



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE MATEMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Longa Dependência em Sequências de DNA: Análise  
de Flutuações Destendenciadas, Teorias das  
Distribuições Estáveis e de Wavelets

Raquel Romes Linhares

Porto Alegre, 24 de Janeiro de 2011

Tese submetida por Raquel Romes Linhares<sup>1</sup> como requisito parcial para a obtenção do grau de Doutorado em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professora Orientadora:

Dr<sup>a</sup>. Sílvia Regina Costa Lopes

Banca Examinadora:

Dr. Artur Oscar Lopes

Dr<sup>a</sup>. Daniela Müller Quevedo

Dr. Eduardo Henrique de Mattos Brietzke

Dr<sup>a</sup>. Simone Maffini Cerezer

Data da Defesa: 24 de Janeiro de 2011

---

<sup>1</sup>Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

## AGRADECIMENTOS

Agradeço inicialmente à DEUS por me guiar sempre. Gostaria de agradecer à minha orientadora, Dr<sup>a</sup>. Sílvia Regina Costa Lopes, pela orientação e principalmente pelo permanente incentivo. À todos os meus colegas pela amizade. À Rosane, nossa querida secretária do Programa de Pós-Graduação pela atenção e paciência. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo auxílio financeiro. Quero agradecer em especial à minha mãe Aparecida Marcia Silva Linhares, ao meu pai Eidmar Oliveira Linhares e ao meu irmão Erick Oliveira Linhares pelo apoio incondicional. Por fim quero agradecer à todos que, de uma forma ou de outra, contribuíram pela realização deste trabalho.

## RESUMO

O método da *análise de flutuações destendenciadas* (*Detrended Fluctuation Analysis* - DFA), proposto por Peng et al. (1994), é um exemplo de metodologia recente, sendo utilizada em um crescente número de aplicações, para identificar longa dependência em séries temporais. Como não existe uma regra específica para a escolha dos números de regressores no método de DFA, apresentamos aqui uma escolha ótima assintótica. Para um ruído Gaussiano fracionário, provamos que o estimador  $\hat{H}_{\text{DFA}}$  tem distribuição Gaussiana exata e assintótica.

O parâmetro mais importante para ser estimado em dados com caudas pesadas é a sua taxa de decaimento  $\alpha$  que determina a probabilidade de ocorrência dos valores extremos da distribuição subjacente. Propomos, neste trabalho, um novo estimador para o parâmetro  $\alpha$ , baseado na função característica empírica e no procedimento de *encolhimento de wavelets*.

Estamos interessados em analisar a longa dependência em sequência de DNA utilizando a metodologia de mudança de regimes, proposta por Liu (2000). Nesta metodologia, se a duração dos regimes de uma série temporal tem uma distribuição de caudas pesadas com parâmetro  $\alpha \in (1, 2)$ , então a série temporal apresenta a característica de longa dependência. Além disso, aplicando-se qualquer transformação linear que preserve a propriedade de variância finita na série temporal, igualmente preservará a propriedade de longa dependência.

Por fim, estudamos as distribuições de distâncias das regiões codantes e não codantes em sequências de DNA.

Concluimos que todas as técnicas apresentadas neste trabalho, para analisar longa dependência em série temporais, envolvendo conceitos de análise de flutuações destendenciadas, distribuições com caudas pesadas e encolhimento de wavelets, mostram a existência de longa dependência em todas sequências de DNA aqui estudadas.

## ABSTRACT

The method of *detrended fluctuation analysis* (DFA), proposed by Peng et al. (1994), is useful in revealing the extent of long-range dependence in time series. Since there is not a specific rule for the choice of the numbers of regressors in DFA method, we present here an asymptotic optimal choice. For a fractional Gaussian noise, we prove the exact and the asymptotic Gaussian distributions for the  $\widehat{H}_{\text{DFA}}$  estimator.

The most important parameter to estimate in a heavy-tailed data is the tail rate of decay  $\alpha$  which determines the probability of occurrence of extreme values of the underlying distribution. Here we propose a novel estimator for  $\alpha$  based on the empirical characteristic function and on the principal of wavelet shrinkage.

Here we are interested in analyzing the *long-range dependence* in several DNA sequences, under the heavy-tail regime switching mechanism, proposed by Liu (2000). In this mechanism, if the duration of the regimes of a given time series has a heavy tail distribution with index parameter  $\alpha \in (1, 2)$ , then there is long-range dependence in this time series, and any functional transformation of the original time series preserving the property of finite variance, also preserves the property of long-range dependence.

We also study the length distribution of coding and noncoding regions in DNA sequences.

We conclude that all techniques presented in this paper to analyze long-range dependence in time series, involving concepts of detrended fluctuation analysis, heavy tail distributions and wavelet shrinkage, show the existence of long-range dependence in all DNA sequences studied here.

# Índice

Nomenclatura	vii
Índice de Figuras	viii
Índice de Tabelas	ix
<b>1 Introdução</b>	<b>1</b>
<b>2 Noções Preliminares</b>	<b>4</b>
2.1 Processos Estocásticos . . . . .	5
2.2 Estimação do Parâmetro de Longa Dependência . . . . .	10
2.2.1 Estimador $\hat{d}_{\text{GPH}}$ . . . . .	11
2.2.2 Estimador $\hat{d}_{\text{W}}$ . . . . .	11
2.2.3 Estimador $\hat{d}_{\text{WS}}$ . . . . .	12
<b>3 Método das Análises de Flutuações Destendenciadas</b>	<b>13</b>
3.1 Propriedades do Estimador $\hat{H}_{\text{DFA}}$ . . . . .	15
3.1.1 Erros i.i.d. com Distribuição Gaussiana . . . . .	15
3.1.2 Erros i.i.d. com Distribuição Não Gaussiana . . . . .	27
3.2 Simulações . . . . .	37
3.2.1 Investigação do Efeito de $g(n)$ no Método DFA . . . . .	38
3.2.2 Comparação entre os Estimadores $\hat{d}_{\text{DFA}}$ e $\hat{d}_{\text{W}}$ . . . . .	43
<b>4 Wavelets</b>	<b>45</b>
4.1 Análise da Wavelet Haar . . . . .	51
4.2 Regularidade das Wavelets . . . . .	56
4.3 Transformadas de Wavelets . . . . .	57
4.4 Análise de Multiresolução . . . . .	59
4.5 Decomposição e Reconstrução . . . . .	66
4.5.1 Decomposição . . . . .	66
4.5.2 Reconstrução . . . . .	69
4.6 Encolhimento de Wavelets . . . . .	70
4.7 Núcleos de Wavelets . . . . .	74
4.8 Wavelets e Processos Estacionários . . . . .	77

4.8.1	Transformadas de Wavelets de Processos Estacionários	77
4.8.2	<i>Whitening</i> de Processos Estacionários . . . . .	78
<b>5</b>	<b>Teoria dos Valores Extremos</b>	<b>81</b>
5.1	Distribuições Estáveis . . . . .	87
5.2	Estimadores para o Parâmetro $\alpha$ . . . . .	90
5.3	Simulações . . . . .	97
<b>6</b>	<b>Análise em Sequências de DNA</b>	<b>103</b>
6.1	Longa Dependência Através de Diferentes Transformações . .	103
6.2	Mudança de Regime . . . . .	105
6.3	Análise da Distribuição do Comprimento . . . . .	110
<b>7</b>	<b>Conclusão</b>	<b>113</b>
	<b>Referências</b>	<b>115</b>
	<b>Anexo A</b>	<b>121</b>

# Nomenclatura

- $\alpha$  Índice de estabilidade de distribuições  $\alpha$ -estáveis.
- $\approx$  Aproximadamente igual.
- $\mathbb{C}^p$  Conjunto de todas as funções reais  $p$  vezes continuamente diferenciáveis.
- $\Gamma'(\cdot)$  Derivada da função matemática gama.
- $\Gamma(\cdot)$  Função matemática gama.
- $\Gamma(\cdot, \cdot)$  Distribuição Gama.
- $\gamma_X(\cdot)$  Função de autocovariância do processo  $X$ .
- $[\cdot]$  Função maior inteiro.
- $\ln(\cdot)$  Função logaritmo na base  $e$ .
- $\log_2(\cdot)$  Função logaritmo na base 2.
- $\mathbb{C}$  Conjunto dos números complexos.
- $\mathbb{N}$  Conjunto dos números naturais:  $\mathbb{N} = \{1, 2, 3, \dots\}$
- $\mathbb{R}$  Conjunto dos números reais.
- $\mathbb{Z}$  Conjunto dos números inteiros.
- $\mathbf{1}_n$  Vetor  $n \times 1$  com todas as componentes iguais à 1.
- $\mathcal{C}(\cdot, \cdot)$  Distribuição de Cauchy.
- $\mathcal{E}(\cdot)$  Distribuição Exponencial.
- $A'$  Transposta da matrix  $A$ .
- $\text{Corr}(\cdot, \cdot)$  Correlação.
- $\text{Cov}(\cdot, \cdot)$  Covariância.



- H Parâmetro H de Hurst.
- $I_A(\cdot)$  Função indicadora do conjunto  $A$ .
- U.M.V.U. Estimador não viciado e uniformemente de mínima variância.
- $\text{Var}(\cdot)$  Variância.
- $\mathcal{N}(\cdot, \cdot)$  Distribuição Gaussiana.
- $\bar{A}$  Fecho do conjunto  $A$
- $\psi'(\cdot)$  Derivada da função digama.
- $\psi(\cdot)$  Função digama.
- $\rho_X(\cdot)$  Função de autocorrelação do processo  $X$ .
- $\zeta(\cdot, \cdot)$  Função zeta de Hurwitz.
- $\text{avar}(\cdot)$  Variância assintótica.
- $d$  Parâmetro fracionário  $d$ .
- $f \sim g$   $f$  é assintótica para  $g$ .
- $I_n(\cdot)$  Função periodograma.
- $L^2(\mathbb{R})$  Conjunto de todas as funções reais de quadrado integrável.
- $S_\alpha(0, 1, 0)$  Classe das distribuições estáveis simétricas com média 0.
- $S_\alpha(\beta, \gamma, \delta)$  Classe das distribuições estáveis.
- $U(0, 1)$  Distribuição Uniforme no intervalo  $(0, 1)$ .
- $X_n \xrightarrow{D} X$  A sequência de variáveis aleatórias  $X_n$  converge para a variável aleatória  $X$  em distribuição.
- $X_n \xrightarrow{P} X$  A sequência de variáveis aleatórias  $X_n$  converge para a variável aleatória  $X$  em probabilidade.
- $X_n \xrightarrow{q.c.} X$  A sequência de variáveis aleatórias  $X_n$  converge para a variável aleatória  $X$  quase certamente.
- DFA Método da Análise de Flutuações Destendenciadas.
- eqm Erro quadrático médio.
- FCE Função Característica Empírica.
- FGN Ruído Gaussiano Fracionário.

# Índice de Figuras

4.1	Wavelets Haar: (a) wavelet mãe; (b) wavelet pai. . . . .	48
4.2	Diferentes Wavelets Mãe: (a) Haar; (b) Daubelets “d12”; (c) Symmlets “s12” e (d) Coiflets “c12”. . . . .	49
4.3	Versões Escaladas e Transladadas da Wavelet Symmlets “s8”: (a) escala 2 e translação 4; (b) escala 4 e translação 8; (c) escala 4 e translação 0 e (d) escala 4 e translação 4. . . . .	50
4.4	Sinal Doppler. . . . .	61
4.5	Transformada Discreta de Wavelets utilizando wavelet “s8”. . . . .	62
4.6	Decomposição de Multiresolução do Sinal Doppler. . . . .	63
4.7	Funções de Encolhimento com $\lambda = 3$ aplicadas na Função Linear $f(x) = x$ : (a) Função Encolhimento Suave; (b) Função Encolhimento Dura. . . . .	72
4.8	Ruído “Glint” de Radar. . . . .	74
4.9	TDW do Sinal “Glint” (baseada na wavelet “s8”). . . . .	75
4.10	TDW do Sinal Estimado (baseada na wavelet “s8”). . . . .	75
5.1	Funções Densidades das Distribuições $\mathcal{N}(0, 1)$ , $\mathcal{C}(1, 0)$ e Lévy(1, 0). . . . .	89
6.1	Histogramas das Séries Temporais de Durações de Regimes de Algumas Sequências de DNA Dadas na Tabela 6.2. . . . .	109
6.2	Histograma para Sequências das Distâncias de: (a) Regiões Codantes e (b) Não Codantes do Cromossomo 21 do <i>Homo sapiens</i> . . . . .	112

# Índice de Tabelas

3.1	Estimador $\hat{d}_{\text{DFA}}$ para ARFIMA(0, $d$ , 0) com Diferentes $g(n)$ . . . . .	39
3.2	Estimador $\hat{d}_{\text{DFA}}$ para ARFIMA(0, $d$ , 0) com Diferentes $g(n)$ . . . . .	40
3.3	Estimador $\hat{H}_{\text{DFA}}$ para FGN com Diferentes $g(n)$ . . . . .	41
3.4	Estimador $\hat{H}_{\text{DFA}}$ para FGN com Diferentes $g(n)$ . . . . .	42
3.5	Resultados dos Estimadores $\hat{d}_{\text{DFA}}$ (com $g(n) = \lfloor n^{0.5} \rfloor$ ) e $\hat{d}_{\text{W}}$ para Processos ARFIMA(0, $d$ , 0), para Diferentes Valores de $d$ e $n$ . . . . .	44
5.1	Resultado da Estimação de $\alpha$ , baseada em $\hat{\alpha}_{\text{fce}}$ , $\hat{\alpha}_{\text{reg}}$ e $\hat{\alpha}_{\text{wave}}$ , para $S_{\alpha}(0, 1, 0)$ com $n = 1000$ . . . . .	98
5.2	Resultado da Estimação de $\alpha$ , baseada em $\hat{\alpha}_{\text{fce}}$ , $\hat{\alpha}_{\text{reg}}$ e $\hat{\alpha}_{\text{wave}}$ , para $S_{\alpha}(0, 1, 0)$ com $n = 5000$ . . . . .	99
5.3	Resultado da Estimação de $\alpha$ , baseada em $\hat{\alpha}_{\text{fce}}$ , $\hat{\alpha}_{\text{reg}}$ e $\hat{\alpha}_{\text{wave}}$ , para para $S_{\alpha}(0, 1, 0)$ com $n = 10000$ . . . . .	100
5.4	Resultado da Estimação de $\alpha$ , baseada em $\hat{\alpha}_{\text{fce}}$ , $\hat{\alpha}_{\text{reg}}$ e $\hat{\alpha}_{\text{wave}}$ , para $S_{\alpha}(0, 1, 0)$ com $n = 50000$ . . . . .	101
5.5	Resultado da Estimação de $\alpha$ , baseada em $\hat{\alpha}_{\text{fce}}$ , $\hat{\alpha}_{\text{reg}}$ e $\hat{\alpha}_{\text{wave}}$ , para $S_{\alpha}(0, 1, 0)$ com $n = 100000$ . . . . .	102
6.1	Estimadores para o parâmetro $d$ , para a sequência de DNA AL163202, usando 30 diferentes transformações. . . . .	104
6.2	Estimadores $\hat{\alpha}_{\text{mle}}$ , $\hat{\alpha}_{\text{fce}}$ , $\hat{\alpha}_{\text{reg}}$ e $\hat{\alpha}_{\text{wave}}$ para o Índice $\alpha$ , das Séries Temporais de Durações de Regimes, de Sequências de DNA. . . . .	108
6.3	Estimador de Hill para as sequências das distâncias de região codante ( $\hat{\alpha}_{\text{cds}}$ ) e não codante ( $\hat{\alpha}_{\text{ncds}}$ ) para os cromossomos do <i>Homo sapiens</i> , com $k = \lfloor n^{0.5} \rfloor$ . . . . .	112

# Capítulo 1

## Introdução

O método da *análise de flutuações destendenciadas* (*Detrended Fluctuation Analysis* - DFA), proposto por Peng et al. (1994), é um exemplo de metodologia recente, sendo utilizada em um crescente número de aplicações, para identificar longa dependência em séries temporais (ver Peng et al., 1994; Liu et al., 1997; Koscielny-Bunde et al., 1998; Buldyrev et al., 1998; Ben-Avraham e Havlin, 2000; Bardet e Kammoun, 2008; Crato et al., 2010, Crato et al., 2011). O método DFA tem como objetivo o cálculo de uma flutuação estatística  $F(l)$ , onde  $l$  representa o tamanho de uma janela, para mapear um conjunto de medidas. Variando o tamanho de  $l$ , as flutuações podem ser caracterizadas através de um expoente de escala obtido a partir da curva ajustada ao gráfico  $\ln(F(l))$  versus  $\ln(l)$ , com  $l \in \{4, 5, \dots, g(n)\}$ . Para aplicar o método DFA, primeiro divide-se a série temporal em blocos com  $l$  observações. Em cada bloco, calculam-se as somas parciais  $\{Y_t\}_{t=1}^l$ , ajusta-se uma reta  $Y_t^l = a + bt$  para estas somas parciais e então calcula-se a soma  $\sum_{t=1}^l (Y_t - \hat{Y}_t^l)^2$ , para cada  $l$  observações. Taqqu et al. (1995) fornecem uma aproximação para a variância amostral dos resíduos obtidos da diferença entre  $Y_t$  e  $\hat{Y}_t^l$  em cada bloco. Os autores provam que a esperança da variância amostral é assintoticamente proporcional à  $l^{2H}$ , onde  $H$  é o parâmetro de Hurst (ver Hurst, 1951). Bardet e Kammoun (2008) constituem uma continuação do trabalho de Taqqu et al. (1995) e mostram, em uma abordagem semiparamétrica, para processos estacionários com longa dependência, considerando valores de  $l \in \{4, 5, \dots, g(n)\}$ , que o estimador DFA de longa dependência é convergente com uma taxa razoável de convergência (mas nem sempre ótima). Crato et al. (2010) provam a consistência do estimador DFA, baseados nas condições usuais *i.i.d.*, considerando  $l \in \{4, 5, \dots, g(n)\}$ . Na literatura a escolha mais utilizada para  $g(n)$  é  $\lfloor \frac{n}{10} \rfloor$ , mas não existe uma regra específica para esta escolha (ver Hu et al., 2001). Nosso primeiro objetivo é mostrar, matematicamente, uma escolha ótima assintótica para  $g(n)$  e provar que, sob certas condições, o estimador DFA tem distribuição assintótica Gaussiana.

Nas duas últimas décadas diversas técnicas estatísticas e probabilísticas foram desenvolvidas, sob a denominação geral de valores extremos, especialmente na área de finanças e seguros (ver Embrechts et al., 2003). Esta popularidade é devida a habilidade de melhor quantificar as probabilidades de ocorrência de eventos raros. O parâmetro mais importante a ser estimado, em dados com caudas pesadas, é a taxa de decaimento  $\alpha$  que determina a probabilidade da ocorrência dos valores extremos da distribuição subjacente. Isto é,  $\alpha$  é o parâmetro tal que  $\lim_{x \rightarrow \infty} \frac{\mathbb{P}(X > x)}{C x^{-\alpha}} = 1$ , para todo  $x \in \mathbb{R}$ , onde  $C$  é uma constante positiva. Diversos métodos vem sendo propostos para a estimação do parâmetro  $\alpha$ . Geralmente, estes métodos estão divididos em quatro categorias: métodos baseados no estimador de Hill (ver Hill, 1975), na máxima verossimilhança (ver Nolan, 2001), em quantis (ver Fama e Roll, 1971) e métodos baseados na função característica (ver Press, 1972 e Koutrouvelis, 1981). Nosso segundo objetivo é apresentar um método de estimação para o parâmetro  $\alpha$ , baseado na função característica empírica (ver Koutrouvelis, 1981), utilizando o procedimento de *encolhimento de wavelets* (ver Donoho e Jonstone, 1994). O procedimento de *encolhimento de wavelets* tem por objetivo a redução do ruído presente no sinal, reduzindo ou zerando a magnitude dos coeficientes de wavelets. Wavelets são nada mais do que ondas pequenas com determinadas propriedades que as tornam adequadas a servirem de base para a decomposição de outras funções, assim como senos e cossenos servem de base para decomposições de Fourier.

Uma sequência de DNA é composta de bases A (adenina), C (citosina), T (timina) e G (guanina). Um dos principais resultados recentes é a existência de longa dependência em uma sequência de DNA (ver Guharay et al., 2000; Cristea, 2002; Stoffer e Rosen, 2007 e Crato et. al, 2010). Para aplicar métodos numéricos a uma sequência de DNA é necessário transformá-la em uma sequência numérica. Do ponto de vista biológico, é comum aplicar a regra SW “*strong-weak pairing*”, a qual transforma C e G em 1 e A e T em 0. Na literatura existem diferentes transformações que aplicam métodos numéricos a uma sequência de DNA (ver Guharay et al., 2000; Cristea, 2002; Garcia e José, 2005; Stoffer e Rosen, 2007; Podobnik et al., 2007 e Crato et al., 2011). Nosso terceiro objetivo é analisar a longa dependência em sequências de DNA utilizando diferentes transformações. Se considerarmos a abordagem de transformação numérica arbitrária para  $\{A, C, T, G\}$  e depois utilizarmos a análise espectral, o resultado irá depender de cada transformação em particular. Portanto, não sabemos se a existência de longa dependência em sequência de DNA é real ou é induzida pela transformação. Nosso quarto objetivo é analisar a longa dependência em sequência de DNA utilizando metodologia de mudança de regimes (ver Hamilton, 1989) proposta por Liu (2000). Nesta metodologia, se a duração dos regimes de uma série temporal tem uma distribuição de cauda pesada com parâmetro  $\alpha \in (1, 2)$ , então a série temporal apresenta a característica de longa dependência. Além disso,

aplicando-se qualquer transformação linear que preserve a propriedade de variância finita na série temporal, igualmente preservará a propriedade de longa dependência. Portanto, nosso interesse é aplicar esta metodologia em sequências de DNA, para saber se a longa dependência nestas sequências é realmente uma propriedade destas sequências, ou ela é induzida pela transformação. Um problema comum na análise de sequências de DNA é como identificar regiões codantes (*cds*) dispersas por toda a sequência e separadas por regiões não codantes (*ncds*). Distinguindo a distribuição para cada uma destas regiões pode ser de grande ajuda para encontrar regiões codantes e não codantes em sequências de DNA. Então, nosso quinto objetivo é analisar as distribuições de distâncias das regiões codantes e não codantes em sequências de DNA. Esta análise nos possibilita identificar diferenças entre estas distribuições e, portanto, distinguir as duas regiões em qualquer sequência de DNA.

Portanto, o nosso objetivo principal é analisar a longa dependência em sequências de DNA, utilizando diferentes técnicas que apresentam conceitos de análise de flutuações destendenciadas, distribuições com caudas pesadas e encolhimento de wavelets.

O Capítulo 2 apresenta algumas definições, métodos e propriedades preliminares necessárias para a compreensão deste trabalho. No Capítulo 3 descrevemos o método DFA (“*Detrended Fluctuation Analysis*”), proposto por Peng et al. (1994), analisamos as suas propriedades estatísticas e no final do capítulo são apresentadas simulações. No Capítulo 4, encontra-se um breve resumo sobre a teoria de wavelets onde apresentamos a metodologia de *encolhimento de wavelets*. Apresentamos no Capítulo 5, uma breve introdução à teoria dos valores extremos, definimos distribuição  $\alpha$ -estáveis, apresentamos alguns estimadores para o parâmetro  $\alpha$ , propomos um novo método de estimação para este parâmetro utilizando o procedimento de *encolhimento de wavelets*, e apresentamos simulações. A metodologia de mudança de regimes, proposta por Liu (2000), e a análise de diversas sequências de DNA, com o uso do software R-project são apresentadas no Capítulo 6. As conclusões finais são apresentadas no Capítulo 7.

# Capítulo 2

## Noções Preliminares

Apresentamos, neste capítulo, algumas definições, métodos e propriedades preliminares necessárias para a compreensão deste trabalho. Apresentamos alguns conceitos básicos para a análise de processos estocásticos e séries temporais, tais como estacionariedade e longa dependência. Descrevemos os processos ARFIMA( $p, d, q$ ) e os processos ruído Gaussiano fracionário. Por fim, apresentamos alguns métodos de estimação para o parâmetro fracionário  $d$  dos processos estocásticos ARFIMA( $p, d, q$ ).

**Definição 2.1 (Função Matemática Gama).** A *função matemática gama* é definida por

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \quad \text{para todo } \alpha > 0. \quad (2.1)$$

Denotamos por  $\Gamma'(\cdot)$ , a derivada da função matemática gama  $\Gamma(\cdot)$ .

**Definição 2.2 (Função Digama).** A *função digama* é definida por

$$\psi(\alpha) = \frac{d}{dx}(\ln(\Gamma(\alpha))) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}, \quad \text{para todo } \alpha > 0,$$

onde  $\Gamma(\cdot)$  é a função matemática gama e  $\Gamma'(\cdot)$  sua derivada. Denotamos por  $\psi'(\cdot)$ , a derivada da função digama  $\psi(\cdot)$ .

**Definição 2.3 (Função Zeta de Hurwitz).** A *função zeta de Hurwitz* é definida formalmente para um argumento complexo  $z$  e um argumento real  $q$  por

$$\zeta(z, q) = \sum_{k=0}^{\infty} (k+q)^{-z}.$$

Esta série é convergente para  $q > 0$  e  $\text{Re}(z) > 1$ , onde  $\text{Re}(z)$  é a parte real de  $z \in \mathbb{C}$ .

**Definição 2.4 (Distribuição Degenerada).** Uma *distribuição degenerada* é a distribuição de probabilidade de uma variável aleatória discreta cujo suporte consiste de somente um valor.

**Definição 2.5 (Distribuição Gama).** Uma variável aleatória  $X$  tem função de *distribuição gama*, se sua função densidade for dada por

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbf{I}_{(0,\infty)}(x),$$

onde  $\alpha$  e  $\beta$  são parâmetros positivos,  $\mathbf{I}_A(\cdot)$  é a função indicadora do conjunto  $A$  e  $\Gamma(\cdot)$  é a função matemática gama. Usamos a notação  $X \sim \Gamma(\alpha, \beta)$ .

Um vetor aleatório tem função de distribuição Gaussiana multivariada se possui a mesma função de distribuição de uma transformação afim de normais-padrão independentes. Ver definição a seguir.

**Definição 2.6 (Distribuição Gaussiana Multivariada).** Sejam  $X_1, \dots, X_m$  variáveis aleatórias independentes e identicamente distribuídas tais que  $X_j \sim \mathcal{N}(0, 1)$  para todo  $j \in \{1, \dots, m\}$  e seja  $\mathbf{Y} = (Y_1, \dots, Y_k)'$  o vetor aleatório obtido de  $\mathbf{X} = (X_1, \dots, X_m)'$  através da transformação

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \boldsymbol{\eta},$$

onde  $\mathbf{A}$  é uma matriz real  $k \times m$  e  $\boldsymbol{\eta}$  é um vetor real  $k$ -dimensional. Então dizemos que  $\mathbf{Y}$  tem função de *distribuição Gaussiana multivariada* com média  $\boldsymbol{\eta}$  e matriz de covariâncias  $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}'$ . A função de *distribuição Gaussiana multivariada singular* é aquela cuja matriz de covariância tem o posto estritamente menor do que a dimensão do vetor aleatório  $\mathbf{Y}$ .

## 2.1 Processos Estocásticos

Na análise de séries temporais, o objetivo inicial é tentar identificar, se possível, de qual processo estocástico foi originado um conjunto de dados (que é uma série temporal), e com isto poder realizar previsões do que se pode esperar no futuro.

**Definição 2.7 (Processo Estocástico).** Sejam  $(\Omega, \mathfrak{F}, \mathbb{P})$  um espaço de probabilidade,  $(S, \mathcal{G})$  um espaço mensurável e  $\{X_t\}_{t \in T}$  uma família de variáveis aleatórias onde, para cada  $t \in T$ ,  $X_t : (\Omega, \mathfrak{F}, \mathbb{P}) \rightarrow (S, \mathcal{G})$  é mensurável. Dizemos que  $\{X_t\}_{t \in T}$  é um *processo estocástico*. O conjunto de índice  $T$  pode ser discreto ou contínuo. No presente texto  $S$  é sempre finito ou enumerável e assim  $\mathcal{G} = \mathcal{P}(S)$ , onde  $\mathcal{P}(S)$  é o conjunto das partes de  $S$ .

A seguir, apresentamos dois modelos de processos estocásticos.



**Definição 2.8 (Ruído Branco).** O processo  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  é denominado um *ruído branco* com média zero e variância  $\sigma_\varepsilon^2$ , denotado por  $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ , se

$$\mathbb{E}(\varepsilon_t) = 0, \quad \text{Var}(\varepsilon_t) = \mathbb{E}(\varepsilon_t^2) = \sigma_\varepsilon^2 \quad \text{e} \quad \gamma_\varepsilon(k) = \begin{cases} \sigma_\varepsilon^2, & k = 0, \\ 0, & k \neq 0. \end{cases}$$

**Definição 2.9 (Processo Gaussiano).** Um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é dito *Gaussiano* se, para qualquer conjunto  $t_1, t_2, \dots, t_n \in \mathbb{Z}$ , as variáveis aleatórias  $X_{t_1}, X_{t_2}, \dots, X_{t_n}$  têm uma função de distribuição normal multivariada.

Um conjunto finito de observações organizadas cronologicamente no tempo é denominada uma *série temporal*, como definimos a seguir.

**Definição 2.10 (Série Temporal).** Uma *série temporal* é uma realização  $\{X_t\}_{t=1}^n$  de um processo estocástico  $\{X_t\}_{t \in T}$ .

A análise das diversas propriedades de um processo estocástico será tão mais complexa quanto mais intensa forem as relações de dependência entre as variáveis. Nesta hierarquia de dificuldades, os mais simples são os processos independentes.

**Definição 2.11 (Processos Independentes).** Seja  $\{X_t\}_{t \in \mathbb{N}}$  um processo estocástico definido em um espaço de probabilidade  $(\Omega, \mathfrak{F}, \mathbb{P})$ , tomando valores em  $S$  (enumerável). Dizemos que o processo estocástico  $\{X_t\}_{t \in \mathbb{N}}$  é independente, se

$$\begin{aligned} \mathbb{P}(X_{t_1} \in A_1, X_{t_2} \in A_2, \dots, X_{t_n} \in A_n) = \\ \mathbb{P}(X_{t_1} \in A_1) \mathbb{P}(X_{t_2} \in A_2) \dots \mathbb{P}(X_{t_n} \in A_n), \end{aligned}$$

para todo  $n \in \mathbb{N}$ , para toda sequência  $t_1 < t_2 < \dots < t_n$ , onde  $t_j \in \mathbb{N}$ , e para toda sequência de conjuntos  $A_1, A_2, \dots, A_n$ , onde  $A_j \subset S$ , para todo  $j = 1, \dots, n$ .

Para obtermos propriedades probabilísticas e estatísticas desejáveis na utilização de modelos para descrever processos físicos, são necessárias algumas suposições. No caso de processos estocásticos, uma suposição geralmente feita é a estacionariedade. Intuitivamente, um processo é estacionário se ele se desenvolve no tempo, mantendo-se em torno da sua média, de modo que a escolha de uma origem não seja importante.

**Definição 2.12 (Processo Estacionário).** Um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é dito *estacionário* se e somente se:

- (a)  $\mathbb{E}(|X_t|^2) < \infty, \forall t \in \mathbb{Z}$
- (b)  $\mathbb{E}(X_t) = m, \forall t \in \mathbb{Z}$

$$(c) \text{Cov}(X_r, X_s) = \text{Cov}(X_{r+k}, X_{s+k}), \forall r, s, k \in \mathbb{Z}.$$

Os processos estritamente estacionários são aqueles em que um deslocamento uniforme de um valor  $k$ , em todos os tempos envolvidos na função de distribuição conjunta, não altera esta função. A seguir, definimos formalmente este tipo de processo.

**Definição 2.13 (Processo Estritamente Estacionário).** Um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é dito *estritamente estacionário* se, e somente se,

$$\mathbb{P}(X_{t_1} \in A_1, X_{t_2} \in A_2, \dots, X_{t_n} \in A_n) =$$

$$\mathbb{P}(X_{t_1+k} \in A_1, X_{t_2+k} \in A_2, \dots, X_{t_n+k} \in A_n),$$

para todo  $n \in \mathbb{N}$ ,  $k > 0$ , para toda sequência  $t_1 < t_2 < \dots < t_n$ , onde  $t_j \in \mathbb{N}$ , e para toda sequência de conjuntos  $A_1, A_2, \dots, A_n$ , onde  $A_j \subset S$ , para todo  $j = 1, \dots, n$ .

**Definição 2.14 (Função de Autocovariância).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estacionário. A *função de autocovariância*  $\gamma_X(\cdot)$  de  $\{X_t\}_{t \in \mathbb{Z}}$  é definida por

$$\begin{aligned} \gamma_X(k) &\equiv \text{Cov}(X_{t+k}, X_t) \\ &= \mathbb{E}[(X_{t+k} - \mathbb{E}(X_{t+k}))(X_t - \mathbb{E}(X_t))], \text{ para todo } t, k \in \mathbb{Z}. \end{aligned}$$

**Definição 2.15 (Função de Autocorrelação).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estacionário. A *função de autocorrelação*  $\rho_X(\cdot)$  de  $\{X_t\}_{t \in \mathbb{Z}}$  é definida por

$$\rho_X(k) \equiv \frac{\gamma_X(k)}{\gamma_X(0)} = \text{Corr}(X_{t+k}, X_t), \text{ para todo } t, k \in \mathbb{Z},$$

onde  $\gamma_X(\cdot)$  é a função de autocovariância de  $\{X_t\}_{t \in \mathbb{Z}}$  e  $\gamma_X(0) = \text{Cov}(X_t, X_t) = \text{Var}(X_t)$ .

Usaremos como estimador de  $\gamma_X(\cdot)$  a função de autocovariância amostral e de  $\rho_X(\cdot)$  a função de autocorrelação amostral definidas a seguir.

**Definição 2.16 (Função de Autocovariância Amostral).** Seja  $\{X_t\}_{t=1}^n$  uma série temporal advinda do processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ . A *função de autocovariância amostral* de ordem  $k$  é definida por

$$\hat{\gamma}_X(k) = \frac{1}{n} \sum_{j=1}^{n-k} (X_{j+k} - \bar{X})(X_j - \bar{X}), \quad 0 \leq k < n, \quad (2.2)$$

e  $\hat{\gamma}_X(k) = \hat{\gamma}_X(-k)$ ,  $-n < k \leq 0$ , onde  $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$  é a *média amostral*.

**Definição 2.17 (Função de Autocorrelação Amostral).** Seja  $\{X_t\}_{t=1}^n$  uma série temporal advinda do processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ . A *função de autocorrelação amostral* de ordem  $k$  é dada por

$$\hat{\rho}_X(k) \equiv \frac{\hat{\gamma}_X(k)}{\hat{\gamma}_X(0)}, \quad |k| < n,$$

onde  $\hat{\gamma}_X(\cdot)$  é dada pela expressão (2.2).

A análise de processos estocásticos estacionários pela sua representação espectral é normalmente referida como análise no domínio da frequência. A análise no domínio do tempo é baseada nas funções de autocovariância e de autocorrelação e é uma análise equivalente à anterior.

**Definição 2.18 (Função Densidade Espectral).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário com função de autocovariância  $\gamma_X(\cdot)$  absolutamente convergente, i.e.,  $\sum_{k \in \mathbb{Z}} |\gamma_X(k)| < \infty$ . A *função densidade espectral* de  $\{X_t\}_{t \in \mathbb{Z}}$  é definida por

$$f_X(w) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_X(k) e^{-iwk}, \quad \text{para todo } w \in (-\pi, \pi].$$

A seguir definimos a *função periodograma*, que é um estimador não-viciado mas, inconsistente, para a função densidade espectral  $f_X(\cdot)$ .

**Definição 2.19 (Função Periodograma).** Sejam  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário e  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir deste processo. A *função periodograma* calculada à partir da série temporal  $\{X_t\}_{t=1}^n$ , é definida por

$$I_n(w) = \frac{1}{2\pi} \left( \hat{\gamma}_X(0) + 2 \sum_{k=1}^{n-1} \hat{\gamma}_X(k) \cos(wk) \right), \quad w \in (-\pi, \pi], \quad (2.3)$$

onde  $\hat{\gamma}_X(\cdot)$  é a função de autocovariância amostral do processo  $\{X_t\}_{t \in \mathbb{Z}}$  dada na Definição 2.16.

A inferência estatística de longa dependência tem sido observada em séries temporais de diferentes áreas de estudos tais como meteorologia, astronomia, hidrologia e economia (ver Beran, 1994). A característica de *longa dependência* em um processo estocástico, significa que mesmo para tempos bastante distantes entre si, a correlação entre duas variáveis aleatórias quaisquer é não desprezível. No caso de processos estocásticos estacionários, a longa dependência pode ser caracterizada no domínio do tempo, onde a função de autocorrelação decai lentamente a zero, ou no domínio da frequência, onde a função densidade espectral tende ao infinito quando a frequência se aproxima de zero. A seguir introduzimos uma definição formal para a propriedade de *longa dependência* e apresentamos importantes processos estocásticos que caracterizam esta propriedade.

**Definição 2.20 (Longa Dependência).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estacionário. Se existe um número real  $d \in (0, 0.5)$  e uma constante  $C_\rho > 0$ , tal que

$$\lim_{t \rightarrow \infty} \frac{\rho_X(t)}{C_\rho t^{2d-1}} = 1,$$

onde  $\rho_X(\cdot)$  é a função de autocorrelação do processo, ou equivalentemente, se existe um número real  $d \in (0, 0.5)$  e uma constante  $C_f > 0$ , tal que

$$\lim_{t \rightarrow 0} \frac{f_X(t)}{C_f t^{-2d}} = 1,$$

onde  $f_X(\cdot)$  é a densidade espectral do processo. Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo estacionário com longa dependência (ou com longa memória).

**Definição 2.21 (Movimento Browniano Fracionário).** O processo estocástico  $\{B_H(t)\}_{\mathbb{R}^+}$  é denominado um movimento Browniano fracionário, se é um processo Gaussiano com média  $\mu = 0$ , incrementos estacionários, variância  $\mathbb{E}(B_H^2(t)) = t^{2H}$  e função de autocovariância dada por

$$\mathbb{E}(B_H(s)B_H(t)) = \frac{1}{2} \left\{ s^{2H} + t^{2H} - |s - t|^{2H} \right\}, \quad \text{para todo } t, s \in \mathbb{R}^+. \quad (2.4)$$

O índice  $H$  é o parâmetro sugerido por Harold Edwin Hurst (1880-1978), para medir longa dependência, nomeado *H de Hurst*.

**Definição 2.22 (Ruído Gaussiano Fracionário).** O processo estocástico  $\{X_t\}_{t \in \mathbb{R}^+}$  é denominado um ruído Gaussiano fracionário (FGN), se é o incremento do movimento Browniano fracionário, ou seja,

$$X_t = B_H(t+1) - B_H(t), \quad \text{para todo } t \in \mathbb{R}^+. \quad (2.5)$$

Se  $\frac{1}{2} < H < 1$ , o processo ruído Gaussiano fracionário  $\{X_t\}_{t \in \mathbb{R}^+}$  é formado por variáveis aleatórias possivelmente correlacionadas e com a característica de longa dependência.

**Definição 2.23 (Processos ARFIMA( $p, d, q$ )).** Sejam  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  um processo ruído branco e  $\mathcal{B}$  o operador defasagem, i.e.,  $\mathcal{B}^k(X_t) = X_{t-k}$ . Se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo estacionário tal que

$$\Phi(\mathcal{B})(1 - \mathcal{B})^d(X_t - \mu) = \Theta(\mathcal{B})\varepsilon_t, \quad t \in \mathbb{Z}, \quad (2.6)$$

para  $d \in (-0.5, 0.5)$ , onde  $\mu$  é a média do processo e  $\Phi(\cdot)$  e  $\Theta(\cdot)$  são os polinômios autoregressivos e média móvel, de ordens  $p$  e  $q$ , respectivamente, então  $\{X_t\}_{t \in \mathbb{Z}}$  é denominado um processo geral com diferenciação fracionária, denotado por ARFIMA( $p, d, q$ ).

Se  $0 < d < 0.5$ , o processo ARFIMA( $p, d, q$ ) apresenta a propriedade de *longa dependência*.

O parâmetro de H de Hurst está relacionado com o parâmetro fracionário  $d$ , pela seguinte equação (ver Taqqu et al., 1995),

$$d = H - \frac{1}{2}. \quad (2.7)$$

## 2.2 Estimação do Parâmetro de Longa Dependência

Nos últimos anos, um grande número de autores vem propondo métodos para a estimação do parâmetro fracionário  $d$ . Dentro da classe semiparamétrica, o método mais popular é referido como método GPH, proposto por Geweke e Porter-Hudak em 1983. O método mais conhecido na classe paramétrica é o proposto por Fox e Taqqu (1986), onde a função de máxima verossimilhança é obtida de forma aproximada.

Nesta seção, apresentamos alguns métodos de estimação para o parâmetro fracionário  $d$ : o método de regressão utilizando a função periodograma proposto por Geweke e Porter-Hudak (1983); o método da máxima verossimilhança (ver Fox e Taqqu, 1986), utilizando a aproximação sugerida por Whittle (1953) e um novo método utilizando a teoria do envelope espectral (ver Stoffer et al., 1993).

**Definição 2.24 (Estimador Não Viciado, U.M.V.U. e Consistente).**

Sejam  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário e  $\{X_t\}_{t=1}^n$  uma amostra deste processo. Seja  $\theta \in \mathbb{R}$  um parâmetro qualquer do processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  e seja  $T(X_1, \dots, X_n)$  um estimador qualquer para  $\theta$ .

i) Se  $\mathbb{E}(T(X_1, \dots, X_n)) = \theta$ , então  $T(X_1, \dots, X_n)$  é um *estimador não viciado* para  $\theta$ .

ii) Se  $\text{Var}(T^*(X_1, \dots, X_n)) \leq \text{Var}(T(X_1, \dots, X_n))$ , para quaisquer estimadores  $T(X_1, \dots, X_n)$  e  $T^*(X_1, \dots, X_n)$  não viciados de  $\theta$ , então  $T^*(X_1, \dots, X_n)$  é dito ser um *estimador U.M.V.U.* (“*Uniformly Minimum Variance Unbiased*”) para  $\theta$ .

iii) Seja  $\{T_n(X_1, \dots, X_n)\}_{n \geq 1}$  uma sequência de estimadores para  $\theta$ , que são gerados similarmente para cada  $n \geq 1$ . Então,  $T_n$  é um estimador *consistente* para  $\theta$ , se e somente se,

$$\lim_{n \rightarrow \infty} \mathbb{P}[|T_n(X_1, \dots, X_n) - \theta| \geq \epsilon] = 0,$$

para todo  $\epsilon > 0$ .

### 2.2.1 Estimador $\hat{d}_{\text{GPH}}$

Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico ARFIMA( $p, d, q$ ), definido na expressão (2.6). O primeiro método de estimação do parâmetro  $d$ , baseado na função periodograma  $I_n(\cdot)$ , (ver Definição 2.19) foi proposto por Geweke e Porter-Hudak (1983). Estes autores obtém o estimador para  $d$  considerando a regressão de  $\ln(I_n(\omega_j))$  versus  $\ln(2 \sin(\frac{\omega_j}{2}))^2$ . Portanto, o estimador  $\hat{d}_{\text{GPH}}$  é dado por

$$\hat{d}_{\text{GPH}} = - \frac{\sum_{j=1}^{g(n)} (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^{g(n)} (x_j - \bar{x})^2}, \quad (2.8)$$

onde  $y_j = \ln(I_n(w_j))$ ,  $x_j = \ln(2 \sin(\frac{w_j}{2}))^2$ ,  $\bar{x} = \frac{1}{g(n)} \sum_{j=1}^{g(n)} x_j$ , com  $g(n) = n^\beta$ , para algum  $\beta \in (0, 1)$ , e  $w_j$  a  $j$ -ésima frequência de Fourier.

### 2.2.2 Estimador $\hat{d}_{\text{W}}$

Apresentamos aqui o estimador de máxima verossimilhança aproximada sugerido por Fox e Taquq (1986). Será denotado por  $\hat{d}_{\text{W}}$  pois ele é baseado em uma aproximação da função de verossimilhança sugerida por Whittle (1953). O estimador  $\hat{d}_{\text{W}}$ , sob certas condições de regularidade, é consistente e tem distribuição assintótica normal. Além disso, é estimador não viciado e de variância mínima.

Seja  $\{X_t\}_{t=1}^n$  uma série temporal obtida do processo ARFIMA( $p, d, q$ ) definido pela expressão (2.6). O estimador para  $d$ , utilizando o método de verossimilhança, é o valor

$$\boldsymbol{\eta} = (\sigma_X^2, d, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q) \quad (2.9)$$

que minimiza a função

$$\mathcal{Q}(\boldsymbol{\eta}) = \sum_{j=1}^{\lfloor \frac{n-1}{2} \rfloor} \frac{I_n(w_j)}{f_X(w_j, \boldsymbol{\eta})}, \quad (2.10)$$

onde  $\boldsymbol{\eta}$  é o vetor de parâmetros desconhecidos dado em (2.9),  $f_X(\cdot, \boldsymbol{\eta})$  é a função densidade espectral de  $\{X_t\}_{t \in \mathbb{Z}}$ ,  $\lfloor \cdot \rfloor$  é a função maior inteiro,  $w_j = \frac{2\pi j}{n}$  é a  $j$ -ésima frequência de Fourier,  $j \in \{1, \dots, \lfloor \frac{n-1}{2} \rfloor\}$ , e  $I_n(\cdot)$  é a função periodograma (ver Definição 2.19).

Ver Fox e Taquq (1986) e Beran (1994), para um estudo mais completo do estimador  $\hat{d}_{\text{W}}$ .

### 2.2.3 Estimador $\hat{d}_{\text{WS}}$

Nesta subseção, propomos um estimador para  $d$ , obtido pelo método da máxima verossimilhança, baseado na teoria do envelope espectral (ver Stoffer et al., 1993). A seguir, definimos a função envelope espectral.

**Definição 2.25 (Função Envelope Espectral).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico categórico, ou seja, que toma valores em um espaço finito  $\mathcal{C} = \{c_1, c_2, \dots, c_k\}$ . Seja  $p_j = \mathbb{P}(X_t = c_j) > 0$ , para todo  $j = 1, 2, \dots, k$ . Para  $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k) \in \mathbb{R}^k$  denotamos por  $\{X_t(\boldsymbol{\beta})\}_{t \in \mathbb{Z}}$  o processo correspondente a transformação, onde  $c_j$  é dado por  $\beta_j$ ,  $j = 1, 2, \dots, k$ . Assumimos que  $\{X_t(\boldsymbol{\beta})\}_{t \in \mathbb{Z}}$  é um processo estacionário e que  $f_X(\omega; \boldsymbol{\beta})$  é a função densidade espectral deste processo. A *função envelope espectral* é definida por

$$\lambda(\omega) = \max_{\boldsymbol{\beta}} \left\{ \frac{f_X(\omega; \boldsymbol{\beta})}{\sigma^2(\boldsymbol{\beta})} \right\}, \quad \omega \in (0, \pi], \quad (2.11)$$

para todo  $\boldsymbol{\beta}$  não proporcional a  $\mathbf{1}_k$  (vetor  $k \times 1$  cujas componentes são todas iguais à 1).

O nome *função envelope espectral* é bastante apropriado já que  $\lambda(\cdot)$  envolve o espectro normalizado de qualquer processo transformado  $\{X_t(\boldsymbol{\beta})\}_{t \in \mathbb{Z}}$ . Isto é, dado o vetor  $\boldsymbol{\beta}$  normalizado então  $\{X_t(\boldsymbol{\beta})\}_{t \in \mathbb{Z}}$  tem função densidade  $f_X(\omega; \boldsymbol{\beta}) \leq \lambda(\omega)$ , onde a desigualdade vale se e somente se  $\boldsymbol{\beta}$  é proporcional a  $\boldsymbol{\beta}(\omega)$ . Para mais detalhes ver Stoffer et al. (1993).

Nós propomos aqui um estimador para  $d$ , denotado por  $\hat{d}_{\text{WS}}$ , que é o valor que minimiza a função dada pela expressão (2.10), onde a função periodograma  $I_n(\cdot)$  é substituída pela função envelope espectral  $\lambda(\cdot)$  dada pela expressão (2.11).

