



Trabalho de Conclusão de Curso

**Processos Auto-regressivos Bidimensionais de Primeira  
Ordem com Inovações Gaussianas e Não-Gaussianas**

Raphael Cabral de Mello

3 de dezembro de 2020

Raphael Cabral de Mello

**Processos Auto-regressivos Bidimensionais de Primeira Ordem com  
Inovações Gaussianas e Não-Gaussianas**

Trabalho de Conclusão apresentado à comissão de Graduação do Departamento de Estatística da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientador(a): Profa. Dr. Sílvia Regina Costa  
Lopes

Porto Alegre  
Dezembro de 2020

Raphael Cabral de Mello

**Processos Auto-regressivos Bidimensionais de Primeira Ordem com  
Inovações Gaussianas e Não-Gaussianas**

Este Trabalho foi julgado adequado para obtenção dos créditos da disciplina Trabalho de Conclusão de Curso em Estatística e aprovado em sua forma final pela Orientador(a) e pela Banca Examinadora.

Orientador(a): \_\_\_\_\_  
Profa. Dr. Sílvia Regina Costa Lopes, IME - UFRGS  
Doutor(a) pela University of Maryland, College Park,  
USA

Banca Examinadora:

Profa. Dra. Márcia Helena Barbian,  
Doutor pela Universidade Federal de Minas Gerais – Belo Horizonte, MG

Prof. Dr. Marcio Valk,  
Doutor pela Universidade Estadual de Campinas – Campinas, SP

Porto Alegre  
Dezembro de 2020

# Agradecimentos

Agradeço, primeiramente, a todos que participaram, direta ou indiretamente, da construção deste trabalho.

Agradeço à minha mãe, Moira, que sempre esteve do meu lado, me ajudando e apoiando em minhas escolhas.

Agradeço à Vanessa, que tornou este período mais agradável com sua disponibilidade e carinho.

Agradeço, em especial, à professora Sílvia e ao LPCM pela oportunidade de aprendizado e apoio durante a iniciação científica e durante o período de escrita deste trabalho.

Agradeço a Josiane Stein, a Roberto Molina de Souza e Maicon Josué Karling pela constante ajuda e retirada de dúvidas.

E, não menos importante, a meus colegas e amigos, que tornaram esta etapa mais agradável.

# Resumo

Este trabalho analisa o desempenho de métodos de estimação clássicos e Bayesianos para processos auto-regressivos bivariados de primeira ordem. As inovações do processo, denotado VAR(1), são provenientes de variáveis aleatórias, independentes e identicamente distribuídas, com distribuições Gaussianas,  $t$ -Student e  $\alpha$ -Estável. É conhecida a não existência das funções de autocovariância e autocovariância cruzada quando as inovações são provenientes de distribuições  $\alpha$ -Estáveis. Por esta razão, neste trabalho, é também abordada as aplicações das funções codiferença e codiferença cruzada, tanto nas simulações como na aplicação de dados reais. Na fundamentação teórica, apresentam-se, tanto os métodos de estimação, como a identificação de dados  $\alpha$ -Estáveis. Para isso, os testes de adequação de modelos são também analisados. Aplicamos os métodos apresentados a dados reais de ações da *Coca-Cola Company* e *PepsiCo Inc.*

**Palavras-Chave:** Distribuição  $\alpha$ -Estável, Séries Temporais, Inferência Clássica e Bayesiana, Testes de Adequação de Modelo.

# Abstract

This work analyzes the performance of classical and Bayesian methods for estimating bivariate first-order auto-regressive processes, denoted by VAR(1). The innovation processes are formed by independent and identically distributed random variables, with Gaussian,  $t$ -Student, and  $\alpha$ -Stable distributions. It is known the non-existence of the autocovariance and cross-autocovariance functions when the innovation comes from an  $\alpha$ -Stable distribution. For this reason, this work also employed the codifference and cross-codifference functions, both in simulations and real data analysis. In the theoretical foundation, we present both methods of estimating, as well as the identification of Stable data. For this, goodness-of-fit tests are also analyzed. The methods shown in this work are applied to a real *Coca-Cola Company* and *PepsiCo Inc.* stock dataset.

**Keywords:**  $\alpha$ -stable Distribution, Time Series, Classical and Bayesian Inference, Goodness-of-fit Tests.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>13</b>
<b>2</b>	<b>Conceitos e Resultados Preliminares</b>	<b>15</b>
2.1	Processos Estocásticos e Séries Temporais	15
2.2	Distribuição $\alpha$ -Estável	21
2.2.1	Univariada	21
2.2.2	Multivariada	25
2.3	Processos Auto-regressivos de Ordem $p$	26
2.3.1	AR( $p$ ) Univariado	26
2.3.2	AR( $p$ ) Bivariado	28
2.4	Medidas de Correlação	29
2.4.1	Coefficiente de Correlação de Pearson	29
2.4.2	Função Codiferença	29
2.4.3	Função Codiferença Cruzada	31
2.5	Testes de Adequação	31
2.5.1	Testes de Função de Distribuição Empírica	32
2.5.2	Teste por Quantis de McCulloch	33
2.6	Inferência Bayesiana	33
<b>3</b>	<b>Estimação</b>	<b>35</b>
3.1	Estimação de Processos VAR	35
3.2	Estimação de Parâmetros	36
3.2.1	Distribuição Gaussiana	36
3.2.2	Distribuição $t$ -Student	36
3.2.3	Distribuição $\alpha$ -Estável	37
<b>4</b>	<b>Estimações Clássica e Bayesiana de um VAR(1) Bidimensional</b>	<b>38</b>
4.1	Estimação Clássica	38
4.1.1	Inovações Gaussianas	38
4.1.2	Inovações $t$ -Student	40
4.1.3	Inovações $\alpha$ -Estáveis	40

<b>4.2</b>	<b>Estimação Bayesiana</b>	<b>41</b>
4.2.1	Inovações Gaussianas	42
4.2.2	Inovações $t$ -Student	42
4.2.3	Inovações $\alpha$ -Estáveis	42
<b>5</b>	<b>Aplicação a Dados Reais</b>	<b>50</b>
5.1	Apresentação dos Dados	50
5.2	Análise e Estimação	51
<b>6</b>	<b>Conclusão</b>	<b>56</b>

# Lista de Figuras

Figura 2.1:	(a) Função densidade de uma $\alpha$ -Estável para os valores de $\beta \in \{-1, 0, 1\}$ com $\alpha = 1.1$ , $\sigma = 1$ e $\delta = 0$ ; (b) Função densidade de uma $\alpha$ -Estável para os valores de $\sigma \in \{1, 3, 5\}$ com $\alpha = 1.5$ , $\beta = 0$ e $\delta = -5$ ; (c) Função densidade de uma $\alpha$ -Estável para os valores de $\delta \in \{-3, 0, 3\}$ com $\alpha = 0.5$ , $\sigma = 1$ e $\beta = 0$ ; (d) Função densidade de uma $\alpha$ -Estável para os valores de $\alpha \in \{0.8, 1.2, 1.6\}$ com $\sigma = 2$ , $\beta = 0$ e $\delta = 3$ . . . . .	23
Figura 2.2:	(a) Série temporal $\{X_t\}_{t=1}^{300}$ referente ao processo AR(1) dado na expressão (2.16); (b) Função de autocorrelação amostral; (c) Função de autocorrelação parcial amostral. . . . .	27
Figura 4.1:	Exemplo de séries geradas de um VAR(1) bidimensional definido em (2.17) com inovações Gaussianas quando $T = 10,000$ , $\mu = (0, 0)$ , $\sigma^2$ é dado pela expressão (4.1), $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ . . . . .	39
Figura 4.2:	Funções de Autocovariância (FAC) e de Autocovariância Parcial (FACP) e seus intervalos a 95% de confiança, em azul, para duas séries Gaussianas quando $\mu = (0, 0)$ , $\sigma^2$ é dado pela expressão (4.1), $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ : (a) FAC; (b) FACP. . . . .	44
Figura 4.3:	Exemplo de séries geradas de um VAR(1) bidimensional definido em (2.17) com inovações $t$ -Student, onde $T = 10.000$ , $gl = (4, 4)$ , $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ . . . . .	45
Figura 4.4:	Funções de Autocovariância (FAC) e de Autocovariância Parcial (FACP) e seus intervalos a 95% de confiança, em azul, para duas séries $t$ -Student, quando $T = 10,000$ , $gl = (4, 4)$ , $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ : (a) FAC; (b) FACP. . . . .	46
Figura 4.5:	Exemplo de séries geradas de um VAR(1) bidimensional definido em (2.17) com inovações $\alpha$ -Estáveis, onde $T = 10,000$ , $\alpha = (1.3, 1.3)$ , $\beta = (0, 0)$ , $\delta = (0, 0)$ , $\sigma = (0.3, 0.2)$ , $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ . . . . .	47

Figura 4.6:	Funções de Codiferença e seus intervalos a 95% de confiança, em azul, para duas séries $\alpha$ -Estáveis quando $\alpha = (1.3, 1.3)$ , $\beta = (0, 0)$ , $\delta = (0, 0)$ , $\sigma = (0.3, 0.2)$ , $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ : (a) Codiferença para a Componente 1; (b) Codiferença para a Componente 2. . . . .	48
Figura 4.7:	Codiferença Cruzada para $k$ entre -10 e 10 e seu intervalo a 95% de confiança, em azul, de um conjunto de séries geradas de um VAR(1) bidimensional definido em (2.17), com inovações $\alpha$ -Estáveis, quando $T = 10,000$ , $\alpha = (1.3, 1.3)$ , $\beta = (0, 0)$ , $\delta = (0, 0)$ , $\sigma = (0.3, 0.2)$ , $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ . . . . .	49
Figura 5.1:	Valores diários das ações das duas empresas analisadas no período de 01 de janeiro de 2015 a 29 de outubro de 2019, com $n = 1,215$ observações: (a) Coca-Cola Company; (b) PepsiCo Inc. . . . .	50
Figura 5.2:	Séries temporais resultantes de uma primeira diferença: (a) Coca-Cola Company; (b) PepsiCo Inc. . . . .	51
Figura 5.3:	Q-Q plots Gaussianos para as séries temporais $Y_t$ : (a) Coca-Cola Company; (b) PepsiCo Inc. . . . .	52
Figura 5.4:	Q-Q plots $t$ -Student para as séries temporais $Y_t$ : (a) Coca-Cola Company; (b) PepsiCo Inc. . . . .	52
Figura 5.5:	Q-Q plots $\alpha$ -Estáveis para as séries temporais $Y_t$ : (a) Coca-Cola Company; (b) PepsiCo Inc. . . . .	52
Figura 5.6:	Codiferença Cruzada para $k$ entre -10 e 10, das séries provenientes dos valores diferenciados das ações das empresas Coca-Cola Company e PepsiCo Inc. As linhas azuis representam o intervalo a 95% de confiança. . . . .	54

# Lista de Tabelas

Tabela 4.1:	Resultado das estimações em $re = 1,000$ simulações de Monte Carlo de processos VAR(1) bidimensional com inovações Gaussianas quando $T = 10,000$ , $\mu = (0, 0)$ , $\sigma^2$ é dado pela expressão (4.1), $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ . . . . .	39
Tabela 4.2:	Resultado das estimações em $re = 1,000$ simulações de Monte Carlo de processos VAR(1) bidimensional com inovações $t$ -Student, quando $T = 10,000$ , $gl = (4, 4)$ , $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ . . . . .	40
Tabela 4.3:	Resultado das estimações clássicas em $re = 1,000$ simulações de Monte Carlo de processos VAR(1) bidimensional com inovações $\alpha$ -Estáveis, quando $T = 10,000$ , $\alpha = (1.3, 1.3)$ , $\beta = (0, 0)$ , $\delta = (0, 0)$ , $\sigma = (0.3, 0.2)$ , $\phi_1 = (0.2, 0.1)$ e $\phi_2 = (0.1, 0.2)$ . . . . .	41
Tabela 4.4:	Resultado das estimações da função de codiferença cruzada em processos VAR(1) bidimensional com inovações $\alpha$ -Estáveis, quando $T = 10,000$ , $\alpha = (1.3, 1.3)$ , $\beta = (0, 0)$ , $\delta = (0, 0)$ , $\sigma = (0.3, 0.2)$ , $\phi_1 = (0.2, 0.1)$ , $\phi_2 = (0.1, 0.2)$ e $k \in \{-2, -1, 0, 1, 2\}$ . . . . .	41
Tabela 4.5:	Resultado das estimações Bayesianas em processos VAR(1) bidimensional com inovações Gaussianas, quando $T = 10,000$ , $\mu = (0, 0)$ , $\sigma^2 = (1, 1)$ , $\phi_1 = (0.2, 0.1)$ , $\phi_2 = (0.1, 0.2)$ e $re = 500$ . . . . .	42
Tabela 4.6:	Resultado das estimações Bayesianas em processos VAR(1) bidimensional com inovações $t$ -Student, quando $T = 10,000$ , $gl = (4, 4)$ , $\phi_1 = (0.2, 0.1)$ , $\phi_2 = (0.1, 0.2)$ e $re = 500$ . . . . .	43
Tabela 4.7:	Resultado das estimações Bayesianas em processos VAR(1) bidimensional com inovações $\alpha$ -Estáveis, quando $T = 10,000$ , $\alpha = (1.3, 1.3)$ , $\beta = (0, 0)$ , $\delta = (0, 0)$ , $\sigma = (0.3, 0.2)$ , $\phi_1 = (0.2, 0.1)$ , $\phi_2 = (0.1, 0.2)$ e $re = 500$ . . . . .	43
Tabela 5.1:	Resultado dos testes KS, AD, MKS e MC de adequação de modelos para normalidade, para os dados diferenciados das ações das empresas Coca-Cola Company e PepsiCo Inc. . . . .	53
Tabela 5.2:	Resultado dos testes KS, AD, MKS e MC de adequação de modelos para $\alpha$ -Estáveis, para os dados das ações, diferenciadas, das empresas Coca-Cola Company e PepsiCo Inc. . . . .	53

Tabela 5.3: Resultados das estimações dos modelos VAR( $p$ ),  $p = 1, 2, 3$  para os dados diários diferenciados das ações da Coca-Cola Company e da PepsiCo Inc, com seus respectivos AIC e BIC. O coeficiente  $\phi_{j,k}$  corresponde ao componente  $j = (pep, ko)$  no lag  $k$ . . . . . 55

# 1

## Introdução

Este trabalho tem como objetivo o estudo dos métodos de estimação clássicos e Bayesianos de processos auto-regressivos bivariados estacionários quando as inovações tem origem em variáveis aleatórias, independentes e identicamente distribuídas, com distribuições Gaussianas,  $t$ -Student e  $\alpha$ -Estável.

A decisão pelo uso das distribuições Gaussianas e não-Gaussianas, foi fortemente motivada pela presença de caudas pesadas em séries temporais financeiras. É conhecido que o mercado financeiro apresenta comportamento com mudanças bruscas e temporárias, permitindo uma melhor adequação das distribuições  $\alpha$ -Estáveis, por exemplo.

A distribuição  $\alpha$ -Estável, de acordo com Samorodnitsky e Taqqu (1994), é caracterizada por não possuir forma fechada e ter variância infinita quando o parâmetro de estabilidade  $\alpha \in (0, 2)$ , é menor que dois, o que dificulta sua utilização. É uma distribuição generalizada e que, dependendo do valor do parâmetro de estabilidade, pode se reduzir a distribuições mais usuais como Gaussiana, Cauchy e Lévy.

A estimação dos parâmetros do processo foi majoritariamente realizada por métodos de máxima verossimilhança. Para os parâmetros da  $\alpha$ -Estável, no entanto, foi utilizado o método por quantis abordado por McCulloch (1986). Para detecção da  $\alpha$ -Estável, foram implementados os testes de adequação apresentados por Beaulieu et al. (2014). Mediu-se a dependência entre as variáveis  $\alpha$ -Estáveis, entre si e entre dois diferentes processos, através das funções de Codiferença e Codiferença Cruzada.

O trabalho é apresentado em cinco capítulos, iniciando por esta introdução ao estudo. O segundo capítulo aborda as teorias e fundamentações aqui utilizadas. Em seguida, o terceiro capítulo apresenta as técnicas de estimação utilizadas nos Capítulos 4 e 5. O quarto capítulo inicia a aplicação a dados onde realizou-se simulações de Monte Carlo de processos auto-regressivos bivariados com as inovações de interesse e seus respectivos parâmetros. Em seguida, comparou-se os resultados obtidos pelos métodos de estimação clássicos e Bayesianos, abordados por Ashcar et al. (2013). Por fim, o quinto capítulo aplica a teoria aqui apresentada para os dados diários das ações das empresas *Coca-Cola Company* e *PepsiCo Inc.*, no período de 01 de Janeiro de

2015 até 29 de Outubro de 2019.

## 2

# Conceitos e Resultados Preliminares

## 2.1 Processos Estocásticos e Séries Temporais

Nesta seção apresentamos algumas definições básicas necessárias para um estudo de séries temporais. Uma *série temporal* é um conjunto de observações ordenadas no tempo.

Dentre os diversos motivos que nos levam a analisar séries temporais, definidas em Morettin e Tolo (2004), podemos destacar os seguintes:

- investigar o mecanismo gerador da série temporal, de modo a aproximar como ela foi gerada;
- fazer previsões dos valores futuros da série temporal, para podermos nos prevenir a respeito do comportamento que ela poderá assumir no futuro;
- descrever o comportamento da série temporal através da análise de seu gráfico para identificar características como tendência, ciclos, variações sazonais e valores anômalos;
- procurar periodicidades relevantes nos dados. Uma das técnicas que pode ser utilizada para este fim é a análise espectral.

Tais análises são baseadas na utilização de *modelos probabilísticos*, também denominados *modelos estocásticos*, constituindo o desafio de encontrar o modelo mais simples e que melhor descreva a série temporal analisada.

Ao iniciarmos a análise de uma série temporal, devemos estabelecer um método de como tratar com a natureza aleatória do fenômeno a ser modelado. A maneira usual é supor que cada observação da série temporal seja uma variável aleatória  $X_t$ , onde  $t$  designa o tempo em que a observação foi coletada. Assim, a série temporal  $\{X_t\}_{t=1}^n$  nada mais é do que uma realização (ou parte de uma realização) de um processo estocástico.

A partir de agora, introduzimos algumas definições preliminares, necessárias para o estudo realizado neste trabalho. Os conceitos de estacionariedade, função de distribuição  $n$ -dimensional

e as definições das funções de autocovariância e autocorrelação são de suma importância na busca de modelos para séries temporais. Iniciamos com a definição de processo estocástico.

**Definição 2.1.1.** Seja  $T$  um conjunto numérico qualquer. Um *processo estocástico* é uma família de variáveis aleatórias  $\{X_t\}_{t \in T}$  em um espaço de probabilidade  $(\Omega, \mathcal{A}, \mathbb{P})$ , em que  $\Omega$  é o espaço amostral,  $\mathcal{A}$  é a  $\sigma$ -álgebra da classe de eventos aleatórios e  $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$  é a função que associa probabilidade a um evento qualquer  $A \in \mathcal{A}$ .

**Observação 2.1.1.** 1. Em geral, consideramos o conjunto  $T$  como sendo o conjunto dos inteiros  $\mathbb{Z}$  ou dos reais  $\mathbb{R}$ .

