

Ilka Oliveira Torres

**Identificação e estimação de modelos ARMA
com inovações Alpha-estáveis**

Brasília

2016

Ilka Oliveira Torres

Identificação e estimação de modelos ARMA com inovações Alpha-estáveis

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Universidade de Brasília - UnB

Departamento de Estatística

Programa de Pós-Graduação

Orientador: Dr. Jhames Matos Sampaio

Brasília

2016

Ilka Oliveira Torres

Identificação e estimação de modelos ARMA com inovações Alpha-estáveis/
Ilka Oliveira Torres. – Brasília, 2016-
85 p. : il. (algumas color.) ; 30 cm.

Orientador: Dr. Jhames Matos Sampaio

Dissertação de Mestrado – Universidade de Brasília - UnB
Departamento de Estatística
Programa de Pós-Graduação, 2016.

1. Séries Temporais. 2. Distribuições α -estáveis. 3. Modelo ARMA. 4.
Função codiferença. I. Jhames Matos Sampaio. II. Universidade de Brasília. III.
Departamento de Estatística. IV. Identificação e estimação de modelos ARMA
com inovações α -estáveis

Ilka Oliveira Torres

Identificação e estimação de modelos ARMA com inovações Alpha-estáveis

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Trabalho aprovado. Brasília, 29 de julho de 2016:

Banca examinadora:

- Dr. Jhames Matos Sampaio(Orientador)
- Dr. Lucas Moreira
- Dra. Cátia Regina Gonçalves
- Dr. Raul Yukihiro Matsushita (Suplente)

Brasília
2016

Este trabalho é dedicado a todos que me disseram “vai dar certo”.

Agradecimentos

Agradeço primeiramente a Deus por nunca me deixar desistir.

Ao apoio prestado por minha irmã que aguentou meus momentos de estresse, me consolou nos momentos de tristeza e celebrou comigo em cada conquista. Sem você ao meu lado não teria conquistado metade do que sou hoje, obrigada!

Aos meus pais que me criaram para voar sozinha mas nunca deixaram de voar ao meu lado.

Aos meus amigos pela compreensão da minha ausência aos eventos que não pude comparecer e pelos momentos de descontração em meio a pressão do mestrado.

Aos meus chefes por flexibilizarem meu horário para cursar as disciplinas do mestrado.

Ao meu orientador, Jhames Matos Sampaio, por aceitar meu convite de orientação e compartilhar seus conhecimentos em séries temporais.

Aos professores Lucas Moreira e Cira Etheowalda Guevara Otiniano pelas sugestões durante o exame de qualificação.

Aos membros da banca examinadora, por aceitarem o convite.

E ao serviço prestado por todos os funcionários do Departamento de Estatística.

*“Se você deseja chegar a algum lugar,
precisa saber aonde quer ir e como chegar lá.”*

Norman Vicente Peale

Resumo

No mercado financeiro é usual que os dados apresentem características não-gaussianas como assimetria e caudas pesadas. Nestes casos, os resultados de uma análise baseada na suposição de normalidade podem ser insatisfatórios. Uma alternativa ao se trabalhar com este tipo de dados é a utilização das distribuições estáveis. Neste trabalho aplicamos o uso do modelo ARMA com inovações α -estáveis para ajustar séries do mercado financeiro. O método de identificação da ordem é baseado na função codiferença proposta por [Kokoszka e Taqqu \(1994\)](#) e [Yang et al. \(2001\)](#), uma vez que a função de autocorrelação não existe para modelos com inovações α -estáveis dado que o segundo momento não existe para $\alpha < 2$. Com base na função característica empírica, [Rosadi e Deistler \(2011\)](#) apresentam o estimador para a função codiferença e mostram sua consistência para o caso do modelo ARMA com inovações α -estáveis simétrica. A estimação dos parâmetros foi realizada através do método de máxima verossimilhança condicional e, por fim, foi realizada uma aplicação empírica a partir do preço de fechamento das ações da Usinas Siderúrgicas de Minas Gerais (Usiminas).

Palavras-chaves: Séries Temporais, Distribuições α -estáveis, Modelo ARMA, Função codiferença.

Abstract

In the financial market is usual that the data present non-Gaussian features like asymmetry and heavy tails. In these cases, the results of an analysis based on normality assumption may be unsatisfactory. An alternative when working with this type of data is the use of stable distributions. In this paper we apply the use of ARMA model with α -stable innovations to adjust time series of the financial market. The method to identify the model order is based on the codiferença function, proposed by [Kokoszka e Taqqu \(1994\)](#) and [Yang et al. \(2001\)](#), since the autocorrelation function does not exist for models with α -stable innovations given that the second moment do not exist for $\alpha < 2$. Based on the empirical characteristic function, [Rosadi e Deistler \(2011\)](#) proposes the estimator for codiferença function and proof consistency in the case of the ARMA symmetric α -stable innovations. The conditional maximum likelihood method was applied on the estimation of parameters and, finally, was carried out a empirical application with the closing stock prices of the Steel Mills of Minas Gerais (Usiminas).

Key-words: Time series, α -stable distribution, ARMA model, Codifference function.

Lista de ilustrações

Figura 1 – Exemplo de série estacionária.	2
Figura 2 – FAC e FACP amostrais de um AR(1) simulado.	8
Figura 3 – FAC e FACP amostrais de um MA(2) simulado.	8
Figura 4 – FAC e FACP amostrais de um ARMA(2,2) simulado.	9
Figura 5 – Densidade da α -estável variando o índice de estabilidade α e parâmetro de assimetria.	16
Figura 6 – Densidade da α -estável variando o parâmetro de escala e locação.	17
Figura 7 – Gráficos $\mathfrak{R}\hat{I}(s,k)$ versus s	40
Figura 8 – $\mathfrak{R}\hat{I}(s,k)$ versus s nos lags 1, 2 e 3.	41
Figura 9 – Gráficos de $W_1(i,j)$ para $s_i, s_j \in [0,01,1]$ e alguns valores de α	42
Figura 10 – Gráficos quantil-quantil do log retorno	49
Figura 11 – Densidade do log retorno e densidade estimada para os parâmetros estimados.	50
Figura 12 – Gráfico da densidade do log retorno em diferentes níveis de agregação	51
Figura 13 – Gráfico da codiferença amostral estimada ($\mathfrak{R}\hat{I}(1)$)	52
Figura 14 – Gráfico Codiferença padronizada estimada para $s=(0.01,0.1,0.2)$, 0.01 e (0.01,0.06,0.11) e FAC	53
Figura 15 – Autocorrelação parcial dos log retornos	53
Figura 16 – Codiferença, FAC e FACP dos resíduos do ajuste ARMA(3,1).	55
Figura 17 – Codiferença, FAC e FACP dos resíduos do ajuste AR(1).	56
Figura 18 – Codiferença, FAC e FACP dos resíduos do ajuste ARMA(1,1).	57
Figura 19 – Gráficos quantil-quantil do resíduos do ajuste ARMA(3,1).	58
Figura 20 – Gráficos quantil-quantil do resíduos do ajuste AR(1).	58
Figura 21 – Gráficos quantil-quantil do resíduos do ajuste ARMA(1,1).	58

Lista de tabelas

Tabela 1 – Valores da função codiferença $I(.)$ e sua estimativa $\hat{I}(.)$ para séries simulada.	45
Tabela 2 – Resumo estatístico do log retorno	48
Tabela 3 – Testes de normalidade para o log retorno	48
Tabela 4 – Parâmetros da α -estável estimado para os log retornos.	49
Tabela 5 – Estimativa de α para diferentes níveis de agregação.	51
Tabela 6 – Ajuste do modelo ARMA(3,1) para erros normais e α -estáveis.	54
Tabela 7 – Teste Box-Pierce-Ljung para os resíduos do ajuste ARMA(3,1).	54
Tabela 8 – Ajuste do modelo AR(1) para erros normais e α -estáveis.	55
Tabela 9 – Teste Box-Pierce-Ljung para os resíduos do ajuste AR(1).	56
Tabela 10 – Ajuste do modelo ARMA(1,1) para erros normais e α -estáveis	57
Tabela 11 – Teste Box-Pierce-Ljung para os resíduos do ajuste ARMA(1,1).	57
Tabela 12 – Teste Kolmogorov-Smirnov para os resíduos dos modelos ajustados.	59
Tabela 13 – Valores da função codiferença $I(.)$ e sua estimativa $\hat{I}(.)$ para séries simulada - Modelo 2	82
Tabela 14 – Valores da função codiferença $I(.)$ e sua estimativa $\hat{I}(.)$ para séries simulada - Modelo 3	83
Tabela 15 – Valores da função codiferença $I(.)$ e sua estimativa $\hat{I}(.)$ para séries simulada - Modelo 4	85

Sumário

	Agradecimentos	v
	Lista de ilustrações	ix
	Lista de tabelas	x
	Sumário	xi
1	CONCEITOS BÁSICOS	1
1.1	Séries Temporais e Processos Estocásticos	1
1.2	Processos ARMA(p,q)	4
1.2.1	Condições de estacionariedade e invertibilidade	5
1.2.2	Identificação da ordem do Modelo ARMA	6
1.2.3	Estimação dos parâmetros do Modelo ARMA	9
1.2.3.1	Método dos Momentos	10
1.2.3.2	Método de Máxima Verossimilhança	11
1.2.3.3	Método de Máxima Verossimilhança Condicional	11
2	DISTRIBUIÇÕES ESTÁVEIS	13
2.1	Definições	13
2.2	Parâmetros da distribuição α-estável	16
2.3	Propriedades das variáveis aleatórias α-estáveis	17
2.4	Estimação dos parâmetros	19
2.5	Processos ARMA com inovações α-estáveis	21
3	IDENTIFICAÇÃO DA ORDEM PARA ARMA(P,Q) COM INOVAÇÕES α-ESTÁVEIS	25
3.1	Função codiferença	25
3.1.1	Definição	25
3.1.2	Estimação	27
3.1.3	Propriedades assintóticas do estimador	27
3.2	Função codiferença na identificação da ordem	30
4	SIMULAÇÕES	33
4.1	Simulação de uma variável aleatória α-estável	33
4.2	Resultados para estimativa da função codiferença	39

4.2.1	Considerações práticas	39
4.2.2	Resultados da simulação	43
5	APLICAÇÃO	48
6	CONCLUSÃO	60
6.1	Trabalhos futuros	60
	REFERÊNCIAS	62
	APÊNDICE A – DEMONSTRAÇÃO DO TEOREMA 3.2	65
	APÊNDICE B – A DISTRIBUIÇÃO LIMITE DA FUNÇÃO CO- DIFERENÇA AMOSTRAL	68
	APÊNDICE C – DEMONSTRAÇÃO DO COROLÁRIO 3.1 . .	77
	APÊNDICE D – CONTINUAÇÃO RESULTADOS DA SIMU- LAÇÃO	81

Introdução

Uma das áreas da estatística que está mais presente em nosso dia-a-dia são séries temporais, como quando acompanhamos a cotação do dólar, índice de desemprego e o preço da cesta básica, por exemplo. As técnicas de análise de séries temporais visam representá-las sob a forma de uma expressão numérica e reconhecer padrões de comportamento, essas expressões numéricas são denominadas por modelo.

O modelo de séries temporais a ser apresentado no trabalho é o processo autorregressivo e de médias móveis, ARMA(p,q), que consiste na junção dos processos autorregressivo de ordem p , representado por AR, e médias móveis de ordem q , MA. O modelo ARMA(p,q) pode ser escrito da seguinte forma:

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q} \quad (1)$$

onde X_t é a observação no tempo t e $\{a_t\} \sim RB(0, \sigma^2)$ no caso gaussiano.

Os estudos clássicos das séries temporais são voltados para modelos gaussianos, entretanto, estes modelos não permitem grandes flutuações e são, portanto, frequentemente inadequados para modelos com alta variabilidade ([SAMORODNITSKY; TAQQU, 1994](#)). Os modelos com inovações α -estáveis não apresentam estas limitações. A família das distribuições estáveis são mais flexíveis permitindo uma maior versatilidade no ajuste à curva dos dados. Dizemos que a variável Y_t tem distribuição α -estável com parâmetros $(\alpha, \beta, \sigma, \mu)$ se sua função característica pode ser escrita da forma:

$$\ln \varphi_Y(t) = \begin{cases} i\mu t - \sigma^\alpha |t|^\alpha [1 - i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2}], & \text{se } \alpha \neq 1, \\ i\mu t - \sigma |t| [1 + i\beta \frac{2}{\pi} \operatorname{sign}(t) \ln |t|], & \text{se } \alpha = 1. \end{cases}$$

Em que $\alpha \in (0,2]$ é conhecido como parâmetro de estabilidade e tem maior influência sob o decaimento das caudas, elas ficam mais pesadas para valores baixos de α . Os outros parâmetros são o de assimetria $\beta \in [-1,1]$, de escala $\sigma \in (0, +\infty)$ e de posição $\mu \in \mathbb{R}$. Para $\alpha \leq 1$, $E|Y_t|^p = \infty$ e $E|Y_t|^p < \infty$ para $0 < p < \alpha$. Exceto para o caso normal ($\alpha = 2$), a distribuição estável tem variância infinita.

No modelo ARMA com inovações α -estáveis, a representação é semelhante à equação (1), onde a distribuição de a_t é α -estável. Como a distribuição α -estável não tem segundo momento finito para $\alpha < 2$, não é recomendável calcular a função de covariância para descrever a estrutura de dependência do processo X_t com inovações α -estáveis. Para contornar esta situação, utilizamos a função codiferença proposta por [Kokoszka e Taqqu \(1994\)](#) e [Yang et al. \(2001\)](#) para descrever a estrutura de

dependência e auxiliar na identificação da ordem.

$$\begin{aligned} \tau(h) = \tau(s, -s; h) = & -\ln E \exp\{is(X_{t+h} - X_t)\} + \ln E \exp\{isX_{t+h}\} \\ & + \ln E \exp\{-isX_t\}, \end{aligned} \quad (2)$$

onde $s \in \mathbb{R}$ e $h \in \mathbb{Z}$.

Por ser calculada a partir da função característica e existir para todo valor de s , não é necessário que os momentos de $\{X_t\}$ sejam finitos.

Este trabalho tem por objetivo identificar a ordem e estimar os parâmetros para dados com as seguintes características, caudas pesadas e assimetria. Nestes casos é indicado o uso do modelo ARMA com inovações da família α -estável, onde será utilizada a função codiferença na identificação da ordem e o método de máxima verossimilhança condicional na estimação dos parâmetros. Ao final vamos comparar os resultados do ajuste com erros gaussianos e erros estáveis através do erro quadrático médio e análise de resíduos. O conteúdo dos capítulos é apresentado de forma gradual até atingir o objetivo final no Capítulo 5.

O Capítulo 1 consiste em uma síntese sobre o contexto geral de séries temporais e processos estocásticos, passando desde a definição do modelo ARMA, passos para a identificação da ordem por meio da função de autocorrelação e autocorrelação parcial até a aplicação do método de máxima verossimilhança para estimação dos parâmetros no caso do ARMA gaussiano. O Capítulo 2 é voltado para as distribuições estáveis, nele são apresentadas suas definições, propriedades, exemplos e métodos de estimação dos parâmetros. Este capítulo finaliza com a apresentação do processo ARMA com inovações α -estáveis.

Os Capítulos 3 e 4 foram baseados no artigo [Rosadi e Deistler \(2011\)](#). O Capítulo 3 expõe um estudo sobre o estimador da função codiferença e suas propriedades assintóticas e é finalizado com a indicação do uso da função codiferença para identificação da ordem de modelos com inovações α -estáveis. O Capítulo 4 engloba a simulação da variável α -estável segundo [Weron \(1996\)](#) e as considerações práticas para obter uma boa estimativa da função codiferença.

No Capítulo 5 todos os resultados e conceitos apresentados até aqui são utilizados para analisar o preço das ações da Usinas Siderúrgicas de Minas Gerais S.A (Usiminas). Por último, o Capítulo 6 sintetiza os resultados apresentados no Capítulo 5 e sugere algumas propostas de trabalhos futuros.

1 Conceitos Básicos

Neste capítulo serão apresentados os conceitos básicos sobre séries temporais e processos estocásticos. A primeira seção descreve a definição de série temporal e processo estocástico, seguido pelo conceito de estacionariedade e medidas de dependência. Na segunda seção discorreremos sobre o modelo ARMA, incluindo a identificação da ordem e estimação dos parâmetros.

1.1 Séries Temporais e Processos Estocásticos

Série temporal é um conjunto de observações Y_t coletadas em um espaço de tempo $\{t \in T\}$, em que T é um conjunto de índices. Quando o conjunto de índices T é discreto, ou seja, as observações são coletadas em intervalos de tempos espaçados, a série temporal é dita *discreta*. Caso a coleta seja realizada de forma contínua num intervalo $T \in [a, b]$ dizemos que a série temporal é *contínua*.

Abaixo citamos alguns exemplos de séries temporais discretas:

- Registro da temperatura máxima diária em Brasília;
- Índices diários da Bovespa;
- Número de clientes atendidos em um dia.

Os registros de maré representam um exemplo de série temporal contínua.

Processo estocástico $\{X(t), t \in T\}$ é um conjunto de variáveis aleatórias ordenadas no tempo, onde para cada $t \in T$, $X(t)$ é uma variável aleatória e T é um conjunto ordenado de índices. Portanto, série temporal é uma realização de um processo estocástico.

Definição 1.1. O processo $\{X_t, t \in T\}$ diz-se *estritamente estacionário* se todas as distribuições finito-dimensionais (1.1) permanecem as mesmas sob translações no tempo, ou seja,

$$F(x_1, \dots, x_n; t_{1+h}, \dots, t_{n+h}) = F(x_1, \dots, x_n; t_1, \dots, t_n) \quad (1.1)$$

para quaisquer $t_1, \dots, t_{n+h} \in T$.

Portanto, a partir da Definição 1.1 podemos afirmar que uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) formam uma série estritamente estacionária.

Graficamente, dizemos que a série é estacionária quando esta tem comportamento aleatório em torno da média ao longo do tempo. O ruído branco é um exemplo de série temporal estacionária.

Definição 1.2. O processo $\{a_t\}$ é um ruído branco com média 0 e variância σ^2 , se e somente se, $\{a_t\}$ tem média zero e covariância igual a:

$$\text{Cov}(a_{t+h}, a_t) = \begin{cases} \sigma^2 & , \text{ se } h = 0 \\ 0 & , \text{ se } h \neq 0. \end{cases}$$

Ao longo deste trabalho o ruído branco será denotado por $a_t \sim RB(0, \sigma^2)$. A figura abaixo apresenta a simulação de uma série estacionária $RB(0,2)$ e sua função de autocovariância.

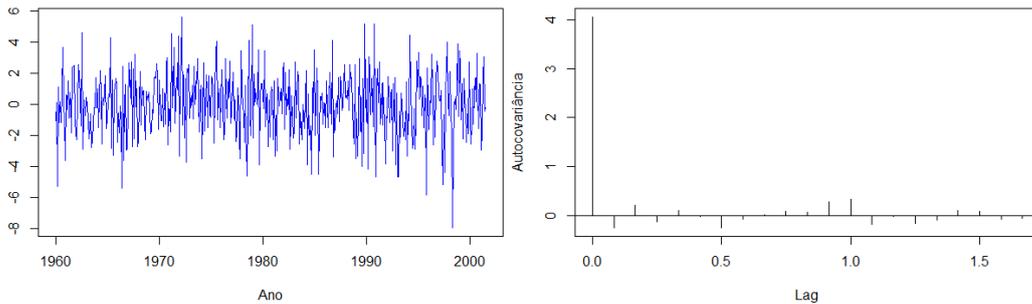
(a) $RB(0,2)$ simulado.(b) FAC do $RB(0,2)$ simulado.

Figura 1 – Exemplo de série estacionária.

A detecção de série estacionária a partir do cálculo das distribuições finito-dimensionais em todas as translações de tempo segundo a Definição 1.1 demanda muito trabalho e gasto computacional. Apesar do gráfico da série temporal auxiliar nesta tarefa, faz-se necessário um resultado numérico para sustentar a análise do gráfico. A Definição 1.3 permite que estes resultados sejam obtidos de forma mais simples.

Definição 1.3. O processo $\{X_t, t \in T\}$ diz-se (fracamente) estacionário se, e somente se:

1. $E[X_t] = \mu_X(t) = \mu$, independente de t ;
2. $E[X^2(t)] < \infty$, para todo $t \in T$;

3. $\gamma_X(h) = \gamma_X(t+h, t) = Cov(X_{t+h}, X_t)$ é independente de t para cada h .

Onde $\gamma_X(h)$ é a função de autocovariância no lag h .

Note que o processo $\{X_t\}$ ser estritamente estacionário não implica que ele seja fracamente estacionário e se o processo $\{X_t\}$ é estritamente estacionário, ele será fracamente estacionário somente se seu segundo momento for finito.

Grande parte das séries temporais apresentam alguma dependência entre os valores observados em tempos diferentes. As funções de autocovariância e autocorrelação são medidas que mensuram o grau de dependência em cada intervalo de tempo da série. Os gráficos destas funções são ferramentas fundamentais na identificação da ordem do modelo e detecção de sazonalidade (MORETTIN; TOLOI, 2006).

Definição 1.4. *Seja $\{X_t, t \in T\}$ um processo estacionário. A função de autocovariância (FACV) de X_t é dada por:*

$$\gamma_X(h) = Cov(X_{t+h}, X_t).$$

E a função de autocorrelação (FAC) é dada por:

$$\rho_x(h) \equiv \frac{\rho_X(h)}{\rho_X(0)} = Corr(X_{t+h}, X_t).$$

Para calcular a FAC e FACV a partir da Definição 1.4 é necessário conhecer os parâmetros do processo estacionário $\{X_t\}$, entretanto, ao trabalharmos com um conjunto de dados observados raramente este modelo é conhecido. Dada a importância destas medidas na seleção de modelos temporais, estabelecer seus estimadores é fundamental. Desse modo, é bastante natural estimarmos a FACV por meio da função de autocovariância amostral apresentada abaixo:

Definição 1.5. *Seja o vetor $X = (x_1, \dots, x_n)$ as observações de uma série temporal. A média amostral de X é dada por:*

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t.$$

A função de autocovariância amostral é dada por:

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x})$$

e a função de autocorrelação amostral é dada por:

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)},$$

para $-n < h < n$.

Pode-se mostrar que para a série i.i.d. com variância finita, as funções de autocorrelação amostral $\hat{\rho}(h)$, para $h > 0$, são aproximadamente i.i.d. com distribuição $N(0,1/n)$ para n grande. Portanto, aproximadamente 95% das funções de autocorrelação amostral devem cair no intervalo $\pm 1.96/\sqrt{n}$, uma vez que 1.96 é o 0.975 quantil da normal padronizada (BROCKWELL; DAVIS, 1991).

Outra medida de dependência muito utilizada na identificação da ordem dos modelos é a função de autocorrelação parcial (FACP), correspondente à correlação entre X_t e X_{t-k} removendo o efeito das observações $X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$. A FACP é denotada pela função $\alpha(\cdot)$ e definida pelas equações:

$$\begin{aligned}\alpha(0) &= 1 \text{ e} \\ \alpha(h) &= \phi_{hh}, \quad h \geq 1.\end{aligned}$$

Em que ϕ_{kk} é o último componente de

$$\phi_h = \Gamma_h^{-1} \gamma_h,$$

para $\Gamma_h = [\gamma(i-j)]_{i,j=1}^h$ e $\gamma = [\gamma(1), \dots, \gamma(h)]^T$. Podemos calcular a FACP amostral $\hat{\alpha}(h)$ através das seguintes equações:

$$\begin{aligned}\hat{\alpha}(0) &= 1 \text{ e} \\ \hat{\alpha}(h) &= \hat{\phi}_{hh}, \quad h \geq 1.\end{aligned}$$

Onde $\hat{\phi}_{kk}$ é o último componente de

$$\hat{\phi}_h = \hat{\Gamma}_h^{-1} \hat{\gamma}_h.$$

Para uma amostra suficientemente grande de um processo AR(p), Quenouille (1949) mostra que, as FACP estimadas de ordem $p+1, p+2, \dots$ são, aproximadamente, i.i.d com

$$\text{Var}(\hat{\phi}_{kk}) \approx \frac{1}{N}, \quad k \geq p+1$$

e distribuição aproximada normal com média 0 e variância $\frac{1}{N}$. Portanto, podemos criar um intervalo de confiança na forma $\pm 1.96/\sqrt{n}$, uma vez que 1.96 é o 0.975 quantil da normal padronizada.

1.2 Processos ARMA(p,q)

O processo autorregressivo e de médias móveis (ARMA) consiste na “junção” dos processos autorregressivo de ordem p, representado por AR(p), e médias móveis de ordem q, MA(q). Esta classe de modelo é recorrente em áreas como economia e

ciências físicas e geofísicas. A inclusão de termos autorregressivos e de médias móveis surgiu como uma alternativa para simplificar um modelo com grande número de parâmetros (MORETTIN; TOLOI, 2006).

Definição 1.6. $\{X_t\}$ é um processo ARMA(p,q), se $\{X_t\}$ é estacionário e $\forall t$

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}. \quad (1.2)$$

Onde $\{a_t\} \sim RB(0, \sigma^2)$.

A equação (1.2) pode ser reescrita de forma mais sintética utilizando o operador retroativo B .

$$\Phi(B)X_t = \Theta(B)a_t, \quad (1.3)$$

onde $\Phi(\cdot)$ e $\Theta(\cdot)$ são polinômios de ordem p e q:

$$\begin{aligned} \Phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \\ \Theta(B) &= 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \end{aligned}$$

e B é o operador retroativo ($B^j X_t = X_{t-j}$).

Nos casos em que $\Theta(B) \equiv 1$ o processo $\{X_t\}$ é um processo AR(p) e se $\Phi \equiv 1$ o processo se torna um MA(q).

1.2.1 Condições de estacionariedade e invertibilidade

Dizemos que o processo ARMA(p,q) é **causal** se pudermos reescrevê-lo como uma soma dos erros passados, ou seja, se existem constantes $\{\psi_j\}$ tais que $\sum_{j=0}^{\infty} |\psi_j| < \infty$ e

$$X_t = \frac{\Theta(B)}{\Phi(B)} a_t = \psi(B) a_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}, \text{ para todo } t. \quad (1.4)$$

De fato, a causalidade de um processo ARMA(p,q) se confunde com a estacionariedade, uma vez que sendo o processo causal ele também será estacionário (MORETTIN; TOLOI, 2006).

Observe que a equação (1.4) é a representação de um processo ARMA(p,q) na forma de um processo de médias móveis com ordem infinita, $MA(\infty)$.

Outra condição para determinar se um processo ARMA(p,q) é estacionário consiste em analisar as raízes do polinômio $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$. O processo será estacionário se as soluções da equação $\Phi(B) = 0$ estiverem fora do círculo unitário.

O processo ARMA(p,q) é dito **invertível** se puder ser reescrito com função das observações passadas, isto é, se existirem constantes $\{\pi_j\}$ tais que $\sum_{j=0}^{\infty} |\pi_j| < \infty$ e

$$a_t = \frac{\Phi(B)}{\Theta(B)} X_t = \pi(B) X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \text{ para todo } t. \quad (1.5)$$

A equação (1.5) é a representação do processo ARMA(p,q) escrito na forma de um processo autorregressivo de ordem infinita $AR(\infty)$. De forma análoga, pode-se mostrar que a condição para que um processo ARMA(p,q) seja invertível é que as raízes da equação $\Theta(B) = 0$ estejam fora do círculo unitário.

1.2.2 Identificação da ordem do Modelo ARMA

Segundo a metodologia de [Box e Jenkins \(1970\)](#), a modelagem de uma série temporal baseia-se no seguinte processo iterativo:

1. Identificação da ordem do modelo;
2. Estimação dos parâmetros;
3. Diagnóstico do modelo.

Se na fase de checagem os resíduos do modelo estimado não constituírem um ruído branco o processo é reiniciado. Antes de iniciar a fase de identificação da ordem do modelo é preciso verificar se a série é estacionária. Se os dados apresentarem comportamento não estacionário (com tendência e/ou sazonalidade, por exemplo), é necessário aplicar uma transformação até que a condição de estacionariedade seja satisfeita ([MORETTIN; TOLOI, 2006](#)).

A tendência e sazonalidade podem ser identificadas através do gráfico da série temporal ou pela função de autocorrelação. A FAC tem decaimento lento em caso de dados com tendência e comportamento periódico para séries com sazonalidade. Para eliminar estes efeitos podemos aplicar a diferenciação ou decompor a série em componentes de sazonalidade, tendência e resíduo aleatório ([MORETTIN; TOLOI, 2006](#)). Em alguns casos pode ser necessário estabilizar a variância. Uma forma de identificar a heterocedasticidade é quando a amplitude dos valores da série vai aumentando ao longo do tempo. O método mais conhecido para estabilizar a variância é a transformação de [Box e Cox \(1964\)](#):

$$y_t = \begin{cases} \frac{x_t^\lambda - c}{\lambda} & \text{se } \lambda \neq 0, \\ \log(x_t) & \text{se } \lambda = 0. \end{cases}$$

Em que λ e c são parâmetros a serem estimados. A transformação logarítmica é apropriada se o desvio padrão da série (ou outra medida de dispersão) for proporcional à média.

Nos estudos financeiros é comum que seja utilizada a transformação log retorno ao invés da série de preços dos ativos. A vantagem de se trabalhar com o log retorno é que a série fica livre de escala e tem propriedades estatísticas mais interessantes (como estacionariedade e heterocedasticidade condicional). Seja P_t o preço de um ativo no instante t , o retorno líquido simples entre os instantes $t - 1$ e t é dado por:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1 \Rightarrow 1 + R_t = \frac{P_t}{P_{t-1}}.$$

Dizemos que $1 + R_t$ é o retorno bruto simples, geralmente R_t é expressado em percentual, denominado taxa de retorno em relação ao período t . Quando os retornos são pequenos, ou seja, os preços nos instantes $t - 1$ e t são próximos, temos que $R_t \ll 1$ e $\ln(1 + R_t) \approx R_t$. Portanto, o log retorno é definido por:

$$r_t = \log \frac{P_t}{P_{t-1}} = \ln(1 + R_t) = \ln(P_t) - \ln(P_{t-1}) \approx R_t. \quad (1.6)$$

Concluída a fase de verificação dos dados, podemos dar início à identificação da ordem do modelo. Segundo [Morettin e Tolo \(2006\)](#), as ferramentas mais utilizadas nesta fase são os gráficos das funções de autocorrelação e autocorrelação parcial. A partir da comparação do comportamento das FAC e FACP amostrais com o comportamento teórico destas funções, elegemos a ordem para o modelo. As Figuras 2, 3 e 4 exemplificam o comportamento teórico dos modelo AR(p), MA(q) e ARMA(p,q) através de séries simuladas.

Um resultado importante apresentado por [Box e Jenkins \(1970\)](#) é que para um processo AR(p) a FAC decai exponencialmente e a FACP apresenta um “corte” após o lag $h > p$. Na Figura 2 apresentamos a FAC amostral e a FACP amostral de um AR(1) simulado.

Para um processo MA(q), a FAC apresenta um “corte” após o lag $h > q$ e a FACP decai exponencialmente. Na Figura 3 apresentamos a FAC amostral e a FACP amostral de um MA(2) simulado.

Finalmente, para um processo ARMA(p,q), a FAC é infinita em extensão com decaimento exponencial e/ou senoides amortecidas após o lag $q - p$ enquanto a FACP se comporta como a de um MA puro. Na Figura 4 apresentamos a FAC amostral e a FACP amostral de um ARMA(2,2) simulado.

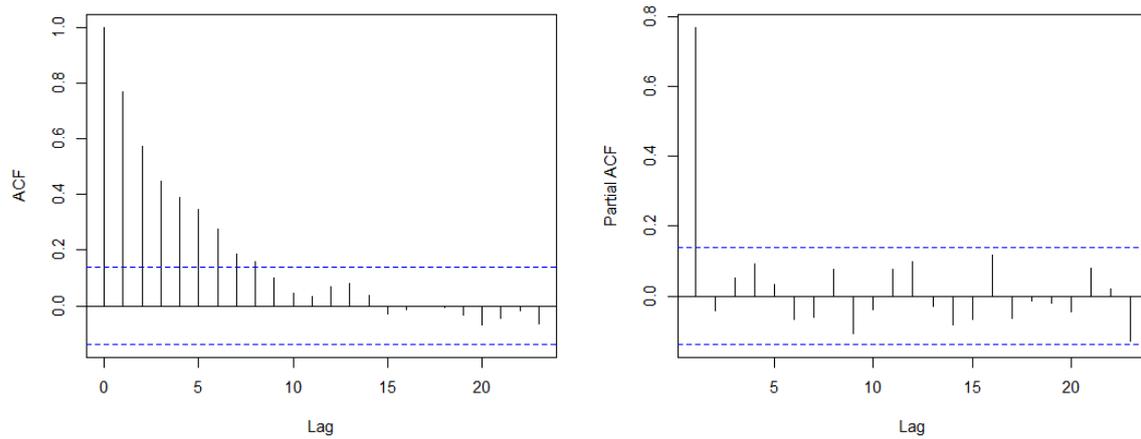


Figura 2 – FAC e FACP amostrais de um AR(1) simulado.

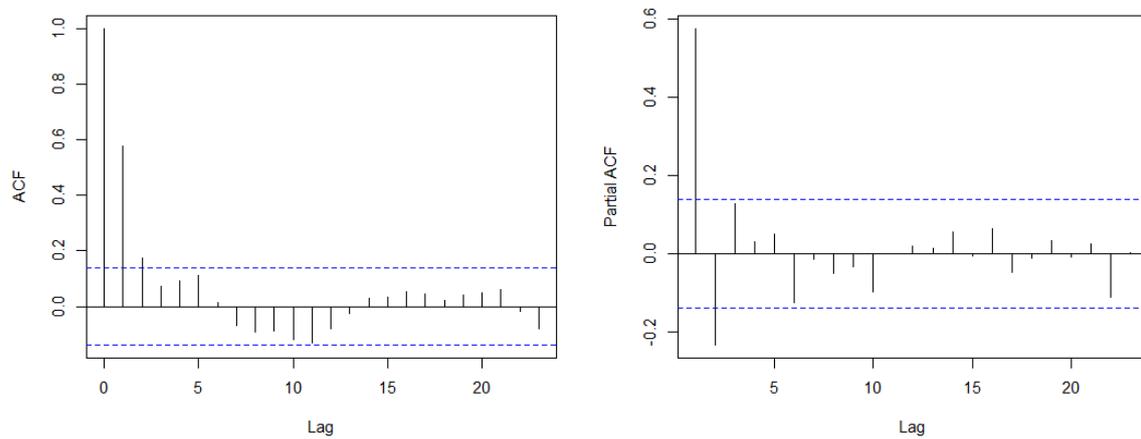


Figura 3 – FAC e FACP amostrais de um MA(2) simulado.

