

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

RAFAEL SOARES PAIXÃO

**MÉTODO ZERO-VARIANCE PARA MONTE CARLO HAMILTONIANO APLICADO A MODELOS  
GARCH UNIVARIADOS E MULTIVARIADOS**

Tese 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 Doutor em Estatística - Programa Interinstitucional de Pós-Graduação em Estatística UFSCar-USP.

Orientador: Prof. Dr. Ricardo Sandes Ehlers

**São Carlos  
Junho de 2021**



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

RAFAEL SOARES PAIXÃO

**ZERO-VARIANCE METHOD FOR HAMILTONIAN MONTE CARLO APPLIED TO UNIVARIATE AND  
MULTIVARIATE GARCH MODELS**

Doctoral 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 Doctorate Joint Graduate Program in Statistics UFSCar-USP

Advisor: Prof. Dr. Ricardo Sandes Ehlers

**São Carlos  
June 2021**



# UNIVERSIDADE FEDERAL DE SÃO CARLOS

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

---

## Folha de Aprovação

---

Defesa de Tese de Doutorado do candidato Rafael Soares Paixão, realizada em 13/05/2021.

### Comissão Julgadora:

Prof. Dr. Ricardo Sandes Ehlers (USP)

Prof. Dr. Luiz Koodi Hotta (UNICAMP)

Prof. Dr. Marcio Poletti Laurini (USP)

Profa. Dra. Verónica Andrea González López (UNICAMP)

Prof. Dr. Carlos Antonio Abanto-Valle (UFRJ)

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

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa Interinstitucional de Pós-Graduação em Estatística.

*Dedico este trabalho à minha família e ao meu companheiro, que foram meu maior apoio nos momentos de angústia.*



# AGRADECIMENTOS

---

---

*O presente trabalho foi realizado com apoios da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 e do Centro de Ciências Matemáticas Aplicadas à Indústria (CeMEAI).*

Agradeço inicialmente a Deus; aos meus familiares: meu pai, Ramilson; minha irmã, Samara; meu irmão, Francisco Ramilson; meu cunhado, Fernando Alisson; e ao meu companheiro, Paulo Henrique, por me darem forças em vários momentos de fraqueza. Sem eles, talvez eu não teria como ter concluído este trabalho.

Agradeço ao meu orientador, Prof. Dr. Ricardo Sandes Ehlers, por ter me aceitado como orientando e pela confiança depositada em mim.

A toda a comunidade do ICMC e UFSCar; a todos os colegas de curso do doutorado, principalmente ao Rogério, ao Marco Pollo, ao Juan, ao Carlos, ao Helio, ao Fabiano, pelas conversas que tivemos e lutas que compartilhamos e pelo meu amigo "Guto" pela amizade e apoio.

Por fim, agradeço aos meus antigos professores, que sempre foram a minha inspiração: à Profa. Dra. Clélia Toloí, minha orientadora de mestrado; e ao Prof. João Maurício Mota, por ter sido a base de tudo que aprendi na área de Probabilidade.





*“The only constant in the universe is change.”*  
*(Mary Pipher)*



# RESUMO

PAIXÃO, R. S. **Método *Zero-Variance* para Monte Carlo Hamiltoniano aplicado a modelos GARCH univariados e multivariados**. 2021. 124 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2021.

Este trabalho de doutorado desenvolve, compara e aplica métodos Monte Carlo via Cadeias de Markov (MCMC) para estimação de parâmetros em modelos GJR-GARCH univariados e multivariados. Especificamente, os seguintes problemas são abordados: (i) concepção de uma abordagem de estimação puramente bayesiana; (ii) desenvolvimento de um método bayesiano para maior eficiência computacional na estimação de parâmetros; e (iii) escolha flexível de distribuições de probabilidade de resíduos para modelos GJR-GARCH. Como resultado das investigações dos problemas mencionados, este trabalho apresenta quatro contribuições. A primeira diz respeito a uma abordagem de inferência bayesiana para modelos GJR-GARCH univariados e multivariados. A segunda consiste no estudo de três distribuições de probabilidade de resíduos, uma delas tendo sido utilizada de forma inovadora para casos multivariados. A terceira combina duas técnicas, o algoritmo Hamiltoniano Monte Carlo (HMC) e o método *Zero-Variance*, para possibilitar a estimação de parâmetros em modelos GJR-GARCH com maior eficácia dos estimadores, bem como com maior eficiência computacional. Por fim, a quarta apresenta resultados de estudos de simulação e de uma aplicação em dados reais, no contexto de índices de bolsas de valores mundiais, mostram que as contribuições propostas solucionam os problemas abordados eficaz e eficientemente, avançando o estado da arte de modelos GARCH univariados e multivariados.

**Palavras-chave:** Inferência Bayesiana, Monte Carlo Hamiltoniano, *Zero-Variance*, GARCH.



# ABSTRACT

PAIXÃO, R. S. **Zero-Variance method for Hamiltonian Monte Carlo applied to univariate and multivariate GARCH models**. 2021. 124p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2021.

This PhD work develops, compares and applies Monte Carlo Markov Chains (MCMC) methods for parameter estimation in univariate and multivariate GJR-GARCH models. Specifically, the following problems are addressed: **(i)** conception of a purely bayesian estimation approach; **(ii)** development of a bayesian method for higher computational efficiency in parameter estimation; and **(iii)** flexible selection of residual probability distributions for GJR-GARCH models. As a result from the investigations of the aforementioned problems, this work presents four contributions. The first corresponds to a bayesian inference approach for univariate and multivariate GJR-GARCH models. The second consists of studying three residual probability distributions, one of which having been inovatively employed for multivariate cases. The third combines two techniques, namely the Hamiltonian Monte Carlo (HMC) algorithm and the Zero-Variance method, to allow parameter estimation in GJR-GARCH models with higher estimator efficiency, as well as higher computational performance. Finally, the fourth presents results from simulation studies and an application over real-world data, in the context of worldwide stock market indexes, show that the proposed contributions solve the addressed problems effective and efficiently, advancing the state of the art of univariate and multivariate GARCH models.

**Keywords:** Bayesian Inference, Hamiltonian Monte Carlo, Zero-Variance, GARCH.



# LISTA DE ILUSTRAÇÕES

---

---

Figura 1 – Log-retornos diários da série S&P500.	26
Figura 2 – Função de autocorrelação dos log-retornos diários da série S&P500.	26
Figura 3 – Gráfico da série diária do S&P500.	76
Figura 4 – Gráfico dos log-retornos diários da série S&P500.	76
Figura 5 – Gráfico da função de autocorrelação dos log-retornos diários.	77
Figura 6 – Histograma dos log-retornos diários.	77
Figura 7 – Q-Q plot dos log-retornos diários.	78
Figura 8 – Gráfico da divergência de Kullback-Leibler para o modelo GJR-GARCH(1,1)- GED.	83
Figura 9 – Gráficos das séries temporais.	85
Figura 10 – Gráficos dos log-retornos diários das séries temporais.	86
Figura 11 – Gráficos da função de auto-correlação dos log-retornos diários.	87
Figura 12 – Comparação dos métodos para dois parâmetros do modelo DCC-GJR-GARCH(1,1)- GED.	90
Figura 13 – Gráfico da divergência de Kullback-Leibler para o modelo DCC-GJR-GARCH(1,1)- t.	91
Figura 14 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro $\omega$ .	109
Figura 15 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro $\phi$ .	110
Figura 16 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro $\beta$ .	110
Figura 17 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro $\mu$ .	111
Figura 18 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro $v$ .	111





# LISTA DE ALGORITMOS

---

---

Algoritmo 1 – Procedimento do método HMC para geração de dados com distribuição	
posteriori . . . . .	44
Algoritmo 2 – Metropolis-Hastings . . . . .	62



# LISTA DE TABELAS

Tabela 1 – Descrição dos valores da função perda. . . . .	35
Tabela 2 – Resultado das simulações do modelo GJR-GARCH(1,1)-NORMAL. . . . .	64
Tabela 3 – Resultado das simulações do modelo GJR-GARCH(1,1)-t. . . . .	65
Tabela 4 – Resultado das simulações do modelo GJR-GARCH(1,1)-GED. . . . .	66
Tabela 5 – Resultados do estimador HMC para o modelo DCC-GJR-GARCH(1,1)- NORMAL. . . . .	69
Tabela 6 – Resultados do estimador ZV-HMC-L para o modelo DCC-GJR-GARCH(1,1)- NORMAL. . . . .	70
Tabela 7 – Resultados do estimador ZV-HMC-Q para o modelo DCC-GJR-GARCH(1,1)- NORMAL. . . . .	70
Tabela 8 – Resultados do estimador HMC para o modelo DCC-GJR-GARCH(1,1)-t. . . . .	71
Tabela 9 – Resultados do estimador ZV-HMC-L para o modelo DCC-GJR-GARCH(1,1)-t. . . . .	71
Tabela 10 – Resultados do estimador ZV-HMC-Q para o modelo DCC-GJR-GARCH(1,1)- t. . . . .	72
Tabela 11 – Resultados do estimador HMC para o modelo DCC-GJR-GARCH(1,1)-GED. . . . .	72
Tabela 12 – Resultados do estimador ZV-HMC-L para o modelo DCC-GJR-GARCH(1,1)- GED. . . . .	73
Tabela 13 – Resultados do estimador ZV-HMC-Q para o modelo DCC-GJR-GARCH(1,1)- GED. . . . .	73
Tabela 14 – Estatísticas básicas da série diária dos log-retornos do S&P500 (%). . . . .	78
Tabela 15 – Resultado das simulações da posteriori do modelo GJR-GARCH(1,1)-NORMAL. . . . .	80
Tabela 16 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q do modelo GJR-GARCH(1,1)- NORMAL. . . . .	80
Tabela 17 – Resultado das simulações da posteriori do modelo GJR-GARCH(1,1)-t. . . . .	81
Tabela 18 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q do modelo GJR-GARCH(1,1)-t. . . . .	81
Tabela 19 – Resultado das simulações da posteriori do modelo GJR-GARCH(1,1)-GED. . . . .	82
Tabela 20 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q do modelo GJR-GARCH(1,1)- GED. . . . .	82
Tabela 21 – Critérios de seleção de modelos para o caso univariado. . . . .	82
Tabela 22 – Estatísticas básicas das séries diárias dos log-retornos (%). . . . .	88
Tabela 23 – Critérios de seleção de modelos para o caso multivariado. . . . .	91
Tabela 24 – Médias estimadas dos retornos. . . . .	93
Tabela 25 – Matriz de variâncias e covariâncias estimadas dos índices. . . . .	93

Tabela 26 – Comparação dos índices individuais com a carteira ótima. . . . .	93
Tabela 27 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-NORMAL.	113
Tabela 28 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-NORMAL.	114
Tabela 29 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)- NORMAL. . . . .	115
Tabela 30 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)- NORMAL. . . . .	116
Tabela 31 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-t. . . .	117
Tabela 32 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-t. . . .	118
Tabela 33 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)-t.	119
Tabela 34 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)-t.	120
Tabela 35 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-GED.	121
Tabela 36 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-GED.	122
Tabela 37 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)- GED. . . . .	123
Tabela 38 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)- GED. . . . .	124

# LISTA DE ABREVIATURAS E SIGLAS

---

---

ADCC-GARCH	<i>Asymmetric DCC-GARCH</i>
APARCH	<i>Asymmetric Power ARCH</i>
ARCH	<i>Autoregressive Conditional Heteroscedasticity</i>
B3	Brasil Bolsa Balcão
BSE SENSEX	<i>Bombay Stock Exchange Sensitive Index</i>
CAC 40	<i>Cotation Assistée en Continu 40</i>
CCC-GARCH	<i>Constant Conditional Correlation GARCH</i>
DAX	<i>Deutscher Aktienindex</i>
DCC-GARCH	<i>Dynamic Conditional Correlation GARCH</i>
DCC-GJR-GARCH	<i>Dynamic Conditional Correlation GJR-GARCH</i>
DIC	Critério de Informação de Desvio
EAIC	Esperança da posteriori do Critério de Informação de Akaike
EBIC	Esperança da posteriori do Critério de Informação Bayesiano
EGARCH	<i>Exponential GARCH</i>
ESS	<i>Effective Sample Size</i>
f.d.p.	Função Densidade de Probabilidade
FBTS	<i>Full Bayesian Significance Test</i>
FTSE 100	<i>Financial Times Stock Exchange 100</i>
GARCH	<i>Generalized Autoregressive Conditional Heteroscedasticity</i>
GED	<i>Generalized Error Distribution</i>
GJR-GARCH	<i>Glosten-Jagannathan-Runkle GARCH</i>
HMC	<i>Hamiltonian Monte Carlo</i>
HPD	<i>Highest Posterior Density</i>
IBOVESPA	Índice Bovespa
KL	<i>Kullback-Leibler</i>
LKJ	Lewandowski-Kurowicka-Joe
LOOIC	Critério de Informação <i>Leave-One-Out cross-validation</i>
MC	Monte Carlo
MCMC	<i>Markov Chain Monte Carlo</i>
NASDAQ	<i>National Association of Securities Dealers Automated Quotations</i>
NYSE	<i>New York Stock Exchange</i>

S&P500     *Standard & Poor's 500*  
TARCH     *Threshold ARCH*  
VAR        *Vector AutoRegression*  
VaR        *Value at Risk*  
VEC-GARCH *Vector Error Correction GARCH*  
WAIC       *Critério de Informação de Watanabe*  
ZV-MCMC   *estimador Zero-Variance do método MCMC*

# SUMÁRIO

---

---

<b>1</b>	<b>INTRODUÇÃO</b>	25
<b>2</b>	<b>INFERÊNCIA BAYESIANA</b>	31
2.1	Teorema de Bayes	31
2.2	Teoria da Decisão	32
2.3	Estimação	33
2.4	Região de Credibilidade	34
2.5	Teste de Hipótese	34
2.5.1	<i>Hipóteses Plenas</i>	35
2.5.2	<i>Hipóteses Precisas</i>	36
2.6	Computação Bayesiana	36
2.6.1	<i>Métodos MCMC</i>	37
2.7	<i>Método Zero-Variance</i>	39
2.8	<i>Método Hamiltonian Monte Carlo</i>	42
2.9	Critérios para Seleção de Modelos	44
<b>3</b>	<b>MODELO GARCH ASSIMÉTRICO</b>	47
3.1	GJR-GARCH	47
3.1.1	<i>Condições de Estacionariedade</i>	48
3.1.2	<i>Função de Verossimilhança</i>	50
3.1.3	<i>Distribuições das Inovações</i>	51
3.2	DCC-GJR-GARCH	52
3.2.1	<i>Função de Verossimilhança</i>	54
3.2.2	<i>Distribuições das Inovações</i>	54
3.3	Estimação Bayesiana	55
3.4	Análise de Pontos Influentes	58
<b>4</b>	<b>SIMULAÇÃO</b>	61
4.1	Caso Univariado	62
4.2	Caso Multivariado	67
4.3	Considerações Finais	69
<b>5</b>	<b>APLICAÇÃO</b>	75
5.1	Caso Univariado: Índice S&P500	75

<b>5.1.1</b>	<b><i>Estatísticas Básicas</i></b>	75
<b>5.1.2</b>	<b><i>Ajuste dos Modelos</i></b>	79
<b>5.2</b>	<b>Caso Multivariado: Índices das Principais Bolsas de Valores Mundiais</b>	82
<b>5.2.1</b>	<b><i>Estatísticas Básicas</i></b>	84
<b>5.2.2</b>	<b><i>Ajuste dos Modelos</i></b>	88
<b>5.2.3</b>	<b><i>Exemplo de Aplicação: Seleção de Carteira de Investimentos</i></b>	92
<b>6</b>	<b>CONCLUSÃO</b>	95
<b>6.1</b>	<b>Trabalhos Futuros</b>	97
	<b>REFERÊNCIAS</b>	99
<b>APÊNDICE A</b>	<b>DERIVADAS DO LOGARITMO DA FUNÇÃO DENSIDADE DAS DISTRIBUIÇÕES POSTERIORI</b>	103
<b>A.1</b>	<b>GJR-GARCH</b>	103
<b>A.2</b>	<b>DCC-GJR-GARCH</b>	105
<b>APÊNDICE B</b>	<b>SIMULAÇÕES DAS DISTRIBUIÇÕES POSTERIORI DO CASO UNIVARIADO</b>	109
<b>APÊNDICE C</b>	<b>SIMULAÇÕES DAS DISTRIBUIÇÕES POSTERIORI DO CASO MULTIVARIADO</b>	113



---

## INTRODUÇÃO

---

A análise de séries temporais tem sido objeto de estudos em diversas áreas, como na Engenharia, Física e Ciências da Terra. Entretanto, diante da crescente quantidade e complexidade de atividades financeiras nas últimas décadas, houve uma maior necessidade de compreender séries de retornos financeiros, com o intuito de fazer previsões e mensurar riscos. Nesse contexto, alguns modelos estatísticos foram desenvolvidos para este fim.

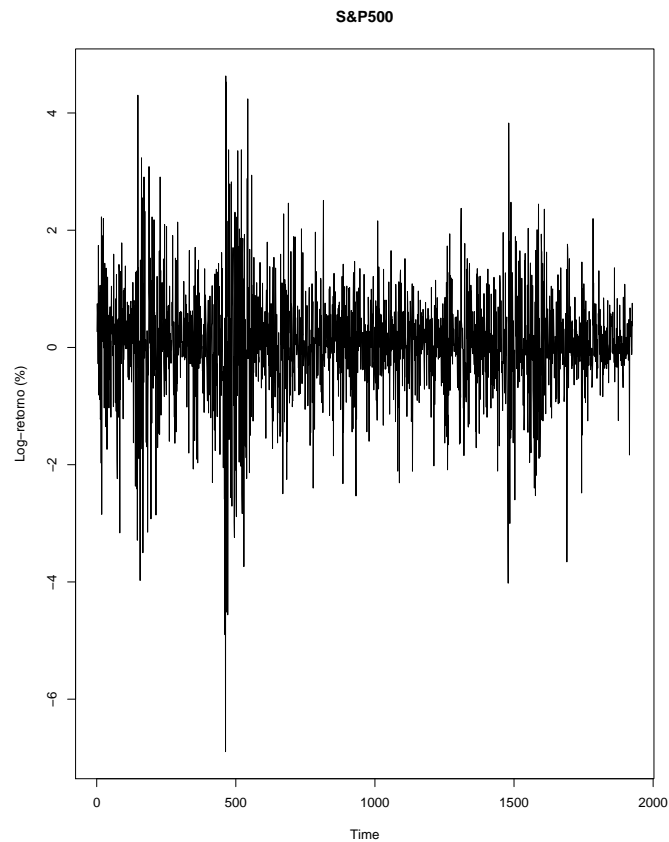
Algumas das características de retornos financeiros são que sua média não condicionada geralmente é constante e que existe uma correlação pequena entre retornos. Além disso, retornos financeiros apresentam um comportamento dinâmico ao longo do tempo. Isto é, existem períodos de tempo com grandes variações nos valores, bem como períodos com pequenas variações, o que aponta para mudanças impactantes na variância da série ao longo do tempo. Essa variação no comportamento é denominada **volatilidade**. Uma característica importante é que, por se tratar de uma variação instantânea, a volatilidade não é uma característica observável. Em termos estatísticos, a volatilidade pode ser definida como variância condicionada ao passado, portanto sendo uma característica essencial para a compreensão do comportamento da série.

Para ilustrar o comportamento da volatilidade em retornos financeiros, as Figuras 1 e 2 mostram, respectivamente, os log-retornos diários e sua função de autocorrelação do índice *Standard & Poor's 500* (S&P500) no período entre 06/10/2009 a 02/06/2017. Pode-se observar nas figuras a existência de períodos de grandes oscilações na série e uma pequena autocorrelação, que são características comuns em retornos financeiros, conforme mencionado anteriormente.

A fim de compreender a volatilidade de uma série, é necessária a utilização de modelos heterocedásticos condicionais, ou seja, modelos com variância condicional não constante. Dentre os diversos modelos para volatilidade propostos na literatura, os modelos da família *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH) (BOLLERSLEV, 1986) estão entre os mais utilizados, juntamente a modelos de volatilidade estocástica.

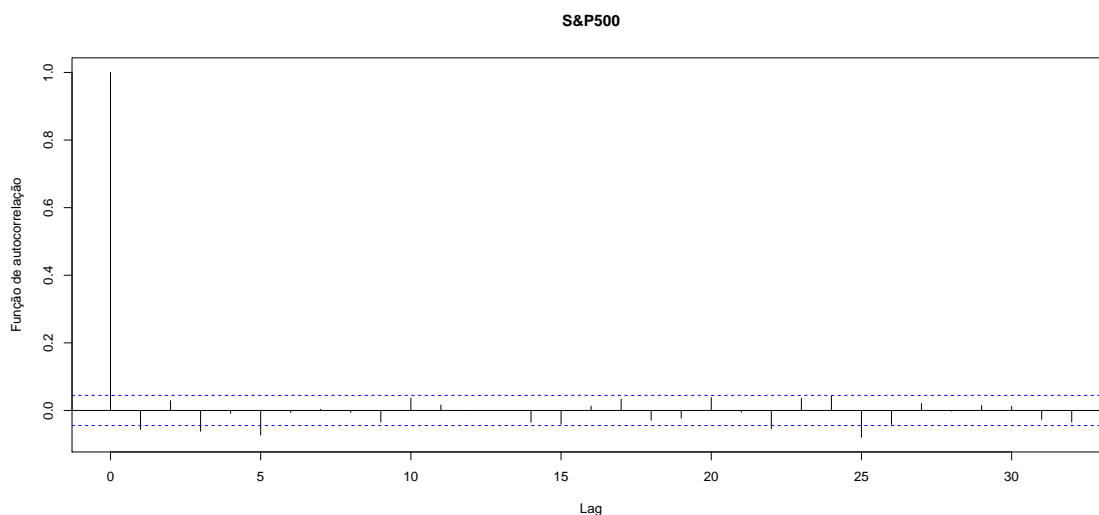
Inicialmente, foram propostos os modelos *Autoregressive Conditional Heteroscedasticity*

Figura 1 – Log-retornos diários da série S&amp;P500.



Fonte: Elaborada pelo autor.

Figura 2 – Função de autocorrelação dos log-retornos diários da série S&amp;P500.



Fonte: Elaborada pelo autor.

(ARCH) (ENGLE, 1982), cuja ideia básica é que a volatilidade depende linearmente do quadrado dos retornos passados. Em seguida, Bollerslev (1986) propôs uma classe de modelos em que

a volatilidade é uma combinação linear dos quadrados dos retornos passados e da volatilidade passada. Estes modelos são classificados como parte da família GARCH, que generaliza os modelos ARCH anteriormente propostos por Engle (1982). Os modelos ARCH e GARCH são considerados parcimoniosos, pois em geral descrevem a volatilidade com poucos parâmetros. Entretanto, em algumas situações, tais modelos não são apropriados para o ajuste, pois estudos empíricos mostram que a volatilidade de retornos financeiros tem um comportamento assimétrico em relação à volatilidade, o que não é o caso dos modelos ARCH e GARCH tradicionais, pois estes assumem que a volatilidade é simétrica.

No contexto de modelos assimétricos para retornos financeiros, diversos modelos GARCH assimétricos foram propostos na literatura. Um dos mais prominentes é o modelo *Exponential GARCH* (EGARCH) (NELSON, 1991), em que o logaritmo da volatilidade depende conjuntamente do passado da série e da volatilidade passada. Glosten, Jagannathan e Runkle (1993) propuseram que a volatilidade dependesse também do (de)crecimento dos valores anteriores da série. Isto é, valores negativos na série impactam diferentemente a volatilidade futura se comparados a valores positivos. Este modelo é denominado Glosten-Jagannathan-Runkle GARCH (GJR-GARCH). Outros exemplos de modelos assimétricos são o modelo *Threshold ARCH* (TARCH) (ZAKOIAN, 1994) e o modelo *Asymmetric Power ARCH* (APARCH) (DING; GRANGER; ENGLE, 1993).

Além da modelagem da volatilidade, a compreensão da correlação entre séries financeiras distintas é de grande importância, pois geralmente retornos financeiros possuem uma dependência linear entre eles. Nesse contexto, inúmeras extensões multivariadas dos modelos GARCH foram propostas na literatura. Além da área financeira, diversas aplicações têm sido beneficiadas por modelos multivariados GARCH, tais como a investigação de efeitos colaterais em estudos de contágio (TSE; TSUI, 2002; BAE; KAROLYI; STULZ, 2003). Uma revisão sobre modelos multivariados foi realizada por Bauwens, Laurent e Rombouts (2006).

No contexto de modelos multivariados para séries temporais, o primeiro modelo GARCH para matrizes de covariância condicional, conhecido como *Vector Error Correction GARCH* (VEC-GARCH) (BOLLERSLEV; ENGLE; WOOLDRIDGE, 1988), é uma generalização simples do modelo GARCH univariado. Em seguida, Engle, Ng e Rothschild (1990) introduziram a noção de fatores em modelos GARCH, com base na ideia de que a volatilidade é explicada por múltiplos fatores. Um dos modelos mais utilizados é o *Constant Conditional Correlation GARCH* (CCC-GARCH), proposto por Bollerslev (1990). Essencialmente, este modelo considera que a correlação de uma série multivariada é constante ao longo do tempo. Já o modelo *Dynamic Conditional Correlation GARCH* (DCC-GARCH) (ENGLE, 2002; TSE; TSUI, 2002) considera que a correlação depende do passado da série. Uma extensão multivariada do modelo assimétrico GJR-GARCH é o modelo *Asymmetric DCC-GARCH* (ADCC-GARCH) (CAPPIELLO; ENGLE; SHEPPARD, 2006), que também é uma extensão do modelo DCC-GARCH.

Em grande parte dos trabalhos envolvendo retornos financeiros, utiliza-se a distribuição

normal padrão como distribuição para as inovações de modelos da família GARCH, tanto no caso univariado como no multivariado. Entretanto, estudos empíricos mostram que, em geral, a distribuição das inovações financeiros possui uma cauda mais pesada, indicando que a distribuição das inovações deve pertencer a uma família de distribuições que possua essa característica. Alguns autores propuseram distribuições de cauda pesada para as inovações, tais como a distribuição  $t$  de Student e a *Generalized Error Distribution* (GED) (BAILLIE; BOLLERSLEV, 2002).

Sob o ponto de vista da inferência bayesiana, poucos trabalhos na literatura utilizam modelos da família GARCH, principalmente no caso de modelos multivariados. Um dos trabalhos na linha bayesiana aplica modelos DCC-GARCH com distribuição assimétrica para as inovações (FIORUCI; EHLERS; LOUZADA, 2014). Outro trabalho foca em modelos ADCC-GARCH com uma abordagem semiparamétrica e distribuição de inovações desconhecida (VIRBICKAITĖ; AUSÍN; GALEANO, 2016). Apesar de ambos os trabalhos utilizarem a inferência bayesiana, sua abordagem pode ser considerada mista, pois alguns parâmetros são estimados pela inferência clássica. A pouca utilização da inferência bayesiana nesse contexto da literatura ocorre devido à complexidade de modelos da família GARCH, bem como ao custo operacional de métodos computacionais baseados em simulação *Markov Chain Monte Carlo* (MCMC).

Uma das características das cadeias de *Markov* geradas por métodos MCMC, principalmente o método *Metropolis-Hastings* (HASTINGS, 1970), é que elas apresentam uma alta dependência. Conseqüentemente, torna-se necessário ter grandes cadeias para gerar bons estimadores bayesianos. Diante disso, foram propostos outros algoritmos MCMC que possibilitam gerar cadeias com baixa dependência, como o método *Hamiltonian Monte Carlo* (HMC) (NEAL, 2011), amplamente usado desde a criação da linguagem Stan (Stan Development Team, 2016). Entretanto, simular cadeias com baixa dependência exige um grande custo computacional. Por isso, Assaraf e Caffarel (1999) propuseram o método *Zero-Variance*, que consiste em encontrar um novo estimador MCMC com menor variância. Em outras palavras, este método é capaz de encontrar bons estimadores mesmo com cadeias de menor tamanho, o que leva a um menor custo computacional.

Neste contexto, este trabalho de doutorado foi realizado com base nos seguintes objetivos:

1. Estudar e aplicar modelos da família GARCH assimétricos, com foco nos modelos GJR-GARCH e DCC-GJR-GARCH;
2. Analisar o impacto da utilização de inovações seguindo as distribuições **normal** e **t**, bem como propor o uso da distribuição **GED** para o caso multivariado dos modelos estudados;
3. Realizar a estimação dos parâmetros de forma puramente bayesiana, o que até então não havia sido abordado pela literatura;

4. Propor uma metodologia para estimação **eficaz e eficiente**: **(i)** eficaz em termos de possibilitar a geração de cadeias com baixa dependência; e **(ii)** eficiente em termos de desempenho computacional.

O restante deste documento está organizado conforme a seguir. O Capítulo [2](#) descreve conceitos importantes da inferência bayesiana, com foco na teoria da decisão e nos métodos computacionais HMC e *Zero-Variance*. O Capítulo [3](#) apresenta os modelos GJR-GARCH e DCC-GJR-GARCH, juntamente à análise de influência de cada modelo. Os Capítulos [4](#) e [5](#) relatam, respectivamente, o estudo de simulação e exemplos de aplicação de cada modelo estudado. Por fim, o Capítulo [6](#) apresenta as considerações finais e propostas de trabalhos futuros.



---

# INFERÊNCIA BAYESIANA

---

A inferência estatística tem como objetivo estudar as características de uma população através da informação sobre uma quantidade de interesse, representada por um parâmetro, cujo valor é desconhecido. A tarefa de inferência pode ser realizada através de uma dentre duas possíveis abordagens: bayesiana e clássica (esta última também chamada de frequentista) (DEGROOT *et al.*, 1986). A inferência bayesiana descreve a incerteza sobre o parâmetro através de modelos probabilísticos. Especialmente, assume-se que a probabilidade é subjetiva, ou seja, cada pesquisador definirá um modelo probabilístico que melhor descreve o parâmetro em questão conforme sua visão de mundo. Este é um paradigma diferente daquele utilizado na inferência clássica, em que assume-se que os parâmetros são constantes e a probabilidade é frequentista.

O restante deste capítulo apresenta a ideia geral sobre inferência bayesiana e descreve duas técnicas computacionais de estimação bayesiana: os métodos *Zero-Variance* e *Hamiltonian Monte Carlo* (HMC).

## 2.1 Teorema de Bayes

Seja  $\theta$  um vetor de parâmetros de interesse cujos valores são desconhecidos;  $\pi(\theta)$  a distribuição de probabilidade que resume toda informação prévia sobre  $\theta$ , que é definida como **distribuição a priori**; e  $\pi(\mathbf{X} | \theta)$  é a verossimilhança, que representa a população de interesse e o desempenho em relação a  $\theta$ . O objetivo da inferência bayesiana é atualizar a informação sobre  $\theta$  após observar  $\mathbf{X}$ . Isso ocorre segundo o teorema de Bayes (DEGROOT *et al.*, 1986), definido a seguir:

$$\begin{aligned}
\pi(\theta | \mathbf{X}) &= \frac{\pi(\mathbf{X} | \theta)\pi(\theta)}{\pi(\mathbf{X})} \\
&= \frac{\pi(\mathbf{X} | \theta)\pi(\theta)}{\int \pi(\mathbf{X} | \theta)\pi(\theta)d\theta} \\
&\propto \pi(\mathbf{X} | \theta)\pi(\theta),
\end{aligned} \tag{2.1}$$

onde  $\pi(\theta | \mathbf{X})$  é denominada **distribuição posteriori**.

A distribuição posteriori representa toda a informação atualizada sobre o parâmetro por meio de um modelo probabilístico. Portanto, toda inferência feita sobre  $\theta$  deve, obrigatoriamente, utilizar essa distribuição. Entretanto, às vezes é preciso resumir a informação através de **(i)** um valor numérico (estimação pontual); **(ii)** uma região (intervalo ou região de credibilidade); ou **(iii)** uma decisão referente ao parâmetro (teste de hipótese). Essencialmente, todo problema de inferência bayesiana pode ser encarado como um dos três problemas de decisão recém-mencionados.

A próxima seção descreve um ramo da teoria estatística denominado *teoria da decisão*, com enfoque nas soluções de cada problema de decisão citado anteriormente.

## 2.2 Teoria da Decisão

O objetivo da teoria da decisão (DEGROOT, 2005) é indicar qual é a melhor ação dentre as disponíveis segundo algum critério. Para formular um problema de decisão, deve-se especificar cada um de seus elementos, que são:

- $\mathbf{X} \in \chi$ : uma quantidade desconhecida que expressa os dados a serem observados, onde  $\chi$  é o domínio de todos os possíveis dados;
- $\mathcal{A}$ : o conjunto das alternativas (ações) disponíveis;
- $\Theta$ : o conjunto de possíveis ocorrências sobre os dados;
- $\mathcal{P}$ : a medida de probabilidade sobre  $\Theta$ ;
- $l$ : uma função  $\mathcal{A} \times \Theta \rightarrow \mathbb{R}^+$  definida como função perda.

A função perda  $l(a, \theta)$  é interpretada como a perda sofrida ao tomar a decisão  $a \in \mathcal{A}$  dado que o valor do parâmetro é  $\theta \in \Theta$ . Com base nos dados  $\mathbf{X}$ , pode-se tomar uma decisão utilizando um procedimento ou regra de decisão definida como:

**Definição 1** (Regra de Decisão). O procedimento ou regra de decisão é uma função  $\delta : \chi \rightarrow \mathcal{A}$ . A regra  $\delta$  significa que, se a amostra  $\mathbf{X}$  for observada, toma-se a decisão  $\delta(\mathbf{X}) \in \mathcal{A}$ .



Observe-se que existe um conjunto inumerável de regras de decisão. Entretanto, não é necessário o conhecimento de todas, pois o interesse é encontrar a regra de decisão ótima  $\delta^*$ , que é definida conforme a seguir:

$$\delta^* = \operatorname{argmin}_{a \in \mathcal{A}} \mathbb{E}[l(a, \theta) | \mathbf{X}]. \quad (2.2)$$

Em outras palavras, a melhor regra de decisão é aquela que, para cada dado observado, escolhe a ação com menor perda esperada (referente à distribuição posteriori). Com os procedimentos apresentados nas seções a seguir, pode-se encontrar o estimador de Bayes (Seção 2.3) — o nome é uma homenagem a Thomas Bayes, conhecido por ter formulado o caso especial do teorema de Bayes —, a região de credibilidade (Seção 2.4) ou tomar uma decisão em um teste de hipótese (Seção 2.5).

## 2.3 Estimação

O problema de estimação consiste em escolher um valor aproximado do parâmetro do modelo a partir dos dados. Este trabalho de doutorado utiliza a função perda quadrática, que é dada por

$$l(\mathbf{a}, \theta) = (\theta - \mathbf{a})^T (\theta - \mathbf{a}), \quad (2.3)$$

onde  $\mathbf{a}, \theta \in \mathbb{R}^d$ . Esta função é uma das mais utilizadas na teoria estatística e está associada à técnica dos mínimos quadrados e ao erro quadrático médio. Nesse contexto, é possível encontrar o melhor estimador conforme a seguir.

**Teorema 1.** Considere-se um estimador arbitrário denominado  $\hat{\theta}$ . Se  $l(\mathbf{a}, \theta) = (\theta - \mathbf{a})^T (\theta - \mathbf{a})$ , então o melhor estimador  $\hat{\theta}^*$  é tal que

$$\hat{\theta}^* = \mathbb{E}[\theta | \mathbf{X}]. \quad (2.4)$$

*Demonstração.*

$$\begin{aligned} \hat{\theta}^* &= \operatorname{arg min}_{a \in \mathcal{A}} \mathbb{E}[l(\mathbf{a}, \theta) | \mathbf{X}] \\ &= \operatorname{arg min}_{a \in \mathcal{A}} \mathbb{E}[(\theta - \mathbf{a})^T (\theta - \mathbf{a}) | \mathbf{X}] \\ &= \operatorname{arg min}_{a \in \mathcal{A}} (\mathbb{E}[\theta^T \theta | \mathbf{X}] - \mathbb{E}[\theta^T \mathbf{a} | \mathbf{X}] - \mathbb{E}[\mathbf{a}^T \theta | \mathbf{X}] + \mathbb{E}[\mathbf{a}^T \mathbf{a} | \mathbf{X}]) \\ &= \operatorname{arg min}_{a \in \mathcal{A}} (\mathbb{E}[\theta^T \theta | \mathbf{X}] - \mathbb{E}[\theta | \mathbf{X}]^T \mathbb{E}[\theta | \mathbf{X}] + \mathbb{E}[\theta | \mathbf{X}]^T \mathbb{E}[\theta | \mathbf{X}] \\ &\quad - \mathbb{E}[\theta | \mathbf{X}]^T \mathbf{a} - \mathbf{a}^T \mathbb{E}[\theta | \mathbf{X}] + \mathbf{a}^T \mathbf{a}) \\ &= \operatorname{arg min}_{a \in \mathcal{A}} (\operatorname{Var}[\theta | \mathbf{X}] + (\mathbf{a} - \mathbb{E}[\theta | \mathbf{X}])^T (\mathbf{a} - \mathbb{E}[\theta | \mathbf{X}])) \\ &= \mathbb{E}[\theta | \mathbf{X}]. \end{aligned} \quad (2.5)$$

□

Assim, utilizando-se a perda quadrática, o melhor estimador para  $\theta$  é o vetor de médias da posteriori.

## 2.4 Região de Credibilidade

Diferentemente do problema de estimação, em que se está interessado em escolher um número aproximado do parâmetro, o problema da região de credibilidade consiste em encontrar uma região paramétrica na qual o parâmetro provavelmente está contido. Porém, busca-se a menor região possível para que haja uma maior precisão com respeito ao valor do parâmetro que é a região de credibilidade de maior probabilidade à posteriori.

Matematicamente, seja  $\mathcal{A}$  é a família de subconjuntos de espaços paramétricos ( $\mathcal{A} \subset \Theta$ ). Então, utiliza-se a seguinte função perda:

$$l(a, \theta) = k \int_{\mathbf{X} \in \mathcal{A}} \partial \mathbf{X} - I(\theta \in a), \quad (2.6)$$

onde  $I$  é a função indicadora. A interpretação da função perda na Equação 2.6 é que a perda diminui quando o parâmetro está dentro do intervalo, bem como aumenta à medida que a região de credibilidade aumenta. Formalmente, tem-se o seguinte teorema.

**Teorema 2.** Se  $l(\delta(\mathbf{X}), \theta) = k \int_{\mathbf{X} \in \delta(\mathbf{X})} \partial \mathbf{X} - I(\theta \in \delta(\mathbf{X}))$ , então a melhor região de credibilidade é dada por

$$\delta^*(\mathbf{X}) = \{\theta \in \Theta : \pi(\theta | \mathbf{X}) \geq k\}. \quad (2.7)$$

A região  $\delta^*(\mathbf{X})$  é denominada *Highest Posterior Density* (HPD) de  $\pi(\theta | \mathbf{X})$ .

## 2.5 Teste de Hipótese

O problema de teste de hipótese consiste em selecionar uma dentre duas proposições disjuntas e mutuamente exclusivas no espaço paramétrico, conhecidas como **hipótese nula** e **hipótese alternativa** e denotadas por:

$$\begin{cases} H_0 : \theta \in \Theta_0 \\ H_1 : \theta \in \Theta_1 \end{cases}, \quad (2.8)$$

onde  $\Theta_0$  e  $\Theta_1$  são, respectivamente, o espaço paramétrico referente à hipótese nula e o espaço paramétrico referente à hipótese alternativa. Tem-se também que  $\Theta_0 \cup \Theta_1 = \Theta$  e  $\Theta_0 \cap \Theta_1 = \emptyset$ . As hipóteses podem ser classificadas conforme a seguir.

- **Hipóteses plenas.**  $\Theta_i, i \in \{0, 1\}$ , possui a mesma dimensão do espaço paramétrico. Exemplo: quando  $\Theta = \mathbb{R}$ ,  $H_0 : \theta \in \mathbb{R}^+$  e  $H_1 : \theta \in \mathbb{R}^-$ .
- **Hipóteses precisas.** Um dos  $\Theta_i, i \in \{0, 1\}$ , possui dimensão menor do que a dimensão do espaço paramétrico. Exemplo: quando  $\Theta = \mathbb{R}$ ,  $H_0 : \theta = 0$  e  $H_1 : \theta \in \mathbb{R}^*$ .

Em um teste de hipótese, é possível cometer dois tipos de erros: **(i)** Tipo I – rejeição de  $H_0$  quando esta é verdadeira; e **(ii)** Tipo II – aceitação de  $H_0$  quando esta é falsa. Portanto, a função de perda deve compreender ambos os tipos de erros.

### 2.5.1 Hipóteses Plenas

Quando as duas hipóteses são plenas, utiliza-se neste trabalho a função de perda (0-1-c) apresentada na Tabela I. A função interpreta que não há perda ao selecionar-se a hipótese correta, enquanto o valor  $c$  representa o quão mais grave é o erro Tipo I em relação ao erro Tipo II. Se  $c > 1$ , o erro Tipo I é mais grave do que o erro Tipo II; se  $c < 1$ , o erro Tipo I é menos grave do que o erro Tipo II; e se  $c = 1$ , ambos os erros são igualmente graves. Com essa função perda, é possível encontrar a regra ótima.

Tabela 1 – Descrição dos valores da função perda.

	$\theta \in \Theta_0$	$\theta \in \Theta_1$
Aceita $H_0$	0	1
Rejeita $H_0$	$c$	0

Fonte: Elaborada pelo autor.

**Teorema 3.** No problema do teste de hipótese com a função perda (0-1-c), a regra de decisão  $\delta^*(\mathbf{X})$  ótima é dada por:

$$\delta^*(\mathbf{X}) = \begin{cases} 0, & \text{se } \mathbb{P}(\theta \in \Theta_0 | \mathbf{X}) > (1+c)^{-1} \\ 1, & \text{se } \mathbb{P}(\theta \in \Theta_1 | \mathbf{X}) < (1+c)^{-1} \end{cases}, \quad (2.9)$$

onde  $\delta^*(\mathbf{X}) = 0$  quando se aceita  $H_0$  e  $\delta^*(x) = 0$  quando se aceita  $H_1$ .

