
Modelagem da volatilidade em séries temporais
financeiras via modelos GARCH com abordagem
Bayesiana

Karen Fiorella Aquino Gutierrez

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

KAREN FIORELLA AQUINO GUTIERREZ

**MODELAGEM DA VOLATILIDADE EM SÉRIES TEMPORAIS FINANCEIRAS VIA
MODELOS GARCH COM ABORDAGEM BAYESIANA**

Dissertação apresentada ao Departamento de Estatística – DEs-UFSCar e ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestra em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística UFSCar-USP.
VERSÃO REVISADA.

Orientador: Prof. Dr. Ricardo Sandes Ehlers
Coorientador: Prof. Dr. Marinho Gomes de Andrade Filho

**São Carlos
Agosto de 2017**

FEDERAL UNIVERSITY OF SÃO CARLOS
EXACT SCIENCES AND TECHNOLOGY CENTER
JOIN GRADUATE PROGRAM IN STATISTICS UFSCar-USP

KAREN FIORELLA AQUINO GUTIERREZ

**MODELING OF VOLATILITY IN FINANCIAL TIME SERIES USING GARCH MODELS WITH
BAYESIAN APPROACH**

Master dissertation submitted to the Departamento de Estatística – DEs-UFSCar and to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master joint Graduate Program in Statistics UFSCar-USP. FINAL VERSION.

Advisor : Prof. Dr. Ricardo Sandes Ehlers
Co-advisor: Prof. Dr. Marinho Gomes de Andrade Filho

**São Carlos
August 2017**



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado da candidata Karen Fiorella Aquino Gutierrez, realizada em 18/07/2017:

Prof. Dr. Ricardo Sandes Ehlers
USP

Profa. Dra. Maria Sílvia de Assis Moura
UFSCar

Prof. Dr. Mauricio Enrique Zevallos Herencia
UNICAMP

Dedicado a Adriano.

AGRADECIMENTOS

Agradeço a minha mãe por ter confiado em mim sempre, a meu namorado pelo apoio constante e a cada uma das pessoas que me transmitiu essa vontade de querer fazer ciência. Também agradeço ao professor Ricardo Sandes Ehlers, por ter me permitido trabalhar com ele.

Agradecimento especial é direcionado ao Centro de Ciências Matemáticas Aplicadas à Indústria¹ da Universidade de São Paulo (CeMEAI), por disponibilizar o cluster Euler para fins de meu projeto.

¹ <<http://www.cemeai.icmc.usp.br/>>

“the application of GARCH to long time series of stock-return data will yield a high measure of persistence because of the presence of deterministic shifts in the unconditional variance and the subsequent failure of the econometrician to model these shifts.”
(Christopher G. Lamoureux and William D. Lastrapes)

RESUMO

K. FIORELLA AQUINO G. **Modelagem da volatilidade em séries temporais financeiras via modelos GARCH com abordagem Bayesiana**. 2017. 84 p. Dissertação (Mestrado em Estatística – Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

Nas últimas décadas a volatilidade transformou-se num conceito muito importante na área financeira, sendo utilizada para mensurar o risco de instrumentos financeiros. Neste trabalho, o foco de estudo é a modelagem da volatilidade, que faz referência à variabilidade dos retornos, sendo esta uma característica presente nas séries temporais financeiras. Como ferramenta fundamental da modelação usaremos o modelo GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*), que usa a heterocedasticidade condicional como uma medida da volatilidade. Considerar-se-ão duas características principais a ser modeladas com o propósito de obter um melhor ajuste e previsão da volatilidade, estas são: a assimetria e as caudas pesadas presentes na distribuição incondicional da série dos retornos. A estimação dos parâmetros dos modelos propostos será feita utilizando a abordagem Bayesiana com a metodologia MCMC (*Markov Chain Monte Carlo*) especificamente o algoritmo de Metropolis-Hastings.

Palavras-chave: Séries Temporais, Inferência Bayesiana, Volatilidade, Modelos GARCH, Distribuições assimétricas, MCMC.

ABSTRACT

K. FIORELLA AQUINO G. **Modeling of volatility in financial time series using GARCH models with Bayesian approach.** 2017. 84 p. Dissertação (Mestrado em Estatística – Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

In the last decades volatility has become a very important concept in the financial area, being used to measure the risk of financial instruments. In this work, the focus of study is the modeling of volatility, that refers to the variability of returns, which is a characteristic present in the financial time series. As a fundamental modeling tool, we used the GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*) model, which uses conditional heteroscedasticity as a measure of volatility. Two main characteristics will be considered to be modeled with the purpose of a better adjustment and prediction of the volatility, these are: heavy tails and an asymmetry present in the unconditional distribution of the return series. The estimation of the parameters of the proposed models is done by means of the Bayesian approach with an MCMC (*Markov Chain Monte Carlo*) methodology, specifically the Metropolis-Hastings algorithm.

Keywords: Time series, Bayesian inference, Volatility, GARCH models, Asymmetric distributions, MCMC.

LISTA DE ILUSTRAÇÕES

Figura 1 – Preços de Fechamento diários do Ibovespa desde Janeiro 2, 2009 até Dezembro 31, 2014.	31
Figura 2 – Retornos diários do Ibovespa desde Janeiro 2, 2009 até Dezembro 31, 2014.	32
Figura 3 – Características dos retornos do Ibovespa.	34
Figura 4 – Caudas da distribuição normal padrão, da T-Student padronizada com 5 graus de liberdade, da GED padronizada com $\delta = 1$ (Laplace) e com $\delta = 1.5$	45
Figura 5 – Distribuição T-Student com $\nu = 5$ e com γ igual a 0.6 ou 0.5, distribuição GED com $\delta = 1.5$ e com γ igual a 0.5 ou 2 e distribuição Normal com γ igual a 0.5 ou 1.5.	51
Figura 6 – Série de IBOVESPA de 03/01/2012 até 30/12/2016. (a) Gráfico da série; (b) Gráfico dos retornos; (c) Gráfico das autocorrelações dos retornos e (d) Gráfico das autocorrelações do quadrado dos retornos.	65
Figura 7 – Série de S&P500 de 03/01/2012 até 30/12/2016. (a) Gráfico da série; (b) Gráfico dos retornos; (c) Gráfico das autocorrelações dos retornos e (d) Gráfico das autocorrelações do quadrado dos retornos.	67
Figura 8 – Retornos e Volatilidades estimadas usando o modelo GARCH(1,1) com erros com distribuição T-Student assimétricos padronizados.	69
Figura 9 – Retornos e Volatilidades estimadas usando o modelo GARCH(1,1) com erros com distribuição Normal assimétricos padronizados.	70
Figura 10 – Retornos e Volatilidades estimadas usando o modelo GARCH(1,1) com erros com distribuição GED assimétricos padronizados.	71
Figura 11 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série IBOVESPA. Modelo ajustado: GARCH(1,1) com erros T-Student assimétricos padronizados.	72
Figura 12 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série IBOVESPA. Modelo ajustado: GARCH(1,1) com erros Normal assimétricos padronizados.	73
Figura 13 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série IBOVESPA. Modelo ajustado: GARCH(1,1) com erros GED assimétricos padronizados.	74
Figura 14 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série S&P500. Modelo ajustado: GARCH(1,1) com erros T-Student assimétricos padronizados.	75

Figura 15 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série S&P500. Modelo ajustado: GARCH(1,1) com erros Normal assimétricos padronizados.	76
Figura 16 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série S&P500. Modelo ajustado: GARCH(1,1) com erros GED assimétricos padronizados.	77

LISTA DE ALGORITMOS

Algoritmo 1 – Metropolis-Hastings com um bloco	58
Algoritmo 2 – Metropolis-Hastings em blocos	59

LISTA DE TABELAS

Tabela 1 – Estatística descritiva da série de retornos.	30
Tabela 2 – Primeiros momentos absolutos das distribuições assimétricas geradas.	46
Tabela 3 – Transformações adotadas dos parâmetros.	53
Tabela 4 – Viés e Raiz do Erro Quadrático Médio da estimação posteriori.	60
Tabela 5 – Viés e Raiz do Erro Quadrático Médio da estimação posteriori.	61
Tabela 6 – Viés e Raiz do Erro Quadrático Médio da estimação posteriori.	61
Tabela 7 – Estimativas do GARCH(1,1) com erros T-Student assimétricos padronizados de IBOVESPA.	65
Tabela 8 – Estimativas do GARCH(1,1) com erros Normal assimétricos padronizados de IBOVESPA.	66
Tabela 9 – Estimativas do GARCH(1,1) com erros GED assimétricos padronizados de IBOVESPA.	66
Tabela 10 – Estimativas do GARCH(1,1) com erros T-Student assimétricos padronizados de S&P500.	67
Tabela 11 – Estimativas do GARCH(1,1) com erros Normal assimétricos padronizados de S&P500.	68
Tabela 12 – Estimativas do GARCH(1,1) com erros GED assimétricos padronizados de S&P500.	68

SUMÁRIO

1	INTRODUÇÃO	23
1.1	Motivação	23
1.2	Revisão Bibliográfica	24
1.3	Apresentação dos Capítulos	26
2	SÉRIES DE TEMPO FINANCEIRAS	29
2.1	Retornos	29
2.2	Fatos Estilizados dos Retornos	31
2.2.1	<i>Não auto-correlacionados</i>	31
2.2.2	<i>Agrupamentos de volatilidades ao longo do tempo</i>	32
2.2.3	<i>Distribuição incondicional assimétrica e de caudas pesadas</i>	33
2.2.4	<i>Efeito de alavanca</i>	33
2.3	Inferência Bayesiana	35
2.4	Métodos MCMC	36
2.4.1	<i>Gibbs sampling</i>	36
2.4.2	<i>Metropolis-Hastings</i>	37
2.4.3	<i>Avaliação da Convergência</i>	39
3	MODELOS GARCH	41
3.1	Modelo GARCH	42
3.1.1	<i>Propriedades da distribuição incondicional</i>	42
3.2	Estimação dos Parâmetros	48
3.2.1	<i>Função de verossimilhança</i>	48
3.2.2	<i>Distribuições a Priori</i>	49
3.2.3	<i>Distribuição a Posteriori</i>	50
4	ESTUDO DE SIMULAÇÃO	53
4.1	Algoritmo Metropolis-Hastings	53
4.1.1	<i>Distribuição Proposta</i>	56
4.1.2	<i>Esquema do Algoritmo</i>	57
4.2	Estudo de Simulação	58
4.2.1	<i>Medidas do Erro de Estimação</i>	59
5	APLICAÇÃO	63

5.1	Série IBOVESPA	64
5.2	Série S&P500	66
5.3	Estimação da Volatilidade	67
6	CONCLUSÕES	79
	REFERÊNCIAS	81

INTRODUÇÃO

1.1 Motivação

A análise de séries temporais financeiras trabalha com dados coletados no mercado financeiro, por exemplo com preços de ativos financeiros. Devido à alta instabilidade que enfrenta o mercado financeiro por causa de fatores econômicos, políticos e sociais os preços dos ativos oscilam fortemente. Portanto, se define a esta variabilidade dos preços como **volatilidade** a qual não é diretamente observável. Com o propósito de modelar a volatilidade se define a variância condicional como uma medida estatística dela, logrando assim obter mais informação que possa nos ajudar a prever o comportamento futuro dos dados no mercado financeiro.

Uma das características da volatilidade é que ela tende a se apresentar em grupos numa série financeira, isto nós leva a pensar na existência de um grau de dependência no tempo. De modo que é necessário recorrer aos chamados modelos heterocedásticos condicionais, os quais fazem diferença entre a **variância condicional** variável no tempo e condicionada às informações disponíveis até o instante dado e a **variância incondicional** que é constante em uma série observada. Estes modelos e suas extensões são muito empregados, pois possuem propriedades que conseguem explicar alguns dos fatos estilizados ou características inerentes nas séries temporais financeiras.

O primeiro em utilizar esse tipo de modelos foi [Engle \(1982\)](#) que propõe modelar e prever a variância condicional, com os modelos ARCH (*Autoregressive Conditional Heteroscedasticity*). Uma generalização dos modelos ARCH foi proposta por [Bollerslev \(1986\)](#), os chamados modelos GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*), os quais são mais parcimoniosos do que os modelos ARCH no sentido de que descrevem a variância condicional com menos parâmetros, por conseguinte, são os mais empregados na prática financeira. Uma suposição muito comum nos modelos ajustados a dados de séries financeiras é que os *retornos*, que são calculados como a primeira diferença do logaritmo da série, estejam condicionalmente

distribuídos como uma normal. Mas com esta especificação os modelos não serão capazes de capturar eficientemente as caudas pesadas e a assimetria que caracterizam as séries financeiras. Portanto considerando o dito anteriormente, nesta dissertação nós apresentaremos a utilização de distribuições de probabilidade com caudas mais pesadas do que a normal para a distribuição dos erros e além disso vamos a inserir um parâmetro de assimetria nessas distribuições com o objetivo de obter uma melhor estimativa da série.

A inferência nos modelos ARCH e GARCH majoritariamente tem sido feita utilizando a abordagem da estatística clássica. Tem havido menos progresso na análise deste tipo de modelos desde uma abordagem Bayesiana. Isto ocorre devido a complexidade dos modelos e ao custo computacional que suas análises precisam. Portanto, com este trabalho pretende-se contribuir para o estudo da modelagem da volatilidade utilizando os modelos GARCH desde uma perspectiva Bayesiana, considerando distribuições assimétricas com caudas pesadas na distribuição do termo do erro e fazendo uso dos métodos MCMC para a estimação do modelo.

1.2 Revisão Bibliográfica

Antes da proposta de Engle, se trabalhava com métodos estatísticos que assumiam volatilidade constante, por falta de uma alternativa que modele as flutuações aleatórias da volatilidade ao longo do tempo. Portanto, a descoberta de [Engle \(1982\)](#) foi um grande avanço. Ele apresentou o conceito de heterocedasticidade condicional autorregressiva que capturava com precisão as propriedades de muitas séries de tempo financeiras. Seu modelo ARCH proposto, tem como ideia básica, que a volatilidade em um instante de tempo t não é constante e depende dos retornos passados. [Bollerslev \(1986\)](#) propôs uma generalização do ARCH que considera que a volatilidade depende além dos valores passados dos retornos, também dos valores passados da própria volatilidade. Estes modelos são os chamados modelos GARCH.

Numerosas extensões e refinamentos do modelo GARCH têm sido propostos para capturar os fatos estilizados adicionais observados nas séries de retornos financeiras, ver [Bollerslev \(2009\)](#). Como por exemplo a evidência empírica neste tipo de séries, de que em geral choques negativos tem um impacto maior sobre a volatilidade da série que choques positivos, mostrando assim, um certo grau de assimetria na volatilidade. Essa característica levou a alguns pesquisadores a propor extensões para o modelo, [Nelson \(1991\)](#) introduziu os modelos Exponencial GARCH (EGARCH), [Glosten, Jagannathan e Runkle \(1993\)](#) introduziram os modelos GJR-GARCH, [Ding, Granger e Engle \(1993\)](#) introduziram os modelos ARCH com potência assimétrica (APARCH) que foi um dos modelos mais promissores nesta linha. Os modelos APARCH representam uma generalização de modelos de classe ARCH e GARCH propostos na literatura, ver [Laurent \(2004\)](#). Notar que nos modelos apresentados acima a assimetria é considerada diretamente na formulação do modelo para modelar a existência de assimetria no impacto dos retornos passados sobre a volatilidade atual. Além disso, as distribuições empíricas

dos retornos dos ativos ademais de ter caudas pesadas podem ser assimétricas. Para lidar com esta última característica de assimetria dos retornos, que segundo estudos empíricos não pode ser capturada modelando a assimetria no processo de volatilidade, se propôs considerar certo grau de assimetria na distribuição dos erros. No artigo de [Ehlers \(2012\)](#) se compararam modelos GARCH com diferentes distribuições assimétricas para os termos dos erros. Em [Pipien \(2006\)](#) são avaliados alguns métodos propostos na literatura para inserir assimetria em distribuições originalmente contínuas, unimodais e simétricas ao redor da moda. [Ehlers \(2012\)](#) utilizou o método de [Fernández e Steel \(1998\)](#), que propõe introduzir assimetria nas distribuições mediante fatores de escala inversos nos valores positivos e negativos da variável. Obtendo assim, uma distribuição assimétrica ajustada apenas a um parâmetro de forma, o qual define o grau de assimetria.

Por outro lado, considerando captar outras características como as mudanças de comportamento ao longo do tempo das séries. Foram propostos os modelos com mudanças de regimes estocásticos, como os modelos Regime Switching ARCH (SWARCH), os modelos Regime switching GARCH (RS-GARCH), os modelos Markov Switching GARCH (MS-GARCH) ver [Ardia \(2009\)](#), [Bauwens, Preminger e Rombouts \(2010\)](#) e entre outros.

Além dos modelos já citados, para modelar a volatilidade, existem os modelos de Volatilidade Estocástica (VE) introduzidos por [Taylor \(1982\)](#), os quais tem sido uma alternativa aos modelos de tipo ARCH de [Engle \(1982\)](#). Estes modelos podem ser mais realistas ao assumir que a variância condicional varia estocasticamente, por isso é que em alguns casos estes modelos são conhecidos por modelar melhor a volatilidade do que os modelos GARCH, ver [Kim, Shephard e Chib \(1998\)](#) e [Barossi-Filho, Achcar e Souza \(2010\)](#). A estimativa dos parâmetros não é fácil neste tipo de modelos, devido a que não é possível obter a função da verossimilhança de forma analítica pois as volatilidades aparecem como variáveis latentes. Muitos métodos de estimação tem sido propostos na literatura, sendo o método MCMC proposto por [Jacquier, Polson e Rossi \(2002\)](#) um dos mais eficientes. Devido à evidência empírica de assimetria na distribuição dos retornos financeiros ver por exemplo [Jondeau e Rockinger \(2003\)](#), tem surgido extensões do modelo VE como as apresentadas por [Cappuccio, Lubian e Raggi \(2006\)](#), [Tsiotas \(2012\)](#) e [Abanto-Valle, Lachos e Dey \(2013\)](#).

Com respeito à aplicação do enfoque Bayesiano para a estimação dos modelos tipo ARCH, uma das primeiras propostas foi implementada por [Geweke \(1989\)](#) usando a integração de Monte Carlo com amostragem de importância. Anos depois, com o mesmo objetivo foi utilizado o algoritmo de Metropolis em [Geweke et al. \(1994\)](#). Continuando a pesquisa de [Bauwens e Lubrano \(1998\)](#), que apresentaram a inferência Bayesiana dos modelos GARCH usando o algoritmo Amostrador de Gibbs. No mesmo ano, [Nakatsuma \(1998\)](#) propõe um algoritmo de amostragem baseado em Cadeias de Markov para os modelos GARCH. Depois de um longo período, [Ardia \(2006\)](#) propõe a estimação Bayesiana do parcimonioso mas efetivo modelo GARCH(1,1) com erros Normais usando a metodologia proposta por [Nakatsuma \(1998\)](#) para

amostrar valores dos parâmetros diretamente da distribuição a posteriori conjunta, construindo assim uma cadeia de Markov através do algoritmo de Metropolis-Hastings. Este algoritmo foi primeiro proposto por [Metropolis *et al.* \(1953\)](#) e anos mais tarde generalizado por [Hastings \(1970\)](#) resultando no algoritmo de Metropolis-Hastings. No ano seguinte ao trabalho de [Ardia Takaishi \(2007\)](#) usou o algoritmo de Monte Carlo Híbrido para a estimação Bayesiana dos modelos GARCH. Devido às evidências na literatura de que muitas séries financeiras tendem a ter curtose observada ainda maior do que aquela considerada por um modelo GARCH com erros normais, [Ardia e Hoogerheide \(2010\)](#) implementou o pacote **bayesGARCH** no software R baseado no procedimento descrito em [Ardia \(2008\)](#) onde se faz a estimação Bayesiana de um modelo GARCH(1,1) com erros T-Student através do algoritmo de Metropolis-Hastings por blocos. Além da distribuição T-Student que é geralmente usada, há outras distribuições de probabilidade na literatura como a GED (*Generalized Error Distribution*), também conhecida como distribuição Exponencial Potência ver [Nelson \(1991\)](#).

Com base a tudo o exposto, a motivação da presente dissertação é modelar a volatilidade de índices do mercado acionário, empregando as metodologias propostas desde uma perspectiva Bayesiana. Tendo como objetivo principal contribuir à comparação de algoritmos para análise de modelos GARCH com distribuições assimétricas.

1.3 Apresentação dos Capítulos

Neste capítulo se apresentou a motivação para o planejamento do projeto, além da revisão bibliográfica da literatura existente nesta área de estudos.

No [Capítulo 2](#), definem-se alguns conceitos básicos em finanças que nos ajudarão a entrar em contexto com os termos de retornos e volatilidade. Se faz uma pequena introdução à inferência bayesiana e também se apresenta a metodologia Monte Carlo via Cadeias de Markov (MCMC) que é essencialmente integração de Monte Carlo usando Cadeias de Markov. Existem diversos métodos para simular cadeias de Markov, aqui apresentaremos os mais conhecidos o algoritmo Metropolis-Hastings e o algoritmo Gibbs sampling que em resumo é um caso particular do algoritmo Metropolis-Hastings.

No [Capítulo 3](#), define-se o modelo GARCH(p,q) que é a principal ferramenta nesta dissertação para alcançar o objetivo de modelar a volatilidade. Como também se estudam as propriedades da distribuição incondicional dos retornos, com base a estos resultados se propõe inserir assimetria e considerar distribuições com caudas pesadas na distribuição dos erros do modelo. Portanto, apresentaremos o método de [Fernández e Steel \(1998\)](#) utilizado neste trabalho para transformar a distribuição simétrica, continua e unimodal do erro em uma distribuição assimétrica ajustada por um parâmetro de forma que descreve o grau de assimetria. Para a estimação dos parâmetros do modelo aplicando a abordagem Bayesiana, se calcula a função de verossimilhança condicional e se estabelecem as distribuições a prioris dos parâmetros portanto

pelo teorema de Bayes obtemos a distribuição a posteriori conjunta dos parâmetros, fazendo uso da metodologia MCMC calculamos as médias amostrais para aproximar as medias populacionais obtendo assim os parâmetros estimados.

No [Capítulo 4](#), explica-se detalhadamente como se implementou o algoritmo e se faz a avaliação do desempenho do mesmo trabalhando com dados artificias.

No [Capítulo 5](#), apresenta-se a aplicação do modelo GARCH(1,1) com distribuição nos erros T-Student, Normal e GED assimétricas e padronizadas a duas séries de dados reais: IBOVESPA e S&P500. A inferência dos modelos é baseada nas amostras geradas pelo primeiro algoritmo definido no [Capítulo 4](#).

Para finalizar, no [Capítulo 6](#) faz-se uma avaliação dos resultados obtidos e apresentamos as conclusões do trabalho apresentado.

Por último, se apresentam as referências bibliográficas.

SÉRIES DE TEMPO FINANCEIRAS

Neste capítulo são apresentados alguns conceitos básicos que serão usados no transcorrer dos próximos capítulos. Assim, com respeito aos conceitos relacionados a séries temporais em finanças na [seção 2.1](#) se define o termo **retorno**, na [seção 2.2](#) se apresentam as principais características das séries de retornos conhecidas como **fatos estilizados**, através da análise de uma série com dados reais. A [seção 2.3](#) contém noções básicas de inferência Bayesiana e finalmente na [seção 2.4](#) se definem os métodos de Monte Carlo via Cadeias de Markov (MCMC).

A análise de séries temporais financeiras está focada na avaliação de ativos ao longo do tempo. Uma característica chave que distingue a análise de séries temporais financeiras de outras análises de séries temporais é a volatilidade, um elemento de incerteza que não é diretamente observável. Como resultado dessa incerteza adicional, se recorre aos chamados modelos heterocedásticos condicionais.

