Universidade Estadual de Campinas

Instituto de Matemática, Estatística e Computação Científica

Departamento de Estatística

Dissertação de Mestrado

Comportamento assintótico do Primeiro Retorno de uma sequência gerada por variáveis aleatórias independentes e identicamente distribuídas

por

Rodrigo Lambert †

Mestrado em Estatística - Campinas - SP

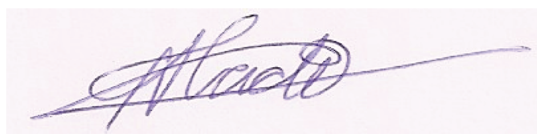
Orientador: Prof. Dr. Miguel Natalio Abadi

†Este trabalho contou com apoio financeiro do CNPQ

Comportamento assintótico do Primeiro Retorno de uma sequência gerada por variáveis aleatórias independentes e identicamente distribuídas

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por **Rodrigo Lambert** e aprovada pela comissão julgadora.

Campinas, 30 de julho de 2010.



Prof. Dr: Miguel Natalio Abadi
Orientador

Banca Examinadora:

Prof. Dr. Miguel Natalio Abadi.

Prof. Dr. Jefferson Antonio Galves.

Prof. Dr. Serguei Popov.

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para obtenção do Título de **MESTRE em ESTATÍSTICA .**

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP
Bibliotecária: Maria Fabiana Bezerra Müller – CRB8 / 6162

Lambert, Rodrigo
L172c Comportamento assintótico do primeiro retorno de uma sequência gerada por variáveis aleatórias independentes e identicamente distribuídas/Rodrigo Lambert-- Campinas, [S.P. : s.n.], 2010.

Orientador : Miguel Natalio Abadi
Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1.Teoria ergodica. 2.Grandes desvios. 3.Sequências (Matemática). I. Abadi, Miguel Natalio. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Título em inglês: Convergence in distribution of the overlapping function: the IID case

Palavras-chave em inglês (Keywords): 1. Ergodic theory. 2. Large deviations. 3. Sequences (Mathematics).

Área de concentração: Probabilidade

Titulação: Mestre em Estatística

Banca examinadora: Prof. Dr. Miguel Natalio Abadi (IMECC – UNICAMP)
Prof. Dr. Jefferson Antonio Galves (IME – USP)
Prof. Dr. Sergei Popov (IMECC - UNICAMP)

Data da defesa: 30/07/2010

Programa de Pós-Graduação: Mestrado em Estatística

Dissertação de Mestrado defendida em 30 de julho de 2010 e aprovada

Pela Banca Examinadora composta pelos Profs. Drs.

Prof(a). Dr(a). MIGUEL NATALIO ABADI

Prof(a). Dr(a). JEFFERSON ANTONIO GALVES

Sergueí Popov

Prof(a). Dr(a). SERGUEI POPOV

DEDICATÓRIA

À minha mãe Thelma, que
me ensinou a entender
claramente a frase
tão difícil de explicar:
Mãe é mãe!

Agradecimentos

- Primeiramente, agradeço a minha mãe, por sua dedicação e sua presença durante toda minha vida.

- A minha avó Rita e minha tia Elizabeth, que me deram a oportunidade de saber como é ser criado por três mães.

- Ao meu pai, pelo apoio e pela força que me deu, e por ser um exemplo de que estudar é um grande caminho para a vida.

- A todos os meus amigos de Goiânia, que me convencem a cada dia que a distância pode ser superada pela amizade. Em particular, aos 5 guerreiros que rasgaram mais de mil quilômetros de estrada para assistir a uma simples defesa de dissertação de mestrado. David, João Gabriel(Dil), Diogo(Dente), Fábio e Felipe: tudo o que eu puder escrever aqui não conseguirá representar o meu sentimento de gratidão e dívida para com vocês.

- Aos amigos goianos que estão ou estiveram geograficamente mais próximos: Lucas, Edcarlos, Dylene, Thiago, Tiago e Joaquim.

- Ao meu irmão Diogo, e ao meu primo Aristarcho, que sempre me deram força e apoio.

- A minhas irmãs Larissa, Sofia e Clara, e minha prima Juliana.

- Ao meu orientador e grande amigo Miguel Abadi, por sua paciência, incentivo, e por nunca ter fugido em momento algum da sua posição de orientador, sempre se esforçando para me ensinar a fazer matemática e me tornar um pesquisador um dia.

- Ao casal gaúcho mais simpático do Brasil Karen e Márcio, por estarem sempre presentes nesse período, e por sempre me levarem alegria até quando parecia impossível.
- Ao Aldo, que me ajudou a aprender algo de estatística.
- Ao Guilherme, a quem se devem as simulações do Apêndice dessa dissertação. Um cara que mostra talento tanto para estatística quanto para matemática, e certamente tem um grande caminho pela frente.
- Aos meus amigos de repúbica, que acabaram por constituir a minha nova família: Marquinhos, Luciano, Vitão, Elen, Linão, Welington(Chorão), Anderson(Cabecinha). Sempre que precisarem de mim, podem contar comigo.
- A minha amiga e sempre orientadora Denise Duarte, por sempre se preocupar comigo e ter me ensinado os primeiros passos do trabalho de pesquisa.
- Ao meus amigos de turma, em especial Karina, Bia, Zé, Márcio e Lorena.
- Ao CNPQ pelo apoio financeiro.
- A Deus e aos meus anjos da guarda, que me guiaram com saúde e alegria durante esses anos.

Resumo

Seja χ um alfabeto finito ou enumerável, e considere o espaço de todas as seqüências finitas compostas por concatenação de símbolos desse alfabeto. A essas seqüências daremos o nome de palavras. Denotaremos por χ^n conjunto de todas as palavras de tamanho n. No presente trabalho, consideramos uma função que leva cada palavra de tamanho n em um número inteiro entre 0 e $n-1$. Essa função é definida pelo maior tamanho possível de uma sobreposição da palavra com uma cópia dela mesma transladada, e é chamada de função de sobreposição. A ela daremos o nome de S_n. A relevância da função de sobreposição foi colocada em evidência, entre outros casos, na análise estatística da Recorrência de Poincaré, e possui relação explícita com a entropia do processo. Nesse trabalho, provamos a convergência da distribuição da função de sobreposição, quando a seqüência é escolhida de acordo com relação a n variáveis aleatórias independentes e identicamente distribuídas no alfabeto χ. Também apresentamos um limitante para a velocidade dessa convergência. Como conseqüência, mostramos também a convergência da esperança e da variância da função de sobreposição.

Abstract

We consider the set of finite sequencies of length n over a finite or contable alphabet χ. We consider the function defined over χ^n which gives the size of the maximum overlap of a given sequence with a (shifted) copy of itself. That function will be denoted by *overlapping function*. We prove the convergence of the distribution of this function when the sequence is chosen according to a product measure, with identically distributed marginals. We give a point-wise upper bound for the velocity of this convergence. As a byproduct, we show the convergence of te mean and the variance of the overlapping function.

Sumário

Introdução

Seja n um número inteiro positivo. Considere o espaço de todas as sequências de tamanho n com símbolos tomados de um alfabeto χ finito ou enumerável. A essas sequências daremos o nome de *palavras* (ver Definição 1.1.1). No presente trabalho, nós consideramos uma função S_n definida em χ^n e tomando valores em $\{0, 1, \cdots, n-1\}$. Para cada palavra, essa função assume o valor do maior tamanho possível de uma sobreposição que essa palavra pode ter com uma cópia dela mesma transladada. Se não houver sobreposição, a função assume o valor zero (ver Definição 2.1.1).

A função S_n está relacionada com a função τ_n, que também está definida sobre χ^n e assume valores em $\{1, 2, \cdots, n\}$, e indica a posição do *primeiro encaixe* que a palavra realizou com uma cópia dela transladada(ver Definição 1.2.1). Essa relação é dada pela equação $S_n = n - \tau_n$.

A relevância da função τ_n (e portanto de S_n) foi colocada em evidência na análise estatística da *Recorrência de Poincaré*. Para provar a convergência do número de ocorrência de palavras (digamos, de tamanho n) quando n diverge, para a distribuição de Poisson, é necessário qua a palavra não se sobreponha com ela mesma [16]. Ou, pelo menos, que a proporção das palavras que se sobrepõem seja pequena com relação a n [1]. Se esse não for o caso, a distribuição limite tem lei Poisson composta [15]. Também temos na literatura algumas aproximações para esse limite em [20, 21, 22].

Tal relevância também aparece quando consideramos o tempo decorrido até a primeira ocorrência da palavra. Esse tempo é conhecido como *Tempo de Ocorrência (hitting time)*. É sabido que o tempo de ocorrência pode ser bem aproximado por uma lei exponencial com parâmetro dado pela medida da palavra, no caso em que a palava não possui sobreposição com ela mesma, ou quando tal sobreposição é "pequena". Mas, quando a palavra apresenta sobreposição (que não seja "pequena") com ela mesma, o parâmetro deve ser corrigido por um fator. Tal fator é a probabilidade da palavra não aparecer duas vezes consecutivas. E essa probabilidade é dada pelas propriedades de sobreposição da palavra [2, 12].