2. Neste trabalho, consideramos o conjunto  $T = \{1, 2, \dots, n\}$  quando estivermos com uma realização de tamanho  $n$  do processo em estudo, ou seja, quando utilizarmos amostras.

3. Neste trabalho, consideramos o conjunto  $T$  como sendo  $\mathbb{Z}$ , ou seja, os processos são à tempo discreto.

**Definição 2.1.2.** Um espaço de probabilidade  $\mathcal{H} = \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$  é um espaço de Hilbert se satisfizer as seguintes condições:

- a.  $\mathbb{E}(X^2) = \int_{\Omega} X(w)^2 \mathbb{P}(dw) < \infty$ , para todo  $X \in \mathcal{H}$ ,
- b. a multiplicação por escalar e adição de vetores é a usual,
- c.  $\mathbb{E}(\zeta X)^2 = \zeta^2 \mathbb{E}(X^2)$ , para todo  $\zeta \in \mathbb{R}$  e  $X \in \mathcal{H}$ ,
- d.  $\mathbb{E}(X + Y)^2 \leq 2\mathbb{E}(X^2) + 2\mathbb{E}(Y^2)$ , para todo  $X, Y \in \mathcal{H}$ .

Definimos o produto interno em  $\mathcal{H}$  como sendo

$$\langle X, Y \rangle = \mathbb{E}(XY), \quad \text{para todo } X, Y \in \mathcal{H}.$$

**Definição 2.1.3.** Seja  $T$  um conjunto qualquer. Uma *série temporal*  $\{X_t\}_{t=1}^n$  é uma amostra finita de um processo estocástico  $\{X_t\}_{t \in T}$ .

A seguir, definimos a *função de distribuição  $n$ -dimensional* (ou *função de distribuição conjunta*) para um vetor aleatório  $(X_{t_1}, \dots, X_{t_n})$ .

**Definição 2.1.4.** Seja  $(X_{t_1}, \dots, X_{t_n})$  um vetor aleatório (ou variável aleatória  $n$ -dimensional), cujas componentes são variáveis aleatórias definidas no mesmo espaço de probabilidade  $(\Omega, \mathcal{A}, \mathbb{P})$ . A *função de distribuição  $n$ -dimensional* (ou *função de distribuição conjunta*) do vetor é definida por

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = \mathbb{P}(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n), \quad (2.1)$$

quaisquer que sejam  $n \in \mathbb{N}$ ,  $t_i \in T$ ,  $x_i \in \mathbb{R}$ , para todo  $i = 1, 2, \dots, n$ .

Um processo estocástico  $\{X_t\}_{t \in T}$  estará especificado se conhecermos as distribuições finito-dimensionais, dadas em (2.1), para todo  $n \geq 1$ .

As *funções de autocovariância* e de *autocorrelação* de um processo estocástico são definidas a seguir. Estas duas funções fornecem o grau de interdependência entre as variáveis aleatórias e são importantes na identificação de um modelo para uma série temporal.

**Definição 2.1.5.** Considere um processo estocástico  $\{X_t\}_{t \in T}$  onde, para todo  $t \in T$ ,  $\text{Var}(X_t) < \infty$ . A *função de autocovariância* (FAC) do processo, denotada por  $\gamma_X(\cdot, \cdot)$ , é dada por

$$\gamma_X(r, s) \equiv \text{Cov}(X_r, X_s) = \mathbb{E}[(X_r - \mathbb{E}(X_r))(X_s - \mathbb{E}(X_s))], \quad r, s \in T,$$

onde  $\mathbb{E}(X_t) \equiv \mu_t$  é a *esperança matemática* da variável aleatória  $X_t$ , para todo  $t \in T$ .

**Definição 2.1.6.** Um processo estocástico  $\{X_t\}_{t \in T}$  com  $\text{Var}(X_t) < \infty$ , para todo  $t \in T$ . A *função de autocorrelação* do processo, denotada por  $\rho_X(\cdot, \cdot)$ , é dada por

$$\rho_X(r, s) = \frac{\gamma_X(r, s)}{\sqrt{\text{Var}(X_r)}\sqrt{\text{Var}(X_s)}}, \quad r, s \in T,$$

onde  $\text{Var}(X_t) \equiv \gamma_X(t, t)$ , para todo  $t \in T$ .

Na busca de modelos para descrever séries temporais, é necessário utilizar suposições que permitam simplificar as séries temporais analisadas. Dentre as suposições que devem ser realizadas, a mais comum diz respeito à estacionariedade do processo.

De acordo com Brockwell e Davis (1991) e Morettin e Toloi (2004), um processo estacionário  $\{X_t\}_{t \in \mathbb{Z}}$  é aquele no qual a origem temporal ( $t = 0$ ) não é importante. Em outras palavras, as características de  $X_{t+h}$  são as mesmas de  $X_t$ , para todo  $t, h \in \mathbb{Z}$ .

A literatura apresenta dois tipos de estacionariedade: *estacionariedade forte* e *estacionariedade fraca*. As definições destas propriedades são dadas a seguir.

**Definição 2.1.7.** Um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é *fortemente estacionário* (ou *estritamente estacionário*) se todas as distribuições  $n$ -dimensionais (2.1) permanecem as mesmas sob translações do tempo, ou seja,

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = F_{t_1+h, \dots, t_n+h}(x_1, \dots, x_n),$$

para todo  $t_1, \dots, t_n, h \in \mathbb{Z}$ .

**Definição 2.1.8.** Um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é *fracamente estacionário* (ou somente *estacionário*) se

- a.  $\mathbb{E}|X_t|^2 < \infty$ , para todo  $t \in \mathbb{Z}$ ;
- b.  $\mathbb{E}(X_t) = \mu_X$ , uma constante independente de  $t$ ;
- c.  $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ , para quaisquer  $r, s, t \in \mathbb{Z}$ .

**Definição 2.1.9.** Um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é *Gaussiano* se, para qualquer conjunto  $t_1, \dots, t_n \in \mathbb{Z}$ , as variáveis aleatórias  $X_{t_1}, \dots, X_{t_n}$  têm distribuição normal  $n$ -dimensional.

**Observação 2.1.2.** 1. Um processo  $\{X_t\}_{t \in \mathbb{Z}}$  com  $\mathbb{E}|X_t|^2 < \infty$ , fracamente estacionário não necessita ser fortemente estacionário, mas a recíproca é sempre verdadeira. Como um processo Gaussiano é determinado pelas médias e pela matriz de variâncias-covariâncias, se ele for fracamente estacionário, será também fortemente estacionário.

2. Se o processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  for estacionário, temos que  $\gamma_X(r, s) = \gamma_X(r - s, 0)$ , para todo  $r, s \in \mathbb{Z}$ . Assim, podemos redefinir a função de autocovariância de um processo estocástico estacionário em termos de apenas uma variável:

$$\begin{aligned} \gamma_X(r, s) &= \gamma_X(r - s, s - s) = \gamma_X(r - s, 0) \\ &\equiv \gamma_X(h), \text{ para todo } r, s, h \in \mathbb{Z}. \end{aligned}$$

A função  $\gamma_X(h)$  será chamada de *função de autocovariância* de ordem  $h$  (ou também chamado de *lag*  $h$ ). De forma análoga, definimos a função de autocorrelação de ordem  $h$  como sendo

$$\rho_X(h) \equiv \frac{\gamma_X(h)}{\gamma_X(0)}, \quad \text{para todo } h \in \mathbb{Z},$$

onde  $\gamma_X(0) = \text{Cov}(X_t, X_t) = \text{Var}(X_t)$ , para todo  $t \in \mathbb{Z}$ .

3. As funções de autocovariância e de autocorrelação de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  possuem as seguintes propriedades:

- $\gamma_X(h) = \text{Cov}(X_{t+h}, X_t) = \text{Cov}(X_t, X_{t-h})$ , para todo  $t, h \in \mathbb{Z}$ ;
- $\gamma_X(h) = \gamma_X(-h)$ , para todo  $h \in \mathbb{N}$ ;
- $|\gamma_X(h)| \leq \gamma_X(0)$ , para todo  $h \in \mathbb{N}$ ;
- $\rho_X(h) = \rho_X(-h)$ , para todo  $h \in \mathbb{N}$ ;
- $|\rho_X(h)| \leq \rho_X(0) = 1$ , para todo  $h \in \mathbb{N}$ .

**Definição 2.1.10.** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário com média zero e com função de autocovariância  $\gamma_X(\cdot)$  tal que  $\gamma_X(h) \rightarrow 0$ , quando  $|h| \rightarrow \infty$ . A *função de autocorrelação parcial* (FACP), denotada por  $\phi_X(k, j)$ ,  $j = 1, \dots, k$ ,  $k \in \mathbb{N}$ , é o coeficiente na equação

$$\mathcal{P}_{\overline{\text{sp}}(X_1, \dots, X_k)}(X_{k+1}) = \sum_{j=1}^k \phi_X(k, j) X_{k+1-j},$$

onde  $\mathcal{P}_{\overline{\text{sp}}(X_1, \dots, X_k)}(X_{k+1})$  é a projeção ortogonal de  $X_{k+1}$  no subespaço fechado  $\overline{\text{sp}}(X_1, \dots, X_k)$  gerado pelas  $k$  observações anteriores.

A partir da equação

$$\langle X_{k+1} - \mathcal{P}_{\overline{sp}(X_1, \dots, X_k)}(X_{k+1}), X_j \rangle = 0, \quad \text{para } j = 1, \dots, k,$$

onde  $\langle \cdot, \cdot \rangle$  é o produto interno, obtemos os coeficientes  $\phi_X(k, j)$ , para  $j \in \{1, \dots, k\}$ . Através da Definição 2.1.2, obtemos o sistema de equações

$$\begin{bmatrix} 1 & \rho_X(1) & \rho_X(2) & \cdots & \rho_X(k-1) \\ \rho_X(1) & 1 & \rho_X(1) & \cdots & \rho_X(k-2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_X(k-1) & \rho_X(k-2) & \rho_X(k-3) & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_X(k, 1) \\ \phi_X(k, 2) \\ \vdots \\ \phi_X(k, k) \end{bmatrix} = \begin{bmatrix} \rho_X(1) \\ \rho_X(2) \\ \vdots \\ \rho_X(k) \end{bmatrix}, \quad (2.2)$$

onde  $\rho_X(\cdot)$  é a função de autocorrelação do processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  dada pela Definição 2.1.6. Os coeficientes  $\phi_X(\cdot, \cdot)$  são unicamente determinados pelo sistema (2.2).

Utilizando a regra de Cramér sucessivamente para  $k \in \mathbb{N}$ , obtemos a *função de autocorrelação parcial de ordem k* do processo  $\{X_t\}_{t \in \mathbb{Z}}$  dada por

$$\phi_X(k, k) = \frac{\begin{vmatrix} 1 & \rho_X(1) & \rho_X(2) & \cdots & \rho_X(k-2) & \rho_X(1) \\ \rho_X(1) & 1 & \rho_X(1) & \cdots & \rho_X(k-3) & \rho_X(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_X(k-1) & \rho_X(k-2) & \rho_X(k-3) & \cdots & \rho_X(1) & \rho_X(k) \end{vmatrix}}{\begin{vmatrix} 1 & \rho_X(1) & \cdots & \rho_X(k-2) & \rho_X(k-1) \\ \rho_X(1) & 1 & \cdots & \rho_X(k-3) & \rho_X(k-2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_X(k-1) & \rho_X(k-2) & \cdots & \rho_X(1) & 1 \end{vmatrix}}.$$

Um dos processos estocásticos mais simples é aquele cujas variáveis aleatórias são independentes e identicamente distribuídas. Definimos os processos independentes e identicamente distribuídos e ruído branco a seguir.

**Definição 2.1.11.** O processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é dito ser composto por *variáveis aleatórias independentes e identicamente distribuídas* (i.i.d.), denotado por  $X_t \sim IID(0, \sigma^2)$ , se valer para toda variável  $X_{t_1}, \dots, X_{t_n}$ , a condição

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = F_{t_1}(x_1) \cdots F_{t_n}(x_n)$$

e se as variáveis aleatórias  $X_t, t \geq 1$ , possuírem a mesma distribuição.

**Definição 2.1.12.** O processo estocástico  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  é um *ruído branco* com média zero e variância  $\sigma_\varepsilon^2$ , denotado por  $\varepsilon_t \sim RB(0, \sigma_\varepsilon^2)$ , se é i.i.d. e se

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

**Exemplo 2.1.1.** Seja  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  um processo independente com média zero e variância  $\sigma_\varepsilon^2$ . Defina  $\{X_t\}_{t \in \mathbb{Z}}$  por

$$X_t = \beta_0 + \beta_1 t + \varepsilon_t,$$

em que  $\beta_0, \beta_1 \in \mathbb{R}$  são constantes fixas. Desejamos mostrar que  $\{X_t\}_{t \in \mathbb{Z}}$  é não estacionário. Observe que

$$\mathbb{E}(X_t) = \mathbb{E}(\beta_0 + \beta_1 t + \varepsilon_t) = \beta_0 + \beta_1 t,$$

pois  $\mathbb{E}(\varepsilon_t) = 0$ .

Desta forma, mostramos que o processo  $X_t$  não é estacionário, pois sua média é uma constante que depende de  $t$ .

△

**Exemplo 2.1.2.** Considere  $X_t = \sum_{j=1}^n f_j(t)\varepsilon_j$ , para todo  $t \in \mathbb{Z}$ , uma variável aleatória onde  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  é um ruído branco, como apresentado na Definição 2.1.12, e  $\{f_j\}_{j=1}^n$  uma coleção de funções reais integráveis. Considere que exista uma função  $g : \mathbb{Z} \rightarrow \mathbb{R}$  satisfazendo  $0 < g(0) < \infty$  e

$$\sum_{j=1}^n f_j(a)f_j(b) = g(|b - a|), \forall a, b, \in \mathbb{Z}.$$

Vamos mostrar que, neste caso,  $\{X_t\}_{t \in \mathbb{Z}}$  é fracamente estacionário.

Primeiramente, será calculada a esperança do processo, com o intuito de provar que ela é uma constante independente de  $t$ .

$$\mathbb{E}(X_t) = \mathbb{E} \left[ \sum_{j=1}^n f_j(t)\varepsilon_j \right] = \sum_{j=1}^n f_j(t)\mathbb{E}(\varepsilon_j) = 0,$$

pois  $\mathbb{E}(\varepsilon_t) = 0$ , para todo  $t \in \mathbb{Z}$ .

O processo  $X_t$  possui média constante e não dependente de  $t$ .

O segundo passo para provar a estacionariedade do processo é concluir que o seu segundo momento é finito. Observe que

$$\begin{aligned} \mathbb{E}(X_t^2) &= \mathbb{E} \left[ \left( \sum_{j=1}^n f_j(t)\varepsilon_j \right)^2 \right] = \mathbb{E} \left[ \sum_{j=1}^n \sum_{k=1}^n f_j(t)f_k(t)\varepsilon_j\varepsilon_k \right] = \\ &= \sum_{j=1}^n \sum_{k=1}^n f_j(t)f_k(t)\mathbb{E}(\varepsilon_j\varepsilon_k) = \sum_{j=1}^n f_j(t)^2\sigma_\varepsilon^2 = g(0)\sigma_\varepsilon^2. \end{aligned}$$

Como  $g(0) < \infty$  e  $\sigma_\varepsilon^2 < \infty$ , concluímos que  $\mathbb{E}(|X_t|^2) = \mathbb{E}(X_t^2) < \infty$ .

Observe que

$$\begin{aligned} Cov(X_t, X_{t+h}) &= Cov\left(\sum_{j=1}^n f_j(t)\varepsilon_j, \sum_{k=1}^n f_k(t+h)\varepsilon_k\right) = \\ &= \sum_{j=1}^n \sum_{k=1}^n f_j(t)f_k(t+h)Cov(\varepsilon_j, \varepsilon_k). \end{aligned}$$

Agora,

$$Cov(\varepsilon_j, \varepsilon_k) = \begin{cases} \sigma_\varepsilon^2, & j = k, \\ 0, & j \neq k. \end{cases} \quad (2.3)$$

Então, se  $j = k$

$$Cov(X_t, X_{t+h}) = \sum_{j=1}^n \sum_{k=1}^n f_j(t)f_k(t+h)\sigma_\varepsilon^2 = g(|t+h-t|)\sigma_\varepsilon^2 = g(|h|)\sigma_\varepsilon^2.$$

E se  $j \neq k$

$$Cov(X_t, X_{t+h}) = 0,$$

pois  $Cov(\varepsilon_j, \varepsilon_k) = 0$ , para  $j \neq k$ .

Por fim,

$$Cov(X_t, X_{t+h}) = \begin{cases} g(|h|)\sigma_\varepsilon^2, & h = 0, \\ 0, & h \neq 0. \end{cases}$$

Desta forma, mostramos que a covariância do processo não depende de  $t$  e é finita. Portanto o processo  $X_t$  é fracamente estacionário.

△

## 2.2 Distribuição $\alpha$ -Estável

### 2.2.1 Univariada

As distribuições  $\alpha$ -Estáveis fazem parte de uma família que admite assimetria e caudas mais pesadas, além de outras características matemáticas. Essa classe foi caracterizada por Paul Lévy na década de 1920 (veja Samorodnitsky e Taqqu, 1994). A falta de fórmulas fechadas para as funções de distribuição e densidade, exceto em três casos, tem sido o maior inconveniente no uso dessa família. Porém, hoje possuímos programas mais confiáveis para o cálculo de funções densidade, funções de distribuição e quantis, muito utilizados em problemas aplicados.

A denominação *estável* provém do parâmetro de estabilidade  $\alpha \in (0, 2]$ , cujo valor define inversamente a probabilidade de valores extremos na distribuição. Este parâmetro também é chamado de *índice de estabilidade* ou *expoente característico*.

A notação utilizada para denominar uma distribuição  $\alpha$ -Estável segue a forma  $S_\alpha(\beta, \sigma, \delta)$ . O parâmetro  $\beta \in [-1, 1]$  refere-se a simetria, onde valores positivos indicam assimetria à direita. De forma análoga, valores negativos indicam assimetria à esquerda, e nulo ( $\beta = 0$ ) indica simetria. O parâmetro  $\sigma > 0$  designa a escala da distribuição e  $\delta \in \mathbb{R}$  é o parâmetro de locação.

Há diversas parametrizações utilizadas ao longo dos anos por diferentes autores em diversas áreas de atuação. A fim de explicitar qual parametrização está sendo utilizada e evitar confusões, iremos utilizar a parametrização 1 (ver Nolan, 2015).

Uma consequência das distribuições  $\alpha$ -Estáveis é que nem todos os momentos existem. Isso decorre que, quando  $0 < \alpha < 2$ , não temos segundo momento finito, e se  $\alpha \leq 1$ , a média é infinita. Dessa forma, podemos entender que o parâmetro de locação  $\delta$  não será necessariamente a média do processo. O mesmo aplica-se ao parâmetro  $\sigma$  de escala.

