Universidade Federal do Rio Grande do Sul

Instituto de Matemática

Programa de Pós-Graduação em Matemática

Estimação em Processos Advindos da Solução da Equação de Langevin

Dissertação de Mestrado

Douglas Rodrigues Pinto

Porto Alegre, 27 de setembro de 2010

Dissertação submetida por Douglas Rodrigues Pinto [*] , como requisito parcial para a obtenção do grau de Mestre em Ciência Matemática, pelo Programa de Pós-Graduação em Matemática, do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professor Orientador:
Prof. Dra. Sílvia Regina Costa Lopes

Banca Examinadora:
Prof. Dr. Artur Oscar Lopes
Prof. Dr. Ary Vasconcelos Medino
Prof. Dr. Rafael Rigão Souza

[*]Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

DEDICATÓRIA

À vida, pelas oportunidades que vem me concedendo e por me permitir chegar tão longe.

Agradecimentos

Quero agradecer àqueles que sempre me incentivaram e nunca me deixaram desistir: minha família. Primeiramente à minha mãe, que foi minha primeira professora; minhas irmãs Karen e Francine, que apesar da distância nunca me abandonaram; ao Renato, que muito colaborou na minha criação. Um agradecimento especial para a Karla Backes, que sempre procurou me alegrar nos momentos mais difíceis do mestrado e nunca permitiu que eu perdesse o foco. Um grande agradecimento aos meus amigos Fagner, Regiton, Patrícia Guidolin, Thaísa Tamusiunas, Diego Leban, Diego Marcon, Nicolau, Avelino, que tive o imenso prazer de conhecer na faculdade, mas que levarei para toda a vida. Um abraço especial ao meu colega de quarto Rodrigo Lima, que sempre colaborou para que eu pudesse estudar. Agradecimentos especiais para as doutorandas Taiane e Raquel, que deram valiosas sugestões na elaboração deste texto.

Academicamente, agradeço à professora Cydara Ripoll, pelo incentivo dado desde os primeiros semestres da graduação; aos professores Eduardo Brietzke, Alexandre Baraviera, Artur Lopes e Leonardo Bonorino, por sempre acreditarem no meu potencial. Um agradecimento especial à minha orientadora Sílvia Regina Costa Lopes, não só pelos ensinamentos matemáticos, mas por ter sido praticamente uma mãe durante esse último ano de mestrado.

Resumo

O presente trabalho apresenta os resultados das estimações realizadas nos processos advindos da solução da equação de Langevin clássica, com ruído movimento Browniano, e generalizada, cujo ruído possui distribuição α-estável. Realizamos a estimação do parâmetro λ vinculado à função memória presente em tais processos e do parâmetro α nos casos em que o ruído é um processo com distribuição α-estável. A estimação de λ é obtida através do método de máxima verossimilhança, do método dos mínimos quadrados e da função de autocovariância amostral do processo. Discutimos os tipos de ruídos que podem ser associados à Equação de Langevin e apresentamos o Teorema de Existência e Unicidade de Kannan e a Fórmula de Representação de Kannan, que provam a existência e unicidade da solução para a equação de Langevin generalizada e apresenta critérios para a forma de tal solução.

Abstract

This paper presents the results of the estimations made in the process arising from the solution of classical Langevin equation with Brownian motion noise, and generalized, with α-stable noise. We perform the estimation of the parameter λ related to the memory function of such processes and the parameter α when the noise is a process with α-stable distribution. The estimation of λ is obtained from the maximum likelihood method, the ordinary least squares method and from the sample autocovariance function of the process. We discuss the types of noise that may be associated to the Langevin equation and present the Existence and Uniqueness theorem of Kannan and Kannan representation formula, which prove the existence and uniqueness of the solution for the generalized Langevin equation

and presents criteria for the form of such solution.

Índice

1 Introdução **1**

2 Conceitos Preliminares **4**

2.1 Processos Estocásticos e σ-álgebra 5

2.2 Teoria Espectral de Processos Estacionários 13

2.3 Processos Estocásticos com Longa Dependência 17

2.4 Movimento Browniano 18

2.5 Processos de Lévy 20

2.6 Distribuições Estáveis 21

2.7 Auto-Similaridade 25

2.8 Movimento Browniano Fracionário 28

2.9 Martingales 32

2.10 Integral de Riemann-Stieltjes 36

2.11 Integral de Itô 39

2.11.1 Integral de Itô para Processos Simples 41

2.11.2 Caso Geral da Integral de Itô 44

2.11.3 Lema de Itô 49

3 Equação de Langevin **53**

3.1 Equação de Langevin Clássica e sua Solução 54

3.2 Forma Discreta da Solução da Equação de Langevin Clássica 62

3.3 Equação de Langevin Clássica Fracionária 64

3.4 Equação de Langevin Generalizada e sua Solução 66

4 Estimação **70**

4.1 Estimação de λ na Solução da Equação de Langevin Clássica 70

4.1.1 Estimadores Advindos da Função de Autocorrelação Amostral 70

4.1.2 Estimador de Mínimos Quadrados 71

4.1.3 Estimador de Máxima Verossimilhança 75

4.2 Estimação de α na Solução da Equação de Langevin com Ruído α-estável 77

5 Simulações **80**

5.1 Estimação de λ na Solução da Equação de Langevin Clássica 80

5.2 Estimação de λ na Solução da Equação de Langevin com Ruído α-estável 92

5.3 Estimação de α na Solução da Equação de Langevin com Ruído α-estável 105

6 Aplicação **111**

7 Conclusão **114**

Referências Bibliográficas **116**

Anexo A **122**

A.1 Espaços de Hilbert e Espaços de Banach 122

Capítulo 1

Introdução

O estudo das Equações de Langevin iniciou com o matemático francês Paul Langevin, em 1908, com o objetivo de modelar a dinâmica do movimento irregular e imprevisível de uma partícula imersa em um meio fluido. A equação de Langevin clássica é dada por

$$\begin{cases} m\, dX_t = -\lambda X_t dt + \eta_t \\ X\,(0) = 0, \end{cases} \tag{1.1}$$

onde m representa a massa da partícula, λ é a constante de atrito e $\{\eta_t\}_{t\geq 0}$ é um processo ruído, proveniente da colisão entre a partícula e as moléculas do meio fluído. Neste trabalho, bem como na literatura, adotamos $m = 1$ e, para todo $t \geq 0$, $\eta_t = \sigma dB_t$, onde $\{B_t\}_{t\geq 0}$ é o processo movimento Browniano e σ^2 é conhecido como parâmetro de difusão. Essa foi a primeira equação diferencial estocástica, e seu estudo motivou muito do que se conhece hoje em Probabilidade.

O processo estocástico $\{X_t\}_{t\geq 0}$ solução da equação (1.1) é chamado processo de Ornstein-Uhlenbeck, nome esse dado devido aos físicos teóricos Leonard Ornstein e George Eugene Uhlenbeck.

O estudo dos processos de Ornstein-Uhlenbeck foi reimpulsionado pelo advento de séries temporais financeiras como uma importante ferramenta para modelar taxas de juros, taxas de câmbio, preços de mercadorias, etc.

Surgiram diversas derivações do processo de Ornstein-Uhlenbeck, tais como os modelos propostos por Barndorff-Nielsen e Shephard (2000), que apresentam a aplicação destes processos com ruído dado pelo processo de Lévy para modelar séries temporais nas áreas de finanças e econometria.

Devido a presença de longa dependência, outra variação de grande aplicabilidade é o processo de Ornstein-Uhlenbeck fracionário (ver Önalan, 2010), cujo ruído é o movimento Browniano fracionário.

Uma generalização do modelo proposto por Langevin surge em 1965, com Hazime Mori, e em 1966, com Ryogo Kubo, conhecida como equação de Langevin generalizada, para uma partícula de massa unitária, dada por

$$\begin{cases} V'(t) = -\int_0^t \gamma(t-s)\, V(s)\ ds + \eta(t) \\ V(0) = V_0 \end{cases} \tag{1.2}$$

onde $V = \{V(t)\}_{t\geq 0}$ é um processo estocástico que representa a velocidade da partícula e $\eta = \{\eta(t)\}_{t\geq 0}$ é um processo ruído branco, representando a força aleatória resultante das colisões entre as moléculas e a partícula.

A função $\gamma(\cdot)$, que aparece no interior da integral em (1.2), é chamada de função memória, e representa o efeito de memorização na evolução do movimento da partícula. Essa equação foi estudada via Teoria da Probabilidade por Kannan (1977), considerando-se o ruído $\{\eta_t\}_{t\geq 0}$ como sendo um processo estocástico com segundo momento finito. Trabalhos mais recentes na área realizados por Medino (2005) tem suposto o modelo (1.2) com ruído possuindo o segundo momento infinito (processos de Lévy, por exemplo).

O objetivo desse trabalho é a estimação do parâmetro vinculado à função memória presente nos processos estocásticos obtidos a partir da solução da equação de Langevin generalizada. A estimação é obtida através do método de máxima verossimilhança, do método dos mínimos quadrados e da função de autocovariância amostral do processo. Os estimadores de máxima verossimilhança e o de mínimos quadrados foram propostos por Kutoyants (2004). Neste trabalho comparamos o

estimador de máxima verossimilhança com os dois estimadores propostos por Valdivieso et al. (2009) e que são obtidos através da função de autocorrelação amostral. Desejamos também analisar a existência de longa dependência (ver Lopes, 2008) nestes processos.

Iniciamos este trabalho introduzindo os conceitos necessários ao estudo dos processos estocásticos resultantes da solução da equação de Langevin clássica e generalizada. Apresentamos também diversas famílias de processos estocásticos, como o movimento Browniano, os processos de Lévy, os processos auto-similares e o movimento Browniano fracionário. No final do capítulo apresentamos uma breve introdução à integral de Itô.

O Capítulo 3 trata do estudo dos processsos estocásticos advindos da solução da equação de Langevin, inicialmente no caso clássico, com ruído movimento Browniano e movimento Browniano fracionário. Na Seção 3.4 apresentamos a equação de Langevin generalizada e enunciamos o Teorema de Existência e Unicidade de Kannan e a Fórmula de Representação de Kannan, que apresentam critérios para a obtenção da solução da equação de Langevin generalizada.

No Capítulo 5 apresentamos simulações para a estimação do parâmetro de processos obtidos da solução da equação de Langevin, tanto no caso clássico como no caso generalizado.

No Capítulo 6, estimamos os parâmetros do processo de Ornstein-Uhlenbeck modelado a partir dos dados extraídos das geleiras da Groelândia para estudar a temperatura na Terra ao longo dos séculos.

As conclusões finais referentes a este trabalho são apresentadas no Capítulo 7.

Capítulo 2

Conceitos Preliminares

Neste capítulo, apresentamos noções básicas e preliminares, além dos resultados necessários para a análise de processos estocásticos e equações diferenciais estocásticas.

Ao longo deste trabalho utilizamos a seguinte notação:

1) $\xrightarrow{\mathcal{D}}$ denota convergência em distribuição;

2) $\overset{\mathcal{D}}{=}$ denota igualdade em distribuição;

3) as derivadas parciais são denotadas por

$$f_{i}\left(u,v\right)=\left[\frac{\partial}{\partial x_{i}}f\left(x_{1},x_{2}\right)\right]_{(x_{1},x_{2})=(u,v)},\ i=1,2,$$

$$f_{ij}\left(u,v\right)=\left[\frac{\partial}{\partial x_{i}}\frac{\partial}{\partial x_{j}}f\left(x_{1},x_{2}\right)\right]_{(x_{1},x_{2})=(u,v)},\ i,j=1,2.$$

4) $\mathbb{Z}_{+}$ é o conjunto dos números inteiros não negativos, ou seja, $\mathbb{Z}_{+}=\mathbb{N}\cup\{0\}$.

2.1 Processos Estocásticos e σ-álgebra

Definição 2.1 (σ-álgebra)**.** Dado um espaço amostral Ω, uma σ-álgebra $\mathcal{F}$ em Ω é uma coleção de subconjuntos de Ω satisfazendo as seguintes condições

i) $\Omega \in \mathcal{F}$.

ii) Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$.

iii) Se $A_1, A_2, \cdots \in \mathcal{F}$, então

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}.$$

Exemplo 2.1. Seja $\Omega = X$, onde X é uma conjunto qualquer não vazio. Então, $\mathcal{F} = \{\emptyset, X\}$ é uma σ-álgebra. Basta notarmos que os ítens i) ii) e iii) da Definição 2.1 são trivialmente satisfeitos, já que, em Ω, $X^c = \emptyset$ e $\emptyset^c = X$. Essa é a menor σ-álgebra em Ω contendo X.

Exemplo 2.2. Seja $\Omega = X$, onde X é um conjunto enumerável não vazio. Então, $\mathcal{F} = \mathcal{P}(X)$ é uma σ-álgebra, onde $\mathcal{P}(X)$ é o conjunto das partes de X. Essa é a maior σ-álgebra em Ω contendo X.

Exemplo 2.3. Seja Ω em conjunto qualquer não vazio. Se $\mathcal{F}$ e $\mathcal{G}$ são duas σ-álgebras de subconjuntos de Ω, então $\mathcal{F} \cap \mathcal{G}$ também é uma σ-álgebra.

Portanto, a σ-álgebra de subconjuntos de Ω pode ser um conjunto muito pequeno, como no Exemplo 2.1, ou muito grande, como mostra o Exemplo 2.2. Dado um subconjunto $\mathcal{C} \subset \Omega$, estamos interessados na σ-álgebra que contenha $\mathcal{C}$, mas que seja a menor σ-álgebra possível.

Definição 2.2. Seja $\mathcal{C}$ uma classe não vazia de subconjuntos de Ω. A *σ-álgebra gerada por* $\mathcal{C}$, denotada por $\sigma(\mathcal{C})$, é a menor σ-álgebra que contém $\mathcal{C}$.

Exemplo 2.4. A σ-álgebra gerada pela classe dos subconjuntos abertos de $\mathbb{R}^n$ é chamada σ-álgebra de Borel e seus elementos são chamados de borelianos.

Exemplo 2.5. Extendendo-se a idéia do Exemplo 2.3, pode-se mostrar que se $\{A_i\}_{i\in J}$ é um conjunto de σ-álgebras de subconjuntos de Ω, onde J é um conjunto qualquer de índices, então $\bigcap_{i\in J} A_i$ também é uma σ-álgebra.
Seja Ω um conjunto enumerável e $\mathcal{C}$ uma classe de subconjuntos não vazia de Ω. Queremos obter a menor σ-álgebra contendo $\mathcal{C}$. Sabemos que existe pelo menos uma σ-álgebra contendo $\mathcal{C}$, que é $\mathcal{P}(\Omega)$. Se $\{A_i\}_{i\in J}$ é o conjunto de todas as σ-álgebras contendo $\mathcal{C}$, então $\bigcap_{i\in J} A_i$ trivialmente é a menor σ-álgebra de subconjuntos de Ω contendo $\mathcal{C}$.

Neste trabalho nos interessam as σ-álgebras geradas por processos estocásticos.

Definição 2.3 (Espaço de Probabilidade)**.** Um *espaço de probabilidade* é a terna $(\Omega, \mathcal{F}, \mathbb{P})$, onde

a) Ω é um conjunto não vazio;

b) $\mathcal{F}$ é uma σ-álgebra de subconjuntos de Ω. Seus elementos são chamados de *eventos aleatórios*;

c) $\mathbb{P}$ é uma probabilidade definida sobre os elementos de $\mathcal{F}$.

Definição 2.4. Uma *variável aleatória* X em um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ é uma função real definida no espaço Ω tal que $[X \leq x] := \{\omega \in \Omega : X(\omega) \leq x\}$ é evento aleatório para todo $x \in \mathbb{R}$.

Definição 2.5 (Processo Estocástico)**.** Um *processo estocástico* é uma coleção de variáveis aleatórias

$$\{X_t\}_{t\in\mathcal{T}} = (X_t(\omega), t \in \mathcal{T}, \omega \in \Omega), \tag{2.1}$$

todas elas definidas no mesmo espaço $(\Omega, \mathcal{F}, \mathbb{P})$, onde $\mathcal{T} \subset \mathbb{R}$ é um conjunto qualquer de índices.

Observação 2.1. Quando $\mathcal{T} = \{0, \cdots, n\}$, a sequência $\{X_t\}_{t=0}^{n}$ é chamada *Série Temporal*.

Definição 2.6 (σ-álgebra Gerada por um Processo Estocástico)**.** Seja $X = \{X_t\}_{t\in\mathcal{T}}$ um processo estocástico. A *σ-álgebra gerada por X*, denotada por $\sigma(X)$, é a menor σ-álgebra contendo todos os conjuntos da forma

$$\{\omega \in \Omega : X_t(\omega) \text{ pertence a } \mathcal{C}\}, \text{ onde } t \in \mathcal{T} \text{ e } \mathcal{C} \text{ é um boreliano.}$$

Para uma variável aleatória, ou um vetor aleatório ou um processo estocástico X em Ω, a σ-álgebra $\sigma(X)$ gerada por X contém as informações essenciais sobre a estrutura de X como uma função de $\omega \in \Omega$. Como X gera a σ-álgebra, dizemos que X *carrega a informação* $\sigma(X)$.

Sejam X e Y duas variáveis aleatórias ou dois processos estocásticos em um mesmo espaço $(\Omega, \mathcal{F}, \mathbb{P})$, onde $\mathcal{F}$ é a σ-álgebra em Ω. Então dizemos que

- A informação de Y está contida em $\mathcal{F}$ ou Y não contém mais informações do que aquelas contidas em $\mathcal{F}$ se $\sigma(Y) \subset \mathcal{F}$.
- X contém mais informações que Y se $\sigma(Y) \subset \sigma(X)$.

Para compreendermos um pouco mais a estrutura e a definição de martingales, útil neste trabalho e a ser dada na Seção 2.9, é importante tratarmos da esperança condicional de uma variável aleatória X condicionada a outra variável Y. Para leitores com familiaridade na teoria da medida, recomendamos o livro do Billingsley (1995). Caso contrário, James (2008) apresenta a esperança condicional sem a necessidade de um prévio conhecimento da teoria da medida.

Definição 2.7 (Esperança Condicional)**.** Sejam X e Y variáveis aleatórias em $(\Omega, \mathcal{F}, \mathbb{P})$. A *esperança condicional de X dado que $Y = y$* é dada por

$$\mathbb{E}(X|Y = y) = \int x\, dF_X(x|Y = y).$$

Exemplo 2.6. Seja X uma variável aleatória discreta, que assume os valores $x_1, x_2, \cdots, x_n$. Seja Y uma variável aleatória tal que $\mathbb{P}(Y = y) > 0$. Então, a

esperança condicional de X dado $Y = y$ é dada por

$$\begin{aligned}\mathbb{E}(X|Y=y) &= \sum_{i=1}^{n} x_i \frac{\mathbb{P}(\{\omega : X(\omega) = x_i\} \cap Y = y)}{\mathbb{P}(Y=y)} \\ &= \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i | Y = y).\end{aligned}$$

Exemplo 2.7. Seja X variável aleatória contínua, com função densidade $f_X(\cdot)$. Seja Y uma variável aleatória tal que $\mathbb{P}(Y = y) > 0$. Então, a esperança condicional de X dado $Y = y$ é dada por

$$\mathbb{E}(X|Y=y) = \frac{1}{\mathbb{P}(Y=y)} \int_{-\infty}^{\infty} x f_{X,Y}(x,y)\, dx.$$

Proposição 2.1. *Sejam X_1, X_2, X variáveis aleatórias quaisquer e Y uma variável aleatória qualquer. Então, valem as seguintes propriedades*

i) A esperança condicional é linear, ou seja, dadas a e b constantes reais, então

$$\mathbb{E}(aX_1 + bX_2|Y=y) = a\mathbb{E}(X_1|Y=y) + b\mathbb{E}(X_2|Y=y);$$

ii) As esperanças de $\mathbb{E}(X|Y=y)$ e de X são as mesmas, ou seja

$$\mathbb{E}[\mathbb{E}(X|Y=y)] = \mathbb{E}(X);$$

iii) Se X e Y são independentes, então

$$\mathbb{E}(X|Y=y) = \mathbb{E}(X) \;\; e \;\; \mathbb{E}(Y|X=x) = \mathbb{E}(Y).$$

Demonstração: Ver Billingsley (1995), página 447. □

Definição 2.8 (Esperança Condicional em Relação à uma σ-Álgebra)**.** Sejam X e Y variáveis aleatórias quaisquer. A variável aleatória Y é chamada a *esperança condicional de X dada a σ-álgebra $\mathcal{F}$* se

i) $\sigma(Y) \subset \mathcal{F}$;

ii) Y satisfaz a relação

$$\mathbb{E}\left(XI_A\right) = \mathbb{E}\left(YI_A\right), \text{ para todo } A \in \mathcal{F}.$$

Teorema 2.1. *Sejam X e Y variáveis aleatórias (ou vetores aleatórios ou processos estocásticos) em $(\Omega, \mathcal{F}, \mathbb{P})$ e $\sigma(Y)$ a σ-álgebra gerada por Y. Então, vale a igualdade*

$$\mathbb{E}\left(X|Y\right) = \mathbb{E}\left(X|\sigma\left(Y\right)\right).$$

Demonstração: Ver Billingsley (1995), página 447. □

Assim, a esperança de uma variável aleatória condicionada a uma outra variável aleatória é o mesmo que condicionar a uma σ-álgebra. Logo, é razoável herdar algumas propriedades da teoria de probabilidade.

Proposição 2.2. *Sejam X e Y variáveis aleatórias quaisquer. Então, valem as seguintes propriedades.*

i) Se X e a σ-álgebra $\mathcal{F}$ são independentes, então

$$\mathbb{E}\left(X|\mathcal{F}\right) = \mathbb{E}\left(X\right); \tag{2.2}$$

ii) Se $\sigma(X)$ é a σ-álgebra gerada por X e $\sigma(X) \subset \mathcal{F}$, onde $\mathcal{F}$ é σ-álgebra qualquer, então

$$\mathbb{E}\left(X|\mathcal{F}\right) = X. \tag{2.3}$$

Em particular, se X é uma função de Y, ou seja, $\sigma(X) \subset \sigma(Y)$, então

$$\mathbb{E}\left(X|Y\right) = X; \tag{2.4}$$

iii) Seja $\mathcal{F}$ uma σ-álgebra. Então,

$$\mathbb{E}\left(X\right) = \mathbb{E}\left[\mathbb{E}\left(X|\mathcal{F}\right)\right]; \tag{2.5}$$

iv) Sejam $\mathcal{F}$ e $\mathcal{G}$ duas σ-álgebras, tais que $\mathcal{F} \subset \mathcal{G}$. Então,

$$\mathbb{E}\left(X|\mathcal{F}\right) = \mathbb{E}\left(\mathbb{E}\left(X|\mathcal{G}\right)|\mathcal{F}\right) = \mathbb{E}\left(\mathbb{E}\left(X|\mathcal{F}\right)|\mathcal{G}\right). \tag{2.6}$$

Demonstração: Ver Billingsley (1995), página 447. □

Teorema 2.2 (Teorema da Convergência Dominada de Lebesgue)**.** *Se* $\{X_n\}_{n\in\mathbb{N}}$ *é uma sequência de variáveis aleatórias uniformemente limitada, e* $X = \lim_{n\to\infty} X_n$ *com probabilidade 1, então*

$$\mathbb{E}(X) = \mathbb{E}\left(\lim_{n\to\infty} X_n\right) = \lim_{n\to\infty} \mathbb{E}(X_n). \tag{2.7}$$

Demonstração: Ver Billingsley (1995), página 77. □

Dado um processo estocástico $\{X_t\}_{t\in\mathcal{T}}$, onde $\mathcal{T} \in \mathbb{R}$ é um conjunto de índices quaisquer, nos interessa conhecer qual o nível de dependência entre as variáveis aleatórias geradas pelo processo ao longo do tempo. Esse nível de dependência é calculado pela função de autocovariância.

Definição 2.9 (Função de Autocovariância)**.** Seja $\{X_t\}_{t\in\mathcal{T}}$ um processo estocástico, onde $\mathcal{T} \subset \mathbb{R}$ é um conjunto de índices quaisquer, tal que $Var(X_t) < \infty$, para todo $t \in \mathcal{T}$. Então, a *função de autocovariância* $\gamma_X(\cdot,\cdot)$ de $\{X_t\}_{t\in\mathcal{T}}$ é definida por

$$\gamma_X(r,s) = Cov(X_r, X_s) = \mathbb{E}\left[(X_r - \mathbb{E}(X_r))(X_s - \mathbb{E}(X_s))\right],\ r,s \in \mathcal{T}. \tag{2.8}$$

Definição 2.10 (Estacionariedade Fraca)**.** O processo estocástico $\{X_t\}_{t\in\mathbb{Z}}$ é *fracamente estacionário* se

i) $\mathbb{E}(|X_t|^2) < \infty$, para todo $t \in \mathbb{Z}$,

ii) $\mathbb{E}(X_t) = c$, para todo $t \in \mathbb{Z}$,

iii) $\gamma_X(r,s) = \gamma_X(r+h, s+h)$, para todo $r,s,t \in \mathbb{Z}$.

Quando o processo estocástico é fracamente estacionário, considerando $h = -s$ no item iii) temos que

$$\gamma_X(r,s) = \gamma_X(r-s, s-s) = \gamma_X(r-s, 0),\ \text{para todo } r,s \in \mathbb{Z}.$$

Portanto, a função de autocovariância depende apenas da distância entre as variáveis aleatórias consideradas. Assim, é conveniente redefinir a função de autocovariância para os processos fracamente estacionários como uma função de apenas uma variável.

Definição 2.11. Seja $\{X_t\}_{t\in\mathbb{Z}}$ um processo estocástico fracamente estacionário. A *função de autocovariância* $\gamma_X(\cdot)$ é definida como

$$\gamma_X(h) := \gamma_X(h,0) = Cov(X_{t+h}, X_t) = \mathbb{E}(X_{t+h}X_t) - \mathbb{E}(X_{t+h})\,\mathbb{E}(X_t),$$

para todo $t, h \in \mathbb{Z}$.

Definição 2.12 (Função de Autocorrelação)**.** Seja $\{X_t\}_{t\in\mathbb{Z}}$ um processo estocástico fracamente estacionário. A *função de autocorrelação* $\rho_X(\cdot)$ é definida como

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

onde $\gamma_X(0) = Cov(X_t, X_t) = Var(X_t)$.

Definição 2.13 (Estacionariedade Forte)**.** O processo estocástico $\{X_t\}_{t\in\mathbb{Z}}$ é um processo estacionário forte se

$$(X_{t_1}, X_{t_2}, \cdots, X_{t_k})' \overset{\mathcal{D}}{=} (X_{t_1+h}, X_{t_2+h}, \cdots, X_{t_k+h})',$$

para toda escolha de $k \in \mathbb{N}$, para todo $t_1, \cdots, t_k, h \in \mathbb{Z}$.

Observação 2.2. Ao longo do texto, sempre que falarmos em estacionariedade estamos nos referindo à estacionariedade forte, salvo nos casos em que especificarmos o contrário.

Em particular, considerando $k = 1$ na Definição 2.13, temos que se um processo $\{X_t\}_{t\in\mathbb{Z}}$ é estacionário, então X_t tem a a mesma distribuição, para cada $t \in \mathbb{Z}$. Logo, se $\mathbb{E}(|X_t|^2) < \infty$, então $\mathbb{E}(X_t)$ e $Var(X_t)$ são constantes, para todo $t \in \mathbb{Z}$. Ainda, considerando $k = 2$ na Definição 2.13, temos que X_{t+h} e X_t tem a mesma distribuição conjunta, o que implica que a covariância é a mesma para todo $h \in \mathbb{Z}$. Então, se um processo é estacionário e possui segundo momento finito, isso implica que ele é fracamente estacionário. No entanto a recíproca não é verdadeira, como mostra o exemplo a seguir.

Os processos ergódigos (ver Dudley, 2004, página 268) são casos particulares dos processos estacionários fortes. Sob tais condições, o Teorema Ergódico de Birkhoff

assegura que se pode obter médias espaciais a partir de médias temporais. No Capítulo 5 estimamos propriedades dos processos estocásticos através de médias temporais. Sendo assim, estamos implicitamente assumindo que os processos aqui considerados são ergódicos.

Exemplo 2.8. Seja $\{X_t\}_{t\in\mathbb{Z}_+}$ um processo estocástico definido por

$$X_t = \begin{cases} W, \text{ se t é par,} \\ Y, \text{ se t é impar,} \end{cases}$$

onde $W \sim \mathcal{N}(1,1)$ e $Y \sim \mathcal{E}(1)$ são independentes. Então, $\{X_t\}_{t\in\mathbb{Z}_+}$ é fracamente estacionário, com

$$\gamma_X(h) = \begin{cases} 1, \text{ se } h = 0, \\ 0, \text{ se } h \neq 0. \end{cases}$$

No entanto, no caso particular $t = 1$ e $t = 2$, temos que $X_1 = Y$ e $X_2 = W$, o que implica que X_1 e X_2 têm distribuições diferentes. Logo, $\{X_t\}_{t\in\mathbb{Z}_+}$ não pode ser estacionário.

Temos um caso especial em que estacionariedade fraca (Definição 2.10) implica em estacionariedade (Definição 2.13).

Definição 2.14 (Processo Gaussiano)**.** O processo estocástico $\{X_t\}_{t\in\mathbb{Z}}$ é um *processo Gaussiano* se e somente se para toda escolha de k e para todo $t_1, t_2, \cdots, t_k \in \mathbb{Z}$, $(X_1, X_2, \cdots, X_k)'$ tem distribuição normal multivariada.

Então, se $\{X_t\}_{t\in\mathbb{Z}}$ é um processo Gaussiano fracamente estacionário, não é difícil ver que, $\{X_t\}_{t\in\mathbb{Z}}$ é também um processo estacionário. Basta notar que para todo $k \in \mathbb{N}$ e para todo $h, t_1, \cdots, t_k \in \mathbb{Z}$, temos que os vetores aleatórios $(X_{t_1}, X_{t_2}, \cdots, X_{t_k})'$ e $(X_{t_1+h}, X_{t_2+h}, \cdots, X_{t_k+h})'$ tem a mesma média e a mesma matriz de covariância. Logo, possuem a mesma distribuição.

Definição 2.15. Sejam $X_1, X_2, \cdots, X_n$ observações de um processo estocástico $\{X_t\}_{t\in\mathcal{T}}$.

i) A *média amostral* de $\{X_t\}_{t\in\mathcal{T}}$ é

$$\overline{X} := \frac{1}{n}\sum_{t=1}^{n} X_t. \tag{2.9}$$

ii) A *função de autocovariância amostral* de $\{X_t\}_{t\in\mathcal{T}}$ é

$$\widehat{\gamma}_X(h) := \frac{1}{n}\sum_{t=1}^{n-|h|}\left(X_{t+|h|}-\overline{X}\right)\left(X_t-\overline{X}\right),\ -n<h<n. \tag{2.10}$$

iii) A *função de autocorrelação amostral* de $\{X_t\}_{t\in\mathcal{T}}$ é

$$\widehat{\rho}_X(h) = \frac{\widehat{\gamma}_X(h)}{\widehat{\gamma}_X(0)},\ h\in\mathbb{N}. \tag{2.11}$$

2.2 Teoria Espectral de Processos Estacionários

Seja $\{X_t\}_{t\in\mathbb{Z}}$ um processo real estacionário com função de autocovariância absolutamente somável, ou seja, $\sum_{k=-\infty}^{\infty}|\gamma_X(k)|<\infty$. Então, pode-se mostrar (ver Brockwell e Davis, 1991) que a transformada de Fourier de $\gamma_X(\cdot)$ existe e é dada por

$$\begin{aligned}
f_X(\omega) &= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_X(k)\,e^{-i\omega k} = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_X(k)\left[\cos(\omega k)+i\,\mathrm{sen}(\omega k)\right]\\
&= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_X(k)\cos(\omega k)\\
&= \frac{1}{2\pi}\gamma_X(0)+\frac{1}{2\pi}\sum_{k=-\infty}^{-1}\gamma_X(k)\cos(\omega k)+\frac{1}{2\pi}\sum_{k=1}^{\infty}\gamma_X(k)\cos(\omega k)\\
&= \frac{1}{2\pi}\gamma_X(0)+\frac{1}{\pi}\sum_{k=1}^{\infty}\gamma_X(k)\cos(\omega k),\ -\pi\le\omega\le\pi,
\end{aligned} \tag{2.12}$$

onde usamos as igualdades

$$\gamma_X(k)=\gamma_X(-k),\ \mathrm{sen}(-\omega k)=-\mathrm{sen}(\omega k)\ \text{ e }\ \cos(-\omega k)=\cos(\omega k).$$

Definição 2.16 (Função Densidade Espectral)**.** Seja $\{X_t\}_{t\in\mathbb{Z}}$ um processo real estacionário tal que $\sum_{k=-\infty}^{\infty}|\gamma_X(k)|<\infty$. A função $f_X(\cdot)$ obtida em (2.12) é chamada *função densidade espectral* do processo $\{X_t\}_{t\in\mathbb{Z}}$.

Através do Teorema de Herglotz (ver Brockwell e Davis, 1991, página 117) mostra-se que a função de autocovariância pode ser obtida a partir de sua transformada de Fourier $f_X(\cdot)$ aplicando-se a transformada inversa de Fourier

$$\gamma_X(k) = \int_{-\pi}^{\pi} f_X(\omega)\, e^{i\omega k}\, d\omega. \tag{2.13}$$

Proposição 2.3. *Seja $\{X_t\}_{t\in\mathbb{Z}}$ um processo real estacionário com função de autocovariância absolutamente somável. A transformada de Fourier $f_X(\cdot)$ de $\gamma_X(\cdot)$, definida em (2.12), possui as seguintes propriedades*

i) $f_X(\cdot)$ é função real contínua não negativa;

ii) $f_X(\cdot)$ é periódica de período 2π;

iii) A variância de X_t, para todo $t \in \mathbb{Z}$, é dada por

$$Var(X_t) = \gamma_X(0) = \int_{-\pi}^{\pi} f_X(\omega)\, d\omega. \tag{2.14}$$

Demonstração: Ver Brockwell e Davis (1991).

Exemplo 2.9. Seja $\gamma_X(\cdot)$ a função de autocovariância de um processo estacionário, dada por

$$\gamma_X(h) = \begin{cases} 1, \text{ se } h = 0, \\ \rho, \text{ se } h = \pm 1, \\ 0, \text{ caso contrário,} \end{cases}$$

onde $|\rho| \leq \frac{1}{2}$.

Como $\gamma_X(\cdot)$ é absolutamente somável, temos que a função densidade espectral do processo estocástico é dada por

$$\begin{aligned} f_X(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{i\omega k} \gamma_X(k) \\ &= \frac{1}{2\pi}\gamma_X(0) + \frac{1}{\pi}\sum_{k=1}^{\infty} \gamma_X(k)\cos(\omega k) \\ &= \frac{1}{2\pi}\left[1 + 2\rho\cos(\omega)\right]. \end{aligned}$$

Nota que a função de densidade espectral, obtida na equação (2.12), é dada somente para os casos em que a função de autocovariância é absolutamente somável. Caso contrário, temos o seguinte resultado.

