# Tempo de espera para a ocorrência de palavras em ensaios de Markov

***Mariele Parteli Florencio***

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura:______________________

# Tempo de espera para a ocorrência de palavras em ensaios de Markov

**Mariele Parteli Florencio**

***Orientador:*** **Prof. Dr. Renato Jacob Gava**

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Mestre em Estatística – Interinstitucional de Pós-Graduação em Estatística. *VERSÃO REVISADA*

**USP/UFSCar – São Carlos**
**Maio de 2016**

Ficha catalográfica elaborada pelo DePT da Biblioteca Comunitária UFSCar
Processamento Técnico
com os dados fornecidos pelo(a) autor(a)

F632t

Florencio, Mariele Parteli
Tempo de espera para a ocorrência de palavras em ensaios de Markov / Mariele Parteli Florencio. -- São Carlos : UFSCar, 2016.
37 p.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2016.

1. Palavras. 2. Tempo de espera. 3. cadeias de Markov. I. Título.

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

## Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado da candidata Mariele Parteli Florêncio, realizada em 06/04/2016:

Prof. Dr. Renato Jacob Gava
UFSCar

Prof. Dr. Manuel Alejandro González Navarrete
USP

Valdivino Vargas Júnior
Prof. Dr. Valdivino Vargas Júnior
UFG

# Waiting time for the occurrence of patterns in Markov chains

**Mariele Parteli Florencio**

*Advisor:* **Prof. Dr. Renato Jacob Gava**

Master dissertation submitted to the *Instituto de Ciências Matemáticas e de Computação* - ICMC-USP and to the *Departamento de Estatística* – DEs-UFSCar, in partial fulfillment of the requirements for the degree of the Master Joint Graduate Program in Statistics DEs-UFSCar/ICMC-USP. *FINAL VERSION*.

**USP/UFSCar – São Carlos**
**May 2016**

# Resumo

Consideremos uma sequência de lançamentos de moedas em que denotamos o resultado de cada lançamento por H, se der cara, ou por T, se der coroa. Formemos uma palavra apenas com H's e T's, por exemplo, *HHHHH* ou *HTHTH*. Quantas vezes arremessaremos uma mesma moeda até que uma das duas palavras acima ocorrerá? Por exemplo, dadas as sequências

*THTHHHHH* e *TTHTTHTHTH*.

O número de vezes que arremessamos a moeda até que *HHHHH* e *HTHTH* ocorreram pela primeira vez é oito e dez, respectivamente. Podemos generalizar a ideia acima para um número finito de palavras em um alfabeto finito qualquer.

Assim, o nosso principal objetivo dessa dissertação é encontrarmos a distribuição do tempo de espera até que um membro de uma coleção finita de palavras seja observado em uma sequência de ensaios de Markov de letras de um alfabeto finito. Mais especificamente, as letras de um alfabeto finito são geradas por uma cadeia de Markov até que uma das palavras de uma coleção finita ocorra.

Além disso encontraremos a probabilidade de que determinada palavra ocorra antes das demais palavras pertencentes a um mesmo conjunto finito. Por último encontraremos a função geradora de probabilidade do tempo de espera.

**Palavras-chave:** palavras, tempo de espera e cadeias de Markov.

# Abstract

Consider a sequence of independent coin flips where we denote the result of any landing for H, if coming up head, or T, otherwise. Create patterns with H's and T's, for example, *HHHHH* or *HTHTH*. How many times do we have to land the same coin until one such two patterns happens? For example, let the sequences being

$$THTHHHHH \text{ and } TTHTTHTHTH.$$

The number of times that we landed the coin until *HHHHH* and *HTHTH* happens it was eight and ten times respectively. We can generalize this idea for a finite number of patterns in any finite set.

Then, the first of all interest of this dissertation is to find the distribution of the waiting time until a member of a finite colection of patterns is observed in a sequence of Markov chains of letters in from finite set. More specifically the letters in a finite set are generated by Markov chain until one of the patterns in any finite set happens.

Besides that, we will find the probability of a pattern happen before of all patterns in the same finite set. Finally we will find the generator function of probability of waiting time.

**Keywords:** patterns, waiting time and Markov chains.

# Sumário


**1 Introdução** **4**

**2 Preliminares** **6**

2.1 Função geradora de probabilidade . . . . . . . . . . . . . . . . . . . . . . . . . 6

2.2 Martingais . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

2.3 Cadeias de Markov . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13

**3 Ocorrência de palavras em cadeias de Markov** **18**

3.1 Valor esperado . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18

3.2 Probabilidade de parada . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31

3.3 Função geradora de probabilidade . . . . . . . . . . . . . . . . . . . . . . . . . 35

**4 Conclusão** **38**

**Referências Bibliográficas** **39**

# Capítulo 1

# Introdução

Nessa dissertação estudamos principalmente o artigo Pozdnyakov[6], cujo o principal objetivo é encontrar a distribuição do tempo de espera até que um elemento de um conjunto finito de palavras é observado em uma sequência de cadeias de Markov de letras de um alfabeto finito.

Por exemplo, considere uma cadeia de Markov com o seguinte espaço de estados $\Omega = \{1, 2\}$ e distribuição inicial igual a $\mathbb{P}(Z = 1) = \mathbb{P}(Z = 2) = 1/2$ e matriz de transição

$$\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix},$$

a matriz de transição. Dado conjunto de palavras $\mathcal{C} = \{11, 12\}$. Na realização 22212 a palavra 12 ocorreu antes de 11 e o tempo de espera é 5.

Observe que $\mathbb{P}(21 \text{ ocorrer antes de } 11) = 1 - 1/4 = 3/4$, pois se a palavra 11 não ocorrer antes de 21 nas duas primeiras realizações não poderá ocorrer mais. O que significa que os valores esperados de ocorrência das palavras não são os mesmos, ainda que tenham o mesmo tamanho. Sem essa observação um leitor desavisado poderia considerar o problema mais simples do que ele realmente é.

Com o objetivo de escrever uma dissertação autocontida começamos com o Capítulo 2, em que expomos resultados importantes para a compreensão dos artigos Pozdnyakov[6] e Gava e Salotti[9] que não são totalmente claros a primeira vista.

Na primeira seção 3.1 do Capítulo 3 é onde se encontram as ideias centrais e os principais resultados, que inclui a ideia de adaptar o método de martingais que é introduzido por Li[7] em 1980 para variáveis aleatórias independentes para cadeias de Markov. Apresentamos essas ideias cuidadosamente de forma que mesmo um aluno de graduação possa entendê-las, não economizamos em exemplos e detalhes para uma melhor compreensão do tema abordado.

Nessa seção se encontra a Proposição 3.1.7, onde demonstramos que dadas as Hipóteses 3.1.1 para o conjunto de palavra $\mathcal{C}$ o tempo de espera $\tau$ é finito, que é um resultado não trivial que não havia sido demonstrado nos artigos que estudamos. Além disso, construímos a matriz de apostas (3.3) para o Exemplo 3.1.8 que não estava completa no artigo Pozdnyakov[6] e era necessária para encontrar o valor esperado de $\tau$. Usando o software R mostramos como encontrar o valor esperado de $\tau$ para o Exemplo 3.1.8 computacionalmente detalhando todos os passos.

Na seção 3.2 estudamos o artigo Gava e Salotti[9] para encontrar a probabilidade de que uma palavra ocorra antes das demais. Como resultado adicional ao artigo expomos o cálculo das probabilidades do Exemplo 3.2.2 detalhadamente utilizando o software Maple.

Na última seção 3.3 retomamos o artigo Pozdnyakov[6] para encontrar a função geradora de probabilidade de $\tau$. No Exemplo 3.3.2 construímos a matriz definida em (3.9) e o vetor definido em (3.10) que são necessárias para encontrar a função geradora de $\tau$ para o Exemplo 3.1.8 e não são apresentadas no artigo Pozdnyakov[6], além de novamente mostrar os cálculos feitos no software Maple.

A distribuição do tempo de espera é a chave para muitos problemas reais em questões em vários campos: controle de qualidade, teste de hipótese, biologia molecular, sequência de DNA entre outras. Assim a ocorrência de palavras tem sido estudadas extensivamente como pode ser observado nas nossas referências bibliográficas e mais profundamente no artigo Pozdnyakov[6].

# Capítulo 2

# Preliminares

Nesse capítulo faremos uma revisão rápida de conceitos básicos que usaremos ao longo do trabalho. Mais especificamente introduziremos os conceitos: função geradora, cadeias de Markov e martingais. Na Seção 2.1 seguiremos o livro Gut[2] e tais conceitos serão usados na seção 3.3. Na seção 2.2 introduziremos o conceito de martingais de maneira formal para tal usaremos o livro Gut[1], para enunciar e demonstrar o Teorema 2.2.9 usaremos o livro Williams[4]. Na seção 2.3 seguiremos o livro Norris[5]. Os conceitos de martingais e cadeias de Markov formam a base para todos os resultados apresentados no Capítulo 3.

## 2.1 Função geradora de probabilidade

Todos os teoremas e demonstrações dessa seção são encontrados na seção 7 do capítulo 4 do livro Gut[2].

**Definição 2.1.1.** *Seja $X$ uma variável aleatória inteira não negativa. A função geradora de probabilidade de $X$ é,*

$$g_X(t) = \mathbb{E}(t^X) = \sum_{n=0}^{\infty} t^n \mathbb{P}(X = n).$$

Observe que a função geradora está bem definida ao menos para $|t| < 1$, já que é uma série de potências com coeficientes em $[0, 1]$. Note também que $g_X(1) = \sum_{n=0}^{\infty} \mathbb{P}(X = n) = 1$. A função geradora é importante para caracterizar uma variável aleatória, este fato pode ser notado nos seguintes resultados.

**Teorema 2.1.2.** *Sejam $X$ e $Y$ variáveis aleatórias inteiras não-negativas. Se $g_x = g_y$, então $p_x = p_y$.*

**Teorema 2.1.3.** *Sejam $X_1, X_2, ..., X_n$ variáveis aleatórias inteiras não negativas e independentes. E seja $S_n = X_1 + X_2 + ... + X_n$, então,*

$$g_{S_n}(t) = \prod_{k=1}^{n} g_{X_k}(t).$$

**Corolário 2.1.4.** *Sob as condições do Teorema 2.1.3, se $X_1, X_2, ..., X_n$ são identicamente distribuídas então,*

$$g_{S_n}(t) = (g_X(t))^n. \tag{2.1}$$

Ao menos para $|t| < 1$ podemos garantir que as derivadas da função geradora existem. Escrevendo a derivada explicitamente obtemos,

$$g_X^{'}(t) = \sum_{n=1}^{\infty} nt^{n-1}\mathbb{P}(X = n)$$

e

$$g_X^{''}(t) = \sum_{n=2}^{\infty} n(n-1)t^{n-2}\mathbb{P}(X = n).$$

E generalizando para $k = 1, 2, ...,$

$$g_X^{(k)}(t) = \sum_{n=k}^{\infty} n(n-1)...(n-k+1)t^{n-k}\mathbb{P}(X = n).$$

Colocando $t = 0$ obtemos $g_X^{(n)}(0) = n!\mathbb{P}(X = n)$, isto é,

$$\mathbb{P}(X = n) = \frac{g_X^{(n)}(0)}{n!}.$$

Observe que se tivermos a função geradora podemos encontrar a distribuição de probabilidade, por isso que esta função é denominada geradora de probabilidade.

Devemos tomar cuidado pois nem sempre as derivadas estão bem definidas para $t \nearrow 1$. Feita essa ressalva podemos enunciar o teorema abaixo.

**Teorema 2.1.5.** *Seja $X$ uma variável aleatória inteira não negativa, e suponha que $\mathbb{E}|X|^k < \infty$ para algum $k = 1, 2, ...$ então,*

$$\mathbb{E}(X(X-1)...(X-k+1)) = g_X^{(k)}(1).$$

A derivada em $t = 1$ pode ser interpretada como o limite quando $t \nearrow 1$. Entretanto por simplicidade escrevemos apenas $g_X^{(k)}(1)$.

