



**UNICAMP**

UNIVERSIDADE ESTADUAL DE  
CAMPINAS

Instituto de Matemática, Estatística e  
Computação Científica

JEAN SABINO MAGALHÃES DINIZ

**Estimação Robusta de Modelos GARCH e DCC  
com Mudança de Regime Markoviano**

Campinas

2022

Jean Sabino Magalhães Diniz

# **Estimação Robusta de Modelos GARCH e DCC com Mudança de Regime Markoviano**

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Estatística.

Orientador: Luiz Koodi Hotta

Este exemplar corresponde à versão final da Dissertação defendida pelo aluno Jean Sabino Magalhães Diniz e orientada pelo Prof. Dr. Luiz Koodi Hotta.

Campinas

2022

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Matemática, Estatística e Computação Científica  
Ana Regina Machado - CRB 8/5467

D615e      Diniz, Jean Sabino Magalhães, 1996-  
Estimação robusta de modelos GARCH e DCC com mudança de regime markoviano / Jean Sabino Magalhães Diniz. – Campinas, SP : [s.n.], 2022.

Orientador: Luiz Koodi Hotta.

Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Métodos estatísticos robustos. 2. Modelo de volatilidade multivariada. 3. Modelos markovianos ocultos. 4. Valores estranhos (Estatística). 5. Valor em Risco (VaR). I. Hotta, Luiz Koodi, 1952-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Robust estimation of Markov-Switching GARCH and DCC models

**Palavras-chave em inglês:**

Robust statistical methods

Multivariate volatility model

Hidden Markov models

Outliers (Statistics)

Value at risk (VaR)

**Área de concentração:** Estatística

**Titulação:** Mestre em Estatística

**Banca examinadora:**

Luiz Koodi Hotta [Orientador]

Mauricio Enrique Zevallos Herencia

Pedro Luiz Valls Pereira

**Data de defesa:** 10-08-2022

**Programa de Pós-Graduação:** Estatística

**Identificação e informações acadêmicas do(a) aluno(a)**

- ORCID do autor: <https://orcid.org/0000-0002-1488-0729>

- Currículo Lattes do autor: <http://lattes.cnpq.br/4531741306416593>

**Dissertação de Mestrado defendida em 10 de agosto de 2022 e aprovada  
pela banca examinadora composta pelos Profs. Drs.**

**Prof(a). Dr(a). LUIZ KOODI HOTTA**

**Prof(a). Dr(a). MAURICIO ENRIQUE ZEVALLOS HERENCIA**

**Prof(a). Dr(a). PEDRO LUIZ VALLS PEREIRA**

A Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação do Instituto de Matemática, Estatística e Computação Científica.



# Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Agradeço a Deus, familiares e amigos pelo apoio durante todo o curso.

Agradeço fortemente ao meu orientador Prof. Dr. Luiz Koodi Hotta pela paciência, compreensão, por tudo que me ensinou e por fazer jus à palavra orientador.

*“Instinto é algo vital para duas  
existências que tem poder  
igual, mas um quer algo além.  
É algo vital pra se tornar um rei.”  
(Bleach)*

# Resumo

O estudo da volatilidade se demonstrou importante em diversos campos da área financeira. Nesse contexto, os modelos GARCH tem sido frequentemente utilizado e estudado na literatura devido a sua capacidade de captar alguns dos fatos estilizados das séries temporais financeiras. Entretanto, em diversa dessas séries também é observado a presença de *outliers*. E nesses casos, vários autores mostram que os estimadores comuns tendem a não ter uma boa performance. Para lidar com esses casos, há várias metodologias presentes, se destacando a estimação robusta que utiliza simultaneamente um mecanismo de filtro robusto no cálculo da variância condicional e M-estimadores como função objetivo. No caso multivariado, isto é, quando estamos analisando a volatilidade de mais de um ativo financeiro, o modelo DCC está entre os mais utilizados na literatura. E nesse caso também é observado que, na presença de *outliers*, a estimação dos parâmetros desse modelo é afetada negativamente e os mecanismos robustos utilizados no modelo GARCH também apresentaram desempenho satisfatórios.

No entanto, em algumas séries temporais financeiras existe a mudança de estrutura na dinâmica da volatilidade. E nesses casos, os modelos GARCH tendem a se tornarem inadequados. Sendo assim, o modelo frequentemente utilizado para esses casos é o modelo GARCH com mudança de regime markoviano. Na literatura há vários métodos para a estimação dos parâmetros desse modelo, sendo o mais frequente a estimação via filtro de Hamilton. Porém, a estimação robusta a *outliers* não foi abordada na literatura. Dito isso, essa dissertação avalia as técnicas robustas utilizadas no modelo GARCH adaptada aos modelos GARCH e DCC com mudança de regime markoviano. Os estimadores são avaliados pela precisão, desempenho do filtro e a capacidade preditiva na presença de *outliers* aditivos através de simulação. Eles também são avaliados através de aplicações a dados empíricos.

**Palavras-chave:** Métodos estatísticos robustos, Modelo de volatilidade multivariada, Modelos markovianos ocultos, Valores estranhos (Estatística), Valor em Risco (VaR).

# Abstract

The study of volatility proved to be important in several fields of the financial area. In this context, GARCH models have been frequently used and studied in the literature due to their ability to capture some of the stylized facts of financial time series. However, in various of these series, the presence of outliers is also observed. And in these cases, several authors show that common estimators tend not to perform well. To deal with these cases there are several methodologies, highlighting the robust estimation that simultaneously uses a robust filter mechanism to calculate the conditional variance and M-estimators as an objective function. In the multivariate case, the DCC model is among the most used in the literature. And in this case, it is observed that, in the presence of outliers, the estimation of the parameters of this model is negatively affected and the robust mechanisms used in the GARCH model also presented satisfactory performance.

However, in diverse financial time series there is a change of structure in the volatility dynamics. And in these cases, GARCH models tend to become inadequate and Markov-Switching GARCH models are frequently used. In the literature, there are several methods for estimating the parameters of these models, the most frequent being the estimation via Hamilton filter. However, robust estimation of outliers has not been addressed in the literature. That said, this dissertation evaluates the robust techniques used in the GARCH model adapted to Markov-Switching GARCH and DCC models. The estimators are evaluated by accuracy, filter performance and predictive ability in the presence of additive outliers through simulation. They are also evaluated through application to empirical data.

**Keywords:** Robust statistical methods, Multivariate volatility model, Hidden Markov models, Outliers (Statistics), Value at risk (VaR).

# Lista de ilustrações

- Figura 1 – Probabilidades filtradas para uma série simulada de tamanho 3000 contaminada com 30 *outliers* de tamanho 5 vezes o desvio padrão condicional do modelo MS(2)-GARCH(1,1) com parâmetros  $\alpha_0^{(1)} = 2$ ,  $\alpha_1^{(1)} = 0,1$ ,  $\beta_1^{(1)} = 0,6$ ,  $\alpha_0^{(2)} = 0,3$ ,  $\alpha_1^{(2)} = 0,35$ ,  $\beta_1^{(2)} = 0,2$ ,  $p_{11} = 0,96$ ,  $p_{22} = 0,98$ . A linha preta representa a probabilidade filtrada, a linha azul representa o estado real e a linha tracejada em vermelho representa a posição do *outlier*. O 1 representa o regime de alta volatilidade, isto é, <sup>(1)</sup> e 0 representa o regime de baixa volatilidade, isto é, <sup>(2)</sup>. . . . . 68
- Figura 2 – Proporção de classificação correta dos diferentes estimadores para o modelo MS-GARCH nas observações contaminadas. No eixo horizontal estão as porcentagens de *outliers*. Preenchido em branco o estimador QML-n, com bolas o estimador QML-t e em linhas diagonais o estimador Robusto. Os quadros da esquerda são referentes à quando o regime correto é de alta volatilidade, ou seja, o regime (1), os da direita para o regime de baixa volatilidade, regime (2). Nos quadros superiores são *outliers* de tamanho 3 vezes o desvio padrão condicional, nos inferiores 5 vezes o desvio padrão condicional. . . . . 69
- Figura 3 – Proporção de classificação correta dos diferentes estimadores para o modelo MS-GARCH para todas as observações. No eixo horizontal estão as porcentagens de *outliers*. Preenchido em branco o estimador QML-n, com bolas o estimador QML-t e em linhas diagonais o estimador Robusto. Os quadros da esquerda são referentes à quando o regime correto é de alta volatilidade, ou seja, o regime (1), os da direita para o regime de baixa volatilidade, regime (2). Nos quadros superiores são *outliers* de tamanho 3 vezes o desvio padrão condicional, nos inferiores 5 vezes o desvio padrão condicional. . . . . 69
- Figura 4 – Proporção de classificação correta dos diferentes estimadores para o modelo IS-DCC nas observações contaminadas. No eixo horizontal estão as porcentagens de *outliers*. Preenchido em branco o estimador QML-n, com bolas o estimador QML-t e em linhas diagonais o estimador Robusto. Os quadros da esquerda são referentes à quando o regime correto é de maior persistência na correlação, ou seja, o regime (1), os da direita para o regime de menor persistência na correlação, regime (2). Nos quadros superiores são *outliers* de tamanho 3 vezes o desvio padrão condicional, nos inferiores 5 vezes o desvio padrão condicional. . . . . 70

Figura 5 – Proporção de classificação correta dos diferentes estimadores para o modelo IS-DCC para todas as observações. No eixo horizontal estão as porcentagens de *outliers*. Preenchido em branco o estimador QML-n, com bolas o estimador QML-t e em linhas diagonais o estimador Robusto. Os quadros da esquerda são referentes à quando o regime correto é de maior persistência na correlação, ou seja, o regime (1), os da direita para o regime de menor persistência na correlação, regime (2). Nos quadros superiores são *outliers* de tamanho 3 vezes o desvio padrão condicional, nos inferiores 5 vezes o desvio padrão condicional. 70

Figura 6 – Boxplot das estimativas dos parâmetros do regime de alta volatilidade das replicações Monte Carlo do modelo MS-GARCH(1,1). O eixo x representa a porcentagem de *outliers* e o eixo y representa os valores das estimativas dos parâmetros. No quadro esquerdo são os valores para os tamanhos de *outliers* com  $d = 3$  e no direito  $d = 5$ . Em branco o estimador QML-n, verde claro o estimador QML-t e verde escuro o estimador Robusto. A linha tracejada em vermelho representa o valor real do parâmetro. Os parâmetros em cada gráfico são: (a)  $\alpha_0^{(1)}$  (b)  $\alpha_1^{(1)}$  (c)  $\beta_1^{(1)}$  (d)  $p_{11}$ . <sup>(1)</sup> representa o regime de alta persistência e <sup>(2)</sup> representa o regime de baixa persistência. . . . . 71

Figura 7 – Boxplot das estimativas dos parâmetros do regime de baixa volatilidade das replicações Monte Carlo do modelo MS-GARCH(1,1). O eixo x representa a porcentagem de *outliers* e o eixo y representa os valores das estimativas dos parâmetros. No quadro esquerdo são os valores para os tamanhos de *outliers* com  $d = 3$  e no direito  $d = 5$ . Em branco o estimador QML-n, verde claro o estimador QML-t e verde escuro o estimador Robusto. A linha tracejada em vermelho representa o valor real do parâmetro. Os parâmetros em cada gráfico são: (a)  $\alpha_0^{(2)}$  (b)  $\alpha_1^{(2)}$  (c)  $\beta_1^{(2)}$  (d)  $p_{22}$ . . . . . 72

Figura 8 – Boxplot das estimativas dos parâmetros da correlação das replicações Monte Carlo do modelo IS-DCC(1,1). O eixo x representa a porcentagem de *outliers* e o eixo y representa os valores das estimativas dos parâmetros. No quadro esquerdo são os valores para os tamanhos de *outliers* com  $d = 3$  e no direito  $d = 5$ . Em branco o estimador QML-n, verde claro o estimador QML-t e verde escuro o estimador Robusto. A linha tracejada em vermelho representa o valor real do parâmetro. Os parâmetros em cada gráfico são: (a)  $a_1^{(1)}$  (b)  $a_1^{(2)}$  (c)  $b_1^{(1)}$  (d)  $b_1^{(2)}$ . . . . . 74

|  |    |
|--|----|
| Figura 9 – Boxplot das estimativas dos parâmetros da matriz de transição das replicações Monte Carlo do modelo IS-DCC(1,1). O eixo x representa a porcentagem de <i>outliers</i> e o eixo y representa os valores das estimativas dos parâmetros. No quadro esquerdo são os valores para os tamanhos de <i>outliers</i> com $d = 3$ e no direito $d = 5$ . Em branco o estimador QML-n, verde claro o estimador QML-t e verde escuro o estimador Robusto. A linha tracejada em vermelho representa o valor real do parâmetro. Os parâmetros em cada gráfico são: (a) $p_{11}$ (b) $p_{22}$ . . . . . | 75 |
| Figura 10 – Evolução diária das taxas de câmbio em relação ao dólar americano no gráfico superior e no inferior os respectivos log-retornos (a) Euro (b) Libra (c) Yen (d) Franco suíço. . . . .   | 92 |
| Figura 11 – Do lado esquerdo a função de auto-correlação amostral dos log-retornos e do lado direito a função de auto-correlação amostral dos log-retornos ao quadrado. . . . .  | 93 |
| Figura 12 – Probabilidades estimadas de estar no regime de alta volatilidade pelo os três estimadores para o Yen, utilizando a primeira janela de estimação. Os estimadores são: (a) Robusto (b) QML-n (c) QML-t. . . . .  | 94 |
| Figura 13 – Estimativa do VaR para as quatro moedas. A linha de cor preta representa o log retorno filtrado por um AR(1), a linha representa o VaR estimado pelo estimador robusto, a linha em azul por QML-n e a linha verde por QML-t. Do lado esquerdo o VaR 5% e do direito VaR 1%. . .  | 95 |
| Figura 14 – Probabilidades estimadas de estar no regime de alta correlação pelo os três estimadores para o par Euro-Libra, utilizando a primeira janela de estimação. Os estimadores são: (a) Robusto (b) QML-n (c) QML-t. . .   | 96 |
| Figura 15 – Estimativa do VaR para os dois pares de moedas. A linha de cor preta representa o log retorno da carteira, a linha representa o VaR estimado pelo estimador robusto, a linha em azul por QML-n e a linha verde por QML-t. Do lado esquerdo o VaR 5% e do direito VaR 1%. . . . .   | 97 |

# Lista de tabelas

|          |   |    |
|----------|---|----|
| Tabela 1 | – Resultados das estimativas dos parâmetros de aplicações de aplicações apresentados por Haas, Mittnik e Paoletta (2004) e Paoletta et al. (2012); e processo gerador de dados (PGD) de simulações utilizadas por Augustyniak (2014), Billio, Casarin e Osuntuyi (2016) e Bauwens, Preminger e Rombouts (2010) para o modelo MS(2)-GARCH(1,1). $V^{(i)} = \alpha_0^{(i)} / (1 - \alpha_1^{(i)} - \beta_1^{(i)})$ para $\alpha_1^{(i)} + \beta_1^{(i)} < 1$ e $\infty$ para $\alpha_1^{(i)} + \beta_1^{(i)} \geq 1$ , EM representa o algoritmo EM e Bay. é estimação via técnicas bayesianas. | 59 |
| Tabela 2 | – Resultados das estimativas dos parâmetros de aplicações para o modelo IS(2)-DCC(1,1) apresentados por Lee e Song (2009).  | 60 |
| Tabela 3 | – Viés e RMSE, para os estimadores do modelo MS(2)-GARCH(1,1) com parâmetros $\alpha_0^{(1)} = 2$ , $\alpha_1^{(1)} = 0,1$ , $\beta_1^{(1)} = 0,6$ , $\alpha_0^{(2)} = 0,3$ , $\alpha_1^{(2)} = 0,35$ , $\beta_1^{(2)} = 0,2$ , $p_{11} = 0,96$ , $p_{22} = 0,98$ , $per^{(i)}$ é a persistência do regime $i$ , $^{(1)}$ representa o regime de alta volatilidade e $^{(2)}$ representa o regime de baixa volatilidade e $T = 3000$ . Em negrito são os que apresentaram os melhores resultados.   | 65 |
| Tabela 4 | – Viés e RMSE, para os estimadores do modelo IS(2)-DCC(1,1) com parâmetros $a_1^{(1)} = 0,04$ , $b_1^{(1)} = 0,93$ , $a_1^{(2)} = 0,2$ , $b_1^{(2)} = 0,6$ , $p_{11} = 0,96$ , $p_{22} = 0,98$ , $per^{(i)}$ é a persistência do regime $i$ , $^{(1)}$ representa o regime de alta persistência e $^{(2)}$ representa o regime de baixa persistência e $T = 3000$ . Em negrito são os que apresentaram os melhores resultados.  | 73 |
| Tabela 5 | – Estatísticas descritivas das taxas de câmbio.   | 78 |
| Tabela 6 | – Estimativas dos parâmetros do modelo MS(2)-GARCH(1,1) para a série do log-retorno da taxa de câmbio do Euro em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, $t$ e o estimador robusto.   | 83 |
| Tabela 7 | – Estimativas dos parâmetros do modelo MS(2)-GARCH(1,1) para a série do log-retorno da taxa de câmbio da Libra em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, $t$ e o estimador robusto.  | 84 |
| Tabela 8 | – Estimativas dos parâmetros do modelo MS(2)-GARCH(1,1) para a série do log-retorno da taxa de câmbio do Yen em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, $t$ e o estimador robusto.  | 84 |



|  |    |
|--|----|
| Tabela 9 – Estimativas dos parâmetros do modelo MS(2)-GARCH(1,1) para a série do log-retorno da taxa de câmbio do Franco Suíço em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto. . . . .   | 85 |
| Tabela 10 – Back-testing para o VaR 5% e 1% para as quatro moedas utilizando o modelo MS(2)-GARCH(1,1) sob várias estimações. Janela móvel de tamanho 3000 iniciando em 02-07-2010. P-valores maiores que 0,05 estão em negrito. UC significa o teste de cobertura incondicional de Kupiec (1995). CC e IND são, respectivamente os testes de cobertura condicional e independência de Christoffersen (1998). DQ é o teste de Engle e Manganelli (2004). . . . . | 87 |
| Tabela 11 – Estimativas dos parâmetros do modelo IS(2)-DCC(1,1) para a série bivariado dos log-retornos da taxa de câmbio do Euro e Libra em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto. . . . .  | 88 |
| Tabela 12 – Estimativas dos parâmetros do modelo IS(2)-DCC(1,1) para a série bivariado dos log-retornos da taxa de câmbio do Euro e Yen em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto, . . . . .  | 89 |
| Tabela 13 – Estimativas dos parâmetros do modelo IS(2)-DCC(1,1) para a série bivariado dos log-retornos da taxa de câmbio do Libra e Franco Suíço em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto. . . . .  | 89 |
| Tabela 14 – Back-testing para o VaR 5% e 1% para as os casos de Euro-Libra, Euro-Yen e Libra-Franco Suíço. Janela móvel de tamanho 3000 iniciando em 02-07-2010. P valores menores que 0,05 estão em negrito. UC significa o teste de cobertura incodicional de Kupiec (1995). CC e IND são, respectivamente os testes de cobertura condicional e independência de Christoffersen (1998). DQ é o teste de Engle e Manganelli (2004). . . .                       | 90 |

# Sumário

|            |   |           |
|------------|---|-----------|
| <b>1</b>   | <b>INTRODUÇÃO</b>   | <b>16</b> |
| <b>2</b>   | <b>MODELOS GARCH COM 1 REGIME</b>   | <b>21</b> |
| <b>2.1</b> | <b>Estimadores Não Robustos</b>   | <b>24</b> |
| 2.1.1      | Estimação por máxima quase-verossimilhança                                | 24        |
| <b>2.2</b> | <b>Estimadores Robustos dos Modelos GARCH</b>                             | <b>27</b> |
| 2.2.1      | BIP-GARCH   | 28        |
| <b>2.3</b> | <b>Previsão das Volatilidades</b>   | <b>30</b> |
| <b>3</b>   | <b>MODELOS GARCH UNIVARIADOS COM MUDANÇA DE RE-<br/>GIME MARKOVIANO</b>   | <b>32</b> |
| <b>3.1</b> | <b>Modelo de Haas, Mittnik e Paoella (2004)</b>                           | <b>34</b> |
| <b>3.2</b> | <b>Estimadores Não Robustos</b>   | <b>37</b> |
| 3.2.1      | Estimação via filtro de Hamilton  | 37        |
| 3.2.2      | Observações sobre a estimação por máxima-verossimilhança                  | 39        |
| <b>3.3</b> | <b>Estimação Robusta</b>  | <b>40</b> |
| <b>3.4</b> | <b>Previsão das Volatilidades</b>   | <b>41</b> |
| <b>4</b>   | <b>MODELOS GARCH MULTIVARIADOS COM MUDANÇA DE RE-<br/>GIME MARKOVIANO</b> | <b>44</b> |
| <b>4.1</b> | <b>Modelo DCC com 1 Regime</b>  | <b>45</b> |
| 4.1.1      | Estimação por máxima quase-verossimilhança                                | 47        |
| 4.1.2      | Estimadores robustos  | 49        |
| <b>4.2</b> | <b>Modelos DCC com Mudança de Regime</b>                                  | <b>52</b> |
| 4.2.1      | Modelo de Lee (2010)  | 53        |
| 4.2.2      | Estimação via filtro de Hamilton  | 54        |
| 4.2.3      | Estimação robusta   | 56        |
| <b>5</b>   | <b>SIMULAÇÃO</b>  | <b>57</b> |
| <b>5.1</b> | <b>Processos Geradores de Dados</b>                                       | <b>57</b> |
| 5.1.1      | Caso univariado   | 58        |
| 5.1.2      | Caso bivariado  | 59        |
| <b>5.2</b> | <b>Estimadores Considerados</b>   | <b>60</b> |
| <b>5.3</b> | <b>Valores Iniciais</b>   | <b>61</b> |
| <b>5.4</b> | <b>Resultados</b>   | <b>62</b> |
| 5.4.1      | Análise do desempenho dos filtros   | 62        |

|            |  |            |
|------------|--|------------|
| 5.4.1.1    | Caso univariado . . . . .                                | 62         |
| 5.4.1.2    | Caso bivariado . . . . .                                 | 63         |
| 5.4.2      | Estimação dos modelos . . . . .                          | 64         |
| 5.4.2.1    | Caso univariado . . . . .                                | 64         |
| 5.4.2.2    | Caso bivariado . . . . .                                 | 66         |
| <b>5.5</b> | <b>Conclusões . . . . .</b>                              | <b>66</b>  |
| <b>6</b>   | <b>APLICAÇÕES . . . . .</b>                              | <b>76</b>  |
| <b>6.1</b> | <b>Análise Descritiva . . . . .</b>                      | <b>77</b>  |
| <b>6.2</b> | <b>Medidas de Desempenho das Previsões . . . . .</b>     | <b>79</b>  |
| 6.2.1      | Testes de desempenho de previsão . . . . .               | 80         |
| <b>6.3</b> | <b>Análise Univariada . . . . .</b>                      | <b>82</b>  |
| 6.3.1      | Estimação das probabilidades dos regimes . . . . .       | 85         |
| 6.3.2      | Análise das estimativas do VaR e das previsões . . . . . | 85         |
| <b>6.4</b> | <b>Análise Bivariada . . . . .</b>                       | <b>86</b>  |
| 6.4.1      | Estimação das probabilidades dos regimes . . . . .       | 88         |
| 6.4.2      | Análise das estimativas do VaR e das previsões . . . . . | 89         |
| <b>6.5</b> | <b>Comentários Finais . . . . .</b>                      | <b>91</b>  |
| <b>7</b>   | <b>CONSIDERAÇÕES FINAIS . . . . .</b>                    | <b>98</b>  |
|            | <b>REFERÊNCIAS . . . . .</b>                             | <b>100</b> |

# 1 Introdução

O estudo da volatilidade tem sido uma das áreas de pesquisa mais frequentes na econometria financeira nas últimas décadas. Do ponto de vista financeiro, segundo Hull (2022, Glossário de termos, pág., 849), podemos definir a volatilidade com sendo uma medida de incerteza do retorno realizado sobre um ativo. Além disso, existem duas de volatilidades que são comumente utilizados, a volatilidade implícita e a histórica. De acordo com Hull (2022, Seção 15.11, pág. 358), a volatilidade implícita é a volatilidade implicada por um preço de uma opção, que são derivativos<sup>1</sup> que representam contratos onde se negocia o direito de compra e venda de um determinado ativo por um período, além, claro, do direito de “não fazer nada”. Já a volatilidade histórica no tempo  $t$  é geralmente definido como o desvio padrão dos retornos observados em um dado período  $\{t - p + 1, \dots, t\}$ .

Naturalmente, uma das utilidades de se estimar a volatilidade é utilizá-la para precificar opções. Para gestores e investidores experientes, o uso apropriado desse mecanismo pode trazer lucros aos seus portfólios de investimentos. Podemos ver, por exemplo, que em março de 2021 investidores com opções de venda de ações da Petrobras podem ter lucrado cerca de R\$ 18 milhões com a queda dos papéis da companhia (NETO, 2021). Hull (2022, Capítulo 12) traz também alguns exemplos de negociações envolvendo opções. Nesse contexto de precificação de opções, no início da década de 1970, Black e Scholes (1973) propõem uma fórmula que fornece o prêmio de uma opção europeia a partir do preço atual do ativo, o tempo de expiração da opção e alguns parâmetros acessíveis, enquanto Merton (1973) elaborou uma estrutura teórica que serve como base para critérios de precificação de opções (OLIVEIRA, 1998). A importância e relevância dos métodos propostos por eles, resultaram no Prêmio Nobel da economia de 1997. O que se observa é que, nesse método e em diversos outros utilizados para precificação, a volatilidade emprega um papel importante.

Existem outras aplicações onde a volatilidade é relevante, podemos destacar algumas medidas de risco em um portfólio como o Valor-em-Risco (VaR - *Value at Risk* em inglês), que pode ser definido como uma perda que não excederá um nível de confiança especificado. Uma aplicação relevante é mostrada por Hull (2022, História de Negócios 22.1, pág. 533). Em 1988, o Comitê de Supervisão Bancária de Basileia publicou um documento conhecido como Basileia I, que é um acordo entre os reguladores sobre calcular o capital que cada banco é obrigado a manter para fins de risco de crédito. Já em 1996 foi publicado uma emenda que exigia que os bancos mantivessem capital para proteção contra risco de mercado e não apenas risco de crédito. A emenda diferencia entre a carteira negociável do banco e a sua carteira de investimento em títulos. A Emenda de 1996 calcula

<sup>1</sup> Derivativos são instrumentos financeiros cujo valor é “derivado” do valor de um ativo.

o capital para a carteira negociável usando a medida VaR. Com o passar do tempo, vieram a Basileia II, Basileia II.5, Basileia III. E em 2012, o Comitê da Basileia trouxe para debate a substituição do VaR por ES (*Expected Shortfall* em inglês), que é a perda esperada durante um período de  $N$  dias, desde que a perda seja pior que a perda do VaR.

Do ponto de vista empírico, diversos autores observam que as séries temporais financeiras apresentam alguns fatos estilizados como cita Francq e Zakoian (2019, Seção 1.3, pág. 7). Nesse cenário, Engle (1982) propõe os modelos autorregressivos condicionalmente heterocedásticos (ARCH) para modelar a variância condicional da série de inflação Reino Unido. Esse modelo é dado de forma que a volatilidade no tempo  $t$  é obtida em função do quadrado das  $p$  observações anteriores da série. No entanto, em diversas séries, ao estimar os modelos ARCH observou-se que ordem  $p$  era extremamente alta. Posto isso, Bollerslev (1986) apresenta uma extensão do modelo ARCH, GARCH (*Generalized ARCH* em inglês) também com intuito de modelar a volatilidade e capturar suas características. Nos modelos GARCH, a variância condicional é expressa como uma função linear dos valores passados ao quadrado da série e das  $q$  volatilidades anteriores.

Existem diversas maneiras para estimar os parâmetros dos modelos GARCH, tanto bayesianas quanto frequentistas. A mais popular entre elas é o método de máxima quase-verossimilhança. Francq e Zakoian (2019, Seção 7, pág. 175) mostram que, sob certas condições, esses estimadores apresentam ótimas propriedades assintóticas, como consistência, por exemplo. E ao utilizar esse método de estimação, diversos autores utilizam a distribuição Normal e t-Student nas perturbações (inovações). Porém, mesmo não sendo listado como um dos fatos estilizados das séries temporais financeiras, diversas séries apresentam valores atípicos, que aqui serão chamados de *outliers*. E mesmo na presença de apenas um deles, tanto a estimação do modelo quanto a previsão das volatilidades podem ser afetadas, levando a estimadores com grande viés e aumento da variância.

Para lidar com *outliers* na estimação dos modelos GARCH, existem duas metodologias que são utilizadas com mais frequência: a primeira é detectar os *outliers* e tratá-los de alguma forma apropriada; e a segunda é utilizar estimadores robustos a *outliers*. Hotta e Tsay (2012) apresentam, baseado no modelo ARCH e no modelo GARCH com  $p = q = 1$ , um estudo do impacto do *outliers* na estimativa de máxima verossimilhança além de um método de detecção de *outliers*. Em relação aos estimadores robustos, vários autores apresentaram deferentes métodos para estimar os parâmetros do modelo GARCH de maneira robusta a *outliers*. Por exemplo, Sakata e White (1998) utilizam os S-estimadores, Muler e Yohai (2008) apresentam um estimador baseado nos M-estimadores aliado à um mecanismo de filtro robusto para reduzir o efeito de *outliers* nas variâncias condicionais. Boudt, Danielsson e Laurent (2013) propõem algumas modificações na função objetivo do M-Estimador e do mecanismo usado por Muler e Yohai (2008). Na presença de *outliers* esses estimadores têm um ganho significativo em relação aos estimadores não robustos,

como mostram Carnero, Peña e Ruiz (2012) para o estimador de Muler e Yohai (2008). Adicionalmente Trucíos, Hotta e Ruiz (2017) apresentam uma modificação para previsão robusta da volatilidade em modelos GARCH.

Entretanto, muitas séries temporais financeiras apresentam quebras estruturais ao longo do tempo. Normalmente essas quebras são associadas a eventos políticos ou crises financeiras, como a do *subprime* em 2009, ou até mesmo em séries muito longas, devido ao fato de que as vezes não é coerente assumir que os mesmos parâmetros do modelo GARCH expliquem a série por tanto tempo. Diante disso, apesar dos modelos GARCH serem capazes de capturar diversos aspectos das séries temporais financeiras, na presença de quebras estruturais, tais modelos se tornam inapropriados. Hillebrand (2005) aponta que esse efeito adverso independe do método de estimação. Ou seja, há um erro de especificação do modelo. E, nesse cenário, Hwang e Pereira (2008) mostram que a persistência estimada da volatilidade condicional em grandes amostras pode ser excessiva, na presença de quebras estruturais nos parâmetros ARCH e GARCH.

Hamilton (1989) e Cai (1994) propõem o modelo autorregressivo heterocedástico condicional com mudança de regime markoviano (MS-ARCH). O modelo MS-ARCH pode ser descrito como um modelo ARCH onde os parâmetros dependem do estado não observável de uma Cadeia de Markov com espaço de estados discreta. Este modelo possui diferentes equações para caracterizar os comportamentos das séries temporais em diferentes regimes. Ao permitir a alternância entre essas estruturas, este modelo é capaz de capturar padrões dinâmicos mais complexos (KUAN, 2002). E uma vez que os regimes não são observáveis, na estimação por máxima-verossimilhança do modelo, precisamos inferir a probabilidade de estar no tempo em um determinado regime. Para isso, o mecanismo utilizado é o filtro de Hamilton (1989).

Ao estender essa metodologia para os modelos GARCH, foi observado que era inviável a estimação por máxima-verossimilhança devido uma vez que ela dependia de todos as combinações de caminho possíveis para ser computada. Esse problema pode ser contornado via estimação bayesiana, assim como mostram Das, Yoo et al. (2004) e Bauwens, Preminger e Rombouts (2010). No entanto, Gray (1996) foi o primeiro autor que lidou com o problema da dependência do caminho tornando possível a estimação por máxima-verossimilhança. Posteriormente, Haas, Mittnik e Paolella (2004) apresentaram uma nova abordagem aos modelos GARCH univariados com mudanças de regime markoviano com uma interpretação diferente do modelo apresentado por Gray (1996).

A utilização desses modelos GARCH como mudança de regime markoviano se tornou frequente na literatura. O motivo disso é que em diversos casos o modelo GARCH com mudança de regime markoviano tem melhor desempenho do que aqueles com um regime. Assim, Ardia et al. (2018) descobrem que os modelos GARCH com mudança de regime markoviano fornecem melhores previsões de VaR e ES comparado aos GARCH

com um único regime em diversos casos. Além disso, eles citam que as melhorias são mais acentuadas quando o mecanismo de troca de Markov é aplicado a especificações simples, como o modelo GARCH-Normal.

Porém, assim como outros filtros habituais, o filtro de Hamilton (1989) tem mal desempenho na presença de *outliers* como mostra Calvet, Czellar e Ronchetti (2015). Na prática, alguns autores observam que ao estimar o modelo, um regime tende a capturar o *outliers*. Diante disso, o objetivo desse trabalho é apresentar alternativas robustas na estimação por máxima-verossimilhança do modelo GARCH com mudança de regime markoviano. Para fazer isso, iremos ter como base os filtros robustos apresentados por Petrus (1999) e Calvet, Czellar e Ronchetti (2015). Além disso, teremos como referência de objetivos da robustez, Huber e Ronchetti (2009, Seção 1.2, pág. 5): o estimador deve ter uma eficiência razoavelmente boa no modelo assumido; deve ser robusto no sentido de que pequenos desvios das suposições do modelo devem prejudicar o desempenho apenas ligeiramente; e desvios um pouco maiores do modelo não devem causar uma catástrofe. Neste sentido, também serão consideradas outras modificações nos métodos de estimação de modelos GARCH com mudanças de regime.

Até o momento em que esse trabalho foi escrito, não foram encontrados na literatura trabalhos que lidem com a estimação robusta de modelos GARCH com mudança de regime. Diante disso, nessa dissertação iremos abordar metodologias robustas na estimação dos parâmetros desse modelo. O objetivo é analisar o desempenho de métodos robustos aplicados ao modelo GARCH e métodos robustos aplicado a filtros na estimação dos parâmetros do modelo GARCH com mudança de regime.

No Capítulo 2 será apresentado o modelo GARCH com 1 regime, bem como sua interpretação e suas propriedades. Além disso serão discutidos métodos diferentes de estimação desse modelo. Primeiramente se abordado a estimação por quase-verossimilhança e suas propriedades. Em seguida será discutida a sobre estimação robusta desse modelo, tratando com detalhes a estimação apresentado por Muler e Yohai (2008) e Boudt, Danielsson e Laurent (2013).

No Capítulo 3 serão apresentados os modelos GARCH com mudança de regime markoviano, assim como suas propriedades, interpretação e vantagens em relação ao GARCH em determinados cenários. Dentre as diferentes especificações desse modelo, utilizaremos a apresentada por Haas, Mittnik e Paoletta (2004). Para estimação desse modelo, iremos abordar, no caso não robusto, o estimador de quase-verossimilhança através do filtro de Hamilton (1989) e algumas propriedades. Além disso, iremos abordar uma alternativa de estimação robusta que consiste em adaptar as técnicas utilizadas por Boudt, Danielsson e Laurent (2013) no caso de mudança de regime.

No Capítulo 4 será apresentado os modelos DCC com mudança de regime markoviano. Nesse trabalho iremos utilizar a especificação utilizada por Lee (2010). Assim

como no caso univariado, iremos abordar, no caso não robusto, o estimador de quase-verossimilhança através do filtro de Hamilton (1989). E para o caso robusto, iremos adaptar as técnicas utilizadas por Boudt, Danielsson e Laurent (2013) no caso de mudança de regime.

No Capítulo 5 será mostrado um estudo com simulações Monte Carlo para avaliar a performance dos estimadores dos parâmetros do modelo GARCH com mudança de regime. Avaliaremos tanto a estimação via quase-verossimilhança através do filtro de Hamilton (1989) quanto alternativas robustas. Neste capítulo iremos focar na avaliação da precisão dos parâmetros e na capacidade do filtro em identificar os estados. No caso univariado, observamos que, de maneira geral, os estimadores de quase máxima verossimilhança (QML) utilizando a t-Student (QML-t) e o robusto, o filtro de Hamilton têm desempenho razoável na identificação dos regimes, diferente do QML utilizando a distribuição normal (QML-n). Já, na estimação dos parâmetros da variância condicional, notamos que os estimadores tendem a estimar bem alguns parâmetros, porém não todos.

No Capítulo 6 serão apresentadas aplicações às séries de dados reais. Iremos trabalhar com a série das moedas do Euro, Yen, Franco Suíço e da Libra, todas em relação ao dólar americano. Iremos verificar as estimativas dos parâmetros e a capacidade de previsão um passo a frente do VaR. Neste trabalho analisamos VaR 1% e 5%. Para o caso do VaR com 1% vimos que os estimadores QML normal e robusto têm desempenho melhor. Entretanto, para o VaR 5% o estimador QML-t tem performance melhor.

Por fim, no Capítulo 7 será discutido as principais conclusões de todos os resultados encontrados nas simulações e nas aplicações em relação aos estimadores considerados.



## 2 Modelos GARCH com 1 Regime

Assim como cita Francq e Zakoian (2019, Seção 1.3, pág. 6), a modelagem de séries temporais financeiras é algo complicado devido a variedade de ativos financeiros, a frequência da série observada e da existência de falta de observações. De modo geral, o interesse nesse caso é avaliar o preço ou retorno de ativos num determinado período. No entanto, diversos autores preferem utilizar retornos ao invés do preço, uma vez que o retorno de um ativo é um resumo completo e sem escala e também as séries dos retornos são mais fáceis de manusear do que as séries dos preços Tsay (2010, Seção 1.1, pág. 2). Seja  $p_t$  o preço de um ativo no tempo  $t$ , os tipos mais comuns de retornos utilizados são a variação relativa do preço,  $y_t = (p_t - p_{t-1})/p_{t-1}$  e o log-retorno,  $y_t = \log(p_t/p_{t-1})$ . Além disso, os investidores estão diretamente interessados nos lucros, isto é, retornos. Empiricamente foi observado que grande parte das séries financeiras dos retornos apresentavam alguns “fatos estilizados”, bem como descrevem Francq e Zakoian (2019, Seção 1.3, pág. 7):

- (i) Os retornos não são correlacionados serialmente, ou têm baixa correlação;
- (ii) Os retornos ao quadrado ou o valor absoluto dos retornos são geralmente autocorrelacionados e com alta persistência.
- (iii) A série dos retornos apresentam conglomerados de volatilidade. Isto é, valores grandes de  $|y_t|$  são seguidos de valores grandes e valores pequenos de  $|y_t|$  são seguidos de valores pequenos.
- (iv) A distribuição incondicional de  $y_t$  frequentemente tem cauda mais pesada que a da distribuição normal.
- (v) Valores negativos de  $y_t$  comumente causam mais impacto na volatilidade nos dias seguintes do que valores positivos com a mesma magnitude.

Nesse cenário, os modelos autorregressivos condicionalmente heterocedásticos (ARCH) e sua extensão GARCH (*Generalized ARCH*) foram introduzidos por Engle (1982) e por Bollerslev (1986), respectivamente, com intuito de explicar e prever a volatilidade em cenário econômicos e financeiros. Apesar de, nesses artigos iniciais, os modelos foram aplicados para explicar a volatilidade da série de inflação Reino Unido e dos Estados Unidos, esses modelos se expandem para diversas outras aplicações. Engle (2001) mostra aplicações destes modelos na análise de portfólios, principalmente utilizando medidas como o VaR. Um dos motivos de sua popularidade é que, como cita Francq e Zakoian (2019, Seção 2, pág. 17), o modelo GARCH é capaz de capturar os principais fatos estilizados que caracterizam as séries financeiras.

**Definição 1.** Conforme Bollerslev (1986), dizemos que  $y_t$  segue um modelo GARCH( $p, q$ ) se:

$$\begin{aligned} y_t &= \sqrt{h_t} z_t, \\ h_t &= \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \end{aligned} \quad (2.1)$$

em que  $\{z_t\}$  é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d) com média 0 e variância 1;  $\alpha_0 > 0$ ;  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$  são as condições suficientes para que  $h_t$  seja positiva. Se,  $\beta_j = 0$  para todo  $j$ , então  $y_t$  assume um modelo ARCH( $p$ ). Neste modelo temos que valores altos de  $y_{t-1}^2$  ou de  $h_{t-1}$  implicam que  $h_t$  seja alto. Isto é, valores altos da volatilidade condicional  $t - 1$  são seguidos de valores altos em  $t$ , caracterizando conglomerados de volatilidade, um dos fatos estilizados da volatilidade.

