Universidade Federal de Minas Gerais

Instituto de Ciências Exatas

Departamento de Matemática

Matemática e Embaralhamento de Cartas: de Mágicas a Cadeias de Markov

Paula Mendes Soares Fialho

Orientador: Bernardo Nunes Borges de Lima

Apoio: Capes

Belo Horizonte - MG

2016

Para Sérgio e Maria,

meus pais.

Agradecimentos

Gostaria de agradecer primeiramente à minha mãe, Maria, por ter me ensinado a gostar de Matemática, acreditado em mim, me apoiado durante toda a minha vida acadêmica e por todo o esforço para que eu alcançasse os meus sonhos.

Agradeço ao meu pai, Sérgio, e à minha irmã, Renata, pela paciência, pelo encorajamento e por torcerem tanto por mim.

Agradeço ao meu marido, Tiago, por me apoiar, por estar sempre disposto a me ajudar e por todo o companheirismo.

Agradeço a todos os meus amigos, por tornarem esta jornada tão mais prazerosa. Em especial agradeço ao meu amigo Marcos, por estar ao meu lado desde a graduação, me ajudando nos momentos de dificuldades e me incentivando sempre.

Agradeço ao meu orientador, Bernardo, pela paciência, por ter acreditado em mim e por ter me guiado, me dando a oportunidade de ir além do que eu imaginava. Serei eternamente grata.

Agradeço à Fump, por ter me mantido durante a graduação, e agradeço à Capes pelo apoio financeiro durante o mestrado.

Por fim, agradeço a Deus por ter colocado cada uma dessas pessoas na minha vida, sem elas eu não teria chegado até aqui.

Resumo

Este trabalho tem como objetivo estudar o modelo matemático do embaralhamento *Riffle Shuffle*, que denotaremos por Embaralhamento Canônico. O modelo matemático para o embaralhamento canônico é chamado de *modelo de Gilbert-Shannon-Reeds (GSR)*, em homenagem aos matemáticos que o desenvolveram.

A partir da análise de uma mágica introduziremos o conceito de sequências levantadoras, que está intimamente ligado ao modelo GSR. Com este conceito será possível calcular a probabilidade de obtermos permutações específicas das cartas de um baralho, após a realização de embaralhamentos canônicos.

O restante do trabalho será dedicado a encontrar o número de embaralhamentos canônicos consecutivos que aproxima a distribuição GSR da distribuição uniforme. Desenvolviremos um estudo sobre Cadeias de Markov, relacionando-as com o embaralhamento canônico.

PALAVRAS-CHAVE: Cadeias de Markov, Mágica, Modelo GSR, Riffle Shuffle.

Abstract

This work aims to study the mathematical model of *Riffle Shuffle*, that we will denote by canonical shuffle. The mathematical model for the canonical shuffle is called *Gilbert-Shannon-Reeds (GSR) model*, in honor of the mathematicians who developed it.

Based on an analysis of a magic, we will introduce the concept of rising sequence, which is closely linked to the GSR model. With this concept we will be able to calculate the probability of getting a specific permutation of cards, after performing several canonical shuffles.

The remaining of this work will be dedicated to find the number of consecutive canonical shuffles that approaches the GSR distribution from the uniform distribution. We will develop a study of Markov Chains, relating them with the canonical shuffle.

KEYWORDS: GSR Model, Magic, Markov Chains, Riffle Shuffle.

Sumário

Lista de Figuras **vi**

Lista de Tabelas **vii**

Introdução **1**

1 O modelo de Gilbert-Shannon-Reeds (GSR) **3**

1.1 Algumas definições 3

1.2 Mágica Premo 9

1.2.1 A performance 9

1.2.2 Análise da mágica 10

1.2.3 Um pouco sobre Charles Thornton Jordan 16

1.3 A distribuição GSR 17

1.3.1 Sequências levantadoras e o modelo GSR 18

1.3.2 Generalizando o modelo GSR 20

2 Cadeias de Markov e embaralhamentos **27**

2.1 O passeio aleatório em S_n 28

2.2 O passeio aleatório inverso 33

2.3 A distância da uniformidade 45

Apêndice **54**

Referências Bibliográficas **55**

Lista de Figuras

1.1 Pilha identidade 7

1.2 Intercalando os montes 8

1.3 Pilha resultante a partir do embaralhamento canônico 8

1.4 Sequências de variáveis aleatórias 8

1.5 Pilha resultante a partir do embaralhamento canônico inverso 9

1.6 *ab*-embaralhamento inverso 24

2.1 Diferença entre duas distribuições 40

Lista de Tabelas

1.1 Probabilidade de sucesso na realização da mágica 16

2.1 Comparação entre o passeio aleatório e o passeio inverso em S_n 38

2.2 Distância de Variação Total para um baralho de tamanho 52 54

Introdução

Uma das mais divertidas aplicações da Matemática é em mágicas. Matemática e mágicas compartilham uma relação de contribuição mútua, a partir da Matemática é possível criar bons truques de mágica e a partir de bons truques podemos encontrar príncipios que são aplicados em Matemática avançada. Considerando especificamente mágicas com baralhos, o segredo geralmente está na maneira de se embaralhar, e é também neste ponto em que a Matemática se encontra.

Este texto tem como objetivo estudar o modelo matemático de embaralhamento de cartas mais comum, o *Riffle Shuffle*, e questões relacionadas com a sua realização. Ao longo desta dissertação vamos responder a duas questões:

1. Qual a probabilidade de obtermos o baralho em uma ordem desejada, após um certo número de embaralhamentos?

2. Como embaralhar bem um baralho?

O ponto de partida para responder estas duas questões será a análise de uma mágica, conhecida como Premo. As referências principais para todo o texto serão o artigo de Bayer e Diaconis, *Trailing the Dovetail Shuffle to its Lair* [1], o texto de Hilário e Oliveira, *A Matemática de Embaralhar Cartas* [4], e o Capítulo 8 do livro *Markov Chains and Mixing Times* [6], de Lewin, Peres e Wilmer.

No Capítulo 1 apresentaremos as primeiras definições que serão utilizadas constantemente ao longo deste texto, dentre elas o modelo GSR e as Sequências Levantadoras. Em seguida, analisaremos a mágica Premo e faremos uma breve biografia de Charles Thornton Jordan, o criador da mágica inspiradora para o desenvolvimento da mágica Premo, tendo como referência os livros *Magical Mathematics: The mathematical ideas*

that animate great magical tricks [2], de Diaconis e Graham, e *Mathematics, magic and mystery* [3], de Gardner. Com o objetivo de responder à primeira pergunta citada acima, apresentaremos uma generalização para o modelo GSR, e vamos oferecer quatro descrições alternativas deste modelo a fim de poder analisar vários embaralhamentos canônicos consecutivos. As principais referências para este capítulo serão [1] e [4].

O Capítulo 2 será dedicado a responder à segunda pergunta citada acima, para isso desenvolviremos um estudo sobre Cadeias de Markov. Inicialmente apresentaremos vários conceitos referentes às Cadeias de Markov, como distribuição estacionária e irredutibilidade, e mostraremos como estes conceitos se relacionam tanto com o embaralhamento canônico quanto com o embaralhamento canônico inverso. Em seguida definiremos uma maneira de medirmos a distância entre duas distribuições de probabilidade e, no final da Seção 2, mostraremos que podemos analisar o embaralhamento canônico inverso ao invés do embaralhamento canônico. Finalizamos este trabalho oferecendo uma cota superior e uma cota inferior para o número de embaralhamentos canônicos consecutivos que embaralham bem um baralho. As referências para este capítulo serão [4] e [6].

Capítulo 1

O modelo de Gilbert-Shannon-Reeds (GSR)

1.1 Algumas definições

O objeto principal deste trabalho é um baralho, e denotaremos por *tamanho do baralho* o número de cartas que o baralho possui. Ao longo deste texto vamos considerar um baralho de tamanho n, com $n \in \mathbb{N}$, $n \neq 0$, sendo que todas as suas cartas são diferentes entre si.

Definição 1 (Pilha e Montes)**.** *Ao dizer que um baralho de tamanho n está em uma pilha, vamos considerar que as cartas estão umas sobre as outras e todas com as faces voltadas para baixo. Ao separarmos de uma pilha uma quantidade k de cartas, $k \in \{0, 1, \cdots, n-1, n\}$, formamos dois montes, um de tamanho k e outro de tamanho $n - k$.*

Algumas observações devem ser feitas sobre a Definição 1. Primeiramente observamos que montes de tamanho zero são permitidos, chamaremos estes montes de montes vazios. Outra observação é que uma pilha pode ser dividida em mais de dois montes da seguinte maneira: inicialmente separamos k_1 cartas, formando um monte com k_1 cartas e outro com $n - k_1$ cartas, em seguida separamos k_2 cartas do monte que possui $n - k_1$ cartas, formando três montes de tamanhos k_1, k_2 e $n - k_1 - k_2$. Repentindo este processo, é possível dividir uma pilha em quantos montes desejarmos.

Neste texto iremos considerar pilhas ordenadas, ou seja, quando o baralho estiver em uma pilha, dizemos que a carta do topo está na primeira posição, a carta logo abaixo da carta do topo está na segunda posição, e assim sucessivamente, terminando com a última carta, que está no fundo da pilha e vamos dizer que esta carta está na n-ésima posição. Representaremos uma pilha por

$$b = (b_1, b_2, ..., b_n),$$

onde b_i é a carta que está na i-ésima posição. Duas pilhas $b = (b_1, \cdots, b_n)$ e $c = (c_1, \ldots, c_m)$ são consideradas iguais se, e somente se, $n = m$ e $c_i = b_i$ para todo $i = 1, \cdots, n$.

Com a intenção de simplificar a notação, ao longo do texto consideraremos que um baralho de tamanho n é composto pelas cartas $1, 2, \cdots, n$, e denotaremos por *pilha identidade* a pilha $id = (1, 2, \ldots, n)$. Considerando um baralho de tamanho n, a seguir definimos duas operações elementares que podemos realizar com uma pilha qualquer.

Definição 2 (Corte reto)**.** *Consiste em separar uma pilha em dois montes, dividindo a pilha em duas partes simplesmente separando uma quantidade do topo e deixando as demais cartas na pilha. Se a pilha é representada por* $(b_1, b_2, ..., b_n)$*, o corte reto irá formar dois montes* $(b_1, ..., b_k)$ *e* $(b_{k+1}, ..., b_n)$*, com* $0 \leq k \leq n$*,* $k \in \mathbb{N}$*.*

Definição 3 (Corte)**.** *Primeiramente realizamos um corte reto e em seguida colocamos o monte retirado do topo no fundo do outro monte. Ou seja, a partir do corte reto formamos dois montes* $(b_1, b_2, ..., b_k)$ *e* $(b_{k+1}, ..., b_n)$*, e, ao colocar o primeiro monte sob o segundo, formamos uma nova pilha* $(b_{k+1}, ..., b_n, b_1, b_2, ..., b_k)$*.*

Feitas as definições e convenções sobre o baralho, podemos começar a falar sobre embaralhamentos em geral. Quando somos convidados a embaralhar um baralho na realidade estamos sendo convidados a modificar a posição original entre as cartas. Um baralho é considerado *bem embaralhado* quando as cartas estão bem misturadas, no sentido de que a pilha obtida após o embaralhamento não guarde memória da sua ordenação anterior.

Considerando um baralho de tamanho n em uma pilha $b = (b_1, \cdots, b_n)$ qualquer, ao embaralhar levamos cada carta b_i, $i \in \{1, 2, \cdots, n\}$, em uma nova posição, que

denotaremos por $\sigma(b_i)$, com $1 \leq \sigma(b_i) \leq n$. Assim, um embaralhamento qualquer leva uma pilha $b = (b_1, \cdots, b_n)$, em outra pilha $c = (c_1, \cdots, c_n)$, sendo que para cada carta $b_i, i \in \{1, \cdots, n\}$, existe um $j \in \{1, \cdots, n\}$ tal que $\sigma(b_i) = j$, ou seja, a carta que está na j- ésima posição na pilha c, representada c_j, correspode a carta b_i da pilha b. Desta forma, é natural associar embaralhamentos a permutações, mais detalhes desta associação podem ser vistos em [4].

A partir da associação embaralhamento-permutação, temos que o conjunto de todos os embaralhamentos possíveis de um baralho de tamanho n é o conjunto de permutações de n elementos, ou seja, o grupo simétrico, conhecido como S_n, que possui $n!$ elementos. Optamos por usar a mesma notação utilizada em [1]: dado um baralho de tamanho $n = 8$, suponhamos que o baralho esteja disposto na pilha identidade

$$id = (1, 2, 3, 4, 5, 6, 7, 8).$$

Se após a realização de um embaralhamento a pilha inicial é levada na nova pilha

$$c = (3, 7, 1, 2, 8, 5, 4, 6),$$

então associamos este embaralhamento à permutação $\varphi \in S_8$

$$\varphi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 7 & 1 & 2 & 8 & 5 & 4 & 6 \end{pmatrix}.$$

Considerando a associação embaralhamento-permutação, temos que o conjunto de todos embaralhamentos juntamente com a operação composição, ou seja, embaralhar repetidamente, é um grupo, e como tal, possui um elemento neutro. Chamaremos este elemento neutro de *permutação identidade*, que equivale a não modificar a ordenação das cartas. Assim,

$$id = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Seja b uma pilha qualquer, $b = (b_1, b_2, \cdots, b_n)$. Esta pilha b pode ser obtida a partir da pilha identidade da seguinte maneira: consideramos que inicialmente tinhamos a pilha identidade $(1, \cdots, n)$ e que após um embaralhamento ρ obtemos a pilha $(b_1, b_2, \cdots, b_n)$,

$$\rho = \begin{pmatrix} 1 & 2 & \cdots & n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}.$$

Desta forma, além de denotar embaralhamentos por permutações, também podemos denotar cada pilha pela permutação que a gera a partir da pilha identidade. Neste caso, por exemplo, podemos denotar b por ρ. Portanto, constatamos que tanto pilhas quanto embaralhamentos podem ser vistos como elementos de S_n. Com isso, podemos apresentar uma definição mais formal de um embaralhamento qualquer.

Definição 4. *Seja ρ uma pilha qualquer de S_n. Um embaralhamento é uma aplicação $\rho \to (x \circ \rho)$, de S_n em S_n, em que a pilha ρ é levada na pilha $(x \circ \rho) \in S_n$, com $x \in S_n$.*

Observe que, na definição acima, se $\rho = id$, o embaralhamento $x \in S_n$ é levado na pilha $x \in S_n$, o que é coerente com o que foi discutido anteriormente.

Outra característica de um grupo é que cada um de seus elementos possui um elemento inverso. Para o caso de embaralhamentos, suponha que uma pilha $\varphi \in S_n$ seja embaralhada segundo um embaralhamento $\rho \in S_n$ e obtemos a pilha $\theta \in S_n$, ou seja,

$$\rho \circ \varphi = \theta.$$

Em seguida podemos realizar outro embaralhamento λ na pilha θ de forma a recolocar as cartas em sua pilha original, assim obtendo novamente a pilha φ. Portanto, temos que

$$\varphi = \lambda \circ \theta = \lambda \circ \rho \circ \varphi,$$

ou seja, $\lambda \circ \rho = id$. Denotaremos um embaralhamento λ, tal que $\lambda \circ \rho = id$, por ρ^{-1} e o chamaremos de *embaralhamento inverso de ρ*.

