

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito: 18.02.2005

Assinatura:

Modelos estocásticos com
heterocedasticidade para séries temporais
em finanças

Sandra Cristina de Oliveira

Orientador: Prof. Dr. Marinho Gomes de Andrade Filho

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências – Ciências de Computação e Matemática Computacional.

USP – São Carlos
Julho de 2005

*Aos meus pais,
Edgar e Regina,
e ao meu irmão Junior.*

Agradecimentos

A Deus, por ter me dado forças para iniciar e concluir mais uma etapa.

Ao Prof. Dr. Marinho G. Andrade Filho pela dedicada orientação, pelo incentivo à pesquisa e amizade durante o desenvolvimento deste trabalho.

Ao Prof. Dr. Carlos A. Diniz, ao Prof. Dr. Luiz K. Hotta e ao Prof. Dr. Marcos N. Arenales pelas valiosas sugestões apresentadas em meu Exame de Qualificação realizado em 02/12/2002.

Ao Prof. Dr. Guilherme A. Barreto e ao amigo Ricardo G. Silva pela contribuição dedicada a este trabalho.

`A minha família pelo amor, paciência e incentivo.

`A Prof^a. Dra. Vera L. Tomazella e ao Prof. Carlos A. Santos pela amizade, companheirismo e apoio constantes nos momentos tristes e felizes desta jornada.

`As amigas Valeria M. Ferreira e Juliana Cobre pela amizade e solidariedade especialmente no início e no término deste trabalho.

Aos amigos Viviane, Ornella, Renata, Rúbia, Helton, Selene, Leandro, Gilmar, Denise, Rodrigo, Lauro, Lilian, Josemar, Ricardo, Silvio, Selma, Juliana, Mônica, Pio, Thiago, Gleice, Cíntia, Veridiana, Cláudia, Led, Wendell, Igor, Guido, Ana Laura, Gecirlei, Fabrizio, Ulisses, Willian, Cillene, Josmar, Juliano e Wruck pelos bons momentos compartilhados.

*`A minha avó Paulina Z. Guanho, ao meu avô Geraldo Oliveira, ao meu tio Nelson Guanho e ao amigo Juan R. Cid pelos conselhos, incentivo e carinho. São lembrados com *saudades* hoje e sempre...*

Aos coordenadores da UNESP-Tupã, Prof. Dr. Elias J. Simon e Prof. Dr. José Matheus Perosa, e aos demais professores e amigos, pelo apoio e companheirismo na finalização do doutorado.

Aos professores, funcionários e colegas da USP e da UEM que, de alguma maneira, contribuíram para a realização deste trabalho.

`A CAPES e ao PCD-UNESP pelo auxílio financeiro.

Resumo

Neste trabalho desenvolvemos um estudo sobre modelos auto-regressivos com heterocedasticidade (ARCH) e modelos auto-regressivos com erros ARCH (AR-ARCH). Apresentamos os procedimentos para a estimação dos modelos e para a seleção da ordem dos mesmos. As estimativas dos parâmetros dos modelos são obtidas utilizando duas técnicas distintas: a inferência Clássica e a inferência Bayesiana. Na abordagem de Máxima Verossimilhança obtivemos intervalos de confiança usando a técnica Bootstrap e, na abordagem Bayesiana, adotamos uma distribuição a priori informativa e uma distribuição a priori não-informativa, considerando uma reparametrização dos modelos para mapear o espaço dos parâmetros no espaço real. Este procedimento nos permite adotar distribuição a priori normal para os parâmetros transformados. As distribuições a posteriori são obtidas através dos métodos de simulação de Monte Carlo em Cadeias de Markov (MCMC). A metodologia é exemplificada considerando séries simuladas e séries do mercado financeiro brasileiro.

Abstract

In this work we present a study of autoregressive conditional heteroskedasticity models (ARCH) and autoregressive models with autoregressive conditional heteroskedasticity errors (AR-ARCH). We also present procedures for the estimation and the selection of these models. The estimates of the parameters of those models are obtained using both Maximum Likelihood estimation and Bayesian estimation. In the Maximum Likelihood approach we get confidence intervals using Bootstrap resampling method and in the Bayesian approach we present informative prior and non-informative prior distributions, considering a reparametrization of those models in order to map the space of the parameters into real space. This procedure permits to choose prior normal distributions for the transformed parameters. The posterior distributions are obtained using Monte Carlo Markov Chain methods (MCMC). The methodology is exemplified considering simulated and Brazilian financial time series.

SUMÁRIO

Capítulo 1: Introdução	01
1.1 Introdução e Revisão Bibliográfica	01
1.2 Motivação	04
1.3 Objetivos	05
1.4 Organização da tese	05
Capítulo 2: Modelos Auto-Regressivos com Heterocedasticidade	07
2.1 Modelos $AR(p)$ - $ARCH(q)$	07
2.1.1 Introdução	07
2.1.2 Função de Verossimilhança	09
2.2 Modelos $ARCH(q)$	15
2.2.1 Introdução	15
2.2.2 Função de Verossimilhança	16
Capítulo 3: Inferência Bayesiana para Modelos $AR(p)$-$ARCH(q)$ e $ARCH(q)$	18
3.1 Inferência Bayesiana para Modelos $AR(p)$ - $ARCH(q)$	18
3.1.1 Distribuição a Priori Informativa	19
3.1.2 Distribuição a Priori Não-Informativa	22
3.2 Inferência Bayesiana para Modelos $ARCH(q)$	22
3.2.1 Distribuição a Priori Informativa	23
3.2.2 Distribuição a Priori Não-Informativa	24

Capítulo 4: Algoritmos para a Estimação dos Parâmetros de Modelos AR(p)-ARCH(q) e ARCH(q)	25
4.1 Algoritmos para a Abordagem de Máxima Verossimilhança dos Modelos AR(p)-ARCH(q)	25
4.1.1 Método de Newton	25
4.1.2 Procedimento Bootstrap Paramétrico	26
4.1.3 Critério de Akaike e Critério de Schwarz	28
4.2 Algoritmos para a Abordagem Bayesiana dos Modelos AR(p)-ARCH(q)	39
4.2.1 O Algoritmo Metropolis-Hastings	29
4.2.2 Critério da Densidade Preditiva Ordenada (CPO)	32
4.3 Algoritmos para a Abordagem de Máxima Verossimilhança dos Modelos ARCH(q)	33
4.3.1 Método de Newton	33
4.3.2 Procedimento Bootstrap Paramétrico	33
4.3.3 Critério de Akaike e Critério de Schwarz	34
4.4 Algoritmos para a Abordagem Bayesiana dos Modelos ARCH(q)	34
4.4.1 O Algoritmo Metropolis-Hastings	34
4.4.2 Critério da Densidade Preditiva Ordenada (CPO)	35
Capítulo 5: Aplicações	37
5.1 Introdução	37
5.2 Resultados obtidos com o Ajuste de Modelos ARCH(q)	38
5.2.1 Série Gerada	39
5.2.2 Série Índice Bovespa (IBovespa)	45
5.2.3 Série Telebrás	51
5.3 Resultados obtidos com o Ajuste de Modelos AR(p)-ARCH(q)	58
5.3.1 Série Gerada	58
5.3.2 Série Índice Bovespa (IBovespa)	66
5.3.3 Série Telebrás	73
Capítulo 6: Um Estudo Comparativo entre as Abordagens Bayesiana e de Máxima Verossimilhança para Modelos ARCH(q)	80
6.1 Introdução	80
6.2 Aplicações	82

Capítulo 7: Conclusões e Perspectivas	87
7.1 Conclusões	87
7.2 Perspectivas	88
Apêndices	90
A.1 Modelos Auto-Regressivos	90
A.2 Método de Simulação Bootstrap	93
A.3 Inferência Bayesiana	96
A.4 Gráficos da Seção 5.2	100
A.5 Gráficos da Seção 5.3	103
A.6 Gráficos do Capítulo 6	107
A.7 Análise de Resíduos e Volatilidade Estimada	109
Bibliografia	116

Lista de Tabelas

Tabela 5.1: Critérios de seleção de modelos - Modelo ARCH(q)	40
Tabela 5.2: Estimativas de MV e Intervalos de Confiança de 95%	40
Tabela 5.3: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI	41
Tabela 5.4: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI	43
Tabela 5.5: Critérios de seleção de modelos - Série IBovespa	47
Tabela 5.6: Estimativas de MV e Intervalos de Confiança de 95% - Série IBovespa	47
Tabela 5.7: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI	48
Tabela 5.8: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI	48
Tabela 5.9: Critérios de seleção de modelos - Série Telebrás	53
Tabela 5.10: Estimativas de MV e Intervalos de Confiança de 95% - Série Telebrás	53
Tabela 5.11: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI	54
Tabela 5.12: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI	54
Tabela 5.13: Critérios de seleção de modelos - Modelo AR(p)-ARCH(q)	60
Tabela 5.14: Estimativas de MV e Intervalos de Confiança de 95%	60
Tabela 5.15: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI	61
Tabela 5.16: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI	63
Tabela 5.17: Estimativas de MV e Intervalos de Confiança de 95% - Série IBovespa	67
Tabela 5.18: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI	69
Tabela 5.19: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI	69
Tabela 5.20: Critérios de seleção de modelos - Série IBobespa	72
Tabela 5.21: Estimativas de MV e Intervalos de Confiança de 95% - Série Telebrás	74
Tabela 5.22: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI	75
Tabela 5.23: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI	76
Tabela 5.24: Critérios de seleção de modelos - Série Telebrás	78
Tabela 6.1: Parâmetros dos Modelos Simulados	82
Tabela 6.2: Parâmetros da Transformação (3.4)	82
Tabela 6.3: Estimativas de MV e Bayesianas - Modelos ARCH(2)	83
Tabela 6.3 (cont.): Estimativas de MV e Bayesianas - Modelos ARCH(2)	84
Tabela 6.4: Estimativas de MV e Bayesianas - Modelo ARCH(5)	86

Lista de Figuras

Figura 5.1: Modelo ARCH(2) com $\alpha_0 = 0,05$, $\alpha_1 = 0,40$ e $\alpha_2 = 0,20$	39
Figura 5.2: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t - Modelo ARCH(2)	39
Figura 5.3: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t^2 - Modelo ARCH(2)	40
Figura 5.4: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Modelo ARCH(2)	41
Figura 5.5: Valores gerados dos parâmetros considerando distribuição a priori informativa - Modelo ARCH(2)	41
Figura 5.6: Análise de Resíduos - Modelo ARCH(2): Abordagem Bayesiana com distribuição a priori informativa	42
Figura 5.7: Volatilidade Estimada - Modelo ARCH(2)	42
Figura 5.8: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Modelo ARCH(2)	43
Figura 5.9: Valores gerados dos parâmetros considerando distribuição a priori não-informativa - Modelo ARCH(2)	43
Figura 5.10: Análise de resíduos - Modelo ARCH(2): Abordagem Bayesiana com distribuição a priori não-informativa	44
Figura 5.11: Volatilidade Estimada - Modelo ARCH(2)	44
Figura 5.12: Preço e Retorno - Série IBovespa	45
Figura 5.13: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t - Série IBovespa	46
Figura 5.14: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t^2 - Série IBovespa	46
Figura 5.15: Distribuição Empírica dos EMV - Série IBovespa	48
Figura 5.16: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Série IBovespa	49
Figura 5.17: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série IBovespa	49
Figura 5.18: Análise de resíduos - Série IBovespa: Abordagem Bayesiana com distribuição a priori informativa	50
Figura 5.19: Volatilidade Estimada - Série IBovespa	50
Figura 5.20: Preço e Retorno - Série Telebrás	51
Figura 5.21: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t - Série Telebrás	52

Figura 5.22: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t^2 - Série Telebrás	52
Figura 5.23: Distribuição Empírica dos EMV - Série Telebrás	53
Figura 5.23 (cont.): Distribuição Empírica dos EMV - Série Telebrás	54
Figura 5.24: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Série Telebrás	55
Figura 5.25: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série Telebrás	55
Figura 5.25 (cont.): Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série Telebrás	56
Figura 5.26: Análise de resíduos - Série Telebrás: Abordagem Bayesiana com distribuição a priori informativa	56
Figura 5.27: Volatilidade Estimada - Série Telebrás	57
Figura 5.28: Modelo AR(1)-ARCH(3) com $\alpha_0 = 0,05, \alpha_1 = 0,40, \alpha_2 = 0,25, \alpha_3 = 0,10$ e $\beta_1 = 0,50$	58
Figura 5.29: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t - Modelo AR(1)-ARCH(3)	59
Figura 5.30: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t - Modelo AR(1)-ARCH(3)	59
Figura 5.31: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t^2 - Modelo AR(1)-ARCH(3)	60
Figura 5.32: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Modelo AR(1)-ARCH(3)	61
Figura 5.33: Valores gerados dos parâmetros considerando distribuição a priori informativa - Modelo AR(1)-ARCH(3)	62
Figura 5.34: Análise de resíduos - Modelo AR(1)-ARCH(3): Abordagem Bayesiana com distribuição a priori informativa	62
Figura 5.35: Volatilidade Estimada - Modelo AR(1)-ARCH(3)	63
Figura 5.36: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Modelo AR(1)-ARCH(3)	64
Figura 5.37: Valores gerados dos parâmetros considerando distribuição a priori não-informativa - Modelo AR(1)-ARCH(3)	64
Figura 5.38: Análise de resíduos - Modelo AR(1)-ARCH(3)	65
Figura 5.39: Volatilidade Estimada - Modelo AR(1)-ARCH(3)	65
Figura 5.40: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t - Série IBovespa	66
Figura 5.41: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t^2 - Série IBovespa	66
Figura 5.42: Distribuição Empírica dos EMV - Série IBovespa	68
Figura 5.43: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Série IBovespa	70
Figura 5.44: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série IBovespa	70

Figura 5.44 (cont.): Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série IBovespa	71
Figura 5.45: Análise de resíduos - Série IBovespa: Abordagem Bayesiana com distribuição a priori informativa	72
Figura 5.46: Volatilidade Estimada - Série IBovespa	72
Figura 5.47: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t - Série Telebrás	73
Figura 5.48: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t^2 - Série Telebrás	74
Figura 5.49: Distribuição Empírica dos EMV - Série Telebrás	75
Figura 5.50: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Série Telebrás	76
Figura 5.51: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série Telebrás	77
Figura 5.52: Análise de resíduos - Série Telebrás: Abordagem Bayesiana com distribuição a priori informativa	78
Figura 5.53: Volatilidade Estimada - Série Telebrás	78
Figura 6.1: Histograma das Estimativas de MV - Modelo M1	83
Figura 6.2: Histograma das Estimativas Bayesianas - Modelo M1	83
Figura 6.3: Histograma das Estimativas de MV - Modelo ARCH(5)	85
Figura 6.4: Histograma das Estimativas Bayesianas - Modelo ARCH(5)	86

CAPÍTULO 1

Introdução

1.1. Introdução e Revisão Bibliográfica

A análise gráfica de séries financeiras revela que estas apresentam uma elevada taxa de mudança da variância em alguns períodos de tempo. A raiz quadrada desta taxa (desvio-padrão) é chamada de *volatilidade*. Entender como a volatilidade muda com o tempo é fundamental para o mercado financeiro, influenciando na avaliação do risco de investimentos e na apreciação de ativos. Ela determina o grau de variação do preço do ativo no futuro, ou seja, um baixo valor da mesma implica em pequenas alterações no futuro (baixo risco), enquanto que um alto valor implica em variações significativas (alto risco).

Seja p_t o preço de um determinado ativo no instante t , normalmente um dia de negócio. Suponha, primeiramente, que não haja dividendos pagos no período. A variação de preços entre o instantes $t-1$ e t é dada por $\Delta p_t = p_t - p_{t-1}$ e a variação relativa de preços ou *retorno líquido simples* deste ativo, entre os mesmos instantes, é definido por:

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} = \frac{\Delta p_t}{p_{t-1}} \quad (1.1)$$

Podemos notar que $r_t = \frac{p_t}{p_{t-1}} - 1$. O retorno $1 + r_t = \frac{p_t}{p_{t-1}}$ é chamado de *retorno bruto simples*. Usualmente r_t é expresso em porcentagem, relativamente ao período (um dia, um mês, um ano, etc.), sendo também chamado de *taxa de retorno*.

Denotando $P_t = \ln p_t$, temos o *retorno composto continuamente* ou simplesmente *log-retorno*, dado por:

$$R_t = \ln\left(\frac{p_t}{p_{t-1}}\right) = \ln(1 + r_t) = P_t - P_{t-1} \quad (1.2)$$

Esta definição é comumente utilizada e, neste trabalho, R_t será chamado simplesmente de *retorno* (Morettin & Toloï, 2004). Podemos observar que $R_t = \Delta \ln(p_t)$, ou seja, tomamos o logaritmo dos preços e depois a primeira diferença. Na prática é preferível trabalhar com retornos, que são livres de escala, do que com preços, pois os primeiros têm propriedades estatísticas mais interessantes (como estacionariedade e ergodicidade).

A caracterização das propriedades estatísticas das séries de retornos é essencial para uma correta aplicação de modelos aos dados, que permitam inferir a respeito das características deste retorno, sobretudo no que diz respeito à média e à variância, que vão determinar o retorno esperado e a previsão da volatilidade para os próximos períodos. Esta determinação é fundamental para a decisão de investimentos não apenas no ativo como também em seus derivativos.

Em geral, os retornos são não-autocorrelacionados e os quadrados dos mesmos autocorrelacionados; apresentam agrupamentos de volatilidade ao longo do tempo; a distribuição não-condicional destas séries tem caudas mais pesadas que as de uma distribuição normal e, embora seja aproximadamente simétrica, em geral, a distribuição é leptocúrtica. Além disso, algumas séries de retornos são não-lineares, ou seja, respondem de maneira diferente a choques grandes ou pequenos, ou ainda, choques positivos ou negativos (Morettin & Toloï, 2004).

Existe uma ampla variedade de modelos não-lineares para a estimação da volatilidade de séries de retornos de ativos financeiros, e os mais difundidos na literatura são os modelos auto-regressivos com heterocedasticidade (ARCH), propostos por Engle (1982), e sua extensão, os modelos ARCH Generalizados (GARCH), propostos por Bollerslev (1986). Tais modelos caracterizam uma dependência não-linear entre os retornos, função da dependência serial da variância condicional. Posteriormente, outros modelos discretos foram desenvolvidos, dentre eles, os modelos IGARCH, introduzidos por Engle & Bollerslev (1986), em que se considera que a variância não-condicional da série é infinita; os modelos

ARCH-M apresentados por Engle et al (1987), nos quais a esperança condicional do retorno é uma função explícita da variância condicional; e os modelos EGARCH propostos por Nelson (1991) que, ao contrário dos modelos GARCH, levam em conta não apenas o tamanho das séries de retornos, mas também o fato de serem positivos ou negativos. Uma ampla revisão das propriedades destes modelos pode ser encontrada em Bollerslev et al (1992) e Karlis and Xekalaki (2003).

Diversos trabalhos sobre a modelagem de séries temporais financeiras em países emergentes e, em particular no Brasil, podem ser encontrados em revistas especializadas. Dentre eles destacamos, Duarte et al (1996) que comparam modelos de volatilidade de preço e volatilidade implícita; Issler (1999) que compara os resultados obtidos usando modelos da família ARCH em séries de retornos de ativos de diferentes tipos (ação, bônus, moeda e *commodity*); Pereira et al (1999) que comparam modelos da família GARCH, volatilidade estocástica e GARCH com parâmetros variantes no tempo; Costa & Baidya (2001) que analisam as propriedades estatísticas das principais séries de retornos de ações brasileiras; Silveira et al (2001) que propõem um método alternativo à família GARCH na estimação da volatilidade destas séries.

Destacamos ainda alguns trabalhos cujo enfoque está em modelos de regressão com erros ARCH: Weiss (1984) deriva as distribuições assintóticas dos estimadores de Máxima Verossimilhança (EMV) e dos estimadores de mínimos quadrados (EMQ) de modelos ARMA com erros ARCH e Pantula (1988) considera processos AR com erros ARCH (AR-ARCH) mostrando a consistência e a normalidade assintótica dos EMV e apresentando as distribuições assintóticas dos EMQ e dos EMQ Generalizados. Maecker (1991) constrói estimadores eficientes para o parâmetro auto-regressivo e considera a estimação de parâmetros em processos AR(1) com erros ARCH(1), sob a suposição de densidade simétrica para os erros. Dutta (1999) constrói as estatísticas do teste Rao e do teste Wald para testar parâmetros de regressão de modelos com erros ARCH e Ha & Lee (2002) consideram um teste de constância do coeficiente de modelos AR-ARCH.

Sob um enfoque Bayesiano, um dos primeiros trabalhos propostos foi o de Geweke (1989) para modelos da família ARCH, em que um caso particular de reparametrização permitiu a utilização de distribuições a priori não-informativas. As estimativas dos parâmetros foram obtidas usando algoritmos de simulação de Monte Carlo. Posteriormente, uma abordagem semi-paramétrica para modelos ARCH foi sugerida por Koop (1994) que usou as

vantagens da metodologia Bayesiana de Geweke (1989). Nakatsuma & Tsurumi (1996) propõem uma abordagem Bayesiana e a comparam com a de Máxima Verossimilhança para modelos ARMA-GARCH. No contexto de modelos de componentes não observados, Giakoumatos et al (1998) propõem uma abordagem Bayesiana para os modelos ARCH usando uma amostragem de variáveis auxiliares. Mais recentemente, uma abordagem Bayesiana foi proposta para modelos GARCH em Migon & Mazucheli (1999), dentro da classe de modelos dinâmicos. Nakatsuma (2000) utiliza distribuições a priori normais para os parâmetros de modelos ARMA-GARCH e o algoritmo Metropolis-Hastings na determinação de sumários a posteriori. Polasek & Kozumi (2000) propõem uma abordagem Bayesiana para modelos VAR-VARCH e Polasek (2001) sugere uma estrutura hierárquica, também sob um enfoque Bayesiano, para modelos PAR-ARCH, usando métodos de simulação de Monte Carlo em cadeias de Markov (MCMC), no entanto, afirma que o esforço computacional desta modelagem deve ser avaliado.

1.2. Motivação

Como é considerado que a volatilidade em um dado instante de tempo depende dos valores passados da série, a determinação de estimadores de Máxima Verossimilhança dos parâmetros de modelos da família ARCH requer a maximização de uma função não-linear. Portanto, as estimativas só podem ser obtidas numericamente. Engle (1982) sugere o uso do método de Newton como um método iterativo para o cálculo das estimativas de Máxima Verossimilhança. Este procedimento relaxa as restrições impostas aos parâmetros, que asseguram estacionariedade na covariância. Por outro lado, a determinação de estimadores de Máxima Verossimilhança, com tais restrições, envolve muitas dificuldades (Geweke, 1986a).

Além disso, procedimentos para identificação, ajuste e diagnóstico dos modelos, assim como previsão de valores de séries econométricas, necessitam de propriedades da teoria assintótica. Como os modelos estão muito distantes da linearidade, as propriedades assintóticas destes estimadores só se verificam para séries muito longas e, em geral, são mais apropriadas na presença de distribuição simétrica para os erros e de distribuição normal para os dados. Uma alternativa para fazer a estimação destes modelos, contornando tais dificuldades, é considerar uma abordagem Bayesiana (Apêndice A.3).

Na seção 1.1. foram citados alguns trabalhos desenvolvidos na literatura, os quais utilizam a metodologia Bayesiana na estimação de parâmetros de modelos da família ARCH. Neste trabalho apresentamos um estudo de dois modelos com heterocedasticidade que têm por objetivo descrever a volatilidade existente em séries financeiras. Neste estudo mostramos os procedimentos para a estimação e para a seleção da ordem de modelos ARCH e de modelos AR-ARCH. As estimativas são obtidas utilizando a abordagem de Máxima Verossimilhança e a abordagem Bayesiana. Na abordagem de Máxima Verossimilhança obtivemos intervalos de confiança da forma padrão e usando o método de simulação Bootstrap (Apêndice A.2), e na abordagem Bayesiana, propomos uma reparametrização dos modelos estudados e adotamos distribuição a priori informativa normal para os parâmetros transformados. Além disso, consideramos a distribuição a priori não-informativa de Geweke (1989) similarmente reparametrizada, e usamos métodos de simulação MCMC para a obtenção dos resumos a posteriori.

1.3. Objetivos

O objetivo deste trabalho é mostrar que a abordagem Bayesiana proposta para a estimação de parâmetros de modelos ARCH e de modelos AR-ARCH é uma alternativa viável e mais robusta que a abordagem tradicional de Máxima Verossimilhança. Especificamente, temos o intuito de avaliar o desempenho da abordagem Bayesiana na estimação de parâmetros de modelos ARCH e de modelos AR-ARCH, observando o efeito do aumento da ordem dos modelos na qualidade das estimativas dos parâmetros destes processos e a estabilidade na convergência dos algoritmos de simulação MCMC. A metodologia é exemplificada considerando séries simuladas e séries do mercado financeiro brasileiro.

1.4. Organização da tese

Os demais capítulos deste trabalho estão estruturados da seguinte forma:

No capítulo 2 apresentamos as principais propriedades teóricas associadas aos modelos AR-ARCH e aos modelos ARCH.

No capítulo 3 apresentamos uma abordagem Bayesiana para os modelos AR-ARCH e para os modelos ARCH, com as distribuições a priori propostas e respectivas distribuições a posteriori obtidas.

No capítulo 4 apresentamos os algoritmos desenvolvidos para calcular as estimativas dos parâmetros dos modelos usando a metodologia estudada neste trabalho.

No capítulo 5 mostramos uma comparação das estimativas de parâmetros de modelos ARCH e de modelos AR-ARCH obtidas através da abordagem Bayesiana, com as estimativas obtidas através da abordagem de Máxima Verossimilhança, no estudo de séries simuladas e de séries do mercado financeiro brasileiro.

No capítulo 6 apresentamos uma comparação das estimativas de parâmetros de modelos ARCH obtidas através da abordagem Bayesiana (usando distribuição a priori informativa) com as obtidas através da abordagem de Máxima Verossimilhança, avaliando a raiz quadrada do erro quadrático médio e do erro percentual absoluto médio entre os valores reais dos parâmetros e os valores estimados por cada uma das técnicas.

No último capítulo apresentamos as conclusões obtidas neste trabalho e algumas propostas que poderão ser desenvolvidas posteriormente a fim de darem continuidade ao estudo aqui desenvolvido.

Um apêndice foi introduzido com o objetivo de descrever pré-requisitos básicos para o entendimento deste trabalho e, de apresentar alguns resultados e figuras complementares aos que constam nos capítulos 5 e 6.

Os programas desenvolvidos para a obtenção dos resultados dos exemplos de aplicação dos modelos propostos foram implementados no software MATLAB - versão 5.3 e podem ser obtidos diretamente com os autores deste trabalho.

CAPÍTULO 2

Modelos Auto-Regressivos com Heterocedasticidade

Nesta seção, apresentamos as principais propriedades teóricas associadas aos modelos AR-ARCH e aos modelos ARCH, bem como as estimativas de Máxima Verossimilhança para os parâmetros dos mesmos.

2.1. Modelos AR(p)-ARCH(q)

2.1.1. Introdução

O modelo de regressão, proposto por Engle (1982), com uma média não-nula e expresso como uma combinação linear de variáveis exógenas, tem uma estrutura que pode ser resumida como:

$$y_t = \mathbf{x}_t \boldsymbol{\beta} + z_t \quad (2.1)$$

$$y_t | \Omega_{t-1} \sim P(\mathbf{x}_t \boldsymbol{\beta}, h_t) \quad (2.2)$$

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j z_{t-j}^2 \quad (2.3)$$

$$z_t = y_t - \mathbf{x}_t \boldsymbol{\beta} \quad (2.4)$$

em que y_t representa uma série de retornos, $P(\cdot)$ é uma distribuição paramétrica, usualmente a Normal ou t -Student, \mathbf{x}_t representa um vetor de variáveis exógenas que pode incluir valores de y_t defasados no tempo, $\boldsymbol{\beta}$ é um vetor de parâmetros desconhecidos e Ω_{t-1} representa o conjunto de informações disponíveis até o instante $t-1$. Neste trabalho vamos considerar somente os modelos AR(p)-ARCH(q) tais que $\mathbf{x}_t = (y_{t-1}, y_{t-2}, \dots, y_{t-p})'$. Assim, temos:

$$y_t = \sum_{i=1}^p \beta_i y_{t-i} + z_t \quad (2.5)$$

Seja z_t um erro que satisfaz o modelo

$$z_t = h_t^{1/2} \varepsilon_t \quad (2.6)$$

em que $\{\varepsilon_t, t \geq 0\}$ é uma seqüência de ruído branco *i.i.d.* $N(0,1)$ independente de \mathbf{x}_t .

A distribuição condicional para a taxa de retorno utilizada neste trabalho é a normal que, apesar de ser uma representação bastante simplificada do processo gerador de dados dos retornos condicionais, é amplamente utilizada na estimação dos modelos de volatilidade (por ex.: Ferreira, 2001; Oliveira et al, 2003). Portanto, a distribuição condicionada de y_t é normal com média $\mathbf{x}_t \boldsymbol{\beta}$ e variância h_t , ou seja, $y_t | \Omega_{t-1} \sim N(\mathbf{x}_t \boldsymbol{\beta}, h_t)$. Uma interpretação para o modelo em estudo é que os distúrbios na regressão linear seguem um processo ARCH(q).

Para que o modelo (2.1)-(2.4) seja plausível ($h_t > 0$ para todo t), devemos ter $\alpha_0 > 0$ e $\alpha_j \geq 0, j = 1, \dots, q$. Além disso, o processo y_t tem variância finita e, portanto, tem covariância estacionária se, e somente se, todas as raízes dos polinômios $1 - \sum_{i=1}^p \beta_i l^i$ e $1 - \sum_{j=1}^q \alpha_j l^j$ estiverem fora do círculo de raio unitário (Engle, 1982: Teorema 2). Quando estas condições são satisfeitas, pode ser mostrado que a variância incondicional de y_t é dada por:

$$\frac{\alpha_0}{\left(1 - \sum_{j=1}^q \alpha_j\right) \left(1 - \sum_{i=1}^p \beta_i \rho_i\right)}$$

em que $\rho_i = \frac{\gamma_i}{\gamma_0}$, sendo $\gamma_0 = V(y_t) = E(y_t^2)$ e $\gamma_i = E(y_t y_{t-i})$, $i = 1, \dots, p$. Portanto, a condição

suficiente para que o processo tenha covariância estacionária é que $\sum_{j=1}^q \alpha_j < 1$ e $\sum_{i=1}^p |\beta_i| < 1$.

2.1.2. Função de Verossimilhança

Consideremos uma trajetória observada $\mathbf{Y} = \{y_t, t = 1, 2, \dots, T\}$ do processo y_t e que \mathbf{x}_t envolve somente os “ p ” valores passados de y_t . Então, a função de verossimilhança de y_t , $t = p + q + 1, \dots, T$, condicionada às $p + q$ primeiras observações (assumindo que elas são conhecidas) pode ser escrita como:

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta} | y_1, y_2, \dots, y_{p+q}) \propto \prod_{t=p+q+1}^T \left(\frac{1}{h_t} \right)^{1/2} \exp \left\{ - \frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t} \right\} \quad (2.7)$$

sendo $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$ e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$. Assim, a função de verossimilhança pode ser maximizada com respeito aos parâmetros desconhecidos $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$. Para simplificar a notação, em vez de escrevermos $L(\boldsymbol{\alpha}, \boldsymbol{\beta} | y_1, y_2, \dots, y_{p+q})$ para representar a função de verossimilhança condicionada às $p + q$ primeiras observações, escreveremos somente $L(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Destacamos ainda que, o uso da função de verossimilhança condicionada no lugar da função de verossimilhança exata, pode ser feito sem grandes perdas de precisão nas estimativas, quando a série y_t é grande. Esta aproximação geralmente é considerada devido às grandes vantagens práticas no cálculo dos estimadores (Box et al, 1994).