Corolário 2.1. *Uma função* $\gamma_X(\cdot) : \mathbb{Z} \to \mathbb{C}$ *é a função de autocovariância de um processo estacionário* $\{X_t\}_{t\in\mathbb{Z}}$ *se e somente se ocorre uma das seguintes situações:*

i) $\gamma_X(h) = \int_{-\pi}^{\pi} e^{i\omega h}\, dF_X(\omega)$, *para todo* $h \in \mathbb{Z}$, *onde* F_X *é uma função limitada definida em* $[-\pi, \pi]$, *contínua à direita, não decrescente e com* $F_X(-\pi) = 0$;

ii) $\sum_{i,j=1}^{n} a_i \gamma_X(i-j)\,\overline{a}_j \geq 0$, *para todo* $n \in \mathbb{Z}_+$ *e para todo* $a = (a_1, \cdots, a_n)' \in \mathbb{C}^n$.

A função $F_X(\cdot)$ *é chamada* função de distribuição espectral *e sua respectiva densidade, quando ela existe, é chamada* função de densidade espectral.

Demonstração: Ver Brockwell e Davis (2006), página 119. □

Definição 2.17 (Função Periodograma)**.** Seja $\{X_t\}_{t=1}^{n}$ uma série temporal gerada a partir de um processo estocástico $\{X_t\}_{t\in\mathbb{R}}$, com função de autocovariância não necessariamente absolutamente convergente. A função

$$I_n(\omega) = \frac{1}{n}\left|\sum_{t=1}^{n} X_t e^{-it\omega}\right|^2, \text{ para todo } \omega \in (0, \pi], \tag{2.15}$$

é chamada *função periodograma* da série temporal $\{X_t\}_{t=1}^{n}$.

Brockwell e Davis (1991) mostram que (2.15) é equivalente a

$$I_n(\omega) = \begin{cases} n\left|\overline{X}\right|^2, \text{ se } \omega = 0, \\ \sum_{|k|<n} \hat{\gamma}_X(k)\, e^{-it\omega}, \text{ se } \omega \in (0, \pi], \end{cases} \tag{2.16}$$

onde $\overline{X}$ é a média amostral e $\hat{\gamma}_X(\cdot)$ é a função de autocovariância amostral definidos, respectivamente, em (2.9) e (2.11).

Da expressão (2.16), um estimador natural para a função densidade espectral $f_X(\omega)$ é $\dfrac{I_n(\omega)}{2\pi}$, para todo $\omega \in (0, \pi]$.

Na prática, a função periodograma só pode ser calculada para um número finito de frequências $\omega \in (0, \pi]$. Pode-se mostrar que (2.15) fica completamente determinada por seus valores nas frequências de Fourier $\omega_j = \frac{2\pi j}{n}$, para $j \in \left\{0, 1, \cdots, \left\lfloor \frac{n}{2} \right\rfloor\right\}$.

Brockwell e Davis (1991) provam que $\dfrac{I_n(\cdot)}{2\pi}$ é um estimador assintoticamente não-viciado, porém inconsistente, para a função densidade espectral, isto é,

$$\lim_{n\to\infty} \mathbb{E}\left(\frac{I_n(\omega)}{2\pi}\right) = f_X(\omega) \text{ e}$$
$$\lim_{n\to\infty} Var\left(\frac{I_n(\omega)}{2\pi}\right) \neq 0,$$

para $\omega \in (0, \pi]$.

Em vista disso, aplicamos uma suavização na função periodograma, através de uma *função de ponderação*, obtendo assim a *função periodograma suavizado*, que é um estimador não-viciado e consistente para a função densidade espectral (Brockwell e Davis, 1991).

Definição 2.18 (Função Periodograma Suavizado)**.** Seja $\{X_t\}_{t=1}^n$ uma série temporal gerada a partir de um processo estocástico $\{X_t\}_{t\in\mathbb{R}}$, com função de autocovariância não necessariamente absolutamente convergente. A *função periodograma suavizado*, denotada por $\tilde{I}_n(\cdot)$, é dada por

$$\tilde{I}_n(\omega) = \frac{1}{2\pi} \sum_{|k|\leq m} \Lambda\left(\frac{k}{m}\right) \hat{\gamma}_X(k)\, e^{-ik\omega}, \text{ para } \omega \in (0, \pi], \tag{2.17}$$

onde $m = n^\beta$, para $0 \leq \beta \leq 1$ e $\Lambda(\cdot)$ é uma função par, contínua por partes e satisfazendo as condições

i) $0 \leq \Lambda(x) \leq \Lambda(0) = 1$;

ii) $\Lambda(-x) = \Lambda(x)$, para todo x;

iii) $\Lambda(x) = 0$, para $|x| > 1$.

A função $\Lambda(\cdot)$ é chamada *função de ponderação* ou *núcleo*. Nesse trabalho, utilizamos o núcleo da Bartlett, que é dado por

$$\Lambda_m(h) = \begin{cases} 1 - \frac{|h|}{m}, \text{ se } |h| \leq m, \\ 0, \text{ se } |h| > m, \end{cases} \tag{2.18}$$

onde m é chamado ponto de truncamento.

2.3 Processos Estocásticos com Longa Dependência

A primeira noção de longa dependência em análise de dados surgiu a partir de 1906 com Harold Edwin Hurst, durante seu trabalho para predizer o nível de inundação do rio Nilo, no Egito. Ele desenvolveu um teste para a dependência de longo alcance e encontrou correlações significativas a longo prazo entre as vazões do Rio Nilo e descreveu as correlações em termos de leis de potência. Essa estatística ficou conhecida como estatística R/S. Nos anos 60, surgem estudos mais importantes nessa área, realizados por Mandelbrot (1963) e Mandelbrot e Wallis (1969). Foi devido à influência dessas publicações e dos artigos subsequentes de Mandelbrot que a longa dependência tornou-se também associada com o comportamento fractal. Atualmente tem-se intensificado os estudos dos processos de longa dependência, devido à sua grande aplicabilidade na área de finanças (ver Olbermann et al., 2006 e Lopes, 2008), econometria, sequenciamento de DNA (ver Lopes e Nunes, 2006), etc.

Assim como apresentamos neste trabalho, é padrão na teoria de processos estocásticos aplicar a noção de longa dependência somente à processos estacionários. No entanto, artigos recentes tem expandido o estudo à processos não-estacionários (ver Lopes et al., 2004).

Há diversos critérios diferentes para determinar se um processo possui longa dependência. Apresentamos neste trabalho os critérios de longa dependência pro-

postos por Lopes e Nunes (2006) e Lopes (2008).

Definição 2.19 (Longa Dependência)**.** Seja $\rho_X(\cdot)$ a função de autocorrelação de um processo estocástico estacionário $\{X_t\}_{t\in\mathbb{Z}}$. O processo $\{X_t\}_{t\in\mathbb{Z}}$ possui *longa dependência* se existe um número real $d \in \left(0, \frac{1}{2}\right)$ tal que

$$\rho_X(h) \approx C\, h^{2d-1},\ C \text{ não nulo, quando } h \to \infty, \tag{2.19}$$

ou, equivalentemente, se $f_X(\cdot)$ é a função de densidade espectral do processo, o processo $\{X_t\}_{t\in\mathbb{Z}}$ possui *longa dependência* se existe um número real $d \in \left(0, \frac{1}{2}\right)$ tal que

$$f_X(\omega) \approx \omega^{-2d}, \text{ quando } \omega \to 0.$$

Observação 2.3. Palma (2007) mostra que, nos casos em que o processo estocástico $\{X_t\}_{t\in\mathbb{Z}}$ pode ser escrito na forma

$$X_t = \sum_{j=0}^{\infty} \psi_j \xi_{t-j}, \tag{2.20}$$

onde $\psi_0 = 1$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ e $\{\xi_t\}_{t\in\mathbb{Z}}$ é um ruído branco com variância constante, então, se $\gamma_X(\cdot)$ é a função de autocovariância de $\{X_t\}_{t\in\mathbb{Z}}$, a ocorrência de (2.19) implica que

$$\sum_{h=-\infty}^{\infty} |\gamma_X(h)| = \infty. \tag{2.21}$$

2.4 Movimento Browniano

A primeira noção de movimento Browniano surgiu a partir das observações do célebre botânico Robert Brown, em 1827, enquanto estudava a vida das plantas no mares do sul. Ele observou que os grãos de pólem, em uma substância aquosa, mantinham um rápido movimento errático. O movimento Browniano tem um papel fundamental na Teoria da Probabilidade, nos Processos Estocásticos, na Física, etc, e é uma ferramenta fundamental nesse trabalho. Nesta seção iniciamos com a definição e algumas propriedades importantes desse processo.

Definição 2.20 (Movimento Browniano)**.** Um processo estocástico $B = \{B_t\}_{t\geq 0}$ é chamado *movimento Browniano ou processo de Wiener* se

i) $B_0 = 0$ com probabilidade 1 e B_t é contínua em $t = 0$;

ii) possui incrementos independentes, isto é, para toda escolha de $t_i \in [0, \infty)$, para $i = 1, \cdots, n$, com $t_1 < \cdots < t_n$ e $n \geq 1$, temos que

$$B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}}$$

são variáveis aleatórias independentes;

iii) os incrementos $B_t - B_s$ tem distribuição $\mathcal{N}(0, t-s)$, para todo $t > s$.

Da Definição 2.20 obtemos que o processo $B = \{B_t\}_{t\geq 0}$ possui incrementos estacionários, $B_t \sim \mathcal{N}(0, t)$, para todo $t > 0$, e $B = \{B_t\}_{t\geq 0}$ possui trajetórias contínuas quase certamente (ver Karlin e Taylor, 1975, capítulo 7).

Proposição 2.4. *Seja $B = \{B_t\}_{t\geq 0}$ o processo movimento Browniano. Então, para todo $s, t \geq 0$ temos que a função de autocorrelação de $\{B_t\}_{t\geq 0}$ é dada por*

$$\gamma_B(s,t) = \mathbb{E}(B_s B_t) = min(s,t). \tag{2.22}$$

Demonstração: Considere $s, t \geq 0$ tais que $s \leq t$. Então, da independência dos incrementos do movimento Browniano, temos que

$$\begin{aligned}
\gamma_B(s,t) &= \mathbb{E}(B_s B_t) \\
&= \mathbb{E}\left[B_s(B_t - B_s) + (B_s)^2\right] \\
&= \mathbb{E}[B_s(B_t - B_s)] + \mathbb{E}\left(B_s^2\right) \\
&= \mathbb{E}[(B_s - B_0)(B_t - B_s)] + Var(B_s) \\
&= \mathbb{E}(B_s - B_0)\,\mathbb{E}(B_t - B_s) + s = s.
\end{aligned}$$

Análogo para o caso $t \leq s$. □

2.5 Processos de Lévy

Nesta seção vamos definir os processos de Lévy e apresentar algumas das suas propriedades.

Definição 2.21 (Processo de Lévy)**.** Um processo estocástico $\{X_t\}_{t\geq 0}$ é um *processo de Lévy* se

i) $X_0 = 0$;

ii) possui incrementos independentes e estacionários;

iii) possui trajetórias contínuas à direita, com limites à esquerda finitos.

Exemplo 2.10. Descrevemos aqui três exemplos de processos de Lévy.

i) O processo movimento Browniano é um processo de Lévy.

ii) Seja $X \sim \mathcal{G}(\alpha, \beta)$ uma variável aleatória com distribuição gama. Sua função densidade é dada por

$$f_X(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}},\ x > 0,$$

onde $\Gamma(\cdot)$ denota a função gama usual.

O processo gama com parâmetros μ e ν, denotado por $\{G_{\mu,\nu}(t)\}_{t\geq 0}$, é um processo contínuo, com incrementos independentes e estacionários, tais que, para todo t e todo $h > 0$,

$$G_{\mu,\nu}(t+h) - G_{\mu,\nu}(t) \sim \mathcal{G}\left(\mu^2 h, \frac{\nu}{\mu}\right).$$

Logo, o processo gama $\{G_{\mu,\nu}(t)\}_{t\geq 0}$ é um processo de Lévy.

iii) Um exemplo de processo de Lévy não contínuo é o processo de Poisson (ver Karlin e Taylor, 1975, páginas 22 e 118), cuja função de distribuição contém saltos.

2.6 Distribuições Estáveis

As distribuições estáveis são uma classe de distribuições probabilísticas, caracterizadas por Paul Lévy em meados de 1920 ao estudar a soma de variáveis aleatórias independentes e identicamente distribuídas. Tem como principais características a possibilidade de possuir cauda pesada (uma distribuição possui *cauda pesada* se $\mathbb{P}(|X| > x) \sim x^{-\varsigma}$, quando $x \to \infty$, $0 < \varsigma < 2$) e assimetria. Uma das consequências de uma variável aleatória possuir cauda pesada é a não existência de todos os momentos finitos. Nesse trabalho, o *r-ésimo momento absoluto* de uma variável aleatória X é definido como

$$\mathbb{E}(|X|^r) = \int_0^\infty \mathbb{P}(|X|^r > x)\, dx, \tag{2.23}$$

onde $r \in \mathbb{R}$.

Um dos incovenientes das distribuições estáveis é a ausência de uma fórmula fechada para a grande maioria das funções de distribuição e densidade. No entanto, com o avanço dos recursos computacionais, temos hoje programas confiáveis que nos possibilitam a aplicação das distribuições estáveis em diversos problemas práticos, como, por exemplo, sistemas econômicos (Mandelbrot, 1963), sistemas de comunicação (Stuck e Kleiner, 1974), etc.

Definição 2.22 (Estabilidade, Estabilidade Estrita e Estabilidade Simétrica)**.** Uma variável aleatória X é *estável* se dadas duas amostras aleatórias independentes X_1 e X_2 de X e quaisquer a, b constantes reais positivas, existem c, $d \in \mathbb{R}$ tais que

$$aX_1 + bX_2 \stackrel{\mathcal{D}}{=} cX + d. \tag{2.24}$$

Se $d = 0$ para toda escolha de a e b, dizemos que X é *estritamente estável*.

No caso em que X é estável e $X \stackrel{\mathcal{D}}{=} -X$, dizemos que X é *simetricamente estável*.

Embora a Definição 2.22 seja extremamente intuitiva, temos uma definição equivalente de estabilidade mais interessante, pois nos permite conhecer a estrutura

da variável aleatória. A demonstração da equivalência pode ser vista em Gnedenko e Kolmogorov (1954).

Definição 2.23. Uma variável aleatória X é *α-estável* se existe um índice de estabilidade $\alpha \in (0, 2]$, um parâmetro de assimetria $\beta \in [-1, 1]$, um parâmetro de escala $\sigma^* > 0$ e um parâmetro de locação $\mu \in \mathbb{R}$ tal que X tem função característica dada por

$$\varphi_X(t) = \mathbb{E}\left(e^{itX}\right) = \begin{cases} exp\left\{-\sigma^*|t|\left[1 + i\beta\frac{2}{\pi}S(t)\ln(|t|)\right] + i\mu t\right\}, \text{ se } \alpha = 1, \\ exp\left\{-\sigma^{*\alpha}|t|^\alpha\left[1 - i\beta S(t)\tan\left(\frac{\pi\alpha}{2}\right)\right] + i\mu t\right\}, \text{ caso contrário,} \end{cases} \tag{2.25}$$

onde $S(\cdot)$ é a função sinal, definida por

$$S(t) = \begin{cases} -1, \text{ se } t < 0, \\ 0, \text{ se } t = 0, \\ 1, \text{ se } t > 0. \end{cases} \tag{2.26}$$

Como a distribuição estável requer quatro parâmetros, quando X é α-estável dizemos que $X \sim S_\alpha(\sigma^*, \beta, \mu)$.

No caso em que $\beta = \mu = 0$, dizemos que X é α-estável simétrica, denotada por SαS, e sua função característica é dada por

$$\varphi_X(t) = e^{-\sigma^{*\alpha}|t|^\alpha}. \tag{2.27}$$

Exemplo 2.11. Algumas distribuições conhecidas são α-estáveis e possuem uma forma fechada para representar sua função densidade.

i) Se X é uma variável aleatória com função de distribuição $\mathcal{N}(\mu, \sigma^2)$, então $X \sim S_2\left(\frac{\sigma}{\sqrt{2}}, \beta, \mu\right)$. Nota que no caso que $\alpha = 2$, temos sempre uma distribuição Gaussiana, onde μ é a média, $\frac{\sigma}{\sqrt{2}}$ é a variância e β é irrelevante.

ii) Se X tem distribuição Cauchy(μ, σ), ou seja, se X tem função densidade

$$f_X(x) = \frac{1}{\pi}\frac{\sigma}{(x-\mu)^2 + \sigma^2}, \; x \in \mathbb{R}, \; \mu \in \mathbb{R} \text{ e } \sigma > 0, \tag{2.28}$$

então $X \sim S_1(\sigma, 0, \mu)$.

iii) Se X tem distribuição Lévy(σ, μ), ou seja, se X tem função densidade

$$f_X(x) = \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}} \frac{1}{(x-\mu)^{\frac{3}{2}}} e^{-\frac{\sigma}{2(x-\mu)}}, \text{ para todo } x \in \mathbb{R}, \ x > \mu, \tag{2.29}$$

então $X \sim S_{\frac{1}{2}}(\sigma, 1, \mu)$.

Uma característica essencial para as distribuições α-estáveis é que a soma de variáveis aleatórias independentes α-estáveis também é α-estável, embora apresente diferentes parâmetros, como mostra a proposição a seguir.

Proposição 2.5. *Sejam X_1 e X_2 variáveis aleatórias independentes tais que $X_i \sim S_\alpha(\sigma^*{}_i, \beta_i, \mu_i)$, $i = 1, 2$, conforme a Definição 2.23. Então, $X_1 + X_2 \sim S_\alpha(\sigma^*, \beta, \mu)$, onde*

$$\sigma^* = (\sigma^{*\alpha}_1 + \sigma^{*\alpha}_2)^{\frac{1}{\alpha}}, \ \beta = \frac{\beta_1 \sigma^{*\alpha}_1 + \beta_2 \sigma^{*\alpha}_2}{\sigma^{*\alpha}_1 + \sigma^{*\alpha}_2} \ \ e \ \ \mu = \mu_1 + \mu_2.$$

Demonstração: Suponha que $\alpha \neq 1$. Se $\varphi_{X_i}(\cdot)$ representa a função característica de X_i, para $i = 1, 2$, então pela independência obtemos que

$$\begin{aligned}
\ln\left(\varphi_{X_1+X_2}(t)\right) &= \ln\left(\varphi_{X_1}(t)\,\varphi_{X_2}(t)\right) = \ln\left(\varphi_{X_1}(t)\right) + \ln\left(\varphi_{X_2}(t)\right) \\
&= -(\sigma^{*\alpha}_1 + \sigma^{*\alpha}_2)|t|^\alpha + i|t|^\alpha S(t) \tan\left(\frac{\pi\alpha}{2}\right)(\beta_1 \sigma^{*\alpha}_1 + \beta_2 \sigma^{*\alpha}_2) \\
&\quad + it(\mu_1 + \mu_2) \\
&= -(\sigma^{*\alpha}_1 + \sigma^{*\alpha}_2)|t|^\alpha \left[1 - i\frac{\beta_1 \sigma^{*\alpha}_1 + \beta_2 \sigma^{*\alpha}_2}{\sigma^{*\alpha}_1 + \sigma^{*\alpha}_2} S(t) \tan\left(\frac{\pi\alpha}{2}\right)\right] \\
&\quad + it(\mu_1 + \mu_2).
\end{aligned}$$

Logo,

$$\varphi_{X_1+X_2}(t) = e^{-(\sigma^{*\alpha}_1 + \sigma^{*\alpha}_2)|u|^\alpha \left[1 - i\frac{\beta_1 \sigma^{*\alpha}_1 + \beta_2 \sigma^{*\alpha}_2}{\sigma^{*\alpha}_1 + \sigma^{*\alpha}_2} S(t) \tan\left(\frac{\pi\alpha}{2}\right)\right] + it(\mu_1 + \mu_2)}.$$

No caso em que α for igual a um, a demonstração é análoga. □

A seguir apresentamos algumas importantes propriedades das distribuições estáveis.

Proposição 2.6. *Seja $X \sim S_\alpha(\sigma^*, \beta, \mu)$. Valem as seguintes propriedades.*

i) Se $a \in \mathbb{R}$ é uma constante, então $X + a \sim S_\alpha(\sigma^*, \beta, \mu + a)$.

ii) Se a é uma constante real não nula, então

$$\begin{aligned} aX &\sim S_1\left(|a|\sigma^*, S(a)\beta, a\mu - \frac{2a}{\pi}\ln(|a|)\sigma^*\beta\right), \text{ se } \alpha = 1, \\ aX &\sim S_\alpha(|a|\sigma^*, S(a)\beta, a\mu), \text{ caso contrário.} \end{aligned} \tag{2.30}$$

iii) Se $0 < \alpha < 2$, então $X \sim S_\alpha(\sigma^*, \beta, 0)$ se e somente se $-X \sim S_\alpha(\sigma^*, -\beta, 0)$.

iv) $X \sim S_\alpha(\sigma^*, \beta, \mu)$ é simétrica sobre μ se e somente se $\beta = 0$.

v) Se $\alpha \neq 1$, então $X \sim S_\alpha(\sigma^*, \beta, \mu)$ é estritamente estável se e somente se $\mu = 0$. Se $\alpha = 1$, então $X \sim S_\alpha(\sigma^*, \beta, \mu)$ é estritamente estável se e somente se $\beta = 0$.

Demonstração:

i) Suponha que $\alpha \neq 1$. Pela propriedade das funções características, temos que $\varphi_{X+a}(t) = e^{ita}\varphi_X(t)$. Logo, de (2.25) temos que

$$\begin{aligned} \varphi_{X+a}(t) &= exp\{ita\}\, exp\left\{-\sigma^{*\alpha}|t|^\alpha\left[1 - i\beta S(t)\tan\left(\frac{\pi\alpha}{2}\right)\right] + it\mu\right\} \\ &= exp\left\{(ita) - \sigma^{*\alpha}|u|^\alpha\left[1 - i\beta S(t)\tan\left(\frac{\pi\alpha}{2}\right)\right] + it\mu\right\} \\ &= exp\left\{-\sigma^{*\alpha}|t|^\alpha\left[1 - i\beta S(t)\tan\left(\frac{\pi\alpha}{2}\right)\right] + it(\mu + a)\right\}. \end{aligned}$$

Portanto, $X + a \sim S_\alpha(\sigma^*, \beta, \mu + a)$. O caso $\alpha = 1$ é análogo.

ii) Suponha que $\alpha \neq 1$. Da propriedade das funções características, temos que $\varphi_{aX}(t) = \varphi_X(at)$. Então, de (2.25) obtemos que

$$\begin{aligned} \ln(\varphi_{aX}(t)) &= \ln(\varphi_X(at)) \\ &= -|at|^\alpha\sigma^{*\alpha}\left[1 - i\beta S(at)\tan\left(\frac{\pi\alpha}{2}\right)\right] + i(at)\mu \\ &= -(\sigma^*|a|)^\alpha|t|^\alpha\left[1 - i\beta S(a)S(t)\tan\left(\frac{\pi\alpha}{2}\right)\right] + it(\mu a). \end{aligned}$$

Logo,

$$\varphi_{aX}(t) = exp\left\{-(\sigma^*|a|)^\alpha|t|^\alpha\left[1 - i\beta S(a)S(t)\tan\left(\frac{\pi\alpha}{2}\right)\right] + it\mu a\right\}.$$

O caso $\alpha = 1$ é análogo.

iii) Decorre da expressão (2.25).

iv) Das propriedades das funções características, sabemos que uma variável aleatória X é simétrica se e somente se $\varphi_X(\cdot)$ é real. Por (2.25), temos que isso ocorre se e somente se $\beta = 0$ e $\mu = 0$, o que prova a primeira afirmação. Utilizando o item i), temos demonstrada a segunda afirmação. □

Como citamos anteriormente, uma importante característica dos processos α-estáveis é a possibilidade de não existir todos os momentos finitos. A proposição a seguir apresenta esse resultado.

Proposição 2.7. *Seja* $X \sim S_\alpha(\sigma^*, \beta, \mu)$*, com* $0 < \alpha < 2$*. Se* $\mathbb{E}(|X|^r)$ *é definido por* (2.23)*, então*

$$\mathbb{E}(|X|^r) < \infty, \text{ para todo } 0 < r < \alpha,$$
$$\mathbb{E}(|X|^r) = \infty, \text{ para todo } r \geq \alpha.$$

Demonstração: Ver Samorodinitsky e Taqqu (1994), página 18. □

Quando $0 < \alpha < 2$, $X \sim S_\alpha(\sigma^*, \beta, \mu)$ tem variância infinita. No caso em que $0 < \alpha \leq 1$, $\mathbb{E}(|X|) = \infty$, ou seja, a esperança não está definida.

Quando $1 < \alpha \leq 2$, Samorodinitsky e Taqqu (1994) mostram que a média de $X \sim S_\alpha(\sigma^*, \beta, \mu)$ é definida pelo parâmetro de locação μ, ou seja, $\mathbb{E}(X) = \mu$.

2.7 Auto-Similaridade

A idéia básica de auto-similaridade é muito antiga, embora somente tenha sido introduzida em um contexto teórico por Kolmogorov (1941). Sua relevância na Estatística foi observada nos trabalhos de Mandelbrot e van Ness (1968) e Mandelbrot e Wallis (1969).

Nessa seção consideramos $T = \mathbb{R}$ ou $T = \mathbb{R}_+$.

Definição 2.24 (Processo Auto-Similar)**.** Um processo estocástico real $\{X_t\}_{t \in T}$ é *auto-similar com índice* $H > 0$ (denotamos por H-ss) se para qualquer $p \geq 1$,

$t_1, \cdots, t_p \in T$ e qualquer $a > 0$,

$$(X(at_1), X(at_2), \cdots, X(at_p))' \overset{\mathcal{D}}{=} \left(a^H X(t_1), a^H X(t_2), \cdots, a^H X(t_p)\right)'. \quad (2.31)$$

Por abuso de notação, quando (2.31) ocorre, dizemos que

$$\{X(at)\}_{t\in T} \overset{\mathcal{D}}{=} \left\{a^H X(t)\right\}_{t\in T}. \quad (2.32)$$

Quando considerarmos a variável aleatória unidimencional $X(t)$ para um t fixo, diremos que $X(at) \overset{\mathcal{D}}{=} a^H X(t)$.

Exemplo 2.12. O movimento Browniano $\{B_t\}_{t\in T}$, apresentado na Seção 2.4, é um processo auto-similar com índice $H = \frac{1}{2}$ (ver Samorodnitsky e Taqqu, 1994, página 331).

Considere $\{X(t)\}_{t\in T}$ um processo H-ss, tal que $0 \in T$. Então, para todo $a > 0$, temos que

$$X(0) = X(a0) \overset{\mathcal{D}}{=} a^H X(0). \quad (2.33)$$

Logo, para todo processo H-ss, temos que $X(0) = 0$ quase certamente.

O próximo resultado apresenta uma importante relação entre os processos H-ss e os processos estacionários.

Proposição 2.8. *Se $\{X(t)\}_{t\geq 0}$ é um processo H-ss, então o processo $\{Y(t)\}_{t\in\mathbb{R}}$ dado por*

$$Y(t) = e^{-tH} X\left(e^t\right) \quad (2.34)$$

é estacionário.

Analogamente, se $\{Y(t)\}_{t\in\mathbb{R}}$ é um processo estocástico estacionário, então o processo $\{X(t)\}_{t\geq 0}$ dado por

$$X(t) = t^H Y(\ln(t)) \quad (2.35)$$

é um processo H-ss.

Demonstração: Sejam $\theta_1, \cdots, \theta_k \in \mathbb{R}$. Se $\{X(t)\}_{t\geq 0}$ é um processo H-ss, então, para quaisquer $t_1, \cdots, t_k \in \mathbb{R}$ e $h > 0$, temos

$$\begin{aligned}
\sum_{j=1}^{k} \theta_j Y(t_j + h) &= \sum_{j=1}^{k} \theta_j e^{-(t_j+h)H} X\left(e^{t_j+h}\right) = \sum_{j=1}^{k} \theta_j e^{-t_j H} e^{-hH} X\left(e^{t_j} e^{h}\right) \\
&\overset{\mathcal{D}}{=} \sum_{j=1}^{k} \theta_j e^{-t_j H} e^{-hH} e^{hH} X\left(e^{t_j}\right) = \sum_{j=1}^{k} \theta_j e^{-t_j H} X\left(e^{t_j}\right) \\
&= \sum_{j=1}^{k} \theta_j Y(t_j).
\end{aligned}$$

Logo,

$$\sum_{j=1}^{k} \theta_j Y(t_j + h) \overset{\mathcal{D}}{=} \sum_{j=1}^{k} \theta_j Y(t_j), \tag{2.36}$$

ou seja, $\{Y(t)\}_{t\in\mathbb{R}}$ é um processo estacionário.

Analogamente, temos que se $\{Y(t)\}_{t\in\mathbb{R}}$ é um processo estacionário, então para quaisquer $t_1, \cdots, t_k > 0$ e $a > 0$, considerando $X(t) = t^H Y(\ln(t))$, para todo $t \geq 0$, temos que

$$\begin{aligned}
\sum_{j=1}^{k} \theta_j X(a t_j) &= \sum_{j=1}^{k} \theta_j (a t_j)^H Y(\ln(a t_j)) = \sum_{j=1}^{k} \theta_j a^H t_j^H Y(\ln(a) + \ln(t_j)) \\
&\overset{\mathcal{D}}{=} \sum_{j=1}^{k} \theta_j a^H {t_j}^H Y(t_j) = \sum_{j=1}^{k} \theta_j a^H X(t_j).
\end{aligned}$$

Portanto, $\{X(t)\}_{t\geq 0}$ é um processo H-ss. □

Exemplo 2.13. Seja $\{B_t\}_{t\geq 0}$ o processo movimento Browniano. Pelo exemplo 2.12 temos que $\{B_t\}_{t\geq 0}$ é um processo H-ss com $H = \frac{1}{2}$, Gaussiano e com média 0. Logo, denotando $B_t = B(t)$, temos que o processo $\{Y(t)\}_{t\in\mathbb{R}}$, onde

$$Y(t) = e^{\frac{-t}{2}} B\left(e^t\right), \text{ para todo } t \in \mathbb{R},$$

é um processo Gaussiano estacionário com média zero e função de autocovariância dada por

$$\begin{aligned}
\gamma_B(r, s) &= \mathbb{E}[Y(r) Y(s)] = e^{-\frac{1}{2}(r+s)} \mathbb{E}[B(e^r) B(e^s)] \\
&= e^{-\frac{1}{2}(r+s)} min(e^r, e^s) \\
&= e^{-\frac{1}{2}|s-r|}.
\end{aligned}$$

Definição 2.25. Se um processo $\{X(t)\}_{t\in T}$ possui incrementos estacionários e é auto-similar com parâmetro H, diremos que $\{X(t)\}_{t\in T}$ é um *processo H-sssi*.

Um exemplo de processo H-sssi é o processo movimento Browniano, com $H = \frac{1}{2}$.

O corolário a seguir nos mostra a relação entre os processos que são α-estáveis e H-sssi.

Corolário 2.2. *Seja $\{X(t)\}_{t\in\mathbb{R}}$ um processo não degenerado α-estável, $0 < \alpha \leq 2$, H-sssi. Então,*

$$\begin{aligned} \alpha < 1 &\Rightarrow 0 < H \leq \frac{1}{\alpha}, \\ \alpha \geq 1 &\Rightarrow 0 < H \leq 1. \end{aligned}$$

Demonstração: Ver Samorodnitsky e Taqqu (1994), página 316. □

2.8 Movimento Browniano Fracionário

O movimento Browniano fracionário é uma família de processos Gaussianos, indexado por um parâmetro $H \in (0, 1)$, chamado de parâmetro de Hurst. Foi introduzido por Kolmogorov (1940) e posteriormente estudado por Mandelbrot e Van Ness (1968). A primeira aplicação desse processo foi realizada por Hurst (1951) para estudar a capacidade de armazenamento a longo prazo dos reservatórios ao longo do rio Nilo. Posteriormente, Mandelbrot usou esses processos para modelar séries temporais aplicadas à economia e, mais recentemente, Leland e Taqqu (1994) estudaram aplicações na área de modelagem do tráfego de telecomunicações. Uma importante característica do processo estocástico gerado pelo movimento Browniano fracionário é, sob certas circunstâncias, a presença de longa dependência.

Definição 2.26 (Movimento Browniano Fracionário)**.** Um *movimento Browniano fracionário* com parâmetro de Hurst ou índice de similaridade $0 < H < 1$ é um processo $B^H = \left\{B_t^H\right\}_{t\in\mathbb{R}}$ tal que

i) $\left\{B_t^H\right\}_{t\in\mathbb{R}}$ é um processo Gaussiano;

ii) $B_0^H = 0$;

iii) a função de autocovariância é dada por

$$\gamma_{B^H}(r,s) = \frac{\sigma^2}{2}\left(r^{2H} + s^{2H} - |r-s|^{2H}\right), \text{ onde } \sigma > 0. \tag{2.37}$$

Uma consequência da Definição 2.26 é que $B_t^H \sim \mathcal{N}\left(0, \sigma^2 t^{2H}\right)$, para todo $t \in \mathbb{R}$ e $\sigma > 0$. Observe que para $t = 1$, temos $\sigma^2 = Var\left[B^H(1)\right]$. O resultado a seguir apresenta uma forte ligação entre os processos auto-similares estacionários de variança finita e o processo Browniano fracionário.

Lema 2.1. *Seja $\{X(t)\}_{t\in\mathbb{R}}$ um processo H-sssi de variância finita não degenerado. Então,*

i) $0 < H \leq 1$;

ii) $X(0) = 0$ *quase sempre;*

iii) a função de autocovariância é dada por

$$\gamma_X(r,s) = \frac{Var(X(1))}{2}\left(r^{2H} + s^{2H} - |r-s|^{2H}\right); \tag{2.38}$$

iv) se $0 < H < 1$, *então* $E(X(t)) = 0$, *para todo* $t \in R$;

v) se $H = 1$, *temos que* $X(t) = tX(1)$.

Demonstração: Ver Samorodnitsky e Taqqu (1994), página 319. □

Assim, é possivel definir o processo movimento Browniano fracionário com base nos processos H-sssi, equivalente à Definição 2.26.

Definição 2.27. Um processo Gaussiano H-sssi, com $0 < H < 1$ é chamado *movimento Browniano fracionário* com parâmetro de Hurst ou índice de similaridade H, e é denotado por $\left\{B_t^H\right\}_{t\in\mathbb{R}}$.

Proposição 2.9. *Seja $B^H = \left\{B_t^H\right\}_{t\in\mathbb{R}}$ o processo movimento Browniano fracionário. Então,*

i) os incrementos de B^H são estacionários;

ii) as trajetórias de B^H são não diferenciáveis quase certamente. No entanto, são Hölder contínuas quase sempre, ou seja, existe uma constante $c > 0$ tal que

$$|B_r^H - B_s^H| \leq c|r-s|^{H-\varepsilon}, \text{ para todo } \varepsilon > 0;$$

iii) se $H \in \left(\frac{1}{2}, 1\right)$, então B^H apresenta longa dependência.

Demonstração: Ver Nualart (2006). □

Proposição 2.10. *Seja $\{Z_n\}_{n\geq 1}$ o processo formado a partir dos incrementos de um processo movimento Browniano fracionário, onde, para cada $n \geq 1$, $Z_n = B_n^H - B_{n-1}^H$. Então,*

i) $Z_n \sim \mathcal{N}(0, \sigma^2)$;

ii) o processo $\{Z_n\}_{n\geq 1}$ é estacionário e possui função de autocovariância dada por

$$\gamma_z(h) = \frac{\sigma^2}{2}\left[(h+1)^{2H} + (h-1)^{2H} - 2h^{2H}\right];$$

iii) para $H > \frac{1}{2}$, o processo $\{Z_n\}_{n\geq 1}$ possui longa dependência.