Como mencionado anteriormente, embaralhar significa mudar a posição inicial das cartas, e isto pode ser feito de diversas maneiras, sendo uma delas o *Riffle Shuffle*. O Riffle Shuffle consiste em realizar um corte reto em uma pilha e em seguida intercalar os dois montes juntamente, a fim de formar uma única pilha novamente. Este é o método mais comum de embaralhamento de cartas, por isso, neste trabalho denotaremos o Riffle Shuffle por *embaralhamento canônico*.

Um modelo matemático preciso para o embaralhamento canônico foi introduzido pelos matemáticos Edgar N. Gilbert e Claude Shannon em 1955, e independentemente por Jim Reeds em 1981. De acordo com [1], temos a seguinte definição:

Definição 5 (O modelo de Gilbert-Shannon-Reeds: modelo GSR)**.** *Dado um baralho de tamanho n em uma pilha, realizamos um corte reto dividindo a pilha em dois montes, monte 1 e monte 2, onde o número de cartas do monte 1 segue uma distribuição binomial de parâmetros n e* $\frac{1}{2}$*. Isto é, a probabilidade de que k cartas sejam separadas,* $0 \leq k \leq n$*, é de* $\frac{\binom{n}{k}}{2^n}$*. Em seguida intercalamos os dois montes a fim de formar uma nova pilha, adicionando a cada vez uma carta vinda do monte 1 ou do monte 2 com probabilidade proporcional ao número de cartas em cada monte. Ou seja, se temos A cartas no monte 1 e B cartas no monte 2, a próxima carta a ser colocada na nova pilha que está se formando tem probabilidades* $\frac{A}{A+B}$ *de vir do monte 1 e* $\frac{B}{A+B}$ *de vir do monte 2.*

Exemplo 1 (Embaralhamento canônico)**.** *Considere um baralho de tamanho* $n = 8$*, a partir da pilha identidade (Figura 1.1), escolhendo* $k = 5$*, realizamos um embaralhamento canônico (Figura 1.2) intercalando os montes* $(1, 2, 3, 4, 5)$ *e* $(6, 7, 8)$*, obtendo a pilha* $(1, 2, 6, 3, 4, 7, 8, 5)$*, como mostra a Figura 1.3.*

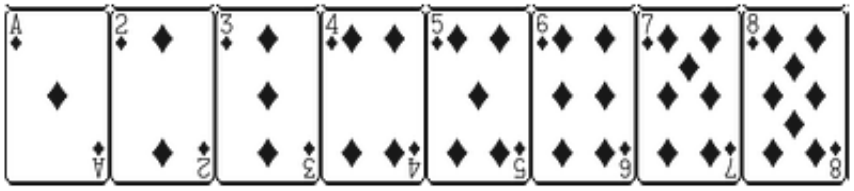

Figura 1.1: Pilha identidade

A seguir definimos o embaralhamento inverso de um embaralhamento canônico, de acordo com [4].

Definição 6 (Embaralhamento canônico inverso)**.** *Consideremos um baralho de tamanho n em uma pilha* $b = (b_1, \cdots, b_n)$*, e uma sequência* (W_i)*,* $i \in \{1, \cdots, n\}$*, de variáveis aleatórias i.i.d., com distribuição uniforme em* $\{0, 1\}$*. Associamos a cada carta* b_i *o valor* W_i*, e em seguida, posicionamos as cartas* b_i *tais que* $W_i = 0$ *por cima das cartas*

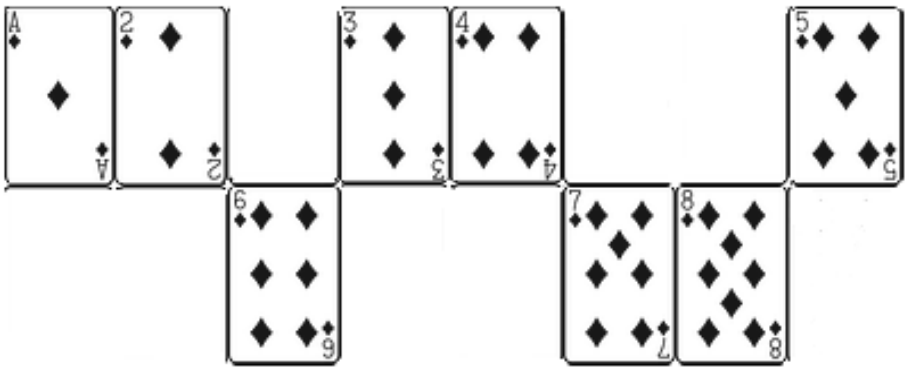

Figura 1.2: Intercalando os montes

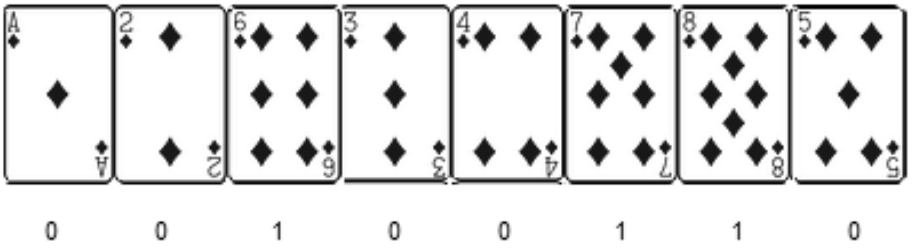

Figura 1.3: Pilha resultante a partir do embaralhamento canônico

b_j, *tais que* $W_j = 1$, *preservando a ordem relativa entre as cartas associadas a valores iguais.*

As imagens a seguir mostram a realização de um embaralhamento canônico inverso. A partir da pilha identidade para $n = 8$, atribuindo as variáveis aleatórias $W_1 = 1, W_2 = 1, W_3 = 0, W_4 = 1, W_5 = 0, W_6 = 0, W_7 = 1$ e $W_8 = 0$, ou seja, associando à pilha identidade a sequência $(1, 1, 0, 1, 0, 0, 1, 0)$, obtemos a nova pilha $(3, 5, 6, 8, 1, 2, 4, 7)$. Veja Figura 1.4 e Figura 1.5.

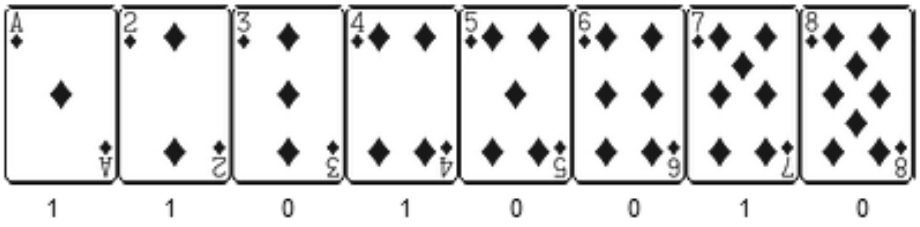

Figura 1.4: Sequências de variáveis aleatórias

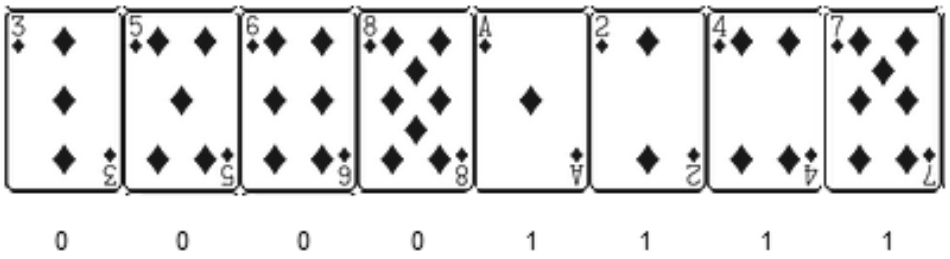

Figura 1.5: Pilha resultante a partir do embaralhamento canônico inverso

Considerando a pilha representada pela Figura 1.3, obtida a partir de um embaralhamento canônico, ao realizarmos o embaralhamento canônico inverso dado por $W_1 = 0, W_2 = 0, W_3 = 1, W_4 = 0, W_5 = 0, W_6 = 1, W_7 = 1$ e $W_8 = 0$ obtemos novamente a pilha identidade. Portanto, a sequência $(0, 0, 1, 0, 0, 1, 1, 0)$ corresponde ao inverso do embaralhamento canônico de um baralho de tamanho 8, cuja pilha resultante seja $(1, 2, 6, 3, 4, 7, 8, 5)$.

1.2 Mágica Premo

As definições feitas na seção anterior nos possibilitam a apresentação e a análise de uma importante mágica com cartas, chamada *Premo*. Essa mágica é o ponto de partida para o trabalho realizado em [1], e a atual seção será dedicada à descrição e à análise da mágica de acordo com esta referência.

1.2.1 A performance

O mágico convida um espectador para ser seu voluntário na realização da mágica. O mágico entrega ao voluntário um baralho em uma pilha e pede para que este dê um corte, dois embaralhamentos canônicos consecutivos e em seguida mais um corte. Durante esse processo o mágico não toca no baralho e nem ao menos olha para o baralho. Em seguida o mágico convida o voluntário a olhar a carta posicionada no topo da pilha, memorizá-la e recolocá-la na pilha, desta vez não no topo, mas aproximadamente no meio da pilha. O espectador é convidado a realizar mais um embaralhamento canônico

e finalizar com um corte. O mágico pega o baralho e abre a pilha em uma mesa, com as faces das cartas viradas para cima. Para a surpresa da audiência, ele rapidamente encontra a carta que o voluntário memorizou.

1.2.2 Análise da mágica

Começaremos a análise da mágica pelo baralho. Para a realização desta mágica não importa a quantidade exata de cartas, a numeração das cartas ou os naipes a serem utilizados, o único fator importante é a ordenaração inicial das cartas, ou seja, a pilha inicial. Para facilitar a análise, consideremos que o baralho esteja inicialmente na pilha identidade $(1, 2, \cdots, n)$.

Antes de considerarmos a mágica completa, vamos investigar o efeito que um embaralhamento canônico produz na ordem das cartas. Pela Definição 5, para a realização de um embaralhamento canônico primeiramente devemos efetuar um corte reto na pilha, e assim dividimos o baralho em dois: um monte composto pelas cartas $1, 2, ..., k$, e outro compostos pelas cartas $k+1, k+2, ..., n$, com $0 \leq k \leq n$, $k \in \mathbb{N}$. Ao intercalarmos esses dois montes obtemos uma nova pilha, mas observe que as cartas $1, 2, ..., k$ permanecem na mesma ordem relativa na pilha, assim como as cartas $k+1, k+2, ..., n$, reveja a Figura 1.3, o que nos leva a seguinte definição.

Definição 7 (Sequências levantadoras)**.** *Dado um baralho em uma pilha qualquer, uma sequência levantadora é um conjunto maximal de cartas consistindo em sucessivos valores apresentados em ordem.*

Exemplo 2. *Suponha que temos as cartas* $1, 2, ..., 10$ *na pilha*

$$b = (5, 1, 3, 8, 2, 4, 9, 10, 6, 7)\,.$$

Para esse arranjo de cartas temos 4 sequências levantadoras, que são $(5, 6, 7)$*,* $(1, 2)$*,* $(3, 4)$ *e* $(8, 9, 10)$*.*

As sequências levantadoras foram descobertas independentemente pelos mágicos Charles Oswald Williams e Charles Thornton Jordan no início do Século XX. Cada elemento de S_n, ou seja, cada permutação de um baralho de tamanho n, possui r

sequências levantadoras, com $r \in \{1, 2, \cdots, n\}$. A permutação que possui $r = 1$ é a permutação identidade, e a permutação com $r = n$ é a permutação $(n, n-1, \cdots, 3, 2, 1)$.

Fixada uma pilha, notamos que duas sequências levantadoras desta pilha não se intersectam, e, sendo assim, podemos representar uma pilha qualquer de S_n como a união das suas sequências levantadoras de maneira única, porém, a união de sequências levantadoras não determinam uma única pilha. Ilustramos esta afirmação com o Exemplo 2, não é possível representarmos a pilha b a partir de suas sequências levantadoras de uma maneira distinta de $\{(A, 2), (3, 4), (5, 6, 7), (8, 9, 10)\}$, a menos de ordem. Porém, observe que $\{(A, 2), (3, 4), (5, 6, 7), (8, 9, 10)\}$ também pode expressar a pilha $c = (A, 5, 3, 2, 8, 4, 9, 6, 10, 7)$, sendo que $b \neq c$.

Como para a mágica consideramos a pilha inicial sendo a pilha identidade, inicialmente temos apenas uma sequência levantadora. Após a realização de um embaralhamento canônico, a nova pilha poderá apresentar uma ou duas sequências levantadoras. No corte reto que compõe o embaralhamento canônico, escolhendo $k = 0$ ou $k = n$ não há como intercalarmos as cartas e portanto o baralho irá manter sua ordenação inicial, $(1, 2, ..., n)$, e a nova pilha permanecerá com apenas uma sequência levantadora. Entretanto, caso $1 \leq k \leq n - 1$, o corte reto irá quebrar a sequência levantadora em duas partes, uma parte com as cartas de 1 a k, e a outra com as cartas de $k + 1$ a n. Se ao intercalar apenas sobrepormos as cartas, obtemos novamente a pilha identidade, e portanto uma sequência levantadora, mas ao intercalarmos de outras maneiras obtemos pilhas com duas sequências levantadoras, uma vez que as cartas $1, 2, ..., k$ permanecerão na mesma ordem relativa na nova pilha, assim como as cartas $k + 1, k + 2, ..., n$.

Esse processo irá se repetir a cada nova realização de um embaralhamento canônico, o corte reto poderá quebrar cada sequências levantadora da pilha em duas e, ao intercalarmos as cartas, teremos no máximo o dobro do número de sequências levantadoras que tinhamos anteriormente. Sendo assim, após a realização de t embaralhamentos canônicos partindo da pilha identidade, a pilha resultante terá no máximo 2^t sequências levantadoras.

De volta a análise da mágica, suponhamos que o baralho não seja cortado e que a carta do topo seja movida somente após o último embaralhamento canônico. Após três

embaralhamentos canônicos o baralho terá no máximo oito sequências levantadoras, e ao movermos a carta do topo para o meio do baralho criamos mais uma dessas sequências, consistindo apenas da carta movida. Assim, quando o mágico abre o baralho com as faces das cartas viradas para cima, ele buscará uma carta que está em uma sequência levantadora composta apenas por uma única carta, ou seja, a carta que se encontra após o seu sucessor e antes do seu antecessor.

Agora vamos observar o efeito que um corte produz na ordem das cartas. Supondo que o baralho esteja inicialmente na pilha identidade, ao efetuarmos um corte a nova ordem será $(k+1, ..., n, 1, 2, ..., k-1, k)$, com $k \in \{1, 2, ..., n\}$, ou seja, se considerarmos a carta 1 como a sucessora da carta n, podemos ver o baralho arranjado em um loop e com um corte estamos apenas rotacionando o loop. De volta a mágica, quando o espectador corta o baralho uma vez o mágico passa a não saber mais o início do loop, mas ainda assim, após um embaralhamento canônico teremos as sequências levantadoras, a saber uma começando com $k+1$ e a outra terminando com k, que é o que realmente interessa para a realização desta mágica. Os demais cortes irão rotacionar o loop novamente e também poderão criar e modificar as sequências levantadoras, tornando assim mais difícil a tarefa do mágico de encontar a carta que o espectador moveu.

Como mencionado anteriormente, uma maneira de verificar qual carta foi movida pelo voluntário é procurar pela carta que está antes de sua antecessora e depois de sua sucessora, e isto pode ser feito atribuindo um contador a cada carta.