Uma situação similar acontece quado consideramos o tempo decorrido até a primeira ocorrência da palavra, colocando como condição inicial a ocorrência da própria palavra. Este tempo é chamado de *Tempo de Retorno (Return Time)*. Neste caso, acontece o mesmo. É sabido que o tempo de retorno também pode ser bem aproximado por uma lei exponencial. Mas quando a palavra apresenta sobreposição com ela mesma, essa lei deixa de ser exponencial, passando a ser uma combinação convexa de uma medida degenerada na origem e uma exponencial [9]. Neste caso, o mesmo fator citado no parágrafo anterior aparece na combinação convexa e na lei exponencial.

Até onde sabemos, os primeiros a observarem que a medida de todas as palavras que possuem "grandes"sobreposições converge para zero foram Collet, Galves & Schmitt, em 1999 [9]. Nesse trabalho, os autores provaram o decaimento exponencial dessa medida quando "grande"significa maior ou igual a $n - n/3$. Tal resultado vale para *processos misturadores* com função ψ de decaimento exponencial. Mais tarde, o mesmo resultado foi generalizado em [2] para processos misturadores com função ϕ de decaimento exponencial. No último caso, "grande"significa ser maior ou igual a $n - cn$, onde c é uma constante que depende da cardinalidade do alfabeto.

A distribuição de τ_n é desconhecida. Ou seja, não há uma forma explícita para

$$F_{\tau_n}(y) := \mu(\tau_n \leq y).$$

Em 2002, usando a *Complexidade de Kolmogorov*, Saussol, Troubetzkoy e Vaienti provaram, para um processo ergódico com entropia positiva, que $\lim_{n\to\infty} \frac{\tau_n}{n} = c \geq 1$ quase certamente [24].

Em 2003, usando o *Teorema de Shannon, Mc-Millan e Breiman*, Afraimovich, Chazottes e Saussol provaram que vale o mesmo resultado ($\tau_n/n \to c \geq 1 \quad q.c.$) [8].

Abadi e Vaienti, em 2008, encontraram a função de grandes desvios e relacionaram τ_n com a entropia de Rényi do processo gerador das palavras.

Sendo assim, nosso trabalho tem uma ambição interessante no que diz respeito a *Sistemas Dinâmicos*. Essa ambição consiste no fato de que um conhecimento sobre τ_n implica em conhecer também a *Entropia de Rényi* do processo.

Para o nosso contexto, o resultado apresentado em [8] e [24] diz que o conjunto de palavras que encaixam em tempo finito assintoticamente tem medida nula.

Note que, para τ_n (ver 1.2.1)temos:

$$\tau_n \in \{1, 2, \cdots, n\}$$

E isso implica que:

$$\frac{\tau_n}{n} \leq 1$$

Então, para nós, tal resultado se resume a:

$$\lim_{n \to \infty} \frac{\tau_n}{n} = 1 \qquad \mu - q.c.$$

A expressão acima diz que, assintoticamente, τ_n cresce na mesma velocidade que n. Isto nos sugere que a variável aleatória τ_n apresenta comportamento assintótico defectivo. Ou seja, temos então uma sequência de variáveis aleatórias não *estocasticamente limitadas*. Aí vemos a necessidade de definir convenientemente uma outra sequência de variáveis aleatórias que nos forneça informação sobre o *Tempo de Encaixe*. Definimos, portanto, S_n (ver Definição 2.1.1). A seguir, para uma medida produto, provamos a convergência da distribuição de S_n para uma distribuição limite, cuja cauda decai exponencialmente quando n diverge. Além disso, apresentamos um limite superior para a velocidade dessa convergência. Tal limite superior é não-uniforme. Ele depende do vetor de medidas p com marginais identicamente distribuídas, do tamanho do alfabeto χ, e de k, onde k é um inteiro positivo.

Essa convergência não-uniforme nos permite obter como corolário a convergência, em particular, da esperança, e em geral, de todos os momentos de ordem polinomial de S_n.

Além disso, obtemos limitantes inferior e superior para $E(S_n)$.

Também apresentaremos limitantes inferior e superior para a esperança de S_n.

Essa dissertação é organizada da seguinte maneira:

No Capítulo 1, apresentaremos o problema, daremos as definições iniciais e algumas ferramentas que iremos usar.

No Capítulo 2, provaremos a existência da distribuição limite.

No Capítulo 3, exibiremos limites inferior e superior para $E(S_n)$.

No Apêndice, comentaremos sobre alguns problemas em aberto.

Capítulo 1

Palavras e o Tempo de Encaixe

Quando falamos sobre *Palavras*, várias definições e exemplos podem vir à cabeça.

Se fôssemos a diversos pontos distintos da comunidade científica, e perguntássemos qual a definição de *palavra*, teríamos uma enriquecedora gama de respostas.

Um *Linguista* poderia citar a origem do nome *Palavra* (do latim *"parabola"*, que por sua vez deriva do grego *"parabole"*). Depois disso, comentaria sobre a diferença entre *termo*(forma escrita) e *vocábulo*(forma falada), com a intenção de tornar ainda mais clara a definição. Finalmente, ele definiria *palavra* como sendo um conjunto de letras de um alfabeto.

Já um *Cientista da Computação* poderia defini-la como sendo a unidade de informação usada por um tipo particular de computador. Motivado pelo seu interesse, ele daria alguns exemplos, falaria de *comprimento de palavras de cada máquina(bits)*, de expansão binária de números decimais, entre outras coisas. Assim, ele te deixaria mais bem-informado a respeito da definição de *palavra*.

Para um *Matemático*, essa tende a ser uma pergunta mais simples. O motivo dessa simplicidade é que, para nós, independente de seu significado ou aplicabilidade, uma *palavra* é, antes de nada (e nada além de), uma sequência(finita ou infinita) de elementos de um conjunto(finito ou infinito). Portanto, para o matemático são também *palavras* a *palavra* do *Linguista*, a do *Cientista da Computação*, as sequências de *RNA* em *Biologia*, e, naturalmente, as sequências de números.

A seguir, daremos uma definição matemática de *palavra*, tentando esclarecer a forma como enxergamos a matéria-prima essencial do presente trabalho.

1.1 Definições e notação

Definição 1.1.1. *Seja* χ *um conjunto finito com cardinal denotado por* $|\chi|$*, ao qual chamaremos de* Alfabeto*. Diremos que uma* palavra *é uma sequência finita de símbolos desse alfabeto.*

Formalmente:

$$x_1^n = (a_1 a_2 \cdots a_n) \quad a_i \in \chi, i = 1, 2, \cdots, n.$$

Onde os índices 1 *e* n *em* x *indicam onde começa e onde termina essa palavra de* n *símbolos.*

O conjunto de todas as palavras com símbolos do alfabeto χ será denotado por χ^n.

Em *Sistemas Dinâmicos*, essa definição é equivalente à definição de *cilindro*.

No presente texto, usaremos a notação x ou w para tratar de palavras. No nosso trabalho, consideramos apenas as palavras geradas por uma sequência de variáveis aleatórias independentes e indentimamente distribuídas no alfabeto χ.

Exemplo 1.1.2. 010100 é uma palavra de tamanho $n = 6$ no alfabeto $\chi = \{0, 1\}$. $\square$

Aqui, consideraremos sempre um alfabeto χ tal que $2 \leq |\chi|$, com $\mu(a_i) > 0, \forall a_i \in \chi$. As palavras serã formadas então por uma realização de uma sequência de variáveis aleatórias independente e identicamente distribuídas em χ.

Na seção seguinte, trataremos do *Tempo de Encaixe*, que é uma definição essencial do presente trabalho.

1.2 O Tempo de Encaixe

Definição 1.2.1. *Seja* $x_1^n \in \chi^n$*. Dizemos que o* Tempo de Encaixe *de* x_1^n *é a variável aleatória* $\tau_n : \chi^n \longrightarrow \{1, 2, \cdots, n\}$ *tal que:*

$$\tau_n(x_1^n) = inf\left\{k > 0 : x_1^{n-k} = x_{k+1}^n\right\}$$

Em palavras: dizemos que o *Tempo de Encaixe* de uma palavra é o tempo que ela leva para realizar um encaixe por sobreposição com ela mesma, deslocando, em cada unidade de tempo, a palavra em uma unidade para a direita. Em termos mais simples, é o número de translações à direita necessárias, feitas em uma cópia da palavra, para que ela realize uma sobreposição com a original.

Quando a condição acima não é satisfeita para nenhum k tal que $0 < k < n$, diremos simplesmente que o tempo de encaixe é igual ao tamanho da palavra. Ou seja:

$$\tau_n(x_1^n) = n$$

Essa definição fica mais clara nos dois exemplos seguintes.