# Capítulo 3

## Método das Análises de Flutuações Destendenciadas

O método das análises de flutuações destendenciadas (“*Detrended Fluctuation Analysis*” - DFA) foi proposto por Peng et al. (1994). Para utilizar este método em uma série temporal qualquer é preciso seguir os seguintes passos. Dada uma série temporal  $\{X_t\}_{t=1}^n$ , o primeiro passo do método DFA consiste em calcular a soma parcial

$$Y_t = \sum_{j=1}^t X_j, \quad (3.1)$$

para cada  $t \in \{1, 2, \dots, n\}$ . No segundo passo divide-se a série temporal  $\{Y_t\}_{t=1}^n$  em  $\lfloor \frac{n}{l} \rfloor$  blocos não sobrepostos ( $\lfloor \cdot \rfloor$  é a função maior inteiro), onde cada bloco contém exatamente  $l$  observações. No terceiro passo, para cada bloco ajusta-se, pelo método dos mínimos quadrados, uma reta aos dados. Denotamos por  $\widehat{Y}_t^l$ , para  $t = 1, \dots, n$ , a ordenada  $y$  dos segmentos de reta ajustados nos blocos de tamanho  $l$ . Em seguida, destendencia-se a série temporal  $\{Y_t\}_{t=1}^n$ , ou seja, em cada bloco calcula-se

$$Z_t^l = Y_t - \widehat{Y}_t^l, \quad \text{para todo } t \in \{1, \dots, n\}. \quad (3.2)$$

O desvio padrão dos resíduos obtidos da diferença entre  $Y_t$  e  $\widehat{Y}_t^l$  é denominado de *raiz da flutuação média quadrática* e é definido por

$$F(l) = \sqrt{\frac{1}{\tilde{n}} \sum_{t=1}^{\tilde{n}} (Z_t^l)^2}, \quad (3.3)$$

onde  $\tilde{n} = l \cdot \lfloor \frac{n}{l} \rfloor$ . Por fim, para cada  $l \in \{4, 5, \dots, g(n)\}$ , calcula-se a *raiz da flutuação média quadrática*, dada na expressão (3.3). Observe que  $F(l)$  aumenta quando  $l$  cresce. Uma relação linear em um gráfico  $\ln$  versus  $\ln$  indica presença de escala

$$F(l) \sim \varphi l^\beta. \quad (3.4)$$



Aplicando logaritmo em ambos os lados da expressão (3.4) obtemos

$$\ln(F(l)) \sim \ln(\varphi) + \beta \ln(l). \quad (3.5)$$

A expressão (3.5) é da forma de uma equação de regressão linear simples, dada por

$$y_j = a + bx_j + \epsilon_j, \quad j \in \{1, \dots, m\}, \quad (3.6)$$

onde

$$y_j = \ln(F(l)), \quad a = \ln(\varphi), \quad b = H, \quad x_j = \ln(l), \quad l = j + 3, \quad (3.7)$$

com  $l \in \{4, 5, \dots, g(n)\}$  e  $m = g(n) - 3$ . Não consideramos  $l = 3$  pelo aumento significativo no tempo computacional. A relação da escala  $\beta$  com o parâmetro  $H$  do efeito de Hurst (ver Hurst, 1951) da série temporal  $\{X_t\}_{t=1}^n$  é dada por  $\beta = H$ . Segue-se então que o estimador de  $H$ , pelo método dos mínimos quadrados baseado na regressão linear (3.6), é dado por

$$\hat{H}_{\text{DFA}} = \frac{\sum_{j=1}^m (x_j - \bar{x})y_j}{\sum_{j=1}^m (x_j - \bar{x})^2}, \quad (3.8)$$

onde  $y_j = \ln(F(j + 3))$ ,  $x_j = \ln(j + 3)$ ,  $\bar{x} = \frac{1}{m} \sum_{j=1}^m x_j$  e  $m = g(n) - 3$ .

O estimador para parâmetro fracionário  $d$ , pelo método DFA, é dado por

$$\hat{d}_{\text{DFA}} = \hat{H}_{\text{DFA}} - \frac{1}{2}, \quad (3.9)$$

onde  $\hat{H}_{\text{DFA}}$  é dado pela expressão 3.8.

Note que as equações dadas pelo modelo de regressão linear em (3.6), podem ser escritas na forma vetorial por

$$\mathbf{Y} = (y_1, \dots, y_m)' \equiv \boldsymbol{\xi} + \boldsymbol{\epsilon}, \quad (3.10)$$

onde

$$\boldsymbol{\xi} = (\xi_1, \dots, \xi_m)' \equiv (a + bx_1, \dots, a + bx_m)' \quad (3.11)$$

e

$$\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_m)',$$

onde  $X'$  designa o *transposto* do vetor  $X$ .

**Observação 3.1.** Na literatura da área, não existe uma regra específica para a escolha de  $g(n)$ . Então, determinamos, neste Capítulo, uma escolha ótima assintótica para o número de regressores  $m$  na equação (3.6).

### 3.1 Propriedades do Estimador $\hat{H}_{\text{DFA}}$

Apresentamos, nesta seção, propriedades estatísticas do estimador  $\hat{H}_{\text{DFA}}$ , dado na expressão (3.8), para os casos em que a sequência dos erros  $\{\epsilon_j\}_{j=1}^m$ , na expressão (3.6), tem função de distribuição Gaussiana ou não.

#### 3.1.1 Erros i.i.d. com Distribuição Gaussiana

Seja  $W = \{\mathbf{t} \in \mathbb{R}^m : \mathbf{t} = c_1 \mathbf{1}_m + c_2 \mathbf{x}'\}$ ,  $c_1, c_2 \in \mathbb{R}$  um espaço linear, gerado pelos vetores  $\mathbf{1}_m$  (vetor  $m \times 1$  de 1's) e  $\mathbf{x}' = (x_1, \dots, x_m)'$ . Note que  $\boldsymbol{\xi} \in W$  (ver expressão 3.11). Sejam  $\mathbf{v}_1, \dots, \mathbf{v}_m$ , vetores de dimensão  $m$ , que formam uma base ortonormal de  $\mathbb{R}^m$ , tal que  $\mathbf{v}_1, \dots, \mathbf{v}_r$  geram  $W$ .

Sabemos que uma base é ortonormal se,

$$\begin{aligned} \mathbf{v}_k' \mathbf{v}_j &= 0, & k \neq j \\ &= 1, & k = j. \end{aligned}$$

Logo, para todo vetor  $\mathbf{t}$ , tem-se que

$$\mathbf{t} = \sum_{j=1}^m (\mathbf{t}' \mathbf{v}_j) \mathbf{v}_j, \quad (3.12)$$

e se definimos a norma  $\|\cdot\|$ , por  $\|\mathbf{t}\|^2 = \sum_{j=1}^m (t_j)^2$ , então

$$\|\mathbf{t}\|^2 = \sum_{j=1}^m (\mathbf{t}' \mathbf{v}_j)^2. \quad (3.13)$$

Como  $\mathbf{v}_1, \dots, \mathbf{v}_r$  geram  $W$ , temos que  $\mathbf{t} = (t_1, \dots, t_m)' \in W$  se, e somente se,  $\mathbf{t}' \mathbf{v}_j = 0$ , para todo  $j = r+1, \dots, m$ . Então,  $\mathbf{t} \in W$  se, e somente se,

$$\mathbf{t} = \sum_{j=1}^r (\mathbf{t}' \mathbf{v}_j) \mathbf{v}_j. \quad (3.14)$$

Seja  $\mathbf{v}_k = (v_{k1}, \dots, v_{km})'$ . Definimos

$$Z_k = Y' \mathbf{v}_k = \sum_{j=1}^m Y_j v_{kj} \quad (3.15)$$

e

$$\eta_k = \boldsymbol{\xi}' \mathbf{v}_k = \sum_{j=1}^m \xi_j v_{kj}, \quad (3.16)$$

onde  $\mathbf{Y}$  e  $\boldsymbol{\xi}$ , são dados pela expressão (3.10). Multiplicando  $\mathbf{v}_k$  pelo lado direito, em ambos os lados da primeira igualdade da expressão (3.15), segue-se pela expressão (3.12), que

$$\mathbf{Y} = \sum_{k=1}^m Z_k \mathbf{v}_k = \boldsymbol{\xi} + \boldsymbol{\epsilon}. \quad (3.17)$$

Multiplicando  $\mathbf{v}_k$  pelo lado direito, em ambos os lados da primeira igualdade da expressão (3.16), segue-se pela expressão (3.12), que

$$\boldsymbol{\xi} = \sum_{k=1}^m \eta_k \mathbf{v}_k,$$

como  $\boldsymbol{\xi} \in W$ , então

$$\eta_k = 0, \quad r + 1 \leq k \leq m,$$

e

$$\boldsymbol{\xi} = \sum_{k=1}^r \eta_k \mathbf{v}_k. \quad (3.18)$$

**Teorema 3.1.** *Para todo  $k \in \{1, \dots, m\}$ , sejam  $Z_k$  e  $\eta_k$  definidos pelas expressões (3.15) e (3.16), respectivamente. Então,  $Z_k$  é um estimador U.M.V.U. de  $\eta_k$ .*

**Demonstração:** Ver Bickel e Doksum (1977), página 262. □

Tomando os primeiros vetores ortonormais como sendo

$$\mathbf{v}_1 = \left( \frac{1}{\sqrt{m}}, \dots, \frac{1}{\sqrt{m}} \right)' \quad \text{e} \quad \mathbf{v}_2 = (\tau_1, \dots, \tau_m),$$

onde

$$\tau_k = \frac{(x_k - \bar{x})}{\left( \sum_{j=1}^m (x_j - \bar{x})^2 \right)^{\frac{1}{2}}}, \quad \text{para } k = 1, \dots, m,$$

temos que  $\mathbf{v}_1$  e  $\mathbf{v}_2$  geram  $W$  e

$$\hat{\boldsymbol{\xi}} = \sum_{j=1}^2 (\mathbf{Y}' \mathbf{v}_j) \mathbf{v}_j = \left( \sum_{j=1}^m \frac{Y_j}{\sqrt{m}} \right) \mathbf{v}_1 + \left( \sum_{j=1}^m Y_j \tau_j \right) \mathbf{v}_2, \quad (3.19)$$

produz o vetor dos valores ajustados, dado na seguinte definição.

**Definição 3.1 (Vetor dos Valores Ajustados).** Seja  $\epsilon = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com função de distribuição  $\mathcal{N}(0, \sigma_\epsilon^2)$ . Então, o *vetor dos valores ajustados*, denotado por  $\hat{\xi} = (\hat{\xi}_1, \dots, \hat{\xi}_m)$ , é definido por

$$\hat{\xi}_j = \bar{Y} + (x_j - \bar{x}) \frac{\sum_{k=1}^m (x_k - \bar{x}) y_k}{\sum_{k=1}^m (x_k - \bar{x})^2}, \quad \text{para todo } j = 1, \dots, m. \quad (3.20)$$

A esperança e a variância de  $\hat{\xi}_j$ , para  $j = 1, \dots, m$ , são dadas por

$$\mathbb{E}(\hat{\xi}_j) = \xi \text{ e } \text{Var}(\hat{\xi}_j) = \sigma_\epsilon^2 \left( \frac{1}{m} + \frac{(x_j - \bar{x})^2}{\sum_{k=1}^m (x_k - \bar{x})^2} \right), \quad (3.21)$$

onde  $\bar{x} = \frac{1}{m} \sum_{k=1}^m x_k$  e  $\bar{y} = \frac{1}{m} \sum_{k=1}^m y_k$ , para todo  $j \in \{1, \dots, m\}$ .

**Proposição 3.1.** *O valor ajustado  $\hat{\xi}_j$ , dado na expressão (3.20), é um estimador U.M.V.U. para  $\xi_j$ , para todo  $j \in \{1, \dots, m\}$ .*

**Demonstração:** Para todo  $j \in \{1, \dots, m\}$ , seja  $\hat{\xi}_j$  o *valor ajustado*, dado na Definição 3.1. Note que,

$$\hat{\xi}_j = \sum_{k=1}^m Z_k v_{kj}, \quad 1 \leq j \leq m. \quad (3.22)$$

Basta substituir a expressão (3.15) em (3.19). Logo, pelo Teorema 3.1,  $\hat{\xi}_j$  é um estimador U.M.V.U. para  $\xi_j$ , para todo  $j \in \{1, \dots, m\}$ . □

**Proposição 3.2.** *Seja  $\epsilon = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com função de distribuição  $\mathcal{N}(0, \sigma_\epsilon^2)$ . Seja  $\psi(\cdot)$  uma função real linear definida por*

$$\psi(\xi) = \sum_{j=1}^m w_j \xi_j,$$

onde  $\xi$  é dado na expressão (3.11). Então,

$$\psi(\hat{\xi}) = \sum_{j=1}^m w_j \hat{\xi}_j, \quad (3.23)$$

é um estimador U.M.V.U. para a função  $\psi(\boldsymbol{\xi})$ , com esperança dada por

$$\mathbb{E}(\psi(\hat{\boldsymbol{\xi}})) = \psi(\boldsymbol{\xi}),$$

onde  $\hat{\boldsymbol{\xi}}$  é o vetor dos valores ajustados dado na Definição 3.1.

**Demonstração:** Pela Proposição 3.1, para todo  $j \in \{1, \dots, m\}$ , temos que  $\hat{\xi}_j$  é um estimador U.M.V.U. de  $\xi_j$ . Como a propriedade de estimador U.M.V.U. é preservada por operações lineares sobre estimadores U.M.V.U., segue-se que

$$\psi(\hat{\boldsymbol{\xi}}) = \sum_{j=1}^m w_j \hat{\xi}_j,$$

é um estimador U.M.V.U. para  $\psi(\boldsymbol{\xi})$ .

Pela Proposição 3.1, para todo  $j \in \{1, \dots, m\}$ , temos que  $\hat{\xi}_j$  é um estimador U.M.V.U. para  $\xi_j$ . Logo,  $\mathbb{E}(\hat{\xi}_j) = \xi_j$ . Segue-se então que

$$\mathbb{E}(\psi(\hat{\boldsymbol{\xi}})) = \mathbb{E}\left(\sum_{j=1}^m w_j \hat{\xi}_j\right) = \sum_{j=1}^m w_j \mathbb{E}(\hat{\xi}_j) = \sum_{j=1}^m w_j \xi_j = \psi(\boldsymbol{\xi}).$$

□

**Proposição 3.3.** *Seja  $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com função de distribuição  $\mathcal{N}(0, \sigma_\epsilon^2)$ . Seja  $\psi(\hat{\boldsymbol{\xi}})$  o estimador dado na expressão (3.23). Se a função  $\psi(\cdot)$  é linear, então  $\psi(\hat{\boldsymbol{\xi}})$  sempre pode ser escrito como uma função linear das observações,*

$$\psi(\hat{\boldsymbol{\xi}}) = \sum_{j=1}^m a_j y_j, \quad (3.24)$$

para algum  $\mathbf{a} = (a_1, \dots, a_m)' \in \mathbb{R}^m$ . A variância do estimador  $\psi(\hat{\boldsymbol{\xi}})$  é dada por

$$\text{Var}(\psi(\hat{\boldsymbol{\xi}})) = \sigma_\epsilon^2 \sum_{j=1}^m a_j^2, \quad (3.25)$$

onde  $\sigma_\epsilon^2 = \text{Var}(\epsilon_j)$ , para  $j \in \{1, \dots, m\}$ .

**Demonstração:** Pela expressão (3.15), para todo  $k \in \{1, \dots, n\}$ ,  $Z_k$  é uma função linear de  $Y_k$ , e pela expressão (3.22),  $\psi(\hat{\boldsymbol{\xi}})$  é uma função linear de  $Z_k$ . Então,

$$\psi(\hat{\boldsymbol{\xi}}) = \sum_{j=1}^m a_j y_j,$$

para algum  $\mathbf{a} = (a_1, \dots, a_m)' \in \mathbb{R}^m$ . A variância de  $\psi(\hat{\boldsymbol{\xi}})$  é dada por

$$\text{Var}(\psi(\hat{\boldsymbol{\xi}})) = \sum_{j=1}^m \text{Var}(a_j y_j) = \sum_{j=1}^m a_j^2 \text{Var}(y_j) = \sigma_\epsilon^2 \sum_{j=1}^m a_j^2.$$

□

O Teorema 3.2, a seguir, mostra que, sob certas condições, qualquer estimador obtido como uma combinação linear do vetor dos valores ajustados, tem distribuição normal.

**Teorema 3.2.** *Seja  $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com função de distribuição  $\mathcal{N}(0, \sigma_\epsilon^2)$ . Qualquer estimador  $\psi(\hat{\boldsymbol{\xi}})$ , dado pela expressão (3.23), tem distribuição normal tal que  $\mathbb{E}(\psi(\hat{\boldsymbol{\xi}})) = \psi(\boldsymbol{\xi})$  e a variância é dada pela expressão (3.25).*

**Demonstração:** A normalidade de  $\psi(\hat{\boldsymbol{\xi}})$ , segue-se da representação (3.24), e da propriedade aditiva da função de distribuição Gaussiana.

□

**Lema 3.1.** *Seja  $\hat{H}_{\text{DFA}}$  o estimador dado na expressão (3.8). Então,  $\hat{H}_{\text{DFA}}$  é uma combinação linear do vetor dos valores ajustados  $\hat{\boldsymbol{\xi}} = (\hat{\xi}_1, \dots, \hat{\xi}_m)'$ , dada por  $\psi(\hat{\boldsymbol{\xi}}) = \frac{\hat{\xi}_2 - \hat{\xi}_1}{x_2 - x_1} = \hat{\xi}_2 \frac{1}{x_2 - x_1} - \hat{\xi}_1 \frac{1}{x_2 - x_1}$ .*

**Demonstração:** Seja  $\hat{H}_{\text{DFA}}$  o estimador dado na expressão (3.8). O valor ajustado  $\hat{\xi}_j$  é dado pela expressão (3.20), para todo  $j \in \{1, \dots, m\}$ . Segue-se que

$$\hat{\xi}_2 - \hat{\xi}_1 = (x_2 - x_1) \frac{\sum_{j=1}^m (x_j - \bar{x}) y_j}{\sum_{j=1}^m (x_j - \bar{x})^2} = (x_2 - x_1) \hat{H}_{\text{DFA}}.$$

Portanto,

$$\hat{H}_{\text{DFA}} = \frac{\hat{\xi}_2 - \hat{\xi}_1}{x_2 - x_1}, \quad (3.26)$$

onde  $x_1 = \ln(4)$  e  $x_2 = \ln(5)$ .

□

O próximo teorema apresenta a condição necessária para que o estimador  $\hat{H}_{\text{DFA}}$ , dado na expressão (3.8), tenha função de distribuição Gaussiana exata.

**Teorema 3.3.** *Seja  $\epsilon = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com função de distribuição  $\mathcal{N}(0, \sigma_\epsilon^2)$ . Então, o estimador  $\hat{H}_{\text{DFA}}$ , dado pela expressão (3.8), tem distribuição normal com esperança e variância dadas, respectivamente, por*

$$\mathbb{E}(\hat{H}_{\text{DFA}}) = H \quad e \quad \text{Var}(\hat{H}_{\text{DFA}}) = \frac{\sigma_\epsilon^2}{\sum_{j=1}^m (x_j - \bar{x})^2},$$

onde  $x_j = \ln(j + 3)$  e  $\sigma_\epsilon^2 = \text{Var}(\epsilon_j)$ , para  $j \in \{1, \dots, m\}$ .

**Demonstração:** Pela definição do estimador  $\hat{H}_{\text{DFA}}$  em (3.8), temos que

$$\hat{H}_{\text{DFA}} = \frac{\sum_{j=1}^m (x_j - \bar{x})y_j}{\sum_{j=1}^m (x_j - \bar{x})^2} = \sum_{j=1}^m a_j y_j,$$

onde  $a_j = \frac{x_j - \bar{x}}{\sum_{j=1}^m (x_j - \bar{x})^2}$ ,  $y_j = \ln(F(j + 3))$  e  $x_j = \ln(j + 3)$ , para  $j \in \{1, \dots, m\}$ .

Pelo Lema 3.1 temos que  $\hat{H}_{\text{DFA}}$  é uma combinação linear do vetor dos valores ajustados  $\hat{\xi}$ , dada pela expressão (3.26). Portanto, pelo Teorema 3.2 o estimador  $\hat{H}_{\text{DFA}}$  é normalmente distribuído com

$$\mathbb{E}(\hat{H}_{\text{DFA}}) = \mathbb{E}(\psi(\hat{\xi})) = \psi(\xi) = \frac{\xi_2 - \xi_1}{x_2 - x_1} = \frac{a + Hx_2 - (a + Hx_1)}{x_2 - x_1} = H$$

e

$$\text{Var}(\hat{H}_{\text{DFA}}) = \sigma_\epsilon^2 \sum_{j=1}^m a_j^2 = \sigma_\epsilon^2 \frac{\sum_{j=1}^m (x_j - \bar{x})^2}{\left[ \sum_{j=1}^m (x_j - \bar{x})^2 \right]^2} = \frac{\sigma_\epsilon^2}{\sum_{j=1}^m (x_j - \bar{x})^2},$$

onde  $x_j = \ln(j + 3)$  e  $\sigma_\epsilon^2 = \text{Var}(\epsilon_j)$ , para  $j \in \{1, \dots, m\}$ . □

Para alguns resultados, que apresentamos a seguir, é interessante definir a flutuação média quadrática definida em cada bloco.

**Definição 3.2 (Flutuação Média Quadrática Parcial).** A *flutuação média quadrática parcial do  $j$ -ésimo bloco* é definida por

$$F_j^2(l) = \frac{1}{l} \sum_{t=l(j-1)+1}^{lj} (Z_t^l)^2, \quad \text{para todo } j \in \{1, \dots, \lfloor \frac{n}{l} \rfloor\}, \quad (3.27)$$

onde  $Z_t^l$  é dada pela expressão (3.2).

**Observação 3.2.** Note que a flutuação média quadrática é dada pela média das flutuações parciais

$$F^2(l) = \frac{1}{\lfloor \frac{n}{l} \rfloor} \sum_{j=1}^{\lfloor \frac{n}{l} \rfloor} F_j^2(l), \quad (3.28)$$

onde  $F(\cdot)$  é dada pela expressão (3.3).

**Definição 3.3.** Sejam  $f(\cdot)$  e  $g(\cdot)$  duas funções definidas sobre um mesmo conjunto  $D$ , que contém números arbitrariamente grandes. Dizemos que  $f(x)$  é *assintótica para  $g(x)$* , quando  $x \rightarrow \infty$ , se  $f(x) \neq 0$  e  $g(x) \neq 0$ , para todo  $x$  suficientemente grande, e

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1.$$

Denotamos por  $f(x) \sim g(x)$ .

**Observação 3.3.** Observe que a relação  $f(x) \sim g(x)$ , dada na Definição 3.3, pode ser reformulada como

$$f(x) = g(x)(1 + o(1)), \quad \text{quando } x \rightarrow \infty.$$

Ou seja,  $f(x) - g(x) = o(g(x))$ .

Vimos que, para aplicar o método DFA, primeiro divide-se a série temporal em blocos com  $l$  observações. Em cada bloco, calculam-se as somas parciais  $\{Y_t\}_{t=1}^l$ , ajusta-se uma reta  $Y_t^l = a + bt$  para estas somas parciais e então calcula-se a soma  $\sum_{t=1}^l (Y_t - \widehat{Y}_t^l)^2$ , para cada  $l$  observações. O Teorema 3.4 de Taqqu et al. (1995), a seguir, fornece uma aproximação para a variância amostral dos resíduos obtidos da diferença entre  $Y_t$  e  $\widehat{Y}_t^l$  em cada bloco. O objetivo é que a esperança da variância amostral seja assintoticamente proporcional à  $l^{2H}$ .

**Teorema 3.4.** *Seja  $\{X_t\}_{t \in \mathbb{R}^+}$  um processo ruído Gaussiano fracionário. Considere  $\{X_t\}_{t=1}^n$  uma série temporal advinda deste processo. Então, para todo  $j \in \{1, 2, \dots, \lfloor \frac{n}{l} \rfloor\}$ , tem-se que  $\mathbb{E}(F_j^2(l)) \sim C_H l^{2H}$ , quando  $l \rightarrow \infty$ , ou seja,*

$$\lim_{l \rightarrow \infty} \frac{\mathbb{E}(F_j^2(l))}{l^{2H}} = C_H, \quad (3.29)$$



onde  $F_j(l)$  é dada pela expressão (3.27) e

$$C_H = \left( \frac{2}{2H+1} + \frac{1}{H+2} - \frac{2}{H+1} \right). \quad (3.30)$$

**Demonstração:** Ver Taqqu et al. (1995). □

O Teorema 3.5 a seguir apresenta as condições necessárias para que a flutuação média quadrática parcial  $F_j^2(l)$ , dada pela expressão (3.27), tenha função de distribuição  $\Gamma(\frac{l}{2}, \frac{l}{2\sigma_l^2})$ , onde  $\sigma_l^2$  é a variância teórica das variáveis aleatórias  $Z_1^l, Z_2^l, \dots, Z_n^l$ .

**Teorema 3.5.** *Suponha que as variáveis aleatórias  $Z_1^l, Z_2^l, \dots, Z_n^l$ , dadas pela expressão (3.2), sejam independentes e identicamente distribuídas com função de distribuição comum  $\mathcal{N}(0, \sigma_l^2)$ . Então, para todo  $j \in \{1, 2, \dots, \lfloor \frac{n}{l} \rfloor\}$ ,  $F_j^2(l)$ , dada pela expressão (3.27), tem função de distribuição  $\Gamma(\frac{l}{2}, \frac{l}{2\sigma_l^2})$ .*

**Demonstração:** Ver Linhares (2007). □

Sob as condições dadas no Teorema 3.5, o Corolário 3.1 a seguir mostra que a flutuação média quadrática parcial  $F_j^2(l)$  é um estimador não viciado para a variância teórica  $\sigma_l^2$  das variáveis aleatórias  $Z_1^l, Z_2^l, \dots, Z_n^l$ , e é consistente, quando  $l \rightarrow \infty$ .

**Corolário 3.1.** *Suponha que as variáveis aleatórias  $Z_1^l, Z_2^l, \dots, Z_n^l$ , dadas pela expressão (3.2), sejam independentes e identicamente distribuídas com função de distribuição comum  $\mathcal{N}(0, \sigma_l^2)$ . Então, para todo  $j \in \{1, \dots, \lfloor \frac{n}{l} \rfloor\}$ ,  $F_j^2(l)$ , dada pela expressão (3.27), tem esperança e variância dadas, respectivamente, por*

$$\mathbb{E}(F_j^2(l)) = \sigma_l^2 \quad e \quad \text{Var}(F_j^2(l)) = \frac{2\sigma_l^4}{l}, \quad (3.31)$$

sempre que  $0 < \sigma_l^4 < \infty$ .

**Demonstração:** Ver Linhares (2007). □

O Teorema 3.6 a seguir apresenta as condições necessárias para que  $\hat{H}_{\text{DFA}}$ , dado pela expressão (3.8), seja um estimador não viciado e uniformemente de mínima variância.

**Teorema 3.6.** *Seja  $\epsilon = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com função de*

distribuição  $\mathcal{N}(0, \sigma_\epsilon^2)$ . Então,  $\widehat{H}_{\text{DFA}}$ , dado pela expressão (3.8), é um estimador U.M.V.U. com esperança e variância dadas, respectivamente, por

$$\mathbb{E}(\widehat{H}_{\text{DFA}}) = H \quad e \quad \text{Var}(\widehat{H}_{\text{DFA}}) = \frac{\sigma_\epsilon^2}{m \sum_{j=1}^m (x_j - \bar{x})^2}, \quad (3.32)$$

onde  $x_j = \ln(j+3)$  e  $\sigma_\epsilon^2 = \text{Var}(\epsilon_j)$ , para  $j \in \{1, \dots, m\}$ .

**Demonstração:** Ver Linhares (2007). □

O comportamento da soma e do produto de variáveis aleatórias, uma convergindo em distribuição e outra em probabilidade, é o objeto do próximo teorema.

**Teorema 3.7 (Teorema de Slutsky).** *Sejam  $\{X_n\}_{n \in \mathbb{N}}$  e  $\{Y_n\}_{n \in \mathbb{N}}$  sequências de variáveis aleatórias, tais que  $X_n \xrightarrow{\mathcal{D}} X$  e  $Y_n \xrightarrow{P} c$ , onde  $c$  é uma constante. Então,*

- (a)  $X_n + Y_n \xrightarrow{\mathcal{D}} X + c$ ,
- (b)  $X_n Y_n \xrightarrow{\mathcal{D}} cX$ ,
- (c) Se  $c \neq 0$ ,  $\frac{X_n}{Y_n} \xrightarrow{\mathcal{D}} \frac{X}{c}$ , desde que  $\mathbb{P}(Y_n \neq 0) = 1$ .

**Demonstração:** Ver Rohatgi (1976). □

O Teorema 3.8 a seguir é consequência do Teorema de Slutsky.

**Teorema 3.8.** *Sejam  $Y_1, Y_2, \dots$  variáveis aleatórias tais que  $\sqrt{n}(Y_n - u) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2)$ . Se a função  $g(\cdot)$  é derivável em  $u$ , então  $\sqrt{n}(g(Y_n) - g(u)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(g'(u))^2)$ .*

**Demonstração:** Ver Rohatgi (1976). □

O Teorema 3.9 mostra que, sob certas condições, a sequência  $\{\epsilon_j\}_{j=1}^m$ , dada na expressão (3.6), é formada por variáveis aleatórias independentes e identicamente distribuídas com função de distribuição  $\mathcal{N}(0, \frac{1}{2n})$ .

**Teorema 3.9.** *Seja  $\{X_t\}_{t \in \mathbb{R}^+}$  um processo ruído Gaussiano fracionário. Considere  $\{X_t\}_{t=1}^n$  uma série temporal advinda deste processo. Suponha que, para todo  $l \in \{4, 5, \dots, g(n)\}$ , as variáveis aleatórias  $Z_1^l, Z_2^l, \dots, Z_n^l$ , dadas pela expressão (3.2), sejam independentes e identicamente distribuídas com função de distribuição comum  $\mathcal{N}(0, \sigma_l^2)$ . Suponha que  $\lim_{g(n), n \rightarrow \infty} \left\lfloor \frac{n}{g(n)} \right\rfloor = \infty$ . Então,  $\widehat{H}_{\text{DFA}}$ , dado pela expressão (3.8), é um estimador*

(a) U.M.V.U. para o parâmetro  $H$ ,

(b) com função de distribuição Gaussiana, com esperança e variância dadas, respectivamente, por

$$\mathbb{E}(\widehat{H}_{\text{DFA}}) = H \quad e \quad \text{Var}(\widehat{H}_{\text{DFA}}) = \frac{1}{\frac{g(n)-3}{2n \sum_{j=1}^{g(n)-3} (x_j - \bar{x})^2}}, \quad (3.33)$$

onde  $x_j = \ln(j+3)$ , para  $j \in \{1, \dots, g(n)-3\}$ .

**Demonstração:** Para todo  $l \in \{4, 5, \dots, g(n)\}$  e  $j \in \{1, 2, \dots, \lfloor \frac{n}{l} \rfloor\}$ , denotamos

$$Z_j(l) \equiv \frac{F_j^2(l)}{C_H l^{2H}} \quad e \quad Z(l) \equiv \frac{F^2(l)}{C_H l^{2H}}, \quad (3.34)$$

onde  $C_H$  é dado pela expressão (3.30). Note que

$$Z(l) \equiv \frac{F^2(l)}{C_H l^{2H}} = \frac{1}{\lfloor \frac{n}{l} \rfloor} \sum_{j=1}^{\lfloor \frac{n}{l} \rfloor} \frac{F_j^2(l)}{C_H l^{2H}} = \frac{1}{\lfloor \frac{n}{l} \rfloor} \sum_{j=1}^{\lfloor \frac{n}{l} \rfloor} Z_j(l), \quad (3.35)$$

para todo  $l \in \{4, 5, \dots, g(n)\}$ .

Pelo Teorema 3.4, temos que

$$\lim_{l \rightarrow \infty} \frac{\mathbb{E}(F_j^2(l))}{l^{2H}} = C_H. \quad (3.36)$$

Então, da expressão (3.36) temos que

$$\lim_{l \rightarrow \infty} \mathbb{E}(Z_j(l)) = \lim_{l \rightarrow \infty} \frac{\mathbb{E}(F_j^2(l))}{C_H l^{2H}} = \frac{1}{C_H} \lim_{l \rightarrow \infty} \frac{\mathbb{E}(F_j^2(l))}{l^{2H}} = 1.$$

Portanto,

$$\mathbb{E}(Z_j(l)) \sim 1.$$

Pela Observação 3.3, temos que

$$\mathbb{E}(Z_j(l)) = 1 + o(1).$$

Pelo Corolário 3.1, obtemos que

$$\text{Var}(F_j^2(l)) = \frac{2\sigma_l^4}{l} = \frac{2(\mathbb{E}(F_j^2(l)))^2}{l}.$$

Logo,

$$\begin{aligned} \lim_{l \rightarrow \infty} \frac{\text{Var}(Z_j(l))}{\frac{2}{l}} &= \frac{l}{2} \lim_{l \rightarrow \infty} \frac{\text{Var}(F_j^2(l))}{(C_H l^{2H})^2} = \frac{l}{2} \lim_{l \rightarrow \infty} \frac{2}{l} \left( \frac{\mathbb{E}(F_j^2(l))}{C_H l^{2H}} \right)^2 \\ &= \frac{1}{(C_H)^2} \lim_{l \rightarrow \infty} \left( \frac{\mathbb{E}(F_j^2(l))}{l^{2H}} \right)^2 = \left( \frac{1}{C_H} \lim_{l \rightarrow \infty} \frac{\mathbb{E}(F_j^2(l))}{l^{2H}} \right)^2 = 1. \end{aligned}$$

Portanto,

$$\text{Var}(Z_j(l)) \sim \frac{2}{l}$$

e, pela Observação 3.3, temos que

$$\text{Var}(Z_j(l)) = \frac{2}{l}(1 + o(1)).$$

Segue-se que, quando  $l \rightarrow \infty$ , temos variáveis aleatórias independentes  $\{Z_j(l)\}_{j \geq 1}$ , com  $\mathbb{E}(Z_j(l)) = 1 + o(1)$  e  $\text{Var}(Z_j(l)) = \frac{2}{l}(1 + o(1))$ . Portanto, pelo Teorema Central do Limite, se  $\lim_{l, n \rightarrow \infty} \lfloor \frac{n}{l} \rfloor = \infty$  então,

$$\sqrt{k} \cdot \frac{\frac{1}{k} \sum_{j=1}^k Z_j(l) - 1}{\sqrt{\frac{2}{l}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \quad (3.37)$$

onde  $k = \lfloor \frac{n}{l} \rfloor$ . Pelo Teorema 3.8, temos que

$$\begin{aligned} \sqrt{k} \left( \frac{1}{k} \sum_{j=1}^k Z_j(l) - 1 \right) &\xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{2}{l}\right) \\ \Leftrightarrow \sqrt{k}(Z(l) - 1) &\xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{2}{l}\right), \end{aligned}$$

onde  $Z(l)$  é dado na expressão (3.35).

Pelo Teorema 3.8, tomando-se  $g(x) = \ln(x)$  e  $u = 1$ , tem-se  $g'(u)|_{u=1} = \frac{1}{u}|_{u=1} = 1$ , segue-se que

$$\begin{aligned} \sqrt{k}[\ln(Z(l))] &\xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{2}{l}\right) \Leftrightarrow \sqrt{k} \frac{[\ln(Z(l))]}{\sqrt{\frac{2}{l}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \\ \Leftrightarrow \sqrt{\frac{k \cdot l}{2}}[\ln(Z(l))] &\xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \Leftrightarrow \sqrt{\frac{\tilde{n}}{2}}[\ln(Z(l))] \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \end{aligned}$$

onde  $\tilde{n} = k \cdot l$  e  $k = \lfloor \frac{n}{l} \rfloor$ .

Nas equações a seguir utilizamos o símbolo  $\approx$  para indicar aproximadamente igual. Quando  $\lim_{l, n \rightarrow \infty} \lfloor \frac{n}{l} \rfloor = \infty$  temos então que

$$\sqrt{\frac{\tilde{n}}{2}}[\ln(Z(l))] \approx \tilde{\epsilon}_l, \quad (3.38)$$

onde  $\tilde{\epsilon}_l \sim \mathcal{N}(0, 1)$ . Note que a expressão (3.38) é equivalente a

$$\ln(Z(l)) \approx \frac{\sqrt{2}}{\sqrt{\tilde{n}}} \tilde{\epsilon}_l = \frac{2}{\sqrt{2\tilde{n}}} \tilde{\epsilon}_l.$$

Como  $\tilde{n} \approx n$ , temos que

$$\frac{\ln(Z(l))}{2} \approx \frac{1}{\sqrt{2\tilde{n}}} \tilde{\epsilon}_l \approx \frac{1}{\sqrt{2n}} \tilde{\epsilon}_l. \quad (3.39)$$

Como  $\tilde{\epsilon}_l$  tem distribuição  $\mathcal{N}(0, 1)$ , segue-se que

$$\epsilon_l := \frac{1}{\sqrt{2n}} \tilde{\epsilon}_l, \quad (3.40)$$

tem distribuição  $\mathcal{N}(0, \frac{1}{2n})$ . Portanto, pelas expressões (3.39) e (3.40)

$$\frac{\ln(Z(l))}{2} \approx \epsilon_l,$$

onde  $\epsilon_l$  tem distribuição  $\mathcal{N}(0, \frac{1}{2n})$ .

Aplicando logaritmo em ambos os lados da expressão  $Z(l)$ , dada em (3.34), temos que

$$\begin{aligned} \ln(Z(l)) &= \ln(F^2(l)) - \ln(C_H l^{2H}) \\ \Leftrightarrow \ln(F^2(l)) &= \ln(Z(l)) + \ln(C_H l^{2H}) \\ \Leftrightarrow 2 \ln(F(l)) &= \ln(Z(l)) + \ln(C_H) + 2H \ln(l) \\ \Leftrightarrow \ln(F(l)) &= \frac{\ln(C_H)}{2} + H \ln(l) + \frac{\ln(Z(l))}{2}. \end{aligned}$$

Se  $l \in \{4, 5, \dots, g(n)\}$ , então  $l \leq g(n)$ . Assim  $\lfloor \frac{n}{g(n)} \rfloor \leq \lfloor \frac{n}{l} \rfloor$ . Sob a hipótese de que  $\lim_{g(n), n \rightarrow \infty} \left\lfloor \frac{n}{g(n)} \right\rfloor = \infty$ , temos que  $\lim_{l, n \rightarrow \infty} \left\lfloor \frac{n}{l} \right\rfloor = \infty$ . Portanto,

$$\ln(F(l)) \approx \frac{\ln(C_H)}{2} + H \ln(l) + \epsilon_l,$$

onde  $\epsilon_l$  tem distribuição  $\mathcal{N}(0, \frac{1}{2n})$ .

Logo, pelo Teorema 3.6 obtemos a parte (a). A demonstração da parte (b) segue do Teorema 3.3. □

O corolário a seguir, sob certas condições, determina uma escolha ótima para  $g(n)$ . Conseqüentemente, fornece uma escolha ótima assintótica para o número de regressores  $m$  na equação (3.6).

**Corolário 3.2.** *Seja  $\{X_t\}_{t \in \mathbb{R}^+}$  um processo ruído Gaussiano fracionário. Considere  $\{X_t\}_{t=1}^n$  uma série temporal advinda deste processo. Suponha que, para todo  $l \in \{4, 5, \dots, g(n)\}$ , as variáveis aleatórias  $Z_1^l, Z_2^l, \dots, Z_{\tilde{n}}^l$ , dadas pela expressão (3.2), sejam independentes e identicamente distribuídas com função de distribuição comum  $\mathcal{N}(0, \sigma_l^2)$ . Suponha que  $g(n) = \lfloor n^\beta \rfloor$ ,  $0 < \beta < 1$ . Então,  $\hat{H}_{\text{DFA}}$ , dado pela expressão (3.8), é um estimador*

(a) U.M.V.U. para o parâmetro  $H$ ,

(b) com função de distribuição Gaussiana, com esperança e variância dadas, respectivamente, por

$$\mathbb{E}(\hat{H}_{\text{DFA}}) = H \quad e \quad \text{Var}(\hat{H}_{\text{DFA}}) = \frac{1}{2n \sum_{j=1}^{g(n)-3} (x_j - \bar{x})^2}, \quad (3.41)$$

onde  $x_j = \ln(j+3)$ , para  $j \in \{1, \dots, g(n)-3\}$  e  $g(n) = \lfloor n^\beta \rfloor$ .

**Demonstração:** Considere  $g(n) = \lfloor n^\beta \rfloor$ . Então,

$$\left\lfloor \frac{n}{g(n)} \right\rfloor = \left\lfloor \frac{n}{n^\beta} \right\rfloor = \lfloor n^{1-\beta} \rfloor.$$

Como  $\lim_{g(n), n \rightarrow \infty} \left\lfloor \frac{n}{g(n)} \right\rfloor = \lim_{g(n), n \rightarrow \infty} \lfloor n^{1-\beta} \rfloor = \infty$ , o resultado segue do Teorema 3.9. □

### 3.1.2 Erros i.i.d. com Distribuição Não Gaussiana

O modelo usual de regressão linear simples é dado pela expressão (3.6), com  $\mathbb{E}(\epsilon_j) = 0$  e  $\text{Var}(\epsilon_j) = \sigma_\epsilon^2$ . Para obter propriedades estatísticas deste modelo, é necessário assumir condições sobre os erros  $\epsilon_j$ . Como os  $\epsilon_j$ 's são não observados, isto deve ser feito indiretamente utilizando os resíduos. Para o modelo linear de mínimos quadrados, o vetor dos resíduos ordinários  $\mathbf{e} \equiv (e_1, e_2, \dots, e_m)$  é dado por

$$e_j = y_j - \hat{y}_j, \quad j \in \{1, 2, \dots, m\}. \quad (3.42)$$

A relação entre  $e_j$  e  $\epsilon_j$ , para cada  $j \in \{1, 2, \dots, m\}$ , é dada por

$$e_j = \epsilon_j - \sum_{k=1}^m v_{jk} \epsilon_k, \quad j \in \{1, 2, \dots, m\}, \quad (3.43)$$

onde

$$\mathbf{V} = (v_{jk})_{j,k=1}^m \equiv \mathbf{X}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X},$$

com  $\mathbf{A}'$  a transposta da matriz  $\mathbf{A}$ .

**Observação 3.4.** Para um modelo de regressão linear simples, dado pela expressão (3.6), tem-se que os elementos da diagonal da matriz  $\mathbf{V}$  são dados por

$$v_{jj} = \frac{1}{m} + \frac{(x_j - \bar{x})^2}{\sum_{k=1}^m (x_k - \bar{x})^2}, \quad j \in \{1, 2, \dots, m\} \quad (3.44)$$

(ver Cook e Weisberg, 1994).

A expressão (3.43) mostra que a relação entre  $\mathbf{e}$  e  $\boldsymbol{\epsilon}$  depende somente de  $\mathbf{V}$ . Se os  $v_{jk}$ 's são suficientemente pequenos,  $\mathbf{e}$  servirá como um substituto razoável para  $\boldsymbol{\epsilon}$ , caso contrário a utilidade do  $\mathbf{e}$  pode ser limitada.

A distribuição do vetor dos resíduos ordinários  $\mathbf{e}$  segue-se imediatamente da expressão (3.43), ou seja, se  $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_\epsilon^2)$  então  $\mathbf{e}$  tem uma distribuição Gaussiana singular (ver Definição 2.6) com  $\mathbb{E}(\mathbf{e}) = 0$  e  $\text{Var}(\mathbf{e}) = (\mathbf{I} - \mathbf{V})\sigma_\epsilon^2$ .

**Proposição 3.4.** *Sejam  $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros e  $\boldsymbol{\beta} = (a, b)$  o vetor dos parâmetros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com uma certa função de distribuição não Gaussiana de média 0 e variância  $\sigma_\epsilon^2$ . Uma condição necessária e suficiente para o estimador de mínimos quadrados de qualquer combinação linear  $d_1\hat{a} + d_2\hat{b}$  ser assintoticamente normal, com média  $d_1a + d_2b$  e variância  $\sigma_\epsilon^2$  é que  $\lim_{m \rightarrow \infty} \max_{1 \leq j \leq m} v_{jj} = 0$ .*

**Demonstração:** Ver Huber (1973), página 803. □

O teorema a seguir apresenta uma condição necessária e suficiente para que o estimador  $\hat{\mathbf{H}}_{\text{DFA}}$ , dado na expressão (3.8), seja assintoticamente normal, quando a sequência  $\{\epsilon_j\}_{j=1}^m$  dada em (3.6) é formada por variáveis aleatórias independentes e identicamente distribuídas com uma certa função de distribuição não Gaussiana.

**Teorema 3.10.** *Seja  $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com uma certa função de distribuição não Gaussiana de média 0 e variância  $\sigma_\epsilon^2$ . Uma condição necessária e suficiente para que o estimador  $\hat{\mathbf{H}}_{\text{DFA}}$ , dado na expressão (3.8) seja assintoticamente normal, com esperança  $\mathbf{H}$  e variância  $\sigma_\epsilon^2$  é que  $\lim_{m \rightarrow \infty} \max_{4 \leq j \leq m+3} v_{jj} = 0$ , onde*

$$v_{jj} = \frac{1}{m} + \frac{(\ln(j) - \bar{X}_m)^2}{\sum_{k=1}^m (\ln(k+3) - \bar{X}_m)^2} \quad (3.45)$$

$$e \bar{X}_m = \frac{1}{m} \sum_{k=1}^m \ln(k+3), \text{ para todo } j \in \{4, 5, \dots, m+3\}.$$

**Demonstração:** Pelas equações (3.6) e (3.7), o estimador dos mínimos quadrados é dado por  $\hat{b} = \hat{H}_{\text{DFA}}$ . Pela Proposição 3.4 e substituindo (3.7) em (3.44), segue-se o resultado.  $\square$

**Lema 3.2.** Para todo  $m \in \mathbb{N}$ ,  $m \geq 1$ , tem-se que

$$4^{m-2}(m+3)^m \geq 4^{m-1}(m+4)^{m-1}. \quad (3.46)$$

**Demonstração:** Basta provar que

$$\frac{(m+4)^{m-1}}{(m+3)^m} \leq \frac{1}{4}, \quad \forall m \in \mathbb{N}, m \geq 1.$$

Consideramos a sequência de reais positivos  $X_m = \frac{(m+4)^{m-1}}{(m+3)^m}$ . É fácil ver que a sequência  $(X_m)_{m \geq 1}$  é decrescente, pois

$$\frac{X_{m+1}}{X_m} = \frac{(m+5)^m}{(m+4)^{m+1}} \cdot \frac{(m+3)^m}{(m+4)^{m-1}} = \left( \frac{(m+5)(m+3)}{(m+4)^2} \right)^m.$$

Como

$$\frac{(m+5)(m+3)}{(m+4)^2} = \frac{m^2 + 8m + 15}{m^2 + 8m + 16} < 1,$$

segue-se que  $\frac{X_{m+1}}{X_m} < 1$ , ou seja,  $(X_m)_{m \geq 1}$  é sequência decrescente. Portanto,  $X_m \leq X_1 = \frac{1}{4}$ , para todo  $m \in \mathbb{N}$ ,  $m \geq 1$ , provando o lema.  $\square$

O Lema 3.3, apresenta uma ferramenta básica, para a prova do teorema que apresentaremos a seguir.

**Lema 3.3.** Para todo  $m \in \mathbb{N}$ , tem-se que

$$4^2 \cdot 5^2 \cdot \dots \cdot (m+3)^2 \geq 4^m (m+3)^m. \quad (3.47)$$

**Demonstração:** A prova será realizada por indução matemática. Observe que a desigualdade (3.47) vale para  $m = 1$ . De fato,

$$(1+3)^2 = 4^1(1+3)^1.$$

Vamos mostrar que, se a expressão (3.47) vale para  $k$ , então valerá para  $k+1$ . Pela hipótese de indução, temos que

$$4^2 \cdot 5^2 \cdot \dots \cdot (k+2)^2 \cdot (k+3)^2 \geq 4^k (k+3)^k. \quad (3.48)$$

Pelo Lema 3.2, o lado direito da desigualdade (3.48) é dado por

$$4^k (k+3)^k = 4^2 \cdot 4^{k-2} (k+3)^k \geq 4^2 \cdot 4^{k-1} (k+4)^{k-1} = 4^{k+1} (k+4)^{k-1}. \quad (3.49)$$



Portanto, pelas expressões (3.48) e (3.49), segue-se que

$$4^2 \cdot 5^2 \cdot \dots \cdot (k+2)^2 \cdot (k+3)^2 \geq 4^{k+1}(k+4)^{k-1}. \quad (3.50)$$

Multiplicando ambos os lados da desigualdade (3.50) por  $(k+4)^2$  temos

$$4^2 \cdot 5^2 \cdot \dots \cdot (k+2)^2 \cdot (k+3)^2 \cdot (k+4)^2 \geq 4^{k+1}(k+4)^{k+1},$$

valendo a desigualdade (3.47) para  $k+1$ . Logo, por indução, vale a desigualdade (3.47), para todo  $m \geq 1$ . □

**Lema 3.4.** *Considere as componentes da diagonal da matriz  $\mathbf{V} = (v_{jk})_{j,k=1}^m$  definidas na expressão (3.45), para todo  $j \in \{4, 5, \dots, m+3\}$ . Então, para todo  $m \geq 1$ , tem-se que*

$$\max_{4 \leq j \leq m+3} v_{jj} = v_{44}. \quad (3.51)$$

**Demonstração:** Note que, na expressão (3.45), para obtermos o máximo dos  $v_{jj}$ , para todo  $j \in \{4, 5, \dots, m+3\}$ , precisamos apenas da parte  $(\ln(j) - \bar{X}_m)^2$ , pois o restante da expressão não depende de  $j$ . Logo, para obtermos a expressão (3.51), basta provar que

$$\max_{4 \leq j \leq m+3} (\ln(j) - \bar{X}_m)^2 = (\ln(4) - \bar{X}_m)^2, \quad (3.52)$$

onde  $\bar{X}_m = \frac{1}{m} \sum_{k=1}^m \ln(k+3)$ .

