Modelo GARCH com Mudança de Regime Markoviano para séries financeiras

William Gonzalo Rojas Durán

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título
de
Mestre em Ciências

Programa: Estatística
Orientador: Prof. Dr. Airlane Pereira Alencar

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CNPq

São Paulo, Março de 2014

Modelo GARCH com Mudança de Regime Markoviano para séries financeiras

Esta versão final da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 24/03/2014.

Comissão Julgadora:

- Prof^{a}. Dr^{a}. Airlane Pereira Alencar (orientadora) - IME-USP
- Prof. Dr. Pedro Alberto Morettin - IME-USP
- Prof^{a}. Dr^{a}. Thelma Sáfadi - UFLA

Agradecimentos

Primeiramente, agradeço a minha familia que de alguma ou outra forma me ajudaram neste caminho.

Agradeço a LARC por seu apoio incondicional em todo momento.

Agradeço a Professora Airlane pela ajuda constante neste trabalho.

Aos Professores Pedro Morettin e Thelma Sáfadi pelas correções feitas no trabalho.

A meus amigos do IME, Elisângela, Josemir, Agatha, Ana Paula, Andressa, Alejandra, Fábio Oki.

Aos professores do IME.

Resumo

ROJAS, D. W. G. **Modelo GARCH com Mudança de Regime Markoviano para séries financeiras**. 2013. 120 f. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2013.

Neste trabalho analisaremos a utilização dos modelos de mudança de regime markoviano para a variância condicional. Estes modelos podem estimar de maneira fácil e inteligente a variância condicional não observada em função da variância anterior e do regime. Isso porque, é razoável ter coeficientes variando no tempo dependendo do regime correspondentes à persistência da variância (variância anterior) e às inovações. A noção de que uma série econômica possa ter alguma variação na sua estrutura é antiga para os economistas. Marcucci (2005) comparou diferentes modelos com e sem mudança de regime em termos de sua capacidade para descrever e predizer a volatilidade do mercado de valores dos EUA. O trabalho de Hamilton (1989) foi uns dos mais importantes para o desenvolvimento de modelos com mudança de regime. Inicialmente mostrou que a série do PIB dos EUA pode ser modelada como um processo que tem duas formas diferentes, uma na qual a economia encontra-se em crescimento e a outra durante a recessão. O câmbio de uma fase para outra da economia pode seguir uma cadeia de Markov de primeira ordem.

Utilizamos as séries de índice Bovespa e S&P500 entre janeiro de 2003 e abril de 2012 e ajustamos o modelo GARCH(1,1) com mudança de regime seguindo uma cadeia de Markov de primeira ordem, considerando dois regimes. Foram consideradas as distribuições gaussiana, t de Student e generalizada do erro (GED) para modelar as inovações.

A distribuição t de Student com mesmo grau de liberdade para ambos os regimes e graus distintos se mostrou superior à distribuição normal para caracterizar a distribuição dos retornos em relação ao modelo GARCH com mudança de regime. Além disso, verificou-se um ganho no percentual de cobertura dos intervalos de confiança para a distribuição normal, bem como para a distribuição t de Student com mesmo grau de liberdade para ambos os regimes e graus distintos, em relação ao modelo GARCH com mudança de regime quando comparado ao modelo GARCH usual.

Palavras-chave: Mudança de Regime Markoviano, Volatilidade, modelos GARCH.

Abstract

ROJAS, D. W. G. **Markov Regime Switching GARCH Model for Financial Series**. 2010. 120 f. Thesis (Master) - Mathematics and Statistics Institute, University of São Paulo, São Paulo, 2013.

In this work we analyze heterocedastic financial data using Markov regime switching models for conditional variance. These models can estimate easily the unobserved conditional variance as function of the previous variance and the regime. It is reasonable to have time-varying coefficients corresponding to the persistence of variance (previous variance) and innovations. The economic series notion may have some variation in their structure is usual for economists. Marcucci (2005) compared different models with and without regime switching in terms of their ability to describe and predict the volatility of the U.S. market. The Hamilton's (1989) work was the most important one in the regime switching models development. Initially showed that the series of U.S. GDP can be modeled as a process that has two different forms one in which the economy is growing and the other during the recession. The change from one phase to another economy can follow a Markov first order chain.

We use the Bovespa series index and S&P500 between January 2003 and April 2012 and fitted the GARCH (1,1) models with regime switching following a Markov first order chain, considering two regimes. We considered Gaussian distribution, Student-t and generalized error (GED) to model innovations.

The t-Student distribution with the same freedom degree for both regimes and distinct degrees showed higher than normal distribution for characterizing the distribution of returns relative to the GARCH model with regime switching. In addition, there was a gain in the percentage of coverage of the confidence intervals for the normal distribution, as well as the t-Student distribution with the same freedom degree for both regimes and distinct degrees related to GARCH model with regime switching when compared to the usual GARCH model.

Keywords: Markov Regime Switching, Volatility, GARCH models.

Sumário

Lista de Figuras **ix**

Lista de Tabelas **xiii**

1 Introdução **1**

2 Conceitos **3**

2.1 Séries Temporais Estacionárias e Ergódicas . . . 3

2.1.1 Processos Lineares . . . 4

2.2 Séries Temporais Não Estacionárias . . . 5

2.2.1 Modelos da família ARIMA . . . 5

2.3 Martingal e Sequência de Diferenças de Martingal . . . 8

2.4 Variância de Longo Prazo . . . 8

2.5 Retornos . . . 9

3 Modelos GARCH **11**

3.1 Introdução . . . 11

3.2 Modelos ARCH . . . 11

3.3 Testes para efeitos com Heteroscedasticidade Condicional . . . 13

3.4 Modelo GARCH . . . 14

3.5 Modelo GARCH e Fatos Estilizados . . . 15

3.6 Estimação de Modelos GARCH . . . 15

3.6.1 Estimador de Máxima Verossimilhança Condicional . . . 16

3.6.2 Estimadores de Mínimos Desvios Absolutos . . . 17

3.7 Diagnóstico de Modelos GARCH . . . 18

3.8 Extensões dos Modelos GARCH . . . 20

3.9 Distribuição de Erros Não-Gaussianos . . . 21

3.10 Seleção e Comparação de Modelos GARCH . . . 22

3.11 Predição de Modelos GARCH . . . 24

4 Modelo GARCH com Mudança de Regime **25**

4.1 Modelo GARCH com mudança de regime Markoviano . . . 25

4.2 MRS-GARCH . . . 26

4.3 Estimação . . . 27

4.4 Probabilidade Suavizada . . . 28

5 Aplicação em Séries Reais **31**
5.1 Aplicações 31
5.2 IBOVESPA 31
5.3 Análise de Resíduos-IBOVESPA 39
5.3.1 Análise de Resíduos para o modelo MRS-GARCH-N 40
5.3.2 Análise de Resíduos para o modelo MRS-GARCH-t2 42
5.3.3 Análise de Resíduos para o modelo MRS-GARCH-t 44
5.3.4 Análise de Resíduos para o modelo MRS-GARCH-GED 46
5.4 S&P500 51
5.5 Análise de Resíduos-S&P500 58
5.5.1 Análise de Resíduos para o modelo MRS-GARCH-N 58
5.5.2 Análise de Resíduos para o modelo MRS-GARCH-t2 60
5.5.3 Análise de Resíduos para o modelo MRS-GARCH-GED 62

6 Conclusões **67**

A Mistura de Distribuições **69**

Referências Bibliográficas **73**

Lista de Figuras

3.1 Gráfico da Série do índice Bovespa e os Log-retornos. 12
3.2 Gráfico da função de autocorrelação e autocorrelação parcial da Série dos log-retornos do índice Bovespa 12
3.3 Gráfico de quantis dos Resíduos padronizados do modelo GARCH(1,1) da série de retornos do Ibovespa 20
3.4 Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados - Modelo GARCH(1,1) da série de retornos do Ibovespa 20
3.5 Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados da Distribuição Normal e t-Student 23
3.6 Gráfico de quantis dos Resíduos Padronizados da Distribuição Normal e t-Student . 23

5.1 Gráfico da Série do índice Bovespa e os Log-retornos. 31
5.2 Gráfico de quantis e Histograma dos log-retornos da Série do índice Bovespa. 32
5.3 Função de autocorrelação dos log-retornos e log-retornos ao quadrado-IBOVESPA. . 33
5.4 Probabilidade suavizada do Regime 1 - Ibovespa - Normal. 35
5.5 Estimativa do desvio padrão condicional para os retornos do Ibovespa no Regime 1 e 2 com erro gaussiano. 36
5.6 Probabilidade suavizada do Regime 1 - Ibovespa - t. 36
5.7 Estimativa do desvio padrão condicional para os retornos do Ibovespa no Regime 1 e 2 com erro t de Student. 37
5.8 Probabilidade suavizada do Regime 1 - Ibovespa - t2. 37
5.9 Estimativa do desvio padrão condicional para os retornos do Ibovespa no Regime 1 e 2 com erro t de Student-t2. 38
5.10 Probabilidade suavizada do Regime 1 - Ibovespa - GED. 38
5.11 Estimativa do desvio padrão condicional para os retornos do Ibovespa no Regime 1 e 2 com erro GED. 39
5.12 Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo Normal-Ibovespa. 40
5.13 Resíduos padronizados nos Regimes 1 e 2 do modelo Normal Ibovespa. 40
5.14 Gráfico de quantis dos Resíduos padronizados do modelo Normal-Ibovespa. 41
5.15 Histograma dos resíduos padronizados do modelo Normal-Ibovespa. 41
5.16 Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo t2-Ibovespa. 42
5.17 Resíduos padronizados nos Regimes 1 e 2 do modelo t2-Ibovespa. 42

5.18 Gráfico de quantis dos Resíduos padronizados do modelo t2-Ibovespa. 43
5.19 Histograma dos Resíduos padronizados do modelo t2-Ibovespa. 43
5.20 Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo t-Ibovespa. 44
5.21 Resíduos padronizados nos Regimes 1 e 2 do modelo t-Ibovespa. 44
5.22 Gráfico de quantis dos Resíduos padronizados do modelo t-Ibovespa. 45
5.23 Histograma dos Resíduos padronizados do modelo t-Ibovespa. 45
5.24 Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo GED-Ibovespa. 46
5.25 Resíduos padronizados nos Regimes 1 e 2 do modelo GED-Ibovespa. 46
5.26 Gráfico de quantis dos Resíduos padronizados do modelo GED-Ibovespa. 47
5.27 Histograma dos Resíduos padronizados do modelo GED-Ibovespa. 47
5.28 Retornos e Intervalos de confiança (95%)-Normal-Ibovespa. 48
5.29 Retornos e Intervalos de confiança (95%)-t-Ibovespa. 48
5.30 Retornos e Intervalos de confiança (95%)-t2-Ibovespa. 49
5.31 Retornos e Intervalos de confiança (95%)-GED-Ibovespa. 49
5.32 Gráfico da Série S&P500 e os Log-retornos. 51
5.33 Gráfico de quantis e Histograma dos log-retornos da Série S&P500. 51
5.34 Função de autocorrelação dos log-retornos e log-retornos ao quadrado-S&P500. . . . 52
5.35 Probabilidade suavizada do Regime 1 - Normal - S&P500. 55
5.36 Estimativa do desvio padrão condicional para os retornos da S&P500 no Regime 1 e 2 com erro gaussiano. 55
5.37 Probabilidade suavizada do Regime 1- t2 - S&P500. 56
5.38 Estimativa do desvio padrão condicional para os retornos da S&P500 no Regime 1 e 2 com erro t de Student. 56
5.39 Probabilidade suavizada do Regime 1 - GED - S&P500. 57
5.40 Estimativa do desvio padrão condicional para os retornos da S&P500 no Regime 1 e 2 com erro GED. 57
5.41 Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo Normal-S&P500. 58
5.42 Resíduos padronizados nos Regimes 1 e 2 do modelo Normal-S&P500. 59
5.43 Gráfico de quantis dos Resíduos padronizados do modelo Normal-S&P500. 59
5.44 Histograma dos Resíduos padronizados do modelo Normal-S&P500. 59
5.45 Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo t2-S&P500. 60
5.46 Resíduos padronizados nos Regimes 1 e 2 do modelo t2-S&P500. 60
5.47 Gráfico de quantis dos Resíduos padronizados do modelo t2-S&P500. 61
5.48 Histograma dos Resíduos padronizados do modelo t2-S&P500. 61
5.49 Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo GED-S&P500. 62
5.50 Resíduos padronizados nos Regimes 1 e 2 do modelo GED-S&P500. 62
5.51 Gráfico de quantis dos Resíduos padronizados do modelo GED-S&P500. 63
5.52 Histograma dos Resíduos padronizados do modelo GED-S&P500. 63

5.53 Retornos e Intervalos de confiança (95%)-Normal-S&P500. 64
5.54 Retornos e Intervalos de confiança (95%)-t2-S&P500. 64
5.55 Retornos e Intervalos de confiança (95%)-GED-S&P500. 65

Lista de Tabelas

3.1 Coeficientes Estimados do modelo GARCH(1,1) 19
3.2 Testes para diagnóstico do modelo GARCH(1,1) ajustado 19
3.3 AIC, BIC e log da verossimilhança dos modelos GARCH(1,1) ajustados para a série de retornos Ibovespa 23

5.1 Medidas descritivas do retorno-Ibovespa 32
5.2 Estimativas de Máxima Verossimilhança erros padrões do Modelo MRS - GARCH - Ibovespa 34
5.3 Porcentagem de Cobertura do intervalos de confiança para os log-retornos do Ibovespa 50
5.4 Medidas descritivas do retorno-S&P500 52
5.5 Estimativas de Máxima Verossimilhança e erros padrões do Modelo MRS - GARCH - S&P500 54
5.6 Porcentagem de Cobertura do intervalos de confiança para os log-retornos da S&P500 65

Capítulo 1

Introdução

Neste trabalho analisaremos a utilização dos modelos de mudança de regime markoviano para a variância condicional. Estes modelos podem estimar de maneira fácil e inteligente a variância condicional não observada em função da variância anterior e do regime. No entanto, é razoável ter coeficientes variando no tempo correspondentes à persistência da variância (variância anterior) e às inovações.

Séries temporais apresentam mudanças de comportamento ao longo do tempo que se caracterizam por diversos regimes. Um exemplo é quando o mercado se encontra em dois possíveis regimes: recessão e expansão do mercado. Esses regimes não são variáveis medidas diretamente e, portanto, constituem variáveis aleatórias. O trabalho de Hamilton (1989) foi uns dos mais importantes para o desenvolvimento de modelos com mudança de regime. Quandt (1972) apresenta um modelo de regressão com mudanças independentes de regime e Goldfeld e Quandt (1973), amplia a análise para modelos com mudanças de regimes markovianas.

A principal motivação deste trabalho é propor um modelo de volatilidade com mudança de regime markoviana, regime este associado aos parâmetros do modelo.

Adotamos o modelo MRS-GARCH introduzido por Klaassen (2002) usando as distribuições normal, t-Student e a distribuição generalizada do erro (GED). Além disso, o parâmetro ν (graus de libertade) se deixa variar entre diferentes regimes, para capturar a curtose. De uma forma geral, Dueker (1997) mostrou que no modelo com mudança de regime apenas alguns parâmetros são dependentes do estado. O maior interesse deste é analisar o uso destes modelos para uma aplicação em finanças, ou seja, utilizá-los na construção de regras de mercado.

No Capítulo 2 apresentamos alguns conceitos básicos utilizados no decorrer do todo o texto.

O Capítulo 3 aborda os modelos GARCH e suas extenções. Apresentamos testes para efeitos ARCH, métodos de estimação dos parâmetros dos modelos e algumas ferramentas de diagnóstico dos modelos ajustados. Os modelos GARCH usual e com mudança de regime são descritos no Capítulo 4, com ênfase no estudo do modelo utilizado. Encontram-se neste capítulo as ferramentas de diagnóstico dos ajustes dos modelos. Aplicações a dados reais podem ser encontradas no Capítulo 5. As séries utilizadas para as aplicações são séries de cotações de preços de ações de duas grandes bolsas mundiais (Ibovespa e S&P500) desde 2003 até 2012.

Finalmente, apresentamos algumas conclusões a respeito do trabalho realizado.

Capítulo 2

Conceitos

Neste capítulo serão apresentados alguns conceitos e definições que servirão de base para o desenvolvimento dos outros capítulos. Mais detalhes podem ser encontrados em Toloi (2006), Fuller (1996) e Hamilton e Susmel (1994).

2.1 Séries Temporais Estacionárias e Ergódicas

Um processo estocástico $y = \{y(t,w), t \in T, w \in \Omega\}$ é uma família de variáveis aleatórias ($v.a.$) definidas num espaço de probabilidade (Ω, Λ, P) e indexadas pelos elementos de um conjunto de parâmetros T, $T \subset [0, \infty)$.
Na realidade, um processo estocástico é uma função de dois argumentos, $y(t,w), t \in T, w \in \Omega$. Para cada $t \in T$, a função $y(t, \cdot)$ é mensurável relativamente a Λ. Por outro lado, para cada $w \in \Omega$ fixado, obtemos uma função $y(\cdot, w) : T \to \mathbb{R}$ que é chamada trajetória, realização, função amostral do processo ou série temporal.
Para facilitar a notação, representaremos $y(t,w)$ por y_t.
O processo estocástico y é estacionário de segunda ordem, fracamente estacionário, ou ainda, estacionário em sentido amplo se

- $E(y_t) = \mu$, constante, para todo $t \in T$;
- $\gamma(y_t, y_{t-j}) = E[(y_t - \mu)(y_{t-j} - \mu)] = \gamma_j$, só depende da defasagem j, para todo t e qualquer j;
- $E\left(y_t^2\right) < \infty$, para qualquer t.

Seja y_t um processo estacionário e γ_j a autocovariância de y_t de j-ésima ordem ou de lag j, calculada para todo j. A função de autocovariância γ_j satisfaz as seguintes propriedades:

- $\gamma_0 = var(y_t) > 0$;
- $\gamma_0 \geq |\gamma_j|$;
- $\gamma_{-j} = \gamma_j$;
- γ_j é não negativa definida, no sentido que:

$$\sum_{l=1}^{n} \sum_{k=1}^{n} a_l a_k \gamma_{j_l - j_k} \geq 0,$$

para quaisquer números reais $a_1, ..., a_n$ e $j_1, ..., j_n$.
A função de autocorrelação (FAC) de y_t é definida por

$$\rho_j = \frac{cov(y_t, y_{t-j})}{\sqrt{var(y_t) var(y_{t-j})}} = \frac{\gamma_j}{\gamma_0}, j = 0, \pm 1, \cdots.$$

Intuitivamente, um processo estacionário é definido por sua média, variância e função de autocorrelação.
Qualquer função de um processo estacionário também é um processo estacionário, isto é, se y_t é estacionária, então $z_t = g(y_t)$ também é estacionária, para qualquer função $g(\cdot)$.
A autocovariância amostral de lag j e a autocorrelação amostral de lag j são definidas para processos estacionários por

$$\widehat{\gamma}_j = \frac{1}{T}\sum_{t=j+1}^{T}(y_t - \overline{y})(y_{t-j} - \overline{y}) \tag{2.1}$$

$$\widehat{\rho}_j = \frac{\widehat{\gamma}_j}{\widehat{\gamma}_0}, \tag{2.2}$$

onde $\overline{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$ é a média amostral.
Dizemos que um processo estocástico y_t é ergódico se os momentos amostrais convergirem em probabilidade para os momentos populacionais, ou seja, $\overline{y} \xrightarrow{p} \mu, \widehat{\gamma}_j \xrightarrow{p} \gamma_j, \widehat{\rho}_j \xrightarrow{p} \rho_j$.

2.1.1 Processos Lineares

Em situações em que se pretende utilizar modelos para descrever séries temporais, é necessário introduzir suposições simplificadoras, que nos conduza a analisar determinadas classes de processos estocásticos. Assim, podemos ter:

- processos estacionários ou não-estacionários, de acordo com a independência ou não relativamente à origem dos tempos;
- processos normais (Gaussianos) ou não-normais, de acordo com as funções de densidade de probabilidade que caracterizam os processos;
- processos Markovianos ou não-Markovianos, de acordo com a independência dos valores do processo, em dado instante, de seus valores em instantes precedentes.

Mais detalhes encontram-se em Toloi (2006).
Seja y_t um processo estocástico. O processo y_t pode ser representado por um processo linear, ou processo de médias móveis de ordem infinita, da seguinte forma

$$\begin{aligned} y_t &= \mu + \sum_{k=0}^{\infty}\psi_k\epsilon_{t-k}, \\ \psi_0 &= 1, \sum_{k=0}^{\infty}\psi_k^2 < \infty, \\ \epsilon_t &\sim RB\left(0, \sigma^2\right), \end{aligned} \tag{2.3}$$

onde ϵ_t é um ruído branco. Dizemos que $\{\epsilon_t, t \in T\}$ é um ruído branco se as variáveis aleatórias ϵ_t são não correlacionadas, isto é, $Cov\{\epsilon_t, \epsilon_s\} = 0, t \neq s$. Para detalhes ver **?**.
Para o processo linear (2.3), pode-se mostrar que

$$\begin{aligned} E[y_t] &= \mu, \\ \gamma_0 &= var(y_t) = \sigma^2\sum_{k=0}^{\infty}\psi_k^2, \\ \gamma_j &= cov(y_t, y_{t-1}) = \sigma^2\sum_{k=0}^{\infty}\psi_k\psi_{k+j}, \\ e & \\ \rho_j &= \frac{\sum_{k=0}^{\infty}\psi_k\psi_{k+j}}{\sum_{k=0}^{\infty}\psi_k^2}. \end{aligned}$$

Portanto, as autocorrelações de qualquer processo estacionário e ergódico y_t são determinadas pelos pesos $\{\psi_j\}$ da representação (2.3).

2.2 Séries Temporais Não Estacionárias

Uma vez que um processo estacionário possui momentos invariantes, um processo não estacionário deve ter alguma dependência nos momentos. As formas mais comuns de não estacionariedade são causadas pela dependência temporal na média e na variância. Iremos apresentar somente os processos integrados de ordem 1, mais usados em finanças.

Processo Integrados

Chamamos y_t de um processo integrado de ordem 1 ($y_t \sim I(1)$) se possui a forma

$$y_t = y_{t-1} + u_t, \tag{2.4}$$

onde u_t é um processo estacionário. Desta forma, tem-se que a primeira diferença de y_t é $y_t = u_t$ e é estacionário. Devido a esta propriedade, o processo $I(1)$ é chamado também de processo de diferença estacionária. Uma vez que o processo estacionário u_t não precisa sofrer diferenças, é chamado de processo integrado de ordem zero ($u_t \sim I(0)$).

O processo y_t é $I(d)$ se $\Delta^d y_t \sim I(0)$, ou seja, após d diferenças torna-se $I(0)$. Em séries temporais financeiras raramente são modelados processos $I(d)$ com $d > 2$.

2.2.1 Modelos da família ARIMA

Modelos $AR(p)$
O modelo autoregressivo de ordem p ($AR(p)$) é dado por

$$y_t - \mu = \phi_1 (y_{t-1} - \mu) + \cdots + \phi_p (y_{t-p} - \mu) + \epsilon_t,$$

ou, utilizando-se a notação de operadores,

$$\phi(Z)(y_t - \mu) = \epsilon_t,$$

onde Z é o operador de translação para o passado, definido como $Z^i y_t = y_{t-i}$, e $\Phi(Z) = 1 - \phi_1 Z - \cdots - \phi_p Z^p$ e $\mu = E[y_t]$.
Pode-se mostrar que um processo $AR(p)$ é sempre invertível, e é estacionário e ergódico sob a condição de que as raízes da equação característica,

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0, \tag{2.5}$$

estejam fora do círculo unitário. Uma condição necessária para estacionariedade muito útil na prática é que $|\phi_1 + \cdots + \phi_p| < 1$.
A função de autocovariância de ordem j de um processo $AR(p)$ estacionário é dado por

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \cdots + \phi_p \gamma_{j-p},$$

com $\gamma_0 = var(y_t)$. A função de autocorrelação de ordem j do processo é dada por

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \cdots + \phi_p \rho_{j-p}.$$

Função de Autocorrelação Parcial

A função de autocorrelação parcial (FACP) é um importante instrumento para ajudar na identificação da ordem p de modelos $AR(p)$. A obtenção da FACP baseia-se na estimação da sequência de modelos $AR(1), AR(2), ..., AR(p)$:

$$\begin{aligned} z_t &= \phi_{11} z_{t-1} + \epsilon_{1t} \\ z_t &= \phi_{21} z_{t-1} + \phi_{22} z_{t-2} + \epsilon_{2t} \\ &\vdots \\ z_t &= \phi_{p1} z_{t-1} + \phi_{p2} z_{t-2} + \cdots + \phi_{pp} z_{t-p} + \epsilon_{pt}, \end{aligned}$$

em que $z_t = y_t - \mu$. Os últimos coeficientes em cada $AR(p)$ ($\phi_{jj}, j = 1, ..., p$) são chamados coeficientes de autocorrelação parcial.
Para um processo $AR(p)$, o p-ésimo coeficiente de autocorrelação parcial é diferente de zero, e o restante é igual a zero para $j > p$.

Modelos $MA(q)$

O modelo de médias móveis de ordem q ($MA(q)$) tem a forma

$$y_t - \mu = \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q},$$

onde $\epsilon_t \sim RB\left(0, \sigma^2\right)$.

Em notação de operadores temos

$$(y_t - \mu) = \theta(L)\, \epsilon_t,$$

onde $\theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$. O modelo $MA(q)$ é estacionário e ergódico desde que $\theta_1, \cdots, \theta_q$ sejam finitos, e é invertível se todas as raízes da equação característica,

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q = 0, \tag{2.6}$$

estiverem fora do círculo unitário. A propriedade de invertibilidade torna possível escrever o modelo como um $AR(\infty)$.

A função de autocovariância de um processo $MA(q)$ é dada por:

$$\begin{aligned} \gamma_0 &= \sigma^2 \left(1 + \theta_1^2 + \cdots + \theta_q^2\right); \\ \gamma_j &= \begin{cases} \sigma^2 \left(\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_{q-j}\theta_q\right), & \text{para} \quad j = 1, 2, ..., q \\ 0, & \text{c.c} \end{cases} \end{aligned}$$

Observa-se que a FAC de um processo $MA(q)$ é diferente de zero para o lag q e igual a zero para $j > q$. Dessa forma, a FAC pode ser usada como um importante instrumento na identificação deste tipo de processo.

A função de autocorrelação do processo pode ser escrita como

$$\rho_j = \frac{\left(\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_{q-j}\theta_q\right)}{1 + \theta_1^2 + \cdots + \theta_q^2}.$$

A função de autocorrelação parcial de um processo $MA(q)$ invertível apresenta decaimento exponencial.

Modelos $ARMA\,(p,q)$

O modelo autoregressivo e de médias móveis ($ARMA\,(p,q)$) tem a forma

$$y_t - \mu = \phi_1\,(y_t - \mu) + \cdots + \phi_p\,(y_t - \mu) + \epsilon_t + \theta_1\epsilon_{t-1} + \cdots + \theta_q\epsilon_{t-q}, \tag{2.7}$$
$$\tag{2.8}$$

em que $\epsilon_t \sim RB\left(0,\sigma^2\right)$. Pode ser escrito na forma da regressão com intercepto c

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \epsilon_t + \theta_1\epsilon_{t-1} + \cdots + \theta_q\epsilon_{t-q},$$

ou representado utilizando-se o operador L

$$\phi\,(L)\,(y_t - \mu) = \theta\,(L)\,\epsilon_t,$$

e é estacionário e ergódico se as raízes da equação característica $\phi\,(z) = 0$ estiverem fora do círculo unitário, e é invertível se as raízes da equação característica $\theta\,(z) = 0$ estiverem fora do círculo unitário. Assume-se a condição de que os polinômios $\phi\,(z)$ e $\theta\,(z)$ não tenham fatores em comum, pois nesse caso teríamos redundância de parâmetros e a ordem do modelo poderia ser reduzida.

A média de um processo estacionário e ergódico ($ARMA\,(p,q)$) é igual a

$$\mu = \frac{c}{1 - \phi_1 - \cdots - \phi_p},$$

e suas autocovariâncias e autocorrelações satisfazem as seguintes relações recursivas

$$\begin{aligned} \gamma_j &= \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} + \cdots + \phi_p\gamma_{j-p}, j > q, \\ \rho_j &= \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \cdots + \phi_p\rho_{j-p}, j > q. \end{aligned} \tag{2.9}$$

A forma geral da FAC de um processo ($ARMA\,(p,q)$) é complicada. Para detalhes, ver Hamilton (1994). Em geral, para o processo ($ARMA\,(p,q)$), a FAC tem o mesmo comportamento da FAC de um processo $AR\,(p)$ para $p > q$, e a FACP tem comportamento parecido com a FACP de um processo $MA\,(q)$ para $q > p$. Dessa forma, ambas FAC e FACP podem apresentar decaimento exponencial.
Na prática, modelos ($ARMA\,(p,q)$) de ordens altas são difíceis de serem estimados e, por este motivo, são raramente usados para analisar dados financeiros. Segundo Zivot (2003), modelos ($ARMA\,(p,q)$) de ordens baixas, com p e q menores do que três, são geralmente suficientes na análise de séries temporais financeiras.

Modelos ($ARIMA\,(p,d,q)$)

A especificação do modelo ($ARMA\,(p,q)$) (2.6) assume que y_t é estacionário e ergódico. Se y_t possuir alguma tendência como o preço do barril de petróleo, então y_t deve sofrer alguma transformação com o objetivo de eliminar esta tendência.
O modelo ($ARIMA\,(p,d,q)$) autoregressivo integrado e de médias móveis foi proposto para a modelagem de processos nos quais pelo menos uma raiz da equação característica $\phi\,(z) = 0$ se encontra sobre o círculo unitário. Neste caso o processo é não estacionário na média.
Se houver uma raíz unitária em y_t, então a primeira diferença $\Delta y_t = y_t - y_{t-1}$ elimina tal tendência, onde $\Delta = 1 - L$ é o operador diferença. Se houver duas raízes unitárias em y_t então haverá tendência linear em Δy_t, mas a segunda diferença $\Delta^2 y_t = \left(1 - 2L + L^2\right) y_t = y_t - 2y_{t-1} + y_{t-2}$ não apresentará tendência. O modelo ($ARIMA\,(p,d,q)$) apresenta d diferenças para eliminar as tendências e depois é ajustado um modelo $ARMA(p,q)$.

O modelo $(ARIMA\,(p,d,q))$ pode ser representado por

$$\phi\,(L)\,\Delta^d\,(y_t-\mu)=\theta\,(L)\,\epsilon_t,$$

onde $\Delta^d=(1-L)^d$.

2.3 Martingal e Sequência de Diferenças de Martingal

Seja $\{y_t\}$ uma sequência de variáveis aleatórias e seja $I_t=\{y_t,y_{t-1},\cdots\}$ um conjunto de informações baseadas no passado histórico de $\{y_t\}$. A sequência $\{y_t,I_t\}$ é chamada de um martingal se

- $I_{t-1}\subset I_t$;
- $E\,[|y_t|]<\infty$;
- $E\,[y_t|I_{t-1}]=y_{t-1}$ (propriedade martingal)

Um exemplo comum de um martingal é o modelo de passeio aleatório

$$y_t=y_{t-1}+\epsilon_t,\epsilon_t\sim RB\left(0,\sigma^2\right), \tag{2.10}$$

onde y_0 é um valor inicial fixado. Se $I_t=\{y_t,...,y_0\}$ então $E\,[y_t|I_{t-1}]=y_{t-1}$, dado que $E\,[\epsilon_t|I_{t-1}]=0$.
Seja $\{\epsilon_t\}$ uma sequência de variáveis aleatórias com um conjunto de informações associadas a I_t. A sequência $\{\epsilon_t,I_t\}$ é chamada de uma sequência de diferenças de martigal (martingale difference sequence (MDS)) se

- $I_{t-1}\subset I_t$;
- $E\,[\epsilon_t|I_{t-1}]=0$; (propriedade MDS)

Se $\{y_t,I_t\}$ é um martingal, um MDS $\{\epsilon_t,I_t\}$ pode ser construído definindo-se

$$\epsilon_t=y_t-E\,[y_t|I_{t-1}]\,.$$

Por construção, um MDS é um processo não correlacionado.
Para $k>0$ temos

$$\begin{aligned}E\,[\epsilon_t\epsilon_{t-k}] &= E\,[E\,[\epsilon_t\epsilon_{t-k}|I_{t-1}]]\\ &= E\,[\epsilon_{t-k}E\,[\epsilon_t|I_{t-1}]]\\ &= 0.\end{aligned} \tag{2.11}$$

Embora um processo MDS seja não correlacionado, ele não é necessariamente independente, isto é, pode exitir dependência em momentos de maior ordem de ϵ_t.
Processos MDS possuem muitos resultados úteis de convergência (lei de los grandes números, teorema central do limite, etc.) White (1984), Hamilton e Susmel (1994) e Hayashi (2000) apresentam muitos destes resultados para a análise de séries temporais financeiras.