Exemplo 1.2.2. Seja a palavra $x_1^6 = 010101$:

$$\begin{array}{lcc:cccc:cc}
\textit{Início}: & 0 & 1 & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & & \\
1^o Passo: & & 0 & 1 & 0 & 1 & 0 & 1 & \\
2^o Passo: & & & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & 0 & 1
\end{array}$$

Portanto, como a palavra realizou um encaixe no segundo passo (ou segunda translação), temos que o tempo de encaixe de x_1^6 é igual a 2. $\square$

Exemplo 1.2.3. Seja a palavra $x_1^4 = 0111$:

$$
\begin{array}{lccccccc}
\mathit{Início}: & 0 & 1 & 1 & 1 & & & \\
1^o Passo: & & 0 & 1 & 1 & 1 & & \\
2^o Passo: & & & 0 & 1 & 1 & 1 & \\
3^o Passo: & & & & 0 & 1 & 1 & 1 \\
4^o Passo: & & & & & 0 & 1 & 1 \quad 1
\end{array}
$$

Portanto, como a palavra não realizou encaixe, temos que o tempo de encaixe de x_1^4 é igual a 4, que é o tamanho da palavra. □

1.3 Definição dinâmica

Considere x como sendo um processo qualquer (ou uma palavra infinita) num alfabeto finito χ. Diremos que $x \in \chi^{\mathbb{N}}$. Então, podemos escrevê-lo da seguinte maneira:

$$x = x_1^n x_{n+1}^{\infty}.$$

Note que $x_1^n \in \chi^n$ é a parte finita, e $x_{n+1}^{\infty} \in \chi^{\mathbb{N}}$ é a parte infinita de x, que foram concatenadas, gerando x.

Defina agora uma translação T da seguinte maneira:

$$T : \chi^{\mathbb{N}} \longrightarrow \chi^{\mathbb{N}}$$

$$T(x) = T(x_1^n x_{n+1}^{\infty}) = x_2^n x_{n+1}^{\infty} = x_2^{\infty}$$

onde T retira da palavra a primeira letra. Aplicando recursivamente T, temos que para $1 \leq m$:

$$T^m(x) = x_{m+1}^{\infty}.$$

Então, finalmente, definimos $\tau_n(x)$ da seguinte maneira:

$$\tau_n(x) = \inf_{\{y \in \chi^{\mathbb{N}} : y_1^n = x_1^n\}} \{m \geq 1 : T^m(y) = y\},$$

onde o ínfimo acima é tomado entre todos as palavras y de tamanho infinito, e que tenham os mesmos n primeiros símbolos de x.

É claro que y sempre pode ser tomado de maneira que $1 \leq \tau_n \leq n$.

Note que essa definição de *Tempo de Retorno* é equivalente a de *Tempo de Encaixe*.

Vejamos um exemplo:

Exemplo 1.3.1. Seja a palavra $x = x_1^6 x_7^{\infty}$, onde $x_1^6 = 010101$. Então, se tomarmos $y = 010101....$, então teremos que $T^2(y) = y$, e portanto $\tau_n(x) = 2$.

O resultado acima coincide com o caso do exemplo 1.2 $\square$

1.4 Classificação periódica das palavras

Uma forma de entender melhor o problema do tempo de encaixe é tentar identificar algum comportamento periódioco nas palavras. Note que, quando temos uma palavra formada por "bloquinhos"que se repetem, temos um encaixe iminente. A seguir, definiremos uma ferramenta fundamental deste trabalho:

Definição 1.4.1. *Seja $0 < k < n$. Definimos $B_n(k)$ como sendo o conjunto de todas as palavras de χ^n que são k-periódicas. Ou seja:*

$$B_n(k) = \{x_1^n \in \chi^n : x_1^n = (\underbrace{x_1, \ldots, x_k}_{1}, \underbrace{x_1, \ldots, x_k}_{2}, \ldots, \underbrace{x_1, \ldots, x_k}_{\lfloor n/k \rfloor}, \underbrace{x_1, \ldots, x_r}_{1})\}$$

Portanto, o conjunto $B_n(k)$ pode ser visto como o conjunto de todas as palavras que podem ser escritas como concatenação de $\lfloor n/k \rfloor$ palavras $x_1^k \in \chi^k$ com uma palavra

$x_1^r \in \chi^r$. Aqui, sempre consideramos $r < k$, o resto da divisão de n por k. Note que o tamanho da palavra não precisa necessariamente ser um múltiplo de k. Somente a palavra precisa apresentar certo "comportamento periódico".

Falando ede forma mais simples, a palavra é formada por "bloquinhos"de tamanho k.

Exemplo 1.4.2. Tome o conjunto de palavras de tamanho $n = 4$. Denotaremos esse conjunto por C_4. Daí:

$C_4 = \{0000; 0001; 0010; 0011; 0100; 0101; 0110; 0111; 1000; 1001; 1010; 1011;$
$1100; 1101; 1110; 1111\}$

Além disso:

$B_4(1) = \{0000; 1111\}$

$B_4(2) = \{0000; 1111; 0101; 1010\}$

$B_4(3) = \{0000; 1111; 0010; 0100; 0110; 1001; 1011; 1101\}$

$B_4(4) = C_4$ $\square$

Capítulo 2

A distribuição de $n - \tau_n$

No presente capítulo, iremos definir uma nova variável aleatória, que determina o tamanho máximo de um encaixe que a palavra pode fazer com uma cópia dela mesma. Para palavras geradas por sequências de variáveis aleatórias independentes no alfabeto χ, com uma medida produto, provaremos que essa variável aleatória possui distribuição limite.

2.1 Definindo S_n

O Teorema apresentado em [8, 24] nos diz que, a menos de um conjunto de medida nula, τ_n cresce na mesma velocidade que n, . Então, $\forall x > 0$ fixado, temos que $\mu(\tau_n < x) \to 0$, quando n diverge.

Definindo $\mu_n(x) = \mu(\tau_n \leq x) = F_{\tau_n}(x)$, temos então que $\{\tau_n\}_{n\in\mathbb{N}}$ não forma uma sequência de variáveis aleatórias *estocasticamente limitadas*, e portanto $\{\mu_n\}_{n\in\mathbb{N}}$ não é *uniformemente compacta.*

Isso nos leva então à nossa próxima definição:

Definição 2.1.1. *Seja*

$$\tau_n : \chi^n \longrightarrow \{1, 2, \cdots, n\}$$

como definido no texto.

Definimos então uma nova sequência $S_n : \chi^n \longrightarrow \{0, 1, \cdots, n-1\}$, *da seguinte*

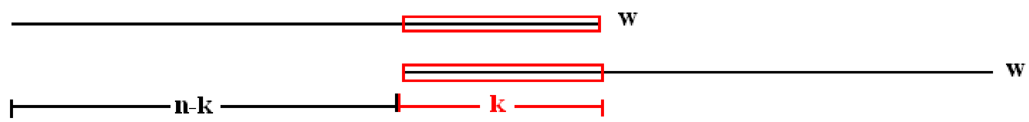


Figura 2.1: $\tau_n(w) = n - k$ é o tempo que a palavra leva para encaixar, enquanto $S_n(w) = k$ é o tamanho desse encaixe.

maneira:

$$S_n(w) = n - \tau_n(w).$$

Uma vez que τ_n é o número mínimo de translações necessárias para que ocorra um encaixe, S_n define o maior tamanho de uma sobreposição.

A figura 2.1 torna clara a relação entre τ_n e S_n.

Vamos agora dar um corolário do teorema apresentado em [8] e [24], que diz respeito a S_n:

Corolário 2.1.2. *Seja $S_n : \chi^{\mathbb{N}} \longrightarrow \{0, 1, \cdots, n-1\}$ como definido no texto. Então:*

$$\lim_{n\to\infty} \frac{S_n}{n} = 0 \qquad \mu - q.c.$$

Demonstração.

$$\lim_{n\to\infty} \frac{S_n}{n} = \lim_{n\to\infty} \frac{n - \tau_n}{n}$$

$$= 1 - \lim_{n\to\infty} \frac{\tau_n}{n} = 0.$$

Onde a última parte se deve ao teorema apresentado em [8] e [24], que foi comentado na introdução. □

No Capítulo 3, mostraremos que $\limsup_{n\to\infty} E(S_n) < +\infty$. Usando um resultado que pode ser encontrado em [19], temos que que a sequência $\{\mu_{S_n}\}_{n\in\mathbb{N}}$ é *uniformemente compacta*.

No presente capítulo, definiremos os conjuntos $R_n(k) := B_n(n-k)$, que serão as palavras com uma sobreposição de tamanho k. O motivo para essa mudança será mera conveniência. A primeira notação faz menção direta ao que foi feito em [1]. A segunda se refere à simbologia sugerida por André Toom e Andréa Vanessa Rocha em [23].

2.2 Ferramentas auxiliares

Definição 2.2.1. *Definimos o conjunto $R_n(k)$ como sendo o conjunto de palavras que tem uma sobreposição de tamanho k. Ou seja:*

$$R_n(k) = \{x_1^n \in \chi^n \quad : \quad x_1^k = x_{n-k}^n\}.$$

Note que vale a seguinte "dualidade":

$$R_n(k) = B_n(n-k).$$

Para $0 < k < n$, seja:

$$\bigcup_{j=1}^{k} R_n(j)$$

o conjunto de palavras de tamanho n com alguma sobreposição de tamanho $s \leq k$.