Demonstração:

i) Da distribuição de B_t^H, decorre que $\{Z_n\}_{n\geq 1}$ é um processo Gaussiano com média zero e variância dada por

$$\begin{aligned} Var(Z_n) &= \mathbb{E}\left[\left(B_n^H - B_{n-1}^H\right)^2\right] \\ &= \mathbb{E}\left[\left(B_n^H\right)^2\right] + \mathbb{E}\left[\left(B_{n-1}^H\right)^2\right] - 2\mathbb{E}\left(B_n^H B_{n-1}^H\right) \\ &= \sigma^2 n^{2H} + \sigma^2 (n-1)^{2H} \\ &\quad -2\frac{\sigma^2}{2}\left(n^{2H} + (n-1)^{2H} - |n-(n-1)|^{2H}\right) \\ &= \sigma^2. \end{aligned}$$

Ou seja, $Z_n \sim \mathcal{N}(0, \sigma^2)$, para todo $n \geq 1$.

ii) A função de autocovariância do processo $\{Z_n\}_{n\geq 1}$ é dada por

$$\begin{aligned}
\gamma_{_Z}(h) &= \mathbb{E}(Z_n Z_{n+h}) = \mathbb{E}\left[\left(B_n^H - B_{n-1}^H\right)\left(B_{n+h}^H - B_{n+h-1}^H\right)\right] \\
&= \mathbb{E}\left(B_n^H B_{n+h}^H - B_n^H B_{n+h-1}^H - B_{n-1}^H B_{n+h}^H\right) \\
&\quad +\mathbb{E}\left(B_{n-1}^H B_{n+h-1}^H\right) \\
&= \mathbb{E}\left(B_n^H B_{n+h}^H\right) - \mathbb{E}\left(B_n^H B_{n+h-1}^H\right) - \mathbb{E}\left(B_{n-1}^H B_{n+h}^H\right) \\
&\quad +\mathbb{E}\left(B_{n-1}^H B_{n+h-1}^H\right) \\
&= \gamma_{_{B^H}}(n, n+h) - \gamma_{_{B^H}}(n, n+h-1) - \gamma_{_{B^H}}(n-1, n+h) \\
&\quad +\gamma_{_{B^H}}(n-1, n+h-1) \\
&= \frac{\sigma^2}{2}\left(n^{2H} + (n+h)^{2H} - h^{2H}\right) \\
&\quad -\frac{\sigma^2}{2}\left(n^{2H} + (n+h-1)^{2H} - (h-1)^{2H}\right) \\
&\quad -\frac{\sigma^2}{2}\left[(n-1)^{2H} + (n+h)^{2H} + (h+1)^{2H}\right] \\
&\quad +\frac{\sigma^2}{2}\left[(n-1)^{2H} + (n+h-1)^{2H} - h^{2H}\right] \\
&= \frac{\sigma^2}{2}\left[(h+1)^{2H} + (h-1)^{2H} - 2h^{2H}\right].
\end{aligned}$$

Do item *i)*, decorre que $\{Z_n\}_{n\geq 1}$ é um processo Gaussiano fracamente estacionário, ou seja, $\{Z_n\}_{n\geq 1}$ é estacionário.

iii) Aplicando a expansão em polinômios de Taylor na função de autocovariância do processo $\{Z_n\}_{n\geq 1}$, obtemos

$$\begin{aligned}
(h+1)^{2H} &= h^{2H} + 2Hh^{2H-1} + H(2H-1)h^{2H-2} + r_1(h), \\
(h-1)^{2H} &= h^{2H} - 2Hh^{2H-1} + H(2H-1)h^{2H-2} + r_2(h),
\end{aligned}$$

onde $\frac{r_i(h)}{h} \to 0$, quando $h \to \infty$, para $i = 1, 2$. Então, obtemos que

$$\begin{aligned}
\gamma_{_Z}(h) &= \frac{\sigma^2}{2}\left[(h+1)^{2H} + (h-1)^{2H} - 2h^{2H}\right] \\
&\approx \sigma^2 H(2H-1)h^{2H-2} \to 0, \text{ quando } h \to \infty.
\end{aligned}$$

Logo, se $H > \frac{1}{2}$, temos que $\sum_{h=1}^{\infty} \gamma_Z(h) = \infty$, ou seja, o processo $\{Z_n\}_{n\geq 1}$ possui longa dependência. □

Proposição 2.11. *Seja $B^H = \left\{B_t^H\right\}_{t\in\mathbb{R}}$ o processo movimento Browniano fracionário. Se*

i) $H > \frac{1}{2}$, então os incrementos de B^H são positivamente correlacionados;

ii) $H < \frac{1}{2}$, então os incrementos de B^H são negativamente correlacionados;

iii) $H = \frac{1}{2}$, então B^H é o movimento Browniano visto na Seção 2.4.

Demonstração: Decorre da demonstração da Proposição 2.10, item iii), onde obtemos que

$$\gamma_Z(h) \approx \sigma^2 H(2H-1) h^{2H-2} \to 0, \text{ quando } h \to \infty.$$

□

Observação 2.4. O parâmetro H vinculado ao movimento Browniano fracionário possui relação com o parâmetro de longa dependência d visto em (2.19). Mais precisamente, $H = \frac{1}{2} + d$.

2.9 Martingales

A noção de martingales é muito importante para compreendermos e fundamentarmos a integral estocástica de Itô. Assumimos que $\{\mathcal{F}_t\}_{t\geq 0}$ é uma coleção de σ-álgebras no mesmo espaço Ω e que $\mathcal{F}_t \subset \mathcal{F}$, para todo $t \in \mathbb{R}_+$, onde $\mathcal{F}$ é σ-álgebra em Ω.

Definição 2.28 (Filtragem)**.** Uma coleção $\{\mathcal{F}_t\}_{t\geq 0}$ de σ-álgebras em Ω é uma *filtragem* se

$$\mathcal{F}_s \subset \mathcal{F}_t, \text{ para todo } 0 \leq s \leq t.$$

Definição 2.29 (Processo Adaptado a uma Filtragem)**.** Um processo estocástico $\{X_t\}_{t\geq 0}$ é dito ser um *processo adaptado a uma filtragem* $\{\mathcal{F}_t\}_{t\geq 0}$ se

$$\sigma\left(X_t\right) \subset \mathcal{F}_t, \text{ para todo } t \geq 0,$$

onde $\sigma\left(X\right)$ é a σ-álgebra gerada por X.

Ou seja, se um processo $\{X_t\}_{t\geq 0}$ é adaptado a uma filtragem $\{\mathcal{F}_t\}_{t\geq 0}$, isso significa que X_t não carrega mais informações que $\mathcal{F}_t$.

O processo estocástico $\{X_t\}_{t\geq 0}$ é sempre adaptado à *filtragem natural gerada* por si mesmo, isto é, $\{X_t\}_{t\geq 0}$ é adaptado à filtragem $\{\mathcal{F}_t\}_{t\geq 0}$, onde cada $\mathcal{F}_t$ é dado por

$$\mathcal{F}_t = \sigma\left(X_s, s \leq t\right).$$

Por exemplo, se $\{B_t\}_{t\geq 0}$ é o processo movimento Browniano e $\{\mathcal{F}_t\}_{t\geq 0}$ a filtragem natural correspondente, então todo processo estocástico da forma

$$X_t = f\left(t, B_t\right), \text{ para todo } t \geq 0,$$

é adaptado à filtragem $(\mathcal{F}_t, t \geq 0)$, onde f é função de duas variáveis.

Se um processo estocástico $\{X_t\}_{t\geq 0}$ é adaptado à filtragem natural Browniana $\{\mathcal{F}_t\}_{t\geq 0}$, dizemos que $\{X_t\}_{t\geq 0}$ é *adaptado ao movimento Browniano.* Isso significa que $\{X_t\}_{t\geq 0}$ é função de B_s, para todo $s \leq t$.

Definição 2.30 (Martingale à Tempo Discreto)**.** O processo estocástico $\{X_n\}_{n\in\mathbb{Z}_+}$ é chamado *martingale à tempo discreto com respeito à filtragem* $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$ se

i) $\mathbb{E}\left(|X_n|\right) < \infty$, para todo $n \in \mathbb{Z}_+$.

ii) $\{X_n\}_{n\in\mathbb{Z}_+}$ é adaptado à $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$.

iii) $\mathbb{E}\left(X_{n+k}|\mathcal{F}_n\right) = X_n$, para todo $n, k \in \mathbb{Z}_+$, isto é, X_n é o melhor preditor de X_{n+k} dado $\mathcal{F}_n$.

Na verdade, no item iii) da Definição 2.30 é suficiente requerer

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

De fato, invocando a igualdade (2.6), obtemos que

$$\begin{aligned}\mathbb{E}\left(X_{n+1}|\mathcal{F}_n\right) &= \mathbb{E}\left(\mathbb{E}\left(X_{n+2}|\mathcal{F}_{n+1}\right)|\mathcal{F}_n\right) = \mathbb{E}\left(X_{n+2}|\mathcal{F}_n\right) \\ &= \mathbb{E}\left(\mathbb{E}\left(X_{n+3}|\mathcal{F}_{n+2}\right)|\mathcal{F}_n\right) = \mathbb{E}\left(X_{n+3}|\mathcal{F}_n\right) \\ &= \cdots = \mathbb{E}\left(X_{n+k}|\mathcal{F}_n\right). \end{aligned} \tag{2.39}$$

A soma de variáveis aleatórias independentes com média zero é um clássico exemplo de martingale à tempo discreto.

Exemplo 2.14. Sejam $X_0 = 0$ e $X_1, X_2, \cdots$ variáveis aleatórias independentes, com $\mathbb{E}\left(|X_k|\right) < \infty$ e $\mathbb{E}\left(X_k\right) = 0$, para todo $k \geq 0$. Sejam $S_0 = 0$ e $S_n = \sum_{i=1}^{n} X_i$, para $n \geq 1$. Então, considerando

$$\mathcal{F}_m = \sigma\left(S_r, r \leq m\right),$$

temos que $\{S_i\}_{i=0}^{n}$ é martingale à tempo discreto com respeito à filtragem $\{\mathcal{F}_i\}_{i=0}^{n}$. De fato, temos que

$$\mathbb{E}\left(|S_n|\right) \leq \mathbb{E}\left(|X_1|\right) + \cdots + \mathbb{E}\left(|X_n|\right) < \infty,$$

o que prova o item i). Ainda,

$$\begin{aligned}\mathbb{E}\left(S_{n+1}|\mathcal{F}_n\right) &= \mathbb{E}\left(S_n + X_{n+1}|\mathcal{F}_n\right) = \mathbb{E}\left(S_n|\mathcal{F}_n\right) + \mathbb{E}\left(X_{n+1}|\mathcal{F}_n\right) \\ &= S_n + \mathbb{E}\left(X_{n+1}\right) \\ &= S_n \end{aligned}$$

onde a terceira igualdade vale pela independência das variáveis X_i, para todo $i \in \mathbb{Z}_+$.

Podemos facilmente estender a idéia de martingale à tempo contínuo.

Definição 2.31 (Martingale à Tempo Contínuo)**.** Um processo $X = \{X_t\}_{t\geq 0}$ é chamado *martingale à tempo contínuo com respeito à filtragem* $\{\mathcal{F}_t\}_{t\geq 0}$, e denotado por $(X, (\mathcal{F}_t))$, se

i) $\mathbb{E}(|X_t|) < \infty$, para todo $t \geq 0$.

ii) X é adaptado a $\{\mathcal{F}_t\}_{t\geq 0}$.

iii) $\mathbb{E}(X_t|\mathcal{F}_s) = X_s$, para todo $0 \leq s < t$, isto é, X_s é o melhor preditor de X_t dado $\mathcal{F}_s$.

Uma importante propriedade é que todo martingale tem esperança constante. De fato, usando o item iii) da Definição 2.31, temos que $\mathbb{E}(X_t|\mathcal{F}_s) = X_s$, para $s < t$, e pela igualdade (2.5), obtemos

$$\mathbb{E}(X_s) = \mathbb{E}(\mathbb{E}(X_t|\mathcal{F}_s)) = \mathbb{E}(X_t), \text{ para todo } s \text{ e } t. \tag{2.40}$$

Um clássico exemplo de martingale é o movimento Browniano.

Exemplo 2.15. Seja $B = \{B_t\}_{t\geq 0}$ o processo movimento Browniano dado na Definição 2.20. Seja

$$\mathcal{F}_t = \sigma(B) = \sigma(B_s, s \leq t).$$

Essa é a menor σ-álgebra contendo as informações essenciais sobre a estrutura de B em $[0, t]$. Queremos calcular $\mathbb{E}(B_t|\mathcal{F}_s) = \mathbb{E}(B_t|B_k, k \leq s)$, para $s \leq t$, isto é,

$$\mathbb{E}(B_t|\mathcal{F}_s) = \mathbb{E}((B_t - B_s) + B_s|\mathcal{F}_s) = \mathbb{E}(B_t - B_s|\mathcal{F}_s) + \mathbb{E}(B_s|\mathcal{F}_s). \tag{2.41}$$

Como o movimento Browniano possui incrementos independentes, então temos que $(B_t - B_s)$ é independente de $(B_s - B_0) = B_s$. Logo, $\mathcal{F}_s$ não acrescenta nenhuma informação em $(B_t - B_s)$, o que implica que $(B_t - B_s)$ é independente de $\mathcal{F}_s$. Logo, $\mathbb{E}(B_t - B_s|\mathcal{F}_s) = \mathbb{E}(B_s - B_s) = 0$. Assim, a expressão (2.41) pode ser reescrita como

$$\mathbb{E}\left(B_t|\mathcal{F}_s\right) = \mathbb{E}\left(B_s|\mathcal{F}_s\right) = B_s.$$

Como $\mathbb{E}\left(B_t\right) = 0$, para todo $t \geq 0$, então, de fato, o movimento Browniano é uma martingale à tempo contínuo.

2.10 Integral de Riemann-Stieltjes

A partir desta seção, sempre que falarmos em uma partição $\tau_n = \{s_0, s_1, s_2, \cdots, s_n\}$ de um intervalo real finito $[a, b]$, estamos convencionando que $a = s_0 < s_1 < s_2 < \cdots < s_n = b$.

Seja $\tau_n = \{s_0, s_1, s_2, \cdots, s_n\}$ uma partição de tamanho n do intervalo $[0, t]$ e seja $\pi_n = \{y_1, y_2, \cdots, y_n\}$ uma partição intermediária de τ_n tal que

$$\pi_n : s_{i-1} \leq y_i \leq s_i, \text{ para } i = 1, \cdots, n.$$

Dadas f e g duas funções reais definidas em $[0, t]$, definimos a *soma de Riemann-Stieltjes* correspondente a τ_n e π_n como sendo

$$S_n := S_n(\tau_n, \pi_n) = \sum_{i=1}^{n} f\left(y_i\right)\left[g\left(s_i\right) - g\left(s_{i-1}\right)\right]. \tag{2.42}$$

Se existe o limite

$$S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{i=1}^{n} f\left(y_i\right)\left[g\left(s_i\right) - g\left(s_{i-1}\right)\right],$$

quando $\sup\limits_{0\leq j\leq n} |s_j - s_{j-1}| \to 0$, e S é independente da escolha da partição τ_n e da sua partição intermediária π_n, então S é chamada de *integral de Riemann-Stieltjes de f com respeito a g em* $[0, t]$. Escrevemos

$$S = \int_0^t f\left(s\right) dg\left(s\right). \tag{2.43}$$

Nos interessa saber agora quais são as condições suficientes para a existência da integral de Riemann-Stieltjes. Primeiramente, estendemos a definição de função de variação limitada.

Definição 2.32 (Função de p-Variação Limitada)**.** Uma função real $f : [0, t] \to \mathbb{R}$ tem *variação p-limitada* para algum $p > 0$ se

$$\sup_{\tau} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|^p < \infty,$$

onde o supremo é tomado sobre todas as partições τ de $[0, t]$.

Observação 2.5. Note que no caso especial em que $p = 1$, f é dita ser de *variação limitada*.

Em geral, na literatura é apresentado que a condição suficiente para a existência da integral de Riemann-Stieltjes (2.43) é que f e g não tenham pontos de descontinuidade em comum no intervalo $[0, t]$, f seja contínua e g seja de variação limitada.

Observação 2.6. A variação limitada de g não é um critério necessário para a existência da integral de Riemann-Stieltjes (2.43).

No entanto, nesse trabalho, estamos interessados no caso em que $g(x) = B_x$, isto é, que o integrador seja um movimento Browniano. A proposição a seguir mostra que o movimento Browniano não é um processo de variação limitada.

Proposição 2.12. *Seja* $\{B_t\}_{t \geq 0}$ *um processo movimento Browniano. Então,*

$$\sup_{\tau} \sum_{i=1}^{n} |B_{s_i} - B_{s_{i-1}}| = \infty,$$

onde o supremo é tomado sobre todas as partições τ *de* $[0, t]$.

Demonstração: Ver Mikosch (1998), página 189. □

Logo, precisamos de uma condição mais forte para justificar a existência da integral de Riemann-Stieltjes para o caso em que g é um movimento Browniano. Nesse caso, Dudley e Norvaisa (1998) nos apresentam um critério mais abrangente para averiguar a existência da integral de Riemann-Stieltjes.

Proposição 2.13. *A integral de Riemann-Stieltjes, dada em* (2.43)*, existe se as seguintes condições estão satisfeitas*

i) As funções f e g não têm descontinuidade no mesmo ponto $x \in [0,t]$;

ii) A função f tem variação p-limitada e a função g tem variação q-limitada, para algum $p > 0$ e $q > 0$, tais que $p^{-1} + q^{-1} > 1$.

Demonstração: Ver Dudley e Norvaisa (1998), página 136. □

A partir dessa nova condição suficiente, conseguimos provar a existência da integral de Riemann-Stieltjes para diversos casos quando temos um movimento Browniano como integrador.

Proposição 2.14. *Seja f uma função determinística e diferenciável, com derivada limitada em $[0,t]$. Então, existe a integral de Riemann-Stieltjes*

$$\mathcal{I}(t) = \int_0^t f(s)\, dB_s(\omega), \tag{2.44}$$

para toda trajetória amostral Browniana $B_t(\omega)$.

Demonstração: É conhecido que trajetórias aleatórias do movimento Browniano tem variação p-limitada em qualquer intervalo finito, desde que $p > 2$, e ilimitada para $p \leq 2$ (ver Taylor, 1972).

Então, pelas condições de existência da integral de Riemann-Stieltjes, dada $f(\cdot)$ uma função determinística em $[0,t]$, podemos definir a integral $\mathcal{I}(t)$ desde que f seja de q-variação limitada, para $q < 2$, ou seja, basta que f seja de variação limitada ($q = 1$).

Suponhamos que f seja função diferenciável, com derivada limitada em $[0,t]$. Logo, pelo teorema do valor médio, existe uma constante $K > 0$ tal que

$$|f(s) - f(r)| \leq K(s-r), \text{ para todo } r, s \in \mathbb{R}, r < s.$$

Então,

$$\sup_{\tau} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \leq K \sum_{i=1}^{n} (x_i - x_{i-1}) = K(x_n - x_0) = Kt < \infty,$$

já que $x_0 = 0$ e $x_n = t$, provando que f é de variação limitada. Portanto, a integral $\mathcal{I}(\cdot)$ dada por (2.44) existe. □

2.11 Integral de Itô

As Proposições 2.13 e 2.14 sugerem que podemos definir a integral de Riemann-Stieltjes $\mathcal{I}(\cdot)$ para diversas funções f. No entanto, há casos em que não é possível o cálculo da integral de Riemann-Stieltjes. A seguir apresentamos um exemplo clássico, quando o integrando e o integrador são um processo movimento Browniano.

Exemplo 2.16. Denote $B = \{B_t\}_{t\geq 0}$ um processo movimento Browniano. Desejamos calcular

$$\mathcal{I}(t) = \int_0^t B_s(\omega)\, dB_s(\omega). \tag{2.45}$$

Consideramos $\tau_n = \{t_0, t_1, t_2, \cdots, t_n\}$ uma partição de tamanho n do intervalo $[0, t]$, e seja $\pi_n = \{y_1, y_2, \cdots, y_n\}$ uma partição intermediária de τ_n tal que

$$\pi_n : t_{i-1} \leq y_i \leq t_i, \text{ para } i = 1, \cdots, n.$$

Como podemos selecionar qualquer partição intermediária, consideramos $y_i = t_i$ para $i = 1, \cdots, n$, ou seja, tomamos os pontos à direita da partição. Logo, $\pi_n = \{t_1, t_2, \cdots, t_n\}$. Obtemos a soma de Riemann-Stieltjes correspondente, dada por

$$S_n := S_n(\tau_n, \pi_n) = \sum_{i=1}^{n} B_{t_i}\left(B_{t_i} - B_{t_{i-1}}\right).$$

Suponha que existe o limite

$$S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{i=1}^{n} B_{t_i}\left(B_{t_i} - B_{t_{i-1}}\right),$$

quando $\sup_{0\leq i\leq n} |t_i - t_{i-1}| \to 0$. Dizemos que $S = S(\tau_n, \pi_n)$ é a integral de Riemann-Stieltjes de $\mathcal{I}(\cdot)$ quando o valor de S independe da escolha da partição τ_n e sua partição intermediária π_n.

Aplicando as propriedades da esperança de uma variável aleatória, as propriedades do movimento Browniano e o Teorema 2.2 (Convergência de Lebesgue),

obtemos que

$$\begin{aligned}
\mathbb{E}(S) &= \mathbb{E}\left[\lim_{n\to\infty}\sum_{i=1}^{n} B_{t_i}\left(B_{t_i}-B_{t_{i-1}}\right)\right] = \lim_{n\to\infty}\mathbb{E}\left[\sum_{i=1}^{n} B_{t_i}\left(B_{t_i}-B_{t_{i-1}}\right)\right] \\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}\left[B_{t_i}\left(B_{t_i}-B_{t_{i-1}}\right)\right] \\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\left[\mathbb{E}\left(B_{t_i}^2\right)-\mathbb{E}\left(B_{t_i}B_{t_{i-1}}\right)\right] \\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\left[Var\left(B_{t_i}\right)-\gamma_B\left(t_i,t_{i-1}\right)\right] = \lim_{n\to\infty}\sum_{i=1}^{n}\left(t_i-t_{i-1}\right) \\
&= \lim_{n\to\infty}\left(t_n-t_0\right)=t.
\end{aligned}$$

Considerando a mesma partição $\tau_n = \{t_0, t_1, \cdots, t_n\}$ adotada anteriormente, definimos uma partição intermediária $\overline{\pi}_n = \{t_0, \cdots, t_{n-1}\}$, ou seja, tomamos os pontos à esquerda de cada intervalo $[t_{i-1}, t_i]$, para $i = 1, \cdots, n$.

Suponha que existe o limite

$$\overline{S} = \lim_{n\to\infty}\overline{S}_n = \lim_{n\to\infty}\sum_{i=1}^{n} B_{t_{i-1}}\left(B_{t_i}-B_{t_{i-1}}\right).$$

Aplicando as propriedades da esperança de uma variável aleatória, as propriedades do movimento Browniano e o Teorema 2.2 (Convergência de Lebesgue), obtemos que

$$\begin{aligned}
\mathbb{E}\left(\overline{S}\right) &= \mathbb{E}\left[\lim_{n\to\infty}\sum_{i=1}^{n} B_{t_{i-1}}\left(B_{t_i}-B_{t_{i-1}}\right)\right] = \lim_{n\to\infty}\mathbb{E}\left[\sum_{i=1}^{n} B_{t_{i-1}}\left(B_{t_i}-B_{t_{i-1}}\right)\right] \\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}\left[B_{t_{i-1}}\left(B_{t_i}-B_{t_{i-1}}\right)\right] \\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}\left[\left(B_{t_{i-1}}-B_0\right)\left(B_{t_i}-B_{t_{i-1}}\right)\right] \\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}\left(B_{t_{i-1}}-B_0\right)\mathbb{E}\left(B_{t_i}-B_{t_{i-1}}\right)=0.
\end{aligned}$$

Logo, temos que $S \neq \overline{S}$, ou seja, o valor de S depende da partição intermediária escolhida. Portanto, $\mathcal{I}(\cdot)$ não pode ser calculada pela integral de Riemann-Stieltjes.

Este exemplo nos mostra a necessidade de introduzirmos um novo método de integração, que contemple um número maior de casos. Há uma grande variedade de integrais estocásticas que solucionam esse tipo de questão. Apresentamos, neste trabalho, as integrais de Itô.

Vamos definir a integral de Itô inicialmente para um tipo de processo chamado *processo simples*. No Lema 2.2, mostramos que dada uma função que cumpra certas condições, existe uma sequência de processos simples que converge em média quadrática para ela. Assim, definimos a integral de Itô para essa função como o limite das integrais dessa sequência.

2.11.1 Integral de Itô para Processos Simples

Seja $B = \{B_t\}_{t\geq 0}$ o processo movimento Browniano. Seja

$$\mathcal{F}_t = \sigma\left(B_s, s \leq t\right),\, t \geq 0,$$

a filtragem natural correspondente. Todos os processos a partir daqui serão considerados no intervalo $[0, T]$.

Definição 2.33 (Processo Simples)**.** O processo $C = \{C_t\}_{t\in[0,T]}$ é dito *simples* se existe uma partição $\tau_n = \{s_0, s_1, \cdots, s_n\}$ de $[0, T]$ e uma sequência $\{Z_i\}_{i=1}^n$ de variáveis aleatórias tais que, para cada $t \in [0, T]$, o processo é dado por

$$C_t = \begin{cases} Z_n,\ se\, t = T \\ \\ Z_i,\ se\, s_{i-1} \leq t < s_i,\ i = 1, 2, \cdots, n. \end{cases}$$

A sequência $\{Z_i\}_{i=1}^n$ é adaptada à $\{\mathcal{F}_t\}_{t\in[0,T]}$, isto é, Z_i é função do movimento Browniano até t_{i-1}, e satisfaz $\mathbb{E}\left(Z_i^2\right) < \infty$.

Exemplo 2.17. Considere a função determinística

$$f_n\left(t\right) = \begin{cases} \dfrac{n-1}{n},\ se\, t = T \\ \\ \dfrac{i-1}{n},\ se\, \dfrac{i-1}{n} \leq t < \dfrac{i}{n},\ \text{para todo } i = 1, \cdots, n. \end{cases}$$

Observe que $f(\cdot)$ é uma função escada. Logo, $C = \{f_n(t)\}_{t\in[0,T]}$ é um processo simples.

Exemplo 2.18. A partir do processo movimento Browniano $B = \{B_t\}_{t\in[0,T]}$, definimos o processo

$$C_t = \begin{cases} B_{t_{n-1}},\ se\ t = T, \\ B_{t_{i-1}},\ se\ t_{i-1} \leq t \leq t_i,\ \text{para todo } i = 1, \ldots, n, \end{cases}$$

para uma partição τ_n de $[0, T]$. Este é um processo simples. De fato, suas trajetórias são constantes por partes, e C_t é uma função do movimento Browniano até o tempo t.

Definição 2.34 (Integral de Itô para um Processo Simples)**.** A *integral de Itô para um processo simples* C em $[0, t]$, $t_{k-1} \leq t \leq t_k$ é dada por

$$I_t(C) \equiv \int_0^t C_s\, dB_s := \sum_{i=1}^{k-1} Z_i \nabla_i(B) + Z_k\left(B_t - B_{t_{k-1}}\right), \tag{2.46}$$

onde $\nabla_i(B) := B_{t_i} - B_{t_{i-1}}$, para $i = 1, \cdots, k-1$.

Portanto, o valor da integral (2.46) nada mais é do que a soma de Riemann-Stieltjes do processo C, calculada nos pontos à esquerda dos intervalos da partição $[t_{i-1}, t_i]$, com respeito ao movimento Browniano.

Proposição 2.15. *O processo estocástico* $\{I_t(C)\}_{t\in[0,T]}$ *dado em* (2.46) *é um martingale com respeito à filtragem natural Browniana* $\{\mathcal{F}_t\}_{t\in[0,T]}$.

Demonstração: Basta provar os três itens da Definição 2.31. Temos que, pela expressão (2.46),

$$\mathbb{E}(I_t(C)) = \mathbb{E}\left(\int_0^t C_s\, dB_s\right) = \mathbb{E}\left(\sum_{i=1}^{k-1} Z_i \nabla_i(B) + Z_k\left(B_t - B_{t_{k-1}}\right)\right).$$

Como Z_i e $\nabla_i(B)$ são independentes, temos que

$$\mathbb{E}(I_t(C)) = \sum_{i=1}^{k-1} \mathbb{E}(Z_i \nabla_i(B)) + \mathbb{E}\left[Z_k\left(B_t - B_{t_{k-1}}\right)\right] = 0,$$

pois os incrementos do movimento Browniano tem esperança zero. Logo, o item i) da Definição 2.31 está satisfeito.

Agora, notamos que, para $s, t \in [t_{k-1}, t_k]$, $s < t$, temos

$$\begin{aligned} I_t(C) &= \int_0^t C_s\, dB_s = \sum_{i=1}^{k-1} Z_i \nabla_i(B) + Z_k\left(B_t - B_{t_{k-1}}\right) \\ &= I_{k-1}(C) + Z_k\left(B_t + B_s - B_s - B_{t_{k-1}}\right) \\ &= I_{k-1}(C) + Z_k\left(B_s - B_{t_{k-1}}\right) + Z_k(B_t - B_s) = I_s(C) + Z_k(B_t - B_s), \end{aligned}$$

onde $B_t - B_s$ é independente de $\mathcal{F}_s$ e $I_s(C)$ e Z_k são funções do movimento Browniano até o tempo s. Então,

$$\begin{aligned} \mathbb{E}\left(I_t(C)\,|\mathcal{F}_s\right) &= \mathbb{E}\left[I_s(C) + Z_k(B_t - B_s)|\,\mathcal{F}_s\right] = \mathbb{E}\left[I_s(C) + Z_k(B_t - B_s)\right] \\ &= I_s(C) + Z_k\,\mathbb{E}(B_t - B_s) = I_s(C), \end{aligned}$$

e o item iii) da Definição 2.31 está satisfeito.

Já o item ii) vem do fato que, até o tempo t, as variáveis aleatórias $\{Z_i\}_{i=1}^k$, $\{\nabla_i(B)\}_{i=1}^k$, e $B_t - B_{t_{k-1}}$ em (2.46) são todas funções do movimento Browniano até o tempo t. □

Da demonstração da Proposição 2.15, extraímos um resultado que nos será de grande importância posteriormente.

Corolário 2.3. *A integral estocástica $I_t(\cdot)$ de Itô, definida em* (2.46)*, tem esperança zero.*

A seguir, apresentamos algumas propriedades básicas da integral estocástica de Itô para processos simples.

Teorema 2.3. *Sejam c_1 e c_2 constantes reais e $C^{(1)}$ e $C^{(2)}$ processos simples em [0,T]. Então, valem as seguintes propriedades.*

i) A integral estocástica de Itô é linear*, isto é,*

$$\int_0^t \left[c_1 C_s^{(1)} + c_2 C_s^{(2)}\right] dB_s = c_1 \int_0^t C_s^{(1)}\, dB_s + c_2 \int_0^t C_s^{(2)}\, dB_s.$$

ii) A integral estocástica de Itô é linear em intervalos adjacentes*, isto é,*

$$\int_0^T C_s\, dB_s = \int_0^t C_s\, dB_s + \int_t^T C_s\, dB_s.$$

iii) A integral estocástica de Itô para processos simples satisfaz a propriedade da isometria

$$\mathbb{E}\left(\int_0^t C_s\, dB_s\right)^2 = \int_0^t \mathbb{E}\left({C_s}^2\right)\, ds, \text{ para todo } t \in [0,T]. \tag{2.47}$$

iv) O processo $I_t(C)$ é contínuo.

Demonstração: Ver Mikosch (1998), página 106. □

2.11.2 Caso Geral da Integral de Itô

Na seção anterior, introduzimos a integral estocástica de Itô para processos simples, isto é, para processos cujas trajetórias são funções escada. No que segue, seja $C = \{C_t\}_{t\in[0,T]}$ o processo integrador.

Suposição 2.1. Vamos supor que as seguintes condições sobre C estão satisfeitas:

H1. C é adaptado ao movimento Browniano em $[0,T]$, isto é, C_t é função de B_s, para todo s, t, com $s \leq t$;

H2. A integral $\int_0^T \mathbb{E}\left(C_s^2\right)\, ds$ é finita.

Observe que as condições acima são trivialmente satisfeitas pelos processos simples. Outra classe de integrandos admissíveis consiste na das funções determinísticas $c(\cdot)$ em $[0,T]$ tais que $\int_0^T c^2(t)\, dt < \infty$. Isso inclui as funções contínuas em $[0,T]$.

O Lema 2.2 a seguir nos garante que os processos simples $\left\{C^{(n)}\right\}_{n\in\mathbb{N}}$ convergem, em média quadrática, para C. Como $C^{(n)}$ são processos simples, para cada $n \in \mathbb{N}$, então podemos calcular a integral estocástica $I_t\left(C^{(n)}\right)$ de Itô, dada em (2.46), para todo $n \in \mathbb{N}$ e todo $t \in [0,T]$. Mais ainda, a expressão (2.48) nos garante que $\left\{C^{(n)}\right\}_{n\in\mathbb{N}}$ é uma sequência de Cauchy em L^2.

Lema 2.2. *Se* $C = \{C_t\}_{t\in[0,T]}$ *é um processo satisfazendo as condições H1 e H2 da Suposição 2.1, então existe uma sequência* $\left\{C^{(n)}\right\}_{n\in\mathbb{N}}$ *de processos simples tal que*

$$\int_0^T \mathbb{E}\left[C_s - C_s^{(n)}\right]^2 ds \to 0. \tag{2.48}$$

Demonstração: Ver Kloeden e Platen (1992), lema 3.2.1. □

O Lema 2.5 apresenta uma cota superior para

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\left[I_t\left(C^{(n+k)}\right) - I_t\left(C^{(n)}\right)\right]^2\right),$$

que será útil na aplicação da desigualdade de Chebyshev a seguir.

Lema 2.3 (Desigualdade de Chebyshev)**.** *Seja* X *uma variável aleatória e* c *uma constante real tal que* $\mathbb{E}\left[(X-c)^2\right]$ *é finito. Então,*

$$\mathbb{P}\left[|X-c| \geq \varepsilon\right] \leq \frac{\mathbb{E}\,(X-c)^2}{\varepsilon^2}, \text{ para todo } \varepsilon > 0. \tag{2.49}$$

Demonstração: Ver Billingsley (1995), página 80. □

Lema 2.4 (Lema de Borel-Cantelli)**.** *Sejam* $A_1, A_2, \cdots$ *eventos aleatórios definidos em* $(\Omega, \mathcal{F}, \mathbb{P})$.

i) Se $\sum_{n=1}^{\infty}\mathbb{P}\,(A_n)$ *converge, então* $\mathbb{P}\left(lim\sup_n A_n\right) = 0$.

ii) Se $A_1, A_2, \cdots$ *são eventos aleatórios independentes e* $\sum_{n=1}^{\infty}\mathbb{P}\,(A_n)$ *diverge, então* $\mathbb{P}\left(lim\sup_n A_n\right) = 1$.