Começamos com os conceitos básicos de retornos de ativos, pois a maioria dos estudos financeiros envolve retornos, em vez de preços de ativos. Tomando como referência o livro de [Tsay \(2005\)](#), existem dois motivos principais para o uso de retornos. O primeiro é que para investidores médios, o retorno de um ativo é um resumo completo e sem escala da oportunidade de investimento. O segundo motivo é porque as séries de retornos são mais fáceis de analisar do que as séries de preços porque as primeiras possuem melhores propriedades estatísticas.

2.1 Retornos

Os dados financeiros, como os preços diários de um determinado ativo financeiro formam uma série de tempo financeira. Estes ativos financeiros podem ser por exemplo: ações de uma empresa ou comódites. Uma análise estatística direta dos preços dos ativos é difícil, porque os preços consecutivos são altamente correlacionados e as variâncias dos preços variam com o tempo. Portanto, as séries dos preços em geral não são estacionárias. Entendendo-se de modo bastante

geral que uma série de tempo diz-se estacionária se não houver mudança sistemática na média (sem tendência), se não houver mudança sistemática na variância e se as variações estritamente periódicas foram removidas. A definição formal de estacionariedade pode se encontrar no livro de [Chatfield \(2013\)](#). Então, assumindo que o intervalo de tempo entre observações consecutivas da série de preços é constante e que não haja dividendos pagos no período. O tipo de variação usualmente utilizada pelos pesquisadores na análise deste tipo de séries e que também será usada neste trabalho é o **retorno composto continuamente** (ou também chamado simplesmente **log-retorno** ou **retorno**) é definido como:

$$\log\left(\frac{P_t}{P_{t-1}}\right) = \log P_t - \log P_{t-1} = p_t - p_{t-1},$$

considerando a P_t o preço de um ativo no instante de tempo t (comumente um dia de negócio) e p_t como o logaritmo na base e do preço deste ativo ($p_t = \log P_t$).

Para um análise mais aprofundado da definição de retorno podem se estudar os livros de [Morettin \(2006\)](#) e de [Taylor \(2008\)](#).

Em seguida, a título de ilustração tomaremos a série dos valores de fechamento diário do Índice da Bolsa de Valores de São Paulo (**IBOVESPA**) no período de 2 de Janeiro de 2009 a 31 de dezembro de 2014 com um total de 1498 observações, que é considerado o mais importante indicador do desempenho médio das cotações das ações negociadas na Bolsa de Valores de São Paulo. A [Figura 1](#) e a [Figura 2](#), apresentam respectivamente a evolução dos preços e dos retornos de fechamento diário do Ibovespa.

A seguir, na [Tabela 1](#) se apresentam as estatísticas descritivas dos retornos do Ibovespa. Onde pode-se constatar que os retornos apresentam excesso de curtose amostral por ser maior do que 3 (valor referencial da curtose da distribuição normal), indicando que a distribuição da série apresenta caudas pesadas. Além disso, apresenta pequena assimetria negativa e média próxima de zero.

Tabela 1 – Estatística descritiva da série de retornos.

Estatística	Valor
média ($\times 10^{-4}$)	1,45
máximo	0,06
mínimo	-0,08
variância ($\times 10^{-4}$)	2,29
curtose	4,75
assimetria	-0,03
número de observações	1497

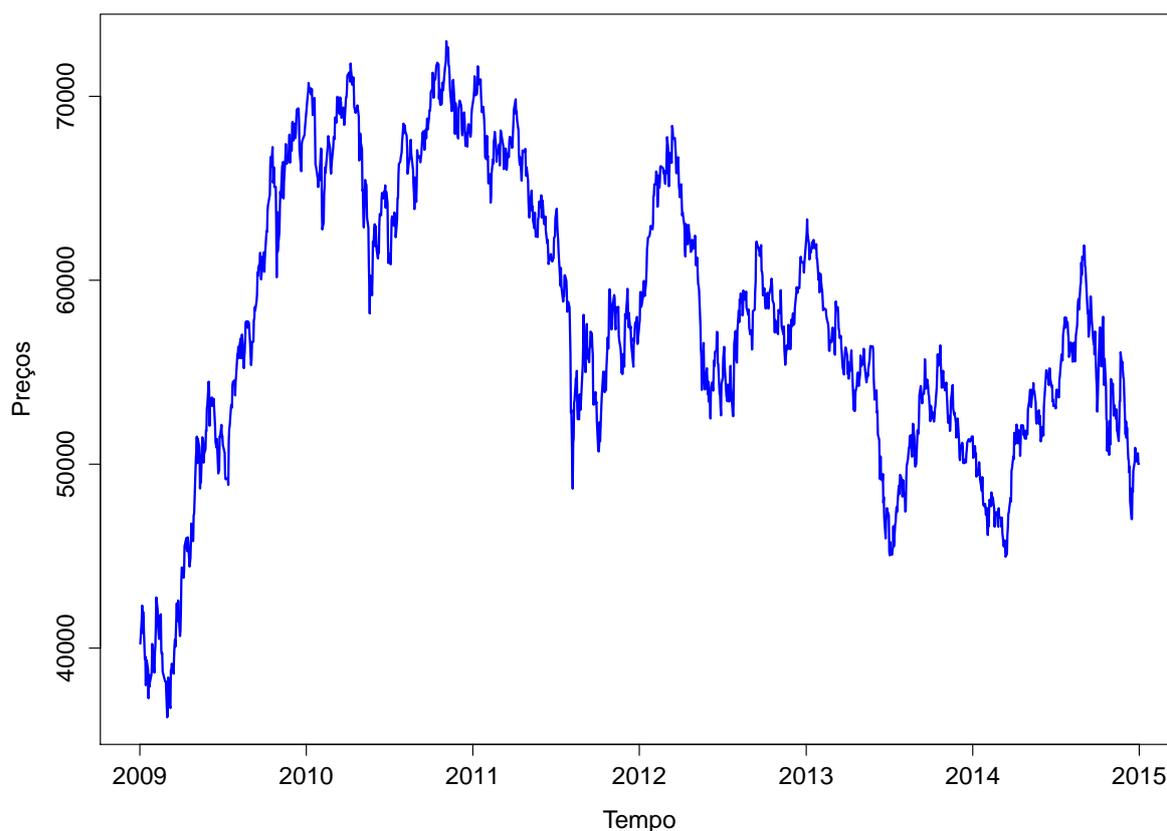


Figura 1 – Preços de Fechamento diários do Ibovespa desde Janeiro 2, 2009 até Dezembro 31, 2014.

2.2 Fatos Estilizados dos Retornos

A continuação os principais fatos estilizados dos retornos, serão apresentados. Para um estudo mais aprofundado veja [Straumann \(2005\)](#).

2.2.1 Não auto-correlacionados

Na análise usual de séries temporais, a dependência entre as observações consecutivas é frequentemente medida através das autocovariâncias ou autocorrelações, para mais detalhes veja [Chatfield \(2013\)](#). As contrapartidas amostrais de ditas quantidades são as autocovariâncias e autocorrelações amostrais. Assim, analisando a série dos retornos, geralmente apresentam o seguinte comportamento peculiar:

- Os retornos apresentam autocorrelações amostrais pequenas, não significativas nas primeiras defasagem.
- Existe autocorrelação amostral significativa nos quadrados dos retornos.

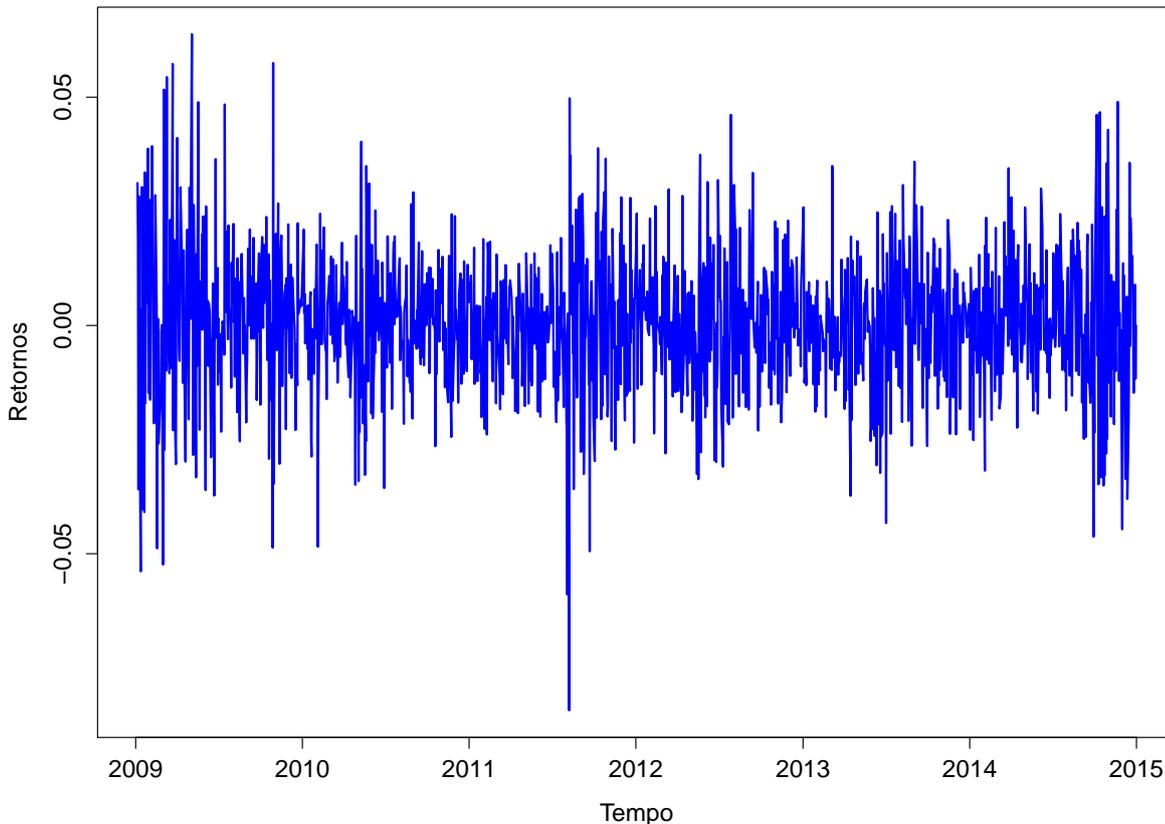


Figura 2 – Retornos diários do Ibovespa desde Janeiro 2, 2009 até Dezembro 31, 2014.

- As autocorrelações amostrais dos log-retornos absolutos são diferentes de zero para um grande número de defasagens.

Essas características são exemplificadas na [Figura 3 \(a\)](#) e na [Figura 3 \(b\)](#), considerando série dos retornos do Ibovespa. A [Figura 3 \(a\)](#) mostra que não existe autocorrelação dos retornos em todos os defasagens e a [Figura 3 \(b\)](#) mostra que as autocorrelações do quadrado dos retornos são diferentes de zero para diversos defasagens.

2.2.2 Agrupamentos de volatilidades ao longo do tempo

Na realidade, os mercados financeiros reagem nervosamente na presença de distúrbios políticos, crises econômicas, guerras, catástrofes naturais ou ante qualquer acontecimento que se creia uma ameaça para a sociedade humana. Durante esses períodos de estresse os preços dos ativos financeiros variam fortemente, comparados com os preços em períodos normais. De modo que, numa série de retornos a volatilidade, entendida como o grau de variabilidade dos mercados, evolui no tempo. Para exemplificar, na [Figura 2](#), a simples vista é possível detectar períodos de stress da série de tempo do Ibovespa. São os períodos onde a volatilidade dos retornos se

incrementa, por exemplo, no gráfico se observa uma grande instabilidade dos retornos a finais do ano 2011 que foi o período da crise econômica que atravessou Brasil reflexo da crise na Europa.

Em termos estatísticos, a variância de um retorno num dado instante de tempo depende de retornos passados e de outras informações disponíveis até aquele instante, de modo que temos que definir a variância condicional dado o passado, que é variável no tempo e não coincide com a variância global (incondicional) da série observada. Portanto, para se levar em conta a presença de grupos de volatilidade em uma série financeira é necessário recorrer a modelos ditos condicionalmente heterocedásticos.

2.2.3 *Distribuição incondicional assimétrica e de caudas pesadas*

Quando a distribuição incondicional empírica dos retornos é desenhada, podemos observar que geralmente não se assemelha a uma distribuição normal. Testes clássicos normalmente levam à rejeição da suposição de normalidade. Mais precisamente, as densidades tem caudas pesadas (decrecendo a zero mais lentamente do que $\exp\{-x^2/2\}$) e estão acentuadamente atingindo ao pico em zero: elas são chamadas leptocúrticas. Para medir o grau de leptocurtose pode-se utilizar o coeficiente de curtose, definido como a razão entre o momento de quarto ordem amostral e a variância amostral ao quadrado. A curtoses é igual a 3 para observações independentes e identicamente distribuídas com distribuição normal, este coeficiente é muito maior do que 3 para a série de retornos. Por exemplo para a série de retornos apresentada na [Figura 2](#), a [Figura 3 \(c\)](#) compara uma estimação do núcleo da densidade dos retornos (linha tracejada vermelha) com a densidade de uma distribuição normal com média e variância igual à media e variância dos retornos (linha azul). O pico em torno de zero aparece claramente, mas a espessura das caudas é mais difícil de visualizar. A existência de caudas pesadas nesta série pode se visualizar na [Figura 3 \(d\)](#), em que é apresentado o gráfico quantil-quantil da normal. Observe que os pontos caem ao longo da linha no meio do gráfico, mas curvam-se nas extremidades. O qual é um indicador de que a série de retornos têm valores mais extremos do que seria esperado se realmente vieram de uma distribuição Normal. Além disso, existe evidência empírica que a distribuição incondicional dos retornos pode mostrar frequentemente algum grau de assimetria. Por exemplo, no caso dos dados analisados anteriormente há indícios de assimetria negativa (fraca) nos retornos de Ibovespa como é descrito na [Tabla 1](#).

2.2.4 *Efeito de alavanca*

O efeito de alavanca, assinalado por [Black \(1976\)](#), faz referência à assimetria do impacto dos retornos passados positivos e negativos sobre a volatilidade atual. Os retornos negativos, que correspondem a diminuições dos preços, tendem a aumentar a volatilidade por uma quantidade maior do que os retornos positivos da mesma magnitude, que correspondem a aumentos dos preços. Mas no modelo GARCH, a variância condicional depende do quadrado dos retornos passados então erros passados positivos e negativos têm o mesmo efeito na volatilidade atual,

portanto não é capaz de explicar este fato estilizado. Por causa deste efeito, surgiram outros modelos com a capacidade de captar a assimetria na volatilidade, tais como os modelos EGARCH (*Exponential GARCH*), TARCH (*Threshold ARCH*) e APARCH (ARCH com potência assimétrica). Em estes tipos de modelos a assimetria é tomada em conta na formulação do modelo, ou seja a assimetria é considerada no processo gerador das volatilidades.

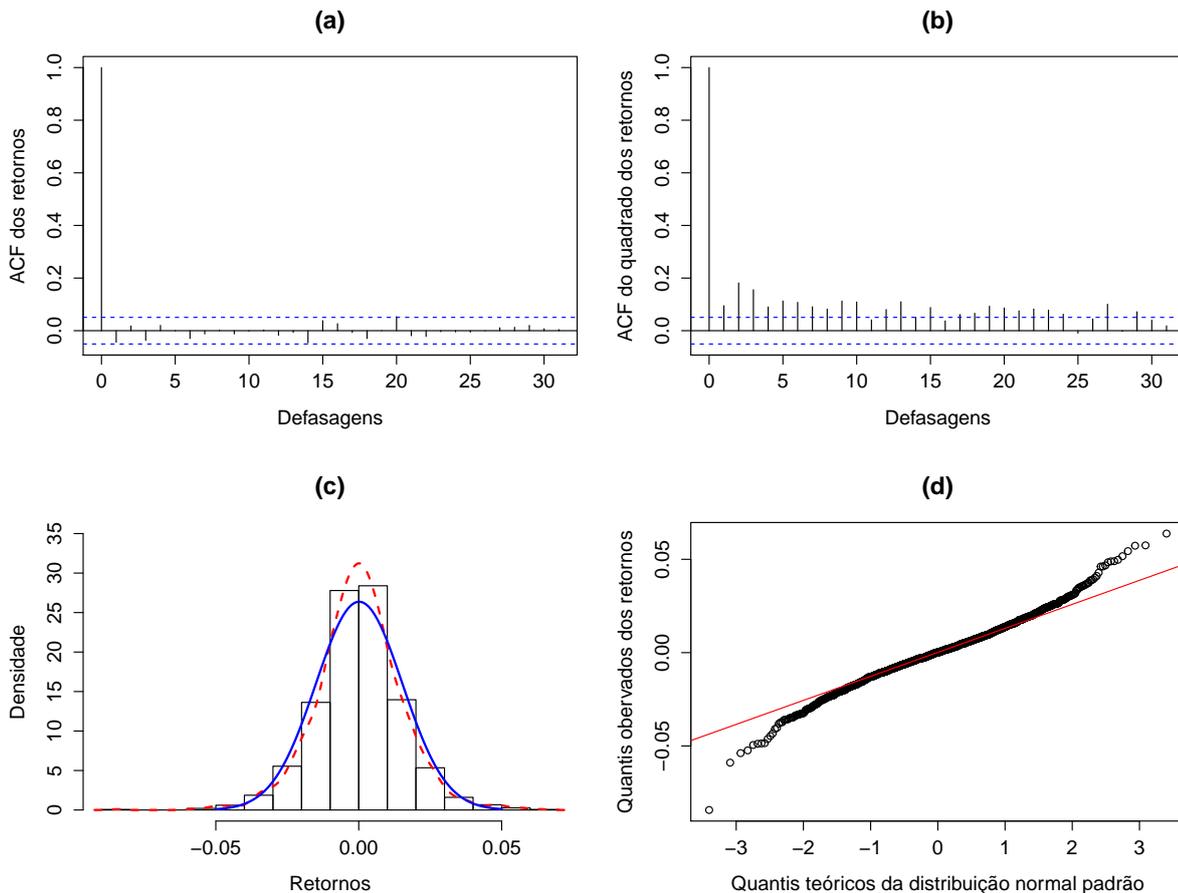


Figura 3 – Características dos retornos do Ibovespa.

Os fatos estilizados das séries dos retornos que acabamos de apresentar, podem ser atribuídos a [Mandelbrot \(1963\)](#) e a [Fama \(1965\)](#). Com base na análise de muitas séries de retornos de ações, mostraram ditas características apresentadas nas séries de variações de preços de dados financeiros. Nos seguintes capítulos, vamos apresentar e implementar o modelo GARCH considerando certo grau de assimetria na distribuição dos erros, com o fim de capturar a assimetria na distribuição dos retornos que segundo estudos empíricos não pode ser capturada pelo tipo de modelos apresentados no parágrafo anterior.

Na próxima seção, se apresenta um resumo de inferência Bayesiana. Observe que usaremos os termos *distribuição* e *densidade* alternadamente.

2.3 Inferência Bayesiana

Nesta seção apresentamos noções básicas da inferência Bayesiana. Para um estudo aprofundado deste tema sugerimos estudar os livros de [Casella e Berger \(2010\)](#) e [Migon, Gamerman e Louzada \(2014\)](#). A diferença entre a abordagem Clássica e Bayesiana reside na natureza matemática de $\boldsymbol{\theta}$, o vetor de parâmetros. Na abordagem Clássica, assume-se que existe um valor verdadeiro e fixo para $\boldsymbol{\theta}$. Por outro lado, na abordagem Bayesiana considera-se a $\boldsymbol{\theta}$ como uma variável aleatória a qual é caracterizada por uma **distribuição a priori** denotada por $p(\boldsymbol{\theta})$ que representa a incerteza inicial acerca do vetor de parâmetros antes de ter observado \mathbf{y} . Sendo \mathbf{y} um vetor de observações cujos valores inicialmente são incertos e descritos através de uma densidade de probabilidade $p(\mathbf{y}|\boldsymbol{\theta})$, conhecida como a **função de verossimilhança** do modelo. Se assumirá a situação canônica em que as observações de uma amostra aleatória do vetor são identicamente distribuídas como $p(\mathbf{y}|\boldsymbol{\theta})$ e independentes (condicionalmente ao conhecimento de $\boldsymbol{\theta}$). Com a especificação de $p(\boldsymbol{\theta})$ e $p(\mathbf{y}|\boldsymbol{\theta})$, podemos obter o modelo probabilístico

$$p(\boldsymbol{\theta}, \mathbf{y}) = p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}).$$

Uma vez que os dados \mathbf{y} contêm informação acerca de $\boldsymbol{\theta}$, pode-se usar \mathbf{y} para atualizar a informação que temos acerca de $\boldsymbol{\theta}$. Portanto, usando o teorema de Bayes determina-se a **distribuição a posteriori** $p(\boldsymbol{\theta}|\mathbf{y})$, a qual é uma descrição probabilística do conhecimento de $\boldsymbol{\theta}$ depois de observar os dados, e é dada por

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int_{\Theta} p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) d\boldsymbol{\theta}}, \quad (2.1)$$

na maioria das aplicações, a integral do denominador da [Ecuación 2.1](#) não possui forma analítica fechada. Mas como pode-se ver a função no denominador da [Ecuación 2.1](#) não depende de $\boldsymbol{\theta}$, o qual leva a

$$p(\boldsymbol{\theta}|\mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}). \quad (2.2)$$

Em muitos cenários não é possível ou não é desejável trabalhar com distribuições a prioris conjugadas ou com a prioris que conduzam a uma forma analítica fechada da distribuição a posteriori. Nestes casos, pode ser possível utilizar aproximações analíticas ou numéricas para a distribuição a posteriori. Outra alternativa consiste em resumir a inferência obtendo amostras aleatórias da distribuição a posterior. Às vezes, é possível obter essas amostras por simulação direta, mas muitas vezes esse não é o caso e, portanto, métodos como Monte Carlo via cadeias Markov (MCMC) são usados.

Na seguinte seção apresentamos os principais métodos de simulação de Monte Carlo com Cadeias de Markov (MCMC).

2.4 Métodos MCMC

Estes métodos baseados em simulação estocástica geram amostras da distribuição a posteriori (nossa distribuição objetivo) quando o número de realizações é grande. Para um estudo mais detalhado dos métodos MCMC ver o livro de [Gamerman e Lopes \(2006\)](#).

A amostragem MCMC baseia-se na construção de uma cadeia de Markov ergódica com distribuição estacionária $p(\boldsymbol{\theta}|\mathbf{y})$ e uma função de transição $p(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$. Uma Cadeia de Markov é dita ser ergódica se todos seus estados são estados ergódicos, entenda-se que um estado ergódico é um estado recorrente que é aperiódico. Sendo $\boldsymbol{\theta}^{[0]}, \boldsymbol{\theta}^{[1]}, \dots, \boldsymbol{\theta}^{[j]}, \dots$ as realizações da cadeia, cujo espaço de estados será o espaço paramétrico Θ . Quando o valor de j aumenta as realizações se tornam cada vez mais próximas das realizações da distribuição estacionária $p(\boldsymbol{\theta}|\mathbf{y})$ e podem ser consideradas amostras de $p(\boldsymbol{\theta}|\mathbf{y})$. Para um conhecimento mais aprofundado deste tema ler o capítulo 6 de [Robert \(2004\)](#).

Uma vez que uma amostra de $p(\boldsymbol{\theta}|\mathbf{y})$ é obtida pode se calcular estimativas das características da distribuição a posteriori. Por exemplo, a média a posteriori de $\boldsymbol{\theta}$ é estimada pela média aritmética dos valores amostrais, as médias a posterioris de transformações $t(\boldsymbol{\theta})$ são estimadas pela média das transformações $t(\boldsymbol{\theta}^{[j]})$ dos valores amostrais. Este método é usado quando os métodos de amostragem direto ou indireto da distribuição a posteriori não podem ser aplicados, portanto a simulação por cadeias de Markov é capaz de aproximar a amostragem de praticamente qualquer distribuição a posteriori.