Sabemos que a função logaritmo é monótona crescente. Assim,

$$\ln(4) \leq \ln(j) \leq \ln(m+3),$$

para todo  $j \in \{4, 5, \dots, m+3\}$ . Logo,

$$\ln(4) \leq \bar{X}_m \leq \ln(m+3).$$

Note que,  $\ln(j) \leq \bar{X}_m$  ou  $\bar{X}_m \leq \ln(j)$ , para todo  $j \in \{4, 5, \dots, m+3\}$ . Considerando  $j \in \{4, 5, \dots, m+3\}$ , tal que  $\ln(j) \leq \bar{X}_m$ , tem-se que

$$\ln(4) \leq \ln(j) \leq \bar{X}_m. \quad (3.53)$$

Adicionando  $(-\bar{X}_m)$  na expressão (3.53), obtemos

$$\ln(4) - \bar{X}_m \leq \ln(j) - \bar{X}_m \leq 0.$$

Portanto,

$$(\ln(j) - \bar{X}_m)^2 \leq (\ln(4) - \bar{X}_m)^2. \quad (3.54)$$

Considerando  $j \in \{4, 5, \dots, m+3\}$ , tal que  $\bar{X}_m \leq \ln(j)$ , tem-se que

$$\bar{X}_m \leq \ln(j) \leq \ln(m+3). \quad (3.55)$$

Adicionando  $(-\bar{X}_m)$  na expressão (3.55), obtemos

$$0 \leq \ln(j) - \bar{X}_m \leq \ln(m+3) - \bar{X}_m.$$

Portanto,

$$(\ln(j) - \bar{X}_m)^2 \leq (\ln(m+3) - \bar{X}_m)^2. \quad (3.56)$$

Pelas expressões (3.54) e (3.56), temos

$$\max_{4 \leq j \leq m+3} (\ln(j) - \bar{X}_m)^2 = (\ln(4) - \bar{X}_m)^2,$$

ou

$$\max_{4 \leq j \leq m+3} (\ln(j) - \bar{X}_m)^2 = (\ln(m+3) - \bar{X}_m)^2.$$

Então, para provar a expressão (3.52), é suficiente obter que

$$(\ln(4) - \bar{X}_m)^2 \geq (\ln(m+3) - \bar{X}_m)^2. \quad (3.57)$$

Pelo Lema 3.3, para todo  $m \geq 1$ , tem-se que

$$4^2 \cdot 5^2 \cdot \dots \cdot (m+3)^2 \geq 4^m (m+3)^m,$$

ou seja,

$$\begin{aligned} & [4 \cdot 5 \cdot \dots \cdot (m+3)]^{\frac{2}{m}} \geq 4(m+3) \\ & \Leftrightarrow \ln([4 \cdot 5 \cdot \dots \cdot (m+3)]^{\frac{2}{m}}) \geq \ln(4(m+3)) \\ & \Leftrightarrow \frac{2}{m} \sum_{k=1}^m \ln(k+3) \geq \ln(4(m+3)) \Leftrightarrow 2\bar{X}_m \geq \ln(4) + \ln(m+3) \\ & \Leftrightarrow \ln(m+3) - \bar{X}_m \leq \bar{X}_m - \ln(4), \end{aligned}$$

onde  $\bar{X}_m = \frac{1}{m} \sum_{k=1}^m \ln(k+3)$ . Como  $\ln(m+3) \geq \bar{X}_m$ , ou seja,  $\ln(m+3) - \bar{X}_m \geq 0$ , segue-se que

$$(\ln(m+3) - \bar{X}_m)^2 \leq (\bar{X}_m - \ln(4))^2 = (\ln(4) - \bar{X}_m)^2.$$

Logo, vale a expressão (3.57) e, portanto,  $\max_{4 \leq j \leq m+3} v_{jj} = v_{44}$ .

□

O Teorema 3.11, a ser apresentado a seguir, fornece a normalidade assintótica do estimador  $\hat{H}_{\text{DFA}}$ , dado pela expressão (3.8), quando a sequência  $\{\epsilon_j\}_{j=1}^m$  dada em (3.6) é formada por variáveis aleatórias independentes e identicamente distribuídas com uma certa função de distribuição não Gaussiana. Para a prova do teorema, necessitamos do cálculo de diversos limites que serão dados na observação a seguir.

**Observação 3.5.** Apresentamos, a seguir, o cálculo dos limites que aparecem na prova do Teorema 3.11.

(i)  $\lim_{x \rightarrow 0^+} \frac{(\ln(x) - 1)^2 + 1}{\frac{1}{x}} = 0$ . De fato, utilizando a regra de L'Hôpital duas vezes, temos que

$$\begin{aligned} \lim_{x \rightarrow 0^+} \frac{(\ln(x) - 1)^2 + 1}{\frac{1}{x}} &= \lim_{x \rightarrow 0^+} \frac{2(\ln(x) - 1) \cdot \frac{1}{x}}{-\frac{1}{x^2}} \\ &= \lim_{x \rightarrow 0^+} \frac{2(\ln(x) - 1)}{-\frac{1}{x}} \\ &= \lim_{x \rightarrow 0^+} \frac{2 \cdot \frac{1}{x}}{\frac{1}{x^2}} = \lim_{x \rightarrow 0^+} 2x = 0. \end{aligned} \quad (3.58)$$

(ii)  $\lim_{m \rightarrow \infty} \frac{2}{m+7} \left( \ln \left( \frac{k}{m+7} \right) \right)^2 = 0$ , para todo  $k \in \mathbb{R}^+$ . De fato, utilizando a regra de L'Hôpital duas vezes, temos que

$$\begin{aligned} \lim_{m \rightarrow \infty} \frac{2}{m+7} \left( \ln \left( \frac{k}{m+7} \right) \right)^2 &= \lim_{m \rightarrow \infty} \frac{2 \left( \ln \left( \frac{k}{m+7} \right) \right)^2}{m+7} \\ &= \lim_{m \rightarrow \infty} 4 \cdot \ln \left( \frac{k}{m+7} \right) \cdot \frac{m+7}{k} \cdot \frac{(-k)}{(m+7)^2} \\ &= \lim_{m \rightarrow \infty} \frac{(-4) \cdot \ln \left( \frac{k}{m+7} \right)}{m+7} \\ &= \lim_{m \rightarrow \infty} (-4) \cdot \frac{m+7}{k} \cdot \frac{(-k)}{(m+7)^2} \\ &= \lim_{m \rightarrow \infty} \frac{4}{m+7} = 0. \end{aligned} \quad (3.59)$$

(iii)  $\lim_{m \rightarrow \infty} \left( \frac{2}{m+7} \cdot S_1 \right) = 0$ , onde

$$S_1 = \left( \ln \left( \frac{2 \cdot 1}{m+7} \right) \right)^2 + \left( \ln \left( \frac{2 \cdot 2}{m+7} \right) \right)^2 + \left( \ln \left( \frac{2 \cdot 3}{m+7} \right) \right)^2. \quad (3.60)$$

De fato, pelo item (ii), desta observação, segue-se que

$$\begin{aligned} \lim_{m \rightarrow \infty} \left( \frac{2}{m+7} \cdot S_1 \right) &= \lim_{m \rightarrow \infty} \frac{2}{m+7} \left[ \left( \ln \left( \frac{2}{m+7} \right) \right)^2 + \left( \ln \left( \frac{4}{m+7} \right) \right)^2 \right. \\ &\quad \left. + \left( \ln \left( \frac{6}{m+7} \right) \right)^2 \right] = 0. \end{aligned} \quad (3.61)$$

(iv)  $\lim_{m \rightarrow \infty} \frac{2}{m+7} \left( \ln \left( \frac{2(m+k)}{m+7} \right) \right)^2 = 0$ , para todo  $k \in \mathbb{R} \cap (-m, \infty)$ . De fato,

$$\begin{aligned} \lim_{m \rightarrow \infty} \frac{2}{m+7} \left( \ln \left( \frac{2(m+k)}{m+7} \right) \right)^2 &= \lim_{m \rightarrow \infty} \frac{2}{m+7} \cdot \lim_{m \rightarrow \infty} \left( \ln \left( \frac{2 + \frac{2k}{m}}{1 + \frac{7}{m}} \right) \right)^2 \\ &= 0 \cdot (\ln(2))^2 = 0. \end{aligned} \quad (3.62)$$

(v)  $\lim_{m \rightarrow \infty} \left( \frac{2}{m+7} \cdot S_2 \right) = 0$ , onde

$$\begin{aligned} S_2 &= \left( \ln \left( \frac{2(m+4)}{m+7} \right) \right)^2 + \left( \ln \left( \frac{2(m+5)}{m+7} \right) \right)^2 \\ &\quad + \left( \ln \left( \frac{2(m+6)}{m+7} \right) \right)^2 + (\ln(2))^2. \end{aligned} \quad (3.63)$$

De fato, pelo item (iv), desta observação, segue-se que

$$\lim_{m \rightarrow \infty} \left( \frac{2}{m+7} \cdot S_2 \right) = \lim_{m \rightarrow \infty} \frac{2}{m+7} \left[ \left( \ln \left( \frac{2(m+4)}{m+7} \right) \right)^2 \right. \quad (3.64)$$

$$\begin{aligned} &\quad + \left( \ln \left( \frac{2(m+5)}{m+7} \right) \right)^2 \\ &\quad \left. + \left( \ln \left( \frac{2(m+6)}{m+7} \right) \right)^2 + (\ln(2))^2 \right] = 0. \end{aligned} \quad (3.65)$$

(vi)  $\lim_{m \rightarrow \infty} \frac{\ln(m)}{\sqrt{m+7}} = 0$ . De fato, utilizando a regra de L'Hôpital duas vezes, temos que

$$\begin{aligned} \lim_{m \rightarrow \infty} \frac{\ln(m)}{\sqrt{m+7}} &= \lim_{m \rightarrow \infty} \frac{\frac{1}{m}}{\frac{1}{2\sqrt{m+7}}} = 2 \lim_{m \rightarrow \infty} \frac{\sqrt{m+7}}{m} \\ &= \lim_{m \rightarrow \infty} \frac{1}{\sqrt{m+7}} = 0. \end{aligned} \quad (3.66)$$

(vii)  $\lim_{m \rightarrow \infty} \frac{\ln[(m+1)(m+2)(m+3)]}{m\sqrt{m+7}} = 0$ . De fato, utilizando a regra de

L'Hôpital uma vez, temos que

$$\begin{aligned}
\lim_{m \rightarrow \infty} \frac{\ln[(m+1)(m+2)(m+3)]}{m\sqrt{m+7}} &= \lim_{m \rightarrow \infty} \frac{\ln[m^3 + 6m^2 + 11m + 6]}{m\sqrt{m+7}} \\
&= \lim_{m \rightarrow \infty} \frac{3m^2 + 12m + 11}{m^3 + 6m^2 + 11m + 6} \cdot \frac{1}{\sqrt{m+7} + \frac{m}{2\sqrt{m+7}}} \\
&= \lim_{m \rightarrow \infty} \frac{3m^2 + 12m + 11}{m^3 + 6m^2 + 11m + 6} \cdot \frac{2\sqrt{m+7}}{2(m+7) + m} \\
&= \lim_{m \rightarrow \infty} \frac{(3m^2 + 12m + 11)(2\sqrt{m+7})}{(m^3 + 6m^2 + 11m + 6)(3m + 14)} = 0.
\end{aligned}$$

**Teorema 3.11.** *Seja  $\epsilon = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com uma certa função de distribuição não Gaussiana com média 0 e variância  $\sigma_\epsilon^2$ . Então, o estimador  $\hat{H}_{\text{DFA}}$ , dado em (3.8), é assintoticamente normal com média  $H$  e variância  $\sigma_\epsilon^2$ .*

**Demonstração:** Pelo Teorema 3.10 e pelo Lema 3.4, vistos anteriormente, basta provar que  $\lim_{m \rightarrow \infty} v_{44} = 0$ , ou seja, é suficiente provar que

$$\lim_{m \rightarrow \infty} \frac{(\ln(4) - \bar{X}_m)^2}{\sum_{k=1}^m (\ln(k+3) - \bar{X}_m)^2} = 0,$$

onde  $\bar{X}_m = \frac{1}{m} \sum_{k=1}^m \ln(k+3)$ .

Note que, para todo  $k \in \{4, 5, \dots, m+3\}$ ,

$$\begin{aligned}
\ln(k) - \bar{X}_m &= \ln(k) - \ln[4 \cdot 5 \cdot \dots \cdot (m+3)]^{\frac{1}{m}} \\
&= \ln\left(\frac{k}{[4 \cdot 5 \cdot \dots \cdot (m+3)]^{\frac{1}{m}}}\right). \tag{3.67}
\end{aligned}$$

Pela desigualdade entre as médias aritmética e geométrica tem-se que

$$\begin{aligned}
[4 \cdot 5 \cdot \dots \cdot (m+3)]^{\frac{1}{m}} &< \frac{4 + 5 + \dots + (m+3)}{m} \\
&= \frac{(m+7)m}{2m} = \frac{m+7}{2}. \tag{3.68}
\end{aligned}$$

Logo, pelas expressões (3.67) e (3.68) obtemos

$$\ln(k) - \bar{X}_m \geq \ln\left(\frac{2k}{m+7}\right), \quad \text{para todo } k \in \{4, 5, \dots, m+3\}.$$

Assim,

$$\sum_{k=1}^m (\ln(k+3) - \bar{X}_m)^2 \geq \sum_{k=1}^m \left( \ln\left(\frac{2(k+3)}{m+7}\right) \right)^2.$$

Logo,

$$L = \frac{(\ln(4) - \bar{X}_m)^2}{\sum_{k=1}^m (\ln(k+3) - \bar{X}_m)^2} < \frac{(\ln(4) - \bar{X}_m)^2}{\sum_{k=1}^m \left( \ln\left(\frac{2(k+3)}{m+7}\right) \right)^2}.$$

Observe que se

$$\lim_{m \rightarrow \infty} \frac{(\ln(4) - \bar{X}_m)^2}{\sum_{k=1}^m \left( \ln\left(\frac{2(k+3)}{m+7}\right) \right)^2} = 0, \quad (3.69)$$

então

$$\lim_{m \rightarrow \infty} \frac{(\ln(4) - \bar{X}_m)^2}{\sum_{k=1}^m (\ln(k+3) - \bar{X}_m)^2} = 0.$$

Portanto, basta provar a expressão (3.69). Note que

$$\begin{aligned} \sum_{k=1}^m \left( \ln\left(\frac{2(k+3)}{m+7}\right) \right)^2 &= \sum_{k=1}^{m+7} \left( \ln\left(\frac{2 \cdot k}{m+7}\right) \right)^2 - (S_1 + S_2) \\ &= \frac{m+7}{2} \cdot \left( \frac{2}{m+7} \sum_{k=1}^{m+7} \left( \ln\left(\frac{2 \cdot k}{m+7}\right) \right)^2 - \frac{2}{m+7} (S_1 + S_2) \right), \end{aligned} \quad (3.70)$$

onde  $S_1$  e  $S_2$  são dados, respectivamente, pelas expressões (3.60) e (3.63). Utilizando o limite de somas de Riemman da função  $\ln^2(x)$  no intervalo  $(0, 2]$ , com partição  $\frac{2}{m+7}$ , e o item (i) da Observação 3.5, obtemos

$$\begin{aligned} \lim_{m \rightarrow \infty} \frac{2}{m+7} \sum_{k=1}^{m+7} \left( \ln\left(\frac{2 \cdot k}{m+7}\right) \right)^2 &= \int_{0+}^2 (\ln(x))^2 dx \\ &= x[(\ln(x) - 1)^2 + 1] \Big|_{0+}^2 \\ &= 2[(\ln(2) - 1)^2 + 1] - \lim_{x \rightarrow 0+} \frac{(\ln(x) - 1)^2 + 1}{\frac{1}{x}} \\ &= 2[(\ln(2) - 1)^2 + 1] + 0 \approx 2.188317. \end{aligned} \quad (3.71)$$

Por outro lado, pelos itens (iii) e (v) da Observação 3.5, temos que

$$\lim_{m \rightarrow \infty} \left( \frac{2}{m+7} \cdot S_1 \right) = 0 \quad \text{e} \quad \lim_{m \rightarrow \infty} \left( \frac{2}{m+7} \cdot S_2 \right) = 0. \quad (3.72)$$

Finalmente, pelas expressões (3.70), (3.71) e (3.72), segue-se que

$$\begin{aligned}
& \lim_{m \rightarrow \infty} \frac{(\ln(4) - \bar{X}_m)^2}{\sum_{k=1}^m \left( \ln \left( \frac{2(k+3)}{m+7} \right) \right)^2} \\
&= \lim_{m \rightarrow \infty} \frac{(\ln(4) - \bar{X}_m)^2}{\frac{m+7}{2} \left( \frac{2}{m+7} \sum_{k=1}^{m+7} \left( \ln \left( \frac{2 \cdot k}{m+7} \right) \right)^2 - \frac{2}{m+7} (S_1 + S_2) \right)} \\
&= \frac{1}{2.188317} \lim_{m \rightarrow \infty} \frac{2(\ln(4) - \bar{X}_m)^2}{m+7} \\
&= 1.094159 \lim_{m \rightarrow \infty} \frac{(\ln(4) - \bar{X}_m)^2}{m+7}. \tag{3.73}
\end{aligned}$$

Note que

$$\begin{aligned}
\ln(4) - \bar{X}_m &= \ln(4) - \frac{1}{m} \sum_{k=1}^m \ln(k+3) \\
&= \ln(4) - \frac{1}{m} \sum_{k=1}^m \ln(k) + \left( \frac{1}{m} \ln(1) + \frac{1}{m} \ln(2) + \frac{1}{m} \ln(3) \right) \\
&\quad - \left( \frac{1}{m} \ln(m+1) + \frac{1}{m} \ln(m+2) + \frac{1}{m} \ln(m+3) \right) \\
&= \ln(4) - \frac{1}{m} \sum_{k=1}^m \ln \left( \frac{k}{m} \right) + \ln(m) + \left( \frac{1}{m} \ln(1) + \frac{1}{m} \ln(2) + \frac{1}{m} \ln(3) \right) \\
&\quad - \left( \frac{1}{m} \ln(m+1) + \frac{1}{m} \ln(m+2) + \frac{1}{m} \ln(m+3) \right) \\
&= \ln(4) - \int_{0^+}^1 \ln(x) dx + \ln(m) + \left( \frac{1}{m} \ln(1) + \frac{1}{m} \ln(2) + \frac{1}{m} \ln(3) \right) \\
&\quad - \left( \frac{1}{m} \ln(m+1) + \frac{1}{m} \ln(m+2) + \frac{1}{m} \ln(m+3) \right) \\
&= \ln(4) + 1 + \ln(m) + \frac{\ln(1 \cdot 2 \cdot 3)}{m} - \frac{\ln[(m+1)(m+2)(m+3)]}{m}. \tag{3.74}
\end{aligned}$$

Utilizando a expressão (3.74), temos que

$$\begin{aligned}
\lim_{m \rightarrow \infty} \frac{(\ln(4) - \bar{X}_m)^2}{m+7} &= \lim_{m \rightarrow \infty} \left( \frac{\ln(4) - \bar{X}_m}{\sqrt{m+7}} \right)^2 \\
&= \left( \lim_{m \rightarrow \infty} \frac{\ln(4) + 1 + \ln(m) + \frac{\ln(1 \cdot 2 \cdot 3)}{m} - \frac{\ln[(m+1)(m+2)(m+3)]}{m}}{\sqrt{m+7}} \right)^2 \\
&= \left( \lim_{m \rightarrow \infty} \frac{\ln(4) + 1}{\sqrt{m+7}} + \lim_{m \rightarrow \infty} \frac{\ln(m)}{\sqrt{m+7}} + \lim_{m \rightarrow \infty} \frac{\ln(6)}{m\sqrt{m+7}} \right. \\
&\quad \left. - \lim_{m \rightarrow \infty} \frac{\ln[(m+1)(m+2)(m+3)]}{\sqrt{m+7}} \right)^2 = 0. \tag{3.75}
\end{aligned}$$

Para o cálculo de alguns limites dados na expressão (3.75), ver itens (vi) e (vii) da Observação 3.5.

Logo, pelas expressões (3.73) e (3.75), temos que

$$\lim_{m \rightarrow \infty} \frac{(\ln(4) - \bar{X}_m)^2}{\sum_{k=1}^m \left( \ln \left( \frac{2(k+3)}{m+7} \right) \right)^2} = 0.$$

Portanto,

$$\lim_{m \rightarrow \infty} \frac{(\ln(4) - \bar{X}_m)^2}{\sum_{k=1}^m (\ln(k+3) - \bar{X}_m)^2} = 0$$

e concluímos a demonstração. □

Apresentamos no Anexo A, algumas propriedades adicionais para o método DFA. Sob certas condições, mostramos que a função densidade da variável aleatória  $Y = \ln(F(l))$ , forma uma família exponencial de um ou dois parâmetros, onde  $F(l)$  é dada pela expressão (3.3). Utilizando a informação de Fisher (ver Definição A.2), calculamos a variância do estimador  $\hat{H}_{\text{DFA}}$ , dado na expressão (3.8).

## 3.2 Simulações

Nesta seção apresentamos uma investigação do efeito de  $g(n)$  no método DFA, apresentado no início do Capítulo 3. Apresentamos também uma comparação entre os estimadores  $\hat{d}_{\text{DFA}}$  (ver expressão (3.9)) e  $\hat{d}_{\text{W}}$  (ver Subseção 2.2.2). Os resultados foram obtidos de simulações utilizando programas na linguagem R.



### 3.2.1 Investigação do Efeito de $g(n)$ no Método DFA

Para investigar o efeito de  $g(n) = \lfloor n^\beta \rfloor$ , para algum  $0 < \beta < 1$ , dado na expressão (3.6), no método da análise de flutuações destendenciadas, geramos séries temporais obtidas de processos ARFIMA(0,  $d$ , 0) e ruído Gaussiano fracionário (FGN), com 200 replicações e tamanho  $n \in \{1000, 5000, 10000\}$ .

Consideramos diferentes valores para  $g(n)$ , incluindo o valor mais utilizado na literatura, dado por  $\lfloor \frac{n}{10} \rfloor$  (ver Hu et al., 2001). Para cada  $g(n)$ , calculamos o valor médio ( $\hat{d}_{\text{DFA}}$ ), o vício, o erro quadrático médio (eqm) e a variância (var). Os resultados dos experimentos, são dados nas Tabelas 3.1 a 3.4.

As Tabelas 3.1 e 3.2, apresentam os resultados para séries temporais de processos ARFIMA(0,  $d$ , 0) com  $d \in \{0.1, 0.2, 0.3, 0.4\}$ , enquanto que as Tabelas 3.3 e 3.4, apresentam resultados para séries temporais obtidas de processos ruído Gaussiano fracionário (FGN) com  $H \in \{0.6, 0.7, 0.8, 0.9\}$ .

Podemos observar nas Tabelas 3.1 a 3.4, que para cada  $g(n)$  fixo, quando o tamanho da amostra cresce, o erro quadrático médio e a variância tendem a zero. Em nosso experimento,  $g(n) = \lfloor n^{0.3} \rfloor$ , parece gerar maior vício para as estimativas dos parâmetros  $H$  e  $d$ . Excessão ocorre quando  $n = 5000$  e  $d = 0.4$  onde  $g(n) = \lfloor n^{0.3} \rfloor$  apresenta melhores resultados. Observe que a escolha mais utilizada na literatura,  $g(n) = \lfloor \frac{n}{10} \rfloor$ , gera boas estimativas para os parâmetros  $H$  e  $d$ . No entanto,  $g(n) = \lfloor n^\beta \rfloor$  apresenta melhores resultados do que  $g(n) = \lfloor \frac{n}{10} \rfloor$ , para  $d \in \{0.1, 0.2, 0.3, 0.4\}$  se  $\beta \in \{0.5, 0.6, 0.7\}$  e para  $H \in \{0.6, 0.7, 0.8, 0.9\}$  se  $\beta = 0.7$ . Para  $n$  suficientemente grande, a escolha  $g(n) = \lfloor n^\beta \rfloor$ , sob o ponto de vista computacional, tem a vantagem de exigir um número menor de regressores no método DFA.

**Tabela 3.1:** Estimador  $\hat{d}_{\text{DFA}}$  para ARFIMA(0,  $d$ , 0) com Diferentes  $g(n)$ .

<b>ARFIMA(0, 0.1, 0)</b>						
$n = 1000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.2391	0.1554	0.1222	0.1068	0.0995	0.1015
vício	0.1391	0.0554	0.0222	0.0068	0.0005	0.0015
eqm	0.0229	0.0042	0.0014	0.0012	0.0018	0.0017
var	0.0035	0.0012	0.0009	0.0011	0.0018	0.0017
$n = 5000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.1743	0.1243	0.1076	0.1004	0.0948	0.0918
vício	0.0743	0.0243	0.0076	0.0004	-0.0052	-0.0082
eqm	0.0058	0.0008	0.0003	0.0004	0.0009	0.0012
var	0.0003	0.0002	0.0003	0.0004	0.0008	0.0011
$n = 10000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.1560	0.1177	0.1034	0.0986	0.0970	0.0970
vício	0.0560	0.0177	0.0034	-0.0014	-0.0030	-0.0030
eqm	0.0032	0.0004	0.0001	0.0003	0.0007	0.0009
var	0.0001	0.0001	0.0001	0.0003	0.0007	0.0009
<b>ARFIMA(0, 0.2, 0)</b>						
$n = 1000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.3031	0.2332	0.2099	0.2018	0.1939	0.1973
vício	0.1031	0.0332	0.0099	0.0018	-0.0061	-0.0027
eqm	0.0144	0.0025	0.0014	0.0015	0.0018	0.0018
var	0.0038	0.0014	0.0013	0.0015	0.0018	0.0018
$n = 5000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.2454	0.2057	0.1976	0.1955	0.1907	0.1883
vício	0.0454	0.0057	-0.0024	-0.0045	-0.0093	-0.0117
eqm	0.0023	0.0002	0.0003	0.0005	0.0010	0.0014
var	0.0003	0.0002	0.0002	0.0004	0.0010	0.0012
$n = 10000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.2300	0.2026	0.1966	0.1959	0.1965	0.1953
vício	0.0300	0.0026	-0.0034	-0.0041	-0.0035	-0.0047
eqm	0.0010	0.0001	0.0002	0.0003	0.0007	0.0012
var	0.0001	0.0001	0.0002	0.0003	0.0007	0.0012

**Tabela 3.2:** Estimador  $\hat{d}_{\text{DFA}}$  para ARFIMA(0,  $d$ , 0) com Diferentes  $g(n)$ .

<b>ARFIMA(0, 0.3, 0)</b>						
$n = 1000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.3625	0.3088	0.2937	0.2890	0.2839	0.2868
vício	0.0625	0.0088	-0.0063	-0.0110	-0.0161	-0.0132
eqm	0.0086	0.0014	0.0014	0.0020	0.0026	0.0024
var	0.0047	0.0013	0.0014	0.0019	0.0024	0.0022
$n = 5000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.3211	0.2930	0.2893	0.2934	0.2921	0.2900
vício	0.0211	-0.0070	-0.0107	-0.0066	-0.0079	-0.0100
eqm	0.0008	0.0003	0.0004	0.0006	0.0011	0.0015
var	0.0003	0.0003	0.0003	0.0006	0.0011	0.0014
$n = 10000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.3102	0.2935	0.2925	0.2938	0.2959	0.2969
vício	0.0102	-0.0065	-0.0075	-0.0062	-0.0041	-0.0031
eqm	0.0003	0.0002	0.0002	0.0004	0.0008	0.0013
var	0.0002	0.0001	0.0002	0.0003	0.0008	0.0013
<b>ARFIMA(0, 0.4, 0)</b>						
$n = 1000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.4370	0.3972	0.3846	0.3821	0.3823	0.3846
vício	0.0370	-0.0028	-0.0154	-0.0179	-0.0177	-0.0154
eqm	0.0053	0.0012	0.0016	0.0022	0.0031	0.0016
var	0.0040	0.0012	0.0014	0.0019	0.0028	0.0014
$n = 5000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.4017	0.3842	0.3859	0.3888	0.3886	0.3859
vício	0.0017	-0.0158	-0.0141	-0.0112	-0.0114	-0.0141
eqm	0.0004	0.0005	0.0006	0.0007	0.0014	0.0006
var	0.0004	0.0003	0.0004	0.0006	0.0013	0.0004
$n = 10000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\hat{d}_{\text{DFA}}$	0.3952	0.3852	0.3912	0.3961	0.3948	0.3912
vício	-0.0048	-0.0148	-0.0088	-0.0039	-0.0052	-0.0088
eqm	0.0002	0.0004	0.0003	0.0004	0.0008	0.0003
var	0.0002	0.0002	0.0003	0.0004	0.0008	0.0003

**Tabela 3.3:** Estimador  $\widehat{H}_{\text{DFA}}$  para FGN com Diferentes  $g(n)$ .

<b>H= 0.6</b>						
$n = 1000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.7648	0.6767	0.6380	0.6143	0.6027	0.6160
vício	0.1648	0.0767	0.0380	0.0143	0.0027	0.0160
eqm	0.0300	0.0070	0.0025	0.0016	0.0021	0.0018
var	0.0029	0.0011	0.0010	0.0014	0.0021	0.0015
$n = 5000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.6963	0.6419	0.6203	0.6103	0.6036	0.6036
vício	0.0963	0.0419	0.0203	0.0103	0.0036	0.0036
eqm	0.0095	0.0020	0.0007	0.0005	0.0008	0.0011
var	0.0003	0.0002	0.0003	0.0004	0.0008	0.0011
$n = 10000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.6805	0.6344	0.6158	0.6078	0.6019	0.6009
vício	0.0805	0.0344	0.0158	0.0078	0.0019	0.0009
eqm	0.0066	0.0013	0.0004	0.0004	0.0006	0.0008
var	0.0001	0.0001	0.0002	0.0003	0.0006	0.0008
<b>H=0.7</b>						
$n = 1000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.8583	0.7747	0.7389	0.7208	0.7078	0.7212
vício	0.1583	0.0747	0.0389	0.0208	0.0078	0.0212
eqm	0.0293	0.0067	0.0026	0.0019	0.0023	0.0022
var	0.0043	0.0011	0.0011	0.0015	0.0022	0.0018
$n = 5000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.7918	0.7386	0.7168	0.7058	0.6999	0.7012
vício	0.0918	0.0386	0.0168	0.0058	-0.0001	0.0012
eqm	0.0088	0.0017	0.0005	0.0005	0.0009	0.0011
var	0.0004	0.0002	0.0003	0.0005	0.0009	0.0011
$n = 10000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.7722	0.7298	0.7134	0.7048	0.7009	0.6989
vício	0.0722	0.0298	0.0134	0.0048	0.0009	-0.0011
eqm	0.0053	0.0010	0.0004	0.0004	0.0007	0.0010
var	0.0001	0.0001	0.0002	0.0004	0.0007	0.0010

**Tabela 3.4:** Estimador  $\widehat{H}_{\text{DFA}}$  para FGN com Diferentes  $g(n)$ .

<b>H= 0.8</b>						
$n = 1000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.9522	0.8694	0.8349	0.8172	0.8036	0.8170
vício	0.1522	0.0694	0.0349	0.0172	0.0036	0.0170
eqm	0.0275	0.0063	0.0024	0.0019	0.0023	0.0022
var	0.0044	0.0015	0.0012	0.0016	0.0023	0.0019
$n = 5000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.8872	0.8346	0.8147	0.8057	0.8011	0.8016
vício	0.0872	0.0346	0.0147	0.0057	0.0011	0.0016
eqm	0.0080	0.0015	0.0006	0.0007	0.0012	0.0014
var	0.0004	0.0003	0.0004	0.0007	0.0012	0.0014
$n = 10000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.8694	0.8286	0.8123	0.8038	0.7995	0.7983
vício	0.0694	0.0286	0.0123	0.0038	-0.0005	-0.0017
eqm	0.0050	0.0009	0.0004	0.0004	0.0007	0.0012
var	0.0001	0.0001	0.0002	0.0004	0.0007	0.0012
<b>H=0.9</b>						
$n = 1000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	1.0481	0.9681	0.9346	0.9188	0.9053	0.9188
vício	0.1481	0.0681	0.0346	0.0188	0.0053	0.0188
eqm	0.0270	0.0062	0.0028	0.0025	0.0028	0.0025
var	0.0051	0.0015	0.0016	0.0021	0.0028	0.0021
$n = 5000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.9834	0.9351	0.9151	0.9052	0.8958	0.8971
vício	0.0834	0.0351	0.0151	0.0052	-0.0042	-0.0029
eqm	0.0073	0.0015	0.0006	0.0006	0.0010	0.0012
var	0.0004	0.0003	0.0003	0.0006	0.0010	0.0012
$n = 10000$						
$g(n)$	$n^{0.3}$	$n^{0.4}$	$n^{0.5}$	$n^{0.6}$	$n^{0.7}$	$\frac{n}{10}$
$\widehat{H}_{\text{DFA}}$	0.9686	0.9269	0.9094	0.9010	0.8970	0.8983
vício	0.0686	0.0269	0.0094	0.0010	-0.0030	-0.0017
eqm	0.0048	0.0009	0.0003	0.0003	0.0009	0.0013
var	0.0001	0.0001	0.0002	0.0003	0.0008	0.0013

### 3.2.2 Comparação entre os Estimadores $\hat{d}_{\text{DFA}}$ e $\hat{d}_{\text{W}}$

O objetivo desta Subseção, é comparar o estimador  $\hat{d}_{\text{DFA}}$ , com o melhor estimador proposto na literatura para o parâmetro  $d$  que é aquele obtido pela aproximação da função de verossimilhança ( $\hat{d}_{\text{W}}$ ). Neste estudo, para fazer uma comparação entre os estimadores  $\hat{d}_{\text{DFA}}$  e  $\hat{d}_{\text{W}}$ , consideramos  $g(n) = n^{0.5}$  e geramos séries temporais obtidas de processos ARFIMA(0,  $d$ , 0). A escolha de  $g(n) = n^{0.5}$ , foi baseada nos resultados apresentados na Subseção 3.2.1. O estudo é baseado em 200 replicações, com tamanhos  $n \in \{1000, 5000, 10000, 20000, 30000\}$ . Os valores utilizados para o parâmetro  $d$  foram  $d \in \{0.1, 0.2, 0.3, 0.4\}$ . A Tabela 3.5 apresenta os estimadores  $\hat{d}_{\text{DFA}}$  e  $\hat{d}_{\text{W}}$  obtidos com seus respectivos valor médio, vício, erro quadrático médio (eqm) e variância (var).

Na Tabela 3.5 os estimadores  $\hat{d}_{\text{DFA}}$  e  $\hat{d}_{\text{W}}$  fornecem estimativas similares, mas o módulo do vício do estimador  $\hat{d}_{\text{W}}$ , é, em geral, menor do que o módulo do vício do  $\hat{d}_{\text{DFA}}$ . Note que à medida que  $n$  aumenta,  $\text{var}(\hat{d}_{\text{DFA}})$  e  $\text{var}(\hat{d}_{\text{W}})$  estão muito próximas, ambas tendendo a zero. A vantagem do estimador  $\hat{d}_{\text{DFA}}$  sobre o  $\hat{d}_{\text{W}}$  é que o primeiro pode ser utilizado para qualquer processo, inclusive os não estacionários.

**Tabela 3.5:** Resultados dos Estimadores  $\hat{d}_{\text{DFA}}$  (com  $g(n) = \lfloor n^{0.5} \rfloor$ ) e  $\hat{d}_{\text{W}}$  para Processos ARFIMA(0,  $d$ , 0), para Diferentes Valores de  $d$  e  $n$ .

$n = 1000$								
$d$	$\hat{d}_{\text{DFA}}$	vício	eqm	var	$\hat{d}_{\text{W}}$	vício	eqm	var
0.1	0.1245	0.0245	0.0017	0.0011	0.0912	-0.0088	0.0009	0.0009
0.2	0.2083	0.0083	0.0013	0.0013	0.1921	-0.0079	0.0009	0.0008
0.3	0.2929	-0.0071	0.0013	0.0013	0.2915	-0.0085	0.0009	0.0008
0.4	0.3812	-0.0188	0.0020	0.0016	0.3895	-0.0105	0.0009	0.0008
$n = 5000$								
$d$	$\hat{d}_{\text{DFA}}$	vício	eqm	var	$\hat{d}_{\text{W}}$	vício	eqm	var
0.1	0.1080	0.0080	0.0003	0.0002	0.0978	-0.0022	0.0002	0.0002
0.2	0.1958	-0.0042	0.0003	0.0003	0.1968	-0.0032	0.0002	0.0001
0.3	0.2918	-0.0082	0.0004	0.0003	0.2964	-0.0036	0.0002	0.0001
0.4	0.3875	-0.0125	0.0005	0.0004	0.3992	-0.0008	0.0001	0.0001
$n = 10000$								
$d$	$\hat{d}_{\text{DFA}}$	vício	eqm	var	$\hat{d}_{\text{W}}$	vício	eqm	var
0.1	0.1035	0.0035	0.0002	0.0001	0.0983	-0.0017	0.0001	0.0001
0.2	0.1977	-0.0023	0.0002	0.0002	0.1996	-0.0004	0.0001	0.0001
0.3	0.2917	-0.0083	0.0003	0.0002	0.2984	-0.0016	0.0001	0.0001
0.4	0.3921	-0.0079	0.0002	0.0002	0.3987	-0.0013	0.0001	0.0001
$n = 20000$								
$d$	$\hat{d}_{\text{DFA}}$	vício	eqm	var	$\hat{d}_{\text{W}}$	vício	eqm	var
0.1	0.1091	0.0091	0.0002	0.0001	0.0995	-0.0005	0.0000	0.0000
0.2	0.2019	0.0019	0.0001	0.0001	0.1992	-0.0008	0.0000	0.0000
0.3	0.2960	-0.0040	0.0001	0.0001	0.2996	-0.0004	0.0000	0.0000
0.4	0.3920	-0.0080	0.0002	0.0001	0.3995	-0.0005	0.0000	0.0000
$n = 30000$								
$d$	$\hat{d}_{\text{DFA}}$	vício	eqm	var	$\hat{d}_{\text{W}}$	vício	eqm	var
0.1	0.1012	0.0012	0.0001	0.0001	0.0993	-0.0007	0.0000	0.0000
0.2	0.1957	-0.0957	0.0093	0.0001	0.1993	-0.0007	0.0099	0.0000
0.3	0.2930	-0.0070	0.0002	0.0001	0.2995	-0.0005	0.0000	0.0000
0.4	0.3756	-0.0244	0.0020	0.0014	0.3831	-0.0169	0.0017	0.0014

# Capítulo 4

## Wavelets

Apresentamos neste capítulo, um breve resumo sobre a teoria de wavelets, pois no Capítulo 5, propomos um método de estimação para o índice do decaimento da cauda  $\alpha$  de distribuições estáveis, utilizando encolhimento de wavelets.

Na prática, a maioria dos sinais são funções no domínio do tempo. Quando plotamos um sinal, a variável independente é o tempo e a outra variável dependente é geralmente a amplitude, assim obtemos uma representação do tempo versus a amplitude do sinal. Para a maioria das aplicações relativas ao processamento de sinais, esta representação não é sempre a melhor para o sinal. Em muitos casos, a informação mais distinta está escondida no índice da frequência do sinal. O espectro de um sinal é basicamente as componentes de frequência (componentes espectrais) desse sinal. O espectro de um sinal mostra quais frequências estão presentes neste sinal.

Se uma variável oscila muito rapidamente, dizemos que é de alta frequência, se oscila pouco dizemos que é de baixa frequência e se esta variável nunca oscila então ela tem frequência zero. Logo, perguntamos como medimos a frequência, ou como encontramos o índice da frequência de um sinal? A resposta é a *transformada de Fourier* (TF). Tomando a transformada de Fourier de um sinal no domínio do tempo, obtemos a representação da frequência versus a amplitude do sinal.

Os sinais cujos índices de frequências não mudam com o tempo são chamados sinais estacionários. A transformada de Fourier fornece o índice espectral do sinal, mas não dá nenhuma informação a respeito de quando as componentes espectrais aparecem. Conseqüentemente, a TF não é uma técnica apropriada para o sinal não estacionário, com uma exceção: a TF pode ser utilizada para sinais não estacionários, se estamos somente interessados em saber quais são as componentes espectrais existentes no sinal, sem no entanto saber quando estas ocorrem. Entretanto, se esta informação é necessária, a TF não é a ferramenta apropriada para esta situação.

Suponhamos que dois sinais têm as mesmas componentes espectrais, mas com uma diferença. Por exemplo, um dos sinais tem quatro componentes de



frequência a todo tempo, e o outro tem as mesmas quatro componentes de frequência em tempos diferentes. Embora os dois sinais sejam completamente diferentes, a TF é a mesma para cada um deles. Isto, nos diz obviamente que não podemos utilizar a TF para sinais não estacionários.

Quando é necessária a localização no tempo das componentes espectrais, é importante uma transformação que forneça a representação do tempo versus frequência. A transformada de Wavelet é uma das transformações que fornece esta representação. Esta transformada é útil para analisar os sinais não estacionários, isto é, sinais cuja resposta de frequência varia no tempo.

Nesta seção apresentamos as wavelets, a transformada de Wavelets (TW) e uma introdução à regressão não paramétrica. O fato básico sobre wavelets é que elas são localizadas no tempo (ou espaço), contrariamente ao que ocorre com as funções trigonométricas. A idéia na análise de wavelets, é aproximar sinais (funções) através de uma combinação linear de wavelets. Existe um grande número de aplicações de wavelets nas diferentes áreas do conhecimento, tais como na compressão de dados e no processamento de sinais e de imagens. No campo da estatística, as wavelets são utilizadas na estimação de funções densidades, na regressão não-paramétrica, na estimação do espectro de processos estacionários e do espectro evolutivo de processos não estacionários.

Existem duas funções que são muito importantes na análise de wavelets, a wavelet mãe e a wavelet pai. Estas wavelets geram uma família de funções que podem reconstruir um sinal.

**Definição 4.1 (Wavelet Mãe).** A *wavelet mãe*  $\psi(\cdot)$  é uma função real  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  e a *wavelet pai* (ou *função escala*)  $\phi(\cdot)$  é uma função real  $\phi : \mathbb{R} \rightarrow \mathbb{R}$ , tais que

$$\int_{\mathbb{R}} \psi(t) dt = 0 \text{ e } \int_{\mathbb{R}} \phi(t) dt = 1$$

e que satisfazem a condição de integrabilidade, isto é,  $\psi, \phi \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ . Tipicamente, é exigido que a wavelet mãe seja limitada, centrada em torno da origem e deve decair para zero, quando  $|t| \rightarrow \infty$ .

O par  $(\psi, \phi)$  forma uma família. A relação entre as duas funções  $\psi(\cdot)$  e  $\phi(\cdot)$  é dada pelas equações de dilatação,

$$\begin{aligned} \psi(t) &= \sqrt{2} \sum_k h_k \phi(2t - k), \\ \phi(t) &= \sqrt{2} \sum_k g_k \phi(2t - k). \end{aligned} \tag{4.1}$$

Os coeficientes  $g_k$  e  $h_k$  são coeficientes de filtros passa baixo e passa alto, respectivamente, satisfazendo  $h_k = (-1)^k g_{1-k}$ .

Para explorar propriedades no domínio de Fourier utilizamos a função de transferência  $m_0(\cdot)$ , definida a seguir.

**Definição 4.2 (Função de Transferência).** A *função de transferência* é definida por

$$m_0(\omega) = \frac{1}{\sqrt{2}} \sum_{k \in \mathbb{Z}} g_k e^{-ik\omega}, \quad \text{para todo } \omega \in [-\pi, \pi], \quad (4.2)$$

onde  $g_k$  é um filtro passa baixo.

Dadas as wavelets  $\psi(\cdot)$  e  $\phi(\cdot)$  construímos uma sequência de wavelets através das translações e dilatações das wavelets pai e mãe dadas por

$$\begin{aligned} \psi_{j,k}(t) &= 2^{-j/2} \psi(2^{-j}t - k), \\ \phi_{j,k}(t) &= 2^{-j/2} \phi(2^{-j}t - k). \end{aligned} \quad (4.3)$$

Existem várias famílias de wavelets, sendo que as mais utilizadas, na prática, tem suporte compacto. As wavelets ortogonais mais utilizadas em geral são: Haar, Daubelets, Symmlets e Coiflets (ver Figura 4.2).

O exemplo mais simples e antigo de wavelets é a *função Haar* (ver Definição 4.3). A *wavelet Haar*, é uma função com partes constantes, que preserva descontinuidades.

**Definição 4.3 (Wavelet Haar).** A *função wavelet mãe Haar* é definida por

$$\psi(t) = \begin{cases} 1, & \text{se } t \in [0, \frac{1}{2}) \\ -1, & \text{se } t \in [\frac{1}{2}, 1) \\ 0, & \text{caso contrário.} \end{cases} \quad (4.4)$$

e a *função wavelet pai Haar* ou *função escala de Haar* é definida por

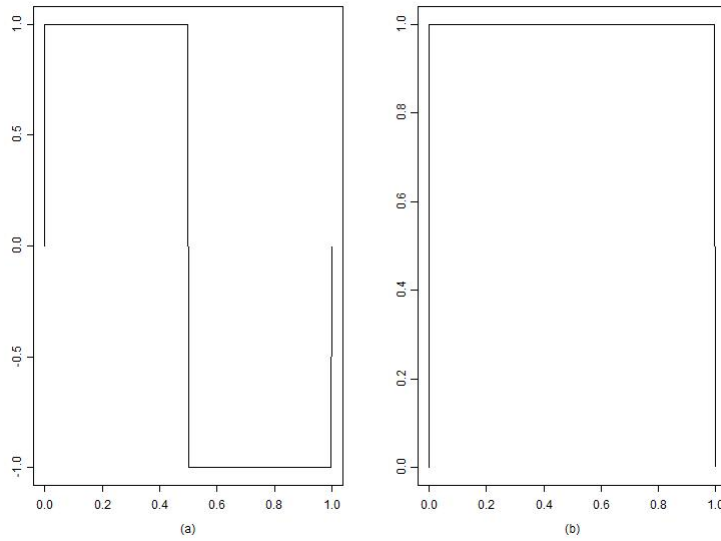
$$\phi(t) = \begin{cases} 1, & \text{se } t \in [0, 1) \\ 0, & \text{caso contrário.} \end{cases} \quad (4.5)$$

**Observação 4.1.** A relação entre as duas wavelets  $\psi(\cdot)$  e  $\phi(\cdot)$  da Definição 4.3 é dada por

$$\psi(t) = \phi(2t) - \phi(2t - 1).$$

Os gráficos das wavelets mãe e pai da Haar são apresentados na Figura 4.1.