*Demonstração.* Nesse caso, existem duas ações: (i) não rejeitar  $H_0$  e (ii) rejeitar  $H_0$ . Aplicando-se a Equação 2.2, tem-se que

$$\delta^*(\mathbf{X}) = \begin{cases} 0, & \text{se } \mathbb{E}[l(0, \theta) | \mathbf{X}] < \mathbb{E}[l(1, \theta) | \mathbf{X}] \\ 1, & \text{se } \mathbb{E}[l(0, \theta) | \mathbf{X}] > \mathbb{E}[l(1, \theta) | \mathbf{X}] \end{cases}.$$

Então,

$$\begin{aligned} \mathbb{E}[l(0, \theta) | \mathbf{X}] &= \mathbb{E}[0 I(\theta \in \Theta_0) | \mathbf{X}] + \mathbb{E}[I(\theta \in \Theta_1) | \mathbf{X}] \\ &= \mathbb{P}[\theta \in \Theta_1 | \mathbf{X}] \end{aligned}$$

e

$$\begin{aligned} \mathbb{E}[l(1, \theta) | \mathbf{X}] &= \mathbb{E}[c I(\theta \in \Theta_0) | \mathbf{X}] + \mathbb{E}[0 I(\theta \in \Theta_1) | \mathbf{X}] \\ &= c\mathbb{P}[\theta \in \Theta_0 | \mathbf{X}]. \end{aligned}$$

Então,  $\delta^*(\mathbf{X}) = 0$  se

$$\begin{aligned} \mathbb{E}[l(0, \theta) | \mathbf{X}] &< \mathbb{E}[l(1, \theta) | \mathbf{X}], \\ \mathbb{P}[\theta \in \Theta_1 | \mathbf{X}] &< c\mathbb{P}[\theta \in \Theta_0 | \mathbf{X}], \\ 1 - \mathbb{P}[\theta \in \Theta_0 | \mathbf{X}] &< c\mathbb{P}[\theta \in \Theta_0 | \mathbf{X}] \text{ e} \\ \mathbb{P}[\theta \in \Theta_0 | \mathbf{X}] &> (1+c)^{-1}. \end{aligned}$$

□

Similarmente,  $\delta^*(\mathbf{X}) = 1$  se  $\mathbb{P}[\theta \in \Theta_0 | \mathbf{X}] < (1+c)^{-1}$ . Em resumo, o Teorema 3 indica que se aceita  $H_0$  quando a probabilidade da distribuição posteriori é suficientemente grande, bem como se aceita  $H_0$  se a probabilidade da posteriori for maior do que  $\frac{1}{2}$  quando  $c = 1$ .

## 2.5.2 Hipóteses Precisas

Existe um problema inerente a hipóteses precisas em que, quando a distribuição posteriori é contínua, a probabilidade de  $\theta \in \Theta_0$  é nula ( $\mathbb{P}[\theta \in \Theta_0 | \mathbf{X} = \mathbf{X}] = 0$ ). Portanto, o uso da função perda 0-1-c sempre leva à rejeição de  $H_0$ , o que indica que essa função não é adequada para hipóteses precisas.

Madruga, Esteves e Wechsler (2001) mostra que o teste, apresentado a seguir, é um procedimento bayesiano, em que a regra de decisão é obtida através de um HPD, conforme apresenta o seguinte teorema.

**Teorema 4** (*Full Bayesian Significance Test* (FBST)). Existe uma função utilidade (MADRUGA; ESTEVES; WECHSLER, 2001) tal que o melhor teste de hipótese possui a seguinte regra:

1. Constrói-se um HPD para  $\theta | \mathbf{X}$  com credibilidade  $1 - \alpha$ .
2. Rejeita-se  $H_0$  se nenhum ponto de  $H_0$  estiver no HPD.

Outra maneira de usar o FBST é pelo cálculo da evidência (e-valor ou  $EV$ ) dado por

$$\text{e-valor} = \mathbb{P}[\{\theta \in \Theta : \pi(\theta \geq k | \mathbf{X} = \mathbf{X})\}], \quad (2.10)$$

onde  $k = \max_{\theta \in \Theta_0} \pi(\theta \geq k | \mathbf{X})$ . Portanto, não se rejeita  $H_0$  quando  $EV > \alpha$ . Observe-se que  $\alpha$  é determinado pelo pesquisador (subjetivo) e, além disso, quanto menor o e-valor, maior a evidência de que a hipótese nula é falsa. Em outras palavras, o e-valor é uma medida de evidência diferente do valor p. Uma das vantagens do teste de hipótese bayesiano em relação ao clássico é a flexibilização controlada em ambos os tipos de erro (Tipo I e Tipo II), estando a critério do pesquisador, enquanto o caso clássico só permite tal controle no erro Tipo I.

## 2.6 Computação Bayesiana

Conforme discutido anteriormente, na inferência bayesiana, toda informação atualizada (após considerar os dados obtidos) é representada pela distribuição posteriori, obtida pelo teorema de Bayes:

$$\pi(\theta | \mathbf{X}) = \frac{\pi(\mathbf{X} | \theta)\pi(\theta)}{\int \pi(\mathbf{X} | \theta)\pi(\theta) d\theta}$$

Porém, em alguns casos, como nos modelos estudados nesta tese (Capítulo 3), não é possível encontrar diretamente  $\int \pi(x | \theta)\pi(\theta)\partial\theta$ , nem mesmo por aproximação via métodos determinísticos. Mais especificamente, suponha-se uma situação em que se deseja estimar um parâmetro (ou uma função dele). Sabe-se pela Seção 2.3 que o melhor estimador bayesiano é a média da posteriori, que é dada por:

$$\mathbb{E}[g(\theta) | \mathbf{X}] = \int g(\theta)\pi(\theta | \mathbf{X})\partial\theta. \quad (2.11)$$

Devido à falta de elementos para identificar totalmente a distribuição posteriori, não é possível encontrar diretamente o estimador bayesiano, uma vez que a integral da Equação 2.11 não pode ser resolvida. Diante disso, a abordagem comumente empregada consiste em utilizar métodos computacionais, tais como o Monte Carlo (MC), que são baseados em simulações que geram a distribuição posteriori. Assim, torna-se possível realizar a análise bayesiana substituindo-se a Equação 2.11 por:

$$\hat{\mathbb{E}}[g(\theta) | \mathbf{X}] = \sum_{i=1}^N \frac{g(\theta_i^*)}{N}, \quad (2.12)$$

onde  $\theta_i^*$  representa os dados gerados pelo método computacional em questão e  $N$  é o número total de elementos gerados. Note-se que essa abordagem pode ser similarmente aplicada tanto para testes de hipótese quanto para região de credibilidade. O método computacional ideal seria o método Monte Carlo. Entretanto, esse método exige o conhecimento da distribuição, que não é o caso no contexto desta tese. Diante disso, optou-se por aplicar um método **Monte Carlo via Cadeias de Markov** (*Markov Chain Monte Carlo* (MCMC)). Assim, a próxima seção apresenta os estimadores que utilizam o método MCMC.

### 2.6.1 Métodos MCMC

A ideia geral dos métodos MCMC é gerar uma amostra da distribuição posteriori e calcular as estimativas amostrais dos parâmetros através da Equação 2.12, assim como ocorre no método Monte Carlo original. Entretanto, diferentemente do método Monte Carlo, as técnicas de simulação são baseadas em cadeias de *Markov* e, portanto, os dados gerados não são independentes.

Os métodos MCMC requerem que a cadeia de *Markov* gerada (isto é, os dados simulados) possua uma única distribuição estacionária, que é a distribuição posteriori. Para isso, é necessário que a cadeia gerada seja homogênea, irredutível e aperiódica (para mais detalhes, ver [Robert e Casella \(2005\)](#)). Geralmente, a cadeia gerada é influenciada pelos valores das iterações iniciais e, à medida que mais iterações ocorrem, a influência inicial passa a diminuir. Assim, na prática, as iterações iniciais são descartadas como se tivessem sido uma amostra de aquecimento (comumente conhecida como *burn-in*). Estimadores MCMC são definidos conforme a seguir.

**Definição 2.** Sejam  $\theta_1^*, \theta_2^*, \dots, \theta_N^*$  gerados por uma cadeia de *Markov* ergótica com distribuição estacionária  $\pi(\theta | \mathbf{X})$  e seja  $g$  uma função. Então, o estimador MCMC de  $\mathbb{E}[g(\theta) | \mathbf{X}]$  é dado

com base na Equação 2.12:

$$\hat{\mu}_{g(\theta)} = \sum_{i=1}^N \frac{g(\theta_i^*)}{N}.$$

O estimador  $\hat{\mu}_{g(\theta)}$  é não-viesado para  $\mathbb{E}[g(\theta) | \mathbf{X}]$ , embora geralmente seja viesado para  $g(\theta)$ , e possui variância

$$\text{Var}(\hat{\mu}_{g(\theta)}) = \sum_{i=1}^N \frac{\sigma_g^2 \tau_\pi}{N}, \quad (2.13)$$

onde  $\sigma_g^2 = \text{Var}(g(\theta))$ ,  $\tau_\pi = 1 + 2 \sum_{k=1}^{\infty} \rho_\pi(k)$  e  $\rho_\pi(\cdot)$  é a função de autocorrelação da cadeia. Observe-se que a variância do estimador MCMC aumenta quando existe uma autocorrelação positiva na cadeia gerada. Portanto, a escolha do método é fundamental, pois se um método tende a gerar cadeias com uma grande autocorrelação, é necessário ter uma grande quantidade de elementos na cadeia para amortizar esse problema. Entretanto, manipular uma cadeia maior implica um maior custo computacional. Além disso, métodos MCMC que geram cadeias inerentemente com baixa autocorrelação também têm um elevado custo computacional, mesmo com cadeias curtas.

Diante desses problemas, surgiu a necessidade de ter um novo estimador não-viesado para  $\mathbb{E}[g(\theta) | \mathbf{X}]$ , definido conforme a seguir. Seja uma nova função  $\tilde{g}(\theta)$  denotada por

$$\tilde{g}(\theta) = g(\theta) + \mathbf{a}^T \mathbf{w}(\theta), \quad (2.14)$$

onde  $\mathbf{a} \in \mathbb{R}^{n_a}$ ,  $\mathbf{w}$  é uma função de  $\mathbb{R}^d \rightarrow \mathbb{R}^{n_a}$  e  $\mathbb{E}(\mathbf{w}(\theta)) = \mathbf{0}$ . O termo  $\mathbf{w}(\theta)$  consiste em variáveis de controle. Quanto mais variáveis de controle existirem, melhor será o estimador. Portanto, tem-se o estimador  $\hat{\mu}_{\tilde{g}(\theta)} = \sum_{i=1}^N \frac{\tilde{g}(\theta_i^*)}{N}$  para  $\mathbb{E}[g(\theta) | \mathbf{X}]$  com variância

$$\sigma_{\tilde{g}}^2 = \sigma_g^2 + \mathbf{a}^T \mathbf{a} \sigma_w^2 + \mathbf{a}^T \sigma_{g,\mathbf{w}}, \quad (2.15)$$

onde  $\sigma_{\tilde{g}}^2 = \text{Var}[\tilde{g}(\theta)]$ ,  $\sigma_g^2 = \text{Var}[g(\theta)]$  e  $\sigma_{g,\mathbf{w}} = \text{Cov}[g(\theta), \mathbf{w}(\theta)]$ . Idealmente, o estimador deve ter a mínima variância possível. De modo a minimizar a variância definida pela Equação 2.15, tem-se a seguinte proposição.

**Proposição 1.** Seja  $\tilde{g}(\theta)$  dado pela Equação 2.14. O valor ótimo de  $\mathbf{a}^T$  que minimiza a variância de  $\tilde{g}(\theta)$  é definido por

$$\mathbf{a} = -\frac{\sigma_{g,\mathbf{w}}}{\sigma_g^2}. \quad (2.16)$$

Portanto, o novo estimador não-viesado de  $\mathbb{E}[g(\theta) | \mathbf{X}]$ , denominado  $\hat{\mu}_{\tilde{g}(\theta)}$ , possui variância

$$\text{Var}(\hat{\mu}_{\tilde{g}(\theta)}) = \sum_{i=1}^N \frac{\sigma_{\tilde{g}}^2 \tau_\pi}{N} \leq \sum_{i=1}^N \frac{\sigma_g^2 \tau_\pi}{N} = \text{Var}(\hat{\mu}_{g(\theta)}), \quad (2.17)$$

o que indica que o estimador  $\hat{\mu}_{\tilde{g}(\theta)}$  é mais eficiente do que  $\hat{\mu}_{g(\theta)}$ . Portanto,  $\hat{\mu}_{\tilde{g}(\theta)}$  é o estimador utilizado nesta tese. Conforme mencionado no Capítulo 1, um dos objetivos desta tese é indicar

um estimador eficiente (isto é, com variância mínima) que demande o menor custo computacional possível. Para diminuir essa variância, é necessário:

1. propor um método MCMC que produza uma cadeia de Markov com pouca autodependência, porém com eficiência computacional; e
2. encontrar uma função  $\mathbf{w}$  com  $\mathbb{E}[\mathbf{w}(\boldsymbol{\theta})] = \mathbf{0}$  que reduza a variância do estimador apresentado em relação à  $\boldsymbol{\theta}$ .

Com relação ao item 1, o método MCMC selecionado foi o *Hamiltonian Monte Carlo* (HMC) (NEAL, 2011). A principal vantagem do método HMC é a capacidade de gerar cadeias com pouca autodependência, principalmente se comparado ao método mais comumente usado, chamado *Metropolis-Hastings* (HASTINGS, 1970). Com respeito ao item 2, para a escolha da função  $\mathbf{w}$  e do estimador correspondente, empregou-se o método *Zero-Variance* introduzido por Assaraf e Caffarel (1999). As duas próximas seções apresentam, respectivamente, o método *Zero-Variance* e o método HMC.

## 2.7 Método Zero-Variance

O método *Zero-Variance* consiste em substituir a função original  $g(\boldsymbol{\theta})$  de parâmetros de interesse por uma função auxiliar  $\tilde{g}(\boldsymbol{\theta})$  que possui a mesma média, mas com variância menor. Assaraf e Caffarel (1999) propõem o princípio de *Zero-Variance* com base nas variáveis de controle  $\mathbf{w}(\boldsymbol{\theta})$  e definem a função  $\tilde{g}(\boldsymbol{\theta})$  como

$$\tilde{g}(\boldsymbol{\theta}) = g(\boldsymbol{\theta}) + \frac{H[\boldsymbol{\psi}(\boldsymbol{\theta})]}{\sqrt{\pi(\boldsymbol{\theta} | X)}}, \quad (2.18)$$

sendo  $H(\boldsymbol{\psi}(\boldsymbol{\theta}))$  um hamiltoniano Schrödinger que é dado por

$$H[\boldsymbol{\psi}(\boldsymbol{\theta})] = -\frac{1}{2} \Delta_{\boldsymbol{\theta}} [\boldsymbol{\psi}(\boldsymbol{\theta})] + \frac{\boldsymbol{\psi}(\boldsymbol{\theta})}{2\sqrt{\pi(\boldsymbol{\theta} | X)}} \Delta_{\boldsymbol{\theta}} [\sqrt{\pi(\boldsymbol{\theta} | X)}], \quad (2.19)$$

onde  $\Delta_{\boldsymbol{\theta}} := \sum_{i=1}^d \frac{\partial}{\partial \theta_i}$  é o operador de Laplace e  $\boldsymbol{\psi}(\boldsymbol{\theta})$  é denominado função experimental (*trial function*). A condição

$$\int H[\boldsymbol{\psi}(\boldsymbol{\theta})] \sqrt{\pi(\boldsymbol{\theta} | X)} \partial \boldsymbol{\theta} = 0 \quad (2.20)$$

é satisfeita para qualquer função integrável  $\boldsymbol{\psi}(\boldsymbol{\theta})$ . Assim, pela Equação 2.14, tem-se que  $\hat{\mu}_{\tilde{g}(\boldsymbol{\theta})}$  é um estimador não-viesado assintoticamente de  $\mu_{g(\boldsymbol{\theta})}$ , sendo  $\mathbf{a}^T \mathbf{w}(\boldsymbol{\theta}) = \frac{H[\boldsymbol{\psi}(\boldsymbol{\theta})]}{\sqrt{\pi(\boldsymbol{\theta} | X)}}$ .

Mira, Solgi e Imparato (2013) propuseram um conjunto de funções experimentais definidas como

$$\boldsymbol{\psi}(\boldsymbol{\theta}) = P(\boldsymbol{\theta}) \sqrt{\pi(\boldsymbol{\theta} | X)}, \quad (2.21)$$

onde  $P(\theta)$  é um polinômio arbitrário. Então, desenvolvendo-se  $\frac{H[\psi(\theta)]}{\sqrt{\pi(\theta|X)}}$ , tem-se que

$$\begin{aligned}
\frac{H[\psi(\theta)]}{\sqrt{\pi(\theta|X)}} &= -\frac{1}{2\sqrt{\pi(\theta|X)}} \left( \Delta_{\theta}[\psi(\theta)] - \frac{\psi(\theta)}{\sqrt{\pi(\theta|X)}} \Delta_{\theta}[\sqrt{\pi(\theta|X)}] \right) \\
&= -\frac{1}{2\sqrt{\pi(\theta|X)}} \left( \Delta_{\theta}[P(\theta)\sqrt{\pi(\theta|X)}] - P(\theta) \Delta_{\theta}[\sqrt{\pi(\theta|X)}] \right) \\
&= -\frac{1}{2\sqrt{\pi(\theta|X)}} \left( \sum_{i=1}^d \left( \frac{\partial P(\theta)}{\partial^2 \theta_i} \sqrt{\pi(\theta|X)} + 2 \frac{\partial P(\theta)}{\partial \theta_i} \frac{\partial \sqrt{\pi(\theta|X)}}{\partial \theta_i} + P(\theta) \frac{\partial \sqrt{\pi(\theta|X)}}{\partial^2 \theta_i} \right) \right. \\
&\quad \left. - \sum_{i=1}^d P(\theta) \frac{\partial \sqrt{\pi(\theta|X)}}{\partial^2 \theta_i} \right) \\
&= -\frac{1}{2\sqrt{\pi(\theta|X)}} \sum_{i=1}^d \left( \frac{\partial P(\theta)}{\partial^2 \theta_i} \sqrt{\pi(\theta|X)} + 2 \frac{\partial P(\theta)}{\partial \theta_i} \frac{\partial \sqrt{\pi(\theta|X)}}{\partial \theta_i} \right) \\
&= -\frac{1}{2} \sum_{i=1}^d \frac{\partial P(\theta)}{\partial^2 \theta_i} - \sum_{i=1}^d \frac{\partial P(\theta)}{\partial \theta_i} \frac{1}{\sqrt{\pi(\theta|X)}} \frac{\partial \sqrt{\pi(\theta|X)}}{\partial \theta_i} \\
&= -\frac{1}{2} \Delta_{\theta}[P(\theta)] + \sum_{i=1}^d \frac{\partial P(\theta)}{\partial \theta_i} - \frac{1}{2} \frac{\partial \log \pi(\theta|X)}{\partial \theta_i}.
\end{aligned}$$

Portanto, a função auxiliar é

$$\tilde{g}(\theta) = g(\theta) - \frac{1}{2} \Delta_{\theta}[P(\theta)] + \nabla_{\theta}[P(\theta)]^T \mathbf{z}(\theta), \quad (2.22)$$

onde  $\mathbf{z}(\theta) := \nabla_{\theta}[\log(\pi(\theta|X))]$ ,  $\nabla_{\theta} = (\frac{\partial}{\partial \theta_1}, \frac{\partial}{\partial \theta_2}, \dots, \frac{\partial}{\partial \theta_{n_{\theta}}})^T$  é a função gradiente e  $n_{\theta}$  é o total de parâmetros.

Assim, com base na Equação [2.22](#), a fim de encontrar as variáveis de controle  $\mathbf{w}(\theta)$ , é necessário definir uma função experimental (no caso, o grau do polinômio  $P(\theta)$ ) e calcular a primeira e segunda derivadas de tal função, bem como a derivada do logaritmo da distribuição de interesse que, nesse caso, é a distribuição posteriori.

Esta tese aborda os casos em que a função experimental é linear e quadrática.

### **Polinômio Linear $P(\theta)$**

Seja  $P(\theta)$  um polinômio linear

$$P(\theta) = \mathbf{a}^T \theta, \quad (2.23)$$

onde  $\mathbf{a} \in \mathbb{R}^d$ . Nesse caso, o gradiente e o operador de *Laplace* são, respectivamente,

$$\Delta_{\theta}[P(\theta)] = \mathbf{a}^T, \quad \nabla_{\theta}[P(\theta)] = 0. \quad (2.24)$$

Portanto, a função auxiliar é

$$\tilde{g}(\theta) = g(\theta) + \mathbf{a}^T \mathbf{z}(\theta) \quad (2.25)$$



e o estimador *Zero-Variance* do método MCMC (ZV-MCMC) é dado por

$$\tilde{\mu}_{\tilde{g}(\theta)} = \hat{\mu}_{g(\theta)} + \hat{\mathbf{a}}^T \mathbf{z}(\theta^*), \quad (2.26)$$

onde  $\hat{\mathbf{a}} = \widehat{\text{Var}}[\mathbf{z}(\theta)]^{-1} \widehat{\text{Cov}}[g(\theta), \mathbf{z}(\theta)]$ ,  $\widehat{\text{Var}}[\mathbf{z}(\theta)]$  e  $\widehat{\text{Cov}}[g(\theta), \mathbf{z}(\theta)]$  são, respectivamente, a variância e a matriz de covariância amostrais.

### **Polinômio de Segundo Grau $P(\theta)$**

Seja  $P(\theta)$  um polinômio de segundo grau

$$P(\theta) = \mathbf{c}^T \theta + \frac{1}{2} \theta^T \mathbf{B} \theta, \quad (2.27)$$

onde  $\mathbf{c} \in \mathbb{R}^d$  e  $\mathbf{B}$  é uma matriz simétrica. Então, o gradiente e o operador de *Laplace* da função experimental quadrática são

$$\Delta_{\theta}[P(\theta)] = (\mathbf{c} + \mathbf{B}\theta)^T, \quad \nabla_{\theta}[P(\theta)] = \text{tr}(\mathbf{B}), \quad (2.28)$$

onde  $\text{tr}(\cdot)$  é o traço de uma matriz. Consequentemente, a função auxiliar é dada por

$$\tilde{g}(\theta) = g(\theta) - \frac{1}{2} \text{tr}(\mathbf{B}) + (\mathbf{c} + \mathbf{B}\theta)^T \mathbf{z}(\theta) \quad (2.29)$$

e pode ser reduzida para a forma

$$\tilde{g}(\theta) = g(\theta) + \mathbf{a}^T \mathbf{w}(\theta), \quad (2.30)$$

onde  $\mathbf{a}$  e  $\mathbf{w}(\theta)$  são vetores com  $d(d+3)/2$  elementos e são definidos conforme a seguir:

- $\mathbf{a} := (\mathbf{c}^T, \mathbf{d}^T, \mathbf{b}^T)$ , onde  $\mathbf{d} := \text{diag}(\mathbf{B})$  é a diagonal da matriz  $\mathbf{B}$  e  $\mathbf{b}$  é um vetor de  $d(d-1)/2$  elementos com  $b_k = \mathbf{B}(i, j)$  e  $k = (2d-j)(j-1)/2 + (i-j)$ ,  $\forall i > j$ ; e
- $\mathbf{w} := (\mathbf{z}^T, \mathbf{u}^T, \mathbf{v}^T)$ , onde  $\mathbf{u} := \theta \circ \mathbf{z} - \frac{1}{2} \mathbf{1}$ ,  $\circ$  é o produto de Hadamard,  $\mathbf{1}$  é um vetor de 1s e  $\mathbf{v}$  é um vetor de  $d(d-1)/2$  elementos com  $v_k = \theta_i z_j + \theta_j z_i$ , onde  $k = (2d-j)(j-1)/2 + (i-j)$ ,  $\forall i > j$ .

Observe-se que o estimador *Zero-Variance* para o polinômio de segundo grau é similar ao linear substituindo-se  $\mathbf{z}(\theta)$  por  $\mathbf{w}(\theta)$  na Equação [2.26](#).

### **Estimador Não-Viesado**

A Equação [2.20](#) não é suficiente para garantir que o estimador  $\hat{\mu}_{\tilde{g}(\theta)}$  seja assintoticamente não-viesado. Portanto, é necessário incluir algumas condições para satisfazer essa propriedade, apresentadas nas proposições a seguir (para mais detalhes, ver [Mira, Solgi e Imparato \(2013\)](#)).

**Proposição 2.** Seja  $\pi$  uma função densidade de uma variável aleatória e sejam  $l$  e  $u$ , respectivamente, os limites inferior e superior do suporte da variável aleatória. Então, a condição suficiente para que ZV-MCMC seja não-viesado é dada por

$$\pi(l) \frac{\partial P(l)}{\partial x} = \pi(u) \frac{\partial P(u)}{\partial x}. \quad (2.31)$$

**Proposição 3.** Seja  $\pi$  uma função densidade de um vetor aleatório, onde o suporte da distribuição  $\Omega$  é um conjunto aberto com limite regular  $\partial\Omega$ , cuja primeira e segunda derivadas são contínuas. Então, se  $\psi(\theta) = P(\theta) \sqrt{\pi(\theta | X)}$ , a condição suficiente para que ZV-MCMC seja não-viesado é dada por

$$\pi(\mathbf{X}) \frac{\partial P(\mathbf{X})}{\partial x_i} = 0, \mathbf{X} \in \partial\Omega. \quad (2.32)$$

Através da Proposição 3, conclui-se que, nos casos em que a função experimental é um polinômio de primeiro ou segundo grau, o estimador ZV-MCMC é não-viesado se a função densidade for igual a 0 para todo limite do suporte da distribuição. Portanto, o estimador ZV-MCMC não é indicado quando o estimador bayesiano é próximo do limite do suporte paramétrico. Observe-se o seguinte exemplo referente à distribuição exponencial.

**Exemplo 1.** Sejam  $g(\theta) = \theta$  e  $\theta | \mathbf{X} \sim \exp(\lambda)$ , ou seja,  $\pi(\theta | \mathbf{X}) = \lambda \exp(-\lambda\theta)$ . Se  $P(\theta)$  for um polinômio de primeiro grau, tem-se com base na Equação 2.26 que o estimador ZV-MCMC é dado por  $\tilde{\mu}_\theta = \hat{\mu}_\theta + a\lambda$ . Note-se que o estimador é viesado, exceto quando  $a = 0$ . No caso em que  $P(\theta) = a_0 + a_1x + a_2x^2$ , a Proposição 2 é satisfeita quando  $a_1 = 0$ . Portanto, o valor ótimo que minimiza a variância do estimador implica  $a_2 = \frac{1}{2\lambda}$ .

Para contornar o possível problema de o estimador ZV-MCMC ser viesado, no contexto de inferência bayesiana, basta que as regiões paramétricas não tenham restrição. Porém, caso tenham, pode-se realizar uma reparametrização que garanta a ausência de restrições paramétricas.

## 2.8 Método *Hamiltonian Monte Carlo*

O método *Hamiltonian Monte Carlo* (HMC) consiste em um método MCMC alternativo que combina alternadamente os métodos *Gibbs* (SPIEGELHALTER *et al.*, 1994, 2003) e *Metropolis-Hasting*. Uma de suas principais características é a habilidade de evitar o comportamento de passeio aleatório, uma vez que evita-se gerar cadeias com grande autodependência. No processo de geração de uma cadeia de *Markov*, este método gera um novo estado da cadeia através de uma simulação de trajetória que segue uma dinâmica hamiltoniana (NEAL, 2011).

Considere-se um vetor aleatório  $\theta \in \mathbb{R}^d$  (parâmetros) e  $\mathbf{r} \in \mathbb{R}^d$  um vetor aleatório auxiliar e independente, onde  $\mathbf{r} \sim N_d(0, M)$ . Então, a função de densidade conjunta de  $(\theta, \mathbf{r})$  é dada por

$$\pi(\theta, \mathbf{r}) \propto \exp(-H(\theta, \mathbf{r})), \quad (2.33)$$

onde  $H(\boldsymbol{\theta}, \mathbf{r}) = U(\boldsymbol{\theta}) + K(\mathbf{r})$  é uma função hamiltoniana,  $U(\boldsymbol{\theta}) = -\log(\pi(\boldsymbol{\theta} | \mathbf{x})\pi(\boldsymbol{\theta}))$  e  $K(\mathbf{r}) = \mathbf{r}M^{-1}\mathbf{r}$ . Analogamente à Física, o termo  $U(\boldsymbol{\theta})$  é chamado de energia potencial e o termo  $K(\mathbf{r})$  é chamado de energia cinética.

Este método propõe valores candidatos para que  $(\boldsymbol{\theta}, \mathbf{r})$  sejam gerados em dois estados, antes de serem submetidos ao passo de aceitação do método *Metropolis-Hastings*. No primeiro estado, um valor de  $\mathbf{r}$  é simulado a partir de uma distribuição normal com média  $\mathbf{0}$  e matriz de covariância  $M$  independente de  $\boldsymbol{\theta}$ . No segundo estado, simula-se um sistema conjunto  $(\boldsymbol{\theta}, \mathbf{r})$  que obedeça à dinâmica hamiltoniana. O sistema é desenvolvido através das equações hamiltonianas a seguir:

$$\frac{\partial \boldsymbol{\theta}}{\partial t} = \frac{H(\boldsymbol{\theta}, \mathbf{r})}{\partial \mathbf{r}} = \frac{\partial K(\mathbf{r})}{\partial \mathbf{r}} \quad (2.34)$$

$$\frac{\partial \mathbf{r}}{\partial t} = -\frac{H(\boldsymbol{\theta}, \mathbf{r})}{\partial \boldsymbol{\theta}} = -\frac{\partial U(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}. \quad (2.35)$$

As equações hamiltonianas podem ser aproximadas através da discretização do tempo, usando um pequeno passo de tamanho  $\varepsilon$  através do método *leapfrog*. Este método é utilizado para resolver as equações hamiltonianas por meio da aplicação de  $L$  passos, cada um definido por

$$\begin{aligned} r_i(t + \varepsilon/2) &= r_i(t) - (\varepsilon/2) \frac{\partial U(\boldsymbol{\theta}(t))}{\partial \theta_i} \\ \theta_i(t + \varepsilon) &= \theta_i(t + \varepsilon/2) - \varepsilon \frac{\partial K(\mathbf{r}(t + \varepsilon/2))}{\partial r_i} \\ r_i(t + \varepsilon) &= r_i(t + \varepsilon/2) - (\varepsilon/2) \frac{\partial U(\boldsymbol{\theta}(t + \varepsilon))}{\partial \theta_i}. \end{aligned}$$

Ao final, é simulado o estado  $(\boldsymbol{\theta}^*, \mathbf{r}^*)$ , que corresponde ao valor do sistema no tempo fictício  $L\varepsilon$ . Em seguida, utiliza-se a etapa de aceitação do método *Metropolis-Hastings*, em que o estado  $(\boldsymbol{\theta}^*, \boldsymbol{\omega}^*)$  é aceito como o próximo estado da cadeia de *Markov* com a seguinte probabilidade:

$$\mathbb{P}(\boldsymbol{\theta}, \boldsymbol{\omega}; \boldsymbol{\theta}^*, \boldsymbol{\omega}^*) = \min\{1, \exp\{H(\boldsymbol{\theta}, \boldsymbol{\omega}) - H(\boldsymbol{\theta}^*, \boldsymbol{\omega}^*)\}\}. \quad (2.36)$$

Para a execução do método HMC, deve-se definir a escolha de três parâmetros.

- $\varepsilon$ , referente à probabilidade de aceitação apresentada na Equação 2.36. Quanto menor for seu valor, maiores serão a probabilidade de aceitação e a autodependência da cadeia.
- $L$ , referente à autodependência da cadeia. Quanto maior for seu valor, menor será essa autodependência, porém maior será o custo computacional.
- $M$ , que é a combinação dos dois parâmetros recém-mencionados.

Apesar de existirem alguns algoritmos para tentar encontrar o valor ideal para cada um desses parâmetros, não há um consenso na literatura quanto ao critério de escolha. Portanto, esta tese

adota o critério de atribuição manual de valores que tragam o melhor custo-benefício em termos de autodependência e desempenho computacional.

O Algoritmo 1 descreve o procedimento do método HMC para a geração de dados que seguem uma distribuição posteriori.

---

**Algoritmo 1** – Procedimento do método HMC para geração de dados com distribuição posteriori

---

Dados  $\theta^0$ ,  $\varepsilon$ ,  $L$ ,  $\pi$ ,  $M$  e  $S$

**Para**  $i = 1$  a  $S$  **faça**

$$\theta^i \leftarrow \theta^{i-1}$$

$$\text{Simule } r^0 \sim N(0, M)$$

$$\text{Simule } u \sim \text{Uniforme}(0, 1)$$

$$\leftarrow \theta^i$$

$$\tilde{r} \leftarrow r^0$$

**Para**  $j = 1$  a  $L$  **faça**

$$\tilde{\theta}, \tilde{r} \leftarrow \text{Leapfrog}(\tilde{\theta}, \tilde{r}, \varepsilon, M)$$

**Fim do para**

**Se**  $(\log(\pi(\tilde{\theta})) - \frac{1}{2}\tilde{r}^T M \tilde{r} - \log(\pi(\theta)) + \frac{1}{2}r^{0T} M r^0) > \log(u)$  **faça**

$$\theta^m \leftarrow \tilde{\theta}$$

**Fim do para**

**função**  $\text{Leapfrog}(\theta, r, \varepsilon, M)$

$$\tilde{r} \leftarrow r + \frac{\varepsilon}{2} \nabla_{\theta} \pi(\theta)$$

$$\tilde{\theta} \leftarrow \theta + \varepsilon M^{-1} \tilde{r}$$

$$\tilde{r} \leftarrow r + \frac{\varepsilon}{2} \nabla_{\theta} \pi(\tilde{\theta})$$

**retorne**  $\tilde{\theta}, \tilde{r}$

---

## 2.9 Critérios para Seleção de Modelos

Uma tarefa recorrente na Estatística diz respeito à escolha do modelo mais apropriado dentre um conjunto de modelos. Busca-se realizar essa escolha com base em algum critério pré-definido. Idealmente, escolhe-se o modelo mais parcimonioso, ou seja, o modelo que possui o menor número possível de parâmetros que expliquem suficientemente o fenômeno estudado. Existe uma vasta gama de critérios de escolha de modelo, porém esta tese limitou-se a um conjunto de critérios que são baseados na distribuição posteriori. Esses critérios são os seguintes:

- Esperança da posteriori do Critério de Informação de Akaike (EAIC) (BROOKS, 2002):

$$EAIC(M) = -2\mathbb{E}_{\theta|\mathbf{X}}[\log(\pi(\mathbf{X} | \theta))] + 2np_m,$$

- Esperança da posteriori do Critério de Informação Bayesiano (EBIC) (CARLIN; LOUIS, 2000):

$$EBIC(M) = -2\mathbb{E}_{\theta|\mathbf{X}}[\log(\pi(\mathbf{X} | \theta))] + 2p_m \log(n),$$

- Critério de Informação de Desvio (DIC) (SPIEGELHALTER *et al.*, 2002):

$$DIC(M) = -2\mathbb{E}_{\theta|\mathbf{X}}[\log(\pi(\mathbf{X} | \theta))] + 2p_D,$$

- Critério de Informação de Watanabe (WAIC) (WATANABE, 2010):

$$WAIC(M) = -2\mathbb{E}_{\theta|\mathbf{X}}[\log(\pi(\mathbf{X} | \theta))] + 2p_{waic},$$

- Critério de Informação *Leave-One-Out cross-validation* (LOOIC) (VEHTARI; GELMAN; GABRY, 2016):

$$LOOIC(M) = -2\mathbb{E}_{\theta|\mathbf{X}}\left[\sum_i^n \log(\pi^{-i}(\mathbf{X} | \theta))\right],$$

onde  $p_m$  é o total de parâmetros do modelo,  $p_D = \mathbb{E}_{\theta|\mathbf{X}}[\log(\pi(\mathbf{X} | \theta))] - 2\log(\pi(\mathbf{X} | \hat{\theta}))$ ,  $p_{waic} = \sum_i^n \text{Var}_{\theta|\mathbf{X}}(\log(\pi(\mathbf{X} | \theta)) \text{ e } \pi^{-i}(\mathbf{X} | \theta))$  é a função de verossimilhança sem a  $i$ -ésima informação. As estimativas  $p_D$  e  $p_{waic}$  representam o número ideal de parâmetros do modelo, com a diferença de que  $p_{waic}$  utiliza a variância da posteriori do logaritmo da função de verossimilhança. Isso torna a estimativa  $p_{waic}$  mais apropriada, porém com um custo computacional maior.

Os critérios AIC e BIC são bastante utilizados na abordagem clássica, porém na abordagem bayesiana é mais indicado utilizar a esperança da posteriori desses critérios, conforme apresentado anteriormente. O critério DIC é amplamente usado na inferência bayesiana, principalmente quando métodos MCMC são empregados. Os critérios WAIC e LOOIC podem ser interpretados como uma aproximação conveniente para validação cruzada. No caso em que  $\mathbf{X}$  é independente, o critério LOOIC converge assintoticamente para o critério WAIC. Em resumo, quanto menor for o valor de cada critério, melhor será o ajuste do modelo.



## MODELO GARCH ASSIMÉTRICO

Os modelos da família *Autoregressive Conditional Heteroscedasticity* (ARCH), primeiramente introduzidos por Engle (1982) e generalizados por Bollerslev (1986), têm sido alvo de grande interesse nas últimas décadas. Esses modelos, cujo propósito é voltado a séries temporais financeiras, apresentam o conceito de volatilidade, que é a variância condicionada ao passado da série. Nos modelos clássicos *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH) propostos por Bollerslev (1986), a variância condicional consiste em uma função linear dos valores do passado da série, considerando sua volatilidade e seus dados. As extensões a estes modelos incluíram expressões mais sofisticadas, tais como modelos de limiar para capturar a assimetria da variância condicional, como é o caso do modelo Glosten-Jagannathan-Runkle GARCH (GJR-GARCH), introduzido por Glosten, Jagannathan e Runkle (1993).

Este capítulo apresenta duas classes de modelo: (i) uma univariada, que corresponde ao modelo GJR-GARCH, e (ii) uma multivariada, denominada *Dynamic Conditional Correlation* GJR-GARCH (DCC-GJR-GARCH), que é um caso especial do modelo apresentado por Cappiello, Engle e Sheppard (2006).

### 3.1 GJR-GARCH

Seja  $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^T$  um processo estocástico. O modelo GJR-GARCH(p,q) é definido como

$$\begin{aligned}
 X_t &= Y_t + \mu, \\
 Y_t &= \sqrt{h_t} \varepsilon_t, \\
 h_t &= \omega + \sum_{i=1}^p (\alpha_i + \phi_i I_{t-i}) Y_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \\
 \varepsilon_t &\sim D(0, 1, \nu),
 \end{aligned} \tag{3.1}$$

com os parâmetros  $\omega > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$ ,  $\phi_i \geq -\beta$ ,  $\mu \in \mathbb{R}$ ,  $\forall i = 1, \dots, p$  e  $\forall j = 1, \dots, q$ . Adicionalmente,  $I_{t-1}$  é uma função indicadora cujos valores são:  $I_{t-1} = 1$  se  $X_{t-1} < 0$  e  $I_{t-1} = 0$  caso contrário. As inovações  $\varepsilon_t$  são independentes e identicamente distribuídos.  $D(0, 1, \nu)$  é uma distribuição de probabilidade com média 0, variância 1 e parâmetro de forma  $\nu$ . Especificamente, no caso em que  $\phi = 0$ , o modelo da Equação 3.1 é um modelo GARCH.

Seja  $\mathcal{F}_{t-1} = (X_1, X_2, \dots, X_{t-1})^T$  o passado do processo estocástico no instante  $t$ . Ao calcular-se a esperança e a variância condicional dado o passado ( $\mathcal{F}_{t-1}$ ), tem-se que

$$\begin{aligned} \mathbb{E}(X_t | \mathcal{F}_{t-1}) &= \mathbb{E}(Y_t + \mu | \mathcal{F}_{t-1}) \\ &= \mathbb{E}(\sqrt{h_t} \varepsilon_t | \mathcal{F}_{t-1}) + \mu \\ &= \mu \end{aligned} \tag{3.2}$$

e

$$\begin{aligned} \text{Var}(X_t | \mathcal{F}_{t-1}) &= \text{Var}(Y_t + \mu | \mathcal{F}_{t-1}) \\ &= \text{Var}(\sqrt{h_t} \varepsilon_t | \mathcal{F}_{t-1}) \\ &= h_t. \end{aligned} \tag{3.3}$$

Portanto, o parâmetro  $\mu$  representa a média do processo estocástico, enquanto  $h_t$  é a variância condicional de  $X_t$  dado o passado do processo estocástico ( $\mathcal{F}_{t-1}$ ). Note-se que  $h_t$  é uma variável não observável.

### 3.1.1 Condições de Estacionariedade

Ao utilizar uma série temporal para fazer inferência, é importante saber se o processo estocástico que a gerou é estacionário. Portanto, para qualquer modelo aplicado a uma série temporal, é necessário ter o conhecimento das condições de estacionariedade. Nesse contexto, são apresentadas a seguir as condições de estacionariedade do modelo GJR-GARCH(p,q).

Inicialmente, é importante relembrar a definição de um processo fracamente estacionário ou estacionário de segunda ordem.

**Definição 3.** Diz-se que um processo estocástico  $Z = \{Z_t, t \in T\}$  é fracamente estacionário ou estacionário de segunda ordem se, e somente se,

- (i)  $\mathbb{E}[Z_t] = \mu_t = \mu$ , constante para todo  $t \in T$ ;
- (ii)  $\mathbb{E}[Z_t^2] < \infty$ , para todo  $t \in T$ ;
- (iii)  $\gamma(t_1, t_2) = \text{Cov}[Z_{t_1}, Z_{t_2}]$  é função de  $|t_1 - t_2|$ .



**Proposição 4.** Um processo GJR-GARCH(p,q) é fracamente estacionário se

$$\sum_{i=1}^p (\alpha_i + p\phi_i) + \sum_{j=1}^q \beta_j < 1, \quad (3.4)$$

onde  $p = \mathbb{E}[\varepsilon_t^2 \mid \varepsilon_t > 0]$ .