Seja $\bigcup_{j=1}^{n-1} R_n(j)$ o conjunto das palavras que possuem alguma sobreposição.

Definição 2.2.2. *Se para cada $a_i \in \chi$ temos $\mu(a_i) = p_i$, então defina:*

$$m_\ell = \sum_{i=1}^{|\chi|} p_i^\ell,$$

$$\rho = \max_i p_i,$$

motivados pela definição de norma $\mathcal{L}_p$ e norma do sup, *respectivamente.*

Note que $\rho^\ell \leq m_\ell$, pois como $\rho \in \{p_1, p_2, \cdots, p_{|\chi|}\}$, então $\rho = p_k$, para algum $k \in \{1, 2, \cdots, |\chi|\}$. Portanto:

$$m_\ell = \rho^\ell + \sum_{i \neq k} p_i^\ell \geq \rho^\ell.$$

O lema a seguir nos dá condições para que possamos "encolher" uma palavra de tamanho n, tirando dela as duas letras centrais no caso em que n for par, e as três letras centrais no caso em que n for ímpar.

Lema 2.2.3. *Seja $k \leq \lfloor \frac{n}{2} \rfloor - 1$. Então:*

$$\mu\left(\bigcup_{j=k}^{\lfloor \frac{n}{2} \rfloor - 1} R_n(j)\right) = \mu\left(\bigcup_{j=k}^{\lfloor \frac{n}{2} \rfloor - 1} R_{2(\lfloor \frac{n}{2} \rfloor - 1)}(j)\right).$$

Demonstração. $x_1^n \in \bigcup_{j=k}^{\lfloor \frac{n}{2} \rfloor - 1} R_n(j)$ se e somente se $x_1^n \in R_n(j)$, para algum j tal que $k \leq j \leq \lfloor \frac{n}{2} \rfloor - 1$. Logo

$$x_1^n = w_1 w_2 w_1,$$

onde w_1 é uma palavra de tamanho j, w_2 é uma palavra de tamanho $2n - j$, e elas são independentes. Agora escrevemos $w_2 = w_{2,1} w_{2,2} w_{2,3}$, onde $w_{2,2}$ é a palavra central de w_2, e que tem tamanho 2 no caso onde n é ímpar ou tamanho 3 no caso em que n é par. Ela pode ser escrita da seguinte forma:

$$w_{2,2} = x_{\lfloor \frac{n}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor + 1},$$

com $w_{2,1}$ e $w_{2,3}$ palavras de tamanho $\lfloor \frac{n-2j}{2} \rfloor - 1$. Agora, defina $\tilde{w} = w_1 w_{2,1} w_{2,3} w_1 \in R_{2\lfloor \frac{n}{2} \rfloor - 1}(j)$, que é independente de $w_{2,2}$. Portanto:

$$\mu\left(\bigcup_{j=k}^{\lfloor \frac{n}{2} \rfloor - 1} R_n(j)\right) = \sum_{x_1^n \in R_n(j)} \mu(x_1^n) = \sum_{w_1 w_{2,1} \in \chi^{n-2}} \sum_{w_{2,2} \in \chi^i} \mu(w_1 w_{2,1} w_{2,1} w_1) \mu(w_{2,2}).$$

Somando independentemente cada termo, o primeiro termo resulta em $R_{2\lfloor \frac{n}{2} \rfloor - 1}(j)$, e o segundo termo soma 1, pois se refere a todas as palavras $w_{2,2}$ possíveis.

$\square$

O próximo lema nos diz que a medida total do conjunto de palavras de tamanho n com uma sobreposição grande converge exponencialmente para zero.

Lema 2.2.4. *Vale a seguinte desigualdade:*

$$\mu\left(\bigcup_{j=\lfloor \frac{n}{2} \rfloor}^{n-1} R_n(j)\right) \leq \frac{\left(m_2^{\frac{1}{2}}\right)^{\lceil \frac{n}{2} \rceil}}{1 - m_2^{\frac{1}{2}}}$$

Demonstração. Usando a σ-subaditividade de μ e a "dualidade"

$$R_n(j) = B_n(n-j),$$

temos a seguinte expressão:

$$\mu\left(\bigcup_{j=\lfloor \frac{n}{2} \rfloor}^{n-1} R_n(j)\right) \leq \sum_{j=\lfloor \frac{n}{2} \rfloor}^{n-1} \mu\left(R_n(j)\right) = \sum_{j=1}^{\lceil \frac{n}{2} \rceil} \mu(B_n(j)). \tag{2.1}$$

Agora, note que, se $w \in B_n(j)$, então temos a relação $n = j\lfloor \frac{n}{j} \rfloor + r$, onde $0 \leq r < j$, e r indica o resto da divisão de n por j. Portanto:

$$w = \underbrace{w_j w_j \cdots w_j}_{\lfloor \frac{n}{j} \rfloor \quad vezes} w_r \qquad w_j \in \chi^j, \quad w_r \in \chi^r.$$

Além disso, temos também que:

$$w_j = w_r w_{j-r}$$

onde w_{j-r} é uma palavra em χ^{j-r}. Portanto:

$$\mu(B_n(j)) = \sum_{w_j \in \chi^j} \left(\mu(w_j)\right)^{\lfloor \frac{n}{j} \rfloor} \mu(w_r) = \sum_{w_r \in \chi^r} \left(\mu(w_r)\right)^{\lfloor \frac{n}{j} \rfloor + 1} \sum_{w_{j-r} \in \chi^{j-r}} \left(\mu(w_{j-r})\right)^{\lfloor \frac{n}{j} \rfloor}.$$

Pela definição de m_ℓ, temos que o último termo da expressão acima é igual a:

$$m^r_{\lfloor \frac{n}{j} \rfloor + 1} m^{j-r}_{\lfloor \frac{n}{j} \rfloor}.$$

Temos também que vale:

$$m^r_{\lfloor \frac{n}{j} \rfloor + 1} m^{j-r}_{\lfloor \frac{n}{j} \rfloor} \leq m^j_{\lfloor \frac{n}{j} \rfloor} = \left(m^{\frac{1}{\lfloor \frac{n}{j} \rfloor}}_{\lfloor \frac{n}{j} \rfloor} \right)^{\lfloor \frac{n}{j} \rfloor j} \leq \left(m^{\frac{1}{\lfloor \frac{n}{j} \rfloor}}_{\lfloor \frac{n}{j} \rfloor} \right)^{n-j}.$$

Por uma desigualdade clássica em norma $\mathcal{L}_p$, que pode ser vista em [19], temos que, para $1 \leq j \leq \frac{n}{2}$ vale:

$$m^{\frac{1}{\lfloor \frac{n}{j} \rfloor}}_{\lfloor \frac{n}{j} \rfloor} \leq m_2^{\frac{1}{2}}.$$

Portanto, a última parte da soma 2.1 satisfaz:

$$\sum_{j=1}^{\lceil \frac{n}{2} \rceil} \left(m_2^{\frac{1}{2}} \right)^{n-j} = \frac{\left(m_2^{\frac{1}{2}} \right)^{n-1} - \left(m_2^{\frac{1}{2}} \right)^{\lceil \frac{n}{2} \rceil + 1}}{1 - m_2^{\frac{1}{2}}} \leq \frac{\left(m_2^{\frac{1}{2}} \right)^{n-1}}{1 - m_2^{\frac{1}{2}}}$$

E assim, o lema está provado. $\square$

O lema a seguir nos fornece a taxa exponencial para a convergência, que usaremos para provar o Teorema principal deste trabalho.

Lema 2.2.5. *Valem as seguintes desigualdades:*

a) $\displaystyle\sum_{i=\frac{n}{2}+1}^{\infty} \mu\left(R_{2i}(i) \cap R_{2i}(k)\right) \leq \left(\frac{m_4}{m_2^2} \right)^k \frac{m_2^{\frac{n}{2}+1}}{1 - m_2}$

b) $\displaystyle\sum_{j=\frac{n}{2}}^{n-1} \left(R_n(j) \cap R_n(k)\right) \leq C \epsilon^n \delta^k,$

onde C, $\epsilon < 1$, e $\delta < 1$ são constantes positivas que dependem do vetor p.

