

Joseane Gregório Gomes

Um modelo probabilístico para o problema da irreversibilidade dos gases

São José do Rio Preto

2018

Joseane Gregório Gomes

Um modelo probabilístico para o problema da irreversibilidade dos gases

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista Júlio de Mesquita Filho”, Campus de São José do Rio Preto, Polo de Bauru.
Financiadora: CAPES
Orientador: Prof. Dr. Fabiano Borges da Silva

São José do Rio Preto

2018

Gomes, Joseane Gregório.

Um modelo probabilístico para o problema da irreversibilidade dos gases / Joseane Gregório Gomes. -- São José do Rio Preto, 2018
85 f. : il.

Orientador: Fabiano Borges da Silva

Dissertação (mestrado profissional) – Universidade Estadual Paulista (Unesp), Instituto de Biociências, Letras e Ciências Exatas, São José do Rio Preto

1. Matemática (Ensino médio) - Estudo e ensino. 2. Probabilidades – Estudo e ensino. 3. Markov, Processos de. 4. Processo estocástico. 5. Matrizes (Matemática). 6. Matemática – Metodologia. I. Título.

CDU – 51(07)

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Câmpus de São José do Rio Preto

Joseane Gregório Gomes

Um modelo probabilístico para o problema da irreversibilidade dos gases

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista Júlio de Mesquita Filho”, Campus de São José do Rio Preto, Polo de Bauru.
Financiadora: CAPES

Comissão Examinadora

Prof. Dr. Fabiano Borges da Silva

UNESP- Câmpus de Bauru

Orientador

Prof^a Dr^a. Cristiane Alexandra Lázaro

UNESP- Câmpus Bauru

Prof. Dr. Régis Leandro Braguim Stábile

IFSP- Câmpus Birigui

Bauru

05 de setembro de 2018

Agradecimentos

Agradeço em primeiro lugar a Deus, que sempre iluminou os meus pensamentos e guiou meus passos até aqui.

Aos meus amados pais Antonio e Rosemeire, que sempre me mostraram a importância dos estudos e não mediram esforços para me apoiar nesta jornada.

Aos meus irmãos Juliane, Antonio Jr e João, que sempre torceram pelo meu sucesso e ao meu namorado Rafael, pelo carinho, força, e palavras de conforto nos momentos difíceis.

Ao meu orientador Prof. Dr. Fabiano Borges da Silva, pela paciência, orientação e incentivo, que permitiram a conclusão deste trabalho.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Neste trabalho apresentamos uma introdução aos processos estocásticos de Markov discretos e suas propriedades, e como uma aplicação, estudamos um modelo probabilístico para o problema da irreversibilidade dos gases, ou modelo da urna de Ehrenfest. Por fim, apresentamos uma modificação deste modelo, cuja abordagem é adaptada para o Ensino Médio.

Palavras-chave: Cadeia de Markov. Processo estocástico. Irreversibilidade dos gases. Probabilidade. Matrizes de transição.

Abstract

In this work we present an introduction to discrete Markov stochastic processes and their properties, and as an application, we study a probabilistic model for the problem of irreversibility of gases, or model of Ehrenfest urn. Finally, we present a modification of this model, whose approach is adapted for High School.

Keywords: Markov chain. Stochastic processes. Irreversibility of gases. Probability. Transition matrices.

Lista de Figuras

4.1	Diagrama de transição do exemplo da porta.	42
4.2	Diagrama de transição do exemplo do sapo.	45
5.1	Físico Austríaco Paul Ehrenfest (1880-1933)	47
5.2	Momento inicial do experimento com o gás.	51
5.3	Experimento no tempo igual à 1.	51
5.4	Experimento em equilíbrio termodinâmico.	52
5.5	Estado do sistema no tempo n	53
5.6	Estado do sistema no tempo $n - 1$	57
5.7	Dinâmica do sistema regido por Q_ϵ	65
6.1	O modelo da urna para o sorteio das bolinhas.	67
6.2	O nosso modelo de urna para o sorteio de bolinhas.	73
6.3	Cronômetro.	73
6.4	Momento inicial do sorteio.	74
6.5	1º lançamento simultâneo dos dois dados.	74
6.6	1º lançamento de dados realizado.	75
6.7	Realizando o 1º sorteio.	75
6.8	A primeira bolinha sorteada.	76
6.9	1ª substituição das bolinhas.	76
6.10	Após a substituição da 1ª bolinha.	77
6.11	O 2º lançamento simultâneo dos dados.	77
6.12	2º lançamento de dados realizado.	78
6.13	2º sorteio da bolinha realizado.	78

6.14	2ª substituição das bolinhas.	79
6.15	3º lançamento simultâneo dos dados.	79
6.16	Encerrando nosso sorteio.	80
6.17	Tempo de realização do nosso experimento.	80

Sumário

Introdução	10
1 Teoria da Probabilidade	12
1.1 Espaço de Probabilidade	12
1.1.1 Probabilidade	14
1.2 Propriedades das Probabilidades.....	15
1.3 Probabilidade Condicional	17
1.4 Variáveis Aleatórias	20
2 Matrizes Estocásticas	22
2.1 Uma revisão sobre matrizes e suas operações para fixar notações.	22
2.1.1 Adição de matrizes	24
2.1.2 Multiplicação de uma matriz por um número	25
2.2 Matrizes Estocásticas e uma versão do Teorema de Perron-Frobenius . . .	30
3 Processos Estocásticos	34
3.1 Álgebra de eventos	34
3.2 Processo estocástico	36
3.3 Probabilidade no espaço das sequências	37
4 Cadeias de Markov	39
4.1 Probabilidade de transição	39
5 O problema do gás de Ehrenfest	47
5.1 Biografia de Paul Ehrenfest.....	47

5.2	Introdução ao problema do gás de Ehrenfest	50
5.3	Tempo de Permanência	61
5.4	Uma pequena perturbação no modelo	64
6	Sorteio de bolinhas: uma abordagem voltada para uma aula de Matemática no Ensino Médio	66
6.1	O modelo para o sorteio das bolinhas	67
6.2	Uma simulação para o sorteio de bolinhas	73
6.3	Sugestão de plano de aula	81
	Conclusão	83
	Referências	84

Introdução

A contextualização de conteúdos matemáticos sempre foi e ainda hoje é um dos maiores desafios para os professores da disciplina em nosso país. É preciso que nossos alunos vejam as aplicações de tais conteúdos nas mais diversas áreas do conhecimento e também no seu dia-a-dia, para que a aprendizagem se torne significativa, e tanto professor quanto aluno alcancem os seus objetivos acadêmicos. E é em busca dessa contextualização que apresentaremos as cadeias de Markov como um instrumento que nos possibilita interagir os conhecimentos sobre matrizes, sistemas lineares e probabilidade, que são ensinados ao longo do Ensino Médio.

Resumidamente, dizemos que uma cadeia de Markov é um processo estocástico sem memória que caracteriza-se por seu estado futuro depender apenas do seu estado presente (que é conhecido), não sofrendo influências de seus estados passados. Recebe esse nome em homenagem ao seu idealizador Andrei Andreyevich Markov e possui diversas aplicações em Física, Química, Estatística, Economia, Biologia, Genética, Ciências Sociais, Medicina, dentre outros.

Este trabalho busca oferecer aos leitores um embasamento teórico sobre cadeias de Markov e apresentar um problema originário da Física que buscou nas cadeias de Markov uma solução para tal, e por fim propor para professores de Matemática uma abordagem mais simples do problema apresentado, para que possam aplicar em suas aulas/oficinas de Matemática no Ensino Médio. O trabalho está organizado da seguinte forma:

No Capítulo 1 introduzimos os principais conceitos da Teoria de Probabilidade, como por exemplo, as definições de álgebra, σ -álgebra, probabilidade condicional e variável aleatória.

No Capítulo 2 fizemos uma introdução ao estudo de matrizes e suas operações e apresentamos as matrizes de probabilidade de transição, as quais são também chamadas de matrizes estocásticas. Além disso, apresentamos uma versão do Teorema de Perron-Frobenius (Teorema 2), o qual garante convergência de matrizes estocásticas regulares. Este resultado será útil para estudar o comportamento assintótico das distribuições de probabilidade, na abordagem apresentada no Capítulo 6. Já no capítulo 3 foi feita uma breve

introdução aos processos estocásticos discretos, onde o tempo t em que o processo evolui é discreto ($t = 0, 1, 2, \dots$) e o espaço de estados (valores que podem ocorrer no decorrer do processo) é finito (ou enumerável).

No capítulo 4, explicamos o que é uma cadeia de Markov de maneira formal (Definição 13). A ideia de que o futuro do processo só depende do estado presente e não do passado, é formalizada pela equação (4.2).

No capítulo 5 apresentamos um pouco da história do físico Paul Ehrenfest e apresentamos o problema da irreversibilidade dos gases e um modelo probabilístico para tal problema. Boa parte do capítulo foi baseado em Friedli (2011).

A maior contribuição do trabalho consiste em provar que a distribuição de Bernoulli é de fato uma distribuição invariante para a matriz de transição Q do nosso modelo proposto (Teorema 3). A distribuição invariante será útil para constatar que existe a possibilidade dos gases retornarem ao seu estado inicial, porém o tempo de retorno é tão alto (para 100 partículas, por exemplo, pode ultrapassar a idade do Universo- Seção 5.3), que o processo é considerado irreversível. No último capítulo apresentamos um problema que foi pensado durante as discussões entre a autora e orientador deste trabalho, para que o problema da irreversibilidade se torna-se mais acessível e de fácil aplicação em sala de aula ou oficinas de Matemática. O nosso problema consiste em um modelo de sorteio de bolinhas distribuídas em duas cores diferentes: vermelhas e azuis. Apenas as vermelhas iniciam-se na caixa. Antes de realizar o sorteio, lançam-se simultaneamente dois dados e, caso a face voltada para cima de ambos seja qualquer resultado diferente de 6 e 6, realizamos o sorteio de uma bolinha de dentro da caixa. A cada sorteio realizado, a bolinha sorteada é substituída por uma bolinha de cor “oposta”. O sorteio apenas se encerra quando a caixa voltar a ter todas as bolinhas vermelhas dentro dela. Para a resolução deste problema, fizemos uma perturbação na matriz de transição Q do capítulo 5, para que fosse possível aplicar o Teorema 2, gerando um repertório de conteúdo pertinente ao Ensino Médio.

Capítulo 1

Teoria da Probabilidade

Trataremos neste capítulo sobre alguns resultados da teoria da probabilidade, abordando inclusive a probabilidade condicional, muito utilizada nas Cadeias de Markov. Algumas definições como a de álgebra, σ -álgebra e variável aleatória também serão apresentadas. Para este capítulo usaremos como texto base James (1996), portanto, algumas das definições e propriedades aqui apresentadas foram extraídas do livro indicado.

1.1 Espaço de Probabilidade

Ao realizarmos um experimento podemos obter um conjunto Ω de resultados possíveis, e este conjunto será chamado de **espaço amostral**.

Exemplo 1. *Ao sortear de uma urna com dez bolas, todas numeradas de 1 a 10, uma única bola, podemos obter os seguintes resultados: 1, 2, 3, 4, 5, 6, 7, 8, 9 ou 10. Ou seja, o espaço amostral para este experimento será:*

$$\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

Ao realizar um experimento, há certos conjuntos que atribuímos probabilidades de ocorrer ou não. E a estes conjuntos denominamos de **eventos aleatórios**. Por exemplo, no experimento apresentado anteriormente, pode-se observar alguns eventos:

A= “observa-se um número par”

B= “observa-se um número ímpar”

C= “observa-se um número maior do que 5”

D= “observa-se um múltiplo do 3”

E cada um desses eventos pode ser identificado à um subconjunto de Ω que são:

$$A = \{2,4,6,8\}, B = \{1,3,5,7,9\}, C = \{6,7,8,9\}, D = \{3,6,9\}.$$

Temos ainda que, Ω é o evento certo e \emptyset é o evento impossível de ocorrer.

Indicaremos por \mathcal{A} a *classe* dos eventos aleatórios, que deve possuir certas propriedades essenciais para a teoria de probabilidade. São elas:

A1. $\Omega \in \mathcal{A}$;

A2. Se $A \in \mathcal{A}$ então $A^c \in \mathcal{A}$;

A3. Se $A \in \mathcal{A}$ e $B \in \mathcal{A}$, então $A \cup B \in \mathcal{A}$.

Definição 1. *Seja Ω um conjunto não-vazio. Uma classe \mathcal{A} de subconjuntos de Ω que satisfazem as propriedades A1, A2 e A3 é chamada de **álgebra** de subconjuntos de Ω .*

Como consequência das propriedades A_1, A_2, A_3 , das leis de Morgan e ainda do fato de que se $A, B \in \mathcal{A}$, então temos que $A \cap B = (A^c)^c \cap (B^c)^c = (A^c \cup B^c)^c \in \mathcal{A}$ segue que:

A4. $\emptyset \in \mathcal{A}$ (pois $\emptyset = \Omega^c$);

A5. $\forall n \in \mathbb{N}$ e \forall sequência de subconjuntos $A_1, A_2, A_3, \dots, A_n \in \mathcal{A}$, temos $\bigcup_{i=1}^n A_i \in \mathcal{A}$ e $\bigcap_{i=1}^n A_i \in \mathcal{A}$.

Considere a seguinte propriedade para eventos aleatórios:

A3'. Se $A_n \in \mathcal{A}$ para $n = 1, 2, 3, \dots$, então $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Definição 2. *Uma classe \mathcal{A} de subconjuntos de um conjunto não-vazio Ω que satisfaça A1, A2, A3' é chamada **σ -álgebra** de subconjuntos de Ω .*

Observação: Uma σ -álgebra é sempre uma álgebra, pois A3 é consequência de A3', já que $A \cup B = A \cup B \cup B \cup B \dots \in \mathcal{A}$ se \mathcal{A} é σ -álgebra.

Proposição 1. *Seja \mathcal{A} uma σ -álgebra de subconjuntos de Ω . Se $A_1, A_2, A_3, \dots \in \mathcal{A}$, então $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$.*

Demonstração: Basta notar que:

$$\bigcap_{i=1}^{\infty} A_i = (\bigcup_{i=1}^{\infty} A_i^c)^c$$

e aplicar a A2, A3'.

Podemos dizer, então, que uma σ -álgebra é fechada para um número *enumerável* de aplicações das operações \cup , \cap , e c .

Exemplo 2. (*σ -álgebra de eventos aleatórios*) *Se Ω for finito ou enumerável, então uma σ -álgebra de Ω é formada por todas as partes de \mathcal{A} , isto é, σ -álgebra = $\mathcal{P}(\Omega)$. Por exemplo, no experimento do início deste capítulo (o de sortear uma bola em uma urna com dez bolas numeradas de 1 a 10), temos que $\mathcal{A} = \mathcal{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, \{3\}, \dots, \{1, 2\}, \{1, 3\}, \dots, \Omega\}$. A classe \mathcal{A} tem $2^{10} = 1024$ elementos, de modo que há 1024 eventos aleatórios associados a este experimento (se Ω tem n elementos, $\mathcal{P}(\Omega)$ tem 2^n).*

1.1.1 Probabilidade

Para a definição de probabilidade de um evento aleatório, pode-se utilizar algumas definições clássicas como, por exemplo, o de resultados equiprováveis, probabilidade geométrica ou até mesmo o de *frequência relativa*: onde $\mathbb{P}(A)$ é definida como o limite da frequência relativa da ocorrência de A em n repetições independentes do experimento, com n tendendo ao infinito, ou seja:

$$\mathbb{P}(A) = \lim_{n \rightarrow \infty} \frac{\Phi(n)}{n}.$$

Se $\Phi(n)$ é o número de ocorrências de A em n “ensaios” independentes do experimento, então definição dada acima é chamada de “frequentista” ou “estatística” de probabilidade. Vamos supor agora que, para todo $A \in \mathcal{A}$ seja associado um número real $\mathbb{P}(A)$, chamado de probabilidade de A , de modo que sejam satisfeitos os axiomas a seguir:

Axioma 1. $\mathbb{P}(A) \geq 0$;

Axioma 2. $\mathbb{P}(\Omega) = 1$;

Axioma 3. Se $A_1, A_2, A_3, \dots, A_n \in \mathcal{A}$ são disjuntos, dois a dois, então:

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i).$$

A construção axiomática de probabilidade dada acima deve-se ao matemático soviético Kolmogorov (1903-1987). Uma função $\mathbb{P} : \mathcal{A} \rightarrow \mathbb{R}$ satisfazendo os Axiomas 1,2 e 3 é chamada de *probabilidade finita aditiva*.

Axioma 3'. Se $A_1, A_2, A_3, \dots \in \mathcal{A}$ formam uma sequência disjunta, então:

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

Definição 3. Uma função \mathbb{P} definida numa σ -álgebra \mathcal{A} satisfazendo os Axiomas 1,2 e 3' é chamada de *medida de probabilidade em \mathcal{A}* , ou simplesmente, de *probabilidade em \mathcal{A}* . Quando \mathbb{P} não satisfaz o Axioma 2, dizemos apenas que \mathbb{P} é uma medida em \mathcal{A} .

Um espaço de probabilidade é um trio $(\Omega, \mathcal{A}, \mathbb{P})$ onde, Ω é um conjunto não vazio, \mathcal{A} é uma σ -álgebra de subconjuntos de Ω e \mathbb{P} é uma probabilidade em \mathcal{A} .

