac{1}{k+1}$.

Referências Bibliográficas

- [1] K. B. Athreya and P. E. Ney. Branching processes. Dover, 2004.
- [2] K.L. Chung. Markov Chains with Stationary Transition Probabilities. Springer, New York, 1967.
- [3] R. Durrett. Probability: Theory and Examples. Duxbury Press, 1988.
- [4] G.R. Grimmett and D.R. Stirzaker. Probability and Random Processes. Oxford University Press, 2005.
- [5] C.M. Grinstead and J.L. Snell. Introduction to Probability. AMS, 1997.
- [6] B.R. James. Probabilidade: um curso em nível intermediário. Projeto Euclides. IMPA, Rio de Janeiro, 1981.
- [7] J.G. Kemeny, J.L. Snell, and A.W. Knapp. Denumerable Markov chains. Springer, 1976.
- [8] G. Lawler. Introduction to Stochastic Processes. Chapman and Hall, 2006.
- [9] G.F. Lawler and V. Limic. Random walk: a modern introduction, volume 123 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2010.
- [10] P. Révész. Random walk in random and nonrandom environments. World Scientific Publishing Co. Inc., 1990.
- [11] E. Seneta. Non-Negative Matrices and Markov Chains. Springer Series in Statistics, 1976.
- [12] F. Spitzer. Principles of random walks. Springer-Verlag, New York, 1976.