Como cita Francq e Zakoian (2019, Seção 2.3.2, pág. 39), às vezes é útil considerar a representação ARCH( $\infty$ ) de um processo GARCH( $p, q$ ). Esta representação permite que a variância condicional  $h_t$  seja escrita explicitamente como uma função de seu passado infinito. Considerando o modelo GARCH (1,1) e satisfeitas as condições de positividade,  $\alpha_0 > 0$ ;  $\alpha_1 \geq 0$ ,  $\beta_1 \geq 0$ , e considerando também que  $\beta_1 < 1$ , obtemos que:

$$h_t = \frac{\alpha_0}{1 - \beta_1} + \alpha_1 \sum_{i=1}^{\infty} \beta_1^{i-1} y_{t-i-1}^2. \quad (2.2)$$

E, assim como ressalta Haas, Mittnik e Paoletta (2004), a representação ARCH( $\infty$ ) revela o papel dos parâmetros  $\alpha_1$  e  $\beta_1$  no modelo GARCH(1,1). Sendo assim, segundo Campbell, Lo e MacKinlay (1997, Seção 12.2, pág. 483) e Haas, Mittnik e Paoletta (2004), temos que  $\alpha_1$  mede a magnitude em que um choque de volatilidade no tempo  $t$  alimenta a volatilidade do período  $t + 1$ ,  $\beta_1$  é um parâmetro de inércia e indica a memória na variância,  $\alpha_1 + \beta_1$  mede a taxa na qual esse efeito morre ao longo do tempo, ou seja, a persistência da variância e o impacto total de um choque de unidade nas variâncias futuras é  $\alpha_1(1 - \beta_1)^{-1}$ .

De acordo com Francq e Zakoian (2019, Seção 2.2.2, pág. 27), a principal ferramenta para estudar a estacionariedade estrita do modelo GARCH é o expoente de Lyapunov  $\gamma$ . Seja  $\{\mathbf{A}_t, t \in \mathbb{Z}\}$  uma sequência de matrizes aleatórias  $(p + q) \times (p + q)$  estritamente estacionária, tal que  $E[\log^+(\mathbf{A}_t)]$  é finito. Temos que o expoente de Lyapunov é dado por

$$\gamma = \lim_{t \rightarrow \infty} \frac{1}{t} E[\log \|\mathbf{A}_t \mathbf{A}_{t-1} \dots \mathbf{A}_1\|]. \quad (2.3)$$

Portanto, fazendo  $A_t$  como sendo:

$$A_t = \begin{bmatrix} \alpha_1 z_t^2 & \dots & \alpha_p z_t^2 & \beta_1 z_t^2 & \dots & \beta_p z_t^2 \\ 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots \\ 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ \alpha_1 & \dots & \alpha_p & \beta_1 & \dots & \beta_p \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad (2.4)$$

obtém-se que uma condição necessária e suficiente para a existência da estacionariedade estrita para o GARCH(p,q) é que o expoente de Lyapunov  $\gamma$  da matriz (2.4) seja menor que 0.

Porém, o mais comum é utilizar a estacionariedade de segunda ordem. Neste caso, a condição necessária e suficiente para a existência da estacionariedade de segunda ordem para o GARCH(p,q) é que  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ . E quando  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j = 1$ , o modelo é chamado de IGARCH(p,q) ou GARCH integrado.

Nessas condições, temos que a média e a variância condicional de  $y_t$ , isto é,  $E[y_t]$  e  $Var[y_t]$  são dadas por:

$$\begin{aligned} E[y_t] &= E[E(y_t|\mathcal{I}_{t-1})] = E[E(\sqrt{h_t}z_t|\mathcal{I}_{t-1})] = E[\sqrt{h_t}E(z_t|\mathcal{I}_{t-1})] = E[\sqrt{h_t}E(z_t)] = 0, \\ Var[y_t] &= E[y_t^2] = E[E(y_t^2|\mathcal{I}_{t-1})] = E[E(h_t z_t^2|\mathcal{I}_{t-1})] = E[h_t E(z_t^2|\mathcal{I}_{t-1})] = E[h_t E(z_t^2)] \\ &= E[h_t] = \alpha_0 + \sum_{i=1}^p \alpha_i E[h_t] + \sum_{j=1}^q \beta_j E[h_t] = \alpha_0 + E[h_t] \left( \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j \right) \\ &= \frac{\alpha_0}{1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j}, \end{aligned}$$

em que  $\mathcal{I}_{t-1}$  denota toda a informação até o tempo  $t - 1$ .

Analogamente, obtemos que a função de autocovariância incondicional de  $y_t$  é obtida por:

$$E[y_t y_{t+\tau}] = E[E(y_t y_{t+\tau}|\mathcal{I}_{t-1})] = \begin{cases} Var[y_t], & \text{se } \tau = 0 \\ 0, & \text{se } \tau \geq 1. \end{cases}$$

Já, para a função de autocorrelação de  $y_t^2$ , um mecanismo útil é que o processo  $y_t^2$  pode

ser escritos na forma de um ARMA( $q, \max(p,q)$ ), isto é,

$$y_t^2 = \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i)y_{t-i}^2 = \alpha_0 + (y_t^2 - h_t) - \sum_{i=1}^q \beta_i(y_{t-i}^2 - h_{t-i}).$$

Através disso, podemos obter a função de autocorrelação de  $y_t^2$ . A forma da função de autocorrelação para os modelos ARMA está disponível em Morettin e Tolo (2018, Seção 5.2.4, pág 133).

Como citam Francq e Zakoian (2019, Seção 2.4.2, pág. 45), no caso dos processos GARCH, é interessante notar a diferença entre as caudas das distribuições marginais e condicionais. Para isso, utiliza-se a curtose. Assumindo que exista o quarto momento de  $y_t$ , o excesso de curtose do modelo GARCH( $p,q$ ), quando  $z_t$  segue distribuição Normal, é dado por:

$$\kappa_{y_t} = \frac{E[y_t^4]}{\{E[y_t^2]\}^2} - 3 = \frac{6a + \kappa_{z_t}(1 + 3a)}{1 - a(\kappa_{z_t} + 2)},$$

em que  $\kappa_{z_t}$  é o excesso de curtose da distribuição normal e  $a = \sum_{i=1}^{\infty} \psi_i^2$ , tal que  $\psi_i$  é dado por:

$$1 + \sum_{i=1}^{\infty} \psi_i k^i = \left\{ 1 - \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) k^i \right\}^{-1} \left( 1 - \sum_{i=1}^p \beta_i k^i \right).$$

## 2.1 Estimadores Não Robustos

A estimação dos parâmetros dos modelos GARCH pode ser feita de várias maneiras. Baillie e Chung (2001) estimam os parâmetros de um modelo GARCH(1,1) utilizando um estimador de mínima distância para as autocorrelações do processo ao quadrado. Pérez-Cruz et al. (2003) utilizam a metodologia de máquina de suporte vetoriais para estimar os parâmetros do modelo GARCH. Aknouche e Guerbyenne (2006) desenvolvem três algoritmos online recursivos, baseados em um esquema de mínimos quadrados de dois estágios para estimar modelos GARCH. Ardia (2008, Seção 3) apresentam as metodologias para a estimativa Bayesiana de modelos GARCH. Entretanto, a mais comum delas é por máxima quase-verossimilhança (QML). Como citam Francq e Zakoian (2019, Seção 7, pág. 175), o método QML é particularmente relevante para modelos GARCH porque fornece estimadores consistentes e assintoticamente normais para processos GARCH estritamente estacionários sob condições de regularidade moderada.

### 2.1.1 Estimação por máxima quase-verossimilhança

Primeiramente, considere  $\mathbf{y}_T = \{y_1, y_2, \dots, y_T\}$ ,  $\boldsymbol{\theta}' = (\boldsymbol{\alpha}', \boldsymbol{\beta}')$ , com  $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_p)'$  e  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)'$ , de modo que ele pertença ao espaço paramétrico

de forma  $\Theta \subset (0, +\infty) \times [0, \infty)^{p+q}$  - frequentemente o espaço no qual o modelo GARCH é estacionário. Para obter a verossimilhança do modelo GARCH(p,q) definido em (2.1), é necessário especificar a distribuição de  $\{z_t\}$ . O nome quase-verossimilhança vem do fato de utilizarmos, por exemplo, quando sabemos que a distribuição não é normal. Para o caso da Normal, dados os valores iniciais, por exemplo,  $\phi_0 = \{y_0, \dots, y_{1-p}, h_0, \dots, h_{1-q}\}$ , obtemos que a quase-verossimilhança pode ser expressa como:

$$\begin{aligned} L(\boldsymbol{\theta}|\mathbf{y}_T, \phi_0) &= f(y_t|\mathcal{I}_{t-1}, \boldsymbol{\theta}, \phi_0) f(y_{t-1}|\mathcal{I}_{t-2}, \boldsymbol{\theta}, \phi_0) \dots f(y_1|\boldsymbol{\theta}, \phi_0) \\ &= \prod_{t=1}^T \frac{1}{\sqrt{2\pi h_t}} \exp\left\{-\frac{y_t^2}{2h_t}\right\}, \end{aligned} \quad (2.5)$$

em que  $\mathbf{y}_T$  representa o conjunto de observações em toda a série e  $h_t$  pode ser obtido de maneira recursiva através de (2.1). Assim, temos que a estimativa por máxima quase-verossimilhança  $\hat{\boldsymbol{\theta}}$ , quando assumimos a distribuição normal padrão (QML-n), é dada por:

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta}|\mathbf{y}_T, \phi_0). \quad (2.6)$$

No entanto, neste caso maximizar a função de quase-verossimilhança é equivalente a maximizar seu logaritmo. Consequentemente, a estimativa  $\hat{\boldsymbol{\theta}}$  de  $\boldsymbol{\theta}$  por QML-n, é usualmente obtida maximizando a log-verossimilhança:

$$\ell(\boldsymbol{\theta}|\mathbf{y}_T, \phi_0) = \log(L(\boldsymbol{\theta}|\mathbf{y}_T, \phi_0)) = -\frac{1}{2} \sum_{t=1}^T \left[ \ln(h_t) + \frac{y_t^2}{h_t} + \ln(2\pi) \right] \propto -\frac{1}{2} \sum_{t=1}^T \left[ \ln(h_t) + \frac{y_t^2}{h_t} \right]. \quad (2.7)$$

E, de acordo com Francq e Zakoian (2019, Seção 7.1.1, pág. 179), considere as condições:

- C1 -  $\boldsymbol{\theta} \in \Theta$  e  $\Theta$  é compacto;
- C2 - O modelo é estacionário em  $\boldsymbol{\theta}$ ;
- C3 -  $z_t^2$  tem distribuição não degenerada e  $E[z_t^2] = 1$ ;
- C4 - Seja  $\mathcal{A}_{\boldsymbol{\theta}}(v) = \sum_{i=1}^p \alpha_i v^i$  e  $\mathcal{B}_{\boldsymbol{\theta}}(v) = \sum_{j=1}^q \beta_j v^j$ . Se  $q > 0$ ,  $\mathcal{A}_{\boldsymbol{\theta}}(v)$  e  $\mathcal{B}_{\boldsymbol{\theta}}(v)$  têm raízes distintas,  $\mathcal{A}_{\boldsymbol{\theta}}(1) \neq 0$  e  $\alpha_{0p} + \beta_{0q} \neq 0$ .
- C5 -  $\boldsymbol{\theta} \in \Theta^\circ$ , em que  $\Theta^\circ$  denota o interior de  $\Theta$ ;
- C6 -  $E[z_t^4] < \infty$ .

Sob as condições C1-C4 temos que o estimador de quase-verossimilhança é consistente e sob C1-C6, o estimador converge em distribuição, i.e.,

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\mathcal{L}} N\left(0, (\kappa_z - 1)J^{-1}\right),$$

em que,

$$J := E \left( \frac{\partial^2 l(\boldsymbol{\theta} | \mathbf{y}_t, \phi_0)}{\partial \boldsymbol{\theta} \boldsymbol{\theta}'} \right).$$

É importante frisar que a escolha dos valores iniciais em amostras finitas pode produzir valores diferentes nas estimativas, principalmente quando o tamanho da amostra é pequeno. Pelagatti e Lisi (2009) mostram a precisão das estimativas QML-n para o modelo GARCH(1,1) sob diferentes escolhas de valores iniciais.

Outro ponto a ser citado é sobre as escolhas das ordens dos modelos GARCH. Em problemas reais, a escolha mais comum é utilizar o modelo GARCH(1,1) por questões de parcimônia e pelo fato de que ele já consegue captar a maioria dos fatos estilizados das séries temporais financeiras. No entanto, caso o modelo GARCH(1,1) não seja o único de interesse, há métodos que se destacam na identificação da ordem do modelo. Um deles é utilizar algum tipo critério, dentre esses se destacam os critérios de informação, como AIC e BIC, ou alguma função de perda, como a perda quadrática. Outra maneira é identificar a ordem do modelo é através do comportamento da função de autocorrelação e autocorrelação parcial do quadrado dos valores observados.

Complementarmente, alguns autores mostram que a presença de *outliers* afetam as estimativas dos modelos GARCH. Para os modelos GARCH, segundo Hotta e Tsay (2012), estudam dois tipos de *outlier*: aditivo e de inovação. Um *outlier* aditivo afeta apenas uma única observação, como um erro de digitação ou de gravação, por exemplo. Por outro lado, um *outlier* de inovação afeta muitas observações e, muitas vezes, significa uma perturbação externa à série que tem um impacto que diminui gradualmente.

Na presença de *outliers* aditivos, diversos autores mostram a má performance do estimador de quase-verossimilhança para os modelos GARCH. Para o modelo GARCH(1,1), Carnero, Peña e Ruiz (2007) mostram que o estimador QML não apresenta bom desempenho tanto na estimativa dos parâmetros como nos seus desvios padrões assintóticos na presença de um único *outlier* e de dois *outliers* consecutivos. Além disso, eles mostram que os valores discrepantes podem ter um forte efeito sobre o tamanho e o poder dos testes comuns de homocedasticidade condicional com base em autocorrelações de quadrados. Mostram também que, um *outlier* isolado muito grande leva a testes que nunca rejeitam a hipótese nula de homocedasticidade, independentemente de a série não contaminada ser homocedástica ou heterocedástica. Por outro lado, dois ou mais *outliers* consecutivos grandes levam a testes que sempre rejeitam a hipótese nula de homocedasticidade, mesmo que a série não contaminada seja verdadeiramente homocedástica.

## 2.2 Estimadores Robustos dos Modelos GARCH

Para lidar com *outliers* na estimação dos modelos GARCH, existem duas metodologias que são utilizadas com mais frequência: a primeira é detectar os *outliers* e tratá-los de alguma forma apropriada, a segunda é utilizar estimadores robustos a *outliers*. Em relação aos métodos de identificação, podemos destacar algumas metodologias que mostram bom desempenho para a detecção de *outliers* nos modelos GARCH. Hotta e Tsay (2012) propõem uma metodologia baseada nos multiplicadores de Lagrange para identificação de *outliers*, tanto de nível quanto de inovação, para os modelos ARCH(p) e GARCH(1,1). Crosato e Grossi (2019) propõem um teste utilizando o mecanismo para a detecção de *outliers* em modelos GARCH (1,1) unindo a um estimador, que também utiliza esse mecanismo, para corrigir automaticamente o efeito dos *outliers* na estimativa dos parâmetros do modelo.

Já para estimadores robustos, existem na literatura inúmeras metodologias diferentes para lidar com os efeitos dos *outliers* nos modelos GARCH, dando enfoque principalmente em *outliers* aditivos. Para esses estimadores, também podemos destacar alguns que apresentam boa performance na presença de *outliers*. Sakata e White (1998) propõem S-estimadores para modelos GARCH e mostraram que eles são resistentes aos efeitos de *outliers* além disso eles mostram algumas propriedades desses estimadores. Park (2002) propõe um modelo robusto usando estimativa de desvio mínimo absoluto e mostra através de simulações Monte Carlo de modelos AR(1)-GARCH(1,1) que, conforme a magnitude dos *outliers* aumenta, seu estimador robusto tem desempenho melhor na previsão um passo a frente do que o modelo GARCH estimado por QML. Lee e Song (2009) propõem um estimador de mínima divergência de potência de densidade (MDPDE)<sup>1</sup> para o modelo GARCH com inovações gaussianas, eles mostram as propriedades assintóticas do estimador e complementarmente eles apresentam através de simulações Monte Carlo do modelo GARCH(1,1) que o estimador tem performance melhor em relação à estimação por QML na presença de *outliers*. Para outros testes de *outliers* e estimadores robustos, veja, por exemplo Hotta e Trucíos (2018).

Contudo, Muler e Yohai (2008) salientam que as estimativas dos métodos robustos citados acima podem não serem robustos na presença de uma fração de *outliers* na amostra, uma vez que eles são baseados na variância condicional, que são muito sensíveis a *outliers*. Sendo assim, eles apresentaram um modelo que utilizam uma modificação dos M-Estimadores e usam uma especificação modificada da variância condicional do modelo GARCH que diminui o efeito dos *outliers*.

---

<sup>1</sup> ver Basu et al. (1998)

### 2.2.1 BIP-GARCH

Para ganhar robustez Muler e Yohai (2008) utilizam um mecanismo semelhante a filtros robustos na estimativa da variância condicional. Esse mecanismo, chamado de BIP (de *bounded innovation propagation*), restringe a propagação do efeito *outliers* presente no tempo  $t - 1$  na variância condicional do tempo  $t$ . Assim, para o modelo GARCH(p,q), temos que  $h_t$  é calculado como:

$$h_t^* = \alpha_0 + \sum_{i=1}^p \alpha_i h_{t-i}^* r_k \left( \frac{y_{t-i}^2}{h_{t-i}^*} \right) + \sum_{i=1}^q \beta_j h_{t-i}^*, \quad \text{em que} \quad (2.8)$$

$$r_k(u) = \begin{cases} u, & \text{se } u \leq k \\ k, & \text{se } u > k. \end{cases}$$

No entanto, Boudt, Danielsson e Laurent (2013) observaram em suas simulações que ao utilizar a especificação do BIP em (2.8) as estimativas dos parâmetros do modelo GARCH apresentavam pequenos vieses. Assim eles propõem uma modificação para garantir que a esperança condicional à  $\mathcal{I}_{t-2}$  de  $c_{k_{\delta,N}} h_{t-1}^{(k)} r_{k_{\delta,N}}(y_{t-1}^2/h_{t-1}^{(k)})$  seja igual a esperança condicional de  $y_{t-1}^2$ . A novidade está na introdução do fator de correção  $c_{k_{\delta,1}}$ . A variância condicional passa a ser dada como:

$$h_t^* = \alpha_0 + \sum_{i=1}^p \alpha_i h_{t-i}^* c_{k_{\delta,1}} r_{k_{\delta,1}} \left( \frac{y_{t-i}^2}{h_{t-i}^*} \right) + \sum_{i=1}^q \beta_j h_{t-i}^*, \quad \text{em que} \quad (2.9)$$

$$r_{k_{\delta,1}}(u) = \begin{cases} u, & \text{se } u \leq k_{\delta,1} \\ k_{\delta,1}, & \text{se } u > k_{\delta,1}. \end{cases}$$

em que

$$c_{k_{\delta,1}} = \frac{E[W]}{E[r_{k_{\delta,1}}(W)W]} = \frac{1}{F_{\chi_{N+2}^2}(\chi_N^2(\delta)) + (1 - \delta)\chi_N^2(\delta)},$$

tal que  $W$  é uma variável aleatória qui-quadrado com  $N$  graus de liberdade,  $F_{\chi_{N+2}^2}(\cdot)$  é a função de distribuição acumulada da distribuição qui-quadrado com  $N + 2$  graus de liberdade,  $\chi_N^2(\delta)$  é valor corresponde ao quantil  $\delta$  da distribuição qui-quadrado com  $N$  graus de liberdade e  $k_{\delta,1}$  é o quantil  $\delta$  da distribuição qui-quadrado com um grau de liberdade. A notação  $k_{\delta,1}$  foi introduzida para ficar bem claro o significado do limiar, a partir do qual suspeitamos que o valor observado está afetado por um *outlier*. A tabela 1 do artigo de Boudt, Danielsson e Laurent (2013) apresenta os valores de  $c_{k_{\delta,1}}$  para diferentes valores de  $N$  e  $\delta$ . Como estamos no caso univariado, temos que  $N = 1$ . Porém, posteriormente utilizaremos a mesma notação para o caso multivariado.

Tanto Boudt, Danielsson e Laurent (2013) e Muler e Yohai (2008) sugerem nas equações (2.8) e (2.9) substituir pelo próprio valor  $k$  quando  $u > k$ . No entanto, outros autores propõem outros valores como, por exemplo, Carnero, Peña e Ruiz (2012)



que utilizaram 1. Isto é, quando tiver uma observação atípica, ela é substituída por  $E[y_{t-1}^2] = h_{t-1}$ . Para os casos em que se utiliza outro valor para  $k_{\delta,1}$  em (2.9), quando  $u > k_{\delta,1}$ , digamos  $k_{\delta^*,1}$ , temos que o valor de  $c_{k_{\delta,1}}$  é dado por:

$$c_{k_{\delta,1}} = \frac{1}{F_{W_3}(k_{\delta,1}) + k_{\delta^*,1}(1 - F_{W_1}(k_{\delta,1}))},$$

em que  $W_N$  é uma variável aleatória qui-quadrado com  $N$  graus de liberdade. Para o caso de Carnero, Peña e Ruiz (2012) que substituem por 1, temos que  $\delta^* = 0,6826895$ .

Daí, temos que o parâmetro  $\delta$  funciona como um *trade-off* de robustez e eficiência, se o valor escolhido for muito alto  $h_t^*$  tende a coincidir com  $h_t$ , e para valores pequenos  $h_t^*$  se torna mais robusto a *outliers*. Assim, se  $y_t$  segue um modelo GARCH, mas contém alguns *outliers*, podemos esperar que as estimativas dos parâmetros utilizando a forma da variância condicional dada por (2.9) se ajustem melhor do que a forma usual (2.1). E, quando  $z_t$  segue uma distribuição normal,  $y_t^2/h_t$  segue uma distribuição qui-quadrado com 1 grau de liberdade. Assim, é natural detectar uma ocorrência de *outlier* em  $y_t^2$ , se  $y_t^2/h_t$  exceder o valor  $k_{\delta,1}$ , o qual representa o quantil  $\delta$  da distribuição qui-quadrado com um grau de liberdade (BOUDT; DANIELSSON; LAURENT, 2013). Ou seja, a fórmula para o valor de correção  $c_{k_{\delta,1}}$  e os valores tabelados são para o caso em que a distribuição dos  $z_t$  em (2.1) tem distribuição normal.

Boudt, Danielsson e Laurent (2013) propõem utilizar *variance target*, isto é,  $\alpha_0$  é reparametrizado como  $\hat{h}(1 - \alpha_1 - \beta_1)$  - com  $\hat{h}$  sendo um estimador consistente da variância incondicional  $h$  e estimado antes dos demais parâmetros. Sem *outliers*  $\hat{h}$  pode ser obtido simplesmente pela variância amostral, porém esse estimador é extremamente afetado na presença de *outliers*. Assim, o ideal nesse cenário é utilizar um estimador robusto a *outliers* para a variância incondicional. Para isso, Boudt, Danielsson e Laurent (2013) utilizam um estimador chamado *reweighted variance estimator*.

O *reweighted variance estimator*, quando  $z_t \sim N(0, 1)$ , é dado por:

$$\hat{h} = 1,318 \frac{\sum_{i=1}^T y_i J_i}{\sum_{i=1}^T J_i},$$

em que

$$J_t = I \left[ \frac{y_t^2}{\text{MAD}_t^2} \leq \chi_1^2(95\%) \right],$$

tal que  $I$  é a função indicadora e o MAD de uma sequência de observações  $y_1, \dots, y_T$  é definido como  $1,486 \times \text{mediana}_i(|y_i - \text{mediana}_j(y_j)|)$ , em que 1,486 é um fator de correção para garantir que o MAD é um estimador de escala consistente para a distribuição normal.

Por fim, Boudt, Danielsson e Laurent (2013) estimam os parâmetros do modelo BIP-GARCH através de um estimador que minimiza uma função objetivo  $\rho$  avaliada no logaritmo das observações padronizadas ao quadrado. Isto é,

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \frac{1}{T} \sum_{t=1}^T \rho \left( \ln \frac{y_t^2}{h_t} \right). \quad (2.10)$$

Com base na comparação de várias funções  $\rho$  candidatas, Boudt, Danielsson e Laurent (2013) recomendam uma associada à densidade da t-Student com 4 graus de liberdade. Para o caso mais geral, associada à densidade da t-Student com  $\nu$  graus de liberdade temos

$$\rho(z) = -z + \sigma_{1,4} \rho_{t_{1,\nu}}(\exp(z)), \quad (2.11)$$

em que

$$\rho_{t_{1,\nu}}(u) = (1 + \nu) \log \left( 1 + \frac{u}{\nu - 2} \right) \quad (2.12)$$

e

$$\sigma_{t_{N,\nu}} = \frac{N}{E[\rho'_{t_{N,\nu}}(W)W]}, \quad (2.13)$$

tal que  $W$  é uma variável aleatória qui-quadrada com  $N$  grau de liberdade.  $N$  é a dimensão do vetor de retornos. No caso do modelo GARCH temos  $N = 1$ . Apresentamos na forma geral porque as funções definidas em (2.12) e (2.13) serão utilizadas mais tarde na estimação de modelos DCC  $N$ -dimensionais.

## 2.3 Previsão das Volatilidades

A principal utilidade dos modelos GARCH e suas variações é prever volatilidades futuras para aplicações financeiras, como gerenciamento de riscos, como valor em risco e perda esperada, betas variáveis no tempo, índice de Sharpe condicionais, alocação de ativos com covariâncias que variam no tempo, avaliação de opções com volatilidade dinâmica etc. Por questões de simplicidade e notação, aqui iremos considerar o modelo GARCH(1,1) (2.1) para ilustrar como é dada previsão da volatilidade  $\tau$  passos a frente. Suponha que o modelo tenha sido estimado ao longo do período de tempo  $t = 1, 2, \dots, T$ . A previsão da volatilidade no tempo  $T$ , dada a informação até este tempo,  $\tau$  passos a frente, denotada por  $h_{T+\tau|T}$ , com menor erro médio quadrático, é dada por  $E[h_{T+\tau}|\mathcal{I}_T]$  e pode ser calculado como:

$$h_{T+\tau|T} = \alpha_0 + \alpha_1 y_{T+\tau-1|T}^2 + \beta_1 h_{T+\tau-1|T} \quad (2.14)$$

em que, analogamente,  $y_{T+\tau|T}^2$  é a previsão de  $y_t^2$   $\tau$ -passos a frente, condicionada a  $\mathcal{I}_T$ ; ou seja,  $y_{T+\tau|T}^2 \equiv E[y_{T+\tau}^2 | \mathcal{I}_T]$ .  $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1)'$  pode ser estimado por (2.6). Assim como na estimação dos parâmetros, a presença de *outliers* afeta a previsão da volatilidade. Para lidar com isso, uma alternativa para a previsão  $\tau$  passos a frente da variância condicional é apresentada por Boudt, Danielsson e Laurent (2013), que é dada por:

$$h_{T+\tau|T} = \alpha_0 + \alpha_1 c_{k_{\delta,1}} h_{T+\tau-1|T} r_{k_{\delta,1}}(y_{T+\tau-1|T}^2 / h_{T+\tau-1|T}) + \beta_1 h_{T+\tau-1|T}, \quad (2.15)$$

em que  $r_{k_{\delta,1}}(\cdot)$  é a função definida em (2.9). E, como cita Boudt, Danielsson e Laurent (2013), na presença de *outliers*,  $y_t^2$  não é observado, e é substituído por  $h_t$  quando se utiliza o mecanismo de  $r_{k_{\delta,1}}$ . Assim, como as previsões das volatilidades são baseadas em expectativas de retornos quadrados futuros sob a suposição de um modelo sem *outliers*, eles afetam as previsões não apenas por meio de um viés potencial na estimativa dos parâmetros, mas também porque os retornos defasados são usados para prever as variações futuras. O efeito desses valores discrepantes nas variações futuras é ilimitado em (2.14), mas limitado em (2.15). Adicionalmente, Carnero, Peña e Ruiz (2012) observam que, na presença de *outliers*, estimadores não robustos dos parâmetros do modelo GARCH geralmente levam a vieses nas estimativas das volatilidades.

### 3 Modelos GARCH Univariados com Mudança de Regime Markoviano

Os modelos GARCH são um dos mais populares nas aplicações de séries temporais financeiras. Isto porque eles são capazes de captar com alguns dos fatos estilizados dessas séries, como a não correlação dos retornos, conglomerados de volatilidade e excesso de curtose em relação a distribuição normal. O tamanho da amostra tem grande influência na estimação desses modelos em problemas empíricos, dado que diversas evidências empíricas sugerem que os comportamentos das variáveis econômicas e financeiras nas séries temporais podem apresentar padrões diferentes ao longo do tempo. Por exemplo, Lamoureux e Lastrapes (1990), Horváth, Kokoszka e Zhang (2006) e outros autores exibem indicativos empíricos de mudanças estruturais para retornos de ações. Nesse cenário, assim como cita Gray (1996), muitos modelos populares produzem resultados ruins quando ajustados aos dados.

Lamoureux e Lastrapes (1990), Hillebrand (2005), Hwang e Pereira (2008) e outros autores mostram que, na presença de quebras estruturais nos parâmetros dos processos ARCH e GARCH, ao ajustar o modelo GARCH, a persistência estimada, mesmo em grandes amostras, pode ser excessiva. Além disso, Hillebrand (2005) afirma que esse efeito negativo independe do método de estimação do modelo GARCH. Adicionalmente, alguns autores observaram este fato de maneira empírica, como Mikosch e Stărică (2004), que estimam um modelo GARCH em uma amostra que exhibe mudanças estruturais em sua variância condicional e obteve persistência praticamente igual a um (BILLIO; CASARIN; OSUNTUYI, 2016).

Para lidar com essa situação, Cai (1994) e Hamilton e Susmel (1994) propõem um modelo ARCH univariado com mudança de regime markoviano (MS-ARCH). A ideia da abordagem MS para modelar heterocedasticidade é que a distribuição do processo ARCH depende de um estado não observado. O modelo MS-ARCH apresentado por Cai (1994), por exemplo, é dado por:

$$y_t = \sqrt{h_t} z_t, \quad \text{em que}$$

$$h_t = \alpha_0^{(\Delta_t)} + \sum_{i=1}^p \alpha_i y_{t-i}^2,$$

em que  $\Delta_t$  uma cadeia de Markov ergódica homogênea em um conjunto finito  $S = \{1, \dots, K\}$ .

Entretanto, eles observam que, nessa especificação, a estimação por verossi-

milhança para o modelo GARCH não era factível devido à dependência do caminho do modelo, uma vez que o cálculo da função de máxima verossimilhança para uma amostra de tamanho  $T$  requer uma integração sobre todos as  $K^T$  possíveis combinações até a  $T$ -ésima observação, o que é inviável na prática.

Gray (1996) foi o primeiro autor que lidou com o problema da dependência do caminho. Primeiro, ele observou que a densidade condicional de  $y_t$  é uma mistura de normais com peso de mistura ( $p_{t-1}(\Delta_t = j)$ ) variando ao longo do tempo, isto é

$$f(y_t|\mathcal{I}_{t-1}) = \sum_{k=1}^K p_{t-1}(\Delta_t = j) f_N(y_t|\mathcal{I}_{t-1}, \Delta = k)$$

em que  $f_N(\cdot)$  representa a densidade de uma distribuição Normal. Com a finalidade de facilitar a notação, considere que  $h_t^{(k)}$  seja a variância condicional quando  $\Delta_t = k$ , ou seja,

$$h_t^{(k)} = \alpha_0^{(k)} + \alpha_1^{(k)} y_{t-1}^2 + \beta_1^{(k)} h_{t-1}^{(k)}. \quad (3.1)$$

Daí, tem-se que a variância condicional de  $y_{t-1}$  dado  $\mathcal{I}_{t-2}$  é dada por  $h_{t-1} = \sum_{k=1}^K p_{t-2}(\Delta_{t-1} = k) h_{t-1}^{(k)}$ . Assim, Gray (1996) substituiu a expressão de  $h_{t-1}$  em (3.1). Daí, a variância condicional específica de cada regime é dada por:

$$h_t^{(k)} = \alpha_0^{(k)} + \alpha_1^{(k)} y_{t-1}^2 + \beta_1^{(k)} h_{t-1}. \quad (3.2)$$

Note que, substituindo  $h_{t-1}$  em (3.2), a expressão  $h_{t-1}$  não depende do caminho, sendo assim, a estimação por verossimilhança é tratável. Abramson e Cohen (2007) mostram alguns resultados inferenciais para esse modelo.

No entanto, Haas, Mittnik e Paoletta (2004) ressaltam que a interpretação do modelo proposto por Gray (1996) é problemática. Tal como eles mencionam, quando especificamos um modelo com mudança de regime, estamos interessados em captar a dinâmica na variância condicional nos períodos de alta e baixa volatilidade ao longo da série. Assim sendo, temos interesse na relação dos parâmetros do modelo GARCH correspondente a variância do seu regime  $h_t^{(j)}$ , semelhantemente ao modelo GARCH com um único regime descrito na Seção 2. Contudo, usando a variância condicional (3.2),  $\beta_1^{(k)}$  não pode ser interpretado como sendo a inércia de  $h_t^{(k)}$ , visto que  $h_t$  é definido como sendo a média das variâncias condicionais de todos os  $K$  regimes em relação às probabilidades de estar no regime  $k$ . E, desde que  $\beta \neq 0$ ,  $h_t^{(k)}$  será afetada por choques conforme  $h_t$  varia mesmo que  $\alpha_1^{(k)}$  seja igual a zero.

Diante disso, Haas, Mittnik e Paoletta (2004) apresentam uma nova abordagem aos modelos GARCH univariados com mudança de regime markoviano no qual possui uma interpretação clara em relação a variância condicional do regime e possui uma estimação exequível e que será apresentada na subseção 3.1.

Uma outra alternativa para o modelo MS-GARCH(1,1) é apresentada por Francq, Roussignol e Zakoian (2001) que segue a seguinte especificação:

$$y_t = \sqrt{h_t} z_t, \quad \text{em que} \quad (3.3)$$

$$h_t = \alpha_{0,\Delta_t} + \alpha_{1,\Delta_t} y_{t-1}^2 + \beta_{1,\Delta_t} h_{t-1}, \quad (3.4)$$

em que  $\alpha_{0,\Delta_t}$ ,  $\alpha_{1,\Delta_t}$ ,  $\beta_{1,\Delta_t}$  são os parâmetros do processo GARCH(1,1) definido em (2.1) do regime associados ao regime  $\Delta_t$ .  $\alpha_{0,\Delta_t} > 0$ ,  $\alpha_{0,\Delta_t} \geq 0$ ,  $\alpha_{0,\Delta_t} \geq 0$  são condições para que  $h_{t,\Delta_t}$  seja positivo e  $\{z\}_t$  é uma sequência de variáveis aleatórias i.i.d. com média 0 e variância 1. Bauwens, Preminger e Rombouts (2010) avaliam aspectos inferenciais para o modelo (3.4) e mostram que, se  $\sum_{k=1}^K \pi_k E[\log(\alpha_1^{(k)} z_t^2 + \beta_1^{(k)})] < 0$ , então o processo é estacionário. Complementarmente, eles apresentam a estimação bayesiana desse modelo.

Apesar das diferentes especificações do modelo GARCH com mudança de regime, nesse trabalho iremos utilizar o modelo proposto por Haas, Mittnik e Paoletta (2004) apresentado na subseção 3.1. Além disso, aqui estaremos interessados apenas na estimação por verossimilhança desse modelo.

### 3.1 Modelo de Haas, Mittnik e Paoletta (2004)

**Definição 2.** *O modelo MS(K)-GARCH(1,1) proposto por Haas, Mittnik e Paoletta (2004) pode ser expresso da forma:*

$$y_t = \sqrt{h_t^{(\Delta_t)}} z_t, \quad (3.5)$$

$$h_t^{(k)} = \alpha_0^{(k)} + \alpha_1^{(k)} y_{t-1}^2 + \beta_1^{(k)} h_{t-1}^{(k)}, \quad (3.6)$$

em que  $h_t^{(k)}$  são as variâncias condicionais do regime  $k$ . Note que, nessa definição, o termo que acompanha o parâmetro  $\beta_1^{(k)}$ , ou seja,  $h_{t-1}^{(k)}$ , é o que diferencia este modelo do proposto por Gray (1996).  $\alpha_0^{(k)} > 0$ ,  $\alpha_1^{(k)} \geq 0$ ,  $\beta_1^{(k)} \geq 0$  para todo  $k = 1, \dots, K$  são as condições suficientes para que  $h_t^{(k)}$  seja positiva.  $\{z_t\}$  é uma sequência de variáveis aleatórias i.i.d. com média 0 e variância 1, comumente escolhida a  $N(0,1)$ . E, por fim,  $\Delta_t$  é uma cadeia de Markov com espaço de estados finitos  $\mathcal{S} = \{1, \dots, K\}$  e uma matriz de transição  $K \times K$   $\mathbf{P}$  dada por:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1K} \\ p_{21} & p_{22} & \dots & p_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{K1} & p_{K2} & \dots & p_{KK} \end{bmatrix}, \quad (3.7)$$

no qual,  $p_{ij} = P(\Delta_t = j | \Delta_{t-1} = i)$ , a soma de cada linha é igual a 1 e todos os elementos da matriz  $p_{ij}$  são maiores que zero. Assumindo que a cadeia de Markov seja irredutível e aperiódica, obtemos que distribuição estacionária  $\boldsymbol{\pi}_\infty = (\pi_\infty^{(1)}, \pi_\infty^{(2)}, \dots, \pi_\infty^{(k)})'$  pode ser obtida por  $\boldsymbol{\pi}_\infty = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{v}$ , em que

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_k - \mathbf{P} \\ \mathbf{1}'_k \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} \mathbf{0}_k \\ \mathbf{1}_k \end{bmatrix}, \quad (3.8)$$

$\mathbf{I}_k$  é a matriz  $k \times k$  identidade, e  $\mathbf{0}_k$  e  $\mathbf{1}_k$  são vetores  $k \times 1$  de zeros e uns, respectivamente.

Note que, nesse modelo, a variância condicional  $h_t^{(k)}$  depende apenas dos respectivos parâmetros do seu regime  $k$ , em contraste a dada por Gray (1996) em (3.2). Portanto, caso  $\max\{\beta_1^{(1)}, \dots, \beta_1^{(K)}\} < 1$ , similarmente aos modelos GARCH com um único regime, temos que  $\alpha_1^{(k)}$  representa o impacto de um choque em  $t-1$  em  $h_t^{(k)}$ ;  $\beta_1^{(k)}$  representa a memória em  $h_t^{(k)}$ ; e  $\alpha_1^{(k)}(1 - \beta_1^{(k)})^{-1}$  reflete o impacto total de um choque de unidade nas variações futuras do regime  $k$ .

Para estudar a estacionariedade de segunda ordem do modelo definido em (3.3), primeiro, considere a matriz  $\mathbf{M}$  de ordem  $K^2 \times K^2$ :

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{21} & \dots & \mathbf{M}_{k1} \\ \mathbf{M}_{12} & \mathbf{M}_{22} & \dots & \mathbf{M}_{k2} \\ \vdots & \vdots & & \vdots \\ \mathbf{M}_{1k} & \mathbf{M}_{2k} & \dots & \mathbf{M}_{kk} \end{bmatrix}, \quad (3.9)$$

em que  $\mathbf{M}_{ij} = p_{ij}(\boldsymbol{\beta} + \boldsymbol{\alpha}_1 \mathbf{e}'_k)$  são matrizes  $K \times K$ ,  $\boldsymbol{\beta} = \text{diag}(\beta_1^{(1)}, \beta_1^{(2)}, \dots, \beta_1^{(K)})'$ ,  $\boldsymbol{\alpha}_1 = (\alpha_1^{(1)}, \alpha_1^{(2)}, \dots, \alpha_1^{(K)})'$ ,  $p_{ij}$  são as probabilidades de transição definida em (3.7) e  $\mathbf{e}'_k$  é um vetor  $K \times 1$  que contém zeros, exceto no  $k$ -ésimo elemento. De acordo com Haas, Mittnik e Paoletta (2004), o processo (3.3) é fracamente estacionário se, e somente se,  $\rho(\mathbf{M}) < 1$ , em que  $\rho(\cdot)$  é o maior autovalor em módulo da matriz  $\mathbf{M}$ . Similarmente ao modelo GARCH com 1 regime, a condição de estacionariedade estrita é complicada. No entanto, Liu (2006) apresenta as condições para a estacionariedade estrita. Além disso, Haas, Mittnik e Paoletta (2004) citam que uma condição necessária para a estacionariedade fraca do processo é que, para todo  $k$ ,  $\beta_1^{(k)} < 1$ . No entanto, tais condições são referentes ao modelo MS(K)-GARCH(1,1), para o modelo MS(K)-GARCH(p,q) Abramson e Cohen (2007) apresentam as condições de estacionariedade estrita assim como a de segunda ordem.

Nessas condições de estacionariedade, temos que a média e variância incondicional de  $y_t$  são dadas por,

$$E[y_t] = 0;$$

$$Var[y_t] = E[y_t^2] = \sum_{j=1}^k \pi_j E[y_t^2 | \Delta_t = j] = (\text{vec } \mathbf{P})' (\mathbf{I}_{k^2} - \mathbf{M})^{-1} (\boldsymbol{\pi}_\infty \otimes \boldsymbol{\alpha}_0),$$

em que  $\boldsymbol{\alpha}_0 = (\alpha_0^{(1)}, \alpha_0^{(2)}, \dots, \alpha_0^{(K)})'$ .