A identificação da ordem de um modelo ARMA(p,q) se torna mais subjetiva uma vez que a diferença entre o número de parâmetros $q - p$ pode ser satisfeita para diferentes valores (por exemplo os modelos ARMA(2,1) e ARMA(3,2) possuem a mesma diferença). Uma saída para este problema é estimar modelos ARMA de baixa ordem (de modo a preservar a parcimônia do modelo) e adicionar algum critério de seleção que tenha como base a verossimilhança. Recentemente os critérios AIC (*Akaike Information Criterion*) e BIC (*Bayes Information Criteria*) têm sido utilizados para auxiliar nesta tarefa. O AIC é definido como:

$$AIC(p,q) = \ln(L_p) + 2((p + q + 1) + 1), \quad (1.7)$$

em que L_p é a função de máxima verossimilhança do modelo. O objetivo é ajustar o modelo com AIC mínimo. [Hurvich e Tsai \(1989\)](#) propõem uma correção para AIC

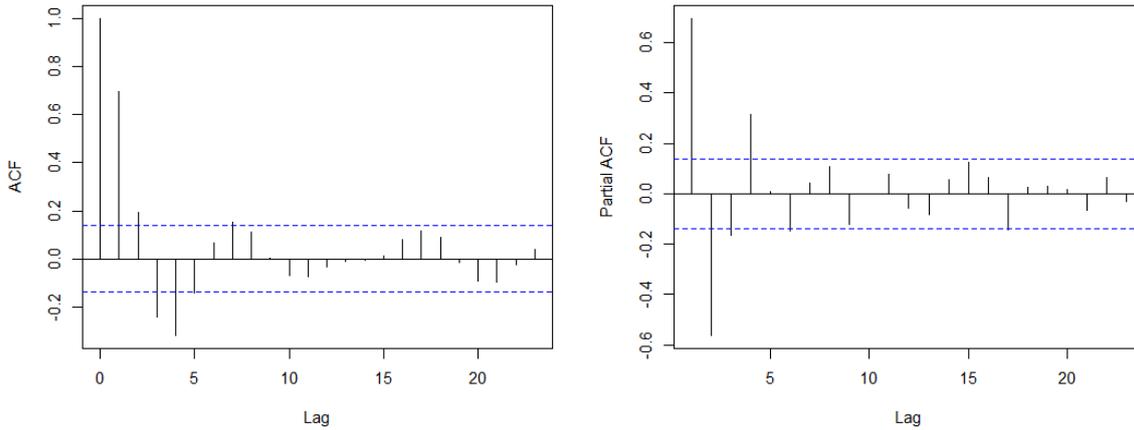


Figura 4 – FAC e FACP amostrais de um ARMA(2,2) simulado.

a fim de melhorar seu comportamento no sentido de diminuir a probabilidade de selecionar uma ordem maior que a verdadeira. Esta correção é dada por

$$AIC_c(p,q) = AIC(p,q) + \frac{2(p+q+1)(p+q+2)}{N - (p+q+1) - 1}. \quad (1.8)$$

Simulações mostram que esta correção é vantajosa quando o número de observações é pequeno ou quando a soma $p+q$ é uma fração “moderadamente grande” do tamanho da amostra. Akaike (1977), Rissanen (1978) e Schwarz et al. (1978) sugerem minimizar o Critério de Informação Bayesiana (BIC), dado por

$$BIC(p,q) = -2 \ln(L_p) + (p+q+1) \ln(N), \quad (1.9)$$

em que L_p é a função de máxima verossimilhança do modelo e N é o número de observações. Hannan (1980) e Hannan (1982) mostram que, sob determinadas condições, as estimativas de p e q que minimizam a expressão (1.9) são fortemente consistentes.

Entre um conjunto de modelos, vamos selecionar valores de p e q para o modelo ajustado que minimiza os valores de $AIC(p,q)$ e $BIC(p,q)$. Há uma evidência empírica de que o critério AIC está propenso para a escolher modelos superparametrizados enquanto o critério BIC tende a corrigir este problema (BROCKWELL; DAVIS, 1991).

1.2.3 Estimação dos parâmetros do Modelo ARMA

O próximo passo na análise de modelos após a identificação da ordem é a estimação dos parâmetros. Precisamos estimar $\Phi = (\phi_1, \phi_2, \dots, \phi_p)$, $\Theta = (\theta_1, \theta_2, \dots, \theta_q)$

e σ^2 , sendo assim, temos $p + q + 1$ parâmetros assumindo que $\mu = 0$. Serão apresentados os dois principais métodos de estimação para parâmetros de modelos de séries temporais: o método de momentos e o de máxima verossimilhança.

1.2.3.1 Método dos Momentos

Este é método mais simples para obter as estimativas de (Φ, Θ, σ^2) e consiste em substituir nas equações que relacionam as autocorrelações (ou autocovariâncias) aos parâmetros do modelo e os momentos teóricos pelos correspondentes momentos amostrais e resolver as equações restantes. Como os estimadores deste método são menos eficientes para os modelos MA(q) e ARMA(p,q) (BROCKWELL; DAVIS, 1991), serão apresentados apenas os estimadores para um modelo AR(p) com média zero:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + a_t. \quad (1.10)$$

Utilizando as equações que relacionam as funções de autocovariância e os parâmetros do AR(p), temos:

$$\begin{aligned} \gamma(0) &= \phi_1 \gamma(1) + \dots + \phi_p \gamma(p) + \sigma^2, \\ \gamma(h) &= \phi_1 \gamma(h-1) + \dots + \phi_p \gamma(h-p), h = 1, 2, \dots, p. \end{aligned}$$

Escrevendo as expressões em notação matricial segue que

$$\begin{aligned} \sigma^2 &= \gamma(0) - \Phi^T \gamma_{\mathbf{p}}, \\ \mathbf{\Gamma}_{\mathbf{p}} \Phi &= \gamma_{\mathbf{p}}. \end{aligned}$$

onde

$$\begin{aligned} \mathbf{\Gamma}_{\mathbf{p}} &= [\gamma(i-j)]_{i,j=1,\dots,p}, \\ \Phi &= (\phi_1, \phi_2, \dots, \phi_p)^T, \\ \gamma_{\mathbf{p}} &= (\gamma(1), \dots, \gamma(p))^T. \end{aligned}$$

Substituindo $\gamma(h)$ pela função de autocovariância amostral, $\hat{\gamma}(h)$, obtemos a solução:

$$\begin{aligned} \hat{\sigma}^2 &= \hat{\gamma}(0) - \hat{\gamma}_{\mathbf{p}}^T \hat{\mathbf{\Gamma}}_{\mathbf{p}}^{-1} \hat{\gamma}_{\mathbf{p}}, \\ \hat{\Phi} &= \hat{\mathbf{\Gamma}}_{\mathbf{p}}^{-1} \hat{\gamma}_{\mathbf{p}}. \end{aligned} \quad (1.11)$$

Estas equações são conhecidas como *estimadores de Yule-Walker* e podem ser expressas em termos da função de autocorrelação:

$$\begin{aligned} \hat{\sigma}^2 &= \hat{\gamma}(0)(1 - \hat{\rho}_{\mathbf{p}}^T \hat{\mathbf{R}}_{\mathbf{p}}^{-1} \hat{\rho}_{\mathbf{p}}), \\ \hat{\Phi} &= \hat{\mathbf{R}}_{\mathbf{p}}^{-1} \hat{\rho}_{\mathbf{p}}. \end{aligned} \quad (1.12)$$

onde

$$\hat{\mathbf{R}}_{\mathbf{p}} = [\hat{\rho}(i-j)]_{i,j=1,2,\dots,p}$$

é a matriz das autocorrelações amostral, e

$$\hat{\rho}_{\mathbf{p}} = (\hat{\rho}(1), \dots, \hat{\rho}(p))^T$$

é o vetor de autocorrelações amostrais. As soluções das equações (1.11) e (1.12) podem ser obtidas utilizando o algoritmo de Durbin-Lewison (MORETTIN; TOLOI, 2006).

Proposição 1.1. *A distribuição dos parâmetros do modelo causal AR(p) descrito em (1.10) converge assintoticamente (quando $n \rightarrow \infty$) em distribuição para uma normal:*

$$\sqrt{n}(\hat{\Phi} - \Phi) \xrightarrow{d} N(0, \sigma^2 \mathbf{\Gamma}_{\mathbf{p}}^{-1}) \text{ e } \hat{\sigma}^2 \xrightarrow{p} \sigma^2.$$

Em muitos casos, a variância dos estimadores de momentos é muito maior do que a obtida em outros métodos de estimação, como o de máxima verossimilhança, por exemplo. No entanto, os estimadores de Yule-Walker para Φ do modelo AR(p) tem aproximadamente a mesma distribuição assintótica dos estimadores de máxima verossimilhança para o modelo correspondente (BROCKWELL; DAVIS, 1991).

1.2.3.2 Método de Máxima Verossimilhança

Seja $\boldsymbol{\theta} = (\Phi, \Theta, \sigma^2)$ o vetor dos parâmetros populacionais do processo ARMA(p,q) descrito na equação (1.2) e suponha que a distribuição da variável aleatória $\mathbf{X} = (X_1, \dots, X_n)^T$ seja gaussiana. A função de verossimilhança é dado por:

$$L(\boldsymbol{\theta}|x_1, \dots, x_n) = (2\pi)^{-n/2} |\mathbf{\Gamma}_{\mathbf{n}}|^{-1/2} \exp\left\{-\frac{1}{2} \mathbf{X}^T \mathbf{\Gamma}_{\mathbf{n}}^{-1} \mathbf{X}\right\}, \quad (1.13)$$

em que, $\mathbf{\Gamma}$ é a matriz de autocovariâncias e não singular. Tomando o logaritmo de $L(\boldsymbol{\theta}|x_1, \dots, x_n)$, obtemos

$$\ell(\boldsymbol{\theta}|x_1, \dots, x_n) \propto -\frac{1}{2} \ln |\mathbf{\Gamma}_{\mathbf{n}}| - \frac{1}{2} \mathbf{X}^T \mathbf{\Gamma}_{\mathbf{n}}^{-1} \mathbf{X}. \quad (1.14)$$

Os estimadores de máxima verossimilhança dos parâmetros (Φ, Θ, σ^2) são os valores que maximizam a função $\ell(\Phi, \Theta, \sigma^2|x_1, \dots, x_n)$.

1.2.3.3 Método de Máxima Verossimilhança Condicional

Considere o processo ARMA(p,q) descrito na equação (1.2), onde todas as raízes de $\Phi(B) = 0$ e $\theta(B) = 0$ estão fora do disco unitário e $a_t \sim N(0, \sigma^2)$. O método de estimação por máxima verossimilhança condicional consiste na aproximação da função de verossimilhança condicionado aos valores iniciais de x_t 's e a_t 's. Onde a $(p+1)$ -ésima observação é igual a

$$X_{p+1} = \phi_1 X_p + \phi_2 X_{p-1} + \dots + \phi_p X_1 + a_{p+1} + \theta_1 a_p + \dots + \theta_q a_{p-q+1}. \quad (1.15)$$

Condicionado a $X_1 = x_1, X_2 = x_2, \dots, X_p = x_p$ e tomando $a_p = a_{p-1} = \dots = a_{p-q+1} = 0$ temos

$$Y_{p+1} \sim N((\phi_1 Y_p + \phi_2 Y_{p-1} + \dots + \phi_p Y_1), \sigma^2).$$

Então a verossimilhança condicional calculada de $t = p + 1, \dots, n$ é dado por

$$\begin{aligned} \ell^*(\boldsymbol{\theta}) &= \log f(x_n, x_{n-1}, \dots, x_{p+1} | x_p, \dots, y_1, a_p = a_{p-1} = \dots = a_{p-q+1} = 0; \boldsymbol{\theta}) \\ &= -\frac{n-p}{2} \log(2\pi) - \frac{n-p}{2} \log(\sigma^2) - \sum_{t=p+1}^n \frac{a_t^2}{2\sigma^2}, \end{aligned} \quad (1.16)$$

onde a sequência $\{a_{p+1}, a_{p+2}, \dots, a_n\}$ pode ser calculada a partir de $\{x_1, x_2, \dots, x_n\}$ na forma iterativa

$$a_t = x_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} - \theta_1 a_{t-1} - \dots + \theta_q a_{t-q},$$

para $t = p + 1, p + 2, \dots, n$. Observe que a partir de uma amostra de tamanho n a estimativa de um ARMA(p,q) pelo método de máxima verossimilhança condicional utiliza apenas $n - p$ observações da amostra.

A partir da equação (1.16), temos que a função de log verossimilhança é dado por

$$-\frac{n^*}{2} \log(2\pi) - \frac{n^*}{2} \log(\sigma^2) - \sum_{t=t^*}^n \frac{a_t^2}{2\sigma^2},$$

onde $n^* = n - p$ e $t^* = p + 1$.

Apresentamos este método no caso dos erros com distribuição gaussiana mas é possível aplicá-lo para outras distribuições, como para o modelo ARMA(p,q) com inovações α -estáveis.

2 Distribuições estáveis

Nos estudos sobre séries temporais, a base teórica para estimação da ordem e parâmetros de modelos gaussianos está consolidada. Porém, em diversas áreas, os dados empíricos não apresentam as características de curva simétrica e mesocúrtica como no caso normal. Em áreas como economia e finanças ([EMBRECHTS et al., 1997](#)), processamento de sinal ([NIKIAS; SHAO, 1995](#)) e engenharia telegráfica ([RESNICK et al., 1997](#)) os dados são assimétricos e/ou com caudas pesadas. Nestes casos, os métodos aplicados para dados gaussianos podem não produzir resultados próximos a realidade e satisfatórios.

Entre as décadas de 1920 e 1930, Paul Lévy e Aleksander Y. Khinchine desenvolveram a teoria das distribuições estáveis. Esta família de distribuições, possui grande versatilidade quanto a forma da função de densidade de acordo com a variação de seus quatro parâmetros.

2.1 Definições

Em [Samorodnitsky e Taqqu \(1994\)](#) são apresentadas quatro definições equivalentes para a distribuição estável. As duas primeiras discorrem sobre a propriedade de estabilidade, ou seja, as propriedades da família de distribuições estáveis é preservada sob convolução. Na terceira definição, as distribuições estáveis são escritas no contexto do Teorema Limite Central. Esta aproximação das distribuições estáveis por uma soma normalizada de variáveis aleatórias i.i.d. é muito útil nos resultados de modelagem de muitos efeitos aleatórios pequenos. Por fim, a última definição apresenta a função característica das distribuições estáveis.

Definição 2.1. *Dizemos que X tem distribuição estável se para qualquer valor de A e B positivos existem um valor positivo C e um número real D tais que:*

$$AX_1 + BX_2 \stackrel{d}{=} CX + D, \quad (2.1)$$

onde X_1 e X_2 são cópias independentes de X e $\stackrel{d}{=}$ denota igualdade em distribuição.

Uma variável aleatória X é chamada de *estritamente estável* se $D = 0$ na equação (2.1). Uma variável aleatória estável X é *simétrica* se sua distribuição é simétrica, isto é, se X e $-X$ tem a mesma distribuição. Segue que, uma distribuição estável simétrica tem que ser estritamente estável.

Teorema 2.1. *Para qualquer variável aleatória estável X , existe um número $0 < \alpha \leq 2$ tal que C na equação (2.1) satisfaz:*

$$C^\alpha = A^\alpha + B^\alpha, \quad (2.2)$$

onde α é chamado índice de estabilidade. Uma variável aleatória estável X com índice α é chamada de α -estável.

A demonstração do Teorema 2.1 pode ser encontrada em Feller (1971), capítulo VI.1.

Exemplo 2.1. *Seja X uma variável aleatória gaussiana com média μ e variância σ^2 , então X é estável com $\alpha = 2$ pois*

$$AX_1 + BX_2 \sim N((A+B)\mu, (A^2 + B^2)\sigma^2),$$

ou seja, segundo o Teorema 2.1, $C = (A^2 + B^2)^{1/2}$ e $D = (A + B - C)$.

Segue a segunda definição para variável estável.

Definição 2.2 (equivalente a Definição 2.1). *Dizemos que a variável aleatória X tem distribuição estável se para qualquer $n \geq 2$ existem um número positivo c_n e um número real d_n , tais que:*

$$X_1 + X_2 + \cdots + X_n \stackrel{d}{=} C_n X + D_n, \quad (2.3)$$

onde $X_1 + X_2 + \cdots + X_n$ são cópias independentes de X .

O resultado da Definição 2.2 pode ser encontrado a partir da Definição 2.1 por indução. É fácil mostrar que a implicação reversa também é verdadeira (ver Feller (1971), Teorema VI.1). Assim como na Definição 2.1, se a constante $d_n = 0$, dizemos que X é estritamente estável. A próxima definição declara que as distribuições estáveis são as únicas que podem ser obtidas na forma de limites de somas normalizadas de variáveis aleatórias i.i.d.

A partir de Feller (1971), Teorema VI.1.1, temos que na equação (2.3), necessariamente, $C_n = n^{1/\alpha}$ para algum $0 < \alpha \leq 2$. Este α é o mesmo que aparece na equação (2.2).

Definição 2.3 (equivalente as Definições 2.1 e 2.2). *Dizemos que a variável aleatória X tem distribuição estável se ela tem domínio de atração, isto é, se existe uma sequência de variáveis i.i.d. Y_1, Y_2, \dots e uma sequência de números positivos $\{d_n\}$ e uma sequência de valores reais $\{a_n\}$, tais que:*

$$\frac{Y_1 + Y_2 + \cdots + Y_n}{d_n} + a_n \xrightarrow{d} X, \quad (2.4)$$

onde \xrightarrow{d} denota convergência em distribuição.

Observe que quando a v.a. X tem distribuição Gaussiana e a sequência Y_1, Y_2, \dots é i.i.d com variância finita, a Definição 2.3 equipara à declaração do Teorema Limite Central. Quando $d_n = n^{1/\alpha}$, dizemos que a sequência de variáveis i.i.d. Y_1, Y_2, \dots pertencem ao domínio de atração *normal* de X . Em geral, $d_n = n^{1/\alpha}h(n)$ onde $h(x)$, para $x > 0$ é uma função que varia lentamente no infinito, isto é, $\lim_{x \rightarrow \infty} h(ux)/h(x) = 1$ para todo $u > 0$ (FELLER, 1971, XVII.5). A função $h(x) = \ln(x)$ varia lentamente no infinito, por exemplo.

Por não possuir uma representação simples para a função de distribuição, a quarta definição das distribuições estáveis apresentada por Samorodnitsky e Taquq (1994) é a mais utilizada e corresponde a sua representação em termos da função característica.

Definição 2.4 (equivalente as Definições 2.1, 2.2 e 2.3). *Dizemos que a variável aleatória X tem distribuição α -estável ou estável se existem os parâmetros $\alpha \in (0,2]$, $\beta \in [-1,1]$, $\sigma > 0$ e $\mu \in \mathbb{R}$ tais que a função característica de X tem a forma:*

$$\ln \varphi_X(t) = \begin{cases} i\mu t - \sigma^\alpha |t|^\alpha [1 - i\beta \text{ sinal}(t) \tan \frac{\pi\alpha}{2}], & \text{se } \alpha \neq 1, \\ i\mu t - \sigma |t| [1 + i\beta \frac{2}{\pi} \text{ sinal}(t) \ln |t|], & \text{se } \alpha = 1, \end{cases} \quad (2.5)$$

onde $\alpha \in (0,2]$, $\beta \in [-1,1]$, $\sigma > 0$ e $\mu \in \mathbb{R}$. E a função $\text{sinal}(t)$ é definida por:

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

Demonstrar que a Definição 2.4 implica na Definição 2.2 é uma tarefa simples e pode ser consultada em Gnedenko e Kolmogorov (1954), Seção 34.

A notação $X \sim S_\alpha(\sigma, \beta, \mu)$ é adotada quando X tem distribuição α -estável e será utilizada no decorrer do trabalho.

A função densidade da distribuição α -estável coincide com distribuições conhecidas quando seus parâmetros assumem determinados valores. Estas distribuições são apresentadas nos exemplos a seguir.

Exemplo 2.2. *Seja a variável aleatória $X \sim S_\alpha(\sigma, \beta, \mu)$, para $\alpha = 1/2$ e $\beta = \pm 1$, a equação (2.5) é igual a*

$$\phi(t) = \exp\{i\mu t - \sqrt{\sigma|t|}(1 - \text{sinal}(t))\},$$

que coincide com a função característica de uma distribuição Lévy com parâmetros μ e σ .

Exemplo 2.3. Nos casos em que $\beta = 0$ e $\alpha = 1$, a equação (2.5) é igual a

$$\phi(t) = \exp\{i\mu t - \sigma|t|\},$$

que coincide com a função característica de uma distribuição Cauchy com parâmetros μ e σ .

Exemplo 2.4. Quando fazemos $\beta = 0$ e $\alpha = 2$,

$$\phi(t) = \exp\{i\mu t - \sigma^2 t^2\},$$

que coincide com a função característica de uma distribuição normal com média μ e variância $2\sigma^2$.

2.2 Parâmetros da distribuição α -estável

Cada parâmetro da distribuição α -estável exerce papel diferente na curva da função de densidade. O índice de estabilidade denotado por $\alpha \in (0,2]$ mensura o peso das caudas (mais pesadas para valores pequenos). O parâmetro $\beta \in [-1,1]$ determina o índice de assimetria, sendo que para $\beta < 0$ ($\beta > 0$) a distribuição é assimétrica à esquerda (direita) e simétrica quando $\beta = 0$. O parâmetro de escala é denotado por σ , onde $\sigma > 0$ e $\mu \in \mathbb{R}$ é o parâmetro de locação.

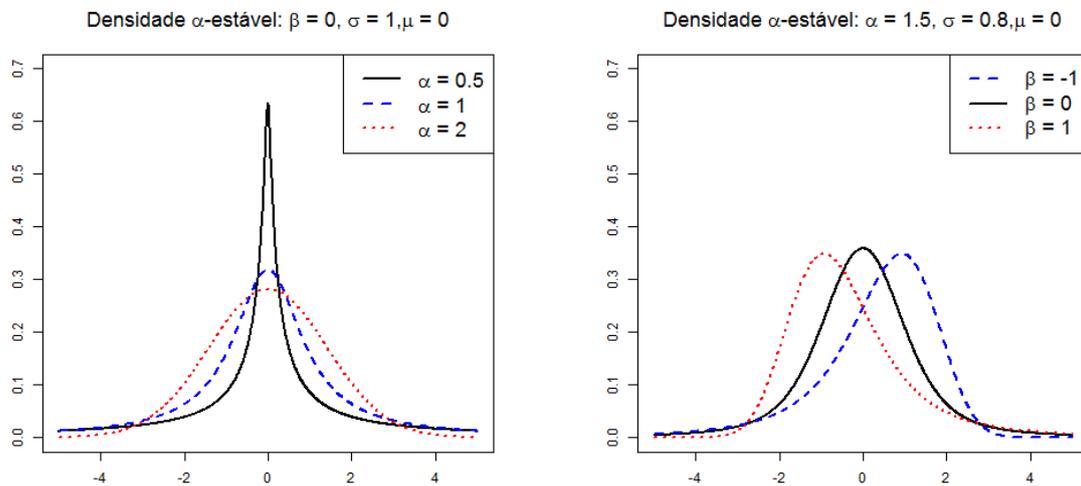


Figura 5 – Densidade da α -estável variando o índice de estabilidade o e parâmetro de assimetria.

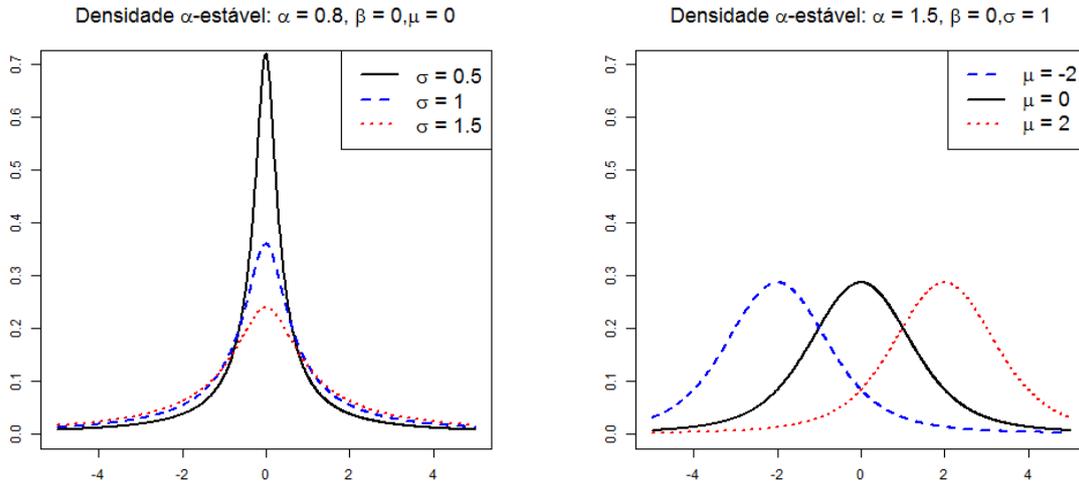


Figura 6 – Densidade da α -estável variando o parâmetro de escala e locação.

2.3 Propriedades das variáveis aleatórias α -estáveis

Ilustrado os efeitos de cada parâmetro na forma da distribuição estável, esta seção consiste na apresentação de algumas propriedades relacionadas aos parâmetros da α -estável

Propriedade 2.1. *Sejam as variáveis aleatórias X_1 e X_2 independentes, onde $X_i \sim S_\alpha(\sigma_i, \beta_i, \mu_i)$, $i = 1, 2$. Então $X_1 + X_2 \sim S_\alpha(\sigma, \beta, \mu)$, onde:*

$$\begin{aligned}\sigma &= (\sigma_1^\alpha + \sigma_2^\alpha)^{1/\alpha}, \\ \beta &= \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}, \\ \mu &= \mu_1 + \mu_2.\end{aligned}$$

Dem. *Para o caso $\alpha \neq 1$. Por independência:*

$$\begin{aligned}\ln [E \exp it(X_1 + X_2)] &= \ln(E \exp itX_1) + \ln(E \exp itX_2) \\ &= -(\sigma_1^\alpha + \sigma_2^\alpha) |t|^\alpha + i |t|^\alpha \operatorname{sinal}(t) \tan\left(\frac{\pi\alpha}{2}\right) (\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha) + it(\mu_1 + \mu_2) \\ &= -(\sigma_1^\alpha + \sigma_2^\alpha) |t|^\alpha \left[1 - i \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha} \operatorname{sinal}(t) \tan\left(\frac{\pi\alpha}{2}\right) + it(\mu_1 + \mu_2) \right].\end{aligned}$$

A demonstração para $\alpha = 1$ é análoga. ■

Propriedade 2.2. *Sejam $X \sim S_\alpha(\sigma, \beta, \mu)$ e c uma constante real, então $X + c \sim S_\alpha(\sigma, \beta, \mu + c)$.*

Dem. *Este resultado pode ser encontrado facilmente através da função característica*

descrita na Definição 2.4.

$$\begin{aligned} \ln [E \exp it(X + c)] &= \ln(E \exp itX) + \ln(E \exp itc) \\ &= \begin{cases} it(\mu + c) - \sigma^\alpha |t|^\alpha [1 - i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2}], & \text{se } \alpha \neq 1, \\ it(\mu + c) - \sigma |t| [1 + i\beta \frac{2}{\pi} \operatorname{sign}(t) \ln |t|], & \text{se } \alpha = 1. \end{cases} \end{aligned}$$

■

Propriedade 2.3. *Sejam $X \sim S_\alpha(\sigma, \beta, \mu)$ e c uma constante diferente de zero. Então:*

$$\begin{cases} cX \sim S_\alpha(|c|\sigma, \operatorname{sign}(c)\beta, c\mu) & \text{se } \alpha \neq 1 \\ cX \sim S_1(|c|\sigma, \operatorname{sign}(c)\beta, c\mu - \frac{2}{\pi}c(\ln |c|)\sigma\beta) & \text{se } \alpha = 1. \end{cases}$$

Dem. Por (2.5), temos que para $\alpha \neq 1$,

$$\begin{aligned} \ln [E \exp it(cX)] &= -|tc|^\alpha \sigma^\alpha \left(1 - i\beta \operatorname{sign}(ct) \tan\left(\frac{\pi\alpha}{2}\right)\right) + i(\mu c)t \\ &= (\sigma |c|)^\alpha |t|^\alpha (1 - i\beta \operatorname{sign}(c) \operatorname{sign}(t) \tan\left(\frac{\pi\alpha}{2}\right)) + i(\mu c)t. \end{aligned}$$

A demonstração para $\alpha = 1$ é análoga. ■

Observe que quando $\alpha = 1$, a multiplicação de uma constante afeta o parâmetro de locação de forma não-linear e passa a depender dos parâmetros σ e β . Neste caso, σ deixa de ter a função de parâmetro de escala.

Propriedade 2.4. *Seja uma variável aleatória com distribuição α -estável tal que o parâmetro $\mu = 0$, temos que para qualquer $0 < \alpha \leq 2$,*

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

Propriedade 2.5. *Uma variável aleatória $X \sim S_\alpha(\sigma, \beta, \mu)$ é simétrica se e somente se $\beta = 0$ e $\mu = 0$. Dizemos que X é simétrica em relação a μ se e somente se $\beta = 0$.*

Dem. Para um variável aleatória ser simétrica, é necessário e suficiente que sua função característica seja real. Por (2.5), isto ocorre se e somente se $\beta = 0$ e $\mu = 0$. Da Propriedade 2.2 segue a segunda afirmação. ■

A notação $S_\alpha S$ é adotada quando a variável aleatória α -estável é simétrica, ou seja, $X \sim S_\alpha(\sigma, 0, 0)$. Em adição, quando uma variável é simétrica, a função característica assume apenas valores reais.

Propriedade 2.6. *Seja $X \sim S_\alpha(\sigma, \beta, \mu)$ e $\alpha \neq 1$. Então X é estritamente estável se e somente se $\mu = 0$.*

Dem. *Sejam X_1 e X_2 cópias independentes de X e A e B , constantes positivas. Usando as propriedades 2.1 e 2.3*

$$AX_1 + BX_2 \sim S_\alpha(\sigma(A^\alpha + B^\alpha)^{1/\alpha}, \beta, \mu(A + B)).$$

Definimos $C = (A^\alpha + B^\alpha)^{1/\alpha}$ em (2.1). Pelas propriedades 2.2 e 2.3,

$$CX + D \sim S_\alpha(\sigma(A^\alpha + B^\alpha)^{1/\alpha}, \beta, \mu(A^\alpha + B^\alpha)^{1/\alpha} + D),$$

e, portanto, temos que $AX_1 + BX_2 \stackrel{d}{=} CX + D$ com $D = 0$ se e somente se $\mu = 0$. ■

Para $\mu \neq 0$, a variável $Y = X - \mu$ é estritamente estável quando $\alpha \neq 1$. Para $\alpha = 1$ consultar Samorodnitsky e Taqqu (1994)

Propriedade 2.7. *Seja $X \sim S_\alpha(\sigma, \beta, \mu)$ com $0 < \alpha \leq 2$. Então:*

$$E|X|^p < \infty \text{ para qualquer } 0 < p < \alpha,$$

$$E|X|^p = \infty \text{ para qualquer } p \geq \alpha.$$

Portanto, a variável aleatória $X \sim S_\alpha(\sigma, \beta, \mu)$ tem esperança finita apenas para $1 < \alpha \leq 2$ e segundo momento finito para $\alpha = 2$. Perceba que não é possível calcular medidas como covariância e correlação para $\alpha < 2$, pois estas medidas dependem da variância finita.

Propriedade 2.8. *Quando $1 < \alpha \leq 2$ o parâmetro de locação é igual a média.*

Propriedade 2.9. *Seja $X \sim S_\alpha(1, \beta, 0)$. Então*

$$Z = \begin{cases} \sigma X + \mu & \text{se } \alpha \neq 1, \\ \sigma X + \frac{2}{\pi} \beta \sigma \log \sigma + \mu & \text{se } \alpha = 1, \end{cases}$$

tem distribuição $S_\alpha(\sigma, \beta, \mu)$.

2.4 Estimação dos parâmetros

Nesta seção serão apresentados dois métodos de estimação que contornam o problema da difícil representação de uma fórmula fechada para a função de densidade da variável α -estável. O primeiro método apresentado depende apenas dos quantis e o segundo é derivado a partir da função característica.