A seguir serão apresentados os métodos MCMC mais comumente utilizados, o algoritmo Gibbs sampling e o algoritmo Metropolis-Hastings.

2.4.1 Gibbs sampling

[Geman e Geman \(1984\)](#) são os criadores deste método no contexto de processamento de imagens. Mas, [Gelfand e Smith \(1990\)](#) foram os primeiros autores em apontar com sucesso para a comunidade estatística em geral que o esquema de amostragem elaborado por [Geman e Geman \(1984\)](#) para as distribuições de Gibbs, poderia de fato ser usado para uma série de outras distribuições a posteriores. Neste algoritmo MCMC a função de transição da cadeia está formada pelas distribuições condicionais completas $p(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k, \mathbf{y})$, $i = 1, \dots, k$, considerando que são conhecidas e podem ser amostradas.

O esquema de amostragem baseado em gerações sucessivas é descrito como segue:

1. Inicializar o contador de iteração da cadeia para $j = 1$ e especificar um valor inicial $\boldsymbol{\theta}^{[0]} = (\theta_1^{[0]}, \dots, \theta_k^{[0]})'$;
2. Dado $\boldsymbol{\theta}^{[j-1]}$, gerar o próximo valor $\boldsymbol{\theta}^{[j]} = (\theta_1^{[j]}, \dots, \theta_k^{[j]})'$ através da simulação sucessiva

dos valores:

$$\begin{aligned}\theta_1^{[j]} &\sim p(\theta_1|\theta_2^{[j-1]}, \dots, \theta_k^{[j-1]}, \mathbf{y}), \\ \theta_2^{[j]} &\sim p(\theta_2|\theta_1^{[j]}, \theta_3^{[j-1]}, \dots, \theta_k^{[j-1]}, \mathbf{y}), \\ &\vdots \\ \theta_k^{[j]} &\sim p(\theta_k|\theta_1^{[j]}, \dots, \theta_{k-1}^{[j]}, \mathbf{y});\end{aligned}$$

3. Alterar o contador de j para $j + 1$, voltar para o passo 2 e continuar o procedimento até alcançar a convergência.

Enquanto o número de iterações aumenta, espera-se que a cadeia se aproxime a sua distribuição estacionária. Após a convergência, os valores resultantes formam uma amostra de $p(\boldsymbol{\theta}|\mathbf{y})$.

Observe que não é necessário conhecer a distribuição conjunta, mas é necessário conhecer as distribuições condicionais completas. Se as distribuições condicionais completas coincidirem com alguma distribuição de probabilidade conhecida na literatura a partir da qual pode-se facilmente gerar os valores, então deve-se usar o Gibbs sampling. No caso, de uma ou mais distribuições condicionais completas não coincidirem com alguma distribuição de probabilidade conhecida, podem-se utilizar outros métodos de simulação como o algoritmo Metropolis-Hastings para simular a distribuição conjunta.

2.4.2 Metropolis-Hastings

Aqui apresentaremos as cadeias de Markov geradas pelo algoritmo Metropolis-Hastings (M-H), cujo nome deriva dos dois artigos básicos para a caracterização deste método [Metropolis et al. \(1953\)](#) e [Hastings \(1970\)](#). Este algoritmo é usado geralmente quando a geração não iterativa de $p(\boldsymbol{\theta}|\mathbf{y})$ é muito complicada ou custosa.

Então, considerando uma distribuição $p(\boldsymbol{\theta}|\mathbf{y})$ a partir da qual uma amostra será extraída via cadeias de Markov. A função de transição da cadeia $p(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ onde $\boldsymbol{\theta}^*$ é um novo estado da cadeia, deve ser construída de tal forma que $p(\boldsymbol{\theta}|\mathbf{y})$ seja a distribuição estacionária da cadeia. Isto se reduz a encontrar cadeias reversíveis onde a função de transição satisfaça a condição de reversibilidade da cadeia. [Hastings \(1970\)](#) propôs definir a probabilidade de aceitação $\alpha(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ de tal maneira que quando combinada com $q(\boldsymbol{\theta}^*|\boldsymbol{\theta})$ chamada de distribuição proposta se defina uma cadeia reversível.

Em termos práticos, a simulação de uma amostra de $p(\boldsymbol{\theta}|\mathbf{y})$ usando a cadeia de Markov definida por este método tem o seguinte esquema geral:

1. Inicializar o contador de iteração para $j = 1$ e especificar um valor inicial $\boldsymbol{\theta}^{[0]}$;

2. Gerar um novo valor candidato θ^* a partir da densidade proposta $q(\cdot|\theta^{[j-1]})$;
3. Calcular a probabilidade de aceitação $\alpha(\theta^{[j-1]}, \theta^*)$ e gerar $u \sim U(0, 1)$;
4. Se $u < \alpha(\theta^{[j-1]}, \theta^*)$ então o novo valor é aceito e fazemos $\theta^{[j]} = \theta^*$, caso contrario rejeite e faça $\theta^{[j]} = \theta^{[j-1]}$ de modo que a cadeia não se mova;
5. Alterar o contador de j para $j + 1$, voltar para o passo 2 e continuar o procedimento até alcançar a convergência.

Observe que a distribuição proposta $q(\cdot|\theta^{[j-1]})$ é usada para gerar um novo valor θ^* dado o valor atual $\theta^{[j-1]}$ e $\alpha(\theta^{[j-1]}, \theta^*)$ é a probabilidade de aceitação do movimento da cadeia de $\theta^{[j-1]}$ a θ^* , que é dada por

$$\alpha(\theta^{[j-1]}, \theta^*) = \min \left\{ \frac{p(\theta^*|\mathbf{y})q(\theta^{[j-1]}|\theta^*)}{p(\theta^{[j-1]}|\mathbf{y})q(\theta^*|\theta^{[j-1]})}, 1 \right\}.$$

Embora a distribuição proposta q possa ser escolhida arbitrariamente, na prática a escolha da proposta é crucial para a convergência da cadeia. A cadeia deve sempre ser capaz de atravessar todo o espaço de parâmetros para convergir à distribuição de equilíbrio. Os movimentos da cadeia determinados por q , devem ser estabelecidos de tal maneira que forneçam deslocamentos consideráveis do estado atual mas com uma probabilidade substancial, determinada por α , de ser aceitos.

Existem casos especiais do algoritmo quanto à seleção da distribuição proposta, que são apresentados a seguir:

- **Cadeias independentes**

O algoritmo resultante é conhecido como o *algoritmo Metropolis-Hastings independente*, quando a distribuição proposta não depende do valor do estado atual da cadeia, *i.e.*, $q(\theta^*|\theta^{[j-1]}) = q(\theta^*)$.

- **Cadeias simétricas**

O algoritmo resultante é conhecido como o *algoritmo de Metropolis*, quando a distribuição proposta é simétrica em torno da iteração anterior, *i.e.*, $q(\theta^*|\theta^{[j]}) = q(\theta^{[j]}|\theta^*)$ para todos os valores de θ^* e $\theta^{[j]}$. Este é o algoritmo original proposto por [Metropolis et al. \(1953\)](#).

- **Cadeias com passeio aleatório**

O algoritmo resultante é conhecido como o *algoritmo Metropolis com passeio aleatório*, é um caso especial do algoritmo de Metropolis baseado em um passeio aleatório. Consiste de uma distribuição proposta a qual só depende da distância entre θ^* e $\theta^{[j-1]}$, *i.e.*, $q(\theta^*|\theta^{[j-1]}) = q(|\theta^{[j-1]} - \theta^*|)$.

Uma característica importante deste algoritmo é que só precisamos conhecer a distribuição de interesse parcialmente, já que a probabilidade de aceitação não se altera. Isto é fundamental em aplicações Bayesianas quando não se conhece completamente a posteriori.

Finalizando, observe-se que na forma do que o algoritmo M-H foi apresentado, o vetor $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)'$ é atualizado em um único bloco em cada iteração de modo que todos os elementos são alterados simultaneamente. Um caso mais geral do algoritmo e que em alguns casos pode ser computacionalmente mais eficiente, consiste em dividir $\boldsymbol{\theta}$ em p blocos $\{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_p\}$, cada bloco pode conter um ou mais elementos e é gerado por sua própria distribuição proposta, assim dentro de cada iteração teremos o algoritmo aplicado p vezes. Note que o Amostrador de Gibbs é um caso especial do algoritmo M-H por blocos, onde cada componente de $\boldsymbol{\theta}$ é atualizado sequencialmente e onde as distribuições propostas são as distribuições condicionais completas. Para este caso, os valores candidatos são sempre aceitos.

2.4.3 Avaliação da Convergência

Conforme é demonstrado na teoria, uma amostra da distribuição de interesse $p(\boldsymbol{\theta}|\mathbf{y})$ só é obtida quando o número de iterações da cadeia é infinito. Do ponto de vista da implementação dos métodos, isso não é alcançável e um valor obtido em uma iteração suficientemente grande é tomado. Devido às dificuldades na determinação de quão grande essa iteração deveria de ser, se optou por estudar as características de convergência da cadeia. Havendo surgido abordagens teóricas e empíricas que tentam decidir se a convergência pode ser seguramente assumida. Os métodos empíricos que fazem a análise das propriedades estatísticas da cadeia observada para diagnosticar a convergência são os mais usados na literatura.

Para avaliar a convergência das cadeias observadas neste trabalho se faz uso do método de Geweke (1992), que sugeriu a utilização de testes sobre as médias do parâmetro amostrado em dois intervalos de tempo diferentes da cadeia para verificar a convergência, com base nas séries temporais definidas pela sequência de valores amostrados do parâmetro ou qualquer transformação real dele. Este método ao igual que outros estão implementados no pacote CODA de Cowles e Carlin (1996) executável no software R.

Após a execução da cadeia e sua convergência com a distribuição estacionária. Como a cadeia não é inicializada na distribuição estacionária, uma prática comum é usar um período de aquecimento, ver Gilks, Richardson e Spiegelhalter (1996). A cadeia é rodada por $L + M$ iterações, sendo as primeiras L iterações iniciais descartadas (*burn-in* ou período de aquecimento). Espera-se que depois deste período de aquecimento a cadeia tenha esquecido os valores iniciais e convergido para a distribuição de equilíbrio. Assim, a amostra resultante de tamanho M , será uma amostra da distribuição de equilíbrio. Logo, para eliminar as possíveis autocorrelações da cadeia seleciona-se uma nova amostra cada k iterações a partir da amostra resultante. Portanto,

uma amostra $\{\boldsymbol{\theta}^{[j]}\}_{j=1}^J$ da densidade a posteriori conjunta $p(\boldsymbol{\theta}|\mathbf{y})$ está disponível, onde J indica o tamanho da amostra final.

De forma que, pela integração de Monte Carlo, podemos aproximar o esperado posteriori de qualquer função dos parâmetros do modelo $f(\boldsymbol{\theta})$ pela média amostral dos dados gerados na simulação,

$$\mathbb{E}_{\boldsymbol{\theta}|\mathbf{y}} [f(\boldsymbol{\theta})] \approx \frac{1}{J} \sum_{j=1}^J f(\boldsymbol{\theta}^{[j]}).$$

Sob certas condições, a média amostral converge quase certamente ao valor esperado a posteriori pela lei forte dos grandes números, mesmo que as extrações sejam geradas por um método MCMC, ver Tierney (1994). Alguns casos particulares da função dos parâmetros permitem obter características da distribuição a posteriori conjunta. Por exemplo, quando $f(\boldsymbol{\theta}) = \boldsymbol{\theta}$ obtém-se o vetor de médias a posteriori $\bar{\boldsymbol{\theta}}$; para $f(\boldsymbol{\theta}) = (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}) (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})'$, sendo $\bar{\boldsymbol{\theta}}$ a média amostral, obtemos a matriz de variâncias a posteriori.

MODELOS GARCH

Os modelos ARCH (*Autoregressive Conditional Heteroscedasticity*) foram introduzidos por [Engle \(1982\)](#) e sua generalização o modelo GARCH (*Generalized Autoregressive Conditional Heteroscedasticity*) foi sugerida por [Bollerslev \(1986\)](#). Este tipo de modelos reconhece explicitamente a diferença entre a **variância incondicional** e a **variância condicional** (isto é, variância condicionada ao passado), permitindo a esta última mudar ao longo do tempo. Nos modelos GARCH, a variância condicional é expressada como uma função linear do quadrado dos valores passados da série e de seus variâncias condicionais passadas, a evidência empírica mostra que este modelo permite modelar a volatilidade de maneira mais parcimoniosa (isto é, com menor número de parâmetros) comparado com o modelo ARCH. Além disso, é capaz de capturar alguns dos principais fatos estilizados que caracterizam às séries de retornos financeiros, tal como foram descritos no [Capítulo 2](#).

Uma grande quantidade de pesquisas teóricas e empíricas sobre estes modelos foram produzidas com o transcorrer do tempo. A maior parte de trabalhos desenvolvidos foram no contexto da inferência clássica, e se basearam no método de (quase-) verossimilhança e no método generalizado dos momentos, ver [Bollerslev, Chou e Kroner \(1992\)](#)). Mais recentemente, os métodos MCMC têm sido usados na abordagem Bayesiana. Assim por exemplo, [Bauwens e Lubrano \(1998\)](#) usaram o Gibbs sampler para realizar inferência Bayesiana nos modelos GARCH e [Nakatsuma \(2000\)](#) propôs um novo método MCMC para a inferência e estimação do modelo ARCH ou GARCH.

A seguir, o restante do capítulo está organizado da seguinte forma. A [seção 3.1](#) explica brevemente o modelo GARCH e algumas de suas propriedades básicas. A [seção 3.2](#) apresenta o cálculo da função de verossimilhança e as distribuições a priori consideradas para fazer inferência Bayesiana.

3.1 Modelo GARCH

O modelo GARCH(p,q) define a variância condicional no tempo t de y_t (retorno de um ativo ou de um índice financeiro no tempo t), denotado por h_t , como uma função linear dos quadrados dos q retornos passados e das p variâncias condicionais passadas.

Considerando uma série de retornos $\mathbf{y} = \{y_t, t = 1, \dots, T\}$, o modelo tem a seguinte expressão:

$$y_t = \varepsilon_t \sqrt{h_t}, \quad \varepsilon_t \sim D(0, 1) \quad (3.1)$$

$$h_t = \omega + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \quad (3.2)$$

em que h_t é a variância condicional (não observável) de y_t dada a informação prévia $I_{t-1} = \{y_{t-1}, y_{t-2}, \dots\}$, os ε_t são independentes e identicamente distribuídos e $D(0, 1)$ denota a distribuição com média 0 e variância 1. As restrições para garantir a positividade de h_t e estacionariedade do processo são respetivamente $\omega > 0$, $\alpha_i \geq 0$ com $i = 1, \dots, p$, $\beta_j \geq 0$ com $j = 1, \dots, q$ e $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$.

Aplicando a definição obtém-se que a esperança e variância condicional y_t são dados por:

$$\begin{aligned} E[y_t | I_{t-1}] &= \sqrt{h_t} E[\varepsilon_t | I_{t-1}] = \sqrt{h_t} E[\varepsilon_t] = 0 \\ \text{Var}[y_t | I_{t-1}] &= E[y_t^2 | I_{t-1}] - E^2[y_t | I_{t-1}] = E[y_t^2 | I_{t-1}] = h_t \text{Var}[\varepsilon_t | I_{t-1}] = h_t. \end{aligned}$$

Logo, podemos utilizar os resultados anteriores para calcular a esperança e variância incondicional de y_t :

$$\begin{aligned} E[y_t] &= E[E[y_t | I_{t-1}]] = E[0] = 0 \\ \text{Var}[y_t] &= E[\text{Var}[y_t | I_{t-1}]] + \text{Var}[E[y_t | I_{t-1}]] = E[h_t] + \text{Var}[0] = E[h_t] = E[y_t^2] \end{aligned}$$

Pela demonstração apresentada no apêndice A1 pagina 323 do artigo de [Bollerslev \(1986\)](#), temos que,

$$E[y_t^2] = \frac{\omega}{1 - \left(\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j\right)}. \quad (3.3)$$

Da [Ecuación 3.3](#), entende-se que para existir a variância incondicional de y_t é necessário que $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$, garantindo assim a estacionariedade do processo GARCH(p,q).

3.1.1 Propriedades da distribuição incondicional

A distribuição incondicional de y_t não é conhecida explicitamente. A seguir, apresentamos algumas propriedades desta distribuição fazendo uso de seus momentos.

Uma forma fácil de medir o peso das caudas da distribuição é usando o coeficiente de curtose. Sendo definido, para uma distribuição centrada (com média zero), como a razão do momento de quarta ordem e o quadrado do momento de segunda ordem. Este coeficiente é igual a 3 numa distribuição Normal, o qual é tomado como referência na comparação com outras distribuições.

No modelo GARCH(p,q) definido pela [Ecuación 3.1](#) e a [Ecuación 3.2](#), os momentos condicionais de ordem $2k$ são proporcionais a h_t^k :

$$E\left(y_t^{2k}|I_{t-1}\right) = h_t^k E\left(\varepsilon_t^{2k}\right),$$

observe que o coeficiente de curtose da distribuição condicional é proporcional ao coeficiente de curtose de ε_t . Assim, o coeficiente de curtose da distribuição incondicional é

$$k_y = \frac{E\left(y_t^4\right)}{\left\{\text{Var}\left(y_t\right)\right\}^2} = \frac{E\left\{E\left(y_t^4|I_{t-1}\right)\right\}}{\left\{E\left(y_t^2\right)\right\}^2} = \frac{E\left(h_t^2\right)}{\left\{E\left(y_t^2\right)\right\}^2} k_\varepsilon,$$

em que $k_\varepsilon = E\left(\varepsilon_t^4\right)$ denota o coeficiente de curtose de ε_t .

Considerando o modelo GARCH(1,1), com

$$y_t = \varepsilon_t \sqrt{h_t}, \quad \varepsilon_t \sim D(0,1) \quad (3.4)$$

$$h_t = \omega + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1}, \quad (3.5)$$

o cálculo de k_y depende da esperança do quarto momento incondicional e da esperança da variância incondicional do retorno. O cálculo do momento de quarto ordem é

$$E\left(y_t^4\right) = E\left(\varepsilon_t^4\right) E\left(h_t^2\right). \quad (3.6)$$

Logo, descompondo h_t^2

$$h_t^2 = \omega^2 + \alpha_1^2 y_{t-1}^4 + \beta_1^2 h_{t-1}^2 + 2\omega\alpha_1 y_{t-1}^2 + 2\omega\beta_1 h_{t-1} + 2\alpha_1\beta_1 y_{t-1}^2 h_{t-1}$$

Tomando esperado, lembrando que $y_{t-1}^2 = \varepsilon_{t-1}^2 h_{t-1}$, $E\left(\varepsilon_{t-1}^2\right) = 1$ e $E\left(y_{t-1}^2\right) = E\left(h_{t-1}\right)$

$$\begin{aligned} E\left(h_t^2\right) &= \omega^2 + \alpha_1^2 E\left(y_{t-1}^4\right) + \beta_1^2 E\left(h_{t-1}^2\right) + 2\omega\alpha_1 E\left(y_{t-1}^2\right) + 2\omega\beta_1 E\left(h_{t-1}\right) \\ &\quad + 2\alpha_1\beta_1 E\left(\varepsilon_{t-1}^2\right) E\left(h_{t-1}^2\right) \end{aligned}$$

$$= \omega^2 + \alpha_1^2 E\left(y_{t-1}^4\right) + \beta_1^2 E\left(h_{t-1}^2\right) + 2\omega\left(\alpha_1 + \beta_1\right) E\left(y_{t-1}^2\right) + 2\alpha_1\beta_1 E\left(h_{t-1}^2\right)$$

$$= \omega^2 + \alpha_1^2 E\left(y_{t-1}^4\right) + 2\omega\left(\alpha_1 + \beta_1\right) E\left(y_{t-1}^2\right) + \left(2\alpha_1\beta_1 + \beta_1^2\right) E\left(h_{t-1}^2\right)$$

Substituindo $E(y_t^2) = E(y_{t-1}^2) = \omega / \{1 - (\alpha_1 + \beta_1)\}$ e assumindo o quarto momento do retorno constante $E(y_t^4) = E(y_{t-1}^4) = \mu_4$ na [Ecuación 3.6](#), temos

$$\mu_4 = k_\varepsilon \left\{ \omega^2 + \alpha_1^2 \mu_4 + 2\omega(\alpha_1 + \beta_1) \frac{\omega}{1 - (\alpha_1 + \beta_1)} + (\beta_1^2 + 2\alpha_1\beta_1) \frac{\mu_4}{k_\varepsilon} \right\}$$

E assim

$$\mu_4 = \frac{\omega^2 (1 + \alpha_1 + \beta_1)}{(1 - k_\varepsilon \alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1) (1 - \alpha_1 - \beta_1)}$$

Portanto, a curtose k_y é igual a

$$\begin{aligned} k_y &= \frac{\omega^2 (1 + \alpha_1 + \beta_1)}{(1 - k_\varepsilon \alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1) (1 - \alpha_1 - \beta_1)} \frac{(1 - \alpha_1 - \beta_1)^2}{\omega^2} \\ &= \frac{1 - (\alpha_1 + \beta_1)^2}{1 - (\alpha_1 + \beta_1)^2 - \alpha_1^2 (k_\varepsilon - 1)} k_\varepsilon, \end{aligned}$$

note que cumprindo a restrição seguinte $\alpha_1^2 k_\varepsilon < 1 - (\alpha_1 + \beta_1)^2 + \alpha_1^2$, k_y será maior que k_ε . Então, pode-se concluir que no modelo GARCH(1,1) os retornos tem caudas mais pesadas do que a distribuição do erro ε_t .

O excesso do coeficiente de curtose de y_t e ε_t , com respeito ao coeficiente da distribuição Normal, é dado por

$$k_y^* = k_y - 3 = \frac{6\alpha_1^2 + k_\varepsilon^* \left\{ 1 - (\alpha_1 + \beta_1)^2 + 3\alpha_1^2 \right\}}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2 - k_\varepsilon^* \alpha_1^2},$$

onde, $k_\varepsilon^* = k_\varepsilon - 3$. O excesso da curtose de y_t se incrementa com o incremento do excesso da curtose do erro ε_t .

Observe que mesmo se $\varepsilon_t \sim N(0, 1)$ a distribuição incondicional de y_t não é Normal. Além disso, devemos tomar em conta que existe evidência na literatura que muitas séries de retornos financeiros tendem a ter a curtose observada ainda maior do que aquela implicada por um modelo GARCH com erros normais. Portanto, devemos de considerar erros com distribuições com caudas mais pesadas, assim pelo demonstrado anteriormente vai se incrementar a curtose da distribuição incondicional do retorno y_t . Sendo assim, alguns autores têm proposto distribuições com caudas mais pesadas do que a Normal para os erros ε_t . A distribuição mais comumente utilizada é a T-Student padronizada que tem um parâmetro de forma que modela a curtose, outra distribuição que vem sendo usada é a distribuição GED (*Generalized Error Distribution*) padronizada também conhecida como distribuição exponencial potência (ver [Nelson \(1991\)](#)), que ao igual que a distribuição T-Student depende de um parâmetro de forma que modela a curtose. A seguir, a [Figura 4](#) compara as caudas das distribuições que serão tomadas em conta para a distribuição do erro.