Para apresentar a condição para a existência da curtose do processo definido em (3.3), defina a matriz  $K^3 \times K^3$   $\mathbf{Q}$  que contém as matrizes  $K^2 \times K^2$   $\mathbf{Q}_{ij}$  tal que:

$$\mathbf{Q}_{ji} = p_{ji}[3(\boldsymbol{\alpha}_1 \otimes \boldsymbol{\alpha}_1)\text{vec}[\text{diag}(\mathbf{e}_i)]' + (\boldsymbol{\alpha}_1 \mathbf{e}_i') \otimes \boldsymbol{\beta} + \boldsymbol{\beta} \otimes (\boldsymbol{\alpha}_1 \mathbf{e}_i') + \boldsymbol{\beta} \otimes \boldsymbol{\beta}].$$

A condição para a existência do quarto momento incondicional do processo  $y_t$  é dada por  $\rho(\mathbf{Q}) < 1$ . E satisfeita essa condição, obtém-se que o excesso de curtose do processo é dado por:

$$\kappa_{y_t} = \frac{E[y_t^4]}{\{E[y_t^2]\}^2} - 3 = \frac{3\boldsymbol{\omega}'\boldsymbol{\eta}}{[(\text{vec } \mathbf{P})'(I_{k^2} - \mathbf{M})^{-1}(\boldsymbol{\pi}_\infty \otimes \boldsymbol{\alpha}_0)]^2} - 3.$$

em que  $\boldsymbol{\omega} = [\text{vec}(\text{diag}(p_1))', \dots, \text{vec}(\text{diag}(p_k))']'$ , tal que  $p_i$  é a  $i$ -ésima coluna da matriz  $\mathbf{P}$  (3.7); e  $\boldsymbol{\eta} = (I_{k^3} - \mathbf{Q})^{-1}[\mathbf{A}_\infty^2 + \mathbf{R}(I_{k^2} - \mathbf{M})^{-1}\mathbf{A}_\infty^1]$ , de modo que  $\mathbf{R}$  é uma matriz  $K^3 \times K^2$  com que possui como elementos matrizes  $K^2 \times K$   $\mathbf{R}_{ji} = p_{ji}[(\boldsymbol{\alpha}_1 \mathbf{e}_i') \otimes \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_0 \otimes (\boldsymbol{\alpha}_1 \mathbf{e}_i') + \boldsymbol{\alpha}_0 \otimes \boldsymbol{\beta} + \boldsymbol{\beta} \otimes \boldsymbol{\alpha}_0]$  e

$$\mathbf{A}_t = \begin{bmatrix} \mathbf{A}_t^1 \\ \mathbf{A}_t^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\pi}_t \otimes \boldsymbol{\alpha}_0 \\ \boldsymbol{\pi}_t \otimes (\boldsymbol{\alpha}_0 \otimes \boldsymbol{\alpha}_0) \end{bmatrix},$$

em que  $\boldsymbol{\pi}_t = (\pi_t^{(1)}, \dots, \pi_t^{(K)})'$  é a distribuição de probabilidade sob o espaço de estados  $\mathcal{S}$  de  $\Delta_t$ .

Supondo que  $\rho(\mathbf{Q}) < 1$ , quando  $K = 2$ , temos que a função de autocovariância do processo (3.3) ao quadrado  $y_t^2$  para  $l \geq 1$  é dada por:

$$\text{cov}(y_t^2, y_{t-l}^2) = (\text{vec } \mathbf{P})'(\mathbf{M}^{l-1}\boldsymbol{\Theta}\mathbf{P}^*\boldsymbol{\eta} + (\mathbf{I}_4 - \mathbf{M})^{-1}(\mathbf{I}_4 - \mathbf{M}^l)[\mathbf{C}_1\tilde{\boldsymbol{\sigma}}^{(2)} \otimes \boldsymbol{\alpha}_0] + (\delta\mathbf{I}_4 - \mathbf{M})^{-1}[(\mathbf{C}_2\tilde{\boldsymbol{\sigma}}^{(2)}) \otimes \boldsymbol{\alpha}_0]),$$

em que  $\delta = p_{11} + p_{22} - 1$  é o grau de persistência devido aos efeitos da cadeia de Markov;  $\tilde{\boldsymbol{\sigma}}^{(2)} = [\text{diag}(p_1), \dots, [\text{diag}(p_2)](\mathbf{I}_{K^2} - \mathbf{M})^{-1}\mathbf{A}_\infty^{-1}]$ ,

$$\boldsymbol{\Theta} = \begin{bmatrix} 3\boldsymbol{\alpha}_1 \mathbf{e}_1' + \boldsymbol{\beta} & \mathbf{0}_{K \times K} & \dots & \mathbf{0}_{K \times K} \\ \mathbf{0}_{K \times K} & 3\boldsymbol{\alpha}_1 \mathbf{e}_2' & \dots & \mathbf{0}_{K \times K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{K \times K} & \mathbf{0}_{K \times K} & \dots & 3\boldsymbol{\alpha}_1 \mathbf{e}_k' \end{bmatrix}, \quad \mathbf{C}_1 = \begin{bmatrix} \pi_\infty^1 & \pi_\infty^1 \\ \pi_\infty^2 & \pi_\infty^2 \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} \pi_\infty^2 & -\pi_\infty^1 \\ -\pi_\infty^2 & \pi_\infty^1 \end{bmatrix}$$

$$\text{e } \mathbf{P}^* = \begin{bmatrix} p_{11}\mathbf{I}_K & \mathbf{0}_{K \times K} & \dots & \mathbf{0}_{K \times K} & \dots & p_{k1}\mathbf{I}_K & \mathbf{0}_{K \times K} & \dots & \mathbf{0}_{K \times K} \\ \mathbf{0}_{K \times K} & p_{12}\mathbf{I}_K & \dots & \mathbf{0}_{K \times K} & \dots & \mathbf{0}_{K \times K} & p_{k1}\mathbf{I}_K & \dots & \mathbf{0}_{K \times K} \\ \vdots & \vdots & \ddots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{K \times K} & \mathbf{0}_{K \times K} & \dots & p_{1k}\mathbf{I}_K & \dots & \mathbf{0}_{K \times K} & \mathbf{0}_{K \times K} & \dots & p_{kk}\mathbf{I}_K \end{bmatrix}.$$



Podemos observar que a função de autocovariância depende das potências de  $\mathbf{M}$  e  $\delta$ , isto é, a função de autocovariância é controlada pelo decaimento exponencial de taxa  $\max\{\rho(\mathbf{M}), \delta\}$ .

## 3.2 Estimadores Não Robustos

Assim como nos modelos GARCH com um regime, existem diferentes maneiras de estimar os parâmetros do modelo MS-GARCH. Como citado, a estimação por máxima verossimilhança do modelo definido em (3.3) e (3.4) é difícil na prática devido a dependência de caminho. Uma alternativa para lidar com esse problema é a estimação Bayesiana como apresentado por Das, Yoo et al. (2004) e Bauwens, Preminger e Rombouts (2010) entre outros. Outro método de estimação para esse modelo foi apresentada por Francq e Zakoïan (2008), no qual eles estimam os parâmetros via método dos momentos generalizados usando as autocovariâncias das potências do processo ao quadrado. Todavia, aqui estamos interessados na estimação por máxima verossimilhança, mais especificamente no modelo definido por Haas, Mittnik e Paoletta (2004). Para isso, temos como objetivo maximizar:

$$L(\boldsymbol{\theta}|\mathbf{y}_T) = \prod_{t=1}^T f(y_t|\mathcal{I}_{t-1}, \boldsymbol{\theta}) = \prod_{t=1}^T \sum_{k=1}^K [f(y_t|\Delta_t = k, \mathcal{I}_{t-1}, \boldsymbol{\theta})P(\Delta_t = k|\mathcal{I}_{t-1})], \quad (3.10)$$

em que  $\boldsymbol{\theta} = (\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, p_{11}, \dots, p_{kk})'$ ,  $\mathcal{I}_{t-1}$  é o conjunto de todas informações disponíveis até o tempo  $t - 1$  e  $f(\cdot)$  é a densidade condicional de  $y_t$ . No entanto, como o regime é não observado, precisamos inferir a probabilidade de estar no tempo  $t$  em determinado regime  $P(\Delta_t = k|\mathcal{I}_{t-1})$ . Para isso, o mecanismo aqui utilizado é o filtro de Hamilton (1989), que propõe um filtro iterativo não linear para estimar probabilidades dos estados de um modelo autorregressivo com mudança de regime markoviano. Sendo assim, através do filtro podemos obter a estimativa dos parâmetros do modelo MS-GARCH pelo método de máxima quase-verossimilhança como veremos a seguir.

### 3.2.1 Estimação via filtro de Hamilton

Defina  $\boldsymbol{\xi}_t$  como sendo o vetor de probabilidades condicionais de estar em cada regime no tempo  $t$ , isto é  $(P(\Delta_t = 1|\mathcal{I}_{t-1}), \dots, P(\Delta_t = K|\mathcal{I}_{t-1}))'$  e  $\mathbf{h}_t$  como sendo o vetor das variâncias condicionais de cada regime, ou seja,  $(h_t^{(1)}, \dots, h_t^{(K)})'$ . Então, dado os valores iniciais  $(y_0, \mathbf{h}_0, \boldsymbol{\xi}_0)'$  (ou  $(y_1, \mathbf{h}_1, \boldsymbol{\xi}_1)'$ ), a estimação de  $\boldsymbol{\xi}_t$  pelo filtro de Hamilton procede da seguinte maneira:

- 1) Dado o vetor de probabilidades condicionais  $\boldsymbol{\xi}_{t-1}$  de estar em cada regime no tempo

$t - 1$ :

$$\boldsymbol{\xi}_{t-1} = \begin{bmatrix} P(\Delta_{t-1} = 1 | \mathcal{I}_{t-1}) \\ P(\Delta_{t-1} = 2 | \mathcal{I}_{t-1}) \\ \vdots \\ P(\Delta_{t-1} = k | \mathcal{I}_{t-1}) \end{bmatrix}. \quad (3.11)$$

2) Calcular as variâncias condicionais de cada regime:

$$h_t^{(k)} = \alpha_0^{(k)} + \alpha_1^{(k)} y_{t-1}^2 + \beta_1^{(k)} h_{t-1}^{(k)}. \quad (3.12)$$

Aqui, sob suposição de normalidade, calcular o vetor das densidades condicionais de cada regime.

$$\boldsymbol{\eta}_t = \begin{bmatrix} f(y_t | \Delta_t = 1, \mathcal{I}_{t-1}) \\ f(y_t | \Delta_t = 2, \mathcal{I}_{t-1}) \\ \vdots \\ f(y_t | \Delta_t = k, \mathcal{I}_{t-1}) \end{bmatrix} = \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \\ \vdots \\ \eta_{kt} \end{bmatrix} = \frac{1}{\sqrt{2\pi}} \begin{bmatrix} 1/h_t^{(1)} \exp \left\{ -y_t^2 / (2h_t^{(1)}) \right\} \\ 1/h_t^{(2)} \exp \left\{ -y_t^2 / (2h_t^{(2)}) \right\} \\ \vdots \\ 1/h_t^{(k)} \exp \left\{ -y_t^2 / (2h_t^{(k)}) \right\} \end{bmatrix}. \quad (3.13)$$

No entanto, apesar de a maioria das aplicações utilizar inovações com distribuição normal nos regimes, vale salientar que a suposição de normalidade não é a única opção. Em algumas aplicações, alguns autores sugerem o uso outras distribuições, como a t-Student (PAOLELLA et al., 2012). E, se a distribuição das inovações não for normal, mas leptocúrtica, o uso da normalidade dentro do regime pode afetar seriamente a identificação do estado do regime e, conseqüentemente, também a estimação dos parâmetros para cada regime. Isto é, o modelo com distribuição normal nos regimes tenderá a detectar mudanças de regime com muita frequência devido a uma observação atípica (PAOLELLA et al., 2012).

3) Do passo anterior podemos calcular a densidade condicional da  $t$ -ésima observação como sendo:

$$f(y_t | \mathcal{I}_{t-1}) = \sum_{i=1}^k \sum_{j=1}^k p_{ij} P(\Delta_{t-1} = i | \mathcal{I}_{t-1}) \eta_{jt} = \mathbf{1}' (\mathbf{P} \boldsymbol{\xi}_{t-1} \odot \boldsymbol{\eta}_t), \quad (3.14)$$

em que  $\odot$  representa o produto de Hadamard e  $\mathbf{P}$  é a matriz de transição definida em (3.7).

4) Atualizar as probabilidades de estar em cada regime da seguinte forma:

$$\boldsymbol{\xi}_t = \frac{\mathbf{P} \boldsymbol{\xi}_{t-1} \odot \boldsymbol{\eta}_t}{\mathbf{1}' (\mathbf{P} \boldsymbol{\xi}_{t-1} \odot \boldsymbol{\eta}_t)}. \quad (3.15)$$

Como resultado da iteração de 1) a 4), podemos calcular a função quase-verossimilhança condicional ao longo de toda amostra. Assim, a estimativa de  $\boldsymbol{\theta}$  é obtida maximizando a verossimilhança ou a log-verossimilhança através de algum método numérico. De tal forma que:

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \Theta} \sum_{t=2}^T \log f(y_t | \mathcal{I}_{t-1}) = \arg \max_{\boldsymbol{\theta} \in \Theta} \sum_{t=2}^T \log(\mathbf{1}'(\mathbf{P}\boldsymbol{\xi}_{t-1} \odot \boldsymbol{\eta}_{jt})). \quad (3.16)$$

Assim como nos modelos com um regime, a escolha dos valores iniciais pode ter influência nas estimativas dos parâmetros. Para os valores iniciais de  $\boldsymbol{\xi}_1$  frequentemente utiliza-se a distribuição estacionária estimada da cadeia de Markov (3.8) ou então fazê-lo como sendo equiprovável, ou considerados como parâmetros e estimados por máxima verossimilhança. Para os valores de  $\mathbf{h}_1$ , podemos assumir como sendo função dos parâmetros, de tal forma que para o regime  $k$  temos que  $h_1^k = \alpha_{0k}/(1 - \alpha_{1k} - \beta_{1k})$ .

### 3.2.2 Observações sobre a estimação por máxima-verossimilhança

Até o momento em que este trabalho está sendo escrito, ainda não encontramos na literatura as propriedades assintóticas, como consistência e distribuição assintótica do estimador de QML para o modelo MS-GARCH que segue a especificação proposta por Haas, Mittnik e Paoletta (2004).

Como cita Paoletta et al. (2012), semelhantemente aos modelos de mistura GARCH, frequentemente a maximização da log-verossimilhança (3.16) é feita por um método de otimização baseadas na matriz Hessiana com derivadas determinadas numericamente. E nisso pode haver vários máximos locais plausíveis da função de verossimilhança, e apenas por meio do uso de vários valores iniciais pode-se obter o que é provavelmente o máximo global. Para o modelo especificado em (3.4), Augustyniak (2014) também salienta o impacto dos valores iniciais nas estimativas dos parâmetros. E, apesar de ser mais complexo, ele propõe algoritmo que pode reduzir a sensibilidade aos valores iniciais. Outro ponto a ser citado, é que, devido à natureza da função de verossimilhança de uma mistura, ela pode possuir singularidades (valores de verossimilhança infinitos).

Outro ponto que é importante discutir é o número de regimes  $K$  a definido. Andreou e Ghysels (2009) mostram diferentes métodos estatísticos para testar a hipótese de haver quebra estrutura. Porém, Paoletta et al. (2012) afirmam que esses testes padrões e também testes baseados em razões de verossimilhança não têm bom desempenho e que critérios de informação como AIC e BIC são razoavelmente bons nesse contexto. No entanto, em problemas empíricos, levando em conta a interpretabilidade, geralmente o número de regimes escolhidos costuma ser 2, uma vez que em modelos com  $K > 2$ , pelo menos um dos regime tende a ter uma probabilidade estacionária muito pequena e, conseqüentemente, seus parâmetros estão sujeitos a um erro de estimação bastante grande (PAOLELLA et al., 2012).

### 3.3 Estimação Robusta

Para lidar com o impacto na estimação do modelo causado pelos *outliers*, nessa seção iremos propor um estimador que utiliza três modificações ao apresentado anteriormente: i) o método BIP dado pela Equação (2.9) na Seção 2.2 para minimizar o efeito de *outliers* nas estimativas da variância condicional de cada regime; ii) a utilização da distribuição t-Student com 4 graus de liberdade nas densidades condicionais; e iii) uma função objetivo dada igual à do Boudt, Danielsson e Laurent (2013) no caso univariado.

A motivação de utilizar técnicas robustas nessas três partes é dada pelos aspectos desejáveis em um estimador robusto descrito por Huber e Ronchetti (2009, pg. 25). A utilização e a motivação do BIP são dadas de forma semelhante ao do caso univariado. Já em relação ao uso da distribuição t-Student com 4 graus de liberdade nas densidades condicionais é embasada pelas técnicas utilizadas em filtros robustos. Diferentes autores mencionam que os filtros habituais, assim como o de Hamilton, não são robustos, mesmo nos casos mais modestos apresentam uma performance desapropriada. Petrus (1999), Zou, Chan e Ng (2000), e Calvet, Czellar e Ronchetti (2015) mostram o mau desempenho de diversos filtros como o de mínimos quadrados e bayesianos, por exemplo, na presença de *outliers*.

Para entender a representação dos *outliers* nos filtros, seguiremos as definições apresentadas por Calvet, Czellar e Ronchetti (2015). Seja  $y_t \in \mathbb{R}$  a observação com densidade condicional  $f(y_t|\Delta_t, \mathcal{I}_{t-1})$ . Assim como em (3.15), pela regra de Bayes, as probabilidades de estar em cada regime inferidas pelo filtro satisfazem:

$$\lambda(\Delta_t|y_t, \mathcal{I}_{t-1}) = \frac{f(y_t|\Delta_t, \mathcal{I}_{t-1})\lambda(\Delta_t|\mathcal{I}_{t-1})}{f(y_t|\mathcal{I}_{t-1})}, \quad (3.17)$$

em que  $f(y_t|\mathcal{I}_{t-1}) = \int f(y_t|\Delta_t, \mathcal{I}_{t-1})\lambda(\Delta_t|\mathcal{I}_{t-1})d\Delta_t$ . Assim, seja uma família de densidades de observação contaminadas por *outliers*, denotada por  $f_{cont}(\cdot|\Delta_t, \mathcal{I}_{t-1}, \omega)$ , parametrizada por  $\omega$ , no qual  $\omega$  pertence a um intervalo  $\mathcal{D}$  na reta dos reais contendo zero. Se  $\omega = 0$ , a densidade contaminada coincide com a densidade de observação não contaminada, isto é,  $f_{cont}(y_t|\Delta_t, \mathcal{I}_{t-1}, \omega) = f(y_t|\Delta_t, \mathcal{I}_{t-1})$  para todo  $\Delta_t$  e  $\mathcal{I}_{t-1}$ . Quando a contaminação não é levada em consideração, a aplicação ingênua da regra de Bayes leva à fórmula (3.17).

Posto isso, para tornar as probabilidades inferidas pelo filtro robustas, tanto Calvet, Czellar e Ronchetti (2015) e Petrus (1999) utilizam densidades robustas. Calvet, Czellar e Ronchetti (2015) utiliza uma baseada na limitação da derivada do log da densidade e Petrus (1999) em uma que é baseada na função objetivo de Huber,  $\rho(u) = u^2I(|u| \leq k) + k|u|I(|u| < k)$ . No entanto, para o nosso trabalho, necessitamos de uma densidade que seja capaz de ser estendida de maneira simples para o caso multivariado. Assim, de maneira análoga a utilizada por Boudt, Danielsson e Laurent (2013), usaremos uma t-Student com 4 graus de liberdade.

Sendo assim, a proposta de estimação robusta será modificando a estimação apresentada da Subseção 3.2.1 de tal forma que, no passo 2) utilizaremos o mecanismo BIP para minimizar o efeito de *outliers* nas estimativas da variância condicional de cada regime e utilizaremos a distribuição t com 4 graus de liberdade no cálculo da densidade condicional de cada regime; e, por fim, ao invés de maximizar a log-verossimilhança como em (3.16), utilizaremos uma modificação dela cominada a função  $\rho$  dada em (2.11). Assim sendo, temos que a estimativa de  $\theta$  é obtida minimizando:

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \sum_{t=1}^T \log \left[ \sum_{i=1}^k \sum_{j=1}^k p_{ij} P(\Delta_{t-1} = i | \mathcal{I}_{t-1}) \frac{1}{\sqrt{h_t^{(j)}}} \left( 1 + \frac{y_t^2}{2h_t^{(j)}} \right)^{\frac{-(1+4)\sigma_{t1,4}}{2}} \right]. \quad (3.18)$$

em que  $\sigma_{tN,\nu}$  é dado por (2.13).

### 3.4 Previsão das Volatilidades

Assim, como cita Ardia et al. (2019), uma das principais aplicações dos modelos MS-GARCH é em finanças quantitativas, quando os investidores desejam alocar seu patrimônio em uma série de oportunidades de investimento de risco. Para isso, prever volatilidades futuras para calcular indicadores como Valor em Risco e Perda Esperada, por exemplo, é de grande importância. E como cita Abramson e Cohen (2007), em algumas aplicações a previsão de volatilidade um passo à frente supera significativamente o modelo GARCH com um único regime.

Como ressalta Paoella et al. (2012), uma característica interessante dos modelos MS é que as esperanças condicionais podem ser calculadas analiticamente, o que está em contraste com muitos outros modelos de séries temporais não lineares. Posto isso, aqui iremos utilizar a forma da previsão apresentada em Paoella et al. (2012) para o modelo MS(K)-GARCH(1,1). Para entender como funciona a previsão da volatilidade, iremos escrever (3.6) em forma matricial:

$$\mathbf{X}_t = \Psi + \mathbf{C}_{\Delta_{t-1}, t-1} \mathbf{X}_{t-1} \quad (3.19)$$

em que  $\mathbf{X}_t = \{h_t^{(1)}, \dots, h_t^{(k)}, y_{t-1}^2\}'$ ,  $\mathbf{C}_{\Delta_t, t} = \alpha \mathbf{1}'_{\Delta_t} z_t^2 + \beta$ , tal que  $\mathbf{1}^{(k)}$  é um vetor unitário em  $\mathbb{R}^{K+1}$ , para  $k = 1, \dots, K$ , e

$$\Psi = \begin{bmatrix} \alpha_0^{(1)} \\ \vdots \\ \alpha_0^{(K)} \\ 0 \end{bmatrix}, \quad \alpha = \begin{bmatrix} \alpha_1^{(1)} \\ \vdots \\ \alpha_1^{(K)} \\ 1 \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1^{(1)} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \beta_1^{(k)} & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix}. \quad (3.20)$$

Além disso, para qualquer função  $f : S \rightarrow M_{n \times n'}(\mathbb{R})$ , tal que  $M_{n \times n'}(\mathbb{R})$  é o espaço de matrizes reais  $n \times n'$ , e  $S$  é o espaço de estado de  $\{\Delta_t\}$ , defina a matriz

$$\mathbb{P}_f^{(d)} = \begin{bmatrix} p_{11}^{(d)} f(1) & \dots & p_{K1}^{(d)} f(1) \\ \vdots & \dots & \vdots \\ p_{1K}^{(d)} f(k) & \dots & p_{KK}^{(d)} f(K) \end{bmatrix}, \quad (3.21)$$

e defina também

$$C_{mm}(j) = E(C_{jt}^{\otimes m}), \quad j = 1, \dots, K, \quad m \in \mathbb{N}, \quad (3.22)$$

em que  $A^{\otimes m}$  denota o m-ésima potência Kronecker de  $A^2$ .

Como cita, Paoletta et al. (2012) as das volatilidades previsões podem ser calculadas por meio de um esquema recursivo simples. Assim, a substituição repetida da Equação (3.19) produz:

$$\begin{aligned} \mathbf{X}_{t+\tau+1} &= \Psi + C_{\Delta_{t+\tau}, t+\tau} \mathbf{X}_{t+\tau} \\ &= \Psi + C_{\Delta_{t+\tau}, t+\tau} \Psi + C_{\Delta_{t+\tau}, t+\tau} C_{\Delta_{t+\tau-1}, t+\tau-1} \mathbf{X}_{t+\tau-1} \\ &\vdots \\ &= \sum_{i=1}^{\tau} \left\{ \prod_{l=1}^{i-1} C_{\Delta_{t+\tau+1-l}, t+\tau+1-l} \right\} \Psi + \prod_{l=1}^{\tau} \left\{ C_{\Delta_{t+\tau+1-l}, t+\tau+1-l} \right\} \mathbf{X}_{t+1}, \end{aligned}$$

em que  $\prod_{l=1}^0 C_{\Delta_{t+\tau+1-l}, t+\tau+1-l} \stackrel{\text{def}}{=} \mathbf{I}_{K+1}$ . E seja  $\Delta_t = \{\Delta_s : s \leq t\}$ , então

$$E_t(\mathbf{X}_{t+\tau+1} | \Delta_{t+\tau}) = \sum_{i=1}^{\tau} \left\{ \prod_{l=1}^{i-1} C_{11}(\Delta_{t+\tau+1-l}) \right\} \Psi + \left\{ \prod_{l=1}^{\tau} C_{11}(\Delta_{t+\tau+1-l}) \right\} \mathbf{X}_{t+1}, \quad (3.23)$$

em que  $E_t(\cdot)$  denota a esperança dado  $\mathcal{I}_t$ . Aplicando alguns lemas mostrados em Paoletta et al. (2012), obtém-se que:

$$E_t(\mathbf{X}_{t+\tau+1}) = (\mathbf{1}_K \otimes I_{k+1}) M(\tau), \quad \tau \geq 0, \quad (3.24)$$

em que  $M(0) = \pi_t \otimes \mathbf{X}_{t+1}$  e

$$\begin{aligned} M(\tau) &= \sum_{i=1}^{\tau} \mathbb{P}_{C_{11}}^{i-1}(\pi_{t+\tau+1-i} \otimes \Psi) + \mathbb{P}_{C_{11}}^d(\pi_t \otimes \mathbf{X}_{t+1}) \\ &= \pi_{t+\tau} \otimes \Psi + \mathbb{P}_{C_{11}} \left\{ \sum_{i=1}^{\tau-1} \mathbb{P}_{C_{11}}^{i-1}(\pi_{t+(\tau-1)+1-i} \otimes \Psi) + \mathbb{P}_{C_{11}}^{d-1}(\pi_t \otimes \mathbf{X}_{t+1}) \right\} \\ &= \pi_{t+\tau} \otimes \Psi + \mathbb{P}_{C_{11}} M(\tau-1), \quad \tau \geq 1. \end{aligned}$$

Portanto, as previsões de volatilidade podem ser calculadas como

$$E_t(y_{t+\tau}^2) = (\mathbf{1}_K \otimes e_{K+1})' M(\tau), \quad (3.25)$$

em que

$$M(\tau) = \begin{cases} \pi_t \otimes \mathbf{X}_{t+1}, & \text{se } \tau = 0 \\ \pi_{t+\tau} \otimes \mathbf{\Psi} + \mathbb{P}_{C_{11}} M(\tau - 1), & \text{se } \tau \geq 1. \end{cases} \quad (3.26)$$

Na prática, estimamos  $\pi_t$  na Equação (3.26) via  $\xi_t$  na Equação (3.15).

## 4 Modelos GARCH Multivariados com Mudança de Regime Markoviano

Discutimos anteriormente que os modelos GARCH são usualmente aplicados a séries temporais financeiras, principalmente para modelar a dinâmica da volatilidade, bem como na sua previsão para aplicação de medidas de risco como VaR e ES. Porém, frequentemente também estamos interessados em avaliar não apenas o comportamento de um único ativo, mas sim de um portfólio contendo vários ativos e/ou para analisar a dependência entre um grupo de ativos. Isto é, vimos como analisar e prever as séries financeiras univariadas, mas é interessante ver como ela se relacionam entre si ao longo do tempo. E estender a modelagem univariada para o caso multivariado proporciona mais ferramentas de decisão na análise de risco como índice Sharpe, carteira de mínima variância, *turnover* etc.

Segundo Tsay (2014, Seção 7, pág 399), seja o caso em que  $\mathbf{y}_t$  é um vetor de dimensão  $N \times 1$  com  $N$  ativos, o modelo GARCH pode ser estendido para o caso multivariado da seguinte forma:

$$\mathbf{y}_t = \mathbf{H}_t^{1/2} \mathbf{z}_t \quad (4.1)$$

em que  $\mathbf{z}_t$  é uma sequência de vetores aleatórios independentes e identicamente distribuídos, tais que  $E[\mathbf{z}_t] = \mathbf{0}$  e  $Cov[\mathbf{z}_t] = \mathbf{I}_N$ .  $\mathbf{H}_t^{1/2}$  denota a matriz  $N \times N$  raiz quadrada definida positiva da matriz de covariância condicional  $\mathbf{H}_t$ . Especificamente, seja  $\mathbf{H}_t = \mathbf{P}_t \mathbf{\Lambda}_t \mathbf{P}_t'$  representa a decomposição espectral de  $\mathbf{H}_t$ , em que  $\mathbf{\Lambda}_t$  é a matriz diagonal dos autovalores de  $\mathbf{H}_t$  e  $\mathbf{P}_t'$  denota a matriz ortonormal de autovetores. Então,  $\mathbf{H}_t^{1/2} = \mathbf{P}_t \mathbf{\Lambda}_t^{1/2} \mathbf{P}_t'$ , pode ser obtida através da decomposição de Cholesky. Além disso como cita Francq e Zakoian (2019), a escolha de uma especificação para  $\mathbf{H}_t$  é obviamente mais delicada do que na estrutura univariada porque: (i)  $\mathbf{H}_t$  deve ser (quase certamente) simétrico e definido positivo para todo  $t$ ; (ii) a especificação deve ser simples o suficiente para ser passível de estudo probabilístico (existência de soluções, estacionariedade etc.), embora seja de generalidade suficiente; (iii) a especificação deve ser parcimoniosa o suficiente para permitir uma estimação viável. No entanto, o modelo deve ser capaz de capturar a, possivelmente sofisticada, dinâmica na estrutura de covariância.

Diante disso, surgiram diversas especificações de  $\mathbf{H}_t$  na literatura. Nesse trabalho não entraremos em detalhe em todos eles, no entanto, Francq e Zakoian (2019) e Almeida, Hotta e Ruiz (2018) apresentam uma boa revisão sobre esse tópico. De qualquer forma, podemos listar o modelo VEC-GARCH (*vector GARCH* em inglês) proposto por Bollerslev, Engle e Wooldridge (1988), que é a generalização mais direta do GARCH univariado, em que cada covariância condicional é uma função de variâncias condicionais



defasadas, bem como de produtos cruzados defasados de todos os componentes do vetor de retornos. No entanto, este modelo não é muito parcimonioso, visto que possui um total de  $(N + 1) + (1/2)N(N + 1) + (1/4)N^2(N + 1)^2(p + q)$  parâmetros. Semelhante a esse modelo temos que o modelo BEKK (Baba-Engle-Kraft-Kroner) de Engle e Kroner (1995) que também usa muitos parâmetros e que na especificação mais simples possui  $N^2 + [N(N + 1)/2]$ .

Há também os modelos fatoriais como o O-GARCH de (*Orthogonal GARCH* em inglês) Alexander e Chibumba (1997) e o modelo GO-GARCH (*Generalized Orthogonal-GARCH* em inglês) de Weide (2002). Como cita Orskaug (2009), a ideia de modelos fatoriais vem da teoria econômica. Nesta classe, as matrizes de covariância condicional são motivadas pela parcimônia. Supõe-se que o processo é gerado por um (pequeno) número de fatores heterocedásticos não observados, portanto, esses modelos são chamados de modelos fatoriais. Esses fatores podem ser estudados e pode-se fazer suposições de que algumas características dos dados são capturadas, de forma semelhante à análise de componentes principais. Esta abordagem tem a vantagem de reduzir a dimensionalidade do problema quando o número de fatores relativos à dimensão do vetor de retorno é pequeno.

Porém, nesse trabalho estamos interessados nos modelos de modelos de correlações condicionais, mais especificamente no modelo DCC (*Dynamic Conditional Correlation* em inglês) proposto por Engle (2002).

## 4.1 Modelo DCC com 1 Regime

A ideia do modelo DCC é que a matriz  $\mathbf{H}_t$ , pode ser decomposta pela matriz dos desvios padrões condicionais, digamos  $\mathbf{D}_t$ , e pela matriz de correlação,  $\mathbf{R}_t$ , que variam ao longo do tempo. Isto é, segundo Engle (2002), seja  $\mathbf{y}_t$  um vetor de dimensão  $N \times 1$  com  $N$  ativos, temos que:

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t \quad , \quad (4.2)$$

Apesar de ser possível representar o modelo DCC para quaisquer ordens  $P, Q \geq 1$ , neste trabalho utilizaremos  $P = Q = 1$ , tanto por questões de parcimônia e devido ao fato de que, em muitas aplicações essa ordem é utilizada.

Denote a matriz  $\mathbf{D}_t$  por

$$\mathbf{D}_t = \begin{bmatrix} \sqrt{h_{1t}} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{h_{2t}} & 0 & \dots & 0 \\ 0 & 0 & \sqrt{h_{3t}} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sqrt{h_{Nt}} \end{bmatrix},$$

em que  $\sqrt{h_{it}}$  representa o desvio padrão condicional de  $y_{it}$  do  $i$ -ésimo ativo definida por (2.1) ou por algum outro modelo.

A matriz de correlação  $\mathbf{R}_t$  que pode ser expressa por:

$$\mathbf{R}_t = \begin{bmatrix} 1 & \rho_{12t} & \rho_{13t} & \dots & \rho_{1Nt} \\ \rho_{21t} & 1 & \rho_{23t} & \dots & \rho_{2Nt} \\ \rho_{31t} & \rho_{32t} & 1 & \dots & \rho_{3Nt} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{N1t} & \rho_{N2t} & \rho_{N3t} & \dots & 1 \end{bmatrix}.$$

O vetor de retornos padronizados é dado por:

$$\boldsymbol{\epsilon}_t = \mathbf{D}_t^{-1} \mathbf{y}_t.$$

No modelo DCC temos:

$$\begin{aligned} \mathbf{R}_t &= \mathbf{Q}_t^{*-1} \mathbf{Q}_t \mathbf{Q}_t^{*-1} \\ \mathbf{Q}_t &= (1 - a - b) \bar{\mathbf{Q}} + a \boldsymbol{\epsilon}_{t-1} \boldsymbol{\epsilon}'_{t-1} + b \mathbf{Q}_t, \end{aligned}$$

em que parâmetros  $a$  e  $b$  possuem as restrições  $a, b \geq 0$  e  $a + b < 1$  para positividade de  $\mathbf{H}_t$ . Além disso, como enfatiza Aielli (2013) temos que a persistência do processo de correlação é uma função crescente de  $a + b$ . Quanto maior o impacto do inovações, medidas por  $a$ , maior será a variabilidade do processo de correlação. Além disso, temos que se se houver uma raiz unitária nos parâmetros de correlação dinâmica (ou seja, para  $a + b = 1$ ), os elementos de  $\mathbf{Q}_t$  são pensados como processos GARCH integrados.

Nesse trabalho não iremos mostrar quais são as condições suficientes e necessárias para estacionariedade do processo DCC; estas condições podem ser encontradas em Fermanian e Malongo (2017). Satisfeita as condições de estacionariedade, temos que  $\bar{\mathbf{Q}} = \text{Cov}[\boldsymbol{\epsilon}_{t-1} \boldsymbol{\epsilon}'_{t-1}] = E[\boldsymbol{\epsilon}_{t-1} \boldsymbol{\epsilon}'_{t-1}]$ , tal que  $\boldsymbol{\epsilon}_{t-1}$  representa o vetor  $N \times 1$  de retornos padronizados dos ativos, tal quantidade pode ser estimada através de

$$\widehat{\bar{\mathbf{Q}}} = \frac{1}{T} \sum_{i=1}^T \hat{\boldsymbol{\epsilon}}_i \hat{\boldsymbol{\epsilon}}_i', \quad (4.3)$$

em que  $\hat{\boldsymbol{\epsilon}}_t = \widehat{\mathbf{D}}_t^{-1} \mathbf{y}_t$ . Por fim,  $\mathbf{Q}_t^{*-1}$  é uma matriz diagonal cujo elementos são a raiz quadrada dos elementos da diagonal de  $\mathbf{Q}_t$ .

E, de acordo com Tsay (2014, Seção 7.7, pág 428), os modelos DCC tiram vantagem do fato de que as matrizes de correlação são mais fáceis de manipular do que as matrizes de covariância. A ideia dos modelos DCC é interessante e atraente porque embora haja muitos parâmetros a estimação pode ser realizada em dois passos. Como vimos antes, a matriz  $\bar{\mathbf{Q}}$  pode ser estimada pela Equação (4.3) e depois temos apenas que estimar os parâmetros  $a$  e  $b$ . A estimação do modelo DCC, incluindo as estimativas das variâncias condicionais, quando elas são estimadas por modelos GARCH é apresentado a seguir.

#### 4.1.1 Estimação por máxima quase-verossimilhança

De maneira similar aos modelos GARCH univariados, existem formas diferentes de se estimar os parâmetros do modelo DCC definido em (4.2). Por exemplo, Nascimento et al. (2020) apresentam uma abordagem bayesiana para a estimação dos parâmetros do modelo. No entanto, aqui estaremos interessados na estimação por verossimilhança do modelo, de maneira mais específica utilizaremos a quase-verossimilhança. Aqui não iremos fazer nenhuma suposição sobre a distribuição de  $\mathbf{z}_t$ , porém iremos utilizar a função de quase verossimilhança da distribuição gaussiana multivariada.

Posto isso, temos que a distribuição conjunta de  $\mathbf{z}_t$ , condicionada a toda a informação até o tempo  $t - 1$ ,  $\mathcal{I}_{t-1}$ , é dada por:

$$f(\mathbf{z}_t | \mathcal{I}_{t-1}) = \prod_{t=1}^T \frac{1}{(2\pi)^{N/2}} \exp \left\{ -\frac{1}{2} \mathbf{z}_t' \mathbf{z}_t \right\},$$

assim temos que a verossimilhança do modelo é dada por:

$$L(\boldsymbol{\theta} | \mathbf{y}_t) = \prod_{t=1}^T \frac{1}{(2\pi)^{N/2} |\mathbf{H}_t|^{1/2}} \exp \left\{ -\frac{1}{2} \boldsymbol{\epsilon}_t' \mathbf{H}_t^{-1} \boldsymbol{\epsilon}_t \right\}. \quad (4.4)$$

Fazendo  $\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t$  e tomando o logaritmo em (4.4), como mostra Orskaug (2009), obtemos então que:

$$\begin{aligned} \ell(\boldsymbol{\theta} | \mathbf{y}_t) &= \ln(L(\boldsymbol{\theta} | \mathbf{y}_t)) = -\frac{1}{2} \sum_{t=1}^T (N \ln(2\pi) + \ln(|\mathbf{H}_t|) + \boldsymbol{\epsilon}_t' \mathbf{H}_t^{-1} \boldsymbol{\epsilon}_t) \\ &= -\frac{1}{2} \sum_{t=1}^T (N \ln(2\pi) + \ln(|\mathbf{D}_t \mathbf{R}_t \mathbf{D}_t|) + \boldsymbol{\epsilon}_t' \mathbf{D}_t^{-1} \mathbf{R}_t^{-1} \mathbf{D}_t^{-1} \boldsymbol{\epsilon}_t) \\ &= -\frac{1}{2} \sum_{t=1}^T (N \ln(2\pi) + 2 \ln(|\mathbf{D}_t|) + \ln(\mathbf{R}_t) + \boldsymbol{\epsilon}_t' \mathbf{D}_t^{-1} \mathbf{R}_t^{-1} \mathbf{D}_t^{-1} \boldsymbol{\epsilon}_t). \end{aligned} \quad (4.5)$$

Consequentemente, a estimativa  $\hat{\boldsymbol{\theta}}$  de  $\boldsymbol{\theta}$ , é obtida maximizando a log-verossimilhança  $\ell(\boldsymbol{\theta})$ . No entanto, na maioria das vezes o modelo DCC é estimado em dois estágios. No primeiro,

estimamos os parâmetros dos modelos univariados presente em  $\mathbf{D}_t$  - frequentemente se utiliza o modelo GARCH(1,1). No segundo estágio estimaremos  $\mathbf{R}_t$  utilizando os valores de  $\mathbf{D}_t$  obtidos no primeiro estágio.

### Estágio 1

Como mostra Orskaug (2009), no primeiro estágio,  $\mathbf{R}_t$  é substituído pela matriz de identidade  $\mathbf{I}_N$  em (4.5), o que resulta na função de quase-verossimilhança:

$$\begin{aligned} \ell(\boldsymbol{\theta}|\mathbf{y}_t) &= -\frac{1}{2} \sum_{t=1}^T (N \ln(2\pi) + 2 \ln(|\mathbf{D}_t|) + \ln(\mathbf{R}_t) + \boldsymbol{\epsilon}'_t \mathbf{D}_t^{-1} \mathbf{R}_t^{-1} \mathbf{D}_t^{-1} \boldsymbol{\epsilon}_t) \\ &= \sum_{i=1}^N \left( -\frac{1}{2} \sum_{t=1}^T \left[ \ln(h_{it}) + \frac{y_{it}^2}{h_{it}} \right] + \text{constante} \right). \end{aligned}$$

Nota-se que a log-verossimilhança acima é dada pelo somatório das log-verossimilhanças dos N modelos GARCH univariados. Nesse caso então temos que os parâmetros dos diferentes GARCH univariados os modelos podem ser determinados separadamente. Logo nessa etapa obtemos  $\hat{\mathbf{D}}_t$ , restando estimar os parâmetros da correlação.