**Exemplo 2.1.6.** *Suponha que para $k = 1, 2, 3, ...$ $X$ tenha função de probabilidade dada por*

$$p(k) = \frac{C}{k^2}.$$

*Podemos encontrar um valor para $C$ fazendo,*

$$\sum_{k=1}^{\infty} p(k) = \sum_{k=1}^{\infty} \frac{C}{k^2} \Leftrightarrow C^{-1} = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

*A divergência da série harmônica nos diz que a distribuição não tem média finita,*

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} k\mathbb{P}(X = x) = \sum_{k=1}^{\infty} k\frac{C}{k^2} = \infty.$$

*Podemos encontrar a função geradora para* $|t| \leq 1$ *que é dada por,*

$$g_X(t) = \mathbb{E}(t^k) = \sum_{k=1}^{\infty} t^k\mathbb{P}(X = x) = \sum_{k=1}^{\infty} t^k\frac{C}{k^2} = C\sum_{k=1}^{\infty} \frac{t^k}{k^2} = \frac{6}{\pi^2}\sum_{k=1}^{\infty} \frac{t^k}{k^2}.$$

Derivando em relação a $t$ e fazendo $t \nearrow 1$ obtemos,

$$g_X'(t) = \frac{6}{\pi^2}\sum_{k=1}^{\infty} \frac{t^{k-1}}{k} = -\frac{6}{\pi^2}\frac{log(1-t)}{t} \to +\infty.$$

Isso mostra que embora a função geradora exista para $t = 1$ a sua derivada existe apenas para valores extritamente menor do que 1, mas não para o limite de $t \nearrow 1$.

**Corolário 2.1.7.** *Seja* $X$ *como no Teorema 2.1.5 então:*

1. $\mathbb{E}(|X|) < \infty \Rightarrow \mathbb{E}(X) = g_X'(1)$;
2. $\mathbb{E}(X^2) < \infty \Rightarrow \mathbb{V}ar(X) = g_X''(1) + g_X'(1) - (g_X'(1))^2$.

## 2.2 Martingais

Vamos introduzir o conceito de martingal de maneira formal. O ponto de partida é o espaço de probabilidade tradicional $(\Delta, \mathcal{F}, \mathbb{P})$, adicionando a sequência $\{\mathcal{F}\}_{n\geq 0}$ de sub-$\sigma$-álgebras de $\mathcal{F}$, que chamamos de filtração, isto é

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \ldots \subset \mathcal{F}.$$

Se $n$ é interpretado como o tempo (discreto), então $\mathcal{F}_n$ contém toda a informação até o tempo $n$. Também introduzimos $\mathcal{F}_\infty = \sigma\{\cup_n \mathcal{F}_n\} \subset \mathcal{F}$, que é o conjunto das $\sigma$-álgebras geradas por $\cup_n \mathcal{F}_n$. Note que a união de uma sequência de $\sigma$-álgebras não é necessariamente uma $\sigma$-álgebra.

**Definição 2.2.1.** *A sequência* $\{X_n\}_{n\geq 0}$ *de variáveis aleatórias é* $\mathcal{F}_n$*-adaptada se* $X_n$ *é* $\mathcal{F}_n$*-mensurável para todo* $n$*. Se* $\mathcal{F}_n = \sigma\{X_0, ..., X_n\}$*, chamamos de sequência adaptada e chamamos* $\{\mathcal{F}_n\}_{n\geq 0}$ *a sequência de* $\sigma$*-álgebras naturais ou de filtração natural.*

Observe que se $X_n$ é adaptado, o valor $X_n$ é conhecido no tempo n.

**Definição 2.2.2.** *Uma sequência* $\{X_n\}_{n\geq 0}$ *integrável* ($\mathbb{E}(|X_n|) < \infty, \forall n$) *e* $\mathcal{F}_n$*-adaptada é chamada de martingal se*

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n \text{ quase certamente ou q.c. } \forall n \geq 0.$$

*É chamada de submartingal se*

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geq X_n \text{ q.c. } \forall n \geq 0.$$

*É chamada de supermartingal se*

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n \text{ q.c. } \forall n \geq 0.$$

*Chamamos* $\{X_n\}$ *de* $L^P$*-martingal se* $\mathbb{E}|X_n|^P < \infty$ *para todo* $n$. *E* $\{X_n\}$ *de* $L^P$*-limitado se além disso,* $sup_n \mathbb{E}|X_n|^P < \infty$.

**Exemplo 2.2.3.** *Seja* $\{Y_n\}_{n\geq 1}$ *uma sequência de variáveis aleatórias não-negativas e independentes com* $\mathbb{E}(Y_n) = 1$ *para todo* $n \geq 1$. *Defina* $Y_0 = 1$ *e*

$$X_n = Y_1 Y_2 \ldots Y_n.$$

*Então, para* $n \geq 0$

$$\begin{aligned}
\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \mathbb{E}(Y_1 \ldots Y_n Y_{n+1}|\mathcal{F}_n) \\
&= \mathbb{E}(Y_1 \ldots Y_n Y_{n+1}|\mathcal{F}_n) \\
&= Y_1 Y_2 \ldots Y_n \, \mathbb{E}(Y_{n+1}) \\
&= X_n,
\end{aligned}$$

*ou seja, a sequência* $\{X_n\}_{n\geq 1}$ *é martingal.*

**Exemplo 2.2.4.** *Seja* $\{Y_n\}_{n\geq 1}$ *uma sequência de variáveis aleatórias i.i.d. tal que* $\mathbb{P}(Y_n = 1) = p$ *e* $\mathbb{P}(Y_n = -1) = 1 - p = q$. *Considere* $S_0 = 0$ *e, para* $n \geq 1$, *defina*

$$S_n = \sum_{i=1}^{n} Y_n.$$

*A sequência* $\{S_n\}_{n\geq 0}$ *é chamada passeio aleatório simples de parâmetro* $p$. *Podemos imaginar que* $S_n$ *descreve a posição de uma partícula no tempo* $n$: *em* $n = 0$ *ela está posicionada na origem e a cada instante de tempo* $n \geq 1$, *ela dá um passo à direita de tamanho 1 com probabilidade* $p$ *ou dá um passo à esquerda de tamanho 1 com probabilidade* $q$. *A palavra simples refere-se ao tamanho unitário do salto da partícula a cada instante de tempo.*

*Observe que* $\mathbb{E}(S_n) = n(2p-1)$ *e que* $S_{n+1} = S_n + Y_{n+1}$. *Defina* $X_n = S_n - n(2p-1)$ *para* $n \geq 0$. *Mostremos que* $\{X_n\}_{n\geq 0}$ *é martingal*

$$\begin{aligned}
\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \mathbb{E}(S_{n+1} - (n+1)(2p-1)|\mathcal{F}_n) \\
&= \mathbb{E}(S_n + Y_{n+1} - (n+1)(2p-1)|\mathcal{F}_n) \\
&= \mathbb{E}(S_n - n(2p-1)|\mathcal{F}_n) + \mathbb{E}(Y_{n+1}|\mathcal{F}_n) - (2p-1) \\
&= X_n.
\end{aligned}$$

**Proposição 2.2.5.** *Seja* $\{X_n\}_{n\geq 0}$ *um martingal. Então para todo* $n \geq 0$, $\{-X_n\}_{n\geq 0}$ *também é um martingal.*

**Prova:**

Primeiro observe que $\mathbb{E}(|-X_n|) = \mathbb{E}(|X_n|) < \infty$, para todo $n \geq 0$. E por último,

$$\mathbb{E}(-X_{n+1}|\mathcal{F}_n) = -\mathbb{E}(X_{n+1}|\mathcal{F}_n) = -X_n$$

para todo $n \geq 0$.

□

**Definição 2.2.6.** *Uma função* $\tau : \Omega \to \{0, 1, 2, \ldots\}$ *é chamada de tempo de parada se,*

1. $\{\tau \leq n\} = \{w : \tau(w) \leq n\} \in \mathcal{F}_n$, $\forall n \leq \infty$

   *equivalentemente;*

2. $\{\tau = n\} = \{w : \tau(w) = n\} \in \mathcal{F}_n$, $\forall n \leq \infty$.

*Note que* $\tau$ *pode ser infinito.*

Iremos demonstrar a equivalência de (1) e (2). Se $\tau$ tem a propriedade (1), então

$$\{\tau = n\} = \{\tau \leq n\} - \{\tau \leq n-1\} \in \mathcal{F}_n.$$

Se $\tau$ tem a propriedade (2), então para $k \leq n$, $\{\tau = k\} \in \mathcal{F}_k \subset \mathcal{F}_n$ e

$$\{\tau \leq n\} = \cup_{0\leq k\leq n} \{\tau = k\} \in \mathcal{F}_n.$$

Intuitivamente, uma variável aleatória inteira não-negativa $\tau$ é um tempo de parada para a sequência $\{X_n\}_{n\geq 0}$ se o evento $\{\tau \leq n\}$ depende apenas das variáveis aleatórias $X_1, ..., X_n$, isto é, se conhecemos a realização das variáveis aleatórias $X_1, ..., X_n$, então saberemos a ocorrência ou não de $\{\tau \leq n\}$.

**Exemplo 2.2.7.** *Considere uma variável aleatória discreta* $Z$ *que assume valores em um alfabeto finito* $\Omega$. *Seja* $\{Z_n\}_{n\geq 1}$ *uma sequência de variáveis aleatórias independentes e identicamente distribuídas com a mesma distribuição de* $Z$. *Chamamos de palavra qualquer sequência finita de elementos de* $\Omega$. *Considere a palavra* $A_1 = l_1...l_m$, *denotamos por* $\tau_{A_1}$ *o tempo de espera até que* $A_1$ *surja pela primeira vez na sequência* $\{Z_n\}_{n\geq 1}$, *ou seja,*

$$\tau_{A_1} = \min \{n \geq 1; Z_n = l_m, Z_{n-1} = l_{m-1}, ..., Z_{n-m+1} = l_1\}.$$

*Em seguida, considere uma coleção finita* $\mathcal{C} = \{A_1, \ldots, A_k\}$ *de palavras e defina*

$$\tau = \min \{\tau_{A_1}, \ldots, \tau_{A_k}\}.$$

$\tau$ *é o tempo de espera até que qualquer uma das* $k$ *palavras pertencentes a* $\mathcal{C}$ *ocorram. Então* $\tau_{A_1}$ *e* $\tau$ *são tempos de parada, já que a ocorrência ou não de uma das* $k$ *palavras* $A_1, \ldots, A_k$ *até o tempo* $n$ *depende apenas dos* $n$ *primeiros ensaios.*

**Exemplo 2.2.8.** *Seja* $\{Y_n\}_{n\geq 1}$ *uma sequência de variáveis aleatórias i.i.d. tal que* $\mathbb{P}(Y_n = 1) = 1/2 = \mathbb{P}(Y_n = -1)$. *Considere* $S_0 = 0$ *e para* $n \geq 1$ *defina*

$$S_n = \sum_{i=1}^{n} Y_n.$$

*Já vimos que* $S_n$ *é um passeio aleatório simples, quando* $p = 1/2$, *dizemos que o passeio é simétrico. Além disso,* $S_n$ *é martingal.*

*Agora considere* $\max\{\phi\} = 0$ *e defina*

$$\delta = \max \{1 \leq n \leq 15; S_n = 1\}.$$

*A variável* $\delta$ *nos dá a última visita ao vértice 1 para* $n \in \{0, 1, \ldots, 15\}$. *Observe que o evento* $\{\delta \leq 5\}$ *não depende apenas dos 5 primeiros ensaios. De modo geral,* $\{\delta \leq i\}$ *não depende apenas dos i primeiros ensaios. Portanto* $\delta$ *não é tempo de parada.*

Agora iremos enunciar o Teorema da parada ótima de Doob da seção 10.10 do livro Willians[4].

**Teorema 2.2.9.** ***Teorema da parada ótima de Doob***

*Sejam* $\tau$ *um tempo de parada e* $\{X_n\}_{n\geq 0}$ *um supermartingal. Então* $X_\tau$ *é integrável e*

$$\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$$

*em cada uma das seguintes situações:*

1. $\tau$ *é limitado (para algum* $N \in \mathbb{N}$, $\tau(w) \leq N$, $\forall w$*);*

2. $X$ *é limitada (para algum* $K$ *em* $\mathbb{R}^+$, $|X_n(w)| \leq K$ *para todo* $n$ *e todo* $w$*) e* $\tau$ *é quase certamente finito;*

3. $\mathbb{E}(\tau) < \infty$, *e para* $K$ *em* $\mathbb{R}^+$

$$|X_n(w) - X_{n-1}(w)| \leq K \ \forall (n, w).$$

*Se alguma das condições (1)-(3) são válidas e $\{X_n\}_{n\geq 0}$ é um martingal então,*

$$\mathbb{E}(X_\tau) = \mathbb{E}(X_0).$$

Para demonstrar o Teorema acima precisamos enunciar o clássico Teorema da convergência dominada ver o Teorema na seção 5.9 do livro Willians[4],

**Teorema 2.2.10.** ***Teorema da convergência dominada***
*Seja $\{X_n\}_{n\geq 0}$ uma sequência de variáveis aleatórias e $X$ uma variável aleatória. Suponha que $X_n \to X$ quase certamente, ou seja, para $n \to \infty$*

$$\mathbb{P}(X_n \to X) = 1.$$

*Se $|X_n(w)| \leq Y(w)$, $\forall (n, w)$, em que $\mathbb{E}(Y) < \infty$ então*

$$\mathbb{E}(|X_n - X|) \to 0.$$

*De modo que,*

$$\mathbb{E}(X_n) \to \mathbb{E}(X).$$

A demonstração do Teorema da convergência dominada pode ser encontrado no livro Williams [4] na página 54. Enunciado o Teorema acima, temos todas a ferramentas para demonstrar o 2.2.9.