Uma propriedade importante de variáveis aleatórias Gaussianas consiste em que a soma é, também, uma variável aleatória Gaussiana. Proveniente disso, se  $X$  é Gaussiana, para  $X_1$  e  $X_2$  cópias independentes de  $X$ , ou seja, provém da mesma distribuição de  $X$  e não apresentam dependência. Para quaisquer constantes positivas  $a$  e  $b$ ,

$$aX_1 + bX_2 \stackrel{d}{=} cX + d, \quad (2.4)$$

para  $c$  positivo e  $d \in \mathbb{R}$ . O símbolo  $\stackrel{d}{=}$  denota igualdade em distribuição, isto é, ambas as expressões têm a mesma lei de distribuição. Em outras palavras, a equação (2.4) afirma que a forma de  $X$  é preservada (em escala e locação) sob adição. Na classe de distribuições  $\alpha$ -Estáveis, essa propriedade permanece, como segue.

**Definição 2.2.1.** Uma variável aleatória  $X$  é *estável* ou *fracamente estável* se para  $X_1$  e  $X_2$  sendo cópias independentes de  $X$  e para quaisquer constantes positivas  $a$  e  $b$ , a equação (2.4) for válida para qualquer valor positivo de  $c$  e  $d \in \mathbb{R}$ . A variável aleatória é *estritamente estável* ou *fortemente estável* se a equação (2.4) for válida para  $d = 0$  e para qualquer valor de  $a$  e  $b$ . Uma variável aleatória é *simetricamente estável* se for estável e simetricamente distribuída em torno de 0, ou seja,  $X \stackrel{d}{=} -X$ .

Na Figura 2.1 observamos gráficos da função densidade de uma  $\alpha$ -Estável conforme ocorrem mudanças em seus parâmetros.

Em seguida, apresentamos a definição de variáveis aleatórias  $X \sim S_\alpha(\beta, \sigma, \delta)$  e sua função característica.

**Definição 2.2.2.** Uma variável aleatória não-degenerada  $X$  é  $\alpha$ -Estável se e somente se para todo  $n > 1$  existirem constantes  $c_n > 0$  e  $d_n \in \mathbb{R}$  tais que

$$X_1 + \cdots + X_n \stackrel{d}{=} c_n X + d_n, \quad (2.5)$$

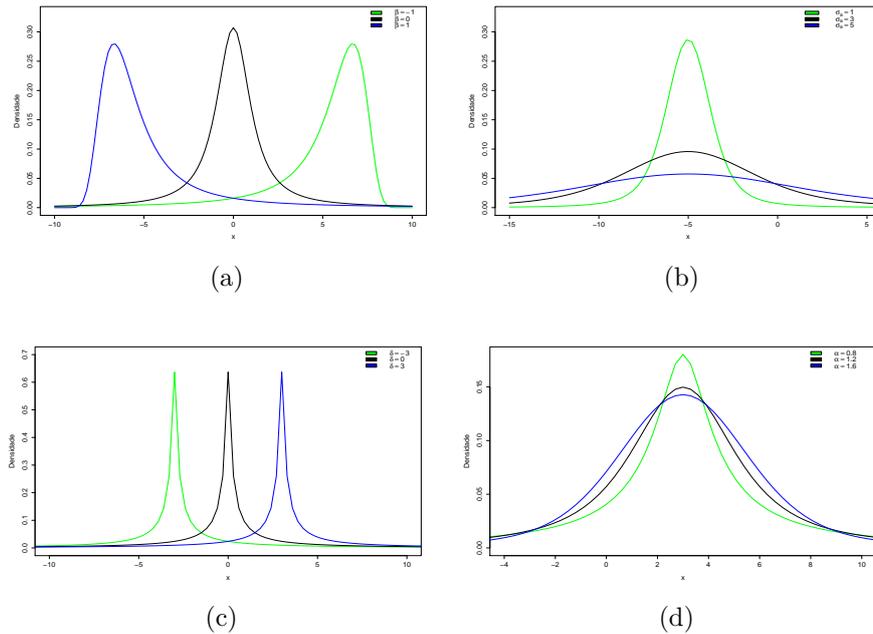


Figura 2.1: (a) Função densidade de uma  $\alpha$ -Estável para os valores de  $\beta \in \{-1, 0, 1\}$  com  $\alpha = 1.1$ ,  $\sigma = 1$  e  $\delta = 0$ ; (b) Função densidade de uma  $\alpha$ -Estável para os valores de  $\sigma \in \{1, 3, 5\}$  com  $\alpha = 1.5$ ,  $\beta = 0$  e  $\delta = -5$ ; (c) Função densidade de uma  $\alpha$ -Estável para os valores de  $\delta \in \{-3, 0, 3\}$  com  $\alpha = 0.5$ ,  $\sigma = 1$  e  $\beta = 0$ ; (d) Função densidade de uma  $\alpha$ -Estável para os valores de  $\alpha \in \{0.8, 1.2, 1.6\}$  com  $\sigma = 2$ ,  $\beta = 0$  e  $\delta = 3$ .

onde  $X_1, \dots, X_n$  são cópias idênticas e independentes de  $X$ .

Outra forma de caracterizar uma variável aleatória  $\alpha$ -Estável é dada a seguir.

**Definição 2.2.3.** Uma variável aleatória  $X$  é  $\alpha$ -Estável se e somente se  $X \stackrel{d}{=} \sigma Z + \delta$ , onde  $0 < \alpha \leq 2$ ,  $-1 \leq \beta \leq 1$ ,  $\sigma > 0$ ,  $\delta \in \mathbb{R}$ , com função característica dada por

$$\varphi_X(t) = \mathbb{E}(e^{itX}) = \begin{cases} \exp\left\{i\delta t - \sigma^\alpha |t|^\alpha \left[1 - i\beta \frac{t}{|t|} \tan\left(\frac{\pi\alpha}{2}\right)\right]\right\}, & \alpha \neq 1, \\ \exp\left\{i\delta t - \sigma^\alpha |t| \left[1 + i\beta \frac{2t}{|t|\pi} \ln |t|\right]\right\}, & \alpha = 1, \end{cases} \quad (2.6)$$

Estas distribuições são simétricas em torno de zero quando  $\beta = 0$  e  $\delta = 0$ . Neste caso, a função característica de  $\sigma Z$  tem uma expressão mais simples dada por

$$\varphi(t) = \exp(-\sigma^\alpha |t|^\alpha). \quad (2.7)$$

Quando  $\alpha \neq 1$ , a função característica (2.6) pode ser escrita como:

$$\varphi_X(t) = \exp\left\{-\tilde{\sigma}^\alpha |t|^\alpha e^{-i\tilde{\beta}(\text{sign}(t))\frac{\pi}{2}K(\alpha)} + i\delta t\right\}, \quad (2.8)$$

onde

$$\tilde{\sigma} = \sigma \left(1 + \beta^2 \tan^2 \alpha \frac{\pi}{2}\right)^{\frac{1}{2\alpha}},$$

$$K(\alpha) = \begin{cases} \alpha, & \text{se } \alpha < 1 \\ \alpha - 2, & \text{se } \alpha > 1, \end{cases}$$

$$\tilde{\beta} = \begin{cases} \frac{2}{\pi\alpha} \arctan(\beta \tan \alpha \frac{\pi}{2}), & \text{se } 0 < \alpha < 1 \\ \frac{2}{\pi(\alpha-2)} \arctan(\beta \tan \frac{\pi(\alpha-2)}{2}), & \text{se } 1 < \alpha < 2, \end{cases}$$

e

$$\text{sign}(t) = \begin{cases} -1, & t < 0, \\ 0, & t = 0 \\ 1, & t > 0. \end{cases}$$

Pode-se mostrar que, se  $X \sim \mathcal{N}(\mu, \sigma^2)$  então  $X \sim S_2(0, \frac{\sigma}{\sqrt{2}}, \mu)$ . Da mesma forma, uma distribuição de Cauchy  $\mathcal{C}(\sigma, \delta)$  é estável e é denotada por  $S_1(0, \sigma, \delta)$  e uma distribuição de Lévy  $\mathcal{L}(\sigma, \delta)$  é também estável com notação  $S_{0.5}(1, \sigma, \delta)$ . Apenas nestes três casos a variável aleatória  $\alpha$ -Estável tem função de distribuição com forma fechada e conhecida. Em qualquer outra situação não existe fórmula explícita para as funções de distribuição e densidade de uma  $\alpha$ -Estável geral.

Através da função característica da distribuição, podemos definir algumas propriedades básicas.

**Proposição 2.2.1.** Sejam  $X_1$  e  $X_2$  variáveis aleatórias independentes com  $X_i \sim S_\alpha(\beta_i, \sigma_i, \delta_i)$ ,  $i = 1, 2$ , então  $X_1 + X_2 \sim S_\alpha(\beta, \sigma, \delta)$  onde

$$\sigma = (\sigma_1^\alpha + \sigma_2^\alpha)^{1/\alpha}, \quad \beta = (\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha) / (\sigma_1^\alpha + \sigma_2^\alpha), \quad \delta = \delta_1 + \delta_2. \quad (2.9)$$

**Proposição 2.2.2.** Dado uma variável  $X \sim S_\alpha(\beta, \sigma, \delta)$  e uma constante real  $a$ ,  $X + a \sim S_\alpha(\beta, \sigma, \delta + a)$ .

**Proposição 2.2.3.** Dado uma variável  $X \sim S_\alpha(\beta, \sigma, \delta)$  e uma constante real não-nula  $a$ ,

$$\begin{aligned} aX &\sim S_\alpha(\text{sign}(a)\beta, |a|\sigma, a\delta), & \text{se } \alpha \neq 1 \\ aX &\sim S_1(\text{sign}(a)\beta, |a|\sigma, a\delta - (2/\pi)a(\ln|a|)\sigma\beta), & \text{se } \alpha = 1. \end{aligned} \quad (2.10)$$

**Proposição 2.2.4.** Para qualquer  $0 < \alpha < 2$ ,

$$X \sim S_\alpha(\beta, \sigma, 0) \iff -X \sim S_\alpha(-\beta, \sigma, 0). \quad (2.11)$$

**Definição 2.2.4.** Uma variável  $X \sim S_\alpha(\beta, \sigma, \delta)$  é dita ser *simétrica* se e somente se  $\beta = 0$  e  $\delta = 0$ . É *simétrica em torno de  $\delta$*  se e somente se  $\beta = 0$ .

**Proposição 2.2.5.** Seja  $X \sim S_\alpha(\beta, \sigma, \delta)$  com  $\alpha \neq 1$ . Dizemos que  $X$  é estritamente estável se e somente se  $\delta = 0$ .

**Proposição 2.2.6.** Dizemos que  $X \sim S_1(\beta, \sigma, \delta)$  é estritamente estável se e somente se  $\beta = 0$ .

**Proposição 2.2.7.** Seja  $X \sim S_\alpha(\beta, \sigma, \delta)$  com  $0 < \alpha < 2$ . Então

$$\begin{aligned}\mathbb{E}|X|^p &< \infty, \quad \text{para qualquer } 0 < p < \alpha, \\ \mathbb{E}|X|^p &= \infty, \quad \text{para qualquer } p \geq \alpha.\end{aligned}$$

O fato de variáveis  $\alpha$ -Estáveis, com  $\alpha < 2$ , terem segundo momento infinito significa que muitas das técnicas válidas para os casos Gaussianos não se aplicam. Além disso, quando  $\alpha \leq 1$  temos  $E|X| = \infty$ .

## 2.2.2 Multivariada

A definição de estabilidade é definida em Samorodnitsky e Taqqu (1994).

**Definição 2.2.5.** Um vetor  $\underline{X} = (X_1, X_2, \dots, X_d)$  é um *vetor  $\alpha$ -Estável* em  $\mathbb{R}^d$  se para quaisquer números positivos A e B existem um número positivo C e um vetor  $D \in \mathbb{R}^d$  tais que

$$A\underline{X}^{(1)} + B\underline{X}^{(2)} \stackrel{d}{=} C\underline{X} + D, \quad (2.12)$$

onde  $\underline{X}^{(1)}$  e  $\underline{X}^{(2)}$  são cópias independentes de  $\underline{X}$ .

Equivalentemente, para  $0 < \alpha < 2$ ,  $\underline{X} = (X_1, X_2, \dots, X_d)$  é um *vetor  $\alpha$ -Estável* em  $\mathbb{R}^d$  se e somente se existem uma medida finita  $\Gamma$  na esfera unitária  $S_d$  de  $\mathbb{R}^d$  e um vetor  $\underline{\delta}^0$  em  $\mathbb{R}^d$  tais que

$$\varphi_{\underline{X}}(\underline{t}) = \begin{cases} \exp \left\{ - \int_{S_d} |(\underline{t}, \underline{s})|^\alpha (1 - i \operatorname{sign}(\underline{t}, \underline{s}) \tan((\pi\alpha)/2)) \Gamma(d\underline{s}) + i(\underline{t}, \underline{\delta}^0) \right\}, & \alpha \neq 1, \\ \exp \left\{ - \int_{S_d} |(\underline{t}, \underline{s})|^\alpha (1 + i(2/\pi) \operatorname{sign}(\underline{t}, \underline{s}) \ln(|(\underline{t}, \underline{s})|)) \Gamma(d\underline{s}) + i(\underline{t}, \underline{\delta}^0) \right\}, & \alpha = 1. \end{cases} \quad (2.13)$$

O par  $(\Gamma, \underline{\delta}^0)$  é único e  $S_d = \{\underline{s} \in \mathbb{R}^d; \|\underline{s}\| = 1\}$  é a esfera unitária em  $\mathbb{R}^d$ . A medida  $\Gamma$  é chamada de *medida espectral* do vetor aleatório  $\alpha$ -Estável  $\underline{X}$ .

**Observação 2.2.1.** a. Segue da Definição 2.2.5 que um vetor aleatório  $\underline{X}$  é  $\alpha$ -Estável se possui a função característica  $\varphi_{\underline{X}}(\cdot)$  dada em (2.13).

b. A Definição 2.2.5 se mantém se  $\mathbb{R}^d$  é substituído por um espaço de Hilbert separável  $H$  com produto escalar  $(\cdot, \cdot)$ . Por exemplo,  $H$  pode ser  $\ell^2$ , o espaço de sequências reais  $X = \{x_1, x_2, \dots\}$  com  $\|x\|^2 = \sum_{k=1}^{\infty} x_k^2 < \infty$  e produto escalar  $(x, y) = \sum_{k=1}^{\infty} x_k y_k$ . Este teorema com  $H = \ell^2$  pode ser utilizado para demonstrar que processos estáveis sempre possuem representação integral.

c. A função característica de  $\tilde{X}$  é dada por  $\varphi_{\tilde{X}}(t) = \mathbb{E}(e^{i(t, \tilde{X})}) = \mathbb{E}(e^{i \sum_{k=1}^d t_k X_k})$ .

## 2.3 Processos Auto-regressivos de Ordem $p$

Processos auto-regressivos (AR) (Morettin e Toloi, 2004) são representações de processos aleatórios, nos quais se especifica que a variável de interesse depende linearmente de seus próprios valores passados e de um termo estocástico, a inovação. São frequentemente utilizados para estimação e previsão de séries temporais.

### 2.3.1 AR( $p$ ) Univariado

Um *processo auto-regressivo de ordem  $p$*  é um processo estocástico definido como

$$X_t = c + \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t, \quad \text{para todo } t \in \mathbb{Z}, \quad (2.14)$$

onde  $\varepsilon_t \sim \text{RB}(0, \sigma_\varepsilon^2)$ , isto é,  $\varepsilon_t$  é um ruído branco com média zero e variância  $\sigma_\varepsilon^2$ , e  $\boldsymbol{\eta} = (c, \phi_1, \phi_2, \dots, \phi_p, \sigma_\varepsilon^2)$  é o vetor de parâmetros.

Definindo o operador auto-regressivo estacionário de ordem  $p$  dado por

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad \text{onde } B^k(X_t) = X_{t-k}, \quad \text{para todo } k, t \in \mathbb{Z},$$

podemos reescrever a expressão (2.14) como sendo

$$\phi(B)X_t = \varepsilon_t. \quad (2.15)$$

**Exemplo 2.3.1.** Para ilustrar um processo AR(1) Gaussiano, suponhamos  $\{X_t\}_{t=1}^n$  uma amostra de tamanho  $n = 300$  gerada pelo *software* estatístico R de um processo AR(1) Gaussiano dado por

$$X_t = 5 + 0.5X_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots, 300, \quad (2.16)$$

onde  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  é o processo de variáveis aleatórias i.i.d. também gerados pelo *software* com distribuição normal de média 0 e variância 1, ou seja,  $\varepsilon_t \sim \mathcal{N}(0, 1)$ , para cada  $t \in \mathbb{Z}$ .

A Figura 2.2 apresenta uma amostra  $\{X_t\}_{t=1}^{300}$ , de tamanho amostral  $n = 300$ , do processo dado na expressão (2.16), observando que essa amostra se desenvolve aleatoriamente no tempo ao redor de uma média constante, neste caso, algo próximo de 10 (ver Figura 2.2(a)). Observando a função de autocorrelação (ver Figura 2.2(b)) nota-se um decaimento rápido, comportamento característico de processos AR, enquanto que a função de autocorrelação parcial (ver Figura 2.2(c)) apresenta um corte brusco no *lag* 1, indicando que, de fato, os dados gerados são advindos de um AR(1).

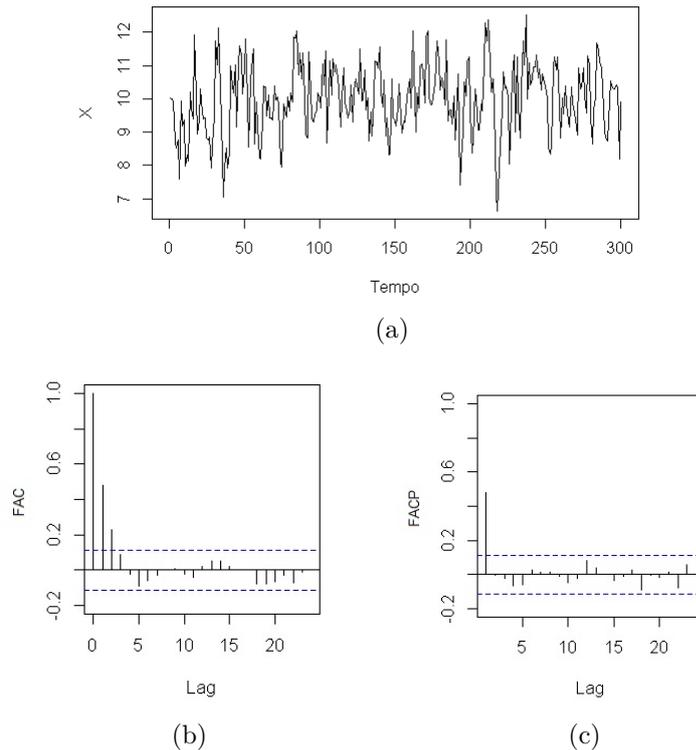


Figura 2.2: (a) Série temporal  $\{X_t\}_{t=1}^{300}$  referente ao processo AR(1) dado na expressão (2.16); (b) Função de autocorrelação amostral; (c) Função de autocorrelação parcial amostral.

△

Neste trabalho, serão abordadas inovações com três tipos de distribuições: a Gaussiana, a  $t$ -Student e a  $\alpha$ -Estável.

### Inovações Gaussianas

Um processo AR( $p$ ) com inovações Gaussianas é dado por (2.14), onde  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  é i.i.d. com distribuição  $\mathcal{N}(0, \sigma_\varepsilon^2)$ . Neste caso, o vetor de parâmetros é dado por  $\boldsymbol{\eta} = (c, \phi_1, \phi_2, \dots, \phi_p, \sigma_\varepsilon^2)$ .