Uma das preocupações atuais, tem sido a evidência empírica de assimetria nos retornos financeiros, não tratada pelos modelos GARCH da proposta de [Bollerslev \(1986\)](#). Para levar

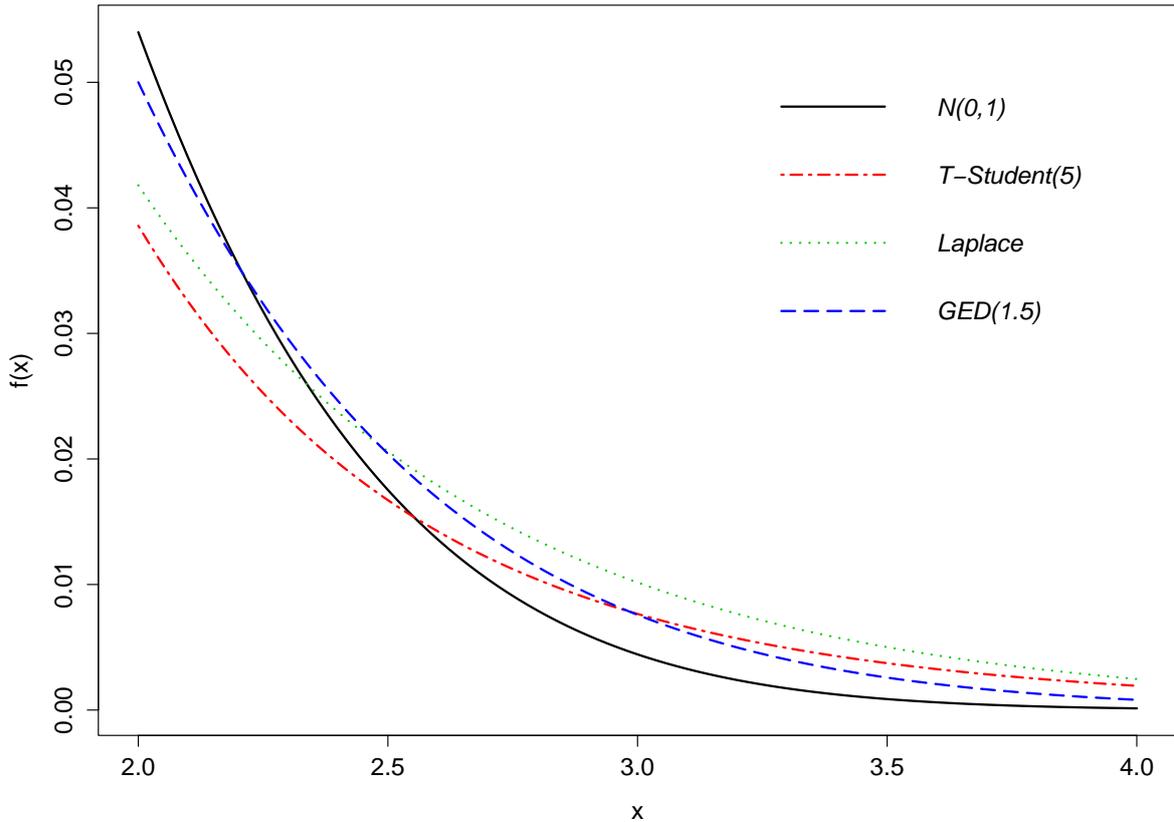


Figura 4 – Caudas da distribuição normal padrão, da T-Student padronizada com 5 graus de liberdade, da GED padronizada com $\delta = 1$ (Laplace) e com $\delta = 1.5$.

em conta esta característica adicional dos retornos se propôs permitir certo grau de assimetria na distribuição dos erros, como no artigo de Lambert e Laurent (2001). Portanto, as distribuições dos erros serão modificadas para que se tornem distribuições assimétricas. Então, vamos a ter as distribuições T-Student assimétrica, Normal assimétrica e GED assimétrica. Existem diversas propostas para inserir assimetria, mas devido a sua simplicidade e generalidade neste trabalho vamos a adotar a proposta de [Fernández e Steel \(1998\)](#), que apresentaram um método geral para transformar qualquer distribuição univariada, contínua, unimodal e simétrica (com respeito a 0) em uma assimétrica mudando a escala em cada lado da moda. Eles propõem a seguinte classe de de distribuições assimétricas indexadas por um parâmetro de forma $\gamma > 0$, o qual descreve o grau de assimetria,

$$s(x|\gamma) = \frac{2}{\gamma + 1/\gamma} \left\{ f(\gamma x) I_{(-\infty, 0)}(x) + f\left(\frac{x}{\gamma}\right) I_{[0, \infty)}(x) \right\}$$

Observe que $\gamma = 1$ produz a distribuição simétrica ou seja $s(x|\gamma = 1) = f(x)$, e os valores de $\gamma > 1$ ($\gamma < 1$) indicam assimetria à direita (esquerda). Além disso, a moda desta densidade permanece em zero independentemente do valor particular de γ . Os momentos de ordem r de

$s(x|\gamma)$ dependem de γ e são dados por,

$$E(X^r|\gamma) = m_r \frac{\gamma^{r+1} + (-1)^r/\gamma^{r+1}}{\gamma + 1/\gamma},$$

em que

$$m_r = 2 \int_0^\infty x^r f(x) dx, \quad (3.7)$$

é o r -ésimo momento absoluto de $f(x)$ sobre a reta real positiva. Portanto, não é difícil verificar que a média e variância de $s(x|\gamma)$ são dadas por

$$\begin{aligned} \mu_\gamma &= m_1(\gamma - 1/\gamma), \\ \sigma_\gamma^2 &= (m_2 - m_1^2)(\gamma^2 + 1/\gamma^2) + 2m_1^2 - m_2. \end{aligned} \quad (3.8)$$

Note também que, se $f(x)$ é uma distribuição padronizada, então $m_2 = 1$.

Para que o processo de inovação tenha novamente uma média zero e variância unitária, usamos a reparametrização para padronizar. Depois de calcular μ_γ e σ_γ^2 definimos $x^* = \sigma_\gamma x + \mu_\gamma$ e a função de densidade de uma distribuição assimétrica padronizada pode ser expressada como,

$$f(x|\gamma) = \frac{2\sigma_\gamma}{\gamma + 1/\gamma} \left\{ f^*(\gamma x^*) I_{(-\infty, 0)}(x^*) + f^*\left(\frac{x^*}{\gamma}\right) I_{[0, \infty)}(x^*) \right\},$$

em que $f^*(x)$ é a f.d.p. simétrica padronizada, em cujo caso $m_2 = 1$.

A seguir, a [Tabla 2](#) mostra as expressões para o primeiro momento absoluto m_1 das distribuições Normal assimétrica, T-Student assimétrica e GED assimétrica.

Tabela 2 – Primeiros momentos absolutos das distribuições assimétricas geradas.

	m_1
Normal	$2/\sqrt{2\pi}$
T-Student	$\Gamma((v-1)/2) \sqrt{v-2} / \Gamma(v/2) \sqrt{\pi}$
GED	$\Gamma(2/\delta) / [\Gamma(1/\delta)\Gamma(3/\delta)]^{1/2}$

Como exemplo, se aplicarmos o método apresentado na distribuição Normal padronizada cuja f.d.p. é dada por

$$f^*(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right), \quad x \in \mathbb{R},$$

então a distribuição Normal assimétrica padronizada é

$$f(x|\gamma) = \left(\frac{2}{\pi}\right)^{1/2} \frac{\sigma_\gamma}{\gamma + 1/\gamma} \exp\left(-\frac{1}{2}x^{*2}\right), \quad (3.9)$$

onde x^* é dado por

$$x^* = \begin{cases} (x\sigma_\gamma + \mu_\gamma)\gamma & \text{se } x\sigma_\gamma + \mu_\gamma < 0, \\ (x\sigma_\gamma + \mu_\gamma)/\gamma & \text{se } x\sigma_\gamma + \mu_\gamma \geq 0. \end{cases} \quad (3.10)$$

Neste caso, sabe-se que $m_2 = 1$ e aplicando a [Ecuación 3.7](#) obtemos $m_1 = (2/\pi)^{1/2}$. Desta forma, a partir das expressões da [Ecuación 3.8](#) obtemos a média e a variância da distribuição assimétrica da distribuição Normal padronizada

$$\begin{aligned} \mu_\gamma &= (2/\pi)^{1/2}(\gamma - 1/\gamma), \\ \sigma_\gamma^2 &= (\gamma^2 + \gamma^{-2} - 1) - \mu_\gamma^2. \end{aligned}$$

A seguir apresentaremos os resultados da aplicação do método às distribuições T-Student e GED padronizadas, estas distribuições têm um parâmetro adicional de ajuste o qual modela a curtoses.

A distribuição T-Student padronizada com ν graus de liberdade e com f.d.p. dada por

$$f^*(x|\nu) = \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)\sqrt{\pi(\nu-2)}} \left(1 + \frac{x^2}{\nu-2}\right)^{-(\nu+1)/2}, \quad x \in \mathbb{R}, \quad \nu > 2,$$

logo a f.d.p. da distribuição T-Student assimétrica padronizada é

$$f(x|\gamma, \nu) = \frac{2}{\sqrt{\pi}} \frac{\sigma_\gamma}{\gamma + 1/\gamma} \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)(\nu-2)^{1/2}} \left(1 + \frac{x^{*2}}{\nu-2}\right)^{-\frac{\nu+1}{2}}, \quad (3.11)$$

com x^* igual que na [Ecuación 3.10](#), mas considerando a média e a variância da distribuição assimétrica da distribuição T-Student padronizada que são:

$$\begin{aligned} \mu_\gamma &= \frac{\Gamma((\nu-1)/2)\sqrt{\nu-2}}{\Gamma(\nu/2)\sqrt{\pi}}(\gamma - 1/\gamma), \\ \sigma_\gamma^2 &= (\gamma^2 + \gamma^{-2} - 1) - \mu_\gamma^2. \end{aligned}$$

A distribuição GED (*generalized error distribution*) padronizada com δ como parâmetro de forma e com f.d.p. dada por

$$f^*(x|\delta) = \frac{\delta}{\lambda 2^{1+1/\delta} \Gamma(1/\delta)} \exp\left(-\frac{|x/\lambda|^\delta}{2}\right), \quad x \in \mathbb{R}, \quad 0 < \delta \leq \infty,$$

onde $\Gamma(\cdot)$ é a função gama e $\lambda = \left[2^{-2/\delta} \Gamma(1/\delta) / \Gamma(3/\delta)\right]^{1/2}$. Esta distribuição reduz-se a uma distribuição de Laplace (ou dupla exponencial) se $\delta = 1$, a uma distribuição normal padronizada se $\delta = 2$ e a uma distribuição uniforme no intervalo $(-2\sqrt{3}, 2\sqrt{3})$ se $\delta \rightarrow \infty$. E tem caudas mais pesadas quando $\delta < 2$.

Então a f.d.p da distribuição GED assimétrica padronizada é

$$f(x|\gamma, \delta) = \frac{\sigma_\gamma}{\gamma + 1/\gamma} \left[\frac{\Gamma(3/\delta)}{\Gamma(1/\delta)} \right]^{1/2} \frac{1}{\Gamma((\delta + 1)/\delta)} \exp \left\{ - \left[\frac{\Gamma(3/\delta)}{\Gamma(1/\delta)} x^{*2} \right]^{\delta/2} \right\}, \quad (3.12)$$

com x^* igual que na [Ecuación 3.10](#), mas considerando a média e a variância da distribuição assimétrica da distribuição GED padronizada que são:

$$\begin{aligned} \mu_\gamma &= \frac{\Gamma(2/\delta)}{[\Gamma(1/\delta)\Gamma(3/\delta)]^{1/2}} (\gamma - 1/\gamma), \\ \sigma_\gamma^2 &= (\gamma^2 + \gamma^{-2} - 1) - \mu_\gamma^2. \end{aligned}$$

Observe que para cada uma das distribuições assimétricas padronizadas obtidas anteriormente, o γ^2 é igual à proporção de massas de probabilidade acima e abaixo da moda da distribuição e portanto é uma medida da assimetria. Para mostrar como muda a distribuição original depois de aplicado o método de [Fernández e Steel \(1998\)](#), a [Figura 5](#) apresenta as distribuições padronizadas e suas respectivas distribuições assimétricas obtidas aplicando o método. Em cada um dos três gráficos se faz a comparação da distribuição padronizada e de suas distribuições assimétricas padronizadas, geradas variando o parâmetro de assimetria tomando os casos quando for maior ou menor que um.

3.2 Estimação dos Parâmetros

A estimação dos parâmetros será feita usando a abordagem Bayesiana. Desta forma tem que se que definir uma distribuição a priori dos parâmetros, que especifica a incerteza inicial, calcular a verossimilhança para assim se obter a distribuição a posteriori dos parâmetros e portanto por fim fazer inferência dos mesmos.

3.2.1 Função de verossimilhança

Para uma amostra de observações $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$, a função de verossimilhança pode ser escrita como

$$\begin{aligned} p(y_1, \dots, y_n) &= p(y_1) \prod_{t=2}^n p(y_t | y_{t-1}, \dots, y_1) \\ &= p(y_1, \dots, y_r) \prod_{t=r+1}^n p(y_t | y_{t-1}, \dots, y_1). \end{aligned}$$

Então, em um modelo GARCH(p, q) tomando em conta que cada y_t tem uma distribuição de probabilidade que pode ser descrita como uma função de uma coleção parâmetros θ e de s observações passadas, sendo $\theta = (\theta_0, \theta_1, \dots, \theta_{p+q})' = (\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q)'$ e $s = \max(p, q)$,

temos

$$\begin{aligned} p(y_1, \dots, y_n | \boldsymbol{\theta}) &= p(y_1, \dots, y_s | \boldsymbol{\theta}) p(y_{s+1:n} | \boldsymbol{\theta}, y_{1:s}) \\ &= p(y_1, \dots, y_s | \boldsymbol{\theta}) \prod_{t=s+1}^n p(y_t | y_{t-1}, \dots, y_{t-s}, \boldsymbol{\theta}). \end{aligned}$$

Nós usaremos a verossimilhança condicional $p(y_{s+1:n} | \boldsymbol{\theta}, y_{1:s})$ denotada por $l(\boldsymbol{\theta})$, como uma aproximação da verossimilhança. Ver [Prado e West \(2010\)](#) para entender mais um pouco do uso da verossimilhança condicional em outros modelos de séries temporais. Portanto

$$p(y_1, \dots, y_n | \boldsymbol{\theta}) \approx \prod_{t=s+1}^n p(y_t | y_{t-1}, \dots, y_{t-s}, \boldsymbol{\theta}). \quad (3.13)$$

Observe que dada a [Ecuación 3.1](#) a distribuição condicional da cada retorno y_t depende da distribuição do erro ε_t . Logo, por teoria de transformação de variáveis aleatórias temos que

$$p(y_t | I_{t-1}) = (h_t)^{-1/2} p_\varepsilon\left(y_t / \sqrt{h_t}\right) \quad (3.14)$$

sendo p_ε a função de densidade de probabilidade de ε_t . Agora, podemos obter a função de verossimilhança condicionada do modelo substituindo o resultado da [Ecuación 3.14](#) na [Ecuación 3.13](#). Então:

$$l(\boldsymbol{\theta}) = \prod_{t=s+1}^n \frac{1}{\sqrt{h_t}} p_\varepsilon\left(\frac{y_t}{\sqrt{h_t}}\right),$$

dependendo da distribuição que tome o erro ε_t e a especificação dos parâmetros p e q no modelo, vai se poder calcular esta verossimilhança condicional.

Considerando que na maioria das aplicações empíricas verifica-se que a especificação $p = q = 1$ simples é capaz de reproduzir a dinâmica da volatilidade de dados financeiros, sendo para a maioria das séries financeiras o modelo mais parcimonioso para descrever a volatilidade. Então, nesta dissertação vamos nos limitar a aplicar o modelo GARCH(1,1). Por outro lado, vai se aplicar a proposta de inserir assimetria à distribuição do erro portanto vamos trabalhar com as distribuições assimétricas padronizadas obtidas na seção anterior.

3.2.2 Distribuições a Priori

A abordagem Bayesiana considera $\boldsymbol{\theta}$ como um vetor de variáveis aleatórias o qual é caracterizado por uma função de densidade a priori $p(\boldsymbol{\theta})$. O vetor aleatório $\boldsymbol{\theta}$, contém além dos parâmetros do modelo GARCH(1,1) os parâmetros de assimetria e do peso nas caudas. Assumindo independência a priori entre os parâmetros, cada a priori é especificada com a ajuda de hiperparâmetros os quais são assumidos conhecidos e constantes.

Portanto, para completar a especificação do modelo utilizaremos prioris normais truncadas sobre

os parâmetros do GARCH ω , α e β que foram propostas por [Ardia \(2006\)](#), desta forma temos:

$$\begin{aligned}\omega &\sim N(\mu_\omega, \sigma_\omega^2) I_{(\omega>0)} \\ \alpha &\sim N(\mu_\alpha, \sigma_\alpha^2) I_{(0<\alpha<1)} \\ \beta &\sim N(\mu_\beta, \sigma_\beta^2) I_{(0<\beta<1)},\end{aligned}$$

sendo μ_ω , μ_α , μ_β , σ_ω^2 , σ_α^2 e σ_β^2 hiperparâmetros.

No caso de usar o parâmetro de assimetria teríamos as distribuições dos erros assimétricas, como: T-Student assimétrica, GED assimétrica ou Normal assimétrica sendo necessário estimar o parâmetro γ . Seguindo a proposta de [Fernández e Steel \(1998\)](#), usaremos uma distribuição a priori Gama(a, b) para γ^2 sendo igual à razão de probabilidade que indica o grau de assimetria da distribuição com respeito a zero, *i.e.* $\gamma^2 = Pr(X \geq 0)/Pr(X < 0)$. Como é razoável escolher uma priori que seja simétrica, calculamos os hiperparâmetros a e b tais que cumpram o seguinte: $E(\gamma) = 1$, com simples operações se deduz que $b = [\Gamma(a + 1/2)/\Gamma(a)]^2$ assim fixando o valor de a pode se controlar a variância priori e a probabilidade priori de γ sobre $(0, 1)$. [Fernández e Steel \(1998\)](#) encontraram que fixando um $a = 1/2$ temos um $b \approx 0.32$ o que nos leva a obter $Var(\gamma) \approx 0.57$ e $Pr(0 < \gamma < 1) \approx 0.58$ o que é considerado uma escolha razoável. Esta particular eleição é equivalente a estabelecer a priori seguinte:

$$\gamma \sim N(0, 0.64^{-1}) I_{(\gamma>0)},$$

em que 0.64^{-1} indica a variância da distribuição.

Também se utilizarmos as distribuições dos erros com caudas pesadas, teremos que estimar o parâmetro ν ou δ , para uma distribuição T-Student ou uma distribuição GED, respectivamente. Neste caso, serão usadas prioris com distribuições Normais truncadas,

$$\begin{aligned}\nu &\sim N(\mu_\nu, \sigma_\nu^2) I_{(\nu>2)}, \\ \delta &\sim N(\mu_\delta, \sigma_\delta^2) I_{(\delta>0)},\end{aligned}$$

com hiperparâmetros μ_ν , μ_δ , σ_ν^2 e σ_δ^2 .

3.2.3 Distribuição a Posteriori

Aplicando o teorema de Bayes calculamos a distribuição a posteriori dos parâmetros, como é comum, deixaremos a constante de normalização para focarmos só no núcleo da distribuição. E procederemos à utilização de métodos de simulação estocástica como o MCMC apresentado no [Capítulo 2](#), que consiste na extração de amostras da distribuição a posteriori construindo eficazmente cadeias de Markov ao longo do tempo, para assim poder obter uma estimativa dos parâmetros.

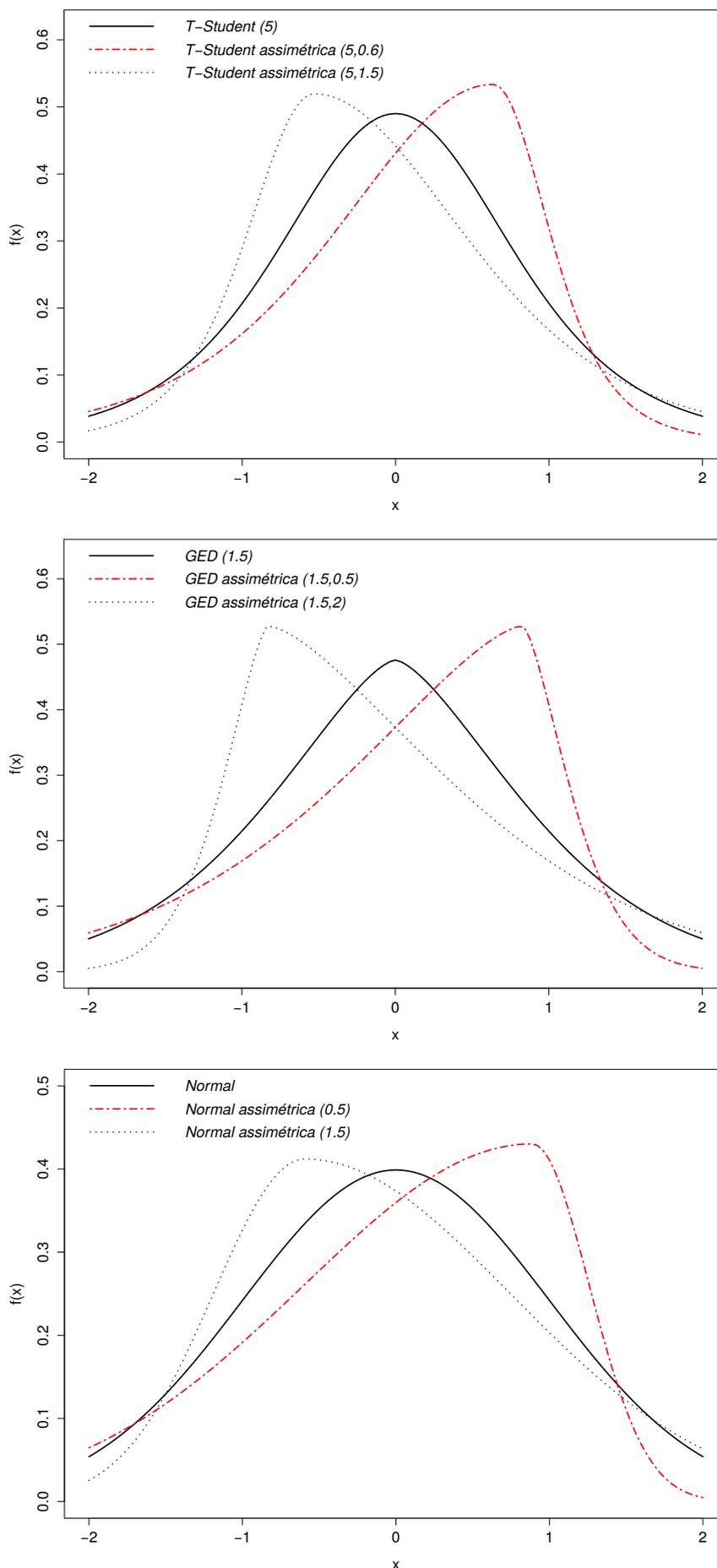


Figura 5 – Distribuição T-Student com $\nu = 5$ e com γ igual a 0.6 ou 0.5, distribuição GED com $\delta = 1.5$ e com γ igual a 0.5 ou 2 e distribuição Normal com γ igual a 0.5 ou 1.5.

ESTUDO DE SIMULAÇÃO

No presente capítulo vamos a apresentar o algoritmo programado e empregado para a utilização da estimação. Além disso, faremos uma análise do desempenho do mesmo.

4.1 Algoritmo Metropolis-Hastings

Para obter uma amostra da distribuição posteriori dos parâmetros do modelo, utilizaremos o algoritmo Metropolis-Hastings do tipo passeio aleatório em que o estado da cadeia de Markov é atualizado acrescentando ruído ao estado atual. Para desenvolver o algoritmo optamos por considerar a distribuição do ruído igual a uma distribuição Normal multivariada simétrica em torno de $\mathbf{0}$, que é muito usada na prática, então a distribuição proposta será uma distribuição Normal multivariada centrada no estado atual da cadeia atualizando assim todos os parâmetros em um só bloco. O esquema de atualização se implementa de maneira mais eficiente simulando todos os parâmetros em linha reta, portanto vamos considerar transformações em todos os parâmetros tornando os reais como o espaço paramétrico. A [Tabela 3](#) descreve as transformações dos parâmetros adotadas.