A primeira contribuição nos métodos de estimação dos parâmetros para a distribuição estável foi de Fama e Roll (1968). O método era baseado nos quantis e restrito às distribuições α -estáveis simétricas ($\beta = 0, \mu = 0$) quando $\alpha > 1$. A extensão

deste método para a estimação de todos os quatro parâmetros de uma variável α -estável foi proposta por [McCulloch \(1986\)](#). Neste método, são definidas as quantidades

$$\nu_\alpha = \frac{q_{.95} - q_{.05}}{q_{.75} - q_{.25}}$$

e

$$\nu_\beta = \frac{q_{.95} + q_{.05} - 2q_{.5}}{q_{.95} - q_{.05}}.$$

Observe que as quantidades ν_α e ν_β não dependem de σ e μ . A estimação dos parâmetros α e β é realizada a partir da comparação dos valores tabelados para as respectivas quantidades teóricas. A estimação do parâmetro de escala σ deriva da quantidade

$$\nu_\sigma = \frac{q_{.75} - q_{.25}}{\sigma}.$$

Esta quantidade foi tabulada como função de α e β , então

$$\hat{\sigma} = \frac{\hat{q}_{.75} - \hat{q}_{.25}}{\nu_\sigma(\hat{\alpha}, \hat{\beta})}.$$

O parâmetro de locação é o último a ser estimado. Considere a seguinte transformação para o parâmetro de escala

$$\zeta = \begin{cases} \mu + \beta\sigma \tan \frac{\pi\alpha}{2}, & \text{se } \alpha \neq 1, \\ \mu, & \text{se } \alpha = 1. \end{cases}$$

A quantidade

$$\nu_\zeta = \frac{\zeta - q_{.5}}{\sigma},$$

foi tabulada como função de α e β , então $\hat{\zeta} = q_{.5} + \hat{\sigma}\nu_\zeta(\hat{\alpha}, \hat{\beta})$ e

$$\hat{\mu} = \hat{\zeta} - \hat{\beta}\hat{\sigma} \tan \frac{\pi\alpha}{2}.$$

Os estimadores apresentados são consistentes e tem boa precisão ([ADLER et al., 1998](#)). Porém, a estimação é restrita para $\alpha \in [0.6, 2]$ e $\beta \in [-1, 1]$. O próximo método de estimação é baseado na função característica, onde [Press \(1972\)](#) propõe diversos métodos de estimação utilizando a função característica amostral, definida para uma amostra aleatória de tamanho n por

$$\hat{\phi}(t) = \frac{1}{n} \sum_{j=1}^n \exp(itx_j). \quad (2.6)$$

O método exposto em [Press \(1972\)](#), tem simples aplicação mas depende da escolha apropriada dos valores de t . Em [Koutrouvelis \(1980\)](#), foi proposto um procedimento

diferente utilizando a função característica. Observou-se que ao aplicar a função logaritmo na função característica temos:

$$\ln \phi(t) = -|\sigma t|^\alpha + i \left[\mu t + |\sigma t|^\alpha \text{sinal}(t) \beta \tan \frac{\pi \alpha}{2} \right], \text{ para } \alpha \neq 1.$$

Este resultado é separado em parte real e imaginária,

$$\Re[\ln \phi(t)] = -\sigma |t|^\alpha \text{ e } \Im[\ln \phi(t)] = \mu t + |\sigma t|^\alpha \text{sinal}(t) \beta \tan \frac{\pi \alpha}{2}.$$

Em seguida é calculada a função característica empírica (2.6), $\Re[\ln \hat{\phi}(t)]$ e $\Im[\ln \hat{\phi}(t)]$. A estimação dos parâmetros é obtida ao fazer uma regressão das partes imaginárias e reais da função característica empírica. Os pontos nos quais a função característica é calculada são determinados com base em uma tabela proposta pelo autor. Este método foi aprimorado em Koutrouvelis (1981) por meio de uma regressão GLS iterada onde as covariâncias dos erros da regressão fazem parte do peso.

2.5 Processos ARMA com inovações α -estáveis

Na Seção 1.2 do Capítulo 1, foi apresentado um processo ARMA(p,q) gaussiano. Geralmente, as inovações propostas para o processo ARMA são distribuições em que a_t é não-gaussiana com variância finita e a probabilidade de assumir valores extremos baixa (o uso da distribuição t-student, por exemplo). Neste trabalho será proposto o processo ARMA com inovações α -estáveis, com o objetivo de conseguir melhor ajuste para séries com caudas pesadas.

O processo ARMA(p,q) com inovações α -estáveis é definido por:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (2.7)$$

onde $\{a_t\} \sim S_\alpha(\sigma, \beta, \mu)$ e supomos que as inovações a_t 's são i.i.d com $0 < \alpha < 2$. As distribuições finito-dimensionais das variáveis X_t 's dependem dos coeficientes $\phi_1, \phi_2, \dots, \phi_p$ e $\theta_1, \theta_2, \dots, \theta_q$.

Considerando a equação (2.7) com coeficientes reais $\phi_0, \phi_1, \phi_2, \dots, \phi_p, \theta_0, \theta_1, \theta_2, \dots, \theta_q$, estes podem se escritos na forma dos polinômios

$$\begin{aligned} \Phi(z) &= 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p, \\ \Theta(z) &= 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q, \end{aligned}$$

onde z é uma variável complexa. O processo ARMA(p,q) com inovações α -estáveis da equação (2.7) pode ser expressado por:

$$\Phi(B)X_t = \Theta(B)a_t, \quad n = \dots, -1, 0, 1, \dots, \quad (2.8)$$

onde B é o operador retroativo, ou seja, $B^j X_t = X_{t-j}$. No caso do modelo gaussiano, solucionamos a equação (2.7) mostrando que $X_t = \Phi(B)^{-1}\Theta(B)a_t$ é bem definido. Naturalmente, supomos que os polinômios $\Phi(z)$ e $\Theta(z)$ não tem raízes comuns.

O Teorema a seguir mostra que, assim como no caso gaussiano, a solução de um processo ARMA com inovações α -estáveis existe se, e somente se, as raízes do polinômio $\Phi(z)$ não pertencem ao disco unitário $\{z : |z| \leq 1\}$.

Teorema 2.2. *A equação (2.7) tem solução única na forma*

$$X_t = \frac{\Theta(B)}{\Phi(B)} a_t = \sum_{j=0}^{\infty} c_j a_{t-j}, \quad (2.9)$$

para todo $t \in \mathbb{Z}$. Onde os coeficientes c_j 's satisfazem $|c_j| < Q^{-1}$, $Q > 1$, se e somente se as raízes do polinômio $\Phi(z)$ não pertencem ao disco unitário $\{z : |z| \leq 1\}$. Dizemos então que a sequência $\{X_t, n \in \mathbb{Z}\}$ é estacionária e α -estável. Os c_j 's são coeficientes da expansão de séries $\frac{\Theta(B)}{\Phi(B)}$, $|z| < 1$

Dem. Suponha que $\Phi(B)$ não tenha raízes em $\{z : |z| \leq 1\}$. A função

$$C(z) = \frac{\Theta(B)}{\Phi(B)}$$

é, portanto, analítica no disco $\{z : |z| < R\}$, onde $R > 1$ é o raio de convergência da série $C(z) = \sum_{j=0}^{\infty} c_j z^j$. Dado que $1/R = \limsup_{j \rightarrow \infty} |c_j|^{1/j}$, para qualquer $1 < Q < R$, $|c_j| < Q^{-j}$ eventualmente. Usando a relação $\Phi(z)C(z) = \Theta(z)$, na qual é válida para $|z| \leq 1$, e o fato que a série $\sum_{j=0}^{\infty} c_j z^j$ converge absolutamente para $|z| \leq 1$, temos o seguinte sistema de equações:

$$\begin{cases} c_0 = 1, \\ c_1 - \phi_1 c_0 = \theta_1, \\ c_2 - \phi_1 c_1 - \phi_2 c_0 = \theta_2, \\ \vdots \\ c_q - \phi_1 c_{q-1} - \phi_2 c_{q-2} - \cdots - \phi_q c_0 = \theta_q, \\ c_s - \phi_1 c_{s-1} - \phi_2 c_{s-2} - \cdots - \phi_s c_0 = 0, s > q, \end{cases} \quad (2.10)$$

considerando que $\phi_i = 0$ se $i > p$. A partir de (2.10) temos que os coeficientes c_j 's são reais e dado que $|c_j| < Q^{-j}$ eventualmente, o processo (2.9) é bem definido e, de fato, converge absolutamente quase certamente.

Para averiguar que o processo (2.9) com os coeficientes c_j 's definidos unicamente pelo sistema de equações (2.10) satisfaz (2.7), basta utilizarmos a relação (2.10)

e o fato de que o processo (2.9) converge absolutamente quase certamente. Rearranjando os termos da forma

$$\sum_{j=0}^{\infty} c_j a_{t-j} - \phi_1 \sum_{j=0}^{\infty} c_j a_{t-1-j} - \dots - \phi_p \sum_{j=0}^{\infty} c_j a_{t-p-j}$$

produz (2.7).

Para provar a convergência, supomos que o sistema de equações (2.7) tem solução única na forma (2.9) com os coeficientes c_j 's satisfazendo $|c_j| < Q^{-j}$ eventualmente para algum $Q > 1$. Queremos mostrar que $\Phi \neq 0$ para $|z| \leq 1$. Considere a série $C(z) = \sum_{j=0}^{\infty} c_j z^j$ que, sob as condições citadas anteriormente, converge absolutamente e uniformemente no unitário $\{z : |z| \leq 1\}$. Definindo

$$\tilde{\Theta}(z) := \Phi(z)C(z) =: \sum_{j=0}^{\infty} \tilde{\theta}_j z^j, \quad |z| \leq 1 \quad (2.11)$$

obtemos

$$\begin{cases} \tilde{\theta}_0 = c_0, \\ \tilde{\theta}_s = c_s - \phi_1 c_{s-1} - \dots - \phi_p c_0, \quad s \geq 1. \end{cases} \quad (2.12)$$

Uma vez que, para qualquer t , a série $\sum_{j=0}^{\infty} c_j a_{t-j}$ converge absolutamente quase certamente, (2.7) e (2.12) implicam

$$\sum_{j=0}^{\infty} \theta_j a_{t-j} = X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \sum_{j=0}^{\infty} \tilde{\theta}_j a_{t-j}, \quad (2.13)$$

quase certamente, que por sua vez, gera $\tilde{\theta}_j = \theta_j$ para $j = 0, 1, \dots, q$ e $\tilde{\theta}_j = 0$ para $j > q$. Então $\tilde{\Theta}(z) = \Theta(z)$ e, segundo a Definição (2.11),

$$\Phi(z) = \frac{\Theta(z)}{C(z)}, \quad |z| \leq 1.$$

Como $C(z)$ é limitada por $\{z : |z| \leq 1\}$, $\Phi(z) = 0$ implica $\Theta(z) = 0$. Mas $\Phi(z)$ e $\Theta(z)$ não tem raízes em comum, então $\Phi(z) \neq 0$, para todos $|z| \leq 1$, provando a convergência. ■

Como consequência imediata do Teorema 2.2, notamos que o processo $\{X_t\}$ é uma combinação linear de variáveis aleatórias α -estáveis, portanto, também tem distribuição α -estável com mesmo índice de estabilidade. Se os coeficientes c_j 's são reais e satisfazem $|c_j| < cQ^{-j}$, para algum $c > 0$ e $Q > 1$, dizemos que o processo $\{X_t\}$ é estritamente estacionária e α -estável.

Os coeficientes c_j 's são obtidos através da identificação dos coeficientes de $C(z) = \sum_{j=0}^{\infty} c_j z^j$ com os da expansão da série $\Theta(z)/\Phi(z)$. Este procedimento é o mesmo do caso gaussiano.

O processo ARMA(p, q) é invertível se existe uma sequência de constantes $\{\tilde{c}_j\}$ tal que $\sum_{j=0}^{\infty} |\tilde{c}_j| < \infty$ e $\sum_{j=0}^{\infty} \tilde{c}_j X_{t-j} = a_t$, $n \in \mathbb{Z}$, onde a convergência em probabilidade é verdadeira. Escrever uma série na forma invertível é particularmente útil, pois permite que a série X_t seja expressa em função das observações anteriores X_j , $j \leq t$. O Teorema a seguir fornece a condição para invertibilidade.

Teorema 2.3. *Suponha que as raízes de $\Theta(B)$ estão fora do disco unitário $\{z : |z| \leq 1\}$. Então o processo ARMA(p, q) é invertível, ou seja,*

$$a_t = \sum_{j=0}^{\infty} \tilde{c}_j X_{t-j}$$

onde $n \in \mathbb{Z}$ e \tilde{c}_j 's são coeficientes da expansão de série $\Theta(z)^{-1}\Phi(z)$, $|z| < 1$.

As condições de invertibilidade e estacionariedade do processo ARMA com inovações α -estáveis são basicamente iguais as de um processo gaussiano, porém, a identificação da ordem e estimação dos parâmetros do processo ARMA com inovações α -estáveis requer a adoção de métodos que levem em consideração a variância infinita e a difícil representação da função de densidade da distribuição estável. As soluções para contornar estas adversidades serão apresentadas nos capítulos a seguir.

3 Identificação da ordem para ARMA(p,q) com inovações α -estáveis

No Capítulo 1 foi apresentado o método para identificação da ordem para um modelo ARMA gaussiano. No caso dos processos ARMA com inovações α -estáveis o momento de segunda ordem não existe para $\alpha < 2$, portanto a FAC e FACP não estão definidas. Neste caso, a principal ferramenta para identificar a ordem são as funções de codiferença e codiferença padronizada, medidas de dependência para processo estacionário proposta por Kokoszka e Taqqu (1994) e Yang et al. (2001), com o auxílio dos critérios de informação AIC e BIC. Apresentamos na Seção 3.1 estas medidas de dependência e suas propriedades. Em seguida, na Seção 3.2 abordamos a possibilidade da utilização da função codiferença na identificação da ordem dos modelos.

3.1 Função codiferença

3.1.1 Definição

A função codiferença é proposta em Kokoszka e Taqqu (1994) e Yang et al. (2001) segue a expressão abaixo:

$$\begin{aligned} \tau(k) = \tau(s, -s; k) = & -\ln E \exp\{is(X_{t+k} - X_t)\} + \ln E \exp\{isX_{t+k}\} \\ & + \ln E \exp\{-isX_t\}, \end{aligned} \quad (3.1)$$

onde $s \in \mathbb{R}$ e $k \in \mathbb{Z}$.

Devido a equação (3.1) ser calculada a partir da função característica, temos que a função codiferença existe para todo valor de s , não sendo necessário que os momentos de $\{X_t\}$ sejam finitos. Uma propriedade bastante importante e interessante está no fato de que no caso gaussiano a função de codiferença é proporcional à função de covariância:

$$\tau(s, -s; k) = -s^2 \text{Cov}(X_{t+k}, X_t) = -s^2 \gamma(k).$$

A função codiferença padronizada $I(k)$ segundo Rosadi e Deistler (2011) é definida por:

$$I(k) = \frac{\tau(s, -s; k)}{\tau(s, -s; 0)}. \quad (3.2)$$

Vale a correspondência $I(k) = \rho(k)$ no caso gaussiano, onde $\rho(k)$ é a função de autocorrelação. Note que, em geral, $\tau(-k) = \tau(k)^*$, onde $\tau(k)^*$ é o conjugado de $\tau(k)$.

A partir da Propriedade 2.10.5 em [Samorodnitsky e Taqqu \(1994\)](#), a codiferença padronizada $I(k)$ varia nos intervalos:

$$\begin{aligned} 0 \leq I(k) \leq 1 & \quad \text{se } 0 < \alpha \leq 1, \\ 1 - 2^{\alpha-1} \leq I(k) \leq 1 & \quad \text{se } 1 < \alpha \leq 2. \end{aligned}$$

Observe que para $\alpha = 2$ o intervalo de $I(k)$ coincide com intervalo de $\rho(k)$, $-1 \leq \rho(k) \leq 1$.

A estrutura de dependência da função codiferença (3.1) é explicada pelo Teorema a seguir.

Teorema 3.1. *Considere dois processos ARMA(p, q) independentes, $\{X_t\}$ e $\{Y_t\}$, com os mesmos parâmetros de distribuição. Se $\{X_t\}$ tem mais autocovariância que $\{Y_t\}$, digamos, $I^{(x)}(s, -s; k) \geq I^{(y)}(s, -s; k)$ para todo k , então o processo $\{X_t\}$ é menos autodependente que $\{Y_t\}$.*

Dem. *Primeiramente considere*

$$\mu_k = -\ln \mathbf{E} \left[e^{is(X_t - X_{t-k})} \right] \quad e \quad \nu_k = -\ln \mathbf{E} \left[e^{is(Y_t - Y_{t-k})} \right]. \quad (3.3)$$

$I^{(x)}(s, -s; k) \geq I^{(y)}(s, -s; k)$ se, somente se

$$\begin{aligned} -\ln \mathbf{E} \left[e^{is(X_t - X_{t-k})} \right] + \ln \mathbf{E} \left[e^{isX_t} \right] + \ln \mathbf{E} \left[e^{-isX_{t-k}} \right] &\geq \\ -\ln \mathbf{E} \left[e^{is(Y_t - Y_{t-k})} \right] + \ln \mathbf{E} \left[e^{isY_t} \right] + \ln \mathbf{E} \left[e^{-isY_{t-k}} \right] & \end{aligned}$$

e portanto de (3.3) segue que

$$\mu_k \geq \nu_k$$

ou

$$\mu_k^{-1} \nu_k \geq 1. \quad (3.4)$$

Observe que $\frac{X_t - X_{t-k}}{\mu_k}$ e $\frac{Y_t - Y_{t-k}}{\nu_k}$ tem a mesma distribuição, desse modo, para todo $c > 0$ temos

$$\begin{aligned} P(|X_t - X_{t-k}| > c) &= P\left(\left|\frac{X_t - X_{t-k}}{\mu_k}\right| > \frac{c}{\mu_k}\right) \\ &= P\left(\left|\frac{Y_t - Y_{t-k}}{\nu_k}\right| > \frac{c}{\mu_k}\right) \\ &= P(|Y_t - Y_{t-k}| > c\mu_k^{-1}\nu_k) \\ &\stackrel{(3.4)}{\geq} P(|Y_t - Y_{t-k}| > c). \end{aligned}$$

A desigualdade acima significa que Y_t e Y_{t-k} tem menor probabilidade de serem diferentes do que X_t e X_{t-k} . Concluimos assim que o processo $\{Y_t\}$ é mais autodependente que o processo $\{X_t\}$. ■

3.1.2 Estimação

A função codiferença pode ser estimada a partir da função característica empírica proposta por Yu (2004). Dada uma amostra X_1, X_2, \dots, X_N , o estimador para a função codiferença no lag $k \in \mathbb{Z}$ e $s \in \mathbb{R}$ é definida por (ROSADI, 2007):

$$\hat{\tau}(s, -s; k) = \sqrt{\frac{N-k}{N}} \times [-\ln \phi(s, -s; k) + \ln \phi(s, 0; k) + \ln \phi(0, -s; k)] \quad (3.5)$$

em que $u, v \in \mathbb{R}$ e

$$\phi(u, v; k) = \begin{cases} (N-k)^{-1} \sum_{N-k}^{t=1} \exp\{i(uX_{t+k} + vX_t)\} & \text{para } k \geq 0 \\ (N+k)^{-1} \sum_{N+k}^{t=1} \exp\{i(uX_{t-k} + vX_t)\} & \text{para } k < 0. \end{cases} \quad (3.6)$$

Consequentemente, o estimador de $I(k)$ é dado por:

$$\hat{I}(s, -s; k) = \frac{\hat{\tau}(s, -s; k)}{\hat{\tau}(s, -s; 0)}. \quad (3.7)$$

A estimação da função codiferença ocorre de maneira discreta, ou seja, é calculado o valor da função codiferença em r pontos $s_1 < s_2 < \dots < s_r$, onde $s_i \in \mathbb{R}$, $s_i \neq 0$, $i = 1, \dots, r$. Definimos os vetores $\mathbf{s} = [s_1, \dots, s_r]^T$,

$$\begin{aligned} \hat{\tau}(\mathbf{s}, k) &= [\hat{\tau}(s_1, -s_1; k), \hat{\tau}(s_2, -s_2; k), \dots, \hat{\tau}(s_r, -s_r; k)]^T \text{ e} \\ \hat{I}(\mathbf{s}, k) &= [\hat{I}(s_1, -s_1; k), \hat{I}(s_2, -s_2; k), \dots, \hat{I}(s_r, -s_r; k)]^T. \end{aligned}$$

Os vetores $\hat{\tau}(\mathbf{s}, k)$ e $\hat{I}(\mathbf{s}, k)$ são denotados por vetores de função codiferença e codiferença padronizada, respectivamente. Note que, pela unicidade, o fator $\sqrt{(N-k)/N}$ na equação (3.5) pode ser substituído por um, assim como o denominador $(N-k)$ em (3.6) pode ser trocado por N sem prejudicar as propriedades assintóticas do estimador. Note ainda que $\hat{\tau}(-k) = \hat{\tau}(k)$, possibilitando que a análise seja restrita para o caso $k \geq 0$. Outros dois estimadores semelhantes para a função codiferença foram propostos em Yang et al. (2001) e Hong (1999).

3.1.3 Propriedades assintóticas do estimador

Para os teoremas a seguir, considere o processo linear estacionário univariado $\{X_t, t \in \mathbb{Z}\}$, onde \mathbb{Z} denota o conjunto dos números inteiros, dado por:

$$X_t = \sum_{j=0}^{\infty} c_j a_{t-j}, \quad (3.8)$$

atendendo as condições:

(C1) Os coeficientes c_j 's são reais e satisfazem $|c_j| < Q^{-j}$ para algum $c > 0$, $Q > 1$;

(C2) As inovações a_t 's são i.i.d. com distribuição α -estável simétrica, isto é, a_t tem função característica na forma

$$E[\exp(isa_t)] = \exp(-\sigma^\alpha |s|^\alpha).$$

Sob estas condições, a desigualdade $\sum_{j=0}^{\infty} |c_j|^\alpha < \infty$ é mantida e a soma infinita (3.8) é bem definida no sentido da convergência quase certa. Além disso, o processo $\{X_t\}$ é $S\alpha S$ estritamente estacionário com índice de estabilidade α e parâmetro de escala $\sigma_X = \sigma(\sum_{j=0}^{\infty} |c_j|^\alpha)^{1/\alpha}$, segundo o Teorema 2.2.

Em particular, a função codiferença $\tau(k)$ para um processo linear (3.8), sob as condições (C1) e (C2), é dada por (KOKOSZKA; TAQQU, 1994):

$$\tau(k) = \sigma^\alpha |s|^\alpha \left[\sum_{j=0}^{\infty} (|c_{j+k} - c_j|^\alpha - |c_{j+k}|^\alpha - |-c_j|^\alpha) \right], \quad k \geq 0. \quad (3.9)$$

Note que sob as condições (C1) e (C2), a função codiferença padronizada (3.2) não depende da escolha de s .

As propriedades assintóticas da codiferença amostral estão resumidas nos dois teoremas a seguir.

Teorema 3.2. *Seja $\{X_t, t \in \mathbb{Z}\}$ um processo estacionário linear (3.8) satisfazendo as condições (C1) e (C2). Para $\mathbf{s} \in \mathbb{R}^r$, $s_i \neq 0$, $i = 1, \dots, r$, $k \in \{0, 1, 2, \dots\}$ a codiferença amostral $\hat{\tau}(\mathbf{s}, k)$ e a codiferença amostral padronizada $\hat{I}(\mathbf{s}, k)$ são estimadores (fracamente) consistentes para $\tau(\mathbf{s}, k)$ e $I(\mathbf{s}, k)$, respectivamente.*

A demonstração do Teorema 3.2 pode ser consultada no Apêndice A. A distribuição assintótica da função codiferença amostral (e função codiferença amostral padronizada) do processo linear (3.8) pode ser derivada usando Teorema Limite Central para funções características empíricas, (HESSE, 1990). Para facilitar, vamos dividir a função codiferença amostral em parte real e imaginária:

$$\Re \hat{\tau}(\mathbf{s}, k) = [\Re \hat{\tau}(s_1, -s_1; k), \dots, \Re \hat{\tau}(s_r, -s_r; k)]^T$$

e

$$\Im \hat{\tau}(\mathbf{s}, k) = [\Im \hat{\tau}(s_1, -s_1; k), \dots, \Im \hat{\tau}(s_r, -s_r; k)]^T.$$

Como $\hat{\tau}(s, -s; 0)$ é sempre real por definição, temos portanto

$$\Re\hat{\tau}(\mathbf{s}, k) = \begin{bmatrix} \Re\hat{\tau}(s_1, -s_1; k)/\hat{\tau}(s_1, -s_1; 0) \\ \Re\hat{\tau}(s_2, -s_2; k)/\hat{\tau}(s_2, -s_2; 0) \\ \vdots \\ \Re\hat{\tau}(s_r, -s_r; k)/\hat{\tau}(s_r, -s_r; 0) \end{bmatrix},$$

e

$$\Im\hat{\tau}(\mathbf{s}, k) = \begin{bmatrix} \Im\hat{\tau}(s_1, -s_1; k)/\hat{\tau}(s_1, -s_1; 0) \\ \Im\hat{\tau}(s_2, -s_2; k)/\hat{\tau}(s_2, -s_2; 0) \\ \vdots \\ \Im\hat{\tau}(s_r, -s_r; k)/\hat{\tau}(s_r, -s_r; 0) \end{bmatrix}.$$

O Teorema a seguir apresenta um resultado sobre a distribuição assintótica da função codiferença padronizada amostral. A demonstração do Teorema 3.3 pode ser consultada no Apêndice B.

Teorema 3.3. *Seja $\{X_t, t \in \mathbb{Z}\}$ um processo estacionário linear (3.8) satisfazendo as condições (C1) e (C2). Então para $k \in \{1, 2, \dots\}$,*

$$\left[\begin{pmatrix} \Re\hat{I}(\mathbf{s}, 1) \\ \Im\hat{I}(\mathbf{s}, 1) \end{pmatrix}, \begin{pmatrix} \Re\hat{I}(\mathbf{s}, 2) \\ \Im\hat{I}(\mathbf{s}, 2) \end{pmatrix}, \dots, \begin{pmatrix} \Re\hat{I}(\mathbf{s}, k) \\ \Im\hat{I}(\mathbf{s}, k) \end{pmatrix} \right]^T$$

tem distribuição

$$AN \left(\left[\begin{pmatrix} l_r I(1) \\ l_r 0 \end{pmatrix}, \begin{pmatrix} l_r I(2) \\ l_r 0 \end{pmatrix}, \dots, \begin{pmatrix} l_r I(k) \\ l_r 0 \end{pmatrix} \right]^T, N^{-1} \mathbf{W} \right).$$

Onde AN é a notação para indicar que a variável segue distribuição normal assintoticamente. A matriz de variância e covariâncias \mathbf{W} é dada em (B.19) e $l_r = [1, 1, \dots, 1]^T \in \mathbb{R}^r$.

O Colorário 3.1, a seguir, trata-se da aplicação do Teorema 3.3 no caso de observações i.i.d..

Colorário 3.1. *Seja $\{X_t, t \in \mathbb{Z}\}$ uma sequência i.i.d satisfazendo a condição a_t é i.i.d. com distribuição α -estável simétrica. Então para $k \in 1, 2, \dots$*

$$\Re\hat{I}(\mathbf{s}, k) \sim AN(0, N^{-1} \mathbf{W}_1) \quad (3.10)$$

e

$$\Im\hat{I}(\mathbf{s}, k) \sim AN(0, N^{-1} \mathbf{W}_2), \quad (3.11)$$

onde os elementos (i,j) das matrizes \mathbf{W}_1 e \mathbf{W}_2 são dados por

$$W_1(i,j) = \frac{f_{ij}}{g_{ij}}, \quad W_2(i,j) = \frac{h_{ij}}{g_{ij}}, \quad i,j = 1, \dots, r,$$

onde

$$\begin{aligned} f_{ij} &= e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i - s_j|^\alpha)} \left\{ \frac{1}{2} e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i - s_j|^\alpha)} - 1 \right\} \\ &\quad + e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i + s_j|^\alpha)} \left\{ \frac{1}{2} e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i + s_j|^\alpha)} - 1 \right\} + 1, \\ h_{ij} &= e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i - s_j|^\alpha)} \left\{ \frac{1}{2} e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i - s_j|^\alpha)} - 1 \right\} \\ &\quad + e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i + s_j|^\alpha)} e \left\{ 1 - \frac{1}{2} e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i + s_j|^\alpha)} \right\} \end{aligned}$$

e

$$g_{ij} = 4\sigma^{2\sigma} |s_i|^\alpha |s_j|^\alpha.$$

A demonstração pode ser consultada no Apêndice C.

3.2 Função codiferença na identificação da ordem

O segundo momento não definido para $\alpha < 2$ é a principal diferença entre o processo ARMA α -estável (2.7) e o ARMA gaussiano. Sem o segundo momento finito, a função de autocorrelação populacional não está definida para o processo ARMA α -estável, porém, as funções de autocovariância e autocorrelação amostrais estão bem definidas para qualquer conjunto de observações. Em Davis e Resnick (1986) mostra-se que para modelos lineares com variância infinita, a FAC amostral converge e, se padronizada corretamente, tem uma distribuição limite. Um exemplo de função limite pode ser consultada em Adler et al. (1998), onde os autores mostram que para um processo linear MA(q), $X_t = \sum_{j=0}^q \psi_j a_{t-j}$, com inovações SaS i.i.d., $\alpha < 2$ e $k > q$, vale

$$(n/\ln(n))^{1/\alpha} (\tilde{\rho}(k) - \rho(k)) \xrightarrow{d} \left(1 + 2 \sum_{j=0}^q |\rho(j)|^\alpha \right)^{1/\alpha} G_1/G_0, \quad (3.12)$$

onde $\tilde{\rho}(k) = \sum_{j=1}^{n-k} X_j X_{j+k} / \sum_{j=1}^n X_j^2$, $\rho(j) = \sum_{j=1}^{n-k} (X_j - \bar{X})(X_{j+k} - \bar{X}) / \sum_{j=1}^n (X_j - \bar{X})^2$, para $k = 1, 2, \dots$ e G_0 e G_1 são variáveis aleatórias independentes com distribuição $G_0 \sim S_{\alpha/2}(C_{\alpha/2}^{-2/\alpha}, 1, 0)$ e $G_1 \sim S_\alpha(C_\alpha^{-1/\alpha}, 0, 0)$, sendo

$$C_\alpha = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha) \cos(\frac{\pi\alpha}{2})} & \text{se } \alpha \neq 1, \\ \frac{2}{\pi} & \text{se } \alpha = 1. \end{cases}$$

Motivado pelo resultado em (3.12) Rosadi (2007) desenvolveu um método semelhante para identificação da ordem para casos de processos médias móveis com variância infinita.

Outras estratégias eram utilizadas para identificação da ordem para processos médias móveis puros com variância infinita considerados na literatura:

1. Segundo [Rosenfeld \(1976\)](#), caso o pesquisador desconheça o fato de que os dados são originados de um processo com caudas pesadas, a técnica utilizada será o plot do gráfico da FAC amostral em vários lags e comparar com os quantis $1,96\sqrt{n}$ da distribuição normal.
2. Segundo [Adler et al. \(1998\)](#), o pesquisador deve plotar o gráfico de $\rho(k)_\alpha = (n/\ln(n))^{1/\alpha}\rho(k)$ em vários lags e comparar com quantis 2,5% e 97,5% da distribuição de G_1/G_0 .

Além dos recursos acima, algumas generalizações da função de autocovariância foram sugeridas na literatura. Por exemplo, [Samorodnitsky e Taqqu \(1994\)](#) propuseram a *autovariância* e a *função de codiferença*, esta pode ser encontrada também em [Kokoszka e Taqqu \(1994\)](#), e a *função dinâmica* apresentada por [Janicki e Weron \(1993\)](#). Neste trabalho, a função codiferença será a medida utilizada para mensurar a dependência do modelo com inovações α -estáveis.

A utilização da função codiferença é estendida para identificação da ordem de modelos ARMA com inovações α -estáveis em [Rosadi e Deistler \(2011\)](#). O método dos autores funciona de forma semelhante ao apresentado no Capítulo 1 utilizando a FAC amostral. Mas antes de seguir com o método, é preciso fazer algumas observações sobre a função codiferença e a codiferença amostral. A partir da equação (3.9) e o fato de que todos os coeficientes c_j 's são números reais, temos que a função codiferença está definida no conjunto dos números reais apesar do seu estimador (3.5) estar definido no conjunto dos números complexos. Como o estimador de $I(k)$ é consistente, uma possibilidade é usar somente a parte real do estimador.

Outra diferença entre a função codiferença padronizada e seu estimador é que este depende do vetor $\mathbf{s} = \{s_1, \dots, s_r\}$. Aparentemente, $\hat{I}(\cdot)$ é definido para todo $\mathbf{s} > \mathbf{0}$, contudo, em caso de amostra finita, a precisão do estimador dos valores populacionais depende da escolha de \mathbf{s} ([ROSADI; DEISTLER, 2011](#)). No Capítulo 4 será discutido com mais detalhes os critérios para a seleção dos valores de \mathbf{s} .

A identificação da ordem para processos ARMA com inovações α -estáveis é realizada a partir do gráfico da parte real de $\hat{I}(k)$ com o intervalo de confiança calculado a partir da distribuição definida em (3.10). Consideramos que a parte real de $\hat{I}(\cdot)$ é “aproximadamente zero” para os valores dentro do intervalo de confiança, a partir de certo lag. Entretanto, este método de identificação da ordem apresentado não deve ser interpretado da mesma forma que os resultados de um processo gaussiano, dado

que tanto a função codiferença como a FAC para processos com variância infinita não tem uma base teórica sólida que justifique a utilização das mesmas (ROSADI, 2007).

4 Simulações

Iniciamos este capítulo apresentando as ferramentas disponíveis para a simulação de uma variável aleatória α -estável. Embora estimar os parâmetros de uma distribuição α -estável não seja uma tarefa simples, a simulação se dá de maneira bastante direta segundo a técnica de [Weron \(1996\)](#). Além do algoritmo para a simulação, o autor apresenta a demonstração da igualdade da lei das variáveis α -estáveis e da transformação não-linear das variáveis independentes uniforme e exponencial. Em seguida serão apresentadas algumas considerações práticas, idealizadas em [Rosadi e Deistler \(2011\)](#), sobre o estimador da função codiferença (3.5), dos quais os resultados de simulações de Monte Carlo direcionam à escolha dos pontos do vetor \mathbf{s} a fim de obter a melhor estimativa para a função codiferença (3.1).

4.1 Simulação de uma variável aleatória α -estável

A função característica da variável aleatória α -estável apresentada na equação (2.5) está escrita na forma canônica, entretanto, esta representação não é contínua nos pontos $\alpha = 1$ e $\beta \neq 0$. Entretanto, [Zolotarev \(1986\)](#) estabelece,

$$\mu_1 = \begin{cases} \mu + \beta\sigma^\alpha \tan \frac{\pi\alpha}{2}, & \text{se } \alpha \neq 1, \\ \mu, & \text{se } \alpha = 1, \end{cases} \quad (4.1)$$

resultando na expressão

$$\ln \varphi_X(t) = \begin{cases} i\mu_1 t - \sigma^\alpha [|t|^\alpha - it\beta(|t|^{\alpha-1} - 1) \tan \frac{\pi\alpha}{2}], & \text{se } \alpha \neq 1, \\ i\mu_1 t - \sigma |t| [1 + i\beta \frac{2}{\pi} \text{ sinal}(t) \ln |t|], & \text{se } \alpha = 1, \end{cases} \quad (4.2)$$

que é conjuntamente contínua para α e β . A desvantagem desta fórmula é que μ_1 não pode ser interpretado como parâmetro de locação. A Definição 4.1 representa uma outra forma de expressar a função característica, cujo uso pode ser justificado por considerações de natureza analítica ([ZOLOTAREV, 1986](#)).