2.4 Variância de Longo Prazo

Seja y_t uma série temporal estacionária e ergódica. Pelo Teorema Central do Limite

$$\sqrt{T}\,(\overline{y}-\mu)\xrightarrow{d}N\left(0,\sum_{j=-\infty}^{\infty}\gamma_j\right),\quad T\to\infty$$

ou $\overline{y}$ se distribui assintoticamente com distribuição $N\left(\mu, \frac{1}{T}\sum_{j=-\infty}^{\infty}\gamma_j\right)$, onde T é o tamanho da amostra, $\overline{y}$ é a média amostral de y_t e γ_j são as autocovariâncias.
A variância de longo prazo (vlp) de y_t é dada pela multiplicação de T pela variância assintótica (avar) da média amostral:

$$vlp\left(y_t\right) = T \cdot avar\left(\overline{y}\right) = \sum_{j=-\infty}^{\infty}\gamma_j.$$

Dado que $\gamma_{-j} = \gamma_j$, a variância de longo prazo de y_t pode ser escrita como

$$vlp\left(y_t\right) = \gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j. \tag{2.12}$$

Estimando a Variância de Longo Prazo

Se y_t é um processo linear temos que

$$\sum_{j=-\infty}^{\infty}\gamma_j = \sigma^2\left(\sum_{j=0}^{\infty}\psi_j\right)^2 = \sigma^2\psi\left(1\right)^2,$$

e, portanto

$$vlp\left(y_t\right) = \sigma^2\psi\left(1\right)^2. \tag{2.13}$$

Se $y_t \sim ARMA\left(p,q\right)$ então

$$\phi\left(1\right) = \frac{1+\theta_1+\cdots+\theta_q}{1-\phi_1-\cdots-\phi_p} = \frac{\theta\left(1\right)}{\phi\left(1\right)}$$

e, portanto

$$vlp\left(y_t\right) = \frac{\sigma^2\theta\left(1\right)^2}{\phi\left(1\right)^2}. \tag{2.14}$$

Uma estimativa consistente para $vlp\left(y_t\right)$ pode ser calculada pela substituição de parâmetros estimados do modelo $\left(ARMA\left(p,q\right)\right)$ apropriado em (2.10). De forma alternativa, pode-se aproximar o modelo $\left(ARMA\left(p,q\right)\right)$ por um processo $\left(AR\left(p^*\right)\right)$

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_{p^*}y_{t-p^*} + \epsilon_t,$$

onde p^* é escolhido de tal forma que ϵ_t seja não correlacionado. Assim, a estimativa da variância de longo prazo autoregressiva fica dada por

$$vlp_{AR}\left(y_t\right) = \frac{\sigma^2}{\phi^*\left(1\right)^2}. \tag{2.15}$$

2.5 Retornos

Definiremos retornos e log-retornos e apresentaremos suas principais caraterísticas como apresentado em Toloi (2006). Para avaliar o risco de uma carteira de ativos no mercado financeiro é comum medir as variações de preços destes ativos. Na prática, é preferível trabalhar com retornos do que com preços pois os retornos são livres de escala e têm propriedades estatísticas interessantes

(como estacionariedade e ergodicidade). Os modelos apresentados neste trabalho têm como objetivo modelar a volatilidade dos retornos de séries financeiras, como retornos de preços de ações, por exemplo.
Seja P_t o preço de um ativo no instante t. A variação de preços entre os instantes $t-1$ e t é dada por $\Delta P_t = P_t - P_{t-1}$ e a variação relativa de preços ou retorno simples deste ativo é definido por

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}}. \tag{2.16}$$

De (2.14) temos que $R_t = \frac{P_t}{P_{t-1}} - 1$. Chamamos $1 + R_t = \frac{P_t}{P_{t-1}}$ de retorno bruto simples ou taxa de retorno.
Definimos o retorno composto continuamente, ou $log - retorno$ ou simplesmente $retorno$, como

$$r_t = \ln \frac{P_t}{P_{t-1}} = \ln(1 + R_t) = \ln(P_t) - \ln(P_{t-1}) = p_t - p_{t-1}, \tag{2.17}$$

onde $p_t = \ln(P_t)$. Nesse trabalho utilizaremos somente o retorno r_t do índice Ibovespa.

Fatos Estilizados sobre os Retornos

Séries financeiras apresentam algumas características que são comuns a outras séries temporais, como tendências, sazonalidade, pontos influentes (atípicos), heterocedasticidade condicional e não-linearidade. Os retornos financeiros, por outro lado, apresentam outras características peculiares que outras séries não apresentam. Retornos raramente apresentam tendências ou sazonalidades, com exceção eventualmente de retornos intra-diários, e séries de taxas de câmbio ou taxa de juros que podem apresentar tendências que variam com o tempo. Podemos resumir os principais fatos estilizados relativos a retornos financeiros:

1. retornos são geralmente não autocorrelacionados;
2. os quadrados dos retornos são autocorrelacionados;
3. séries de retornos apresentam agrupamentos de volatilidades ao longo do tempo;
4. a distribuição (não-condicional) dos retornos apresenta caudas mais pesadas do que uma distribuição normal;
5. algumas séries de retornos são não-lineares (respondem de maneira diferente a choques grandes ou pequenos, ou a choques negativos ou positivos).

Utilizaremos essas características para analisar as séries de retornos da Ibovespa. Os conceitos serão importantes ao introduzirmos os modelos de volatilidade no capítulo seguinte.

Capítulo 3

Modelos GARCH

3.1 Introdução

O comportamento variável do mercado financeiro é usualmente referido como "volatilidade". A volatilidade tem-se transformado num conceito muito importante em diferentes áreas da teoria e da prática financeira, como gerenciamento de risco, seleção de portifólio, derivação de preços, etc. Em termos estatísticos, a volatilidade é medida pela variância ou pelo desvio padrão condicional.

Nesse capítulo vamos introduzir a classe dos modelos GARCH *(Generalized Autoregressive Conditional Heteroskedasticity)* desenvolvidos por Engle (1989), Bollerslev (1986), Nelson (1991), e outros, que são capazes de modelar a volatilidade variando no tempo e de capturar muitos dos fatos estilizados do comportamento da volatilidade observada em séries temporais financeiras. Também será apresentado o ajuste do modelo GARCH para os retornos do Ibovespa e os métodos de avaliação desse ajuste.

3.2 Modelos ARCH

O Índice Bovespa é composto pelas ações mais negociadas na bolsa de São Paulo e os pesos das ações mudam ao longo do tempo, mas as que tem maior peso são Petrobras, Eletrobras, etc. Na Figura 3.1 temos os índices e os $log-retornos$ diários do índice Bovespa, no período de 2 de janeiro de 2003 a 9 de abril de 2012, para um total de 2294 observações. Pode-se perceber que os períodos de alta (crise em 2008) e de baixa volatilidade estão agrupados (fato estilizado). Essa tendência é típica de muitas séries temporais financeiras. Para confirmar este comportamento são apresentadas na Figura 3.2 as funções de autocorrelação dos retornos e dos retornos ao quadrado da série de log retornos do índice Bovespa.

Observando os gráficos das funções de autocorrelação percebemos que não há autocorrelações significantes para qualquer defasagem da série dos log-retornos. Por outro lado, a autocorrelação da série dos log-retornos ao quadrado se mostra significante pelo menos até a defasagem 30. Uma vez que os log-retornos ao quadrado medem o momento de segunda ordem da série original, este resultado indica que a variância dos log-retornos da série índice Bovespa condicionada ao passado histórico pode sofrer alterações ao longo do tempo, ou seja, a série de log-retornos pode mostrar a presença de heteroscedasticidade condicional.

A heteroscedasticidade condicional da série dos log-retornos ao quadrado pode ser modelada usando um processo autorregressivo simples (AR) para resíduos ao quadrado. Seja y_t uma série temporal estacionária tal como log-retornos financeiros e não exista autocorrelação significante em y_t então y_t pode ser expresso como sua média acrescida de um ruído branco:

$$y_t = c + \epsilon_t, \tag{3.1}$$

onde c é a média de y_t e ϵ_t são erros não correlacionados com média zero.

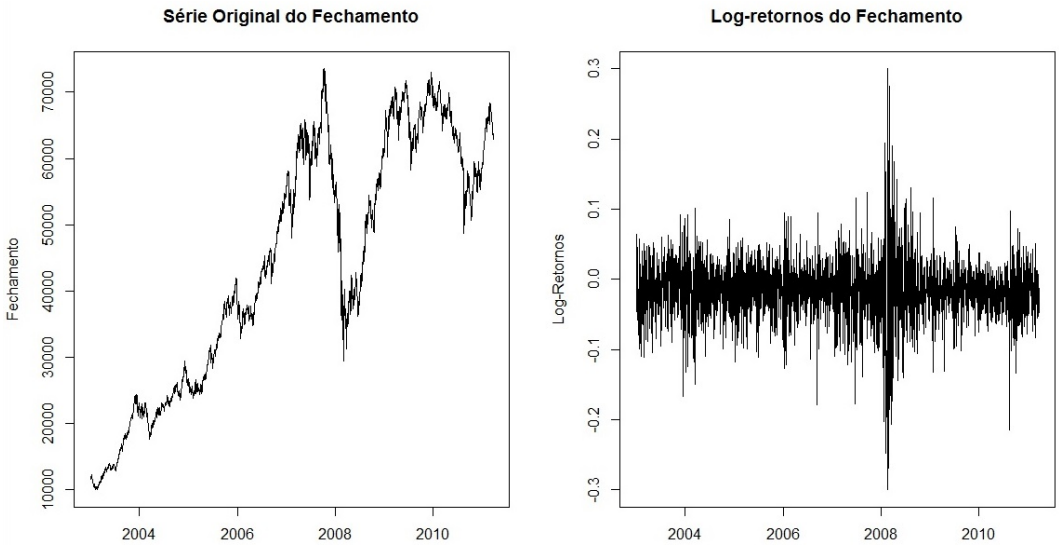

Figura 3.1: *Gráfico da Série do índice Bovespa e os Log-retornos.*

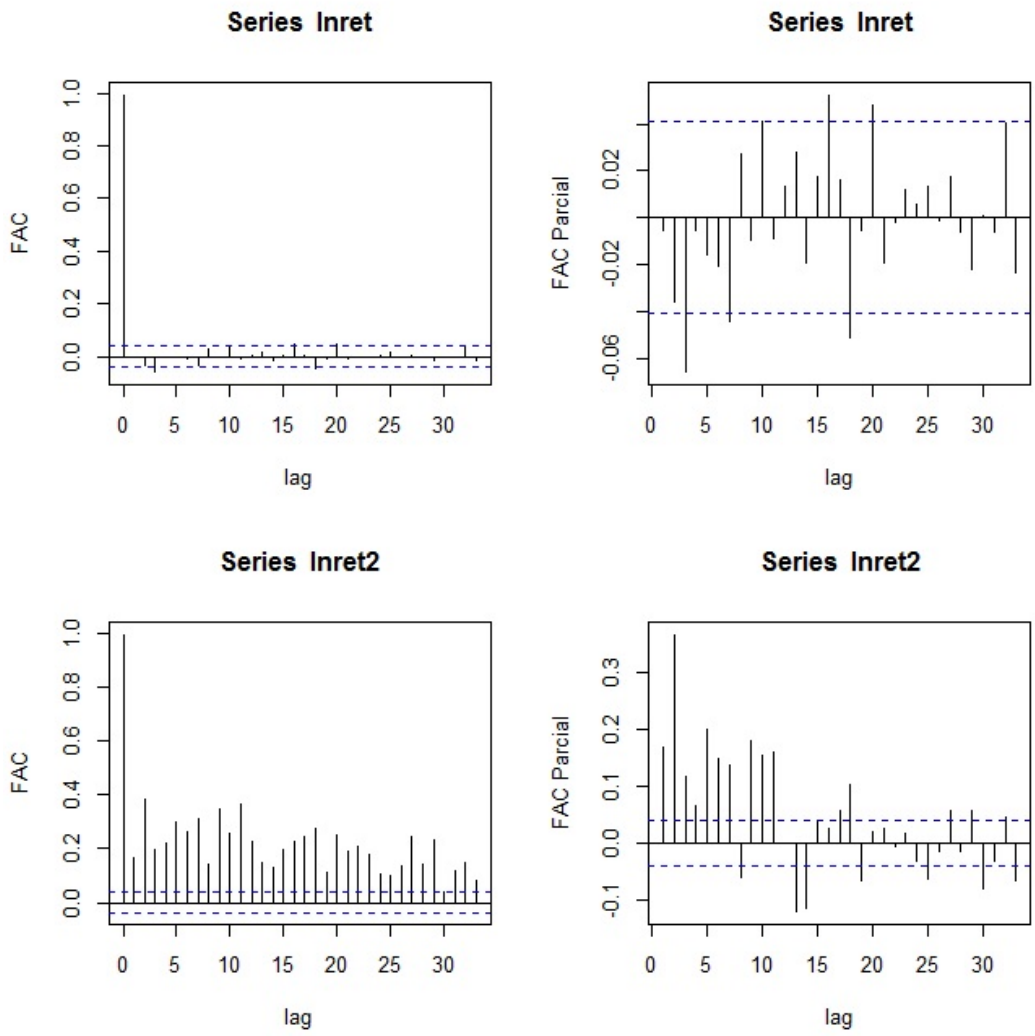

Figura 3.2: *Gráfico da função de autocorrelação e autocorrelação parcial da Série dos log-retornos do índice Bovespa*

Para incluir o efeito da hereroscedasticidade condicional, seja

$$Var_{t-1}(\epsilon_t) = \sigma_t^2,$$

em que $Var_{t-1}(\cdot)$ é a variância condicional à informação até o tempo $t-1$, e

$$\sigma_t^2 = a_0 + a_1\epsilon_{t-1}^2 + \cdots + a_p\epsilon_{t-p}^2, \tag{3.2}$$

em que p indica o lag mais distante do instante t, de modo que as informações em p ainda influenciam σ_t^2. Como ϵ_t tem média zero, $Var_{t-1}(\epsilon_t) = E_{t-1}\left(\epsilon_t^2\right) = \sigma_t^2$, a equação acima pode ser reescrita como:

$$\epsilon_t^2 = a_0 + a_1\epsilon_{t-1}^2 + \cdots + a_p\epsilon_{t-p}^2 + u_t, \tag{3.3}$$

em que $u_t = \epsilon_t^2 - E_{t-1}\left(\epsilon_t^2\right)$ é um processo diferença de martingal com média zero, mas com variância não constante. A equação (3.3) representa um processo $AR(p)$ para ϵ_t^2, e o modelo em (3.1) e (3.2) é conhecido como o modelo autorregressivo com heteroscedasticidade condicional de Engle (1989), que é usualmente chamado de modelo $ARCH(p)$.

Uma formulação equivalente para o modelo $ARCH(p)$ é

$$\begin{aligned} y_t &= c + \epsilon_t, \\ \epsilon_t &= z_t\sigma_t, \\ \sigma_t^2 &= a_0 + a_1\epsilon_{t-1}^2 + \cdots + a_p\epsilon_{t-p}^2, \end{aligned}$$

em que z_t é uma sequência de variáveis aleatórias independentes e identicamente distribuidas com média 0 e variância 1 com uma distribuição especificada. No modelo ARCH básico a distribuição de z_t é a normal padrão (iid). Entretanto, também é usual adotar distribuições com caudas mais pesadas como a t-Student e a distribuição do erro generalizada.

3.3 Testes para efeitos com Heteroscedasticidade Condicional

Antes de estimar um modelo ARCH para uma série temporal financeira, é interessante testar a presença de efeitos ARCH nos resíduos padronizados. Se não houver efeitos ARCH nos resíduos, então o modelo ARCH é desnecessário e mal especificado.
Uma vez que o modelo ARCH pode ser escrito como um modelo AR em termos de resíduos quadráticos como em (3.3), um teste simples assintótico de Multiplicador de Lagrange (LM) para modelos ARCH pode ser construído baseado na regressão auxiliar (3.3). Sob a hipótese nula de que não há efeitos ARCH: $a_1 = a_2 = \cdots = a_p = 0$, a estatística do teste terá distribuição assintótica qui quadrado:

$$LM = T \cdot R^2 \xrightarrow{a} \chi^2_{(p)},$$

onde T é o tamanho da amostra e R^2 é o quadrado do coeficiente de correlação múltipla da regressão auxiliar da equação (3.3) usando os resíduos estimados, ou seja, substituindo σ_t^2 por $\widehat{\epsilon}_t^2$ e os erros defasados ϵ_{t-j} por $\widehat{\epsilon}_{t-j}$
Outro teste muito utilizado para verificar a correlação entre os resíduos ao quadrado é o teste de $Box - Pierce - Ljung$ (Box e Pierce, 1970) para ϵ_t^2. Apesar desse teste não detectar quebras estruturais específicas no comportamento de um ruído branco, pode indicar se esses valores são muito altos. Uma modificação deste teste foi proposta por Ljung e Box (1978).
Esse teste tem $H_0 : \rho_1 = \cdots = \rho_K = 0$ onde ρ_k é a autocorrelação de lag k dos resíduos $\widehat{\epsilon}_t$. A

estatística do teste é

$$Q(K) = n(n+2)\sum_{k=1}^{K}\frac{\widehat{r}_k^2}{(n-k)},$$

terá sob H_0 uma distribuição χ^2 com $K-p-q$ graus de liberdade, em que, p e q indicam as ordens do modelo GARCH(p,q) e K refere-se às primeiras K correlações amostrais.

3.4 Modelo GARCH

Como foi mostrado anteriormente, uma maneira adequada para modelar a volatilidade é a utilização dos modelos ARCH. No entanto, tem-se que na prática geralmente é preciso um número grande de lags p para se obter um bom ajuste no modelo, o que implica um número grande de parâmetros. Um modelo mais parcimonioso foi proposto por Bollerslev (1986) e substitui o modelo AR em (3.2) pela formulação:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i\epsilon_{t-i}^2 + \sum_{j=1}^{q} b_j\sigma_{t-j}^2, \tag{3.4}$$

onde $a_0 > 0$ e os coeficientes $a_i\,(i = 1, \cdots, p)$ e $b_j\,(j = 1, \cdots, q)$ são todos não negativos para garantir que a variância condicional σ_t^2 seja sempre positiva. O modelo (3.4) junto com (3.1) é conhecido como o modelo GARCH(p, q) "generalized ARCH". Quando $q = 0$, o modelo GARCH se reduz ao modelo ARCH(p).
No modelo GARCH(p, q) a variância condicional de ϵ_t^2, σ_t^2, depende do quadrado dos erros nos p períodos anteriores e da variância condicional nos q períodos anteriores.

Representação ARMA do Modelo GARCH

Assim como um modelo ARCH pode ser escrito como um modelo AR em termos de erros quadráticos, um modelo GARCH poder ser escrito na forma de um modelo ARMA de erros quadráticos. Considere o modelo GARCH(1,1):

$$\sigma_t^2 = a_0 + a_1\epsilon_{t-1}^2 + b_1\sigma_{t-1}^2. \tag{3.5}$$

Sendo $E_{t-1}\left(\epsilon_t^2\right) = \sigma_t^2$, a equação acima pode ser escrita como:

$$\epsilon_t^2 = a_0 + (a_1 + b_1)\,\epsilon_{t-1}^2 + u_t - b_1 u_{t-1}, \tag{3.6}$$

que é um modelo ARMA(1,1) onde $u_t = \epsilon_t^2 - E_{t-1}\left(\epsilon_t^2\right)$ é ruído branco de média zero e variância não constante.
Dada a representação ARMA de um modelo GARCH, muitas propriedades do modelo GARCH seguem facilmente da correspondência de um processo ARMA para ϵ_t^2. Por exemplo, para o modelo GARCH$(1, 1)$ ser estacionário é preciso $a_1 + b_1 < 1$. Assumindo a estacionariedade do modelo GARCH$(1, 1)$, pode-se mostrar que a variância incondicional de ϵ_t é dada por $Var(\epsilon_t) = E\left(\epsilon_t^2\right) = \frac{a_0}{1-a_1-b_1}$, pois de (3.6) temos:

$$E\left(\epsilon_t^2\right) = a_0 + (a_1 + b_1)\,E\left(\epsilon_{t-1}^2\right)$$

e portanto

$$E\left(\epsilon_t^2\right) = a_0 + (a_1 + b_1)\,E\left(\epsilon_t^2\right),$$

baseado na suposição de que ϵ_t^2 é estacionário.

Para o modelo GARCH(p,q) em (3.4), os erros quadráticos ϵ_t^2 têm comportamento de um processo ARMA(max(p,q),q). A estacionariedade requer que $\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j < 1$ e a variância incondicional de ϵ_t é dada por

$$\overline{\sigma}^2 = Var\left(\epsilon_t\right) = \frac{a_0}{1 - \left(\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j\right)}. \tag{3.7}$$

3.5 Modelo GARCH e Fatos Estilizados

Na prática, existem alguns "fatos estilizados"sobre a volatilidade de séries temporais financeiras. Bollerslev, Engle e Nelson (1994) apresentam uma visão completa sobre esses fatos. Usando a representação ARMA dos modelos GARCH é possível mostrar que o modelo GARCH é capaz de explicar muitos desses fatos estilizados. Nessa seção serão detalhados dois importantes fatos estilizados: agrupamento de volatilidades *(volatility clustering)* e caudas pesadas.

No modelo GARCH(1,1) dado em (3.5) a estimativa do coeficiente b_1 é normalmente um número próximo de $0,9$ para séries temporais financeiras diárias ou semanais. Dessa maneira, temos que valores grandes de σ_{t-1}^2 são seguidos por valores grandes de σ_t^2, e valores pequenos de σ_{t-1}^2, por sua vez, são seguidos por valores pequenos de σ_t^2. Seguindo o mesmo raciocínio para a representação ARMA em (3.6), temos que grandes/pequenas mudanças em ϵ_{t-1}^2 são seguidas por grandes/pequenas mudanças em ϵ_t^2.

É fato conhecido que a distribuição de muitas séries temporais financeiras geralmente possuem caudas mais pesadas que a distribuição normal, ou seja, grandes mudanças ocorrem com mais frequência do que em uma distribuição normal. Bollerslev (1986) descreve as condições para a existência do momento de quarta ordem de um processo GARCH(1,1). Assumindo que o momento de quarta ordem exista, Bollerslev (1986) mostra que a curtose de um processo GARCH(1,1) é maior do que 3, que é a curtose de uma distribuição normal. He (1999a) e He (1999b) extenderam o resultado de Bollerslev para os modelos GARCH(p,q). Portanto, um modelo GARCH pode reproduzir as caudas pesadas observadas em séries temporais financeiras.

3.6 Estimação de Modelos GARCH

O objetivo dessa seção é estimar um modelo GARCH. Conforme apresentado anteriormente, o modelo geral GARCH(p,q) é da forma

$$\begin{aligned} y_t &= c + \epsilon_t \\ \epsilon_t &= z_t \sigma_t \\ \sigma_t^2 &= a_0 + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2, \end{aligned} \tag{3.8}$$

para $t = 1, \cdots, T$, onde $\sigma_t^2 = Var_{t-1}\left(\epsilon_t\right)$.

Serão introduzidos dois tipos de estimadores para os parâmetros c, a_0, a_i e b_j: estimador de máxima verossimilhança condicional e estimador de mínimos desvios absolutos. O primeiro é o estimador mais conhecido e é bastante usado na área de operações financeiras. O estimador de mínimos desvios absolutos é interessante quando tratamos de erros com caudas pesadas. Nas analises desse trabalho utilizaremos apenas o estimador de máxima verossimilhança condicional.

3.6.1 Estimador de Máxima Verossimilhança Condicional

Assim como na estimação dos modelos ARMA, os estimadores mais usados para modelos ARCH/GARCH são aqueles derivados de uma função Gaussiana de máxima verossimilhança (condicional). Por exemplo, se z_t em (3.8) tem distribuição normal e $q = 0$ (ou seja, um modelo ARCH puro), a função log-verossimilhança condicional negativa baseada nas observações $\epsilon_1, \cdots, \epsilon_T$, ignorando-se as constantes, é igual a

$$l\left(\sigma_t^2 | \epsilon_1, \cdots, \epsilon_T\right) = -\sum_{t=p+1}^{T} \left(\log \sigma_t^2 + \epsilon_t^2/\sigma_t^2\right), \tag{3.9}$$

onde $\sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2$. Os estimadores de máxima verossimilhança são obtidos minimizando-se a função em (3.9). Podemos perceber que esta função de verossimilhança é baseada na função densidade de probabilidade condicional de $\epsilon_{p+1}, \cdots, \epsilon_T$, dados, $\epsilon_1, \cdots, \epsilon_p$, uma vez que a função de densidade incondicional, que envolve a densidade conjunta de $\epsilon_1, \cdots, \epsilon_p$, é mais difícil de ser obtida.

Para o modelo GARCH geral, ou seja, $q > 0$ no modelo (3.4) a variância condicional σ_t^2 não pode ser expressa em termos de um número finito de observações passadas $\epsilon_{t-1}, \epsilon_{t-2}, \cdots$. Por indução, podemos escrever

$$\sigma_t^2 = \frac{a_0}{1 - \sum_{j=1}^{q} b_j} + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} a_i \sum_{k=1}^{\infty} \sum_{j_1=1}^{q} \cdots \sum_{j_k=1}^{q} b_{j_1} \cdots b_{j_k} \epsilon_{t-i-j_1-\cdots-j_k}^2, \tag{3.10}$$

onde a soma múltipla desaparece quando $q = 0$. Nota-se que a soma múltipla acima converge com probabilidade 1 desde que cada a_i e b_j seja não negativo, e desde que o valor esperado da série múltipla seja finito. Na prática, a expressão (3.10) é substituída por uma versão truncada

$$\begin{aligned} \widetilde{\sigma_t}^2 = \frac{a_0}{1 - \sum_{j=1}^{q} b_j} + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2 + \\ + \sum_{i=1}^{p} a_i \sum_{k=1}^{\infty} \sum_{j_1=1}^{q} \cdots \sum_{j_k=1}^{q} b_{j_1} \cdots b_{j_k} \\ \epsilon_{t-i-j_1-\cdots-j_k}^2 I\left(t - i - j_1 - \cdots - j_k \geq 1\right), t > p. \end{aligned} \tag{3.11}$$

Temos que, quando $q = 0$, $\widetilde{\sigma_t}^2 = \sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2$.
Considere a notação $a = (a_1, \cdots, a_p)^T$ e $b = (b_1, \cdots, b_q)^T$. O estimador de máxima verossimilhança (condicional) $\left(\widehat{a}_0, \widehat{a}, \widehat{b}\right)$ é definido maximizando-se a expressão

$$l_\nu\left(a_0, a, b\right) = -\sum_{t=\nu}^{T} \left(\log \widetilde{\sigma_t}^2 + \epsilon_t^2/\widetilde{\sigma_t}^2\right), \tag{3.12}$$

onde $\nu > p$ é um número inteiro. Em geral $\nu = p + 1$, onde ν são os graus de liberdade.
Supondo que $f(\cdot)$ seja a função densidade de probabilidades de z_t conhecida, temos que os estimadores de máxima verossimilhança são obtidos maximizando-se

$$l_\nu\left(a_0, a, b\right) = -\sum_{t=\nu}^{T} \left\{\log \widetilde{\sigma_t}^2 - 2 \log f\left(\epsilon_t/\widetilde{\sigma_t}\right)\right\}, \tag{3.13}$$

em vez de maximizar a função (3.12). Além da distribuição normal, algumas distribuições frequentemente usadas são:

- distribuição t com ν graus de liberdade com densidade apresentada em Marcucci (2005):

$$f(\epsilon_t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi}\Gamma(\frac{\nu}{2})}(\nu-2)^{-\frac{1}{2}}(h_t)^{-\frac{1}{2}}\left[1+\frac{\epsilon_t^2}{h_t(\nu-2)}\right]^{-\frac{(\nu+1)}{2}} \; ; \; \epsilon_t \in \Re,$$

onde $\nu > 2$ pode ser tratado como um parâmetro contínuo.

- distribuição Gaussiana generalizada com densidade apresentada em Marcucci (2005):

$$f(\epsilon_t) = \nu\left\{\lambda 2^{-\frac{2}{\nu}}\Gamma\left(1/\nu\right)\right\}^{-1}\exp\left\{-\frac{1}{2}\left|\frac{x}{\lambda}\right|^{\nu}\right\},$$

onde $\lambda = \left\{2^{-\frac{2}{\nu}}\Gamma\left(\frac{1}{\nu}\right)/\Gamma\left(\frac{3}{\nu}\right)\right\}^{1/2}$ e $0 < \nu < 2$.
Quando $\nu = 1$, a distribuição Gaussiana generalizada se reduz à função $f(x) = \exp\left\{-\sqrt{2}\,|x|\right\}/\sqrt{2}$. Todas as distribuições acima foram normalizadas para ter média 0 e variância 1, e todas elas possuem caudas mais pesadas do que a distribuição normal. Para comparar dois ou mais modelos ajustados, usamos os valores AIC e BIC para cada um dos modelos ajustados, que são calculados da seguinte maneira:

$$AIC = l_\nu\left(\widehat{a}_0, \widehat{a}, \widehat{b}\right) + 2\left(p+q+1\right), \tag{3.14}$$

e

$$BIC = l_\nu\left(\widehat{a}_0, \widehat{a}, \widehat{b}\right) + 2\left(p+q+1\right)\log\left(T-\nu+1\right), \tag{3.15}$$

onde $l_\nu(\cdot)$ segue de (3.13).

3.6.2 Estimadores de Mínimos Desvios Absolutos

O estimador apresentado anteriormente é derivado da maximização da verossimilhança Gaussiana ou de uma aproximação da verossimilhança Gaussiana. Sabe-se que os L_1-estimadores são mais robustos do que os L_2-estimadores, no que diz respeito às distribuições com caudas pesadas. Algumas evidências empíricas sugerem que este método é mais apropriado para séries financeiras pois estas apresentam caudas mais pesadas que a distribuição normal. Para detalhes ver Mandelbrot (1963), Fama (1965), Mittnik (1998) e Mittnik (2000).
A ideia deste estimador implica uma reparametrização do modelo (3.4), onde $E(z_t) = 0$ e a mediana de z_t^2, ao invés da variância de z_t, é igual a 1. Sob essa nova formulação, os parâmetros a_0 e a_i's diferem daqueles da formulação original por um fator constante, enquanto os parâmetros b_j's não se alteram. Seja

$$\epsilon_t^2/\sigma_t^2 = 1 + e_{t,1}, \tag{3.16}$$

onde $e_{t,1} = \left(z_t^2 - 1\right)$, que possui mediana igual a zero. Assim, temos que o estimador de mínimos desvios absolutos $\widehat{\theta}_1$, que é um L_1-estimador baseado na regressão (3.16), pode ser obtido minimizando-se

$$\sum_{t=\nu}^{T}\left|\epsilon_t^2/\widetilde{\sigma_t^2} - 1\right|, \tag{3.17}$$

onde $\widetilde{\sigma}_t^2$ é definido em (3.12) e $\nu = p+1$ se $q = 0$ e $\nu > p+1$ se $q > 0$. Peng. L. (2003) mostraram que este estimador é viesado. Para resolver este problema, Peng e Yao (2003) definiram uma forma modificada para o estimador de mínimos desvios absolutos $\left(\widehat{\theta}_2\right)$, que resulta da minimização da

expressão:

$$\sum_{t=\nu}^{T} \left|\log\left(\epsilon_t^2\right) - \log\left(\widetilde{\sigma}_t^2\right)\right|, \tag{3.18}$$

motivados pelo modelo de regressão

$$\log\left(\epsilon_t^2\right) = \log\left(\sigma_t^2\right) + e_{t,2}, \tag{3.19}$$

onde os erros $e_{t,2} = \log\left(z_t^2\right)$ são i.i.d com mediana zero.
Peng e Yao (2003) demonstraram que sob condições muito amenas, os estimadores de mínimos desvios absolutos são assintoticamente normais com razão de convergência padrão $T^{1/2}$, independentemente do fato de a distribuição de z_t possuir caudas pesadas ou não. Esta é uma diferença marcante entre estes estimadores e os estimadores de máxima verossimilhança condicional derivados de (3.13), que podem apresentar convergência lenta quando z_t possui caudas pesadas.
Outro estimador proposto por Peng e Yao (2003) foi motivado pela equação de regressão

$$\epsilon_t^2 = \sigma_t^2 + e_{t,3}, \tag{3.20}$$

onde $e_{t,3} = \sigma_t^2\left(z_t - 1\right)$ possui mediana igual a zero. Assim, o estimador $\widehat{\theta_3}$ é obtido minimizando-se:

$$\sum_{t=\nu}^{T} \left|\epsilon_t^2 - \widetilde{\sigma_t^2}\right|. \tag{3.21}$$

Intuitivamente, é preferível usar o estimador $\widehat{\theta_2}$ ao $\widehat{\theta_3}$ pois os termos de erros $e_{t,2}$ do modelo de regressão (3.20) são independentes e identicamente distribuídos enquanto que os erros $e_{t,3}$ do modelo (3.21) não são independentes.
Em Peng. L. (2003) foram realizadas comparações entre os três estimadores de mínimos desvios absolutos com o estimador de máxima verossimilhança Gaussiana, $\widehat{\theta_{ML}}$, para modelos simulados ARCH(2) e GARCH(1,1), tomando os erros z_t com distribuição normal padrão ou t-Student padronizada com $d = 3$ ou $d = 4$ graus de liberdade. A performance do estimador de máxima verossimilhança Gaussiana piora conforme as caudas da distribuição dos erros tendem a ser mais pesadas. Entretanto, esse comportamento não ocorre para os estimadores de mínimos desvios absolutos uma vez que estes apresentam-se mais robustos para caudas pesadas.