### Estágio 2

Na segunda etapa, inicialmente estimamos  $\bar{\mathbf{Q}}$  através da Equação (4.3), e depois estimamos  $(a, b)$  usando a log-verossimilhança especificada em (4.5), dados os parâmetros estimados da etapa 1. A função de quase verossimilhança de segundo estágio obtida pela Equação (4.5) substituindo  $\mathbf{D}_t$  por  $\hat{\mathbf{D}}_t$ .

Além disso, sobre as condições apresentada por Engle e Sheppard (2001), temos que o estimador  $\hat{\boldsymbol{\theta}}$  é consistente, isto é,  $\hat{\boldsymbol{\theta}} \xrightarrow{P} \boldsymbol{\theta}$ . Além disso, sua distribuição assintótica é dada por:

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} N\left(\mathbf{0}, \mathbf{A}_0^{-1} \mathbf{B}_0 \mathbf{A}_0'^{-1}\right),$$

em que

$$\begin{aligned} \mathbf{A}_0 &= \begin{bmatrix} \nabla_{\phi\phi} \ln(f_1(\boldsymbol{\phi}_0)) & 0 \\ \nabla_{\phi\psi} \ln(f_2(\boldsymbol{\theta}_0)) & \nabla_{\psi\psi} \ln(f_2(\boldsymbol{\theta}_0)) \end{bmatrix} \quad \text{e} \\ \mathbf{B}_0 &= \text{Var} \left[ \sum_{t=1}^T (T^{-1/2} \nabla'_{\phi} \ln(f_1(\mathbf{y}_t, \boldsymbol{\phi}_0)), \nabla'_{\psi} \ln(f_1(\mathbf{y}_t, \boldsymbol{\phi}_0, \boldsymbol{\psi}_0)) \right]. \end{aligned}$$

Assim como no caso univariado, alguns autores mostram que a presença de *outliers* pode afetar bastante as estimativas dos modelos GARCH multivariado. Como cita LaBarr (2010), *outliers* em um modelo DCC-GARCH afetam a estimativa das volatilidades univariadas pela definição do modelo. Além disso, devido à natureza de correlação, esse

*outliers* podem afetar as estimativas da volatilidade das outras séries dentro do mesmo conjunto de observações. Alguns autores mostram o mal desempenho do estimador de máxima verossimilhança na presença de *outliers*; por exemplo, Boudt, Danielsson e Laurent (2013) e Grané, Veiga e Martín-Barragán (2014).

#### 4.1.2 Estimadores robustos

Dentre os estimadores robustos propostos na literatura, aqui apresentaremos o proposto por Boudt, Danielsson e Laurent (2013). A ideia desse estimador é de certa forma semelhante a da estimação do modelo GARCH univariado proposto por Boudt, Danielsson e Laurent (2013). No entanto, esse estimador foi baseado no modelo cDCC apresentado em Aielli (2013). Apesar da estimação apresentada ser referente ao modelo cDCC, podemos apresentá-la para o caso DCC. Para lidar com os efeitos dos *outliers*, Boudt, Danielsson e Laurent (2013) propõem três modificações do modelo DCC:

- (a) A ideia da primeira modificação é controlar os efeitos dos *outliers* nos modelos univariados. Para isso, o objetivo é utilizar os métodos robustos apresentados anteriormente;
- (b) Estimar  $\bar{\mathbf{Q}}$  com um estimador de correlação robusto; e
- (c) Substituição do modelo DCC por uma especificação robusta dele.

Semelhantemente a diversos métodos robustos que citamos aqui, a ideia é limitar o efeito do *outliers* em todos os pontos da estimação do modelo. Para a especificação apresentada por Boudt, Danielsson e Laurent (2013), denotada por BIP-DCC, iremos começar pela terceira modificação, isto é, pela especificação da matriz  $\mathbf{Q}_t$ . Para isso, aqui é utilizado um mecanismo parecido com (2.9). De acordo com Boudt, Danielsson e Laurent (2013), o modelo DCC padrão sem *outliers*,  $\boldsymbol{\epsilon}_{t-1}$  é condicionalmente distribuído normalmente com média zero e matriz de covariância  $\mathbf{R}_{t-1}$ . Para detectar saltos em  $\boldsymbol{\epsilon}_{t-1}$ , eles utilizam o quadrado da distância de Mahalanobis (MD):

$$d_{t-1} = \boldsymbol{\epsilon}'_{t-1} \mathbf{R}_{t-1} \boldsymbol{\epsilon}_{t-1}. \quad (4.6)$$

A distância MD é distribuída condicionalmente como uma variável aleatória qui-quadrado com N graus de liberdade. Se qualquer um dos componentes em  $\mathbf{y}_{t-1}$  for um *outliers*, a distância MD será elevada. Portanto, se  $d_{t-1}$  exceder um quantil alto da distribuição  $\chi^2_{(N)}$  (denotado por  $k_{\delta, N}$ ), é provável que  $\mathbf{y}_{t-1}$  seja um valor extremo e deverá ser reduzido. Assim, o BIP-DCC assume a forma

$$\mathbf{Q}_t = (1 - a - b)\bar{\mathbf{Q}} + aw_{k_{\delta, N}}(d_{t-1})\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1} + b\mathbf{Q}_t \quad (4.7)$$

em que  $w_{k_{\delta,N}}(u) = c_{\delta,N} \min(1, 1/k_{\delta,N})$ . O fator de correção  $c_{\delta,N}$  é tal que  $E[w_{k_{\delta,N}}(u)(\mathbf{z}'\mathbf{z})\mathbf{z}\mathbf{z}'] = \mathbf{I}_N$  se  $\mathbf{y} \sim N(\mathbf{0}, \mathbf{I}_N)$ . A escolha de  $\delta$  é baseada em uma troca de eficiência versus robustez. E, como salientam Boudt, Danielsson e Laurent (2013), se todas as observações de retorno seguirem o modelo DCC, então o modelo BIP-DCC induz um viés e erros quadráticos médios maiores nos parâmetros estimados. Quanto mais altos os valores de  $k_{\delta,N}$ , mais próximo o modelo BIP-DCC está do modelo DCC e, portanto, menor é o viés de especificação incorreta. No entanto, grandes valores de  $k_{\delta,N}$  também implicam que o efeito dos extremos nas correlações se torna maior.

Como mostramos anteriormente, comumente  $\bar{\mathbf{Q}}$  é estimado através da correlação amostral. Mas, na presença de *outliers*, tal estimativa provavelmente será afetada negativamente. Assim, na segunda modificação, Boudt, Danielsson e Laurent (2013) substituem a matriz de correlação simples por um estimador que seja mais robusto na presença de tais *outliers*. Eles propõem um estimador de correlação robusto e reponderado que é proporcional à correlação amostral das observações para as quais nenhum *outliers* foi detectado usando uma estatística de teste multivariada baseada em estimativas de correlação local.

A estimativa da matriz de correlação apresentada por eles é calculada em duas etapas. Primeiro, é calculado a matriz de correlação de Spearman  $\mathbf{RC}_t$  em uma janela local em torno de cada observação  $\epsilon_t$ . Mais precisamente, seja  $\epsilon_{t:1}, \dots, \epsilon_{t:K+1}$  como observações ordenadas no tempo na janela  $t$ ; em seguida, ranqueie cada série de componentes na janela local. Denote a série de vetores contendo esses ranques por  $L_{t:1}, \dots, L_{t:K+1}$ . A matriz de correlação de Spearman bruta  $\mathbf{C}_t$  para a janela em torno de  $\epsilon_t$  é a matriz de correlação de amostra de  $L_{t:1}, \dots, L_{t:K+1}$ . Moran (1948) mostra que esta matriz de correlação precisa ser corrigida da seguinte forma para garantir a consistência:

$$\mathbf{SC}_t = 2 \operatorname{sen} \left( \frac{1}{6} \pi \mathbf{C}_t \right)$$

Como cita Boudt, Danielsson e Laurent (2013), as vantagens da matriz de correlação de Spearman em relação a outros estimadores de correlação robustos incluem sua simplicidade computacional, alta eficiência e robustez *outliers*.

A segunda etapa é calcular o estimador de correlação reponderado robusto:

$$\mathbf{RC} = \frac{c_{\delta,N}}{\sum_{t=1}^T L_t} \sum_{t=1}^T \epsilon_t \epsilon_t' L_t$$

com pesos  $L_t = I[\epsilon_t' \mathbf{RC}_t^{-1} \epsilon_t \leq \chi_N^2(\delta)]$  O estimador de correlação reponderado de  $\bar{\mathbf{Q}}$  é dado por

$$\hat{\bar{\mathbf{Q}}}_{RW} = \operatorname{diag}(\mathbf{RC}_{11}^{-1/2} \dots \mathbf{RC}_{NN}^{-1/2}) \mathbf{RC} \times \operatorname{diag}(\mathbf{RC}_{11}^{-1/2} \dots \mathbf{RC}_{NN}^{-1/2}).$$

De maneira similar à ideia usada por Muler e Yohai (2008), Boudt, Danielsson e Laurent (2013) utilizam M-estimadores para estimar  $a$  e  $b$ . Eles definem os estimadores M para modelos de correlação condicional como minimizadores da soma do valor médio de uma função  $\rho$ , avaliada nas distâncias de Mahalanobis ao quadrado e o valor médio de o log do determinante das matrizes de correlação. Isto é, aqui estamos estimando o modelo em 2 etapas: na primeira estimamos os GARCH univariados utilizando métodos robustos; na segunda, além das especificações descritas acima, utilizamos o seguinte estimador para  $\boldsymbol{\theta} = (a, b)$ :

$$\hat{\boldsymbol{\theta}} = \operatorname{argmim} M_T(\boldsymbol{\theta}, \rho) \equiv \frac{1}{T} \sum_{t=1}^T \left[ \ln \det \mathbf{R}_t + \sigma \rho \left( \boldsymbol{\epsilon}_t' \mathbf{R}_t^{-1} \boldsymbol{\epsilon}_t \right) \right] \quad (4.8)$$

em que  $\sigma$  é um fator de correção. Se  $\rho(z) = z$  e  $\sigma = 1$ , voltamos para o estimador QML-n apresentado anteriormente. A escolha de  $\rho$  utilizada por Boudt, Danielsson e Laurent (2013) de tal forma que seja menos afetada por *outliers* é dada por:

$$\rho(z) = -z + \sigma_{N,\nu} \rho_{t_{N,\nu}}(\exp(z)), \quad (4.9)$$

em que

$$\rho_{t_{N,\nu}}(u) = (N + \nu) \log \left( 1 + \frac{u}{\nu - 2} \right) \quad (4.10)$$

e

$$\sigma_{t_{N,\nu}} = \frac{N}{E[\rho'_{t_{N,\nu}}(W)W]}, \quad (4.11)$$

tal que  $W$  é uma variável aleatória qui-quadrada com  $N$  grau de liberdade,  $N$  é a dimensão do vetor de retornos e  $\nu$  são os graus de liberdade da distribuição t-Student. Boudt, Danielsson e Laurent (2013) recomendam utilizar os graus de liberdade  $\nu = 4$ , assim para  $N = 2$ , por exemplo,  $\sigma_{2,4} = 0,8258$ . Os demais valores tabelados de  $\sigma_{t_{N,\nu}}$  podem ser encontrados na Tabela 1 em Boudt, Danielsson e Laurent (2013).

## 4.2 Modelos DCC com Mudança de Regime

Em séries temporais muito longas também há possibilidades de mudanças de regimes não só nos parâmetros das variâncias condicionais das séries univariadas, mas também nos parâmetros das correlações. Além disso, é importante dizer que as mudanças de regime da correlação ou das variâncias condicionais nem sempre caminham juntas. Na literatura, existem diversos modelos DCC com mudança de regime. Abaixo citaremos algumas delas, mas nesse trabalho utilizaremos o modelo de Lee (2010) que será apresentado na Seção 4.2.1.

Pelletier (2006) primeiro apresenta um modelo de correlação dinâmica com mudança de regime (RSDC). Semelhante ao modelo DCC as matrizes de covariâncias são decompostas em desvios padrão e correlações, mas essas correlações são dinâmicas; e as transições entre os regimes são governadas por uma cadeia de Markov. Em outras palavras, ele escreve as matrizes de correlações como:

$$\mathbf{R}_t = \sum_{k=1}^K \mathcal{I}_{\Delta_t=k} \mathbf{R}_k,$$

em que  $\Delta_t$  é uma cadeia de Markov não observada com espaço de estados finitos  $S = \{1, \dots, K\}$  e uma matriz de transição  $\mathbf{P}$   $K \times K$  com elementos  $p_{ij} = P(\Delta_t = j | \Delta_{t-1} = i)$ .  $\mathcal{I}$  é uma função indicadora. As matrizes  $\mathbf{R}_t$   $K \times K$  são matrizes de correlação. Assim, para cada regime temos uma matriz de correlação diferente. Além disso, Pelletier (2006) apresenta uma versão parcimoniosa que é dada por:

$$\mathbf{R}_t = \mathbf{R} \lambda^{(k)} + \mathbf{I}_K (1 - \lambda^{(k)}),$$

em que  $\mathbf{R}$  é uma matriz de correlação fixa ao longo do tempo,  $\mathbf{I}_K$  é uma matriz de identidade  $K \times K$ ,  $\lambda^{(k)} \in [0, 1]$  é um processo aleatório univariado governado por um processo de Markov não observado com espaço de estados  $\{1, 2, \dots, K\}$ .

E, como cita Pelletier (2006), para os elementos fora da diagonal, apenas o produto de  $\mathbf{R}$  e  $\lambda$  pode ser identificado. Para resolver este problema de identificação, podemos considerar dois conjuntos naturais de restrições. A primeira é:

$$1 = \lambda^{(1)} > \lambda^{(2)}, \dots, \lambda^{(K-1)} > \lambda^{(K)}.$$

Pelletier (2006) também restringe o  $\lambda^{(K)}$ 's para serem uma sequência decrescente para remover a possibilidade de reclassificar o regime  $i$  como regime  $j$  e vice-versa.

Haas (2010) apresenta uma forma de recursão simples para previsões de covariância para tal modelo. Para o modelo MS(K)-CCC, Haas e Liu (2018) apresentam as propriedades de estacionariedade do processo assim como a estrutura dos momentos.



Complementarmente, eles usam essa metodologia para modelar a dinâmica da distribuição conjunta do mercado de ações global e dos retornos das ações imobiliárias.

Para o modelo DCC com mudança de regime, primeiro discutiremos o modelo apresentado por Billio e Caporin (2005). Além, é claro, de nesse modelo eles considerarem correlação constante, existem algumas diferenças entre esse modelo e o proposto por Pelletier (2006). Dito isso, o modelo proposto por Billio e Caporin (2005) pode ser expresso como:

$$\mathbf{Q}_t = (1 - a^{(k)} - b^{(k)})\bar{\mathbf{Q}}^{(k)} + a^{(k)}\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1} + b^{(k)}\mathbf{Q}_{t-1},$$

em que  $\bar{\mathbf{Q}}^{(k)}$  é a correlação incondicional no regime  $k$ . Isto é, aqui tanto a matriz de correlação incondicional quanto os parâmetros que conduzem a dinâmica da correlação podem ser dependentes do regime. E note que, assim como no modelo apresentado por Pelletier (2006), aqui estamos considerando a estrutura de mudança de regime nas correlações, excluindo qualquer efeito sobre as variâncias. Uma vantagem disso é que essa restrição permite considerar um procedimento de estimativa em duas etapas. Além disso, como cita Billio e Caporin (2005), um modelo de mudança de regime markoviana se tornaria altamente instável devido ao grande número de parâmetros.

E semelhante ao caso univariado, a estimação desse modelo seguindo a abordagem frequentista por verossimilhança não era factível devido à dependência do caminho do modelo. Dito isso, eles modificam a matriz  $\mathbf{Q}_t$  fazendo ela como sendo:

$$\mathbf{Q}_t^{ik} = (1 - a^{(k)} - b^{(k)})\bar{\mathbf{Q}}^{(k)} + a^{(k)}\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1} + b^{(k)}\mathbf{Q}_{t-1}^{li},$$

em que o sobrescrito  $k$ ,  $i$  e  $l$  se referem ao estado em  $t$ ,  $t - 1$  e  $t - 2$ , respectivamente. Segundo Billio e Caporin (2005), a estrutura dinâmica de  $\mathbf{Q}_t$  induz a dependência do regime atual com todos os regimes anteriores. Portanto, para a estimação por verossimilhança, alguma aproximação é necessária.

Porém, a aproximação utilizada por Billio e Caporin (2005) remete a dificuldades de interpretação semelhante ao caso univariado apresentado por Gray (1996). Dito isso, aqui estamos interessado em uma extensão do caso univariado apresentado por Haas, Mittnik e Paoletta (2004), em que eles resolvem o problema de dependência de caminho fazendo a variância condicional de regime depender apenas dela mesma.

#### 4.2.1 Modelo de Lee (2010)

Seja, o processo:

$$\begin{aligned} \mathbf{y}_t &= \left(\mathbf{H}_t^{(\Delta t)}\right)^{1/2} \mathbf{z}_t \\ \mathbf{H}_t^{(\Delta t)} &= \mathbf{D}_t \mathbf{R}_t^{(\Delta t)} \mathbf{D}_t \end{aligned} \tag{4.12}$$

em que  $\Delta_t$  é uma cadeia de Markov não observada com espaço de estados finitos  $S = \{1, \dots, K\}$  e uma matriz de transição  $\mathbf{P}$   $K \times K$  com elementos  $p_{ij} = P(\Delta_t = j | \Delta_{t-1} = i)$ . As demais componentes são definidas de maneira semelhante ao caso em que não há mudança de regime definida em (4.2).

**Definição 3.** *O modelo IS(K)-DCC(1,1) proposto por Lee (2010) seguindo o processo (4.12), possui matriz de covariância dos retornos padronizados como sendo:*

$$\mathbf{Q}_t^{(k)} = (1 - a^{(k)} - b^{(k)})\bar{\mathbf{Q}} + a^{(k)}\epsilon_{t-1}\epsilon'_{t-1} + b^{(k)}\mathbf{Q}_{t-1}^{(k)}, \quad (4.13)$$

em que os parâmetros a serem estimados  $a^{(k)}$  e  $b^{(k)}$  possuem as restrições  $a^{(k)}, b^{(k)} \geq 0$  e  $a^{(k)} + b^{(k)} < 1$  para positividade de  $\mathbf{H}_t^{(\Delta_t)}$ . Assim como no caso univariado apresentado por Haas, Mittnik e Paoletta (2004), essa especificação corrige o problema de dependência do caminho, bem como aqui os parâmetros  $a^{(k)}$  e  $b^{(k)}$  têm a mesma interpretação do caso em que não há mudança de regime, porém aqui a interpretação se para o seu respectivo regime. A nomenclatura IS vem do termo em inglês *Independent Switching*, que faz referência ao fato de que aqui  $\mathbf{Q}_t^{(k)}$  depende somente de  $\mathbf{Q}_{t-1}^{(k)}$ .

É importante notar que  $\bar{\mathbf{Q}}$  é o mesmo para todos os regimes, podendo ser estimado de maneira acessível pela amostra, fazendo  $\bar{\mathbf{Q}} = \text{Cov}[\epsilon_{t-1}\epsilon'_{t-1}] = E[\epsilon_{t-1}\epsilon'_{t-1}]$ , por exemplo. E também, até o momento não encontramos na literatura as propriedades desse modelo de forma detalhada igual ao caso univariado. No entanto, vale destacar que a estacionariedade dos DCC's de cada regime é uma condição suficiente para a estacionariedade do IS-DCC.

#### 4.2.2 Estimação via filtro de Hamilton

A estimação do modelo IS(K)-DCC(1,1) por verossimilhança se dá também com o auxílio do filtro de Hamilton. Então, defina  $\boldsymbol{\xi}_t$  como sendo o vetor de probabilidades condicionais de estar em cada regime no tempo  $t$ , isto é  $\boldsymbol{\xi}_t = (P(\Delta_t = 1 | \mathcal{I}_t), \dots, P(\Delta_t = K | \mathcal{I}_t))'$  e  $\mathbf{Q}_t$  como sendo o vetor das matrizes de correlações condicionais de cada regime, ou seja,  $(\mathbf{Q}_t^{(1)}, \dots, \mathbf{Q}_t^{(K)})'$ . Então, dado os valores iniciais  $(\mathbf{y}_0, \mathbf{Q}_0, \boldsymbol{\xi}_0)'$  (ou  $(\mathbf{y}_1, \mathbf{Q}_1, \boldsymbol{\xi}_1)'$ ), o filtro de Hamilton procede da seguinte maneira:

- 1) Dado o vetor de probabilidades condicionais  $\boldsymbol{\xi}_{t-1}$  de estar em cada regime no tempo  $t - 1$ :

$$\boldsymbol{\xi}_{t-1} = \begin{bmatrix} P(\Delta_{t-1} = 1 | \mathcal{I}_{t-1}) \\ P(\Delta_{t-1} = 2 | \mathcal{I}_{t-1}) \\ \vdots \\ P(\Delta_{t-1} = k | \mathcal{I}_{t-1}) \end{bmatrix}. \quad (4.14)$$

2) Calcular matriz de covariância de cada regime:

$$\begin{aligned} \mathbf{Q}_t^{(k)} &= (1 - a^{(k)} - b^{(k)})\bar{\mathbf{Q}} + a^{(k)}\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1} + b^{(k)}\mathbf{Q}_{t-1}^{(k)} \\ \mathbf{R}_t^{(k)} &= \text{diag} \left\{ \left( \mathbf{Q}_t^{(k)} \right) \right\}^{-1/2} \mathbf{Q}_t^{(k)} \text{diag} \left\{ \left( \mathbf{Q}_t^{(k)} \right) \right\}^{-1/2} \\ \mathbf{H}_t^{(k)} &= \mathbf{D}_t \mathbf{R}_t^{(k)} \mathbf{D}_t \end{aligned} \quad (4.15)$$

Aqui, sob suposição de normalidade, calcular o vetor das densidades condicionais de cada regime.

$$\boldsymbol{\eta}_t = \begin{bmatrix} f(\mathbf{y}_t | \Delta_t = 1, \mathcal{I}_{t-1}) \\ f(\mathbf{y}_t | \Delta_t = 2, \mathcal{I}_{t-1}) \\ \vdots \\ f(\mathbf{y}_t | \Delta_t = K, \mathcal{I}_{t-1}) \end{bmatrix} = \frac{1}{(2\pi)^{m/2}} \begin{bmatrix} \frac{1}{|\mathbf{H}_t^{(1)}|^{1/2}} \exp \left\{ \frac{-1}{2} \mathbf{y}'_t \left( \mathbf{H}_t^{(1)} \right)^{-1} \mathbf{y}_t \right\} \\ \frac{1}{|\mathbf{H}_t^{(2)}|^{1/2}} \exp \left\{ \frac{-1}{2} \mathbf{y}'_t \left( \mathbf{H}_t^{(2)} \right)^{-1} \mathbf{y}_t \right\} \\ \vdots \\ \frac{1}{|\mathbf{H}_t^{(K)}|^{1/2}} \exp \left\{ \frac{-1}{2} \mathbf{y}'_t \left( \mathbf{H}_t^{(K)} \right)^{-1} \mathbf{y}_t \right\} \end{bmatrix}. \quad (4.16)$$

Aqui também temos que, apesar de a maioria das aplicações utilizar inovações com distribuição normal multivariada nos regimes, vale salientar que a suposição de normalidade não é a única opção. Em algumas aplicações o uso da t-Student multivariadas e outras distribuições são sugeridas por autores. E se os regimes não forem normais, mas leptocúrticos, o uso da normalidade dentro do regime pode afetar seriamente a identificação do processo do regime e, portanto, também processos GARCH dentro do regime. Isto é, o modelo com distribuição normal multivariada nos regimes tenderá a detectar mudanças de regime com muita frequência devido a uma observação atípica.

3) Do passo anterior podemos calcular a densidade condicional da t-ésima observação como sendo:

$$f(\mathbf{y}_t | \mathcal{I}_{t-1}) = \sum_{i=1}^k \sum_{j=1}^k p_{ij} P(\Delta_{t-1} = K | \mathcal{I}_{t-1}) \eta_{jt} = \mathbf{1}' (\mathbf{P} \boldsymbol{\xi}_{t-1} \odot \boldsymbol{\eta}_t), \quad (4.17)$$

em que  $\odot$  representa o produto de Hadamard e  $\mathbf{P}$  é a matriz de transição da cadeia de Markov.

4) Atualizar as probabilidades de estar em cada regime da seguinte forma:

$$\boldsymbol{\xi}_t = \frac{\mathbf{P} \boldsymbol{\xi}_{t-1} \odot \boldsymbol{\eta}_t}{\mathbf{1}' (\mathbf{P} \boldsymbol{\xi}_{t-1} \odot \boldsymbol{\eta}_t)}. \quad (4.18)$$

Como resultado da iteração de 1) a 4), podemos calcular a função de quase-verossimilhança condicional para cada  $t$ . Assim, a estimativa de  $\boldsymbol{\theta}$  é obtida maximizando a verossimilhança ou a log-verossimilhança através de algum método numérico. De tal forma que:

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \Theta} \sum_{t=2}^T \log f(\mathbf{y}_t | \mathcal{I}_{t-1}) = \arg \max_{\boldsymbol{\theta} \in \Theta} \sum_{t=2}^T \log(\mathbf{1}'(\mathbf{P}\boldsymbol{\xi}_{t-1} \odot \boldsymbol{\eta}_{jt})). \quad (4.19)$$

### 4.2.3 Estimação robusta

O efeito dos *outliers* nos filtros robustos para o caso multivariado pode ser estendido de maneira similar ao caso univariado através da notação apresentada por Calvet, Czellar e Ronchetti (2015) e descrita na subseção 3.3. Para o caso multivariado, Calvet, Czellar e Ronchetti (2015) também utilizam densidades robustas baseada na limitação da derivada do log da densidade de  $f(\mathbf{y}_t | \Delta_t = k, \mathcal{I}_{t-1})$ .

Portanto, de maneira semelhante ao caso univariado, iremos propor um estimador utiliza com as três modificações apresentadas anteriormente: o método BIP dado na Equação (4.7) para minimizar o efeito de *outliers* nas estimativas das correlações condicionais de cada regime; a utilização da distribuição t-Student multivariada com 4 graus de liberdade nas densidades condicionais; e uma função objetivo dada igual a do Boudt, Danielsson e Laurent (2013) no caso multivariado.

Aqui também utilizaremos uma modificação da log-verossimilhança combinada com a função  $\rho$  dada em (2.11). Assim sendo, temos que a estimativa de  $\boldsymbol{\theta}$  é obtida minimizando:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \Theta} \sum_{t=1}^T \log \left[ \sum_{i=1}^k \sum_{j=1}^k p_{ij} P(\Delta_{t-1} = i | \mathcal{I}_{t-1}) \frac{1}{|\mathbf{H}_t^{(j)}|^{1/2}} \left( 1 + \frac{\mathbf{y}_t' (\mathbf{H}_t^{(j)})^{-1} \mathbf{y}_t}{2} \right)^{\frac{-(N+4)*\sigma_{tN,4}}{2}} \right]. \quad (4.20)$$

em que  $\sigma_{tN,\nu}$  é dado por (2.13).

## 5 Simulação

Neste capítulo iremos discutir o desempenho dos estimadores dos modelos MS-GARCH e IS-DCC na presença ou não de *outliers* aditivos. Para ambos os modelos iremos utilizar 2 regimes e ordem (1,1) para os modelos GARCH e DCC nos dois regimes. Assim como cita Francq e Zakoian (2019, Seção 8.5, pág 235), o motivo para escolha da ordem dos modelos GARCH é dado pelo fato de que o modelo GARCH(1,1) é de longe o mais amplamente utilizado por profissionais que desejam estimar a volatilidade dos retornos diários; e esta prática é motivada pela crença comum de que o GARCH (1,1) é suficiente para capturar as propriedades de séries financeiras e modelos de ordem superior podem ser desnecessariamente complicados. Já, a escolha de 2 regimes é motivada devido que em aplicações para dados financeiros, por razões de parcimônia e interpretabilidade, muitas vezes é razoável restringir a priori o número de componentes para ser pequeno, por exemplo,  $k = 2$  (PAOLELLA et al., 2012).

### 5.1 Processos Geradores de Dados

Iremos considerar um tamanho de amostra de  $T = 3000$  e para cada modelo iremos gerar 1000 replicações Monte Carlo. E, dado a escolha  $p_{11} = 0,96$  e  $p_{22} = 0,98$ , obtemos que a distribuição estacionária da cadeia de Markov é  $\boldsymbol{\pi}_\infty = (1/3, 2/3)'$ . Então, com estes tamanhos amostrais, a média em cada regime são 1000 e 2000. Além disso, quando falamos de *outliers*, iremos considerar apenas o caso em que *outliers* aditivos contaminam os dados. Para o caso univariado, a série  $y_t^*$  observada, com *outliers* aditivos, é definida da seguinte forma:

$$y_t^* = y_t + \text{sign}(y_t) I_A(t) d_t \sqrt{h_t^{(\Delta t)}},$$

em que  $y_t$  é dado por (3.5),  $\text{sign}(\cdot)$  representa a função sinal,  $I_A(t)$  é uma variável indicadora representando a posição dos *outliers*, sendo  $A$  o conjunto de observações afetados por *outliers*,  $d_t$  é uma constante pertencente ao  $\mathbb{R}_+$  indicando o tamanho dos *outliers* em termos da volatilidade condicional, e  $\sqrt{h_t^{(\Delta t)}}$  está definido em (3.6). Já para o caso multivariado,

$$\mathbf{y}_t^* = \mathbf{y}_t + \text{sign}(\mathbf{y}_t) \mathbf{I}_{A_t}(t) \mathbf{d}_t,$$

em que  $\mathbf{y}_t$  é dado por (4.12). São utilizadas as notações  $\mathbf{I}_{A_t}(t)$  e  $\mathbf{d}_t$  porque no tempo  $t$  não necessariamente todas as séries são afetadas, e nem com *outliers* do mesmo tamanho. No caso multivariado, estamos interessado em avaliar o impacto dos *outliers* exclusivamente na estimação dos parâmetros do modelo DCC. Sendo assim, aqui geramos as séries utilizando a variância condicional igual a 1, assim como utiliza Boudt, Danielsson e Laurent (2013), por exemplo.

Alguns autores utilizam *outliers* isolados e consecutivos em pontos estratégicos, como no início, meio e fim da série, para avaliar os seus respectivos impactos. No entanto, neste trabalho iremos avaliar como os estimadores se comportam quando uma pequena fração de *outliers* contamina a série, assim como fazem Muler e Yohai (2008) e Boudt, Danielsson e Laurent (2013). Diante disso, avaliaremos os cenários em que a fração de *outliers*, digamos  $\epsilon$ , são dadas por  $\epsilon = 1\%$ ,  $\epsilon = 5\%$ ,  $\epsilon = 10\%$  e o caso em que não há *outliers*. É importante salientar que iremos fazer com que a posição dos *outliers* seja dada de maneira aleatória por meio de um processo de amostragem aleatória simples sem reposição. Por fim, iremos considerar dois valores para  $d$ , 3 e 5.

### 5.1.1 Caso univariado

O desempenho dos estimadores será analisado através de simulações Monte Carlo. Os dados foram gerados através de:

$$y_t = \sqrt{h_t^{(\Delta_t)}} z_t \text{ em que,} \quad (5.1)$$

$$h_t^{(1)} = 2,0 + 0,10 y_{t-1}^2 + 0,6 h_{t-1}^{(1)} \quad (5.2)$$

$$h_t^{(2)} = 0,3 + 0,35 y_{t-1}^2 + 0,2 h_{t-1}^{(2)}, \quad (5.3)$$

em que  $z_t \sim N(0, 1)$ ,  $p_{11} = 0,96$  e  $p_{22} = 0,98$ . Em cada replicação foi considerado um *burn-in* de 500 observações.

A Tabela 1 apresenta valores encontrados em aplicações do modelo seguindo a especificação apresentada por Haas, Mittnik e Paoletta (2004) e valores de simulações utilizando a especificação dada por (3.4). Diante disso, apesar das formas diferentes de definir o modelo, utilizaremos para as simulações os mesmos valores de parâmetros do processo gerador apresentado por Bauwens, Preminger e Rombouts (2010), Augustyniak (2014) e Billio, Casarin e Osuntuyi (2016).

A variância incondicional do primeiro regime é 10 vezes maior do que a do segundo. Isto é, temos que  $\alpha_0^{(1)}/(1 - \alpha_1^{(1)} - \beta_1^{(1)}) = 20/3$  e  $\alpha_0^{(2)}/(1 - \alpha_1^{(2)} - \beta_1^{(2)}) = 2/3$ . Portanto, a escolha dos parâmetros para o processo gerador de dados se justifica no ponto de vista teórico, devido ao fato de que quando estamos trabalhando com modelo de mudanças de regime, temos interesse em captar um regime de alta e de baixa volatilidade, então se ambos os regimes forem parecidos um modelo GARCH com um único regime poderia explicar a variância e ainda ser mais parcimonioso. Além disso, as escolhas desses parâmetros permitem evitar impasses na hora de identificar os regimes nas simulações.

Tabela 1 – Resultados das estimativas dos parâmetros de aplicações de aplicações apresentados por Haas, Mittnik e Paolella (2004) e Paolella et al. (2012); e processo gerador de dados (PGD) de simulações utilizadas por Augustyniak (2014), Billio, Casarin e Osuntuyi (2016) e Bauwens, Preminger e Rombouts (2010) para o modelo MS(2)-GARCH(1,1).  $V^{(i)} = \alpha_0^{(i)} / (1 - \alpha_1^{(i)} - \beta_1^{(i)})$  para  $\alpha_1^{(i)} + \beta_1^{(i)} < 1$  e  $\infty$  para  $\alpha_1^{(i)} + \beta_1^{(i)} \geq 1$ , EM representa o algoritmo EM e Bay. é estimação via técnicas bayesianas.

| Dados                                      | Estim.    | $\alpha_0^{(1)}$ | $\alpha_1^{(1)}$ | $\beta_1^{(1)}$ | $\alpha_0^{(2)}$ | $\alpha_1^{(2)}$ | $\beta_1^{(2)}$ | $p_{11}$ | $p_{22}$ | $V^{(1)}$ | $V^{(2)}$ |
|--|-----------|------------------|------------------|-----------------|------------------|------------------|-----------------|----------|----------|-----------|-----------|
| Aplicações Haas, Mittnik e Paolella (2004) |           |                  |                  |                 |                  |                  |                 |          |          |           |           |
| Yen  | QML-N     | 0,003            | 0,023            | 0,945           | 0,097            | 0,227            | 0,818           | 0,744    | 0,285    | 0,094     | $\infty$  |
| Libra                                      | QML-N     | 0,001            | 0,037            | 0,927           | 0,010            | 0,071            | 0,947           | 0,642    | 0,402    | 0,028     | $\infty$  |
| Singapura                                  | QML-N     | 0,001            | 0,049            | 0,902           | 0,004            | 0,106            | 0,952           | 0,873    | 0,501    | 0,020     | $\infty$  |
| Aplicações Paolella et al. (2012)          |           |                  |                  |                 |                  |                  |                 |          |          |           |           |
| França                                     | QML-N     | 0,008            | 0,028            | 0,959           | 0,204            | 0,066            | 0,934           | 0,991    | 0,909    | 0,615     | $\infty$  |
| França                                     | QML-t     | 0,009            | 0,028            | 0,960           | 0,235            | 0,078            | 0,915           | 0,993    | 0,928    | 0,692     | 29,375    |
| Alemanha                                   | QML-N     | 0,008            | 0,071            | 0,917           | 1,676            | 0,645            | 0,629           | 0,970    | 0,044    | 0,727     | $\infty$  |
| Alemanha                                   | QML-t     | 0,002            | 0,015            | 0,977           | 0,039            | 0,122            | 0,878           | 0,982    | 0,980    | 0,250     | $\infty$  |
| Libra                                      | QML-N     | 0,005            | 0,029            | 0,958           | 0,019            | 0,076            | 0,965           | 0,987    | 0,882    | 0,384     | $\infty$  |
| Libra                                      | QML-t     | 0,003            | 0,011            | 0,978           | 0,037            | 0,110            | 0,878           | 0,990    | 0,989    | 0,273     | 2,846     |
| Simulações                                 |           |                  |                  |                 |                  |                  |                 |          |          |           |           |
| PGD  | EM e Bay. | 2,00             | 0,100            | 0,600           | 0,300            | 0,350            | 0,200           | 0,960    | 0,980    | 0,667     | 6,667     |

### 5.1.2 Caso bivariado

No estudo do desempenho dos estimadores no caso bivariado, nas simulações Monte Carlo foi utilizado o modelo:

$$\mathbf{y}_t = \left( \mathbf{H}_t^{(\Delta t)} \right)^{1/2} \mathbf{z}_t, \tag{5.4}$$

$$\mathbf{Q}_t^{(1)} = (1 - 0,04 - 0,93)\bar{\mathbf{Q}} + 0,04\epsilon_{t-1}\epsilon'_{t-1} + 0,93\mathbf{Q}_{t-1}^{(1)} \tag{5.5}$$

$$\mathbf{Q}_t^{(2)} = (1 - 0,20 - 0,60)\bar{\mathbf{Q}} + 0,20\epsilon_{t-1}\epsilon'_{t-1} + 0,60\mathbf{Q}_{t-1}^{(2)}, \tag{5.6}$$

$$\tag{5.7}$$

em que  $\mathbf{z}_t \sim N_2(\mathbf{0}, \mathbf{I}_2)$ ,  $p_{11} = 0,96$ ,  $p_{22} = 0,98$  e  $\bar{\mathbf{Q}}$  é uma matriz  $2 \times 2$  com 1 na diagonal principal e 0,864 fora dela. Foi utilizado um *burn-in* de 3000 observações. O regime 1 é o regime de maior persistência, 0,97 contra 0,80 do regime 2. Além disso, no regime 2 temos uma maior influência na dinâmica da correlação condicional.

A Tabela 2 apresenta resultados da aplicação utilizada por Lee e Song (2009) do modelo IS(2)-DCC(1,1) para a série semanal dos preços à vista e futuros de algumas commodities. De maneira geral, não há algum padrão nas estimativas dos parâmetros. Mas, podemos notar que há um regime de “alta persistência” na correlação e um de baixa. Embora, em algumas aplicações foram encontrados alguns regimes com persistência muito baixa, para nossas simulações utilizaremos um regime com maior persistência, 0,93 e outro com menor persistência, 0,8.

Aqui também é importante citar sobre a escolha desse processo. De maneira

semelhante ao caso univariado, estamos interessados em regimes distintos entre si, pois se ambos os regimes forem parecidos um modelo DCC com um único regime poderia explicar a dinâmica de correlação. Por isso, foram escolhidos dois regimes com dinâmicas e variâncias não condicionais bem distintas.

Tabela 2 – Resultados das estimativas dos parâmetros de aplicações para o modelo IS(2)-DCC(1,1) apresentados por Lee e Song (2009).

|                     | Estim. | $a_1^{(1)}$ | $b_1^{(1)}$ | $a_1^{(2)}$ | $b_1^{(2)}$ | $p_{11}$ | $p_{22}$ |
|---------------------|--------|-------------|-------------|-------------|-------------|----------|----------|
| Trigo               | QML-N  | 0,021       | 0,299       | 0,467       | 0,533       | -        | -        |
| Milho               | QML-N  | 0,048       | 0,175       | 0,342       | 0,658       | -        | -        |
| Cacau               | QML-N  | 0,371       | 0,627       | 0,029       | 0,962       | -        | -        |
| Café                | QML-N  | 0,216       | 0,003       | 0,113       | 0,887       | -        | -        |
| Óleo cru            | QML-N  | 0,000       | 0,010       | 0,455       | 0,543       | -        | -        |
| Gás natural         | QML-N  | 0,097       | 0,135       | 0,000       | 0,858       | -        | -        |
| Óleo de aquecimento | QML-N  | 0,006       | 0,000       | 0,444       | 0,544       | -        | -        |
| Platina             | QML-N  | 0,132       | 0,000       | 0,273       | 0,724       | -        | -        |

## 5.2 Estimadores Considerados

Sabemos que, um estimador robusto deve ter uma eficiência razoavelmente boa (ótima ou quase ótima) no modelo assumido; deve ser robusto no sentido de que pequenos desvios das suposições do modelo devem prejudicar o desempenho apenas ligeiramente; e desvios um pouco maiores do modelo não devem causar uma catástrofe. Vimos que no processo de estimação podemos “robustificar” diversos passos, desde o filtro até a função objetivo. Assim, para avaliar a performance dos estimadores, consideramos 3 modelos possíveis: QML-normal, QML-t definidos nas Seções 3.2 e 4.2.2 para o caso univariado e multivariado, respectivamente, e o Robusto (Rob) definido nas Seções 3.3 e 4.2.3 para o caso univariado e multivariado, respectivamente.

Para o caso univariado, foi feita uma alteração na função  $r_{k_{\delta,1}}(u)$ , definida na Equação (2.9), para que os estimadores produzam resultados melhores na estimação dos parâmetros. Na literatura, alguns autores propõem substituir o próprio valor do  $k_{\delta,1}$  quando  $x > k_{\delta,1}$  como Muler e Yohai (2008) e Boudt, Danielsson e Laurent (2013), por exemplo. Entretanto, outros sugerem utilizar 1 (valor esperado) quando  $x > k_{\delta,1}$  como, por exemplo, Carnero, Peña e Ruiz (2012). Este valor corresponde à aproximadamente  $k_{0,69,1}$ . Nesta dissertação sugerimos a utilização do ponto médio entre o  $\delta = 0,95$  e  $0,69$ . Assim, quando  $x > k_{\delta,1}$ , substituiremos por  $k_{0,82,1}$ . Já para o caso multivariado, utilizaremos  $\delta = 0,95$  para a função  $w_{k_{\delta,2}}(u)$ .