As wavelets Daubelets foram as primeiras wavelets ortogonais com suporte compacto. As Daubelets são totalmente assimétricas e as wavelets Symmlets foram construídas para serem as mais simétricas possíveis. As wavelets Coiflets foram construídas para serem quase simétricas e possuem



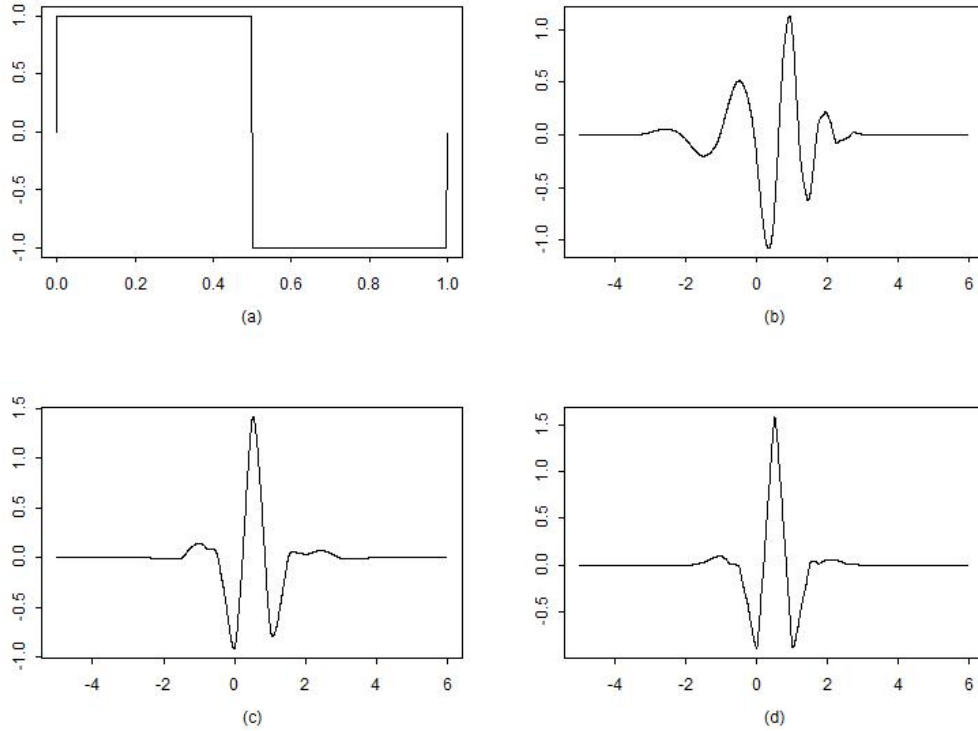
**Figura 4.1:** Wavelets Haar: (a) wavelet mãe; (b) wavelet pai.

propriedades adicionais (momentos nulos para as funções). Todas foram introduzidas por Ingrid Daubechies, que é uma das pioneiras no estudo de wavelets. A Figura 4.2 apresenta 4 diferentes wavelets mãe ortogonais “Haar”, “d12”, “s12” e “c12” onde a primeira letra da wavelet indica o nome: **d** para Daublet, **s** para Symmlet e **c** para Coiflet. O número da wavelet indica a largura e a suavização.

**Observação 4.2.** As wavelets Haar, Daublet, Symmlet e Coiflet são dadas no pacote “wmtsa” do software R-project.

A razão central na escolha de uma wavelet é usar aquela que melhor combine com as características do sinal que se está analisando. A escolha de uma wavelet necessita de um compromisso entre diferentes propriedades tais como suavização, localização espacial, localização de frequência, habilidade para representar funções polinomiais locais, ortogonalidade e simetria.

As funções  $\psi_{j,k}(\cdot)$  e  $\phi_{j,k}(\cdot)$ , dadas na expressão (4.3), são versões escaladas e transladadas de  $\psi(\cdot)$  e  $\phi(\cdot)$ , com escala  $2^j$  e parâmetro de translação  $2^j k$ , respectivamente. O fator escala  $2^j$  é chamado também de *fator dilatação* e o parâmetro  $2^j k$  é chamado também de *posição*. Quando  $j$  cresce, o fator escala  $2^j$  aumenta e as funções  $\psi_{j,k}(\cdot)$  e  $\phi_{j,k}(\cdot)$  tendem a ficar mais curtas e largas. Assim, o fator  $2^j$  é uma medida de escala, ou largura, das funções  $\psi_{j,k}(\cdot)$  e  $\phi_{j,k}(\cdot)$ . O parâmetro da translação  $2^j k$  é combinado com o parâmetro de escala  $2^j$  no sentido de que quando as funções  $\psi_{j,k}(\cdot)$  e  $\phi_{j,k}(\cdot)$  tendem a ficar mais largas, suas etapas de translação são correspondentemente maiores.



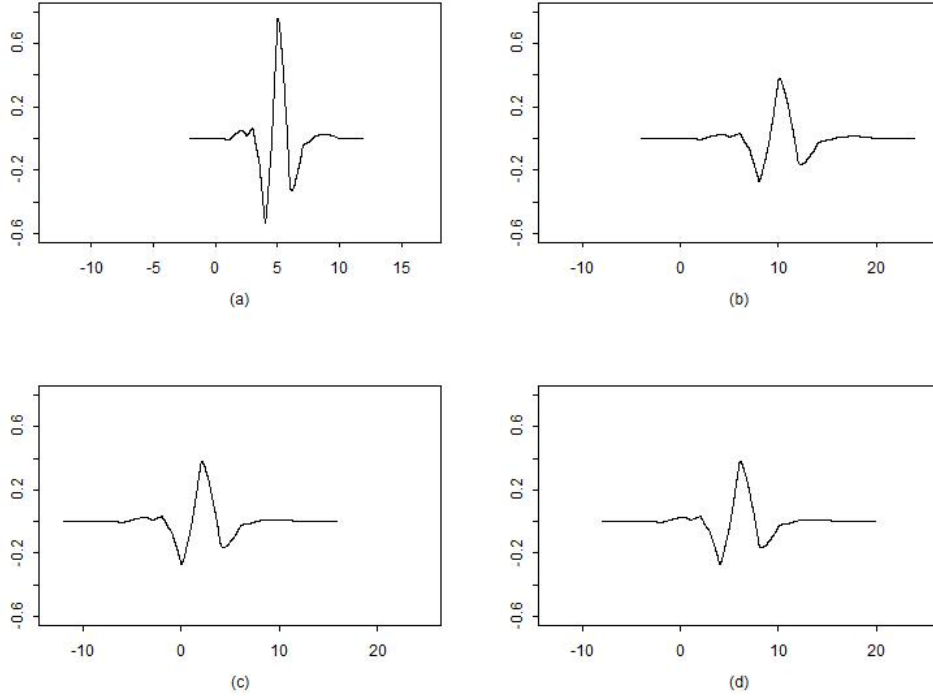
**Figura 4.2:** Diferentes Wavelets Mãe: (a) Haar; (b) Daubelets “d12”; (c) Symmlets “s12” e (d) Coiflets “c12”.

**Exemplo 4.1.** A Figura 4.3 apresenta 4 versões escaladas e transladadas da wavelet mãe “s8”:  $\psi_{1,2}(\cdot)$  (escala 2 e translação 4),  $\psi_{2,2}(\cdot)$  (escala 4 e translação 8),  $\psi_{2,0}(\cdot)$  escala 4 e translação 0 e  $\psi_{2,1}(\cdot)$  (escala 4 e translação 4), respectivamente.

A aproximação de um sinal contínuo  $f(\cdot) \in L^2(\mathbb{R})$  por uma série ortogonal de wavelets é dada por

$$\begin{aligned}
 f(t) \approx & \sum_k s_{J,k} \phi_{J,k}(t) + \sum_k d_{J,k} \psi_{J,k}(t) \\
 & + \sum_k d_{J-1,k} \psi_{J-1,k}(t) + \cdots + \sum_k d_{1,k} \psi_{1,k}(t), \quad (4.6)
 \end{aligned}$$

onde  $J$  é o número de componentes de multiresoluções (ver Definição 4.9) e  $k$  varia de 1 ao número de coeficientes do componente específico. Os coeficientes  $s_{J,k}, d_{J,k}, \dots, d_{1,k}$  são os *coeficientes da transformada de wavelets* e as funções  $\psi_{j,k}(\cdot)$  e  $\phi_{j,k}(\cdot)$  são dadas na expressão (4.3).



**Figura 4.3:** Versões Escaladas e Transladadas da Wavelet Symmlets “s8”: (a) escala 2 e translação 4; (b) escala 4 e translação 8; (c) escala 4 e translação 0 e (d) escala 4 e translação 4.

A aproximação dada na expressão (4.6) é uma aproximação de séries ortogonais pois as funções da base são ortogonais por construção. De fato,

$$\begin{aligned}
 \int \phi_{J,k}(t)\phi_{J,k'}(t)dt &= \delta_{kk'} \\
 \int \psi_{j,k}(t)\phi_{J,k'}(t)dt &= 0 \\
 \int \psi_{j,k}(t)\psi_{j',k'}(t)dt &= \delta_{jj'}\delta_{kk'}
 \end{aligned} \tag{4.7}$$

onde

$$\delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & i \neq j. \end{cases} \tag{4.8}$$

**Observação 4.3.** Os coeficiente  $s_{J,k}$ 's capturam as oscilações de baixa frequência, enquanto que os coeficientes  $d_{j,k}$ , para  $j = 1, 2, \dots, J$ , capturam

oscilações de alta frequência. Os coeficientes  $d_{j,k}$ , para  $j = 1, 2, \dots, J - 1$  capturam escalas finas (detalhes) e os  $s_{J,k}$  e  $d_{J,k}$  representam escalas grossas (suaves).

Os coeficientes das wavelets são dados por

$$\begin{aligned} s_{J,k} &\approx \int \phi_{J,k}(t) f(t) dt, \\ d_{j,k} &\approx \int \psi_{j,k}(t) f(t) dt, \end{aligned} \quad (4.9)$$

para todo  $j = 1, 2, \dots, J$  e  $k = 1, 2, \dots, \lfloor \frac{n}{2^j} \rfloor$ .

## 4.1 Análise da Wavelet Haar

**Definição 4.4 (Sistema Ortonormal Completo).** Uma sequência  $\{f_j\}_{j \in \mathbb{Z}}$  é dita ser um *sistema ortonormal completo* se as  $f_j$ 's são ortogonais por pares e se  $\|f_j\| = 1$ , para todo  $j \in \mathbb{Z}$ , e a única função ortogonal para cada  $f_j$  é a função zero.

**Teorema 4.1.** O conjunto  $\{\psi_{j,k}, j, k \in \mathbb{Z}\}$  forma um sistema ortonormal completo para  $L^2(\mathbb{R})$ , onde as funções  $\psi_{j,k}(\cdot)$ , para todo  $j = 1, 2, \dots, J$  e  $k = 1, 2, \dots, \frac{n}{2^j}$ , são dadas pela expressão (4.3) utilizando a wavelet mãe Haar.

**Demonstração:** Ver Daubechies (1992) e Ogden (1997). Para mostrar este teorema é necessário provar que o conjunto  $\{\psi_{j,k}, j, k \in \mathbb{Z}\}$  é ortonormal e que qualquer função  $f \in L^2(\mathbb{R})$  pode ser arbitrariamente bem aproximada por uma combinação linear finita das funções  $\psi_{j,k}(\cdot)$ , para todo  $j = 1, 2, \dots, J$  e  $k = 1, 2, \dots, \frac{n}{2^j}$ . □

**Definição 4.5 (Espaço  $V_j$ ).** O espaço das funções escadas no nível  $j \in \mathbb{Z}^+$ , denotado por  $V_j$ , é definido pelo espaço gerado pelo conjunto

$$A = \{\phi(2^j t + 1), \phi(2^j t), \phi(2^j t - 1), \phi(2^j t - 2)\}.$$

O conjunto  $V_j$ , para cada  $j$ , é um espaço de partes de funções constantes com descontinuidades contidas no conjunto

$$\{\dots, -1/2^j, 0, 1/2^j, 2/2^j, 3/2^j, \dots\}.$$

**Teorema 4.2.** Uma função  $f(t) \in V_0$  se e somente se  $f(2^j t) \in V_j$  e uma função  $f(t) \in V_j$  se e somente se  $f(2^{-j} t) \in V_0$ .

**Demonstração:** Se uma função  $f \in V_0$ , então  $f(t)$  é uma combinação linear de  $\phi(t-k)$ , para todo  $k \in \mathbb{Z}$ . Portanto, é uma combinação linear de  $\phi(2^j t - k)$ , para todo  $k \in \mathbb{Z}$ , isto é,  $f(2t) \in V_j$ . A prova do outro item é similar.  $\square$

**Observação 4.4.** O conjunto das funções  $\{2^{j/2}\phi(2^j t - k); k \in \mathbb{Z}\}$  é uma base ortonormal de  $V_j$ .

**Teorema 4.3.** *Seja  $W_j$  um espaço de funções dadas por*

$$\sum_{k \in \mathbb{Z}} a_k \psi(2^j t - k),$$

onde apenas um número finito de  $a_k \in \mathbb{R}$  não são nulos. O conjunto  $W_j$  é o complemento ortogonal de  $V_j$  em  $V_{j+1}$  e

$$V_{j+1} = V_j \oplus W_j.$$

**Demonstração:** Para provar este teorema, deve-se mostrar que cada função em  $W_j$  é ortogonal a toda função em  $V_j$  e que qualquer função em  $V_{j+1}$  ortogonal a  $V_j$  pertence a  $W_j$ . Ver Boggess e Narcowich (2001).  $\square$

**Teorema 4.4.** *O espaço  $L^2(\mathbb{R})$  é dado por  $V_0 \oplus W_0 \oplus W_1 \oplus \dots$ . Em particular, cada  $f \in L^2(\mathbb{R})$  pode ser escrita unicamente por*

$$f = f_0 + \sum_{j=0}^{\infty} \omega_j,$$

onde  $f_0 \in V_0$  e  $\omega_j \in W_j$ , para todo  $j \in \mathbb{N}$ .

**Demonstração:** Ver Boggess e Narcowich (2001), página 165.  $\square$

**Lema 4.1.** *As seguintes relações valem, para todo  $t \in \mathbb{R}$ ,*

$$\phi(2^j t) = (\psi(2^{j-1} t) + \phi(2^{j-1} t))/2 \quad (4.10)$$

$$\phi(2^j t - 1) = (\phi(2^{j-1} t) - \psi(2^{j-1} t))/2. \quad (4.11)$$

**Demonstração:** Ver Boggess e Narcowich (2001).  $\square$

**Exemplo 4.2.** Considere a função  $f : \mathbb{R} \rightarrow \mathbb{R}$ , dada por

$$f(t) = 2\phi(4t) + 2\phi(4t - 1) + \phi(4t - 2) - \phi(4t - 3). \quad (4.12)$$

Queremos decompor  $f$  em componentes  $W_1, W_0$  e  $V_0$ . Pelas expressões (4.10) e (4.11) com  $j = 2$  obtemos as seguintes equações

$$\begin{aligned}
\phi(4t) &= \frac{1}{2}[\psi(2t) + \phi(2t)] \\
\phi(4t - 1) &= \frac{1}{2}[\phi(2t) - \psi(2t)] \\
\phi(4t - 2) &= \phi(4(t - 1/2)) = \frac{1}{2}[\psi(2(t - 1/2)) + \phi(2(t - 1/2))] \\
\phi(4t - 3) &= \frac{1}{2}[\phi(2(t - 1/2)) - \psi(2(t - 1/2))]. \tag{4.13}
\end{aligned}$$

Substituindo as equações da expressão (4.13) na expressão (4.12) obtemos

$$\begin{aligned}
f(t) &= [\psi(2t) + \phi(2t)] + [\phi(2t) - \psi(2t)] \\
&+ \frac{1}{2}[\psi(2(t - 1/2)) + \phi(2(t - 1/2))] \\
&- \frac{1}{2}[\phi(2(t - 1/2)) - \psi(2(t - 1/2))] \\
&= 2\phi(2t) + \psi(2(t - 1/2)) \\
&= 2\phi(2t) + \psi(2t - 1). \tag{4.14}
\end{aligned}$$

O componente  $W_1$  da função  $f(t)$  é  $\psi(2t - 1)$  e o componente  $V_1$  da função  $f(t)$  é  $2\phi(2t)$ , utilizando a equação  $\phi(2t) = \frac{1}{2}[\phi(t) + \psi(t)]$  o componente  $V_1$  pode ser decomposto em componentes  $V_0$  e  $W_0$ . Assim, segue-se o resultado

$$f(t) = \psi(2t - 1) + \psi(t) + \phi(t).$$

**Teorema 4.5 (Decomposição Haar).** *Suponha que, para todo  $j = 1, 2, \dots, J$ , a função  $f_j(\cdot)$  é dada por*

$$f_j(t) = \sum_{k \in \mathbb{Z}} a_k^j \phi(2^j t - k) \in V_j.$$

Então,  $f_j$  pode ser decomposta por

$$f_j = \omega_{j-1} + f_{j-1},$$

onde

$$\omega_{j-1} = \sum_{k \in \mathbb{Z}} b_k^{j-1} \psi(2^{j-1} t - k) \in W_{j-1}$$

$$f_{j-1} = \sum_{k \in \mathbb{Z}} a_k^{j-1} \phi(2^{j-1} t - k) \in V_{j-1},$$

com

$$b_k^{j-1} = \frac{a_{2k}^j - a_{2k+1}^j}{2} \quad a_k^{j-1} = \frac{a_{2k}^j + a_{2k+1}^j}{2}.$$



**Demonstração:** Seja a função  $f_j(\cdot)$  dada por

$$f_j(t) = \sum_{k \in \mathbb{Z}} a_k^j \phi(2^j t - k) \in V_j.$$

Dividindo a função  $f_j(\cdot)$  em partes par e ímpar, temos que

$$f_j(t) = \sum_{k \in \mathbb{Z}} a_{2k} \phi(2^j t - 2k) + \sum_{k \in \mathbb{Z}} a_{2k+1} \phi(2^j t - 2k - 1). \quad (4.15)$$

Utilizando as expressões (4.10) e (4.11) e substituindo  $t$  por  $t - k2^{1-j}$

$$\phi(2^j t - 2k) = \frac{1}{2} [\psi(2^{j-1} t - k) + \phi(2^{j-1} t - k)] \quad (4.16)$$

$$\phi(2^j t - 2k - 1) = \frac{1}{2} [\phi(2^{j-1} t - k) - \psi(2^{j-1} t - k)]. \quad (4.17)$$

Substituindo as expressões (4.16) e (4.17) em (4.15), obtemos

$$\begin{aligned} f_j(t) &= \sum_{k \in \mathbb{Z}} \frac{a_{2k}}{2} [\psi(2^{j-1} t - k) + \phi(2^{j-1} t - k)] \\ &+ \sum_{k \in \mathbb{Z}} \frac{a_{2k+1}}{2} [\phi(2^{j-1} t - k) - \psi(2^{j-1} t - k)] \\ &= \sum_{k \in \mathbb{Z}} \left[ \left( \frac{a_{2k} - a_{2k+1}}{2} \right) \psi(2^{j-1} t - k) + \left( \frac{a_{2k} + a_{2k+1}}{2} \right) \phi(2^{j-1} t - k) \right] \\ &= \omega_{j-1} + f_{j-1}. \end{aligned} \quad (4.18)$$

□

O procedimento utilizado no Teorema 4.5 pode ser repetido com o  $j$  substituído por  $j - 1$  para decompor  $f_{j-1}$  como  $\omega_{j-2} + f_{j-2}$ . Repetindo este procedimento, obtemos a decomposição

$$f_j = \omega_{j-1} + \omega_{j-2} + \cdots + \omega_0 + f_0.$$

**Teorema 4.6 (Reconstrução Haar).** *Suponha que a função  $f$  é dada por*

$$f = f_0 + \omega_0 + \omega_1 + \cdots + \omega_{j-1}$$

onde

$$f_0(t) = \sum_{k \in \mathbb{Z}} a_k^0 \phi(t - k) \in V_0 \quad e \quad \omega_{j'}(t) = \sum_{k \in \mathbb{Z}} b_k^{j'} \psi(2^{j'} t - k) \in W_{j'}$$

para  $0 \leq j' < j$ . Então,

$$f(t) = \sum_{l \in \mathbb{Z}} a_l^j \phi(2^j t - l) \in V_j,$$

onde  $a_l^{j'}$  são determinados recursivamente por  $j' = 1$ , depois por  $j' = 2$  e assim até  $j' = j$ , pelo algoritmo

$$a_l^{j'} = \begin{cases} a_k^{j'-1} + b_k^{j'-1}, & \text{se } l = 2k \\ a_k^{j'-1} - b_k^{j'-1}, & l = 2k + 1. \end{cases}$$

**Demonstração:** Seja

$$f(t) = f_0(t) + \omega_0(t) + \omega_1(t) + \cdots + \omega_{j-1}(t)$$

com

$$f_0(t) = \sum_{k \in \mathbb{Z}} a_k^0 \phi(t - k) \in V_0 \quad \text{e} \quad \omega_l(t) = \sum_{k \in \mathbb{Z}} b_k^l \psi(2^l t - k) \in W_l,$$

para  $0 \leq l < j - 1$ . Pelas definições de  $\phi(\cdot)$  e  $\psi(\cdot)$ , obtemos as equações

$$\begin{aligned} \phi(t) &= \phi(2t) + \phi(2t - 1) \\ \psi(t) &= \phi(2t) - \phi(2t - 1). \end{aligned} \tag{4.19}$$

Substituindo  $t$  por  $2^{j-1}t$  na expressão (4.19),

$$\begin{aligned} \phi(2^{j-1}t) &= \phi(2^j t) + \phi(2^j t - 1) \\ \psi(2^{j-1}t) &= \phi(2^j t) - \phi(2^j t - 1). \end{aligned}$$

Utilizando a primeira equação da expressão (4.19) e substituindo  $t$  por  $t - k$ , obtemos que

$$\begin{aligned} f_0(t) &= \sum_{k \in \mathbb{Z}} a_k^0 \phi(t - k) \\ &= \sum_{k \in \mathbb{Z}} a_k^0 \phi(2t - 2k) + a_k^0 \phi(2t - 2k - 1). \end{aligned}$$

Então,

$$f_0(t) = \sum_{k \in \mathbb{Z}} \hat{a}_1^0 \phi(2t - l), \tag{4.20}$$

onde

$$\hat{a}_l^1 = \begin{cases} a_k^0, & \text{se } l = 2k \\ a_k^0, & \text{se } l = 2k + 1. \end{cases}$$

Similarmente, utilizando a segunda equação da expressão (4.19) para  $\psi(t - k)$ , podemos escrever  $\omega_0(t) = \sum_{k \in \mathbb{Z}} b_k^0 \psi(t - k)$  como

$$\omega_0(t) = \sum_{k \in \mathbb{Z}} \hat{b}_1^0 \phi(2t - l), \tag{4.21}$$

onde

$$\hat{b}_l^1 = \begin{cases} b_k^0, & \text{se } l = 2k \\ -b_k^0, & \text{se } l = 2k + 1. \end{cases}$$

Combinando (4.20) com (4.21), obtemos

$$f_0(t) + \omega_0(t) = \sum_{l \in \mathbb{Z}} a_l^1 \phi(2t - l), \quad (4.22)$$

onde

$$a_l^1 = \hat{a}_l^1 + \hat{b}_l^1 = \begin{cases} a_k^0 + b_k^0, & \text{se } l = 2k \\ a_k^0 - b_k^0, & \text{se } l = 2k + 1. \end{cases}$$

Adicionando  $\omega_1(t) = \sum_{k \in \mathbb{Z}} b_k^1 \psi(2t - l)$  em (4.22) e utilizando (4.19) com  $t$  substituído por  $2t - k$  temos

$$f_0(t) + \omega_0(t) + \omega_1(t) = \sum_{l \in \mathbb{Z}} a_l^2 \phi(2^2 t - l),$$

onde

$$a_l^2 = \begin{cases} a_k^1 + b_k^1, & \text{se } l = 2k \\ a_k^1 - b_k^1, & \text{se } l = 2k + 1. \end{cases}$$

Note que os coeficientes  $a_l^0$  e  $b_l^0$  determinam o coeficiente  $a_l^1$ . Então, o  $a_l^1$  e  $b_l^1$  determinam o coeficiente  $a_l^2$  e assim recursivamente. □

## 4.2 Regularidade das Wavelets

Uma propriedade importante das bases de wavelets é a diversidade. Pode-se construir wavelet com diferente suavização, simetria e sustentação. A regularidade (suavização) das wavelets está relacionada com a taxa de decaimento das funções escalas e com o número de momentos nulos das funções escalas e das wavelets mãe.

O Teorema 4.7 a seguir relaciona regularidade de wavelets com o número de momentos nulos e a forma da função transferência  $m_0(\omega)$  (ver Definição 4.2). A prova é baseada em argumentos de séries de Taylor e em propriedades da função escala de wavelets. Para mais detalhes ver Daubechies (1992).

Sejam

$$\mathcal{M}_k = \int x^k \phi(t) dt \quad \text{e} \quad \mathcal{N}_k = \int x^k \psi(t) dt,$$

os  $k$ -ésimos momentos das funções escala e wavelet mãe, respectivamente.

**Teorema 4.7.** *Seja  $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$ ,  $j, k \in \mathbb{Z}$  um sistema ortonormal (ver Definição 4.4) de funções em  $L^2(\mathbb{R})$ . Considere que  $\psi \in \mathbb{C}^{N-1}(\mathbb{R})$  e*

$$|\psi(t)| \leq \frac{C_1}{(1+|t|)^\alpha}, \quad \alpha > N,$$

onde as derivadas  $\psi^{(k)}(t)$  são limitadas para  $k \leq N - 1$ . Então,  $\psi$  tem  $N$  momentos nulos, isto é,

$$\mathcal{N}_k = 0, \quad 0 \leq k \leq N - 1.$$

Se também

$$|\phi(t)| \leq \frac{C_2}{(1+|t|)^\alpha}, \quad \alpha > N$$

então, a função associada  $m_0(t)$  é necessariamente da forma

$$m_0(\omega) = \left( \frac{1 + e^{-i\omega}}{2} \right)^N \cdot \mathcal{L}(\omega), \quad (4.23)$$

onde  $\mathcal{L}$  é uma função em  $\mathbb{C}^{N-1}$ ,  $2\pi$  periódica.

**Demonstração:** Ver Vidakovic (1992), página 80. □

**Observação 4.5.** Uma função pertence ao espaço  $\mathbb{C}^{N-1}$  se é  $N - 1$  vezes diferenciável.

**Definição 4.6 (AM  $r$ -regular).** A análise de multiresolução (ou função escala) é dita ser  $r$ -regular se, para todo  $\alpha \in \mathbb{Z}$ ,

$$|\phi^{(k)}(t)| \leq \frac{C}{(1+|t|)^\alpha},$$

para  $k = 0, 1, \dots, r$ .

## 4.3 Transformadas de Wavelets

A transformada de wavelets apareceu em sua forma contínua com os trabalhos de dois pesquisadores franceses, o geofísico Morlet e o físico Grossmann (ver Morlet et al., 1982 e e Grossmann e Morlet, 1985).

**Definição 4.7 (Transformada Contínua de Wavelets).** Seja  $f \in L^2(\mathbb{R})$ . A transformada contínua de wavelets para a função  $f(\cdot)$ , com respeito a  $\psi$ , é definida por

$$W_\psi f(a, b) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt, \quad a, b \in \mathbb{R}, a \neq 0. \quad (4.24)$$

Note que a transformada contínua de wavelets é o produto interno de  $f$  pela wavelet  $\psi_{a,b} = |a|^{-1/2}\psi\left(\frac{x-b}{a}\right)$ .

**Exemplo 4.3 (Transformada Contínua da Haar).** A transformada contínua de wavelets pode ser simples. Seja

$$\psi_{a,b}^{Haar}(x) = \frac{1}{\sqrt{a}} \left[ \mathbf{I}_{\left[b, \frac{a}{2}+b\right)}(x) - \mathbf{I}_{\left[\frac{a}{2}+b, a+b\right)}(x) \right], \quad a \in \mathbb{R}^+, b \in \mathbb{R}, \quad (4.25)$$

onde  $\mathbf{I}_A(\cdot)$  é a função indicadora do conjunto  $A$ . Seja  $F$  a função primitiva de  $f$ , isto é,  $F' = f$ . Então,

$$W_\psi f(a, b) = \langle f, \psi_{a,b}^{Haar} \rangle = \frac{2}{\sqrt{a}} \left[ F\left(\frac{a}{2} + b\right) - \frac{F(b) + F(a+b)}{2} \right], \quad (4.26)$$

onde  $\langle f, \psi_{a,b}^{Haar} \rangle$  é o produto interno de  $f$  pela wavelet Haar  $\psi_{a,b}^{Haar}$ .

A transformada discreta de wavelets utiliza dilatações e translações da wavelet mãe para capturar as características locais no tempo e na frequência (ver Definição 4.8).

**Definição 4.8 (Transformada Discreta de Wavelets).** Seja  $\mathbf{X} = (X_1, X_2, \dots, X_T)'$  uma amostra i.i.d. de um processo estocástico, onde  $J = \lfloor \log_2(T) \rfloor$ . Definimos a *transformada discreta de wavelets* (TDW) de  $\mathbf{X}$ , com respeito a wavelet mãe  $\psi(\cdot)$ , como sendo

$$d_{j,k}^\psi = \sum_{t=1}^T X_t \psi_{j,k}(t), \quad (4.27)$$

a qual denotaremos simplesmente por  $d_{j,k}$ , omitindo a dependência sobre  $\psi(\cdot)$ . Esta transformada é calculada para  $j = 1, 2, \dots, J$  e  $k = 1, 2, \dots, \lfloor \frac{n}{2^j} \rfloor$ . A expressão (4.27) é matematicamente equivalente à multiplicação por uma matriz ortogonal  $W$ , isto é,

$$d = WX.$$

Supondo-se condições de fronteira apropriadas, a transformada é ortogonal e podemos invertê-la para obter a *transformada inversa discreta de wavelets* (ITWD) dada por

$$X = W'd,$$

onde  $W'$  denota a transposta de  $W$ .

A *transformada discreta de wavelets* (TDW) e a *transformada inversa discreta de wavelets* (ITDW) são calculadas pelo algoritmo piramidal de Mallat

(ver Mallat, 1989). Estes algoritmos envolvem filtros passa baixo e passa alto, junto com um operador *down-sampling* (*decimation*) ou *up-sampling* (*zero-padding*). Na verdade existem vários algoritmos piramidais: o Capítulo 3 de Meyer (1993) apresenta os algoritmos piramidais através de uma perspectiva histórica.

**Exemplo 4.4.** Seja o vetor  $(1, 2)$  e seja  $M(1, 2)$  um ponto em  $\mathbb{R}^2$  com as coordenadas dadas pelo vetor de dados. A rotação das coordenadas do eixo por um ângulo de  $\frac{\pi}{4}$  pode ser interpretada como uma transformada discreta de wavelets na base de wavelets Haar. A matrix de rotação é dada por

$$W = \begin{pmatrix} \cos \frac{\pi}{4} & \sin \frac{\pi}{4} \\ \sin \frac{\pi}{4} & \cos \frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix},$$

e a transformada discreta de wavelets de  $(1, 2)'$  é  $W \cdot (1, 2)' = (\frac{3}{\sqrt{2}}, -\frac{1}{\sqrt{2}})'$ . Note que a distância ao quadrado do ponto da origem é preservada,  $1^2 + 2^2 = (\frac{1}{2})^2 + (\frac{\sqrt{3}}{2})^2$ , pois  $W$  é uma rotação.

## 4.4 Análise de Multiresolução

A análise de multiresolução permite analisar os dados disponíveis em várias escalas de resolução. Formalmente, uma análise de multiresolução é uma sequência crescente de subespaços fechados  $\{V_j, j \in \mathbb{Z}\}$  que aproximam  $L^2(\mathbb{R})$  (ver Definição 4.9).

**Definição 4.9 (Análise de Multiresolução).** Seja  $\{V_j, j \in \mathbb{Z}\}$  uma sequência de subespaços fechados em  $L^2(\mathbb{R})$ . A coleção de espaços  $\{V_j, j \in \mathbb{Z}\}$  é chamada uma análise de multiresolução (AMR) com função escala  $\phi$  se

1.  $V_j \subset V_{j+1}$
2.  $\overline{\bigcup V_j} = L^2(\mathbb{R})$  ( $\bar{A}$  significa fecho do conjunto  $A$ )
3.  $\bigcap V_j = \{0\}$
4.  $f(t) \in V_j \Leftrightarrow f(2^{-j}t) \in V_0$
5. A função  $\phi \in V_0$  e o conjunto  $\{\phi(n-k), k \in \mathbb{Z}\}$  é uma base ortonormal de  $V_0$ . Se  $V_j$  for gerado por  $\{\phi_{j,k}\}$ , então  $V_{j+1}$  será gerado por  $\{\phi_{j+1,k}\}$ . Portanto,  $\phi_{j+1,k}(t) = \sqrt{2}\phi_{j,k}(2t)$ , para todo  $j, k \in \mathbb{Z}$ .

**Exemplo 4.5.** Os espaços  $V_j = \{f \in L^2(\mathbb{R}) : f \text{ é constante em } (2^{-j}k, 2^{-j}(k+1)), k \in \mathbb{Z}\}$ , para todo  $j \in \mathbb{Z}$ , formam uma aproximação de multiresolução.

A idéia que está por trás da análise de multiresolução é a seguinte. Considere uma função  $f$  de  $L^2(\mathbb{R})$ . Desejamos obter aproximações a  $f$  em vários níveis de resolução. Cada subespaço  $V_j$  será constituído por funções aproximantes, sendo que a melhor aproximação é obtida considerando-se a projeção ortogonal de  $f$  sobre cada  $V_j$ . O fato de  $V_j \subset V_{j+1}$  significa que ao passar do nível de resolução  $j$  (escala  $2^{-j}$ ) para o nível de resolução  $(j+1)$ , ganhamos informações. À medida que a resolução aumenta, a função aproximada converge para a função original e obtemos  $\bigcup V_j = L^2(\mathbb{R})$ . Por outro lado, quando aproximamos  $f$  a níveis de resolução cada vez menores, perdemos informação, a aproximação de  $f$  converge para a função nula e obtemos  $\bigcap V_j = \{0\}$ .

Para todo  $t \in \mathbb{R}$ , as funções

$$\begin{aligned} S_J(t) &= \sum_k s_{J,k} \phi_{J,k}(t), \\ D_j(t) &= \sum_k d_{j,k} \psi_{j,k}(t), \end{aligned} \quad (4.28)$$

onde  $j = 1, 2, \dots, J$ ,  $J$  é o número de componentes de multiresoluções e  $k = 1, 2, \dots, n/2^j$ , são chamadas *senal suave* e *senal detalhe*, respectivamente. Observe que a aproximação de séries ortogonais de wavelets para uma função contínua  $f(\cdot)$  dada na expressão (4.6) é expressa em termos destes sinais

$$f(t) \approx S_J(t) + D_J(t) + D_{J-1}(t) + \dots + D_1(t). \quad (4.29)$$

Os termos na aproximação dada pela expressão (4.29) constituem uma decomposição do sinal  $f(\cdot)$  em componentes ortogonais  $S_J(t)$  e  $D_j(t)$  para  $j \in \{1, \dots, J\}$  em diferentes escalas. A aproximação é chamada *decomposição de multiresolução* (DMR), pois os termos em diferentes escalas representam componentes do sinal  $f(\cdot)$  em diferentes resoluções.

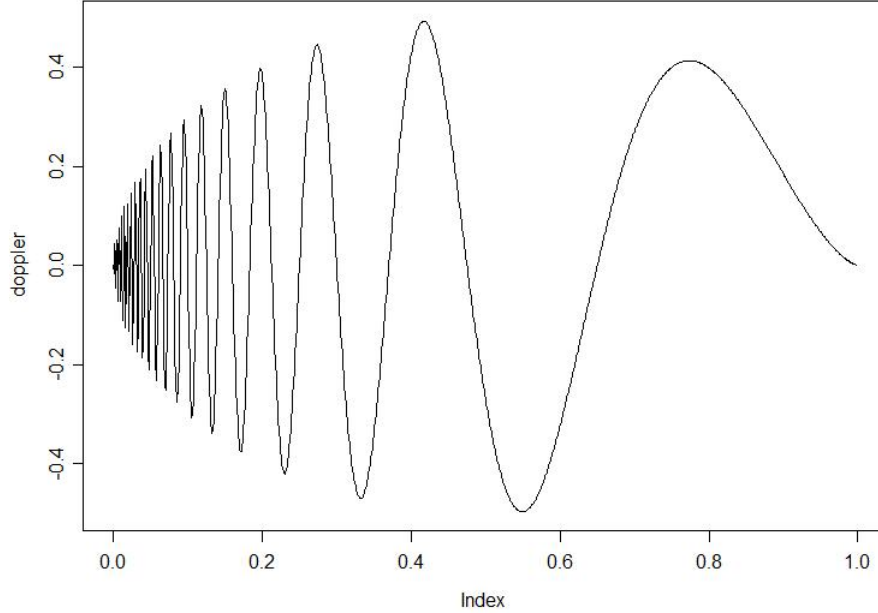
**Exemplo 4.6.** A Figura 4.4 apresenta o sinal Doppler que é definido por

$$f(t) = \sqrt{t(1-t)} \operatorname{sen}\left(\frac{2.1\pi}{t+0.05}\right), \text{ para todo } 0 \leq t \leq 1. \quad (4.30)$$

A Figura 4.5 apresenta a transformada discreta de wavelets do sinal Doppler, utilizando a wavelet s8. O sinal original é plotado na primeira linha e os coeficientes são plotados nas linhas seguintes partindo daqueles de escala fina  $d_{1,k}$ , apresentados na segunda linha até os de escala grossa  $d_{J,k}$  e  $s_{J,k}$ , apresentados nas duas últimas linhas. Na Figura 4.5, o sinal original é na verdade reconstruído pelos coeficientes por meio da transformada inversa discreta de wavelets (ITWD) dada na Definição 4.8.

A Figura 4.6 apresenta uma análise de multiresolução do sinal Doppler, utilizando a wavelet s8. As características finas (as oscilações de alta frequência no início do sinal) são capturadas principalmente pelos componentes

de detalhes de escala fina  $D_1$  e  $D_2$ . Os componentes de escala grossa  $D_6$  e  $S_6$  correspondem as baixas oscilações de frequência para o final da série.



**Figura 4.4:** Sinal Doppler.

**Teorema 4.8.** *Suponha que  $\{V_j; j \in \mathbb{Z}\}$  é uma análise de multiresolução com função escala  $\phi$ . Então para todo  $j \in \mathbb{Z}$ , o conjunto de funções*

$$\{\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k); k \in \mathbb{Z}\}$$

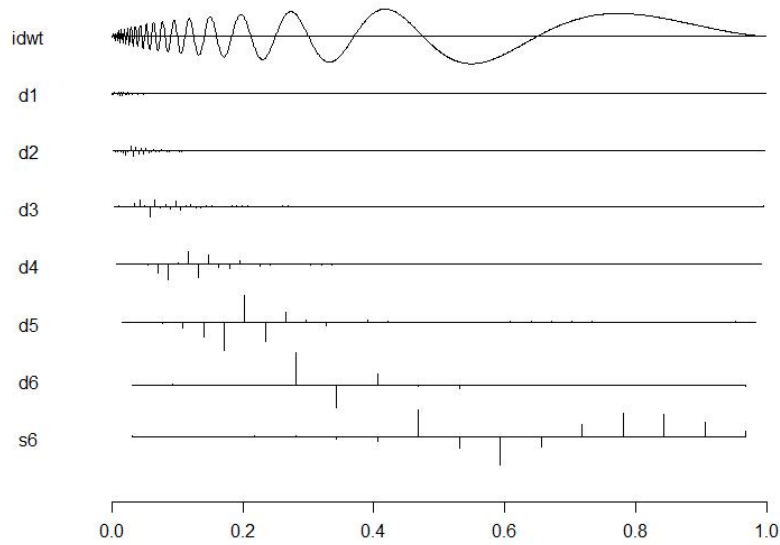
*é uma base ortonormal para  $V_j$ .*

**Demonstração:** Para mostrar que  $\{\phi_{j,k}; k \in \mathbb{Z}\}$  gera  $V_j$ , devemos mostrar que qualquer função  $f(t) \in V_j$  pode ser escrita como combinação linear de  $\{\phi(2^j t - k); k \in \mathbb{Z}\}$ . Usando a condição 4 da Definição 4.9, temos que a função  $f(2^{-j}t) \in V_0$  e portanto  $f(2^{-j}t)$  é uma combinação linear de  $\{\phi(t - k); k \in \mathbb{Z}\}$ . Substituindo  $t$  por  $2^j t$ , temos que  $f(t)$  é uma combinação linear de  $\{\phi(2^j t - k); k \in \mathbb{Z}\}$ .

Para mostrar que  $\{\phi_{j,k}; k \in \mathbb{Z}\}$  é ortonormal, devemos mostrar que

$$\langle \phi_{j,k}, \phi_{j,l} \rangle_{L^2} = \delta_{k,l} = \begin{cases} 0, & \text{se } k \neq l \\ 1, & \text{se } k = l \end{cases} \quad (4.31)$$





**Figura 4.5:** Transformada Discreta de Wavelets utilizando wavelet “s8”.

ou seja,

$$2^j \int_{-\infty}^{\infty} \phi(2^j t - k) \overline{\phi(2^j t - l)} dx = \delta_{k,l}.$$

Fazendo mudança de variáveis  $y = 2^j t$  ( $dy = 2^j dt$ ). Obtemos

$$2^j \int_{-\infty}^{\infty} \phi(2^j t - k) \overline{\phi(2^j t - l)} dt = \int_{-\infty}^{\infty} \phi(y - k) \overline{\phi(y - l)} dy = \delta_{k,l}$$

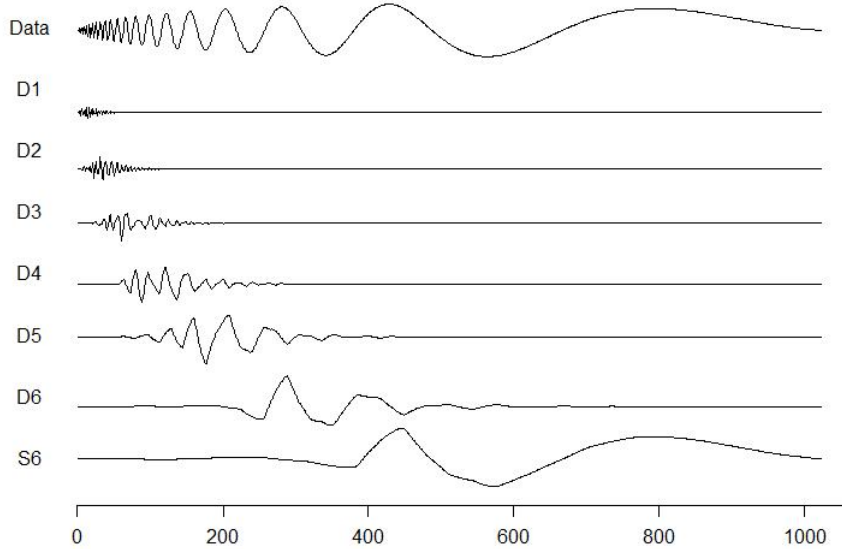
pela propriedade dada na Definição 4.9, condição 5.

□

Dado um espaço vetorial  $V$  com produto interno, o teorema a seguir fornece a representação da projeção ortonormal de um vetor  $\mathbf{v} \in V$  em um subespaço vetorial  $V_0$  com base ortonormal.

**Teorema 4.9.** *Seja  $V$  um espaço com produto interno e seja  $V_0$  um subespaço de dimensão  $N$  com base ortonormal  $\{e_1, e_2, \dots, e_N\}$ . A projeção ortonormal de um vetor  $\mathbf{v} \in V$  em  $V_0$  é dada por*

$$\mathbf{v}_0 = \sum_{j=1}^N \alpha_j e_j \quad \text{com} \quad \alpha_j = \langle \mathbf{v}, e_j \rangle.$$



**Figura 4.6:** Decomposição de Multiresolução do Sinal Doppler.

**Demonstração:** Ver Boggens and Narcowich (2001), página 16. □

O teorema a seguir, prova a equação central na análise de multiresolução, a relação escala.

**Teorema 4.10.** *Suponha que  $\{V_j; j \in \mathbb{Z}\}$  é uma análise de multiresolução com função escala  $\phi$ . Então, vale a seguinte relação*

$$\phi(t) = \sum_{k \in \mathbb{Z}} p_k \phi(2t - k) \quad \text{onde} \quad p_k = 2 \int_{-\infty}^{\infty} \phi(t) \overline{\phi(2t - k)} dt.$$

Além disso, temos também que

$$\phi(2^{j-1}t - l) = \sum_{k \in \mathbb{Z}} p_{k-2l} \phi(2^j t - k), \quad (4.32)$$

ou seja,

$$\phi_{j-1,l} = 2^{-1/2} \sum_{k \in \mathbb{Z}} p_{k-2l} \phi_{j,k} \quad (4.33)$$

onde  $\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k)$ .

**Demonstração:** Note que  $\phi(t) = \sum \tilde{p}_k \phi_{1,k}(t)$  vale para algum  $\tilde{p}_k$  pois  $\phi(t) \in V_0 \subset V_1$ , que é gerado por  $\{\phi_{1,k}, k \in \mathbb{Z}\}$ . Como o conjunto  $\{\phi_{1,k}, k \in \mathbb{Z}\}$  forma uma base ortonormal de  $V_1$ , o  $\tilde{p}_k$  pode ser determinado usando o Teorema 4.9

$$\tilde{p}_k = \langle \phi, \phi_{1,k} \rangle_{L^2} = 2^{1/2} \int_{-\infty}^{\infty} \phi(t) \overline{\phi(2t - k)} dt.$$

Portanto,

$$\phi(t) = \sum_{k \in \mathbb{Z}} \tilde{p}_k \phi_{1,k}(t) = \sum_{k \in \mathbb{Z}} \tilde{p}_k 2^{1/2} \phi(2t - k).$$

Seja  $p_k = 2^{1/2} \tilde{p}_k = 2 \int_{-\infty}^{\infty} \phi(t) \overline{\phi(2t - k)} dt$ . Portanto, temos

$$\phi(t) = \sum_{k \in \mathbb{Z}} p_k \phi(2t - k).$$

Para obtermos a equação (4.32), substituímos  $t$  por  $2^{j-1}t - l$  em  $\phi$  e ajustamos o índice da soma na série resultante,

$$\begin{aligned} \phi(2^{j-1}t - l) &= \sum_{k \in \mathbb{Z}} p_k \phi(2(2^{j-1}t - l) - k) \\ &= \sum_{k \in \mathbb{Z}} p_k \phi(2^j t - (2l + k)) \\ &= \sum_{k' \in \mathbb{Z}} p_{k' - 2l} \phi(2^j t - k'). \end{aligned}$$

A equação (4.33) segue-se da equação (4.32) multiplicada por  $2^{(j-1)/2}$ .  $\square$

**Exemplo 4.7.** Os valores de  $p_k$  para o sistema Haar são  $p_0 = p_1 = 1$  e os outros  $p_k$ 's são zero (ver expressão (4.19)).

Sabemos que  $V_j$  é um subconjunto de  $V_{j+1}$ . A fim de realizar o algoritmo de decomposição no caso geral, precisamos decompor  $V_{j+1}$  em uma soma ortogonal direta de  $V_j$  e seu complemento ortogonal  $W_j$  (como foi feito no sistema Haar). Além disso, precisamos construir uma função  $\psi$ , na qual, gera o espaço  $W_j$  (como foi feito no sistema Haar). O teorema a seguir, mostra que uma vez especificada a função  $\phi$ , a relação escala dada no Teorema 4.10, pode ser utilizada para construir a função  $\psi$  que gera o espaço  $W_j$ .

**Teorema 4.11.** *Suponha que  $\{V_j; j \in \mathbb{Z}\}$  é uma análise de multiresolução com função escala dada por*

$$\phi(t) = \sum_{k \in \mathbb{Z}} p_k \phi(2t - k),$$

onde o coeficiente  $p_k$ , para todo  $k \in \mathbb{Z}$ , é dado no Teorema 4.10. Seja  $W_j$  gerado por  $\{\psi(2^j t - k); k \in \mathbb{Z}\}$ , onde

$$\psi(t) = \sum_{k \in \mathbb{Z}} (-1)^k \overline{p_{1-k}} \phi(2t - k). \quad (4.34)$$

Então,  $W_j \subset V_{j+1}$  é um complemento ortogonal de  $V_j$  em  $V_{j+1}$ . Além disso,  $\{\psi_{j,k}(t) | \psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k), k \in \mathbb{Z} \text{ e } t \in \mathbb{R}\}$  forma uma base ortonormal para  $W_j$ .

**Demonstração:** Ver Boggens and Narcowich (2001), página 190. □

**Observação 4.6.** Para a função escala de Haar, os coeficientes  $p_0$  e  $p_1$  são ambos iguais a 1. Pelo Teorema 4.11 temos que

$$\psi(t) = \phi(2t) - \phi(2t - 1),$$

como foi dado na Observação 4.1.

O próximo teorema mostra que toda função  $f \in L^2(\mathbb{R})$ , pode ser expressa unicamente como uma soma  $\sum_{k=-\infty}^{\infty} \mathbf{w}_k$ , com  $\mathbf{w}_k \in W_k$ , onde os  $\mathbf{w}_k$ 's são mutualmente ortogonais.