Para mais detalhes encontram-se em Peng. L. (2003).

3.7 Diagnóstico de Modelos GARCH

As seções anteriores apresentaram métodos para estimação dos modelo GARCH. Para diagnosticar os modelos ajustados são usados basicamente dois métodos: a análise gráfica e a análise inferencial.
Na análise das estatísticas resumo, podemos analisar os erros padrão e os níveis descritivos (valores p) das estatísticas para avaliar se os coeficientes do modelo são estatisticamente iguais a zero.
Além disso, existem testes variados para os resíduos padronizados $\widehat{\epsilon_t}/\widehat{\sigma_t}$. Para a série de log-retornos dos preços do índice Bovespa, as Tabelas 3.1 e 3.2 apresentam os coeficientes ajustados por meio do método de máxima verossimilhança condicional e alguns testes para resíduos padronizados do modelo GARCH(1,1): testes de normalidade Jarque (1980) e Shapiro (1965), teste de Ljung-Box para resíduos padronizados e para o quadrado dos resíduos padronizados e teste do multiplicador de Lagrange para efeito ARCH nos resíduos padronizados.
No teste Ljung-Box para os resíduos padronizados, não rejeitamos a hipótese nula de autocorrelação ($p = 0,465$). Por outro lado, no teste Ljung-Box para o quadrado dos resíduos padronizados, não rejeitamos a hipótese da não existência de autocorrelação ($p = 0,465$). Portanto, o modelo

captura com sucesso a estrutura de correlação serial na variância condicional mas não obtivemos um resultado satisfatório para a média condicional.
Além dos testes para autocorrelação nos resíduos padronizados, pode-se aplicar o teste do multiplicador de Lagrange para efeitos ARCH nos resíduos padronizados para verificar se existe algum efeito ARCH. Não rejeitamos a hipótese nula da não existência de efeitos ARCH ($p = 0,3117$).

Tabela 3.1: *Coeficientes Estimados do modelo GARCH(1,1)*

Coeficiente	Valor Estimado	Erro Padrão	Estat t	Valor p
C	0,10800	0,03196	3,379	0,0007
A	0,06877	0,01814	3,792	0,0001
ARCH(1)	0,07383	0,01050	7,028	0,0000
GARCH(1)	0,90369	0,01354	66,760	0,0002

O modelo GARCH básico assume distribuição normal para os erros ϵ_t. Se o modelo está corretamente especificado então os resíduos padronizados $\hat{\epsilon}_t/\hat{\sigma}_t$ devem se comportar como uma variável aleatória normal padrão. Os testes de normalidade Jarque-Bera e Shapiro-Wilks levam a outras conclusões, pois ambos rejeitam a suposição de normalidade ($p < 0,0001$). Para comparar os resíduos padronizados com os valores esperados se esses tivessem distribuição normal, construímos o gráfico quantil quantil (qqplot) apresentado na Figura 3.3. Nele podemos perceber que nas caudas há um significante desvio no gráfico de dispersão com relação à linha que corresponde ao caso de quantis normais, e portanto a suposição de normalidade dos resíduos não é apropriada.

Tabela 3.2: *Testes para diagnóstico do modelo GARCH(1,1) ajustado*

Testes de Normalidade:		
Jarque-Bera	P-valor	
110,0797	<0,0001	
Shapiro-Wilk	P-valor	
0,9921	<0,0001	
Teste Ljung-Box para os resíduos padronizados:		
Estatística	P-valor	χ^2-g.l
9,7236	0,4651	10
Teste Ljung-Box para o quadrado dos resíduos padronizados:		
Estatística	P-valor	χ^2-g.l
13,0993	0,2182	10
Teste do multiplicador de Lagrange para os resíduos padronizados:		
Estatística	P-valor	χ^2-g.l
13,8301	0,3117	10

Outros gráficos também podem ser utilizados para visualizar o ajuste do modelo. Por exemplo, na Figura 3.4 apresentamos a função de autocorrelação dos retornos padronizados e do quadrado dos retornos padronizados.

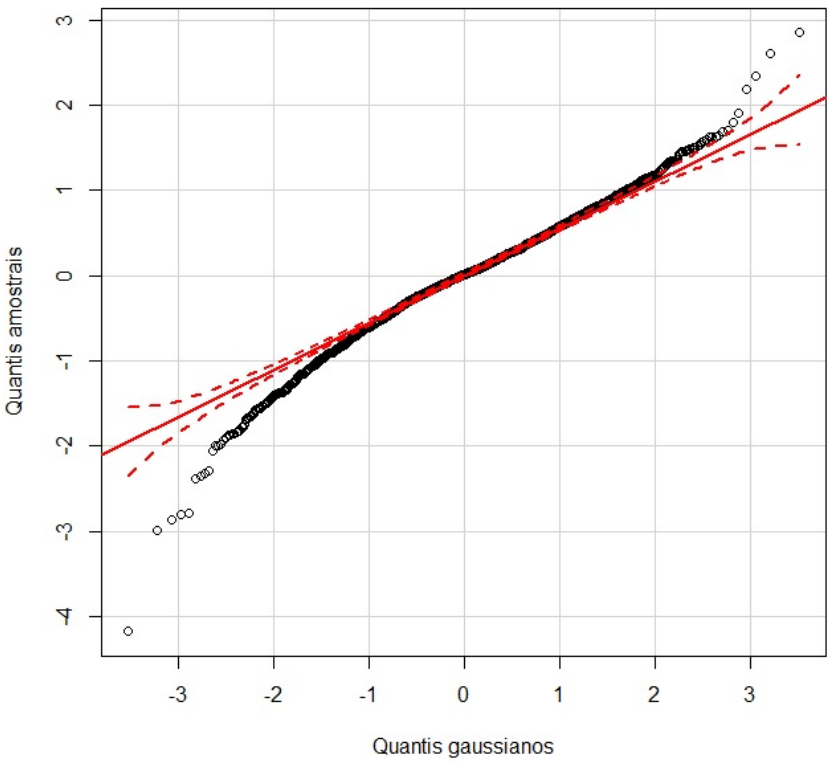

Figura 3.3: *Gráfico de quantis dos Resíduos padronizados do modelo GARCH(1,1) da série de retornos do Ibovespa*

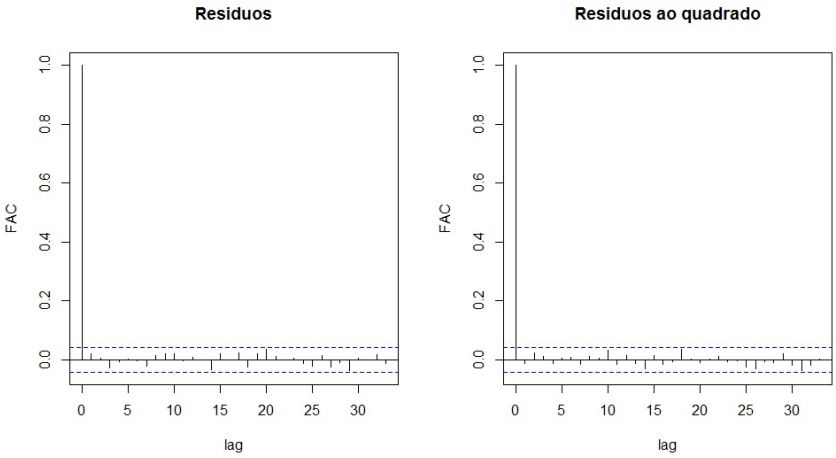

Figura 3.4: *Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados - Modelo GARCH(1,1) da série de retornos do Ibovespa*

3.8 Extensões dos Modelos GARCH

Em muitos casos, o modelo GARCH básico (3.4) fornece um modelo razoavelmente bom para analisar séries temporais financeiras e para estimar a volatilidade condicional. No entanto, existem alguns aspectos do modelo que podem ser melhorados para que, assim, ocorra uma melhor captura das características e da dinâmica de uma particular série temporal.

Efeitos de Alavancagem Assimétrica e Impactos de Informações Externas

No modelo GARCH básico (3.4), uma vez que apenas os resíduos ao quadrado ϵ_{t-i}^2 entram na equação, os sinais dos resíduos ou choques não têm efeitos na volatilidade condicional. Entretanto,

um fato estilizado da volatilidade financeria é que más notícias (choques negativos) tendem a oferecer maior impacto na volatilidade do que boas notícias (choques positivos). Black (1976), atribui esse efeito ao fato de que más notícias tendem a baixar os preços das ações, e portanto aumentam a alavancagem da ação causando maior volatilidade nessa ação. Baseado nessa conjectura, o impacto assimétrico das notícias, das informações externas, é normalmente chamado de efeito de alavancagem ou efeito alavanca (leverage effect). Um modo de incluir esse efeito é utilizar os modelos com efeito do sinal.

Em investimentos financeiros espera-se que alto risco resulte em altos retornos. Apesar da teoria moderna de precificação de ativos ou ações não implicar uma relação simples, ela sugere que existam algumas interações entre retornos esperados e risco que são medidas pela volatilidade. Engle (1987), Lilien e Robins propuseram extender o modelo GARCH básico de modo que a volatilidade condicional pudesse gerar um prêmio de risco *(risk premium)*, que é parte dos retornos esperados. Este modelo GARCH extendido é usualmente chamado de modelo GARCH-in-the-mean, ou modelo GARCH-M.

O modelo GARCH-M extende a equação da média condicional (3.1) conforme segue:

$$y_t = c + \alpha g\left(\sigma_t\right) + \epsilon_t$$

onde $g\left(\cdot\right)$ pode ser uma função arbitrária da volatilidade σ_t. Por exemplo, algumas possíveis funções $g\left(\sigma_t\right)$ podem ser σ_t, σ_t^2 ou $log\left(\sigma_t^2\right)$.

Variáveis Exógenas na Equação Geral da Média Condicional

Até o momento a equação da média condicional ficou restrita a uma constante nos modelos GARCH, com exceção do modelo GARCH-M onde a volatilidade leva em conta a equação da média como uma variável explicativa.
A forma geral da média condicional é dada por

$$y_t = c + \sum_{i=1}^{r} \phi_i y_{t-i} + \sum_{j=1}^{s} \theta_j \epsilon_{t-j} + \sum_{l=1}^{L} \beta_l x_{t-l} + \epsilon_t \tag{3.22}$$

onde x_t é um vetor $k \times 1$ de variáveis exógenas, e β_t é o vetor $k \times 1$ de coeficientes.

3.9 Distribuição de Erros Não-Gaussianos

Até o momento, foi utilizada a suposição da distribuição normal para os erros. No entanto, é conhecido o fato de que séries temporais financeiras possuem caudas pesadas, portanto é de interesse usar distribuições que possuem caudas mais pesadas do que as da distribuição normal. Descreveremos aqui três possiveis distribuições de erros com caudas pesadas para o ajuste de modelos GARCH: a distribuição t-Student, a distribuição dupla e a distribuição do erro generalizado.

Se uma variável aleatória u_t tem distribuição t-Student com ν graus de liberdade e com parâmetro de escala s_t, a função densidade de probabilidade de u_t é dada por

$$f\left(u_t\right) = \frac{\Gamma\left[\left(\nu+1\right)/2\right]}{\left(\pi\nu\right)^{1/2}\Gamma\left(\nu/2\right)} \frac{s_t^{-1/2}}{\left[1+\mu_t^2/\left(s_t\nu\right)\right]^{(\nu+1)/2}},$$

onde $\Gamma(\cdot)$ é a função gama. A variância de u_t é dada por:

$$Var(u_t) = \frac{s_t \nu}{\nu - 2}, \nu > 2$$

Se o termo de erro ϵ_t em um modelo GARCH segue uma distribuição t-Student com ν graus de liberdade e $Var_{t-1}(\epsilon_t) = \sigma_t^2$, o parâmetro de escala s_t pode ser escrito como

$$s_t = \frac{\sigma_t^2 (\nu - 2)}{\nu}.$$

Portanto, a função de log verossimilhança de um modelo GARCH com distribuição t-Student para os erros pode ser facilmente construída baseada na função densidade acima.

Distribuição do Erro Generalizada

Nelson (1991) propôs o uso da distribuição do erro generalizada (GED) para capturar as caudas pesadas geralmente observadas na distribuição de séries temporais financeiras. Se a variável aleatória u_t é uma GED com média zero e variância unitária, a função densidade de u_t é dada por:

$$f(u_t) = \frac{\nu exp\left[-(1/2)|u_t/\lambda|^{\nu}\right]}{\lambda \cdot 2^{(\nu+1)/\nu}\Gamma(1/\nu)},$$

onde

$$\lambda = \left[\frac{2^{-2/\nu}\Gamma(1/\nu)}{\Gamma(3/\nu)}\right]^{1/2}$$

e ν é um parâmetro positivo que determina o comportamento das caudas dessa distribuição. Quando $\nu = 2$, a densidade acima se reduz a densidade normal padrão; quando $\nu < 2$, a densidade possui caudas mais densas do que da distribuição normal; quando $\nu > 2$, a densidade possui caudas mais leves do que a da normal.
Quando $\nu = 1$, a densidade de GED se reduz à distribuição exponencial dupla:

$$f(\mu_t) = \frac{1}{\sqrt{2}} \exp^{-\sqrt{2}|\mu_t|}.$$

Baseado nas funções densidade apresentadas acima, a função de log verossimilhança do modelo GARCH com GED ou de erros com distribuição exponencial dupla podem ser facilmente construídas.

3.10 Seleção e Comparação de Modelos GARCH

As seções anteriores ilustraram extensões dos modelos GARCH. Selecionar o melhor modelo para determinado conjunto de dados pode ser uma tarefa muito complicada. O diagnóstico de modelos baseado nos resíduos padronizados e curvas de impacto para os efeitos de alavancagem pode ser usado para comparar a eficácia de diferentes aspectos dos modelos GARCH. Além disso, uma vez que os modelos GARCH podem ser tratados como modelos ARMA para resíduos quadráticos, os critérios tradicionais de seleção de modelos como o critério de informação de Akaike (AIC) e o critério da informação Bayesiana (BIC) também podem ser usados para selecionar os melhores modelos.
Considerando o ajuste de um modelo $GARCH(1,1)$ simples e um modelo com distribuição t-Student para a série de log-retornos do índice Bovespa, temos que o AIC e BIC do modelo com distribuição t-Student é menor do que o modelo com distribuição normal, o que sugere que a distribuição t-Student seja melhor que a distribuição normal. Os valores de BIC, AIC e log da verossimilhança

dos modelos ajustados são apresentados na Tabela 3.3.

Tabela 3.3: *AIC, BIC e log da verossimilhança dos modelos GARCH(1,1) ajustados para a série de retornos Ibovespa*

	Distribuição Normal	Distribuição t-Student
AIC	3,855	3,851
BIC	3,865	3,861
log-Verossimilhança	-4418,49	-4413,98

Podemos comparar graficamente os ajustes dos modelos. Na Figura 3.5 são apresentadas as FAC do quadrado dos resíduos padronizados dos modelos ajustados. Este gráfico sugere que os dois modelos são apropriados para modelar a volatilidade condicional. Podemos comparar também os gráficos QQ-plot dos resíduos padronizados, que estão apresentados na Figura 3.6. A análise gráfica reforça a conclusão obtida na análise dos critérios tradicionais de seleção de modelos, sugerindo também que a distribuição t-Student é melhor que a distribuição normal.

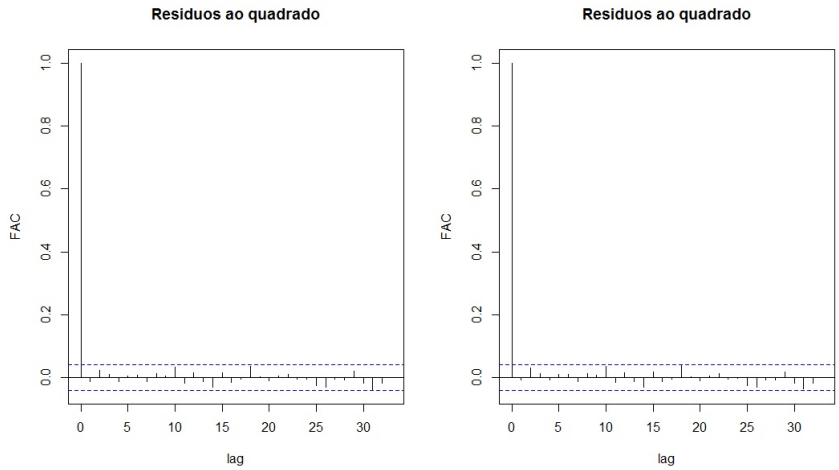

Figura 3.5: *Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados da Distribuição Normal e t-Student*

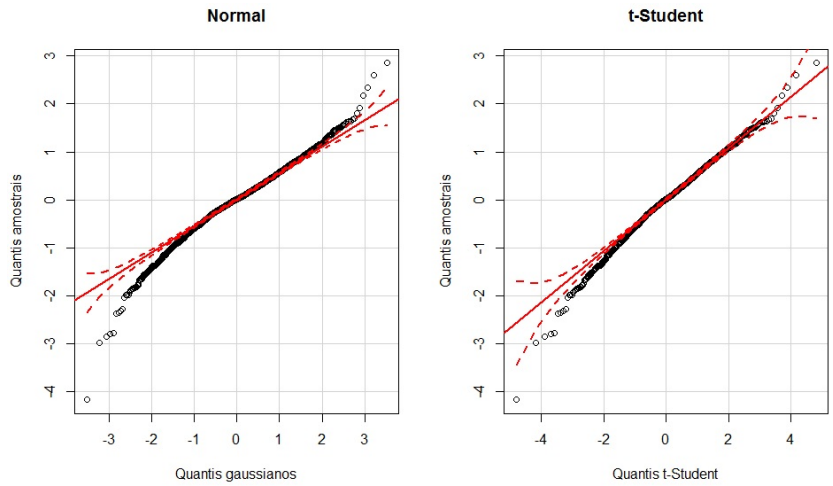

Figura 3.6: *Gráfico de quantis dos Resíduos Padronizados da Distribuição Normal e t-Student*

3.11 Predição de Modelos GARCH

Uma tarefa importante na modelagem da volatilidade condicional é a geração de previsões precisas tanto para valores futuros de uma série temporal financeira quanto para sua volatilidade condicional. Uma vez que a média condicional do modelo geral GARCH dado em (3.17) assume a forma de um ARMA tradicional, a previsão de valores futuros de uma série temporal pode ser obtida seguindo-se a abordagem tradicional para predição de modelos ARMA. Portanto, levando-se em conta também a variância condicional variando com o tempo, modelos GARCH podem gerar previsões adequadas para os valores futuros, especialmente para horizontes pequenos. Esta seção ilustra como prever a volatilidade usando modelos GARCH.
Para simplificar, consideremos o modelo básico GARCH(1,1):

$$\sigma_t^2 = a_0 + a_1\epsilon_{t-1}^2 + b_1\sigma_{t-1}^2,$$

com $t = 1, 2, \cdots, T$. Para obter $E_T\left[\sigma_{T+k}^2\right]$, que é a previsão da volatilidade futura σ_{T+k}^2, para $k > 0$, dada a informação no tempo T, a partir da equação anterior temos:

$$E_T\left[\sigma_{T+1}^2\right] \quad = \quad a_0 + a_1 E_T\left[\epsilon_T^2\right] + b_1 E_T\left[\sigma_T^2\right] = a_0 + a_1\epsilon_T^2 + b_1\sigma_T^2,$$

onde ϵ_T^2 e σ_T^2 são os valores obtidos depois da estimação. Para $T + 2$ temos:

$$E_T\left[\sigma_{T+2}^2\right] = a_0 + a_1 E_T\left[\epsilon_{T+1}^2\right] + b_1 E_T\left[\sigma_{T+1}^2\right] = a_0 + (a_1 + b_1)\, E_T\left[\sigma_{T+1}^2\right],$$

onde $E_T\left[\epsilon_{T+1}^2\right] = E_T\left[\sigma_{T+1}^2\right]$. Seguindo o mesmo raciocínio, pode-se obter a equação de previsão da volatilidade condicional:

$$E_T\left[\sigma_{T+k}^2\right] = a_0 \sum_{i=1}^{k-2} (a_1 + b_1)^i + (a_1 + b_1)^{k-1}\, E_T\left[\sigma_{T+1}^2\right], \tag{3.23}$$

para $k \geq 2$. Para $k \to \infty$, a previsão da volatilidade em (4.32) se aproxima da variância incondicional $a_0/\left(1 - a_1 - b_1\right)$ se o processo GARCH for estacionário (ou seja, se $a_1 + b_1 < 1$).
O algoritmo de predição (3.23) produz uma previsão para a variância condicional σ_{T+k}^2. A previsão para a volatilidade condicional, σ_{T+k}, é definida pela raiz quadrada da previsão de σ_{T+k}^2.

Capítulo 4

Modelo GARCH com Mudança de Regime

4.1 Modelo GARCH com mudança de regime Markoviano

A característica principal dos modelos com mudança de regime é a possibilidade de alguns, ou todos, os parâmetros mudarem para cada regime (ou estado) de acordo com um processo de Markov, que é regido por uma variável de estado, denotada por s_t. A ideia é ter uma mistura de distribuições com características diferentes, dos quais o modelo extrai o valor atual da variável, de acordo com o estado (não observado), que poderia ter determinado o seu valor. A variável estado evolui de acordo com uma cadeia de Markov estacionária de primeira ordem, com probabilidade de transição dada por

$$Pr(s_t = j|s_{t-1} = i) = p_{ij}, \tag{4.1}$$

que indica a probabilidade de mudança do estado i no tempo $t-1$ para o estado j no instante t. Geralmente, estas probabilidades são escritas na matriz de transição

$$P = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix},$$

considerando a existência de somente dois regimes. A probabilidade incondicional de estar no k-ésimo regime é dada por $P(s_t = 1) = \pi_1$, em que $\pi_1 = (1 - p_{22})/(2 - p_{11} - p_{22})$.

O modelo GARCH(1,1) com 2 regimes pode ser escrito na forma geral como:

$$r_t|\zeta_{t-1} = f\left(\theta_t^{(i)}\right), \quad \text{se} \quad s_t = i, \tag{4.2}$$

em que $i = 1$ ou 2 e $f(\cdot)$ representa uma das possíveis densidades condicionais que podem ser assumidas normal, t-Student ou GED, e $\theta_t^{(i)}$ denota o vetor de parâmetros no i-ésimo regime que caracteriza a distribuição.

O vetor de parâmetros é dado da seguinte forma:

$$\theta_t^{(i)} = \left(\mu_t^{(i)}, h_t^{(i)}, \nu_t^{(i)}\right),$$

onde $\mu_t^{(i)} \equiv E\left(r_t|\zeta_{t-1}\right)$ é a média condicional ou parâmetro de locação, $h_t^{(i)} \equiv Var\left(r_t|\zeta_{t-1}\right)$ é a variância condicional ou parâmetro de escala, $\nu_t^{(i)}$ é outro parâmetro da distribuição condicional e

ζ_{t-1} é a informação até o instante $t-1$. Por isso, a família de funções de densidade de r_t é uma família com parâmetros de locação ou escala que mudam ao longo do tempo de acordo com o regime $s_t = i$.

Portanto, o MRS-GARCH consiste de quatro elementos: a média condicional, a variância condicional, o parâmetro da distribuição condicional e o período de transição. A equação da média condicional, que é geralmente modelada por meio de um passeio aleatório, aqui é simplesmente modelada como

$$r_t = \mu_t^{(i)} + \epsilon_t, \tag{4.3}$$

onde $i = 1, 2$, $\epsilon_t = \eta_t \left[h_t^{(i)}\right]^{1/2}$ e η_t é um processo com média zero e variância unitária. A principal razão para esta escolha é o nosso foco na modelagem de volatilidade.

Uma informação que pode ser util é a duração média de cada regime dado que estamos num regime específico. (Hamilton (1989)).
Definindo W como a duração num determinado regime podemos observar que:

$$\begin{aligned} w = 1; &\quad \text{se} \quad s_t = i, s_{t+1} = 1, s_{t+2} = 2; p(w = 1|s_t = i; s_{t+1} = i) = 1 - p_{ii} \\ w = 2; &\quad \text{se} \quad s_t = s_{t+1} = i, s_{t+2} = i; p(w = 2|s_t = s_{t+1} = i; s_{t+2} = i) = p_{ii}\left(1 - p_{ii}\right). \end{aligned}$$

Observe que a variável aleatória W tem distribuição geométrica e, portanto a duração esperada do regime $s_t = i$ com $i = 1, 2$ é dada por:

$$E\left[w\right] = \frac{1}{1 - p_{ii}}.$$

4.2 MRS-GARCH

Recentemente, Klaassen (2002) sugere o uso da média condicional e variância condicional dos regimes do passado, levando em conta também o atual. Klaassen (2002) adota a seguinte expressão para a variância condicional

$$h_t^{(i)} = \alpha_0^{(i)} + \alpha_1^{(i)}\epsilon_{t-1}^2 + \beta_1^{(i)} E_{t-1}\left\{h_{t-1}^{(i)}|s_t\right\}, \tag{4.4}$$

onde a esperança condicional ao regime s_t é

$$E_{t-1}\left\{h_{t-1}^{(i)}|s_t\right\} = \sum_{j=1}^{2} \widetilde{p}_{ij,t-1}\left[\left(\mu_{t-1}^{(j)^2} + h_{t-1}^{(j)}\right)\right] - \left[\sum_{j=1}^{2} \widetilde{p}_{ji,t-1}\mu_{t-1}^{(j)}\right]^2.$$

e as probabilidades condicionais são dadas por

$$\widetilde{p}_{ij,t} = Pr\left(s_t = j|s_{t+1} = i, \zeta_{t-1}\right) = \frac{p_{ij} Pr\left(s_t = j|\zeta_{t-1}\right)}{Pr\left(s_{t+1} = i|\zeta_{t-1}\right)} = \frac{p_{ji}p_{j,t}}{p_{i,t+1}},$$

onde $i, j = 1, 2$, e ζ_{t-1} denota a informação até o instante $t-1$.

Na literatura é usual utilizar o método de máxima verossimilhança para estimar parâmetros do modelo de mudança de regime markoviano. O essencial é obter a probabilidade a priori $p_{1,t} = Pr\left[s_t = 1|\zeta_{t-1}\right]$ ou seja, a probabilidade do regime 1 no tempo t dada a informação no instante

$t-1$, o qual é dada por

$$p_{j,t} = Pr\left(s_t = j|\zeta_{t-1}\right) = \sum_{i=1}^{2} p_{ij}\left[\frac{f\left(r_{t-1}|s_{t-1}=i\right)p_{i,t-1}}{\sum_{k=1}^{2} f\left(r_{t-1}|s_{t-1}=k\right)p_{k,t-1}}\right],$$

onde p_{ij} são as probabilidades de transição e $f(\cdot)$ é a densidade de probabilidade dos log-retornos. Então a função de log-verossimilhança pode ser escrita como

$$l = \sum_{t=-R+w+1}^{T+w} \log\left[p_{1,t} f\left(r_t|s_t=1\right) + \left(1-p_{1,t}\right) f\left(r_t|s_t=2\right)\right],$$

onde $w = 0, 1, ..., n$, e $f(\cdot|s_t = i)$ é a distribuição condicional dado que o regime i ocorre no tempo t.

4.3 Estimação

Nesta seção derivamos a função de verossimilhança do modelo GARCH com mudança de regime e pode-se encontrar em Gray (1996a) que esta tem uma estrutura recursiva de primeira ordem. Seja $p_{t-1}(r_t)$ que denota a densidade avaliada no retorno igual a r_t. Isso pode ser expressado como

$$p_{t-1}(r_t) = \sum_{s_t=1,2} p_{t-1}(r_t|s_t)\cdot p_{t-1}(s_t). \tag{4.5}$$

O primeiro termo do lado direiro $p_{t-1}(r_t|s_t)$ denota a densidade do retorno do preço da ação até o tempo t avaliado no valor r_t condicionada em ζ_{t-1} e o regime tendo valor s_t. Esta densidade segue as formulas (4.2) e (4.3), entretanto não é simples calcular a variância condicional em (4.3). Isso requer a integração do regime $\tilde{s}_{t-1}$ em $E_{t-1}\left[Var_{t-2}\left\{\epsilon_{t-1}|\tilde{s}_{t-1}\right\}|s_t\right]$, porque $Var_{t-2}\left\{\epsilon_{t-1}|\tilde{s}_{t-1}\right\}$ depende somente de s_{t-1} e precisamos apenas de $p_{t-1}(s_{t-1}|s_t)$, a probabilidade que o regime anterior era s_{t-1} dado que o regime atual é s_t e dada a informação ζ_{t-1}:

$$p_{t-1}(s_{t-1}|s_t) = \frac{p_{t-1}(s_{t-1})\cdot p_{t-1}(s_t|s_{t-1})}{p_{t-1}(s_t)} \tag{4.6}$$

onde

$$p_{t-1}(s_t) = \sum_{s_t=1,2} p_{t-1}(s_{t-1})\cdot p_{t-1}(s_t|s_{t-1}). \tag{4.7}$$

A probabilidade de mudança de regime $p_{t-1}(s_t|s_{t-1})$ segue de (4.1).
O termo restante em (4.6) e (4.7) é $p_{t-1}(s_{t-1})$. Esta probabilidade é crucial já que todos os regimes de probabilidade deste trabalho podem ser derivados dele. Usando a tecnica similar à de Gray (1996a), a seguinte formula mostra que estas probabilidades tem uma estrutura recursiva de primeira ordem, o que simplifica substancialmente o seu cálculo:

$$\begin{aligned} p_{t-1}(s_{t-1}) &= p_{t-2}(s_{t-1}|r_{t-1}) \\ &= \frac{p_{t-2}(r_{t-1}|s_{t-1})\cdot p_{t-2}(s_{t-1})}{p_{t-2}(r_{t-1})} \\ &= \frac{p_{t-2}(r_{t-1}|s_{t-1})\cdot \sum_{s_{t-2}=1,2} p_{t-2}(s_{t-2})\cdot p_{t-2}(s_{t-1}|s_{t-2})}{p_{t-2}(r_{t-1})} \end{aligned} \tag{4.8}$$

Por isso, as variáveis para calcular $p_{t-1}(s_{t-1})$ são seus valores anteriores $p_{t-2}(s_{t-2})$ e a constante $p_{t-2}(s_{t-1}|s_{t-2})$ para $s_{t-2} = 1, 2$ e as densidades anteriores $p_{t-2}(r_{t-1}|s_{t-1})$ e $p_{t-2}(s_{t-1})$. Isto faz o cálculo de $p_{t-1}(s_{t-1})$ um processo recursivo de primeira ordem. O segundo termo do lado direito de

(4.5), $p_{t-1}(s_t)$, é a probabilidade condicional que o regime atual é s_t. Isso é dado por (4.7).