Demonstração: Ver Billingsley (1995), página 59. □

Lema 2.5. *Seja* $C = \{C_t\}_{t\in[0,T]}$ *um processo satisfazendo as condições H1 e H2 da Suposição 2.1. Seja* $\left\{C^{(n)}\right\}_{n\in\mathbb{N}}$ *uma sequência de processos simples tal que valha a expressão* (2.48)*. Então, para todo* k, $n \in \mathbb{N}$, *temos que*

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\left[I_t\left(C^{(n+k)}\right) - I_t\left(C^{(n)}\right)\right]^2\right) \leq 4\mathbb{E}\left(\int_0^T\left[C_t^{(n+k)} - C_t^{(n)}\right]^2 dt\right). \tag{2.50}$$

Demonstração: Denotamos $I(\cdot)$ como sendo o processo $\{I_t(\cdot)\}_{t\in[0,T]}$. Então, como $I\left(C^{(n+k)}\right)$ e $I\left(C^{(n)}\right)$ são martingales com respeito ao movimento Browniano. Assim, $I\left(C^{(n+k)}\right) - I\left(C^{(n)}\right) = I\left(C^{(n+k)} - C^{(n)}\right)$ é uma martingale com respeito a $\{\mathcal{F}_t\}_{t\geq 0}$, que é a filtragem natural do processo movimento Browniano. Para uma martingale $(M, (\mathcal{F}_t))$ em $[0,T]$, a inequação de Doob (ver Dellacherie e Meyer, 1982) nos diz que

$$\mathbb{E}\left(\sup_{0\leq t\leq T} M_t^2\right) \leq 4\,\mathbb{E}\,(M_T^2).$$

Em particular,

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\left[I_t\left(C^{(n+k)} - C^{(n)}\right)\right]^2\right) \leq 4\,\mathbb{E}[I_T\left(C^{(n+k)} - C^{(n)}\right)]^2. \tag{2.51}$$

Pela propriedade da isometria dos martingales, aplicado ao processo simples $I_T\left(C^{(n+k)} - C^{(n)}\right)$, temos que

$$\mathbb{E}[I_T\left(C^{(n+k)} - C^{(n)}\right)]^2 = \mathbb{E}\left(\int_0^T \left[C_t^{(n+k)} - C_t^{(n)}\right]^2 dt\right). \tag{2.52}$$

De (2.51) e (2.52), segue o resultado desejado. □

Como $\left\{C^{(n)}\right\}_{n\in\mathbb{N}}$ é uma sequência de Cauchy, isso implica que, em particular, o lado direito da desigualdade (2.50) converge a zero. Logo, podemos encontrar uma subsequência $(k_n)_{n\in\mathbb{N}}$ dos naturais tal que

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\left[I_t\left(C^{(k_{n+1})} - C^{(k_n)}\right)\right]^2\right) \leq \frac{1}{2^n}. \tag{2.53}$$

A desigualdade (2.53), em conjunto com a desigualdade de Chebyshev, implica que, para qualquer $\varepsilon > 0$,

$$\begin{aligned}
& \sum_{n=1}^{\infty} \mathbb{P}\left(\sup_{0\leq t\leq T} |I_t\left(C^{(k_{n+1})}\right) - I_t\left(C^{(k_n)}\right)| > \varepsilon\right) \\
\leq\ & \frac{1}{\varepsilon^2}\sum_{n=1}^{\infty} \mathbb{E}\left(\sup_{0\leq t\leq T}\left[I_t\left(C^{(k_{n+1})}\right) - I_t\left(C^{(k_n)}\right)\right]^2\right) \\
<\ & \frac{1}{\varepsilon^2}\sum_{n=1}^{\infty}\frac{1}{2^n} < \frac{1}{\varepsilon^2} < \infty.
\end{aligned}$$

Logo, pelo Lema de Borel-Cantelli, temos que o processo $\left\{I\left(C^{(k_n)}\right)\right\}_{n\in\mathbb{N}}$ converge no sentido L^2 uniformemente em $[0,T]$ para um processo estocástico $I(C)$ com probabilidade 1. Se tomarmos $k\to\infty$ em (2.50), temos que $I_t\left(C^{n+k}\right)\to I_t(C)$ e que $C_t^{n+k}\to C_t$. Logo, concluímos que

$$\mathbb{E}\left(\sup_{0\le t\le T}\left[I_t(C)-I_t\left(C^{(n)}\right)\right]^2\right)\le 4\mathbb{E}\left(\int_0^T\left[C_t-C_t^{(n)}\right]^2 dt\right). \tag{2.54}$$

Pelo Lema 2.2, o lado direito da desigualdade (2.54) converge a zero quando $n\to\infty$. Então, temos que o limite $I(C)$ não depende de uma escolha particular de uma subsequência de $\left\{C^{(n)}\right\}_{n\in\mathbb{N}}$.

Definição 2.35 (Integral Estocástica de Itô)**.** Seja $C=\{C_t\}_{t\in[0,T]}$ um processo satisfazendo as condições H1 e H2 da Suposição 2.1. O processo limite $I(C)=\{I_t(C)\}_{t\in[0,T]}=\lim_{n\to\infty} I\left(C^{(n)}\right)$ é chamado *integral estocástica de Itô de C*, e é denotado por

$$I_t(C)=\int_0^t C_s\,dB_s,\ \text{para cada } t\in[0,T].$$

Assim como no caso em que o integrando é uma função simples, no caso geral também temos que a integral estocástica de Itô é um martingale. Isto será mostrado na proposição a seguir.

Proposição 2.16. *O par* $(I(C),(\mathcal{F}_t))$ *constitui um martingale.*

Demonstração: Sabemos, pela Proposição 2.15, que $\left(I\left(C^{(n)}\right),(\mathcal{F}_t)\right)$ constitui um martingale, para todo $n\in\mathbb{N}$. Então, vale que

$$\mathbb{E}\left(I_t\left(C^{(n)}\right)|\mathcal{F}_s\right)=I_s\left(C^{(n)}\right),\ \text{para } s<t.$$

Assim, passando se necessário para uma subsequência $(k_n)_{n\in\mathbb{N}}$ tal que $I_t\left(C^{(k_n)}\right)$ converge uniformemente para $I_t(C)$, com probabilidade 1, obtemos

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\left(I_t\left(C^{(k_n)}\right)|\mathcal{F}_s\right) &= \lim_{n\to\infty}\left[I\left(C^{(k_n)}\right)\right]\\ &= I_s(C).\end{aligned}$$

Mas, pelo Teorema 2.2 (Convergência de Lebesgue), temos

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\left(I_t\left(C^{(k_n)}\right)|\mathcal{F}_s\right) &= \mathbb{E}\left(\lim_{n\to\infty}\left(I_t\left(C^{(k_n)}\right)|\mathcal{F}_s\right)\right)\\ &= \mathbb{E}\left(I_t\left(C\right)|\mathcal{F}_s\right)\end{aligned}$$

Logo,

$$\mathbb{E}\left(I_t\left(C\right)|\mathcal{F}_s\right) = I_s\left(C\right).$$

Portanto, $\left(I\left(C\right),\left(\mathcal{F}_t\right)\right)$ é um martingale. □

Uma importante característica da integral estocástica dada na Definição 2.35 é que ela herda diversas propriedades da integral de Itô para processos simples.

Proposição 2.17. *Seja $I\left(C\right)$ a integral estocástica de Itô dada na Definição 2.35. Então, valem as seguintes propriedades*

i) O processo estocástico $\left\{I_t\left(C\right)\right\}_{t\in[0,T]}$ tem esperança zero, para todo $t\in[0,T]$;

ii) A integral estocástica de Itô é linear*, isto é,*

$$\int_0^t\left[c_1C_s^{(1)}+c_2C_s^{(2)}\right]\,dB_s = c_1\int_0^t C_s^{(1)}\,dB_s + c_2\int_0^t C_s^{(2)}\,dB_s;$$

iii) A integral estocástica de Itô é linear em intervalos adjacentes*, isto é,*

$$\int_0^T C_s\,dB_s = \int_0^t C_s\,dB_s + \int_t^T C_s\,dB_s;$$

iv) A integral estocástica de Itô satisfaz a propriedade da isometria*,isto é,*

$$\mathbb{E}\left(\int_0^t C_s\,dB_s\right)^2 = \int_0^t\mathbb{E}\left({C_s}^2\right)\,ds,\ \textit{para todo } t\in[0,T]\,; \tag{2.55}$$

v) O processo $\left\{I_t\left(C\right)\right\}_{t\in[0,T]}$ é contínuo.

Demonstração: Ver Mikosch (1998), página 192. □

2.11.3 Lema de Itô

Na última seção vimos que a integral $I_t(C)$ está bem definida, mas, com excessão dos casos em que C_s é um processo simples, calcular essa integral pode ser uma tarefa um tanto árdua sem ferramentas adicionais. Nessa seção, apresentamos uma ferramenta que será bastante útil nessa tarefa, a saber, o Lema de Itô.

Vamos iniciar relembrando a regra da cadeia para funções determinísticas na forma integral.

Proposição 2.18. *Sejam f e g duas funções reais diferenciáveis. Então,*

$$f(g(t)) - f(g(0)) = \int_0^t f'(g(s))\, g'(s)\, ds = \int_0^t f'(g(s))\, dg(s),$$

onde f' e g' são as derivadas de f e g, respectivamente.

Demonstração: Ver Lima (2004). □

A validade dessa proposição segue basicamente do fato que, sob certas condições, podemos escrever $f(g(t))$ na expansão de Taylor através de

$$\begin{aligned} f(g(t) - dg(t)) - f(g(s)) &= f'(g(s))\, dg(t) \\ &\quad + \frac{1}{2} f''(g(s))\, [dg(t)] + \cdots, \end{aligned} \tag{2.56}$$

onde $dg(t) = g(t + dt) - g(t)$.

Ainda sob condições apropriadas, o termo de segunda ordem e os de ordem superior na expansão (2.56) são desprezíveis, para um dt suficientemente pequeno. Agora, vamos assumir que f é função duas vezes diferenciável, e consideramos $g(t) = B_t$, onde $B = \{B_t\}_{t \geq 0}$ denota o processo movimento Browniano. Escrevemos $dB_t = B_{t+dt} - B_t$ para os incrementos de B em $[t, t + dt]$. Então, reescrevemos a expressão (2.56) como

$$f(B_t + dB_t) - f(B_t) = f'(B_t)\, dB_t + \frac{1}{2} f''(B_t)\, [dB_t]^2 + \cdots. \tag{2.57}$$

Diferentemente do caso determinístico, a contribuição do termo de segunda ordem não é desprezível (ver Mikosch, 1998).

Integrando ambos os lados de (2.57) e desprezando os termos de ordem 3 ou maiores, obtemos, para s < t,

$$\begin{aligned}\int_s^t df\left(B_x\right) &:= f\left(B_t\right) - f\left(B_s\right) \\ &= \int_s^t f^{'}\left(B_x\right) dB_x + \frac{1}{2}\int_s^t f^{''}\left(B_x\right) dx. \end{aligned} \tag{2.58}$$

A primeira integral de (2.58) é a integral de Itô de $f^{'}(B)$, e a segunda integral pode ser interpretada como a equação de Riemann de $f^{''}(B)$.

Teorema 2.4 (Forma Simples do Lema de Itô)**.** *Seja f função duas vezes diferenciável, com derivadas contínuas. Então, a* forma simples do Lema de Itô *é dada por*

$$f\left(B_t\right) - f\left(B_s\right) = \int_s^t f^{'}\left(B_x\right) dB_x + \frac{1}{2}\int_s^t f^{''}\left(B_x\right) dx, \quad \textit{para todo } s < t. \tag{2.59}$$

Demonstração: Ver Protter (1992), página 78. □

A seguir, damos um exemplo da aplicação da forma simples do lema de Itô.

Exemplo 2.19. Seja $f : \mathbb{R} \to \mathbb{R}$ a função dada por $f(t) = t^2$. Então, $f^{'}(t) = 2t$ e $f^{''}(t) = 2$. Pelo Teorema 2.4, o lema de Itô resulta em

$$\begin{aligned} B_t^2 - B_s^2 &= \int_s^t 2B_x\, dB_x + \frac{1}{2}\int_s^t 2\, dx \\ &= 2\int_s^t B_x\, dB_x + \int_s^t dx, \text{ para todo } s < t. \end{aligned} \tag{2.60}$$

Fazendo $s = 0$, a primeira integral do lado direito da expressão (2.60) pode ser escrita por

$$\int_0^t B_x\, dB_x = \frac{1}{2}\left[\left(B_x^2 - x\right)\right]_0^t = \frac{1}{2}\left(B_t^2 - t\right).$$

De maneira análoga, podemos estender o Lema de Itô de diveras maneiras. As demonstrações de tais extensões podem ser encontradas em Protter (1992).

Teorema 2.5 (Extensão I do Lema de Itô)**.** *Seja $f(t,x)$ função cujas derivadas parciais de segunda ordem são contínuas. Então, para todo $s<t$,*

$$\begin{aligned} f(t,B_t)-f(s,B_s) &= \int_s^t \left[f_1(x,B_x)+\frac{1}{2}f_{22}(x,B_x)\right]dx \\ &\quad +\int_s^t f_2(x,B_x)\,dB_x, \end{aligned} \tag{2.61}$$

onde

$$\begin{aligned} f_i(t,x) &= \left[\frac{\partial}{\partial x_i}f(x_1,x_2)\right]_{(x_1,x_2)=(t,x)},\ i=1,2, \\ f_{22}(t,x) &= \left[\frac{\partial^2}{(\partial x_2)^2}f(x_1,x_2)\right]_{(x_1,x_2)=(t,x)}, \end{aligned}$$

são as derivadas parcias de f.

Apresentamos a seguir uma versão mais geral do Lema de Itô. Para isso, consideramos a seguinte definição.

Definição 2.36 (Processo de Itô)**.** Se X_t é dado por

$$X_t = X_0+\int_0^t A_s^{(1)}\,ds+\int_0^t A_s^{(2)}\,dB_s, \tag{2.62}$$

onde $A^{(1)}=\left(A_t^{(1)},\ t\geq 0\right)$ e $A^{(2)}=\left(A_t^{(1)},\ t\geq 0\right)$ são processos adaptados ao movimento Browniano, então dizemos que $X=(X_t,\ t\geq 0)$ é um *processo de Itô.*

Então, utilizando argumentos similares àqueles para justificar o Lema de Itô (ver Teorema 2.4), obtemos uma nova extensão do Teorema 2.5.

Teorema 2.6 (Extensão II do Lema de Itô)**.** *Seja $X=(X_t,\ t\geq 0)$ um processo de Itô com representação dada pela expressão* (2.62) *e seja $f(t,x)$ uma função cujas derivadas parciais de segunda ordem são contínuas. Então, para todo $s<t$,*

$$\begin{aligned} f(t,X_t)-f(s,X_s) &= \int_s^t \left[f_1(y,X_y)+A_y^{(1)}f_2(y,X_y)+\frac{1}{2}\left[A_y^{(2)}\right]^2 f_{22}(y,X_y)\right]dy \\ &\quad +\int_s^t A_y^{(2)}f_2(y,B_y)\,dB_y, \end{aligned} \tag{2.63}$$

onde $A_y^{(1)}$ *e* $A_y^{(2)}$ *estão definidos na expressão* (2.62) *e*

$$\begin{aligned} f_i(t,x) &= \left[\frac{\partial}{\partial x_i} f(x_1,x_2)\right]_{(x_1,x_2)=(t,x)}, \ i=1,2, \\ f_{22}(t,x) &= \left[\frac{\partial^2}{(\partial x_2)^2} f(x_1,x_2)\right]_{(x_1,x_2)=(t,x)}, \end{aligned}$$

são as derivadas parcias de f.

Com o Lema de Itô, e suas respectivas extensões, possuímos ferramentas suficientes para o estudo da solução da equação de Langevin, a ser visto no Capítulo 3.1.

Capítulo 3

Equação de Langevin

Para complementar e generalizar os resultados sobre o movimento Browniano e a dinâmica do movimento irregular e imprevisível de uma partícula imersa em um meio líquido obtidos por Albert Einstein, o físico francês Paul Langevin iniciou uma série de estudos a respeito.

Paul Langevin iniciou escrevendo a equação do movimento da partícula Browniana de acordo com a segunda lei de Newton, assumindo que a sua trajetória é influenciada por duas forças: uma força de atrito, regida por um coeficiente λ, e uma força de flutuação η_t, que é devido ao impacto das moléculas do líquido com a partícula. Langevin assume ainda que a força de atrito segue a lei de Stokes, que afirma que a força de atrito desacelerando uma partícula esférica é proporcional ao coeficiente de viscosidade do fluido. Assume ainda que η_t é independente da velocidade X_t, e que a variação de η_t é extremamente rápida quando comparada com a variação de X_t. Dessa forma, é obtida a equação de Langevin apresentada em (3.1) a seguir.

Uma variação da equação de Langevin clássica é a equação de Langevin clássica fracionária, que apresenta um ruído movimento Browniano fracionário. Nesse caso, o processo estocástico gerado a partir da solução da equação de Langevin clássica

fracionária não possui mais uma função de covariância com decaimento exponencial, mas sim um decaimento mais lento, que permite a aplicação da teoria de séries temporais com característica de longa dependência.

Em 1965 com Hazime Mori e em 1966 com Ryogo Kubo, surge uma generalização do modelo proposto por Langevin, a chamada equação de Langevin generalizada.

3.1 Equação de Langevin Clássica e sua Solução

Consideramos o processo estocástico $\{X_t\}_{t\in[0,T]}$ dado pela equação diferencial estocástica linear

$$\begin{cases} dX_t = -\lambda X_t dt + \sigma dB_t \\ X(0) = X_0, \end{cases} \tag{3.1}$$

onde λ e σ são constantes positivas, $\{B_t\}_{t\in[0,T]}$ é o movimento Browniano e X_0 é independente de B_t, para todo $t \in [0,T]$.

Podemos reescrever a equação diferencial (3.1) na forma

$$X_t = X_0 - \lambda \int_0^t X_s \, ds + \sigma \int_0^t dB_s. \tag{3.2}$$

Observe que (3.2) é um processo de Itô. Assim, realizamos a mudança de variável

$$Y_t = e^{\lambda t} X_t, \text{ para todo } t \in [0,T]. \tag{3.3}$$

Os processos $\{X_t\}_{t\in[0,T]}$ e $\{Y_t\}_{t\in[0,T]}$ têm a mesma condição inicial $X_0 = Y_0$.

Assim, aplicando o Teorema 2.6, visto na Subseção 2.11.3, com $f(t,x) = e^{\lambda t}x$, $A_s^{(1)} = -\lambda X_s$ e $A_s^{(2)} = \sigma$, notamos que

$$f_{_1}(s, X_s) = \lambda f(s, X_s), \quad f_{_2}(s, X_s) = e^{\lambda s}, \quad f_{_{22}} = 0.$$

Assim, substituindo na expressão (2.63), obtemos

$$\begin{aligned} Y_t - Y_0 &= \int_0^t \left[f_1(s, X_s) - \lambda X_s f_2(s, X_s) + \frac{1}{2}\sigma^2 f_{22}(s, X_s) \right] ds + \int_0^t \sigma f_2(s, X_s)\, dB_s \\ &= \int_0^t \left(\lambda e^{\lambda s} X_s - \lambda e^{\lambda s} X_s \right) ds + \int_0^t \sigma e^{\lambda s}\, dB_s \\ &= \int_0^t \left(\lambda Y_s - \lambda Y_s \right)\, ds + \int_0^t \sigma e^{\lambda s}\, dB_s \\ &= \int_0^t \sigma e^{\lambda s}\, dB_s. \end{aligned} \tag{3.4}$$

Logo, da igualdade (3.4), temos

$$Y_t = Y_0 + \int_0^t \sigma e^{\lambda s}\, dB_s.$$

Pela expressão (3.3) e do fato que $X_0 = Y_0$, obtemos a solução da equação (3.1) dada por

$$X_t = X_0 e^{-\lambda t} + \sigma e^{-\lambda t} \int_0^t e^{\lambda s}\, dB_s. \tag{3.5}$$

Definição 3.1 (Processo de Ornstein-Uhlenbeck)**.** O processo $\{X_t\}_{t\in[0,T]}$, dado pela expressão (3.5), solução da equação diferencial (3.1), para uma condição inicial X_0, é chamado *processo de Ornstein-Uhlenbeck (ou processo O-U).*

A seguir apresentamos algumas propriedades do processo $\{X_t\}_{t\in[0,T]}$, dado pela expressão (3.5). Apresentamos também sua forma recursiva, que será essencial para obtermos uma forma discreta desse processo.

Proposição 3.1. *Seja $\{X_t\}_{t\in[0,T]}$ o processo O-U dado na expressão* (3.5)*. Então,*

i) a esperança do processo é dada por

$$\mathbb{E}\left(X_t\right) = e^{-\lambda t}\, \mathbb{E}(X_0), \text{ para todo } t \in [0, T];$$

ii) a variância do processo é dada por

$$Var\left(X_t\right) = e^{-2\lambda t} \left[Var(X_0) - \frac{\sigma^2}{2\lambda} \right] + \frac{\sigma^2}{2\lambda}, \text{ para todo } t \in [0, T];$$

iii) a forma recursiva da equação (3.5) *é dada por*

$$X_{t+h} = e^{-\lambda h}\left(X_t + \sigma e^{-\lambda t}\int_t^{t+h} e^{\lambda s}\, dB_s\right),\ \textit{para } h \in \mathbb{R} \textit{ tal que } h > 0; \quad (3.6)$$

iv) se $X_0 \sim \mathcal{N}\left(0, \frac{\sigma^2}{2\lambda}\right)$, *então* $X_t \sim \mathcal{N}\left(0, \frac{\sigma^2}{2\lambda}\right)$;

v) se $X_0 \sim \mathcal{N}\left(0, \frac{\sigma^2}{2\lambda}\right)$, *o processo é estacionário e tem função de autocovariância dada por*

$$\begin{aligned}\gamma_X(h) &= e^{-\lambda|h|}\left[e^{-2\lambda t}\left(Var(X_0) - \frac{\sigma^2}{2\lambda}\right) + \frac{\sigma^2}{2\lambda}\right] \\ &= \frac{\sigma^2 e^{-\lambda|h|}}{2\lambda}, \textit{ para todo } h \in [0,T]; \end{aligned} \quad (3.7)$$

vi) a função de autocorrelação do processo é dada por

$$\rho_X(h) = Corr(X_{t+h}, X_t) = e^{-\lambda|h|}, \textit{ para todo } t, t+h \in [0,T]; \quad (3.8)$$

vii) a função de densidade espectral do processo é dada por

$$f_X(\omega) = \frac{\sigma^2}{4\pi\lambda}\left(\frac{e^{2\lambda}-1}{1+e^{2\lambda}-2e^{\lambda}\cos(\omega)}\right), \textit{ para todo } \omega \in [0,\pi]. \quad (3.9)$$

Demonstração:

i) Da expressão (3.5) obtemos que

$$\begin{aligned}\mathbb{E}(X_t) &= \mathbb{E}\left(X_0 e^{-\lambda t} + \sigma e^{-\lambda t}\int_0^t e^{\lambda s}\, dB_s\right) \\ &= \mathbb{E}\left(X_0 e^{-\lambda t}\right) + \mathbb{E}\left(\sigma e^{-\lambda t}\int_0^t e^{\lambda s}\, dB_s\right) \\ &= e^{-\lambda t}\mathbb{E}(X_0),\end{aligned}$$

pois, como visto na Proposição 2.17, a integral estocástica de Itô tem esperança zero.

ii) Para calcular a variância do processo, utilizamos a expressão (3.5), e obtemos

$$\begin{aligned}
Var\left(X_t\right) &= \mathbb{E}\left({X_t}^2\right) - \left[\mathbb{E}(X_t)\right]^2 \\
&= \mathbb{E}\left(X_0 e^{-\lambda t} + \sigma e^{-\lambda t}\int_0^t e^{\lambda s}\, dB_s\right)^2 - \left[e^{-\lambda t}\mathbb{E}\left(X_0\right)\right]^2 \\
&= \mathbb{E}\left({X_0}^2 e^{-2\lambda t} + 2X_0\sigma e^{-2\lambda t}\int_0^t e^{\lambda s}\, dB_s + e^{-2\lambda t}\left(\sigma\int_0^t e^{\lambda s}\, dB_s\right)^2\right) \\
&\quad -e^{-2\lambda t}\left[\mathbb{E}\left(X_0\right)\right]^2 \\
&= e^{-2\lambda t}\mathbb{E}\left({X_0}^2\right) + 2\sigma e^{-2\lambda t}\mathbb{E}\left(X_0\int_0^t e^{\lambda s}\, dB_s\right) \\
&\quad +e^{-2\lambda t}\sigma^2\mathbb{E}\left(\int_0^t e^{\lambda s}\, dB_s\right)^2 - e^{-2\lambda t}\left[\mathbb{E}\left(X_0\right)\right]^2.
\end{aligned}$$

Temos que X_0 e B_t são independentes, para todo $t \in [0,T]$. Logo, obtemos que

$$\mathbb{E}\left(X_0\int_0^t e^{\lambda s}\, dB_s\right) = \mathbb{E}\left(X_0\right)\mathbb{E}\left(\int_0^t e^{\lambda s}\, dB_s\right) = 0.$$

Pelo Teorema 2.3, podemos aplicar a propriedade da isometria da integral de Itô. Assim, para todo $t \in [0,T]$, obtemos

$$\begin{aligned}
Var\left(X_t\right) &= e^{-2\lambda t}\mathbb{E}\left(X_0^2\right) + e^{-2\lambda t}\sigma^2\int_0^t e^{2\lambda s}\, ds - e^{-2\lambda t}\left[\mathbb{E}\left(X_0\right)\right]^2 \\
&= e^{-2\lambda t}Var\left(X_0\right) + e^{-2\lambda t}\sigma^2\left[\frac{e^{2\lambda s}}{2\lambda}\right]_0^t \\
&= e^{-2\lambda t}Var\left(X_0\right) + e^{-2\lambda t}\left(\frac{\sigma^2}{2\lambda}e^{2\lambda t} - \frac{\sigma^2}{2\lambda}\right) \\
&= e^{-2\lambda t}\left(Var\left(X_0\right) - \frac{\sigma^2}{2\lambda}\right) + \frac{\sigma^2}{2\lambda}.
\end{aligned}$$

iii) Para obter a forma recursiva, basta notar que

$$\begin{aligned}
X_{t+h} &= e^{-\lambda(t+h)}\left(X_0 + \sigma\int_0^{t+h} e^{\lambda s}\,dB_s\right)\\
&= e^{-\lambda h}\left(X_0 e^{-\lambda t} + \sigma e^{-\lambda t}\int_0^{t+h} e^{\lambda s}\,dB_s\right)\\
&= e^{-\lambda h}\left[X_0 e^{-\lambda t} + \sigma e^{-\lambda t}\left(\int_0^{t} e^{\lambda s}\,dB_s + \int_t^{t+h} e^{\lambda s}\,dB_s\right)\right]\\
&= e^{-\lambda h}\left[X_0 e^{-\lambda t} + \sigma e^{-\lambda t}\int_0^{t} e^{\lambda s}\,dB_s + \left(\sigma e^{-\lambda t}\int_t^{t+h} e^{\lambda s}\,dB_s\right)\right]\\
&= e^{-\lambda h}\left(X_t + \sigma e^{-\lambda t}\int_t^{t+h} e^{\lambda s}\,dB_s\right).
\end{aligned}$$

iv) Temos que $\displaystyle\int_0^t e^{\lambda s}\,dB_s$ é um processo Gaussiano (ver Klebaner, 1999, página 107). De fato, basta verificar que, assumindo $Y_t^m = \displaystyle\sum_{i=1}^{m} e^{\lambda s_i}\left(B_{s_i} - B_{s_{i-1}}\right)$, onde $\tau_m = \{s_0, \cdots, s_m\}$ é uma partição igualmente espaçada do intervalo $[0,t]$, temos que $Y_t^m \to Y_t$ no sentido L^2, onde $Y_t = \displaystyle\int_0^t e^{\lambda s}\,dB_s$.

Para todo $t \in [0,T]$, $Y_t^m = \displaystyle\sum_{i=1}^{m} e^{\lambda s_i}\left(B_{s_i} - B_{s_{i-1}}\right)$ é a soma de processos com distribuição normal. Portanto, também é normal. Como convergência em L^2 implica em convergência em distribuição, temos que $Y_t^m \xrightarrow{\mathcal{D}} Y_t$, ou seja, $\displaystyle\int_0^t e^{\lambda s}\,dB_s$ tem distribuição normal. Portanto, X_t é a soma de dois processos Gaussianos e, consequentemente, é um processo Gaussiano, com média e variância obtidas nos itens i) e ii).

v) Para obter a função de autocovariância do processo, calculamos primeiramente $\mathbb{E}\left(X_t X_{t+h}\right)$, dada por

$$\mathbb{E}\left[\left(X_0 e^{-\lambda t} + \sigma e^{-\lambda t}\int_0^t e^{\lambda s}\,dB_s\right)\left(X_0 e^{-\lambda(t+h)} + \sigma e^{-\lambda(t+h)}\int_0^{t+h} e^{-\lambda s}\,dB_s\right)\right].$$

Multiplicando e aplicando a propriedade de linearidade da esperança, obtemos

$$\begin{aligned}\mathbb{E}\left(X_tX_{t+h}\right) &= \mathbb{E}\left(e^{-\lambda(2t+h)}X_0^2\right)+\mathbb{E}\left(e^{-\lambda t}X_0e^{-\lambda(t+h)}\sigma\int_0^{t+h}e^{\lambda s}\,dB_s\right)\\ &\quad+\mathbb{E}\left(e^{-\lambda t}X_0e^{-\lambda(t+h)}\sigma\int_0^{t}e^{\lambda s}\,dB_s\right)\\ &\quad+\mathbb{E}\left[\left(\sigma e^{-\lambda t}\int_0^{t}e^{\lambda s}\,dB_s\right)\left(\sigma e^{-\lambda(t+h)}\int_0^{t+h}e^{\lambda s}\,dB_s\right)\right].\end{aligned}$$

Pela independência de X_0 e B_t, e como a integral de Itô tem esperança zero, obtemos

$$\begin{aligned}\mathbb{E}\left(X_tX_{t+h}\right) &= \mathbb{E}\left(e^{-\lambda(2t+h)}X_0^2\right)\\ &\quad+\sigma^2e^{-\lambda(2t+h)}\mathbb{E}\left[\left(\int_0^{t}e^{\lambda s}\,dB_s\right)\left(\int_0^{t+h}e^{\lambda s}\,dB_s\right)\right]\\ &= \mathbb{E}\left(e^{-\lambda(2t+h)}{X_0}^2\right)\\ &\quad+\sigma^2e^{-\lambda(2t+h)}\mathbb{E}\left[\left(\int_0^{t}e^{\lambda s}\,dB_s\right)\left(\int_0^{t}e^{\lambda s}\,dB_s+\int_t^{t+h}e^{\lambda s}\,dB_s\right)\right]\\ &= \mathbb{E}\left(e^{-\lambda(2t+h)}{X_0}^2\right)\\ &\quad+\sigma^2e^{-\lambda(2t+h)}\left[\mathbb{E}\left(\int_0^{t}e^{\lambda s}\,dB_s\right)^2\right]\\ &\quad+\sigma^2e^{-\lambda(2t+h)}\left[\mathbb{E}\left(\int_0^{t}e^{\lambda s}\,dB_s\right)\left(\int_t^{t+h}e^{\lambda s}\,dB_s\right)\right]. \qquad (3.10)\end{aligned}$$

Na expressão (3.10) aplicamos a propriedade da isometria, dada por (2.55), isto é,

$$\mathbb{E}\left(\int_0^{t}e^{\lambda s}\,dB_s\right)^2=\int_0^{t}e^{2\lambda s}\,ds \qquad (3.11)$$

e a propriedade de martingales, isto é,

$$\mathbb{E}\left[\left(\int_0^{t}e^{\lambda s}\,dB_s\right)\left(\int_t^{t+h}e^{\lambda s}\,dB_s\right)\right]=0. \qquad (3.12)$$

Assim, utilizando (3.11) e (3.12), a expressão (3.10) pode ser reescrita como

$$\begin{aligned}\mathbb{E}\left(X_tX_{t+h}\right) &= e^{-\lambda(2t+h)}\mathbb{E}\left(X_0^2\right)+\sigma^2e^{-\lambda(2t+h)}\int_0^{t}e^{2\lambda s}\,ds\\ &= e^{-\lambda(2t+h)}\mathbb{E}\left(X_0^2\right)+\sigma^2e^{-\lambda(2t+h)}\left[\frac{e^{2\lambda s}}{2\lambda}\right]_0^{t}\\ &= e^{-\lambda(2t+h)}\mathbb{E}\left(X_0^2\right)+\frac{\sigma^2e^{-\lambda h}}{2\lambda}-\frac{\sigma^2e^{-\lambda(2t+h)}}{2\lambda}. \qquad (3.13)\end{aligned}$$

Então, a partir da expressão (3.13) e o item i) desta proposição, concluímos que

$$\begin{aligned}
\gamma_X(h) &= Cov(X_{t+h}, X_t) \\
&= e^{-\lambda(2t+h)}\mathbb{E}\left(X_0^2\right) + \frac{\sigma^2 e^{-\lambda h}}{2\lambda} \\
&\quad -\frac{\sigma^2 e^{-\lambda(2t+h)}}{2\lambda} - \left(e^{-\lambda t}\mathbb{E}(X_0)\, e^{-\lambda(t+h)}\mathbb{E}(X_0)\right) \\
&= e^{-\lambda(2t+h)}\left(\mathbb{E}\left({X_0}^2\right) - [\mathbb{E}(X_0)]^2\right) + \frac{\sigma^2 e^{-\lambda h}}{2\lambda} - \frac{\sigma^2 e^{-\lambda(2t+h)}}{2\lambda} \\
&= e^{-\lambda(2t+h)}\left(Var(X_0) - \frac{\sigma^2}{2\lambda}\right) + \frac{\sigma^2 e^{-\lambda h}}{2\lambda} \\
&= e^{-\lambda h}\left[e^{-2\lambda t}\left(Var(X_0) - \frac{\sigma^2}{2\lambda}\right) + \frac{\sigma^2}{2\lambda}\right].
\end{aligned}$$

Como $\gamma_X(-h) = \gamma_X(h)$ e $Var(X_0) = \frac{\sigma^2}{2\lambda}$, então

$$\begin{aligned}
\gamma_X(h) &= e^{-\lambda|h|}\left[e^{-2\lambda t}\left(Var(X_0) - \frac{\sigma^2}{2\lambda}\right) + \frac{\sigma^2}{2\lambda}\right] \\
&= \frac{\sigma^2 e^{-\lambda|h|}}{2\lambda}. \qquad (3.14)
\end{aligned}$$

É fácil ver que o processo $\{X_t\}_{t\in[0,T]}$ é fracamente estacionário. Então, como $\{X_t\}_{t\in[0,T]}$ é um processo Gaussiano fracamente estacionário, também é estacionário.

vi) Temos que a função de autocorrelação do processo é dada por

$$\rho_X(h) = Corr(X_{t+h}, X_t) = \frac{\gamma_X(h)}{\gamma_X(0)}, \text{ para todo } h \in [0,T],$$

onde $\gamma_X(\cdot)$ é a função de autocovariância do processo $\{X_t\}_{t\in[0,T]}$.

