



Universidade de Brasília  
Instituto de Ciências Exatas  
Departamento de Estatística

Dissertação de Mestrado

MODELOS COMBINADOS AR-GARCH GOVERNADOS  
POR DISTRIBUIÇÕES ESTÁVEIS

por

Thiago do Rêgo Sousa

Orientadora: Profa. Dra. Cira Etheowalda Guevara Otiniano

Julho de 2013

Thiago do Rêgo Sousa

MODELOS COMBINADOS AR-GARCH GOVERNADOS  
POR DISTRIBUIÇÕES ESTÁVEIS

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

Universidade de Brasília

Brasília, Julho de 2013

TERMO DE APROVAÇÃO

Thiago do Rêgo Sousa

MODELOS COMBINADOS AR-GARCH GOVERNADOS  
POR DISTRIBUIÇÕES ESTÁVEIS

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

Data da defesa: 26 de Julho de 2013

Orientadora:

---

Profa. Cira Etheowalda Guevara Otiniano, PhD  
Departamento de Estatística, UnB

Comissão Examinadora:

---

Profa. Chang Chung Yu Dorea, PhD  
Departamento de Matemática, UnB

---

Prof. Raul Yukihiro Matsushita, PhD  
Departamento de Estatística, UnB

---

Prof. Donald Matthew Pianto, PhD  
Departamento de Estatística, UnB

Brasília, Julho de 2013

## Ficha Catalográfica

**SOUSA, THIAGO**

Modelos Combinados AR-GARCH governados por distribuições estáveis,  
(UnB - IE, Mestre em Estatística, 2013).

Dissertação de Mestrado - Universidade de Brasília.

Departamento de Estatística - Instituto de Ciências Exatas.

1. ARMA
2. GARCH
3. Distribuições estáveis
4. Distribuição GEV
5. Máxima Verossimilhança
6. MCMC
7. Estacionariedade
8. Expoente de Lyapunov.

É concedida à Universidade de Brasília a permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação pode ser reproduzida sem a autorização por escrito do autor.

Thiago do Rêgo Sousa

*A Deus criador de todas  
as coisa e dono  
de todo o conhecimento.*

# Agradecimentos

Agradeço a Deus por me ajudar a trilhar todo o caminho necessário para chegar até aqui.

Aos amigos e familiares por nossos momentos de descontração e alegria, que me ajudaram a continuar firme com esse propósito de concluir o curso de Mestrado: Robson, Adélio, Rodrigo, Caio, Lucas, Jociel e Marcus. Em especial aos meus amigos da PP: Luís, Douglas e Eurico.

A minha mãe, a quem devo minha vida e todas as minhas conquistas. Sem ela eu não estaria aqui, e muito menos teria chegado tão longe.

Ao meu pai que, mesmo sem tantas condições, lutou para me proporcionar uma vida com dignidade. E é por isso que tenho o prazer e orgulho de ter estudado tanto e mostrado a ele o resultado desse esforço. Tenho muito orgulho de ser filho de um homem que luta para viver a cada dia.

Ao meu Irmão Júnior que me ajudou a trilhar esse caminho de sucesso profissional e pessoal. A minha namorada Larizza por estar sempre presente em todos os momentos e também, pelo amor e compreensão durante esse período.

A minha orientadora, Cira, que além de compartilhar seus conhecimentos, se mostrou ser um exemplo singeleza, simplicidade e competência. Também a professora Sílvia Regina que além de ter sido minha coorientadora, mostrou ser uma pessoa de uma generosidade admirável. Ao professor Rathie que compartilhou seu imenso conhecimento acadêmico comigo. Aos professores João Carlos e Baigori com quem compartilhei ótimas conversas neste período.

Ao professor John P.Nolan que gentilmente cedeu seu programa de distribuições estáveis para utilização nesta dissertação.

# Sumário

<b>Lista de Figuras</b>	<b>5</b>
<b>Lista de Tabelas</b>	<b>8</b>
<b>Resumo</b>	<b>9</b>
<b>Abstract</b>	<b>10</b>
<b>Introdução</b>	<b>11</b>
<b>1 Preliminares</b>	<b>16</b>
1.1 Conceitos Básicos de Séries Temporais . . . . .	16
1.2 Outros Conceitos Importantes . . . . .	22
1.3 Distribuições de Probabilidade de Interesse . . . . .	23
1.3.1 Distribuições Normal e t-Student . . . . .	24
1.3.2 Distribuições estáveis . . . . .	24
1.3.3 Distribuição GEV . . . . .	29
<b>2 Modelos de Séries Temporais</b>	<b>31</b>
2.1 Modelos lineares ARMA e ARIMA . . . . .	31
2.2 Modelos não lineares ARCH e suas extensões . . . . .	34
<b>3 Modelos Combinados (ARMA-GARCH)</b>	<b>40</b>
3.1 O modelo GEV-GARCH . . . . .	42
3.2 O modelo STABLE-GARCH e algumas generalizações . . . . .	45
3.3 Estimação de Modelos ARMA-GARCH . . . . .	47

<b>4</b>	<b>Estimação do Modelo GEV-GARCH</b>	<b>54</b>
4.1	Estimação do modelo GEV-GARCH pelo método da Máxima Verossimilhança . . . . .	56
4.2	Estimação Bayesiana do Modelo GEV-GARCH . . . . .	56
<b>5</b>	<b>Estimação do Modelo STABLE-GARCH</b>	<b>63</b>
5.1	Estimação pelo método da Máxima Verossimilhança do modelo STABLE-GARCH . . . . .	66
5.2	Estimação Bayesiana do Modelo STABLE-GARCH . . . . .	67
<b>6</b>	<b>Provas de estacionariedade</b>	<b>72</b>
<b>7</b>	<b>Aspectos Computacionais</b>	<b>83</b>
<b>8</b>	<b>Aplicação com Dados de Finanças</b>	<b>94</b>
8.1	Seleção do modelo . . . . .	98
8.2	Diagnóstico do Modelo . . . . .	102
8.3	Comparação dos modelos GEV-GARCH e STABLE-GARCH . . . . .	105
8.4	Previsão de parâmetros futuros . . . . .	109
<b>9</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>115</b>
9.1	Conclusões . . . . .	115
9.2	Trabalhos Futuros . . . . .	116
	Referências Bibliográficas . . . . .	117



# Lista de Figuras

1.1	Gráfico distribuições Gaussiana, de Cauchy e de Lévy com $\mu = 0$ e $\sigma = 1$ .	27
3.1	Simulação de 1000 observações do modelo AR(1)-GARCH(1,1). Lado esquerdo: Modelo com inovações seguindo uma distribuição GEV com parâmetros $\theta = (0.2, 0.3, 0.2, 0.3, 0.2, 0.2)$ . Lado direito: Modelo com os mesmos parâmetros exceto que a distribuição das inovações segue um modelo Normal.	44
3.2	Simulação de 1000 observações do modelo AR(1)-GARCH(1,1) com inovações GEV e parâmetro $\theta$ .	44
3.3	Simulação de 1000 observações do modelo AR(1)-GARCH(1,1). Lado esquerdo: Inovações seguindo uma distribuição $\alpha$ -estável com parâmetros $\varphi = (0.2, 0.3, 0.2, 0.3, 0.2, 1.8, 0)$ . Lado direito: Modelo com os mesmos parâmetros porém com inovações seguindo uma distribuição Normal.	46
3.4	Simulação de 1000 observações do modelo AR(1)-GARCH(1,1) com inovações estáveis e parâmetro $\varphi$ .	46
4.1	Estimação MCMC com o Algoritmo 9.1 para estimação dos parâmetros de uma série temporal simulada do modelo GEV-GARCH com parâmetros $\theta_1$ .	58
4.2	Convergência da cadeia simulada com o algoritmo 9.1 para simulação da distribuição a posteriori do parâmetro $a_1 = 0.32$ do modelo GEV-GARCH.	59
4.3	Convergência da cadeia simulada utilizando a rotina <code>MCMCmetrop1R</code> para simulação da distribuição a posteriori do parâmetro $a_1 = 0.32$ do modelo GEV-GARCH.	60

5.1	Estimação MCMC com o Algoritmo 9.2 para estimação dos parâmetros de uma série temporal simulada do modelo STABLE-GARCH com parâmetros $\phi_1$ . . . . .	69
5.2	Convergência da cadeia simulada com o algoritmo 9.2 para simulação da distribuição a posteriori do parâmetro $a_1 = 0.32$ do modelo STABLE-GARCH. . . . .	69
5.3	Convergência da cadeia simulada utilizando a rotina <code>MCMCmetrop1R</code> para simulação da distribuição a posteriori do parâmetro $a_1 = 0.32$ do modelo STABLE-GARCH. . . . .	71
8.1	Retornos das séries temporais analisadas. . . . .	95
8.2	Estrutura de autocorrelação para os dados de log-retornos da empresa Microsoft. . . . .	96
8.3	Estrutura de autocorrelação para os dados de log-retornos da empresa Google. . . . .	96
8.4	Estrutura de autocorrelação para os dados de log-retornos do índice S&P500. . . . .	97
8.5	Autocorrelação para os resíduos ajustados do modelo ARMA(0,1)-GARCH(1,1) com resíduos estáveis para os dados da Microsoft. . . .	104
8.6	Autocorrelação para os resíduos ajustados do modelo ARMA(2,2)-GARCH(1,2) com resíduos GEV para os dados da Microsoft. . . . .	104
8.7	Quantis de 95% estimados dos retornos máximos da série S&P500. Os pontos sólidos representam os quantis reais. Gráfico acima: As linhas azul e cinza representam os quantis estimados através dos modelos GEV e normal. Gráfico abaixo: As linhas amarelo e cinza representam os quantis estimados através dos modelos estável e normal. . . . .	107
8.8	Quantis de 95% estimados dos retornos máximos da série S&P500. Os pontos sólidos representam os quantis reais e os gráficos azul e amarelo representam os quantis estimados através dos modelos GEV e estável, respectivamente. . . . .	108
8.9	Simulação e previsão da volatilidade no modelo GEV-GARCH descrito na Tabela 9.18 do Apêndice. . . . .	113

8.10	Simulação e previsão da volatilidade no modelo STABLE-GARCH descrito na Tabela 9.21 do Apêndice. . . . .	113
9.1	MCMC GEV-GARCH com $\theta = (0.01, 0.45, 0.08, 0.21, 0.32, 0.08)$ . . . . .	122
9.2	MCMC GEV-GARCH com $\theta = (0.01, 0.8, 0.02, 0.01, 0.1, 0.2)$ . . . . .	123
9.3	MCMC GEV-GARCH com $\theta = (0.05, 0.5, 0.1, 0.05, 0.2, 0.3)$ . . . . .	124
9.4	MCMC STABLE-GARCH com $\varphi_1 = (0.01, 0.45, 0.08, 0.21, 0.32, 1.8, 0.5)$ . . . . .	132
9.5	MCMC STABLE-GARCH com $\varphi_2 = (0.01, 0.8, 0.02, 0.01, 0.1, 1.9, 0.5)$ . . . . .	133
9.6	MCMC STABLE-GARCH com $\varphi_3 = (0.05, 0.5, 0.1, 0.05, 0.2, 1.8, 0.5)$ . . . . .	134

# Lista de Tabelas

4.1	Conjunto de parâmetros escolhido para simulação e estimação do modelo GEV-GARCH. . . . .	55
4.2	Comparação da estimação Bayesiana GEV-GARCH deste trabalho com os resultados de Zhao et al. (2011) . . . . .	62
5.1	Conjunto de parâmetros escolhido para simulação e estimação do modelo STABLE-GARCH. . . . .	66
7.1	Comparação na estimação dos parâmetros para os dados dem2gbp no modelo normal-GARCH(1,1). . . . .	87
7.2	Comparação na estimação dos parâmetros para os dados dem2gbp no modelo t-Student-GARCH(1,1). . . . .	88
7.3	Comparação na estimação dos parâmetros para os dados dem2gbp nos modelo GARCH(1,1) com distribuição condicional GEV e estável. . .	89
7.4	Comparação na estimação dos parâmetros para os dados dem2gbp no modelo normal-ARMA(1,1)-APARCH(1,1). . . . .	90
7.5	Comparação na estimação dos parâmetros para os dados dem2gbp no modelo t-Student-ARMA(1,1)-APARCH(1,1). . . . .	91
7.6	Comparação na estimação dos parâmetros para os dados dem2gbp no modelo ARMA(1,1)-APARCH(1,1) com distribuição GEV ou estável. . . . .	92
7.7	Comparação do tempo computacional em segundos entre as rotinas <code>garchFit</code> e <code>garchFitExtremes</code> com distribuição condicional normal. . . . .	93
8.1	Séries de dados reais utilizadas neste trabalho. . . . .	94
8.2	Sumário das estatísticas das séries temporais de retornos $r_t$ . . . . .	95
8.3	Teste de Ljung-Box para retornos ao quadrado . . . . .	98

8.4	Testes de estacionariedade em séries temporais . . . . .	101
8.5	Ajuste do modelo ARMA-GARCH com distribuição condicional estável para os dados da Microsoft. . . . .	102
8.6	Ajuste do modelo ARMA-GARCH com distribuição condicional GEV para os dados da Microsoft. . . . .	103
8.7	Ajuste do modelo ARMA-GARCH com distribuição condicional estável.	105
8.8	Ajuste do modelo ARMA-GARCH com distribuição condicional GEV.	106
8.9	Previsão de parâmetros futuros para o modelo GARCH(1,1) com inovações normal. . . . .	111
8.10	Previsão de parâmetros futuros para o modelo ARMA(1,1)-GARCH(1,1) com inovações normal. . . . .	112
8.11	Previsão de parâmetros futuros para o modelo ARMA(1,1)-GARCH(1,1) com inovações GEV e estável utilizando a rotina <code>garchFitExtremes</code> . . . . .	114
9.1	Resultados da estimação dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações GEV utilizando MLE. . . . .	121
9.2	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo AR(1)-GARCH(1,1) com inovações GEV utilizando MLE. . . . .	127
9.3	Resultados da estimação dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações GEV utilizando a metodologia Bayesiana. . . . .	128
9.4	Teste de estacionariedade para as cadeias simuladas da distribuição a posteriori dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações GEV. . . . .	129
9.5	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo AR(1)-GARCH(1,1) com inovações GEV utilizando a metodologia Bayesiana. . . . .	130
9.6	Resultados da estimação dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações estáveis utilizando MLE. . . . .	131
9.7	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo AR(1)-GARCH(1,1) com inovações estáveis utilizando MLE. . . . .	137
9.8	Resultados da estimação dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações estáveis utilizando a metodologia Bayesiana. . . . .	138

9.9	Teste de estacionariedade para as cadeias simuladas da distribuição a posteriori dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações estáveis. . . . .	139
9.10	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo AR(1)-GARCH(1,1) com inovações estáveis utilizando a metodologia Bayesiana. . . . .	140
9.11	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-APARCH(2,2) com inovações estáveis. . . . .	141
9.12	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-APARCH(2,2) com inovações estáveis. . . . .	142
9.13	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-GARCH(2,2) com inovações estáveis. . . . .	143
9.14	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(1,1)-GARCH(1,1) com inovações estáveis. . . . .	144
9.15	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-APARCH(2,2) com inovações GEV. . . . .	145
9.16	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-GARCH(2,2) com inovações GEV. . . . .	146
9.17	Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(1,1)-GARCH(1,1) com inovações GEV. . . . .	147
9.18	Comparação da volatilidade realizada com a volatilidade estimada com dados simulados do modelo ARMA(1,1)-GARCH(1,1) com inovações GEV e parâmetros (0.1, 0.2, 0.2, 0.05, 0.1, 0.1, 0.2). . . . .	148
9.19	Comparação da volatilidade realizada com a volatilidade estimada com dados simulados do modelo ARMA(2,2)-GARCH(2,2) com inovações GEV e parâmetros (0.1, 0.2, 0.3, 0.2, 0.5, 0.1, 0.1, 0.2, 0.1, 0.1, 0.1). . . . .	149
9.20	Comparação da volatilidade realizada com a volatilidade estimada com dados simulados do modelo ARMA(1,1)-GARCH(1,1) com inovações estáveis e parâmetros (0.1, 0.2, 0.2, 0.05, 0.1, 1.8, 0). . . . .	150
9.21	Comparação da volatilidade realizada com a volatilidade estimada com dados simulados do modelo ARMA(2,2)-GARCH(2,2) com inovações estáveis e parâmetros (0.1, 0.2, 0.3, 0.2, 0.5, 0.1, 0.1, 0.2, 0.1, 0.1, 1.9, 0.3). . . . .	151

# Resumo

Neste trabalho estendemos a aplicação do modelo combinado AR-GARCH governado por distribuições GEV e apresentado por Zhao et. al. (2011) para um modelo governado por distribuições estáveis, já que estas distribuições podem ser utilizadas para modelar dados de finanças, incluindo os eventos extremos. Além de estimação pelo método Bayesiano explorada por Zhao et. al. (2011), estimamos ambos os modelos também com o método clássico da máxima verossimilhança.

Posteriormente investigamos as condições de estacionariedade de um modelo ARMA-power-GARCH com inovações estáveis proposto por Rachev et. al. (2002) e estendemos este modelo derivando as condições de estacionariedade para um modelo assimétrico ARMA-APARCH com inovações estáveis. Esta última generalização nos permitiu implementar uma rotina numérica de estimação de modelos ARMA-APARCH que, ao contrário da conhecida rotina `fGARCH` apresentada por Wurtz et. al. (2006) estima modelos ARMA-APARCH com distribuição condicional GEV ou estável.

**Palavras Chave:** *ARMA, GARCH, Distribuições estáveis, Distribuição GEV, Máxima Verossimilhança, MCMC, Estacionariedade, Expoente de Lyapunov.*

# Abstract

We extend the application of the GEV-GARCH model given by Zhao et al. (2011) to a model driven by stable distributions as they share some similarities in modelling financial data, including extreme events. We perform both Maximum likelihood and Bayesian estimation of these models. Thereafter, we investigate the stationarity conditions of the ARMA-power-GARCH model with stable innovations proposed by Rachev et. al. (2002) and prove the stationarity conditions for the asymmetric model ARMA-APARCH with stable innovations. The last result allowed us to construct a numerical routine to estimate the parameters of an ARMA-APARCH model following stable and GEV distributions.

**key words:** *ARMA, GARCH, stable Distribution, GEV Distribution, Maximum Likelihood, MCMC, Stationarity, Lyapunov exponent.*



# Introdução

A maioria dos estudos empíricos envolvendo o mercado financeiro não focam diretamente nos preços, mas sim em seus retornos. A razão é que, em geral, os preços se comportam como passeios aleatórios ao contrário do que acontece com os retornos que são estacionários.

A princípio, não existe razão alguma para explicar por que ativos financeiros se comportam dessa maneira. Segundo Jondeau et al. (2007), a modelagem de dados financeiros é uma tentativa de capturar e explorar o fato de que a distribuição empírica dos retornos financeiros tem caudas mais pesadas do que as de uma distribuição normal e parece ter assimetria negativa.

Vários modelos já foram propostos para a modelagem da dependência temporal existente em dados financeiros, dos quais podemos citar os modelos ARMA e ARIMA presentes na maioria dos livros clássicos sobre o tema, por exemplo em Brockwell e Davis (1992, 1996).

Apesar dos retornos serem aparentemente estacionários e fracamente auto-correlacionados, o quadrado dos retornos podem ser serialmente correlacionados. Essa correlação serial nos quadrados dos retornos foi inicialmente modelada por Engle (1982) com os modelos ARCH (AutoRegressive Conditional Heteroskedasticity), e posteriormente por Bollerslev (1986) com o modelo ARCH generalizado (GARCH).

Uma abordagem bastante utilizada consiste em combinar um modelo linear ARMA com um modelo GARCH para ajustar dados estacionários, veja, por exemplo, Nakatsuma T. (2000), Brummelhuis R. (2007) e Wurtz et al. (2006). Isso porquê mesmo após o ajuste de um modelo linear ARMA, os resíduos apresentam volatilidade significativa.

Os modelos ARMA-GARCH são objeto de estudo de vários trabalhos envolvendo

a modelagem de dados financeiros (vide Wurtz et al., 2006). Na prática, esses modelos são utilizados em conjunto com distribuições de probabilidade conhecidas, tais como a normal e t-Student. Outro fato importante é que a aplicação dessas distribuições de probabilidade se limita ao espaço paramétrico em que possuem variância finita.

A ampla utilização dos modelos ARMA-GARCH para modelagem de dados que exibem volatilidade culminou no desenvolvimento de várias aplicações computacionais para estimação dos parâmetros desses modelos. O trabalho de Wurtz et al. (2006) introduz a rotina `garchFit` e a compara com outras rotinas computacionais de estimação de modelos ARMA-GARCH. Tanto a rotina de Wurtz et al. quanto as rotinas citadas neste artigo permitem a especificação de várias distribuições condicionais, incluindo a distribuição Normal e t-Student. Porém, nenhuma delas é capaz de estimar modelos com distribuição condicional estável ou GEV.

Zhao et al. (2011) utilizou a distribuição GEV combinada com o modelo ARMA-GARCH para modelar dados de finanças que exibem comportamentos extremos. Zhao et al. comparou seu modelo com o modelo clássico com inovações normal e concluiu que o modelo GEV é mais poderoso para explicar o comportamento de natureza extremal. Para realizar a estimação deste modelo Zhao et al. trabalhou especificamente com o modelo AR(1)-GARCH(1,1) e utilizou o método de estimação Bayesiana.

Os modelos ARMA-GARCH com distribuição estável foram abordados pela primeira vez em Rachev et al. (2002). A equação desse processo

$$\sigma_t = \omega + \sum_{i=1}^p \alpha_i |\varepsilon_{t-i}| + \sum_{j=1}^q \beta_j \sigma_{t-j}$$

se diferencia de maneira substancial da definição de um processo GARCH padrão

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2,$$

em que a potência do processo  $\sigma_t$  é quadrática. Isso acontece porque as distribuições estáveis têm variância infinita quando  $0 < \alpha < 2$ . Rachev et al. ainda impuseram  $\alpha > 1$  para assegurarem que as inovações possuam média finita.

Embora o modelo de Rachev et al. seja aplicado na prática para modelagem de dados financeiros (veja Omelchenko, 2009 e Lindsey, 2004), não existem rotinas computacionais disponíveis para estimação destes modelos.

Nosso principal interesse neste trabalho é de estudar o modelo ARMA-GARCH combinado com distribuições estáveis, pois esta distribuição tem sido amplamente utilizada para modelagem de dados financeiros, incluindo os eventos extremos (Nolan, 1999).

Neste trabalho iremos focar nos aspectos teóricos e computacionais envolvidos durante a estimação dos modelos ARMA-GARCH com distribuição condicional estável e GEV. Para isto, organizamos esta dissertação em 9 capítulos.

No Capítulo 1, definimos conceitos básicos da teoria de séries temporais e algumas distribuições de probabilidade de interesse. Também apresentamos alguns conceitos e teoremas que serão utilizados no Capítulo 6.

O capítulo 2 define os modelos estacionários ARMA, GARCH e APARCH (asymmetric power ARCH). Este último modelo será abordado com mais detalhes nos capítulos 6 e 7. No capítulo 3, definimos os modelos combinados ARMA-GARCH, e tratamos dos aspectos teóricos e práticos envolvidos durante a estimação de parâmetros utilizando as metodologias clássica e bayesiana.

Os capítulos 4 e 5 apresentam os resultados da estimação de parâmetros dos modelos com distribuição condicional GEV de Zhao et al. (2011) e estável de Rachev et al. (2002). O capítulo 4 apresenta também uma comparação dos nossos resultados da estimação do modelo GEV com os obtidos no trabalho de Zhao et al. (2011).

No capítulo 6, analisamos em detalhes as condições de estacionariedade do modelo introduzido por Rachev et al. (2002). Este modelo é, na verdade, uma generalização do processo GARCH, denotada por power-GARCH. A equação de volatilidade possui a flexibilidade de ter o expoente não fixo e assume a forma:

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i |\varepsilon_{t-i}|^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta.$$

Como a distribuição das inovações é estável, a estacionariedade do modelo irá depender do expoente  $\alpha$  (não confundir com os coeficientes  $\alpha_i$  da equação anterior) da cauda. Além da restrição  $\alpha > 1$ , Rachev et al. (2002) impôs ainda que  $0 < \delta < \alpha$ , pois neste caso o momento de ordem  $\delta$  da distribuição das inovações existe.

O processo power-GARCH é uma generalização intermediária entre o processo GARCH e APARCH. Este último possui a flexibilidade do parâmetro  $\delta$  e ainda per-

mite a inclusão de parâmetro(s) de assimetria  $\gamma_i$  na equação das inovações

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta.$$

De acordo com Wurtz et al. (2006) a rotina `garchFit` é também capaz de estimar os parâmetros dos modelos ARMA-APARCH que, em alguns casos, são capazes de modelar melhor a assimetria existente em dados reais (vide Zhou, 2009).

Com base no trabalho de Rachev et al. (2002), ficamos limitados à estimação de modelos power-GARCH. Devido à alta expansão na utilização dos modelos APARCH, seria de grande valor uma rotina capaz de estimar modelos ARMA-GARCH/APARCH com inovações estáveis.

Com essa motivação, investigamos os detalhes envolvidos na demonstração da estacionariedade do modelo ARMA-power-GARCH no trabalho de Rachev et al. (2002). A demonstração se baseia na aplicação do Teorema 3.2 de Bungeoun e Piccard (1992) que estabelece a estacionariedade de modelos autorregressivos com representação matricial na forma

$$X_{n+1} = A_{n+1}X_n + B_{n+1}, \quad n \in \mathbb{Z}.$$

Rachev et al. (2002) apresenta uma representação matricial para o processo ARMA-power-GARCH e, posteriormente verifica as hipóteses do Teorema de Bungeoun e Piccard (1992) para concluir a estacionariedade estrita deste processo.

De forma análoga, utilizamos as idéias contidas no trabalho de Rachev et al. (2002) para reescrever o processo ARMA-APARCH com inovações estáveis em forma matricial, e demonstramos alguns lemas para concluir na Proposição 6.1 que este modelo também satisfaz as condições do Teorema 3.2 de Bungeoun e Piccard (1992). De posse desse resultado temos condições de realizar a estimação ARMA-APARCH com distribuições estáveis que possuem variância infinita.

No capítulo 7, desenvolvemos uma rotina computacional de estimação dos modelos ARMA-GARCH/APARCH com distribuição condicional estável, GEV, normal, t-Student e skew t-Student. Juntamente com essa rotina incluímos outras rotinas de estimação em um pacote do software R denominado `garchFitExtremes`. Comparamos nossos resultados da estimação com a conhecida rotina `garchFit` para os casos de ajuste com distribuições normal, t-Student e skew t-Student. Os casos com distribuição estável e GEV não puderam ser comparados já que esta última rotina não

permite a estimação destes modelos. Além da estimação desses modelos de séries temporais que citamos, o pacote `garchFitExtremes` possui ainda funções de simulação e previsão de parâmetros futuros para o modelo ARMA-GARCH.

O capítulo 8 utiliza os resultados e conceitos apresentados até aqui para análise de três conjuntos de dados reais. Realizamos também um pequeno estudo de simulação para avaliar a previsão da volatilidade com os modelos com distribuição condicional estável e GEV.

Por último, o capítulo 9 conclui e apresenta algumas propostas de trabalhos futuros.

# Capítulo 1

## Preliminares

Nesta seção definimos alguns conceitos básicos sobre processos estocásticos e séries temporais. Apresentamos posteriormente algumas das estatísticas mais comuns para estimação dos parâmetros e funções de uma série temporal estacionária. Na seção seguinte apresentamos alguns conceitos e Teoremas que serão utilizados no capítulo 6. Na última seção definimos as distribuições estáveis e GEV, além de apresentar algumas de suas principais propriedades. Essas distribuições são objeto de estudo deste trabalho.

### 1.1 Conceitos Básicos de Séries Temporais

Uma série temporal é um conjunto de observações  $x_t$ , cada uma registrada em um instante específico  $t$ . Uma série temporal discreta é tal que o conjunto dos tempos  $T$  que a série é observada é um conjunto enumerável. Podemos ter também séries temporais contínuas que podem ser obtidas quando registramos os valores de  $x_t$  em um determinado intervalo de tempo. Um exemplo clássico de série temporal apresentado em Brockwell e Davis (1992) mostra a evolução da população dos Estados Unidos entre os anos 1790 e 1980. Outros exemplos de séries temporais incluem o preço de ações negociadas na bolsa de valores ou as cotações de taxas de câmbio entre duas moedas estrangeiras.

Segundo Brockwell e Davis (1992) o primeiro passo na análise de uma série temporal é a seleção de um modelo matemático adequado para as observações.

Suponha que cada observação  $x_t$  é a realização de uma determinada variável aleatória  $X_t$ . A série temporal  $\{x_t, t \in T_0\}$  é então a realização de uma família de variáveis aleatórias  $\{X_t, t \in T_0\}$ . Isso sugere que podemos modelar dados com um processo estocástico  $\{X_t, t \in T\}$  em que  $T \supseteq T_0$ . A definição formal de um processo estocástico é como se segue:

**Definição 1.1.** Um processo estocástico é uma família de variáveis aleatórias  $\{X_t, t \in T\}$  definidas em um espaço de probabilidade  $(\Omega, \Sigma, P)$

O espaço de probabilidade mencionado acima consiste em três elementos. O espaço amostral  $\Omega$  que é o conjunto (não vazio) de todas as realizações possíveis de um determinado experimento. A  $\sigma$ -álgebra  $\Sigma$  de subconjuntos de  $\Omega$  e a função  $P : \Sigma \rightarrow [0, 1]$  que associa probabilidades aos eventos ou elementos de  $\Sigma$ . O elemento  $\Sigma$  precisa satisfazer as seguintes propriedades:

- $\Omega \in \Sigma$ .
- Se  $A \in \Sigma$ , então  $A^c \in \Sigma$ .
- Se  $A_1, A_2, \dots, \in \Sigma$ , então  $A_1 \cup A_2 \cup \dots \in \Sigma$ .
- Além disso, para cada  $t$ , a função  $X_t : \Omega \rightarrow I$  é tal que para cada  $B$  subconjunto de  $I$ , tem-se que  $X_t^{-1}(B) = \{w \in \Omega : X_t(w) \in B\}$  é um subconjunto de  $\Sigma$ .

Quando trabalhamos com séries temporais o conjunto de índices  $T$  normalmente é igual a  $\{0, \pm 1, \pm 2, \dots\}$ ,  $[0, \infty)$  ou  $(-\infty, \infty)$ . Os processos estocásticos cujo conjunto  $T$  não é um subconjunto de  $\mathbb{R}$  também são importantes, mas não serão tratados aqui. Vale ressaltar também que para cada  $t \in T$ ,  $X_t = X(t, \cdot)$  é de fato uma função com domínio no conjunto  $\Omega$ . Por outro lado, para cada  $\omega \in \Omega$ ,  $X(\cdot, \omega)$  é uma função com domínio em  $T$  chamada de trajetória amostral ou realização do processo.

Um exemplo bastante utilizado na literatura é o de um passeio aleatório. Se definirmos o passeio como o processo  $\{S_t, t = 0, 1, 2, \dots\}$  é definido de tal forma que  $S_0 = 0$  e :

$$S_t = \sum_{i=1}^t X_i \text{ para } t \geq 1,$$

em que  $X_t$  é uma sequência i.i.d. (independente e identicamente distribuídos) de variáveis aleatórias satisfazendo  $P(X_t = 1) = P(X_t = -1) = 1/2$ .

**Definição 1.2.** (Função de distribuição de um processo estocástico). Seja  $\mathfrak{S}$  o conjunto de vetores  $\{\mathbf{t} = (t_1, \dots, t_n)' \in T^n : t_1 < t_2 < \dots < t_n, n = 1, 2, \dots\}$ . Então as funções de distribuição do processo  $\{X_t, t \in T\}$  são  $\{F_{\mathbf{t}}(\cdot), \mathbf{t} \in \mathfrak{S}\}$  definidas da seguinte forma:

$$F_{\mathbf{t}}(\mathbf{x}) = P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n), \quad \mathbf{x} = (x_1, \dots, x_n)' \in \mathbf{R}^n. \quad (1.1)$$

**Teorema 1.1.** (Teorema de Kolmogorov). As distribuições de probabilidade  $\{F_{\mathbf{t}}(\cdot), \mathbf{t} \in \mathfrak{S}\}$  são funções de distribuição de um processo estocástico se, e somente se, para cada  $n \in \{1, 2, \dots\}$ ,  $\mathbf{t} = (t_1, \dots, t_n)' \in \mathfrak{S}$  e  $1 \leq i \leq n$ ,

$$\lim_{x_i \rightarrow \infty} F_{\mathbf{t}}(\mathbf{x}) = F_{\mathbf{t}(\mathbf{i})}(\mathbf{x}(\mathbf{i})) \quad (1.2)$$

onde  $\mathbf{t}(\mathbf{i})$  e  $\mathbf{x}(\mathbf{i})$  são os vetores obtidos quando retiramos a  $i$ -ésima componente dos vetores  $\mathbf{t}$  e  $\mathbf{x}$ , respectivamente.

De acordo com Brockwell e Davis (1996) a condição anterior é simplesmente uma condição de consistência sobre a função  $F_{\mathbf{t}}(\cdot)$  que deve ter função de distribuição marginal coincidindo com a distribuição marginal de dimensão menor.

**Definição 1.3.** Se o processo  $\{X_t, t \in T\}$  é tal que  $Var(X_t) < \infty$  para cada  $t \in T$ , então a função de autocovariância  $\gamma_{\mathbf{x}}(\cdot, \cdot)$  do processo  $\{X_t\}$  é definida como:

$$\gamma_{\mathbf{x}}(r, s) = Cov(X_r, X_s) = E[(X_r - EX_r)(X_s - EX_s)] \quad r, s \in T. \quad (1.3)$$

**Definição 1.4.** A série temporal  $\{X_t, t \in \mathbb{Z}\}$  é dita como *fracamente* estacionária ou simplesmente *estacionária* se:

- $E|X_t|^2 < \infty$  para todo  $t \in \mathbb{Z}$
- $EX_t = c$  para todo  $t \in \mathbb{Z}$
- $\gamma_X(r, s) = \gamma_X(r + t, s + t)$  para todo  $r, s, t \in \mathbb{Z}$

Esse tipo de estacionariedade é geralmente chamado de *estacionariedade de segunda ordem* ou simplesmente *estacionário*. A partir deste momento quando mencionarmos o termo estacionariedade estaremos nos referindo a estacionariedade fraca.



Se o processo é estacionário, então a função de autocorrelação  $\gamma_X(r, s) = \gamma_X(r - s, 0)$  e portanto podemos redefinir essa função da seguinte forma:

$$\gamma_X(h) = \gamma_X(h, 0) = \text{Cov}(X_{t+h}, X_t) \quad \text{para todo } t, h \in \mathbb{Z}.$$

O valor  $h$  é referido na literatura como o *lag* em que estamos avaliando o processo.

**Definição 1.5.** A função de autocorrelação (ACF) do processo  $\{X_t, t \in T\}$  é definida em termos da função de autocovariância como:

$$\rho_X(h) = \gamma_X(h)/\gamma_X(0) = \text{Corr}(X_{t+h}, X_t) \quad \text{para todo } t, h \in \mathbb{Z}. \quad (1.4)$$

**Definição 1.6.** (Estacionariedade forte). A série temporal  $\{X_t, t \in \mathbb{Z}\}$  é dita ser fortemente estacionária se a distribuição conjunta dos vetores  $(X_{t_1}, \dots, X_{t_k})'$  e  $(X_{t_1+h}, \dots, X_{t_k+h})'$  é a mesma para qualquer inteiro positivo  $k$  e para todo  $t_1, \dots, t_k, h \in \mathbb{Z}$ . Esse tipo de estacionariedade é também referida na literatura como estacionariedade estrita. Segundo Brockwell e Davis (1996) esse conceito pode ser entendido intuitivamente como a propriedade de que os gráficos da realização de uma série temporal em dois intervalos diferentes de mesmo tamanho tenham o mesmo comportamento estatístico.

Se o processo  $\{X_t\}$  é fortemente estacionário podemos aplicar a definição 1.6 com  $k = 1$  e notar que a distribuição de  $X_t$  é a mesma para cada  $t \in \mathbb{Z}$ . Se  $E|X_t|^2 < \infty$  isso implica que  $EX_t$  e  $\text{Var}(X_t)$  são constantes. Por outro lado, para  $k = 2$ , podemos concluir que  $X_{t+h}$  e  $X_t$  tem a mesma distribuição e como consequência tem a mesma função de autocovariância para todo  $h \in \mathbb{Z}$ . Portanto, se um processo é estritamente estacionário com segundo momento finito, então ele será fracamente estacionário.

**Definição 1.7.** O processo  $\{X_t\}$  é uma série temporal gaussiana se, e somente se, a função de distribuição do processo é normal multivariada.

Vale ressaltar que se o processo  $\{X_t\}$  é estacionário e gaussiano, então ele também será estritamente estacionário. O seguinte exemplo pode ser encontrado em Brockwell e Davis (1992) e ilustra um processo estacionário que será depois abordado de uma forma mais geral.

Suponha, por exemplo, que temos uma sequência de variáveis aleatórias i.i.d.  $\varepsilon_t$  com variância  $\sigma^2$  e defina o processo  $X_t = \varepsilon_t + b\varepsilon_{t-1}$ . Então a função de autocovariância de  $X_t$  é dada por:

$$\begin{aligned} \text{Cov}(X_{t+h}, X_t) &= \text{Cov}(\varepsilon_{t+h} + b\varepsilon_{t+h-1}, \varepsilon_t + b\varepsilon_{t-1}) \\ &= \begin{cases} (1+b^2)\sigma^2 & \text{se } h = 0, \\ b\sigma^2 & \text{se } h = \pm 1, \\ 0, & \text{se } |h| > 1, \end{cases} \end{aligned}$$

e portanto o processo é estacionário.

**Definição 1.8.** Seja  $\{X_t\}$  um processo estacionário com média zero e função de autocovariância  $\gamma_X(\cdot)$  tal que  $\gamma_X(h) \rightarrow 0$  quando  $h \rightarrow \infty$ . A função de autocorrelação parcial (PACF) de uma série temporal estacionária é definida por  $\Lambda(k) = \phi_{kk}$ ,  $k \geq 1$ , com:

$$P_{\bar{s}p\{X_1, X_2, \dots, X_k\}} X_{k+1} = \sum_{j=1}^k \phi_{kj} X_{k+1-j},$$

em que os coeficientes  $\phi_{kk}$  se referem a projeção ortogonal de  $X_t$  no subespaço  $\bar{s}p\{X_1, X_2, \dots, X_k\}$  gerado pelas observações anteriores ao tempo  $t$ . As projeções  $P_{\bar{s}p\{1, X_2, \dots, X_k\}} X_{k+1}$  e  $P_{\bar{s}p\{1, X_2, \dots, X_k\}} X_1$  são encontradas através das seguintes equações válidas para  $k \geq 1$ :

$$\left\langle X_{k+1} - P_{\bar{s}p\{X_1, X_2, \dots, X_k\}} X_{k+1}, X_j \right\rangle = 0, \quad j = k, \dots, 1,$$

A aplicação do operador esperança nas últimas equações nos permite reescrever essas equações através do seguinte sistema:

$$\begin{bmatrix} \rho_X(0) & \rho_X(1) & \rho_X(2) & \cdots & \rho_X(k-1) \\ \rho_X(1) & \rho_X(0) & \rho_X(1) & \cdots & \rho_X(k-2) \\ \vdots & & & \ddots & \vdots \\ \rho_X(k-1) & \rho_X(k-2) & \rho_X(k-3) & & \rho_X(0) \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_X(1) \\ \rho_X(2) \\ \vdots \\ \rho_X(k) \end{bmatrix}. \quad (1.5)$$

**Propriedade 1.1.** A função de autocorrelação possui as seguintes propriedades:

- $\gamma_X(0) \geq 0$ ,

- $\gamma_X(h) = \gamma_X(-h)$  para todo  $h$ ,
- $|\gamma_X(h)| \leq \gamma_X(0)$  para todo  $h$ ,
- $\gamma_X(h)$  é uma função não negativa definida, ou seja,

$$\sum_{i,j}^n v_i \gamma_X(i-j) v_j \geq 0,$$

para todo inteiro positivo  $n$  e todo vetor real  $\mathbf{v} = (v_1, \dots, v_n)'$ .

De fato, a função de autocovariância é caracterizada por ser uma função real par nos inteiros e não-negativa definida. A prova dessa afirmação pode ser encontrada em Brockwell e Davis (1992).

A seguir apresentamos os estimadores que são normalmente utilizadas para estimação de parâmetros de um processo estacionário.

**Teorema 1.2.** *O estimador da média  $\mu$  de um processo estacionário é a média amostral dada por:*

$$\bar{X}_n = n^{-1}(X_1 + X_2 + \dots + X_n). \quad (1.6)$$

Este estimador é não viesado e seu erro médio quadrático pode ser expressado em termos da função de autocovariância:

$$E(\bar{X}_n - \mu)^2 = n^{-1} \sum_{h=-n}^n \left(1 - \frac{|h|}{n}\right) \gamma_X(h),$$

e portanto, se  $\gamma_X(h) \rightarrow 0$  quando  $h \rightarrow 0$ , o erro quadrático médio tende para zero e assim concluímos que esse estimador converge em média para  $\mu$ .

As funções de autocovariância e autocorrelação amostral são definidas como, respectivamente:

$$\hat{\gamma}_X(h) = n^{-1} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X}_n)(X_t - \bar{X}_n) \quad (1.7)$$

e,

$$\hat{\rho}_X(h) = \frac{\hat{\gamma}_X(h)}{\hat{\gamma}_X(0)}. \quad (1.8)$$

Segundo Brockwell e Davis (1992) os estimadores  $\hat{\gamma}_X(h)$  e  $\hat{\rho}_X(h)$  são viesados, até mesmo quando trocamos o fator  $n^{-1}$  pelo fator  $(n-h)^{-1}$  que é normalmente feito

quando ajustamos o estimador em seus graus de liberdade. Porém, para amostras grandes eles são praticamente não viesados e, portanto são largamente utilizados para estimação dessas funções.

De forma Similar a estimativa da função parcial de autocorrelação pode ser encontrada utilizando o estimador da função de autocorrelação.

**Definição 1.9.** A função amostral de autocorrelação parcial  $\hat{\Lambda}(k)$  no lag  $k$  do processo  $\{X_1, X_2, \dots, X_n\}$  é definida (desde que  $X_i \neq X_j$  para algum  $i$  e  $j$  da seguinte forma:

$$\hat{\Lambda}(k) = \hat{\phi}_{kk}, \quad 1 \leq k < n, \quad (1.9)$$

em que os coeficientes  $\hat{\phi}_{kk}$  são determinados através da equação (1.5) com  $\rho_X(j)$  substituído pelo seu respectivo estimador  $\hat{\rho}_X(j)$ .