### Inovações $t$ -Student

Um processo AR( $p$ ) com inovações  $t$ -Student é dado por (2.14), onde a inovação é dada por  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ , i.i.d. com distribuição  $t(gl)$ . Neste caso, o vetor de parâmetros é dado por  $\boldsymbol{\eta} = (c, \phi_1, \phi_2, \dots, \phi_p, gl)$ .

### Inovações $\alpha$ -Estáveis

Um processo AR( $p$ ) com inovações  $\alpha$ -Estáveis é dado por (2.14), onde  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  é uma sequência de variáveis aleatórias i.i.d. seguindo distribuição  $\mathcal{S}_\alpha(\beta, \sigma, \delta)$ , onde  $\alpha \in (0, 2)$ ,  $\beta \in (-1, 1)$ ,  $\sigma > 0$  e  $\delta \in \mathbb{R}$  são, respectivamente, os parâmetros de estabilidade, simetria, escala e locação. O vetor de parâmetros do modelo é dado por  $\boldsymbol{\eta} = (c, \phi_1, \phi_2, \dots, \phi_p, \alpha, \beta, \sigma, \delta)$ .

### 2.3.2 AR( $p$ ) Bivariado

O processo auto-regressivo vetorial (VAR) é um dos modelos mais flexíveis e fáceis de manusear na análise de séries temporais multivariadas. Uma extensão natural do modelo univariado, o modelo VAR se mostrou especialmente útil na descrição do comportamento dinâmico de séries econômicas e financeiras.

**Definição 2.3.1.** Seja  $\underline{X}_t = (X_{1t}, X_{2t}, \dots, X_{kt})$  um vetor de tamanho  $k$  de variáveis advindas do processo  $\{X_t\}_{t \in \mathbb{Z}}$ . O modelo VAR de ordem  $p$  (VAR( $p$ )) é dado por

$$\underline{X}_t = c + \phi_1 \underline{X}_{t-1} + \phi_2 \underline{X}_{t-2} + \dots + \phi_p \underline{X}_{t-p} + \varepsilon_t, \quad t = 1, \dots, T \quad (2.17)$$

onde  $\phi_j$  são matrizes de coeficientes (de ordem  $k \times k$ ),  $\varepsilon_t$  é um processo de ruído branco não observável, de média zero e matriz de covariância  $\Sigma$  e  $T$  é o tempo máximo.

A seguir, damos um exemplo de um VAR(1) bidimensional.

**Exemplo 2.3.2.** Um processo bivariado VAR(1) tem a forma

$$\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} X_{1t-1} \\ X_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (2.18)$$

ou

$$\begin{aligned} X_{1t} &= c_1 + \phi_{11} X_{1t-1} + \phi_{12} X_{2t-1} + \varepsilon_{1t} \\ X_{2t} &= c_2 + \phi_{21} X_{1t-1} + \phi_{22} X_{2t-1} + \varepsilon_{2t} \end{aligned} \quad (2.19)$$

onde  $Cov(\varepsilon_{1t}, \varepsilon_{2s}) = \sigma_{12}$ , para  $t = s$ , e 0 caso contrário.

△

O processo também pode ser escrito na forma

$$\Phi(B) \underline{X}_t = c + \varepsilon_t \quad (2.20)$$

onde  $\Phi(B) = I_k - \Phi_1 B - \dots - \Phi_p B^p$ . O VAR( $p$ ) é estacionário se as raízes de

$$\det(I_k - \Phi_1 z - \dots - \Phi_p z^p) = 0$$

estão fora do círculo unitário (possuem módulo maior do que um) ou, equivalentemente, se os auto-valores da matriz ( $k^2 \times k^2$ )

$$F = \begin{pmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_k \\ I_k & 0 & \dots & 0 \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & I_k & 0 \end{pmatrix} \quad (2.21)$$

possuem módulo inferior a um. Assumindo que o processo foi inicializado a um tempo infinito no passado, então o  $\text{VAR}(p)$  é estacionário e ergódico com médias, variâncias e autocovariâncias independentes do tempo.

## 2.4 Medidas de Correlação

Quando trabalhamos com séries multivariadas, um dos componentes mais importante da análise é a correlação entre elas. Nesta seção abordaremos alguns dos estimadores que serão aplicados nos Capítulos 4 e 5.

### 2.4.1 Coeficiente de Correlação de Pearson

**Definição 2.4.1.** Sejam  $\underline{X} = (X_1, X_2, \dots, X_n)$  e  $\underline{Y} = (Y_1, Y_2, \dots, Y_n)$  variáveis aleatórias com  $n$  observações. O *coeficiente de correlação de Pearson* é definido como

$$\rho = \frac{\sum_j (X_j - \bar{X})(Y_j - \bar{Y})}{\sqrt{\sum_j (X_j - \bar{X})^2 \sum_j (Y_j - \bar{Y})^2}} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}, \quad (2.22)$$

onde  $\bar{X} = \frac{1}{n} \sum_j^n X_j$  e  $\bar{Y} = \frac{1}{n} \sum_j^n Y_j$  são as médias aritméticas de ambas as variáveis.

O coeficiente de correlação de Pearson quantifica a associação linear entre duas variáveis aleatórias, ou seja, quantifica o quão bem uma linha reta se ajustaria a uma nuvem de pontos com posições em  $(X, Y)$ .

Algumas propriedades são:

- Se  $\rho > 0$ , a correlação é positiva, ou seja, o crescimento de uma variável acarreta no crescimento da outra;
- Se  $\rho < 0$ , temos correlação negativa. Neste caso, o crescimento de uma variável indica o decréscimo da outra;
- Se  $\rho = 0$ , a dependência entre as duas variáveis é nula.

### 2.4.2 Função Codiferença

Consideramos neste trabalho processos estacionários, isto é, para todo  $k \geq 0, t_1, t_2, \dots, t_k \geq 0$ , as distribuições finito-dimensionais do vetor  $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$  não dependem de  $h \geq 0$ .

**Definição 2.4.2.** Dadas duas variáveis aleatórias quaisquer  $X_1$  e  $X_2$ , a *função codiferença* é definida como

$$\tau(X_1, X_2) = \ln \left\{ \frac{\mathbb{E}[\exp(i(X_1 - X_2))]}{\mathbb{E}[\exp(i(X_1))]\mathbb{E}[\exp(-i(X_2))]} \right\}. \quad (2.23)$$

**Propriedade 2.4.1.** Algumas propriedades da *função codiferença* são dadas por

- Se  $X_1$  e  $X_2$  são variáveis aleatórias independentes, então  $\tau(X_1, X_2) = 0$ .
- Se  $X_1$  e  $X_2$  são variáveis aleatórias Gaussianas, então  $\tau(X_1, X_2) = Cov(X_1, X_2)$ .
- A função codiferença está bem definida mesmo quando o processo não possui segundo momento finito.

Se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo estocástico, então a *função codiferença* é dada por

$$\tau_X(h, t) = \tau(X_t, X_{t+h}), \quad (2.24)$$

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

Uma definição ainda mais geral para a codiferença, similar à definição de Rosadi e Deistler (2009) para processos estocásticos quaisquer, é dada por

$$\tau_X(s; h, t) = \ln\{\mathbb{E}(\exp(is(X_{t+h} - X_t)))\} - \ln\{\mathbb{E}[\exp(is(X_{t+h}))]\} - \ln\{\mathbb{E}[\exp(-is(X_t))]\}, \quad (2.25)$$

onde  $s \in \mathbb{R}$  e  $h, t \in \mathbb{Z}$ . Quando  $s = 1$ , a expressão (2.25) se reduz à (2.24). Assim, utiliza-se  $s$  como forma de parametrizar a função codiferença.

**Observação 2.4.1.** Se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo estacionário, então a expressão (2.25) não depende de  $t$ . Nesta situação, a codiferença será denotada por  $\tau_X(s; h)$ .

Consideramos o estimador proposto por Rosadi e Deistler (2009) para processos  $\alpha$ -Estáveis. Seja  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo estacionário e  $\{X_j\}_{j=1}^n$  uma amostra de tamanho  $n$  deste processo. Como a função codiferença é definida a partir de funções características, a função codiferença empírica pode ser estimada utilizando funções características empíricas. O estimador, proposto por Rosadi e Deistler (2009), para a função codiferença na defasagem  $h$  é dado por

$$\begin{aligned} \hat{\tau}_X(s; h) = & \sqrt{\frac{n-h}{n}} \left[ \ln\left(\frac{1}{n-h} \sum_{t=1}^{n-h} e^{is(X_{t+h} - X_t)}\right) - \ln\left(\frac{1}{n-h} \sum_{t=1}^{n-h} e^{isX_{t+h}}\right) \right. \\ & \left. - \ln\left(\frac{1}{n-h} \sum_{t=1}^{n-h} e^{-isX_t}\right) \right], \end{aligned} \quad (2.26)$$

para todo  $h \in \{0, \dots, n\}$ .

### 2.4.3 Função Codiferença Cruzada

Desenvolvendo o logaritmo na equação (2.23), obtemos:

$$CD(X) = \ln\{\mathbb{E} \exp(i(X_1 - X_2))\} - \ln\{\mathbb{E} \exp(iX_1)\} - \ln\{\mathbb{E} \exp(-iX_2)\}. \quad (2.27)$$

A partir da equação (2.27), podemos definir a *função Codiferença Cruzada* (medida similar à covariância cruzada, quando existem os momentos de ordem dois) e que pode ser aplicada a processos com variância infinita.

**Definição 2.4.3.** Se  $\underline{X}_t = (X_{1t}, X_{2t})$  é um processo bidimensional, então a *função codiferença cruzada* é definida como:

$$CD(X_{1t}, X_{2(t+k)}) = \ln\{\mathbb{E} \exp(i(X_{1t} - X_{2(t+k)}))\} - \ln\{\mathbb{E} \exp(iX_{1t})\} - \ln\{\mathbb{E} \exp(-iX_{2(t+k)})\}, \quad t = 1, \dots, T; \quad k \in \mathbb{R}. \quad (2.28)$$

Grzesiek, Teuerle e Wylomańska (2019) definem um estimador para a *função Codiferença Cruzada* para processos bidimensionais estacionários.

**Definição 2.4.4.** Se a série bidimensional  $\underline{X}(t)$  for originária de um processo estacionário, a *função Codiferença Cruzada* pode ser estimada a partir de uma trajetória através da seguinte estatística

$$\widehat{CD}(X_{1t}, X_{2(t+k)}) = \ln\{\hat{\Phi}(1, -1; k)\} - \ln\{\hat{\Phi}(1, 0; k)\} - \ln\{\hat{\Phi}(0, -1; k)\}, \quad (2.29)$$

onde o estimador da função característica empírica  $\Phi(u, v; k)$  é dado por

$$\hat{\Phi}(u, v; k) = \begin{cases} \frac{1}{n-k} \sum_{t=1}^{n-k} \exp\{i(uX_{1t} + vX_{2(t+k)})\}, & k \geq 0, \\ \frac{1}{n+k} \sum_{t=1}^{n+k} \exp\{i(uX_{1(t-k)} + vX_{2t})\}, & k < 0. \end{cases} \quad (2.30)$$

## 2.5 Testes de Adequação

Considere uma amostra aleatória i.i.d. e ordenada  $\underline{X} = (X_{(1)}, X_{(2)}, \dots, X_{(n)})$ , proveniente de uma distribuição desconhecida qualquer  $\mathbb{P}$ . Os testes apresentados nesta seção permitem testar se a amostra pertence a uma função de distribuição conhecida  $\mathbb{P}_0$ , ou seja

$$\begin{cases} H_0 : \mathbb{P} = \mathbb{P}_0 \\ H_1 : \mathbb{P} \neq \mathbb{P}_0. \end{cases} \quad (2.31)$$

Seja  $\underline{Y} = (Y_{(1)}, Y_{(2)}, \dots, Y_{(n)})$  a amostra padronizada dada por

$$Y_{(j)} = \frac{X_{(j)} - \bar{X}}{s_X}, \quad j = 1, \dots, n, \quad (2.32)$$

onde  $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$  e  $s_X = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})^2}$ .

Podemos definir a *função de distribuição empírica acumulada*  $Y$  através de

$$F_n(y) = \frac{1}{n} \sum_{j=1}^n I(Y_{(j)} \leq y), \quad (2.33)$$

onde  $I(A)$  indica a função indicadora no conjunto  $A$ .

### 2.5.1 Testes de Função de Distribuição Empírica

Quando trabalhamos com a função de distribuição  $\alpha$ -Estável, Beaulieu et al. (2014) mostram que não é necessário considerar os parâmetros  $\sigma$  e  $\delta$ , permitindo que escrevamos as hipóteses em (2.31) da forma

$$\begin{cases} H_0 : & \alpha = \hat{\alpha} \text{ e } \beta = \hat{\beta} \\ H_1 : & \alpha \neq \hat{\alpha} \text{ ou } \beta \neq \hat{\beta}. \end{cases} \quad (2.34)$$

Para isso, é preciso que os dados sejam padronizados de forma que se  $X \sim S_\alpha(\beta, \sigma, \delta)$  então  $Y \sim S_\alpha(\beta, 1, 0)$ .

Os valores hipotéticos  $\hat{\alpha}$  e  $\hat{\beta}$  são obtidos a partir do método de estimação apresentado no Capítulo 3.

Considere  $F_j^n = F_n(Y_{(j)})$  a função de distribuição empírica dos dados  $Y$  dados em (2.32) com função de distribuição empírica  $F_n(\cdot)$ , dada em (2.33). Seja  $F(\cdot; \alpha, \beta)$  a função de distribuição de uma variável aleatória com distribuição  $\alpha$ -Estável onde  $\sigma = 1$  e  $\delta = 0$ , índice de estabilidade  $\alpha$  e parâmetro de simetria  $\beta$ . Os testes de adequação utilizados nesse trabalho são os seguintes

- Teste de Kolmogorov-Smirnov (KS):

$$D_n = \max_{j \in \{1, \dots, n\}} [\hat{F}_j - F_{j-1}^n, F_j^n - \hat{F}_j]; \quad (2.35)$$

- Teste de Anderson-Darling (AD):

$$A_n^2 = -n - \sum_{j=1}^n \frac{(2j-1)}{n} [\ln(\hat{F}_j) + \ln(1 - \hat{F}_{n-j+1})]; \quad (2.36)$$

- Teste Modificado de Kolmogorov-Smirnov (MKS), Beaulieu et al. (2014), que considera maior peso às caudas do que o teste original:

$$MKS = \sqrt{n} \max_{j \in \{1, \dots, n\}} \left[ \frac{|F_j^n - \hat{F}_j|}{\hat{F}_j(1 - \hat{F}_j) + 1/n} \right]. \quad (2.37)$$

## 2.5.2 Teste por Quantis de McCulloch

Explorado em Beaulieu et al. (2014), este teste utiliza o método de estimação dos parâmetros  $\alpha$  e  $\beta$ , parâmetros de uma  $\alpha$ -Estável, proposto por McCulloch (1986).

O teste por quantil de McCulloch (MC) tem as seguintes estatísticas

$$\hat{\phi}_1(\alpha, \beta) = |\phi_1 - \bar{\phi}_1(\alpha, \beta)|, \quad \hat{\phi}_2(\alpha, \beta) = |\phi_2 - \bar{\phi}_2(\alpha, \beta)|, \quad (2.38)$$

onde

$$\phi_1 = \frac{\hat{Y}([95]) - \hat{Y}([5])}{\hat{Y}([75]) - \hat{Y}([25])}, \quad \phi_2 = \frac{\hat{Y}([95]) + \hat{Y}([5]) - 2\hat{Y}([50])}{\hat{Y}([95]) - \hat{Y}([5])}, \quad (2.39)$$

$\bar{\phi}_1(\alpha, \beta)$  e  $\bar{\phi}_2(\alpha, \beta)$  correspondem aos valores hipotéticos impondo a hipótese (2.34) e  $\hat{Y}([q])$  significa o  $q$ -ésimo percentil de  $Y$ . As expressões  $\bar{\phi}_1(\alpha, \beta)$  e  $\bar{\phi}_2(\alpha, \beta)$  são estimadas por simulação, com o seguinte algoritmo:

- Considere  $re$  amostras de tamanho  $n$  de uma variável aleatória padronizada com função de distribuição  $\alpha$ -Estável impondo as hipóteses (2.34);
- Para cada amostra, calcule os quantis e construa os indicadores  $\phi_1$  e  $\phi_2$  dados em (2.39);
- As médias dos  $re$  valores simulados de  $\phi_1$  e  $\phi_2$  correspondem a  $\bar{\phi}_1(\alpha, \beta)$  e  $\bar{\phi}_2(\alpha, \beta)$ , respectivamente.

## 2.6 Inferência Bayesiana

Seja  $(A_1, A_2, \dots, A_n)$  uma sequência de eventos mutuamente exclusivos e que constituem o espaço amostral  $\Omega$ , ou seja,  $\bigcup_{j=1}^n A_j = \Omega$  e  $A_j \cap A_l = \emptyset$ , para  $j \neq l$ , tal que

$$\mathbb{P}\left(\bigcup_{j=1}^n A_j\right) = \sum_{j=1}^n \mathbb{P}(A_j) = 1. \quad (2.40)$$

Logo, para um evento qualquer  $B$ ,  $B \subset \Omega$ , temos

$$\mathbb{P}(A_l|B) = \frac{\mathbb{P}(B|A_l)\mathbb{P}(A_l)}{\sum_{j=1}^k \mathbb{P}(B|A_j)\mathbb{P}(A_j)}, \quad (2.41)$$

para todo  $l \in \{1, \dots, n\}$  e  $\mathbb{P}(A|B)$  significa “Probabilidade de ocorrência do evento  $A$  dada a ocorrência do evento  $B$ ”. De acordo com Linden et al. (2014),  $\mathbb{P}(A_l)$  é a probabilidade *a priori* e  $\mathbb{P}(A_l|B)$  é a probabilidade *a posteriori* atualizada através dos resultados conhecidos do evento  $B$ .

Assumimos agora um vetor de dados  $\underline{Y} = (Y_1, Y_2, \dots, Y_n)$  e um conjunto  $\Theta$  de parâmetros desconhecidos. Assumimos que  $\theta_l$ , para  $l \in \{1, \dots, n\}$ , advém de uma distribuição associada

a  $Y$ . Temos que  $f(Y|\theta)$  é a verossimilhança de  $Y$  e  $\pi(\theta)$  é a distribuição *a priori* para  $\theta$ . A distribuição *a posteriori* de  $\theta_l$  dado  $Y$  é obtida por

$$\pi(\theta_l|Y) = \frac{f(Y|\theta_l)\pi(\theta_l)}{\sum_{j=1}^n f(Y|\theta_j)\pi(\theta_j)}. \quad (2.42)$$

**Exemplo 2.6.1.** Dado  $X$  um vetor de  $n$  variáveis aleatórias, cada uma delas com distribuição  $\mathcal{N}(\mu, \sigma^2)$ , para estimar seus parâmetros precisamos definir a distribuição *a priori* a ser utilizada para cada parâmetro. Sabendo que  $\mu \in (-\infty, \infty)$  e  $\sigma^2 \in (0, \infty)$ , podemos utilizar  $\pi(\mu|X) = \mathcal{U}(\{\bar{X} - 3s_X\}, \{\bar{X} + 3s_X\})$  e  $\pi(\sigma^2|X) = \mathcal{U}(0, 3s_X)$ , sendo  $\bar{X}$  e  $s_X$  os estimadores de máxima verossimilhança para  $\mu$  e  $\sigma$ , de forma a evitar influenciar o resultado com uma distribuição *a priori não informativa*.