### 5.3 Valores Iniciais

Como comenta Paoella et al. (2012), pode haver vários máximos locais “plausíveis” da função de verossimilhança do modelo MS-GARCH. Uma alternativa para lidar com isso é através do uso de vários valores iniciais para encontrar o máximo global. Assim, na nossa simulação, para cada replicação em que avaliamos a performance dos estimadores, levamos em consideração diversos valores iniciais distintos e para cada um deles avaliamos a função de verossimilhança. Para os 3 valores iniciais que gerarem a maior verossimilhança (caso seja uma função de perda, o menor), estimamos os parâmetros. Essa simulação está sendo feita utilizando o software *R* e a otimização é feita usando a função *solnp*.

Portanto, para cada replicação teremos três estimativas de parâmetros geradas por três valores iniciais distintos. Dentre esses três valores, utilizamos o seguinte mecanismo para determinar qual estimativa é correspondente ao máximo global. O mecanismo segue a seguinte ideia:

- Quando as 3 estimativas convergem para o mesmo lugar, isto é, os parâmetros estimados são os todos iguais e a verossimilhança final encontrada for a mesma, selecionar o respectivo valor;
- Quando as 3 estimativas não convergem para o mesmo lugar:
  - Selecionar o que possui maior verossimilhança final;
  - Caso não haja diferença entre as verossimilhanças, escolher a média das 3 estimativas.

Aqui, consideramos que os parâmetros e a verossimilhança final sejam diferentes se existir uma diferença maior que  $10^{-4}$  ou nos parâmetros ou na verossimilhança final, exceto no intercepto do regime de alta volatilidade, que será diferente se existir uma diferença maior que  $10^{-2}$ .

Esses casos em que as 3 estimativas não convergem para o mesmo lugar são de certa forma exceções. Para exemplificar isso geramos 1000 replicações Monte Carlo do processo definido em (5.3) e geramos o mecanismo acima utilizando a estimativa de máxima verossimilhança. Dessas replicações, apenas 27 (2,7%) não convergiram para o mesmo lugar. E, desses casos, em apenas 14 (1,4%) convergiram para o mesmo valor de verossimilhança, mesmo com valores de parâmetros diferentes.

De maneira análoga, seguimos esse mecanismo para a utilização dos valores iniciais da estimação do modelo IS-DCC. A única diferença é que aqui não teremos um parâmetro representando o intercepto. Assim, consideramos que os parâmetros e a verossimilhança final sejam diferentes se existir uma diferença maior que  $10^{-4}$  ou nos parâmetros ou na verossimilhança final.

## 5.4 Resultados

Nessa seção avaliaremos o desempenho dos estimadores para os processos geradores de dados considerados. A avaliação será dada de duas formas. A primeira, apresentada na Seção 5.4.1, será em relação ao desempenho do filtro de Hamilton. A segunda, apresentada na Seção 5.4.2, será sobre a precisão na estimação dos parâmetros do modelo. A análise será dividida para o caso univariado e multivariado.

### 5.4.1 Análise do desempenho dos filtros

Na geração dos dados, esperamos que 1/3 deles estejam em um regime o restante no outro. No entanto, checar apenas a distribuição estacionária e verificar a proporção não necessariamente é o ideal, pois a proporção pode estar correta em lugares diferentes. Assim, como estamos lidando com dados simulados, conseguimos saber, para cada simulação, em que regime uma determinada observação se encontra. Ademais, como discutido anteriormente, via filtro de Hamilton conseguimos estimar a probabilidade de estar ou não em um determinado regime. A Figura 1 apresenta para uma série simulada contaminada com 1% de *outliers* com  $d = 5$  do modelo MS-GARCH, em que o processo gerador dos dados é dado pela Equação (5.3), as probabilidades estimadas pelo filtro para os três estimadores. As linhas vermelhas indicam as posições dos *outliers*. Para essa série podemos ver, por exemplo, que, quando os *outliers* ocorrem em um regime de baixa volatilidade, as estimativas dos estimadores robusto e QML-t foram pouco afetados. Já, para o estimador QML-n, as probabilidades estimadas de estar no regime de alta volatilidade são altas.

Portanto, para capturar esse acerto de classificação, analisamos a inferência das probabilidades para as observações contaminadas por *outliers* e para série toda em geral. Para isso, estabelecemos com um limiar de 0,5 para classificar se uma observação está em um regime ou não. Por exemplo, caso a probabilidade inferida de uma observação estar no regime 1 seja 0,6, dizemos que naquele instante estamos no regime 1 e assim comparamos com o que foi simulado.

#### 5.4.1.1 Caso univariado

A Figura 2 apresenta a proporção de classificação correta, nas observações contaminadas, dos diferentes estimadores para o modelo MS-GARCH, em que o processo gerador dos dados é dado pela Equação (5.3). Como esperado, o desempenho dos filtros é melhor no regime de alta volatilidade. No regime de alta volatilidade, com baixa fração de contaminação (1%), tanto os estimadores QML quanto o estimador robusto, classificam corretamente o regime de alta volatilidade em mais de 84% das observações. Também, como esperado, esta performance tende a melhorar com maior contaminação e maior

tamanho do *outlier* já que o efeito do *outlier* é no sentido de classificar a observação como regime de alta volatilidade. Assim, quando a fração de *outliers* é de 5%, nos três métodos a percentagem de classificação correta em aproximadamente 85% dos casos.

Para o caso do regime da baixa volatilidade, com 1% de contaminação, quando se utiliza o estimador QML-n, a percentagem de classificação correta é menor do que 14% dos casos, mesmo com apenas com tamanho de *outlier* igual a 3, em contraste com os estimadores QML-t e Robusto, quando o acerto é maior do que 75% quando  $d = 3$  e maior do que 48% quando  $d = 5$ . Quando o grau de contaminação aumenta para 5%, o desempenho dos estimadores piora. O desempenho dos estimadores QML-t e robusto ainda continuam razoáveis para  $d = 3$  com mais de 60%, mas menores do que 17% quando  $d = 5$ . Por fim, quando a proporção de *outliers* é de 10% e o  $d = 3$ , podemos observar que o estimador robusto tem desempenho semelhante com contaminação de 5% e melhor do que o QM-t. Porém, quando  $d = 5$  todos tem desempenho ruim.

De forma geral pode-se dizer que os três estimadores têm bom desempenho no regime de alta volatilidade. No regime de baixa volatilidade o filtro com o estimador QML-n tem péssimo desempenho e que, quando  $d = 3$  o filtro com estimador robusto tem desempenho razoável, melhor do que com o estimador QML-t. Já, com tamanho do *outlier* igual a  $d = 5$  o desempenho do filtro utilizado os três estimadores são muito ruins quando a contaminação é maior ou igual a 5%

A Figura 3 apresenta a proporção de classificação correta dos diferentes estimadores para o modelo MS-GARCH para todas as observações. Para o regime de baixa volatilidade, observamos que todos os estimadores tendem a classificar bem todas as observações. No geral, para esse regime, podemos observar que o estimador robusto tem um desempenho ligeiramente superior aos estimadores por QML. Já para o regime de alta volatilidade, podemos observar um desempenho superior do estimador QML-t e Robusto em relação ao estimador QML-n, exceto para o caso em a proporção de *outliers* é de 10% e  $d = 5$ .

#### 5.4.1.2 Caso bivariado

A Figura 4 a proporção de classificação correta dos diferentes estimadores para o modelo IS-DCC em que o processo gerador dos dados é dado pela Equação (5.3), nas observações contaminadas. De maneira geral, utilizando tanto o estimador QML quanto o estimador robusto, ambos os casos classificam corretamente o regime 1 de alta persistência para mais de 75% para os dois valores de  $d$ . No entanto, para o caso do regime 2, de menor persistência, notamos que o estimador QML-n tem desempenho ruim para ambos os valores de  $d$ . O estimador QML-t tem um desempenho razoável quando  $d = 3$  e a percentagem de *outliers* é de 1% ou 5%, mas insatisfatório quando  $d = 5$  ou a percentagem é de 10%. Por fim, o estimador robusto tem um desempenho razoável todos os cenários,

exceto quando  $d = 5$  e a porcentagem é de 5% ou 10%.

A Figura 5 apresenta a proporção de classificação correta dos diferentes estimadores para o modelo IS-DCC para todas as observações. Para o regime de 1, todos os estimadores têm desempenho razoável, porém não tão bom quanto no caso univariado. E, de maneira geral, todos eles tendem a classificar mal o regime 2 independente do cenário considerado. Em geral o estimador robusto tem melhor desempenho nos regimes de alta persistência (regime 1), enquanto o estimador QML-t no outro regime.

## 5.4.2 Estimação dos modelos

### 5.4.2.1 Caso univariado

A Tabela 3 mostra o viés e o RMSE das estimativas, enquanto as Figuras 6 e 7 apresentam os boxplots para as estimativas dos parâmetros do modelo MS-GARCH para o regime de alta e baixa volatilidade, respectivamente, das replicações Monte Carlo. É esperado que na ausência de *outliers*,  $\epsilon = 0\%$ , o estimador QML-n possua um desempenho melhor que os demais. E conforme a presença de *outliers* aumenta, esperamos que essa relação de performance se inverta, o que foi confirmado nas simulações. Em relação aos parâmetros das equações das variâncias do modelo MS-GARCH, observamos que, em termos de RMSE, conforme a fração e o tamanho dos *outliers* aumentam, pior o desempenho dos estimadores. Além disso, no geral, o estimador QML-n tem desempenho pior que o estimador QML-t e Robusto. Observamos também que o estimador robusto se destaca na maior parte dos casos na estimação desses parâmetros.

No regime de alta volatilidade (Tabela 3 e 6) mostra que o desempenho do estimador QML-n, a partir de 5% com  $d = 3$  e 1% com  $d = 5$  deteriora bastante. Quando  $d = 3$  o desempenho do estimador robusto é melhor que o QML-t. Quando  $d = 5$ , a partir de 5% as estimativas dos parâmetros  $\beta_1$ ,  $p_{11}$  e  $p_{22}$  deteriora bastante para todos os estimadores. Quando  $d = 5$ , para os parâmetros  $\alpha_0$  e  $\alpha_1$ , não temos vantagem em relação aos estimadores.

Em relação ao regime de baixa volatilidade (Tabela 3 e 7), podemos observar que o estimador robusto e o QML-t têm desempenho superior ao QML-n, exceto para o parâmetro  $\alpha_1$  quando há 1% de contaminação por *outliers*. Além disso, podemos observar que o estimador robusto é melhor que o estimador QML-t.

Por fim, com relação aos estimadores das probabilidades da matriz de transição relacionado a cadeia de Markov, observamos que o estimador QML-t foi o que obteve melhores resultados. Notamos que o RMSE do estimador QML-t é metade do estimador Robusto, exceto quando a proporção de *outliers* é de 10%. A partir de 5% de contaminação o desempenho do estimador QML-n são bastante ruins. Nota-se também que, como esperado pela Figura , o estimador QML-n tende a não se manter no mesmo regime. Esse efeito

no estimador QML-n pode ser observado a partir de 5% de contaminação, enquanto nos outros dois estimadores podemos observar a partir de 10% e  $d = 5$ .

Tabela 3 – Viés e RMSE, para os estimadores do modelo MS(2)-GARCH(1,1) com parâmetros  $\alpha_0^{(1)} = 2, \alpha_1^{(1)} = 0, 1, \beta_1^{(1)} = 0, 6, \alpha_0^{(2)} = 0, 3, \alpha_1^{(2)} = 0, 35, \beta_1^{(2)} = 0, 2, p_{11} = 0, 96, p_{22} = 0, 98$ ,  $\text{per}^{(i)}$  é a persistência do regime  $i$ ,  $^{(1)}$  representa o regime de alta volatilidade e  $^{(2)}$  representa o regime de baixa volatilidade e  $T = 3000$ . Em negrito são os que apresentaram os melhores resultados.

|                                 |      | Estim. | $\alpha_0^{(1)}$ | $\alpha_1^{(1)}$ | $\beta_1^{(1)}$ | $\alpha_0^{(2)}$ | $\alpha_1^{(2)}$ | $\beta_1^{(2)}$ | $p_{11}$      | $p_{22}$      | $\text{per}^{(1)}$ | $\text{per}^{(2)}$ |
|---------------------------------|------|--------|------------------|------------------|-----------------|------------------|------------------|-----------------|---------------|---------------|--------------------|--------------------|
| $\varepsilon = 0\%$             | Viés | QML-N  | 0,247            | <b>0,002</b>     | -0,045          | <b>0,000</b>     | <b>-0,002</b>    | <b>0,003</b>    | <b>-0,001</b> | <b>0,000</b>  | -0,042             | <b>0,001</b>       |
|                                 |      | QML-t  | 0,789            | 0,051            | -0,032          | 0,001            | 0,197            | 0,172           | 0,010         | 0,007         | <b>0,019</b>       | 0,369              |
|                                 |      | Rob    | <b>0,018</b>     | 0,043            | <b>-0,012</b>   | -0,041           | 0,146            | 0,082           | -0,017        | 0,012         | 0,030              | 0,228              |
|                                 | RMSE | QML-N  | 1,231            | <b>0,049</b>     | 0,213           | <b>0,041</b>     | <b>0,056</b>     | <b>0,096</b>    | <b>0,008</b>  | <b>0,005</b>  | 0,190              | <b>0,078</b>       |
|                                 |      | QML-t  | 2,110            | 0,132            | 0,244           | 0,067            | 0,228            | 0,220           | 0,013         | 0,030         | 0,217              | 0,376              |
|                                 |      | Rob    | <b>0,960</b>     | 0,082            | <b>0,172</b>    | 0,064            | 0,163            | 0,118           | 0,020         | 0,013         | <b>0,144</b>       | 0,241              |
| $\varepsilon = 1\%$<br>$d = 3$  | Viés | QML-N  | 0,197            | <b>0,002</b>     | -0,021          | 0,036            | <b>-0,050</b>    | <b>0,015</b>    | -0,014        | -0,011        | -0,019             | <b>-0,036</b>      |
|                                 |      | QML-t  | 1,111            | 0,013            | -0,028          | 0,077            | 0,163            | 0,114           | <b>0,008</b>  | <b>0,006</b>  | <b>-0,015</b>      | 0,277              |
|                                 |      | Rob    | <b>0,072</b>     | 0,040            | <b>-0,003</b>   | <b>-0,026</b>    | 0,143            | 0,083           | -0,018        | 0,012         | 0,037              | 0,226              |
|                                 | RMSE | QML-N  | 1,737            | <b>0,081</b>     | 0,272           | 0,074            | <b>0,110</b>     | 0,219           | 0,019         | 0,015         | 0,229              | <b>0,156</b>       |
|                                 |      | QML-t  | 2,511            | 0,112            | 0,262           | 0,108            | 0,204            | 0,193           | <b>0,012</b>  | <b>0,008</b>  | 0,232              | 0,293              |
|                                 |      | Rob    | <b>1,060</b>     | 0,083            | <b>0,178</b>    | <b>0,058</b>     | 0,163            | <b>0,125</b>    | 0,021         | 0,013         | <b>0,149</b>       | 0,241              |
| $\varepsilon = 1\%$<br>$d = 5$  | Viés | QML-N  | 0,347            | <b>0,001</b>     | 0,037           | <b>-0,012</b>    | <b>-0,059</b>    | 0,100           | -0,075        | -0,024        | <b>0,038</b>       | <b>0,042</b>       |
|                                 |      | QML-t  | 1,796            | -0,014           | -0,101          | 0,146            | 0,130            | <b>0,053</b>    | <b>0,006</b>  | <b>0,002</b>  | -0,115             | 0,183              |
|                                 |      | Rob    | <b>-0,020</b>    | 0,045            | <b>0,007</b>    | -0,023           | 0,140            | 0,080           | -0,019        | 0,012         | 0,051              | 0,220              |
|                                 | RMSE | QML-N  | 2,724            | 0,166            | 0,293           | 0,094            | <b>0,142</b>     | 0,313           | 0,184         | 0,029         | 0,208              | <b>0,203</b>       |
|                                 |      | QML-t  | 3,270            | 0,139            | 0,314           | 0,177            | 0,204            | 0,229           | <b>0,011</b>  | <b>0,007</b>  | 0,297              | 0,223              |
|                                 |      | Rob    | <b>1,008</b>     | <b>0,088</b>     | <b>0,172</b>    | <b>0,057</b>     | 0,160            | <b>0,121</b>    | 0,022         | 0,013         | <b>0,149</b>       | 0,235              |
| $\varepsilon = 5\%$<br>$d = 3$  | Viés | QML-N  | -0,863           | 0,069            | 0,222           | -0,206           | -0,285           | 0,518           | -0,309        | -0,122        | 0,291              | 0,233              |
|                                 |      | QML-t  | 1,482            | <b>-0,008</b>    | 0,066           | 0,309            | <b>-0,112</b>    | 0,100           | <b>0,001</b>  | <b>-0,005</b> | -0,073             | <b>-0,012</b>      |
|                                 |      | Rob    | <b>0,450</b>     | 0,037            | <b>0,010</b>    | <b>0,044</b>     | 0,122            | <b>0,091</b>    | -0,022        | 0,012         | <b>0,048</b>       | 0,213              |
|                                 | RMSE | QML-N  | <b>1,058</b>     | 0,102            | 0,231           | 0,211            | 0,288            | 0,528           | 0,340         | 0,124         | 0,295              | 0,246              |
|                                 |      | QML-t  | 3,316            | 0,104            | 0,309           | 0,364            | 0,185            | 0,332           | <b>0,011</b>  | <b>0,012</b>  | 0,270              | <b>0,230</b>       |
|                                 |      | Rob    | 1,463            | <b>0,098</b>     | <b>0,193</b>    | <b>0,085</b>     | <b>0,159</b>     | <b>0,166</b>    | 0,026         | 0,014         | <b>0,159</b>       | 0,234              |
| $\varepsilon = 5\%$<br>$d = 5$  | Viés | QML-N  | 2,868            | 0,099            | 0,191           | -0,191           | -0,279           | 0,513           | -0,559        | -0,116        | 0,290              | 0,234              |
|                                 |      | QML-t  | <b>-0,099</b>    | -0,080           | <b>0,132</b>    | 0,029            | -0,340           | 0,516           | <b>-0,035</b> | -0,036        | <b>0,052</b>       | <b>0,176</b>       |
|                                 |      | Rob    | -0,583           | <b>0,024</b>     | 0,138           | <b>0,010</b>     | <b>0,056</b>     | <b>0,168</b>    | -0,093        | <b>-0,007</b> | 0,162              | 0,224              |
|                                 | RMSE | QML-N  | 5,481            | 0,169            | <b>0,234</b>    | 0,199            | 0,289            | 0,535           | 0,585         | 0,116         | 0,290              | <b>0,248</b>       |
|                                 |      | QML-t  | 2,497            | <b>0,086</b>     | 0,318           | 0,244            | 0,342            | 0,589           | <b>0,044</b>  | 0,039         | 0,287              | 0,327              |
|                                 |      | Rob    | <b>1,398</b>     | 0,126            | 0,264           | <b>0,124</b>     | <b>0,149</b>     | <b>0,283</b>    | 0,180         | <b>0,027</b>  | <b>0,215</b>       | 0,271              |
| $\varepsilon = 10\%$<br>$d = 3$ | Viés | QML-N  | -0,508           | 0,099            | 0,195           | -0,227           | -0,312           | 0,552           | -0,413        | -0,221        | 0,295              | 0,240              |
|                                 |      | QML-t  | <b>0,440</b>     | -0,028           | 0,097           | <b>0,155</b>     | -0,319           | 0,422           | -0,033        | -0,043        | 0,069              | <b>0,103</b>       |
|                                 |      | Rob    | 0,884            | <b>0,023</b>     | <b>0,041</b>    | 0,160            | <b>0,051</b>     | <b>0,118</b>    | <b>-0,028</b> | <b>0,008</b>  | <b>0,063</b>       | 0,169              |
|                                 | RMSE | QML-N  | <b>0,658</b>     | 0,107            | <b>0,199</b>    | 0,229            | 0,312            | 0,555           | 0,419         | 0,222         | 0,295              | 0,245              |
|                                 |      | QML-t  | 3,224            | <b>0,055</b>     | 0,301           | 0,418            | 0,322            | 0,534           | 0,049         | 0,061         | 0,298              | 0,321              |
|                                 |      | Rob    | 2,304            | 0,112            | 0,242           | <b>0,206</b>     | <b>0,137</b>     | <b>0,236</b>    | <b>0,036</b>  | <b>0,016</b>  | <b>0,189</b>       | <b>0,214</b>       |
| $\varepsilon = 10\%$<br>$d = 5$ | Viés | QML-N  | 2,761            | 0,081            | 0,211           | -0,227           | -0,333           | 0,606           | -0,576        | -0,188        | 0,292              | 0,273              |
|                                 |      | QML-t  | <b>0,166</b>     | <b>0,013</b>     | 0,275           | -0,205           | -0,333           | 0,625           | <b>-0,367</b> | -0,172        | <b>0,288</b>       | <b>0,292</b>       |
|                                 |      | Rob    | 1,946            | 0,085            | <b>0,205</b>    | <b>-0,184</b>    | <b>0,112</b>     | <b>0,314</b>    | -0,630        | <b>-0,148</b> | 0,290              | 0,426              |
|                                 | RMSE | QML-N  | 3,023            | 0,090            | <b>0,215</b>    | 0,228            | 0,333            | 0,607           | 0,578         | 0,188         | 0,292              | <b>0,275</b>       |
|                                 |      | QML-t  | <b>0,950</b>     | <b>0,036</b>     | 0,277           | 0,207            | 0,333            | 0,626           | <b>0,379</b>  | 0,173         | <b>0,289</b>       | 0,294              |
|                                 |      | Rob    | 2,974            | 0,136            | 0,231           | <b>0,186</b>     | <b>0,129</b>     | <b>0,321</b>    | 0,634         | <b>0,149</b>  | 0,290              | 0,427              |

### 5.4.2.2 Caso bivariado

A Tabela 4 mostra o viés e o RMSE e a Figuras 8 e 9 apresenta os boxplots das estimativas dos parâmetros do modelo IS-DCC das replicações Monte Carlo para os estimadores estudados. Quando não há *outliers*, espera-se que o estimador de QML-n possua um desempenho melhor que os estimadores QML-t e Robusto. Entretanto, podemos observar que na estimação dos parâmetros da matriz de transição da cadeia de Markov, essa relação se inverte. Em comparação, o RMSE do estimador QML-t, é a metade da do estimador QML-n.

Na presença de *outliers*, notamos que, em relação as estimativas dos parâmetros das correlações de cada regime, o estimador robusto tende a ter resultados melhores que os demais, em relação a viés e RMSE. Em relação aos parâmetros da matriz de transição da cadeia de Markov, observamos que o estimador QML-t tem desempenho melhor que os demais.

De maneira mais detalhada, notamos que tanto o estimador QML-n quanto o QML-t tem desempenho ruim quando a contaminação é a partir de 1%. Além disso, podemos notar que esse efeito negativo é mais acentuado no regime de alta correlação, onde as estimativas apresentam uma maior variabilidade. Em relação ao regime de alta persistência, podemos observar que a partir de 5% de contaminação a maioria das estimativas de  $a_1$  para os estimadores QML tendem a ir para próximo de 0, enquanto o robusto apresenta bons resultados. Já  $b_1$  vemos que a maioria das estimativas são bem menores que os valores reais dos parâmetros.

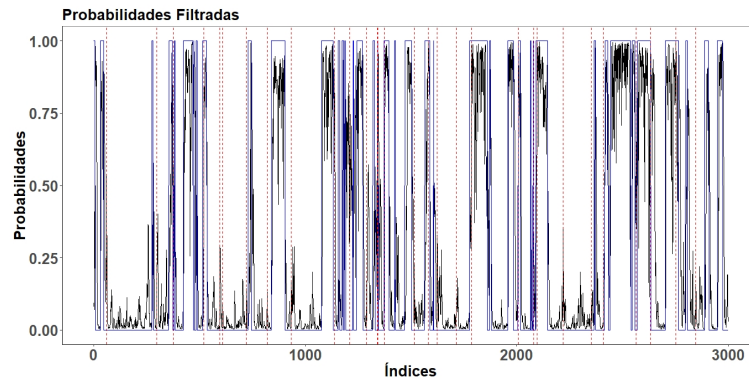
Em relação ao regime de baixa persistência, na presença de *outliers*, o estimador QML-n tem desempenho ruim. Para o parâmetro  $a_1$ , o estimador QML-t e o robusto tem desempenho parecido, exceto com contaminação de 10%, quando o estimador QML-t apresenta menor variabilidade. Para o parâmetro  $b_1$ , podemos notar que o estimador QML-t não apresenta bons resultados, mesmo na ausência de *outliers*, e só piora quando a contaminação aumenta. Embora o estimador robusto também não apresente bons resultados na ausência de *outliers*, seu desempenho é muito melhor que os outros dois estimadores, para o parâmetro  $b_1$ .

## 5.5 Conclusões

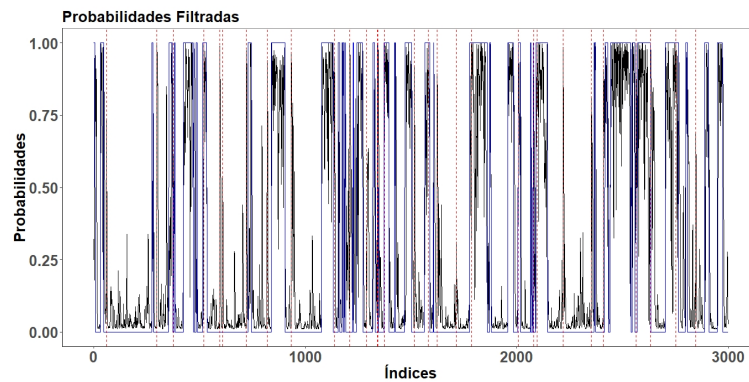
Na literatura podemos encontrar alguns resultados de simulações Monte Carlo em relação a estimação dos parâmetros do modelo MS-GARCH. Augustyniak (2014) apresenta os resultados das simulações utilizando a especificação dada por Gray (1996); Billio, Casarin e Osuntuyi (2016) mostra os resultados das simulações estimando os parâmetros estimando via métodos bayesianos. Em ambos os casos podemos notar que há uma dificuldade em estimar os parâmetros  $\alpha_0^{(1)}$ ,  $\beta_1^{(1)}$  e  $\beta_1^{(2)}$ , mesmo na ausência de *outliers*.

Em nossas simulações também notamos que, de maneira geral, esses parâmetros foram os mais problemáticos. Note também, que o mecanismo BIP tende a funcionar melhor quando a porcentagem de *outliers* é mais alta. Além disso, vimos que o estimador QML-t teve um bom desempenho na estimação das probabilidades da matriz de transição da cadeia de Markov.

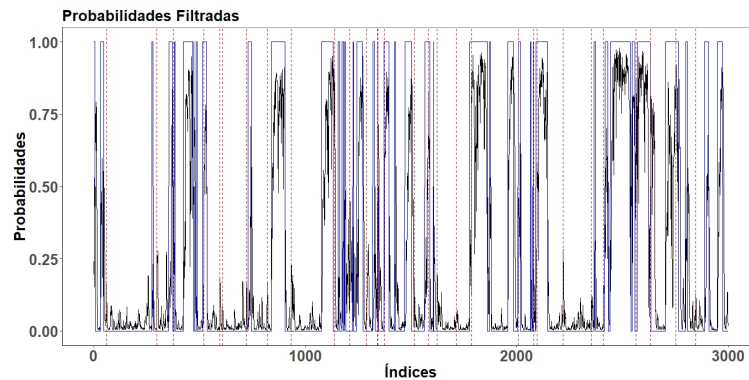
Em relação ao modelo IS-DCC, não encontramos resultados na literatura para a comparação. No entanto notamos que devido a especificação do modelo, principalmente pelo fato de  $\bar{Q}$  ser o mesmo para ambos os regimes, há uma certa dificuldade em identificar os regimes. Além disso, observamos que, na presença de *outliers*, o estimador robusto teve bom desempenho na estimação dos parâmetros da equação da correlação. Aqui também o estimador QML-t teve um bom desempenho na estimação das probabilidades da matriz de transição da cadeia de Markov.



(a) Robusto



(b) QML-n



(c) QML-t

Figura 1 – Probabilidades filtradas para uma série simulada de tamanho 3000 contaminada com 30 *outliers* de tamanho 5 vezes o desvio padrão condicional do modelo MS(2)-GARCH(1,1) com parâmetros  $\alpha_0^{(1)} = 2$ ,  $\alpha_1^{(1)} = 0, 1$ ,  $\beta_1^{(1)} = 0, 6$ ,  $\alpha_0^{(2)} = 0, 3$ ,  $\alpha_1^{(2)} = 0, 35$ ,  $\beta_1^{(2)} = 0, 2$ ,  $p_{11} = 0, 96$ ,  $p_{22} = 0, 98$ . A linha preta representa a probabilidade filtrada, a linha azul representa o estado real e a linha tracejada em vermelho representa a posição do *outlier*. O 1 representa o regime de alta volatilidade, isto é, <sup>(1)</sup> e 0 representa o regime de baixa volatilidade, isto é, <sup>(2)</sup>.



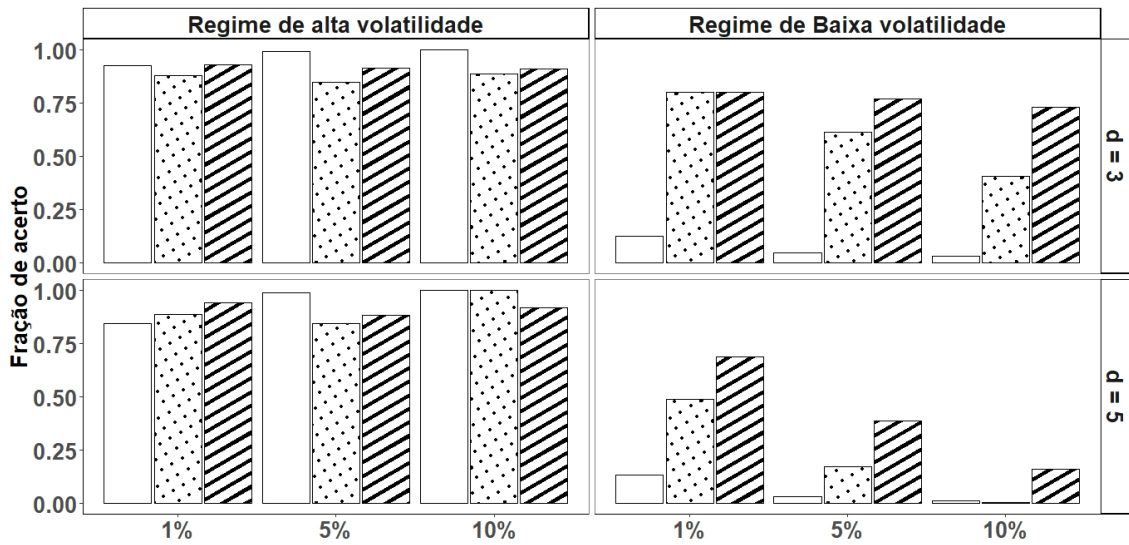


Figura 2 – Proporção de classificação correta dos diferentes estimadores para o modelo MS-GARCH nas observações contaminadas. No eixo horizontal estão as porcentagens de *outliers*. Preenchido em branco o estimador QML-n, com bolas o estimador QML-t e em linhas diagonais o estimador Robusto. Os quadros da esquerda são referentes à quando o regime correto é de alta volatilidade, ou seja, o regime (1), os da direita para o regime de baixa volatilidade, regime (2). Nos quadros superiores são *outliers* de tamanho 3 vezes o desvio padrão condicional, nos inferiores 5 vezes o desvio padrão condicional.

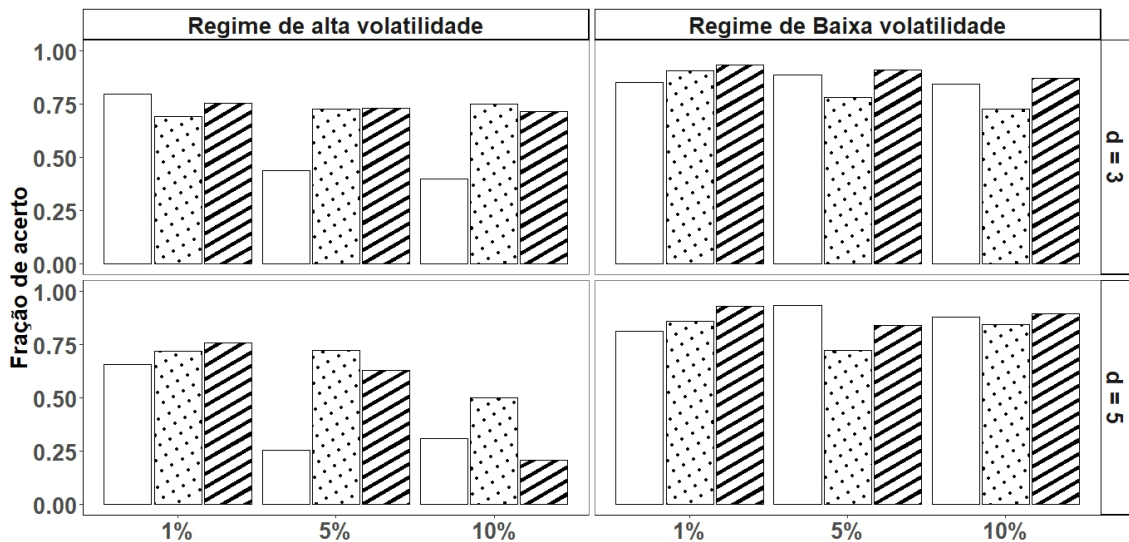


Figura 3 – Proporção de classificação correta dos diferentes estimadores para o modelo MS-GARCH para todas as observações. No eixo horizontal estão as porcentagens de *outliers*. Preenchido em branco o estimador QML-n, com bolas o estimador QML-t e em linhas diagonais o estimador Robusto. Os quadros da esquerda são referentes à quando o regime correto é de alta volatilidade, ou seja, o regime (1), os da direita para o regime de baixa volatilidade, regime (2). Nos quadros superiores são *outliers* de tamanho 3 vezes o desvio padrão condicional, nos inferiores 5 vezes o desvio padrão condicional.

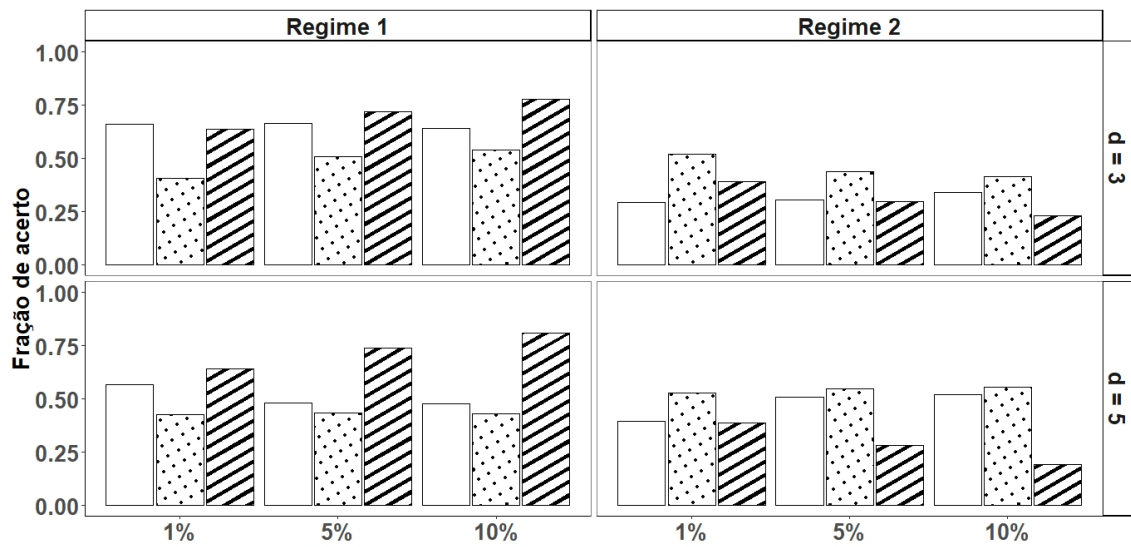


Figura 4 – Proporção de classificação correta dos diferentes estimadores para o modelo IS-DCC nas observações contaminadas. No eixo horizontal estão as porcentagens de *outliers*. Preenchido em branco o estimador QML-n, com bolas o estimador QML-t e em linhas diagonais o estimador Robusto. Os quadros da esquerda são referentes à quando o regime correto é de maior persistência na correlação, ou seja, o regime (1), os da direita para o regime de menor persistência na correlação, regime (2). Nos quadros superiores são *outliers* de tamanho 3 vezes o desvio padrão condicional, nos inferiores 5 vezes o desvio padrão condicional.

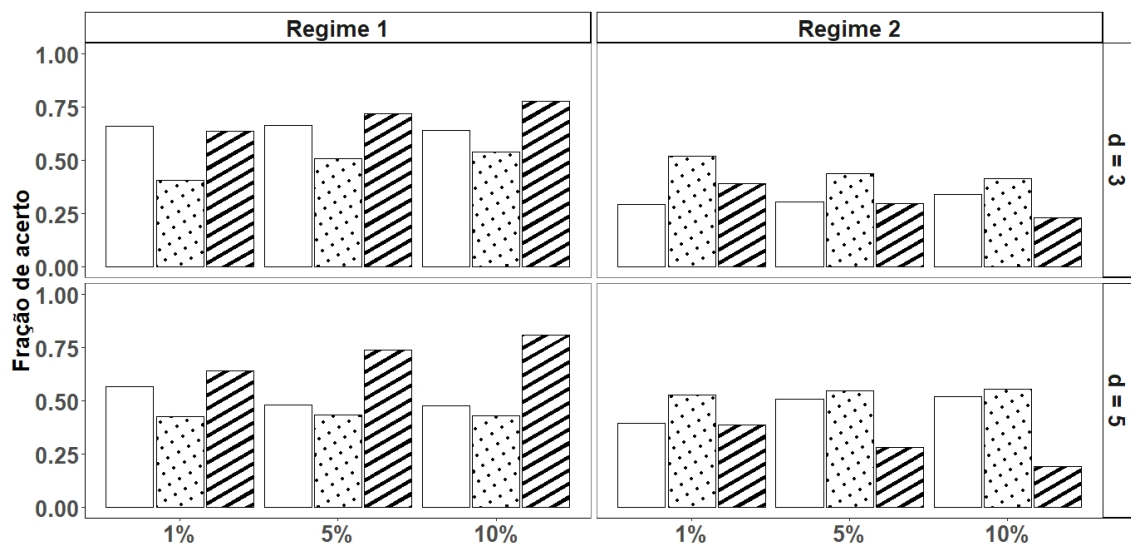


Figura 5 – Proporção de classificação correta dos diferentes estimadores para o modelo IS-DCC para todas as observações. No eixo horizontal estão as porcentagens de *outliers*. Preenchido em branco o estimador QML-n, com bolas o estimador QML-t e em linhas diagonais o estimador Robusto. Os quadros da esquerda são referentes à quando o regime correto é de maior persistência na correlação, ou seja, o regime (1), os da direita para o regime de menor persistência na correlação, regime (2). Nos quadros superiores são *outliers* de tamanho 3 vezes o desvio padrão condicional, nos inferiores 5 vezes o desvio padrão condicional.