Em [1], o seguinte contador é apresentado: seja $\sigma(i)$ a posição da carta i, denotamos por $d(i, j)$ o menor inteiro positivo obtido por $\sigma(j) - \sigma(i) \pmod{n}$, e a cada carta i associamos o contador

$$c(i) = d(i-1, i) + d(i, i+1) - 1.$$

Idealmente a carta movida será a única com o contador maior ou igual a n. Este fato é apenas brevemente mencionado em [1], por completude, apresentamos o seguinte teorema no qual provamos que idealmente o contador da carta movida será maior ou igual a n, porém o contador das demais cartas pode variar de acordo com a sequência levantadora a qual cada carta pertence.

Teorema 1. *Considere a mágica Premo descrita na Seção 1.2.1 . Seja i a carta movida*

pelo voluntário durante a mágica, temos que:

a) $c(i) \geq n$*;*

b) $c(j)$ *pode ser maior, menor ou igual a* n*, para toda carta* $j \neq i$*.*

Demonstração. a) Seja $\sigma(i) = y$. Observe que se i está antes de seu antecessor e depois do seu sucessor, temos que $\sigma(i-1) = y+z$ e $\sigma(i+1) = y-x$, com $x > 0$ e $z > 0$. Uma vez que o baralho possui n cartas, temos que

$$x + z + 1 \leq n. \tag{1.1}$$

Denotando $a \equiv b$ se $a = b \pmod{n}$, obtemos então

$$d(i-1, i) \equiv y - y - z \equiv -z \equiv n - z$$

e

$$d(i, i+1) \equiv y - x - y \equiv -x \equiv n - x,$$

portanto $c(i) = n - z + n - x - 1 = 2n - (x + z + 1) \geq n$, onde a desigualdade decorre da Desigualdade 1.1. Ou seja, se a carta movida está após o seu sucessor e antes do seu antecessor, o seu contador será maior ou igual a n, e portanto mostramos a primeira parte do teorema.

b) Para a segunda parte, vamos analisar qual será o contador das demais cartas. Idealmente as demais cartas irão pertencer a uma sequência levantadora composta por mais cartas. Seja j a carta que estamos analisando e considere $x > 0$, $z > 0$, com $x, z \in \mathbb{N}$.

Caso 1: As cartas $j-1$, j e $j+1$ estão na mesma sequência levantadora. Sejam $\sigma(j) = y$, $\sigma(j-1) = y - x$ e $\sigma(j+1) = y + z$, temos que $x + z + 1 \leq n$, uma vez que o baralho possui n cartas. Pela definição de $d(.,.)$, temos que

$$d(j-1, j) = y - y + x \equiv x$$

e

$$d(j, j+1) = y + z - y \equiv z,$$

e portanto, $c(j) = x+z-1 = (x+z+1)-2 \leq n-2 < n$, onde a primeira desigualdade também segue da Desigualdade 1.1. Ou seja, para estas cartas o contador é estritamente menor que n.

Caso 2: As cartas $j-1$ e j pertencem à mesma sequência levantadora, mas a carta $j+1$ não. Neste caso consideremos dois subcasos.

2.1) A carta $j+1$ está antes de $j-1$. Sejam $\sigma(j) = y$, $\sigma(j-1) = y-x$ e $\sigma(j+1) = y-z$, com $0 < x < z$. Logo,

$$d(j-1, j) = y - y + x \equiv x$$

e

$$d(j, j+1) = y - z - y \equiv -z \equiv n - z.$$

Portanto, $c(j) = x + n - z - 1 < n - 1 < n$, visto que $x < z$. Logo, para estas cartas o contador é estritamente menor que n.

2.2) A carta $j+1$ está entre as cartas $j-1$ e j. Suponhamos que $\sigma(j) = y$, $\sigma(j-1) = y-z$ e $\sigma(j+1) = y-x$, com $0 < x < z$. Logo,

$$d(j-1, j) \equiv y - y + z \equiv z$$

e

$$d(j, j+1) \equiv y - x - y \equiv -x \equiv n - x,$$

e como $x < z$, temos que $c(j) = z + n - x - 1 \geq n$.

Caso 3: As cartas j e $j+1$ estão na mesma sequência levantadora, mas a carta $j-1$ não. Para este caso também temos dois subcasos.

3.1) A carta $j-1$ está entre as cartas j e $j+1$, isto é, $\sigma(j) = y$, $\sigma(j-1) = y+x$ e $\sigma(j+1) = y+z$, com $0 < x < z$. Temos então que

$$d(j-1, j) \equiv y - y - x \equiv -x \equiv n - x$$

e

$$d(j, j+1) = y + z - y \equiv z.$$

Como $z > x$, $c(j) = n - x + z - 1 \geq n$. Assim o contador será maior que n.

3.2) A carta $j-1$ está depois da carta $j+1$. Sejam $\sigma(j) = y$, $\sigma(j-1) = y+z$ e $\sigma(j+1) = y+x$, com $0 < x < z$. Portanto,

$$d(j-1, j) \equiv y - y - z \equiv -z \equiv n - z$$

e

$$d(j, j+1) \equiv y + x - y \equiv x.$$

Logo, $c(j) = n-z+x-1 < n$, uma vez que $x < z$. Ou seja, para estas cartas o contador é estritamente menor que n. Assim finalizamos a última parte do teorema. □

Pelo teorema acima, notamos que este truque pode falhar, pois podemos ter mais de uma carta com o contador maior ou igual a n, e portanto, mais de uma opção para a carta movida. Além disso, quando o voluntário move a carta do topo ele pode não retirá-la de sua sequência levantadora, ou também pode retirá-la de uma sequência levantadora, mas colocá-la em outra sequência levantadora, e portanto o seu contador também poderá ser menor ou igual a n.

Bayer e Diaconis em [1] realizaram vários experimentos de Monte Carlo que apontaram que o truque é realizado com maior sucesso quando a carta é movida após o último embaralhamento canônico. Eles programaram um computador para embaralhar as cartas t vezes, $t \in \mathbb{N}$, de acordo com o modelo GSR, cortar o baralho segundo a distribuição uniforme, mover a carta do topo para uma posição escolhida segundo a distribuição binomial e, por fim, cortar novamente. O computador calculava o contador de cada carta e selecionava a carta com o contador mais alto. Caso houvesse mais de uma carta com o maior contador, o computador selecionava uma delas de modo equiprovável.

A seguir apresentamos uma tabela extraída de [1] referente ao experimento citado acima. A tabela é baseada em 1.000.000 de experimentos de Monte Carlo, e cada entrada (i, j) expressa a probabilidade de com j embaralhamentos o mágico dizer corretamente a carta movida, sendo permitidas i tentativas.

Na Tabela 1.1 as colunas indicam o número t de embaralhamentos canônicos consecutivos, com $2 \leq t \leq 12$, e cada linha representa a quantidade de tentativas de acerto que são permitidadas. Segundo a Tabela 1.1, notamos que com três embaralhamentos canônicos temos uma probabilidade de 0,839 de acertar a carta movida com apenas

t	2	3	4	5	6	7	8	9	10	11	12
1	0,997	0,839	0,228	0,088	0,042	0,028	0,023	0,021	0,020	0,020	0,019
2	1,000	0,943	0,471	0,168	0,083	0,057	0,047	0,042	0,040	0,039	0,039
3	1,000	0,965	0,590	0,238	0,123	0,085	0,070	0,063	0,061	0,059	0,058
13	1,000	0,998	0,884	0,617	0,427	0,334	0,290	0,270	0,260	0,254	0,252
26	1,000	0,999	0,975	0,835	0,688	0,596	0,548	0,524	0,513	0,505	0,503

Tabela 1.1: Probabilidade de sucesso na realização da mágica

uma tentativa, e essa probabilidade aumenta para 0,943 se são permitidas duas tentativas. Uma observação interessante é que mesmo com oito embaralhamentos canônicos consecutivos, após 26 tentativas ainda temos uma probabilidade de acerto considerável maior que $\frac{1}{2}$.

O fator que torna importante a mágica Premo é que ela nos mostra que com poucos embaralhamentos canônicos o baralho ainda guarda memória de sua ordenação inicial, possibilitando assim que o mágico consiga "adivinhar" qual carta foi movida pelo voluntário. O objetivo do Capítulo 2 será encontrar o número de embaralhamentos canônicos necessários para que o baralho não seja mais capaz de guardar a memória da sua ordem inicial, ou seja para que o baralho esteja bem embaralhado.

1.2.3 Um pouco sobre Charles Thornton Jordan

Charles Thornton Jordan nasceu no final do Século XIX em Berkeley, Califórnia. Além de ser um construtor de rádios habilidoso e criador de galinhas, ele tinha como hobby inventar mágicas e participar de desafios publicados em jornais. Participando destes desafios, Jordan ganhou diversos prêmios e com o passar do tempo foi proibido de envolver-se neste tipo de concurso, mas isso não o impediu de concorrer. Jordan e sua equipe contrataram um mágico da época, Blackledge, para se apresentar durante os desafios no lugar deles, mas Jordan e sua equipe forneciam as respostas para Blackledge.

Jordan foi o primeiro grande inventor de truques matemáticos com cartas e também o primeiro mágico a usar o princípio por trás das sequências de de Bruijn. Porém, ele era muito tímido para fazer performaces em público, o que o levou a criar mágicas para

vendê-las a outros mágicos. Um dos mais celebrados truques de Jordan é o *Leitor de mentes a longa distância*, esse truque foi anunciado na principal revista de mágicas da época, Sphinix, em maio de 1916 e o truque estava sendo vendido por uma quantia de 25 vezes o valor da revista.

A performace desta mágica é descrita em [3], como se segue: *O mágico envia por correio um baralho comum de cartas, em forma de pilha, a uma pessoa, solicitando-lhe que realize algumas operações com o baralho. Primeiramente a pessoa deve cortar o baralho quantas vezes desejar, realizar um embaralhamento canônico e cortar novamente quantas vezes desejar. Em seguida a pessoa deve dar um corte reto dividindo o baralho em dois montes, selecionar uma carta no centro de um dos montes, memorizá-la, e coloca-lá no centro do outro monte. A pessoa escolhe um dos dois montes, realiza mais um embaralhamento canônico e retorna este monte ao mágico, sem revelar se nele está ou não contida a carta memorizada. Em seguida, o mágico envia-lhe o nome correto da carta que a pessoa havia escolhido.*

O truque *Leitor de mentes a longa distância* foi a base para a mágica Premo, pois em sua realização também se utiliza o conceito de sequências levantadoras. Será a partir deste conceito que na próxima seção responderemos à seguinte questão: qual a probabilidade de obtermos o baralho em uma pilha desejada, após um certo número de embaralhamentos canônicos?

1.3 A distribuição GSR

Esta seção será dividida em duas partes, ambas relacionadas ao modelo GSR. Na primeira parte relacionaremos esse conceito com as sequências levantadoras, calculando a probabilidade de obtermos uma configuração com um certo número de sequências levantadoras, a partir de um único embaralhamento canônico, assim como feito em [4]. Já na segunda parte vamos generalizar o modelo GSR, oferecendo descrições alternativas, de acordo com [1], a fim de poder analisar consecutivos embaralhamentos canônicos.

1.3.1 Sequências levantadoras e o modelo GSR

Nesta seção estudaremos a distribuição de probabilidade induzida em S_n pelo embaralhamento canônico. Denotaremos por $P(\varphi, \theta)$ a probabilidade de com apenas um embaralhamento canônico levarmos a pilha φ na pilha θ.

Relembramos que, para obter a permutação θ a partir da permutação φ, precisamos embaralhar o baralho segundo a permutação $\psi = \theta \circ \varphi^{-1}$. Assim, definimos a distribuição de probabilidade induzida em S_n por $Q : S_n \to S_n$,

$$Q(\psi) = Q(\theta \circ \varphi^{-1}) = P(\varphi, \theta). \tag{1.2}$$

No caso em que a pilha inicial é a pilha identidade, temos

$$Q(\theta) = P(id, \theta).$$

Chamaremos esta distribuição de *distribuição GSR*.

O teorema a seguir relaciona a distribuição GSR com sequências levantadoras.

Teorema 2. *A probabilidade de que um baralho, inicialmente na pilha identidade, apresente a permutação ρ após a realização de um embaralhamento canônico é:*

$$Q(\rho) = P(id, \rho) = \begin{cases} (n+1)\frac{1}{2^n}, & \text{se } \rho \text{ apresenta uma sequência levantadora} \\ \frac{1}{2^n}, & \text{se } \rho \text{ apresenta duas sequências levantadoras} \\ 0, & \text{se } \rho \text{ apresenta mais de duas sequências levantadoras.} \end{cases}$$

Demonstração. Na Seção 1.2.2 mostramos que com um embaralhamento canônico geramos no máximo duas sequências levantadoras, assim, se a configuração ρ apresenta três ou mais sequências levantadoras, não é possível obtê-la com apenas um embaralhamento canônico.

Para gerar uma configuração com apenas uma sequência levantadora, após o embaralhamento canônico devemos obter a configuração inicial do baralho novamente. Logo, o embaralhamento canônico dado deve ser equivalente a permutação identidade, sendo assim, $\rho = id$.

Pela Definição 5, temos que é possível obter a mesma configuração inicial de dois modos distintos. Primeiramente, no modelo GSR podemos escolher $k = 0$ ou $k = n$ e

assim teriamos um monte vazio e outro com n cartas, dessa forma a única maneira de intercalar os dois montes seria retomando a configuração inicial. Cada um desses dois casos tem probabilidade igual a $\frac{1}{2^n}$. Outra maneira seria dividir a pilha em montes de k e $n-k$ cartas, com $k \in \{1, 2, \cdots, n\}$, e depois apenas sobrepor o monte com k cartas sobre o monte com $n-k$ cartas. Escolhemos k com probabilidade $\frac{\binom{n}{k}}{2^n}$ e em seguida devemos sobrepor os dois montes.

Segundo a Definição 5, ao intecalar os dois montes adicionamos a cada vez uma carta vinda do monte 1 ou do monte 2 com probabilidade proporcional ao número de cartas em cada monte, assim, como no primeiro monte temos k cartas e no segundo monte temos $n-k$ cartas, para toda intercalação possível teremos a probabilidade

$$\frac{k(k-1)\cdots 1(n-k)(n-k-1)\cdots 1}{n(n-1)\cdots(n-k)\cdots 1} = \frac{k!(n-k)!}{n!} = \binom{n}{k}^{-1}.$$

Portanto, após o corte reto, a probabilidade de intercalar montes com tamanhos k e $n-k$ de maneira a formar uma configuração específica é $\binom{n}{k}^{-1}$.

Ou seja, para essa maneira, a probabilidade de retornar a configuração inicial é

$$\frac{\binom{n}{k}}{2^n}\binom{n}{k}^{-1} = \frac{1}{2^n},$$

e observe que temos $n-1$ opções para k. Portanto, a probabilidade de um embaralhamento canônico gerar uma configuração com apenas uma sequência levantadora é:

$$\frac{1}{2^n} + \frac{1}{2^n} + \sum_{k=1}^{n-1} \frac{1}{2^n} = (n+1)\frac{1}{2^n}.$$

Suponhamos agora que após um embaralhamento canônico o baralho apresente uma permutação fixa que possua duas sequências levantadoras, uma com as cartas de 1 a k e a outra com as cartas de $k+1$ a n. Neste caso o baralho foi cortado na k-*ésima* carta, o que ocorre com probabilidade $\frac{\binom{n}{k}}{2^n}$. Após o corte temos que intercalar as cartas de maneira a formar a configuração apresentada pelo baralho, e, como mostrado anteriormente, isso acontece com probabilidade $\binom{n}{k}^{-1}$.