O modelo de estimação Bayesiana apresentado no Exemplo 2.6.1 é denominado *Empirical Bayes*, cujas distribuições são estimadas utilizando de estatísticas obtidas a partir de amostras, como definido em Linden et al. (2014). Uma priori não informativa possui total ignorância em relação ao parâmetro, ou seja, é proporcional à uma constante. De acordo com Linden et al. (2014), é importante a utilização de prioris não informativas pois a informação pode conflitar com a verossimilhança, afetando e viesando os resultados.

# 3

## Estimação

Neste capítulo definimos os métodos de estimação a serem utilizados no Capítulo 4. Na Seção 3.1, definimos os métodos de estimação por máxima verossimilhança de processos vetor auto-regressivos (VAR). Em seguida, na Seção 3.2, definimos as distribuições de probabilidade utilizadas no estudo e os métodos de estimação de seus parâmetros.

### 3.1 Estimação de Processos VAR

Consideremos a equação descrita em (2.17). Assuma que o modelo VAR( $p$ ) possui estacionariedade fraca e não há restrições nos parâmetros do modelo. Cada equação do modelo pode ser escrita como

$$X_{jt} = c + \phi_{j1}X_{jt-1} + \cdots + \phi_{jp}X_{jt-p} + \varepsilon_{jt}, \text{ para todo } p \in \mathbb{N} - \{0\}; j = 1, \dots, k \text{ e } t = 1, \dots, T, \quad (3.1)$$

onde  $X$  é uma matriz com  $T$  observações de  $k$  conjuntos de dados,  $\phi_j$  é um vetor de parâmetros de tamanho  $p$  e  $\varepsilon_{jt}$  é o erro, com tamanho  $T$  e matriz de covariância  $\sigma^2 I_T$ . Como cada equação possui as mesmas variáveis explicativas, elas podem ser estimadas separadamente pelo método dos mínimos quadrados ordinários sem perda de eficiência em relação ao método generalizado.

A seguir definimos o método dos mínimos quadrados para  $X_{jt}$ , quando  $j = 1$  e  $p = 1$ .

**Observação 3.1.1.** Considerando a equação

$$X_{1t} = c_1 + \phi_{11}X_{1t-1} + \phi_{12}X_{2t-1} + \varepsilon_{1t}, \text{ para todo } t = 1, \dots, T, j = 1 \text{ e } p = 1. \quad (3.2)$$

Definimos

$$X_1 = \begin{bmatrix} 1 & X_{10} & X_{20} \\ 1 & X_{11} & X_{21} \\ \vdots & \vdots & \vdots \\ 1 & X_{1T-1} & X_{2T-1} \end{bmatrix}, \text{ matriz } (T \times 3),$$

$$\pi_1 = [c_1, \phi_{11}, \phi_{12}]',$$

$$\varepsilon_1 = [\varepsilon_{1t}, \dots, \varepsilon_{1T}]',$$

Logo, conforme Morettin e Tolo (2004),  $X = X_{11}\pi_1 + \varepsilon_1$  e o estimador de mínimos quadrados é dado por

$$\hat{\pi}_1 = (X'X)^{-1}X'X_1. \quad (3.3)$$

Considerando a segunda equação da (2.16), podemos repetir o mesmo processo e obter a equação para o estimador  $\hat{\pi}_2$ .

## 3.2 Estimação de Parâmetros

Nesta seção estudamos as três distribuições utilizadas neste trabalho, e a estimação de seus respectivos parâmetros. Iniciamos na Subseção 3.2.1 com a Distribuição Gaussiana, a distribuição mais utilizada em séries temporais. Em seguida, na Subseção 3.2.2, vemos a distribuição *t*-Student, uma distribuição com características semelhantes à Gaussiana, mas com caudas mais pesadas. Por fim, na Subseção 3.2.3, temos a distribuição  $\alpha$ -Estável, o objetivo deste trabalho.

### 3.2.1 Distribuição Gaussiana

**Definição 3.2.1.** Uma variável é dita *Normal* (ou *Gaussiana*) quando sua função característica segue a forma

$$\varphi_X(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}} \quad (3.4)$$

e seus parâmetros  $\mu$  e  $\sigma^2$  representam a média e a variância da variável, respectivamente.

**Definição 3.2.2.** Dada uma amostra de dados  $\underline{X} = (X_1, \dots, X_n)$  advinda de uma  $\mathcal{N}(\mu, \sigma^2)$ , os estimadores de máxima verossimilhança para  $\mu$  e  $\sigma^2$  são definidos como

$$\hat{\mu} = \frac{\sum_{j=1}^n X_j}{n} \quad \text{e} \quad \hat{\sigma}^2 = \sum_{j=1}^n \frac{(X_j - \bar{X})^2}{n-1}. \quad (3.5)$$

### 3.2.2 Distribuição *t*-Student

**Definição 3.2.3.** Uma variável  $X$  advém da distribuição *t*-Student se, e somente se, sua função de densidade pode ser escrita da seguinte forma

$$f_X(x) = \frac{\Gamma\left(\frac{gl+1}{2}\right)}{\sqrt{\pi gl} \Gamma\left(\frac{gl}{2}\right)} \left(1 + \frac{x^2}{gl}\right)^{-\frac{gl+1}{2}}, \quad (3.6)$$

onde  $gl$  é o número de graus de liberdade e  $\Gamma(\cdot)$  é a função Gamma dada por  $\Gamma(x) = x\Gamma(x-1)$ .

### 3.2.3 Distribuição $\alpha$ -Estável

Definimos uma variável  $\alpha$ -Estável na Seção 2.2. Os métodos clássicos de estimação foram definidos em McCulloch (1986) e expandidos para processos auto-regressivos de média móvel (ARMA) por Adler et al. (1998).

**Lema 3.2.1.** *Suponhamos  $n$  amostras aleatórias de uma distribuição  $\alpha$ -Estável  $S_\alpha(\beta, \sigma, \delta)$ , onde  $\sigma$  e  $\delta$  são conhecidos e fixos. Seja  $X([q])$  o  $q$ -ésimo quantil populacional e  $\hat{X}([q])$  o quantil amostral correspondente. Então  $\hat{X}([q])$  é um estimador consistente para  $X([q])$ .*

**Prova:** Ver McCulloch (1986).

Definimos

$$\begin{aligned}\hat{v}_\alpha &= (\hat{X}([95]) - \hat{X}([5]) / (\hat{X}([75]) - \hat{X}([25]))) \\ \hat{v}_\beta &= (\hat{X}([95]) + \hat{X}([5]) - 2\hat{X}([50])) / (\hat{X}([95]) - \hat{X}([5])),\end{aligned}\tag{3.7}$$

onde  $\hat{X}([q])$  corresponde ao  $q$ -ésimo quantil amostral.

Ambos  $\hat{v}_\alpha$  e  $\hat{v}_\beta$  são estatísticas consistentes dos índices  $v_\alpha = \phi_1(\alpha, \beta)$  e  $v_\beta = \phi_2(\alpha, \beta)$ , respectivamente. Estes são independentes dos valores de  $\sigma$  e  $\delta$ .

As relações  $v_\alpha = \phi_1(\alpha, \beta)$  e  $v_\beta = \phi_2(\alpha, \beta)$  podem ser invertidas de modo que

$$\alpha = \psi_1(v_\alpha, v_\beta) \text{ e } \beta = \psi_2(v_\alpha, v_\beta),\tag{3.8}$$

onde  $\psi$  é a função log-característica das distribuições  $\alpha$ -Estáveis, para  $\alpha \neq 1$ , dada por

$$\psi(t) = i\delta t - |\sigma t|^\alpha \exp\left\{-i\beta \frac{\pi}{2} \min(\alpha, 2-\alpha) \text{sign}(t)\right\}.$$

Podemos, assim, estimar consistentemente  $\alpha$  e  $\beta$  através de

$$\hat{\alpha} = \psi_1(\hat{v}_\alpha, \hat{v}_\beta) \text{ e } \hat{\beta} = \psi_2(\hat{v}_\alpha, \hat{v}_\beta).\tag{3.9}$$

Adler et al. (1998) demonstram que estes estimadores apresentam os melhores resultados em um processo ARMA quando aplicados diretamente na série. Aqui, neste trabalho, utilizamos estes estimadores, dados em (3.9), para os processos VAR(1) no Capítulo 4.

## 4

# Estimações Clássica e Bayesiana de um VAR(1) Bidimensional

Nesta capítulo tratamos das estimações dos parâmetros de um *processo VAR(1) bidimensional* quando as inovações forem Gaussianas, *t*-Student e  $\alpha$ -Estável. Utilizamos o software *R* e os pacotes *mvtnorm* e *alphastable* para simulação das distribuições multivariadas, *portes* para simulação de processos VAR e *vars* para a estimação.

## 4.1 Estimação Clássica

Nesta seção, apresentamos resultados provenientes de estimações utilizando métodos clássicos, como definidos na Seção 2.4. A partir destes e das séries definidas anteriormente, analisamos a veracidade e eficiência destes métodos quando aplicados a dados com distribuições Gaussianas, *t*-Student e  $\alpha$ -Estável.

Para todos os casos, geramos amostras de tamanho  $T = 10,000$  e  $\phi = \begin{bmatrix} 0.2 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$ , seguindo a expressão (2.17) onde definimos um processo VAR(1) bidimensional.

### 4.1.1 Inovações Gaussianas

Iniciamos com inovações Gaussianas e geramos um processo VAR(1), conforme expressão (2.17), com  $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$  onde

$$\sigma^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{ com } t \in 1, \dots, T, \quad T = 10,000 \text{ e } re = 1,000. \quad (4.1)$$

Utilizando as equações (3.5) e (3.3), a Tabela 4.1 abaixo fornece as estimações para o vetor de parâmetros  $\boldsymbol{\eta} = (\mu, \sigma, \phi_1, \phi_2)$  com suas respectivas médias (*média*), desvios padrões (*dp*), erros quadráticos médios (*eqm*) e vícios (*vício*), obtidas através de simulações de Monte Carlo, ao

invés da distribuição assintótica dos estimadores. Ressaltamos que, neste trabalho, a matriz de variâncias-covariâncias é dada em (4.1) e o vetor das médias  $\mu$  é  $(0,0)$ .

Tabela 4.1: Resultado das estimações em  $re = 1,000$  simulações de Monte Carlo de processos VAR(1) bidimensional com inovações Gaussianas quando  $T = 10,000$ ,  $\mu = (0,0)$ ,  $\sigma^2$  é dado pela expressão (4.1),  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ .

Componente	Estatística	$\mu_1 = 0$	$\sigma_1 = 1$	$\phi_{11} = 0.2$	$\phi_{12} = 0.1$
1	<i>média</i>	-0.0001	1.0006	0.2000	0.0997
	<i>dp</i>	0.0129	0.0132	0.0097	0.0095
	<i>eqm</i>	0.0129	0.9994	0.0097	0.0095
	<i>vício</i>	0.0001	0.0006	0.0000	0.0002
	Estatística	$\mu_2 = 0$	$\sigma_2 = 1$	$\phi_{21} = 0.1$	$\phi_{22} = 0.2$
2	<i>média</i>	0.0271	1.0267	0.0999	0.1995
	<i>dp</i>	0.0077	0.0072	0.0096	0.0097
	<i>eqm</i>	0.0282	0.0276	0.0096	0.0098
	<i>vício</i>	0.0271	0.0267	0.0001	0.0005

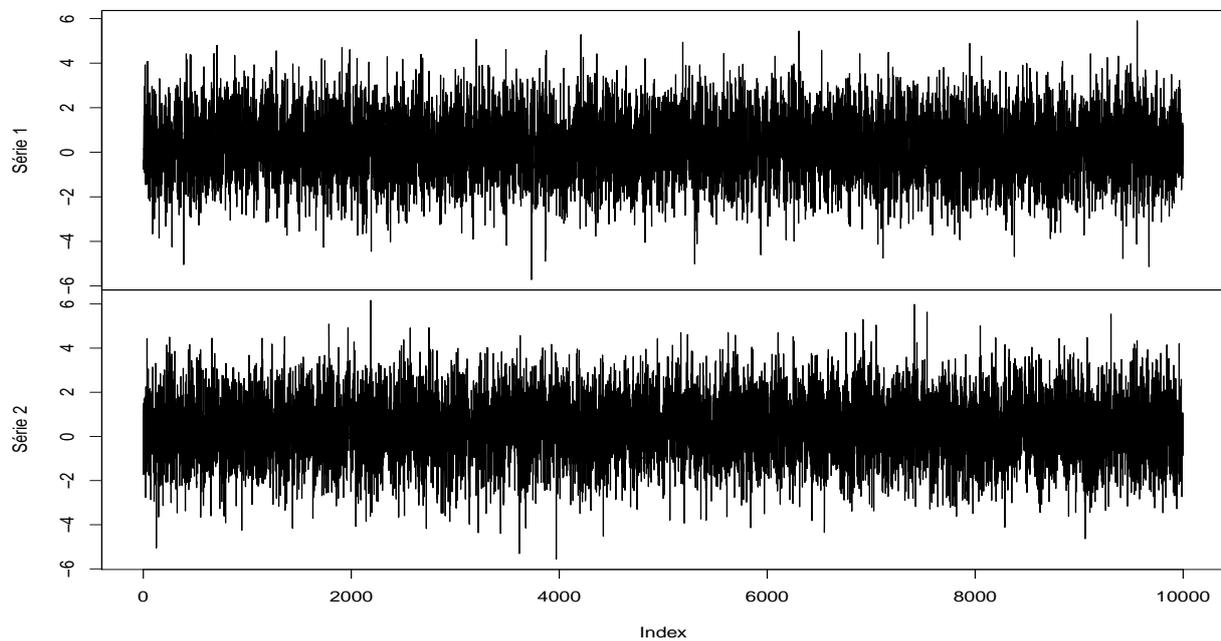


Figura 4.1: Exemplo de séries geradas de um VAR(1) bidimensional definido em (2.17) com inovações Gaussianas quando  $T = 10,000$ ,  $\mu = (0,0)$ ,  $\sigma^2$  é dado pela expressão (4.1),  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ .

### 4.1.2 Inovações $t$ -Student

Geramos um VAR(1), conforme expressão (2.17), com  $\varepsilon_t \sim t(gl, \sigma^2)$  onde  $gl = 4$  e  $\sigma^2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ ,  $T = 10,000$  e  $re = 1,000$ .

A Tabela 4.2 fornece as estimações para o vetor de parâmetros  $\boldsymbol{\eta} = (gl, \phi_1, \phi_2)$  com suas respectivas médias (*média*), desvios padrões (*dp*), erros quadráticos médios (*eqm*) e vícios (*vício*), obtidas através de simulações de Monte Carlo, ao invés da distribuição assintótica dos estimadores.

Tabela 4.2: Resultado das estimações em  $re = 1,000$  simulações de Monte Carlo de processos VAR(1) bidimensional com inovações  $t$ -Student, quando  $T = 10,000$ ,  $gl = (4, 4)$ ,  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ .

Componente	Estatística	$gl_1 = 4$	$\phi_{11} = 0.2$	$\phi_{21} = 0.1$
1	<i>média</i>	4.3053	0.2001	0.0997
	<i>dp</i>	0.1936	0.0100	0.0097
	<i>eqm</i>	0.3615	0.0100	0.0097
	<i>vício</i>	0.3053	0.0001	0.0002
	Estatística	$gl_2 = 4$	$\phi_{21} = 0.1$	$\phi_{22} = 0.2$
2	<i>média</i>	4.3052	0.0999	0.2001
	<i>dp</i>	0.1821	0.0100	0.0095
	<i>eqm</i>	0.3554	0.0100	0.0095
	<i>vício</i>	0.3052	0.0001	0.0001

### 4.1.3 Inovações $\alpha$ -Estáveis

Geramos mil processos VAR(1) com  $\varepsilon_t \sim S_\alpha(\beta, \sigma, \delta)$  onde  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $T = 10,000$  e  $re = 1,000$ . Utilizando as equações (3.9) e (3.3), a Tabela 4.3 a seguir fornece as estimações para o vetor de parâmetros  $\boldsymbol{\eta} = (\alpha, \beta, \phi_1, \phi_2)$  com suas respectivas médias (*média*), desvios padrões (*dp*), erros quadráticos médios (*eqm*) e vícios (*vício*), obtidas através de simulações de Monte Carlo, ao invés da distribuição assintótica dos estimadores.

A Figura 4.6 abaixo apresenta a função codiferença para duas séries temporais geradas de um VAR(1) bidimensional definido em (2.17), com  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ . Observe que a Componente 1 apresenta dependências significativas nos lags 1, 10 e 19 enquanto que a Componente 2 apresenta dependências significativas nos lags 1 e 15.

Observando apenas uma simulação do processo VAR(1) bidimensional com inovações  $\alpha$ -Estáveis,  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ , a Figura 4.7 mostra a codiferença cruzada para diversos valores de  $k$ . Observamos uma significância de 5% dos valores para  $|k| \leq 2$ , com tendência decrescente.

Tabela 4.3: Resultado das estimações clássicas em  $re = 1,000$  simulações de Monte Carlo de processos VAR(1) bidimensional com inovações  $\alpha$ -Estáveis, quando  $T = 10,000$ ,  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ .

Componente	Estatística	$\alpha_1 = 1.3$	$\beta_1 = 0$	$\phi_{11} = 0.2$	$\phi_{21} = 0.1$
1	<i>média</i>	1.3074	0.0979	0.1998	0.0992
	<i>dp</i>	0.0166	0.1414	0.0065	0.0262
	<i>eqm</i>	0.0181	0.1719	0.0065	0.0262
	<i>vício</i>	0.0074	0.0978	0.0002	0.0007
	Estatística	$\alpha_2 = 1.3$	$\beta_1 = 0$	$\phi_{21} = 0.1$	$\phi_{22} = 0.2$
2	<i>média</i>	1.3076	0.0887	0.0997	0.2001
	<i>dp</i>	0.0165	0.1335	0.0094	0.0086
	<i>eqm</i>	0.0181	0.1602	0.0094	0.0086
	<i>vício</i>	0.0076	0.0887	0.0003	0.0001

Tabela 4.4: Resultado das estimações da função de codiferença cruzada em processos VAR(1) bidimensional com inovações  $\alpha$ -Estáveis, quando  $T = 10,000$ ,  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $\phi_1 = (0.2, 0.1)$ ,  $\phi_2 = (0.1, 0.2)$  e  $k \in \{-2, -1, 0, 1, 2\}$ .

Estatística	$k = -2$	$k = -1$	$k = 0$	$k = 1$	$k = 2$
<i>média</i>	0.0990	0.3010	0.3310	0.2980	0.0980
<i>dp</i>	0.0640	0.0560	0.0570	0.0580	0.0680

Através da Tabela 4.4 e da Figura 4.7, podemos ver uma maior dependência dos dados quando tratamos dos lags  $|k| \leq 1$ , como espera-se de um processo VAR bidimensional de primeira ordem.