**Teorema 4.12.** *Seja  $\{V_j, j \in \mathbb{Z}\}$  uma análise de multiresolução com função escala  $\phi$ . Seja  $W_j$  o complemento ortogonal de  $V_j$  em  $V_{j+1}$ . Então,*

$$L^2(\mathbb{R}) = \dots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \dots$$

*Em particular, cada  $f \in L^2(\mathbb{R})$  pode ser expressa unicamente como uma soma  $\sum_{k=-\infty}^{\infty} \mathbf{w}_k$ , com  $\mathbf{w}_k \in W_k$ , onde os  $\mathbf{w}_k$ 's são mutualmente ortogonais.*

*Equivalentemente, o subconjunto de todas wavelets,  $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ , forma uma base ortonormal para  $L^2(\mathbb{R})$ .*

**Demonstração:** Substituindo  $2^j t - l$  por  $t$  na expressão (4.34), multiplicando ambos os lados por  $2^{j/2}$  e ajustando o índice da soma, obtemos que  $\psi_{j,l}(t) = 2^{j/2} \psi(2^j t - l)$  tem a expansão

$$\psi_{j,l} = 2^{-1/2} \sum_{k \in \mathbb{Z}} (-1)^k \overline{p_{1-k+2l}} \phi_{j+1,k}.$$

Pelo Teorema 4.11, o conjunto  $\{\psi_{j-1,k}\}_{k \in \mathbb{Z}}$  é uma base ortonormal para o espaço  $W_{j-1}$ , o qual é um complemento ortogonal de  $V_{j-1}$  em  $V_j$  (portanto,

$V_j = W_{j-1} \oplus V_{j-1}$ ). Por decomposições ortogonais sucessivas, temos que

$$\begin{aligned} V_j &= W_{j-1} \oplus V_{j-1} \\ &= W_{j-1} \oplus W_{j-2} \oplus V_{j-2} \\ &\quad \dots \\ &= W_{j-1} \oplus W_{j-2} \oplus \dots \oplus W_0 \oplus V_0. \end{aligned} \quad (4.35)$$

Como o conjunto  $V_j$  foi também definido para  $j < 0$ , segue-se que

$$\begin{aligned} V_j &= W_{j-1} \oplus W_{j-2} \oplus \dots \oplus W_0 \oplus W_{-1} \oplus V_{-1} \\ &\quad \dots \\ &= W_{j-1} \oplus W_{j-2} \oplus \dots \oplus W_0 \oplus W_{-1} \oplus W_{-2} \dots \end{aligned} \quad (4.36)$$

Temos que  $V_j \subset V_{j+1}$  e a união de todos os  $V_j$  é o espaço  $L^2(\mathbb{R})$ . Portanto, para  $j$  tendendo à infinito, temos que

$$L^2(\mathbb{R}) = \dots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \dots .$$

□

## 4.5 Decomposição e Reconstrução

Nesta seção, descrevemos os algoritmos de decomposição e reconstrução associados com a análise de multiresolução. Estes algoritmos são análogos aos que foram apresentados para a wavelet Haar na Seção 4.1.

### 4.5.1 Decomposição

A fim de fazer processamento de um sinal, tal como codificação e compressão de dados, um algoritmo eficiente é necessário para decompor o sinal em partes que contém a informação sobre o comportamento oscilatório do sinal. Se utilizamos a análise de multiresolução, precisamos desenvolver um algoritmo que decompõe o sinal em seus componentes  $W_j$ , pois estes componentes fornecem a informação necessária. Existem três etapas principais para decompor um sinal  $f$ : iniciação, interação e terminação.

**Iniciação:** Esta etapa envolve duas partes. Primeiramente, temos que decidir qual o espaço de aproximação  $V_j$  que possibilita melhor ajuste às informações disponíveis sobre  $f$ . Isto é feito pela nossa escolha na análise do multiresolução. Na segunda parte, escolhemos  $f_j \in V_j$  tal que é o melhor ajuste para  $f$ . A melhor aproximação para  $f$  em  $V_j$ , no sentido de energia, é  $P_j f$ , a projeção ortogonal de  $f$  em  $V_j$ . Como  $2^{j/2}\phi(2^j t - k)$  é ortonormal, segue-se, pelo Teorema 4.9, que

$$P_j f(t) = \sum_{k \in \mathbb{Z}} a_k^j \phi(2^j t - k), \quad \text{onde} \quad a_k^j = 2^j \int_{-\infty}^{\infty} f(t) \overline{\phi(2^j t - k)} dt. \quad (4.37)$$

A informação do sinal nem sempre é suficiente para determinar exatamente os coeficientes  $a_k^j$ . Então, temos que aproximá-los utilizando a regra de quadratura dada pelo Teorema 4.13.

**Teorema 4.13.** *Seja  $\{V_j, j \in \mathbb{Z}\}$  uma análise de multiresolução com função escala  $\phi$  (de suporte compacto). Seja  $f \in L^2(\mathbb{R})$  contínua. Então,*

$$a_k^j = 2^j \int_{-\infty}^{\infty} f(t) \overline{\phi(2^j t - k)} dt \approx m f(k/2^j),$$

para  $j$  suficientemente grande, onde  $m = \int \overline{\phi(t)} dt$ .

**Demonstração:** Como  $\phi$  tem suporte compacto, o conjunto onde  $\phi$  é diferente de zero está contido em um intervalo da forma  $\{|u| \leq M\}$ . Então, o intervalo de integração para  $a_k^j$  em (4.37) é  $\{t; |2^j t - k| \leq M\}$ . Fazendo a mudança de variável  $u = 2^j t - k$ , obtemos

$$a_k^j = \int_{-M}^M f(2^{-j}u + 2^{-j}k) \overline{\phi(u)} du. \quad (4.38)$$

Quando  $j$  é grande,  $2^{-j}u + 2^{-j}k \approx 2^{-j}k$ , para  $u \in [-M, M]$ . Portanto,  $f(2^{-j}u + 2^{-j}k) \approx f(2^{-j}k)$ , para todo  $u \in [-M, M]$ , pois  $f$  é uniformemente contínua em qualquer intervalo finito. Em particular, podemos aproximar a integral dada em (4.38) substituindo  $f(2^{-j}u + 2^{-j}k)$  por  $f(2^{-j}k)$  para obter

$$a_k^j \approx f(k/2^j) \int_{-M}^M \overline{\phi(u)} du.$$

Relembrando que  $\phi$  é zero fora do intervalo  $[-M, M]$ , temos que

$$\int_{-M}^M \overline{\phi(u)} du = \int_{-\infty}^{\infty} \overline{\phi(u)} du = m.$$

Portanto,  $a_k^j = m f(k/2^j)$ . □

**Interação:** Após a etapa de inicialização, temos  $f \approx f_j \in V_j$ . Pelo Teorema 4.12, podemos começar com  $f_j$  e decompor  $f_j$  em uma soma de  $f_{j-1} \in V_{j-1}$  e  $w_{j-1} \in W_{j-1}$ , isto é,  $f_j = w_{j-1} + f_{j-1}$  onde

$$w_{j-1} = \sum_{l \in \mathbb{Z}} b_l^{j-1} \psi(2^{j-1}t - l)$$

$$f_{j-1} = \sum_{l \in \mathbb{Z}} a_l^{j-1} \phi(2^{j-1}t - l).$$

Os coeficientes  $a_l^{j-1}$  e  $b_l^{j-1}$  são determinados, de forma recursiva, a partir de  $a_j$  por

$$\begin{cases} a_l^{j-1} = 2^{-1} \sum_{k \in \mathbb{Z}} \overline{p_{k-2l}} a_k^j \\ b_l^{j-1} = 2^{-1} \sum_{k \in \mathbb{Z}} (-1)^k p_{1-k+2l} a_k^j. \end{cases} \quad (4.39)$$

Repetimos o argumento com  $f_{j-1}$ , depois com  $f_{j-2}$  e assim por diante. Se pararmos a decomposição no nível zero, então obtemos

$$f \approx f_j = w_{j-1} + w_{j-2} + \cdots + w_0 + f_0. \quad (4.40)$$

**Definição 4.10 (Convolução Discreta).** Dadas duas sequências  $\mathbf{x} = (\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots)$  e  $\mathbf{y} = (\cdots, y_{-2}, y_{-1}, y_0, y_1, y_2, \cdots)$ , a *convolução*  $\mathbf{x} * \mathbf{y}$  é definida por

$$(\mathbf{x} * \mathbf{y})_k = \sum_{n \in \mathbb{Z}} x_{k-n} y_n, \quad (4.41)$$

contanto que a série envolvida em (4.41) seja absolutamente convergente.

Sejam  $h$  e  $l$  as sequências dadas por

$$h_k := \frac{1}{2} (-1)^k p_{k+1}$$

e

$$l_k := \frac{1}{2} \overline{p_{-k}}.$$

Sejam  $H$  e  $L$  os filtros definidos por  $H(\mathbf{x}) = h * \mathbf{x}$  e  $L(\mathbf{x}) = l * \mathbf{x}$ , respectivamente. Considere  $\mathbf{x} = a^j$  e note que  $L(a^j)_l = \frac{1}{2} \sum_{k \in \mathbb{Z}} \overline{p_{k-l}} a_k^j$ . Comparando com a equação (4.39), vemos que  $a_l^{j-1} = L(a^j)_{2l}$ . Similarmente,  $b_l^{j-1} = H(a^j)_{2l}$ .

**Definição 4.11 (Operador *Downsampling*).** Seja  $\mathbf{x} = (\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots)$  uma sequência real. Definimos o operador *downsampling*  $D$  através de

$$D\mathbf{x} = (\cdots, x_{-2}, x_0, x_2, \cdots)$$

ou  $(D\mathbf{x})_l = x_{2l}$ , para todo  $l \in \mathbb{Z}$ .

Utilizando a Definição 4.11 podemos agora formular a etapa iterativa no algoritmo utilizando filtros discretos:

$$\begin{cases} \text{Forma de Convolução: } a^{j-1} = D(l * a_j) & \text{e } b^{j-1} = D(h * a^j) \\ \text{Forma de Operador: } a^{j-1} = DL a^j & \text{e } b^{j-1} = DH a^j. \end{cases} \quad (4.42)$$

Os dois filtros que utilizamos,  $h$  e  $l$ , são chamados de *filtros decomposição passa alto* e *decomposição passa baixo*, respectivamente.

**Terminação:** Existem diversos critérios para terminar a decomposição. O mais simples é aquele em que decomparamos até que esgotemos o número finito de amostras que tomamos. Geralmente, a escolha do ponto de parada depende extremamente do que desejamos realizar.

## 4.5.2 Reconstrução

Uma vez que o sinal  $f$  foi decomposto, alguns componentes  $W_{j'}$  podem ser modificados. Se queremos filtrar ruído, então os componentes  $W_{j'}$  de  $f$  correspondente às frequências não desejadas, podem ser retirados e o sinal resultante ficará significativamente com menos ruído. Se o objetivo é compressão de dados, os componentes  $W_{j'}$  que são pequenos podem ser retirados, sem mudança significativa no sinal. Apenas os componentes significantes  $W_{j'}$  precisam ser transmitidos para obter uma adequada compressão de dados. Em ambos os casos, como os componentes  $W_{j'}$  são modificados, precisamos de um algoritmo de reconstrução para reconstruir o sinal comprimido ou filtrado em termos dos elementos da base  $\phi(2^j t - l)$ , de  $V_j$ . A idéia é reconstruir  $f_j \approx f$  (ver expressão (4.40)) utilizando  $f_{j'} = f_{j'-1} + w_{j'-1}$ , começando em  $j' = 1$ . Novamente, dividimos o algoritmo em três partes: inicialização, interação e terminação.

**Inicialização:** O que temos disponível é um conjunto de possíveis coeficientes modificados, onde inclui os coeficientes de aproximação  $\{a_k^0\}$  e os coeficientes de detalhes  $\{b_k^{j'}\}$ , para  $j' = 0, \dots, j$ . Estes coeficientes aparecem nas formas

$$f_0(t) = \sum_{k \in \mathbb{Z}} a_k^0 \phi(t - k) \in V_0 \quad (4.43)$$

$$w_{j'}(t) = \sum_{k \in \mathbb{Z}} b_k^{j'} \psi(2^{j'} t - k) \in W_{j'}, \quad \text{para } 0 \leq j' < j, \quad (4.44)$$

onde

$$a_k^j = \sum_{l \in \mathbb{Z}} p_{k-2l} a_l^{j-1} + \sum_{l \in \mathbb{Z}} (-1)^k \overline{p_{1-k+2l}} b_l^{j-1}. \quad (4.45)$$

**Interação:** Formulamos esta etapa em termos de filtros discretos. Sejam  $\tilde{h}$  e  $\tilde{l}$  seqüências dadas por

$$\tilde{h}_k := \overline{p_{1-k}} (-1)^k \quad (4.46)$$

$$\tilde{l}_k := p_k. \quad (4.47)$$

Definimos dois filtros de seqüências, denotados por  $\tilde{H}$  e  $\tilde{L}$ , via  $\tilde{H}(\mathbf{x}) = \tilde{h} * \mathbf{x}$  e  $\tilde{L}(\mathbf{x}) = \tilde{l} * \mathbf{x}$ . A fórmula da reconstrução dada pela equação (4.45) fornece  $a_k^j$  como uma soma dos dois termos  $\sum_{l \in \mathbb{Z}} p_{k-2l} a_l^{j-1}$  e  $\sum_{l \in \mathbb{Z}} (-1)^k \overline{p_{1-k+2l}} b_l^{j-1}$ . Utilizando os filtros de seqüências  $\tilde{H}$  e  $\tilde{L}$ , podemos escrever (4.45) como

$$a_k^j = \sum_{l \in \mathbb{Z}} \tilde{l}_{k-2l} a_l^{j-1} + \sum_{l \in \mathbb{Z}} \tilde{h}_{k-2l} b_l^{j-1}. \quad (4.48)$$



Isto é quase uma soma de convoluções, a única diferença é que o índice para a convolução é  $k - l$  no lugar de  $k - 2l$ . Em outras palavras, a expressão (4.48) é uma convolução sem os termos ímpares. Os termos ímpares podem ser recolocados na expressão (4.48) simplesmente multiplicando-os eles por zero

$$\begin{aligned} a_k^j &= \cdots + \tilde{l}_{k+4}a_{-2}^{j-1} + \tilde{l}_{k+3} \cdot 0 + \tilde{l}_{k+2}a_{-1}^{j-1} + \tilde{l}_{k+1} \cdot 0 + \tilde{l}_k a_0^{j-1} + \tilde{l}_{k-1} \cdot 0 + \cdots \\ &+ \tilde{h}_{k+4}b_{-2}^{j-1} + \tilde{h}_{k+3} \cdot 0 + \tilde{h}_{k+2}b_{-1}^{j-1} + \tilde{h}_{k+1} \cdot 0 + \tilde{h}_k b_0^{j-1} + \cdots \end{aligned} \quad (4.49)$$

Para colocar esta soma na forma de uma convolução, definimos a seguir o operador *upsampling* (ver Definição 4.12).

**Definição 4.12 (Operador *upsampling*).** Seja  $\mathbf{x} = (\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots)$  uma sequência real. Definimos o operador *upsampling*  $U$  através de

$$U\mathbf{x} = (\cdots, x_{-2}, 0, x_{-1}, 0, x_0, 0, x_1, 0, x_2, 0, \cdots)$$

ou

$$(U\mathbf{x})_k = \begin{cases} 0, & \text{se } k \text{ é ímpar} \\ x_{k/2}, & \text{se } k \text{ é par.} \end{cases}$$

Utilizando a Definição 4.12, podemos agora formular a etapa iterativa no algoritmo de reconstrução utilizando filtros discretos:

$$\begin{cases} \text{Forma de Convolução: } a^j = \tilde{l} * (Ua^{j-1}) + \tilde{h} * (Ub^{j-1}) \\ \text{Forma de Operador: } a^j = \tilde{L}Ua^{j-1} + \tilde{H}Ub^{j-1}. \end{cases} \quad (4.50)$$

**Terminação:** Os algoritmos de decomposição e reconstrução utilizam os coeficientes escalas  $p_k$ , mas não as fórmulas reais para  $\phi(\cdot)$  e  $\psi(\cdot)$ . Para plotar o sinal reconstruído  $f(t) = \sum_l a_l^j \phi(2^j t - l)$ , podemos aproximar os

valores de  $f$  em  $t = l/2^j$  por  $a_l^j$  (pelo Teorema 4.13, com  $\int \phi(t)dt = 1$ ). Assim as fórmulas para  $\phi(\cdot)$  e  $\psi(\cdot)$  não são utilizadas no gráfico do sinal reconstruído. De qualquer modo,  $\phi(\cdot)$  e  $\psi(\cdot)$  são importantes para o sucesso dos algoritmos de decomposição e reconstrução.

## 4.6 Encolhimento de Wavelets

Nesta seção, discutimos estimação por encolhimento de wavelets que é uma das aplicações mais importantes da teoria de wavelets na Estatística. Encolhimento de wavelets refere-se geralmente às reconstruções obtidas dos coeficientes encolhidos. A metodologia é baseada no princípio do encolhimento de wavelets, que tem por objetivo a redução do ruído presente no sinal, reduzindo ou zerando a magnitude dos coeficientes de wavelets. Donoho e

Johnstone (1994) desenvolveram uma importante teoria para a regressão não-paramétrica e suavizada baseada no princípio de encolhimento de wavelets. Para mais detalhes ver também Vidakovic (1999).

Nos problemas não-paramétricos, desejamos estimar um sinal  $f(t_j)$  a partir de alguns dados  $y_j$  com ruídos. Suponhamos que estes dados são obtidos a partir da expressão

$$y_j = f(t_j) + \epsilon_j, \quad j = 1, 2, \dots, n, \quad (4.51)$$

onde  $f(t_j) = f_j$  é um sinal discreto e  $\epsilon_j$  são erros independentes e normalmente distribuídos, isto é,  $\epsilon_j \sim N(0, \sigma^2)$  e  $t_j = (j - 1)/n$ . A regressão paramétrica usual exige que  $f(\cdot)$  obedeça a um modelo particular (por exemplo,  $f(\cdot)$  é uma função linear de  $p$  parâmetros  $\theta_1, \dots, \theta_p$  como um  $AR(p)$ ). A estimação é feita utilizando-se o modelo de mínimos quadrados ordinários. Na situação do modelo (4.51), assumimos que  $f(\cdot)$  pertence à uma classe de funções, satisfazendo certas propriedades de regularidades. Se  $f_j = f(t_j)$  então a finalidade é estimar  $f = (f_1, \dots, f_n)'$ , com o menor erro quadrático médio possível. Para uma classe de funções  $\mathfrak{F}$ , desejamos encontrar  $\hat{f} = \hat{f}(y_1, \dots, y_n)$  que atinja o risco mínimo

$$R(n, \mathfrak{F}) = \inf_{\hat{f}} \sup_f \mathfrak{R}(\hat{f}, f), \quad (4.52)$$

onde  $\mathfrak{R}(\hat{f}, f) = (1/n)\mathbb{E}\|\hat{f} - f\|_2^2$  e  $\|f\|_2^2 = \sum_{j=1}^n f_j^2$ .

O procedimento *encolhimento de wavelets* consiste de três etapas:

1. Aplicar a *transformada discreta de wavelet*, com  $J$  níveis, para o sinal  $f(\cdot)$ , obtendo coeficientes contaminados por ruídos  $d_1, d_2, \dots, d_J, s_J$ .
2. Usar limiares (*threshold*) para diminuir os coeficientes, tornando nulos aqueles coeficientes abaixo de determinado valor, nas escalas  $j$  mais finas, para obtermos novos coeficientes  $\tilde{d}_1 \equiv \delta_{\lambda_1 \sigma_1}(d_1), \dots, \tilde{d}_j \equiv \delta_{\lambda_j \sigma_j}(d_j)$ , onde  $\lambda_1, \dots, \lambda_j$  são os limiares. A *função encolhimento*  $\delta_c(x)$  encolhe  $x$  para zero e é parametrizada por  $c = \lambda \sigma$ , onde  $\lambda$  é o limiar (ver Definição 4.14) e  $\sigma$  é uma estimativa da escala do ruído.
3. Aplicar a *transformada inversa discreta de wavelets* usando os coeficientes detalhes  $\tilde{d}_1, \dots, \tilde{d}_j, d_{j+1}, \dots, d_J, s_J$  para obter o estimador encolhimento de wavelets de  $f(\cdot)$ .

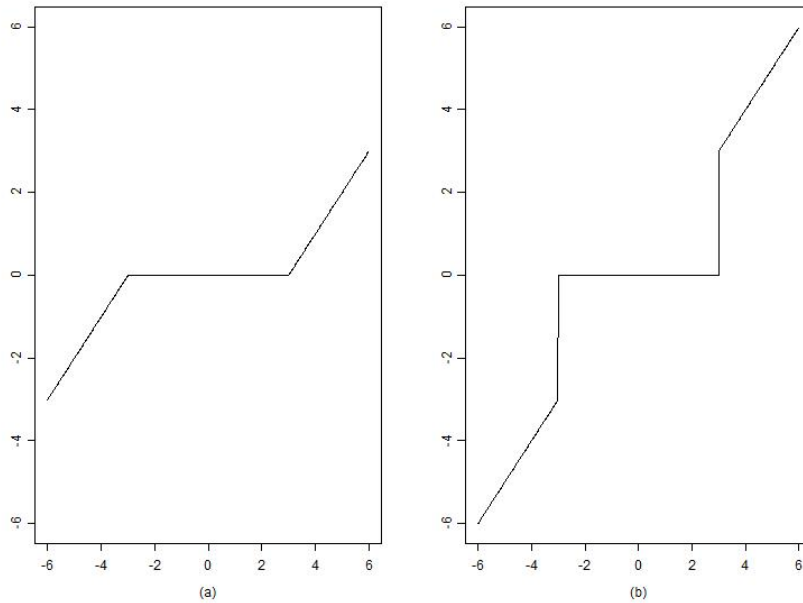
**Definição 4.13 (Função de Encolhimento).** Existem duas formas básicas para a função de encolhimento  $\delta_c(\cdot)$ . A *função de encolhimento suave* (*soft shrinkage function*) é definida por

$$\delta_c^S(x) = \begin{cases} 0, & \text{se } |x| \leq \lambda \\ \text{sgn}(x)(|x| - \lambda), & \text{se } |x| > \lambda \end{cases} \quad (4.53)$$

e a função de encolhimento dura (*hard shrinkage function*) é definida por

$$\delta_c^H(x) = \begin{cases} 0, & \text{se } |x| \leq \lambda \\ x, & \text{se } |x| > \lambda. \end{cases} \quad (4.54)$$

Existe uma descontinuidade na função de encolhimento dura enquanto que a de encolhimento suave é uma função contínua. A motivação para o uso da função de encolhimento suave vem do princípio que o ruído afeta todos os coeficientes das wavelets. Além disso, a continuidade da função a torna preferível por razões estatísticas. A Figura 4.7 apresenta um exemplo de como as duas funções de encolhimento agem sobre a função  $f(x) = x$ .



**Figura 4.7:** Funções de Encolhimento com  $\lambda = 3$  aplicadas na Função Linear  $f(x) = x$ : (a) Função Encolhimento Suave; (b) Função Encolhimento Dura.

Existem diferentes regras para selecionar o limiar  $\lambda_j$  da função de encolhimento  $\delta_{\lambda_j \sigma_j}(\cdot)$ , algumas delas são motivadas da estatística (ver Morettin, 1999). Donoho e Johnstone (1995) propõem o limiar universal  $\lambda$  baseado no resultado da teoria de valores extremos de variáveis aleatórias normais.

**Teorema 4.14.** *Seja  $Z_1, \dots, Z_n$  uma sequência de variáveis aleatórias independentes e identicamente distribuídas com distribuição normal. Defina*

$$A_n = \left\{ \max_{j=1, \dots, n} |Z_j| \leq \sqrt{2 \log(n)} \right\}.$$

Então,

$$\pi_n = \mathbb{P}(A_n) \rightarrow 0, \text{ quando } n \rightarrow \infty.$$

Além disso, se

$$B_n(t) = \left\{ \max_{j=1, \dots, n} |Z_j| > t + \sqrt{2 \log(n)} \right\},$$

então  $\mathbb{P}(B_n(t)) < e^{-\frac{t^2}{2}}$ .

Informalmente, o Teorema 4.14 diz que as variáveis aleatórias  $Z_j$ 's são quase limitadas por  $\pm \sqrt{2 \log(n)}$ . Isto motiva o uso do seguinte limiar (ver Definição 4.14).

**Definição 4.14 (Limiar Universal).** O *limiar universal*, é definido por

$$\lambda = \lambda_j = \hat{\sigma} \sqrt{2 \log(n)},$$

que não depende da escala, onde  $\hat{\sigma}$  é uma estimativa do nível do ruído e  $n$  é o tamanho do sinal.

O nível do ruído  $\sigma$  é desconhecido e deve ser estimado a partir dos dados. Quase todo método para estimar a variância do ruído envolve os coeficientes de wavelets. Três possibilidades podem ser consideradas para o calcular o fator de escala  $\sigma_j$ , para todo  $j = 1, 2, \dots, J$ .

- a) Os coeficientes  $\mathbf{d}_1$ , da escala mais fina, são usados para estimar um fator de escala para todos os níveis:  $\sigma_j = \hat{\sigma}(\mathbf{d}_1)$ , para todo  $j = 1, 2, \dots, J$ .
- b) Consideram-se todas as escalas e um estimador de  $\sigma$  baseado em todos os coeficientes de wavelets:  $\sigma_j = \hat{\sigma}(\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_J)$ .
- c) Estima-se  $\sigma_j$  para cada escala:  $\sigma_j = \hat{\sigma}(\mathbf{d}_j)$ , para  $j = 1, 2, \dots, J$ .

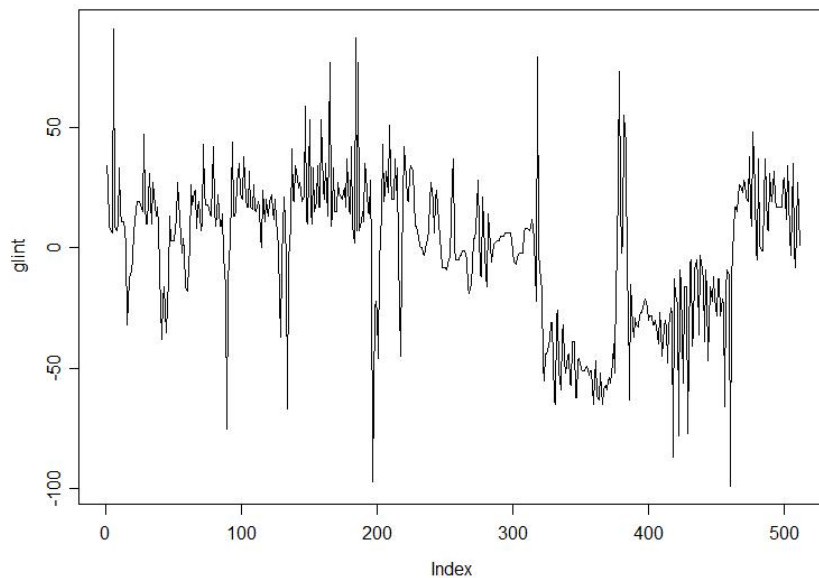
Em qualquer caso, pode-se usar a variância dos coeficientes ou, então, o estimador proposto por Donoho et al. (1995), dado por

$$\hat{\sigma} = \frac{\text{med}\{|d_{J-1,k}| : 0 \leq k < 2^J\}}{0.6745},$$

onde *med* indica a mediana.

**Exemplo 4.8 (Eliminando Ruído Utilizando Wavelets).** Para ilustrar o procedimento de encolhimento de wavelets vamos considerar o sinal “glint” dado na Figura 4.8. O sinal “glint” representa medidas do ruído do radar de “glint”. Para eliminar o ruído do sinal “glint” aplicamos o procedimento de encolhimento no sinal e depois aplicamos a transformada discreta

de wavelet no sinal estimado. A Figura 4.9 apresenta a transformada discreta de wavelets do sinal “glint” baseada na wavelet s8 enquanto que a Figura 4.10 apresenta a transformada discreta de wavelets do sinal “glint” estimado, obtido pelo procedimento de encolhimento de wavelets utilizando a wavelet s8. A transformada discreta de wavelets do sinal estimado é muito similar à transformada discreta de wavelet do sinal “glint”. A principal diferença é que algumas escalas finas  $d_{1,k}$  e  $d_{2,k}$  foram encolhidas para zero. Note que a razão heurística para o procedimento de encolhimento de wavelets segue-se de dois princípios: diminuir os coeficientes, tornando nulos aqueles coeficientes abaixo de um determinado valor. O procedimento elimina o ruído enquanto recupera as principais características do sinal.

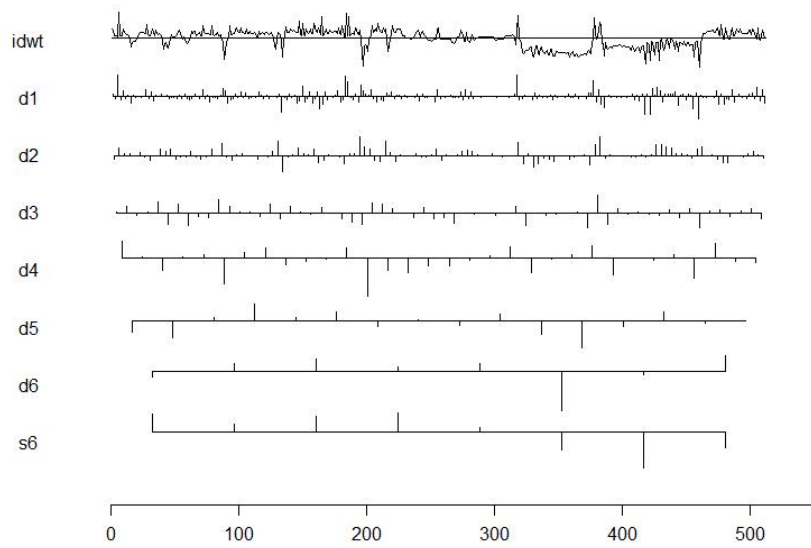


**Figura 4.8:** Ruído “Glint” de Radar.

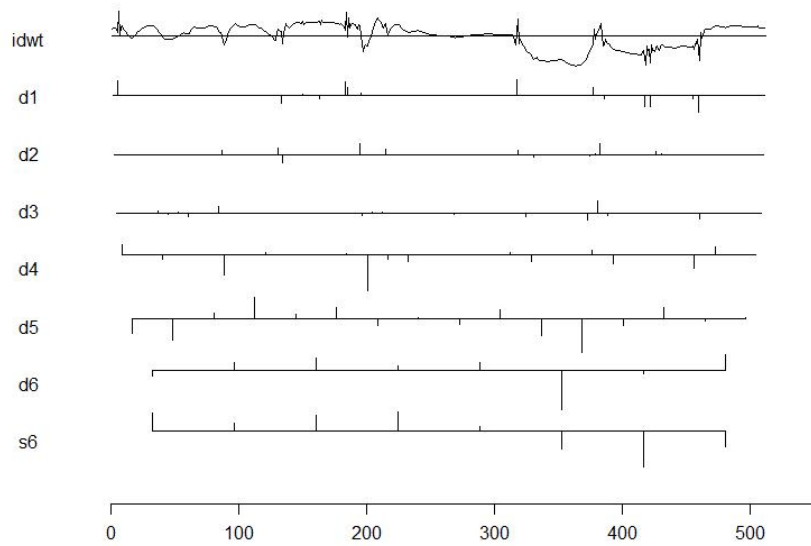
## 4.7 Núcleos de Wavelets

Dado  $V$  um espaço de  $L^2$ , para qualquer função  $f \in V$ , apresentamos nesta seção a projeção de  $f$  no espaço  $V_j$ , para todo  $j = 1, 2, \dots, J$ .

**Definição 4.15 (Função de Reprodução de Núcleo).** A função  $\mathbb{K}(s, t)$  de duas variáveis  $s$  e  $t$ , é chamada *função de reprodução de núcleo* para o espaço de funções  $\mathcal{H}$  (de Hilbert) se



**Figura 4.9:** TDW do Sinal “Glint” (baseada na wavelet “s8”).



**Figura 4.10:** TDW do Sinal Estimado (baseada na wavelet “s8”).

- (i) para um  $t$  fixo,  $\mathbb{K}(s, t)$  é uma função em  $\mathcal{H}$ .
- (ii) para qualquer função  $f \in \mathcal{H}$  e qualquer  $t$ ,  $\mathbb{K}(s, t)$  tem a seguinte propriedade

$$f(t) = \langle f(s), \mathbb{K}(s, t) \rangle.$$

**Teorema 4.15.** *Seja  $V$  um subespaço de  $L^2$  e seja  $\{e_1, e_2, \dots\}$  uma base ortonormal de  $V$ . Então,  $V$  é um “reproducing kernel Hilbert space” com o núcleo (algumas vezes chamado de núcleo de Bergman)*

$$\mathbb{K}(s, t) = \sum_{k=1}^{\infty} e_k(s)e_k(t). \quad (4.55)$$

Para qualquer função  $f \in V$ ,

$$f(t) = \int f(s)\mathbb{K}(s, t)ds.$$

**Demonstração:** Ver Vidakovic (1999), página 29. □

Pelo Teorema 4.15, segue-se que

$$\mathbb{K}(s, t) = \sum_k \phi(s - k)\phi(t - k) \quad (4.56)$$

é uma reprodução de núcleo de  $V_0$ . Pela auto-similaridade dos subespaços de multiresolução

$$\mathbb{K}_j(s, t) = 2^j \mathbb{K}(2^j s, 2^j t) \quad (4.57)$$

é uma função de reprodução de núcleo de  $V_j$ . Então, a projeção de  $f$  no espaço  $V_j$  é dada por

$$\text{Proj}_{V_j} f(t) = \int 2^j \mathbb{K}(2^j s, 2^j t) f(s) ds.$$

**Lema 4.2.** *Para todo  $j \in \{1, 2, \dots, J\}$ ,  $\int \mathbb{K}_j(s, t) dt = 1$ .*

**Demonstração:** Como  $\sum_k \phi(t - k) = 1$  então, para todo  $j \in \{1, 2, \dots, J\}$ ,

$$\begin{aligned} 1 &= \int \phi(t) dt = \sum_k \phi(2^j t - k) \int \phi(t) dt \\ &= \sum_k \phi(2^j s - k) \cdot 2^j \int \phi(2^j t - k) dt = \int 2^j \mathbb{K}(2^j s, 2^j t) dt \\ &= \int \mathbb{K}_j(s, t) dt. \end{aligned} \quad (4.58)$$

□

O seguinte resultado, é útil na exploração dos estimadores de núcleos de wavelets.

**Teorema 4.16.** *Seja  $\mathbb{K}_j(s, t)$  um núcleo de wavelets do espaço  $V_j$  gerado pela função escala  $\phi$ , para todo  $j \in \{1, 2, \dots, J\}$ .*

(a) *Se  $\phi$  tem um decaimento exponencial, i.e.,  $\phi(t) \leq e^{-a|t|}$ , para algum  $a$  positivo, então  $|\mathbb{K}_j(s, t)| \leq C2^j e^{-a2^j|s-t|/2}$ .*

(b) *Se  $\phi$  tem decaimento algébrico, i.e.,  $\phi(t) \leq C_N/(1 + |t|)^N$ , para algum  $N > 1$ , então  $|\mathbb{K}_j(s, t)| \leq C_N 2^j / (1 + 2^j|s - t|)^N \leq C_N 2^j$ , para  $N > 1$ .*

**Demonstração:** Ver Kelly, Kon e Raphael (1994). □

## 4.8 Wavelets e Processos Estacionários

Nesta seção apresentamos resultados que relacionam processos estacionários de tempo contínuo com a teoria de wavelets. Consideramos apenas série de wavelet e transformada discreta de wavelets.

### 4.8.1 Transformadas de Wavelets de Processos Estacionários

Seja  $\{X(t)\}_{t \in \mathbb{R}}$  um processo estacionário com função de autocorrelação  $\gamma_X(h)$ . A transformada discreta de wavelets de  $\{X(t)\}_{t \in \mathbb{R}}$  é um conjunto aleatório discreto

$$\{d_{jk}, j, k \in \mathbb{Z}\} = \left\{ \int_{\mathbb{R}} X(t) \psi_{j,k}(t) dt, j, k \in \mathbb{Z} \right\}. \quad (4.59)$$

Este conjunto está bem definido se as integrais em (4.59) estão bem definidas e

$$\int_{\mathbb{R}} \sqrt{\gamma_X(t)} |\psi_{j,k}(t)| dt < \infty. \quad (4.60)$$

Assim, se vale a expressão (4.60) então

$$\mathbb{E}(d_{j,k} d_{j',k'}) = \int \int_{\mathbb{R}^2} \gamma_X(t, s) \psi_{j,k}(t) \psi_{j',k'}(s) dt ds.$$

Qualquer nível em uma transformada de wavelets de um processo estacionário é também uma sequência estacionária (ver Lema 4.3), mas não vale o inverso (ver Exemplo 4.9).

**Lema 4.3.** *Seja  $\{X(t)\}_{t \in \mathbb{R}}$  um processo fracamente estacionário. Para  $l, n \in \mathbb{Z}$  e  $j \geq l$ , a sequência  $\{d_{j, 2^{j-l}k+n}, k \in \mathbb{Z}\}$  é também um processo fracamente estacionário.*



**Demonstração:** Ver Vidakovic (1999). □

**Exemplo 4.9.** Considere o processo estocástico definido por

$$X(t) = \sum_k Z_k \psi(t - k),$$

onde os  $Z_k$ 's são variáveis aleatórias uniformemente distribuídas no intervalo  $(-1, 1)$ , i.é.,  $Z_k \sim U(-1, 1)$ , para todo  $k \in \mathbb{Z}$ . É um processo de segunda ordem mas não é necessariamente estacionário. Os correspondentes níveis na decomposição de wavelets de  $\{X(t)\}_{t \in \mathbb{R}}$  são claramente estacionários pois todos eles são zeros ou  $Z_k, k \in \mathbb{Z}$  (o nível correspondente ao  $W_0$ ).

**Teorema 4.17.** *Seja  $\{X(t)\}_{t \in \mathbb{R}}$  um processo estacionário tal que a função de autocovariância  $\gamma_X(t, s)$  é limitada e contínua em  $\mathbb{R}^2$ . Sejam  $d_{j,k}$  definidos para todo  $j, k$ . Se  $\{X(t)\}_{t \in \mathbb{R}}$  é fracamente estacionário, então o processo formado pelo conjunto  $\{d_{j,k}, j, k \in \mathbb{Z}\}$  é fracamente estacionário.*

**Demonstração:** Ver Averkamp e Houdré (1998). □

## 4.8.2 Whitening de Processos Estacionários

Nesta seção apresentamos resultados que relacionam processos estacionários de tempo contínuo e base de wavelets do tipo Meyer.

**Definição 4.16 (Wavelets Tipo Meyer).** *As funções escala e wavelet mãe das wavelets tipo Meyer (ver Walter, 1994) são definidas no domínio da frequência por*

$$\phi(t) = \frac{1}{\pi} \int_0^{4\pi/3} \Phi(\omega) \cos(\omega t) d\omega,$$

e

$$\psi(t) = -\frac{1}{\pi} \int_{2\pi/3}^{4\pi/3} \Phi(\omega/2) \Phi(\omega - 2\pi) \cos \omega(t - 1/2) d\omega,$$

onde

$$\Phi(\omega) = \left( \int_{\omega-\pi}^{\omega+\pi} d\mathbb{P} \right)^{1/2}, \quad \Psi(\omega) = e^{-i\omega/2} \left( \int_{\omega/2-\pi}^{\omega/2+\pi} d\mathbb{P} \right)^{1/2}$$

e  $\mathbb{P}$  é alguma medida de probabilidade sobre o intervalo  $[-\frac{\pi}{3}, \frac{\pi}{3}]$ .

**Observação 4.7.** A ortogonalidade das wavelets tipo Meyer é verificada facilmente por

$$\begin{aligned} \sum_k |\Phi(\omega + 2k\pi)|^2 &= \sum_k \left( \int_{\omega+2k\pi-\pi}^{\omega+2k\pi+\pi} d\mathbb{P} \right) \\ &= \sum_k \int_{\omega-(2k-1)\pi}^{\omega+(2k+1)\pi} d\mathbb{P} = \int_{\mathbb{R}} d\mathbb{P} = 1. \end{aligned} \quad (4.61)$$

Seja  $\{X(t)\}_{t \in \mathbb{R}}$ , um processo estacionário e seja  $\{X_m(t)\}_{t \in \mathbb{R}}$  sua projeção no espaço de multiresolução  $V_m$ , i.e.,

$$X_m(t) = \sum_k c_{m,k} \phi_{m,k}(t), \quad (4.62)$$

onde  $c_{m,k} = \int X(t) \phi_{m,k}(t) dt$ . Se a wavelet  $\phi$  é r-regular (ver Definição 4.6),

$$\mathbb{E}[X(t) - X_m(t)]^2 \rightarrow 0, \quad \text{quando } m \rightarrow \infty. \quad (4.63)$$

Certamente, desde que

$$\begin{aligned} \mathbb{E}[X(t) - X_m(t)]^2 &= \mathbb{E}[X^2(t)] - 2\mathbb{E}[X(t)X_m(t)] + \mathbb{E}[X_m^2(t)] \\ &= \gamma_X(0) - 2 \int \gamma_X(t-s) \mathbb{K}_m(t,s) ds \\ &\quad + \int \int \gamma_X(u-s) \mathbb{K}_m(s,t) \mathbb{K}_m(t,u) dud s \end{aligned} \quad (4.64)$$

e

$$\int \gamma_X(u-s) \mathbb{K}_m(t,s) ds \rightarrow \gamma_X(u-t), \quad \text{quando } m \rightarrow \infty,$$

uniformemente em conjuntos limitados. Então, a expressão (4.63) é válida.

A declaração informal “*wavelets whiten data*” pode ser formalizada se os dados tem uma estrutura de dependência estacionária. Considere uma representação equivalente de (4.62) dada por

$$X_m(t) = \sum_{j=-\infty}^{m-1} \sum_k d_{j,k} \psi_{m,k}(t),$$

onde

$$d_{j,k} = \int_{\mathbb{R}} X(t) \psi_{j,k}(t) dt. \quad (4.65)$$

Então,

$$\begin{aligned} \mathbb{E}(d_{j,k} d_{j',k'}) &= \frac{1}{2\pi} \int_{\mathbb{R}} \hat{\gamma}_X(\omega) \Psi\left(\frac{\omega}{2^j}\right) \overline{\Psi\left(\frac{\omega}{2^{j'}}\right)} \\ &\quad \times e^{-i\omega k 2^{-j}} e^{i\omega k' 2^{-j'}} 2^{-j/2} 2^{-j'/2} d\omega, \end{aligned} \quad (4.66)$$

onde  $\Psi(\cdot)$  e  $\hat{\gamma}_X(\cdot)$  são transformadas de Fourier de  $\psi(\cdot)$  e  $\gamma_X(\cdot)$ , respectivamente. Explorando (4.66), Walter (1994) provou o resultado apresentado no Teorema 4.18 a seguir.

**Teorema 4.18.** *Seja  $\{X(t)\}_{t \in \mathbb{R}}$ , um processo estacionário. Consideremos uma base de wavelets do tipo Meyer, tal que as transformadas de Fourier de  $\psi(\cdot)$  e  $\gamma_X(\cdot)$ , denotadas por  $\Psi(\cdot)$  e  $\hat{\gamma}_X(\cdot)$ , respectivamente, pertencem ao espaço  $\mathbb{C}^p$ ,  $p > 1$ . Então, os coeficientes  $d_{j,k}$  e  $d_{j',k'}$  definidos em (4.65)*

(i) são correlacionados, se  $|j - j'| > 1$ ,

(ii) tem arbitrariamente correlação pequena, se  $|j - j'| = 1$  e

(iii) tem correlação da ordem de  $O(|k - k'|^{-p})$ , se as escalas  $j$  e  $j'$  coincidem.

**Demonstração:** Ver Vidakovic (1999), página 275.

□

# Capítulo 5

## Teoria dos Valores Extremos

A teoria dos valores extremos desempenha um papel fundamental na modelagem de eventos associados a probabilidades muito pequenas ou eventos raros. Problemas relacionados a valores extremos existem na astronomia, meteorologia, engenharia naval, oceanografia, controle de qualidade, em estatísticas populacionais e na economia. Os fundamentos da teoria dos valores extremos foram inicialmente propostos por Fisher e Tippett (1928). Estes autores definiram os três tipos possíveis de distribuições assintóticas de valores extremos, conhecidas como tipo I ou de Gumbel, tipo II ou de Fréchet e tipo III ou de Weibull. Essas distribuições são frequentemente utilizadas para estimar probabilidades de ocorrência de eventos acima de valores pré-estabelecidos.

Gnedenko (1943) provou as condições necessárias e suficientes para a existência das distribuições assintóticas de valores extremos. Também determinou que as caudas dessas distribuições, ou seja, a parte que trata dos valores máximos e mínimos menos frequentes, podem ser modeladas por algum tipo de distribuição contínua. As caudas da distribuição de Gumbel correspondem às distribuições exponencial, gama, normal ou log-normal; as da distribuição de Fréchet seguem uma distribuição de Cauchy, Pareto ou t de Student e as da distribuição de Weibull seguem uma distribuição uniforme.

Sejam  $X_1, X_2, \dots$  variáveis aleatórias i.i.d. representando riscos ou perdas com uma função de distribuição acumulada desconhecida (CDF)  $F(x) = \mathbb{P}(X_j \leq x)$ , para todo  $j \in \mathbb{N}$ . Uma perda é tratada como um número positivo e os eventos extremos ocorrem quando as perdas tomam valores na cauda direita da distribuição  $F(\cdot)$ . Definimos  $M_n = \max(X_1, \dots, X_n)$  como a maior perda de uma amostra de tamanho  $n$ . Uma importante parte da teoria dos valores extremos é focada na função de distribuição de  $M_n$ . Com a hipótese de que as variáveis aleatórias são independentes e identicamente distribuídas, a função de distribuição acumulada de  $M_n$  é dada por

$$\mathbb{P}(M_n \leq x) = \mathbb{P}(X_1 \leq x, \dots, X_n \leq x) = \prod_{j=1}^n F(x) = F^n(x).$$

Como  $F^n(\cdot)$  é desconhecida e a função de distribuição empírica não é sempre um bom estimador para  $F^n(\cdot)$ , uma aproximação assintótica para  $F^n(\cdot)$ , baseada no Teorema de *Fisher-Tippett* (ver Fisher e Tippett, 1928), é usada para fazer inferência sobre  $M_n$ . Além disso, como  $F^n(x) \rightarrow 0$  ou  $1$ , quando  $n \rightarrow +\infty$  e  $x$  é fixo, a aproximação assintótica é baseada na variável máximo normalizada, isto é,

$$Z_n = \frac{M_n - \mu_n}{\sigma_n}, \quad (5.1)$$

onde  $\sigma_n$  e  $\mu_n$  são seqüências de números reais tais que  $\sigma_n > 0$  é interpretada como uma medida da escala e  $\mu_n$  como uma medida da posição. O Teorema de Fisher e Tippett (1928) diz que se o máximo normalizado dado em (5.1) converge para uma função de distribuição; a distribuição resultante deve ser uma *distribuição generalizada de valores extremos (GVE)*, dada na expressão (5.2), que inclui como casos particulares, as distribuições de Gumbel, Fréchet e Weibull.

O Teorema 5.1 a seguir indica que o máximo normalizado  $(M_n - \mu_n)/\sigma_n$  converge em distribuição para uma variável aleatória com distribuição nas famílias *Gumbel (tipo I)*, *Fréchet (tipo II)* e *Weibull (tipo III)*. Estas três classes de distribuição são denominadas coletivamente de *distribuições de valores extremos*. Cada família tem um parâmetro de locação  $b$  e escala  $a$ . O Teorema 5.1 diz que, quando  $M_n$  pode ser normalizado com seqüências apropriadas, a variável normalizada correspondente tem uma distribuição que deve ser um dos três tipos das distribuições de valores extremos.

**Teorema 5.1.** *Se existem seqüências de constantes  $(\sigma_n)_{n \in \mathbb{N}}$  e  $(\mu_n)_{n \in \mathbb{N}}$ , com  $\sigma_n > 0$ , para todo  $n \in \mathbb{N}$ , tais que*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \frac{M_n - \mu_n}{\sigma_n} \leq z \right) = G(z),$$

onde  $G$  é uma função de distribuição não degenerada, então  $G(\cdot)$  pertence a uma das seguintes famílias:

$$\begin{aligned} \text{I: } G(z) &= \exp \left\{ - \exp \left[ - \left( \frac{z - \mu}{\sigma} \right) \right] \right\}, \quad z \in \mathbb{R}; \\ \text{II: } G(z) &= \begin{cases} 0, & \text{se } z \leq \mu, \\ \exp \left\{ - \left( \frac{z - \mu}{\sigma} \right)^{-\alpha} \right\}, & \text{se } z > \mu; \end{cases} \\ \text{III: } G(z) &= \begin{cases} \exp \left\{ - \left[ - \left( \frac{z - \mu}{\sigma} \right) \right]^\alpha \right\}, & \text{se } z < \mu, \\ 1, & \text{se } z \geq \mu; \end{cases} \end{aligned}$$

para parâmetros  $\sigma, \mu \in \mathbb{R}$ , com  $\sigma > 0$ , e no caso das famílias II e III,  $\alpha > 0$ .