Temos a seguir alguns exemplos de espaços de medida de probabilidade:

i) Seja Ω um conjunto qualquer, tome $\mathcal{A} = \{\emptyset, \Omega\}$. Definindo $\mathbb{P}(\emptyset) = 0$ e $\mathbb{P}(\Omega) = 1$ temos que $(\Omega, \mathcal{A}, \mathbb{P})$ é um espaço de probabilidade trivial.

ii) Considere o conjunto $\Omega = \{1, 2, 3, \dots\}$. Escolhe-se como medida de probabilidade $\mathbb{P}(k) = \frac{1}{2^k}$, $k = 1, 2, 3, \dots$ e \mathcal{A} é formada por subconjuntos de Ω . Temos que, para $A \in \mathcal{A}$, $\mathbb{P}(A)$ será dada pela soma de probabilidades de eventos elementares $\omega_i \in A, i = 1, 2, 3, \dots$. Portanto, $\mathbb{P}(A) \geq 0$, para todo A , e $\sum_{i=1}^{\infty} \mathbb{P}(k) = \frac{1}{1 - \frac{1}{2}} = 1$, logo $\mathbb{P}(\Omega) = 1$.

1.2 Propriedades das Probabilidades

Seja $(\Omega, \mathcal{A}, \mathbb{P})$ um espaço de probabilidade e $A, A_i \in \mathcal{A}, i = 1, 2, \dots$. Então, as propriedades abaixo são consequências dos axiomas citados acima:

Propriedade 1. $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$.

De fato, pelo Axioma 2, $\mathbb{P}(\Omega) = 1$ e $\Omega = A \cup A^c$ e, como A e A^c são disjuntos, segue do Axioma 3 que

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c).$$

Logo, $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$.

Como caso especial, para $\mathbb{P}(\emptyset) = 1 - \mathbb{P}(\Omega) = 0$.

Propriedade 2. $0 \leq \mathbb{P}(A) \leq 1$.

De fato, Axioma 1 temos que $0 \leq \mathbb{P}(A)$. Por outro lado, pela Propriedade 1, temos que $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$ e, como $\mathbb{P}(A^c) \geq 0$, segue que $\mathbb{P}(A) \leq 1$. Logo, $0 \leq \mathbb{P}(A) \leq 1$.

Propriedade 3. Se $A_1 \subset A_2 \implies \mathbb{P}(A_1) \leq \mathbb{P}(A_2)$.

De fato, observe que $A_2 = A_1 \cup (A_2 - A_1)$. Do Axioma 1 e do Axioma 3 temos que

$$\mathbb{P}(A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2 - A_1) \geq \mathbb{P}(A_1).$$

Logo, $\mathbb{P}(A_1) \leq \mathbb{P}(A_2)$.

Propriedade 4. $\mathbb{P}(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \mathbb{P}(A_i)$.

De fato, pelo Axioma 3, temos $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2 \cap A_1^c) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2)$. Pela Propriedade 3, pois $A_2 \cap A_1^c \subset A_2$. Completaremos a prova por indução.

Observe que, para $n = 2$, já está provado acima. Suponhamos, então, a validade para n ($\mathbb{P}(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \mathbb{P}(A_i)$) e demonstraremos para $n + 1$.

Note que,

$$\mathbb{P}(\bigcup_{i=1}^{n+1} A_i) = \mathbb{P}((\bigcup_{i=1}^n A_i) \cup A_{n+1}).$$

Pela hipótese de indução e pelo caso $n = 2$ segue que:

$$\begin{aligned}
 \mathbb{P}\left(\bigcup_{i=1}^{n+1} A_i\right) &= \mathbb{P}\left(\left(\bigcup_{i=1}^n A_i\right) \cup A_{n+1}\right) \\
 &\leq \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) + \mathbb{P}(A_{n+1}) \\
 &\leq \sum_{i=1}^n \mathbb{P}(A_i) + \mathbb{P}(A_{n+1}) \\
 &= \sum_{i=1}^{n+1} \mathbb{P}(A_i).
 \end{aligned}$$

Como para $n + 1$ é verdadeiro, a nossa hipótese de indução é válida para todo n .

Propriedade 5. $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

De fato, temos que:

$A_1 \cup A_2 \cup A_3 \cup \dots = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus (A_1 \cup A_2)) \cup (A_4 \setminus (A_1 \cup A_2 \cup A_3)) \cup \dots$, onde $A_1, (A_2 \setminus A_1), (A_3 \setminus (A_1 \cup A_2)), (A_4 \setminus (A_1 \cup A_2 \cup A_3)) \dots$ são disjuntos.

Dessa maneira, pelo Axioma 3' temos que

$$\begin{aligned}
 \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) &= \mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \dots) \\
 &= \mathbb{P}(A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus (A_1 \cup A_2)) \cup \dots) \\
 &= \mathbb{P}(A_1) + \mathbb{P}(A_2 \setminus A_1) + \mathbb{P}(A_3 \setminus (A_1 \cup A_2)) + \dots \\
 &\leq \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) + \dots \\
 &= \sum_{i=1}^{\infty} \mathbb{P}(A_i)
 \end{aligned}$$

A desigualdade acima deve-se ao fato de que $(A_1 \setminus A_2) \subset A_2, (A_3 \setminus (A_1 \cup A_2)) \subset A_3, \dots$ e também em decorrência da Propriedade 3.

1.3 Probabilidade Condicional

Definição 4. *Seja $(\Omega, \mathcal{A}, \mathbb{P})$ um espaço de probabilidade. Se $B \in \mathcal{A}$ e $\mathbb{P}(B) > 0$, a probabilidade condicional de $A \in \mathcal{A}$, dado, B é definida por*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

A probabilidade de A dado B representa, intuitivamente, a probabilidade do evento A acontecer, sabendo que B ocorreu de antemão.

Exemplo 3. *Do lançamento de um dado sejam os eventos:*

$A = \{\text{o número obtido é maior do que } 2\};$

$B = \{\text{o número obtido é ímpar}\}.$

Sabendo que B ocorreu por exemplo, temos que:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{3, 5\})}{\mathbb{P}(\{1, 3, 5\})} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3}$$

Veremos, a seguir, que uma probabilidade condicional é de fato uma probabilidade.

Proposição 2. *Seja $(\Omega, \mathcal{A}, \mathbb{P})$ um espaço de probabilidade. Fixemos $B \in \mathcal{A}$ um evento de probabilidade positiva $\mathbb{P}(B) > 0$. Então, $\mathbb{P}(\cdot|B)$ é uma probabilidade em \mathcal{A} .*

Demonstração. Temos que

$$\mathbb{P}(\emptyset|B) = \frac{\mathbb{P}(\emptyset \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(B)} = 0$$

e também que

$$\mathbb{P}(\Omega|B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1.$$

Além disso, $0 \leq \mathbb{P}(A|B) \leq 1$, pois $(A \cap B) \subset B$ e, assim, $\mathbb{P}(A \cap B) \leq \mathbb{P}(B)$.

Falta apenas verificar a aditividade. Sejam A_1, A_2, A_3, \dots eventos disjuntos. Logo, os eventos $A_1 \cap B, A_2 \cap B, A_3 \cap B, \dots$ também são disjuntos. Portanto,

$$\mathbb{P}(\cup_{k=1}^{\infty} A_k|B) = \frac{\mathbb{P}(\cup_{k=1}^{\infty} (A_k \cap B))}{\mathbb{P}(B)} = \frac{\sum_{k=1}^{\infty} \mathbb{P}(A_k \cap B)}{\mathbb{P}(B)} = \sum_{k=1}^{\infty} \mathbb{P}(A_k|B).$$



O próximo resultado mostra qual é a relação entre a probabilidade de A acontecer dado que B ocorreu, e a probabilidade de B acontecer dado que A ocorreu.

Teorema 1. (Teorema de Bayes) *Sejam B_1, B_2, \dots eventos, todos de probabilidade positiva, que particionam o espaço Ω (são disjuntos e sua união é igual a Ω). Dado um evento A , vale que:*

$$a) \mathbb{P}(A) = \sum_{k=1}^{\infty} \mathbb{P}(A|B_k)\mathbb{P}(B_k);$$

b) *Suponha que $\mathbb{P}(A) > 0$. Então, para qualquer índice $i \in \mathbb{N}$, temos*

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{k=1}^{\infty} \mathbb{P}(A|B_k)\mathbb{P}(B_k)}.$$

Demonstração. Como B_1, B_2, \dots particionam o espaço Ω , temos que $\Omega = \cup_{k=1}^{\infty} B_k$. Logo, $A = A \cap \Omega = \cup_{k=1}^{\infty} (A \cap B_k)$. Como os eventos B_k são disjuntos, $A \cap B_k$ também são disjuntos. Temos que $\mathbb{P}(A \cap \Omega) = \sum_{k=1}^{\infty} \mathbb{P}(A \cap B_k)$ (pela aditividade da probabilidade), o que mostra o item a) do enunciado. E, como consequência, temos que

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(B_i \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{k=1}^{\infty} \mathbb{P}(A|B_k)\mathbb{P}(B_k)}.$$

■

Exemplo 4. *Uma vacina tem 85% de eficácia na imunização contra um novo vírus da gripe, que acomete 40% da população não vacinada. Supondo que após uma campanha de vacinação onde 75% da população foi atingida, um paciente chegue ao hospital com o vírus em questão, não sabendo informar se tomou ou não a vacina. Qual é a probabilidade de que este paciente a tenha tomado?*

Sejam os eventos:

$A = \{\text{paciente foi infectado pelo vírus}\}.$

$B = \{\text{paciente tomou a vacina}\}.$

Queremos calcular $\mathbb{P}(B|A)$ e, para isso, temos os seguintes dados:

$$\mathbb{P}(A|B) = 0,15; \quad \mathbb{P}(A|B^c) = 0,40; \quad \mathbb{P}(B) = 0,75.$$

De 1 temos:

$$\begin{aligned} \mathbb{P}(B|A) &= \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)} \\ &= \frac{0,15 \times 0,75}{0,15 \times 0,75 + 0,40 \times 0,25} \\ &= \frac{0,1125}{0,1125 + 0,1} \\ &\approx 0,53. \end{aligned}$$

Definição 5. *Dois eventos A e B são independentes se $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

No caso de B ter probabilidade positiva, é o mesmo que dizer que $\mathbb{P}(A|B) = \mathbb{P}(A)$. Traduzindo, se A e B são independentes, o fato de saber que B ocorreu não altera a probabilidade de A ocorrer.

Definição 6. *Os eventos $A_1, A_2, A_3, \dots \in \mathcal{A}$ são independentes, dois a dois, se para quaisquer $i \neq j$, vale*

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j).$$

1.4 Variáveis Aleatórias

Uma variável aleatória (v.a.) é uma função cujo domínio é Ω (espaço amostral) e associa um número real ao resultado de um experimento. Formalmente, temos:

Definição 7. *Seja $(\Omega, \mathcal{A}, \mathbb{P})$ um espaço de probabilidade. Diremos que uma função $X : \Omega \rightarrow \mathbb{R}$ é uma variável aleatória se o conjunto $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A}$ (é um evento), para todo $x \in \mathbb{R}$.*

Exemplo 5. *Considere que uma moeda seja lançada três vezes sucessivamente. Seja C cara e K coroa. O espaço amostral deste experimento é:*

$$\Omega = \{(C, C, C); (C, C, K); (C, K, C); (K, C, C); (C, K, K); (K, C, K); (K, K, C); (K, K, K)\},$$

e podemos definir uma variável aleatória X como sendo o número de caras obtidas nos três lançamentos. Assim, teríamos por exemplo, que $X((C, C, C)) = 3$, $X((K, C, C)) = 2$. Note que, $X(\omega) \in \{0, 1, 2, 3\}, \forall \omega \in \Omega$.

Dizemos que X é uma **variável aleatória discreta** se o número de valores possíveis de X for um conjunto finito ou enumerável $\{x_1, x_2, x_3, \dots\} \subset \mathbb{R}$, ou seja

$$X(\omega) \in \{x_1, x_2, x_3, \dots\}, \forall \omega \in \Omega.$$

O Exemplo 5 é um caso de variável aleatória discreta.

Dizemos que X é uma **variável aleatória contínua** se os valores possíveis de X for qualquer valor numérico em um determinado intervalo, ou, coleções de intervalos.

Um exemplo de variável aleatória contínua seria uma função $X : [0, 1] \rightarrow \mathbb{R}$ tal que $X(w) = w$, onde esta indicaria um valor numérico para um ponto w escolhido aleatoriamente no intervalo $[0, 1]$. Neste exemplo ainda seria necessário definir a σ -álgebra, mas não faremos isto por não ser o objetivo deste trabalho.

Capítulo 2

Matrizes Estocásticas

O conteúdo matrizes é proposto pelo Currículo do Estado de São Paulo: Matemática e suas tecnologias (2011) para o 2º ano do Ensino Médio, onde além de compreender o significado de matrizes e suas operações o aluno deve saber expressar por meio delas situações relativas a fenômenos físicos ou geométricos que possam aparecer no cotidiano. Estas situações garantem uma melhor compreensão entre o conteúdo aprendido e sua importância.

Neste capítulo apresentaremos definições de matrizes e suas operações, que serão necessárias para a compreensão dos próximos capítulos. Para aqueles que necessitam ou querem se aprofundar em tal conteúdo, recomendamos Boldrini (1980). O leitor familiarizado com propriedades básicas sobre matrizes, pode ir direto para a Seção 2.2 deste capítulo, onde definiremos matrizes estocásticas.

2.1 Uma revisão sobre matrizes e suas operações para fixar notações.

Sejam $m \geq 1$ e $n \geq 1$ ($m, n \in \mathbb{Z}$). Chamamos de matriz $m \times n$ à uma sequência de números reais, distribuídos em m linhas e n colunas formando uma tabela, onde cada número que compõe esta matriz será chamado de *termo* e representado por a_{ij} . Representaremos a matriz da seguinte maneira:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \dots & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \dots & \dots & a_{3n} \\ a_{41} & a_{42} & a_{43} & a_{44} & \dots & \dots & a_{4n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & a_{m4} & \dots & \dots & a_{mn} \end{pmatrix}$$

Indicaremos por $M_{m \times n}(\mathbb{R})$ o conjunto das matrizes $m \times n$ com termos (entradas) reais. Se $m \neq n$ chamaremos a matriz M de **matriz retangular**, já se $m = n$ temos que M é uma **matriz quadrada** e podemos indicá-la por M_n . E são estas últimas, que utilizaremos com frequência no presente trabalho.

Em uma matriz quadrada os termos $a_{11}, a_{22}, a_{33}, \dots, a_{nn}$ são chamados de *diagonal principal da matriz*. Temos ainda matrizes que possuem uma única linha, chamadas de **matriz linha** e matrizes que possuem uma única coluna, chamadas de **matriz coluna**.

A matriz O descrita abaixo é chamada de **matriz nula**:

$$O = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Se $A = (a_{ij})$ então a matriz $(-A) = (-a_{ij})$, é chamada de **matriz oposta de A**.

Exemplo:

Para

$$A = \begin{pmatrix} 2 & 5 & -7 & 8 \\ -4 & 9 & 11 & -2 \\ 7 & -10 & 15 & 5 \end{pmatrix}$$

temos que

$$(-A) = \begin{pmatrix} -2 & -5 & 7 & -8 \\ 4 & -9 & -11 & 2 \\ -7 & 10 & -15 & -5 \end{pmatrix}.$$

2.1.1 Adição de matrizes

Sejam $A = (a_{ij})$ e $B = (b_{ij})$ matrizes $m \times n$. Indicaremos a **soma** de A com B por $A + B$, cujos termos são da forma $a_{ij} + b_{ij}$, ou seja:

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} & a_{14} + b_{14} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} & a_{24} + b_{24} & \dots & a_{2n} + b_{2n} \\ a_{31} + b_{31} & a_{32} + b_{32} & a_{33} + b_{33} & a_{34} + b_{34} & \dots & a_{3n} + b_{3n} \\ a_{41} + b_{41} & a_{42} + b_{42} & a_{43} + b_{43} & a_{44} + b_{44} & \dots & a_{4n} + b_{4n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & a_{m3} + b_{m3} & a_{m4} + b_{m4} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

A operação descrita acima, chama-se **adição de matrizes** que transforma cada par (A, B) de matrizes do mesmo tipo em uma única matriz $A+B$.

Na adição de matrizes valem as seguintes propriedades:

1. $A + (B + C) = (A + B) + C$ (*associativa*);
2. $A + B = B + A$ (*comutativa*);
3. Temos que a matriz nula O é tal que, $O + A = A$ (*existência do elemento neutro*);
4. A matriz $(-A)$ é tal que, $A + (-A) = O$ (*existência da oposta de uma matriz*).

2.1.2 Multiplicação de uma matriz por um número

Seja a matriz $A = (a_{ij})$ e α um número real, o **produto** de α por A é uma matriz real, $m \times n$, onde cada termo de A é multiplicado por α . Ou seja,

$$\alpha A = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \alpha a_{13} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \alpha a_{23} & \dots & \alpha a_{2n} \\ \alpha a_{31} & \alpha a_{32} & \alpha a_{33} & \dots & \alpha a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \alpha a_{m3} & \dots & \alpha a_{mn} \end{pmatrix}.$$

Para quaisquer números reais α e β , e quaisquer matrizes $A \in M_{m \times n}(\mathbb{R})$ e $B \in M_{m \times n}(\mathbb{R})$, temos as seguintes propriedades:

1. $(\alpha\beta)A = \alpha(\beta A)$;
2. $(\alpha + \beta)A = \alpha A + \beta A$;
3. $\alpha(A + B) = \alpha A + \alpha B$;
4. $1 A = A$.