Para mais detalhes sobre os conceitos e teoremas apresentados neste capítulo veja Brockwell e Davis (1992) ou Brockwell e Davis (1996). As próximas Definições serão utilizadas para demonstração de algumas proposições no capítulo 7.

## 1.2 Outros Conceitos Importantes

**Definição 1.10.** (Desigualdade de Jensen) Seja  $f$  uma função côncava definida na reta. Se a variável aleatória  $X$  é integrável, então:

$$f(E(X)) \geq E(f(X)). \quad (1.10)$$

Nas definições seguintes iremos denotar por  $M(n)$  o conjunto de todas as matrizes quadradas com entradas em  $\mathbb{R}$ .

**Definição 1.11.** Um vetor  $v \neq 0$  pertencente a  $\mathbb{R}^n$  é chamado de autovetor de  $A \in M(n)$  se existe  $\lambda \in \mathbb{R}$  satisfazendo:

$$Av = \lambda v, \quad (1.11)$$

Aqui  $\lambda$  é chamado de autovalor de  $A$  correspondente a  $v$ .

Vale observar também que se  $\lambda$  é um autovalor da matriz  $A$  então a seguinte equação deverá ser satisfeita:

$$\det(A - \lambda I) = 0. \quad (1.12)$$

**Definição 1.12.** Sejam  $\lambda_1, \lambda_2, \dots, \lambda_n$  os autovalores de uma matriz  $A \in M(n)$ . Então o raio espectral de  $A$  é definido como sendo:

$$R(A) = \max_i (|\lambda_i|) \quad (1.13)$$

**Definição 1.13.** Uma norma em  $\mathbb{R}^n$  é uma função  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  definida de tal forma que para todo  $c \in \mathbb{R}$ ,  $u$  e  $v \in \mathbb{R}^n$  valem:

- $\|cv\| = |c|\|v\|$ ,
- $\|u + v\| \leq \|u\| + \|v\|$ ,
- Se  $\|v\| = 0$  então  $v$  é um vetor nulo.

Além da conhecida norma euclidiana, temos ainda varias outras funções que definem normas no espaço  $\mathbb{R}^n$  dentre as quais destacamos a norma do máximo que é dada pela seguinte equação:

$$\|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}. \quad (1.14)$$

**Definição 1.14.** A norma de uma matriz  $A \in M(n)$  é definida através da seguinte equação:

$$\|M\| = \sup\{\|Mx\|/\|x\|, x \in \mathbb{R}^n, x \neq 0\} \quad (1.15)$$

.

**Definição 1.15.** O maior expoente de Lyapunov associado a uma sequência i.i.d. de matrizes aleatórias  $\{A_n, n \in \mathbb{Z}\}$  é definido da seguinte forma:

$$\zeta = \inf_{n \rightarrow \infty} E\left(\frac{1}{n+1} \log \|A_t A_{t-1} \dots A_{t-n}\|\right), \quad (1.16)$$

quando  $E(\log^+ \|A_t\|) < \infty$  em que  $\log^+(x) = \max(\log(x), 0)$ .

### 1.3 Distribuições de Probabilidade de Interesse

Nesta seção, definimos brevemente as conhecidas distribuições Normal e t-Student, pois serão utilizadas nas análises que realizamos nos Capítulos 8 e 9. Posteriormente iremos descrever as principais de definições e propriedades sobre as distribuições GEV e estável utilizadas nos modelos de séries temporais que iremos abordar neste trabalho.

### 1.3.1 Distribuições Normal e t-Student

**Definição 1.16.** Uma variável aleatória  $Z$  possui distribuição normal com parâmetro de locação  $\mu \in \mathbb{R}$  e parâmetro de escala  $\sigma > 0$  se possui função densidade de probabilidade (f.d.p.):

$$f_Z(z|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}}.$$

**Definição 1.17.** Uma variável aleatória  $Z$  possui distribuição t-Student padrão se possui f.d.p. dada por:

$$f_Z(z|\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi(\nu-2)}\Gamma(\frac{\nu}{2})} \frac{1}{\left(1 + \frac{z^2}{\nu-2}\right)^{\frac{\nu+1}{2}}},$$

em que  $\nu > 2$  é o parâmetro de forma e  $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$  é a função Beta.

**Definição 1.18.** O grau de assimetria  $\text{skew}(Z)$  e a curtose  $\text{curt}(Z)$  de uma variável aleatória  $Z$  com  $E(Z^4) < \infty$  podem ser calculados pelas seguintes expressões:

$$\text{skew}(Z) = \frac{E(Z^3)}{E(Z^2)^{3/2}}, \quad \text{curt}(Z) = \frac{E(Z^4)}{E(Z^2)^2} - 3.$$

Tanto o coeficiente de assimetria, quanto o de curtose são iguais a zero em uma distribuição normal. De acordo com Wurtz (2006) a assimetria e a kurtose da distribuição t-Student são dadas por:

$$\text{skew}(Z) = \frac{E(Z^3)}{E(Z^2)^{3/2}} = 0, \quad \text{curt}(Z) = \frac{E(Z^4)}{E(Z^2)^2} - 3 = \frac{6}{\nu - 4}$$

Das expressões acima podemos notar que a distribuição t-Student tem caudas mais pesadas que a da distribuição Normal, pois seu coeficiente de Kurtose foi sempre maior que zero.

### 1.3.2 Distribuições estáveis

Uma das primeiras justificativas para utilização das distribuições estáveis foi trazida por Mandelbrot (1963), que defendeu a hipótese de que os retornos dos preços de ativos financeiros seguem uma distribuição estável, ou seja, devemos modelar retornos utilizando distribuições cuja variância não existe.

**Definição 1.19.** Uma variável aleatória  $Z$  é dita ter distribuição estável se ela tiver um domínio de atração, isto é, se existe uma sequência de variáveis aleatórias i.i.d.  $Y_1, Y_2, \dots$ , uma sequência de números positivos  $c_n$  e uma sequência de números reais  $d_n$ , tal que:

$$c_n(Y_1 + Y_2 + \dots + Y_n) - d_n \xrightarrow{d} Z \quad (1.17)$$

A notação  $\xrightarrow{d}$  denota convergência em distribuição

Embora bastante intuitiva, esta primeira definição não é tão utilizada na prática. A Definição mais conhecida e extremamente útil para derivação de propriedades desta distribuição é dada em termos de sua função característica.

Por não apresentarem uma forma fechada simples, as distribuições estáveis são geralmente definidas em termos de sua função característica.

**Definição 1.20.** Dizemos que  $Z$  tem distribuição estável, denotada por  $Z \sim S(\alpha, \beta, \sigma, \mu)$ , se sua função característica  $\phi(Z)$  for dada por:

$$\phi_Z(u) = Ee^{iuZ} = \exp\{-|\sigma u|^\alpha [1 - i\beta \operatorname{sgn}(u)w(\alpha, t) + i\mu u]\}, \quad (1.18)$$

em que  $w(\alpha, u) = \tan(\frac{\pi\alpha}{2})$  se  $\alpha \neq 1$  e  $-(2/\pi) \ln |u|$  se  $\alpha = 1$ .

Uma distribuição estável univariada é caracterizada por quatro parâmetros. Estes parâmetros são o índice de estabilidade  $\alpha$ , o parâmetro de escala  $\sigma$ , o parâmetro de assimetria  $\beta$  e o parâmetro de posição  $\mu$ . A distribuição estável é gaussiana (normal) quando  $\alpha = 2$ , e neste caso,  $\sigma$  é proporcional ao desvio padrão,  $\beta$  pode ser tomado como zero e  $\mu$  é a média. Essa distribuição é simétrica quando  $\beta = 0$  e assimétrica à esquerda (direita) quando  $\beta > 0$  ( $\beta < 0$ ).

Existem também outras parametrizações dos parâmetros desta distribuição que nos levam a definições ligeiramente diferentes da função característica de uma variável aleatória estável. Algumas destas representações apresentam propriedades importantes de continuidade da função característica em determinados parâmetros (veja Samorodnitsky e Taqqu, 2000). Uma outra reparametrização das distribuições estáveis conhecida com parametrização  $S_0$  é de importância crucial para o desenvolvimento de rotinas numéricas para o cálculo da função de densidade. Seguindo a parametrização

$S_0$  (veja Nolan 1997):

$$\phi_Z^*(u) = \begin{cases} \exp\{-|\sigma u|^\alpha \{1 + i\beta \operatorname{sgn}(u) \tan \frac{\pi\alpha}{2} [\sigma|u|^{1-\alpha} - 1] + i\mu_0 u\}\} & \alpha \neq 1 \\ \exp\{-\sigma|u| [1 + i\beta \operatorname{sgn}(u) \frac{2}{\pi} \ln(\sigma|u|) + i\mu_0 u]\} & \alpha = 1. \end{cases} \quad (1.19)$$

Quando  $\alpha = 2$  temos a distribuição Normal com média 0 e variância 2 e quando  $\alpha = 1$  temos a distribuição de Cauchy. Outro fato importante é que para  $\delta > 0$ ,  $E(|Z|^\delta) < \infty$  se, e somente se  $\delta < \alpha$ . Portanto, somente para o caso  $\alpha > 1$ , a média de  $Z$  existe e em todos os casos em que  $\alpha < 2$  a variância não existe.

Distribuições estáveis com  $\alpha < 2$  dividem várias propriedades com a distribuição gaussiana, mas também diferem da distribuição gaussiana de várias formas significantes. Quando  $\alpha < 2$  por exemplo, as caudas da distribuição decaem como uma função potência. Isto significa que a variável aleatória estável exibe muito mais variabilidade do que uma Gaussiana, ou seja, é muito mais provável de assumir valores distantes da média.

A alta variabilidade das distribuições estáveis é uma das múltiplas razões que as tornam importantes no campo da modelagem. Distribuições estáveis têm sido usadas para modelar diversos fenômenos tais como campos gravitacionais de estrelas, distribuição da temperatura em reatores nucleares, taxas de câmbio, preços no mercado de estoque e precipitação anual de chuvas.

A teoria das distribuições estáveis univariadas foi essencialmente desenvolvida nas décadas de 1920 e 1930 por Paul Lévy e Aleksander Yakovlevic Khinchine, o que pode ser visto em Zolotarev (1986). Ela é abordada em detalhes em livros clássicos de probabilidade como Gnedenko and Kolmogorov (1954) e Feller (1971). Denotamos uma distribuição estável por  $S(\alpha, \sigma, \beta, \mu)$  e escrevemos:

$$X \sim S(\alpha, \sigma, \beta, \mu).$$

Também denotaremos por:

$$X \sim S_\alpha S$$

quando  $X$  for *simétrica  $\alpha$ -estável*.

**Exemplos:**



1. A distribuição Gaussiana  $S(2, \sigma, 0, \mu) = N(\mu, 2\sigma^2)$  tem função densidade conhecida e dada por:

$$f(x|\mu, \sigma) = \frac{1}{2\sqrt{\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{4\sigma^2}}, \quad -\infty \leq \mu \leq \infty, \quad \sigma \geq 0.$$

2. A distribuição de Cauchy  $S(1, \sigma, 0, \mu)$  também tem função densidade de probabilidade conhecida e dada por:

$$f(x|\mu, \sigma) = \frac{\sigma}{\pi((x-\mu)^2 + \sigma^2)}, \quad -\infty \leq \mu \leq \infty, \quad \sigma \geq 0.$$

3. A distribuição de Levy  $S(1/2, \sigma, 1, \mu)$  tem função densidade dada por:

$$f(x|\mu, \sigma) = \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}} \frac{1}{(x-\mu)^{\frac{3}{2}}} \exp\left\{-\frac{\sigma}{2(x-\mu)}\right\}, \quad x > \mu.$$

Esta última distribuição é concentrada no intervalo  $(\mu, \infty)$ .

A Figura 1.1 mostra a forma dessas distribuições para mesmos valores dos parâmetros  $\mu$  e  $\sigma$ .

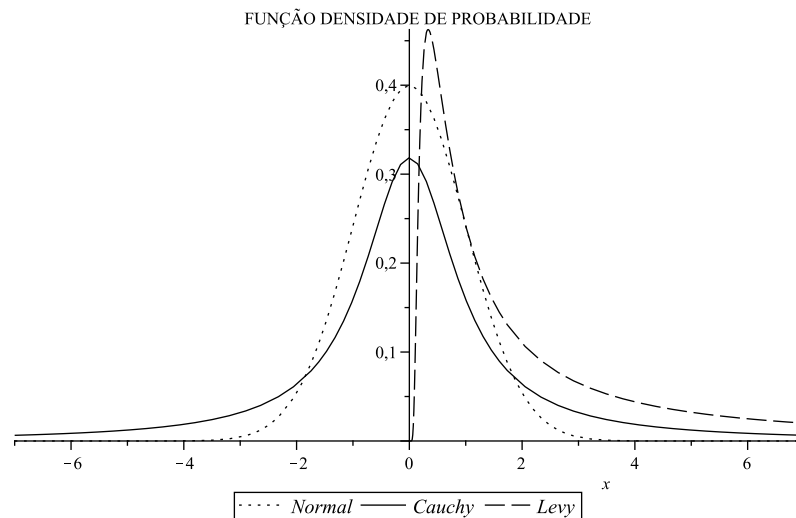


Figura 1.1: Gráfico distribuições Gaussiana, de Cauchy e de Lévy com  $\mu = 0$  e  $\sigma = 1$ .

Note que as caudas da distribuição de Cauchy são bem mais “pesadas” que as caudas da distribuição normal. As propriedades de funções características serão usadas para derivação desses resultados importantes envolvendo distribuições estáveis.

**Propriedade 1.2.** *Sejam  $Z_1$  e  $Z_2$  variáveis aleatórias independentes tais que  $Z_i \sim S(\alpha, \sigma_i, \beta_i, \mu_i), i = 1, 2$ . Então:*

$$Z_1 + Z_2 \sim S(\alpha, \sigma, \beta, \mu)$$

em que:

$$\sigma = (\sigma_1^\alpha + \sigma_2^\alpha)^{\frac{1}{\alpha}}, \quad \beta = \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}, \quad \mu = \mu_1 + \mu_2.$$

**Propriedade 1.3.** *Seja  $Z \sim S(\alpha, \sigma, \beta, \mu)$  e seja  $a$  uma constante real diferente de zero e  $b \in \mathbb{R}$ . Então:*

- $aZ \sim S(\alpha, |a|\sigma, \text{sign}(a)\beta, a\mu)$ , se  $\alpha \neq 1$ .
- $aZ \sim S(1, |a|\sigma, \text{sign}(a)\beta, a\mu - \frac{2}{\pi}a(\ln|a|)\sigma\beta)$ , se  $\alpha = 1$ .
- $Z + b \sim S(\alpha, \sigma, \beta, \mu + b)$ .

**Propriedade 1.4.** *Para qualquer  $0 < a < 2$ ,*

$$Z \sim S(\alpha, \sigma, \beta, 0) \Leftrightarrow -Z \sim S(\alpha, \sigma, -\beta, 0) \quad (1.20)$$

**Propriedade 1.5.** *Seja  $Z \sim S(\alpha, \sigma, \beta, \mu)$  com  $0 < \alpha < 2$ . Então:*

$$\begin{cases} \lim_{\lambda \rightarrow \infty} \lambda^\alpha P\{Z > \lambda\} &= C_\alpha \frac{1+\beta}{2} \sigma^\alpha \\ \lim_{\lambda \rightarrow \infty} \lambda^\alpha P\{Z < -\lambda\} &= C_\alpha \frac{1-\beta}{2} \sigma^\alpha. \end{cases}$$

em que:

$$C_\alpha = \left( \int_0^\infty x^{-\alpha} \sin x dx \right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} &, \text{ se } \alpha \neq 1 \\ 2/\pi &, \text{ se } \alpha = 1. \end{cases}$$

Para a demonstração desses dois últimos resultados veja Samorodnitsky (1994, p. 16). A seguinte propriedade nos diz que a convergência da função geradora de momentos de uma variável aleatória estável depende somente do parâmetro de estabilidade  $\alpha$ .

**Propriedade 1.6.** *Seja  $Z \sim S(\alpha, \sigma, \beta, \mu)$  com  $0 < \alpha < 2$ . Então:*

$$\begin{cases} E|Z|^\delta < \infty & \text{para qualquer } 0 < \delta < \alpha, \\ E|Z|^\delta = \infty & \text{para qualquer } \delta \geq \alpha. \end{cases}$$

### 1.3.3 Distribuição GEV

A teoria de valores extremos lida com a modelagem de eventos cuja ocorrência é gerada por uma sequência de máximos ou de mínimos. O foco principal está no índice  $\xi$  que é a característica das caudas das distribuições que apresentaremos a seguir. Segundo Jondeau et al. (2007) assumir que retornos financeiros seguem uma distribuição de valor extremo é muitas vezes uma suposição forte.

Seja  $Z_1, \dots, Z_n$  uma sequência de variáveis aleatórias a partir da qual definimos:

$$M_n = -\min(-Z_1, \dots, -Z_n) = \max(Z_1, \dots, Z_n),$$

em que podemos notar que precisamos olhar somente para a cauda superior da distribuição de  $Z_t$ . Observe também que se  $Z_t$  é uma amostra i.i.d., então se  $F_Z(\cdot)$  é a função de distribuição acumulada de  $Z_t$ , segue que:

$$\begin{aligned} P(M_n \leq z) &= P[\max(Z_1, \dots, Z_n) \leq z] = P(Z_1 \leq z, \dots, Z_n \leq z) \\ &= \prod_{t=1}^n P(Z_t \leq z) = [F_Z(z)]^n. \end{aligned}$$

Em suma, esta última equação mostra que se assumirmos que as observações são i.i.d., então a distribuição dos máximos de uma amostra finita pode ser obtida facilmente da distribuição de  $F_X(\cdot)$ . Se denotarmos  $z_F = \sup\{z \in \mathbb{R}; F_Z(z) < 1\} < \infty$  como sendo o último ponto da amostra ordenada, podemos notar que:

$$P(M_n \leq z) = \begin{cases} 0, & \text{para todo } z < z_F, \\ 1, & \text{para todo } z > z_F, \end{cases}$$

e, portanto, obtemos a distribuição degenerada de  $M_n$ . Porém, esta função não é de utilidade prática e este fato deu origem ao trabalho que culminou no Teorema de Fisher e Tippett de 1928 (descrito em Jondeau et al., 2007) que caracteriza o limite da distribuição do máximo.

**Teorema 1.3.** *Seja  $Z_t$  uma sequência de variáveis aleatórias i.i.d. Se existe um parâmetro de locação  $\mu_n \in \mathbf{R}$ , e um parâmetro de escala  $\sigma_n > 0$ , e alguma função de distribuição acumulada  $H$  tal que o limite dos extremos normalizados  $Y_n = (M_n - \mu_n)/\sigma_n$  converge para  $H$ , então:*

$$\lim_{n \rightarrow \infty} Pr\left(\frac{M_n - \mu_n}{\sigma_n} \leq y\right) = H(y), \quad \forall y \in \mathbb{R}, \quad (1.21)$$

então  $H$  pertença a uma das seguintes distribuições:

$$\begin{aligned} \text{Weibull: } \Psi_\alpha(y) &= \begin{cases} \exp(-(-y)^\alpha), & \text{se } y \leq 0 \text{ e } \alpha > 0 \\ 1, & \text{se } y > 0. \end{cases} \\ \text{Gumbel: } \Lambda_\alpha(y) &= \exp(-y^{-\alpha}), \text{ para } y \in \mathbb{R}, \\ \text{Fréchet: } \Sigma_\alpha(y) &= \begin{cases} 0, & \text{se } y \leq 0 \text{ e } \alpha > 0 \\ \exp(-y^{-\alpha}), & \text{se } y > 0. \end{cases} \end{aligned} \quad (1.22)$$

Como observado em Jondeau et al. (2007) a distribuição de Gumbel apresenta caudas mais leves que as caudas da distribuição de Fréchet. Essas três distribuições podem ser combinadas em uma única distribuição conhecida como distribuição GEV.

**Definição 1.21.** Uma variável aleatória  $Z$  é dita ter distribuição GEV se possui função de distribuição com parâmetros  $\xi \in \mathbb{R}$ ,  $\sigma > 0$  e  $\mu \in \mathbb{R}$  dada por:

$$F_Z(z|\xi, \sigma, \mu) = \begin{cases} \exp\{-[1 + \xi(z - \mu)/\sigma]^{-1/\xi}\}, & \text{se } \xi \neq 0, 1 + \xi(z - \mu)/\sigma > 0 \\ \exp\{-\exp[-(z - \mu)/\sigma]\}, & \text{se } \xi = 0. \end{cases} \quad (1.23)$$

A distribuição GEV é na verdade a generalização de três importantes distribuições da teoria de valores extremos: distribuição de Fréchet ( $\xi > 0$ ), de Gumbel ( $\xi = 0$ ) e de Weibull ( $\xi < 0$ ). O parâmetro  $\xi$  é chamado de *índice da cauda da distribuição* e  $1/\xi$  é chamado de *índice de forma da distribuição*. A função densidade de probabilidade (f.d.p.) da distribuição GEV pode ser facilmente derivada das equações anteriores.

**Definição 1.22.** Uma variável aleatória  $Z$  com distribuição GEV definida anteriormente tem f.d.p. dada por:

$$f_Z(z|\xi, \sigma, \mu) = \begin{cases} \frac{1}{\sigma} \left(1 + \xi \frac{z_i - \mu}{\sigma}\right)^{-\frac{1}{\xi} - 1} \exp\left(-\left(1 + \xi \left(\frac{z_i - \mu}{\sigma}\right)^{-\frac{1}{\xi}}\right)\right), & \xi \neq 0 \\ \frac{1}{\sigma} \exp\left(-\frac{z_i - \mu}{\sigma} - e^{-\frac{z_i - \mu}{\sigma}}\right), & \xi = 0. \end{cases} \quad (1.24)$$

Uma descrição mais completa sobre distribuições de valores extremos pode ser encontrada no capítulo 7 de Jondeau et al. (2007).

No próximo capítulo iremos definir os modelos de séries temporais ARMA e ARCH, juntamente com suas extensões.

# Capítulo 2

## Modelos de Séries Temporais

Nesta seção definiremos modelos básicos de séries temporais que utilizaremos durante o desenvolvimento deste trabalho. Explicitamos também alguns teoremas importantes para utilização destes modelos e definiremos algumas distribuições de probabilidade largamente utilizadas em conjunto com esses modelos.

### 2.1 Modelos lineares ARMA e ARIMA

**Definição 2.1.** O processo  $\{\varepsilon_t\}$  é dito ser um ruído branco com média 0 e variância  $\sigma^2$ , denotado por

$$\{\varepsilon_t\} \sim \text{WN}(0, \sigma^2), \quad (2.1)$$

se, e somente se,  $\{\varepsilon_t\}$  tem média zero e função de autocovariância igual a:

$$\gamma_X(h) = \begin{cases} \sigma^2, & \text{se } h = 0 \\ 0, & \text{se } h \neq 0. \end{cases}$$

De acordo com Brockwell e Davis (1992) uma grande classe de modelos estacionários pode ser gerada utilizando uma sequência de ruídos brancos combinada com uma série de equações de diferenças lineares. O modelo ARMA definido a seguir é talvez o mais conhecido e estudado dentre eles.

**Definição 2.2.** (Processos ARMA). O processo  $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$  é dito ser um processo ARMA( $m, n$ ) se  $\{X_t\}$  é estacionário e para todo  $t$ , temos

$$X_t - a_1 X_{t-1} - \dots - a_m X_{t-m} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_n \varepsilon_{t-n}, \quad (2.2)$$

em que  $\{\varepsilon_t\} \stackrel{i.i.d}{\sim} \text{WN}(0, \sigma^2)$ . Além do mais, dizemos que  $\{X_t\}$  é um processo ARMA( $m, n$ ) com média  $\mu$  se  $\{X_t - \mu\}$  é um processo ARMA( $m, n$ ).

As equações descritas em (2.2) podem ser representadas simbolicamente de uma forma mais compacta pela seguinte notação:

$$a(\mathcal{B})X_t = b(\mathcal{B})\varepsilon_t, \quad t = 0, \pm 1, \pm 2, \dots, \quad (2.3)$$

em que  $a(v)$  e  $b(v)$  são polinômios de grau  $m$  e  $n$  definidos da seguinte forma:

$$a(v) = 1 - a_1v - \dots - a_mv^m \quad (2.4)$$

e,

$$b(v) = 1 - b_1v - \dots - b_nv^n \quad (2.5)$$

e  $\mathcal{B}$  é o operador de atraso definido por:  $\mathcal{B}^j X_t = X_{t-j}, j = 0, \pm 1, \pm 2, \dots$

Os polinômios  $a(v)$  e  $b(v)$  são geralmente conhecidos na literatura de séries temporais como polinômios autorregressivo e de média móvel, respectivamente.

**Teorema 2.1.** *O processo ARMA definido em (2.2) tem uma única solução estacionária se, e somente se:*

$$a(v) = 1 - a_1v - \dots - a_mv^m \neq 0 \quad \text{para todo } |v| = 1. \quad (2.6)$$

*Demonstração.* Observe primeiro que a região definida por todos os complexos  $v$  tais que  $|v| = 1$  é definida como sendo um círculo unitário no plano complexo. Se  $a(v) \neq 0$  para todo  $v$  no círculo unitário, então existe  $\epsilon > 0$  tal que:

$$\frac{1}{a(v)} = \sum_{j=-\infty}^{\infty} \chi_j v^j \quad \text{para } 1 - \epsilon < |v| < 1 + \epsilon,$$

e  $\sum_{j=-\infty}^{\infty} \chi_j < \infty$ . Então podemos definir o polinômio  $1/a(\mathcal{B})$  como sendo a seguinte série de operadores passados:

$$\frac{1}{a(\mathcal{B})} = \sum_{j=-\infty}^{\infty} \chi_j \mathcal{B}^j.$$

Aplicando o operador  $\chi(\mathcal{B}) = 1/a(\mathcal{B})$  nos dois lados da equação (2.3), obtemos

$$X_t = \chi(\mathcal{B})a(\mathcal{B})X_t = \chi(\mathcal{B})b(\mathcal{B})\varepsilon_t = \psi(\mathcal{B})\varepsilon_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j},$$

onde  $\psi(\mathcal{B}) = \chi(\mathcal{B})b(\mathcal{B}) = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon^j$ , e portanto,  $\psi(\mathcal{B})\varepsilon_t$  é a única solução estacionária das equações (2.2).  $\square$

**Definição 2.3.** (Causalidade). Um processo ARMA( $m, n$ ) é causal se existem constantes  $\{\psi_j\}$  tais que  $\sum_{j=0}^{\infty} |\psi_j| < \infty$  e

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \text{ para todo } t. \quad (2.7)$$

A causalidade dos processos ARMA pode ser entendida como a capacidade de expressar o processo  $\{X_t\}$  como uma função causal do processo  $\{\varepsilon_t\}$ , ou seja, utilizando somente valores passados deste último processo. Segundo Brockwell e Davis (1996), a causalidade do processo  $\{X_t\}$  é equivalente a seguinte condição:

$$a(v) = 1 - a_1v - \dots - a_mv^m \neq 0 \text{ para todo } |v| \leq 1. \quad (2.8)$$

Outra propriedade importante é a *invertibilidade* de um processo temporal. O processo  $\{X_t\}$  é dito ser invertível se podemos expressar  $\varepsilon_t$  em termos de  $X_s$ ,  $s \leq t$ .

**Teorema 2.2.** Um processo ARMA( $m, n$ ) é invertível se existem constantes  $\{\pi_j\}$  tais que  $\sum_{j=0}^{\infty} |\pi_j| < \infty$  e

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \text{ para todo } t. \quad (2.9)$$

A condição de invertibilidade é equivalente a seguinte equação:

$$b(\varepsilon) = 1 - b_1\varepsilon - \dots - b_q\varepsilon^q \neq 0 \text{ para todo } t. \quad (2.10)$$

Os modelos ARMA são muito importantes pois representam uma grande parte das séries estacionárias de segunda ordem existentes. Uma generalização dessa classe, que incorpora também séries não estacionárias, está contida nos modelos autorregressivos integrados de média movel (ARIMA). Esses processos, nada mais são do que processos que podem ser reduzidos a processos ARMA quando aplicamos equações de diferenças um número finito de vezes.

**Definição 2.4.** Seja  $d$  um inteiro não negativo, então o processo  $\{X_t\}$  é um processo ARIMA( $m, d, n$ ) se  $Y_t = (1 - B)^d X_t$  é um processo ARMA( $m, n$ ) causal.

A representação dos processos ARMA definida na equação (2.3) nos permite obter uma representação mais compacta para os processos ARIMA:

$$a^*(\mathcal{B})X_t = a(\mathcal{B})(1 - \mathcal{B})^d X_t = b(\mathcal{B})\varepsilon_t, \quad \{\varepsilon_t\} \stackrel{i.i.d}{\sim} \text{WN}(0, \sigma^2),$$

De acordo com Brockwell e Davis (2002) os modelos ARIMA são úteis também para representar dados com tendências.

## 2.2 Modelos não lineares ARCH e suas extensões

Embora bastante úteis, os modelos ARMA assumem variância condicional constante, portanto, sua utilização não é adequada para modelagem de dados que possuem heterocedasticidade condicional. Engle (1982) ao introduzir os modelos ARCH ressaltou essa propriedade com o seguinte exemplo. Considere inicialmente um modelo autoregressivo AR(1) descrito por:

$$X_t = a_1 X_{t-1} + \varepsilon_t \quad (2.11)$$

em que  $\varepsilon_t$  é um ruído branco com  $Var(\varepsilon_t) = \sigma^2$ . A média condicional de  $X_t$  é  $a_1 X_{t-1}$  enquanto a média não condicional  $E(X_t)$  é tal que  $(1 - a_1)E(X_t) = 0$  e portanto  $E(X_t) = 0$  assumindo  $a_1 \neq 1$ . Por outro lado, a variância condicional de  $X_t$  é  $\sigma^2$ , enquanto que a variância não condicional é  $\sigma^2/(1 - a_1^2)$ . Segundo Engle (1982), esses tipos de processos poderiam trazer melhores resultados para previsão se fossem capazes de utilizar informação do passado para interferir na variância condicional. Para superar essa dificuldade e permitir que a informação passada influencie na variância condicional, Engle (1982) introduziu os modelos ARCH (Autoregressive Conditional Heteroskedastic), que foram posteriormente generalizados por Bollerslev (1986) nos modelos GARCH (Generalized Autoregressive Conditional Heteroskedastic). Engle (1982) definiu as inovações  $\varepsilon_t$  da equação (2.11) como sendo um processo autoregressivo com heterocedasticidade condicional da seguinte forma:

$$\varepsilon_t = z_t \sigma_t, \quad (2.12)$$

em que  $z_t$  é um processo i.i.d. com média zero e variância unitária. Embora esse processo seja não correlacionado, sua variância condicional agora pode variar com o tempo e é dada por  $Var(\varepsilon_t | \Omega_{t-1}) = \sigma_t$ . O modelo autorregressivo com heterocedasticidade condicional ARCH( $p$ ) de Engle (1982) para o processo  $\{X_t\}$  pode ser definido como a solução das equações

$$\varepsilon_t = z_t \sigma_t, \quad \{z_t\} \stackrel{i.i.d.}{\sim} \text{WN}(0, 1), \quad (2.13)$$

em que  $\sigma_t$  é uma função positiva de  $\{\varepsilon_s, s < t\}$ , definida por:

$$\varepsilon_t = \omega + \sum_{i=0}^p \alpha_i \varepsilon_{t-i}^2, \quad (2.14)$$



com  $\omega > 0$  e  $\alpha_j \geq 0, j = 1, 2, \dots, p$ . O processo  $\sigma_t$  pode ser entendido como a variância condicional de  $\varepsilon_t$  dado  $\{\varepsilon_s, s < t\}$ .

Essa abordagem pode ser também generalizada para as conhecidas extensões dos modelos ARCH, incluindo os modelos GARCH introduzidos por Bollerslev (1986).

**Definição 2.5.** (Bollerslev, 1986) O processo  $\{\varepsilon_t\}$  segue um modelo GARCH( $p, q$ ) se satisfaz as seguintes equações:

$$\begin{aligned} \varepsilon_t &= z_t \sigma_t, \text{ onde } z_t \stackrel{i.i.d}{\sim} \mathcal{D}_\vartheta(0, 1), \\ \sigma_t^2 &= \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \\ &= \omega + \alpha(\mathcal{B})\varepsilon_t^2 + \beta(\mathcal{B})\sigma_t^2, \end{aligned} \quad (2.15)$$

em que  $\mathcal{D}_\vartheta(0, 1)$  é a distribuição de probabilidade das inovações com média zero e variância igual a 1. Aqui,  $\vartheta$  se refere aos parâmetros adicionais que a distribuição de probabilidade pode ter para ajuste da forma e assimetria. Observe que o modelo GARCH se reduz ao modelo ARCH quando todos os coeficientes  $\beta_j$  são iguais a zero. De acordo com Wurtz et al. (2006), de forma semelhante ao que acontece com os modelos ARMA, o modelo GARCH geralmente nos leva a uma representação mais parcimoniosa capturando a dependência da variância condicional nos coeficientes  $\beta_j$  ao invés de utilizarmos mais coeficientes no modelo original ARCH para ajuste de dados reais. Bollerslev (1986) estabeleceu as condições para que os modelos GARCH sejam estacionários, ou seja,  $E(\varepsilon_t)$  e  $Var(\varepsilon_t)$  devem ser constantes e portanto não dependentes do tempo. De fato, Bollerslev mostrou que se  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$  teremos  $E(\varepsilon_t) = 0$ ,  $Var(\varepsilon_t) = \omega / (1 - \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j)$  e  $Cov(\varepsilon_t, \varepsilon_s) = 0$ ,  $t \neq s$ , ou seja, o processo é não correlacionado e estritamente estacionário. Esses fatos são resumidos no próximo teorema:

**Teorema 2.3.** (Bollerslev, 1986) O processo GARCH( $p, q$ ) da definição 2.5 é fracamente estacionário com:

$$E(\varepsilon_t) = 0, \quad var(\varepsilon_t) = \omega \left( 1 - \sum_{i=1}^p \alpha_i \sum_{j=1}^q \beta_j \right)^{-1} \quad e \quad cov(\varepsilon_t, \varepsilon_s) = 0 \quad \text{para } t \neq s,$$

se, e somente se,

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1.$$

*Demonstração.* A prova desse teorema pode ser encontrada em Bollerslev (1986). Conforme a definição dada por Engle (1982) temos:

$$\varepsilon_t = z_t \sigma_t, \quad n_t \stackrel{i.i.d}{\sim} N(0, 1). \quad (2.16)$$

Como veremos durante a prova desse teorema, só é requerido que a distribuição das inovações tenha segundo momento finito, e portanto, podemos utilizar outras distribuições que satisfazem essa propriedade. Fazendo substituições sucessivas na equação da volatilidade temos

$$\begin{aligned} \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i z_{t-i}^2 \sigma_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 \\ &= \omega + \sum_{i=1}^q \alpha_i z_{t-i}^2 \left( \omega + \sum_{i=1}^q \alpha_i z_{t-i-j}^2 \sigma_{t-i-j}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i-j}^2 \right) + \\ &\quad + \sum_{i=1}^p \beta_i \left( \omega + \sum_{i=1}^q \alpha_i z_{t-i-j}^2 \sigma_{t-i-j}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i-j}^2 \right) \\ &\quad \vdots \\ &= \omega \sum_{k=0}^{\infty} M(t, k), \end{aligned}$$

em que o termo  $M(t, k)$  envolve todos os termos da forma:

$$\prod_{i=1}^q \alpha_i^{u_i} \prod_{j=1}^p \beta_j^{v_j} \prod_{l=1}^n \eta_{t-l}^2,$$

para

$$\sum_{i=1}^q u_i + \sum_{j=1}^p v_j = k, \quad \sum_{i=1}^q a_i = n,$$

e

$$1 \leq S_1 < S_2 < \dots < S_n \leq \max\{kq, (k-1)q + p\}.$$

Isso pode ser feito pois, cada soma envolvendo o termo  $\sigma_t^2$  pode sempre ser separada em  $\omega$  mais uma outra parte envolvendo  $\sigma_{t-j}^2$ . Portanto,

$$\begin{aligned} M(t, 0) &= 1, \\ M(t, 1) &= \sum_{i=1}^q \alpha_i z_{t-i}^2 + \sum_{i=1}^p \beta_i \\ M(t, 2) &= \sum_{j=1}^q \alpha_j z_{t-j}^2 \left( \sum_{i=1}^q \alpha_i z_{t-i-j}^2 + \sum_{i=1}^p \beta_i \right) + \\ &\quad + \sum_{i=1}^p \beta_i \left( \sum_{i=1}^q \alpha_i z_{t-i-j}^2 + \sum_{i=1}^p \beta_i \right), \end{aligned}$$

e em geral temos:

$$M(t, k + 1) = \sum_{i=1}^q \alpha_i z_{t-i}^2 M(t - i, k) + \sum_{i=1}^p \beta_i M(t - i, k). \quad (2.17)$$

Como o processo  $z_t^2$  é um processo i.i.d., os momentos  $M(t, k)$  não dependem de  $t$ , e podemos afirmar que:

$$E(M(t, k)) = E(M(s, k)) \text{ para todo } k, t, s.$$

Aplicando o operador esperança nos momentos  $M(t, k + 1)$  obtemos:

$$\begin{aligned} E(M(t, k + 1)) &= \left( \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right) E(M(t, k)) \\ &\quad \vdots \\ &= \left( \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right)^{k+1} E(M(t, 0)) \\ &= \left( \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right)^{k+1}. \end{aligned}$$

Finalmente podemos calcular a esperança de  $\varepsilon_t^2$  como:

$$\begin{aligned} E(\varepsilon_t^2) &= \omega E\left( \sum_{k=0}^{\infty} M(t, k) \right) \\ &= \omega \sum_{k=0}^{\infty} E\left( M(t, k) \right) \\ &= \omega \sum_{k=0}^{\infty} \left( \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j \right)^k, \\ &= \omega \left( 1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j \right)^{-1}, \end{aligned}$$

pois esta última expressão é exatamente uma série geométrica e sua convergência está condicionada a:

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1,$$

e portanto,  $\varepsilon_t^2$  converge quase certamente. Como  $\varepsilon_t = z_t \sigma_t$ , temos que

$$E(\varepsilon_t) = E\{E(\varepsilon_t | \varepsilon_s, s < t)\} = 0$$

e para  $t \neq s$  temos:

$$Cov(\varepsilon_t, \varepsilon_s) = E\{\varepsilon_s E(\varepsilon_t | \varepsilon_s, s < t)\} = 0$$

□

O próximo modelo foi introduzido por Ding, Granger e Engle (1993) e generaliza o processo GARCH de Bollerslev.

**Definição 2.6.** (Ding, Granger e Engle ,1993) O processo  $\varepsilon_t$  segue um modelo APARCH( $p, q$ ) se é definido da seguinte forma:

$$\begin{aligned} \varepsilon_t &= z_t \sigma_t, \quad z_t \stackrel{i.i.d}{\sim} \mathcal{D}_\vartheta(0, 1), \\ \sigma_t^\delta &= \omega + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta, \\ \omega &> 0, \quad \delta \geq 0, \\ \alpha_i &\geq 0, \quad i = 1, \dots, p, \\ -1 &< \gamma_i < 1, \quad i = 1, \dots, p, \\ \beta_j &\geq 0, \quad j = 1, \dots, q. \end{aligned}$$

Este modelo foi originalmente introduzido por Ding, Granger, e Engle (1993) e é de extrema importância pois nos permite levar em consideração outros expoentes na equação de variância e também a assimetria existente no processo.

**Teorema 2.4.** *O modelo APARCH( $p, q$ ) da definição 2.6 é fracamente estacionário se:*

$$\sum_{i=1}^p \lambda_i \alpha_i + \sum_{j=1}^q \beta_j < 1, \quad (2.18)$$

onde  $\lambda_i = E(|z| + \gamma_i z)^\delta$  e  $z_t \stackrel{i.i.d}{\sim} \mathcal{D}_\vartheta(0, 1)$ .

A prova da estacionariedade deste modelo segue basicamente os mesmos passos da demonstração do Teorema (2.3) e será omitido aqui. É importante notar que para  $\gamma_i \neq 0$  e/ou  $\delta \neq 2$  a constante  $\lambda_i$  irá depender da distribuição de inovações do processo.

Como foi salientado anteriormente, cada modelo de série temporal tem inovações seguindo uma determinada distribuição de probabilidades, e portanto, é necessário

definirmos algumas dessas distribuições. Sem dúvida a distribuição Normal é a distribuição mais conhecida, e juntamente com a distribuição t-Student fazem parte de uma grande variedade de trabalhos acadêmicos publicados na área de séries temporais. A distribuição Normal, por exemplo, foi utilizada em conjunto com o modelo ARCH de Engle (1983) para explicar o comportamento da inflação da Inglaterra em seu artigo de 1983. Já a distribuição t-Student mostrou ter a capacidade de capturar melhor a kurtose em modelos quando é utilizada em conjunto com os modelos ARCH ou suas extensões.

No próximo capítulo apresentamos os modelos combinados ARMA-GARCH, bastante comuns em aplicações envolvendo a modelagem de dados financeiros. Posteriormente iremos abordar os modelos GEV-GARCH (introduzido por Zhao et al. ,2011) e o modelo STABLE-GARCH (Introduzido por Rachev et al. ,2002).

# Capítulo 3

## Modelos Combinados (ARMA-GARCH)

Segundo Morettim (2004), podemos considerar inicialmente o ajuste de um modelo ARMA a uma série de retornos para eliminar a correlação serial entre as observações e, em seguida, verificar se os resíduos do modelo apresentam heterocedasticidade condicional e ajustar um modelo ARCH (ou alguma de suas extensões) a estes resíduos. Esses modelos são conhecidos de uma forma geral como modelos ARMA-GARCH e serão definidos a seguir.

De acordo com Wurtz et al. (2006) um modelo ARMA-GARCH de uma série temporal univariada  $x_t$  pode ser representado na forma:

$$X_t = E(X_t|\Omega_{t-1}) + \varepsilon_t \quad (3.1)$$

em que  $E(\cdot|\cdot)$  é a esperança condicional,  $\Omega_{t-1}$  é o conjunto de todas as informações até o tempo  $t - 1$  e  $\varepsilon_t$  são os resíduos ou as inovações da série temporal. Em um modelo ARMA-GARCH a esperança condicional é modelada de acordo com um processo ARMA, e as inovações são modeladas de acordo com um processo GARCH.