Denotamos a média do logaritmo natural da função de verossimilhança por:

$$l(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{T} \sum_{t=p+q+1}^T l_t(\boldsymbol{\alpha}, \boldsymbol{\beta}) \quad (2.8)$$

em que $l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})$ indica o logaritmo natural da verossimilhança da t -ésima observação. Usando as expressões (2.1-2.4), temos:

$$l_t(\boldsymbol{\alpha}, \boldsymbol{\beta}) = -\frac{1}{2} \log h_t - \frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t} \quad (2.9)$$

O gradiente da função $l(\boldsymbol{\alpha}, \boldsymbol{\beta})$ com respeito a $\boldsymbol{\alpha}$ é dado por:

$$\frac{\partial l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\alpha}} = \frac{1}{T} \sum_{t=p+q+1}^T \frac{\partial l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\alpha}} \quad (2.10)$$

tal que $\frac{\partial l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\alpha}}$ é um vetor cujos componentes são:

$$\frac{\partial l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_j} = \frac{1}{2h_t} \left(\frac{\partial h_t}{\partial \alpha_j} \right) \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \right), \quad j = 0, 1, \dots, q \quad (2.11)$$

Da equação (2.3) temos:

$$\frac{\partial h_t}{\partial \alpha_j} = \begin{cases} 1 & j = 0 \\ z_{t-j}^2 & j = 1, 2, \dots, q, \quad t \geq p + q + 1 \end{cases} \quad (2.12)$$

Portanto,

$$\frac{\partial l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\alpha}} = \frac{1}{2h_t} \mathbf{v}_t \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \right) \quad (2.13)$$

em que $\mathbf{v}_t = (1, z_{t-1}^2, \dots, z_{t-q}^2)$, sendo $z_t = y_t - \mathbf{x}_t \boldsymbol{\beta}$ e $t = p + q + 1, \dots, T$. Assim, o gradiente de $l(\boldsymbol{\alpha}, \boldsymbol{\beta})$ com respeito a $\boldsymbol{\alpha}$ pode ser calculado pela substituição da expressão (2.13) em (2.10), resultando em:

$$\frac{\partial l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\alpha}} = \frac{1}{T} \sum_{t=p+q+1}^T \frac{1}{2h_t} \mathbf{v}_t \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \right) \quad (2.14)$$

Simbolizando,

$$\tilde{\mathbf{v}}_t = \frac{(1, z_{t-1}^2, \dots, z_{t-q}^2)}{h_t} \quad t = p + q + 1, \dots, T \quad (2.15)$$

$$\tilde{\mathbf{v}}' = (\tilde{\mathbf{v}}'_{p+q+1}, \dots, \tilde{\mathbf{v}}'_T)$$

$$f_t = \left(\frac{z_t^2}{h_t} - 1 \right) \quad t = p + q + 1, \dots, T \quad (2.16)$$

$$\mathbf{f}' = (f'_{p+q+1}, \dots, f'_T)$$

tal que $z_t = y_t - \mathbf{x}_t \boldsymbol{\beta}$, podemos escrever $\frac{\partial l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\alpha}}$ na forma matricial como:

$$\frac{\partial l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\alpha}} = \frac{1}{2T} \tilde{\mathbf{v}}' \mathbf{f} \quad (2.17)$$

e, denotando a Hessiana de $l(\boldsymbol{\alpha}, \boldsymbol{\beta})$ com respeito a $\boldsymbol{\alpha}$ por $H_\alpha = \left\{ \frac{\partial^2 l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_i \partial \alpha_j}, i, j = 0, 1, \dots, q \right\}$

temos:

$$\frac{\partial^2 l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_i \partial \alpha_j} = \frac{1}{T} \sum_{t=p+q+1}^T \frac{\partial^2 l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_i \partial \alpha_j} \quad (2.18)$$

tal que

$$\begin{aligned} \frac{\partial^2 l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_i \partial \alpha_j} &= \frac{\partial}{\partial \alpha_i} \left[\frac{1}{2h_t} \frac{\partial h_t}{\partial \alpha_j} \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \right) \right] = \\ &= -\frac{1}{2h_t^2} \left(\frac{\partial h_t}{\partial \alpha_i} \right) \left(\frac{\partial h_t}{\partial \alpha_j} \right) \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \right) + \frac{1}{2h_t} \frac{\partial^2 h_t}{\partial \alpha_i \partial \alpha_j} \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \right) + \frac{1}{2h_t} \left(\frac{\partial h_t}{\partial \alpha_j} \right) \left(-\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t^2} \frac{\partial h_t}{\partial \alpha_i} \right) \end{aligned} \quad (2.19)$$

Da equação (2.3) temos que $\frac{\partial^2 h_t}{\partial \alpha_i \partial \alpha_j} = 0$. Assim, a expressão (2.19) pode ser

simplificada para:

$$\frac{\partial^2 l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_i \partial \alpha_j} = -\frac{1}{2h_t^2} \left(\frac{\partial h_t}{\partial \alpha_i} \right) \left(\frac{\partial h_t}{\partial \alpha_j} \right) \left(2 \frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \right) \quad (2.20)$$

A matriz de informação $I_{\alpha\alpha} = \{I_{ij}\}, i, j = 0, 1, \dots, q$ tem seus elementos definidos como:

$$I_{ij} = -E \left[\frac{1}{T} \sum_{t=p+q+1}^T \frac{\partial^2 l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_i \partial \alpha_j} \right] \quad i, j = 0, 1, \dots, q \quad (2.21)$$

Tomando as esperanças condicionais da Hessiana H_α e, sendo

$$E \left[2 \frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \mid \Omega_{t-1} \right] = 2 \frac{E[(y_t - \mathbf{x}_t \boldsymbol{\beta})^2 \mid \Omega_{t-1}]}{h_t} - 1 = 1, \quad \text{os elementos da matriz de}$$

informação $\{I_{ij}\}, i, j = 0, 1, \dots, q$ podem ser consistentemente estimados por:

$$\hat{I}_{ij} = \frac{1}{T} \sum_{t=p+q+1}^T \frac{1}{2h_t^2} \left\{ \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \alpha_j} \right\} \quad i, j = 0, 1, \dots, q \quad (2.22)$$

Assim, pelas expressões definidas em (2.15), podemos escrever a matriz de informação como:

$$\hat{I}_{\alpha\alpha} = \frac{1}{2T} \sum_{t=p+q+1}^T \tilde{\mathbf{v}}_t' \tilde{\mathbf{v}}_t = \frac{1}{2T} \tilde{\mathbf{v}}' \tilde{\mathbf{v}} \quad (2.23)$$

Em relação ao gradiente da função $l(\boldsymbol{\alpha}, \boldsymbol{\beta})$ com respeito a $\boldsymbol{\beta}$, temos que:

$$\frac{\partial l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{1}{T} \sum_{t=p+q+1}^T \frac{\partial l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \quad (2.24)$$

tal que $\frac{\partial l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$ é um vetor cujos componentes são:

$$\frac{\partial l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_i} = \frac{(y_t - \mathbf{x}_t \boldsymbol{\beta}_i) \mathbf{x}_t'}{h_t} + \frac{1}{2h_t} \left(\frac{\partial h_t}{\partial \beta_i} \right) \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta}_i)^2}{h_t} - 1 \right) \quad i=1, \dots, p \quad (2.25)$$

O primeiro termo da equação (2.25) é a condição de primeira ordem para uma correção heterocedástica exógena e o segundo termo resulta do fato de h_t também ser uma função de $\boldsymbol{\beta}$. Assim,

$$\frac{\partial l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{1}{T} \sum_{t=p+q+1}^T \left[\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta}) \mathbf{x}_t'}{h_t} - \frac{1}{h_t} \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \right) \sum_{j=1}^q \alpha_j (y_{t-j} - \mathbf{x}_{t-j} \boldsymbol{\beta}) \mathbf{x}_{t-j}' \right] \quad (2.26)$$

que pode ser reescrito como (Engle, 1982):

$$\frac{\partial l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{1}{T} \sum_{t=1}^{T-p-1} \mathbf{x}_t' (y_t - \mathbf{x}_t \boldsymbol{\beta}) \left[\frac{1}{h_t} - \sum_{j=1}^q \frac{\alpha_j}{h_{t+j}^2} [(y_{t+j} - \mathbf{x}_{t+j} \boldsymbol{\beta})^2 - h_{t+j}] \right] \quad (2.27)$$

Fazendo $s_t = \frac{1}{h_t} - \sum_{j=1}^q \frac{\alpha_j}{h_{t+j}^2} [(y_{t+j} - \mathbf{x}_{t+j} \boldsymbol{\beta})^2 - h_{t+j}]$, o gradiente da função $l(\boldsymbol{\alpha}, \boldsymbol{\beta})$ em

relação a $\boldsymbol{\beta}$ é dado por:

$$\frac{\partial l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{1}{T} \sum_{t=1}^{T-p-1} \mathbf{x}_t' (y_t - \mathbf{x}_t \boldsymbol{\beta}) s_t \quad (2.28)$$

e, denotando a Hessiana de $l(\boldsymbol{\alpha}, \boldsymbol{\beta})$ com respeito a $\boldsymbol{\beta}$ por $H_\beta = \left\{ \frac{\partial^2 l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_i \partial \beta_j}, i, j = 1, \dots, p \right\}$

temos:

$$\frac{\partial^2 l(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_i \partial \beta_j} = \frac{1}{T} \sum_{t=1}^{T-p-1} \frac{\partial^2 l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_i \partial \beta_j} \quad (2.29)$$

tal que

$$\frac{\partial^2 l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_i \partial \beta_j} = -\frac{\mathbf{x}'_t \mathbf{x}_t}{h_t} - \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \beta_i} \frac{\partial h_t}{\partial \beta_j} \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} \right) - \frac{2(y_t - \mathbf{x}_t \boldsymbol{\beta}) \mathbf{x}'_t}{h_t^2} \frac{\partial h_t}{\partial \beta_i} + \left(\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \beta_j} \left[\frac{1}{2h_t} \frac{\partial h_t}{\partial \beta_i} \right] \quad (2.30)$$

A matriz de informação $I_{\beta\beta} = \{I_{ij}\}, i, j = 1, \dots, p$ tem seus elementos definidos como:

$$I_{ij} = -E \left[\frac{1}{T} \sum_{t=1}^{T-p-1} \frac{\partial^2 l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_i \partial \beta_j} \right] \quad i, j = 1, \dots, p \quad (2.31)$$

Tomando as esperanças condicionais da Hessiana H_β , os dois últimos termos da equação (2.30) desaparecem, pois h_t é uma função inteiramente do passado e, portanto,

independente de $y_t - \mathbf{x}_t \boldsymbol{\beta}$. Logo, temos que $E \left[\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t} \mid \Omega_{t-1} \right] = \frac{E \left[(y_t - \mathbf{x}_t \boldsymbol{\beta})^2 \mid \Omega_{t-1} \right]}{h_t} = 1$.

Segundo Engle (1982), esta afirmação é válida mesmo que \mathbf{x}_t inclua variáveis exógenas com valores de y_t defasados no tempo. Assim, os elementos da matriz de informação $\{I_{ij}\}, i, j = 1, \dots, p$ podem ser consistentemente estimados por:

$$\hat{I}_{ij} = \frac{1}{T} \sum_{t=1}^{T-p-1} E \left[\frac{\mathbf{x}'_t \mathbf{x}_t}{h_t} + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \beta_j} \frac{\partial h_t}{\partial \beta_i} \right] \quad i, j = 1, \dots, p \quad (2.32)$$

Conseqüentemente, temos:

$$\hat{I}_{\beta\beta} = \frac{1}{T} \sum_{t=p+q+1}^T \left[\frac{\mathbf{x}'_t \mathbf{x}_t}{h_t} + 2 \sum_{j=1}^q \alpha_j^2 \frac{(y_{t-j} - \mathbf{x}_{t-j} \boldsymbol{\beta})^2}{h_t^2} \mathbf{x}'_{t-j} \mathbf{x}_{t-j} \right] \quad (2.33)$$

cujo termo $\mathbf{x}'_t \mathbf{x}_t$ pode ser posto em evidência resultando em:

$$\hat{I}_{\beta\beta} = \frac{1}{T} \sum_{t=1}^{T-p-1} \mathbf{x}'_t \mathbf{x}_t \left[\frac{1}{h_t} + 2(y_t - \mathbf{x}_t \boldsymbol{\beta})^2 \sum_{j=1}^q \frac{\alpha_j^2}{h_{t+j}^2} \right] \quad (2.34)$$

Assim, podemos reescrever a equação (2.34), definindo a matriz de informação como:

$$\hat{I}_{\beta\beta} = \frac{1}{T} \sum_{t=1}^{T-p-1} \mathbf{x}'_t \mathbf{x}_t r_t^2 \quad (2.35)$$

sendo $r_t^2 = \frac{1}{h_t} + 2(y_t - \mathbf{x}_t \boldsymbol{\beta})^2 \sum_{j=1}^q \frac{\alpha_j^2}{h_{t+j}^2}$.

De forma similar, os blocos fora da diagonal da matriz de informação podem ser expressos como:

$$\begin{aligned} I_{\alpha\beta} &= -E \left[\frac{1}{T} \sum_{t=p+q+1}^T \frac{\partial^2 l_t(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_i \partial \beta_j} \right] \\ &= \frac{1}{T} \sum_{t=p+q+1}^T E \left[\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \beta_j} \right] \end{aligned} \quad (2.36)$$

para $i = 0, 1, \dots, q$ e $j = 1, \dots, p$.

Engle (1982: *Teorema 4*) mostra que se um modelo de regressão ARCH(q) é simétrico e regular, então, $I_{\alpha\beta} = 0$, ou seja, o bloco fora da diagonal da matriz de informação é zero. Portanto, a estimação dos parâmetros $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$ pode ser feita separadamente sem perda da eficiência assintótica, e as variâncias também podem ser calculadas separadamente.

Como os estimadores de Máxima Verossimilhança de modelos AR(p)-ARCH(q) não têm expressões explícitas, devemos estimá-los por procedimentos iterativos. Estes procedimentos são apresentados no Capítulo 4.

Engle (1982) considera que, sob as condições de Crowder (1976), os estimadores de Máxima Verossimilhança $\hat{\boldsymbol{\alpha}}$ e $\hat{\boldsymbol{\beta}}$ são assintoticamente distribuídos, com distribuições limitantes:

$$\sqrt{T}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \xrightarrow{D} N(0, I_{\alpha\alpha}^{-1}) \quad (2.37)$$

$$\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{D} N(0, I_{\beta\beta}^{-1})$$

Weiss (1984) e Pantula (1988) apresentam formalmente a consistência e a normalidade assintótica destes estimadores.

2.2. Modelos ARCH(q)

2.2.1. Introdução

Se no modelo de regressão proposto por Engle (1982), definido pelas expressões (2.1)-(2.4), considerarmos $\beta = \mathbf{0}$, então, temos uma nova estrutura que pode ser resumida como:

$$y_t = z_t \quad (2.38)$$

$$z_t | \Omega_{t-1} \sim P(0, h_t) \quad (2.39)$$

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j z_{t-j}^2 \quad (2.40)$$

em que y_t representa uma série de retornos, $P(\cdot)$ é uma distribuição paramétrica, usualmente a Normal ou t -Student e Ω_{t-1} representa o conjunto de informações disponíveis até o instante $t-1$. Seja z_t um processo que satisfaz o modelo

$$z_t = h_t^{1/2} \varepsilon_t \quad (2.41)$$

tal que $\{\varepsilon_t, t \geq 0\}$ é uma seqüência de ruído branco *i.i.d.* $N(0,1)$. Uma interpretação para o modelo definido em (2.38)-(2.41) é que os retornos na regressão linear seguem um processo ARCH(q).

Da mesma forma que foi definido na seção 2.1., para que o modelo (2.38)-(2.40) seja plausível ($h_t > 0$ para todo t), devemos ter $\alpha_0 > 0$ e $\alpha_j \geq 0$ para $j = 1, \dots, q$. Além disso, o processo z_t tem variância finita e, portanto, tem covariância estacionária se, e somente se, todas as raízes do polinômio $1 - \sum_{j=1}^q \alpha_j l^j$ estiverem fora do círculo de raio unitário, sendo l o

operador translação. Quando esta condição é satisfeita, pode ser mostrado que a variância incondicional de z_t é dada por $\frac{\alpha_0}{\left(1 - \sum_{j=1}^q \alpha_j\right)}$. Portanto, a condição suficiente para que o

processo tenha covariância estacionária é que $\sum_{j=1}^q \alpha_j < 1$.

Algumas representações alternativas de (2.40) podem ser utilizadas para a identificação da ordem q do modelo ARCH. Para entender melhor a dinâmica do modelo ARCH(q), Harvey (1997) considerou a seguinte relação:

$$z_t^2 = h_t + (z_t^2 - h_t) \quad (2.42)$$

sendo $h_t = \alpha_0 + \sum_{j=1}^q \alpha_j z_{t-j}^2$ e $z_t = h_t^{1/2} \varepsilon_t$.

Portanto,

$$z_t^2 = \alpha_0 + \sum_{j=1}^q \alpha_j z_{t-j}^2 + (h_t \varepsilon_t^2 - h_t) \quad (2.43)$$

e, considerando $w_t = h_t (\varepsilon_t^2 - 1)$, a equação (2.43) pode ser reescrita como:

$$z_t^2 = \alpha_0 + \sum_{j=1}^q \alpha_j z_{t-j}^2 + w_t \quad (2.44)$$

Podemos notar que a equação (2.44) representa um modelo AR(q) para z_t^2 com um ruído w_t (Apêndice A.1). Assim, a identificação da ordem do modelo ARCH(q) pode ser baseada na função de auto-correlação (FAC) e auto-correlação parcial (FACP) do processo ao quadrado. No entanto, a utilização do modelo (2.44) para inferência dos parâmetros do modelo ARCH(q) é impraticável, uma vez que não se conhece a distribuição do ruído w_t , dificultando a obtenção da função de verossimilhança para este modelo.

2.2.2. Função de Verossimilhança

Dada uma trajetória observada $\mathbf{Z} = \{z_t, t = 1, 2, \dots, T\}$ do processo z_t , a função de verossimilhança de z_t , $t = q + 1, \dots, T$ condicionada às q primeiras observações (assumindo que elas são conhecidas) é dada por:

$$L(\boldsymbol{\alpha} | z_1, z_2, \dots, z_q) \propto \prod_{t=q+1}^T \left(\frac{1}{h_t} \right)^{1/2} \exp \left\{ -\frac{z_t^2}{2h_t} \right\} \quad (2.45)$$

sendo $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$. Assim, a função de verossimilhança pode ser maximizada com respeito ao parâmetro desconhecido $\boldsymbol{\alpha}$. Para simplificar a notação, em vez de escrevermos $L(\boldsymbol{\alpha} | z_1, z_2, \dots, z_q)$ para representar a função de verossimilhança condicionada às q primeiras observações, escreveremos somente $L(\boldsymbol{\alpha})$.

Denotamos a média do logaritmo natural da função de verossimilhança por:

$$l(\boldsymbol{\alpha}) = \frac{1}{T} \sum_{t=q+1}^T l_t(\boldsymbol{\alpha}) \quad (2.46)$$

em que $l_t(\boldsymbol{\alpha})$ indica o logaritmo da verossimilhança da t -ésima observação, ou seja,

$$l_t(\boldsymbol{\alpha}) = -\frac{1}{2} \log h_t - \frac{z_t^2}{2h_t} \quad (2.47)$$

O gradiente da função $l(\boldsymbol{\alpha})$ é dado por:

$$\frac{\partial l(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = \frac{1}{T} \sum_{t=q+1}^T \frac{\partial l_t(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \quad (2.48)$$

De maneira similar ao procedimento realizado na seção 2.1.2., temos que:

$$\frac{\partial l_t(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = \frac{1}{2h_t} \mathbf{v}_t \left(\frac{z_t^2}{h_t} - 1 \right) \quad (2.49)$$

sendo $\tilde{\mathbf{v}}_t = \frac{(1, z_{t-1}^2, \dots, z_{t-q}^2)}{h_t}$, $t = q+1, \dots, T$. Assim, substituindo a expressão (2.49) em (2.48), o

gradiente de $l(\boldsymbol{\alpha})$ resulta em:

$$\frac{\partial l(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = \frac{1}{T} \sum_{t=q+1}^T \frac{1}{2h_t} \mathbf{v}_t \left(\frac{z_t^2}{h_t} - 1 \right) \quad (2.50)$$

A matriz de informação $I = \{I_{ij}\}, i, j = 0, 1, \dots, q$ é definida como:

$$I_{ij} = -E \left[\frac{1}{T} \sum_{t=q+1}^T \frac{\partial^2 l_t(\boldsymbol{\alpha})}{\partial \alpha_i \partial \alpha_j} \right] \quad i, j = 0, 1, \dots, q \quad (2.51)$$

Usando a definição de $\tilde{\mathbf{v}}' = (\tilde{\mathbf{v}}'_{q+1}, \dots, \tilde{\mathbf{v}}'_T)$, podemos escrever a matriz de informação como:

$$\hat{I} = \frac{1}{2T} \sum_{t=q+1}^T \tilde{\mathbf{v}}'_t \tilde{\mathbf{v}}_t = \frac{1}{2T} \tilde{\mathbf{v}}' \tilde{\mathbf{v}} \quad (2.52)$$

Sob as condições descritas por Engle (1982), os estimadores de Máxima Verossimilhança $\hat{\boldsymbol{\alpha}}$ são assintoticamente normais com distribuição limitante:

$$\sqrt{T}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \xrightarrow{D} N(0, I^{-1}) \quad (2.53)$$

Uma ampla revisão das propriedades destes modelos pode ser encontrada em Bollerslev et al (1992).

CAPÍTULO 3

Inferência Bayesiana para Modelos AR(p)-ARCH(q) e ARCH(q)

No capítulo 2 vimos que os estimadores de Máxima Verossimilhança de modelos AR(p)-ARCH(q) e ARCH(q) não têm expressões explícitas e, portanto, devem ser estimados por procedimentos iterativos. É comum a ocorrência de problemas numéricos na abordagem de Máxima Verossimilhança destes modelos, causados principalmente pelo mau condicionamento da matriz de informação. Este fato causa imprecisões na inversão matricial requerida, particularmente para valores mais elevados da ordem dos modelos. Além disso, outra dificuldade nesta abordagem são as restrições de estacionariedade impostas aos parâmetros. Para contornar tais dificuldades, vamos propor a abordagem Bayesiana, uma vez que não há necessidade do uso de técnicas de otimização numérica. Uma vantagem a mais desta abordagem é a possibilidade de incorporar a experiência de especialistas na área de finanças, o que geralmente é uma questão relevante na análise de séries econômicas e financeiras.

3.1. Inferência Bayesiana para os modelos AR(p)-ARCH(q)

Consideremos uma trajetória observada de retornos $\mathbf{Y} = \{y_t, t = 1, 2, \dots, T\}$. A abordagem Bayesiana para a inferência dos parâmetros dos modelos AR(p)-ARCH(q) parte da combinação da função de verossimilhança desta trajetória, $L(\boldsymbol{\alpha}, \boldsymbol{\beta})$, com uma distribuição a priori para os parâmetros, $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta})$, através da regra de Bayes (Gelman et al, 1995; Apêndice A2) :

$$\pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y}) \propto L(\boldsymbol{\alpha}, \boldsymbol{\beta})\pi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \quad (3.1)$$

tal que $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ e $L(\boldsymbol{\alpha}, \boldsymbol{\beta})$ é a função definida pela equação (2.7), dada por:

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta} | y_1, y_2, \dots, y_{p+q}) \propto \prod_{t=p+q+1}^T \left(\frac{1}{h_t} \right)^{1/2} \exp \left\{ -\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t} \right\}$$

A expressão $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y})$ é chamada de distribuição a posteriori de $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$ e diz como as variáveis aleatórias $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$ estão distribuídas após os dados terem sido observados.

Como podemos observar, a análise Bayesiana requer a especificação de uma distribuição conjunta a priori para $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta})$, que reflete o conhecimento prévio sobre a distribuição de $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$. Nas seções 3.1.1 e 3.1.2 a seguir propomos distribuições a priori para a inferência dos parâmetros de modelos AR(p)-ARCH(q).

3.1.1. Distribuição a priori informativa

Neste trabalho consideramos $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \propto \pi(\boldsymbol{\beta} | \boldsymbol{\alpha})\pi(\boldsymbol{\alpha})$, que é uma estrutura hierárquica para distribuição a priori dos parâmetros $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$. Alguns autores têm sugerido estruturas hierárquicas nas últimas décadas e, para modelos de regressão, algumas aplicações podem ser encontradas na literatura (Polasek & Kozumi, 2000; Polasek, 2001 e Nakatsuma, 2000).

Quando temos alguma informação prévia sobre o objeto de estudo, podemos propor uma distribuição a priori informativa para os parâmetros do modelo. Neste trabalho, vamos supor que $\beta_i, i = 1, \dots, p$ são independentes, tais que:

$$\pi(\boldsymbol{\beta} | \boldsymbol{\alpha}) \propto N(\mathbf{b}, \boldsymbol{\Sigma}) \quad (3.2)$$

tal que

$$\boldsymbol{\Sigma} = \left(\frac{\alpha_0}{1 - \sum_{j=1}^q \alpha_j} \right) \begin{pmatrix} \xi_1 & & & \\ & \xi_2 & & \\ & & \ddots & \\ & & & \xi_p \end{pmatrix} = \begin{pmatrix} \frac{\alpha_0 \xi_1}{1 - \sum_{j=1}^q \alpha_j} & & & \\ & \frac{\alpha_0 \xi_2}{1 - \sum_{j=1}^q \alpha_j} & & \\ & & \ddots & \\ & & & \frac{\alpha_0 \xi_p}{1 - \sum_{j=1}^q \alpha_j} \end{pmatrix} \quad (3.3)$$

com os hiperparâmetros $\mathbf{b} = (b_1, \dots, b_p)'$ e $\xi_i, i = 1, \dots, p$ conhecidos (Polasek, 2001). A variância assumida para a distribuição a priori normal, $\pi(\boldsymbol{\beta} | \boldsymbol{\alpha})$, é proposta com base na informação que temos sobre a variância incondicional de um processo ARCH(q) associada a um termo $\xi_i, i = 1, \dots, p$, que pode ser definido a partir de informações disponíveis sobre $\boldsymbol{\beta}$.

Para a escolha de uma distribuição a priori para o vetor de parâmetros $\boldsymbol{\alpha}$, vamos considerar primeiramente que todos os $\alpha_j, j = 0, 1, \dots, q$ são independentes com uma distribuição a priori definida no intervalo $[a_j, b_j]$ com $a_j > 0$ e $b_j < 1$, tal que $a_j \leq \alpha_j \leq b_j$, visto que a restrição para modelos ARCH(q), $\sum_{j=1}^q \alpha_j < 1$, exige que $\alpha_j \in (0, 1), j = 1, 2, \dots, q$. Podemos definir ainda $a_0 \leq \alpha_0 \leq b_0$, com $a_0 > 0$ e $b_0 \leq E(z_t^2)$, $z_t = y_t - \mathbf{x}_t \boldsymbol{\beta}$. Uma vez que desejamos gerar valores para α_j próximos das médias de cada intervalo $\frac{(a_j + b_j)}{2}$ com mais frequência do que valores próximos dos limites a_j e b_j , adotamos uma reparametrização (Ferreira, 2001) dada por:

$$\phi_j = \log \left(\frac{\alpha_j - a_j}{b_j - \alpha_j} \right) \quad j = 0, 1, 2, \dots, q \quad (3.4)$$

em que ϕ_j é um componente do vetor de parâmetros $\boldsymbol{\varphi} = (\phi_0, \phi_1, \dots, \phi_q)'$ e a_j e b_j são escolhidos com base em alguma informação a priori, por exemplo, estudos anteriores sobre a série analisada. Esta reparametrização conduz à escolha de uma transformação que mapeia os intervalos $(-\infty, +\infty)$ no domínio (a_j, b_j) e vice-versa.

Então, podemos escrever a função de verossimilhança, expressa na equação (2.7), como:

$$L(\boldsymbol{\varphi}, \boldsymbol{\beta}) \propto \prod_{t=p+q+1}^T \left(\frac{1}{h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})} \right)^{1/2} \exp \left\{ -\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})} \right\} \quad (3.5)$$

sendo que, a volatilidade h_t é escrita em função dos parâmetros α_j devidamente transformados em ϕ_j , e de $\boldsymbol{\beta}$. Denotando por $\boldsymbol{\Sigma}(\boldsymbol{\varphi}, \boldsymbol{\beta})$ uma matriz diagonal expressa por $\boldsymbol{\Sigma}(\boldsymbol{\varphi}, \boldsymbol{\beta}) = \text{diag}(h_{p+q+1}(\boldsymbol{\varphi}, \boldsymbol{\beta}), \dots, h_T(\boldsymbol{\varphi}, \boldsymbol{\beta}))$, podemos reescrever a equação (3.5) como:

$$L(\boldsymbol{\varphi}, \boldsymbol{\beta}) \propto [\det(\boldsymbol{\Sigma}(\boldsymbol{\varphi}, \boldsymbol{\beta}))]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \mathbf{Z}' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\varphi}, \boldsymbol{\beta}) \mathbf{Z}\right\} \quad (3.6)$$

sendo $\mathbf{Z} = (z_{p+q+1}, \dots, z_T)'$ e $z_t = y_t - \mathbf{x}_t \boldsymbol{\beta}$.

Supondo que $\phi_j, j = 0, 1, 2, \dots, q$ são independentes e têm distribuição normal com média zero e variância conhecida σ_j^2 , ou seja, $\pi(\phi_j) \sim Normal(0, \sigma_j^2)$. Assim, a distribuição conjunta a priori para $\boldsymbol{\varphi} = (\phi_0, \phi_1, \dots, \phi_q)'$ e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ é dada por:

$$\pi(\boldsymbol{\varphi}, \boldsymbol{\beta}) \propto \prod_{i=1}^p \left(\frac{1}{k(\boldsymbol{\varphi}) \xi_i} \right)^{1/2} \exp\left\{-\frac{(\beta_i - \mathbf{b})^2}{2k(\boldsymbol{\varphi}) \xi_i}\right\} \times \prod_{j=1}^q \left(\frac{1}{\sigma_j} \right)^{1/2} \exp\left\{-\frac{\phi_j^2}{2\sigma_j^2}\right\} \quad (3.7)$$

sendo $k(\boldsymbol{\varphi}) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i}$ e $\alpha_j = \frac{b_j e^{\phi_j} + a_j}{1 + e^{\phi_j}}, j = 0, 1, \dots, q$.

Combinando as expressões (3.5) e (3.7), podemos escrever a distribuição conjunta a posteriori para $\boldsymbol{\varphi}$ e $\boldsymbol{\beta}$ como:

$$\pi(\boldsymbol{\varphi}, \boldsymbol{\beta} | \mathbf{Y}) \propto \prod_{t=p+q+1}^T \left(\frac{1}{h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})} \right)^{1/2} \exp\left\{-\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})}\right\} \times \prod_{i=1}^p \left(\frac{1}{k(\boldsymbol{\varphi}) \xi_i} \right)^{1/2} \exp\left\{-\frac{(\beta_i - \mathbf{b})^2}{2k(\boldsymbol{\varphi}) \xi_i}\right\} \times \prod_{j=1}^q \left(\frac{1}{\sigma_j} \right)^{1/2} \exp\left\{-\frac{\phi_j^2}{2\sigma_j^2}\right\} \quad (3.8)$$

As distribuições condicionais a posteriori para $\boldsymbol{\varphi}$ e $\boldsymbol{\beta}$, obtidas a partir da expressão (3.8), são dadas, respectivamente, por:

$$\pi(\phi_j | \boldsymbol{\varphi}_{-j}, \boldsymbol{\beta}, \mathbf{Y}) \propto \prod_{t=p+q+1}^T \left(\frac{1}{h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})} \right)^{1/2} \exp\left\{-\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})}\right\} \times \prod_{i=1}^p \left(\frac{1}{k(\boldsymbol{\varphi}) \xi_i} \right)^{1/2} \exp\left\{-\frac{(\beta_i - \mathbf{b})^2}{2k(\boldsymbol{\varphi}) \xi_i}\right\} \times \pi(\phi_j) \quad (3.9)$$

$$\pi(\beta_i | \beta_{-i}, \boldsymbol{\varphi}, \mathbf{Y}) \propto \prod_{t=p+q+1}^T \left(\frac{1}{h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})} \right)^{1/2} \exp\left\{-\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})}\right\} \times \left(\frac{1}{k(\boldsymbol{\varphi}) \xi_i} \right)^{1/2} \exp\left\{-\frac{(\beta_i - \mathbf{b})^2}{2k(\boldsymbol{\varphi}) \xi_i}\right\} \quad (3.10)$$

tal que $\boldsymbol{\varphi}_{-j}$ é um vetor com os parâmetros do modelo exceto o parâmetro ϕ_j , ou seja, $\boldsymbol{\varphi}_{-j} = (\phi_0, \phi_1, \dots, \phi_{j-1}, \phi_{j+1}, \dots, \phi_q)$, β_{-i} é um vetor com os parâmetros do modelo exceto o

parâmetro β_i , isto é, $\beta_{-i} = (\beta_0, \beta_1, \dots, \beta_{i-1}, \beta_{i+1}, \dots, \beta_p)$, $k(\boldsymbol{\varphi}) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i}$ e $\alpha_j = \frac{b_j e^{\phi_j} + a_j}{1 + e^{\phi_j}}$.

3.1.2. Distribuição a priori não-informativa

Quando se tem pouco ou nenhum conhecimento prévio sobre a distribuição dos parâmetros de um modelo, adotamos uma distribuição a priori não-informativa. Geweke (1989) propõe uma distribuição a priori não-informativa para os parâmetros $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$ de um modelo de regressão com erros ARCH, definida como:

$$\pi(\boldsymbol{\alpha}, \boldsymbol{\beta}) \propto \begin{cases} \left(\frac{1 - \sum_{j=1}^q \alpha_j}{\alpha_0} \right) & , \alpha_0 > 0, \alpha_j \geq 0, j = 1, \dots, q \text{ e } -\infty < \beta_i < \infty, i = 1, \dots, p \\ 0 & , \text{ caso contrário} \end{cases} \quad (3.11)$$

Condicional a $\alpha_j = 0, j = 1, \dots, q$, esta é a distribuição a priori invariante de Jeffreys (Box & Tiao, 1973) para o modelo de regressão linear normal.

Usando a reparametrização proposta na seção 3.1.1., definida pela equação 3.4, temos:

$$\pi(\boldsymbol{\varphi}, \boldsymbol{\beta}) \propto \begin{cases} \left(\frac{1}{k(\boldsymbol{\varphi})} \right)^{\frac{1}{2}} & , -\infty < \phi_j < \infty, j = 0, 1, \dots, q \text{ e } -\infty < \beta_i < \infty, i = 1, \dots, p \\ 0 & , \text{ caso contrário} \end{cases} \quad (3.12)$$

sendo $\boldsymbol{\varphi} = (\phi_0, \phi_1, \dots, \phi_q)'$, $k(\boldsymbol{\varphi}) = \frac{\alpha_0}{1 - \sum_{j=1}^q \alpha_j}$ e $\alpha_j = \frac{b_j e^{\phi_j} + a_j}{1 + e^{\phi_j}}, j = 0, 1, \dots, q$.