Definição 4.1. *Dizemos que a variável aleatória X é α -estável se somente se sua função característica é dada por*

$$\ln \varphi_X(t) = \begin{cases} i\mu t - \sigma_2^\alpha |t|^\alpha \exp\{i\beta_2 \text{ sinal}(t) \frac{\pi}{2} K(\alpha)\}, & \text{se } \alpha \neq 1, \\ i\mu t - \sigma_2 |t| [\frac{\pi}{2} + i\beta_2 \text{ sinal}(t) \ln |t|], & \text{se } \alpha = 1, \end{cases} \quad (4.3)$$

onde

$$K(\alpha) = \alpha - 1 + \text{sinal}(1 - \alpha) = \begin{cases} \alpha, & \text{se } \alpha < 1, \\ \alpha - 2, & \text{se } \alpha > 1. \end{cases} \quad (4.4)$$

Os parâmetros σ_2 e β_2 estão relacionados aos σ e β da equação (2.4) da seguinte forma: Para $\alpha \neq 1$, β_2 é de tal modo que

$$\tan\left(\beta_2 \frac{\pi K(\alpha)}{2}\right) = \beta \tan \frac{\pi\alpha}{2}, \quad (4.5)$$

e o novo parâmetro de escala

$$\sigma_2 = \sigma \left(1 + \beta^2 \tan^2 \frac{\pi\alpha}{2}\right)^{1/(2\alpha)}. \quad (4.6)$$

Para $\alpha = 1$, $\beta_2 = \beta$ e $\sigma_2 = \frac{2}{\pi}\sigma$.

A representação da integral da função densidade de probabilidade para todos os valores dos parâmetros α e β dada em Zolotarev (1986) será apresentada na Proposição 4.1.

Proposição 4.1. (ZOLOTAREV, 1986, Nota 1, pág 78). Sejam

$$\epsilon(\alpha) = \text{sinal}(1 - \alpha),$$

$$\gamma_0 = -\frac{\pi}{2}\beta_2 \frac{K(\alpha)}{\alpha},$$

$$C(\alpha, \beta_2) = 1 - \frac{1}{4}(1 + \beta_2 K(\alpha)/\alpha)(1 + \epsilon(\alpha)),$$

$$U_\alpha(\gamma, \gamma_0) = \left(\frac{\text{sen } \alpha(\gamma - \gamma_0)}{\cos \gamma}\right)^{\alpha/(1-\alpha)} \frac{\cos(\gamma - \alpha(\gamma - \gamma_0))}{\cos \gamma}$$

e

$$U_1(\gamma_1, \beta_2) = \frac{\frac{\pi}{2} + \beta_2 \gamma}{\cos \gamma} \exp\left(\frac{1}{\beta_2} \left(\frac{\pi}{2} + \beta_2 \gamma\right) \tan \gamma\right).$$

Então a função de distribuição $F(x, \alpha, \beta_2)$ de uma variável aleatória estável padronizada, cuja função característica seja descrita pela equação (4.3) pode ser escrita como:

- Se $\alpha \neq 1$ e $x > 0$ então

$$F(x, \alpha, \beta_2) = C(\alpha, \beta_2) + \frac{\epsilon(\alpha)}{\pi} \int_{\gamma_0}^{\frac{\pi}{2}} \exp\{-x^{\alpha/(1-\alpha)} U_\alpha(\gamma, \gamma_0)\} d\gamma, \quad (4.7)$$

- Se $\alpha = 1$ e $\beta_2 > 0$ então

$$F(x, 1, \beta_2) = \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp\{-e^{-x/\beta_2} U_1(\gamma, \beta_2)\} d\gamma. \quad (4.8)$$

Os casos $\alpha \neq 1$, $x < 0$ e $\alpha = 1$, $\beta_2 < 0$ podem ser reduzidos para os casos correspondentes $\alpha \neq 1$, $x > 0$ e $\alpha = 1$, $\beta_2 > 0$ com o auxílio da seguinte igualdade:

$$F(-x, \alpha, \beta_2) + F(x, \alpha, -\beta_2) = 1, \quad (4.9)$$

que é válida para as duas representações da variável estável, (2.5) e (4.3), para qualquer x real e quaisquer parâmetros admissíveis α e β_2 (ou β).

A complexidade em relação ao problema da simulação de uma sequência de variáveis aleatórias estáveis está ligada ao fato de não existir uma expressão analítica para a inversa da função de densidade. As únicas exceções são para os casos específicos em que, a depender dos parâmetros, a função de distribuição da α -estável coincide com as distribuições Gaussiana, Cauchy e Lévy. A solução para o problema foi publicada no artigo de [Kanter et al. \(1975\)](#), onde o um método direto de simulação geravam variáveis aleatórias $S_\alpha(1,1,0)$ para $\alpha < 1$. [Chambers et al. \(1976\)](#) foram os primeiros a descrever as fórmulas. Entretanto, eles não deram as demonstrações, apenas indicaram o artigo de [Zolotarev \(1966\)](#) onde as expressões (4.7) e (4.8) podem ser encontradas. A falta destas demonstrações explícitas acarretou em algumas imprecisões na literatura. [Weron \(1996\)](#) esclarece esta situação e apresenta as demonstrações e o algoritmo para a simulação de um variável aleatória α -estável.

Lema 4.1. *Sejam γ_0 e $U_\alpha(\gamma, \gamma_0)$ definidos segundo a Proposição 4.1. Para $\alpha \neq 1$ e $\gamma_0 < \gamma < \frac{\pi}{2}$, X é uma v.a. com distribuição $S_\alpha(1, \beta_2, 0)$ (de acordo com a representação (4.3)) se e somente se para $x > 0$,*

$$\frac{1}{\pi} \int_{\gamma_0}^{\frac{\pi}{2}} \exp\{-x^{\alpha/(\alpha-1)} U_\alpha(\gamma, \gamma_0)\} d\gamma = \begin{cases} P(0 < X \leq x), & \text{se } \alpha < 1, \\ P(X \geq x), & \text{se } \alpha > 1. \end{cases} \quad (4.10)$$

Dem. *Parte 1: Para $0 < \alpha < 1$. A partir de (4.7) temos que*

$$\begin{aligned} F(x, \alpha, \beta_2) &= P(X \leq x) \\ &= \frac{1 - \beta_2}{2} + \frac{1}{\pi} \int_{\gamma_0}^{\frac{\pi}{2}} \exp\{-x^{\alpha/(\alpha-1)} U_\alpha(\gamma, \gamma_0)\} d\gamma \\ &= \frac{1 - \beta_2}{2} + P(0 < X \leq x), \end{aligned}$$

como $\alpha < 1$, $\frac{1-\beta_2}{2} = P(X \leq 0)$ ([ZOLOTAREV, 1986, Nota 2, pág 79](#)).

Parte 2: Quando $1 < \alpha \leq 2$. A partir de 4.7 temos que

$$\begin{aligned} F(x, \alpha, \beta_2) &= P(X \leq x) \\ &= 1 - \frac{1}{\pi} \int_{\gamma_0}^{\frac{\pi}{2}} \exp\{-x^{\alpha/(\alpha-1)} U_\alpha(\gamma, \gamma_0)\} d\gamma \\ &= 1 - P(X \geq x). \end{aligned}$$

■

Teorema 4.1. *Seja γ_0 definido como na Proposição 4.1. Seja γ uniformemente distribuída no intervalo $(-\frac{\pi}{2}, \frac{\pi}{2})$ e seja W uma v.a. independente com distribuição exponencial de média 1. Então*

- para $\alpha \neq 1$

$$X = \frac{\text{sen } \alpha(\gamma - \gamma_0)}{(\cos \gamma)^{1/\alpha}} \left(\frac{\cos(\gamma - \alpha(\gamma - \gamma_0))}{W} \right)^{(1-\alpha)/\alpha} \quad (4.11)$$

segue a distribuição $S_\alpha(1, \beta_2, 0)$ e

- para $\alpha = 1$

$$X = \left(\frac{\pi}{2} + \beta_2 \gamma \right) \tan \gamma - \beta_2 \ln \left(\frac{W \cos \gamma}{\frac{\pi}{2} + \beta_2 \gamma} \right) \quad (4.12)$$

segue a distribuição $S_1(1, \beta, 0)$

para a representação (4.3).

Dem. Quando $\gamma > \gamma_0$, então o lado direito de (4.11) é positivo e pode ser expresso por

$$\left(\frac{a(\gamma)}{W} \right)^{(1-\alpha)/\alpha}, \quad (4.13)$$

onde

$$a(\gamma) = \left(\frac{\text{sen } \alpha(\gamma - \gamma_0)}{\cos \gamma} \right)^{\alpha/(1-\alpha)} \frac{\cos(\gamma - \alpha(\gamma - \gamma_0))}{\cos \gamma}. \quad (4.14)$$

Para $0 < \alpha < 1$: A equação (4.11) implica que $X > 0$ se e somente se $\gamma > \gamma_0$. Dado que $\frac{1-\alpha}{\alpha}$, podemos escrever

$$\begin{aligned} P(0 < X \leq x) &= P(0 < X \leq x, \gamma > \gamma_0) \\ &= P(0 < (a(\gamma)/W)^{(1-\alpha)/\alpha} \leq x, \gamma > \gamma_0) \\ &= P(W \geq x^{\alpha/(\alpha-1)} a(\gamma), \gamma > \gamma_0) \\ &= E_\gamma \exp\{-x^{\alpha/(\alpha-1)} a(\gamma)\} I_{(\gamma > \gamma_0)} \\ &= \frac{1}{\pi} \int_{\gamma_0}^{\frac{\pi}{2}} \exp\{-x^{\alpha/(\alpha-1)} a(\gamma)\} d\gamma. \end{aligned}$$

A partir do Lema 4.1 e (4.9) concluímos que $X \sim S_\alpha(1, \beta_2, 0)$.

Para $1 < \alpha \leq 2$: Dado que $\frac{\alpha-1}{\alpha}$, para $x > 0$ podemos escrever

$$\begin{aligned} P(X \geq x) &= P(X \geq x, \gamma > \gamma_0) \\ &= P((a(\gamma)/W)^{(1-\alpha)/\alpha} \geq x, \gamma > \gamma_0) \\ &= P((W/a(\gamma))^{(\alpha-1)/\alpha} \geq x, \gamma > \gamma_0) \\ &= P(W \geq x^{\alpha/(\alpha-1)} a(\gamma), \gamma > \gamma_0) \\ &= E_\gamma \exp\{-x^{\alpha/(\alpha-1)} a(\gamma)\} I_{(\gamma > \gamma_0)} \\ &= \frac{1}{\pi} \int_{\gamma_0}^{\frac{\pi}{2}} \exp\{-x^{\alpha/(\alpha-1)} a(\gamma)\} d\gamma. \end{aligned}$$

A partir do Lema 4.1 e (4.9) concluímos que $X \sim S_\alpha(1, \beta_2, 0)$.

Para $\alpha = 1$. Quando $\beta_2 = 0$, o lado direito de (4.12) se reduz a $\frac{\pi}{2} \tan \gamma$ e segue uma distribuição Cauchy (segundo a representação (4.3)). Quando $\beta_2 \neq 0$, pode ser expressa por

$$\beta_2 \ln \left(\frac{a_1(\gamma)}{W} \right), \quad (4.15)$$

onde

$$a_1(\gamma) = \frac{\frac{\pi}{2} + \beta_2 \gamma}{\cos \gamma} \exp \left(\frac{1}{\beta_2} \left(\frac{\pi}{2} + \beta_2 \gamma \right) \tan \gamma \right). \quad (4.16)$$

Temos que para $\beta_2 > 0$,

$$\begin{aligned} P(X \leq x) &= P(\beta_2 \ln(a_1(\gamma)/W) \leq x) \\ &= P(W \geq e^{-x/\beta_2} a_1(\gamma)) \\ &= E_\gamma \exp\{-e^{-x/\beta_2} a_1(\gamma)\} \\ &= \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp\{-e^{-x/\beta_2} a_1(\gamma)\} d\gamma. \end{aligned}$$

A partir de (4.8) e (4.9), concluímos que para todo β_2 , $X \sim S_\alpha(1, \beta_2, 0)$. Com isso a demonstração está completa. ■

A partir da aplicação do Teorema 4.1 podemos apresentar um algoritmo para simulação de uma variável aleatória $X \sim S_\alpha(1, \beta, 0)$ de acordo com a representação (2.5). Desse modo, para $\alpha \in (0, 2]$ e $\beta \in [-1, 1]$, os passos são:

- simular uma v.a. $V \sim U\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ e uma v.a. W com distribuição exponencial independente com média 1.
- para $\alpha \neq 1$ calcular:

$$X = S_{\alpha,\beta} \times \frac{\text{sen}(\alpha(V + B_{\alpha,\beta}))}{(\cos V)^{1/\alpha}} \times \left(\frac{\cos(V - \alpha(V + B_{\alpha,\beta}))}{W} \right)^{(1-\alpha)/\alpha}, \quad (4.17)$$

onde

$$B_{\alpha,\beta} = \frac{\arctan(\beta \tan \frac{\pi\alpha}{2})}{\alpha} e$$

$$S_{\alpha,\beta} = \left[1 + \beta^2 \tan^2 \frac{\pi\alpha}{2} \right]^{1/(2\alpha)}.$$

- para $\alpha = 1$ calcular:

$$X = \frac{2}{\pi} \left[\left(\frac{\pi}{2} + \beta V \right) \tan V - \beta \ln \left(\frac{\frac{\pi}{2} W \cos V}{\frac{\pi}{2} + \beta V} \right) \right]. \quad (4.18)$$

A função $B_{\alpha,\beta}$ é encarregada pela mudança de β_2 para β e substitui γ_0 em (4.11). E a função $S_{\alpha,\beta}$ realiza a mudança de σ_2 para σ .

A fórmula (4.17) foi inicialmente apresentada por Janicki e Weron (1993), no entanto eles descreveram as funções de transformações equivalentes a $B_{\alpha,\beta}$ e $S_{\alpha,\beta}$ de forma incorreta e computacionalmente mais complicada, respectivamente. Além disso, não apresentaram uma fórmula para o caso $\alpha = 1$. Na simulação de uma v.a. α -estável proposta por Chambers et al. (1976) uma fórmula para $\alpha \neq 1$, equivalente a (4.17) é proposta para a representação (4.3). Weron (1996) publicou uma correção para essa fórmula alegando que o logaritmo apresentava o valor $\frac{\frac{\pi}{2} W \cos V}{\frac{\pi}{2} + \beta V}$ quando deveria ser $\frac{W \cos V}{\frac{\pi}{2} + \beta V}$. Contudo, mais tarde, o autor publicou uma correção retratando sua colocação e informando que a fórmula de Chambers et al. (1976) estava correta.

As fórmulas indicadas acima permitem a simulação de variáveis aleatórias α -estáveis padronizadas. Aplicando as Propriedades 2.2 e 2.3 podemos gerar uma variável α -estável para todos os valores admissíveis de seus parâmetros α, σ, β e μ :

Se $X \sim S_\alpha(1, \beta, 0)$, então

$$Y = \begin{cases} \sigma X + \mu, & \text{se } \alpha \neq 1, \\ \sigma X + \frac{2}{\pi} \beta \sigma \ln \sigma + \mu, & \text{se } \alpha = 1. \end{cases}$$

tem distribuição $S_\alpha(\sigma, \beta, \mu)$.

4.2 Resultados para estimativa da função codiferença

Um estudo de simulações é conduzido nesta seção e consiste na reprodução dos resultados da Seção 3 de [Rosadi e Deistler \(2011\)](#). A apresentação está dividida em duas partes: as considerações práticas que orientam sobre a escolha dos pontos $\mathbf{s} = (s_1, \dots, s_r)$ para o cálculo de uma boa estimativa de $I(k)$ e os resultados referentes à análise de simulações Monte Carlo para vários cenários.

4.2.1 Considerações práticas

Vimos no Capítulo 3 que a estimativa da função codiferença, $\hat{I}(\cdot)$, é calculada para cada valor do vetor $\mathbf{s} = (s_1, \dots, s_r)$ e a escolha do número de pontos, r , e suas posições tem impacto na precisão da estimativa do valor populacional de $I(\cdot)$. De acordo com [Koutrouvelis \(1980\)](#) e [Kogon e Williams \(1998\)](#), para obter uma boa estimativa da função característica a partir da função característica empírica, a localização dos pontos s_1, \dots, s_r devem ser próximos de zero, mas não iguais a zero. Uma ferramenta para nos auxiliar na escolha do intervalo em que os valores de $s_1 < s_2 < \dots < s_r$ devem estar consiste na construção de gráficos para $\Re\hat{I}(s, k)$ versus s com valores entre $0,01 \leq s \leq 2$ e alguns valores de $k > 0$. Estes gráficos irão indicar os intervalos para \mathbf{s} onde o viés é menor.

Para exemplificar, calculamos $\hat{I}(\cdot)$ para 10 séries simuladas do processo $X_t = a_t + c_1 a_{t-1} + c_2 a_{t-2}$ onde $c_1 = 2$, $c_2 = 1,111$ e as inovações a_t 's são i.i.d. com distribuição SaS com $\sigma = 1$. Para os gráficos da Figura 7 foram simuladas séries com 3 valores de α e dois diferentes tamanhos de amostra: $N=100$ e $N=1000$.

Observa-se na Figura 7 que para valores pequenos de α , $\Re\hat{I}(s, k)$ tem comportamento mais instável e quando o tamanho da amostra aumenta o efeito errático diminui. Podemos concluir que o estimador da função codiferença, ao que tudo indica, depende do tamanho da amostra e do parâmetro α .

Nos gráficos da Figura 8 foi simulado da mesma forma que os gráficos da Figura 7 com uma amostra de tamanho 1000, variando o parâmetro α e para os lags $k = 1, 2$ e 3 .

Observando os lags na estimação de $\Re\hat{I}(s, k)$ percebe-se que o efeito errático sofre um leve aumento à medida que k aumenta, mas a redução do α tem efeito maior sob este comportamento. Com o auxílio dos gráficos simulados, podemos concluir que a escolha do intervalo de \mathbf{s} é mais influenciado pelos parâmetros dos dados em si (α e tamanho da amostra) do que a escolha do lag k . Além disso, recomenda-se que o limite inferior de \mathbf{s} seja igual a $s_1 = 0,01$ e a escolha do limite superior, s_r , seja até

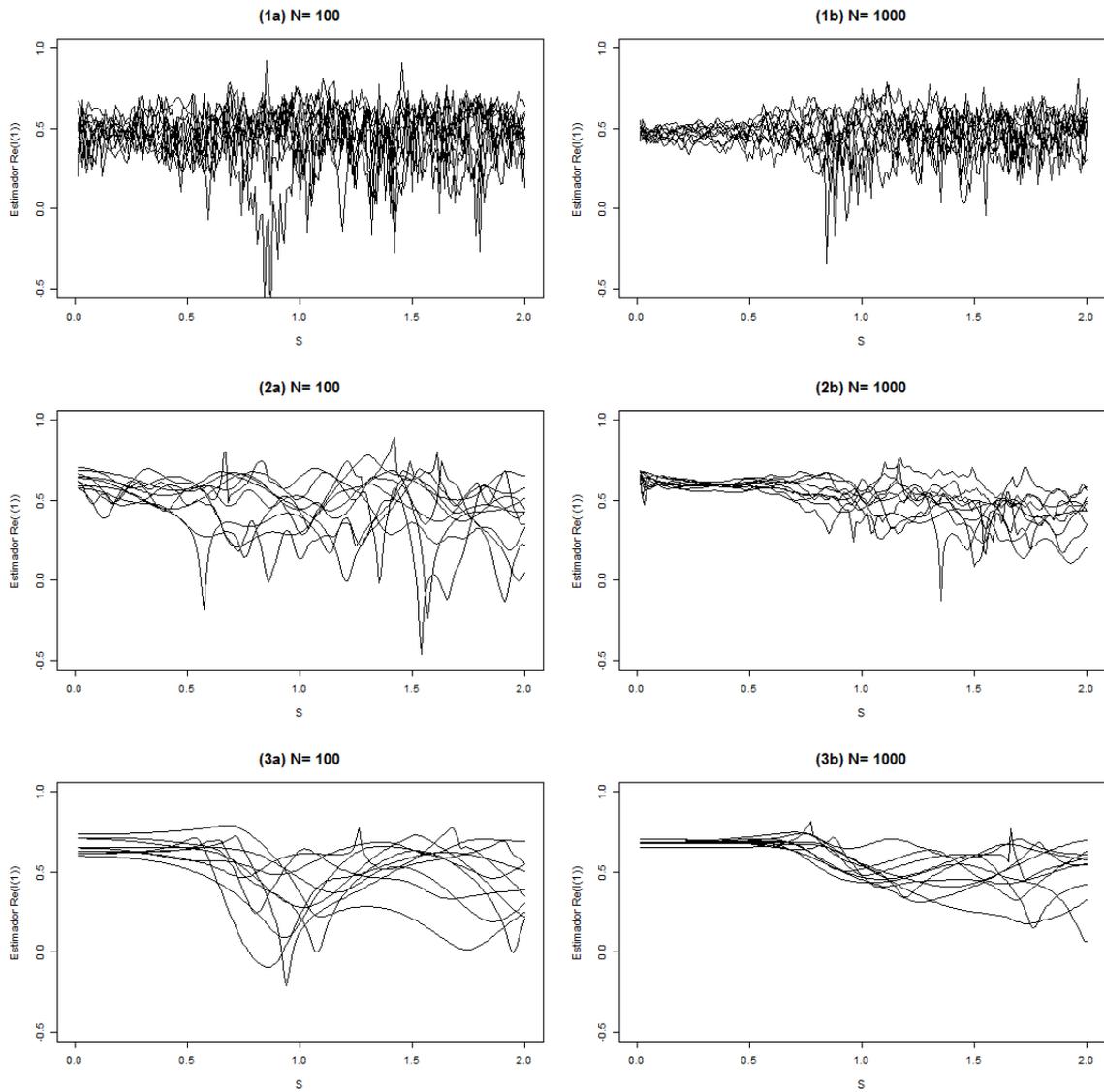


Figura 7 – Gráficos $\Re\hat{I}(s, k)$ versus s . Onde $\alpha = 2$ (**1a-1b**), $\alpha = 1,5$ (**2a-2b**) e $\alpha = 0,8$ (**3a-3b**).

onde a gráfico de $\Re\hat{I}(s, k)$ versus s mantém comportamento retilíneo.

Definido os limites para os valores de s , vamos discutir sobre o número pontos r para produzir uma boa estimação de $\Re\hat{I}(s, k)$. Quando $\alpha = 2$, temos que $\Re\hat{I}(s, k)$ tem comportamento relativamente suave, portanto, não é mais vantajoso escolher valores de s_i 's muito próximos. Nos casos em que $\alpha < 2$, o comportamento errático de $\Re\hat{I}(s, k)$ se intensifica à proporção que α diminui, justificando a seleção de s_i 's suficientemente próximos. [Rosadi e Deistler \(2011\)](#) recomendam as seguintes distâncias (d) entre s_i 's:

Especialmente nos casos i.i.d., podemos mostrar que estas escolhas são suficientes, dado α , escolhendo a menor distância entre os s_i 's que não reduz significativamente o determinante da matriz de covariâncias (3.10). Todavia, como o comporta-

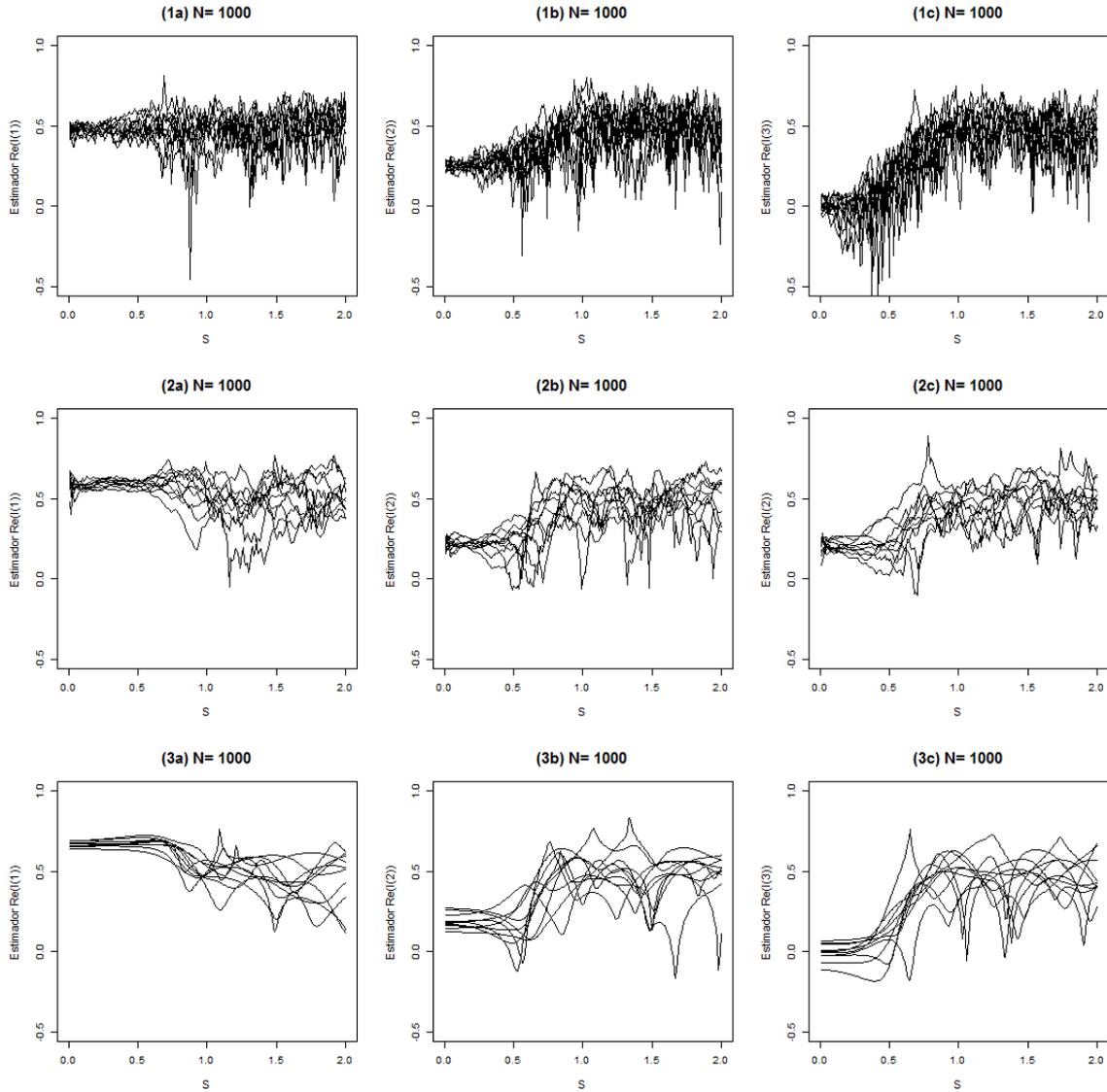


Figura 8 – $\Re\hat{I}(s,k)$ versus s nos lags 1, 2 e 3. Onde $\alpha = 2$ (**1a-1c**), $\alpha = 1,5$ (**2a-2c**) e $\alpha = 0,8$ (**3a-3c**).

$$\begin{aligned}
 & \text{se } \alpha \leq 1, \quad d=0,01, \\
 & \text{se } 1 < \alpha \leq 1,5, \quad 0,01 < d \leq 0,05, \\
 & \text{se } 1,5 < \alpha < 2, \quad 0,05 < d \leq 0,1, \\
 & \text{se } \alpha = 2, \quad d = 0,1 \text{ ou maior.}
 \end{aligned}$$

mento errático da função codiferença amostral é tipicamente relacionado ao valor de α , na prática não é necessário conhecer o valor de α .

Os s_i 's podem ser igualmente espaçados, ou seja, $\mathbf{s} = (s_1 = 0,01, 0,01 + i \frac{s_b - s_1}{r-1}, s_b)$ para $i = 1, 2, \dots, r-2$ ou com espaçamento desigual desde de que sejam suficientemente próximos de acordo com o comportamento de $\Re\hat{I}(s,k)$.

Note que ao decidir os valores de s_i 's a estimativa de $\Re\hat{I}(s,k)$ é calculada para

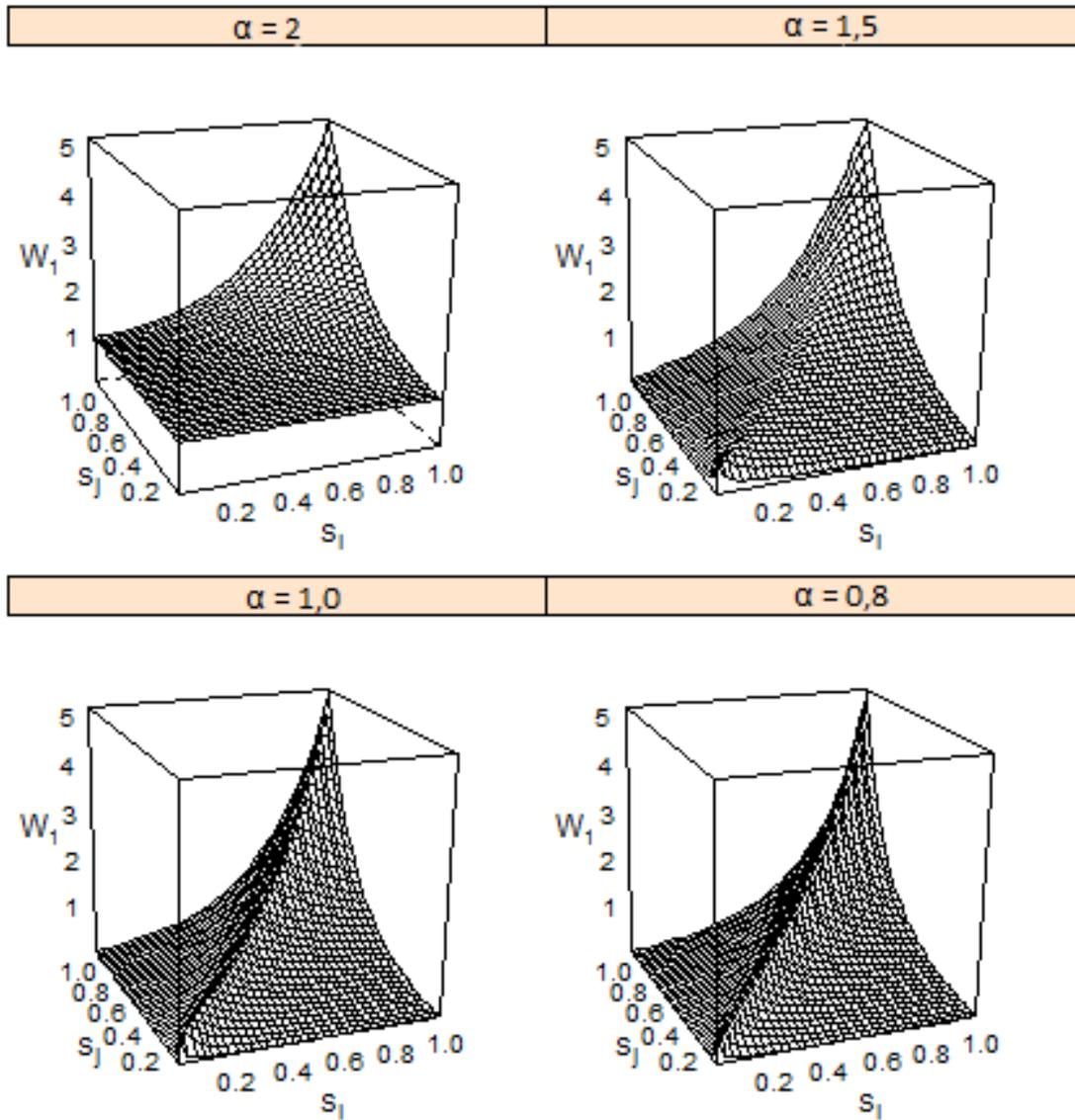


Figura 9 – Gráficos de $W_1(i,j)$ para $s_i, s_j \in [0,01,1]$ e alguns valores de α .

cada ponto, resultando no vetor $\mathfrak{R}\hat{I}(\mathbf{s},k) = [\mathfrak{R}\hat{I}(s_1, -s_1; k), \dots, \mathfrak{R}\hat{I}(s_r, -s_r; k)]^T$. A estimativa final de $\mathfrak{R}\hat{I}(s,k)$ pode ser definida como uma média ponderada das estimativas nos pontos s_1, \dots, s_r , ou seja, $\hat{I}(s,k) = \sum_{i=1}^r w_i \hat{I}(s_i, -s_i; k)$ onde $\sum_{i=1}^r w_i = 1$. A princípio, podemos utilizar a média simples, com $w_i = 1/r$, ou a média ponderada exponencialmente negativa, onde $w_i = \exp(-s_i^2) / \sum_{j=1}^r \exp(-s_j^2)$. No caso i.i.d. obtemos, ao calcular a média das estimativas em diferentes pontos, que a variância do estimador será menor ou igual a variância calculada em um único ponto, como pode ser observado na Figura 9. Isso mostra que para $\alpha = 2$ não existe diferença, em termos assintóticos, entre variância e a estimativa de $\hat{I}(\cdot)$ em um ou mais pontos, enquanto para $\alpha < 2$ a diferença é significativa, especialmente quando α é pequeno. Além disso, podemos concluir, a partir da Figura 9, que quanto menor α menor será

a covariância entre os pontos. Para os casos de amostra finita, este fato é consistente com o típico comportamento errático do gráfico $\mathfrak{R}\hat{I}(\cdot)$ para algumas amostras não i.i.d.'s, como apresentado nas Figuras 7 a 8. Baseado nos resultados das simulações, sugere-se que a escolha do número de pontos r ocorra da seguinte forma: para $\alpha = 2$ (ou seja, para os gráficos de $\mathfrak{R}\hat{I}(k)$, $k > 0$ com comportamento suave), observou-se que $r = 1$ é suficiente, enquanto que para $\alpha < 2$ (ou seja, gráficos de $\mathfrak{R}\hat{I}(k)$, $k > 0$ com comportamento errático), pelo menos dois pontos devem ser escolhidos, e o número de pontos deve ser maior à medida que o valor de α diminui. É importante destacar que a partir dos resultados da simulação, observou-se que a precisão do estimador é mais sensível a localização dos pontos do que a quantidade de pontos.