Recentemente, Zhao et al. (2011) propôs um modelo para valores extremos assumindo que as observações seguem uma distribuição GEV segundo um processo AR(1)-GARCH(1,1). Como as distribuições estáveis podem ser utilizadas para modelagem de dados de finanças, incluindo eventos extremos, o objetivo deste trabalho é estudar o modelo proposto por Zhao et al. (2011) que possui inovações seguindo uma distribuição GEV e estender sua aplicação a modelos governados por distribuições

estáveis.

**Definição 3.1.** A série temporal  $X_t$  é definida como um processo ARMA( $m, n$ )-GARCH( $p, q$ ) se

$$\begin{aligned} X_t &= \mu + \sum_{i=1}^m a_i X_{t-i} + \sum_{j=1}^n b_j \varepsilon_{t-j} + \varepsilon_t \\ \varepsilon_t &= z_t \sigma_t, \text{ em que } z_t \stackrel{i.i.d.}{\sim} \mathcal{D}_\vartheta(0, 1), \\ \sigma_t^2 &= \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \end{aligned} \tag{3.2}$$

em que  $\mu$  é a média do processo,  $a_1, \dots, a_m$  são os coeficientes autoregressivos,  $b_1, \dots, b_n$  são os coeficientes de média móvel,  $\alpha_1, \dots, \alpha_p \geq 0$ ,  $\beta_1, \dots, \beta_q \geq 0$ ,  $\omega > 0$  e  $\mathcal{D}_\vartheta(0, 1)$  é a distribuição de probabilidade das inovações com média zero e variância igual a 1.

De acordo com Wurtz et al. (2006), de forma semelhante ao que acontece com os modelos ARMA, o modelo GARCH geralmente nos leva a uma representação mais parcimoniosa capturando a dependência da variância condicional nos coeficientes  $\beta_j$  ao invés de utilizarmos mais coeficientes no modelo original ARCH para ajuste de dados reais. Bollerslev (1986) estabeleceu as condições para que os modelos GARCH sejam estacionários, ou seja,  $E(\varepsilon_t)$  e  $Var(\varepsilon_t)$  devem ser constantes e portanto não dependentes do tempo. De acordo com Wurtz et al. (2006) o modelo combinado ARMA-GARCH é estacionário quando ambos os modelos ARMA e GARCH analisados individualmente são estacionários, ou seja, satisfazem os Teoremas 2.1 e 2.3.

De acordo em Wurtz et al. (2006) a especificação de modelagem da volatilidade pode ser substituída por um modelo mais abrangente, a saber o modelo APARCH da Definição 2.6.

Os APARCH são largamente aplicados em conjunto com distribuições condicionais com variância finita, tais como normal, t-Student, GED (Wurtz, 2006).

Neste trabalho iremos investigar o modelo ARMA-GARCH com inovações GEV proposto por Zhao et al. (2011) para modelagem de valores extremos.

**Definição 3.2.** A série temporal  $X_t$  segue um processo ARMA( $m, n$ )-APARCH( $p, q$ )

se  $X_t$  satisfaz as seguintes equações:

$$\begin{aligned}
X_t &= \mu_t + \varepsilon_t = \mu + \sum_{i=1}^m a_i x_{t-i} + \sum_{j=1}^n b_j \varepsilon_{t-j} + \varepsilon_t \\
\varepsilon_t &= z_t \sigma_t, \text{ em que } z_t \stackrel{i.i.d.}{\sim} \mathcal{D}_\vartheta(0, 1), \\
\sigma_t^\delta &= \omega + \sum_{i=1}^p \alpha_i (\varepsilon_{t-i} - \gamma_i |\varepsilon_{t-i}|)^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta,
\end{aligned} \tag{3.3}$$

em que  $\mu$  é a média do processo,  $a_1, \dots, a_m$  são os coeficientes autoregressivos,  $b_1, \dots, b_n$  são os coeficientes de média móvel,  $\alpha_1, \dots, \alpha_p \geq 0$ ,  $\beta_1, \dots, \beta_q \geq 0$ ,  $\omega > 0$ ,  $-1 < \gamma_1, \dots, \gamma_p \leq 1$ ,  $\delta > 0$  e  $\mathcal{D}_\vartheta(0, 1)$  é a f.d.p. da variável aleatória que representa as inovações do processo.

Na próxima seção definimos o modelo proposto por Zhao et al. (2011).

### 3.1 O modelo GEV-GARCH

O modelo GEV-GARCH proposto por Zhao et al. (2011) assume que as observações seguem uma distribuição GEV com locação e escala variando com o tempo, e portanto, de extrema importância para a modelagem de eventos extremos que apresentem dependência temporal.

**Definição 3.3.** (Zhao et al., 2011) O modelo GEV-GARCH assume que as observações  $X_t$  são advindas de uma distribuição GEV seguindo um processo AR(1)-GARCH(1,1) com vetor de parâmetros  $\theta = (\mu, a_1, \omega, \alpha_1, \beta_1, \xi)$ . Dessa forma, dada a informação passada, podemos escrever:

$$X_t \sim GEV(\xi, \sigma_t, \mu_t), \quad \sigma_t > 0, \quad 1 + \xi(x_t - \mu_t)/\sigma_t > 0, \tag{3.4}$$

com dependência condicional dada por:

$$\begin{aligned}
X_t &= \mu + a_1 X_{t-1} + \varepsilon_t, \\
\sigma_t^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \\
\varepsilon_t &= X_t - \mu_t,
\end{aligned} \tag{3.5}$$

em que  $|a_1| < 1, \omega > 0, \alpha_1 > 0, \beta_1 > 0$ .



**Teorema 3.1.** *O modelo AR(1)-GARCH(1,1) proposto por Zhao et al. (2011) é fracamente estacionário se, e somente se,  $\alpha_1 + \beta_1 < 1$ .*

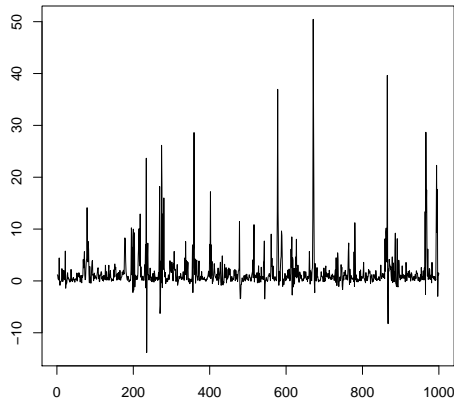
*Demonstração.* Embora escrito de maneira diferente, esse modelo é exatamente o mesmo modelo GARCH da Definição 2.5. Se assumimos que  $z_t$  é uma sequência de variáveis aleatórias i.i.d. com f.d.p.  $f_Z(z_t|\xi, 1, 0)$ , podemos definir o processo GARCH como sendo  $\varepsilon_t = \sigma_t z_t$  e finalmente escrevemos  $X_t$  como sendo um processo AR-GARCH da Definição 3.1, ou seja,  $X_t = \mu_t + \varepsilon_t$  onde  $\mu_t = \mu + a_1 X_{t-1}$ . Como este modelo só se diferencia pelo fato de utilizar outra distribuição com segundo momento finito para a equação das inovações, as condições de estacionariedade do modelo seguem diretamente do Teorema 2.3.  $\square$

Um fato que pode ser observado nesse tipo de modelo é que a forma de dependência temporal assume simetria em relação a variância condicional, ou seja, a escala da distribuição GEV responde da mesma forma não levando em consideração se os choques são positivos ou negativos. Uma conclusão natural é que poderíamos utilizar uma extensão do modelo ARCH conhecida como Asymmetric Power ARCH (APARCH) para superar essa suposição de simetria na dependência da variância condicional. Utilizando as propriedades da distribuição GEV podemos obter expressões para a esperança e a variância de  $X_t$ :

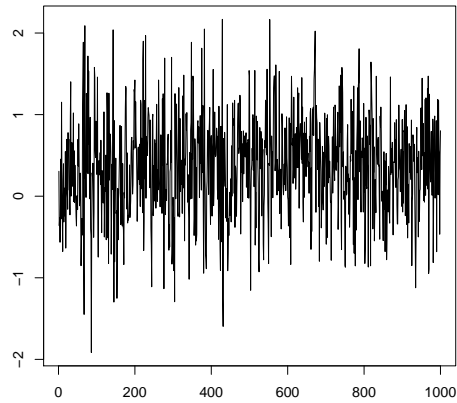
$$E(X_t) = \mu_t + \sigma_t \frac{(\Gamma(1 - \xi) - 1)}{\xi}, \quad \xi < 1 \quad (3.6)$$

$$Var(X_t) = \sigma_t^2 \frac{(\Gamma(1 - 2\xi) - \Gamma^2(1 - \xi))}{\xi^2}, \quad \xi < 0.5 \quad (3.7)$$

A Figura 3.1 mostra a simulação de uma série temporal seguindo um modelo AR(1)-GARCH(1,1) com inovações GEV e outro com inovações seguindo uma distribuição normal. Claramente notamos que as duas séries temporais exibem vários períodos onde se concentra em altas ou em baixas, ambas de grande magnitude, fenômeno este que é conhecido como *conglomerados de volatilidade*. Ao contrário do modelo GARCH padrão (com inovações seguindo uma distribuição normal), o modelo com inovações GEV apresenta choques de magnitude significativamente maiores que o modelo normal. Para melhor visualização deste modelo, na Figura 3.2 mostramos os gráficos de mais duas séries.

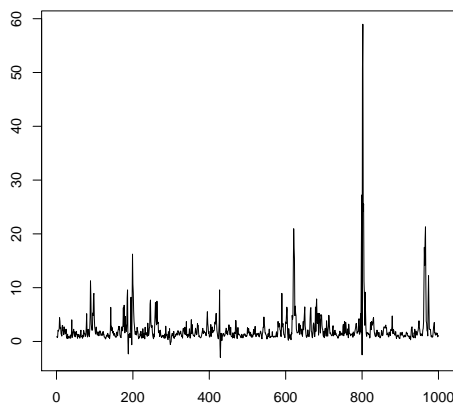


(a) Modelo GEV-GARCH

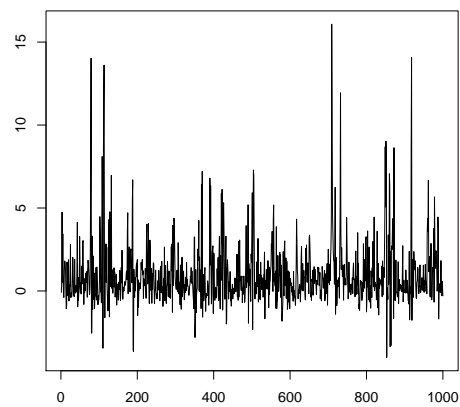


(b) Modelo Normal

Figura 3.1: Simulação de 1000 observações do modelo AR(1)-GARCH(1,1). Lado esquerdo: Modelo com inovações seguindo uma distribuição GEV com parâmetros  $\theta = (0.2, 0.3, 0.2, 0.3, 0.2, 0.2)$ . Lado direito: Modelo com os mesmos parâmetros exceto que a distribuição das inovações segue um modelo Normal.



(a)  $\theta = (0.1, 0.4, 0.1, 0.5, 0.5, 0.2)$



(b)  $\theta = (0.4, 0.4, 0.1, 0.1, 0.1, 0.1)$

Figura 3.2: Simulação de 1000 observações do modelo AR(1)-GARCH(1,1) com inovações GEV e parâmetro  $\theta$ .

## 3.2 O modelo STABLE-GARCH e algumas generalizações

O seguinte modelo foi introduzido por Rachev et al. (1995).

**Definição 3.4.** (Rachev et al., 1995) Considere que, dada a informação passada, a distribuição das observações  $X_t$  com parâmetros  $\varphi = (\mu, a_1, \omega, \alpha_1, \beta_1, \alpha, \beta)$  pode ser escrita como:

$$X_t \sim S(\alpha, \beta, \sigma_t, \mu_t), \quad (3.8)$$

em que  $0 < \alpha \leq 2$ ,  $-1 \leq \beta \leq 1$ ,  $-\infty < X_t < \infty$ . A dependência condicional do modelo é dada por:

$$\begin{aligned} X_t &= \mu + a_1 X_{t-1} + \varepsilon_t, \\ \sigma_t &= \omega + \alpha_1 \sigma_{t-1} + \beta_1 |\varepsilon_{t-1}|, \\ \varepsilon_t &= X_t - \mu_t, \end{aligned} \quad (3.9)$$

onde  $|a_1| < 1$ ,  $\omega_0 > 0$ ,  $\alpha_1 > 0$ ,  $\beta_1 > 0$ .

Nesse artigo Rachev et al. (1995) provou que a condição de estacionariedade forte deste modelo é

$$\lambda \alpha_1 + \beta_1 < 1,$$

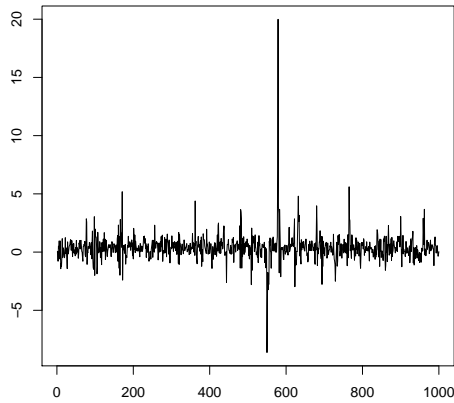
onde

$$\lambda = \Gamma(1 - 1/\alpha) \frac{\pi}{2}$$

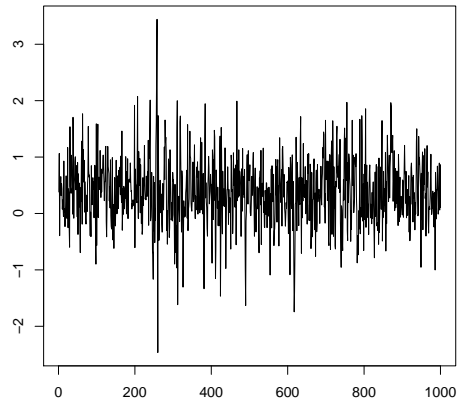
Existe uma diferença substancial entre o modelo da Definição 3.4 e o modelo GARCH de Bollerslev da Definição 2.5. A dependência temporal da variância condicional não é mais uma relação quadrática.

Na Figura 3.3 mostramos um gráfico da simulação do modelo AR(1)-GARCH(1,1) com inovações estáveis e outro modelo com inovações normal. Adicionalmente, na Figura 3.4 mostramos o gráfico de outras duas séries do modelo da Definição 3.4.

Modelos mais gerais que os AR-GARCH são os STABLE-power-GARCH cuja estacionariedade foi demonstrada em Rachev et al. (2002). Este modelo é denotado stable Paretian power-GARCH process no trabalho de Rachev.

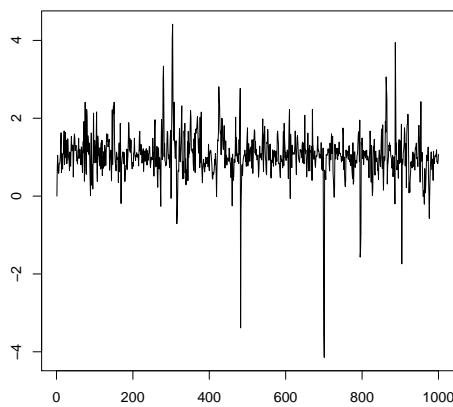


(a) Modelo STABLE-GARCH

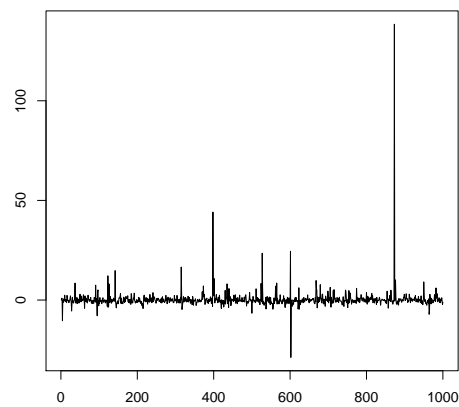


(b) Modelo Normal

Figura 3.3: Simulação de 1000 observações do modelo AR(1)-GARCH(1,1). Lado esquerdo: Inovações seguindo uma distribuição  $\alpha$ -estável com parâmetros  $\varphi = (0.2, 0.3, 0.2, 0.3, 0.2, 1.8, 0)$ . Lado direito: Modelo com os mesmos parâmetros porém com inovações seguindo uma distribuição Normal.



(a)  $\varphi = (0.1, 0.4, 0.1, 0.5, 0.5, 1.8, 0)$



(b)  $\varphi = (0.4, 0.4, 0.1, 0.1, 0.1, 1.8, 0.5)$

Figura 3.4: Simulação de 1000 observações do modelo AR(1)-GARCH(1,1) com inovações estáveis e parâmetro  $\varphi$ .

**Definição 3.5.** (Rachev et al., 2002) O processo  $X_t$  é chamado STABLE-power-GARCH se pode ser escrito como um processo ARMA-APARCH da Definição 3.1 com  $\gamma_i = 0, \forall i = 1, \dots, p$  e distribuição condicional estável.

Esse modelo será explorado no Capítulo 6 onde exploramos com mais detalhes a demonstração dada por Rachev et al. (2002) sobre a estacionariedade desses modelos.

### 3.3 Estimação de Modelos ARMA-GARCH

Após a etapa de especificação do modelo o nosso objetivo é a estimação dos parâmetros. Tanto a inferência clássica quanto a inferência bayesiana tem seus pilares baseados na função de verossimilhança dos dados.

A metodologia desse trabalho consiste em utilizar os métodos de máxima verossimilhança (MLE) e Markov Chain Monte Carlo (MCMC) para a estimação dos parâmetros do modelo. A metodologia Bayesiana, em contraste com a inferência clássica, assume a priori que podemos incluir no processo inferencial qualquer informação disponível advinda de algum especialista.

O método clássico de estimação, também conhecido como método da Máxima Verossimilhança (MLE) consiste em encontrar o conjunto de parâmetros  $\theta$  que melhor se ajustam aos dados, e são geralmente obtidos minimizando-se o negativo da função log-verossimilhança. Alguns dos parâmetros do modelo são definidos em um subconjunto limitado e por isso precisamos de alguma rotina numérica que resolva esse problema para otimização de funções cujo domínio tenha restrições. O software R oferece as rotinas `nlminb()` e `optim(method="L-BFGS-B")` que podem ser utilizados para a otimização da verossimilhança. Um dos principais problemas desse tipo de abordagem é que muitas vezes precisamos de uma estimativa inicial do valor dos parâmetros para evitar que o método convirja para uma solução que é localmente ótima, mas que não é a solução que estamos interessados. Em outras ocasiões o método de otimização pode também procurar por soluções fora do espaço paramétrico, e por isso devemos penalizar a avaliação da verossimilhança de forma adequada. Uma solução natural para esse problema é realizar a otimização iniciando o algoritmo em pontos diferentes do espaço paramétrico e escolher a solução ótima dentre as soluções individuais.

Podemos também utilizar outra metodologia de estimação dos parâmetros e utilizar esses valores retornados como valores iniciais do algoritmo de otimização da função de verossimilhança.

Segundo Wurtz (2006) uma outra solução válida é realizar a estimação com um método de otimização livre, e utilizar essa solução em um método de otimização limitada. Em alguns casos podemos também melhorar nossas estimativas se possuímos uma boa estimativa da solução do modelo é utilizar um algoritmo de otimização livre que irá ajustar melhor nossas estimativas.

Em contraste com a metodologia clássica o método Bayesiano depende somente da estimação da distribuição a posteriori dos parâmetros do modelo. Como as soluções analíticas só existem em alguns casos, o método de simulação Monte Carlo via cadeias de Markov, mais conhecido como *Markov chain Monte Carlo* (MCMC) é geralmente utilizado para simular valores da distribuição posteriori. O método MCMC para gerar valores da distribuição posteriori  $p(\theta|x)$  consiste em selecionar um ponto inicial  $\theta^{(0)}$  e então, para cada iteração  $i$  (não confundir com o indexador da série temporal), simular um valor  $\theta^{(i)}$  de uma outra distribuição de probabilidade conhecida  $J_t(\theta^{(i)}|\theta^{(i-1)})$  de tal forma que a distribuição da cadeia  $\theta^{(0)}, \theta^{(1)}, \dots, \theta^{(i)}, \dots$  convirja para a distribuição  $p(\theta|x)$ . Em especial, tratamos de dois importantes algoritmos de simulação estocástica, o primeiro deles é o *Metropolis-Hastings (M-H)* e o segundo é o método de amostragem de Gibbs, mais conhecido como amostrador de Gibbs ou *Gibbs sample*. O algoritmo Metropolis-Hastings pertence a família dos algoritmos de simulação de cadeias de Markov e é utilizado quando não conseguimos simular diretamente de uma determinada distribuição condicional completa. Esse algoritmo utiliza uma regra de aceitação/rejeição de novos valores da amostra para que a cadeia tenha distribuição convergindo para a distribuição de interesse. O algoritmo 3.1 descreve de maneira sucinta os passos do método Metropolis-Hastings para simulação de uma distribuição qualquer.

Em problemas onde o conjunto paramétrico é multidimensional o método de amostragem de Gibbs é largamente utilizado devido sua eficiência e simplicidade. Em cada passo  $i$  da cadeia, o algoritmo de Gibbs atualiza cada componente condicionalmente as componentes anteriores no passo  $t$  e condicionalmente as componentes posteriores na iteração  $i - 1$ , pois estas ainda não foram atualizadas. O algoritmo 3.2 exemplifica

---

**Algorithm 3.1** Metropolis-Hastings

---

**Input:**  $p(x|\theta)$ ,  $J_t(\cdot|\cdot)$  (distribuição proposta),  $\eta$  (tamanho da cadeia).

**Output:**  $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(\eta)}$

- 1: Simule ou defina um valor inicial  $\theta^{(0)}$ , com  $p(\theta^{(0)}|x) > 0$
- 2: **for**  $t = 1, 2, \dots, \eta$  **do**
- 3:     Simule  $\theta^*$  de uma distribuição  $J_t(\theta^*|\theta^{(i-1)})$  (distribuição proposta)
- 4:     Calcule a razão das densidades,

$$r = \frac{p(x|\theta^*)/J_t(\theta^*|\theta^{(i-1)})}{p(x|\theta^{(i-1)})/J_t(\theta^{(i-1)}|\theta^*)}$$

- 5:     Simule  $s$  de uma distribuição  $U(0, 1)$
  - 6:     **if**  $s < \min(1, r)$  **then**
  - 7:          $\theta^{(i)} = \theta^*$
  - 8:     **else**
  - 9:          $\theta^{(i)} = \theta^{(i-1)}$
  - 10:    **end if**
  - 11: **end for**
- 

o funcionamento do amostrador de Gibbs na simulação da distribuição a posteriori de  $\theta = (\alpha, \beta, \gamma)$ . Em aplicações de modelos multiparamétricos em que a distribuição a posteriori dos parâmetros não é uma distribuição simples é natural pensar em utilizar uma combinação do amostrador de Gibbs com o método Metropolis-Hastings para simular valores da distribuição a posteriori. De acordo com Gelman (2003), quando utilizamos o algoritmo Metropolis-Hastings, o parâmetro de dispersão da distribuição proposta  $J_i(\cdot|\cdot)$  precisa ser ajustado para que obtenhamos uma boa taxa de aceitação de novos valores da amostra.

Voltando ao modelo ARMA-GARCH, podemos definir a função de verossimilhança em termo das inovações  $z_t$  que são variáveis aleatórias i.i.d. com função densidade conhecida. Para uma amostra  $\{X_1, X_2, \dots, X_n\}$  de uma série temporal seguindo o modelo especificado, a função de log-verossimilhança é dada por:

$$\mathcal{L}_N(\theta) = \ln \prod_t \mathcal{D}_\theta(X_t, E(X_t|\Omega_{t-1}), \sigma_t) \quad (3.10)$$

Ainda com relação a estimação Bayesiana devemos ressaltar que, após a simulação

---

**Algorithm 3.2** Gibbs sampler

---

**Input:**  $p(\theta|x)$ ,  $\eta$ (Tamanho da cadeia),  $\theta = (\alpha, \beta, \gamma)$

**Output:**  $\theta^{(0)} = (\alpha^{(0)}, \beta^{(0)}, \gamma^{(0)}), \dots, \theta^{(\eta)} = (\alpha^{(\eta)}, \beta^{(\eta)}, \gamma^{(\eta)})$

- 1: Simule ou defina um valor inicial  $\theta^{(0)} = (\alpha^{(0)}, \beta^{(0)}, \gamma^{(0)})$ , com  $p(\theta^{(0)}|x) > 0$
  - 2: **for**  $i = 1, 2, \dots, \eta$  **do**
  - 3:   Simule  $\alpha^{(i)}$  da distribuição  $p(\alpha|\beta^{(i-1)}, \gamma^{(i-1)}, x)$
  - 4:   Simule  $\beta^{(i)}$  da distribuição  $p(\beta|\alpha^{(i)}, \gamma^{(i-1)}, x)$
  - 5:   Simule  $\gamma^{(i)}$  da distribuição  $p(\gamma|\alpha^{(i)}, \beta^{(i)}, x)$
  - 6: **end for**
- 

da distribuição dos parâmetros a posteriori, o processo inferencial para realizar estimativas dos parâmetros, intervalos de confiança e testes de hipóteses é realizado de forma simples e direta. Porém, precisamos garantir que a que a cadeia simulada convergiu para uma distribuição estacionária.

Para superar essa dificuldade, Heidenberg e Welch (1983) desenvolveram um método para testar a convergência de cadeias simuladas por métodos de amostragem MCMC.

O cálculo desta integral pode ser realizado através de alguma técnica de integração numérica já que é possível avaliar o integrando em qualquer ponto sem maiores dificuldades. Se a hipótese nula for rejeitada, o teste é repetido após descartar 10% das interações. Se a hipótese nula for rejeitada novamente, o processo é repetido descartando-se 10% das observações. Este processo é repetido até que uma porção da cadeia (de tamanho maior ou igual a 50%) passe no teste de estacionariedade, ou até que 50% das iterações sejam descartadas e a hipótese nula ainda seja rejeitada. Este teste de estacionariedade é mono caudal e a rejeição da hipótese nula ocorre quando o  $p$ -valor é maior que  $1 - \alpha$ .

Por outro lado, se a cadeia passar pelo teste de estacionariedade, outro teste é realizado para assertar se a cadeia de Markov simulada é suficiente para estimar com precisão a média do parâmetro em questão. A parte da cadeia escolhida para estimação é submetida a um teste baseado no intervalo de confiança para os parâmetros do modelo. A estatística calculada aqui é chamada de **relative half-width (RHW)** do intervalo de confiança.

O RHW mede a acurácia de um intervalo de confiança de nível  $1 - \alpha$  através da



razão entre o erro padrão da média e a própria média. Em suma, nós podemos parar a simulação da cadeia quando a variabilidade da média estabilizar com respeito a média real. Para realizarmos esse teste, precisamos selecionar um nível  $\alpha$  de confiança e uma tolerância predeterminada  $\epsilon$ . Por conseguinte, se o RHW calculado for maior que  $\epsilon$ , podemos concluir que não possuímos dados suficientes para estimar corretamente a média do parâmetro com nível  $1 - \alpha$  e tolerância  $\epsilon$ . Para mais detalhes sobre os métodos de inferência bayesiana veja Gelman (2003).

Já na abordagem clássica precisamos realizar a estimação da matriz Hessiana para obter estimativas da distribuição assintótica dos erros padrões das estimativas obtidas por máxima verossimilhança (veja Casella 2009). É importante ressaltar que os dois métodos de otimização disponíveis no software R possuem a opção de retornar uma estimativa da matriz Hessiana, porém podemos utilizar outra metodologia para estimação desta matriz. Como a matriz Hessiana se relaciona diretamente com a segunda derivada da função de verossimilhança, é natural que se aproxime essa matriz através da própria Definição de derivada. De forma sucinta, sendo  $\mathcal{L}_N(\theta)$  a função de verossimilhança do vetor de parâmetros  $\theta = (\theta_1, \theta_2, \dots)$  podemos estimar as derivadas parciais de segunda ordem da seguinte maneira:

$$\frac{\partial \mathcal{L}_N}{\partial \theta_i} = \lim_{\delta_i \rightarrow 0} \frac{\mathcal{L}_N(\dots, \theta_i + \delta_i, \dots) - \mathcal{L}_N(\dots, \theta_i - \delta_i, \dots)}{2\delta_i}.$$

Aproximando essa expressão para  $\delta_i$  suficientemente pequeno, derivamos uma aproximação para a segunda derivada parcial de  $\mathcal{L}_N(\theta)$  dada por:

$$\frac{\partial^2 \mathcal{L}_N}{\partial \theta_j \partial \theta_i} \approx \frac{\mathcal{L}_N(\dots, \theta_i + \delta_i, \dots, \theta_j + \delta_j, \dots) - \mathcal{L}_N(\dots, \theta_i - \delta_i, \dots, \theta_j + \delta_j, \dots)}{4\delta_j \delta_i} - \frac{\mathcal{L}_N(\dots, \theta_i + \delta_i, \dots, \theta_j - \delta_j, \dots) + \mathcal{L}_N(\dots, \theta_i - \delta_i, \dots, \theta_j - \delta_j, \dots)}{4\delta_j \delta_i}.$$

Se quisermos uma aproximação para essa derivada basta escolher  $\delta_i$  e  $\delta_j$  suficientemente pequenos. Wurtz et al. (2006) utilizou essa aproximação para a matriz Hessiana escolhendo  $\delta_i$  e  $\delta_j$  como sendo uma determinada porcentagem dos reais valores dos parâmetros  $\theta_i$  e  $\theta_j$  e escolhemos utilizar essa mesma abordagem para estimação dos modelos ARMA( $m, n$ )-GARCH( $p, q$ ).

Para a estimação dos parâmetros de uma série temporal  $X_t$  segundo o modelo ARMA-GARCH, precisamos inferir a série da variância condicional  $\sigma_t$ , pois a função

de verossimilhança é definida em termos das inovações  $z_t = \varepsilon_t/\sigma_t$ . Dado uma série temporal  $X_t = E(X_t|\Omega_{t-1}) + \varepsilon_t$  com vetor de parâmetros  $\theta$  procedemos da seguinte forma para o cálculo de  $\mathcal{L}_N(\theta)$ :

- Inicializamos o processo  $\varepsilon_t$  com  $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_{\max(m,n)}$  e procedemos recursivamente o cálculo recursivo de  $\hat{\varepsilon}_t$ ,  $t \geq \max(m,n)$  utilizando a equação  $\varepsilon_t = X_t - E(X_t|\Omega_{t-1})$ , substituindo  $\theta$  na média condicional do modelo.
- Inicializamos a variância condicional  $\hat{\sigma}_1, \dots, \hat{\sigma}_{\max(p,q)}$  e procedemos recursivamente o cálculo de  $\hat{\sigma}_t$ ,  $t \geq \max(p,q)$  utilizando a equação de volatilidade descrita em para o modelo GARCH, ou em (XX) para o modelo APARCH.
- Calculamos as inovações  $\hat{z}_t = \hat{\varepsilon}_t/\hat{\sigma}_t$  para computação da função de verossimilhança em  $\theta$ .

Por exemplo, se  $\varepsilon_t = z_t\sigma_t$  segue um processo GARCH com vetor de parâmetros  $\theta = (\omega, \alpha_1, \beta_1, \sigma^2)$  com  $z_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$ , a função de verossimilhança fica:

$$\begin{aligned} \mathcal{L}_N(\varphi) &= \ln \prod_t \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{\varepsilon_t^2}{2\sigma_t^2}} \\ &= \ln \prod_t \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{z_t^2}{2}} \\ &= -\frac{1}{2} \sum_t [\ln(2\pi) + \ln(\sigma_t^2) + z_t^2] \\ &= -\frac{1}{2} \sum_t \left[ \ln(2\pi) + \ln(\hat{\sigma}_t^2) + \left( \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t} \right)^2 \right] \end{aligned}$$

, onde  $\hat{\sigma}_t$  e  $\hat{\varepsilon}_t$  são calculados conforme descrito anteriormente. Desta forma, a função  $\mathcal{L}_N(\theta)$  define a verossimilhança para o modelo e pode ser utilizada para realizar inferências sobre o vetor de parâmetros  $\theta$ .

Neste trabalho utilizamos a abordagem clássica para estimação dos parâmetros do modelo através de métodos de otimização computacional no software R. Posteriormente, seguimos com a utilização da metodologia Bayesiana através do método MCMC Metropolis-Hastings Gibbs sampler para simular valores da distribuição a posteriori dos parâmetros do modelo.

Os resultados de simulação nesta etapa foram praticamente iguais aos resultados obtidos no trabalho de Zhao et al. (2011), porém o tempo computacional dos algorit-

mos que desenvolvemos não foi satisfatório. Felizmente, o desempenho das rotinas implementadas no pacote `MCMCPack` do software R foram semelhantes as nossas rotinas, porém com tempo computacional bastante inferior, já que a função `MCMCmetrop1R` implementada neste pacote utiliza o método Metropolis-Hastings de outra maneira.

Essa função na verdade, simula os valores de uma distribuição normal multivariada, com matriz de covariância  $V$  que pode ser ajustada pelo usuário. Em sua forma padrão, o método utiliza uma metodologia clássica para inferir quais seriam os verdadeiros parâmetros do modelo através da maximização da função de verossimilhança, gerando uma estimativa  $\hat{\theta}$  dos parâmetros e  $\hat{V}$  da matriz de covariância de  $\theta$ . Com isso, o algoritmo pode utilizar essa informação para ajustar a distribuição de onde estamos simulando os valores propostos. Finalmente, o algoritmo inicia a cadeia a partir de  $\hat{\theta}$  e simula valores de uma distribuição  $\mathcal{N}(\hat{\theta}, \hat{V})$ . Durante a estimação Bayesiana simulamos cadeias de tamanho 50000 e posteriormente aplicamos o teste de estacionariedade de Heidleberger e Welch (1983) para garantir que a cadeia convergiu para a distribuição de interesse.

As estimações realizadas nos próximos capítulos serão realizadas com dados simulados de séries temporais de tamanho 2500. Para cada um dos conjuntos de parâmetros utilizados mostramos a convergência das cadeias e apresentamos as medidas resumo das estimativas pelo método da verossimilhança e bayesiano. Para verificar a estabilidade dos métodos de estimação realizamos a simulação de 100 séries temporais e calculamos a média aritmética e o RMSE (Root Mean Square Error) das estimativas. Os resultados serão apresentados nos Gráficos e Tabelas que seguem nos próximos capítulos.

# Capítulo 4

## Estimação do Modelo

### GEV-GARCH

Neste capítulo, exploramos os aspectos computacionais envolvidos durante a estimação do modelo GEV-GARCH com as metodologias bayesiana e de máxima verossimilhança.

A maximização da função de verossimilhança pode ser feita utilizando a expressão para a densidade GEV descrita na Definição 1.21 assumindo-se  $\xi \neq 0$ .

**Definição 4.1.** Seja  $\{z_1, z_2, \dots, z_n\}$  uma amostra i.i.d. da distribuição GEV. Então a log-verossimilhança tem a seguinte forma:

$$\mathcal{L}_n(\theta|z_1, z_2, \dots, z_n) = \sum_{i=1}^n l_i(\theta), \quad (4.1)$$

onde  $l_i(\theta)$  denota a log-verossimilhança de uma só observação. Se  $\xi \neq 0$  temos:

$$l_i(\theta) = -\log(\sigma) - \left(\frac{1}{\xi} + 1\right) \log\left(1 + \xi \frac{z_i - \mu}{\sigma}\right) - \left(1 + \xi \frac{z_i - \mu}{\sigma}\right)^{-\frac{1}{\xi}}$$

e se  $\xi = 0$  obtemos:

$$l_i(\theta) = -\log(\sigma) - \frac{z_i - \mu}{\sigma} - e^{-\frac{z_i - \mu}{\sigma}}$$

Conforme ressalta Jondeau et al. (2007) a função de verossimilhança não linear e, portanto, a estimação deve ser realizada utilizando procedimentos de otimização não lineares. Durante a estimação devemos evitar que o algoritmo de otimização se mova para parâmetros que violem as restrições do modelo.

Além disso, precisamos evitar erros de cálculos computacionais quando avaliamos a função de densidade para  $\xi$  próximo de zero. Para resolver esta última limitação, estabelecemos um limite adequado para decidir qual das expressões na Definição 1.21 utilizar para o cálculo da verossimilhança. Esse procedimento é necessário pois a expressão geral quando  $\xi \neq 0$  é instável para o cálculo da densidade quando temos valores de  $\xi$  próximos de zero. Ainda com relação a distribuição GEV, vale ressaltar que existe uma restrição quanto ao utilização da função de verossimilhança em algoritmos de estimação. Segundo Jondeau et al. (2007) as propriedades assintóticas das estimativas de máxima verossimilhança da distribuição GEV valem para  $\xi > -0.5$ . Porém, essa dificuldade não limita a aplicação deste método na área de retornos financeiros já que tipicamente observamos quase sempre valores de  $\xi > 0$ .

Após todas essas considerações a respeito do modelo ARMA-GARCH, podemos definir agora a função de verossimilhança para o modelo AR(1)-GARCH(1,1) e utilizá-la em conjunto com algum método de otimização não linear.

Para analisar o desempenho do modelo e verificar a estabilidade da estimação dos parâmetros utilizamos dados simulados neste primeiro momento. Para efeitos de comparação com os resultados obtidos em Zhao et al. (2011), iremos utilizar três conjuntos de parâmetros diferentes para estimação,  $\theta_1$ ,  $\theta_2$  e  $\theta_3$  descritos na Tabela 4.1.

Tabela 4.1: Conjunto de parâmetros escolhido para simulação e estimação do modelo GEV-GARCH.

	$\mu$	$a_1$	$\omega$	$\alpha_1$	$\beta_1$	$\xi$
$\theta_1$	0.21	0.32	0.01	0.45	0.08	0.08
$\theta_2$	0.01	0.10	0.01	0.80	0.02	0.20
$\theta_3$	0.05	0.20	0.05	0.50	0.10	0.30

## 4.1 Estimação do modelo GEV-GARCH pelo método da Máxima Verossimilhança

A Tabela 9.1 contém as estimativas obtidas na estimação de três séries temporais simuladas de tamanho 2500. Os resultados da estimação mostraram que todos os parâmetros são significativos e não rejeitamos nenhuma das hipóteses, a um nível de 95% de confiança de que algum parâmetro seja diferente do valor estimado. Os erros padrões estimados foram pequenos em relação aos valores estimados, o que nos permite inferir que teremos intervalos de confianças com amplitude pequena para as estimativas dos parâmetros do modelo.

Para avaliar a estabilidade da estimação dos parâmetros do modelo construímos a Tabela 9.2. Neste estudo simulamos 100 séries temporais de tamanho 2500. Nosso conjunto paramétrico de estudo descrito na Tabela 4.1 considera somente valores positivos para  $\xi$ , pois raramente observamos valores negativos para essa variável em dados financeiros. O valor estimado de cada parâmetro é a média aritmética sobre as 100 séries temporais simuladas. Também apresentamos o valor de RMSE (root Mean Square Error) dos estimadores em relação ao valor real dos parâmetros. Observamos que o valor esperado para os estimadores está bem próximo do valor real e que a variação devido a amostra é pequena. Outra observação é que à medida que o parâmetro de forma  $\xi$  aumenta, o valor do RMSE diminui sugerindo que a estimação tem uma performance melhor quando temos caudas mais pesadas.

## 4.2 Estimação Bayesiana do Modelo GEV-GARCH

O trabalho de Zhao et al. (2011) utiliza inferência bayesiana para a estimação do modelo AR(1)-GARCH(1,1) que acabamos de mencionar. Zhao et al. (2011) separou o conjunto paramétrico  $\theta = (\mu, a_1, \omega, \alpha_1, \beta_1, \xi)$  em dois subconjuntos distintos:  $\theta_a = (\xi, \mu)$  que pode assumir qualquer valor em  $\mathcal{R}^2$  e  $\theta_b = (a_1, \omega, \alpha_1, \beta_1)$  que assume valores reais em um subconjunto limitado de  $\mathcal{R}^4$ . O objetivo do trabalho Zhao et al. (2011) era verificar a estabilidade das estimativas dos parâmetros deste modelo sem assumir a priori que possuímos informação sobre o processo inferencial, ou seja, sobre a distribuição de probabilidade dos parâmetros  $\theta$ . Dessa forma o autor utilizou

distribuições a priori normais para os parâmetros em  $\theta_a$  e distribuições uniformes em  $\theta_b$ .

Embora seja difícil assumir distribuições priori difusas para representar a falta de informação na estimação do modelo, o autor assumiu, por exemplo, distribuição a priori uniforme em  $[0, 0.2]$  para o parâmetro  $\omega$  na estimação dos parâmetros  $\theta_1, \theta_2$  e  $\theta_3$  caso em que  $\omega$  assumia valores iguais a, 0.01 e 0.05 e isso já pode ser considerado uma informação relevante para estimação do modelo. Por isso, optamos por assumir que os parâmetros do conjunto  $\theta_b$  são uniformemente distribuídos em todo o conjunto paramétrico do modelo, e não nos restringindo a um subconjunto deste. Assumimos também que o algoritmo inicia-se em um valor fixo que não tem, a princípio relação alguma com o verdadeiro valor dos parâmetros a serem estimados.

Utilizamos o algoritmo Metropolis-Hastings combinada com o amostrador de Gibbs para realizar a atualização de cada componente de  $\theta$  em cada passo da cadeia. Ao contrário do método utilizado em Zhao et al. (2011) que atualiza primeiro o parâmetro  $\xi$  da distribuição GEV, escolhemos atualizar os parâmetros sem nenhum critério, ou seja, atualizamos os parâmetros na seguinte ordem:  $\mu, a_1, \omega, \alpha_1, \beta_1$  e  $\xi$ .

Denotando por  $\pi(\cdot)$  a distribuição a priori do parâmetro  $\theta_i$  em  $\theta$  temos:

$$\pi(\theta_i) \propto I_{(\theta_i)}.$$

Dessa forma, a distribuição a posteriori dos parâmetros para o modelo  $X_t \sim \text{GEV}(\xi, \sigma_t, \mu_t)$  de uma série simulada  $X_1, X_2, \dots, X_n$  será dada por:

$$p(\theta|X) \propto \prod_{t=1}^n f_X(X_t|\xi, \sigma_t, \mu_t) \times \prod_{i=1}^6 I_{(\theta_i)}, \quad (4.2)$$

em que  $f_X(X_t|\xi, \sigma_t, \mu_t)$  é utilizada aqui para denotar a f.d.p. de uma distribuição GEV. É importante salientar que o algoritmo utilizado para realizar a amostragem precisa garantir que não sejam escolhidos parâmetros que violem algumas das restrições do modelo, pois isso pode interferir na convergência da cadeia simulada para a distribuição de interesse.