Então, combinando as expressões (2.7) e (3.12), podemos escrever a distribuição conjunta a posteriori para $\boldsymbol{\varphi}$ e $\boldsymbol{\beta}$ como:

$$\pi(\boldsymbol{\varphi}, \boldsymbol{\beta} | \mathbf{Y}) \propto \left(\frac{1}{k(\boldsymbol{\varphi})} \right)^{1/2} \times \prod_{t=p+q+1}^T \left(\frac{1}{h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})} \right)^{1/2} \exp \left\{ -\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t(\boldsymbol{\varphi}, \boldsymbol{\beta})} \right\} \quad (3.13)$$

3.2. Inferência Bayesiana para os modelos ARCH(q)

Considerando uma trajetória observada de retornos $\mathbf{Z} = \{z_t, t = 1, 2, \dots, T\}$, a abordagem Bayesiana para a inferência dos parâmetros dos modelos ARCH(q) parte da combinação da função de verossimilhança dessa trajetória, $L(\boldsymbol{\alpha})$, com a distribuição a priori para os parâmetros, $\pi(\boldsymbol{\alpha})$, resultando na distribuição a posteriori:

$$\pi(\mathbf{a} | \mathbf{Z}) \propto L(\mathbf{a})\pi(\mathbf{a}) \quad (3.16)$$

tal que $\mathbf{a} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$ e $L(\mathbf{a})$ é a função, definida pela expressão (2.45), dada por:

$$L(\mathbf{a} | z_1, z_2, \dots, z_q) \propto \prod_{t=q+1}^T \left(\frac{1}{h_t} \right)^{1/2} \exp \left\{ -\frac{z_t^2}{2h_t} \right\}$$

3.2.1. Distribuição a priori informativa

Na seção 3.1.1. foi proposta uma distribuição a priori informativa com estrutura hierárquica para os parâmetros de modelos AR(p)-ARCH(q). Como as restrições pertinentes aos modelos ARCH(q) são similares às dos modelos AR(p)-ARCH(q), vamos propor o mesmo procedimento de inferência. Assim, considerando que todos os $\alpha_j, j = 0, 1, \dots, q$ são independentes com uma distribuição a priori definida no intervalo $[a_j, b_j]$ e, usando a reparametrização proposta na equação (3.4), a função de verossimilhança é escrita da mesma forma que na expressão (2.45), mas com a volatilidade h_t reescrita em função dos parâmetros α_j , transformados em ϕ_j por (3.4), ou seja:

$$L(\boldsymbol{\phi}) \propto \prod_{t=q+1}^T \left(\frac{1}{h_t(\boldsymbol{\phi})} \right)^{1/2} \exp \left\{ -\frac{z_t^2}{2h_t(\boldsymbol{\phi})} \right\} \quad (3.17)$$

sendo $\boldsymbol{\phi} = (\phi_0, \phi_1, \dots, \phi_q)'$. Definindo $\boldsymbol{\Sigma}(\boldsymbol{\phi}) = \text{diag}(h_{q+1}(\boldsymbol{\phi}), \dots, h_T(\boldsymbol{\phi}))$, podemos reescrever (3.17) como:

$$L(\boldsymbol{\phi}) \propto [\det(\boldsymbol{\Sigma}(\boldsymbol{\phi}))]^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{Z}' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\phi}) \mathbf{Z} \right\} \quad (3.18)$$

Supondo que os $\phi_j, j = 0, 1, \dots, q$ são normalmente distribuídos, isto é, $\pi(\phi_j) \sim \text{Normal}(0, \sigma_j^2)$, a distribuição conjunta a priori para $\boldsymbol{\phi} = (\phi_0, \phi_1, \dots, \phi_q)'$ é dada por:

$$\pi(\boldsymbol{\phi}) = \prod_{j=0}^q \pi(\phi_j) \quad (3.19)$$

Combinando as expressões (3.18) e (3.19), podemos escrever a distribuição conjunta a posteriori para $\boldsymbol{\phi}$ como:

$$\pi(\boldsymbol{\varphi} | \mathbf{Z}) \propto [\det(\boldsymbol{\Sigma}(\boldsymbol{\varphi}))]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \mathbf{Z}' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\varphi}) \mathbf{Z}\right\} \left(\prod_{j=0}^q \pi(\phi_j)\right) \quad (3.20)$$

Assim, as distribuições condicionais a posteriori para $\phi_j, j = 0, 1, \dots, q$ obtidas da expressão (3.20), são dadas por:

$$\pi(\phi_j | \boldsymbol{\varphi}_{-j}, \mathbf{Z}) \propto L(\boldsymbol{\varphi}) \pi(\phi_j) \quad (3.21)$$

onde $\boldsymbol{\varphi}_{-j}$ é um vetor com os parâmetros do modelo exceto o parâmetro ϕ_j , ou seja, $\boldsymbol{\varphi}_{-j} = (\phi_0, \phi_1, \dots, \phi_{j-1}, \phi_{j+1}, \dots, \phi_q)$.

3.2.2. Distribuição a priori não-informativa

De forma análoga à da seção 3.1.2., propomos uma distribuição a priori não-informativa para os parâmetros $\boldsymbol{\alpha}$ de um modelo ARCH(q), baseada naquela proposta por Geweke (1989). Usando a reparametrização definida pela expressão (3.4), temos:

$$\pi(\boldsymbol{\varphi}) \propto \begin{cases} \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{\frac{1}{2}}, & -\infty < \phi_j < \infty, j = 0, 1, \dots, q \\ 0 & , \text{ caso contrário} \end{cases} \quad (3.22)$$

sendo $\boldsymbol{\varphi} = (\phi_0, \phi_1, \dots, \phi_q)'$, $k(\boldsymbol{\varphi}) = \frac{\alpha_0}{1 - \sum_{j=1}^q \alpha_j}$ e $\alpha_j = \frac{b_j e^{\phi_j} + a_j}{1 + e^{\phi_j}}, j = 0, 1, \dots, q$.

Então, combinando as expressões (2.45) e (3.22), podemos escrever a distribuição conjunta a posteriori para $\boldsymbol{\varphi}$ como:

$$\pi(\boldsymbol{\varphi} | \mathbf{Z}) \propto \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{1/2} \times \prod_{t=q+1}^T \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t(\boldsymbol{\varphi})}\right\} \quad (3.23)$$

As distribuições condicionais descritas neste capítulo são utilizadas no algoritmo de simulação MCMC, Metropolis-Hastings (Chib & Greenberg, 1995), para gerarmos uma amostra representativa das distribuições a posteriori obtidas. Os algoritmos construídos para cada caso apresentado nas seções anteriores estão enunciados no capítulo 4, a seguir.

CAPÍTULO 4

Algoritmos para a Estimação dos Parâmetros de Modelos AR(p)-ARCH(q) e ARCH(q)

Neste capítulo apresentamos os algoritmos propostos na literatura para a obtenção das estimativas de Máxima Verossimilhança e das estimativas Bayesianas, na inferência dos parâmetros de processos AR(p)-ARCH(q) e de processos ARCH(q).

4.1. Algoritmos para abordagem de Máxima Verossimilhança dos modelos AR(p)-ARCH(q)

Nesta seção enunciamos os algoritmos para abordagem de Máxima Verossimilhança dos modelos AR(p)-ARCH(q). Para obtermos as estimativas de Máxima Verossimilhança usamos o Método de Newton, e para melhorarmos os intervalos de confiança padrão propomos o uso da técnica Bootstrap. Apresentamos ainda os critérios de seleção de modelos de Akaike (AIC) e de Schwarz (BIC).

4.1.1. Método de Newton

Devido à diagonalidade do bloco da matriz de informação, expressa em 2.35, a estimação de α e β pode ser considerada separadamente sem perda da eficiência assintótica.

As estimativas de Máxima Verossimilhança para α e β são obtidas através do método de Newton. O procedimento recomendado por Engle (1982) é de, inicialmente, estimarmos β

por mínimos quadrados e obtermos os resíduos. Destes resíduos, uma estimativa de α pode ser construída. Baseada nesta estimativa $\hat{\alpha}$, uma nova estimativa de β pode ser encontrada.

Para obtermos os estimadores de Máxima Verossimilhança para o parâmetro α , usamos a seguinte equação recursiva:

$$\hat{\alpha}^{(k+1)} = \hat{\alpha}^{(k)} + [\hat{I}_{\alpha\alpha}^{(k)}]^{-1} \left[\frac{\partial l^{(k)}}{\partial \alpha} \right]_{\alpha=\hat{\alpha}^{(k)}} \quad (4.1)$$

sendo que $\hat{I}_{\alpha\alpha}$ é a matriz de informação definida em (2.23):

$$\hat{I}_{\alpha\alpha} = \frac{1}{2T} \sum_{t=p+q+1}^T \tilde{\mathbf{v}}_t' \tilde{\mathbf{v}}_t = \frac{1}{2T} \tilde{\mathbf{v}}' \tilde{\mathbf{v}}$$

em que $\tilde{\mathbf{v}}_t = \frac{(1, z_{t-1}^2, \dots, z_{t-q}^2)}{h_t}$, $t = p+q+1, \dots, T$ e $\tilde{\mathbf{v}}' = (\tilde{\mathbf{v}}'_{p+q+1}, \dots, \tilde{\mathbf{v}}'_T)$.

Para uma dada estimativa de α , o método iterativo pode ser usado para melhorar a estimativa de β . Logo, a equação recursiva para β é dada por:

$$\hat{\beta}^{(k+1)} = \hat{\beta}^{(k)} + [\hat{I}_{\beta\beta}^{(k)}]^{-1} \left[\frac{\partial l^{(k)}}{\partial \beta} \right]_{\beta=\hat{\beta}^{(k)}} \quad (4.2)$$

sendo que $\hat{I}_{\beta\beta}$ é a matriz de informação expressa em (2.35):

$$\hat{I}_{\beta\beta} = \frac{1}{T} \sum_{t=1}^{T-p-1} \mathbf{x}'_t \mathbf{x}_t r_t^2$$

em que $r_t^2 = \frac{1}{h_t} + 2e_t^2 \sum_{j=1}^q \frac{\alpha_j^2}{h_{t+j}^2}$.

Nas expressões descritas anteriormente, $e_t = y_t - \mathbf{x}_t \hat{\beta}$ é o resíduo (estimativa de z_t) e $h_t = \hat{\alpha}_0 + \sum_{j=1}^q \hat{\alpha}_j e_{t-j}^2$ é a variância condicional estimada.

4.1.2. Procedimento Bootstrap Paramétrico

Pequenos problemas amostrais surgem quando a matriz de informação \hat{I} é usada na estimação de intervalos de confiança para as estimativas de parâmetros de modelos AR(p)-ARCH(q). Com o objetivo de corrigir intervalos largos obtidos através da abordagem de Máxima Verossimilhança para os parâmetros α_j , $j = 0, 1, \dots, q$, os quais podem ter limitantes

negativos, usamos o procedimento Bootstrap paramétrico para obter melhores intervalos para os parâmetros estimados. A técnica implementada segue a base da que foi apresentada por Gonçalves & White (2003) e Gonçalves & Kilian (2003). Maiores detalhes sobre o procedimento Bootstrap podem ser encontrados no Apêndice A.2.

Para a implementação da técnica Bootstrap paramétrico é necessário que as estimativas de Máxima Verossimilhança sejam calculadas pelas expressões (4.1) e (4.2) até a convergência. Então, usando as estimativas $\hat{\alpha}$ e $\hat{\beta}$ e assumindo $\varepsilon_t \sim N(0,1)$, obtemos uma amostra bootstrap $\{\hat{\alpha}^{*(j)} = \alpha_0^*(j), \alpha_1^*(j), \dots, \alpha_q^*(j), j=1, \dots, B\}$ e $\{\hat{\beta}^{*(j)} = \beta_0^*(j), \beta_1^*(j), \dots, \beta_p^*(j), j=1, \dots, B\}$, usando o seguinte procedimento:

Algoritmo Bootstrap

Passo 1: Obtenha as estimativas de Máxima Verossimilhança $\hat{\alpha} = (\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_q)'$ e $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_p)'$.

Passo 2: Adote um período de aquecimento $t_0 \gg p + q$, um tamanho amostral T e faça o contador $j = 1$.

Passo 3: Faça $z_t = 0$ para $t = 1, \dots, q$ e, em seguida, faça $t = q + 1$.

Passo 4: Calcule $h_t = \hat{\alpha}_0 + \hat{\alpha}_1 z_{t-1} + \dots + \hat{\alpha}_q z_{t-q}$.

Passo 5: Gere $\varepsilon_t \sim N(0,1)$ e calcule $z_t = \sqrt{h_t} \varepsilon_t$.

Passo 6: Faça $\mathbf{x}_t = \mathbf{0}$ para $t = 1, \dots, p$ e, em seguida, faça $t = p + 1$.

Passo 7: Calcule $s_t = \hat{\beta}_1 \mathbf{x}_{t-1} + \hat{\beta}_2 \mathbf{x}_{t-2} + \dots + \hat{\beta}_p \mathbf{x}_{t-p} = \mathbf{x}_t \hat{\beta}$.

Passo 8: Calcule $y_t = s_t + z_t$ para $t > t_0$ até T .

Passo 9: Faça $t = t + 1$ e volte ao passo 4 até $t = T$.

Passo 10: Use a amostra $\{y_{t_0}, \dots, y_T\}$ para calcular as estimativas de Máxima Verossimilhança Bootstrap $\{\hat{\alpha}^{*(j)} = \alpha_0^*(j), \alpha_1^*(j), \dots, \alpha_q^*(j)\}$ e $\{\hat{\beta}^{*(j)} = \beta_0^*(j), \beta_1^*(j), \dots, \beta_p^*(j)\}$.

Passo 11: Faça $j = j + 1$ e volte ao passo 3 até $j = B$.

Portanto, repetindo este procedimento B vezes, obtemos uma amostra bootstrap de $\{\hat{\alpha}^{*(j)} = \alpha_0^*(j), \alpha_1^*(j), \dots, \alpha_q^*(j), j = 1, \dots, B\}$ e $\{\hat{\beta}^{*(j)} = \beta_0^*(j), \beta_1^*(j), \dots, \beta_p^*(j), j = 1, \dots, B\}$. Supondo que uma transformação $u_1^* = f(\hat{\alpha}^* - \hat{\alpha})$ e $u_2^* = f(\hat{\beta}^* - \hat{\beta})$ existe, e que esta faça a distribuição simétrica bootstrap próxima de zero, então, a relação simétrica $u_{\frac{\alpha}{2}}^* = -u_{1-\frac{\alpha}{2}}^*$ para os quantis de u^* pode ser usada e, depois de uma transformação inversa, leva aos intervalos:

$$r_{\left(\frac{\alpha}{2}\right)}^{*(B+1)} \leq \alpha \leq r_{\left(1-\frac{\alpha}{2}\right)}^{*(B+1)} \quad (4.3)$$

$$r_{\left(\frac{\alpha}{2}\right)}^{*(B+1)} \leq \beta \leq r_{\left(1-\frac{\alpha}{2}\right)}^{*(B+1)} \quad (4.4)$$

Estes intervalos são chamados de intervalos percentis bootstrap (Efron, 1979). Desta forma, calculamos intervalos de confiança através da metodologia percentil. Neste trabalho, os intervalos são dados pelos percentis 2,5% e 97,5% das estimativas amostrais.

4.1.3. Critério de Akaike (AIC) e Critério de Schwarz (BIC)

Uma vez escolhidos vários modelos e calculados seus parâmetros, o passo seguinte é escolher o modelo mais “parcimonioso”, ou seja, aquele que tem o menor número de parâmetros possível e menor variância dos resíduos. A seleção do modelo é de extrema importância, pois a escolha de um valor “pequeno” demais para a ordem do modelo pode nos levar a uma representação inadequada da série, enquanto que um valor “grande” pode nos levar a um grau de complexidade desnecessário.

A seleção de modelos e a identificação da ordem dos mesmos podem ser feitos baseados no critério de informação de Akaike (AIC) e no critério de informação Bayesiano de Schwarz (BIC). Para os modelos AR(p)-ARCH(q) esses critérios apresentam as seguintes formulações (Akaike, 1974; Schwarz, 1978):

$$AIC = -\frac{2}{T}l + \frac{2M}{T} \quad (4.5)$$

$$BIC = -\frac{2}{T}l + \frac{M \ln(T)}{T} \quad (4.6)$$

sendo que $M = p + q + 1$ é o número de parâmetros, T é o tamanho da série em questão e l é a média do logaritmo natural da função de verossimilhança definida na equação (2.9).

Para séries com um número de observações T muito grande, como é o caso de séries financeiras, o critério de Schwarz (BIC) mostra-se mais adequado que o de Akaike (AIC).

Os critérios AIC e BIC, expressos pelas equações (4.5) e (4.6), respectivamente, qualificam o modelo pelo chamado “critério da parcimônia”, escolhendo o modelo pelo compromisso entre o ajuste adequado (Máxima Verossimilhança) e o menor número de parâmetros M . O modelo que apresenta o menor valor de AIC e/ou BIC é o modelo mais adequado para representar os dados.

4.2. Algoritmos para abordagem Bayesiana dos modelos AR(p)-ARCH(q)

Nesta seção enunciamos os algoritmos para abordagem Bayesiana dos modelos AR(p)-ARCH(q). Para obtermos as estimativas Bayesianas usamos métodos de simulação de Monte Carlo em Cadeias de Markov (MCMC), tal como o algoritmo Metropolis-Hastings. Apresentamos ainda o critério de seleção de modelos baseado na densidade Preditiva Ordenada (CPO).

4.2.1. O Algoritmo Metropolis-Hastings

O Metropolis-Hastings é um esquema iterativo usado para gerar amostra de uma densidade de forma indireta, quando não se sabe gerar diretamente desta densidade. O algoritmo consiste em escolher um núcleo $\gamma(\cdot)$, e gerar deste núcleo, usando um critério de aceitação do valor gerado para garantir que a amostra obtida é representativa da amostra gerada. Maiores detalhes podem ser encontrados em Chib & Greenberg (1995); Gamerman (1996) e no Apêndice A.3.

Para obtermos uma amostra representativa da distribuição conjunta a posteriori dos parâmetros dos modelos AR(p)-ARCH(q), usamos as distribuições condicionais dos mesmos, a fim de implementarmos o algoritmo Metropolis-Hastings, que pode ser resumido como segue:

Algoritmo Metropolis-Hastings (Distribuição a priori informativa)

Passo 1: Forneça valores iniciais arbitrários $\boldsymbol{\phi}^{(0)} = (\phi_j^{(0)}), j = 0, 1, \dots, q$ e calcule $\boldsymbol{\alpha}^{(0)} = (\alpha_j^{(0)}), j = 0, 1, \dots, q$ correspondente. Forneça valores iniciais arbitrários $\boldsymbol{\beta}^{(0)} = (\beta_k^{(0)}), k = 1, \dots, p$. Faça $i = 1$.

Passo 2: Gere um novo valor γ_1 a partir da densidade condicional $\pi_j(\phi_j | \phi_{-j}, \boldsymbol{\beta}, \mathbf{Y})$, definida em (3.9), fazendo:

$$[\boldsymbol{\phi}^c, \boldsymbol{\beta}] = (\phi_0^{(i-1)}, \phi_1^{(i-1)}, \dots, \phi_{j-1}^{(i-1)}, \gamma_1, \phi_{j+1}^{(i-1)}, \dots, \phi_q^{(i-1)}, \beta_1^{(i-1)}, \dots, \beta_p^{(i-1)})$$

Passo 3: Calcule a probabilidade de aceitação do novo valor gerado γ_1 :

$$\lambda(\phi_j^{(i-1)}, \gamma_1) = \begin{cases} \min \left\{ 1, \frac{\pi(\boldsymbol{\phi}^c, \boldsymbol{\beta} | \mathbf{Y})}{\pi(\boldsymbol{\phi}^{(i-1)}, \boldsymbol{\beta}^{(i-1)} | \mathbf{Y})} \right\} & , \text{ se } \pi(\boldsymbol{\phi}^{(i-1)}, \boldsymbol{\beta}^{(i-1)} | \mathbf{Y}) \neq 0 \\ 1 & , \text{ caso contrário} \end{cases}$$

Passo 4: Gere uma variável aleatória uniforme $u_1 \sim U(0,1)$ e faça:

$$\phi_j^{(i)} = \begin{cases} \gamma_1 & , \text{ se } u_1 \leq \lambda(\phi_j^{(i-1)}, \gamma_1) \\ \phi_j^{(i-1)} & , \text{ caso contrário} \end{cases}$$

Passo 5: Repita os passos 2 a 4 para $j = 0, 1, \dots, q$.

Passo 6: Gere um novo valor γ_2 da densidade condicional $\pi_k(\beta_k | \beta_{-k}, \boldsymbol{\phi}, \mathbf{Y})$, definida em (3.10), fazendo:

$$[\boldsymbol{\phi}, \boldsymbol{\beta}^c] = (\phi_0^{(i)}, \phi_1^{(i)}, \dots, \phi_q^{(i)}, \beta_1^{(i-1)}, \dots, \beta_{k-1}^{(i-1)}, \gamma_2, \beta_{k+1}^{(i-1)}, \dots, \beta_p^{(i-1)})$$

Passo 7: Calcule a probabilidade de aceitação do novo valor gerado γ_2 :

$$\lambda(\beta_k^{(i-1)}, \gamma_2) = \begin{cases} \min \left\{ 1, \frac{\pi(\boldsymbol{\phi}, \boldsymbol{\beta}^c | \mathbf{Y})}{\pi(\boldsymbol{\phi}^{(i)}, \boldsymbol{\beta}^{(i-1)} | \mathbf{Y})} \right\} & , \text{ se } \pi(\boldsymbol{\phi}^{(i)}, \boldsymbol{\beta}^{(i-1)} | \mathbf{Y}) \neq 0 \\ 1 & , \text{ caso contrário} \end{cases}$$

Passo 8: Gere uma variável aleatória uniforme $u_2 \sim U(0,1)$ e faça:

$$\beta_k^{(i)} = \begin{cases} \gamma_2 & , \text{ se } u_2 \leq \lambda(\beta_k^{(i-1)}, \gamma_2) \\ \beta_k^{(i-1)} & , \text{ caso contrário} \end{cases}$$

Passo 9: Repita os passos 6 a 8 para $k = 1, \dots, p$.

Passo 10: Faça $i = i + 1$ e repita os passos 2 a 9 até a convergência estar garantida.

Após a convergência os parâmetros estimados no espaço de parâmetros original

$$\hat{\alpha}_j, j = 0, 1, \dots, q \text{ são recuperados pela transformação inversa } \hat{\alpha}_j = \frac{b_j e^{\hat{\phi}_j} + a_j}{1 + e^{\hat{\phi}_j}}.$$

Algoritmo Metropolis-Hastings (Distribuição a priori não-informativa)

Quando estamos considerando distribuição a priori não-informativa, não temos na expressão da posteriori em (3.12), uma distribuição conhecida para gerarmos parâmetros

candidatos. Neste caso, utilizamos um núcleo Gaussiano $N(\boldsymbol{\theta})$, de forma que, a distribuição a posteriori pode ser escrita como:

$$\pi^*(\boldsymbol{\theta} | \mathbf{Y}) = \frac{\pi(\boldsymbol{\theta} | \mathbf{Y})N(\boldsymbol{\theta})}{N(\boldsymbol{\theta})}, \quad N(\boldsymbol{\theta}) \neq 0 \quad (4.7)$$

sendo $\boldsymbol{\theta} = (\boldsymbol{\varphi}, \boldsymbol{\beta})$.

A probabilidade de aceitação do candidato gerado (probabilidade de movimentação na cadeia gerada) é calculada por:

$$\lambda(\boldsymbol{\theta}_j^{(i-1)}, \gamma_j) = \begin{cases} \pi^*(\boldsymbol{\theta}^{(i)} | \mathbf{Y}) / \pi^*(\boldsymbol{\theta}^{(i-1)} | \mathbf{Y}) & , \text{ se } \pi^*(\boldsymbol{\theta}^{(i-1)} | \mathbf{Y}) \neq 0 \\ 1 & , \text{ caso contrário} \end{cases} \quad (4.8)$$

O candidato γ_j é gerado do núcleo e formamos o novo vetor de parâmetros:

$$\boldsymbol{\theta}^{(i)} = (\boldsymbol{\theta}_0^{(i-1)}, \boldsymbol{\theta}_1^{(i-1)}, \dots, \boldsymbol{\theta}_{j-1}^{(i-1)}, \gamma_j, \boldsymbol{\theta}_{j+1}^{(i-1)}, \dots, \boldsymbol{\theta}_{p+q}^{(i-1)}) \quad (4.9)$$

Podemos notar que, considerando um núcleo Gaussiano, temos:

$$\frac{\pi^*(\boldsymbol{\theta}^{(i)} | \mathbf{Y})}{\pi^*(\boldsymbol{\theta}^{(i-1)} | \mathbf{Y})} = \frac{\pi(\boldsymbol{\theta}^{(i)} | \mathbf{Y})}{\pi(\boldsymbol{\theta}^{(i-1)} | \mathbf{Y})} \quad (4.10)$$

Da mesma forma descrita para a distribuição a priori informativa, os parâmetros estimados $\hat{\alpha}_j, j = 0, 1, \dots, q$ podem ser recuperados pela transformação inversa

$$\hat{\alpha}_j = \frac{b_j e^{\hat{\phi}_j} + a_j}{1 + e^{\hat{\phi}_j}}.$$

Uma característica que devemos levar em consideração na utilização do algoritmo Metropolis-Hastings é o fato deste gerar amostras correlacionadas, uma vez que, se o candidato gerado é rejeitado, adotamos o valor anterior como candidato. A avaliação deste tipo de procedimento é refletida na taxa de rejeição e a convergência é verificada usando o critério de Geweke (Geweke, 1992; Apêndice A.3).

Para assegurar que o algoritmo Metropolis-Hastings convirja para uma distribuição de equilíbrio $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y})$, devemos observar as condições de regularidade da distribuição a posteriori $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y})$, as quais são asseguradas no nosso caso.

Assumindo uma função de perda quadrática, as estimativas Bayesianas de $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$ podem ser expressas como a avaliação da esperança de uma função de interesse $g(\boldsymbol{\alpha}, \boldsymbol{\beta})$, com relação a distribuição a posteriori $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y})$, dada por:

$$E[g(\boldsymbol{\alpha}, \boldsymbol{\beta})] = \int \int_{\boldsymbol{\alpha} \boldsymbol{\beta}} g(\boldsymbol{\alpha}, \boldsymbol{\beta}) \pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y}) d\boldsymbol{\alpha} d\boldsymbol{\beta} \quad (4.11)$$

A solução exata da integral múltipla em (4.11) não pode ser calculada analiticamente para os modelos que estamos considerando neste trabalho. No entanto, sempre que $E[g(\boldsymbol{\alpha}, \boldsymbol{\beta})]$ existe, uma aproximação para a expressão (4.11) pode ser obtida por simulação de Monte Carlo (Geweke, 1986c).

4.2.2. Critério da Densidade Preditiva Ordenada (CPO)

Na abordagem Bayesiana dos modelos AR(p)-ARCH(q) utilizamos como critério para seleção dos modelos ajustados, o critério da densidade Preditiva Ordenada (Carlin & Chib, 1995; Oliveira, 1998). Neste critério usamos a distribuição de y_{T+m} condicionada aos dados \mathbf{Y} e aos parâmetros $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$ e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$. Portanto, a densidade preditiva é dada por:

$$\pi(y_{T+m} | \mathbf{Y}) = \int \int_{\boldsymbol{\alpha} \boldsymbol{\beta}} \pi(y_{T+m} | \mathbf{Y}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y}) d\boldsymbol{\alpha} d\boldsymbol{\beta} \quad (4.12)$$

$$c_m = \pi(y_{T+m} | \mathbf{Y}) = \int \int_{\boldsymbol{\alpha} \boldsymbol{\beta}} \left(\frac{1}{2\pi h_{T+m}} \right)^{1/2} \exp\left\{ -\frac{1}{2} \frac{z_{T+m}^2}{h_{T+m}} \right\} \pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y}) d\boldsymbol{\alpha} d\boldsymbol{\beta}$$

em que $z_{T+m} = y_{T+m} - \mathbf{x}_{T+m} \boldsymbol{\beta}$, $\pi(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y})$ é a distribuição a posteriori para os parâmetros $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$, e a integral em (4.12) é uma integral múltipla no espaço dos parâmetros.

A equação (4.12) também pode ser avaliada usando as observações $\alpha_j^{(i)}$, $j = 0, 1, \dots, q$ e $\beta_k^{(i)}$, $k = 1, \dots, p$ obtidas por simulação MCMC e considerando h_{T+m} como uma função dos parâmetros $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$. Neste caso, a estimativa de Monte Carlo da densidade preditiva é dada por:

$$\hat{c}_m = \frac{1}{M} \sum_{i=1}^M \frac{1}{\sqrt{2\pi h_{T+m}(\boldsymbol{\alpha}^{(i)}, \boldsymbol{\beta}^{(i)})}} \exp\left\{ -\frac{1}{2} \frac{z_{T+m}^2}{h_{T+m}(\boldsymbol{\alpha}^{(i)}, \boldsymbol{\beta}^{(i)})} \right\} \quad (4.13)$$

Um baixo valor de \hat{c}_m indica que a observação z_{T+m} é improvável para o modelo considerado. Assim, o critério de seleção de modelo baseado na densidade Preditiva Ordenada (CPO) consiste em escolher o modelo l que apresenta o maior valor de $\hat{c}(l) = \prod_{m=T+1}^{T+K} \hat{c}_m(l)$.

4.3. Algoritmos para abordagem de Máxima Verossimilhança dos modelos ARCH(q)

4.3.1. Método de Newton

De modo análogo ao apresentado na seção 4.1.1., um método iterativo para calcularmos os estimadores de Máxima Verossimilhança para os parâmetros α_j , $j = 0, 1, \dots, q$ de modelos ARCH(q), consiste em iteragir a equação recursiva:

$$\boldsymbol{\alpha}^{(k+1)} = \boldsymbol{\alpha}^{(k)} + [\hat{I}^{(k)}]^{-1} \left[\frac{\partial l^{(k)}}{\partial \boldsymbol{\alpha}} \right]_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}^{(k)}} \quad (4.14)$$

em que \hat{I} é a matriz de informação dada pela equação (2.52), avaliada com a estimativa $\boldsymbol{\alpha}^{(k)}$:

$$\hat{I} = \frac{1}{2T} \sum_{t=q+1}^T \tilde{\mathbf{v}}_t' \tilde{\mathbf{v}}_t = \frac{1}{2T} \tilde{\mathbf{v}}' \tilde{\mathbf{v}}$$

sendo $\tilde{\mathbf{v}}_t = \frac{(1, z_{t-1}^2, \dots, z_{t-q}^2)}{h_t}$, $t = q+1, \dots, T$ e $\tilde{\mathbf{v}}' = (\tilde{\mathbf{v}}_{q+1}', \dots, \tilde{\mathbf{v}}_T')$.

4.3.2. Procedimento Bootstrap Paramétrico

Da mesma forma que a da seção 4.1.2., para a implementação da técnica Bootstrap paramétrico, é necessário que as estimativas de Máxima Verossimilhança sejam calculadas pela equação (4.14) até a obtenção da convergência. Usando as estimativas $\hat{\boldsymbol{\alpha}}$ e assumindo $\varepsilon_t \sim N(0,1)$, obtemos uma amostra Bootstrap $\{\hat{\boldsymbol{\alpha}}^{*(j)} = \alpha_0^*(j), \alpha_1^*(j), \dots, \alpha_q^*(j), j = 1, \dots, B\}$. Resumidamente, temos:

Algoritmo Bootstrap

Passo 1: Obtenha as estimativas de Máxima Verossimilhança $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_q)'$.

Passo 2: Adote um período de aquecimento $t_0 \gg q$, um tamanho amostral T e faça o contador $j = 1$.

Passo 3: Faça $z_t = 0$ para $t = 1, \dots, q$ e, em seguida, faça $t = q + 1$.

Passo 4: Calcule $h_t = \hat{\alpha}_0 + \hat{\alpha}_0 z_{t-1} + \dots + \hat{\alpha}_q z_{t-q}$.

Passo 5: Gere $\varepsilon_t \sim N(0,1)$ e calcule $z_t = \sqrt{h_t} \varepsilon_t$.

Passo 6: Faça $t = t + 1$ e volte ao passo 4 até $t = T$.

Passo 7: Use a amostra $\{z_{t_0}, \dots, z_T\}$ para calcular as estimativas de Máxima Verossimilhança Bootstrap $\{\hat{\alpha}^{*(j)} = \alpha_0^*(j), \alpha_1^*(j), \dots, \alpha_q^*(j)\}$.

Passo 8: Faça $j = j + 1$ e volte ao passo 3 até $j = B$.

Portanto, repetindo este procedimento B vezes, obtemos uma amostra bootstrap de $\{\hat{\alpha}^{*(j)} = \alpha_0^*(j), \alpha_1^*(j), \dots, \alpha_q^*(j), j = 1, \dots, B\}$ e, supondo que uma transformação $u^* = f(\hat{\alpha}^* - \hat{\alpha})$ existe e faz a distribuição simétrica bootstrap próxima de zero, temos a relação simétrica $u_{\frac{\alpha}{2}}^* = -u_{1-\frac{\alpha}{2}}^*$ para os quantis de u^* , que depois de uma transformação inversa, leva aos intervalos percentis bootstrap:

$$r_{\left(\frac{\alpha}{2}\right)}^{*(B+1)} \leq \mathbf{u} \leq r_{\left(1-\frac{\alpha}{2}\right)}^{*(B+1)} \quad (4.15)$$

4.3.3. Critério de Akaike (AIC) e Critério de Schwarz (BIC)

A seleção de modelos e a identificação da ordem dos mesmos também podem ser feitos baseados no critério de informação de Akaike (AIC) e no critério de informação Bayesiano de Schwarz (BIC). Para os modelos ARCH(q) estes critérios apresentam formulações similares às apresentadas na seção 4.1.3. pelas equações (4.5) e (4.6), onde $M = q + 1$ é o número de parâmetros e l é a média do logaritmo natural da função de verossimilhança definida pela equação (2.47). O modelo que apresenta o menor valor de AIC e/ou BIC é o modelo mais adequado para representar os dados.