Portanto, a probabilidade de que um embaralhamento canônico gere uma permutação específica com duas sequências levantadoras é

$$\frac{\binom{n}{k}}{2^n}\binom{n}{k}^{-1} = \frac{1}{2^n}.$$

□

O Teorema 2 afirma que a probabilidade de obtermos as configurações σ e ρ a partir da realização de um embaralhamento canônico são as mesmas, se σ e ρ possuem o mesmo número de sequências levantadoras. Além disso, este teorema diz que fixado o tamanho do baralho, a probabilidade de obtermos uma permutação ρ específica de S_n após a realização de um embaralhamento canônico depende apenas do número de sequências levantadoras de ρ.

1.3.2 Generalizando o modelo GSR

Ao realizarmos um embaralhamento canônico, inicialmente dividimos o baralho em dois montes. Vamos generalizar o embaralhamento canônico permitindo que o baralho seja cortado em a montes, com $a \geq 2$, $a \in \mathbb{N}$, e depois intercalar estes a montes. Este procedimento gera descrições alternativas para o modelo GSR, que apresentaremos abaixo.

Nas seguintes descrições, consideremos $a \geq 2$, $a \in \mathbb{N}$.

a) Descrição geométrica: consideramos o intevalo $[0, 1]$ da reta. Colocamos n pontos neste intervalo de maneira independente e com distribuição uniforme, e feito isto, rotulamos os pontos de acordo com a ordem: $x_1 < x_2 < ... < x_n$. Ou seja, chamamos de x_1 o ponto mais próximo de zero, enquanto x_n é o ponto mais próximo de um. O mapa

$$x \to ax(\text{mod } 1)$$

leva o intervalo $[0, 1]$ nele mesmo rearranjando os pontos x_i, $1 \leq i \leq n$. A esse procedimento damos o nome de *a-embaralhamento*.

Observe que o embaralhamento canônico é um 2-embaralhamento, no qual os intervalos $[0, \frac{1}{2}]$ e $[\frac{1}{2}, 1]$ são "esticados" e intercalados. A transformação

$$f : [0, 1] \to [0, 1]\,,$$

$$f(x) = 2x(\text{mod } 1),$$

é conhecida como a *transformação do padeiro*.

b) Descrição da máxima entropia: todas as possíveis maneiras de cortar um baralho em a montes e intercalá-los são igualmente prováveis, permitindo que haja montes vazios inclusive, ou seja, formados por nenhuma carta.

c) Descrição inversa: considerando um baralho embaralhado, todas as possíveis maneiras de colocá-lo separadamente em a montes são igualmente prováveis, permitindo que haja montes vazios.

Uma maneira de realizarmos um a-embaralhamento inverso é a seguinte: dado um baralho de tamanho n em uma pilha, retiramos as cartas uma a uma, virando-as de modo a ficar com a face voltada para cima e em seguida escolhemos um dos a montes, dispostos lado a lado, para colocá-la de modo uniforme e independente. Após distribuirmos todas as cartas, sobrepomos os a montes da esquerda para a direita e o baralho, completo mais uma vez, é virado de modo a ficar com as faces voltadas para baixo novamente.

d) Descrição sequencial: escolhemos inteiros j_i, com $0 \leq j_i \leq n$ e $1 \leq i \leq a$, de acordo com a distribuição multinomial

$$P(j_1, ..., j_a) = \frac{\binom{n}{j_1,...,j_a}}{a^n},$$

sendo $\sum_{i=1}^{a} j_i = n$.

Escolhidos os $j_i's$, cortamos as j_1 primeiras cartas do baralho, em seguida cortamos as j_2 próximas cartas, e assim por diante, produzindo assim a montes, cada um deles com j_i cartas, com $i = 1, 2, \cdots, a$. Embaralhamos os montes de tamanho j_1 e j_2 segundo o modelo GSR de modo a formar um único monte, em seguida embaralhamos o monte resultante com o monte que possui j_3 cartas, e assim sucessivamente até obter uma pilha de tamanho n. Esse procedimento é equivalente a embaralhar todos os a montes juntamente de uma só vez.

O método apresentado na descrição sequencial é comum e eficaz para embaralharmos muitas cartas. Por exemplo, em cassinos nos jogos em que são utilizados dois baralhos, 104 cartas, é comum dividir o baralho em quatro montes, A, B, C e D, e embaralhar os montes A e B juntamente, e os montes C e D juntamente e depois embaralhar os dois montes resultantes.

O seguinte teorema relaciona as quatro descrições alternativas para o modelo GSR apresentadas acima. Esse teorema é originalmente apresentado em [1].

Teorema 3. *As quatro descrições alternativas para o modelo GSR geram a mesma distribuição de probabilidade. Ainda mais, um a-embaralhamento seguido de um b-embaralhamento é equivalente a um ab-embaralhamento.*

Demonstração. Mostraremos que cada uma das descrições resulta em um número multinomial de cartas em cada monte, e já sabemos que isto é por definição a descrição sequencial.

Considere a descrição inversa, para cada uma das n cartas devemos escolher independentemente um dos a montes para colocá-la com probabilidade $\frac{1}{a}$ para cada monte. Denotando por X_i o número de cartas do monte i, com $1 \leq i \leq a$, temos que,

$$P\left(X_1 = j_1; X_2 = j_2; \cdots; X_a = j_a\right) =$$

$$\binom{n}{j_1} \times \frac{1}{a^{j_1}} \times \binom{n-j_1}{j_2} \times \frac{1}{a^{j_2}} \times \cdots \times \binom{n-j_1-j_2-\cdots-j_{a-1}}{j_a} \times \frac{1}{a^{j_a}} =$$

$$\binom{n}{j_1, ..., j_a} \frac{1}{a^n}.$$

Para a descrição geométrica, o tamanho dos montes é determinado pelo número de pontos em cada intervalo $\left[\frac{i-1}{a}, \frac{i}{a}\right]$, com $i \in \{1, 2, \cdots, a\}$. Novamente, temos n pontos para distribuir de maneira independente e com distribuição uniforme no intervalo $[0, 1]$, e portanto temos que a probabilidade de cada ponto cair em um intervalo $\left[\frac{i-1}{a}, \frac{i}{a}\right]$ é $\frac{1}{a}$. De maneira análoga à descrição inversa, encontramos a distribuição multinomial para o número de cartas de cada monte.

Considerando agora a descrição da máxima entropia, o número de possíveis intercalações a partir de um dado corte é multinomial, existe uma bijeção com as maneiras de se dividir o baralho em a montes com o tamanho do monte correspondente. Isto mostra a primeira parte.

Dado os tamanhos dos montes, a descrição da máxima entropia assegura que todas as formas de intercalar são igualmente prováveis, assim como na descrição inversa. Consideremos a descrição sequencial, quando embaralhamos os dois primeiros montes,

de tamanhos j_1 e j_2, a probabilidade de aparecer qualquer sequência específica da esquerda para a direita é

$$\frac{j_1(j_1-1)\cdots 1 \cdot j_2(j_2-1)\cdots 1}{(j_1+j_2)(j_1+j_2-1)\cdots 1} = \binom{j_1+j_2}{j_1}^{-1}.$$

Quando essas cartas são embaralhadas com o terceiro monte, de tamanho j_3, analogamente ao feito acima, temos que todas as $\binom{j_1+j_2+j_3}{j_3}^{-1}$ posições para as cartas são igualmente prováveis, e esse processo continua para cada monte sucessivo.

Pelo já mostrado, a regra do produto para uma sequência de embaralhamentos vale em cada um desses modelos uma vez que vale para um deles. Isso sai facilmente da descrição inversa: primeiramente dividimos o baralho, carta a carta, em a montes e depois empilhamos estes montes de forma que o primeiro monte fica sobre o segundo monte e assim por diante até terminar com o a-ésimo monte, o último. Em seguida dividimos o baralho novamente, carta a carta, mas desta vez em b montes, então as cartas que estavam no primeiro monte serão divididas em b montes, as cartas que estavam no segundo monte serão divididas em b montes, e assim por diante, até finalizar com as cartas do a-ésimo monte, o que é equivalente a produzir ab montes. A Figura 1.6 ilustra um ab-embaralhamento.

□

No Teorema 2 analisamos a realização de um embaralhamento canônico, gostaríamos de também poder analisar diversos embaralhamentos canônicos consecutivos. Sejam $P^t(\varphi, \theta)$ a probabilidade de com t embaralhamentos canônicos levarmos a pilha φ na pilha θ, e

$$Q^t(\theta \circ \varphi^{-1}) = P^t(\varphi, \theta)$$

a distribuição de probabilidade induzida por t embaralhamentos canônicos consecutivos.

Com a generalização do modelo GSR e o Teorema 3, agora já possuímos todas as ferramentas necessárias para responder à primeira questão a ser abordada neste texto: "quantos embaralhamentos são necessários para chegar em uma configuração específica do baralho?". A resposta encontra-se em [1] na forma do seguinte teorema e seus dois corolários.

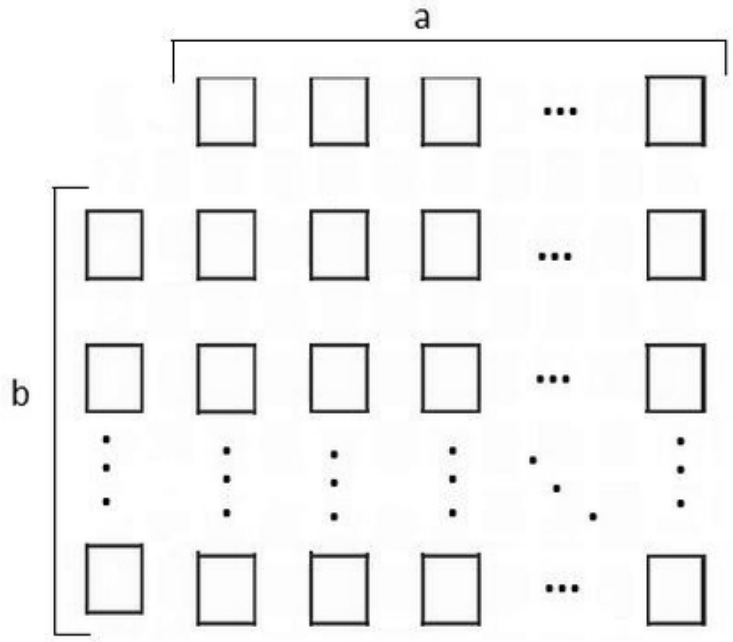

Figura 1.6: *ab*-embaralhamento inverso

Teorema 4. *A probabilidade de que um a-embaralhamento resulte na permutação φ é*

$$\frac{\binom{a+n-r}{n}}{a^n}, \tag{1.3}$$

onde r é o número de sequências levantadoras em φ.

Demonstração. Para provar este teorema, usaremos a descrição da máxima entropia. Segundo esta descrição, todas as maneiras de se cortar um baralho em a montes e intercalá-los é igualmente provável. Logo, a probabilidade será determinada pelo número de maneiras de cortar um baralho ordenado em a montes, sendo que φ é uma possível intercalação destes a montes.

Como cada monte permanece em ordem quando as cartas são embaralhadas juntamente, cada sequência levantadora no baralho embaralhado, ou seja, em φ, será uma união de montes. Sendo assim, nosso objetivo é saber o número de maneiras de obtermos r sequências levantadoras a partir de a montes. Já sabemos que pelo menos um corte dividiu cada par de sucessivas sequências levantadoras de φ, e observe que os cortes restantes podem ser realocados arbitrariamente. Assim, as n cartas formam

$n+1$ espaçamentos onde $a-r$ cortes podem ser alocados arbitrariamente, mas observe que esse processo equivale a termos $n+a-r$ objetos, n do tipo 1 (cartas) e $a-r$ do tipo 2 (cortes) e escolhermos $a-r$ posições para colocar os objetos do tipo 2, e sabemos que existem $\binom{a+n-r}{n}$ maneiras de se fazer isso.

Portanto, o número de maneiras de cortar o baralho em a montes de modo que φ seja uma intercalação possível é $\binom{a+n-r}{n}$. Como existem a^n possíveis $a-$embaralhamentos para um baralho de tamanho n, temos que a probabilidade de se obter a permutação φ com um a-embaralhamento é

$$\frac{\binom{a+n-r}{n}}{a^n}.$$

□

O próximo corolário relaciona o teorema acima a vários embaralhamentos canônicos consecutivos, já o Corolário 2 representa a forma mais usual de se embaralhar um baralho.

Corolário 1. *Se a um baralho aplicarmos uma sequência de t embaralhadas do tipo $a_1, a_2, \cdots, a_t$, então a chance de obtermos a pilha φ é dada por*

$$\frac{\binom{a+n-r}{n}}{a^n},$$

sendo que r é o número de sequências levantadoras em φ e $a = a_1 a_2 \cdots a_t$.

Demonstração. Pelo Teorema 3, temos que um a_1-embaralhamento, seguido de um a_2-embaralhamento, $\cdots$, seguido de um a_t-embaralhamento é equivalente a um $a_1 a_2 \cdots a_t$-embaralhamento, neste caso um a-embaralhamento. Basta agora aplicar o Teorema 4. □

Corolário 2. *Se um baralho de tamanho n inicialmente na pilha identidade é embaralhado t vezes segundo o embaralhamento canônico, então a probabilidade do baralho apresentar a pilha φ é*

$$P^t(id, \varphi) = \frac{\binom{2^t+n-r}{n}}{2^{tn}},$$

sendo que r é o número de sequências levantadoras em φ.

Demonstração. Segue do Corolário 1, lembrando que um embaralhamento canônico é equivalente a um 2-embaralhamento. □

O Corolário 2 generaliza o Teorema 2, pois mostra que, fixado o tamanho do baralho, a probabilidade de obtermos uma permutação específica de S_n a partir de consecutivos embaralhamentos canônicos também depende apenas do número de sequências levantadoras da permutação resultante. Relembrando a primeira pergunta, "qual a probabilidade de obtermos o baralho em uma ordem desejada após um certo número de embaralhamentos?", o Corolário 2 é a melhor resposta, pois reflete a maneira mais usual de se embaralhar, através de embaralhamentos canônicos. Respondida à primeira pergunta, no próximo capítulo vamos apresentar as ferramentas necessárias para se responder a próxima questão a ser abordada neste trabalho: como embaralhar bem um baralho?

Capítulo 2

Cadeias de Markov e embaralhamentos

No capítulo anterior, o Teorema 2 nos mostra que após a realização de um embaralhamento canônico temos a mesma probabilidade de obter permutações com o mesmo número de sequências levantadoras, mas probabilidades diferentes para permutações com diferentes números de sequências levantadoras. O Corolário 2 oferece um resultado análogo, mas para a realização de t embaralhamentos canônicos consecutivos. Ou seja, a distribuição GSR, tanto para um ou diversos embaralhamentos canônicos, não oferece o mesmo peso a todas as possíveis permutações de um baralho com n cartas. A distribuição que oferece o mesmo peso para todos os elementos do seu espaço amostral é conhecida como distribuição uniforme.