Portanto,

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \frac{\frac{\sigma^2 e^{-\lambda|h|}}{2\lambda}}{\frac{\sigma^2}{2\lambda}} = e^{-\lambda|h|}, \text{ para todo } h \in [0,T]. \qquad (3.15)$$

Logo, obtemos o resultado desejado.

vii) Como o processo $\{X_t\}_{t\in[0,T]}$ é real, temos por (2.12) que a função densidade espectral é dada por

$$\begin{aligned} f_X(\omega) &= \frac{1}{2\pi}\gamma_X(0) + \frac{1}{\pi}\sum_{k=1}^{\infty}\gamma_X(k)\cos(\omega k) \\ &= \frac{1}{2\pi}\gamma_X(0) + \frac{1}{\pi}\sum_{k=1}^{\infty}\frac{\sigma^2 e^{-\lambda k}}{2\lambda}\cos(\omega k) \\ &= \frac{\sigma^2}{4\pi\lambda} + \frac{\sigma^2}{2\pi\lambda}\sum_{k=1}^{\infty}e^{-\lambda k}\cos(\omega k) \\ &= \frac{\sigma^2}{4\pi\lambda} + \frac{\sigma^2}{2\pi\lambda}\left(\frac{e^{\lambda}\cos(\omega)-1}{1+e^{2\lambda}-2e^{\lambda}\cos(\omega)}\right) \\ &= \frac{\sigma^2}{4\pi\lambda}\left(1+\frac{2e^{\lambda}\cos(\omega)-2}{1+e^{2\lambda}-2e^{\lambda}\cos(\omega)}\right) \\ &= \frac{\sigma^2}{4\pi\lambda}\left(\frac{e^{2\lambda}-1}{1+e^{2\lambda}-2e^{\lambda}\cos(\omega)}\right), \end{aligned}$$

pois $\sum_{k=1}^{\infty}e^{-\lambda k}\cos(\omega k) = \frac{e^{\lambda}\cos(\omega)-1}{1+e^{2\lambda}-2e^{\lambda}\cos(\omega)}$ (ver Gradshteyn e Ryzhik, 2000, página 47).

□

Nos itens v) e vi) da Proposição 3.1 apresentamos as funções de autocovariância e de autocorrelação definidas para um intervalo de tempo real, que é um contexto diferente ao que estamos habituados nos processos estocásticos discretos, onde $\gamma_X(h)$ e $\rho_X(h)$ estão definidos apenas para $h \in \mathbb{Z}$.

Quando estudamos aplicações dos processos estocásticos contínuos, em particular o processo $\{X_t\}_{t\in[0,T]}$ apresentado na Definição 3.1, utilizamos amostras obtidas deste processo, separadas por um espaço de tempo fixo $h \in \mathbb{R}_+$. Isto é, analisamos o processo discreto

$$S = \{X_0,\, X_h,\, X_{2h}, \cdots,\, X_{kh}, \cdots,\, X_T\} = \left\{S_0,\, S_1,\, \cdots,\, S_{\frac{T}{h}}\right\} \equiv \{S_t\}_{t=0}^{\left\lfloor \frac{T}{h} \right\rfloor}. \quad (3.16)$$

Dado um elemento $X_t \in S$, os k-ésimos elementos seguintes são

$$X_t,\, X_{t+h},\, X_{t+2h},\, X_{t+3h}, \cdots,\, X_{t+kh},$$

para $h \in \mathbb{R}_+$. Assim, para analisar a função de autocovariância entre os elementos de S, calculamos, para $h \in \mathbb{R}_+$,

$$\gamma_{_S}(k) = Cov(S_t, S_{t+k}) = Cov(X_t, X_{t+kh}) = \gamma_{_X}(kh), \text{ para todo } k \in \mathbb{Z}.$$

Obtemos, assim, a seguinte proposição.

Proposição 3.2. *Seja $\{X_t\}_{t\in[0,T]}$ o processo advindo da solução da equação de Langevin clássica. Considerando $h \in \mathbb{R}_+$, seja S o conjunto de observações discretas do processo $\{X_t\}_{t\in[0,T]}$, dado na expressão* (3.16)*. Então,*

i) a função de autocovariância de S, para $h \in \mathbb{R}_+$ fixo, é dada por

$$\gamma_{_S}(k) = \frac{\sigma e^{-\lambda|k|h}}{2\lambda}, \text{ para todo } k \in \mathbb{Z}; \tag{3.17}$$

ii) a função de autocorrelação de S, para $h \in \mathbb{R}_+$ fixo, é dada por

$$\rho_{_S}(k) = e^{-\lambda|k|h}, \text{ para todo } k \in \mathbb{Z}. \tag{3.18}$$

Demonstração: A demonstração é semelhante à dos itens iv) e v) da Proposição 3.1, bastando substituir h por kh nas expressões. □

3.2 Forma Discreta da Solução da Equação de Langevin Clássica

Para facilitar a conexão entre a teoria e a implementação computacional do processo $\{X_t\}_{t\in[0,T]}$, dado pela expressão (3.6), é importante obtermos uma forma discreta da solução da equação de Langevin clásssica. De Valdivieso et al. (2009) obtemos a seguinte proposição.

Proposição 3.3. *Seja $\{Z_t\}_{t\geq 0}$ um processo de Lévy. Então, para todo t, $h > 0$,*

$$e^{-\lambda t}\int_t^{t+h} e^{\lambda s}\, dZ_s \overset{\mathcal{D}}{=} \int_0^h e^{\lambda s}\, dZ_s. \tag{3.19}$$

Demonstração: A demonstração desta proposição segue da regra de integração por partes para semimartingales (ver Protter, 1990). No caso em que $Z_t = B_t$, para $t \geq 0$, ou seja, quando o processo considerado é o movimento Browniano, temos, pela fórmula de integração por partes, que

$$\int_a^b f(s)\, dB_s = f(b)\, B_b - f(a)\, B_a - \int_a^b B_s f'(s)\, ds.$$

Logo, obtemos que

$$\begin{aligned} e^{-\lambda t}\int_t^{t+h} e^{\lambda s} dB_s &= \int_t^{t+h} e^{\lambda(s-t)} dB_s = e^{\lambda h} B_{t+h} - B_t - \int_t^{t+h} B_s \lambda e^{\lambda(s-t)} ds \\ &= e^{\lambda h} B_{t+h} - B_t - \int_0^h B_{t+u} e^{\lambda u} d(\lambda u) \\ &= e^{\lambda h} B_{t+h} - B_t - \int_0^h (B_{t+u} - B_t)\, e^{\lambda u} d(\lambda u) - \int_0^h B_t e^{\lambda u} d(\lambda u) \\ &= e^{\lambda h}(B_{t+h} - B_t) - \int_0^h (B_{t+h} - B_t)\, e^{\lambda u} d(\lambda u), \end{aligned}$$

onde, na terceira igualdade acima, consideramos $u = s - t$. Como $B_{t+h} - B_t \overset{\mathcal{D}}{=} B_h$ e $B_{t+u} - B_t \overset{\mathcal{D}}{=} B_u$, então

$$e^{-\lambda t}\int_t^{t+h} e^{\lambda s} dB_s \overset{\mathcal{D}}{=} e^{\lambda h} B_h - \int_0^h B_u \lambda e^{\lambda u} du = \int_0^h e^{\lambda s} dB_s,$$

novamente pela fórmula de integração por partes. □

Como o processo movimento Browniano é um processo de Lévy, da Proposição 3.3 podemos reescrever a solução (3.6) na forma

$$X_{t+h} \overset{\mathcal{D}}{=} e^{-\lambda h}\left(X_t + \sigma \int_0^h e^{\lambda s}\, dB_s\right). \tag{3.20}$$

Pela Proposição 2.14, considerando que $f(x) = e^{\lambda x}$ é diferenciável e com derivada limitada em $[0, t]$, a integral de Riemann-Stieltjes $\int_0^h e^{\lambda s}\, dB_s$ existe e está bem definida. Desta forma, (3.20) nos sugere considerar uma expressão discretizada dada por

$$X_{t+h} = e^{-\lambda h} \lim_{m\to\infty}\left(X_t + \sigma \sum_{i=0}^{m} e^{\lambda s_{i-1}}\left(B_{s_i} - B_{s_{i-1}}\right)\right) \tag{3.21}$$

onde a sequência de números reais $(s_i)_{i=0}^m$ é uma partição, igualmente espaçada, do intervalo $[0, h]$.

3.3 Equação de Langevin Clássica Fracionária

Consideramos a equação integral estocástica

$$X_t^H = X_0^H - \lambda \int_0^t X_s^H \, ds + \sigma B_t^H, \tag{3.22}$$

onde $t \in [0, T]$, σ e λ são constantes positivas e $\left\{B_t^H\right\}_{t\in[0,T]}$ é o processo movimento Browniano fracionário com parâmetro de Hurst $H \in (0, 1]$. A equação (3.22) é chamada *equação de Langevin Clássica Fracionária*, e sua solução é dada por

$$X_t^H = e^{-\lambda t}\left(X_0^H + \sigma \int_0^t e^{\lambda s} \, dB_s^H\right), \tag{3.23}$$

conforme verificamos na Proposição 3.4 a seguir.

Proposição 3.4. *Seja $\left\{B_t^H\right\}_{t\in[0,T]}$ o processo movimento Browniano fracionário com parâmetro de Hurst $H \in (0, 1]$ e $X_0^H \in L_0(\Omega)$. Sejam $-\infty \leq a < \infty$ e λ, $\sigma > 0$. Então, para quase todo $\omega \in \Omega$, temos que*

i) a integral

$$\int_a^t e^{\lambda s} \, dB_s^H$$

é integrável à Riemann-Stieltjes, para todo $t > a$, e vale que

$$\int_a^t e^{\lambda s} \, dB_s^H = e^{\lambda t} B_t^H - e^{\lambda a} B_a^H - \lambda \int_a^t e^{\lambda s} B_s^H \, ds;$$

ii) para $t > a$, a função

$$y(t) = \int_a^t e^{\lambda s} \, dB_s^H$$

é contínua.

iii) o único processo estocástico contínuo que soluciona a equação de Langevin clássica fracionária, definida em (3.22)*, é dada por*

$$X_t^H = e^{-\lambda t}\left(X_0^H + \sigma \int_0^t e^{\lambda s} \, dB_s^H\right), \tag{3.24}$$

onde $t > 0$. O processo estocástico $\left\{X_t^H\right\}_{t\in[0,T]}$, gerado por (3.24)*, é chamado de* processo de Ornstein-Uhlenbeck fracionário.

No caso especial em que $X_0^H = \sigma \int_{-\infty}^{0} e^{\lambda s}\, dB_s^H$, *temos que a única solução contínua de* (3.22) *é dada por*

$$X_t^H = \sigma \int_{-\infty}^{t} e^{-\lambda(t-s)}\, dB_s^H, \tag{3.25}$$

para $t > 0$.

Demonstração: Os itens i) e ii) podem ser vistos em Cheridito et al. (2003). Para provar o item iii), observe que uma função contínua X_t^H soluciona (3.22) se e somente se a função

$$z(t) = \int_0^t X_u^H\, du, \tag{3.26}$$

para $t > 0$, soluciona a equação diferencial linear

$$\begin{cases} z'(t) = -\lambda z(t) + X_0^H + \sigma B_t^H \\ z(0) = 0, \end{cases} \tag{3.27}$$

cuja única solução é facilmente obtida, e é dada por

$$z(t) = e^{-\lambda t} \int_0^t e^{\lambda s} \left(X_0^H + \sigma B_s^H\right) ds, \tag{3.28}$$

para $t > 0$. Logo, substituindo em (3.27) obtemos que

$$X_t^H = -\lambda e^{-\lambda t} \int_0^t e^{\lambda s} \left(X_0^H + \sigma B_s^H\right) ds + X_0^H + \sigma B_t^H. \tag{3.29}$$

Pelo item i),

$$\int_0^t e^{\lambda s}\, dB_s^H \;=\; e^{\lambda t} B_t^H - e^{\lambda 0} B_0^H - \lambda \int_0^t B_u^H e^{\lambda u}\, du,$$

o que implica que

$$\int_0^t B_u^H e^{\lambda u}\, du = -\frac{1}{\lambda} \left(\int_0^t e^{\lambda s} dB_s^H + e^{\lambda t} B_t^H \right).$$

Substituindo em (3.29),

$$\begin{aligned}
X_t^H &= -\lambda e^{-\lambda t}\left[X_0^H\int_0^t e^{\lambda u}\,du+\frac{\sigma}{\lambda}\left(e^{\lambda t}B_t^H-\int_0^t e^{\lambda s}\,dB_s^H\right)\right]+X_0^H+\sigma B_t^H\\
&= -\lambda e^{-\lambda t}\left[\frac{X_0^H}{\lambda}\left(e^{\lambda t}-1\right)+\frac{\sigma}{\lambda}\left(e^{\lambda t}B_t^H-\int_0^t e^{\lambda s}\,dB_s^H\right)\right]+X_0^H+\sigma B_t^H\\
&= -e^{-\lambda t}\left[X_0^H\left(e^{\lambda t}-1\right)+\sigma\left(e^{\lambda t}B_t^H-\int_0^t e^{\lambda s}\,dB_s^H\right)\right]+X_0^H+\sigma B_t^H\\
&= -X_0^H+e^{-\lambda t}X_0^H-\sigma B_t^H+\sigma e^{-\lambda t}\int_0^t e^{\lambda s}\,dB_s^H+X_0^H+\sigma B_t^H\\
&= e^{-\lambda t}\left(X_0^H+\sigma\int_0^t e^{\lambda s}\,dB_s^H\right).
\end{aligned}$$

□

No caso em que $X_0^H=\sigma\displaystyle\int_{-\infty}^0 e^{\lambda s}\,dB_s^H$, segue imediatamente da estacionariedade e da normalidade dos incrementos do movimento Browniano fracionário que (3.24) é um processo estocástico estacionário Gaussiano.

O teorema a seguir apresenta a forma aproximada da função de autocovariância do processo de Ornstein-Uhlenbeck fracionário.

Teorema 3.1. *Sejam* $H\in\left\{\left(0,\frac{1}{2}\right)\cup\left(\frac{1}{2},1\right]\right\}$, $N\in\mathbb{N}$ *e* $\left\{X_t^H\right\}_{t\in[0,T]}$ *o processo de Ornstein-Uhlenbeck fracionário com a condição inicial* $X_0^H=\sigma\displaystyle\int_{-\infty}^0 e^{-\lambda s}\,dB_s^H$ *dado em* (3.25)*. Então, para* t *fixo e* $h\to\infty$*, a função de autocovariância de* $\left\{X_t^H\right\}_{t\in[0,T]}$ *é dada por*

$$\gamma_{_X}(h)=\frac{1}{2}\sigma^2\sum_{n=1}^{N}\lambda^{-2n}\left[\prod_{k=0}^{2n-1}(2H-k)\right]h^{2H-2n}+o\left(h^{2H-2N-2}\right).$$

Demonstração: Ver Cheridito et al. (2003), página 8. □

3.4 Equação de Langevin Generalizada e sua Solução

A equação de Langevin generalizada é uma equação íntegro-diferencial estocástica, surgida em 1965 com Hazime Mori e em 1966 com Ryogo Kubo, que modela a

dinâmica de uma partícula Browniana em um meio fluido no qual a massa das moléculas não seja muito menor comparada à massa da partícula. Sua equação, para uma partícula de massa unitária, é dada por

$$\begin{cases} V'(t) = -\int_0^t \gamma(t-s)\, V(s)\, ds + \eta(t) \\ V(0) = V_0, \end{cases} \tag{3.30}$$

onde $V = \{V(t)\}_{t\geq 0}$ é um processo estocástico que representa a velocidade da partícula e $\eta = \{\eta(t)\}_{t\geq 0}$ é um processo ruído branco, representando a força aleatória resultante das colisões entre as moléculas e a partícula. A função $\gamma(\cdot)$, que aparece no interior da integral em (3.30), é chamada de *função memória*, e representa o efeito de memorização na evolução do movimento da partícula.

Quando a função memória é dada por

$$\gamma(t-s) = \gamma\delta(t-s) = \begin{cases} \lambda, \text{ se } t = s, \\ 0, \text{ caso contrário,} \end{cases} \tag{3.31}$$

onde $\lambda > 0$ e $\eta = \{\eta(t)\}_{t\geq 0}$ é o processo movimento Browniano, recaímos na equação de Langevin Clássica que estudamos na Seção 3.1.

Em 1977, Kannan estudou as equações de Langevin generalizadas utilizando o contexto da análise no espaço de Hilbert $L_2(\Omega, \mathcal{F}, \mathbb{P})$ das variáveis aleatórias quadrado integráveis, sendo $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade completo.

Assim, dada a equação generalizada de Langevin apresentada em (3.30), o autor supôs que

i) o ruído $\eta(t)$ é uma função $\eta : (0, \infty) \to L_2(\Omega, \mathcal{F}, \mathbb{P})$, ou seja, $\mathbb{E}[\eta^2(t)]$ é finita, para todo $t \geq 0$;

ii) $\eta(t)$ é uma função localmente integrável, para todo $t \geq 0$;

iii) $\mathbb{E}(V_0^2)$ é finita.

A abordagem adotada por Kannan tem como vantagem contextualizar o estudo das equações de Langevin generalizadas na teoria das equações íntegro-diferenciais em espaços de Banach, mais especificamente, no espaço de Hilbert $L^2(\Omega, \mathcal{F}, \mathbb{P})$ das variáveis aleatórias quadrado integráveis.

O teorema a seguir apresenta critérios suficientes para que a equação (3.30) possua um solução e que ela seja única.

Teorema 3.2 (Teorema de Existência e Unicidade de Kannan)**.** *Sejam* $V_0 \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, $\eta \in \mathcal{L}[\mathbb{R}_+, L^2(\Omega, \mathcal{F}, \mathbb{P})]$ *e* $\gamma \in \mathcal{L}_\infty[\mathbb{R}_+, \mathbb{R}]$, *onde* $\mathcal{L}$ *é o espaço de todas as funções localmente integráveis de* $\mathbb{R}_+$ *em* $L^2(\Omega, \mathcal{F}, \mathbb{P})$ *e* $\mathcal{L}_\infty$ *é o espaço de todas as funções reais essencialmente limitadas em* $[0, T]$, *para cada* $T > 0$. *Então,* (3.30) *possui uma única solução em média quadrática.*

Demonstração: Ver Kannan (1977). □

Kannan (1977) obtém uma forma explícita para tal solução.

Teorema 3.3 (Fórmula de Representação de Kannan)**.** *Supondo válida as hipóteses do Teorema 3.2, então* $V(t)$ *é uma solução em média quadrática da equação de Langevin generalizada* (3.30) *se e somente se*

$$V(t) = \rho(t) V_0 + \int_0^t \rho(t-s)\, d\eta(s), \tag{3.32}$$

onde a função $\rho(\cdot)$ *satisfaz a seguinte equação íntegro-diferencial determinística*

$$\begin{cases} \rho'(t) = -\displaystyle\int_0^t \gamma(t-s)\, \rho(s)\, ds \\ \rho(0) = 1. \end{cases} \tag{3.33}$$

Demonstração: Ver Kannan (1977). □

Exemplo 3.1. Consideramos a equação de Langevin generalizada

$$\begin{cases} X_t' = -\displaystyle\int_0^t \gamma(t-s)\, X_s\, ds + dB_t \\ X(0) = X_0, \end{cases} \tag{3.34}$$

para $t \in [0, T]$, com função memória dada por

$$\gamma(t-s) = \begin{cases} k(\lambda), \text{ se } t = s \\ 0, \text{ caso contrário,} \end{cases} \tag{3.35}$$

onde $k(\lambda)$ é uma constante real positiva indexada por um parâmetro λ positivo.

Pelo Teorema 3.3, temos que a solução de (3.34) é dada por

$$X_t = \rho(t) X_0 + \int_o^t \rho(t-s)\, dB_s,$$

onde a função $\rho(\cdot)$ satisfaz a equação íntegro-diferencial

$$\begin{cases} \rho'(t) = -\int_0^t \gamma(t-s)\,\rho(s)\, ds \\ \rho(0) = 1. \end{cases} \tag{3.36}$$

Substituindo (3.35) na equação (3.36), obtemos a equação diferencial

$$\begin{cases} \rho'(t) = -k(\lambda)\,\rho(t) \\ \rho(0) = 1, \end{cases}$$

cuja solução é dada por

$$\rho(t) = e^{-k(\lambda)t}. \tag{3.37}$$

Logo, a solução de (3.34) é dada por

$$X_t = e^{-k(\lambda)\,t} X_0 + e^{-k(\lambda)\,t} \int_0^t e^{k(\lambda)s} dB_s. \tag{3.38}$$

A equação de Langevin Clássica (3.1), considerada no Capítulo 3, é equivalente à equação (3.34), apresentada no Exemplo 3.1, com $k(\lambda) = \lambda$.

Capítulo 4

Estimação

Neste capítulo apresentamos os estimadores que utilizamos nas simulações apresentadas no Capítulo 5. Para a estimação do parâmetro λ, na solução da equação de Langevin clássica, apresentamos os estimadores advindos da função de autocorrelação amostral, o de mínimos quadrados (lse) e o de máxima verossimilhança (mle). Na Seção 4.2 apresentamos um método para estimar o índice de estabilidade α do ruído do processo advindo da solução da equação de Langevin generalizada quando ele tem distribuição α-estável.

4.1 Estimação de λ na Solução da Equação de Langevin Clássica

4.1.1 Estimadores Advindos da Função de Autocorrelação Amostral

Seja $\{X_t\}_{t\in[0,T]}$ o processo advindo da solução da equação de Langevin clássica. Considerando $h \in \mathbb{R}$, seja S o conjunto de observações discretas do processo

$\{X_t\}_{t\in[0,T]}$, tal que

$$S = \{X_0,\, X_h,\, X_{2h}, \cdots,\, X_{kh}, \cdots,\, X_T\} = \left\{S_0,\, S_1,\, \cdots,\, S_{\frac{T}{h}}\right\} \equiv \{S_t\}_{t=0}^{\left\lfloor \frac{T}{h} \right\rfloor}.$$

Na Proposição 3.2 vimos que a função de autocorrelação de S é dada por

$$\rho_{_S}(k) = e^{-\lambda|k|h}. \tag{4.1}$$

Isolando o parâmetro λ em (4.1), obtemos o estimador $\hat{\lambda}$ em termos da função de autocorrelação amostral do processo $\{S_t\}_{t=0}^{\left\lfloor \frac{T}{h} \right\rfloor}$ dado por

$$\hat{\lambda} = -\frac{\ln\left(|\hat{\rho}_{_S}(k)\,|\right)}{|k|h}. \tag{4.2}$$

Considerando $k = 1$ (o processo é um autoregressivo de ordem 1), obtemos o primeiro estimador de λ proposto por Valdivieso et al. (2009)

$$\hat{\lambda}_1 = -\frac{\ln\left(|\hat{\rho}_{_S}(1)\,|\right)}{h}, \tag{4.3}$$

onde $\hat{\rho}_{_S}(\cdot)$ é a função de autocorrelação amostral (2.11) para o processo $\{S_t\}_{t=0}^{\left\lfloor \frac{T}{h} \right\rfloor}$.

Outro estimador proposto por Valdivieso et al. (2009) consiste em calcular o λ que minimiza a soma do quadrado da diferença entre a função de autocorrelação amostral de $\{S_t\}_{t=0}^{\left\lfloor \frac{T}{h} \right\rfloor}$ e $e^{-\lambda kh}$, ou seja,

$$\hat{\lambda}_2 = arg\min_{\lambda} \sum_{k=1}^{m^*} \left(\hat{\rho}_{_S}(k) - e^{-\lambda kh}\right)^2, \text{ para algum } m^* \text{ tal que } m^* < n, \tag{4.4}$$

onde $\hat{\rho}_{_S}(\cdot)$ é a função de autocorrelação amostral (2.11) e $n = \left\lfloor \frac{T}{h} \right\rfloor$.

4.1.2 Estimador de Mínimos Quadrados

No método de regressão por mínimos quadrados, o objetivo é escrever uma sequência de variáveis aleatórias $\{X_t\}_{t=1}^{n}$ como uma combinação afim de uma sequência de dados $\{Y_t\}_{t=1}^{n}$, ou seja

$$X_t = aY_t + b + \varepsilon_t, \tag{4.5}$$

onde $\{\varepsilon_t\}_{t=1}^{n}$ é o conjunto de erros, independente e identicamente distribuídos.

O objetivo é estimar a e b de forma a minimizar a soma dos erros ao quadrado, ou seja, calcular

$$\min_{a,b} \sum_{t=1}^{n} \varepsilon_t^2 = \min_{a,b} \sum_{t=1}^{n} \left(X_t - aY_t - b\right)^2, \tag{4.6}$$

o que equivale a solucionar as equações

$$\frac{\partial \left(X_t - aY_t - b\right)^2}{\partial a} = 0 \quad \text{e} \quad \frac{\partial \left(X_t - aY_t + b\right)^2}{\partial b} = 0. \tag{4.7}$$

Da segunda equação em (4.7) obtemos que

$$\begin{aligned}
& \sum_{t=1}^{n} 2\left(X_t - aY_t - b\right)(-1) = 0 \Leftrightarrow \sum_{t=1}^{n} X_t - \hat{a} \sum_{t=1}^{n} Y_t - n\,\hat{b} = 0 \\
\Leftrightarrow \quad & n\,\overline{X} - \hat{a}\, n\,\overline{Y} = n\,\hat{b} \\
\Leftrightarrow \quad & \hat{b} = \overline{X} - \hat{a}\,\overline{Y},
\end{aligned}$$

onde $\overline{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$ e $\overline{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t$. Da primeira equação em (4.7) obtemos que

$$\begin{aligned}
& \sum_{t=1}^{n} 2\left(X_t - aY_t - b\right)(-Y_t) = 0 \Leftrightarrow \sum_{t=1}^{n} X_t Y_t - \hat{a} \sum_{t=1}^{n} Y_t^2 - \hat{b} \sum_{t=1}^{n} Y_t = 0 \\
\Leftrightarrow \quad & \sum_{t=1}^{n} X_t Y_t - \hat{a} \sum_{t=1}^{n} Y_t^2 - \left(\overline{X} - \hat{a}\,\overline{Y}\right) \sum_{t=1}^{n} Y_t = 0 \\
\Leftrightarrow \quad & \hat{a} \left(\sum_{t=1}^{n} Y_t^2 - n\,\overline{Y}^2 \right) = \sum_{t=1}^{n} X_t Y_t - n\,\overline{X}\,\overline{Y} \\
\Leftrightarrow \quad & \hat{a} = \frac{\displaystyle\sum_{t=1}^{n} X_t Y_t - n\,\overline{X}\,\overline{Y}}{\displaystyle\sum_{t=1}^{n} Y_t^2 - n\,\overline{Y}^2} = \frac{\displaystyle\sum_{t=1}^{n} \left(X_t - \overline{X}\right)\left(Y_t - \overline{Y}\right)}{\displaystyle\sum_{t=1}^{n} \left(Y_t - \overline{Y}\right)^2}.
\end{aligned} \tag{4.8}$$

Neter e Wasserman (1974) mostram que o estimador não viciado da variância σ_ε^2 do processo ruído $\{\varepsilon_t\}_{t\in[0,T]}$ é dado por

$$\hat{\sigma}_\varepsilon^2 = \frac{\displaystyle\sum_{t=1}^{n} \left(X_t - aY_t - b\right)^2}{n-2}, \tag{4.9}$$

ou, equivalentemente,

$$\hat{\sigma}_{\varepsilon}^{2} = \frac{1}{(n-2)} \left(\sum_{t=1}^{n} X_t^2 - \frac{1}{n} \left(\sum_{t=1}^{n} X_t \right)^2 - \hat{a} \right).$$

Considere o processo de Ornstein-Uhlenbeck $\{X_t\}_{t\in[0,T]}$, para algum $T \in \mathbb{R}_+$, solução da equação de Langevin clássica (3.1), apresentado em (3.5). Na Proposição 3.1 obtemos uma forma recursiva para a solução (3.5), dada por

$$X_{t+h} = e^{-\lambda h} \left(X_t + \sigma e^{-\lambda t} \int_t^{t+h} e^{\lambda s}\, dB_s \right), \text{ para } h \in \mathbb{R}_+. \tag{4.10}$$

Considerando

$$\varepsilon_{t+h} = \sigma e^{-\lambda(t+h)} \int_t^{t+h} e^{\lambda s}\, dB_s,\ a = e^{-\lambda h} \text{ e } b = 0, \tag{4.11}$$

podemos reescrever (4.10) como

$$X_{t+h} = aX_t + b + \varepsilon_{t+h}, \text{ para } h \in \mathbb{R}_+, \tag{4.12}$$

ou seja,

$$X_t = aX_{t-h} + \varepsilon_t, \text{ para } h \in \mathbb{R}_+. \tag{4.13}$$

Apresentamos a seguir um importante lema que mostra que o processo $\{\varepsilon_t\}_{t\in[0,T]}$, dado em (4.11), é um ruído branco Gaussiano.

Lema 4.1. *O processo $\{\varepsilon_t\}_{t\in[0,T]}$ é um ruído branco Gaussiano.*

Demonstração: Observe que as variáveis aleatórias ε_t, para todo $t \in [0,T]$, possuem distribuição Gaussiana, tais que

$$\mathbb{E}(\varepsilon_t) = \mathbb{E}\left(\sigma e^{-\lambda t} \int_{t-h}^{t} e^{\lambda h} dB_s \right) = \sigma e^{-\lambda t} \mathbb{E}\left(\int_{t-h}^{t} e^{\lambda h} dB_s \right) = 0 \tag{4.14}$$

e

$$\begin{aligned} Var(\varepsilon_t) &= Var\left(\sigma e^{-\lambda t} \int_{t-h}^{t} e^{\lambda s}\, dB_s \right) \\ &= \sigma^2 e^{-2\lambda} Var\left(\int_{t-h}^{t} e^{\lambda s}\, dB_s \right). \end{aligned} \tag{4.15}$$

Como a integral de Itô tem esperança zero, aplicando a propriedade da isometria (2.55) na expressão (4.15), obtemos

$$\begin{aligned} Var\left(\varepsilon_t\right) &= \sigma^2 e^{-2\lambda t}\mathbb{E}\left[\left(\int_{t-h}^{t} e^{\lambda s}\, dB_s\right)^2\right] \\ &= \sigma^2 e^{-2\lambda t}\int_{t-h}^{t} e^{2\lambda s}\, ds = \sigma^2 e^{-2\lambda t}\left(\frac{e^{2\lambda t}}{2\lambda} - \frac{e^{2\lambda(t-h)}}{2\lambda}\right) \\ &= \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda h}\right), \end{aligned}$$

ou seja, a variância de ε_t, para todo $t \in [0,T]$, não depende de t. Logo, $\{\varepsilon_t\}_{t\in[0,T]}$ é um processo estocástico constituído por variáveis aleatórias independentes e identicamente distribuídas, com distribuição comum Gaussiana. Ou seja, $\{\varepsilon_t\}_{t\in[0,T]}$ é um ruído branco Gaussiano. □

Voltando à equação de regressão linear dada pela expressão (4.5), consideramos $X = \{X_t\}_{t=1}^{T}$ e $Y = \{X_{t-1}\}_{t=1}^{T}$. Desta forma, o estimador $\hat{a}$ em (4.8) pode ser reescrito por

$$\hat{a} = \frac{\displaystyle\sum_{t=1}^{T}\left(X_t - \overline{X}\right)\left(X_{t-1} - \overline{Y}\right)}{\displaystyle\sum_{t=1}^{T}\left(X_{t-1} - \overline{Y}\right)^2},$$

onde $\overline{X} = \frac{1}{T}\displaystyle\sum_{t=1}^{T} X_t$ e $\overline{Y} = \frac{1}{T}\displaystyle\sum_{t=1}^{T} X_{t-1}$.

Como $a = e^{i\lambda h}$, o estimador de mínimos quadrados para o parâmetro λ, quando $h = 1$, é dado por

$$\hat{\lambda} = -\ln\left(\frac{\displaystyle\sum_{t=1}^{T}\left(X_t - \overline{X}\right)\left(X_{t-1} - \overline{Y}\right)}{\displaystyle\sum_{t=1}^{T}\left(X_{t-1} - \overline{Y}\right)^2}\right). \tag{4.16}$$

No caso geral, para $h \in \mathbb{R}_+$, temos que

$$\hat{\lambda} = -\frac{1}{h} \ln \left(\frac{\sum_{k=1}^{\lfloor \frac{T}{h} \rfloor} \left(X_{k\,h} - \overline{X} \right) \left(X_{(k-1)h} - \overline{Y} \right)}{\sum_{k=1}^{\lfloor \frac{T}{h} \rfloor} \left(X_{(k-1)h} - \overline{Y} \right)^2} \right). \tag{4.17}$$

Da demonstração do Lema 4.1 obtemos que

$$Var\left(\varepsilon_t\right) = \frac{\sigma^2}{2\lambda} \left(1 - e^{2\lambda h}\right),$$

o que implica em um estimador de σ^2 dado por

$$\hat{\sigma}^2 = \frac{2\hat{\lambda} \widehat{Var}\left(\varepsilon_t\right)}{1 - e^{-2\hat{\lambda} h}}, \tag{4.18}$$

onde

$$\widehat{Var}\left(\varepsilon_t\right) = \frac{1}{(n-2)} \left(\sum_{k=1}^{\frac{T}{h}} X_{k\,h}^2 - \frac{1}{n} \left(\sum_{k=1}^{\frac{T}{h}} X_{k\,h} \right)^2 - \hat{a} \right).$$

4.1.3 Estimador de Máxima Verossimilhança

Seja $\{X_t\}_{t \in [0,T]}$ o processo advindo da solução da equação de Langevin clássica. Considerando $\{S_t\}_{t=0}^{\lfloor \frac{T}{h} \rfloor}$ o processo definido em (3.16), dado o ponto inicial $S_0 = s_0$, temos que

$$\mathbb{E}\left(S_1 | S_0\right) = \mathbb{E}\left(e^{-\lambda h} s_0 + e^{-\lambda h} \sigma \int_0^h e^{\lambda s}\, dB_s \right) = e^{-\lambda h} s_0.$$

Com cálculos equivalentes aos efetuados em (4.15) obtemos que

$$Var\left(S_1 | S_0\right) = Var\left(e^{-\lambda h} \sigma \int_0^h e^{\lambda s}\, dB_s \right) = \frac{\sigma^2}{2\lambda} \left(1 - e^{-2\lambda h}\right).$$

Logo, $S_1 | S_0 \sim \mathcal{N}\left(e^{-\lambda h} s_0,\, \frac{\sigma^2}{2\lambda} \left(1 - e^{-2\lambda h}\right) \right)$.

Consequentemente, estendendo para um $0 < t \leq \lfloor \frac{T}{h} \rfloor$ qualquer, obtemos que

$$S_t | S_{t-1} \sim \mathcal{N}\left(\mu s_{t-1},\, \phi^2\right), \tag{4.19}$$

onde $\mu = e^{-\lambda h}$ e $\phi^2 = \dfrac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda h}\right)$.

Logo, a função de verossimilhança do processo $\{S_t|S_{t-1}\}_{t=1}^{\lfloor \frac{T}{h} \rfloor}$ é dada por

$$L\left(\mu, \phi^2\right) = \prod_{i=1}^{n} f_{S_i|S_{i-1}}\left(s_i\right) = \left(2\pi\phi^2\right)^{-\frac{n}{2}} \prod_{i=1}^{n} exp\left\{\frac{\left(s_i - \mu s_{i-1}\right)^2}{2\phi^2}\right\}, \tag{4.20}$$

onde $n = \lfloor \frac{T}{h} \rfloor$.