## 4.2 Estimação Bayesiana

Nesta seção, apresentamos resultados provenientes de estimações utilizando métodos Bayesianos, como definidos na Seção 2.6, por meio de *Amostrador de Gibbs* (Linden et al., 2014). A partir das séries definidas anteriormente, analisamos a veracidade e eficiência destes métodos quando aplicados a dados com distribuições Gaussiana,  $t$ -Student e  $\alpha$ -Estável.

Para todos os casos, geramos amostras de tamanho  $T = 10,000$  e  $\phi = \begin{bmatrix} 0.2 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$ , seguindo a expressão (2.17) onde definimos um processo VAR(1) bidimensional. Além disto, as distribuições a priori para  $\phi_{jk}$ ,  $j$  e  $k$  no conjunto  $\{1, 2\}$ , denominadas  $\pi(\phi_{jk})$ , são distribuições Gaussianas de média 0 e precisão 0.1, o *burn-in*, quantidade de observações descartadas previamente à convergência da cadeia, é igual a 100 e todas as cadeias aqui estimadas apresentaram convergência.

### 4.2.1 Inovações Gaussianas

Geramos um processo VAR(1) com  $\varepsilon_t \sim \mathcal{N}(\mu, \sigma^2)$  onde  $\mu = (0, 0)$ ,  $\sigma^2 = (1, 1)$ ,  $T = 10,000$ . Utilizando métodos Bayesianos de estimação onde  $\pi(\mu)$  é uma distribuição Gaussiana com média 0 e precisão 0.1 e  $\pi(\sigma^2)$  é uma distribuição Uniforme no intervalo  $[0, 100]$ , a Tabela 4.5 abaixo fornece as estimações para o vetor de parâmetros  $\eta = (\mu, \sigma^2, \phi_1, \phi_2)$  com suas respectivas médias (*média*), desvios padrões (*dp*) e vícios (*vício*), obtidas através de simulações de Monte Carlo.

Tabela 4.5: Resultado das estimações Bayesianas em processos VAR(1) bidimensional com inovações Gaussianas, quando  $T = 10,000$ ,  $\mu = (0, 0)$ ,  $\sigma^2 = (1, 1)$ ,  $\phi_1 = (0.2, 0.1)$ ,  $\phi_2 = (0.1, 0.2)$  e  $re = 500$ .

Componente	Estatística	$\mu_1 = 0$	$\sigma_1^2 = 1$	$\phi_{11} = 0.2$	$\phi_{12} = 0.1$
1	<i>média</i>	0.0141	1.0438	0.1898	0.0755
	<i>dp</i>	0.0186	0.0264	0.0181	0.0174
	<i>vício</i>	0.0141	0.0438	0.0102	0.0245
	Estatística	$\mu_2 = 0$	$\sigma_2^2 = 1$	$\phi_{21} = 0.1$	$\phi_{22} = 0.2$
2	<i>média</i>	-0.0071	1.0493	0.1045	0.2062
	<i>dp</i>	0.0192	0.0270	0.0179	0.0174
	<i>vício</i>	0.0071	0.0493	0.0045	0.0062

Observando ambas as Tabelas 4.1 e 4.5 temos que os vícios e desvios padrões das estimações dos métodos clássicos foram inferiores aos dos métodos Bayesianos.

### 4.2.2 Inovações *t*-Student

Geramos um processo VAR(1) com  $\varepsilon_t \sim t(gl)$  onde  $gl = (4, 4)$ ,  $T = 10,000$ . Utilizando métodos Bayesianos de estimação onde  $\pi(gl)$  é uma distribuição Exponencial com parâmetro de taxa 0.0333, a Tabela 4.6 abaixo fornece as estimações para o vetor de parâmetros  $\eta = (gl, \phi_1, \phi_2)$  com suas respectivas médias (*média*), desvios padrões (*dp*) e vícios (*vício*), obtidas através de simulações de Monte Carlo.

Observando ambas as Tabelas 4.2 e 4.6 temos que os vícios do método Bayesiano se apresentaram, em média, inferiores ao do método clássico quando tratamos dos graus de liberdade, apesar de seus desvios serem superiores. Quando observamos, no entanto, os resultados das estimações do processo VAR(1), a estimação clássica apresenta menores vício e desvio padrão.

### 4.2.3 Inovações $\alpha$ -Estáveis

Geramos um processo VAR(1) com  $\varepsilon_t \sim S_\alpha(\beta, \sigma, \delta)$  onde  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $T = 10,000$ . Utilizando métodos Bayesianos de estimação onde  $\pi(\alpha)$  é

Tabela 4.6: Resultado das estimações Bayesianas em processos VAR(1) bidimensional com inovações  $t$ -Student, quando  $T = 10,000$ ,  $gl = (4, 4)$ ,  $\phi_1 = (0.2, 0.1)$ ,  $\phi_2 = (0.1, 0.2)$  e  $re = 500$ .

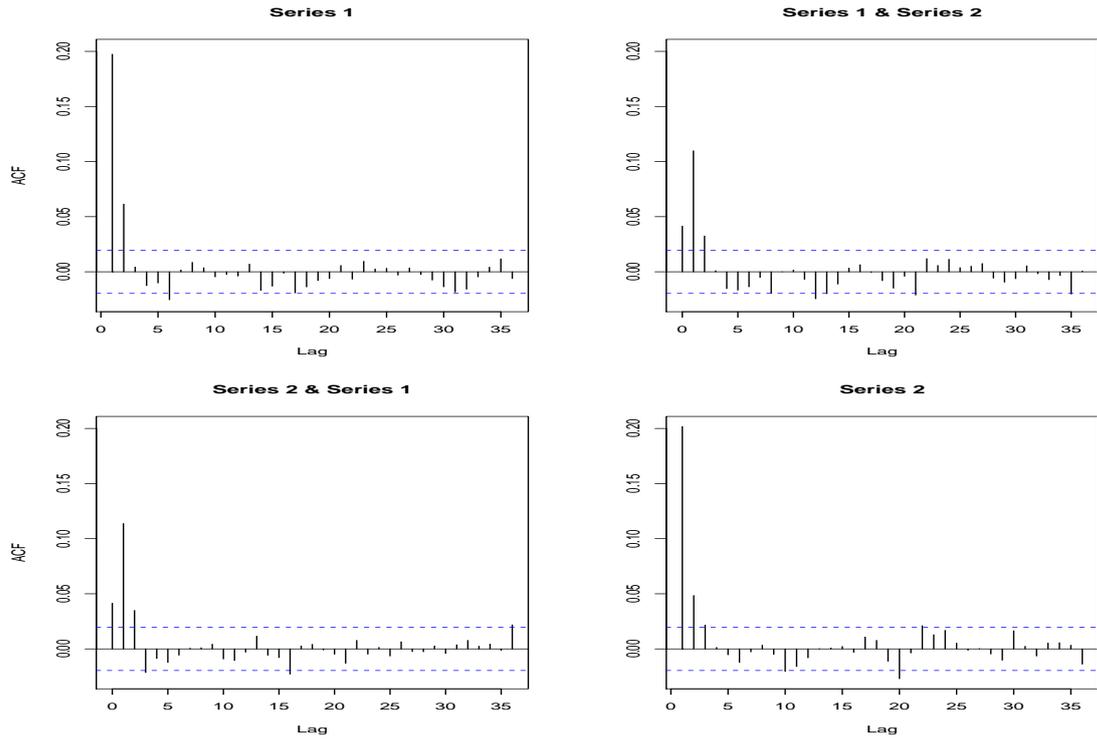
Componente	Estatística	$gl_1 = 4$	$\phi_{11} = 0.2$	$\phi_{12} = 0.1$
1	<i>média</i>	3.8859	0.2128	0.1090
	<i>dp</i>	0.2957	0.0183	0.0185
	<i>vício</i>	0.1141	0.0128	0.0090
	Estatística	$gl_2 = 4$	$\phi_{21} = 0.1$	$\phi_{22} = 0.2$
2	<i>média</i>	4.1031	0.0775	0.1935
	<i>dp</i>	0.3363	0.0185	0.0180
	<i>vício</i>	0.1031	0.0225	0.0065

uma distribuição Gaussiana com média 1.3 e precisão 1 truncada no intervalo  $(1, 2)$  e  $\pi(\beta)$  é uma distribuição Gaussiana com média 0 e precisão 1, truncada no intervalo  $(-1, 1)$ , a Tabela 4.7 abaixo fornece as estimações para o vetor de parâmetros  $\eta = (\alpha, \beta, \phi_1, \phi_2)$  com suas respectivas médias (*média*), desvios padrões (*dp*) e vícios (*vício*), obtidas através de simulações de Monte Carlo.

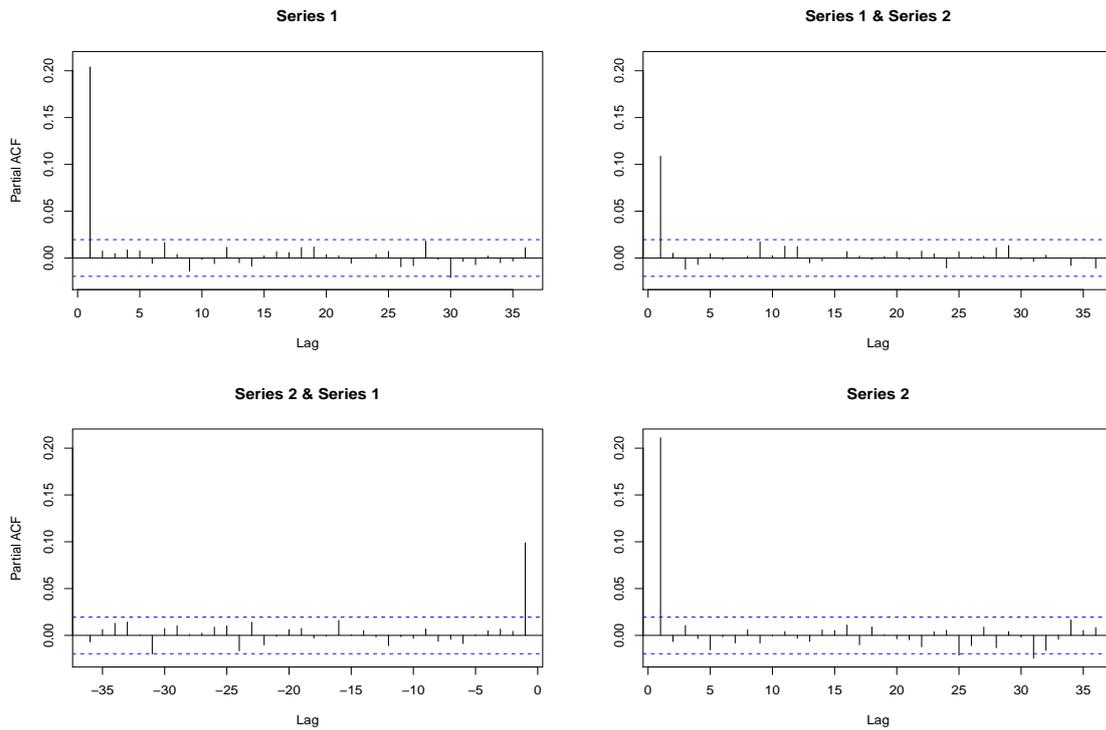
Tabela 4.7: Resultado das estimações Bayesianas em processos VAR(1) bidimensional com inovações  $\alpha$ -Estáveis, quando  $T = 10,000$ ,  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $\phi_1 = (0.2, 0.1)$ ,  $\phi_2 = (0.1, 0.2)$  e  $re = 500$ .

Componente	Estatística	$\alpha_1 = 1.3$	$\beta_1 = 0$	$\phi_{11} = 0.2$	$\phi_{12} = 0.1$
1	<i>média</i>	1.3453	-0.2159	0.1824	0.0987
	<i>dp</i>	0.0055	0.0193	0.0375	0.0130
	<i>vício</i>	0.0453	0.2159	0.0176	0.0013
	Estatística	$\alpha_2 = 1.3$	$\beta_1 = 0$	$\phi_{21} = 0.1$	$\phi_{22} = 0.2$
2	<i>média</i>	1.3226	-0.2102	0.0877	0.2020
	<i>dp</i>	0.0052	0.0185	0.0366	0.0130
	<i>vício</i>	0.0226	0.2102	0.0123	0.0020

Observamos através das Tabelas 4.3 e 4.7 que, para todas as estimações, o vício do método Bayesiano é superior ao do método clássico e os desvios padrões são menores para a maioria das estimações Bayesianas. Assim, podemos dizer que o método clássico obteve melhor resultado.



(a)



(b)

Figura 4.2: Funções de Autocovariância (FAC) e de Autocovariância Parcial (FACP) e seus intervalos a 95% de confiança, em azul, para duas séries Gaussianas quando  $\mu = (0, 0)$ ,  $\sigma^2$  é dado pela expressão (4.1),  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ : (a) FAC; (b) FACP.

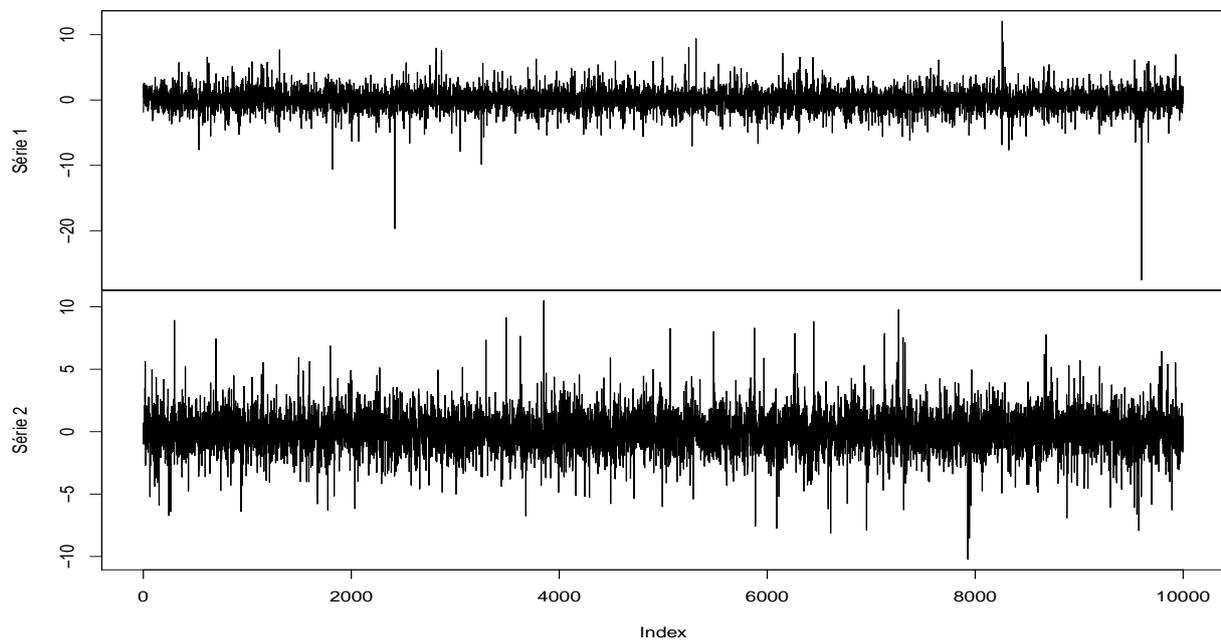
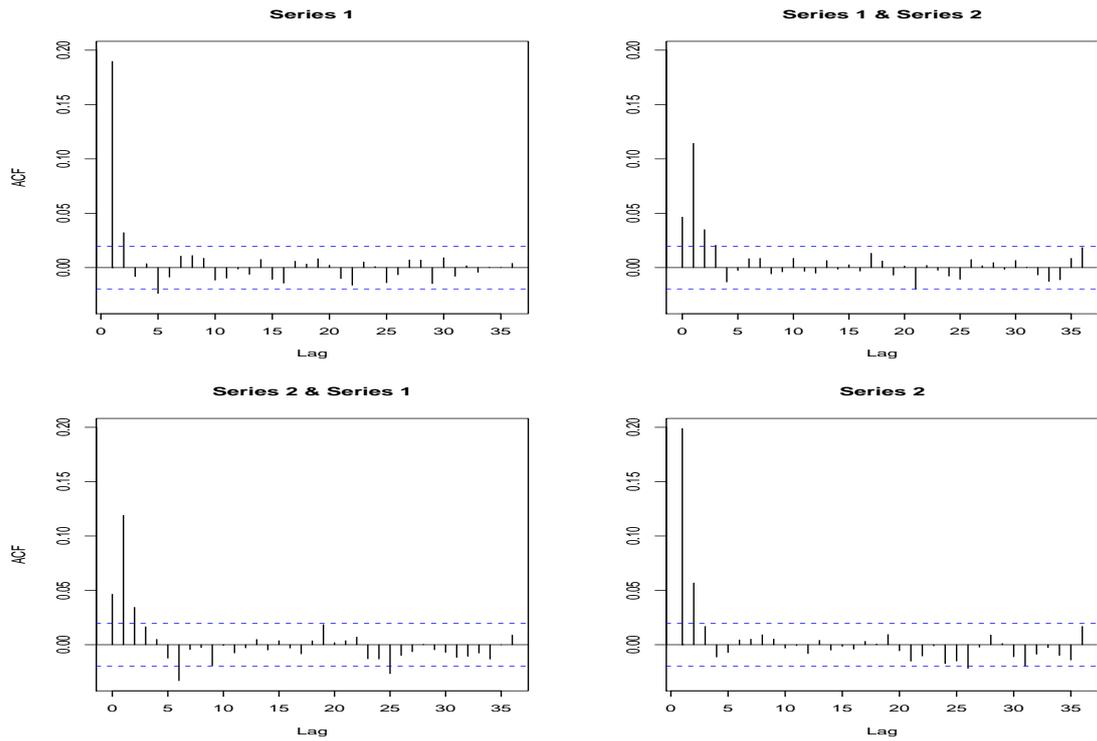
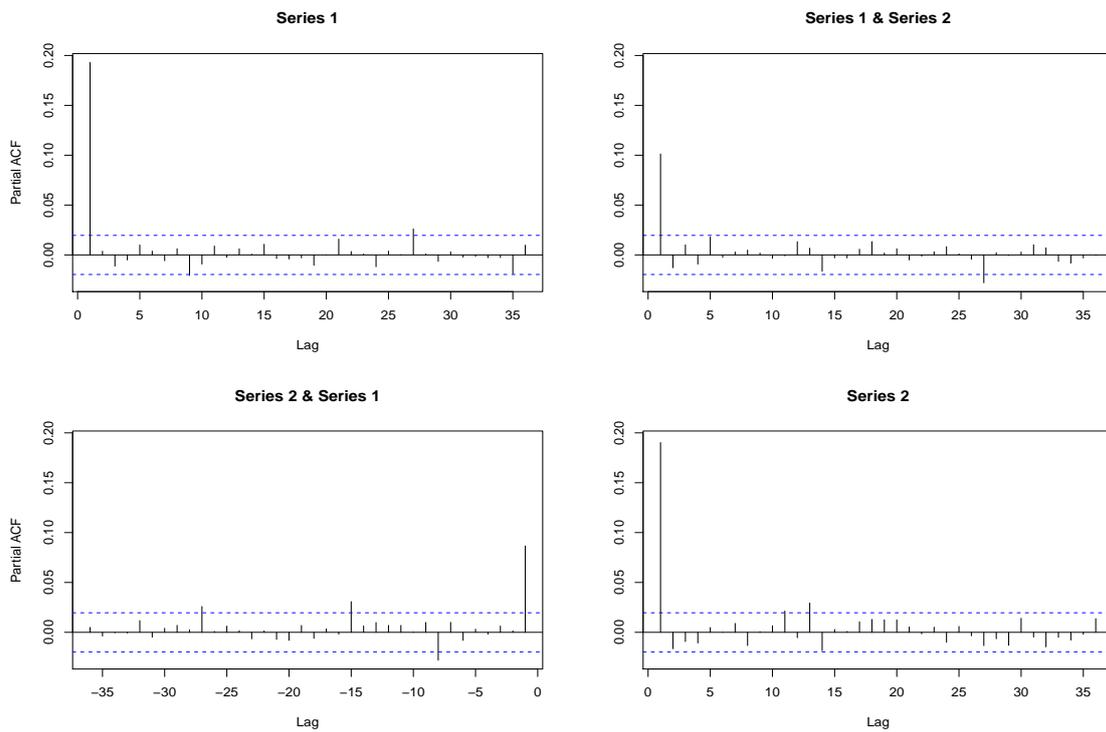


Figura 4.3: Exemplo de séries geradas de um VAR(1) bidimensional definido em (2.17) com inovações  $t$ -Student, onde  $T = 10.000$ ,  $gl = (4, 4)$ ,  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ .



(a)



(b)

Figura 4.4: Funções de Autocovariância (FAC) e de Autocovariância Parcial (FACP) e seus intervalos a 95% de confiança, em azul, para duas séries *t*-Student, quando  $T = 10,000$ ,  $gl = (4, 4)$ ,  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ : (a) FAC; (b) FACP.

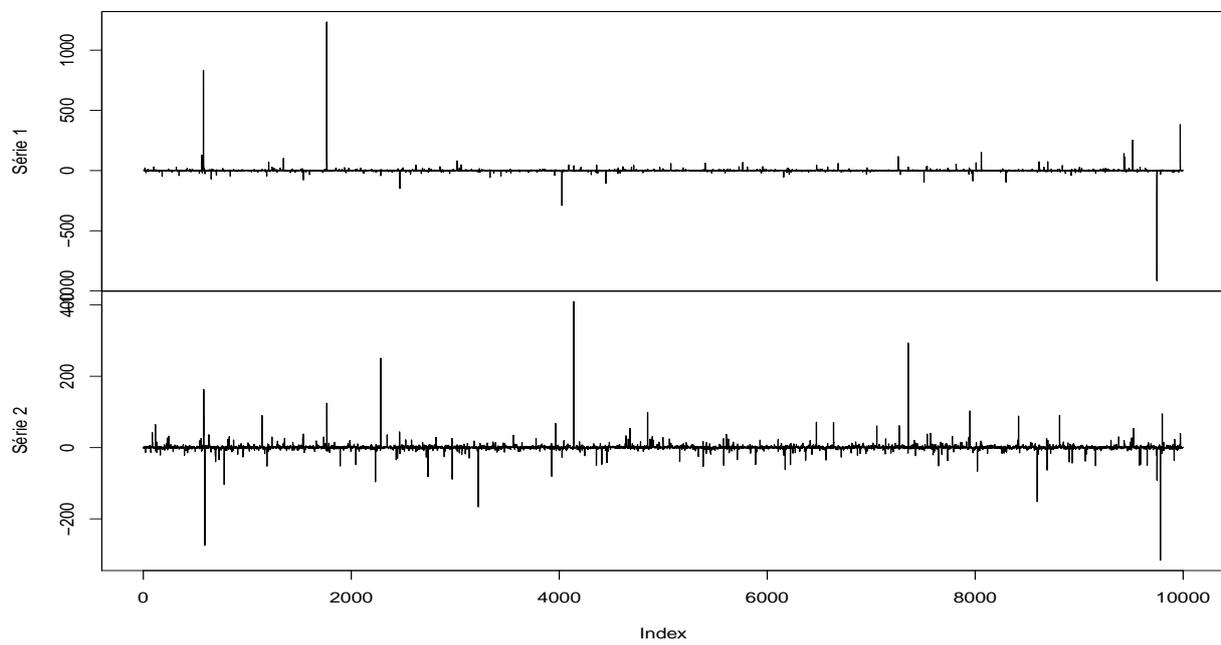
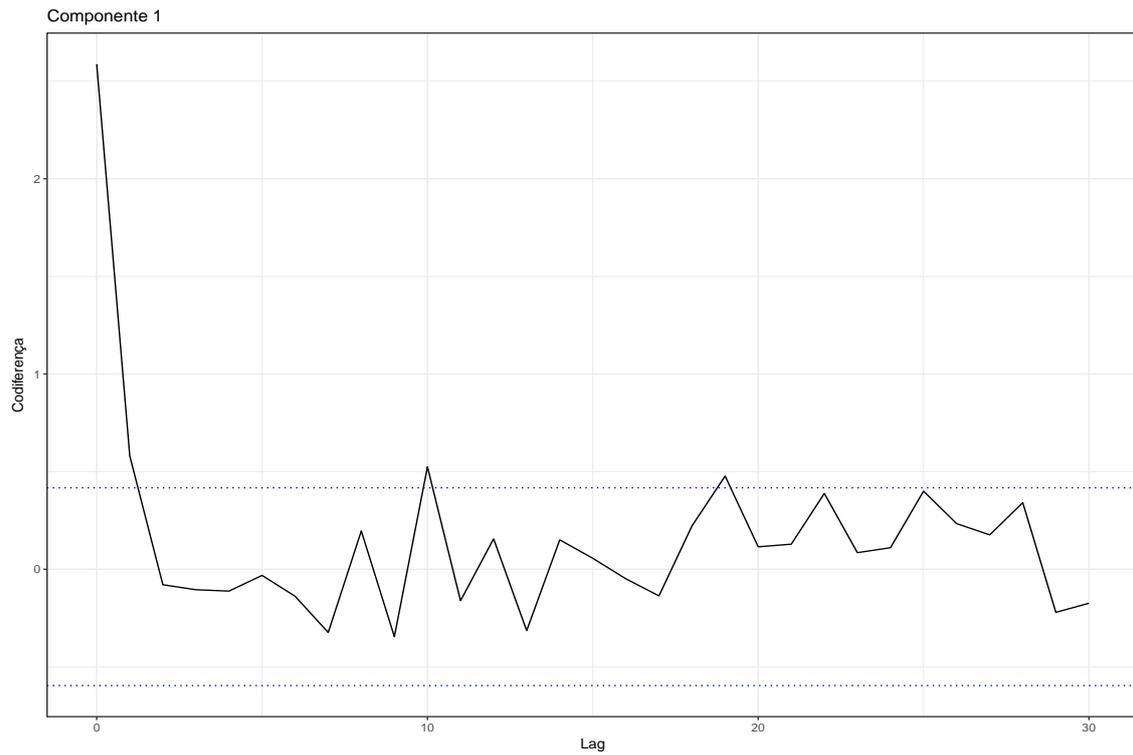
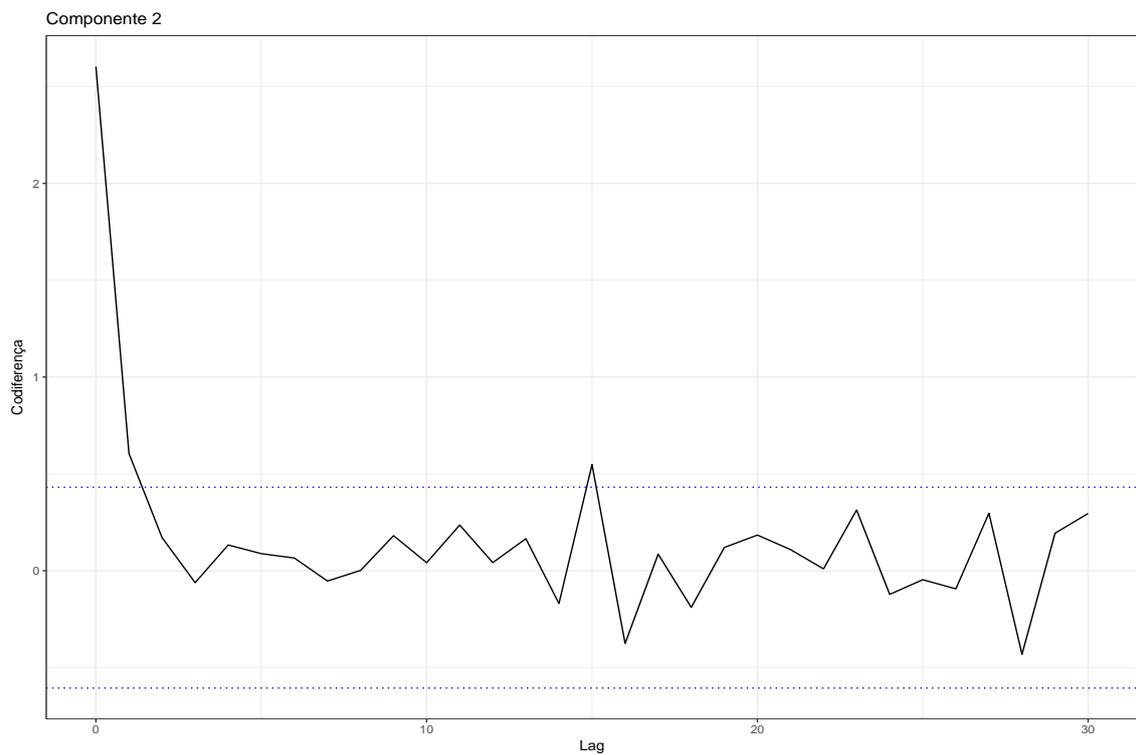


Figura 4.5: Exemplo de séries geradas de um VAR(1) bidimensional definido em (2.17) com inovações  $\alpha$ -Estáveis, onde  $T = 10,000$ ,  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ .



(a)



(b)

Figura 4.6: Funções de Codiferença e seus intervalos a 95% de confiança, em azul, para duas séries  $\alpha$ -Estáveis quando  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ : (a) Codiferença para a Componente 1; (b) Codiferença para a Componente 2.

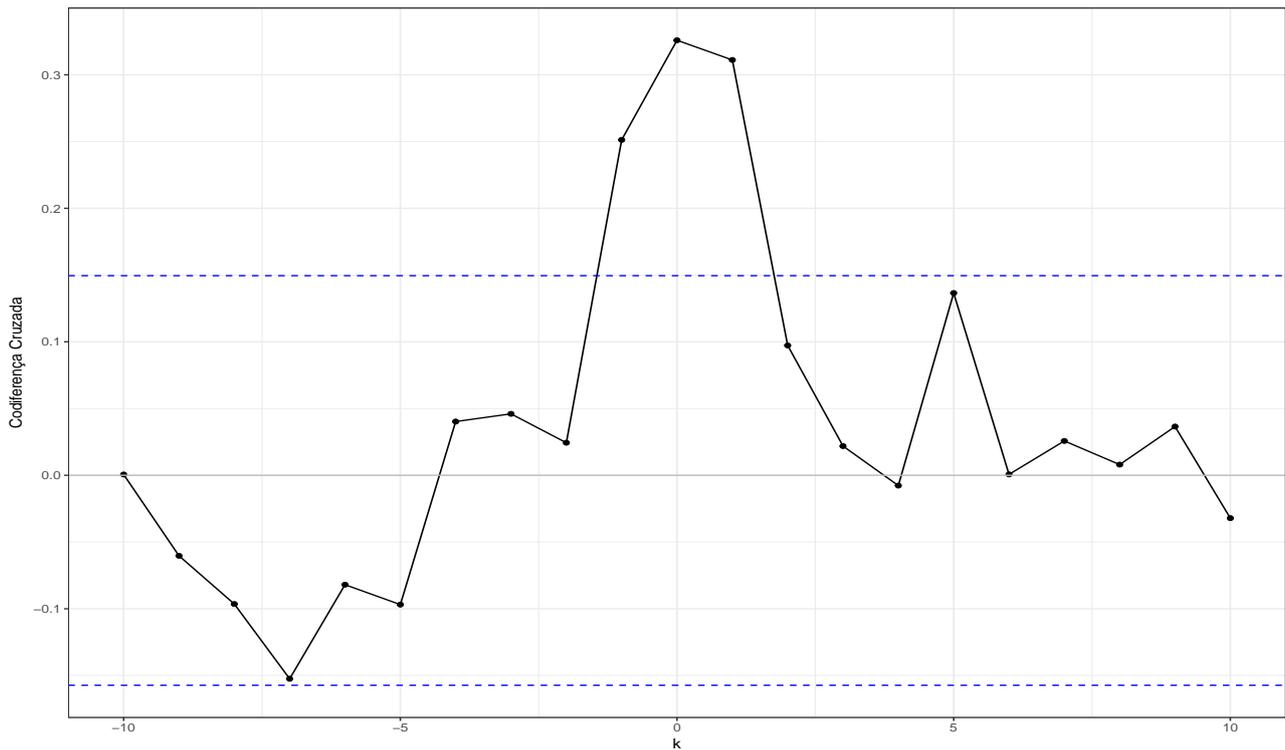


Figura 4.7: Codiferença Cruzada para  $k$  entre -10 e 10 e seu intervalo a 95% de confiança, em azul, de um conjunto de séries geradas de um VAR(1) bidimensional definido em (2.17), com inovações  $\alpha$ -Estáveis, quando  $T = 10,000$ ,  $\alpha = (1.3, 1.3)$ ,  $\beta = (0, 0)$ ,  $\delta = (0, 0)$ ,  $\sigma = (0.3, 0.2)$ ,  $\phi_1 = (0.2, 0.1)$  e  $\phi_2 = (0.1, 0.2)$ .

# 5

## Aplicação a Dados Reais

Neste Capítulo trabalhamos com os métodos apresentados no Capítulo 3 aplicados em dados reais. A intenção é comparar, novamente, os métodos clássico e Bayesiano na estimação dos parâmetros dos modelos.

### 5.1 Apresentação dos Dados

Utilizamos dados diários das ações de duas empresas com produtos semelhantes, a *Coca-Cola Company* (*ko*) e a *PepsiCo Inc* (*pep*). Os dados foram obtidos pelo Yahoo Finance (<https://finance.yahoo.com/>), através da biblioteca de funções financeiras *ffn*, disponível em Python.

As duas séries temporais iniciam em 01 de Janeiro de 2015, com informações para todos os dias úteis até 29 de Outubro de 2019. Isso significa um total de 1,215 observações através de um período de quase quatro anos. Por se tratar de produtos semelhantes, observamos uma correlação de aproximadamente 0.9487. A Figura 5.1 apresenta as séries originais.

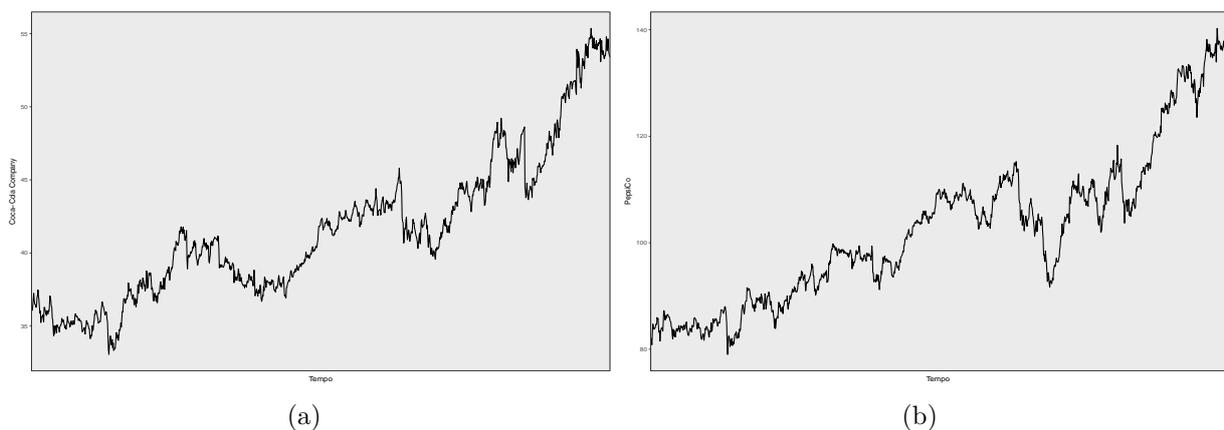


Figura 5.1: Valores diários das ações das duas empresas analisadas no período de 01 de janeiro de 2015 a 29 de outubro de 2019, com  $n = 1,215$  observações: (a) Coca-Cola Company; (b) PepsiCo Inc.

São séries com clara tendência crescente e possível sazonalidade devido as suas características de origem. Como os métodos estudados são para processos auto-regressivos estacionários, aplicamos uma primeira diferença, ou seja, transformamos os dados tais que

$$Y_t = X_t - X_{t-1} \quad (5.1)$$

onde  $X_t$  é a série original e  $Y_t$  possui  $n - 1$  observações.

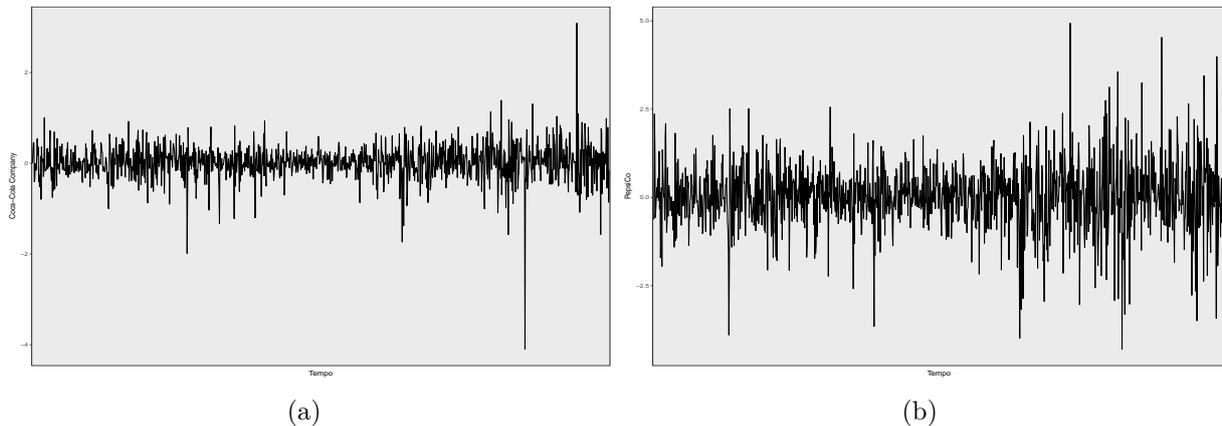


Figura 5.2: Séries temporais resultantes de uma primeira diferença: (a) Coca-Cola Company; (b) PepsiCo Inc.

A série temporal apresentada na Figura 5.2(a) possui picos típicos de uma distribuição  $\alpha$ -Estável enquanto que a série temporal apresentada na Figura 5.2(b) possui picos maiores e mais frequentes. Mesmo com este processo de uma primeira diferença, a correlação entre as séries se mantém alta, com valor de 0.6948. No entanto, observamos que as duas componentes, resultantes de uma primeira diferença, não possuem mais tendência crescente.