Definição 8 (Distribuição uniforme)**.** *A distribuição uniforme sobre S_n é a distribuição de probabilidade π que satisfaz $\pi(\varphi) = \frac{1}{n!}$, para cada $\varphi \in S_n$.*

Embaralhar segundo a distribuição uniforme é equivalente a afirmar que a pilha resultante pode ser dada por qualquer uma das $n!$ permutações de S_n com a mesma probabilidade, e assim o baralho provavelmente não guardará memória de sua ordenação anterior.

Neste capítulo encontraremos o número de embaralhamentos canônicos consecutivos necessários para que a distribuição induzida esteja próxima da distribuição uniforme. Nesta segunda parte do trabalho desenvolveremos um estudo sobre Cadeias de

Markov, baseado em [6], que irá nos possibilitar relacionar Cadeias de Markov com o embaralhamento canônico, e seguindo a estratégia usada em [4] e [6], mostraremos que para analisar o embaralhamento canônico basta analisar o embaralhamento canônico inverso.

A fim de evitar ambiguidades nas definições que faremos sobre Cadeias de Markov, passaremos a denotar elementos de S_n por letras do nosso alfabeto ao invés de letras do alfabeto grego. Também passaremos a omitir o sinal de composição entre duas permutações, sempre que estiver claro a operação que desejamos realizar.

2.1 O passeio aleatório em S_n

Iniciaremos essa seção com a definição de Cadeias de Markov finitas, de acordo com [6].

Definição 9 (Cadeias de Markov finitas)**.** *Considere um conjunto finito Ω. Uma Cadeia de Markov finita é um processo que se move entre os elementos de Ω da seguinte maneira: quando o processo se encontra em $x \in \Omega$, a próxima posição é escolhida de acordo com uma distribuição de probabilidade fixa $P(x, .)$. Mais precisamente, uma sequência de variáveis aleatórias $(X_0, X_1, \cdots)$ é uma Cadeia de Markov com espaço de estados Ω e matriz de transição $\boldsymbol{P}$ se para todo $x, y \in \Omega$, todo $t \geq 1$ e para todos os eventos*

$$H_{t-1} = \bigcap_{s=0}^{t-1} \{X_s = x_s\},$$

satisfazendo $P(H_{t-1} \cap \{X_t = x\}) > 0$, tem-se

$$P(X_{t+1} = y | H_{t-1} \cap \{X_t = x\}) = P(X_{t+1} = y | X_t = x) = P(x, y) \tag{2.1}$$

A Equação 2.1 é chamada de *Propriedade de Markov*, que significa que a probabilidade de no tempo $t + 1$ a cadeia estar no estado y não depende dos valores $(X_1, X_2, \cdots, X_{t-1})$, mas apenas de X_t. Ou seja, a probabilidade condicional da cadeia estar em algum estado no futuro, dado o passado e o presente, depende apenas do presente.

Considerando a matriz de transição $\mathbf{P}$ de uma Cadeia de Markov, a x-ésima linha de $\mathbf{P}$ é a distribuição $P(x, .)$. Sendo assim, em cada linha de $\mathbf{P}$ as entradas são maiores

ou iguais a zero e

$$\sum_{y \in \Omega} P(x, y) = 1 \ \forall x \in \Omega.$$

Observamos que o embaralhamento canônico induz uma Cadeia de Markov finita, uma vez que cada embaralhamento canônico produz uma permutação de S_n, que é um conjunto finito, e a probabilidade de passarmos de uma pilha para outra depende apenas da pilha atual, mais precisamente do número de suas sequências levantadoras, como mostra o Corolário 2. Como para o embaralhamento canônico o espaço de estados $\Omega = S_n$ é um grupo, temos a seguinte definição:

Definição 10 (Passeios aleatórios em grupos)**.** *Seja $(G, \circ)$ um grupo finito. Dada uma distribuição de probabilidade μ em G, definimos o passeio aleatório em G como uma Cadeia de Markov cujo espaço de estados é G e que move do seu estado atual para o próximo compondo o estado atual pela esquerda por um elemento de G de acordo com μ. Isto é, a matriz de transição $\boldsymbol{P}$ para essa cadeia tem entradas*

$$P(g, h) = \mu(h \circ g^{-1}).$$

Como dito acima, para o embaralhamento canônico temos $G = S_n$, e μ é equivalente à distribuição GSR, Q definida na Equação 1.2. Logo, pela Definição 10, o passeio aleatório em S_n com a distribuição incremento GSR é uma Cadeia de Markov com espaço de estados S_n, que move compondo a permutação atual pela esquerda por um elemento aleatório de S_n.

A matriz de transição $\mathbf{P}$ do passeio aleatório em S_n é uma matriz $n! \times n!$, sendo que cada entrada $(a_{x,y})$ é a probabilidade dada por $Q(z)$, onde z é a permutação tal que

$$z \circ x = y.$$

Cada entrada da primeira linha de $\mathbf{P}$ é a probabilidade de, o baralho estando na permutação identidade id, após a realização de um embaralhamento canônico o baralho passar a apresentar uma nova pilha x. No capítulo anterior encontramos a primeira linha da matriz de transição $\mathbf{P}$ do passeio aleatório em S_n, cada uma de suas entradas é dada pelo Teorema 2.

Seja $(X_0, X_1, \cdots)$ uma Cadeia de Markov finita com espaço de estados Ω e matriz de transição $\mathbf{P}$, e seja μ_t a distribuição de X_t, ou seja,

$$\mu_t(x) = P(X_t = x) \ \forall x \in \Omega.$$

Condicionando nos possíveis estados antecessores, temos que, $\forall\ y \in \Omega$,

$$\begin{aligned}\mu_t(y) &= \sum_{x\in\Omega} P(X_{t-1} = x)P(x, y) \\ &= \sum_{x\in\Omega} \mu_{t-1}(x)P(x, y).\end{aligned}$$

Considerando que o processo se inicia com a distribuição μ_0, a distribuição do próximo passo, μ_1, pode ser obtida multiplicando-se μ_0 por $\mathbf{P}$ pela direita: $\mu_1 = \mu_0\mathbf{P}$. Repetindo o mesmo processo, temos que $\mu_t = \mu_{t-1}\mathbf{P}$, para todo $t \geq 1$, e, portanto, para qualquer distribuição μ_0, temos que, para todo $t \geq 0$,

$$\mu_t = \mu_0\mathbf{P}^t \tag{2.2}$$

Suponhamos que a distribuição inicial esteja concentrada em apenas um estado $x \in \Omega$, ou seja, se $z \in \Omega$,

$$P(X_0 = z) = \begin{cases} 0, & \text{se } z \neq x \\ 1, & \text{se } z = x. \end{cases}$$

Denotamos por P_x essa probabilidade e temos $P_x(X_t = y) = \mathbf{P}^t(x, y)$.

Considerando o passeio aleatório no grupo S_n, a matriz de transição $\mathbf{P}^t$ equivale a uma sequência de t embaralhamentos canônicos consecutivos, sendo que cada uma de suas entradas $(a_{x,y})$ expressa a probabilidade de o baralho estando na permutação x passar a apresentar a permutação y após a realização de t embaralhamentos canôncos consecutivos. Observe que a primeira linha da matriz $\mathbf{P}^t$ também já foi calculada no capítulo anterior, no Corolário 2.

Definida a matriz de transição $\mathbf{P}^t$, podemos introduzir as seguintes classificações das Cadeias de Markov.

Definição 11 (Cadeia de Markov Irredutível)**.** *Uma Cadeia de Markov no espaço de estados Ω, com matriz de transição $\boldsymbol{P}$ é chamada irredutível se para dois estados x e y quaisquer em Ω, existe um inteiro t tal que $\boldsymbol{P}^t(x, y) > 0$.*

Definição 12 (Cadeia de Markov Aperiódica)**.** *Considere uma Cadeia de Markov com espaço de estados* Ω*. Se* $x \in \Omega$*, definimos* $T(x) = \{t \geq 1; \boldsymbol{P}^t(x,x) > 0\}$ *como o conjunto dos tempos em que é possível que a Cadeia retorne ao estado* x*. O período do estado* x *é o maior divisor comum de* $T(x)$*. Dizemos que a Cadeia é aperiódica quando o período de todo* $x \in \Omega$ *é igual a um.*

É fácil ver que o passeio aleatório em S_n é uma cadeia aperiódica, pois para todo $x \in S_n$ temos que $P(x,x) > 0$. Este fato foi provado na demonstração do Teorema 2 ao calcularmos como obter a mesma pilha após um embaralhamento canônico. Sendo assim, $1 \in T(x)\ \forall\ x \in S_n$.

Teorema 5. *O passeio aleatório em* S_n *é uma Cadeia de Markov irredutível.*

Demonstração. Sem perda de generalidade, podemos supor que o passeio começa na permutação identidade, logo, basta mostrar que

$$\forall x \in S_n, \exists\ t \in \mathbb{N},\ \text{tal que } \mathbf{P}^t(id, x) > 0.$$

Pelo Teorema 2, temos que para todo $x \in S_n$, tal que x possui uma ou duas sequências levantadoras, $\mathbf{P}(id, x) > 0$, sendo assim, já provamos que para as permutações de S_n com uma ou duas sequências levantadoras existe t tal que $\mathbf{P}^t(id, x) > 0$, neste caso $t = 1$.

Agora consideremos uma permutação com r sequências levantadoras. Pelo Corolário 2, tomando t tal que

$$\frac{\binom{2^t+n-r}{n}}{2^{tn}} > 0,$$

ou seja, $2^t - r \geq 0$, temos então que $\mathbf{P}^t(id, x) > 0$. Ou seja, se x possui r sequências levantadoras, tomando $t \geq \log_2 r$, $t \in \mathbb{N}$, temos $\mathbf{P}^t(id, x) > 0$. Sendo assim, para toda permutação $x \in S_n$ podemos encontrar um t tal que $\mathbf{P}^t(id, x) > 0$, ou seja, o passeio aleatório em S_n é uma Cadeia de Markov irredutível. □

Voltando a nossa atenção para distribuições de probabilidade, em uma Cadeia de Markov qualquer com espaço de estados Ω, uma distribuição de probabilidade μ em Ω será identificada como um vetor linha, e para todo evento $A \in \Omega$, escrevemos

$$\mu(A) = \sum_{x \in A} \mu(x).$$

Definição 13 (Distribuição Estacionária)**.** *Seja* $(X_1, X_2, \cdots)$ *uma Cadeia de Markov com matriz de transição* $\boldsymbol{P}$*. Uma distribuição de probabilidade* ν *em* Ω *é uma distribuição estacionária de* $\boldsymbol{P}$ *se* ν *satisfaz* $\nu = \nu\boldsymbol{P}$*.*

Também podemos escrever $\nu = \nu\mathbf{P}$ elemento a elemento da seguinte maneira: para todo $y \in \Omega$,

$$\nu(y) = \sum_{x\in\Omega} \nu(x)P(x, y).$$

Observe que se ν é a distribuição estacionária para uma Cadeia de Markov com matriz de transição $\mathbf{P}$ e $\mu_0 = \nu$, então, pela Equação 2.2, $\mu_t = \nu$ para todo $t \geq 0$.

Relembrando a distribuição uniforme definida no ínicio deste capítulo, o teorema a seguir assegura que a distribuição uniforme é a distribuição estacionária para um passeio aleatório em um grupo, em particular, temos que a distribuição uniforme é a distribuição estacionária para o passeio aleatório em S_n.

Teorema 6. *Seja* $\boldsymbol{P}$ *a matriz de transição de um passeio aleatório em um grupo finito* G *e* π *a distribuição uniforme em* G*. Temos que* π *é a distribuição estacionária para* $\boldsymbol{P}$*.*

Demonstração. Seja μ a distribuição de probabilidade do passeio aleatório. Para todo $g \in G$, temos que

$$\begin{aligned}
\sum_{h\in G} \pi(h)\mathbf{P}(h, g) &= \sum_{h\in G} \frac{1}{|G|}\mathbf{P}(h, g) \\
&= \frac{1}{|G|} \sum_{k\in G} \mathbf{P}(k^{-1}g, g) \\
&= \frac{1}{|G|} \sum_{k\in G} \mu(k) \\
&= \frac{1}{|G|} \\
&= \pi(g),
\end{aligned}$$

onde $k = gh^{-1}$. Logo, $\pi\mathbf{P} = \pi$, ou seja, a distribuição uniforme é a distribuição estacionária para o passeio aleatório em um grupo.

□

Finalizamos esta seção com um resultado originalmente apresentado em [1] que relaciona sequências levantadoras e Cadeias de Markov. Não apresentaremos a demonstração deste teorema pois durante a sua prova não utilizamos resultados relevantes ao estudo realizado nesta dissertação. A base da demostração é o Lema 1 da referência [7].

Teorema 7. *Considere o passeio aleatório em S_n com distribuição incremento GSR, iniciando na permutação identidade e procedendo por sucessivos a-embaralhamentos independentes. Então o número de sequências levantadoras r forma uma Cadeia de Markov.*

Em um trabalho não publicado, Ira M. Gessel desenvolveu fórmulas para as linhas da matriz de transição desta Cadeia. No artigo [5], Jonsson e Trefethen apresentam a seguinte fórmula para as entradas da Cadeia de Markov induzida pelas sequências levantadoras, produzidas a partir de sucessivos embaralhamentos canônicos,

$$p_{ij} = \frac{\binom{n+1}{2i-j}}{2^n} \alpha_j {\alpha_i}^{-1},$$

sendo i e j o número de sequências levantadoras, com $1 \leq i, j \leq n$, e α_i e α_j o número de permutações com i e j sequências levantadoras respectivamente.

2.2 O passeio aleatório inverso

Toda a seção anterior foi desenvolvida com a motivação de relacionar o embaralhamento canônico com Cadeias de Markov. Nesta seção relacionaremos o embaralhamento canônico inverso com Cadeias de Markov a fim de mostrar que a distância, conceito que ainda será introduzido na Definição 17, entre a distribuição induzida pelo embaralhamento canônico inverso e a distribuição uniforme é igual à distância entre a distribuição induzida pelo embaralhamento canônico e a distribuição uniforme. Iniciamos esta seção introduzindo o conceito de distribuição inversa em um grupo.

Definição 14 (Distribuição inversa)**.** *Para uma distribuição de probabilidade μ em um grupo G, a distribuição inversa μ^* é definida por $\mu^*(g) = \mu(g^{-1})$, para todo $g \in G$.*

Mostraremos que essa distribuição é coerente com o embaralhamento canônico inverso, ou seja, que a distribuição de probabilidade induzida pelo embaralhamento canônico inverso a partir da permutação identidade é dada por $Q^*(x) = Q(x^{-1})$:

$$Q^*(x) = \begin{cases} (n+1)\frac{1}{2^n}, & \text{se } x^{-1} \text{ apresenta uma sequência levantadora} \\ \frac{1}{2^n}, & \text{se } x^{-1} \text{ apresenta duas sequências levantadoras} \\ 0, & \text{caso contrário.} \end{cases}$$

Relembrando a Definição 6, temos 2^n possíveis sequências binárias $(W_1, \cdots, W_n)$ e, como a distribuição é uniforme em $\{0, 1\}$ para cada variável aleatória W_i, cada uma destas sequências tem probabilidade $\frac{1}{2^n}$.

Começamos com o caso de uma sequência levantadora. Observe que a permutação cuja inversa possui apenas uma sequência levantadora é a permutação identidade, cuja inversa é ela mesma. No embaralhamento canônico inverso, as cartas i com $W_i = 0$ são colocadas por cima das cartas com $W_i = 1$, observe que temos $n + 1$ sequências (W_i) que não modificam a ordenação das cartas, que são as sequências

$$W_1 = 0, W_2 = 0, \cdots, W_j = 0, W_{j+1} = 1, \cdots, W_n = 1, \tag{2.3}$$

com $j = 0, 1, \cdots, n$. Portanto, temos probabilidade de $\frac{n+1}{2^n}$ de obtermos a permutação identidade após a realização de um embaralhamento canônico inverso.