Dessa forma, a função de log-verossimilhança é dada por

$$\mathcal{L}\left(\mu, \phi^2\right) = -\frac{n}{2}\ln\left(2\pi\phi^2\right) - \sum_{i=1}^{n} \frac{\left(s_i - \mu s_{i-1}\right)^2}{2\phi^2}.$$

Para estimar os parâmetros μ e ϕ^2, desejamos maximizar a função $\mathcal{L}\left(\mu, \phi^2\right)$, o que equivale a solucionar as equações

$$\frac{\partial \mathcal{L}\left(\mu, \phi^2\right)}{\partial \mu} = 0 \text{ e } \frac{\partial \mathcal{L}\left(\mu, \phi^2\right)}{\partial \phi^2} = 0.$$

Por exemplo, para obter o estimador $\hat{\mu}$, temos que

$$\frac{\partial \mathcal{L}\left(\mu, \phi^2\right)}{\partial \mu} = 0 \Leftrightarrow \sum_{i=1}^{n} \frac{\partial\left[\left(s_i - \mu s_{i-1}\right)^2\right]}{\partial \mu} = 0$$

$$\Leftrightarrow \sum_{i=1}^{n} 2\left(s_i - \hat{\mu} s_{i-1}\right)\left(-s_{i-1}\right) = 0 \Leftrightarrow \sum_{i=1}^{n}\left(s_i s_{i-1} - \hat{\mu} s_{i-1}^2\right) = 0$$

$$\Leftrightarrow \hat{\mu}\sum_{i=1}^{n} s_{i-1}^2 = \sum_{i=1}^{n} s_i s_{i-1} \Leftrightarrow \hat{\mu} = \frac{\sum_{i=1}^{n} s_i s_{i-1}}{\sum_{i=1}^{n} s_{i-1}^2}.$$

Da expressão (4.19), como $\mu = e^{-\lambda h}$, temos que

$$\hat{\lambda} = -\frac{1}{h}\ln\left(\left|\frac{\sum_{i=1}^{n} s_i s_{i-1}}{\sum_{i=1}^{n} s_{i-1}^2}\right|\right), \tag{4.21}$$

ou, escrevendo em termos do processo $\{X_{th}\}_{t=1}^{\lfloor \frac{T}{h} \rfloor}$,

$$\hat{\lambda} = -\frac{1}{h}\ln\left(\left|\frac{\sum_{i=1}^{n} X_{ih} X_{(i-1)h}}{\sum_{i=1}^{n} X_{(i-1)h}^2}\right|\right), \tag{4.22}$$

onde $n = \lfloor \frac{T}{h} \rfloor$.

Observação 4.1. No caso em que os ruídos possuem distribuição normal, temos que os métodos de estimação por mínimos quadrados e de máxima verossimilhança são equivalentes.

4.2 Estimação de α na Solução da Equação de Langevin com Ruído α-estável

Seja $\{X_t\}_{t\in[0,T]}$ o processo advindo da solução da equação de Langevin com ruído α-estável, dado por

$$X_{t+h} = e^{-\lambda h} X_t + \sigma e^{-\lambda h} \int_0^h e^{\lambda s}\, dZ_s, \tag{4.23}$$

onde $\{Z_t\}_{t\in[0,T]}$ é um processo α-estável, ou seja, $Z \sim S_\alpha(\sigma^*, \beta, \mu)$.

Da expressão (4.23) obtemos que

$$X_{t+h} - e^{-\lambda h} X_t = \sigma e^{-\lambda h} \int_0^h e^{\lambda s}\, dZ_s,$$

ou seja,

$$e^{\lambda h} X_{t+h} - X_t = \sigma \int_0^h e^{\lambda s}\, dZ_s. \tag{4.24}$$

Aproximando a integral $\displaystyle\int_0^h e^{\lambda s}\, dZ_s$ por um somatório, obtemos a forma discreta

$$\sigma \int_0^h e^{\lambda s}\, dZ_s \approx \sigma \sum_{i=1}^{m} e^{\lambda s_i} \left(Z_{s_i} - Z_{s_{i-1}}\right),$$

para m suficientemente grande, onde a sequência de números reais $(s_i)_{i=0}^m$ é uma partição, igualmente espaçada, do intervalo $[0, h]$. Consideramos aqui $s_i = \dfrac{1}{i}$, ou seja, a partição é formada por pontos equidistantes.

Aplicando as Proposições 2.5 e 2.6, obtemos que $\sigma \displaystyle\sum_{i=1}^{m} e^{\lambda s_i} \left(Z_{s_i} - Z_{s_{i-1}}\right)$ também é um processo α-estável, com o mesmo índice de estabilidade α do processo $\{Z_t\}_{t\in[0,T]}$.

Dessa forma, a partir da expressão (4.24), o objetivo é estimar o valor do parâmetro α do processo ruído $\{Z_t\}_{t\in[0,T]}$, através de um algoritmo de dois passos, descrito a seguir.

Algoritmo de Dois Passos para Estimar α:

Passo 1: Estimamos o parâmetro λ, obtendo $\hat{\lambda}$;

Passo 2: Estimamos o parâmetro α através da expressão $e^{\hat{\lambda}h}X_{t+h} - X_t = \sigma \int_0^h e^{\hat{\lambda}s}\, dZ_s$.

Para estimar o parâmetro λ, podemos utilizar qualquer um dos estimadores propostos na Seção 4.1. Para estimar o parâmetro α, apresentamos o estimador proposto por Press (1972).

Vimos na Definição 2.23 que se $X \sim S_\alpha(\sigma, \beta, \mu)$, então sua função característica é dada por

$$\varphi_X(t) = \mathbb{E}\left(e^{itX}\right) = \begin{cases} exp\left\{-\sigma|t|\left[1 + i\beta\frac{2}{\pi}S(t)\ln(|t|)\right] + i\mu t\right\}, \text{ se } \alpha = 1, \\ exp\left\{-\sigma^\alpha|t|^\alpha\left[1 - i\beta S(t)\tan\left(\frac{\pi\alpha}{2}\right)\right] + i\mu t\right\}, \text{ caso contrário,} \end{cases} \tag{4.25}$$

onde $S(\cdot)$ é a função sinal, definida por

$$S(t) = \begin{cases} -1, \text{ se } t < 0, \\ 0, \text{ se } t = 0, \\ 1, \text{ se } t > 0. \end{cases} \tag{4.26}$$

Para todo $\alpha \neq 1$ obtemos que

$$\begin{aligned} |\varphi_X(t)| &= |exp\left\{-\sigma^\alpha|t|^\alpha\left[1 - i\beta S(t)\tan\left(\frac{\pi\alpha}{2}\right)\right] + i\mu t\right\}| \\ &= |exp\left(-\sigma^\alpha|t|^\alpha\right) exp\left\{-i\sigma^\alpha|t|^\alpha\beta S(t)\tan\left(\frac{\pi\alpha}{2}\right) + i\mu t\right\}| \\ &= e^{-\sigma^\alpha|t|^\alpha}. \end{aligned}$$

Ou seja,

$$\sigma^\alpha|t|^\alpha = -\ln(|\varphi_X(t)|). \tag{4.27}$$

Dados $0 < t_1 < t_2$, da expressão (4.27) obtemos o seguinte sistema de duas equações.

$$\begin{cases} \sigma^{\alpha}|t_1|^{\alpha} = -\ln\left(|\varphi_X\left(t_1\right)|\right) \\ \sigma^{\alpha}|t_2|^{\alpha} = -\ln\left(|\varphi_X\left(t_2\right)|\right). \end{cases} \tag{4.28}$$

Dividindo a primeira equação do sistema (4.28) pela segunda equação do mesmo sistema, obtemos

$$\left(\frac{|t_2|}{|t_1|}\right)^{\alpha} = \frac{\ln\left(|\varphi_X\left(t_2\right)|\right)}{\ln\left(|\varphi_X\left(t_1\right)|\right)}. \tag{4.29}$$

Aplicando o logarítmo neperiano em ambos os lados da igualdade (4.29), obtemos um estimador para α, dado por

$$\hat{\alpha} = \frac{\ln\left(\frac{\ln\left(|\hat{\varphi}_X(t_1)|\right)}{\ln\left(|\hat{\varphi}_X(t_2)|\right)}\right)}{\ln\left(\frac{t_1}{t_2}\right)}, \tag{4.30}$$

onde $\hat{\varphi}_X\left(\cdot\right)$ é a função característica amostral do processo $\{X_t\}_{t\in[0,T]}$, dada por

$$\hat{\varphi}_X\left(t\right) = \frac{1}{n}\sum_{j=1}^{n} e^{itX_j}. \tag{4.31}$$

Capítulo 5

Simulações

Nesse capítulo apresentamos simulações do processo advindo da solução da equação de Langevin clássica e da equação de Langevin generalizada. Apresentamos os resultados da estimação do parâmetro λ através dos estimadores propostos por Valdivieso et al. (2009), do estimador de máxima verossimilhança e do estimador de mínimos quadrados. Através das funções de autocorrelação e periodograma, analisamos a existência de longa dependência no processo dado pela expressão (3.5). Na Seção 5.3 apresentamos os resultados da estimação do parâmetro α no processo advindo da solução da equação de Langevin generalizada com ruído α-estável. A parte computacional deste capítulo foi realizada através do aplicativo R-project.

5.1 Estimação de λ na Solução da Equação de Langevin Clássica

Para a simulação de processo (3.5) no intervalo $[0, T]$, consideramos a sua versão recursiva e discreta dada por

$$X_{t+h} = e^{-\lambda h} \lim_{m\to\infty} \left(X_t + \sigma \sum_{i=1}^{m} e^{\lambda s_i} \left(B_{s_i} - B_{s_{i-1}} \right) \right), \tag{5.1}$$

onde a sequência de números reais $(s_i)_{i=0}^m$ é uma partição, igualmente espaçada, do intervalo $[0, h]$, apresentada em (3.21).

Como o processo obtido da solução da equação de Langevin clássica é contínuo, é interessante simular o processo para pontos tão próximos quanto se queira. Assim, consideramos o processo (3.5) na forma

$$X_{th} = e^{-\lambda h}\left(X_{(t-1)h} + \sigma\sum_{i=1}^{m} e^{\lambda s_i}\left(B_{s_i} - B_{s_{i-1}}\right)\right), \tag{5.2}$$

para m suficientemente grande, onde a sequência de números reais $(s_i)_{i=0}^m$ é uma partição, igualmente espaçada, do intervalo $[0, h]$. Consideramos aqui $s_i = \dfrac{1}{i}$, ou seja, a partição é formada por pontos equidistantes.

Os estimadores (4.3) e (4.4) propostos por Valdivieso et al. (2009), denotados por $\hat{\lambda}_{1,j}$ e $\hat{\lambda}_{2,j}$, respectivamente, são definidos por

$$\hat{\lambda}_{1,j} = -\frac{\ln\left(\left|\hat{\rho}_X\left(1\right)\right|\right)}{h},$$

$$\hat{\lambda}_{2,j} = arg\min_{\lambda}\sum_{k=1}^{m^*}\left(\hat{\rho}_X\left(k\right) - e^{-\lambda k h}\right)^2, \text{ para algum } m^* = m^*\left(n\right) \text{ tal que } m^* < n,$$

onde $\hat{\rho}_X\left(\cdot\right)$ é a função de autocorrelação amostral (2.11) e o j indica a aplicação na j-ésima replicação do processo $\{X_{th}\}_{t=0}^n$, onde $n = \left\lfloor \frac{T}{h} \right\rfloor$.

O estimador de máxima verossimilhança definido em (4.22), denotado por $\hat{\lambda}_{3,j}$, é definido por

$$\hat{\lambda}_{3,j} = -\frac{1}{h}\ln\left(\left|\frac{\sum_{i=1}^{n} X_{ih}X_{(i-1)h}}{\sum_{i=1}^{n} X_{(i-1)h}^2}\right|\right),$$

onde $n = \left\lfloor \frac{T}{h} \right\rfloor$ e j indica a aplicação na j-ésima replicação do processo.

Consideramos re o número de replicações do processo simulado. Portanto, de-

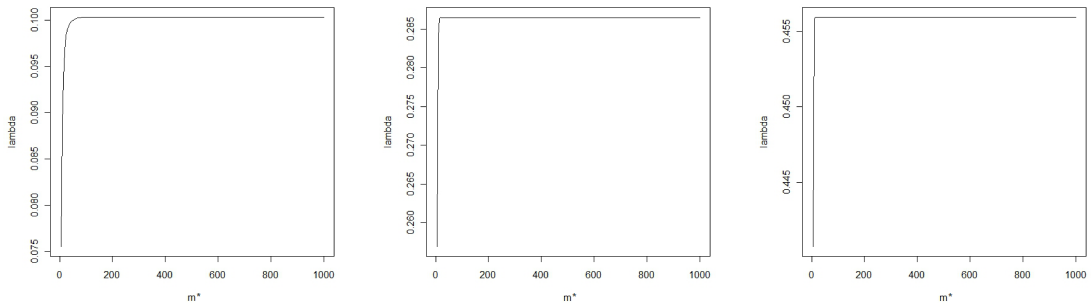
finimos os estimadores

$$\hat{\lambda}_1 = \frac{1}{re}\sum_{j=1}^{re}\hat{\lambda}_{1,j}, \tag{5.3}$$

$$\hat{\lambda}_2 = \frac{1}{re}\sum_{j=1}^{re}\hat{\lambda}_{2,j}, \tag{5.4}$$

$$\hat{\lambda}_3 = \frac{1}{re}\sum_{j=1}^{re}\hat{\lambda}_{3,j}. \tag{5.5}$$

O estimador $\hat{\lambda}_2$ depende de um ponto de truncamento m^*, que por sua vez depende do tamanho amostral n. Dessa forma, para encontrarmos o m^* ideal, plotamos o gráfico de m^* versus $\hat{\lambda}_2\,(m^*)$.



(a) (b) (c)

Figura 5.1: *Análise do Estimador* $\hat{\lambda}_2\,(m^*)$ *para Séries Temporais do Processo* $\{X_t\}_{t=1}^{1000}$*, com* $h = 1$*,* $\sigma = 1$ *e* $m = 1000$*: (a)* $\lambda = 0.1$*; (b)* $\lambda = 0.4$*; (c)* $\lambda = 0.9$*.*

A Figura 5.1 mostra que, para $h = 1$, $\hat{\lambda}_2\,(m^*)$ é uma função crescente e, com excessão do caso $\lambda = 0.1$, onde é possível avaliar pela Figura 5.1 que o m^* ideal localiza-se entre 20 e 50, a cota superior da função ficou muito aquém do esperado. Ou seja, m^* depende do valor de λ. Dessa forma, consideramos $m^* = 1000$, ou seja, o maior possível.

As simulações do processo (5.2) foram realizadas para $t \in [0, T]$. Consideramos três valores distintos de h: 1, 0.1 e 0.05. No caso $h = 1$, temos que $n = 1000$, onde n

representa o tamanho amostral. Para $h = 0.1$ e $h = 0.5$ obtemos, respectivamente, $n = 10000$ e $n = 20000$.

O parâmetro λ varia entre 0.1 e 0.9, com incrementos de tamanho 0.1. Consideramos o parâmetro σ fixo e igual a 1. O processo é simulado com 100 e 500 replicações.

Apresentamos os gráficos do processo simulado, da função de autocorrelação amostral e da função periodograma suavizado, onde a função de ponderação utilizada é o núcleo de Bartlett (2.18), para $\lambda \in \{0.1, 0.4, 0.9\}$. Todas as tabelas deste capítulo apresentam os valores estimados, com vício e erro médio quadrático (denotado por mse), onde os estimadores $\hat{\lambda}_i$, para $i \in \{1, 2, 3\}$ são obtidos à partir das expressões (5.3), (5.4) e (5.5), respectivamente, e

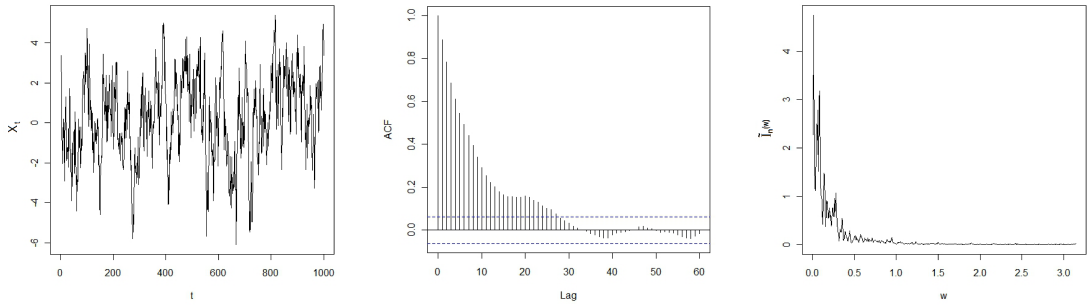
$$\text{vício}\left(\hat{\lambda}_i\right) = \frac{1}{re}\sum_{j=1}^{re}\left(\hat{\lambda}_{i,j} - \lambda\right) \text{ e}$$

$$\text{mse}\left(\hat{\lambda}_i\right) = \frac{1}{re}\sum_{j=1}^{re}\left(\hat{\lambda}_{i,j} - \lambda\right)^2,$$

para todo $i \in [1, 2, 3\}$.

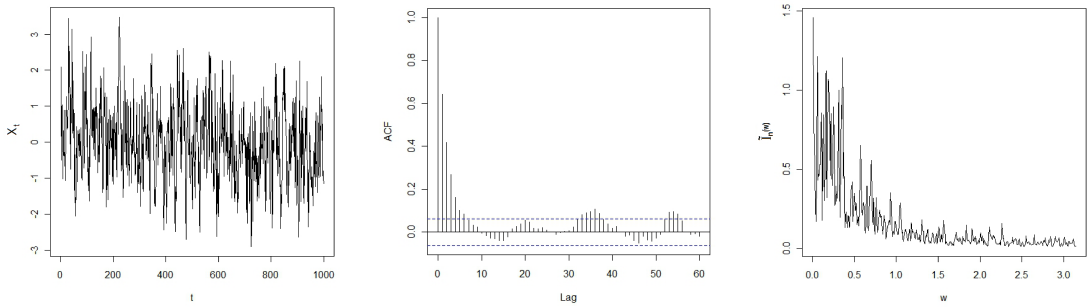
A seguir, apresentamos as simulações para os casos $n \in \{1000, 10000, 20000\}$.

Caso n = 1000.



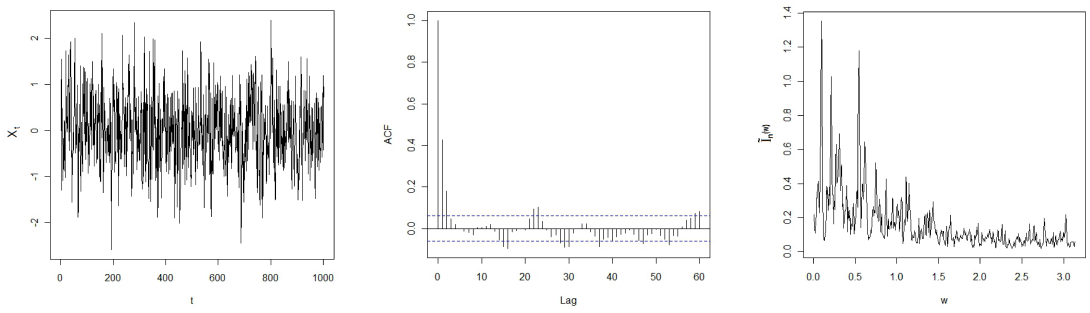
(a) (b) (c)

Figura 5.2: Geração do Processo (5.2) para $\lambda = 0.1$, $n = 1000$, $h = 1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.3: Geração do Processo (5.2) para $\lambda = 0.4$, $n = 1000$, $h = 1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.4: Geração do Processo (5.2) para $\lambda = 0.9$, $n = 1000$, $h = 1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.

Tabela 5.1: *Estimação do Parâmetro* λ*, quando* $n = 1000$*,* $\sigma = 1$*,* $h = 1$*,* $m = 1000$ *e* $m^* = 1000$.

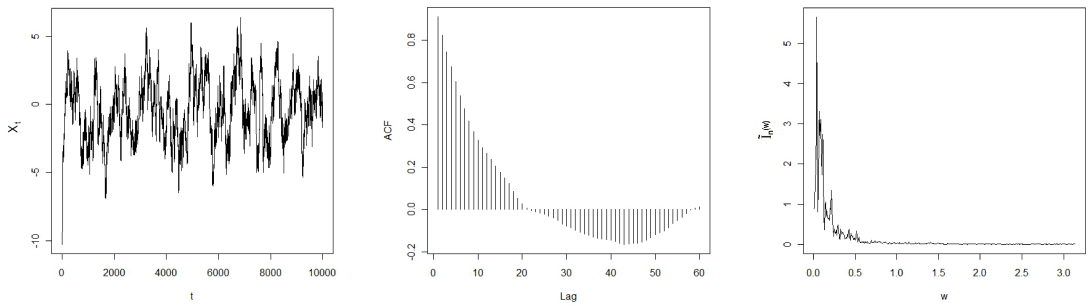
λ	$\hat{\lambda}_1$	vício$\left(\hat{\lambda}_1\right)$	mse$\left(\hat{\lambda}_1\right)$	$\hat{\lambda}_2$	vício$\left(\hat{\lambda}_2\right)$	mse$\left(\hat{\lambda}_2\right)$	$\hat{\lambda}_3$	vício$\left(\hat{\lambda}_3\right)$	mse$\left(\hat{\lambda}_3\right)$
$re = 100$									
0.1	0.1063	0.0063	0.0003	0.1006	0.0016	0.0003	0.1022	0.0022	0.0002
0.2	0.2071	0.0071	0.0006	0.1757	-0.0242	0.0012	0.2037	0.0037	0.0005
0.3	0.3072	0.0072	0.0009	0.2449	-0.0550	0.0040	0.3040	0.0040	0.0008
0.4	0.4022	0.0022	0.0016	0.3002	-0.0997	0.0113	0.3982	-0.0017	0.0016
0.5	0.5049	0.0049	0.0016	0.3569	-0.1430	0.0217	0.5009	0.0009	0.0015
0.6	0.6085	0.0085	0.0026	0.4098	-0.1901	0.0377	0.6040	0.0039	0.0025
0.7	0.7013	0.0013	0.0024	0.4610	-0.2389	0.0585	0.6970	-0.0029	0.0024
0.8	0.8037	0.0037	0.0049	0.5114	-0.2885	0.0858	0.7999	-0.0000	0.0049
0.9	0.9183	0.0183	0.0070	0.5623	-0.3376	0.1171	0.9140	0.0140	0.0067
$re = 500$									
0.1	0.1045	0.0045	0.0002	0.0985	-0.0014	0.0004	0.1016	0.0016	0.0002
0.2	0.2047	0.0047	0.0005	0.1772	-0.0227	0.0012	0.2011	0.0047	0.0005
0.3	0.3050	0.0050	0.0008	0.2415	-0.0584	0.0043	0.3015	0.0015	0.0008
0.4	0.4073	0.0073	0.0011	0.3046	-0.0953	0.0103	0.4039	0.0039	0.0011
0.5	0.5061	0.0061	0.0016	0.3619	-0.1380	0.0204	0.5023	0.0023	0.0016
0.6	0.6077	0.0077	0.0023	0.4151	-0.01848	0.0359	0.6037	0.0037	0.0023
0.7	0.7084	0.0084	0.0031	0.4641	-0.2358	0.0573	0.7041	0.0041	0.0031
0.8	0.8125	0.0125	0.0041	0.5148	-0.2851	0.0836	0.8082	0.0082	0.0042
0.9	0.9095	0.0095	0.0051	0.5646	-0.3353	0.1148	0.9051	0.0051	0.0051

Como podemos observar nas Figuras 5.2, 5.3 e 5.4, os processos simulados possuem uma função de autocorrelação amostral com decaimento exponencial mais lento quando $\lambda = 0.1$ e mais rápido à medida que aumenta o valor de λ. Através da função periodograma suavizado fica evidenciado a ausência de longa dependência.

Na Tabela 5.1 podemos verificar que os estimadores $\hat{\lambda}_1$ e $\hat{\lambda}_3$ apresentam bons resultados, com um vício e erro médio quadrático (mse) muito pequenos. Observe que, conforme aumentamos o valor de λ, há um pequeno aumento no vício, tanto para o estimador $\hat{\lambda}_1$ quanto para $\hat{\lambda}_3$. Observamos que o aumento do número de replicações não obteve melhora significativa na estimação.

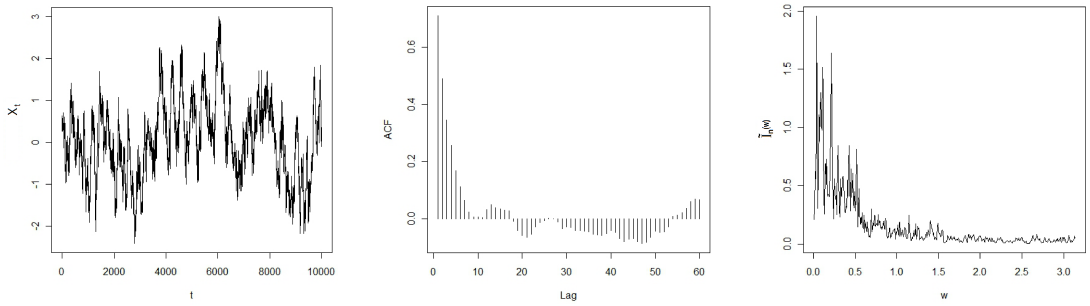
No caso do estimador $\hat{\lambda}_2$, obtivemos um resultado razoável apenas para $\lambda = 0.1$. Para os demais valores de λ, os resultados não são satisfatórios, pois apresentam vício e mse muito elevados.

Caso n = 10000.



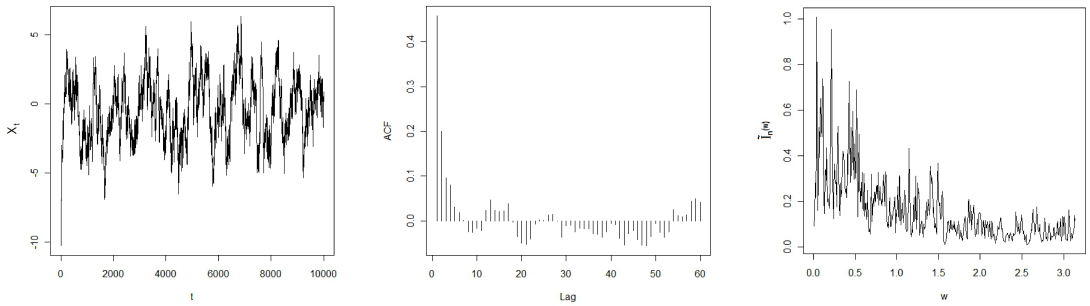
(a) (b) (c)

Figura 5.5: Geração do Processo (5.2) para $\lambda = 0.1$, $n = 10000$, $h = 0.1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

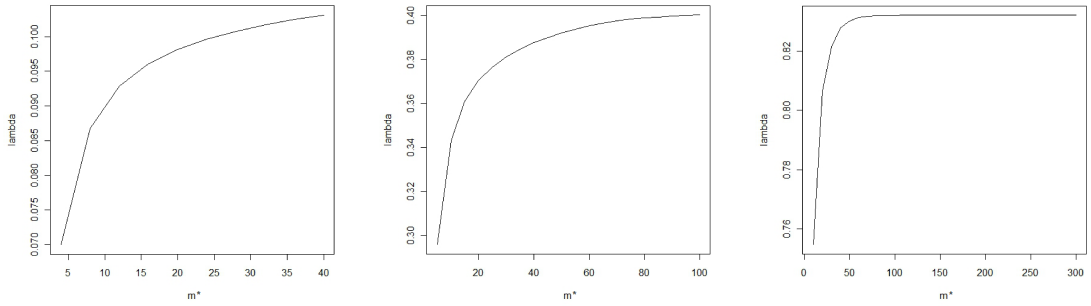
Figura 5.6: Geração do Processo (5.2) para $\lambda = 0.4$, $n = 10000$, $h = 0.1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.7: Geração do Processo (5.2) para $\lambda = 0.9$, $n = 10000$, $h = 0.1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.

Novamente, para encontrarmos o m^* ideal para o estimador $\hat{\lambda}_2$, plotamos o gráfico de m^* versus $\hat{\lambda}_2\,(m^*)$.



(a) (b) (c)

Figura 5.8: *Análise do Estimador* $\hat{\lambda}_2\,(m^*)$ *para Séries Temporais do Processo* $\{X_{th}\}_{t=1}^{10000}$, *com* $h = 0.1$, $\sigma = 1$ *e* $m = 1000$*: (a)* $\lambda = 0.1$*; (b)* $\lambda = 0.4$*; (c)* $\lambda = 0.9$.

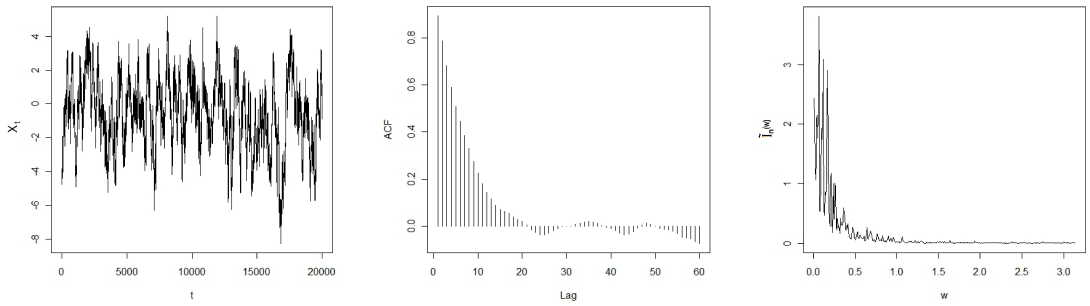
Nesse caso, temos que para $\lambda = 0.1$ um m^* entre 25 e 35 apresenta um bom resultado, enquanto que para $\lambda = 0.4$, para valores de m^* muito grandes, $\hat{\lambda}_2$ se aproxima de 0.4, enquanto que no caso $\lambda = 0.9$, temos que mesmo para valores elevados de m^* o estimador não se aproxima do seu valor nominal. Dessa forma, novamente não há um m^* ideal que contemple todos os casos. Assim, consideramos $m^* = 2000$.

Tabela 5.2: *Estimação do Parâmetro λ, quando $n = 10000$, $\sigma = 1$, $h = 0.1$, $m^* = 2000$ e $m = 1000$.*

λ	$\hat{\lambda}_1$	vício$\left(\hat{\lambda}_1\right)$	mse$\left(\hat{\lambda}_1\right)$	$\hat{\lambda}_2$	vício$\left(\hat{\lambda}_2\right)$	mse$\left(\hat{\lambda}_2\right)$	$\hat{\lambda}_3$	vício$\left(\hat{\lambda}_3\right)$	mse$\left(\hat{\lambda}_3\right)$
$re = 100$									
0.1	0.1027	0.0027	0.0002	0.1076	0.0076	0.0006	0.0997	-0.0002	0.0001
0.2	0.2046	0.0046	0.0004	0.2050	0.0050	0.0011	0.2014	0.0014	0.0003
0.3	0.3063	0.0063	0.0007	0.3070	0.0070	0.0015	0.3033	0.0033	0.0006
0.4	0.4134	0.0134	0.0012	0.4009	0.0009	0.0030	0.4098	0.0098	0.0011
0.5	0.5054	0.0054	0.0008	0.4875	-0.0124	0.0029	0.5027	0.0027	0.0008
0.6	0.6011	0.0011	0.0010	0.5650	-0.0349	0.0035	0.5985	-0.0014	0.0010
0.7	0.7020	0.0020	0.0013	0.6641	-0.0358	0.0056	0.6963	-0.0006	0.0013
0.8	0.8060	0.0060	0.0016	0.7615	-0.0384	0.0054	0.8026	0.0026	0.0016
0.9	0.9103	0.0103	0.0022	0.8346	-0.0653	0.0089	0.9071	0.0071	0.0020
$re = 500$									
0.1	0.1042	0.0042	0.0002	0.1103	0.0103	0.0008	0.1011	0.0011	0.0002
0.2	0.2046	0.0046	0.0004	0.2046	0.0046	0.0013	0.2015	0.0015	0.0003
0.3	0.3062	0.0062	0.0006	0.3031	0.0031	0.0018	0.3033	0.0033	0.0006
0.4	0.4067	0.0067	0.0009	0.3957	-0.0042	0.0024	0.4037	0.0037	0.0009
0.5	0.5039	0.0039	0.0012	0.4876	-0.0123	0.0032	0.5008	0.0008	0.0011
0.6	0.6047	0.0047	0.0013	0.5730	-0.0269	0.0042	0.6016	0.0016	0.0012
0.7	0.7030	0.0030	0.0014	0.6626	-0.0373	0.0053	0.6998	-0.0001	0.0014
0.8	0.8063	0.0063	0.0018	0.7503	-0.0496	0.0071	0.8033	0.0033	0.0018
0.9	0.9048	0.0048	0.0019	0.8384	-0.0615	0.0085	0.9017	0.0014	0.0019

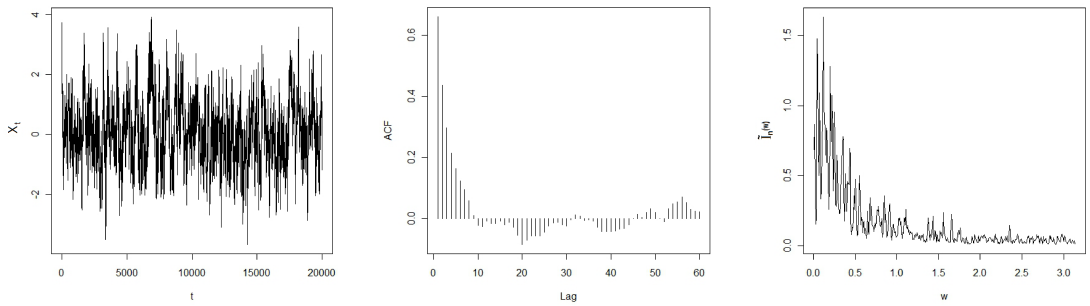
Considerando o espaçamento entre os dados igual a 0.1, esperávamos uma melhora na estimação do parâmetro λ. Podemos observar na Tabela 5.2 que houve melhoras na estimação de λ em todos os casos, com vício e mse menores que os observados na Tabela 5.1. Observe também uma melhora do estimador $\hat{\lambda}_2$, que obteve bons resultados para valores de λ de 0.1 até 0.4. O aumento do número de replicações não apresentou melhoras significativas na estimação do parâmetro.

Caso n = 20000.



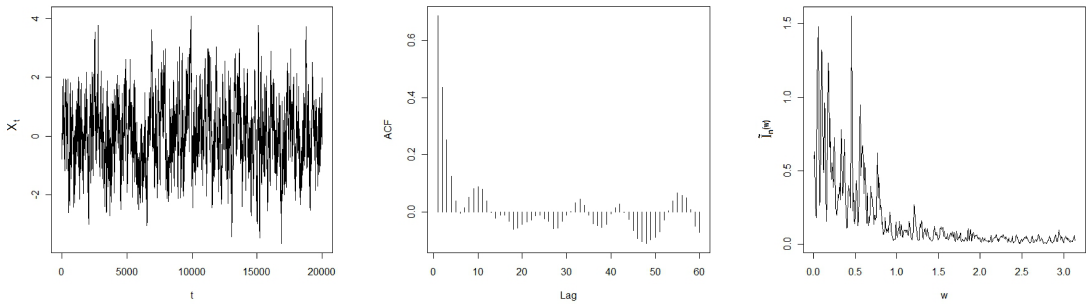
(a) (b) (c)

Figura 5.9: Geração do Processo (5.2) para $\lambda = 0.1$, $n = 20000$, $h = 0.05$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.10: Geração do Processo (5.2) para $\lambda = 0.4$, $n = 20000$, $h = 0.05$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.11: Geração do Processo (5.2) para $\lambda = 0.9$, $n = 20000$, $h = 0.05$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.