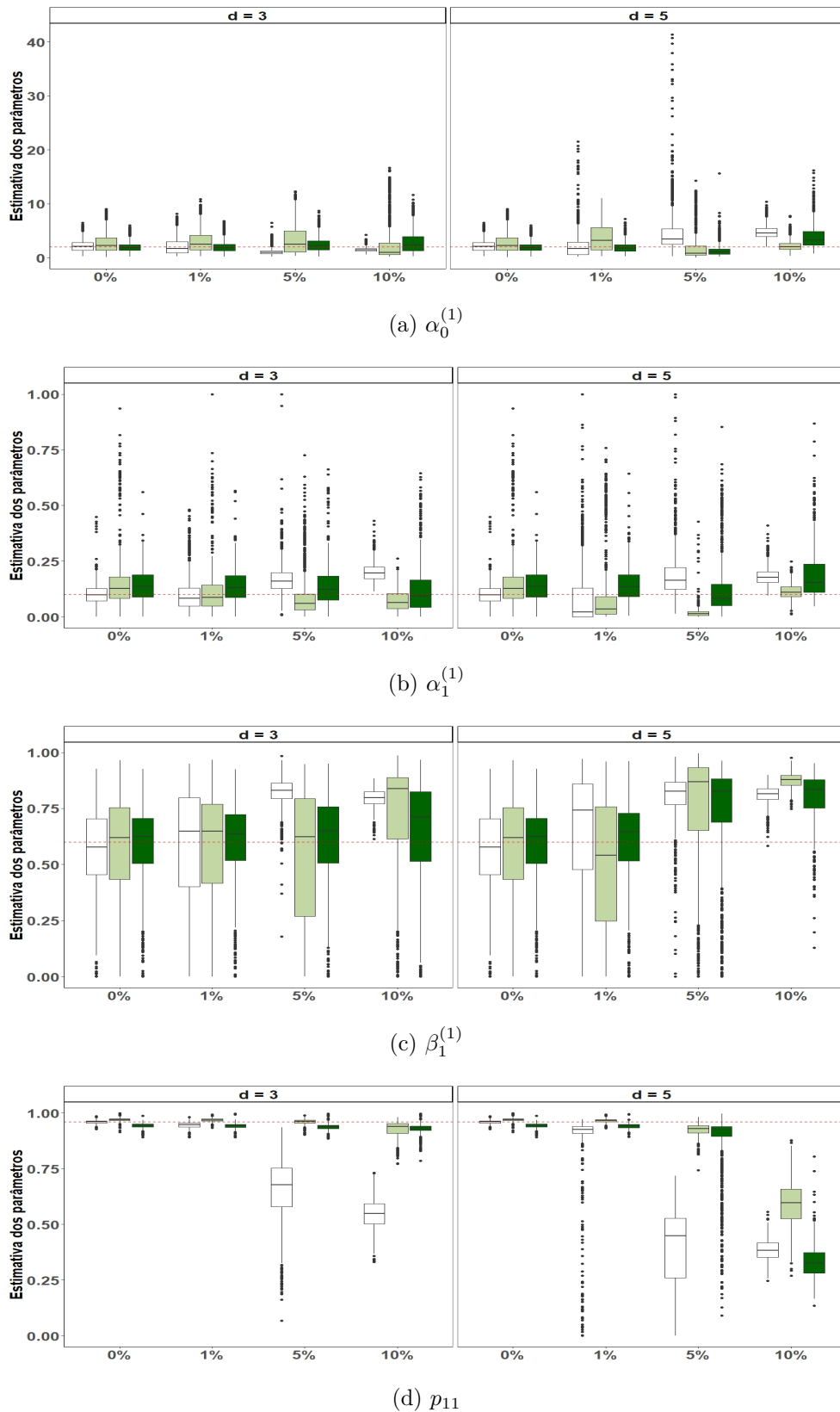


Figura 6 – Boxplot das estimativas dos parâmetros do regime de alta volatilidade das replicações Monte Carlo do modelo MS-GARCH(1,1). O eixo x representa a porcentagem de *outliers* e o eixo y representa os valores das estimativas dos parâmetros. No quadro esquerdo são os valores para os tamanhos de *outliers* com  $d = 3$  e no direito  $d = 5$ . Em branco o estimador QML-n, verde claro o estimador QML-t e verde escuro o estimador Robusto. A linha tracejada em vermelho representa o valor real do parâmetro. Os parâmetros em cada gráfico são: (a)  $\alpha_0^{(1)}$  (b)  $\alpha_1^{(1)}$  (c)  $\beta_1^{(1)}$  (d)  $p_{11}$ . <sup>(1)</sup> representa o regime de alta persistência e <sup>(2)</sup> representa o regime de baixa persistência.

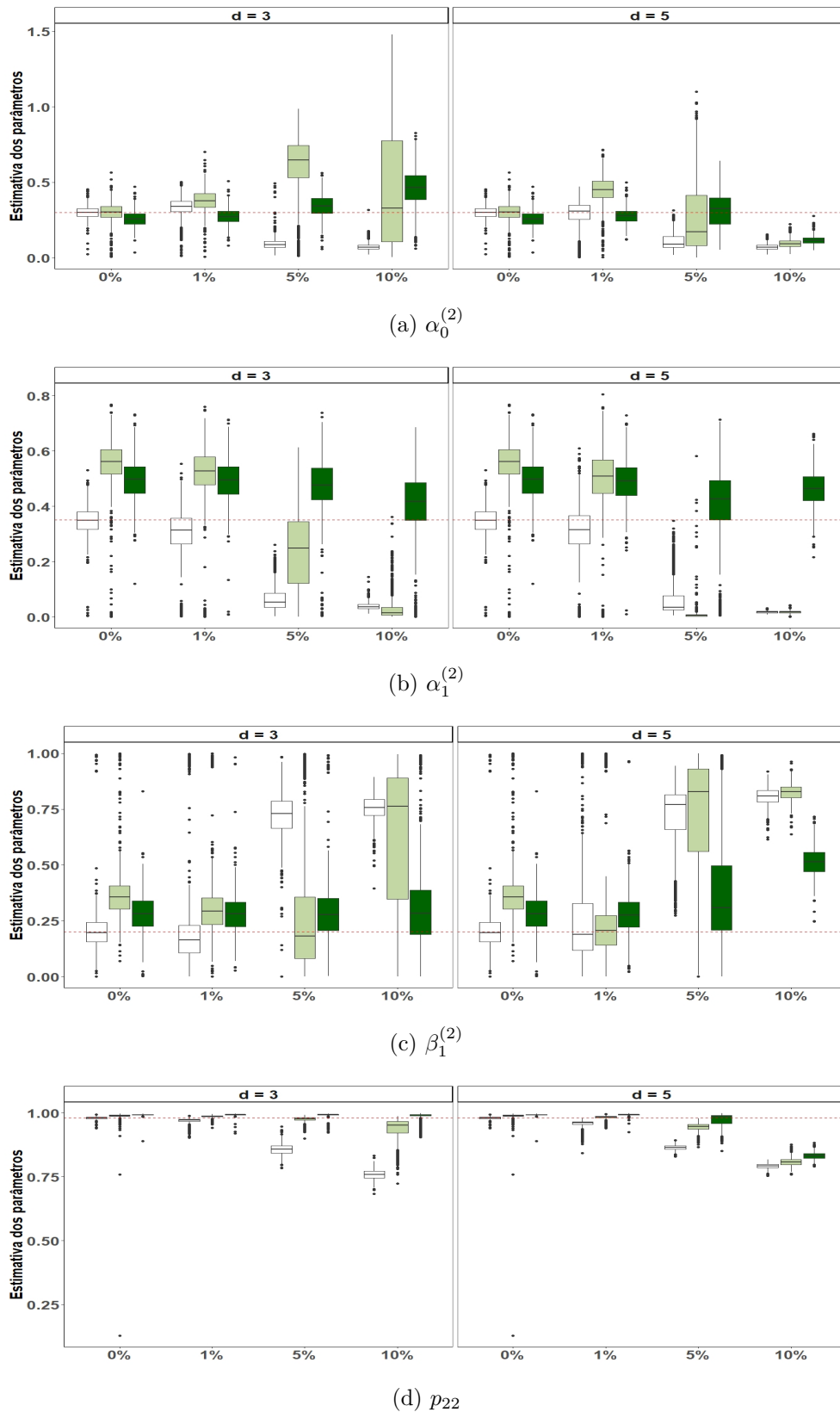
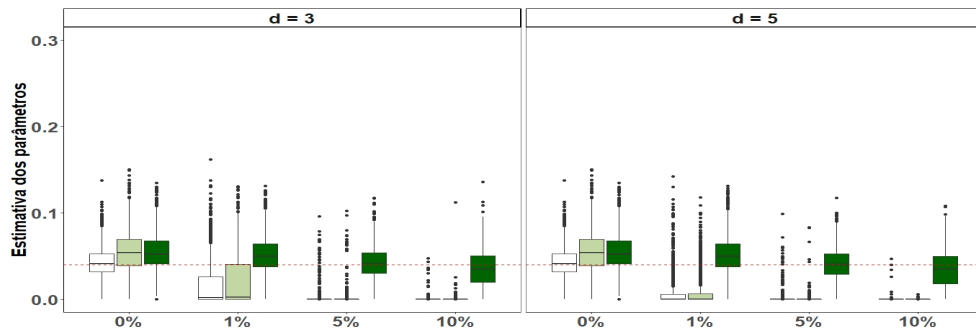


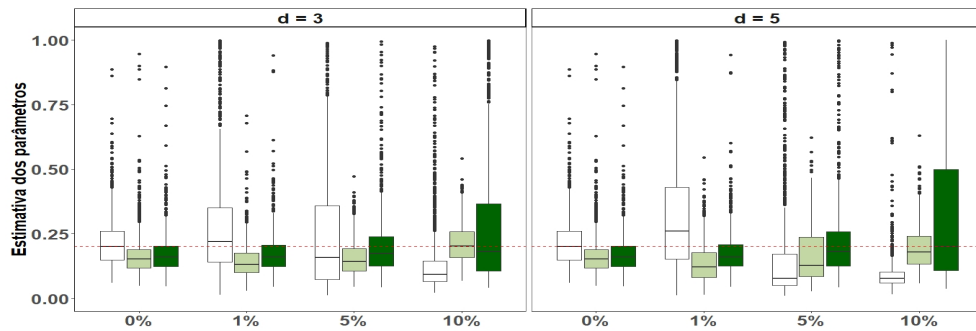
Figura 7 – Boxplot das estimativas dos parâmetros do regime de baixa volatilidade das replicações Monte Carlo do modelo MS-GARCH(1,1). O eixo x representa a porcentagem de *outliers* e o eixo y representa os valores das estimativas dos parâmetros. No quadro esquerdo são os valores para os tamanhos de *outliers* com  $d = 3$  e no direito  $d = 5$ . Em branco o estimador QML-n, verde claro o estimador QML-t e verde escuro o estimador Robusto. A linha tracejada em vermelho representa o valor real do parâmetro. Os parâmetros em cada gráfico são: (a)  $\alpha_0^{(2)}$  (b)  $\alpha_1^{(2)}$  (c)  $\beta_1^{(2)}$  (d)  $p_{22}$ .

Tabela 4 – Viés e RMSE, para os estimadores do modelo IS(2)-DCC(1,1) com parâmetros  $a_1^{(1)} = 0,04$ ,  $b_1^{(1)} = 0,93$ ,  $a_1^{(2)} = 0,2$ ,  $b_1^{(2)} = 0,6$ ,  $p_{11} = 0,96$ ,  $p_{22} = 0,98$ ,  $\text{per}^{(i)}$  é a persistência do regime  $i$ ,  $^{(1)}$  representa o regime de alta persistência e  $^{(2)}$  representa o regime de baixa persistência e  $T = 3000$ . Em negrito são os que apresentaram os melhores resultados.

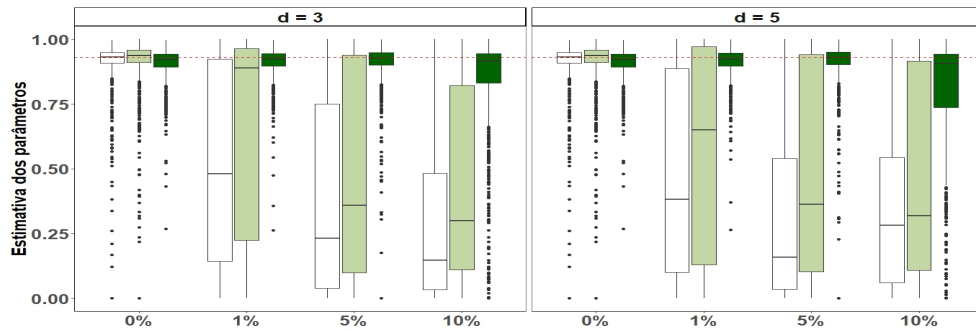
|                                 |      | Estim. | $a_1^{(1)}$   | $b_1^{(1)}$   | $a_1^{(2)}$   | $b_1^{(2)}$   | $p_{11}$      | $p_{22}$      | $\text{per}^{(1)}$ | $\text{per}^{(2)}$ |
|---------------------------------|------|--------|---------------|---------------|---------------|---------------|---------------|---------------|--------------------|--------------------|
| $\varepsilon = 0\%$             | Viés | QML-N  | <b>0,003</b>  | <b>-0,016</b> | <b>0,019</b>  | <b>-0,018</b> | -0,105        | -0,139        | -0,013             | <b>0,000</b>       |
|                                 |      | QML-t  | 0,016         | -0,021        | -0,033        | 0,052         | <b>-0,023</b> | <b>-0,007</b> | <b>-0,005</b>      | 0,020              |
|                                 |      | Rob    | 0,013         | -0,019        | -0,030        | 0,040         | -0,031        | -0,014        | -0,006             | 0,014              |
|                                 | RMSE | QML-N  | <b>0,018</b>  | <b>0,089</b>  | 0,103         | 0,171         | 0,206         | 0,254         | 0,082              | 0,122              |
|                                 |      | QML-t  | 0,029         | 0,114         | 0,092         | <b>0,150</b>  | <b>0,098</b>  | <b>0,098</b>  | 0,103              | 0,108              |
|                                 |      | Rob    | 0,025         | 0,064         | <b>0,087</b>  | 0,160         | 0,109         | 0,100         | <b>0,047</b>       | <b>0,105</b>       |
| $\varepsilon = 1\%$<br>$d = 3$  | Viés | QML-N  | -0,023        | -0,418        | 0,086         | 0,047         | -0,123        | -0,284        | -0,441             | 0,133              |
|                                 |      | QML-t  | -0,019        | -0,292        | -0,053        | 0,081         | <b>-0,005</b> | <b>0,019</b>  | -0,311             | 0,028              |
|                                 |      | Rob    | <b>0,009</b>  | <b>-0,017</b> | <b>-0,028</b> | <b>0,029</b>  | -0,031        | -0,021        | <b>-0,006</b>      | <b>0,010</b>       |
|                                 | RMSE | QML-N  | 0,035         | 0,561         | 0,229         | 0,260         | 0,154         | 0,366         | 0,580              | 0,198              |
|                                 |      | QML-t  | 0,034         | 0,483         | <b>0,089</b>  | 0,280         | <b>0,028</b>  | <b>0,049</b>  | 0,500              | 0,246              |
|                                 |      | Rob    | <b>0,022</b>  | <b>0,066</b>  | 0,093         | <b>0,170</b>  | 0,107         | 0,107         | <b>0,047</b>       | <b>0,117</b>       |
| $\varepsilon = 1\%$<br>$d = 5$  | Viés | QML-N  | 0,037         | 0,599         | 0,269         | 0,266         | 0,149         | 0,372         | -0,498             | 0,160              |
|                                 |      | QML-t  | -0,031        | -0,363        | -0,063        | 0,152         | <b>-0,009</b> | <b>0,018</b>  | -0,395             | 0,089              |
|                                 |      | Rob    | <b>0,008</b>  | <b>-0,015</b> | <b>-0,027</b> | <b>0,026</b>  | -0,033        | -0,023        | <b>-0,006</b>      | <b>0,008</b>       |
|                                 | RMSE | QML-N  | 0,327         | 0,327         | 0,211         | 0,389         | 0,364         | 0,185         | 0,624              | 0,192              |
|                                 |      | QML-t  | 0,036         | 0,537         | 0,098         | 0,290         | <b>0,024</b>  | <b>0,029</b>  | 0,561              | 0,239              |
|                                 |      | Rob    | <b>0,022</b>  | <b>0,061</b>  | <b>0,092</b>  | <b>0,169</b>  | 0,108         | 0,108         | <b>0,046</b>       | <b>0,119</b>       |
| $\varepsilon = 5\%$<br>$d = 3$  | Viés | QML-N  | -0,038        | -0,545        | 0,052         | 0,118         | -0,121        | -0,286        | -0,583             | 0,170              |
|                                 |      | QML-t  | -0,038        | -0,463        | -0,041        | 0,202         | <b>-0,019</b> | <b>0,003</b>  | -0,502             | 0,160              |
|                                 |      | Rob    | <b>-0,001</b> | <b>-0,027</b> | <b>0,002</b>  | <b>-0,017</b> | -0,043        | -0,083        | <b>-0,019</b>      | <b>-0,012</b>      |
|                                 | RMSE | QML-N  | 0,039         | 0,663         | 0,240         | 0,293         | 0,142         | 0,357         | 0,695              | 0,214              |
|                                 |      | QML-t  | 0,039         | 0,604         | <b>0,081</b>  | 0,233         | <b>0,031</b>  | <b>0,035</b>  | 0,634              | 0,181              |
|                                 |      | Rob    | <b>0,018</b>  | <b>0,125</b>  | 0,148         | <b>0,205</b>  | 0,126         | 0,203         | <b>0,093</b>       | <b>0,144</b>       |
| $\varepsilon = 5\%$<br>$d = 5$  | Viés | QML-N  | -0,039        | -0,608        | -0,034        | 0,217         | -0,118        | -0,219        | -0,647             | 0,183              |
|                                 |      | QML-t  | -0,039        | -0,464        | -0,036        | 0,210         | <b>-0,033</b> | <b>-0,013</b> | -0,503             | 0,174              |
|                                 |      | Rob    | <b>-0,003</b> | <b>-0,031</b> | <b>0,025</b>  | <b>-0,031</b> | -0,048        | -0,110        | <b>-0,021</b>      | <b>-0,006</b>      |
|                                 | RMSE | QML-N  | 0,039         | 0,700         | 0,211         | 0,320         | 0,143         | 0,293         | 0,735              | 0,209              |
|                                 |      | QML-t  | 0,040         | 0,601         | <b>0,108</b>  | 0,241         | <b>0,053</b>  | <b>0,062</b>  | 0,632              | 0,183              |
|                                 |      | Rob    | <b>0,018</b>  | <b>0,142</b>  | 0,183         | <b>0,221</b>  | 0,134         | 0,238         | <b>0,104</b>       | <b>0,147</b>       |
| $\varepsilon = 10\%$<br>$d = 3$ | Viés | QML-N  | -0,040        | -0,640        | -0,057        | 0,248         | -0,124        | -0,247        | -0,679             | 0,192              |
|                                 |      | QML-t  | -0,040        | -0,502        | <b>0,014</b>  | 0,152         | <b>-0,040</b> | <b>-0,030</b> | -0,542             | 0,165              |
|                                 |      | Rob    | <b>-0,005</b> | <b>-0,111</b> | 0,089         | <b>-0,041</b> | -0,076        | -0,217        | <b>-0,116</b>      | <b>0,048</b>       |
|                                 | RMSE | QML-N  | 0,040         | 0,715         | 0,161         | 0,301         | 0,136         | 0,287         | 0,751              | 0,203              |
|                                 |      | QML-t  | 0,040         | 0,619         | <b>0,074</b>  | <b>0,171</b>  | <b>0,053</b>  | <b>0,064</b>  | 0,651              | 0,166              |
|                                 |      | Rob    | <b>0,022</b>  | <b>0,263</b>  | 0,273         | 0,279         | 0,148         | 0,338         | <b>0,262</b>       | <b>0,171</b>       |
| $\varepsilon = 10\%$<br>$d = 5$ | Viés | QML-N  | -0,040        | -0,597        | -0,103        | 0,298         | -0,150        | -0,239        | -0,637             | 0,195              |
|                                 |      | QML-t  | -0,040        | -0,482        | <b>-0,005</b> | 0,180         | <b>-0,053</b> | <b>-0,045</b> | -0,522             | 0,175              |
|                                 |      | Rob    | <b>-0,005</b> | <b>-0,142</b> | 0,144         | <b>-0,060</b> | -0,095        | -0,308        | <b>-0,148</b>      | <b>0,084</b>       |
|                                 | RMSE | QML-N  | 0,040         | 0,663         | 0,138         | 0,316         | 0,162         | 0,266         | 0,699              | 0,201              |
|                                 |      | QML-t  | 0,040         | 0,611         | <b>0,080</b>  | <b>0,197</b>  | <b>0,073</b>  | <b>0,092</b>  | 0,643              | <b>0,175</b>       |
|                                 |      | Rob    | <b>0,023</b>  | <b>0,291</b>  | 0,338         | 0,310         | 0,163         | 0,416         | <b>0,292</b>       | 0,182              |



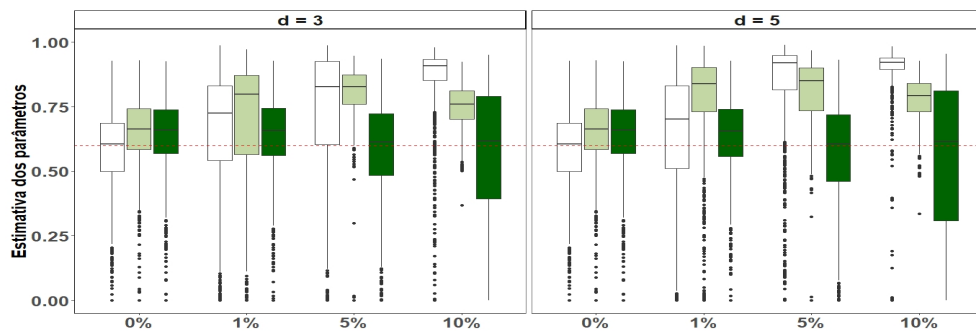
(a)  $a_1^{(1)}$



(b)  $a_1^{(2)}$

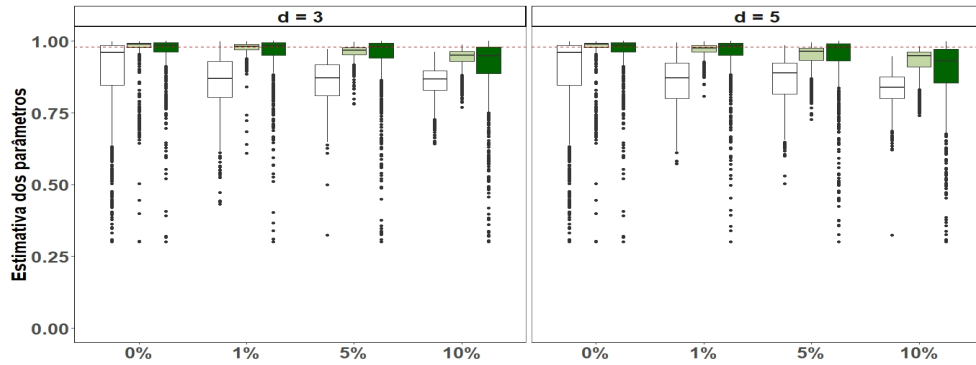


(c)  $b_1^{(1)}$

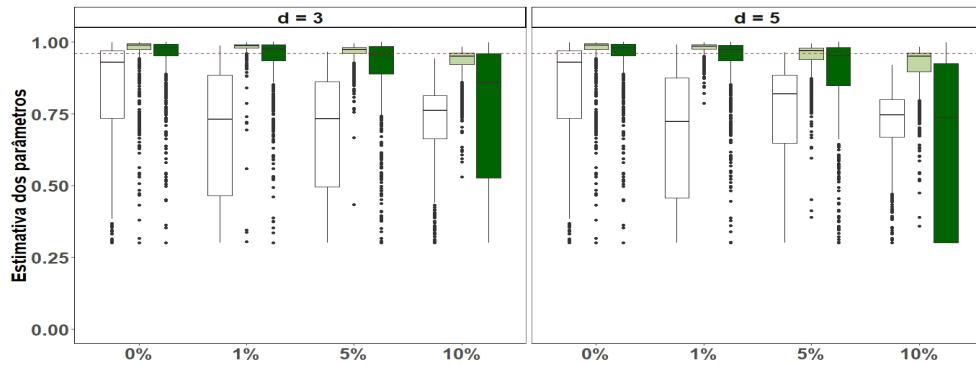


(d)  $b_1^{(2)}$

Figura 8 – Boxplot das estimativas dos parâmetros da correlação das replicações Monte Carlo do modelo IS-DCC(1,1). O eixo x representa a porcentagem de *outliers* e o eixo y representa os valores das estimativas dos parâmetros. No quadro esquerdo são os valores para os tamanhos de *outliers* com  $d = 3$  e no direito  $d = 5$ . Em branco o estimador QML-n, verde claro o estimador QML-t e verde escuro o estimador Robusto. A linha tracejada em vermelho representa o valor real do parâmetro. Os parâmetros em cada gráfico são: (a)  $a_1^{(1)}$  (b)  $a_1^{(2)}$  (c)  $b_1^{(1)}$  (d)  $b_1^{(2)}$ .



(a)  $p_{11}$



(b)  $p_{22}$

Figura 9 – Boxplot das estimativas dos parâmetros da matriz de transição das replicações Monte Carlo do modelo IS-DCC(1,1). O eixo x representa a porcentagem de *outliers* e o eixo y representa os valores das estimativas dos parâmetros. No quadro esquerdo são os valores para os tamanhos de *outliers* com  $d = 3$  e no direito  $d = 5$ . Em branco o estimador QML-n, verde claro o estimador QML-t e verde escuro o estimador Robusto. A linha tracejada em vermelho representa o valor real do parâmetro. Os parâmetros em cada gráfico são: (a)  $p_{11}$  (b)  $p_{22}$ .

## 6 Aplicações

Estudar a volatilidade de ativos financeiros tem sido alvo de diversos autores na literatura. Mais especificamente, modelando seus respectivos retornos, com ênfase na estimação da volatilidade. Para isso, os modelos GARCH se destacam devido a sua capacidade de capturar a maioria dos fatos estilizados das séries temporais financeiras e pela sua simplicidade de estimação como foi dito anteriormente no Capítulo 2. E uma das aplicações mais relevantes desses modelos é a previsão do VaR, bem como mostram Manganeli e Engle (2001).

No entanto, vimos que na presença de mudanças de regime, os modelos GARCH superestimam a persistência do modelo, assim gerando estimativas impróprias a longo prazo. Para esse cenário, vimos que os modelo GARCH com mudança de regime markoviano são mais adequados. Ardia et al. (2018) comparam o desempenho das previsões de modelos de regime único e MS-GARCH sob a perspectiva do VaR para ativos que são frequentemente negociados por gestores de fundos na bolsa americana e observam que os modelos MS-GARCH geram previsões mais precisas do VaR.

Diante disso, o foco aqui é comparar o desempenho das previsões do modelo MS-GARCH estimado por quase-verossimilhança via filtro de Hamilton, apresentada na Seções 3.2 e 4.2.2 com a das versão robusta apresentada nas Seções 3.3 e 4.2.3, na estimação do VaR. Para isso iremos comparar a performance de ambos os estimadores para 5 séries diferentes, sendo elas a série das moedas do Euro, Yen, Franco Suíço e da Libra - todas em relação ao dólar americano e negociadas ao meio-dia em Nova York. Para todas as séries, os dados são de 04/01/1999 a 28/05/2021. Apesar de em 1951, a Alemanha Ocidental, a Bélgica, a França, a Itália, o Luxemburgo e os Países Baixos decidirem criar um mercado comum do carvão e do aço, apenas em 1999 o euro é lançado como moeda escritural nos mercados financeiros e usado para fazer pagamentos eletrônicos. Em 1999, o euro tornou-se a moeda oficial de Portugal, da Alemanha, Áustria, Bélgica, Espanha, Finlândia, França, Irlanda, Itália, Luxemburgo e Países Baixos<sup>1</sup>. Assim, decidimos utilizar as séries a partir dessa data para todas as séries. Para as quatro taxas de câmbio em que iremos utilizar neste trabalho foram obtidas através do site <https://www.federalreserve.gov/releases/H10/hist/>. Lá, os dados da taxa de câmbio bilateral são atualizados todas as segundas-feiras. E as taxas de câmbio a seguir são certificadas pelo *Federal Reserve Bank* de Nova York para fins alfandegários.

A escolha da data final foi definida pela data mais recente quando foi iniciada a análise, com um total de 5845 observações. Esse trabalho tem como referência principal

<sup>1</sup> <https://europa.eu/euroat20/pt/a-historia-do-euro/>



os trabalhos apresentados por Haas, Mittnik e Paolella (2004) e Paolella et al. (2012). Na aplicação apresentada por Haas, Mittnik e Paolella (2004), eles estudam o desempenho do modelo MS-GARCH em termos de previsão para diferentes moedas. São utilizado um total de 6336 observações (Janeiro de 1978 a Junho de 2003), 6313 (Janeiro de 1978 a Junho de 2003), e 5313 (Janeiro de 1981 a Junho de 2003) observações para as taxas de câmbio do Yen, a Libra e o dólar de Singapura contra o dólar americano, respectivamente. Já, Paolella et al. (2012) avaliam o desempenho deste modelo utilizando os índices da França, Alemanha e Reino Unido em um total de 5203 observações (Janeiro de 1990 a Junho de 2010). E, quando avaliaram o desempenho em previsão, Paolella et al. (2012) utilizam janelas móveis de tamanho 2500 enquanto Haas, Mittnik e Paolella (2004) utilizam janelas em expansão iniciando com uma de tamanho 3000.

Outro ponto a se destacar é o cenário social e político econômico do período analisado. Na introdução do modelo MS-GARCH, foi destacada a sua capacidade capturar os períodos mais ou menos voláteis da economia. Diante disso, é importante ressaltar pontos-chaves do ponto de vista econômico e político nesse período. De fato, não faz parte do escopo desse trabalho elencar todos os acontecimentos, mas sim alguns que julgamos serem cruciais. Além de muitos países adotarem o euro, podemos citar que, em 2001, houve o atentado ao World Trade Center e Pentágono. Em complemento, destacamos que, na década de 90, o Japão estava enfrentando um período de incertezas econômicas, em que havia um aumento das taxas de desemprego, falências bancárias etc. Temos também a crise do subprime que ocorreu entre 2008 e 2009 e acarretou um abalo em toda a economia mundial. E a partir de 2020, o mundo enfrenta a pandemia do vírus COVID-19, no qual a rotina e atividades de todo o mundo tiveram de ser adaptados.

Posto isso, neste capítulo será apresentado inicialmente na Seção 6.1 uma análise descritiva das séries e dos retornos que serão utilizadas na modelagem. Posteriormente, na Seção 6.2 será discutido a forma de se obter o VaR tanto para o caso univariado quanto bivariado bem como as métricas para avaliar a qualidade da previsão. Os resultados para os casos univariados e bivariados são apresentados nas Seções 6.3 e 6.4, respectivamente. Finalmente, a Seção 6.5 apresenta as principais conclusões das aplicações.

## 6.1 Análise Descritiva

Para análise que iremos fazer utilizaremos os log-retornos das taxas de câmbio percentual, isto é, seja  $P_t$  a taxa no tempo  $t$ , o log-retorno é definido como sendo  $100 \times \log(P_t/P_{t-1})$ . Nos dias em que o mercado não operou, excluindo os finais de semana, imputamos essas observações utilizando o modelo de nível e tendência utilizando a estimativa suavizada dado pelo filtro de Kalman. A Tabela 5 apresenta algumas medidas descritivas desses retornos. Podemos ver que, exceto o euro, todos apresentam curtose

maior do que a da distribuição normal ( $> 3$ ), destacando o franco suíço e a libra. Também podemos observar que todas elas apresentam baixa assimetria e que, embora as variâncias amostrais sejam aproximadamente a mesma para as quatro séries, a curtose da série do franco suíço, 33,01, é muito maior do que as das outras séries, seguidas da série da libra, 10,79.

Tabela 5 – Estatísticas descritivas das taxas de câmbio.

| Moeda        | Min    | $q_{25}$ | $q_{50}$ | Média | $q_{75}$ | Max  | Variância | Assimetria | Curtose |
|--------------|--------|----------|----------|-------|----------|------|-----------|------------|---------|
| Euro         | -3,00  | -0,31    | 0,00     | 0,00  | 0,31     | 4,62 | 0,33      | 0,15       | 2,69    |
| Libra        | -8,17  | -0,31    | 0,00     | -0,00 | 0,32     | 4,43 | 0,32      | -0,56      | 10,79   |
| Yen          | -5,22  | -0,31    | 0,00     | -0,00 | 0,33     | 3,34 | 0,37      | -0,31      | 4,50    |
| Franco suíço | -13,02 | -0,33    | 0,00     | -0,01 | 0,32     | 4,54 | 0,41      | -1,57      | 33,01   |

As Figuras 10 e 11 apresentam a evolução diária das moedas, os log-retorno das mesmas, a função de auto-correlação amostral dos log-retornos e a função de auto-correlação amostral do quadrado dos log-retornos do Euro, da Libra, do Yen e do Franco suíço, respectivamente. Nestas Figuras podemos observar o comportamento dessas séries no período estudado. Além disso, na série dos retornos podemos ver os conglomerados e de volatilidade. Também conseguimos observar o nível de cada uma. Podemos destacar o nível da série Yen, que já vem caindo sob efeito da crise do subprime e além disto em março 2011 o Japão foi afetado por tsunami e terremoto <sup>2</sup>. Porém, em 2013 o país anunciou um megaplano econômico para tirar país da crise fazendo a moeda voltar ao nível pré-crise. <sup>3</sup>. Podem observar também como a crise do subprime de 2008-2009 afetou essas séries e como esse período é mais volátil para todos as quatro séries. Para série dos log-retornos do franco suíço podemos ver um retorno negativo que se destoa dos demais. Esse retorno é referente ao dia 15 de janeiro de 2015 que teve como principal destaque econômico o fato do Banco Nacional Suíço (SNB) anunciar o abandono de sua política de estabelecer um teto para a moeda do país diante do euro <sup>4</sup>.

Ao olhar as quatro séries em conjunto podemos observar alguns comportamentos similares e distintos entre elas. Podemos observar por exemplo que a série do franco suíço e do Yen vinham caindo e no período da crise do subprime ambas caíram mais ainda, assim como a série da libra. No entanto, a partir de 2013 a série do Yen voltou a crescer, diferente do franco suíço e da libra. A série do euro, de maneira geral, se comporta de maneira semelhante a libra até o período da crise do subprime, ambas estão crescendo, no entanto, a série do euro não teve uma queda tão forte quanto a da libra.

<sup>2</sup> <https://www.bbc.com/portuguese/internacional-55943220>

<sup>3</sup> <http://g1.globo.com/economia/noticia/2013/01/japao-anuncia-megaplano-economico-para-tirar-pais-da-crise.html>

<sup>4</sup> <https://noticias.uol.com.br/ultimas-noticias/agencia-estado/2015/01/15/franco-suico-e-dolar-sobem-ante-euro-apos-decisao-do-banco-central-da-suica.htm>

Em relação aos gráficos das funções de autocorrelação dos retornos das quatro moedas, podemos observar que há correlação significativa na primeira defasagem. E em relação às funções de autocorrelação dos quadrados dos retornos podemos ver que elas são bastante significativas e com alta persistência, principalmente para a série do Euro.

Ademais, como estamos trabalhando com modelos MS-GARCH sem média, é necessário filtrar essas séries para remover a média condicional. O filtro utilizado para todas as séries foi um AR(1), ou seja, seja  $r_t$  um retorno. Assim, temos que o retorno filtrado  $r_t^*$  será dado por:

$$r_t^* = r_t - (\phi r_{t-1}),$$

em que  $\phi$  é o parâmetro associado ao termo auto-regressivo estimado aqui por máxima verossimilhança. Observe que aqui utilizamos um AR(1) sem constante. Uma análise, não mostrada no texto, das funções autocorrelações de  $r_t^*$  mostra que este filtro foi suficiente para explicar a correlação serial.

## 6.2 Medidas de Desempenho das Previsões

O desempenho desses modelos em termos de previsão será avaliado através da medida Valor em Risco (VaR - *Value at Risk*). Aqui para calcular o VaR, utilizaremos de maneira semelhante a ideia apresentada por Ardia et al. (2018) para os modelos MS-GARCH. Para esses modelos temos que a função de densidade de probabilidade condicional um passo à frente de  $y_{T+1}$  é uma mistura de  $K$  distribuições dependentes de regime, isto é,

$$f(y_{T+1}|\mathcal{I}_T) = \sum_{k=1}^K \pi_{k,T+1} f(y_{T+1}|\Delta = k, \mathcal{I}_T),$$

em que  $\pi_{k,T+1}$  atua como peso de mistura é definida como sendo  $\pi_{k,T+1} \equiv \sum_{i=1}^K p_{i,k} \eta_{i,T}$ , tal que  $p_{i,j}$  é a probabilidade da matriz de transição e  $\eta_{i,T}$  é a probabilidade filtrada  $P[\Delta_T = i|\mathcal{I}_{T-1}]$  do tempo  $t$ . Assim, temos que a função de distribuição acumulada de  $y_{T+1}$  é dada por

$$F(y_{T+1}|\mathcal{I}_T) \equiv \int_{-\infty}^{y_{T+1}} f(x|\mathcal{I}_T) dx.$$

Então temos que o VaR é obtido como sendo o quantil ( $\alpha$ ) da densidade preditiva, invertendo numericamente a função de distribuição acumulada preditiva, ou seja,

$$P_T[y_t < -\text{VaR}_{T+1}(\alpha)] = \alpha. \quad (6.1)$$

De maneira mais específica, nesse trabalho estamos interessado em avaliar a performance do modelo em previsões um passo a frente para uma das séries utilizando no

VaR  $\alpha = 0,01, 0,05$ . Para isso, cada modelo é estimado em uma janela móvel, onde os VaR um passo a frente é obtido. Utilizando janelas com 3000 observações na estimação, teremos um total de 2884 valores de VaR no backtesting. Inicialmente, para cada janela, são imputados os valores faltantes utilizando o modelo de nível e tendência utilizando a estimativa suavizada dado pelo filtro de Kalman e depois é aplicado um filtro AR(1) para tornar a média condicional igual a zero (*demeaned series*). Para obter as estimativas do VaR um passo a frente seguimos os seguintes passos, denotando por  $T$  a última :

1. Caso a observação no tempo  $t$  ou  $t + 1$  seja um dado faltante, não faremos previsão do VaR um passo a frente.
2. Caso não haja dados faltante no tempo  $t$  ou  $t + 1$ , faremos todo passo de estimação até o tempo  $t$ . E para prever o VaR no tempo  $t + 1$  utilizamos a previsão 1 passos a frente.

Para o caso multivariado, estamos interessado em avaliar a possível perda de uma carteira em um determinado horizonte de tempo (FRANCO; ZAKOIAN, 2019, Seção 11.3 pág. 332). Para avaliar isso, considere uma carteira com  $N$  ativos e cada log retorno individual  $i$  da carteira dado por  $r_{i,t}$ . Então, de acordo com Panna (2017), temos que o log retorno da carteira no tempo  $t$ , dado a informação disponível no tempo  $t - 1$ , é dado por:

$$S_t = \ln \left( \sum_{i=1}^N w_i e^{r_{i,t}} \right), \quad (6.2)$$

em que  $w_i$  são os pesos de cada ativo na carteira, de tal forma que  $\sum_{i=1}^N w_i = 1$ .

A previsão 1 passo à frente do log retorno de uma carteira, ou seja,  $S_{t+1}$  pode ser obtida através da previsão 1 passo a frente de  $\mathbf{y}_{t+1}$  definido em 4.12, que pode ser obtida de maneira sequencial. Sendo assim, temos que, ponto de vista probabilístico, o VaR, sob um quantil ( $\alpha$ ), é dado por:

$$\text{VaR}_{t+1}(\alpha) = \inf\{x \in \mathbb{R} \mid P[S_{t+1} < x] > 1 - \alpha\}. \quad (6.3)$$

No caso multivariado, utilizaremos os mesmos valores de  $\alpha$ , a quantidade de observações dentro da amostra, os filtros AR(1) e a imputação feita no caso univariado.

### 6.2.1 Testes de desempenho de previsão

O VaR pode ser definido como uma perda que não excederá um nível de confiança especificado assim como está descrito em (6.1). Para verificar se o VaR estimado está em concordância com as perdas da série dos retornos, utilizaremos os métodos

conhecidos como *backtesting*. De maneira geral, se estabelecemos  $\alpha = 5\%$  esperamos que 5% dos retornos excedam o VaR estimado e que a sequência temporal em que as perdas sejam maiores do que o VaR seja independente. E assim como cita Ardia et al. (2018), testamos a precisão das previsões VaR usando a chamada variável de *hit*, denotada por  $H$ . Para definir esta variável, defina inicialmente a variável indicadora  $I_t(\alpha)$ , dos tempos em que a perda é maior do que o VaR, que é dada por:

$$I_t(\alpha) = \begin{cases} 0, & \text{se } y_t < -\text{VaR}_t(\alpha) \\ 1, & \text{caso contrário,} \end{cases}$$

em que  $y_t$  é o retorno do portfólio no caso multivariado, ou da série no caso univariado.

Para testar se a cobertura do VaR está correta, Kupiec (1995) propõe um teste baseado na razão de verossimilhança nomeado de teste de cobertura incondicional. Para esse teste, caso o VaR estiver estimado corretamente,  $I_t(\alpha)$  tem distribuição Bernoulli( $\alpha$ ) independentes; assim,  $H = \sum_{t=1}^T I_t(\alpha) \sim \text{Binomial}(T, \alpha)$ . Daí, seja a hipótese nula  $H_0 : E[H] = n\alpha$ , A estatística do teste  $UC(\alpha)$  é dada por:

$$LR_{UC} = -2 \ln \left[ (1 - \alpha)^{T-H} \alpha^H \right] + 2 \ln \left[ \left( 1 - \frac{H}{T} \right)^{T-H} \left( \frac{H}{T} \right)^H \right].$$

Sob a hipótese nula,  $UC(\alpha) \sim \chi^2(1)$ .