4.2.2 Resultados da simulação

Para investigar o comportamento da função codiferença para amostras finitas, em particular sua dependência em relação a escolha de \mathbf{s} , foram realizadas diversas simulações de Monte Carlo utilizando o software R Core Team (2014) versão 3.1.2. Para a simulação da variável aleatória α -estável foi implementado um algoritmo seguindo a técnica de Weron (1996) descrita na Seção 4.1 e utilizamos a função *arima.sim* do pacote **Stats** para gerar o modelo MA(2) $X_t = a_t + c_1 a_{t-1} + c_2 a_{t-2}$ com inovações α -estáveis. Foram considerados quatro modelos nas simulações:

Modelo	c_1	c_2
M1	2	1.111
M2	-1	0.5
M3	0.55	0.05
M4	-0.4	0.7

De agora em diante cada um dos modelos serão denotados pelas siglas M1, M2, M3 e M4, respectivamente. Quando a distribuição de a_t é gaussiana, os modelos M1 a M3 correspondem aos experimentos I, II e III examinados por Bhansali (1983).

Para avaliar a performance do estimador, simulamos as séries temporais dos modelos M1 a M4 para vários valores de α com $\sigma = 1$ e dois tamanhos de amostra, $N=100$ representando uma amostra “pequena” e $N=1000$ considerada como uma amostra “grande”. A codiferença amostral foi calculada para os lags 1 a 3 e cada experimento foi repetido $T=1000$ vezes. As Figuras 7 e 8 sugerem que $\mathfrak{R}\hat{I}(\cdot)$ é menos viesado no intervalo $0.01 \leq s \leq 0.5$, embora a escolha do intervalo de \mathbf{s} dependa do parâmetro α . Para comparar o impacto da escolha dos pontos \mathbf{s} Rosadi e Deistler (2011) escolhe diferentes conjuntos de $\mathbf{s}_i = \{s_1, \dots, s_r\}$, $i = 1, \dots, 28$. Foram considerados conjuntos com pontos equidistantes e não equidistante. Segue abaixo a lista completa dos pontos escolhidos:

$\mathbf{s}_1 = \{0.01\}$	$\mathbf{s}_{15} = \{0.01, 0.1, 0.2\}$
$\mathbf{s}_2 = \{0.1\}$	$\mathbf{s}_{16} = \{0.01, 0.1, 0.5\}$
$\mathbf{s}_3 = \{0.2\}$	$\mathbf{s}_{17} = \{0.01, 0.1, 1\}$
$\mathbf{s}_4 = \{0.3\}$	$\mathbf{s}_{18} = \{0.1, 0.2, 0.3\}$
$\mathbf{s}_5 = \{0.5\}$	$\mathbf{s}_{19} = \{0.1, 0.3, 0.5\}$
$\mathbf{s}_6 = \{1\}$	$\mathbf{s}_{20} = \{0.01, 0.5, 1\}$
$\mathbf{s}_7 = \{0.01, 0.1\}$	$\mathbf{s}_{21} = \{0.1, 0.5, 1\}$
$\mathbf{s}_8 = \{0.01, 0.2\}$	$\mathbf{s}_{22} = \{0.1, 0.2, 0.3, 0.4, 0.5\}$
$\mathbf{s}_9 = \{0.01, 0.5\}$	$\mathbf{s}_{23} = \{0.1, 0.2, \dots, 1\}$
$\mathbf{s}_{10} = \{0.01, 1\}$	$\mathbf{s}_{24} = \{0.01, 0.06, 0.11, 0.16, 0.21\}$
$\mathbf{s}_{11} = \{0.1, 0.2\}$	$\mathbf{s}_{25} = \{0.01, 0.02, \dots, 0.2\}$
$\mathbf{s}_{12} = \{0.1, 0.5\}$	$\mathbf{s}_{26} = \{0.01, 0.02, \dots, 0.1\}$
$\mathbf{s}_{13} = \{0.1, 1\}$	$\mathbf{s}_{27} = \{0.11, 0.12, \dots, 0.2\}$
$\mathbf{s}_{14} = \{0.5, 1\}$	$\mathbf{s}_{28} = \{0.5, 0.55, \dots, 1\}$

Para cada s_i na repetição $h = 1, \dots, T$ a estimativa final da função codiferença foi calculada como a média ponderada em cada ponto $s_{ij}, j = 1, \dots, r_i$, denotada por $\mathfrak{R}\hat{I}(\cdot)_{ih} = \sum_{j=1}^{r_i} w_{ij} \mathfrak{R}\hat{I}(s_{ij}, -s_{ij}; \cdot)_h$, onde $\mathfrak{R}\hat{I}(s_i, -s_i; \cdot)_h$ é a parte real da função codiferença amostral da repetição h em algum lag nos pontos $s_{ij}, j = 1, \dots, r_i$. Foram utilizados como método de ponderação das estimativas, a média simples e exponencial negativa. Por questão de espaço, será apresentado na Tabela 1 apenas os resultados do M1, mas os resultados para os outros modelos podem ser consultados no Apêndice D. Na tabela, foi registrado apenas os resultados para a melhor escolha de \mathbf{s} , que foi definida como os pontos que minimizam a soma do desvio médio absoluto (DMA) das estimativas no lag 1 entre todos os s_i 's apresentados. O DMA no lag k e s_i é definido por $DMA_{ik} = \frac{1}{T} \sum_{h=1}^T |\mathfrak{R}\hat{I}(\cdot)_{ih} - I(k)|$, $k = 1, 2$ e 3 . Foi calculado também o erro quadrático médio (EQM) mas o resultado foi muito similar ao do DMA, por isso não será apresentado.

Tabela 1 – Valores da função codiferença $I(\cdot)$ e sua estimativa $\hat{I}(\cdot)$ para séries simulada.

N	α	Método	s	I(1)	$\hat{I}^*(1)$	MAD_1	I(2)	$\hat{I}^*(2)$	MAD_2	I(3)	$\hat{I}^*(3)$	MAD_3
100	2	Simples	{0.01, 0.2}	0,67722	0,6631	0,0432	0,1782	0,1404	0,1031	0	-0,0383	0,1145
		Exp.	{0.01}		0,6620	0,0434		0,1502	0,1032	0	-0,0244	0,1105
	1,8	Simples	{0.01}	0,64700	0,6617	0,0445	0,1924	0,1487	0,1002	0	-0,0371	0,1009
		Exp.	{0.01}		0,6563	0,0436		0,1491	0,1011	0	0,0400	0,0927
	1,5	Simples	{0.01, 0.1, 0.2}	0,59903	0,6123	0,0480	0,2134	0,1725	0,0786	0	-0,0273	0,0886
		Exp.	{0.01, 0.1, 1}		0,5995	0,0468		0,2125	0,0722	0	0,0472	0,0828
	1,3	Simples	{0.01, 0.06, 0.11, 0.16, 0.21}	0,56554	0,5776	0,0499	0,2267	0,1858	0,0720	0	-0,0318	0,0803
		Exp.	{0.01, 0.06, 0.11, 0.16, 0.21}		0,5746	0,0487		0,1886	0,0742	0	0,0463	0,0777
	1	Simples	{0.01, 0.02, ..., 0.2}	0,51350	0,5100	0,0463	0,2267	0,1963	0,0714	0	-0,0298	0,0713
		Exp.	{0.01, 0.02, ..., 0.2}		0,5107	0,0475		0,1960	0,0681	0	-0,0322	0,0724
	0,8	Simples	{0.01, 0.02, ..., 0.1}	0,47792	0,4734	0,0448	0,2501	0,2285	0,0548	0	-0,0296	0,0761
		Exp.	{0.01, 0.02, ..., 0.2}		0,4677	0,0457		0,2270	0,0535	0	-0,0269	0,0763
	0,5	Simples	{0.01, 0.02, ..., 0.1}	0,42379	0,4064	0,0446	0,2481	0,2300	0,0616	0	-0,0182	0,0932
		Exp.	{0.01, 0.02, ..., 0.1}		0,4077	0,0446		0,2243	0,0632	0	0,0428	0,0950

Tabela 1 - Continuação.

N	α	Método	s	I(1)	$\hat{I}^*(1)$	MAD_1	I(2)	$\hat{I}^*(2)$	MAD_2	I(3)	$\hat{I}^*(3)$	MAD_3
1000	2	Simples	{0.01, 0.1, 0.2}	0,6772	0,6749	0,0133	0,1782	0,1733	0,0330	0	-0,0045	0,0357
		Exp.	{0.01, 0.06, 0.11, 0.16, 0.21}		0,6759	0,0131		0,1748	0,0313	0	-0,0014	0,0367
	1,8	Simples	{0.1, 0.2, 0.3}	0,6470	0,6452	0,0156	0,1924	0,1891	0,0303	0	-0,0038	0,0352
		Exp.	{0.1, 0.2, 0.3}		0,6459	0,0156		0,1898	0,0304	0	-0,0035	0,0350
	1,5	Simples	{0.1, 0.2, 0.3}	0,5990	0,5992	0,0167	0,2134	0,2086	0,0274	0	-0,0030	0,0358
		Exp.	{0.1, 0.2, 0.3}		0,5983	0,0164		0,2102	0,0264	0	-0,0035	0,0340
	1,3	Simples	{0.01, 0.02, ..., 0.2}	0,5655	0,5663	0,0160	0,2267	0,2233	0,0200	0	-0,0029	0,0259
		Exp.	{0.1, 0.2, 0.3}		0,5642	0,0157		0,2230	0,0249	0	-0,0032	0,0441
	1	Simples	{0.01, 0.02, ..., 0.2}	0,5135	0,5136	0,0140	0,2432	0,2406	0,0179	0	-0,0043	0,0257
		Exp.	{0.01, 0.02, ..., 0.2}		0,5135	0,0139		0,2411	0,0177	0	-0,0044	0,0263
	0,8	Simples	{0.01, 0.02, ..., 0.1}	0,4779	0,4776	0,0137	0,2501	0,2487	0,0157	0	-0,0021	0,0226
		Exp.	{0.01, 0.02, ..., 0.2}		0,4773	0,0132		0,2463	0,0182	0	-0,0034	0,0229
	0,5	Simples	{0.01, 0.02, ..., 0.1}	0,4238	0,4223	0,0131	0,2481	0,2455	0,0179	0	0,4223	0,0131
		Exp.	{0.01, 0.02, ..., 0.1}		0,4225	0,0133		0,2468	0,0186	0	0,4225	0,0133

Tabela com os valores de $I(\cdot)$ e sua estimativa $\hat{I}(\cdot)$ a partir do processo MA(2), onde $X_t = a_t + 2a_{t-1} + 1.111a_{t-2}$ para $T = 1000$ replicações em amostras de tamanho 100 e 1000. O processo a_t é SaS para alguns valores de α e $\sigma = 1$. Na tabela,

$\hat{I}^*(i) = \frac{1}{T} \sum_{j=1}^T \Re \hat{I}(i)_j$, e $MAD_i = \frac{1}{T} \sum_{j=1}^T |\Re \hat{I}(i)_j - I(i)|$, $i = 1, 2$, onde $\Re \hat{I}(i)_j$ é a estimativa no lag i da j -ésima repetição.

Como esperado, observamos que a precisão da estimação é melhorada quando o tamanho da amostra cresce. Além disso, por meio da análise das simulações, os resultados indicam que a precisão das estimativas da função codiferença normalizada depende da escolha dos pontos do vetor \mathbf{s} , onde a escolha ótima dos pontos depende do parâmetro α e o tamanho da amostra N . Quando $\alpha = 2$, surpreendentemente sob a escolha de pontos adequados, foi descoberto que para alguns casos em que a função codiferença normalizada pode oferecer uma estimação melhor para a $I(\cdot)$ do que a estimativa dada pela função de autocorrelação amostral. Outro resultado encontrado quando $\alpha = 2$, é que parece ser vantajoso calcular $\mathfrak{R}\hat{I}(\cdot)$ em mais de um ponto, pois, sob a escolha de pontos adequados, a performance dos métodos de ponderação são aproximadamente os mesmos. Para os modelos MA(2), considerados em todos os casos da simulação, foi encontrado que a precisão é significativamente melhor se os pontos de \mathbf{s} estão no intervalo $0.01 \leq s \leq 0.5$. Em geral, há benefício em termos de precisão na estimação ao se incluir um ponto próximo de zero. Foi observado ainda que quando $\alpha < 1,5$, a escolha de pontos equidistantes com distância entre 0.01 e 0.05 aparenta ser mais adequado. Para $\alpha \geq 1,5$, a distância de 0,1 aparenta ser adequada, pois ao se utilizar uma distância menor não houve ganhos de precisão na estimativa. No geral, conclui-se a partir das simulações que a escolha dos pontos de \mathbf{s} seguem as mesmas linhas descritas na Seção 4.2.1.

5 Aplicação

Sabe-se que no mercado financeiro é comum encontrar séries com caudas pesadas e assimetria, neste caso a utilização do modelo com inovações α -estáveis tende a apresentar um ajuste melhor que o modelo gaussiano (ROSADI; DEISTLER, 2011). Para constatar a afirmação, vamos trabalhar com o log retorno diário, através da fórmula $r_t = \log(\frac{P_t}{P_{t-1}})$, onde P_t é o preço de um ativo no tempo t , $t = 1, \dots, N - 1$.

A série temporal analisada no trabalho são os preços diários do fechamento das ações da Usinas Siderúrgicas de Minas Gerais S.A (Usiminas) extraídos do site do Yahoo Finanças (<http://finance.yahoo.com>). Os dados são relativos ao período de 3 de janeiro de 2000 a 31 de dezembro de 2015 totalizando 3962 observações. A Tabela 2 apresenta um resumo com algumas estatísticas descritivas da série analisada.

Tabela 2 – Resumo estatístico do log retorno

Medida	Resultado
Mínimo	-0.1885
Máximo	0.6668
Média	0.0004
Mediana	0.000
Desv. Pad.	0.0340
Curtose	4.2842
Assimetria	2.2429
N	3961

O índice de assimetria positivo indica assimetria a direita e a curtose maior que 3 evidencia que a curva é leptocúrtica, ou seja, tem caudas pesadas. No caso de normalidade, estas medidas deveriam ser iguais ou próximas de zero. Para confirmar o estado de não normalidade dos dados, foram realizados alguns testes de hipóteses para distribuição normal e o gráfico quantil-quantil.

Tabela 3 – Testes de normalidade para o log retorno

Teste	Estatística	p-valor
Lilliefors (Kolmogorov-Smirnov)	0.05	0.00
Shapiro-Wilk	0.89	0.00
Shapiro-Francia	0.89	0.00
Anderson-Darling	27.97	0.00

Em todos os testes de hipóteses utilizados, a hipótese de que os dados sigam uma distribuição normal foi rejeitada e o gráfico quantil-quantil da distribuição nor-

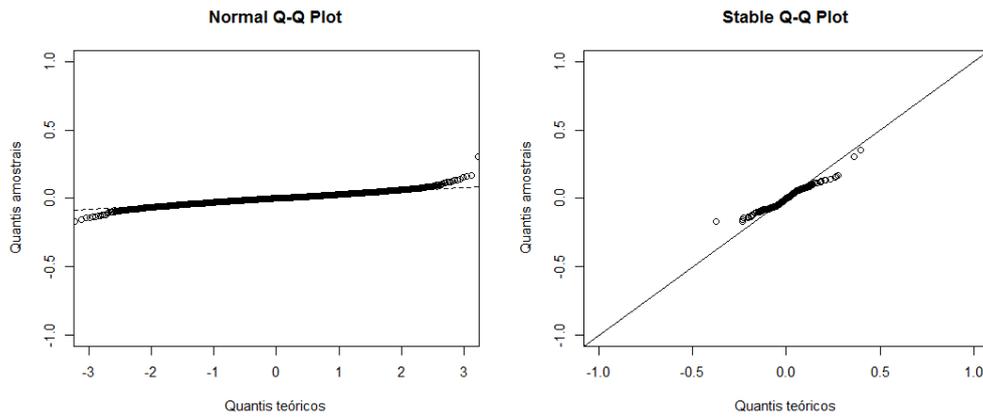


Figura 10 – Gráficos quantil-quantil do log retorno

mal reforça os resultados dos testes indicando fuga de normalidade. Para checar a estabilidade dos dados, estimamos os parâmetros da α -estável dos log retornos para fazer o gráfico quantil-quantil estável. O gráfico apresenta desvio em alguns pontos, mas em sua maioria estão posicionados de forma linear.

Para estimar os parâmetros da distribuição estável tínhamos quatro opções disponíveis no R Core Team (2014) versão 3.1.2: a função *Estim* do pacote **StableEstim** que permite a estimação dos parâmetros pelo Estimador de Máxima Verossimilhança (EMV) e método de regressão proposto por Koutrouvelis (1981) e a função *stableFit* do pacote **fBasics** que realiza a estimação por máxima verossimilhança e método dos quartis de acordo com McCulloch (1986). Porém, as funções que realizam a estimação pelo método de máxima verossimilhança gastaram mais tempo na execução, chegando a 44.25 minutos no caso da função *Estim*. A função mais rápida foi a *stableFit* pelo método de McCulloch, com 0.04 segundos. O critério de escolha da estimativa a ser utilizada ao longo do trabalho foi comparar a densidade da α -estável estimada que melhor se ajustou a densidade empírica dos log retornos. Analisando os gráficos das densidades empíricas (ver Figura 11), todas as estimativas estão próximas à densidade empírica dos logretornos, porém, a de McCulloch foi a que melhor acompanhou a densidade empírica ao longo dos pontos da curva. Portanto, estes serão os parâmetros da α -estável adotado nas próximas fases da análise.

Tabela 4 – Parâmetros da α -estável estimado para os log retornos.

Método (função)	α	β	σ	μ	Tempo de execução
Koutrouvelis (Estim)	1.8124	-0.1718	0.0194	-0.0001	10.97 s
EMV (Estim)	1.8600	-0.1850	0.0201	0.0000	44.24 min
EMV (stableFit)	1.8379	-0.1158	0.0199	0.0007	25.60 min
McCulloch (stableFit)	1.6850	-0.0590	0.0189	0.0002	0.04 s

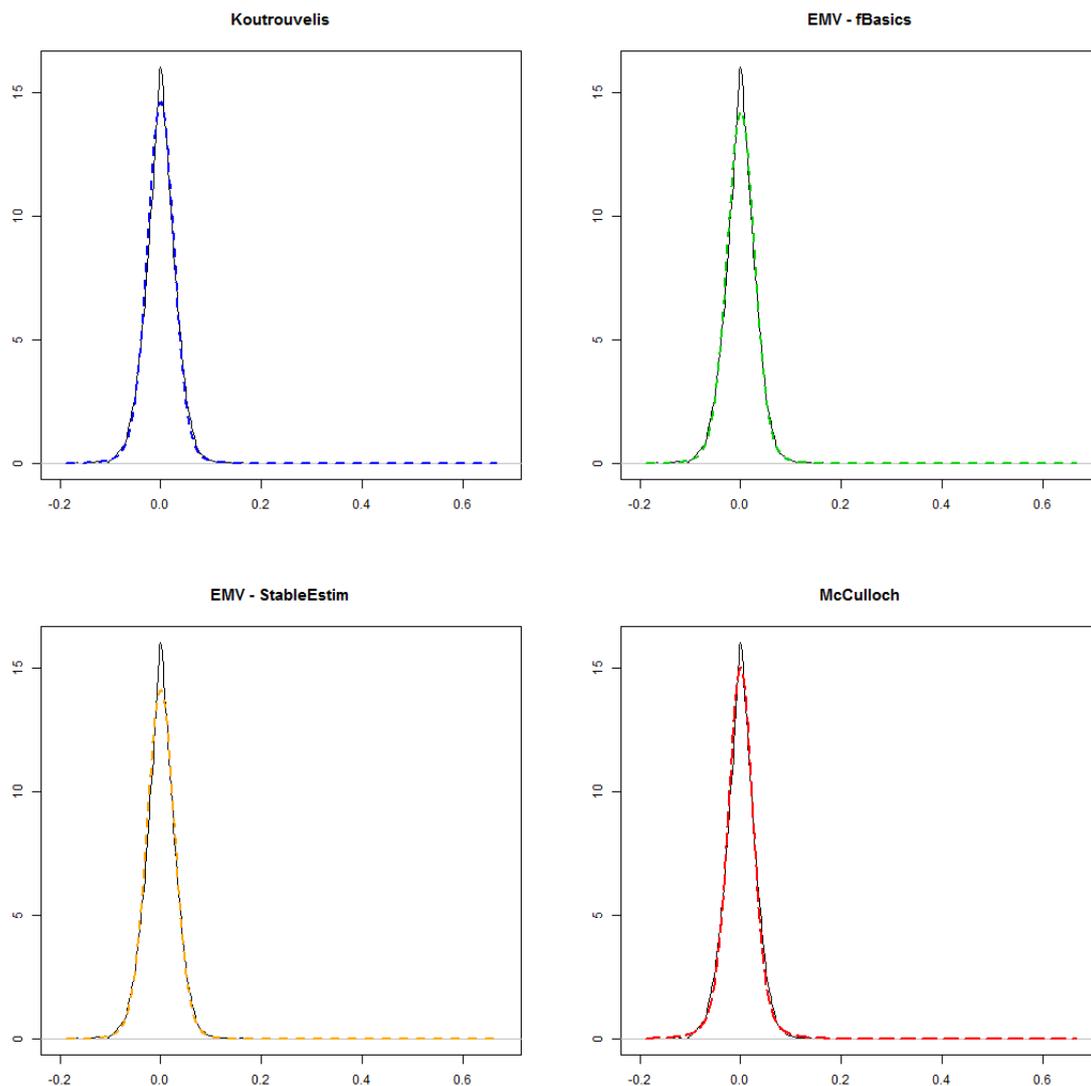


Figura 11 – Densidade do log retorno e densidade estimada para os parâmetros estimados.

Outra ferramenta para verificar a estabilidade dos dados consiste em estimar o índice α para diferentes níveis de agregação (diário, semanal e mensal, por exemplo). Para cada nível de agregação comparamos informalmente se o intervalo de confiança das estimativas de α (média das estimativas mais ou menos duas vezes o desvio padrão das estimativas) estão abaixo de 2,0 e rejeitamos a possibilidade de usar a distribuição estável se eles ultrapassam o limite de 2,0 (PAOLELLA, 2001). Os níveis de agregação de tempo foram $j = 1, \dots, 8$ dias, onde para $j > 1$ foi calculada a média dos j dias sem sobreposição. Os resultados das estimativas estão disponíveis na Tabela 5.

O intervalo de confiança para as estimativas dos índices de estabilidade ($\bar{\hat{\alpha}} \pm 2 * EP(\hat{\alpha})$) de foi 1.74 ± 0.1993 . E o gráfico da densidade dos dados para alguns níveis de

Tabela 5 – Estimativa de α para diferentes níveis de agregação.

Nível de agregação (dias)	N	$\hat{\alpha}$
1	3961	1.68
2	1980	1.80
3	1320	1.71
4	990	1.76
5	792	1.78
6	660	1.81
7	565	1.71
8	495	1.64

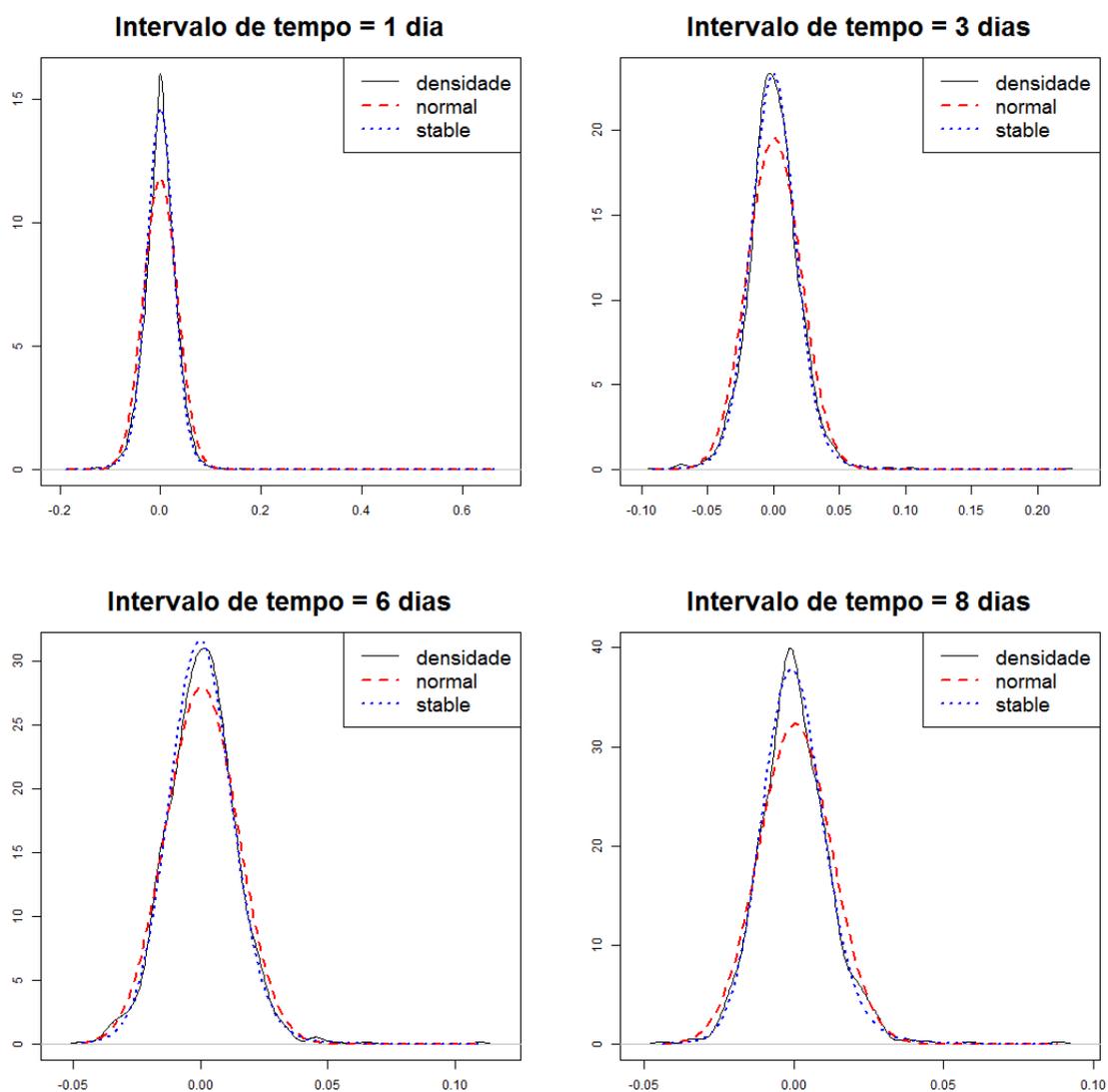


Figura 12 – Gráfico da densidade do log retorno em diferentes níveis de agregação

agregação indicam claramente que a distribuição estável se ajusta melhor aos dados (ver Figura 12).

Dado os indícios dos dados seguirem a distribuição estável, vamos utilizar a

função codiferença padronizada amostral (3.7) na identificação da ordem do modelo. Para obter uma estimativa com boa precisão da função codiferença padronizada, os resultados das simulações evidenciam a importância na escolha dos pontos. Para auxiliar na seleção dos pontos plotamos o gráfico de $\hat{\mathcal{R}}(k)$ no intervalo $s \in [0.01, 2]$, para $k = 1$ e 2 . Nos dois gráficos, a curva tem comportamento suave, indicando que $\hat{\alpha}$ é próximo de 2, em concordância com o resultado da Tabela 5. Portanto, segundo a Figura 13, recomenda-se que os pontos estejam no intervalo $0 < s < 0.5$ com distanciamento de 0.05 a 0.1 entre eles. Como $\alpha > 1.5$, recomenda-se a escolha de pelo menos dois pontos para estimar a função codiferença.

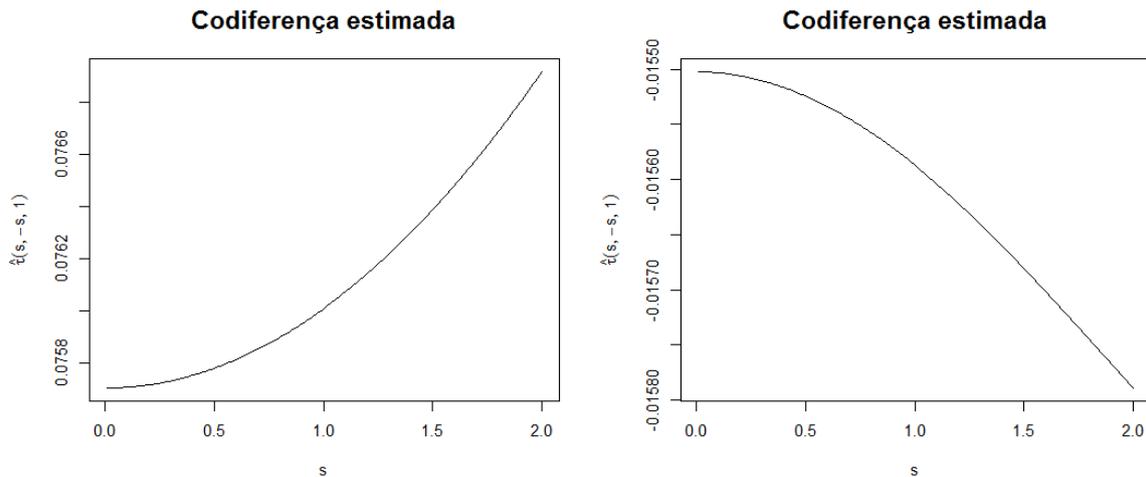


Figura 13 – Gráfico da codiferença amostral estimada ($\hat{\mathcal{R}}(1)$)

Na Figura 14 temos os gráficos de $\hat{\mathcal{R}}(k)$ para $\mathbf{s}_1 = (0.01, 0.1, 0.2)$, $\mathbf{s}_2 = (0.01)$ e $\mathbf{s}_3 = (0.01, 0.06, 0.11)$, sendo que, para \mathbf{s}_3 foi utilizado a média ponderada exponencial e média aritmética para outros casos. Comparando o gráfico da codiferença estimada com a FAC amostral, percebe-se que esta responde à estrutura de dependência dos dados de forma semelhante a FAC amostral. O benefício em utilizar a função codiferença é que por definição ela está definida para todas as funções, ao contrário da FAC que só existe para distribuições com variância finita.

Apesar da FAC e FACP não existirem para séries com caudas pesadas, seus estimadores amostrais podem ser calculados e se mostram excelentes ferramentas na identificação da ordem em séries com distribuição estável (Adler et al, 1998). Os gráficos da codiferença e FAC amostral (ver Figura 14) indicam uma ordem máxima $q = 3$. A FACP amostral apresentada na Figura 15 sugere que a ordem p pode ser igual a 1, 3 ou 15. Com o objetivo de estimar um modelo parcimonioso, vamos testar ajustes com ordem máxima $p = q = 3$.

Pelo ajuste gaussiano, os modelos indicados pelo critério AIC e AICc que conversam com os gráficos de identificação da ordem (FAC e FACP) foi um ARMA(1,3)

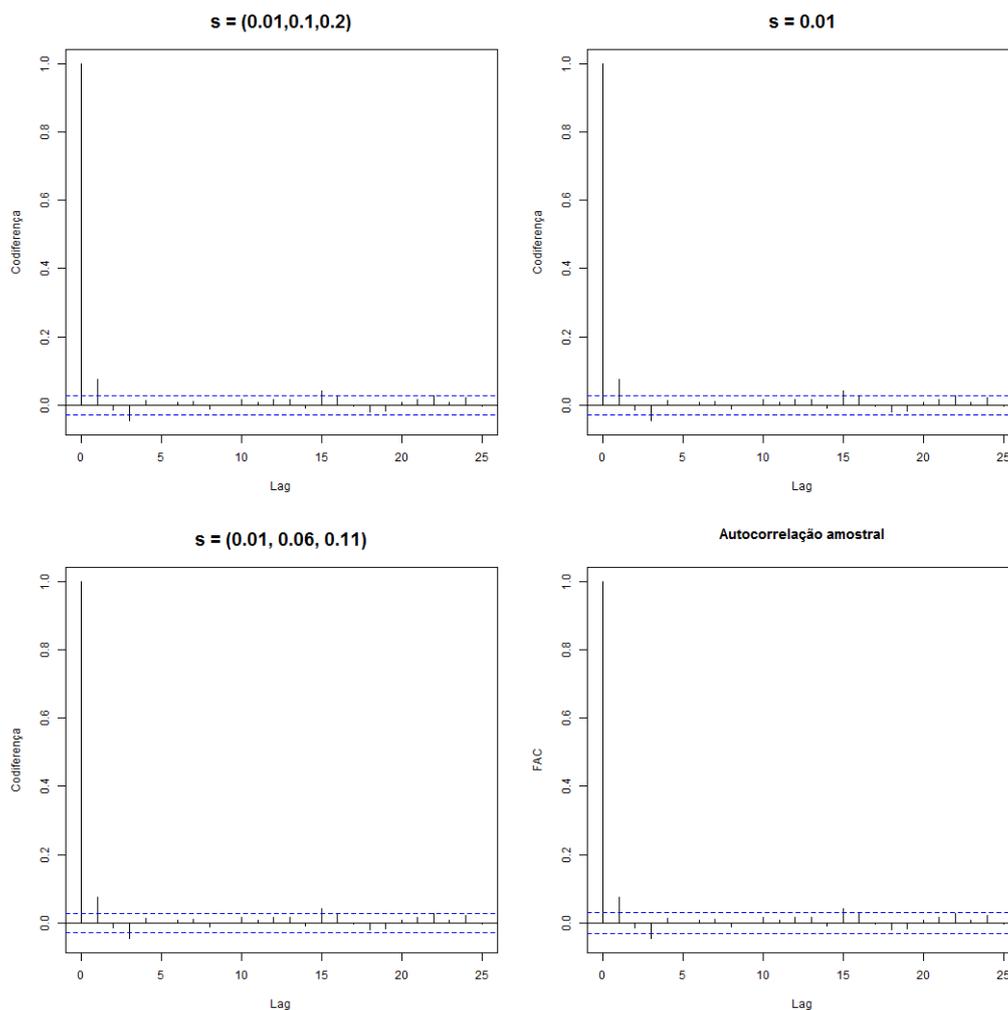


Figura 14 – Gráfico Codiferença padronizada estimada para $s=(0.01,0.1,0.2)$, 0.01 e $(0.01,0.06,0.11)$ e FAC

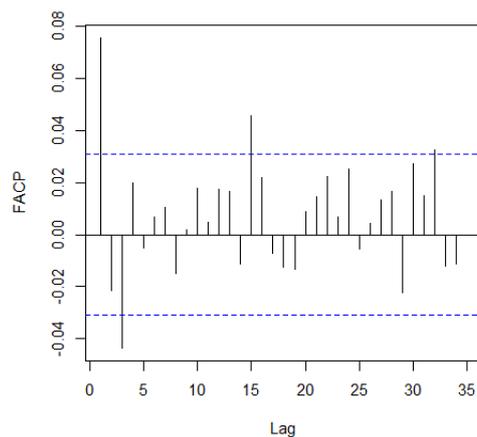


Figura 15 – Autocorrelação parcial dos log retornos

e o critério BIC indica o ajuste de um modelo AR(1). Analisando os gráficos da codiferença e FACP, será ajustado também um ARMA(1,1).