Considere que o conjunto de parâmetros  $\theta = (\theta_1, \dots, \theta_d)$  tenha  $d$  componentes. Para simplificar a notação utilizada nos próximos algoritmos defina  $\theta_{(-j)}^{(i-1)} = (\theta_1^{(i)}, \dots, \theta_{j-1}^{(i)}, \theta_{j+1}^{(i-1)}, \dots, \theta_d^{(i-1)})$ , ou seja, é o vetor de parâmetros **suprimido** da componente  $\theta_j$  cujas componentes anteriores a  $j$  são avaliadas no tempo  $i$  e as posteriores

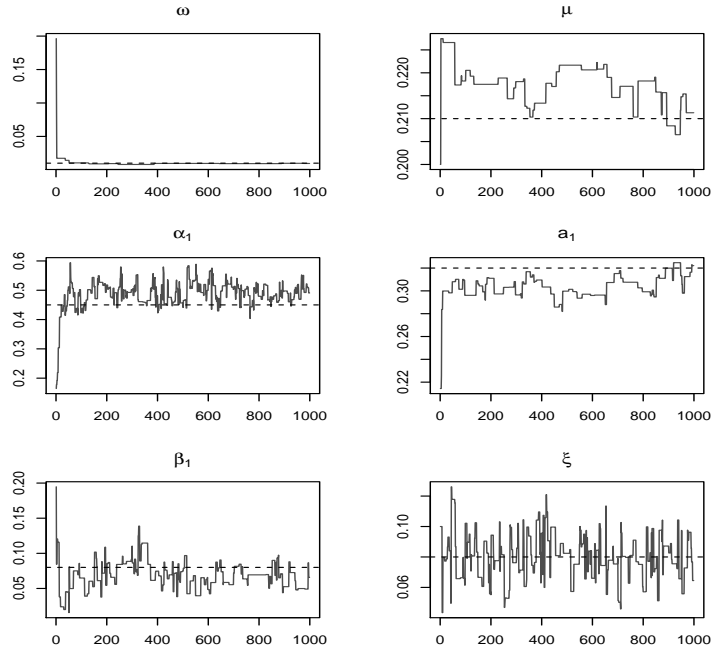


Figura 4.1: Estimaco MCMC com o Algoritmo 9.1 para estimaco dos parâmetros de uma srie temporal simulada do modelo GEV-GARCH com parâmetros  $\theta_1$ .

a  $j$  no tempo  $i - 1$ . J o vetor  $\theta_{(+j)}^{(i-1)} = (\theta_1^{(i)}, \dots, \theta_{j-1}^{(i)}, \theta_j^{(i-1)}, \theta_{j+1}^{(i-1)}, \dots, \theta_d^{(i-1)})$   o vetor de parâmetros **completo** cujas componentes at  $j$ , so avaliadas no tempo  $i$  e as componentes a partir de  $j$  avaliadas no tempo  $i - 1$ . Denote tambm  $p(x|\xi, \sigma_t, \mu_t) = \prod_{t=1}^n f_X(X_t|\xi, \sigma_t, \mu_t)$ . O algoritmo 9.1 descrito no Apndice resume os passos da estimaco MCMC implementados no software R para ajuste do modelo GEV-GARCH.

Inicialmente implementamos o algoritmo 9.1 no software R para realizar a simulao da distribuo a posteriori. A Figura 5.1 mostra os primeiros passos de uma cadeia simulada utilizando o algoritmo 9.1 para estimaco Bayesiana de uma srie simulada com parâmetros  $\theta_1$ . Neste grfico podemos observar a convergncia da cadeia para uma distribuo estacionria pois seus valores comeam a se dispersar em torno do real valor do parâmetro.

Outro fato importante a ser observado  que a convergncia para esta distribuo estacionria no pode depender de onde iniciamos o algoritmo. Este fato  ilustrado na Figura 4.2 que mostra a convergncia de diferentes cadeias simuladas para a mesma distribuo estacionria.

Embora as estimativas estivessem de acordo com o apresentado em Zhao et al. (2011), o tempo computacional deste algoritmo no foi satisfatrio. Por exemplo,



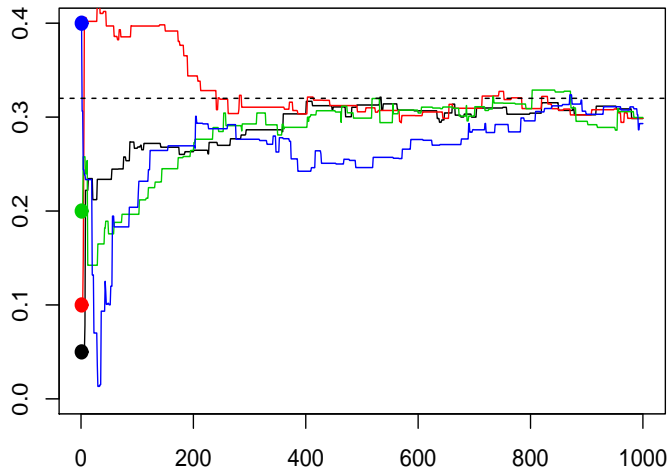


Figura 4.2: Convergência da cadeia simulada com o algoritmo 9.1 para simulação da distribuição a posteriori do parâmetro  $a_1 = 0.32$  do modelo GEV-GARCH.

para simulação de uma cadeia com 50000 passos em um computador Core2 Duo T5800 2.0GHz o algoritmo 9.1 demora em torno de 30 minutos. Como o objetivo de avaliar a estabilidade da estimação inclui a estimação de muitas séries temporais simuladas o tempo computacional é uma variável limitante para utilização deste algoritmo.

Esse problema foi contornado com a utilização da rotina `MCMCmetrop1R` disponível no pacote `MCMCpack` no software R. Essa rotina realiza a simulação de distribuições a posteriori simulando a partir de uma distribuição Normal Multivariada, e portanto, a aceitação ou rejeição de um novo valor da amostra é mais rápida que o método descrito no Algoritmo 1.

Após adaptarmos nossas rotinas implementamos a estimação Bayesiana para o modelo GEV-GARCH da Definição 3.3 utilizando a rotina `MCMCmetrop1R`. Com isso obtivemos resultados condizentes com a estimação que utilizava o algoritmo 9.1 ilustrada na Figura 5.1. Além do mais, obtivemos um excelente ganho de tempo computacional em utilizar a rotina `MCMCmetrop1R`. O tempo para simulação de uma cadeia de tamanho 50000 que antes era de aproximadamente 33 minutos caiu para 6 minutos. Sendo assim, optamos realizar a estimação Bayesiana do modelo GEV-GARCH em conjunto com a rotina `MCMCmetrop1R`. A Figura 4.3 mostra a convergência para a distribuição

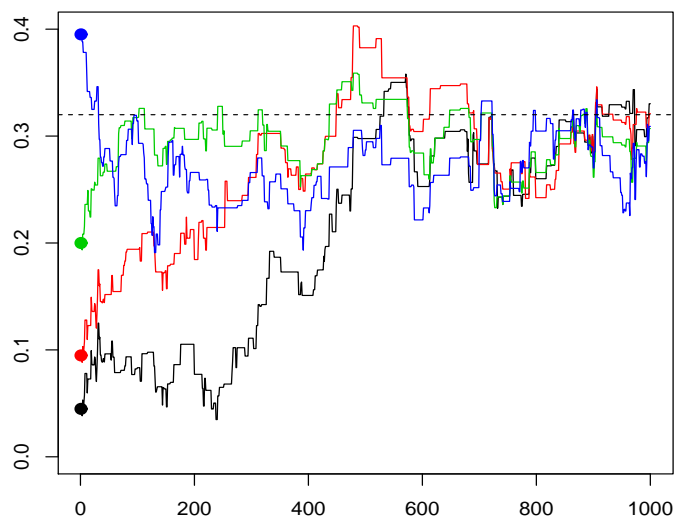


Figura 4.3: Convergência da cadeia simulada utilizando a rotina `MCMCmetrop1R` para simulação da distribuição a posteriori do parâmetro  $a_1 = 0.32$  do modelo GEV-GARCH.

a posteriori do parâmetro  $a_1 = 0.32$  utilizando a rotina `MCMCmetrop1R`. Ao comparar as Figuras 4.2 e 4.3 vemos que as cadeias simuladas com a rotina `MCMCmetrop1R` estão mais mixadas que as cadeias que utilizam o Algoritmo 9.1.

As Figuras 9.1, 9.2 e 9.3 mostram os gráficos das distribuições a posteriori estimadas da simulação de três séries temporais seguindo o modelo AR(1)-GARCH(1,1) com inovações GEV com parâmetros  $\theta_1$ ,  $\theta_2$  e  $\theta_3$  definidos na Tabela 4.1. Em todos os casos descritos a seguir o valor utilizado para inicializar o algoritmo de estimação foi igual a  $\theta_{inicial} = (0.1, 0.1, 0.1, 0.1, 0.1, 0.1)$ . Simulamos cadeias de tamanho 50000 e descartamos uma determinada quantidade de observações de acordo com o recomendado pelo teste de estacionariedade de Heidleberger e Welch (1983) implementado no pacote `boa` do software R. Os resultados do teste de Heidleberger e Welch para as séries estimadas na Tabela 9.3 estão descritos na Tabela 9.4. Esse fato foi confirmado também através do teste que utiliza a estatística HWD, pois em todos os casos o valor desta estatística nos levou aceitar a hipótese de que a cadeia restante possui a variabilidade da média estável em relação ao valor real da média do parâmetro. Cada uma das Figuras ilustra o comportamento da cadeia simulada pelo método Metropolis-Hastings MCMC descrito anteriormente e mostra a distribuição a posteriori simulada

dos parâmetros do modelo.

A Tabela 9.3 mostra o sumário das estimativas das distribuições a posteriori descritas nas Figuras 9.1, 9.2 e 9.3. Aqui o valor estimado do parâmetro é a média da distribuição a posteriori e o intervalo de credibilidade é um intervalo HPD (Máxima densidade a posteriori) com 95% de credibilidade. Todos os intervalos de confiança contém o verdadeiro valor do parâmetro e os valores estimados também foram bem próximos dos valores reais, exceto para o parâmetro  $\beta_0 = 0.01$  (conjunto paramétrico  $\theta_1$ ) cuja estimativa foi 0.012058 com um vício de 0.000788.

Já a Tabela 9.10 apresenta os resultados das estimativas quando simulamos 100 séries temporais deste modelo. Ao contrário do que aconteceu com essa mesma estimação para a estimativa pelo método da verossimilhança, o valor do RMSE não diminuiu com o aumento de  $\xi$ . O valor do RMSE foi também pequeno e praticamente igual aos valores obtidos utilizando-se o método MLE. As estimativas dos parâmetros também foram tomadas como a média das 100 estimativas Bayesianas obtidas da distribuição a posteriori.

Finalmente, a Tabela 4.2 mostra que os resultados obtidos aqui foram semelhantes aos resultados apresentados no trabalho de Zhao et al. (2011). Vale ressaltar que as medidas de RMSE obtidas aqui estão sujeitas a variações devido às séries que foram simuladas e, portanto, não podemos comparar de forma justa o desempenho da metodologia empregada neste trabalho e no trabalho de Zhao et al. (2011).

No próximo capítulo iremos tratar da estimação dos parâmetros do modelo STABLE-GARCH.

Tabela 4.2: Comparação da estimação Bayesiana GEV-GARCH deste trabalho com os resultados de Zhao et al. (2011)

		Simulado	Estimado	Estimado-Zhao	RMSE	RMSE-Zhao
	$\alpha_0$	0.01	0.0100	0.009	<b>0.0007</b>	0.0024
	$\alpha_1$	0.45	0.4509	0.480	<b>0.0282</b>	0.0827
$\theta_1$	$\alpha_2$	0.08	0.0805	0.078	0.0156	0.0136
	$\beta_0$	0.21	0.2114	0.210	0.0073	0.0075
	$\beta_1$	0.32	0.3162	0.321	0.0166	0.0163
	$\xi$	0.08	0.0808	0.072	<b>0.0146</b>	0.0244
	$\alpha_0$	0.01	0.0102	0.010	0.0012	0.0024
	$\alpha_1$	0.80	0.7944	0.793	0.0353	0.0385
$\theta_2$	$\alpha_2$	0.02	0.0209	0.021	0.0042	0.0043
	$\beta_0$	0.10	0.0103	0.010	0.0036	0.0064
	$\beta_1$	0.10	0.0982	0.010	0.0131	0.0122
	$\xi$	0.20	0.1988	0.201	<b>0.0213</b>	0.0168
	$\alpha_0$	0.05	0.0504	0.051	0.0037	0.0047
	$\alpha_1$	0.50	0.4984	0.498	0.0282	0.0196
$\theta_3$	$\alpha_2$	0.10	0.1013	0.099	0.0129	0.0094
	$\beta_0$	0.05	0.0504	0.051	0.0074	0.0113
	$\beta_1$	0.20	0.1995	0.201	0.0144	0.0154
	$\xi$	0.30	0.2961	0.297	0.0166	0.0171

# Capítulo 5

## Estimação do Modelo

### STABLE-GARCH

Neste capítulo iremos comentar sobre os aspectos principais envolvidos no cálculo das densidades estáveis. Posteriormente explicamos como foi realizada a estimação dos parâmetros do modelo STABLE-GARCH e apresentamos os resultados em tabelas e gráficos.

A primeira dificuldade que surge nessa etapa é sobre o cálculo das densidades estáveis, pois precisamos otimizar sobre a expressão de verossimilhança  $\mathcal{L}_N(\varphi) = \ln \prod_t \mathcal{D}_\vartheta(x_t, E(x_t|\Omega_{t-1}), \sigma_t)$  onde  $\mathcal{D}_\vartheta$  representa a função de densidade de uma variável aleatória estável  $S(\alpha, \beta, 1, 0)$ . De acordo com Sheneider (1986), a função densidade de probabilidade de uma variável aleatória com distribuição estável pode ser escrita em termos de uma função especial chamada função-H, também conhecida como Fox H-function. Neste mesmo artigo, Schneider (1986) apresentou uma representação em séries de potências para o cálculo de densidades estáveis. Porém, essa representação em séries é bastante instável, ou seja, não conseguimos estabelecer uma quantidade ideal de termos para realizar o truncamento da série e calcular valores da densidade com acurácia.

Uma observação importante é que esta mesma representação dada por Schneider (1986) pode ser encontrada na pág. 583 de Feller (1971), ou seja, a contribuição mais importante do trabalho de Schneider foi a conexão entre a função H e as distribuições estáveis. Para tentar resolver esse problema, Yilmaz e Alouini (2009) propuseram

algumas rotinas no software Matemática para computação da função H, porém o método não nos permite realizar os cálculos com rapidez fazendo ser impraticável a implementação de algum método de otimização envolvendo densidades estáveis representadas em termos de função H.

Ainda com relação as funções especiais, Penson e Gorsak (2011) proporam uma representação dessas densidades em termos de funções hipergeométricas para o caso em que os parâmetros alpha e beta são racionais. De fato, existe sim uma conexão entre o trabalho de Scheneider (1986) e o trabalho de Penson e Gorsak (2011), pois de acordo com Rathie et al. (2006), para alpha racional, a função H pode ser escrita em termos de uma função G (mais conhecida como Meijer G-function) e esta última pode também ser escrita em termos de funções hipergeométricas que é a representação dada por Penson e Gorsak (2011), ressaltando que neste último trabalho, foram apresentadas expressões para densidades assimétricas. Porém, não obtivemos sucesso em calcular essas densidades pois a velocidade de computação dos valores da densidade não é suficiente para implementação de métodos computacionais intensivos, tais como a otimização da verossimilhança.

Infelizmente o cálculo de densidades estáveis em termos de funções especiais ainda não é tão eficiente como os métodos de integração numérica utilizados para este fim. Atualmente, a melhor alternativa para computação de funções envolvendo distribuições estáveis tem suas bases no trabalho de Nolan (1997). Neste trabalho Nolan investigou a estabilidade numérica das representações existentes para o cálculo de densidades estáveis utilizando procedimentos de integração numérica. De acordo com Nolan (1997) a f.d.p. de uma variável aleatória estável na parametrização  $S_0$  definida na equação (1.19) pode ser expressada como:

- se  $\alpha \neq 1$  e  $z > \varsigma$ :

$$f_Z(z; \alpha, \beta, 1, 0) = \frac{\alpha(z - \varsigma)^{\frac{1}{\alpha-1}}}{\pi|\alpha - 1|} \int_{-\xi}^{\frac{\pi}{2}} V(u; \alpha, \beta) \exp\{-(z - \varsigma)^{\frac{1}{\alpha-1}} V(u; \alpha, \beta)\} du$$

- se  $\alpha \neq 1$  e  $z = \varsigma$ :

$$f_Z(z; \alpha, \beta, 1, 0) = \frac{\Gamma(1 + \frac{1}{\alpha}) \cos(\epsilon)}{\pi(1 + \varsigma^2)^{\frac{1}{2\alpha}}}$$

- se  $\alpha \neq 1$  e  $z < \varsigma$ :

$$f_Z(z; \alpha, \beta, 1, 0) = f_Z(-z; \alpha, -\beta, 1, 0)$$

- quando  $\alpha = 1$ :

$$f_Z(z; 1, \beta, 1, 0) = \begin{cases} \frac{1}{2|\beta|} e^{-\frac{\pi z}{2\beta}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V(u; 1, \beta) \exp\{-e^{-\frac{\pi z}{2\beta}} V(u; 1, \beta)\} du, & \beta \neq 0 \\ \frac{1}{\pi(1+z^2)} & \beta = 0, \end{cases}$$

onde

$$\epsilon = \begin{cases} \frac{1}{\alpha} \arctan(-\zeta) & \alpha \neq 1, \\ \frac{\pi}{2} & \alpha = 1, \end{cases}$$

$$V(u; \alpha, \beta) = \begin{cases} (\cos \alpha \epsilon)^{\frac{1}{\alpha-1}} \left( \frac{\cos u}{\sin \alpha(\xi+u)} \right)^{\frac{\alpha}{\alpha-1}} \frac{\cos[\alpha\xi+(\alpha-1)u]}{\cos u} & \alpha \neq 1 \\ \frac{2}{\pi} \left( \frac{\frac{\pi}{2} + \beta u}{\cos u} \right) \exp\left\{ \frac{1}{\beta} \left( \frac{\pi}{2} + \beta u \right) \tan u \right\} & \alpha = 1, \beta \neq 0. \end{cases}$$

Aqui  $f_Z(z; \alpha, \beta, 1, 0)$  é representada em termos da integral de uma função limitada e não oscilante (veja Nolan, 1997). Essa representação é bastante eficiente e pode ser utilizada em conjunto com métodos de integração numérica para calcular o valor das densidades estáveis com acurácia. De acordo com Nolan (1997) o método é acurado sempre que  $\alpha$  é maior que 0.4. Como em aplicações envolvendo finanças, quase sempre observamos  $\alpha > 1$  essa restrição não será um problema para nosso objeto de estudo. Existem algumas restrições para  $\alpha$  próximo de 1, mas isso somente quando  $|\alpha - 1| < 0.0001$ . Essas rotinas estão implementadas no programa chamado STABLE disponível em <http://academic2.american.edu/~jpnolan/stable/stable.html>.

O trabalho de Nolan foi implementado em um pacote estatístico para utilização no software R chamado `dstable` que realiza o cálculo das densidades estáveis de acordo com as representações descritas em Nolan (1997). Apesar das rotinas deste pacote serem melhores que os métodos usuais que utilizam densidades estáveis, ainda não são suficientemente rápidos para realização dos cálculos das densidades estáveis, fato esse que se deve ao modo como essas integrais foram implementadas.

Felizmente, existe uma versão comercial do programa STABLE para o software R, que implementa, além de funções de score e de distribuição, o cálculo da densidade de forma eficiente. Todas as rotinas foram escritas e compiladas na linguagem C e portanto superam em desempenho as rotinas citadas anteriormente. Utilizaremos essa versão do programa STABLE gentilmente cedida pela empresa Robust Analysis, Inc. e disponível em <http://www.robustanalysis.com>.

Para efeitos de comparação de tempo computacional entre os dois métodos, realizamos um experimento simples que consistiu em simular valores na reta real e calcular o valor da densidade de probabilidade com parâmetros  $S_{1.5}(0.5, 1, 0)$  utilizando a rotina `dstable` e posteriormente a versão comercial do programa `stable` mencionada anteriormente. Simulamos  $10^3$  valores aleatoriamente distribuídos na reta real e calculamos a função de densidade utilizando cada um desses pacotes. Em um computador Core 2 Duo 2Ghz a rotina `dstable` levou 100 segundos para realizar este cálculo, enquanto a versão comercial realizou o mesmo cálculo mais de mil vezes em apenas 0.65 segundos. Embora a rotina comercial para o cálculo da densidade no pacote `dstable` seja muito mais rápida que a rotina do pacote `dstable` implementado no programa R, a rotina comercial não possui tanta acurácia para a computação das densidades, mas já é suficiente para implementação de problemas de otimização. Em nosso experimento o erro máximo entre o cálculo dessas densidades por essas duas rotinas foi de  $10^{-5}$ .

Para realizar as estimativas do modelo utilizamos três conjuntos de parâmetros semelhantes aos escolhidos na estimação do modelo GEV-GARCH. O valor dos parâmetros estão descritos na Tabela 5.1.

Tabela 5.1: Conjunto de parâmetros escolhido para simulação e estimação do modelo STABLE-GARCH.

	$\mu$	$a_1$	$\omega$	$\alpha_1$	$\beta_1$	$\alpha$	$\beta$
$\varphi_1$	0.21	0.32	0.01	0.45	0.08	1.80	0.50
$\varphi_2$	0.01	0.10	0.01	0.80	0.02	1.90	0.50
$\varphi_3$	0.05	0.20	0.05	0.50	0.10	1.80	0.50

## 5.1 Estimação pelo método da Máxima Verossimilhança do modelo STABLE-GARCH

A Tabela 9.6 mostra o resultado da estimação de três séries temporais simuladas do modelo com inovações  $\alpha$ -estáveis. As estimativas estão bem próximas dos reais valores dos parâmetros do modelo e o erro padrão das estimativas foi também pe-



queno relativo ao verdadeiro valor do parâmetro nos fazendo inferir que os intervalos de confiança para os parâmetros do modelo terão amplitude pequena.

A Tabela 9.7 mostra o resultado das estimativas quando simulamos 100 séries temporais seguindo o modelo com inovações estáveis. De maneira análoga ao que foi apresentado na seção anterior sobre estimação do modelo GEV-GARCH o valor estimado de cada parâmetro é a média aritmética sobre as 100 séries temporais simuladas. Novamente o valor das estimativas do parâmetros ficou próximo dos verdadeiros valores e a variação devido a amostra foi pequena. No entanto, o RMSE para estimação do parâmetro de assimetria  $\beta$  da distribuição das inovações foi grande em relação aos outros parâmetros.

## 5.2 Estimação Bayesiana do Modelo STABLE -GARCH

Realizamos a estimação Bayesiana aqui com o objetivo de verificar a estabilidade das estimativas supondo que não possuímos informação a priori dos parâmetros do modelo. Este estudo é de extrema importância pois nos permite explorar os aspectos envolvidos durante a estimação desses modelos com distribuição condicional estável.

Neste caso o conjunto paramétrico é dado por  $\varphi = (\mu, a_1, \omega, \alpha_1, \beta_1, \alpha, \beta)$  e optamos por utilizar distribuição a priori uniforme para cada parâmetro do modelo com intervalo definido de acordo com o espaço paramétrico do modelo em que os parâmetros estão definidos. De maneira similar utilizamos o algoritmo Metropolis-Hastings combinado com o método de amostragem de Gibbs para realizar a atualização de cada componente de  $\varphi$  em cada passo da cadeia. Embora os parâmetros  $\alpha$  e  $\beta$  sejam os mais importantes, não os colocamos em primeiro lugar para realização da simulação a posteriori.

Denotando por  $\pi(\cdot)$  a distribuição a priori do parâmetro  $\varphi_i$  em  $\varphi$  do algoritmo temos:

$$\pi(\varphi_i) \propto \prod_{i=1}^7 I_{(\varphi_i)},$$

Segue então que a distribuição a posteriori dos parâmetros para o modelo  $X_t \sim$

$S_\alpha(\beta, \sigma_t, \mu_t)$  de uma série temporal  $X_1, X_2, \dots, X_n$  é dada por:

$$p(\theta|X) \propto \prod_{t=1}^n f_X(X_t|\alpha, \beta, \sigma_t, \mu_t) \times \prod_{i=1}^7 I_{(\theta_i)} \quad (5.1)$$

, onde  $f(X_t|\alpha, \beta, \sigma_t, \mu_t)$  é a função densidade de probabilidades de uma variável aleatória estável.

Durante a simulação da cadeia, precisamos garantir que as condições de estacionariedade e as restrições no espaço paramétrico dos parâmetros sejam respeitadas. Caso um determinado valor seja simulado e não atenda a essas exigências, ele é rejeitado para compor a cadeia e continuamos com o valor obtido no passo anterior.

O Algoritmo 9.2 descrito em detalhes no Apêndice mostra os passos para simulação da distribuição a posteriori para este modelo.

Inicialmente implementamos o algoritmo 9.2 no software R para realizar essa estimação. De maneira similar ao que aconteceu durante a estimativa do modelo GEV-GARCH, o tempo computacional para simulação das cadeias não foi satisfatório. Esse problema foi contornado com a utilização da rotina `MCMCmetropR1` disponível no pacote `MCMCpack` no software R. Com isso, implementamos a estimação do modelo STABLE-GARCH definindo uma função de verossimilhança para o modelo adequada de forma para permitir sua utilização pela rotina `MCMCmetropR1`. Os próximos gráficos mostram a convergência das cadeias simuladas com o algoritmo 9.2. A Figura 5.1 mostra os primeiros passos da simulação MCMC de todos os parâmetros do modelo e a Figura 5.2 mostra que a convergência de quatro cadeias iniciadas em locais diferentes.

Para utilização desta rotina, fixamos os valores iniciais dos parâmetros para todos os casos e a partir desses valores o algoritmo iniciava a simulação da cadeia. O valor inicial utilizado para todas as estimações descritas a seguir foi  $\phi_{inicial} = (0.1, 0.1, 0.1, 0.1, 0.1, 1.95, 0)$ .

Inicialmente não obtivemos êxito em utilizar a rotina `MCMCmetropR1` para simulação da distribuição a posteriori quando a série simulada tinha o parâmetro  $\alpha$  (índice de estabilidade da distribuição estável) próximo de 2. Mais especificamente isso aconteceu para o conjunto paramétrico  $\phi_2$  fazendo a `MCMCmetropR1` retornar um erro, pois não conseguia ser inicializada. Assim sendo, forçamos a inicialização iniciando do algoritmo com uma matriz Identidade para representar a função de Covariância do vetor de parâmetros que estamos interessados.

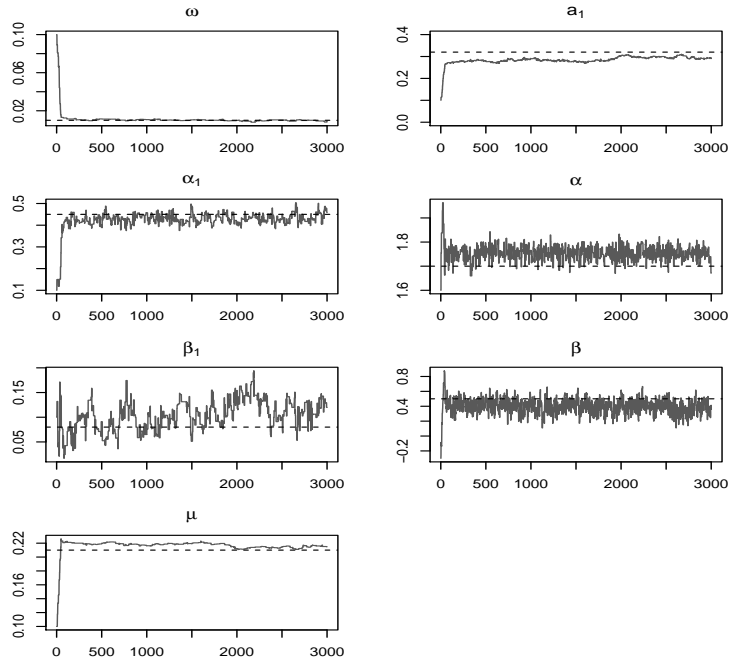


Figura 5.1: Estimação MCMC com o Algoritmo 9.2 para estimação dos parâmetros de uma série temporal simulada do modelo STABLE-GARCH com parâmetros  $\phi_1$ .

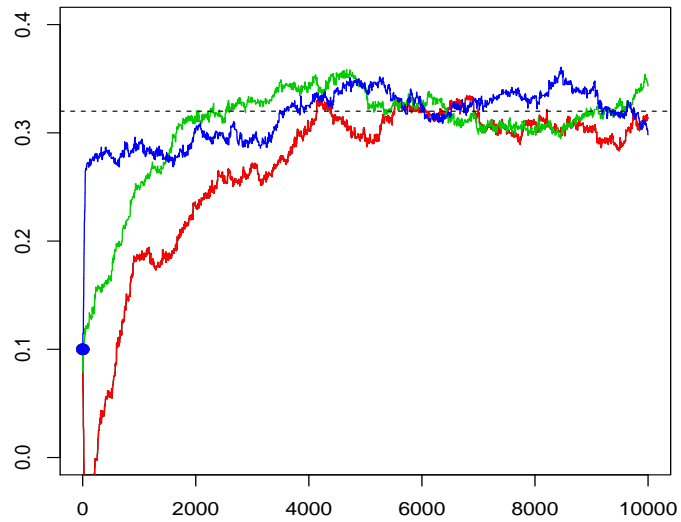


Figura 5.2: Convergência da cadeia simulada com o algoritmo 9.2 para simulação da distribuição a posteriori do parâmetro  $a_1 = 0.32$  do modelo STABLE-GARCH.

Essa abordagem resolve parcialmente nosso problema, pois a convergência para uma distribuição estacionária foi bastante lenta para o parâmetro  $\beta$  (assimetria da distribuição estável). Mesmo simulando cadeias de tamanho 70000 a distribuição a posteriori marginal deste parâmetro não estava convergindo. Em alguns casos a convergência para uma distribuição estacionária só era observada a partir de 30000 passos do algoritmo. Este último problema foi resolvido com a seguinte abordagem:

- Inicializamos o algoritmo com matriz de Covariância igual a uma matriz identidade, porém com diagonal igual a 0.00001. Na prática esse valor pequeno serve para regular a taxa de aceitação de novos elementos da cadeia.
- Simulamos a distribuição a posteriori até 10000 iterações e utilizamos as 2000 últimas observações para estimar a matriz de Covariância da distribuição a posteriori do conjunto paramétrico  $\phi$ .
- Por fim, inicializamos o algoritmo utilizando a matriz  $\hat{V}$  obtida na etapa anterior e simulamos cadeias de tamanho 50000.

A Figura 5.3 mostra a convergência para a distribuição a posteriori do parâmetro  $a_1 = 0.32$  utilizando a rotina `MCMCmetrop1R`. Novamente, comparando as Figuras 5.2 e 5.3 vemos que as cadeias do modelo estável simuladas com a rotina `MCMCmetrop1R` convergem mais rapidamente que as cadeias simuladas com o Algoritmo 9.1.

Apresentamos nas próximas Figuras a convergência das cadeias simuladas para a distribuição posteriori dos parâmetros, bem como as distribuições a posteriori após descartamos uma parte da cadeia através do teste de Heidenberg e Welch (1983). Todas as estimativas apresentadas foram realizadas utilizando a rotina `MCMCmetrop1R` do software R em conjunto com nossa implementação da função de verossimilhança do modelo STABLE-GARCH.

Simulamos cadeias de tamanho 50000 e descartamos uma quantidade de observações iniciais como recomendado pelo teste de estacionariedade de Heidleberger e Welch. As Figuras 9.4, 9.5 e 9.6 mostram a convergência das cadeias simuladas para a distribuição a posteriori dos parâmetros. Ao contrário dos modelos GEV-GARCH estimados anteriormente, as cadeias simuladas para estimação do modelo STABLE-GARCH não estavam convergindo tão rapidamente.

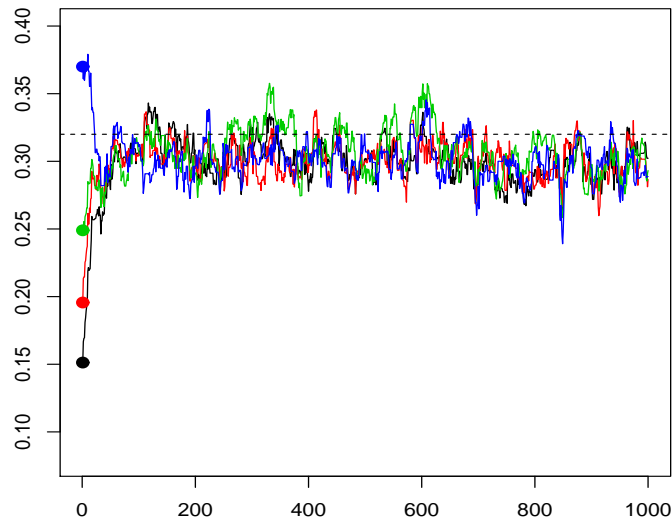


Figura 5.3: Convergência da cadeia simulada utilizando a rotina `MCMCmetrop1R` para simulação da distribuição a posteriori do parâmetro  $a_1 = 0.32$  do modelo STABLE-GARCH.

Para o conjunto de parâmetros  $\varphi_1$  a solução estacionário só podia ser encontrada quando o parâmetro  $\alpha$  da distribuição de inovações é próxima de 1.9 e talvez essa tenha sido a dificuldade da estimativa do modelo Bayesiano. No entanto, essa dificuldade não foi enfrentada na estimação do mesmo modelo pelo método da máxima verossimilhança.

Por último, as Tabelas 9.6 e 9.9 apresentam o sumário das estimativas Bayesianas para avaliar a estabilidade da estimação dos parâmetros do modelo STABLE-GARCH. A Tabela 9.10 mostra os resultados da estimação para simulação de 100 séries temporais de tamanho 2500. De forma análoga a estimação pelo método da máxima verossimilhança a estimação através do método Bayesiano apresentou resultados similares com RMSE alto para a estimação do parâmetro de assimetria  $\beta$  da distribuição de inovações.

No próximo capítulo iremos explorar com mais detalhes os modelos ARMA-power-GARCH com distribuição condicional estável definidos por Rachev et al. (2002).

# Capítulo 6

## Provas de estacionariedade

Os resultados da seção passada mostraram que ambas as metodologias de estimação utilizadas tiveram desempenhos semelhantes. Isso se deve ao fato de em ambos os casos estarmos utilizando somente a função de verossimilhança como escore para estimação dos parâmetros do modelo. No método da máxima verossimilhança essa função é utilizada para o cálculo do gradiente que guia o processo de otimização. Já no método de amostragem MCMC, a função de verossimilhança é utilizada para o cálculo da razão de aceitação/rejeição de novos elementos na amostra. Caso tivéssemos utilizado distribuições a priori informativas para os parâmetros do modelo a metodologia bayesiana teria desempenho diferente do método MLE. Porém, é importante ressaltar que a constatação de sucesso na estimação bayesiana neste tipo de modelo é um sinal positivo na direção de as estimativas seriam melhores caso obtivéssemos informação passada e a incluíssemos nessa metodologia.

Os modelos APARCH implementados por Wurtz et al. (2006) são largamente aplicados em conjunto com distribuições condicionais com variância finita, tais como normal, t-Student, GED (Wurtz et al. , 2006) e GEV (Zhao et al. , 2011). Porém, ainda não foram estabelecidas em que condições os modelos APARCH com distribuição condicional estável são estacionários. Nesse sentido, Rachev et al. (2002) provaram a estacionariedade de um caso particular desse modelo conhecido como power-GARCH. Esse modelo permite uma maior flexibilidade no expoente da equação de volatilidade do modelo da definição 3.8.

De acordo com a Definição 3.2 o modelo power-GARCH é na verdade um processo

ARMA-APARCH com  $\gamma_i = 0$  e distribuição condicional estável. Dessa forma a equação de volatilidade do modelo fica resumida a:

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i |\varepsilon_{t-i}|^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta \quad (6.1)$$

O próximo teorema apresenta as condições de estacionariedade forte do modelo STABLE-power-GARCH de Mitnik et al. (2002), já que neste caso, a distribuição das inovações tem variância infinita, e portanto, não podemos aplicar diretamente o resultado da estacionariedade trazido por Ding, Granger e Engle (1993) no Teorema 2.4.

**Teorema 6.1.** (Rachev et al., 2002) *O modelo STABLE-power-GARCH da Definição 6.1 tem uma única solução estritamente (fortemente) estacionária se:*

$$\lambda_{\alpha,\beta,\delta} \sum_{i=1}^r \alpha_i + \sum_{i=1}^s \beta_j \leq 1, \quad (6.2)$$

em que  $\lambda_{\alpha,\beta,\delta} = E(|z_t|^\delta)$  e o parâmetro  $\alpha$  desse coeficiente corresponde ao parâmetro de cauda da distribuição estável. É importante observar que se colocarmos  $\delta = 2$  na Definição da variância condicional, teremos  $E(|z_t|^\delta) = \infty$ . O parâmetro  $\lambda_{\alpha,\beta,\delta}$  depende da potência  $\delta$ , do índice de estabilidade  $\alpha$  e do parâmetro de assimetria da distribuição e, segundo o mesmo artigo, é dado por:

$$\lambda_{\alpha,\beta,\delta} = \kappa_\delta^{-1} \Gamma\left(1 - \frac{\delta}{\alpha}\right) (1 + \tau_{\alpha,\beta}^2)^{\delta/2\alpha} \cos\left(\frac{\delta}{\alpha} \arctan \tau_{\alpha,\beta}\right)$$

para  $\delta < \alpha$  e  $\alpha > 1$ , em que  $\tau_{\alpha,\beta} = \beta \tan(\alpha\pi/2)$  e :

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

É importante salientar que Rachev et al. (2002) demonstraram a estacionariedade forte dos modelos ARMA-power-GARCH e, portanto, não possui a flexibilidade do parâmetro de assimetria existente nos modelos ARMA-APARCH da Definição 3.2.

Devido à alta expansão na utilização dos modelos APARCH e das distribuições estáveis, é desejável que tenhamos um modelo assimétrico na equação de variância quando utilizamos distribuições com variância infinita.

Vamos então definir o modelo ARMA-APARCH que iremos tratar neste capítulo.

**Definição 6.1.** A série temporal  $X_t$  segue um processo STABLE-APARCH se pode ser escrito como um processo ARMA-APARCH da Definição 3.2 com distribuição condicional estável. Observe que esse modelo generaliza o modelo STABLE-power-GARCH definido em Rachev et al. (2002) pois a equação das inovações assume a seguinte forma

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (\varepsilon_{t-i} - \gamma_i |\varepsilon_{t-i}|)^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta, \quad (6.3)$$

com  $-1 < \gamma_i < 1$  arbitrário.

Em seu trabalho, Rachev et al. (2002) descreveu as principais condições de estacionariedade estrita do modelo Power-GARCH. Estas condições estão descritas a seguir:

- A potência na equação de variância  $\delta$  deve ser menor que o índice de cauda  $\alpha$  da distribuição estável.
- A esperança  $E|z|^\delta$  deve ser finita, onde  $z \sim S(\alpha, \beta, 0, 0)$ .

O próximo resultado generaliza as condições de estacionariedade para que tenhamos um modelo Assimétrico Power-ARCH com distribuição condicional estável. Iremos seguir basicamente os mesmos passos da demonstração apresentada por Rachev et al. (2002) utilizando alguns lemas definidos em Bungeoun e Picardo (1992), em Kesten e Spitzer (1984) e em referências citadas nestes artigos.

Seguindo a estratégia descrita no trabalho de Rachev et al. (2002) obtemos o seguinte Lema.

**Lema 6.1.** *Considere o modelo ARMA-APARCH com inovações estáveis  $z_t \stackrel{i.i.d}{\sim} S(\alpha, \beta, 0, 0)$  da Definição 6.1. Assuma também que a representação do vetor de estados  $\tilde{X}_t = (X_{1,t}, \dots, X_{m,t})$ ,  $m = \max(p, q)$  seja dada por:*

$$\tilde{X}_{k,t} = \begin{cases} \sigma_t^\delta, & \text{se } k = 1, \\ \sum_{i=k}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=k}^q \beta_j \sigma_{t-j}^\delta, & \text{se } k = 2, \dots, m.. \end{cases} \quad (6.4)$$

*Então, esse modelo tem a seguinte representação em espaço de estados:*

$$\tilde{X}_{t+1} = A_t \tilde{X}_t + B, \quad (6.5)$$



onde:

$$A_t = \begin{bmatrix} (|z_t| - \gamma_1 z_t)^\delta \alpha_1 + \beta_1 & 1 & 0 & \cdots & 0 \\ (|z_t| - \gamma_2 z_t)^\delta \alpha_2 + \beta_2 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ (|z_t| - \gamma_{m-1} z_t)^\delta \alpha_{m-1} + \beta_{m-1} & 0 & 0 & & 1 \\ (|z_t| - \gamma_m z_t)^\delta \alpha_m + \beta_m & 0 & 0 & \cdots & 0 \end{bmatrix}, B = \begin{bmatrix} \alpha_0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}. \quad (6.6)$$

*Demonstração.* Como  $m = \max(p, q)$ , então o processo ARMA-APARCH dado anteriormente pode ser reescrito da seguinte maneira:

$$\tilde{X}_{t+1} = F\tilde{X}_t + G(|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + B, \quad (6.7)$$

onde as matrizes F, G e B são definidas por:

$$F = \begin{bmatrix} \beta_1 & 1 & 0 & \cdots & 0 \\ \beta_2 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ \beta_{m-1} & 0 & 0 & & 1 \\ \beta_m & 0 & 0 & \cdots & 0 \end{bmatrix}, G = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{m-1} \\ \alpha_m \end{bmatrix}, B = \begin{bmatrix} \alpha_0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad (6.8)$$

Se tivermos  $r \neq s$ , então fazemos  $\alpha_i = 0$ , para  $i = p + 1, \dots, m$ , em G ou  $\beta_i = 0$ , para  $i = q + 1, \dots, m$ , em F.