No entanto, assim como cita Gaglianone et al. (2011), a propriedade de cobertura incondicional não fornece informações sobre a dependência temporal das violações do VaR, e o teste de Kupiec (1995) ignora a cobertura condicionante, uma vez que as violações podem se agrupar ao longo do tempo, o que também deve invalidar um modelo VaR. Assim, Christoffersen (1998) estende a estatística de razão de verossimilhança  $UC(\alpha)$  de tal forma que a sequência de acertos também deva ser independente ao longo do tempo. Para isso, Christoffersen (1998) assume que o processo de violação  $I_t(\alpha)$  pode ser representado como uma cadeia de Markov com dois estados:

$$\Pi = \begin{pmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{pmatrix}$$

de tal forma que  $\pi_{ij} = P[I_t(\alpha) = j | I_t(\alpha) = i]$ . Logo, sob a hipótese nula,

$$H_0 : \Pi = \begin{pmatrix} 1 - \alpha & \alpha \\ 1 - \alpha & \alpha \end{pmatrix},$$

as estatísticas de teste para cobertura condicional e para independência são dadas respectivamente por:

$$LR_{CC} = -2 \ln \left[ (1 - \alpha)^{T-H} \alpha^H \right] + 2 \ln \left[ (1 - \hat{\pi}_{01})^{n_{00}} \hat{\pi}_{01}^{n_{01}} (1 - \hat{\pi}_{11})^{n_{10}} \hat{\pi}_{11}^{n_{11}} \right] \xrightarrow{d} \chi^2(2)$$

$$LR_{Ind} = -2 \ln \left[ (1 - H/T)^{T-H} (H/T)^H \right] + 2 \ln \left[ (1 - \hat{\pi}_{01})^{n_{00}} \hat{\pi}_{01}^{n_{01}} (1 - \hat{\pi}_{11})^{n_{10}} \hat{\pi}_{11}^{n_{11}} \right] \xrightarrow{d} \chi^2(1),$$

em que  $n_{ij}$  é o total de observações que vão de  $i$  para  $j$ . E como cita Gaglianone et al. (2011), sob a hipótese nula de independência, temos que  $\pi = \pi_{11} = \pi_{01} = (n_{01} + n_{11}) = T$ . Além disso, temos que  $LR_{CC} = LR_{UC} + LR_{Ind}$ .

Também, Engle e Manganelli (2004) propõem um teste baseado em uma regressão linear. Seja  $Hit_t^\alpha \equiv I_t(\alpha) - \alpha$ . considere a seguinte regressão:

$$Hit_t^\alpha = \delta_0 + \sum_{l=1}^L \delta_l Hit_{t-l}^\alpha + \delta_{L+1} VaR_{t-1}(\alpha) + \epsilon, \quad (6.4)$$

em que  $\epsilon \sim N(0, \sigma^2)$ . Agora, denote  $\boldsymbol{\delta} = (\delta_0, \dots, \delta_{L+1})'$  o vetor de todos os parâmetros e  $\hat{\boldsymbol{\delta}}$  sua estimativa por mínimos quadrados ordinários. Então, sob a hipótese nula  $H_0 : \boldsymbol{\delta} = \mathbf{0}$ , a estatística do teste, denotada por  $DQ_\alpha$  é dada por:

$$DQ_\alpha = \frac{\hat{\boldsymbol{\delta}}' \mathbf{Z}' \mathbf{Z} \hat{\boldsymbol{\delta}}}{\alpha(1 - \alpha)},$$

em que  $\mathbf{Z}$  é a matriz das variáveis explicativas. Também, sob a hipótese nula,  $DQ_\alpha$  converge em distribuição para uma  $\chi^2(L+2)$ . Note que esse teste representa o teste de Wald no qual testamos os parâmetros conjuntamente. Geralmente, toma-se  $L = 4$ . Não necessariamente a única variável explicativa em (6.4) precisa ser  $VaR_{t-1}(\alpha)$ . Outras variáveis como os retornos, por exemplo, podem entrar no modelo. No entanto, por praticidade, e seguindo a notação apresentada por Ardia et al. (2018), escolhemos essa definição.

De fato, existe diversos métodos para testar a validade do VaR e não faz parte do escopo deste trabalho listar todos eles. Entre elas, podemos citar a função de perda apresentada por González-Rivera, Lee e Mishra (2004), mas não serão utilizadas neste trabalho.

### 6.3 Análise Univariada

Nesta seção serão apresentados os resultados para o modelo MS(2)-GARCH(1,1) para a estimação robusta e para a estimação por quase-verossimilhança normal e t-Student.

Para melhor entender o cenário em que as previsões estão realizadas, inicialmente apresentamos os resultados das estimações dos parâmetros do modelo para cada uma das quatro moedas na primeira e última janela. A ideia é exemplificar como os parâmetros variam ao longo do tempo, o que de certa forma é algo natural. Essa variação reforça a nossa metodologia de utilizar janela móveis.

A Tabela 6 apresenta os parâmetros estimados para a série do Euro para os três estimadores. Nela podemos observar a existência de grandes mudanças entre os parâmetros da primeira para última janela para todos os estimadores. Além disso, podemos observar que os estimadores QML-n e robusto tendem a estimar a probabilidade estacionária de um regime bem maior que a do outro regime. E para esse regime com maior probabilidade, as estimativas são semelhantes.

Tabela 6 – Estimativas dos parâmetros do modelo MS(2)-GARCH(1,1) para a série do log-retorno da taxa de câmbio do Euro em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto.

|                  | Primeira Janela |        |        | Última Janela |        |        |
|------------------|-----------------|--------|--------|---------------|--------|--------|
|                  | Rob             | QML-n  | QML-t  | Rob           | QML-n  | QML-t  |
| $\alpha_0^{(1)}$ | 0,0014          | 0,0001 | 0,0047 | 0,0133        | 0,0278 | 0,0096 |
| $\alpha_1^{(1)}$ | 0,0459          | 0,0018 | 0,0344 | 0,0774        | 0,0432 | 0,0503 |
| $\beta_1^{(1)}$  | 0,9511          | 0,9911 | 0,9608 | 0,8939        | 0,9485 | 0,9377 |
| $p_{11}$         | 0,8687          | 0,0792 | 0,6738 | 0,9567        | 0,0059 | 0,9926 |
| $\pi_{1\infty}$  | 0,3490          | 0,1977 | 0,4118 | 0,0546        | 0,0854 | 0,3729 |
| $\alpha_0^{(2)}$ | 0,0033          | 0,0032 | 0,0047 | 0,0001        | 0,0006 | 0,0001 |
| $\alpha_1^{(2)}$ | 0,0513          | 0,0328 | 0,0344 | 0,0308        | 0,0323 | 0,0503 |
| $\beta_1^{(2)}$  | 0,9416          | 0,9636 | 0,9608 | 0,9689        | 0,9575 | 0,9377 |
| $p_{22}$         | 0,9296          | 0,7731 | 0,7716 | 0,9975        | 0,9072 | 0,9956 |
| $\pi_{2\infty}$  | 0,6510          | 0,8023 | 0,5882 | 0,9454        | 0,9146 | 0,6271 |

De maneira similar, podemos ampliar a interpretação para as demais moedas. Quando analisamos os parâmetros estimados da Libra, apresentados na Tabela 7, podemos notar que existem diferenças nas estimativas dos parâmetros entre as janelas. Observa-se que os parâmetros estimados por QML-n da matriz de transição são pequenos. Nota-se também que o estimador robusto tende a estimar um regime com probabilidade muito alta na última janela. Além disso, o estimador QML-n tem uma mudança significativa em  $\alpha_1^{(1)}$  que vai de 0,1095 para 0,2379, respectivamente. E aqui também podemos observar que, por exemplo, na última janela, as estimativas dos parâmetros no regime com maior probabilidade na distribuição estacionária são bem parecidas.

A Tabela 8 apresenta os resultados para o Yen. Podemos observar há mudança na estimativa dos parâmetros entre as janelas em alguns parâmetros, principalmente nos parâmetros da matriz de transição. O estimador QML-n não apresenta muita diferença nas estimativas dos parâmetros da equação da variância, porém existe uma diferença em  $p_{11}$  que varia de 0,1061 para 0,2778. No entanto, podemos ver que a probabilidade de se manter no estado 1 ainda é muito baixa. Já em relação ao estimador QML-t, podemos notar que há uma diferença muito grande nos parâmetros do regime 2. Além disso, podemos observar que, na última janela, a estimativa de  $\alpha_1^{(1)}$  é muita alta. Novamente, podemos observar

Tabela 7 – Estimativas dos parâmetros do modelo MS(2)-GARCH(1,1) para a série do log-retorno da taxa de câmbio da Libra em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto.

|                  | Primeira Janela |        |        | Última Janela |        |        |
|------------------|-----------------|--------|--------|---------------|--------|--------|
|                  | Rob             | QML-n  | QML-t  | Rob           | QML-n  | QML-t  |
| $\alpha_0^{(1)}$ | 0,0173          | 0,0269 | 0,0058 | 0,0167        | 0,0650 | 0,0634 |
| $\alpha_1^{(1)}$ | 0,0564          | 0,1095 | 0,0484 | 0,3000        | 0,2379 | 0,0825 |
| $\beta_1^{(1)}$  | 0,8794          | 0,8772 | 0,9458 | 0,6948        | 0,7524 | 0,8098 |
| $p_{11}$         | 0,9625          | 0,2603 | 0,6340 | 0,6165        | 0,1544 | 0,9939 |
| $\pi_{1\infty}$  | 0,4993          | 0,2677 | 0,0193 | 0,0420        | 0,2612 | 0,2824 |
| $\alpha_0^{(2)}$ | 0,0014          | 0,0004 | 0,0036 | 0,0013        | 0,0010 | 0,0019 |
| $\alpha_1^{(2)}$ | 0,0438          | 0,0232 | 0,0359 | 0,0396        | 0,0219 | 0,0290 |
| $\beta_1^{(2)}$  | 0,9507          | 0,9680 | 0,9593 | 0,9550        | 0,9614 | 0,9687 |
| $p_{22}$         | 0,9626          | 0,7296 | 0,9928 | 0,9953        | 0,7010 | 0,9976 |
| $\pi_{2\infty}$  | 0,5007          | 0,7323 | 0,9807 | 0,9580        | 0,7388 | 0,7176 |

que, na última janela, as estimativas dos parâmetros no regime com maior probabilidade na distribuição estacionária são bem parecidas.

Tabela 8 – Estimativas dos parâmetros do modelo MS(2)-GARCH(1,1) para a série do log-retorno da taxa de câmbio do Yen em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto.

|                  | Primeira Janela |        |        | Última Janela |        |        |
|------------------|-----------------|--------|--------|---------------|--------|--------|
|                  | Rob             | QML-n  | QML-t  | Rob           | QML-n  | QML-t  |
| $\alpha_0^{(1)}$ | 0,1199          | 0,0482 | 0,1029 | 0,0767        | 0,0533 | 0,0008 |
| $\alpha_1^{(1)}$ | 0,0014          | 0,0658 | 0,0980 | 0,1096        | 0,0747 | 0,9990 |
| $\beta_1^{(1)}$  | 0,6241          | 0,9284 | 0,6875 | 0,7811        | 0,9149 | 0,0086 |
| $p_{11}$         | 0,9783          | 0,1061 | 0,3105 | 0,9132        | 0,2778 | 0,3290 |
| $\pi_{1\infty}$  | 0,2386          | 0,1743 | 0,1249 | 0,0345        | 0,1531 | 0,0722 |
| $\alpha_0^{(2)}$ | 0,0012          | 0,0030 | 0,0037 | 0,0004        | 0,0027 | 0,0031 |
| $\alpha_1^{(2)}$ | 0,0576          | 0,0222 | 0,0297 | 0,0431        | 0,0327 | 0,0498 |
| $\beta_1^{(2)}$  | 0,9415          | 0,9555 | 0,9675 | 0,9561        | 0,9305 | 0,9478 |
| $p_{22}$         | 0,9932          | 0,8257 | 0,9016 | 0,9969        | 0,8694 | 0,9478 |
| $\pi_{2\infty}$  | 0,7614          | 0,9190 | 0,8751 | 0,9655        | 0,8469 | 0,9278 |

Por fim, a Tabela 9 apresenta os parâmetros referentes à moeda franco suíço. Para esta moeda, observamos mudanças mais extremas nas estimativas dos parâmetros, principalmente em relação à  $\alpha_0^{(1)}$ . Da primeira para a última janela as estimativas vão de 0,0614, 0,0000 e 0,0057 para 0,0603, 0,5053 e 0,1845, para os estimadores robusto, QML-n e QML-t, respectivamente. Além disso, vimos que, novamente, o estimador robusto tende a estimar um regime com maior predominância, com probabilidades estacionárias iguais a 0,9633 e 0,9905 para a primeira e última janela, respectivamente. Para o estimador



QML-t estas probabilidades caem para 0,7019 e 0,7297, valores bem próximos. No caso do estimador QML-n, existe grande variação, com valor igual a 0,7631, próximo à estimativa QML-t, na primeira janela e igual a 0,9883 na última janela, próximo à estimativa robusta. Outra vez vemos que as estimativas dos parâmetros no regime com maior probabilidade na distribuição estacionária são bem parecidas.

Tabela 9 – Estimativas dos parâmetros do modelo MS(2)-GARCH(1,1) para a série do log-retorno da taxa de câmbio do Franco Suíço em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto.

|                  | Primeira Janela |        |        | Última Janela |        |        |
|------------------|-----------------|--------|--------|---------------|--------|--------|
|                  | Rob             | QML-n  | QML-t  | Rob           | QML-n  | QML-t  |
| $\alpha_0^{(1)}$ | 0,0614          | 0,0000 | 0,0057 | 0,0603        | 0,5053 | 0,1845 |
| $\alpha_1^{(1)}$ | 0,0679          | 0,0025 | 0,0304 | 0,0440        | 0,0000 | 0,0977 |
| $\beta_1^{(1)}$  | 0,7529          | 0,9888 | 0,9639 | 0,9494        | 0,9389 | 0,6412 |
| $p_{11}$         | 0,6276          | 0,1539 | 0,5201 | 0,7920        | 0,0675 | 0,9946 |
| $\pi_{1\infty}$  | 0,0367          | 0,2369 | 0,2981 | 0,0095        | 0,0117 | 0,2703 |
| $\alpha_0^{(2)}$ | 0,0019          | 0,0050 | 0,0057 | 0,0009        | 0,0015 | 0,0014 |
| $\alpha_1^{(2)}$ | 0,0330          | 0,0310 | 0,0304 | 0,0491        | 0,0291 | 0,0293 |
| $\beta_1^{(2)}$  | 0,9628          | 0,9647 | 0,9640 | 0,9486        | 0,9602 | 0,0086 |
| $p_{22}$         | 0,9858          | 0,7374 | 0,7962 | 0,9980        | 0,9890 | 0,9980 |
| $\pi_{2\infty}$  | 0,9633          | 0,7631 | 0,7019 | 0,9905        | 0,9883 | 0,7297 |

### 6.3.1 Estimação das probabilidades dos regimes

Para ilustrar as estimativas das probabilidades de estar em cada regime, para a primeira janela da série do Yen, a Figura 12 mostra as estimativas das probabilidades de estar no regime de maior volatilidade. O estimador QML-n estima para quase todas as observações uma baixa probabilidade de estar no regime de mais alta volatilidade. Entretanto, em um número bastante grande de observações isoladas, as estimativas são próximas a um. Como não se espera que ocorra este tipo de comportamento no mercado, isto é uma indicação de que o estimador QML-n é muito instável. As estimativas dos outros dois estimadores são mais estáveis, indicando durante todo o período uma alta probabilidade de estar no regime de baixa volatilidade. Entretanto as estimativas também têm comportamentos diferentes, enquanto as estimativas do estimador QML-t permanecem em um intervalo menor, existe uma estabilidade local maior nas estimativas do estimador robusto.

### 6.3.2 Análise das estimativas do VaR e das previsões

Para a análise do VaR iremos utilizar a série do log retorno filtrada por um AR(1). Ou seja, para o preço  $P_t$  e o seu log retorno  $r_t = \log(P_t/P_{t-1}) * 100$ , iremos fazer o

VaR para  $r_t^* = r_t - (\phi r_{t-1})$  em que  $\phi$  é o parâmetro associado ao termo auto-regressivo estimado. Portanto, estimamos  $\phi$  utilizando as 3000 observações da janela. A próxima observação que será utilizada para comparar com o VaR estimada,  $r_{t+1}$  também será filtrada com o valor de  $\phi$  estimado.

A Tabela 10 apresenta os resultados dos p-valores do backtesting VaR com  $\alpha = 1\%$  e  $5\%$  para os índices das quatro moedas, segundo os três estimadores. A Figura 13 apresenta os valores dos retornos observados e das estimativas do VaR para os três métodos.

Para os testes de cobertura, independência e DQ o ideal é não rejeitar a hipótese nula. Logo, quanto mais próximo de 1 o p-valor é, melhor. Posto isso, para a série do Euro todos os estimadores tiveram um bom desempenho na previsão 1 passo à frente do VaR.

De modo geral vimos que com  $5\%$  o estimador QML-t tem desempenho satisfatório para todas as moedas e para todas as moedas a percentagem de casos com perdas acima do VaR é sempre próximo do valor nominal e em nenhum caso acima. Além disso, o único teste que rejeita a hipótese nula é o do quantil dinâmico para o Yen e o Franco suíço. Porém, os estimadores QML-n e robusto tem desempenho ruim para todas, tendo o robusto um pior desempenho. Observa-se que o estimador QML-n e principalmente o robusto tendem a subestimar os valores do VaR. Para o Franco suíço o estimador QML-n tem desempenho razoável.

Para  $1\%$  a relação se inverte. Aqui o estimador QML-n tem desempenho superior ao do QML-t. Vimos que o estimador robusto teve desempenho razoável para o Euro e Libra. Já o estimador QML-n teve desempenho razoável para todas as moedas, exceto para o Yen. Enquanto o estimador robusto tende a subestimar o valor do VaR, os outros dois tendem a subestimar. Isto pode ser uma indicação da necessidade de uma melhor calibragem da função  $r_{k,\delta,1}(u)$ , definida na Equação (2.9). Duas possibilidades seriam, ser mais difícil considerar uma observação como possível *outlier* e/ou substituir o valor suspeito como *outlier* por um valor maior; isto é, modificar as observações o menor possível.

## 6.4 Análise Bivariada

Novamente, com a ideia de ilustrar o cenário de previsão no qual estamos analisamos, primeiramente apresentamos os resultados das estimações dos parâmetros do modelo na primeira e última janela. Apresentamos os resultados para o modelo IS(2)-DCC(1,1) para a estimação robusta, para a estimação por quase-verossimilhança normal e t-Student.

A Tabela 11 apresenta os parâmetros estimados para a série bivariada do Euro e Libra para os três estimadores considerados nesse trabalho. Nela podemos observar, para

Tabela 10 – Back-testing para o VaR 5% e 1% para as quatro moedas utilizando o modelo MS(2)-GARCH(1,1) sob várias estimações. Janela móvel de tamanho 3000 iniciando em 02-07-2010. P-valores maiores que 0,05 estão em negrito. UC significa o teste de cobertura incondicional de Kupiec (1995). CC e IND são, respectivamente os testes de cobertura condicional e independência de Christoffersen (1998). DQ é o teste de Engle e Manganeli (2004).

| Var          | Estimador | %Hit  | UC           | IND          | CC           | DQ           |
|--------------|-----------|-------|--------------|--------------|--------------|--------------|
| Euro         |           |       |              |              |              |              |
| 5%           | Rob       | 7,042 | 0,000        | <b>0,246</b> | 0,000        | 0,000        |
|              | QML-n     | 6,238 | 0,005        | <b>0,069</b> | 0,004        | 0,008        |
|              | QML-t     | 4,975 | <b>0,953</b> | <b>0,173</b> | <b>0,394</b> | <b>0,785</b> |
| 1%           | Rob       | 1,301 | <b>0,139</b> | <b>0,344</b> | <b>0,214</b> | <b>0,497</b> |
|              | QML-n     | 0,880 | <b>0,530</b> | <b>0,523</b> | <b>0,669</b> | <b>0,949</b> |
|              | QML-t     | 0,612 | 0,032        | <b>0,657</b> | <b>0,091</b> | <b>0,444</b> |
| Libra        |           |       |              |              |              |              |
| 5%           | Rob       | 7,271 | 0,000        | <b>0,371</b> | 0,000        | 0,000        |
|              | QML-n     | 6,659 | 0,000        | <b>0,398</b> | 0,001        | 0,000        |
|              | QML-t     | 4,746 | <b>0,547</b> | <b>0,694</b> | <b>0,772</b> | 0,000        |
| 1%           | Rob       | 1,072 | <b>0,716</b> | 0,036        | <b>0,104</b> | 0,000        |
|              | QML-n     | 0,574 | 0,017        | <b>0,677</b> | <b>0,054</b> | <b>0,360</b> |
|              | QML-t     | 0,536 | 0,009        | <b>0,698</b> | 0,030        | <b>0,315</b> |
| Yen          |           |       |              |              |              |              |
| 5%           | Rob       | 7,310 | 0,000        | <b>0,395</b> | 0,000        | 0,000        |
|              | QML-n     | 5,779 | <b>0,074</b> | <b>0,923</b> | <b>0,203</b> | <b>0,287</b> |
|              | QML-t     | 4,592 | <b>0,333</b> | <b>0,523</b> | <b>0,510</b> | <b>0,767</b> |
| 1%           | Rob       | 1,607 | 0,004        | <b>0,182</b> | 0,007        | 0,004        |
|              | QML-n     | 0,536 | 0,009        | <b>0,698</b> | 0,030        | <b>0,143</b> |
|              | QML-t     | 0,651 | <b>0,055</b> | <b>0,637</b> | <b>0,142</b> | <b>0,569</b> |
| Franco suíço |           |       |              |              |              |              |
| 5%           | Rob       | 7,348 | 0,000        | <b>0,179</b> | 0,000        | 0,000        |
|              | QML-n     | 5,894 | 0,041        | <b>0,749</b> | <b>0,118</b> | <b>0,148</b> |
|              | QML-t     | 4,554 | <b>0,289</b> | <b>0,062</b> | <b>0,100</b> | 0,001        |
| 1%           | Rob       | 1,339 | <b>0,097</b> | <b>0,090</b> | <b>0,060</b> | 0,000        |
|              | QML-n     | 0,918 | <b>0,671</b> | <b>0,505</b> | <b>0,731</b> | <b>0,971</b> |
|              | QML-t     | 0,689 | <b>0,090</b> | <b>0,617</b> | <b>0,210</b> | <b>0,723</b> |

todos os estimadores, mudanças entre as estimativas dos parâmetros da primeira para última janela. Além disso, podemos observar que o estimador robusto tende a estimar um regime com maior probabilidade que o outro e com maior persistência dos regimes ( $p_{11} + p_{22}$ ). Além disso, vimos que em alguns casos, ele tende a estimar um regime com probabilidade estacionária bem maior que a outra.

Complementarmente, a Tabela 12 apresenta os parâmetros estimados para a série bivariada do Euro e Yen para os três estimadores. Nela podemos observar que diferente da Libra,  $\bar{Q}_{12}$  é negativo. Porém, podemos observar que há uma diferença grande entre os parâmetros do regime 2 para todos os parâmetros. Assim como no caso univariado,

podemos observar que, no regime com maior probabilidade as estimativas por QML-n e o Robusto são parecidas.

Por fim, a Tabela 13 apresenta os parâmetros estimados para a série bivariada da Libra e Franco suíço para os três estimadores. Nela também podemos observar que  $\bar{Q}_{12}$  é negativo para as duas janelas. Aqui também temos mudanças nas estimativas entre as janelas. Nas duas janelas, em pelo menos um dos regimes, os estimadores QML-n e QML-t tendem a estimar  $\alpha_i^1$  ou  $\alpha_i^2$  com valores acima de 0,70, exceto para o estimador QML-n na última janela. Valores desta ordem de grandeza normalmente não é esperado porque indica pouca alta dependência com a última observação.

Tabela 11 – Estimativas dos parâmetros do modelo IS(2)-DCC(1,1) para a série bivariado dos log-retornos da taxa de câmbio do Euro e Libra em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto.

|                 | Primeira Janela |        |        | Última Janela |        |        |
|-----------------|-----------------|--------|--------|---------------|--------|--------|
|                 | Rob             | QML-n  | QML-t  | Rob           | QML-n  | QML-t  |
| $\bar{Q}_{12}$  | 0,6817          | 0,6802 | 0,6802 | 0,5517        | 0,5557 | 0,5557 |
| $a_1^{(1)}$     | 0,0233          | 0,0822 | 0,0000 | 0,0059        | 0,0851 | 0,9798 |
| $b_1^{(1)}$     | 0,9284          | 0,9178 | 0,8250 | 0,9763        | 0,9149 | 0,0008 |
| $p_{11}$        | 0,9953          | 0,3399 | 0,9956 | 0,9982        | 0,7668 | 0,7237 |
| $\pi_{1\infty}$ | 0,2034          | 0,3350 | 0,4823 | 0,0526        | 0,3540 | 0,3620 |
| $\bar{Q}_{12}$  | 0,6817          | 0,6802 | 0,6802 | 0,5517        | 0,5557 | 0,5557 |
| $a_1^{(2)}$     | 0,0093          | 0,0066 | 0,3083 | 0,0447        | 0,0024 | 0,0217 |
| $b_1^{(2)}$     | 0,9907          | 0,9929 | 0,6917 | 0,9505        | 0,9968 | 0,9783 |
| $p_{22}$        | 0,9988          | 0,6675 | 0,9959 | 0,9999        | 0,8722 | 0,8432 |
| $\pi_{2\infty}$ | 0,7966          | 0,6650 | 0,5177 | 0,9474        | 0,6460 | 0,6380 |

#### 6.4.1 Estimação das probabilidades dos regimes

De maneira semelhante ao caso univariado, primeiramente iremos trazer as probabilidades inferidas pelo filtro para um par de moedas, aqui escolhemos o par Libra e Euro. Além disso, estimaremos apenas para a primeira janela. A Figura 14 mostra as estimativas das probabilidades de estar no regime (2). Os estimadores QML estimam a probabilidade de estar no regime (2) em quase todo o período, com valores maiores para o estimador QML-t. Entretanto, as estimativas do estimador QML-t são mais suaves que a do estimador QML-n. Já, as estimativas do estimador robusto variam mais ao longo do período, mas com um comportamento muito mais suave. Este comportamento é o esperado em um estimador robusto e é mais adequado se acreditamos que deva existir tanta oscilação entre observações próximas.

Podemos observar a instabilidade dos estimadores QML-n na estimação das

Tabela 12 – Estimativas dos parâmetros do modelo IS(2)-DCC(1,1) para a série bivariado dos log-retornos da taxa de câmbio do Euro e Yen em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto,

|                 | Primeira Janela |         |         | Última Janela |         |         |
|-----------------|-----------------|---------|---------|---------------|---------|---------|
|                 | Rob             | QML-n   | QML-t   | Rob           | QML-n   | QML-t   |
| $\bar{Q}_{12}$  | -0,2808         | -0,2611 | -0,2611 | -0,3091       | -0,3085 | -0,3085 |
| $a_1^{(1)}$     | 0,1166          | 0,4764  | 0,1485  | 0,0952        | 0,1539  | 0,9902  |
| $b_1^{(1)}$     | 0,8588          | 0,5236  | 0,7367  | 0,8438        | 0,8461  | 0,0000  |
| $p_{11}$        | 0,7077          | 0,0001  | 0,9020  | 0,9227        | 0,5135  | 0,7086  |
| $\pi_{1\infty}$ | 0,2741          | 0,1911  | 0,4793  | 0,4665        | 0,4439  | 0,3212  |
| $\bar{Q}_{12}$  | -0,2808         | -0,2611 | -0,2611 | -0,3091       | -0,3085 | -0,3085 |
| $a_1^{(2)}$     | 0,0237          | 0,0146  | 0,1854  | 0,0264        | 0,0050  | 0,2176  |
| $b_1^{(2)}$     | 0,9763          | 0,9848  | 0,8146  | 0,9723        | 0,9887  | 0,7824  |
| $p_{22}$        | 0,8896          | 0,7638  | 0,9098  | 0,9324        | 0,6116  | 0,8621  |
| $\pi_{2\infty}$ | 0,7259          | 0,8089  | 0,5207  | 0,5335        | 0,5561  | 0,6788  |

Tabela 13 – Estimativas dos parâmetros do modelo IS(2)-DCC(1,1) para a série bivariado dos log-retornos da taxa de câmbio do Libra e Franco Suíço em relação ao dólar americano utilizando os estimadores de quase verossimilhança normal, t e o estimador robusto.

|                 | Primeira Janela |         |         | Última Janela |         |         |
|-----------------|-----------------|---------|---------|---------------|---------|---------|
|                 | Rob             | QML-n   | QML-t   | Rob           | QML-n   | QML-t   |
| $\bar{Q}_{12}$  | -0,6282         | -0,5959 | -0,5959 | -0,4326       | -0,3755 | -0,3755 |
| $a_1^{(1)}$     | 0,0300          | 0,8149  | 0,9662  | 0,0623        | 0,0000  | 0,0565  |
| $b_1^{(1)}$     | 0,9700          | 0,0000  | 0,0264  | 0,9349        | 0,7213  | 0,9435  |
| $p_{11}$        | 0,7577          | 0,8539  | 0,5825  | 0,9898        | 0,6848  | 0,9812  |
| $\pi_{1\infty}$ | 0,4045          | 0,0482  | 0,1841  | 0,0097        | 0,3056  | 0,4835  |
| $\bar{Q}_{12}$  | -0,6282         | -0,5959 | -0,5959 | -0,4326       | -0,3755 | -0,3755 |
| $a_1^{(2)}$     | 0,0121          | 0,0139  | 0,0585  | 0,0239        | 0,0478  | 0,7186  |
| $b_1^{(2)}$     | 0,9879          | 0,9861  | 0,9415  | 0,9341        | 0,9522  | 0,2647  |
| $p_{22}$        | 0,8354          | 0,9926  | 0,9058  | 0,9999        | 0,8613  | 0,9824  |
| $\pi_{2\infty}$ | 0,5955          | 0,9518  | 0,8159  | 0,9903        | 0,6944  | 0,5165  |

probabilidades, uma vez que elas variam muito de um regime para o outro. Aqui podemos observar que o QML-t tende a permanecer em apenas um regime.

#### 6.4.2 Análise das estimativas do VaR e das previsões

Para a análise do VaR no caso bivariado iremos utilizar a série do log retorno do portfólio dado em (6.2). Aqui, utilizamos o log retorno  $r_t = \log(P_t/P_{t-1})$  e  $w_1 = w_2 = 0,5$ . No caso bivariado também filtramos as séries por AR(1). Além disso, para a obtenção das

volatilidades das marginais, utilizamos os modelos MS-GARCH estimado pelos respectivos estimadores. Isto é, se estimamos o modelo IS-DCC com o estimador QML-n, estimamos a volatilidade do modelo MS-GARCH através do estimador QML-n.

A Tabela 14 e a Figura 15 apresentam os resultados dos backtesting VaR com  $\alpha = 1\%$  e  $5\%$  para os casos bivariados do Euro com a libra e com o Yen e libra com franco suíço. Para  $5\%$  o estimador QML-n teve um bom desempenho nos três casos, exceto para o caso Euro-Libra, mas mesmo neste caso seu desempenho foi melhor do que os outros dois estimadores. O estimador QML-t teve o pior desempenho com a menor percentagem *Hits* igual a  $7,4\%$ .

Tabela 14 – Back-testing para o VaR  $5\%$  e  $1\%$  para as os casos de Euro-Libra, Euro-Yen e Libra-Franco Suíço. Janela móvel de tamanho 3000 iniciando em 02-07-2010. P valores menores que 0,05 estão em negrito. UC significa o teste de cobertura incondicional de Kupiec (1995). CC e IND são, respectivamente os testes de cobertura condicional e independência de Christoffersen (1998). DQ é o teste de Engle e Manganelli (2004).

| %Var                 | Estimador | %Hit  | UC           | IND          | CC           | DQ           |
|----------------------|-----------|-------|--------------|--------------|--------------|--------------|
| Euro - Libra         |           |       |              |              |              |              |
| 5%                   | Rob       | 7,348 | 0,000        | <b>0,746</b> | 0,000        | 0,000        |
|                      | QML-n     | 6,353 | 0,002        | <b>0,384</b> | 0,007        | 0,031        |
|                      | QML-t     | 7,424 | 0,000        | <b>0,101</b> | 0,000        | 0,000        |
| 1%                   | Rob       | 0,842 | <b>0,404</b> | 0,013        | 0,032        | 0,003        |
|                      | QML-n     | 0,689 | <b>0,090</b> | <b>0,617</b> | <b>0,210</b> | <b>0,656</b> |
|                      | QML-t     | 2,679 | 0,000        | <b>0,927</b> | 0,000        | 0,000        |
| Euro - Yen           |           |       |              |              |              |              |
| 5%                   | Rob       | 7,118 | 0,000        | <b>0,428</b> | 0,000        | 0,000        |
|                      | QML-n     | 5,205 | <b>0,633</b> | <b>0,463</b> | <b>0,682</b> | <b>0,770</b> |
|                      | QML-t     | 8,419 | 0,000        | <b>0,233</b> | 0,000        | 0,000        |
| 1%                   | Rob       | 0,957 | <b>0,823</b> | <b>0,487</b> | <b>0,766</b> | <b>0,963</b> |
|                      | QML-n     | 0,689 | <b>0,090</b> | <b>0,617</b> | <b>0,210</b> | <b>0,740</b> |
|                      | QML-t     | 2,641 | 0,000        | <b>0,495</b> | 0,000        | 0,000        |
| Libra - Franco Suíço |           |       |              |              |              |              |
| 5%                   | Rob       | 7,348 | 0,000        | 0,034        | 0,000        | 0,000        |
|                      | QML-n     | 5,702 | <b>0,107</b> | <b>0,382</b> | <b>0,186</b> | <b>0,435</b> |
|                      | QML-t     | 9,147 | 0,000        | <b>0,620</b> | 0,000        | 0,000        |
| 1%                   | Rob       | 1,263 | <b>0,194</b> | <b>0,071</b> | <b>0,084</b> | <b>0,060</b> |
|                      | QML-n     | 0,651 | <b>0,055</b> | <b>0,637</b> | <b>0,142</b> | <b>0,626</b> |
|                      | QML-t     | 2,755 | 0,000        | 0,043        | 0,000        | 0,000        |

Para o caso de  $1\%$  o estimador robusto teve melhor desempenho para Euro-Libra e Euro-Yen. No caso Libra-Franco Suíço, claramente o estimador QML-t teve o pior desempenho, enquanto não fica claro na comparação entre os outros dois estimadores. Em termos de *Hits*, embora o estimador robusto teve menor diferença percentual em relação ao valor nominal,  $26,5\%$  comparado com  $34,9\%$  do estimador QML-n. Entretanto o estimador robusto tende a subestimar o valor do VaR, enquanto o estimador QML-n tende a subestimar.

Em termos de rejeição das hipóteses nulas nos testes de desempenho das estimativas do VaR, o teste DQ de (ENGLE; MANGANELLI, 2004) rejeita a hipótese nula em 11 dos 24 testes ao nível de 5% enquanto o teste CC de (CHRISTOFFERSEN, 1998) rejeita em 9 dos testes. As rejeições das hipóteses nulas parecem ser mais devido à percentagem de cobertura (10 casos), teste Uc de (KUPIEC, 1995) do que de independência (3 casos), teste IND de (CHRISTOFFERSEN, 1998). Destes três casos, dois são para o estimador robusto e nenhum para o estimador QML-n.

## 6.5 Comentários Finais

O objetivo desse capítulo era exemplificar a performance dos estimadores em séries reais como também apresentar como é feita a previsão um passo a frente dos modelos univariado e bivariado.

Em relação ao caso univariado, vimos que o estimador QML-t teve desempenho bom para o VaR 5%, enquanto que para 1%, o robusto teve melhor desempenho. O estimador robusto proposto tende a subestimar a volatilidade, enquanto os estimadores QML-t e QML-n tendem a subestimar o VaR 5% e superestimar o VaR 1%. Uma justificativa possível para esse problema pode ser o uso do mecanismo BIP.

Em relação ao caso bivariado, podemos ver claramente que o estimador robusto teve melhor desempenho para 1%. Para 5% todos os estimadores tendem a subestimar o VaR. Para 1% podemos observar que o estimador QML-n subestima o VaR e o QML-t superestima.

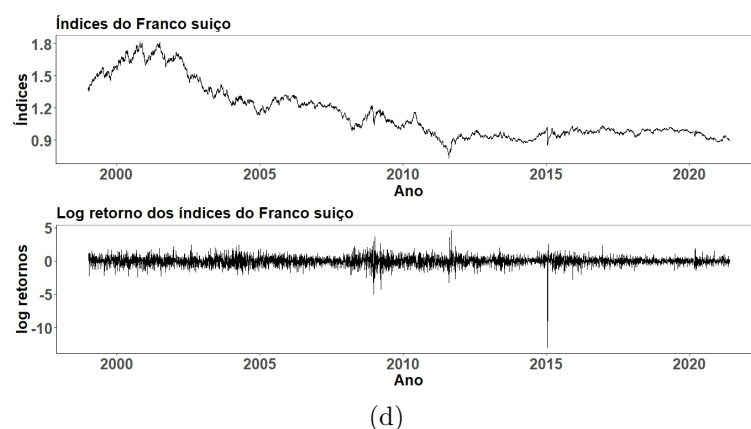
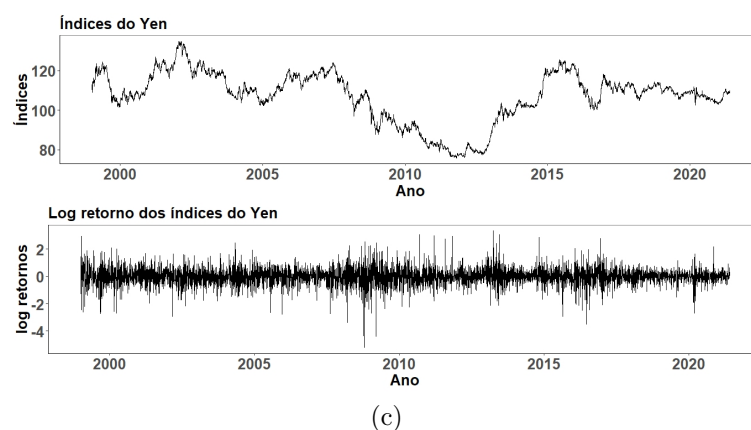
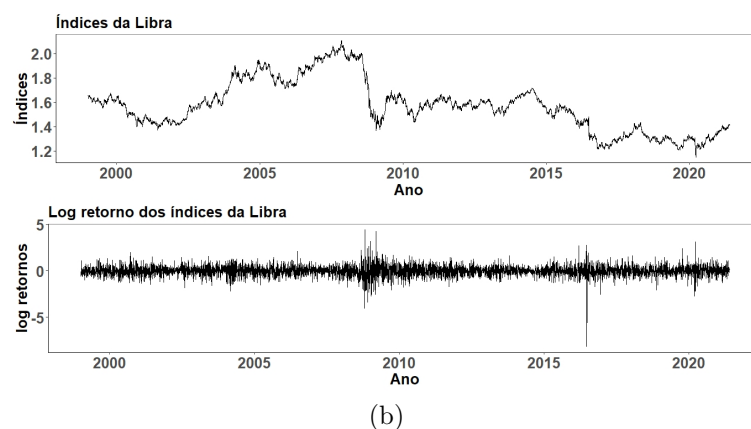
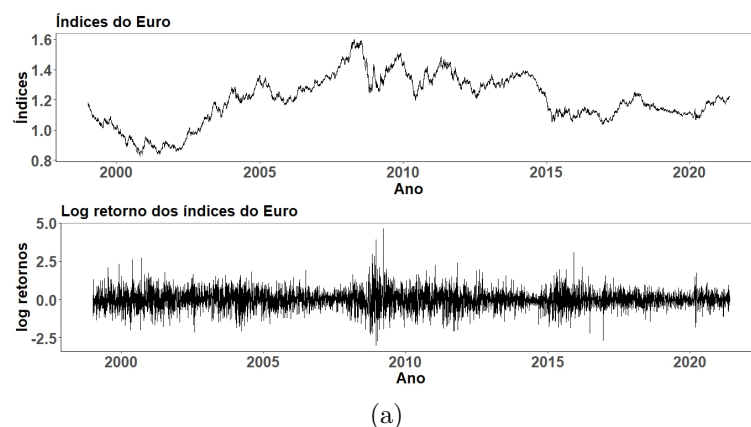
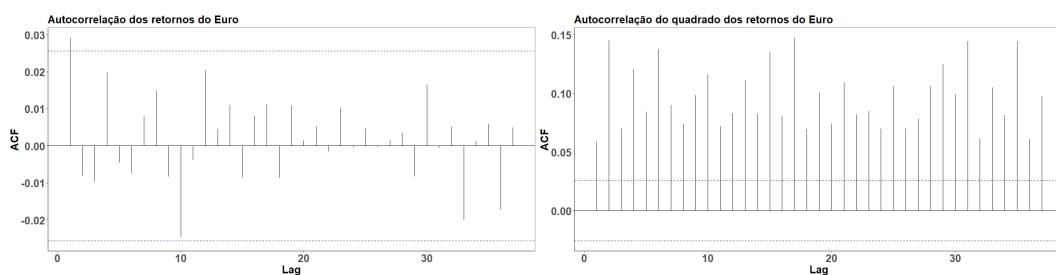
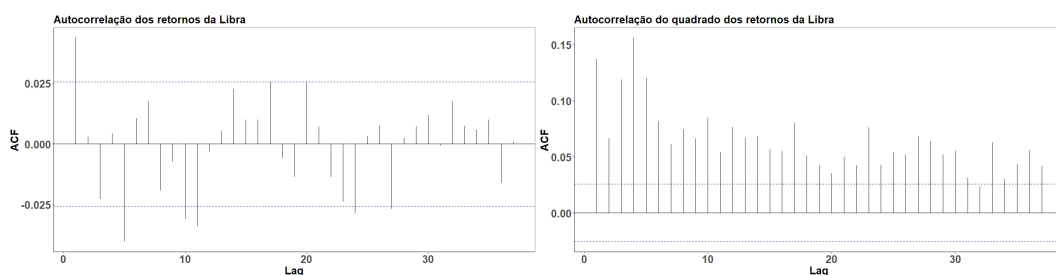


Figura 10 – Evolução diária das taxas de câmbio em relação ao dólar americano no gráfico superior e no inferior os respectivos log-retornos (a) Euro (b) Libra (c) Yen (d) Franco suíço.