**Prova: Teorema da parada ótima de Doob,**
Seja $\tau \wedge n = \min\{\tau, n\}$, como $\{X_n\}_{n\geq 0}$ é um supermartingal e $0 \leq \tau \wedge n \leq n$ então,

$$\mathbb{E}(X_{\tau\wedge n}) \leq \mathbb{E}(X_0) < \infty$$

portanto $X_{\tau\wedge n}$ é integrável.

Note que equivalentemente podemos escrever

$$\mathbb{E}(X_{\tau\wedge n} - X_\tau) \leq 0.$$

1. Como $\tau$ é limitado, então $0 \leq \tau \wedge n \leq \tau$

$$\mathbb{E}(X_\tau) \leq \mathbb{E}(X_{\tau\wedge\tau}) \leq \mathbb{E}(X_0).$$

2. Tomando $Y_n = X_{\tau\wedge n} - X_0$

$$\mathbb{P}(\lim_{n\to\infty} Y_n = X_\tau - X_0) = 1.$$

   Usando a hipótese (2), como $Y_n = X_{\tau\wedge n} - X_0$ então $|Y_n| \leq K$ e $\mathbb{E}(Y_n) \leq 0$, $\forall n \geq 0$. Observe que $\lim_{n\to\infty}(X_{\tau\wedge n} - X_0) = X_\tau - X_0$, pois $n \to \infty$ e pela hipótese (2) $\tau$ é quase certamente finito. Assim aplicando o Teorema da convergência dominada,

$$\lim_{n\to\infty} \mathbb{E}(Y_n) = \mathbb{E}(\lim_{n\to\infty} Y_n) = \mathbb{E}(X_\tau - X_0).$$

Mas $\mathbb{E}(Y_n) \leq 0$, $\forall n \geq 0$ então

$$\mathbb{E}(X_\tau - X_0) \leq 0.$$

E portanto,

$$\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0).$$

3. Tomando novamente $Y_n = X_{\tau\wedge n} - X_0$ e usando a hipótese (3),

$$|Y_n| = |X_\tau - X_0| = |\sum_{k=1}^{\tau\wedge n}(X_k - X_{k-1})| \leq K\tau,$$

como $\mathbb{E}(K\tau) < \infty$, podemos usar novamente o Teorema da convergência dominada e proceder analogamente ao que fizemos para o caso (2).

E por último, suponhamos que $\{X_n\}_{n\geq 0}$ é martingal então $\{-X_n\}_{n\geq 0}$ também é, pelo resultado que acabamos de demonstrar temos que,

$$\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$$

e

$$\mathbb{E}(-X_\tau) \leq \mathbb{E}(-X_0) \Leftrightarrow \mathbb{E}(X_\tau) \geq \mathbb{E}(X_0).$$

O que implica que

$$\mathbb{E}(X_\tau) = \mathbb{E}(X_0).$$

□

## 2.3 Cadeias de Markov

**Definição 2.3.1.** *Um processo estocástico a tempo discreto é uma sequência de variáveis aleatórias $\{Z_n\}_{n\geq 0}$, $n \in \mathbb{N}$, definidas no mesmo espaço de probabilidade e com valores em algum conjunto enumerável $\Omega$ chamado de conjunto de espaço de estados.*

Denomina-se "tempo discreto" devido o subíndice $n$ representar unidades de tempo e $\{Z_n\}_{n\geq 0}$ representar a evolução de um certo fenômeno ao longo desse tempo.

**Exemplo 2.3.2.** *Seja* $\{Z_n\}_{n\geq 0}$ *uma sequência de variáveis aleatórias i.i.d então* $\{Z_n\}_{n\geq 0}$ *é um processo estocástico.*

As cadeias de Markov se caracterizam pela falta de memória, isto é, o comportamento futuro do processo estocástico não depende do passado, apenas do presente. Para definir o conceito formalmente precisamos de algumas considerações:

1. seja $\Omega$ um conjunto enumerável;
2. cada $i \in \Omega$ é chamado de estado;
3. $\lambda = (\lambda_i, i \in \Omega)$ é uma medida de probabilidade, ou seja, $\exists$ uma variável aleatória $Z$ tal que $\lambda_i = \mathbb{P}(Z = i)$ com $\lambda_i \geq 0, \forall i \sum_{i\in\Omega} \lambda_i = 1$;
4. uma matriz $P = (P_{ij}; i, j \in \Omega)$ tal que $P_{ij} \geq 0$ para todo $i, j \in \Omega$ e $\sum_j P_{ij} = 1$ para todo $i \in \Omega$ $P$ é chamada matriz de transição.

**Definição 2.3.3.** *Dizemos que* $\{Z_n\}_{n\geq 1}$ *é uma cadeia de Markov com distribuição inicial* $\lambda$ *e matriz de transição* $P$ *se:*

1. $\mathbb{P}(Z_1 = i_1) = \lambda_{i1}$;
2. $\mathbb{P}(Z_{n+1} = i_{n+1}|Z_n = i_n, ..., Z_0 = i_0) = \mathbb{P}(Z_{n+1} = i_{n+1}|Z_n = i_n) = p_{inin+1}$ *para todo* $n \in \{0, 1, 2, ...\}$ *e* $i_0, ..., i_{n+1} \in \Omega$.

*Se* $\{Z_n\}_{n\geq 0}$ *é uma sequência de variáveis aleatórias satisfazendo (1) e (2) para* $n = 0, 1, ....$ *E usaremos a seguinte notação:* $Markov(\lambda, P)$

**Exemplo 2.3.4.** *Voltando ao Exemplo 2.2.3 colocando* $\{Y_n\}_{n\geq 0} = \{Z_n\}_{n\geq 0}$, *vamos mostrar que o passeio aleatório é uma cadeia de Markov e mais duas propriedades:*

1. $\mathbb{P}(S_n = j|S_0 = a) = \mathbb{P}(S_n = j + b|S_0 = a + b)$;
2. $\mathbb{P}(S_n = j|S_0 = a) = \mathbb{P}(S_{n+m} = j|S_m = a)$;
3. $\{S_n\}$ *é Markov* $(\lambda_a, P)$ *em que* $\Omega = \mathbb{Z}$ *e*

$$\lambda_a = \mathbb{P}(S_0 = a)$$

   *em que* $P = (p_{ij}, i, j \in \mathbb{Z})$ *é tal que*

$$p_{ii+1} = p$$

$$p_{ii-1} = 1 - p$$

$$p_{ij} = 0, \ \ se \ \ j \neq i + 1, i - 1.$$

**Prova:**

*1.*

$$\begin{aligned}\mathbb{P}(S_n = j|S_0 = a) &= \mathbb{P}\left(\sum_{i=1}^{n} Z_i = j - a\right)\\ &= \mathbb{P}\left(\sum_{i=1}^{n} Z_i = (j+b) - (a+b)\right)\\ &= \mathbb{P}\left(\sum_{i=1}^{n} Z_i = (j+b)|S_0 = (a+b)\right);\end{aligned}$$

*2.*

$$\begin{aligned}\mathbb{P}(S_{n+m} = j|S_m = a) &= \mathbb{P}\left(\sum_{i=m+1}^{n+m} Z_i = j - a\right)\\ &= \mathbb{P}\left(\sum_{i=1}^{n} Z_i = j - a\right)\\ &= \mathbb{P}\left(S_n = j|S_0 = a\right);\end{aligned}$$

*3.*

$$\begin{aligned}\mathbb{P}(S_{n+1} = j|S_n = i, ..., S_0 = a) &= \mathbb{P}(S_n + Z_{n+1} = j|S_n = i, ..., S_0 = a)\\ &= \mathbb{P}(Z_{n+1} = j - i)\\ &= \mathbb{P}\left(S_{n+1} = j|S_n = i\right).\end{aligned}$$

□

Iremos enunciar o Teorema 1.1.1 do livro Norris[5].

**Teorema 2.3.5.** $\{Z_n\}_{\geq 0}$ *é uma cadeia de Markov* $(\lambda, P)$ $\Leftrightarrow$ *para todo* $i_1, i_1, ..., i_N \in \Omega$ *e* $N \in \mathbb{N}$,

$$\mathbb{P}(Z_0 = i_0, ..., Z_N = i_N) = \lambda_0 P_{i_0 i_1} ... P_{i_{N-1} i_N}.$$

Antes de demonstrar o teorema precisamos enunciar a seguinte preposição,

**Proposição 2.3.6.** ***(Regra do produto:)*** *Sejam* $A_1, ..., A_n$ *eventos, então*

$$\mathbb{P}(A_n, ..., A_1) = \mathbb{P}(A_n|A_1, ..., A_{n-1})\mathbb{P}(A_{n-1}|A_1, ..., A_{n-2})...\mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$$

**Prova:** É fácil ver que,

$$\mathbb{P}(A_2, A_1) = \mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$$

Suponha por indução que,

$$\mathbb{P}(A_n, ..., A_1) = \mathbb{P}(A_n|A_1, ..., A_{n-1})\mathbb{P}(A_{n-1}|A_1, ..., A_{n-2})...\mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$$

Então para $n + 1$,

$$\begin{aligned}\mathbb{P}(A_{n+1}, ..., A_1) &= \mathbb{P}(A_{n+1}|A_n, ..., A_1)\mathbb{P}(A_n, ..., A_1)\\ &= \mathbb{P}(A_{n+1}|A_n, ..., A_1)\mathbb{P}(A_n|A_{n-1}, ..., A_1)...\mathbb{P}(A_2|A_1)\mathbb{P}(A_1)\end{aligned}$$

□

Agora estamos aptos a demonstrar o Teorema,

**Prova:** Suponha que $\{Z_n\}_{n\geq 0}$ seja Markov $(\lambda, P)$. Então,

$$\begin{aligned}\mathbb{P}(Z_0 = i_0, ..., Z_N = i_N) &= \mathbb{P}(Z_N = i_N|Z_{N-1} = i_{N-1}, ..., Z_0 = i_0)\\ &\times \mathbb{P}(Z_{N-1} = i_{N-1}|Z_{N-2} = i_{N-2}, ..., Z_0 = i_0)\\ &\times \mathbb{P}(Z_1 = i_1|Z_0 = i_0)\mathbb{P}(Z_0 = i_0)\\ &= \lambda_0 P_{i_0 i_1}...P_{i_{N-1} i_N}\end{aligned}$$

Por outro lado, somando em $i_N$, temos que,

$$\mathbb{P}(Z_0 = i_0, ..., Z_{N-1} = i_{N-1}) = \lambda_0 P_{i_0 i_1}...P_{i_{N-2} i_{N-1}}$$

Por indução segue que,

$$\mathbb{P}(Z_0 = i_0, ..., Z_n = i_n) = \lambda_0 P_{i_0 i_1}...P_{i_{n-1} i_n}$$

Em particular $\mathbb{P}(X_0 = i_0) = \lambda_0$, agora observe que,

$$\begin{aligned}\mathbb{P}(Z_{n+1} = i_{n+1}|Z_n = i_n, ..., Z_0 = i_0) &= \frac{\mathbb{P}(Z_{n+1} = i_{n+1}, ..., Z_0 = i_0)}{\mathbb{P}(Z_n = i_n, ..., Z_0 = i_0)}\\ &= \frac{\lambda_0 P_{i_0 i_1}...P_{i_n i_{n+1}}}{\lambda_0 P_{i_0 i_1}...P_{i_{n-1} i_n}}\\ &= P_{i_n i_{n+1}}\end{aligned}$$

O que prova que $\{Z_n\}_{n\geq 0}$ é uma cadeia de Markov $(\lambda, P)$. □

**Definição 2.3.7.** *Uma cadeia de Markov é homogênea se a probabilidade de transição não depende do tempo.*

**Exemplo 2.3.8.** *A cadeia de Markov com os seguintes estados* $\Omega = \{1, 2\}$ *cuja matriz de transição é dada por,*

$$\begin{bmatrix} 1/2 & 1/2 \\ 2/3 & 1/3 \end{bmatrix}$$

*é homogênea, porém a cadeia de Markov com os estados* $\Omega = \{1, 2\}$ *e matriz de transição,*

$$\begin{bmatrix} e^{-n} & 1 - e^{-n} \\ \frac{1}{n} & 1 - \frac{1}{n} \end{bmatrix}$$

*não é homogênea.*

Podemos dividir uma cadeia de Markov partes que facilitem entender a cadeia. Para isso identificamos as classes comunicantes da cadeia. Dizemos que "$i$ vai para $j$´´ e escrevemos $i \to j$ se

$$\mathbb{P}(X_n = j \text{ para algum } n \geq 0 | X_0 = i) > 0$$

Dizemos que $i$ e $j$ são comunicantes se $i \to j$ e $j \to i$. Notação: $i \leftrightarrow j$.
Note que " $\leftrightarrow$ " define uma classe de equivalência $C$:

1. $i \to j$;
2. Se $i \to j$ então $j \to i$;
3. Se $i \to j$ e $j \to k$ então $i \to k$.