Podemos escrever  $(|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta = \sigma_t^\delta (|z_t| - \gamma_i z_t)^\delta = e'_{1,m} X_t (|z_t| - \gamma_i z_t)^\delta$ , onde  $e'_{1,m} = [1, 0, \dots, 0]$  denota o vetor unitário de dimensão  $m \times 1$ . Neste caso a equação de transição (6.5) pode ser expressada em forma matricial da seguinte forma:

$$\tilde{X}_{t+1} = A_t \tilde{X}_t + B, \quad (6.9)$$

com:

$$\begin{aligned} A_t &= F + G e'_{1,m} (|z_t| - \gamma_i z_t)^\delta = \\ &= \begin{bmatrix} (|z_t| - \gamma_1 z_t)^\delta \alpha_1 + \beta_1 & 1 & 0 & \cdots & 0 \\ (|z_t| - \gamma_2 z_t)^\delta \alpha_2 + \beta_2 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ (|z_t| - \gamma_{m-1} z_t)^\delta \alpha_{m-1} + \beta_{m-1} & 0 & 0 & & 1 \\ (|z_t| - \gamma_m z_t)^\delta \alpha_m + \beta_m & 0 & 0 & \cdots & 0 \end{bmatrix} \end{aligned}$$

□

Iremos apresentar agora alguns resultados e lemas necessários para demonstração da estacionariedade do modelo da definição 6.1. A estratégia abordada por Rachev et al. (2002) e também empregada aqui, será a de utilizar representação matricial desse modelo para provar sua estacionariedade. De acordo com Bougerou e Picard (1992) o maior expoente de Lyapunov associado a uma sequência i.i.d. de matrizes aleatórias  $\{A_t, t \in \mathbb{Z}\}$  satisfaz a seguinte restrição:

$$\zeta \leq E(\log \|A_t\|) \quad (6.10)$$

**Lema 6.2.** (Bougerou e Picard, 1992) *Sempre que  $E(\log^+ \|A_t\|)$  for finita temos a seguinte convergência a respeito do maior expoente de Lyapunov:*

$$\zeta \stackrel{q.c.}{=} \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_t A_{t-1} \dots A_{t-n}\|. \quad (6.11)$$

**Lema 6.3.** *A matriz  $A_t$  definida na equação (6.6) satisfaz a restrição:*

$$E(\log^+ \|A_t\|) < \infty. \quad (6.12)$$

*Demonstração.* A primeira etapa é calcular a norma da matriz  $A_t$ . Observe que dado qualquer  $\tilde{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  temos:

$$A_t \tilde{x} = x_1 \begin{bmatrix} (|z_t| - \gamma_1 z_t)^\delta \alpha_1 + \beta_1 \\ (|z_t| - \gamma_2 z_t)^\delta \alpha_2 + \beta_2 \\ \vdots \\ (|z_t| - \gamma_{m-1} z_t)^\delta \alpha_{m-1} + \beta_{m-1} \\ (|z_t| - \gamma_m z_t)^\delta \alpha_m + \beta_m \end{bmatrix} + \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_m \\ 0 \end{bmatrix} \quad (6.13)$$

Para simplificar a notação defina:  $C_i = (|z_t| - \gamma_i z_t)^\delta \alpha_i + \beta_i$ ,  $\alpha^* = \max \alpha_i$  e  $\beta^* = \max \beta_i$ . Então utilizando a norma do máximo ( $\|\tilde{x}\| = \max(|x_1|, \dots, |x_m|)$ ) vemos que para todo  $i \in \{1, \dots, m\}$ :

$$\begin{aligned} |x_1 C_i + x_i| &\leq |x_1| |C_i| + |x_i| \\ &\leq \|\tilde{x}\| (1 + |C_i|) \\ &= \|\tilde{x}\| (1 + C_i) \\ &= \|\tilde{x}\| (1 + (|z_t| - \gamma_i z_t)^\delta \alpha_i + \beta_i) \\ &= \|\tilde{x}\| (1 + (1 - \gamma_i \text{sgn} z_t)^\delta |z_t|^\delta \alpha_i + \beta_i) \\ &\leq \|\tilde{x}\| (1 + 2^\delta |z_t|^\delta \alpha^* + \beta^*), \end{aligned}$$

e portanto:

$$\frac{\|A_t \tilde{x}\|}{\|\tilde{x}\|} \leq \frac{\|\tilde{x}\|(1 + 2^\delta |z_t|^\delta \alpha^* + \beta^*)}{\|\tilde{x}\|} = 1 + 2^\delta |z_t|^\delta \alpha^* + \beta^* \quad (6.14)$$

Utilizando a desigualdade anterior podemos impor um limite superior para  $\log^+ \|A_t\|$  da seguinte forma:

$$\begin{aligned} \log^+ \|A_t\| &= \max(\log \|A_t\|, 0) \\ &\leq \max\{\log(1 + |z_t|^\delta 2^\delta \alpha^* + \beta^*), 0\} \\ &= \log(1 + |z_t|^\delta 2^\delta \alpha^* + \beta^*) \end{aligned} \quad (6.15)$$

Agora, como a função  $\log$  é uma função côncava, podemos aplicar a desigualdade de Jensen na Definição 1.10 para ver que:

$$\begin{aligned} E(\log^+ \|A_t\|) &\leq E \log(1 + |z_t|^\delta 2^\delta \alpha^* + \beta^*) \\ &\leq \log\{E(1 + |z_t|^\delta 2^\delta \alpha^* + \beta^*)\} < \infty, \end{aligned} \quad (6.16)$$

em que esta última desigualdade segue de uma aplicação direta da Propriedade 1.6.  $\square$

**Lema 6.4.** *A sequência de matrizes  $\{A_t, A_{t-1}, \dots, A_{t-m}\}$  definidas em 6.6 são tais que não existe um subconjunto estrito  $I$  de  $\{1, 2, \dots, m\}$  tal que, se  $V$  é um subespaço gerado pelos vetores canônicos  $\{e_i, i \in I\}$ , então:*

$$A_t V + B \subset V, q.c.$$

*Demonstração.* Como o espaço  $V$  é gerado por um subconjunto estrito de  $\{e_1, e_2, \dots, e_m\}$  temos que existe  $i \in \{1, 2, \dots, m\}$  tal que todo elemento  $v \in V$  tem sua  $i$ -ésima coordenada igual a zero. Seja  $i^* = \min\{i : e_i \in V\}$ , ou seja,  $i^*$  é a menor coordenada dos elementos de  $V$  que é quase certamente diferente de zero. Se  $i^* > 1$ , então vemos através da segunda parte da equação (6.13) que existe  $x^* \in V$  tal que  $A_t x^*$  tem a  $(i^* - 1)$ -ésima coordenada diferente de zero e portanto  $A_t x^* \notin V$ . Por outro lado, se  $i = 1$ , tome  $x^*$  em  $V$  com a primeira coordenada diferente de zero. É fácil ver através da primeira parte da equação (6.13) que quase certamente teremos  $A_t x^*$  com todas as entradas diferentes de zero e, portanto  $A_t x^* \notin V$ .  $\square$

**Teorema 6.2.** *(Bougerol e Picard, 1992) Suponha que  $E(\log^+ \|A_t\|) < \infty$  e o Lema 6.4 seja válido. Então:*

i. Se o modelo definido em 6.5 tiver uma solução estritamente estacionária, então o maior expoente de Lyapunov  $\zeta$  associado com a sequência de matrizes  $\{A_t, t \in \mathbb{Z}\}$  é estritamente negativo.

ii. Por outro lado, se  $E(\log^+ \|B_t\|) < \infty$  e  $\zeta < 0$ , então para todo  $t \in \mathbb{Z}$  a série:

$$X_n = B + \sum_{k=1}^{\infty} A_t A_{t-1} \dots A_{t-k+1} B \quad (6.17)$$

converge quase certamente e o processo  $\{X_n, n \in \mathbb{Z}\}$  é a única solução estacionária do modelo definido na equação (6.5).

A demonstração desse teorema pode ser encontrada no artigo de Bugeol e Picard (1992) e será omitida aqui. O resultado deste teorema será utilizado na demonstração da seguinte proposição.

**Proposição 6.1.** *O modelo ARMA-APARCH com distribuição condicional estável da Definição 6.1 tem uma solução estritamente estacionária se, e somente se, o maior expoente de Lyapunov,  $\gamma$ , associado com as matrizes  $\{A_t\}$  é estritamente negativo. A série  $\tilde{X}_t = B + \sum_t^{\infty} A_t A_{t-1} \dots A_{t-i+1} B$  converge quase certamente para todo  $t$ ; e o processo  $\tilde{X}_t, t \in \mathbb{Z}$  é a única solução estritamente estacionária e ergódica da equação  $\tilde{X}_{t+1} = A_t \tilde{X}_t + B$ .*

*Demonstração.* De acordo com o Teorema 6.2, a existência da solução do modelo da Definição 6.1 é equivalente a existência da solução do modelo na equação (6.5) e é determinada pelo maior expoente de Lyapunov associado com a sequência  $\{A_t\}$ .

De acordo com o Lema 6.3 temos  $\zeta \leq E \log \|A_t\| < \infty$ . Como a matriz  $B$  da representação 6.5 é constante temos que  $E(\log^+ \|B\|) < 0$ . Portanto, podemos aplicar o Teorema 6.2 para ver que o processo  $\tilde{X}_t$  converge quase certamente e é a única solução estritamente estacionária da equação 6.5.  $\square$

O próximo resultado explicita as condições de estacionariedade de modelos que podem ser escritos na forma da equação (6.5). Antes de estabelecer esse resultado precisaremos de alguns resultados encontrados em Kesten e Slitzer (1984). De acordo com a equação (1.4) de Kesten e Slitzer (1984) se a sequência  $\{A_t\}$  for formada por matrizes não negativas, então vale a seguinte desigualdade envolvendo o maior

expoente de Lyapunov:

$$\zeta \leq \log\{r(EA_t)\} \quad (6.18)$$

. De acordo com este mesmo artigo temos também o seguinte lema:

**Lema 6.5.** (*Kesten e Slitzer, 1984*) *Suponha que a matriz  $A_t$  satisfaz as seguintes condições:*

*i.  $A_t(i, j) \geq 0$  com probabilidade 1 para todo  $1 \leq i, j \leq m$  e para algum  $d > 0$  temos:*

$$P\{A_t, A_{t-1}, \dots, A_{t-d} > 0 \text{ para todo } 1 \leq i, j \leq m\} > 0,$$

*ii.  $P\{A_t \text{ ter uma coluna ou uma linha toda igual a zero}\} = 0,$*

*iii.  $EA_t(i, j) < \infty$  para  $1 \leq i, j \leq m$  e  $r(EA_t) = 1.$*

*Então temos a seguinte desigualdade envolvendo o maior expoente de Lyapunov e o raio espectral da matriz  $A_t$ :*

$$\zeta < \log\{r(EA_t)\} \quad (6.19)$$

**Lema 6.6.** *A matriz  $A_t$  definida no modelo ARMA-APARCH satisfaz as condições (i) e (ii) do Lema 6.5.*

*Demonstração.* A condição (i) é sempre válida pois quase certamente qualquer linha ou coluna da matriz  $A_t$  tem pelo menos um elemento não nulo. Agora, observe primeiro que de acordo com a equação (6.6) qualquer elemento da primeira coluna dessa matriz é quase certamente não negativo. Para provar a propriedade (ii) precisamos garantir que com probabilidade diferente de zero podemos efetuar um número finito de multiplicações com as matrizes  $A_t, A_{t-1}, \dots$  de forma a obter uma matriz com todos elementos estritamente maiores que zero.

Podemos afirmar que com probabilidade diferente de zero a primeira coluna das matrizes  $A_t, \dots, A_{t-m}$  é estritamente maior que zero. Em particular, para o produto de duas matrizes, se fizermos  $P_2 = A_t A_{t-1}$  iremos obter uma matriz com todos os elementos da segunda coluna estritamente maiores que zero, pois  $A_{t-1}(1, 2) \neq 0$ . Além disso essa matriz terá o elemento  $P_2(1, 3) \neq 0$ . De forma semelhante, a matriz  $P_3 = P_2 A_{t-2}$  terá também todos os elementos da terceira coluna diferentes de zero.

Seguindo este mesmo raciocínio podemos provar por indução que, quase certamente, o produto  $A_t, \dots, A_{t-m}$  será uma matriz com elementos estritamente maiores que zero.  $\square$

**Lema 6.7.** (Rachev et al., 2002) *Seja  $A_t^*$  a matriz definida na equação (6.6) com  $\gamma_i = 0 \forall i$ . Então para todo  $s \in \mathbb{R}$  temos:*

$$\det(Is - EA_t) = s^{p+q-1} \left( 1 - \sum_{i=1}^p \lambda \alpha_i s^{-i} - \sum_{j=1}^q \beta_j s^{-j} \right) \quad (6.20)$$

**Proposição 6.2.** *O modelo ARMA-APARCH com distribuição condicional estável da Definição 6.1 é estacionário se  $\delta < \alpha$  e:*

$$\sum_{i=1}^r \lambda_i \alpha_i + \sum_{j=1}^s \beta_j \leq 1, \quad (6.21)$$

em que  $\lambda_i = E(|z_t| - \gamma_i z_t)$  e  $z \sim S(\alpha, \beta, 1, 0)$  é uma variável aleatória estável padrão.

*Demonstração.* Assim como no trabalho de Rachev et al. (2002) a prova dessa proposição está inteiramente ligada aos resultados trazidos nos trabalhos de Bougerou e Picard (1992) e Kesten e Slitzer (1984). Já que temos uma desigualdade envolvendo o maior expoente de Lyapunov e o raio espectral da matriz  $A_t$  (Lema 6.5) vamos estabelecer um limite superior para o raio espectral dessa matriz. Para isso precisamos olhar para a equação que nos permite encontrar os autovalores da matriz  $A_t$ . Nossa representação da matriz  $EA_t$  é ligeiramente diferente da representação de Rachev et al. (2002), pois ao invés de termos a primeira coluna da matriz  $A_t$  na equação (6.6) com coeficientes  $E|z_t|^\delta \alpha_i + \beta_i$ , temos  $E(z_t - \gamma_i |z_t|)^\delta \alpha_i + \beta_i$ . Se absorvermos a constante  $E(z_t - \gamma_i |z_t|)^\delta$  no coeficiente  $\alpha_i$  podemos aplicar o Lema 6.7 para ver que:

$$\det(Is - EA_t) = s^{p+q-1} \left( 1 - \sum_{i=1}^p (\lambda_i \alpha_i) s^{-i} - \sum_{j=1}^q \beta_j s^{-j} \right).$$

onde  $\lambda_i = E(z_t - \gamma_i |z_t|)^\delta$ . Vale notar que para o caso do modelo power-ARCH de Rachev et al. (2002),  $\lambda_i = \lambda_{\alpha, \beta, \delta} = E|z_t|^\delta$ . Se  $|s| > 1$  e aplicamos a inequação

$|a - b| \geq ||a| - |b||$  então:

$$\begin{aligned}
\det(Is - EA_t) &\geq \left| 1 - \left( \sum_{i=1}^p \lambda_i \alpha_i s^{-i} - \sum_{j=1}^q \beta_j s^{-j} \right) \right| \\
&\geq \left| 1 - \left| \sum_{i=1}^p \lambda_i \alpha_i s^{-i} + \sum_{j=1}^q \beta_j s^{-j} \right| \right| \\
&\geq 1 - \left| \sum_{i=1}^p \lambda_i \alpha_i s^{-i} + \sum_{j=1}^q \beta_j s^{-j} \right| \tag{6.22} \\
&\geq 1 - \left( \sum_{i=1}^p \lambda_i \alpha_i |s^{-i}| + \sum_{j=1}^q \beta_j |s^{-j}| \right) \\
&\geq 1 - \sum_{i=1}^p \lambda_i \alpha_i - \sum_{j=1}^q \beta_j.
\end{aligned}$$

Vamos separar a análise em termos da equação  $\sum_{i=1}^p \lambda_i \alpha_i + \sum_{j=1}^q \beta_j$ .

- i. Se  $\sum_{i=1}^p \lambda_i \alpha_i + \sum_{j=1}^q \beta_j = 1$  então utilizando a desigualdade anterior vemos que essa matriz não tem nenhuma autovalor  $s$  com  $|s| > 1$ . Por outro lado, a equação  $\det(I1 - EA_t) = 0$  é válida e, portanto um dos autovalores dessa matriz é igual a 1. Então o maior autovalor possível que a matriz pode ter é igual a 1 e assim segue da Definição 1.12 que o raio espectral de  $A_t$  é  $r(E(A_t)) = 1$ .
- ii. Agora se  $\sum_{i=1}^p \lambda_i \alpha_i + \sum_{j=1}^q \beta_j < 1$  vemos através da equação (6.22) que novamente a matriz  $A_t$  ela não possui nenhum autovalor  $s$  com  $|s| > 1$  e portanto o raio espectral de  $A_t$  satisfaz  $r(E(A_t)) < 1$ .

Por último, aplicando os Lemas 6.5 e 6.6 para o caso (i) temos que o maior expoente de Lyapunov  $\zeta < \log r(E(A_t)) = \log 1 = 0$ . Por outro lado, aplicando o lema 6.5 vemos que  $\zeta \leq \log r(E(A_t)) < \log 1 < 0$ . Portanto, em ambos os casos o maior expoente de Lyapunov é estritamente negativo.  $\square$

**Proposição 6.3.** *O coeficiente  $\lambda_i$  na equação (6.21) pode ser calculado através da seguinte equação:*

$$\begin{aligned}
\lambda_i &= E(|z| - \gamma_i z)^\delta \\
&= (1 + \gamma)^\delta E(z^\delta | z > 0) P(z > 0) \\
&\quad + (1 - \gamma)^\delta E((-z)^\delta | z < 0) P(z < 0),
\end{aligned}$$

onde  $z \sim S(\alpha, \beta, 1, 0)$ .

*Demonstração.* Defina a variável indicadora  $I_z$  de tal forma que:

$$I_z = \begin{cases} 1, & \text{se } z \geq 0, \\ 0, & \text{if } z < 0. \end{cases}$$

Deste modo podemos aplicar o operador esperança condicional a variável  $I_z$  obtendo:

$$\begin{aligned} E(|z| - \gamma_i z)^\delta &= E\{E[(|z| - \gamma_i z)^\delta | I_z]\} \\ &= E[(z - \gamma_i z)^\delta | I_z = 1]P(I_z = 1) \\ &\quad + E[(-z - \gamma_i z)^\delta | I_z = 0]P(I_z = 0) \\ &= (1 - \gamma)^\delta E(z^\delta | I_z = 1)P(I_z = 1) \\ &\quad + (1 + \gamma)^\delta E[(-z)^\delta | I_z = 0]P(I_z = 0). \end{aligned}$$

□

De posse dos resultados deste capítulo estamos prontos para criar uma rotina computacional para estimação dos modelos ARMA-APARCH com distribuição condicional estável. O próximo capítulo irá tratar dos aspectos computacionais envolvidos na estimação desses modelos.



# Capítulo 7

## Aspectos Computacionais

Neste capítulo, generalizamos o modelo AR(1)-GARCH(1,1) proposto por Zhao (2011) para modelos ARMA( $m, n$ )-APARCH( $p, q$ ) e avaliamos seu desempenho para estimação dos parâmetros do modelo. A implementação desses modelos é testada com alguns modelos ARMA-GARCH para os dados DEMGBP de taxas de câmbio dados por Bollerslev and Ghysels descrito em Wurtz et al. (2006) e disponíveis no pacote `fGarch` do software R. No artigo de Wurtz et al. (2006) as distribuições condicionais normal, t-Student e skew t-student são aplicadas juntamente com os modelos ARMA-GARCH e por esse motivo, também implementamos as mesmas distribuições para efeitos de comparação de nossas estimativas. Os resultados serão apresentados nas Tabelas seguintes.

Além de comparar nossa implementação para estimação dos modelos ARMA-APARCH, chamada aqui de `GarchFitExtremes` iremos compará-la a implementação `GarchFit` de Wurtz et al. (2006). Ao contrário desta última rotina (amplamente utilizada na por pesquisadores na área de séries temporais) que implementa esses modelos com distribuições condicionais normal, t-Student, GED e suas versões assimétricas, a rotina `GarchFitExtremes` é capaz de realizar a estimação de modelos ARMA-APARCH com as seguintes distribuições condicionais: normal, t-Student, skew-t-Student, GEV e estável.

Outro ponto importante é sobre o cálculo das densidades estáveis. A estimação dos modelos de séries temporais da rotina `GarchFitExtremes` depende fortemente da rotina de cálculo rápido das densidades estáveis cedida por John Nolan através do

pacote `stable`.

A Generalização trazida pelos modelos APARCH é salientada no trabalho de Wurtz et al. (2006), pois inclui, além dos modelos ARCH de Engle e GARCH de Bollerslev, cinco outras extensões dos modelos ARCH, a saber:

- Modelo TS-GARCH de Taylor e Shwert (1986) quando  $\delta = 1$  e  $\gamma_i = 0$ .
- Modelo GJR-GARCH de Glosten, Jagannathan e Runkle (1993) quando  $\delta = 2$ .
- Modelo T-ARCH de Zakoian (1993) quando  $\delta = 1$ .
- Modelo N-ARCH de Higgens e Bera (1992) quando  $\gamma_i = 0$  e  $\beta_j = 0$ .
- Modelo Log-ARCH de Geweke e Pentula (1986) quando  $\sigma \rightarrow 0$ .

A prova de todas essas equivalências pode ser encontrada em Ding, Granger e Engle (1993). Atualmente, a implementação da função `garchFit` no R permite a especificação das seguintes distribuições condicionais: normal, t-sudent, GED (Distribuição de erro generalizada) e suas versões assimétricas.

As distribuições assimétricas implementadas no pacote `GarchFit` utilizam o método de Fernandez e Steel (1998) que permite introduzir um parâmetro de assimetria em distribuições unimodais contínuas e simétricas a partir da seguinte formula:

$$f(z|\beta) = \frac{2}{\beta + \frac{1}{\beta}} \left[ f(\beta z)H(-z) + f\left(\frac{z}{\beta}\right)H(z) \right],$$

onde o parâmetro de assimetria satisfaz  $0 < \beta < \infty$ . A função é definida como  $H(z) = (1 + \text{sgn}(z))/2$ . As expressões para média e variância serão dadas em termos do parâmetro de assimetria da distribuição. Se esse parâmetro for igual a 1, então recuperamos a densidade original. As versões assimétricas das distribuições Normal e t-Student estão implementadas na rotina `GarchFit` do R para utilização em conjunto com os modelos ARMA-GARCH/APARCH.

O trabalho de Wurtz (2006) apresenta sugestões de como utilizar o poder da função `filter` no software R para estimação dos modelos ARMA-GARCH. Wurtz mostra o seguinte trecho de código que pode ser utilizado calcular recursivamente a série de volatilidade  $\sigma_t$  crucial para a avaliação da função de verossimilhança do modelo. O seguinte trecho de código está apresentado na página 7 de Wurtz (2006) para o cálculo

da variância condicional:

```
# Filtro na variância Condicional - Com laço - Wurtz et al. (2006):  
> for (i in(uv+1):N )  
+ ed = 0  
+ for (j in 1:u)  
+ ed = ed+alpha[j]*(abs(eps[i-j])-gamma[j]*eps[i-j])^delta  
+ h[i] = omega + ed + sum(beta*h[i-(1:v)])
```

Essa representação não é eficiente, pois utiliza um laço `for` com o tamanho da série temporal e isso tem baixa desempenho computacional quando utilizado no software R. Wurtz et al. (2006) propõe a seguinte representação que é muito mais eficiente pois utiliza a rotina `filter` ao invés do laço `for`.

```
# Filtro na variância Condicional - Wurtz (2006):  
> edelta = (abs(eps)-gamma*eps)^delta  
> edeltat = filter(edelta, filter = c(0, alpha), sides = 1)  
> c = omega/(1-sum(beta))  
> h = c( h[1:uv], c + filter(edeltat[-(1:uv)], filter = beta,  
+ method = "recursive", init = h[uv:1]-c))
```

O código funciona muito bem quando utilizamos em modelos com simétricos, ou seja, quando  $\gamma = 0$ , e este é o caso para o exemplo que Wurtz et al. (2006) apresentou. Porém, quando tentamos utilizar a mesma representação para modelos assimétricos, os valores reportados para o processo  $\sigma_t$  diferem significativamente dos valores retornados pelo método mais lento utilizando-se um laço `for`. Para superar essa dificuldade fizemos um outro procedimento para o cálculo da volatilidade em casos assimétricos que também evita a utilização de um laço com o tamanho da série temporal. O código é o seguinte:

```

# Filtro na Variância Condicional para modelos APARCH:
> h <- rep(0.1, pq)
> edeltat = 0
> for( i in 1:p)
> {
> edelta <- alpha[i]*(abs(z)-gm[i]*z)^delta
> edeltat = edeltat + edelta[(p-(i-1)):(N-i)]
> }
> edeltat = c(h[1:p],edeltat)
> c <- omega/(1-sum(beta))
> h <- c( h[1:pq], c + filter(edeltat[-(1:pq)], filter = beta,
> method = "recursive", init = h[q:1]-c))
> hh <- abs(h)^(1/delta)

```

Embora esse procedimento utilize um laço `for`, esse laço tem tamanho igual a ordem  $p$  do modelo APARCH, e portanto é tão eficiente quanto a representação dada por Wurtz et al. (2006). Apresentamos abaixo o exemplo sugerido no trabalho de Wurtz para ilustrar o comportamento das três rotinas citadas anteriormente.

```

> N = 5; z = round(rnorm(N), digits = 2) / 10
> omega = 0.4; alpha = c(0.1, 0.6); gamma = c(0.7, -0.53)
> beta = c(0.1, 0.54); delta = 1.58
> u = length(alpha); v = length(beta); uv = max(u,v)
> # h - laço for
[1] 0.1000000 0.1000000 0.4954465 0.5070474 0.7238669
> # h - Representação Wurtz
[1] 0.1000000 0.1000000 0.4805949 0.5046803 0.7133731
> # h - Nossa Representação
[1] 0.1000000 0.1000000 0.4954465 0.5070474 0.7238669

```

O exemplo serve para mostrar que o cálculo da variância condicional realizado com a representação eficiente de Wurtz não retornou corretamente os valores da série de volatilidade. Porém, nossa representação retornou exatamente os mesmos valores que os valores calculados com a representação em laço `for`, que contém os valores corretos.

Apesar de bastante utilizadas em aplicações de finanças, as distribuições estáveis e GEV não foram implementadas no pacote `fGARCH` e é por esse motivo que implementamos a estimação de modelos ARMA-GARCH/APARCH com essas distribuições.

Com o objetivo de comparar nossa implementação de estimação no software R utilizaremos o banco de dados `dem2gbp` disponível no pacote `fGarch`. Esses dados contém 1975 observações diárias do período de 2 de Janeiro de 1984 a 31 de Dezembro de 1991 e se refere a taxa de câmbio envolvendo o Marco Alemão e a Libra Britânica. Esses dados foram utilizados para testar e comparar a implementação da rotina `garchFit` com outros softwares de estimação existentes. De acordo com Wurtz (2006), devido a complexidade dos modelos GARCH é evidente que diferentes implementações de software levem a diferentes resultados numéricos demonstrando que podem sim haver diferenças significantes quando utilizamos diferentes softwares na estimação destes modelos. Como as implementações da função `garchFit` já foram testada em relação aos outros pacotes de estimação iremos focar na comparação de nossa implementação com a implementação existente no pacote `garchFit`.

A primeira comparação é em relação ao modelo GARCH(1,1) descrito em Bollerslev (1996). A Tabela 7.7 apresenta os dados da estimação com as duas rotinas citadas anteriormente.

Tabela 7.1: Comparação na estimação dos parâmetros para os dados `dem2gbp` no modelo normal-GARCH(1,1).

Parâmetros	GarchFit	GarchFitExtremes
$\mu$	-0.006190	-0.006197
$\omega$	0.010761	0.010058
$\alpha_1$	0.153134	0.147264
$\beta_1$	0.805974	0.814682
LLH	-1106.608	-1104.243

Podemos concluir que os resultados para estimação do modelo GARCH(1,1) de Bollerslev foi bem similar em nosso algoritmo com os resultados implementados na rotina `garchFit`. Embora haja semelhança nas estimativas de máxima verossimilhança encontrada pelos dois modelos, o valor da verossimilhança encontrada pela rotina `garchFitExtremes` foi superior indicando melhor ajuste nos parâmetros do modelo. Também comparamos estes mesmos resultados utilizando a distribuição condicional t-student. Os resultados estão descritos na Tabela 7.2. Novamente, temos indícios

Tabela 7.2: Comparação na estimação dos parâmetros para os dados dem2gbp no modelo t-Student-GARCH(1,1).

Parâmetros	GarchFit	GarchFitExtremes
$\mu$	0.002240	0.002098
$\omega$	0.002319	0.002272
$\alpha_1$	0.124430	0.120906
$\beta_1$	0.884653	0.887283
$\nu$	4.118430	4.148788
LLH	-989.408	-987.045

de que existe acurácia em nossa rotina de estimação pois os valores dos parâmetros para o modelo GARCH(1,1) de Bollerslev estão concordando com as estimativas encontradas pela rotina `GarchFit`.

De posse desses resultados, iremos utilizar nossa rotina para estimação do mesmo modelo com as distribuições GEV e estável. Para estes casos não podemos comparar os resultados de nossas estimações com a implementação de Wurtz et al. (2006). Até o presente momento não conhecemos rotinas para estimação desses modelos com distribuições GEV e estável.

As estimativas do modelo com distribuição condicional GEV não foram próximas dos valores estimados com o modelo normal, a despeito da média  $\mu$  que também foi próxima de zero. Já o modelo com distribuição condicional estável se ajustou com parâmetros próximos ao modelo t-Student, devido ao fato desta última distribuição também ter caldas mais pesadas que a da distribuição normal. Vale ressaltar ainda que o modelo com distribuição condicional estável também forneceu melhor verossi-

Tabela 7.3: Comparação na estimação dos parâmetros para os dados dem2gbp nos modelo GARCH(1,1) com distribuição condicional GEV e estável.

		Estimativa	Erro Padrão	Valor t	Pr( $t >  t $ )
GEV	$\mu$	-0.065704	0.096664	-0.679744	0.496666
	$\omega$	0.046267	0.048651	0.951000	0.341604
	$\alpha_1$	0.430080	0.153448	2.802776	0.005066
	$\beta_1$	0.522252	0.193567	2.698042	0.006974
	$\xi$	-0.232709	0.020766	-11.20613	< 2.22e-16
	LLH	-1307.219			
	STABLE	$\mu$	0.017839	0.008925	1.99871
$\omega$		0.006275	0.002458	2.55274	0.010688
$\alpha_1$		0.106561	0.016682	6.38744	1.6869e-10
$\beta_1$		0.888101	0.018554	47.8640	< 2.22e-16
$\beta$		-0.360698	0.106160	-3.39768	0.000679
$\alpha$		1.749815	0.030657	57.0759	< 2.22e-16
LLH		-984.7173			

milhança ajustada que o modelo t-Student.

Nossa implementação na função `garchFitExtremes` também nos permite realizar a estimação de modelos APARCH. Iremos comparar nossa implementação do modelo ARMA(1,1)-APARCH(1,1) com distribuições condicionais normal e t-Student. A Tabela 7.4 mostra os resultados obtidos com as duas rotinas em questão. As estimativas dos parâmetros  $\mu, \omega, \alpha_1, \gamma_1$  e  $\beta_1$  estão de acordo com as estimativas obtidas pela rotina original `garchFit`. O maior erro relativo nessas estimativas foi igual a 0.095373 e correspondeu ao parâmetro  $\delta$ . A Tabela 7.5 mostra as estimativas do mesmo modelo quando utilizamos distribuição condicional t-Student. Novamente, a diferença na estimação dos parâmetros  $\mu, \omega, \alpha_1, \gamma_1, \beta_1, \delta$  foi praticamente imperceptível se levarmos em conta o erro relativo dessas estimativas. Já para os parâmetros  $a_1, b_1$  e  $\delta$  não foram tão satisfatórias, porém o maior erro relativo obtido foi 0.220347 para o parâmetro  $a_1$ . A seguir apresentamos as estimativas para o mesmo modelo com distribuições condicionais GEV e estável. Embora em alguns casos o valor da verossimilhança estimada seja ligeiramente menor em nossa implementação na rotina

Tabela 7.4: Comparação na estimação dos parâmetros para os dados dem2gbp no modelo normal-ARMA(1,1)-APARCH(1,1).

Parâmetros	GarchFit	GarchFitExtremes
$\mu$	-0.014198	-0.009919
$a_1$	-0.433343	-0.392834
$b_1$	0.485347	0.444245
$\omega$	0.023692	0.026085
$\alpha_1$	0.179850	0.174844
$\gamma_1$	0.093452	0.105305
$\beta_1$	0.790301	0.797760
$\delta$	1.368196	1.237706
LLH	1099.134	1097.072

`garchFitExtremes` esta diferença pode ser atribuída ao modo como as duas rotinas foram internamente implementadas. Nossa implementação foi escrita somente utilizando a linguagem nativa do software R, enquanto que as rotinas de otimização da rotina original no pacote `fGarch` realizam a otimização com rotinas implementadas em código `Fortran`. Além disso, a implementação de Wurtz et al. (2006) utiliza rotinas SQP (sequential quadratic programming) para realizar a otimização da função de verossimilhança, lançando mão das rotinas de otimização disponíveis no `software R`.

Em contraste, nossas rotinas utilizam dois métodos diferentes de otimização, a saber, a rotina `nlminb` e a rotina `solnp` de otimização SQP implementada no pacote `Rsolnp`. Em praticamente todos os casos não há diferença significativa entre as estimativas utilizando as rotinas `nlminb` ou `solnp`, porém o tempo computacional da rotina `solnp` é, em geral, sempre inferior ao método `nlminb`.

Para avaliar a estabilidade da estimação de modelos ARMA-APARCH iremos realizar a estimação dos de 100 séries temporais simuladas do modelo ARMA(1,1)-APARCH(1,1) com distribuições condicionais GEV e estável. As Tabelas estão descritas no apêndice C. Os resultados da estimação do modelo estável foram bastante satisfatórios, inclusive para o parâmetro de assimetria desta distribuição com RMSE menor que 0.1 na estimação dos modelos com  $\beta \in \{0.5, -0.7\}$ .



Tabela 7.5: Comparação na estimação dos parâmetros para os dados dem2gbp no modelo t-Student-ARMA(1,1)-APARCH(1,1).

Parâmetros	GarchFit	GarchFitExtremes
$\mu$	0.000495	0.000314
$a_1$	-0.505978	-0.394537
$b_1$	0.542778	0.431499
$\omega$	0.005905	0.005964
$\alpha_1$	0.138504	0.135576
$\gamma_1$	0.127676	0.126038
$\beta_1$	0.882806	0.883467
$\delta$	1.343409	1.362619
$\nu$	4.123448	4.205216
LLH	-982.8418	-982.1621

Finalmente, apresentamos uma comparação do tempo computacional para estimação dos modelos utilizando a rotina `garchFit` do software R e nossa rotina `garchFitExtremes`. Utilizamos o banco de dados dem2gbp para estimação dos modelos com ordens diferentes. O tempo computacional foi calculado em um computador com processador Core 2 Duo 2.0 GHz.

No próximo capítulo iremos utilizar nossa rotina `garchFitExtremes` para realizar a estimação de modelos ARMA-GARCH com dados reais. Iremos nos ater somente às técnicas e procedimentos atualmente utilizados para seleção, diagnóstico e previsão de parâmetros futuros com dados reais.

Tabela 7.6: Comparação na estimação dos parâmetros para os dados dem2gbp no modelo ARMA(1,1)-APARCH(1,1) com distribuição GEV ou estável.

		Estimativa	Erro Padrão	Valor t	Pr( $t >  t $ )
GEV	$\mu$	-0.065704	0.010955	-5.993921	< 2.22e-16
	$a_1$	-0.237749	0.200469	-1.186186	0.23555
	$b_1$	0.360630	0.195781	1.842176	0.06545
	$\omega$	0.051264	0.009119	5.622537	< 2.22e-16
	$\alpha_1$	0.305528	0.046328	6.594552	< 2.22e-16
	$\gamma_1$	0.128408	0.057168	2.351064	0.01872
	$\beta_1$	0.677799	0.057168	11.85658	< 2.22e-16
	$\delta$	1.299585	0.286989	4.527556	6e-06
	$\xi$	-0.227203	0.007000	-32.45701	< 2.22e-16
	LLH	1291.527			
STABLE	$\mu$	0.014560	0.008868	1.64193	0.100606
	$a_1$	-0.449642	0.307309	-1.46316	0.143424
	$b_1$	0.483370	0.307180	1.57357	0.115586
	$\omega$	0.002901	0.001752	1.65579	0.097763
	$\alpha_1$	5.25823	0.017170	6.594552	< 2.22e-16
	$\gamma_1$	0.112163	0.072236	1.55274	0.120486
	$\beta_1$	0.898262	0.018176	49.41935	< 2.22e-16
	$\delta$	1.338673	0.191246	6.99974	< 2.22e-16
	$\beta$	-0.329141	0.101034	-3.25773	0.001123
	$\alpha$	1.751208	0.030096	58.18724	< 2.22e-16
LLH	981.44				

Tabela 7.7: Comparação do tempo computacional em segundos entre as rotinas `garchFit` e `garchFitExtremes` com distribuição condicional normal.

Modelo	GarchFit	GarchFitExtremes
GARCH(1,1)	0.27	0.92
APARCH(1,1)	0.73	1.28
AR(1)-GARCH(1,1)	0.44	1.03
AR(1)-APARCH(1,1)	0.95	1.24
ARMA(1,1) - GARCH(1,1)	0.72	1.08
ARMA(1,1) - APARCH(1,1)	2.71	2.79
ARMA(2,2) - GARCH(2,2)	2.71	2.79
ARMA(2,2) - APARCH(2,2)	3.84	3.41

# Capítulo 8

## Aplicação com Dados de Finanças

Nesta seção analisamos os aspectos práticos envolvidos na identificação de modelos GARCH. Esses modelos são em geral aplicados a dados de financeiros envolvendo preços de ativos negociados no mercado de ações. Seja  $P_t$  o preço de um determinado ativo ao final do dia  $t$ . O log retorno desse ativo é definido como sendo  $r_t = \ln(P_t/P_{t-1})$ . Como iremos trabalhar com distribuições de valores extremos, iremos também analisar uma série temporal de máximos retornos diários em gráfico de 15min. As séries temporais que iremos analisar estão descritas na Tabela 8.1:

Tabela 8.1: Séries de dados reais utilizadas neste trabalho.

Série	Período	Frequência	Tamanho
Google Inc. (GOOG)	1 de Janeiro 2004 a 27 Março 2013	diária	2165
Microsoft Corporation (MSFT)	1 de Janeiro 2003 a 27 Março 2013	diária	2535
S&P 500 (SPS)	13 de Dezembro 2004 a 8 Outubro 2009	15 minutos	1224

Os dados da Microsoft e da Google foram obtidos no sítio <http://finance.yahoo.com> e os dados de S&P 500 foram gentilmente cedidos pela professora Sílvia Regina Costa Lopes da UFRGS. A Figura 8.1 mostra o gráfico das séries dos retornos e dos retornos ao quadrado. Este último gráfico nos permite notar a existência de conglomerados de volatilidade. Podemos também constatar que existem correlações significativas para os retornos quadrados e absolutos. As Figuras 8.2, 8.3 e 8.4 mostram as funções de autocorrelação (ACF) e de autocorrelação parcial (PACF) amostrais

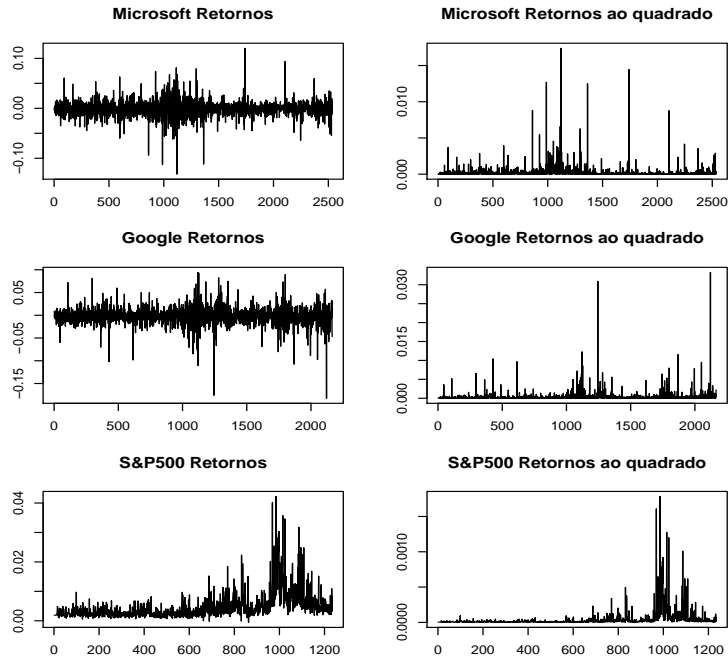


Figura 8.1: Retornos das séries temporais analisadas.

para os dados que estamos analisando. Uma análise preliminar dessas funções nos permite ver que, apesar de não possuir correlação significativa nos retornos, o mesmo não acontece com os retornos ao quadrado, ou com os retornos absolutos. Essa é uma característica dos modelos GARCH, que revela a não constância da volatilidade desses processos. Outra característica importante desses modelos é a formação de conglomerados de volatilidade, que pode ser notada no gráfico destas séries temporais.

A tabela 8.1 apresenta um sumário com algumas estatísticas das séries temporais analisadas.

Tabela 8.2: Sumário das estatísticas das séries temporais de retornos  $r_t$ .

Série	Média	Med.	Min	Max	Desv.Pad.	Assim.	kurt.	JB	KS
Microsoft	-0.0001	0.0003	-0.1319	0.1202	0.0147	-0.2297	10.8692	12525	0.4750
Google	-0.0009	-0.0001	-0.1824	0.0943	0.0188	-1.0331	11.7795	12931	0.4729
S&P500	0.0046	0.0032	-0.0005	0.0422	0.0045	3.4348	16.7254	16748	0.4998

Para todas as séries temporais em questão a curtose calculada foi maior que a

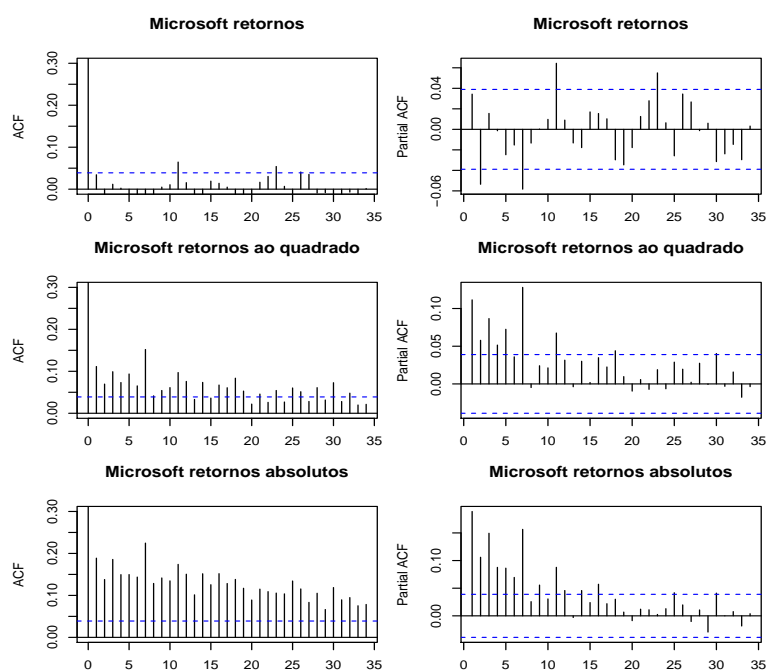


Figura 8.2: Estrutura de autocorrelação para os dados de log-retornos da empresa Microsoft.