Uma análise mais completa é oferecida pela reformulação dos modelos no Teorema 5.1. As famílias Gumbel, Fréchet e Weibull podem ser combinadas em uma família de modelos com distribuição dada por

$$G(z) = \exp \left\{ - \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (5.2)$$

definida no conjunto  $A = \{z \in \mathbb{R} : 1 + \xi(z - \mu)/\sigma > 0\}$ , onde os parâmetros satisfazem  $\mu \in \mathbb{R}$ ,  $\sigma \in \mathbb{R}$ , com  $\sigma > 0$ , e  $\xi \in \mathbb{R}$ . Esta é a família chamada de *distribuição generalizada de valores extremos* (GVE). Essa classe de distribuições tem três parâmetros: parâmetro de locação  $\mu$ ; parâmetro de escala  $\sigma$  e o parâmetro de forma  $\xi$ .

**Observação 5.1.** O tipo II e tipo III das distribuições dos valores extremos correspondem aos casos em que  $\xi > 0$  e  $\xi < 0$ , respectivamente, nesta parametrização. O subconjunto da família GVE com  $\xi = 0$  é interpretado como o limite de (5.2) quando  $\xi \rightarrow 0$  conduzindo à família Gumbel com função de distribuição dada por

$$G(z) = \exp \left\{ - \exp \left[ - \left( \frac{z - \mu}{\sigma} \right) \right] \right\}, \quad z \in \mathbb{R}.$$

A unificação das três famílias das distribuições de valores extremos em uma única classe simplifica implementações estatísticas. Através da inferência sobre  $\xi$ , os dados determinam o tipo de comportamento da cauda e não há necessidade de fazer julgamento prévio sobre qual família de valores extremos deva ser adotada. Por conveniência, temos o Teorema 5.2 em uma versão modificada do Teorema 5.1.

**Teorema 5.2.** *Se existem sequências de constantes  $(\sigma_n)_{n \in \mathbb{N}}$  e  $(\mu_n)_{n \in \mathbb{N}}$ , com  $\sigma_n > 0$ , para todo  $n \in \mathbb{N}$ , tais que*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \frac{M_n - \mu_n}{\sigma_n} \leq z \right) = G(z),$$

onde  $G$  é uma função de distribuição não degenerada, então  $G(z)$  é membro de uma família GVE dada por

$$G(z) = \exp \left\{ - \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad z \in A,$$

onde  $A = \{z \in \mathbb{R} : 1 + \xi(z - \mu)/\sigma > 0\}$ , com  $\mu \in \mathbb{R}$ ,  $\sigma \in \mathbb{R}$ , com  $\sigma > 0$ , e  $\xi \in \mathbb{R}$ .

A seguir, apresentamos três exemplos que mostram aplicação do Teorema 5.2.

**Exemplo 5.1.** Sejam  $X_1, X_2, \dots$  variáveis aleatórias i.i.d. com distribuição  $\mathcal{E}(1)$  cuja função de distribuição é dada por  $F(x) = 1 - e^{-x}$ , para todo  $x \in \mathbb{R}$ , com  $x > 0$ . Neste caso, considerando  $\sigma_n = 1$  e  $\mu_n = \ln(n)$ , para todo  $n \in \mathbb{N}$ , temos que

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n - \mu_n}{\sigma_n} \leq z\right) &= \lim_{n \rightarrow \infty} F^n(z + \ln(n)) \\ &= \lim_{n \rightarrow \infty} [1 - e^{-(z + \ln(n))}]^n \\ &= \lim_{n \rightarrow \infty} \left(1 - \frac{e^{-z}}{n}\right)^n \\ &= \exp(-e^{-z}), \end{aligned} \tag{5.3}$$

para todo  $z \in \mathbb{R}$ . Portanto, a distribuição limite é do tipo Gumbel, correspondente a  $\xi = 0$  na família GVE.

**Exemplo 5.2.** Sejam  $X_1, X_2, \dots$  variáveis aleatórias i.i.d. com distribuição padrão de Fréchet cuja função de distribuição é dada por  $F(x) = \exp(-1/x)$ , para todo  $x \in \mathbb{R}$ , com  $x > 0$ . Considere  $\sigma_n = n$  e  $\mu_n = 0$ , para todo  $n \in \mathbb{N}$ . Então,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n - \mu_n}{\sigma_n} \leq z\right) &= \lim_{n \rightarrow \infty} F^n(nz) \\ &= \lim_{n \rightarrow \infty} (\exp\{-1/(nz)\})^n \\ &= \exp(-1/z), \end{aligned} \tag{5.4}$$

para todo  $z \in \mathbb{R}$ , com  $z > 0$ . Portanto, neste caso o limite é também a distribuição padrão de Fréchet, com  $\xi = 1$  na família GVE.

**Exemplo 5.3.** Sejam  $X_1, X_2, \dots$  variáveis aleatórias i.i.d. com distribuição uniforme  $U(0, 1)$ , cuja função de distribuição é dada por  $F(x) = x$ , para todo  $x \in \mathbb{R}$ , com  $0 \leq x \leq 1$ . Para todo  $z \in \mathbb{R}$ , com  $z > 0$ , suponha que  $n > -z$  e considere  $\sigma_n = 1/n$  e  $\mu_n = 1$ , para todo  $n \in \mathbb{N}$ . Então,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n - \mu_n}{\sigma_n} \leq z\right) &= \lim_{n \rightarrow \infty} F^n(n^{-1}z + 1) \\ &= \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n \\ &= \exp(z). \end{aligned} \tag{5.5}$$

Portanto, a distribuição limite é do tipo Weibull, com  $\xi = -1$  na família GVE.

Uma medida natural dos eventos extremos são os valores de  $X_j$  que excedem um ponto inicial elevado  $u$ . Definimos a *distribuição de excessos* acima do ponto inicial  $u$  como a probabilidade condicional

$$F_u(y) = \mathbb{P}(X \leq u + y | X > u) = \frac{F(y + u) - F(u)}{1 - F(u)}, \quad y > 0. \quad (5.6)$$

**Definição 5.1 (Distribuição Generalizada de Pareto).** A *distribuição generalizada de Pareto* (GPD),  $G_\xi : A \rightarrow [0, 1]$ , é definida por

$$G_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-\frac{1}{\xi}}, & \xi \neq 0, \\ 1 - e^{-x}, & \xi = 0, \end{cases}$$

onde  $A = \{x \in \mathbb{R} : x \geq 0\}$  se  $\xi \geq 0$ , ou  $A = \{x \in \mathbb{R} : 0 \leq x \leq -\frac{1}{\xi}\}$  se  $\xi < 0$ .

O Teorema 5.3 a seguir mostra que, para um alto valor de  $u$ , a distribuição de excessos (5.6) é bem aproximada pela *distribuição generalizada de Pareto* (GPD).

**Teorema 5.3.** *Seja  $(X_1, X_2, \dots, X_n)$  uma sequência de variáveis aleatórias independentes com função de distribuição comum  $F$ , e seja*

$$M_n = \max\{X_1, \dots, X_n\}.$$

*Denotamos por  $X$  um termo arbitrário na sequência  $(X_j)_{j=1}^n$  e suponhamos que  $F$  satisfaz o Teorema 5.2. Então,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq z) \approx G(z),$$

onde

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}$$

para algum  $\mu, \sigma > 0$  e  $\xi$ . Então, para  $u$  suficientemente grande, a função de distribuição de  $X - u$ , condicionada em  $X > u$ , é aproximadamente

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{-\frac{1}{\xi}}, \quad \text{para todo } y \in B, \quad (5.7)$$

onde  $B = \{y \in \mathbb{R} : y > 0 \text{ e } (1 + \xi y/\tilde{\sigma}) > 0\}$ , com

$$\tilde{\sigma} = \sigma + \xi(u - \mu). \quad (5.8)$$

**Demonstração:** Seja  $X$  uma variável aleatória com função de distribuição  $F$ . Pelas suposições do Teorema 5.1, para  $n$  suficientemente grande,

$$F^n(z) \approx \exp\left\{-\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}$$



para algum parâmetro  $\mu, \sigma > 0$  e  $\xi$ . Segue-se que

$$n \ln(F(z)) \approx - \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}}. \quad (5.9)$$

Mas para valores grandes de  $z$ , uma expansão em série de Taylor implica

$$\ln(F(z)) \approx -(1 - F(z)). \quad (5.10)$$

Substituindo (5.10) em (5.9),

$$1 - F(u) \approx \frac{1}{n} \left[ 1 + \xi \left( \frac{u - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}}$$

para  $u$  suficientemente grande. Similarmente, para  $y > 0$ ,

$$1 - F(u + y) \approx \frac{1}{n} \left[ 1 + \xi \left( \frac{u + y - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}}.$$

Logo,

$$\begin{aligned} \mathbb{P}(X > u + y | X > u) &\approx \frac{n^{-1} \left[ \frac{1 + \xi(u + y - \mu)}{\sigma} \right]^{-\frac{1}{\xi}}}{n^{-1} \left[ \frac{1 + \xi(u - \mu)}{\sigma} \right]^{-\frac{1}{\xi}}} \\ &= \left[ 1 + \frac{\xi(u + y - \mu)}{\frac{1 + \xi(u - \mu)}{\sigma}} \right]^{-\frac{1}{\xi}} \\ &= \left[ 1 + \frac{\xi y}{\tilde{\sigma}} \right]^{-\frac{1}{\xi}}, \end{aligned} \quad (5.11)$$

Portanto,

$$\mathbb{P}(X \leq u + y | X > u) \approx 1 - \left[ 1 + \frac{\xi y}{\tilde{\sigma}} \right]^{-\frac{1}{\xi}},$$

onde  $\tilde{\sigma} = \sigma + \xi(u - \mu)$ . □

A seguir apresentamos alguns exemplos em termos da *distribuição de excessos*.

**Exemplo 5.4.** Para uma variável aleatória com distribuição exponencial, sabemos que  $F(x) = 1 - e^{-x}$ , para todo  $x > 0$ . Pelo cálculo direto, a expressão (5.6) nos fornece

$$\frac{1 - F(u + y)}{1 - F(u)} = \frac{e^{-(u+y)}}{e^{-u}} = e^{-y}, \quad \text{para todo } y > 0.$$

Conseqüentemente, o limite da distribuição de excessos é a distribuição exponencial, correspondente a  $\xi = 0$  e  $\tilde{\sigma} = 1$  na família generalizada de Pareto. Este é um resultado exato para todos os pontos  $u > 0$ .

**Exemplo 5.5.** Para uma variável aleatória com distribuição de Fréchet, sabemos que  $F(x) = \exp(-1/x)$ , para todo  $x > 0$ . Note que

$$\begin{aligned} \lim_{u \rightarrow \infty} \frac{\left( \frac{1 - e^{-\frac{1}{u+y}}}{1 - e^{-\frac{1}{u}}} \right)}{\left(1 + \frac{y}{u}\right)^{-1}} &= \lim_{u \rightarrow \infty} \left[ \left( \frac{1 - e^{-\frac{1}{u+y}}}{1 - e^{-\frac{1}{u}}} \right) \cdot \left(1 + \frac{y}{u}\right) \right] \\ &= \lim_{u \rightarrow \infty} \left( \frac{1 - e^{-\frac{1}{u+y}}}{1 - e^{-\frac{1}{u}}} \right) \cdot \lim_{u \rightarrow \infty} \left(1 + \frac{y}{u}\right) = 1. \end{aligned}$$

Segue-se então que a expressão (5.6) nos fornece

$$\frac{1 - F(u + y)}{1 - F(u)} = \frac{1 - e^{-\frac{1}{u+y}}}{1 - e^{-\frac{1}{u}}} \sim \left(1 + \frac{y}{u}\right)^{-1}$$

quando  $u \rightarrow \infty$ , para todo  $y > 0$ . Neste caso, o limite da distribuição de excessos corresponde a distribuição de Pareto com  $\xi = 1$  e  $\tilde{\sigma} = u$ .

**Exemplo 5.6.** Para uma variável aleatória com distribuição uniforme  $U(0, 1)$ , sabemos que  $F(x) = x$ , para  $0 \leq x \leq 1$ . Segue-se que a expressão (5.6) fornece

$$\frac{1 - F(u + y)}{1 - F(u)} = \frac{1 - (u + y)}{1 - u} = 1 - \frac{y}{1 - u},$$

para  $0 \leq y \leq 1 - u$ . Portanto, o limite da distribuição de excessos corresponde ao caso da distribuição generalizada de Pareto com  $\xi = -1$  e  $\tilde{\sigma} = 1 - u$ .

Para mais detalhes ver Embrechts et al. (1997) e Coles (2001).

## 5.1 Distribuições Estáveis

Distribuições estáveis são uma classe rica de distribuições de probabilidade que permitem caudas pesadas e têm fascinantes propriedades matemáticas. A classe foi caracterizada por Paul Lévy em seu estudo das somas de variáveis aleatórias independentes e identicamente distribuídas em 1920. A inexistência de fórmulas fechadas para as funções densidade e de distribuição com exceção de três casos (Gaussiana, Cauchy e Lévy), foi a maior desvantagem no uso de distribuições estáveis. No entanto, hoje existem programas computacionais confiáveis para calcular as funções densidade, de distribuição e os quantis estáveis. Com estes programas, é possível utilizar modelos estáveis em uma variedade de problemas aplicados. Apresentamos, nesta seção, fatos básicos das distribuições estáveis no caso univariado.

Uma importante propriedade de variáveis aleatórias com distribuição normal ou Gaussiana é que a soma de duas variáveis é uma variável aleatória

com distribuição normal. Uma consequência é que se  $X$  é normal, então para  $X_1$  e  $X_2$ , cópias independentes de  $X$ , com  $a$  e  $b$  constantes positivas,

$$aX_1 + bX_2 \stackrel{\mathcal{D}}{=} cX + d \quad (5.12)$$

para uma constante positiva  $c$  e algum  $d \in \mathbb{R}$ . O símbolo  $\stackrel{\mathcal{D}}{=}$  significa igualdade em distribuição. Em palavras, a equação (5.12) diz que a forma de  $X$  é preservada sob a adição. Esta seção trata da classe de distribuições com esta propriedade.

**Definição 5.2 (Variável Aleatória Estável).** Uma variável aleatória  $X$  é *estável* se  $X_1$  e  $X_2$  são cópias independentes de  $X$  e se  $a$  e  $b$  são constantes positivas, então vale a equação (5.12) para  $c > 0$  e  $d \in \mathbb{R}$ . A variável aleatória  $X$  é *estritamente estável* se vale a equação (5.12) com  $d = 0$  para todas as escolhas de  $a$  e  $b$ . Uma variável aleatória é *estável simétrica* se é estável e simetricamente distribuída em torno de 0, isto é,  $X \stackrel{\mathcal{D}}{=} -X$ .

O teorema a seguir mostra a condição que caracteriza uma variável aleatória estável.

**Teorema 5.4.** *A variável aleatória  $X$  é estável se e somente se, para todo  $n > 1$ , existem  $\alpha \in (0, 2]$  e constante real  $d_n \in \mathbb{R}$  tais que*

$$X_1 + \cdots + X_n \stackrel{\mathcal{D}}{=} n^{1/\alpha} X + d_n,$$

onde  $X_1, \dots, X_n$  são variáveis aleatórias independentes e são cópias idênticas de  $X$ . A v.a.  $X$  é estritamente estável se e somente se,  $d_n = 0$ , para todo  $n \in \mathbb{N}$ .

**Demonstração:** Ver Nolan (2009), página 7. □

A seguir apresentamos uma definição equivalente de variável aleatória estável. A função sinal será usada na Definição 5.3 e é definida por

$$\text{sign}(u) = \begin{cases} -1, & u < 0, \\ 0, & u = 0 \\ 1, & u > 0. \end{cases} \quad (5.13)$$

**Definição 5.3.** Uma variável aleatória  $X$  é *estável* se e somente se  $X \stackrel{\mathcal{D}}{=} aZ + b$ , onde  $0 < \alpha \leq 2$ ,  $-1 \leq \beta \leq 1$ ,  $a > 0$ ,  $b \in \mathbb{R}$  e  $Z$  é uma variável aleatória com função característica

$$\varphi_Z(t) = \mathbb{E}(e^{itZ}) = \begin{cases} \exp(-|t|^\alpha [1 - i\beta \tan(\frac{\pi\alpha}{2})(\text{sign}(t))]), & \alpha \neq 1, \\ \exp(-|t| [1 + i\beta \frac{2}{\pi}(\text{sign}(t)) \ln |t|]), & \alpha = 1, \end{cases} \quad (5.14)$$

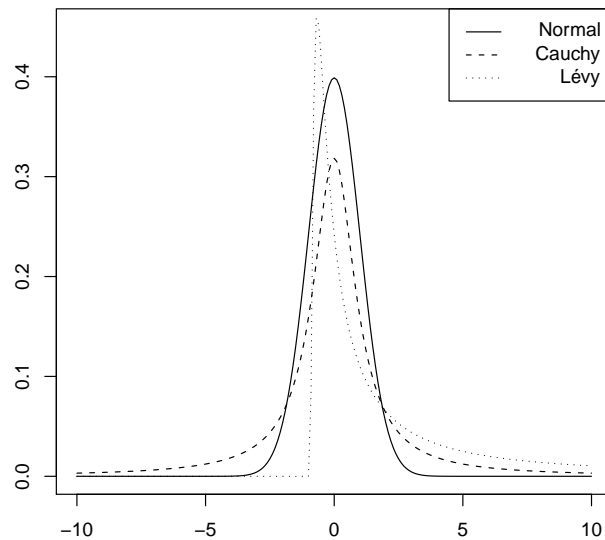
para todo  $t \in \mathbb{R}$ .

As distribuições estáveis são simétricas em torno de zero quando  $\beta = 0$  e  $b = 0$ . Neste caso, a função característica de  $aZ$  (ver Definição 5.3) tem a forma simples dada por

$$\varphi_X(t) = e^{-a^\alpha |t|^\alpha}, \quad \text{para todo } t \in \mathbb{R}.$$

**Observação 5.2.** A Definição 5.3 mostra que, em geral, distribuições estáveis exigem quatro parâmetros para serem descritas: índice de estabilidade  $\alpha \in (0, 2]$ , parâmetro de inclinação  $\beta \in [-1, 1]$ , parâmetro de escala  $\gamma > 0$  e parâmetro de posição  $\delta \in \mathbb{R}$ . Na literatura é usual denotar a classe das distribuições estáveis por  $S_\alpha(\beta, \gamma, \delta)$ .

Existem três casos onde é possível dar a expressão fechada para a função densidade (ver Figura 5.1) e verificar diretamente que a variável aleatória em questão é estável.



**Figura 5.1:** Funções Densidades das Distribuições  $\mathcal{N}(0, 1)$ ,  $\mathcal{C}(1, 0)$  e Lévy(1, 0).

**Exemplo 5.7 (Distribuição Normal ou Gaussiana).** A variável aleatória  $X$  tem distribuição  $\mathcal{N}(\mu, \sigma^2)$  se tem função densidade dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.$$

A função de distribuição acumulada é dada por  $F_X(x) = \mathbb{P}(X \leq x) = \Phi_Z\left(\frac{x-\mu}{\sigma}\right)$ , onde  $\Phi_Z(\cdot)$  é a função de distribuição acumulada de uma variável aleatória normal padronizada. A distribuição  $\mathcal{N}(\mu, \sigma^2)$  é estável com  $\alpha = 2, \beta = 0, a = \sigma/\sqrt{2}, b = \mu$ , ou seja, a variável aleatória  $X \stackrel{D}{=} S_2(0, \sigma/\sqrt{2}, \mu)$ .

**Exemplo 5.8 (Distribuição de Cauchy).** A variável aleatória  $X$  tem distribuição  $\mathcal{C}(\gamma, \delta)$  se tem função densidade dada por

$$f_X(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - \delta)^2}, \quad x \in \mathbb{R}.$$

A distribuição  $\mathcal{C}(\gamma, \delta)$  é estável com  $\alpha = 1, \beta = 0, a = \gamma, b = \delta$ , ou seja, a variável aleatória  $X \stackrel{D}{=} S_1(0, \gamma, \delta)$ .

**Exemplo 5.9 (Distribuição de Lévy).** Considere a variável aleatória  $X \sim \text{Lévy}(\gamma, \delta)$  cuja função densidade é dada por

$$f_X(x) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(x - \delta)^{3/2}} \exp\left\{-\frac{\gamma}{2(x - \delta)}\right\}, \quad \delta < x < \infty.$$

A distribuição de Lévy é estável com  $\alpha = 1/2, \beta = 1, a = \gamma, b = \delta$ , ou seja, a variável aleatória  $X \stackrel{D}{=} S_{\frac{1}{2}}(1, \gamma, \delta)$ .

Existem diferentes parametrizações para as distribuições estáveis. A (0)-parametrização é a mais usual em aplicações, que é uma variação da (M)-parametrização de Zolotarev (1986). Dizemos  $X \sim S_\alpha(\beta, \gamma, \delta; 0)$  se a função característica de  $X$  é dada por  $\varphi_X(t) = \mathbb{E}(e^{itZ}) =$

$$\begin{cases} \exp\left(-\gamma^\alpha |t|^\alpha \left[1 + i\beta \tan\left(\frac{\pi\alpha}{2}\right) (\text{sign}(t)) ((\gamma(|t|)^{1-\alpha} - 1)\right] + i\delta t\right), & \alpha \neq 1, \\ \exp\left(-\gamma |t| \left[1 + i\beta \frac{2}{\pi} (\text{sign}(t)) (\ln(|t|) + \ln(\gamma))\right] + i\delta t\right), & \alpha = 1, \end{cases} \quad (5.15)$$

para todo  $t \in \mathbb{R}$ , onde  $\text{sign}(\cdot)$  é dada pela expressão (5.13).

## 5.2 Estimadores para o Parâmetro $\alpha$

O parâmetro mais importante para ser estimado em dados com caudas pesadas é a sua taxa de decaimento  $\alpha$  que determina a probabilidade da ocorrência dos valores extremos da distribuição subjacente. Isto é,  $\alpha$  é o parâmetro tal que

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(X > x)}{C x^{-\alpha}} = 1,$$

onde  $C$  é uma constante positiva. Diversos métodos vem sendo propostos para a estimação do parâmetro  $\alpha$ . Geralmente, estão divididos em quatro

categorias de métodos: baseados no estimador de Hill, baseados na máxima verossimilhança, baseados em quantis e baseados na função característica.

Nesta seção, apresentamos alguns estimadores para o parâmetro  $\alpha$ : o estimador de Hill (ver Hill, 1975); uma versão suavizada do estimador de Hill, proposta por Resnick e Stărică (1997); o estimador de máxima verossimilhança (ver Nolan, 2001); o estimador baseado na função característica empírica, proposto por Press (1972) e o estimador de regressão baseado na função característica empírica (ver Koutrouvelis, 1981). Por fim, propomos um método de estimação para o parâmetro  $\alpha$ , baseado na função característica empírica, utilizando o procedimento de *encolhimento de wavelets* (ver Capítulo 4).

No método proposto por Hill (1975), não é necessário assumir uma forma paramétrica para a função de distribuição. Este método focaliza apenas no comportamento da cauda. Entretanto, na prática, o estimador de Hill é difícil de ser utilizado, por duas principais razões: vício alto e sensibilidade na escolha do número de estatísticas de ordens utilizadas na estimação. O estimador de Hill geralmente não funciona para distribuições estáveis quando  $\alpha > 1.5$  (ver Resnick e Stărică, 1997 e McCulloch, 1997). Para descrever o estimador de Hill, consideramos a definição a seguir.

**Definição 5.4 (Estimador de Hill).** Seja  $(X_1, \dots, X_n)'$  uma amostra aleatória. Considere as estatísticas de ordem como sendo

$$0 < X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}.$$

Para um número positivo  $k$ , o *estimador de Hill* para  $\alpha$  é definido por

$$\hat{\alpha}_{k,n}^{(H)} = \left( k^{-1} \sum_{j=1}^k (\ln(X_{(j)}) - \ln(X_{(k+1)})) \right)^{-1}. \quad (5.16)$$

O estimador de Hill para  $\alpha$  depende do inteiro  $k$ . Pode ser mostrado (ver Embrechts et al., 1997) que se  $F(\cdot)$  está no domínio da distribuição generalizada de Pareto, então  $\hat{\alpha}_{k,n}^{(H)}$  é assintoticamente normal com variância assintótica dada por

$$avar(\hat{\alpha}_{k,n}^{(H)}) = k^{-1} \alpha^2.$$

**Definição 5.5 (Variação Regular).** Uma função positiva  $L(\cdot)$  mensurável à Lebesgue sobre  $(0, \infty)$ , é de *variação lenta no  $\infty$*  se

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1, \quad \text{para todo } t > 0. \quad (5.17)$$

Denotamos  $L \in \mathcal{R}_0$ . Uma função positiva  $h(\cdot)$  mensurável à Lebesgue sobre  $(0, \infty)$ , é de *variação lenta no  $\infty$  com índice  $\alpha \in \mathbb{R}$*  se

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = t^\alpha, \quad \text{para todo } t > 0. \quad (5.18)$$

Denotamos  $h \in \mathcal{R}_\alpha$ .

O Teorema 5.5, a seguir, descreve as principais propriedades assintóticas do estimador de Hill.

**Teorema 5.5.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estritamente estacionário com função de distribuição marginal  $F_X(\cdot)$  onde, para algum  $\alpha > 0$ , e  $L \in \mathcal{R}_0$  tem-se que*

$$1 - F_X(x) = \mathbb{P}(X > x) = x^{-\alpha} L(x), \quad \text{para todo } x > 0. \quad (5.19)$$

Seja  $\hat{\alpha}^{(H)} = \hat{\alpha}_{k,n}^{(H)}$  o estimador de Hill dado na expressão (5.16).

(a) *(Consistência Fraca) Se o processo  $\{X_t\}_{t \in \mathbb{Z}}$ , constituído por variáveis aleatórias i.i.d., é fracamente dependente e linear e se  $k \rightarrow \infty$  e  $\lim_{n \rightarrow \infty} \frac{k}{n} = 0$ , então*

$$\hat{\alpha}^{(H)} \xrightarrow{P} \alpha.$$

(b) *(Consistência Forte) Se o processo  $\{X_t\}_{t \in \mathbb{Z}}$  é constituído por variáveis aleatórias i.i.d. e se  $\lim_{n \rightarrow \infty} \frac{k}{n} = 0$  e  $\lim_{n \rightarrow \infty} \frac{k}{\ln(\ln(n))} = \infty$ , então*

$$\hat{\alpha}^{(H)} \xrightarrow{q.c.} \alpha. \quad (5.20)$$

(c) *(Normalidade Assintótica) Se o processo  $\{X_t\}_{t \in \mathbb{Z}}$  é constituído por variáveis aleatórias i.i.d. e se as condições dadas nos itens (a) e (b) sobre o  $k$  estão satisfeitas, então*

$$\sqrt{k}(\hat{\alpha}^{(H)} - \alpha) \xrightarrow{D} \mathcal{N}(0, \alpha^2). \quad (5.21)$$

**Demonstração:** Ver Embrechts et al. (1997), página 336. □

Para descrever a versão suavizada do estimador de Hill, proposta por Resnick e Stărică (1997), considere a seguinte definição.

**Definição 5.6 (Estimador de Hill Suavizado).** Seja  $(X_1, \dots, X_n)'$  uma amostra aleatória. Considere as estatísticas de ordem como sendo

$$0 < X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}.$$

Para um número positivo  $k$ , o *estimador de Hill suavizado* para  $\alpha$  é definido por

$$\hat{\alpha}_{k,u}^{(HS)} = \frac{1}{(u-1)k} \sum_{p=k+1}^{uk} \hat{\alpha}_{p,n}^{(H)}, \quad (5.22)$$

onde  $u > 1$  e  $\hat{\alpha}_{p,n}^{(H)}$  é o estimador de Hill dado na Definição 5.4.

Esta versão do estimador de Hill, tem propriedades assintóticas satisfatórias, mas as estimativas variam consideravelmente com a escolha de  $k$  e de  $u$ .

Nolan (2001) descreve o método de máxima verossimilhança, para estimar todos os parâmetros de uma distribuição estável, e mostra que o estimador de máxima verossimilhança para o parâmetro  $\alpha$  é altamente eficiente quando  $\alpha$  está próximo de 2, em contraste com o estimador de Hill nesta mesma região.

**Definição 5.7 (Estimador de Máxima Verossimilhança).** Seja uma amostra aleatória  $(X_1, \dots, X_n)'$  e  $\Theta = \{(\alpha, \beta, \gamma, \delta) : \alpha \in (0, 2], \beta \in [-1, 1], \gamma \in (0, \infty), \delta \in \mathbb{R}\}$  o espaço dos parâmetros. Considerando a (0)-parametrização usual, o *estimador de máxima verossimilhança* para o vetor de parâmetros  $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \delta)'$   $\in \Theta$  é obtido maximizando a função

$$L(\boldsymbol{\theta}) = \sum_{i=1}^n \ln(f_{\mathbf{X}}(X_i; \boldsymbol{\theta})),$$

onde  $f_{\mathbf{X}}(\cdot; \boldsymbol{\theta})$  é a função densidade da distribuição estável.

**Observação 5.3.**

**a)** Em geral, não sabemos a fórmula fechada para a função densidade de distribuições estáveis, e temos que aproximá-la numericamente ou por integração numérica direta, lidando com o aumento do tempo computacional. O método de integração direta, proposto por Nolan (1997) e Nolan (2001), consiste de integração numérica das fórmulas de Zolotarev (1986), para as funções densidade e de distribuição.

**b)** Denotamos por  $\hat{\alpha}_{\text{mle}}$ , o *estimador de máxima verossimilhança* para o parâmetro  $\alpha$ . A variância assintótica para  $\hat{\alpha}_{\text{mle}}$  é dada por

$$\text{avar}(\hat{\alpha}_{\text{mle}}) = \frac{\sigma_{\hat{\alpha}_{\text{mle}}}^2}{n}, \quad (5.23)$$

onde  $\sigma_{\hat{\alpha}_{\text{mle}}}^2$  é a primeira entrada da matrix inversa  $4 \times 4$  da informação de Fisher (ver Definição A.2 no Anexo A).

Press (1972), baseou-se na função característica empírica para estimar o parâmetro  $\alpha$  (ver Definição 5.9). Observe que, pela expressão (5.15), obtemos

$$|\varphi_X(t)| = \exp(-\gamma|t|^\alpha), \text{ para todo } t \in \mathbb{R}.$$

Portanto,

$$\ln(|\varphi_X(t)|) = -\gamma|t|^\alpha, \text{ para todo } t \in \mathbb{R}. \quad (5.24)$$



Dados  $t_1, t_2 \in \mathbb{R}$ ,  $t_1 \neq t_2$ ,

$$\ln(|\varphi_X(t_1)|) = -\gamma|t_1|^\alpha \text{ e } \ln(|\varphi_X(t_2)|) = -\gamma|t_2|^\alpha. \quad (5.25)$$

Dividindo as duas equações, na expressão (5.25), obtemos

$$\left| \frac{t_1}{t_2} \right|^\alpha = \frac{\ln(|\varphi_X(t_1)|)}{\ln(|\varphi_X(t_2)|)}.$$

Portanto,

$$\alpha = \frac{\ln\left(\frac{\ln(|\varphi_X(t_1)|)}{\ln(|\varphi_X(t_2)|)}\right)}{\ln\left(\left|\frac{t_1}{t_2}\right|\right)}.$$

**Definição 5.8 (Estimador Baseado na FCE).** Seja  $(X_1, \dots, X_n)'$  uma amostra aleatória. A *função característica empírica* é dada por

$$\hat{\varphi}_X(t) = \frac{1}{n} \sum_{j=1}^n e^{itX_j}. \quad (5.26)$$

Dados  $t_1, t_2 \in \mathbb{R}$ ,  $t_1 \neq t_2$ , um estimador para  $\alpha$ , baseado na função característica empírica, proposto por Press (1972), denotado por  $\hat{\alpha}_{\text{fce}}$ , é definido por

$$\hat{\alpha}_{\text{fce}} = \frac{\ln\left(\frac{\ln(|\hat{\varphi}_X(t_1)|)}{\ln(|\hat{\varphi}_X(t_2)|)}\right)}{\ln\left(\left|\frac{t_1}{t_2}\right|\right)},$$

onde  $|\hat{\varphi}_X(\cdot)|$  é dado em (5.26).

**Observação 5.4.**

**a)** Note que  $\{\hat{\varphi}_X(t)\}_{t \in \mathbb{R}}$  é um processo estocástico e  $|\hat{\varphi}_X(t)| \leq 1$ , para todo  $t \in \mathbb{R}$ . Portanto, todos os momentos de  $\hat{\varphi}_X(\cdot)$  são finitos, e para todo  $t \in \mathbb{R}$  tem-se que  $\hat{\varphi}_X(t)$  é a média amostral de variáveis aleatórias independentes e identicamente distribuídas. Logo, pela lei dos grandes números,  $\hat{\varphi}_X(\cdot)$  é um estimador consistente para  $\varphi_X(\cdot)$ .

**b)** Ainda é questão de estudo a escolha ótima para  $t_1$  e  $t_2$ . Aqui consideramos  $t_1 = 0.2$  e  $t_2 = 0.8$ . Para detalhes sobre a escolha dos valores  $t_1$  e  $t_2$ , ver Besbeas (2008).

**c)** A variância assintótica de  $\hat{\alpha}_{\text{fce}}$  é dada por

$$\text{avar}(\hat{\alpha}_{\text{fce}}) = \frac{\sigma_{\hat{\alpha}_{\text{fce}}}^2}{n}, \quad (5.27)$$

onde

$$\begin{aligned} \sigma_{\hat{\alpha}_{\text{fce}}}^2 &= \frac{[1 + |\hat{\varphi}_X(2t_1)| - 2|\hat{\varphi}_X(t_1)|^2]}{2[|\hat{\varphi}_X(t_1)| \ln(|\hat{\varphi}_X(t_1)|) \ln\left(\left|\frac{t_1}{t_2}\right|\right)]^2} \\ &+ \frac{[1 + |\hat{\varphi}_X(2t_2)| - 2|\hat{\varphi}_X(t_2)|^2]}{2[|\hat{\varphi}_X(t_2)| \ln(|\hat{\varphi}_X(t_2)|) \ln\left(\left|\frac{t_1}{t_2}\right|\right)]^2} \\ &- \frac{[|\hat{\varphi}_X(t_1 + t_2)| + |\hat{\varphi}_X(t_1 - t_2)| - 2|\hat{\varphi}_X(t_1)||\hat{\varphi}_X(t_2)|]}{[|\hat{\varphi}_X(t_1)||\hat{\varphi}_X(t_2)| \ln(|\hat{\varphi}_X(t_1)|) \ln(|\hat{\varphi}_X(t_2)|) [\ln\left(\left|\frac{t_1}{t_2}\right|\right)]^2}. \end{aligned}$$

Observe que pela expressão (5.24) obtemos

$$\ln(-\ln(|\varphi_X(t)|)) = \ln(\gamma) + \alpha \ln(|t|), \quad t \in \mathbb{R}. \quad (5.28)$$

A equação (5.28) sugere estimar o parâmetro  $\alpha$ , utilizando regressão linear de  $y_j = \ln(-\ln(|\varphi_X(t)|))$  versus  $x_j = \ln(|t_j|)$ , no modelo dado na expressão (3.6), para um conjunto de pontos  $\{t_j\}_{j=1}^m$ , onde  $a = \ln(\gamma)$  e  $\epsilon_j$  denota um erro aleatório, para todo  $j \in \{1, \dots, m\}$ .

**Definição 5.9 (Estimador Utilizando Regressão Baseado na FCE).**

Seja  $(X_1, \dots, X_n)'$  uma amostra aleatória. Dada uma sequência de pontos  $\{t_j\}_{j=1}^m$ ,  $t_j \in \mathbb{R}$ , o estimador baseado na função característica para  $\alpha$ , utilizando regressão linear e denotado por  $\hat{\alpha}_{\text{reg}}$ , é definido por

$$\hat{\alpha}_{\text{reg}} = \frac{\sum_{j=1}^m (x_j - \bar{x})(\ln(-\ln(|\hat{\varphi}_X(t_j)|)))}{\sum_{j=1}^m (x_j - \bar{x})^2} \quad (5.29)$$

e a variância de  $\hat{\alpha}_{\text{reg}}$  é dada por

$$\text{Var}(\hat{\alpha}_{\text{reg}}) = \frac{\text{Var}(\ln(-\ln(|\varphi_X(t_j)|)))}{\sum_{j=1}^m (x_j - \bar{x})^2}, \quad (5.30)$$

onde  $x_j = \ln(|t_j|)$ , para todo  $j \in \{1, \dots, m\}$ ,  $\bar{x} = \frac{1}{m} \sum_{j=1}^m x_j$  e  $\hat{\varphi}_X(\cdot)$  é dado pela expressão (5.26).

Propomos, a seguir, um estimador para  $\alpha$ , baseado no estimador de regressão, utilizando o procedimento de *encolhimento de wavelets* (ver Seção 4.6).

**Definição 5.10 (Estimador Baseado em Wavelets).** Seja  $(X_1, \dots, X_n)'$  uma amostra aleatória. Dada uma sequência de pontos  $\{t_j\}_{j=1}^m$ ,  $t_j \in \mathbb{R}$ , seja  $\hat{\varphi}_X(\cdot)$  o sinal dado pela expressão (5.26). Para obtermos o estimador baseado na função característica para  $\alpha$ , utilizando o procedimento de *encolhimento de wavelets*, denotado por  $\hat{\alpha}_{\text{wave}}$ , é preciso seguir os seguintes passos:

1. Aplicar a *transformada discreta de wavelet*, com  $J$  níveis, para o sinal  $|\hat{\varphi}_X(\cdot)|$ , obtendo coeficientes contaminados por ruídos  $d_1, d_2, \dots, d_J, s_J$ .
2. Usar limiares para diminuir os coeficientes, tornando nulos aqueles coeficientes abaixo de um determinado valor, nas escalas  $j$  mais finas, para obtermos novos coeficientes  $\tilde{d}_1 \equiv \delta_{\lambda_1 \sigma_1}(d_1), \dots, \tilde{d}_j \equiv \delta_{\lambda_j \sigma_j}(d_j)$ . A *função encolhimento*  $\delta_c(x)$  encolhe  $x$  para zero e é parametrizada por  $c = \lambda \sigma$ , onde  $\lambda$  é o limiar (ver Definição 4.14) e  $\sigma$  é uma estimativa da escala do ruído.
3. Aplicar a *transformada inversa discreta de wavelets* usando os coeficientes de detalhes  $\tilde{d}_1, \dots, \tilde{d}_j, d_{j+1}, \dots, d_J, s_J$  para obter o estimador de encolhimento de wavelets para o sinal  $|\varphi_X(\cdot)|$ . Denotamos por  $|\hat{\varphi}_{\text{wave}}(\cdot)|$  o estimador de encolhimento de wavelets para  $|\varphi_X(\cdot)|$ .

Portanto, um estimador para  $\alpha$ , aplicando o procedimento de *encolhimento de wavelets*, é definido por

$$\hat{\alpha}_{\text{wave}} = \frac{\sum_{j=1}^m (x_j - \bar{x}) (\ln(-\ln(|\hat{\varphi}_{\text{wave}}(t_j)|)))}{\sum_{j=1}^m (x_j - \bar{x})^2} \quad (5.31)$$

e a variância de  $\hat{\alpha}_{\text{wave}}$  é dada por

$$\text{Var}(\hat{\alpha}_{\text{wave}}) = \frac{\text{Var}(\ln(-\ln(|\hat{\varphi}_{\text{wave}}(t_j)|)))}{\sum_{j=1}^m (x_j - \bar{x})^2}, \quad (5.32)$$

onde  $x_j = \ln(|t_j|)$ , para  $j \in \{1, \dots, m\}$ ,  $\bar{x} = \frac{1}{m} \sum_{j=1}^m x_j$ ,  $|\hat{\varphi}_{\text{wave}}(\cdot)|$  é o estimador de encolhimento de wavelets para  $|\varphi_X(\cdot)|$ .

Para detalhes sobre a escolha da sequência  $\{t_j\}_{j=1}^m$ , ver Besbeas (2008).

## 5.3 Simulações

Para analisar o estimador  $\hat{\alpha}_{\text{wave}}$ , proposto na Seção 5.2, geramos séries temporais  $\alpha$ -estáveis simétricas de média 0, com 1000 replicações e tamanho amostral  $n \in \{1000, 5000, 10000, 50000, 100000\}$ . Calculamos o valor médio, o vício, erro quadrático médio (eqm) e variância (var) e comparamos com os resultados obtidos pelos estimadores  $\hat{\alpha}_{\text{fce}}$  e  $\hat{\alpha}_{\text{reg}}$ . Nesta análise, não utilizamos o estimador  $\hat{\alpha}_{\text{mle}}$ , por exigir tempo computacional muito grande.

Para obter o estimador  $\hat{\alpha}_{\text{wave}}$ , consideramos a wavelet “s8”, o nível  $J = \lfloor \log_2(n) \rfloor$ , o limiar universal (ver Definição 4.14), a estimativa do nível do ruído  $\sigma = -1$  e a função de encolhimento dura (ver Observação 4.13). Para os estimadores  $\hat{\alpha}_{\text{reg}}$  e  $\hat{\alpha}_{\text{wave}}$ , consideramos a sequência de pontos  $\{t_j\}_{j=1}^{2000}$ , definida por  $t_j = 0.1 + (j - 1) \frac{0.9}{1999}$ , para todo  $j \in \{1, 2, \dots, 2000\}$ .

As Tabelas 5.1 a 5.5, apresentam os resultados das simulações. Note que os estimadores  $\hat{\alpha}_{\text{fce}}$ ,  $\hat{\alpha}_{\text{reg}}$  e  $\hat{\alpha}_{\text{wave}}$  fornecem boas estimativas e são similares, mas o estimador  $\hat{\alpha}_{\text{wave}}$  apresentou melhores resultados para a variância e o erro quadrático médio. Verificou-se menores valores da variância e do erro quadrático médio em quase todo o experimento.

Observa-se também nas Tabelas 5.1 a 5.5, que para todo  $\alpha \in \{0.1, \dots, 2\}$ , o tamanho amostral  $n$  não altera muito nos resultados dos estimadores propostos. Para todo  $n \in \{1000, 5000, 10000, 50000, 100000\}$ , os estimadores  $\hat{\alpha}_{\text{fce}}$ ,  $\hat{\alpha}_{\text{reg}}$  e  $\hat{\alpha}_{\text{wave}}$  apresentaram estimativas similares, com variância e erro quadrático médio sempre próximos de zero.

Para todo  $\alpha \in \{0.1, 0.2, \dots, 1.4\}$ , em quase toda a análise, o estimador  $\hat{\alpha}_{\text{wave}}$  apresentou vício, variância e erro quadrático médio inferiores aos resultados obtidos pelos estimadores  $\hat{\alpha}_{\text{fce}}$  e  $\hat{\alpha}_{\text{reg}}$ , equiparando-se a estes para  $\alpha$  próximo a 2.

**Tabela 5.1:** Resultado da Estimação de  $\alpha$ , baseada em  $\hat{\alpha}_{fce}$ ,  $\hat{\alpha}_{reg}$  e  $\hat{\alpha}_{wave}$ , para  $S_\alpha(0, 1, 0)$  com  $n = 1000$ .

$\alpha$	Estimador	$\hat{\alpha}$	vício	var	eqm
0.1	$\hat{\alpha}_{fce}$	0.0945	-0.0055	0.0022	0.0022
	$\hat{\alpha}_{reg}$	0.0945	-0.0055	0.0022	0.0022
	$\hat{\alpha}_{wave}$	0.0973	-0.0027	0.0003	0.0003
0.2	$\hat{\alpha}_{fce}$	0.1951	-0.0049	0.0023	0.0023
	$\hat{\alpha}_{reg}$	0.1950	-0.0050	0.0023	0.0023
	$\hat{\alpha}_{wave}$	0.1970	-0.0030	0.0006	0.0006
0.3	$\hat{\alpha}_{fce}$	0.3048	0.0048	0.0023	0.0024
	$\hat{\alpha}_{reg}$	0.3051	0.0051	0.0024	0.0024
	$\hat{\alpha}_{wave}$	0.3036	0.0036	0.0008	0.0008
0.4	$\hat{\alpha}_{fce}$	0.4004	0.0004	0.0024	0.0024
	$\hat{\alpha}_{reg}$	0.4000	0.0000	0.0024	0.0024
	$\hat{\alpha}_{wave}$	0.4014	0.0014	0.0011	0.0011
0.5	$\hat{\alpha}_{fce}$	0.5018	0.0018	0.0024	0.0023
	$\hat{\alpha}_{reg}$	0.5016	0.0016	0.0024	0.0023
	$\hat{\alpha}_{wave}$	0.5019	0.0019	0.0011	0.0011
0.6	$\hat{\alpha}_{fce}$	0.5976	-0.0024	0.0031	0.0031
	$\hat{\alpha}_{reg}$	0.5977	-0.0023	0.0031	0.0031
	$\hat{\alpha}_{wave}$	0.5997	-0.0003	0.0016	0.0016
0.7	$\hat{\alpha}_{fce}$	0.7038	0.0038	0.0031	0.0031
	$\hat{\alpha}_{reg}$	0.7036	0.0036	0.0031	0.0031
	$\hat{\alpha}_{wave}$	0.7038	0.0038	0.0017	0.0017
0.8	$\hat{\alpha}_{fce}$	0.8070	0.0070	0.0033	0.0033
	$\hat{\alpha}_{reg}$	0.8069	0.0069	0.0033	0.0033
	$\hat{\alpha}_{wave}$	0.8075	0.0075	0.0020	0.0020
0.9	$\hat{\alpha}_{fce}$	0.8919	-0.0081	0.0038	0.0039
	$\hat{\alpha}_{reg}$	0.8919	-0.0081	0.0038	0.0039
	$\hat{\alpha}_{wave}$	0.8922	-0.0078	0.0024	0.0024
1.0	$\hat{\alpha}_{fce}$	1.0026	0.0026	0.0032	0.0032
	$\hat{\alpha}_{reg}$	1.0026	0.0026	0.0032	0.0032
	$\hat{\alpha}_{wave}$	0.9998	-0.0002	0.0024	0.0024
1.1	$\hat{\alpha}_{fce}$	1.0988	-0.0012	0.0031	0.0031
	$\hat{\alpha}_{reg}$	1.0988	-0.0012	0.0031	0.0031
	$\hat{\alpha}_{wave}$	1.0974	-0.0026	0.0022	0.0022
1.2	$\hat{\alpha}_{fce}$	1.2017	0.0017	0.0034	0.0034
	$\hat{\alpha}_{reg}$	1.2019	0.0019	0.0034	0.0034
	$\hat{\alpha}_{wave}$	1.2025	0.0025	0.0030	0.0030
1.3	$\hat{\alpha}_{fce}$	1.3052	0.0052	0.0033	0.0033
	$\hat{\alpha}_{reg}$	1.3050	0.0050	0.0033	0.0033
	$\hat{\alpha}_{wave}$	1.3070	0.0070	0.0028	0.0029
1.4	$\hat{\alpha}_{fce}$	1.3981	-0.0019	0.0029	0.0029
	$\hat{\alpha}_{reg}$	1.3984	-0.0016	0.0029	0.0029
	$\hat{\alpha}_{wave}$	1.3977	-0.0023	0.0024	0.0024
1.5	$\hat{\alpha}_{fce}$	1.5038	0.0038	0.0029	0.0029
	$\hat{\alpha}_{reg}$	1.5036	0.0036	0.0029	0.0029
	$\hat{\alpha}_{wave}$	1.5041	0.0041	0.0027	0.0027
1.6	$\hat{\alpha}_{fce}$	1.5979	-0.0021	0.0039	0.0039
	$\hat{\alpha}_{reg}$	1.5985	-0.0015	0.0040	0.0039
	$\hat{\alpha}_{wave}$	1.5990	-0.0010	0.0035	0.0035
1.7	$\hat{\alpha}_{fce}$	1.7023	0.0023	0.0037	0.0037
	$\hat{\alpha}_{reg}$	1.7024	0.0024	0.0037	0.0037
	$\hat{\alpha}_{wave}$	1.7021	0.0021	0.0035	0.0035
1.8	$\hat{\alpha}_{fce}$	1.8022	0.0022	0.0023	0.0023
	$\hat{\alpha}_{reg}$	1.8024	0.0024	0.0023	0.0023
	$\hat{\alpha}_{wave}$	1.8027	0.0027	0.0021	0.0021
1.9	$\hat{\alpha}_{fce}$	1.9005	0.0005	0.0015	0.0015
	$\hat{\alpha}_{reg}$	1.9004	0.0004	0.0015	0.0015
	$\hat{\alpha}_{wave}$	1.9010	0.0010	0.0015	0.0015
2.0	$\hat{\alpha}_{fce}$	2.0014	0.0014	0.0001	0.0001
	$\hat{\alpha}_{reg}$	2.0014	0.0014	0.0001	0.0001
	$\hat{\alpha}_{wave}$	2.0014	0.0014	0.0001	0.0001

**Tabela 5.2:** Resultado da Estimação de  $\alpha$ , baseada em  $\hat{\alpha}_{fce}$ ,  $\hat{\alpha}_{reg}$  e  $\hat{\alpha}_{wave}$ , para  $S_\alpha(0, 1, 0)$  com  $n = 5000$ .