Multiplicação de Matrizes

Sejam a matriz $A = (a_{ij})$, de ordem $m \times n$, e a matriz $B = (b_{ij})$, de ordem $n \times p$, o **produto** AB será uma matriz C de ordem $m \times p$, cujo termo c_{ij} será dado por:

$$c_{ij} = \sum_{k=1}^n a_{ik} \cdot b_{kj} = a_{i1} \cdot b_{1j} + \dots + a_{in} \cdot b_{nj}$$

Observações:

1. O produto entre duas matrizes ocorre apenas se o número de colunas da primeira matriz for igual ao número de linhas da segunda matriz.
2. o elemento c_{ij} (i-ésima linha e j-ésima coluna da matriz produto) será obtido multiplicando os elementos da i-ésima linha da primeira matriz pelos elementos correspondentes da j-ésima coluna da segunda matriz, e somando estes produtos.

Exemplos:

$$1. \begin{pmatrix} 1 & 3 & -5 \\ 0 & 4 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + 3 \cdot 5 + (-5) \cdot 7 \\ 0 \cdot 2 + 4 \cdot 5 + 2 \cdot 7 \end{pmatrix} = \begin{pmatrix} -18 \\ 34 \end{pmatrix}$$

$$2. \begin{pmatrix} 4 & 6 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 3 & 9 \\ 4 & -1 \end{pmatrix} = \begin{pmatrix} 4 \cdot 3 + 6 \cdot 4 & 4 \cdot 9 + 6 \cdot (-1) \\ 0 \cdot 3 + (-1) \cdot 4 & 0 \cdot 9 + (-1) \cdot (-1) \end{pmatrix} = \begin{pmatrix} 36 & 30 \\ -4 & 1 \end{pmatrix}$$

Algumas observações para a multiplicação entre matrizes:

1. Em geral, $AB \neq BA$.

Exemplo: Sejam $A = \begin{pmatrix} 2 & 0 \\ -1 & 3 \end{pmatrix}$ e $B = \begin{pmatrix} 4 & 1 \\ 0 & -2 \end{pmatrix}$.

Logo, $AB = \begin{pmatrix} 8 & 2 \\ -4 & -7 \end{pmatrix}$ e $BA = \begin{pmatrix} 7 & 3 \\ 2 & -6 \end{pmatrix}$.

2. $AI_n = I_n A = A$ para todo $A \in M_n(\mathbb{R})$ (admite elemento neutro para a multiplicação), onde I_n é a matriz de ordem n chamada de **matriz identidade**, isto é,

$$I_n = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Respeitando as ordens das matrizes necessárias para a multiplicação, temos que para as matrizes A , B e C , valem as seguintes propriedades:

3. $A(B + C) = AB + AC$ (distributividade à esquerda da multiplicação, em relação a soma);

4. $(A + B)C = AC + BC$ (distributividade à direita da multiplicação, em relação a soma);
5. $(AB)C = A(BC)$ (associatividade).

Além das matrizes e suas propriedades já citadas anteriormente, temos ainda as seguintes operações:

Transposta de uma matriz

Dada uma matriz $A = (a_{ij})$ real, $m \times n$, chamaremos de matriz transposta de A a matriz $A^t = (b_{ij})$, $n \times m$, cujas linhas são as colunas de A.

Exemplos:

i) Se

$$A = \begin{pmatrix} -3 & 2 & 0 \\ 4 & -1 & 8 \end{pmatrix}$$

então

$$A^t = \begin{pmatrix} -3 & 4 \\ 2 & -1 \\ 0 & 8 \end{pmatrix};$$

ii) Se

$$B = \begin{pmatrix} 1 & 5 \\ 5 & 8 \end{pmatrix}$$

então

$$B^t = \begin{pmatrix} 1 & 5 \\ 5 & 8 \end{pmatrix}$$

.

Uma matriz será chamada de **simétrica**, se e somente se, ela for igual a sua transposta. A matriz B do exemplo ii) acima é uma matriz simétrica.

Uma matriz A é dita **anti-simétrica** se $a_{ij} = -a_{ji}$, ou seja, $A^t = -A$.

Propriedades:

- i) A transposta da transposta de uma matriz é a própria matriz, ou seja $A^{tt} = A$.
- ii) A transposta de uma soma é a soma das transpostas, ou seja $(A + B)^t = A^t + B^t$.
- iii) $(\alpha A)^t = \alpha A^t$, onde α é qualquer escalar.

A inversa de uma matriz

Chamaremos de matriz inversa de A (sendo A uma matriz quadrada de ordem n), e denotaremos por A^{-1} , a matriz tal que $A \cdot A^{-1} = A^{-1} \cdot A = I_n$, onde I_n é a matriz identidade de ordem n .

Exemplo:

Se

$$A = \begin{pmatrix} 6 & 2 \\ 11 & 4 \end{pmatrix},$$

então

$$A^{-1} = \begin{pmatrix} 2 & -1 \\ \frac{-11}{2} & 3 \end{pmatrix}.$$

Pois,

$$AA^{-1} = \begin{pmatrix} 6 & 2 \\ 11 & 4 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ \frac{-11}{2} & 3 \end{pmatrix} = \begin{pmatrix} (6 \cdot 2 + 2 \cdot \frac{-11}{2}) & (6 \cdot -1 + 2 \cdot 3) \\ (11 \cdot 2 + 4 \cdot \frac{-11}{2}) & (11 \cdot -1 + 4 \cdot 3) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Observações:

- i) Como a matriz A do exemplo acima admite uma matriz inversa A^{-1} , dizemos que a matriz A é uma **matriz inversível**.
- ii) A matriz A^{-1} encontrada no exemplo acima é única.
- iii) Para determinação de uma matriz inversa são necessários alguns métodos que são apresentados em livros de álgebra linear, e estes métodos podem ser compreendidos facilmente por estudantes do ensino médio que tenham domínio sobre o conteúdo matrizes e suas operações. Não os apresentaremos aqui, pois fogem do objetivo deste trabalho.

Diagonalização de uma matriz

Uma matriz quadrada A é dita **matriz diagonal** quando todos os elementos $a_{ij} = 0$ se $i \neq j$, podendo os elementos a_{ij} para $i = j$ serem iguais a zero ou não.

Exemplos:

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Dadas duas matrizes quadradas, de ordem n , A e B , dizemos que a matriz B é **semelhante** a A , se existe uma matriz P invertível tal que

$$B = P^{-1}AP.$$

Exemplo:

Sejam

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} -4 & -4 \\ 1 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}.$$

Temos que $P^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$.

Logo, a matriz B é semelhante a matriz A , pois:

$$\begin{pmatrix} -4 & -4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}.$$

Uma matriz A , de ordem n , é uma **matriz diagonalizável**, se for semelhante a matriz diagonal. Isto significa que deve existir uma matriz P , tal que $P^{-1}AP$ é a matriz diagonal.

Exemplo:

Sejam

$$A = \begin{pmatrix} 1 & 0 \\ 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}, \quad P = \begin{pmatrix} 3 & 0 \\ 1 & 1 \end{pmatrix} \text{ e } P^{-1} = \begin{pmatrix} \frac{1}{3} & 0 \\ \frac{-1}{3} & 1 \end{pmatrix}.$$

Temos que a matriz A é diagonalizável pois é semelhante a matriz B (matriz diagonal):

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

Observação: Existem métodos para verificação da diagonalização de uma matriz, que podem ser encontrados em livros de álgebra.

2.2 Matrizes Estocásticas e uma versão do Teorema de Perron-Frobenius

Uma **matriz estocástica** ou matriz de probabilidade de transição é uma matriz quadrada cujas colunas são **vetores de probabilidade** (vetores com componentes não-negativos cuja soma é 1). Ou seja, para que uma matriz quadrada T seja uma matriz estocástica, ela deve satisfazer:

1. $T_{ij} \geq 0$, $i, j \in \{1, 2, 3, \dots, n\}$.
2. $\sum_{i=1}^n T_{i,j} = 1$, $j \in 1, 2, \dots, n$.

No contexto de cadeia de Markov, veremos que T_{ij} será uma probabilidade de transição entre os estados i e j (Capítulo 4).

Exemplo:

$$T = \begin{pmatrix} 1 & 0 & 0,15 & 0,25 \\ 0 & 0,2 & 0,11 & 0,32 \\ 0 & 0,3 & 0,34 & 0,13 \\ 0 & 0,5 & 0,4 & 0,3 \end{pmatrix}$$

Uma matriz estocástica T é **regular**, se existe $k, k \in \mathbb{N}$, tal que T^k tem todas suas entradas não-nulas.

Exemplo: Vamos verificar se a matriz estocástica

$$T = \begin{pmatrix} 0 & 1 & 0,1 \\ 0,2 & 0 & 0,4 \\ 0,8 & 0 & 0,5 \end{pmatrix}$$

é regular. Para isto, faremos:

$$T^2 = \begin{pmatrix} 0,28 & 0 & 0,45 \\ 0,32 & 0,2 & 0,22 \\ 0,4 & 0,8 & 0,33 \end{pmatrix}$$

e como nem todas suas entradas são não-nulas, não podemos afirmar, ainda, que a matriz T é regular. Mas, ao continuar o processo temos

$$T^3 = \begin{pmatrix} 0,36 & 0,28 & 0,253 \\ 0,216 & 0,32 & 0,222 \\ 0,424 & 0,4 & 0,525 \end{pmatrix}.$$

Portanto, T é uma matriz estocástica regular.

Observações:

- Nem toda matriz estocástica é regular. Um exemplo é a matriz

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

Temos que A é estocástica, porém, não é regular, pois $A^{2k} = I$ e $A^{2k+1} = A, \forall k \geq 0$.

- Em geral, se T é uma matriz estocástica, então T^k também é estocástica, $\forall k \geq 1$ (verificaremos este fato na proposição seguinte).

Proposição 3. *Produto de matrizes estocásticas é estocástica.*

Demonstração. Devemos mostrar que se $A = (a_{i,j})_{m \times m}$ e $B = (b_{i,j})_{m \times m}$ são matrizes estocásticas, então AB também é estocástica. Para isto, observe que cada entrada i, j da matriz AB é dada por

$$(AB)_{i,j} = \sum_{n=1}^m a_{in} b_{nj}$$

e somando os elementos da j -ésima coluna, obtemos

$$\begin{aligned} \sum_{i=1}^m (AB)_{i,j} &= \sum_{i=1}^m \left(\sum_{j=1}^m a_{i,n} b_{n,j} \right) \\ &= \sum_{n=1}^m \underbrace{\left(\sum_{i=1}^m a_{i,n} \right)}_1 b_{n,j} \\ &= \sum_{n=1}^m b_{nj} \\ &= 1 \end{aligned}$$

■

Teorema 2. *Se T é uma matriz estocástica regular $r \times r$ então:*

- T^n se aproxima de uma matriz M , no sentido de que cada entrada da matriz T^n aproxima-se da entrada correspondente em M ;
- Todas as colunas de M são iguais, sendo dadas por um vetor coluna

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_r \end{bmatrix},$$

com $w_i > 0$, para $i = 1, \dots, r$.

iii) Para qualquer vetor de probabilidades inicial

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_r \end{bmatrix},$$

o vetor de probabilidades $T^n v$ aproxima-se de w , quando $n \rightarrow \infty$;

iv) O vetor w é o único que satisfaz $Tw = w$ e é chamado vetor estacionário.

A demonstração do teorema acima encontra-se em Silva e Rota (2016, p.7-12), e será omitida neste trabalho pois foge dos nossos objetivos.

Porém, faremos aqui algumas observações sobre o teorema acima citado.

Observação 1. O teorema é utilizado para o cálculo de previsões a longo prazo, para o caso onde a matriz de transição é regular, o qual utilizaremos na Seção 5.4. Com ele pode-se perceber que a previsão a longo prazo não dependerá do vetor de probabilidade inicial (item iii).

Observação 2. Este resultado é uma versão do conhecido Teorema de Perron-Frobenius.

Observação 3. No contexto de Cadeia de Markov, T será a matriz de transição (Capítulo 4). Para encontrarmos uma previsão w a longo prazo, basta conhecermos a matriz das probabilidades de transição T e resolvermos o sistema linear $Tw = w$. Sendo assim, é possível explorar com alunos do Ensino Médio este assunto, já que matrizes, probabilidades e sistemas lineares são trabalhados nesta etapa de ensino.

Uma aplicação deste Teorema será utilizado no Exemplo 10 do Capítulo 4.

Capítulo 3

Processos Estocásticos

3.1 Álgebra de eventos

Considere o lançamento de uma moeda não viciada no tempo $n = 1, 2, 3, \dots, T$. As faces da moeda são: cara(C) ou coroa(K). Denotando por Ω o conjunto de todos os possíveis resultados dos lançamentos até o tempo T , temos:

$$\Omega = \{\omega : \omega = (e_1, e_2, \dots, e_T)\} \quad \text{onde, } e_n = C \text{ ou } K, \text{ para } n = 1, 2, \dots, T.$$

O conjunto $E = \{C, K\}$ que representa os possíveis estados da moeda (possíveis resultados em cada lançamento) será chamado de **espaço de estados**. Com o decorrer do tempo, várias informações são reveladas sobre os estados que já ocorreram no experimento do lançamento da moeda. No tempo $n = 2$, por exemplo, se nós conhecemos os resultados, vamos supor que sejam e_1 e e_2 , então a informação obtida é

$$A = \{(e_1, e_2, *, \dots, *)\},$$

onde “*” indica os estados que ainda poderão ocorrer com o decorrer do tempo.

Denotaremos por F_n a **álgebra de eventos** gerada por A , à qual dá a ideia de informação disponível no tempo n , que consiste nos resultados que ocorreram antes e durante o tempo n . No tempo $n = 0$, antes da realização do experimento, não temos

informações e, portanto, $F_0 = \{\emptyset, \Omega\}$.

Exemplo:

Para o caso $n = 1$, vamos supor que o resultado obtido foi cara (C). Tomemos

$$A = \{(C, e_2), e_2 = C \text{ ou } K\} = \{(C, C), (C, K)\}.$$

sendo F_1 a álgebra de eventos para $n = 1$, segue que

$$F_1 = \{\emptyset, \Omega, A, A^c\}.$$

Note que $F_0 \subset F_1$.

Observação: Lembramos que F_n é uma álgebra de subconjuntos de Ω , se possui as seguintes propriedades:

P1. $\Omega \in F_0 \subset F_n$, para $n = 1, 2, \dots, T$.

P2. Se $A \in F_n$, então $A^c \in F_n$.

P3. Se $A \in F_n$ e $B \in F_n$, então $A \cup B \in F_n$.

Alguns exemplos de álgebras de eventos são:

1. $F_0 = \{\Omega, \emptyset\}$ (álgebra trivial).
2. $\{\emptyset, \Omega, A, A^c\}$ (denotada por F_A é a álgebra gerada por A).
3. $\{A : A \in \Omega\}$ (denotada por 2^Ω é a álgebra de todos os subconjuntos de Ω).

Se F é uma álgebra de eventos, então todo conjunto de F é chamado de conjunto mensurável. Se $F = 2^\Omega$, por exemplo, então todo subconjunto de Ω é mensurável.

Definição 8. Uma filtração \mathbb{F} é a coleção de álgebras de eventos F_n , com

$$\mathbb{F} = \{F_0, F_1, \dots, F_n, \dots, F_T\}; \text{ com } F_n \subset F_{n+1}.$$

Utilizamos a filtração \mathbb{F} para modelar um fluxo de informações, obtidas com o decorrer dos experimentos. Com o passar do tempo, obtemos informações mais detalhadas, isto é, partições de Ω cada vez mais refinadas.

Exemplo 6. Como F_0, F_A e 2^Ω são álgebras de eventos e $F_0 \subset F_A \subset 2^\Omega$, temos que $\mathbb{F} = \{F_0, F_A, 2^\Omega\}$ é uma filtração.

3.2 Processo estocástico

Vimos que, se a função X é uma variável aleatória em (Ω, F) , com Ω sendo um conjunto finito, então todos os conjuntos $\{X = x_i\} = \{\omega \in \Omega : X(\omega) = x_i\}$, $i = 1, \dots, k$ pertencem a F . Isto significa que, se conhecemos qual evento da forma $\{X = x_i\}$ ocorreu em F (informação descrita por F), então sabemos qual valor de X ocorreu. Lembrando que se $F = 2^\Omega$, então qualquer função em Ω é uma variável aleatória.

A fim de definir processo estocástico, daremos o seguinte exemplo.

Exemplo 7. *Considere o exemplo do lançamento de uma moeda não viciada, para $n = 1, 2$. Então $\Omega = \{\omega_1 = (C, C), \omega_2 = (C, K), \omega_3 = (K, C), \omega_4 = (K, K)\}$.*

Tome $A = \{\omega_1, \omega_2\}$, o qual representa o evento em que no tempo $n = 1$ o resultado obtido foi cara. Assim, $F_1 = \{\emptyset, \Omega, A, A^c\}$ e tome $F_2 = 2^\Omega$.