Primeiro serão estimados os parâmetros do modelos assumindo que os erros tem distribuição normal e depois considerando erros com distribuição α -estável. Na comparação entre os resultados, é esperado que os ajuste com a α -estável seja o melhor. A estimação foi realizada pelo método de máxima verossimilhança, no caso dos erros normais foi utilizada a função *arima* do pacote *stats* e para os erros α -estáveis implementamos um algoritmo para calcular a função de verossimilhança condicional. A escolha da máxima verossimilhança foi devido a limitação da representação por fórmula fechada da função de densidade da variável α -estável. O resultado da verossimilhança estimada foi maximizado através da função *optim*.

Os parâmetros estimados para o modelo ARMA(3,1) são apresentados na Tabela 6. O diagnóstico foi realizado a partir dos gráficos da função codiferença, FAC e FACP dos resíduos e teste de Box-Pierce-Ljung. A Figura 16 apresenta a codiferença amostral, FAC e FACP para os resíduos do modelo ARMA(3,1) com erros normais na primeira linha e com erros α -estáveis na segunda linha. Todos os gráficos indicam que os resíduos constituem ruído branco, apesar da FACP ultrapassar o limite no lag 15. Entretanto, o teste de Box-Pierce-Ljung (Tabela 7) revela que não há evidências para rejeitar hipótese de que os resíduos sejam não correlacionados.

Tabela 6 – Ajuste do modelo ARMA(3,1) para erros normais e α -estáveis.

Ajuste 1	ar1	ar2	ar3	ma1	AIC	BIC	EQM
Normal	-0.2567	0.0078	-0.0526	0.3340	-15577.33	-15547.91	0.0011
α -estável	-0.0496	0.0066	-0.0430	0.1424	16336.85	16366.27	0.0011

Tabela 7 – Teste Box-Pierce-Ljung para os resíduos do ajuste ARMA(3,1).

Lag	Ajuste Normal	Ajuste α -estável
1	0.99	0.34
2	1.00	0.46
3	1.00	0.65
4	1.00	0.63
5	1.00	0.76
6	1.00	0.84
7	1.00	0.85
8	0.99	0.85
9	1.00	0.91
10	0.98	0.86
11	0.99	0.91
12	0.98	0.88
13	0.95	0.84
14	0.95	0.84
15	0.59	0.44

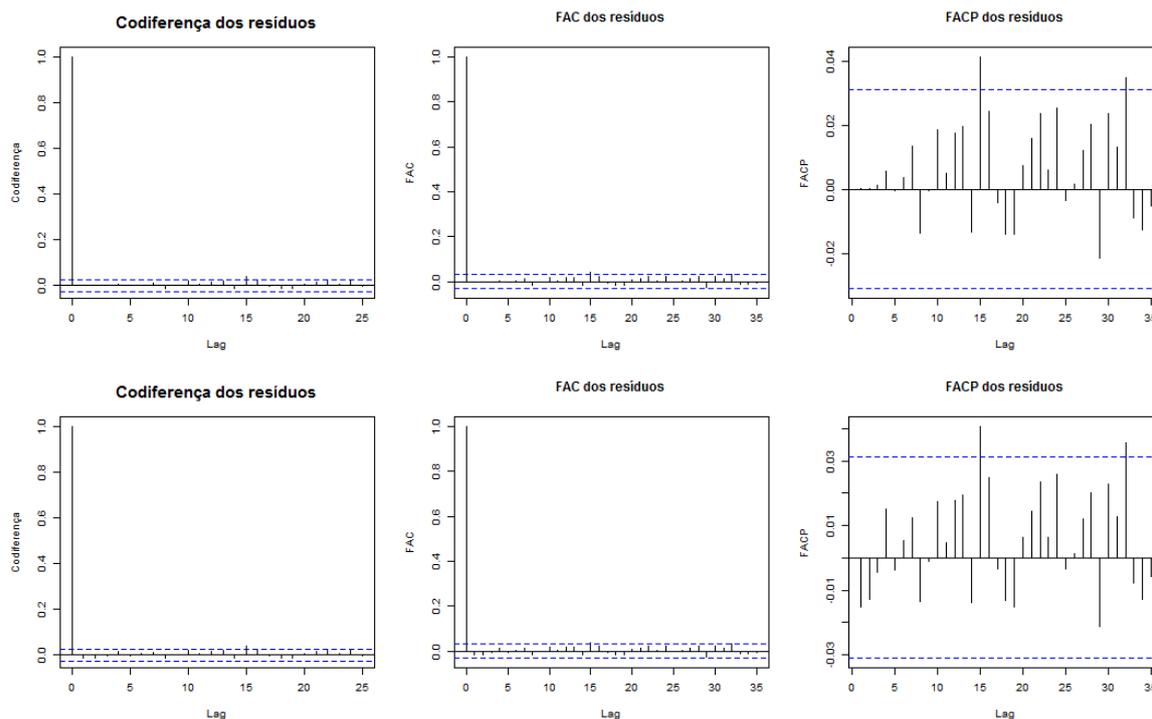


Figura 16 – Codiferença, FAC e FACP dos resíduos do ajuste ARMA(3,1). A primeira linha são os resíduos do ajuste normal e a segunda o ajuste com erros α -estáveis.

A Tabela 8 exibe os resultados dos parâmetros estimados para o modelo selecionado pelo critério BIC, o AR(1). Os gráficos da função codiferença, FAC e FACP amostral dos resíduos de ambos ajustes indicam ruído branco, exceto quando a FACP dos dois resíduos ultrapassam o intervalo de confiança nos lags 3 e 15, resultado detectado pelo teste de hipóteses de Box-Pierce-Ljung (Tabela 9).

Tabela 8 – Ajuste do modelo AR(1) para erros normais e α -estáveis.

Ajuste 2	ar1	AIC	BIC	EQM
Normal	0.0757	-15572.34	-15561.77	0.0011
α -estável	0.0927	16330.81	16341.37	0.0011

O último modelo ajustado foi um ARMA(1,1), escolhido por ser um modelo parcimonioso e atende o critério da análise dos gráficos da função codiferença, FAC e FACP amostrais. A Figura 18 indica que os resíduos dos dois ajustes são ruído branco, mas assim como no ajuste do AR(1) os resíduos para o ajuste com erros normais e α -estáveis ultrapassam o intervalo de confiança no lag 3 e 15, refletindo nos lags 2 e 3 do teste de Box-Pierce-Ljung (Tabela 11).

Continuando a análise dos resíduos, foram produzidos os gráficos quantil-quantil. Os resíduos devem seguir a mesma distribuição dos erros do modelo, portanto,

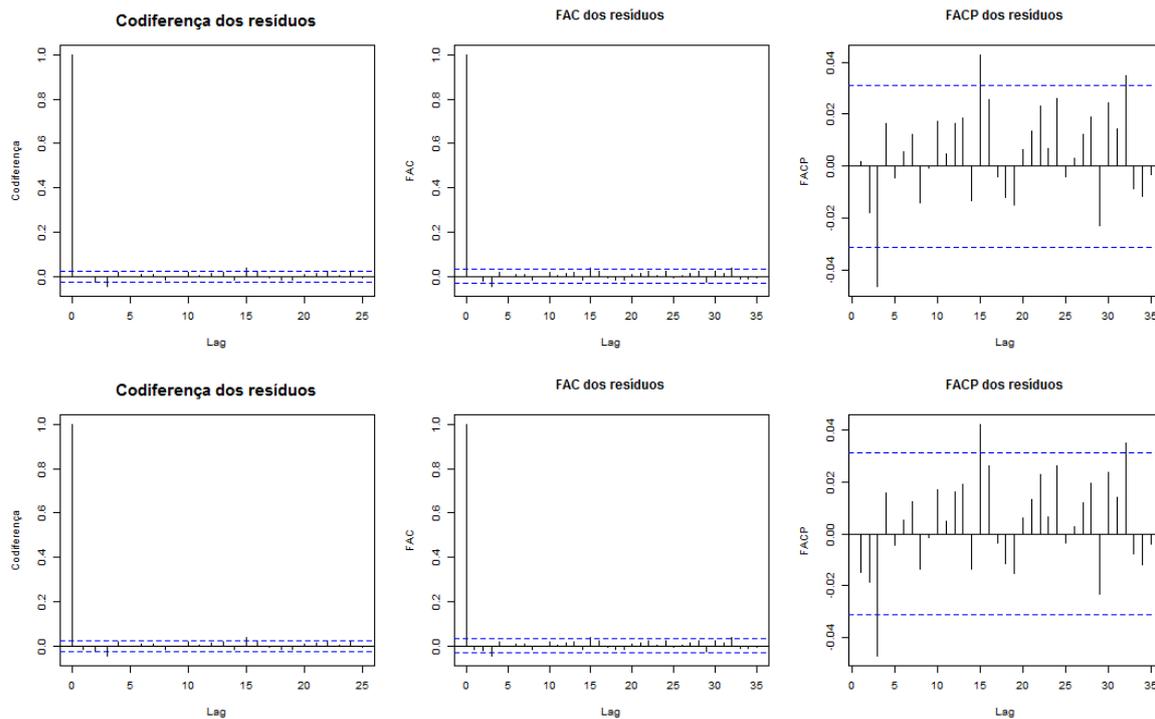


Figura 17 – Codiferença, FAC e FACP dos resíduos do ajuste AR(1). A primeira linha são os resíduos do ajuste normal e a segunda o ajuste com erros α -estáveis.

Tabela 9 – Teste Box-Pierce-Ljung para os resíduos do ajuste AR(1).

Lag	Ajuste Normal	Ajuste α -estável
1	0.92	0.35
2	0.52	0.33
3	0.02	0.01
4	0.03	0.02
5	0.05	0.03
6	0.08	0.05
7	0.11	0.08
8	0.13	0.09
9	0.19	0.14
10	0.19	0.14
11	0.25	0.19
12	0.26	0.20
13	0.26	0.21
14	0.28	0.22
15	0.08	0.06

para os ajustes gaussiano foi gerado o gráfico quantil-quantil com os quantis da normal e no ajuste com inovações α -estáveis o gráfico quantil-quantil com os quantis da distribuição estável.

Tabela 10 – Ajuste do modelo ARMA(1,1) para erros normais e α -estáveis

Ajuste 3	ar1	ma1	AIC	BIC	EQM
Normal	-0.0460	0.1232	-15571.15	-15554.30	0.0011
α -estável	0.0373	0.0562	16333.08	16349.93	0.0011

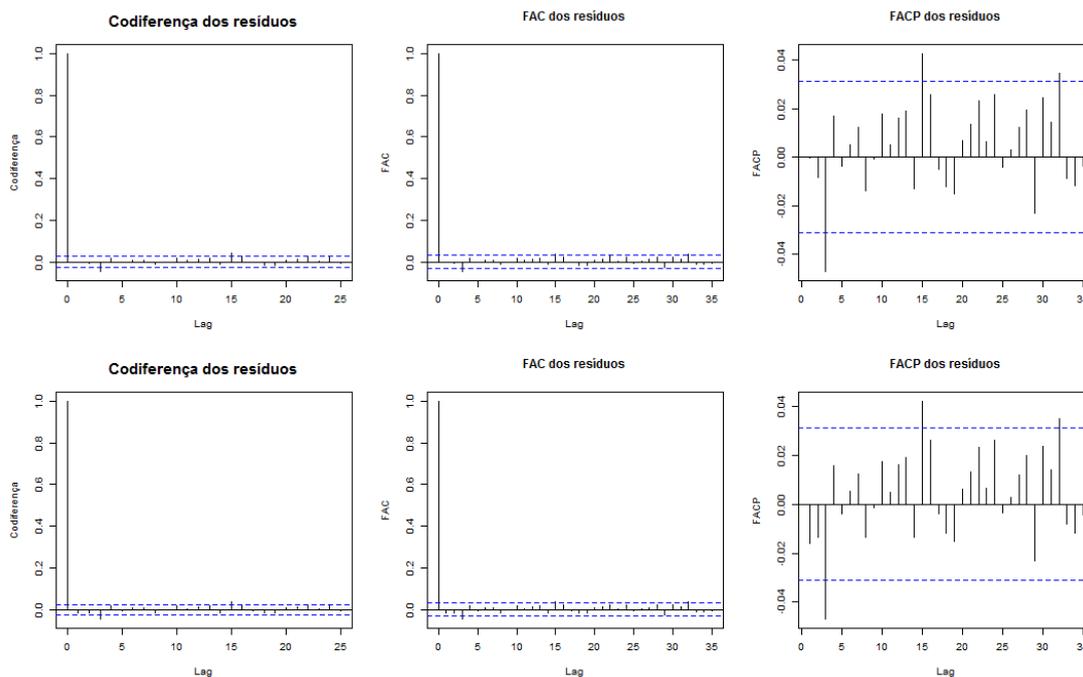


Figura 18 – Codiferença, FAC e FACP dos resíduos do ajuste ARMA(1,1). A primeira linha são os resíduos do ajuste normal e a segunda o ajuste com erros α -estáveis.

Tabela 11 – Teste Box-Pierce-Ljung para os resíduos do ajuste ARMA(1,1).

Lag	Ajuste Normal	Ajuste α -estável
1	0.99	0.31
2	0.87	0.42
3	0.03	0.02
4	0.04	0.02
5	0.07	0.04
6	0.11	0.07
7	0.14	0.09
8	0.17	0.11
9	0.24	0.16
10	0.24	0.17
11	0.30	0.22
12	0.31	0.23
13	0.30	0.23
14	0.33	0.25
15	0.10	0.07

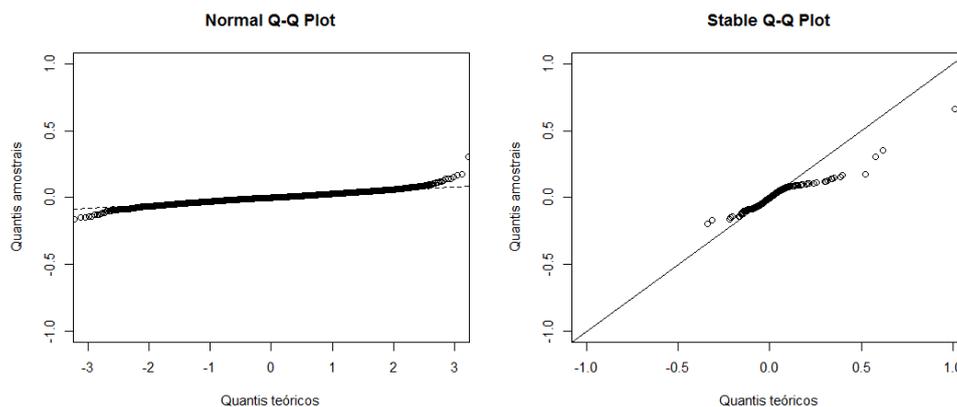


Figura 19 – Gráficos quantil-quantil do resíduos do ajuste ARMA(3,1).

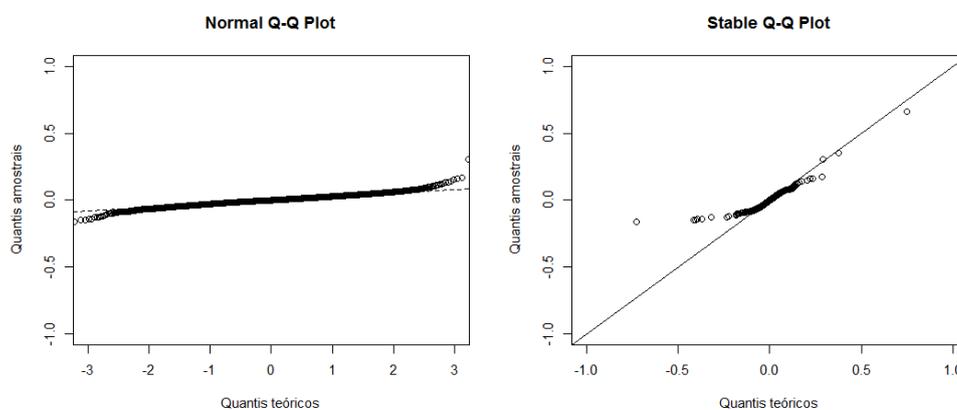


Figura 20 – Gráficos quantil-quantil do resíduos do ajuste AR(1).

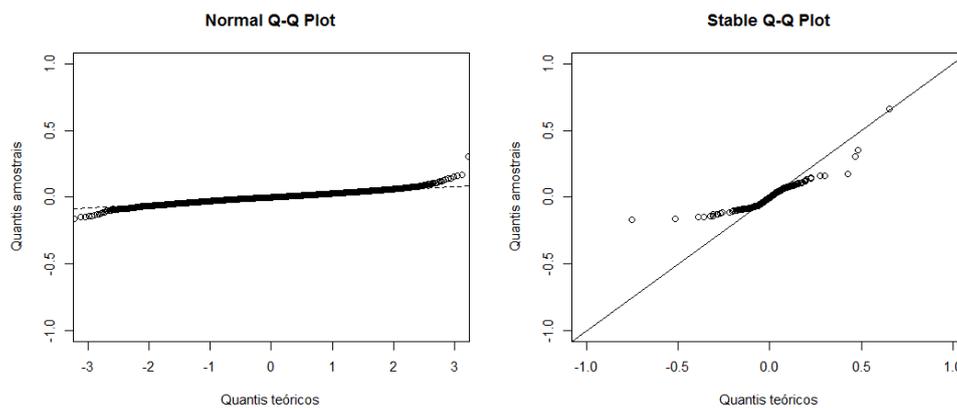


Figura 21 – Gráficos quantil-quantil do resíduos do ajuste ARMA(1,1).

Observa-se nas Figuras 19 a 21 que os resíduos gaussianos não se comportam como o quantis da normal em nenhum dos modelos ajustados. Apesar da fuga de alguns pontos nas caudas, os resíduos dos ajustes com inovações α -estáveis demonstram seguir distribuição estável.

Complementando a análise dos qqplots realizamos o teste Kolmogorov-Smirnov. Sendo a hipótese nula: os dados seguem distribuição gaussiana foi adotada para os resíduos do ajuste gaussiano e para o ajuste com inovações α -estáveis foi testada a hipótese nula que os resíduos seguem a distribuição estável. Os resultados podem ser consultados na Tabela 12.

Tabela 12 – Teste Kolmogorov-Smirnov para os resíduos dos modelos ajustados.

Modelo	Distribuição dos resíduos	p-valor
ARMA(3,1)	Normal	2.2e-16
	α -estável	0.3796
AR(1)	Normal	2.2e-16
	α -estável	0.3886
ARMA(1,1)	Normal	2.2e-16
	α -estável	0.3503

Não há evidências para rejeitar a hipótese nula de distribuição estável para os resíduos dos modelos com inovações α -estáveis, ao contrário dos testes para resíduos gaussianos, onde os p-valores de todos os três modelos indicam que a hipótese nula é rejeitada. Entretanto, é preciso usar o teste Kolmogorov-Smirnov com cautela uma vez que é conhecida sua baixa sensibilidade na região das caudas e os qqplots para os resíduos dos ajustes com inovações α -estáveis apresentam desvio maior na região das caudas em comparação com o resíduos do ajuste com erros gaussianos.

No geral, a julgar pelas ferramentas disponíveis, o melhor modelo para os log retornos dos preços do fechamento das ações da Usiminas é o AR(1), pois, entre os três modelos com inovações α -estáveis ajustados, foi o que apresentou menor AIC e BIC, gráfico quartil-quartil estável com mais pontos sob a reta diagonal e teste Kolmogorov-Smirov com maior p-valor.

6 Conclusão

Neste trabalho estudamos a identificação da ordem e estimação dos parâmetros do modelo ARMA com inovações α -estáveis para dados com caudas pesadas. Para a identificação da ordem, foi utilizado o estimador da função codiferença padronizada proposto por [Rosadi e Deistler \(2011\)](#). Apesar dos estudos de simulação de [Adler et al. \(1998\)](#) apontarem resultados satisfatórios da FAC amostral na identificação da ordem das séries com inovações estáveis, a FAC que tem distribuição limitante não familiar quando $\alpha < 2$ e quantis da distribuição limitante relativamente complicados de serem obtidos. Os estimadores da função codiferença e codiferença padronizada são consistentes e normalmente distribuídos assintoticamente. Os parâmetros do modelo foram estimados pelo método de máxima verossimilhança condicional.

Os dados analisados foram os log retornos dos preços do fechamento das ações da Usiminas no período de 03/01/2000 a 31/12/2012, totalizando 3961 observações. Os resultados da estimação com inovações α -estáveis foram comparados com a estimação tradicional do modelo gaussiano para os três modelos selecionados (AR(1), ARMA(1,1) e ARMA(3,1)) e, como esperado, o ajuste estável apresentou melhor resultado segundo as ferramentas de análise de resíduos, gráfico quantil-quantil e teste Kolmogorov-Smirnov. Sendo que o modelo AR(1) foi o melhor entre os três ajustes, segundo a análise de resíduos.

6.1 Trabalhos futuros

O trabalho de aplicação do ajuste de modelo ARMA com inovações α -estáveis em um conjunto de dados reais abriu caminho para elaboração de estudos futuros para desenvolver ferramentas como a FACP e critérios AIC e BIC voltados para distribuições com caudas pesadas, adaptação de teste hipóteses para detectar a distribuição dos dados que seja mais sensível nas caudas e até mesmo otimizar a rotina de estimação dos parâmetros por máxima verossimilhança condicional. Os tópicos a seguir são sugestões de trabalhos futuros que podem dar continuidade aos resultados apresentados:

- Desenvolver ferramenta semelhante a estrutura da função de autocovariância parcial a partir da função codiferença.
- Implementar estimação de parâmetros para modelo ARMA com inovações t de

student e comparar com o ajuste do modelo ARMA gaussiano e com inovações α -estáveis.

- Estudo sobre critérios de seleção de modelos envolvendo distribuições estáveis.
- Estudo sobre o diagnóstico de modelos com inovações estáveis.

Referências

- ADLER, R. J.; FELDMAN, R. E.; GALLAGHER, C. Analysing stable time series. *A practical guide to heavy tails. Statistical techniques and applications*, p. 133–158, 1998.
- AKAIKE, H. On entropy maximization principle. *Application of statistics*, North-Holland, 1977.
- BHANSALI, R. Estimation of the order of a moving average model from autoregressive and window estimates of the inverse correlation function. *Journal of Time Series Analysis*, Wiley Online Library, v. 4, n. 3, p. 137–162, 1983.
- BOX, G. E.; JENKINS, G. M. *Time series analysis: forecasting and control*. [S.l.]: Holden-Day, 1970.
- BOX, G. E. P.; COX, D. R. An analysis of transformations. *Journal of the Royal Statistical Society*, v. 26, n. 2, p. 211–252, 1964.
- BROCKWELL, P. J.; DAVIS, R. A. *Introduction to time series and forecasting*. New York: Springer-Verlag, 1991.
- CHAMBERS, J. M.; MALLOWS, C. L.; STUCK, B. A method for simulating stable random variables. *Journal of the american statistical association*, Taylor & Francis Group, v. 71, n. 354, p. 340–344, 1976.
- DAVIS, R.; RESNICK, S. Limit theory for the sample covariance and correlation functions of moving averages. *The Annals of Statistics*, JSTOR, p. 533–558, 1986.
- EMBRECHTS, P.; KLÜPPELBERG, C.; MIKOSCH, T. Modelling extremal events, vol. 33 of. *Applications of Mathematics (New York)*, 1997.
- FAMA, E. F.; ROLL, R. Some properties of symmetric stable distributions. *Journal of the American Statistical Association*, Taylor & Francis, v. 63, n. 323, p. 817–836, 1968.
- FELLER, W. An introduction to probability and its applications, vol. ii. *Wiley, New York*, 1971.
- GNEDENKO, B. V.; KOLMOGOROV, A. N. Limit distributions for sums of independent. *Amer. J. Math.*, Addison-Wesley, v. 105, p. 28–35, 1954.
- HANNAN, E. Testing for autocorrelation and akaike's criterion. *Journal of Applied Probability*, JSTOR, p. 403–412, 1982.
- HANNAN, E. J. The estimation of the order of an arma process. *The Annals of Statistics*, JSTOR, p. 1071–1081, 1980.

- HESSE, C. Rates of convergence for the empirical distribution function and the empirical characteristic function of a broad class of linear processes. *Journal of Multivariate Analysis*, Elsevier, v. 35, n. 2, p. 186–202, 1990.
- HONG, Y. Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 94, n. 448, p. 1201–1220, 1999.
- HURVICH, C. M.; TSAI, C.-L. Regression and time series model selection in small samples. *Biometrika*, Biometrika Trust, v. 76, n. 2, p. 297–307, 1989.
- JANICKI, A.; WERON, A. *Simulation and chaotic behavior of alpha-stable stochastic processes*. [S.l.]: CRC Press, 1993. v. 178.
- KANTER, M. et al. Stable densities under change of scale and total variation inequalities. *The Annals of Probability*, Institute of Mathematical Statistics, v. 3, n. 4, p. 697–707, 1975.
- KOGON, S. M.; WILLIAMS, D. B. Characteristic function based estimation of stable distribution parameters. *A practical guide to heavy tails: statistical techniques and applications*, Birkhäuser Boston, MA, p. 311–338, 1998.
- KOKOSZKA, P. S.; TAQQU, M. S. Infinite variance stable arma processes. *Journal of Time Series Analysis*, Wiley Online Library, v. 15, n. 2, p. 203–220, 1994.
- KOUTROUVELIS, I. A. Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, Taylor & Francis, v. 75, n. 372, p. 918–928, 1980.
- KOUTROUVELIS, I. A. An iterative procedure for the estimation of the parameters of stable laws: An iterative procedure for the estimation. *Communications in Statistics-Simulation and Computation*, Taylor & Francis, v. 10, n. 1, p. 17–28, 1981.
- MCCULLOCH, J. H. Simple consistent estimators of stable distribution parameters. *Communications in Statistics-Simulation and Computation*, Taylor & Francis, v. 15, n. 4, p. 1109–1136, 1986.
- MORETTIN, P. A.; TOLOI, C. *Análise de séries temporais*. [S.l.]: Blucher, 2006.
- NIKIAS, C. L.; SHAO, M. *Signal processing with alpha-stable distributions and applications*. [S.l.]: Wiley-Interscience, 1995.
- PAOLELLA, M. S. Testing the stable paretian assumption. *Mathematical and Computer Modelling*, Elsevier, v. 34, n. 9, p. 1095–1112, 2001.
- PRESS, S. J. Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 67, n. 340, p. 842–846, 1972.
- QUENOUILLE, M. H. Approximate tests of correlation in time-series 3. In: CAMBRIDGE UNIV PRESS. *Mathematical Proceedings of the Cambridge Philosophical Society*. [S.l.], 1949. v. 45, n. 03, p. 483–484.

- RESNICK, S. I. et al. Heavy tail modeling and teletraffic data: special invited paper. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 25, n. 5, p. 1805–1869, 1997.
- RISSANEN, J. Modeling by shortest data description. *Automatica*, Elsevier, v. 14, n. 5, p. 465–471, 1978.
- ROSADI, D. Identification of moving average process with infinite variance. *Statistics & probability letters*, Elsevier, v. 77, n. 14, p. 1490–1496, 2007.
- ROSADI, D.; DEISTLER, M. Estimating the codifference function of linear time series models with infinite variance. *Metrika*, Springer, v. 73, n. 3, p. 395–429, 2011.
- ROSENFELD, G. Identification of time series with infinite variance. *Applied Statistics*, JSTOR, p. 147–153, 1976.
- SAMORODNITSKY, G.; TAQQU, M. S. *Stable non-Gaussian random processes: stochastic models with infinite variance*. [S.l.]: CRC press, 1994. v. 1.
- SCHWARZ, G. et al. Estimating the dimension of a model. *The annals of statistics*, Institute of Mathematical Statistics, v. 6, n. 2, p. 461–464, 1978.
- WERON, R. On the chambers-mallows-stuck method for simulating skewed stable random variables. *Statistics & probability letters*, Elsevier, v. 28, n. 2, p. 165–171, 1996.
- YANG, X.; PETROPULU, A. P.; PESQUET, J.-C. Estimating long-range dependence in impulsive traffic flows. In: IEEE. *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*. [S.l.], 2001. v. 6, p. 3413–3416.
- YU, J. Empirical characteristic function estimation and its applications. *Econometric reviews*, Taylor & Francis, v. 23, n. 2, p. 93–123, 2004.
- ZOLOTAREV, V. M. On representation of stable laws by integrals. *Selected translations in mathematical statistics and probability*, American Mathematical Society Providence, RI, v. 6, n. 1, p. 84–5, 1966.
- ZOLOTAREV, V. M. *One-dimensional stable distributions*. [S.l.]: American Mathematical Soc., 1986. v. 65.

APÊNDICE A – Demonstração do Teorema 3.2

Antes de apresentar os lemas que são necessários para provar a consistência do estimador da codiferença, um resultado extraído de Kokoszka e Taqqu (1994) será exposto, no qual mostra que a função codiferença no caso de um modelo ARMA é delimitada por uma função de decaimento exponencial assim como a função de covariância no caso clássico. Em Kokoszka e Taqqu (1994) é considerado uma definição mais geral para a função codiferença (para $\theta_1, \theta_2 \in \mathbb{R}$):

$$\tau_G(\theta_1, \theta_2; k) = \ln E \exp(i(\theta_1 X_{t+k} + \theta_2 X_t)) + \ln E \exp(i\theta_1 X_{t+k}) + \ln E \exp(i\theta_2 X_t). \quad (\text{A.1})$$

A equação (3.1) é um caso especial onde $\theta_1 = s$ e $\theta_2 = -s$.

Teorema A.1. (Kokoszka e Taqqu (1994), Teorema 2.1) *Considere o processo linear $X_t = \sum_{j=0}^{\infty} c_j a_{t-j}$, onde os coeficientes c_j 's são reais e satisfazem $|c_j| < Q^{-j}$ para algum $c > 0$, $Q > 1$ e a_t é i.i.d. com distribuição α -estável simétrica. Então*

$$\limsup_{k \rightarrow \infty} Q^{\alpha k} |\tau_G(k)| \leq 2(1 - Q^\alpha)^{-1/\alpha} |\theta_1|^\alpha \text{ para } 0 < \alpha \leq 1 \quad (\text{A.2})$$

e

$$\limsup_{k \rightarrow \infty} Q^k |\tau_G(k)| \leq \alpha \left(\sum_{j=0}^{\infty} |c_j|^\alpha \right)^{\frac{\alpha-1}{\alpha}} (1 - Q^\alpha)^{-1/\alpha} |\theta_1| |\theta_2|^{\alpha-1} \text{ para } 1 < \alpha \leq 2. \quad (\text{A.3})$$

Para mostrar a consistência do estimador da codiferença, precisamos dos dois lemas a seguir.

Lema A.1. *Seja o processo linear $X_t = \sum_{j=0}^{\infty} c_j a_{t-j}$, onde os coeficientes c_j 's são reais e satisfazem $|c_j| < Q^{-j}$ para algum $c > 0$, $Q > 1$ e a_t é i.i.d. com distribuição α -estável simétrica. Considere ainda a função característica de primeira ordem $\Phi(s) = E \exp(isX_t)$. Para $k \in \{0, 1, 2, \dots\}$ e $s \in \mathbb{R}$, $s \neq 0$*

$$\ln \phi(s, k) = \ln \left((N - k)^{-1} \sum_{t=1}^{N-k} \exp(isX_t) \right)$$

é um estimador consistente de $\ln \Phi(s)$.

Dem. *Seja $y_t = \exp(isX_t)$. Aparentemente, a magnitude de y_t é igual a um, portanto este é um processo estacionário de segunda ordem. Para simplificar a notação,*

ao invés de trabalhar com $\phi(s,k)$, primeiro mostramos a consistência de $\phi^*(s,k) = N^{-1} \sum_{t=1}^N \exp(isX_t)$. Temos que, $\phi^*(s,k)$ é um estimador não-viesado para $\Phi(s) = E(y_t)$. Para mostrar a fraca consistência deste estimador, mostramos que y_t é um processo ergótico em média. Uma condição suficiente para y_t ser ergótico em média, ou seja, $\phi^*(s,k) \rightarrow E(y_t)$ em média quadrática, é que sua função de covariância tende a zero quando t tende a infinito (exemplo, Teorema 7.1.1 em Brockwell e Davis 1987). A função covariância de y_t no lag k pode ser expressa por

$$\begin{aligned} c(k) &= |\Phi(s)|^2 \left(\frac{E(\exp(is(X_{t+k} - X_t)))}{E(\exp(isX_{t+k}))E(\exp(-isX_t))} - 1 \right) \\ &= |\Phi(s)|^2 (\exp(-\tau(k)) - 1). \end{aligned} \quad (\text{A.4})$$

Usando o Teorema A.1, percebe-se que $c(k) \rightarrow 0$ exponencialmente quando $k \rightarrow \infty$. Assim como a convergência em média quadrática implica em convergência em probabilidade, $\phi^*(s,k) \xrightarrow{p} \Phi(s)$. Além disso, quando os coeficientes c_j 's são reais e satisfazem $|c_j| < Q^{-j}$ para algum $c > 0$, $Q > 1$ e a_t é i.i.d. com distribuição α -estável simétrica, temos que $\Phi(s) = \exp(-\sum_{j=0}^{\infty} \sigma^\alpha |sc_j|^\alpha)$ é uma função que assume valores reais. Portanto podemos concluir que $\Re\phi^*(s,k) \xrightarrow{p} \Re\Phi(s)$ e $\Im\phi^*(s,k) \xrightarrow{p} \Im\Phi(s) = 0$.