**Definição 2.3.9.** *Dizemos que uma classe* $C \subset \Omega$ *é fechada se* $i \in C$ *e* $i \to j$*, então* $j \in C$*. Quando* $\Omega$ *é uma única classe, dizemos que a cadeia é irredutível e quando uma classe fechada contiver apenas um elemento* ($C = \{i\}$)*, dizemos que* $i$ *é um estado absorvente.*

# Capítulo 3

# Ocorrência de palavras em cadeias de Markov

O principal objetivo desse capítulo é o tempo de espera até que um elemento de um conjunto finito de palavras é observado em uma sequência estocástica de letras de um alfabeto finito. Na seção 3.1 apresentaremos detalhadamente nossos estudos sobre o artigo Pozdnyakov[6]. Nesse artigo o autor demonstra que a ocorrência de palavras em cadeias de Markov pode ser tratada com a ajuda de métodos de martingais. Esses métodos foram introduzidos por Li[7] em 1980 para variáveis aleatórias independentes e consistiam em uma construção de times de apostas de forma a obter um jogo justo. O senso comum dizia que tais métodos não eram adequados para a situação de realizações de cadeias de Markov, porém o autor demonstra que feitas as corretas mudanças nos times de apostas conseguimos obter o que desejamos.

Na seção 3.2 calcularemos a probabilidade de que uma palavra ocorra antes das demais, para isso estudaremos o artigo Gava e Salotti[9]. E na última seção 3.3 voltamos ao artigo Pozdnyakov[6], em que encontraremos a função geradora de probabilidades do tempo de espera até que determinada palavra ocorra.

## 3.1 Valor esperado

Iremos começar introduzindo o problema formalmente. Seja $\{Z_n\}_{n\geq 1}$ uma cadeia de Markov finita e homogênea com o espaço de estados $\Omega = \{1, 2, ..., K\}$. Observe que o instante inicial da cadeia de Markov será $n = 1$. Suponhamos que a cadeia tenha distribuição inicial $P(Z_1 = k) = p_k$, $1 \leq k \leq K$ e matriz de transição $P = \{p_{ij}\}$ com $1 \leq i, j \leq K$, tal que $p_{ij} = P(Z_{n+1} = j | Z_n = i)$.

Nesse contexto tomaremos o espaço de estados $\Omega$ como nosso alfabeto. E definiremos como palavra uma sequência finita de elementos de $\Omega$, ou seja, $A = l_1 \dots l_p$ é uma palavra de tamanho $p \in \mathbb{N}$. Em seguida consideraremos o conjunto $\mathcal{C} = \{A_1, ..., A_M\}$ de $M$ palavras possivelmente com tamanho diferentes. Denotaremos por $A_i$ qualquer elemento de $\mathcal{C}$, observe que $1 \leq i \leq M$, porém omitiremos a variação de $i$ quando ela não for necessária para o entendimento do leitor. Seja $\tau_{Ai}$ o tempo de espera até que $A_i$ ocorra na série $Z_1Z_2...$, ou seja,

$$\tau_{A_i} = \min \{n \geq 1; Z_n = l_p, ..., Z_{n-p+1} = l_1\}.$$

Por fim, definiremos a seguinte variável aleatória

$$\tau = \min \{\tau_{A_1}, ..., \tau_{A_M}\}.$$

Observe que $\tau$ é um tempo de parada, pois representa o tempo da primeira vez que qualquer um dos padrões $A_i$ ocorreu e a ocorrência ou não de qualquer uma dessas palavras até o tempo $n$ só depende dos $n$ primeiros ensaios. O nosso principal objetivo será encontrar o valor esperado para $\tau$.

**Hipóteses 3.1.1.** *Começaremos listando algumas hipóteses para o conjunto de palavras $\mathcal{C}$:*

1. *para que não haja possibilidade de empate, assumimos que nenhuma sequência de $\mathcal{C}$ é uma subsequência de outra;*

2. *se $\mathbb{P}(\tau = \tau_{A_i}) = 0$ para algum $i$ então não temos motivos para mantermos esse $A_i$ em $\mathcal{C}$. Portanto nossa segunda hipótese é que $\mathbb{P}(\tau = \tau_{A_i}) > 0$ para todo $1 \leq i \leq M$;*

3. *para excluir a possibilidade de todas as sequências de $\mathcal{C}$ se manter em estados transientes, assumimos que $\mathbb{P}(\tau < \infty) = 1$.*

**Exemplo 3.1.2.** *Não podemos ter ambas as sequências* 1212 *e* 12 *no nosso conjunto $\mathcal{C}$ uma ou ambas devem ser excluídas.*

**Exemplo 3.1.3.** *Note que no caso de ensaios independentes a Hipótese 3.1.1* (2) *é consequência da Hipótese 3.1.1* (1). *No entanto não é suficiente quando temos ensaios de cadeias de Markov. Por exemplo, dado uma cadeia de Markov com os seguintes estados $\Omega = \{1, 2, 3\}$ cuja matriz de transição é dada por*

$$\begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 1/3 & 1/3 & 1/3 \end{bmatrix},$$

*a sequência* 213 *tem probabilidade* 0 *de ocorrer.*

**Exemplo 3.1.4.** *A suposição $\mathbb{P}(\tau < \infty) = 1$ não implica que a cadeia seja irredutível. Tomando uma cadeia com os seguintes estados $\Omega = \{1, 2, 3, 4\}$, probabilidades iniciais $p_1 = p_2 = p_3 = p_4 = 1/4$ e matriz de transição*

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

*Seja $\mathcal{C} = \{121, 34\}$, temos que $\mathbb{P}(\tau < \infty) = 1$ e que todos os estados são recorrentes mas a cadeia não é irredutível.*

**Exemplo 3.1.5.** *Tomando uma cadeia com os seguintes estados* $\Omega = \{1, 2, 3\}$*, probabilidades iniciais* $p_1 = p_2 = p_3 = 1/3$ *e matriz de transição*

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{bmatrix}.$$

*Se tomarmos* $\mathcal{C} = \{11\}$*, ainda que a cadeia não seja irredutível existe apenas uma classe de estados recorrentes, o que permite que* $\tau$ *seja finito.*

Provaremos o Lema da seção 10.11 presente no livro Willians[4].

**Lema 3.1.6.** *Suponha que* $\tau$ *é um tempo de parada tal que para algum* $N \in \mathbb{N}$ *e algum* $\epsilon > 0$*, temos para todo* $n \in \mathbb{N}$

$$\mathbb{P}(\tau < n + N|\mathcal{F}_n) > \epsilon \qquad q.c.$$

*Então* $\mathbb{E}(\tau) < \infty$*.*

**Prova:**
Primeiro queremos provar por indução sobre $k = 1, 2, 3...$ que

$$\mathbb{P}(\tau > kN) \leq (1 - \epsilon)^k.$$

Para isso usaremos

$$\mathbb{P}(\tau > kN) = \mathbb{P}(\tau > kN, \tau > (k-1)N).$$

Então para $k = 1$ temos por hipótese

$$\mathbb{P}(\tau > N) = \mathbb{P}(\tau > 0 + N|\mathcal{F}_0) \leq (1 - \epsilon).$$

Suponhamos por indução que $\mathbb{P}(\tau > kN) \leq (1 - \epsilon)^k$. Queremos mostrar que vale para $k + 1$

$$\begin{aligned} \mathbb{P}(\tau > (k+1)N) &= \mathbb{P}(\tau > kN + N|\tau > kN)\mathbb{P}(\tau > kN) \\ &\leq \mathbb{P}(\tau > kN + N|\tau > kN)(1-\epsilon)^k \\ &\leq (1-\epsilon)(1-\epsilon)^k \\ &\leq (1-\epsilon)^{k+1}, \end{aligned}$$

como queríamos demonstrar. Usando o resultado obtido mostraremos que $\mathbb{E}(\tau) < \infty$. Observe que se $i \in \{jN + 1, j(N + 1), \ldots, (j + 1)N\}$, então $\{\tau > i\} \subset \{\tau > jN\}$. Daí

$$\mathbb{P}(\tau > i) \leq \mathbb{P}(\tau > jN), \qquad \text{para todo } i \in \{jN, j(N+1) \ldots, (j+1)N\}.$$

Assim

$$\begin{aligned}\mathbb{E}(\tau) = \sum_{k=0}^{\infty} \mathbb{P}(\tau > k) &= \sum_{j=0}^{\infty} \sum_{i=jN}^{(j+1)N} \mathbb{P}(\tau > i) \\ &\leq \sum_{j>0}^{\infty} N\mathbb{P}(\tau > jN) \\ &\leq N \sum_{j>0}^{\infty} (1-\epsilon)^j \\ &\leq N\frac{1}{\epsilon} < \infty.\end{aligned}$$

Portanto, $\mathbb{E}(\tau) < \infty$.

□

**Proposição 3.1.7.** *Seja $\mathcal{C}$ um conjunto de palavras definido com respeito as hipóteses 3.1.1. Então para $\mathcal{C}$ temos que $\mathbb{E}(\tau) < \infty$.*

**Prova:** Suponha que $|\mathcal{C}| = 1$ e seja $A = l_1 \dots l_m$ a palavra pertencente a $\mathcal{C}$. Como o conjunto de palavras tem apenas uma palavra a cadeia necessariamente deve ter uma única classe de estados recorrentes para que $\tau$ seja finito q.c. O Exemplo 3.1.5 nos mostra que não podemos ter mais de uma classe de estados recorrentes.

Como o estado $l_1$ é recorrente existe $s = s(X_n)$ tal que $\mathbb{P}^s_{X_n l_1} := P^s(X_n, l_1) > 0$. Tomando $N = m + s$ temos

$$\mathbb{P}(\tau \leq n + N|\mathcal{F}_n) \geq \mathbb{P}^s(X_n, l_1)\mathbb{P}(l_1, l_2)\dots\mathbb{P}(l_{m-1}, l_m).$$

Seja $l$ um estado qualquer da cadeia. Pelo mesmo raciocínio anterior, existe $s_l$ tal que $P^{s_l}(l, l_1) > 0$. Defina

$$\underline{p} = \min\left\{\mathbb{P}^{s_l}(l, l_1), l \in \Omega\right\}.$$

Note que $\underline{p} > 0$. Assim,

$$\mathbb{P}(\tau \leq n + N|\mathcal{F}_n) \geq \underline{p}\mathbb{P}(l_1, l_2)\dots\mathbb{P}(l_{m-1}, l_m) := \epsilon$$

Pelo Lema 3.1.6, concluímos que $\mathbb{E}(\tau) < \infty$.

Suponha agora que $|\mathcal{C}| = t$, $t \in \mathbb{N}$. Dado que $X_n = a$ existe ao menos uma palavra $A_i = l_1^{(i)} \dots l_{m_i}^{(i)}$ tal que $\mathbb{P}^{s_i}(a, l_1^{(i)}) > 0$. Defina

$$\begin{aligned}\mathcal{L}_a &= \left\{i\,;\ \text{existe}\ \ s_a^{(i)}\ \text{tal que}\ \mathbb{P}^{s_a^{(i)}}(a, l_1^{(i)}) > 0\right\}, \\ \mathcal{N}_a &= \max\left\{s_a^{(i)}\,;\ \mathbb{P}^{s_a^{(i)}}\left(a, l_1^{(1)}\right) > 0\right\}, \\ \mathcal{P}_a &= \min\left\{\mathbb{P}^{s_a^{(i)}}\left(a, l_1^{(i)}\right)\mathbb{P}\left(l_1^{(i)}, l_2^{(i)}\right), \dots, \mathbb{P}\left(l_{m-1}^{(i)}, l_m^{(i)}\right), i \in \mathcal{L}_a\right\} > 0, \\ \widehat{N} &= \max\left\{\mathcal{N}_a\,;\ a \in \Omega\right\}\ \ \text{e}\ \ \widehat{m} = \max\left\{m_1, \dots, m_t\right\}.\end{aligned}$$

Defina $N = \widehat{N} + \widehat{m}$. Então

$$\mathbb{P}(\tau \leq n + N|\mathcal{F}_n) \geq \min\{\mathcal{P}_a \, ; \, a \in \Omega\} := \epsilon > 0$$

Usando novamente o Lema 3.1.6, concluímos que $\mathbb{E}(\tau) < \infty$. □

É importante salientar que esse resultado é apenas mencionado no artigo Pozdnyakov[6] e que também não havia sido demonstrado nos outros artigos que estudamos. Agora separaremos os tipos de ocorrências das sequências em três grupos da seguinte forma:

1. a palavra $A_i$ ocorre como uma estrutura inicial da sequência $\{Z_n\}_{n\geq 1}$;
2. a palavra $kA_i$, para $(1 \leq k \leq K)$, ocorre como uma estrutura inicial da sequência $\{Z_n\}_{n\geq 1}$;
3. a palavra $kmA_i$, para $(1 \leq k, m \leq K)$, ocorre em algum momento na sequência $\{Z_n\}_{n\geq 1}$.