Tabela 5.3: *Estimação do Parâmetro* λ*, quando* $n = 20000$*,* $\sigma = 1$*,* $h = 0.05$*,* $m^* = 2000$ *e* $m = 1000$.

λ	$\hat{\lambda}_1$	vício$\left(\hat{\lambda}_1\right)$	mse$\left(\hat{\lambda}_1\right)$	$\hat{\lambda}_2$	vício$\left(\hat{\lambda}_2\right)$	mse$\left(\hat{\lambda}_2\right)$	$\hat{\lambda}_3$	vício$\left(\hat{\lambda}_3\right)$	mse$\left(\hat{\lambda}_3\right)$
$re = 100$									
0.1	0.1020	0.0020	0.0002	0.1088	0.0088	0.0007	0.0992	-0.0007	0.0001
0.2	0.2045	0.0045	0.0004	0.2056	0.0056	0.0014	0.2013	0.0013	0.0003
0.3	0.3058	0.0058	0.0005	0.3049	0.0049	0.0015	0.3033	0.0033	0.0005
0.4	0.4019	0.0019	0.0007	0.3967	-0.0032	0.0018	0.3988	-0.0011	0.0007
0.5	0.5051	0.0051	0.0008	0.4993	-0.0006	0.0036	0.5018	0.0018	0.0008
0.6	0.6046	0.0046	0.0010	0.5925	-0.0074	0.0034	0.6011	0.0011	0.0010
0.7	0.7071	0.0071	0.0011	0.6874	-0.0125	0.0036	0.7038	0.0038	0.0011
0.8	0.7997	-0.0002	0.0021	0.7793	-0.0206	0.0053	0.7973	-0.0026	0.0021
0.9	0.9092	0.0092	0.0018	0.8744	-0.0255	0.0071	0.9058	0.0058	0.0018
$re = 500$									
0.1	0.1039	0.0039	0.0002	0.1107	0.0107	0.0007	0.1008	0.0008	0.0002
0.2	0.2068	0.0068	0.0004	0.2093	0.0093	0.0013	0.2038	0.0038	0.0004
0.3	0.3049	0.0049	0.0006	0.3064	0.0065	0.0019	0.3017	0.0017	0.0006
0.4	0.4047	0.0047	0.0008	0.4041	0.0041	0.0026	0.4019	0.0019	0.0008
0.5	0.5056	0.0056	0.0011	0.5009	0.0009	0.0032	0.5024	0.0024	0.0010
0.6	0.6046	0.0046	0.0012	0.5910	-0.0089	0.0037	0.6016	0.0016	0.0012
0.7	0.7086	0.0086	0.0017	0.6891	-0.0108	0.0043	0.7055	0.0055	0.0016
0.8	0.8034	0.0034	0.0017	0.7774	-0.0225	0.0058	0.8003	0.0003	0.0017
0.9	0.9053	0.0053	0.0018	0.8765	-0.0234	0.0058	0.9018	0.0018	0.0018

Considerando o espaçamento entre os dados igual a 0.05, podemos observar na Tabela 5.3 que houve significativa melhora na estimação do parâmetro λ em todos os casos, com vício e mse menores que os observados nas Tabelas 5.1 e 5.2. Observe também que ocorre o melhor desempenho do estimador $\hat{\lambda}_2$, que apresentou bons resultados para valores de λ de 0.1 até 0.6. O aumento do número de replicações não apresentou melhoras significativas na estimação do parâmetro.

5.2 Estimação de λ na Solução da Equação de Langevin com Ruído α-estável

Nessa seção, apresentamos a simulação da solução da equação de Langevin com ruído α-estável, dada por

$$X_{t+h} = e^{-\lambda h} X_t + e^{-\lambda h} \sigma \int_0^h e^{\lambda s} \, dZ_s, \tag{5.6}$$

onde $\{Z_t\}_{t\in\mathbb{R}}$ é um processo α-estável, ou seja, $Z \sim S_\alpha(\sigma^*, \beta, \mu)$.

Também apresentamos os resultados da estimação de λ através dos estimadores (5.3) e (5.4) propostos por Valdivieso et al. (2009) e do estimador de máxima verossimilhança (5.5). Através das funções de autocorrelação e periodograma, analisamos a existência de longa dependência no processo dado pela expressão (5.6).

Para a simulação do processo (5.6) no intervalo $[0, T]$, consideramos a versão recursiva e discreta dada por

$$X_{t+h} = e^{-\lambda h} \lim_{m\to\infty} \left[X_t + \sigma \sum_{i=1}^{m} e^{\lambda s_i} \left(Z_{s_i} - Z_{s_{i-1}} \right) \right], \tag{5.7}$$

onde a sequência de números reais $(s_i)_{i=0}^m$ é uma partição, igualmente espaçada, do intervalo $[0, h]$. Consideramos aqui $s_i = \dfrac{1}{i}$, ou seja, a partição é formada por pontos equidistantes. Em todas as simulações aqui realizadas, $T = 1000$, e n é o tamanho amostral da série temporal simulada no intervalo $[0, T]$.

As simulações do processo (5.6) foram realizadas para $t \in [0, T]$. Consideramos dois valores distintos de h: 1 e 0.1. No caso $h = 1$, temos que $n = 1000$, onde n

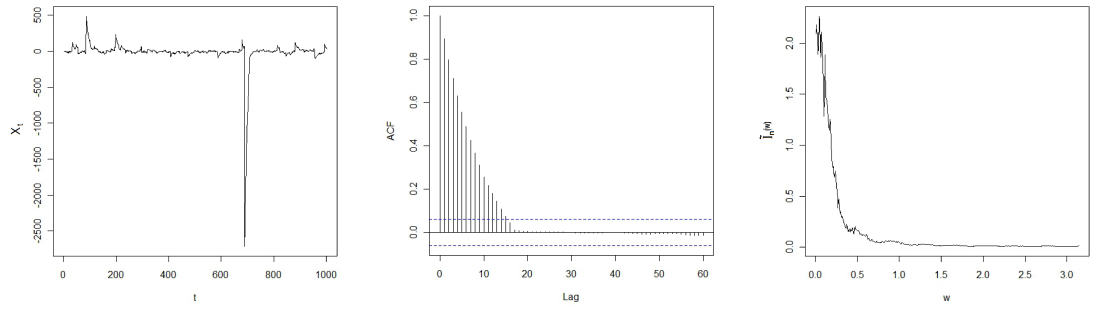
representa o tamanho amostral, e no caso $h = 0.1$ obtemos $n = 10000$.

O parâmetro λ varia entre 0.1 e 0.9, com incrementos de tamanho 0.1. Consideramos o parâmetro σ fixo e igual a 1. O processo é simulado com 100 e 500 replicações. Apresentamos os ruídos α-estáveis com parâmetros $\sigma^* = 1$, $\beta = 0$, $\mu = 0$ e o parâmetro α assumindo dois valores distintos, 1.1 e 1.5, ou seja, $Z_t \sim S_\alpha(1, 0, 0)$, $\alpha \in \{1.1, 1.5\}$.

Apresentamos os gráficos do processo simulado, da função de autocorrelação amostral e da função periodograma suavizado, onde a função de ponderação utilizada é a janela de Bartlett, para λ=0.1, 0.4 e 0.9. Enfim, apresentamos as tabelas com o valores estimados, vício e erro médio quadratico.

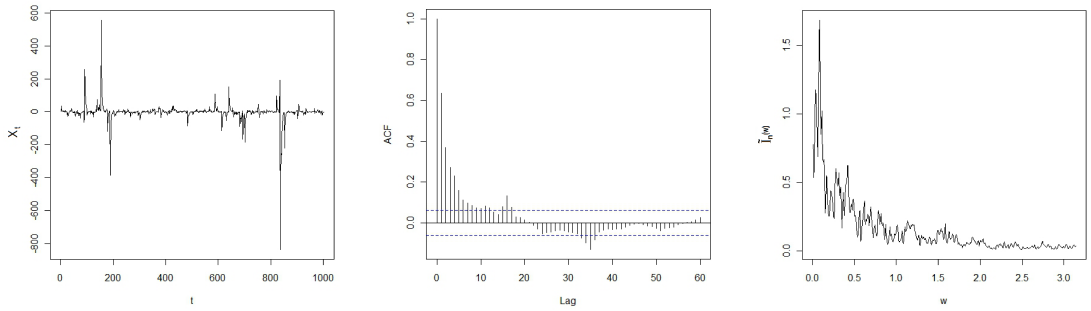
Como na seção anterior, consideramos o estimador $\hat{\lambda}_2$ com $m^* = 1000$.

Caso n = 1000, $\alpha = 1.1$.



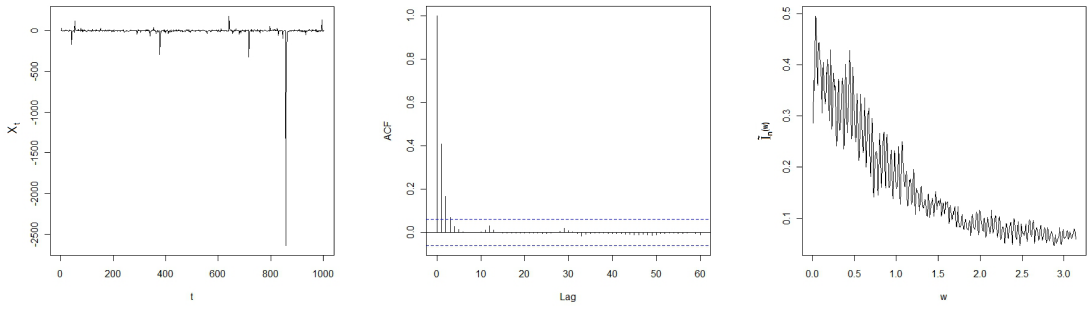
(a) (b) (c)

Figura 5.12: Geração do Processo (5.2) para $\alpha = 1.1$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.1$, $n = 1000$, $h = 1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.13: Geração do Processo (5.2) para $\alpha = 1.1$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.4$, $n = 1000$, $h = 1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



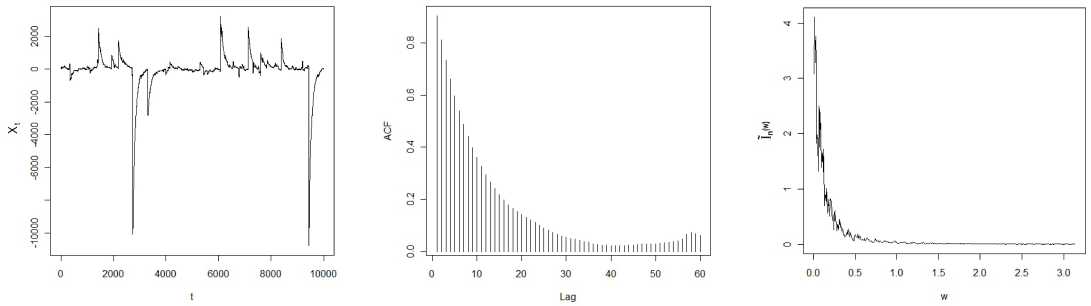
(a) (b) (c)

Figura 5.14: Geração do Processo (5.2) para $\alpha = 1.1$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.9$, $n = 1000$, $h = 1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.

Tabela 5.4: *Estimação do Parâmetro* λ*, quando* $n = 1000$*,* $\sigma = 1$*,* $h = 1$*,* $m^* = 1000$*,* $m = 1000$ *e* $Z \sim S_{1.1}(1, 0, 0)$*.*

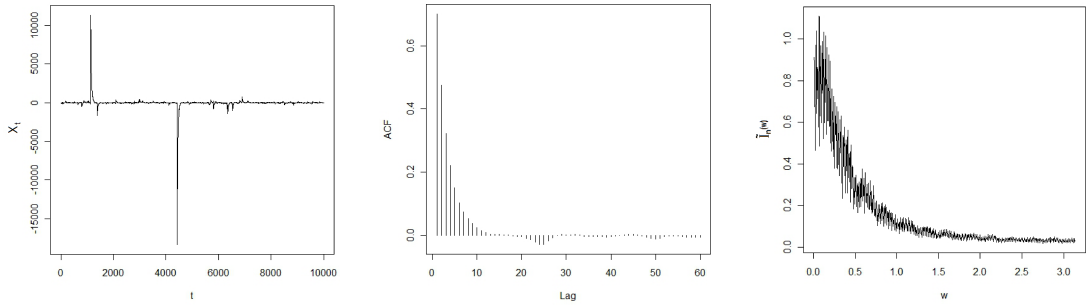
λ	$\hat{\lambda}_1$	vício$\left(\hat{\lambda}_1\right)$	mse$\left(\hat{\lambda}_1\right)$	$\hat{\lambda}_2$	vício$\left(\hat{\lambda}_2\right)$	mse$\left(\hat{\lambda}_2\right)$	$\hat{\lambda}_3$	vício$\left(\hat{\lambda}_3\right)$	mse$\left(\hat{\lambda}_3\right)$
$re = 100$									
0.1	0.1043	0.0043	0.0001	0.0978	-0.0021	0.0003	0.1017	0.0017	0.0001
0.2	0.2065	0.0065	0.0003	0.1740	-0.0259	0.0010	0.2037	0.0037	0.0003
0.3	0.3075	0.0075	0.0013	0.2424	-0.0575	0.0043	0.3050	0.0050	0.0013
0.4	0.4126	0.0125	0.0025	0.3089	-0.0910	0.0098	0.4048	0.0048	0.0011
0.5	0.5047	0.0047	0.0005	0.3558	-0.1441	0.0216	0.5015	0.0015	0.0004
0.6	0.6052	0.0052	0.0021	0.4138	-0.1861	0.0363	0.6019	0.0019	0.0020
0.7	0.7089	0.0089	0.0018	0.4698	-0.2301	0.0540	0.7054	0.0054	0.0017
0.8	0.8146	0.0146	0.0038	0.5216	-0.2783	0.0806	0.8104	0.0104	0.0038
0.9	0.9048	0.0048	0.0055	0.5604	-0.3395	0.1175	0.9012	0.0012	0.0055
$re = 500$									
0.1	0.1045	0.0045	0.0001	0.0991	-0.0008	0.0003	0.1013	0.0013	0.0001
0.2	0.2040	0.0040	0.0003	0.1744	-0.0255	0.0010	0.2008	0.0008	0.0002
0.3	0.3035	0.0035	0.0005	0.2404	-0.0595	0.0040	0.3009	0.0009	0.0005
0.4	0.4045	0.0045	0.0007	0.3029	-0.0970	0.0100	0.4017	0.0017	0.0006
0.5	0.5060	0.0060	0.0019	0.3587	-0.1412	0.0207	0.5018	0.0018	0.0013
0.6	0.6078	0.0078	0.0110	0.4159	-0.1840	0.0400	0.6039	0.0039	0.0104
0.7	0.7025	0.0025	0.0009	0.4605	-0.2394	0.0583	0.6988	-0.0011	0.0008
0.8	0.8115	0.0115	0.0311	0.5138	-0.2861	0.0873	0.7991	-0.0008	0.0017
0.9	0.9046	0.0046	0.0023	0.5598	-0.3401	0.1171	0.9004	0.0004	0.0021

Caso n = 10000, $\alpha = 1.1$.



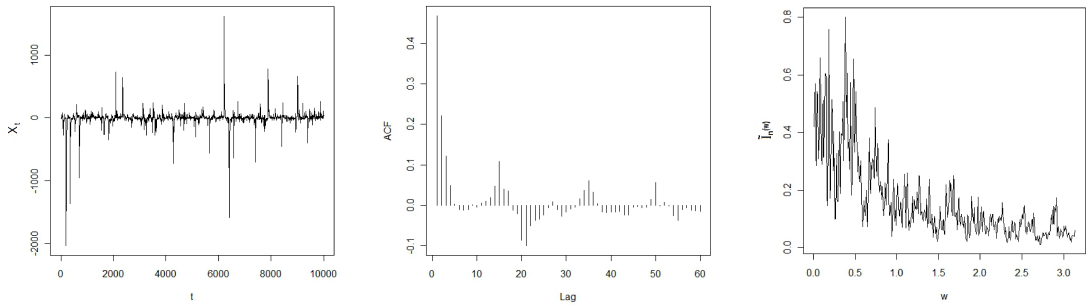
(a) (b) (c)

Figura 5.15: Geração do Processo (5.2) para $\alpha = 1.1$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.1$, $n = 10000$, $h = 0.1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.16: Geração do Processo (5.2) para $\alpha = 1.1$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.4$, $n = 10000$, $h = 0.1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



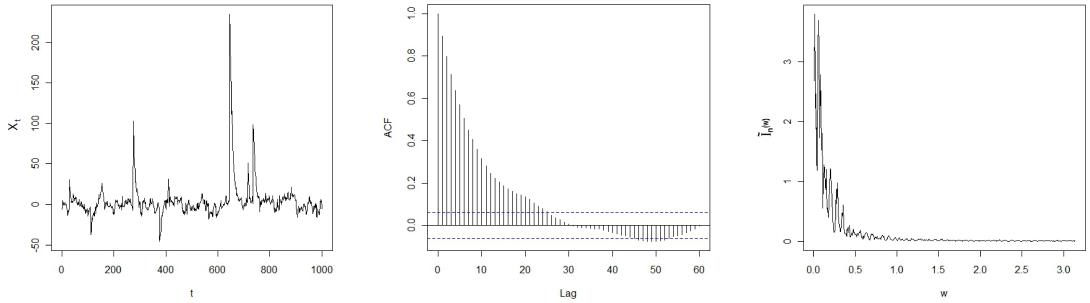
(a) (b) (c)

Figura 5.17: Geração do Processo (5.2) para $\alpha = 1.1$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.9$, $n = 10000$, $h = 0.1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.

Tabela 5.5: *Estimação do Parâmetro* λ*, quando* $n = 10000$*,* $\sigma = 1$*,* $h = 0.1$*,* $m^* = 2000$*,* $m = 1000$ *e* $Z \sim S_{1.1}(1, 0, 0)$.

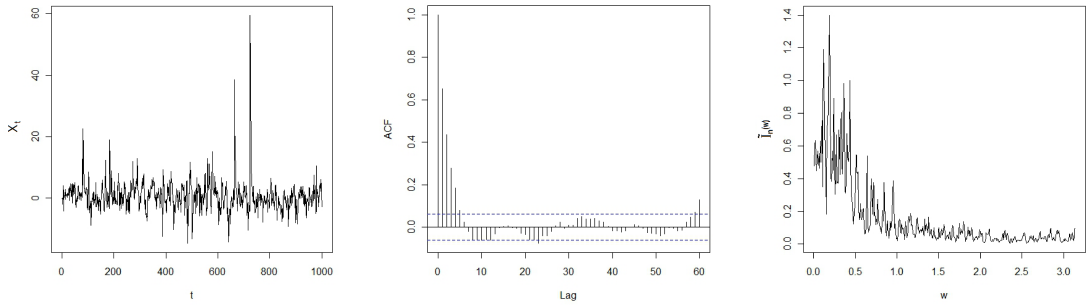
λ	$\hat{\lambda}_1$	vício$\left(\hat{\lambda}_1\right)$	mse$\left(\hat{\lambda}_1\right)$	$\hat{\lambda}_2$	vício$\left(\hat{\lambda}_2\right)$	mse$\left(\hat{\lambda}_2\right)$	$\hat{\lambda}_3$	vício$\left(\hat{\lambda}_3\right)$	mse$\left(\hat{\lambda}_3\right)$
$re = 100$									
0.1	0.1037	0.0037	0.0001	0.1089	0.0089	0.0004	0.1012	0.0012	0.0000
0.2	0.2037	0.0037	0.0004	0.2060	0.0060	0.0012	0.2017	0.0017	0.0004
0.3	0.3033	0.0033	0.0002	0.3001	0.0001	0.0010	0.3008	0.0008	0.0002
0.4	0.4039	0.0039	0.0009	0.3938	-0.0061	0.0021	0.4012	0.0012	0.0009
0.5	0.5048	0.0048	0.0002	0.4844	-0.0155	0.0009	0.5021	0.0021	0.0002
0.6	0.6133	0.0133	0.0052	0.5877	-0.0122	0.0161	0.6048	0.0048	0.0009
0.7	0.7025	0.0025	0.0006	0.6652	-0.0347	0.0043	0.7005	0.0005	0.0005
0.8	0.8017	0.0017	0.0007	0.7466	-0.0533	0.0045	0.7993	-0.0006	0.0005
0.9	0.9010	0.0010	0.0015	0.8409	-0.0590	0.0110	0.8989	-0.0010	0.0015
$re = 500$									
0.1	0.1065	0.0065	0.0026	0.1122	-0.0122	0.0046	0.1014	0.0014	0.0001
0.2	0.2073	0.0073	0.0031	0.2092	-0.0092	0.0050	0.2005	0.0005	0.0002
0.3	0.3043	0.0043	0.0018	0.2993	-0.0006	0.0043	0.2998	-0.0001	0.0003
0.4	0.4057	0.0057	0.0009	0.3952	-0.0047	0.0019	0.4031	0.0031	0.0009
0.5	0.5066	0.0066	0.0037	0.4862	-0.0137	0.0027	0.5014	0.0014	0.0006
0.6	0.6060	0.0060	0.0011	0.5754	-0.0245	0.0028	0.6032	0.0032	0.0010
0.7	0.7045	0.0045	0.0014	0.6586	-0.0417	0.0053	0.7007	0.0007	0.0010
0.8	0.8047	0.0047	0.0009	0.7490	-0.0509	0.0049	0.8020	0.0020	0.0008
0.9	0.9015	0.0015	0.0022	0.8275	-0.0724	0.0076	0.8989	-0.0010	0.0021

Caso n = 1000, $\alpha = 1.5$.



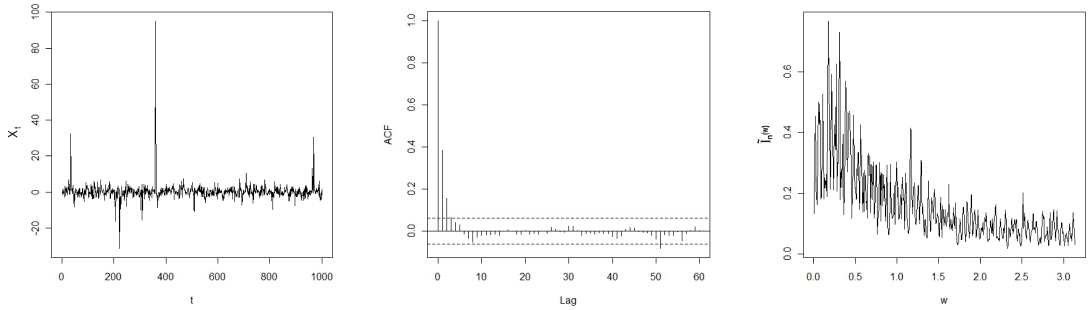
(a) (b) (c)

Figura 5.18: Geração do Processo (5.2) para $\alpha = 1.5$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.1$, $n = 1000$, $h = 1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.19: Geração do Processo (5.2) para $\alpha = 1.5$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.4$, $n = 1000$, $h = 1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



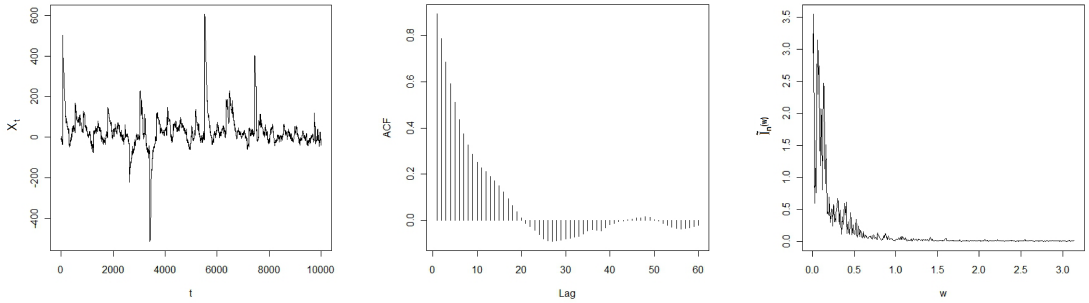
(a) (b) (c)

Figura 5.20: Geração do Processo (5.2) para $\alpha = 1.5$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.9$, $n = 1000$, $h = 1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.

Tabela 5.6: *Estimação do Parâmetro* λ*, quando* $n = 1000$*,* $\sigma = 1$*,* $h = 1$*,* $m^* = 1000$*,* $m = 1000$ *e* $Z \sim S_{1.5}(1, 0, 0)$*.*

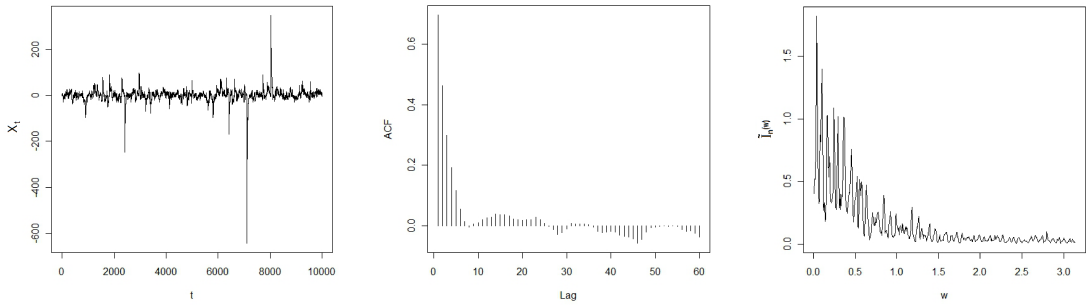
λ	$\hat{\lambda}_1$	vício$\left(\hat{\lambda}_1\right)$	mse$\left(\hat{\lambda}_1\right)$	$\hat{\lambda}_2$	vício$\left(\hat{\lambda}_2\right)$	mse$\left(\hat{\lambda}_2\right)$	$\hat{\lambda}_3$	vício$\left(\hat{\lambda}_3\right)$	mse$\left(\hat{\lambda}_3\right)$
$re = 100$									
0.1	0.1057	0.0057	0.0002	0.1006	0.0006	0.0005	0.1020	0.0020	0.0001
0.2	0.2051	0.0051	0.0004	0.1765	-0.0234	0.0011	0.2020	0.0020	0.0003
0.3	0.3028	0.0038	0.0006	0.2407	-0.0592	0.0044	0.3011	0.0011	0.0005
0.4	0.4120	0.0120	0.0017	0.3094	-0.0905	0.0090	0.4054	0.0054	0.0008
0.5	0.5084	0.0084	0.0015	0.3667	-0.1332	0.0195	0.5037	0.0037	0.0012
0.6	0.6143	0.0143	0.0026	0.4165	-0.1834	0.0357	0.6108	0.0108	0.0024
0.7	0.7006	0.0006	0.0020	0.4634	-0.2365	0.0570	0.6972	-0.0027	0.0020
0.8	0.8050	0.0050	0.0050	0.5116	-0.2883	0.0853	0.8015	0.0015	0.0050
0.9	0.9013	0.0013	0.0023	0.5557	-0.3442	0.1200	0.8983	-0.0016	0.0023
$re = 500$									
0.1	0.1050	0.0050	0.0002	0.0993	-0.0006	0.0004	0.1015	0.0015	0.0001
0.2	0.2056	0.0056	0.0007	0.1760	-0.0239	0.0012	0.2022	0.0022	0.0004
0.3	0.3066	0.0066	0.0005	0.2437	-0.0562	0.0038	0.3032	0.0032	0.0005
0.4	0.4063	0.0063	0.0009	0.3025	-0.0974	0.0104	0.4031	0.0031	0.0008
0.5	0.5042	0.0042	0.0013	0.3569	-0.1430	0.0212	0.5003	0.0003	0.0010
0.6	0.6056	0.0056	0.0021	0.4141	-0.1858	0.0357	0.6017	0.0017	0.0020
0.7	0.7078	0.0078	0.0021	0.4657	-0.2342	0.0562	0.7042	0.0042	0.0020
0.8	0.8088	0.0088	0.0039	0.5157	-0.2842	0.0824	0.8047	0.0047	0.0038
0.9	0.9074	0.0074	0.0036	0.5615	-0.3384	0.1161	0.9037	0.0037	0.0034

Caso n = 10000, $\alpha = 1.5$.



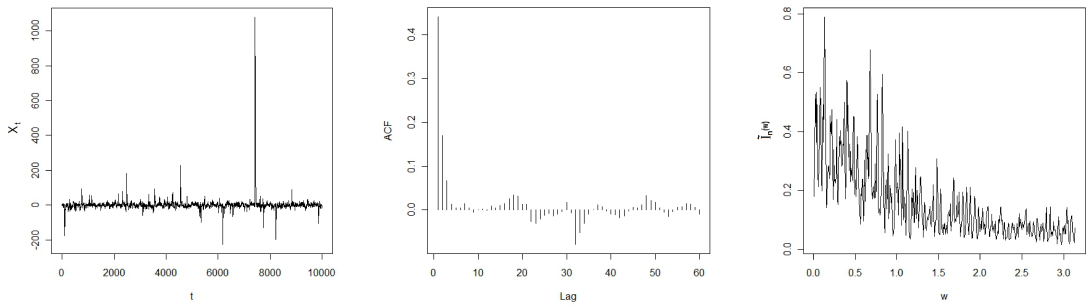
(a) (b) (c)

Figura 5.21: Geração do Processo (5.2) para $\alpha = 1.5$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.1$, $n = 10000$, $h = 0.1$, $\sigma = 1$ e $m = 1000$.: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.22: Geração do Processo (5.2) para $\alpha = 1.5$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.4$, $n = 10000$, $h = 0.1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.



(a) (b) (c)

Figura 5.23: Geração do Processo (5.2) para $\alpha = 1.5$, $\beta = 0$, $\mu = 0$, $\sigma^* = 1$, $\lambda = 0.9$, $n = 10000$, $h = 0.1$, $\sigma = 1$ e $m = 1000$: (a) Série Temporal Simulada; (b) Função de Autocorrelação Amostral; (c) Função Periodograma Suavizado.

Tabela 5.7: *Estimação do Parâmetro* λ*, quando* $n = 10000$*,* $\sigma = 1$*,* $h = 0.1$*,* $m^* = 2000$*,* $m = 1000$ *e* $Z \sim S_{1.5}(1, 0, 0)$*.*

λ	$\hat{\lambda}_1$	vício$\left(\hat{\lambda}_1\right)$	mse$\left(\hat{\lambda}_1\right)$	$\hat{\lambda}_2$	vício$\left(\hat{\lambda}_2\right)$	mse$\left(\hat{\lambda}_2\right)$	$\hat{\lambda}_3$	vício$\left(\hat{\lambda}_3\right)$	mse$\left(\hat{\lambda}_3\right)$
$re = 100$									
0.1	0.1021	0.0021	0.0001	0.1039	0.0039	0.0004	0.0994	-0.0005	0.0001
0.2	0.2064	0.0064	0.0003	0.2095	0.0095	0.0009	0.2036	0.0036	0.0003
0.3	0.3079	0.0079	0.0011	0.3028	0.0028	0.0016	0.3027	0.0027	0.0005
0.4	0.4032	0.0032	0.0006	0.3988	-0.0011	0.0019	0.4003	0.0003	0.0006
0.5	0.5100	0.0100	0.0046	0.4974	-0.0025	0.0052	0.5011	0.0011	0.0017
0.6	0.6078	0.0078	0.0012	0.5736	-0.0263	0.0032	0.6050	0.0050	0.0011
0.7	0.7065	0.0065	0.0013	0.6604	-0.0395	0.0056	0.7033	0.0033	0.0012
0.8	0.8048	0.0048	0.0022	0.7365	-0.0634	0.0067	0.8022	0.0022	0.0022
0.9	0.9098	0.0098	0.0016	0.8373	-0.0626	0.0067	0.9072	0.0072	0.0016
$re = 500$									
0.1	0.1043	0.0043	0.0001	0.1088	0.0088	0.0005	0.1013	0.0013	0.0001
0.2	0.2060	0.0060	0.0004	0.2065	0.0065	0.0010	0.2032	0.0032	0.0004
0.3	0.3048	0.0048	0.0008	0.3013	0.0013	0.0021	0.3010	0.0010	0.0005
0.4	0.4050	0.0050	0.0007	0.3942	-0.0057	0.0018	0.4017	0.0017	0.0006
0.5	0.5046	0.0046	0.0010	0.4842	-0.0157	0.0026	0.5014	0.0014	0.0009
0.6	0.6022	0.0022	0.0010	0.5697	-0.0302	0.0033	0.5997	-0.0002	0.0010
0.7	0.7052	0.0052	0.0013	0.6644	-0.0355	0.0045	0.7027	0.0027	0.0013
0.8	0.8040	0.0040	0.0013	0.7526	-0.0473	0.0056	0.8009	0.0009	0.0011
0.9	0.9045	0.0045	0.0016	0.8354	-0.0645	0.0086	0.9016	0.0016	0.0015

Podemos verificar, em todos os casos analisados, a ausência de longa dependência. Como no caso em que o ruído é o processo movimento Browniano, quanto menor o valor de λ mais lento é o decaimento da função de autocorrelação. Aplicando os mesmos estimadores utilizados na Seção 5.1, obtivemos bons resultados para $\hat{\lambda}_1$ e $\hat{\lambda}_3$, tanto para $h = 1$ quanto para $h = 0.1$. O estimador $\hat{\lambda}_2$ apresentou resultados ruins para $h = 1$, obtendo uma significativa melhora no caso $h = 0.1$. Como podemos verificar nas Tabelas 5.4 a 5.7, o valor de α não influenciou nos resultados. O aumento do número de replicações não representou significativa melhora na estimação.

5.3 Estimação de α na Solução da Equação de Langevin com Ruído α-estável

Seja $\{X_t\}_{t\in[0,T]}$ o processo solução da equação de Langevin com ruído α-estável apresentado em (5.6). Apresentamos a estimação do parâmetro α nestes processos. Para realizar a estimação, utilizamos o método de dois passos apresentado na Seção 4.2. No primeiro passo, estimamos λ através do estimador de máxima verossimilhança (4.22). No segundo passo, utilizamos o método de estimação proposto por Press (1972), considerando $t_1 = 0.2$ e $t_2 = 0.8$.

Nas simulações, consideramos o processo $\{Z_t\}_{t\in[0,T]}$ com distribuição $S_\alpha(1, \beta, 0)$, onde $\alpha \in \{1, 1.1, 1.5, 1.75, 1.9\}$, $\beta \in \{0, 0.5\}$. Utilizamos a forma discreta (5.7), com $m = 1000$ e $h \in \{0.1, 1\}$.

Foram realizadas 100 replicações do processo e o estimador é obtido através da média

$$\hat{\alpha} \;=\; \frac{1}{100}\sum_{j=1}^{100}\hat{\alpha}_j, \tag{5.8}$$

onde $\hat{\alpha}_j$ é o estimador proposto por Press (1972) e apresentado na Seção 4.2, aplicado na j-ésima replicação do processo.