4.4. Algoritmos para abordagem Bayesiana dos modelos ARCH(q)

4.4.1. O Algoritmo Metropolis-Hastings

Com as distribuições condicionais a posteriori obtidas para os parâmetros dos modelos ARCH(q), implementamos o algoritmo Metropolis-Hastings, que pode ser resumido como:

Algoritmo Metropolis-Hastings (Distribuição a priori informativa)

Passo 1: Forneça valores iniciais arbitrários $\boldsymbol{\varphi}^{(0)} = (\phi_j^{(0)}), j = 0, 1, \dots, q$ e faça $i = 1$.

Passo 2: Gere um novo valor γ da densidade condicional $\pi_j(\phi_j | \boldsymbol{\varphi}_{-j}, \mathbf{Z})$, fazendo:

$$\boldsymbol{\varphi}^c = (\phi_0^{(i-1)}, \phi_1^{(i-1)}, \dots, \phi_{j-1}^{(i-1)}, \gamma, \phi_{j+1}^{(i-1)}, \dots, \phi_q^{(i-1)})$$

Passo 3: Calcule a probabilidade de aceitação do novo valor gerado γ :

$$\lambda(\phi_j^{(i-1)}, \gamma) = \begin{cases} \min \left\{ 1, \frac{\pi(\boldsymbol{\varphi}^c | \mathbf{Z})}{\pi(\boldsymbol{\varphi}^{(i-1)} | \mathbf{Z})} \right\} & , \text{ se } \pi(\boldsymbol{\varphi}^{(i-1)} | \mathbf{Z}) \neq 0 \\ 1 & , \text{ caso contrário} \end{cases}$$

Passo 4: Gere uma variável aleatória uniforme $u \sim U(0,1)$ e faça:

$$\phi_j^{(i)} = \begin{cases} \gamma & , \text{ se } u \leq \lambda(\phi_j^{(i-1)}, \gamma) \\ \phi_j^{(i-1)} & , \text{ caso contrário} \end{cases}$$

Passo 5: Repita os passos 2 a 5 para $j = 0, 1, \dots, q$.

Passo 6: Faça $i = i + 1$ e volte ao passo 2 até a convergência estar garantida.

Após a convergência os parâmetros estimados $\hat{\alpha}_j, j = 0, 1, \dots, q$ são recuperados pela

transformação inversa
$$\hat{\alpha}_j = \frac{b_j e^{\hat{\phi}_j} + a_j}{1 + e^{\hat{\phi}_j}}.$$

Podemos notar que a diferença entre o algoritmo enunciado acima e o algoritmo Metropolis-Hastings para os modelos AR(p)-ARCH(q) está no fato de que o segundo tem dois módulos de geração dos parâmetros, um para $\boldsymbol{\alpha}$ e o outro para $\boldsymbol{\beta}$. De resto, os procedimentos são iguais, inclusive no que diz respeito ao uso de um núcleo de geração dos candidatos para o caso da *distribuição a priori não-informativa*, com $\boldsymbol{\theta} = \boldsymbol{\varphi}$.

O estimador Bayesiano, sob uma função de perda quadrática, pode ser obtido por simulação de Monte Carlo, calculando o valor esperado da amostra gerada da distribuição a posteriori $\pi(\boldsymbol{\alpha} | \mathbf{Z})$, de forma similar ao procedimento discutido na seção 4.2.1.

4.4.2. Critério da Densidade Preditiva Ordenada (CPO)

Similarmente, na abordagem Bayesiana dos modelos ARCH(q), utilizamos como critério para seleção dos modelos ajustados, o critério da densidade Preditiva Ordenada.

Portanto, usando a distribuição de z_{T+m} condicionada aos dados \mathbf{Z} e aos parâmetros $\mathbf{a} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$, temos:

$$\pi(z_{T+m} | \mathbf{Z}) = \int_{\mathbf{a}} \pi(z_{T+m} | \mathbf{Z}, \mathbf{a}) \pi(\mathbf{a} | \mathbf{Z}) d\mathbf{a} \quad (4.16)$$

$$c_m = \pi(z_{T+m} | \mathbf{Z}) = \int_{\mathbf{a}} \left(\frac{1}{2\pi h_{T+m}} \right)^{1/2} \exp\left\{ -\frac{1}{2} \frac{z_{T+m}^2}{h_{T+m}} \right\} \pi(\mathbf{a} | \mathbf{Z}) d\mathbf{a}$$

em que $\pi(\mathbf{a} | \mathbf{Z})$ é a distribuição a posteriori para os parâmetros \mathbf{a} e a integral em (4.16) é uma integral múltipla no espaço dos parâmetros.

Avaliando as observações $\alpha_j^{(i)}$, $j = 0, 1, \dots, q$ obtidas por simulação MCMC e considerando h_{T+m} como uma função dos parâmetros \mathbf{a} , a estimativa de Monte Carlo da densidade preditiva é dada por:

$$\hat{c}_m = \frac{1}{M} \sum_{i=1}^M \frac{1}{\sqrt{2\pi h_{T+m}(\mathbf{a}^{(i)})}} \exp\left\{ -\frac{1}{2} \frac{z_{T+m}^2}{h_{T+m}(\mathbf{a}^{(i)})} \right\} \quad (4.17)$$

e o modelo l selecionado dever ser aquele que apresentar o maior valor de $\hat{c}(l) = \prod_{m=T+1}^{T+K} \hat{c}_m(l)$.

CAPÍTULO 5

Aplicações

5.1. Introdução

Neste capítulo apresentamos uma comparação entre a abordagem de Máxima Verossimilhança (MV) e uma abordagem Bayesiana para a inferência dos parâmetros de processos $AR(p)$ - $ARCH(q)$ e de processos $ARCH(q)$. Intervalos de confiança para as estimativas de Máxima Verossimilhança são obtidos da forma padrão e através da técnica Bootstrap, e intervalos de credibilidade para as estimativas Bayesianas são obtidos através da técnica de simulação MCMC.

As estimativas de Máxima Verossimilhança e os intervalos de confiança são comparados com as estimativas Bayesianas e com os intervalos de credibilidade das distribuições a posteriori geradas pelo algoritmo Metropolis-Hastings. De maneira simplificada, estes intervalos indicam o grau de certeza com que se pode encontrar o valor real do parâmetro dentro da faixa por eles especificada. Quanto menor for a largura do intervalo, melhor é o método de estimação. Neste estudo é usado um grau de certeza (nível de confiança e de credibilidade) de 95%. O objetivo é mostrar que a robustez dos estimadores dos parâmetros de modelos $AR(p)$ - $ARCH(q)$ e de modelos $ARCH(q)$, em relação à ordem dos modelos, é maior com o uso da análise Bayesiana do que com o uso da análise de Máxima Verossimilhança, sob diferentes técnicas baseadas em simulação.

Com o propósito de comparar as duas abordagens, vamos apresentar agora séries temporais simuladas e séries temporais do mercado financeiro brasileiro e, através dos intervalos de confiança e dos intervalos de credibilidade, avaliar a robustez de cada abordagem na estimação dos parâmetros dos modelos propostos.

Os resultados para as séries simuladas e para as séries financeiras foram obtidos para os modelos $AR(p)$ - $ARCH(q)$ e para os modelos $ARCH(q)$ selecionados pelos critérios AIC e BIC. Na abordagem Bayesiana usamos o critério da Densidade Preditiva Ordenada (CPO). Os critérios de seleção de modelos, bem como os algoritmos utilizados na obtenção dos resultados, estão enunciados no Capítulo 4.

Com o objetivo de corrigir intervalos largos obtidos através da abordagem de Máxima Verossimilhança, os quais podem ter limitantes negativos, usamos o procedimento Bootstrap paramétrico (seções 4.1.2 e 4.2.2; Apêndice A2) para melhorar os intervalos de confiança dos parâmetros dos modelos ajustados às séries financeiras. Com as estimativas de Máxima Verossimilhança encontradas, fizemos 1000 reamostragens do modelo considerado e obtivemos uma amostra de estimativas de Máxima Verossimilhança. A partir desta amostra, encontramos os intervalos Bootstrap dados pelos percentis 2,5% e 97,5% das estimativas amostrais.

Na implementação do algoritmo Metropolis-Hastings, para cada parâmetro foi simulada uma cadeia com 50000 iterações, descartado 50% dos valores para diminuir o efeito das condições iniciais, e então tomados valores espaçados de 5 em 5, totalizando uma amostra de 5000 observações. A convergência do algoritmo foi verificada pelo critério de Geweke, a um nível de significância de 5% para o teste, sob a hipótese nula H_0 (Geweke, 1992; Apêndice A.3). Assim, a convergência do parâmetro foi considerada para valores obtidos pelo diagnóstico de Geweke entre -1,96 e 1,96.

5.2. Resultados obtidos com o ajuste de Modelos $ARCH(q)$

Primeiramente, vamos apresentar os resultados obtidos para uma série gerada segundo um processo $ARCH(2)$ com 1000 valores e, posteriormente, os resultados obtidos para duas séries reais do mercado financeiro brasileiro: Índice Bovespa (IBovespa) e Telebrás. Todos os conjuntos de dados foram analisados considerando as abordagens propostas para modelos $ARCH(q)$ descritas neste trabalho, sob suposição de normalidade para os dados.

5.2.1. Série Gerada

Selecionar o “melhor modelo” segundo os critérios de seleção apresentados, a análise tem a finalidade de validar os algoritmos desenvolvidos, pois como a série foi gerada segundo um processo ARCH(q) conhecido, espera-se que o modelo escolhido seja o mais próximo do verdadeiro. A Figura 5.1 a seguir ilustra uma série de retornos gerada por um modelo ARCH(2).

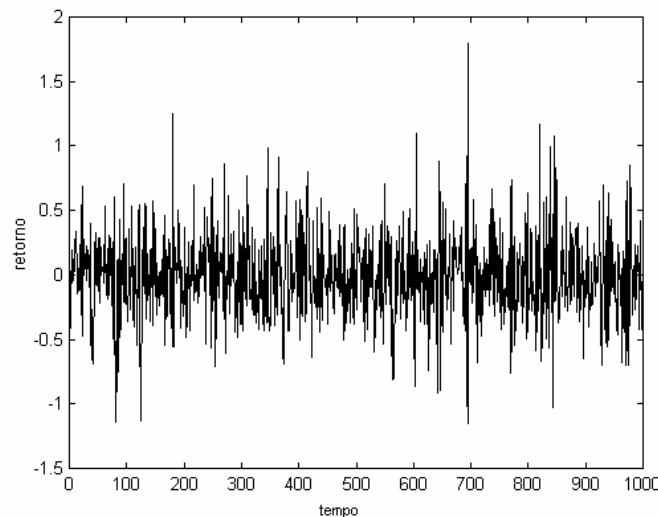


Figura 5.1: Modelo ARCH(2) com $\alpha_0 = 0,05$, $\alpha_1 = 0,40$ e $\alpha_2 = 0,20$.

Os gráficos das figuras a seguir mostram o comportamento da série de retornos gerada. Na Figura 5.2, os gráficos de autocorrelação e de autocorrelação parcial indicam que a série y_t é não-correlacionada. Na Figura 5.3, podemos observar ainda que a série dos quadrados destes retornos, y_t^2 , é correlacionada. Este comportamento pode ser tipicamente associado a processos ARCH(q).

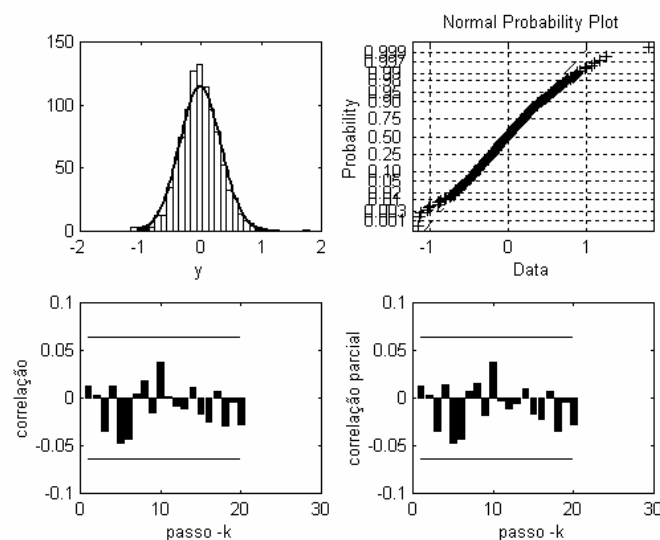


Figura 5.2: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t - Modelo ARCH(2).

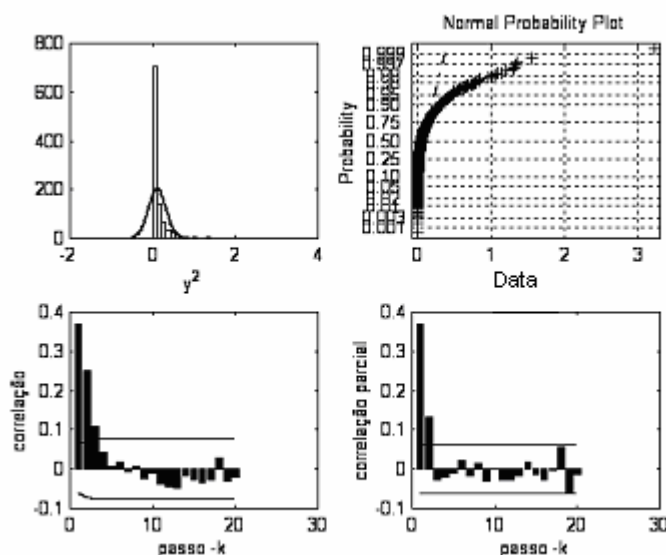


Figura 5.3: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t^2 - Modelo ARCH(2).

A Tabela 5.1 mostra os valores dos critérios de seleção obtidos para algumas ordens de modelos testadas. Como era esperado, concluímos que o modelo ARCH(2) é o modelo mais adequado para representar os dados, pois apresenta os menores valores de AIC e BIC, respectivamente, e os maiores valores para o critério da densidade Preditiva Ordenada (CPO) na análise Bayesiana, considerando distribuição a priori informativa (PI) e distribuição a priori não-informativa (PNI) para os parâmetros do modelo.

Tabela 5.1: Critérios de seleção de modelos - Modelo ARCH(q).

Ordem do modelo	AIC	BIC	CPO – PI	CPO - PNI
$q=1$	0,55149	0,56131	2,97755e-007	3,10616e-007
$q=2$	0,52883	0,54356	2,07838e-006	2,33338e-006
$q=3$	0,53049	0,55012	3,21003e-007	3,98206e-007

Escolhida a ordem do modelo mais adequada, encontramos as estimativas dos parâmetros do modelo selecionado usando a abordagem de Máxima Verossimilhança (MV). Os resultados obtidos são dados na Tabela 5.2 a seguir.

Tabela 5.2: Estimativas de MV e Intervalos de Confiança de 95%.

Parâmetro	Estimativa	Intervalo de Confiança (95%)
α_0	0,05173	[0,04283;0,06063]
α_1	0,42765	[0,30855;0,54674]
α_2	0,15521	[0,07214;0,23827]

Na abordagem Bayesiana, através da distribuição a priori informativa (PI) proposta, obtivemos as estimativas dos parâmetros dadas pela Tabela 5.3. A convergência dos parâmetros foi verificada pelo critério de Geweke (1992), a qual pode ser observada para os mesmos. A última coluna da Tabela 5.3 mostra ainda as taxas (ou probabilidades) de aceitação dos valores dos parâmetros gerados pelo algoritmo Metropolis-Hastings.

Tabela 5.3: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,05179	0,00393	0,05164	[0,04453;0,05977]	1,00810	53,916
α_1	0,40784	0,03697	0,40870	[0,33639;0,47398]	-0,70122	77,316
α_2	0,17973	0,03237	0,17749	[0,12362;0,24764]	0,62440	64,048

Os gráficos da Figura 5.4 representam os histogramas construídos a partir da amostra selecionada para os parâmetros α_0 , α_1 e α_2 , e os da Figura 5.5, os trezentos últimos valores gerados dos mesmos, mostrando a convergência gráfica.

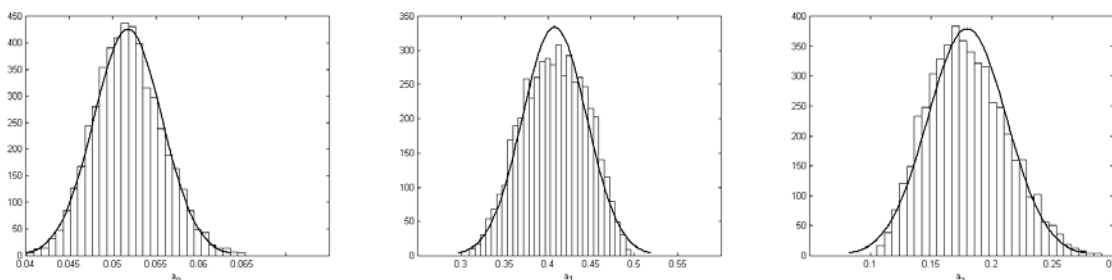


Figura 5.4: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Modelo ARCH(2).

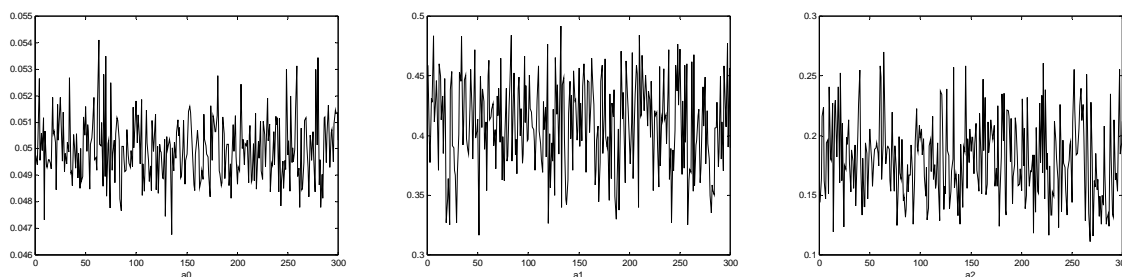


Figura 5.5: Valores gerados dos parâmetros considerando distribuição a priori informativa - Modelo ARCH(2).

Os gráficos da Figura 5.6 resumem a análise de resíduos para o modelo selecionado na abordagem Bayesiana com distribuição a priori informativa. Observamos que os resíduos

estão normalmente distribuídos e pouco correlacionados. Portanto, podemos confirmar que o modelo escolhido é adequado aos dados. A Figura 5.7 mostra a volatilidade obtida a partir das estimativas Bayesianas, usando distribuição a priori informativa (Bayes - PI), comparada com a volatilidade obtida pelas estimativas de Máxima Verossimilhança (MV).

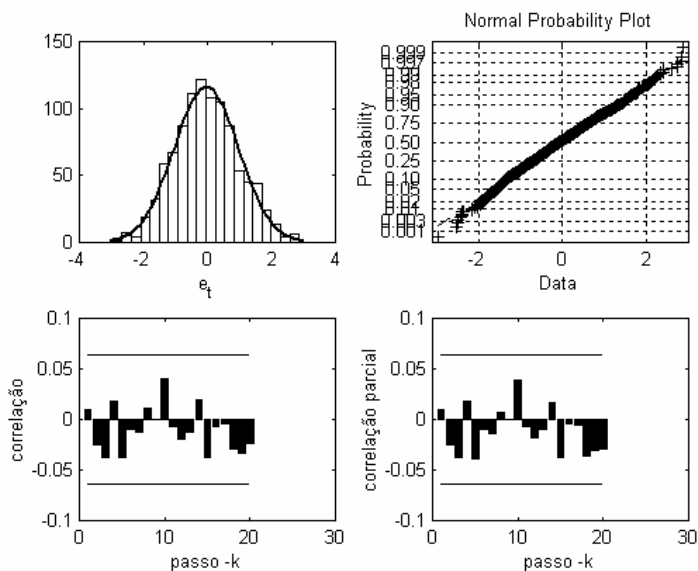


Figura 5.6: Análise de Resíduos - Modelo ARCH(2) Abordagem Bayesiana com distribuição a priori informativa.

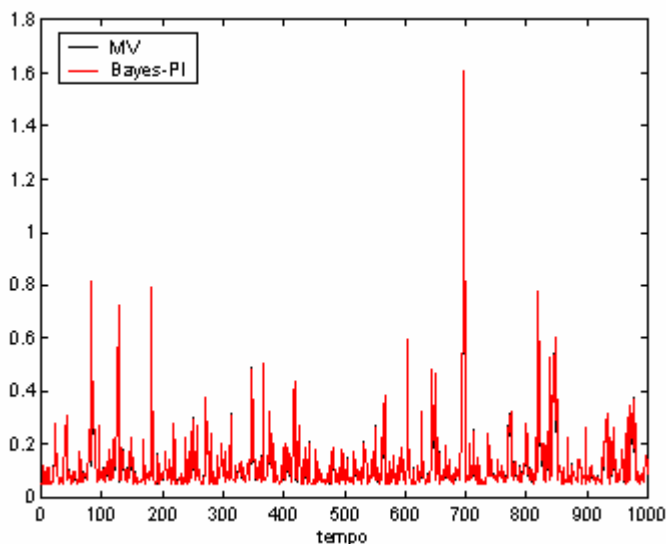


Figura 5.7: Volatilidade Estimada - Modelo ARCH(2).

Para a distribuição a priori não-informativa (PNI), as estimativas dos parâmetros obtidas estão apresentadas na Tabela 5.4. Da mesma forma que para a distribuição a priori informativa, a convergência dos parâmetros foi confirmada pelo critério de Geweke. Os

histogramas construídos a partir da amostra selecionada para os parâmetros do modelo simulado ARCH(2) são dados pela Figura 5.8 e, a Figura 5.9, representa os trezentos últimos valores gerados dos parâmetros, mostrando a convergência gráfica dos mesmos.

Tabela 5.4: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,05068	0,00360	0,05051	[0,04385;0,05781]	1,08262	54,592
α_1	0,41885	0,02162	0,41908	[0,37507;0,45935]	0,25302	88,874
α_2	0,19549	0,02140	0,19536	[0,15506;0,23779]	-0,38957	71,104

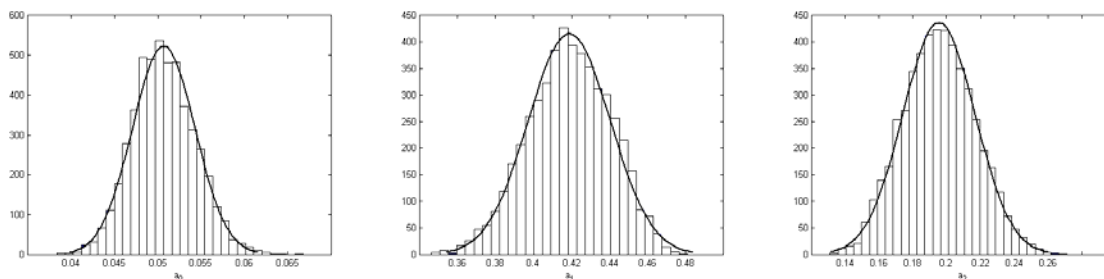


Figura 5.8: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Modelo ARCH(2).

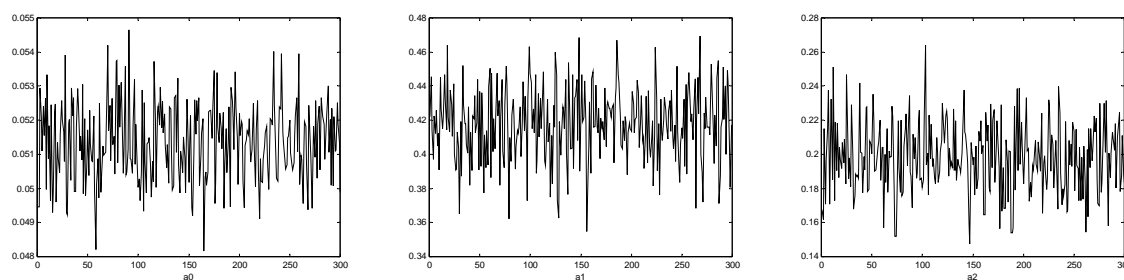
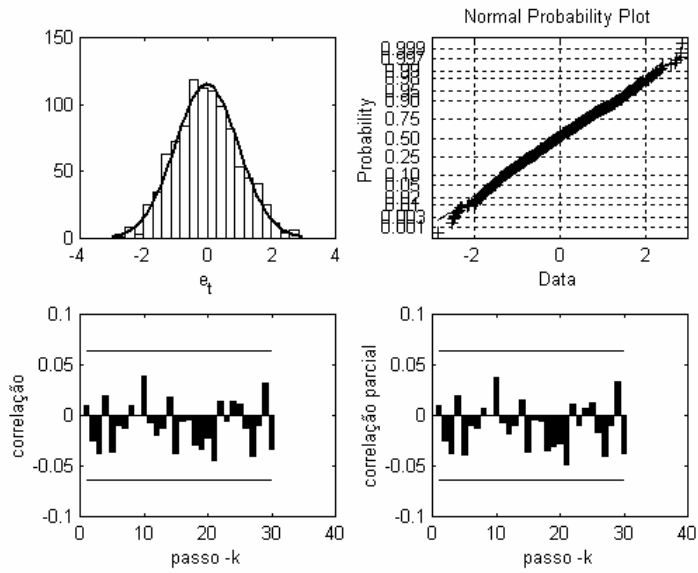


Figura 5.9: Valores gerados dos parâmetros considerando distribuição a priori não-informativa - Modelo ARCH(2).

Os gráficos da Figura 5.10 resumem a análise de resíduos para o modelo selecionado na abordagem Bayesiana com distribuição a priori não-informativa. Observamos que os resíduos para este caso também estão normalmente distribuídos e pouco correlacionados. A Figura 5.11 mostra a volatilidade estimada a partir das estimativas obtidas pela abordagem Bayesiana com distribuição a priori não-informativa (Bayes - PNI) e das estimativas obtidas pela abordagem de Máxima Verossimilhança (MV).



**Figura 5.10: Análise de resíduos - Modelo ARCH(2)
Abordagem Bayesiana com distribuição a priori não-informativa.**

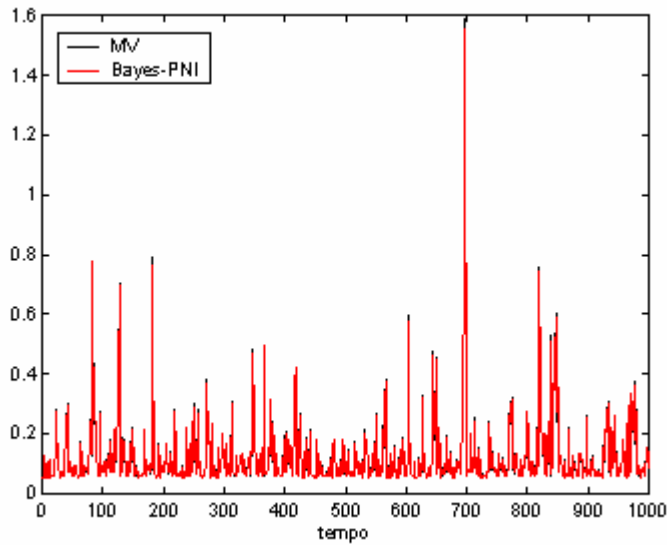


Figura 5.11: Volatilidade Estimada - Modelo ARCH(2).

Os resultados resumidos nas Tabelas 5.2, 5.3 e 5.4 mostram que as estimativas de Máxima Verossimilhança e as estimativas Bayesianas estão próximas, mas que as últimas estimam intervalos de credibilidade com amplitudes menores do que os intervalos de confiança.

5.2.2. Série Índice Bovespa (IBovespa)

O Índice da Bolsa de Valores de São Paulo (IBovespa) é o mais importante indicador do desempenho médio das cotações do mercado de ações brasileiro. Sua relevância vem do fato do IBovespa retratar o comportamento dos principais papéis negociados na Bovespa e também de sua tradição, pois o índice manteve a integridade de sua série histórica e não sofreu modificações metodológicas desde sua implementação em 1968. A finalidade básica do IBovespa é a de servir como indicador médio do comportamento do mercado, ou seja, ele representa o comportamento médio das principais ações transacionadas, e o perfil das negociações à vista observadas nos pregões da Bovespa. Pela importância econômica notória, o comportamento do IBovespa tem sido amplamente estudado na literatura. Dentre os trabalhos desenvolvidos podemos citar Ferreira (2001) e Morettin & Tolo (2004).

A série histórica que vamos analisar apresenta informações sobre índices de fechamento da Bovespa registrados no período de 02/01/1996 a 01/02/1999. Os retornos instantâneos foram calculados através do logaritmo neperiano do quociente entre os valores de fechamento em dois dias consecutivos, expresso pela equação (1.2). Na Figura 5.12 estão os dados originais e os retornos calculados do IBovespa. As Figuras 5.13 e 5.14 mostram o comportamento da série de retornos do IBovespa. O gráfico da função de autocorrelação indica a presença de uma pequena correlação entre os valores de y_t . A existência de correlações pode contribuir para a imprecisão no ajuste de modelos ARCH(q) à série analisada, mas, visto que estas correlações são pequenas ($< |0,15|$), consideramos y_t como um processo não-correlacionado. Podemos observar ainda que a série dos quadrados destes retornos, y_t^2 , é correlacionada. Este comportamento pode ser associado ao de modelos ARCH(q).

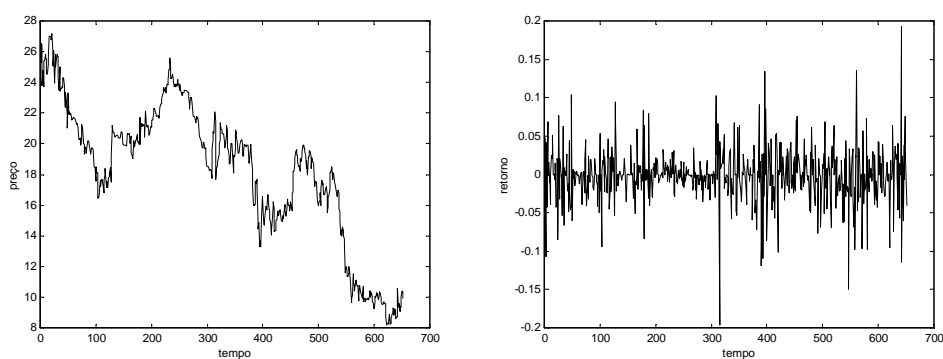


Figura 5.12: Preço e Retorno - Série IBovespa.

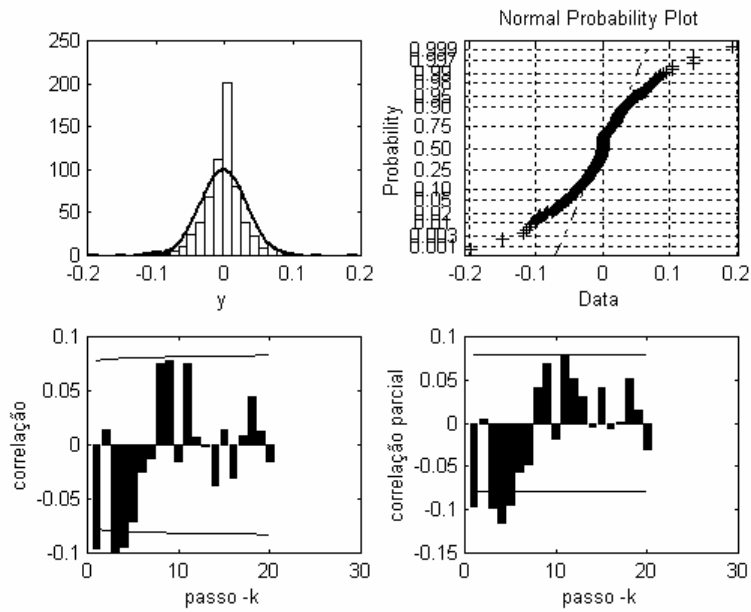


Figura 5.13: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t - Série IBOvespa.

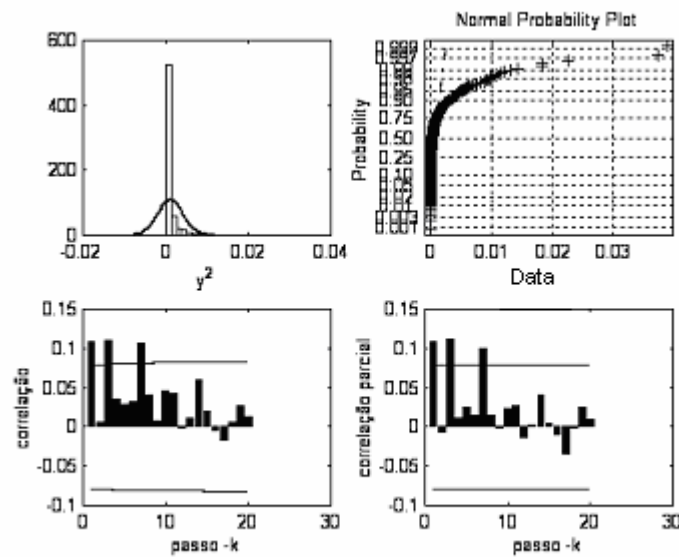


Figura 5.14: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t^2 - Série IBOvespa.