*Demonstração.* A Condição (i) é satisfeita como consequência direta da Equação 3.2. Já as Condições (ii) e (iii) são verificadas conjuntamente a seguir. Dados  $t_1 = t_2 = t$ , tem-se que

$$\begin{aligned} \sigma_t^2 &= \text{Var}[\mathbf{X}_t] \\ &= \text{Var}[Y_t] \\ &= \text{Var}[\mathbb{E}[Y_t \mid \mathcal{F}_{t-1}]] + \mathbb{E}[\text{Var}[Y_t \mid \mathcal{F}_{t-1}]] \\ &= \mathbb{E}[h_t] \\ &= \mathbb{E}\left[\omega + \sum_{i=1}^p (\alpha_i + \phi_i I_{t-1}) Y_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-i}\right] \\ &= \omega + \sum_{i=1}^p \alpha_i \mathbb{E}[Y_{t-i}^2] + \sum_{i=1}^p \phi_i \mathbb{E}[I_{t-1} Y_{t-i}^2] + \sum_{j=1}^q \beta_j \mathbb{E}[h_{t-i}] \\ &= \omega + \sum_{i=1}^p \alpha_i \sigma_{t-i}^2 + \sum_{i=1}^p \phi_i p \sigma_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \\ &= \omega + \sum_{i=1}^p (\alpha_i + p\phi_i) \sigma_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2. \end{aligned}$$

No caso em que  $t_1 < t_2$ , tem-se que

$$\begin{aligned} \gamma(t_1, t_2) &= \text{Cov}[X_{t_1}, X_{t_2}] \\ &= \text{Cov}[Y_{t_1}, Y_{t_2}] \\ &= \text{Cov}[Y_{t_1}, \mathbb{E}[Y_{t_2} \mid Y_{t_1}]] \\ &= \text{Cov}[Y_{t_1}, 0] \\ &= 0. \end{aligned}$$

Assim, é possível concluir que a Condição (iii) é satisfeita, ou seja, as variâncias são iguais. Por fim, com relação à Condição (ii), as variâncias são finitas se a seguinte Equação tiver solução:

$$\sigma^2 = \omega + \sum_{i=1}^p (\alpha_i + p\phi_i) \sigma^2 + \sum_{j=1}^q \beta_j \sigma^2, \quad \sigma^2 > 0, \quad (3.5)$$

que ocorre quando  $\sum_{i=1}^p (\alpha_i + p\phi_i) + \sum_{j=1}^q \beta_j < 1$ .  $\square$

Uma consequência direta da Proposição 4 é que, se o processo  $Y_t$  é estacionário, então  $Y_t$  é não correlacionado. Outra consequência dessa proposição é que, no caso em que  $\mu = 0$ ,

o modelo é um ruído branco, *i.e.* um processo estocástico não correlacionado com média 0 e variância constante. Esta característica é comum em qualquer modelo que pertença à família GARCH e modelos de volatilidade em geral, sendo recorrente em séries temporais financeiras, como ações ou índices de bolsas de valores. Outra consequência é que a variância do processo é dada por

$$\text{Var}(\mathbf{X}_t) = \frac{\omega}{1 + \sum_{i=1}^p (\alpha_i + p\phi_i) + \sum_{j=1}^q \beta_j}. \quad (3.6)$$

**Observação 1.** Se a distribuição dos inovações do modelo for simétrica em torno de 0, então  $p = 0,5$ .

### 3.1.2 Função de Verossimilhança

Através da Equação 3.1, pode-se escrever a Função Densidade de Probabilidade (f.d.p.) condicional de  $\mathbf{X}_t$  dado  $\mathcal{F}_{t-1}$ . Como resultado:

$$f_{\mathbf{X}_t|\mathcal{F}_{t-1}}(\mathbf{X}_t) = h_t^{-1/2} f_{\varepsilon_t} \left( \frac{\mathbf{X}_t - \mu}{\sqrt{h_t}} \right), \quad (3.7)$$

onde  $f_{\varepsilon_t}(\cdot)$  é uma f.d.p. de  $\varepsilon_t$ . Portanto, o cálculo da função de verossimilhança é dado por:

$$\begin{aligned} L(\theta) &= f_{\mathbf{X}}(\mathbf{x}) \\ &= f_{\mathbf{X}|X_0}(\mathbf{x}) f_{X_0}(x_0) \\ &= \left[ \prod_{t=1}^n h_t^{-1/2} f_{\varepsilon_t} \left( \frac{\mathbf{X}_t - \mu}{\sqrt{h_t}} \right) \right] f_{X_0}(x_0), \end{aligned} \quad (3.8)$$

sendo  $\theta = (\omega, \alpha_1, \dots, \alpha_p, \phi_1, \dots, \phi_p, \beta_1, \dots, \beta_q, \nu)^T$  o conjunto de todos os parâmetros.

Vale notar que, na prática, não se pode obter a função de verossimilhança, pois não há conhecimento sobre a distribuição marginal de  $X_0$ . Logo, é necessário utilizar a função de verossimilhança condicional, que é dada por:

$$L_c(\theta) = f_{\mathbf{X}|X_0}(\mathbf{x}) = \prod_{t=1}^n f_{\mathbf{X}_t|\mathcal{F}_{t-1}}(\mathbf{X}_t). \quad (3.9)$$

Então, define-se a função de verossimilhança condicional do modelo GJR-GARCH como

$$L_c(\theta) = \prod_{t=1}^n h_t^{-1/2} f_{\varepsilon_t} \left( \frac{\mathbf{X}_t - \mu}{\sqrt{h_t}} \right), \quad (3.10)$$

sendo  $h_0 = 1$  e  $x_0 = 0$ .

A distribuição marginal de  $\mathbf{X}_t$  de qualquer modelo da família GARCH com  $\varepsilon_t \sim N(0, 1)$  possui cauda mais pesada do que a distribuição normal. Entretanto, há evidências na literatura de que muitas séries financeiras tendem a ter curtose observada maior do que qualquer modelo da família GARCH com erros normais. Portanto, muitos autores propuseram distribuições de caudas pesadas, como por exemplo as distribuições *t de Student* (chamada de *t* no decorrer desta

tese) e *Generalized Error Distribution* (GED) (BAILLIE; BOLLERSLEV, 2002; CHUANG; LU; LEE, 2007), para modelos GARCH. Diante disso, este trabalho estuda esses três casos de distribuições das inovações.

### 3.1.3 Distribuições das Inovações

Esta seção apresenta as distribuições adotadas para as inovações do modelo.

#### Distribuição Normal

A distribuição normal é possivelmente a distribuição mais conhecida na área de Inferência Estatística. Seja  $Z$  uma variável aleatória que segue uma distribuição normal com parâmetros  $\mu = e$  e  $\sigma^2 = 1$ , expressa por  $Z \sim N(0, 1)$ . A variável  $Z$  tem a seguinte f.d.p.:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2}\right), \quad (3.11)$$

Como  $E[Z] = 0$  e  $\text{Var}[Z] = 1$ , assume-se que  $\varepsilon_t \sim N(0, 1) \forall t$ .

#### Distribuição t

A distribuição t é bastante utilizada na Inferência Clássica, principalmente no contexto de intervalos de confiança e testes de hipótese para média amostral de uma distribuição normal. Seja  $Z$  uma variável aleatória que segue uma distribuição t com parâmetros 0, e  $\nu$ , expressa por  $Z \sim t(0, \nu)$ . A f.d.p. de  $Z$  é definida como

$$f_{Z|\nu}(z) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\sigma\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{z^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}, \quad (3.12)$$

com  $\nu \in \mathbb{R}^+$ . Se  $\nu > 2$ , então  $E[Z] = 0$  e  $\text{Var}[Z] = 1$ . Logo, para que  $\varepsilon_t$  seja uma inovações, assume-se que  $\varepsilon_t \sim t(0, \nu) \forall t$ .

#### Distribuição GED

A distribuição GED, também conhecida como *exponential power distribution* de Box e Tiao (1973), é uma generalização da distribuição normal. Seja  $Z$  uma variável aleatória que segue uma distribuição GED com parâmetros 0, e  $\nu$ , expressa por  $Z \sim GED(0, \nu)$ . A f.d.p. de  $Z$  é dada por

$$f_{Z|\nu}(z) = \frac{1}{2^{1+\frac{1}{2\nu}}\Gamma\left(1 + \frac{1}{2\nu}\right)\sigma} \exp\left(-\left(\frac{\Gamma\left(\frac{3}{2\nu}\right)z^2}{\Gamma\left(\frac{1}{2\nu}\right)}\right)^\nu\right), \quad (3.13)$$

com  $\nu \in \mathbb{R}^+$ . No caso em que  $\nu = 1$ , a distribuição GED é uma distribuição normal. Tem-se que  $E[Z] = 0$  e  $\text{Var}[Z] = 1$ . Logo, para que  $\varepsilon_t$  seja uma inovações, assume-se que  $\varepsilon_t \sim GED(0, \nu) \forall t$ .

A próxima seção apresenta o caso multivariado, representado pelo modelo DCC-GJR-GARCH, que é uma extensão do modelo DCC-GARCH (ENGLE, 2002).

## 3.2 DCC-GJR-GARCH

Seja  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_T)$  um processo estocástico multivariado de dimensão  $k$ . Um modelo GARCH multivariado pode ser definido como

$$\begin{aligned}\mathbf{X}_t &= \mathbf{Y}_t - \boldsymbol{\mu} \\ \mathbf{Y}_t &= \mathbf{H}_t^{1/2} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\varepsilon}_t &\sim D_k(\mathbf{0}, \mathbf{I}_k, \nu),\end{aligned}\tag{3.14}$$

onde  $\boldsymbol{\mu}$  é um vetor de parâmetros;  $\mathbf{H}_t^{1/2}$  é uma matriz simétrica positiva definida de dimensão  $k$ , obtida por meio da decomposição de Cholesky de uma matriz  $\mathbf{H}_t$ ;  $\boldsymbol{\varepsilon}_t$  são inovações multivariadas de dimensão  $k$ , que são independentes e identicamente distribuídos; e  $D_k(\mathbf{0}, \mathbf{I}_k, \nu)$  é uma distribuição multivariada de probabilidade com o vetor de médias igual a  $\mathbf{0}$ , matriz de variância e covariância  $\mathbf{I}_k$  (matriz identidade) com parâmetros de forma  $\nu$ .

Semelhantemente ao modelo GJR-GARCH, é possível calcular o vetor de médias e a matriz de variância e covariância dado o passado da série  $\mathcal{F}_{t-1}$ . Tem-se que:

$$\begin{aligned}\mathbb{E}(\mathbf{X}_t \mid \mathcal{F}_{t-1}) &= \mathbb{E}(\mathbf{Y}_t + \boldsymbol{\mu} \mid \mathcal{F}_{t-1}) \\ &= \mathbb{E}(\mathbf{H}_t^{1/2} \boldsymbol{\varepsilon}_t \mid \mathcal{F}_{t-1}) + \boldsymbol{\mu} \\ &= \boldsymbol{\mu}\end{aligned}\tag{3.15}$$

e

$$\begin{aligned}\text{Var}(\mathbf{X}_t \mid \mathcal{F}_{t-1}) &= \text{Var}(\mathbf{Y}_t + \boldsymbol{\mu} \mid \mathcal{F}_{t-1}) \\ &= \text{Var}(\mathbf{H}_t^{1/2} \boldsymbol{\varepsilon}_t \mid \mathcal{F}_{t-1}) \\ &= \mathbf{H}_t.\end{aligned}\tag{3.16}$$

Portanto,  $\mathbf{H}_t$  é a matriz de variâncias e covariâncias condicionais, que são denominadas, respectivamente, volatilidades e covolatilidades. A maneira como é definida a matriz  $\mathbf{H}_t$  é o que, por sua vez, define um modelo GARCH multivariado.

A literatura apresenta diversas propostas para especificar  $H_t$ . Segundo [Silvennoinen e Teräsvirta \(2009\)](#), os modelos GARCH multivariados podem ser divididos em quatro categorias:

- **Modelos de matriz de variância e covariância condicional.** Nesta categoria, a matriz de variância e covariância condicional é obtida diretamente. Exemplos: VEC-GARCH de [Engle e Kroner \(1995\)](#) e BEKK-GARCH de [Bollerslev, Engle e Wooldridge \(1988\)](#);
- **Modelos de fatores.** Estes modelos assumem que o processo  $\mathbf{X}$  é gerado por fatores heterocedásticos não observados ([ENGLE; NG; ROTHSCHILD, 1990](#));

- **Modelos de variâncias e correlações condicionais.** Esta classe modela as variâncias e correlações condicionais ao invés de modelar diretamente a matriz de covariância. Exemplos: CCC-GARCH de Bollerslev (1990) e DCC-GARCH Engle (2002);
- **Modelos não paramétricos e semiparamétricos.** Estes modelos não impõem uma estrutura particular para a matriz de covariância condicional (LINTON, 2009).

Este trabalho estuda o modelo DCC-GJR-GARCH, proposto por Cappiello, Engle e Sheppard (2006). Em relação às quatro classes recém-apresentadas, este é um modelo da terceira classe, baseada nas variâncias e correlações condicionais. O modelo define a matriz  $\mathbf{H}_t$  como:

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t, \quad (3.17)$$

onde

$$\mathbf{D}_t = \text{diag}(h_{11,t}^{1/2}, \dots, h_{kk,t}^{1/2}). \quad (3.18)$$

Assume-se que  $h_{ii,t}$  é definido como um GJR-GARCH  $\forall i = 1, 2, \dots, k$ . Neste trabalho, considere-se que  $h_{ii,t}$  é um GJR-GARCH(1,1). Portanto,  $h_{ii,t}$  pode ser escrito como

$$h_{ii,t} = \omega_i + (\alpha_i + \phi_i I_{i,t-1}) X_{i,t-1}^2 + \beta_i h_{ii,t-1}, \quad (3.19)$$

onde  $\omega_i > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_i \geq 0$ ,  $\phi_i \geq -\beta_i$ ,  $\mu_i \in \mathbb{R}$  e  $I_{i,t-1}$  é uma função indicadora, definida como  $I_{i,t-1} = 1$  se  $Y_{i,t-1} < 0$  e  $I_{i,t-1} = 0$  caso contrário  $\forall i = 1, 2, \dots, k$ . A matriz  $\mathbf{R}_t$  é uma matriz de correlações condicionais, expressa como

$$\begin{aligned} \mathbf{R}_t &= \text{diag}(\mathbf{Q}_t)^{-1/2} \mathbf{Q}_t \text{diag}(\mathbf{Q}_t)^{-1/2} \\ \mathbf{Q}_t &= (1 - \lambda_1 - \lambda_2) \Sigma + \lambda_1 \mathbf{u}_{t-1} \mathbf{u}_{t-1}^T + \lambda_2 \mathbf{Q}_{t-1}, \end{aligned} \quad (3.20)$$

sendo  $\Sigma$  uma matriz simétrica e positiva definida,  $\mathbf{u}_t = \mathbf{D}_t^{-1} \mathbf{X}_t$ , onde  $\odot$  denota o produto de Hadamard.

Na literatura, o parâmetro  $\Sigma$  é usualmente estimado por meio da matriz de variância e covariância amostral, mesmo em casos que utilizam a Inferência Bayesiana. Entretanto, isso é contraintuitivo, pois na Inferência Bayesiana os parâmetros são tratados como variáveis aleatórias. Um dos problemas em se utilizar o estimador baseado em tal matriz é que não há garantias de que a esperança de  $\mathbf{Q}_t$  seja igual à esperança de  $\mathbf{u}_t$ , o que torna esse estimador inadequado.

Outro fato sobre o parâmetro  $\Sigma$  é que, por definição, ele representa a matriz de correlação  $\mathbf{u}_t$  se o processo for estacionário e se a esperança de  $\mathbf{Q}_t$  for igual a  $\Sigma$ . Nesse contexto, para que essa definição seja sempre satisfeita, este trabalho de doutorado inclui a seguinte restrição no modelo DCC-GJR-GARCH:

$$\mathbb{E}(\mathbf{Q}_t) = \mathbb{E}(\mathbf{u}_t \mathbf{u}_t^T). \quad (3.21)$$

Adicionalmente, diferentemente dos trabalhos existentes na literatura até então, este trabalho de doutorado considera um processo de estimação **totalmente bayesiano**.

Por fim, o processo  $\mathbf{X}_t$  é um processo estocástico estacionário se  $h_{ii,t}$  for estacionário e se  $\mathbf{R}_t$  for uma matriz positiva definida. Supondo-se que a distribuição dos  $\boldsymbol{\varepsilon}_t$  seja simétrica e usando-se a Equação 3.4, tem-se as seguintes restrições de estacionariedade:

$$\begin{aligned} 1 - \omega_i - \alpha_i - 0,5\phi_i + \beta_i &< 1, \quad \forall i = 1, \dots, k; \\ 1 - \lambda_1 - \lambda_2 &< 1. \end{aligned} \quad (3.22)$$

### 3.2.1 Função de Verossimilhança

Da Equação 3.14, obtém-se a função das distribuições condicionais da série em função da distribuição das inovações:

$$f_{\mathbf{X}_t | \mathcal{F}_{t-1}}(\mathbf{x}_t) = |\mathbf{H}_t|^{-1/2} f_{\boldsymbol{\varepsilon}_t} \left( \mathbf{H}_t^{-1/2} (\mathbf{x}_t - \boldsymbol{\mu}) \right). \quad (3.23)$$

Semelhantemente ao que ocorre no modelo GJR-GARCH, obtém-se a função de verossimilhança por:

$$L(\boldsymbol{\theta}) = f_{\mathbf{X} | \mathbf{X}_0}(\mathbf{x}) = \left[ \prod_{t=1}^n |\mathbf{H}_t|^{-1/2} f_{\boldsymbol{\varepsilon}_t} \left( \mathbf{H}_t^{-1/2} (\mathbf{x}_t - \boldsymbol{\mu}) \right) \right] f_{\mathbf{X}_0}(\mathbf{x}_0), \quad (3.24)$$

sendo  $\boldsymbol{\theta} = (\boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\phi}, \boldsymbol{\beta}, \Sigma, \boldsymbol{\lambda}, \boldsymbol{\mu}, \nu)^T$  o vetor de parâmetros do modelo DCC-GJR-GARCH. Assim como no modelo GJR-GARCH, a distribuição de probabilidade de  $\mathbf{X}_0$  não é conhecida. Por isso, substitui-se a função de verossimilhança pela função de verossimilhança condicional, que é obtida por:

$$L_c(\boldsymbol{\theta}) = \prod_{t=1}^n |\mathbf{H}_t|^{-1/2} f_{\boldsymbol{\varepsilon}_t} \left( \mathbf{H}_t^{-1/2} (\mathbf{x}_t - \boldsymbol{\mu}) \right), \quad (3.25)$$

onde  $\mathbf{H}_1$  é uma matriz de identidade e  $\mathbf{x}_1$  é igual a zero.

De acordo com a Equação 3.14, é preciso especificar uma distribuição multivariada para  $\boldsymbol{\varepsilon}_t$  que possua vetor de médias igual a zero e matriz de variância e covariância igual à matriz identidade. A próxima seção apresenta algumas distribuições com essas características, que são adotadas neste trabalho.

### 3.2.2 Distribuições das Inovações

As distribuições apresentadas nesta seção consistem em extensões multivariadas das distribuições estudadas no modelo GJR-GARCH.

### Distribuição Normal Multivariada

Seja  $\mathbf{Z}$  um vetor aleatório de dimensão  $k$  que segue uma distribuição normal multivariada padrão, expresso por  $\mathbf{Z} \sim N_k(\mathbf{0}, I_k)$ . A f.d.p. de  $\mathbf{Z}$  é dada por:

$$f_{\mathbf{Z}}(\mathbf{z}) = (2\pi)^{-\frac{k}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{\mathbf{z}^T \mathbf{z}}{2}\right), \quad (3.26)$$

Como  $\mathbb{E}[\mathbf{Z}] = \mathbf{0}$  e  $\text{Var}[\mathbf{Z}]$  é a matriz de identidade, assume-se que  $\boldsymbol{\varepsilon}_t \sim N_k(\mathbf{0}, I_k) \forall t$ .

### Distribuição $t$ Multivariada

Seja  $\mathbf{Z}$  um vetor aleatório de dimensão  $k$  que segue uma distribuição  $t$  multivariada padrão com parâmetro  $\nu$ , expresso por  $\mathbf{Z} \sim t_k(\nu)$ . A f.d.p. de  $\mathbf{Z}$  é dada por:

$$f_{\mathbf{Z}|\boldsymbol{\mu}, \Sigma, \nu}(\mathbf{z}) = \frac{1}{(\pi\nu)^{\frac{k}{2}}} \frac{\Gamma\left(\frac{\nu+k}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} |\Sigma|^{-\frac{1}{2}} \left(1 + \frac{\mathbf{z}^T \mathbf{z}}{\nu-2}\right)^{-\frac{\nu+k}{2}}, \quad (3.27)$$

onde  $\nu > 0$ . Além disso,  $\mathbb{E}[\mathbf{Z}] = \mathbf{0}$  e  $\text{Var}[\mathbf{Z}]$  é uma matriz identidade para  $\nu > 2$ . Portanto, assume-se que  $\boldsymbol{\varepsilon}_t \sim t_k(\nu)$  com  $\nu > 4$ .

### Distribuição GED Multivariada

A distribuição GED multivariada foi introduzida por [Gómez, Gomez-Viilegas e Marin \(1998\)](#), naturalmente como uma extensão multivariada da distribuição GED. Seja  $\mathbf{Z}$  um vetor aleatório de dimensão  $k$  que segue uma distribuição GED multivariada com parâmetro  $\nu$ , expresso por  $\mathbf{Z} \sim GED_k(\nu)$ . A f.d.p. de  $\mathbf{Z}$  é dada por:

$$f_{\mathbf{Z}|\boldsymbol{\mu}, \Sigma, \nu}(\mathbf{z}) = \frac{k\Gamma\left(\frac{k}{2}\right)}{\pi^{\frac{k}{2}} \Gamma\left(1 + \frac{k}{2\nu}\right) 2^{1+\frac{k}{2\nu}}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\left(\frac{\Gamma\left(\frac{k+2}{2\nu}\right) \mathbf{z}^T \mathbf{z}}{k\Gamma\left(\frac{k}{2\nu}\right)}\right)^\nu\right), \quad (3.28)$$

onde  $\nu > 0$ . Temos que  $\mathbb{E}[\mathbf{Z}] = \mathbf{0}$  e  $\text{Var}[\mathbf{Z}]$  é uma matriz identidade. Portanto, assume-se que  $\boldsymbol{\varepsilon}_t \sim GED_k(\nu)$ .

Na literatura, a maioria dos modelos GARCH multivariados apresentados considera a distribuição normal para as inovações, enquanto alguns consideram a distribuição  $t$ . Apesar da existência de trabalhos com enfoque bayesiano, o estudo apresentado neste trabalho de doutorado é o primeiro trabalho bayesiano a abordar a estimação conjunta de todos os parâmetros do modelo.

## 3.3 Estimação Bayesiana

Para o caso univariado, a estimação de parâmetros foca no modelo GJR-GARCH(1,1), apresentado na Seção [3.1](#), devido ao fato de este modelo ser mais parcimonioso com respeito à descrição de sua volatilidade, aliado ao fato de sua distribuição de probabilidade das inovações

seguir uma dentre as distribuições normal, t e GED. Analogamente, para o caso multivariado, a estimação de parâmetros foca no modelo DCC-GJR-GARCH, apresentado na Seção 3.2.

Conforme mencionado no Capítulo 2, na Inferência Bayesiana, é necessário determinar a distribuição priori dos parâmetros do modelo, denotada por  $\pi(\theta)$ . Como a distribuição priori determina o conhecimento inicial sobre os parâmetros, a escolha de tal distribuição depende do domínio de aplicação (conforme apresentado nas simulações do Capítulo 4 e nos experimentos do Capítulo 5). Já a distribuição posteriori dos modelos estudados pode ser expressa por

$$\pi(\theta | \mathbf{X}) \propto \pi(\theta)L_c(\theta), \quad (3.29)$$

onde  $\theta$  é um vetor com os parâmetros do modelo e  $L_c$  é a função de verossimilhança condicional, cuja fórmula foi apresentada para cada modelo nas Equações 3.10 e 3.25.

Nos modelos apresentados, a distribuição posteriori dos parâmetros não é conhecida. Portanto, é preciso utilizar métodos computacionais para fazer as inferências dos modelos. Este trabalho de doutorado utiliza os métodos HMC e *Zero-Variance*, apresentados no Capítulo 2. Ambos os métodos são mais apropriados em casos em que a região paramétrica não possui restrições, diferentemente do que ocorre nos modelos GARCH, pois sabe-se que existem restrições em seus parâmetros, principalmente restrições de estacionariedade. Assim, para que tais métodos possam ser empregados na estimação de parâmetros de modelos GARCH, é necessário realizar transformações paramétricas que os tornem livres de restrições. Diante disso, tem-se uma nova distribuição posteriori para essa reparametrização, que é dada por

$$\pi(\theta^* | \mathbf{X}) \propto |J(\theta^*)| \pi(\theta)L_c(\theta), \quad (3.30)$$

onde  $\theta^* = \tau(\theta)$  é um vetor de parâmetros sem restrições, resultante da aplicação da função biunívoca  $\tau$  sobre o vetor de parâmetros restritos, e  $|J(\theta^*)|$  é o Jacobiano da transformação. Portanto, faz-se a simulação da distribuição posteriori de  $\theta^*$  e, em seguida, aplica-se uma transformação inversa  $\tau^{-1}$  para reobter os parâmetros originais.

Nos dois modelos apresentados, as restrições podem ser separadas em três categorias, descritas a seguir.

### **Restrição 1: Limite Inferior**

Os parâmetros  $\omega$  e  $\nu$  dos modelos estudados apresentam restrição de limite inferior. Sem perda de generalidade, um parâmetro limitado inferiormente pode ser definido conforme a seguir.

**Proposição 5.** Seja  $W$  uma variável aleatória sem restrições. Ao realizar-se a transformação

$$V = \exp(W) + a,$$

obtem-se a variável aleatória  $V$  com limite inferior  $a$ . Uma vez que a função de transformação é biunívoca, aplica-se o seguinte Jacobiano:

$$|J| = \frac{\partial v}{\partial w} = \exp(w).$$



Analogamente, denota-se  $\omega^*$  e  $\nu^*$ , respectivamente, os parâmetros do modelo,  $\omega$  e  $\nu$  sem a restrição de limite inferior, obtidos por meio da transformação definida na Proposição 5.

### Restrição 2: Restrição Simplex

Uma restrição *simplex* é definida da seguinte forma:

$$\left( \sum_{i=1}^n a_i \right) < 1,$$

onde  $a_i > 0$  com  $1 \leq i \leq n$ ,  $i \in \mathbb{Z}$ . Com relação aos modelos estudados, para que os parâmetros  $\alpha$ ,  $\beta$  e  $\phi$  passem a ter uma restrição *simplex*, é necessário realizar a reparametrização a seguir. Sejam  $\alpha^+$  e  $\alpha^-$  definidos por

$$\alpha^- = (\alpha - \phi),$$

$$\alpha^+ = (\alpha + \phi).$$

Então,  $\alpha^-$ ,  $\alpha^+$  e  $\beta$  possuem, conjuntamente, uma restrição *simplex*. Sem perda de generalidade, um conjunto de parâmetros com restrição *simplex* pode ser definido conforme a seguir.

**Proposição 6.** Seja  $\mathbf{W}$  uma variável aleatória multivariada com três parâmetros sem restrições. Além disso, sejam  $\mathbf{V}$  e  $\mathbf{S}$  os seguintes vetores aleatórios com três parâmetros:

$$V_k = \text{logit}^{-1} \left( W_k + \log \left( \frac{1}{4-k} \right) \right), \quad k = 1, 2, 3, \quad (3.31)$$

$$S_k = \left( 1 - \sum_{i=1}^{K-1} x_i \right) V_k. \quad (3.32)$$

Nesse contexto, é possível provar que as duas transformações que definem, respectivamente, os vetores aleatórios  $V$  e  $S$  são biunívocas. Portanto, aplica-se o seguinte Jacobiano:

$$|\det J| = \prod_{i=1}^{K-1} z_k (1 - z_k) \left( 1 - \sum_{i=1}^{K-1} x_i \right) \quad (3.33)$$

Analogamente, denota-se  $\alpha^*$ ,  $\beta^*$  e  $\phi^*$ , respectivamente, os parâmetros do modelo  $\alpha$ ,  $\beta$  e  $\phi$  sem a restrição *simplex*, obtidos por meio das transformações definidas na Proposição 6.

### Restrição 3: Fatores de Cholesky de Matriz de Correlação

Conforme apresentado anteriormente, é requerido que o parâmetro  $\Sigma$  do modelo DCC-GJR-GARCH seja uma matriz de correlação, isto é,  $\Sigma$  deve ser uma matriz positiva definida com a diagonal principal igual a 1. Para que esta restrição seja satisfeita, pode-se reparametrizar  $\Sigma$  por meio da decomposição de Cholesky. Sem perda de generalidade, um conjunto de parâmetros com restrição de fatores de Cholesky de matriz de correlação pode ser definido conforme a seguir.

**Proposição 7.** Suponha-se uma variável aleatória multivariada  $\mathbf{W}$  com dimensão  $(\frac{k}{2}(k-1))$  sem restrições. Seja  $\mathbf{V}$  o vetor aleatório

$$\mathbf{V} = \frac{\exp(2\mathbf{W}) - 1}{\exp(2\mathbf{W}) + 1} \quad (3.34)$$

e seja  $\mathbf{S}$  a matriz aleatória

$$\mathbf{S} = (s_{ij}) \in \mathbb{R}^{k \times k}, \quad (3.35)$$

onde cada elemento de  $S$  é definido por

$$s_{ij} = \begin{cases} 0 & \text{se } i < j \\ \sqrt{1 - \sum_{j^* < j} s_{ij^*}^2} & \text{se } i = j \\ w_{ij} \sqrt{1 - \sum_{j^* < j} s_{ij^*}^2} & \text{se } i > j \end{cases}.$$

Nesse contexto, é possível provar que as duas transformações que definem, respectivamente, o vetor  $V$  e a matriz  $S$  são biunívocas. Portanto, aplica-se o seguinte Jacobiano:

$$|\det J| = \prod_{i>j} \left| \sqrt{1 - \sum_{j^*} x_{ij^*}} \right|$$

Note-se que, para gerar uma matriz de correlação aleatória, basta multiplicar  $S$  por sua matriz transposta.

Analogamente, denota-se  $\Sigma^*$  o parâmetro  $\Sigma$  do modelo sem a restrição de fatores de Cholesky de matriz de correlação, obtido pelas transformações definidas na Proposição 7.

Vale ressaltar que é possível escolher distribuições a priori dos parâmetros sem restrição. Assim, não seria necessário calcular o Jacobiano da transformação. Entretanto, a desvantagem seria a perda da interpretação dos parâmetros, como por exemplo o parâmetro  $\Sigma$ , que representa a matriz de correlação da série temporal. Por isso, optou-se por manter as distribuições a priori dos parâmetros originais.

Conforme apresentado no Capítulo 2, para utilizar-se os métodos HMC e *Zero-Variance* na estimação dos parâmetros dos modelos GARCH estudados, é necessário ter o conhecimento da derivada do logaritmo da função densidade da distribuição posteriori. O Apêndice A apresenta as derivadas de cada modelo estudado.

Na seção a seguir, é apresentada a metodologia proposta para a identificação de possíveis pontos influentes em ambos os modelos.

### 3.4 Análise de Pontos Influentes

Uma observação é considerada influente se ela modifica, de forma substancial, alguma propriedade do modelo ajustado, como por exemplo as estimativas dos parâmetros. Portanto, é

importante utilizar medidas que sejam capazes de mensurar essa influência nos pontos da base de dados. Conforme mencionado no Capítulo 2, toda inferência bayesiana requer o uso de uma distribuição posteriori. Logo, deve-se escolher alguma medida que mensure a “distância” entre duas distribuições.

A divergência de *Kullback-Leibler* (KL) (KULLBACK; LEIBLER, 1951) é uma medida (embora não métrica) comumente utilizada para obter-se a diferença entre a entropia de duas distribuições de probabilidade. No contexto desta tese, a divergência de KL é utilizada para medir a influência da  $k$ -ésima observação sobre a distribuição posteriori.

**Definição 4.** Sejam  $p(x)$  e  $q(x)$  duas funções de densidade de probabilidade. Então, a divergência de Kullback-Leibler é calculada por

$$KL(p(x), q(x)) = \int p(x) \log \left( \frac{p(x)}{q(x)} \right) \partial x. \quad (3.36)$$

Para medir a influência da observação  $\mathbf{X}_l$  sobre a distribuição posteriori, a ideia é aplicar a divergência de KL entre as distribuições posteriori com e sem tal observação, denotadas respectivamente por  $\pi(\theta | \mathbf{X})$  e  $\pi(\theta | \mathbf{X}_{-l})$ . A distribuição sem a observação é dada por

$$\pi(\theta | \mathbf{X}_{-l}) \propto \pi(\mathbf{X}_{-l} | \theta) \pi(\theta), \quad (3.37)$$

onde

1. Para o modelo GJR-GARCH(1,1):

- $\pi(\mathbf{X}_{-l} | \theta) = \prod_{t=2, t \neq l}^n h_{t,(l)}^{-1/2} f_{\varepsilon_t} \left( \frac{\mathbf{X}_t - \mu}{\sqrt{h_{t,(l)}}} \right);$
- $h_{t,(l)} = \omega + \alpha y_{t-1}^2 + \beta h_{t-1,(l)} + \phi I_{t-1} y_{t-i}^2$ , se  $t \neq k + 1$ ;
- $h_{t,(l)} = \omega + (\alpha + \beta + 0,5\phi) h_{t-2,(l)}$ , se  $t = k + 1$ .

2. Para o modelo DCC-GJR-GARCH(1,1):

- $L_c(\theta) = \prod_{t=2, t \neq k}^n \mathbf{H}_{t,(l)}^{-1/2} f_{\varepsilon_t} \left( \mathbf{H}_{t,(l)}^{-1/2} (\mathbf{y}_t - \mu) \right);$
- $\mathbf{H}_{t,(l)} = \mathbf{D}_{t,(l)} \mathbf{R}_{t,(l)} \mathbf{D}_{t,(l)}$ , se  $t \neq l$ ;
- $\mathbf{D}_{t,(l)} = \text{diag}(h_{11,t,(l)}^{1/2}, \dots, h_{kk,t,(l)}^{1/2})$ , se  $t \neq l$ ;
- $\mathbf{R}_{t,(l)} = \text{diag}(\mathbf{Q}_{t,(l)})^{-1/2} \mathbf{Q}_{t,(l)} \text{diag}(\mathbf{Q}_{t,(l)})^{-1/2}$ , se  $t \neq l$ ;
- $h_{ii,t,(l)} = \omega + \alpha (x_{i,t-1} - \mu)^2 + \beta h_{i,t-1,(l)}$ , se  $t \neq k + 1$ ;
- $h_{ii,t,(l)} = \omega + (\alpha + \beta + 0,5\phi) h_{ii,t-2,(l)}$ , se  $t = k + 1$ ;
- $\mathbf{Q}_{t,(l)} = (1 - \lambda_1 - \lambda_2 - 0,5\lambda_3) \Sigma + \lambda_1 \mathbf{u}_{t-1} \mathbf{u}_{t-1}^T + \lambda_2 \mathbf{Q}_{t-1,(l)} + \lambda_3 \mathbf{v}_{t-1} \mathbf{v}_{t-1}^T$ , se  $t \neq l$ ;
- $\mathbf{Q}_{t,(l)} = (1 - \lambda_1 - \lambda_2 - 0,5\lambda_3) \Sigma + (\lambda_1 + \lambda_2 + 0,5\lambda_3) \mathbf{Q}_{t-2,(l)}$ , se  $t = l$ .

Assim, pode-se definir a influência de uma observação da seguinte maneira:

**Definição 5.** Seja  $KLD(l)$  a divergência de KL entre  $\pi(\theta | \mathbf{X})$  e  $\pi(\theta | \mathbf{X}_{-l})$ . Então,

$$KLD(l) = \int \pi(\theta | \mathbf{X}) \log \left( \frac{\pi(\theta | \mathbf{X})}{\pi(\theta | \mathbf{X}_{-l})} \right) \partial \theta = \mathbb{E}_{\theta | \mathbf{X}} \left( \log \left( \frac{\pi(\theta | \mathbf{X})}{\pi(\theta | \mathbf{X}_{-l})} \right) \right). \quad (3.38)$$

De acordo com a Definição 5, a observação mais influente é aquela que possui o maior valor de  $KLD(\cdot)$ . Entretanto, no caso dos modelos estudados, não é possível encontrar os reais valores de  $KLD(\cdot)$  por meio dessa definição, devido ao fato de que não há conhecimento sobre a distribuição posteriori. Porém, existe outra maneira de se calcular o valor de  $KLD(\cdot)$ , conforme apresenta a proposição a seguir.

**Proposição 8.** O valor  $KLD(l)$  pode ser calculado pela seguinte equação:

$$KLD(l) = \log \left( \mathbb{E}_{\theta | \mathbf{X}} \left( \frac{\pi(\mathbf{X}_{-k} | \theta)}{\pi(\mathbf{X} | \theta)} \right) \right) + \mathbb{E}_{\theta | \mathbf{X}} \left( \log \left( \frac{\pi(\mathbf{X} | \theta)}{\pi(\mathbf{X}_{-k} | \theta)} \right) \right) \quad (3.39)$$

*Demonstração.*

$$\begin{aligned} KLD(l) &= \int \pi(\theta | \mathbf{X}) \log \left( \frac{\pi(\theta | \mathbf{X})}{\pi(\theta | \mathbf{X}_{-k})} \right) \partial \theta \\ &= \int \pi(\theta | \mathbf{X}) \log \left( \frac{C_k \pi(\mathbf{X} | \theta) \pi(\theta)}{C \pi(\mathbf{X}_{-k} | \theta) \pi(\theta)} \right) \partial \theta \\ &= \int \pi(\theta | \mathbf{X}) \log \left( \frac{C_k}{C} \right) \partial \theta + \int \pi(\theta | \mathbf{X}) \log \left( \frac{\pi(\mathbf{X} | \theta)}{\pi(\mathbf{X}_{-k} | \theta)} \right) \partial \theta \\ &= \log \left( \frac{C_k}{C} \right) + \mathbb{E}_{\theta | \mathbf{X}} \left( \log \left( \frac{\pi(\mathbf{X} | \theta)}{\pi(\mathbf{X}_{-k} | \theta)} \right) \right) \\ &= \log \left( \int \frac{\pi(\mathbf{X}_{-k} | \theta) \pi(\theta)}{C} \partial \theta \right) + \mathbb{E}_{\theta | \mathbf{X}} \left( \log \left( \frac{\pi(\mathbf{X} | \theta)}{\pi(\mathbf{X}_{-k} | \theta)} \right) \right) \\ &= \log \left( \int \frac{\pi(\mathbf{X}_{-k} | \theta)}{\pi(\mathbf{X} | \theta)} \frac{\pi(\mathbf{X} | \theta) \pi(\theta)}{C} \partial \theta \right) + \mathbb{E}_{\theta | \mathbf{X}} \left( \log \left( \frac{\pi(\mathbf{X} | \theta)}{\pi(\mathbf{X}_{-k} | \theta)} \right) \right) \\ &= \log \left( \int \frac{\pi(\mathbf{X}_{-k} | \theta)}{\pi(\mathbf{X} | \theta)} \pi(\theta | \mathbf{X}) \partial \theta \right) + \mathbb{E}_{\theta | \mathbf{X}} \left( \log \left( \frac{\pi(\mathbf{X} | \theta)}{\pi(\mathbf{X}_{-k} | \theta)} \right) \right) \\ &= \log \left( \mathbb{E}_{\theta | \mathbf{X}} \left( \frac{\pi(\mathbf{X}_{-k} | \theta)}{\pi(\mathbf{X} | \theta)} \right) \right) + \mathbb{E}_{\theta | \mathbf{X}} \left( \log \left( \frac{\pi(\mathbf{X} | \theta)}{\pi(\mathbf{X}_{-k} | \theta)} \right) \right). \end{aligned}$$

□

Conforme a Proposição 8, o cálculo da divergência de KL depende da esperança da log-verossimilhança do modelo, cujo valor não pode ser diretamente obtido. Portanto, é necessário o uso de métodos computacionais para encontrá-lo. Nesta tese, empregou-se o estimador HMC de  $KLD(\cdot)$ .

---

## SIMULAÇÃO

---

Este capítulo apresenta os estudos de simulação para verificar o desempenho do método *Zero-Variance* com estimadores MCMC. Foram realizados dois estudos:

1. Caso univariado com o modelo GJR-GARCH univariado. Este estudo busca avaliar o método *Zero-Variance* com as funções experimentais linear e quadrática, focando na **eficácia** do método em termos da diminuição da variância. Especificamente, foram empregados dois métodos MCMC: **(i)** o método proposto HMC e **(ii)** o método competidor *Metropolis-Hastings*.
2. Caso multivariado com o modelo DCC-GJR-GARCH. Este estudo busca avaliar o método *Zero-Variance* com ambas as funções experimentais do estudo anterior, focando na **eficiência computacional** do método, isto é, seu tempo de execução. Especificamente, foi empregado o método HMC.

Para gerar os conjuntos de dados sintéticos, bem como para simular as distribuições posteriori e executar os algoritmos de estimação do método *Zero-Variance*, foram utilizadas as linguagens de programação R e C++. Particularmente, foram empregados os pacotes Rcpp<sup>1</sup> e RcppArmadillo<sup>2</sup> (EDDELBUETTEL *et al.*, 2011; EDDELBUETTEL, 2013), responsáveis por interligar as linguagens de programação R e C++. Isto é, torna-se possível escrever programas em C++ para serem invocados em uma aplicação R. A principal vantagem da integração dessas duas linguagens é o ganho em desempenho computacional se comparado a um programa escrito puramente em R. A segunda vantagem é a possibilidade de utilizar funções e pacotes de R ao mesmo tempo em que se utiliza funções de C++, porém mantendo-se a sintaxe da programação em R. Por fim, os programas resultantes podem ser executados em ambas as linguagens.

<sup>1</sup> <<https://cran.r-project.org/web/packages/Rcpp/index.html>>

<sup>2</sup> <<https://cran.r-project.org/web/packages/RcppArmadillo/index.html>>

## 4.1 Caso Univariado

Para geração dos conjuntos de dados sintéticos, foram considerados os modelos GJR-GARCH com distribuição de erros normal,  $t$  e GED. Para cada modelo, foram gerados conjuntos de dados de tamanho 200, 500, 1000 e 2000 com parâmetros:  $\omega = 0,05$ ,  $\alpha = 0,05$ ,  $\phi = 0,1$ ,  $\beta = 0,85$ ,  $\mu = 0$ . Quanto aos parâmetros de forma da distribuição dos resíduos, foram utilizados  $\nu = 8$  para  $\varepsilon_t \sim T$  e  $\nu = 0,8$  para  $\varepsilon_t \sim GED$ . A escolha desses valores foi baseada na literatura.