Tendo discutido ambos termos do lado direito de (4.5), podemos agora calcular a densidade de interesse, sendo uma distribuição normal, t-student ou GED. Estas densidades podem então ser usadas para construir a log-verossimilhança $\sum_{t=1}^{T} log(p_{t-1}(r_t))$ com o qual todos os parâmetros do modelo GARCH com mudança de regime podem ser estimados.
Do ponto de vista prático, é importante perceber que a log-verossimilhança tem uma estrutura recursiva de primeira ordem, similar ao modelo padrão, modelo GARCH(1,1) com um regime. Depois de tudo, de (4.6) e (4.7) precisamos apenas da constante $p_{t-1}(s_t|s_{t-1})$e da probabilidade recursiva de primeira ordem $p_{t-1}(s_{t-1})$ em (4.7) para todas as 4 combinações de (s_t, s_{t-1}); então a densidade (4.5) pode ser calculada de (4.7), a mudança anterior r_{t-1}, (4.6) e as variâncias anteriores $V_{t-2}\{\epsilon_{t-1}|s_{t-1}\}$ para $s_{t-1} = 1, 2$. Esta recursividade de primeira ordem de $p_{t-1}(r_t)$ agiliza substancialmente o calculo da log-verossimilhança.
Para começar o processo recursivo, nós definimos o conjunto de variáveis igual à esperança sem o condicionamento do conjunto de informações, isto é a variância incondicional $h_t^{(i)}$ dada em (4.4).

4.4 Probabilidade Suavizada

A probabilidade suavizada que o regime é s_t no tempo t, $p_T(s_t)$, é a probabilidade do regime s_t dada a informação da série inteira e pode ser calculada recursivamente. Mais geralmente, qualquer probabilidade do regime anterior $p_\tau(s_t)$, dado um tempo futuro $\tau \in \{t, t+1, \ldots T\}$, pode ser calculado em uma forma recursiva, a partir da probabilidade anterior $p_{t-1}(s_t)$. Nesta seção será verificado esta afirmação.

Pode-se escrever $p_\tau(s_t)$ para ambos regimes $s_t = 1, 2$ conforme

$$\begin{aligned} p_\tau(s_t) &= p_{\tau-1}(s_t|r_\tau) \\ &= \frac{p_{\tau-1}(r_\tau|s_t) \cdot p_{\tau-1}(s_t)}{\sum_{s_t=1,2} p_{\tau-1}(r_\tau|s_t) \cdot p_{\tau-1}(s_t)}. \end{aligned} \tag{4.9}$$

Suponha primeiro que $\tau = t$. Então $p_\tau(s_t)$ segue diretamente, como $p_{\tau-1}(s_t)$ e $p_{\tau-1}(r_\tau|s_t)$ em (4.9) são conhecidos a partir do processo de estimação (ver seção de estimação).

Assim, vamos supor a partir de agora que $\tau > t$. O cálculo de (4.9) exige duas entradas. A primeira é a probabilidade anterior $p_{\tau-1}(s_t)$, que é conhecido a partir da recursão anterior para ambos os regimes. O segundo ingrediente de (4.9) é a densidade $p_{\tau-1}(r_\tau|s_t)$ para ambos os regimes. O seu cálculo exige uma série de etapas. Primeiramente escrevemos como

$$p_{\tau-1}(r_\tau|s_t) = \sum_{s_\tau=1,2} p_{\tau-1}(r_\tau|s_\tau) \cdot p_{\tau-1}(s_\tau|s_t), \tag{4.10}$$

onde usamos que a distribuição condicional de r_τ dado s_τ não depende do regime anterior s_t. Esta fórmula em si tem dois ingredientes. O primeiro é a densidade $p_{\tau-1}(r_\tau|s_\tau)$ para ambas combinações do regime, que é conhecida do processo de estimação. O segundo termo necessário em (4.10) é o $(\tau - t)$ período à frente para a probabilidade de mudança de regime $p_{\tau-1}(s_\tau|s_t)$ para todos os resultados do regime. Uma vez que tenha sido calculado, ele deve ser salvo, já que será necessário em uma próxima fase recursiva. Fazendo uso da estrutura do processo de regime Markov, ele pode ser escrito em termos de $(\tau - 1 - t)$ probabilidade de mudança em um período à frente

$$p_{\tau-1}(s_\tau|s_t) = \sum_{s_{\tau-1}=1,2} p_{\tau-1}(s_\tau|s_{\tau-1}) \cdot p_{\tau-1}(s_{\tau-1}|s_t). \tag{4.11}$$

Mais uma vez, há dois ingredientes. Primeiro, precisamos $p_{\tau-1}(s_\tau|s_{\tau-1})$ para todas as combinações do regime. Estes são constantes e seguem a partir de (4.1).
O segundo ingrediente de (4.11) é $p_{\tau-1}(s_{\tau-1}|s_\tau)$ para todas as combinações do regime. Temos:

$$\begin{aligned} p_{\tau-1}(s_{\tau-1}|s_t) &= p_{\tau-2}(s_{\tau-1}|s_t, r_{\tau-1}) \\ &= \frac{p_{\tau-2}(r_{\tau-1}|s_{\tau-1}) \cdot p_{\tau-2}(s_{\tau-1}|s_t)}{\sum_{s_{\tau-1}=1,2} p_{\tau-2}(r_{\tau-1}|s_{\tau-1}) \cdot p_{\tau-2}(s_{\tau-1}|s_t)} \end{aligned} \tag{4.12}$$

onde usamos essa densidade condicional de $r_{\tau-1}$ é independente do regime anterior s_t uma vez que $s_{\tau-1}$ é dado. Temos dois ingredientes. Primeiro, a densidade condicional $p_{\tau-2}(r_{\tau-1}|s_{\tau-1})$ para ambos resultados do regime. É conhecida a partir do processo de estimação. Em segundo lugar, precisamos de $(\tau - 1 - t)$ probabilidade de mudança em um período a frente $p_{\tau-2}(s_{\tau-1}|s_t)$ para todas as combinações do regime. Este foi salvo durante a recursão anterior, se $\tau > t+1$. Se $\tau = t+1$, isto é igual a um. Isto completa o algoritmo para calcular (4.10), que é o segundo ingrediente de (4.9). Para cada uma recursividade tem de calcular (4.12), utilizar o resultado para calcular (4.11) e usar esta para calcular (4.10). Usando isso em (4.9) produz a probabilidade posterior $p_\tau(s_t)$ de $p_{\tau-1}(s_t)$. Portanto, a partir da probabilidade anterior $p_{t-1}(s_t)$ pode-se recursivamente calcular a probabilidade posterior $p_\tau(s_t)$.

Capítulo 5

Aplicação em Séries Reais

5.1 Aplicações

Neste capítulo serão apresentadas algumas aplicações a dados reais dos modelos apresentados nos capítulos anteriores. As aplicações foram realizadas em séries de log-retornos dos índices de duas grandes bolsas de valores, Ibovespa e S&P500, obtidos no site yahoo finanças. Chamaremos os log-retornos simplesmente de retornos. O software utilizado para o ajuste do modelo com mudança de regime foi o Matlab e as outras análises foram realizadas no R-project. Ver Zivot (2003) para mais detalhes sobre o uso deste software no estudo de séries financeiras.

5.2 IBOVESPA

O conjunto de dados analisado neste trabalho é o índice do mercado de ações de São Paulo IBOVESPA. O período considerado é de 2 de janeiro de 2003 a 9 de abril de 2012 para um total de 2294 observações. Os dados foram obtidos a partir do site yahoo finanças. A cotação estudada foi o preço de fechamento diário do índice.

Primeiramente, apresentamos nas Figuras 5.1 e 5.2, a série Ibovespa, a série de retornos, gráfico de quantis e o histograma dos retornos.

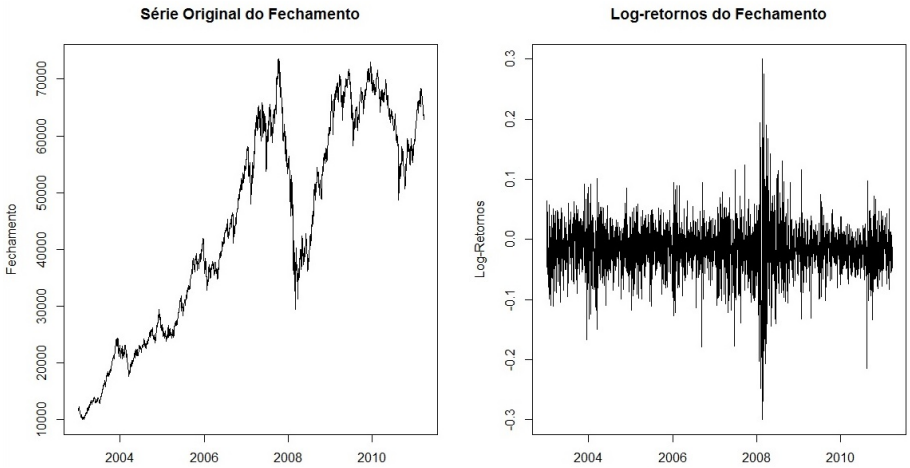

Figura 5.1: *Gráfico da Série do índice Bovespa e os Log-retornos.*

A Tabela 5.1 mostra algumas estatísticas descritivas dos log-retornos do Índice Bovespa. Note que a média dos dados está próxima de zero e o desvio padrão é pequeno, mas os log-retornos variam de -0.12 a 0.13. A média é próxima de zero e, com o desvio padrão pequeno, podemos concluir que os dados se encontram aglomerados em torno do zero, o que pode ser confirmado pelo histograma dos retornos diários do IBOVESPA. Observando o gráfico da série diária do IBOVESPA, podemos ver a que a série é não estacionária, pois apresenta tendência não constante e variabilidade

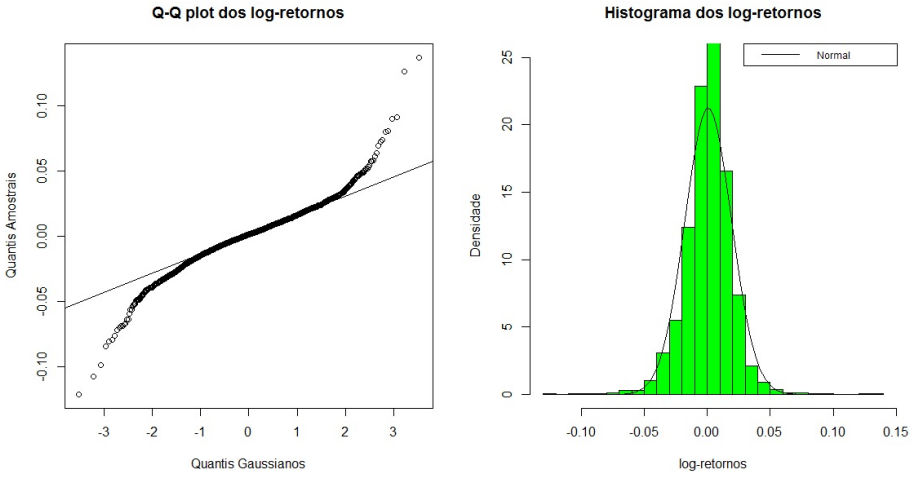

Figura 5.2: *Gráfico de quantis e Histograma dos log-retornos da Série do índice Bovespa.*

Tabela 5.1: *Medidas descritivas do retorno-Ibovespa*

Média	Mediana	D.Padrão	Curtose	Mínimo	Máximo	Assimetria
0.00073	0.00144	0.01876	8.20814	-0.12100	0.136800	-0.11304

não controlada. Já o gráfico dos retornos diários do IBOVESPA, apresenta nível, aparentemente, constante e variabilidade controlada, o que nos leva a pensar que esta série pode ser estacionária. A assimetria e a curtose indicam que os dados não seguem uma distribuição normal. Com base nessas medidas resumo dos retornos diários da IBOVESPA, podemos observar que: a distribuição dos retornos possui uma leve assimetria à esquerda, já que o valor do coeficiente de assimetria é negativo, a medida de curtose da distribuição dos retornos indica que sua curva é mais achatada que a curva da distribuição Normal, ou seja, há uma maior quantidade de dados nas caudas da distribuição dos retornos do que em uma distribuição Normal. A assimetria e a curtose podem ser visualizadas no gráfico de quantis e histograma da Figura 5.2.

Entretanto, na Figura 5.3 observa-se que as autocorrelações dos retornos são não significativas e as autocorrelações dos retornos ao quadrado são significativas. Conforme descrito no Capítulo 3, este comportamento indica que a série de retornos pode mostrar presença de heterocedasticidade condicional nos dados, una vez que os retornos ao quadrado medem o momento de segunda ordem da série original, este resultado indica que a variância dos retornos da série do Ibovespa condicionada ao passado histórico pode sofrer alterações ao longo do tempo.

As estimativas dos parâmetros do modelo MRS-GARCH são apresentadas na Tabela 5.2. Foram ajustados os modelos com distribuição Normal, t de Student e a distribuição generalizada do erro. Para o modelo com distribuição t-Student foi considerado o modelo em que o grau de liberdade é o mesmo nos 2 regimes (t) e com graus de liberdade que podem ser diferentes (t2). Note que as estimativas e os erros padrão são bem semelhantes para os modelos com distribuição t-Student. Em particular, a Tabela 5.2 mostra o desvio padrão condicional dos retornos para cada regime da volatilidade, isto é

$$\sigma^{(i)} = \left(a_0^{(i)} / \left(1 - a_1^{(i)} - b_1^{(i)}\right)\right)^{1/2}.$$

Para os modelos t de Student, os parâmetros da média parecem negativos para o regime 1 e positivos para o regime 2. Isso parece ser o contrário para o modelo normal e para o modelo GED, parece que a média é nula para o regime 1 e positiva para o regime 2. O parâmetro a_1, que mede o efeito do erro anterior na variância, parece não significativo nos modelos normal e t de Student.

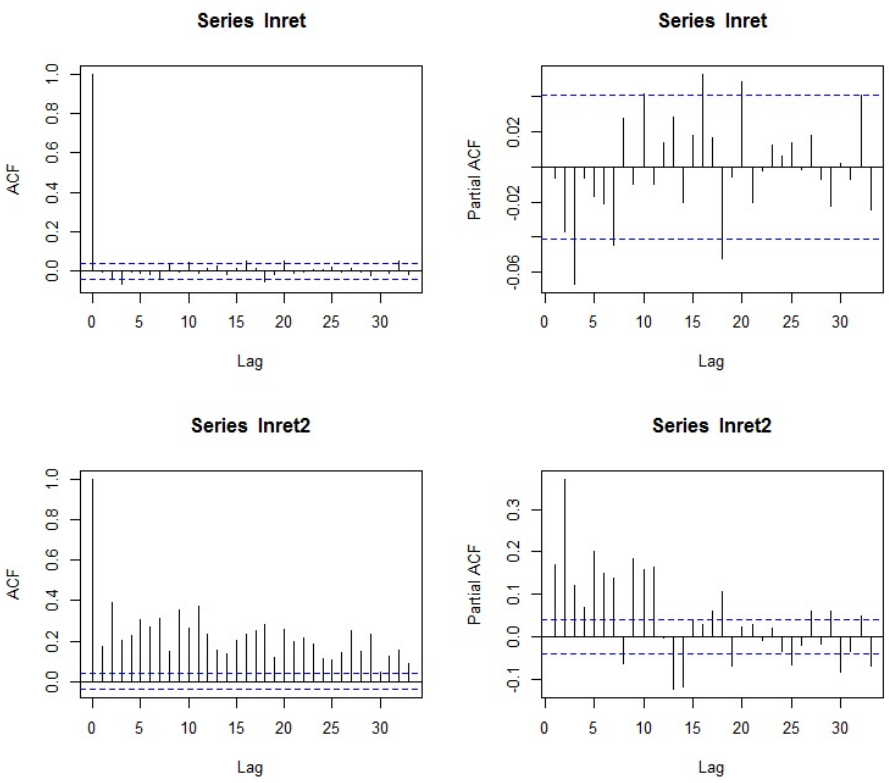

Figura 5.3: *Função de autocorrelação dos log-retornos e log-retornos ao quadrado-IBOVESPA.*

Para o regime 1, a estimativa é quase zero para o modelo normal e está em torno de 0.16 para os modelos t de Student. Entretanto, os valores para os parâmetros b_1 são menores no regime 1, isso significa que temos mais peso para cada erro de ontem e menor inércia para a variância anterior σ^2_{t-1}. A probabilidade de se manter no regime 1, p_{11} é menor (38% para a distribuição t), que a probabilidade de se manter no regime 2, o mesmo acontece para a distribuição t de Student com 2 graus de liberdade.

A Tabela 5.2 também mostra as probabilidades incondicionais de cada modelo MRS-GARCH. A probabilidade incondicional π_1 de estar no primeiro regime varia entre 10% para o modelo t de Student e 66% para o modelo normal. Para o modelo t-Student com 10 e 12 graus de liberdade no primeiro regime, o parâmetro ν indica que a distribuição tem caudas mais pesadas do que a normal. Já para o segundo regime, o valor do parâmetro ν indica que a distribuição é aproximadamente normal. No caso do GED o valor do parâmetro ν está indicando que a distribuição tem caudas mais pesadas do que a normal.

A identificação de regime com caudas mais pesadas no regime 2 da distribuição t de Student nos leva a escolher o modelo t2. O modelo MRS-GARCH com distribuição t-Student em que os graus de liberdade mudam em cada regime implica que a curtose varia como em Hansen (1994) e Dueker (1997). No entanto existe uma diferença entre os trabalhos mencionados. Hansen (1994) sugere um modelo no qual o parâmetro ν varia ao longo do tempo de acordo com uma função de distribuição logística no tempo $t-1$, e Dueker (1997) mostra como um parâmetro depende apenas do regime, no nosso trabalho este parâmetro só depende do regime, juntamente com todos os restantes parâmetros.

Tabela 5.2: *Estimativas de Máxima Verossimilhança erros padrões do Modelo MRS - GARCH - Ibovespa*

	MRS-GARCH-N	MRS-GARCH-t2	MRS-GARCH-t	MRS-GARCH-GED
$\mu^{(1)}$	0.2605	-1.2336	-1.2622	0.0332
	(0.0468)	(0.4775)	(0.5895)	(0.0285)
$\mu^{(2)}$	-0.3768	0.2425	0.2300	0.0419
	(0.1302)	(0.0652)	(0.0651)	(0.0137)
$\sigma^{(1)}$	0.1033	< 0.0001	0.0851	0.1332
	(0.0531)	(0.4204)	(0.6107)	(0.0381)
$\sigma^{(2)}$	0.4053	< 0.0001	< 0.0001	0.0029
	(0.1047)	(0.0572)	(0.0656)	(0.0012)
$a_1^{(1)}$	< 0.0001	0.1768	0.1546	0.1002
	(0.0273)	(0.1275)	(0.1276)	(0.0220)
$a_1^{(2)}$	0.0645	0.0528	0.0520	0.0142
	(0.0262)	(0.0184)	(0.0184)	(0.0052)
$b_1^{(1)}$	0.8457	0.8111	0.8332	0.8278
	(0.0318)	(0.1712)	(0.1701)	(0.0330)
$b_1^{(2)}$	0.9315	0.8797	0.8805	0.9785
	(0.0361)	(0.0272)	(0.0277)	(0.0060)
p_{11}	0.9107	0.3817	0.3799	0.9987
	(0.0285)	(0.1943)	(0.1988)	(0.0011)
p_{22}	0.8197	0.9243	0.9304	0.9989
	(0.0476)	(0.0456)	(0.0466)	(< 0.0001)
$\nu^{(1)}$	——	10.2401	12.1334	1.2667
		(2.2062)	(3.3496)	(0.0309)
$\nu^{(2)}$	——	81.3425	——	——
		(< 0.0001)		
$1000 * Log(L)$	-4.3928	-4.3861	-4.3867	-4.3877
π_1	0.6687	0.1091	0.1009	0.4583
π_2	0.3313	0.8909	0.8991	0.5417
ρ_1	0.8457	0.9879	0.9878	0.9280
ρ_2	0.9960	0.9325	0.9325	0.9927

As probabilidades suavizadas do regime 1 com erro gaussiano e os log-retornos podem ser observados no seguinte gráfico:

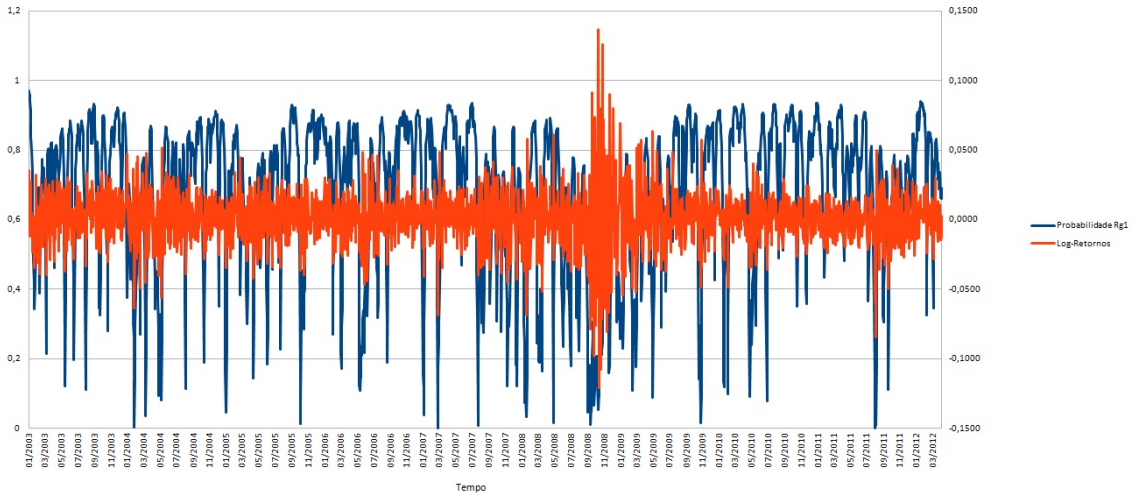

Figura 5.4: *Probabilidade suavizada do Regime 1 - Ibovespa - Normal.*

Em nossa aplicação existem dois regimes, $s_t = 1$ e $s_t = 2$, ou seja, se o processo estiver no regime 1, a observação será retirada de uma distribuição normal com média 0.2605 e desvio padrão 0.1033 e se estiver no segundo regime com média -0.3768 e desvio padrão 0.4053. A Figura 5.4 apresenta a probabilidade suavizada do regime 1 e os log-retornos. As probabilidades de estarmos no regime 1, variam entre 0 e 0.9 e as probabilidades de estarmos em um regime e continuarmos neste são p_{11} e p_{22} iguais a 0.9107 e 0.8197 respetivamente. Com isto temos que a matriz de transição é dada por

$$P = \begin{pmatrix} 0.9107 & 0.1803 \\ 0.0893 & 0.8197 \end{pmatrix}.$$

Temos então que a probabilidade incondicional de estar no k-ésimo regime, neste caso para nosso modelo com 2 regimes é dada por $\pi_1 = 0.6687$ e $\pi_2 = 0.3313$. Ou seja, ao sortearmos uma observação qualquer da série, a probabilidade é maior de estarmos no regime 1 do que no regime 2.
Uma informação que pode ser util é a duração média de cada regime dado que estamos num regime específico. (Hamilton (1989)).
Definindo W como a duração num determinado regime podemos observar que:

$$\begin{aligned} w = 1; &\quad \text{se} \quad s_t = i, s_{t+1} = 1, s_{t+2} = 2; p(w = 1|s_t = i; s_{t+1} = i) = 1 - p_{ii} \\ w = 2; &\quad \text{se} \quad s_t = s_{t+1} = i, s_{t+2} = i; p(w = 2|s_t = s_{t+1} = i; s_{t+2} = i) = p_{ii}\,(1 - p_{ii})\,. \end{aligned}$$

Observe que a variável aleatória W tem distribuição geométrica e, portanto a duração esperada do regime $s_t = i$ com $i = 1, 2$ é dada por:

$$E\,[w] = \frac{1}{1 - p_{ii}}.$$

Logo a duração média de estarmos no regime 1 é

$$p_{11} = \frac{1}{1 - p_{11}} = \frac{1}{1 - 0.9107} = 11 \quad \text{dias.}$$

E para o regime 2 é

$$p_{22} = \frac{1}{1 - p_{22}} = \frac{1}{1 - 0.8197} = 5 \quad \text{dias.}$$

Ao observarmos o desvio padrão condicional estimado pelo modelo com distribuição normal para a série IBOVESPA na Figura 5.5, observa-se que a volatilidade aumentou bastante durante a crise do subprime de 2008 em ambos os regimes.

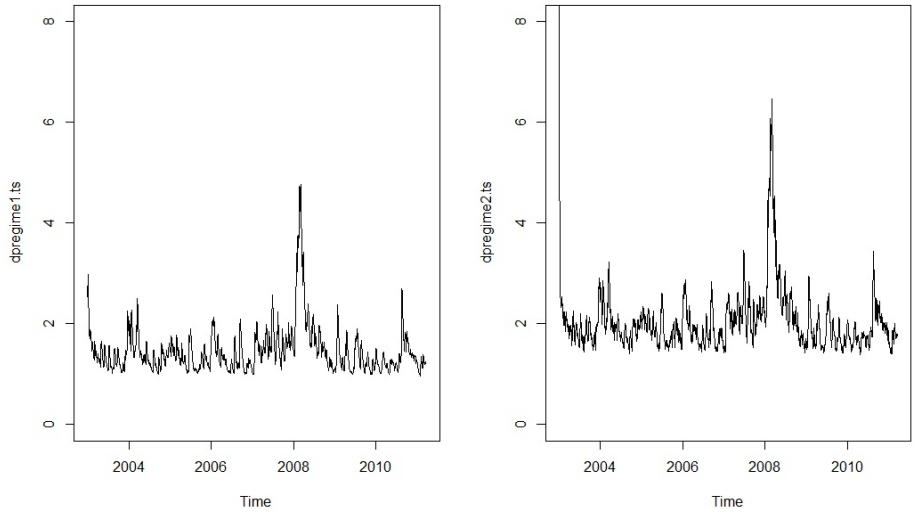

Figura 5.5: *Estimativa do desvio padrão condicional para os retornos do Ibovespa no Regime 1 e 2 com erro gaussiano.*

As probabilidades suavizadas do regime 1 com erro t de Student e os log-retornos podem ser observados no seguinte gráfico:

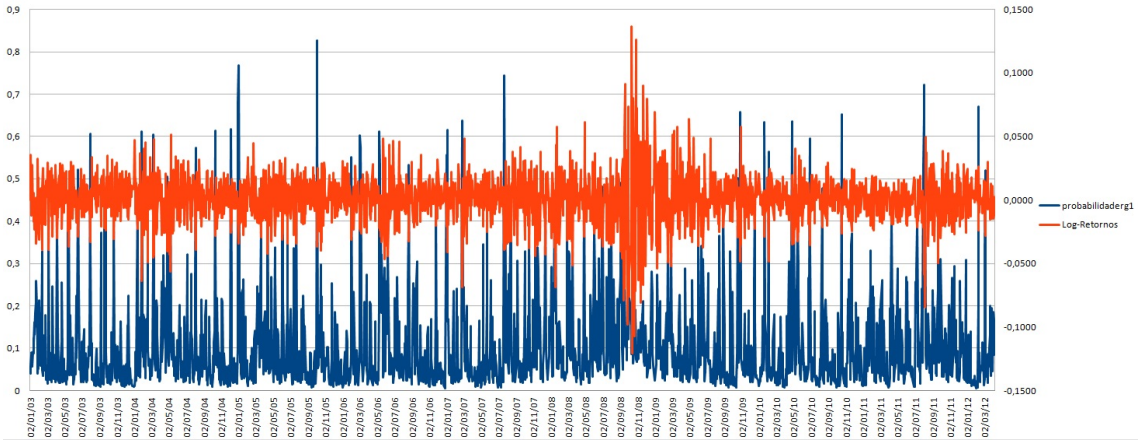

Figura 5.6: *Probabilidade suavizada do Regime 1 - Ibovespa - t.*

Se o processo estiver no regime 1, a observação será retirada de uma distribuição t de Student com média -1.2622 e desvio padrão 0.0651 e se estiver no segundo regime com média 0.2300 e desvio padrão < 0.0001. A Figura 5.6 apresenta a probabilidade suavizada do regime 1 e os log-retornos. As probabilidades de estarmos no regime 1, variam entre 0 e 0.4 e as probabilidade de estarmos em um regime e continuarmos neste p_{11} e p_{22} são 0.3799 e 0.9304. Com isto temos que a matriz de transição é dada por

$$P = \begin{pmatrix} 0.3799 & 0.0696 \\ 0.6201 & 0.9304 \end{pmatrix}.$$

Temos então que a probabilidade incondicional de estar no k-ésimo regime, neste caso para nosso modelo com 2 regimes é dada por $\pi_1 = 0.1009$ e $\pi_2 = 0.8991$. Ou seja, ao sortearmos uma observação qualquer da série, a probabilidade é maior de estarmos no regime 2 do que no regime 1.

A duração média de estarmos no regime 1 é 1 dia e do regime 2 é de 14 dias.

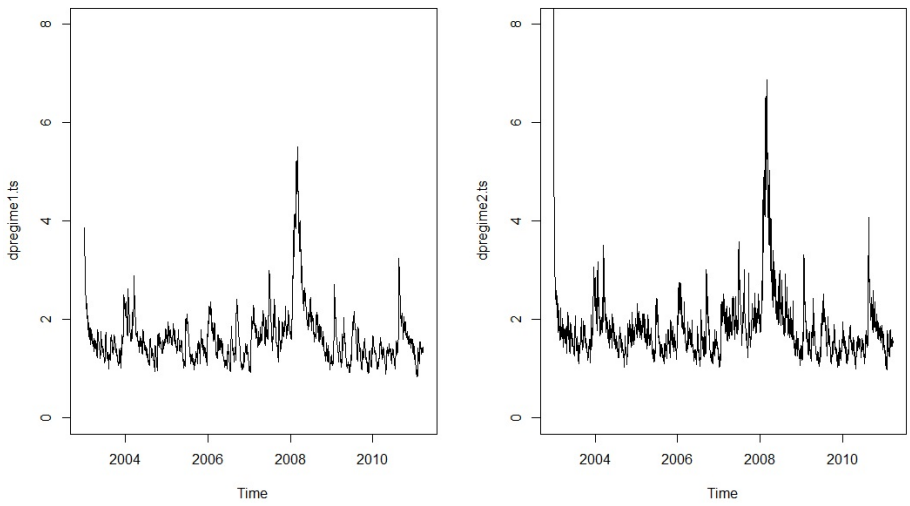

Figura 5.7: *Estimativa do desvio padrão condicional para os retornos do Ibovespa no Regime 1 e 2 com erro t de Student.*

As probabilidades suavizadas do regime 1 com erro t de Student com distintos graus de liberdade e os log-retornos podem ser observados na Figura 5.8:

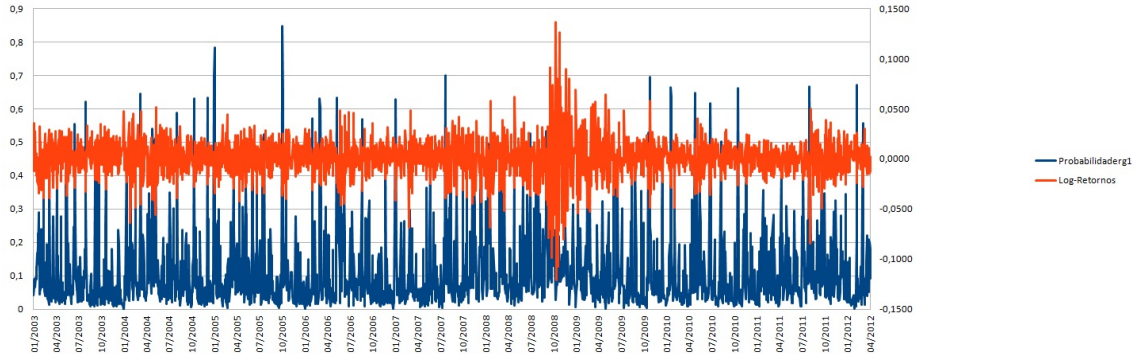

Figura 5.8: *Probabilidade suavizada do Regime 1 - Ibovespa - t2.*

Se o processo estiver no regime 1, a observação será retirada de uma distribuição t de Student com média -1.2336 e desvio padrão < 0.0001 e se estiver no segundo regime com média 0.2425 e desvio padrão < 0.0001. As probabilidades de estarmos no regime 1, variam entre 0 e 0.38 e as probabilidades de estarmos em um regime e continuarmos neste p_{11} e p_{22} são 0.3817 e 0.9243. Com isto temos que a matriz de transição é dada por

$$P = \begin{pmatrix} 0.3817 & 0.0757 \\ 0.6183 & 0.9243 \end{pmatrix}.$$

Temos então que a probabilidade incondicional de estar no k-ésimo regime, neste caso para nosso modelo com 2 regimes é dada por $\pi_1 = 0.1091$ e $\pi_2 = 0.8909$. Ou seja, temos ao sortearmos uma observação qualquer da série, a probabilidade é maior de estarmos no regime 2 do que no regime 1.