Tabela 3 – Transformações adotadas dos parâmetros.

<i>Transformação</i>	<i>Inversa</i>	<i>Descrição (dist. dos erros)</i>
$\phi_1 = g_1(\boldsymbol{\theta}) = \log(v - 2)$	$v = h_1(\boldsymbol{\phi}) = \exp(\phi_1) + 2$	T-Student assimétrica padronizada
$\phi_1 = g_1(\boldsymbol{\theta}) = \log(\delta)$	$\delta = h_1(\boldsymbol{\phi}) = \exp(\phi_1)$	GED assimétrica padronizada
$\phi_2 = g_2(\boldsymbol{\theta}) = \log(\gamma)$	$\gamma = h_2(\boldsymbol{\phi}) = e^{\phi_2}$	Parâmetros do modelo
$\phi_3 = g_3(\boldsymbol{\theta}) = \log(\omega)$	$\omega = h_3(\boldsymbol{\phi}) = e^{\phi_3}$	
$\phi_4 = g_4(\boldsymbol{\theta}) = \log[\alpha/(1 - \alpha)]$	$\alpha = h_4(\boldsymbol{\phi}) = e^{\phi_4} / (1 + e^{\phi_4})$	
$\phi_5 = g_5(\boldsymbol{\theta}) = \log[\beta/(1 - \beta)]$	$\beta = h_5(\boldsymbol{\phi}) = e^{\phi_5} / (1 + e^{\phi_5})$	

Os detalhes se darão supondo que estamos trabalhando com erros com distribuição T-Student ou GED assimétrica padronizada, sabendo que no caso de trabalhar com a distribuição normal assimétrica padronizada só não se considerará o parâmetro da cauda.

Então, denotando $\boldsymbol{\theta} = h(\boldsymbol{\phi}) = (h_1(\boldsymbol{\phi}), \dots, h_5(\boldsymbol{\phi}))'$ com distribuição a posteriori $\pi(\boldsymbol{\theta}|\mathbf{y})$ e $\boldsymbol{\phi} = g(\boldsymbol{\theta}) = (g_1(\boldsymbol{\theta}), \dots, g_5(\boldsymbol{\theta}))' = (\phi_1, \dots, \phi_5)'$ como o vetor de parâmetros transformados. Para conhecer a distribuição a posteriori de $\boldsymbol{\phi}$, usamos a distribuição a posteriori do vetor de parâmetros iniciais $\boldsymbol{\theta}$ em que a distribuição a priori conjunta é calculada assumindo independência a priori entre os parâmetros. Usando o método Jacobiano, obtemos o seguinte

$$\pi(\boldsymbol{\phi}|\mathbf{y}) = \pi_{\boldsymbol{\theta}|\mathbf{y}}(h(\boldsymbol{\phi})) |J_h(\boldsymbol{\phi})|, \quad (4.1)$$

em que:

$$\pi_{\boldsymbol{\theta}|\mathbf{y}}(h(\boldsymbol{\phi})) \propto l(h(\boldsymbol{\phi})) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\boldsymbol{\phi})),$$

$$J_h(\boldsymbol{\phi}) = \begin{vmatrix} \frac{\partial h_1}{\partial \phi_1} & \frac{\partial h_1}{\partial \phi_2} & \dots & \frac{\partial h_1}{\partial \phi_5} \\ \frac{\partial h_2}{\partial \phi_1} & \frac{\partial h_2}{\partial \phi_2} & \dots & \frac{\partial h_2}{\partial \phi_5} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_5}{\partial \phi_1} & \frac{\partial h_5}{\partial \phi_2} & \dots & \frac{\partial h_5}{\partial \phi_5} \end{vmatrix} = \frac{e^{\phi_1 + \dots + \phi_5}}{[(1+e^{\phi_4})(1+e^{\phi_5})]^2},$$

então, substituindo na [Ecuación 4.1](#) temos que

$$\pi(\boldsymbol{\phi}|\mathbf{y}) \propto l(h(\boldsymbol{\phi})) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\boldsymbol{\phi})) \frac{e^{\phi_1 + \dots + \phi_5}}{[(1+e^{\phi_4})(1+e^{\phi_5})]^2}. \quad (4.2)$$

Para desenvolver a verossimilhança condicional avaliada no $h(\boldsymbol{\phi})$, temos que especificar qual será a distribuição do erro, assim:

- Para uma distribuição T-Student assimétrica padronizada, usando a [Ecuación 3.11](#) se tem

$$l(h(\boldsymbol{\phi})) = \prod_{t=2}^n \left(e^{\phi_3} + \frac{e^{\phi_4}}{1+e^{\phi_4}} y_{t-1}^2 + \frac{e^{\phi_5}}{1+e^{\phi_5}} h_{t-1} \right)^{-1/2} \times \prod_{t=2}^n \left\{ \frac{2}{\sqrt{\pi}} \frac{e^{\phi_2} \sigma_\gamma}{1+e^{2\phi_2}} \frac{\Gamma((e^{\phi_1}+3)/2)}{\Gamma((e^{\phi_1}+2)/2) e^{\phi_1/2}} \left[1 + \frac{x_t^{*2}}{e^{\phi_1}} \right]^{-\frac{e^{\phi_1}+3}{2}} \right\},$$

em que

$$x_t^* = \begin{cases} (x_t \sigma_\gamma + \mu_\gamma) e^{\phi_2} & \text{se } x_t < -\mu_\gamma / \sigma_\gamma, \\ (x_t \sigma_\gamma + \mu_\gamma) / e^{\phi_2} & \text{se } x_t \geq -\mu_\gamma / \sigma_\gamma, \end{cases} \quad (4.3)$$

sendo:

$$\begin{aligned} x_t &= y_t / \sqrt{h_t} \quad \text{com } h_t = e^{\phi_3} + (e^{\phi_4} / (1+e^{\phi_4})) y_{t-1}^2 + (e^{\phi_5} / (1+e^{\phi_5})) h_{t-1}, \\ \mu_\gamma &= (e^{\phi_2} - e^{-\phi_2}) m_1 \quad \text{com } m_1 = (\Gamma((e^{\phi_1}+1)/2) e^{\phi_1/2}) / (\Gamma((e^{\phi_1}+2)/2) \sqrt{\pi}), \\ \sigma_\gamma^2 &= e^{2\phi_2} + e^{-2\phi_2} - 1 - \mu_\gamma^2. \end{aligned}$$

- Para uma distribuição GED assimétrica padronizada, usando a [Ecuación 3.12](#) se tem

$$l(h(\boldsymbol{\phi})) = \prod_{t=2}^n \left(e^{\phi_3} + \frac{e^{\phi_4}}{1+e^{\phi_4}} y_{t-1}^2 + \frac{e^{\phi_5}}{1+e^{\phi_5}} h_{t-1} \right)^{-1/2} \times \prod_{t=2}^n \left\{ \frac{e^{\phi_2} \sigma_\gamma}{1+e^{2\phi_2}} \left[\frac{\Gamma(3/e^{\phi_1})}{\Gamma(1/e^{\phi_1})} \right]^{1/2} \frac{1}{\Gamma((e^{\phi_1}+1)/e^{\phi_1})} \exp \left\{ - \left[\frac{\Gamma(3/e^{\phi_1})}{\Gamma(1/e^{\phi_1})} x_t^{*2} \right]^{e^{\phi_1}/2} \right\} \right\},$$

em que x_t^* fica igual que na [Ecuación 4.3](#), mas com as seguintes atualizações

$$\begin{aligned} x_t &= y_t / \sqrt{h_t} \quad \text{com } h_t = e^{\phi_3} + (e^{\phi_4}/(1+e^{\phi_4})) y_{t-1}^2 + (e^{\phi_5}/(1+e^{\phi_5})) h_{t-1}, \\ \mu_\gamma &= (e^{\phi_2} - e^{-\phi_2}) m_1 \quad \text{com } m_1 = \Gamma(2/e^{\phi_1}) / \left[\Gamma(1/e^{\phi_1}) \Gamma(3/e^{\phi_1}) \right]^{1/2}, \\ \sigma_\gamma^2 &= e^{2\phi_2} + e^{-2\phi_2} - 1 - \mu_\gamma^2. \end{aligned}$$

- Para uma distribuição Normal assimétrica padronizada, usando a [Ecuación 3.9](#) se tem

$$l(h(\boldsymbol{\phi})) = \prod_{t=2}^n \left(e^{\phi_3} + \frac{e^{\phi_4}}{1+e^{\phi_4}} y_{t-1}^2 + \frac{e^{\phi_5}}{1+e^{\phi_5}} h_{t-1} \right)^{-1/2} \times \prod_{t=2}^n \left\{ \left(\frac{2}{\pi} \right)^{1/2} \frac{e^{\phi_2} \sigma_\gamma}{1+e^{2\phi_2}} \exp \left\{ - \frac{1}{2} x_t^{*2} \right\} \right\},$$

em que x_t^* fica igual que na [Ecuación 4.3](#), mas com as seguintes atualizações

$$\begin{aligned} x_t &= y_t / \sqrt{h_t} \quad \text{com } h_t = e^{\phi_3} + (e^{\phi_4}/(1+e^{\phi_4})) y_{t-1}^2 + (e^{\phi_5}/(1+e^{\phi_5})) h_{t-1}, \\ \mu_\gamma &= (e^{\phi_2} - e^{-\phi_2}) m_1 \quad \text{com } m_1 = 2/\sqrt{2\pi}, \\ \sigma_\gamma^2 &= e^{2\phi_2} + e^{-2\phi_2} - 1 - \mu_\gamma^2. \end{aligned}$$

Por outro lado, para desenvolver o produto das distribuições a priori, se tomará em conta o produto das seguintes prioris avaliadas em $h(\boldsymbol{\phi})$.

Se a distribuição do erro é T-Student ou GED assimétrica padronizada, se considerará esta a priori

$$\pi_{\theta_1}(h_1(\boldsymbol{\phi})) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp \left\{ - \frac{1}{2} \frac{(e^{\phi_1} + 2 - \mu_v)^2}{\sigma_v^2} \right\} / \left[1 - \Phi \left(\frac{2 - \mu_v}{\sigma_v} \right) \right] I_{(\phi_1 \in \mathbb{R})},$$

$$\pi_{\theta_1}(h_1(\boldsymbol{\phi})) = \frac{1}{\sqrt{2\pi}\sigma_\delta} \exp \left\{ - \frac{1}{2} \frac{(e^{\phi_1} - \mu_\delta)^2}{\sigma_\delta^2} \right\} / \left[1 - \Phi \left(\frac{-\mu_\delta}{\sigma_\delta} \right) \right] I_{(\phi_1 \in \mathbb{R})}.$$

As a priori restantes são tomadas por qualquer das três distribuições dos erros

$$\pi_{\theta_2}(h_2(\boldsymbol{\phi})) = \frac{1}{\sqrt{2\pi}(0.64)^{-1/2}} \exp\left\{-\frac{1}{2} \frac{e^{2\phi_2}}{0.64^{-1}}\right\} / [1 - \Phi(0)] I_{(\phi_2 \in \mathbb{R})},$$

$$\pi_{\theta_3}(h_3(\boldsymbol{\phi})) = \frac{1}{\sqrt{2\pi}\sigma_\omega} \exp\left\{-\frac{1}{2} \frac{(e^{\phi_3} - \mu_\omega)^2}{\sigma_\omega^2}\right\} / \left[1 - \Phi\left(\frac{-\mu_\omega}{\sigma_\omega}\right)\right] I_{(\phi_3 \in \mathbb{R})},$$

$$\pi_{\theta_4}(h_4(\boldsymbol{\phi})) = \frac{1}{\sqrt{2\pi}\sigma_\alpha} \exp\left\{-\frac{1}{2\sigma_\alpha^2} \left(\frac{e^{\phi_4}}{1+e^{\phi_4}} - \mu_\alpha\right)^2\right\} / \left[\Phi\left(\frac{1-\mu_\alpha}{\sigma_\alpha}\right) - \Phi\left(\frac{-\mu_\alpha}{\sigma_\alpha}\right)\right] I_{(\phi_4 \in \mathbb{R})},$$

$$\pi_{\theta_5}(h_5(\boldsymbol{\phi})) = \frac{1}{\sqrt{2\pi}\sigma_\beta} \exp\left\{-\frac{1}{2\sigma_\beta^2} \left(\frac{e^{\phi_5}}{1+e^{\phi_5}} - \mu_\beta\right)^2\right\} / \left[\Phi\left(\frac{1-\mu_\beta}{\sigma_\beta}\right) - \Phi\left(\frac{-\mu_\beta}{\sigma_\beta}\right)\right] I_{(\phi_5 \in \mathbb{R})}.$$

Então, substituindo as igualdades calculadas na [Ecuación 4.2](#) obtemos a distribuição posteriori do vetor transformado $\boldsymbol{\phi}$.

4.1.1 Distribuição Proposta

Existem duas possibilidades para implementar o algoritmo, em diferentes blocos e em um só bloco. A seguir apresentaremos um breve resumo de ambas opções.

Supondo trabalhar com o algoritmo M-H de um só bloco, todo o vetor de parâmetros $\boldsymbol{\phi} = (\phi_1, \dots, \phi_5)$ será atualizado em cada iteração em um só passo. Para isso, se define uma densidade proposta $q(\boldsymbol{\phi}, \cdot)$ que gerará os valores candidatos $\boldsymbol{\phi}'$ dado o valor atual $\boldsymbol{\phi}$. Como o processo de geração do valor candidato surge de um processo de passeio aleatório $\boldsymbol{\phi}' = \boldsymbol{\phi} + \boldsymbol{\varepsilon}$ com $\boldsymbol{\varepsilon} \sim N_5(\mathbf{0}, c\Sigma)$, em que c é uma constante de ajuste para otimizar a taxa de aceitação e Σ é uma aproximação da variância a posteriori aqui usamos a aproximação normal. Então, a função proposta resulta simétrica $q(\boldsymbol{\phi}, \boldsymbol{\phi}') = q(\boldsymbol{\phi}', \boldsymbol{\phi})$ e a probabilidade de aceitar um valor proposto $\boldsymbol{\phi}'$ como novo é

$$\alpha(\boldsymbol{\phi}, \boldsymbol{\phi}') = \min\left\{\frac{\pi(\boldsymbol{\phi}'|\mathbf{y})}{\pi(\boldsymbol{\phi}|\mathbf{y})}, 1\right\},$$

Considerando o resultado da [Ecuación 4.1](#), temos

$$\alpha(\boldsymbol{\phi}, \boldsymbol{\phi}') = \min\left\{\frac{l(h(\boldsymbol{\phi}')) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\boldsymbol{\phi}')) |J_h(\boldsymbol{\phi}')|}{l(h(\boldsymbol{\phi})) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\boldsymbol{\phi})) |J_h(\boldsymbol{\phi})|}, 1\right\}.$$

Por outro lado, quando se trabalha com o algoritmo M-H em blocos se amostra cada bloco dado o valor mais atual dos blocos restantes e um ciclo do algoritmo é completado quando tem se atualizado cada bloco. Neste algoritmo o número de blocos é 5 ao igual que a dimensão do vetor aleatório $\boldsymbol{\phi}$. Então, seja $\boldsymbol{\phi}_{-k} = (\phi_1, \dots, \phi_{k-1}, \phi_{k+1}, \dots, \phi_5)$ o vetor de variáveis sem tomar em conta o ϕ_k em que $k = 1, \dots, 5$ e $q_k(\phi_k, \phi'_k | \boldsymbol{\phi}_{-k})$ a densidade proposta ou a densidade que gera o valor candidato ϕ'_k dado o valor atual ϕ_k do bloco pode depender dos valores atuais dos blocos

restantes. Por definição, temos que a probabilidade de movimento do bloco ϕ_k condicionado a ϕ_{-k} é

$$\alpha_k(\phi_k, \phi'_k | \phi_{-k}) = \min \left\{ \frac{\pi(\phi'_k, \phi_{-k} | \mathbf{y}) q_k(\phi'_k, \phi_k | \phi_{-k})}{\pi(\phi_k, \phi_{-k} | \mathbf{y}) q_k(\phi_k, \phi'_k | \phi_{-k})}, 1 \right\}.$$

O valor candidato ϕ'_k é gerado seguindo o processo $\phi'_k = \phi_k + \varepsilon$ em que ε é uma variável de incremento aleatório e segue uma distribuição $N(0, \sigma_{\phi_k}^2)$ simétrica respeito a zero $p(\varepsilon) = p(-\varepsilon)$. Logo, o ϕ'_k é gerado de uma distribuição $N(\phi_k, \sigma_k^2)$ que não depende dos valores restantes. Daqui resulta que $q_k(\phi'_k, \phi_k | \phi_{-k}) = q_k(\phi_k, \phi'_k | \phi_{-k})$, então a probabilidade de movimento é reduzida a

$$\alpha_k(\phi_k, \phi'_k | \phi_{-k}) = \min \left\{ \frac{\pi(\phi'_k, \phi_{-k} | \mathbf{y})}{\pi(\phi_k, \phi_{-k} | \mathbf{y})}, 1 \right\}.$$

Agora fazendo o cálculo da razão das distribuições a posterioris no respetivo bloco k , observe que a constante de padronização será cancelada quanto se fizer a divisão, portanto a probabilidade de aceitação é equivalente a

$$\alpha_k(\phi_k, \phi'_k | \phi_{-k}) = \min \left\{ \frac{l(h(\phi'_k, \phi_{-k})) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\phi'_k, \phi_{-k})) | J_h(\phi'_k, \phi_{-k})|}{l(h(\phi_k, \phi_{-k})) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\phi_k, \phi_{-k})) | J_h(\phi_k, \phi_{-k})|}, 1 \right\},$$

note que algumas partes da verossimilhança, da priori ou do Jacobiano que não envolvem o parâmetro que está sendo atualizado no bloco são considerados também constantes.

4.1.2 Esquema do Algoritmo

Os algoritmos mencionados anteriormente e que gerarão N amostras do vetor de parâmetros transformados podem ser resumidos como se mostra no [Algoritmo 1](#) e no [Algoritmo 2](#).

Na especificação dos valores iniciais e da variância da distribuição proposta, se utilizou o **método de aproximação normal** onde a distribuição a posteriori de $\phi \sim N(m, V)$ com a moda m , que é geralmente obtida da solução da seguinte equação

$$\frac{\partial \log \pi(\phi | \mathbf{y})}{\partial \phi} = 0,$$

e V que é igual à inversa da matriz de segunda derivada respeito a ϕ de $\log \pi(\phi | \mathbf{y})$ avaliada na moda m , assim

$$V = \left[-\frac{\partial^2 \log \pi_{\phi | \mathbf{y}}(m)}{\partial \phi \partial \phi'} \right]^{-1}, \quad (4.4)$$

para aprofundar nestes métodos de aproximação ler o capítulo 3 do livro de [Gelman e Lopes \(2006\)](#). Ambos cálculos são obtidos facilmente usando a função *optim* do R software. Então,

se toma m como valores iniciais do vetor de parâmetros e no caso de usar um bloco se toma V como $\hat{\Sigma}$ mas no caso do algoritmo em blocos se consideram os valores da diagonal de V como $\hat{\sigma}_k^2$ para cada bloco respetivamente.

Algoritmo 1 – Metropolis-Hastings com um bloco

1: **procedimento**

2: Especificar os valores iniciais para o vetor de parâmetros $\boldsymbol{\theta}^{(0)}$ e transformá-los a $\boldsymbol{\phi}^{(0)} = (\phi_1^{(0)}, \dots, \phi_5^{(0)})$,

3: Seja $j = 1$

4: **enquanto** $j \leq N$ **faça**

5: Gerar o valor proposto $\boldsymbol{\phi}' \sim N_5(\boldsymbol{\phi}^{(j-1)}, c\hat{\Sigma})$,

6: Se $\sum_{j=4}^5 e^{\phi'_j} / (1 + e^{\phi_j}) < 1$, então calcular a probabilidade de movimento

$$\alpha(\boldsymbol{\phi}^{(j-1)}, \boldsymbol{\phi}') = \min \left\{ \frac{l(h(\boldsymbol{\phi}')) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\boldsymbol{\phi}')) |J_h(\boldsymbol{\phi}')|}{l(h(\boldsymbol{\phi}^{(j-1)})) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\boldsymbol{\phi}^{(j-1)})) |J_h(\boldsymbol{\phi}^{(j-1)})|}, 1 \right\},$$

7: Estabelecer

$$\boldsymbol{\phi}^{(j)} = \begin{cases} \boldsymbol{\phi}' & \text{Se } \text{Unif}(0, 1) \leq \alpha(\boldsymbol{\phi}^{(j-1)}, \boldsymbol{\phi}'), \\ \boldsymbol{\phi}^{(j-1)} & \text{caso contrário,} \end{cases}$$

8: Atualizar $j = j + 1$.

9: **fim enquanto**

10: **retorna** $\{\boldsymbol{\phi}^{(1)}, \boldsymbol{\phi}^{(2)}, \dots, \boldsymbol{\phi}^{(N)}\}$.

11: **fim procedimento**

4.2 Estudo de Simulação

Antes de que qualquer modelo seja usado com confiança, a avaliação adequada da magnitude dos erros que podem resultar de sua utilização deve ser realizada. A validação do modelo, em sua forma mais simples, é uma comparação entre valores simulados e observados. Por tanto, vamos avaliar o desempenho (ou qualidade) do modelo GARCH(1,1) com distribuições assimétricas e com caudas pesadas utilizado nesta dissertação para modelar séries de retornos financeiros, usando séries simuladas.

Para obter estatísticas quantitativas do desempenho da estimação da média posteriori do modelo com distribuições do erro GED e Normal assimétricas padronizadas, se geraram dados artificiais para um conjunto de 200 séries (replicações) de dados de tamanho 500, 1000 e 1500 utilizando o pacote *fGarch* considerando a distribuição do Erro Normal padronizada com parâmetros $\alpha = 0.2$, $\beta = 0.7$ e $\omega = 0.5$. Por outro lado, para avaliar o desempenho da estimação da média posteriori do modelo com distribuição do erro T-Student assimétrica padronizada,

Algoritmo 2 – Metropolis-Hastings em blocos1: **procedimento**2: Especificar os valores iniciais para o vetor de parâmetros $\theta^{(0)}$ e transformá-los a $\phi^{(0)} = (\phi_1^{(0)}, \dots, \phi_5^{(0)})$,3: Sejam $j = 1$ e $k = 1$ 4: **enquanto** $j \leq N$ **faça**5: **enquanto** $k \leq 5$ **faça**6: Gerar o valor proposto $\phi'_k \sim N_5(\phi_k^{(j-1)}, \hat{\sigma}_k^2)$,

7: Se respeita-se a condição de estacionariedade, a probabilidade de movimento é

$$\alpha_k(\phi_k^{(j-1)}, \phi'_k | \phi_{-k}) = \min \left\{ \frac{l(h(\phi'_k, \phi_{-k})) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\phi'_k, \phi_{-k})) | J_h(\phi'_k, \phi_{-k})|}{l(h(\phi_k^{(j-1)}, \phi_{-k})) \prod_{i=1}^5 \pi_{\theta_i}(h_i(\phi_k^{(j-1)}, \phi_{-k})) | J_h(\phi_k^{(j-1)}, \phi_{-k})|}, 1 \right\},$$

8: Estabelecer $\phi_k^{(j)} = \begin{cases} \phi'_k & \text{Se } \text{Unif}(0, 1) \leq \alpha_k(\phi_k^{(j-1)}, \phi'_k | \phi_{-k}), \\ \phi_k^{(j-1)} & \text{caso contrário,} \end{cases}$ 9: Atualizar $k = k + 1$.10: **fim enquanto**11: Atualizar $j = j + 1$ e $k = 1$.12: **fim enquanto**13: **retorna** $\{\phi^{(1)}, \phi^{(2)}, \dots, \phi^{(N)}\}$.14: **fim procedimento**

se geraram dados artificiais para um conjunto de 200 séries de dados de tamanho 500, 1000 e 1500 utilizando o pacote *fGarch* considerando a distribuição do erro T-Student padronizada com parâmetros $\alpha = 0.4$, $\beta = 0.3$, $\omega = 2.5$ e $\nu = 8$. Para ambos casos espera-se encontrar um $\gamma \approx 1$.