Para o algoritmo *Metropolis-Hastings*, utilizou-se como distribuição proposta a normal multivariada  $N(\theta', \varepsilon M)$ , ou seja, uma proposta simétrica, onde  $\theta'$  é o estado atual da cadeia,  $M$  é uma matriz positiva definida e  $\varepsilon$  é um número real que está relacionado à probabilidade de aceitação da cadeia. Neste caso, escolheu-se um valor flexível de maneira a garantir que a probabilidade de aceitação fosse próxima de 0,8. De acordo com a literatura, o valor ideal de  $M$  é a inversa da matriz de covariância da posteriori. Como não se tem essa informação previamente, o valor de  $M$  foi estimado através da matriz de variância e covariância amostral de uma cadeia de tamanho 2000, gerada pelo algoritmo *Metropolis-Hastings* com  $M = I_k$ . O Algoritmo 2 descreve o método *Metropolis-Hastings*.

---

### Algoritmo 2 – Metropolis-Hastings

---

Dados  $\theta^0$ ,  $\varepsilon$ ,  $\pi$  e  $S$

**Para**  $i = 1$  a 2000 **faça**

$\theta^i \leftarrow \theta^{i-1}$

Simule  $\tilde{\theta} \sim N(\theta^{i-1}, \varepsilon I_k)$

Simule  $U \sim Uniforme(0, 1)$

**Se**  $(\log(\pi(\tilde{\theta})) - \log(\pi(\theta^{i-1}))) > \log(u)$  **faça**

$\theta^i \leftarrow \tilde{\theta}$

**Fim do para**

Descarte os 1000 primeiros elementos da cadeia

Calcule  $M = cov(\theta)$

**Para**  $i = 1$  a  $S$  **faça**

$\theta^i \leftarrow \theta^{i-1}$

Simule  $\tilde{\theta} \sim N(\theta^{i-1}, \varepsilon M)$

Simule  $U \sim Uniforme(0, 1)$ ;

**Se**  $(\log(\pi(\tilde{\theta})) - \log(\pi(\theta^{i-1}))) > \log(u)$  **faça**

$\theta^i \leftarrow \tilde{\theta}$

**Fim do para**

**Retorne**  $\theta$

---

Para o algoritmo HMC, utilizou-se os seguintes parâmetros:  $M = I_k$ ,  $LF = 20$  e  $\varepsilon$  livre, de modo a garantir que a probabilidade de aceitação ficasse entre 0,7 e 0,9.

Para cada algoritmo, foram geradas duas cadeias de tamanho 2000, das quais foram descartadas as 1000 primeiras iterações como amostra de aquecimento. O restante da primeira cadeia foi utilizada para estimar os parâmetros do método *Zero-Variance*, enquanto a segunda foi utilizada para estimar os parâmetros do modelo. Para a distribuição priori, assumiu-se que os

parâmetros são independentes e seguem uma distribuição normal com média 0 e variância 1000. Portanto, foi escolhida uma distribuição priori com pouca informação, ou seja, uma distribuição com variância alta.

As Tabelas 2, 3 e 4 apresentam os resultados das simulações de cada modelo proposto. Especificamente, as tabelas mostram a média do viés e do erro-padrão Monte Carlo das simulações, que é a raiz quadrada da variância do estimador MCMC (definida na Equação 2.13). Para os três modelos, os conjuntos de dados de tamanho 200 tiveram os maiores valores de viés, especialmente os estimadores do parâmetro  $\beta$ . Este era um fato esperado, pois uma série temporal de tamanho 200 é considerada pequena. Portanto, seria mais indicado, se possível, que a distribuição priori fosse mais bem informativa, ou seja, com variância pequena e com média em torno do parâmetro real. Em geral, notou-se que séries de tamanho 1000 ou maior possuem um viés baixo mesmo com uma distribuição priori pouco informativa, indicando que a escolha da distribuição priori não influencia a estimativa dos parâmetros para séries desses tamanhos.

Nos três modelos, o método *Zero-Variance* teve um desempenho melhor no algoritmo HMC, pois suas estimativas não diferem muito das estimativas do algoritmo HMC puro (sem o uso do método *Zero-Variance*), o que indica uma falta de viés. Porém, no caso do algoritmo *Metropolis-Hastings*, as estimativas diferiram bastante para certos modelos, como as estimativas do parâmetro  $\beta$  para o modelo GJR-GARCH(1,1)-GED ( $N = 1000$  e  $N = 2000$ ). Alguns problemas foram encontrados no método *Zero-Variance* para o algoritmo *Metropolis-Hastings*, principalmente para calcular o parâmetro  $\mathbf{a}$  da Equação 2.16. Um dos problemas foi que, como a cadeia gerada possui uma alta dependência, o determinante de  $\sigma_g^2$  ficou próximo de 0, o que dificulta computacionalmente o cálculo da matriz inversa. Outro problema foi o aumento no viés da estimativa. Portanto, nem sempre o método *Metropolis-Hastings* é indicado para reduzir a variância da estimativa em cadeias com alta dependência.

Essencialmente, as simulações mostram que, em geral, os estimadores do método *Zero-Variance* são muito mais eficazes se comparados a estimadores MCMC puros (sem a aplicação do método *Zero-Variance*), principalmente quando a função experimental é quadrática. Essa maior eficácia ocorre porque os estimadores do método *Zero-Variance* reduzem drasticamente o erro-padrão, indicando que não é necessário ter grandes cadeias para obter-se um erro-padrão pequeno, o que proporciona economia computacional. Por exemplo, o erro-padrão do parâmetro  $\omega$  do modelo GJR-GARCH(1,1)( $N = 1000$ ) diminuiu de 0,00015 (HMC) para 0,00001 (ZV-HMC-Q). Isto é, o erro-padrão do estimador HMC foi 15 vezes maior do que o erro-padrão do estimador ZV-HMC-Q. Em outras palavras, para gerar uma estimativa com o mesmo erro-padrão sem a utilização do método *Zero-Variance*, seria necessário ter uma cadeia de tamanho 450.000 (sendo 1000 elementos usados como amostra de aquecimento).

Tabela 2 – Resultado das simulações do modelo GJR-GARCH(1,1)-NORMAL.

N	Par	HMC		ZV-HMC-L		ZV-HMC-Q		RWM		ZV-RWM-L		ZV-RWM-Q	
		Viés	EP	Viés	EP	Viés	EP	Viés	EP	Viés	EP	Viés	EP
200	$\omega$	0,04480	0,00170	0,04480	0,00147	0,04475	0,00078	0,04479	0,00072	0,03396	0,00053	0,03684	0,00043
	$\alpha$	0,04973	0,00459	0,05019	0,00299	0,05041	0,00132	0,04972	0,00173	0,04331	0,00170	0,06101	0,00152
	$\phi$	0,10084	0,00754	0,10072	0,00465	0,10019	0,00200	0,10084	0,00274	0,10723	0,00247	0,09675	0,00219
	$\beta$	-0,32998	0,01156	-0,33024	0,00963	-0,33014	0,00479	-0,32996	0,00347	-0,27652	0,00315	-0,29693	0,00239
	$\mu$	-0,00459	0,00084	-0,00463	0,00024	-0,00463	0,00013	-0,00459	0,00182	-0,00565	0,00052	-0,00832	0,00055
500	$\omega$	0,01149	0,00071	0,01148	0,00051	0,01149	0,00026	0,01149	0,00020	0,00886	0,00017	0,01108	0,00009
	$\alpha$	0,01750	0,00347	0,01747	0,00235	0,01740	0,00074	0,01749	0,00100	0,01103	0,00077	0,01358	0,00053
	$\phi$	0,03600	0,00529	0,03612	0,00354	0,03607	0,00112	0,03600	0,00120	0,04505	0,00118	0,04486	0,00070
	$\beta$	-0,09709	0,00633	-0,09695	0,00398	-0,09698	0,00192	-0,09707	0,00190	-0,08131	0,00126	-0,09674	0,00054
	$\mu$	-0,00082	0,00054	-0,00077	0,00018	-0,00077	0,00007	-0,00081	0,00119	-0,00175	0,00019	0,00202	0,00017
1000	$\omega$	0,00191	0,00015	0,00192	0,00003	0,00192	0,00001	0,00191	0,00006	0,00175	0,00004	0,00206	0,00001
	$\alpha$	0,00712	0,00167	0,00717	0,00087	0,00707	0,00019	0,00712	0,00045	0,00346	0,00024	0,00449	0,00006
	$\phi$	0,00457	0,00274	0,00457	0,00147	0,00474	0,00038	0,00459	0,00050	0,01044	0,00038	0,00937	0,00009
	$\beta$	-0,01863	0,00172	-0,01873	0,00032	-0,01871	0,00008	-0,01863	0,00084	-0,01761	0,00031	-0,01958	0,00004
	$\mu$	-0,00082	0,00022	-0,00081	0,00008	-0,00082	0,00003	-0,00083	0,00060	-0,00130	0,00005	-0,00003	0,00002
2000	$\omega$	0,00191	0,00015	0,00192	0,00003	0,00192	0,00001	0,00191	0,00006	0,00175	0,00004	0,00206	<0,00001
	$\alpha$	0,00712	0,00167	0,00717	0,00087	0,00707	0,00019	0,00712	0,00045	0,00346	0,00024	0,00449	0,00006
	$\phi$	0,00457	0,00274	0,00457	0,00147	0,00474	0,00038	0,00459	0,00050	0,01044	0,00038	0,00937	0,00009
	$\beta$	-0,01863	0,00172	-0,01873	0,00032	-0,01871	0,00008	-0,01863	0,00084	-0,01761	0,00031	-0,01958	0,00004
	$\mu$	-0,00082	0,00022	-0,00081	0,00008	-0,00082	0,00003	-0,00083	0,00060	-0,00130	0,00005	-0,00003	0,00002

Fonte: Elaborada pelo autor.



Tabela 3 – Resultado das simulações do modelo GJR-GARCH(1,1)-t

N	Par	HMC		ZV-HMC-L		ZV-HMC-Q		RWM		ZV-RWM-L		ZV-RWM-Q	
		Viés	EP	Viés	EP	Viés	EP	Viés	EP	Viés	EP	Viés	EP
200	$\omega$	0,05075	0,00319	0,05271	0,00283	0,05523	0,00224	0,05065	0,00069	0,03266	0,00054	0,03650	0,00048
	$\alpha$	0,06404	0,00728	0,06571	0,00499	0,06780	0,00250	0,06399	0,00182	0,04719	0,00206	0,05936	0,00217
	$\phi$	0,13669	0,01246	0,13821	0,00804	0,13943	0,00408	0,13663	0,00283	0,13325	0,00401	0,14682	0,00419
	$\beta$	-0,35080	0,01447	-0,35216	0,01163	-0,35331	0,00578	-0,35016	0,00325	-0,28458	0,00344	-0,29983	0,00238
	$\mu$	-0,00548	0,00079	-0,00540	0,00025	-0,00534	0,00014	-0,00552	0,00163	-0,00742	0,00058	0,00058	0,00075
	$\nu$	1,65575	0,18313	1,16832	0,07342	1,07581	0,04158	1,65374	0,11214	1,28561	0,07894	1,32999	0,04894
500	$\omega$	0,01524	0,00095	0,01542	0,00073	0,01553	0,00041	0,01523	0,00020	0,01100	0,00019	0,01219	0,00010
	$\alpha$	0,02405	0,00446	0,02455	0,00299	0,02487	0,00109	0,02403	0,00100	0,01584	0,00083	0,01767	0,00068
	$\phi$	0,04640	0,00690	0,04706	0,00461	0,04720	0,00158	0,04648	0,00142	0,05328	0,00129	0,06179	0,00094
	$\beta$	-0,13184	0,00855	-0,13248	0,00603	-0,13315	0,00320	-0,13181	0,00198	-0,10612	0,00145	-0,11453	0,00068
	$\mu$	-0,00146	0,00058	-0,00146	0,00014	-0,00146	0,00006	-0,00147	0,00100	-0,00221	0,00018	0,00351	0,00019
	$\nu$	1,80807	0,14842	1,52539	0,04580	1,51592	0,01736	1,80790	0,07274	1,12321	0,04223	1,69395	0,01867
1000	$\omega$	0,00760	0,00062	0,00770	0,00038	0,00778	0,00020	0,00760	0,00013	0,00645	0,00011	0,00811	0,00003
	$\alpha$	0,01633	0,00264	0,01671	0,00180	0,01679	0,00058	0,01627	0,00044	0,01473	0,00037	0,01745	0,00020
	$\phi$	0,01324	0,00376	0,01380	0,00255	0,01385	0,00085	0,01334	0,00078	0,01686	0,00061	0,01118	0,00028
	$\beta$	-0,06116	0,00509	-0,06170	0,00283	-0,06222	0,00142	-0,06108	0,00119	-0,05559	0,00073	-0,06375	0,00019
	$\mu$	-0,00028	0,00041	-0,00031	0,00008	-0,00031	0,00004	-0,00030	0,00081	-0,00047	0,00011	-0,00198	0,00006
	$\nu$	1,27627	0,10997	1,10137	0,03210	1,09932	0,00850	1,27453	0,05149	0,60261	0,02846	1,28191	0,00547
2000	$\omega$	0,00277	0,00018	0,00275	0,00006	0,00276	0,00002	0,00277	0,00007	0,00247	0,00005	0,00319	0,00001
	$\alpha$	0,00552	0,00135	0,00541	0,00083	0,00549	0,00018	0,00553	0,00028	0,00536	0,00021	0,00573	0,00008
	$\phi$	0,01045	0,00192	0,01028	0,00114	0,01016	0,00026	0,01043	0,00054	0,01017	0,00033	0,01175	0,00012
	$\beta$	-0,02320	0,00175	-0,02296	0,00052	-0,02300	0,00016	-0,02320	0,00078	-0,02265	0,00034	-0,02671	0,00006
	$\mu$	0,00009	0,00025	0,00009	0,00004	0,00009	0,00001	0,00010	0,00060	0,00005	0,00005	-0,00001	0,00002
	$\nu$	0,62568	0,05306	0,63454	0,01529	0,63499	0,00346	0,62569	0,03418	0,37650	0,01391	0,65125	0,00174

Fonte: Elaborada pelo autor.

Tabela 4 – Resultado das simulações do modelo GJR-GARCH(1,1)-GED.

N	Par	HMC		ZV-HMC-L		ZV-HMC-Q		RWM		ZV-RWM-L		ZV-RWM-Q	
		Viés	EP	Viés	EP	Viés	EP	Viés	EP	Viés	EP	Viés	EP
200	$\omega$	0,04351	0,00218	0,04319	0,00185	0,04293	0,00092	0,04351	0,00131	0,04280	0,00109	0,04838	0,00086
	$\alpha$	0,05606	0,00638	0,05583	0,00404	0,05592	0,00177	0,05619	0,00291	0,06365	0,00357	0,10990	0,00262
	$\phi$	0,12018	0,01058	0,12024	0,00655	0,12059	0,00282	0,12028	0,00453	0,19138	0,00608	0,26531	0,00529
	$\beta$	-0,32138	0,01459	-0,32042	0,01172	-0,31849	0,00549	-0,32151	0,00447	-0,32651	0,00439	-0,35433	0,00355
	$\mu$	-0,00536	0,00080	-0,00562	0,00027	-0,00555	0,00015	-0,00537	0,00164	-0,00926	0,00101	-0,00961	0,00241
	$\nu$	0,03255	0,00342	0,03285	0,00230	0,03280	0,00104	0,03268	0,02298	0,04755	0,00951	-0,02442	0,00745
500	$\omega$	0,01319	0,00074	0,01320	0,00053	0,01305	0,00029	0,01326	0,00045	0,02150	0,00039	0,02293	0,00022
	$\alpha$	0,02575	0,00375	0,02487	0,00256	0,02710	0,00084	0,02670	0,00175	0,05007	0,00152	0,08252	0,00100
	$\phi$	0,04456	0,00611	0,04423	0,00411	0,04458	0,00129	0,04439	0,00238	0,08529	0,00315	0,13963	0,00185
	$\beta$	-0,11438	0,00666	-0,11400	0,00423	-0,11380	0,00217	-0,11519	0,00284	-0,18299	0,00220	-0,20120	0,00095
	$\mu$	-0,00096	0,00059	-0,00093	0,00016	-0,00084	0,00006	-0,00095	0,00096	-0,00005	0,00034	-0,02020	0,00058
	$\nu$	0,01762	0,00283	0,01768	0,00073	0,01757	0,00031	0,01742	0,01429	-0,02632	0,00322	-0,07287	0,00154
1000	$\omega$	0,00508	0,00060	0,00515	0,00038	0,00509	0,00021	0,00508	0,00034	0,01486	0,00020	0,01842	0,00007
	$\alpha$	0,01343	0,00247	0,01328	0,00176	0,01322	0,00048	0,01343	0,00087	0,03267	0,00059	0,04156	0,00019
	$\phi$	0,01498	0,00362	0,01552	0,00244	0,01419	0,00079	0,01498	0,00156	0,03621	0,00087	0,05068	0,00037
	$\beta$	-0,04750	0,00482	-0,04640	0,00277	-0,04648	0,00144	-0,04750	0,00177	-0,11817	0,00119	-0,14952	0,00037
	$\mu$	0,00074	0,00039	0,00068	0,00010	0,00070	0,00004	0,00074	0,00077	0,00075	0,00013	-0,01364	0,00020
	$\nu$	0,00593	0,00078	0,00557	0,00033	0,00554	0,00012	0,00593	0,01004	-0,02630	0,00145	-0,02943	0,00041
2000	$\omega$	0,00241	0,00019	0,00237	0,00006	0,00239	0,00007	0,00241	0,00016	0,00449	0,00009	0,00597	0,00002
	$\alpha$	0,00694	0,00164	0,00682	0,00110	0,00614	0,00053	0,00697	0,00054	0,01664	0,00032	0,02207	0,00008
	$\phi$	0,00729	0,00171	0,00647	0,00102	0,00626	0,00034	0,00729	0,00109	0,01575	0,00049	0,02400	0,00013
	$\beta$	-0,02141	0,00174	-0,02144	0,00050	-0,02164	0,00053	-0,02150	0,00111	-0,04658	0,00054	-0,05325	0,00010
	$\mu$	-0,00073	0,00022	-0,00063	0,00004	-0,00066	0,00002	-0,00073	0,00054	-0,00044	0,00006	-0,00992	0,00007
	$\nu$	0,00360	0,00186	0,00273	0,00016	0,00271	0,00011	0,00360	0,00686	-0,02111	0,00070	-0,01191	0,00013

Fonte: Elaborada pelo autor.

## 4.2 Caso Multivariado

Conforme mencionado no início do capítulo, este estudo de simulação tem como principal interesse medir a eficiência computacional do método *Zero-Variance*. Para isso, foi utilizado um critério baseado na estimativa *Effective Sample Size* (ESS), que é uma estimativa do tamanho amostral requerido para que a amostra gerada tenha o mesmo nível de precisão (variância) de uma amostra aleatória simples (para mais detalhes, ver [Kass et al. \(1998\)](#)). Matematicamente, a estimativa ESS é definida por

$$ESS = \frac{M}{1 + 2 \sum_{k=1}^j \rho(k)}, \quad (4.1)$$

onde  $M$  representa o tamanho da amostra gerada,  $\rho(\cdot)$  é a função de autocorrelação e  $j$  é um número suficientemente grande, porém menor do que  $M$ . Além disso, considerou-se um critério de tempo de execução, baseado no valor  $ESS_t = ESS \div t_{exec}$ , onde  $t_{exec}$  é o tempo de execução demandado para gerar a amostra.

Para a geração do conjunto de dados sintéticos, foram considerados os modelos DCC-GJR-GARCH de dimensão 3, com distribuições de resíduos normal,  $t$  e GED. Para cada modelo, foram geradas 500 amostras de tamanho 1500 com parâmetros:

- $\boldsymbol{\omega} = (0,03, 0,03, 0,03)^T$ ,
- $\boldsymbol{\alpha} = (0,05, 0,05, 0,05)^T$ ,
- $\boldsymbol{\beta} = (0,85, 0,85, 0,85)^T$ ,
- $\boldsymbol{\phi} = (0,05, 0,05, 0,05)^T$ ,
- $\boldsymbol{\mu} = (0, 0, 0)^T$ ,
- $\boldsymbol{\lambda} = (0,05, 0,7)^T$  e
- $\Sigma = \begin{bmatrix} 1 & 0,5 & 0,25 \\ 0,5 & 1 & 0,5 \\ 0,25 & 0,5 & 1 \end{bmatrix}$ .

Quanto aos parâmetros de forma das distribuições de resíduos,  $\nu = 6$  para  $\varepsilon_t \sim T$  e  $\nu = 0,5$  para  $\varepsilon_t \sim GED$ . A escolha de todos esses parâmetros foi baseada na literatura, principalmente em casos univariados para os parâmetros  $\boldsymbol{\omega}$ ,  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\beta}$  e  $\boldsymbol{\phi}$ .

Diferentemente do caso univariado, este estudo empregou apenas o algoritmo HMC. Os parâmetros escolhidos foram os seguintes:  $M = I_k$ ,  $LF = 20$  e  $\varepsilon$  livre, de modo a garantir que a probabilidade de aceitação estivesse entre 0,7 e 0,9, pois assim garante-se a eficiência do algoritmo. Foi gerada uma cadeia de tamanho 2000, cujas 1000 primeiras iterações foram geradas como amostra de aquecimento e também utilizadas para estimar os parâmetros do método

*Zero-Variance*. O fato de não se ter descartado as 1000 primeiras iterações trouxe um importante benefício, que foi a economia de tempo computacional para estimar os parâmetros do método *Zero-Variance*. Para a distribuição priori, assumiu-se que os parâmetros são independentes e seguem estas distribuições:

$$\begin{aligned}
 \omega_i &\sim \text{Gama}(1,1), \forall i = 1,2,3 \\
 \alpha_i, \beta_i, \phi_i &\sim \text{Uniforme}(A), \forall i = 1,2,3 \\
 \lambda_i &\sim \text{Uniforme}(B), \forall i = 1,2 \\
 \mu_i &\sim \text{Normal}(0,1), \forall i = 1,2,3 \\
 \Sigma &\sim \text{LKJ corr}(2),
 \end{aligned} \tag{4.2}$$

onde

$$\begin{aligned}
 A = \{(\alpha, \beta, \phi) \mid \alpha + \beta + 0,5\phi < 1, \\
 \alpha > 0, \qquad \qquad \qquad \beta > 0, \phi > -\beta\},
 \end{aligned}$$

$$\begin{aligned}
 B = \{(\lambda_1, \lambda_2) \mid \lambda_1 + \lambda_2 < 1, \\
 \lambda_1 > 0, \qquad \qquad \qquad \lambda_2 > 0\}.
 \end{aligned}$$

Adicionalmente, para o parâmetro de forma  $\nu$ , tem-se que:

$$\begin{aligned}
 \nu &\sim \text{Pareto}(4,3) \text{ para o modelo DCC-GJR-GARCH}(1,1)\text{-}t \\
 \nu &\sim \text{Gama}(2,2) \text{ para o modelo DCC-GJR-GARCH}(1,1)\text{-GED.}
 \end{aligned} \tag{4.3}$$

Na Equação 4.2, note-se a distribuição de Lewandowski-Kurowicka-Joe (LKJ), introduzida por Lewandowski, Kurowicka e Joe (2009) e geralmente usada como distribuição priori para matrizes de correlação.

As Tabelas 5-13 apresentam os resultados das simulações de cada modelo estudado. De modo geral, as médias das estimativas de cada parâmetro mostraram-se bem próximas aos valores reais, indicando que o estimador bayesiano é apropriado para os modelos em questão, mesmo quando a distribuição priori é pouco informativa. Note-se que este fato também foi observado no estudo univariado.

Em relação à eficiência computacional, cada simulação levou em média 2,8 minutos no caso do modelo com distribuição normal, 3,4 minutos no caso da distribuição  $t$  e 3,3 minutos no caso da distribuição GED. Considerando-se que todas as execuções levaram **minutos** para serem finalizadas, pode-se observar que as implementações em C++ tornaram as simulações factíveis, dados os curtos tempos de execução para estimações que são inerentemente onerosas. O desempenho computacional obtido foi bastante superior ao de implementações existentes no ambiente R, que levariam por volta de **horas** para realizarem a mesma estimacão em modelos

tridimensionais. Já com relação à qualidade das amostras, para a maioria dos parâmetros, a eficácia das amostras foi acima de 10%, indicando que o estimador HMC é bastante eficaz para a geração de amostras. Vale ressaltar que houve casos em que a eficácia foi maior do que 1000, indicando que a amostra gerada foi mais eficaz do que uma amostra aleatória simples. O método *Zero-Variance* mostrou-se bastante eficaz e eficiente para todas as simulações realizadas, pois diminuiu a variância dos estimadores em pelo menos uma ordem de magnitude, bem como o fez em torno de **segundos**.

Tabela 5 – Resultados do estimador HMC para o modelo DCC-GJR-GARCH(1,1)-NORMAL.

	Real	Média	Erro-padrão	Desvio-padrão	ESS	ESS <sub>t</sub>
$\omega_1$	0,03	0,07077	0,00461	0,02594	69,04330	24,28259
$\omega_2$	0,03	0,06447	0,00353	0,02243	78,08240	27,28250
$\omega_3$	0,03	0,06986	0,00452	0,02570	69,56141	24,16585
$\alpha_1$	0,05	0,06878	0,00249	0,02523	125,78370	44,13221
$\alpha_2$	0,05	0,06888	0,00212	0,02333	144,19707	50,38101
$\alpha_3$	0,05	0,06790	0,00253	0,02512	122,81465	43,04709
$\beta_1$	0,85	0,72410	0,01286	0,07657	73,78174	25,72471
$\beta_2$	0,85	0,74056	0,01007	0,06779	83,04683	29,04626
$\beta_3$	0,85	0,72712	0,01254	0,07595	73,78112	25,98893
$\phi_1$	0,05	0,06275	0,00213	0,03520	326,57762	114,45353
$\phi_2$	0,05	0,05969	0,00179	0,03219	367,51560	129,43834
$\phi_3$	0,05	0,06205	0,00209	0,03469	320,80128	112,65679
$\mu_1$	0	-0,00022	0,00031	0,01561	2926,11561	1023,84911
$\mu_2$	0	-0,00021	0,00031	0,01556	2819,37771	972,12890
$\mu_3$	0	0,00041	0,00028	0,01563	3448,99482	1213,17576
$\lambda_1$	0,05	0,06010	0,00115	0,01516	200,71918	70,66440
$\lambda_2$	0,7	0,54716	0,02124	0,12896	59,51622	20,94290
$\Sigma_{12}$	0,5	0,49948	0,00150	0,02241	285,56419	99,78563
$\Sigma_{13}$	0,25	0,25067	0,00257	0,02789	147,31796	51,54978
$\Sigma_{23}$	0,5	0,50075	0,00068	0,02243	1273,17007	444,13380

Fonte: Elaborada pelo autor.

### 4.3 Considerações Finais

De modo geral, em ambos os casos univariado e multivariado, o método *Zero-Variance* combinado com o algoritmo HMC mostrou-se bastante apropriado para a estimação de parâmetros dos modelos GJR-GARCH estudados. Este trabalho de doutorado é o primeiro a analisar ambas as técnicas conjuntamente para essa classe de modelos e, com base nos resultados apresentados, pode-se ressaltar que os ganhos em eficiência foram bastante relevantes. Tais ganhos foram ainda maiores em modelos multivariados, em que reduziu-se os tempos de execução de horas para minutos no caso de modelos tridimensionais. Além disso, a metodologia adotada possibilita a estimação de parâmetros de modelos com maior dimensionalidade, ao mesmo tempo em que se mantém tempos de execução razoáveis.

Tabela 6 – Resultados do estimador ZV-HMC-L para o modelo DCC-GJR-GARCH(1,1)-NORMAL.

	<b>Real</b>	<b>Média</b>	<b>Erro-padrão</b>	<b>Desvio-padrão</b>	<b>ESS</b>	<b>ESS<sub>t</sub></b>
$\omega_1$	0,03	0,07092	0,00339	0,02012	174,71246	61,44222
$\omega_2$	0,03	0,06456	0,00233	0,01597	198,07873	69,20759
$\omega_3$	0,03	0,06987	0,00323	0,01968	176,26391	61,23323
$\alpha_1$	0,05	0,06878	0,00108	0,00983	196,93623	69,09461
$\alpha_2$	0,05	0,06891	0,00083	0,00846	216,64100	75,69043
$\alpha_3$	0,05	0,06784	0,00111	0,00987	190,58643	66,79945
$\beta_1$	0,85	0,72376	0,00940	0,05771	179,11199	62,44767
$\beta_2$	0,85	0,74025	0,00666	0,04690	201,30719	70,40707
$\beta_3$	0,85	0,72706	0,00897	0,05665	179,13985	63,09939
$\phi_1$	0,05	0,06278	0,00095	0,01026	272,43129	95,47499
$\phi_2$	0,05	0,05967	0,00072	0,00870	294,36531	103,67251
$\phi_3$	0,05	0,06225	0,00092	0,00998	256,00029	89,89834
$\mu_1$	0	-0,00023	0,00008	0,00161	538,08096	188,27038
$\mu_2$	0	-0,00021	0,00008	0,00168	570,62223	196,74755
$\mu_3$	0	0,00039	0,00008	0,00162	579,70202	203,90405
$\lambda_1$	0,05	0,06011	0,00029	0,00379	284,69485	100,22561
$\lambda_2$	0,7	0,54653	0,00909	0,06810	137,83050	48,49830
$\Sigma_{12}$	0,5	0,49945	0,00014	0,00268	449,23475	156,97393
$\Sigma_{13}$	0,25	0,25060	0,00018	0,00321	378,69385	132,50941
$\Sigma_{23}$	0,5	0,50076	0,00012	0,00283	584,06523	203,73994

Fonte: Elaborada pelo autor.

Tabela 7 – Resultados do estimador ZV-HMC-Q para o modelo DCC-GJR-GARCH(1,1)-NORMAL.

	<b>Real</b>	<b>Média</b>	<b>Erro-padrão</b>	<b>Desvio-padrão</b>	<b>ESS</b>	<b>ESS<sub>t</sub></b>
$\omega_1$	0,03	0,06756	0,00201	0,01526	265,38935	93,26569
$\omega_2$	0,03	0,06166	0,00132	0,01035	287,73384	100,46431
$\omega_3$	0,03	0,06657	0,00176	0,01389	262,17565	91,01706
$\alpha_1$	0,05	0,07317	0,00050	0,00510	266,46123	93,42462
$\alpha_2$	0,05	0,07345	0,00033	0,00378	289,03696	100,91509
$\alpha_3$	0,05	0,07224	0,00048	0,00499	273,20346	95,69102
$\beta_1$	0,85	0,72651	0,00536	0,04166	264,29821	92,08500
$\beta_2$	0,85	0,74195	0,00361	0,02857	284,20095	99,32866
$\beta_3$	0,85	0,72983	0,00469	0,03774	261,05783	91,88978
$\phi_1$	0,05	0,05977	0,00047	0,00546	346,20866	121,25226
$\phi_2$	0,05	0,05694	0,00034	0,00415	349,46325	122,99177
$\phi_3$	0,05	0,05928	0,00046	0,00520	328,57958	115,31027
$\mu_1$	0	0,00003	0,00004	0,00073	529,73602	185,22559
$\mu_2$	0	0,00011	0,00004	0,00072	537,42960	185,17233
$\mu_3$	0	0,00067	0,00004	0,00074	585,31247	205,74033
$\lambda_1$	0,05	0,06459	0,00013	0,00153	251,67613	88,54191
$\lambda_2$	0,7	0,58883	0,00410	0,03676	184,59120	64,90826
$\Sigma_{12}$	0,5	0,49978	0,00007	0,00118	483,84171	168,94578
$\Sigma_{13}$	0,25	0,25053	0,00009	0,00147	384,89083	134,58778
$\Sigma_{23}$	0,5	0,50153	0,00005	0,00116	698,67105	243,55134

Fonte: Elaborada pelo autor.

Tabela 8 – Resultados do estimador HMC para o modelo DCC-GJR-GARCH(1,1)-t.

	<b>Real</b>	<b>Média</b>	<b>Erro-padrão</b>	<b>Desvio-padrão</b>	<b>ESS</b>	<b>ESS<sub>t</sub></b>
$\omega_1$	0,03	0,06249	0,00320	0,02103	81,80007	26,84571
$\omega_2$	0,03	0,06436	0,00337	0,02160	81,71904	26,45588
$\omega_3$	0,03	0,06719	0,00417	0,02421	70,68269	22,95279
$\alpha_1$	0,05	0,07124	0,00253	0,02590	122,75917	40,10309
$\alpha_2$	0,05	0,07075	0,00242	0,02497	127,32106	41,81653
$\alpha_3$	0,05	0,07250	0,00280	0,02731	114,84327	37,48307
$\beta_1$	0,85	0,74297	0,00979	0,06636	85,88676	28,21920
$\beta_2$	0,85	0,73897	0,01003	0,06698	84,66643	27,84801
$\beta_3$	0,85	0,73017	0,01204	0,07401	75,15515	24,77019
$\phi_1$	0,05	0,06221	0,00248	0,03831	271,69822	89,02483
$\phi_2$	0,05	0,06316	0,00233	0,03668	288,06896	95,01996
$\phi_3$	0,05	0,06503	0,00274	0,04051	252,92507	83,13177
$\mu_1$	0	0,00056	0,00045	0,01388	1358,34353	447,29961
$\mu_2$	0	0,00004	0,00050	0,01395	1212,57613	400,15648
$\mu_3$	0	-0,00081	0,00045	0,01400	1345,13320	441,13645
$\lambda_1$	0,05	0,06217	0,00149	0,01752	156,40946	51,51411
$\lambda_2$	0,7	0,55013	0,02487	0,13632	51,09476	16,77681
$\Sigma_{12}$	0,5	0,49858	0,00118	0,02443	548,35942	179,59502
$\Sigma_{13}$	0,25	0,25006	0,00184	0,03019	330,16171	108,02942
$\Sigma_{23}$	0,5	0,49998	0,00127	0,02437	488,46721	158,67131
$\nu$	6	6,48631	0,03518	0,60332	301,09931	98,09866

Fonte: Elaborada pelo autor.

Tabela 9 – Resultados do estimador ZV-HMC-L para o modelo DCC-GJR-GARCH(1,1)-t.

	<b>Real</b>	<b>Média</b>	<b>Erro-padrão</b>	<b>Desvio-padrão</b>	<b>ESS</b>	<b>ESS<sub>t</sub></b>
$\omega_1$	0,03	0,06238	0,00194	0,01380	213,97816	70,22060
$\omega_2$	0,03	0,06434	0,00212	0,01469	202,80761	65,65592
$\omega_3$	0,03	0,06711	0,00275	0,01726	192,68609	62,56790
$\alpha_1$	0,05	0,07122	0,00086	0,00934	231,86945	75,74529
$\alpha_2$	0,05	0,07085	0,00089	0,00915	221,79360	72,84290
$\alpha_3$	0,05	0,07245	0,00106	0,01034	217,93136	71,12750
$\beta_1$	0,85	0,74339	0,00592	0,04280	233,91355	76,85320
$\beta_2$	0,85	0,73900	0,00628	0,04464	216,81652	71,31214
$\beta_3$	0,85	0,73036	0,00796	0,05161	208,95000	68,86538
$\phi_1$	0,05	0,06198	0,00085	0,01064	297,11727	97,35100
$\phi_2$	0,05	0,06303	0,00089	0,01037	285,15655	94,05672
$\phi_3$	0,05	0,06513	0,00109	0,01210	267,84634	88,03370
$\mu_1$	0	0,00056	0,00005	0,00120	687,39720	226,35227
$\mu_2$	0	0,00000	0,00006	0,00128	644,44415	212,66412
$\mu_3$	0	-0,00080	0,00006	0,00126	657,72753	215,69588
$\lambda_1$	0,05	0,06216	0,00036	0,00440	248,98998	82,00365
$\lambda_2$	0,7	0,55019	0,01043	0,07317	128,56519	42,21284
$\Sigma_{12}$	0,5	0,49855	0,00014	0,00288	509,63644	166,90821
$\Sigma_{13}$	0,25	0,25005	0,00018	0,00340	420,99634	137,74766
$\Sigma_{23}$	0,5	0,50009	0,00017	0,00311	407,65926	132,41919
$\nu$	6	6,48619	0,00608	0,09633	311,71343	101,55400

Fonte: Elaborada pelo autor.

Tabela 10 – Resultados do estimador ZV-HMC-Q para o modelo DCC-GJR-GARCH(1,1)-t.

	<b>Real</b>	<b>Média</b>	<b>Erro-padrão</b>	<b>Desvio-padrão</b>	<b>ESS</b>	<b>ESS<sub>t</sub></b>
$\omega_1$	0,03	0,05820	0,00095	0,00907	298,36684	97,83841
$\omega_2$	0,03	0,05987	0,00118	0,01015	292,43106	94,59738
$\omega_3$	0,03	0,06228	0,00161	0,01243	280,27101	90,93631
$\alpha_1$	0,05	0,07247	0,00035	0,00440	283,78970	92,63627
$\alpha_2$	0,05	0,07316	0,00038	0,00430	279,40832	91,69853
$\alpha_3$	0,05	0,07394	0,00046	0,00510	283,44235	92,43849
$\beta_1$	0,85	0,74862	0,00280	0,02670	300,27767	98,58525
$\beta_2$	0,85	0,74394	0,00334	0,02919	296,35632	97,39973
$\beta_3$	0,85	0,73665	0,00447	0,03504	289,04270	95,18919
$\phi_1$	0,05	0,05776	0,00040	0,00520	343,70246	112,53257
$\phi_2$	0,05	0,05900	0,00042	0,00512	338,56759	111,59135
$\phi_3$	0,05	0,06061	0,00055	0,00635	329,26292	108,13456
$\mu_1$	0	0,00076	0,00003	0,00048	554,23047	182,36516
$\mu_2$	0	0,00022	0,00003	0,00052	525,05241	173,12665
$\mu_3$	0	-0,00058	0,00003	0,00054	549,39353	180,02740
$\lambda_1$	0,05	0,06624	0,00016	0,00193	252,03839	82,94360
$\lambda_2$	0,7	0,59160	0,00496	0,04071	175,65069	57,63034
$\Sigma_{12}$	0,5	0,49878	0,00009	0,00135	434,41101	142,16779
$\Sigma_{13}$	0,25	0,24987	0,00012	0,00164	375,11356	122,64502
$\Sigma_{23}$	0,5	0,50084	0,00007	0,00136	457,20287	148,40177
$\nu$	6	6,67058	0,00261	0,04736	454,06832	147,82014

Fonte: Elaborada pelo autor.

Tabela 11 – Resultados do estimador HMC para o modelo DCC-GJR-GARCH(1,1)-GED.

	<b>Real</b>	<b>Média</b>	<b>Erro-padrão</b>	<b>Desvio-padrão</b>	<b>ESS</b>	<b>ESS<sub>t</sub></b>
$\omega_1$	0,03	0,06700	0,00390	0,02305	78,38812	23,47795
$\omega_2$	0,03	0,06493	0,00366	0,02259	75,90861	23,01598
$\omega_3$	0,03	0,07232	0,00490	0,02667	67,22763	20,31827
$\alpha_1$	0,05	0,06935	0,00266	0,02581	119,07428	35,92663
$\alpha_2$	0,05	0,06919	0,00249	0,02499	122,50151	36,77078
$\alpha_3$	0,05	0,07466	0,00303	0,02875	108,98781	32,91450
$\beta_1$	0,85	0,73177	0,01150	0,07087	81,97739	24,68549
$\beta_2$	0,85	0,73699	0,01087	0,06989	79,26897	23,87845
$\beta_3$	0,85	0,71516	0,01391	0,08012	71,81906	21,66581
$\phi_1$	0,05	0,06461	0,00269	0,03912	255,77044	77,05610
$\phi_2$	0,05	0,06321	0,00243	0,03724	269,81591	81,40057
$\phi_3$	0,05	0,06630	0,00300	0,04271	239,88997	72,49534
$\mu_1$	0	-0,00032	0,00110	0,01278	244,12436	73,36918
$\mu_2$	0	-0,00029	0,00105	0,01278	294,44134	88,71568
$\mu_3$	0	0,00017	0,00110	0,01288	255,78777	77,75035
$\lambda_1$	0,05	0,06276	0,00156	0,01800	152,69835	46,02478
$\lambda_2$	0,7	0,54347	0,02534	0,13921	47,56434	14,33757
$\Sigma_{12}$	0,5	0,50031	0,00103	0,02491	695,74725	210,29856
$\Sigma_{13}$	0,25	0,24812	0,00165	0,03077	392,80465	118,69766
$\Sigma_{23}$	0,5	0,49835	0,00145	0,02492	378,14354	114,37895
$\nu$	0,5	0,47171	0,00088	0,01924	497,52875	150,50658

Fonte: Elaborada pelo autor.



Tabela 12 – Resultados do estimador ZV-HMC-L para o modelo DCC-GJR-GARCH(1,1)-GED.

	<b>Real</b>	<b>Média</b>	<b>Erro-padrão</b>	<b>Desvio-padrão</b>	<b>ESS</b>	<b>ESS<sub>t</sub></b>
$\omega_1$	0,03	0,06680	0,00259	0,01586	199,00797	59,60290
$\omega_2$	0,03	0,06493	0,00232	0,01568	187,85965	56,95856
$\omega_3$	0,03	0,07230	0,00341	0,01990	177,41028	53,61591
$\alpha_1$	0,05	0,06943	0,00102	0,00945	216,95721	65,45785
$\alpha_2$	0,05	0,06918	0,00096	0,00927	201,72743	60,55020
$\alpha_3$	0,05	0,07460	0,00125	0,01116	194,95582	58,87581
$\beta_1$	0,85	0,73231	0,00752	0,04789	215,41140	64,86393
$\beta_2$	0,85	0,73704	0,00689	0,04774	197,61199	59,52611
$\beta_3$	0,85	0,71522	0,00969	0,05865	191,74163	57,84165
$\phi_1$	0,05	0,06446	0,00109	0,01137	268,01722	80,74367
$\phi_2$	0,05	0,06304	0,00098	0,01085	246,23007	74,28312
$\phi_3$	0,05	0,06633	0,00126	0,01307	238,25281	71,99698
$\mu_1$	0	-0,00034	0,00005	0,00102	548,50482	164,84362
$\mu_2$	0	-0,00025	0,00005	0,00109	568,00889	171,13520
$\mu_3$	0	0,00014	0,00005	0,00107	553,61503	168,27496
$\lambda_1$	0,05	0,06268	0,00038	0,00453	228,26680	68,79978
$\lambda_2$	0,7	0,54394	0,01005	0,07537	118,44189	35,70186
$\Sigma_{12}$	0,5	0,50037	0,00014	0,00295	542,73185	164,04439
$\Sigma_{13}$	0,25	0,24819	0,00018	0,00347	427,15063	129,07373
$\Sigma_{23}$	0,5	0,49846	0,00017	0,00319	392,82520	118,81678
$\nu$	0,5	0,47174	0,00015	0,00198	245,91701	74,39007

Fonte: Elaborada pelo autor.