Tabela 5.8: *Estimação do Parâmetro α para Diferentes Valores do λ, quando $n = 1000$, $\sigma = 1$, $h = 1$, $m = 1000$ e $Z \sim S_\alpha(1, 0, 0)$.*

α	λ	$\hat{\lambda}$	$\hat{\alpha}$	vício($\hat{\alpha}$)	mse($\hat{\alpha}$)
1	0.1	0.1037	1.0006	0.0006	0.0036
1	0.4	0.4031	0.9986	-0.0014	0.0065
1	0.9	0.9008	0.9265	-0.0735	0.0230
1.1	0.1	0.1039	1.1011	0.0011	0.0043
1.1	0.4	0.4077	1.0846	-0.0154	0.0050
1.1	0.9	0.9047	1.0519	-0.0481	0.0148
1.5	0.1	0.1042	1.4995	-0.0005	0.0037
1.5	0.4	0.4080	1.4947	-0.0053	0.0043
1.5	0.9	0.9011	1.4508	-0.0492	0.0120
1.75	0.1	0.0998	1.7490	-0.0010	0.0029
1.75	0.4	0.4055	1.7569	0.0069	0.0034
1.75	0.9	0.9124	1.6563	-0.0937	0.0208
1.9	0.1	0.1039	1.9039	0.0039	0.0014
1.9	0.4	0.4042	1.9014	0.0014	0.0031
1.9	0.9	0.9084	1.7716	-0.1284	0.0298

Tabela 5.9: *Estimação do Parâmetro α para Diferentes Valores do λ, quando $n = 10000$, $\sigma = 1$, $h = 0.1$, $m = 1000$ e $Z \sim S_\alpha(1, 0, 0)$.*

α	λ	$\hat{\lambda}$	$\hat{\alpha}$	vício($\hat{\alpha}$)	mse($\hat{\alpha}$)
1	0.1	0.1029	0.9984	-0.0016	0.0009
1	0.4	0.4035	1.0012	0.0012	0.0004
1	0.9	0.9020	0.9994	-0.0006	0.0005
1.1	0.1	0.1032	1.0966	-0.0034	0.0005
1.1	0.4	0.4033	1.1049	0.0049	0.0004
1.1	0.9	0.9009	1.1011	0.0011	0.0003
1.5	0.1	0.1014	1.4978	-0.0022	0.0003
1.5	0.4	0.4025	1.4998	-0.0002	0.0003
1.5	0.9	0.9094	1.4993	-0.0007	0.0003
1.75	0.1	0.1045	1.7496	-0.0004	0.0003
1.75	0.4	0.4021	1.7476	-0.0024	0.0003
1.75	0.9	0.9018	1.7521	0.0021	0.0002
1.9	0.1	0.1008	1.8991	-0.0009	0.0002
1.9	0.4	0.4019	1.9010	0.0010	0.0002
1.9	0.9	0.9011	1.8980	-0.0020	0.0001

Tabela 5.10: *Estimação do Parâmetro α para Diferentes Valores do λ, quando $n = 1000$, $\sigma = 1$, $h = 1$, $m = 1000$ e $Z \sim S_\alpha(1, 0.5, 0)$.*

α	λ	$\hat{\lambda}$	$\hat{\alpha}$	vício($\hat{\alpha}$)	mse($\hat{\alpha}$)
1	0.1	0.1032	0.9832	-0.0168	0.0047
1	0.4	0.4014	0.9841	-0.0159	0.0049
1	0.9	0.9084	0.9761	-0.0239	0.0161
1.1	0.1	0.1016	1.1106	0.0106	0.0032
1.1	0.4	0.4019	1.0859	-0.0141	0.0051
1.1	0.9	0.9135	1.0635	-0.0365	0.0169
1.5	0.1	0.1055	1.4928	-0.0072	0.0040
1.5	0.4	0.4051	1.4910	-0.0090	0.0045
1.5	0.9	0.9048	1.4736	-0.0264	0.0163
1.75	0.1	0.1039	1.7495	-0.0005	0.0024
1.75	0.4	0.4010	1.7450	-0.0050	0.0050
1.75	0.9	0.8926	1.6714	-0.0786	0.0230
1.9	0.1	0.1059	1.8997	-0.0003	0.0014
1.9	0.4	0.4049	1.9045	0.0045	0.0030
1.9	0.9	0.9077	1.7522	-0.1478	0.0340

Tabela 5.11: *Estimação do Parâmetro α para Diferentes Valores do λ, quando $n = 10000$, $\sigma = 1$, $h = 0.1$, $m = 1000$ e $Z \sim S_\alpha(1, 0.5, 0)$.*

α	λ	$\hat{\lambda}$	$\hat{\alpha}$	vício($\hat{\alpha}$)	mse($\hat{\alpha}$)
1	0.1	0.1023	0.9883	-0.0117	0.0006
1	0.4	0.4014	0.9833	-0.0167	0.0007
1	0.9	0.8998	0.9831	-0.0169	0.0006
1.1	0.1	0.1047	1.1007	0.0007	0.0006
1.1	0.4	0.4002	1.0988	-0.0012	0.0003
1.1	0.9	0.8987	1.0998	-0.0002	0.0004
1.5	0.1	0.1055	1.4999	-0.0001	0.0003
1.5	0.4	0.3992	1.4991	-0.0009	0.0003
1.5	0.9	0.8955	1.5020	0.0020	0.0003
1.75	0.1	0.1019	1.7505	0.0005	0.0003
1.75	0.4	0.4097	1.7477	-0.0023	0.0002
1.75	0.9	0.9018	1.7521	0.0021	0.0002
1.9	0.1	0.1022	1.8997	-0.0003	0.0001
1.9	0.4	0.4073	1.8983	-0.0017	0.0002
1.9	0.9	0.9014	1.8982	-0.0018	0.0001

Na Tabela 5.8 podemos constatar que a variação de α não causou influência significativa no erro quadrático médio (mse) de $\hat{\alpha}$, que, para cada valor de $\lambda \in \{0.1, 0.4, 0.9\}$, mostrou-se razoavelmente estável. No entanto, é possível verificar que conforme aumentamos o valor de λ, maior o erro quadrático médio de $\hat{\alpha}$. Conforme verificamos na Seção 5.2, conforme aumentamos o valor de λ, maior o mse de $\hat{\lambda}$. É razoável que essa variação do mse de $\hat{\lambda}$ influencie na estimação de α no caso $h = 1$.

No entanto, na Tabela 5.9, onde apresentamos as simulações para processos gerados com $h = 0.1$, o mse do estimador $\hat{\alpha}$ mantém-se praticamente estável para todos os valores de λ e α, o que mostra que para h pequeno a estimação de α

apresenta ótimos resultados através do algoritmo de dois passos.

Nas Tabelas 5.10 e 5.11 apresentamos as estimações para processos gerados com $\beta = 0.5$. Essa modificação no valor de β não causou alterações significativas nas estimações de α.

Capítulo 6

Aplicação

Neste capítulo apresentamos a aplicação da teoria do processo advindo da solução da equação de Langevin clássica nos dados obtidos de perfurações realizadas do gelo da Groelândia pela Greeland Icecore Project (GRIP). Essas perfurações tem se tornado cada vez mais importantes para a compreensão do clima do passado.

A Groelândia está coberta por uma capa de gelo de mais de 2km de altura. Essa capa tem uma estrutura em forma de camadas, e cada camada contém informações sobre as condições atmosféricas existentes na época. Mais interessante ainda, é que podemos obter o ano em que cada camada se formou.

Nas amostras de gelo obtidas na perfuração, estudam-se as bolhas de ar envoltas entre os cristais de gelo, que são amostras da atmosfera daquele tempo, e as concentrações relativas de isótopo de hidrogênio e de oxigênio, que fornecem dados sobre a temperatura da época.

Nesse trabalho analisamos o isótopo ^{18}O contido nos cristais de gelo. Mais precisamente, analisamos o índice $\delta^{18}O$, que é a diferença entre o ^{18}O obtido das perfurações no gelo e o ^{18}O contido em amostras atuais da água do oceano. Ditlevsen et al. (2002) mostram que esses dados podem ser modelados através das equações diferenciais estocásticas, mais precisamente através do processo de

Ornstein-Uhlenbeck, aplicado em determinados sub-intervalos onde podemos supor estacionariedade.

Os dados aqui apresentados remetem à 60000 anos atrás, em intervalos de 20 anos, ou seja, temos 3000 dados. Larios e Sorensen (2010) propõem a análise dividida em três subintervalos: [-60000,-30000], [-30000,-10000] e [-10000,0]. Aqui, dividiremos os dados originais nestes três subintervalos.

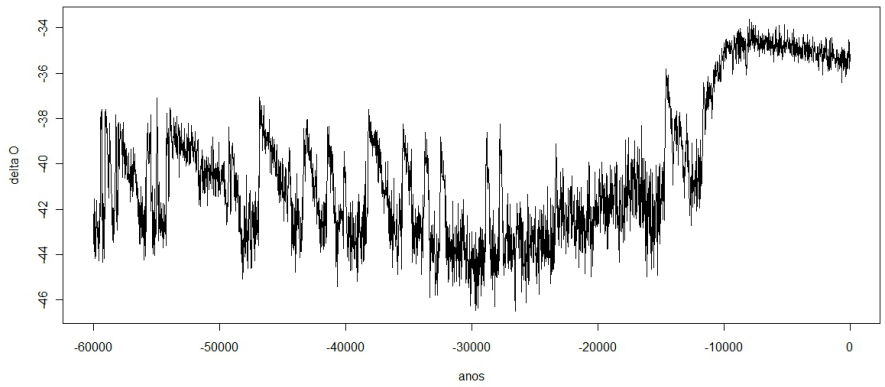
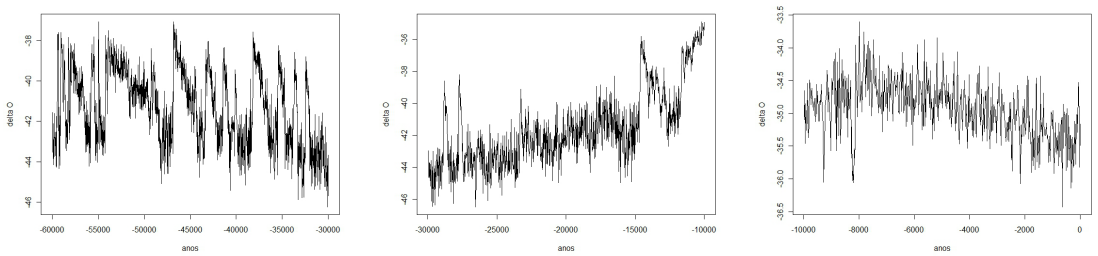


Figura 6.1: $\delta^{18}O$ medido nas amostras extraídas das escavações no gelo da Groelândia.



(a) (b) (c)

Figura 6.2: (a) [-60000,-30000]; (b) [-30000,-10000]; (c) [-10000,0].

Consideramos então as séries temporais $\{X_t\}_{t=1}^{1500}$, $\{Y_t\}_{t=1}^{1000}$ e $\{Z_t\}_{t=1}^{500}$ representando, respectivamente, o processo nos três subintervalos considerados. A Figura

6.2 representa estas três séries temporais, respectivamente, nos gráficos (a), (b) e (c).

Nas estimações consideramos $h = 1$ e utilizamos os estimadores (5.3) e (5.4) para λ e o estimador (4.18) para σ. Para estimar a média, utilizamos o estimador $\overline{X}$ apresentado em (2.9). Os resultados obtidos são apresentados na Tabela 6.1.

Tabela 6.1: *Estimação de λ e σ^2 dos dados $\delta^{18}O$.*

	[-60000,-30000]	[-30000,-10000]	[-10000,0]
$\hat{\lambda}_1$	0.1738	0.1487	0.8455
$\hat{\lambda}_3$	0.1728	0.1411	0.8450
$\hat{\sigma}$	1.0951	1.2907	0.5766
Média	-41.3121	-41.5727	-34.9306

Pela Tabela 6.1 observamos que $\hat{\lambda}_1$ e $\hat{\lambda}_3$ apresentam valores bem próximos entre si para o parâmetro λ em todos os intervalos analisados.

Capítulo 7

Conclusão

Apresentamos, no início deste trabalho, as principais propriedades da integral de Itô, bem como os processos movimento Browniano, de Lévy, α-estáveis e movimento Browniano fracionário, fundamentais para o entendimento e caracterização dos processos advindos da solução da equação de Langevin clássica e generalizada.

Apresentamos na Seção 3.1 a equação de Langevin clássica e sua solução, obtida através das propriedades da integral e do lema de Itô.

Na Seção 3.3 apresentamos brevemente a equação de Langevin clássica com ruído movimento Browniano fracionário e o processo advindo da sua solução, que tem a peculiaridade de sua função de autocorrelação possuir um lento decaimento.

Na Seção 3.4 apresentamos a equação de Langevin generalizada, o teorema de existência e unicidade de Kannan e o teorema de representação de Kannan, que provam a existência e unicidade da solução para a equação de Langevin generalizada e apresenta critérios para a forma de tal solução.

No Capítulo 4 apresentamos os principais estimadores para o parâmetro λ vinculado à função memória da equação de Langevin. Na Seção 4.2 propomos um algoritmo para estimação do parâmetro α no processo advindo da solução da equação de Langevin generalizada com ruído α-estável. O algoritmo é baseado no estimador

proposto por Press (1972), e apresentou resultados animadores nas simulações, com erro quadrático médio relativamente pequeno no caso $h = 1$, e com ótimos resultados quando consideramos $h = 0.1$.

Nas simulações apresentadas no Capítulo 5 obtivemos ótimos resultados para os estimadores (5.3) e (5.5), que apresentaram erro médio quadrático muito pequeno, com significativa melhora nas estimações quando consideramos um espaçamento menor entre os dados. Podemos observar nas tabelas que quanto menor o valor de λ, melhor é o resultado obtido, tanto quando o ruído é um movimento Browniano como quando o ruído tem distribuição α-estável.

Para estudos futuros pretendemos estender a análise para processos advindos da equação de Langevin generalizada com ruído movimento Browniano fracionário e realizar inferências sobre o parâmetro de Hurst destes processos. É interessante também estudar os processos advindos da equação de Langevin generalizada com outros tipos de função memória que não o apresentado em (3.31).

Referências Bibliográficas

[1] Avramidis, A.N., L'Ecuyer, P. e Tremblay, P.A. (2003). "Efficient Simulation of Gamma and Variance-Gamma Processes". *Proceedings of the 2003 Winter Simulation Conference*, 319-326.

[2] Barndorff-Nielsen, O.E. e Shephard, N. (2000). "Modelling by Lévy Process for Financial Econometrics". Em Barndorff-Nielsen, O.E., Mikosch, T. e Resnick, S. (eds.): *Lévy Process: Theory and Application*. New York: Birkhäuser.

[3] Beran, J. (1994). *Statistics for Long-Memory Processes*. New York: Chapman & Hall.

[4] Brockwell, P.J. e Davis, R.A. (1991). *Time Series: Theory and Methods*. New York: Springer.

[5] Billingsley, P. (1995). *Probability and Measure*. New York: Wiley.

[6] Cheridito, P., Kawaguchi, H., Maejima, M. (2003). "Fractional Ornstein-Uhlenbeck Process". *Eletronic Journal of Probability*, **8**: 1-14.

[7] Coffey, W.T., Kalmykov, Yu P., Waldron, J.T. (2004). *The Langevin Equation, With Applications to Stochastic Problems in Physics, Chemistry and Electrical Engineering*. Singapura: World Scientific.

[8] Conway, J.B. (1985). *A Course in Functional Analysis*. New York: Springer Verlag.

[9] Dellacherie, C. e Meyer, P.A. (1982). *Probabilities and Potential B: Theory of Martingales*. Amsterdam: North-Holland.

[10] Ditlevsen, P.D., Ditlevsen, S. e Andersen, K.K. (2002). "The fast climate fluctuations during the stadial and interstadial climate states". *Annals of Glaciology*, **35**: 457-462.

[11] Dudley, R.M. (2004). *Real Analysis and Probability*. Cambridge: Cambridge.

[12] Dudley, R.M. e Norvaisa, R. (1998). *Differentiability of Six Operators on Nonsmooth Functions and p-Variation (Lecture Notes in Mathematics)*. New York: Springer.

[13] Gnedenko, B.V. e Kolmogorov, A.N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Cambridge: Addison-Wesley.

[14] Gradshteyn, I.S. e Ryzhik, I.M. (2000). *Table of Integrals, Series, and Products*. San Diego: Academic Press.

[15] Guerra, J. e Nualart, D. (2008). "Stochastic Differential Equations Driven by Fractional Brownian Motion and a Standard Brownian Motion". *Journal of Stochastic Analysis and Applications*, **26**: 1053-1075.

[16] Hurst, H.E. (1951). "Long-term storage capacity in reservoirs". *Transactions of American Society of Civil Engineers*: **116**: 400-410.

[17] James, B.R. (2008). *Probabilidade: um Curso em Nível Intermediário*. Rio de Janeiro: IMPA.

[18] Jongloed, G., van de Meulen, F. e van der Vaart, A. W. (2005). "Nonparametric Inference for Lévy-Driven Ornstein-Uhlenbeck Processes". *Bernoulli*, **11**: 759-791.

[19] Kannan, D. (1977). "On the Generalized Langevin Equation". *Journal of Mathematical and Physical Sciences*, **11**: 1-24.

[20] Kannan, D. e Arucha-Reid, A.T. (1974). "Randon Integral Equation Formulation of a Generalized Langevin Equation". *Journal of Statistical Physics*, **5**: 209-233.

[21] Karlin, S. e Taylor, H.M. (1975). *A First Course in Stochastic Process*. London: Academic Press.

[22] Klebaner, F.C. (1999). *Introduction to Stochastic Calculus with Applications*. Sassari: World Scientific.

[23] Kloeden, P.E. (1992). *Numerical Solution of Stochastic Differential Equations*. New York: Springer.

[24] Kolmogorov, A.N. (1940). "Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum". *C.R. Dokl. Acad. Sci. USSR N.S.*, **26**: 115-118.

[25] Kolmogorov, A.N. (1941) Local Structure of Turbulence in Fluid for Very Large Reynolds Numbers. *Transl. in Turbulence. S.K. Friedlander e L. Topper*. New York, 151-155.

[26] Kolmogorov, A.N. e Fomin, S.V. (1970). *Introdutory Real Analysis*. New York: Dover.

[27] Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Process*. London: Springer.

[28] Larios, F.B. e Sorensen, M. (2010). "Maximum Likelihood Estimation for Integrated Diffusion Process". Em Chiarella, C. e Novikov, A. (eds.): *Contemporary Quantitative Finance: Essays in Honour of Eckhard Platen*. Heidelberg: Springer.

[29] Leland, W.E., Taqqu, M.S., Willinger, W. e Wilson. D.V. (1994). "On the Self-Similar Nature of Ethernet Traffic". *IEEE/ACM Transactions on Networking*, **2**: 1-15.

[30] Lima, E.L. (2005). *Curso de Análise Vol. 1 e 2*. Rio de Janeiro: IMPA.

[31] Lima, E.L. (2005). *Espaços Métricos*. Rio de Janeiro: IMPA.

[32] Lopes, S.R.C., Olbermann, B.P., Reisen, V.A. (2004). "A Comparison of Estimation Methods in Non-stationary ARFIMA Processes". *Journal of Statistical Computation and Simulation*, **74**(5): 339-347.

[33] Lopes, S.R.C., Nunes, M.A. (2006). "Long Memory Analysis in DNA Sequences". *Physica A: Statistical Mechanics and its Applications*, **361**(2): 569-588.

[34] Lopes, S.R.C. (2008). "Long-Range Dependence in Mean and Volatility: Models, Estimation and Forecasting". *Progress in Probability*, **60**: 497-525.

[35] Mandelbrot, B. (1963). "The Variation of Certain Speculative Prices". *The Journal of Business of the University of Chicago*, **36**: 394-419.

[36] Mandelbrot, B. e Van Ness, J.W. (1968). "Fractional Brownian Motion, Fractional Noices and Applications". *SISM Review*, **10**: 422-437.

[37] Mandelbrot, B. e Wallis, J. (1969). "Computer Experiments With Fractional Gaussian Noises: Part 1, Averages and Variances". *Water Resources Research*, **5**(1): 228-241.

[38] Medino, A.V. (2005). "Índice de Difusão Anômala, Processos de Lévy e Equação de Langevin Generalizada". *Tese de Doutorado*. Programa de Pós-Graduação em Matemática do Instituto de Ciências Exatas da UnB, Brasília.

[39] Mikosch, T. (1998). *Elementary Stochastic Calculus With Finance in View (Advanced Series on Statistical Science e Applied Probability, Vol 6)*. Sassari: World Scientific.

[40] Neter, J. e Wasserman, W. (1974). *Applied Linear Statistical Models*. Illinois: Irwin.

[41] Nualart, D. (2006). "Stochastic calculus with respect to fractional Brownian motion". *Ann Fac. Sci. Toulouse*, **15**: 63-78.

[42] Olbermann, B.P, Lopes, S.R.C., Reisen, V. (2006). "Invariance of the first difference in ARFIMA models". *Journal of Computational Statistics*, **3**: 445-461.

[43] Önalan, Ö. (2010). "Fractional Ornstein-Uhlenbeck Process Driven by Stable Lévy Motion in Finance". *International Research Journal of Finance and Economics*, **42**: 129-138.

[44] Palma, W. (2007). *Long-Memory Time Series: Theory and Methods*. New Jersey: Wiley.

[45] Press, J. (1972). "Estimation in Univariate and Multivariate Stable Distributions". *Journal of the American Statistical Association*, **67**(340): 842-846.

[46] Protter, P.E. (1992). *Stochastic Integration and Differential Equations* New York: Springer.

[47] Samorodnitsky, G. e Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes*. London: Chapman & Hall.

[48] Samorodnitsky, G. (2006). "Long Range Dependence". *Foundations and Trends in Stochastic System*, **1**(3): 163-257.

[49] Stuck, B.W. e Kleiner, B. (1974). "A statistical analysis of telephone noise". *The Bell System Technical Journal*, **53**: 1263-1320.

[50] Taylor, S.J. (1972). "Exact asymptotic estimates of Brownian path variation". *Duke Math Journal*, **39**: 219-241.

[51] Valdivieso, L., Schoutens, W. e Tuerlinckk, F. (2009). "Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type". *Statistical Inference for Stochastic Processes,* **12**: 1-19.

[52] Valdivieso, L. (2005). "Parameter Estimator for an IG-OU Stochastic Volatility Model". *Proceedings of 3rd Acturial and Financial Mathematic Day*, 95-104.

[53] Wei, W.W.S. (1994). *Time Series Analysis*. New York: Addison-Wesley.

[54] Zähle, M. (1998). "Integration with Respect to Fractal Function and Stochastic Calculus I". *Probability Theory and Related Fields,* **111**: 333-374.

Anexo A

A.1 Espaços de Hilbert e Espaços de Banach

No Capítulo 3, Seção 3.4, apresentamos alguns resultados das equações de Langevin no contexto da análise dos espaços de Hilbert e de Banach. Assim, apresentamos alguns resultados e definições preliminares para nos auxiliar no estudo das equações de Langevin generalizada.

Definição A.1. Seja E um espaço vetorial real. Um *produto interno* em E é uma função $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$, tal que, para todo $x, w, y \in E$ e $c \in \mathbb{R}$,

i) $\langle x + w, y \rangle = \langle x, y \rangle + \langle w, y \rangle$;

ii) $\langle cx, y \rangle = c \langle x, y \rangle$;

iii) $\langle x, y \rangle = \langle y, x \rangle$;

iv) $x \neq 0 \Rightarrow \langle x, x \rangle > 0$.

O exemplo mais importante é o produto interno canônico do espaço euclidiano $\mathbb{R}^n$, dado por

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n,$$

onde $x = (x_1, \cdots, x_n)' \in \mathbb{R}^n$ e $y = (y_1, \cdots, y_n)' \in \mathbb{R}^n$.

Teorema A.1 (Desigualdade de Cauchy-Schwarz)**.** *Sejam x e y dois elementos pertencentes à um espaço vetorial munido de um produto interno. Então,*

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle , \tag{A.1}$$

onde a igualdade vale se e somente se x e y são linearmente dependentes.

Demonstração: Ver Lima (2005), página 6. □

Definição A.2 (Norma)**.** Seja L um espaço linear. Uma função $p : L \to \mathbb{R}$ finita é uma *norma* em L, denotada por $p(x) = ||x||$, se

i) $p(x) \geq 0$ para todo $x \in L$, onde $p(x) = 0$ se e somente se $x = 0$;

ii) $p(cx) = |c|p(x)$, para todo $x \in L$ e todo $c \in \mathbb{R}$;

iii) $p(x + y) \leq p(x) + p(y)$, para todo $x,\ y \in L$.

Definição A.3 (Espaço Linear Normado)**.** Um espaço linear L, equipado com uma norma $p(x) = ||x||$, é chamado de *espaço linear normado*.

Exemplo A.1. O conjunto dos números reais é um espaço linear normado, com a norma $||x|| = |x|$, para todo $x \in \mathbb{R}$.

Exemplo A.2. O espaço real n-dimensional $\mathbb{R}^n$ é um espaço linear normado, considerando como norma $||\mathbf{x}|| = \sqrt{\sum_{k=1}^{n} x_k^2}$, para todo $\mathbf{x} = (x_1,\ x_2, \cdots,\ x_n)' \in \mathbb{R}^n$.

Exemplo A.3. Seja $[a, b]$ um intervalo fechado da reta real. O espaço $\mathcal{C}([a, b], \mathbb{R}) = \{f : [a, b] \to \mathbb{R} | f$ é contínua$\}$, equipado com a norma

$$||f|| = \max_{a \leq x \leq b} |f(x)| \tag{A.2}$$

é um espaço linear normado.

Observação A.1. Todo espaço linear L, munido de um produto interno, é um espaço vetorial normado se considerarmos a norma

$$||x|| = \sqrt{\langle x, x \rangle},\ \text{para todo } x \in L. \tag{A.3}$$

A recíproca é verdadeira se e somente se vale a Lei do Paralelogramo (ver Lima, 2005, página 6).

Definição A.4 (Espaço Métrico)**.** Um *espaço métrico* consiste em um par (M, d), onde M é um conjunto qualquer e $d : M \times M \to \mathbb{R}$ é uma função, chamada métrica, com as seguintes propriedades.

i) $d(x, y) \geq 0$, para todo $(x,\ y) \in M \times M$, onde $d(x, y) = 0$ se e somente se $x = y$;

ii) $d(x, y) = d(y, x)$;

iii) $d(x, y) \leq d(x, z) + d(z, y)$, para todo $z \in M$.

Apresentamos a seguir uma ferramenta muito útil para a averiguação do item iii) da Definição A.4.

Proposição A.1. *Sejam* $\boldsymbol{x}$, $\boldsymbol{y}$ *elementos do espaço n-dimensional* M *munido de uma métrica* d*, tal que* $\boldsymbol{x} = (x_1, \cdots, x_n)'$ *e* $\boldsymbol{y} = (y_1, \cdots, y_n)'$*. Então, vale a* desigualdade de Cauchy-Schwarz

$$\left(\sum_{k=1}^{n} x_k y_k\right)^2 \leq \sum_{k=1}^{n} x_k^2 \sum_{k=1}^{n} y_k^2. \tag{A.4}$$

Demonstração: Ver Conway (1985). □

Boa parte dos conjuntos comumente conhecidos são espaços métricos.

Exemplo A.4. O par $(\mathbb{R}, d)$, com $d(x, y) = |x - y|$, é um espaço métrico.

Exemplo A.5. Seja M o conjunto de todas n-uplas ordenadas

$$\mathbf{x} = (x_1, x_2, \cdots, x_n)',$$

onde $x_1, \cdots, x_n \in \mathbb{R}$, munido da métrica

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2},\ \ \mathbf{x}, \mathbf{y} \in M.$$

Então, (M, d) forma uma espaço métrico chamado *espaço Euclidiano n-dimensional*, denotado por $\mathbb{R}^n$.

De fato, o item i) e ii) da Definição A.4 estão trivialmente satisfeitos. Para provar a desigualdade triangular, sejam $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ tais que

$$\mathbf{x} = (x_1, \cdots, x_n)', \ \mathbf{y} = (y_1, \cdots, y_n)', \ \mathbf{z} = (z_1, \cdots, z_n)'.$$

Tomamos então

$$a_k = x_k - y_k, \ b_k = y_k - z_k,$$

para $k = 1, \cdots, n$. Da desigualdade de Cauchy-Schwarz temos que

$$\sum_{k=1}^{n} a_k b_k \leq \sqrt{\sum_{k=1}^{n} a_k^2} \sqrt{\sum_{k=1}^{n} b_k^2}.$$

Logo,

$$\begin{aligned} \sum_{k=1}^{n} (a_k + b_k)^2 &= \sum_{k=1}^{n} a_k^2 + 2 \sum_{k=1}^{n} a_k b_k + \sum_{k=1}^{n} b_k^2 \\ &\leq \sum_{k=1}^{n} a_k^2 + 2 \sqrt{\sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2} + \sum_{k=1}^{n} b_k^2 \\ &= \left(\sqrt{\sum_{k=1}^{n} a_k^2} + \sqrt{\sum_{k=1}^{n} b_k^2} \right)^2. \end{aligned}$$

Portanto,

$$\sqrt{\sum_{k=1}^{n} (a_k + b_k)^2} \leq \sqrt{\sum_{k=1}^{n} a_k^2} + \sqrt{\sum_{k=1}^{n} b_k^2}.$$

Substituindo $a_k = x_k - y_k$ e $b_k = y_k - z_k$, temos a desigualdade

$$\sqrt{\sum_{k=1}^{n} (x_k - z_k)^2} \leq \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2} + \sqrt{\sum_{k=1}^{n} (y_k - z_k)^2},$$

provando a desigualdade triangular.

Exemplo A.6. Para $p \geq 1$, seja

$$l_p = \left\{ (x_n)_{n \geq 1} \, | \, \sum_{k=1}^{\infty} x_k^p < \infty \right\} \tag{A.5}$$

o conjunto de todas as sequências infinitas de números reais satisfazendo a condição de convergência

$$\sum_{k=1}^{\infty} x_k^p < \infty$$

para alguma constante real $p \geq 1$. Tomando a métrica

$$d(\mathbf{x}, \mathbf{y}) = \left(\sum_{k=1}^{\infty} |x_k - y_k|^p \right)^{\frac{1}{p}} \tag{A.6}$$

temos que (l_p, d) forma um espaço métrico, denotado também por $\mathbb{R}_p^n$. Os itens i) e ii) da Definição A.4 são trivialmente satisfeitos. A demonstração do item iii) é um pouco mais trabalhosa, e pode ser vista em Kolmogorov e Fomin (1970).

Observação A.2. Todo espaço linear normado L é um espaço métrico se considerarmos como métrica

$$d(x, y) = ||x - y||, \text{ para todo } x,\ y \in L, \tag{A.7}$$

onde $||\cdot||$ é uma norma qualquer.

A recíproca não é verdadeira. Como exemplo, basta considerar um conjunto aberto de $\mathbb{R}^n$, munido da métrica apresentada no Exemplo A.5.

Definição A.5 (Sequência de Cauchy)**.** Uma sequência $(x_n)_{n\in\mathbb{N}}$ de pontos de um espaço métrico M é chamada de *sequência de Cauchy* se, para todo $\varepsilon > 0$, existe um inteiro N_ε tal que $d(x_m, x_n) < \varepsilon$, para todo $m,\ n > N_\varepsilon$.

Observação A.3. Toda sequência convergente é uma sequência de Cauchy.

De fato, uma sequência $(a_n)_{n\in\mathbb{N}}$ definida em um espaço métrico M é convergente se existe um elemento $a \in M$ tal que para todo $\varepsilon > 0$, existe um $N_\varepsilon \in \mathbb{N}$, de tal forma que para todo $n > N_\varepsilon$, $d(a_n, a) < \varepsilon$.

Assim, considerando $\varepsilon_1 = \frac{\varepsilon}{2}$, existe um $N_{\varepsilon_1} \in \mathbb{N}$ tal que para todo $n > N_{\varepsilon_1}$, $d(a_n, a) < \frac{\varepsilon}{2}$. Pela desigualdade triangular, se $m > n$, então

$$d(a_m, a_n) \leq d(a_m, a) + d(a, a_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Logo, $(a_n)_{n\in\mathbb{N}}$ é uma sequência de Cauchy.

Definição A.6 (Espaço Métrico Completo)**.** Um espaço métrico M é dito *completo* se toda sequência de Cauchy em M converge para um elemento de M.

Exemplo A.7. Um exemplo clássico de espaço métrico completo é a reta real.

Da completude da reta real, temos que o espaço $\mathbb{R}^n$ também é um espaço métrico completo. De fato, seja

$$\mathbf{x}^{(p)} = \left(x_1^{(p)}, x_2^{(p)}, \cdots, x_n^{(p)}\right)', \ p \in \mathbb{N},$$

uma sequência de Cauchy de elementos de $\mathbb{R}^n$. Então, para todo $\varepsilon > 0$, existe um N_ε tal que

$$d\left(\mathbf{x}^{(r)}, \mathbf{x}^{(s)}\right) = \sqrt{\sum_{k=1}^{n}\left(x_k^{(r)} - x_k^{(s)}\right)^2} < \varepsilon,$$

para todo $r, s > N_\varepsilon$. Logo,

$$\sum_{k=1}^{n}\left(x_k^{(r)} - x_k^{(s)}\right)^2 < \varepsilon^2,$$

para todo $r, s > N_\varepsilon$, o que implica que

$$|x_k^{(r)} - x_k^{(s)}| < \varepsilon,$$

para todo $r, s > N_\varepsilon$, para cada $k = 1, 2, \cdots, n$.

Assim, para cada $k = 1, 2, \cdots, n$, temos que $\left(x_k^{(p)}\right)_{p\in\mathbb{N}}$ forma uma sequência de Cauchy em $\mathbb{R}$. Então, se para cada $k = 1, 2, \cdots, n$,

$$x_k = \lim_{p\to\infty} x_k^{(p)},$$

temos

$$\lim_{p\to\infty} \mathbf{x}^{(p)} = \mathbf{x} = (x_1, x_2, \cdots, x_n)' \in \mathbb{R}^n.$$

Definição A.7 (Espaço de Banach)**.** Um espaço linear normado completo, relativo à métrica (A.7), é chamado *espaço de Banach*.

Um importante exemplo de espaço de Banach é o espaço L^p.

Exemplo A.8 (Espaço $L^p(\Omega, \mathcal{F}, \mathbb{P})$)**.** Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade. Para $1 \leq p < \infty$, L^p é o espaço formado por todas as funções $f : \Omega \to \mathbb{R}$ mensuráveis tal que $|f|^p$ é integrável, e para $f \in L^p$, definimos a norma

$$||f||_p = \left[\int |f|^p d\mathbb{P}\right]^{\frac{1}{p}}. \tag{A.8}$$

No caso em que $p = \infty$, $f \in L^\infty$ se existe um c real tal que

$$\mathbb{P}\{\omega \in \Omega : |f(\omega)| \leq c\} = 0. \tag{A.9}$$

Nessa caso, f é dita ser *essencialmente limitada*, e sua norma é dada por

$$||f||_\infty = inf\{c : \mathbb{P}[\omega \in \Omega : |f(\omega)| \leq c] = 0\}. \tag{A.10}$$

Definição A.8 (Espaço de Hilbert)**.** Um *espaço de Hilbert* é um espaço vetorial H, munido de um produto interno e completo em relação à norma definida por esse produto interno.

Exemplo A.9. Os exemplos mais conhecidos de espaços de Hilbert são o espaço Euclidiano $\mathbb{R}^n$, com o produto interno $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$, onde $\mathbf{x} = (x_1, \cdots, x_n)' \in \mathbb{R}^n$ e $\mathbf{y} = (y_1, \cdots, y_n)' \in \mathbb{R}^n$, e o espaço l_2, como definido no Exemplo A.6.