4.2.1 Medidas do Erro de Estimação

A **Raiz do Erro Quadrático Médio** (RMSE em inglês) tem sido usada como uma medida estatística padrão para medir o desempenho do modelo em distintos tipos de estudos de pesquisa. Além desta medida também usaremos o **viés** (Bias em inglês).

O RMSE e o Bias da estimação do parâmetro θ são definidos a seguir:

$$RMSE(\theta) = \sqrt{\frac{1}{M} \sum_{i=1}^M (\theta - \hat{\theta}^{(i)})^2},$$

$$BIAS(\theta) = \left[\frac{1}{M} \sum_{i=1}^M (\hat{\theta}^{(i)}) \right] - \theta.$$

Sendo:

- θ é o valor fixado de cada parâmetro para a geração das 200 séries simuladas.
- $\hat{\theta}$ é a média posteriori do parâmetro calculada para cada uma das 200 séries.
- M é o número de séries simuladas (nosso caso 200), geradas artificialmente com a mesma quantidade de observações e os mesmos parâmetros.

A seguir se apresentam os resultados dos indicadores, usando o [Algoritmo 2](#) M-H com 5 blocos onde cada bloco contém um parâmetro.

A [Tabela 4](#) apresenta o desempenho da estimação da média posteriori. A estimativa para cada série gerada com distribuição do erro T-Student padrão com 8 graus de liberdade, é calculada usando uma cadeia de Markov de comprimento 20 mil gerada de acordo ao algoritmo programado no R software o qual considera o modelo GARCH(1,1) com distribuição do Erro T-Student assimétrica padronizada. Para garantir a convergência à distribuição a posteriori, os primeiros 10 mil valores gerados foram descartados, deixando assim uma sequência de tamanho 10 mil para calcular a estimativa da média posteriori.

Tabela 4 – Viés e Raiz do Erro Quadrático Médio da estimação posteriori.

		<i>T-Student assimétrica padronizada</i>		
		$n = 500$	$n = 1000$	$n = 1500$
BIAS	$\nu = 8$	1.60157	1.12805	0.73767
	$\gamma = 1$	-0.00636	0.00080	-0.00564
	$\omega = 2.5$	0.26949	0.13566	0.09082
	$\alpha = 0.4$	0.01614	0.00945	0.01130
	$\beta = 0.3$	-0.03484	-0.02181	-0.01864
RMSE	$\nu = 8$	3.04065	2.61743	2.01642
	$\gamma = 1$	0.09789	0.08356	0.07964
	$\omega = 2.5$	0.68211	0.55492	0.41608
	$\alpha = 0.4$	0.09841	0.07475	0.06119
	$\beta = 0.3$	0.10692	0.08718	0.06881

As seguintes tabelas: [Tabela 5](#) e [Tabela 6](#) apresentam o desempenho da estimação da média posteriori. A estimativa para cada série gerada com distribuição do erro Normal padrão é calculada usando uma cadeia de Markov de comprimento 20 mil gerada de acordo ao algoritmo programado no R software o qual considera o modelo GARCH(1,1) com distribuição do erro GED ou Normal assimétrica padronizada. Para garantir a convergência à distribuição a posteriori, os primeiros 10 mil valores gerados foram descartados, deixando assim uma sequência de tamanho 10 mil para calcular a estimativa da média posteriori.

Tabela 5 – Viés e Raiz do Erro Quadrático Médio da estimação posteriori.

		<i>GED assimétrica padronizada</i>		
		<i>n = 500</i>	<i>n = 1000</i>	<i>n = 1500</i>
BIAS	$\delta = 2$	0.06900	0.01794	-0.08814
	$\gamma = 1$	0.00924	0.00501	-0.09044
	$\omega = 2.5$	0.41691	0.16305	0.07629
	$\alpha = 0.4$	0.03672	0.01925	-0.00352
	$\beta = 0.3$	-0.1179	-0.04963	-0.09755
RMSE	$\delta = 2$	0.24433	0.15394	0.60586
	$\gamma = 1$	0.07692	0.04723	0.29878
	$\omega = 2.5$	0.60762	0.25481	0.24586
	$\alpha = 0.4$	0.06315	0.04112	0.06788
	$\beta = 0.3$	0.16175	0.08057	0.21740

Tabela 6 – Viés e Raiz do Erro Quadrático Médio da estimação posteriori.

		<i>Normal assimétrica padronizada</i>		
		<i>n = 500</i>	<i>n = 1000</i>	<i>n = 1500</i>
BIAS	$\gamma = 1$	0.00965	0.00462	-0.00162
	$\omega = 2.5$	0.42007	0.15871	0.11625
	$\alpha = 0.4$	0.03543	0.01782	0.01343
	$\beta = 0.3$	-0.11883	-0.04838	-0.03596
RMSE	$\gamma = 1$	0.07586	0.04646	0.04115
	$\omega = 2.5$	0.62950	0.25248	0.19115
	$\alpha = 0.4$	0.06187	0.03995	0.03189
	$\beta = 0.3$	0.16440	0.08058	0.05784

Segundo resultados obtidos dos indicadores, para o modelo com distribuição nos erros T-Student assimétrica padronizada a estimação dos parâmetros ν e ω não teve um bom desempenho mas o erro de estimação vai decrescendo a medida que aumenta o conjunto de dados. No modelo com distribuição nos erros GED assimétrica padronizada os dois indicadores coincidem em notar o mal desempenho na estimação do parâmetro ω mas o RMSE também ficou alto no parâmetro da cauda. Por último, no modelo com distribuição Normal assimétrica padronizada novamente se tem problemas ao estimar o parâmetro ω segundo os dois indicadores. Cabe notar que o erro de estimação em todos os casos analisados acima vai decrescendo conforme a quantidade de dados vai aumentando. Para os demais parâmetros o algoritmo com os três tipos de distribuições do erro obteve bom desempenho.

APLICAÇÃO

Neste capítulo examinaremos o processo da volatilidade de dois importantes índices financeiros, estes são **IBOVESPA** o mais importante indicador do desempenho médio das cotações das ações negociadas na Bolsa de Valores de São Paulo e **S&P500** um índice composto por quinhentos ativos cotados nas bolsas de NYSE (Bolsa de Valores de Nova Iorque) ou NASDAQ (Associação Nacional de Corretores de Títulos de Cotações Automáticas). Para ambos analisaremos no mesmo período diário compreendido entre os anos de 2012 e 2016, utilizando-se o modelo GARCH(1,1) com distribuições nos erros Normal, T-Student e GED assimétricas padronizadas apresentadas no [Capítulo 3](#).

A inferência dos modelos será feita sob abordagem Bayesiana utilizando MCMC. As cadeias de Markov serão geradas pelo [Algoritmo 1](#) apresentado na [subseção 4.1.2](#). A implementação dos algoritmos e as análises das cadeias de Markov foram feitas no software livre R.

Os valores escolhidos para os hiperparâmetros das distribuições a priori são: $\mu_\omega = \mu_\alpha = \mu_\beta = 0$ e $\sigma_\omega^2 = \sigma_\alpha^2 = \sigma_\beta^2 = 100$. Desta forma, obtemos prioris com informação vaga em quase todo espaço paramétrico, exceto para o parâmetro de assimetria, que utiliza a distribuição a priori descrita na [subseção 3.2.2](#).

O cálculo de Σ se faz usando a [Ecuación 4.4](#) do método de aproximação descrito na [subseção 4.1.2](#). Para fixar o valor de c se geraram cadeias de Markov de tamanho 100 mil com diferentes valores de c monitorando a taxa de aceitação da amostra gerada, sendo as primeiras 50mil descartadas como amostra de aquecimento. Não se tem informação exata na literatura sobre quais são as taxas ótimas para este tipo de modelos, portanto se estabeleceu trabalhar com taxas de aceitação dentro de intervalo de 0.70 – 0.75.

As cadeias de Markov geradas são de tamanho 200 mil, sendo as 100 mil primeiras descartadas como amostra de aquecimento para o modelo com distribuição nos erros Normal e T-Student assimétricas padronizadas. Assim a cadeia resultante que será considerada uma amostra da distribuição a posteriori é de tamanho 100 mil. Para o conjunto de dados selecionados temos $n = 1258$, o tempo computacional de cada modelo estimado foi muito longo 7 dias para a T-Student assimétrica padronizada e 6 dias para a Normal assimétrica padronizada. Como as cadeias de Markov geradas pelo modelo com distribuição nos erros GED assimétrica padronizada levam mais tempo, vamos trabalhar com cadeias de Markov de tamanho 100 mil onde as 50 mil primeiras amostras serão descartadas como amostra de aquecimento. Assim este modelo terá uma cadeia resultante de tamanho 50 mil. O tempo computacional deste modelo é de 10 dias. O tempo de consumo muito elevado para executar as simulações, pode-se justificar pelo uso do pacote de alta precisão **Rmpfr** no algoritmo para poder trabalhar com dados muito pequenos ou muito grandes devido a que o algoritmo foi usado sem usar algumas estratégias para evitar trabalhar com os produtórios que irão aparecer no cálculo da função de verossimilhança.

A seguir utilizaremos as distribuições assimétricas para os erros dos modelos GARCH(1,1) e os aplicaremos a dois conjuntos de dados reais.

5.1 Série IBOVESPA

O conjunto de dados utilizados consiste nos valores de fechamento diário do índice, no período de 3 de janeiro de 2012 a 30 de dezembro de 2016, totalizando 1258 observações. Os gráficos da série, dos retornos, das autocorrelações dos retornos e das autocorrelações dos retornos ao quadrado são apresentados na [Figura 6](#). Observe que mesmo tendo uma tendência de decrescimento até o ano 2016 e uma forte tendência de crescimento desde o início do ano 2016 na série (ver a [Figura 6 \(a\)](#)), os retornos são não correlacionados como se observa na [Figura 6 \(c\)](#). Por outro lado, os retornos ao quadrado possuem autocorrelação para algumas defasagens observe na [Figura 6 \(d\)](#), como vimos no [Capítulo 2](#) essas características são comuns em retornos de séries de tempo financeiras.

Observando os gráficos das cadeias obtidas, há evidência que as cadeias geradas para os tipos de modelos com distribuição no erro T-Student e Normal assimétrica padronizada convergiram, a exceção da cadeia obtida para o modelo com distribuição no erro GED assimétrica padronizada. Os traços da Cadeia de Markov, as densidades aproximadas e os gráficos das autocorrelações de todos os parâmetros do modelo GARCH(1,1) para cada tipo de distribuição do erro podem ser observados na [Figura 11](#), na [Figura 12](#) e na [Figura 13](#).

Portanto, fazendo uma análise das amostras obtidas dos modelos com distribuição no erro T-Student e Normal assimétrica padronizada. Temos que os retornos do IBOVESPA apresentam assimetria a direita, veja que uma boa parte da massa de probabilidade do γ está a direita do

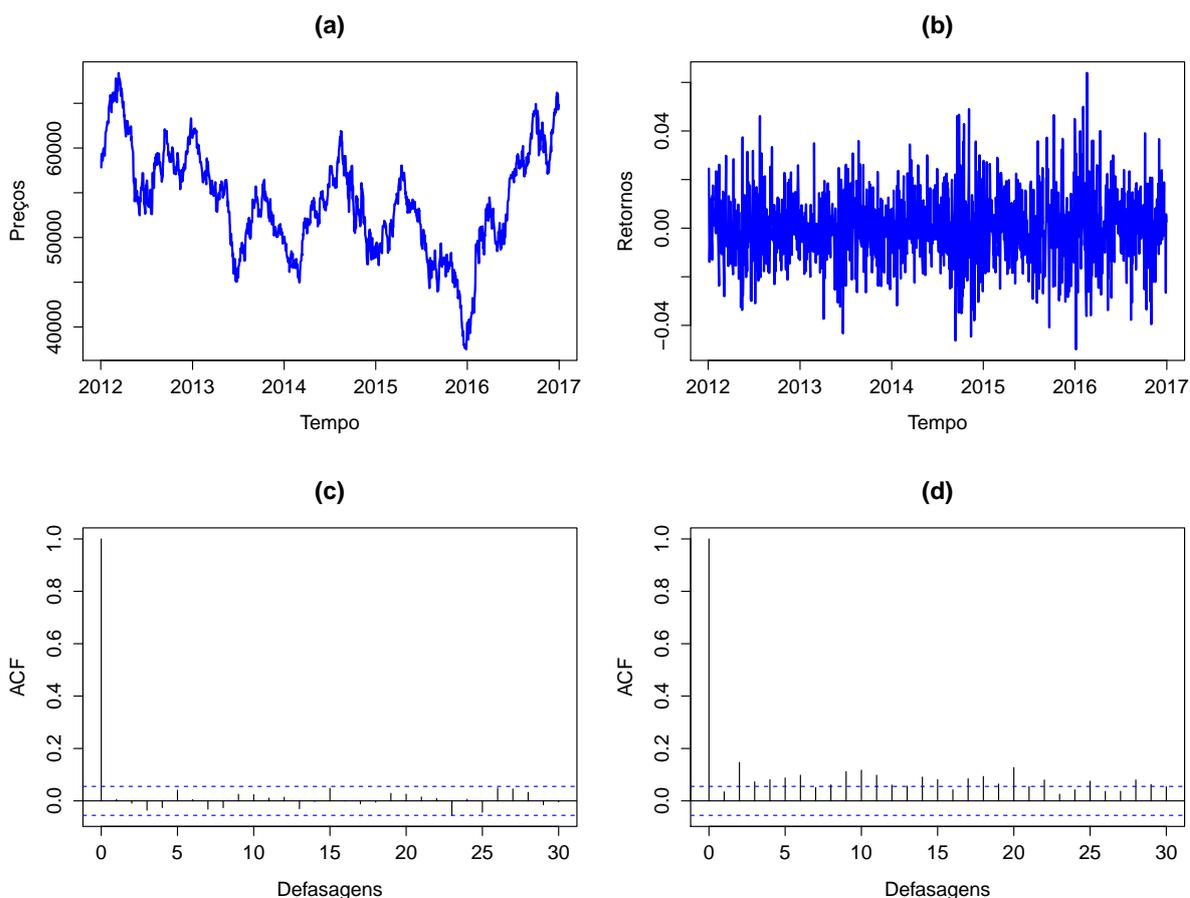


Figura 6 – Série de IBOVESPA de 03/01/2012 até 30/12/2016. **(a)** Gráfico da série; **(b)** Gráfico dos retornos; **(c)** Gráfico das autocorrelações dos retornos e **(d)** Gráfico das autocorrelações do quadrado dos retornos.

1 (valor de simetria). Além disso segundo a estimação do parâmetro que modela a cauda, a distribuição dos retornos se aproxima a uma distribuição normal.

As estimativas das médias a posteriori dos parâmetros dos modelos, estão disponíveis nas seguintes tabelas: [Tabela 7](#), [Tabela 8](#) e [Tabela 9](#) respectivamente.

Tabela 7 – Estimativas do GARCH(1,1) com erros T-Student assimétricos padronizados de IBOVESPA.

Parâmetros	Média	Desv. Pad.	Perc.(2.5%)	Mediana	Perc.(97.5%)
ν	15.65	4.48	8.75	15.00	26.06
γ	1.04	0.04	0.96	1.04	1.12
ω	8.29e-06	3.43e-06	3.33e-06	7.74e-06	16.44e-06
α	0.06	0.01	0.03	0.06	0.09
β	0.91	0.02	0.85	0.91	0.94

Tabela 8 – Estimativas do GARCH(1,1) com erros Normal assimétricos padronizados de IBOVESPA.

Parâmetros	Média	Desv. Pad.	Perc.(2.5%)	Mediana	Perc.(97.5%)
γ	1.04	0.04	0.96	1.04	1.12
ω	7.57e-06	2.96e-06	3.03e-06	7.12e-06	14.84e-06
α	0.06	0.01	0.03	0.05	0.08
β	0.91	0.02	0.86	0.91	0.94

Tabela 9 – Estimativas do GARCH(1,1) com erros GED assimétricos padronizados de IBOVESPA.

Parâmetros	Média	Desv. Pad.	Perc.(2.5%)	Mediana	Perc.(97.5%)
δ	3.68	0.11	3.47	3.68	3.93
γ	1.04	0.04	0.97	1.04	1.12
ω	8.77e-06	2.53e-06	4.88e-06	8.51e-06	14.73e-06
α	0.06	0.02	0.03	0.05	0.09
β	0.91	0.02	0.86	0.91	0.94

5.2 Série S&P500

O conjunto de dados utilizados consiste nos valores de fechamento diário do índice, no período de 3 de janeiro de 2012 a 30 de dezembro de 2016, totalizando 1258 observações. Os gráficos da série, dos retornos, das autocorrelações dos retornos e das autocorrelações dos retornos ao quadrado são apresentados na [Figura 7](#). Observe que há uma forte tendência de crescimento na série ([Figura 7 \(a\)](#)), mesmo assim os retornos não apresentaram autocorrelações como se observa na [Figura 7 \(c\)](#). Por outro lado, os retornos ao quadrado possuem autocorrelação para algumas defasagens observe na [Figura 7 \(d\)](#), como vimos no [Capítulo 2](#) essas características são comuns em retornos de séries de tempo financeiras.

Observando os gráficos das cadeias obtidas, há evidência que as cadeias geradas para os modelos com distribuição no erro T-Student, Normal e GED assimétrica padronizada convergiram. Os traços da Cadeia de Markov, as densidades aproximadas e os gráficos das autocorrelações de todos os parâmetros do modelo GARCH(1,1) para cada tipo de distribuição do erro podem ser observados na [Figura 14](#), na [Figura 15](#) e na [Figura 16](#).

Portanto, fazendo uma análise das amostras obtidas em cada modelo. Cada um dos modelos concluem que os retornos da série do S&P500 apresentam assimetria a esquerda, veja que praticamente toda a massa de probabilidade do γ está a esquerda do 1 (valor de simetria). Além disso segundo as estimações dos parâmetros que modelam as caudas, a distribuição dos retornos tem caudas pesadas.

As estimativas das médias a posteriori dos parâmetros dos modelos, estão disponíveis nas seguintes tabelas: [Tabela 10](#), [Tabela 11](#) e [Tabela 12](#).

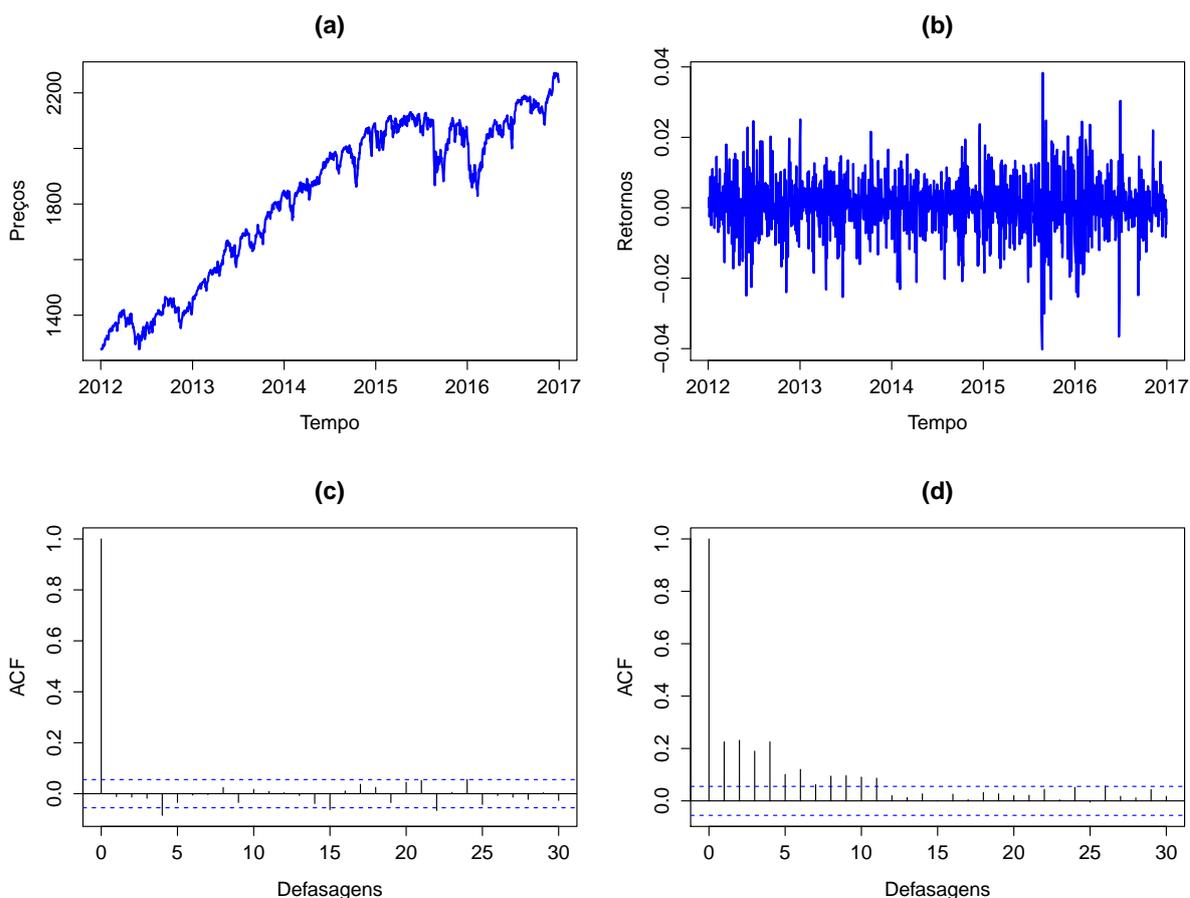


Figura 7 – Série de S&P500 de 03/01/2012 até 30/12/2016. **(a)** Gráfico da série; **(b)** Gráfico dos retornos; **(c)** Gráfico das autocorrelações dos retornos e **(d)** Gráfico das autocorrelações do quadrado dos retornos.

Tabela 10 – Estimativas do GARCH(1,1) com erros T-Student assimétricos padronizados de S&P500.

Parâmetros	Média	Desv. Pad.	Perc.(2.5%)	Mediana	Perc.(97.5%)
ν	6.83	1.54	4.65	6.55	10.63
γ	0.90	0.03	0.84	0.90	0.96
ω	6.95e-06	1.73e-06	4.13e-06	6.76e-06	10.88e-06
α	0.19	0.03	0.12	0.18	0.26
β	0.73	0.04	0.64	0.73	0.80

Com respeito às autocorrelações altas que se observam nas cadeias geradas para alguns parâmetros em cada um dos modelos propostos e para cada uma das séries analisadas, poderiam se melhorar fazendo uso de outras distribuições a priori ou usando outra distribuição proposta.

5.3 Estimação da Volatilidade

Para concluir, a seguir se apresenta a estimação da volatilidade para as duas séries mencionadas acima usando os três tipos de modelos propostos. A [Figura 8](#), mostra a evolução

Tabela 11 – Estimativas do GARCH(1,1) com erros Normal assimétricos padronizados de S&P500.

Parâmetros	Média	Desv. Pad.	Perc.(2.5%)	Mediana	Perc.(97.5%)
γ	0.88	0.03	0.83	0.88	0.94
ω	7.73e-06	1.55e-06	5.13e-06	7.58e-06	11.19e-06
α	0.17	0.03	0.12	0.17	0.23
β	0.72	0.04	0.64	0.72	0.79

Tabela 12 – Estimativas do GARCH(1,1) com erros GED assimétricos padronizados de S&P500.