Demonstração. a) Uma palavra x_1^n pertence a $R_{2i}(i) \bigcap R_{2i}(k)$ se e somente se ela tiver a forma:

$$x_1^n = w_1 w_2 w_1 w_1 w_2 w_1$$

onde w_1 é uma palavra de tamanho k e w_2 é uma palavra de tamanho $i-2k$ (que nesse caso é igual a $\frac{n-4k}{2}$). Portanto:

$$\mu(x_1^6) = (\mu(w_1))^4 (\mu(w_2))^2,$$

e então:

$$\mu(R_n(j) \cap R_n(k)) = \sum_{w_1 \in \chi^k} (\mu(w_1))^4 \sum_{w_2 \in \chi^{i-2k}} (\mu(w_2))^2 = m_4^k m_2^{i-2k}.$$

Agora, vamos somar o resultado em i:

$$\sum_{i=\frac{n}{2}+1}^{\infty} m_4^k m_2^{i-2k} = \frac{m_4^k m_2^{\frac{n}{2}+1-2k} - \lim_{i\to\infty} m_4^k m_2^{i+1-2k}}{1-m_2} = \frac{m_4^k m_2^{\frac{n}{2}+1-2k}}{1-m_2}.$$

b) Pela "dualidade"(2.2.1), a soma no item b) é igual a:

$$\sum_{j=1}^{\frac{n}{2}} \mu(B_n(j) \cap R_n(k))$$

$$= \underbrace{\sum_{j=1}^{\frac{n}{2}-k} \mu(B_n(j) \cap R_n(k))}_{i)} + \underbrace{\sum_{j=\frac{n}{2}-k+1}^{\frac{n-k}{2}} \mu(B_n(j) \cap R_n(k))}_{ii)} + \underbrace{\sum_{j=\frac{n-k}{2}+1}^{\frac{n}{2}} \mu(B_n(j) \cap R_n(k))}_{iii)}.$$

Agora, vamos resolver separadamente cada caso.

i) Suponha que $w \in (B_n(j) \cap R_n(k))$. Como $w \in R_n(k)$, temos que w pode ser escrita na forma

$$w = w_1 w_2 w_1,$$

onde $w_1 \in \chi^k$ e $w_2 \in \chi^{n-2k}$.

Como $w \in B_n(j)$ com $1 \leq j \leq \frac{n}{2} - k$, ela consiste na concatenação de blocos de tamanho j. Além disso, como $B_n(k)$ define uma propriedade periódica das palavras, esses blocos podem ser observados começando de qualquer posição da palavra. Logo, temos que:

$$w_2 \in B_{n-2k}(j).$$

Portanto:

$$\mu(B_n(j) \cap R_n(k)) \leq \sum_{w_1 \in \chi^k} (\mu(w_1))^2 \mu(B_{n-2k}(j)).$$

Usando um argumento combinatório, temos que vale:

$$\sum_{w_1 \in \chi^k} (\mu(w_1))^2 = m_2^k.$$

E, pelo lema 2.2.4, temos:

$$\sum_{j=1}^{\frac{n}{2}-k} \mu(B_{n-2k}(j)) \leq \frac{(m_2^{\frac{1}{2}})^{\lceil \frac{n}{2} \rceil}}{1 - m_2^{\frac{1}{2}}}.$$

ii) Como $2j \leq n - k$ e $w \in B_n(j)$, podemos escrever:

$$w = w_1 w_2 w_1 w_2 \tilde{w_1} w_1,$$

onde w_1 tem tamanho k e w_2 tem tamanho $j - k$. Como $w \in R_n(k)$, primeiro w_1 da concatenação acima aparece. Além disso, $\tilde{w_1}$ tem tamanho $n - k - 2j$. Portanto, usando a independência e fatorizando a medida nós temos:

$$\mu(w) = (\mu(w_1))^3 (\mu(w_2))^2 \mu(\tilde{w_1}).$$

Então:

$$\mu(B_n(j) \cap R_n(k)) \leq \sum_{w_1 \in \chi^k} (\mu(w_1))^3 \sum_{w_1 \in \chi^{j-k}} (\mu(w_2))^2 \rho^{n-k-2j} = m_3^k m_2^{j-k} \rho^{n-k-2j}.$$

Somando ó último termo em j, nós temos:

$$\sum_{j=\frac{n}{2}-k+1}^{\frac{n-k}{2}} \mu(R_n(j) \cap R_n(k)) \leq C(p) m_2^{\frac{n}{2}} \left(\frac{m_3}{m_2^{\frac{3}{2}}} \right)^k .$$

Como $m_l^{\frac{1}{l}}$ é a norma $\mathcal{L}_p$ do vetor p, então temos que a base do último termo da expressão acima é menor que 1.

iii) Tome $w = x_1^n \in B_n(j) \cap R_n(k)$.

Observe que:

- Como $w \in R_n(k)$, temos $x_1^k = x_{n-k+1}^n$.
- Como os blocos podem ser “lidos”tanto de trás para frente quanto da frente para trás, nós temos que $x_{n-k+1}^n = x_{n-j-k}^{n-j-1}$.

Portanto, podemos escrever:

$$w = w_1 w_2 w_1 w_2 \tilde{w}_1 w_1.$$

O tamanho de w_1 é k, o de w_2 é $2j + k - n$, e o tamanho de $\tilde{w}_1$ é $n - 2k - j$.

Novamente utilizando a independência, podemos fatorar a medida e obter:

$$\mu(B_n(j) \cap R_n(k)) \leq \sum_{w_1 \in \chi^k} (\mu(w_1))^3 \sum_{w_2 \in \chi^{j-k}} (\mu(w_2))^2 \rho^{2j+k-n} = m_3^k m_2^{n-2k-j} \rho^{2j+k-n}. \quad (2.2)$$

Somando o último termo em j, obtemos:

$$\sum_{j=\frac{n-k}{2}+1}^{\frac{n}{2}} \mu(R_n(j) \cap R_n(k)) \leq C'(p) m_2^{\frac{n}{2}} \left(\frac{m_3}{m_2^{\frac{3}{2}}} \right)^k .$$

E assim terminamos a demonstração. $\square$

2.3 O Teorema principal

Finalmente, chegamos ao principal teorema do presente texto.

Esse teorema diz, essencialmente, existe uma distribuição limite para S_n, e que a convergência para essa distribuição limite ocorre com taxa exponencial.

Primeiro, vamos definir alguns termos que aparecerão no enunciado:

Sejam:

$$a_{n,k} = \mu\left(\bigcup_{j=\frac{n}{2}}^{n-1} R_n(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-1} R_n(j)\right) + \sum_{i=k+1}^{\frac{n}{2}-1} \mu\left(R_{2i}(i) \backslash \bigcup_{j=k}^{i-1} R_{2i}(j)\right),$$

$$b_{n,k} = \mu\left(\bigcup_{j=\frac{n}{2}}^{n-1} R_n(j) \cap R_n(k) \backslash \bigcup_{j=k+1}^{\frac{n}{2}-1} R_n(j)\right) + \sum_{i=k+1}^{\frac{n}{2}} \mu\left(R_{2i}(i) \cap R_{2i}(k) \backslash \bigcup_{j=k+1}^{i-1} R_{2i}(j)\right),$$

$$a_k = \sum_{i=k+1}^{\infty} \mu\left(R_{2i}(i) \backslash \cup_{j=k}\, i-1 R_{2i}(j)\right),$$

$$b_k = \sum_{i=k+1}^{\infty} \mu\left(R_{2i}(i) \cap R_{2i}(k) \backslash \cup_{j=k+1}^{i-1} R_{2i}(j)\right).$$

Então:

Teorema 2.3.1. *Seja μ uma medida produto em χ^n com marginais identicamente distribuídas. Então:*

a) $\mu(S_n \geq k) = m_2^k + a_{k,n}$.

b) $\mu(S_n = k) = m_2^k - b_{k,n}$.

c) $\lim_{n\to\infty} \mu(S_n \geq k) = m_2^k + a_k$.

d) $\lim_{n\to\infty} \mu(S_n = k) = m_2^k - b_k$.

Além disso, temos que existem constantes positiva c, $\epsilon < 1$ e $\delta < 1$, que dependem do vetor p de probabilidade dos símbolos do alfabeto, tais que:

$$\left|\mu(S_n = k) - (m_2^k - b_k)\right| \leq c\epsilon^n \delta^k.$$

Demonstração. Tome:

$$G_n(k) = \mu(S_n \geq k) = \mu\left(\bigcup_{j=k}^{n-1} R_n(j)\right).$$

Podemos supor, sem perda de generalidade, que $k < \dfrac{n}{2}$.

A princípio, consideraremos o caso em que n é ímpar.