Por exemplo, considere o alfabeto $\Omega = \{1, 2, 3\}$ e a palavra 132. Então nas seguintes realizações: 132, 1132, 12131132; as estruturas 132, 1132 e 31132 correspondem aos grupos 1, 2 e 3 respectivamente. Observe que na realização 12131132, a estrutura é 31132 e ela ocorreu em algum momento na sequência $\{Z_n\}_{n\geq 1}$, por isso pertence ao terceiro grupo.

Além disso, apenas o 132 pertence ao primeiro grupo, para o segundo grupo temos as estruturas: 1132, 2132 e 3132. E por fim, no último grupo estão as estruturas,

11132, 12132, 13132, 21132, 22132, 23132, 31132, 32132 e 33132.

Dados os exemplos acima generalizaremos para uma palavra qualquer. Lembre-se que o alfabeto tem no total $K$ elementos. Para o primeiro grupo existe apenas 1 estrutura, para o segundo grupo existem $K$ estruturas e para o último grupo temos $K^2$ estruturas.

Ou seja, somando todas as opções temos $1 + K + K^2$ estruturas para cada palavra. Como temos $M$ palavras em $\mathcal{C}$, existem no total $(1 + K + K^2)M$ palavras em todos os grupos. As primeiras $1 + K$ estruturas chamaremos de estruturas iniciais e as últimas $K^2$ estruturas serão as estruturas finais.

Para cada estrutura final $kmA_i$ associaremos um time de apostas. Imagine que o cassino realize uma cadeia de Markov $\{Z_n\}_{n\geq 1}$. Então o jogador que entra no tempo $n + 1$ observa o que ocorreu no tempo $n$. Se no tempo $n$ $Z_n = k$, ele aposta uma certa quantia de dinheiro na estrutura $mA_i$. Mas se $Z_n \neq k$ ele aposta em $A_i$. A quantia de dinheiro apostada deve ser a mesma para todos os jogadores pertencentes ao mesmo time de aposta.

Apostar R\$ 1,00 na estrutura $Q = q_1 q_2 ... q_{l'}$, significa que após o jogador observar $Z_n = q_1$, aposta R\$ 1,00 em $q_2$. Se $Z_{n+1} = q_2$ ele ganha $1/(p_{q_1 q_2})$, mas se $Z_{n+1} \neq q_2$ ele perde e deixa do jogo. Suponha que $Z_{n+1} = q_2$, então o jogador deve obrigatoriamente apostar a quantia $1/(p_{q_1 q_2})$ em $q_3$. Se $Z_{n+2} = q_3$ ele ganha $1/(p_{q_1 q_2} p_{q_2 q_3})$, caso contrário ele perde e deixa do jogo.

Em cada rodada surge um novo jogador para cada estrutura final e o jogo prossegue até que uma das sequências de $\mathcal{C}$ finalmente ocorra. Isso significa, tomando novamente o exemplo acima, que no tempo $n+2$ entra um novo jogador, ele então observa se $Z_{n+1} = k$, se sim ele aposta em $mA_i$, caso contrário ele aposta em $A_i$. A soma dos ganhos de cada jogador, que aposta em uma mesma estrutura final $kmA_i$, é o que chamamos de ganho do $i$-ésimo time de aposta.

Nem toda estrutura (final ou inicial) pode ocorrer antes de $\tau$ e na verdade muitas delas não ocorrem durante toda a cadeia de Markov. Por exemplo, dada uma cadeia de Markov com os seguintes estados $\Omega = \{1, 2, 3\}$ cuja matriz de transição é dada por

$$\begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 0 & 1 & 0 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}. \tag{3.1}$$

Embora a sequência 133 seja possível a estrutura final 12133 não ocorre durante toda a cadeia de Markov. Assim da nossa lista de estruturas finais precisamos eliminar casos desse tipo.

**Exemplo 3.1.8.** *Seja $\Omega = \{1, 2, 3\}$ e $\mathcal{C} = \{323, 313, 33\}$. Suponha que*

$$p_1 = 1/3,\ p_2 = 1/3 \text{ e } p_3 = 1/3.$$

*A matriz de transição é dada por*

$$\begin{bmatrix} 3/4 & 0 & 1/4 \\ 0 & 3/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}. \tag{3.2}$$

*As estruturas iniciais são:*

$$323,\ 313,\ 33,\ 1323,\ 2323,\ 1313,\ 2313,\ 133,\ 233.$$

*Note que excluímos:* 3323, 3313 *e* 333. *Pois essas estruturas não podem ocorrer, já que o* 33 *teria aparecido antes que elas fossem completadas. Analogamente excluímos:* 33323, 33313,3333, *das estruturas finais. Também excluímos* 12323, 21323, 12313, 21313, 1233 *e* 2133, *pois as transições* $1 \to 2$ *e* $2 \to 1$ *não são possíveis. Restam assim as seguintes estruturas finais:*

$$11323,\ 22323,\ 11313,\ 22313,\ 1133,\ 2233.$$

*Isto significa que teremos 6 times de apostas. A seguir iremos construir uma tabela resumindo as possíveis realizações que põem fim ao jogo, e os ganhos associados para cada time.*

Observe que como os jogadores são obrigados a prosseguir suas apostas enquanto estão ganhando, antes que uma das palavras de $\mathcal{C}$ seja sorteada não existe a possibilidade de que nenhum jogador ganhe. Por exemplo, tome a seguinte realização 11111323. Até o tempo 4 não existe a possibilidade de nenhum jogador ganhar e mesmo o jogador que entra no tempo 4 só irá ganhar se a estrutura final que ele estiver apostando for a mesma que é sorteada. Por isso, podemos desconsiderar toda a realização anterior à uma das estruturas sorteadas.

Nessas condições, apenas no sentido de explicar com maior clareza como a tabela será construída, tomaremos informalmente um tempo $m$ que começa a ser contado a partir do momento que a primeira letra da estrutura é sorteada. No caso acima nosso $m$ começará a ser contado em $n = 4$, ou seja, quando $n = 4$, $m = 1$.

Para exemplificar como essa tabela é construída, tomaremos dois casos em que apesar das regras do jogo de apostas permanecerem as mesmas, devemos enxergá-las distintamente. No primeiro caso a estrutura sorteada é menor ou igual a estrutura que está sendo apostada. Por exemplo, suponhamos que a palavra sorteada a partir do instante $m = 1$ tenha sido 1133 e tomemos o time de apostas 11323 para calcular seus ganhos. Vamos analisar jogador por jogador:

1. o jogador que entra no tempo $m = 1$ observa que é sorteada a letra 1, então ele obrigatoriamente aposta em 1. Como no tempo $m = 2$ a letra sorteada é 1, ele ganha e continua no jogo. Então, ele é obrigado a apostar em 3, como em $m = 3$ a letra sorteada é 3, ele ganha e prossegue. Apostando em 2, como no tempo $m = 4$ a letra sorteada é 3, ele perde e deixa do jogo;

   Note que nesse caso, não chegaremos a comparar a última letra da estrutura sorteada com a última letra da estrutura apostada, pois a realização acaba antes:

2. o jogador que entra no tempo $m = 2$ observa que é sorteada a letra 1, então ele obrigatoriamente aposta em 1, como no tempo $m = 3$ a letra sorteada é 3, ele perde e deixa do jogo;

3. O jogador que entra no tempo $m = 3$ observa que é sorteada a letra 3, então ele obrigatoriamente aposta em 3 (dado que o jogador aposta na palavra 323), como no tempo $m = 4$ a letra sorteada coincide com a letra apostada, ele ganha $1/p_{33} = 2$.

Agora iremos inverter, a palavra sorteada é 11323 e tomaremos o time de apostas 1133. Note que não faz sentido tomarmos como primeiro jogador o que entra no tempo $m = 1$, porque como dissemos anteriormente, o primeiro jogador só ganhará algo se a estrutura que ele estiver apostando for a mesma que será sorteada. E nesse caso seria impossível, pois como a estrutura final do time de apostas é menor a realização teria terminado antes. Assim, nesse caso igualaremos o tamanho das palavras para começar as apostas. Então, tomaremos como primeiro jogador o que entra no tempo $m = 2$:

1. o jogador que entra no tempo $m = 2$ observa que é sorteada a letra 1, então ele obrigatoriamente aposta em 1, como no tempo $m = 3$ a letra sorteada é 3, ele perde e deixa do jogo;

2. o jogador que entra no tempo $m = 3$ observa que é sorteada a letra 3, então ele obrigatoriamente aposta em 3, como no tempo $m = 4$ a letra sorteada é 2, ele perde e deixa do jogo;

3. o jogador que entra no tempo $m = 4$ observa que é sorteada a letra 2, então ele obrigatoriamente aposta em 3, como no tempo $m = 5$ a letra sorteada é 3, ele ganha $1/p_{23} = 4$.

Segue rapidamente mais dois exemplos para ficar claro que devemos somar os ganhos de cada jogador, para obter o ganho do time de apostas. Tomaremos o time de apostas 11323 e suponhamos que o resultado do sorteio seja o próprio 11323, então esse time ganhou $1/(p_{11}p_{13}p_{32}p_{23}) + 1/p_{23} = 268/3$. Suponhamos agora que o resultado do sorteio seja 22323 então o time de apostas 11323 ganhou $1/(p_{23}p_{32}p_{23}) + 1/p_{23} = 68$. Segue abaixo a tabela que contém os ganhos de todos os times de apostas em todos os sorteios possíveis.

Tabela 3.1: Soma dos ganhos dos times de apostas

| Sorteio \ Apostas | 11323 | 22323 | 11313 | 22313 | 1133 | 2233 |
|---|---|---|---|---|---|---|
| 323 | 4 | 0 | 4 | 0 | 4 | 0 |
| 313 | 0 | 4 | 0 | 4 | 0 | 4 |
| 33 | 2 | 2 | 2 | 2 | 2 | 2 |
| 1323 | 4 | 64 | 4 | 0 | 4 | 0 |
| 2323 | 68 | 0 | 4 | 0 | 4 | 0 |
| 1313 | 0 | 4 | 0 | 68 | 0 | 4 |
| 2313 | 0 | 4 | 64 | 4 | 0 | 4 |
| 133 | 2 | 2 | 2 | 2 | 2 | 10 |
| 233 | 2 | 2 | 2 | 2 | 10 | 2 |
| 11323 | 268/3 | 64 | 4 | 0 | 4 | 0 |
| 22323 | 68 | 256/3 | 4 | 0 | 4 | 0 |
| 11313 | 0 | 4 | 256/3 | 68 | 0 | 4 |
| 22313 | 0 | 4 | 64 | 268/3 | 0 | 4 |
| 1133 | 2 | 2 | 2 | 2 | 38/3 | 10 |
| 2233 | 2 | 2 | 2 | 2 | 10 | 38/3 |

**Observação 3.1.9.** *Estamos usando a palavra "sorteio"como uma maneira informal de designar a realização de uma cadeia de Markov.*

Note que definimos um jogo justo, ou seja, suponha que no tempo $n$ o jogador possua uma fortuna de $F_n := c = c(Z_1, ..., Z_n)$ reais e no tempo $n+1$ ele os aposta na letra $q$, então sua fortuna $F_{n+1}$ no tempo $n+1$ satisfaz a igualdade,

$$F_{n+1} = \frac{c}{\mathbb{P}(Z_{n+1} = q|\mathcal{F}_n)}\mathbb{I}_{\{Z_{n+1}=q\}}.$$

Então quando tomamos a esperança condicional temos que,

$$\mathbb{E}(F_{n+1}|\mathcal{F}_n) = \frac{c}{\mathbb{P}(Z_{n+1} = q|\mathcal{F}_n)}\mathbb{P}(Z_{n+1} = q|\mathcal{F}_n) = c.$$

Ou seja, o jogador espera ter a mesma fortuna após apostar. Seja $X_n$ a variável aleatória que denota o ganho do cassino sobre todos os times na conclusão da $n$-ésima rodada. Como o número de apostas na $n$-ésima rodada depende somente da história até o tempo $n-1$ e já mostramos que o jogo é justo, a sequência $\{X_n\}_{n\geq 0}$ com $X_0 = 0$ é um martingal, ou seja,

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n \text{ para todo } n \geq 0.$$

A fim de encontrar uma fórmula para o lucro do cassino, denotaremos por $K'$ a quantidade de estruturas iniciais e $N'$ a quantidade de estruturas finais. Por $y_j$ o valor que cada time (ligado a uma estrutura final) aposta, note que existem $N'$ times de apostas. E por $W$ a matriz de ganhos de cada time de apostas.