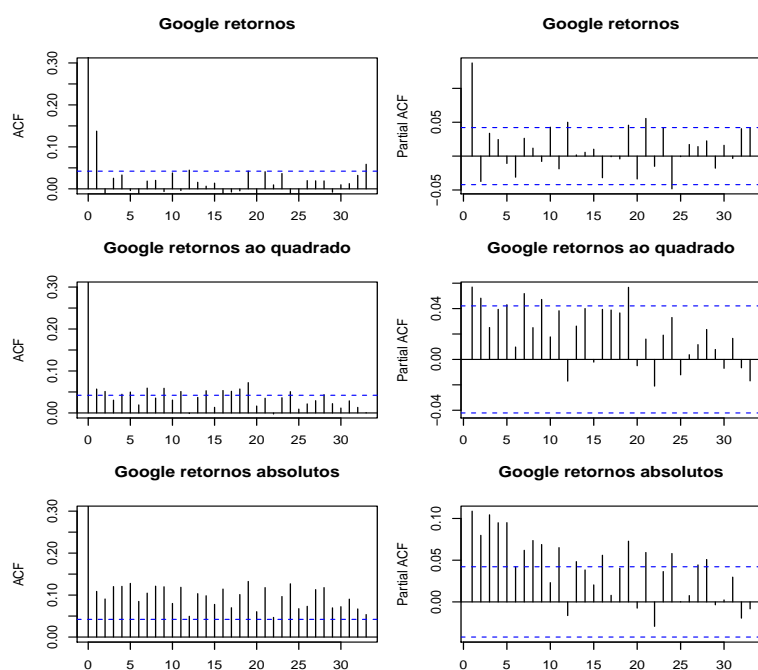


Figura 8.3: Estrutura de autocorrelação para os dados de log-retornos da empresa Google.

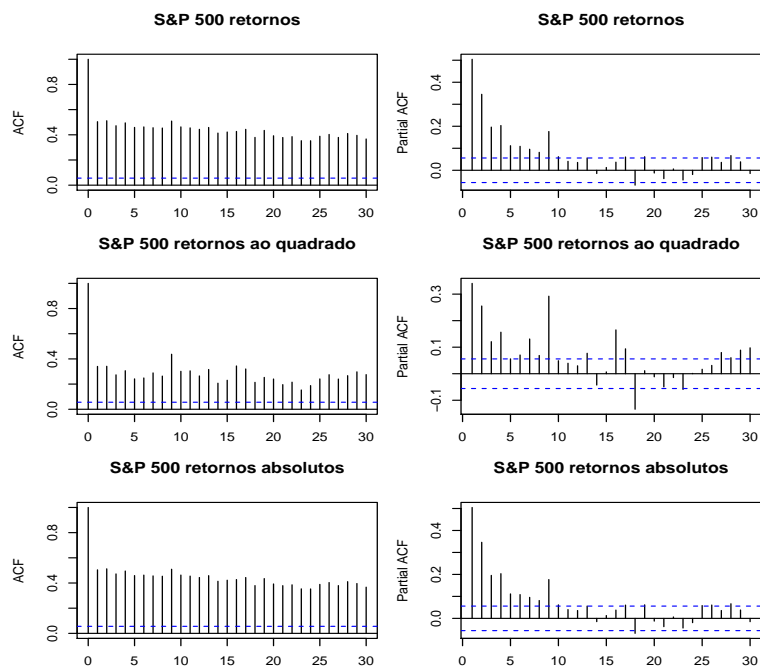


Figura 8.4: Estrutura de autocorrelação para os dados de log-retornos do índice S&P500.

da distribuição normal, o que é confirmada pelos testes de normalidade empregados aqui. Todos os p-valores para o teste de normalidade de Kolmogorov-Smirnov (KS) e Jarque Bera (JB) foram menores que  $2.2e-16$  nos fazendo rejeitar a hipótese de que esses dados seguem uma distribuição normal. Para mais detalhes sobre esses testes veja Brockwell e Davis (1992) e Brockwell e Davis (1996).

Embora nossos dados não sejam i.i.d., essa análise preliminar nos permite inferir que os retornos financeiros não possuem distribuição normal, possuindo assimetria diferente de zero e com kurtose significativamente maior que a da distribuição normal.

A rejeição da hipótese de não normalidade nos retornos de séries financeiras foi também constatado no trabalho de Zivot (2008). Zivot ainda ressalta que, mesmo após a retirada de outliers, o conjunto resultante continua não apresentando características de que são normalmente distribuídos.

Outro fato observado em retornos financeiros é a formação de conglomerados de volatilidade. De acordo com Zivot (2008) podemos utilizar a função de autocorrelação do valor absoluto dos retornos financeiros para testar a hipótese de existência de conglomerados de volatilidade em séries temporais.

É importante salientar que os conglomerados de volatilidade nos retornos financeiros se manifestam através da correlação existente nos retornos ao quadrado e absolu-

tos. Um dos testes comumente utilizados é o teste de Ljung-Box. Conforme descrito em Zivot. E. (2008) a significância da função de autocorrelação pode ser testada utilizando o teste de Ljung-Box:

$$LB(k) = n(n + 2) \sum_{j=1}^k \frac{\hat{\rho}_{X_j}^2}{T - j},$$

onde  $\hat{\rho}_{X_j}$  denota a correlação amostral na defasagem (lag)  $j$  dos retornos ao quadrado ou dos retornos absolutos e  $n$  é o tamanho amostral. Se os dados foram de ruído branco, então a estatística  $LB(k)$  tem distribuição assintótica qui-quadrado com  $p$  graus de liberdade, e portanto, um valor significante dessa estatística é um indicio de que os dados possuem volatilidade variando com o tempo.

A Tabela 8.3 mostra uma análise de nossos dados para teste dos efeitos ARCH utilizando a estatística  $LB(k)$  de Lijung-Box. Os valores da estatística do teste de

Tabela 8.3: Teste de Lijung-Box para retornos ao quadrado

Série	$\rho_X$	1	5	10
Microsoft		31.54360 (1.9e-08)	104.6734 (2.2e-16)	195.7521 (2.2e-16)
Google		7.032500 (8.0e-3)	24.46750 (1.7e-3)	45.21720 (2.0e-06)
S&P500		141.8573 (2.2e-16)	562.6379 (2.2e-16)	1174.424 (2.2e-16)

Lijung-Box para o teste de efeitos ARCH nos resíduos ao quadrado revelam evidência de formação de conglomerados de volatilidade.

## 8.1 Seleção do modelo

Uma vez que os dados foram transformados (por exemplo, por uma combinação da transformação de Box-Cox e de operadores de diferença para remoção de tendências ou de componentes sazonais) e estão prontos para serem ajustados a um modelo



estacionário, nos deparamos com o problema de selecionar apropriadamente a ordem do modelo. Não é vantajoso do ponto de vista de previsão de parâmetros futuros escolher um modelo com ordem alta. De acordo com Brockwell e Davis (2002), o ajuste de modelos com uma grande quantidade de parâmetros normalmente resulta em uma alta verossimilhança, porém se o mesmo modelo for utilizado para realizar previsão, o erro quadrático médio das previsões irá depender também do erro de estimação dos parâmetros do modelo. Estes erros serão grandes para modelos de alta ordem e por isso utilizamos os critérios de seleção de modelos que introduzem um fator penalizador para desencorajar a seleção de modelos com muitos parâmetros.

A identificação dos modelos GARCH é geralmente realizada utilizando-se os métodos tradicionais de seleção de modelos AIC (Akaike information criterion) e BIC (bayesian information criterion). De acordo com Zivot (2008) durante a seleção de modelos GARCH( $p, q$ ), as ordens  $p, q \leq 2$  são tipicamente selecionadas pelos critérios AIC e BIC. O autor também ressalta que os modelos GARCH( $p, q$ ) são geralmente preferidos do que um modelo ARCH( $p$ ) por razões de parcimoniosidade e da estabilidade numérica durante a estimação, pois os modelos GARCH de ordem alta possuem vários máximos e mínimos locais. Zivot termina concluindo que na prática o modelo GARCH(1, 1) é dificilmente superado.

A seleção de modelos para séries temporais é normalmente feita minimizando-se algumas das estatísticas mencionadas anteriormente.

**Definição 8.1.** O critério AIC (Critério de Informação de Akaike) para um modelo com  $r$  parâmetros é dado por:

$$\text{AIC} = -2 \ln \mathcal{L}_N(\theta) + 2r.$$

Em seu livro de introdução as series temporais, Brockwell (2002) explica o funcionamento do software estatístico ITSM2000 que seleciona a ordem de um modelo de série utilizando o critério AICc, que é o critério de informação de Akaike corrigido.

**Definição 8.2.** O critério de informação de Akaike corrigido (AICc) para um modelo com  $r$  parâmetros ajustado em uma amostra de tamanho  $n$  é dado pela seguinte expressão:

$$\text{AICc} = -2 \ln \mathcal{L}_N(\theta) + \frac{2(r+1)n}{(n-r-2)}.$$

**Definição 8.3.** O critério de informação Bayesiano BIC para ajuste de um modelo com  $r$  parâmetros em uma amostra de tamanho  $n$  é dado por:

$$\text{BIC} = -2 \ln \mathcal{L}_N(\theta) + r \ln n.$$

Embora comumente utilizados na literatura, esses critérios dependem do fato de que as distribuições de probabilidade utilizadas no modelo tenham variância finita e portanto, não são adequados para o caso das distribuições estáveis. No entanto, alguns autores utilizam esses critérios para identificação de modelos com distribuições estáveis, já que este critério pode ser entendido intuitivamente como uma penalização para evitar a super parametrização do modelo baseado no valor da verossimilhança ajustada para o modelo. Veja por exemplo, Nolan (1998), Lindsey (2004) e Oral (2012).

Na prática alguns trabalhos aplicam diretamente os modelos GARCH(1,1) a dados de finanças, sem se preocupar com a seleção correta da ordem do modelo. O trabalho de Alberg et al. (2008) aplica diretamente o modelo GARCH(1,1) com as distribuições normal e t-Student, mas não avalia a especificação de outros valores para a ordem do modelo. Já o trabalho de Frain (2009) utiliza as distribuições estáveis para estimação de modelos ARMA-GARCH fixando a ordem desses modelos. Neste mesmo trabalho, o autor mostra que a utilização das distribuições estáveis é mais poderosa para realização do cálculo de uma medida de risco em ativos financeiros conhecida como VAR (Valor em Risco). O trabalho de Zhao et al. (2011) combina a distribuição GEV com o modelo AR(1)-GARCH(1,1), porém aplica esse modelo a dados reais sem se preocupar com fatores de seleção da ordem adequada do modelo aos dados reais.

A seguir explicamos os procedimentos realizados para o ajuste de um modelo adequado aos dados das empresas Microsoft, Google e do índice S&P500.

Procedemos com o teste da hipótese de estacionariedade do modelo e, posteriormente, com o ajuste desses dados ao modelo estacionário ARMA-GARCH. O primeiro teste de estacionariedade utilizado é o de Phillips-Perron (PP) cuja hipótese nula é de que a série tem uma raiz unitária contra a hipótese alternativa de que a série é estacionária. O segundo teste aplicado é conhecido como o teste aumentado de Dickey-Fuller (ADF) para a hipótese nula de que a série tem uma raiz unitária contra

a hipótese alternativa de que a série é estacionária. Esses dois testes implementados no software R através das rotinas, `PP.test` e `adf.test`, respectivamente. A Tabela 8.4 mostra os resultados destes testes aplicados as séries temporais GOOGLE, Microsoft e S&P500.

Tabela 8.4: Testes de estacionariedade em séries temporais

Série	PP	ADF
Microsoft	-48.5813 (0.01)	-13.5451 (0.01)
Google	-40.4347 (0.01)	-11.8551 (0.01)
S&P500	-26.3292 (0.01)	-4.2589 (0.01)

Os resultados mostram que para as séries Microsoft e Google, nós rejeitamos a hipótese de existência de raiz unitária nos testes PP e ADF.

Neste capítulo utilizamos utilizar o critério de seleção AICc para escolher entre o modelo ARMA(m,n)-GARCH(p,q) mais adequado para ajustar nossas séries temporais. Para um determinado conjunto de dados iremos ajustar modelos de ordem até 2, ou seja,  $(m, n, p, q) \in \{0, 1, 2\}^2 \times \{1, 2\} \times \{0, 1, 2\}$ . Apesar dessa abordagem ser computacionalmente mais intensiva do que a usual de ajustar primeiro um modelo ARMA e posteriormente um modelo GARCH, acreditamos ter melhor êxito analisando todas as combinações destes modelos. Todos os modelos foram estimados utilizando a rotina `garchFitExtremes` apresentada no capítulo anterior. As próximas Tabelas mostram o valor das estatísticas AIC, AICc e BIC, bem com o valor da verossimilhança calculada para o modelo ARMA-GARCH. Aqui mostramos os resultados de estimação de vários modelos ARMA-GARCH com a rotina `garchFitExtremes` e explicitamos o valor da verossimilhança calculada para cada modelo.

A Tabela 8.5 mostra um sumário da estimação de vários modelos ARMA-GARCH com distribuição condicional estável para a série de dados da Microsoft. A verossimi-

Tabela 8.5: Ajuste do modelo ARMA-GARCH com distribuição condicional estável para os dados da Microsoft.

Modelo	-LLH	AIC	AICc	BIC
ARMA(0,0)-GARCH(1,0)	-7526.599	-15047.20	-15047.19	-15029.68
ARMA(0,0)-GARCH(1,1)	-7617.101	-15226.20	-15226.19	15202.85
ARMA(0,1)-GARCH(1,0)	-7534.232	-15060.46	-15060.45	-15037.11
ARMA(0,1)-GARCH(1,1)	-7625.920	<b>-15241.84</b>	<b>-15241.80</b>	<b>-15212.65</b>
ARMA(1,0)-GARCH(1,0)	-7534.074	-15060.15	-15060.13	-15036.80
ARMA(1,0)-GARCH(1,1)	-7625.521	-15241.04	-15241.02	-15211.85
ARMA(1,1)-GARCH(1,0)	-7534.307	-15058.61	-15058.59	-15029.42
ARMA(1,1)-GARCH(1,1)	<b>-7626.408</b>	-15240.82	-15240.78	-15205.79

lhança ajustada para o modelo ARMA(0,1)-GARCH(1,1) foi igual a 7625.920 e está bem próximo da verossimilhança ajustada para o modelo ARMA(1,1)-GARCH(1,1) que foi igual a 7626.408. Embora o último modelo tenha maior verossimilhança, essa diferença não foi suficiente para justificar a inclusão de mais um parâmetro no modelo.

Conforme podemos ver através da Tabela 8.6, os três critérios de seleção do modelo foram unânimes na decisão de aceitar o modelo ARMA(0,0)-GARCH(1,1) com distribuição condicional GEV. A verossimilhança ajustada para esse modelo foi igual a 6809.361 e ficou bem próxima da verossimilhança ajustada para o modelo ARMA(1,1)-GARCH(1,1) não justificando a inclusão de mais parâmetros neste modelo.

Após demonstrar a capacidade da rotina `garchFitExtremes` em estimar modelos ARMA-GARCH com diferentes ordens, estamos prontos para utilizar o critério AICc da Definição 8.2 para escolher o modelo adequado para as séries de dados que analisamos neste capítulo. Os modelos selecionados estão descritos nas 8.7 e 8.8.

## 8.2 Diagnóstico do Modelo

Após a etapa de seleção do modelo precisamos verificar se algumas das suposições impostas na Definição dos modelos ARMA-GARCH foram violadas. De acordo com

Tabela 8.6: Ajuste do modelo ARMA-GARCH com distribuição condicional GEV para os dados da Microsoft.

Modelo	-LLH	AIC	AICc	BIC
ARMA(0,0)-GARCH(1,0)	-6672.113	-13338.23	-13338.22	-13320.71
ARMA(0,0)-GARCH(1,1)	-6809.361	<b>-13610.72</b>	<b>-13610.71</b>	<b>-13587.37</b>
ARMA(0,1)-GARCH(1,0)	-6716.853	-13425.71	-13425.69	-13402.35
ARMA(0,1)-GARCH(1,1)	<b>-6809.376</b>	-13608.75	-13608.73	-13579.56
ARMA(1,0)-GARCH(1,0)	-6719.587	-13431.17	-13431.16	-13407.82
ARMA(1,0)-GARCH(1,1)	-6809.332	-13608.66	-13608.64	-13579.48
ARMA(1,1)-GARCH(1,0)	-6720.221	-13430.44	-13430.42	-13401.25
ARMA(1,1)-GARCH(1,1)	-6809.374	-13606.75	-13606.71	-13571.72

Zivot (2008) após selecionarmos um modelo ARMA-GARCH a um determinado conjunto de dados, podemos avaliar se nosso modelo foi adequado através de análises gráficas e alguns testes estatísticos. Se o modelo ARMA-GARCH da Definição 3.1 foi especificado corretamente, então a série de resíduos estimados  $\{\hat{z}_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t}\}$  não deve apresentar correlação serial, heterocedasticidade condicional ou algum tipo de dependência não linear. Vale lembrar que a função de autocorrelação de uma série com volatilidade não constante apresenta a mesma característica de um ruído branco, apesar de apresentar correlações significativas quando analisamos seu valor absoluto. Sendo assim, a função de autocorrelação das séries  $\{z_t\}$ ,  $\{z_t^2\}$  e  $\{|z_t|\}$  devem ser semelhantes a de um ruído branco. A Figura 8.6 apresenta as funções de autocorrelação das séries de resíduos para os dados de retornos da empresa Microsoft. Para o modelo com inovações GEV a função de autocorrelação da série  $\{z_t^2\}$  apresentou valores dentro das faixas esperadas e a função de autocorrelação das séries  $\{z_t\}$  e  $\{|z_t|\}$  apresentaram somente dois valores fora do intervalo de confiança.

De forma similar a Figura 8.5 mostra evidências de que os resíduos do modelo com inovações estáveis para os dados da Microsoft representam uma série i.i.d. Na prática, os resultados dessa análise servem para identificar e diagnosticar se as premissas do modelo foram satisfeitas quando trabalhamos com distribuições estáveis. Porém, temos que ter ciência de que essa ferramenta não é adequada quando estamos

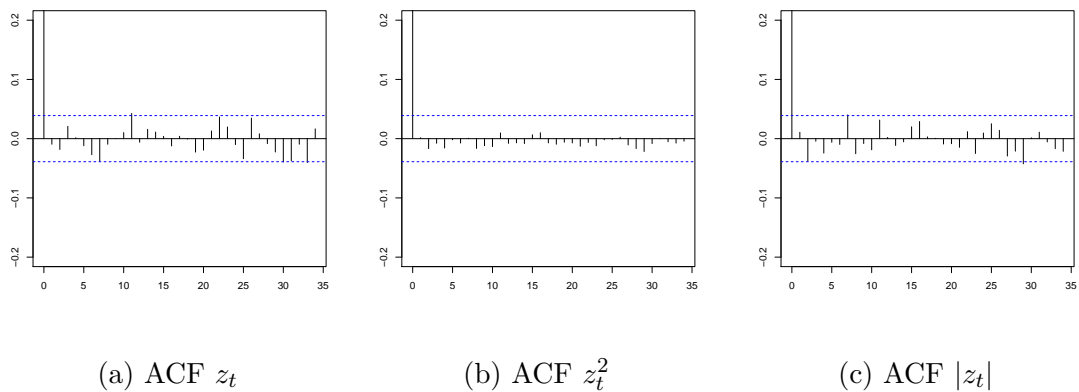


Figura 8.5: Autocorrelação para os resíduos ajustados do modelo ARMA(0,1)-GARCH(1,1) com resíduos estáveis para os dados da Microsoft.

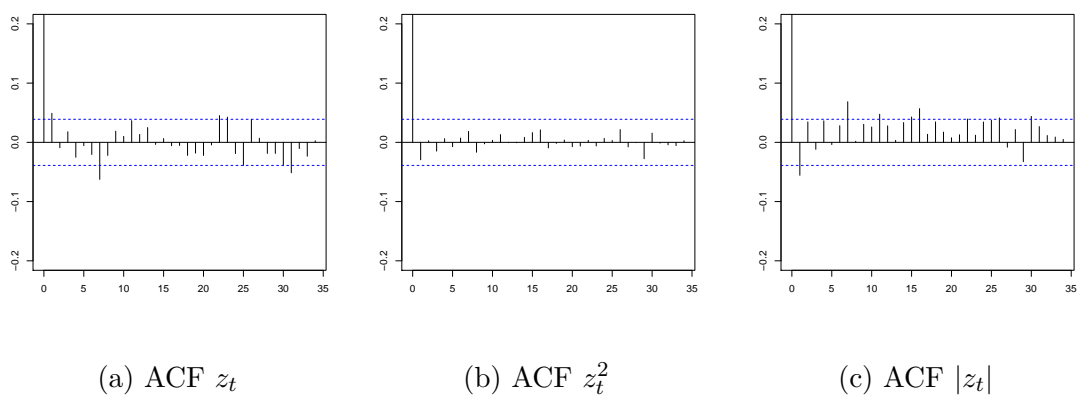


Figura 8.6: Autocorrelação para os resíduos ajustados do modelo ARMA(2,2)-GARCH(1,2) com resíduos GEV para os dados da Microsoft.

Tabela 8.7: Ajuste do modelo ARMA-GARCH com distribuição condicional estável.

Serie	Modelo	-LLH	AIC	AICc	BIC
Microsoft	ARMA(0,1)-GARCH(1,1)	-7625.92	-15237.84	-15237.8	-15196.97
	ARMA(0,1)-APARCH(2,2)	-7668.73	-15313.46	-15313.33	-15243.4
Google	ARMA(0,1)-GARCH(1,1)	-5922.53	-11831.07	-11831.01	-11791.30
	ARMA(0,1)-APARCH(1,2)	-5959.04	-11898.08	-11897.98	-11841.28
S&P500	ARMA(1,1)-GARCH(1,2)	-5860.91	-11703.81	-11703.67	-11657.83
	ARMA(1,1)-APARCH(2,2)	-5933.79	-11841.58	-11841.28	-11775.15

tratando com distribuições que possuem variância infinita e os intervalos de confiança nos gráficos de autocorrelação são obtidos assumindo que os dados são normalmente distribuídos. Na próxima seção analisamos a capacidade dos modelos GEV e estável em captar o comportamento da série de volatilidade em dados financeiros.

### 8.3 Comparação dos modelos GEV-GARCH e STABLE-GARCH

Zhao et al. (2011) utilizou o modelo GEV-GARCH ajustado para estimação dos quantis de 95% para a série temporal ajustada e o comparou ao modelo ajustado com a distribuição Normal. O autor constatou empiricamente que o modelo GEV-GARCH tem uma maior taxa de cobertura desses quantis do que o modelo usual com distribuição normal.

Nesta etapa do trabalho, iremos comparar de forma empírica, a taxa de cobertura dos quantis de 95% dos modelos AR(1)-GARCH(1,1) com distribuições GEV, Normal e estável. É importante salientar que Zhao et al. (2011) utilizou de uma série temporal financeira com dados coletados a cada 5 minutos ajustando os modelos Normal e GEV-GARCH aos máximos retornos diários.

Tabela 8.8: Ajuste do modelo ARMA-GARCH com distribuição condicional GEV.

Serie	Modelo	-LLH	AIC	AICc	BIC
Microsoft	ARMA(2,2)-GARCH(1,2)	-6826.20	-13632.34	-13632.25	-13573.96
	ARMA(2,2)-APARCH(1,2)	-7002.14	-13982.28	-13982.18	-13918.06
Google	ARMA(1,2)-GARCH(1,1)	-5378.98	-10741.97	-10741.9	-10696.52
	ARMA(1,0)-APARCH(1,2)	-5458.77	-10899.55	-10899.47	-10848.43
S&P500	ARMA(1,1)-GARCH(2,1)	-5862.58	-11711.16	-11711.07	-11675.39
	ARMA(1,1)-APARCH(2,2)	-5960.33	-11898.65	-11898.43	-11842.44

A série temporal utilizada aqui são os dados do índice SP500 descritos na tabela 8.1. Esta série é composta por dados coletados a cada 15 minutos e ajustaremos os modelos citados anteriormente aos retornos máximos diários.

Se observarmos a equação 3.4 do modelo GEV-GARCH, vemos que, dado a informação passada, podemos escrever as observações  $X_t$  na forma  $X_t \sim GEV(\mu_t, \xi, \sigma_t)$ . Esta mesma abordagem também pode ser utilizada para calcular os quantis para os modelos Normal e estável. Com isso utilizamos a distribuição de probabilidade destas funções para encontrar os quantis de 95% e compará-los com os quantis da série temporal. Seguindo as diretrizes do trabalho de Zhao et al. (2011) iremos calcular os quantis da série S&P500 através dos retornos das máximas observadas a cada 15 minutos.

Os gráficos a seguir mostram as comparações entre os modelos GEV, normal e estável. A Figura 8.7 nos permitem ver que os modelos GEV e estável tem melhor taxa de cobertura desses quantis do que o modelo Normal, que em geral subestima esses quantis. Analisando a Figura 8.8 podemos notar também que o modelo estável superestima os quantis se comparado ao modelo GEV, indicando que o modelo GEV tem melhor taxa de cobertura que o modelo estável.



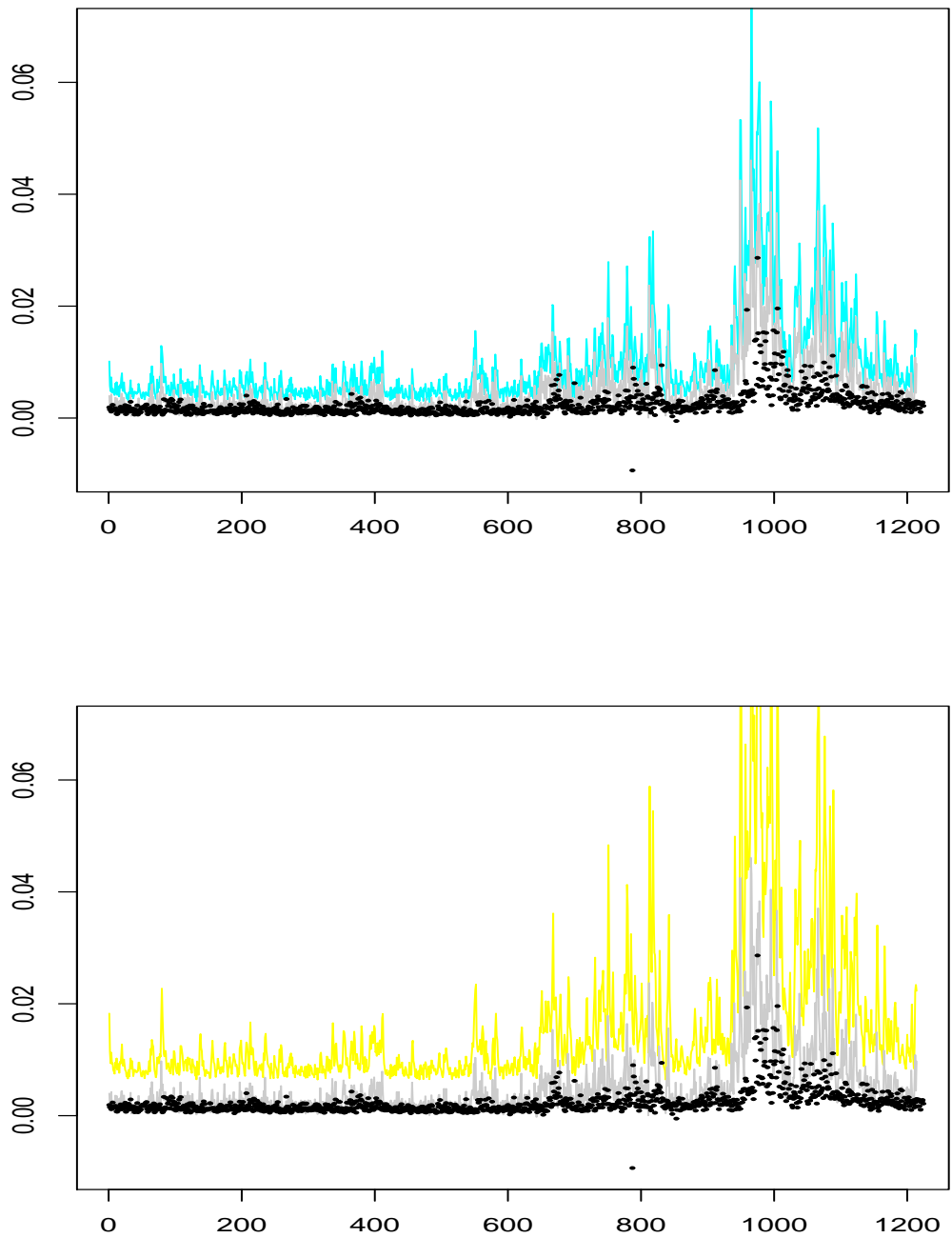


Figura 8.7: Quantis de 95% estimados dos retornos máximos da série S&P500. Os pontos sólidos representam os quantis reais. Gráfico acima: As linhas azul e cinza representam os quantis estimados através dos modelos GEV e normal. Gráfico abaixo: As linhas amarelo e cinza representam os quantis estimados através dos modelos estável e normal.

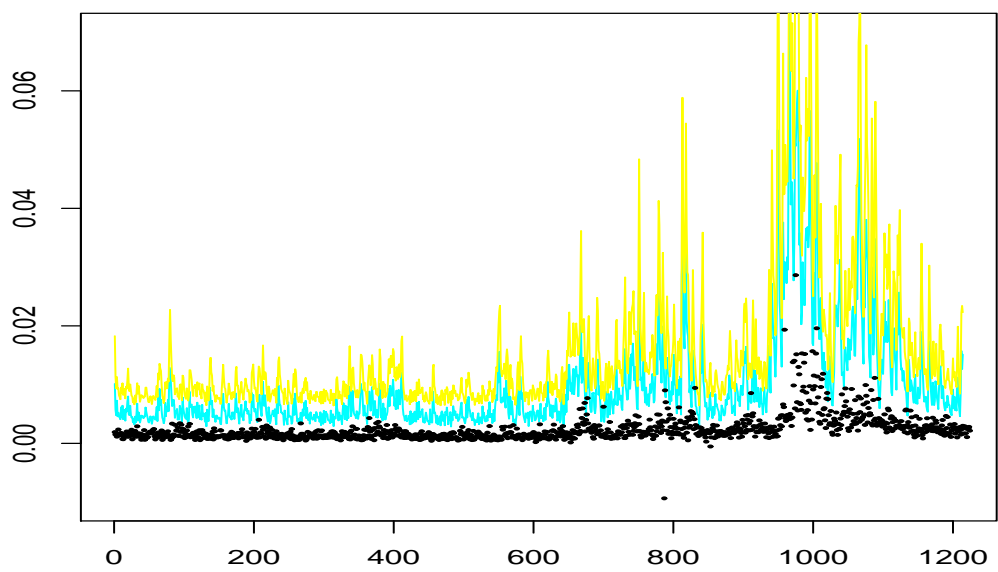


Figura 8.8: Quantis de 95% estimados dos retornos máximos da série S&P500. Os pontos sólidos representam os quantis reais e os gráficos azul e amarelo representam os quantis estimados através dos modelos GEV e estável, respectivamente.

## 8.4 Previsão de parâmetros futuros

Nesta seção iremos tratar da previsão da variância condicional do modelo (também chamada de volatilidade). Segundo Zivot E.(2008) a previsão de volatilidade tem várias aplicações práticas e relevantes na economia, a saber: gerenciamento de risco, mercado de opções, alocação de portfólio, estratégias de mercado e avaliação de modelos. Um dos principais objetivos de modelagem da volatilidade condicional é para gerar previsões acuradas para valores futuros da volatilidade condicional, bem como da média condicional. Vários trabalhos focam em avaliar a habilidade dos modelos GARCH para previsão. No trabalho de Anders Wilhelmsson (2006) o modelo GARCH(1,1) foi combinado com nove diferentes distribuições condicionais para avaliar seu desempenho na previsão de parâmetros futuros. Este trabalho é importante pois, ao invés de focar em diferentes especificações da equação de variância do modelo Wilhelmsson focou na importância de se escolher uma distribuição de inovações adequadas. O autor também argumenta que a habilidade que um modelo tem em realizar previsão deve ser utilizada como critério de seleção de modelos de volatilidade.

De acordo com Wurtz et al. (2006) a variância condicional pode ser predita de forma independente da média condicional. Dado o modelo GARCH(p,q), a previsão h-passos a frente da variância condicional  $\hat{\omega}_{t+h|t}^2$  pode ser calculada recursivamente através da seguinte equação:

$$\hat{\sigma}_{t+h|t}^2 = \hat{\omega} + \sum_{i=1}^q \hat{\alpha}_i \varepsilon_{t+h-i|t}^2 + \sum_{j=1}^p \hat{\beta}_j \sigma_{t+h-i|t}^2, \quad (8.1)$$

onde  $\varepsilon_{t+i|t}^2 = \sigma_{t+i|t}^2$  para  $i > 0$ ,  $\varepsilon_{t+i|t}^2 = \varepsilon_{t+i}^2$  e  $\sigma_{t+i|t}^2 = \sigma_{t+i}^2$  para  $i \leq 0$ . De forma semelhante em um processo APARCH(p,q) a previsão dependerá da distribuição das inovações. Neste caso, a previsão h-passos a frente da variância condicional  $\hat{\omega}_{t+h|t}^\delta$  pode ser calculada pela seguinte expressão:

$$\hat{\sigma}_{t+h|t}^\delta = \hat{\omega} + \sum_{i=1}^p \hat{\alpha}_i E[(|\varepsilon_{t+h-i}| - \hat{\gamma} \varepsilon_{t+h-i})^\delta | \Omega_t] + \sum_{j=1}^q \hat{\beta}_j \sigma_{t+h-j|t}^\delta, \quad (8.2)$$

onde  $E[(|\varepsilon_{t+h-i}| - \hat{\gamma} \varepsilon_{t+h-i})^\delta | \Omega_t] = \lambda_i \sigma_{t+k|t}^\delta$  para  $k > 1$  e  $\lambda_i = E(|z| - \gamma_i z)^\delta$ .

Várias medidas de error podem ser aplicadas para avaliar a habilidade que um certo modelo tem em realizar a previsão de parâmetros futuros. No trabalho de Alberga D. (2008) foram utilizadas as seguintes medidas:

- Erro quadrático médio (MSE):

$$MSE = \frac{1}{h} \sum_{t=S}^{S+h} (\hat{\sigma}_t^2 - \sigma_t^2)^2$$

- Erro absoluto médio (MAE):

$$MAE = \frac{1}{h} \sum_{t=S}^{S+h} |\hat{\sigma}_t^2 - \sigma_t^2|$$

- Erro percentual da média absoluta ajustado:

$$AMAPE = \frac{1}{h+1} \sum_{t=S}^{S+h} \left| \frac{\hat{\sigma}_t^2 - \sigma_t^2}{\hat{\sigma}_t^2 + \sigma_t^2} \right|,$$

onde  $S$  é o tempo que estamos avaliando a primeira previsão do modelo. Iremos utilizar as estatísticas MAE e AMAPE para avaliar o erro de previsão da volatilidade futura. Este procedimento também é utilizado pela rotina de previsão implementada originalmente na função `predict` do software R utilizando de um modelo ARMA-GARCH ajustado com a rotina `garchFit`. De forma similar ao capítulo anterior, utilizaremos os dados de taxa de câmbio descritos no capítulo 5 para testar a previsão com os dados estimados com nossa rotina `garchFitExtremes`. As próximas Tabelas mostram a previsão da série  $\sigma_t$  para os modelos estimados GARCH(1,1) e ARMA(1,1)-GARCH(1,1) com distribuição condicional normal.

Uma análise Tabelas 8.10 acima nos mostra que a previsão do processo  $\sigma_t$  segue a mesma dinâmica para as duas rotinas, porém com pequenas diferenças nos valores previstos. A maior diferença absoluta foi igual a 0.002612 e foi registrada para o a previsão do parâmetro  $\sigma_{10}$ . Na prática essa diferença aumenta a medida que o lag de previsão aumenta, e isso ocorreu devido a diferença nos valores dos parâmetros estimados pelas duas rotinas.

A Tabela 8.11 mostra a previsão da volatilidade para o modelo ARMA(1,1)-GARCH(1,1) estimado com distribuição condicional GEV e estável. A previsão da volatilidade utilizando o modelo estável ficou mais próxima da previsão realizada com o modelo normal como podemos observar na Tabela . Porém essa mesma estimativa foi bastante superior para o modelo GEV. Isso se deve ao fato de a estimação

Tabela 8.9: Previsão de parâmetros futuros para o modelo GARCH(1,1) com inovações normal.

t	$\hat{\sigma}_t$ (garchFit)	$\hat{\sigma}_t$ (garchFitExtremes)
1	0.3833961	0.3817035
2	0.3895421	0.3875761
3	0.3953471	0.3931414
4	0.4008358	0.3984204
5	0.4060303	0.4034323
6	0.4109507	0.4081945
7	0.4156151	0.4127226
8	0.4200402	0.4170311
9	0.4242410	0.4211332
10	0.4282312	0.4250411

do parâmetro  $\omega$  neste modelo ter sido alta. Os valores de  $\omega$  estimados para o modelo ARMA(1,1)-GARCH(1,1) com inovações normal, estável e GEV foram iguais a: 0.010911, 0.006183 e 0.044476, respectivamente.

Iremos mostrar agora nos próximos gráficos a previsão da volatilidade com dados simulados. Esse procedimento é importante pois nos permite avaliar o comportamento de nosso procedimento de previsão para modelos ARMA-GARCH de ordem arbitrária e com distribuição condicional GEV e estável. Procedemos então da seguinte forma:

- Escolhemos um determinado modelo para realizar a simulação dos dados, por exemplo ARMA(1, 1)-GARCH(2, 1) com distribuição condicional estável.
- Simulando uma série temporal de tamanho 2500 deste modelo.
- Utilizamos os primeiros 2040 dados desta série juntamente com a rotina `garchFitExtremes` para realizar a estimação dos parâmetros do modelo.
- Realizamos a previsão da volatilidade de acordo com a equação (8.2) descrita anteriormente.

O próximo Gráfico mostra o comportamento das séries  $\sigma_t$  (simulada) e  $\hat{\sigma}_t$  para o modelo ARMA(1,1)-GARCH(1,1) com distribuição condicional GEV descritos na

Tabela 8.10: Previsão de parâmetros futuros para o modelo ARMA(1,1)-GARCH(1,1) com inovações normal.

t	$\hat{\sigma}_t$ (garchFit)	$\hat{\sigma}_t$ (garchFitExtremes)
1	0.3850127	0.3836229
2	0.3914476	0.3898249
3	0.3975026	0.3956780
4	0.4032068	0.4012078
5	0.4085861	0.4064372
6	0.4136639	0.4113870
7	0.4184612	0.4160759
8	0.4229970	0.4205208
9	0.4272887	0.4247375
10	0.4313521	0.4287400

Tabela 9.18. Os resultados estão descritos de forma detalhada no apêndice D. Em ambos os gráficos podemos observar que a previsão da volatilidade tende a acompanhar a tendência do gráfico de volatilidade simulada. Nas duas simulações do modelo GEV-GARCH mostrados na Figura 8.9 a volatilidade prevista para o segundo passo seguiu em direção a volatilidade simulada. Já com relação aos gráfico da Figura 8.10 não podemos afirmar o mesmo. Porém podemos observar no Gráfico da direita desta mesma Figura que a volatilidade prevista acompanhou a tendência da volatilidade simulada e se aproximando bastante desta nos últimos passos. Nas simulações envolvendo o modelo GEV-GARCH descrita na Tabela 9.18 os valores das estatísticas MAE e AMAPE ficaram em torno de 0.006 e 0.004. Já para o modelo estável esses mesmos valores ficaram em torno de 0.03 e 0.017, que são valores aceitáveis de acordo com a literatura (veja por exemplo, Wilhelmsson, 2006). Os resultados obtidos a partir desse estudo de previsão nos permite ver que ambos os modelos GEV-GARCH e STABLE-GARCH apresentam potenciais de utilização para estimação e previsão com dados reais.

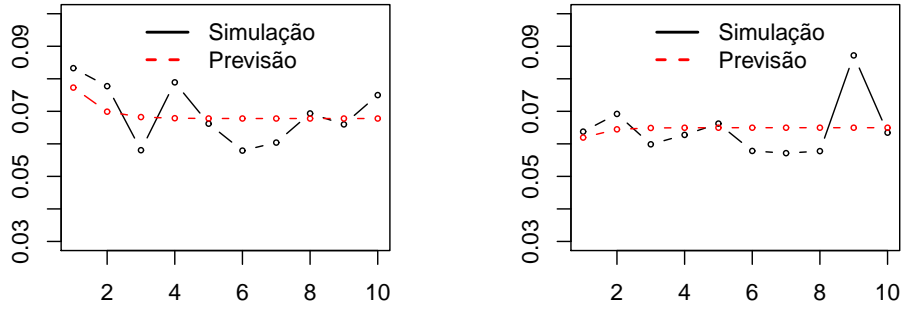


Figura 8.9: Simulação e previsão da volatilidade no modelo GEV-GARCH descrito na Tabela 9.18 do Apêndice.

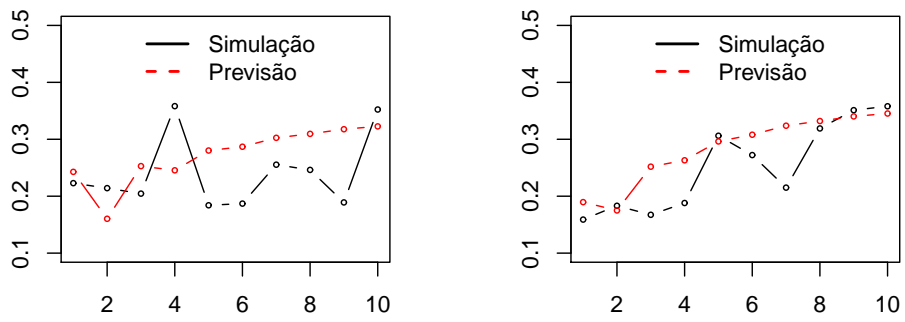


Figura 8.10: Simulação e previsão da volatilidade no modelo STABLE-GARCH descrito na Tabela 9.21 do Apêndice.