Fazendo a decomposição de $\bigcup_{j=k}^{n-1} R_n(j)$:

$$G_n(k) = \mu\left(\bigcup_{j=k}^{n-1} R_n(j)\right) = \mu\left(\bigcup_{j=k}^{\frac{n}{2}-1} R_n(j)\right) + \mu\left(\bigcup_{j=\frac{n}{2}}^{n-1} R_n(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-1} R_n(j)\right).$$

Pelo Lema 2.2.3, o primeiro termo do último lado da igualdade é igual a

$$\mu\left(\bigcup_{j=k}^{\frac{n}{2}-1} R_{2(\lfloor\frac{n}{2}\rfloor-1)}(k)\right).$$

Podemos decompor este último termo, chegando a

$$\mu\left(\bigcup_{j=k}^{(n-2)-1} R_{n-2}(j)\right) - \mu\left(\bigcup_{j=\frac{n}{2}}^{(n-2)-1} R_{n-2}(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-1} R_{n-2}(j)\right).$$

O termo da esquerda é, por definição, $G_{n-2}(k)$. Portanto, concluímos que:

$$\begin{aligned} G_n(k) &= G_{n-2}(k) \\ &+ \mu\left(\bigcup_{j=\frac{n}{2}}^{n-1} R_n(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-1} R_n(j)\right) \\ &- \mu\left(\bigcup_{j=\frac{n}{2}}^{(n-2)-1} R_{n-2}(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-1} R_{n-2}(j)\right). \end{aligned}$$

Repetindo a mesma operação para G_{n-2}, chegamos a:

$$
\begin{aligned}
G_{n-2}(k) &= G_{n-4}(k) \\
&+ \mu\left(\bigcup_{j=\frac{n}{2}-1}^{(n-2)-1} R_{n-2}(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-2} R_{n-2}(j)\right) \\
&- \mu\left(\bigcup_{j=\frac{n}{2}-1}^{(n-4)-1} R_{n-4}(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-2} R_{n-4}(j)\right).
\end{aligned}
$$

Portanto, temos que:

$$
\begin{aligned}
G_{n}(k) &= G_{n-4}(k) \\
&+ \mu\left(\bigcup_{j=\frac{n}{2}}^{n-1} R_{n}(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-1} R_{n}(j)\right) \\
&- \mu\left(\bigcup_{j=\frac{n}{2}}^{(n-2)-1} R_{n-2}(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-2} R_{n-2}(j)\right) \\
&+ \mu\left(\bigcup_{j=\frac{n}{2}-1}^{(n-2)-1} R_{n-2}(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-2} R_{n-2}(j)\right) \\
&- \mu\left(\bigcup_{j=\frac{n}{2}-1}^{(n-4)-1} R_{n-4}(j) \backslash \bigcup_{j=k}^{\frac{n}{2}-2} R_{n-4}(j)\right).
\end{aligned}
$$

Somando as duas linhas do meio da última expressão, temos o seguinte resultado:

$$
\mu\left(R_{n-1}(\frac{n}{2}-1) \backslash \bigcup_{j=k}^{\frac{n}{2}-2} R_{n-2}(j)\right).
$$

Um argumento recursivo aplicado em k nos fornece

$$
\begin{aligned}
G_n(k) &= G_{2k}(k) \\
&+ \mu\left(\bigcup_{j=\frac{n}{2}}^{n-1} R_n(j) \setminus \bigcup_{j=k}^{\frac{n}{2}-1} R_n(j)\right) \\
&+ \sum_{i=k+1}^{\frac{n}{2}-1} \mu\left(R_{2i}(i) \setminus \bigcup_{j=k}^{i-1} R_{2i}(j)\right) \\
&- \mu\left(\bigcup_{j=k+1}^{2k-1} R_{2k}(j) \setminus \bigcup_{j=k}^{k} R_{2k}(j)\right).
\end{aligned}
$$

Finalmente, chegamos a seguinte expressão:

$$
G_n(k) = G_{2k}(k) + \mu\left(\bigcup_{j=\frac{n}{2}}^{n-1} R_n(j) \setminus \bigcup_{j=k}^{\frac{n}{2}-1} R_n(j)\right) + \sum_{i=k+1}^{\frac{n}{2}-1} \mu\left(R_{2i}(i) \setminus \bigcup_{j=k}^{i-1} R_{2i}(j)\right). \tag{2.3}
$$

Pelo lema 2.2.4, temos que o primeiro termo do lado direito da expresão acima tende a zero quando n diverge.

Portanto, temos:

$$
\lim_{n\to\infty} G_n(k) = \mu(R_{2k}(k)) + \sum_{i=k+1}^{\frac{n}{2}-1} \mu\left(R_{2i}(i) \setminus \bigcup_{j=k}^{i-1} R_{2i}(j)\right)
$$

Por outro lado,

$$
\mu(S_n = k) = G_n(k) - G_n(k+1).
$$

Resolvendo esta equação acrescentando 2.3, concluímos que

$$\begin{aligned}
\mu(S_n = k) &= \mu(R_{2k}(k)) - \mu(R_{2(k+1)}(k+1)) \\
&- \mu\left(\bigcup_{j=\frac{n}{2}}^{n-1} R_n(j) \cap R_n(k) \backslash \bigcup_{j=k+1}^{\frac{n}{2}-1} R_n(j)\right) \\
&- \sum_{i=k+2}^{\frac{n}{2}} \mu\left(R_{2i}(i) \cap R_{2i}(k) \backslash \bigcup_{j=k+1}^{i-1} R_{2i}(j)\right) \\
&+ \mu\left(R_{2(k+1)}(k+1) \backslash \bigcup_{j=k}^{k} R_{2(k+1)(j)}\right).
\end{aligned}$$

Calculando o termo da direita na primeira linha da última expressão, o resultado é:

$$\mu\left(R_{2(k+1)}(k+1) \cap R_{2(k+1)(k)}\right).$$

Com um certo abuso de notação, considere uma união em um conjunto vazio de índices como sendo o conjunto vazio.

Teremos então o seguinte resultado:

$$\begin{aligned}
\mu(s_n = k) &= \mu(R_{2k}(k)) \qquad (2.4) \\
&- \mu\left(\bigcup_{j=\frac{n}{2}}^{n-1} R_n(j) \cap R_n(k) \backslash \bigcup_{j=k+1}^{\frac{n}{2}-1} R_n(j)\right) \qquad (2.5) \\
&- \sum_{i=k+1}^{\frac{n}{2}} \mu\left(R_{2i}(i) \cap R_{2i}(k) \backslash \bigcup_{j=k+1}^{i-1} R_{2i}(j)\right). \qquad (2.6)
\end{aligned}$$

Pelo lema 2.2.4, o termo do meio vai a zero quando n diverge. Portanto, o limite

$$q_k = \lim_{n\to\infty} \mu(S_n = k)$$

existe e é igual a:

$$\mu(R_{2k}(k)) - \sum_{i=k+1}^{\infty} \mu\left(R_{2i}(i) \cap R_{2i}(k) \backslash \bigcup_{j=k+1}^{i-1} R_{2i}(j)\right).$$

Além disso, temos como consequência que $|\mu(S_n = k) - q_k|$ é igual a:

$$\left|\sum_{i=\frac{n}{2}+1}^{\infty}\mu\left(R_{2i}(i)\cap R_{2i}(k)\backslash\bigcup_{j=k+1}^{i-1}R_{2i}(j)\right)-\mu\left(\bigcup_{j=\frac{n}{2}}^{n-1}R_n(j)\cap R_n(k)\backslash\bigcup_{j=k+1}^{\frac{n}{2}-1}R_n(j)\right)\right|.$$

Retirando convenientemente conjuntos nos complementares acima, e usando a seguinte desigualdade elementar válida para a e b positivos:

$$|a-b|\leq\max\{a,b\},$$

concluímos que:

$$|\mu(S_n=k)-q_k|\leq\max\left\{\sum_{i=\frac{n}{2}+1}^{\infty}\mu\left(R_{2i}(i)\cap R_{2i}(k)\right),\mu\left(\sum_{j=\frac{n}{2}}^{n-1}\mu(R_n(j)\cap R_n(k))\right)\right\}.$$

No caso em que n é um número par, basta repetir o mesmo argumento acima, apenas trocando $\lfloor\frac{n}{2}\rfloor$ por $\frac{n}{2}$, e a demonstração é idêntica.

Como conclusão, temos que para qualquer inteiro positivo n vale:

$$G_{2n+1}(k)=G_{2n}(k).$$

Além disso, também vale:

$$\mu(S_{2n+1}=k)=\mu(S_{2n}=k).$$

E finalmente, o Lema 2.2.5 finaliza a demonstração. $\square$

Capítulo 3

Limitantes para $E(S_n)$

Seguindo a sequência do trabalho, vamos olhar agora para a $E(S_n)$ quando n diverge.

Primeiramente, daremos alguns corolários do teorema principal .

Depois, apresentaremos limitantes para $E(S_n)$, apresentando alguns exemplos.

A partir daí, mostraremos que $\{\mu_{S_n}\}_{n\in\mathbb{N}}$ é uma sequência *uniformemente compacta.* É claro que o fato de conhecer a forma explícita de tal distribuição talvez nos daria um resultado mais direto. Mas a dificuldade (e a beleza) desse problema se encontra justamente no fato de S_n não possuir uma forma explícita simples de distribuição de probabilidade. Ou seja, não se conhece uma expressao de F_{S_n} distribuição de S_n. Posteriormente, seguindo uma sequência natural, daremos alguns resultados sobre a variância de S_n.

Corolário 3.0.2. *Sejam $w \in \chi^n$, e τ_n como definidos no texto.*

Então:

$$\lim_{n\to\infty} E\left(\frac{\tau_n}{n}\right) = 1$$

Demonstração. Por [8, 24], temos que:

$$\lim_{n\to\infty} \frac{\tau_n}{n} = 1 \quad \mu - q.c.$$

Defina: $Y_n = \dfrac{\tau_n}{n}$.

Como $\tau_n \in \{0, 1, \cdots, n\}$, temos que $\left|\frac{Y_n}{n}\right| = \frac{Y_n}{n} \leq 1$.

Portanto, pelo Teorema da Convergência Dominada, temos que:

$$\lim_{n\to\infty} \int Y_n d\mu = \int \lim_{n\to\infty} Y_n d\mu = 1$$

□

3.1 A convergência de $E(S_n)$

Como consequência direta do teorema 2.3.1, mostraremos a convergência de $E(S_n)$ para um valor Q, onde Q é a esperança de uma variável aleatória com distribuição q_k, definida sobre os inteiros não-negativos.