Para as sequências distintas das do tipo apresentado na Expressão 2.3, após associarmos à cada carta i a variável aleatória W_i e colocarmos as cartas com $W_i = 0$ por cima das cartas com $W_i = 1$, a pilha resultante será dada por uma permutação $x \in S_n$, $x \neq id$. Observe que a posição das cartas que são associadas à variável aleatória $W_i = 0$ formam uma sequência levantadora em x^{-1}, enquanto a posição das cartas que são associadas à variável aleatória $W_j = 1$ formam outra sequência levantadora na permutação x^{-1}, uma vez que preservamos a ordenação entre as cartas associadas às variáveis aleatórias de mesmo valor.

Exemplo 3. *Reveja a Figura 1.4 e a Figura 1.5. As cartas* $3, 5, 6$ *e* 8 *foram associadas ao valor 0 e as cartas* $1, 2, 4$ *e* 7 *foram associadas ao valor 1, obtendo assim a pemutação* $x = (3, 5, 6, 8, 1, 2, 4, 7)$. *Temos que* $x^{-1} = (5, 6, 1, 7, 2, 3, 8, 4)$, *e observe que as cartas*

nas posições $3, 5, 6$ *e* 8 *formam uma sequência levantadora em* x^{-1}*, enquanto as cartas nas posições* $1, 2, 4$ *e* 7 *formam outra sequência levantadora.*

Logo, toda permutação $x \in S_n, x \neq id$, possível de ser obtida através de um embaralhamento canônico inverso é tal que x^{-1} possui duas sequências levantadoras. Observe também que apenas uma das 2^n sequências (W_i) é capaz de produzir uma permutação x em particular. Sendo assim, a probabilidade de obtermos x a partir de um embaralhamento canônico inverso é $\frac{1}{2^n}$, se x^{-1} possui duas sequências levantadoras, mostrando assim que a definição de distribuição inversa é coerente com o embaralhamento canônico inverso.

Definição 15 (Passeio aleatório inverso)**.** *Considerando um passeio aleatório em um grupo* G *com distribuição* μ *e matriz de transição* $\boldsymbol{P}$*, chamaremos de passeio aleatório inverso de* $\boldsymbol{P}$*, ou simplesmente passeio inverso, o passeio aleatório em* G *com distribuição* μ^**, ou seja, com a distribuição inversa de* μ*.*

Em particular, como mostrado acima, para o passeio aleatório inverso em S_n, temos que $\forall\ x \in S_n$,

$$Q^*(x) = Q(x^{-1}),$$

onde Q é a distribuição GSR.

Até o momento conseguimos relacionar o embaralhamento canônico e o embaralhamento canônico inverso através da distribuição GSR, mas ainda não é o bastante. Queremos também relacionar a matriz de transição do passeio aleatório em S_n com distribuição GSR com o passeio aleatório inverso em S_n. Essa relação se dá através do teorema a seguir.

Teorema 8. *Seja* $\boldsymbol{P}$ *a matriz de transição do passeio aleatório em um grupo* G *com distribuição* μ*. Então, a matriz de transição do passeio aleatório com distribuição* μ^**, ou seja, a matriz do passeio inverso, é dada por*

$$\boldsymbol{P}^*(x, y) = \boldsymbol{P}(y, x).$$

Demonstração. Pela Definição 10, a matriz de transição $\mathbf{P}^*$ para o passeio aleatório com distribuição μ^* em um grupo tem entradas

$$\mathbf{P}^*(x, y) = \mu^*(yx^{-1}).$$

Pela definição de distribuição inversa, temos que $\mu^*(yx^{-1}) = \mu(xy^{-1})$, e juntamente com a definição da matriz de transição $\mathbf{P}$, temos que

$$\mathbf{P}^*(x, y) = \mu(xy^{-1}) = \mathbf{P}(y, x).$$

□

O Teorema 8 possui algumas consequências, a primeira delas é que a matriz de transição do passeio inverso em S_n é equivalente à transposta da matriz de transição do passeio aleatório em S_n, ou seja, denotando a transposta de $\mathbf{P}$ por $\mathbf{P}^T$, temos que $\mathbf{P}^T = \mathbf{P}^*$. Como propriedade de matrizes, temos que para uma matriz qualquer C tal que $C = AB$, a transposta de C é dada por $C^T = B^T A^T$. Observamos que a matriz de transição para o tempo $t = 2$ é tal que $\mathbf{P}^2 = \mathbf{PP}$. Sendo assim,

$$\mathbf{P}^{2^T} = \mathbf{P}^T\mathbf{P}^T = \mathbf{P}^*\mathbf{P}^* = \mathbf{P}^{*2}$$

ou seja, a matriz de transição do passeio inverso no tempo $t = 2$ também é a transposta da matriz de transição do passeio aleatório no tempo $t = 2$. Por indução, podemos provar que para $t \in \mathbb{N}$,

$$\mathbf{P}^{t^T} = \mathbf{P}^{*t},$$

e portanto $\mathbf{P}^t(x, y) = \mathbf{P}^{*t}(y, x)\ \forall\ x, y \in S_n$.

O Teorema 8 também nos possibilita mostrar que o passeio aleatório inverso em um grupo faz parte de uma categoria específica de Cadeias de Markov, conhecida como tempo reverso, que definiremos a seguir.

Definição 16 (Reversibilidade e tempo reverso)**.** *Uma Cadeia de Markov com espaço de estados Ω, matriz de transição $\boldsymbol{P}$ e distribuição estacionária ν é chamada reversível se*

$$\nu(x)\boldsymbol{P}(x, y) = \nu(y)\boldsymbol{P}(y, x), \forall x, y \in \Omega.$$

O tempo reverso de uma Cadeia de Markov irredutível com matriz de transição $\boldsymbol{P}$ e distribuição estacionária ν é a cadeia com matriz de transição

$$\boldsymbol{P}_r(x, y) = \frac{\nu(y)\boldsymbol{P}(y, x)}{\nu(x)}.$$

Como a distribuição estacionária para um passeio aleatório em um grupo G é a distribuição uniforme π, onde $\forall x \in G$ temos $\pi(x) = \frac{1}{|G|}$, então pelo Teorema 8

$$\mathbf{P}^*(x,y) = \mathbf{P}(y,x) = \frac{\frac{1}{|G|}\mathbf{P}(y,x)}{\frac{1}{|G|}} = \frac{\pi(y)\mathbf{P}(y,x)}{\pi(x)} = \mathbf{P}_r(x,y).$$

Ou seja, o passeio aleatório inverso é o tempo reverso do passeio aleatório em um grupo.

Por fim, como corolário do Teorema 8 também provamos que o passeio aleatório inverso é uma Cadeia de Markov irredutível.

Corolário 3. *O passeio aleatório inverso em S_n é uma Cadeia de Markov irredutível.*

Demonstração. Sejam x e y permutações quaisquer em S_n, $\mathbf{P}$ e $\mathbf{P}^*$ as matrizes de transição do passeio aleatório em S_n e do passeio inverso em S_n respectivamente. Já sabemos pelo Teorema 5 que existe $t \in \mathbb{N}$ tal que $\mathbf{P}^t(y,x) > 0$, como $\mathbf{P}^*(x,y) = \mathbf{P}(y,x)$, temos que $\mathbf{P}^{*t}(x,y) > 0$. □

Observe que o passeio aleatório inverso também é uma Cadeia aperiódica, Definição 12, pois para todo $x \in S_n$, ao associarmos x à qualquer sequência do tipo 2.3 obtemos a pilha x novamente. Assim $1 \in T(x)\ \forall\ x \in S_n$, e portanto o passeio aleatório inverso é uma Cadeia aperiódica.

O próximo resultado mostra que o passeio aleatório em S_n e o passeio inverso compartilham a mesma distribuição estacionária, a distribuição uniforme.

Teorema 9. *Sejam (X_t) uma Cadeia de Markov irredutível com matriz de transição $\boldsymbol{P}$ e distribuição estacionária ν, e (X_t^*) a cadeia de tempo reverso com matriz de transição $\boldsymbol{P}^*$. Sendo assim, ν é distribuição estacionária de $\boldsymbol{P}^*$ e para $x_0, \cdots, x_t \in \Omega$ temos que*

$$\boldsymbol{P}_\nu(X_0 = x_0, \cdots, X_t = x_t) = \boldsymbol{P}^*_\nu(X_0^* = x_t, \cdots, X_t^* = x_0)$$

Demonstração. Como $\mathbf{P}^*$ é a cadeia do tempo reverso, e do fato de que ν é a distribuição estacionária para $\mathbf{P}$, temos que

$$\sum_{y\in\Omega} \nu(y)\mathbf{P}^*(y,x) = \sum_{y\in\Omega} \nu(y)\frac{\nu(x)\mathbf{P}(x,y)}{\nu(y)} = \nu(x),$$

e portanto ν também é a distribuição estacionária de $\mathbf{P}^*$.

Como $\mathbf{P}^*$ é a matriz de transição da cadeia do tempo reverso da matriz $\mathbf{P}$, para cada i temos

$$\mathbf{P}(x_{i-1}, x_i) = \frac{\nu(x_i)\mathbf{P}^*(x_i, x_{i-1})}{\nu(x_{i-1})}.$$

Logo,

$$\begin{aligned}\mathbf{P}_\nu(X_0 = x_0, \cdots, X_t = x_t) &= \nu(x_0)\mathbf{P}(x_0, x_1)\cdots\mathbf{P}(x_{t-1}, x_t) \\ &= \nu(x_t)\mathbf{P}^*(x_t, x_{t-1})\cdots\mathbf{P}^*(x_1, x_0) \\ &= \mathbf{P}^*_\nu(X_0^* = x_t, \cdots, X_t^* = x_0).\end{aligned}$$

□

Com todo o estudo desenvolvido durante esta seção, criamos um paralelo entre o passeio aleatório induzido pelo embaralhamento canônico e o passeio aleatório inverso, induzido pelo embaralhamento canônico inverso, que pode ser sintetizado na tabela a seguir.

Tabela 2.1: Comparação entre o passeio aleatório e o passeio inverso em S_n

	Passeio aleatório	Passeio inverso
distribuição incremento	Q	Q^*
matriz de transição	$\mathbf{P}(x, y)$	$\mathbf{P}^*(x, y) = \mathbf{P}(y, x)$
irredutível	sim	sim
distribuição estacionária	distribuição uniforme	distribuição uniforme

O objetivo deste capítulo é ver o quão perto podemos chegar da distribuição uniforme a partir da distribuição GSR. Com o paralelo entre o passeio aleatório e o passeio inverso em S_n, provaremos mais adiante que para alcançar o nosso objetivo podemos passar a analisar o embaralhamento canônico inverso, mas antes será necessário introduzir uma maneira de medir a distância entre duas distribuições.

Definição 17 (Distância de Variação Total)**.** *Sejam μ e ν duas distribuições de probabilidade em Ω. A distância de varição total entre μ e ν é dada por*

$$\|\mu - \nu\|_{VT} = \max_{A\subset\Omega} |\mu(A) - \nu(A)|.$$

Em particular, se $\boldsymbol{P}$ *é uma matriz de transição e* ν *sua distribuição estacionária, definimos*

$$d(t) = \max_{x \in \Omega} \left\| P^t(x, .) - \nu \right\|_{VT}$$

como a distância entre a distribuição da Cadeia no tempo t *e a sua distribuição estacionária.*

A definição apresentada de Distância de Variação Total pode não ser fácil de ser aplicada, por ter que analisar subconjuntos de Ω. A seguir apresentaremos uma caracterização alternativa da Distância de Variação Total, que se mostrará mais útil para este trabalho.

Teorema 10. *Sejam* μ *e* ν *duas distribuições de probabilidade no mesmo espaço de estados* Ω*. Então*

$$\|\mu - \nu\|_{VT} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|. \tag{2.4}$$

Demonstração. Seja $B = \{x; \mu(x) \geq \nu(x)\}$ e seja A um evento qualquer em Ω, temos que

$$\begin{aligned}
\mu(A) - \nu(A) &= \sum_{x \in A} \mu(x) - \sum_{x \in A} \nu(x) \\
&= \left(\sum_{x \in A \cap B} \mu(x) + \sum_{x \in A \cap B^c} \mu(x) \right) - \left(\sum_{x \in A \cap B} \nu(x) + \sum_{x \in A \cap B^c} \nu(x) \right) \\
&= \left(\sum_{x \in A \cap B} \mu(x) - \sum_{x \in A \cap B} \nu(x) \right) + \left(\sum_{x \in A \cap B^c} \mu(x) - \sum_{x \in A \cap B^c} \nu(x) \right).
\end{aligned}$$

Observamos que se $x \in A \cap B^c$, então $\mu(x) - \nu(x) < 0$, sendo assim,

$$\mu(A) - \nu(A) \leq \mu(A \cap B) - \nu(A \cap B). \tag{2.5}$$

Por outro lado,

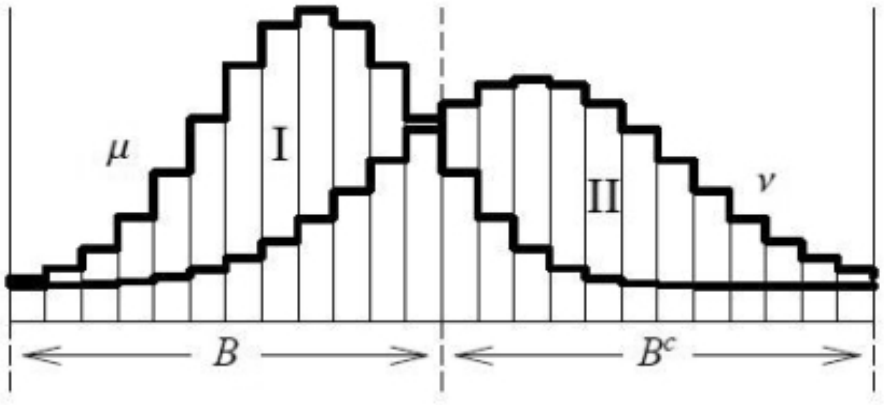

Figura 2.1: Diferença entre duas distribuições

$$\begin{aligned}\mu(B) - \nu(B) &= \sum_{x\in B}\mu(x) - \sum_{x\in B}\nu(x) \\ &= \left(\sum_{x\in A\cap B}\mu(x) + \sum_{x\in A^c\cap B}\mu(x)\right) - \left(\sum_{x\in A\cap B}\nu(x) + \sum_{x\in A^c\cap B}\nu(x)\right) \\ &= \left(\sum_{x\in A\cap B}\mu(x) - \sum_{x\in A\cap B}\nu(x)\right) + \left(\sum_{x\in A^c\cap B}\mu(x) - \sum_{x\in A^c\cap B}\nu(x)\right).\end{aligned}$$

Porém, observe que se $x \in A^c \cap B$, então $\mu(x) - \nu(x) \geq 0$, e portanto

$$\mu(A \cap B) - \nu(A \cap B) \leq \mu(B) - \nu(B). \tag{2.6}$$

A partir das Equações 2.5 e 2.6, concluimos então que $\mu(A) - \nu(A) \leq \mu(B) - \nu(B)$, para todo evento $A \subset \Omega$.