Tabela 8.11: Previsão de parâmetros futuros para o modelo ARMA(1,1)-GARCH(1,1) com inovações GEV e estável utilizando a rotina `garchFitExtremes`.

t	$\hat{\sigma}_t$ (GEV)	$\hat{\sigma}_t$ (estável)
1	0.4975373	0.3271097
2	0.5191184	0.3356404
3	0.5379843	0.3439190
4	0.5545761	0.3519634
5	0.5692380	0.3597896
6	0.5822449	0.3674116
7	0.5938206	0.3748422
8	0.6041500	0.3820926
9	0.6133882	0.3891731
10	0.6216663	0.3960931



# Capítulo 9

## Conclusões e Trabalhos Futuros

### 9.1 Conclusões

Neste trabalho exploramos os aspectos teóricos e computacionais envolvidos na estimação dos modelos ARMA-GARCH com distribuição condicionais GEV e estável. Os resultados do Capítulo 5 deixaram claro que não precisamos impor nenhuma ordem de atualização dos parâmetros do modelo na estimação MCMC, pois os resultados obtidos aqui foram semelhantes aos obtidos por Zhao et al. (2011). Ainda com relação ao trabalho de Zhao et al. (2011) pudemos constatar também que ambos os métodos de estimação (MLE ou bayesiano) tiveram a mesma performance para estimação do modelo. Em relação ao modelo STABLE-GARCH o Capítulo 6 mostrou que a estimação do modelo AR(1)-GARCH(1,1) também pode ser realizada utilizando-se uma distribuição condicional estável.

O Capítulo 7 explorou os aspectos teóricos envolvidos na demonstração da estacionariedade forte dos modelos ARMA-power-GARCH com distribuição condicional estável explorados em Rachev et al. (2002). Posteriormente, indicamos como a demonstração dada por Rachev et al. (2002) pode ser generalizada para derivar as condições de estacionariedade forte dos modelos ARMA-APARCH com distribuição condicional estável. Mostramos também como calcular a constante  $\lambda_i$  crucial para o cálculo da condição de estacionariedade do modelo ARMA-APARCH.

Após a constatação de que o modelo estável também pode ser generalizado para um modelo APARCH, implementamos uma rotina geral de estimação dos modelos

ARMA-GARCH/APARCH com distribuições condicionais Normal, t-student, skew t-student, GEV e estável. Esta rotina apresenta vantagem em relação à conhecida rotina `GarchFit` implementada por Wurtz et al. (2006), pois esta última não contempla de modelos com distribuição condicional GEV ou estável.

Como até o momento não conhecemos nenhuma rotina de estimação de modelos ARMA-GARCH/APARCH com distribuição condicional GEV ou estável, criamos um pacote no software R chamado `garchFitExtremes` com o objetivo de disponibilizar nossas rotinas para utilização pela comunidade acadêmica.

Por fim, no Capítulo 8 aplicamos essas rotinas para estimação de modelos ARMA-GARCH e ARMA-APARCH com distribuição condicional estável e GEV. Indicamos alguns testes que são utilizados pela para o ajuste de dados reais e utilizamos o critério de seleção de modelos de Akaike para escolher a ordem do modelo para ajuste desses dados. Por último indicamos como realizar a previsão da volatilidade utilizando os modelos ARMA-GARCH com distribuição condicional GEV e estável.

## 9.2 Trabalhos Futuros

Indicamos abaixo alguns trabalhos que não foram explorados neste trabalho e podem ser desenvolvidos:

- Estudo sobre critérios de seleção de modelos envolvendo distribuições estáveis.
- Cálculo analítico da expressão  $E\{|z| - \gamma z\}^\delta$ , onde  $-1 < \gamma < 1$ ,  $0 < \delta < \alpha$  e  $z \sim S(\alpha, \beta, 1, 0)$ .
- Previsão de Parâmetros futuros para a série  $X_t$  seguindo um processo ARMA-GARCH/APARCH com distribuição condicional estável e GEV.
- Estudo sobre o diagnóstico de modelos com inovações estáveis.

# Referências Bibliográficas

- [1] Alberga, D. Shalita, H. e Yosefb, R. Estimating stock market volatility using asymmetric GARCH models. *Applied Financial Economics*, 18, 1201-1208, 2008
- [2] Billingsley, P. *Probability and Measure*, Second Edition, New York: John Wiley & Sons, 1986.
- [3] Bougerol, P., Picard, N. Stationarity of GARCH processes and of some non-negative time series. *Journal of Econometrics* 52, 115-127, 1992.
- [4] Brockwell, P. J., e Davis, R. A. *Time Series: Theory and Methods* 2nd ed., Springer, New York, 1992.
- [5] Brockwell, P. J, e Davis, R. A. *Introduction to Time Series and Forecasting*. Springer, New York, 1996.
- [6] Brummelhuis, R., e Kaufmann, R. Time Scaling for GARCH(1,1) and AR(1)-GARCH(1,1) Processes. *Journal of Risk* , 39-94. 2007.
- [7] Casella, G. e Berger, R. L. *Statistical Inference*. China Machine Press. 2<sup>a</sup> Ed. 2009.
- [8] Ding, Z., Granger C.W.J., Engle R.F. A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance* 1, 83-106. 1993.
- [9] Feller, W. *An introduction to Probability Theory and its Applications*. Vol. II. John Wiley. New York, 1971.
- [10] Fernandez, C. and Steel, M. F. J. On Bayesian modeling of fat tails and skewness. *J. Amer. Statist. Assoc.* 93:359-371, 1998.

- [11] Zhou, J. Modelling S&P 500 stock index using ARMA-ASYMMETRIC power ARCH models. Master Thesis in statistics. School of economics and social science. Hogskolan Dalarna, Sweden, 2009.
- [12] Frain C. J. University of Dublin. Studies on the Application of the  $\alpha$ -stable Distribution in Economics, 2009.
- [13] Gelman, Andrew. John B. Carlin, Hal S. Stern e Donald B Rubin. Bayesian data Analysis. Chapman and Hall, 2003.
- [14] Górska K. e Penson K. A. Lévy stable two-sided distributions: Exact and explicit densities for asymmetric case. Physical Review, E 83, 2003.
- [15] Heidelberger P e Welch PD. Simulation run length control in the presence of an initial transient. Opns Res., 31, 1109-44, 1983.
- [16] James, B. R. Probabilidade: Um curso em nível intermediário. Impa, Rio de Janeiro, 2010.
- [17] Jondeau, E., Poon, S.-H., Rockinger, M. Financial Modeling Under Non-Gaussian Distributions. Springer-Verlag, Springer, 2007.
- [18] Kesten, H. and Spitzer, F., Convergence in distribution of products of random matrices, Z. Wahrsch. Verw. Gebiete 67 no. 4, 363-386, 1984.
- [19] Lindsey, J. K. Statistical Analysis of Stochastic Processes in Time. Cambridge, 2004.
- [20] Mandelbrot B. The variation of certain speculative prices. J Bus 36,394-419, 1963.
- [21] Mandelbrot B. New methods in statistical economics. J Polit Econ 61,421-40, 1963.
- [22] Morettin P., A. e Tolo C.M.C. Análise de Séries Temporais. Editora Edgard Blucher, 2004.
- [23] Nakatsuma, T. Bayesian analysis of ARMA-GARCH models: A Markov chain sampling approach. Journal of Econometrics 95, 57-69. 2000.

- [24] Nolan, J. P. Numerical calculations of stable densities and distribution functions. *Communications in Statistics - Stochastic Models* 13: 759-774, 1997.
- [25] Nolan, J. P. Fitting data and assessing goodness-of-fit with stable distributions. Unpublished Manuscript. Washington, DC: American University, 1999.
- [26] Nolan, J. P. Maximum likelihood estimation and diagnostics for stable distribution. In O. E. Barndorff-Nielsen, T. Mikosch, e S. I. Resnick (Eds.), *Levy Processes: Theory and Applications*: 379-400. Boston, Birkhauser, 2000.
- [27] Mittnik S. , Panosrka A. K. , Rachev S. T. Stable GARCH Models for Financial Time Series. *Applied Mathematical Letters* Vol. 8, No. 5, 33-37, 1995.
- [28] Mittnik S. , Paoletta M.S. , Rachev, S.T. Stationarity of stable power-GARCH processes. *Journal of Econometrics* 106, 97-107. 2002.
- [29] Omelchenko V. Parameter Estimation of the Stable GARCH(1,1)-Model. Charles University, WDS'09 Proceedings of Contributed Papers, Part I, 137-142, 2009.
- [30] Oral, E. An empirical analysis of trading volume and return volatility relationship on Istanbul stock exchange national -100 Index. *Journal of Applied Finance & Banking*, vol.2, no.5, 149-158, 2012.
- [31] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>, 2013.
- [32] Rathie, P.N., Dorea, C. C. Y e Matsushita, R. Lévy Distribution, H Function and Applications to Currency Data. *Proceedings of the Seventh International Conference of the Society for Special Functions and Their Applications (Pune Univ., India, Fevereiro 32-23).7*, 17-26. 2006.
- [33] Rathie, P.N., Coutinho M., Sousa T.R. , Rodrigues G.S., Carrijo T.B. Stable and generalized-t distributions and applications. *Communications in Nonlinear Science and Numerical Simulations* 17, 5088-5096, 2012.
- [34] Samorodnitsky, G., Murad, S.T. *Stable non-Gaussian Random Processes*. Chapman & Hall/CRC, 2000.

- [35] Schneider, W. R. Stable Distributions: Fox Function Representation and Generalization. *Stochastic Process in Classical and Quantum*. Baden, 497-511. 1986.
- [36] Szymon Borak et. al. Stable distributions. *Economic Risk*. Berlin, 2005.
- [37] Wei, W.W.S. Time Series analysis: Univariate and Multivariate methods. Second Edition, 2006.
- [38] Wilhelmsson, A. Garch Forecasting Performance under Different Distribution Assumptions. *J. Forecast.* 25, 561-578, 2006.
- [39] Wurtz, D., Y. Chalabi, e L. Luksan. Parameter estimation of ARMA models with GARCH/APARCH errors: An R and SPlus software implementation. *Journal of Statistical Software*, 2006.
- [40] F. Yilmaz e M.-S. Alouini. "Product of the powers of generalized Nakagami-m variates and performance of cascaded fading channels", in *Proceedings of IEEE Global Telecommunications Conference, (GLOBECOM 2009)*, Honolulu, Hawaii, US, 1-8. 2009.
- [41] Zhao X. ,Scarrott C. J. , Oxley L. , Reale M. GARCH dependence in extreme value models with Bayesian inference. *Mathematics an Computers in Simulation*, Vol. 81, Issue 7, 1430-1440. 2011.
- [42] Zivot, E. Practical Issues in the Analysis of Univariate GARCH Models. *Handbook of Financial Time Series*, pp 113-155, 2009.

# Estimação GEV-GARCH

Tabela 9.1: Resultados da estimação dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações GEV utilizando MLE.

		Simulado	Estimado	Erro Padrão	Estat. t	$\Pr(t >  t )$
	$\omega$	0.01	0.012071	0.000799	15.09180	< 2.22e-16
	$\alpha_1$	0.45	0.412298	0.029237	14.10170	< 2.22e-16
$\theta_1$	$\beta_1$	0.08	0.085874	0.019919	4.311001	1.6248e-05
	$\mu$	0.21	0.204151	0.006651	30.69430	< 2.22e-16
	$a_1$	0.32	0.336089	0.014828	22.66470	< 2.22e-16
	$\xi$	0.08	0.062785	0.015142	4.144700	3.4020e-05
	$\omega$	0.01	0.010971	0.000827	13.26230	< 2.22e-16
	$\alpha_1$	0.80	0.814675	0.037655	21.63520	< 2.22e-16
$\theta_2$	$\beta_1$	0.02	0.014679	0.004431	3.312700	0.000923
	$\mu$	0.01	0.011888	0.003462	3.433600	0.000595
	$a_1$	0.10	0.090408	0.010839	8.340700	< 2.22e-16
	$\xi$	0.20	0.218523	0.016580	13.17950	< 2.22e-16
	$\omega$	0.05	0.051413	0.003511	14.642500	< 2.22e-16
	$\alpha_1$	0.50	0.465839	0.027682	16.827008	< 2.22e-16
$\theta_3$	$\beta_1$	0.10	0.097854	0.011734	8.3391000	< 2.22e-16
	$\mu$	0.05	0.050882	0.007869	6.4659000	1.0072e-10
	$a_1$	0.20	0.206386	0.012264	16.828000	< 2.22e-16
	$\xi$	0.30	0.288827	0.016103	17.935800	< 2.22e-16

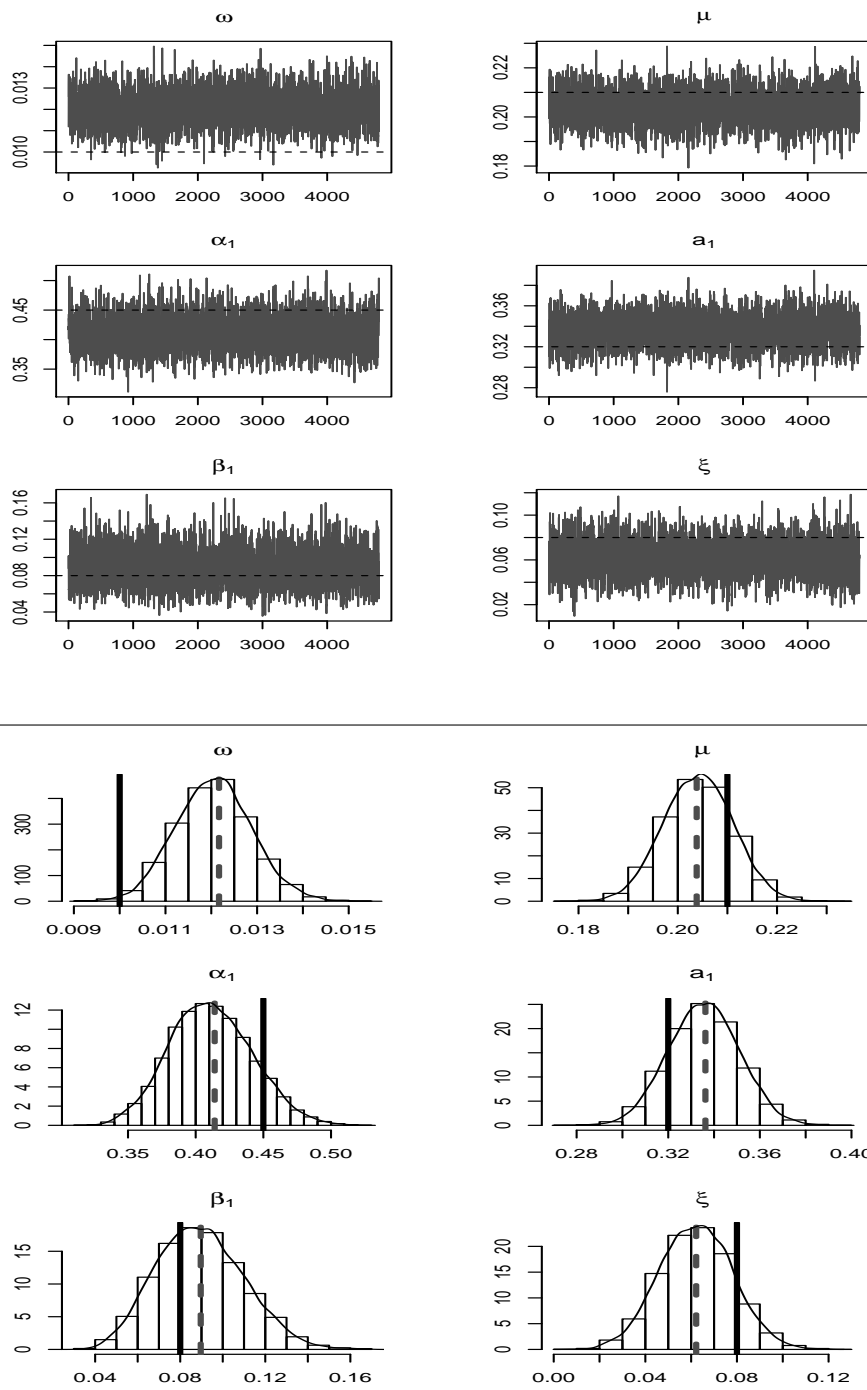


Figura 9.1: GEV-GARCH com  $\theta = (0.01, 0.45, 0.08, 0.21, 0.32, 0.08)$ . Gráfico superior: Cadeia de Markov simulada convergindo para distribuição a posteriori. A linha pontilhada indica o verdadeiro valor do parâmetro. Gráfico inferior: Distribuições a posteriori marginais. A barra horizontal contínua indica o verdadeiro valor do parâmetro e a barra pontilhada o valor estimado.



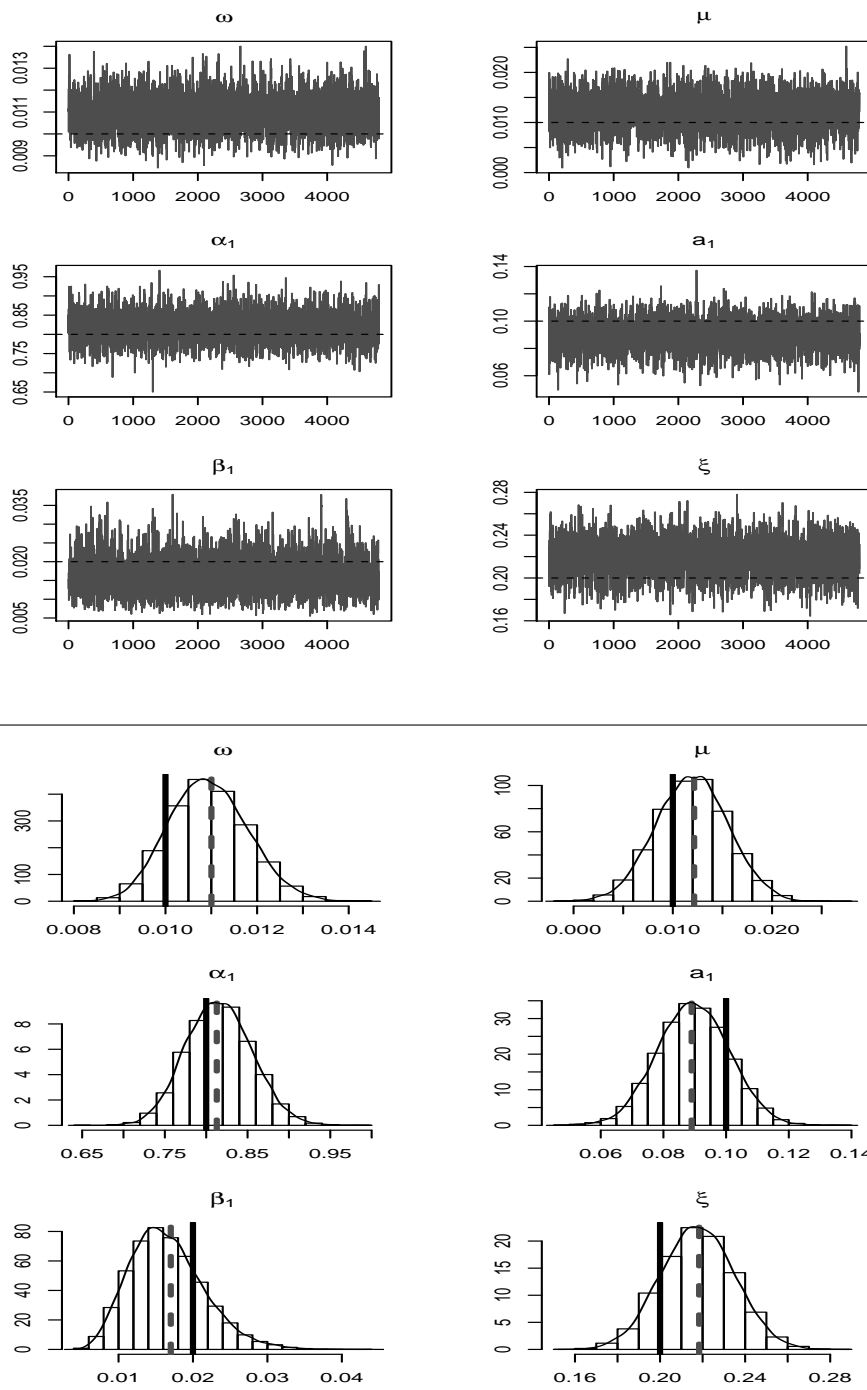


Figura 9.2: GEV-GARCH com  $\theta = (0.01, 0.8, 0.02, 0.01, 0.1, 0.2)$ . Gráfico superior: Cadeia de Markov simulada convergindo para distribuição a posteriori. A linha pontilhada indica o verdadeiro valor do parâmetro. Gráfico inferior: Distribuições a posteriori marginais. A barra horizontal contínua indica o verdadeiro valor do parâmetro e a barra pontilhada o valor estimado.

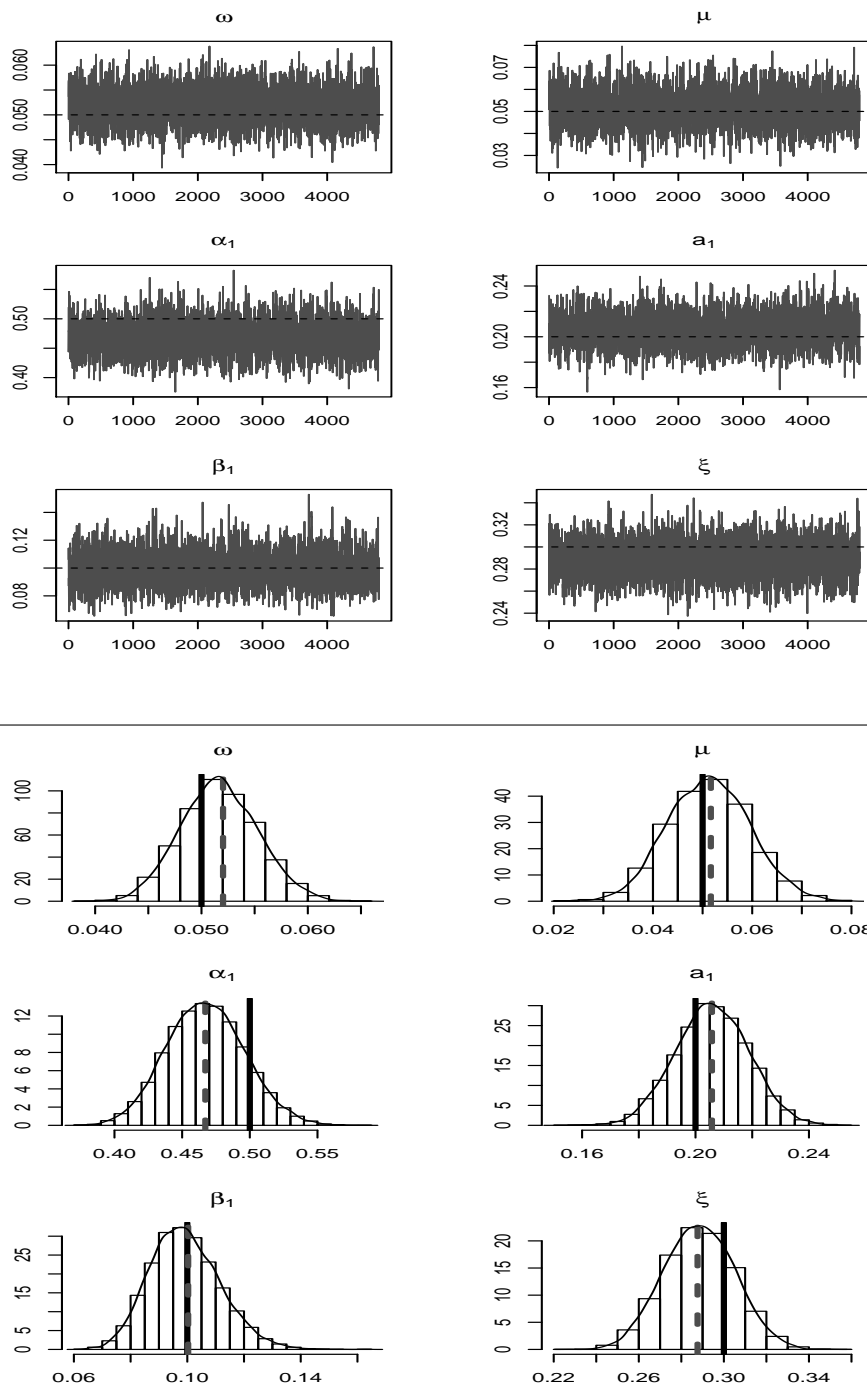


Figura 9.3: GEV-GARCH com  $\theta = (0.05, 0.5, 0.1, 0.05, 0.2, 0.3)$ . Gráfico superior: Cadeia de Markov simulada convergindo para distribuição a posteriori. A linha pontilhada indica o verdadeiro valor do parâmetro. Gráfico inferior: Distribuições a posteriori marginais. A barra horizontal contínua indica o verdadeiro valor do parâmetro e a barra pontilhada o valor estimado.

---

**Algorithm 9.1** Metropolis-Hastings with Gibbs sample - GEV-GARCH

---

**Input:**  $(x_1, x_2, \dots, x_n)$ ,  $\tau = (\tau_1, \dots, \tau_6)$  e  $\eta$  (Tamanho da Cadeia).

**Output:**  $\theta^{(1)}, \dots, \theta^{(i)}$ , com  $\theta^{(i)} = (\alpha_0^{(i)}, \alpha_1^{(i)}, \alpha_2^{(i)}, \beta_0^{(i)}, \beta_1^{(i)}, \xi^{(i)}) = (\theta_1^{(i)}, \dots, \theta_6^{(i)})$

1: Defina  $\theta^0 = (0.1, 0.1, 0.1, 0.1, 0.1, 0.1)$

2: **for**  $i = 1, 2, \dots, \eta$  **do**

3: Simule  $\alpha_0^*$  de  $N(\alpha_0^{(i-1)}, \tau_1)$

4: **if**  $\alpha_0^* \leq 0$  **then**

5:  $\alpha_0^{(i)} = \alpha_0^{(i-1)}$

6: **else**

7: Calcule a razão das densidades,

$$r = \frac{p(x|\alpha_0^*, \theta_{-1}^{(i-1)})/N(\alpha_0^*|\alpha_0^{(i-1)})}{p(x|\theta_{(+1)}^{(i-1)})/N(\alpha_0^{(i-1)}|\alpha_0^*)}$$

8: Aceite  $\alpha_0^*$  com probabilidade  $\min(1, r)$ .

9: **end if**

10: Simule  $\alpha_1^*$  de  $N(\alpha_1^{(i-1)}, \tau_2)$

11: **if**  $\alpha_1^* \leq 0 \parallel \alpha_1^* + \alpha_2^{(i-1)} \geq 1$  **then**

12:  $\alpha_1^{(i)} = \alpha_1^{(i-1)}$

13: **else**

14: Calcule a razão das densidades,

$$r = \frac{p(x|\alpha_1^*, \theta_{-2}^{(i-1)})/N(\alpha_1^*|\alpha_1^{(i-1)})}{p(x|\theta_{(+2)}^{(i-1)})/N(\alpha_1^{(i-1)}|\alpha_1^*)}$$

15: Aceite  $\alpha_1^*$  com probabilidade  $\min(1, r)$ .

16: **end if**

17: Simule  $\alpha_2^*$  de  $N(\alpha_2^{(i-1)}, \tau_3)$

18: **if**  $\alpha_2^* \leq 0 \parallel \alpha_1^{(i)} + \alpha_2^* \geq 1$  **then**

19:  $\alpha_2^{(i)} = \alpha_2^{(i-1)}$

20: **else**

21: Calcule a razão das densidades,

$$r = \frac{p(x|\alpha_2^*, \theta_{-3}^{(i-1)})/N(\alpha_2^*|\alpha_2^{(i-1)})}{p(x|\theta_{(+3)}^{(i-1)})/N(\alpha_2^{(i-1)}|\alpha_2^*)}$$

22: Aceite  $\alpha_2^*$  com probabilidade  $\min(1, r)$ .

23: **end if**

---

24: Simule  $\beta_0^*$  de  $N(\beta_0^{(i-1)}, \tau_4)$

25: Calcule a razão das densidades,

$$r = \frac{p(x|\beta_0^*, \theta_{-4}^{(i-1)})/N(\beta_4^*|\beta_4^{(i-1)})}{p(x|\theta_{(+4)}^{(i-1)})/N(\beta_0^{(i-1)}|\beta_0^*)}$$

26: Aceite  $\beta_0^*$  com probabilidade  $\min(1, r)$ .

27: **if**  $|\beta_1|^* \leq 1$  **then**

28:      $\beta_1^{(i)} = \beta_1^{(i-1)}$

29: **else**

30:     Simule  $\beta_1^*$  de  $N(\beta_1^{i-1}, \tau_5)$

31:     Calcule a razão das densidades,

$$r = \frac{p(x|\beta_1^*, \theta_{-5}^{(i-1)})/N(\beta_1^*|\beta_1^{(i-1)})}{p(x|\theta_{(+5)}^{(i-1)})/N(\beta_1^{(i-1)}|\beta_1^*)}$$

32:     Aceite  $\beta_1^*$  com probabilidade  $\min(1, r)$ .

33: **end if**

34: Simule  $\xi^*$  de  $N(\xi^{i-1}, \tau_6)$

35: Calcule a razão das densidades,

$$r = \frac{p(x|\xi^*, \theta_{-6}^{(i-1)})/N(\xi^*|\xi^{(i-1)})}{p(x|\theta_{(+6)}^{(i-1)})/N(\xi^{i-1}|\xi^*)}$$

36: Aceite  $\xi^*$  com probabilidade  $\min(1, r)$ .

37: **end for**

---

Tabela 9.2: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo AR(1)-GARCH(1,1) com inovações GEV utilizando MLE.

		Simulado	Estimado	RMSE
$\theta_1$	$\omega$	0.01	0.014565	0.037473
	$\alpha_1$	0.45	0.450818	0.059453
	$\beta_1$	0.08	0.079321	0.018056
	$\mu$	0.21	0.220250	0.078336
	$a_1$	0.32	0.332897	0.089601
	$\xi$	0.08	0.079871	0.018498
$\theta_2$	$\omega$	0.01	0.010261	0.001597
	$\alpha_1$	0.80	0.794582	0.041665
	$\beta_1$	0.02	0.020509	0.005398
	$\mu$	0.01	0.010734	0.004125
	$a_1$	0.10	0.095243	0.019899
	$\xi$	0.20	0.198094	0.023523
$\theta_3$	$\omega$	0.05	0.051072	0.006349
	$\alpha_1$	0.50	0.499125	0.029906
	$\beta_1$	0.10	0.100162	0.012253
	$\mu$	0.05	0.052178	0.011911
	$a_1$	0.20	0.195298	0.029277
	$\xi$	0.30	0.297296	0.021280

Tabela 9.3: Resultados da estimação dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações GEV utilizando a metodologia Bayesiana.

		Simulado	Estimado	Erro Padrão	Intervalo de Confiança	
					Inferior	Superior
$\theta_1$	$\omega$	0.01	0.012058	0.000788	0.010509	0.013610
	$\alpha_1$	0.45	0.411326	0.031702	0.354725	0.473343
	$\beta_1$	0.08	0.089060	0.019610	0.052562	0.128447
	$\mu$	0.21	0.204069	0.006570	0.191427	0.217033
	$a_1$	0.32	0.335640	0.014706	0.306561	0.363524
	$\xi$	0.08	0.062392	0.015257	0.032315	0.091354
$\theta_2$	$\omega$	0.01	0.010921	0.000811	0.009348	0.012544
	$\alpha_1$	0.80	0.815530	0.037595	0.742698	0.888486
	$\beta_1$	0.02	0.016371	0.005376	0.007922	0.026120
	$\mu$	0.01	0.011938	0.003417	0.005190	0.018601
	$a_1$	0.10	0.089605	0.011395	0.067426	0.110591
	$\xi$	0.20	0.217883	0.016305	0.187252	0.250871
$\theta_3$	$\omega$	0.05	0.051671	0.003450	0.04522	0.058800
	$\alpha_1$	0.50	0.466934	0.028449	0.41363	0.520602
	$\beta_1$	0.10	0.099000	0.011812	0.07743	0.123192
	$\mu$	0.05	0.051345	0.007900	0.03623	0.067078
	$a_1$	0.20	0.206119	0.012521	0.18078	0.229827
	$\xi$	0.30	0.288456	0.016675	0.25636	0.318900

Tabela 9.4: Teste de estacionariedade para as cadeias simuladas da distribuição a posteriori dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações GEV.

		Teste Estac.	Manter	Descartar	C-von-M	Teste HWD	Estatística HWD
	$\omega$	Passou	40000	10000	0.050694	Passou	2.74e-05
	$\alpha_1$	Passou	45000	5000	0.228186	Passou	0.001463
$\theta_1$	$\beta_1$	Passou	50000	0	0.080811	Passou	0.000609
	$\mu$	Passou	45000	5000	0.100623	Passou	0.000282
	$a_1$	Passou	45000	5000	0.102242	Passou	0.000624
	$\xi$	Passou	45000	5000	0.278419	Passou	0.000666
	$\omega$	Passou	45000	5000	0.065714	Passou	3.23e-05
	$\alpha_1$	Passou	45000	5000	0.018372	Passou	0.001467
$\theta_2$	$\beta_1$	Passou	50000	0	0.262126	Passou	0.000245
	$\mu$	Passou	45000	5000	0.217811	Passou	0.000120
	$a_1$	Passou	50000	0	0.036555	Passou	0.000425
	$\xi$	Passou	45000	5000	0.269272	Passou	0.000682
	$\omega$	Passou	45000	5000	0.144447	Passou	0.000135
	$\alpha_1$	Passou	50000	0	0.359226	Passou	0.001003
$\theta_3$	$\beta_1$	Passou	50000	0	0.053450	Passou	0.000410
	$\mu$	Passou	50000	0	0.347949	Passou	0.000345
	$a_1$	Passou	35000	15000	0.293203	Passou	0.000587
	$\xi$	Passou	50000	0	0.316728	Passou	0.000636

A coluna nomeada como *Teste Estac.* indica se a cadeia passou no teste de estacionariedade e quantas observações devemos utilizar após descartar o burn-in da cadeia. Como todas as cadeias passaram no teste o Half-Width Test foi também realizado e reportou o valor da estatística de Cramer-Von-Misses (C-Von-M).

Tabela 9.5: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo AR(1)-GARCH(1,1) com inovações GEV utilizando a metodologia Bayesiana.

		Simulado	Estimado	RMSE
$\theta_1$	$\omega$	0.01	0.010066	0.000701
	$\alpha_1$	0.45	0.445269	0.030713
	$\beta_1$	0.08	0.084914	0.017127
	$\mu$	0.21	0.209839	0.006122
	$a_1$	0.32	0.320329	0.013813
	$\xi$	0.08	0.081637	0.014740
$\theta_2$	$\omega$	0.01	0.010056	0.000711
	$\alpha_1$	0.80	0.797866	0.042331
	$\beta_1$	0.02	0.022804	0.009676
	$\mu$	0.10	0.010429	0.003046
	$a_1$	0.10	0.097545	0.010214
	$\xi$	0.20	0.200315	0.015574
$\theta_3$	$\omega$	0.05	0.050052	0.004260
	$\alpha_1$	0.50	0.505826	0.025704
	$\beta_1$	0.10	0.100849	0.011281
	$\mu$	0.05	0.050884	0.008593
	$a_1$	0.20	0.199791	0.012996
	$\xi$	0.30	0.300435	0.018021



# Estimação STABLE-GARCH

Tabela 9.6: Resultados da estimação dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações estáveis utilizando MLE.

		Simulado	Estimado	Erro padrão	Estat. t	$Pr(t >  t )$
	$\omega$	0.01	0.009572	0.000634	15.0888	< 2.22e-16
	$\alpha_1$	0.45	0.445496	0.021848	20.3900	< 2.22e-16
$\theta_1$	$\beta_1$	0.08	0.089507	0.031205	2.86830	4.1268e-03
	$\mu$	0.21	0.205235	0.000347	590.2875	< 2.22e-16
	$a_1$	0.32	0.334252	0.000851	392.3186	< 2.22e-16
	$\alpha$	1.80	1.795358	0.023270	77.15280	< 2.22e-16
	$\beta$	0.50	0.645861	0.086850	7.436400	1.0347e-13
	$\omega$	0.01	0.010090	0.000508	19.7864	< 2.22e-16
	$\alpha_1$	0.80	0.806294	0.020814	38.7371	< 2.22e-16
$\theta_2$	$\beta_1$	0.02	0.022670	0.007240	3.1312	0.0017408
	$\mu$	0.01	0.009574	0.000464	20.6191	< 2.22e-16
	$a_1$	0.10	0.102197	0.006062	16.8561	< 2.22e-16
	$\alpha$	1.90	1.908223	0.018044	105.7497	< 2.22e-16
	$\beta$	0.50	0.621496	0.157865	3.9369	0.0000855
	$\omega$	0.05	0.048270	0.003167	15.2394	< 2.22e-16
	$\alpha_1$	0.50	0.486220	0.021717	22.3917	< 2.22e-16
$\theta_3$	$\beta_1$	0.10	0.124405	0.026128	4.76280	1.9094e-06
	$\mu$	0.05	0.045135	0.000589	76.5683	< 2.22e-16
	$a_1$	0.20	0.222621	0.003499	63.6165	< 2.22e-16
	$\alpha$	1.80	1.806035	0.029266	61.7090	< 2.22e-16
	$\beta$	0.50	0.535050	0.096967	5.51780	3.4325e-08

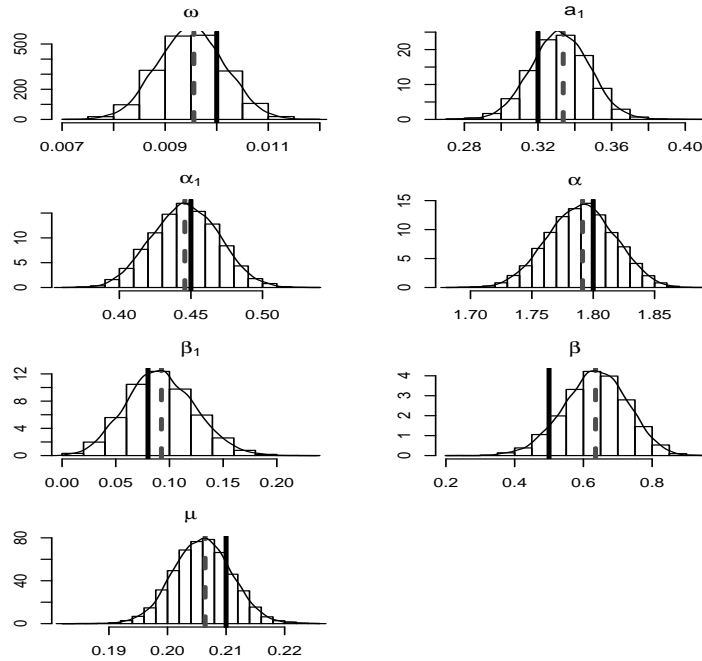
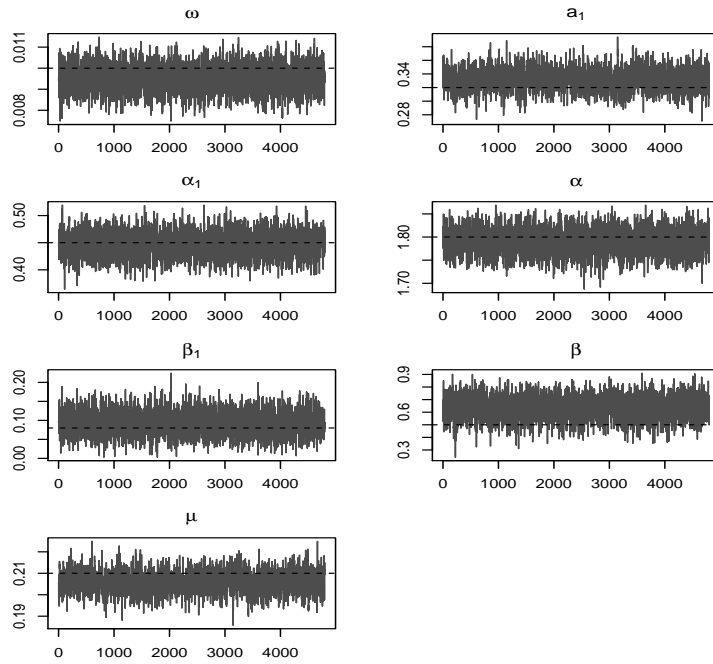


Figura 9.4: STABLE-GARCH com  $\varphi_1 = (0.01, 0.45, 0.08, 0.21, 0.32, 1.8, 0.5)$ . Gráficos superiores: Cadeia de Markov simulada convergindo para a distribuição a posteriori. A linha pontilhada indica o verdadeiro valor do parâmetro. Gráficos inferiores: Distribuições a posteriori marginais. A barra horizontal contínua indica o verdadeiro valor do parâmetro e a barra pontilhada o valor estimado.

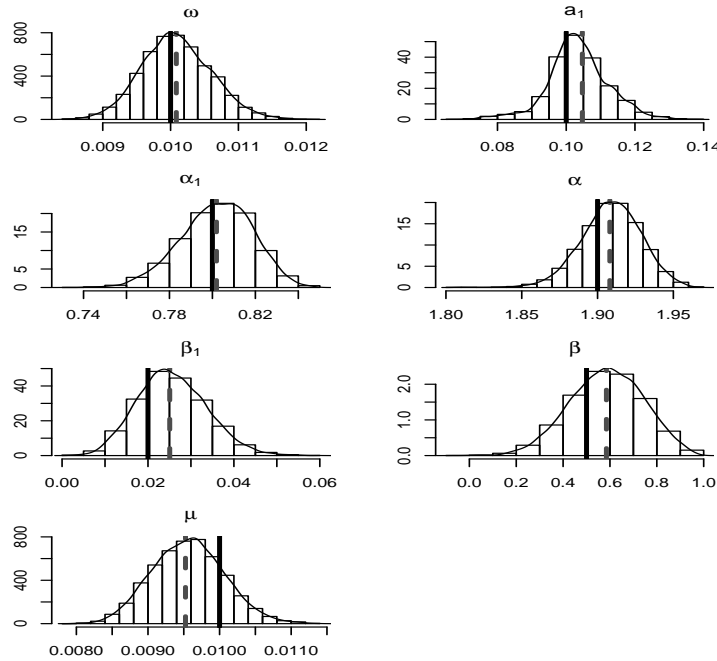
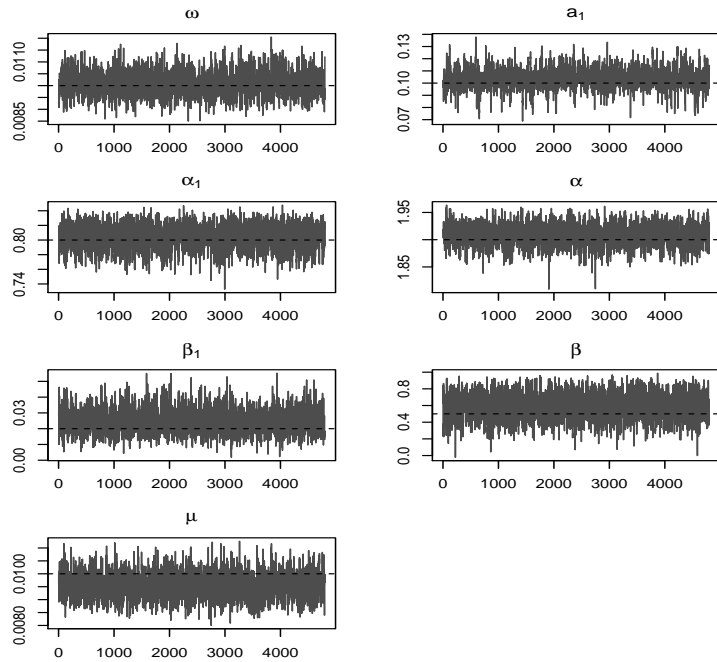


Figura 9.5: STABLE-GARCH com  $\varphi_2 = (0.01, 0.8, 0.02, 0.01, 0.1, 1.9, 0.5)$ . Gráficos superiores: Cadeia de Markov simulada convergindo para a distribuição a posteriori. A linha pontilhada indica o verdadeiro valor do parâmetro. Gráficos inferiores: Distribuições a posteriori marginais. A barra horizontal contínua indica o verdadeiro valor do parâmetro e a barra pontilhada o valor estimado.