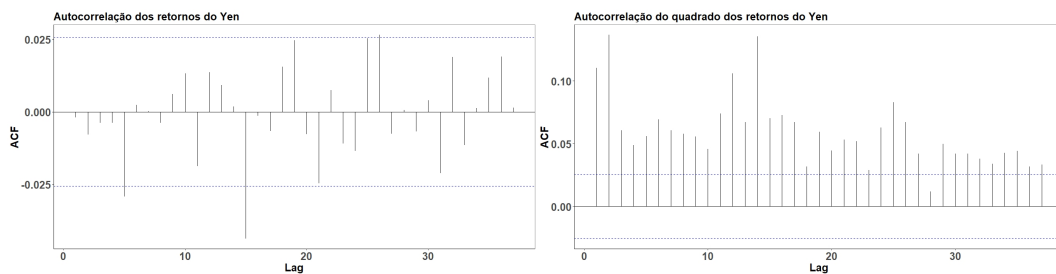




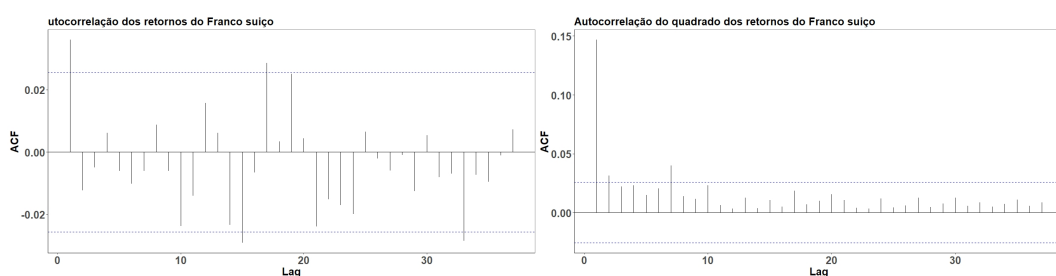
(a) Euro



(b) Libra

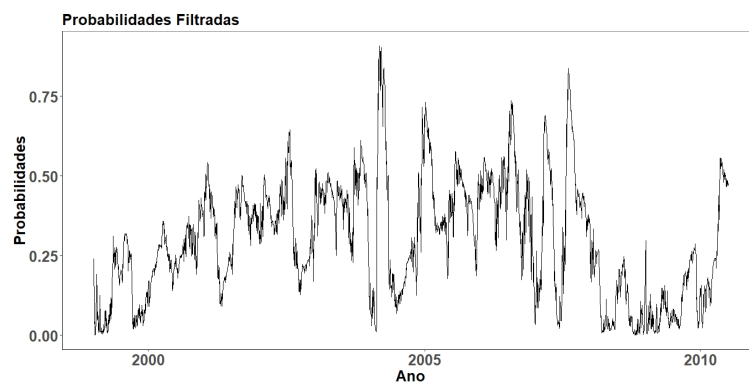


(c) Yen

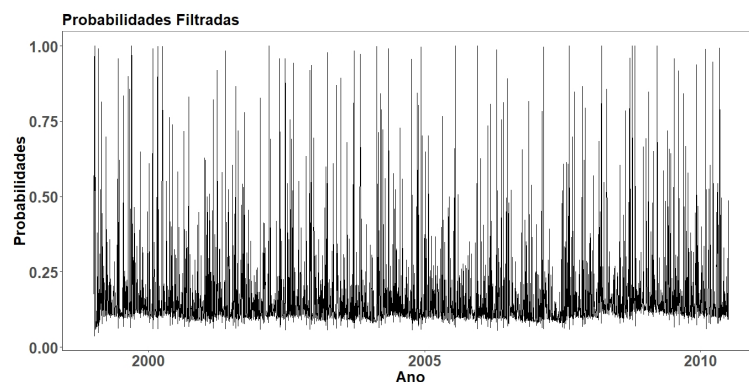


(d) Franco suíço

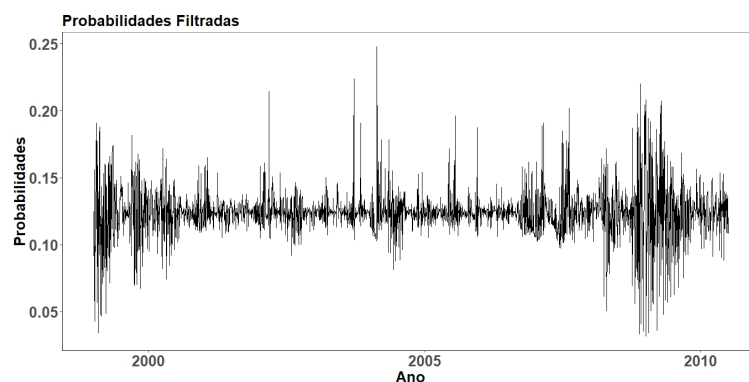
Figura 11 – Do lado esquerdo a função de auto-correlação amostral dos log-retornos e do lado direito a função de auto-correlação amostral dos log-retornos ao quadrado.



(a)

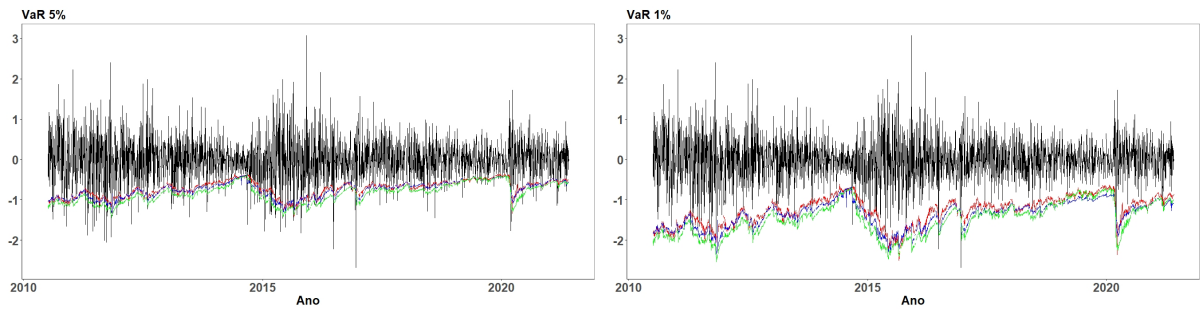


(b)

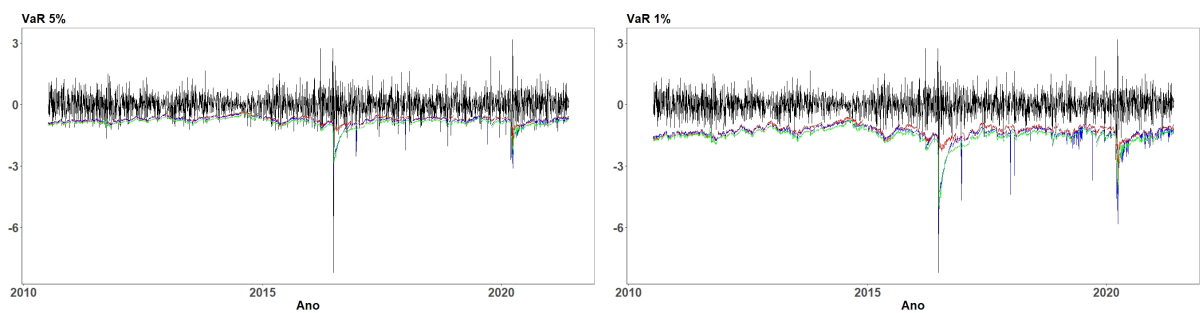


(c)

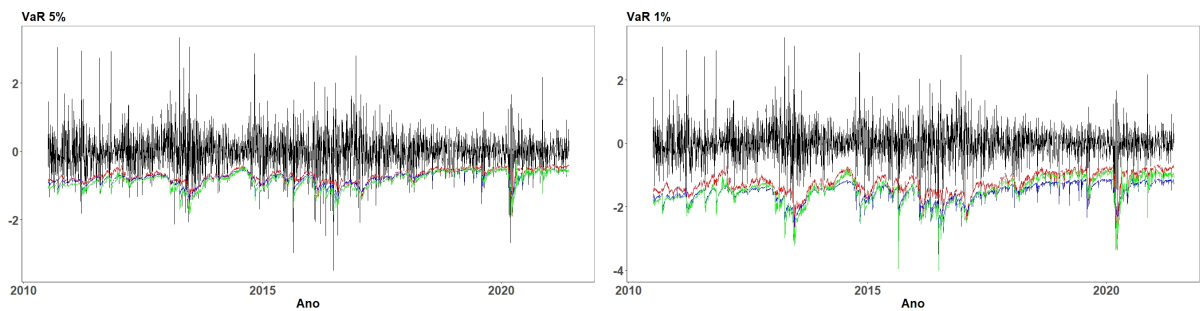
Figura 12 – Probabilidades estimadas de estar no regime de alta volatilidade pelo os três estimadores para o Yen, utilizando a primeira janela de estimação. Os estimadores são: (a) Robusto (b) QML-n (c) QML-t.



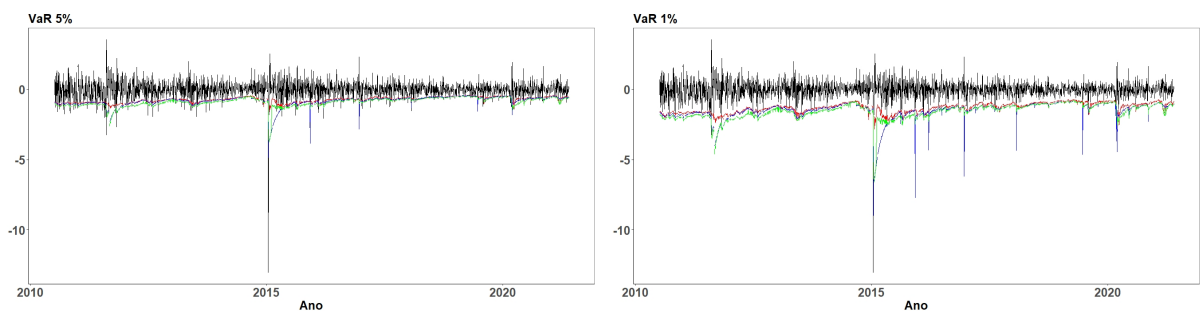
(a) Euro



(b) Libra

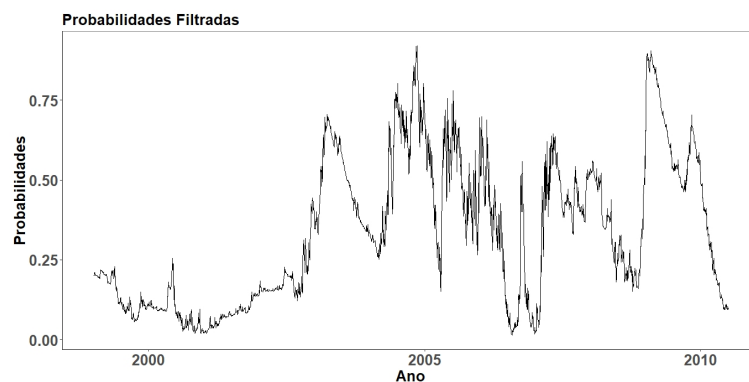


(c) Yen

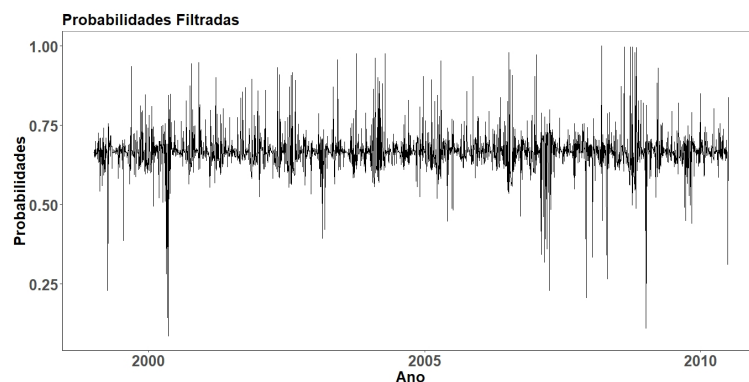


(d) Franco suíço

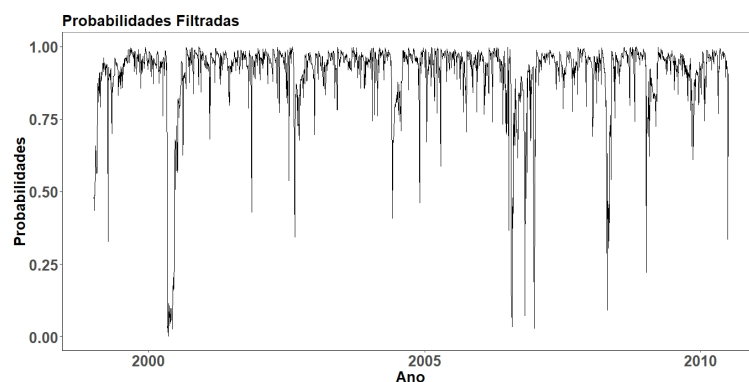
Figura 13 – Estimativa do VaR para as quatro moedas. A linha de cor preta representa o log retorno filtrado por um AR(1), a linha representa o VaR estimado pelo estimador robusto, a linha em azul por QML-n e a linha verde por QML-t. Do lado esquerdo o VaR 5% e do direito VaR 1%.



(a)

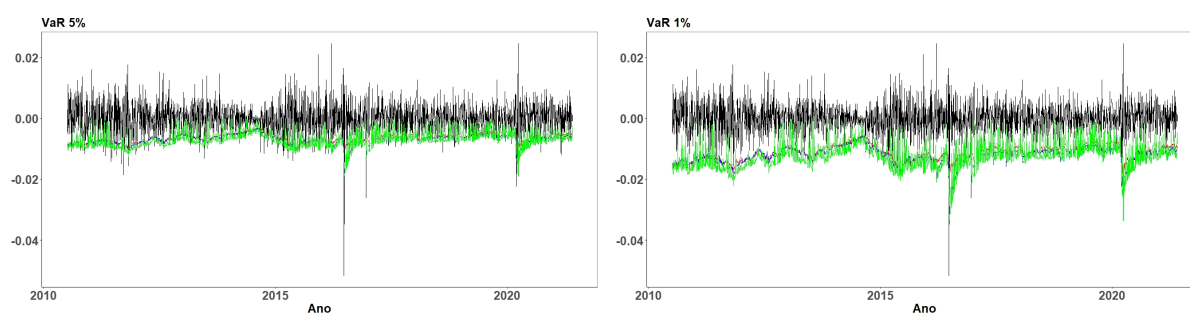


(b)

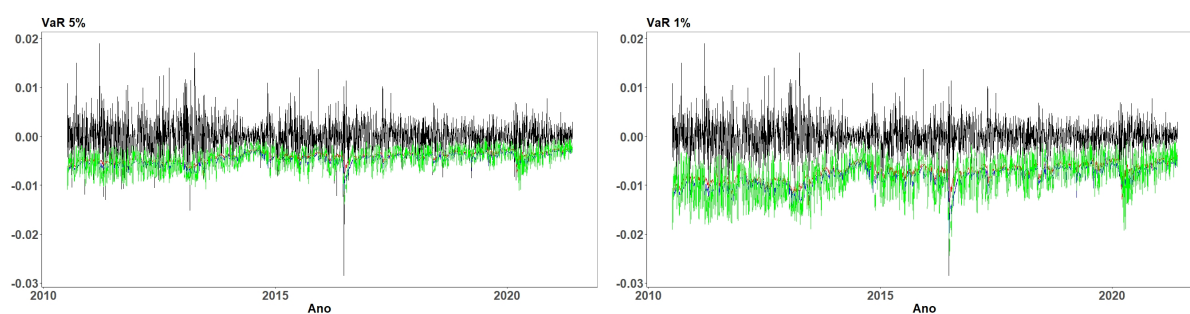


(c)

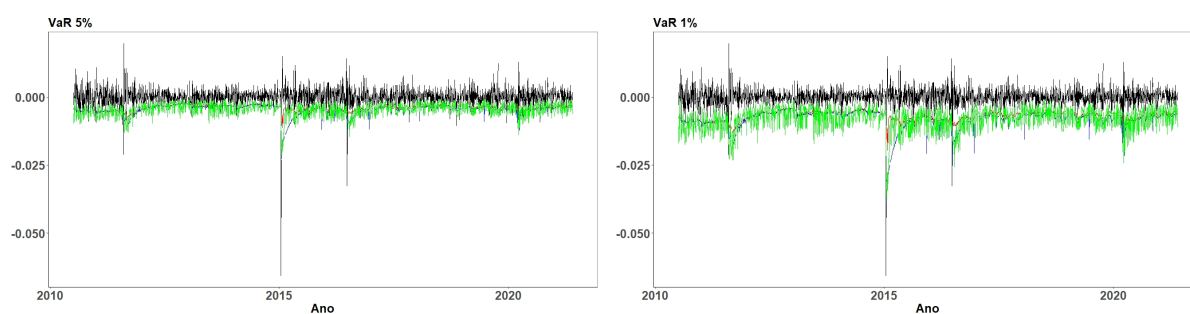
Figura 14 – Probabilidades estimadas de estar no regime de alta correlação pelo os três estimadores para o par Euro-Libra, utilizando a primeira janela de estimação. Os estimadores são: (a) Robusto (b) QML-n (c) QML-t.



(a) Euro - Libra



(b) Euro - Yen



(c) Libra - Franco Suíço

Figura 15 – Estimativa do VaR para os dois pares de moedas. A linha de cor preta representa o log retorno da carteira, a linha representa o VaR estimado pelo estimador robusto, a linha em azul por QML-n e a linha verde por QML-t. Do lado esquerdo o VaR 5% e do direito VaR 1%.

## 7 Considerações Finais

O objetivo desse trabalho é avaliar o desempenho de diversas técnicas de estimação para o modelo GARCH e o modelo DCC com mudança de regime markoviana na presença de outliers aditivos. Uma vez que na literatura não está bem difundido ou não há estudos de técnicas robustas para a estimação desses modelos, este trabalho contribui com estudos de simulação de performance para diversos cenários de contaminação de outliers. O método sugerido adapta os métodos de estimação robusta apresentada por Boudt, Danielsson e Laurent (2013) para o caso de mudança de regime markoviana.

Os resultados desse trabalho podem ser divididos em duas partes, nos resultados das simulações Monte Carlo e nos resultados da Aplicação. Em relação aos resultados da simulação, as principais conclusões que podemos retirar são:

- Para o modelo MS-GARCH, o estimador QML utilizando a distribuição normal como a densidade condicional de cada regime tem desempenho catastrófico conforme a proporção e o tamanho de outliers aumenta. Observou-se que a estimação tem desempenho ruim na capacidade de classificação do filtro e na precisão das estimativas. Utilizando a estimação QML com distribuição t-Student com 4 graus de liberdade como sendo a densidade condicional de cada regime, observamos que ela tem um desempenho razoável na capacidade de classificação do filtro e na estimação de alguns parâmetros das equações de volatilidade. Já a estimação robusta, vimos que ele também tem um desempenho razoável na capacidade de classificação do filtro e na estimação de alguns parâmetros das equações da volatilidade.
- Para o modelo IS-DCC, também foi observado que o estimador QML utilizando a distribuição normal como a densidade condicional de cada regime tem desempenho catastrófico conforme a proporção e o tamanho de outliers aumenta. Observou-se que o estimador tem desempenho ruim tanto capacidade de classificação do filtro e quanto na precisão das estimativas. O estimador QML com distribuição t-Student com 4 graus de liberdade como sendo a densidade condicional de cada regime, teve um desempenho razoável na capacidade de classificação do filtro, porém não tão bom na estimação dos parâmetros das equações da dinâmica de correlação. Já o estimador robusto teve um desempenho mediano na capacidade de classificação do filtro e bom na estimação dos parâmetros das equações da dinâmica de correlação.

Para avaliar a capacidade preditiva dos estimadores apresentados, avaliamos a performance da previsão do VaR um passo a frente para a cotação de 4 moedas em relação ao dólar. As principais conclusões que podemos são:

- Para o caso univariado, o estimador robusto tende a subestimar a volatilidade para o VaR 5% e isso ocorre provavelmente devido ao uso do mecanismo BIP. Para os dados utilizados, o estimador de quase verossimilhança com distribuição t tem desempenho bom na previsão um passo a frente do VaR 5%. Para o VaR 1% o estimador robusto teve melhor desempenho.
- Para o caso bivariado, os três estimadores tendem a subestimar a volatilidade para o VaR 5%. Para o VaR 1% o estimador robusto teve melhor desempenho e os estimadores QML-n e QML-t subestimam e superestimam o VaR, respectivamente.

Como trabalhos futuros podemos citar:

- Estudar o BIP para o caso de mudança de regime;
- Em particular, estudar possíveis mudança na função  $r_{k,\delta,1}(u)$ , definida na Equação (2.9), incluindo diferentes formas funcionais;
- Adaptação de outros estimadores robustos para a estimação dos parâmetros MS-GARCH e IS-DCC;
- Estudar o comportamento de técnicas robustas em relação a filtros diferentes e em relação a estimação Bayesiana;
- Considerar a viabilidade de colocar probabilidades de transição como variantes no tempo, como, por exemplo, em Diebold, Lee e Weinbach (1994), Weinbach (1999), Wang et al. (2020), Psaradakis e Sola (2021) e Wang et al. (2022);
- Comparar com o método de indicador de saturação (*Indicator Saturation Method*) desenvolvido por Santos, Hendry e Johansen (2008) e aplicado, por exemplo, em Bauwens e Sucarrat (2010) e Pretis, Reade e Sucarrat (2018).

## Referências

- ABRAMSON, A.; COHEN, I. On the stationarity of Markov-switching GARCH processes. *Econometric Theory*, p. 485–500, 2007. Citado 3 vezes nas páginas 33, 35 e 41.
- AIELLI, G. P. Dynamic conditional correlation: on properties and estimation. *Journal of Business & Economic Statistics*, v. 31, n. 3, p. 282–299, 2013. Citado 2 vezes nas páginas 46 e 49.
- AKNOUCHE, A.; GUERBYENNE, H. Recursive estimation of GARCH models. *Communications in Statistics-Simulation and Computation*, v. 35, n. 4, p. 925–938, 2006. Citado na página 24.
- ALEXANDER, C.; CHIBUMBA, A. Multivariate orthogonal factor GARCH. *University of Sussex, Mimeo*, 1997. Citado na página 45.
- ALMEIDA, D. de; HOTTA, L. K.; RUIZ, E. MGARCH models: Trade-off between feasibility and flexibility. *International Journal of Forecasting*, v. 34, n. 1, p. 45–63, 2018. Citado na página 44.
- ANDREOU, E.; GHYSELS, E. Structural breaks in financial time series. *Handbook of Financial Time Series*, p. 839–870, 2009. Citado na página 39.
- ARDIA, D. *Financial risk management with Bayesian estimation of GARCH models: Theory and applications*. Primeira ed. Berlin, Heidelberg: Springer, 2008. Citado na página 24.
- ARDIA, D.; BLUTEAU, K.; BOUDT, K.; CATANIA, L. Forecasting risk with Markov-switching GARCH models: A large-scale performance study. *International Journal of Forecasting*, v. 34, n. 4, p. 733–747, 2018. Citado 5 vezes nas páginas 18, 76, 79, 81 e 82.
- ARDIA, D.; BLUTEAU, K.; BOUDT, K.; CATANIA, L.; TROTTIER, D.-A. Markov-switching GARCH models in R: The MSGARCH package. *Journal of Statistical Software*, v. 91, n. 4, 2019. Citado na página 41.
- AUGUSTYNIAK, M. Maximum likelihood estimation of the Markov-switching GARCH model. *Computational Statistics & Data Analysis*, v. 76, p. 61–75, 2014. Citado 5 vezes nas páginas 12, 39, 58, 59 e 66.
- BAILLIE, R. T.; CHUNG, H. Estimation of GARCH models from the autocorrelations of the squares of a process. *Journal of Time Series Analysis*, v. 22, n. 6, p. 631–650, 2001. Citado na página 24.
- BASU, A.; HARRIS, I. R.; HJORT, N. L.; JONES, M. Robust and efficient estimation by minimising a density power divergence. *Biometrika*, v. 85, n. 3, p. 549–559, 1998. Citado na página 27.
- BAUWENS, L.; PREMINGER, A.; ROMBOUTS, J. V. Theory and inference for a Markov switching GARCH model. *The Econometrics Journal*, v. 13, n. 2, p. 218–244, 2010. Citado 6 vezes nas páginas 12, 18, 34, 37, 58 e 59.



- BAUWENS, L.; SUCARRAT, G. General-to-specific modelling of exchange rate volatility: A forecast evaluation. *International Journal of Forecasting*, v. 26, n. 4, p. 885–907, 2010. Citado na página 99.
- BILLIO, M.; CAPORIN, M. Multivariate Markov switching dynamic conditional correlation GARCH representations for contagion analysis. *Statistical methods and applications*, v. 14, n. 2, p. 145–161, 2005. Citado na página 53.
- BILLIO, M.; CASARIN, R.; OSUNTUYI, A. Efficient Gibbs sampling for Markov switching GARCH models. *Computational Statistics & Data Analysis*, v. 100, p. 37–57, 2016. Citado 5 vezes nas páginas 12, 32, 58, 59 e 66.
- BLACK, F.; SCHOLES, M. The pricing of options and corporate liabilities. *Journal of Political Economy*, v. 81, n. 3, p. 637–654, 1973. Citado na página 16.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, v. 31, n. 3, p. 307–327, 1986. Citado 3 vezes nas páginas 17, 21 e 22.
- BOLLERSLEV, T.; ENGLE, R. F.; WOOLDRIDGE, J. M. A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, v. 96, n. 1, p. 116–131, 1988. Citado na página 44.
- BOUDT, K.; DANIELSSON, J.; LAURENT, S. Robust forecasting of dynamic conditional correlation GARCH models. *International Journal of Forecasting*, v. 29, n. 2, p. 244–257, 2013. Citado 16 vezes nas páginas 17, 19, 20, 28, 29, 30, 31, 40, 49, 50, 51, 56, 57, 58, 60 e 98.
- CAI, J. A Markov model of switching-regime Arch. *Journal of Business & Economic Statistics*, v. 12, n. 3, p. 309–316, 1994. Citado 2 vezes nas páginas 18 e 32.
- CALVET, L. E.; CZELLAR, V.; RONCHETTI, E. Robust filtering. *Journal of the American Statistical Association*, v. 110, n. 512, p. 1591–1606, 2015. Citado 3 vezes nas páginas 19, 40 e 56.
- CAMPBELL, J. Y.; LO, A.; MACKINLAY, A. C. *The Econometrics of Financial Markets*. Segunda ed. Princeton, NJ: Princeton University Press, 1997. Citado na página 22.
- CARNERO, M. A.; PEÑA, D.; RUIZ, E. Estimating GARCH volatility in the presence of outliers. *Economics Letters*, v. 114, n. 1, p. 86–90, 2012. Citado 5 vezes nas páginas 18, 28, 29, 31 e 60.
- CARNERO, M. A.; PEÑA, D.; RUIZ, E. Effects of outliers on the identification and estimation of GARCH models. *Journal of Time Series Analysis*, v. 28, n. 4, p. 471–497, 2007. Citado na página 26.
- CHRISTOFFERSEN, P. F. Evaluating interval forecasts. *International Economic Review*, p. 841–862, 1998. Citado 5 vezes nas páginas 13, 81, 87, 90 e 91.
- CROSATO, L.; GROSSI, L. Correcting outliers in GARCH models: a weighted forward approach. *Statistical Papers*, v. 60, n. 6, p. 1939–1970, 2019. Citado na página 27.
- DAS, D.; YOO, B. H. et al. A Bayesian MCMC algorithm for Markov switching GARCH models. *Econometric Society*, p. 1–16, 2004. Citado 2 vezes nas páginas 18 e 37.

- DIEBOLD, F. X.; LEE, J.-H.; WEINBACH, G. C. Regime switching with time-varying transition probabilities. In: HARGREAVES, C. (Ed.). *Nostationary Time Series Analysis and Cointegration*. [S.l.]: Oxford University Press, 1994. p. 1998. Citado na página 99.
- ENGLE, R. GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, v. 15, n. 4, p. 157–168, 2001. Citado na página 21.
- \_\_\_\_\_. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, v. 20, n. 3, p. 339–350, 2002. Citado na página 45.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, p. 987–1007, 1982. Citado 2 vezes nas páginas 17 e 21.
- ENGLE, R. F.; KRONER, K. F. Multivariate simultaneous generalized ARCH. *Econometric Theory*, p. 122–150, 1995. Citado na página 45.
- ENGLE, R. F.; MANGANELLI, S. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, v. 22, n. 4, p. 367–381, 2004. Citado 5 vezes nas páginas 13, 82, 87, 90 e 91.
- ENGLE, R. F.; SHEPPARD, K. *Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH*. [S.l.], 2001. Citado na página 48.
- FERMANIAN, J.-D.; MALONGO, H. On the stationarity of dynamic conditional correlation models. *Econometric Theory*, v. 33, n. 3, p. 636–663, 2017. Citado na página 46.
- FRANCQ, C.; ROUSSIGNOL, M.; ZAKOIAN, J.-M. Conditional heteroskedasticity driven by hidden Markov chains. *Journal of Time Series Analysis*, v. 22, n. 2, p. 197–220, 2001. Citado na página 34.
- FRANCQ, C.; ZAKOIAN, J.-M. *GARCH models: structure, statistical inference and financial applications*. Segunda ed. Hoboken, NJ: John Wiley & Sons, 2019. Citado 8 vezes nas páginas 17, 21, 22, 24, 25, 44, 57 e 80.
- FRANCQ, C.; ZAKOIAN, J.-M. Deriving the autocovariances of powers of Markov-switching GARCH models, with applications to statistical inference. *Computational Statistics & Data Analysis*, v. 52, n. 6, p. 3027–3046, 2008. Citado na página 37.
- GAGLIANONE, W. P.; LIMA, L. R.; LINTON, O.; SMITH, D. R. Evaluating value-at-risk models via quantile regression. *Journal of Business & Economic Statistics*, v. 29, n. 1, p. 150–160, 2011. Citado 2 vezes nas páginas 81 e 82.
- GONZÁLEZ-RIVERA, G.; LEE, T.-H.; MISHRA, S. Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of Forecasting*, v. 20, n. 4, p. 629–645, 2004. Citado na página 82.
- GRANÉ, A.; VEIGA, H.; MARTÍN-BARRAGÁN, B. Additive level outliers in multivariate GARCH models. In: *Topics in Statistical Simulation*. [S.l.: s.n.], 2014. p. 247–255. Citado na página 49.

- GRAY, S. F. Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, v. 42, n. 1, p. 27–62, 1996. Citado 7 vezes nas páginas 18, 32, 33, 34, 35, 53 e 66.
- HAAS, M. Covariance forecasts and long-run correlations in a Markov-switching model for dynamic correlations. *Finance Research Letters*, v. 7, n. 2, p. 86–97, 2010. Citado na página 52.
- HAAS, M.; LIU, J.-C. A multivariate regime-switching GARCH model with an application to global stock market and real estate equity returns. *Studies in Nonlinear Dynamics & Econometrics*, v. 22, n. 3, 2018. Citado na página 52.
- HAAS, M.; MITTNIK, S.; PAOLELLA, M. S. A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics*, v. 2, n. 4, p. 493–530, 2004. Citado 14 vezes nas páginas 12, 18, 19, 22, 33, 34, 35, 37, 39, 53, 54, 58, 59 e 77.
- HAMILTON, J. D. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, p. 357–384, 1989. Citado 4 vezes nas páginas 18, 19, 20 e 37.
- HAMILTON, J. D.; SUSMEL, R. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, v. 64, n. 1-2, p. 307–333, 1994. Citado na página 32.
- HILLEBRAND, E. Neglecting parameter changes in GARCH models. *Journal of Econometrics*, v. 129, n. 1-2, p. 121–138, 2005. Citado 2 vezes nas páginas 18 e 32.
- HORVÁTH, L.; KOKOSZKA, P.; ZHANG, A. Monitoring constancy of variance in conditionally heteroskedastic time series. *Econometric Theory*, p. 373–402, 2006. Citado na página 32.
- HOTTA, L. K.; TRUCÍOS, C. Inference in (M) GARCH models in the presence of additive outliers: Specification, estimation, and prediction. In: LAVOR, C.; NETO, F. d. A. M. G. (Ed.). *Advances in Mathematics and Applications*. [S.l.: s.n.], 2018. p. 179–202. Citado na página 27.
- HOTTA, L. K.; TSAY, R. S. Outliers in GARCH processes. In: BELL, W. R.; HOLAN, S. H.; MCELROY, T. S. (Ed.). *Economic Time Series: Modeling and Seasonality*. [S.l.: s.n.], 2012. p. 337–358. Citado 3 vezes nas páginas 17, 26 e 27.
- HUBER, P. J.; RONCHETTI. *Robust Statistics*. Segunda ed. Hoboken, NJ: John Wiley & Sons, 2009. Citado 2 vezes nas páginas 19 e 40.
- HULL, J. C. *Options, Futures, and Other Derivatives*. Décima primeira ed. United Kingdom: Pearson Education, 2022. Citado na página 16.
- HWANG, S.; PEREIRA, P. L. V. The effects of structural breaks in ARCH and GARCH parameters on persistence of GARCH models. *Communications in Statistics—Simulation and Computation*, v. 37, n. 3, p. 571–578, 2008. Citado 2 vezes nas páginas 18 e 32.
- KUAN, C.-M. Lecture on the Markov switching model. *Institute of Economics Academia Sinica*, p. 1–30, 2002. Citado na página 18.

- KUPIEC, P. Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, v. 3, n. 2, p. 73–84, 1995. Citado 5 vezes nas páginas 13, 81, 87, 90 e 91.
- LABARR, A. D. *Multivariate robust estimation of DCC-GARCH volatility model*. Tese (Doutorado) — North Carolina State University, 2010. Citado na página 48.
- LAMOUREUX, C. G.; LASTRAPES, W. D. Persistence in variance, structural change, and the GARCH model. *Journal of Business & Economic Statistics*, v. 8, n. 2, p. 225–234, 1990. Citado na página 32.
- LEE, H.-T. Regime switching correlation hedging. *Journal of Banking & Finance*, v. 34, n. 11, p. 2728–2741, 2010. Citado 5 vezes nas páginas 14, 19, 52, 53 e 54.
- LEE, S.; SONG, J. Minimum density power divergence estimator for GARCH models. *Test*, v. 18, n. 2, p. 316–341, 2009. Citado 4 vezes nas páginas 12, 27, 59 e 60.
- LIU, J.-C. Stationarity of a Markov-switching GARCH model. *Journal of Financial Econometrics*, v. 4, n. 4, p. 573–593, 2006. Citado na página 35.
- MANGANELLI, S.; ENGLE, R. F. Value at risk models in finance. 2001. Citado na página 76.
- MERTON, R. C. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, p. 141–183, 1973. Citado na página 16.
- MIKOSCH, T.; STĂRICĂ, C. Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *Review of Economics and Statistics*, v. 86, n. 1, p. 378–390, 2004. Citado na página 32.
- MORETTIN, P. A.; TOLOI, C. *Análise de Séries Temporais*. Terceira ed. São Paulo: Blucher, 2018. Citado na página 24.
- MULER, N.; YOHAI, V. J. Robust estimates for GARCH models. *Journal of Statistical Planning and Inference*, v. 138, n. 10, p. 2918–2940, 2008. Citado 8 vezes nas páginas 17, 18, 19, 27, 28, 51, 58 e 60.
- NASCIMENTO, D.; XAVIER, C.; FELIPE, I.; NETO, F. L. Dynamic conditional correlation garch: a multivariate time series novel using a Bayesian approach. *Journal of Modern Applied Statistical Methods*, v. 18, n. 1, p. 6, 2020. Citado na página 47.
- NETO, J. S. *Entenda o que são os contratos de opções de ações que fizeram investidores lucrarem com queda dos papéis da Petrobras*. 2021. Disponível em: <<https://oglobo.globo.com/economia/entenda-que-sao-os-contratos-de-opcoes-de-acoes-que-fizeram-investidores-lucrare-m-com-protect/@normalcr/relax-queda-dos-papeis-da-petrobras-24906475>>. Acesso em: 10 março 2021. Citado na página 16.
- OLIVEIRA, S. P. *Métodos numéricos para precificação de opções*. Dissertação (Mestrado) — Universidade Estadual de Campinas, 1998. Citado na página 16.
- ORSKAUG, E. *Multivariate DCC-GARCH Model: With Various Error Distributions*. Dissertação (Mestrado) — Norwegian University of Science and Technology, 2009. Citado 3 vezes nas páginas 45, 47 e 48.

- PANNA, M. Note on simple and logarithmic return. *APSTRACT: applied studies in agribusiness and commerce*, v. 11, n. 1033-2017-2935, p. 127–136, 2017. Citado na página 80.
- PAOLELLA, M. S.; HAAS, M.; BAUWENS, L.; HAFNER, C. M.; LAURENT, S. Mixture and regime-switching GARCH models. *Wiley Handbooks in Financial Engineering and Econometrics*, n. 3, p. 71–102, 2012. Citado 9 vezes nas páginas 12, 38, 39, 41, 42, 57, 59, 61 e 77.
- PARK, B.-J. An outlier robust GARCH model and forecasting volatility of exchange rate returns. *Journal of Forecasting*, v. 21, n. 5, p. 381–393, 2002. Citado na página 27.
- PELAGATTI, M.; LISI, F. Variance initialisation in GARCH estimation. In: *S. Co. 2009. Sixth Conference. Complex Data Modeling and Computationally Intensive Statistical Methods for Estimation and Prediction*. [S.l.: s.n.], 2009. p. 323. Citado na página 26.
- PELLETIER, D. Regime switching for dynamic correlations. *Journal of Econometrics*, v. 131, n. 1-2, p. 445–473, 2006. Citado 2 vezes nas páginas 52 e 53.
- PÉREZ-CRUZ, F.; AFONSO-RODRIGUEZ, J. A.; GINER, J. et al. Estimating GARCH models using support vector machines\*. *Quantitative Finance*, v. 3, n. 3, p. 163–172, 2003. Citado na página 24.
- PETRUS, P. Robust Huber adaptive filter. *IEEE Transactions on Signal Processing*, v. 47, n. 4, p. 1129–1133, 1999. Citado 2 vezes nas páginas 19 e 40.
- PRETIS, F.; READE, J. J.; SUCARRAT, G. Automated general-to-specific (GETS) regression modeling and indicator saturation for outliers and structural breaks. *Journal of Statistical Software*, v. 86, n. 3, p. 1–44, 2018. Citado na página 99.
- PSARADAKIS, Z.; SOLA, M. Markov-switching models with state-dependent time-varying transition probabilities. *Econometrics and Statistics*, 2021. Citado na página 99.
- SAKATA, S.; WHITE, H. High breakdown point conditional dispersion estimation with application to s&p 500 daily returns volatility. *Econometrica*, p. 529–567, 1998. Citado 2 vezes nas páginas 17 e 27.
- SANTOS, C.; HENDRY, D. F.; JOHANSEN, S. Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, v. 23, n. 2, p. 317–335, 2008. Citado na página 99.
- TRUCÍOS, C.; HOTTA, L. K.; RUIZ, E. Robust bootstrap forecast densities for GARCH returns and volatilities. *Journal of Statistical Computation and Simulation*, v. 87, n. 16, p. 3152–3174, 2017. Citado na página 18.
- TSAY, R. S. *Analysis of financial time series*. Terceira ed. Hoboken, NJ: John Wiley & Sons, 2010. Citado na página 21.
- \_\_\_\_\_. *Multivariate time series analysis: with R and financial applications*. Primeira ed. Hoboken, NJ: John Wiley & Sons, 2014. Citado 2 vezes nas páginas 44 e 47.

- WANG, J.; MA, F.; LIANG, C.; CHEN, Z. Volatility forecasting revisited using markov-switching with time-varying probability transition. *International Journal of Finance & Economics*, v. 27, n. 1, p. 1387–1400, 2022. Citado na página 99.
- WANG, J.; ZHOU, M.; JIN, X.; GUO, X.; QI, L.; WANG, X. Variance minimization hedging analysis based on a time-varying markovian dcc-garch model. *IEEE Transactions on Automation Science and Engineering*, v. 17, n. 2, p. 621–632, 2020. Citado na página 99.
- WEIDE, R. Van der. GO-GARCH: a multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics*, v. 17, n. 5, p. 549–564, 2002. Citado na página 45.
- WEINBACH, G. C. Regime switching with time-varying transition probabilities. *Business Cycles: Durations, Dynamics, and Forecasting*, p. 144, 1999. Citado na página 99.
- ZOU, Y.; CHAN, S.-C.; NG, T.-S. Least mean m-estimate algorithms for robust adaptive filtering in impulse noise. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, v. 47, n. 12, p. 1564–1569, 2000. Citado na página 40.