O modelo ajustado para a série IBOvespa, de acordo com os critérios AIC, BIC e CPO, é o modelo ARCH(3), dado como:

$$z_t = \left(\alpha_0 + \alpha_1 z_{t-1}^2 + \alpha_2 z_{t-2}^2 + \alpha_3 z_{t-3}^2 \right)^{\frac{1}{2}} \varepsilon_t \quad (5.1)$$

A Tabela 5.5 mostra os valores obtidos para os critérios de seleção de modelos considerados:

Tabela 5.5: Critérios de seleção de modelos - Série IBovespa.

Ordem do modelo	AIC	BIC	CPO – PI	CPO – PNI
$q=2$	-3,95768	-3,93704	6,18085e+009	6,98496e+014
$q=3$	-4,03422	-4,00671	6,18086e+017	1,17336e+018
$q=4$	-4,03829	-4,00389	3,61081e+017	3,29299e+017

Escolhida a ordem do modelo, temos o resumo da estimação dos parâmetros do processo ARCH(3) ajustado à série IBovespa. A Tabela 5.6 apresenta os intervalos de confiança dos estimadores de Máxima Verossimilhança e os intervalos de confiança obtidos pela técnica Bootstrap, respectivamente, ambos com 95%.

Tabela 5.6: Estimativas de MV e Intervalos de Confiança de 95% - Série IBovespa.

Parâmetro	EMV	Intervalo de Confiança(95%)	Intervalo de Confiança Bootstrap (95%)
α_0	0,00062	[0,00051;0,00073]	[0,00054;0,00078]
α_1	0,15578	[0,05738;0,25418]	[0,04532;0,27163]
α_2	0,08240	[-0,00414;0,16066]	[0,00885;0,17609]
α_3	0,30628	[0,17219;0,44035]	[0,14654;0,44029]

Podemos observar que o parâmetro α_2 apresenta intervalo com limitante inferior negativo (coluna 3 da Tabela 5.6), o que nos leva erroneamente à suposição de que este parâmetro não é significativamente diferente de zero.

Para corrigir intervalos largos obtidos através da abordagem de Máxima Verossimilhança, que podem levar a limitantes negativos, usamos o procedimento Bootstrap paramétrico para melhorarmos os intervalos de confiança para os parâmetros estimados do modelo ARCH(3) em estudo. Com as estimativas de Máxima Verossimilhança encontradas, fizemos 1000 reamostragens do modelo considerado e obtivemos uma amostra Bootstrap de $\{\hat{\alpha}^{*(j)} = \alpha_0^*(j), \alpha_1^*(j), \dots, \alpha_q^*(j), j = 1, \dots, 1000\}$. Assim, encontramos os intervalos Bootstrap dados pelos percentis 2,5% e 97,5% das estimativas amostrais. Podemos observar que o procedimento melhorou especificamente o intervalo referente ao parâmetro α_2 (coluna 4 da Tabela 5.6).

A partir da amostra de estimativas de Máxima Verossimilhança obtida através do método Bootstrap, construímos os histogramas que mostram a distribuição empírica dos estimadores de Máxima Verossimilhança (EMV), dados pela Figura 5.15 a seguir.

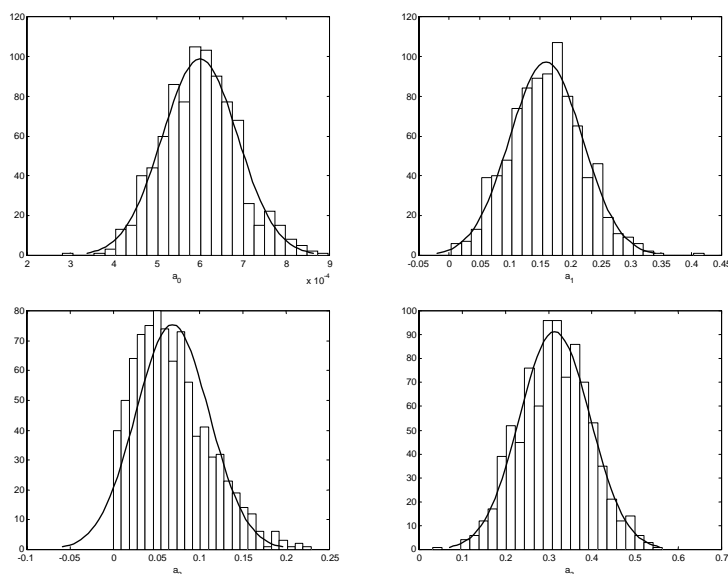


Figura 5.15: Distribuição Empírica dos EMV - Série IBovespa.

A Tabela 5.7 mostra os intervalos de credibilidade de 95% dos estimadores Bayesianos considerando distribuição a priori informativa e, a Tabela 5.8, os resultados referentes à distribuição a priori não informativa.

Tabela 5.7: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,00064	0,00002	0,00057	[0,00059;0,00068]	0,03220	73,658
α_1	0,15857	0,01608	0,16862	[0,12752;0,18999]	0,03714	93,104
α_2	0,08976	0,01279	0,07421	[0,06480;0,11431]	0,02015	80,140
α_3	0,31966	0,02518	0,31329	[0,27042;0,36933]	0,00930	89,138

Tabela 5.8: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação(%)
α_0	0,00061	0,00005	0,00061	[0,00052;0,00070]	-0,43370	55,592
α_1	0,16350	0,01113	0,16344	[0,14187;0,18514]	1,08254	95,058
α_2	0,06065	0,01139	0,06053	[0,03885;0,08257]	-0,77216	92,714
α_3	0,33162	0,02134	0,33172	[0,28888;0,37152]	-1,15740	91,654

Os histogramas que representam a distribuição a posteriori dos estimadores Bayesianos são dados pelas Figuras 5.16 e 5.17, respectivamente.

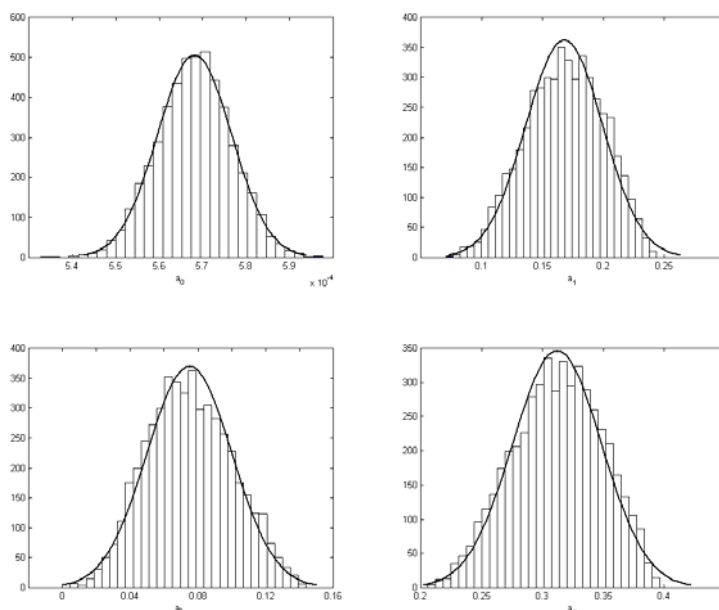


Figura 5.16: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Série IBovespa.

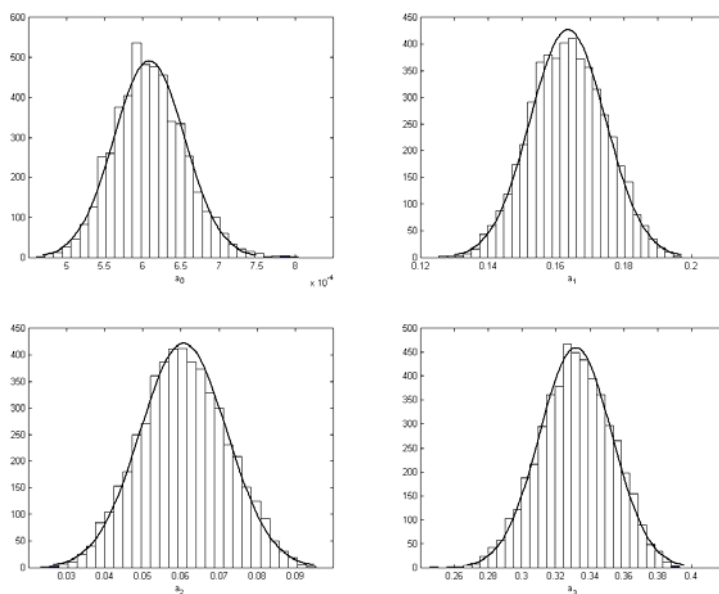
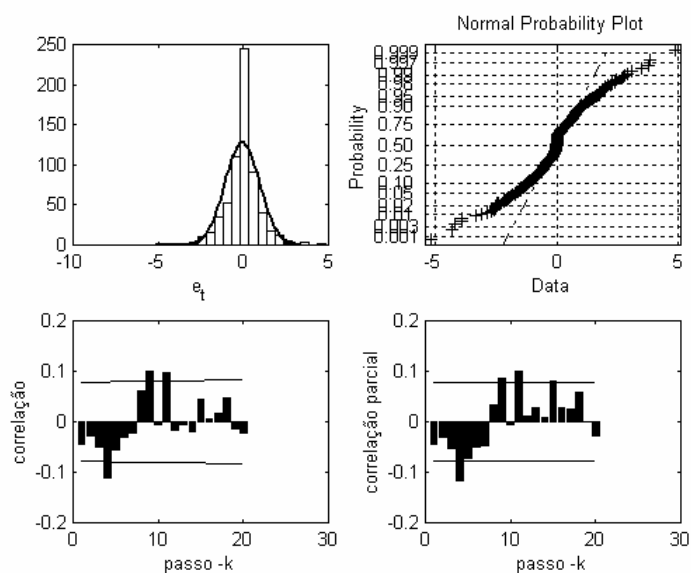


Figura 5.17: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série IBovespa.

Os gráficos referentes aos valores gerados dos parâmetros pelo algoritmo Metropolis-Hastings, mostrando a convergência gráfica dos mesmos, são apresentados no Apêndice A.4.

A análise de resíduos do modelo selecionado para a série IBovespa na abordagem Bayesiana, com distribuição a priori informativa, está na Figura 5.18. A Figura 5.19 mostra a volatilidade estimada a partir das estimativas obtidas pela abordagem Bayesiana com distribuição a priori informativa e das estimativas obtidas pela abordagem de Máxima Verossimilhança. Podemos observar que os resíduos não estão normalmente distribuídos e apresentam uma pequena correlação. Assim, podemos concluir que o modelo escolhido serve como uma aproximação que poderá ser melhorada com o ajuste da série a outro(s) modelo(s) com heterocedasticidade.

Os resíduos obtidos considerando distribuição a priori não-informativa têm um comportamento similar, assim como a volatilidade estimada, e são mostrados no Apêndice A.7.



**Figura 5.18: Análise de resíduos - Série IBovespa
Abordagem Bayesiana com distribuição a priori informativa.**

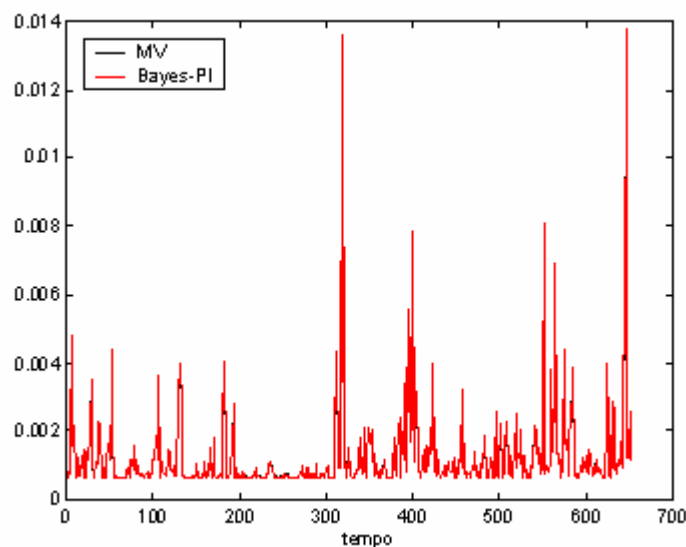


Figura 5.19: Volatilidade Estimada - Série IBovespa.

Na Figura 5.19 observamos ainda que, no período estudado, houve alguns fatos que provocaram alterações (“picos”) no comportamento da volatilidade: Moratória da Rússia, em Agosto de 1998; Crise na Ásia, em Outubro de 1998; e a Desvalorização do Real, em Janeiro de 1999.

Os resultados da estimação dos parâmetros do modelo ARCH(3), ajustado à série IBovespa, obtidos através da abordagem Bayesiana e de Máxima Verossimilhança (Tabelas 5.6, 5.7 e 5.8) mostram que as estimativas de Máxima Verossimilhança e as estimativas Bayesianas estão próximas, mas com uma precisão melhor dos intervalos de credibilidade em relação aos intervalos de confiança. Além disso, observamos a viabilidade do uso do método Bootstrap na melhoria dos intervalos de confiança padrão. Os intervalos Bootstrap fornecem estimativas condizentes com as restrições impostas aos parâmetros de modelos ARCH(q).

5.2.3. Série Telebrás

A série histórica analisada apresenta informações sobre preços das ações da Telebrás negociadas na Bolsa de Valores de São Paulo (Bovespa) no período de 02/01/1992 a 05/01/1996. Assim como a série do IBovespa, o comportamento da série de preços da Telebrás vem sendo amplamente difundida na literatura (Issler, 1999; Costa & Baidya, 2001; Ferreira, 2001). Neste estudo, os retornos instantâneos foram calculados através do logaritmo neperiano do quociente entre os preços das ações em dois dias consecutivos. Os preços e retornos (obtidos a partir da equação (1.2)) das ações da Telebrás são mostrados, a título de ilustração, na Figura 5.20.

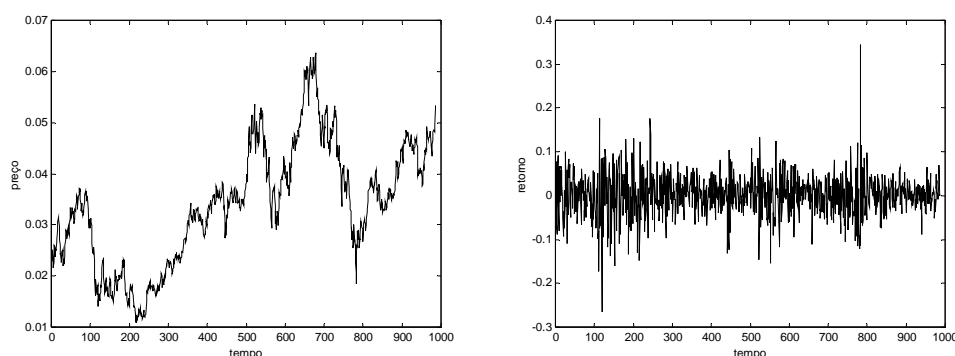


Figura 5.20: Preço e Retorno - Série Telebrás.

Os gráficos da Figura 5.21 e 5.22 mostram o comportamento da série de retornos da Telebrás, e podemos verificar que existe uma pequena correlação entre os valores da série, y_t , porém vamos considerá-la um processo não-correlacionado e ajustá-la a um modelo

ARCH(q). Além disso, podemos observar que a série dos quadrados destes retornos, y_t^2 , mostra-se correlacionada. Este comportamento pode ser associado ao de modelos ARCH(q).

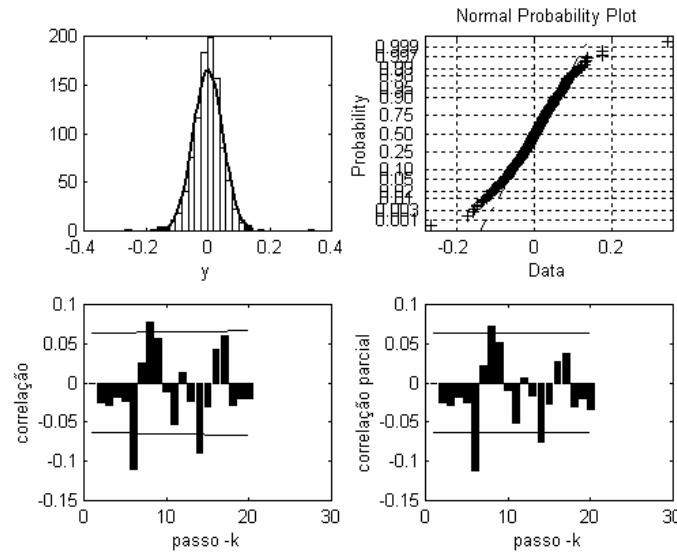


Figura 5.21: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t - Série Telebrás.

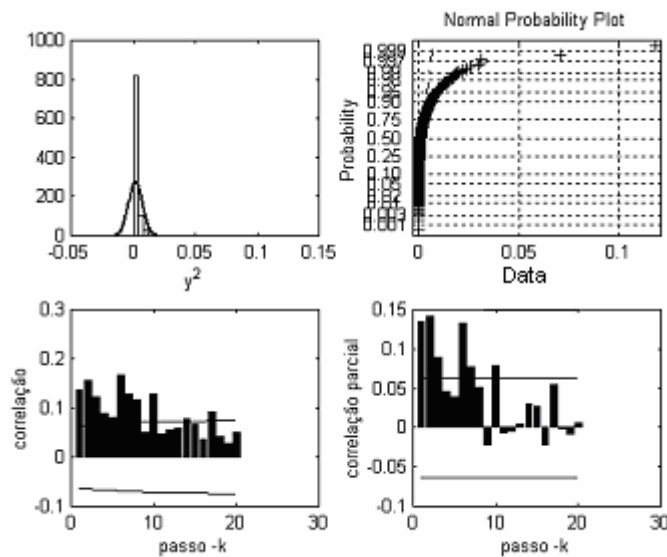


Figura 5.22: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t^2 - Série Telebrás.

O modelo ajustado para a série Telebrás, de acordo com os critérios AIC, BIC e CPO, é o ARCH(7), dado como:

$$z_t = \left(\alpha_0 + \alpha_1 z_{t-1}^2 + \alpha_2 z_{t-2}^2 + \alpha_3 z_{t-3}^2 + \alpha_4 z_{t-4}^2 + \alpha_5 z_{t-5}^2 + \alpha_6 z_{t-6}^2 + \alpha_7 z_{t-7}^2 \right)^{\frac{1}{2}} \varepsilon_t \quad (5.2)$$

A Tabela 5.9 mostra os valores obtidos para os critérios de seleção de modelos considerados:

Tabela 5.9: Critérios de seleção de modelos - Série Telebrás.

Ordem do modelo	AIC	BIC	CPO – PI	CPO – PNI
$q=6$	-3,39285	-3,35811	9,89672e+026	1,16638e+027
$q=7$	-3,39802	-3,35832	1,28314e+027	1,25315e+027
$q=8$	-3,39773	-3,35306	0,90753e+026	7,76120e+026

A Tabela 5.10 mostra os resultados da estimação dos parâmetros do modelo ARCH(7) ajustado à série Telebrás, ilustrando os intervalos dos estimadores de Máxima Verossimilhança (MV) e os intervalos Bootstrap, com 95% de confiança. Podemos observar que o parâmetro α_5 apresenta limite inferior negativo no intervalo de confiança padrão (coluna 3 da Tabela 5.10) e que o método Bootstrap paramétrico melhora as estimativas de Máxima Verossimilhança deste parâmetro (coluna 4 da Tabela 5.10).

Tabela 5.10: Estimativas de MV e Intervalos de Confiança de 95% - Série Telebrás.

Parâmetro	EMV	Intervalo de Confiança (95%)	Intervalo de Confiança Bootstrap (95%)
α_0	0,00066	[0,00046;0,00085]	[0,00042;0,00102]
α_1	0,10622	[0,02666;0,18578]	[0,01178;0,22299]
α_2	0,14749	[0,06009;0,23488]	[0,03109;0,27018]
α_3	0,10513	[0,02448;0,18578]	[0,01157;0,21718]
α_4	0,09410	[0,01533;0,17287]	[0,01155;0,21235]
α_5	0,05602	[-0,01471;0,12675]	[0,00357;0,17233]
α_6	0,10602	[0,02657;0,18546]	[0,01369;0,22360]
α_7	0,11114	[0,03073;0,19156]	[0,01372;0,23595]

A partir da amostra de estimativas obtidas através do procedimento Bootstrap, temos a representação gráfica da distribuição dos estimadores de Máxima Verossimilhança dos parâmetros da série Telebrás, dada pela Figura 5.23.

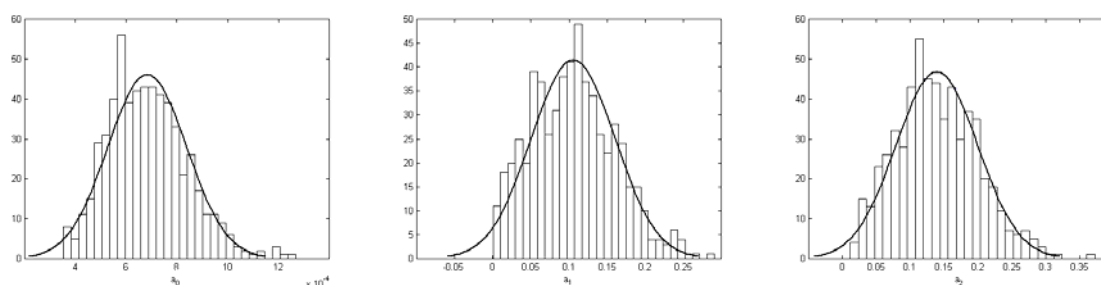


Figura 5.23: Distribuição Empírica dos EMV - Série Telebrás.

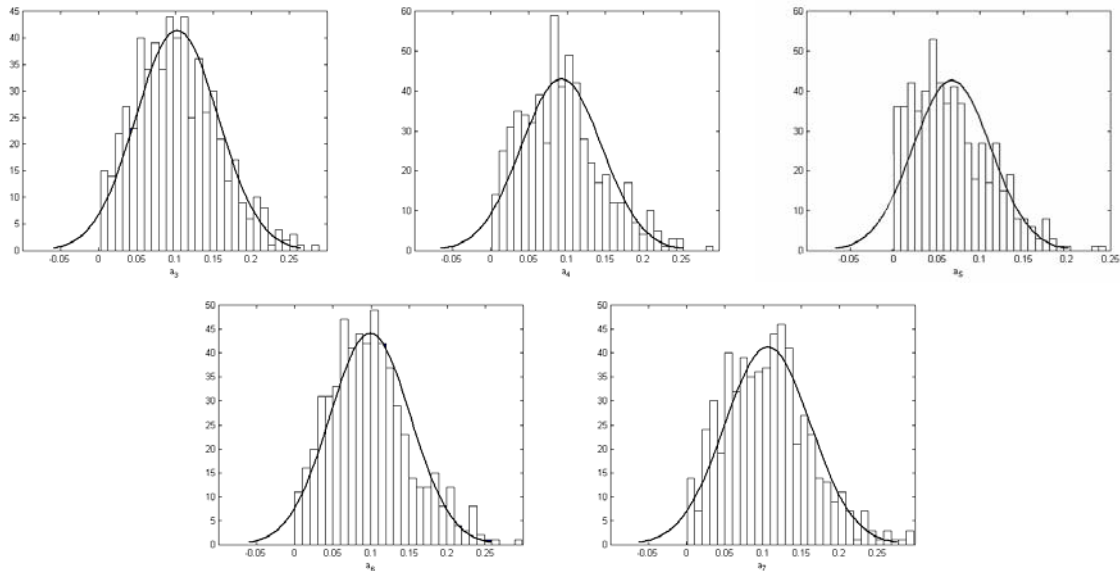


Figura 5.23 (cont.): Distribuição Empírica dos EMV - Série Telebrás.

A Tabela 5.11 e a Tabela 5.12 mostram os intervalos de credibilidade de 95% dos estimadores Bayesianos considerando distribuição a priori informativa e distribuição a priori não-informativa para os parâmetros, respectivamente.

Tabela 5.11: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,00069	0,00008	0,00069	[0,00054;0,00085]	-1,35154	52,332
α_1	0,10234	0,01922	0,10271	[0,06595;0,13691]	0,32766	84,444
α_2	0,15044	0,01952	0,15058	[0,11407;0,18682]	0,92066	86,990
α_3	0,10155	0,01918	0,10199	[0,06566;0,13677]	0,74452	83,814
α_4	0,09868	0,01946	0,09818	[0,06385;0,13534]	1,01998	84,010
α_5	0,05267	0,01903	0,05285	[0,01703;0,08759]	-0,20901	81,420
α_6	0,10252	0,01914	0,10263	[0,06675;0,13739]	0,65443	83,976
α_7	0,10327	0,01893	0,10326	[0,06735;0,13693]	0,40127	82,452

Tabela 5.12: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,00069	0,00003	0,00069	[0,00063;0,00076]	-0,93408	85,208
α_1	0,10279	0,01913	0,10311	[0,06665;0,13787]	0,04851	86,272
α_2	0,15152	0,01952	0,15162	[0,11487;0,18689]	0,75404	87,214
α_3	0,10245	0,01901	0,10279	[0,06611;0,13724]	0,15971	86,554
α_4	0,09897	0,01921	0,09839	[0,06433;0,13574]	1,32577	84,680
α_5	0,05181	0,01871	0,05195	[0,01674;0,08615]	-0,95301	84,320
α_6	0,10296	0,01912	0,10356	[0,06647;0,13743]	-0,25173	85,952
α_7	0,10389	0,01914	0,10462	[0,06775;0,13799]	-0,19425	85,568

A Figura 5.24 mostra a distribuição a posteriori dos parâmetros do modelo ARCH(7) ajustado à série Telebrás considerando distribuição a priori informativa, e a Figura 5.25, a distribuição a posteriori considerando distribuição a priori não-informativa.

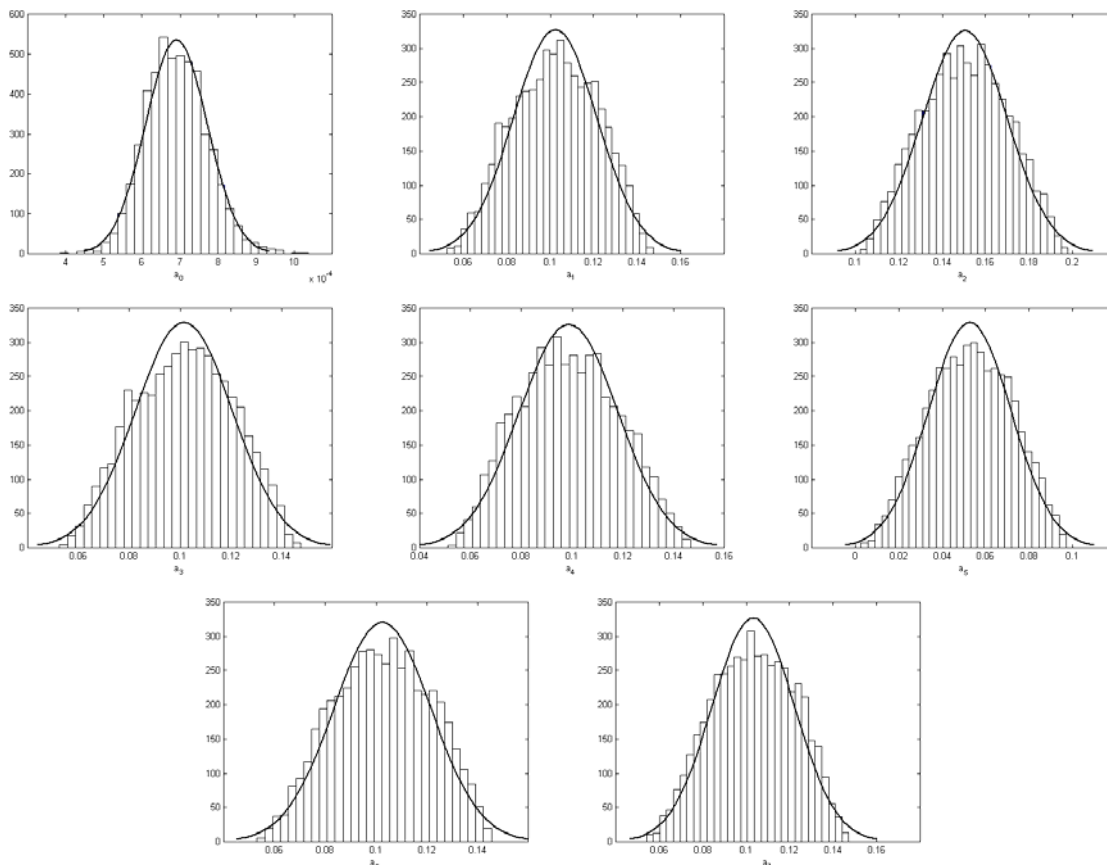


Figura 5.24: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Série Telebrás.

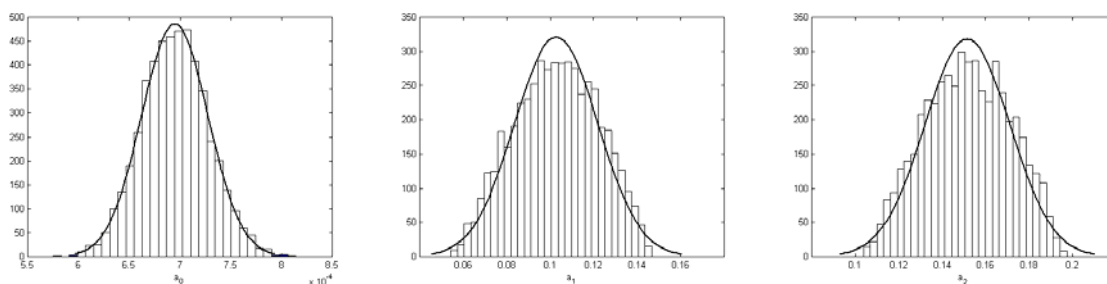


Figura 5.25: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série Telebrás.

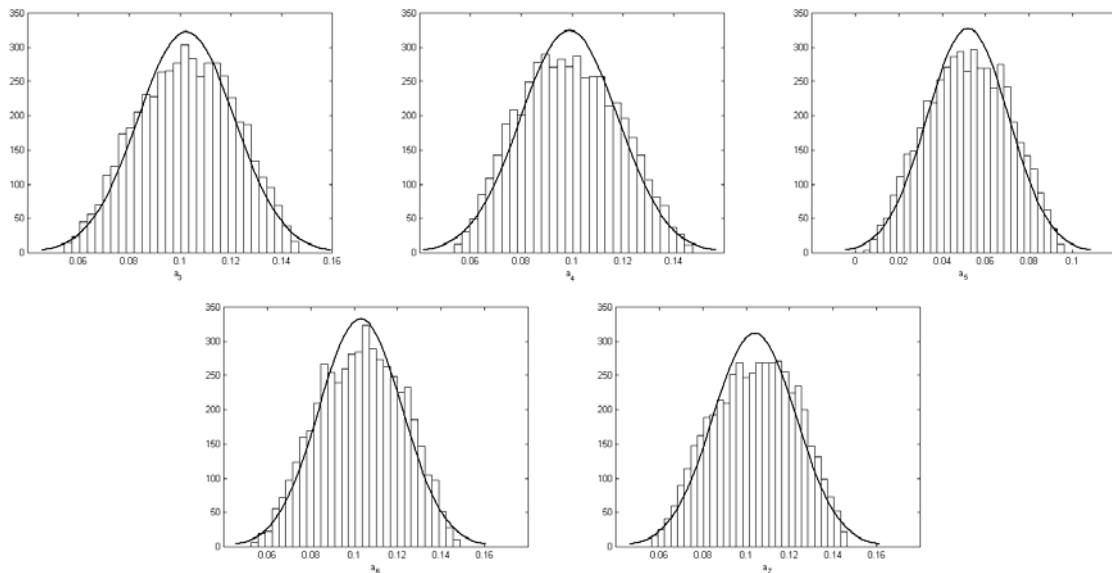


Figura 5.25 (cont.): Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série Telebrás.

A análise de resíduos do modelo selecionado para a série Telebrás na abordagem Bayesiana, com distribuição a priori informativa, está na Figura 5.26. Podemos observar que a distribuição dos resíduos é aproximadamente normal e que praticamente não apresentam correlação. Os resíduos obtidos considerando distribuição a priori não informativa têm um comportamento similar, assim como a volatilidade estimada, e são mostrados no apêndice A.7.

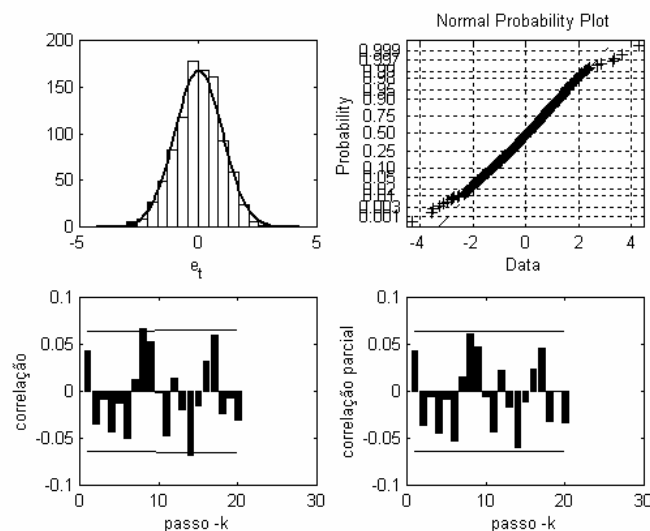


Figura 5.26: Análise de resíduos - Série Telebrás Abordagem Bayesiana com distribuição a priori informativa.