**Exemplo 3.1.10.** *No Exemplo 3.1.8, $K' = 9$ e $N' = 6$. E a matriz $W$ é dada por*

$$\begin{bmatrix} 4 & 0 & 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 & 0 & 4 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 4 & 64 & 4 & 0 & 4 & 0 \\ 68 & 0 & 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 68 & 0 & 4 \\ 0 & 4 & 64 & 4 & 0 & 4 \\ 2 & 2 & 2 & 2 & 2 & 10 \\ 2 & 2 & 2 & 2 & 10 & 2 \\ 268/3 & 64 & 4 & 0 & 4 & 0 \\ 68 & 256/3 & 4 & 0 & 4 & 0 \\ 0 & 4 & 256/3 & 68 & 0 & 4 \\ 0 & 4 & 64 & 268/3 & 0 & 4 \\ 2 & 2 & 2 & 2 & 38/3 & 10 \\ 2 & 2 & 2 & 2 & 10 & 38/3 \end{bmatrix}. \tag{3.3}$$

Observe agora que no tempo de parada $\tau$ o lucro do cassino é dado por

$$X_\tau = \sum_{j=1}^{N'} y_j(\tau - 1) - \sum_{i=1}^{K'}\sum_{j=1}^{N'} W_{ij}y_j\mathbb{I}_{E_i} - \sum_{i=k+1}^{K'+N'}\sum_{j=1}^{N'} W_{ij}y_j\mathbb{I}_{E_i},$$

em que $E_i$ é o evento em que a $i$-ésima estrutura ocorre. Observe primeiro que a cada rodada entra um novo jogador para cada time de aposta, e esse jogador aposta $y_j$ reais. Além disso todos os jogadores pertencentes a uma estrutura final, apostam a mesma quantia. Por último, a primeira letra a ser sorteada é apenas observada pelo apostador. Isso explica a primeira parte da equação

$$\sum_{j=1}^{N'} y_j(\tau - 1),$$

que é o dinheiro que o cassino ganha dos jogadores. A segunda parte da equação representa o dinheiro que o cassino deve pagar aos jogadores. Note que

$$\sum_{i=1}^{K'+N'} \sum_{j=1}^{N'} W_{ij} y_j \mathbb{I}_{E_i} = \sum_{i=1}^{K'} \sum_{j=1}^{N'} W_{ij} y_j \mathbb{I}_{E_i} + \sum_{i=K'+1}^{K'+N'} \sum_{j=1}^{N'} W_{ij} y_j \mathbb{I}_{E_i},$$

ou seja, por conveniência (facilitará as contas futuras), apenas separamos as estruturas iniciais das estruturas finais na matriz $W$. Assim, dado que é sorteada determinada estrutura o cassino deverá pagar para cada time de aposta seus respectivos ganhos.

Quando construímos a tabela $W$, ficou claro que cada $W_{ij}$ não era uma variável aleatória. Pois cada $W_{ij}$ dependia apenas da sobreposição da $i$-ésima estrutura sorteada com a $j$-ésima estrutura final associada a um time de apostas. Porém construímos $W$ de maneira informal, seguindo apenas as regras do jogo de apostas. Agora iremos introduzir uma fórmula para o que fizemos anteriormente. Considere a $kmA$ estrutura final e tomamos $A$ da seguinte forma

$$A = a_1 a_2 ... a_l.$$

Suponha que tenha sido sorteada a estrutura

$$Q = q_1 q_2 ... q_{l'}.$$

Primeiro introduziremos as medidas de sobreposição entre $kmA$ e $Q$

$$\delta_t(Q, kmA) = \begin{cases} \frac{1}{p_{km} p_{ma_1} \cdots p_{a_{t-2} a_{t-1}}}, & se \;\; q_{l'-t} = k, \; q_{l'-t+1} = m, \; q_{l'-t+2} = a_1, \ldots, \; q_{l'} = a_{t-1} \\ 0, & c.c. \end{cases}$$

e

$$\delta'_t(Q, kmA) = \begin{cases} \frac{1}{p_{q_{l'-t} a_1} p_{a_1 a_2} \cdots p_{a_{t-1} a_t}}, & se \;\; q_{l'-t} \neq k, \; q_{l'-t+1} = a_1, \; q_{l'-t+2} = a_2, \ldots, \; q_{l'} = a_t \\ 0, & c.c. \end{cases},$$

então o $kmA$ time de aposta ganha, se cada jogador apostar $R\$1,00$,

$$\sum_{t=1}^{\min(l'-1, l+1)} \delta_t(Q, kmA) + \delta'_t(Q, kmA).$$

A cada rodada entra um novo jogador o que significa que a cada letra da estrutura sorteada temos uma nova aposta acontecendo. Assim variamos $t$, $1 \leq t \leq \min(l'-1, l+1)$, para obter o ganho de cada jogador. Depois somamos todos os ganhos dos jogadores para obter o ganho do time de apostas.

A fórmula começa tomando o último jogador, se $q_{l'-1} = k$ temos que $\delta'_t(Q, kmA) = 0$ então tomamos $\delta_t(Q, kmA)$. Segundo a fórmula se $q_{l'} = m$ o jogador ganha $1/p_{km}$, senão o jogador deixa o jogo com zero reais. Por sua vez, se $q_{l'-1} \neq k$ tomamos $\delta'_t(Q, kmA)$, se $q_{l'} = a_1$ o jogador ganha $1/p_{q_{l'-1}a_1}$, senão deixa o jogo com zero reais.

Assim prosseguimos analogamente até $t = \min(l'-1, l+1)$, que corresponde as apostas do primeiro jogador. Se $\min(l'-1, l+1) = l-1$ então $l'-1 < l+1 \Leftrightarrow l' < l+2$, que significa que a estrutura sorteada é menor ou igual do que a estrutura apostada. Então a fórmula começa comparando a primeira letra da estrutura sorteada com a primeira letra da estrutura apostada, ou seja, em $q_{l'-(l'-1)} = q_1 = k$ ou $q_{l'-(l'-1)} = q_1 \neq k$. E termina ao final da realização, ou seja, para $q_{l'} = a_{l'-2}$ ou $q_{l'} = a_{l'-1}$.

Se $\min(l'-1, l+1) = l+1$ então $l' > l+2$, o que significa que a estrutura sorteada é maior do que a estrutura apostada. Então a fórmula começa comparando as letras onde as estruturas tenham o mesmo tamanho, em $q_{l'-(l+1)} = k$ ou $q_{l'-(l+1)} \neq k$. E termina nas últimas letras de ambas as estruturas $q_{l'} = q_l$ ou $q_{l'} = q_{l+1}$.

Note que no caso em que $l' = l+2$, ou seja, a estrutura sorteada e a estrutura apostada tem o mesmo tamanho, ambas as fórmulas coincidem com $q_1 = k$ ou $q_1 \neq k$.

Agora iremos finalmente deduzir $\mathbb{E}(\tau)$. Para isso iremos supor que para todo $K'+1 \leq i \leq K'+N'$,

$$\sum_{j=1}^{N'} W_{ij} y_j = 1. \tag{3.4}$$

Então

$$X_\tau = \sum_{j=1}^{N'} y_j(\tau - 1) - \sum_{i=1}^{K'} \sum_{j=1}^{N'} W_{ij} y_j \mathbb{I}_{E_i} - \sum_{i=K'+1}^{K'+N'} \mathbb{I}_{E_i}.$$

Seja $\tilde{w_{ij}}$ os elementos da matriz inversa de estruturas finais de $W$. Observe que $\sum_{j=1}^{N'} y_j = \sum_{i=1}^{N'} \sum_{j=1}^{N'} \tilde{w_{ij}}$. Tomando $k = \sum_{i=1}^{N'} \sum_{j=1}^{N'} \tilde{w_{ij}}$. Como $|X_n - X_{n-1}| = |\sum_{j=1}^{N'} y_j|$ então

$$|X_n - X_{n-1}| \leq k, \ \forall n \in \mathbb{N},$$

pelo item (3) do Teorema da parada ótima de Doob (Teorema 2.2.9)

$$\mathbb{E}(X_\tau) = \mathbb{E}(X_1) = 0.$$

Daí

$$0 = \mathbb{E}(X_\tau) = \sum_{j=1}^{N'} y_j(\mathbb{E}(\tau) - 1) - \sum_{i=1}^{K'} \sum_{j=1}^{N'} W_{ij} y_j \pi_i - \left(1 - \sum_{i=1}^{K'} \pi_i\right), \tag{3.5}$$

em que $\pi_i$ é a probabilidade de que a $i$-ésima estrutura inicial ocorra. Resolvendo a equação com respeito a $\mathbb{E}(\tau)$, obtemos o resultado a seguir.

**Teorema 3.1.11.** *Se $(y_1, y_2, ..., y_{N'})$ resolve os sistema linear (3.4) então*

$$\mathbb{E}(\tau) = 1 + \frac{\left(1 - \sum_{i=1}^{K'} \pi_i\right) + \sum_{i=1}^{K'} \pi_i \sum_{j=1}^{N'} y_j W_{ij}}{\sum_{j=1}^{N'} y_j}.$$

**Exemplo 3.1.12.** *No Exemplo 3.1.8 aplicaremos o Teorema 3.1.11 com o intuito de calcular $\mathbb{E}(\tau)$. Começaremos encontrando os valores de $\pi_i$ para $1 \leq i \leq 9$:*

Tabela 3.2: Probabilidades estruturas iniciais

| Estrutura | Probabilidade |
|---|---|
| 323 | $\pi_1 = \frac{1}{3}\frac{1}{4}\frac{1}{4} = \frac{1}{48}$ |
| 313 | $\pi_2 = \frac{1}{3}\frac{1}{4}\frac{1}{4} = \frac{1}{48}$ |
| 33 | $\pi_3 = \frac{1}{3}\frac{1}{2} = \frac{1}{6}$ |
| 1323 | $\pi_4 = \frac{1}{3}\frac{1}{4}\frac{1}{4}\frac{1}{4} = \frac{1}{192}$ |
| 2323 | $\pi_5 = \frac{1}{3}\frac{1}{4}\frac{1}{4}\frac{1}{4} = \frac{1}{192}$ |
| 1313 | $\pi_6 = \frac{1}{3}\frac{1}{4}\frac{1}{4}\frac{1}{4} = \frac{1}{192}$ |
| 2313 | $\pi_7 = \frac{1}{3}\frac{1}{4}\frac{1}{4}\frac{1}{4} = \frac{1}{192}$ |
| 133 | $\pi_8 = \frac{1}{3}\frac{1}{4}\frac{1}{2} = \frac{1}{24}$ |
| 233 | $\pi_9 = \frac{1}{3}\frac{1}{4}\frac{1}{2} = \frac{1}{24}$ |

*Assim $\sum_{i=1}^{9} \pi_i = 0,3123937$. Para prosseguir temos que encontrar os valores de $y_j$ com $1 \leq j \leq 6$, tais que*

$$\sum_{j=1}^{6} W_{ij} y_j = 1.$$

*Usando o software R escreveremos a matriz dos coeficientes (coef)*

$$\begin{bmatrix} 268/3 & 64 & 4 & 0 & 4 & 0 \\ 68 & 256/3 & 4 & 0 & 4 & 0 \\ 0 & 4 & 256/3 & 68 & 0 & 4 \\ 0 & 4 & 64 & 268/3 & 0 & 4 \\ 2 & 2 & 2 & 2 & 38/3 & 10 \\ 2 & 2 & 2 & 2 & 10 & 38/3 \end{bmatrix}.$$