Considere as seguintes funções em Ω :

$$a) X(\omega_1) = X(\omega_2) = 15 \text{ e } X(\omega_3) = X(\omega_4) = 5.$$

X é uma variável aleatória com relação F_1 , já que $\{\omega : X(\omega) = 15\} = \{\omega_1, \omega_2\} = A \in F_1$ e $\{\omega : X(\omega) = 5\} = \{\omega_3, \omega_4\} = A^c \in F_1$.

$$b) Y(\omega_1) = 15, Y(\omega_2) = 75, Y(\omega_3) = 75, \text{ e } Y(\omega_4) = 5.$$

Y não é uma variável aleatória com relação a F_1 , pois $\{\omega : Y(\omega) = 75\} = \{\omega_2, \omega_3\} \notin F_1$. Por outro lado, é claro que Y é uma variável aleatória em F_2 .

Definição 9. *Um processo estocástico adaptado a filtração $\mathbb{F} = \{F_1, \dots, F_T\}$ é uma sequência de variáveis aleatórias X_n tal que cada X_n é uma variável aleatória com relação a (Ω, F_n) para cada $n = 1, 2, 3, \dots, T$.*

Observe que, para cada n , X_n é uma variável aleatória em (Ω, F_T) .

Exemplo 8. *Seja $X_1 = X$ e $X_2 = Y$ como apresentados no Exemplo 7. A sequência $\{X, Y\}$ é um processo estocástico adaptado a filtração $\mathbb{F} = \{F_1, F_2\}$.*

Definição 10. *Sejam $(\Omega, 2^\Omega)$ um espaço amostral com a álgebra de todos os eventos, e X uma variável aleatória com valores x_i , $i = 1, 2, \dots, k$. Considere os conjuntos*

$$A_i = \{\omega : X(\omega) = x_i\} \subseteq \Omega,$$

que formam uma partição de Ω . A álgebra gerada por essa partição é chamada álgebra gerada por X . Esta é a menor álgebra que contém todos os conjuntos da forma $A_i = \{X = x_i\}$ e será denotada por F_X , representando a informação que pode ser extraída da observação de X .

Exemplo 9. Considerando o Exemplo 7, temos $\{\omega : X(\omega) = 15\} = \{\omega_1, \omega_2\} = A$ e $\{\omega : X(\omega) = 5\} = \{\omega_3, \omega_4\} = A^c$. Logo, $F_X = F_1 = \{\emptyset, \Omega, A, A^c\}$.

3.3 Probabilidade no espaço das sequências

Nesta seção daremos uma breve ideia de como podemos construir o espaço de probabilidade, para uma sequência infinita (enumerável) de variáveis aleatórias.

O conjunto Ω pode ser interpretado como o conjunto de todas as sequências formadas de elementos do espaço de estados E , ou seja, estamos levando em conta os valores que as variáveis aleatórias assumem para interpretar Ω . Desta forma, um elemento $\omega \in \Omega$ pode então ser escrito da forma

$$\omega = (e_0, e_1, e_2, \dots),$$

onde cada $e_i \in E$. Assim, cada ponto $\omega \in \Omega$ é chamado de caminho ou trajetória, o espaço Ω pode então ser considerado o espaço das sequências, e o valor e_n em um caminho é chamado “n-ésimo resultado em ω ”. Por exemplo, se $E = \{1, 2\}$, então $\Omega = \{(e_0, e_1, e_2, e_3, \dots); e_i = 1 \text{ ou } 2\}$ e o elemento $(1, 2, 1, 1, 1, 2, \dots) \in \Omega$, por exemplo, é um caminho formado por uma sequência de experimentos.

A função $X_n : \Omega \rightarrow E$, dada por

$$X_n(e_0, e_1, e_2, \dots) = e_n,$$

é chamada de *função saída* ou *avaliação* da trajetória ω .

Fixado n , seja F_n a família de todas as uniões de Ω da forma

$$\{\omega \in \Omega : X_0(\omega) \in \mathcal{E}_0, X_1(\omega) \in \mathcal{E}_1, \dots, X_n(\omega) \in \mathcal{E}_n\},$$

onde $\mathcal{E}_0, \mathcal{E}_1, \dots, \mathcal{E}_n$ são subconjuntos do espaço de estados E . Cada F_n é uma σ -álgebra ¹ e além disso, $F_n \subset F_{n+1}$ forma uma filtração natural na qual o processo X_n é adaptado.

Seja \mathcal{F} a família de conjuntos definida por:

$$\mathcal{F} = \bigcup_{n=0}^{\infty} F_n.$$

Cada elemento em \mathcal{F} é um conjunto de trajetórias para as quais um número finito de entradas das sequências são restritas a pertencer a certos subconjuntos de E , e as demais infinitas entradas são irrestritas. Um conjunto de \mathcal{F} é chamado de *cilindro*. \mathcal{F} é uma álgebra, no entanto, não é uma σ -álgebra. Porém, existe a menor σ -álgebra \mathcal{G} , tal que $\mathcal{F} \subset \mathcal{G}$. Associado a \mathcal{G} existe uma única medida de probabilidade μ , tal que

$$\mu(C_i^m) = \mu\{\omega \in \Omega : X_0(\omega) = e_0, X_1(\omega) = e_1, \dots, X_n(\omega) = e_n\}$$

é dado pelo produto das probabilidades condicionais entre os estados e_0, e_1, \dots, e_n . Cada conjunto C_i^m é chamado de *cilindro básico* de \mathcal{F} e vale notar que cada $C \in F_n$ é da forma $C = \bigcup_i C_i^m$. Ou seja, construímos um espaço de probabilidade $(\Omega, \mathcal{G}, \mu)$, onde μ é dada por

$$\mu(e_0, e_1, \dots, e_n, *, *, *, \dots) = p_0 p_{10} p_{21} \dots p_{n(n-1)},$$

onde p_0 é a probabilidade do processo iniciar no estado e_0 e $p_{ij} = \mu\{\omega \in \Omega : X_i = e_i | X_j = e_j\}$ são probabilidades condicionais. Maiores detalhes podem ser vistos em Brzezniak e Zastawniak (2009), Oliveira, Ribeiro e Silva (2017), Ribeiro (2017) ou Ruffino (2009).

¹A verificação foge dos objetivos do trabalho, por este motivo será omitida.

Capítulo 4

Cadeias de Markov

Uma cadeia de Markov é um processo estocástico $(X_n)_{n \in \mathbb{N}}$, que satisfaz uma determinada propriedade chamada de Markoviana (ver equação (4.2)), que iremos apresentar e explicar mais adiante neste capítulo. Daremos a seguir algumas definições e propriedades básicas das probabilidades de transição.

4.1 Probabilidade de transição

Definição 11. *Definimos a probabilidade de transição em um passo, e denotamos por $p_{ij}(k)$, a probabilidade condicional*

$$p_{ij}(k) = \mathbb{P}(X_{k+1} = i | X_k = j).$$

Traduzindo, é a probabilidade de estar no estado i no tempo $k + 1$, dado que estava no estado j no momento anterior k , para os estados $i, j = 1, 2, \dots$ (estamos assumindo que o espaço de estados E é finito ou enumerável).

Dizemos que uma probabilidade de transição $p_{ij}(k)$ numa cadeia é *homogênea* se ela não depende do tempo k , ou seja, $p_{ij}(n) = p_{ij}(k)$ mesmo para $k \neq n$. Neste caso usaremos simplesmente a notação p_{ij} . Caso as probabilidades de transição dependam do tempo k , dizemos que elas são *não-estacionárias* ou *não homogêneas*. Estudaremos aqui apenas os casos homogêneos.

Para uma cadeia de Markov com um número finito de estados, associa-se uma matriz de transição T , que será dada pelas probabilidades de transição, isto é,

$$T = (p_{ij}).$$

Definição 12. A probabilidade de transição em n -passos ($n \geq 0$), denotada por $p_{ij}^{(n)}$, é a probabilidade de transferência do estado j para o estado i em n etapas de tempo discreto, isto é,

$$p_{ij}^{(n)} = \mathbb{P}(X_n = i | X_0 = j).$$

Para uma quantidade de estados finitos, associa-se uma matriz $T^{(n)}$, onde a ij -ésima posição é dada por $p_{ij}^{(n)}$. Observe que $T^{(0)}$ é a matriz identidade, uma vez que $p_{ii}^{(0)} = 1$ e $p_{ij}^{(0)} = 0$ quando $i \neq j$.

As equações de Chapman-Kolmogorov (para E finito ou enumerável), exibidas abaixo, mostram a relação entre as probabilidades de transição em n -passos, s -passos e $(n - s)$ -passos:

$$p_{ij}^{(n)} = \sum_{k=1}^{\infty} p_{ik}^{(n-s)} p_{kj}^{(s)}, \quad 0 < s < n.$$

Para maiores detalhes sobre a validade desta propriedade, ver Brzezniak e Zastawniak (1999, p.92).

Escrevendo-as, em termos matriciais (para E finito) temos que

$$T^{(n)} = T^{(n-s)}T^{(s)}.$$

Como $T^{(1)} = T$, segue-se então que

$$\begin{aligned} T^{(2)} &= T^{(2-1)}T^{(1)} \\ &= T^{(1)}T^{(1)} \\ &= TT \\ &= T^2. \end{aligned}$$

Repetindo este processo teríamos $T^{(3)} = T^3$, $T^{(4)} = T^4, \dots$ e assim sucessivamente,

concluindo que $T^{(n)} = T^n$, para todo $n \geq 0$. Portanto, pode-se obter as probabilidades $p_{ij}^{(n)}$ por meio da matriz T^n ($T \cdot T \cdot T \dots$ n-vezes).

Outro aspecto interessante em conhecer T^n , é que o vetor v^n de distribuição de probabilidade do processo, no tempo n , é igual ao produto matricial $T^n v^0$, com a distribuição inicial v^0 escrita de forma transposta. Ou seja,

$$v^n = T^n v^0. \quad (4.1)$$

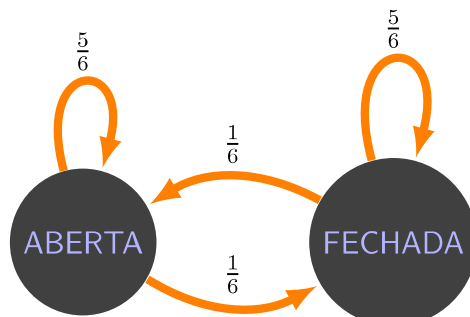
Vejamos um exemplo, para compreensão deste fato.

Exemplo: *Uma criança brinca em seu quarto abrindo e fechando a porta do seu guarda-roupa, lançando sempre um dado a cada minuto. Considere a porta com dois estados: aberta e fechada. Ela lança o dado, se o resultado do lançamento mostrar face 6, ela inverte o estado da porta, caso contrário, deixa como está. Suponha que no início do experimento, a porta está aberta.*

O estado da porta em função do tempo é um processo de Markov. Através deste exemplo podemos explicar o significado: “não tem nenhuma memória”. Quando conhecemos o estado da porta no tempo n , podemos prever sua evolução, em termos de variável aleatória, para todos os tempos futuros, sem exigir conhecimento sobre o estado da porta em tempos menores que n . Traduzindo, o futuro do processo depende do presente, mas não do passado.

O diagrama de transição para este exemplo é apresentado na Figura 4.1. Ele nos mostra que, se a porta estiver aberta, a probabilidade de se manter neste estado é de $\frac{5}{6}$, enquanto que a probabilidade dela ser fechada é de $\frac{1}{6}$. E caso a porta esteja fechada, a probabilidade de se manter fechada é de $\frac{5}{6}$, enquanto que a probabilidade dela ser aberta é de $\frac{1}{6}$.

Figura 4.1: Diagrama de transição do exemplo da porta.



Fonte: O próprio autor (2018).

Denotemos por $x_1^{(n)}$ a probabilidade da porta estar aberta no tempo n e, similarmente, $x_2^{(n)}$ a probabilidade da porta estar fechada no tempo n . Temos então, a seguinte relação de recorrência, obtida à partir das informações anteriores:

$$x_1^{(n+1)} = \frac{5}{6}x_1^{(n)} + \frac{1}{6}x_2^{(n)}, \quad x_2^{(n+1)} = \frac{1}{6}x_1^{(n)} + \frac{5}{6}x_2^{(n)},$$

com $x_1^{(0)} = 1$ e $x_2^{(0)} = 0$, já que assumimos que a porta inicia o experimento aberta. A primeira igualdade vem do fato de que a porta ficará aberta no tempo $n+1$, se ela estava aberta no tempo n e o dado não mostrou a face 6, ou ela estava fechada no tempo n e o dado mostrou a face 6. A segunda igualdade é análoga.

A equação anterior pode ser escrita na forma matricial:

$$v^{(n+1)} = \begin{pmatrix} x_1^{(n+1)} \\ x_2^{(n+1)} \end{pmatrix} = \begin{pmatrix} \frac{5}{6}x_1^{(n)} + \frac{1}{6}x_2^{(n)} \\ \frac{1}{6}x_1^{(n)} + \frac{5}{6}x_2^{(n)} \end{pmatrix} = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \end{pmatrix} \cdot \begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \end{pmatrix} = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \end{pmatrix} \cdot v^{(n)}.$$

Isto é, $v^{n+1} = Tv^n$ e, assim, $v^1 = Tv^0$, $v^2 = Tv^1 = T(Tv^0) = T^2v^0$. E, continuando este processo, chegamos na Equação (4.1).

Consideremos ainda o exemplo da criança que brinca com a porta. Contudo, desta vez, o estado da porta será mudado somente se o dado mostrar face 6 e não ter ocorrido

a face 6 no tempo anterior. Assim, o futuro deste processo dependerá do que ocorreu no passado e, portanto, não é um processo de Markov, pois o estado da porta depende do resultado do lançamento do dado nos tempos anteriores.

Formalmente temos o seguinte:

Definição 13. *Considere um espaço de estados com um número finito (ou enumerável) de elementos $E = \{e_1, \dots, e_n\}$. Um processo estocástico discreto $(X_n)_{n \in \mathbb{N}}$ é **uma cadeia (ou processo) de Markov** se a probabilidade condicional satisfizer*

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n), \quad (4.2)$$

para todo $n \geq 1$ e para toda sequência x_0, x_1, \dots, x_{n+1} de elementos do espaço de estados E . A condição (4.2) indica que o futuro do processo, uma vez conhecido o estado presente, é independente do passado.

Como já mencionamos anteriormente, probabilidades condicionais da forma,

$$p_{ij} = \mathbb{P}(X_{n+1} = e_i | X_n = e_j),$$

são as chamadas probabilidades de transição. E, se para cada i, j ,

$$\mathbb{P}(X_{n+1} = e_i | X_n = e_j) = \mathbb{P}(X_1 = e_i | X_0 = e_j),$$

para todo natural n , diremos que a cadeia de Markov é estacionária ou homogênea.

Assim que especifica-se as probabilidades de transição e a distribuição inicial de probabilidades dos estados, um processo de Markov está completamente definido.

A propriedade (4.2) será referida como propriedade da cadeia de Markov, e esta pode ser estendida para processos de Markov com tempo contínuo e espaço de estados sendo um intervalo da reta real, o que não faremos neste trabalho pois foge do escopo.

Daremos a seguir um exemplo em que ilustra como podemos usar o Teorema 2 no estudo da convergência das distribuições de probabilidade para uma cadeia finita, para um caso particular em que cujas probabilidades de transição são todas positivas.

Iremos considerar um modelo que descreve a dinâmica de um sapo que salta entre duas pedras à beira de um rio. Este modelo poderia ser aplicado em diversas situações. Poderia por exemplo, ser uma partícula de elétrons que salta em duas regiões do átomo.

Exemplo 10. *Suponha que um sapo, à beira de um rio, salta em tempos discretos $n = 0, 1, 2, \dots$ entre duas pedras, representadas pelos estados **1** e **2**, com as condições dadas a seguir:*

- *Se o sapo estiver na pedra **1** no tempo n , ele saltará para **2** com probabilidade p , $0 < p < 1$;*
- *Se o sapo estiver na pedra **2** no tempo n , ele saltará para **1** com probabilidade q , $0 < q < 1$;*
- *No instante inicial ($n=0$) o sapo se encontra na pedra **1**.*

*Seja α_n a probabilidade do sapo estar na pedra **1** no tempo n , e seja $(1 - \alpha_n)$ a probabilidade do sapo estar na pedra **2** no tempo n . O que acontece com α_n quando n cresce?*

Para responder a esta pergunta, observe que:

- *A distribuição do sapo no tempo n , em notação vetorial, é dada por $v_n = (\alpha_n, 1 - \alpha_n)$, sendo $v_0 = (1, 0)$ a distribuição inicial;*
- *Os elementos da matriz de transição são dados por:*

$$a_{11} = \mathbb{P}(A_{n+1}|A_n) = 1 - p, \quad a_{12} = \mathbb{P}(A_{n+1}|B_n) = q,$$

$$a_{21} = \mathbb{P}(B_{n+1}|A_n) = p, \quad a_{22} = \mathbb{P}(B_{n+1}|B_n) = 1 - q,$$

*onde A_n é o evento em que o sapo está na pedra **1** e B_n é o evento em que o sapo está na pedra **2**.*

Logo,

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

e, portanto, T é uma matriz regular.