A Figura 5.27 ilustra a volatilidade estimada. Podemos observar que, no período estudado, houve alguns fatos que provocaram alterações no comportamento da volatilidade, dos quais, destaca-se a Crise do México, em Fevereiro e Março de 1995.

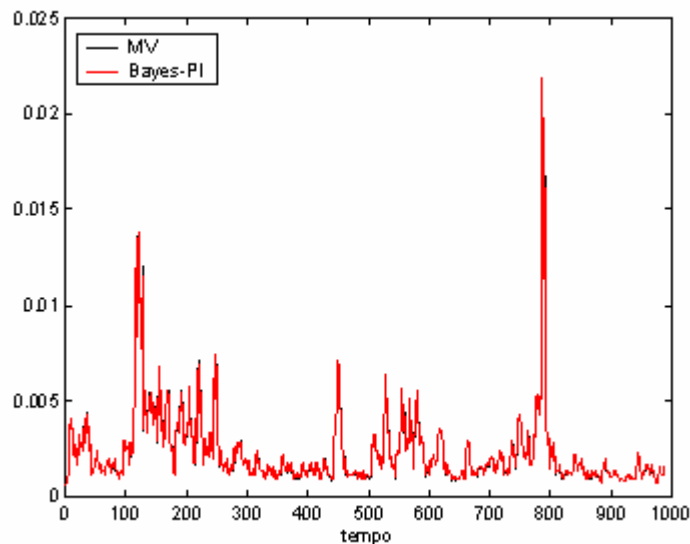


Figura 5.27: Volatilidade Estimada - Série Telebrás.

De acordo com os resultados, concluímos que o modelo escolhido reflete parcialmente o comportamento dos dados, servindo como uma aproximação que poderá ser melhorada com o ajuste da série a outro(s) modelo(s) com heterocedasticidade.

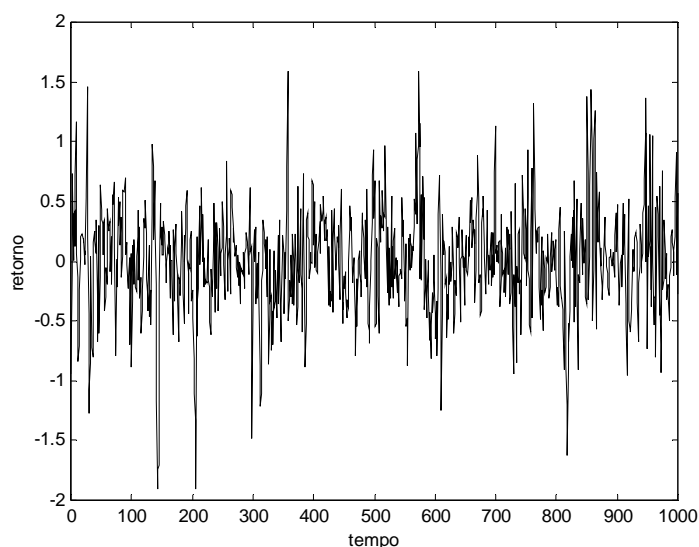
A análise da série de retornos Telebrás mostra que as estimativas de Máxima Verossimilhança e Bayesianas dos parâmetros do modelo ARCH(7) ajustado estão próximas e que os intervalos de credibilidade são mais precisos que os de confiança (Tabelas 5.10, 5.11 e 5.12). Observamos mais uma vez a melhoria dos intervalos de confiança através do método Bootstrap. Podemos notar ainda que, desde que o núcleo do algoritmo Metropolis-Hastings seja bem escolhido, levando em conta as condições do modelo, os intervalos de credibilidade também fornecerão estimativas que refletem tais imposições.

5.3. Resultados obtidos com o ajuste de Modelos AR(p)-ARCH(q)

Primeiramente, vamos apresentar os resultados obtidos para uma série gerada segundo um processo AR(1)-ARCH(3), com 1000 valores e, em seguida, os resultados obtidos para as duas séries do mercado financeiro estudadas na seção 5.2.2. Todos os conjuntos de dados foram analisados considerando as abordagens propostas para modelos AR(p)-ARCH(q) descritas neste trabalho, sob suposição de normalidade para os dados.

5.3.1. Série Gerada

A Figura 5.28 representa uma série de retornos simulada segundo um processo AR(1)-ARCH(3). Na prática, os parâmetros auto-regressivos $\beta_i, i = 1, \dots, p$, de modelos AR(p)-ARCH(q), podem apresentar valores inferiores ao considerado. Neste trabalho consideramos um valor *mediano* ($\beta_1 = 0,50$), dadas as restrições, apenas com a finalidade de confirmar (ou ilustrar) a validação da metodologia estudada.



**Figura 5.28: Modelo AR(1)-ARCH(3) com $\alpha_0 = 0,05, \alpha_1 = 0,40,$
 $\alpha_2 = 0,25, \alpha_3 = 0,10$ e $\beta_1 = 0,50$.**

A Figura 5.29 a seguir mostra o comportamento da série de retornos y_t . Os gráficos de autocorrelação e de autocorrelação parcial indicam que a série é correlacionada, com um comportamento associado aos modelos AR(p).

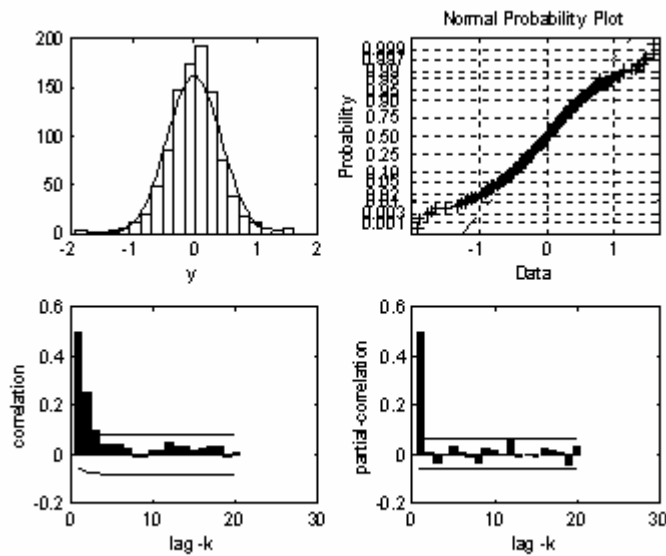


Figura 5.29: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de y_t - Modelo AR(1)-ARCH(3).

Nas Figuras 5.30 e 5.31 podemos observar ainda que a série z_t (obtida por mínimos quadrados) é não-correlacionada, e que z_t^2 é correlacionada. Este comportamento pode ser associado ao de modelos ARCH(q).

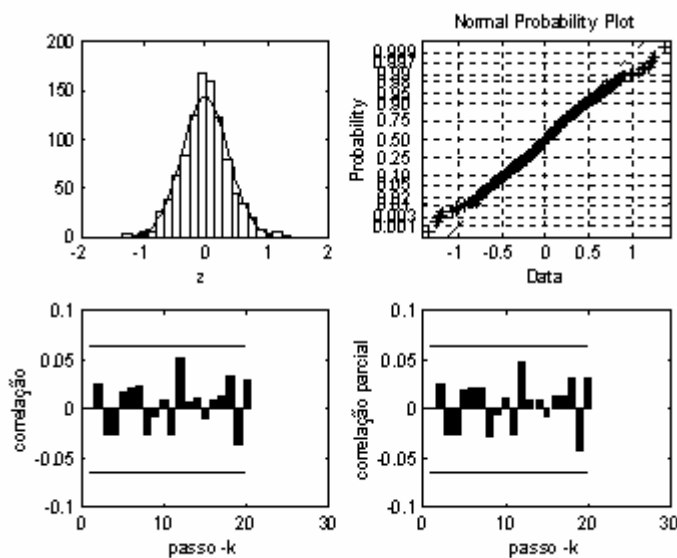


Figura 5.30: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t - Modelo AR(1)-ARCH(3).

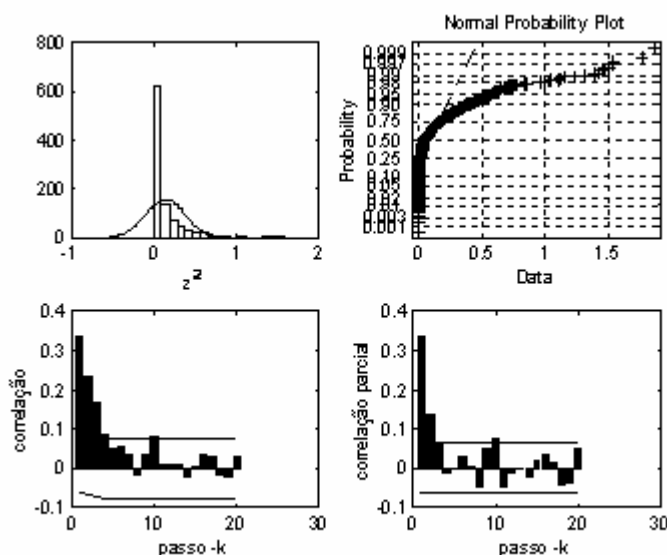


Figura 5.31: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t^2 - Modelo AR(1)-ARCH(3).

Na Tabela 5.13 temos os valores dos critérios de seleção obtidos para algumas ordens de modelos propostas. Como era esperado, concluímos que o modelo AR(1)-ARCH(3) é o modelo mais adequado para representar os dados, pois apresenta os menores valores de AIC e BIC, respectivamente, e os maiores valores para o CPO na análise Bayesiana, com distribuição a priori informativa e distribuição a priori não informativa para os parâmetros do modelo.

Tabela 5.13: Critérios de seleção de modelos – Modelo AR(p)-ARCH(q).

Ordem do modelo	AIC	BIC	CPO - PI	CPO - PNI
$p=1; q=2$	0,72660	0,74623	3,50045e-005	3,12206e-005
$p=1; q=3$	0,71621	0,74075	3,79486e-005	4,28877e-005
$p=1; q=4$	0,71730	0,74675	3,61014e-005	4,11854e-005

Escolhida a ordem do modelo mais adequada aos dados, apresentamos as estimativas dos parâmetros obtidas através da abordagem de Máxima Verossimilhança. Os resultados são mostrados na Tabela 5.14 a seguir.

Tabela 5.14: Estimativas de MV e Intervalos de Confiança de 95%.

Parâmetro	Estimativa	Intervalo de Confiança (95%)
α_0	0,04991	[0,03961;0,06021]
α_1	0,37542	[0,26022;0,49063]
α_2	0,20876	[0,11242;0,30510]
α_3	0,11350	[0,03729;0,18971]
β_1	0,48882	[0,43127;0,54638]

Na abordagem Bayesiana, através da distribuição a priori informativa proposta, as estimativas dos parâmetros obtidas são dadas pela Tabela 5.15. A convergência dos parâmetros foi verificada pelo critério de Geweke, a qual pode ser observada para os mesmos.

Tabela 5.15: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,04950	0,00347	0,04947	[0,04326;0,05635]	0,60617	70,716
α_1	0,38869	0,03553	0,38637	[0,32583;0,46058]	0,53107	76,460
α_2	0,23147	0,03395	0,22909	[0,17403;0,30216]	-0,07655	68,158
α_3	0,10770	0,02912	0,10681	[0,05089;0,16597]	1,36241	64,966
β_1	0,48745	0,02439	0,48749	[0,43953;0,53369]	-0,52467	65,946

Os gráficos da Figura 5.32 mostram a forma das distribuições a posteriori, construídos a partir da amostra selecionada para os parâmetros $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ e β_1 e, os da Figura 5.33, os últimos trezentos valores gerados dos mesmos, mostrando a convergência gráfica.

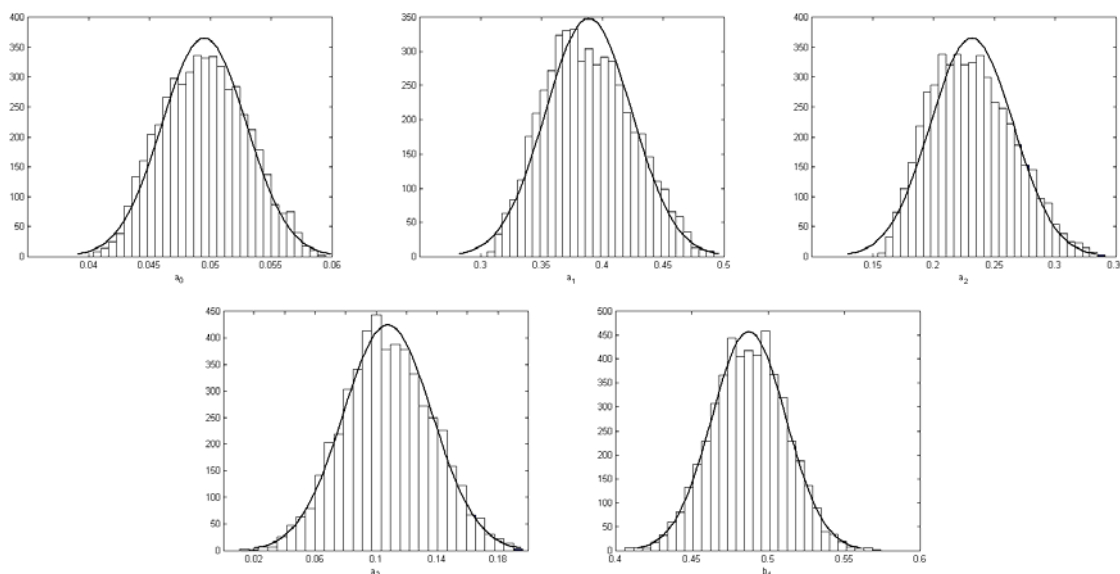


Figura 5.32: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Modelo AR(1)-ARCH(3).

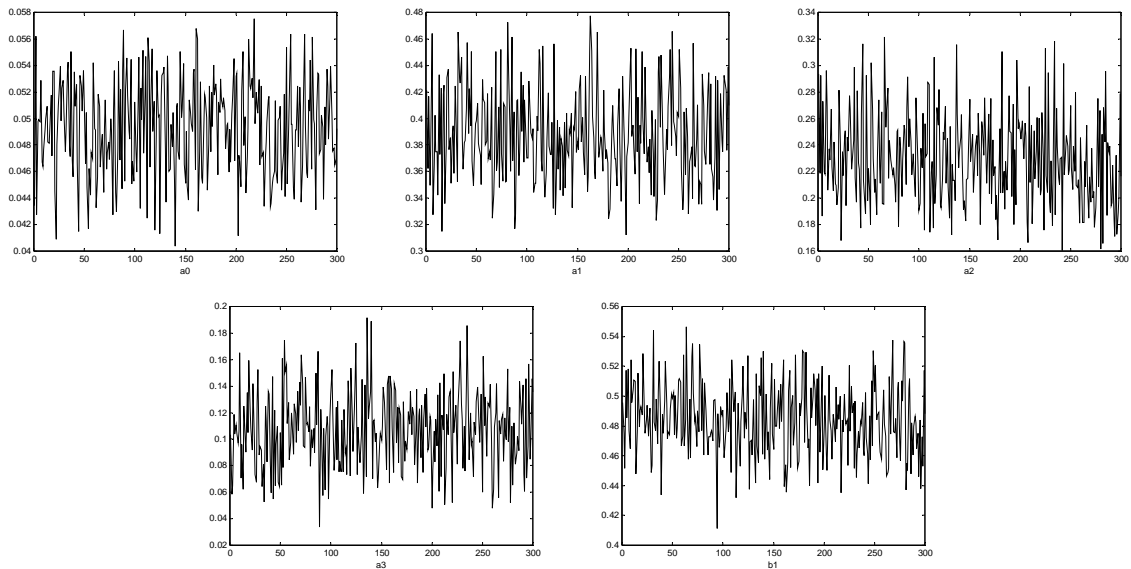


Figura 5.33: Valores gerados dos parâmetros considerando distribuição a priori informativa - Modelo AR(1)-ARCH(3).

Os gráficos da Figura 5.34 resumem a análise de resíduos para o modelo selecionado na abordagem Bayesiana com distribuição a priori informativa. Observamos que os resíduos estão normalmente distribuídos e pouco correlacionados. Portanto, podemos confirmar que o modelo escolhido é adequado. A volatilidade estimada é dada pela Figura 5.35, que mostra a volatilidade obtida a partir das estimativas Bayesianas, usando distribuição a priori informativa, comparada com a volatilidade obtida pelas estimativas de Máxima Verossimilhança. Os resíduos observados considerando distribuição a priori não informativa têm um comportamento similar, assim como a volatilidade estimada, e são mostrados no apêndice A.7.

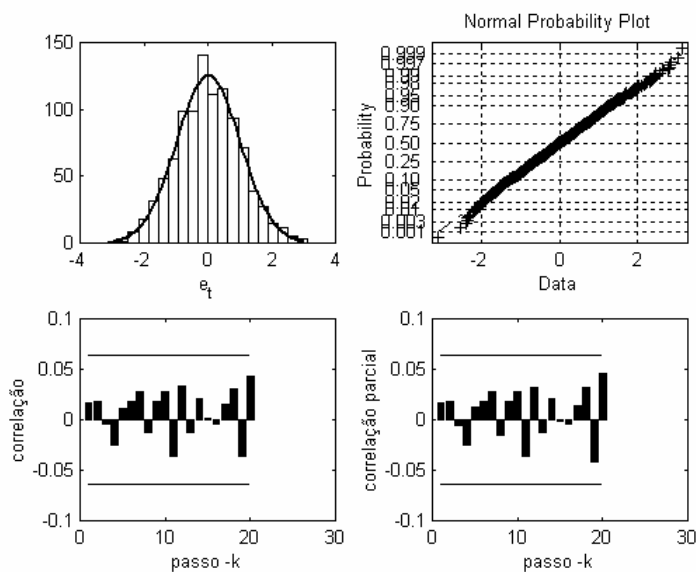


Figura 5.34: Análise de resíduos - Modelo AR(1)-ARCH(3) Abordagem Bayesiana com distribuição a priori informativa.

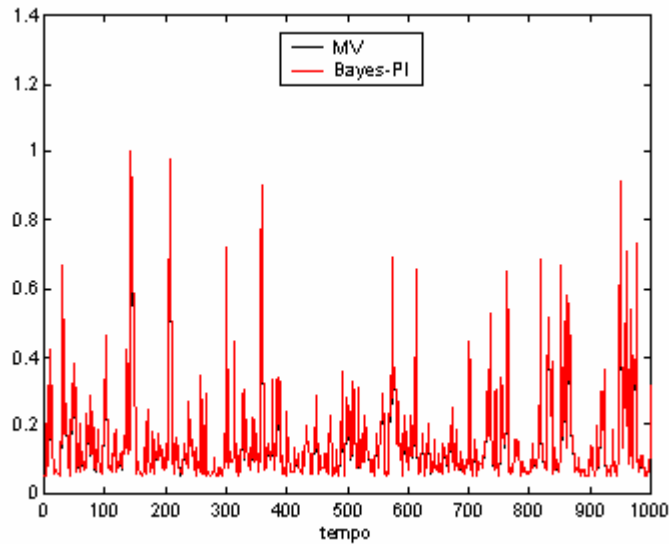


Figura 5.35: Volatilidade Estimada - Modelo AR(1)-ARCH(3).

Para a distribuição a priori não-informativa, as estimativas dos parâmetros obtidas estão apresentadas na Tabela 5.16. A convergência dos parâmetros foi confirmada pelo critério de Geweke, assim como foi observado para a distribuição a priori informativa.

Tabela 5.16: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,04941	0,00416	0,04929	[0,04149;0,05778]	1,37503	53,722
α_1	0,39809	0,03701	0,39713	[0,33192;0,46920]	-1,43647	70,684
α_2	0,23379	0,03406	0,23135	[0,17589;0,30549]	-1,24664	62,142
α_3	0,10606	0,03004	0,10614	[0,04885;0,16501]	-0,25653	65,190
β_1	0,47170	0,05299	0,47308	[0,36698;0,57167]	-0,51638	60,046

Os histogramas construídos a partir da amostra selecionada para os parâmetros do modelo simulado AR(1)-ARCH(3) são dados pela Figura 5.36 a seguir. Podemos observar também os valores gerados dos mesmos na Figura 5.37, mostrando a convergência gráfica.

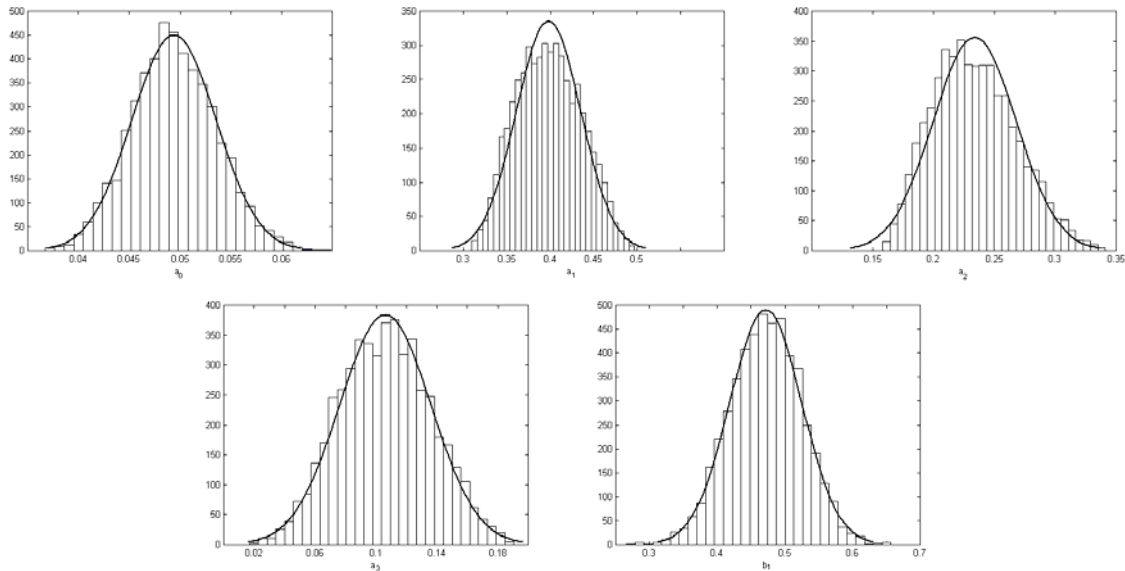


Figura 5.36: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Modelo AR(1)-ARCH(3).

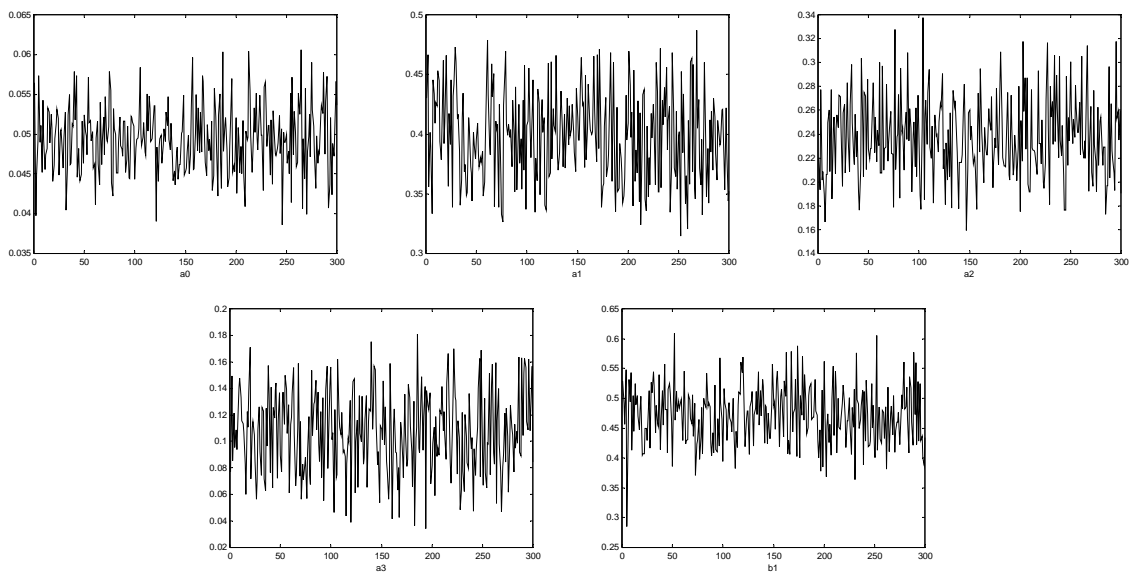


Figura 5.37: Valores gerados dos parâmetros considerando distribuição a priori não-informativa - Modelo AR(1)-ARCH(3).

Os gráficos da Figura 5.38 resumem a análise de resíduos para o modelo selecionado pela abordagem Bayesiana com distribuição a priori não-informativa. Observamos que os resíduos para este caso também estão normalmente distribuídos e pouco correlacionados. Portanto, podemos confirmar a adequação do modelo aos dados. A Figura 5.39 mostra a volatilidade estimada a partir das estimativas obtidas.

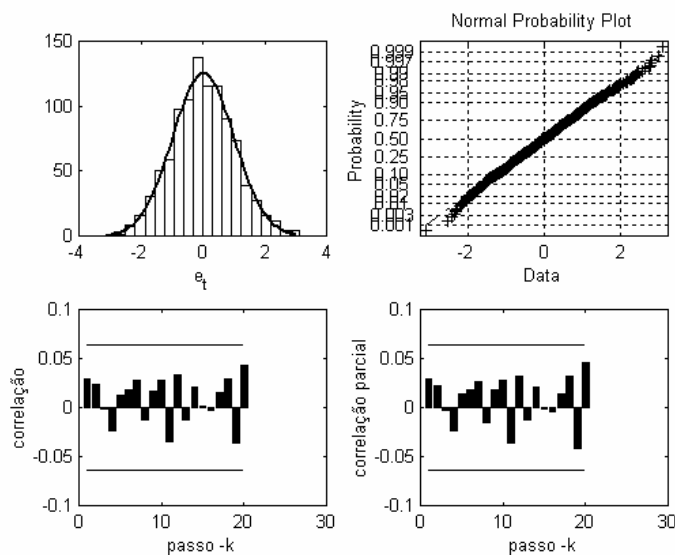


Figura 5.38: Análise de resíduos - Modelo AR(1)-ARCH(3).

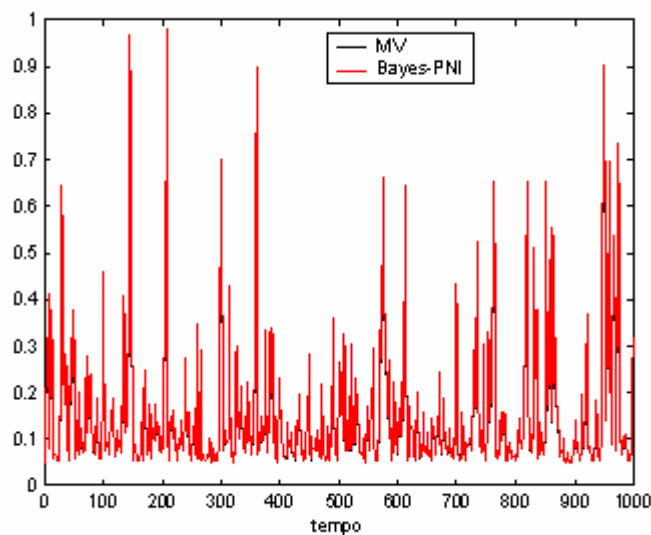


Figura 5.39: Volatilidade Estimada - Modelo AR(1)-ARCH(3).

Os resultados resumidos nas Tabelas 5.14, 5.15 e 5.16 mostram que as estimativas de Máxima Verossimilhança e as estimativas Bayesianas estão próximas, mas que as últimas estimam melhor os parâmetros considerados, principalmente no que diz respeito aos parâmetros α_2 e α_3 , pois os intervalos de credibilidade têm menores amplitudes do que os intervalos de confiança.

5.3.2. Séries Índice Bovespa (IBovespa)

Podemos observar que os gráficos de autocorrelação e de autocorrelação parcial da série de retornos do IBovespa (Figuras 5.13) mostram um comportamento comumente associado aos modelos auto-regressivos (AR), ou seja, as autocorrelações decrescem exponencialmente. Além disso, as Figuras 5.40 e 5.41 mostram ainda que a série z_t (obtida por mínimos quadrados) é pouco correlacionada, e que z_t^2 apresenta correlações superiores às observadas na seção 5.2.2 (ajuste da série ao modelo ARCH(3)).

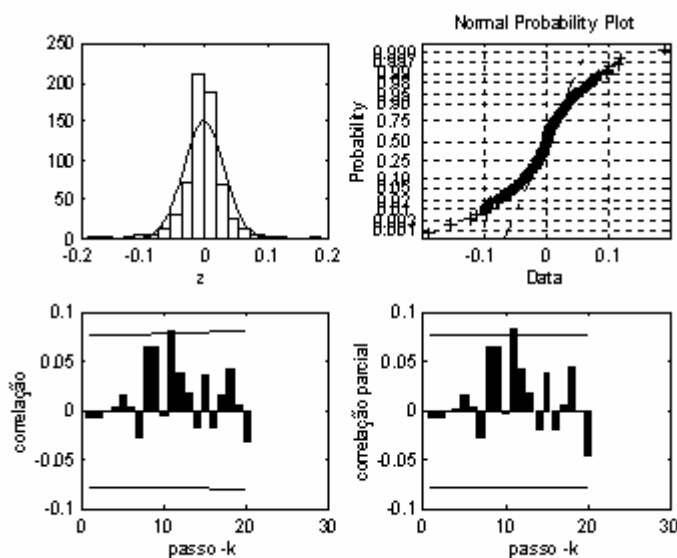


Figura 5.40: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t - Série IBovespa.

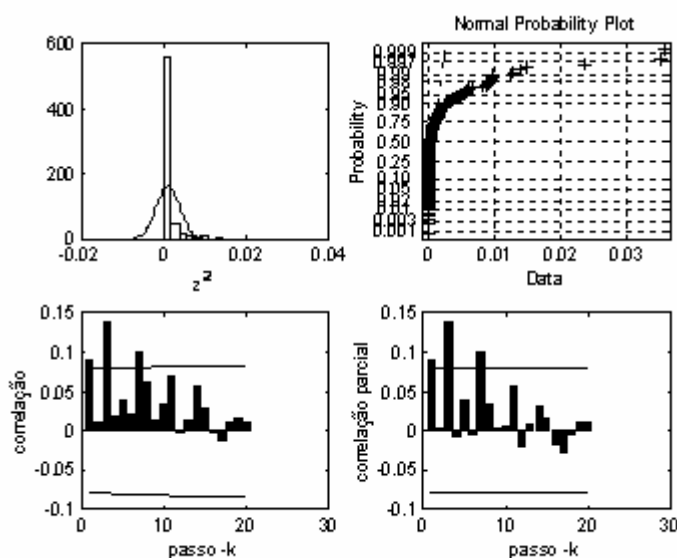


Figura 5.41: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t^2 - Série IBovespa.

Este comportamento nos leva a propor o modelo AR(p)-ARCH(q) com o objetivo de melhorar a qualidade do ajuste. O modelo ajustado para a série analisada, de acordo com os critérios AIC, BIC e da CPO, é o modelo AR(6)-ARCH(3), dado como:

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \beta_4 y_{t-4} + \beta_5 y_{t-5} + \beta_6 y_{t-6} + z_t \quad (5.3)$$

sendo $z_t = (\alpha_0 + \alpha_1 z_{t-1}^2 + \alpha_2 z_{t-2}^2 + \alpha_3 z_{t-3}^2)^{\frac{1}{2}} \varepsilon_t$.

Da mesma forma como foi realizada para modelos ARCH(q), usamos o procedimento Bootstrap paramétrico para melhorar os intervalos de confiança para os parâmetros $\alpha_j, j = 0, \dots, q$ e $\beta_i, i = 1, \dots, p$ do modelo AR(6)-ARCH(3) ajustado à série IBovespa. Com as estimativas de Máxima Verossimilhança encontradas, fizemos 1000 reamostragens do modelo considerado e obtivemos uma amostra Bootstrap de $\{\hat{\alpha}^{*(j)} = \alpha_0^*(j), \alpha_1^*(j), \dots, \alpha_q^*(j), j = 1, \dots, 1000\}$ e de $\{\hat{\beta}^{*(j)} = \beta_0^*(j), \beta_1^*(j), \dots, \beta_p^*(j), j = 1, \dots, 1000\}$. Assim, encontramos os intervalos Bootstrap dados pelos percentis 2,5% e 97,5% das estimativas amostrais.

A Tabela 5.17 apresenta os intervalos de confiança dos estimadores de Máxima Verossimilhança e os intervalos de confiança obtidos pela técnica Bootstrap, respectivamente, ambos com 95%. Podemos observar que o parâmetro α_2 apresenta intervalo de confiança padrão com limitante inferior negativo (coluna 3 da Tabela 5.17), sendo corrigido posteriormente pelo método Bootstrap (coluna 4 da Tabela 5.17).

Tabela 5.17: Estimativas de MV e Intervalos de Confiança de 95% - Série IBovespa.