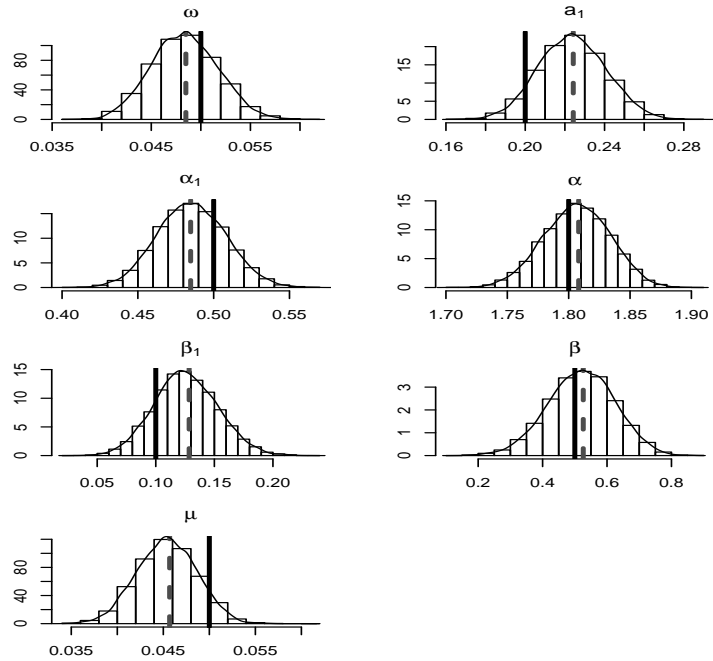
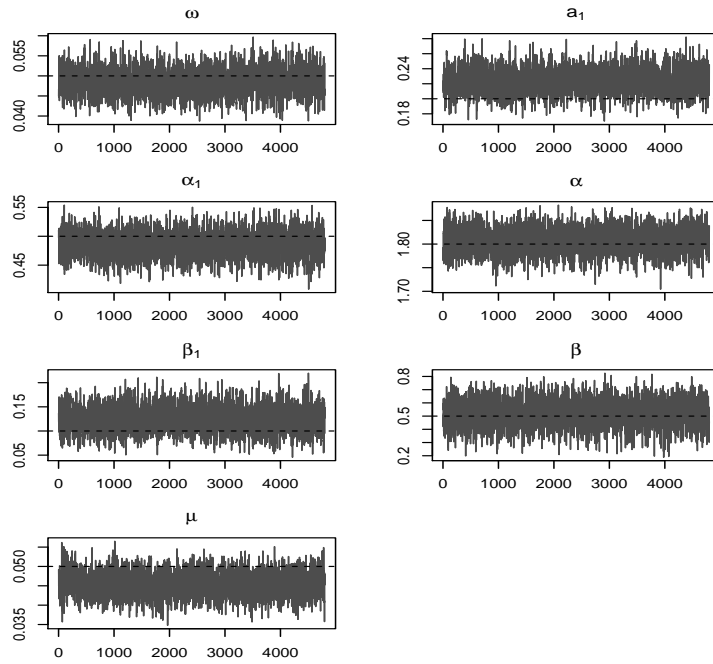


Figura 9.6: STABLE-GARCH com  $\varphi_3 = (0.05, 0.5, 0.1, 0.05, 0.2, 1.8, 0.5)$ . Gráficos superiores: Cadeia de Markov simulada convergindo para a distribuição a posteriori. A linha pontilhada indica o verdadeiro valor do parâmetro. Gráficos inferiores: Distribuições a posteriori marginais. A barra horizontal contínua indica o verdadeiro valor do parâmetro e a barra pontilhada o valor estimado.

---

**Algorithm 9.2** Metropolis-Hastings with Gibbs sample - STABLE-GARCH

---

**Input:**  $(x_1, x_2, \dots, x_n)$ ,  $\tau = (\tau_1, \dots, \tau_6, \tau_7)$  e  $\eta$  (Tamanho da cadeia).

**Output:**  $\theta^{(1)}, \dots, \theta^{(\eta)}$ , com  $\theta^{(i)} = (\alpha_0^{(i)}, \alpha_1^{(i)}, \alpha_2^{(i)}, \beta_0^{(i)}, \beta_1^{(i)}, \alpha^{(i)}, \beta^{(i)})$ .

1: Defina  $\theta^0 = (0.1, 0.1, 0.1, 0.1, 0.1, 1.8, 0)$

2: **for**  $i = 1, 2, \dots, \eta$  **do**

3: Simule  $\alpha_0^*$  de  $N(\alpha_0^{(i-1)}, \tau_1)$

4: **if**  $\alpha_0^* \leq 0$  **then**

5:  $\alpha_0^{(i)} = \alpha_0^{(i-1)}$

6: **else**

7: Calcule a razão das densidades,

$$r = \frac{p(x|\alpha_0^*, \theta_{-1}^{(i-1)})/N(\alpha_0^*|\alpha_0^{i-1})}{p(x|\theta_{+1}^{(i-1)})/N(\alpha_0^{(i-1)}|\alpha_0^*)}$$

8: Aceite  $\alpha_0^*$  com probabilidade  $\min(1, r)$ .

9: **end if**

10: Simule  $\alpha_1^*$  de  $N(\alpha_1^{(i-1)}, \tau_2)$

11: Calcule  $\lambda(2/\pi)\Gamma(1 - 1/\alpha^{(i-1)})$

12: **if**  $\alpha_1^* \leq 0$  ||  $\lambda\alpha_1^* + \alpha_2^{(i-1)} \geq 1$  **then**

13:  $\alpha_1^{(i)} = \alpha_1^{(i-1)}$

14: **else**

15: Calcule a razão das densidades,

$$r = \frac{p(x|\alpha_1^*, \theta_{-2}^{(i-1)})/N(\alpha_1^*|\alpha_1^{i-1})}{p(x|\theta_{+2}^{(i-1)})/N(\alpha_1^{(i-1)}|\alpha_1^*)}$$

16: Aceite  $\alpha_1^*$  com probabilidade  $\min(1, r)$ .

17: **end if**

18: Simule  $\alpha_2^*$  de  $N(\alpha_2^{(i-1)}, \tau_3)$

19: Calcule  $\lambda(2/\pi)\Gamma(1 - 1/\alpha^{(i-1)})$

20: **if**  $\alpha_2^* \leq 0$  ||  $\lambda\alpha_1^{(i-1)} + \alpha_2^* \geq 1$  **then**

21:  $\alpha_2^{(i)} = \alpha_2^{(i-1)}$

22: **else**

23: Calcule a razão das densidades,

$$r = \frac{p(x|\alpha_2^*, \theta_{-3}^{(i-1)})/N(\alpha_2^*|\alpha_2^{i-1})}{p(x|\theta_{+3}^{(i-1)})/N(\alpha_2^{(i-1)}|\alpha_2^*)}$$

24: Aceite  $\alpha_2^*$  com probabilidade  $\min(1, r)$ .

25: **end if**

---

---

26: Simule  $\beta_0^*$  de  $N(\beta_0^{(i-1)}, \tau_4)$

27: Calcule a razão das densidades,

$$r = \frac{p(x|\beta_0^*, \theta_{-4}^{(i-1)})/N(\beta_4^*|\beta_4^{(i-1)})}{p(x|\theta_{+4}^{(i-1)})/N(\beta_0^{(i-1)}|\beta_0^*)}$$

28: Aceite  $\beta_0^*$  com probabilidade  $\min(1, r)$ .

29: **if**  $|\beta_1|^* \leq 1$  **then**

30:  $\beta_1^{(i)} = \beta_1^{(i-1)}$

31: **else**

32: Simule  $\beta_1^*$  de  $N(\beta_1^{i-1}, \tau_5)$

33: Calcule a razão das densidades,

$$r = \frac{p(x|\beta_1^*, \theta_{-5}^{(i-1)})/N(\beta_1^*|\beta_1^{(i-1)})}{p(x|\theta_{+5}^{(i-1)})/N(\beta_1^{(i-1)}|\beta_1^*)}$$

34: Aceite  $\beta_1^*$  com probabilidade  $\min(1, r)$ .

35: **end if**

36: Simule  $\alpha^*$  de  $N(\alpha^{(i-1)}, \tau_6)$

37: Calcule  $\lambda(2/\pi)\Gamma(1 - 1/\alpha^*)$

38: **if**  $\alpha^* < 1.01$  ||  $\alpha^* > 2$  ||  $\lambda\alpha_1^{(i)} + \alpha_1^{(i)} + \geq 1$  **then**

39:  $\alpha^{(i)} = \alpha^{(i-1)}$

40: **else**

41: Calcule a razão das densidades,

$$r = \frac{p(x|\alpha^*, \theta_{-6}^{(i-1)})/N(\alpha^*|\alpha^{i-1})}{p(x|\theta_{+6}^{(i-1)})/N(\alpha^{(i-1)}|\alpha^*)}$$

42: Aceite  $\alpha_2^*$  com probabilidade  $\min(1, r)$ .

43: **end if**

44: Simule  $\beta^*$  de  $N(\beta^{(i-1)}, \tau_7)$

45: Calcule  $\lambda(2/\pi)\Gamma(1 - 1/\beta^*)$

46: **if**  $|\beta| > 0.98$  **then**

47:  $\beta^{(i)} = \beta^{(i-1)}$

48: **else**

49: Calcule a razão das densidades,

$$r = \frac{p(x|\beta^*, \beta_{-7}^{(i-1)})/N(\beta^*|\beta^{i-1})}{p(x|\beta_{+7}^{(i-1)})/N(\beta^{(i-1)}|\beta^*)}$$

50: Aceite  $\beta^*$  com probabilidade  $\min(1, r)$ .

51: **end if**

Tabela 9.7: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo AR(1)-GARCH(1,1) com inovações estáveis utilizando MLE.

		Simulado	Estimado	RMSE
$\theta_1$	$\omega$	0.01	0.010059	0.000707
	$\alpha_1$	0.45	0.446819	0.019922
	$\beta_1$	0.08	0.080364	0.030893
	$\mu$	0.21	0.210501	0.005096
	$a_1$	0.32	0.318587	0.016358
	$\alpha$	1.80	1.803118	0.028536
	$\beta$	0.50	0.516406	0.117259
$\theta_2$	$\omega$	0.01	0.010357	0.000940
	$\alpha_1$	0.80	0.763452	0.070065
	$\beta_1$	0.02	0.032211	0.043768
	$\mu$	0.01	0.009982	0.000792
	$a_1$	0.10	0.106499	0.055180
	$\alpha$	1.90	1.897083	0.034175
	$\beta$	0.50	0.484124	0.252155
$\theta_3$	$\omega$	0.05	0.050466	0.003512
	$\alpha_1$	0.50	0.497222	0.019992
	$\beta_1$	0.10	0.098024	0.025639
	$\mu$	0.05	0.050070	0.003020
	$a_1$	0.20	0.202034	0.015298
	$\alpha$	1.80	1.794382	0.027657
	$\beta$	0.50	0.500529	0.101071

Tabela 9.8: Resultados da estimação dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações estáveis utilizando a metodologia Bayesiana.

		Simulado	Estimado	Erro Padrão	Intervalo de Confiança	
					Inferior	Superior
$\theta_1$	$\omega$	0.01	0.009505	0.000621	0.008337	0.010775
	$\alpha_1$	0.45	0.446963	0.022009	0.403457	0.488811
	$\beta_1$	0.08	0.091957	0.030485	0.033637	0.151435
	$\mu$	0.21	0.205925	0.004838	0.196924	0.215929
	$a_1$	0.32	0.332108	0.015437	0.303533	0.363882
	$\alpha$	1.80	1.789799	0.026251	1.737607	1.839862
	$\beta$	0.50	0.636655	0.090100	0.461694	0.808948
$\theta_2$	$\omega$	0.01	0.010062	0.000486	0.009127	0.011042
	$\alpha_1$	0.80	0.805142	0.016406	0.770955	0.832813
	$\beta_1$	0.02	0.025182	0.008110	0.010462	0.039669
	$\mu$	0.01	0.009550	0.000483	0.008594	0.010454
	$a_1$	0.10	0.103197	0.008406	0.086777	0.122612
	$\alpha$	1.90	1.908750	0.019281	1.870390	1.945810
	$\beta$	0.50	0.592706	0.149997	0.296546	0.873197
$\theta_3$	$\omega$	0.05	0.048385	0.003169	0.042278	0.054702
	$\alpha_1$	0.50	0.485477	0.021375	0.443591	0.526016
	$\beta_1$	0.10	0.125726	0.025703	0.077800	0.179114
	$\mu$	0.05	0.045359	0.003081	0.039688	0.051684
	$a_1$	0.20	0.223647	0.016331	0.192847	0.255601
	$\alpha$	1.80	1.807133	0.025340	1.759032	1.857239
	$\beta$	0.50	0.521345	0.098722	0.332315	0.720121



Tabela 9.9: Teste de estacionariedade para as cadeias simuladas da distribuição a posteriori dos parâmetros do modelo AR(1)-GARCH(1,1) com inovações estáveis.

		Teste Estac.	Manter	Descartar	C-von-M	Teste HWD	Estatística HWD
	$\omega$	Passou	50000	0	0.152077	Passou	2.28e-05
	$\alpha_1$	Passou	25000	25000	0.301989	Passou	0.001146
$\phi_1$	$\beta_1$	Passou	50000	0	0.444079	Passou	0.001126
	$\mu$	Passou	50000	0	0.057387	Passou	0.000250
	$a_1$	Passou	50000	0	0.055832	Passou	0.000811
	$\alpha$	Passou	45000	5000	0.192486	Passou	0.001298
	$\beta$	Passou	50000	0	0.068858	Passou	0.004020
	$\omega$	Passou	50000	0	0.047771	Passou	2.39e-05
	$\alpha_1$	Passou	50000	0	0.055716	Passou	0.000674
$\phi_2$	$\beta_1$	Passou	50000	0	0.156060	Passou	0.000405
	$\mu$	Passou	50000	0	0.091165	Passou	2.03e-05
	$a_1$	Passou	50000	0	0.130223	Passou	0.000636
	$\alpha$	Passou	50000	0	0.272238	Passou	0.000995
	$\beta$	Passou	50000	0	0.106011	Passou	0.005350
	$\omega$	Passou	50000	0	0.056083	Passou	0.000125
	$\alpha_1$	Passou	50000	0	0.408013	Passou	0.000842
$\phi_3$	$\beta_1$	Passou	50000	0	0.188276	Passou	0.001122
	$\mu$	Passou	50000	0	0.089247	Passou	0.000143
	$a_1$	Passou	50000	0	0.110418	Passou	0.000716
	$\alpha$	Passou	50000	0	0.202972	Passou	0.000922
	$\beta$	Passou	50000	0	0.059553	Passou	0.004084

A coluna nomeada como *Teste Estac.* indica se a cadeia passou no teste de estacionariedade e quantas observações devemos utilizar após descartar a parte inicial da cadeia. Como todas as cadeias passaram no teste o *Half-Width Test* foi também realizado e reportou o valor da estatística de *Cramer-Von-Misses* (*C-Von-M*).

Tabela 9.10: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo AR(1)-GARCH(1,1) com inovações estáveis utilizando a metodologia Bayesiana.

		Simulado	Estimado	RMSE
$\theta_1$	$\omega$	0.01	0.010022	0.000653
	$\alpha_1$	0.45	0.444207	0.021076
	$\beta_1$	0.08	0.084949	0.028777
	$\mu$	0.21	0.210386	0.004740
	$a_1$	0.32	0.318889	0.015216
	$\alpha$	1.80	1.795935	0.026635
	$\beta$	0.50	0.507823	0.107869
$\theta_2$	$\omega$	0.01	0.010011	0.000877
	$\alpha_1$	0.80	0.753678	0.084871
	$\beta_1$	0.02	0.049824	0.061930
	$\mu$	0.01	0.010082	0.000526
	$a_1$	0.10	0.099181	0.009174
	$\alpha$	1.90	1.897009	0.025272
	$\beta$	0.50	0.498280	0.173566
$\theta_3$	$\omega$	0.05	0.050615	0.003448
	$\alpha_1$	0.50	0.496373	0.020071
	$\beta_1$	0.10	0.999972	0.025001
	$\mu$	0.05	0.050156	0.003016
	$a_1$	0.20	0.201793	0.014947
	$\alpha$	1.80	1.793039	0.028316
	$\beta$	0.50	0.489071	0.098248

# Aspectos Computacionais

Tabela 9.11: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-APARCH(2,2) com inovações estáveis.

	Simulado	Estimado	RMSE
$\mu$	0.10	0.104749	0.009536
$a_1$	0.50	0.502541	0.024582
$a_2$	0.40	0.397443	0.023127
$b_1$	-0.40	-0.403155	0.023504
$b_2$	0.40	0.402157	0.019651
$\omega$	0.05	0.084491	0.043810
$\alpha_1$	0.10	0.155981	0.064754
$\alpha_2$	0.10	0.150961	0.062041
$\gamma_1$	0.50	0.508365	0.160200
$\gamma_2$	-0.50	-0.560181	0.183680
$\beta_1$	0.10	0.095821	0.089261
$\beta_2$	0.05	0.050094	0.047907
$\delta$	1.30	1.267808	0.222917
$\beta$	-0.70	-0.682651	0.093060
$\alpha$	1.80	1.797829	0.028910

Tabela 9.12: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-APARCH(2,2) com inovações estáveis.

	Simulado	Estimado	RMSE
$\mu$	0.00	-0.004236	0.174858
$a_1$	0.20	0.201691	0.072735
$a_2$	-0.20	-0.206351	0.046882
$b_1$	0.30	0.301405	0.093761
$b_2$	0.50	0.503564	0.028391
$\omega$	0.20	0.298209	0.117561
$\alpha_1$	0.05	0.072010	0.031251
$\alpha_2$	0.10	0.144512	0.053401
$\gamma_1$	0.20	0.183775	0.249802
$\gamma_2$	-0.20	-0.198915	0.132520
$\beta_1$	0.05	0.071745	0.084388
$\beta_2$	0.20	0.172981	0.096978
$\delta$	1.10	1.019102	0.247918
$\beta$	0.50	0.498556	0.083375
$\alpha$	1.60	1.609204	0.038702

Tabela 9.13: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-GARCH(2,2) com inovações estáveis.

	Simulado	Estimado	RMSE
$\mu$	0.1	0.054902	0.045461
$a_1$	0.2	0.203443	0.045727
$a_2$	-0.2	-0.198906	0.039245
$b_1$	0.3	0.295398	0.04455
$b_2$	0.5	0.497522	0.026494
$\omega$	0.1	0.139236	0.045154
$\alpha_1$	0.1	0.140756	0.048278
$\alpha_2$	0.1	0.134487	0.048681
$\beta_1$	0.1	0.115413	0.12539
$\beta_2$	0.1	0.094161	0.096718
$\beta$	0	0.017918	0.203769
$\alpha$	1.9	1.903454	0.022656
$\mu$	0.1	0.045804	0.058307
$a_1$	0.2	0.197849	0.040082
$a_2$	-0.2	-0.192536	0.037634
$b_1$	0.3	0.308166	0.039736
$b_2$	0.5	0.500282	0.021804
$\omega$	0.1	0.1387	0.04417
$\alpha_1$	0.1	0.137148	0.04277
$\alpha_2$	0.1	0.138083	0.049761
$\beta_1$	0.1	0.111581	0.1192
$\beta_2$	0.1	0.085858	0.078253
$\beta$	-0.5	-0.379408	0.18822
$\alpha$	1.7	1.715062	0.032009

Tabela 9.14: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(1,1)-GARCH(1,1) com inovações estáveis.

	Simulado	Estimado	RMSE
$\mu$	0.1	0.054902	0.045461
$a_1$	0.2	0.203443	0.045727
$b_1$	0.3	0.295398	0.04455
$\omega$	0.1	0.139236	0.045154
$\alpha_1$	0.1	0.140756	0.048278
$\beta_1$	0.1	0.115413	0.12539
$\beta$	0	0.017918	0.203769
$\alpha$	1.9	1.903454	0.022656
$\mu$	0.1	0.080337	0.020385
$a_1$	0.3	0.300774	0.034104
$b_1$	0.2	0.196024	0.032871
$\omega$	0.1	0.140706	0.043409
$\alpha_1$	0.2	0.27687	0.081377
$\beta_1$	0.15	0.140878	0.068714
$\beta$	0.5	0.491312	0.09692
$\alpha$	1.8	1.802981	0.02527

Tabela 9.15: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-APARCH(2,2) com inovações GEV.

	Simulado	Estimado	RMSE
$\mu$	0.10	0.099171	0.008888
$a_1$	0.50	0.499481	0.018019
$a_2$	0.40	0.401022	0.017749
$b_1$	-0.40	-0.398776	0.016219
$b_2$	0.40	0.398895	0.014316
$\omega$	0.05	0.055905	0.016546
$\alpha_1$	0.10	0.097183	0.029693
$\alpha_2$	0.10	0.100697	0.033444
$\gamma_1$	0.50	0.445066	0.210093
$\gamma_2$	-0.50	-0.613494	0.327781
$\beta_1$	0.10	0.095959	0.054308
$\beta_2$	0.05	0.049454	0.042078
$\delta$	1.30	1.269226	0.177481
$\xi$	0.20	0.202141	0.015467
$\mu$	0.00	0.001935	0.012034
$a_1$	0.20	0.200421	0.040526
$a_2$	-0.20	-0.197321	0.034729
$b_1$	0.30	0.299255	0.037702
$b_2$	0.50	0.499705	0.018860
$\omega$	0.20	0.221258	0.051102
$\alpha_1$	0.05	0.052011	0.025793
$\alpha_2$	0.10	0.102571	0.034888
$\gamma_1$	0.20	-0.032763	0.550351
$\gamma_2$	-0.20	-0.341750	0.332478
$\beta_1$	0.05	0.060884	0.069265
$\beta_2$	0.20	0.200651	0.076452
$\delta$	1.50	1.351043	0.307075
$\xi$	0.10	0.101032	0.013058

Tabela 9.16: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(2,2)-GARCH(2,2) com inovações GEV.

	Simulado	Estimado	RMSE
$\mu$	0.1	0.05618	0.04655
$a_1$	0.2	0.20078	0.04423
$a_2$	-0.2	-0.20198	0.03628
$b_1$	0.3	0.29946	0.03797
$b_2$	0.5	0.50151	0.02205
$\omega$	0.2	0.19941	0.02234
$\alpha_1$	0.05	0.04852	0.00966
$\alpha_2$	0.1	0.09995	0.01669
$\beta_1$	0.05	0.0671	0.07987
$\beta_2$	0.2	0.189	0.07434
$\xi$	-0.1	0.10245	0.20299
$\mu$	0.1	0.05674	0.0444
$a_1$	0.2	0.19714	0.05707
$a_2$	-0.2	-0.20228	0.04743
$b_1$	0.3	0.30067	0.05296
$b_2$	0.5	0.50004	0.02452
$\omega$	0.1	0.10085	0.0123
$\alpha_1$	0.1	0.09988	0.01911
$\alpha_2$	0.2	0.20233	0.02644
$\beta_1$	0.1	0.09935	0.0846
$\beta_2$	0.1	0.09405	0.06169
$\xi$	-0.2	-0.19972	0.01217



Tabela 9.17: Sumário da simulação de 100 séries temporais de tamanho 2500 do modelo ARMA(1,1)-GARCH(1,1) com inovações GEV.

	Simulado	Estimado	RMSE
$\mu$	0.1	0.071336	0.029476
$a_1$	0.3	0.299885	0.022146
$b_1$	0.4	0.401414	0.020065
$\omega$	0.05	0.049633	0.003965
$\alpha_1$	0.2	0.200334	0.020444
$\beta_1$	0.2	0.202887	0.041746
$\xi$	0.1	0.101097	0.01406
$\mu$	0.1	0.072317	0.028986
$a_1$	0.3	0.294937	0.026564
$b_1$	0.4	0.399256	0.024234
$\omega$	0.05	0.050733	0.005662
$\alpha_1$	0.2	0.199735	0.021919
$\beta_1$	0.2	0.188032	0.063252
$\xi$	-0.1	-0.100019	0.01438

# Previsão de parâmetros futuros

Tabela 9.18: Comparação da volatilidade realizada com a volatilidade estimada com dados simulados do modelo ARMA(1,1)-GARCH(1,1) com inovações GEV e parâmetros  $(\mu, a_1, b_2, \omega, \alpha_1, \beta_2, \xi) = (0.1, 0.2, 0.2, 0.05, 0.1, 0.1, 0.2)$ .

t	$\sigma_t$	$\hat{\sigma}_t$	$\sigma_t$	$\hat{\sigma}_t$
	Simulação 1		Simulação 2	
1	0.083265	0.077293	0.063754	0.061931
2	0.077731	0.069891	0.069197	0.064461
3	0.058047	0.068233	0.059858	0.064878
4	0.078902	0.067879	0.062748	0.064948
5	0.066199	0.067805	0.066226	0.064960
6	0.057921	0.067789	0.057819	0.064962
7	0.060385	0.067786	0.057141	0.064962
8	0.069343	0.067785	0.057763	0.064962
9	0.065977	0.067785	0.087183	0.064962
10	0.074985	0.067785	0.063436	0.064962
MAE	0.006446		0.006095	
AMAPE	0.046959		0.045588	

Tabela 9.19: Comparação da volatilidade realizada com a volatilidade estimada com dados simulados do modelo ARMA(2,2)-GARCH(2,2) com inovações GEV e parâmetros  $(\mu, a_1, a_2, b_1, b_2, \omega, \alpha_1, \alpha_2, \beta_1, \beta_2, \xi) = (0.1, 0.2, 0.3, 0.2, 0.5, 0.1, 0.1, 0.2, 0.1, 0.1, 0.1)$ .

t	$\sigma_t$	$\hat{\sigma}_t$	$\sigma_t$	$\hat{\sigma}_t$
	Simulação 1		Simulação 2	
1	0.810359	0.838193	0.381888	0.496218
2	0.518823	0.544982	0.373898	0.403720
3	0.939085	0.683258	0.371059	0.508229
4	0.675664	0.546479	0.474023	0.481299
5	0.541068	0.594171	0.374269	0.523300
6	0.549311	0.529745	0.377645	0.516689
7	0.491532	0.543133	0.362805	0.535112
8	0.413888	0.512178	0.362084	0.534579
9	0.380325	0.513693	0.423932	0.543176
10	0.363047	0.498326	0.397188	0.544130
MAE	0.093021		0.118766	
MAPE	0.083201		0.131735	

Tabela 9.20: Comparação da volatilidade realizada com a volatilidade estimada com dados simulados do modelo ARMA(1,1)-GARCH(1,1) com inovações estáveis e parâmetros  $(\mu, a_1, a_2, \omega, \alpha_1, \beta_2, \alpha, \beta) = (0.1, 0.2, 0.2, 0.05, 0.1, 1.8, 0)$ .

t	$\sigma_t$	$\hat{\sigma}_t$	$\sigma_t$	$\hat{\sigma}_t$
	Simulação 1		Simulação 2	
1	0.077276	0.104338	0.083364	0.111881
2	0.069376	0.107523	0.080479	0.109390
3	0.110846	0.109115	0.064067	0.108669
4	0.116857	0.109912	0.084996	0.108460
5	0.071424	0.110310	0.066374	0.108400
6	0.061990	0.110510	0.060681	0.108382
7	0.093131	0.110609	0.079765	0.108377
8	0.073302	0.110659	0.066466	0.108376
9	0.062603	0.110684	0.064141	0.108375
10	0.070642	0.110696	0.079001	0.108375
MAE	0.030426		0.035935	
AMAPE	0.168561		0.200537	

Tabela 9.21: Comparação da volatilidade realizada com a volatilidade estimada com dados simulados do modelo ARMA(2,2)-GARCH(2,2) com inovações estáveis e parâmetros  $(\mu, a_1, a_2, b_1, b_2, \omega, \alpha_1, \alpha_2, \beta_1, \beta_2, \alpha, \beta) = (0.1, 0.2, 0.3, 0.2, 0.5, 0.1, 0.1, 0.2, 0.1, 0.1, 1.9, 0.3)$ .

t	$\sigma_t$	$\hat{\sigma}_t$	$\sigma_t$	$\hat{\sigma}_t$
	Simulação 1		Simulação 2	
1	0.222905	0.242732	0.158924	0.189553
2	0.214127	0.160318	0.183135	0.174979
3	0.204537	0.252808	0.167318	0.251859
4	0.358183	0.245442	0.188012	0.263049
5	0.184012	0.280231	0.306156	0.296192
6	0.18699	0.286867	0.272214	0.307986
7	0.255353	0.30254	0.215033	0.323816
8	0.246157	0.309492	0.319019	0.332022
9	0.189057	0.317642	0.351069	0.340152
10	0.352254	0.322651	0.35798	0.345223
MAE	0.069945		0.139283	
MAPE	0.038955		0.081269	

# Pacote `garchFitExtremes`

Para instalar o pacote siga as seguintes instruções:

- Instale o software chamado `Rtools.exe` disponível no seguinte endereço:  
*<http://cran.r-project.org/bin/windows/Rtools/>*
- Coloque o arquivo `garchFitExtremes.tar.gz` no diretório raiz `c:`
- Abra uma janela do prompt de comando do Windows com privilégios de administrador e aponte para o diretório raiz `c:.`
- Digite o seguinte comando R CMD `INSTALL garchFitExtremes`

Após seguir esses passos basta abrir o software R e digitar o comando `library(garchFitExtremes)` para ter acesso as funções deste pacote. O documento apresentado neste Apêndice explica o funcionamento das funções existentes no pacote `garchFitExtremes` implementado neste trabalho.

# Package ‘garchFitExtremes’

July 2, 2013

**Type** Package

**Title** garchFitExtremes

**Version** 1.0

**Date** 2013-06-25

**Depends** fGarch, Rsolnp, fExtremes, stable, skewt

**Author** Thiago do Rego Sousa

**Maintainer** Thiago do Rego Sousa <thiagoestatistico@gmail.com>

**Description**

Package for ARMA-APARCH modelling that allows GEV and stable conditional distributions.

**License** GPL

## R topics documented:

garchFitExtremes-package . . . . .	1
aic.fit . . . . .	2
forecastExtremes . . . . .	4
garchExtremesSIM . . . . .	5
garchFit.new . . . . .	6

---

garchFitExtremes-package

*ARMA-GARCH/APARCH modelling with GEV and stable distributions*

---

**Description**

This package is design to perform maximum likelihood estimation of ARMA-GARCH/APARCH with the following conditional distributions: Normal, t-student, skew t-student, GEV and stable distributions.

**Details**

Package: garchFitExtremes  
 Type: Package  
 Version: 1.0  
 Date: 2013-06-25  
 License: GPL

The most important function of the package is called `garchFit.new`. It allows the user to estimate the model choosing the specific distribution to be used.

**Author(s)**

Thiago do Rego Sousa.

Maintainer: Thiago do Rego Sousa <thiagoestatistico@gmail.com>

**References**

Mittnik S. , Paoletta M.S. , Rachev, S.T. Stationarity of stable power- GARCH processes. Journal of Econometrics 106, 97-107. 2002.

Wurtz, D., Y. Chalabi, e L. Luksan. Parameter estimation of ARMA models with GARCH/APARCH errors: An R and SPlus software implementation. Journal of Statistical Software, 2006.

Zhao X. ,Scarrott C. J. , Oxley L. , Reale M. GARCH dependence in extreme value models with Bayesian inference. Mathematics an Computers in Simulation, Vol. 81, Issue 7, 1430-1440. 2011.

Brockwell, P. J, e Davis, R. A. Introduction to Time Series and Forecasting. Springer, New York, 1996.

Nolan, J. P. Numerical calculations of stable densities and distribution functions. Communications in Statistics - Stochastic Models 13: 759- 774, 1997.

---

aic.fit

*ARMA-APARCH with minimum AIC*

---

**Description**

This function estimates ARMA(m,n)-APARCH(p,q) models with varying order and chooses the order that has the minimum AIC as the best model.

**Usage**

```
aic.fit(data, mMAX, nMAX, pMAX, qMAX, cond.dist,
algorithm = "sqp", APARCH = FALSE, control = NULL)
```



**Arguments**

data	Data with the time series to be estimated. This object can not contain NULL elements.
mMAX	Maximum m order of the ARMA(m,n) equation to be searched.
nMAX	Maximum n of the ARMA(m,n) equation order to be searched.
pMAX	Maximum p order of the GARCH(p,q) equation to be searched.
qMAX	Maximum q order of the GARCH(p,q) equation to be searched.
cond.dist	The conditional distribution of the model to be estimated. This parameter should be one of the following strings: "norm" (Normal Distribution), "t-student" (t-Student distribution), "sstd" (skew t-Student Distribution), "GEV" (Generalized Extreme Value Distribution) or "stable" (Stable Distribution).
algorithm	The algorithm to be used to optimize the likelihood of the process. This parameter should be one of the following strings: "nlminb" ( nlminb R internal Routine for optimization) or "SQP" ( Sequential quadratic programming routine implemented in package RSolnp).
APARCH	You can set this boolean value as FALSE if we want to estimate a pure GARCH model or as TRUE if we want an APARCH model.
control	Control parameters used either in "SQP" routine (see the RSolnp documentation) or in "nlminb" (see the nlminb R documentation).

**Value**

The model with minimum AIC found within the specified range choosen in paramters mMAX, nMAX, pMAX, qMAX.

**Author(s)**

Thiago do Rego Sousa

**References**

Brockwell, P. J, e Davis, R. A. Introduction to Time Series and Forecasting. Springer, New York, 1996.

**Examples**

```
# AIC fit using models from ARMA(0,0)-GARCH(1,0) to ARMA(1,1)-GARCH(1,1)
# with stable conditional distribution
library(fGarch)
data(dem2gbp)
x = dem2gbp[, 1]
aic.fit(data = x,1,1,1,1,cond.dist = "stable")
```

---

forecastExtremes	<i>Forecast of the conditional variance of an ARMA-GARCH/APARCH model.</i>
------------------	--

---

### Description

This functions uses the equations given in Wurtz et. al. (2006) to forecast the conditional variance of the ARMA-GARCH or ARMA-APARCH model.

### Usage

```
forecastExtremes(model, lag)
```

### Arguments

model	This item should be an object returned by the function garchFit.new.
lag	It is the forecast horizon to be used. Note that lag should be greater or equal 1.

### Value

This function returns a vector with the forecasts of the conditional variance of the ARMA-GARCH/APARCH process given in the variable "model" described above.

### Author(s)

Thiago do Rego Sousa.

### References

Wurtz, D., Y. Chalabi, e L. Luksan. Parameter estimation of ARMA models with GARCH/APARCH errors: An R and SPlus software implementation. Journal of Statistical Software, 2006.

### Examples

```
# Forecasting comparison with garchFit routine from package fGarch
# implemented by Wurtz et. al. (2006).
data(dem2gbp)
x = dem2gbp[, 1]
gF.new = garchFit.new(data = x , 1,1,1,1,
cond.dist = "normal", intercept = TRUE,trace = 1,
algorithm = "sqp", printRes = TRUE,get.res = TRUE, APARCH = FALSE)
gF = garchFit(data = x, formula = ~arma(1,1) + garch(1,1),
cond.dist = "norm")
cbind(gF.new$par,gF@fit$par)
gF.predict <- predict(gF);
gF.new.predict <- predict.extremes.sig(model = gF.new, lag = 10)
cbind(gF.predict[[3]],gF.new.predict, abs(gF.predict[[3]]-gF.new.predict))
```

---

garchExtremesSIM	<i>Simulation of ARMA-GARCH/APARCH process with GEV and stable conditional distributions</i>
------------------	--

---

### Description

This functions simulate time series following ARMA-GARCH/APARCH models with several conditional distributions, including GEV and stable distributions. The initial values of the time series are fixed and the recursion formulas of the model are used to simulate the dynamics of the process.

### Usage

```
garchExtremesSIM(N, mu , a , b , omega, alpha, gm, beta,
delta, skew, shape, cond.dist)
```

### Arguments

**N** The size of simulated time series  
**mu , a, b, omega, alpha, gm, beta, delta**  
 These are the parameters of the model to be simulated. They should be passed as real valued vectors depending on the model's order. The parameters will be interpreted following the formula also given in Wurtz et. al. (2006):

$$\begin{cases} X_t = \mu + \sum_{i=1}^m a_i X_{t-i} + \sum_{j=1}^n b_j \varepsilon_{t-j} + \varepsilon_t \\ \varepsilon_t = \sigma_t z_t, \quad z_t \stackrel{iid}{\sim} D(0, 1) , \\ \sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (\varepsilon_{t-i} - \gamma_i |\varepsilon_{t-i}|)^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta \end{cases}$$

**skew** The skew parameter of the conditional distribution. Used only when cond.dist equals "stable" or "sstd".

**shape** Shape parameter of the conditional distribution. Used only when cond.dist equals "stable", "GEV" "t-student" or "sstd".

**cond.dist** The conditional distribution of the model to be estimated. This parameter should be one of the following strings: "norm" (Normal Distribution), "t-student" (t-Student distribution), "sstd" (skew t-Student Distribution), "GEV" (Generalized Extreme Value Distribution) or "stable" (Stable Distribution).

### Value

This functions returns a matrix with two columns. The first column is the simulated process  $X[t]$  and the second one is the  $\sigma[t]$  simulated process.

### Author(s)

Thiago do Rego Sousa.

### References

Brockwell, P. J, e Davis, R. A. Introduction to Time Series and Forecasting. Springer, New York, 1996.

Nolan, J. P. Numerical calculations of stable densities and distribution functions. Communications in Statistics - Stochastic Models 13: 759- 774, 1997.

## Examples

```
# Simulation of a ARMA-APARCH process with stable conditional distribution
x <- garchExtremesSIM(N = 2500, mu = 0.1, a = c(0.2,0.3), b = c(0.2,0.5),
omega = 0.1, alpha = c(0.1,0.2), beta = c(0.1,0.1),
delta = 1, skew = 0.3, shape = 1.9, cond.dis = "stable")
```

---

garchFit.new	<i>ARMA-GARCH/APARCH modelling with GEV and stable distributions</i>
--------------	--

---

## Description

This function uses Maximum Likelihood technique to estimate the parameters of an ARMA-GARCH or ARMA-APARCH model with several conditional distributions. It allows the user to specify the following distributions: "Normal", "t-student", "skew t-student", "GEV" and "stable". The order of the model could be set up to 10 parameters in both ARMA or GARCH order.

## Usage

```
garchFit.new(data, m, n, p, q, intercept = TRUE, printRes = FALSE,
cond.dist = "normal", APARCH = FALSE, algorithm = "sqp",
get.res = FALSE, trace = 0, DEBUG = FALSE, control = NULL)
```

## Arguments

data	Data with the time series to be estimated. This object can not contain NULL elements.
m,n,p,q	These values are used to specify the order of the model. They will be treated as ARMA(m,n)-GARCH(p,q) or ARMA(m,n)-APARCH(p,q)
intercept	This is a boolean variable. If intercept is TRUE then we estimate the model with intercept, otherwise we will not use the intercept variable.
printRes	A boolean variable specifying whether or not the user wants to print the results after the function calling.
cond.dist	The conditional distribution of the model to be estimated. This parameter should be one of the following strings: "norm" (Normal Distribution), "t-student" (t-Student distribution), "sstd" (skew t-Student Distribution), "GEV" (Generalized Extreme Value Distribution) or "stable" (Stable Distribution).
APARCH	You can set this boolean value as FALSE if we want to estimate a pure GARCH model or as TRUE if we want an APARCH model.
algorithm	The algorithm to be used to optimize the likelihood of the process. This parameter should be one of the following strings: "nlminb" (nlminb R internal Routine for optimization) or "SQP" (Sequential quadratic programming routine implemented in package RSolnp).
get.res	If get.res is TRUE the output of the model will include the Residuals of the ARMA model and the Volatility of the GARCH model with the estimated model.
trace	This parameter is used if the "nlminb" algorithm is chosen.
control	Control parameters used either in "SQP" routine (see the RSolnp documentation) or in "nlminb" (see the nlminb R documentation).

**Value**

The function returns an object containing the following:

par	The estimated parameter set for the chosen model.
hessian	The estimated hessian matrix.

**Author(s)**

Thiago do Rego Sousa.

**References**

Mittnik S. , Paoletta M.S. , Rachev, S.T. Stationarity of stable power- GARCH processes. Journal of Econometrics 106, 97-107. 2002.

Wurtz, D., Y. Chalabi, e L. Luksan. Parameter estimation of ARMA models with GARCH/APARCH errors: An R and SPlus software implementation. Journal of Statistical Software, 2006.

Zhao X. ,Scarrott C. J. , Oxley L. , Reale M. GARCH dependence in extreme value models with Bayesian inference. Mathematics an Computers in Simulation, Vol. 81, Issue 7, 1430-1440. 2011.

**Examples**

```
# This examples uses the dataset of the package fGarch to estimate
# an ARMA(1,1)-GARCH(1,1) with stable conditional distribution.
library(fGarch)
data(dem2gbp)
x = dem2gbp[, 1]
gF.new = garchFit.new(data = x , 1,1,1,1,
cond.dist = "stable", intercept = TRUE, APARCH = TRUE, trace = 1,
algorithm = "nllminb", printRes = TRUE, get.res = TRUE)
```