$\alpha$	Estimador	$\hat{\alpha}$	vício	var	eqm
0.1	$\hat{\alpha}_{fce}$	0.1004	0.0004	0.0024	0.0023
	$\hat{\alpha}_{reg}$	0.1004	0.0004	0.0023	0.0023
	$\hat{\alpha}_{wave}$	0.1000	0.0000	0.0003	0.0003
0.2	$\hat{\alpha}_{fce}$	0.2063	0.0063	0.0027	0.0027
	$\hat{\alpha}_{reg}$	0.2062	0.0062	0.0027	0.0027
	$\hat{\alpha}_{wave}$	0.2012	0.0012	0.0005	0.0005
0.3	$\hat{\alpha}_{fce}$	0.2971	-0.0029	0.0021	0.0021
	$\hat{\alpha}_{reg}$	0.2972	-0.0028	0.0021	0.0021
	$\hat{\alpha}_{wave}$	0.2991	-0.0009	0.0007	0.0007
0.4	$\hat{\alpha}_{fce}$	0.4032	0.0032	0.0023	0.0023
	$\hat{\alpha}_{reg}$	0.4032	0.0032	0.0023	0.0023
	$\hat{\alpha}_{wave}$	0.3995	-0.0005	0.0010	0.0010
0.5	$\hat{\alpha}_{fce}$	0.5020	0.0020	0.0024	0.0024
	$\hat{\alpha}_{reg}$	0.5020	0.0020	0.0024	0.0024
	$\hat{\alpha}_{wave}$	0.5022	0.0022	0.0011	0.0011
0.6	$\hat{\alpha}_{fce}$	0.6049	0.0049	0.0033	0.0033
	$\hat{\alpha}_{reg}$	0.6051	0.0051	0.0033	0.0033
	$\hat{\alpha}_{wave}$	0.6000	0.0000	0.0018	0.0018
0.7	$\hat{\alpha}_{fce}$	0.7025	0.0025	0.0029	0.0029
	$\hat{\alpha}_{reg}$	0.7027	0.0027	0.0029	0.0029
	$\hat{\alpha}_{wave}$	0.6997	-0.0003	0.0017	0.0017
0.8	$\hat{\alpha}_{fce}$	0.7949	-0.0051	0.0029	0.0029
	$\hat{\alpha}_{reg}$	0.7947	-0.0053	0.0029	0.0029
	$\hat{\alpha}_{wave}$	0.7944	-0.0056	0.0020	0.0020
0.9	$\hat{\alpha}_{fce}$	0.8963	-0.0037	0.0031	0.0031
	$\hat{\alpha}_{reg}$	0.8963	-0.0037	0.0031	0.0031
	$\hat{\alpha}_{wave}$	0.8981	-0.0019	0.0020	0.0020
1.0	$\hat{\alpha}_{fce}$	1.0031	0.0031	0.0028	0.0028
	$\hat{\alpha}_{reg}$	1.0029	0.0029	0.0028	0.0028
	$\hat{\alpha}_{wave}$	1.0017	0.0017	0.0021	0.0021
1.1	$\hat{\alpha}_{fce}$	1.1006	0.0006	0.0032	0.0032
	$\hat{\alpha}_{reg}$	1.1008	0.0008	0.0032	0.0032
	$\hat{\alpha}_{wave}$	1.1000	0.0000	0.0025	0.0025
1.2	$\hat{\alpha}_{fce}$	1.2058	0.0058	0.0038	0.0038
	$\hat{\alpha}_{reg}$	1.2057	0.0057	0.0037	0.0038
	$\hat{\alpha}_{wave}$	1.2055	0.0055	0.0027	0.0027
1.3	$\hat{\alpha}_{fce}$	1.3069	0.0069	0.0034	0.0035
	$\hat{\alpha}_{reg}$	1.3069	0.0069	0.0034	0.0034
	$\hat{\alpha}_{wave}$	1.3049	0.0049	0.0026	0.0026
1.4	$\hat{\alpha}_{fce}$	1.4002	0.0002	0.0031	0.0031
	$\hat{\alpha}_{reg}$	1.4002	0.0002	0.0031	0.0031
	$\hat{\alpha}_{wave}$	1.4018	0.0018	0.0027	0.0027
1.5	$\hat{\alpha}_{fce}$	1.4994	-0.0006	0.0033	0.0033
	$\hat{\alpha}_{reg}$	1.4996	-0.0004	0.0033	0.0033
	$\hat{\alpha}_{wave}$	1.498	-0.0020	0.0030	0.0030
1.6	$\hat{\alpha}_{fce}$	1.5995	-0.0005	0.0032	0.0032
	$\hat{\alpha}_{reg}$	1.5995	-0.0005	0.0032	0.0032
	$\hat{\alpha}_{wave}$	1.6000	0.0000	0.0029	0.0029
1.7	$\hat{\alpha}_{fce}$	1.7000	0.0000	0.0031	0.0031
	$\hat{\alpha}_{reg}$	1.7004	0.0004	0.0032	0.0031
	$\hat{\alpha}_{wave}$	1.7022	0.0022	0.0030	0.0030
1.8	$\hat{\alpha}_{fce}$	1.8053	0.0053	0.0021	0.0021
	$\hat{\alpha}_{reg}$	1.8054	0.0054	0.0021	0.0021
	$\hat{\alpha}_{wave}$	1.8066	0.0066	0.0020	0.0020
1.9	$\hat{\alpha}_{fce}$	1.9022	0.0022	0.0013	0.0013
	$\hat{\alpha}_{reg}$	1.9023	0.0023	0.0013	0.0013
	$\hat{\alpha}_{wave}$	1.9022	0.0022	0.0014	0.0014
2.0	$\hat{\alpha}_{fce}$	1.9999	-0.0001	0.0001	0.0001
	$\hat{\alpha}_{reg}$	2.0000	0.0000	0.0001	0.0001
	$\hat{\alpha}_{wave}$	1.9999	-0.0001	0.0001	0.0001

**Tabela 5.3:** Resultado da Estimação de  $\alpha$ , baseada em  $\hat{\alpha}_{fce}$ ,  $\hat{\alpha}_{reg}$  e  $\hat{\alpha}_{wave}$ , para para  $S_\alpha(0, 1, 0)$  com  $n = 10000$ .

$\alpha$	Estimador	$\hat{\alpha}$	vício	var	eqm
0.1	$\hat{\alpha}_{fce}$	0.1012	0.0012	0.0026	0.0026
	$\hat{\alpha}_{reg}$	0.1014	0.0014	0.0026	0.0025
	$\hat{\alpha}_{wave}$	0.0980	-0.0020	0.0003	0.0003
0.2	$\hat{\alpha}_{fce}$	0.2033	0.0033	0.0026	0.0026
	$\hat{\alpha}_{reg}$	0.2033	0.0033	0.0026	0.0026
	$\hat{\alpha}_{wave}$	0.2014	0.0014	0.0006	0.0006
0.3	$\hat{\alpha}_{fce}$	0.3011	0.0011	0.0023	0.0023
	$\hat{\alpha}_{reg}$	0.3012	0.0012	0.0023	0.0023
	$\hat{\alpha}_{wave}$	0.3005	0.0005	0.0008	0.0008
0.4	$\hat{\alpha}_{fce}$	0.3948	-0.0052	0.0024	0.0024
	$\hat{\alpha}_{reg}$	0.3948	-0.0052	0.0024	0.0024
	$\hat{\alpha}_{wave}$	0.3998	-0.0002	0.0008	0.0008
0.5	$\hat{\alpha}_{fce}$	0.4988	-0.0012	0.0028	0.0027
	$\hat{\alpha}_{reg}$	0.4989	-0.0011	0.0027	0.0027
	$\hat{\alpha}_{wave}$	0.5013	0.0013	0.0012	0.0012
0.6	$\hat{\alpha}_{fce}$	0.5940	-0.0060	0.0027	0.0027
	$\hat{\alpha}_{reg}$	0.5942	-0.0058	0.0027	0.0027
	$\hat{\alpha}_{wave}$	0.5953	-0.0047	0.0015	0.0015
0.7	$\hat{\alpha}_{fce}$	0.7035	0.0035	0.0028	0.0028
	$\hat{\alpha}_{reg}$	0.7036	0.0036	0.0028	0.0028
	$\hat{\alpha}_{wave}$	0.7040	0.0040	0.0018	0.0018
0.8	$\hat{\alpha}_{fce}$	0.8005	0.0005	0.0030	0.0030
	$\hat{\alpha}_{reg}$	0.8008	0.0008	0.0030	0.0030
	$\hat{\alpha}_{wave}$	0.8007	0.0007	0.0019	0.0019
0.9	$\hat{\alpha}_{fce}$	0.8986	-0.0014	0.0032	0.0032
	$\hat{\alpha}_{reg}$	0.8991	-0.0009	0.0032	0.0032
	$\hat{\alpha}_{wave}$	0.9005	0.0005	0.0021	0.0021
1.0	$\hat{\alpha}_{fce}$	0.9967	-0.0033	0.0033	0.0033
	$\hat{\alpha}_{reg}$	0.9967	-0.0033	0.0033	0.0032
	$\hat{\alpha}_{wave}$	0.9968	-0.0032	0.0024	0.0024
1.1	$\hat{\alpha}_{fce}$	1.0917	-0.0083	0.0034	0.0034
	$\hat{\alpha}_{reg}$	1.0917	-0.0083	0.0034	0.0034
	$\hat{\alpha}_{wave}$	1.0925	-0.0075	0.0024	0.0025
1.2	$\hat{\alpha}_{fce}$	1.2023	0.0023	0.0035	0.0035
	$\hat{\alpha}_{reg}$	1.2024	0.0024	0.0035	0.0035
	$\hat{\alpha}_{wave}$	1.2033	0.0033	0.0027	0.0027
1.3	$\hat{\alpha}_{fce}$	1.2969	-0.0031	0.0037	0.0037
	$\hat{\alpha}_{reg}$	1.2972	-0.0028	0.0037	0.0036
	$\hat{\alpha}_{wave}$	1.2968	-0.0032	0.0030	0.0030
1.4	$\hat{\alpha}_{fce}$	1.4043	0.0043	0.0036	0.0036
	$\hat{\alpha}_{reg}$	1.4041	0.0041	0.0036	0.0036
	$\hat{\alpha}_{wave}$	1.4035	0.0035	0.0029	0.0029
1.5	$\hat{\alpha}_{fce}$	1.5095	0.0095	0.0038	0.0039
	$\hat{\alpha}_{reg}$	1.5095	0.0095	0.0038	0.0039
	$\hat{\alpha}_{wave}$	1.5063	0.0063	0.0032	0.0032
1.6	$\hat{\alpha}_{fce}$	1.5981	-0.0019	0.0038	0.0038
	$\hat{\alpha}_{reg}$	1.5983	-0.0017	0.0038	0.0038
	$\hat{\alpha}_{wave}$	1.5985	-0.0015	0.0033	0.0033
1.7	$\hat{\alpha}_{fce}$	1.6923	-0.0077	0.0033	0.0033
	$\hat{\alpha}_{reg}$	1.6921	-0.0079	0.0033	0.0033
	$\hat{\alpha}_{wave}$	1.6952	-0.0048	0.0031	0.0031
1.8	$\hat{\alpha}_{fce}$	1.8027	0.0027	0.0027	0.0027
	$\hat{\alpha}_{reg}$	1.8026	0.0026	0.0026	0.0026
	$\hat{\alpha}_{wave}$	1.8032	0.0032	0.0027	0.0026
1.9	$\hat{\alpha}_{fce}$	1.9014	0.0014	0.0017	0.0017
	$\hat{\alpha}_{reg}$	1.9016	0.0016	0.0017	0.0017
	$\hat{\alpha}_{wave}$	1.9017	0.0017	0.0015	0.0015
2.0	$\hat{\alpha}_{fce}$	1.9998	-0.0002	0.0001	0.0001
	$\hat{\alpha}_{reg}$	1.9998	-0.0002	0.0001	0.0001
	$\hat{\alpha}_{wave}$	1.9998	-0.0002	0.0001	0.0001

**Tabela 5.4:** Resultado da Estimação de  $\alpha$ , baseada em  $\hat{\alpha}_{fce}$ ,  $\hat{\alpha}_{reg}$  e  $\hat{\alpha}_{wave}$ , para  $S_\alpha(0, 1, 0)$  com  $n = 50000$ .

$\alpha$	Estimador	$\hat{\alpha}$	vício	var	eqm
0.1	$\hat{\alpha}_{fce}$	0.0995	-0.0005	0.0024	0.0024
	$\hat{\alpha}_{reg}$	0.0995	-0.0005	0.0024	0.0023
	$\hat{\alpha}_{wave}$	0.1007	0.0007	0.0002	0.0002
0.2	$\hat{\alpha}_{fce}$	0.2037	0.0037	0.0030	0.0030
	$\hat{\alpha}_{reg}$	0.2039	0.0039	0.0030	0.0030
	$\hat{\alpha}_{wave}$	0.1993	-0.0007	0.0005	0.0005
0.3	$\hat{\alpha}_{fce}$	0.2962	-0.0038	0.0026	0.0026
	$\hat{\alpha}_{reg}$	0.2962	-0.0038	0.0026	0.0026
	$\hat{\alpha}_{wave}$	0.2986	-0.0014	0.0007	0.0007
0.4	$\hat{\alpha}_{fce}$	0.3997	-0.0003	0.0020	0.0020
	$\hat{\alpha}_{reg}$	0.3999	-0.0001	0.0020	0.0020
	$\hat{\alpha}_{wave}$	0.4019	0.0019	0.0010	0.0010
0.5	$\hat{\alpha}_{fce}$	0.5048	0.0048	0.0027	0.0027
	$\hat{\alpha}_{reg}$	0.5048	0.0048	0.0026	0.0027
	$\hat{\alpha}_{wave}$	0.5017	0.0017	0.0013	0.0013
0.6	$\hat{\alpha}_{fce}$	0.599	-0.0010	0.0029	0.0029
	$\hat{\alpha}_{reg}$	0.5992	-0.0008	0.0029	0.0029
	$\hat{\alpha}_{wave}$	0.6002	0.0002	0.0016	0.0016
0.7	$\hat{\alpha}_{fce}$	0.6982	-0.0018	0.0029	0.0029
	$\hat{\alpha}_{reg}$	0.6984	-0.0016	0.0029	0.0029
	$\hat{\alpha}_{wave}$	0.6989	-0.0011	0.0017	0.0017
0.8	$\hat{\alpha}_{fce}$	0.8036	0.0036	0.0030	0.0030
	$\hat{\alpha}_{reg}$	0.8039	0.0039	0.0030	0.0030
	$\hat{\alpha}_{wave}$	0.8036	0.0036	0.0018	0.0018
0.9	$\hat{\alpha}_{fce}$	0.8977	-0.0023	0.0027	0.0027
	$\hat{\alpha}_{reg}$	0.8980	-0.0020	0.0027	0.0027
	$\hat{\alpha}_{wave}$	0.8988	-0.0012	0.0018	0.0018
1.0	$\hat{\alpha}_{fce}$	0.9972	-0.0028	0.0031	0.0031
	$\hat{\alpha}_{reg}$	0.9972	-0.0028	0.0031	0.0031
	$\hat{\alpha}_{wave}$	0.9984	-0.0016	0.0024	0.0024
1.1	$\hat{\alpha}_{fce}$	1.1115	0.0115	0.0032	0.0033
	$\hat{\alpha}_{reg}$	1.1114	0.0114	0.0032	0.0033
	$\hat{\alpha}_{wave}$	1.1095	0.0095	0.0024	0.0025
1.2	$\hat{\alpha}_{fce}$	1.1971	-0.0029	0.0031	0.0031
	$\hat{\alpha}_{reg}$	1.1974	-0.0026	0.0031	0.0031
	$\hat{\alpha}_{wave}$	1.1977	-0.0023	0.0024	0.0024
1.3	$\hat{\alpha}_{fce}$	1.2940	-0.0060	0.0035	0.0035
	$\hat{\alpha}_{reg}$	1.2940	-0.0060	0.0035	0.0035
	$\hat{\alpha}_{wave}$	1.2963	-0.0037	0.0029	0.0029
1.4	$\hat{\alpha}_{fce}$	1.4066	0.0066	0.0040	0.0040
	$\hat{\alpha}_{reg}$	1.4066	0.0066	0.0040	0.0040
	$\hat{\alpha}_{wave}$	1.4080	0.0080	0.0033	0.0033
1.5	$\hat{\alpha}_{fce}$	1.5013	0.0013	0.0037	0.0037
	$\hat{\alpha}_{reg}$	1.5018	0.0018	0.0037	0.0037
	$\hat{\alpha}_{wave}$	1.5027	0.0027	0.0034	0.0034
1.6	$\hat{\alpha}_{fce}$	1.6019	0.0019	0.0033	0.0033
	$\hat{\alpha}_{reg}$	1.6019	0.0019	0.0033	0.0033
	$\hat{\alpha}_{wave}$	1.6029	0.0029	0.0027	0.0027
1.7	$\hat{\alpha}_{fce}$	1.7022	0.0022	0.0027	0.0027
	$\hat{\alpha}_{reg}$	1.7021	0.0021	0.0027	0.0027
	$\hat{\alpha}_{wave}$	1.7034	0.0034	0.0024	0.0024
1.8	$\hat{\alpha}_{fce}$	1.8004	0.0004	0.0030	0.0029
	$\hat{\alpha}_{reg}$	1.8006	0.0006	0.0030	0.0029
	$\hat{\alpha}_{wave}$	1.8009	0.0009	0.0029	0.0029
1.9	$\hat{\alpha}_{fce}$	1.8975	-0.0025	0.0016	0.0016
	$\hat{\alpha}_{reg}$	1.8977	-0.0023	0.0016	0.0016
	$\hat{\alpha}_{wave}$	1.8968	-0.0032	0.0016	0.0016
2.0	$\hat{\alpha}_{fce}$	2.0018	0.0018	0.0001	0.0001
	$\hat{\alpha}_{reg}$	2.0019	0.0019	0.0001	0.0001
	$\hat{\alpha}_{wave}$	2.0018	0.0018	0.0001	0.0001



**Tabela 5.5:** Resultado da Estimação de  $\alpha$ , baseada em  $\hat{\alpha}_{fce}$ ,  $\hat{\alpha}_{reg}$  e  $\hat{\alpha}_{wave}$ , para  $S_\alpha(0, 1, 0)$  com  $n = 100000$ .

$\alpha$	Estimador	$\hat{\alpha}$	vício	var	eqm
0.1	$\hat{\alpha}_{fce}$	0.1037	0.0037	0.0021	0.0021
	$\hat{\alpha}_{reg}$	0.1037	0.0037	0.0021	0.0021
	$\hat{\alpha}_{wave}$	0.1014	0.0014	0.0003	0.0003
0.2	$\hat{\alpha}_{fce}$	0.2041	0.0041	0.0022	0.0022
	$\hat{\alpha}_{reg}$	0.2040	0.0040	0.0022	0.0022
	$\hat{\alpha}_{wave}$	0.1979	-0.0021	0.0005	0.0005
0.3	$\hat{\alpha}_{fce}$	0.2963	-0.0037	0.0028	0.0028
	$\hat{\alpha}_{reg}$	0.2959	-0.0041	0.0028	0.0028
	$\hat{\alpha}_{wave}$	0.2991	-0.0009	0.0007	0.0007
0.4	$\hat{\alpha}_{fce}$	0.4011	0.0011	0.0027	0.0027
	$\hat{\alpha}_{reg}$	0.4008	0.0008	0.0027	0.0027
	$\hat{\alpha}_{wave}$	0.4011	0.0011	0.0010	0.0010
0.5	$\hat{\alpha}_{fce}$	0.5007	0.0007	0.0028	0.0028
	$\hat{\alpha}_{reg}$	0.5007	0.0007	0.0028	0.0028
	$\hat{\alpha}_{wave}$	0.5005	0.0005	0.0013	0.0013
0.6	$\hat{\alpha}_{fce}$	0.6014	0.0014	0.0030	0.0030
	$\hat{\alpha}_{reg}$	0.6013	0.0013	0.0030	0.0030
	$\hat{\alpha}_{wave}$	0.6013	0.0013	0.0016	0.0016
0.7	$\hat{\alpha}_{fce}$	0.7033	0.0033	0.0035	0.0035
	$\hat{\alpha}_{reg}$	0.7038	0.0038	0.0036	0.0036
	$\hat{\alpha}_{wave}$	0.7029	0.0029	0.0020	0.0020
0.8	$\hat{\alpha}_{fce}$	0.8066	0.0066	0.0030	0.0030
	$\hat{\alpha}_{reg}$	0.8064	0.0064	0.0030	0.0030
	$\hat{\alpha}_{wave}$	0.8054	0.0054	0.0019	0.0019
0.9	$\hat{\alpha}_{fce}$	0.8974	-0.0026	0.0028	0.0028
	$\hat{\alpha}_{reg}$	0.8973	-0.0027	0.0028	0.0028
	$\hat{\alpha}_{wave}$	0.8974	-0.0026	0.0020	0.0020
1.0	$\hat{\alpha}_{fce}$	0.9957	-0.0043	0.0031	0.0031
	$\hat{\alpha}_{reg}$	0.9955	-0.0045	0.0030	0.0030
	$\hat{\alpha}_{wave}$	0.9978	-0.0022	0.0021	0.0021
1.1	$\hat{\alpha}_{fce}$	1.0983	-0.0017	0.0036	0.0035
	$\hat{\alpha}_{reg}$	1.0986	-0.0014	0.0036	0.0035
	$\hat{\alpha}_{wave}$	1.0971	-0.0029	0.0026	0.0026
1.2	$\hat{\alpha}_{fce}$	1.2108	0.0108	0.0038	0.0039
	$\hat{\alpha}_{reg}$	1.2109	0.0109	0.0038	0.0039
	$\hat{\alpha}_{wave}$	1.2077	0.0077	0.0031	0.0031
1.3	$\hat{\alpha}_{fce}$	1.2982	-0.0018	0.0030	0.0030
	$\hat{\alpha}_{reg}$	1.2980	-0.0020	0.0030	0.0030
	$\hat{\alpha}_{wave}$	1.3003	0.0003	0.0024	0.0024
1.4	$\hat{\alpha}_{fce}$	1.4022	0.0022	0.0033	0.0033
	$\hat{\alpha}_{reg}$	1.4016	0.0016	0.0033	0.0033
	$\hat{\alpha}_{wave}$	1.4031	0.0031	0.0026	0.0026
1.5	$\hat{\alpha}_{fce}$	1.5047	0.0047	0.0038	0.0038
	$\hat{\alpha}_{reg}$	1.5047	0.0047	0.0038	0.0038
	$\hat{\alpha}_{wave}$	1.5051	0.0051	0.0034	0.0034
1.6	$\hat{\alpha}_{fce}$	1.5989	-0.0011	0.0031	0.0031
	$\hat{\alpha}_{reg}$	1.5991	-0.0009	0.0031	0.0031
	$\hat{\alpha}_{wave}$	1.6010	0.0010	0.0029	0.0029
1.7	$\hat{\alpha}_{fce}$	1.7024	0.0024	0.0032	0.0032
	$\hat{\alpha}_{reg}$	1.7023	0.0023	0.0032	0.0032
	$\hat{\alpha}_{wave}$	1.7020	0.0020	0.0032	0.0032
1.8	$\hat{\alpha}_{fce}$	1.8059	0.0059	0.0024	0.0024
	$\hat{\alpha}_{reg}$	1.8058	0.0058	0.0024	0.0024
	$\hat{\alpha}_{wave}$	1.8071	0.0071	0.0022	0.0022
1.9	$\hat{\alpha}_{fce}$	1.9039	0.0039	0.0014	0.0014
	$\hat{\alpha}_{reg}$	1.9039	0.0039	0.0013	0.0014
	$\hat{\alpha}_{wave}$	1.9050	0.0050	0.0013	0.0013
2.0	$\hat{\alpha}_{fce}$	1.9992	-0.0008	0.0001	0.0001
	$\hat{\alpha}_{reg}$	1.9992	-0.0008	0.0001	0.0001
	$\hat{\alpha}_{wave}$	1.9991	-0.0009	0.0002	0.0002

# Capítulo 6

## Análise em Sequências de DNA

Neste capítulo analisamos o parâmetro de *longa dependência* através de diferentes transformações em sequência de DNA (ver Crato et al., 2011). Analisamos a longa dependência em sequências de DNA utilizando a metodologia de mudança de regimes, proposta por Liu (2000). Por fim, estudamos as distribuições de distâncias das regiões codantes e não codantes em sequências de DNA (ver Crato et al., 2011).

### 6.1 Longa Dependência Através de Diferentes Transformações

Uma sequência de DNA é composta de bases A (adenina), C (citosina), T (timina) e G (guanina). Para aplicar métodos numéricos a uma sequência de nucleotídeos é necessário transformá-la em uma sequência numérica. Do ponto de vista biológico, é comum aplicar a regra SW “*strong-weak pairing*”, a qual transforma C e G em 1 e A e T em 0. Podemos igualmente fazer a transformação de acordo com o alfabeto da purina-pirimidina (regra RY), que é  $A = G = 1$  e  $C = T = -1$ . Esta regra descreve como as purinas (A e G) e as pirimidinas (C e T) são descritas ao longo da sequência.

Guharay et al. (2000) consideram as quatro bases A, C, G, T em um tetraedro onde cada letra é equidistante das demais e as componentes dos vetores somam zero, isto é, transforma A em  $(0.75, -0.25, -0.25, -0.25)$ , C em  $(-0.25, 0.75, -0.25, -0.25)$ , G em  $(-0.25, -0.25, 0.75, -0.25)$  e T em  $(-0.25, -0.25, -0.25, 0.75)$ .

Cristea (2003) propõe a seguinte representação para uma sequência de nucleotídeos: considera quatro vetores de mesmo tamanho, onde cada um é simétrico aos demais, i.e., orientado para os cantos de um tetraedro, e os vetores são colocados em correspondência com os nucleotídeos, na qual destaca-se que os vértices de um tetraedro regular são subconjuntos dos vértices de um cubo. Portanto, esta regra transforma A em  $(1, 1, 1)$ , C em

$(-1, 1, -1)$ , G em  $(-1, -1, -1)$  e T em  $(1, -1, -1)$ .

Para analisar a longa dependência consideramos a sequência de DNA AL163202, correspondente a uma parte do cromossomo 21 do *Homo sapiens* (com 340,000 pares de bases). Consideramos os estimadores  $\hat{d}_W$ ,  $\hat{d}_{GPH}$ ,  $\hat{d}_{DFA}$  e  $\hat{d}_{WS}$ , descritos na Seção 2.2 e no Capítulo 3, calculados para as séries numéricas obtidas para as possíveis transformações  $f : \{A, C, G, T\} \rightarrow \{0, 1, 2, 3\}$ . Note que temos um total de  $4^4$  diferentes transformações, nas quais algumas são equivalentes. Por exemplo,  $\{1, 0, 0, 0\}$ ,  $\{2, 0, 0, 0\}$ ,  $\{3, 0, 0, 0\}$  e  $\{4, 0, 0, 0\}$  tem necessariamente as mesmas correlações e estimativas espectrais. Neste sentido, elas constituem uma classe de equivalência de transformações. Na Tabela 6.1 escolhemos aquelas que são representantes de cada classe, ou seja, escolhemos as transformações relevantes. Embora cada transformação apresente uma propriedade diferente da sequência, note que todas elas apresentam a característica de longa dependência. Para cada sequência, testamos a hipótese  $H_0 : d = 0$  versus  $H_1 : d \neq 0$ , cuja estatística do teste é dada por

$$Z = \frac{\hat{d}_j - d_{H_0}}{\sqrt{\text{Var}(\hat{d}_j)}} = \frac{\hat{d}_j}{\sqrt{\text{Var}(\hat{d}_j)}}, \quad j \in \{W, GPH, DFA\},$$

que tem função de distribuição  $\mathcal{N}(0, 1)$ , onde  $\text{Var}(\hat{d})$  é a variância do estimador  $\hat{d}$ , com nível de significância de 5%, para todos os estimadores propostos.

**Tabela 6.1:** Estimadores para o parâmetro  $d$ , para a sequência de DNA AL163202, usando 30 diferentes transformações.

{A,C,G,T}	$\hat{d}_W$	$\hat{d}_{GPH}$	$\hat{d}_{DFA}$	{A,C,G,T}	$\hat{d}_W$	$\hat{d}_{GPH}$	$\hat{d}_{DFA}$
(0,0,0,1)	0.0766	0.4016	0.1073	(0,1,1,1)	0.0749	0.3015	0.1133
(0,0,1,0)	0.0593	0.2660	0.1033	(0,1,1,2)	0.0923	0.3665	0.0872
(0,1,0,0)	0.0597	0.2581	0.0955	(0,1,1,3)	0.0901	0.3664	0.0892
(0,0,1,1)	0.0804	0.3836	0.0649	(0,0,2,3)	0.0823	0.3980	0.0744
(0,1,0,1)	0.0965	0.3060	0.0817	(0,2,0,3)	0.0947	0.3453	0.0884
(0,1,1,0)	0.0382	0.3137	0.1748	(0,0,3,1)	0.0652	0.2613	0.0776
(0,0,1,2)	0.0820	0.4016	0.0820	(0,0,3,2)	0.0750	0.3377	0.0633
(0,1,0,2)	0.0915	0.3606	0.0931	(0,1,2,1)	0.0559	0.2988	0.1247
(0,1,2,0)	0.0422	0.3081	0.1602	(0,1,2,2)	0.0809	0.3561	0.0870
(0,0,1,3)	0.0808	0.3976	0.0907	(0,1,2,3)	0.0880	0.3875	0.0812
(0,1,0,3)	0.0871	0.3840	0.0982	(0,1,3,1)	0.0526	0.2843	0.1211
(0,1,3,0)	0.0464	0.2931	0.1446	(0,1,3,2)	0.0696	0.3180	0.0908
(0,0,2,1)	0.0704	0.2863	0.0680	(0,1,3,3)	0.0816	0.3772	0.0779
(0,2,0,1)	0.0855	0.2489	0.0790	(0,2,1,2)	0.0907	0.3066	0.0951
(0,2,1,0)	0.0425	0.3024	0.1534	(0,2,1,3)	0.0954	0.3407	0.0888

Melhor do que considerar transformações arbitrárias, a abordagem do envelope espectral (ver Stoffer et al., 1993) seleciona as transformações que

ênfatizam a característica periódica da sequência de DNA. Consideramos também o estimador  $\hat{d}_{WS}$ , ver Definição 2.25, para uma análise utilizando esta abordagem, onde obtivemos  $\hat{d}_{WS} = 0.0722$ . No entanto, não colocamos este resultado na Tabela 6.1 pois ele é um valor único para todas as transformações. Para todos os estimadores apresentados na Tabela 6.1, inclusive o  $\hat{d}_{WS}$ , rejeita-se  $H_0$  ao nível de significância de 5%. Através da Tabela 6.1 observamos que, entre os quatro estimadores propostos para o parâmetro  $d$ , os valores mais altos foram sempre obtidos pelo estimador  $\hat{d}_{GPH}$ . Os estimadores  $\hat{d}_W$ ,  $\hat{d}_{WS}$  e  $\hat{d}_{DFA}$  são da mesma ordem de valor. No entanto, a característica mais importante, é que para todos os estimadores utilizados, rejeita-se  $H_0$  ao nível de significância de 5%, isto é, a sequência de DNA AL163202 apresenta longa dependência.

## 6.2 Mudança de Regime

Nesta seção apresentamos o processo de mudança de regime, proposto por Liu (2000), e suas propriedades. Após a apresentação desta teoria, analisamos diferentes sequências de DNA sob este ponto de vista.

Na literatura existem diferentes transformações que aplicam métodos numéricos a uma sequência de DNA (ver Peng et al., 1992; Stanley et al., 1999; Guharay et al., 2000; Cristea, 2003; Garcia e José, 2005; Lopes e Nunes, 2006; Podobnik et al., 2007 e Crato et al., 2010b). Se consideramos a abordagem de transformação numérica arbitrária para os nucleotídeos  $\{A, C, T, G\}$  e depois utilizamos a análise espectral, o resultado irá depender de cada transformação em particular. Portanto, não sabemos se a existência de longa dependência em sequência de DNA é induzida pela transformação. Nosso objetivo é analisar a longa dependência em várias sequências de DNA utilizando a metodologia de mudança de regimes proposta por Liu (2000). Nesta metodologia, se a duração dos regimes de uma série temporal tem uma distribuição de cauda pesada com parâmetro  $\alpha \in (0, 2)$ , então a série temporal apresenta a característica de longa dependência, e se aplicar qualquer transformação linear que preserva a propriedade de variância finita na série temporal, igualmente preserva a propriedade de longa dependência.

**Definição 6.1 (Processo com Mudança de Regimes).** Sejam  $w_0, w_1, \dots$  variáveis aleatórias independentes e identicamente distribuídas, com média zero e variância finita  $\sigma^2$ , onde a variável aleatória  $w_k$  é constante no  $k$ -ésimo regime. Seja  $I_k(\cdot)$  a função indicadora para o evento onde  $w_k$  sobrevive do período  $k$  até  $t$ , ou seja,

$$I_k(t) = \begin{cases} 1, & \text{se } k < t \leq k + T_k \\ 0, & \text{c.c.} \end{cases}$$

onde  $T_k$  é o tempo em que o  $k$ -ésimo regime termina, chamamos de  $k$ -ésima duração de regime. Assumimos que  $w_k$  é independente de  $T_k$ . Seja  $\{W_t\}_{t \in \mathbb{N}}$  um processo dado por

$$W_t = \sum_{k=0}^{\infty} w_k I_k(t), \quad t \in \mathbb{N},$$

então  $\{W_t\}_{t \in \mathbb{N}}$  é chamado *um processo com mudança de regime*.

**Observação 6.1.** A hipótese de média zero na Definição 6.1 é somente por conveniência analítica. Podemos eliminá-la sem fazer dano ao argumento.

O teorema a seguir mostra que, sob o mecanismo de mudança de regime com índice  $\alpha$ , o processo com mudança de regime tem longa dependência com  $d = 1 - \frac{\alpha}{2}$ .

**Teorema 6.1.** *Seja  $\{W_t\}_{t \in \mathbb{N}}$  um processo com mudança de regime. Suponha que os tempos de mudança de regimes  $\{T_k\}_{k \in \mathbb{N} \cup \{0\}}$  são i.i.d. com distribuição de cauda pesada na forma de*

$$\lim_{t \rightarrow \infty} \frac{\mathbb{P}(T_k > t)}{t^{-\alpha} h(t)} = 1,$$

onde  $1 < \alpha < 2$  e  $h(\cdot)$  é função com pouca variação. Então, o processo com mudança de regime tem longa dependência com parâmetro de longa dependência dado por  $d = 1 - \frac{\alpha}{2}$ .

**Demonstração:** Ver Liu (2000), página 144. □

O próximo teorema mostra que a longa dependência é preservada, quando é aplicada qualquer transformação linear que preserva a propriedade de variância finita no processo  $\{W_t\}_{t \in \mathbb{N}}$  com mudança de regimes.

**Teorema 6.2.** *Seja  $\{W_t\}_{t=1}^n$  uma série temporal obtida de um processo com mudança de regimes  $\{W_t\}_{t \in \mathbb{N}}$ . Suponha que as durações de regimes  $\{T_k\}_{k \in \mathbb{N} \cup \{0\}}$  são i.i.d. com função de distribuição com cauda pesada*

$$\lim_{t \rightarrow \infty} \frac{\mathbb{P}(T_k > t)}{t^{-\alpha} h(t)} = 1,$$

onde  $1 < \alpha < 2$  e  $h(\cdot)$  é função com pouca variação. Então, qualquer transformação linear que preserva a propriedade de variância finita no processo  $\{W_t\}_{t \in \mathbb{N}}$  com mudança de regimes, igualmente preserva a propriedade de longa dependência com mesma magnitude  $d = 1 - \frac{\alpha}{2}$ .

**Demonstração:** Ver Liu (2000), página 145. □

Nesta seção, consideramos a seguinte definição, para uma série temporal que representará uma sequência de DNA.

**Definição 6.2 (Série Temporal de uma Sequência de DNA).** Dada uma sequência de DNA  $\{n_i\}_{i=1}^n$ , a *série temporal*  $\{W_t\}_{t=1}^n$ , obtida desta sequência, é dada por

$$W_t = f(n_t), \quad (6.1)$$

onde  $f(\cdot)$  é qualquer transformação linear que preserva a propriedade de variância finita.

Consideramos que a mudança de regime ocorre em uma sequência de DNA, quando há uma mudança na base da sequência. Seja  $\{w_k\}_{k=1}^m$  os  $m$  diferentes regimes da sequência de DNA. Os regimes são relacionados a uma variável de estado latente  $w$ , que toma o valor  $w_k$  no  $k$ -ésimo regime. Seja  $\{T_k\}_{k=1}^m$  as *durações de regimes* para a sequência de DNA. Nesta seção, investigamos a propriedade de longa dependência, focalizando nossa atenção no parâmetro estável  $\alpha$  da série temporal  $\{T_k\}_{k=1}^m$ , para várias sequências de DNA.

**Exemplo 6.1.** Seja uma parte de uma sequência de DNA de tamanho  $n = 16$  dada por

$$\{n_i\}_{i=1}^{16} = \left\{ \underbrace{\text{C}}_{w_1} \underbrace{\text{AAA}}_{w_2} \underbrace{\text{T}}_{w_3} \underbrace{\text{AAAAAA}}_{w_4} \underbrace{\text{C}}_{w_5} \underbrace{\text{A}}_{w_6} \underbrace{\text{C}}_{w_7} \underbrace{\text{A}}_{w_8} \underbrace{\text{T}}_{w_9} \right\}.$$

Considerando que a mudança de regime em uma sequência de DNA, ocorre quando há uma mudança na base nesta sequência, então os  $\{w_k\}_{k=1}^9$  são os 9 diferentes regimes da sequência  $\{n_i\}_{i=1}^{16}$ . Observe que a série temporal das durações destes 9 regimes, para a sequência  $\{n_i\}_{i=1}^{16}$ , é dada por  $\{T_k\}_{k=1}^9 = \{1, 3, 1, 6, 1, 1, 1, 1, 1\}$ .

Para analisar a longa dependência, consideramos sequências de DNA, apresentadas na Tabela 6.2, correspondentes a partes do cromossomo 21 (de AL163202 até AL163210 na Tabela 6.2) da espécie *Homo sapiens*, as partes do cromossomo X da espécie *Homo sapiens* (de AC004673 até AY286122 na Tabela 6.2) e a sequência completa do cromossomo 1 da espécie *Leishmania braziliensis* (AM494938). Estas sequências estão disponíveis em “*European Bioinformatics Institute*” (EBI, <http://www.ebi.ac.uk/>). Para algumas sequências apresentadas na Tabela 6.2, a Figura 6.2 mostra histogramas das durações de regimes  $\{T_k\}_{k=1}^m$ , das respectivas sequências de DNA. Podemos observar distribuição com cauda pesada em todos os gráficos da Figura 6.2. Uma característica notável na distribuição de duração de regimes é o decaimento da cauda cuja taxa de decaimento é dada pelo parâmetro  $\alpha$ . Pelo Teorema 6.1, dada uma série temporal  $\{T_k\}_{k=1}^m$  de uma sequência de DNA, se  $1 < \alpha < 2$ , então a sequência de DNA apresenta a característica de longa

dependência. E, pelo Teorema 6.2, qualquer transformação linear  $f(\cdot)$  aplicada na sequência de DNA, que preserva variância finita do processo todo, igualmente preserva a propriedade de longa dependência. Para a análise da taxa de decaimento da cauda, consideramos os estimadores  $\hat{\alpha}_{\text{mle}}$ ,  $\hat{\alpha}_{\text{fce}}$ ,  $\hat{\alpha}_{\text{reg}}$  e  $\hat{\alpha}_{\text{wave}}$ , dados na Seção 5.2.

Para todos os estimadores na Tabela 6.2, testamos as hipóteses  $H_0 : \alpha = 2$  versus  $H_1 : \alpha < 2$ . A estatística do teste para os estimadores é dada por

$$Z = \frac{\hat{\alpha} - \alpha}{\sqrt{\text{var}(\hat{\alpha})}},$$

onde  $Z$  tem distribuição  $\mathcal{N}(0, 1)$  e  $\text{var}(\hat{\alpha})$  é a variância (ou variância assintótica) do estimador  $\hat{\alpha}$ , dadas pelas expressões (5.23), (5.27), (5.30) e (5.32), respectivamente. Para todos os estimadores apresentados na Tabela 6.2, rejeita-se  $H_0$  ao nível de significância de 5%.

**Observação 6.2.** Para a Tabela 6.2, consideramos as seguintes notações:

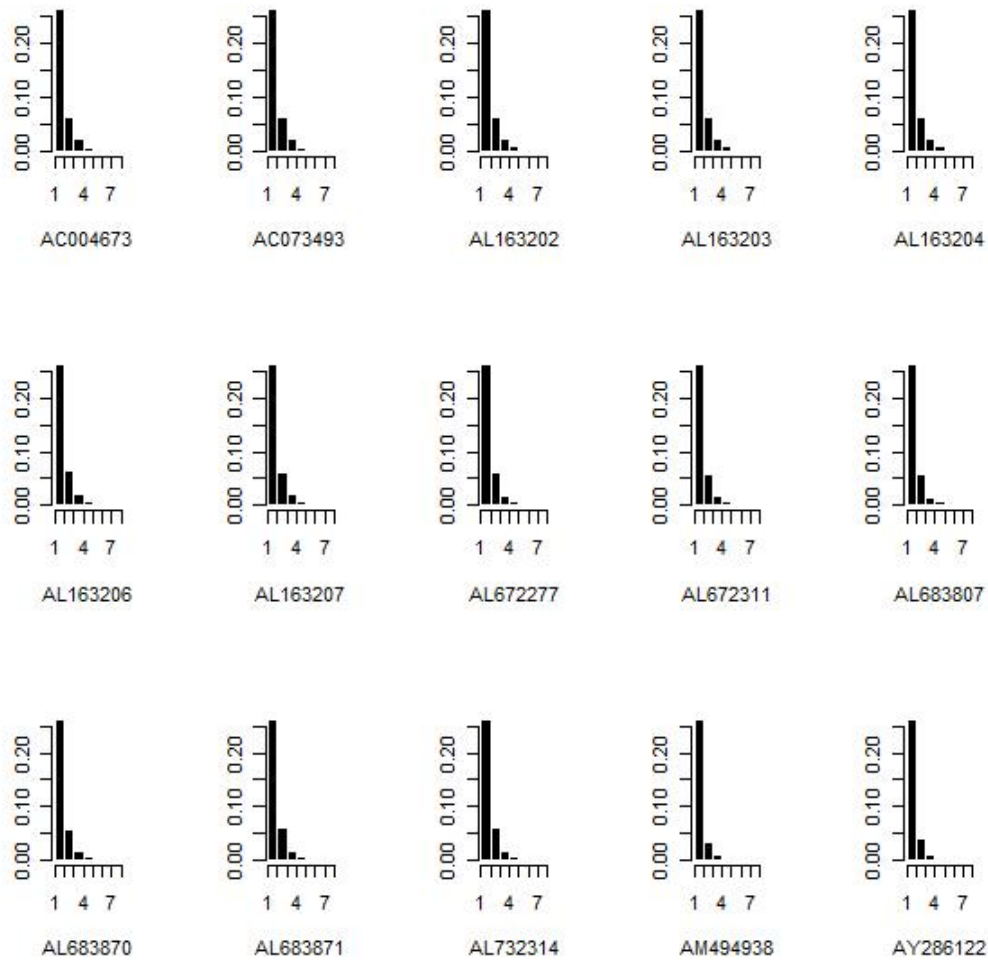
$n$  : número de pares de bases da sequência de DNA;

$m$  : tamanho da série temporal das durações de regimes  $\{T_k\}_{k=1}^m$ .

**Tabela 6.2:** Estimadores  $\hat{\alpha}_{\text{mle}}$ ,  $\hat{\alpha}_{\text{fce}}$ ,  $\hat{\alpha}_{\text{reg}}$  e  $\hat{\alpha}_{\text{wave}}$  para o Índice  $\alpha$ , das Séries Temporais de Durações de Regimes, de Sequências de DNA.

Sequência	$n$	$m$	$\hat{\alpha}_{\text{mle}}$	$\hat{d}_{\text{mle}}$	$\hat{\alpha}_{\text{fce}}$	$\hat{d}_{\text{fce}}$	$\hat{\alpha}_{\text{reg}}$	$\hat{d}_{\text{reg}}$	$\hat{\alpha}_{\text{wave}}$	$\hat{d}_{\text{wave}}$
AL163202	340000	236804	1.9459	0.0270	1.7824	0.1088	1.7814	0.1092	1.7814	0.1092
AL163203	340000	237233	1.9440	0.0280	1.7728	0.1136	1.7731	0.1134	1.7731	0.1134
AL163204	340000	238266	1.9597	0.0200	1.8051	0.0974	1.8035	0.0982	1.8035	0.0982
AL163206	340000	236207	1.9514	0.0243	1.7881	0.1059	1.7881	0.1059	1.7881	0.1059
AL163207	340000	237473	1.9530	0.0235	1.7906	0.1047	1.7908	0.1045	1.7908	0.1045
AL163208	340000	237467	1.9545	0.0227	1.7939	0.1031	1.7940	0.1288	1.7940	0.1288
AL163209	340000	237083	1.9537	0.0231	1.7902	0.1049	1.7900	0.1049	1.7900	0.1049
AL163210	340000	235623	1.9434	0.0283	1.7708	0.1146	1.7704	0.1042	1.7704	0.1042
AC004673	236281	164946	1.9521	0.0239	1.7809	0.1095	1.7797	0.1101	1.7797	0.1101
AC073493	211422	148826	1.9482	0.0259	1.7764	0.1118	1.7744	0.1127	1.7744	0.1127
AL672277	131682	93399	1.9461	0.0269	1.7379	0.1311	1.7390	0.1304	1.7390	0.1304
AL672311	115998	83935	1.9526	0.0237	1.7608	0.1196	1.7546	0.1304	1.7546	0.1304
AL683807	189825	135315	1.9464	0.0268	1.7362	0.1319	1.7336	0.1331	1.7336	0.1331
AL683870	162377	117373	1.9595	0.0202	1.7729	0.1136	1.7723	0.1138	1.7723	0.1138
AL683871	175765	126088	1.9478	0.0261	1.7425	0.1288	1.7423	0.1288	1.7423	0.1288
AL732314	218723	156921	1.9505	0.0247	1.7454	0.1273	1.7431	0.1284	1.7431	0.1284
AL445312	170984	123196	1.9687	0.0156	1.8179	0.0911	1.8187	0.0906	1.8187	0.0906
AL450023	185532	131444	1.9510	0.0245	1.7629	0.1186	1.7617	0.1191	1.7617	0.1191
AL591435	138038	95527	1.9432	0.0284	1.7730	0.1135	1.7726	0.1136	1.7726	0.1136
AL929410	186649	130219	1.9457	0.0271	1.7820	0.1090	1.7836	0.1082	1.7836	0.1082
Z98255	169998	118328	1.9476	0.0262	1.7911	0.1044	1.7915	0.1042	1.7915	0.1042
AL732374	224187	156454	1.9381	0.0309	1.7335	0.1332	1.7316	0.1341	1.7316	0.1341
AY286122	422021	327376	1.9906	0.0004	1.8804	0.0598	1.8805	0.0597	1.8805	0.0597
AM494938	218723	180103	1.9576	0.0212	1.7386	0.1307	1.7392	0.1303	1.7392	0.1303

Podemos observar na Tabela 6.2, que  $\hat{\alpha}_{\text{mle}} \in (1, 2)$ ,  $\hat{\alpha}_{\text{fce}} \in (1, 2)$ ,  $\hat{\alpha}_{\text{reg}} \in (1, 2)$  e  $\hat{\alpha}_{\text{wave}} \in (1, 2)$  a nível de significância de 5%, para todas as sequências de DNA aqui estudadas. Portanto, pelo mecanismo de mudança de regime proposto por Liu (2000), as sequências de DNA, dadas na Tabela



**Figura 6.1:** Histogramas das Séries Temporais de Durações de Regimes de Algumas Sequências de DNA Dadas na Tabela 6.2.