Para encontrarmos o $\lim_{n \rightarrow \infty} \alpha_n$, segundo o Teorema 2, devemos resolver o sistema:

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

onde $\alpha = \lim_{n \rightarrow \infty} \alpha_n$ é a probabilidade do sapo estar em **1** e $1-\alpha$ é a probabilidade dele estar em **2** a longo prazo.

Temos:

$$\begin{cases} (1-p)\alpha + q(1-\alpha) = \alpha \\ p\alpha + (1-q)(1-\alpha) = 1-\alpha \end{cases} \implies \begin{cases} (-p-q)\alpha + q = 0 \\ (p+q)\alpha - q = 0 \end{cases}$$

Obtemos então

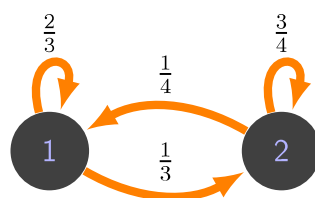
$$\alpha = \frac{q}{p+q}.$$

Portanto, o

$$\lim_{n \rightarrow \infty} \alpha_n = \frac{q}{p+q}.$$

Tomaremos como exemplo $p = \frac{1}{3}$ e $q = \frac{1}{4}$.

Figura 4.2: Diagrama de transição do exemplo do sapo.



Fonte: O próprio autor (2018).

A Figura 4.2 nos mostra que, se o sapo estiver na pedra 1, a probabilidade de que ele se mantenha nesta mesma pedra é de $\frac{2}{3}$, enquanto que a probabilidade do sapo saltar

para a pedra 2 é de $\frac{1}{3}$. Caso o sapo esteja na pedra 2, a probabilidade dele continuar na pedra 2 é de $\frac{3}{4}$, enquanto que a probabilidade dele saltar para a pedra 1 é de $\frac{1}{4}$. Portanto, temos que a distribuição será dada por

$$w = \left(\frac{\frac{1}{4}}{\frac{1}{3} + \frac{1}{4}}, 1 - \frac{\frac{1}{4}}{\frac{1}{3} + \frac{1}{4}} \right) \approx (0.4286, 0.5714).$$

Ou seja, a longo prazo teremos aproximadamente 42,86% de chance do sapo estar em 1 e aproximadamente 57,14% de chance do sapo estar em 2.

Capítulo 5

O problema do gás de Ehrenfest

Apresentaremos à seguir, a biografia de Paul Ehrenfest que propôs um modelo probabilístico para o problema da irreversibilidade dos gases. A biografia apresentada na Seção 5.1, utilizou-se como referência o texto *Physics Biographies: Paul Ehrenfest* (2018), no qual foi feita uma tradução livre.

Nas seções seguintes apresentaremos o problema da irreversibilidade dos gases e o modelo da urna elaborado por Paul Ehrenfest, tendo Friedli (2011) como referência.

5.1 Biografia de Paul Ehrenfest

Figura 5.1: Físico Austríaco Paul Ehrenfest (1880-1933)



Fonte: WIKIPEDIA. Acesso em: janeiro de 2018. ¹

Nascido em 18 de janeiro de 1880 na cidade de Viena na Áustria, **Paul Ehrenfest** era o mais novo dos cinco filhos do casal Sigmund e Johanna Ehrenfest.

Paul desenvolveu interesse nas áreas de Ciências e Matemática, vindo a formar-se em Física Teórica. No ano de 1904 tornou-se doutor com uma dissertação sobre a extensão da mecânica de Hertz, onde teve a supervisão de Ludwig Boltzmann (1844-1906), que inspirou Paul em seus trabalhos. Ainda no ano de 1904 casa-se com a estudante russa de matemática Tatyana Alexeyevna Afanassjewa e tem 4 filhos: Tatyana (matemática e física), Galinka, Paul Jr (físico) e Wassik.

O casal Ehrenfest colaborou com diversos assuntos sobre mecânica estatística de Josiah Willaerd Gibbs (1839-1903) e Boltzmann, através de seus artigos nos primeiros anos de casados. Após anos batalhando para ocupar uma cadeira como professor em uma Universidade, Paul Ehrenfest conquistou uma vaga na cidade de Leiden na Holanda, onde se tornaria sucessor de Hendrik A. Lorentz (1853-1928).

Paul revigorou a universidade ao estabelecer sala de leitura para os estudantes de física, reviver um clube de ciências, colóquios semanais, além de seus esforços para manter o contato intelectual e humano entre os membros da comunidade científica. Foi um notável professor, ganhando elogios de Einstein (1879-1955), que o descreveu como “o melhor professor que já conhecera” e de seus alunos que o apelidaram de “tio Sócrates” por sempre colaborar e ajudar para que cada um desenvolvesse seu próprio talento.

No ano de 1907, Ehrenfest propôs um modelo bem simples, o chamado **o modelo da urna de Ehrenfest** (que será apresentado neste trabalho) no qual mostrava como as leis da probabilidade poderiam produzir uma tendência média em direção ao equilíbrio, embora o comportamento desse modelo fosse reversível no tempo e cada um de seus estados acabariam por se repetir. E isto significava que o Teorema de Boltzmann, onde mostra que colisões moleculares produzirão uma abordagem de equilíbrio com a entropia aumentando monotonamente no tempo, quando interpretadas de forma correta não contradizem as leis reversíveis da mecânica, como argumentava o físico Loschmidt (1821-1895), ou o Teorema de Poincaré, como argumentava matemático Ernst Zermelo (1871-1953).

No ano de 1911, o casal Ehrenfest mostrou em um artigo a estrutura lógica e todas as dificuldades que tinham remanescido desta teoria: fizeram uma clara distinção entre a

abordagem mais antiga (antes de 1877), onde tratavam as moléculas estaticamente, porém tentaram fazer declarações universais válidas sobre o gás como um todo, e posteriormente em um novo trabalho, no qual o próprio gás foi tratado por métodos estatísticos. O papel da hipótese ergódica, relacionar as médias do tempo com as médias em relação a um conjunto foi evidenciado de forma tão clara que, atraiu a atenção dos matemáticos para o problema ergódico. O casal também formulou uma sequência de teoremas que ainda precisavam ser provados antes que o fundamento estatístico da segunda lei da termodinâmica fosse estabelecido firmemente. Porém, encontraram falhas quanto ao tratamento da irreversibilidade e subestimaram a importância dos métodos de conjunto de Gibbs ao lidar com sistemas complexos.

Esta abordagem crítica levou Ehrenfest a sua maior contribuição para a física: o princípio adiabático; que foi muito utilizado e altamente apreciado como um dos poucos guias confiáveis para progredir durante os anos que vigoravam a “antiga teoria quântica”.

Já no final dos anos 20 e início da década de 30, Ehrenfest fez o possível para assegurar a compreensão da nova mecânica quântica e ressaltar suas relações com a física clássica. Provou ainda um resultado que passou a ser chamado de Teorema de Ehrenfest, onde os valores das coordenadas quânticas e momentum obedecem as equações clássicas de movimento. Os seus últimos trabalhos consistiram em uma série de questões fundamentais sobre os aspectos físicos e matemáticos da mecânica quântica.

A contribuição de Ehrenfest para o desenvolvimento da física foi ainda maior devido sua influência sobre outros físicos, em particular, na sua última década de vida. Diversos físicos frequentavam sua casa, tanto que Einstein e Niels Bohr eram seus amigos íntimos, onde conversavam principalmente sobre as ideias fundamentais da física quântica.

Apesar de todo o seu comprometimento e êxito com o seu trabalho na Física e como professor, Ehrenfest sofria de depressão, e isto o levou a um terrível ato no ano de 1933: atirou contra o seu filho mais novo, Wassik, que tinha síndrome de Down e na sequência se suicidou, colocando fim a sua brilhante carreira como físico. Wassik sobreviveu, porém ficou cego de um olho.

5.2 Introdução ao problema do gás de Ehrenfest

Dois compartimentos A e B formam um sistema e estão ligados apenas por uma pequena abertura. Inicialmente, o compartimento A contém N partículas de um gás enquanto que o compartimento B está vazio (na Figura 5.2 $N = 26$). Em cada instante de tempo uma única partícula troca de compartimento, se a partícula estiver no compartimento A vai para o compartimento B e vice-versa (a Figura 5.3 representa o sistema no tempo $n = 1$, onde uma das 26 partículas passou para o compartimento B), sendo que todas as partículas de um mesmo compartimento possuem a mesma probabilidade de passar pela abertura. O sistema evolui, regido pelas leis da mecânica e da termodinâmica, até que depois de certo tempo atinge um estado de equilíbrio macroscópico, quando os compartimentos A e B possuem pressões iguais, e um estado de equilíbrio microscópico, quando o número de partículas nos compartimentos A e B correspondem a $N_A \simeq \frac{N}{2}$ e $N_B \simeq \frac{N}{2}$ (na Figura 5.4, temos $\frac{N}{2} = 13$), onde N_A é o número de partículas no compartimento A , N_B é o número de partículas no compartimento B e N é fixo, da ordem no número de Avogadro:

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

No início do experimento tínhamos $N_A = N$ e $N_B = 0$. Suponhamos agora, que o gás esteja em equilíbrio. Como o número de partículas é finito, é possível que após algum tempo do início do experimento cada partícula retorne ao seu compartimento inicial? Qual a probabilidade disto ocorrer? Quanto tempo levaria para as partículas voltarem ao seu estado inicial?

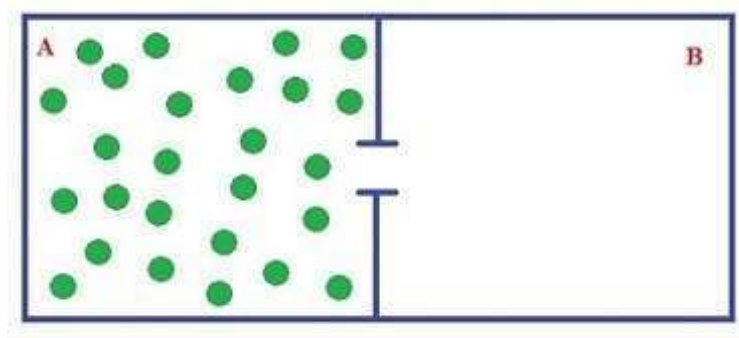
Esta última pergunta foi um dos principais dilemas dos físicos do século XX, pois a Segunda Lei da Termodinâmica diz que quando um gás atinge o seu estado de equilíbrio ele não pode mais voltar ao seu estado inicial, e por outro lado as leis da Mecânica Clássica são reversíveis com respeito ao tempo (se todas as velocidades das partículas do gás são invertidas, então o gás deve voltar ao seu estado inicial). O que contradiz a Segunda Lei da Termodinâmica.

As questões levantadas são clássicas da Física, mais precisamente da área de *mecânica estatística*, que busca estudar o comportamento de sistemas macroscópicos compostos

por um número elevado de partículas, relacionando a Termodinâmica com a Mecânica Clássica, cuja relação não abordaremos neste trabalho, pois foge dos nossos objetivos. Porém, iremos apresentar um modelo simples, porém, nos permitirá responder aos questionamentos acima.

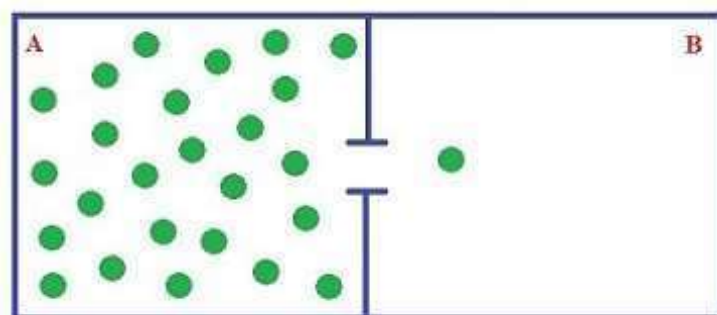
Daremos a seguir algumas ilustrações para o problema que acabamos de descrever:

Figura 5.2: Momento inicial do experimento com o gás.



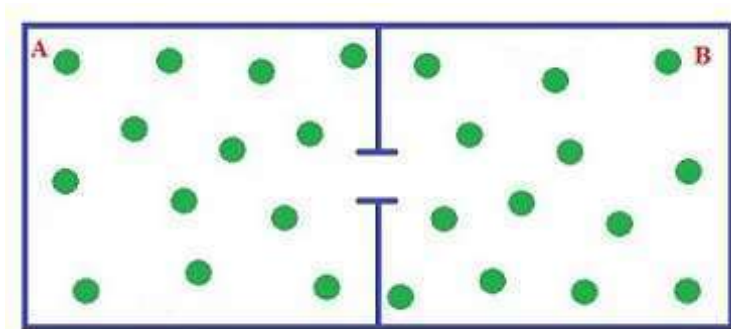
Fonte: O próprio autor (2018).

Figura 5.3: Experimento no tempo igual à 1.



Fonte: O próprio autor (2018).

Figura 5.4: Experimento em equilíbrio termodinâmico.



Fonte: O próprio autor (2018).

Assumiremos as seguintes simplificações em nosso modelo:

- as partículas são indistinguíveis e em quantidade N (N é fixo) e o *estado do sistema* será definido pela variável

$X_n :=$ número de partículas no compartimento B no instante n ;

- o que nos interessa é o número de partículas em cada um dos compartimentos. Suas velocidades ou posições serão desprezadas;
- consideraremos apenas os instantes de tempo discretos, no qual os indicaremos pela letra n . Logo, $n = 0, 1, 2, 3, 4, \dots$ indica o tempo em que a partícula muda de compartimento.

Como o número de partículas no compartimento B pode ser zero, um, dois, três, etc... chegando a no máximo N , temos que $X_n \in E$, onde

$$E := \{0, 1, 2, 3, \dots, N\} \subseteq \mathbb{R},$$

é o **espaço dos estados**. Note que o índice n em X_n é um parâmetro que descreve a variável tempo, enquanto N , que é fixo, é o número de partículas do sistema. Iremos supor que no tempo $n = 0$, a condição inicial é dada por

$$X_0 = 0,$$

ou seja, todas as partículas iniciam o experimento no compartimento A enquanto o compartimento B está vazio. Em seguida “vamos dar regras” para determinar a evolução do número de partículas no compartimento B em função do tempo, isto é, teremos um processo

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

Vamos supor ainda que a cada instante de tempo, uma e apenas uma partícula passará do compartimento A para o B e vice-versa. Em termos de variável X_n isto significa que $X_{n+1} = X_n \pm 1$.

Podemos interpretar X_n como a posição de uma **partícula virtual** que “salta de um ponto para um dos seus pontos vizinhos” (ver Figura 5.5). Quando a partícula virtual se encontrar no estado k no tempo n ($X_n = k$), temos que no tempo $n + 1$ ela se encontrará em $k + 1$ ou em $k - 1$. Logo,

$$\mathbb{P}(X_{n+1} = k + 1 \mid X_n = k) \text{ e } \mathbb{P}(X_{n+1} = k - 1 \mid X_n = k)$$

serão nossas probabilidades de transição.

Figura 5.5: Estado do sistema no tempo n .



Fonte: (FRIEDLI, 2011, pág 6).

De acordo com as regras definidas anteriormente, para o primeiro estado X_1 após a condição inicial $X_0 = 0$, temos que a única coisa que pode acontecer de $n = 0$ para $n = 1$ é de uma das N partículas do compartimento A passar pela abertura em direção

ao compartimento B (vide Figura 5.3, em nossas ilustrações anteriores). Ou seja, $X_1 = 1$. Esse primeiro passo, será representado por

$$\mathbb{P}(X_1 = 1 \mid X_0 = 0) = 1,$$

o que significa que quando $N_B = 0$, uma partícula passará do compartimento A para o compartimento B , com probabilidade igual a 1.

Como assumimos que as partículas são indistinguíveis, segue que elas possuem as mesmas possibilidades quanto a mudança de compartimento. Assim, a partícula que mudará para o compartimento ao lado pode ser aleatoriamente escolhida entre as N partículas contidas no sistema. A probabilidade de se escolher uma partícula em A , dado que temos k partículas em B é de $\frac{N-k}{N}$, logo

$$\mathbb{P}(X_{n+1} = k + 1 \mid X_n = k) = \frac{N - k}{N}.$$

E a probabilidade de se escolher uma partícula em B é de $\frac{k}{N}$, ou seja,

$$\mathbb{P}(X_{n+1} = k - 1 \mid X_n = k) = \frac{k}{N}.$$

Mais uma observação: quando o compartimento B conter todas as partículas ($k = N$), temos que

$$\mathbb{P}(X_{n+1} = N - 1 \mid X_n = N) = 1.$$

No nosso modelo, é intuitivo que o compartimento que se encontra com o maior número de partículas tende a enviar as suas partículas para o compartimento vizinho, ou seja, o sistema busca entrar em equilíbrio, igualando as quantidades em ambos compartimentos.