A duração média de estarmos no regime 1 é 1 dia e do regime 2 é de 13 dias.

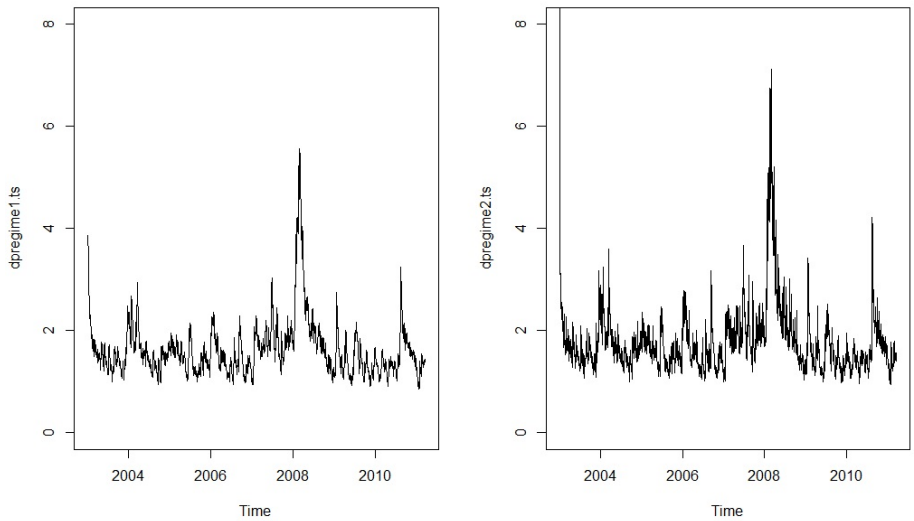

Figura 5.9: *Estimativa do desvio padrão condicional para os retornos do Ibovespa no Regime 1 e 2 com erro t de Student-t2.*

As probabilidades suavizadas do regime 1 com erro GED e os log-retornos podem ser observados no seguinte gráfico:

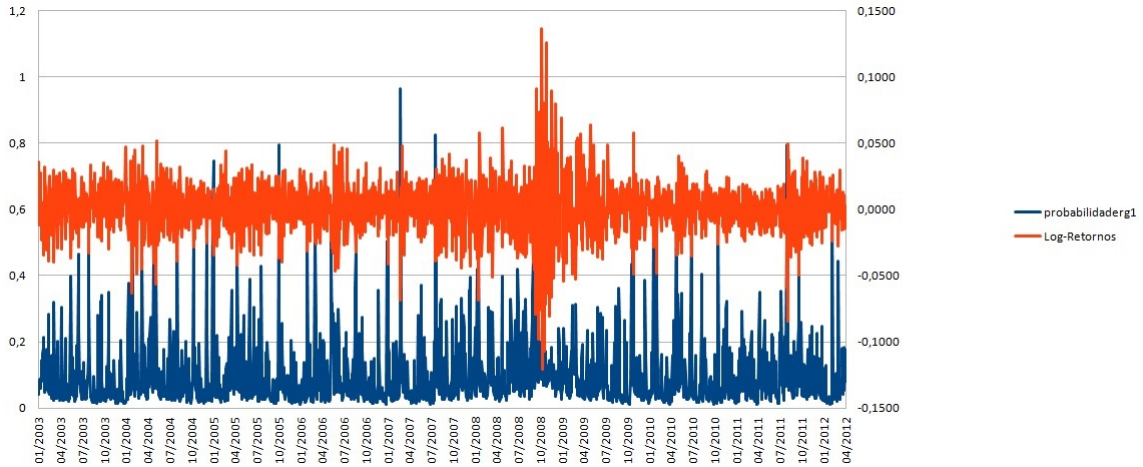

Figura 5.10: *Probabilidade suavizada do Regime 1 - Ibovespa - GED.*

Se o processo estiver no regime 1, a observação será retirada de uma distribuição ged com média 0.0332 e desvio padrão 0.1332 e se estiver no segundo regime com média 0.0419 e desvio padrão 0.0029. A figura 5.10 apresenta a probabilidade suavizada do regime 1 e os log-retornos. As probabilidades de estarmos no regime 1, variam entre 0 e 0.9 e as probabilidades de estarmos em um regime e continuarmos neste p_{11} e p_{22} são 0.9987 e 0.9989 respetivamente, ou seja, os estados seriam absorventes. Com isto temos que a matriz de transição e dada por

$$P = \begin{pmatrix} 0.9987 & 0.0011 \\ 0.0013 & 0.9989 \end{pmatrix}.$$

Temos então que a probabilidade incondicional de estar no k-ésimo regime, neste caso para nosso modelo com 2 regimes é dada por $\pi_1 = 0.4583$ e $\pi_2 = 0.5417$. Ou seja, ao sortearmos uma observação qualquer da série, a probabilidade é só um pouco maior de estarmos no regime 2 do que no regime 1.

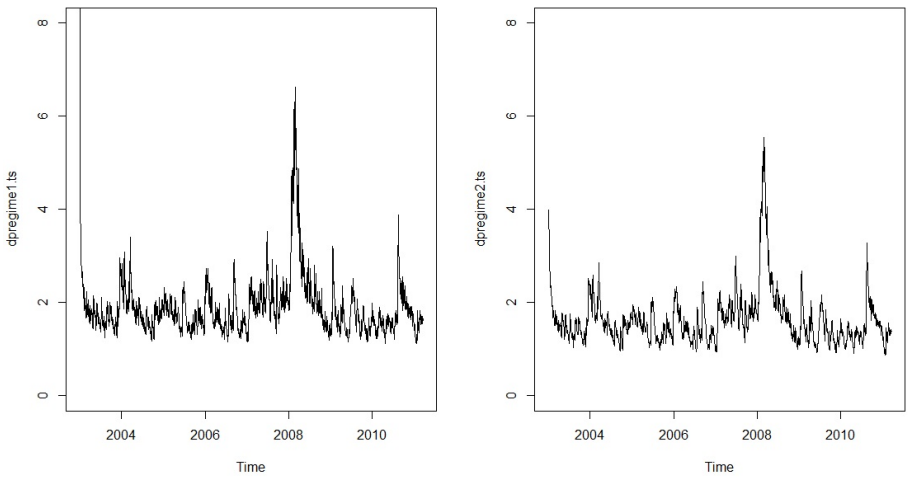

Figura 5.11: *Estimativa do desvio padrão condicional para os retornos do Ibovespa no Regime 1 e 2 com erro GED.*

5.3 Análise de Resíduos-IBOVESPA

O modelo MRS-GARCH com ϵ_t normal, t-Student ou GED foi bem ajustado, se os resíduos padronizados

$$\widetilde{r_t} = r_t/\sqrt{h_t},$$

comportam-se como variáveis aleatorias i.i.d com distribuição normal padrão, t-Student ou GED. Como no caso de modelos GARCH, usualmente supomos que os ϵ_t são normais, ou seguem uma distribuição t-Student, ou ainda, uma distribuição de erro generalizada.

5.3.1 Análise de Resíduos para o modelo MRS-GARCH-N

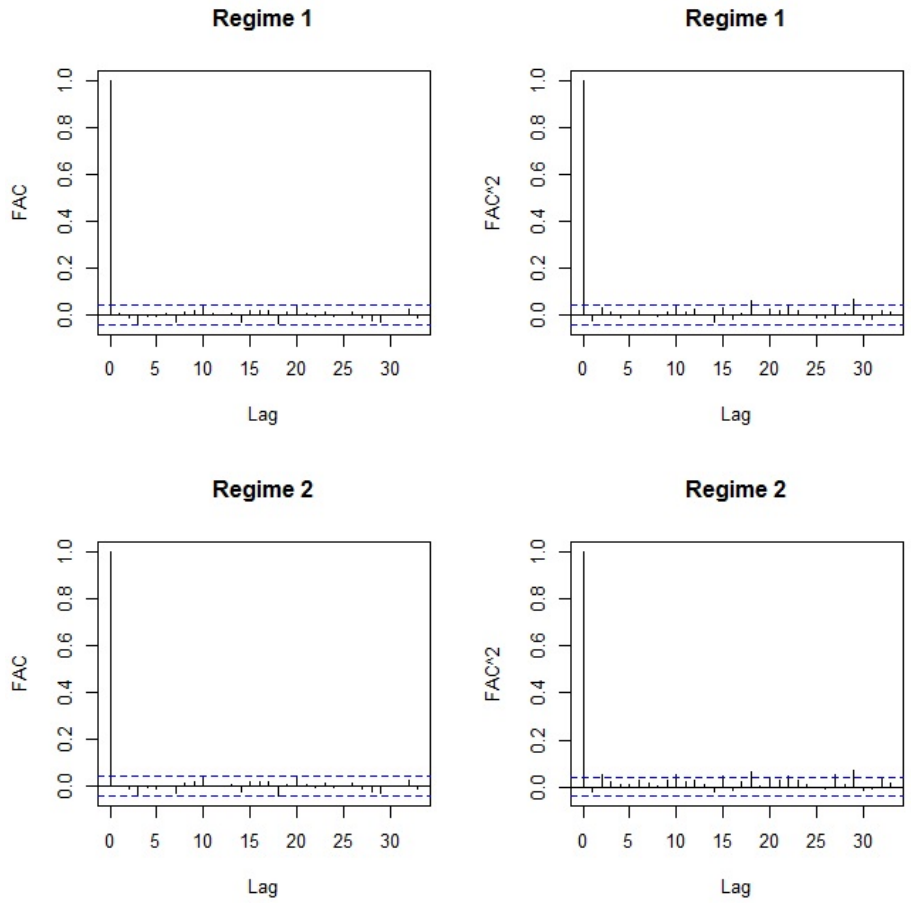

Figura 5.12: *Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo Normal-Ibovespa.*

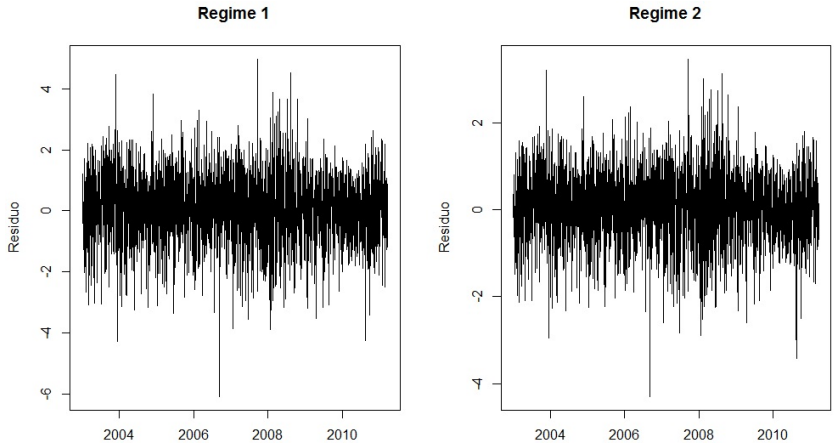

Figura 5.13: *Resíduos padronizados nos Regimes 1 e 2 do modelo Normal Ibovespa.*

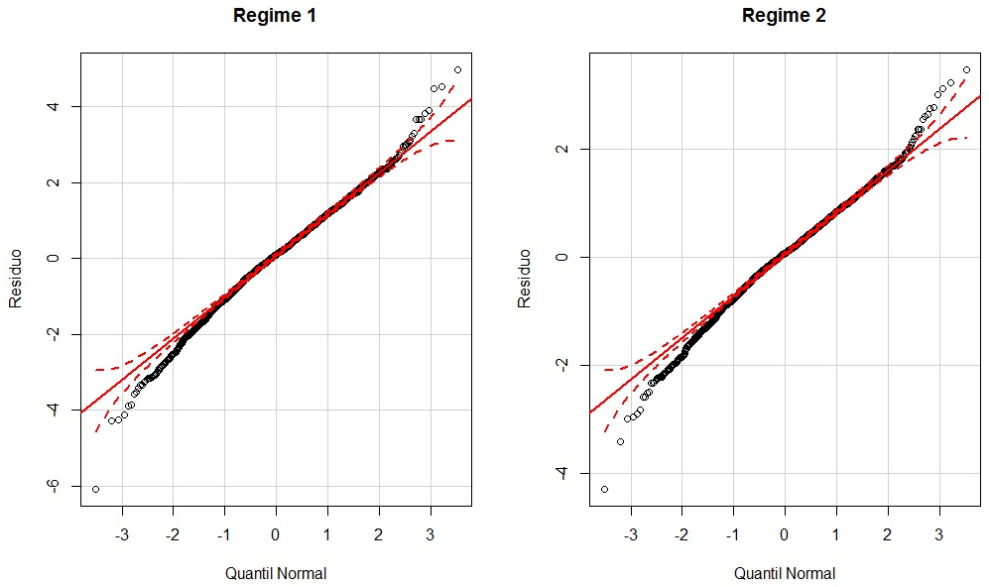

Figura 5.14: *Gráfico de quantis dos Resíduos padronizados do modelo Normal-Ibovespa.*

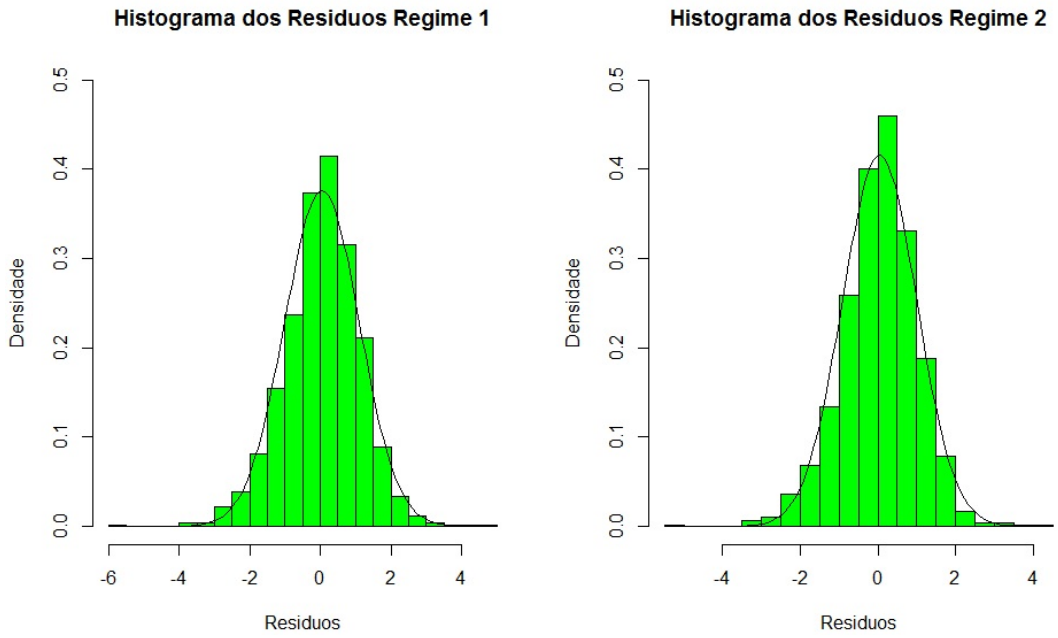

Figura 5.15: *Histograma dos resíduos padronizados do modelo Normal-Ibovespa.*

5.3.2 Análise de Resíduos para o modelo MRS-GARCH-t2

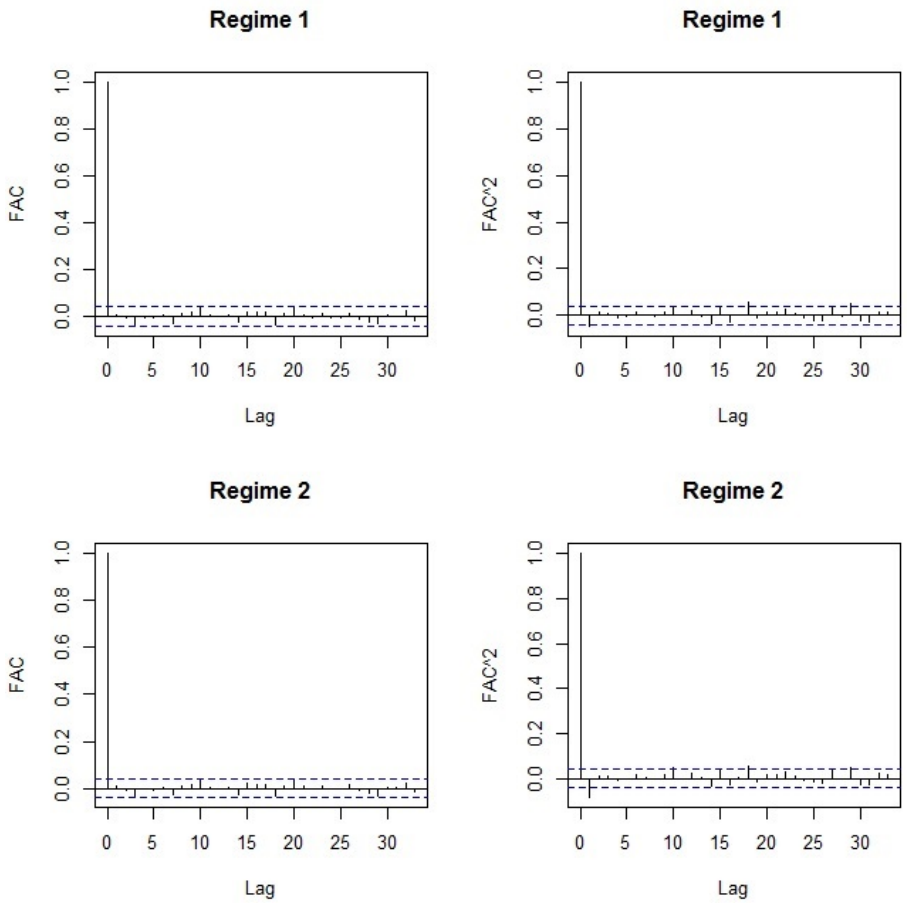

Figura 5.16: *Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo t2-Ibovespa.*

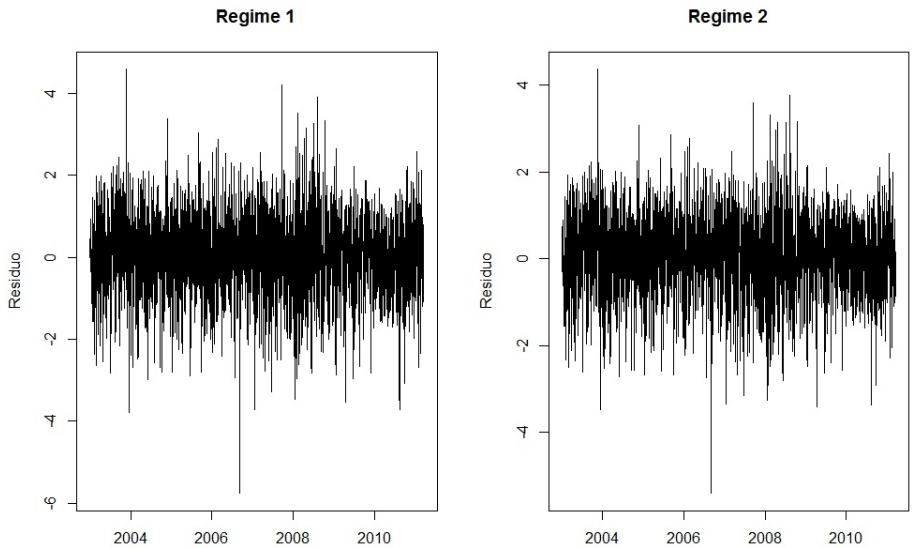

Figura 5.17: *Resíduos padronizados nos Regimes 1 e 2 do modelo t2-Ibovespa.*

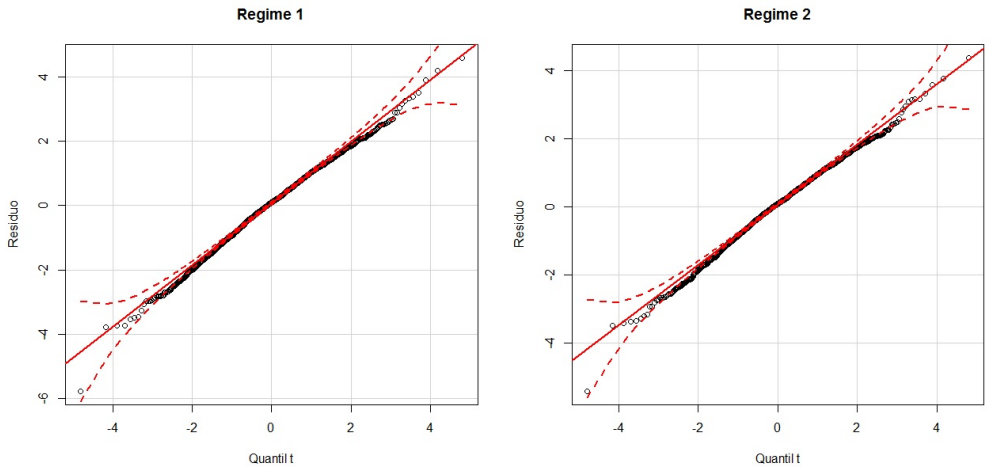

Figura 5.18: *Gráfico de quantis dos Resíduos padronizados do modelo t2-Ibovespa.*

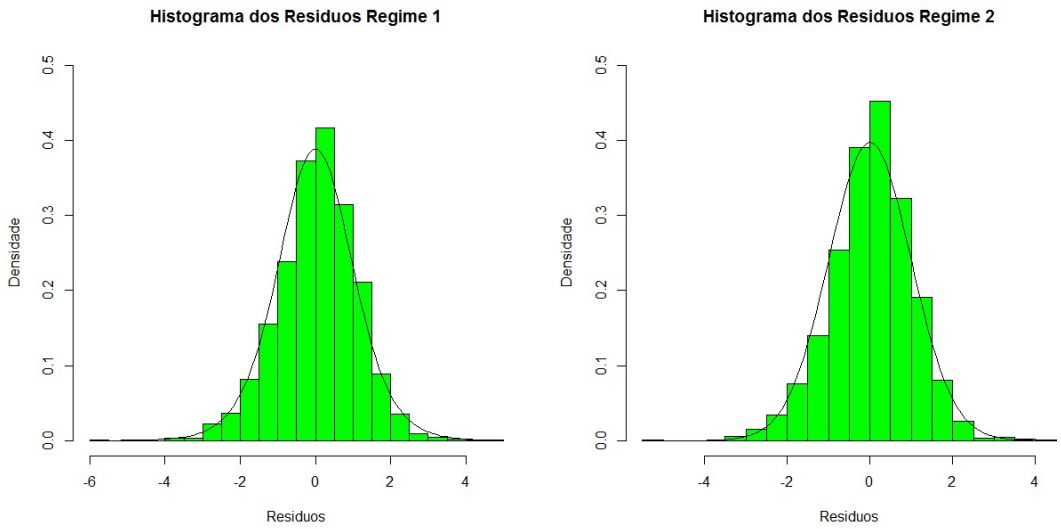

Figura 5.19: *Histograma dos Resíduos padronizados do modelo t2-Ibovespa.*

5.3.3 Análise de Resíduos para o modelo MRS-GARCH-t

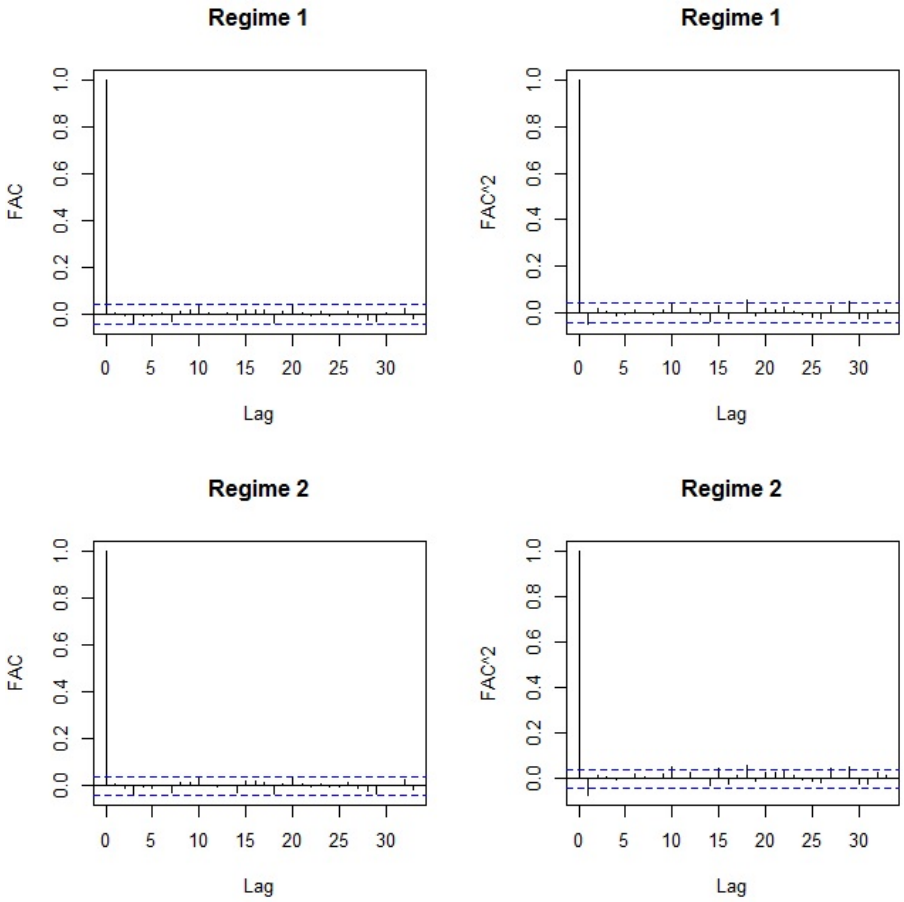

Figura 5.20: *Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo t-Ibovespa.*

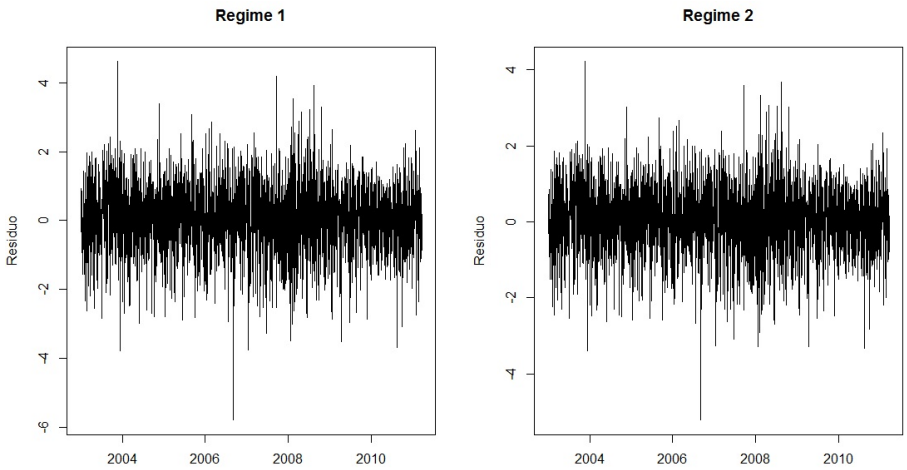

Figura 5.21: *Resíduos padronizados nos Regimes 1 e 2 do modelo t-Ibovespa.*

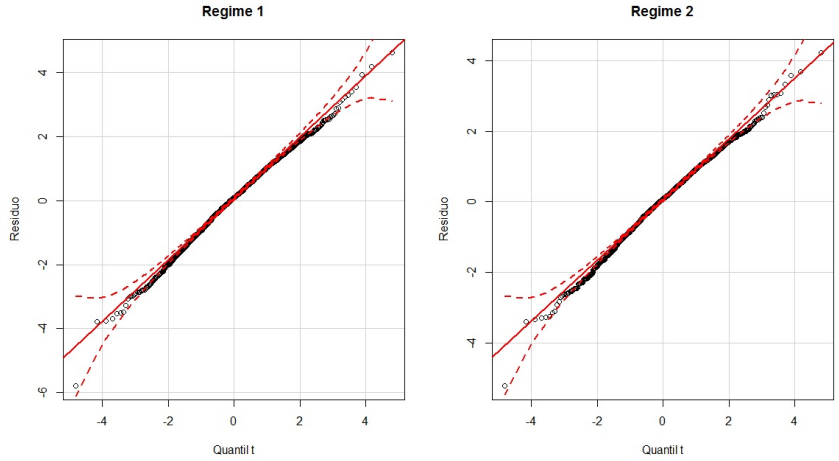

Figura 5.22: *Gráfico de quantis dos Resíduos padronizados do modelo t-Ibovespa.*

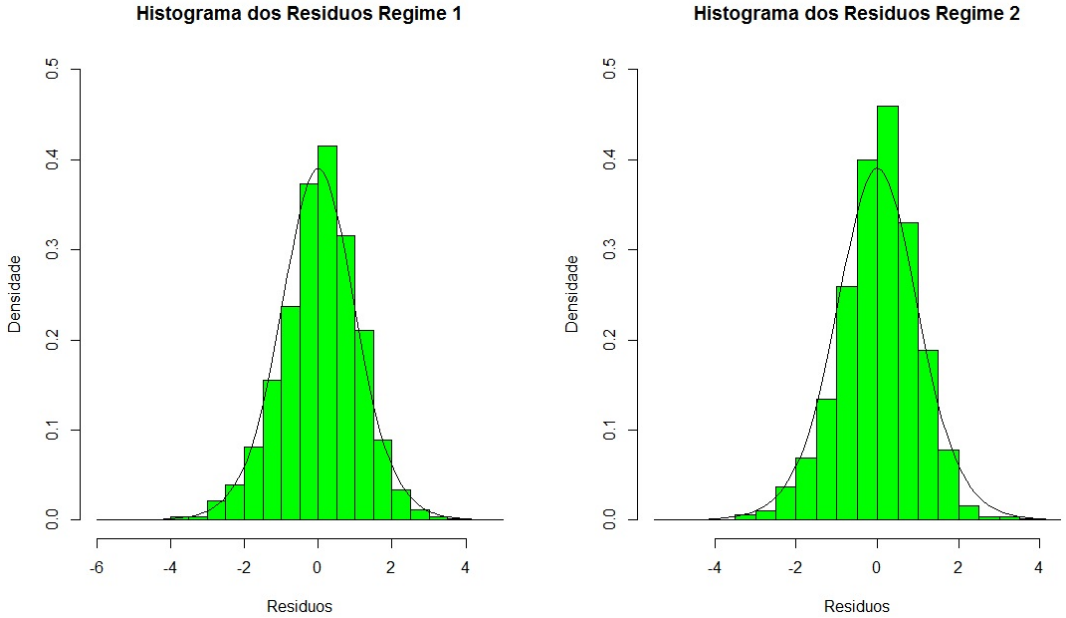

Figura 5.23: *Histograma dos Resíduos padronizados do modelo t-Ibovespa.*

5.3.4 Análise de Resíduos para o modelo MRS-GARCH-GED

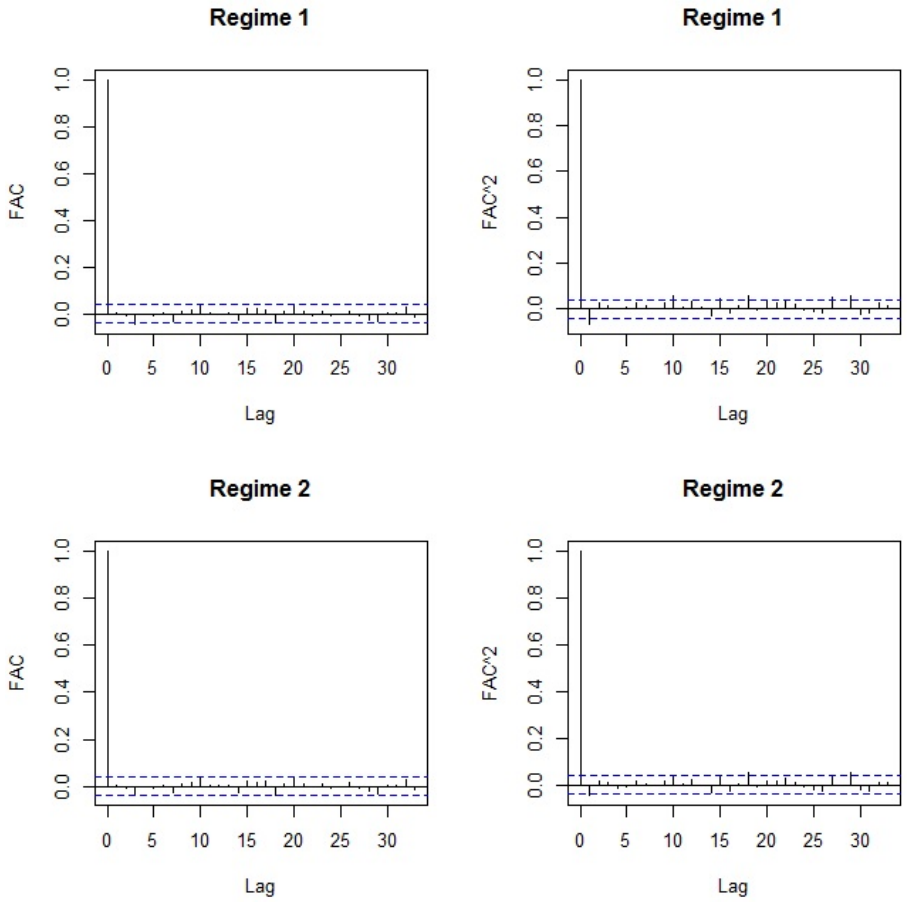

Figura 5.24: *Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo GED-Ibovespa.*