*A matriz dos termos independentes (ind)*

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

*Usando o comando solve(coef,ind) obtemos que os valores de $y_j$ com $1 \leq j \leq 6$ são*

$$\begin{bmatrix} 0,00528169 & 0,00528169 & 0,00528169 & 0,00528169 & 0,04225352 & 0,04225352 \end{bmatrix}.$$

*Somando os valores obtemos que* $\sum_{j=1}^{6} y_j = 0,1056338$. *Agora iremos resolver a seguinte parte da equação*

$$\sum_{i=1}^{9} \pi_i \sum_{j=1}^{6} y_j W_{ij}.$$

*Para tal usaremos novamente o software R. Escrevemos a matriz das estruturas iniciais (matrizestinicial)*

$$\begin{bmatrix} 4 & 0 & 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 & 0 & 4 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 4 & 64 & 4 & 0 & 4 & 0 \\ 68 & 0 & 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 68 & 0 & 4 \\ 0 & 4 & 64 & 4 & 0 & 4 \\ 2 & 2 & 2 & 2 & 2 & 10 \\ 2 & 2 & 2 & 2 & 10 & 2 \end{bmatrix},$$

*os valores que encontramos de* $y_j$ *(apostas)*

$$\begin{bmatrix} 0,00528169 & 0,00528169 & 0,00528169 & 0,00528169 & 0,04225352 & 0,04225352 \end{bmatrix}$$

*com os valores de* $\pi$ *(pi)*

$$\begin{bmatrix} \frac{1}{48} & \frac{1}{48} & \frac{1}{6} & \frac{1}{192} & \frac{1}{192} & \frac{1}{192} & \frac{1}{192} & \frac{1}{24} & \frac{1}{24} \end{bmatrix}.$$

*Usando o comando pi%*%(matrizestinicial%*%apostas) obtemos*

$$\sum_{i=1}^{9} \pi_i \sum_{j=1}^{6} y_j W_{ij} = 0,1012324.$$

*Substituindo todos os valores que encontramos*

$$\mathbb{E}(\tau) = 1 + \frac{\left(1 - \sum_{i=1}^{K'} \pi_i\right) + \sum_{i=1}^{K'} \pi_i \sum_{j=1}^{N'} y_j W_{ij}}{\sum_{j=1}^{N'} y_j} = 1 + \frac{(1 - 0,3123937) + 0,1012324}{0,1056338},$$

*temos por fim*

$$\mathbb{E}(\tau) = 8,466667.$$

*Isso significa que temos que sortear em média entre 8 e 9 letras até que uma das palavras* $\mathcal{C} = \{323, 313, 33\}$ *ocorra.*

## 3.2 Probabilidade de parada

Nessa seção iremos encontrar a probabilidade de que determinada palavra ocorra antes das demais, para isso estudamos o artigo Gava e Salotti[9]. Lembre-se que definimos o espaço de estados por $\Omega = \{1, \ldots, K\}$ e $\{Z_n\}_{n \geq 1}$ uma cadeia de Markov em $\Omega$. A distribuição inicial é $\mathbb{P}(Z_1 = k) = \lambda_k$ e a matriz de transição é $\{p_{ij}\}$, em que $p_{ij} = \mathbb{P}(Z_{n+1} = j | Z_n = i)$. Definimos o conjunto de palavras por $\mathcal{C} = \{A_i\}_{i=1}^M$. Assumimos que:

1. para $i = 1, \ldots, M$, $\mathbb{P}(\tau = \tau_{A_i}) > 0$;
2. $\tau < \infty$ quase certamente.

Observe que as hipóteses 1. e 2. implicam que $\sum_i \mathbb{P}(\tau = \tau_{A_i}) = 1$. Agora precisamos definir mais dois conjuntos,

$$\mathcal{D}' = \{lA_i; l = 1, \ldots, K\}_{i=1}^M \text{ e } \mathcal{C}' = \{lmA_i; l, m = 1, \ldots, K\}_{i=1}^M.$$

Denote por $\mathcal{D}''$ e por $\mathcal{C}''$ a lista de palavras depois de excluir de $\mathcal{D}'$ e de $\mathcal{C}'$ respectivamente, as palavras que só podem ocorrer depois do tempo $\tau$. No Exemplo 3.1.8 $\mathcal{D}' = \{1323, 2323, 1313, 2313, 133, 233, 3323, 3313, 333\}$ e $\mathcal{D}'' = \{1323, 2323, 1313, 2313, 133, 233\}$. Notemos que,

$$\mathbb{P}(\tau = \tau_{A_i}) = \pi_i + \mathbb{P}(B \in \mathcal{D}'' \cup \mathcal{C}'', \text{ isto é, o jogo acaba com B que é associado a } A_i),$$

em que $\mathcal{D}'' \cap \mathcal{C}'' = \phi$. Para $r \in \left\{K', \ldots, K' + N'\right\}$, escreva $\mathbb{P}(E_r) := \pi_r^{(i)}$ para enfatizar a ligação entre a r-ésima estrutura e a palavra $A_i$ a qual ela é associada (em que $E_r$ é o evento que a $r$-ésima estrutura ocorre). Depois assuma que $(z_1, \ldots, z_{N'})$ é a solução do sistema linear

$$z_1 W_{11} + \ldots + z_{N'} W_{1N'} = 1 \tag{3.6}$$

$$z_1 W_{i1} + \ldots + z_{N'} W_{iN'} = 1 \text{ para } i \in \left\{K' + 1, \ldots, K' + N'\right\} \setminus \{r\}. \tag{3.7}$$

Note que o cassino paga para os jogadores

$$S(y_1, \ldots, y_{N'}) = \sum_{i=K'+1}^{K'+N'} \sum_{j=1}^{N'} W_{ij} y_j \mathbb{I}_{E_i}.$$

Então dado (3.6) podemos escrever a quantia que o cassino paga para os jogadores como

$$S(z_1, \ldots, z_{N'}) = 1\mathbb{I}_{E_1} + \sum_{i=2}^{K'} \mathbb{I}_{E_i} \sum_{j=1}^{N'} z_j W_{ij} + \sum_{i > K'; i \neq r} \mathbb{I}_{E_i} + \mathbb{I}_{E_r} \sum_{j=1}^{N'} z_j W_{rj}.$$

Novamente, aplicando o Teorema da parada ótima de Doob (Teorema 2.2.9), tomamos o valor esperado em ambos os lados para $S(z_1, \ldots, z_{N'})$

$$0 = \mathbb{E}\left(X_\tau\right) = \sum_{j=1}^{N'} z_j \left(\mathbb{E}(\tau) - 1\right) - \sum_{i=2}^{K'} \pi_i \sum_{j=1}^{N'} z_j W_{ij} - \left(1 - \sum_{i=2}^{K'} \pi_i\right) + \pi_r^{(i)} \left(1 - \sum_{j=1}^{N'} z_j W_{rj}\right),$$

o que implica que

$$\pi_r^{(i)} = \frac{\sum_{j=1}^{N'} z_j (1 - \mathbb{E}(\tau)) - \sum_{i=2}^{K'} \pi_i (1 - \sum_{j=1}^{N'} z_j W_{ij}) + 1}{1 - \sum_{j=1}^{N'} z_j W_{rj}} \tag{3.8}$$

como queríamos. Note que pelo Teorema 3.1.11 conhecemos $\mathbb{E}(\tau)$, por isso é possível encontrar a solução desse sistema.

**Teorema 3.2.1.** *Seja $\{Z_n\}_{n\geq 1}$ uma cadeia de Markov homogênea em $\{1, \ldots, K\}$. Para $A_i$ suponha que as palavras $B_k, \ldots, B_{k+v} \in \mathcal{D}''$ estão associadas com $A_i$ e correspondem com as estruturas finais $u, u+1, \ldots, u+v$. Suponha também que $B_l, B_{l+1}, \ldots, B_{l+s} \in \mathcal{C}''$ estão associados com $A_i$ e correspondem com as estruturas finais $r, r+1, \ldots, r+s$. Então*

$$\mathbb{P}(\tau = \tau_{A_i}) = \pi_i + \pi_u^{(i)} + \ldots + \pi_{u+v}^{(i)} + \pi_r^{(i)} + \ldots + \pi_{r+s}^{(i)},$$

*em que cada $\pi_l^{(i)}$, $l = r, \ldots, r+s$, demanda uma solução $(z_1^l, \ldots, z_{N'}^l)$ para (3.6) com $l$ ao invés de $r$ tem a forma dada por (3.8).*

**Exemplo 3.2.2.** *Retomando o Exemplo 3.1.8 temos que*

$$\mathcal{D}'' = \{1323, 2323, 1313, 2313, 133, 233\}$$

*e*

$$\mathcal{C}'' = \{11323, 22323, 11313, 22313, 1133, 2233\}.$$

*Considere $A_1 = 323$, $A_2 = 313$ e $A_3 = 33$. Usando que $\mathbb{E}(\tau) = 8,466667$ e o Teorema 3.2.1 iremos calcular $\mathbb{P}(\tau = \tau_{A_1})$, $\mathbb{P}(\tau = \tau_{A_2})$ e $\mathbb{P}(\tau = \tau_{A_3})$. As estruturas foram seguem a ordem da tabela (3.1). Então*

$$\mathbb{P}(\tau = \tau_{A_1}) = \pi_1 + \pi_4^{(1)} + \pi_5^{(1)} + \pi_{10}^{(1)} + \pi_{11}^{(1)}.$$

*Temos que:*

*Para calcular $\pi_{10}^{(1)}$ usamos o software Maple. Começamos resolvendo o sistema 3.6, entramos com a matriz $W_{ij}$ na forma que é definida em (3.6) a qual denominaremos "MatrizInicial"*

$$\begin{bmatrix} 4 & 0 & 4 & 0 & 4 & 0 \\ 68 & 256/3 & 4 & 0 & 4 & 0 \\ 0 & 4 & 256/3 & 68 & 0 & 4 \\ 0 & 4 & 64 & 268/3 & 0 & 4 \\ 2 & 2 & 2 & 2 & 38/3 & 10 \\ 2 & 2 & 2 & 2 & 10 & 38/3 \end{bmatrix}.$$

Tabela 3.3: Probabilidades estruturas iniciais

| Estrutura | Probabilidade |
|---|---|
| 323 | $\pi_1 = \frac{1}{3}\frac{1}{4}\frac{1}{4} = \frac{1}{48}$ |
| 1323 | $\pi_4 = \frac{1}{3}\frac{1}{4}\frac{1}{4}\frac{1}{4} = \frac{1}{192}$ |
| 2323 | $\pi_5 = \frac{1}{3}\frac{1}{4}\frac{1}{4}\frac{1}{4} = \frac{1}{192}$ |

*Usando o comando "LinearAlgebra[MatrixInverse](MatrizInicial);" invertemos a "MatrizInicial" que denotaremos por "InversaMatrizInicial", depois escrevemos a matriz dos termos independente denotada por "Resultado"*

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

*Usando o comando "VetorZ:=InversaMatrizInicial.Resultado;"obtemos o vetor Z, denotado por "VetorZ"*

$$\begin{bmatrix} \frac{16}{79} \\ -\frac{12}{79} \\ \frac{3}{316} \\ \frac{3}{316} \\ \frac{3}{79} \\ \frac{3}{79} \end{bmatrix}.$$

*Para obter* $\sum_{j=1}^{N'} z_j$ *escrevemos um vetor linha de uns denotado por "Auxiliar" e usando o comando "Soma:=Auxiliar.VetorZ;" temos que*

$$\sum_{j=1}^{N'} z_j = \frac{23}{158}.$$

*Para resolver o restante do sistema precisamos escrever o vetor da linha R denominado "VetorR"*

$$\begin{bmatrix} 268/3 & 64 & 4 & 0 & 4 & 0 \end{bmatrix}.$$

*Escrevemos também a matriz* $W_{ij}$ *da estruturas iniciais sem a primeira linha denotada por "SegundaMatriz"*

$$\begin{bmatrix} 0 & 4 & 0 & 4 & 0 & 4 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 4 & 64 & 4 & 0 & 4 & 0 \\ 68 & 0 & 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 68 & 0 & 4 \\ 0 & 4 & 64 & 4 & 0 & 4 \\ 2 & 2 & 2 & 2 & 2 & 10 \\ 2 & 2 & 2 & 2 & 10 & 2 \end{bmatrix}.$$

*E usamos o comando*

$$Final := \frac{(Soma * (1 - 8.66667)ProbInicial.(1 - SegundaMatriz.VetorZ)) + 1}{1 - VetorZ.VetorR}.$$

*Obtemos que* $\pi_{10}^{(1)} = 0.3437500651$. *Procedendo analogamente calculamos* $\pi_{11}^{(1)} = 0.343750$. *Por fim*

$$\mathbb{P}(\tau = \tau_{A_1}) = \pi_1 + \pi_4^{(1)} + \pi_5^{(1)} + \pi_{10}^{(1)} + \pi_{11}^{(1)} = 1/48 + 2 \times 1/192 + 2 \times 0.34375 = 1/10.$$

*Analogamente*

$$\mathbb{P}(\tau = \tau_{A_2}) = \pi_2 + \pi_6^{(2)} + \pi_7^{(2)} + \pi_{12}^{(2)} + \pi_{13}^{(2)} = 1/48 + 2 \times 1/192 + 2 \times 0.34375 = 1/10,$$
$$\mathbb{P}(\tau = \tau_{A_3}) = \pi_3 + \pi_8^{(3)} + \pi_9^{(3)} + \pi_{14}^{(3)} + \pi_{15}^{(3)} = 1/6 + 2 \times 1/24 + 2 \times 0.275 = 1/10.$$

Isso significa que cada uma dessas palavras tem a mesma probabilidade de ocorrer antes da demais que é igual a 1/10.