Tabela 13 – Resultados do estimador ZV-HMC-Q para o modelo DCC-GJR-GARCH(1,1)-GED.

	<b>Real</b>	<b>Média</b>	<b>Erro-padrão</b>	<b>Desvio-padrão</b>	<b>ESS</b>	<b>ESS<sub>t</sub></b>
$\omega_1$	0,03	0,06304	0,00154	0,01131	265,07334	79,33459
$\omega_2$	0,03	0,06172	0,00131	0,01119	254,32895	77,05633
$\omega_3$	0,03	0,06794	0,00194	0,01453	244,08848	73,71589
$\alpha_1$	0,05	0,07193	0,00047	0,00493	256,47173	77,32427
$\alpha_2$	0,05	0,07254	0,00043	0,00469	256,06790	76,80731
$\alpha_3$	0,05	0,07718	0,00060	0,00596	235,08077	70,94155
$\beta_1$	0,85	0,73692	0,00430	0,03221	267,44791	80,47376
$\beta_2$	0,85	0,73948	0,00375	0,03226	257,36004	77,46916
$\beta_3$	0,85	0,72080	0,00536	0,04050	244,67199	73,75709
$\phi_1$	0,05	0,06021	0,00057	0,00620	297,16260	89,46004
$\phi_2$	0,05	0,05918	0,00051	0,00575	283,68625	85,52505
$\phi_3$	0,05	0,06190	0,00068	0,00723	275,87887	83,30892
$\mu_1$	0	-0,00016	0,00003	0,00045	400,83524	120,37450
$\mu_2$	0	-0,00006	0,00003	0,00047	407,98209	122,83710
$\mu_3$	0	0,00031	0,00003	0,00050	400,78245	121,72577
$\lambda_1$	0,05	0,06710	0,00018	0,00204	225,26134	67,84635
$\lambda_2$	0,7	0,58631	0,00478	0,04127	161,38118	48,61238
$\Sigma_{12}$	0,5	0,50067	0,00008	0,00140	445,23188	134,48045
$\Sigma_{13}$	0,25	0,24808	0,00011	0,00169	388,08782	117,12525
$\Sigma_{23}$	0,5	0,49925	0,00008	0,00141	390,84680	118,13523
$\nu$	0,5	0,47539	0,00006	0,00075	282,89106	85,51582

Fonte: Elaborada pelo autor.



---

## APLICAÇÃO

---

Este capítulo apresenta exemplos de aplicação dos modelos descritos no Capítulo 3. No caso univariado, ajustou-se o modelo GJR-GARCH(1,1) considerando-se as seguintes distribuições para os resíduos: normal, t de Student e GED. No caso multivariado, aplicou-se o modelo DCC-GJR-GARCH também com as distribuições de erros normal, t de Student e GED. Após os ajustes, aplicou-se alguns critérios de seleção de modelos que são apresentados a seguir, a fim de identificar, para ambos os casos univariado e multivariado, o modelo com melhor ajuste.

Conforme mencionado anteriormente, para a inferência dos modelos foi utilizada a abordagem bayesiana utilizando o algoritmo HMC e o método da variância zero com função experimental linear e quadrática, que foram apresentados no Capítulo 2.

### 5.1 Caso Univariado: Índice S&P500

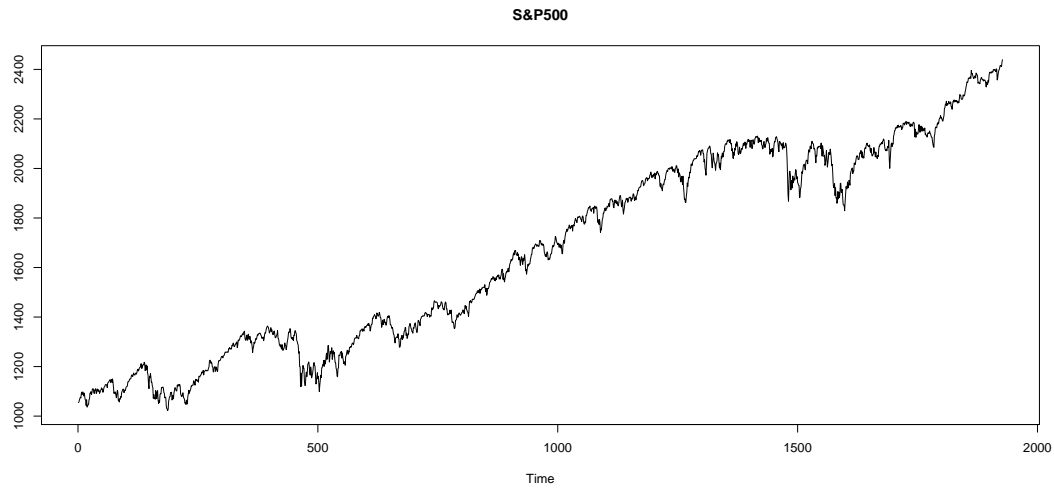
O conjunto de dados utilizado nos experimentos é baseado no indicador Standard & Poor's 500 (S&P500), também conhecido como "S&P". Esse indicador representa o desempenho médio das 500 ações mais importantes do mercado dos Estados Unidos, que são negociadas nas bolsas de valores *New York Stock Exchange* (NYSE) ou *National Association of Securities Dealers Automated Quotations* (NASDAQ). A série temporal que constitui esse conjunto de dados corresponde a valores de 06/10/2009 a 02/06/2017, totalizando 1.926 observações.

#### 5.1.1 Estatísticas Básicas

Esta seção apresenta análises descritivas da série temporal do S&P através de tabelas e figuras. As Figuras 3 e 4 mostram, respectivamente, os gráficos da série e dos log-retornos diários. Analisando-se a Figura 3, é possível observar que a série S&P500 não é estacionária, pois o gráfico apresenta várias tendências de crescimento durante o período, diferentemente do

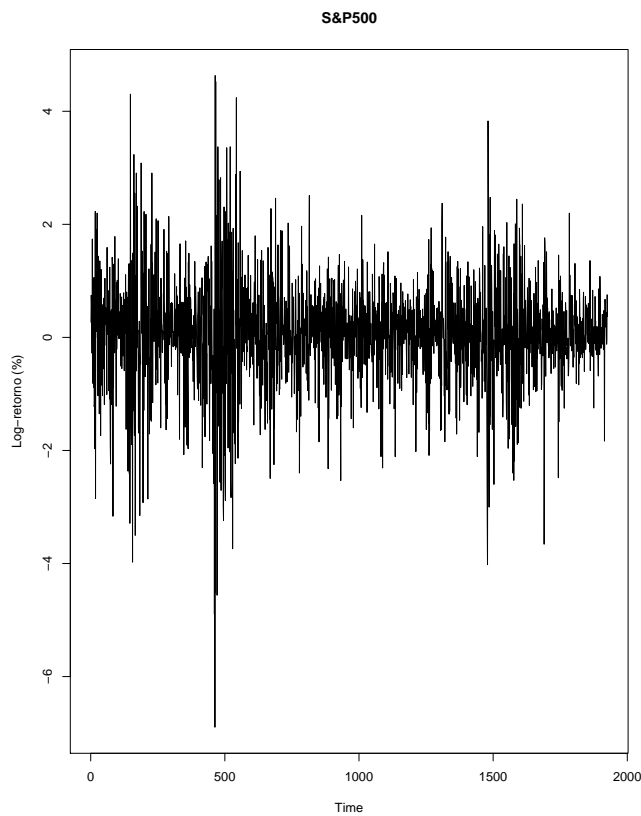
que ocorre com os log-retornos da série, que aparentam ser estacionários de segunda ordem (Figura 4).

Figura 3 – Gráfico da série diária do S&P500.



Fonte: Elaborada pelo autor.

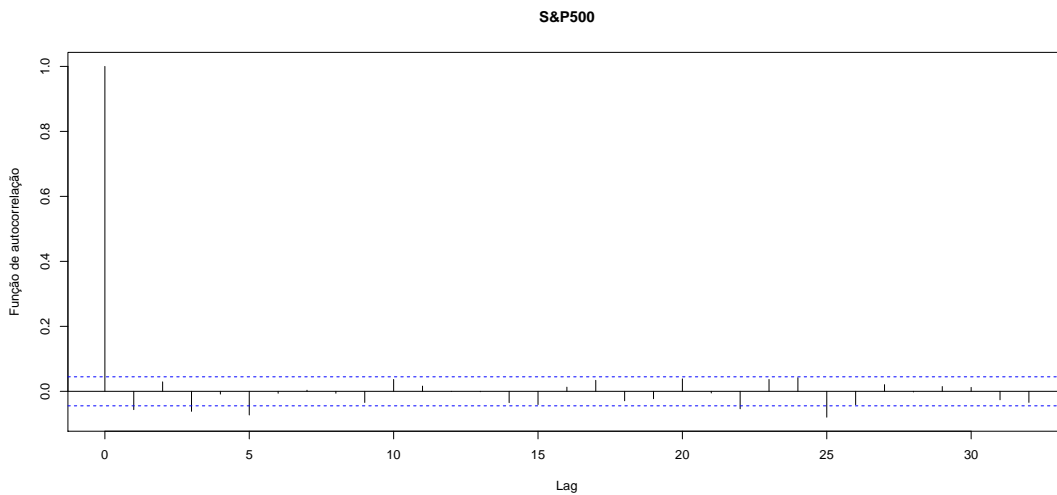
Figura 4 – Gráfico dos log-retornos diários da série S&P500.



Fonte: Elaborada pelo autor.

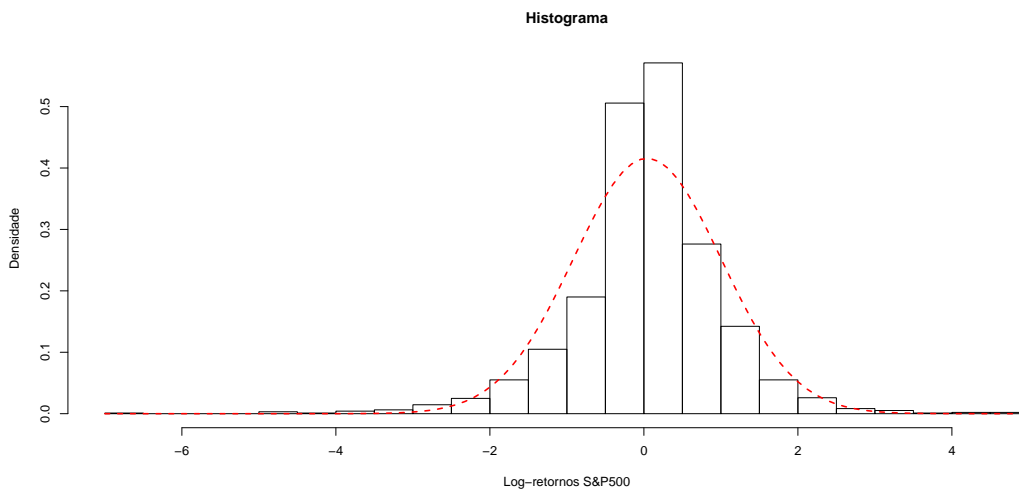
A Figura 5 apresenta a função de autocorrelação dos log-retornos da série S&P500. Pode-se observar que a dependência linear é praticamente inexistente, indicando que a série é um ruído branco, portanto não sendo autocorrelacionada. Essa característica justifica a utilização de modelos da família GARCH.

Figura 5 – Gráfico da função de autocorrelação dos log-retornos diários.



Fonte: Elaborada pelo autor.

Figura 6 – Histograma dos log-retornos diários.

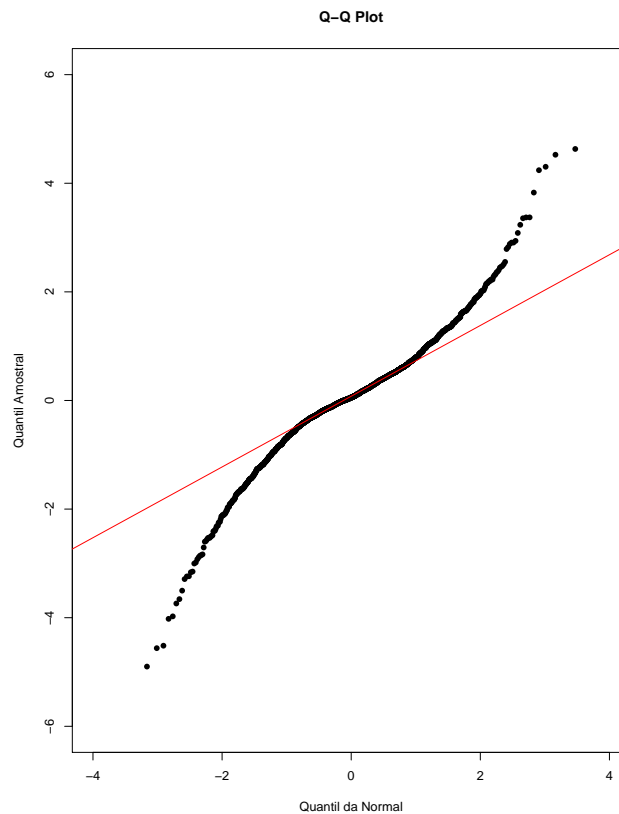


Fonte: Elaborada pelo autor.

De acordo com o histograma apresentado na Figura 6, os log-retornos aparentam um comportamento assimétrico à direita. Além disso, segundo o Q-Q plot da Figura 7, há indícios de que a distribuição da série tem uma cauda mais pesada comparada à distribuição normal.

A Tabela 14 contém as estatísticas básicas da série diária dos log-retornos do S&P500. Observa-se que a série apresenta uma possível assimetria à direita, pois o seu coeficiente de

Figura 7 – Q-Q plot dos log-retornos diários.



Fonte: Elaborada pelo autor.

assimetria estimado é maior do que 0. Além disso, a série apresenta curtose maior do que 3, indicando que a distribuição dos log-retornos possui uma cauda mais pesada do que a distribuição normal, fato já observado através do histograma e do Q-Q plot.

Tabela 14 – Estatísticas básicas da série diária dos log-retornos do S&amp;P500 (%).

	<b>ESTATÍSTICA (%)</b>
Mínimo	-6,89
1º Quartil	-0,36
Mediana	0,05
Média	0,04
3º Quartil	0,52
Máximo	4,63
Desvio-padrão	0,96
Assimetria	-0,44
Curtose	7,32

Fonte: Elaborada pelo autor.

### 5.1.2 Ajuste dos Modelos

Para os ajustes dos modelos GJR-GARCH(1,1), assume-se que os parâmetros são independentes e que seguem uma distribuição normal com média igual a 0 e variância igual a 100. Portanto, foi escolhida uma distribuição priori pouco informativa, visto que a variância é alta em relação aos dados.

Para simular a distribuição priori, foi utilizado o algoritmo HMC com parâmetros  $LF = 20$ ,  $M = I_k$  (matriz identidade) e  $\varepsilon$  ajustável de modo a garantir taxas de aceitação entre 70% e 90%. Para cada modelo foram geradas duas cadeias: (i) uma de tamanho 2.000, utilizada para estimar o vetor de parâmetros  $\mathbf{a}$  da metodologia *Zero-Variance*; e (ii) uma de tamanho 6.000, utilizada para fazer as inferências do modelo. Em ambas as cadeias, as primeiras 1.000 observações foram descartadas como amostra de aquecimento.

As Tabelas 15, 17 e 19 contêm, respectivamente, os resultados das simulações das distribuições posteriori dos modelos GJR-GARCH(1,1) com distribuição de erro normal, t de Student e GED. Em todos os modelos, o parâmetro  $\alpha$  possui um e-valor de 99,99% para o teste  $\alpha = 0$ . Assim, aplicando-se o FBST, não se rejeita a hipótese de que  $\alpha = 0$ . Diante disso, os modelos foram novamente simulados com a distribuição posteriori incluindo a restrição  $\alpha = 0$ . Nota-se que, no caso do modelo GJR-GARCH(1,1)-NORMAL, o parâmetro  $\mu$  também possui, individualmente, uma evidência alta (77,19%) de ser igual a 0, porém o e-valor indica uma baixa evidência ao considerar-se o parâmetro  $\mu$  juntamente a  $\alpha = 0$ . Então, rejeita-se a hipótese de que os parâmetros  $\alpha$  e  $\mu$  sejam ambos iguais a 0. Diferentemente do que ocorre no valor p, o e-valor no teste FBST é uma medida de evidência, podendo ser comparado. Assim, como o e-valor para  $\alpha = 0$  é maior do que o e-valor para  $\mu = 0$ , existe uma maior evidência de  $\alpha$  ser igual a 0 do que  $\mu$ . Observando-se os resultados da segunda simulação das distribuições posteriori com a restrição  $\alpha = 0$ , nota-se que não há evidências estatísticas de que os outros parâmetros sejam iguais a 0. Pode-se verificar a normalidade das inovações no modelo GJR-GARCH(1,1)-GED através do teste  $v = 1$ . Visto que o e-valor para esse teste é  $< 0,01\%$ , rejeita-se a hipótese de que a distribuição das inovações segue uma distribuição normal, ou seja, os dados não seguem o modelo GJR-GARCH(1,1)-NORMAL.

As Tabelas 16, 18 e 20 apresentam as estimativas HMC, ZV-HMC com função experimental linear e ZV-HMC com função experimental quadrática. Analisando-se os valores de erro-padrão de cada estimativa, nota-se que o método *Zero-Variance* é significativamente mais eficiente do que a estimativa HMC, pois suas estimativas apresentam um erro-padrão em torno de 10% do erro-padrão da estimativa HMC. Em termos práticos, para que o estimador HMC tenha a mesma eficiência dos métodos *Zero-Variance*, é necessário que o total de simulações seja por volta de 100 vezes maior do que o atual. Isto é, considerando-se o cenário das simulações realizadas, o estimador HMC precisaria ser executado por pelo menos 500.000 iterações para obter a eficiência dos estimadores *Zero-Variance*, que foram executados por 7.000 iterações.

Tabela 15 – Resultado das simulações da posteriori do modelo GJR-GARCH(1,1)-NORMAL.

Simulação da distribuição posteriori com o parâmetro $\alpha$								
	MÉDIA	ERRO-PADRÃO	2,5%	25%	50%	75%	97,5%	E-VALOR
$\omega$	0,04289	0,00005	0,03163	0,03851	0,04256	0,04685	0,05385	-
$\alpha$	0,00623	0,00023	0,00032	0,00208	0,00451	0,00874	0,01783	99,99%
$\phi$	0,26407	0,00018	0,20437	0,24063	0,26163	0,28587	0,32269	0,07%
$\beta$	0,81459	0,00018	0,77811	0,80285	0,81542	0,82704	0,84253	< 0,01%
$\mu$	0,03091	0,00036	-0,00247	0,01935	0,03078	0,04276	0,05964	77,19%

e-valor $_{\mu, \alpha} < 0,01\%$

Simulação da distribuição posteriori sem o parâmetro $\alpha$								
	MÉDIA	ERRO-PADRÃO	2,5%	25%	50%	75%	97,5%	E-VALOR
$\omega$	0,04245	0,00005	0,03110	0,03814	0,04212	0,04634	0,05317	-
$\phi$	0,26819	0,00137	0,20638	0,24448	0,26647	0,28994	0,32678	< 0,01%
$\beta$	0,81862	0,00009	0,78322	0,80711	0,81929	0,83063	0,84630	< 0,01%
$\mu$	0,03007	0,00017	-0,00261	0,01903	0,03016	0,04159	0,05675	48,66%

Fonte: Elaborada pelo autor.

Tabela 16 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q do modelo GJR-GARCH(1,1)-NORMAL.

	HMC	ERRO-PADRÃO	ZV-HMC-L	ERRO-PADRÃO	ZV-HMC-Q	ERRO-PADRÃO
$\omega$	0,04245	0,00005	0,04240	0,00009	0,042404	<0,00001
$\phi$	0,26819	0,00137	0,26733	0,00003	0,267324	0,00001
$\beta$	0,81862	0,00009	0,81889	0,00001	0,818887	<0,00001
$\mu$	0,03007	0,00017	0,03042	0,00001	0,030420	<0,00001

Fonte: Elaborada pelo autor.

Comparando-se as funções experimentais empregadas nos métodos *Zero-Variance*, observa-se que a função quadrática é a mais eficiente, pois possui valores de erro-padrão ainda menores. Esse comportamento era esperado, pois quanto maior o grau da função, maior a eficiência resultante. Uma consequência da metodologia *Zero-Variance* é que não é necessário ter grandes amostras da distribuição posteriori, o que leva a uma economia significativa de recursos computacionais.

Após a estimação dos parâmetros de cada modelo proposto, selecionou-se o modelo que melhor ajusta os dados. Para isso, foram utilizados todos os critérios de seleção de modelos apresentados no Capítulo 2. A Tabela 21 mostra as métricas referentes aos critérios DIC, EAIC, EBIC, WAIC e LOOIC para cada modelo observado. Analisando-se a tabela, é possível observar que o modelo GJR-GARCH(1,1)-GED possui o menor valor para todos os critérios. Portanto, por ser o modelo que melhor ajusta os dados, esse foi o modelo selecionado.

Após a escolha do modelo mais apropriado, iniciou-se a análise de influência com base na divergência de Kullback-Leibler, apresentada na Seção 3.4, para encontrar possíveis pontos influentes. A Figura 8 mostra o gráfico da divergência de Kullback-Leibler. Observa-se que o ponto amostral 1.690 é o mais influente. O ponto representa o dia 24/06/2016 e tem o valor de  $-3,65\%$ , ou seja, uma queda diária de  $3,65\%$  do índice. Um fato interessante ocorrido nesse dia é que o Reino Unido decidiu deixar a União Europeia através do referendo do Brexit, o que



Tabela 17 – Resultado das simulações da posteriori do modelo GJR-GARCH(1,1)-t.

Simulação da distribuição posteriori com o parâmetro $\alpha$								
	MÉDIA	ERRO-PADRÃO	2,5%	25%	50%	75%	97,5%	E-VALOR
$\omega$	0,03772	0,00008	0,02469	0,03253	0,03732	0,04226	0,05102	-
$\alpha$	0,00728	0,00014	0,00053	0,00275	0,00541	0,01011	0,01991	99,99%
$\phi$	0,32830	0,00071	0,24152	0,29668	0,32689	0,35981	0,40396	< 0,01%
$\beta$	0,79973	0,00038	0,75474	0,78574	0,80006	0,81419	0,83409	< 0,01%
$\mu$	0,05480	0,00025	0,02592	0,04458	0,05483	0,06495	0,07952	8,33%
$\nu$	5,84457	0,00505	4,54306	5,28191	5,74298	6,28803	7,25804	-
Simulação da distribuição posteriori sem o parâmetro $\alpha$								
	MÉDIA	ERRO-PADRÃO	2,5%	25%	50%	75%	97,5%	E-VALOR
$\omega$	0,03774	0,00005	0,02523	0,03248	0,03714	0,04229	0,05093	-
$\phi$	0,33399	0,00039	0,25206	0,30183	0,33144	0,36373	0,41219	< 0,01%
$\beta$	0,80293	0,00025	0,75973	0,78842	0,80409	0,81786	0,83761	< 0,01%
$\mu$	0,05374	0,00020	0,02333	0,04297	0,05362	0,06439	0,08029	3,35%
$\nu$	5,90272	0,00270	4,56380	5,31907	5,80448	6,38982	7,37363	-

Fonte: Elaborada pelo autor.

Tabela 18 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q do modelo GJR-GARCH(1,1)-t.

	HMC	ERRO-PADRÃO	ZV-HMC-L	ERRO-PADRÃO	ZV-HMC-Q	ERRO-PADRÃO
$\omega$	0,03774	0,00005	0,03787	0,00003	0,03783	0,00001
$\phi$	0,33399	0,00039	0,33494	0,00023	0,33482	0,00015
$\beta$	0,80293	0,00025	0,80258	0,00003	0,80265	0,00002
$\mu$	0,05374	0,00020	0,05352	0,00004	0,05352	0,00001
$\nu$	5,90272	0,00270	5,87163	0,00302	5,87241	0,00150

Fonte: Elaborada pelo autor.

levou a uma instabilidade financeira nesse período. O segundo ponto de maior influência é 1.925, que representa o dia 17/05/2017 e teve uma queda de 4,06%. Nesse dia, houve o início da crise do governo Temer devido ao vazamento do áudio do dono da JBS, cujas ações são negociadas no mercado americano, o que justifica tal queda, embora com um menor impacto em relação ao fato histórico do Brexit. Apesar de tais pontos serem os mais influentes, optou-se por mantê-los na base de dados, pois ambos remetem a ocorrências importantes, o que não justifica sua remoção.

Com base nos resultados apresentados, conclui-se que o modelo melhor ajustado para os log-retornos da série S&P500, denotada por  $X_t$ , é dado por

$$\begin{aligned}
 \mathbf{X}_t &= Y_t + 0,04603\%, \\
 Y_t &= \sqrt{h_t} \varepsilon_t, \\
 h_t &= 0,04047 + 0,31297 I_{t-1} Y_{t-1}^2 + 0,80429 h_{t-1}, \\
 \varepsilon_t &\sim GED(0, \sigma_{ged}^2, 0,64755),
 \end{aligned} \tag{5.1}$$

onde  $\sigma_{ged}^2 = 0,348304$ . Observe-se que o parâmetro  $\phi = 0,31297$ , portanto maior do que 0. Em termos práticos, se houver queda na ação S&P500, haverá um aumento da volatilidade, ou seja, a série ficará mais instável. De fato, esse comportamento já foi observado com base nos pontos

Tabela 19 – Resultado das simulações da posteriori do modelo GJR-GARCH(1,1)-GED.

Simulação da distribuição posteriori com o parâmetro $\alpha$								
	MÉDIA	ERRO-PADRÃO	2,5%	25%	50%	75%	97,5%	E-VALOR
$\omega$	0,04085	0,00013	0,02682	0,03545	0,04029	0,04573	0,05440	-%
$\alpha$	0,00719	0,00011	0,00047	0,00246	0,00522	0,00975	0,02072	99,99%
$\phi$	0,30951	0,00053	0,22469	0,27741	0,30675	0,33979	0,39056	< 0,01%
$\beta$	0,79975	0,00045	0,75323	0,78507	0,80063	0,81531	0,83548	< 0,01%
$\mu$	0,04607	0,00019	0,01804	0,03676	0,04615	0,05560	0,06893	15,17%
$\nu$	0,64646	0,00021	0,59229	0,62761	0,64560	0,66526	0,69469	< 0,01%
Simulação da distribuição posteriori sem o parâmetro $\alpha$								
	MÉDIA	ERRO-PADRÃO	2,5%	25%	50%	75%	97,5%	E-VALOR
$\omega$	0,04050	0,00012	0,02761	0,03533	0,04001	0,04505	0,05328	-
$\phi$	0,31299	0,00071	0,23639	0,28216	0,30986	0,34178	0,39047	< 0,01%
$\beta$	0,80467	0,00024	0,75923	0,79027	0,80573	0,82004	0,83974	< 0,01%
$\mu$	0,04555	0,00030	0,01587	0,03571	0,04527	0,05558	0,07064	7,07%
$\nu$	0,64755	0,00041	0,59075	0,62658	0,64684	0,66708	0,69992	< 0,01%

Fonte: Elaborada pelo autor.

Tabela 20 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q do modelo GJR-GARCH(1,1)-GED.

	HMC	ERRO-PADRÃO	ZV-HMC-L	ERRO-PADRÃO	ZV-HMC-Q	ERRO-PADRÃO
$\omega$	0,04050	0,00012	0,04046	0,00001	0,04047	<0,00001
$\phi$	0,31299	0,00071	0,31295	0,00014	0,31297	0,00006
$\beta$	0,80467	0,00024	0,80436	0,00004	0,80429	0,00001
$\mu$	0,04555	0,00030	0,04597	0,00004	0,04603	0,00001
$\nu$	0,64755	0,00041	0,64730	0,00004	0,64727	0,00002

Fonte: Elaborada pelo autor.

Tabela 21 – Critérios de seleção de modelos para o caso univariado.

	DIC	EAIC	EBIC	WAIC	LOOIC
GJR-GARCH(1,1)-NORMAL	4687,4	4691,5	4713,8	4689,9	4689,9
GJR-GARCH(1,1)-T	4594,8	4600,1	4627,9	4595,4	4595,4
GJR-GARCH(1,1)-GED	<b>4592,4</b>	<b>4597,5</b>	<b>4625,3</b>	<b>4593,1</b>	<b>4593,1</b>

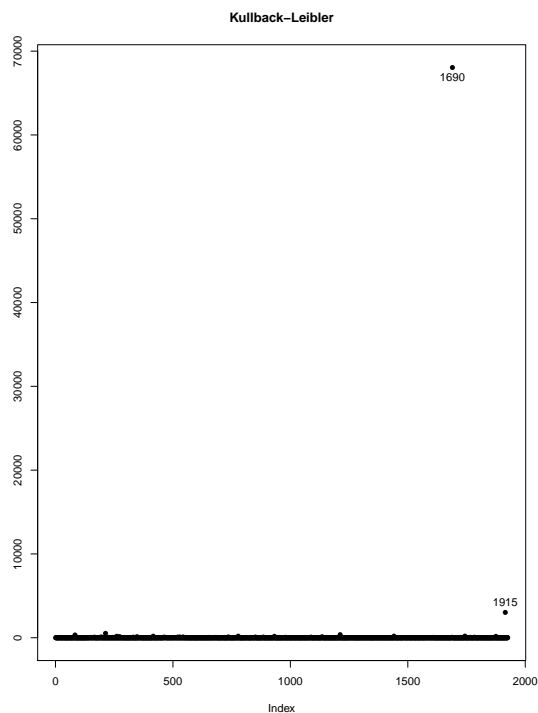
Fonte: Elaborada pelo autor.

influentes. Já com relação ao parâmetro  $\beta = 0,80429$ , uma vez que tal valor é mais próximo de 1 do que  $\frac{1}{2}$ , pode-se concluir que o índice é estável. Isto é, a volatilidade é pouco afetada por fatores externos, tais como políticos e econômicos. Para mais detalhes, pode-se analisar o Apêndice [B](#), que contém o restante dos gráficos da simulação da posteriori do modelo.

## 5.2 Caso Multivariado: Índices das Principais Bolsas de Valores Mundiais

O conjunto multivariado de dados utilizado nos experimentos é baseado nos principais índices econômicos das bolsas de valores mundiais:

Figura 8 – Gráfico da divergência de Kullback-Leibler para o modelo GJR-GARCH(1,1)-GED.



Fonte: Elaborada pelo autor.

- **Dow Jones Industrial Average.** Um dos principais índices estadunidenses (juntamente aos índices SP&500 e NASDAQ Composite) — referenciado apenas como “Dow Jones” no decorrer deste capítulo. O índice é composto pelas ações das 30 maiores empresas dos Estados Unidos.
- **DAX.** O índice *Deutscher Aktienindex* (DAX) compreende as 30 companhias alemãs com maior desempenho na Bolsa de Valores de Frankfurt.
- **BSE SENSEX.** O índice *Bombay Stock Exchange Sensitive Index* (BSE SENSEX) corresponde ao desempenho financeiro das 30 companhias mais rentáveis e bem estabelecidas na Bolsa de Valores de Mumbai, da Índia.
- **FTSE 100.** O índice *Financial Times Stock Exchange 100* (FTSE 100) abrange as 100 companhias com maior capitalização de mercado na Bolsa de Valores de Londres, no Reino Unido.
- **Nikkei 225.** Este índice mensura o desempenho financeiro das ações de 225 empresas japonesas na Bolsa de Valores de Tóquio.
- **CAC 40.** O índice *Cotation Assistée en Continu 40* (CAC 40) compreende as 40 companhias francesas com maior capitalização de mercado na Bolsa de Valores de Paris (*Euronext Paris*).

- **NIFTY 50.** Este índice mensura o desempenho das ações das 50 maiores companhias indianas na Bolsa de Valores Nacional da Índia.
- **IBOVESPA.** O Índice Bovespa (IBOVESPA) corresponde ao desempenho financeiro de aproximadamente 70 companhias brasileiras cujas ações são negociadas no Brasil Bolsa Balcão (B3).

A série temporal multivariada, formada pelos índices mencionados, consiste em 2.872 registros diários no período de 04/01/2010 a 01/01/2021. Os dados foram obtidos por meio da plataforma *Google Finance*<sup>1</sup>, que provê dados abertos sobre o mercado de ações.

Inicialmente, esta seção apresenta uma análise descritiva da base de dados, seguida pelos resultados dos ajustes dos modelos e, por fim, uma aplicação do modelo melhor ajustado em um cenário de escolha de carteira de ações.

### 5.2.1 Estatísticas Básicas

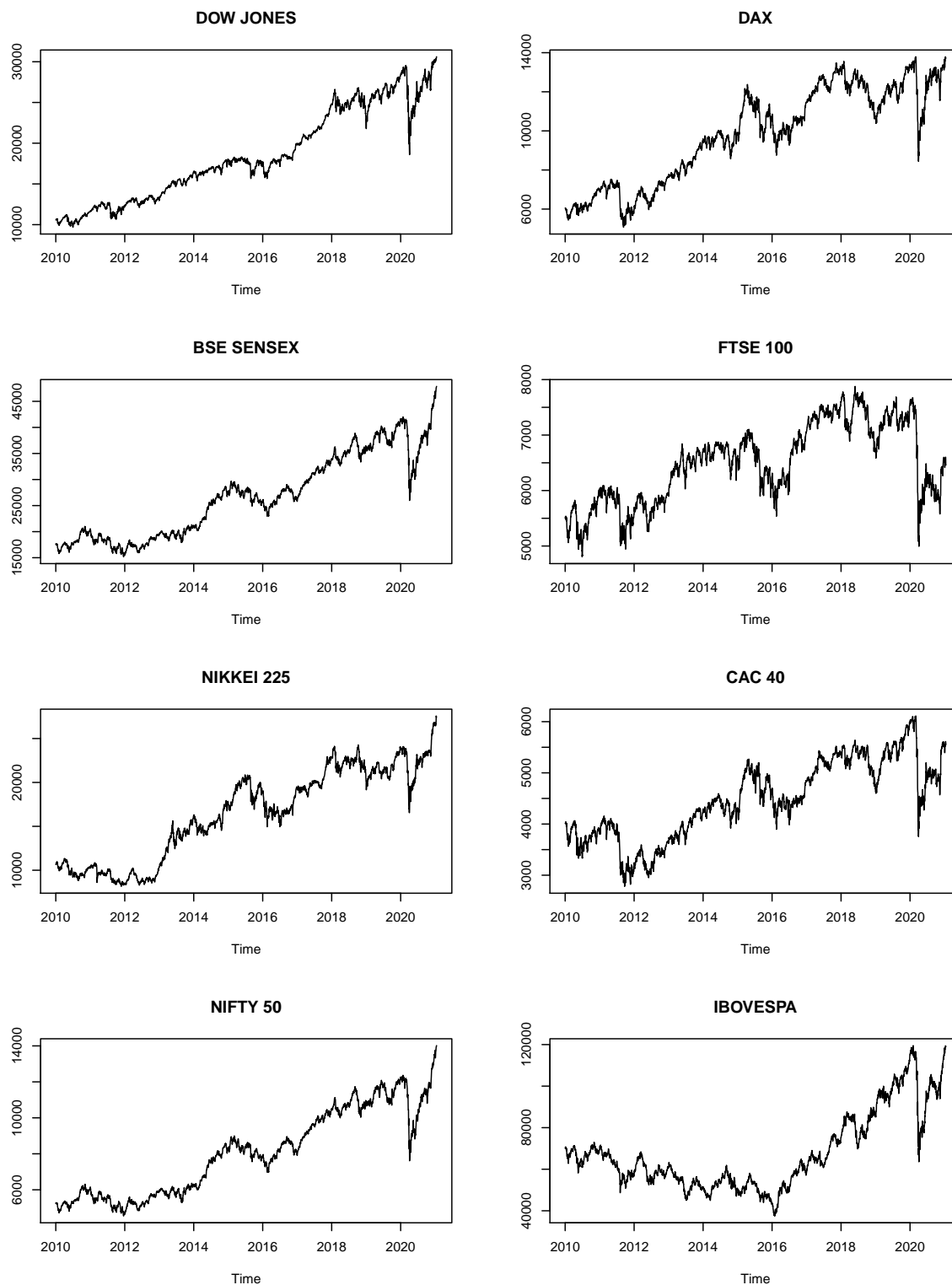
As Figuras 9 e 10 mostram, respectivamente, os gráficos das séries temporais e dos log-retornos diários correspondentes. Analisando-se os gráficos da Figura 9, é possível observar que as séries não são estacionárias, pois apresentam várias tendências de crescimento durante o período analisado. Vale ressaltar que todas as séries apresentaram um decaimento durante o início da pandemia de COVID-19, mas cujo desempenho foi recuperado com o passar do tempo, exceto para o índice FTSE 100. Já na Figura 10, os log-retornos das séries aparentam ser estacionários, porém pode-se observar que as séries são heterocedásticas, isto é, as variâncias (volatilidade) são distintas para cada intervalo de tempo. Além disso, em geral, observa-se um possível aumento da volatilidade no período inicial da pandemia, exceto para o índice Nikkei 225.

A Figura 11 apresenta a função de autocorrelação estimada dos log-retornos para cada índice individualmente. Analisando-se as funções, observa-se um comportamento geral de fraca autodependência observada. O único caso de destaque é o índice Dow Jones, pois apresenta uma maior autodependência, porém menor do que 0,2, o que indica uma autodependência ainda fraca. Portanto, é plausível o uso da modelagem da família GARCH multivariada para as séries utilizadas.

Por fim, a Tabela 22 apresenta as estatísticas básicas das séries de log-retornos. Todas as séries apresentam média acima de 0, ou seja, caso um investimento seja feito em suas respectivas bolsas de valores, é esperado que o investimento será rentável, por menor que seja o lucro. O índice mais estável, *i.e.* com menor desvio padrão (“DP” na tabela), é o FTSE 100, enquanto o mais instável é o IBOVESPA. Nesse contexto, o desvio padrão está refletindo as oscilações econômicas e políticas de cada país. Observa-se que cada série apresenta uma possível assimetria

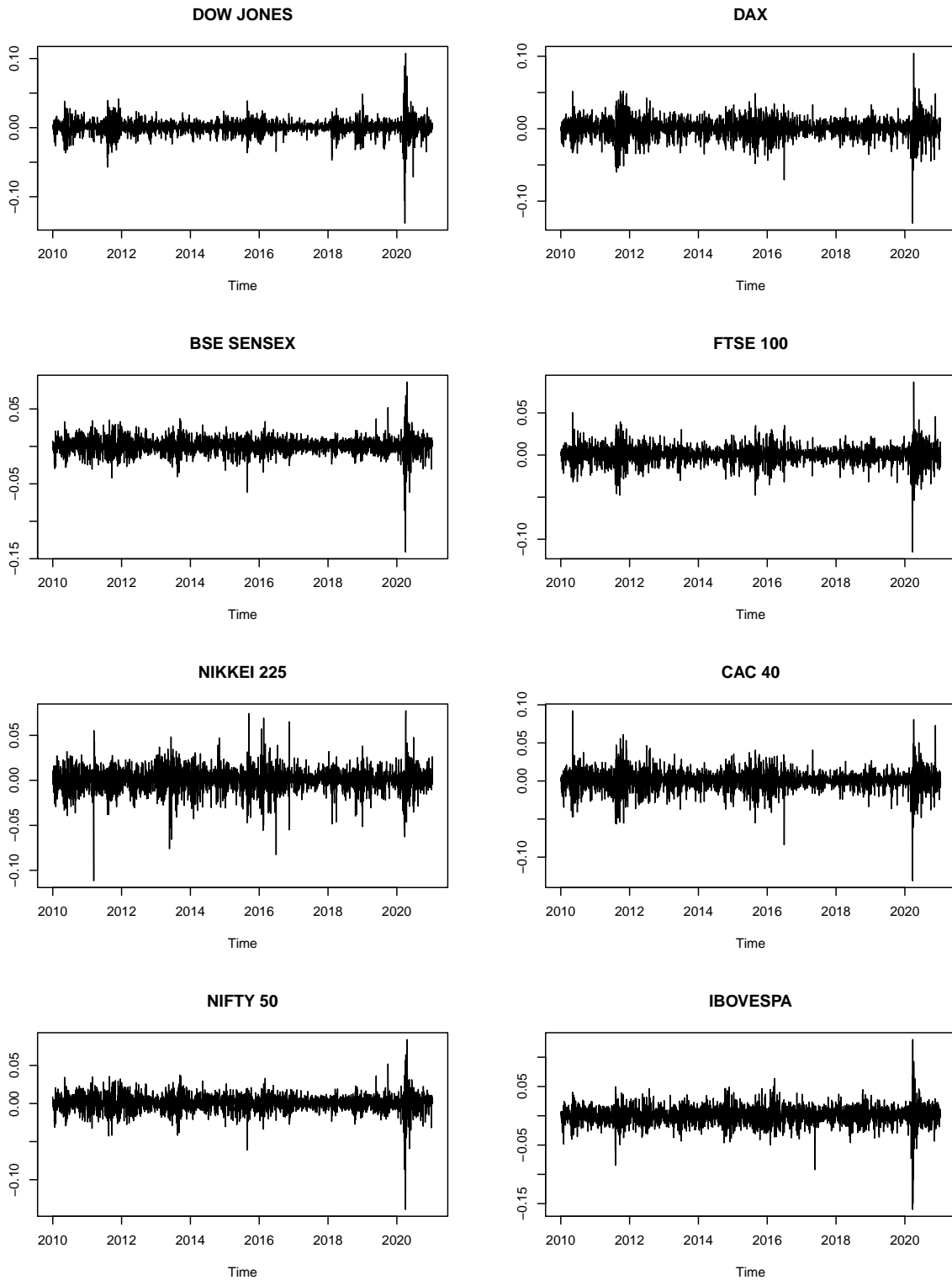
<sup>1</sup> <<https://www.google.com/finance>>

Figura 9 – Gráficos das séries temporais.