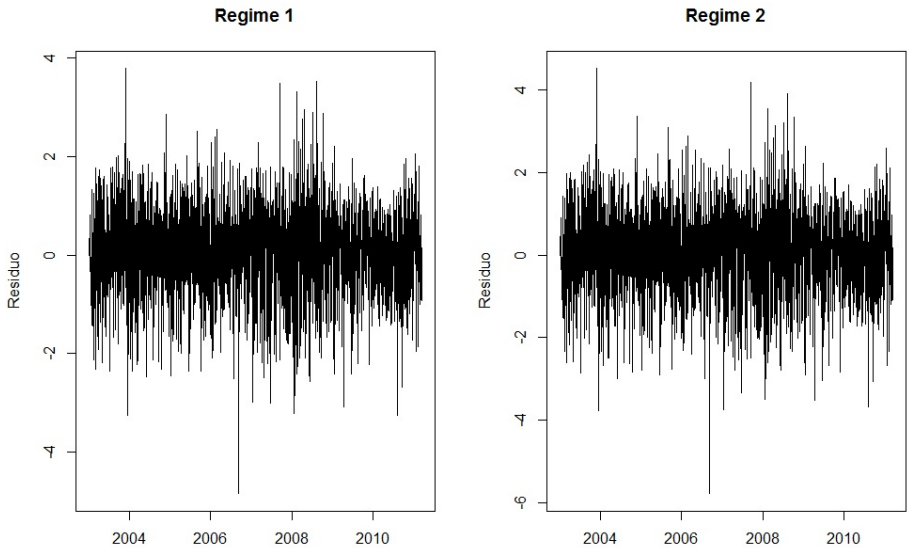

Figura 5.25: *Resíduos padronizados nos Regimes 1 e 2 do modelo GED-Ibovespa.*

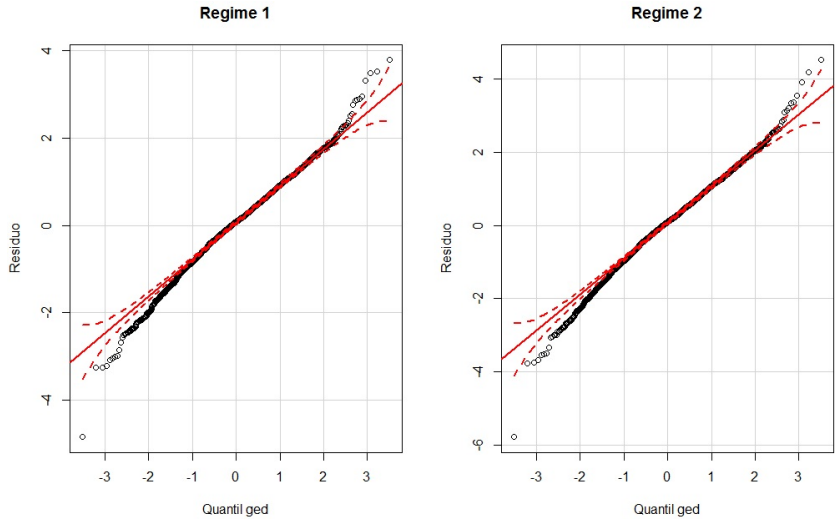

Figura 5.26: *Gráfico de quantis dos Resíduos padronizados do modelo GED-Ibovespa.*

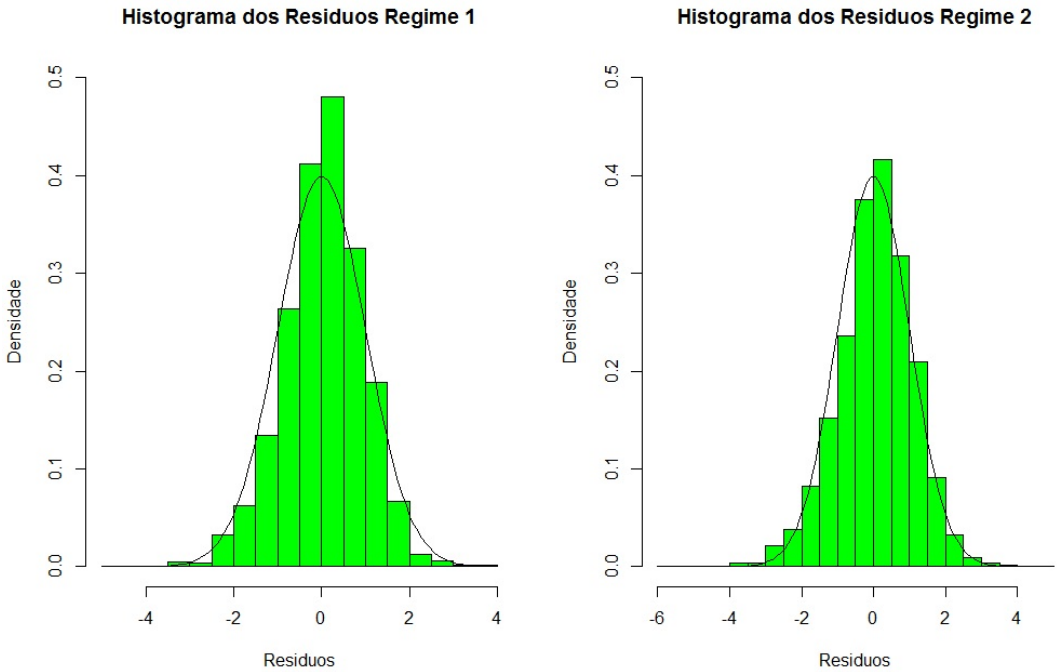

Figura 5.27: *Histograma dos Resíduos padronizados do modelo GED-Ibovespa.*

A análise de resíduos tem o objetivo de verificar a validade das suposições dos modelos de forma geral. Assim, a escolha do melhor modelo depende dessa análise de resíduos, utilizando os resíduos do regime 1 e 2. Os modelos supõem que os erros são não correlacionados e seguem a distribuição especificada.

Portanto, verificamos se os resíduos padronizados são não correlacionados, se os resíduos padronizados ao quadrado também são não correlacionados e se seguem a distribuição normal, t de Student ou GED dependendo do modelo. Os gráficos de quantis indicam que somente os modelos com distribuição t de Student com distintos graus de liberdade são adequados e por isso esses modelos foram escolhidos para os log retornos do Ibovespa. A função de autocorrelação parcial dos resíduos padronizados não indicam nenhuma quebra de comportamento de ruído branco nos resíduos.

Em seguida, avaliaremos o desempenho dos modelos com base na cobertura dos intervalos de confiança para os log retornos.

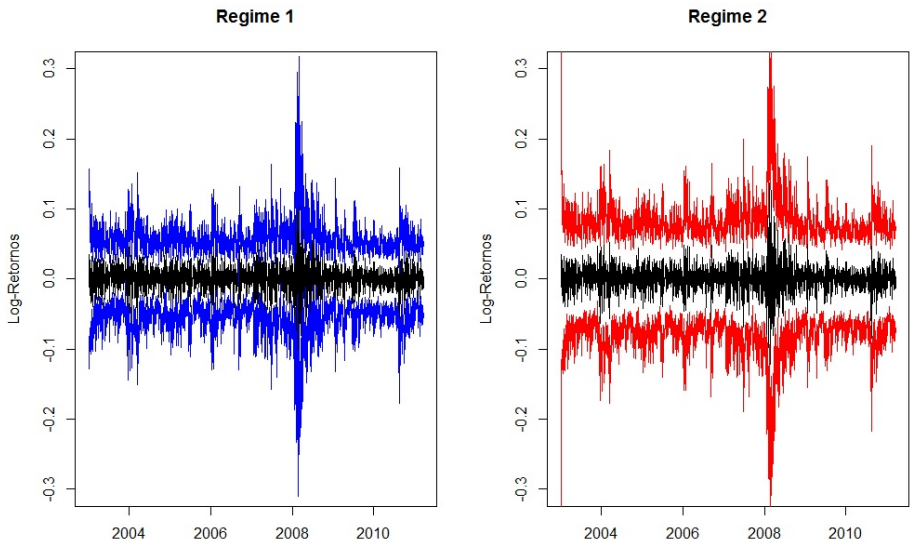

Figura 5.28: *Retornos e Intervalos de confiança (95%)-Normal-Ibovespa.*

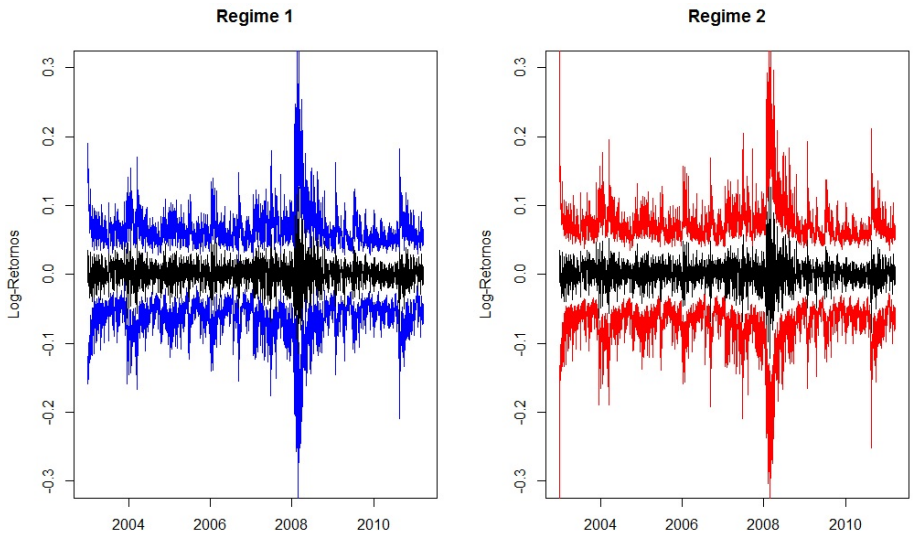

Figura 5.29: *Retornos e Intervalos de confiança (95%)-t-Ibovespa.*

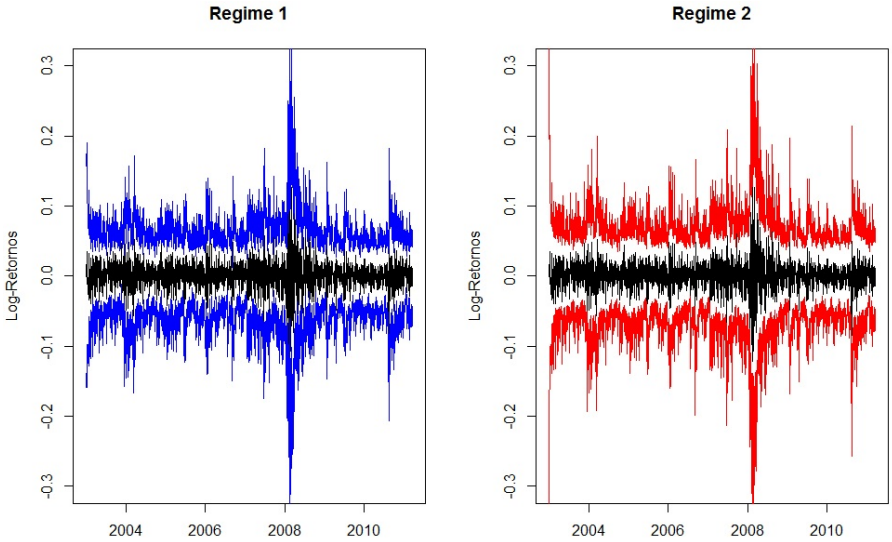

Figura 5.30: *Retornos e Intervalos de confiança (95%)-t2-Ibovespa.*

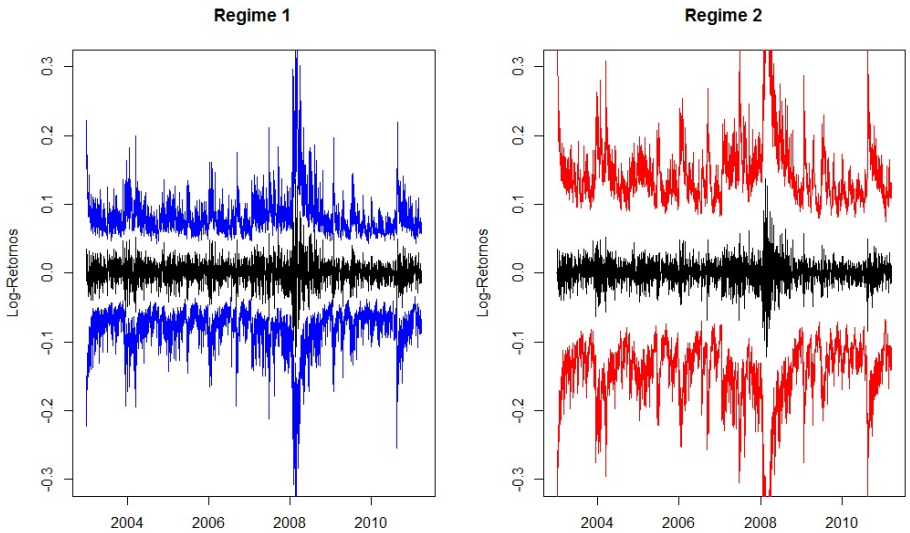

Figura 5.31: *Retornos e Intervalos de confiança (95%)-GED-Ibovespa.*

Tabela 5.3: *Porcentagem de Cobertura do intervalos de confiança para os log-retornos do Ibovespa*

	Cobertura (%)	Comp. Médio	Mínimo	Máximo
GARCH(1,1)-N	95	6.79	4.24	25.46
2R.GARCH(1,1)-N	98	6.23	3.81	25.32
GARCH(1,1)-t	98	8.69	5.40	32.40
2R.GARCH(1,1)-t	99	7.24	3.74	24.45
GARCH(1,1)-t	98	8.69	5.40	32.40
2R.GARCH(1,1)-t2	99	7.25	3.80	24.70
GARCH(1,1)-GED	97	7.80	5.10	28.81
2R.GARCH(1,1)-GED	98	7.64	4.55	35.70

Na Tabela 5.3 mostra-se o modelo Garch usual como GARCH(1,1) para cada distribuição, e 2R.GARCH(1,1) para o modelo Garch com dois regimes.

Os intervalos de confiança foram obtidos da mesma forma que são calculados para a média de uma população com variância conhecida com o respetivo quantil para cada distribuição especificada, utilizando as estimativas da média e da variância em cada regime. Aqui os intervalos de confiança foram calculados para os log-retornos do Ibovespa.
A cobertura foi calculada utilizando o valor do log-retorno maior do que o limite inferior de confiança para observar se ele se encontra dento do intervalo ou não. A porcentagem de cobertura é simplesmente calculada como a média desses valores. Os intervalos apresentaram cobertura maior que 95% e as maiores coberturas foram para as distribuições t de Student com distintos graus de liberdade. O comprimento foi calculado fazendo a diferença entre o limite superior de confiança com o limite inferior de confiança e o comprimento médio é simplesmente a média do comprimento. Aquí o Mínimo é o valor mínimo no comprimento e o máximo é o valor máximo do comprimento. Pode-se observar que os valores para os modelos com 2 regimes tem menor comprimento médio do que os modelos com 1 regime.

5.4 S&P500

O conjunto de dados analisado neste trabalho é o índice S&P500. O período considerado é de 2 de janeiro de 2003 a 9 de abril de 2012 para um total de 2333 observações. Os dados foram obtidos a partir do site yahoo finanças. A cotação estudada foi o preço de fechamento diário do índice.

Primeiramente, apresentamos nas Figuras 5.32 e 5.33, a série S&P500, a série de retornos, o gráfico de quantis e histograma dos retornos.

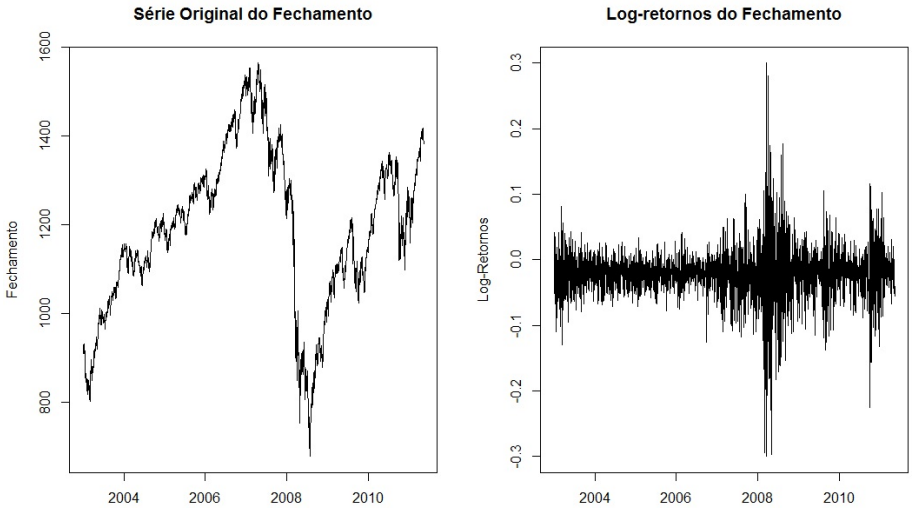

Figura 5.32: *Gráfico da Série S&P500 e os Log-retornos.*

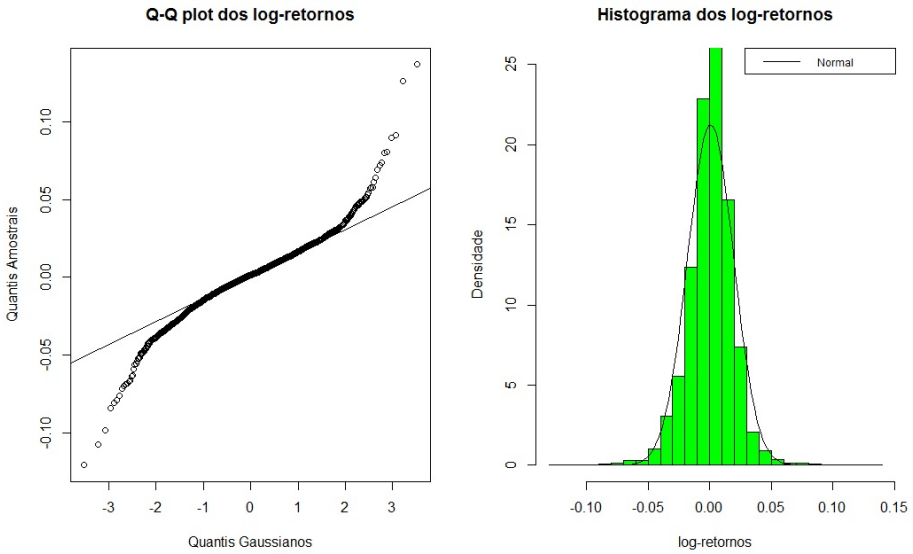

Figura 5.33: *Gráfico de quantis e Histograma dos log-retornos da Série S&P500.*

A Tabela 5.4 mostra algumas estatísticas descritivas dos log-retornos do Índice S&P500. Note que a média dos dados está próxima de zero e o desvio padrão é pequeno, mas os log-retornos variam de -0.09 a 0.10. A média é próxima de zero e, com o desvio padrão pequeno, podemos concluir que os dados se encontram aglomerados em torno do zero, o que pode ser confirmado pelo histograma dos retornos diários do S&P500. Observando o gráfico da série diária do S&P500, podemos ver a que a série é não estacionária, pois apresenta tendência não constante e variabilidade não controlada. Já o gráfico dos retornos diários do S&P500, apresenta nível, aparentemente, constante e variabilidade controlada. A assimetria e a curtose indicam que os dados não seguem uma distribuição normal. Com base nessas medidas resumo dos retornos diários da S&P500, podemos observar que: a distribuição dos retornos possui uma leve assimetria à esquerda, já que o valor do coeficiente de assimetria é

Tabela 5.4: *Medidas descritivas do retorno-S&P500*

Média	Mediana	D.Padrão	Curtose	Mínimo	Máximo	Assimetria
0.00017	0.00081	0.01341	9.95099	-0.09451	0.10950	-0.29579

não nulo e negativo e a medida de curtose da distribuição dos retornos indica que sua curva é mais achatada que a curva da distribuição Normal, ou seja, há uma maior quantidade de dados nas caudas da curva da distribuição dos retornos do que em uma curva da distribuição Normal. A assimetria e a curtose podem ser visualizadas no gráfico de quantis e histograma da Figura 5.33.

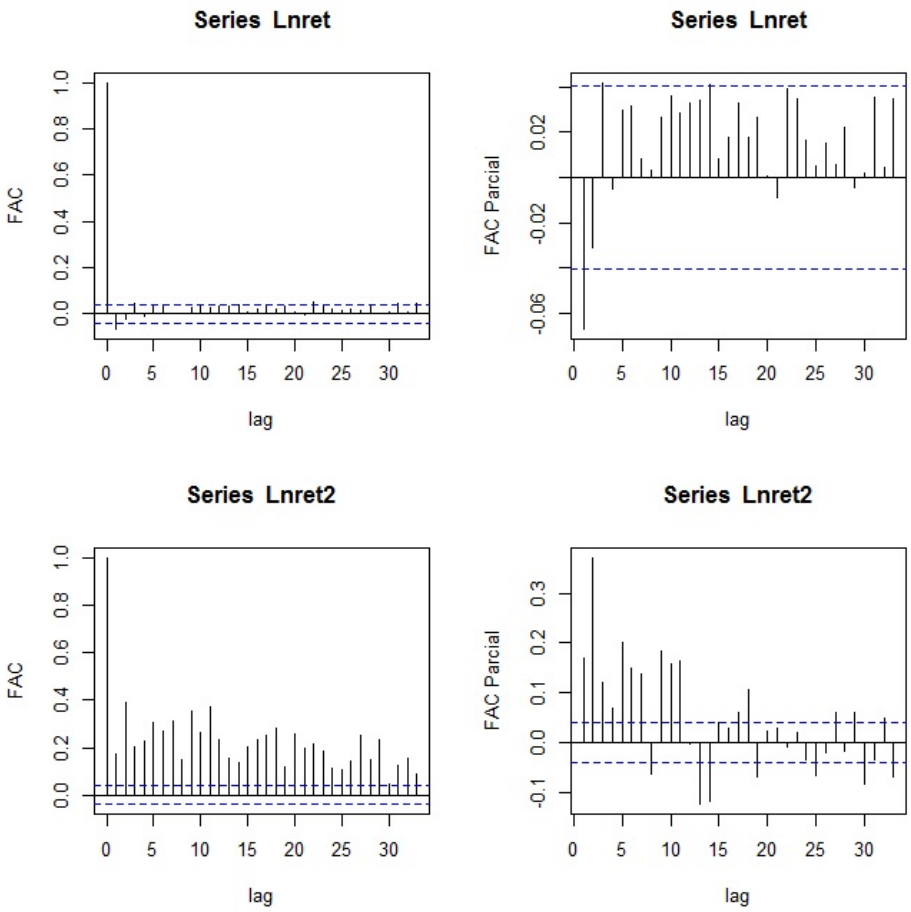

Figura 5.34: *Função de autocorrelação dos log-retornos e log-retornos ao quadrado-S&P500.*

Entretanto, na Figura 5.34 observa-se que as autocorrelações dos retornos são não significativas e as autocorrelações dos retornos ao quadrado são significativas. Conforme descrito no Capítulo 3, este comportamento indica que a série de retornos pode mostrar presença de heterocedasticidade condicional nos dados, una vez que os retornos ao quadrado medem o momento de segunda ordem da série original, este resultado indica que a variância dos retornos da série da S&P500 condicionada ao passado histórico pode sofrer alterações ao longo do tempo.

As estimativas dos parâmetros do modelo MRS-GARCH são apresentadas na Tabela 5.5. Foram ajustados os modelos com distribuição Normal, t de Student e a distribuição generalizada do erro e para o modelo com distribuição t de Student foi considerado o modelo em que o grau de liberdade pode ser diferente nos 2 regimes (t2). Em particular, a Tabela 5.5 mostra o desvio padrão condicional dos retornos para cada regime da volatilidade, isto é

$$\sigma^{(i)} = \left(a_0^{(i)} / \left(1 - a_1^{(i)} - b_1^{(i)}\right)\right)^{1/2}.$$

Para o modelo t de Student (t2), o parâmetro da média é significativo para o regime 1 e não significativo para o regime 2, no caso da distribuição normal e GED os parâmetros da média são negativos para o regime 1. O parâmetro a_1, que mede o efeito do erro anterior na variância, parece ser significativo para os três modelos considerados.

Para o regime 1, as estimativas de a_1 são maiores do que no regime 2 que são quase zero para o modelo normal e GED. Entretanto, os valores para os parâmetros b_1 são maiores para o regime 1 no modelo normal e GED, já para o modelo t de Student esse valor é menor o que significa que temos mais peso para cada erro de ontem e menor inercia i.e para a variância anterior σ_{t-1}^2. A probabilidade de se manter no regime 1, p_{11} é menor (49% para a distribuição t2), que a probabilidade de se manter no regime 2, o mesmo acontece para a distribuição normal e GED.

A Tabela 5.5 também mostra as probabilidades incondicionais de cada modelo MRS-GARCH. A probabilidade incondicional π_1 de estar no primeiro regime varia entre 27% para o modelo normal e 47% para o modelo t de Student com 2 graus de liberdade. Para o modelo t-Student com 2 graus de liberdade no primeiro regime, o parâmetro ν indica que a distribuição tem caudas mais pesadas do que a normal, já para o segundo regime, o valor do parâmetro ν indica que a distribuição é aproximadamente normal. No caso do GED o valor do parâmetro ν está indicando que a distribuição tem caudas mais pesadas do que o normal.
A identificação de regime com caudas mais pesadas no regime 2 da distribuição t de Student nos leva a escolher o modelo t2.

Nesta aplicação pode-se percerber que não encontram-se as estimativas para o modelo t de Student, pois no processo de otimização os valores estavam convergindo para números complexos e não estava conseguindo achar o mínimo local, logo não foi considerada esta distribuição para esse estudo.

Tabela 5.5: *Estimativas de Máxima Verossimilhança e erros padrões do Modelo MRS - GARCH - S&P500*

	MRS-GARCH-N	MRS-GARCH-t2	MRS-GARCH-GED
$\mu^{(1)}$	-0.3116	0.1651	-0.2380
	(0.0828)	(0.0248)	(0.0976)
$\mu^{(2)}$	0.1021	0.0070	0.1111
	(0.0204)	(0.0289)	(0.0210)
$\sigma^{(1)}$	0.1338	0.0256	0.1370
	(0.0334)	(0.1195)	(0.0469)
$\sigma^{(2)}$	0.0247	0.0199	0.0224
	(0.0056)	(0.0567)	(0.0071)
$a_1^{(1)}$	0.0738	0.2504	0.0761
	(0.0240)	(0.0513)	(0.0337)
$a_1^{(2)}$	< 0.0001	0.0248	< 0.0001
	(0.0185)	(0.0281)	(0.0218)
$b_1^{(1)}$	0.9238	0.7108	0.9213
	(0.0303)	(0.1419)	(0.0421)
$b_1^{(2)}$	0.8980	0.9631	0.9018
	(0.0178)	(0.0808)	(0.0206)
p_{11}	0.9046	0.4932	0.9088
	(0.0218)	(0.1831)	(0.0267)
p_{22}	0.9631	0.5391	0.9620
	(0.0084)	(0.1005)	(0.0107)
$\nu^{(1)}$	——	2.2599	1.5120
		(0.0893)	(0.0739)
$\nu^{(2)}$	——	172.5236	——
		< 0.0001	
$1000 * Log(L)$	-3.2836	-3.2561	-3.2663
π_1	0.2789	0.4762	0.2941
π_2	0.7211	0.5238	0.7059
ρ_1	0.9976	0.9612	0.9280
ρ_2	0.8980	0.9879	0.9927

As probabilidades suavizadas do regime 1 com erro gaussiano e os log-retornos podem ser observados no seguinte gráfico:

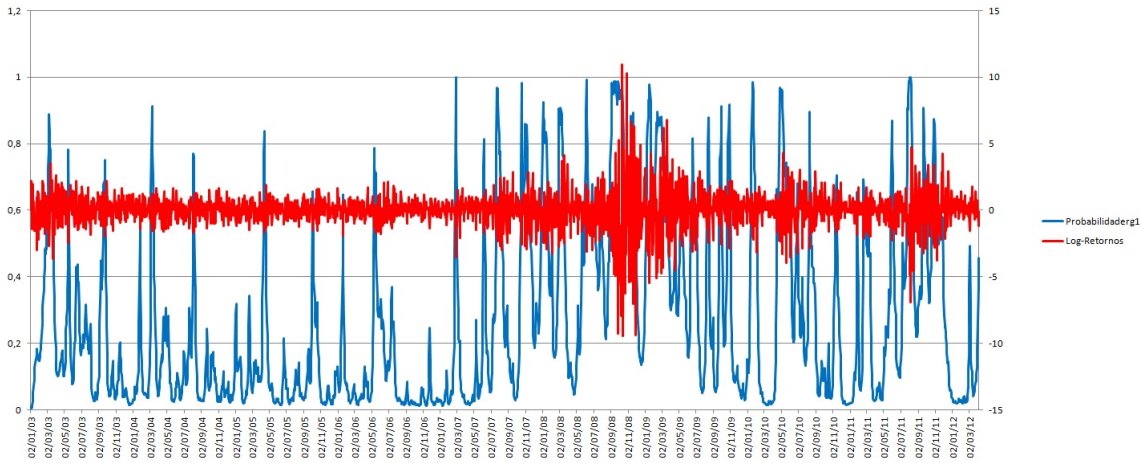

Figura 5.35: *Probabilidade suavizada do Regime 1 - Normal - S&P500.*

Em nossa aplicação existem dois regimes, $s_t = 1$ e $s_t = 2$, ou seja, se o processo estiver no regime 1, a observação será retirada de uma distribuição normal com média -0.3116 e desvio padrão 0.1338 e se estiver no segundo regime com média 0.1021 e desvio padrão 0.0247. A Figura 5.35 apresenta a probabilidade suavizada do regime 1 e os log-retonos. As probabilidades de estarmos no regime 1, variam entre 0 e 0.9, parece haver períodos com menor volatilidade como o ínicio da série até 2007. As probabilidades de estarmos em um regime e continuarmos neste são p_{11} e p_{22} iguais a 0.9046 e 0.9631 respetivamente. Com isto temos que a matriz de transição e dada por

$$P = \begin{pmatrix} 0.9046 & 0.0369 \\ 0.0954 & 0.9631 \end{pmatrix}.$$

Temos então que a probabilidade incondicional de estar no k-ésimo regime, neste caso para nosso modelo com 2 regimes é dada por $\pi_1 = 0.2789$ e $\pi_2 = 0.7211$. Ou seja, ao sortearmos uma observação qualquer da série, a probabilidade é maior de estarmos no regime 2 do que no regime 1.