Parâmetro	EMV	Intervalo de Confiança (95%)	Intervalo de Confiança Bootstrap (95%)
α_0	0,00049	[0,00040;0,00059]	[0,00037;0,00065]
α_1	0,21102	[0,09948;0,32256]	[0,08195;0,34170]
α_2	0,06866	[-0,00531;0,14263]	[0,00373;0,18377]
α_3	0,39070	[0,24089;0,54051]	[0,21960;0,55264]
β_1	-0,08928	[-0,17469;-0,00387]	[-0,17719;-0,00669]
β_2	-0,02663	[-0,10604;0,05277]	[-0,11362;0,05983]
β_3	-0,13199	[-0,22227;-0,04172]	[-0,22148;-0,04683]
β_4	-0,16463	[-0,23376;-0,09549]	[-0,24596;-0,08307]
β_5	-0,12197	[-0,18924;-0,05469]	[-0,19640;-0,04441]
β_6	-0,07454	[-0,14217;-0,00691]	[-0,14866;0,00162]

A partir das amostras obtidas através do método Bootstrap, construímos os histogramas que mostram a distribuição empírica dos estimadores de Máxima Verossimilhança, dados pela Figura 5.42 a seguir.

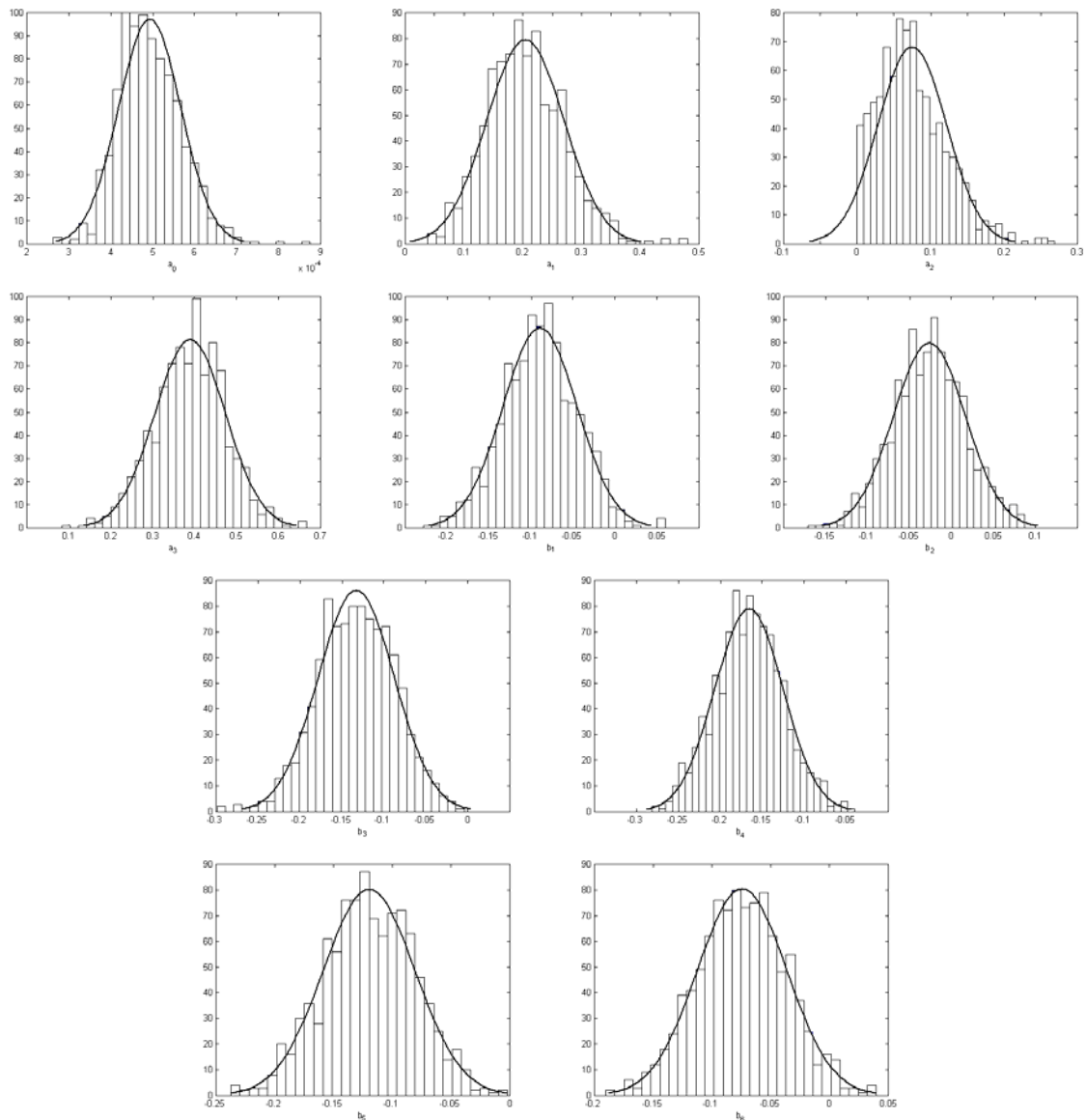


Figura 5.42: Distribuição Empírica dos EMV - Série IBOvespa.

A Tabela 5.18 mostra as estimativas Bayesianas e os intervalos de credibilidade de 95% dos parâmetros considerando distribuição a priori informativa, e a Tabela 5.19 apresenta os resultados referentes à distribuição a priori não-informativa. A convergência dos parâmetros foi verificada pelo critério de Geweke, a qual pode ser observada para os mesmos.

Tabela 5.18: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,00048	0,00004	0,00049	[0,00041;0,00057]	0,22296	55,232
α_1	0,21184	0,04731	0,21049	[0,12234;0,30339]	0,61625	68,000
α_2	0,08894	0,03336	0,08682	[0,03141;0,15847]	-1,18057	71,980
α_3	0,39839	0,05213	0,39724	[0,30219;0,49826]	-0,68755	78,622
β_1	-0,08992	0,01955	0,08998	[-0,12862;-0,05166]	-1,86104	69,274
β_2	-0,02868	0,01926	0,02869	[-0,06696;0,00945]	1,35876	68,850
β_3	-0,13119	0,01978	0,13151	[-0,17048;-0,09243]	0,36285	69,014
β_4	-0,16006	0,01834	0,16009	[-0,19625;-0,12469]	0,97973	69,402
β_5	-0,12139	0,01837	0,12142	[-0,15717;-0,08537]	-0,12622	68,886
β_6	-0,07123	0,01855	0,07159	[-0,10760;-0,03429]	1,19000	69,274

Tabela 5.19: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,00048	0,00003	0,00048	[0,00042;0,00056]	-0,86412	65,788
α_1	0,21579	0,04142	0,21648	[0,13543;0,29508]	-1,33013	75,024
α_2	0,08565	0,02706	0,08445	[0,03649;0,14026]	-0,40036	78,604
α_3	0,41482	0,04587	0,41589	[0,32397;0,49956]	1,31147	79,974
β_1	-0,09571	0,02879	-0,09571	[-0,15259;-0,03821]	0,30797	85,904
β_2	-0,01679	0,02853	-0,01664	[-0,07421;0,03983]	-0,57483	77,570
β_3	-0,13483	0,02876	-0,13548	[-0,19079;-0,07799]	1,63708	83,690
β_4	-0,16456	0,02804	-0,16457	[-0,21835;-0,10972]	1,61275	83,392
β_5	-0,14604	0,02958	-0,14617	[-0,20397;-0,08949]	-0,61480	54,870
β_6	-0,08765	0,02993	-0,08733	[-0,14678;-0,02814]	-0,98251	68,838

Os histogramas que representam a distribuição a posteriori dos estimadores Bayesianos são dados pelas Figuras 5.43 e 5.44, respectivamente. Os gráficos referentes aos valores gerados dos parâmetros pelo algoritmo Metropolis-Hastings, mostrando a convergência gráfica dos mesmos, são apresentados no Apêndice A.5.

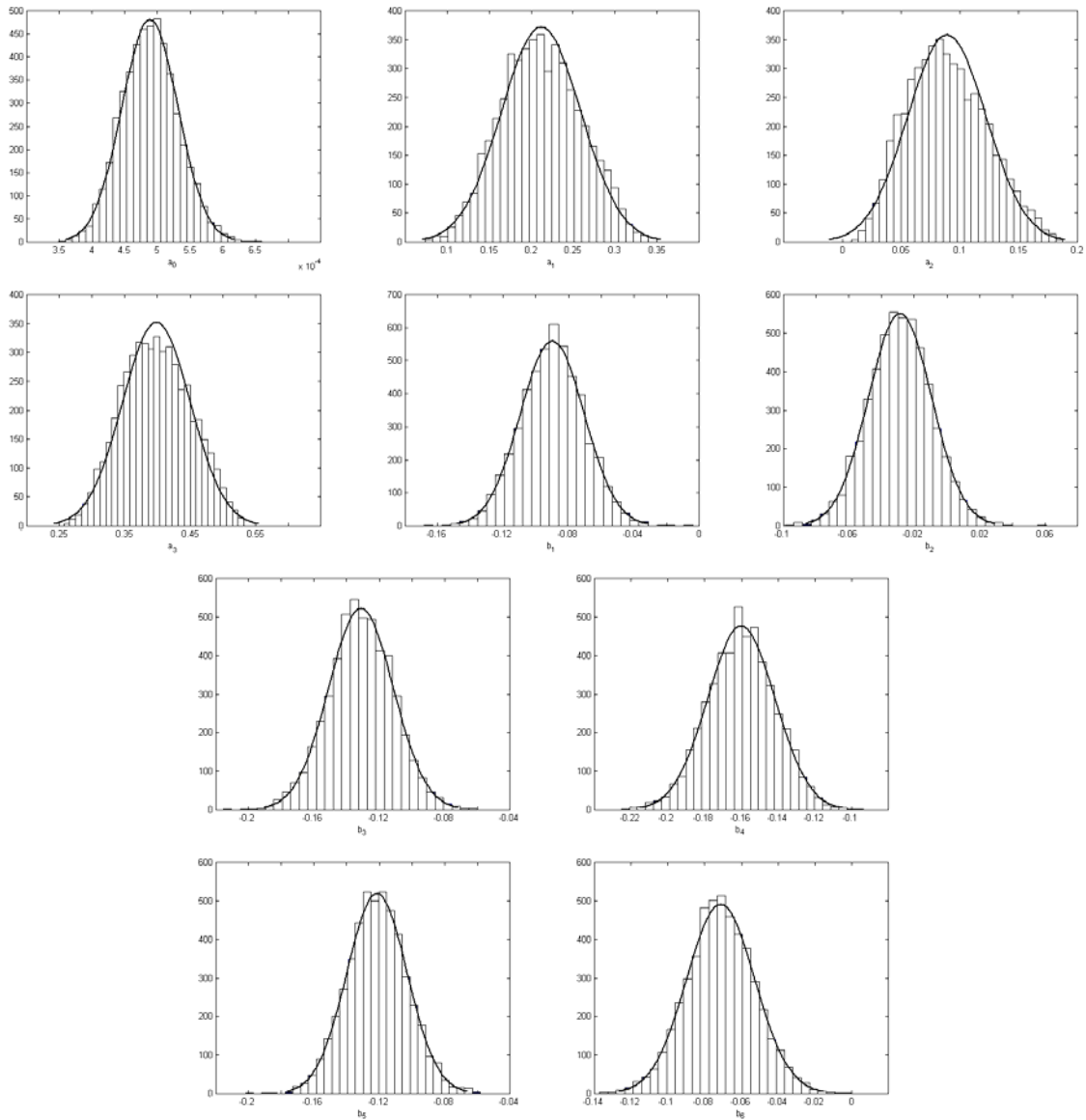


Figura 5.43: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Série IBovespa.

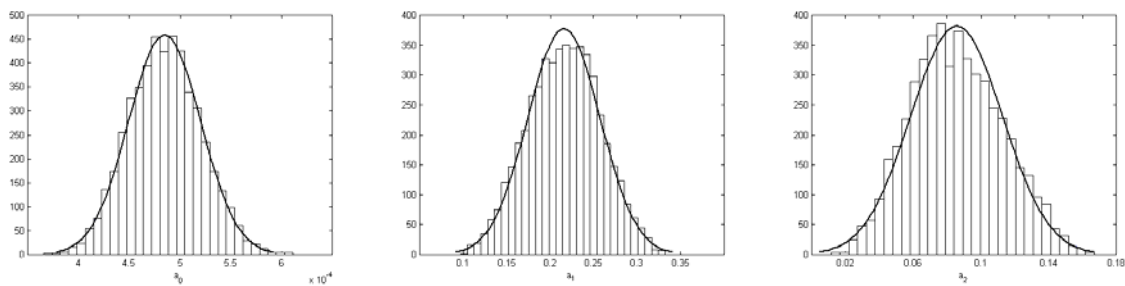


Figura 5.44: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série IBovespa.

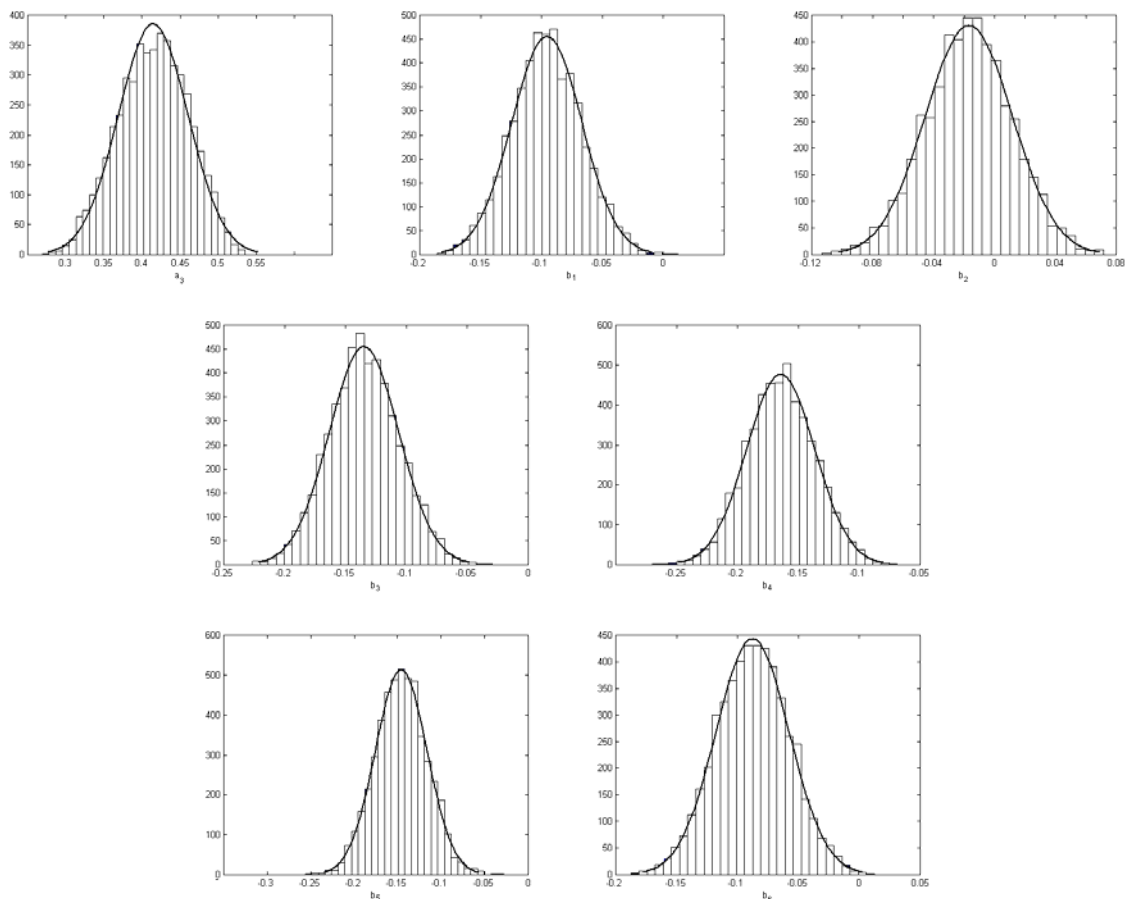
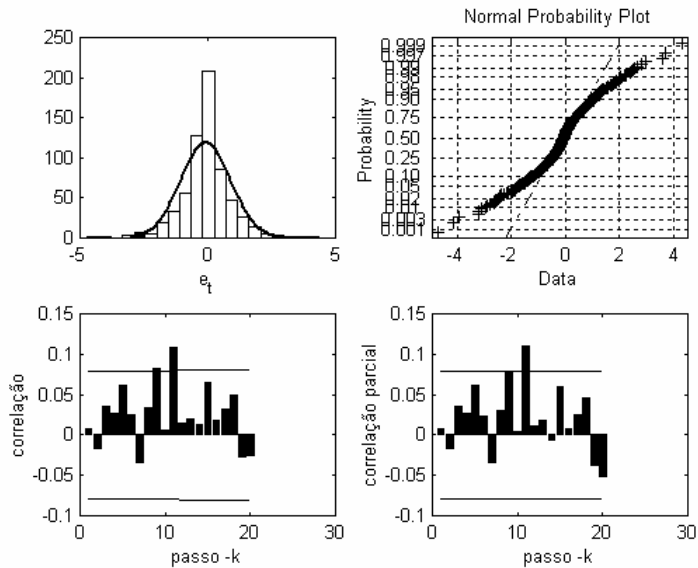


Figura 5.44 (cont.): Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série IBovespa.

A análise de resíduos do modelo selecionado para a série IBovespa na abordagem Bayesiana, com distribuição a priori informativa, está na Figura 5.45 e a volatilidade estimada a partir das estimativas obtidas pela abordagem Bayesiana com distribuição a priori informativa e das estimativas obtidas pela abordagem de Máxima Verossimilhança, está na Figura 5.46. Os resíduos encontrados considerando distribuição a priori não-informativa têm um comportamento similar, assim como a volatilidade estimada, e são dados no Apêndice A.7.

Podemos observar que os resíduos não estão normalmente distribuídos e apresentam uma pequena correlação, porém o comportamento dos mesmos mostra-se melhor (menos correlacionados) do que aqueles obtidos com o ajuste da série a um modelo ARCH(3), como pode ser visto nas Figuras 5.18 e 5.45.



**Figura 5.45: Análise de resíduos - Série IBOvespa
Abordagem Bayesiana com distribuição a priori informativa.**

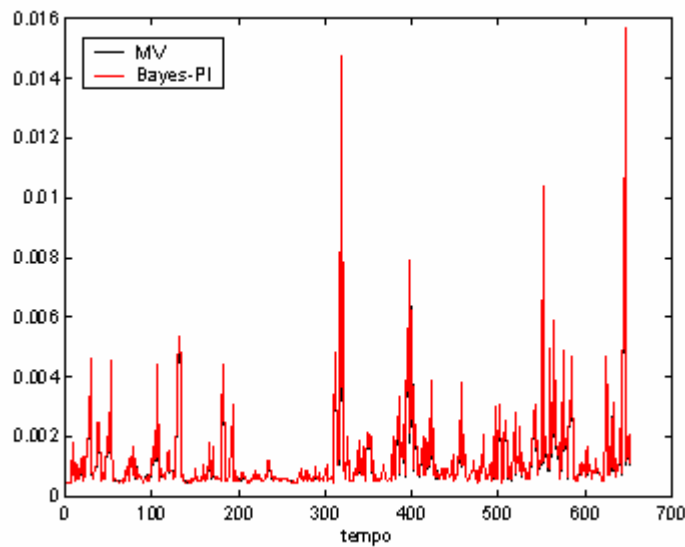


Figura 5.46: Volatilidade Estimada - Série IBOvespa.

A Tabela 5.20 mostra os valores dos critérios de seleção de modelos obtidos pelo ajuste da série de retornos do IBOvespa a um modelo ARCH(q) e, posteriormente, a um modelo AR(p)-ARCH(q).

Tabela 5.20: Critérios de seleção de modelos – Série IBOvespa.

Modelo	AIC	BIC	CPO – PI	CPO - PNI
ARCH(3)	-4,03422	-4,00671	6,18086e+017	1,17336e+018
AR(6)-ARCH(3)	-4,11931	-4,05052	1,32448e+019	1,45479e+019

De acordo com os valores obtidos através dos critérios de seleção, podemos concluir que o modelo que melhor se ajusta à série analisada é o modelo AR(6)-ARCH(3), confirmando a conclusão descrita através da análise gráfica dos resíduos.

Os resultados mostram ainda que as estimativas de Máxima Verossimilhança e Bayesianas dos parâmetros do modelo AR(6)-ARCH(3) ajustado à série IBovespa estão próximas umas das outras, exceto pelo parâmetro α_2 que apresenta uma pequena diferença entre as duas abordagens. Além disso, os intervalos de confiança são significativamente melhores com o uso do método Bootstrap, porém menos precisos que os intervalos Bayesianos.

5.3.3. Série Telebrás

Os gráficos da função de autocorrelação e da função de autocorrelação parcial da série de retornos da Telebrás (Figura 5.21) mostram que as correlações decrescem exponencialmente com componentes senoidais amortecidas. Além disso, as Figuras 5.47 e 5.48 mostram ainda que a série z_t (obtida por mínimos quadrados) é pouco correlacionada, e que z_t^2 apresenta correlações superiores às observadas na seção 5.2.3 (ajuste da série ao modelo ARCH(7)).

Com o objetivo de tentar remover a correlação serial existente nas observações da série de retornos da Telebrás, apresentamos um novo estudo onde ajustamos os mesmos a um modelo AR(p)-ARCH(q).

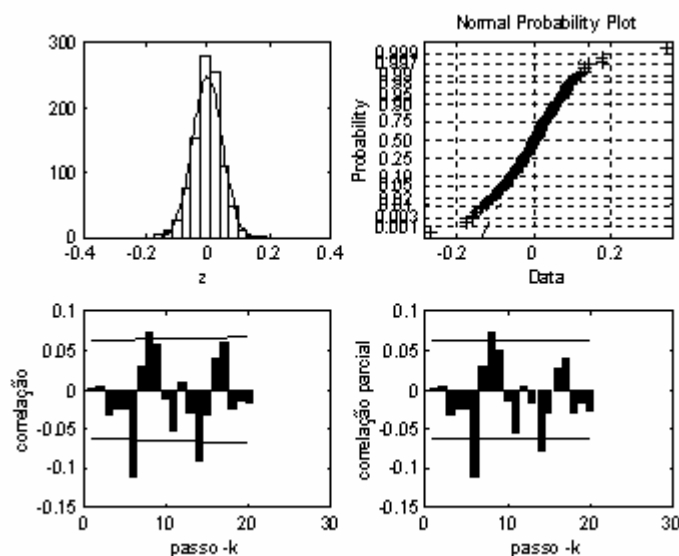


Figura 5.47: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t - Série Telebrás.

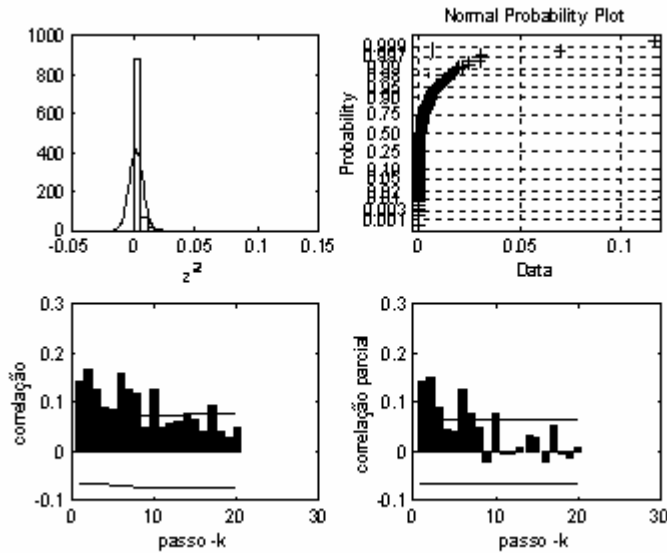


Figura 5.48: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de z_t^2 - Série Telebrás.

O modelo ajustado para a série analisada, de acordo com os critérios AIC, BIC e da CPO, é o AR(2)-ARCH(6), dado como:

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + z_t \quad (5.4)$$

onde $z_t = \left(\alpha_0 + \alpha_1 z_{t-1}^2 + \alpha_2 z_{t-2}^2 + \alpha_3 z_{t-3}^2 + \alpha_4 z_{t-4}^2 + \alpha_5 z_{t-5}^2 + \alpha_6 z_{t-6}^2 \right)^{\frac{1}{2}} \varepsilon_t$.

A Tabela 5.21 mostra os intervalos de Máxima Verossimilhança (MV) e os intervalos Bootstrap, com 95% de confiança, para os parâmetros estimados do modelo AR(2)-ARCH(6) ajustado à série Telebrás, onde podemos observar que o intervalo para o parâmetro α_4 é melhorado pelo método Bootstrap (colunas 3 e 4).

Tabela 5.21: Estimativas de MV e Intervalos de Confiança de 95% – Série Telebrás.

Parâmetro	EMV	Intervalo de Confiança (95%)	Intervalo de Confiança Bootstrap (95%)
α_0	0,00074	[0,00055;0,00094]	[0,00048;0,00111]
α_1	0,11692	[0,03608;0,19775]	[0,01742;0,24901]
α_2	0,18587	[0,09263;0,27912]	[0,04372;0,31449]
α_3	0,09161	[0,01272;0,17051]	[0,00869;0,21984]
α_4	0,06093	[-0,01205;0,13391]	[0,00339;0,17673]
α_5	0,07363	[0,00100;0,14626]	[0,00485;0,18499]
α_6	0,15535	[0,06795;0,24276]	[0,02675;0,28273]
β_1	0,06269	[-0,00235;0,12773]	[-0,02370;0,15100]
β_2	-0,05240	[-0,11902;0,01422]	[-0,14718;0,03726]

A partir das amostras obtidas através do procedimento Bootstrap temos a forma da distribuição dos estimadores de Máxima Verossimilhança dos parâmetros da série Telebrás, dada pela Figura 5.49 a seguir.

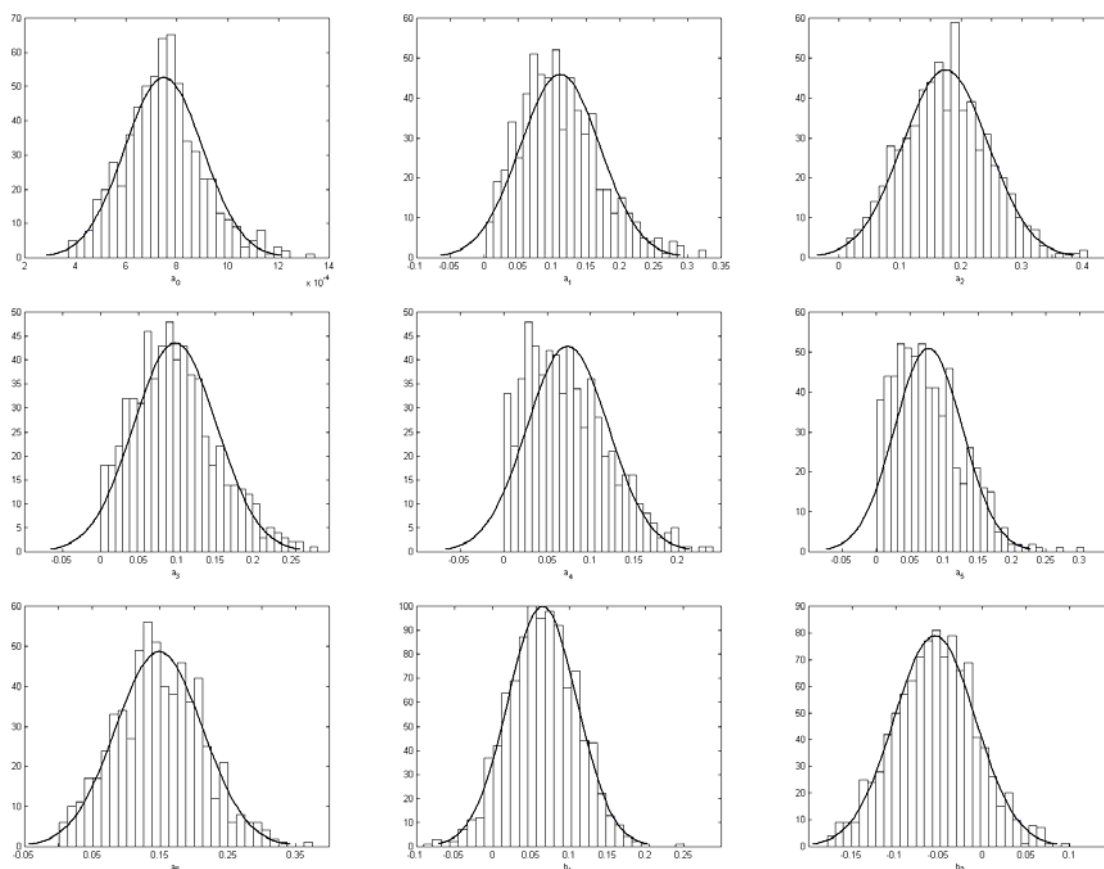


Figura 5.49: Distribuição Empírica dos EMV - Série Telebrás.

A Tabela 5.22 e a Tabela 5.23 mostram os resultados da estimação e os intervalos de credibilidade de 95% considerando distribuição a priori informativa e distribuição a priori não-informativa para os parâmetros, respectivamente.

Tabela 5.22: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,00074	0,00007	0,00074	[0,00061;0,00088]	0,23059	57,366
α_1	0,11040	0,03013	0,11017	[0,05301;0,16850]	-0,67837	63,810
α_2	0,19314	0,03449	0,19199	[0,13035;0,26161]	0,61309	71,940
α_3	0,09771	0,03181	0,09682	[0,03902;0,16067]	-1,23758	66,794
α_4	0,06836	0,02569	0,06707	[0,02310;0,12058]	-1,91867	72,618
α_5	0,08466	0,03061	0,08351	[0,03095;0,14825]	1,25154	62,970
α_6	0,15354	0,03192	0,15228	[0,09404;0,21680]	1,62437	68,744
β_1	0,06180	0,02104	0,06192	[0,02107;0,10267]	0,30245	71,646
β_2	-0,05200	0,02143	-0,05239	[-0,09534;-0,01046]	0,00558	71,772

Tabela 5.23: Estimativas Bayesianas e Intervalos de Credibilidade de 95% - PNI.

Parâmetro	Média	Desvio Padrão	Mediana	Intervalo de Credibilidade (95%)	Critério de Geweke	Taxa de Aceitação (%)
α_0	0,00073	0,00007	0,00073	[0,00060;0,00087]	0,50496	57,128
α_1	0,10763	0,02739	0,10716	[0,05655;0,16132]	-1,23008	71,514
α_2	0,20052	0,03056	0,19989	[0,14188;0,25976]	-1,48922	74,478
α_3	0,09787	0,02866	0,09652	[0,04545;0,15520]	0,04028	70,924
α_4	0,06922	0,02252	0,06843	[0,02791;0,11458]	1,01646	74,640
α_5	0,08933	0,02761	0,08796	[0,04004;0,14706]	0,51479	66,084
α_6	0,15165	0,02837	0,15126	[0,09785;0,20656]	0,12208	73,090
β_1	0,06746	0,04306	0,06768	[-0,01469;0,15223]	0,84597	88,172
β_2	-0,07918	0,04378	-0,07993	[-0,16431;0,01131]	0,48671	66,778

A Figura 5.50 apresenta a forma da distribuição a posteriori dos parâmetros do modelo AR(2)-ARCH(6) ajustado à série Telebrás considerando distribuição a priori informativa e, a Figura 5.51, considerando distribuição a priori não-informativa. Os valores gerados dos parâmetros para ambas as distribuições a priori podem ser vistos no Apêndice A.5.

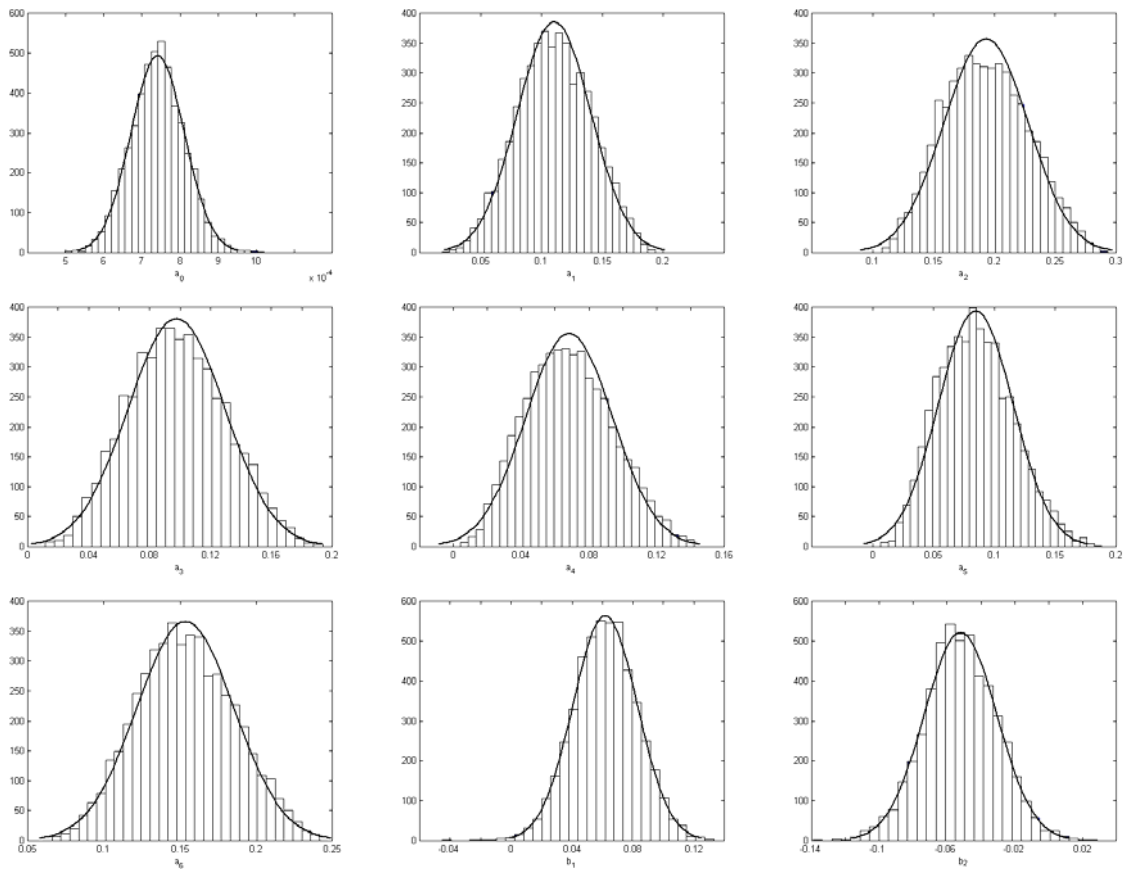


Figura 5.50: Distribuição a posteriori dos parâmetros considerando distribuição a priori informativa - Série Telebrás.