Parâmetros	Média	Desv. Pad.	Perc.(2.5%)	Mediana	Perc.(97.5%)
δ	3.34	0.08	3.20	3.34	3.51
γ	0.93	0.03	0.87	0.93	0.99
ω	7.67e-06	1.84e-06	4.67e-06	7.50e-06	11.75e-06
α	0.18	0.035	0.18	0.18	0.25
β	0.71	0.05	0.62	0.72	0.80

dos valores absolutos dos retornos versus a raiz quadrada da estimação da volatilidade indicada pela linha vermelha acompanhada de duas linhas tracejadas de cor azul que representam os percentis 0.025 e 0.975 das amostras de tamanho 100 mil das volatilidades estimadas (raiz quadrada) para cada período de tempo, obtidas usando as amostras de tamanho 100 mil geradas pelo algoritmo M-H com um só bloco para cada parâmetro do modelo. Para a série Ibovespa como para a série S&P500, usando o modelo GARCH(1,1) com erros com distribuição T-Student assimétrica padronizada. Nas seguintes figuras: [Figura 9](#) e [Figura 10](#), observa-se o mesmo que na figura anteriormente mencionada mas usando o modelo GARCH(1,1) com erros com distribuição Normal assimétrica padronizada e usando o modelo GARCH(1,1) com erros com distribuição GED assimétrica padronizada (neste caso só muda o tamanho amostral a 50 mil) respectivamente.

Observa-se que no caso dos três tipos de modelos usados, a volatilidade da série S&P500 foi a que melhor foi estimada e o intervalo formado pelos percenties 0.025 e 0.975 indica baixa incerteza das estimativas da raiz quadrada das volatilidades.

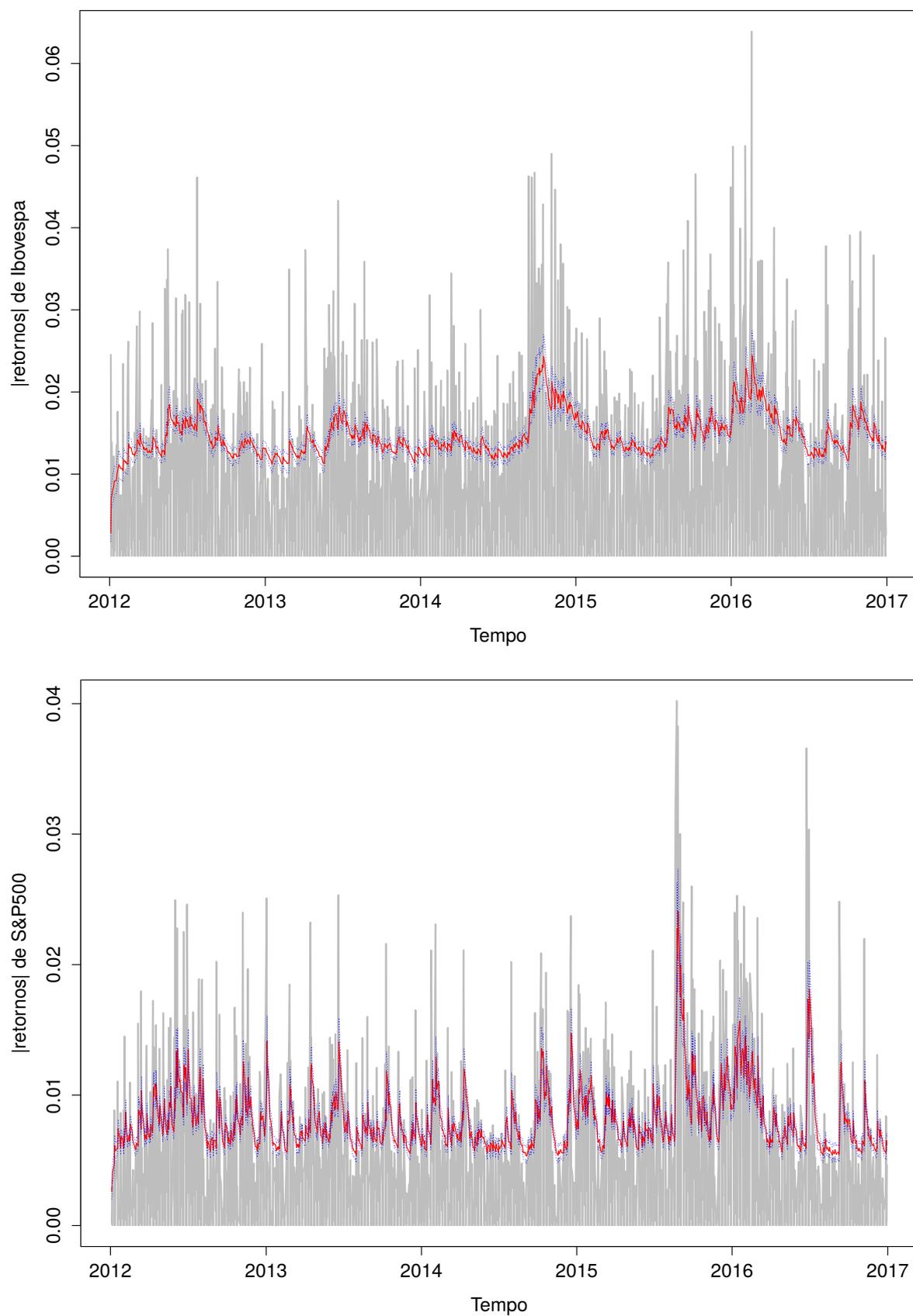


Figura 8 – Retornos e Volatilidades estimadas usando o modelo GARCH(1,1) com erros com distribuição T-Student assimétricos padronizados.

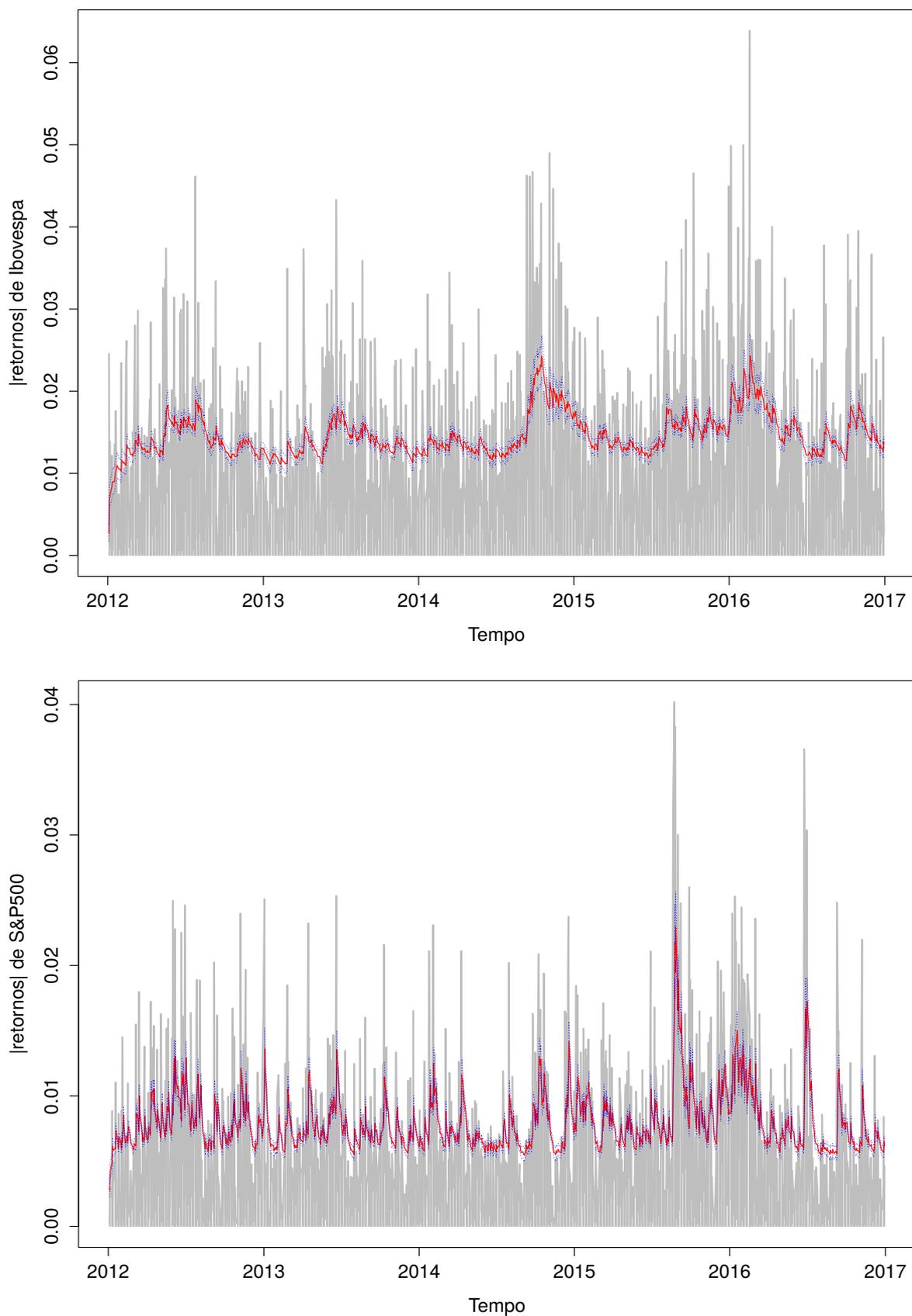


Figura 9 – Retornos e Volatilidades estimadas usando o modelo GARCH(1,1) com erros com distribuição Normal assimétricos padronizados.

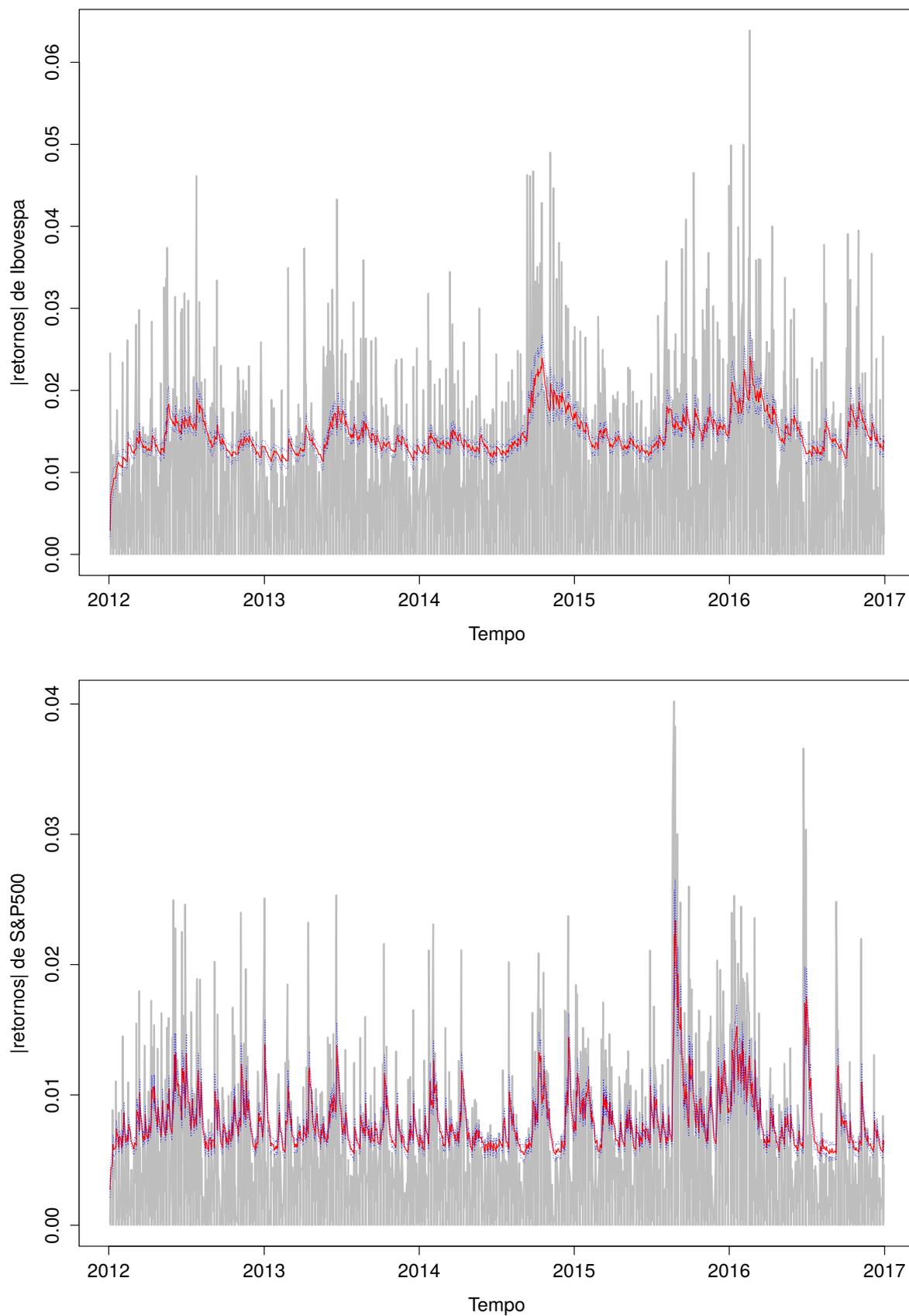


Figura 10 – Retornos e Volatilidades estimadas usando o modelo GARCH(1,1) com erros com distribuição GED assimétricos padronizados.

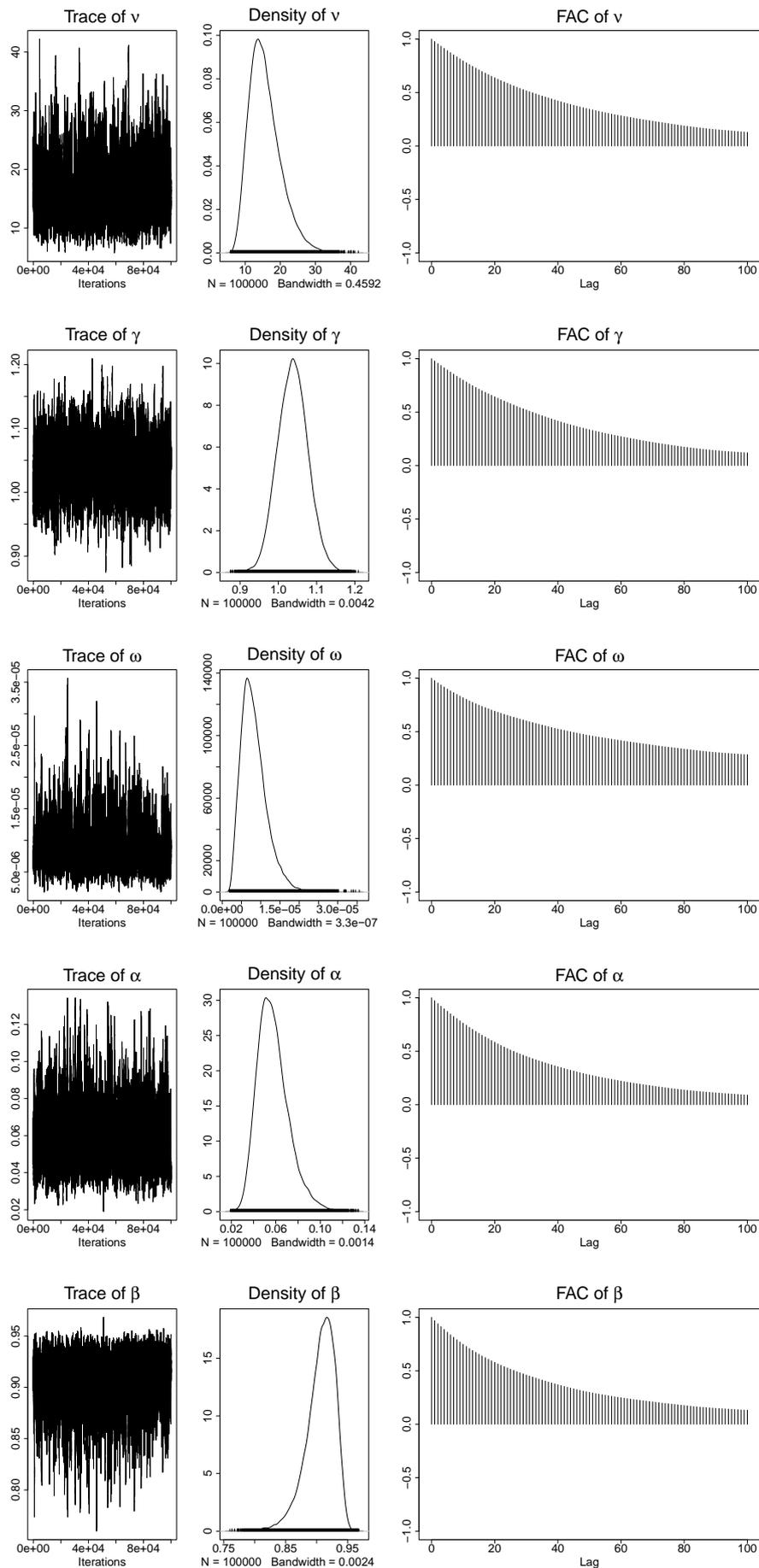


Figura 11 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série IBOVESPA. Modelo ajustado: GARCH(1,1) com erros T-Student assimétricos padronizados.

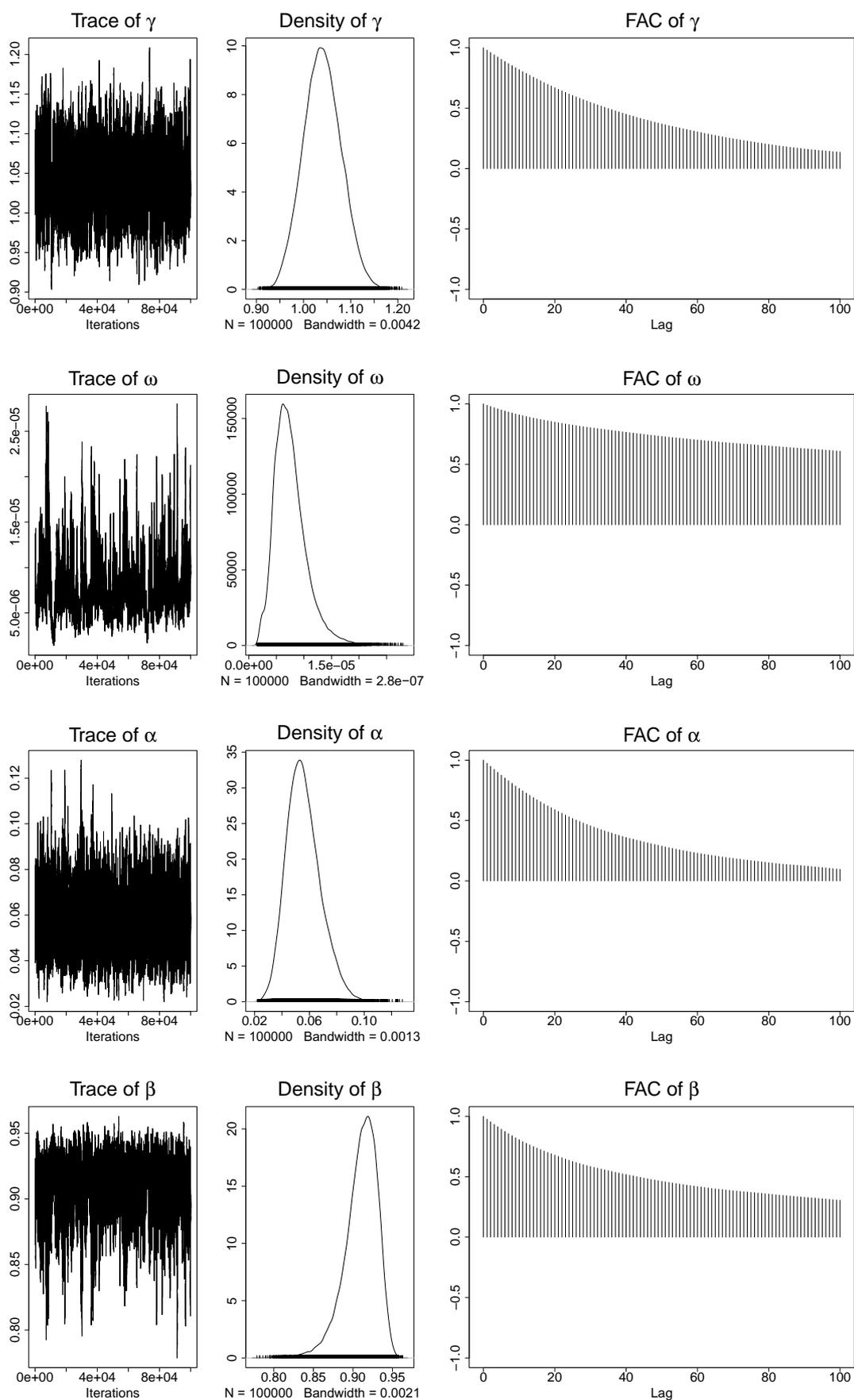


Figura 12 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série IBOVESPA. Modelo ajustado: GARCH(1,1) com erros Normal assimétricos padronizados.

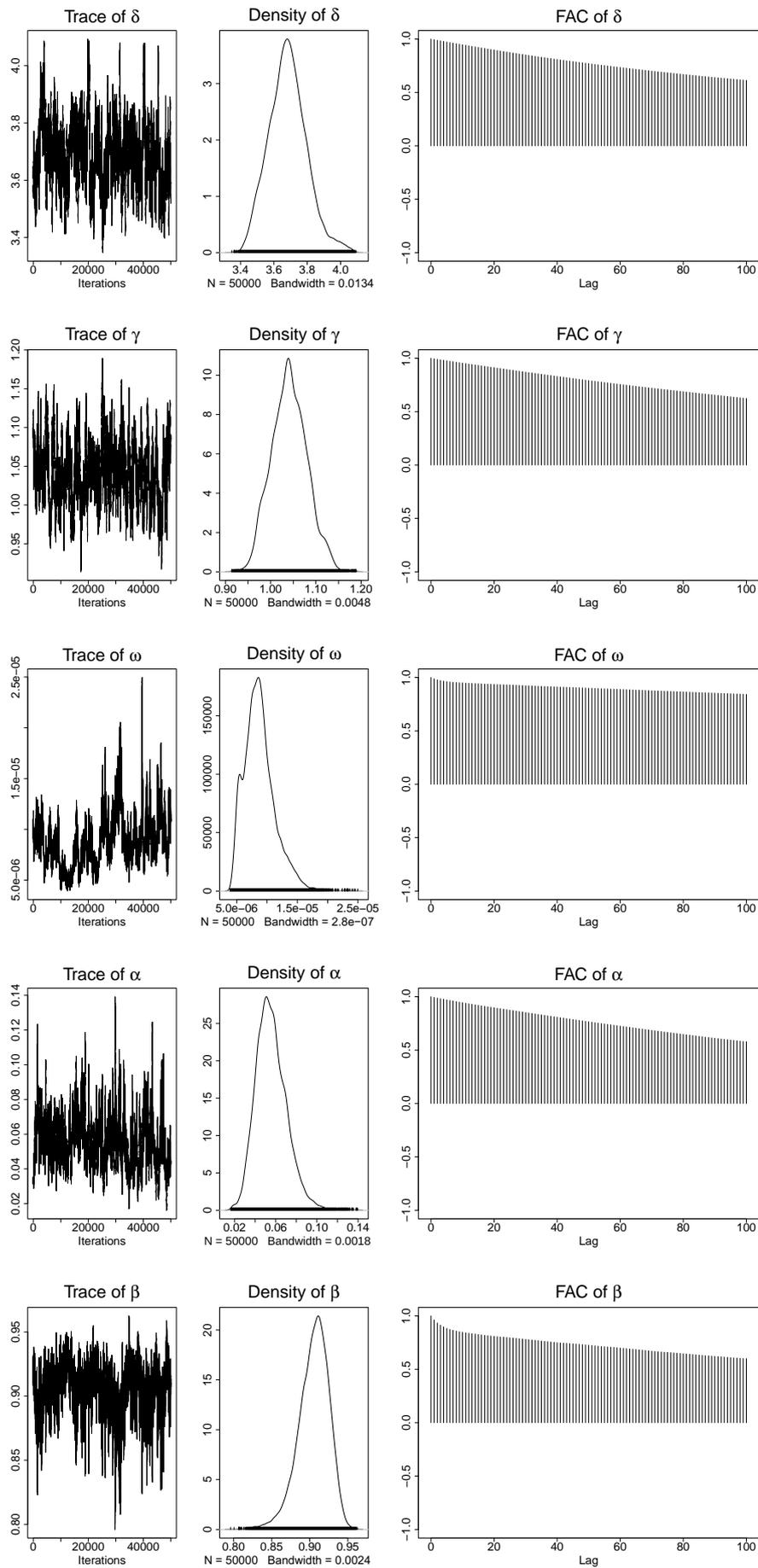


Figura 13 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série IBOVESPA. Modelo ajustado: GARCH(1,1) com erros GED assimétricos padronizados.