A duração média de estarmos no regime 1 é 10 dias e do regime 2 é 27 dias.

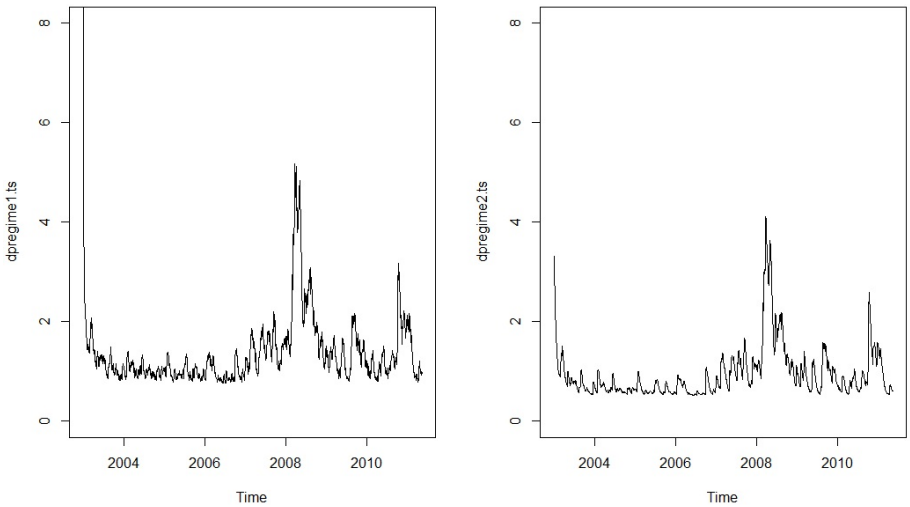

Figura 5.36: *Estimativa do desvio padrão condicional para os retornos da S&P500 no Regime 1 e 2 com erro gaussiano.*

As probabilidades suavizadas do regime 1 com erro t de Student distintos graus de liberdade e os log-retornos podem ser observados no seguinte gráfico:

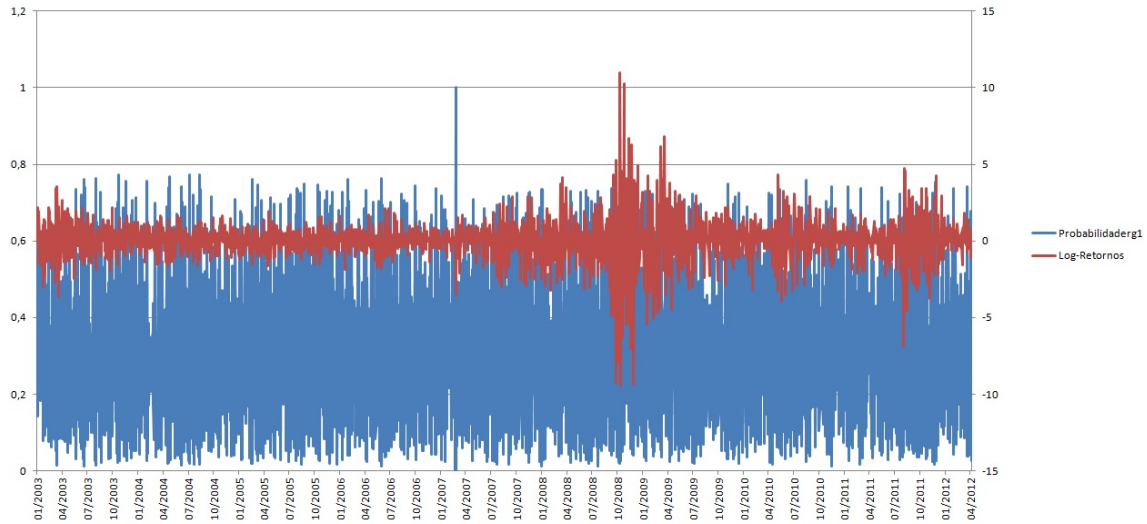

Figura 5.37: *Probabilidade suavizada do Regime 1- t2 - S&P500.*

Se o processo estiver no regime 1, a observação será retirada de uma distribuição t de Student com média 0.1651 e desvio padrão 0.0256 e se estiver no segundo regime com média 0.0070 e desvio padrão 0.0199. A Figura 5.37 apresenta a probabilidade suavizada do regime 1 e os log-retornos. As probabilidades de estarmos no regime 1 variam entre 0 e 0.49, as probabilidades neste caso oscilam bastante e as probabilidades de estarmos em um regime e continuarmos neste p_{11} e p_{22} são 0.4932 e 0.5391, respetivamente. Com isto temos que a matriz de transição e dada por

$$P = \begin{pmatrix} 0.4932 & 0.4609 \\ 0.5068 & 0.5391 \end{pmatrix}.$$

Temos então que a probabilidade incondicional de estar no k-ésimo regime, neste caso para nosso modelo com 2 regimes é dada por $\pi_1 = 0.4762$ e $\pi_2 = 0.5238$. Ou seja, ao sortearmos uma observação qualquer da série, a probabilidade é maior de estarmos no regime 2 do que no regime 1.

A duração média de estarmos no regime 1 é 1 dia e do regime 2 é de 2 dias.

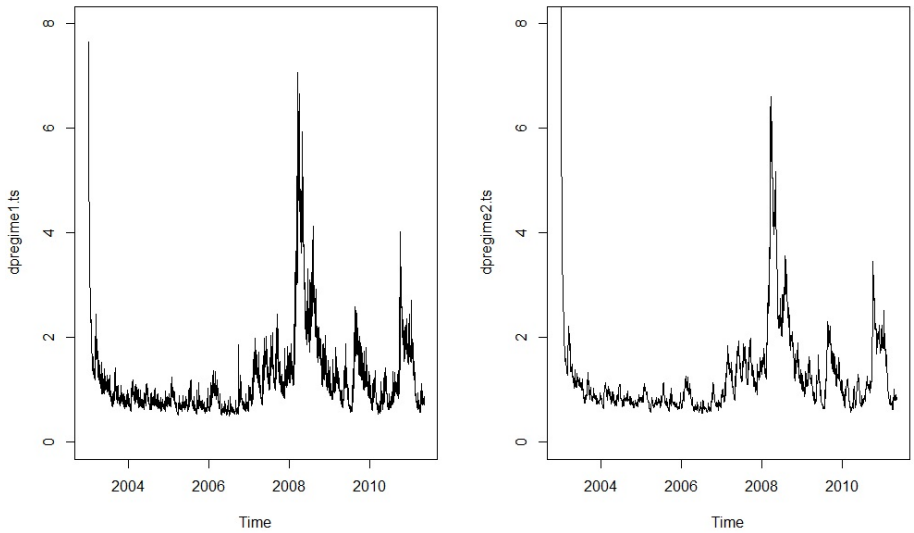

Figura 5.38: *Estimativa do desvio padrão condicional para os retornos da S&P500 no Regime 1 e 2 com erro t de Student.*

As probabilidades suavizadas do regime 1 com erro GED e os log-retornos podem ser observados no seguinte gráfico:

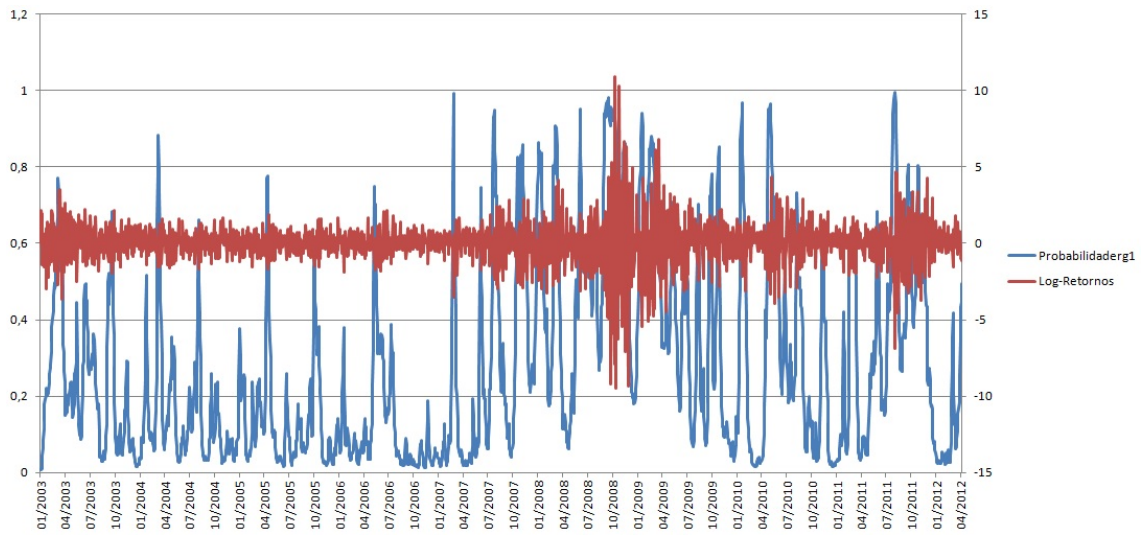

Figura 5.39: *Probabilidade suavizada do Regime 1 - GED - S&P500.*

Se o processo estiver no regime 1, a observação será retirada de uma distribuição ged com média -0.2380 e desvio padrão 0.1370 e se estiver no segundo regime com média 0.1111 e desvio padrão 0.0224. A figura 5.39 apresenta a probabilidade suavizada do regime 1 e os log-retornos. As probabilidades de estarmos no regime 1, variam entre 0 e 0.9, parece haver períodos com menor volatidade como o ínicio da série até 2007, pode-se perceber que o gráfico das probabilidades suavizadas 5.35 e 5.39 tem uma performance semelhante. As probabilidades de estarmos em um regime e continuarmos neste p_{11} e p_{22} são 0.9088 e 0.9620 respetivamente. Com isto temos que a matriz de transição e dada por

$$P = \begin{pmatrix} 0.9088 & 0.0380 \\ 0.0912 & 0.9620 \end{pmatrix}.$$

Temos então que a probabilidade incondicional de estar no k-ésimo regime, neste caso para nosso modelo com 2 regimes é dada por $\pi_1 = 0.2941$ e $\pi_2 = 0.7059$. Ou seja, ao sortearmos uma observação qualquer da série, a probabilidade é maior de estarmos no regime 2 do que no regime 1.

A duração média de estarmos no regime 1 é 10 dias e do regime 2 é de 26 dias.

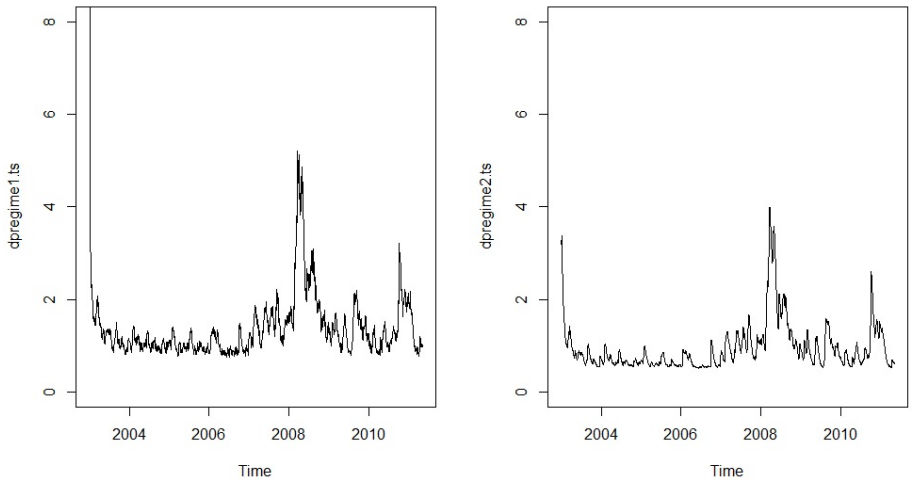

Figura 5.40: *Estimativa do desvio padrão condicional para os retornos da S&P500 no Regime 1 e 2 com erro GED.*

5.5 Análise de Resíduos-S&P500

O modelo MRS-GARCH com ϵ_t normal, t-Student ou GED foi bem ajustado, se os resíduos padronizados

$$\widetilde{r}_t = r_t/\sqrt{h_t},$$

comportam-se como variáveis aleatorias i.i.d com distribuição normal padrão, t-Student ou GED. Como no caso de modelos GARCH, usualmente supomos que os ϵ_t são normais, ou seguem uma distribuição t-Student, ou ainda, uma distribuição de erro generalizada.

5.5.1 Análise de Resíduos para o modelo MRS-GARCH-N

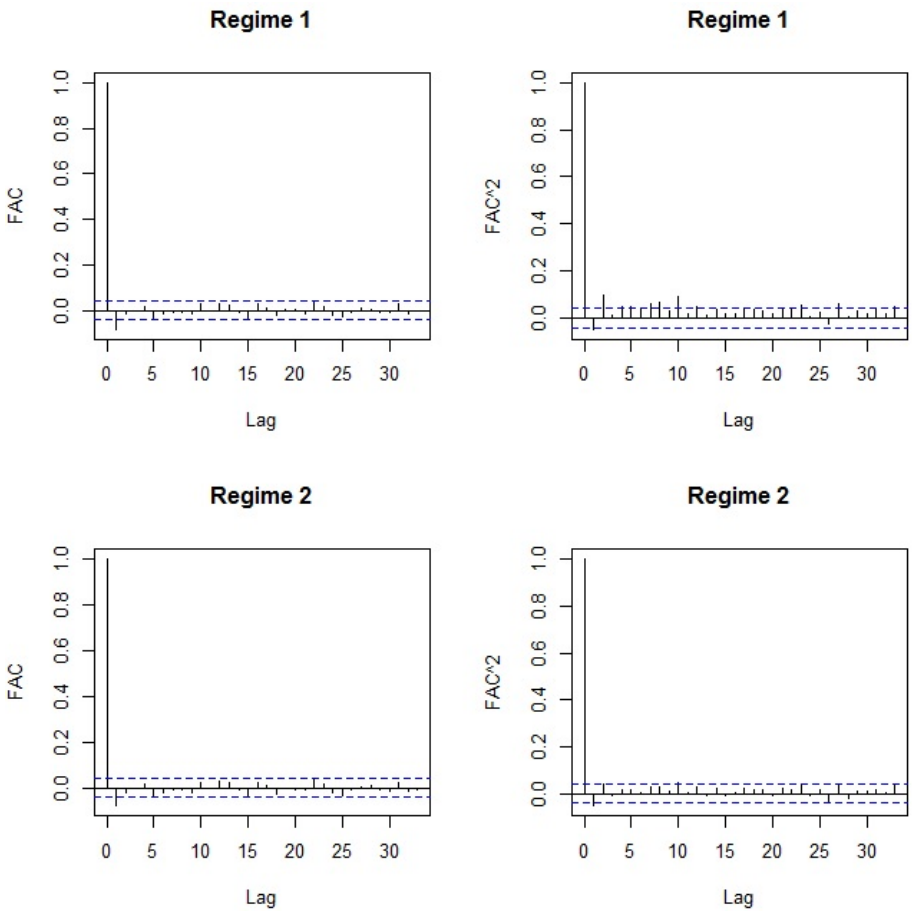

Figura 5.41: *Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo Normal-S&P500.*

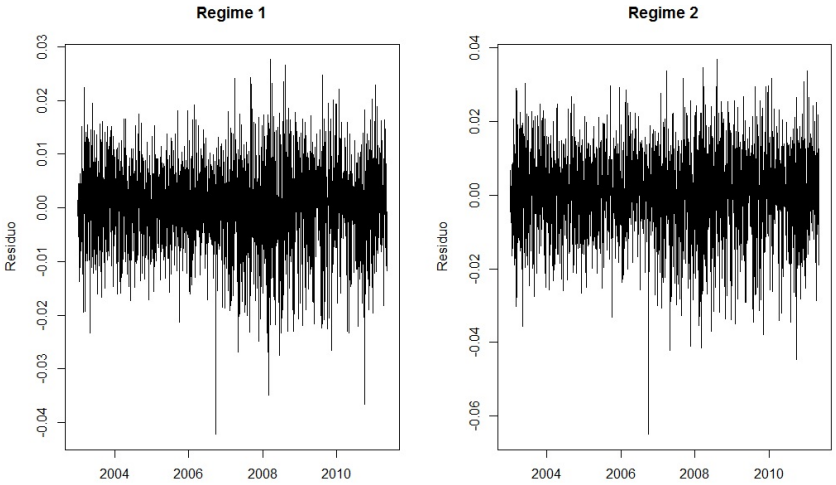

Figura 5.42: *Resíduos padronizados nos Regimes 1 e 2 do modelo Normal-S&P500.*

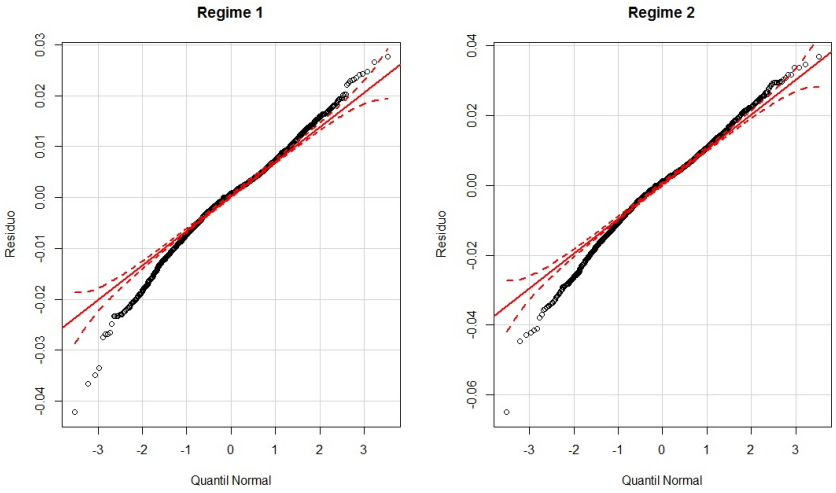

Figura 5.43: *Gráfico de quantis dos Resíduos padronizados do modelo Normal-S&P500.*

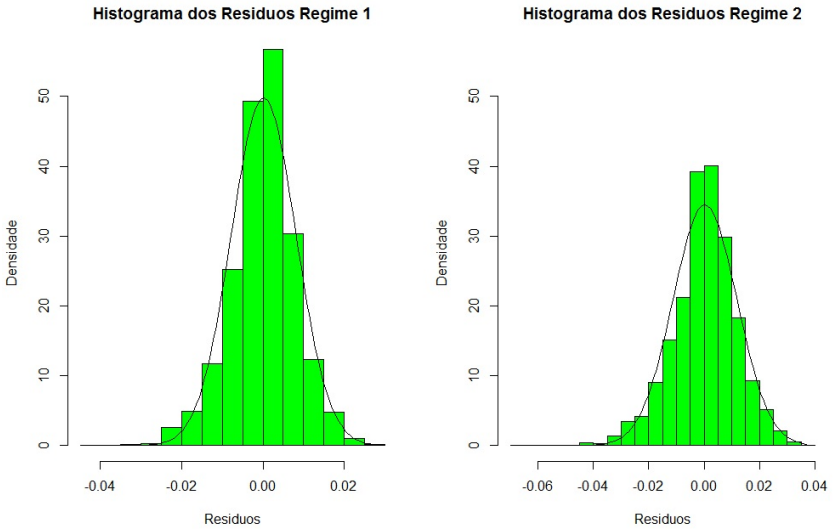

Figura 5.44: *Histograma dos Resíduos padronizados do modelo Normal-S&P500.*

5.5.2 Análise de Resíduos para o modelo MRS-GARCH-t2

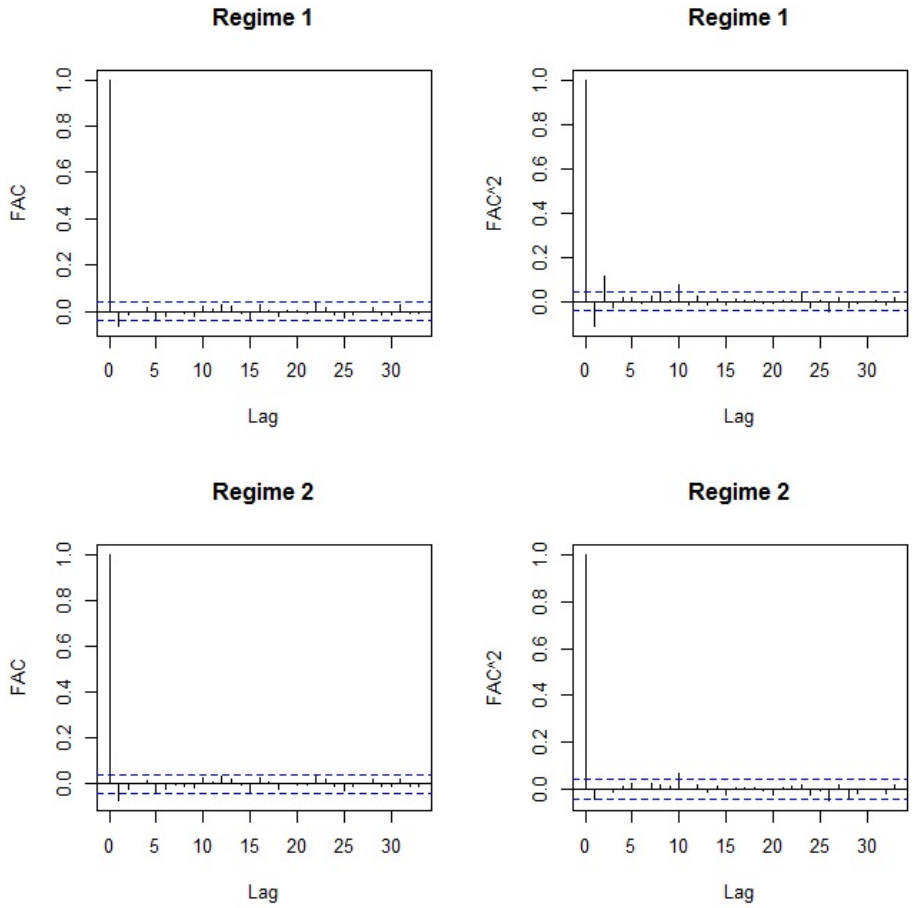

Figura 5.45: *Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo t2-S&P500.*

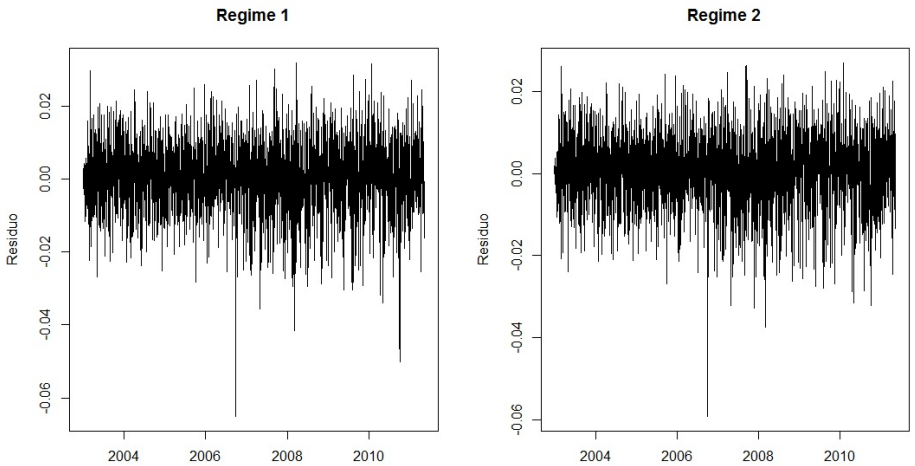

Figura 5.46: *Resíduos padronizados nos Regimes 1 e 2 do modelo t2-S&P500.*

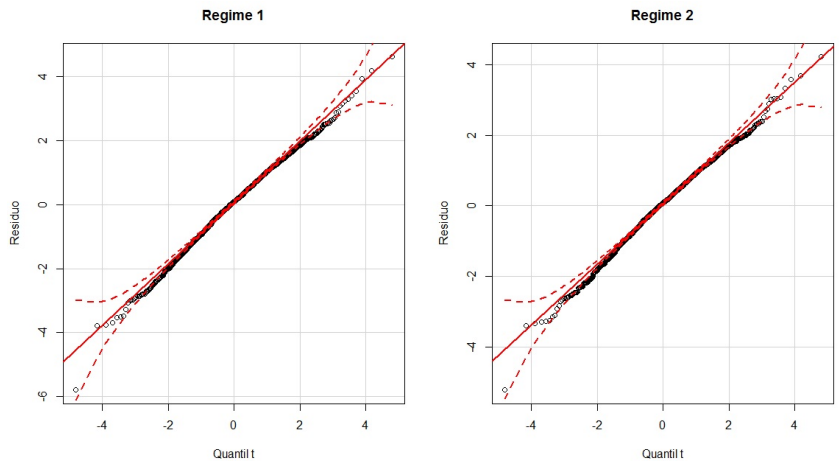

Figura 5.47: *Gráfico de quantis dos Resíduos padronizados do modelo t2-S&P500.*

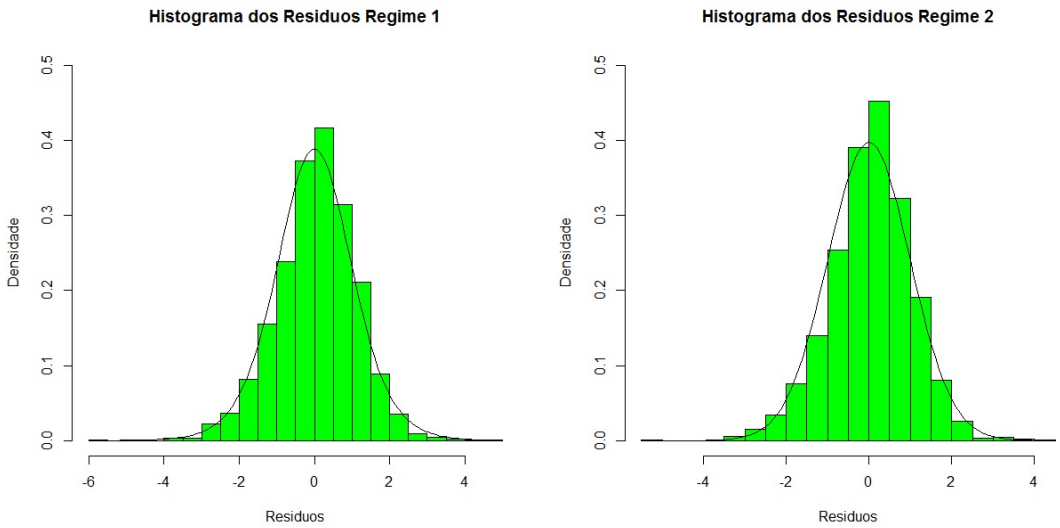

Figura 5.48: *Histograma dos Resíduos padronizados do modelo t2-S&P500.*

5.5.3 Análise de Resíduos para o modelo MRS-GARCH-GED

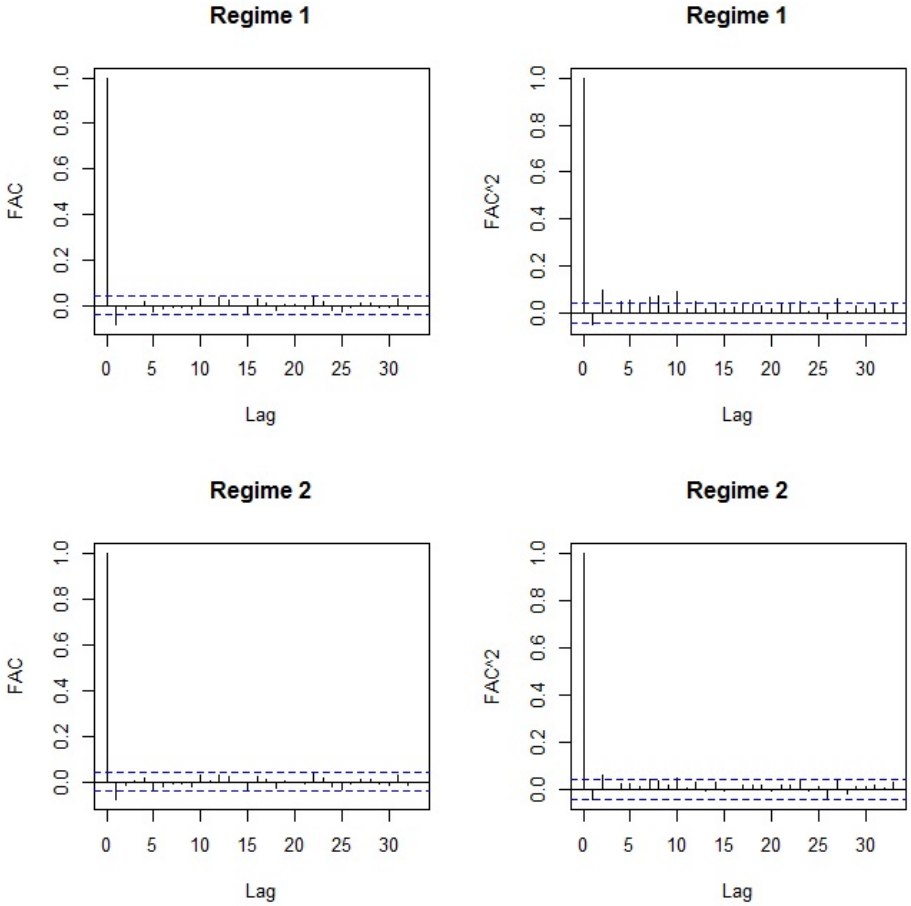

Figura 5.49: *Função de autocorrelação e função de autocorrelação ao quadrado dos resíduos padronizados do modelo GED-S&P500.*

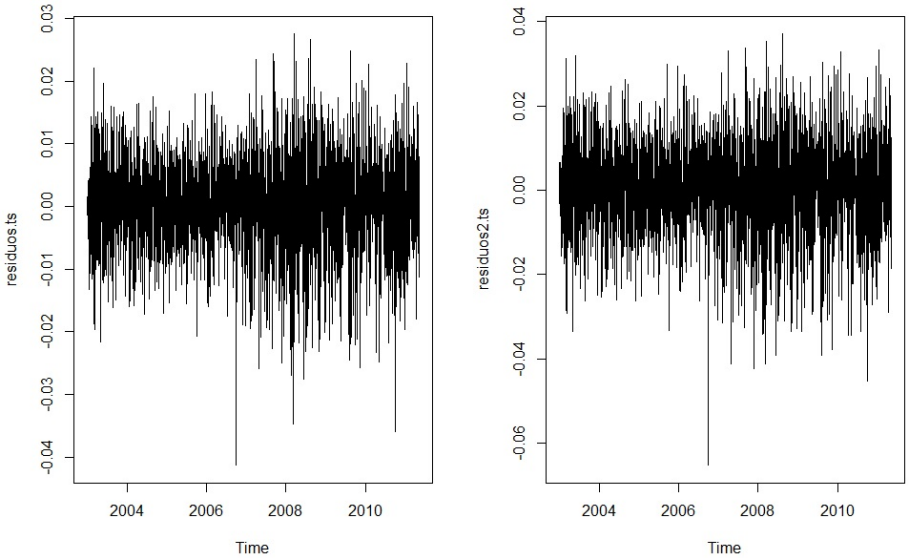

Figura 5.50: *Resíduos padronizados nos Regimes 1 e 2 do modelo GED-S&P500.*

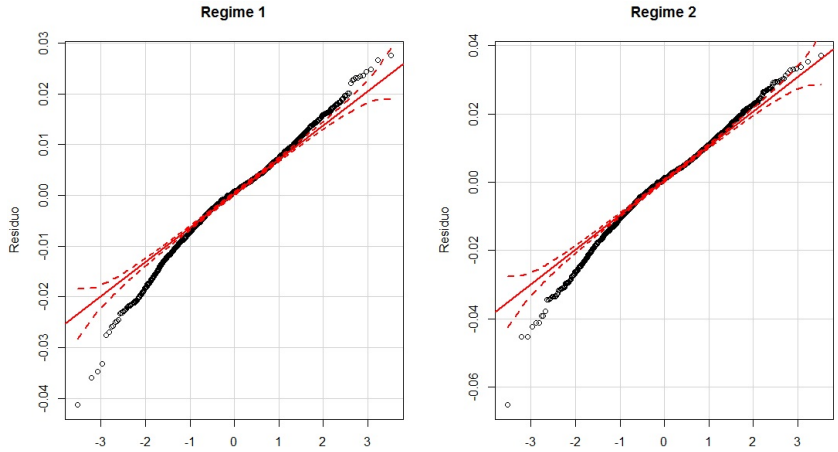

Figura 5.51: *Gráfico de quantis dos Resíduos padronizados do modelo GED-S&P500.*

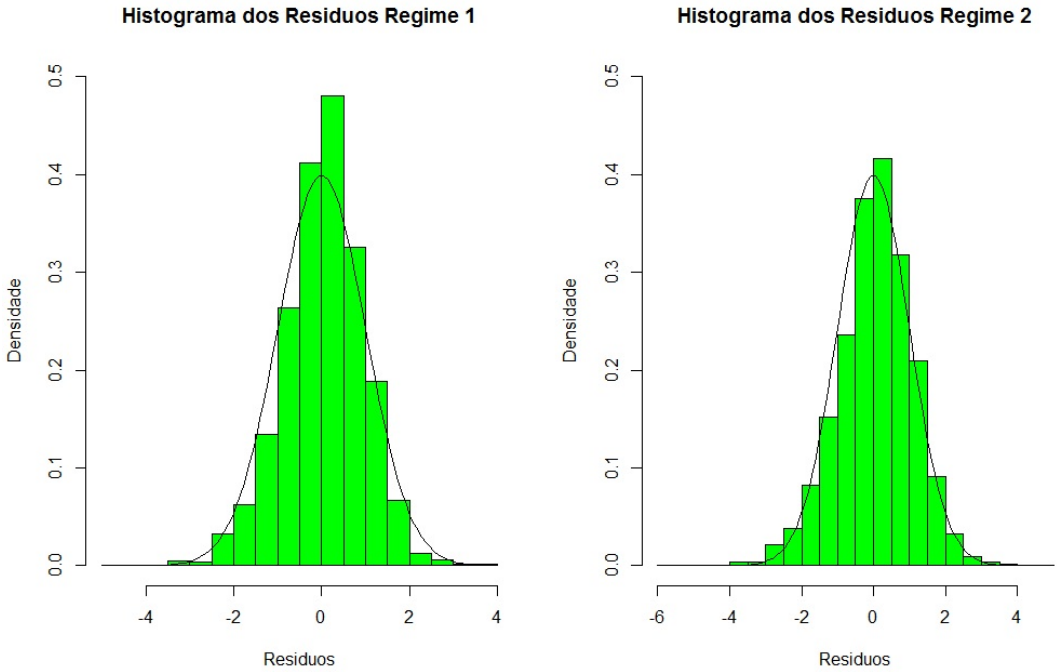

Figura 5.52: *Histograma dos Resíduos padronizados do modelo GED-S&P500.*

Os gráficos de quantis indicam que somente os modelos com distribuição t de Student com distintos graus de liberdade são adequados e por isso esse modelo foi escolhido para os log retornos da S&P500. A função de autocorrelação parcial dos resíduos padronizados não indicam nenhuma quebra de comportamento de ruído branco nos resíduos.