Corolário 3.1.1. *Seja Q a esperança de uma variávela aleatória com distribuição q_k sobre os números inteiros não-negativos. Então:*

$$|E(S_n) - Q| \leq C\epsilon^n.$$

Demonstração. Como $S_n \in \{0, 1, \cdots, n-1\}$ temos que:

$$\sum_{k=0}^{\infty} \mu(S_n = k) = \sum_{k=0}^{n-1} \mu(S_n = k).$$

Além disso, como q_k converge exponencialmente para zero, temos que kq_k é somável.

Pelo Teorema 2.3.1, existem números $0 < \epsilon < 1$ e $0 < \delta < 1$ tais que:

$$|\mu(S_n = k) - q_k| \leq \epsilon^n \delta^k$$

para todo n e k inteiro positivo, então:

$$\left|\sum_{k=0}^{\infty} \mu(S_n = k) - \sum_{k=0}^{\infty} kq_k\right| \leq \sum_{k=0}^{\infty} k|\mu(S_n = k) - q_k| \leq \sum_{k=0}^{\infty} kC\epsilon^n\delta^k.$$

Note que o último termo da expressão acima pode ser escrito da forma:

$$C\epsilon^n \sum_{k=0}^{\infty} k\delta^k = C\epsilon^n \delta \sum_{k=1}^{\infty} k\delta^{k-1} = C\epsilon^n \delta \sum_{k=1}^{\infty} \frac{\partial}{\partial \delta}\delta^k = C\epsilon^n \delta \frac{\partial}{\partial \delta} \sum_{k=1}^{\infty} \delta^k$$

$$= C\epsilon^n \delta \frac{\partial}{\partial \delta}\left(\frac{1}{1-\delta}\right) = C\epsilon^n \delta \left(\frac{1}{(1-\delta)^2}\right) = \frac{C\delta}{(1-\delta)^2}\epsilon^n.$$

Como esse último termo tende a zero quando n diverge, concluímos a demonstração.

□

3.2 Limitantes para Q

Nesta seção, vamos dar um outro corolário do Teorema 2.3.1, que nos fornece limitantes superior e inferior para $Q = \lim_{n\to\infty} E(S_n)$. A Seguir, daremos umá prévia de um dos pontos que será discutido no Apêndice do presentee trabalho.

Corolário 3.2.1. *Seja Q o valor esperado de q_k, onde $Q = Q(p)$, e p é o vetor de probabilidades dos símbolos do alfabeto χ. Então valem as seguintes desigualdades:*

$$\frac{m_2}{1-m_2} \leq Q \leq \frac{m_2}{(1-m_2)^2}.$$

Demonstração. Vamos aqui usar a seguinte identidade, que pode ser encontrada em [19]. Seja X uma variável aleatória assumindo valores inteiros positivos. Então:

$$E(X) = \sum_{k=1}^{\infty} \mu(X \geq k).$$

Usando a equação acima, temos que:

$$Q = \sum_{k=1}^{\infty} \lim_{n\to\infty} [\mu(S_n \geq k)] = \sum_{k=1}^{\infty} \left[\lim_{n\to\infty} G_n(k)\right].$$

Por **??**, temos que:

$$\mu(R_{2k}(k)) \leq \lim_{n\to\infty} G_n(k) \leq \mu(R_{2k}) + \sum_{i=k+1}^{\infty} \mu(R_{2i}(i)). \tag{3.1}$$

Usando o mesmo argumento combinatório do capítulo anterior, temos que:

$$\mu(R_{2i}(i)) = m_2 i.$$

Isto nos dá:

$$\sum_{i=k+1}^{\infty} \mu(R_{2i}(i)) = \frac{m_2^{k+1}}{1-m_2}$$

Sabemos que:

$$Q = \sum_{k=1}^{\infty} \lim_{n\to\infty} G_n(k).$$

Somando em k os termos de 3.1, chegamos a:

$$\sum_{k=1}^{\infty} m_2^k \le Q \le \sum_{k=1}^{\infty} m_2^k + \sum_{k=1}^{\infty} \frac{m_2^{k+1}}{1-m_2}$$

$$\Rightarrow \frac{m_2}{1-m_2} \le Q \le \frac{m_2}{(1-m_2)^2}.$$

E assim, o corolário está demonstrado.

$\square$

Apêndice A

Problemas em aberto

A.1 O caso Bernoulli e a razão áurea

Vamos aqui tratar do caso em que o alfabeto possui dois símbolos, digamos $\chi = \{0, 1\}$ tal que $\mu(0) = p$ e $\mu(1) = 1 - p$.

Como um caso particular do Corolário 3.2.1, temos que no caso do alfabeto com dois símbolos, e distribuição Bernoulli, vale a desigualdade:

$$\psi_1(p) = \frac{m_2}{1 - m_2} \leq Q \leq \frac{m_2}{(1 - m_2)^2} = \psi_2(p).$$

Note que, nesse caso, temos o seguinte valor para m_2:

$$m_2 = p^2 + (1 - p)^2 = 2p^2 - 2p + 1.$$

A seguir, apresentaremos alguns resultados desenvolvidos por Guilherme Ludwig, sob a orientação de Miguel Abadi.

Uma parte desse trabalho(que será aqui apresentada) constituiu em simular amostras de palavras de tamanho 250 no alfabeto $\chi = \{0, 1\}$. As probabilidades de ocorrência de cada símbolo são p e $1 - p$. Ou seja, estamos tratando do caso de uma *medida de Bernoulli.*

Em seguida, foi calculada $E(S_n)$ para cada caso.

O gráfico seguinte mostra os dados simulados e o gráfico das curvas $\psi_1(p)$ e $\psi_2(p)$, num interessante confrontamento entre o que é *determinístico* e o que é *aleatório.*

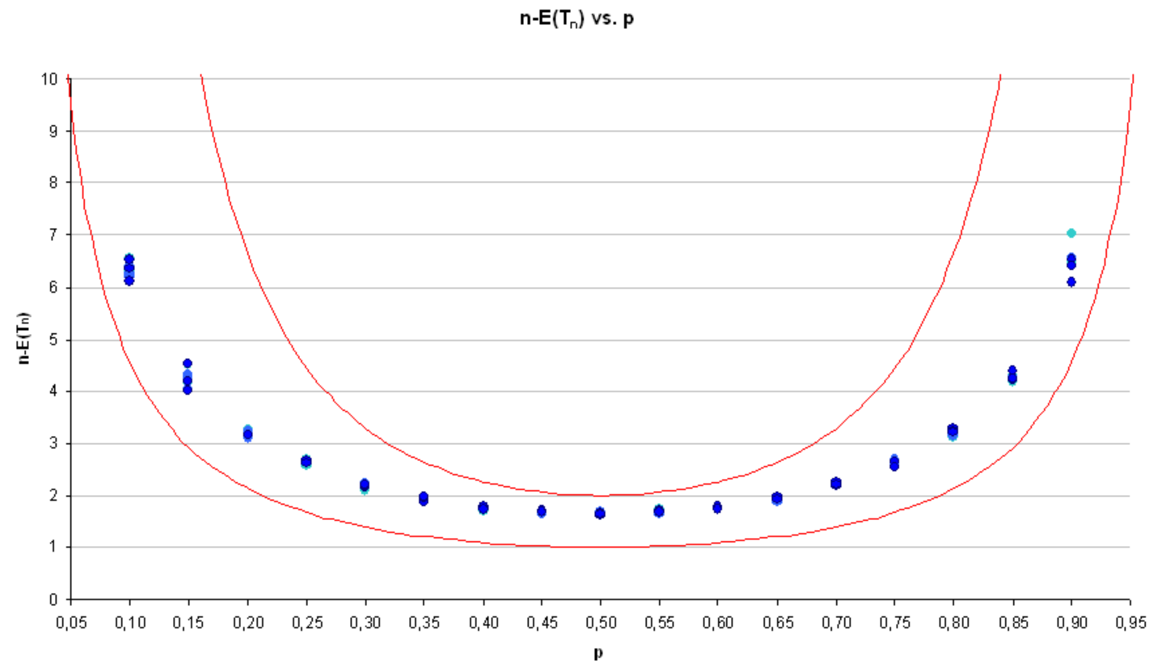


Figura A.1: Relação entre S_n e p no caso $|\chi| = 2$

Note que, quando p se aproxima de 0 ou de 1, a entropia do processo tende a zero. Nesse caso, o nível de complexidade diminiu, e as palavras tendem a realizar encaixes em tempo menor, portanto sobreposições de tamanho maior. Isso se deve ao fato de um mesmo símbolo ocorrer com probabilidade grande. Por exemplo, se $\mu(1) \to 1$ (o que implica $\mu(0) \to 0$), então as palavras terão grande probabilidade de ser do tipo (1111111111111...), e tal palavra realiza um encaixe com $\tau_n = 1$. No caso inverso, $\mu(1) \to 1$ (o que implica $\mu(0) \to 1$), acontece o mesmo com S_n.