Tomando o valor principal da função $\ln(\cdot)$ no domínio complexo, podemos ver que é uma função contínua e bem definida no conjunto \mathbb{C} menos a linha real negativa. Como $|c_j| < cQ^{-j}$ para algum $c > 0$, $Q > 1$, concluímos que $\Re\Phi(s)$ é sempre maior que zero que implica, com probabilidade convergindo para 0, que $\Re\phi^*(s,k)$ será menor ou igual a 0. Portanto, sem perda de generalidade, podemos restringir a definição das partes reais e imaginárias de $\ln\phi^*(s,k)$ apenas ao lado direito do plano onde $\Re\phi^*(s,k) > 0$, e igual a zero caso contrário. A partir desta consideração, obtemos que $\Re \ln \phi^*(s,k) = 1/2 \ln((\Re\phi^*(s,k))^2 + (\Im\phi^*(s,k))^2)$ e $\Im \ln \phi^*(s,k) = \arctan\left(\frac{\Im\phi^*(s,k)}{\Re\phi^*(s,k)}\right)$. A partir da continuidade da função logarítmica no domínio considerado, podemos deduzir $\Re \ln \phi^*(s,k) \xrightarrow{p} \Re \ln \Phi(s) = \ln \Phi(s)$ e $\Im \ln \phi^*(s,k) = \arg \phi^*(s,k) \xrightarrow{p} 0$ quando $N \rightarrow \infty$. Em outras palavras, obtemos $\ln \phi^*(s,k) \xrightarrow{p} \ln \Phi(s)$. Para completar a demonstração, é suficiente mostrar $\phi^*(s,k) - \phi(s,k) \rightarrow p0$. Por suposição do modelo, $\Re\Phi(s) > 0$, então $E|\Re\phi^*(s,k) - \Re\phi(s,k)| < 2\frac{k}{N-k}$ e $E|\Im\phi^*(s,k) - \Im\phi(s,k)| < 2\frac{k}{N-k}$, e portanto podemos concluir que $\phi^*(s,k) - \phi(s,k) = o_p(1)$. ■

Lema A.2. Seja o processo linear $X_t = \sum_{j=0}^{\infty} c_j a_{t-j}$, onde os coeficientes c_j 's são reais e satisfazem $|c_j| < Q^{-j}$ para algum $c > 0$, $Q > 1$ e a_t é i.i.d. com distribuição α -estável simétrica. E para $k \in \{0,1,2,\dots\}$ e $s \in \mathbb{R}$, $s \neq 0$, seja $\Phi(s, -s; k) = E \exp(is(X_{t+k} - x_t))$ sua função característica de segunda ordem calculada no ponto $(s, -s)$. Quando $N \rightarrow \infty$

$$\ln \phi(s, -s; k) \xrightarrow{p} \ln \Phi(s, -s; k), \quad (\text{A.5})$$

onde $\phi(s, -s; k)$ é dado em (3.6).

Dem. Para esta demonstração podemos proceder de forma semelhante ao lema anterior. Para simplificar, ao invés de trabalhar com $\phi(s, -s; k)$, primeiro vamos mostrar a consistência de $\phi^*(s, -s; k) = N^{-1} \sum_{t=1}^N \exp(is((X_{t+k} - X_t))$. Uma condição suficiente para que y_t seja a autocovariância ergótica (Proakis e Manolakis 1996, p.A10), ou seja, $\phi^*(s, -s; k) \rightarrow \Phi(s, -s; k)$ em média quadrática é que

$$E \exp(is(X_t - X_{t+k} - X_{t+n} + X_{t+n+k})) \xrightarrow{p} |\Phi(s, -s; k)|^2$$

a medida em que $n \rightarrow \infty$, onde o índice n denota o lag da covariância entre a função de autocovariância amostral. Consequentemente, temos que

$$\begin{aligned} & E \exp(is(X_t - X_{t+k} - X_{t+n} + X_{t+n+k})) \\ &= |\Phi(s, -s; k)|^2 \frac{E \exp(is((X_t - X_{t+k}) - (X_{t+n} - X_{t+n+k})))}{E \exp(is(X_t - X_{t+k})) E \exp(is(X_{t+n+k} - X_{t+n}))} \\ &= |\Phi(s, -s; k)|^2 \exp(-C_n). \end{aligned}$$

Onde

$$\begin{aligned} C_n &= -\ln E \exp(is((X_t - X_{t+k}) - (X_{t+n} - X_{t+n+k}))) \\ &\quad + \ln E \exp(is(X_t - X_{t+k})) + \ln E \exp(is(X_{t+n+k} - X_{t+n})). \end{aligned} \quad (\text{A.6})$$

Aplicando a técnica similar a usada em (3.9), podemos escrever C_n como

$$\begin{aligned} C_n &= \sigma^\alpha \left[\sum_{j=0}^{\infty} |s(c_j - c_{j+k} - c_{j+n+k})|^\alpha - |s(c_{j+n+k} - c_{j+n})|^\alpha - |s(c_j - c_{j+k})|^\alpha \right] \\ &= \sigma^\alpha \left[\sum_{j=0}^{\infty} |s(k_j - k_{j+k})|^\alpha - |s(k_{j+n})|^\alpha - |s(k_j)|^\alpha \right], \end{aligned}$$

onde $k_j = c_j - c_{j+k}$. Esta expressão é a função codiferença $\tau_G(n)$ para os coeficientes k_j 's e os parâmetros $\theta_1 = -s$, $\theta_2 = s$. Como $|c_j| < cQ^{-j}$ para algum $c > 0$, $Q > 1$, então $|k_j| < c_1Q^{-j}$ para algum $c_1 = 2c > 0$, $Q > 1$. Portanto, por (A.2) e (A.3), podemos concluir que $\exp(-C_n)$ vai convergir para 1 exponencialmente. Ou seja, $E \exp(is(X_t - X_{t+k} - X_{t+n} + X_{t+n+k})) \xrightarrow{p} |\Phi(s, -s; k)|^2$ para $n \rightarrow \infty$, e obtemos a convergência em média quadrática de $\phi^*(s, -s; k)$ para $\Phi(s, -s; k)$ e portanto $\phi^*(s, -s; k) \xrightarrow{p} \Phi(s, -s; k)$. No restante da demonstração, podemos proceder da mesma forma da demonstração do lema anterior e obtemos $\Phi(s, -s; k) = \exp(-\sigma^\alpha(\sum_{j=0}^{k-1} |sc_j|^\alpha - \sum_{j=0}^{\infty} |s(c_{j+k} - c_j)|^\alpha))$ é também uma função real e estritamente maior que 0. ■

Dem. (Teorema A.1) Para k finito e $N \rightarrow \infty$ obtemos $\sqrt{1 - k/N} \rightarrow 1$, então usando os resultados dos Lemas A e A.2, temos que quando $N \rightarrow \infty$, para $i = 1, \dots, r$

$$\hat{\tau}(s_i, -s_i; k) \rightarrow p - \ln \Phi(s_i, -s_i; k) + \ln \Phi(s_i) + \ln \Phi(-s_i) = \tau(s_i, -s_i; k). \quad (\text{A.7})$$

■

APÊNDICE B – A distribuição limite da função codiferença amostral

Neste anexo será derivada a distribuição assintótica da função codiferença amostral de um processo linear. A demonstração é dada por uma série de proposições, onde os principais resultados são apresentados no Teorema B.1 e na demonstração do Teorema 3.3. A demonstração será conduzida de forma similar a aproximação para a obtenção da função limitante da distribuição para FAC amostral no caso clássico, o exemplo pode ser consultado em Brockwell e Davis (1987), Teorema 7.2.1.

Para simplificar a notação, ao invés de trabalhar com $\hat{\tau}(s_i, -s_i; k)$, $i = 1, \dots, r$, vamos considerar um estimador similar $\hat{\tau}^*(s_i, -s_i; k)$,

$$\hat{\tau}^*(s_i, -s_i; k) = \ln \phi^*(s_i, -s_i; k) + \ln \phi^*(s_i, 0; k) + \ln \phi^*(0, -s_i; k), \quad (\text{B.1})$$

onde $\phi^*(u, v; k) = N^{-1} \sum_{t=1}^N \exp(i(uX_{t+k} + vX_t))$, $u, v \in \mathbb{R}$. O resultado necessário será apresentado no Teorema B.1.

Proposição B.1. *Seja $\{X_t, t \in \mathbb{Z}\}$ um processo estacionário linear (3.8) satisfazendo as condições (C1) e (C2). Então se $p \geq 0$ e $q \geq 0$,*

$$\lim_{N \rightarrow \infty} N \operatorname{cov} \left(\begin{pmatrix} \Re \hat{\tau}^*(\mathbf{s}, p) \\ \Im \hat{\tau}^*(\mathbf{s}, p) \end{pmatrix}, \begin{pmatrix} \Re \hat{\tau}^*(\mathbf{s}, q) \\ \Im \hat{\tau}^*(\mathbf{s}, q) \end{pmatrix} \right) = \lambda \mathbf{L}_2^p \mathbf{V}_{pq} \mathbf{L}_2^q \lambda^T,$$

onde as matrizes $\lambda, \mathbf{L}_2^p, k = p, q$ e \mathbf{V}_{pq} são dadas nas equações (B.3), (B.10) e (B.12) logo abaixo. Onde $\operatorname{cov}(X, Y)$ denota a covariância entre X e Y .

Dem. Para chegar a estrutura completa da matriz de variância e covariância do estimador, consideramos a seguinte representação para $\hat{\tau}^*(\mathbf{s}, p)$

$$\begin{pmatrix} \Re \hat{\tau}^*(\mathbf{s}, p) \\ \Im \hat{\tau}^*(\mathbf{s}, p) \end{pmatrix} = \begin{pmatrix} \Re \hat{\tau}^*(s_1, -s_1, k) \\ \Re \hat{\tau}^*(s_2, -s_2, k) \\ \vdots \\ \Re \hat{\tau}^*(s_r, -s_r, k) \\ \Im \hat{\tau}^*(s_1, -s_1, k) \\ \Im \hat{\tau}^*(s_2, -s_2, k) \\ \vdots \\ \Im \hat{\tau}^*(s_r, -s_r, k) \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{Y} \\ \mathbf{X} \end{pmatrix}, \quad (\text{B.2})$$

onde

$$\lambda = \begin{pmatrix} \mathbf{I}_r \otimes \lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_r \otimes \lambda_1 \end{pmatrix} \quad (\text{B.3})$$

e

$$\lambda_1 = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} \Re \ln Y_1^k \\ \Re \ln Y_2^k \\ \vdots \\ \Re \ln Y_r^k \end{pmatrix}, \mathbf{X} = \begin{pmatrix} \Im \ln Y_1^k \\ \Im \ln Y_2^k \\ \vdots \\ \Im \ln Y_r^k \end{pmatrix}.$$

Aqui \mathbf{I}_r denota a matriz identidade de tamanho r , onde definimos

$$Y_i^k = \begin{pmatrix} \phi^*(0, -s_i; k) \\ \phi^*(s_i, 0; k) \\ \phi^*(s_i, -s_i; k) \end{pmatrix} = \begin{pmatrix} \phi_1(s_i, k) \\ \phi_2(s_i, k) \\ \phi_3(s_i, k) \end{pmatrix}$$

e aplicando a função logarítmica e separando os componentes temos,

$$\Re \ln Y_i^k = \begin{pmatrix} \Re \ln \phi_1(s_i, k) \\ \Re \ln \phi_2(s_i, k) \\ \Re \ln \phi_3(s_i, k) \end{pmatrix}$$

e a parte imaginária é obtida de forma similar. Vamos denotar

$$EY_i^k = \begin{pmatrix} E\phi_1(s_i, k) \\ E\phi_2(s_i, k) \\ E\phi_3(s_i, k) \end{pmatrix} = \begin{pmatrix} \Phi_1(s_i, k) \\ \Phi_2(s_i, k) \\ \Phi_3(s_i, k) \end{pmatrix}.$$

Note que $\Phi(u, v; k) = E(\exp(i(uX_{t+k} + vX_t)))$, $u, v \in \mathbb{R}$. Usando o Teorema do valor médio, podemos expandir a função codiferença em

$$\begin{pmatrix} \Re \hat{\tau}^*(\mathbf{s}, k) \\ \Im \hat{\tau}^*(\mathbf{s}, k) \end{pmatrix} = \lambda \{ \mathbf{L}_1^k + \bar{\mathbf{L}}_2^k \mathbf{Z}_N^k \}, \quad (\text{B.4})$$

onde

$$\mathbf{L}_1^k = \begin{pmatrix} \Re \mathbf{L}_1^k \\ \Im \mathbf{L}_1^k \end{pmatrix}, \mathbf{Z}_N^k = \begin{pmatrix} \Re \mathbf{Z}_N^k \\ \Im \mathbf{Z}_N^k \end{pmatrix} = \begin{pmatrix} \Re \varphi_N^k - \Re \psi_N^k \\ \Im \varphi_N^k - \Im \psi_N^k \end{pmatrix}$$

com

$$\Re \mathbf{L}_1^k = \begin{pmatrix} \Re \ln EY_1^k \\ \Re \ln EY_2^k \\ \vdots \\ \Re \ln EY_r^k \end{pmatrix}, \Re \varphi_N^k = \begin{pmatrix} \Re Y_1^k \\ \Re Y_2^k \\ \vdots \\ \Re Y_r^k \end{pmatrix}, \Re \psi_r^k = \begin{pmatrix} \Re EY_1^k \\ \Re EY_2^k \\ \vdots \\ \Re EY_r^k \end{pmatrix}$$

e similarmente para as partes imaginárias, e $\bar{\mathbf{L}}_2^k = (\bar{d}_2^{ij})_{i,j=1,\dots,6}$ denota o Jacobiano de (B.2), que é calculado em $\mathbf{c}(\|\mathbf{c} - \psi_N^k\| < \|\varphi_N^k - \psi_N^k\|)$. A partir na suposição C2,

temos que

$$\Phi_3(s_i, k) = \Phi(s_i, -s_i; k) = \exp\left(-\sum_{j=0}^{k-1} \sigma^\alpha |s_i c_j|^\alpha - \sum_{j=0}^{\infty} \sigma^\alpha |s_i (c_{j+k} - c_j)|^\alpha\right) \quad (\text{B.5})$$

e $\Phi_1(s_i, k) = \Phi_2(s_i, k)$, ou seja,

$$\Phi(s_i, 0; k) = \Phi(0, -s_i; k) = \exp\left(-\sum_{j=0}^{\infty} \sigma^\alpha |s_i c_j|^\alpha\right). \quad (\text{B.6})$$

A partir das identidade (B.5) e (B.6) e posteriormente aplicando a suposição C1, temos que os elementos de $\Re\psi_N^k$ são sempre estritamente maiores ou iguais a 0. Portanto, com a probabilidade de convergência para 0, os elementos de $\Re\varphi_N^k$ será menor ou igual a 0. Conseqüentemente, sem mudar a distribuição limitante do estimador, podemos restringir a definição dos componente reais e imaginários de $\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$ em (B.2) apenas na metade direita do plano onde os elementos de $\Re(\varphi)^k_N > 0$, e igual a 0 caso contrário. Então podemos concluir que a matriz Jacobiana $\bar{\mathbf{L}}_2^k$ é bem definida. A partir do Teorema 3.2, $\bar{\mathbf{L}}_2^k$ vai convergir em probabilidade para \mathbf{L}_2^k , onde

$$\mathbf{L}_2^k = \nabla \mathbf{L}_1^k.$$

Onde ∇g denota o Jacobiano de g . A partir de (B.5) e (B.6), temos as seguintes identidades

$$\Re\Phi(s_i, -s_i; k) = E \cos(s_i(X_{t+k} - X_t)) = \Phi(s_i, -s_i; k), \quad (\text{B.7})$$

$$\Re\Phi(s_i, 0; k) = E \cos(s_i X_{t+k}) = \Phi(s_i, 0; k), \quad (\text{B.8})$$

$$\Re\Phi(0, -s_i; k) = E \cos(-s_i X_t) = \Phi(0, -s_i; k) \quad (\text{B.9})$$

e $\Im\Phi(s_i, -s_i; k) = E \text{sen}(s_i(X_{t+k} - X_t)) = 0$, $\Im\Phi(s_i, 0; k) = E \text{sen}(s_i X_{t+k}) = 0$ e $\Im\Phi(0, -s_i; k) = E \text{sen}(s_i X_t) = 0$. Usando estas identidades, após algumas manipulações algébricas, obtemos diretamente

$$\mathbf{L}_2^k = \begin{pmatrix} \mathbf{I}_r \mathbf{d}^k & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_r \mathbf{d}^k \end{pmatrix}, \quad (\text{B.10})$$

onde $(\mathbf{d}^k)^T = [\mathbf{d}^{k1}, \mathbf{d}_2^k, \dots, \mathbf{d}_r^k]$, e o os elementos \mathbf{d}_i^k , $i = 1, \dots, r$ são

$$\mathbf{d}_i^k(1,1) = (\Re\Phi(0, -s_i; k))^{-1},$$

$$\mathbf{d}_i^k(2,2) = (\Re\Phi(s_i, 0; k))^{-1},$$

$$\mathbf{d}_i^k(3,3) = (\Re\Phi(s_i, -s_i; k))^{-1}$$

e igual a 0 caso contrário. A matriz de variância e covariância assintótica obtida a partir de (B.4) é dada por

$$\lim_{N \rightarrow \infty} N \operatorname{cov} \left(\begin{pmatrix} \Re \hat{\tau}^*(s, -s; p) \\ \Im \hat{\tau}^*(s, -s; p) \end{pmatrix}, \begin{pmatrix} \Re \hat{\tau}^*(s, -s; q) \\ \Im \hat{\tau}^*(s, -s; q) \end{pmatrix} \right) = \lambda \mathbf{L}_2^p \mathbf{V}_{pq} \mathbf{L}_2^q \lambda^T, \quad (\text{B.11})$$

onde

$$\mathbf{V}_{pq} = \begin{pmatrix} \mathbf{V}_{pq}^{RR} & \mathbf{V}_{pq}^{RI} \\ \mathbf{V}_{pq}^{IR} & \mathbf{V}_{pq}^{II} \end{pmatrix} = \lim_{N \rightarrow \infty} N \begin{pmatrix} \operatorname{cov}(\Re \mathbf{Z}_N^p, \Re \mathbf{Z}_N^q) & \operatorname{cov}(\Re \mathbf{Z}_N^p, \Im \mathbf{Z}_N^q) \\ \operatorname{cov}(\Im \mathbf{Z}_N^p, \Re \mathbf{Z}_N^q) & \operatorname{cov}(\Im \mathbf{Z}_N^p, \Im \mathbf{Z}_N^q) \end{pmatrix}. \quad (\text{B.12})$$

A matriz \mathbf{V}_{pq} pode ser obtida aplicando do Teorema 3.2 e Observação 2.6 em Hesse (1990). Seus elementos podem ser derivados da mesma forma que a matriz de variância e covariância obtida no Teorema 1 de Hesse (1990). Isto é possível porque pode ser mostrado que todos os elementos de \mathbf{V}_{pq} (na forma da soma absoluta dos elementos) é finita. Consequentemente, pode-se aplicar a propriedade da média amostral de um processo ergótico (por exemplo, Teorema 7.1.1 em Brockwell e Davis, 1987). Note que neste caso em particular, obtemos que todos os elementos de \mathbf{V}_{pq} em relação a $\operatorname{cov}(\Re \mathbf{Z}_N^p, \Im \mathbf{Z}_N^q)$ e $\operatorname{cov}(\Im \mathbf{Z}_N^p, \Re \mathbf{Z}_N^q)$ são iguais a zero. Os elementos em relação a $\operatorname{cov}(\Re \mathbf{Z}_N^p, \Re \mathbf{Z}_N^q)$ e $\operatorname{cov}(\Im \mathbf{Z}_N^p, \Im \mathbf{Z}_N^q)$ pode ser mostrado que é finito usando as identidades (B.5) e (B.6) e aplicando uma aproximação similar a obtida nas equações (A.4) e (A.6), e ainda a aplicação do Teorema A.1, ou em alguns casos, a equação 2.7 em Kokoszka e Taquq (1994) como os passos similares aos da demonstração do Teorema A.1. Contudo, estes detalhes foram omitidos. ■

Proposição B.2. *Seja $X_t, t \in \mathbb{Z}$ um processo de médias móveis de ordem m , $X_t = \sum_{j=0}^m c_j a_{t-j}$, e satisfaz as condições (C1) e (C2). Então para $h \in 1, 2, \dots, s \in \mathbb{R}, s \neq 0$*

$$\left[\begin{pmatrix} \Re \hat{\tau}^*(\mathbf{s}, 0) \\ \Im \hat{\tau}^*(\mathbf{s}, 0) \end{pmatrix}, \dots, \begin{pmatrix} \Re \hat{\tau}^*(\mathbf{s}, h) \\ \Im \hat{\tau}^*(\mathbf{s}, h) \end{pmatrix} \right]^T$$

é

$$AN \left(\left[\begin{pmatrix} \hat{\tau}(\mathbf{s}, 0) \\ \mathbf{0} \end{pmatrix}, \dots, \begin{pmatrix} \hat{\tau}(\mathbf{s}, 0) \\ \mathbf{0} \end{pmatrix} \right]^T, N^{-1} \mathbf{M} \right).$$

Onde AN é a notação para indicar que a variável segue distribuição normal assintoticamente. E \mathbf{M} é a matriz de covariância.

$$\mathbf{M} = \left[\lambda \mathbf{L}_2^p \mathbf{V}_{pq} \mathbf{L}_2^q \lambda^T \right]_{p,q=0,\dots,h}$$

e as matrizes $\lambda, \mathbf{L}_2^p, k = p, q$ e \mathbf{V}_{pq} foram apresentadas na Proposição B.1.

Dem. Para mostrar esta relação, definimos o vetor \mathbf{Y}_t^T

$$\mathbf{Y}_t^T = (\mathbf{Z}_t, \mathbf{Z}_{t+1}, \dots, \mathbf{Z}_{t+h}),$$

onde

$$\mathbf{Z}_{t+k} = \begin{pmatrix} \mathbf{X}_1^k \\ \mathbf{X}_2^k \\ \vdots \\ \mathbf{X}_r^k \end{pmatrix}$$

e para $j = 1, \dots, r$

$$\mathbf{X}_j^k = \begin{pmatrix} \exp(-is_j X_t) \\ \exp(is_j X_{t+k}) \\ \exp(is_j (X_{t+k} - X_t)) \end{pmatrix}.$$

Por definição, \mathbf{Z}_{t+k} é uma sequência $m+k$ dependente e portanto \mathbf{Y}_t é uma sequência $m+h$ dependente. Definimos a seguir

$$\zeta_{t+j} = (\xi_t, \xi_{t+1}, \dots, \xi_{t+h}),$$

onde

$$\xi_{t+j} = \begin{pmatrix} \Re(\ln N^{-1} \sum_{t=1}^N Z_{t+j}) \\ \Im(\ln N^{-1} \sum_{t=1}^N Z_{t+j}) \end{pmatrix}$$

e

$$\Re(\ln N^{-1} \sum_{t=1}^N Z_{t+j}) = \begin{pmatrix} \Re(\ln N^{-1} \sum_{t=1}^N X_1^j) \\ \Re(\ln N^{-1} \sum_{t=1}^N X_2^j) \\ \vdots \\ \Re(\ln N^{-1} \sum_{t=1}^N X_r^j) \end{pmatrix}.$$

E para $l = 1, \dots, r$

$$\Re(\ln N^{-1} \sum_{t=1}^N X_t^j) = \begin{pmatrix} \Re(\ln N^{-1} \sum_{t=1}^N \exp(-is_l X_t)) \\ \Re(\ln N^{-1} \sum_{t=1}^N \exp(is_l X_{t+j})) \\ \Re(\ln N^{-1} \sum_{t=1}^N \exp(-is_l (X_{t+j} - X_t))) \end{pmatrix}$$

(similarmente para parte imaginária. Note que o somatório e o valor principal de $\ln(\cdot)$ são os mesmo para as partes reais e imaginária), então temos

$$\begin{pmatrix} \Re(\ln N^{-1} \sum_{t=1}^N Y_t) \\ \Im(\ln N^{-1} \sum_{t=1}^N Y_t) \end{pmatrix} \lambda^T = \zeta_t \lambda^T = \left[\begin{pmatrix} \Re \hat{\tau}^*(\mathbf{s}, 0) \\ \Im \hat{\tau}^*(\mathbf{s}, 0) \end{pmatrix}, \dots, \begin{pmatrix} \Re \hat{\tau}^*(\mathbf{s}, h) \\ \Im \hat{\tau}^*(\mathbf{s}, h) \end{pmatrix} \right]^T,$$

onde λ é dado em (B.3). Portanto, temos que mostrar que quando $N \rightarrow \infty$

$$\mathbf{a}^T (\zeta_t \lambda^T)^T \sim AN \left(\mathbf{a}^T \left[\begin{pmatrix} \Re \tau(\mathbf{s}, 0) \\ \mathbf{0} \end{pmatrix}, \dots, \begin{pmatrix} \Re \tau(\mathbf{s}, h) \\ \mathbf{0} \end{pmatrix} \right]^T, N^{-1} \mathbf{a}^T \mathbf{M} \mathbf{a} \right) \quad (\text{B.13})$$

para todos os vetores $\mathbf{a} = (a_0, \dots, a_h)^T \in \mathbb{R}^{h+1}$ tais que $\mathbf{a}^T \mathbf{M} \mathbf{a} > 0$. Para qualquer \mathbf{a} , a sequência $\mathbf{a}^T (\zeta_t \lambda^T)^T$ é $(m+h)$ dependente e de acordo com a Proposição B.1

$$\lim_{N \rightarrow \infty} N \operatorname{var}(\mathbf{a}^T (\zeta_t \lambda^T)^T) = \mathbf{a}^T \mathbf{M} \mathbf{a} > 0,$$

onde \mathbf{M} é a matriz de covariância

$$\mathbf{M} = \left[\lambda \mathbf{L}_2^p \mathbf{V}_{pq} \mathbf{L}_2^q \lambda^T \right]_{p,q=0,\dots,h}$$

e os vetores λ , \mathbf{L}_2^p , \mathbf{L}_2^q , e a matriz \mathbf{V}_{pq} são das na Proposição B.1, apresentada acima. Podemos concluir que $\mathbf{a}^T (\zeta_t \lambda^T)^T$ satisfaz as condições do Teorema do Limite Central para um processo m dependente (Teorema 6.4.2 Brockwell e Davis, 1987), então por este Teorema, para $N \rightarrow \infty$, chegamos ao resultado em (B.13). A relação $\mathfrak{S}\tau(\mathbf{s}, j) = 0, j = 0, 1, \dots, h$ podem ser obtidas diretamente a partir das identidade (B.5) e (B.6).

■

Proposição B.3. Se a Proposição B.2 é verdadeira para X_t , $t \in \mathbb{Z}$ sendo um processo estacionário (3.8) satisfazendo as condições (C1) e (C2).

Dem. Para provar a Proposição B.3, vamos aplicar o resultado da Proposição B.2 em uma sequência truncada $X_{tm} = \sum_{j=0}^m c_j a_{t-j}$ e depois derivar o resultado para X_t fazendo $m \rightarrow \infty$. Para $0 \leq p \leq h$, definimos

$$\hat{\tau}_m^*(s, -s; p) = -\ln \phi_m^*(s, -s; p) + \ln \phi_m^*(s, 0; p) + \phi_m^*(0, -s; p), \quad (\text{B.14})$$

onde $\phi_m^*(u, v; p) = N^{-1} \sum_{t=1}^N \exp(i(uX_{(t+p)m} + vX_{tm}))$. Então, pela Proposição B.2

$$N^{-1/2} \left[\left(\Re \hat{\tau}_m^*(\mathbf{s}, 0) - \Re \tau_m(\mathbf{s}, 0) \right), \dots, \left(\Re \hat{\tau}_m^*(\mathbf{s}, h) - \Re \tau_m(\mathbf{s}, h) \right) \right]^T \Rightarrow Y_m,$$

onde $Y_m \sim N(0, M_m)$. Onde M_m é a matriz de covariância

$$\begin{aligned} \mathbf{M}_m &= \begin{pmatrix} \operatorname{cov}(\Re \hat{\tau}_m^*(\mathbf{s}, p), \Re \hat{\tau}_m^*(\mathbf{s}, q)) & \operatorname{cov}(\Re \hat{\tau}_m^*(\mathbf{s}, p), \Im \hat{\tau}_m^*(\mathbf{s}, q)) \\ \operatorname{cov}(\Im \hat{\tau}_m^*(\mathbf{s}, p), \Re \hat{\tau}_m^*(\mathbf{s}, q)) & \operatorname{cov}(\Im \hat{\tau}_m^*(\mathbf{s}, p), \Im \hat{\tau}_m^*(\mathbf{s}, q)) \end{pmatrix}_{p,q=0,\dots,h} \\ &= \left[\lambda \mathbf{L}_2^p(m) \mathbf{V}_{pq}^m \mathbf{L}_2^q(m) \lambda^T \right]_{p,q=0,\dots,h}, \end{aligned}$$

onde λ é definido em (B.3) e a matriz Jacobiana $\mathbf{L}_2^q(m)$ e a matriz \mathbf{V}_{pq}^m são definidas para X_{tm} nas equações (B.10) e (B.12), respectivamente. Quando $m \rightarrow \infty$, $\mathbf{M}_m \rightarrow \mathbf{M}$, onde \mathbf{M} é definida como M_m , basta substituir X_{tm} por X_t . Então

$$Y_m \Rightarrow \text{onde } Y \sim N(0, \mathbf{M})$$

a demonstração está completa aplicando a Proposição 6.3.9 em Brockwell e Davis (1987) desde que possamos mostrar que

$$\lim_{m \rightarrow \infty} \limsup_{N \rightarrow \infty} P(N^{1/2} |\Re \hat{\tau}_m^*(\mathbf{s}, p) - \Re \tau_m(\mathbf{s}, p) - \Re \hat{\tau}^*(\mathbf{s}, p) + \Re \hat{\tau}(\mathbf{s}, p)| > \epsilon) = 0 \quad (\text{B.15})$$

para $p = 0, 1, \dots, h$ (repetindo os mesmo passos para a parte imaginária). A probabilidade em (B.15) é limitada por

$$\begin{aligned} \epsilon^{-2} N \operatorname{var}(\Re \hat{\tau}_m^*(\mathbf{s}, p) - \Re \hat{\tau}^*(\mathbf{s}, p)) &= \epsilon^{-2} [N \operatorname{var}(\Re \hat{\tau}_m^*(\mathbf{s}, p)) + N \operatorname{var}(\Re \hat{\tau}^*(\mathbf{s}, p)) \\ &\quad - 2N \operatorname{cov}(\Re \hat{\tau}_m^*(\mathbf{s}, p), \Re \hat{\tau}^*(\mathbf{s}, p))]. \end{aligned}$$

A partir dos cálculos da Proposição B.1 e após notar que Teorema 3.2 e a Observação 2.6 em Hesse (1990) pode ser aplicada para o processo de média móvel finito ajustando alguns dos coeficientes c_j 's para ser igual a zero, obtemos

$$\lim_{m \rightarrow \infty} \limsup_{N \rightarrow \infty} N \operatorname{var}(\Re \hat{\tau}_m^*(\mathbf{s}, p)) = \lim_{N \rightarrow \infty} \operatorname{var}(\Re \hat{\tau}^*(\mathbf{s}, p)) = m_{pp}^{RR},$$

onde m_{pp}^{RR} é a matriz de covariância entre os elementos reais no bloco (p, q) da matriz de covariância \mathbf{M} . Além disso, usando os mesmos passos para que, dada na prova da Proposição B.1, pode ser mostrado que

$$\lim_{m \rightarrow \infty} \limsup_{N \rightarrow \infty} N \operatorname{cov}(\Re \hat{\tau}_m^*(\mathbf{s}, p), \Re \hat{\tau}^*(\mathbf{s}, p)) = m_{pp}^{RR},$$

então,

$$\lim_{m \rightarrow \infty} \limsup_{N \rightarrow \infty} \epsilon^{-2} N \operatorname{var}(\Re \hat{\tau}_m^*(\mathbf{s}, p) - \Re \hat{\tau}^*(\mathbf{s}, p)) = 0.$$

Resultados semelhantes podem ser obtidos para a parte imaginária. Isto demonstra (B.15). ■

Teorema B.1. *Seja $\{X_t, t \in \mathbb{Z}\}$ um processo estacionário linear (3.8) satisfazendo as condições C1 e C2. Então para $h \in 1, 2, \dots, s \in \mathbb{R}, s \neq 0$*

$$\left[\begin{pmatrix} \Re \hat{\tau}(\mathbf{s}, 0) \\ \Im \hat{\tau}(\mathbf{s}, 0) \end{pmatrix}, \dots, \begin{pmatrix} \Re \hat{\tau}(\mathbf{s}, h) \\ \Im \hat{\tau}(\mathbf{s}, h) \end{pmatrix} \right]^T \text{ é } AN \left(\left[\begin{pmatrix} \tau(\mathbf{s}, 0) \\ \mathbf{0} \end{pmatrix}, \dots, \begin{pmatrix} \tau(\mathbf{s}, 0) \\ \mathbf{0} \end{pmatrix} \right]^T, N^{-1} \mathbf{M} \right),$$

onde \mathbf{M} é dada na Proposição B.2.