Em seguida, avaliaremos o desempenho dos modelos com base na cobertura dos intervalos de confiança para os log retornos.

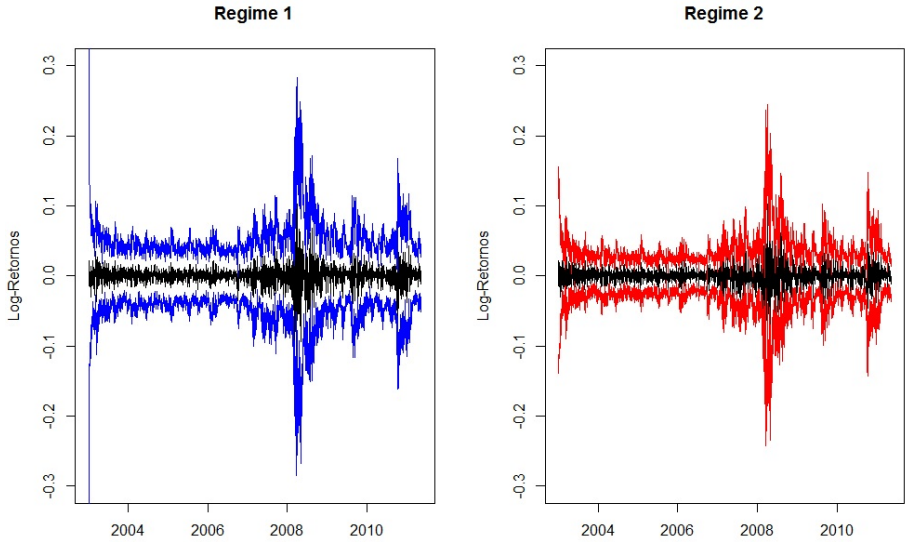

Figura 5.53: *Retornos e Intervalos de confiança (95%)-Normal-S&P500.*

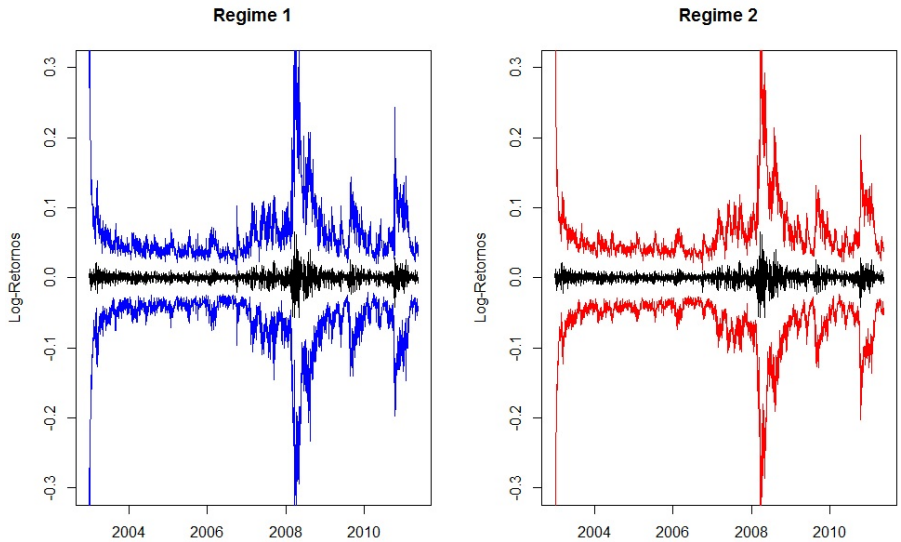

Figura 5.54: *Retornos e Intervalos de confiança (95%)-t2-S&P500.*

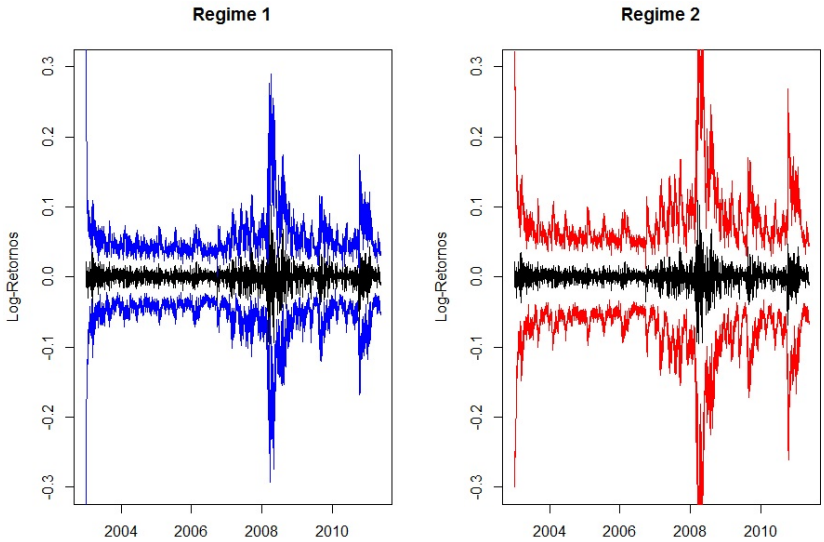

Figura 5.55: *Retornos e Intervalos de confiança (95%)-GED-S&P500.*

Tabela 5.6: *Porcentagem de Cobertura do intervalos de confiança para os log-retornos da S&P500*

	Cobertura (%)	Comp. Médio	Mínimo	Máximo
GARCH(1,1)-N	95	4.41	2.03	20.43
2R.GARCH(1,1)-N	99	5.40	3.00	79.90
2R.GARCH(1,1)-t2	99	7.27	3.12	46.63
GARCH(1,1)-GED	98	5.26	2.81	24.01
2R.GARCH(1,1)-GED	99	5.63	3.02	79.18

Na Tabela 5.6 mostra-se o modelo Garch usual como GARCH(1,1) para cada distribuição, e 2R.GARCH(1,1) para o modelo Garch com dois regimes.
Os intervalos apresentaram cobertura maior que 95% e as maiores coberturas foram para as três distribuições utilizadas com dois regimes. Pode-se observar que os valores para os modelos com 2 regimes tem maior comprimento médio do que os modelos com 1 regime.

Capítulo 6

Conclusões

Neste trabalho apresentamos os modelos GARCH e GARCH com mudança de regime para modelar a volatilidade de séries financeiras. Os modelos de volatilidade com mudança de regime apresentaram duas séries temporais não observadas, a série de volatilidade condicional e dos regimes. Podemos considerar as mudanças Markovianas de regime, que parece ser bastante razoável em situações práticas, já que em um regime hoje, deve influenciar o regime amanhã. Logo, o modelo GARCH com mudança de regime markoviana é bastante versátil para modelar comportamentos variados.

O objetivo principal deste trabalho foi pesquisar a eficácia do uso de modelos com mudança de regime Markoviano para aplicá-los às séries financeiras. Cabe ressaltar o fato de que estas séries, Ibovespa e S&P500, foram muito influenciadas pela crise financeira do subprime em 2008, que mudaram bastante a dinâmica destas séries imediatamente após a crise mencionada. Os modelos que analisamos neste trabalho conseguiram detectar um regime com maior efeito do erro passado e outro com maior inércia para a variância anterior.

Além disso, estes modelos fornecem-nos informações úteis a respeito dos parâmetros de cada uma das fases. Os modelos com mudança de regime permitiram modelar séries que apresentaram mudanças em seus parâmetros e estimar probabilidades de cada regime ao longo do tempo.

Na literatura tem-se mostrado que os modelos com mudança de regime markoviano para séries financeiras com distribuição t de Student se ajustam bem as séries. Aqui neste trabalho pode-se observar que o modelo com distribuição t de Student com distintos graus de liberdade conseguiu corrigir a dependência da volatilidade (ver plot dos resíduos para cada distribuição do erro), o que quer dizer que o modelo proposto se ajustou bem às séries. Também as funções de autocorrelação e autocorrelação parcial dos resíduos padronizados e resíduos padronizados ao quadrado não indicam nenhuma quebra de comportamento de ruído branco nos resíduos.

Em ambas aplicações, para os log-retornos do Ibovespa e do S&P500, o modelo GARCH(1,1) com mudança de regime e erros com distribuição t de Student com graus de liberdade distintos apresentou bom ajuste e todas as suposições do modelo parecem satisfeitas. Entretanto, observando as probabilidades estimadas (suavizadas) dos regimes, não é possível identificar períodos em que o regime 1 deve prevalecer ou o regime 2. Além disso, as séries dos desvios padrões condicionais estimadas são muito semelhantes, o que se torna a distinção entre os regimes um tanto quanto confusa. Essa semelhança entre os regime não parece ocorrer ao analisarmos séries de atividade econômica como em Alencar (2013) e Correa (2004), nesses trabalhos é possível identificar claramente os períodos de aceleração e recessão econômica e os modelos com mudança de regime são utilizados para tal identificação.

Apesar da distribuição normal e GED não apresentarem bom ajuste ao considerarmos os gráficos de quantis dos resíduos padronizados de cada regime, para os retornos do S&P500 as probabilidades

suavizadas parecem identificar mais claramente períodos em que ocorrem cada regime.

Através da análise de resíduos do modelo GARCH com mudança de regime verificou-se que a distribuição t-Student com distintos graus de liberdade se mostrou superior a distribuição normal e GED para caracterizar a distribuição dos retornos. A porcentagem de cobertura dos intervalos de confiança para a distribuição normal do modelo GARCH com mudança de regime foi superior em relação ao modelo GARCH usual, o mesmo aconteceu para a distribuições t-Student com distintos graus de liberdade, o que pode-se tornar interessante sua utilização no mercado financeiro.

A maior conclusão e contribuição desta dissertação é a ótima performance dos modelos de mudanca de regime tanto para captar a dinâmica das séries em análise quanto para a realização de regras de mercado. A aplicação destes modelos ao mercado financeiro tem um grande campo de estudo que com certeza ira se difundir muito no futuro próximo.

Apêndice A

Mistura de Distribuições

Mistura de Normais A densidade de $y_t/s_t = j$ é dada por

$$f\left(y_t/s_t = j;\theta\right) = \frac{1}{\sqrt{2\pi}\sigma_j}\exp\left\{-\frac{\left(y_t-\mu_j\right)^2}{2\sigma_j^2}\right\},$$

em que $j = 1, 2, ..., k$ e θ é o vetor de parâmetros, μ_{s_t} e $\sigma_{s_t}^2$ são dados segundo o valor da variável estado da natureza.

Por Bayes

$$p\left(y_t, s_t = j;\theta/Y_t\right) = f\left(y_t/s_t = j;\theta\right)p\left(s_t = j;\theta\right)$$

onde $P\left(s_t = j;\theta\right) = \pi_j$

$$p\left(y_t, s_t = j;\theta/Y_t\right) = \frac{\pi_j}{\sqrt{2\pi}\sigma_j}\exp\left\{-\frac{\left(y_t-\mu_j\right)^2}{2\sigma_j^2}\right\}$$

A distribuição não condicional de y_t será dada pela soma das probabilidades ergodicas em cada regime

$$f\left(y_t;\theta\right) = \sum_{j=1}^{k} p\left(y_t, s_t = j;\theta/Y_t\right)$$

Neste caso temos dois regimes, logo temos

$$f\left(y_t;\theta\right) = \frac{\pi_1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{\left(y_t-\mu_1\right)^2}{2\sigma_1^2}\right\} + \frac{\pi_2}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{\left(y_t-\mu_2\right)^2}{2\sigma_2^2}\right\}$$

A soma de duas normais tem como distribuição

$$r_t^{(1)}\pi_1 + r_t^{(2)}\pi_2 \rightarrow MisturadeNormais\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\begin{pmatrix}\pi_1\\ \pi_2\end{pmatrix}\begin{pmatrix}\sigma_1^2\\ \sigma_2^2\end{pmatrix}$$

Vamos a encontrar sua média e sua variância respetivamente. Então,

$$\mu = E\left[r_t^{(1)}\pi_1 + r_t^{(2)}\pi_2\right] = \pi 1 E\left[r_t^{(1)}\right] + \pi_2 E\left[r_t^{(2)}\right] = \pi_1\mu_1 + \pi_2\mu_2 = \pi_1\mu_1 + \left(1-\pi_1\right)\mu_2$$

$$Var = V\left[r_t^{(1)}\pi_1 + r_t^{(2)}\pi_2\right] = \pi 1^2 V\left[r_t^{(1)}\right] + \pi_2^2 V\left[r_t^{(2)}\right] = \pi_1^2\sigma_1^2 + \pi_2^2\sigma_2^2 = \pi_1^2\sigma_1^2 + \left(1-\pi_1\right)^2\sigma_2^2$$

Logo,

$$r_t^{(1)}\pi_1 + r_t^{(2)}\pi_2 \to MN\left(\pi_1\mu_1 + \pi_2\mu_2, \pi_1^2\sigma_1^2 + \pi_2^2\sigma_2^2\right)$$

Mistura de t-Student

$$Q(\theta) = \sum_{h=1}^{H} p_h t_d\left(\theta/\mu_h, \sum_h, \nu\right)$$

onde $p_h, (h = 1, ..., H)$ são as probabilidades de mistura de componentes t-Student, $0 \leq p_h \leq 1, \sum_{h=1}^{H} p_h = 1$, e $t_d\left(\theta/\mu_h, \sum_h, \nu\right)$ é a densidade d-dimensional t-Student com vetor de modo μ_h, matriz escala $\sum_h$, e ν graus de liberdade.

$$t_d\left(\theta/\mu_h, \sum_h, \nu\right) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\pi\nu)^{d/2}} \times \left(det\sum_h\right)^{-1/2}\left(1 + \frac{(\theta-\mu_h)'\sum_h^{-1}(\theta-\mu_h)}{\nu}\right)^{\frac{-(\nu+d)}{2}}$$

Nosso caso a distribuição t-Student de nosso interesse tem uma dimensão. Então com $d = 1, \sum_h = \sigma_h^2$ e $\theta = x$ temos que:

$$t_d\left(\theta/\mu_h, \sum_h, \nu\right) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\pi\nu)^{1/2}}\frac{1}{\sigma_h}\left(1 + \frac{(x-\mu_h)^2}{\nu\sigma_h^2}\right)^{\frac{-(\nu+1)}{2}}$$

Para cada regime temos as seguintes funções de distribuição,

$$t_1 = \frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)(\pi\nu_1)^{1/2}}\frac{1}{\sigma_1}\left(1 + \frac{(x-\mu_1)^2}{\nu_1\sigma_1^2}\right)^{\frac{-(\nu_1+1)}{2}}$$

$$t_2 = \frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)(\pi\nu_2)^{1/2}}\frac{1}{\sigma_2}\left(1 + \frac{(x-\mu_2)^2}{\nu_2\sigma_2^2}\right)^{\frac{-(\nu_2+1)}{2}}$$

logo,

$$\begin{aligned} q(\theta) &= p_1 t_1 + p_2 t_2 \\ &= \frac{p_1\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)(\pi\nu_1)^{1/2}}\frac{1}{\sigma_1}\left(1 + \frac{(x-\mu_1)^2}{\nu_1\sigma_1^2}\right)^{\frac{-(\nu_1+1)}{2}} + \frac{(1-p_1)\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)(\pi\nu_2)^{1/2}}\frac{1}{\sigma_2}\left(1 + \frac{(x-\mu_2)^2}{\nu_2\sigma_2^2}\right)^{\frac{-(\nu_2+1)}{2}} \end{aligned}$$

Então temos,

$$T = \begin{matrix} T_1 \\ T_2 \end{matrix} \to T_{2,\nu_1,\nu_2}\left(\mu_1, \mu_2, \sum_h\right)$$

Para saber a forma fechada da densidade $q(\theta)$ temos que fazer $x = p_1 t_1$ e $y = p_2 t_2$. Chamemos $z = x + y$.

$$p(Z \leq z) = p(x + y \leq z) = \int_{x=0}^{x=z}\int_{y=0}^{y=z-x} f(x, y)\, dxdy \tag{A.1}$$

Primeiramente vamos calcular a $f(x,y)$.

$$\begin{aligned} f(x,y) &= \frac{\Gamma\left(\frac{\nu_1+1}{2}\right)p_1}{\Gamma\left(\frac{\nu_1}{2}\right)(\pi\nu_1)^{1/2}}\frac{1}{\sigma_1}\left(1+\frac{(x-\mu_1)^2}{\nu_1\sigma_1^2}\right)^{\frac{-(\nu_1+1)}{2}} \times \frac{\Gamma\left(\frac{\nu_2+1}{2}\right)(1-p_1)}{\Gamma\left(\frac{\nu_2}{2}\right)(\pi\nu_2)^{1/2}}\frac{1}{\sigma_2}\left(1+\frac{(x-\mu_2)^2}{\nu_2\sigma_2^2}\right)^{\frac{-(\nu_2+1)}{2}} \\ &= \frac{\Gamma\left(\frac{\nu_1+1}{2}\right)\Gamma\left(\frac{\nu_2+1}{2}\right)p_1(1-p_1)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)(\pi\nu_1)^{1/2}(\pi\nu_2)^{1/2}}\frac{1}{\sigma_1\sigma_2}\left(1+\frac{(x-\mu_1)^2}{\nu_1\sigma_1^2}\right)^{\frac{-(\nu_1+1)}{2}}\left(1+\frac{(x-\mu_2)^2}{\nu_2\sigma_2^2}\right)^{\frac{-(\nu_2+1)}{2}} \end{aligned} \tag{A.2}$$

Substituindo A2 em A1

$$\begin{aligned} &= \frac{\Gamma\left(\frac{\nu_1+1}{2}\right)\Gamma\left(\frac{\nu_2+1}{2}\right)p_1(1-p_1)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)(\pi\nu_1)^{1/2}(\pi\nu_2)^{1/2}}\frac{1}{\sigma_1\sigma_2}\int_{x=0}^{x=z}\int_{y=0}^{y=z-x}\left(1+\frac{(x-\mu_1)^2}{\nu_1\sigma_1^2}\right)^{\frac{-(\nu_1+1)}{2}} \\ &\qquad \left(1+\frac{(x-\mu_2)^2}{\nu_2\sigma_2^2}\right)^{\frac{-(\nu_2+1)}{2}} dydx \\ &= \frac{\Gamma\left(\frac{\nu_1+1}{2}\right)\Gamma\left(\frac{\nu_2+1}{2}\right)p_1(1-p_1)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)(\pi\nu_1)^{1/2}(\pi\nu_2)^{1/2}}\frac{1}{\sigma_1\sigma_2}\int_{x=0}^{x=z}\left(1+\frac{(x-\mu_1)^2}{\nu_1\sigma_1^2}\right)^{\frac{-(\nu_1+1)}{2}} \\ &\qquad \left[\int_{y=0}^{y=z-x}\left(1+\frac{(x-\mu_2)^2}{\nu_2\sigma_2^2}\right)^{\frac{-(\nu_2+1)}{2}} dy\right] dx \end{aligned}$$

Vamos a calcular

$$\int_{y=0}^{y=z-x}\left(1+\frac{(x-\mu_2)^2}{\nu_2\sigma_2^2}\right)^{\frac{-(\nu_2+1)}{2}} dy$$

Defina $c=\frac{\nu+1}{2}, b=\nu_2\sigma_2^2, b>0 e a=\mu_2$ Logo,

$$\int_{y=0}^{y=z-x}\left(1+\frac{(y-a)^2}{b}\right)^{-c} dy$$

Encontraremos uma primitiva que tenha forma fechada,

$$\int\left(1+\frac{(y-a)^2}{b}\right)^{-c} dy$$

Defina, $\frac{y-a}{\sqrt{b}}=w \Rightarrow \frac{dy}{\sqrt{b}}=dw$, Então

$$\begin{aligned} \int\left(1+\frac{(y-a)^2}{b}\right)^{-c} dy &= \int\left(1+w^2\right)^{-c}\sqrt{b}dw = \sqrt{b}\int\frac{1}{(1+w^2)^c}dw \\ &= \sqrt{b}\left[w2F_1\left(\frac{1}{2},c;\frac{3}{2};-w^2\right)+cte\right] \\ &= \sqrt{b}\left[\frac{y-a}{\sqrt{b}}2F_1\left(\frac{1}{2},\frac{\nu_2+1}{2};\frac{3}{2};-\frac{(y-a)^2}{b}\right)+cte\right] \\ &= (y-a)\,2F_1\left(\frac{1}{2},\frac{\nu_2+1}{2};\frac{3}{2};-\frac{(y-a)^2}{b}\right)+cte \end{aligned}$$

$$= (y-\mu_2)\, 2F_1\left(\frac{1}{2}, \frac{\nu_2+1}{2}; \frac{3}{2}; -\frac{(y-\mu_2)^2}{b}\right) + cte$$

Então,

$$\int_{y=0}^{y=z-x}\left(1+\frac{(y-a)^2}{b}\right)^{-c} dy = (y-\mu_2)\, 2F_1\left(\frac{1}{2}, \frac{\nu_2+1}{2}; \frac{3}{2}; -\frac{(y-\mu_2)^2}{b}\right)\Bigg|_{y=0}^{y=z-x}$$

$$= (z-x-\mu_2)\, 2F_1\left(\frac{1}{2}, \frac{\nu_2+1}{2}; \frac{3}{2}; -\frac{(z-x-\mu_2)^2}{b}\right) - (-\mu_2)\, 2F_1\left(\frac{1}{2}, \frac{\nu_2+1}{2}; \frac{3}{2}; -\frac{(-\mu_2)^2}{b}\right)$$

Definamos agora,

$$2F_1(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\int_0^1 \frac{t^{b-1}(1-t)^{c-b-1}}{(1-tz)^a} dt$$

Em nosso caso,

$$2F_1\left(\frac{1}{2}, \frac{\nu_2+1}{2}; \frac{3}{2}; -\frac{(z-x-\mu_2)^2}{\nu_2\sigma_2^2}\right) = \frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{\nu_2+1}{2}\right)\Gamma\left(\frac{3}{2}-\frac{\nu_2+1}{2}\right)}\int_0^1 \frac{t^{\frac{\nu_2+1}{2}-1}(1-t)^{\frac{3}{2}-\frac{\nu_2+1}{2}-1}}{\left(1-t\left(-\frac{(z-x-\mu_2)^2}{\nu_2\sigma_2^2}\right)\right)^{1/2}} dt$$

$$= \frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{\nu_2+1}{2}\right)\Gamma\left(\frac{3}{2}-\frac{\nu_2+1}{2}\right)}\int_0^1 t^{\frac{\nu_2+1}{2}-1}(1-t)^{\frac{3}{2}-\frac{\nu_2+1}{2}-1}\left(1-t\left(-\frac{(z-x-\mu_2)^2}{\nu_2\sigma_2^2}\right)\right)^{-1/2} dt$$

$$= \frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{\nu_2+1}{2}\right)\Gamma\left(\frac{3}{2}-\frac{\nu_2+1}{2}\right)}\int_0^1 t^{\left(\frac{\nu_2+1}{2}\right)-1}(1-t)^{\left(\frac{3}{2}-\frac{\nu_2+1}{2}\right)-1}\left(1-t\underbrace{\left(-\frac{(z-x-\mu_2)^2}{\nu_2\sigma_2^2}\right)}_{cte}\right)^{-1/2} dt \quad \text{(A.3)}$$

Pode-se Observar que (3) tem forma de uma distribuição beta,

$$f(t) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} t^{a-1}(1-t)^{b-1}, 0 < t < 1$$

$$f(t) = \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} t^{a-1}(1-t)^{b-1} dt = 1$$

más para isso temos que fazer uma mudança de variável para a base $(1-tz)$ em nosso caso, z é uma constante, já que não depende de t.

Referências Bibliográficas

Alencar (2013) C. M. C.; Morettin P. A. Alencar, Airlane P.; Toloi. *State space Markov switching models using wavelets.* Studies in Nonlinear Dynamics Econometrics, v. 0, p. 1-18, 2013. Citado na pág. 67

Bollerslev (1986) T. Bollerslev. *Generalized Autoregressive Conditional Heteroskedasticity.* Journal of Econometrics, 31, 307-327. Citado na pág. 11, 14

Correa (2004) O. R. Correa, D. S. A. ; Hillbrecht. *Ciclos Internacionais de Négocios: Uma Análise de Mudança de Regime Markoviano para Brasil, Argentina e Estados Unidos.* Banco Central do Brasil. Citado na pág. 67

Dueker (1997) M. J. Dueker. *Markov Switching in GARCH Processes and Reverting Stock Market Volatility.* Journal of Business and Economic Statistics. 15(1), 26-34., first ed. Citado na pág. 1, 33

Engle (1987) Lilien D. M. Robins R. P. Engle, R. F. *Estimating Time Varying Risk Premia in the Term-Structure: the ARCH-M Model.* Econometrica, 55(2), 391-407. Citado na pág. 21

Engle (1989) R. F. Engle. *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance Of United Kingdom Inflation.* Econometrica, 57(4), 987-1007. Citado na pág. 11, 13

Fama (1965) E. F. Fama. *The behavior of stock market prices.* Journal of Business, 38, 34-105. Citado na pág. 17

Fuller (1996) W. A. Fuller. *Introduction to Statistical Time Series.* John Wiley and Sons, New York, second ed. Citado na pág. 3

Gray (1996a) S.F. Gray. *Modeling the conditional distribution of interest rates as a Regime-Switching process.* Journal of Financial Economics,42,27-62. Citado na pág. 27

Hamilton (1989) J. D. Hamilton. *A New Approach to the Economic Analysis with of Non-Stationary Time Series and the Bussines Cycle.* Econometrica, 57: 357-380. Citado na pág. iii, 1, 26, 35

Hamilton e Susmel (1994) J. D. Hamilton e R. Susmel. *Autoregressive Conditional Heterocedasticity and Changes in Regime.* Journal of Econometrics. 64(1-2),307-333. Citado na pág. 3, 8

Hansen (1994) B. E. Hansen. *Autoregressive Conditional Density Estimation.* Journal of Applied Econometrics. Citado na pág. 33

Hayashi (2000) F. Hayashi. Econometrics. Princeton University Press, New Jersey. Citado na pág. 8

He (1999a) Terasvirta T. He, C. *Properties of Moments of a Family of GARCH Processes.* Journal of Econometrics, 92, 173-192. Citado na pág. 15

He (1999b) Terasvirta T. He, C. *Fourth Moment Structure of the GARCH(p,q) Process.* Econometric Theory, 15, 824-846. Citado na pág. 15

Jarque (1980) Anil K. Jarque, Carlos M.; Bera. *Efficient tests for normality, homoscedasticity and serial independence of regression residuals.* Economics Letters 6 (3): 255259. Citado na pág. 18

Klaassen (2002) F. Klaassen. *Improving GARCH Volatility Forecasts.* Empirical Economics, 27(2), 363-394. Citado na pág. 1, 26

Mandelbrot (1963) B. B. Mandelbrot. *The variation of certain speculative prices.* Journal of Business, 36, 394-419. Citado na pág. 17

Marcucci (2005) Juri. Marcucci. *Forecasting Stock Market Volatility with Regime-Switching GARCH Models.* Studies in Nonlinear Dynamics and Econometrics, Volume 9, Issue 4, Article 6. Citado na pág. iii, 16, 17

Mittnik (2000) Rachev S. T. Mittnik, S. *Stable Paretian Models in Finance.* Wiley, New York. Citado na pág. 17

Mittnik (1998) Rachev S. T. Paolella M. S. Mittnik, S. *Stable Paretian Models in Finance: Some empirical and theorical aspect, in R. J. Adler, R. E. Feldman and M. S. Taqqu.* A Practial Guide to Heavy Tails: Statistical Techniques and Applications, 79-110, Birkhauser, Boston. Citado na pág. 17

Morettin (2006) Toloi C. M. C. Morettin, P.A. *Análise de Séries Temporais.* Edgar Blucher, São Paulo. Citado na pág. 3, 4, 9

Nelson (1991) D. B. Nelson. *Condicional Heteroscedasticity in Asset Returns: A New Approach.* Econometrica, 59(2),347-370. Citado na pág. 11

Peng. L. (2003) Q. Peng. L., Yao. *Least absolute deviation estimation for ARCH and GARCH models.* Biometrika, 90 (4), 967-975. Citado na pág. 17, 18

Shapiro (1965) M. B. Shapiro, S. S.; Wilk. *An analysis of variance test for normality (complete samples).* Biometrika 52 (3-4): 591611. Citado na pág. 18

White (1984) H. White. *Asymptotic Theory for Econometricians.* Academic Press, San Diego. Citado na pág. 8

Zivot (2003) Wang J. Zivot, E. *Modelling Financial Time Séries with S-PLUS.* Springer, New York. Citado na pág. 7, 31