Podemos escrever as probabilidades de transição como entradas de uma matriz Q , $(N + 1) \times (N + 1)$, cujos elementos $q_{ij}, (i, j \in E)$ são as probabilidades de transição, isto

é,

$$q_{ij} := \mathbb{P}(X_{n+1} = i \mid X_n = j).$$

Assim,

$$Q = \begin{pmatrix} 0 & \frac{1}{N} & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 \\ 1 & 0 & \frac{2}{N} & 0 & \dots & 0 & \dots & \vdots & \vdots & \vdots & \vdots \\ 0 & \frac{N-1}{N} & 0 & \frac{3}{N} & \dots & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{N-2}{N} & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{N-3}{N} & \dots & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & \frac{k}{N} & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & \frac{N-k}{N} & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & \frac{N-3}{N} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \frac{N-2}{N} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \dots & \frac{3}{N} & 0 & \frac{N-1}{N} & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \frac{2}{N} & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \frac{1}{N} & 0 \end{pmatrix} \quad (5.1)$$

Observações:

- Q é uma matriz de transição pois $q_{ij} \geq 0$, $\forall i, j$ e ainda $\sum q_{ij} = 1$, $\forall j$.
- A sequência X_0, X_1, X_2, \dots de variáveis aleatórias é uma cadeia de Markov com espaço de estados $E = \{0, 1, \dots, N\}$ e a matriz de transição Q .

Como dito no início deste capítulo, a evolução do sistema com a condição inicial $N_A = N$ e $N_B = 0$ tende a se aproximar de um estado de equilíbrio, onde $N_A \simeq N_B \simeq \frac{N}{2}$.

Para provar este comportamento à partir da Cadeia de Markov X_n e de sua matriz de

transição Q , representaremos a condição inicial $X_0 = 0$ por um vetor linha

$$\mu^{(0)} = (\mu_0^{(0)}, \mu_1^{(0)}, \mu_2^{(0)}, \dots, \mu_N^{(0)}) = (1, 0, 0, \dots, 0, 0).$$

E isto significa que, no tempo $n = 0$, temos:

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

Assim, a probabilidade contida no vetor $\mu^{(0)}$, que está concentrada no estado 0 no tempo 0, ao longo do tempo vai se espalhando pelo sistema, como iremos verificar. As distribuições iniciais ($n = 1$ e $n = 2$), são fáceis de se determinar, observando as probabilidades de transição.

No tempo $n = 1$ temos que

$$\mu^{(1)} = (\mu_0^{(1)}, \mu_1^{(1)}, \mu_2^{(1)}, \dots, \mu_N^{(1)}) = (0, 1, 0, \dots, 0, 0).$$

Já no tempo $n = 2$, a distribuição é dada por

$$\mu^{(2)} = (\mu_0^{(2)}, \mu_1^{(2)}, \mu_2^{(2)}, \dots, \mu_N^{(2)}) = \left(\frac{1}{N}, 0, \frac{N-1}{N}, \dots, 0\right).$$

De um modo geral, a distribuição

$$\mu^{(n)} = (\mu_0^{(n)}, \mu_1^{(n)}, \mu_2^{(n)}, \dots, \mu_N^{(n)}),$$

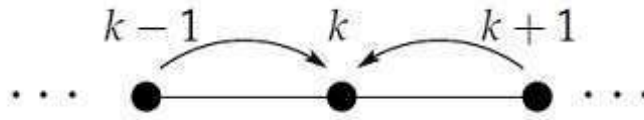
para $n \geq 3$, é um pouco mais complicada de se determinar. Faremos a seguir algumas observações necessárias para constatar que $\mu_k^{(n)} = Q\mu_k^{(n-1)}$ (o que já é esperado pelas observações do Capítulo 4).

Lembremos que a interpretação para $\mu_k^{(n)}$ é a seguinte:

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

Note que se a partícula virtual estiver em k no tempo n , então ou ela estava em $k - 1$ ou estava em $k + 1$ no tempo $n - 1$, como é ilustrado na Figura 5.6 a seguir.

Figura 5.6: Estado do sistema no tempo $n - 1$.



Fonte: (FRIEDLI, 2011, pág 9).

Portanto, temos a seguinte igualdade para $1 \leq k \leq N - 1$:

$$\begin{aligned} \mu_k^{(n)} &= \mathbb{P}(X_n = k) \\ &= \mathbb{P}(X_{n-1} = k - 1)\mathbb{P}(X_n = k | X_{n-1} = k - 1) + \mathbb{P}(X_{n-1} = k + 1)\mathbb{P}(X_n = k | X_{n-1} = k + 1) \\ &= q_{k,k-1}\mu_{k-1}^{(n-1)} + q_{k,k+1}\mu_{k+1}^{(n-1)}. \end{aligned}$$

Como $q_{k,l} = 0$, para $l \neq k - 1$ e $l \neq k + 1$, segue que

$$\mu_k^{(n)} = \sum_{j \in E} q_{kj}\mu_j^{(n-1)}.$$

Ou ainda, em notação matricial:

$$\mu_k^{(n)} = Q\mu_k^{(n-1)}. \tag{5.2}$$

Note ainda que, para $k = 0$ e $k = N$, temos

- $\mu_0^{(n)} = \frac{1}{N}\mu_1^{(n-1)}$;
- $\mu_N^{(n)} = \frac{1}{N}\mu_{N-1}^{(n-1)}$;

onde $\mu_1^{(n-1)}$ é a probabilidade de no tempo $n - 1$ termos 1 partícula no compartimento B ,

e $\mu_{N-1}^{(n-1)}$ é a probabilidade de no tempo $n - 1$ termos $N - 1$ partícula no compartimento B .

Em (5.2) obtivemos, como era de se esperar, a distribuição no tempo n à partir da distribuição do tempo $n - 1$ é dado pelo *produto matricial da matriz de transição com a distribuição do tempo $n - 1$* , isto é,

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

onde Q^n é a n -ésima potência de Q .

A fim de fazer algumas previsões em nosso modelo de distribuição de partículas, seria interessante estudar o comportamento assintótico de $\mu^{(n)}$, ou seja, saber se existe

$$\lim_{n \rightarrow \infty} \mu^{(n)}.$$

Vamos supor que o limite $\lim_{n \rightarrow \infty} \mu^{(n)} = w$ existe. Isto é, para cada $k \in E$,

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

O vetor w , com componentes w_k , é uma distribuição de probabilidade em E , e deve ser interpretado da seguinte forma: para um tempo n grande, a probabilidade da partícula estar em k é w_k . É interessante notar que em termos de interpretação da partícula virtual, significa que a partícula virtual jamais para de saltar, mesmo para um tempo n grande, porém a “chance” dela estar no estado k no tempo n é de w_k .

Observe que quando μ_k^n converge para w_k temos que

$$(Qw)_k = \left(Q \lim_{n \rightarrow \infty} \mu^n \right)_k = \left(\lim_{n \rightarrow \infty} \mu^{n+1} \right)_k = w_k,$$

Ou seja,

$$Qw = w. \tag{5.3}$$

Uma distribuição que satisfaz (5.3) é chamada de **distribuição invariante** em relação

a Q . Recebe este nome pois se a distribuição inicial é w , então ele não mudará ao longo do tempo, isto é, $Q^n w = w$.

Esta medida invariante é única? Bom, isto não é verdade para qualquer cadeia de Markov. Porém, para o nosso modelo, a cadeia é finita (estados finitos) e irredutível (de um estado pode-se alcançar qualquer outro), e portanto a medida invariante é única (ALLEN, 2010, p.68).

Teorema 3. *A matriz Q dada em (5.1) possui uma distribuição invariante $w = (w_0, w_1, w_2, \dots, w_N)$, dada pela **distribuição de Bernoulli**²:*

$$w_k = \frac{1}{2^N} \binom{N}{k}, \text{ para } k = 0, 1, 2, \dots, N.$$

Demonstração. Tomemos $w = \frac{1}{2^N} (\binom{N}{0}, \binom{N}{1}, \dots, \binom{N}{N})$ e verifiquemos que $Qw = w$.

Para isto, notemos primeiramente que

$$(Qw)_k = q_{k0}w_0 + q_{k1}w_1 + q_{k2}w_2 + \dots + q_{kN}w_N = \sum_{j=0}^N q_{kj}w_j,$$

onde cada q_{kj} é diferente de zero quando $j = k + 1$ e $j = k - 1$, e q_{kj} é igual à zero para os demais casos. Vamos agora mostrar que $(Qw)_k = w_k$ para $1 \leq k \leq N - 1$. Temos que

$$\begin{aligned} \sum_{j=0}^N q_{kj}w_j &= q_{k,k-1}w_{k-1} + q_{k,k+1}w_{k+1} \\ &= \frac{1}{2^N} \left[\frac{N - (k - 1)}{N} \binom{N}{k - 1} + \frac{(k + 1)}{N} \binom{N}{k + 1} \right] \\ &= \frac{1}{2^N} \left[\frac{N - (k - 1)}{N} \frac{N!}{(N - (k - 1))!(k - 1)!} + \frac{(k + 1)}{N} \frac{N!}{(N - (k + 1))!(k + 1)!} \right] \\ &= \frac{1}{2^N} \left[\frac{(N - 1)!}{(N - k)!(k - 1)!} + \frac{(N - 1)!}{N - (k + 1)! k!} \right] \end{aligned}$$

²Recebe este nome em homenagem ao matemático suíço Jakob Bernoulli, e é a distribuição discreta de espaço amostral $\{0,1\}$ com probabilidade de sucesso p igual a 1 e probabilidade de não-sucesso $q = 1 - p$ igual a 0.

$$\begin{aligned}
&= \frac{1}{2^N} \left[\frac{(N-1)!}{[(N-k)(N-(k+1))!](k-1)!} + \frac{(N-1)!}{[(N-(k+1))!]k(k-1)!} \right] \\
&= \frac{1}{2^N} \left[\frac{(N-1)!k + (N-1)!(N-k)}{(N-(k+1))!(N-k)(k-1)!k} \right] \\
&= \frac{1}{2^N} \left[\frac{(N-1)!N}{(N-(k+1))!(N-k)(k-1)!k} \right] \\
&= \frac{1}{2^N} \left[\frac{N!}{(N-k)!k!} \right] \\
&= \frac{1}{2^N} \binom{N}{k} \\
&= w_k.
\end{aligned}$$

Vamos analisar o que acontece no resultado obtido acima quando $k = 0$ ou $k = N$.

Para $k = 0$ temos:

$$\begin{aligned}
(Qw)_0 &= \underbrace{q_{0,0}w_0}_{=0} + q_{0,1}w_1 + \underbrace{q_{0,2}w_2}_{=0} + \underbrace{\dots}_{=0} + \underbrace{q_{0,N}w_N}_{=0} \\
&= q_{0,1}w_1 \\
&= \frac{1}{N} \frac{1}{2^N} \binom{N}{1} \\
&= \frac{1}{2^N} \frac{1}{N} \frac{N!}{(N-1)!} \\
&= \frac{1}{2^N} \frac{N!}{N!} \\
&= \frac{1}{2^N} \frac{N!}{(N-0)!0!} \\
&= \frac{1}{2^N} \binom{N}{0} \\
&= \frac{1}{2^N} \\
&= w_0.
\end{aligned}$$

Para $k = N$ temos:

$$\begin{aligned}
(Qw)_N &= \underbrace{q_{N,0}w_0}_{=0} + \underbrace{q_{N,1}w_1}_{=0} + \underbrace{q_{N,2}w_2}_{=0} + \underbrace{\dots}_{=0} + q_{N,N-1}w_{N-1} + \underbrace{q_{N,N}w_N}_{=0} \\
&= q_{N,N-1}w_{N-1} \\
&= \frac{1}{N} \frac{1}{2^N} \binom{N}{N-1}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2^N} \frac{1}{N} \frac{N!}{(N - (N - 1))!(N - 1)!} \\
&= \frac{1}{2^N} \frac{1}{N} \frac{N!}{(N - N + 1)!(N - 1)!} \\
&= \frac{1}{2^N} \frac{1}{N} \frac{N!}{(N - 1)!1!} \\
&= \frac{1}{2^N} \frac{1}{N} \frac{N(N - 1)!}{(N - 1)!1!} \\
&= \frac{1}{2^N} \frac{N!}{N!} \\
&= \frac{1}{2^N} \frac{N!}{(N - N)!N!} \\
&= \frac{1}{2^N} \binom{N}{N} \\
&= \frac{1}{2^N} \\
&= w_N.
\end{aligned}$$

■

5.3 Tempo de Permanência

Como no nosso modelo existe uma única medida invariante $w = (w_0, w_1, \dots, w_N)$, segue de um resultado conhecido para cadeias de Markov (BRZEZNIAK e ZASTAWNIAK 1999, p.212) que

$$w_k = \frac{1}{m_k} \quad (5.4)$$

onde m_k é o tempo médio de recorrência para o estado k .

Formalmente, a média m_k é dada por

$$m_k = \sum_{n=0}^{\infty} n f_{kk}^{(n)},$$

onde

$$f_{kk}^{(n)} = \mathbb{P}(\{w \in \Omega : \xi_n(w) = k, \xi_j(w) \neq k, j = 1, \dots, n - 1 \mid \xi_0(w) = k\})$$

ou seja, é a probabilidade de primeiro retorno ao estado k . Neste trabalho usaremos a medida invariante para determinar as médias m_k 's.

A componente w_k , da distribuição invariante, representa assintoticamente a *fração de tempo* no qual a partícula virtual passou em um estado k ao longo de sua trajetória. Observando, por exemplo, o caso para o qual $N = 10$, temos:

$$w_5 = \frac{1}{2^{10}} \cdot \binom{10}{5} = \frac{1}{2^{10}} \cdot \frac{10!}{5!2^5} \simeq 0.246$$

Isso significa que, ao longo da evolução, o sistema passa em torno de 25 % do seu tempo no estado em que o número de partículas no compartimento A e no compartimento B estão igualmente distribuídas ($N_A = N_B = 5$). Mas por outro lado, temos

$$w_0 = \frac{1}{2^{10}} \cdot \binom{10}{0} = \frac{1}{2^{10}} \cdot \frac{10!}{10!0!} = \frac{1}{2^{10}} \simeq 0,001.$$

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

Note ainda que

$$m_5 = \frac{1}{0,246} \approx 4,07$$

ou seja, depois de deixar o estado 5, leva em média 4,07 unidades de tempo para retornar ao estado 5. Enquanto que, $m_0 = 1024$, e portanto após deixar o estado 0, leva em média 1024 unidades de tempo para retornar ao estado 0.

Para estimar alguns outros resultados, suporemos que n seja medido em segundos e vamos estimar o tempo, em média, de retorno:

- Caso o sistema inicie com 26 partículas no compartimento A e 0 no compartimento B , ou seja, $N = 26$ (como mostra a Figura 5.2), temos

$$m_0 = \frac{1}{w_0} = 2^{26} \text{ segundos} \simeq 67108864 \text{ segundos} \simeq 2,13 \text{ anos}$$

- Iniciando com 40 partículas, todas no compartimento A ($N = 40$), temos

$$m_0 = \frac{1}{w_0} = 2^{40} \text{ segundos} \simeq 1,1 \text{ trilhões segundos} \simeq 35000 \text{ anos}$$

- Para $N = 100$, com as 100 partículas iniciando no compartimento A , temos o

seguinte:

$$m_0 = \frac{1}{w_0} = 2^{100} \text{ segundos} \simeq 1,26765 \times 10^{30} \text{ segundos} \simeq 4 \times 10^{22} \text{ anos.}$$

Observe que, conforme aumentamos, aos poucos, o número de partículas iniciais no sistema, o tempo médio de retorno aumentou consideravelmente. O resultado para que 100 partículas retornem ao seu estado inicial é um tempo muito longo, ultrapassando até mesmo a idade do Universo (aproximadamente $1,37 \times 10^{10}$ anos), e este tempo só tende a aumentar se considerarmos que, em uma situação real, o número de partículas é de $N \simeq 10^{23}$ (número de Avogrado).

Este modelo explica, de certa forma, o uso do termo irreversibilidade: pois apesar de o tempo de retorno, para o estado inicial ser um número finito, este número é tão alto que para escalas humanas, o processo é considerado irreversível.

Observação: Para verificar o que ocorre com $\frac{N}{2}$, quando N é um número muito grande, usamos a **fórmula de Stirling**³ :

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

Sendo assim, a probabilidade de existir o mesmo número de partículas nos compartimentos A e B é aproximada por

$$w_{\frac{N}{2}} = \frac{1}{2^N} \frac{N!}{(N/2)!^2} \sim \sqrt{\frac{2}{\pi N}}.$$

Portanto, quando $N \rightarrow \infty$, $w_{\frac{N}{2}}$ converge para 0, ou seja, ter *exatamente* o mesmo número de partículas em ambos os compartimentos é um evento raro. Porém, a probabilidade de todas as partículas estarem em B é de

$$w_0 = \frac{1}{2^N}.$$

Ou seja, a fração de tempo passado no estado 0 é ínfima comparada com a fração de

³recebe este nome em homenagem à James Stirling (1662-1728) matemático escocês, e estabelece uma aproximação assintótica para o fatorial de um número.

tempo no estado $\frac{N}{2}$, pois $\frac{1}{2^N}$ é bem menor que $\sqrt{\frac{2}{\pi N}}$ para N grande.

Por exemplo, para $N = 100$ temos que $\frac{1}{2^{100}} \approx 7,9 \times 10^{-31}$ e $\sqrt{\frac{2}{100\pi}} \approx 0,08$.