Note que temos:

$$\lim_{p \to 0} \psi_1(p) = \lim_{p \to 1} \psi_1(p) = +\infty$$

Além disso:

$$\lim_{p \to 0} \psi_2(p) = \lim_{p \to 1} \psi_2(p) = +\infty$$

Observe também que a curva de cima (dada por $\psi_2(p)$) diverge mais rápido do que a curva de baixo (ψ_2), nos casos em que p tende a 0 e a 1.

No caso $p = 1 - p = \frac{1}{2}$, o valor de $E(S_n)$ fica entre 1 e 2. As simulações nos dizem que esse valor está próximo de $1,64$.

Uma pergunta interessante a se fazer é:

Seria a *razão áurea* um bom candidato a limite? Ou seja, seria razoável pensar que:

$$\lim_{n\to\infty} E(S_n) = \frac{1+\sqrt{5}}{2},$$

no caso em que $p = \frac{1}{2}$?

A.2 $Q(p)$ e a entropia

Outra pergunta interessante que ainda permanece sem resposta é sobre a monotonicidade de $E(S_n)$ em relação à entropia do processo.

Essencialmente, gostaríamos de poder dizer se, para uma família de processos parametrizados por:

$$p = (p_1, p_2, \cdots, p_m)$$

$$\tilde{p} = (\tilde{p}_1, \tilde{p}_2, \cdots, \tilde{p}_m)$$

seria possível fazer uma afirmação do tipo

$$H(p) \leq H(\tilde{p}) \Leftrightarrow Q(p) \leq Q(\tilde{p})?$$

Ou do tipo

$$H(p) \leq H(\tilde{p}) \Leftrightarrow Q(p) \geq Q(\tilde{p})?$$

O seguinte gráfico mostra a mesma simulação do gráfico anterior, com a diferença que a linha vermelha representa a entropia do processo.

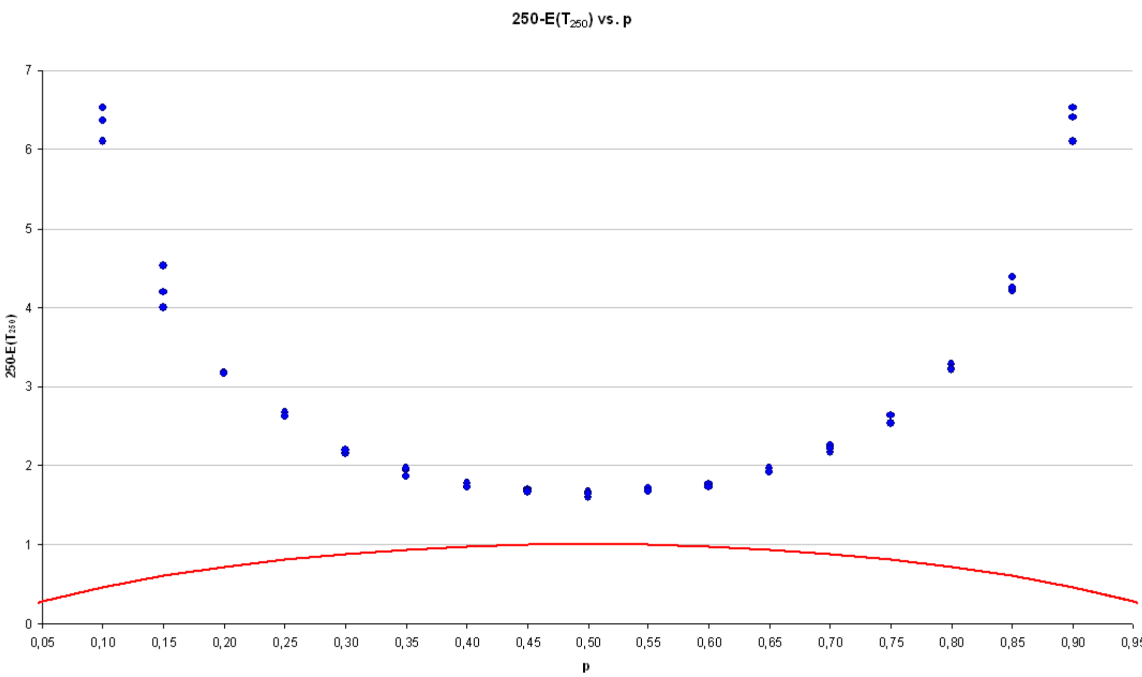


Figura A.2: A linha vermelha representa a entropia do processo, e os pontos, os dados simulados.

Referências Bibliográficas

[1] Abadi, M. & Vaienti, S.; **Large Deviations for Short Recurrence**, Discrete and Continuous Dynamical Systems - Series A (DCDS-A), Volume 21, Number 3, pp.729 – 747, 2008.

[2] Abadi, M.; **Exponential approximation for hitting times in mixing processes**, Mathematical Physics Electronic Journal (7) No.2 2001.

[3] Abadi, M.; **Sharp error therms and necessary conditions for exponencial hitting times in mixing processes**, Annals of Probability (32) 1A, (2004) 243-264.

[4] Abadi, M.: **Poisson approximations via Chen-Stein for non-Markov processes.** In and out of equilibrium. 2, 1–19, Progress in Probability, 60, Birkhauser, Basel, 2008.

[5] Abadi, M.; **Hitting, returning and the short correlation function**, Bulletin of the Brazilian Mathematical Society - (37) 4 (2006), 593-609.

[6] Abadi, M. & Saussol,B. (2010);**Hitting and returning into rare events for all alpha-mixing processes** `http://arxiv.org/PS_cache/arxiv/pdf/1003/1003.4856v1.pdf`

[7] Abadi, M. & Vergne, N. (2008) **Sharp error terms for return time statistics under mixing condition,** Journal of Theoretical Probability 21 (2008), no. 3, 729–747

[8] Afraimovich, V., Chazottes J.-R. & Saussol, B.; **Pointwise dimensions for Poincaré recurrence associated with maps and special flows**, Discrete and Continuous Dynamical Systems, Volume 9, pp.263 – 280, 2003.

[9] Collet, P., Galves, A. & Schmitt,B. **Repetition times for gibbsian sources,** Nonlinearity 12 (1999), 1225–1237.

[10] Collet, P., Galves, A. & Schmitt,B.**Unpredictability of the occurrence time of a long laminar period in a model of temporal intermittency** Annales de L'Institut Henri Poincaré (Physique Théorique), 57 n^o 3, 319-331, 1992.

[11] Galves, A. & Schmitt, B.; **Ocurrance times of rare events for mixing dynamical systems.** Annales de L'Institut Henri Poincaré (Physique Théorique), section A, tome 52 n^o 3(1990). p.267-281.

[12] Galves, A. & Schmitt, B.; **Inequalities for hitting times in mixing dynamical systems**, Random and Computational Dynamics (1997), 5, 4, 337-347.

[13] Galves, A., Maume-Deschamps, V. & Schmitt, B.; **Exponential inequalities for VMLC empirical trees,** ESAIM: Probability and Statistics, Vol. 12, 2008 12, 219-229.

[14] Haydn, N. & Vaienti, S.; **The Renyi entropy function and the large deviation of short return times,** (Preprint). To appear in Ergodic Theory and Dynamical Systems.

[15] Haydn, N. & Vaienti,S.; **The limiting distribution and error terms for return time of dynamical systems,** Discrete and Continuous Dynamical Systems 10 (2004), 584–616.

[16] M. Hirata, **Poisson law for Axiom A diffeomorphisms.** Ergodic Theory and Dynamical Systems 13, no. 3, (1993) 533–556.

[17] Hirata, M., Saussol, B. & Vaienti, S.; **Statistics of return times: a general framework and new applications,** Communications in Mathematical Physics, 206 (1999), 33–55.

[18] Kac, M. **On the notion of recurrence in discrete stochastic processes,** Bulletin of the American Mathematical Society 53, (1947) 1002-1010.

[19] Resnick, Sidney I. **A Probability Path**. Birkhauser, 1998.

[20] Reinert, G. & S. Schbath, S.; **Compound Poisson and Poisson process approximations for occurrences of multiple words in markov chains.** Journal of Computational Biology 5 (1998) 223-254.

[21] Reinert, G. & Schbath, S.; **Compound Poisson approximations for occurrences of multiple words.** Statistics in Genetics and Molecular Biology, (F. Seillier, ed.). IMS Lecture Notes-Monograph Series. Vol. 33, 1999.

[22] Roquain, E. & Schbath, S.; **Improved compound Poisson approximation for the number of occurrences of multiple words in a stationary Markov chain.** Advances in Applied Probability 39 (2007) 1-13.

[23] Rocha, Andréa Vanessa **Substitution Operators**, Tese (doutorado) Universidade Federal de Pernambuco. CCEN. Matemática Computacional, 2009.

[24] Saussol, B., Troubetzkoy, S. & Vaienti, S.; **Recurrence, dimensions and Lyapunov exponents**, Journal of Statistical Physics, 106, pp.623 – 634, 2002.