6.2, apresentam a característica de longa dependência. Segue-se, pelo Teorema 6.2, que qualquer transformação linear  $f(\cdot)$  aplicada na sequência de DNA, que preserve variância finita, igualmente preserva a propriedade de longa dependência da sequência de DNA. Portanto, a longa dependência nas vinte e quatro sequências de DNA da Tabela 6.2 é realmente propriedade da sequência e não da transformação utilizada.



## 6.3 Análise da Distribuição do Comprimento

A *região codante* de uma sequência de DNA qualquer é a sequência de nucleotídeos que, após transcrição e tradução, leva à síntese de uma proteína com uma determinada sequência de aminoácidos. A *região não codante* de uma sequência de DNA qualquer, localizada a montante da codante, inclui elementos ou pequenas sequências, envolvidas na regulação da transcrição. Nesta seção consideramos o estimador de Hill para analisarmos as distribuições de distâncias das regiões codantes e não codantes em sequências de DNA. Esta análise nos ajuda a identificar diferenças entre estas distribuições e a identificar existência de longa dependência em regiões não codantes de sequências de DNA. Um problema comum na análise de sequências de DNA é como identificar regiões codantes (*cds*) dispersas por toda a sequência e separadas por regiões não codantes (*ncds*). Distinguindo a distribuição para cada uma destas regiões pode ser uma ajuda para encontrar e separar regiões codantes e não codantes em sequências de DNA.

Nosso interesse aqui encontra-se na modelagem probabilística e na inferência estatística para as partes extremas das caudas das distribuições de  $X_t$  para regiões codantes (*cds*), e de  $Y_t$  para regiões não codantes (*ncds*), de algumas sequências de DNA (ver Seção 6.3) estimando o parâmetro  $\alpha$ .

**Definição 6.3 (Sequência das Distâncias de Região Codante ou não Codante).** Seja  $\{n_j\}_{j=1}^n$  uma sequência de DNA. Ordenando as distâncias dos segmentos codantes (ou não codantes), de acordo com sua ordem no genoma completo, a sequência obtida é chamada *sequência das distâncias de região codante (ou não codante)* ou *sequência dos comprimentos de região codante (ou não codante)*.

**Observação 6.3.** Pela Definição 6.3, obtemos a série temporal  $\{X_t\}_{t=1}^n$  (analogamente, a série  $\{Y_t\}_{t=1}^n$ ) derivada da *sequência das distâncias de região codante (ou não codante)* dadas por

$$\begin{aligned} X_t &= \text{a distância do segmento codante na posição } t \\ Y_t &= \text{a distância do segmento não codante na posição } t, \end{aligned}$$

onde as *distâncias dos segmentos codantes (ou não codantes)* significa o número de bases no segmento.

A Figura 6.2 mostra os histogramas das *sequências das distâncias de regiões codantes e não codantes*, respectivamente, para o cromossomo 21 do *Homo sapiens*. A cauda observada na Figura 6.2 ilustra quão difícil pode ser atribuir probabilidades aos eventos extremos usando toda a série de dados. Podemos notar diferenças entre os histogramas das *sequência das distâncias de região codante e não codante*.

Nosso interesse é analisar a distribuição das regiões codantes e não codantes, considerando, respectivamente, as sequências das distâncias de região

codante e não codante (ver Definição 6.3). A Tabela 6.3 apresenta, respectivamente, o estimador de Hill  $\hat{\alpha}_{cds}$  para a *sequência das distâncias de região codante* e  $\hat{\alpha}_{ncds}$  para a *sequência das distâncias de região não codante* para todos os cromossomos do *Homo sapiens*. Este conjunto de dados está disponível em “*European Bioinformatics Institute*” (EBI, <http://www.ebi.ac.uk/>). Como cada sequência começa e termina com segmento codante, sempre teremos um segmento a mais (coluna 3) do que na região não codante (coluna 5). O parâmetro  $\alpha$  avalia a taxa de decaimento da cauda, a qual determina a probabilidade de ocorrência de eventos extremos e a existência dos momentos para a distribuição subjacente. Aqui  $k$  é dado por  $n^{0.5}$ , onde  $n$  é o tamanho da amostra. A escolha de  $k = n^{0.5}$  foi baseada nas condições dada pelo Teorema 5.5.

Para todo o estimador de Hill  $\hat{\alpha}_i$ ,  $i \in \{cds, ncds\}$ , apresentado na Tabela 6.3 com  $\hat{\alpha}_i < 2$ , testamos as hipóteses  $H_0 : \alpha = 2$  versus  $H_1 : \alpha < 2$ . E para todo o estimador de Hill  $\hat{\alpha}_i$ ,  $i \in \{cds, ncds\}$ , apresentado na Tabela 6.3 com  $\hat{\alpha}_i > 2$ , testamos as hipóteses  $H_0 : \alpha = 2$  versus  $H_2 : \alpha > 2$ . A estatística dos testes é dada por

$$\begin{aligned} Z &= \frac{\hat{\alpha}_i - \alpha}{\sqrt{\text{Var}(\hat{\alpha}_i)}} = \frac{\hat{\alpha}_i - \alpha}{\frac{\alpha}{\sqrt{k}}} \\ &= \sqrt{k} \left( \frac{\hat{\alpha}_i}{\alpha} - 1 \right), \end{aligned}$$

que tem distribuição  $\mathcal{N}(0, 1)$ , com nível de significância de 5%.

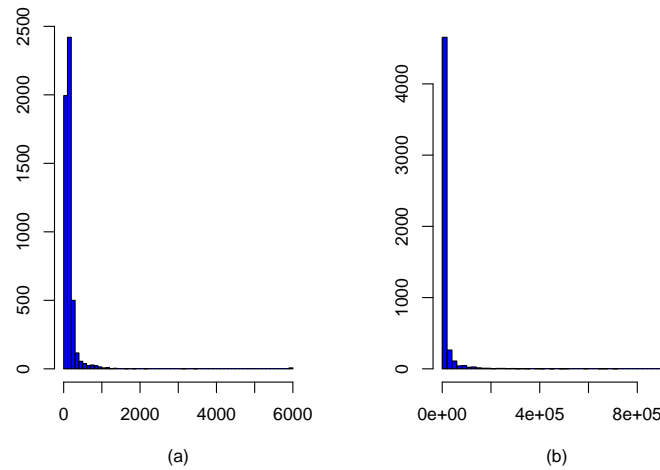
Pela Tabela 6.3, observamos que regiões codantes apresentam estimativas para o parâmetro  $\alpha$  geralmente maiores do que aquelas para as regiões não codantes. Apenas para dois casos (cromossomos 15 e 18) a primeira estimativa é menor que a segunda correspondente. Embora seja necessário uma melhor investigação, podemos suspeitar que o parâmetro  $\alpha$  é geralmente maior do que 2 para regiões codantes e menor do que 2 para regiões não codantes.

Observa-se na Tabela 6.3, que para as regiões codantes, o parâmetro  $\alpha$  é significativamente maior ou igual a 2. Logo, se  $\alpha < 2$  podemos suspeitar que a distribuição subjacente é a de uma região não codante, não valendo o contrário.

**Tabela 6.3:** Estimador de Hill para as sequências das distâncias de região codante ( $\hat{\alpha}_{c ds}$ ) e não codante ( $\hat{\alpha}_{nc ds}$ ) para os cromossomos do *Homo sapiens*, com  $k = \lfloor n^{0.5} \rfloor$ .

Cromossomo	Sequência	Nº de Codantes	$\hat{\alpha}_{c ds}$	Nº de Não Codantes	$\hat{\alpha}_{nc ds}$
1	CM000252	45119	2.6216*	45118	1.8538
2	CM000253	35685	2.0099	35684	1.8062
3	CM000254	27943	2.1347	27942	1.9630
4	CM000255	16325	1.9904	16324	1.8526
5	CM000256	19582	2.2984*	19581	2.0941
6	CM000257	23586	2.3579*	23585	1.6658*
7	CM000258	22096	2.4524*	22095	1.9517
8	CM000259	14854	2.3956*	14853	2.0456
9	CM000260	18837	1.9965	18836	1.7615
10	CM000261	20235	2.2170	20234	1.7593
11	CM000262	25522	2.2307	25521	1.8608
12	CM000263	25970	3.0919*	25969	2.2118
13	CM000264	9149	1.7409	9148	1.7010
14	CM000265	15114	2.2924	15113	1.4330*
15	CM000266	15032	1.9473	15031	2.5036*
16	CM000267	19098	1.9741	19097	1.7487
17	CM000268	26740	2.2214	26739	1.8042
18	CM000269	7317	2.2275	7316	2.6834*
19	CM000270	23421	2.6367*	23420	2.3779*
20	CM000271	13203	2.4240*	13202	1.6034*
21	CM000272	5243	2.1805	5242	1.6884
22	CM000273	5243	2.1805	5242	1.6884
X	CM000274	15604	2.3963*	15603	1.5784*
Y	CM000275	361	2.1131	360	2.0556

Nota: \* *rejeita-se a hipótese  $H_0$  ao nível de significância de 5%.*



**Figura 6.2:** Histograma para Sequências das Distâncias de: (a) Regiões Codantes e (b) Não Codantes do Cromossomo 21 do *Homo sapiens*.

# Capítulo 7

## Conclusão

Descrevemos, no Capítulo 3, o método da análise de flutuações destendenciadas (DFA) e analisamos as propriedades estatísticas do estimador DFA, para os casos em que a sequência dos erros  $\{\epsilon_j\}_{j=1}^m$ , no modelo de regressão, tem função de distribuição Gaussiana ou não. Apresentamos condição necessária para que o estimador DFA tenha função de distribuição Gaussiana exata. Mostramos que  $\lfloor n^\beta \rfloor$ , para algum  $0 < \beta < 1$ , é uma escolha ótima assintótica para  $m = g(n)$  e, sob certas condições, provamos a normalidade assintótica do estimador DFA. Para investigar o efeito de  $g(n) = \lfloor n^\beta \rfloor$ , para algum  $0 < \beta < 1$ , no método da análise de flutuações destendenciadas, geramos séries temporais obtidas de processos ARFIMA(0,  $d$ , 0) e ruído Gaussiano fracionário (FGN). Consideramos diferentes valores para  $g(n)$ , incluindo aquele mais utilizado na literatura, dado por  $\lfloor \frac{n}{10} \rfloor$ . Em nosso experimento,  $g(n) = \lfloor n^{0.3} \rfloor$ , parece gerar maior vício para os estimadores  $\hat{H}_{\text{DFA}}$  e  $\hat{d}_{\text{DFA}}$ . Observamos que a escolha mais utilizada na literatura,  $g(n) = \lfloor \frac{n}{10} \rfloor$ , gera boas estimativas para os parâmetros  $H$  e  $d$ . No entanto,  $g(n) = \lfloor n^\beta \rfloor$  apresenta melhores resultados do que  $g(n) = \lfloor \frac{n}{10} \rfloor$ , para  $d \in \{0.1, 0.2, 0.3, 0.4\}$  se  $\beta \in \{0.5, 0.6, 0.7\}$  e para  $H \in \{0.6, 0.7, 0.8, 0.9\}$  se  $\beta = 0.7$ . Para  $n$  suficientemente grande, a escolha  $g(n) = \lfloor n^\beta \rfloor$ , sob o ponto de vista computacional, tem a vantagem de exigir um número menor de regressores no método DFA. Para  $n$  suficientemente grande, a escolha  $g(n) = \lfloor n^\beta \rfloor$ , para finalidades computacionais, tem a vantagem de exigir um número menor de regressores no método DFA.

No Capítulo 4, apresentamos um resumo sobre a teoria de wavelets e descrevemos a metodologia de *encolhimento de wavelets*. Foi introduzida a teoria dos valores extremos no Capítulo 5 onde definimos distribuição  $\alpha$ -estáveis, apresentamos alguns estimadores para o parâmetro  $\alpha$  e propomos um novo método de estimação utilizando o procedimento de *encolhimento de wavelets*, para obter o estimador  $\hat{\alpha}_{\text{wave}}$ . Para analisar o estimador  $\hat{\alpha}_{\text{wave}}$ , geramos séries temporais  $\alpha$ -estáveis simétricas de média 0 e comparamos com os estimadores baseados na função característica empírica,  $\hat{\alpha}_{\text{fce}}$  e  $\hat{\alpha}_{\text{reg}}$ . Em

quase todo o experimento, verificou-se menores valores da variância e do erro quadrático médio para o estimador  $\hat{\alpha}_{\text{wave}}$ .

A análise de diversas sequências de DNA, com o uso do software R-project, foi apresentada no Capítulo 6. Na Seção 6.1, analisamos a longa dependência em sequências de DNA utilizando diferentes transformações. As sequências de DNA analisadas, utilizando os métodos de estimação propostos na Seção 2, mostram a existência de longa dependência. Mas se consideramos a abordagem de transformação numérica arbitrária para  $\{A, C, T, G\}$  e depois utilizamos a análise espectral, o resultado irá depender de cada transformação em particular. Portanto, para saber se a longa dependência em sequência de DNA é propriedade da sequência, ou se ela é induzida pela transformação, analisamos na Seção 6.2, a longa dependência em sequência de DNA utilizando a metodologia de mudança de regimes, proposta por Liu (2000). Nesta metodologia, se a duração dos regimes de uma série temporal tem uma distribuição de cauda pesada com parâmetro  $\alpha \in (1, 2)$ , então a série temporal apresenta a característica de longa dependência, e se aplicar qualquer transformação linear que preserve a propriedade de variância finita no processo, a série temporal igualmente preserva a propriedade de longa dependência. Através desta metodologia, as sequências de DNA analisadas, apresentaram a característica de longa dependência comprovando que esta é realmente uma propriedade da sequência. Um problema comum na análise de sequências de DNA é como identificar regiões codantes (*cds*) dispersas por toda a sequência e separadas por regiões não codantes (*ncds*). Distinguindo a distribuição para cada uma destas regiões pode ser uma ajuda para encontrar regiões codantes e não codantes em sequências de DNA. Então, analisamos na Seção 6.3, as distribuições de distâncias das regiões codantes e não codantes em sequências de DNA e concluímos que se  $\alpha < 2$  podemos suspeitar, que a distribuição subjacente é a de uma região não codante, não valendo o contrário.

Por fim, concluímos que todas as técnicas apresentadas neste trabalho, com o objetivo principal de analisar a longa dependência em séries temporais, envolvendo conceitos de análise de flutuações destendenciadas, distribuições com caudas pesadas e encolhimento de wavelets, mostram a existência de longa dependência em todas sequências de DNA aqui estudadas.

Como futuros trabalhos, pretendemos nos aprofundar nas questões teóricas e aplicadas das técnicas e métodos aqui estudados.

# Referências

- Almeida, J.S.; J.A. Carriço, A. Marezek, P.A. Noble e M. Fletcher (2001). “Analysis of Genomic Sequences by Chaos Game Representation”. *Bioinformatics*, Vol. **17**(5), 429-437.
- Averkamp, R. e C. Houdré (1998). A note on the Discrete Wavelet Transform of Second-Order Processes. Preprint at School of Mathematics, Georgia Institute of Technology, Atlanta.
- Bardet, J.M. e I. Kammoun (2008). “Asymptotic Properties of the Detrended Fluctuation Analysis of Long-Range Dependence Processes”. *IEEE Transactions on Information Theory*, Vol. **54**(5), 2041-2052.
- Ben-Avraham, D. e S. Havlin (2000). *Diffusion and Reactions in Fractals and Disordered Systems*. Cambridge: Cambridge University Press.
- Beran, J. (1994). *Statistics for Long Memory Processes*. New York: Chapman & Hall.
- Besbeas, P. e B.J.T. Morgan (2008). “Improved Estimation of Stable Laws”. *Statistics and Computing*, Vol. **18**, 219-231.
- Bickel, P.J. e K.A. Doksum (1977). *Mathematical Statistics: Basic Ideas and Selected Topics*. San Francisco: Holden-Day.
- Billingsley, P. (1995). *Probability and Measure*. New York: John Wiley and Sons.
- Boggess, A. e F.J. Narcowich (2001). *A First Course in Wavelets with Fourier Analysis*. Texas: Prentice Hall.
- Box, G.E.P.; G.M. Jenkins e G.C. Reinsel (1994). *Time Series Analysis: Forecasting and Control*. New Jersey: Prentice Hall.
- Brockwell, P.J. e R.A. Davis (1991). *Time Series: Theory and Methods*. New York: Springer-Verlag.

- Buldyrev, S.V.; N.V. Dokholyan, A.L. Goldberger, S. Havlin, C.-K. Peng, H.E. Stanley e G.M. Viswanathan (1998). "Analysis of DNA Sequences Using Methods of Statistical Physics". *Physica A*, Vol. **249**, 430-438.
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer: London.
- Cook, R. D. e S. Weisberg (1983). *Residuals and Influence in Regression*. New York: Chapman & Hall.
- Crato, N., R.R. Linhares e S.R.C. Lopes (2010). "Statistical Properties of Detrended Fluctuation Analysis". *Journal of Statistical Computation and Simulation*, Vol. **80**(6), 625-641.
- Crato, N., R.R. Linhares e S.R.C. Lopes (2011). " $\alpha$ -Stable Laws for Noncoding Regions in DNA Sequences". *Journal of Applied Statistics*, Vol. **38**(2), 261-271.
- Cristea, P.D. (2003). "Large Scale Features in DNA Genomic Signals". *Journal on Applied Signal Processing*, Vol. **83**, 871-888.
- Daubechies. I. (1992). *Ten Lectures on Wavelets*. Number 61 in CBMS-NSF Series in Applied Mathematics. Philadelphia: SIAM.
- Donoho, D.L. e I.M. Johnstone (1994). "Ideal Spatial Adaptation via Wavelet Shrinkage". *Biometrika*, Vol. **81**, 425-455.
- Donoho, D. e I.M. Johnstone (1995). "Adapting to Unknown Smoothness via Wavelet Shrinkage". *J. Amer. Statist. Assoc.*, Vol. **90**, 1200-1224.
- Donoho, D.; I.M., Johnstone; Kerkyacharian, G. e Picard, D. (1995). "Wavelet Shrinkage: Asymptopia?". *Journal of the Royal Statistical Society*, Vol **57**, 301-369.
- Embrechts, P.; C. Klüppelberg e T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer-Verlag.
- Fama, E. e R. Roll (1971). "Parameter Estimates for Symetric Stable Distributions". *Journal of the American Statistical Association*, Vol. **66**(334), 331-338.
- Fisher, R.A. e L.H. Tippett (1928). *Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample*. Proc. Cambridge Philos. Soc., Vol. **24**, 180-190.

- Fox, R. e M.S. Taqqu (1986). “Large-sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series”. *The Annals of Statistics*, Vol. **14**, 517-532.
- Garcia, J.A. e M.V. José (2005). “Mathematical Properties of DNA Sequences from Coding and Noncoding Regions”. *Revista Mexicana de Fisica*, Vol. **51**(2), 122-130.
- Geweke, J. e S. Porter-Hudak (1983). “The Estimation and Application of Long Memory Time Series Model”. *Journal of Time Series Analysis*, Vol. **4**(4), 221-238.
- Gnedenko, B.V. (1943). “Sur la Distribution Limite du Terme Maximum d’une Serie Aleatoire”. *Annals of Mathematics*, Vol. **44**, 423-453.
- Grossmann. A. e J. Morlet (1985). *Decomposition of Functions into Wavelets of Constant Shape and Related Transforms*. In L. Streit, editor, Mathematics and Physics, Lectures on Recent Results. World Scientific, River Edge, NJ.
- Guharay, S.; B.R. Hunt, J.A. Yorke e O.R. White (2000). “Correlations in DNA Sequences Across the Three Domains of Life”. *Physica D: Non-linear Phenomena*, Vol. **146**(1-4), 388-396.
- Hamilton, J.D. (1989). “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle”. *Econometrica*, Vol. **57**, 357-384.
- Hill, B.M. (1975). “A Simple General Approach to Inference about the Tail of a Distribution”. *The Annals of Statistics*, Vol. **3**, 1163-1174.
- Hu, K.; P.C. Ivanov, Z. Chen, P. Carpena e H.E. Stanley (2001). “Effect of Trends on Detrended Fluctuation Analysis”. *Physical Review E*, Vol. **64**, 011114.
- Huber, P.J. (1973). “Robust Regression: Asymptotics, Conjectures and Monte Carlo”. *The Annals of Statistics*, Vol. **1**, 799-821.
- Hurst, H.R. (1951). “Long-Term Storage in Reservoirs”. *Trans. Am. Soc. Civil Eng.*, Vol. **116**, 770-799.
- Hurst, H.E.; R.P. Black e Y.M. Simaika (1965). *Long-Term Storage: An Experimental Study*. London: Constable.
- Kantelhardt, J. W.; E. Koscielny-Bunde, H.H.A. Rego, S. Havlin e A. Bunde (2001). “Detecting long-range correlations with detrended fluctuation analysis”. *Physica A: Statistical Mechanics and its Applications*, Vol. **295**, 441-454.



- Kelly, S.; M. Kon e L. Raphael (1994). “Local Convergence for Wavelet Expansions”. *J. Funct. Anal.*, Vol. **126**, 102-138.
- Koscielny-Bunde, E; H.E. Roman, A. Bunde, S. Havlin e H.-J. Schellnhuber (1998). “Long-Range Power-Law Correlations in Local Daily Temperature Fluctuations”. *Phil. Mag. B*, Vol. **77**, 1331.
- Koutrouvelis, I.A. (1981). “An Iterative Procedure for the Estimation of the Parameters of Stable Laws”. *Communication in Statistics Simulation an Computation*, Vol. **10**(1), 17-28.
- Linhares, R.R. (2007). *Propriedades Estatísticas do Método da Análise de Flutuações Destendenciadas em Seqüências de DNA*. Dissertação de Mestrado, Programa de Pós-Graduação em Matemática, Instituto de Matemática, UFRGS, Porto Alegre.
- Liu, M. (2000). “Modeling Long Memory in Stock Marcket Volatility”. *Journal of Econometrics*, Vol. **99**, 139-171.
- Liu, Y.H.; P. Cizeau, M. Meyer, C.-K. Peng e H.E. Stanley (1997). “Correlations in Economic Time Series”. *Physica A*, Vol. **245**, 437-440.
- Lopes, S.R.C. e B.V.M. Mendes (2006). “Bandwidth Selection in Classical and Robust Estimation of Long Memory”. *International Journal of Statistics and Systems*, Vol. **1**(2), 167-190.
- Lopes, S.R.C. e M.A. Nunes (2006). “Long Memory Analysis in DNA Sequences”. *Physica A: Statistical Mechanics and its Applications*, Vol. **361**(2), 569-588.
- Mallat, S. “A Theory of Multiresolution Signal Decomposition: The Wavelet Representation.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. **11**(7), 674-693.
- McCulloch, J.H. (1986). “Simple Consistet Estimators of Stable Distribution Parameters”. *Communications in Statistics: Simulation and Computation*, Vol. **15**, 1109-1136.
- McCulloch, J.H. (1997). “Measuring Tail Thickness to Estimate the Stable Index Alpha”. *Journal of Business and Economic Statistics*, Vol. **15**, 74-81.
- Meyer, Y. (1993). *Wavelets: Algorithms and Applications*. S.I.A.M, 3600 University City Science Center, Philadelphia, 19104-2688.
- Morettin, P.A. (1999). *Ondas e Ondaletas, da Análise de Fourier à Análise de Ondaletas*. Edusp.

- Morlet, J.; G. Arens, I. Fourgeau e D. Giard (1982). “Wave Propagation and Sampling Theory”. *Geophysics*, Vol. **47**, 203-236.
- Nolan, J.P. (1997). “Numerical Computation of Stable Densities and Distribution Functions”. *Communications in Statistics: Stochastic Models*, Vol. **13**, 759-774.
- Nolan, J.P. (2001). “Maximum Likelihood Estimation of Stable Parameters”. In *O.E. Barndorff-Nielsen, T. Mikosch, and S.I. Resnick (Eds.), Lévy Processes: Theory and Applications*. Boston: Birkhäuser.
- Nolan, J.P. (2009). “Stable Distributions: Models for Heavy Tailed Data”. Disponível na página <http://academic2.american.edu/~jpnolan/stable/chap1.pdf>.
- Ogden. R.T. (1997). *Essential Wavelets for Statistical Applications and Data Analysis*. Cambridge: Birkhäuser.
- Peng, C.; S.V. Buldyrev, A.L. Goldberger, S. Havlin, F. Sciortino, M. Simons e H.E. Stanley (1992). “Long-range Correlations in Nucleotide Sequences”. *Nature*, Vol. **356**, 168-170.
- Peng, C.; S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley e A.L. Goldberger (1994). “Mosaic Organization of DNA Nucleotides”. *Physical Review E*, Vol.**49**(5), 1685-1689.
- Podobnik, B.; J. Shaoc, N.V. Dokholyand, V. Zlatice, H.E. Stanley e I. Grossef (2007). “Similarity and Dissimilarity in Correlations of Genomic DNA”. *Physica A*, Vol. **373**, 497-502.
- Press, S.J. (1972). “Estimation in Univariate and Multivariate Stable Distributions”. *JASA*, Vol **67**, 842-846.
- Resnick, S. e C. Stărică (1997). “Smooth The Hill Estimator”. *Applied Probability Trust*, Vol. **29**, 271-293.
- Rohatgi, V. K. (1976). *Introduction to probability theory and mathematical statistics*. New York: Wiley.
- Stanley, H.E.; S.V. Buldyrev, A.L. Goldberger, S. Havlin, C.K. Peng e M. Simons (1999). “Scaling Features of Concoding DNA”. *Physica A: Statistical Mechanics and its Applications*, Vol. **273**(1), 1-18.
- Stoffer, D.S.; D.E. Tyler e A.J. McDougall (1993). “Spectral Analysis for Categorical Time Series: Scaling and the Spectral Envelope”. *Biometrika*, Vol. **80**, 611-622.

Stoffer, D.S. e O. Rosen (2007). “Automatic Estimation of Multivariate Spectra via Smoothing Splines”. *Biometrika*, Vol. **94**, 335-345.

Taqqu, M.S.; V. Teverovsky e W. Willinger (1995). “Estimators for Long Range Dependence: An Empirical Study”. *Fractals*, Vol. **3**(4), 785-798.

Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. New York: Wiley.

Walter, G.G. (1994). *Wavelets and Other Orthogonal Systems with Applications*. CRC Press Inc., Boca Raton, FL.

Whittle, P. (1953). *Hypothesis Testing in Time Series Analysis*. New York: Hafner.

Zolotarev, V. (1986). *One-Dimensional Stable Distributions*. Providence, RI: American Mathematical Society.

# Anexo A

Neste anexo, apresentamos algumas propriedades adicionais para o método DFA utilizando a função densidade da variável aleatória  $Y = \ln(F(l))$ , onde  $F(l)$  é dada pela expressão (3.3). Sob certas condições, calculamos a função densidade da variável aleatória  $Y = \ln(F(l))$  e mostramos que esta função forma uma família exponencial. Por fim, calculamos a variância do estimador  $\hat{H}$ , dado na expressão (3.8), utilizando a informação de Fisher (ver Definição A.2).

**Teorema A.1.** *Suponha que, para todo  $l \in \{4, 5, \dots, g(n)\}$ , as variáveis aleatórias  $Z_1^l, Z_2^l, \dots, Z_{\tilde{n}}^l$ , dadas pela expressão (3.2), sejam independentes e identicamente distribuídas com função de distribuição comum  $\mathcal{N}(0, \sigma_l^2)$ . Então, para todo  $l \in \{4, 5, \dots, g(n)\}$ , a variável aleatória  $Y = \ln(F(l))$  tem função densidade dada por*

$$f_Y(y) = 2 \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \exp\{2\alpha y - \beta e^{2y}\}, \quad \text{para todo } y \in \mathbb{R}, \quad (\text{A.1})$$

onde  $\alpha = \frac{\tilde{n}}{2}$  e  $\beta = \frac{\tilde{n}}{2\sigma_l^2}$ .

**Demonstração:** Seja  $X = F^2(l)$  e  $Y = \ln(F(l))$ . Definindo  $G_0 = (0, \infty)$ ,  $g(x) = \frac{1}{2} \ln(x)$  e  $G = \mathbb{R}$ , vemos que as condições do teorema do Método do Jacobiano (ver Billingsley, 1995) estão satisfeitas. Assim,

$$f_Y(y) = f_X(g^{-1}(y)) |(g^{-1})'(y)|, \quad \text{para todo } y \in G. \quad (\text{A.2})$$

Como  $g^{-1}(y) = e^{2y}$ ,  $(g^{-1})'(y) = 2e^{2y}$  e  $F^2(l)$  tem distribuição  $\Gamma(\frac{\tilde{n}}{2}, \frac{\tilde{n}}{2\sigma_l^2})$  (ver Definição 2.5), então a função densidade de  $Y$  é dada por

$$\begin{aligned} f_Y(y) &= 2 \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} e^{2(\alpha-1)y} e^{-\beta e^{2y}} e^{2y} \\ &= 2 \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \exp\{2\alpha y - \beta e^{2y}\}, \quad \text{para todo } y \in \mathbb{R}, \end{aligned} \quad (\text{A.3})$$

onde  $\alpha = \frac{\tilde{n}}{2}$  e  $\beta = \frac{\tilde{n}}{2\sigma_l^2}$ . □

**Observação A.1.** Apresentamos as expressões de duas integrais que serão necessárias na prova do Corolário A.1 a seguir.

$$(i) \int_0^{\infty} \ln(v)v^{\alpha-1}e^{-v}dv = \Gamma'(\alpha),$$

$$(ii) \int_0^{\infty} [\ln(v)]^2v^{\alpha-1}e^{-\beta v}dv = \frac{\Gamma(\alpha)}{\beta^\alpha}((\psi(\alpha) - \ln \beta)^2 + \zeta(2, \alpha)),$$

onde  $\Gamma(\cdot)$  é a função matemática Gama e  $\Gamma'(\cdot)$  sua derivada (ver Definição 2.1),  $\psi(\cdot)$  é a função digama (ver Definição 2.2) e  $\zeta(\cdot, \cdot)$  é a função zeta de Hurwitz (ver Definição 2.3). Para mais detalhes, ver Gradshteyn e Ryzhik (2000), página 572.

**Corolário A.1.** *Suponha que, para todo  $l \in \{4, 5, \dots, g(n)\}$ , as variáveis aleatórias  $Z_1^l, Z_2^l, \dots, Z_{\tilde{n}}^l$ , dadas pela expressão (3.2), sejam independentes e identicamente distribuídas com função de distribuição comum  $\mathcal{N}(0, \sigma_l^2)$ . Então, para todo  $l \in \{4, 5, \dots, g(n)\}$ ,*

$$\mathbb{E}[\ln(F(l))] = \frac{1}{2} \left[ \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \ln(\beta) \right],$$

$$\mathbb{E}[(\ln(F(l)))^2] = \frac{1}{4} \left[ \left( \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \ln(\beta) \right)^2 + \zeta(2, \alpha) \right],$$

e

$$\text{Var}(\ln(F(l))) = \frac{1}{4} \zeta \left( 2, \frac{\tilde{n}}{2} \right),$$

onde  $\alpha = \frac{\tilde{n}}{2}$ ,  $\beta = \frac{\tilde{n}}{2\sigma_l^2}$ ,  $\tilde{n} = l \cdot \lfloor \frac{n}{l} \rfloor$ ,  $\Gamma(\cdot)$  é a função matemática Gama e  $\Gamma'(\cdot)$  sua derivada (ver Definição 2.1), e  $\zeta(\cdot, \cdot)$  é a função zeta de Hurwitz (ver Definição 2.3).

**Demonstração:** Seja  $Y = \ln(F(l))$ . Então,

$$\begin{aligned} \mathbb{E}(Y) &= \int_{-\infty}^{\infty} y \left( 2 \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \exp\{2\alpha y - \beta e^{2y}\} \right) dy \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_{-\infty}^{\infty} 2y \cdot \exp\{2\alpha y - \beta e^{2y}\} dy. \end{aligned} \quad (\text{A.4})$$

Pelo Teorema de mudança de variável obtemos que

$$\mathbb{E}(Y) = \frac{1}{2} \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} \ln(u) u^{\alpha-1} e^{-\beta u} du. \quad (\text{A.5})$$

Fazendo  $v = \beta u$  e  $dv = \beta du$ , pela expressão (2.1) e pelo item (i) da Observação A.1, tem-se que

$$\begin{aligned}
\mathbb{E}(Y) &= \frac{1}{2} \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \ln\left(\frac{v}{\beta}\right) \left(\frac{v}{\beta}\right)^{\alpha-1} e^{-v} \frac{dv}{\beta} \\
&= \frac{1}{2} \frac{1}{\Gamma(\alpha)} \int_0^\infty [\ln(v) - \ln(\beta)] v^{\alpha-1} e^{-v} dv \\
&= \frac{1}{2} \frac{1}{\Gamma(\alpha)} \int_0^\infty \ln(v) v^{\alpha-1} e^{-v} dv - \frac{1}{2} \frac{\ln(\beta)}{\Gamma(\alpha)} \int_0^\infty v^{\alpha-1} e^{-v} dv \\
&= \frac{1}{2} \left[ \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \ln(\beta) \right] = \frac{1}{2} \left[ \frac{\Gamma'(\frac{\tilde{n}}{2})}{\Gamma(\frac{\tilde{n}}{2})} - \ln\left(\frac{\tilde{n}}{2\sigma_l^2}\right) \right]. \tag{A.6}
\end{aligned}$$

Observe que

$$\begin{aligned}
\mathbb{E}(Y^2) &= \int_{-\infty}^\infty y^2 \left( 2 \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \exp\{2\alpha y - \beta e^{2y}\} \right) dy \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_{-\infty}^\infty 2y^2 \cdot \exp\{2\alpha y - \beta e^{2y}\} dy. \tag{A.7}
\end{aligned}$$

Fazendo  $v = e^{2y}$ ,  $dv = 2e^{2y} dy$  e  $y = \frac{\ln(v)}{2}$ , pelo item (ii) da Observação A.1, obtemos que

$$\begin{aligned}
\mathbb{E}(Y^2) &= \frac{1}{4} \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty (\ln(v))^2 v^{\alpha-1} e^{-\beta v} dv \\
&= \frac{1}{4} \frac{\beta^\alpha}{\Gamma(\alpha)} \left[ \frac{\Gamma(\alpha)}{\beta^\alpha} \left( (\psi(\alpha) - \ln \beta)^2 + \zeta(2, \alpha) \right) \right] \\
&= \frac{1}{4} [(\psi(\alpha) - \ln \beta)^2 + \zeta(2, \alpha)] \\
&= \frac{1}{4} \left[ \left( \psi\left(\frac{\tilde{n}}{2}\right) - \ln\left(\frac{\tilde{n}}{2\sigma_l^2}\right) \right)^2 + \zeta\left(2, \frac{\tilde{n}}{2}\right) \right] \\
&= \frac{1}{4} \left[ \left( \frac{\Gamma'(\frac{\tilde{n}}{2})}{\Gamma(\frac{\tilde{n}}{2})} - \ln\left(\frac{\tilde{n}}{2\sigma_l^2}\right) \right)^2 + \zeta\left(2, \frac{\tilde{n}}{2}\right) \right]. \tag{A.8}
\end{aligned}$$

Segue-se então que

$$\begin{aligned}
\text{Var}(Y) &= \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 \\
&= \frac{1}{4} \left[ \left( \frac{\Gamma'(\tilde{n}/2)}{\Gamma(\tilde{n}/2)} - \ln\left(\frac{\tilde{n}}{2\sigma_l^2}\right) \right)^2 + \zeta\left(2, \frac{\tilde{n}}{2}\right) \right] \\
&\quad - \left[ \frac{1}{2} \left( \frac{\Gamma'(\frac{\tilde{n}}{2})}{\Gamma(\frac{\tilde{n}}{2})} - \ln\left(\frac{\tilde{n}}{2\sigma_l^2}\right) \right) \right]^2 \\
&= \frac{1}{4} \zeta\left(2, \frac{\tilde{n}}{2}\right). \tag{A.9}
\end{aligned}$$

□

**Teorema A.2.** *Suponha que, para todo  $l \in \{4, 5, \dots, g(n)\}$ , as variáveis aleatórias  $Z_1^l, Z_2^l, \dots, Z_{\tilde{n}}^l$ , dadas pela expressão (3.2), sejam independentes e identicamente distribuídas com função de distribuição comum  $\mathcal{N}(0, \sigma_l^2)$ . Seja  $\epsilon = (\epsilon_1, \dots, \epsilon_m)'$  o vetor de erros no modelo de regressão linear dado pela expressão (3.6). Assuma que  $\epsilon_j$ , para todo  $j \in \{1, \dots, m\}$ , é variável aleatória independente e identicamente distribuída com função de distribuição  $\mathcal{N}(0, \sigma_\epsilon^2)$ . Então,  $\hat{H}$ , dado pela expressão (3.8), é um estimador U.M.V.U. para o parâmetro  $H$ , com esperança e variância dadas, respectivamente, por*

$$\mathbb{E}(\hat{H}) = H \quad e \quad \text{Var}(\hat{H}) = \frac{\zeta\left(2, \frac{\tilde{n}}{2}\right)}{4 \sum_{j=1}^m (x_j - \bar{x})^2},$$

onde  $\zeta(\cdot, \cdot)$  é a função zeta de Hurwitz (ver Definição 2.3).

**Demonstração:** Utilizando resultados do Teorema 3.6 e do Corolário A.1 tem-se que

$$\text{Var}(\hat{H}) = \frac{\text{Var}(\ln(F(l)))}{\sum_{j=1}^m (x_j - \bar{x})^2} = \frac{\zeta\left(2, \frac{\tilde{n}}{2}\right)}{4 \sum_{j=1}^m (x_j - \bar{x})^2},$$

onde  $\tilde{n} = l \cdot \lfloor \frac{n}{l} \rfloor$ . □

**Definição A.1 (Família Exponencial de  $k$  Parâmetros).** Uma família de distribuições de um modelo  $\{\mathbb{P}_\theta : \theta \in \Theta\}$  é dita ser uma *família exponencial de  $k$  parâmetros*, se existem funções reais  $c_1(\cdot), \dots, c_k(\cdot)$  e  $d(\cdot)$  sobre  $\Theta \subseteq \mathbb{R}^k$ , funções reais  $T_1(\cdot), \dots, T_k(\cdot)$  e  $S(\cdot)$  sobre  $\mathbb{R}^m$ , e um conjunto  $A \in \mathbb{R}^m$  tal que a função densidade (ou função de probabilidade) é dada por

$$f(\mathbf{x}, \theta) = \exp\left\{ \sum_{j=1}^k c_j(\theta) T_j(\mathbf{x}) + d(\theta) + S(\mathbf{x}) \right\} I_A(\mathbf{x}), \quad (\text{A.10})$$

onde  $\mathbf{x} = (x_1, \dots, x_m)'$  e  $I_A(\cdot)$  é a função indicadora do conjunto  $A$ .

**Observação A.2.** Na Definição A.1, as funções  $c_j(\cdot)$ ,  $d(\cdot)$ ,  $S(\cdot)$  e  $T_j(\cdot)$ , para todo  $j \in \{1, \dots, k\}$ , não são únicas e o conjunto  $A$  não pode depender de  $\theta$ .

**Exemplo A.1.** Seja  $\{\mathbb{P}_\theta : \theta \in \Theta\}$  a família das distribuições com função densidade dada

$$f_Y(y, \theta) = 2 \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \exp\{2\alpha y - \beta e^{2y}\}, \quad \text{para todo } y \in \mathbb{R}, \quad (\text{A.11})$$

onde  $\Theta = \{(\alpha, \beta) : \alpha > 0, \beta > 0\}$ . Logo, para todo  $y \in \mathbb{R}$ , tem-se que

$$\begin{aligned}
f_Y(y, \theta) &= 2 \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \exp\{2\alpha y - \beta e^{2y}\} \\
&= \exp\left\{ \ln\left[ \frac{2\beta^\alpha}{\Gamma(\alpha)} e^{2\alpha y - \beta e^{2y}} \right] \right\} \\
&= \exp\left[ \ln\left( \frac{2\beta^\alpha}{\Gamma(\alpha)} \right) + \ln\left( e^{2\alpha y - \beta e^{2y}} \right) \right] \\
&= \exp\left[ \ln\left( \frac{2\beta^\alpha}{\Gamma(\alpha)} \right) + 2\alpha y - \beta e^{2y} \right]. \tag{A.12}
\end{aligned}$$

(i) Se  $\alpha$  é desconhecido e  $\beta$  é conhecido, então pela Definição A.1,  $\mathbb{P}_\theta$  forma uma família exponencial de um parâmetro com

$$c(\alpha) = 2\alpha, \quad d(\alpha) = \ln\left[ \frac{2\beta^\alpha}{\Gamma(\alpha)} \right], \quad T(y) = y, \quad S(y) = -\beta_0 e^{2y}, \quad A = \mathbb{R}.$$

(ii) Se  $\alpha$  é conhecido e  $\beta$  é desconhecido, então pela Definição A.1,  $\mathbb{P}_\theta$  forma uma família exponencial de um parâmetro com

$$c(\beta) = -\beta, \quad d(\beta) = \ln\left[ \frac{2\beta^\alpha}{\Gamma(\alpha)} \right], \quad T(y) = e^{2y}, \quad S(y) = 2\alpha y, \quad A = \mathbb{R}.$$

(iii) Se  $\alpha$  e  $\beta$  são desconhecidos, então pela Definição A.1,  $\mathbb{P}_\theta$  forma uma família exponencial de dois parâmetros com

$$\begin{aligned}
c_1(\theta) = 2\alpha, \quad c_2(\theta) = -\beta, \quad d(\theta) = \ln\left[ \frac{2\beta^\alpha}{\Gamma(\alpha)} \right] \\
T_1(y) = y, \quad T_2(y) = e^{2y}, \quad S(y) = 0, \quad A = \mathbb{R}.
\end{aligned}$$

**Observação A.3.** Considere as hipóteses de regularidade para uma família  $\{\mathbb{P}_\theta : \theta \in \Theta\}$  dadas por

**H1)** O conjunto  $A = \{\mathbf{x} \in \mathbb{R}^m : f(\mathbf{x}, \theta) > 0\} \subseteq \mathbb{R}^m$  não depende de  $\theta$ . Para todo  $\mathbf{x} \in A$ ,  $\theta \in \Theta$ ,  $\frac{\partial}{\partial \theta} \ln(f(\mathbf{x}, \theta))$  existe e é finita.

**H2)** Se  $T$  é uma estatística tal que  $\mathbb{E}_\theta(|T|) < \infty$ , para todo  $\theta \in \Theta$ , então as operações de integração e diferenciação por  $\theta$  podem ser trocadas por  $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} T(\mathbf{x}) f(\mathbf{x}, \theta) dx$ , isto é,

$$\frac{\partial}{\partial \theta} \left[ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} T(\mathbf{x}) f(\mathbf{x}, \theta) dx \right] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} T(\mathbf{x}) \frac{\partial}{\partial \theta} f(\mathbf{x}, \theta) dx, \tag{A.13}$$

sempre que o lado direito de (A.13) for finito.



**Proposição A.1.** Se  $f(\mathbf{x}, \theta) = \exp\{c(\theta)T(\mathbf{x}) + d(\theta) + S(\mathbf{x})\}I_A(\mathbf{x})$  é uma família exponencial de um parâmetro e  $c(\theta)$  tem derivada sobre  $\Theta$  diferente de zero, então valem as hipóteses de regularidade dadas na Observação A.3.

**Demonstração:** Ver Bickel e Doksum (1977), página 127.  $\square$

**Definição A.2 (Informação de Fisher).** Seja  $\{\mathbb{P}_\theta : \theta \in \Theta\}$ , uma família de distribuições com função densidade  $f(\mathbf{x}, \theta)$ . Suponhamos que vale a hipótese H1, dada na Observação A.3, então a *informação de Fisher*, o qual denotamos por  $I(\theta)$ , é definido por

$$I(\theta) = \mathbb{E} \left( \frac{\partial}{\partial \theta} \ln(f(\mathbf{X}, \theta)) \right)^2 = \text{Var} \left( \frac{\partial}{\partial \theta} \ln(f(\mathbf{X}, \theta)) \right).$$

**Teorema A.3 (Desigualdade da Informação de Fisher).** Sejam  $\mathbf{X} = (X_1, \dots, X_m)'$  uma amostra de uma população, com função densidade  $f(\mathbf{x}, \theta)$ ,  $\theta \in \Theta$  e  $T(\mathbf{X})$  uma estatística tal que  $\text{Var}_\theta(T(\mathbf{X})) < \infty$ , para todo  $\theta \in \Theta$ . Denotamos  $\mathbb{E}_\theta(T(\mathbf{X}))$  por  $\xi(\theta)$ . Suponhamos que valem as hipóteses H1 e H2 dadas na Observação A.3 e  $0 < I(\theta) < \infty$ . Então,

$$\text{Var}_\theta(T(\mathbf{X})) \geq \frac{[\xi'(\theta)]^2}{I(\theta)},$$

onde  $I(\theta)$  é o número da informação de Fisher (ver Definição A.2).

**Demonstração:** Ver Bickel e Doksum (1977), página 129.  $\square$

**Teorema A.4.** Se  $\{\mathbb{P}_\theta : \theta \in \Theta\}$  é uma família exponencial de um parâmetro dada na Definição A.1, onde  $c(\theta)$  tem derivada contínua e diferente de zero, então  $T(\mathbf{X})$  alcança o limite da desigualdade de informação.

**Demonstração:** Ver Bickel e Doksum (1977), página 130.  $\square$

**Exemplo A.2.** Seja  $\{\mathbb{P}_\theta : \theta \in \Theta\}$  a família das distribuições com função densidade dada pela expressão (A.11). No item (i) do Exemplo A.1, vimos que, se  $\alpha$  é desconhecido e  $\beta$  é conhecido, então  $\mathbb{P}_\theta$  forma uma família exponencial de um parâmetro com

$$c(\alpha) = 2\alpha, \quad d(\alpha) = \ln \left[ \frac{2\beta^\alpha}{\gamma(\alpha)} \right], \quad T(y) = y, \quad S(y) = -\beta_0 e^{2y}, \quad A = \mathbb{R}.$$

Note que a derivada de  $c(\cdot)$  em relação a  $\alpha$ , é constante e diferente de zero. Logo, pela Proposição A.1, valem as hipóteses H1 e H2, dadas na Observação A.3. Portanto, pelo Teorema A.4,

$$\text{Var}(Y) = \frac{[\xi'(\alpha)]^2}{I(\alpha)}, \quad (\text{A.14})$$

onde  $\xi(\alpha) = \mathbb{E}(Y)$  e  $I(\alpha)$  é o número da informação de Fisher. Pelo Corolário A.1, temos que,

$$\xi(\alpha) = \mathbb{E}(Y) = \frac{1}{2} \left[ \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \ln(\beta) \right] = \frac{1}{2} [\psi(\alpha) - \ln(\beta)],$$

segue-se que,

$$\xi'(\alpha) = \frac{1}{2} \psi'(\alpha) = \frac{1}{2} \zeta(2, \alpha). \quad (\text{A.15})$$

Calculando o número da informação de Fisher,

$$\begin{aligned} I(\alpha) &= \text{Var} \left( \frac{\partial}{\partial \alpha} \ln(f(Y, \alpha)) \right) \\ &= \text{Var} \left[ \frac{\partial}{\partial \alpha} \left( \ln \left( \frac{2\beta^\alpha}{\Gamma(\alpha)} \right) + 2\alpha Y - \beta e^{2y} \right) \right] \\ &= \text{Var} \left( \ln(\beta) - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + 2Y \right) \\ &= 4 \cdot \text{Var}(Y). \end{aligned} \quad (\text{A.16})$$

Pelas expressões (A.14), (A.15) e (A.16), obtemos que

$$\begin{aligned} \text{Var}(Y) &= \frac{[\frac{1}{2}\zeta(2, \alpha)]^2}{4 \cdot \text{Var}(Y)} \\ \Leftrightarrow 4 \cdot (\text{Var}(Y))^2 &= \left[ \frac{1}{2}\zeta(2, \alpha) \right]^2 \Leftrightarrow 2 \cdot \text{Var}(Y) = \frac{1}{2}[\zeta(2, \alpha)] \\ \Leftrightarrow \text{Var}(Y) &= \frac{1}{4}[\zeta(2, \alpha)]. \end{aligned} \quad (\text{A.17})$$

Considere as hipóteses dadas no Teorema A.2. Então, pelo Teorema 3.6 e pela expressão (A.17), segue-se que

$$\text{Var}(\widehat{H}) = \frac{\text{Var}(Y)}{\sum_{j=1}^m (x_j - \bar{x})^2} = \frac{\zeta(2, \alpha)}{4 \sum_{j=1}^m (x_j - \bar{x})^2}.$$