5.4 Uma pequena perturbação no modelo

A matriz definida em (5.1) não satisfaz as condições do Teorema 2. Faremos então, uma pequena perturbação no modelo para adaptar ao teorema. Assumiremos que existe uma pequena possibilidade ($2\epsilon > 0$) de se manter no estado atual quando o tempo evolui.

Sendo assim, temos a seguinte modificação da matriz Q na qual chamaremos de Q_ϵ , para $0 < \epsilon < \frac{1}{N}$:

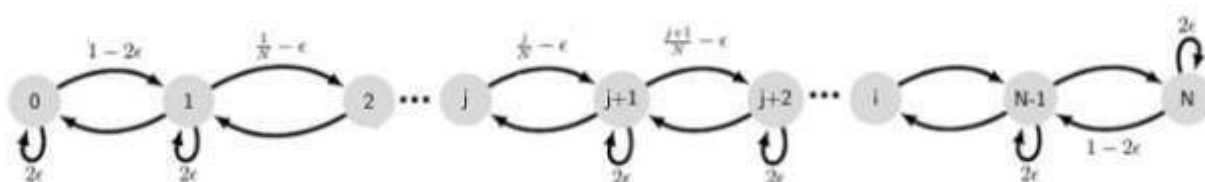
$$\begin{pmatrix} 2\epsilon & \frac{1}{N} - \epsilon & 0 & \dots & 0 & \dots & 0 & 0 & 0 \\ 1 - 2\epsilon & 2\epsilon & \frac{2}{N} - \epsilon & \dots & 0 & \dots & \vdots & \vdots & \vdots \\ 0 & \frac{N-1}{N} - \epsilon & 2\epsilon & \dots & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \frac{N-2}{N} - \epsilon & \dots & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & \frac{k}{N} - \epsilon & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 2\epsilon & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & \frac{N-k}{N} - \epsilon & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & \dots & \frac{N-2}{N} - \epsilon & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & 2\epsilon & \frac{N-1}{N} - \epsilon & 0 \\ 0 & 0 & 0 & \dots & 0 & \dots & \frac{2}{N} - \epsilon & 2\epsilon & 1 - 2\epsilon \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & \frac{1}{N} - \epsilon & 2\epsilon \end{pmatrix} \quad (5.5)$$

onde Q_ϵ é uma pequena perturbação de (5.1), pois o $\lim_{\epsilon \rightarrow 0} Q_\epsilon = Q$. A única diferença, com relação ao modelo anterior, é que a partícula terá uma pequena probabilidade de ficar

onde está antes de saltar para um de seus vizinhos, se a dinâmica do sistema for regida por Q_ϵ . Em particular, para um par de pontos i, j quaisquer no espaço de estado E , temos que a probabilidade $(Q_\epsilon)_{i,j}^N$ é maior ou igual a probabilidade da partícula virtual saltar $N - |j - i|$ vezes em j para depois saltar $|j - i|$ vezes em direção a i . Pois esta é apenas uma das trajetórias possíveis de se sair de j e chegar em i , enquanto $(Q_\epsilon)_{i,j}^N$ representa a probabilidade de sair de j e chegar em i por diversos caminhos possíveis (vide Figura 5.7). Cada probabilidade de transição, quando saltamos de j em direção a i , é maior que $\left(\frac{1}{N} - \epsilon\right)$, e portanto, a probabilidade desta trajetória particular (saltar $N - |j - i|$ vezes em j para depois saltar $|j - i|$ vezes em direção a i) é maior ou igual a

$$(2\epsilon)^{N-|j-i|} \left(\frac{1}{N} - \epsilon\right)^{|j-i|}$$

Figura 5.7: Dinâmica do sistema regido por Q_ϵ .



Fonte: O próprio autor (2018).

Portanto, Q_ϵ satisfaz a condição de regularidade do Teorema 2. Então, existe a distribuição invariante w_ϵ e $\lim_{n \rightarrow \infty} Q_\epsilon^n \mu = w_\epsilon$ para qualquer distribuição inicial μ . Usaremos este modelo Q_ϵ no próximo capítulo, com uma abordagem que possa ser realizada com alunos do Ensino Médio.

Capítulo 6

Sorteio de bolinhas: uma abordagem voltada para uma aula de Matemática no Ensino Médio

Neste capítulo apresentaremos uma aplicação do modelo da urna de Ehrenfest voltada para as aulas de Matemática no Ensino Médio (principalmente no 2º ano, onde alguns dos conceitos aprendidos até aqui são ensinados, visando uma melhor visualização e compreensão, e também no 3º ano, onde esta aplicação pode ser utilizada para retomada de tais conceitos).

Para isto, iremos propor o sorteio de bolinhas, vermelhas e azuis, todas do mesmo tamanho, que estarão dentro de uma caixa. Serão realizados vários sorteios, e em cada um destes sorteios, iremos alterar (substituir) a cor da bolinha sorteada. A grosso modo, as bolinhas farão o papel das partículas dos gases e as cores farão o papel dos compartimentos A e B no modelo probabilístico para o problema da irreversibilidade dos gases (apresentada no capítulo anterior). Este modelo permite que o professor instigue os seus alunos com questionamentos do tipo:

Após um longo período de tempo em foi iniciado o sorteio com bolinhas, todas de mesma cor, qual é a probabilidade de que a caixa volte a ter todas as bolinhas do início do sorteio?"

Esta aplicação surgiu através dos estudos e conversas sobre o Modelo da Urna de

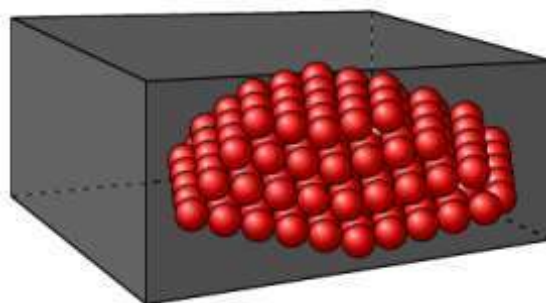
Ehrenfest, entre a autora deste trabalho e seu orientador, que buscavam um modelo aplicável nas aulas/oficinas de Matemática para alunos do Ensino Médio.

6.1 O modelo para o sorteio das bolinhas

Considere uma caixa preta na qual o espectador que sorteará as bolinhas não consiga visualizar o seu interior tendo apenas uma abertura para passagem das mãos, logo o sorteio não sofrerá influências externas. Considere ainda que a quantidade de bolinhas iniciais dentro da caixa é fixa, e que todas as bolinhas possuem mesma massa, tamanho e material, sendo indistinguíveis ao tato. Suporemos que nosso sorteio inicie apenas com bolinhas vermelhas (vide [Figura 6.1](#)).

Caso a bolinha sorteada seja vermelha, o que certamente irá acontecer em um primeiro momento, ela será substituída por uma de cor azul (considere que fora da caixa há bolinhas vermelhas e bolinhas azuis, porém estas são utilizadas apenas para substituição e em nada modificam a quantidade de bolinhas existente dentro da caixa). Após a substituição, a bolinha azul vai para a caixa e a bolinha vermelha que foi sorteada é descartada, ficando de fora da caixa. É realizado um novo sorteio. Se a bolinha sorteada for vermelha, será substituída por outra de cor azul ou, se a bolinha sorteada for da cor azul, será substituída por uma vermelha. A bolinha substituta vai para caixa e a sorteada é descartada. E assim um novo sorteio é realizado repetindo-se inúmeras vezes o processo descrito.

Figura 6.1: O modelo da urna para o sorteio das bolinhas.



Fonte: O próprio autor (2018).

A variável

$X_n :=$ número de bolinhas azuis dentro da caixa no instante n

representa o estado do sistema. Como o número de bolinhas azuis dentro da caixa pode ser zero, um, dois, três, ou quatro, temos que a imagem de X_n é $\{0, 1, 2, 3, 4\} \subseteq \mathbb{R}$.

No tempo $n = 0$, todas as bolinhas serão vermelhas e não há bolinhas azuis no início do sorteio. Logo, a condição inicial é dada por $X_0 = 0$.

Como a cada instante, uma e apenas uma bolinha será sorteada e substituída por outra bolinha de outra cor, temos que $X_{n+1} = X_n \pm 1$; e nosso primeiro passo será representado por:

$$\mathbb{P}(X_1 = 1 \mid X_0 = 0) = 1,$$

já que de $n = 0$ para $n = 1$ a única coisa que pode ocorrer é uma das quatro bolinhas vermelhas ser sorteada e na sequência ser substituída por uma azul ($X_1 = 1$).

Para $n = 2$ temos duas situações:

1. uma das três bolinhas vermelhas dentro da urna ser sorteada e substituída por uma azul. Representaremos esse passo através da probabilidade:

$$\mathbb{P}(X_2 = 2 \mid X_1 = 1) = \frac{3}{4}.$$

2. a única bolinha azul dentro da urna ser sorteada e substituída por uma vermelha.

Para esta situação temos:

$$\mathbb{P}(X_2 = 0 \mid X_1 = 1) = \frac{1}{4}.$$

Caso a segunda situação ocorra, nosso experimento volta ao estado inicial e será dado como encerrado. Se a primeira situação ocorrer, continuamos o sorteio dando início ao terceiro passo, e assim por diante. Supondo que a quantidade de bolinha azul dentro da caixa seja k no tempo n ($X_n = k$), então no tempo $n + 1$ ela será de $k + 1$ ou

$k - 1$, e suas probabilidades serão respectivamente de $\mathbb{P}(X_{n+1} = k + 1 \mid X_n = k)$ e $\mathbb{P}(X_{n+1} = k - 1 \mid X_n = k)$, onde

$$\mathbb{P}(X_{n+1} = k + 1 \mid X_n = k) = \frac{4 - k}{4}$$

e

$$\mathbb{P}(X_{n+1} = k - 1 \mid X_n = k) = \frac{k}{4}.$$

Assim como acontece no modelo da Urna de Ehrenfest, as bolinhas da cor que está em maior quantidade dentro da caixa tende a ser sorteadas e conseqüentemente serem substituídas pelas bolinhas da outra cor. O número de bolinhas vermelhas (N_V) e o número de bolinhas azuis (N_A) tende a se aproximar de um estado de equilíbrio onde $N_V \simeq N_A \simeq 2$. Escreveremos as probabilidades de transição como entradas de uma matriz 5×5 , denotada por Q :

$$Q = \begin{pmatrix} 0 & \frac{1}{4} & 0 & 0 & 0 \\ 1 & 0 & \frac{2}{4} & 0 & 0 \\ 0 & \frac{3}{4} & 0 & \frac{3}{4} & 0 \\ 0 & 0 & \frac{2}{4} & 0 & 1 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \end{pmatrix}$$

Observe que Q possui as mesmas propriedades da Matriz (5.1) descrita no capítulo anterior. Para que Q satisfaça as condições do Teorema 2, vamos dar uma nova regra ao nosso sorteio:

- Antes de sortear uma bolinha, joga-se dois dados. Caso ambas as faces voltadas para cima seja 6, nenhuma bolinha será retirada da urna, realizando assim um novo lançamento dos dados. Para quaisquer resultados diferentes do que foi imposto no lançamento dos dados, o sorteio de uma bolinha deve ocorrer.

Sendo assim, a matriz Q sofre uma modificação (proposta em 5.5), tornando-se

$$Q_\epsilon = \begin{pmatrix} 2 \cdot \frac{1}{36} & \frac{1}{4} - \frac{1}{36} & 0 & 0 & 0 \\ 1 - 2 \cdot \frac{1}{36} & 2 \cdot \frac{1}{36} & \frac{2}{4} - \frac{1}{36} & 0 & 0 \\ 0 & \frac{3}{4} - \frac{1}{36} & 2 \cdot \frac{1}{36} & \frac{3}{4} - \frac{1}{36} & 0 \\ 0 & 0 & \frac{2}{4} - \frac{1}{36} & 2 \cdot \frac{1}{36} & 1 - 2 \cdot \frac{1}{36} \\ 0 & 0 & 0 & \frac{1}{4} - \frac{1}{36} & 2 \cdot \frac{1}{36} \end{pmatrix}$$

onde $\epsilon = \frac{1}{36}$ é a probabilidade de cair duas faces 6 no lançamento simultâneo de dois dados.

Logo,

$$Q_\epsilon = \begin{pmatrix} \frac{1}{18} & \frac{2}{9} & 0 & 0 & 0 \\ \frac{17}{18} & \frac{1}{18} & \frac{17}{36} & 0 & 0 \\ 0 & \frac{13}{18} & \frac{1}{18} & \frac{13}{18} & 0 \\ 0 & 0 & \frac{17}{36} & \frac{1}{18} & \frac{17}{18} \\ 0 & 0 & 0 & \frac{2}{9} & \frac{1}{18} \end{pmatrix}.$$

Vamos agora verificar que Q_ϵ é regular, ou seja, existe $n \in \mathbb{N}$ tal que Q_ϵ^n tem todas as entradas não-nulas:

Para $n = 2$,

$$Q_\epsilon^2 = \begin{pmatrix} \frac{23}{108} & \frac{2}{81} & \frac{17}{162} & 0 & 0 \\ \frac{17}{162} & \frac{359}{648} & \frac{17}{324} & \frac{221}{648} & 0 \\ \frac{221}{324} & \frac{13}{162} & \frac{37}{54} & \frac{13}{162} & \frac{221}{324} \\ 0 & \frac{221}{648} & \frac{17}{324} & \frac{359}{648} & \frac{17}{162} \\ 0 & 0 & \frac{17}{162} & \frac{2}{81} & \frac{23}{108} \end{pmatrix}.$$

Para $n = 3$,

$$Q_\epsilon^3 = \begin{pmatrix} \frac{205}{5832} & \frac{121}{972} & \frac{17}{972} & \frac{221}{2916} & 0 \\ \frac{2057}{3888} & \frac{1073}{11664} & \frac{1241}{2916} & \frac{221}{3888} & \frac{3757}{11664} \\ \frac{221}{1944} & \frac{949}{1458} & \frac{83}{729} & \frac{949}{1458} & \frac{221}{1944} \\ \frac{3757}{11664} & \frac{221}{3888} & \frac{1241}{2916} & \frac{1073}{11664} & \frac{2057}{3888} \\ 0 & \frac{221}{2916} & \frac{17}{972} & \frac{121}{972} & \frac{205}{5832} \end{pmatrix}.$$

Para $n = 4$,

$$Q_\epsilon^4 = \begin{pmatrix} \frac{12547}{104976} & \frac{359}{13122} & \frac{5015}{52488} & \frac{221}{13122} & \frac{3757}{52488} \\ \frac{6103}{52488} & \frac{90289}{209952} & \frac{2465}{26244} & \frac{26741}{69984} & \frac{3757}{52488} \\ \frac{65195}{104976} & \frac{1885}{13122} & \frac{1811}{2916} & \frac{1885}{13122} & \frac{65195}{104976} \\ \frac{3757}{52488} & \frac{26741}{69984} & \frac{2465}{26244} & \frac{90289}{209952} & \frac{6103}{52488} \\ \frac{3757}{52488} & \frac{221}{13122} & \frac{5015}{52488} & \frac{359}{13122} & \frac{12547}{104976} \end{pmatrix}.$$

Portanto, Q_ϵ é regular e satisfaz as condições do Teorema (2). Sendo assim, resolvendo o sistema $Q_\epsilon w_\epsilon = w_\epsilon$, temos

$$\begin{pmatrix} \frac{1}{18} & \frac{2}{9} & 0 & 0 & 0 \\ \frac{17}{18} & \frac{1}{18} & \frac{17}{36} & 0 & 0 \\ 0 & \frac{13}{18} & \frac{1}{18} & \frac{13}{18} & 0 \\ 0 & 0 & \frac{17}{36} & \frac{1}{18} & \frac{17}{18} \\ 0 & 0 & 0 & \frac{2}{9} & \frac{1}{18} \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}$$

que resulta no sistema de equações

$$\begin{cases} \frac{1}{18}w_0 + \frac{2}{9}w_1 + 0w_2 + 0w_3 + 0w_4 = w_0 \\ \frac{17}{18}w_0 + \frac{1}{18}w_1 + \frac{17}{36}w_2 + 0w_3 + 0w_4 = w_1 \\ 0w_0 + \frac{13}{18}w_1 + \frac{1}{18}w_2 + \frac{13}{18}w_3 + 0w_4 = w_2 \\ 0w_0 + 0w_1 + \frac{17}{36}w_2 + \frac{1}{18}w_3 + \frac{17}{18}w_4 = w_3 \\ 0w_0 + 0w_1 + 0w_2 + \frac{2}{9}w_3 + \frac{1}{18}w_4 = w_4 \end{cases}$$

Resolvendo este sistema de equações pelo método da substituição¹, e deixando todos os valores em função de w_0 , temos que $w = (w_0, \frac{17}{4}w_0, \frac{13}{2}w_0, \frac{17}{4}w_0, w_0)$ e do fato de

$$\begin{cases} w_0 + w_1 + w_2 + w_3 + w_4 = 1 \\ w_0 + \frac{17}{4}w_0 + \frac{13}{2}w_0 + \frac{17}{4}w_0 + w_0 = 1 \end{cases}$$

¹o leitor interessado ou que ainda tem dúvidas a respeito de resoluções de sistemas de equações lineares pode se aprofundar utilizando um bom livro de Matemática do Ensino Médio ou até mesmo livros de Álgebra Linear.

resulta que $w_0 = \frac{1}{17}$. Logo,

$$w_\epsilon = \left(\frac{1}{17}, \frac{1}{4}, \frac{13}{34}, \frac{1}{4}, \frac{1}{17} \right).$$

Vamos agora calcular o tempo estimado para que as quatro bolinhas dentro da caixa retornem ao seu estado inicial, ou seja, tenhamos todas vermelhas. Utilizando a Equação 5.4 e supondo que o tempo para a realização do sorteio é de 10 segundos, temos que o tempo médio de retorno para o estado 0 (inicial) é dado por

$$m_0 = \frac{1}{w_0} = \frac{1}{\frac{1}{17}} = 17 \text{ unidades de tempo} = 17 \cdot 10 \text{ seg} = 170 \text{ seg} \simeq 2,83 \text{ min.}$$

Ou seja, caso um professor leve para sua sala de aula este experimento iniciando com 4 bolinhas vermelhas, é possível que em 1 aula (com duração média de 50 minutos) os alunos vejam todas as bolinhas dentro da caixa retornarem ao seu estado inicial.