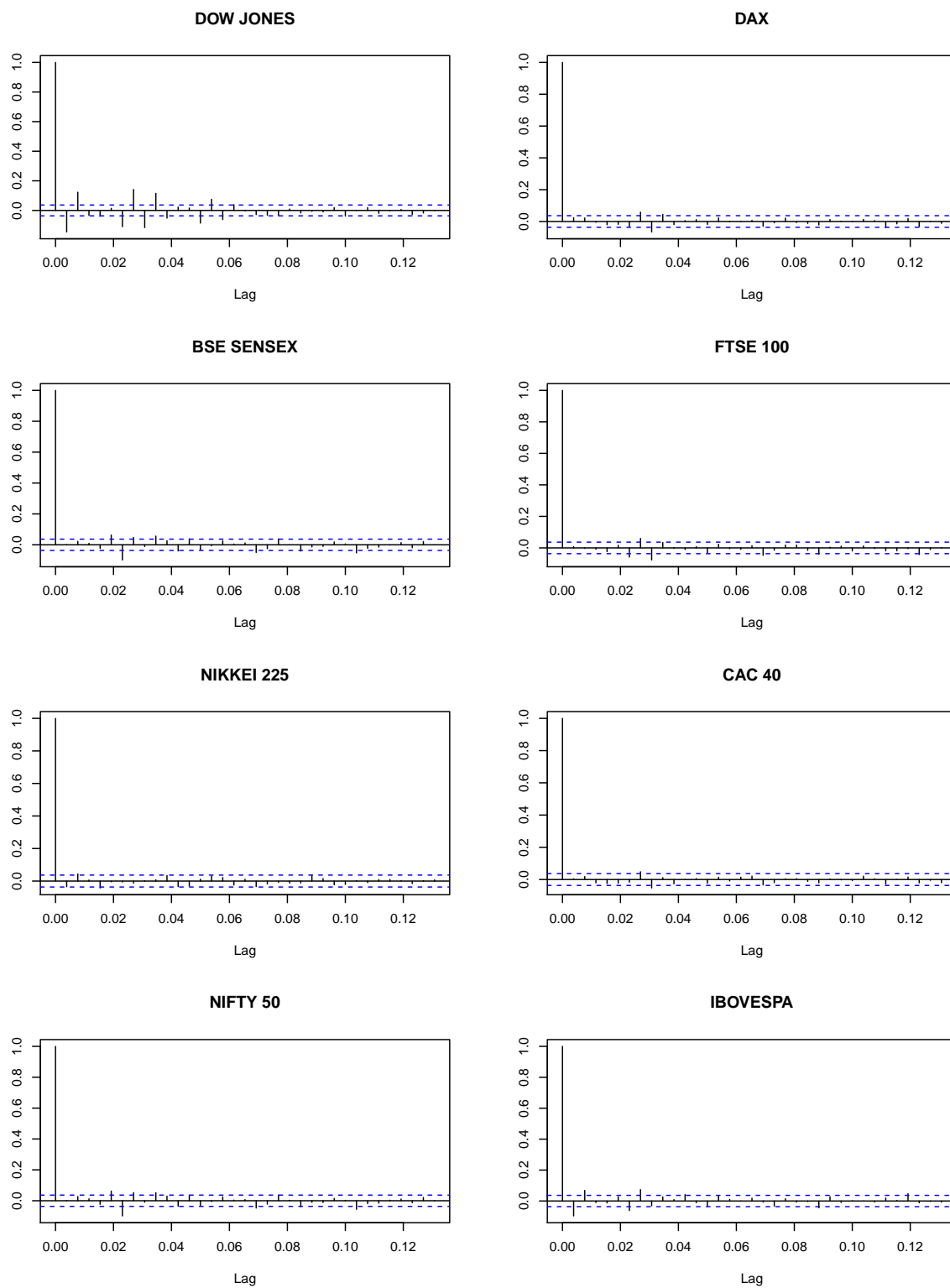
Fonte: Elaborada pelo autor.

Figura 10 – Gráficos dos log-retornos diários das séries temporais.



Fonte: Elaborada pelo autor.

Figura 11 – Gráficos da função de auto-correlação dos log-retornos diários.



Fonte: Elaborada pelo autor.

à direita, pois seus coeficientes de assimetria (“*Skew*” na tabela) estimados são negativos. Além disso, as séries apresentam curtose (“*Kurt*” na tabela) maior do que 3, indicando que a distribuição de probabilidade dos log-retornos possui uma cauda mais pesada em comparação à distribuição normal.

Tabela 22 – Estatísticas básicas das séries diárias dos log-retornos (%).

	<b>DOW JONES</b>	<b>DAX</b>	<b>BSE SENSEX</b>	<b>FTSE 100</b>	<b>NIKKEI 225</b>	<b>CAC 40</b>	<b>NIFTY 50</b>	<b>IBOVESPA</b>
Mín.	-13,84	-13,05	-14,10	-11,51	-11,15	-13,10	-13,90	-15,99
1° Q	-0,32	-0,51	-0,45	-0,46	-0,55	-0,57	-0,46	-0,74
2° Q	0,04	0,05	0,01	0,01	0	0,02	0,01	0
Média	0,04	0,03	0,03	0,01	0,03	0,01	0,03	0,02
3° Q	0,49	0,65	0,58	0,53	0,69	0,64	0,58	0,82
Máx.	10,76	10,41	8,59	8,67	7,73	9,22	8,40	13,02
DP	1,08	1,28	1,07	1,03	1,30	1,30	1,08	1,55
Skew	-0,97	-0,54	-1,06	-0,67	-0,47	-0,53	-1,05	-0,86
Kurt	25,84	11,09	20,10	13,36	9,03	11,26	18,78	15,92

Fonte: Elaborada pelo autor.

## 5.2.2 Ajuste dos Modelos

Para essa análise, os modelos foram ajustados considerando-se os log-retornos centuplicados. Assumiu-se que os parâmetros dos modelos DCC-GJR-GARCH(1,1) são independentes e seguem as seguintes distribuições marginais:

$$\begin{aligned}
 \omega_i &\sim \text{gama}(1, 1), \forall i = 1, 2, \dots, 8 \\
 \alpha_i, \beta_i, \phi_i &\sim \text{uniforme}(A), \forall i = 1, 2, \dots, 8 \\
 \lambda_i &\sim \text{uniforme}(B), \forall i = 1, 2 \\
 \mu_i &\sim \text{normal}(0, 1), \forall i = 1, 2, \dots, 8 \\
 \Sigma &\sim \text{LKJ corr}(2)
 \end{aligned} \tag{5.2}$$

onde

$$\begin{aligned}
 A = \{(\alpha, \beta, \phi) \mid &\alpha + \beta + 0,5\phi < 1, \\
 &\alpha > 0, \\
 &\beta > 0, \\
 &\phi > -\beta\},
 \end{aligned}$$

$$\begin{aligned}
 B = \{(\lambda_1, \lambda_2) \mid &\lambda_1 + \lambda_2 < 1 \\
 &\lambda_1 > 0, \\
 &\lambda_2 > 0\}.
 \end{aligned}$$



A principal vantagem da utilização dessa distribuição priori é que ela já incorpora as restrições do modelo, principalmente dos parâmetros  $\alpha$ ,  $\beta$  e  $\phi$ . Além disso, para os modelos DCC-GJR-GARCH(1,1)-t e DCC-GJR-GARCH(1,1)-GED, existe um parâmetro adicional  $\nu$ , definido conforme a seguir:

$$\begin{aligned} \nu &\sim \text{pareto}(4,3) \text{ para o modelo DCC-GJR-GARCH(1,1)-t} \\ \nu &\sim \text{gama}(2,2) \text{ para o modelo DCC-GJR-GARCH(1,1)-GED.} \end{aligned} \quad (5.3)$$

Analogamente ao caso univariado, foi utilizado o algoritmo HMC para simular a distribuição posteriori com parâmetros  $LF = 80$ ,  $M =$  matriz identidade e  $\varepsilon$  ajustável para garantir taxas de aceitação entre 70% e 90%. Observe-se que, no caso multivariado, o parâmetro  $LF$  é maior. Tal valor foi escolhido porque, comparando-se ao caso univariado, detectou-se que valores pequenos de  $LF$  tornam a estimação menos eficiente no caso multivariado. Esse fato é ocasionado pelo maior número de parâmetros a serem estimados. Para cada modelo foi gerada uma cadeia de 10.000 pontos, dos quais 5.000 foram considerados como amostra de aquecimento e utilizados para estimar o parâmetro  $\beta$  da metodologia *Zero-Variance*. Os outros 5.000 pontos foram utilizados para realizar a inferência dos modelos.

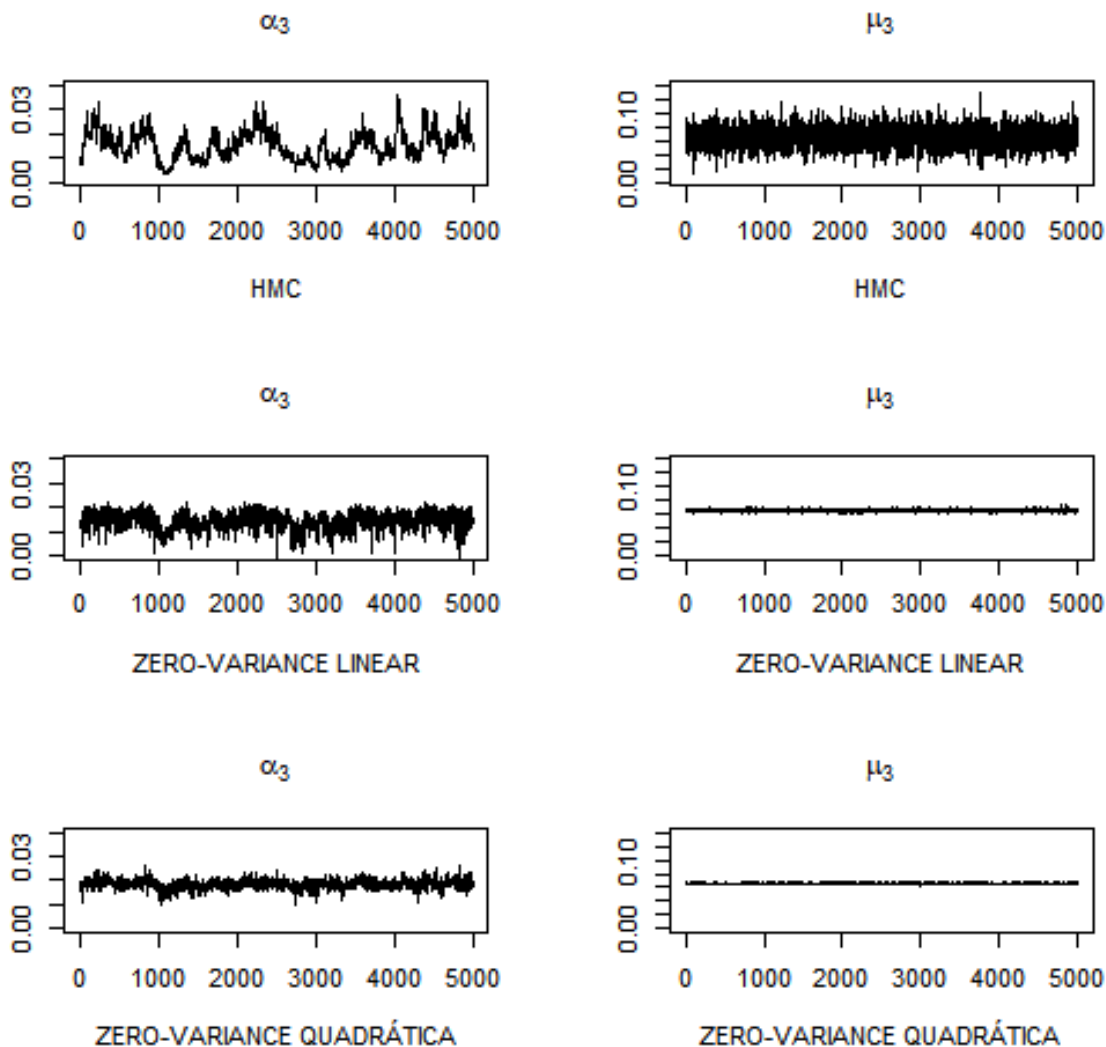
As Tabelas 28, 27, 31, 32, 35 e 36, apresentadas no Apêndice C, contêm as estatísticas descritivas referentes aos resultados do ajuste de cada modelo. O algoritmo HMC mostrou-se bastante eficiente para a maioria dos parâmetros. O único caso para o qual houve menos eficiência foi o conjunto de parâmetros  $\alpha$ . Isso pode ser explicado por alguns fatores.

1. Devido à restrição natural do modelo, existe uma alta dependência dos outros dois conjuntos de parâmetros,  $\beta$  e  $\phi$ .
2. Como os valores de  $\beta$  são próximos de 0.90, os valores de  $\alpha$  serão, no máximo, 0.1.
3. Visto que os valores de  $\alpha$  são muito próximos de 0, diminui-se a eficiência do algoritmo.

Apesar desses fatores, o erro-padrão da estimativa do algoritmo HMC é baixo. Portanto, tais fatores tiveram um baixo impacto negativo na estimação.

As Tabelas 29-30, 33-34 e 37-38 apresentam, respectivamente, as estimativas HMC, ZV-HMC com função experimental linear e ZV-HMC com função experimental quadrática. De forma semelhante ao caso univariado, o método *Zero-Variance* foi em média  $100\times$  mais eficiente para ambas as funções, com maior destaque à quadrática, se comparado ao estimador HMC. No contexto das simulações, para que o estimador HMC obtenha a mesma eficiência estatística dos estimadores *Zero-Variance*, é necessário executar uma cadeia de tamanho 500.000. Na prática, a utilização do método *Zero-Variance* possibilita uma grande economia de tempo computacional, pois no caso de modelos multivariados o tempo de processamento é elevado, devido ao grande número de parâmetros (70-71 nos modelos dessas simulações).

Figura 12 – Comparação dos métodos para dois parâmetros do modelo DCC-GJR-GARCH(1,1)-GED.



Fonte: Elaborada pelo autor.

A Figura 12 apresenta as cadeias geradas para dois parâmetros de cada método avaliado. Um dos parâmetros corresponde a um exemplo cuja cadeia foi menos eficiente, enquanto o outro corresponde a um exemplo de maior eficiência. Nota-se que, em ambos os casos, a metodologia *Zero-Variance* possibilitou resultados mais eficientes, ou seja, diminuiu o erro-padrão.

Após a estimação, selecionou-se o modelo de melhor ajuste aos dados com base nos mesmos critérios adotados no caso univariado, conforme apresentado pela Tabela 23. De acordo com os resultados, optou-se por selecionar o modelo DCC-GJR-GARCH-t.

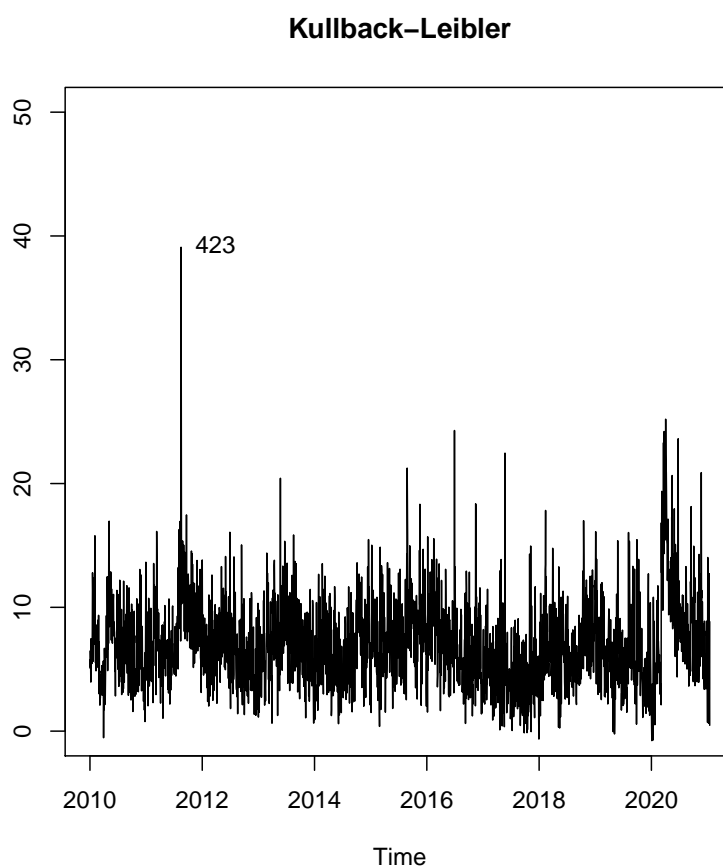
Após a escolha do modelo mais apropriado, iniciou-se a análise de influência com base na divergência de Kullback-Leibler para encontrar possíveis pontos influentes. A Figura 13 mostra o gráfico da divergência de Kullback-Leibler. Observa-se que o ponto amostral 423 é o mais influente. O ponto representa o dia 18/08/2011, em que houve uma queda de quase todos

Tabela 23 – Critérios de seleção de modelos para o caso multivariado.

	DIC	EAIC	EBIC	WAIC	LOOIC
DCC-GJR-GARCH-NORMAL	41309,2	41381,0	41798,3	41376,8	41366,5
DCC-GJR-GARCH-t	40159,9	<b>40225,6</b>	<b>40648,9</b>	<b>40158,3</b>	<b>40158,3</b>
DCC-GJR-GARCH-GED	<b>40158,2</b>	40226,3	40649,6	40164,4	40164,5

Fonte: Elaborada pelo autor.

Figura 13 – Gráfico da divergência de Kullback-Leibler para o modelo DCC-GJR-GARCH(1,1)-t.



Fonte: Elaborada pelo autor.

os índices avaliados. Um fato ocorrido nesse dia foi o temor de uma recessão mundial<sup>2</sup>, pois o banco Morgan Stanley reduziu as previsões de crescimento global para 2011 e 2012. Esse fato foi consequência de outro acontecimento no dia 02/08/2011, em que o Congresso americano autorizou aumentar o limite da dívida interna<sup>3</sup>. Embora esse ponto seja o mais influente, pode-se observar no gráfico que ele não se encontra muito distante dos outros pontos. Além disso, esse ponto remete a fatos relevantes para a época. Por isso, optou-se por mantê-lo na base de dados.

Diante dos resultados apresentados, o modelo escolhido para representar a série temporal

<sup>2</sup> <<https://bit.ly/3jWckQv>>

<sup>3</sup> <<https://exame.com/mundo/os-principais-acontecimentos-de-2011/>>

foi o DCC-GJR-GARCH(1,1)-t, cujos parâmetros foram estimados conforme a metodologia *Zero-Variance* com função experimental quadrática. Os detalhes referentes a valores dos parâmetros são apresentados nas Tabelas 33 e 34 do Apêndice C. Analisando-se os parâmetros estimados, conclui-se que todos os índices possuem uma média positiva, ou seja, as ações envolvidas são, em média, rentáveis. Quanto ao conjunto de parâmetros  $\phi$ , pode-se observar que todos são maiores do que 0, indicando que cada queda nas ações ocasiona um grande aumento na volatilidade dos índices. Já os parâmetros  $\beta$  estão próximos de 0.90, apontando que os índices são estáveis, isto é, fatores externos afetam pouco sua volatilidade. Com relação ao parâmetro  $\Sigma$ , que corresponde à matriz de correlação dos índices, observou-se uma alta correlação entre os índices avaliados. Isso é explicado pela conectividade entre as economias dos países envolvidos, que é intrínseca ao fenômeno da globalização.

### 5.2.3 Exemplo de Aplicação: Seleção de Carteira de Investimentos

Suponha-se um cenário em que se deseja selecionar a carteira mundial de investimentos mais rentável no momento. Especificamente, o objetivo para este problema é encontrar uma carteira que permite o maior rendimento esperado com o menor risco possível. Em termos matemáticos, considerando-se o modelo selecionado na seção anterior, busca-se uma carteira com a maior média e a menor volatilidade. Nesse contexto, tem-se um problema de otimização não linear, que pode ser definido conforme a seguir:

$$\text{Maximizar } \textit{média da carteira} - \psi(\textit{volatilidade da carteira}), \quad (5.4)$$

onde  $\psi$  é o parâmetro do problema de otimização. Seja  $\rho$  um vetor em que cada elemento representa a proporção que uma determinada ação tem em uma determinada carteira. A média e a volatilidade da carteira são dadas por

$$\begin{aligned} \mathbb{E}(\textit{Carteira}) &= \rho^T \mathbb{E}(Y_{t+1} | \mathcal{F}_t) \\ \text{Var}(\textit{Carteira}) &= \rho^T \text{Var}(Y_{t+1} | \mathcal{F}_t) \rho. \end{aligned}$$

Portanto, o problema de otimização 5.4 é matematicamente expresso por

$$\begin{aligned} \text{Maximizar } & \rho^T \mathbb{E}(Y_{t+1} | \mathcal{F}_t) - \psi \rho^T \text{Var}(Y_{t+1} | \mathcal{F}_t) \rho \\ \text{Sujeito a } & \rho^T \mathbf{1}_k = 1 \\ & \rho \geq \mathbf{0}, \end{aligned} \quad (5.5)$$

O parâmetro  $\psi$  denota o coeficiente de aversão ao risco. Essencialmente, quanto menor o valor de  $\psi$ , menor será o interesse na volatilidade da carteira escolhida. Este é um clássico problema de

otimização de carteira que utiliza a média e a variância (*mean-variance portfolio optimization*). Para a resolução desse problema, os valores de  $\mathbb{E}(Y_{t+1} | \mathcal{F}_t)$  e  $\text{Var}(Y_{t+1} | \mathcal{F}_t)$ , inicialmente desconhecidos, foram estimados pelo modelo selecionado. As Tabelas 24 e 25 apresentam, respectivamente, as médias e a matriz de variâncias e covariâncias estimadas dos índices.

Tabela 24 – Médias estimadas dos retornos.

	DOW JONES	DAX	BSE SENSEX	FTSE 100	NIKKEI 225	CAC 40	NIFTY 50	IBOVESPA
MÉDIA	0,06356	0,06914	0,06846	0,03486	0,06076	0,04901	0,06890	0,04794

Fonte: Elaborada pelo autor.

Tabela 25 – Matriz de variâncias e covariâncias estimadas dos índices.

	DOW JONES	DAX	BSE SENSEX	FTSE 100	NIKKEI 225	CAC 40	NIFTY 50	IBOVESPA
DOW JONES	0,19614	0,19119	0,07104	0,14932	0,05083	0,19389	0,07203	0,19229
DAX	0,19119	0,51756	0,18946	0,33369	0,12868	0,47239	0,19153	0,22837
BSE SENSEX	0,07104	0,18946	0,51077	0,15265	0,13807	0,18715	0,51354	0,13514
FTSE 100	0,14932	0,33369	0,15265	0,34062	0,10626	0,34850	0,15394	0,19974
NIKKEI 225	0,05083	0,12868	0,13807	0,10626	0,46393	0,12939	0,13926	0,06051
CAC 40	0,19389	0,47239	0,18715	0,34850	0,12939	0,51354	0,18807	0,24352
NIFTY 50	0,07203	0,19153	0,51354	0,15394	0,13926	0,18807	0,52270	0,13825
IBOVESPA	0,19229	0,22837	0,13514	0,19974	0,06051	0,24352	0,13825	0,82590

Fonte: Elaborada pelo autor.

Na aplicação, foi escolhido  $\psi = 1$ , apontando que o maior interesse está em diminuir o risco da carteira. Para a implementação, foi utilizado o pacote **Rsolnp** (YE, 1987; GHALANOS; THEUSSSL, 2014) do programa R.

Após a execução do modelo, a carteira ótima obtida como resultado foi:

$$\text{CARTEIRA} = (0,6491 * \text{DOW JONES}) + (0,1504 * \text{BSE SENSEX}) + (0,2005 * \text{NIKKEI 225})$$

Tabela 26 – Comparação dos índices individuais com a carteira ótima.

	MÉDIA	VOLATILIDADE	VaR (99%)
DOW JONES	0,05842	0,19586	-1,03854
DAX	0,06278	0,51699	-1,71945
BSE SENSEX	0,06580	0,50281	-1,69181
FTSE 100	0,03140	0,34108	-1,41621
NIKKEI 225	0,05829	0,46272	-1,62780
CAC 40	0,04377	0,51600	-1,73675
NIFTY 50	0,06644	0,51466	-1,71176
IBOVESPA	0,04276	0,82921	-2,21437
<b>CARTEIRA</b>	<b>0,05954</b>	<b>0,14773</b>	<b>-0,89317</b>

Fonte: Elaborada pelo autor.

Observa-se que a carteira ótima possui apenas três índices, com o índice DOW JONES apresentando a maior proporção. Em termos práticos, dado o período considerado pela análise, pode-se obter uma carteira mais rentável e com menor risco ao investir-se nas bolsas de valores dos Estados Unidos (DOW JONES), Índia (BSE SENSEX) e Japão (NIKKEI). Note-se que nenhuma das ações da Europa foi escolhida. Isso pode ser explicado pelo maior impacto da pandemia de COVID-19 nos países europeus, alguns dos quais foram epicentros da pandemia, também com destaque ao surgimento da nova variante do Coronavírus no Reino Unido.

A Tabela 26 apresenta algumas estatísticas dos índices juntamente à carteira ótima. Vale ressaltar a estatística *Value at Risk* (VaR), que representa o máximo risco potencial em ações na bolsa de valores, isto é, o máximo que se pode perder ao investir na respectiva bolsa. Note-se que a carteira ótima não possui a maior rentabilidade (a maior sendo da bolsa de valores de Tóquio), porém possui o menor risco, pois tanto sua volatilidade quanto VaR apresentam os menores valores. De todos os índices, a bolsa representada pelo índice IBOVESPA possui o maior risco, porém não é a mais rentável. Portanto, dado o cenário considerado neste exemplo de aplicação, comprar ações da bolsa de valores do Brasil é o investimento menos indicado.

---

## CONCLUSÃO

---

Modelos da família GARCH são amplamente utilizados em diversas áreas envolvendo séries temporais financeiras. Grande parte dos modelos GARCH multivariados aplicados na literatura faz uso da inferência clássica. Vale ressaltar que, embora haja uma parcela de trabalhos abordando a inferência bayesiana, a inferência clássica também acaba sendo utilizada conjuntamente. Portanto, até o momento da execução deste trabalho de doutorado, poucos trabalhos na literatura armazenam modelos GARCH multivariados de forma puramente bayesiana.

Outra limitação encontrada na literatura diz respeito às distribuições de probabilidade utilizadas para os resíduos dos modelos GARCH. Em geral, emprega-se a distribuição normal, que não é adequada para dados de séries temporais financeiras por estas apresentarem resíduos com distribuição de cauda pesada. Diante disso, alguns trabalhos propuseram o uso de outras distribuições, com destaque para a  $t$  de Student e a GED. Entretanto, existe uma dificuldade inerente a estas distribuições no contexto de modelos GARCH, referente à identificação do parâmetro de forma da distribuição, principalmente na abordagem clássica.

Adicionalmente, na inferência bayesiana computacional, técnicas MCMC para modelos GARCH multivariados usualmente demandam elevados custos computacionais, mesmo em casos de dimensões baixas e moderadas. Um exemplo é a linguagem Stan, que tende a ser infactível para modelos multivariados.

Neste contexto, este trabalho de doutorado buscou abordar os seguintes problemas:

1. Conceber uma metodologia puramente bayesiana para realizar estimações com o menor custo computacional possível, porém sem comprometer a qualidade da estimacão;
2. Analisar e escolher a distribuição de probabilidade que melhor ajusta os dados, dentre as distribuições **normal**, **t** e **GED**; e
3. Aplicação de modelos assimétricos da família GARCH, focando-se nos modelos GJR-GARCH e DCC-GJR-GARCH.

Como resultado, foram desenvolvidas as seguintes contribuições.

1. De maneira a encontrar um estimador apropriado, combinou-se duas técnicas: **(i)** HMC, cujas principais contribuições são sua rápida convergência para a distribuição posteriori e a geração de cadeias com baixa dependência, mesmo em casos de modelos multivariados; e **(ii)** *Zero-Variance*, que propõe um novo estimador eficiente em termos de custo computacional, pois não demanda grandes cadeias.
2. Com relação à distribuição dos resíduos, foi proposta a distribuição GED multivariada, que até então não havia sido considerada como critério de escolha para modelos GARCH multivariados.
3. Aplicou-se os modelos GJR-GARCH e DCC-GJR-GARCH em um estudo de simulação com as técnicas HMC e *Zero-Variance* combinadas. Os resultados foram satisfatórios em termos de **eficácia**, pois os estimadores foram próximos dos valores reais, e de **eficiência**, visto que o custo computacional foi baixo, demandando apenas **segundos** de execução.
4. Os modelos GJR-GARCH e DCC-GJR-GARCH também foram aplicados em um estudo de caso real, com dados temporais sobre índices das principais bolsas de valores mundiais. Os resultados deste estudo de caso foram os seguintes: **(i)** identificou-se o melhor modelo ajustado para os casos univariado e multivariado; **(ii)** analisou-se os pontos influentes e encontrou-se os fatos históricos correspondentes; **(iii)** propôs-se uma metodologia para encontrar a carteira ideal para investimentos.

A implementação dos algoritmos propostos foi realizada através da linguagem C++ aliada à linguagem R, por meio das bibliotecas Rcpp e RcppArmadillo. Essas bibliotecas têm como principal vantagem a união da praticidade do ambiente R, principalmente em relação a funções e modelos estatísticos já embutidos, com o superior desempenho computacional da linguagem C++. O arcabouço utilizado para a implementação possibilita futuras inclusões de funcionalidades conforme a necessidade. Adicionalmente, a principal contribuição da biblioteca RcppArmadillo (ou simplesmente da biblioteca Armadillo no ambiente C++) foi de desempenho, visto que é otimizada para cálculos matriciais. Tais cálculos são essenciais para modelos multivariados, pois envolvem operações onerosas como a computação de matrizes inversas multidimensionais, decomposição de Cholesky, bem como cálculo de determinante.

Por meio dessas contribuições, este trabalho de doutorado avança o estado da arte para modelos GARCH na abordagem bayesiana baseada em técnicas computacionais. Além disso, cria-se oportunidades para trabalhos futuros, alguns dos quais são elencados a seguir.



## 6.1 Trabalhos Futuros

- Propor uma nova distribuição de probabilidade para inovações em modelos GARCH, como por exemplo a distribuição *Slash*, bem como investigar distribuições assimétricas (para mais detalhes, ver [Bauwens e Laurent \(2005\)](#), [Harvey e Sucarrat \(2014\)](#));
- Utilização das técnicas HMC e *Zero-Variance* em outros modelos GARCH, tais como modelos de fatores ou o modelo VEC-GARCH ([BURDA; MAHEU, 2013](#); [BURDA; BÉLISLE, 2019](#));
- Realizar extensões dos modelos estudados, como a combinação do modelo *Vector Autoregression* (VAR) com o modelo DCC-GARCH (VAR-DCC-GARCH) ([TODA, 1991](#)). Além de modelar a volatilidade, este modelo modela as séries temporais. Em outras palavras, este modelo é uma extensão do modelo ARMA-GARCH ([LIU; ERDEM; SHI, 2011](#)).



## REFERÊNCIAS

---

- ASSARAF, R.; CAFFAREL, M. Zero-variance principle for monte carlo algorithms. **Physical review letters**, APS, v. 83, n. 23, p. 4682, 1999. Citado nas páginas 28 e 39.
- BAE, K.-H.; KAROLYI, G. A.; STULZ, R. M. A new approach to measuring financial contagion. **Review of Financial studies**, Soc Financial Studies, v. 16, n. 3, p. 717–763, 2003. Citado na página 27.
- BAILLIE, R. T.; BOLLERSLEV, T. The message in daily exchange rates: a conditional-variance tale. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 20, n. 1, p. 60–68, 2002. Citado nas páginas 28 e 51.
- BAUWENS, L.; LAURENT, S. A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 23, n. 3, p. 346–354, 2005. Citado na página 97.
- BAUWENS, L.; LAURENT, S.; ROMBOUTS, J. V. Multivariate garch models: a survey. **Journal of applied econometrics**, Wiley Online Library, v. 21, n. 1, p. 79–109, 2006. Citado na página 27.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of econometrics**, Elsevier, v. 31, n. 3, p. 307–327, 1986. Citado nas páginas 25, 26 e 47.
- \_\_\_\_\_. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. **The review of economics and statistics**, JSTOR, p. 498–505, 1990. Citado nas páginas 27 e 53.
- BOLLERSLEV, T.; ENGLE, R. F.; WOOLDRIDGE, J. M. A capital asset pricing model with time-varying covariances. **Journal of political Economy**, The University of Chicago Press, v. 96, n. 1, p. 116–131, 1988. Citado nas páginas 27 e 52.
- BOX, G. E.; TIAO, G. C. **Bayesian Inference in Statistical Analysis**. [S.l.], 1973. Citado na página 51.
- BROOKS, S. **Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde**. [S.l.]: WILEY-BLACKWELL 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2002. Citado na página 44.
- BURDA, M.; BÉLISLE, L. Copula multivariate garch model with constrained hamiltonian monte carlo. **Dependence Modeling**, De Gruyter, v. 7, n. 1, p. 133–149, 2019. Citado na página 97.
- BURDA, M.; MAHEU, J. M. Bayesian adaptively updated hamiltonian monte carlo with an application to high-dimensional bekk garch models. **Studies in Nonlinear Dynamics & Econometrics**, De Gruyter, v. 17, n. 4, p. 345–372, 2013. Citado na página 97.
- CAPPIELLO, L.; ENGLE, R. F.; SHEPPARD, K. Asymmetric dynamics in the correlations of global equity and bond returns. **Journal of Financial econometrics**, Oxford Univ Press, v. 4, n. 4, p. 537–572, 2006. Citado nas páginas 27, 47 e 53.

- CARLIN, B. P.; LOUIS, T. A. **Bayes and empirical Bayes methods for data analysis**. [S.l.]: Chapman & Hall/CRC Boca Raton, FL, 2000. v. 17. Citado na página [44](#).
- CHUANG, I.-Y.; LU, J.-R.; LEE, P.-H. Forecasting volatility in the financial markets: a comparison of alternative distributional assumptions. **Applied Financial Economics**, Taylor & Francis, v. 17, n. 13, p. 1051–1060, 2007. Citado na página [51](#).
- DEGROOT, M. H. **Optimal statistical decisions**. [S.l.]: John Wiley & Sons, 2005. v. 82. Citado na página [32](#).
- DEGROOT, M. H.; SCHERVISH, M. J.; FANG, X.; LU, L.; LI, D. **Probability and statistics**. [S.l.]: Addison-Wesley Reading, MA, 1986. v. 2. Citado na página [31](#).
- 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 [27](#).
- EDDELBUETTEL, D. **Seamless R and C++ Integration with Rcpp**. New York: Springer, 2013. ISBN 978-1-4614-6867-7. Citado na página [61](#).
- EDDELBUETTEL, D.; FRANÇOIS, R.; ALLAIRE, J.; USHEY, K.; KOU, Q.; RUSSEL, N.; CHAMBERS, J.; BATES, D. Rcpp: Seamless r and c++ integration. **Journal of Statistical Software**, v. 40, n. 8, p. 1–18, 2011. Citado na página [61](#).
- ENGLE, R. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 20, n. 3, p. 339–350, 2002. Citado nas páginas [27](#), [51](#) e [53](#).
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. **Econometrica: Journal of the Econometric Society**, v. 50, p. 987–1007, 1982. Citado nas páginas [26](#), [27](#) e [47](#).
- ENGLE, R. F.; KRONER, K. F. Multivariate simultaneous generalized arch. **Econometric theory**, Cambridge Univ Press, v. 11, n. 01, p. 122–150, 1995. Citado na página [52](#).
- ENGLE, R. F.; NG, V. K.; ROTHSCCHILD, M. Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. **Journal of Econometrics**, Elsevier, v. 45, n. 1-2, p. 213–237, 1990. Citado nas páginas [27](#) e [52](#).
- FIORUCI, J. A.; EHLERS, R. S.; LOUZADA, F. Bayesdccgarch-an implementation of multivariate garch dcc models. **arXiv preprint arXiv:1412.2967**, 2014. Citado na página [28](#).
- GHALANOS, A.; THEUSSL, S. **Rsolnp: General Non-linear Optimization Using Augmented Lagrange Multiplier Method**. [S.l.], 2014. R package version 1.15. Citado na página [23](#).
- GLOSTEN, L. R.; JAGANNATHAN, R.; RUNKLE, D. E. On the relation between the expected value and the volatility of the nominal excess return on stocks. **The journal of finance**, Wiley Online Library, v. 48, n. 5, p. 1779–1801, 1993. Citado nas páginas [27](#) e [47](#).
- GÓMEZ, E.; GOMEZ-VIILEGAS, M.; MARIN, J. A multivariate generalization of the power exponential family of distributions. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 27, n. 3, p. 589–600, 1998. Citado na página [55](#).

- HARVEY, A.; SUCARRAT, G. Egarch models with fat tails, skewness and leverage. **Computational Statistics & Data Analysis**, Elsevier, v. 76, p. 320–338, 2014. Citado na página [97](#).
- 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 [28](#) e [39](#).
- KASS, R. E.; CARLIN, B. P.; GELMAN, A.; NEAL, R. M. Markov chain monte carlo in practice: a roundtable discussion. **The American Statistician**, Taylor & Francis Group, v. 52, n. 2, p. 93–100, 1998. Citado na página [67](#).
- KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. **The annals of mathematical statistics**, JSTOR, v. 22, n. 1, p. 79–86, 1951. Citado na página [59](#).
- LEWANDOWSKI, D.; KUROWICKA, D.; JOE, H. Generating random correlation matrices based on vines and extended onion method. **Journal of multivariate analysis**, Elsevier, v. 100, n. 9, p. 1989–2001, 2009. Citado na página [68](#).
- LINTON, O. B. Semiparametric and nonparametric arch modeling. In: **Handbook of financial time series**. [S.l.]: Springer, 2009. p. 157–167. Citado na página [53](#).
- LIU, H.; ERDEM, E.; SHI, J. Comprehensive evaluation of arma–garch (-m) approaches for modeling the mean and volatility of wind speed. **Applied Energy**, Elsevier, v. 88, n. 3, p. 724–732, 2011. Citado na página [97](#).
- MADRUGA, M. R.; ESTEVES, L. G.; WECHSLER, S. On the bayesianity of pereira-stern tests. **Test**, Springer, v. 10, n. 2, p. 291–299, 2001. Citado na página [36](#).
- MIRA, A.; SOLGI, R.; IMPARATO, D. Zero variance markov chain monte carlo for bayesian estimators. **Statistics and Computing**, Springer, v. 23, n. 5, p. 653–662, 2013. Citado nas páginas [39](#) e [41](#).
- NEAL, R. M. MCMC using Hamiltonian dynamics. In: **Handbook of Markov Chain Monte Carlo**. [S.l.]: Boca Raton: Chapman and Hall-CRC Press, 2011. v. 2, p. 113–162. Citado nas páginas [28](#), [39](#) e [42](#).
- NELSON, D. B. Conditional heteroskedasticity in asset returns: A new approach. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 347–370, 1991. Citado na página [27](#).
- ROBERT, C. P.; CASELLA, G. **Monte Carlo Statistical Methods (Springer Texts in Statistics)**. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005. ISBN 0387212396. Citado na página [37](#).
- SILVENNOINEN, A.; TERÄSVIRTA, T. Multivariate garch models. **Handbook of financial time series**, Springer, p. 201–229, 2009. Citado na página [52](#).
- SPIEGELHALTER, D.; THOMAS, A.; BEST, N.; GILKS, W.; LUNN, D. Bugs: Bayesian inference using gibbs sampling. mrc biostatistics unit, cambridge, england. URL: <http://www.mrc-bsu.cam.ac.uk/bugs>, 1994, 2003. Citado na página [42](#).
- SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. V. D. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 64, n. 4, p. 583–639, 2002. Citado na página [45](#).

- Stan Development Team. **RStan: the R interface to Stan**. 2016. R package version 2.14.1. Disponível em: <http://mc-stan.org/>. Citado na página 28.
- TODA, H. **Vector autoregression and causality**. Tese (Doutorado) — Yale University, 1991. Citado na página 97.
- TSE, Y. K.; TSUI, A. K. C. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 20, n. 3, p. 351–362, 2002. Citado na página 27.
- VEHTARI, A.; GELMAN, A.; GABRY, J. Practical bayesian model evaluation using leave-one-out cross-validation and waic. **Statistics and Computing**, Springer, p. 1–20, 2016. Citado na página 45.
- VIRBICKAITĖ, A.; AUSÍN, M. C.; GALEANO, P. A bayesian non-parametric approach to asymmetric dynamic conditional correlation model with application to portfolio selection. **Computational Statistics & Data Analysis**, Elsevier, v. 100, p. 814–829, 2016. Citado na página 28.
- WATANABE, S. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. **Journal of Machine Learning Research**, v. 11, n. Dec, p. 3571–3594, 2010. Citado na página 45.
- YE, Y. **Interior Algorithms for Linear, Quadratic, and Linearly Constrained Non-Linear Programming**. Tese (Doutorado) — Department of ESS, Stanford University, 1987. Citado na página 93.
- ZAKOIAN, J.-M. Threshold heteroskedastic models. **Journal of Economic Dynamics and control**, Elsevier, v. 18, n. 5, p. 931–955, 1994. Citado na página 27.

# DERIVADAS DO LOGARITMO DA FUNÇÃO DENSIDADE DAS DISTRIBUIÇÕES POSTERIORI

## A.1 GJR-GARCH

O modelo GJR-GARCH(1,1) é definido como

$$\begin{aligned} X_t &= Y_t + \mu, \\ Y_t &= \sqrt{h_t} \varepsilon_t, \\ h_t &= \omega + (\alpha + \phi I_{t-1}) Y_{t-1}^2 + \beta h_{t-1}, \\ \varepsilon_t &\sim D(0, 1, \nu), \end{aligned}$$

com os parâmetros  $\omega > 0$ ,  $\alpha \geq 0$ ,  $\phi \geq 0$ ,  $\beta \geq 0$ ,  $\mu \in \mathbb{R}$  e  $I_{t-1}$  igual a uma função indicadora, onde  $I_{t-1} = 1$  se  $X_{t-1} < 0$  e  $r_{i,t-1} = 0$  caso contrário. Os resíduos  $\varepsilon_t$  são independentes e identicamente distribuídos e  $D(0, 1, \nu)$  é uma distribuição de probabilidade com média 0, variância 1 e parâmetro  $\nu$ .