Pelo mesmo raciocínio acima podemos mostrar que

$$\nu(A) - \mu(A) \leq \nu(B^c) - \mu(B^c).$$

Observamos na Figura 2.1, extraída de [6], que a região I tem área $\mu(B) - \nu(B)$ e a região II, $\nu(B^c) - \mu(B^c)$. Como as áreas totais abaixo de μ e ν têm valor 1, temos que as regiões I e II possuem a mesma área. Logo, a cota superior para $\mu(B) - \nu(B)$ e

$\nu(B^c) - \mu(B^c)$ é a mesma, e portanto,

$$\begin{aligned}\|\mu - \nu\|_{VT} &= \frac{1}{2}\left[\mu(B) - \nu(B) + \nu(B^c) - \mu(B^c)\right] \\ &= \frac{1}{2}\sum_{x\in\Omega}|\mu(x) - \nu(x)|.\end{aligned}$$

□

A demonstração do Teorema 10 também mostra que

$$\|\mu - \nu\|_{VT} = \sum_{x\in\Omega;\ \mu(x)\geq\nu(x)}[\mu(x) - \nu(x)], \tag{2.7}$$

oferencendo assim uma terceira caracterização para a Distância de Variação Total.

O teorema a seguir mostra que tanto o passeio aleatório em S_n quanto o passeio aleátorio inverso em S_n apresentam a mesma distância da distribuição uniforme quando o passeio começa na permutação identidade. Este teorema é muito útil, pois nos possibilita analisar o embaralhamento canônico inverso ao invés do embaralhamento canônico.

Teorema 11. *Sejam $\boldsymbol{P}$ e $\boldsymbol{P}^*$ matrizes de transição de um passeio aleatório em um grupo G com distribuição μ e distribuição inversa μ^*, respectivamente. Seja π a distribuição uniforme em G, então, para todo $t \geq 0$, temos que*

$$\left\|P^t(id,.) - \pi\right\|_{VT} = \left\|P^{*t}(id,.) - \pi\right\|_{VT}.$$

Demonstração. Pelo Teorema 10, temos que

$$\left\|P^t(id,.) - \pi\right\|_{VT} = \frac{1}{2}\sum_{a\in G}\left|P^t(id,a) - |G|^{-1}\right|,$$

mas, pelo Teorema 8 temos que $P(x,y) = P^*(y,x)$, $\forall x, y \in G$. Logo $P^t(x,y) = P^{*t}(y,x)$, e portanto

$$\left\|P^t(id,.) - \pi\right\|_{VT} = \frac{1}{2}\sum_{a\in G}\left|P^{*t}(a,id) - |G|^{-1}\right|. \tag{2.8}$$

Pela definição de passeio aleatório inverso, temos que

$$P^*(a,id) = \mu^*(id\circ a^{-1}) = \mu^*(a^{-1}) = \mu^*(a^{-1}\circ id^{-1}) = P^*(id,a^{-1}). \tag{2.9}$$

Pelas Equações 2.8 e 2.9, concluimos que

$$\begin{aligned}\left\|P^t(id,.)-\pi\right\|_{VT} &= \frac{1}{2}\sum_{a\in G}\left|P^t(id,a)-|G|^{-1}\right| \\ &= \frac{1}{2}\sum_{a\in G}\left|P^{*t}(id,a^{-1})-|G|^{-1}\right|.\end{aligned}$$

Renomeando o somatório obtemos

$$\begin{aligned}\left\|P^t(id,.)-\pi\right\|_{VT} &= \frac{1}{2}\sum_{a\in G}\left|P^{*t}(id,a)-|G|^{-1}\right| \\ &= \left\|P^{*t}(id,.)-\pi\right\|_{VT},\end{aligned}$$

onde a segunda igualdade segue do Teorema 10. □

Com a Distância de Variação Total, definimos uma maneira de medir a distância entre duas distribuições. A seguir vamos introduzir um parâmetro que mede o tempo necessário para que a distância entre uma Cadeia de Markov e sua distribuição estacionária seja suficientemente pequena.

Definição 18 (Tempo de mistura)**.** *O tempo de mistura é definido por*

$$t_{mix}(\epsilon) = \min\left\{t; d(t) \leq \epsilon\right\},$$

lembrando que $d(t) = \max_{x\in\Omega}\left\|P^t(x,.)-\mu\right\|_{VT}$. *Denotamos por* $t_{mix} = t_{mix}(\frac{1}{4})$.

Com a definição acima obtemos o seguinte corolário do Teorema 11.

Corolário 4. *Se* t_{mix} *é o tempo de mistura de um passeio aleatório em um grupo e* t^*_{mix} *é o tempo de mistura do passeio inverso, então* $t_{mix} = t^*_{mix}$.

Provado que a distância da uniformidade é a mesma tanto para o passeio aleatório, quanto para o passeio aleatório inverso, a partir de agora vamos então passar a analisar apenas o embaralhamento canônico inverso. A seguir examinaremos o efeito que este embaralhamento produz na ordenação das cartas, de acordo com [4].

Considere a pilha identidade $(1, 2, \cdots, n)$, de acordo com a Definição 6. Para o primeiro embaralhamento canônico inverso, teremos cartas associadas às variáveis

aleatórias com valores 0 ou 1, sendo que as cartas associadas à variável aleatória $W_i = 0$ serão posicionadas por cima das cartas associadas à variável aleatória $W_i = 1$. Para o segundo embaralhamento, a cada carta i será associada uma variável aleatória Z_i tomando valores 0 ou 1, assim cada carta será associada a uma sequência de dois dígitos, cada um deles podendo ser zero ou um. Observe que as cartas associadas à sequência 00 estarão por cima das associadas à 10, que estarão por cima das com sequência 01 e por fim as associadas à sequência 11.

Após t embaralhamentos canônicos inversos consecutivos, cada carta estará associada a uma sequência de t dígitos, cada um dos dígitos sendo zero ou um. Seja i uma carta qualquer do baralho de tamanho n, denotaremos por

$$s(i) = (s_1(i), \cdots, s_t(i)) \in \{0,1\}^t,$$

a sequência de dígitos atribuída a i após t embaralhamentos canônicos inversos, onde $s_j(i)$ é o dígito atribuido a carta i no j-ésimo embaralhamento canônico inverso. A carta i estará posicionada antes da carta j após t embaralhamentos se:

1. $s(i) = s(j)$ e $i < j$, ou seja i já estava a frente da carta j na pilha original.

2. $s(i) \neq s(j)$ e existe $u \in \{1, 2, \cdots, t\}$ tal que $s_u(i) = 0$ e $s_u(j) = 1$, sendo que $s_v(i) = s_v(j)\ \forall v \in \{u+1, \cdots, t\}$.

Exemplo 4. *Se após 5 embaralhamentos canônicos inversos tivermos* $s(i) = (1,0,0,1,1)$ *e* $s(j) = (0,0,1,1,1)$*, a carta* i *estará a frente da carta* j*, uma vez que* $s_3(i) = 0$ *e* $s_3(j) = 1$*, sendo que o embaralhamento de número* 3 *é o último no qual se atribui dígitos diferentes a* i *e* j*.*

Considerando o baralho inicialmente na pilha identidade, após a realização de um embaralhamento canônico inverso as cartas associadas ao mesmo valor preservarão a ordem entre si, da mesma forma, após dois embaralhamentos canônicos inversos, as cartas associadas à mesma sequência de dígitos conservam a mesma ordem relativa entre si, de maneira crescente de cima para baixo. Este comportamento irá se repetir a cada novo embaralhamento canônico inverso, ou seja, as cartas associadas à mesma sequência de dígitos conservarão a ordem relativa entre si. Portanto, enquanto atribuirmos sequências iguais às diferentes cartas, existirão permutações que não poderão

ser alcançadas, a saber, aquelas nas quais as cartas que possuem a mesma sequência aparecem em ordem decrescente. Sendo assim, a distribuição uniforme só poderá ser alcançada a partir do momento em que todas as cartas estejam associadas a sequências distintas. Dessa forma, é interessante saber qual é a probabilidade de se obter sequências distintas para cada carta após um número qualquer de embaralhamentos canônicos inversos consecutivos. O teorema abaixo é a resposta para esta questão.

Teorema 12. *Considere um baralho de tamanho n em uma pilha. A probabilidade de que, com t embaralhamentos canônicos inversos, cada carta possua uma sequência de zeros e uns distinta, isto é $s(i) \neq s(j)$, $\forall i \neq j$, é dada por*

$$\prod_{i=0}^{n-1} \left(1 - \frac{i}{2^t}\right).$$

Demonstração. Após a realização de t embaralhamentos canônicos inversos temos 2^t sequências possíveis, sendo que cada uma das cartas pode apresentar qualquer uma das sequências, logo existem 2^{tn} maneiras de atribuirmos uma sequência a cada carta. Porém, gostaríamos que cada carta possuísse uma sequência própria, distinta das demais. Temos

$$(2^t)(2^t - 1)(2^t - 2) \cdots (2^t - n + 1)$$

maneiras de atribuirmos sequências a todas as cartas sem que haja repetição. Logo, a probabilidade de que com t embaralhamentos canônicos inversos cada carta possua uma sequência distinta é

$$\frac{(2^t)(2^t - 1)(2^t - 2) \cdots (2^t - n + 1)}{2^{tn}} =$$

$$1 \left(1 - \frac{1}{2^t}\right) \left(1 - \frac{2}{2^t}\right) \cdots \left(1 - \frac{n-1}{2^t}\right) = \prod_{i=0}^{n-1} \left(1 - \frac{i}{2^t}\right).$$

□

Na próxima seção iremos encontrar o valor t para o qual todas as cartas possuam sequências distintas com alta probabilidade.

2.3 A distância da uniformidade

O principal objetivo desta seção é encontrar cotas para o tempo de mistura do embaralhamento canônico inverso, e, pelo Corolário 4, também do embaralhamento canônico. A cota superior é dada pelo teorema a seguir.

Teorema 13. *Considere um baralho de tamanho n, vale a seguinte cota superior para o tempo de mistura do embaralhamento canônico*

$$t_{mix} \leq 2 \log_2 \frac{4n}{3},$$

para n suficientemente grande.

As definições e os teoremas que serão apresentados a seguir são as ferramentas necessárias para a demonstração do Teorema 13.

Definição 19. *Considere uma Cadeia de Markov finita no espaço de estados Ω, com matriz de transição $\boldsymbol{P}$. Um mapa de representação aleatória da matriz de transição $\boldsymbol{P}$ é uma função $f : \Omega \times \Lambda \longrightarrow \Omega$, juntamente com uma variável aleatória Z que toma valores em Λ, satisfazendo*

$$P(f(x, Z) = y) = \boldsymbol{P}(x, y).$$

A seguir mostramos que toda matriz de transição possui um mapa de representação aleatória, em particular, o passeio aleatório inverso em S_n.

Teorema 14. *Toda matriz de transição em um espaço de estados finito possui um mapa de representação aleatória.*

Demonstração. Considere uma Cadeia de Markov finita com espaço de estados $\Omega = \{x_1, \cdots, x_n\}$ e matriz de transição $\mathbf{P}$. Escolhemos nossa variável aleatória Z com distribuição uniforme em $\Lambda = [0, 1]$. Definimos

$$F_{j,k} = \sum_{i=1}^{k} \mathbf{P}(x_j, x_i),$$

e seja $f(x_j, z) = x_k$, quando $F_{j,k-1} < z \leq F_{j,k}$. Logo, temos que

$$P(f(x_j, Z) = x_k) = P(F_{j,k-1} < Z \leq F_{j,k}) = \mathbf{P}(x_j, x_k).$$

□

Observe que se $Z_1, Z_2, \cdots$ são variáveis aleatórias i.i.d. com a mesma distribuição que a variável aleatória Z definida na Definição 19 e X_0 com distribuição μ, $\mu(x) = P(X_0 = x)$, então a sequência $(X_0, X_1, \cdots)$ definida por

$$X_n = f(X_{n-1}, Z_n),$$

com $n = 1, 2, \cdots$, é uma Cadeia de Markov com matriz de transição $\mathbf{P}$ e distribuição inicial μ, uma vez que

$$\begin{aligned}
P(X_{t+1} &= x_{t+1} | X_0 = x_0, \cdots, X_t = x_t) \\
&= P(f(X_t, Z_{t+1}) = x_{t+1} | X_0 = x_0, \cdots, X_t = x_t) \\
&= P(f(x_t, Z_{t+1}) = x_{t+1} | X_0 = x_0, \cdots, X_t = x_t) \\
&= P(f(x_t, Z_{t+1}) = x_{t+1}) \\
&= P(f(x_t, Z) = x_{t+1}) \\
&= \mathbf{P}(x_t, x_{t+1}).
\end{aligned}$$

Ou seja, podemos aplicar o mapa f a uma sequência i.i.d $(Z_t)_{t=1}^{\infty}$ de tal forma que a sequência $(X_t)_{t=0}^{\infty}$ definida indutivamente por $X_0 = x$ e $X_t = f(X_{t-1}, Z_t)$ seja uma Cadeia de Markov com matriz de transição $\mathbf{P}$ começando de x.

O conceito de mapa de representação aleatória é necessário para definirmos o tempo de parada aleatório.

Definição 20 (Tempo de parada e Tempo de parada aleatório)**.** *Dada uma sequência $(X_t)_{t=0}^{\infty}$ de variáveis aleatórias em Ω, uma variável aleatória τ que assume valores em $\{0, 1, 2, \cdots, \infty\}$ é um tempo de parada para (X_t) se, para cada $t \in \{0, 1, \cdots\}$, existe um conjunto $B_t \subset \Omega^{t+1}$ tal que $\{\tau = t\} = \{(X_0, \cdots, X_t) \in B_t\}$. Um tempo de parada aleatório para uma Cadeia de Markov é o tempo de parada para a sequência (Z_t) definida acima.*

Aplicando sucessivos embaralhamentos canônicos inversos a um baralho de tamanho n inicialmente na pilha identidade, temos um passeio aleatório (X_t) em S_n. Se x é a permutação atual na qual o passeio se encontra, temos que

$$\mathbf{P}(x,y) = \begin{cases} \frac{1}{2^n}, & \text{se } y \text{ pode ser obtido ao associarmos a } x \text{ uma sequência de zeros e uns} \\ \frac{n+1}{2^n}, & \text{se } y = x \\ 0, & \text{caso contrário.} \end{cases}$$

uma vez que cada sequências em $\{0,1\}^n$, distintas daquelas como na Expressão **??**, produz uma permutação distinta ao ser aplicada a x, e as sequências do tipo **??** produzem a permutação x novamente.

Seja (Z_t) uma sequência de variáveis aleatórias independentes e uniformes em $\{0,1\}^n$. Se o estado atual do passeio é x, denotando por $x \circ z$ o embaralhamento canônico inverso regido pela sequência $z \in \{0,1\}^n$ a partir da permutação x, temos que

$$P(x \circ Z = y) = \mathbf{P}(x,y).$$

Assim, o passeio aleatório inverso é determidado pelas sequências

$$Z_t = (s_t(1), s_t(2), \cdots, s_t(n)),$$

lembrando que $s_t(i)$ é o dígito associado à carta i no t-ésimo embralhamento canônico inverso. Assim, temos um mapa de representação aleatória para o passeio inverso em S_n.

Definição 21. *Seja*

$$\tau = \min\left\{t \in \mathbb{N}; (s_1(i), s_2(i), \cdots, s_t(i)) \neq (s_1(j), s_2(j), \cdots, s_t(j)) \ \forall \ i \neq j\right\},$$

o primeiro tempo em que cada carta i possui uma sequência binária $s(i)$ distinta.