Repare ainda que, o fato de termos feito uma modificação em nosso experimento (adicionado a regra do lançamento de dados onde, caso ambas as faces fossem 6 o sorteio da bolinha não seria realizado) e por consequência uma modificação em nossa matriz de probabilidade de transição, fez com que o tempo de retorno aumentasse. Isso se deve ao fato de que existe a probabilidade de não efetuarmos o sorteio e as bolinhas permanecerem em seu estado atual.

Caso essa modificação em nosso modelo não existisse, w seria o mesmo encontrado em 3, e o tempo médio de retorno seria de

$$m_0 = \frac{1}{w_0} = 2^4 \text{ unidade de tempo} = 16 \text{ unidade de tempo} = 16 \cdot 10 \text{ seg} = 160 \text{ seg} \approx 2,7 \text{ min.}$$

Para o nosso modelo, iniciando com 4 bolinhas da cor vermelha, e com a modificação proposta, pudemos verificar que houve uma diferença de 1 unidade de tempo (no nosso caso 10 segundos) para o tempo médio de retorno. É interessante que o professor comente com seus alunos que, a medida que aumenta a quantidade de bolinhas iniciais no sorteio, o tempo de retorno aumenta “consideravelmente bem mais”.

6.2 Uma simulação para o sorteio de bolinhas

Nesta seção faremos uma simulação do modelo de sorteio de bolinhas, descrito na seção anterior, apresentando algumas etapas do processo por meio de fotografias.

Para a realização deste experimento foram utilizados materiais de baixo custo, e que na maioria das vezes, encontra-se disponível nas escolas (caixa de papelão, papel de cor preta, tecido, bolinhas de isopor de mesmo tamanho, tinta guachê e dois dados), a Figura 6.2. Utilizou-se também para a realização deste experimento um cronômetro (Figura 6.3) para o registro do tempo para realização do nosso modelo.

Figura 6.2: O nosso modelo de urna para o sorteio de bolinhas.



Fonte: O próprio autor (2018).

Figura 6.3: Cronômetro.



Fonte: O próprio autor (2018).

Para dar início ao nosso experimento, iremos colocar as 4 bolinhas vermelhas dentro da caixa, deixando de fora as quatro bolinhas azuis (Figura 6.4).

Figura 6.4: Momento inicial do sorteio.



Fonte: O próprio autor (2018).

Antes da realização de qualquer sorteio, iremos lançar simultaneamente dois dados (como mostra a Figura 6.5). Caso a face voltada para cima de ambos os dados seja 6, não realizamos o sorteio, devendo novamente serem lançados os dados. Em qualquer outra situação que ocorra durante o lançamento dos dados, diferente de ambas faces 6, o sorteio de uma bolinha dentro da caixa deve ser realizado.

Figura 6.5: 1º lançamento simultâneo dos dois dados.



Fonte: O próprio autor (2018).

Figura 6.6: 1º lançamento de dados realizado.



Fonte: O próprio autor (2018).

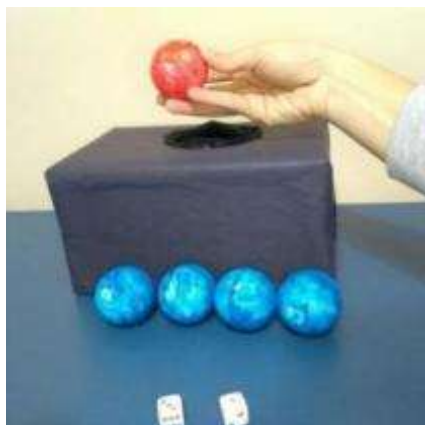
Como visto na Figura 6.6 o primeiro lançamento simultâneo dos dados resultou nas seguintes faces: 3 e 1. Logo, devemos realizar o sorteio de uma bolinha da caixa (Figura 6.7).

Figura 6.7: Realizando o 1º sorteio.



Fonte: O próprio autor (2018).

Figura 6.8: A primeira bolinha sorteada.



Fonte: O próprio autor (2018).

Como era de se esperar, e já descrito na seção anterior, como todas as bolinhas vermelhas iniciam-se dentro da caixa, a única possibilidade de ocorrer neste primeiro sorteio, é o de se retirar uma bolinha vermelha de dentro da caixa (como visto na Figura 6.8).

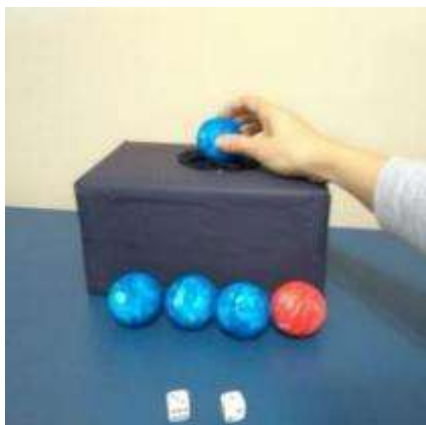
O próximo passo agora é substituir a bolinha vermelha sorteada por uma bolinha azul (Figura 6.9). A bolinha azul vai para dentro da caixa, enquanto a vermelha fica do lado de fora (Figura 6.10).

Figura 6.9: 1ª substituição das bolinhas.



Fonte: O próprio autor (2018).

Figura 6.10: Após a substituição da 1ª bolinha.



Fonte: O próprio autor (2018).

Figura 6.11: O 2º lançamento simultâneo dos dados.



Fonte: O próprio autor (2018).

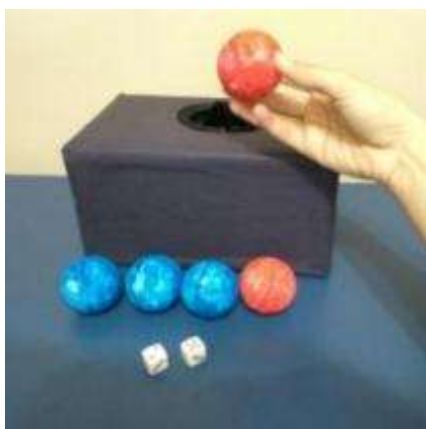
Após a substituição e reposição da bolinha de “cor oposta” dentro da caixa, deve-se realizar um novo lançamento de dados (vide Figura 6.11), seguindo as mesmas regras do início do sorteio. Neste segundo lançamento as faces voltadas para cima foram 2 e 5 (como mostra a Figura 6.12). Sendo assim, realizamos o segundo sorteio de uma bolinha. Agora nossas probabilidades de sorteio são as seguintes, levando em conta que temos um total de quatro bolinhas dentro da caixa (3 vermelhas e 1 azul): $\frac{3}{4}$ de chance de sortearmos uma bolinha vermelha e $\frac{1}{4}$ de chance de sortear a bolinha azul.

Figura 6.12: 2º lançamento de dados realizado.



Fonte: O próprio autor (2018).

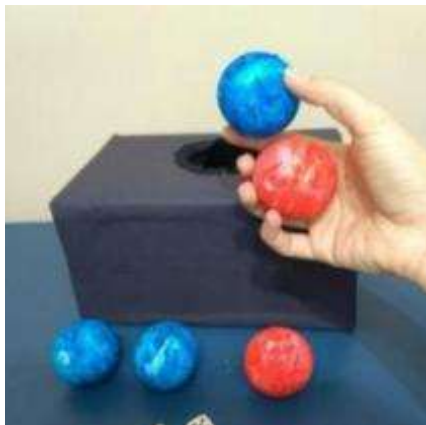
Figura 6.13: 2º sorteio da bolinha realizado.



Fonte: O próprio autor (2018).

Como podemos ver na Figura 6.13, no 2º sorteio foi retirada de dentro da urna uma bolinha vermelha, que será substituída por uma bolinha da cor azul (Figura 6.14). Sendo assim, agora dentro da caixa haverá 2 bolinhas vermelhas e 2 bolinhas azuis. E para o próximo sorteio as chances agora são de 50% para cada cor.

Figura 6.14: 2ª substituição das bolinhas.



Fonte: O próprio autor (2018).

Figura 6.15: 3º lançamento simultâneo dos dados.



Fonte: O próprio autor (2018).

O nosso sorteio continuou, e no 3º lançamento simultâneo dos dados (Figura 6.15) as faces foram diferentes de 6 e 6, nos permitindo a realização do sorteio da bolinha. A 3ª bolinha sorteada foi da cor azul. Sendo assim, após a sua substituição, a caixa passou a ter 3 bolinhas vermelhas e 1 bolinha azul, logo para nosso próximo sorteio, as chances de se retirar uma bolinha vermelha eram maiores. O que não ocorreu.

Após o 4º lançamento simultâneo dos dados apresentar faces diferentes de 6 e 6, um novo sorteio foi realizado. Apesar da probabilidade de se sortear uma bolinha vermelha

fosse maior (75%), a bolinha que saiu de dentro da urna foi da cor azul. Sendo assim, esta bolinha azul foi substituída por uma de cor vermelha e a caixa voltou ao seu estado inicial: com as 4 bolinhas vermelhas dentro da caixa (vide Figura 6.16). Encerrando assim o nosso experimento, após o cronômetro marcar quase 2 minutos e 15 segundos (como mostra a Figura 6.17).

Figura 6.16: Encerrando nosso sorteio.



Fonte: O próprio autor (2018).

Figura 6.17: Tempo de realização do nosso experimento.



Fonte: O próprio autor (2018).

Observação: O tempo de quase 2 minutos e 15 segundos, poderia ter sido bem menor (pois durante a realização estavam sendo tiradas fotos, o que atrasava os lançamentos

dos dados e sorteio das bolinhas), quanto poderia ter sido bem maior, caso as cores das bolinhas ficassem “oscilando” durante o sorteio (hora uma bolinha seria sorteada, hora uma bolinha azul seria sorteada).

6.3 Sugestão de plano de aula

Apresentamos a seguir uma sugestão de plano de aula para aplicação do modelo de sorteio das bolinhas apresentada na seção anterior.

Plano de Aula:

Tema: O sorteio de bolinhas e o tempo médio de retorno para o seu estado inicial.

Objetivos Gerais:

- Interpretar informações contidas em enunciados de situações-problema, com o objetivo de caracterizar a necessidade de mobilizar o raciocínio matemático;

- Utilizar elementos de matrizes para organizar e justificar a resolução de situações-problema envolvendo cadeias de Markov.

- Calcular e associar um vetor de probabilidade à situações-problema envolvendo cadeias de Markov.

- Verificar o tempo médio de retorno ao estado inicial de um experimento ou situação.

Objetivos Específicos:

- Adquirir conhecimentos básicos sobre probabilidade, em particular, da probabilidade condicional;

- Adquirir conhecimentos básicos sobre matrizes;

- Saber expressar, por meio de matrizes, situações relativas a cadeias de Markov;

- Introduzir o conceito de matrizes estocásticas;

- Saber resolver sistemas de equações lineares

- Compreender o significado das cadeias de Markov;

- Conhecer e compreender o problema da irreversibilidade dos gases (modelo da urna de Ehrenfest);

- Compreender o modelo do sorteio de bolinhas e sua relação com o modelo da urna de Ehrenfest.

Conteúdo:

- Noções de probabilidade;
- Matrizes: Representação de uma matriz, operações com matrizes, representação matricial de sistemas lineares, matrizes estocásticas;
- Cadeias de Markov: processos estocásticos, processos de Markov, matriz das probabilidades de transição, matrizes regulares, vetor de probabilidade a longo prazo;

Metodologia:

- Aula expositiva do conteúdo proposto;
- Apresentação do problema da irreversibilidade dos gases;
- Apresentação do problema do sorteio das bolinhas;
- Realização do experimento “sorteio das bolinhas”, para verificação do tempo médio de retorno para o estado inicial.

Recursos Didáticos:

- Quadro, giz e sala de informática.

Recursos materiais:

- Caixa de papelão;
- Papel (na cor preta);
- TNT (na cor preta);
- Bolinhas de isopor (todas de mesmo tamanho);
- Tintas guachê: nas cores vermelho e azul;
- 2 dados.

Avaliação:

- Observação da participação do aluno na resolução dos problemas sugeridos;
- Observação da participação do aluno na construção e realização do sorteio das bolinhas;
- Entrega de relatório referente as atividades em sala de aula.

Conclusão

Sabemos o quanto a Matemática é mal interpretada por boa parte de nossos alunos, e isto se deve, na minha opinião, pelo fato de que muitos enxergam a Matemática como a disciplina do “decorar fórmulas”, e não entendem o significado de tais resultados, o que acarreta em grandes defasagens e até mesmo num trauma com a disciplina. O professor deve então buscar ferramentas para tornar as aulas mais atrativas e significativas para o seu aluno, mostrar que a Matemática pode sim ser aprendida e não simplesmente decorada. Uma dessas ferramentas que o professor pode utilizar é a associação do conteúdo ensinado com o cotidiano do aluno, ou algo que mesmo não sendo do seu cotidiano, que lhe desperte interesse.

A autora sendo professora do ensino básico, durante a elaboração de todo este trabalho, buscou associar todo o conteúdo aprendido com o conteúdo que leciona aos seus alunos. Desta maneira, e através de conversas com o seu orientador, pensou-se no modelo do sorteio de bolinhas. Este experimento, que é uma inspiração na urna de Ehrenfest, visa a aprendizagem dos conceitos previstos no Currículo do Estado de São Paulo para o 2º ano do Ensino Médio na disciplina de Matemática, além de explicar os conceitos matemáticos envolvidos por trás de um problema da Física.

Dessa forma, espera-se que os alunos possam adquirir conceitos básicos (ou até mesmo avançados) de probabilidade, matrizes, resolução de sistemas de equações lineares, além de introduzir os conceitos (básicos) de matrizes estocásticas e cadeias de Markov. Espera-se ainda que, as dificuldades quanto aos conteúdos trabalhados sejam sanados e que os alunos apresentem uma melhora quanto a forma que enxergam a disciplina de Matemática e sintam-se motivados a buscar novos desafios ou explorar outros temas da área.

Mesmo não sendo possível a aplicação em sala de aula do modelo do sorteio de bolinhas durante a elaboração deste trabalho, a autora acredita que não encontraria grandes dificuldades para a sua execução, pois como já foi dito, boa parte do conteúdo necessário está previsto para o 2º ano do Ensino Médio e os materiais utilizados para a elaboração de tal modelo são todos de baixo custo.

Referências

- [1] ALLEN, L.J.S. **An introduction to stochastic processes with applications to biology**. Chapman and Hall/CRC, 2010.
- [2] BOLDRINI, J.L. et al. **Álgebra Linear**. 3. ed Campinas: Universidade Estadual de Campinas, Departamento de Matemática, 1980.
- [3] BRZEZNIAK, Z.; ZASTAWNIAK, T. **Basic stochastic processes: a course through exercises**. London: Springer, 1999. (Springer undergraduate mathematics series).
- [4] FRIEDLI, S. **Dinâmica Estocástica: três exemplos**. Universidade Federal de Minas Gerais, ano 2011.
- [5] JAMES, B. R. **Probabilidade: um curso em nível intermediário**. 2. ed. Rio de Janeiro: IMPA, 1996.
- [6] **PHYSICS BIOGRAPHIES: PAUL EHRENFEST**. Disponível em www.encyclopedia.com/people/science-and-technology/physics-biographies/paul-ehrenfest . Acesso em: janeiro, 2018.
- [7] OLIVEIRA, A.S.L.; RIBEIRO, T.S.G.; SILVA, F.B. **Cadeia de Markov: modelo probabilístico e convergência das distribuições de probabilidade**. CQD-Revista Eletrônica Paulista de Matemática, Bauru, v.11ic, p.49-61, ano 2017.
- [8] RIBEIRO, T.S.G. **Processos de Markov discretos: exemplos voltados para o ensino médio**. 2017. 73f. Dissertação de Mestrado - Universidade Estadual Paulista (UNESP), Bauru, 2017.

- [9] RUFFINO, P. R. C. **Uma iniciação aos sistemas dinâmicos estocásticos**. 2. ed. Rio de Janeiro: IMPA, 2009.
- [10] São Paulo (Estado). Secretaria da Educação. **Currículo do Estado de São Paulo: matemática e suas tecnologias**. São Paulo, 2011.
- [11] SILVA, F. B; ROTA, I.S. **Convergência de matrizes estocásticas regulares**. CQD–Revista Eletrônica Paulista de Matemática, Bauru, volume 8, p. 4-14, ano 2016.