O logaritmo da distribuição posteriori é dado por

$$\log(\pi(\boldsymbol{\theta} | \mathbf{X})) = \log(K) + \log(\pi(\mathbf{X} | \boldsymbol{\theta})) + \log(\pi(\boldsymbol{\theta})),$$

$$\log(\pi(\boldsymbol{\theta})) = -0,5 * (\boldsymbol{\theta} - \mathbf{a})^T B(\boldsymbol{\theta} - \mathbf{a}),$$

$$\log(\pi(\mathbf{X} | \boldsymbol{\theta})) = \begin{cases} c_n - 0,5 \sum_{t=2}^T (\log(h_t) + \delta_t^2) & , \text{ se } e_t \sim N(0, 1) \\ c_t - 0,5 \sum_{t=2}^T \left( \log(h_t) + (\nu + 1) \log \left( 1 + \frac{\delta_t^2}{\nu - 2} \right) \right) & , \text{ se } e_t \sim t(0, \frac{\nu-2}{\nu}, \nu) \end{cases}, \text{ onde}$$

- $\delta_t^2 = \frac{y_t^2}{h_t}$ ,
- $c_n = -\frac{(T-1)k}{2} \log(2\pi)$ ,
- $c_t = (T-1) \left( -\frac{1}{2} \log(\pi) - \frac{1}{2} \log(v-2) + \log(\Gamma(\frac{v+1}{2})) - \log(\Gamma(\frac{v}{2})) \right)$ ,
- $c_{ged} = -(T-1) \left( \log(\Gamma(1 + \frac{1}{2v})) + (1 + \frac{1}{2v}) \log(2) + \log(\sigma_{ged}) \right)$ ,
- $c_{tgen} = (T-1) \left( \log(\eta) + \log\left(\Gamma\left(v + \frac{1}{2\eta}\right)\right) + v \log(v) - \log(\Gamma(v)) - \log\left(\Gamma\left(\frac{1}{2\eta}\right)\right) - \log(\sigma_{tgen}) \right)$ ,
- $\sigma_{ged}^2 = \frac{\Gamma(\frac{1}{2v})}{2^{1/v} \Gamma(\frac{3}{2v})} e$
- $\sigma_{tgen}^2 = \frac{\Gamma(\frac{1}{2\eta}) \Gamma(v)}{v^{1/\eta} \Gamma(\frac{3}{2\eta}) \Gamma(v - \frac{1}{\eta})}$ .

## Derivadas

Derivadas da função densidade da distribuição posteriori:

1. Modelo GJR-GARCH(1,1)-NORMAL:

$$\frac{\partial \log(\pi(\mathbf{X} | \theta))}{\partial \theta} = -0,5 \sum_{t=2}^T \left( \frac{1}{h_t} \frac{\partial h_t}{\partial \theta} + \frac{\partial \delta_t^2}{\partial \theta} \right) \forall \theta \in \{\omega, \alpha, \phi, \beta, \mu\},$$

2. Modelo GJR-GARCH(1,1)-t:

$$\begin{aligned} \frac{\partial \log(\pi(\mathbf{X} | \theta))}{\partial \theta} &= -0,5 \sum_{t=2}^T \left( \frac{1}{h_t} \frac{\partial h_t}{\partial \theta} + \left( \frac{v+1}{v-2} \right) \left( 1 + \frac{\delta_t^2}{v-2} \right)^{-1} \frac{\partial \delta_t^2}{\partial \theta} \right) \forall \theta \in \{\omega, \alpha, \phi, \beta, \mu\}, \\ \frac{\partial \log(\pi(\mathbf{X} | \theta))}{\partial v} &= \frac{T-1}{2} \left( \frac{1}{v-2} + \Psi\left(\frac{v+1}{2}\right) - \Psi\left(\frac{v}{2}\right) \right) \\ &\quad - 0,5 \sum_{t=2}^T \left( \log\left(1 + \frac{\delta_t^2}{v-2}\right) - \frac{(v-1)\delta_t^2}{(v-2)^2} \left(1 + \frac{\delta_t^2}{v-2}\right)^{-1} \right), \end{aligned}$$

3. Modelo GJR-GARCH(1,1)-GED:

$$\begin{aligned} \frac{\partial \log(\pi(\mathbf{X} | \theta))}{\partial \theta} &= -0,5 \sum_{t=2}^T \left( \frac{1}{h_t} \frac{\partial h_t}{\partial \theta} + \frac{v \delta_t^{2(v-1)}}{\sigma_{ged}^{2v}} \frac{\partial \delta_t^2}{\partial \theta} \right) \forall \theta \in \{\omega, \alpha, \phi, \beta, \mu\}, \\ \frac{\partial \log(\pi(\mathbf{X} | \theta))}{\partial v} &= (T-1) \left( +\frac{1}{2v^2} \Psi\left(1 + \frac{1}{2v}\right) + \frac{\log(2)}{2v^2} - \frac{\partial \log(\sigma_{tgen})}{\partial v} \right) \\ &\quad - 0,5 \sum_{t=2}^T \left( \left( \frac{\delta_t^2}{\sigma_{ged}^2} \right)^v \log\left(\frac{\delta_t^2}{\sigma_{ged}^2}\right) \delta_t^2 \frac{\partial \sigma_{ged}^{-2}}{\partial v} \right), \end{aligned}$$



onde

- $\Psi()$  é uma função digama;
- $g(\delta_t^2) = \left( v + \left( \frac{\delta_t^2}{\sigma_{tgen}^2} \right)^\eta \right)$ ;
- $\frac{\partial \delta_t^2}{\partial \theta} = -\frac{y_t^2}{2h_t} \frac{\partial h_t^2}{\partial \theta}$  se  $\theta \in \{\omega, \alpha, \phi, \beta\}$ ;
- $\frac{\partial \delta_t^2}{\partial \mu} = -\frac{y_t^2}{2h_t} \frac{\partial h_t^2}{\partial \mu} - \frac{2(y_t - \mu)}{h_t}$ ;
- $\frac{\partial h_t}{\partial \omega} = 1 + \beta \frac{\partial h_{t-1}}{\partial \omega}$ ;
- $\frac{\partial h_t}{\partial \alpha} = Y_{t-1}^2 + \beta \frac{\partial h_{t-1}}{\partial \alpha}$ ;
- $\frac{\partial h_t}{\partial \phi} = I_{t-1} Y_{t-1}^2 + \beta \frac{\partial h_{t-1}}{\partial \phi}$ ;
- $\frac{\partial h_t}{\partial \beta} = h_{t-1} + \beta \frac{\partial h_{t-1}}{\partial \beta}$ ;
- $\frac{\partial h_t}{\partial \mu} = -2(\alpha + \phi I_{t-1}) Y_{t-1} + \beta \frac{\partial h_{t-1}}{\partial \mu}$ ;
- $\frac{\partial \log(\sigma_{ged}^2)}{\partial v} = -\frac{1}{2v^2} \Psi\left(\frac{1}{2v}\right) + \frac{\log(2)}{v^2} + \frac{3}{2v^2} \Psi\left(\frac{3}{2v}\right)$ ;
- $\frac{\partial \log(\sigma_{tgen}^2)}{\partial v} = \Psi(v) - \frac{1}{v\eta} - \Psi\left(v - \frac{1}{\eta}\right)$ ; e
- $\frac{\partial \log(\sigma_{tgen}^2)}{\partial \eta} = -\frac{1}{2\eta^2} \Psi\left(\frac{1}{2\eta}\right) + \frac{\log(v)}{\eta^2} + \frac{3}{2\eta^2} \Psi\left(\frac{3}{2\eta}\right) + \frac{1}{\eta^2} \Psi\left(v - \frac{1}{\eta}\right)$ .

## A.2 DCC-GJR-GARCH

O modelo DCC-GJR-GARCH é dado por

$$\begin{aligned} \mathbf{X}_t &= \mathbf{Y}_t - \mu \\ \mathbf{Y}_t &= \mathbf{H}_t^{1/2} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\varepsilon}_t &\sim D_k(\mathbf{0}, \mathbf{I}_k, v), \end{aligned}$$

onde

$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t$ ,  $\mathbf{D}_t = \text{diag}(h_{11,t}^{1/2}, \dots, h_{kk,t}^{1/2})$ ,  $\mathbf{R}_t = \text{diag}(\mathbf{Q}_t)^{-1/2} \mathbf{Q}_t \text{diag}(\mathbf{Q}_t)^{-1/2}$ ,  $\mathbf{Q}_t = (1 - \lambda_1 - \lambda_2 - 0,5\lambda_3)\Sigma + \lambda_1 \mathbf{u}_{t-1} \mathbf{u}_{t-1}^T + \lambda_2 \mathbf{Q}_{t-1} + \lambda_3 \mathbf{v}_{t-1} \mathbf{v}_{t-1}^T$ ,  $\mathbf{R}_t = \text{diag}(\mathbf{Q}_t)^{-1/2} \mathbf{Q}_t \text{diag}(\mathbf{Q}_t)^{-1/2}$ ,  $h_{ii,t} = \omega_i + (\alpha_i + \phi_i I_{i,t-1}) X_{i,t-1}^2 + \beta_i h_{ii,t-1}$ ,  $\mathbf{u}_t = \mathbf{D}_t^{-1} \mathbf{Y}_t$ ,  $\mathbf{v}_t = \mathbf{v}_t \odot \mathbf{I}_{\mathbf{u}_t < 0}$ , onde  $\odot$  denota o produto Hadamard e  $\mathbf{I}_{\mathbf{u}_t < 0}$  é um vetor cujo  $i$ -ésimo elemento é igual a 1 se  $u_{i,t} < 0$  e igual a 0 caso contrário.

O logaritmo da distribuição posteriori é dado por

$$\log(\pi(\theta | \mathbf{X})) = \log(K) + \log(\pi(\mathbf{X} | \theta)) + \log(\pi(\theta)),$$

$$\log(\pi(\theta)) = -0,5(\theta - \mathbf{a})^T \mathbf{B}(\theta - \mathbf{a}),$$

$$\log(\pi(\mathbf{X} | \theta)) = \begin{cases} c_n - 0,5 \sum_{t=2}^T (\log(|\mathbf{H}_t|) + \delta_t^2) & , \text{ se } e_t \sim N_k(0, \mathbf{I}_k) \\ c_t - 0,5 \sum_{t=2}^T \left( \log(|\mathbf{H}_t|) + (\nu + k) \log \left( 1 + \frac{\delta_t^2}{\nu - 2} \right) \right) & , \text{ se } e_t \sim t_k(0, \frac{\nu-2}{\nu} \mathbf{I}_k, \nu) \\ c_{ged} - 0,5 \sum_{t=2}^T \left( \log(|\mathbf{H}_t|) + \left( \frac{\delta_t^2}{\sigma_{ged}^2} \right)^\nu \right) & , \text{ se } e_t \sim GED_k(0, \sigma_{ged}^2 \mathbf{I}_k, \nu) \\ c_{tgen} - \sum_{t=2}^T \left( \frac{\log(|\mathbf{H}_t|)}{2} + \left( \nu + \frac{k}{2\eta} \right) \log \left( \nu + \left( \frac{\delta_t^2}{\sigma_{tgen}^2} \right)^\eta \right) \right) & , \text{ se } e_t \sim tgen_k(0, \sigma_{tgen}^2 \mathbf{I}_k, \nu, \eta). \end{cases}$$

onde

- $\delta_t^2 = \mathbf{y}_t^T \mathbf{H}_t^{-1} \mathbf{y}_t$ ;
- $c_n = -\frac{(T-1)k}{2} \log(2\pi)$ ;
- $c_t = (T-1) \left( -\frac{k}{2} \log(\pi) - \frac{k}{2} \log(\nu - 2) + \log(\Gamma(\frac{\nu+k}{2})) - \log(\Gamma(\frac{\nu}{2})) \right)$ ;
- $c_{ged} = (T-1) \left( \log(k) \log \Gamma\left(\frac{k}{2}\right) - \frac{k}{2} \log(\pi) - \log(\Gamma(1 + \frac{k}{2\nu})) + (1 + \frac{k}{2\nu}) \log(2) - k \log(\sigma_{ged}) \right)$ ;
- $c_{tgen} = (T-1) \left( \log(\eta) + \log \Gamma\left(\frac{k}{2}\right) + \log \left( \Gamma\left(\nu + \frac{k}{2\eta}\right) \right) + \nu \log(\nu) - \frac{k}{2} \log(\pi) - \log(\Gamma(\nu)) \right. \\ \left. - \log \left( \Gamma\left(\frac{k}{2\eta}\right) \right) - \frac{k}{2} \log(\sigma_{tgen}) \right)$ ;
- $\sigma_{ged}^2 = \frac{k\Gamma(\frac{k}{2\nu})}{2^{1/\nu} \Gamma(\frac{k+2}{2\nu})}$ ; e
- $\sigma_{tgen}^2 = \frac{k\Gamma(\frac{k}{2\eta})\Gamma(\nu)}{\nu^{1/\eta} \Gamma(\frac{k+2}{2\eta})\Gamma(\nu - \frac{1}{\eta})}$ .

## Derivadas

Derivadas da função densidade da distribuição posteriori:

1. Modelo DCC-GJR-GARCH(1,1)-NORMAL:

$$\frac{\partial l(\theta)}{\partial \theta_i} = -\frac{1}{2} \sum_{t=2}^T \left( \frac{\log(|\mathbf{H}_t|)}{\partial \theta_i} + \frac{\partial \delta_t^2}{\partial \theta_i} \right), \forall \theta_i \in \theta,$$

## 2. Modelo DCC-GJR-GARCH(1,1)-t:

$$\begin{aligned}\frac{\partial l(\theta)}{\partial \theta_i} &= -\frac{1}{2} \sum_{t=2}^T \left( \frac{\log(|\mathbf{H}_t|)}{\partial \theta_i} + \frac{\nu+k}{\nu-2} \left(1 + \frac{\delta_t^2}{\nu-2}\right)^{-1} \frac{\partial \delta_t^2}{\partial \theta_i} \right), \forall \theta_i \in (\omega, \alpha, \phi, \beta, \mu, \Sigma)^T; \\ \frac{\partial l(\theta)}{\partial \nu} &= \frac{T-1}{2} \left( \psi\left(\frac{\nu+k}{2}\right) - \psi\left(\frac{\nu}{2}\right) - \frac{k}{\nu-2} \right) \\ &\quad - \frac{1}{2} \sum_{t=2}^T \left( \log\left(1 + \frac{\delta_t^2}{\nu-2}\right) - \frac{(\nu+1)\delta_t^2}{(\nu-2)^2} \left(1 + \frac{\delta_t^2}{\nu-2}\right)^{-1} \right),\end{aligned}$$

## 3. Modelo DCC-GJR-GARCH(1,1)-GED:

$$\begin{aligned}\frac{\partial l(\theta)}{\partial \theta_i} &= -\frac{1}{2} \sum_{t=2}^T \left( \frac{\log(|\mathbf{H}_t|)}{\partial \theta_i} + \frac{\nu}{2 \exp(c_\nu)^\nu} (\delta_t^2)^{\nu-1} \frac{\partial \delta_t^2}{\partial \theta_i} \right) \forall \theta_i \in (\omega, \alpha, \phi, \beta, \mu, \Sigma)^T; \\ \frac{\partial l(\theta)}{\partial \nu} &= -\frac{(T-1)k}{2} \left( \frac{1}{\nu^2} \psi\left(1 + \frac{k}{2\nu}\right) + \frac{\log(2)}{\nu^2} - \frac{\partial c_\nu}{\partial \nu} \right) \\ &\quad + \frac{1}{2} \sum_{t=2}^T \left( \left( \frac{\delta_t^2}{\exp(c_\nu)} \right)^{\nu+1} (\log(\delta^2) - c_\nu) \frac{\partial c_\nu}{\partial \nu} \right),\end{aligned}$$

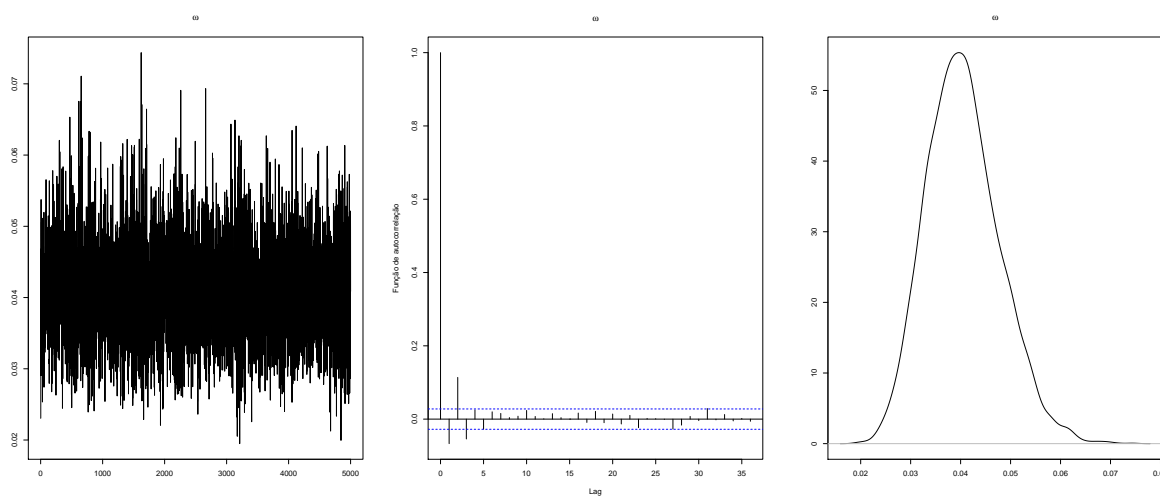
onde

- $\theta = (\omega, \alpha, \phi, \beta, \mu, \lambda, \Sigma, \nu, \eta)^T$ ;
- $g(\delta_t^2) = \left( \nu + \left( \frac{\delta_t^2}{\sigma_{tgen}^2} \right)^\eta \right)$ ;
- $\log(|\mathbf{H}_t|) = \text{tr} \left( \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \theta^*} \right) \forall \theta^* \in \{\omega, \alpha, \phi, \beta, \mu, \lambda, \Sigma\}$ ;
- $\delta_t^2 = \mathbf{y}_t^T \mathbf{H}_t^{-1} \mathbf{y}_t$ ;
- $\frac{\partial \delta_t^2}{\partial \theta^*} = -\mathbf{y}_t^T \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \theta^*} \mathbf{H}_t^{-1} \mathbf{y}_t \forall \theta^* \in \{\omega, \alpha, \phi, \beta, \mu, \lambda, \Sigma\}$ ;
- $\frac{\partial \mathbf{H}_t}{\partial \theta^*} = \frac{\partial \mathbf{D}_t}{\partial \theta^*} \mathbf{R}_t \mathbf{D}_t + \mathbf{D}_t \frac{\partial \mathbf{R}_t}{\partial \theta^*} \mathbf{D}_t + \mathbf{D}_t \mathbf{R}_t \frac{\partial \mathbf{D}_t}{\partial \theta^*} \forall \theta^* \in \{\omega, \alpha, \phi, \beta, \mu, \lambda, \Sigma\}$ ;
- $\frac{\partial \mathbf{R}_t}{\partial \theta^*} = \frac{\partial \mathbf{DQ}_t^{-1/2}}{\partial \theta^*} \mathbf{Q}_t \mathbf{DQ}_t^{-1/2} + \mathbf{DQ}_t^{-1/2} \frac{\partial \mathbf{Q}_t}{\partial \theta^*} \mathbf{DQ}_t^{-1/2} + \mathbf{DQ}_t^{-1/2} \mathbf{Q}_t \frac{\partial \mathbf{DQ}_t^{-1/2}}{\partial \theta^*}$ ;
- $\mathbf{DQ}_t^{-1/2} = \text{diag}(\mathbf{Q}_t)^{-1/2}$ ;
- $\frac{\partial h_{ii,t}}{\partial \omega_i} = 1 + \beta_i \frac{\partial h_{ii,t-1}}{\partial \omega_i} \forall i = 1, 2, \dots, k$ ;
- $\frac{\partial h_{ii,t}}{\partial \alpha_i} = y_{i,t-1}^2 + \beta_i \frac{\partial h_{ii,t-1}}{\partial \alpha_i} \forall i = 1, 2, \dots, k$ ;

- $\frac{\partial h_{ii,t}}{\partial \phi_i} = \mathbb{I}_{ii,t-1} y_{i,t-1}^2 + \beta_i \frac{\partial h_{t-1}}{\partial \phi_i} \quad \forall i = 1, 2, \dots, k;$
- $\frac{\partial h_{ii,t}}{\partial \beta_i} = h_{ii,t-1} + \beta_i \frac{\partial h_{ii,t-1}}{\partial \beta_i} \quad \forall i = 1, 2, \dots, k;$
- $\frac{\partial h_{ii,t}}{\partial \mu_i} = -2(\alpha_i + \mathbb{I}_{t-1} \phi_i) y_{i,t-1} + \beta_i \frac{\partial h_{ii,t-1}}{\partial \mu_i} \quad \forall i = 1, 2, \dots, k;$
- $\frac{\partial \mathcal{Q}_t}{\partial \theta^*} = \lambda_1 \frac{\partial \mathbf{u}_{t-1} \mathbf{u}_{t-1}^T}{\partial \theta^*} + \lambda_2 \frac{\partial \mathcal{Q}_{t-1}}{\partial \theta^*} + \lambda_3 \frac{\partial \mathbf{v}_{t-1} \mathbf{v}_{t-1}^T}{\partial \theta^*};$
- $\frac{\partial \mathcal{Q}_t}{\partial \lambda_1} = -\Sigma + \mathbf{u}_{t-1} \mathbf{u}_{t-1}^T + \lambda_2 \frac{\partial \mathcal{Q}_{t-1}}{\partial \lambda_1};$
- $\frac{\partial \mathcal{Q}_t}{\partial \lambda_2} = -\Sigma + \lambda_2 \mathcal{Q}_{t-1} + \lambda_2 \frac{\partial \mathcal{Q}_{t-1}}{\partial \lambda_2};$
- $\frac{\partial \mathcal{Q}_t}{\partial \lambda_3} = -\Sigma + \lambda_2 \frac{\partial \mathcal{Q}_{t-1}}{\partial \lambda_3} + \mathbf{v}_{t-1} \mathbf{v}_{t-1}^T;$
- $\frac{\partial \mathcal{Q}_t}{\partial \sigma_{i,j}} = (1 - \lambda_1 - \lambda_2 - 0,5\lambda_3) \frac{\partial \Sigma}{\partial \sigma_{i,j}} + \lambda_2 \frac{\partial \mathcal{Q}_{t-1}}{\partial \lambda_3} \quad \forall \sigma_{i,j} \in \Sigma; \mathbf{e}$
- $\frac{\partial \log(\sigma_{ged}^2)}{\partial v} = -\frac{k}{2v^2} \Psi\left(\frac{k}{2v}\right) + \frac{\log(2)}{v^2} + \frac{k+2}{2v^2} \Psi\left(\frac{k+2}{2v}\right).$

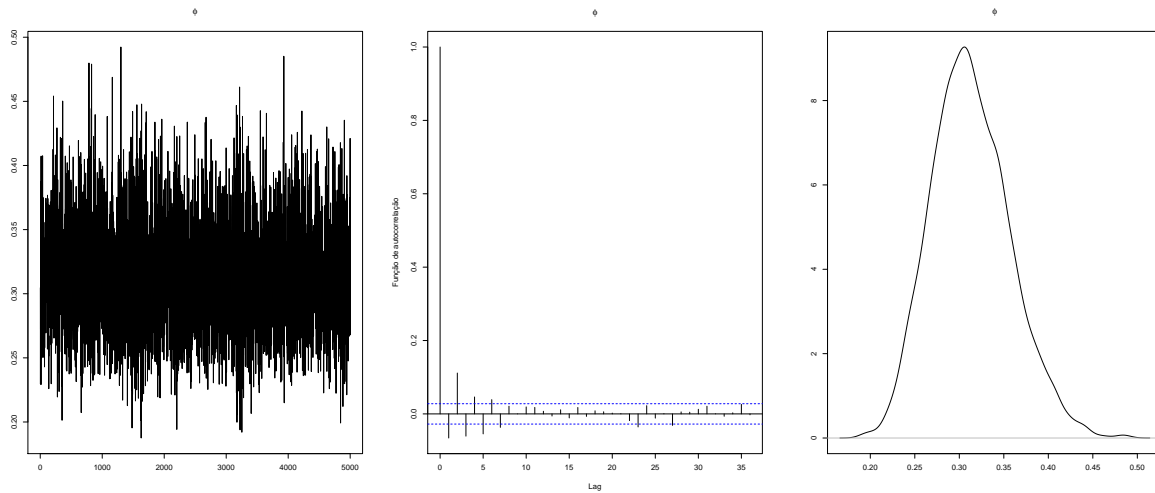
## SIMULAÇÕES DAS DISTRIBUIÇÕES POSTERIORI DO CASO UNIVARIADO

Figura 14 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro  $\omega$ .



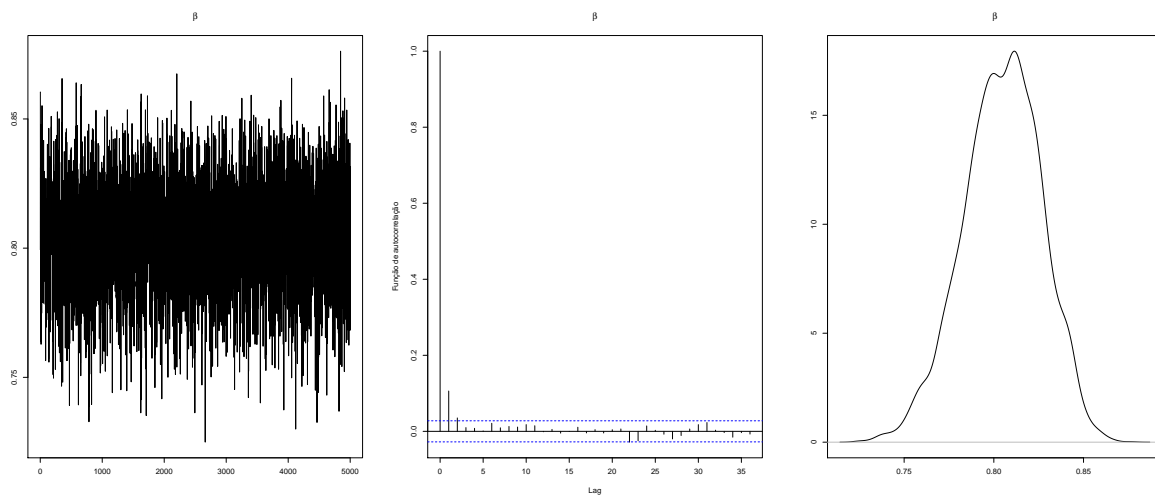
Fonte: Elaborada pelo autor.

Figura 15 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro  $\phi$ .



Fonte: Elaborada pelo autor.

Figura 16 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro  $\beta$ .



Fonte: Elaborada pelo autor.

Figura 17 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro  $\mu$ .

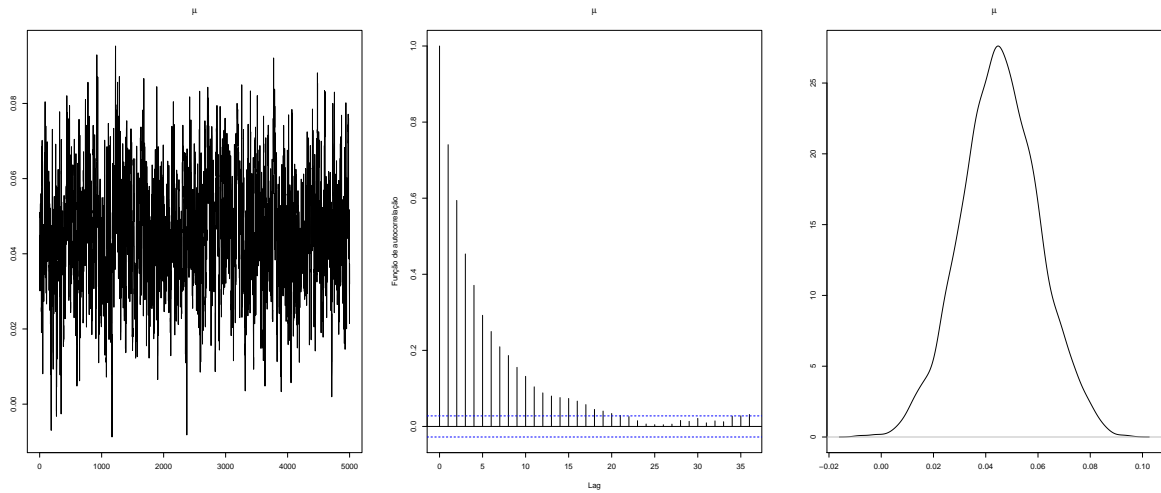
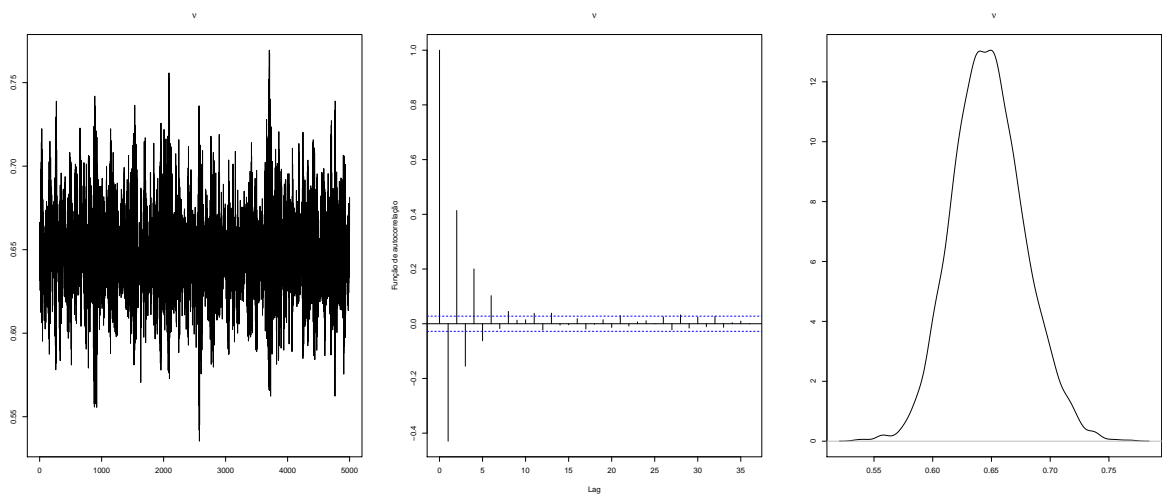


Figura 18 – Traços da simulação da posteriori, função de autocorrelação da simulação e função densidade aproximada do parâmetro  $\nu$ .



Fonte: Elaborada pelo autor.





## SIMULAÇÕES DAS DISTRIBUIÇÕES POSTERIORI DO CASO MULTIVARIADO

Tabela 27 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-NORMAL.

	Média	SD	Naive SE	TM SE	2,5%	25%	50%	75%	97,5%	n_eff
$\Sigma_{12}$	0,62475	0,01278	0,00018	0,00037	0,59933	0,61611	0,62471	0,63385	0,64942	1169,04176
$\Sigma_{13}$	0,25171	0,01913	0,00027	0,00036	0,214	0,23862	0,25169	0,26522	0,28879	2834,32018
$\Sigma_{14}$	0,59961	0,01321	0,00019	0,00049	0,57353	0,5906	0,60001	0,60872	0,6251	736,98009
$\Sigma_{15}$	0,2024	0,01999	0,00028	0,00027	0,16366	0,18922	0,20226	0,21616	0,24202	5679,74583
$\Sigma_{16}$	0,6336	0,01255	0,00018	0,00037	0,60891	0,62511	0,63388	0,64201	0,65774	1161,41379
$\Sigma_{17}$	0,25202	0,01902	0,00027	0,00036	0,21444	0,23901	0,25208	0,2653	0,28843	2740,32734
$\Sigma_{18}$	0,48073	0,01618	0,00023	0,00034	0,44875	0,46968	0,48091	0,49207	0,51214	2283,74883
$\Sigma_{23}$	0,39842	0,01716	0,00024	0,00033	0,36505	0,38648	0,39858	0,4101	0,43182	2737,89894
$\Sigma_{24}$	0,80677	0,00747	0,00011	0,00027	0,79248	0,80183	0,80685	0,81188	0,82105	783,18019
$\Sigma_{25}$	0,298	0,01866	0,00026	0,00027	0,26137	0,2855	0,29794	0,31019	0,33458	4620,2817
$\Sigma_{26}$	0,92199	0,00312	4e-05	0,00014	0,9156	0,91996	0,92207	0,92414	0,92788	474,17661
$\Sigma_{27}$	0,39663	0,0171	0,00024	0,00033	0,36357	0,38494	0,39684	0,40834	0,4298	2768,22016
$\Sigma_{28}$	0,36591	0,01835	0,00026	0,00035	0,32936	0,35371	0,3662	0,37824	0,4014	2682,19906
$\Sigma_{34}$	0,3943	0,01647	0,00023	0,00027	0,36183	0,38359	0,39425	0,40543	0,42636	3617,37493
$\Sigma_{35}$	0,31019	0,0186	0,00026	0,00021	0,27293	0,29792	0,31027	0,32308	0,34616	7685,76443
$\Sigma_{36}$	0,39504	0,01671	0,00024	0,00033	0,36295	0,3837	0,39507	0,40645	0,42778	2588,80376
$\Sigma_{37}$	0,99335	0,00027	0	2e-05	0,99282	0,99317	0,99335	0,99353	0,99387	181,10144
$\Sigma_{38}$	0,23575	0,01929	0,00027	0,00027	0,19738	0,22289	0,23573	0,24912	0,27218	5069,67225
$\Sigma_{45}$	0,30358	0,01895	0,00027	0,00025	0,26571	0,29115	0,30351	0,31615	0,3405	5660,0016
$\Sigma_{46}$	0,8417	0,0061	9e-05	0,00034	0,82937	0,83756	0,8418	0,8459	0,85325	318,55513
$\Sigma_{47}$	0,3916	0,01652	0,00023	0,00028	0,35852	0,38098	0,39178	0,40257	0,42392	3399,99008
$\Sigma_{48}$	0,39585	0,01738	0,00025	0,00033	0,36182	0,38393	0,39619	0,40751	0,42956	2712,92394
$\Sigma_{56}$	0,30259	0,01853	0,00026	0,00027	0,26623	0,29038	0,30241	0,31466	0,33891	4885,0979
$\Sigma_{57}$	0,30973	0,01855	0,00026	0,00021	0,27252	0,29739	0,30975	0,32249	0,345	7702,13685
$\Sigma_{58}$	0,12186	0,02092	3e-04	0,00025	0,08125	0,1078	0,12153	0,13651	0,16293	7219,45404
$\Sigma_{67}$	0,39249	0,01667	0,00024	0,00033	0,36033	0,38132	0,39233	0,40373	0,42547	2607,43852
$\Sigma_{68}$	0,3899	0,01792	0,00025	0,00035	0,35487	0,37758	0,39043	0,40211	0,4243	2605,42913
$\Sigma_{78}$	0,23785	0,01926	0,00027	0,00026	0,19971	0,2247	0,23772	0,25128	0,27409	5374,48155

Fonte: Elaborada pelo autor.

Tabela 28 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-NORMAL.

	Média	SD	Naive SE	TM SE	2,5%	25%	50%	75%	97,5%	n_eff
$\omega_1$	0,04139	0,00447	6e-05	0,00022	0,03277	0,03842	0,04124	0,04429	0,0506	430,14053
$\omega_2$	0,05141	0,00613	9e-05	4e-04	0,04064	0,04704	0,05098	0,05535	0,06443	231,52063
$\omega_3$	0,05038	0,00507	7e-05	0,00032	0,0411	0,04683	0,05014	0,05369	0,06094	248,98522
$\omega_4$	0,04474	0,00568	8e-05	0,00036	0,03435	0,04069	0,04449	0,04834	0,05635	245,10397
$\omega_5$	0,08593	0,01324	0,00019	0,00093	0,06186	0,07664	0,08515	0,09446	0,11444	202,91791
$\omega_6$	0,06682	0,00715	1e-04	0,00047	0,05418	0,06174	0,06635	0,07154	0,08142	235,6886
$\omega_7$	0,05186	0,0053	7e-05	0,00034	0,04216	0,04814	0,05168	0,05537	0,06309	243,93646
$\omega_8$	0,10685	0,02027	0,00029	0,0019	0,07237	0,09231	0,10472	0,1198	0,1497	113,30761
$\alpha_1$	0,06393	0,01248	0,00018	0,00094	0,04075	0,05547	0,06342	0,07193	0,08998	176,0613
$\alpha_2$	0,01002	0,0054	8e-05	0,00099	0,00132	0,00581	0,00943	0,01341	0,02221	30,03485
$\alpha_3$	0,00328	0,00287	4e-05	0,00108	0,00025	0,0012	0,00246	0,00443	0,0113	7,10888
$\alpha_4$	0,01155	0,00547	8e-05	0,00091	0,00368	0,00737	0,01052	0,01497	0,02339	35,8765
$\alpha_5$	0,04462	0,00991	0,00014	8e-04	0,02609	0,03792	0,04426	0,05109	0,06509	154,2161
$\alpha_6$	0,0249	0,00692	1e-04	0,00068	0,01366	0,01965	0,02418	0,02939	0,03974	102,8629
$\alpha_7$	0,00374	0,00294	4e-05	0,00102	0,00037	0,0017	0,00294	0,00506	0,01122	8,37947
$\alpha_8$	0,03111	0,01013	0,00014	0,00117	0,01269	0,02458	0,03065	0,03715	0,05345	75,42858
$\beta_1$	0,79932	0,01449	2e-04	0,00077	0,77079	0,7898	0,79942	0,80931	0,82718	355,36861
$\beta_2$	0,90807	0,00858	0,00012	0,00074	0,89012	0,90234	0,90863	0,91432	0,92335	135,83952
$\beta_3$	0,89172	0,00808	0,00011	0,00055	0,8753	0,88648	0,89189	0,89732	0,90702	214,21178
$\beta_4$	0,88213	0,01207	0,00017	0,001	0,85704	0,87453	0,88268	0,89052	0,90389	147,07828
$\beta_5$	0,83091	0,01496	0,00021	0,00096	0,79952	0,82084	0,83173	0,84121	0,85798	242,84554
$\beta_6$	0,8794	0,01009	0,00014	0,00075	0,85818	0,87297	0,8797	0,88656	0,89764	180,78849
$\beta_7$	0,89098	0,00832	0,00012	0,00052	0,87433	0,88537	0,89124	0,89672	0,90672	259,08976
$\beta_8$	0,87499	0,01609	0,00023	0,00144	0,84052	0,86517	0,87606	0,88585	0,9043	124,86662
$\phi_1$	0,17235	0,02148	3e-04	0,00095	0,13276	0,15733	0,17184	0,18659	0,21637	513,91837
$\phi_2$	0,07895	0,00825	0,00012	0,00063	0,06333	0,0733	0,07875	0,08437	0,0957	171,13078
$\phi_3$	0,09285	0,00797	0,00011	0,00038	0,07774	0,08741	0,09263	0,09812	0,10922	435,63963
$\phi_4$	0,09582	0,01027	0,00015	0,00048	0,07661	0,08879	0,09566	0,10241	0,11685	464,95995
$\phi_5$	0,15026	0,02049	0,00029	0,00101	0,11113	0,1361	0,14952	0,16376	0,19247	413,03403
$\phi_6$	0,08528	0,01009	0,00014	0,00055	0,06566	0,0784	0,08526	0,09192	0,10572	336,03582
$\phi_7$	0,09177	0,00801	0,00011	0,00039	0,07709	0,08623	0,09138	0,09704	0,10815	419,70206
$\phi_8$	0,08441	0,01658	0,00023	0,00098	0,05208	0,07318	0,08385	0,0958	0,11679	285,12696
$\mu_1$	0,04645	0,01234	0,00017	0,00025	0,02165	0,03821	0,04643	0,0546	0,07072	2468,84893
$\mu_2$	0,04146	0,0182	0,00026	0,00038	0,00611	0,02923	0,04151	0,05367	0,07814	2339,21477
$\mu_3$	0,04839	0,01544	0,00022	0,00047	0,01883	0,0382	0,04825	0,05864	0,07924	1081,77531
$\mu_4$	0,01054	0,01439	2e-04	0,00031	-0,01767	0,00105	0,01052	0,02031	0,03874	2188,08595
$\mu_5$	0,0391	0,01928	0,00027	0,00036	0,00109	0,02627	0,03916	0,05216	0,07634	2792,96045
$\mu_6$	0,02166	0,01759	0,00025	0,00041	-0,01157	0,00956	0,02165	0,03339	0,05666	1842,77844
$\mu_7$	0,0492	0,01556	0,00022	0,00047	0,01948	0,03899	0,04899	0,05946	0,08002	1089,19458
$\mu_8$	0,02643	0,0231	0,00033	0,00039	-0,01857	0,01038	0,02645	0,04184	0,07219	3455,0112
$\lambda_1$	0,03954	0,00322	5e-05	0,00016	0,03366	0,03732	0,03944	0,04158	0,04612	410,31134
$\lambda_2$	0,67744	0,04571	0,00065	0,00381	0,58159	0,64692	0,68102	0,71131	0,75263	144,19153

Fonte: Elaborada pelo autor.

Tabela 29 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)-NORMAL.