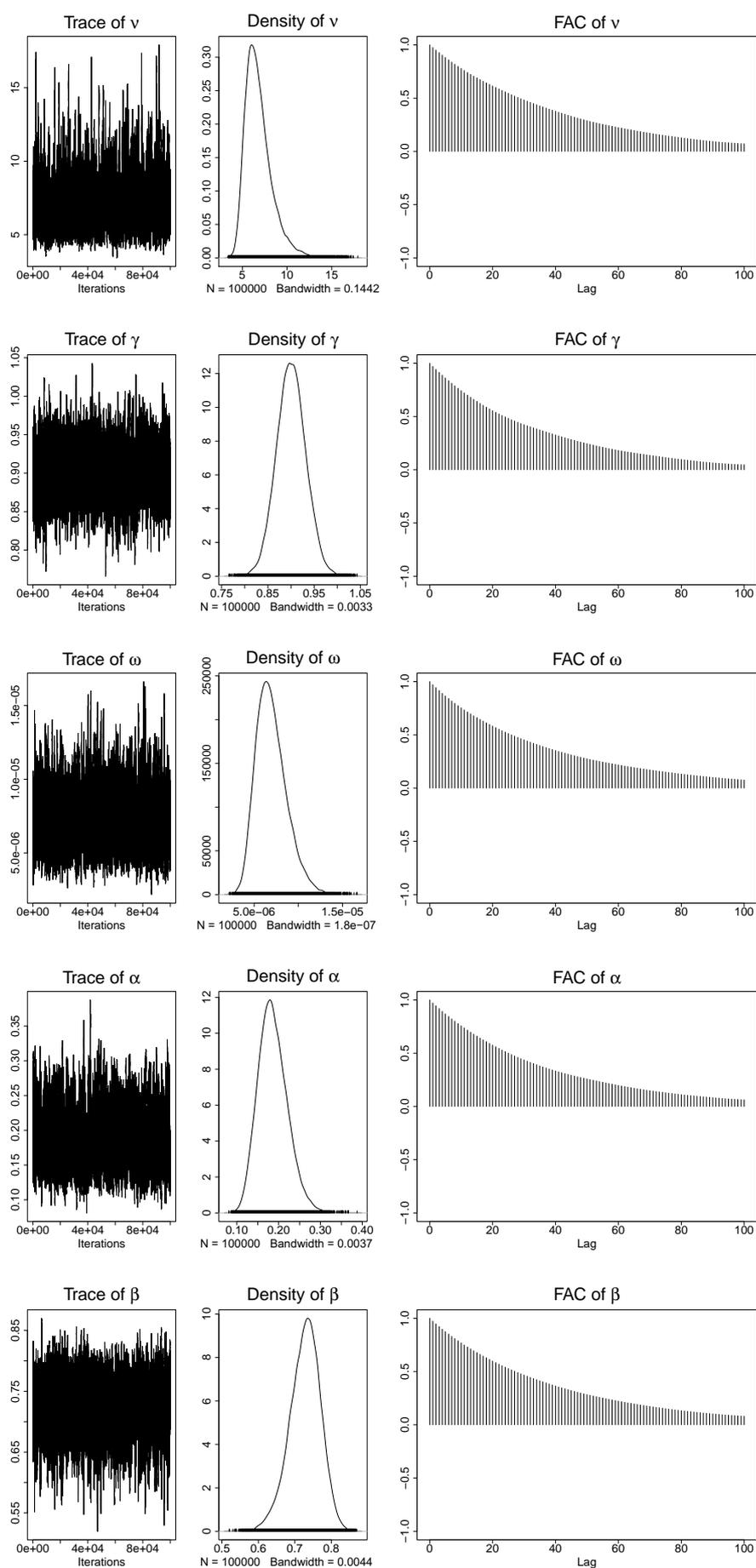


Figura 14 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série S&P500. Modelo ajustado: GARCH(1,1) com erros T-Student assimétricos padronizados.

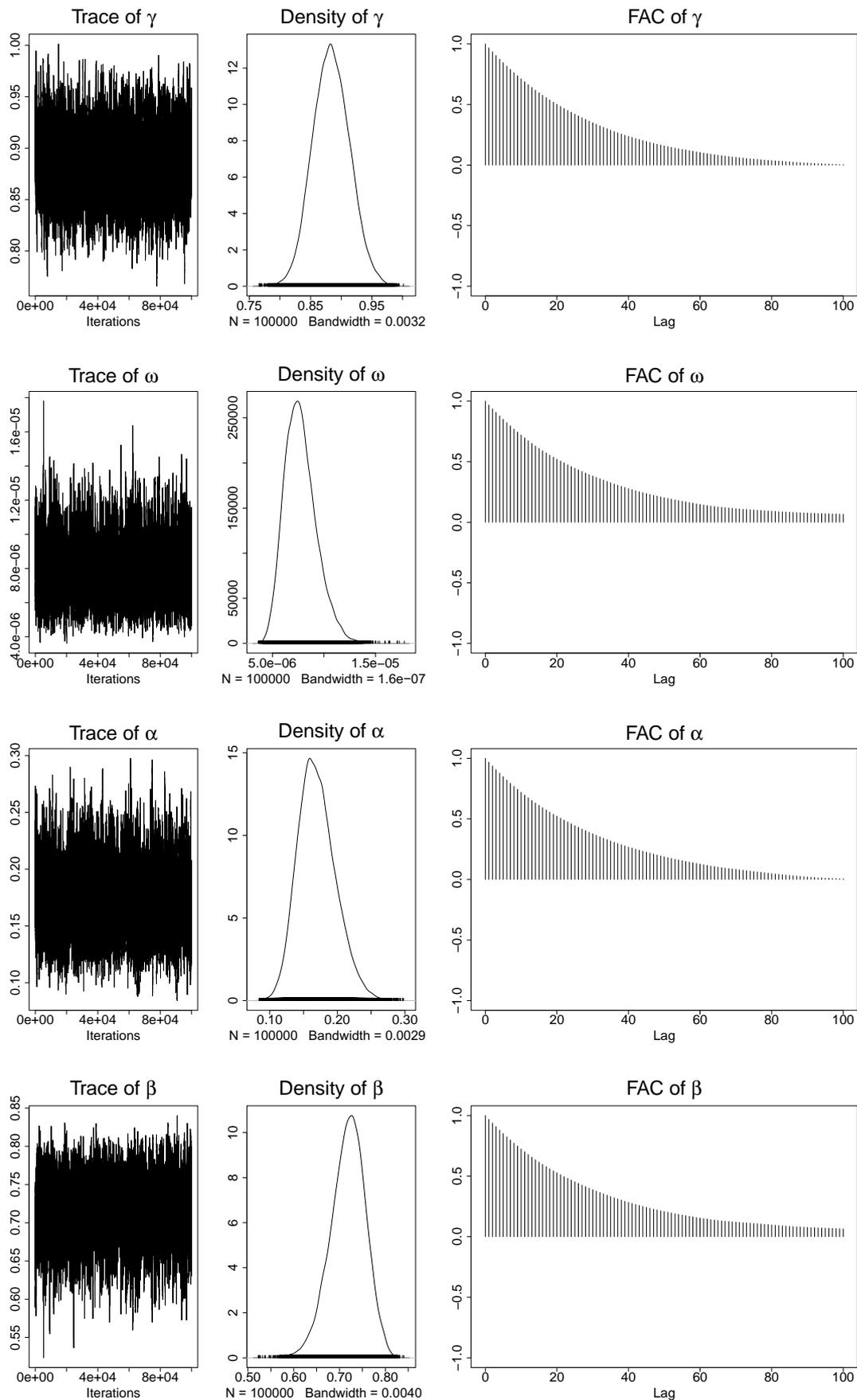


Figura 15 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série S&P500. Modelo ajustado: GARCH(1,1) com erros Normal assimétricos padronizados.

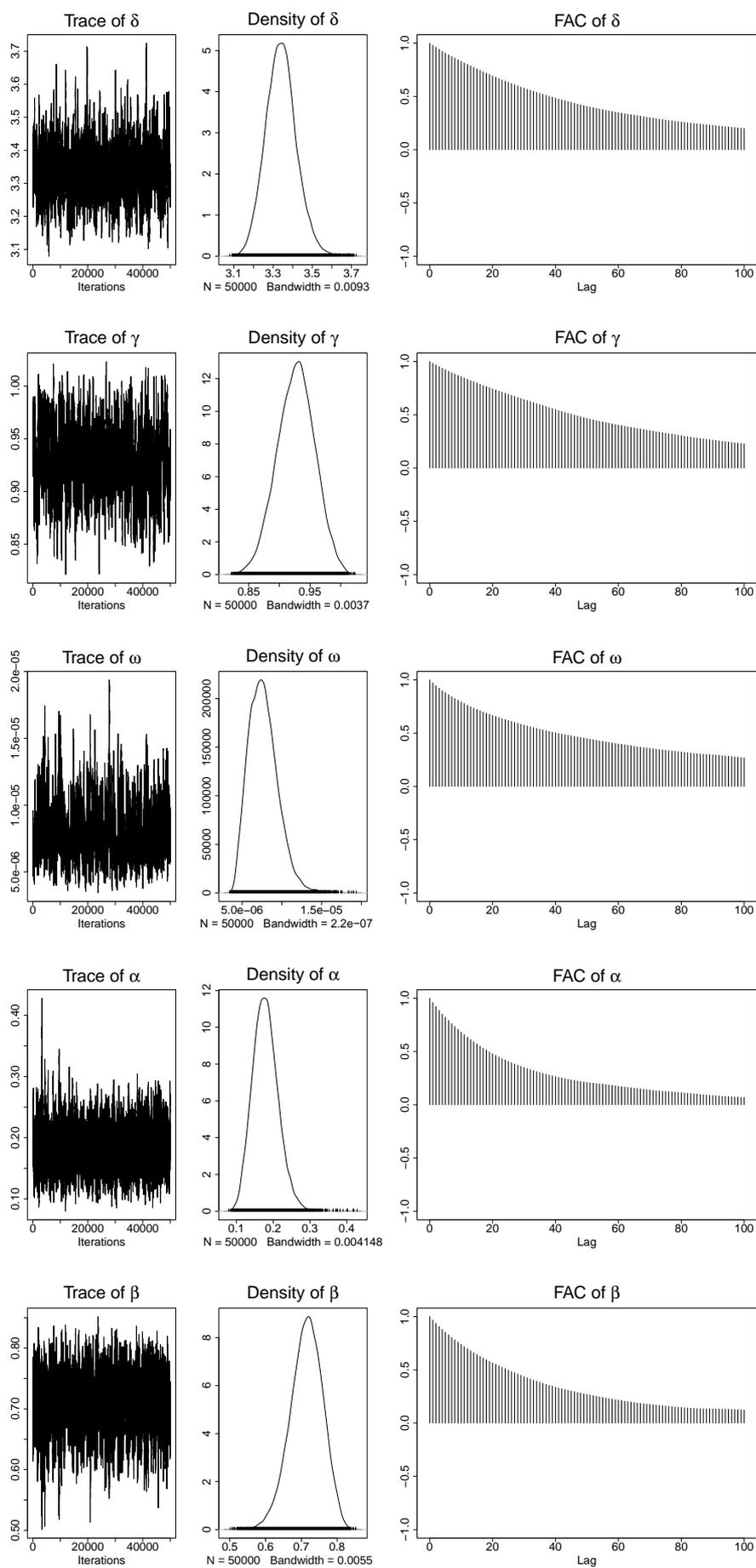


Figura 16 – Traços da cadeia, densidades aproximadas e autocorrelações dos parâmetro para a série S&P500. Modelo ajustado: GARCH(1,1) com erros GED assimétricos padronizados.

CONCLUSÕES

Nesta dissertação estudamos o modelo GARCH(1,1) considerando uma nova extensão do modelo, que considera assimetria na distribuição dos erros estocásticos do modelo. Na literatura as distribuições de probabilidade mais utilizadas para os erros do modelo GARCH são a Normal, a T-Student e a GED em suas versões padronizadas. Para inserir assimetria à distribuição do erro aplicamos o método de Fernandez & Steel (1998), obtendo assim versões assimétricas das distribuições usadas geralmente.

A abordagem Bayesiana para a estimação desses modelos gerados, traz consigo algumas vantagens como facilidade de interpretação dos parâmetros e possibilidade de inserir informação a priori. Para a simulação das cadeias se analisaram os resultados do algoritmo de Metropolis-Hasting em blocos e num só bloco, os resultados melhor comportados foram os do algoritmo de um só bloco, foi por isso que na aplicação se usou este algoritmo. As autocorrelações de alguns parâmetros ficaram altas, poderia-se dever ao tipo de distribuição proposta ou à distribuição a priori utilizada. Segundo nossa experiência, rodar este tipo de algoritmos no software R leva muito mais tempo do que demoraria se fosse programado num software de baixo nível. Portanto se recomenda usar outro software para implementar este tipo de algoritmos.

A implementação dos três tipos de modelos, corroborou a existência de um dos novos fatos estilizados declarado em algumas pesquisas empíricas: a assimetria da distribuição dos retornos. A generalização deste tipo de modelos GARCH ao nível multivariado já foi desenvolvida mas a utilização desta metodologia em outros modelos que são variantes do modelo GARCH pode se propor como um trabalho futuro.

REFERÊNCIAS

ABANTO-VALLE, C.; LACHOS, V.; DEY, D. K. Bayesian estimation of a skew-student-t stochastic volatility model. **Methodology and Computing in Applied Probability**, Springer, p. 1–18, 2013. Citado na página 25.

ARDIA, D. Bayesian estimation of the GARCH(1,1) model with normal innovations. **Student**, v. 5, n. 3–4, p. 283–298, 2006. URL <http://ssrn.com/abstract=1543409>. Citado nas páginas 25 e 50.

_____. **Financial Risk Management with Bayesian Estimation of GARCH Models: Theory and Applications**. Berlin, Germany: Springer-Verlag, 2008. v. 612. (Lecture Notes in Economics and Mathematical Systems, v. 612). ISBN 978-3-540-78656-6, e-ISBN 978-3-540-78657-3, doi 10.1007/978-3-540-78657-3. Citado na página 26.

_____. Bayesian estimation of a markov-switching threshold asymmetric garch model with student-t innovations. **The Econometrics Journal**, Wiley Online Library, v. 12, n. 1, p. 105–126, 2009. Citado na página 25.

ARDIA, D.; HOOGERHEIDE, L. F. Estimation of the garch (1, 1) model with student-t innovations. **R Journal**, v. 2, n. 2, 2010. Citado na página 26.

BAROSSO-FILHO, M.; ACHCAR, J. A.; SOUZA, R. M. d. Modelos de volatilidade estocástica em séries financeiras: uma aplicação para o ibovespa. **Economia Aplicada**, SciELO Brasil, v. 14, n. 1, p. 25–40, 2010. Citado na página 25.

BAUWENS, L.; LUBRANO, M. Bayesian inference on garch models using the gibbs sampler. **The Econometrics Journal**, Wiley Online Library, v. 1, n. 1, p. 23–46, 1998. Citado nas páginas 25 e 41.

BAUWENS, L.; PREMINGER, A.; ROMBOUTS, J. V. Theory and inference for a markov switching garch model. **The Econometrics Journal**, Wiley Online Library, v. 13, n. 2, p. 218–244, 2010. Citado na página 25.

BLACK, F. Studies of stock price volatility changes. **Business and Economic Statistics**, Proceedings from the American Statistical Association, p. 177–181, 1976. Citado na página 33.

BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of econometrics**, Elsevier, v. 31, n. 3, p. 307–327, 1986. Citado nas páginas 23, 24, 41, 42 e 44.

_____. Glossary to arch (garch). In: CITESEER. **Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle**. [S.l.], 2009. Citado na página 24.

BOLLERSLEV, T.; CHOU, R. Y.; KRONER, K. F. Arch modeling in finance: A review of the theory and empirical evidence. **Journal of econometrics**, Elsevier, v. 52, n. 1, p. 5–59, 1992. Citado na página 41.

- CAPPUCCIO, N.; LUBIAN, D.; RAGGI, D. Investigating asymmetry in us stock market indexes: evidence from a stochastic volatility model. **Applied Financial Economics**, Taylor & Francis, v. 16, n. 6, p. 479–490, 2006. Citado na página 25.
- CASELLA, G.; BERGER, R. L. **Inferência estatística**. [S.l.]: São Paulo: Cengage Learning, 2010. Citado na página 35.
- CHATFIELD, C. **The analysis of time series: an introduction**. [S.l.]: CRC press, 2013. Citado nas páginas 30 e 31.
- COWLES, M. K.; CARLIN, B. P. Markov chain monte carlo convergence diagnostics: a comparative review. **Journal of the American Statistical Association**, Taylor & Francis, v. 91, n. 434, p. 883–904, 1996. Citado na página 39.
- DING, Z.; GRANGER, C. W.; ENGLE, R. F. A long memory property of stock market returns and a new model. **Journal of empirical finance**, Elsevier, v. 1, n. 1, p. 83–106, 1993. Citado na página 24.
- EHLERS, R. S. Computational tools for comparing asymmetric garch models via bayes factors. **Mathematics and Computers in Simulation**, Elsevier, v. 82, n. 5, p. 858–867, 2012. Citado na página 25.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 987–1007, 1982. Citado nas páginas 23, 24, 25 e 41.
- FAMA, E. F. The behavior of stock-market prices. **The Journal of Business**, JSTOR, v. 38, n. 1, p. 34–105, 1965. Citado na página 34.
- FERNÁNDEZ, C.; STEEL, M. F. On bayesian modeling of fat tails and skewness. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 93, n. 441, p. 359–371, 1998. Citado nas páginas 25, 26, 45, 48 e 50.
- GAMERMAN, D.; LOPES, H. F. **Markov chain Monte Carlo: stochastic simulation for Bayesian inference**. [S.l.]: CRC Press, 2006. Citado nas páginas 36 e 57.
- GELFAND, A. E.; SMITH, A. F. Sampling-based approaches to calculating marginal densities. **Journal of the American statistical association**, Taylor & Francis Group, v. 85, n. 410, p. 398–409, 1990. Citado na página 36.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, n. 6, p. 721–741, 1984. Citado na página 36.
- GEWEKE, J. Exact predictive densities for linear models with arch disturbances. **Journal of Econometrics**, Elsevier, v. 40, n. 1, p. 63–86, 1989. Citado na página 25.
- _____. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: **IN BAYESIAN STATISTICS**. [S.l.]: University Press, 1992. p. 169–193. Citado na página 39.
- GEWEKE, J. *et al.* Bayesian comparison of econometric models. **Federal Reserve bank of Minneapolis working paper**, Citeseer, n. 532, 1994. Citado na página 25.

- GILKS, W.; RICHARDSON, S.; SPIEGELHALTER, D. **Markov Chain Monte Carlo in Practice**, ser. **Interdisciplinary Statistics Series**. [S.l.]: Chapman and Hall, 1996. Citado na página 39.
- GLOSTEN, L.; JAGANNATHAN, R.; RUNKLE, D. On the relation between expected value and the volatility of the nominal excess return on stock. **Journal of Finance**, v. 48, p. 1779–1801, 1993. Citado na página 24.
- HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. **Biometrika**, Biometrika Trust, v. 57, n. 1, p. 97–109, 1970. Citado nas páginas 26 e 37.
- JACQUIER, E.; POLSON, N. G.; ROSSI, P. E. Bayesian analysis of stochastic volatility models. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 20, n. 1, p. 69–87, 2002. Citado na página 25.
- JONDEAU, E.; ROCKINGER, M. Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. **Journal of Economic Dynamics and Control**, Elsevier, v. 27, n. 10, p. 1699–1737, 2003. Citado na página 25.
- KIM, S.; SHEPHARD, N.; CHIB, S. Stochastic volatility: likelihood inference and comparison with arch models. **The Review of Economic Studies**, Oxford University Press, v. 65, n. 3, p. 361–393, 1998. Citado na página 25.
- LAURENT, S. Analytical derivatives of the aparch model. **Computational Economics**, Springer, v. 24, n. 1, p. 51–57, 2004. Citado na página 24.
- MANDELROT, B. The variation of certain speculative prices. **The Journal of Business**, The University of Chicago Press, v. 36, n. 4, p. 394–419, 1963. Citado na página 34.
- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. **The journal of chemical physics**, AIP Publishing, v. 21, n. 6, p. 1087–1092, 1953. Citado nas páginas 26, 37 e 38.
- MIGON, H. S.; GAMERMAN, D.; LOUZADA, F. **Statistical inference: an integrated approach**. [S.l.]: CRC press, 2014. Citado na página 35.
- MORETTIN, P. A. **Econometria financeira: um curso em séries temporais financeiras**. [S.l.]: ABE, 2006. Citado na página 30.
- NAKATSUMA, T. A markov-chain sampling algorithm for garch models. **Studies in Nonlinear Dynamics and Econometrics**, v. 3, n. 2, p. 107–117, 1998. Citado na página 25.
- _____. Bayesian analysis of arma–garch models: A markov chain sampling approach. **Journal of Econometrics**, Elsevier, v. 95, n. 1, p. 57–69, 2000. Citado na página 41.
- NELSON, D. B. Conditional heteroskedasticity in asset returns: A new approach. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 347–370, 1991. Citado nas páginas 24, 26 e 44.
- PIPIEN, M. Bayesian comparison of garch processes with skewness mechanism in conditional distributions. **arXiv preprint physics/0606253**, 2006. Citado na página 25.
- PRADO, R.; WEST, M. **Time series: modeling, computation, and inference**. [S.l.]: CRC Press, 2010. Citado na página 49.

ROBERT, C. P. **Monte carlo methods**. [S.l.]: Wiley Online Library, 2004. Citado na página 36.

STRAUMANN, D. **Estimation in Conditionally Heteroscedastic Time Series Models**. [S.l.]: Springer, 2005. Citado na página 31.

TAKAISHI, T. Bayesian estimation of garch model by hybrid monte carlo. **arXiv preprint physics/0702240**, 2007. Citado na página 26.

TAYLOR, S. **Modelling financial time series**. World Scientific Publishing, 2008. Citado na página 30.

TAYLOR, S. J. Financial returns modelled by the product of two stochastic processes—a study of the daily sugar prices 1961-75. **Time series analysis: theory and practice**, North-Holland, v. 1, p. 203–226, 1982. Citado na página 25.

TIERNEY, L. Markov chains for exploring posterior distributions. **the Annals of Statistics**, JSTOR, p. 1701–1728, 1994. Citado na página 40.

TSAY, R. S. **Analysis of financial time series**. [S.l.]: John Wiley & Sons, 2005. v. 543. Citado na página 29.

TSIOTAS, G. On generalised asymmetric stochastic volatility models. **Computational Statistics & Data Analysis**, Elsevier, v. 56, n. 1, p. 151–172, 2012. Citado na página 25.