## 5.2 Análise e Estimação

Para a criação de um Q-Q Plot, realizamos a estimação dos parâmetros da série utilizando o método clássico de McCulloch (1986), definido no Capítulo 3. Obtivemos as estimativas para  $Y_t$  de  $\hat{\alpha}_{Coca} = 1.5311$  e  $\hat{\alpha}_{Pepsi} = 1.6566$ , que parecem nos indicar que uma função de distribuição  $\alpha$ -Estável se adequa melhor aos dados do que uma Gaussiana ou uma  $t$ -Student.

Iniciamos com uma análise descritiva dos dados para verificar se o modelo VAR(1) bidimensional com inovações  $\alpha$ -Estáveis é o melhor entre os modelos estudados. Para isso, utilizamos os testes de adequação de modelos, apresentados na Seção 2.5.

Observamos, a partir das Figuras 5.3 e 5.4, que os dados parecem se adequar melhor a uma distribuição não-Gaussiana.

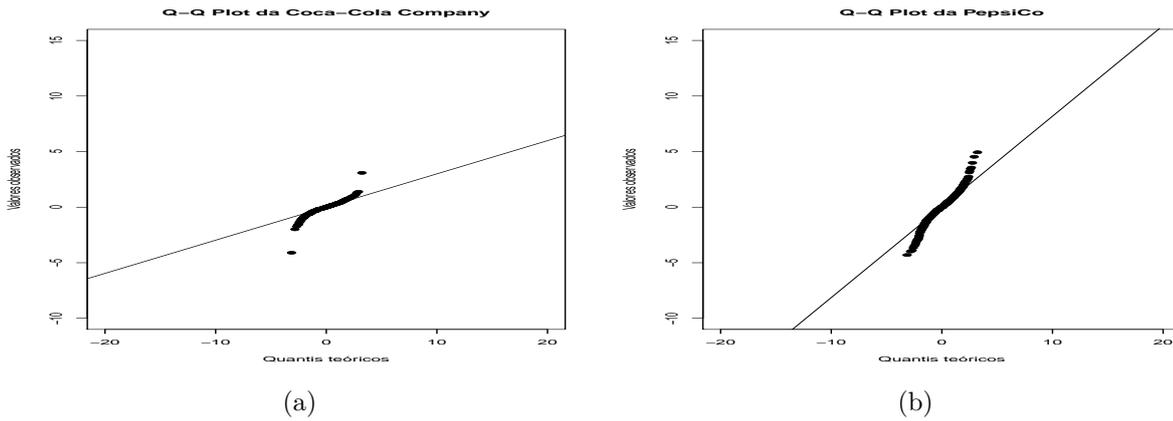


Figura 5.3: Q-Q plots Gaussianos para as séries temporais  $Y_t$ : (a) Coca-Cola Company; (b) PepsiCo Inc.

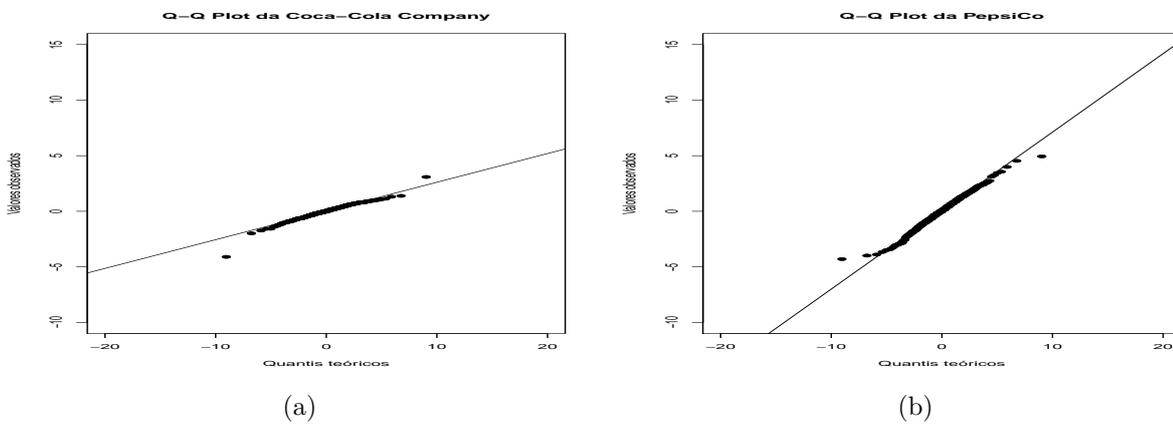


Figura 5.4: Q-Q plots  $t$ -Student para as séries temporais  $Y_t$ : (a) Coca-Cola Company; (b) PepsiCo Inc.

Em comparação, podemos ver os seguintes gráficos para a distribuição  $\alpha$ -Estável.

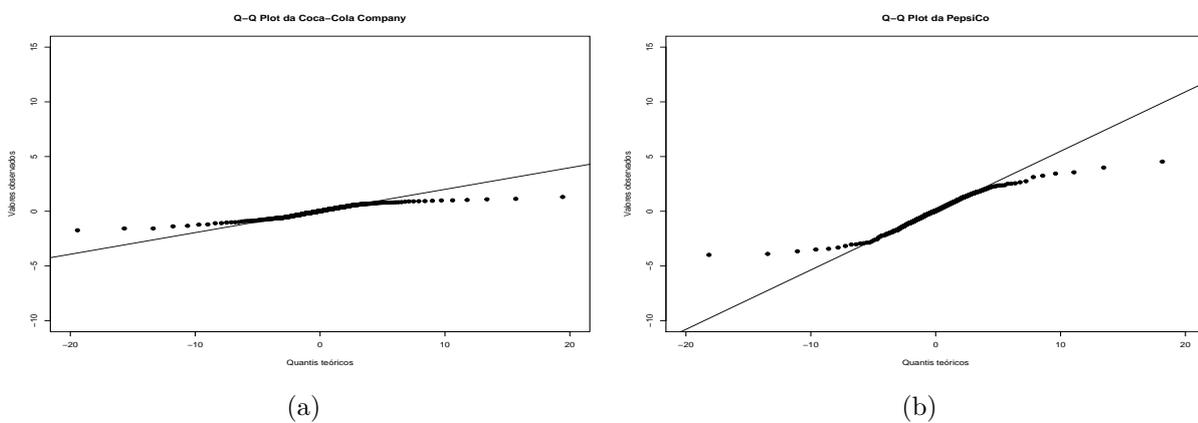


Figura 5.5: Q-Q plots  $\alpha$ -Estáveis para as séries temporais  $Y_t$ : (a) Coca-Cola Company; (b) PepsiCo Inc.

Tabela 5.1: Resultado dos testes KS, AD, MKS e MC de adequação de modelos para normalidade, para os dados diferenciados das ações das empresas Coca-Cola Company e PepsiCo Inc.

Componente	Teste	Estatística	$p$ -Valor
Coca-Cola Company	KS	0.0810	0.001
	AD	12.2000	0.001
	MKS	179.5270	0.001
	MC	0.9990	0.001
PepsiCo Inc.	KS	0.0580	0.001
	AD	8.0800	0.001
	MKS	190.1400	0.001
	MC	0.9990	0.001

Tabela 5.2: Resultado dos testes KS, AD, MKS e MC de adequação de modelos para  $\alpha$ -Estáveis, para os dados das ações, diferenciadas, das empresas Coca-Cola Company e PepsiCo Inc.

Componente	Teste	Estatística	$p$ -Valor
Coca-Cola Company	KS	0.1400	1.000
	AD	68.3000	0.135
	MKS	34.4500	0.943
	MC	0.2780	0.931
PepsiCo Inc.	KS	0.1260	1.000
	AD	57.2300	0.999
	MKS	32.9500	0.880
	MC	0.1580	0.962

Os gráficos apresentados na Figura 5.5 mostram que os dados se comportam semelhantemente a uma distribuição  $\alpha$ -Estável com os parâmetros estimados. Há um erro presente na cauda da distribuição que pode ser explicado pelo tamanho amostral e pela variabilidade intrínseca de um processo  $\alpha$ -Estável.

Em seguida, apresentamos os resultados para os testes de adequação de modelos para os dados padronizados, estudados na Seção 2.5.

A Tabela 5.1 apresenta os resultados dos testes de adequação de modelos. Utilizamos os testes KS, AD, MKS e MC para a normalidade. A hipótese nula ( $H_0$ ) de todos os testes apresentados nesta tabela é de que os dados são provenientes de uma distribuição Gaussiana. Equivalentemente, a hipótese alternativa ( $H_1$ ) é de que os dados são provenientes de outra distribuição não-Gaussiana. Observamos que, para todos os testes utilizados, os  $p$ -valores são inferiores a 0.05. Portanto, os testes indicam que os dados devem ser modelados por uma distribuição não-Gaussiana. Em seguida, testamos para a  $\alpha$ -Estável.

Na Tabela 5.2, a hipótese nula ( $H_0$ ) dos testes apresentados é de que os dados são pro-

venientes de uma distribuição  $\alpha$ -Estável com  $\hat{\alpha} = (1.5311, 1.6566)$  e  $\hat{\beta} = (0, 0)$ , para os dados diferenciados das ações da Coca-Cola Company e da PepsiCo Inc., respectivamente. A hipótese alternativa ( $H_1$ ) é de que os dados são provenientes de uma  $\alpha$ -Estável com outros vetores de parâmetros. Observamos que os testes não rejeitam  $H_0$ , indicando que os dados possivelmente podem ser modelados por uma  $\alpha$ -Estável com  $\alpha = \hat{\alpha}$  e  $\beta = \hat{\beta}$ .

Tendo em vista que os testes mostram que as duas séries são bem modeladas por uma distribuição  $\alpha$ -Estável, calculamos a codiferença cruzada, estudada na Seção 2.4.3. O objetivo é estudar a relação de dependência entre as duas séries. Através da Figura 5.6 podemos ver que há uma moderada dependência entre as séries para  $k = 0$  ( $\widehat{CD}(ko, pep) = 0.2074$ ), ou seja, para  $t_{ko} = t_{pep}$ . Sabendo que o modelo VAR(1) bidimensional define uma dependência entre cada uma das suas duas componentes entre si mesmas e entre elas nos instantes  $t$  e  $t - 1$ , os dados parecem se adequar bem ao modelo. Observamos as estimações dos parâmetros na Tabela 5.3, onde comparamos três modelos VAR( $p$ ) com  $p = 1, 2, 3$ . Pela Tabela 5.3 concluímos que os dados são provenientes de um VAR(1) bidimensional pois para  $p = 1$  obtivemos os menores valores para BIC. Apesar dos valores de AIC serem superiores para  $p = 1$ , selecionamos o modelo mais parcimonioso, ou seja, menos complexo pois este apresentará o menor erro nas estimativas, conforme Morettin e Toloi (2014).

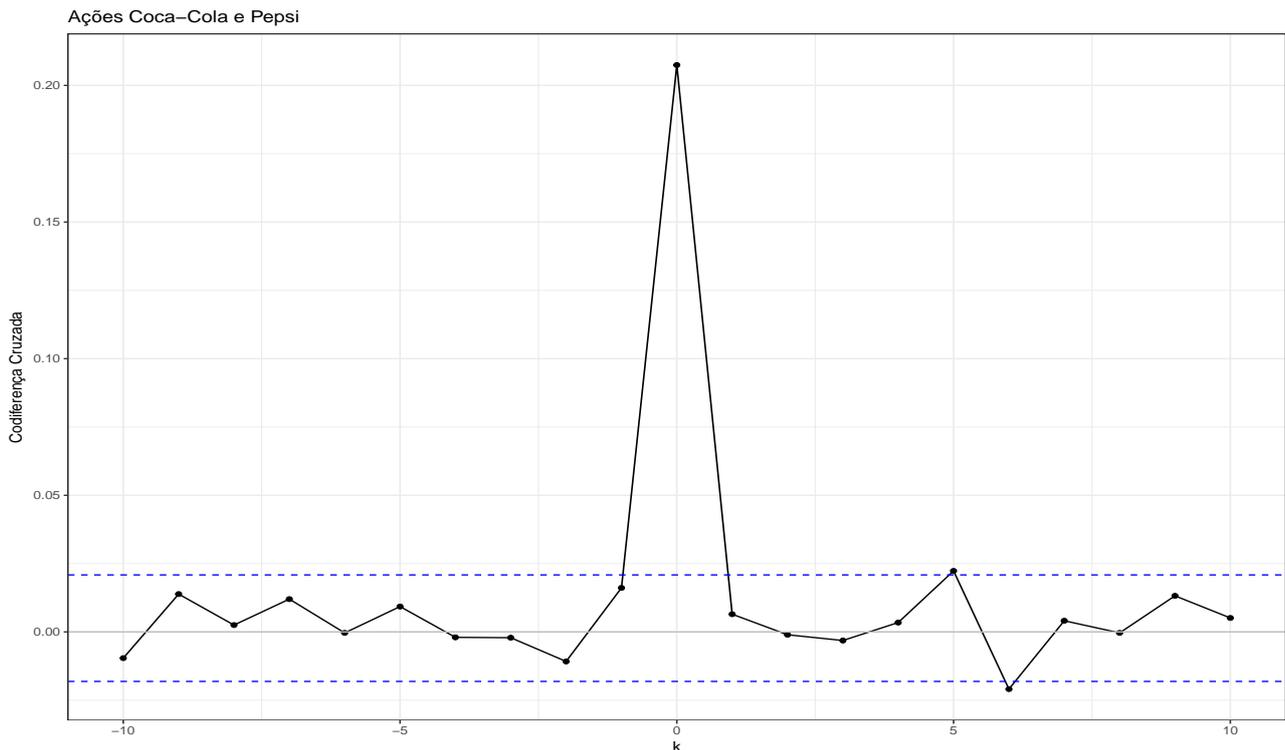


Figura 5.6: Codiferença Cruzada para  $k$  entre -10 e 10, das séries provenientes dos valores diferenciados das ações das empresas Coca-Cola Company e PepsiCo Inc. As linhas azuis representam o intervalo a 95% de confiança.

Tabela 5.3: Resultados das estimações dos modelos VAR( $p$ ),  $p = 1, 2, 3$  para os dados diários diferenciados das ações da Coca-Cola Company e da PepsiCo Inc, com seus respectivos AIC e BIC. O coeficiente  $\phi_{j,k}$  corresponde ao componente  $j = (pep, ko)$  no lag  $k$ .

$p$	Componente	$\phi_{pep,1}$	$\phi_{ko,1}$	$\phi_{pep,2}$	$\phi_{ko,2}$	$\phi_{pep,3}$	$\phi_{ko,3}$	AIC	BIC
1	Pepsi	0.0111	-0.1827					3650.7960	3671.1990
	Coca-Cola	0.0122	-0.0388						
2	Pepsi	0.0061	-0.1731	-0.0661	0.0274			3637.8410	3678.6420
	Coca-Cola	0.0082	-0.0324	-0.0326	-0.0373				
3	Pepsi	0.0112	-0.1804	-0.0629	0.0208	0.0493	-0.0830	3635.9800	3697.1700
	Coca-Cola	0.0095	-0.0343	-0.0323	-0.0385	0.0184	-0.0409		

## 6

# Conclusão

Este trabalho trata de métodos de estimação clássicos e Bayesianos de processos autorregressivos bivariados de primeira ordem, cujas inovações seguem as distribuições Gaussiana,  $t$ -Student e  $\alpha$ -Estável. Através de simulações de Monte Carlo, observou-se vício, desvio padrão e erro quadrático médio e percebeu-se que as estimações são adequadas. Através destas estatísticas, conclui-se que o método de estimação clássico apresentou melhor resultado para as séries estudadas pois apresentou menor desvio padrão, vício e erro quadrático médio.

Abordou-se testes de adequação capazes de detectar o comportamento estável nas séries. Apresentamos e aplicamos as funções Codiferença e Codiferença Cruzada como medidas de dependência para os casos  $\alpha$ -Estáveis. As distribuições  $\alpha$ -Estáveis apresentam segundo momento infinito e, portanto, as funções de autocovariância e autocovariância cruzada não estão bem definidas.

Analisou-se dados diários de ações da *Coca-Cola Company* e da *PepsiCo Inc.* e aplicou-se os métodos estudados. Observou-se, a partir dos resultados obtidos, que estes beneficiam-se das inovações  $\alpha$ -Estáveis, em relação aos casos Gaussianos e  $t$ -Student, quando para as duas séries temporais é aplicada a primeira diferença.

Por fim, conclui-se que os resultados deste trabalho asseguram a aplicabilidade das estimações de séries temporais com inovações Gaussianas e não-Gaussianas. Com base no estudo realizado até o momento e nos resultados obtidos, temos informações para comparação das metodologias e continuidade do objetivo em futuros trabalhos.

# Referências

- Achcar, J.A., S.R.C. Lopes, J. Mazucheli e R.R. Linhares (2013). “A Bayesian Approach for Stable Distributions: Some Computational Aspects”. *Open Journal of Statistics*, Vol. **03**, 268-277.
- Achcar, J.A. e S.R.C. Lopes (2016). “Linear and Non-Linear Regression Models Assuming a Stable Distribution”. *Colombian Journal of Statistics*, Vol. **39(1)**, 109-128.
- Adler, R.J., R.E. Feldman e C. Gallangher (1998). “Analysing stable time series”. Em: R.J. Adler, R.E. Feldman e M.S. Taqqu (eds.) *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, 133-158.
- Brockwell, P.J. e R.A. Davis (1991). *Time Series: Theory and Methods*. New York: Springer-Verlag.
- Grzesiek, A., M. Teuerle e A. Wylomańska (2019). “Cross-codifference for bidimensional VAR(1) models with infinite variance”. *arXiv preprint arXiv:1902.02142*.
- Linden, W. van der, V. Dose e U. von Toussaint (2014). *Bayesian Probability Theory: Applications in the Physical Sciences*. Cambridge: Cambridge University Press.
- McCulloch, J.H. (1986). “Simple Consistent Estimator of Stable Distribution Parameters”. *Communications in Statistics - Simulation and Computation*, Vol. **15**, 1109-1136.
- Morettin, P.A. e C.M.C. Tolo (2004). *Análise de Séries Temporais*. ABE - Projeto Fisher. São Paulo: Edgard Blücher.
- Rosadi, D. e M. Deistler (2009). “Estimating the codifference function of linear time series models with infinite variance”. *Metrika*, Vol. **73**, 395-429.
- Samorodnitsky, G. e M.S. Taqqu (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. New York: Chapman & Hall.
- Stein, J., S.R.C. Lopes e A.V. Medino (2016). “Continuous Processes Derived from the Solution of Generalized Langevin Equation: Theoretical Properties and Some Simulations”. *Journal of Statistical Computation and Simulation*, Vol. **86**, 2819-2845.
- Stein, J. (2015). “Estimação e Previsão em Processos Advindos da Solução da Equação Generalizada de Langevin com Ruído  $\alpha$ -Estável”. *Tese de Doutorado defendida no PPG-MAT, UFRGS, Porto Alegre*.
- Tsay, R.S. (2014). *Multivariate Time Series Analysis: With R and Financial Applications*. New Jersey: Wiley.