Observe que τ não é uma função do passeio aleatório (X_t), mas da sequência (Z_t) de zeros e uns atribuida em cada tempo t. A partir da Definição 20, podemos agora definir um Tempo Estacionário Forte.

Definição 22 (Tempo Estacionário Forte)**.** *Um tempo estacionário forte para uma Cadeia de Markov (X_t) com distribuição estacionária ν é um tempo de parada aleatório τ possivelmente dependendo do estado inicial x, tal que*

$$P_x(\tau = t, X_\tau = y) = P_x(\tau = t)\nu(y).$$

Ou seja, (X_t) tem distribuição ν e é independente de τ.

Podemos relacionar os conceitos definidos acima com o embaralhamento canônico inverso, como mostra o teorema a seguir.

Teorema 15. *Seja τ como definido na Definição 21, então τ é um tempo estacionário forte para o passeio aleatório inverso em S_n.*

Demonstração. Suponhamos que $\tau = t$. Pela definição de τ, no tempo t cada carta possui uma sequência binária $s(i)$ distinta. Como para cada carta i os termos de $s(i)$ são variáveis aleatórias escolhidas independentemente, cada uma com distribuição uniforme em $\{0, 1\}$, temos que cada sequência com t variáveis aleatórias é igualmente provável. Como para t todas as sequências são distintas, a permutação é totalmente determinada por $s(i)$, com $i = 1, 2, \cdots, n$. Sendo assim, a permutação de cartas no tempo τ é uniforme, não importando o valor de τ.

Ou seja, se $x \in S_n$, a probabilidade de que $\tau = t$ e que no tempo t o passeio esteja em x é dado por

$$P(\tau = t, X_\tau = x) = P(\tau = t)\pi(x),$$

e portanto τ é um tempo estacionário forte. □

Assim, podemos utilizar a definição de tempo estacionário forte para analisar o embaralhamento canônico inverso. A seguir apresentaremos algumas propriedades de um tempo estacionário forte que serão úteis na demonstração do Teorema 13.

Lema 1. *Seja (X_t) uma Cadeia de Markov irredutível com distribuição estacionária ν. Se τ é um tempo estacionário forte para (X_t), então para todo $t \geq 0$,*

$$P_x(\tau \leq t, X_t = y) = P_x(\tau \leq t)\nu(y).$$

Demonstração. Sejam $Z_1, Z_2, \cdots$ a sequência i.i.d. usada no mapa de representação aleatória de (X_t). Para qualquer $q \in \{1, \cdots, t\}$,

$$P_x(\tau = q, X_t = y) = \sum_{z \in \Omega} P_x(X_t = y | \tau = q, X_q = z) P_x(\tau = q, X_q = z). \tag{2.10}$$

Pela definição de tempo estacionário forte, τ é um tempo de parada para (Z_t), logo o evento $\{\tau = q\}$ é da forma $\{(Z_1, \cdots, Z_q) \in B\}$ para algum $B \subset \Omega^q$. Também, para inteiros $q, u \geq 0$, existe uma função $f_u : \Omega^{u+1} \to \Omega$, tal que

$$X_{q+u} = f_u(X_q, Z_{q+1}, \cdots, Z_{q+u}).$$

Uma vez que os vetores $(Z_1, \cdots, Z_q)$ e $(Z_{q+1}, \cdots, Z_t)$ são independentes, temos que

$$\begin{aligned} P_x(X_t = y|\tau = q, X_q = z) &= P_x(f_{t-q}(z, Z_{q+1}, \cdots, Z_t) = y|(X_1, \cdots, X_q) \in B, X_q = z) \\ &= P^{t-q}(z, y). \end{aligned}$$

Pela definição de tempo estacionário forte, temos que

$$P_x(\tau = q, X_q = z) = P_x(\tau = q)\nu(z), \tag{2.11}$$

logo, substituindo (2.11) na Equação (2.10), temos

$$P_x(\tau = q, X_t = y) = \sum_{z \in \Omega} P^{t-q}(z, y)\nu(z)P_x(\tau = q).$$

Como ν é a distribuição estacionária para esta Cadeia, temos que ν satisfaz $\nu = \nu P^{t-q}$. Portanto, o lado direito da equação acima é igual a $\nu(y)P_x(\tau = q)$. Sendo assim,

$$\begin{aligned} P_x(\tau \leq t, X_t = y) &= \sum_{q \leq t} P_x(\tau = q, X_t = y) \\ &= \sum_{q \leq t} \nu(y)P_x(\tau = q) \\ &= \nu(y) \sum_{q \leq t} P_x(\tau = q) \\ &= \nu(y)P_x(\tau \leq t). \end{aligned}$$

□

Na Definição 17 definimos $d(t) = \max_{x \in \Omega} \|P^t(x, .) - \nu\|_{VT}$ como a distância entre a distribuição da Cadeia no tempo t e a distribuição estacionária. O teorema a seguir relaciona esta distância ao tempo estacionário forte.

Teorema 16. *Se τ é um tempo estacionário forte, então*

$$d(t) \leq \max_{x \in \Omega} P_x(\tau > t).$$

Demonstração. Fixe $x \in \Omega$. Usando a Equação (2.7), a terceira caracterização de Distância de Variação Total, temos que

$$d(t) = \left\|P^t(x,.) - \nu\right\|_{VT} = \sum_{y\in\Omega;\ P^t(x,y)<\nu(y)} \nu(y) - P^t(x,y)$$
$$= \sum_{y\in\Omega;\ P^t(x,y)<\nu(y)} \nu(y)\left(1 - \frac{P^t(x,y)}{\nu(y)}\right)$$
$$\leq \max_y \left\{1 - \frac{P^t(x,y)}{\nu(y)}\right\}.$$

Agora observe que para todo $y \in \Omega$ temos

$$1 - \frac{P^t(x,y)}{\nu(y)} = 1 - \frac{P_x(X_t = y)}{\nu(y)} \leq 1 - \frac{P_x(X_t = y, \tau \leq t)}{\nu(y)}.$$

Pelo Lema 1, sabemos que $P_x(X_t = y, \tau \leq t) = \nu(y)P_x(\tau \leq t)$, logo

$$1 - \frac{P^t(x,y)}{\nu(y)} \leq 1 - \frac{\nu(y)P_x(\tau \leq t)}{\nu(y)} = P_x(\tau > t).$$

Sendo assim,

$$d(t) = \left\|P^t(x,.) - \nu\right\|_{VT} \leq \max_y \left\{1 - \frac{P^t(x,y)}{\nu(y)}\right\}$$
$$\leq \max_{x\in\Omega} P_x(\tau > t).$$

□

Agora estamos prontos para demonstrar o Teorema 13.

Demonstração. (Teorema 13).

Seja τ como na Definição 21, se a carta i é diferente da carta j, então

$$(s_1(i), s_2(i), \cdots, s_\tau(i)) \neq (s_1(j), s_2(j), \cdots, s_\tau(j)).$$

Se $\tau \leq t$ então diferentes sequências de variáveis aleatórias foram associadas a cada uma das n cartas após t embaralhamentos canônicos inversos. Portanto, pelo Teorema 12

$$P(\tau \leq t) = \prod_{i=0}^{n-1}\left(1 - \frac{i}{2^t}\right).$$

Seja $c(n) = n2^{\frac{-t}{2}}$, iremos omitir o n e denotar $c(n)$ apenas por c para simplificar a notação. Temos então que $t = 2\log_2\left(\frac{n}{c}\right)$, logo $2^t = \frac{n^2}{c^2}$, assim temos

$$\begin{aligned}\log\prod_{i=0}^{n-1}\left(1-\frac{i}{2^t}\right) &= \sum_{i=0}^{n-1}\log\left(1-\frac{i}{2^t}\right)\\ &= -\sum_{i=1}^{n-1}\left\{\frac{c^2 i}{n^2}+O\left(\frac{i}{n^2}\right)^2\right\}\\ &= -\frac{c^2 n(n-1)}{2n^2}+O\left(\frac{n^3}{n^4}\right)\\ &= -\frac{c^2}{2}+O\left(\frac{1}{n}\right).\end{aligned}$$

Logo, como tanto c quanto $P(\tau \leq t)$ são funções em n, tirando o limite quando n tende a infinito,

$$\lim_{n\to\infty}\frac{P(\tau\leq t)}{e^{-\frac{c^2}{2}}}=1,$$

obtemos um valor assintótico para $P(\tau \leq t)$.

O Teorema 16 nos diz que $d(t) \leq \max_{x\in\Omega} P_x(\tau > t)$, pelo o que fizemos até o momento, temos então que

$$d(t)\leq 1-e^{-\frac{c^2}{2}}.$$

Como $t_{mix} = \min\left\{t, d(t) \leq \frac{1}{4}\right\}$, tomando qualquer valor de c tal que c seja menor que $\sqrt{2\log(\frac{4}{3})}$, teremos uma cota para o t_{mix}. Um valor conveniente é $c = \frac{3}{4}$, e como definimos acima que $t = 2\log_2(\frac{n}{c})$, temos o resultado. □

Considerando $n = 52$ no Teorema 13, temos que o tempo de mistura é menor ou igual a 12 embaralhamentos canônicos.

Para finalizar este trabalho, vamos encontrar uma cota inferior para o tempo de mistura.

Teorema 17. *Fixe $0 < \epsilon, \delta < 1$ e considere sucessivos embaralhamentos canônicos inversos em um baralho de tamanho n. Então, para n suficientemente grande, temos*

$$t_{mix}(\epsilon)\geq(1-\delta)\log_2 n.$$

Para demonstrar este teorema precisamos do seguinte lema:

Lema 2. *Seja (X_t) uma Cadeia de Markov irredutível e aperiódica, com espaço de estados Ω, e suponha que π, a distribuição uniforme, seja a distribuição estacionária*

para esta Cadeia. Denotamos por $d_{out}(x) = |\{y; \boldsymbol{P}(x,y) > 0\}|$ *o número de estados acessíveis por* x *com apenas um passo e* Ω_t^x *o conjunto dos estados que podem ser alcançados a partir de* x *com* t *passos e seja* $\Delta = \max_{x\in\Omega} d_{out}(x)$*. Se* $\Delta^t < (1-\epsilon)\,|\Omega|$*, então*

$$t_{mix}(\epsilon) \geq \frac{\log|\Omega|\,(1-\epsilon)}{\log\Delta}.$$

Demonstração. Observe que $|\Omega_t^x| \leq \Delta^t$. Como $\Delta^t < (1-\epsilon)\,|\Omega|$, pela definição de Distância de Variação Total temos que

$$d(t) = \left\|\mathbf{P}^t(x,.) - \pi\right\|_{VT} \geq \mathbf{P}^t(x,\Omega_t^x) - \pi(\Omega_t^x) \geq 1 - \frac{\Delta^t}{|\Omega|} > \epsilon.$$

Portanto,

$$t_{mix}(\epsilon) \geq \frac{\log\left[|\Omega|\,(1-\epsilon)\right]}{\log\Delta}.$$

□

Demonstração. (Teorema 17).

Observe que para o passeio aleatório inverso, para cada passo existem no máximo 2^n estados que podem ser alcançados, uma vez que o próximo passo será determinado a partir de uma sequência de n digitos, cada um deles sendo zero ou um. Usando a mesma notação do Lema 2, temos que $\log_2\Delta \leq n$. Como S_n possui $n!$ elementos, pelo Lema 2 temos que

$$t_{mix}(\epsilon) \geq \frac{\log_2\left[n!(1-\epsilon)\right]}{n}.$$

Usando a Fórmula de Stirling

$$\log_2 n! = [1+o(1)]\,n\log_2 n,$$

temos

$$\begin{aligned} t_{mix}(\epsilon) &\geq \frac{\log_2\left[n!(1-\epsilon)\right]}{n} \\ &= \frac{\log_2 n!}{n} + \frac{\log_2(1-\epsilon)}{n} \\ &= \log_2 n\,[1+o(1)] + \frac{\log_2(1-\epsilon)}{n} \\ &= \log_2 n\left[1+o(1)+\frac{\log_2(1-\epsilon)}{n\log_2 n}\right]. \end{aligned}$$

Seja $0 < \delta < 1$, observe que para n suficientemente grande temos $o(1) \geq \frac{-\epsilon}{2}$, e o mesmo ocorre para $\frac{\log_2(1-\epsilon)}{n \log_2 n}$. Logo

$$-\left[o(1) + \frac{\log_2(1-\epsilon)}{n \log_2 n}\right] \leq \delta.$$

□

Assim para um baralho de tamanho $n = 52$, temos que $t_{mix} \geq 6$.

A cota superior e a cota inferior para o t_{mix} mostram que o número de embaralhamentos canônicos necessários para que a distribuição induzida por sucessivos embaralhamentos canônicos esteja próxima da distribuição uniforme está no conjunto $\{6, 7, \cdots 12\}$.

Apêndice

No artigo [1] os autores apresentam uma estratégia diferente para fazer a aproximação da distribuição uniforme. A partir do Corolário 2 e da Equação 2.7, a terceira caracterização de Distância de Variação Total, eles encontram

$$d(t) = \sum_{n \in X} \left[\binom{2^t + n - r}{n} 2^{-tn} - \frac{1}{n!} \right] \alpha_r,$$

em que α_r é o número de permutações de n cartas com r sequências levantadoras e $X = \left\{ n \in \mathbb{N}, \binom{2^t+n-r}{n} 2^{-tn} > \frac{1}{n!} \right\}$. Com esta fórmula é possível calcular o valor exato para a Distância de Variação Total. A tabela a seguir, extraída de [1], apresenta este valor para um baralho de tamanho $n = 52$, considerando t embaralhamentos, com $1 \leq t \leq 10$.

t	1	2	3	4	5	6	7	8	9	10
d(t)	1,000	1,000	1,000	1,000	0,924	0,614	0,334	0,167	0,085	0,043

Tabela 2.2: Distância de Variação Total para um baralho de tamanho 52

O famoso resultado de que com 7 embaralhamentos canônicos um baralho está bem embaralhado vem do fato de se escolher $t_{mix} = \min \{t, d(t) \leq 0,334\}$. Então, pela Tabela 2.2 temos que $t_{mix} = 7$.

Referências Bibliográficas

[1] Bayer, D.; Diaconis, P., *Trailing the Dovetail Shuffle to its Lair*, The Annals of Applied Probability , Vol.2, No. 2, 294-313, 1992.

[2] Diaconis, P.;Graham, R., *Magical Mathematics: The mathematical ideas that animate great magical tricks*, Princeton University Press, 2012.

[3] Gardner, M., *Mathematics, magic and mistery*, Dover Publications, INC., New York, 1956.

[4] Hilário, M. R.; Oliveira, R. I., *A Matemática de Embaralhar Cartas*, notas de minicursos da VI Bienal da Sociedade Brasileira de Matemática, 2012.

[5] Jonsson, G. F; Trefethen, L. N., *A numerical analyst looks at the cutoff phenomenon in card shuffling and other Markov Chains*, Numerical Analysis 1997, 150-178, Longman, 1998.

[6] Levin, D. A.; Peres, Y.; Wilmer, E. L.,*Markov Chains and Mixing Times*, American Mathematical Society, 2008.

[7] Rogers, L. C. G.; Pitman J. W., *Markov Functions*, The Annals of Probability, Vol. 9, pp. 573-582, No. 4, 1981.