	<b>HMC</b>	<b>Erro-padrão</b>	<b>ZV-HMC-L</b>	<b>Erro-padrão</b>	<b>ZV-HMC-Q</b>	<b>Erro-padrão</b>
$\omega_1$	0,04139	0,00022	0,04144	0,00001	0,04162	0,00000
$\omega_2$	0,05141	0,00040	0,05127	0,00014	0,05256	0,00004
$\omega_3$	0,05038	0,00032	0,04998	0,00003	0,05063	0,00001
$\omega_4$	0,04474	0,00036	0,04399	0,00009	0,04495	0,00002
$\omega_5$	0,08593	0,00093	0,08554	0,00007	0,08545	0,00003
$\omega_6$	0,06682	0,00047	0,06646	0,00011	0,06801	0,00003
$\omega_7$	0,05186	0,00034	0,05143	0,00004	0,05214	0,00001
$\omega_8$	0,10685	0,00190	0,10576	0,00025	0,10528	0,00008
$\alpha_1$	0,06393	0,00094	0,06445	0,00008	0,06452	0,00003
$\alpha_2$	0,01002	0,00099	0,00959	0,00037	0,01192	0,00009
$\alpha_3$	0,00328	0,00108	0,00378	0,00027	0,00682	0,00009
$\alpha_4$	0,01155	0,00091	0,01073	0,00024	0,01268	0,00005
$\alpha_5$	0,04462	0,00080	0,04333	0,00012	0,04349	0,00003
$\alpha_6$	0,02490	0,00068	0,02437	0,00021	0,02691	0,00006
$\alpha_7$	0,00374	0,00102	0,00424	0,00029	0,00724	0,00011
$\alpha_8$	0,03111	0,00117	0,03120	0,00025	0,03241	0,00015
$\beta_1$	0,79932	0,00077	0,79879	0,00006	0,79762	0,00002
$\beta_2$	0,90807	0,00074	0,90829	0,00033	0,90554	0,00010
$\beta_3$	0,89172	0,00055	0,89196	0,00022	0,88908	0,00007
$\beta_4$	0,88213	0,00100	0,88377	0,00030	0,88072	0,00005
$\beta_5$	0,83091	0,00096	0,83193	0,00005	0,83125	0,00002
$\beta_6$	0,87940	0,00075	0,88005	0,00021	0,87683	0,00006
$\beta_7$	0,89098	0,00052	0,89125	0,00023	0,88830	0,00008
$\beta_8$	0,87499	0,00144	0,87562	0,00019	0,87500	0,00010
$\phi_1$	0,17235	0,00095	0,17292	0,00011	0,16897	0,00003
$\phi_2$	0,07895	0,00063	0,07952	0,00027	0,07685	0,00005
$\phi_3$	0,09285	0,00038	0,09235	0,00016	0,08927	0,00005
$\phi_4$	0,09582	0,00048	0,09598	0,00012	0,09376	0,00003
$\phi_5$	0,15026	0,00101	0,15119	0,00011	0,14854	0,00004
$\phi_6$	0,08528	0,00055	0,08582	0,00017	0,08282	0,00004
$\phi_7$	0,09177	0,00039	0,09124	0,00013	0,08828	0,00006
$\phi_8$	0,08441	0,00098	0,08438	0,00023	0,08102	0,00011
$\mu_1$	0,04645	0,00025	0,04614	0,00006	0,04690	0,00002
$\mu_2$	0,04146	0,00038	0,04093	0,00011	0,04216	0,00004
$\mu_3$	0,04839	0,00047	0,04774	0,00007	0,04850	0,00004
$\mu_4$	0,01054	0,00031	0,01016	0,00008	0,01104	0,00003
$\mu_5$	0,03910	0,00036	0,03880	0,00006	0,03975	0,00003
$\mu_6$	0,02166	0,00041	0,02092	0,00011	0,02227	0,00004
$\mu_7$	0,04920	0,00047	0,04858	0,00007	0,04935	0,00004
$\mu_8$	0,02643	0,00039	0,02625	0,00008	0,02756	0,00003
$\lambda_1$	0,03954	0,00016	0,03943	0,00002	0,04011	0,00001
$\lambda_2$	0,67744	0,00381	0,67928	0,00075	0,66863	0,00021

Fonte: Elaborada pelo autor.

Tabela 30 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)-NORMAL.

	HMC	Erro-padrão	ZV-HMC-L	Erro-padrão	ZV-HMC-Q	Erro-padrão
$\Sigma_{12}$	0,62475	0,00037	0,62489	0,00005	0,62417	0,00002
$\Sigma_{13}$	0,25171	0,00036	0,25218	0,00006	0,25140	0,00002
$\Sigma_{14}$	0,59961	0,00049	0,59998	0,00004	0,59929	0,00002
$\Sigma_{15}$	0,20240	0,00027	0,20247	0,00004	0,20198	0,00001
$\Sigma_{16}$	0,63360	0,00037	0,63387	0,00005	0,63313	0,00002
$\Sigma_{17}$	0,25202	0,00036	0,25252	0,00006	0,25171	0,00002
$\Sigma_{18}$	0,48073	0,00034	0,48065	0,00005	0,48016	0,00002
$\Sigma_{23}$	0,39842	0,00033	0,39919	0,00006	0,39892	0,00002
$\Sigma_{24}$	0,80677	0,00027	0,80710	0,00003	0,80712	0,00001
$\Sigma_{25}$	0,29800	0,00027	0,29821	0,00004	0,29815	0,00002
$\Sigma_{26}$	0,92199	0,00014	0,92210	0,00001	0,92219	0,00001
$\Sigma_{27}$	0,39663	0,00033	0,39738	0,00005	0,39707	0,00002
$\Sigma_{28}$	0,36591	0,00035	0,36583	0,00004	0,36551	0,00002
$\Sigma_{34}$	0,39430	0,00027	0,39476	0,00006	0,39450	0,00002
$\Sigma_{35}$	0,31019	0,00021	0,31050	0,00004	0,31082	0,00001
$\Sigma_{36}$	0,39504	0,00033	0,39590	0,00006	0,39580	0,00002
$\Sigma_{37}$	0,99335	0,00002	0,99337	0,00000	0,99341	0,00000
$\Sigma_{38}$	0,23575	0,00027	0,23575	0,00004	0,23561	0,00002
$\Sigma_{45}$	0,30358	0,00025	0,30386	0,00005	0,30387	0,00002
$\Sigma_{46}$	0,84170	0,00034	0,84202	0,00002	0,84219	0,00001
$\Sigma_{47}$	0,39160	0,00028	0,39205	0,00006	0,39178	0,00002
$\Sigma_{48}$	0,39585	0,00033	0,39593	0,00005	0,39567	0,00002
$\Sigma_{56}$	0,30259	0,00027	0,30290	0,00005	0,30288	0,00002
$\Sigma_{57}$	0,30973	0,00021	0,31005	0,00004	0,31037	0,00001
$\Sigma_{58}$	0,12186	0,00025	0,12170	0,00006	0,12135	0,00001
$\Sigma_{67}$	0,39249	0,00033	0,39332	0,00006	0,39318	0,00002
$\Sigma_{68}$	0,38990	0,00035	0,38985	0,00005	0,38963	0,00002
$\Sigma_{78}$	0,23785	0,00026	0,23785	0,00005	0,23767	0,00002

Fonte: Elaborada pelo autor.

Tabela 31 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-t.

	Média	SD	Naive SE	TM SE	2,5%	25%	50%	75%	97,5%	n_eff
$\omega_1$	0,02793	0,00339	5e-05	2e-04	0,02184	0,02552	0,02776	0,0301	0,03503	300,83376
$\omega_2$	0,04067	0,0055	8e-05	0,00041	0,03069	0,03705	0,04036	0,04404	0,05218	176,39198
$\omega_3$	0,06509	0,0084	0,00012	0,00071	0,05048	0,05917	0,06437	0,07026	0,08362	141,77045
$\omega_4$	0,03321	0,00522	7e-05	0,00049	0,02416	0,02952	0,03286	0,03631	0,04475	115,71005
$\omega_5$	0,06762	0,01272	0,00018	0,00105	0,04731	0,05856	0,06591	0,07563	0,09628	146,98236
$\omega_6$	0,0513	0,00625	9e-05	0,00046	0,03985	0,04694	0,0511	0,05535	0,06456	187,61176
$\omega_7$	0,06755	0,00887	0,00013	0,00075	0,05185	0,06127	0,06682	0,07301	0,08729	139,71553
$\omega_8$	0,08928	0,01658	0,00023	0,00149	0,06048	0,07732	0,08838	0,10031	0,12352	123,73208
$\alpha_1$	0,01557	0,0102	0,00014	0,00212	0,00242	0,00721	0,01379	0,02149	0,04086	23,10203
$\alpha_2$	0,00707	0,00473	7e-05	0,00099	0,00044	0,00354	0,00642	0,0099	0,01786	22,84971
$\alpha_3$	0,01743	0,00624	9e-05	0,00108	0,00687	0,01211	0,01759	0,02163	0,03039	33,45375
$\alpha_4$	0,01044	0,00651	9e-05	0,00126	0,00166	0,00555	0,00908	0,01367	0,026	26,58887
$\alpha_5$	0,03086	0,01121	0,00016	0,00144	0,01072	0,02276	0,03045	0,03849	0,05331	60,85993
$\alpha_6$	0,01438	0,00502	7e-05	6e-04	0,00554	0,01068	0,01421	0,01779	0,0249	70,99994
$\alpha_7$	0,0168	0,00604	9e-05	0,00109	0,00686	0,0119	0,01698	0,02094	0,02921	30,55192
$\alpha_8$	0,02651	0,01008	0,00014	0,00134	0,00901	0,01911	0,02593	0,03276	0,04907	56,16436
$\beta_1$	0,85718	0,0132	0,00019	0,00122	0,82988	0,84855	0,85805	0,86639	0,88154	116,37288
$\beta_2$	0,92164	0,00808	0,00011	0,00065	0,90395	0,91669	0,92206	0,92696	0,93636	153,76262
$\beta_3$	0,87367	0,01272	0,00018	0,00106	0,84536	0,86537	0,87507	0,88284	0,89565	143,60266
$\beta_4$	0,90327	0,01208	0,00017	0,00146	0,87684	0,89581	0,90438	0,9117	0,9248	68,26374
$\beta_5$	0,85385	0,01641	0,00023	0,00119	0,81868	0,84345	0,8549	0,86525	0,88398	191,24396
$\beta_6$	0,9007	0,00886	0,00013	0,00066	0,88166	0,89526	0,90093	0,90676	0,91667	181,43652
$\beta_7$	0,8719	0,01316	0,00019	0,0011	0,84293	0,86345	0,87348	0,88119	0,89528	144,30417
$\beta_8$	0,89405	0,01308	0,00019	0,0011	0,86766	0,88506	0,89461	0,90373	0,91681	142,19577
$\phi_1$	0,16261	0,01751	0,00025	0,00087	0,12998	0,15097	0,1618	0,17388	0,19941	406,18603
$\phi_2$	0,06582	0,00838	0,00012	0,00055	0,04973	0,0601	0,06565	0,07127	0,08269	231,30688
$\phi_3$	0,06325	0,01085	0,00015	0,00074	0,04354	0,05574	0,06241	0,07046	0,08613	214,8054
$\phi_4$	0,07711	0,01038	0,00015	0,00083	0,05757	0,06997	0,0771	0,08403	0,09767	157,7161
$\phi_5$	0,14976	0,02267	0,00032	0,00136	0,10797	0,13365	0,14872	0,16567	0,19579	279,11039
$\phi_6$	0,07568	0,00963	0,00014	0,00064	0,05774	0,06908	0,07548	0,08204	0,09506	227,77448
$\phi_7$	0,06478	0,01089	0,00015	0,00078	0,04514	0,05722	0,0639	0,07205	0,0876	194,12538
$\phi_8$	0,07331	0,01606	0,00023	0,00115	0,04387	0,0625	0,07237	0,08336	0,10819	194,13448
$\mu_1$	0,06274	0,01152	0,00016	0,00021	0,0399	0,05496	0,06273	0,07053	0,08506	3112,69695
$\mu_2$	0,06816	0,0171	0,00024	0,00037	0,03405	0,05682	0,06818	0,0796	0,10181	2110,61506
$\mu_3$	0,06841	0,01526	0,00022	0,00026	0,03841	0,05826	0,06839	0,07896	0,09815	3559,4167
$\mu_4$	0,03414	0,01351	0,00019	3e-04	0,00821	0,02476	0,03413	0,04332	0,06064	2064,22051
$\mu_5$	0,05959	0,01829	0,00026	0,00043	0,02257	0,04711	0,05941	0,0718	0,09635	1838,63071
$\mu_6$	0,04805	0,01688	0,00024	0,00032	0,01493	0,03667	0,04804	0,05945	0,08048	2777,35436
$\mu_7$	0,06879	0,01546	0,00022	0,00026	0,03862	0,05853	0,06868	0,07926	0,09903	3452,69322
$\mu_8$	0,04629	0,02191	0,00031	0,00034	0,0032	0,03164	0,0465	0,0611	0,0891	4100,30705
$\lambda_1$	0,03269	0,00352	5e-05	0,00018	0,02643	0,03029	0,03249	0,03491	0,04018	364,60514
$\lambda_2$	0,73667	0,05071	0,00072	0,00535	0,61901	0,70711	0,74461	0,77356	0,81365	89,96555
$\nu$	9,40211	0,50647	0,00716	0,01635	8,46522	9,04486	9,38041	9,73242	10,4123	959,36477

Fonte: Elaborada pelo autor.

Tabela 32 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-t.

	<b>Média</b>	<b>SD</b>	<b>Naive SE</b>	<b>TM SE</b>	<b>2,5%</b>	<b>25%</b>	<b>50%</b>	<b>75%</b>	<b>97,5%</b>	<b>n_eff</b>
$\Sigma_{12}$	0,5993	0,01418	2e-04	0,00041	0,57041	0,58993	0,59991	0,60888	0,62619	1178,46956
$\Sigma_{13}$	0,22282	0,02106	3e-04	0,00031	0,18104	0,20844	0,22344	0,2372	0,26307	4694,75744
$\Sigma_{14}$	0,57744	0,01464	0,00021	0,00051	0,54846	0,56748	0,57768	0,58767	0,60474	829,88592
$\Sigma_{15}$	0,16645	0,0218	0,00031	0,00023	0,12316	0,1522	0,16638	0,18133	0,20913	9154,14654
$\Sigma_{16}$	0,61069	0,01375	0,00019	0,00035	0,58296	0,60149	0,61116	0,62007	0,63704	1571,66552
$\Sigma_{17}$	0,22333	0,02103	3e-04	0,00031	0,18145	0,20908	0,22389	0,23761	0,26372	4732,00116
$\Sigma_{18}$	0,47741	0,0168	0,00024	0,00027	0,44407	0,46616	0,4775	0,48867	0,50984	3824,05645
$\Sigma_{23}$	0,36697	0,01873	0,00026	0,00024	0,32934	0,35464	0,36725	0,37959	0,40319	5902,40698
$\Sigma_{24}$	0,79409	0,00825	0,00012	0,00023	0,77776	0,7886	0,79424	0,7998	0,80988	1267,0058
$\Sigma_{25}$	0,26105	0,02091	3e-04	0,00021	0,22019	0,24686	0,26124	0,27531	0,30152	9799,1958
$\Sigma_{26}$	0,91593	0,0036	5e-05	0,00013	0,90883	0,91347	0,91599	0,91845	0,92282	740,58128
$\Sigma_{27}$	0,36675	0,0187	0,00026	0,00024	0,32893	0,35435	0,36701	0,37939	0,40227	6070,17034
$\Sigma_{28}$	0,34883	0,0194	0,00027	0,00035	0,31055	0,33549	0,34921	0,36219	0,38565	3077,11939
$\Sigma_{34}$	0,36465	0,02048	0,00029	0,00026	0,324	0,3511	0,36533	0,37845	0,40349	6364,54797
$\Sigma_{35}$	0,28172	0,02071	0,00029	0,00022	0,24058	0,26748	0,28216	0,29585	0,32146	8533,43872
$\Sigma_{36}$	0,36413	0,01969	0,00028	0,00027	0,32453	0,35076	0,36455	0,37753	0,40211	5302,80197
$\Sigma_{37}$	0,99382	0,00028	0	1e-05	0,99325	0,99364	0,99383	0,99401	0,99435	1275,72921
$\Sigma_{38}$	0,20688	0,0219	0,00031	0,00026	0,1635	0,19233	0,20726	0,22195	0,24896	7200,2605
$\Sigma_{45}$	0,26577	0,02037	0,00029	0,00021	0,22531	0,25228	0,26571	0,27956	0,30467	9097,00482
$\Sigma_{46}$	0,83258	0,00704	1e-04	0,00023	0,81861	0,82783	0,83266	0,83737	0,84628	964,11351
$\Sigma_{47}$	0,36353	0,02042	0,00029	0,00025	0,32263	0,34987	0,36422	0,3773	0,40266	6454,6983
$\Sigma_{48}$	0,37595	0,01937	0,00027	0,00033	0,3374	0,36299	0,37664	0,38944	0,41256	3391,51682
$\Sigma_{56}$	0,26376	0,02004	0,00028	0,00022	0,22511	0,25054	0,26364	0,27752	0,30234	8476,87738
$\Sigma_{57}$	0,2809	0,0208	0,00029	0,00022	0,24023	0,26648	0,28141	0,29522	0,32098	8674,41956
$\Sigma_{58}$	0,0969	0,02233	0,00032	0,00018	0,05305	0,08218	0,09685	0,11213	0,14106	15890,46628
$\Sigma_{67}$	0,3617	0,01963	0,00028	0,00027	0,32221	0,34825	0,36227	0,37501	0,3995	5398,22791
$\Sigma_{68}$	0,37352	0,01892	0,00027	0,00033	0,33466	0,36121	0,37375	0,38631	0,40977	3333,81995
$\Sigma_{78}$	0,20926	0,02196	0,00031	0,00024	0,16564	0,19476	0,20959	0,22418	0,25172	8559,49902

Fonte: Elaborada pelo autor.

Tabela 33 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)-t.

	<b>HMC</b>	<b>Erro-padrão</b>	<b>ZV-HMC-L</b>	<b>Erro-padrão</b>	<b>ZV-HMC-Q</b>	<b>Erro-padrão</b>
$\omega_1$	0,02793	0,00020	0,02807	0,00009	0,02857	0,00004
$\omega_2$	0,04067	0,00041	0,04037	0,00013	0,04173	0,00002
$\omega_3$	0,06509	0,00071	0,06538	0,00012	0,06638	0,00004
$\omega_4$	0,03321	0,00049	0,03300	0,00009	0,03422	0,00004
$\omega_5$	0,06762	0,00105	0,06822	0,00013	0,06849	0,00005
$\omega_6$	0,05130	0,00046	0,05096	0,00010	0,05257	0,00003
$\omega_7$	0,06755	0,00075	0,06786	0,00013	0,06887	0,00004
$\omega_8$	0,08928	0,00149	0,08824	0,00016	0,08809	0,00007
$\alpha_1$	0,01557	0,00212	0,01766	0,00085	0,02223	0,00041
$\alpha_2$	0,00707	0,00099	0,00653	0,00040	0,01010	0,00005
$\alpha_3$	0,01743	0,00108	0,01725	0,00035	0,02040	0,00014
$\alpha_4$	0,01044	0,00126	0,01093	0,00033	0,01473	0,00010
$\alpha_5$	0,03086	0,00144	0,03154	0,00033	0,03408	0,00013
$\alpha_6$	0,01438	0,00060	0,01333	0,00025	0,01795	0,00012
$\alpha_7$	0,01680	0,00109	0,01662	0,00035	0,01976	0,00013
$\alpha_8$	0,02651	0,00134	0,02743	0,00044	0,02960	0,00011
$\beta_1$	0,85718	0,00122	0,85555	0,00081	0,85084	0,00032
$\beta_2$	0,92164	0,00065	0,92214	0,00030	0,91872	0,00004
$\beta_3$	0,87367	0,00106	0,87325	0,00027	0,86934	0,00011
$\beta_4$	0,90327	0,00146	0,90354	0,00035	0,89920	0,00012
$\beta_5$	0,85385	0,00119	0,85301	0,00011	0,85106	0,00005
$\beta_6$	0,90070	0,00066	0,90160	0,00021	0,89721	0,00009
$\beta_7$	0,87190	0,00110	0,87147	0,00026	0,86754	0,00011
$\beta_8$	0,89405	0,00110	0,89452	0,00024	0,89308	0,00010
$\phi_1$	0,16261	0,00087	0,16216	0,00026	0,15712	0,00012
$\phi_2$	0,06582	0,00055	0,06652	0,00028	0,06225	0,00005
$\phi_3$	0,06325	0,00074	0,06394	0,00028	0,06081	0,00012
$\phi_4$	0,07711	0,00083	0,07636	0,00022	0,07214	0,00009
$\phi_5$	0,14976	0,00136	0,14912	0,00037	0,14293	0,00014
$\phi_6$	0,07568	0,00064	0,07672	0,00023	0,07148	0,00011
$\phi_7$	0,06478	0,00078	0,06547	0,00027	0,06236	0,00012
$\phi_8$	0,07331	0,00115	0,07170	0,00037	0,06746	0,00010
$\mu_1$	0,06274	0,00021	0,06278	0,00004	0,06356	0,00003
$\mu_2$	0,06816	0,00037	0,06801	0,00007	0,06914	0,00004
$\mu_3$	0,06841	0,00026	0,06813	0,00005	0,06846	0,00003
$\mu_4$	0,03414	0,00030	0,03405	0,00005	0,03486	0,00003
$\mu_5$	0,05959	0,00043	0,05998	0,00007	0,06076	0,00004
$\mu_6$	0,04805	0,00032	0,04785	0,00007	0,04901	0,00005
$\mu_7$	0,06879	0,00026	0,06859	0,00005	0,06890	0,00003
$\mu_8$	0,04629	0,00034	0,04665	0,00007	0,04794	0,00003
$\lambda_1$	0,03269	0,00018	0,03279	0,00004	0,03267	0,00002
$\lambda_2$	0,73667	0,00535	0,73429	0,00158	0,73409	0,00072
$\nu$	9,40211	0,01635	9,41671	0,00339	9,56327	0,00099

Fonte: Elaborada pelo autor.

Tabela 34 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)-t.

	<b>HMC</b>	<b>Erro-padrão</b>	<b>ZV-HMC-L</b>	<b>Erro-padrão</b>	<b>ZV-HMC-Q</b>	<b>Erro-padrão</b>
$\Sigma_{12}$	0,59930	0,00041	0,59965	0,00005	0,59931	0,00003
$\Sigma_{13}$	0,22282	0,00031	0,22325	0,00006	0,22287	0,00003
$\Sigma_{14}$	0,57744	0,00051	0,57779	0,00005	0,57738	0,00002
$\Sigma_{15}$	0,16645	0,00023	0,16672	0,00005	0,16692	0,00002
$\Sigma_{16}$	0,61069	0,00035	0,61103	0,00005	0,61056	0,00002
$\Sigma_{17}$	0,22333	0,00031	0,22376	0,00006	0,22340	0,00003
$\Sigma_{18}$	0,47741	0,00027	0,47756	0,00004	0,47745	0,00002
$\Sigma_{23}$	0,36697	0,00024	0,36759	0,00006	0,36758	0,00003
$\Sigma_{24}$	0,79409	0,00023	0,79433	0,00004	0,79463	0,00001
$\Sigma_{25}$	0,26105	0,00021	0,26144	0,00006	0,26164	0,00002
$\Sigma_{26}$	0,91593	0,00013	0,91600	0,00002	0,91626	0,00001
$\Sigma_{27}$	0,36675	0,00024	0,36732	0,00006	0,36734	0,00003
$\Sigma_{28}$	0,34883	0,00035	0,34907	0,00006	0,34882	0,00003
$\Sigma_{34}$	0,36465	0,00026	0,36513	0,00008	0,36539	0,00003
$\Sigma_{35}$	0,28172	0,00022	0,28218	0,00008	0,28271	0,00002
$\Sigma_{36}$	0,36413	0,00027	0,36470	0,00008	0,36478	0,00003
$\Sigma_{37}$	0,99382	0,00001	0,99383	0,00000	0,99387	0,00000
$\Sigma_{38}$	0,20688	0,00026	0,20710	0,00007	0,20748	0,00003
$\Sigma_{45}$	0,26577	0,00021	0,26616	0,00007	0,26664	0,00003
$\Sigma_{46}$	0,83258	0,00023	0,83274	0,00003	0,83314	0,00001
$\Sigma_{47}$	0,36353	0,00025	0,36397	0,00008	0,36426	0,00003
$\Sigma_{48}$	0,37595	0,00033	0,37624	0,00006	0,37626	0,00003
$\Sigma_{56}$	0,26376	0,00022	0,26409	0,00007	0,26438	0,00002
$\Sigma_{57}$	0,28090	0,00022	0,28135	0,00008	0,28188	0,00002
$\Sigma_{58}$	0,09690	0,00018	0,09703	0,00007	0,09711	0,00002
$\Sigma_{67}$	0,36170	0,00027	0,36225	0,00008	0,36234	0,00003
$\Sigma_{68}$	0,37352	0,00033	0,37374	0,00007	0,37356	0,00003
$\Sigma_{78}$	0,20926	0,00024	0,20947	0,00007	0,20983	0,00003

Fonte: Elaborada pelo autor.



Tabela 35 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-GED.

	<b>Média</b>	<b>SD</b>	<b>Naive SE</b>	<b>TM SE</b>	<b>2,5%</b>	<b>25%</b>	<b>50%</b>	<b>75%</b>	<b>97,5%</b>	<b>n_eff</b>
$\omega_1$	0,02876	0,00349	5e-05	2e-04	0,02235	0,02632	0,02864	0,03104	0,03606	303,0108
$\omega_2$	0,04059	0,0056	8e-05	0,00044	0,03039	0,03691	0,04027	0,04404	0,05248	163,17058
$\omega_3$	0,06223	0,00782	0,00011	0,00063	0,04866	0,05685	0,06162	0,06691	0,07982	154,07578
$\omega_4$	0,03324	0,00525	7e-05	0,00049	0,02426	0,02945	0,03283	0,03655	0,04473	112,6197
$\omega_5$	0,06621	0,01263	0,00018	0,00106	0,0459	0,05725	0,06464	0,07386	0,09467	142,75944
$\omega_6$	0,052	0,00638	9e-05	0,00047	0,04019	0,04756	0,0518	0,05607	0,06546	180,77516
$\omega_7$	0,06457	0,00825	0,00012	0,00068	0,04986	0,05884	0,06398	0,06961	0,08319	148,17285
$\omega_8$	0,0902	0,01701	0,00024	0,00156	0,05995	0,07825	0,08921	0,101	0,12621	119,0128
$\alpha_1$	0,02246	0,01096	0,00015	0,00193	0,00534	0,0145	0,0211	0,02879	0,04882	32,35611
$\alpha_2$	0,007	0,00497	7e-05	0,00115	0,00039	0,00322	0,00608	0,01013	0,01786	18,56637
$\alpha_3$	0,01531	0,00593	8e-05	0,00118	0,00528	0,0104	0,01514	0,0192	0,02794	25,2453
$\alpha_4$	0,01023	0,00651	9e-05	0,00128	0,00172	0,00527	0,00873	0,01371	0,02552	26,02267
$\alpha_5$	0,03462	0,01157	0,00016	0,00138	0,01413	0,02665	0,03432	0,04241	0,05811	70,74965
$\alpha_6$	0,01532	0,00524	7e-05	0,00067	0,00604	0,01151	0,01512	0,01869	0,0265	60,40474
$\alpha_7$	0,01468	0,00569	8e-05	0,00108	0,00551	0,01016	0,01455	0,01845	0,02684	27,57626
$\alpha_8$	0,02674	0,01034	0,00015	0,0014	0,00859	0,0192	0,02584	0,0335	0,0493	54,56662
$\beta_1$	0,85224	0,01352	0,00019	0,00128	0,82476	0,84334	0,85286	0,86137	0,87768	111,92013
$\beta_2$	0,92169	0,00835	0,00012	0,00071	0,90333	0,91658	0,92209	0,92732	0,93677	137,28118
$\beta_3$	0,87705	0,01189	0,00017	0,00096	0,85098	0,8694	0,87827	0,88562	0,89793	154,50561
$\beta_4$	0,90344	0,01222	0,00017	0,00151	0,87651	0,89562	0,90465	0,91226	0,92463	65,53521
$\beta_5$	0,85504	0,01616	0,00023	0,00122	0,82056	0,84495	0,85605	0,86604	0,88444	175,66782
$\beta_6$	0,89969	0,00902	0,00013	0,00068	0,88014	0,89419	0,89994	0,90569	0,91649	178,09201
$\beta_7$	0,87536	0,01229	0,00017	0,00098	0,84831	0,86751	0,87655	0,88405	0,89713	156,92381
$\beta_8$	0,89264	0,01348	0,00019	0,00115	0,86436	0,88382	0,89336	0,90232	0,91639	136,84917
$\phi_1$	0,15777	0,01795	0,00025	0,00091	0,12429	0,14565	0,15722	0,16924	0,19492	385,25204
$\phi_2$	0,06544	0,00832	0,00012	0,00063	0,04964	0,0598	0,06524	0,07082	0,08207	176,27175
$\phi_3$	0,06586	0,01064	0,00015	0,00075	0,04603	0,05871	0,06523	0,07284	0,08792	203,37638
$\phi_4$	0,07614	0,01017	0,00014	8e-04	0,05675	0,06915	0,07618	0,08316	0,09642	160,6272
$\phi_5$	0,143	0,02209	0,00031	0,0013	0,10243	0,12748	0,14193	0,15819	0,1877	288,71434
$\phi_6$	0,07449	0,00968	0,00014	0,00066	0,05613	0,06801	0,07426	0,08101	0,09401	213,12964
$\phi_7$	0,06723	0,01061	0,00015	0,00076	0,04796	0,05994	0,06661	0,07421	0,08922	197,22479
$\phi_8$	0,07392	0,01635	0,00023	0,00108	0,04435	0,06313	0,07272	0,08392	0,10956	230,25854
$\mu_1$	0,05755	0,01106	0,00016	0,00021	0,0357	0,05016	0,0575	0,06487	0,07962	2699,5665
$\mu_2$	0,06163	0,01635	0,00023	0,00036	0,02952	0,05058	0,06192	0,07303	0,09251	2050,65388
$\mu_3$	0,06563	0,01459	0,00021	0,00028	0,03784	0,05583	0,06556	0,07539	0,0945	2679,16206
$\mu_4$	0,03071	0,01297	0,00018	4e-04	0,00615	0,02176	0,03082	0,03916	0,05704	1052,97369
$\mu_5$	0,05713	0,01816	0,00026	0,00045	0,02059	0,0451	0,05671	0,06944	0,09342	1619,03652
$\mu_6$	0,04265	0,01613	0,00023	3e-04	0,01081	0,03187	0,04259	0,05321	0,07436	2955,28036
$\mu_7$	0,0663	0,0147	0,00021	0,00028	0,03788	0,05647	0,06624	0,07607	0,09483	2733,10847
$\mu_8$	0,04103	0,02135	3e-04	0,00034	-0,00063	0,02652	0,04083	0,05567	0,083	4049,03777
$\lambda_1$	0,0334	0,00355	5e-05	0,00019	0,02699	0,03092	0,03324	0,03564	0,04087	367,35594
$\lambda_2$	0,72875	0,05169	0,00073	0,00546	0,61256	0,69795	0,73588	0,7667	0,8082	89,71588
$\nu$	0,4491	0,01125	0,00016	0,00012	0,42726	0,44166	0,44891	0,45634	0,47179	9027,07587

Fonte: Elaborada pelo autor.

Tabela 36 – Resultado das simulações da posteriori para DCC-GJR-GARCH(1,1)-GED.

	<b>Média</b>	<b>SD</b>	<b>Naive SE</b>	<b>TM SE</b>	<b>2,5%</b>	<b>25%</b>	<b>50%</b>	<b>75%</b>	<b>97,5%</b>	<b>n_eff</b>
$\Sigma_{12}$	0,59922	0,01433	2e-04	0,00042	0,5705	0,5896	0,59955	0,60896	0,62635	1163,7087
$\Sigma_{13}$	0,22344	0,02082	0,00029	0,00033	0,18236	0,20943	0,22403	0,23785	0,26279	4064,7561
$\Sigma_{14}$	0,57708	0,01489	0,00021	0,00052	0,54757	0,56719	0,57731	0,58728	0,60553	809,72132
$\Sigma_{15}$	0,16695	0,02172	0,00031	0,00022	0,12283	0,15261	0,16697	0,1816	0,20944	9467,79876
$\Sigma_{16}$	0,6106	0,01413	2e-04	0,00036	0,5819	0,6011	0,61091	0,61991	0,63793	1556,75696
$\Sigma_{17}$	0,22402	0,0208	0,00029	0,00032	0,18258	0,21007	0,2244	0,23826	0,26372	4125,61867
$\Sigma_{18}$	0,47491	0,0171	0,00024	0,00025	0,44164	0,46341	0,47516	0,48633	0,50888	4535,16098
$\Sigma_{23}$	0,36688	0,01891	0,00027	0,00029	0,32802	0,35475	0,36676	0,37987	0,40344	4266,27871
$\Sigma_{24}$	0,7942	0,00818	0,00012	0,00023	0,77857	0,78853	0,79431	0,79975	0,8099	1249,43228
$\Sigma_{25}$	0,26136	0,02093	3e-04	0,00021	0,22011	0,24727	0,26156	0,27574	0,3021	9964,08346
$\Sigma_{26}$	0,9161	0,00357	5e-05	0,00014	0,90932	0,91371	0,91612	0,91858	0,92309	657,26806
$\Sigma_{27}$	0,36646	0,01889	0,00027	0,00029	0,32735	0,35411	0,36654	0,37904	0,40323	4386,922
$\Sigma_{28}$	0,34803	0,01907	0,00027	0,00034	0,31134	0,33451	0,34812	0,36115	0,38432	3232,87381
$\Sigma_{34}$	0,36442	0,02042	0,00029	0,00032	0,32271	0,35135	0,3645	0,37817	0,40557	4033,17025
$\Sigma_{35}$	0,28109	0,02008	0,00028	0,00022	0,24118	0,26765	0,28107	0,29479	0,32035	8508,69679
$\Sigma_{36}$	0,36435	0,01952	0,00028	0,00029	0,32397	0,35177	0,36493	0,37757	0,40188	4572,43637
$\Sigma_{37}$	0,9937	0,00029	0	1e-05	0,9931	0,99351	0,9937	0,9939	0,99423	1159,46419
$\Sigma_{38}$	0,20747	0,02176	0,00031	0,00024	0,16403	0,19349	0,20777	0,22179	0,2495	8144,19827
$\Sigma_{45}$	0,26562	0,02094	3e-04	0,00021	0,22299	0,25207	0,26567	0,2794	0,30635	9920,42203
$\Sigma_{46}$	0,83265	0,00694	1e-04	0,00023	0,81895	0,82794	0,83276	0,83731	0,84634	945,68798
$\Sigma_{47}$	0,36312	0,02034	0,00029	0,00032	0,32234	0,35035	0,3632	0,37671	0,40364	4040,99737
$\Sigma_{48}$	0,3757	0,0195	0,00028	0,00033	0,33723	0,36229	0,37661	0,3892	0,41232	3471,14067
$\Sigma_{56}$	0,26475	0,02059	0,00029	0,00021	0,22324	0,25122	0,26489	0,2785	0,30485	9939,82816
$\Sigma_{57}$	0,28046	0,02013	0,00028	0,00022	0,24088	0,26699	0,28063	0,29421	0,31952	8449,38472
$\Sigma_{58}$	0,09699	0,02281	0,00032	0,00016	0,05213	0,0821	0,09729	0,11205	0,14252	20815,6531
$\Sigma_{67}$	0,36185	0,01953	0,00028	0,00029	0,3216	0,34939	0,36206	0,37499	0,39901	4591,46563
$\Sigma_{68}$	0,37276	0,01857	0,00026	0,00032	0,33566	0,36039	0,37293	0,38537	0,40873	3472,46171
$\Sigma_{78}$	0,20988	0,02182	0,00031	0,00024	0,1661	0,19593	0,2102	0,22432	0,2516	8443,6184

Fonte: Elaborada pelo autor.

Tabela 37 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)-GED.

	HMC	Erro-padrão	ZV-HMC-L	Erro-padrão	ZV-HMC-Q	Erro-padrão
$\omega_1$	0,02876	0,00020	0,02888	0,00006	0,02944	0,00003
$\omega_2$	0,04059	0,00044	0,04031	0,00014	0,04176	0,00003
$\omega_3$	0,06223	0,00063	0,06250	0,00012	0,06377	0,00003
$\omega_4$	0,03324	0,00049	0,03300	0,00012	0,03412	0,00004
$\omega_5$	0,06621	0,00106	0,06699	0,00013	0,06788	0,00009
$\omega_6$	0,05200	0,00047	0,05171	0,00009	0,05352	0,00003
$\omega_7$	0,06457	0,00068	0,06485	0,00013	0,06614	0,00004
$\omega_8$	0,09020	0,00156	0,08944	0,00021	0,08957	0,00006
$\alpha_1$	0,02246	0,00193	0,02419	0,00065	0,02802	0,00021
$\alpha_2$	0,00700	0,00115	0,00666	0,00043	0,01016	0,00006
$\alpha_3$	0,01531	0,00118	0,01511	0,00039	0,01849	0,00015
$\alpha_4$	0,01023	0,00128	0,01093	0,00037	0,01450	0,00010
$\alpha_5$	0,03462	0,00138	0,03539	0,00024	0,03708	0,00012
$\alpha_6$	0,01532	0,00067	0,01457	0,00023	0,01923	0,00008
$\alpha_7$	0,01468	0,00108	0,01446	0,00038	0,01780	0,00014
$\alpha_8$	0,02674	0,00140	0,02759	0,00043	0,02920	0,00013
$\beta_1$	0,85224	0,00128	0,85089	0,00057	0,84670	0,00022
$\beta_2$	0,92169	0,00071	0,92213	0,00033	0,91869	0,00005
$\beta_3$	0,87705	0,00096	0,87673	0,00035	0,87252	0,00012
$\beta_4$	0,90344	0,00151	0,90371	0,00038	0,89969	0,00012
$\beta_5$	0,85504	0,00122	0,85408	0,00010	0,85211	0,00004
$\beta_6$	0,89969	0,00068	0,90045	0,00021	0,89594	0,00007
$\beta_7$	0,87536	0,00098	0,87506	0,00034	0,87083	0,00012
$\beta_8$	0,89264	0,00115	0,89301	0,00027	0,89178	0,00009
$\phi_1$	0,15777	0,00091	0,15745	0,00021	0,15243	0,00007
$\phi_2$	0,06544	0,00063	0,06589	0,00029	0,06174	0,00005
$\phi_3$	0,06586	0,00075	0,06653	0,00029	0,06331	0,00012
$\phi_4$	0,07614	0,00080	0,07515	0,00021	0,07106	0,00007
$\phi_5$	0,14300	0,00130	0,14208	0,00023	0,13634	0,00013
$\phi_6$	0,07449	0,00066	0,07520	0,00022	0,06991	0,00008
$\phi_7$	0,06723	0,00076	0,06789	0,00027	0,06476	0,00012
$\phi_8$	0,07392	0,00108	0,07232	0,00035	0,06854	0,00012
$\mu_1$	0,05755	0,00021	0,05768	0,00003	0,05842	0,00002
$\mu_2$	0,06163	0,00036	0,06172	0,00005	0,06278	0,00004
$\mu_3$	0,06563	0,00028	0,06546	0,00003	0,06580	0,00002
$\mu_4$	0,03071	0,00040	0,03062	0,00004	0,03140	0,00003
$\mu_5$	0,05713	0,00045	0,05751	0,00007	0,05829	0,00005
$\mu_6$	0,04265	0,00030	0,04261	0,00006	0,04377	0,00004
$\mu_7$	0,06630	0,00028	0,06611	0,00003	0,06644	0,00002
$\mu_8$	0,04103	0,00034	0,04145	0,00006	0,04276	0,00003
$\lambda_1$	0,03340	0,00019	0,03352	0,00003	0,03357	0,00002
$\lambda_2$	0,72875	0,00546	0,72606	0,00141	0,72472	0,00055
$\nu$	0,44910	0,00012	0,44914	0,00009	0,45220	0,00003

Fonte: Elaborada pelo autor.

Tabela 38 – Estimativas HMC, ZV-HMC-L e ZV-HMC-Q para DCC-GJR-GARCH(1,1)-GED.

	<b>HMC</b>	<b>Erro-padrão</b>	<b>ZV-HMC-L</b>	<b>Erro-padrão</b>	<b>ZV-HMC-Q</b>	<b>Erro-padrão</b>
$\Sigma_{12}$	0,59922	0,00042	0,59962	0,00005	0,59924	0,00003
$\Sigma_{13}$	0,22344	0,00033	0,22381	0,00006	0,22341	0,00004
$\Sigma_{14}$	0,57708	0,00052	0,57747	0,00005	0,57708	0,00003
$\Sigma_{15}$	0,16695	0,00022	0,16716	0,00005	0,16721	0,00003
$\Sigma_{16}$	0,61060	0,00036	0,61100	0,00005	0,61050	0,00003
$\Sigma_{17}$	0,22402	0,00032	0,22443	0,00006	0,22403	0,00004
$\Sigma_{18}$	0,47491	0,00025	0,47502	0,00004	0,47479	0,00002
$\Sigma_{23}$	0,36688	0,00029	0,36747	0,00006	0,36742	0,00003
$\Sigma_{24}$	0,79420	0,00023	0,79443	0,00004	0,79470	0,00002
$\Sigma_{25}$	0,26136	0,00021	0,26173	0,00006	0,26174	0,00003
$\Sigma_{26}$	0,91610	0,00014	0,91618	0,00002	0,91643	0,00001
$\Sigma_{27}$	0,36646	0,00029	0,36702	0,00006	0,36699	0,00003
$\Sigma_{28}$	0,34803	0,00034	0,34836	0,00006	0,34810	0,00003
$\Sigma_{34}$	0,36442	0,00032	0,36489	0,00008	0,36512	0,00003
$\Sigma_{35}$	0,28109	0,00022	0,28152	0,00006	0,28200	0,00003
$\Sigma_{36}$	0,36435	0,00029	0,36493	0,00008	0,36500	0,00003
$\Sigma_{37}$	0,99370	0,00001	0,99371	0,00000	0,99375	0,00000
$\Sigma_{38}$	0,20747	0,00024	0,20760	0,00007	0,20799	0,00003
$\Sigma_{45}$	0,26562	0,00021	0,26591	0,00006	0,26624	0,00003
$\Sigma_{46}$	0,83265	0,00023	0,83281	0,00003	0,83320	0,00002
$\Sigma_{47}$	0,36312	0,00032	0,36358	0,00008	0,36383	0,00003
$\Sigma_{48}$	0,37570	0,00033	0,37605	0,00006	0,37604	0,00003
$\Sigma_{56}$	0,26475	0,00021	0,26502	0,00006	0,26516	0,00003
$\Sigma_{57}$	0,28046	0,00022	0,28088	0,00007	0,28136	0,00003
$\Sigma_{58}$	0,09699	0,00016	0,09714	0,00007	0,09720	0,00002
$\Sigma_{67}$	0,36185	0,00029	0,36242	0,00008	0,36250	0,00003
$\Sigma_{68}$	0,37276	0,00032	0,37306	0,00006	0,37290	0,00003
$\Sigma_{78}$	0,20988	0,00024	0,20999	0,00007	0,21037	0,00003

Fonte: Elaborada pelo autor.