Dem. Para mostrar a convergência dos estimadores $\Re \hat{\tau}(\mathbf{s}, j)$ e $\Im \hat{\tau}(\mathbf{s}, j)$ para os respectivos limites $\Re \hat{\tau}^*(\mathbf{s}, j)$ e $\Im \hat{\tau}^*(\mathbf{s}, j)$, com $0 \leq j \leq h$, é suficiente mostrar que quando $N \rightarrow \infty$

$$N^{1/2} \lambda_2 \begin{pmatrix} \Re \phi^*(s_k, -s_k; j) \\ \Re \phi^*(s_k, 0; j) \\ \Re \phi^*(0, -s_k; j) \end{pmatrix} - \lambda_2 \begin{pmatrix} \Re \phi(s_k, -s_k; j) \\ \Re \phi(s_k, 0; j) \\ \Re \phi(0, -s_k; j) \end{pmatrix} = o_p(1)$$

(e similarmente para parte imaginária), onde $\phi^*(u, v; j) = N^{-1} \sum_{t=1}^N \exp(i(uX_{t+j} + vX_t))$, $\phi(u, v; j) = (N-j)^{-1} \sum_{t=1}^{N-j} \exp(i(uX_{t+j} + vX_t))$ e $\lambda_2 = \begin{bmatrix} -1 & 1 & 1 \end{bmatrix}$. O resultado necessário, em seguida, decorre do Teorema de Slutsky (por exemplo, o Teorema 5.1.1 em Lehmann 1999).

Após álgebra simples temos, para $0 \leq j \leq h$,

$$\begin{aligned}
& N^{1/2} E \left| \lambda_2 \begin{pmatrix} \Re \phi^*(s_k, -s_k; j) \\ \Re \phi^*(s_k, 0; j) \\ \Re \phi^*(0, -s_k; j) \end{pmatrix} - \lambda_2 \begin{pmatrix} \Re \phi(s_k, -s_k; j) \\ \Re \phi(s_k, 0; j) \\ \Re \phi(0, -s_k; j) \end{pmatrix} \right| \\
&= N^{1/2} E \left| \lambda_2 \begin{pmatrix} \left(\frac{j}{(N-j)} \frac{1}{N} \sum_{t=1}^N \cos(is_k(X_{t+j} - X_t)) - \frac{1}{(N-j)} \sum_{t=N-j+1}^N \cos(is(X_{t+j} - X_t)) \right) \\ \left(\frac{j}{(N-j)} \frac{1}{N} \sum_{t=1}^N \cos(is_k X_{t+j}) - \frac{1}{(N-j)} \sum_{t=N-j+1}^N \cos(is_k X_{t+j}) \right) \\ \left(\frac{j}{(N-j)} \frac{1}{N} \sum_{t=1}^N \cos(-is_k X_t) - \frac{1}{(N-j)} \sum_{t=N-j+1}^N \cos(-is X_t) \right) \end{pmatrix} \right| \\
&\leq 6j(N-j)^{-1/2} \left(\frac{N}{N-j} \right)^{1/2}.
\end{aligned}$$

O resultado desejado é obtido a partir de $3j(N-j)^{-1/2} \rightarrow 0$ e $N/(N-j) \rightarrow 1$ quando $N \rightarrow \infty$. Utilizando os mesmos argumentos, resultados similares podem ser obtidos para a parte imaginária. A conclusão do Teorema em seguida, segue da Proposição B.3. ■

Dem (Teorema 3.3). Seja $\mathbf{g}(\cdot)$ uma função no espaço $\mathbb{R}^{(h+1) \times 2r}$ para $\mathbb{R}^{h \times 2r}$ definida por

$$\begin{aligned}
& \mathbf{g} \left(\left[\begin{pmatrix} \Re \hat{\tau}(\mathbf{s}, 0) \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \Re \hat{\tau}(\mathbf{s}, 1) \\ \Im \hat{\tau}(\mathbf{s}, 1) \end{pmatrix}, \dots, \begin{pmatrix} \Re \hat{\tau}(\mathbf{s}, h) \\ \Im \hat{\tau}(\mathbf{s}, h) \end{pmatrix} \right]^T \right) \\
&= \left[\begin{pmatrix} \Re \hat{I}(\mathbf{s}, 1) \\ \Im \hat{I}(\mathbf{s}, 1) \end{pmatrix}, \begin{pmatrix} \Re \hat{I}(\mathbf{s}, 2) \\ \Im \hat{I}(\mathbf{s}, 2) \end{pmatrix}, \dots, \begin{pmatrix} \Re \hat{I}(\mathbf{s}, h) \\ \Im \hat{I}(\mathbf{s}, h) \end{pmatrix} \right]^T,
\end{aligned}$$

onde para $0 < j \leq h$ e $\hat{\tau}(0) \neq 0$, temos que $\Re \hat{I}(s_i, -s_i; j) = \frac{\Re \hat{\tau}(s_i, -s_i; j)}{\hat{\tau}(s_i, -s_i; 0)}$ e $\Im \hat{I}(s_i, -s_i; j) = \frac{\Im \hat{\tau}(s_i, -s_i; j)}{\hat{\tau}(s_i, -s_i; 0)}$, para $i = 1, \dots, r$. Ao aplicar o método delta e Teorema B.1 acima, podemos mostrar que

$$\left[\begin{pmatrix} \Re \hat{I}(\mathbf{s}, 1) \\ \Im \hat{I}(\mathbf{s}, 1) \end{pmatrix}, \begin{pmatrix} \Re \hat{I}(\mathbf{s}, 2) \\ \Im \hat{I}(\mathbf{s}, 2) \end{pmatrix}, \dots, \begin{pmatrix} \Re \hat{I}(\mathbf{s}, h) \\ \Im \hat{I}(\mathbf{s}, h) \end{pmatrix} \right]^T$$

tem distribuição normal assintoticamente com média

$$\begin{aligned}
& \mathbf{g} \left(\left[\begin{pmatrix} \Re \tau(\mathbf{s}, 0) \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \Re \tau(\mathbf{s}, 1) \\ \mathbf{0} \end{pmatrix}, \dots, \begin{pmatrix} \Re \tau(\mathbf{s}, h) \\ \mathbf{0} \end{pmatrix} \right]^T \right) \\
&= \left[\begin{pmatrix} l_r I(1) \\ l_r 0 \end{pmatrix}, \begin{pmatrix} l_r I(2) \\ l_r 0 \end{pmatrix}, \dots, \begin{pmatrix} l_r I(h) \\ l_r 0 \end{pmatrix} \right]^T
\end{aligned}$$

e variância $N^{-1} \mathbf{DMD}^T$. Aqui a matriz \mathbf{M} é como indicada na Proposição B.2, \mathbf{D} é a matriz Jacobiana de $\mathbf{g}(\cdot)$ e $l_r = [1, 1, \dots, 1]^T \in \mathbb{R}^r$. Para obter os elementos de matriz \mathbf{D} , seguimos os seguintes passos. Primeiro, observe que a função codiferença na lag 0 é uma função real. Portanto, para $0 \leq j \leq h$ e $\tau(0) \neq 0$, temos que $\Re I(j) = \frac{\Re \tau(j)}{\tau(0)} = I(j)$

e $\Im I(j) = \frac{\Re \tau(j)}{\tau(0)} = 0$. É simples obter a matriz Jacobiana \mathbf{D} como

$$\mathbf{D} = \begin{bmatrix} D_{11} & D_{12} & \mathbf{0} & \cdots & \mathbf{0} \\ D_{21} & \mathbf{0} & D_{23} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ D_{h1} & \mathbf{0} & \mathbf{0} & \cdots & D_{h(h+1)} \end{bmatrix}, \quad (\text{B.16})$$

onde

$$D_{l1} = \begin{bmatrix} D_{l1}^{11} & \mathbf{0}_r \\ \mathbf{0}_r & \mathbf{0}_r \end{bmatrix} \quad (\text{B.17})$$

e

$$D_{l(l+1)} = \begin{bmatrix} D_{l(l+1)}^{11} & \mathbf{0}_r \\ \mathbf{0}_r & D_{l(l+1)}^{11} \end{bmatrix} \quad (\text{B.18})$$

para $l = 1, \dots, h$, onde

$$D_{l1}^{11} = \mathbf{I}_r \left[\frac{-I(l)}{\tau(s_1, -s_1; 0)}, \frac{-I(l)}{\tau(s_2, -s_2; 0)}, \dots, \frac{-I(l)}{\tau(s_r, -s_r; 0)} \right]^T$$

e

$$D_{l(l+1)}^{11} = \mathbf{I}_r \left[\frac{1}{\tau(s_1, -s_1; 0)}, \frac{1}{\tau(s_2, -s_2; 0)}, \dots, \frac{1}{\tau(s_r, -s_r; 0)} \right]^T.$$

Aqui \mathbf{I}_r denota a matriz identidade de tamanho r . Vamos denotar w_{ij} , para $i, j = 0, 1, \dots, h$ como o bloco (i, j) da matriz \mathbf{M} . Temos que

$$\begin{aligned} w_{ij} &= \begin{bmatrix} \text{cov}(\Re \hat{I}(\mathbf{s}, i), \Re \hat{I}(\mathbf{s}, j)) & \text{cov}(\Re \hat{I}(\mathbf{s}, i), \Im \hat{I}(\mathbf{s}, j)) \\ \text{cov}(\Im \hat{I}(\mathbf{s}, i), \Re \hat{I}(\mathbf{s}, j)) & \text{cov}(\Im \hat{I}(\mathbf{s}, i), \Im \hat{I}(\mathbf{s}, j)) \end{bmatrix} \\ &= D_{i1} m_{00} D_{j1} + D_{i(i+1)} m_{i0} D_{j1} + D_{i1} m_{0j} D_{j(j+1)} + D_{i(i+1)} m_{ij} D_{j(j+1)} \\ &= \begin{bmatrix} D_{i1}^{11} m_{00}^{RR} D_{j1}^{11} + D_{i(i+1)}^{11} m_{i0}^{RR} D_{j1}^{11} + D_{i1}^{11} m_{0j}^{RR} D_{j(j+1)}^{11} + D_{i(i+1)}^{11} m_{ij}^{RR} D_{j(j+1)}^{11} & \mathbf{0}_r \\ \mathbf{0}_r & D_{i(i+1)}^{11} m_{ij}^{RR} D_{j(j+1)}^{11} \end{bmatrix}. \end{aligned} \quad (\text{B.19})$$

Aqui m_{ij}^{RR} e m_{ij}^{II} denotam as partições de m_{ij} correspondente as partes reais e imaginárias, respectivamente. ■

APÊNDICE C – Demonstração do Corolário 3.1

Como o $MA(0)$ é um caso especial do processo linear (3.8), aplicando o Teorema 3.3, pode-se concluir a normalidade assintótica de

$$\left[\begin{pmatrix} \Re \hat{I}(\mathbf{s}, 1) \\ \Im \hat{I}(\mathbf{s}, 1) \end{pmatrix}, \begin{pmatrix} \Re \hat{I}(\mathbf{s}, 2) \\ \Im \hat{I}(\mathbf{s}, 2) \end{pmatrix}, \dots, \begin{pmatrix} \Re \hat{I}(\mathbf{s}, h) \\ \Im \hat{I}(\mathbf{s}, h) \end{pmatrix} \right]^T$$

para $k \in 1, 2, \dots$. A função codiferença do processo i.i.d. X_t é

$$\begin{aligned} \tau(k) = \tau(s, -s; k) &= -\ln E \exp\{is(X_{t+k} - X_t)\} + \ln E \exp\{isX_{t+k}\} + \ln E \exp\{-isX_t\}, \\ &= \begin{cases} -2\sigma^\alpha |s|^\alpha & \text{para } k = 0, \\ 0 & \text{para } k > 0. \end{cases} \end{aligned}$$

que nos permite concluir que as partes reais e imaginárias de $I(k) = 0$ para qualquer $k > 0$.

A partir de (B.19), temos que $w_{kk}, k > 0$ reduz a

$$w_{kk} = D_{k(k+1)} m_{kk} D_{k(k+1)} \quad (\text{C.1})$$

onde a matriz $D_{k(k+1)}$ é dada em (B.18) com

$$D_{l(l+1)}^{11} = \mathbb{I}_r \left[\frac{1}{-2\sigma^\alpha |s_1|^\alpha}, \frac{1}{-2\sigma^\alpha |s_2|^\alpha}, \dots, \frac{1}{-2\sigma^\alpha |s_r|^\alpha} \right]^T$$

e onde

$$m_{kk} = \lambda \mathbf{L}_2^k \mathbf{V}_{kk} \mathbf{L}_2^k \lambda^T \quad (\text{C.2})$$

com λ dado em (B.3), e os elementos da matriz \mathbb{L}_2^k e da matriz de covariância \mathbf{V}_{kk} serão apresentados abaixo. Vamos denotar

$$\mathbf{V}_{kk}^{RR}(i, j) = [\text{cov}(\Re \phi_p(s_i, k), \Re \phi_q(s_j, k))]_{p, q=1, 2, 3}$$

e

$$\mathbf{V}_{kk}^{II}(i, j) = [\text{cov}(\Im \phi_p(s_i, k), \Im \phi_q(s_j, k))]_{p, q=1, 2, 3},$$

como o bloco (i, j) da matriz \mathbf{V}_{kk}^{RR} e \mathbf{V}_{kk}^{II} , respectivamente. Usando as identidades (B.7)-(B.9) (e depois as respectivas identidades para as partes imaginárias) obtemos os componentes a seguir

$$\begin{aligned}
\text{cov}(\Re(\phi_1(s_i, p)), \Re(\phi_1(s_j, q))) &= \text{cov}(\cos(-s_i X_t), \cos(-s_j X_t)) \\
&= \frac{1}{2} \{ e^{-\sigma^\alpha |s_i + s_j|^\alpha} + e^{-\sigma^\alpha |s_i - s_j|^\alpha} \} - e^{-\sigma^\alpha (|s_i|^\alpha + |s_j|^\alpha)}, \\
\text{cov}(\Re(\phi_1(s_i, p)), \Re(\phi_2(s_j, q))) &= \text{cov}(\Re(\phi_2(s_i, p)), \Re(\phi_1(s_j, q))) \\
&= \text{cov}(\Re(\phi_1(s_i, p)), \Re(\phi_1(s_j, q))), \\
\text{cov}(\Re(\phi_2(s_i, p)), \Re(\phi_2(s_j, q))) &= \text{cov}(\Re(\phi_1(s_i, p)), \Re(\phi_1(s_j, q))), \\
\text{cov}(\Re(\phi_1(s_i, p)), \Re(\phi_3(s_j, q))) &= \text{cov}(\cos(-s_i X_t), \cos(-s_j (X_{t+q} - X_t))) \\
&\quad + \text{cov}(\cos(-s_i X_{t+q}), \cos(-s_j (X_{t+q} - X_t))) \\
&= e^{-\sigma^\alpha (|s_j|^\alpha + |s_i - s_j|^\alpha)} + e^{-\sigma^\alpha (|s_j|^\alpha + |s_i + s_j|^\alpha)} \\
&\quad - 2e^{-\sigma^\alpha (|s_i|^\alpha + |2s_j|^\alpha)}, \\
\text{cov}(\Re(\phi_3(s_i, p)), \Re(\phi_1(s_j, q))) &= \text{cov}(\cos(-s_j X_t), \cos(s_i (X_{t+q} - X_t))) \\
&\quad + \text{cov}(\cos(-s_j X_{t+p}), \cos(s_i (X_{t+p} - X_t))) \\
&= e^{-\sigma^\alpha (|s_i|^\alpha + |s_i - s_j|^\alpha)} + e^{-\sigma^\alpha (|s_i|^\alpha + |s_i + s_j|^\alpha)} \\
&\quad - 2e^{-\sigma^\alpha (|s_j|^\alpha + |2s_i|^\alpha)}, \\
\text{cov}(\Re(\phi_2(s_i, p)), \Re(\phi_3(s_j, q))) &= \text{cov}(\cos(s_i X_{t+q}), \cos(s_j (X_{t+q} - X_t))) \\
&\quad + \text{cov}(\cos(s_i X_{t+p}), \cos(s_j (X_{t+p+q} - X_{t+p}))) \\
&= e^{-\sigma^\alpha (|s_j|^\alpha + |s_i - s_j|^\alpha)} + e^{-\sigma^\alpha (|s_j|^\alpha + |s_i + s_j|^\alpha)} \\
&\quad - 2e^{-\sigma^\alpha (|s_i|^\alpha + |2s_j|^\alpha)}, \\
\text{cov}(\Re(\phi_3(s_i, k)), \Re(\phi_2(s_j, k))) &= \text{cov}(\cos(s_j X_{t+k}), \cos(s_i (X_{t+k} - X_t))) \\
&\quad + \text{cov}(\cos(s_j X_{t+k}), \cos(s_i (X_{t+2k} - X_{t+k}))) \\
&= e^{-\sigma^\alpha (|s_i|^\alpha + |s_i - s_j|^\alpha)} + e^{-\sigma^\alpha (|s_i|^\alpha + |s_i + s_j|^\alpha)} \\
&\quad - 2e^{-\sigma^\alpha (|s_j|^\alpha + |2s_i|^\alpha)}, \\
\text{cov}(\Re(\phi_3(s_i, p)), \Re(\phi_3(s_j, q))) &= \text{cov}(\cos(s_i (X_{t+p} - X_t)), \cos(s_j (X_{t+q} - X_t))) \\
&\quad + \text{cov}(\cos(s_i (X_{t+p+q} - X_{t+q})), \cos(s_j (X_{t+q} - X_t))) \\
&\quad + \text{cov}(\cos(s_i (X_{t+p} - X_t)), \cos(s_j (X_{t+p+q} - X_{t+p}))) + c_{\Re}^{pq},
\end{aligned}$$

onde

$$c_{\Re}^{pq} = \begin{cases} 0 & \text{se } p = q, \\ \text{cov}(\cos(s_i (X_{t+q} - X_{t+q-p})), \cos(s_j (X_{t+q} - X_t))) & \text{se } q > p, \\ \text{cov}(\cos(s_i (X_{t+p} - X_t)), \cos(s_j (X_{t+p} - X_{t+p-q}))) & \text{se } p > q. \end{cases}$$

Quando $p = q$, temos

$$\begin{aligned} \text{cov}(\Re(\phi_3(s_i, p)), \Re(\phi_3(s_j, q))) &= \frac{1}{2}e^{-2\sigma^\alpha |s_i + s_j|^\alpha} + \frac{1}{2}e^{-2\sigma^\alpha |s_i - s_j|^\alpha} - 3e^{-\sigma^\alpha (2|s_i|^\alpha + 2|s_j|^\alpha)} \\ &\quad + e^{-\sigma^\alpha (|s_i|^\alpha + |s_j|^\alpha + |s_i - s_j|^\alpha)} + e^{-\sigma^\alpha (|s_i|^\alpha + |s_j|^\alpha + |s_i + s_j|^\alpha)}. \end{aligned}$$

e para $p \neq q$

$$\begin{aligned} \text{cov}(\Re(\phi_3(s_i, p)), \Re(\phi_3(s_j, q))) &= 2e^{-\sigma^\alpha (|s_i|^\alpha + |s_j|^\alpha + |s_i - s_j|^\alpha)} + 2e^{-\sigma^\alpha (|s_i|^\alpha + |s_j|^\alpha + |s_i + s_j|^\alpha)} \\ &\quad - 4e^{-\sigma^\alpha (2|s_i|^\alpha + 2|s_j|^\alpha)}, \end{aligned}$$

$$\begin{aligned} \text{cov}(\Im(\phi_1(s_i, p)), \Im(\phi_1(s_j, q))) &= \text{cov}(\text{sen}(-s_i X_t), \text{sen}(-s_j X_t)) \\ &= \frac{1}{2}\{e^{-\sigma^\alpha |s_i - s_j|^\alpha} - e^{-\sigma^\alpha |s_i + s_j|^\alpha}\}, \end{aligned}$$

$$\begin{aligned} \text{cov}(\Im(\phi_1(s_i, p)), \Im(\phi_2(s_j, q))) &= \text{cov}(\Im(\phi_2(s_i, p)), \Im(\phi_1(s_j, q))) \\ &= -\text{cov}(\Im(\phi_1(s_i, p)), \Im(\phi_1(s_j, q))), \end{aligned}$$

$$\text{cov}(\Im(\phi_2(s_i, p)), \Im(\phi_2(s_j, q))) = \text{cov}(\Im(\phi_1(s_i, p)), \Im(\phi_1(s_j, q))),$$

$$\begin{aligned} \text{cov}(\Im(\phi_3(s_i, p)), \Im(\phi_3(s_j, q))) &= \text{cov}(\text{sen}(s_i(X_{t+p} - X_t)), \text{sen}(s_j(X_{t+q} - X_t))) \\ &\quad + \text{cov}(\text{sen}(s_i(X_{t+p+q} - X_{t+q})), \text{sen}(s_j(X_{t+q} - X_t))) \\ &\quad + \text{cov}(\text{sen}(s_i(X_{t+p+q} - X_{t+p})), \text{sen}(s_j(X_{t+p} - X_t))) + \mathcal{C}_{\Im}^{pq}, \end{aligned}$$

onde

$$\mathcal{C}_{\Im}^{pq} = \begin{cases} 0 & \text{se } p = q, \\ \text{cov}(\text{sen}(s_i(X_{t+q} - X_{t+q-p})), \text{sen}(s_j(X_{t+q} - X_t))) & \text{se } q > p, \\ \text{cov}(\text{sen}(s_i(X_{t+p} - X_t)), \text{sen}(s_j(X_{t+p} - X_{t+p-q}))) & \text{se } p > q. \end{cases}$$

Quando $p = q$, temos

$$\begin{aligned} \text{cov}(\Im(\phi_3(s_i, k)), \Im(\phi_3(s_j, k))) &= \frac{1}{2}e^{-2\sigma^\alpha |s_i - s_j|^\alpha} \\ &\quad - \frac{1}{2}e^{-2\sigma^\alpha |s_i + s_j|^\alpha} + e^{-\sigma^\alpha (|s_i|^\alpha + |s_j|^\alpha + |s_i + s_j|^\alpha)} \\ &\quad - e^{-\sigma^\alpha (|s_i|^\alpha + |s_j|^\alpha + |s_i - s_j|^\alpha)} \end{aligned}$$

e $\text{cov}(\Im(\phi_3(s_i, k)), \Im(\phi_3(s_j, k))) = 0$ para $p \neq q$. Os outros elementos são iguais a zero. Os elementos de \mathbf{L}_2^k são dados em (B.10), onde os elementos de $\mathbf{d}_i^k, i = 1, \dots, r$ são

$$\begin{aligned} d_i^k(1,1) &= (\Re\Phi(0, -s_i; k))^{-1} = e^{\sigma^\alpha |s_i|^\alpha}, \\ d_i^k(2,2) &= (\Re\Phi(s_i, 0; k))^{-1} = e^{\sigma^\alpha |s_i|^\alpha}, \\ d_i^k(3,3) &= (\Re\Phi(s_i, -s_i; k))^{-1} = e^{2\sigma^\alpha |s_i|^\alpha}. \end{aligned}$$

A partir da equação (C.2) obtemos

$$m_{kk}^{RR} = \text{cov}(\Re\hat{\tau}(\mathbf{s}, k), \Re\hat{\tau}(\mathbf{s}, k)) = (\mathbf{L}_r \otimes \lambda_1) d^k V_{kk}^{RR} d^k (\mathbf{L}_r \otimes \lambda_1^T)$$

e

$$m_{kk}^{II} = \text{cov}(\mathfrak{S}\hat{\tau}(\mathbf{s},k), \mathfrak{S}\hat{\tau}(\mathbf{s},k)) = (\mathbf{I}_r \otimes \lambda_1) d^k V_{kk}^{RR} d^k (\mathbf{I}_r \otimes \lambda_1^T).$$

Logo, o bloco (i,j) de m_{kk}^{RR} e m_{kk}^{II} é encontrado a partir de

$$m_{kk}^{RR}(i,j) = \lambda_1 d_i^k V_{kk}^{RR}(i,j) d_j^k \lambda_1^T$$

e

$$m_{kk}^{II} = \lambda_1 d_i^k V_{kk}^{II}(i,j) d_j^k \lambda_1^T$$

que após um pouco de álgebra simples, obtemos

$$\begin{aligned} m_{kk}^{RR}(i,j) &= e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i - s_j|^\alpha)} \left\{ \frac{1}{2} e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i - s_j|^\alpha)} - 1 \right\} \\ &\quad + e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i + s_j|^\alpha)} \left\{ \frac{1}{2} e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i + s_j|^\alpha)} - 1 \right\} + 1, \\ m_{kk}^{II}(i,j) &= e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i - s_j|^\alpha)} \left\{ \frac{1}{2} e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i - s_j|^\alpha)} - 1 \right\} \\ &\quad + e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i + s_j|^\alpha)} \left\{ 1 - \frac{1}{2} e^{\sigma^\alpha(|s_i|^\alpha + |s_j|^\alpha - |s_i + s_j|^\alpha)} \right\}. \end{aligned}$$

O resultado necessário segue diretamente da equação (C.1)

APÊNDICE D – Continuação resultados da Simulação

Tabela 13 – Valores da função codiferença $I(\cdot)$ e sua estimativa $\hat{I}(\cdot)$ para séries simulada - Modelo 2

N	α	Método	s	I(1)	$\hat{I}^*(1)$	MAD_1	I(2)	$\hat{I}^*(2)$	MAD_2	I(3)	$\hat{I}^*(3)$	MAD_3
100	1	Simples	{0.01}	-0,6667	-0,3341	0,3395	0,3333	0,2031	0,1383	0	0,0014	0,0671
		Exp.	{0.01}		-0,3303	0,3428		0,2050	0,1371		-0,0081	0,0710
	2	Simples	{0.01, 0.1}	-0,6667	-0,6630	0,0436	0,2222	0,2078	0,0942	0	0,0040	0,1116
		Exp.	{0.01}		-0,6594	0,0435		0,2145	0,0929		-0,0019	0,1093
1000	1	Simples	{0.01}	-0,6667	-0,0339	0,6328	0,3333	0,1995	0,1340	0	-0,0001	0,0219
		Exp.	{0.01}		-0,0317	0,6349		0,1998	0,1336		-0,0004	0,0207
	2	Simples	{0.01, 0.1}	-0,6667	-0,6671	0,0135	0,2222	0,2205	0,0303	0	0,0026	0,0356
		Exp.	{0.01, 0.02, ..., 0.1}		-0,6662	0,0137		0,2212	0,0295		0,0000	0,0371

⁽¹⁾Tabela com os valores de $I(\cdot)$ e sua estimativa $\hat{I}(\cdot)$ a partir do processo MA(2), onde $X_t = a_t - a_{t-1} + 0.5a_{t-2}$ para $T = 1000$ replicações em amostras de tamanho 100 e 1000. O processo a_t é S α S para alguns valores de α e $\sigma = 1$. Na tabela,

$\hat{I}^*(i) = \frac{1}{T} \sum_{j=1}^T \Re \hat{I}(i)_j$, e $MAD_i = \frac{1}{T} \sum_{j=1}^T |\Re \hat{I}(i)_j - I(i)|$, $i = 1, 2$, onde $\Re \hat{I}(i)_j$ é a estimativa no lag i da j -ésima repetição.

⁽²⁾ A função $I(\cdot)$ não é definida quando α é não inteiros e pelo menos um coeficiente c_j é negativo.

Tabela 14 – Valores da função codiferença $I(\cdot)$ e sua estimativa $\hat{I}(\cdot)$ para séries simulada - Modelo 3

N	α	Método	s	I(1)	$\hat{I}^*(1)$	MAD_1	I(2)	$\hat{I}^*(2)$	MAD_2	I(3)	$\hat{I}^*(3)$	MAD_3
100	2	Simples	{0.01, 0.1, 0.2}	0,4425	0,4301	0,0600	0,0383	0,0095	0,0897	0	-0,0140	0,0911
		Exp.	{0.01}		0,4293			0,0140	0,0899		-0,0049	0,0883
	1,8	Simples	{0.01}	0,4317	0,4263	0,0571	0,0345	0,0104	0,0827	0	-0,0175	0,0828
		Exp.	{0.01, 0.1, 0.5}		0,4199			0,0064	0,0819		-0,0021	0,0820
	1,5	Simples	{0.01, 0.5}	0,4128	0,4065	0,0484	0,0300	0,0113	0,0707	0	-0,0204	0,0702
		Exp.	{0.01, 0.5}		0,4122			0,0112	0,0701		-0,0213	0,0686
	1,3	Simples	{0.01, 0.1, 0.2}	0,3985	0,4004	0,0453	0,0287	0,0120	0,0643	0	-0,0157	0,0635
		Exp.	{0.01, 0.1, 0.5}		0,4008			0,0085	0,0626		-0,0207	0,0611
	1	Simples	{0.01, 0.02, ..., 0.2}	0,3750	0,3705	0,0399	0,0313	0,0102	0,0563	0	-0,0199	0,0549
		Exp.	{0.01, 0.02, ..., 0.2}		0,3695			0,0369	0,0565		0,0135	0,0526
	0,8	Simples	{0.01, 0.02, ..., 0.2}	0,3590	0,3504	0,0364	0,0384	0,0176	0,0558	0	-0,0185	0,0520
		Exp.	{0.01, 0.02, ..., 0.2}		0,3531			0,0211	0,0559		0,0170	0,0525
0,5	Simples	{0.01, 0.02, ..., 0.1}	0,3381	0,3297	0,0342	0,0633	0,0404	0,0672	0	-0,0273	0,0730	
	Exp.	{0.01, 0.02, ..., 0.2}		0,3294			0,0384	0,0666		-0,0193	0,0717	

Tabela 14 - Continuação.

N	α	Método	s	I(1)	$\hat{I}^*(1)$	MAD_1	I(2)	$\hat{I}^*(2)$	MAD_2	I(3)	$\hat{I}^*(3)$	MAD_3
1000	2	Simple	{0.01, 0.2}	0,4425	0,4410	0,0183	0,0383	0,0374	0,0294	0	-0,0005	0,0286
		Exp.	{0.2}		0,4395	0,0189		0,0303	0,0264		-0,0038	0,0284
	1,8	Simple	{0.1, 0.3, 0.5}	0,4317	0,4310	0,0172	0,0300	0,0288	0,0231	0	-0,0027	0,0232
		Exp.	{0.01, 0.2}		0,4345	0,0165		0,0270	0,0188		-0,0031	0,0191
	1,5	Simple	{0.1, 0.3, 0.5}	0,4128	0,4117	0,0161	0,0313	0,0294	0,0184	0	-0,0027	0,0185
		Exp.	{0.01, 0.02, ..., 0.2}		0,4144	0,0161		0,0366	0,0178		-0,0027	0,0188
	1,3	Simple	{0.01, 0.02, ..., 0.2}	0,3985	0,3992	0,0148	0,0633	0,0615	0,0194	0	-0,0025	0,0213
		Exp.	{0.1, 0.3, 0.5}		0,3976	0,0146		0,0365	0,0290		0,0000	0,0306
	1	Simple	{0.01, 0.02, ..., 0.2}	0,3750	0,3745	0,0122	0,0345	0,0337	0,0256	0	-0,0015	0,0256
		Exp.	{0.01, 0.02, ..., 0.2}		0,3751	0,0123		0,0288	0,0211		-0,0030	0,0209
	0,8	Simple	{0.01, 0.02, ..., 0.2}	0,3590	0,3576	0,0110	0,0287	0,0258	0,0231	0	-0,0044	0,0225
		Exp.	{0.01, 0.02, ..., 0.1}		0,3580	0,0109		0,0302	0,0174		-0,0003	0,0181
	0,5	Simple	{0.01, 0.02, ..., 0.1}	0,3381	0,3381	0,0107	0,0384	0,0360	0,0203	0	-0,0036	0,0181
		Exp.	{0.01, 0.02, ..., 0.1}		0,3375	0,0106		0,0603	0,0216		-0,0034	0,0211

Tabela com os valores de $I(\cdot)$ e sua estimativa $\hat{I}(\cdot)$ a partir do processo MA(2), onde $X_t = a_t + 0.55a_{t-1} + 0.05a_{t-2}$ para $T = 1000$ replicações em amostras de tamanho 100 e 1000. O processo a_t é SaS para alguns valores de α e $\sigma = 1$. Na tabela,

$\hat{I}^*(i) = \frac{1}{T} \sum_{j=1}^T \Re \hat{I}(i)_j$, e $MAD_i = \frac{1}{T} \sum_{j=1}^T |\Re \hat{I}(i)_j - I(i)|$, $i = 1, 2$, onde $\Re \hat{I}(i)_j$ é a estimativa no lag i da j -ésima repetição.

Tabela 15 – Valores da função codiferença $I(\cdot)$ e sua estimativa $\hat{I}(\cdot)$ para séries simulada - Modelo 4

N	α	Método	s	I(1)	$\hat{I}^*(1)$	MAD_1	I(2)	$\hat{I}^*(2)$	MAD_2	I(3)	$\hat{I}^*(3)$	MAD_3
100	1	1 {0.01, 0.1}	-0,2353	-0,1361	0,1455	0,4118	0,3517	0,0792	0	-0,0168	0,0643	
100	1	2 {0.01, 0.2}		-0,1360	0,1362		0,3530	0,0729		-0,0220	0,0694	
100	2	1 {0.01, 0.2}	-0,4121	-0,4236	0,0894	0,4242	0,4091	0,0669	0	-0,0065	0,1044	
100	2	2 {0.01, 0.1, 0.2}		-0,4159	0,0880		0,4082	0,0716		-0,0068	0,1009	
1000	1	1 {0.01, 0.2}	-0,2353	-0,0175	0,2178	0,4118	0,3354	0,0764	0	-0,0007	0,0203	
1000	1	2 {0.01}		-0,0321	0,2151		0,3352	0,0796		-0,0018	0,0198	
1000	2	1 {0.01, 0.1}	-0,4121	-0,4107	0,0287	0,4242	0,4236	0,0214	0	0,0014	0,0331	
1000	2	2 {0.01, 0.06, 0.11, 0.16, 0.21}		-0,4130	0,0284		0,4228	0,0218		0,0011	0,0330	

Tabela com os valores de $I(\cdot)$ e sua estimativa $\hat{I}(\cdot)$ a partir do processo MA(2), onde $X_t = a_t - 0.4a_{t-1} + 0.7a_{t-2}$ para $T = 1000$ replicações em amostras de tamanho 100 e 1000. O processo a_t é S α S para alguns valores de α e $\sigma = 1$. Na tabela,

$\hat{I}^*(i) = \frac{1}{T} \sum_{j=1}^T \Re \hat{I}(i)_j$, e $MAD_i = \frac{1}{T} \sum_{j=1}^T |\Re \hat{I}(i)_j - I(i)|$, $i = 1, 2$, onde $\Re \hat{I}(i)_j$ é a estimativa no lag i da j -ésima repetição.

⁽²⁾ A função $I(\cdot)$ não é definida quando α é não inteiros e pelo menos um coeficiente c_j é negativo.