## 3.3 Função geradora de probabilidade

Para encontrar a função geradora de $\tau$ ($\mathbb{E}(t^\tau)$, $0 \leq t \leq 1$), retomaremos o artigo Pozdnyakov[6]. Começamos introduzindo as mesmas estruturas e os mesmos times de apostas, porém mudaremos a quantidade das apostas iniciais. Mais especificamente, um apostador do j-ésimo time que começa sua aposta na n-ésima rodada apostará inicialmente $y_j t^n$. Definimos por $t^\tau W_{ij}(t) y_j$ o total de ganhos do j-ésimo time na i-ésima estrutura.

Como antes, a observação mais importante é que $W_{ij}(t)$ não é uma variável aleatória e é totalmente determinada pela relação entre o j-ésimo time de aposta e a i-ésima estrutura final. Se $X_n$ novamente denota o lucro do cassino no tempo $n$ então $X_\tau$ é dado por

$$\begin{aligned} X_\tau &= (t^2 + ... + t^\tau) \sum_{j=1}^{N'} y_j - \sum_{i=1}^{K'} \mathbb{I}_{E_i} \sum_{j=1}^{N'} t^\tau y_j W_{ij}(t) - \sum_{i=K'+1}^{N'+K'} \mathbb{I}_{E_i} \sum_{j=1}^{N'} t^\tau y_j W_{ij}(t) \\ &= \frac{t^2 - tt^\tau}{1-t} \sum_{j=1}^{N'} y_j - \sum_{i=1}^{K'} \mathbb{I}_{E_i} \sum_{j=1}^{N'} t^\tau y_j W_{ij}(t) + \sum_{i=1}^{K'} \mathbb{I}_{E_i} t^\tau - \sum_{i=1}^{K'} \mathbb{I}_{E_i} t^\tau - \sum_{i=K'+1}^{N'+K'} \mathbb{I}_{E_i} \sum_{j=1}^{N'} t^\tau y_j W_{ij}(t) \end{aligned}$$

em que $E_i$, $1 \leq i \leq K' + N'$ é o evento que a i-ésima estrutura ocorra. Suponha que para todo $0 < t < 1$ podemos encontrar $(y_1, ..., y_{N'})$ tal que

$$\textstyle\sum_{j=1}^{N'} W_{ij}(t) y_j = 1, \text{ para todo } K' + 1 < i < K' + N'.$$

Então $X_\tau$ é dado por

$$\begin{aligned} X_\tau &= \frac{t^2 - tt^\tau}{1-t} \sum_{j=1}^{N'} y_j - \sum_{i=1}^{K'} t^\tau \mathbb{I}_{E_i} \left( \sum_{j=1}^{N'} y_j W_{ij}(t) - 1 \right) - t^\tau \\ &= \frac{t^2 - tt^\tau}{1-t} \sum_{j=1}^{N'} y_j - \sum_{i=1}^{K'} t^{|A_i|} \mathbb{I}_{E_i} \left( \sum_{j=1}^{N'} y_j W_{ij}(t) - 1 \right) - t^\tau, \end{aligned}$$

em que $|A_i|$, $1 \leq i \leq K'$ é o valor de $\tau$ quando a i-ésima estrutura inicial ocorre. Note que $|A_i|$ não é uma variável aleatória. Aplicando o Teorema da parada ótima de Doob (Teorema 2.2.9) obtemos

$$0 = \frac{t^2 - t\mathbb{E}(t^\tau)}{1-t} \sum_{j=1}^{N'} y_j - \sum_{i=1}^{K'} t^{|A_i|} \pi_i \left( \sum_{j=1}^{N'} y_j W_{ij}(t) - 1 \right) - \mathbb{E}(t^\tau).$$

Depois de alguma álgebra segue o resultado.

**Teorema 3.3.1.** *Se* $(y_1, y_2, ..., y_{N'})$ *resolve o sistema linear (3.4), então*

$$\mathbb{E}(t^\tau) = \frac{t^2/(1-t) \sum_{j=1}^{N'} y_j - \sum_{i=1}^{K'} t^{|A_i|} \pi_i \left( \sum_{j=1}^{N'} y_j W_{ij}(t) - 1 \right)}{1 + t/(1-t) \sum_{j=1}^{N'} y_j}.$$

**Exemplo 3.3.2.** *Voltando ao Exemplo 3.1.8, vamos aplicar o Teorema 3.3.1 para encontrar a função geradora de probabilidades. Nesse exemplo a matriz $W(t)$ é dada por*

$$W(t) = \begin{bmatrix} W^I(t) \\ W^F(t) \end{bmatrix}. \tag{3.9}$$

*Em que as matrizes $W^I$ e $W^F$ serão descritas a seguir. Para resolver a equação*

$\sum_{j=1}^{N'} W_{ij}(t)y_j = 1$ *para todo* $K' + 1 < i < K' + N'$,

*usamos o aplicativo Maple. Primeiro escrevemos a matriz de estruturas finais denotada por "MatrizFinal"*

$$W^F(t) = \begin{bmatrix} \frac{256}{3t^3}+4 & \frac{64}{t^2} & 4 & 0 & 4 & 0 \\ \frac{64}{t^2}+4 & \frac{256}{3t^3} & 4 & 0 & 4 & 0 \\ 0 & 4 & \frac{256}{3t^3} & \frac{64}{t^2}+4 & 0 & 4 \\ 0 & 4 & \frac{64}{t^2} & \frac{256}{3t^3}+4 & 0 & 4 \\ 2 & 2 & 2 & 2 & \frac{32}{3t^2}+2 & \frac{8}{t}+2 \\ 2 & 2 & 2 & 2 & \frac{8}{t}+2 & \frac{32}{3t^2}+2 \end{bmatrix}.$$

*Usando o comando "LinearAlgebra[MatrixInverse](MatrizFinal)" invertemos a matriz de estruturas finais. E então multiplicamos a Inversa da matriz de estruturas final por um vetor coluna de uns, obtendo então o vetor Y denotado por "VetorY"*

$$\begin{bmatrix} \frac{3}{8}\frac{t^3}{3t^3+12t^2+24t+32} \\ \frac{3}{8}\frac{t^3}{3t^3+12t^2+24t+32} \\ \frac{3}{8}\frac{t^3}{3t^3+12t^2+24t+32} \\ \frac{3}{8}\frac{t^3}{3t^3+12t^2+24t+32} \\ \frac{3t^2}{3t^3+12t^2+24t+32} \\ \frac{3t^2}{3t^3+12t^2+24t+32} \end{bmatrix}.$$

*A soma de todos os valores do vetor Y denotado por "Soma" é*

$$\sum_{j=1}^{6} y_j = \frac{3}{2}\frac{t^2(t+4)}{3t^3+12t^2+24t+32}.$$

*Para terminar de resolver a equação presente no Teorema 3.3.1 escrevemos a matriz das estruturas iniciais denotada por " MatrizInicial"*

$$W^I(t) = \begin{bmatrix} 4 & 0 & 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 & 0 & 4 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 4 & \frac{64}{t^2} & 4 & 0 & 4 & 0 \\ \frac{64}{t^2}+4 & 0 & 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & \frac{64}{t^2}+4 & 0 & 4 \\ 0 & 4 & \frac{64}{t^2} & 4 & 0 & 4 \\ 2 & 2 & 2 & 2 & 2 & \frac{8}{t}+2 \\ 2 & 2 & 2 & 2 & \frac{8}{t}+2 & 2 \end{bmatrix}.$$

*Com a matriz de probabilidades iniciais multiplicadas por $t^{A_i}$ denotada por "ProbInicial"*

$$\begin{bmatrix} \frac{t^3}{48} & \frac{t^3}{48} & \frac{t^2}{6} & \frac{t^4}{192} & \frac{t^4}{192} & \frac{t^4}{192} & \frac{t^4}{192} & \frac{t^3}{24} & \frac{t^3}{24} \end{bmatrix}. \tag{3.10}$$

*Por fim calculamos a função geradora de probabilidades usando o comando*

$$\frac{\frac{t^2}{(1-t)} * Soma - ProbInicial.(MatrizInicial.VetorY - \bar{1})}{1 + \frac{t}{(1-t)} * Soma}.$$

*Cujo o resultado é dado por*

$$\mathbb{E}(t^\tau) = \frac{t^2(16-t^2)}{96-9t(8+t^2)} = \frac{1}{6}t^2 + \frac{1}{8}t^3 + \frac{1}{12}t^4 + \frac{5}{64}t^5 + \frac{9}{128}t^6 + \frac{31}{512}t^7 + O(t^8).$$

*Isso significa que a $\mathbb{P}(\tau = 2) = 1/6$, $\mathbb{P}(\tau = 3) = 1/8$, $\mathbb{P}(\tau = 4) = 1/12$ e segue respectivamente para todos os valores de $\tau$.*

# Capítulo 4

# Conclusão

Nessa dissertação estudamos principalmente o artigo Pozdnyakov[6], cujo objetivo central foi o de encontrar a distribuição do tempo de espera $\tau$ para ocorrência de palavras em cadeias de Markov. Começamos com o Capítulo 2 em que apresentamos um resumo de conceitos de função geradora de probabilidade, martingais e cadeias de Markov que foram fundamentais para o desenvolvimento dessa dissertação no sentido que o texto seja autocontido. No Capítulo 3 encontramos a distribuição do valor esperado de $\tau$, a probabilidade de uma das palavras de $\mathcal{C}$ ocorra antes das demais e por último encontramos a função geradora de probabilidade de $\tau$.

Como podemos observar no artigo Pozdnyakov[6] a distribuição do tempo de espera é a chave para muitos problemas reais em questões de vários campos da ciência como: controle de qualidade, teste de hipótese, biologia molecular, sequência de DNA entre outras. Por isso é necessário que tenhamos uma fórmula simples e objetiva para calcular $\mathbb{E}(\tau)$ sem para isso dependermos da derivada da função geradora de probabilidade. Por esse motivo esse trabalho não poderia ter sido resumido a última seção 3.3 como um leitor curioso poderia se indagar.

Observe que para encontrarmos $\mathbb{E}(\tau)$ precisamos assumir que a solução da equação (3.4) exista, como pesquisa futura podemos tentar demonstrar que essa solução sempre existe, assim como foi feito por Gerber e Li[10] em 1981 para variáveis aleatórias i.i.d..

# Referências Bibliográficas


[1] Gut, A. *An intermediate course in probability.* New York, Springer, 2009.

[2] Gut, A. *Probability: A graduate course in probability.* New York, Springer, 2005.

[3] Ross, S. *Stochastic Processes.*Segunda Edição, New York, Wiley, 1996.

[4] Williams, D. *Probability and Martingales.* Cambridge, Cambridge University Press, 1991.

[5] Norris, J. R. *Markov Chains.* Cambridge, Cambridge University Press, 1998.

[6] Pozdnyakov, V. *On ocurrence of patterns in Markov chain: method of gambling teams.* Stat. Probab. Letters, 78, 2762-2767, 2008.

[7] Li, S.Y.R. *A martigale approach to the study of ocurrence of sequence patterns in repeated experiments.*Ann. Prob. 8, 1171-1176, 1980.

[8] Glaz, J., Kulldorff, M., Pozdnyakov, V., Steele, J.M. *Gambling teams and waiting times for patterns in two-state Markov chains.*J. Appl. Prob. 43, 127-140, 2006.

[9] Gava, R.J., Salotti, D. *Stopping probabilities for patterns in Markov chains.*J. Appl. Prob. 51, 287-292, 2014.

[10] Gerber, H.U. e Li, S.Y.R. *The occurrence of sequence patterns in repeated experiments and hitting times in a markov chain.* Stochastic Processes and their Applications, 11(1), 101-108, 1981.