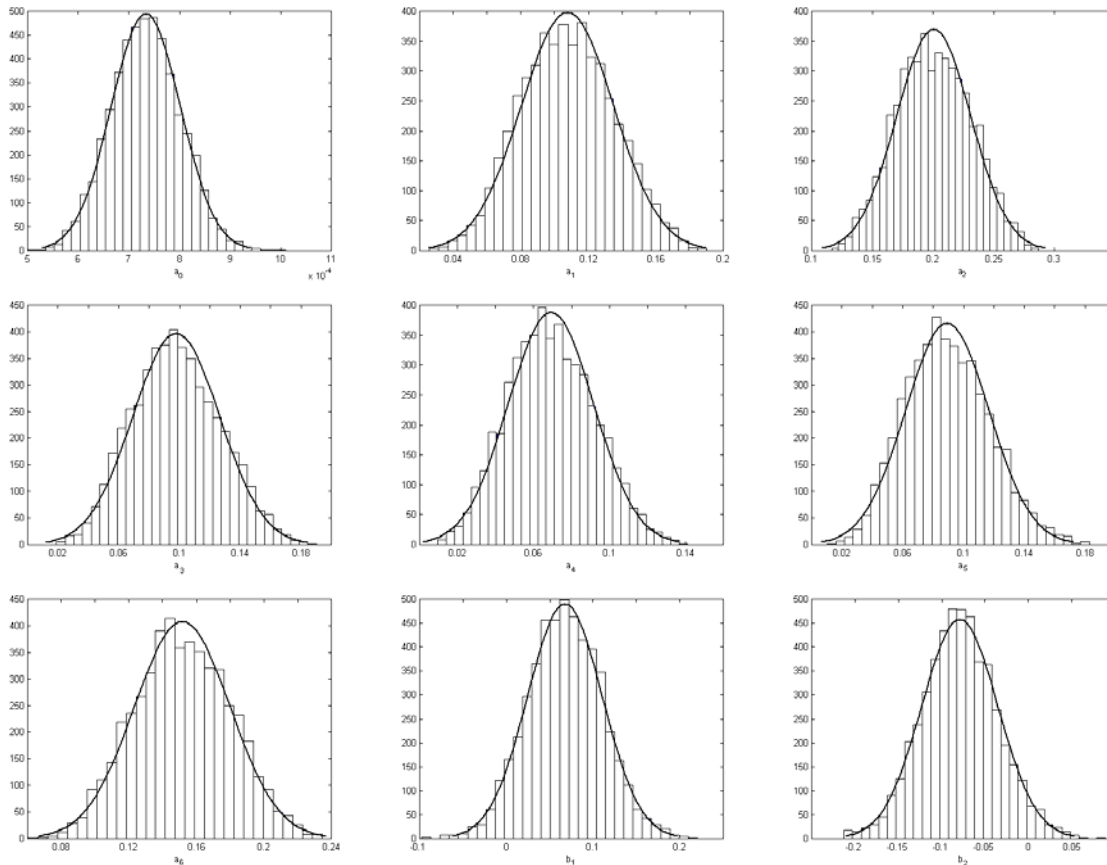


Figura 5.51: Distribuição a posteriori dos parâmetros considerando distribuição a priori não-informativa - Série Telebrás.

A análise gráfica dos resíduos do modelo selecionado para a série Telebrás na abordagem Bayesiana com distribuição a priori informativa está na Figura 5.52. A volatilidade estimada a partir das estimativas obtidas pela abordagem Bayesiana com distribuição a priori informativa e das estimativas obtidas pela abordagem de Máxima Verossimilhança, está na Figura 5.53. A distribuição dos resíduos é aproximadamente normal, e os mesmos praticamente não apresentam correlação. Os resíduos obtidos considerando distribuição a priori não-informativa têm um comportamento similar, assim como a volatilidade estimada, e são dados no apêndice A.7.

Podemos observar ainda que os resíduos mostram-se menos correlacionados do que aqueles obtidos com o ajuste da série a um modelo ARCH(7).

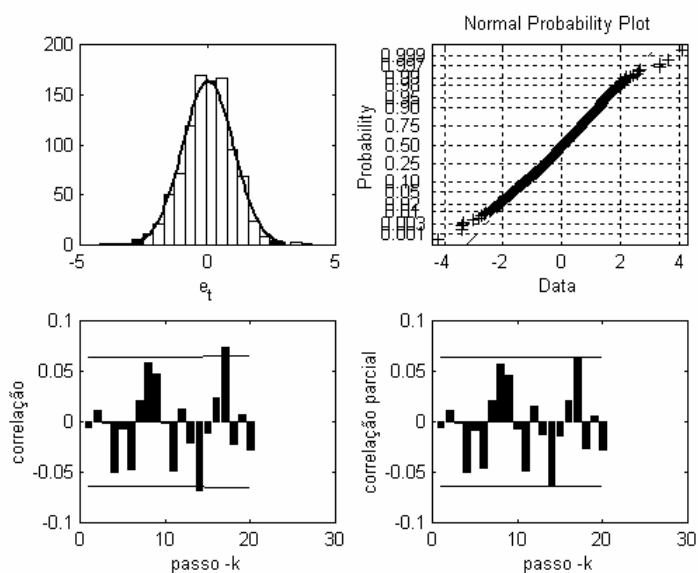


Figura 5.52: Análise de resíduos - Série Telebrás
Abordagem Bayesiana com distribuição a priori informativa.

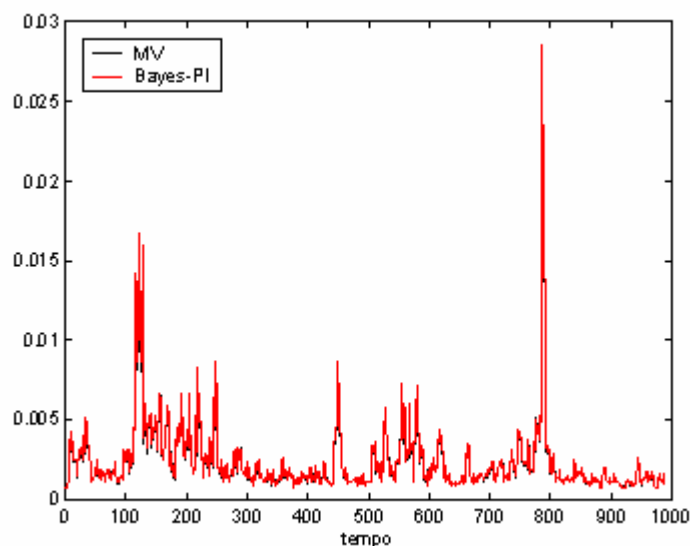


Figura 5.53: Volatilidade Estimada - Série Telebrás.

A Tabela 5.24 a seguir mostra os valores dos critérios de seleção dos modelos ajustados à série Telebrás.

Tabela 5.24: Critérios de seleção de modelos - Série Telebrás.

Modelo	AIC	BIC	CPO - PI	CPO - PNI
ARCH(7)	-3,39802	-3,35432	1,28314e+027	1,25315e+027
AR(2)-ARCH(6)	-3,39955	-3,35887	1,33424e+027	1,34514e+027

Os valores calculados para os critérios de seleção mostram que o modelo AR(2)-ARCH(6) se ajusta melhor aos dados do que o modelo ARCH(7), confirmando a conclusão descrita através da análise gráfica dos resíduos.

A análise da série de retornos Telebrás mostra ainda que as estimativas de Máxima Verossimilhança e Bayesianas dos parâmetros do modelo AR(2)-ARCH(6) ajustado à série estão próximas. Observamos também que os intervalos de confiança foram melhorados pelo método Bootstrap e que os mesmos foram mais precisos do que os Bayesianos para alguns parâmetros, no caso de distribuição a priori não-informativa. Porém, no caso de distribuição a priori informativa, podemos melhorar o ajuste do modelo AR(2)-ARCH(6) de forma significativa. Este fato mostra que a utilização de informação a priori pode nos levar a um modelo mais representativo da série, uma vez que os critérios AIC, BIC e o CPO concordam com o modelo ajustado.

CAPÍTULO 6

Um Estudo Comparativo entre as Abordagens Bayesiana e de Máxima Verossimilhança para Modelos ARCH(q)

6.1. Introdução

Neste capítulo apresentamos um estudo que tem por objetivo comparar a abordagem Bayesiana e a abordagem de Máxima Verossimilhança na inferência dos parâmetros de processos ARCH(q). Usando séries simuladas de modelos ARCH(q), esta comparação é feita através das estimativas Bayesianas obtidas das distribuições a posteriori geradas com técnicas de simulação MCMC e das estimativas de Máxima Verossimilhança dos parâmetros obtidas numericamente. Neste estudo mostramos *empiricamente* o desempenho (ou robustez) do método de estimação com relação à ordem do modelo e a desvios das condições de estacionariedade.

Para a estimação dos parâmetros α_j destes modelos, foram simuladas 500 séries independentes (seqüências diferentes do ruído branco ε_t expresso em (2.41)) para os modelos ARCH(2) e ARCH(5). Cada série tem 1000 valores originalmente, dos quais são descartados os primeiros 500 valores para evitar o efeito das condições iniciais. Os 500 valores restantes foram usados para estimar os parâmetros do modelo ARCH(q), sendo que no final cada um dos parâmetros tem 500 estimativas diferentes geradas pela abordagem de Máxima Verossimilhança. Nesta abordagem, através da técnica Bootstrap, usamos os percentis 2,5% e

97,5% das estimativas pontuais de cada parâmetro como estimativa do intervalo com 95% de confiança para cada um dos parâmetros. Na abordagem Bayesiana, cada um dos parâmetros também tem 500 estimativas diferentes geradas pelo método MCMC. Nesta abordagem, pelo mesmo procedimento, usamos os percentis 2,5% e 97,5% das estimativas obtidas pelo método MCMC, geradas da distribuição a posteriori conjunta. A média aritmética das estimativas Bayesianas representa o valor final do parâmetro $\hat{\alpha}_j, j = 0,1,\dots,q$ (seção 4.4.1). Além disso, a abordagem Bayesiana foi realizada a partir da distribuição a priori informativa proposta para modelos ARCH(q), expressa pela equação (3.19) na seção 3.2.1, e do algoritmo Metropolis-Hastings enunciado na seção 4.4.1.

Devemos ressaltar que a comparação feita neste trabalho só pode ser aplicada em situações hipotéticas, nas quais os parâmetros verdadeiros dos modelos são conhecidos. Mesmo não sendo realista, esta situação ilustra a potencialidade de cada técnica quanto à precisão na estimação dos parâmetros na ausência dos valores verdadeiros dos mesmos.

O desempenho das duas abordagens em estudo é comparado através da raiz quadrada do erro quadrático médio (RMSE) e, o erro percentual absoluto médio (MAPE) entre os valores reais dos parâmetros (α_j) e os valores estimados ($\hat{\alpha}_j$) por cada uma das técnicas. Os erros de estimação RMSE e MAPE são dados, respectivamente, por:

$$RMSE_j = \sqrt{\frac{1}{K} \sum_{i=1}^K (\alpha_j - \hat{\alpha}_j^{(i)})^2} \quad (6.1)$$

$$MAPE_j = \frac{100}{K} \sum_{i=1}^K \left| \frac{\alpha_j - \hat{\alpha}_j^{(i)}}{\alpha_j} \right| \quad (6.2)$$

em que $j = 0,1,\dots,q$ e $K = 500$, ou seja, é o tamanho da série de estimativas.

A qualidade das estimativas também é avaliada através dos intervalos de credibilidade (abordagem Bayesiana) e dos intervalos de confiança (abordagem de Máxima Verossimilhança). Como já mencionado antes, estes intervalos indicam o grau de certeza com que se pode encontrar o valor real do parâmetro dentro da faixa por eles especificada. Assim, quanto menor for a largura do intervalo, melhor é o método de estimação.

6.2. Aplicações

Os coeficientes dos modelos ARCH(2) e ARCH(5) simulados estão na Tabela 6.1. Devemos notar que o modelo ARCH(2) - M3 viola a restrição de estacionaridade na covariância ($\alpha_1 + \alpha_2 < 1$). Este modelo foi adotado propositalmente para verificarmos a robustez das duas técnicas nesta situação. Os parâmetros da transformação usada na técnica Bayesiana, definida pela expressão 3.4, são dados na Tabela 6.2.

Tabela 6.1: Parâmetros dos Modelos Simulados.

Modelo	α_0	α_1	α_2	α_3	α_4	α_5
ARCH(2) - M1	1,00	0,80	0,15	-	-	-
ARCH(2) - M2	1,00	0,50	0,20	-	-	-
ARCH(2) - M3	1,00	0,90	0,20	-	-	-
ARCH(5)	1,00	0,50	0,20	0,15	0,08	0,05

Tabela 6.2: Parâmetros da Transformação (3.4).

	M1	M2	M3	ARCH(5)
a_0, b_0	0,85;1,15	0,85;1,15	0,85;1,15	0,85;1,15
a_1, b_1	0,65;0,95	0,35;0,65	0,75;1,05	0,35;0,65
a_2, b_2	0,00;0,30	0,00;0,40	0,00;0,40	0,00;0,40
a_3, b_3	-	-	-	0,00;0,30
a_4, b_4	-	-	-	0,00;0,16
a_5, b_5	-	-	-	0,00;0,10
σ_i^2	1,00	1,00	1,00	1,00

Na implementação do algoritmo Metropolis-Hastings, para cada parâmetro foi simulada uma cadeia com 5000 iterações, descartado 50% dos valores para diminuir o efeito das condições iniciais, e então tomados valores espaçados de 5 em 5, totalizando uma amostra de tamanho 500. Este processo é repetido 500 vezes a fim de se obter 500 estimativas para cada um dos parâmetros. A convergência do algoritmo foi verificada pelo critério de Geweke (1992), a qual ocorreu sempre para todos os parâmetros.

Os gráficos da Figura 6.1 apresentam a distribuição empírica (histograma) das estimativas de Máxima Verossimilhança (MV) dos parâmetros do modelo M1. Os gráficos para os modelos M2 e M3 são semelhantes aos do modelo M1 e estão apresentados no Apêndice A.6.

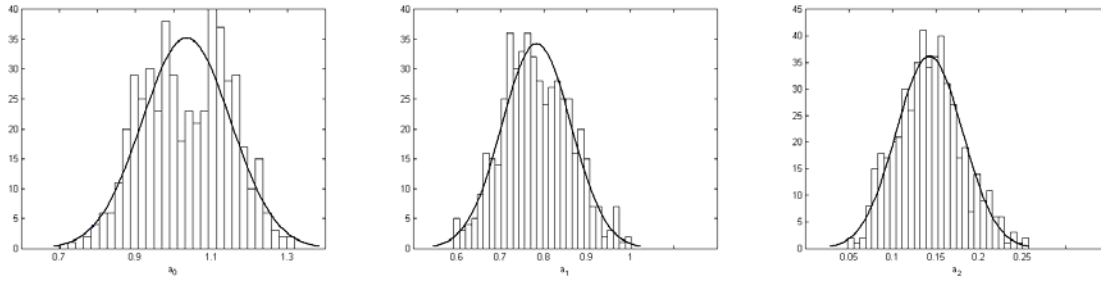


Figura 6.1: Histograma das Estimativas de MV - Modelo M1.

A Figura 6.2 apresenta os histogramas para as amostras geradas via MCMC da distribuição a posteriori dos parâmetros α_0, α_1 e α_2 .

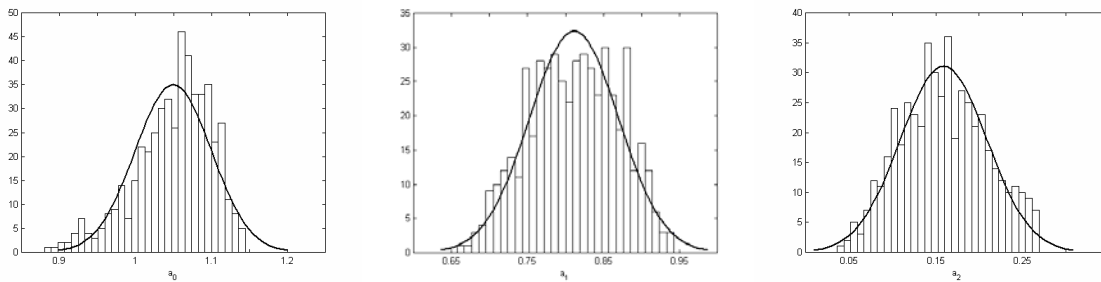


Figura 6.2: Histograma das Estimativas Bayesianas - Modelo M1.

Os erros de estimação para cada um dos modelos ARCH(2) listados na Tabela 6.1 são apresentados na Tabela 6.3, assim como os intervalos de credibilidade de 95% para as estimativas Bayesianas e os intervalos de confiança de 95% para as estimativas de Máxima Verossimilhança (ML).

Tabela 6.3: Estimativas de MV e Bayesianas - Modelos ARCH(2).

Modelo M1	$\hat{\alpha}$	RMSE	MAPE	IC(95%)	
α_0	Bayes	1.0497	0.0215	1.7340	[0.8694,1.1298]
	EMV	1.0342	0.1213	10.1843	[0.8156,1.2410]
α_1	Bayes	0.8114	0.0234	2.3890	[0.6682,0.9309]
	EMV	0.7834	0.0810	8.3347	[0.6376,0.9475]
α_2	Bayes	0.1591	0.0332	18.2335	[0.0065,0.2808]
	EMV	0.1427	0.0389	20.7759	[0.0745,0.2233]

Tabela 6.3 (cont.): Estimativas de MV e Bayesianas - Modelos ARCH(2).

	Modelo M2	$\hat{\alpha}$	RMSE	MAPE	IC(95%)
α_0	Bayes	1.0098	0.1028	1.5850	[0.8099,1.1982]
	EMV	1.0261	0.1665	6.1929	[0.7418,1.4299]
α_1	Bayes	0.5025	0.0920	4.6957	[0.3232,0.6683]
	EMV	0.4836	0.1180	9.2747	[0.2615,0.7058]
α_2	Bayes	0.2202	0.0554	7.6470	[0.1171,0.3279]
	EMV	0.1935	0.0827	16.7062	[0.1851,0.3726]
	Modelo M3	$\hat{\alpha}$	RMSE	MAPE	IC(95%)
α_0	Bayes	1.0041	0.1056	9.4422	[0.7971,1.1490]
	EMV	1.0315	0.2108	9.9490	[0.6622,1.5060]
α_1	Bayes	0.8981	0.0887	6.0951	[0.7309,1.1164]
	EMV	0.8750	0.1543	7.5496	[0.5599,1.1625]
α_2	Bayes	0.2017	0.0416	15.4412	[0.1208,0.3205]
	EMV	0.1949	0.0746	16.9673	[0.0429,0.3401]

Estes resultados mostram que a abordagem Bayesiana apresenta estimativas mais precisas do que as estimativas obtidas pela abordagem de Máxima Verossimilhança e, este comportamento não muda mesmo quando os parâmetros não satisfazem as condições de estacionaridade impostas. O método de Máxima Verossimilhança apresenta um desempenho inferior para o modelo não-estacionário.

Para verificarmos o efeito da ordem do modelo na precisão das estimativas, consideramos um modelo ARCH(5) cujos parâmetros são dados na Tabela 6.1. Os hiperparâmetros das distribuições a priori usados na abordagem Bayesiana são dados na Tabela 6.2. Os gráficos da Figura 6.3 apresentam a distribuição empírica (histograma) das estimativas de Máxima Verossimilhança dos parâmetros do modelo ARCH(5).

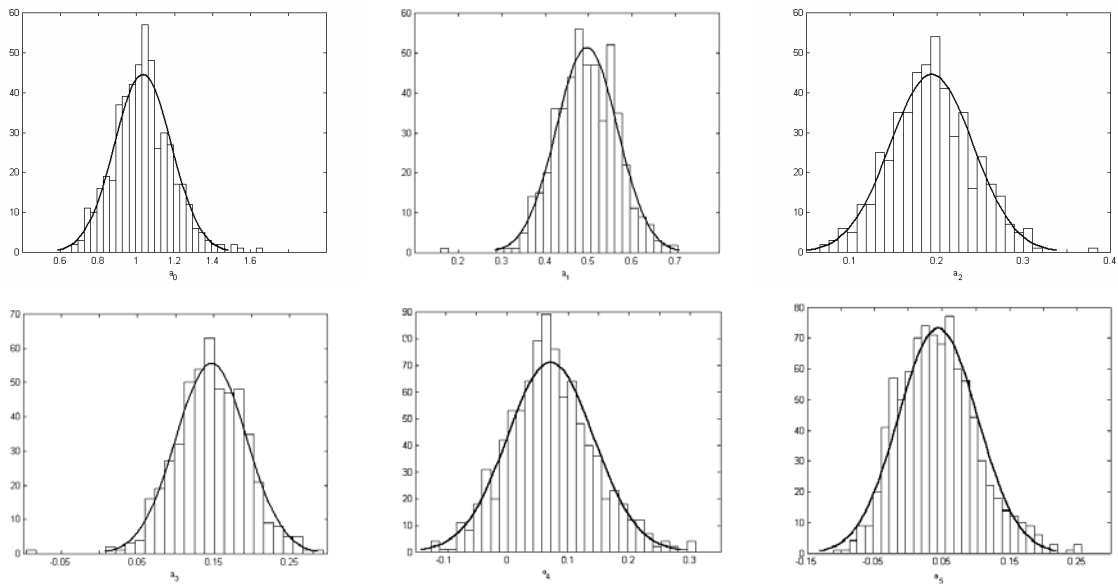


Figura 6.3: Histograma das Estimativas de MV - Modelo ARCH(5).

A Figura 6.4 apresenta os histogramas para as amostras geradas via MCMC da distribuição a posteriori dos parâmetros $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4$ e α_5 do modelo ARCH(5).

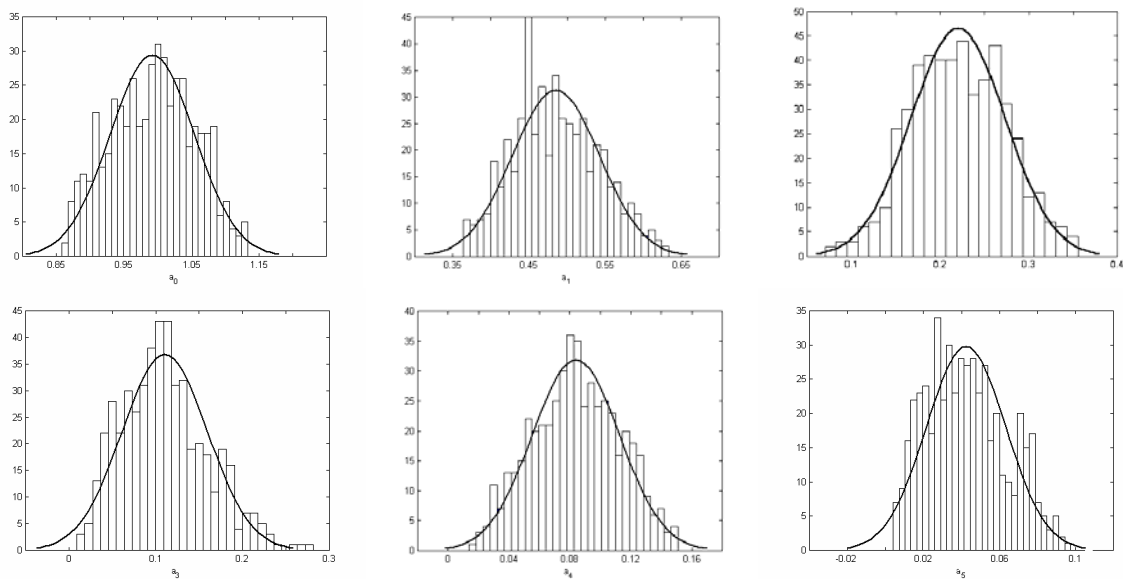


Figura 6.4: Histograma das Estimativas Bayesianas - Modelo ARCH(5).

A Tabela 6.4 mostra as estimativas Bayesianas e de Máxima Verossimilhança obtidas, além de seus respectivos intervalos e, valores do RMSE e do MAPE para os parâmetros do modelo considerado.

Tabela 6.4: Estimativas de MV e Bayesianas - Modelo ARCH(5).

ARCH(5)	$\hat{\alpha}$	RMSE	MAPE	IC(95%)
Bayes	0.9927	0.0139	1.0930	[0.8720,1.1300]
α_0				
EMV	1.0342	0.1537	11.8611	[0.7518,1.3525]
Bayes	0.4855	0.0281	4.6313	[0.3684,0.6304]
α_1				
EMV	0.4958	0.0705	11.3260	[0.3686,0.6356]
Bayes	0.2208	0.0571	4.5685	[0.1157,0.3273]
α_2				
EMV	0.1928	0.0985	19.0299	[0.0103,0.3934]
Bayes	0.1104	0.0319	17.5301	[0.0160, 0.2774]
α_3				
EMV	0.1474	0.0466	24.3752	[0.0630,0.2411]
Bayes	0.0839	0.0148	15.1766	[0.0090, 0.1507]
α_4				
EMV	0.0735	0.0462	38.3743	[0.0055,0.1431]
Bayes	0.0424	0.0068	11.1104	[0.0063, 0.0922]
α_5				
EMV	0.0527	0.1050	55.3151	[-0.0053,0.1103]

De modo similar ao do modelo ARCH(2), os resultados mostram que as estimativas Bayesianas apresentam menores RMSE e MAPE do que as estimativas de Máxima Verossimilhança. Além disso, os intervalos de credibilidade são mais estreitos do que os intervalos de confiança, indicando maior precisão da abordagem Bayesiana em relação à abordagem de Máxima Verossimilhança.

É importante enfatizar que a estimativa de Máxima Verossimilhança do parâmetro α_5 apresenta valor negativo (coluna 5 da Tabela 6.4), violando assim as condições de positividade impostas aos parâmetros de modelos ARCH(q).

Finalmente, comparando os valores do RMSE e do MAPE da Tabela 6.4 com os valores da Tabela 6.3 podemos concluir que as estimativas de Máxima Verossimilhança se tornam menos precisas à medida que a ordem do modelo aumenta.

CAPÍTULO 7

Conclusões e Perspectivas

7.1. Conclusões

Neste trabalho apresentamos um estudo de modelos auto-regressivos com heterocedasticidade (ARCH) e de modelos auto-regressivos com erros ARCH (AR-ARCH), e mostramos os procedimentos para a estimação e para a seleção da ordem dos mesmos. As estimativas foram obtidas utilizando a abordagem de Máxima Verossimilhança e a abordagem Bayesiana. Estes modelos têm por objetivo descrever a volatilidade existente em séries financeiras.

As principais contribuições deste trabalho foram as apresentações de estudos detalhados, tanto do ponto de vista Clássico quanto do Bayesiano, de dois modelos com heterocedasticidade ajustados a séries simuladas e a duas séries específicas do mercado financeiro brasileiro: o IBovespa e a Telebrás.

Do ponto de vista Clássico, mostramos que os intervalos de confiança padrão para parâmetros de modelos ARCH e de modelos AR-ARCH podem apresentar limites inferiores negativos, violando as restrições impostas aos parâmetros dos mesmos, as quais asseguram estacionariedade na covariância. Assim, propomos o uso do método de simulação Bootstrap paramétrico, para estimarmos de forma mais precisa os intervalos de confiança destes modelos.

Na análise Bayesiana consideramos uma reparametrização dos modelos estudados e adotamos distribuição a priori informativa normal para os parâmetros transformados. Além disso, propomos o uso da distribuição a priori não-informativa de Geweke (1989) similarmente reparametrizada. O resumo das distribuições a posteriori foi obtido por método de simulação de Monte Carlo em Cadeias de Markov (MCMC), através do algoritmo Metropolis-Hastings.

Os resultados deste trabalho mostram que, de um modo geral, a abordagem Bayesiana proposta fornece estimativas mais precisas do que a abordagem de Máxima Verossimilhança, sendo mais robusta para ordens elevadas dos modelos.

Além disso, os intervalos de confiança obtidos pela técnica convencional violam restrições impostas aos parâmetros dos modelos estudados, apesar de serem melhorados com o uso do método Bootstrap paramétrico. Este fato não é observado para a abordagem Bayesiana, uma vez que a distribuição a priori, e conseqüentemente, o núcleo do algoritmo Metropolis-Hastings forem bem escolhidos (levando em conta as restrições do modelo), os intervalos de credibilidade também fornecerão estimativas que refletem tais imposições.

Uma outra vantagem da abordagem Bayesiana está na reparametrização proposta, proporcionando a estabilidade da convergência, diferentemente da abordagem de Máxima Verossimilhança. Esta usa o método de Newton para a obtenção das estimativas, cuja solução numérica está bastante sujeita a não-convergência, principalmente para ordens elevadas de modelos ARCH(q) e de modelos AR(p)-ARCH(q).

Portanto, apesar do grande esforço computacional exigido pelos algoritmos de simulação MCMC, a precisão das estimativas obtidas neste trabalho justifica a abordagem Bayesiana como proposta para a estimação de parâmetros de modelos ARCH(q) e de modelos AR(p)-ARCH(q) ajustados às séries temporais do mercado financeiro.

7.2. Perspectivas

Como propostas para a continuidade deste trabalho, podemos considerar a previsão de valores das séries analisadas utilizando métodos de simulação MCMC, além de um estudo comparativo para modelos AR-ARCH visando ressaltar a potencialidade das abordagens de Máxima Verossimilhança e Bayesiana através da raiz quadrada do erro quadrático médio e,

do erro percentual absoluto médio entre os valores reais dos parâmetros e os valores estimados por cada uma das técnicas (como foi realizado para os modelos ARCH no capítulo 6).

Visto que a distribuição não-condicional das séries de retornos, em geral, tem caudas mais pesadas que as de uma distribuição normal e, apesar de simétrica, pode ser caracterizada como leptocúrtica, podemos propor também o uso de outras distribuições para o ruído ε_t , como t -Student e Potência Exponencial.

Além disso, podemos estender o estudo analisando modelos AR com erros GARCH, sob um enfoque Bayesiano. Os modelos AR-GARCH têm uma estrutura similar à estudada na seção 2.1 e pode ser resumida como:

$$y_t = \mathbf{x}_t \boldsymbol{\beta} + z_t \quad (7.1)$$

$$y_t | \Omega_{t-1} \sim P(\mathbf{x}_t, \boldsymbol{\beta}, h_t) \quad (7.2)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i z_{t-i}^2 + \sum_{j=1}^r \gamma_j h_{t-j}^2 \quad (7.3)$$

$$z_t = y_t - \mathbf{x}_t \boldsymbol{\beta} \quad (7.4)$$

em que y_t representa uma série de retornos, $P(\cdot)$ é uma distribuição paramétrica, usualmente a Normal ou t -Student, \mathbf{x}_t representa um vetor de variáveis exógenas que pode incluir valores de y_t defasados no tempo, $\boldsymbol{\beta}$ é um vetor de parâmetros desconhecidos, Ω_{t-1} representa o conjunto de informações disponíveis até o instante $t-1$ e z_t é um erro que satisfaz o modelo

$$z_t = h_t^{1/2} \varepsilon_t \quad (7.5)$$

em que $\{\varepsilon_t, t \geq 0\}$ é uma seqüência de ruído branco *i.i.d.* $N(0,1)$ independente de \mathbf{x}_t .

Como os modelos GARCH(q,r) podem ser escritos como ARCH(∞), o ajuste de modelos AR-GARCH à séries de retornos tende a ser mais vantajoso em relação aos processos AR-ARCH, pois ordens elevadas podem ser evitadas, melhorando ainda mais a qualidade do ajuste e, conseqüentemente, da inferência destes modelos.

APÊNDICE A.1

Modelos Auto-Regressivos

1. Introdução

Se no modelo de regressão proposto por Engle (1982), definido pelas expressões (2.1)-(2.4), z_t é um ruído branco, ou seja, uma seqüência de variáveis aleatórias não correlacionadas, com média zero e variância constante σ^2 , então, temos um modelo linear auto-regressivo de ordem p , $AR(p)$, cujo valor corrente do processo y_t é expresso como uma combinação linear de seus p valores passados $\mathbf{x}_t = (y_{t-1}, y_{t-2}, \dots, y_{t-p})'$ e do ruído branco z_t . Logo, o modelo pode ser escrito da seguinte forma:

$$y_t = \mathbf{x}_t \boldsymbol{\beta} + z_t \quad (1)$$

$$y_t | \Omega_{t-1} \sim P(x_t, \boldsymbol{\beta}, \sigma^2) \quad (2)$$

$$z_t = y_t - \mathbf{x}_t \boldsymbol{\beta} \quad (3)$$

em que $P(\cdot)$ é uma distribuição paramétrica, \mathbf{x}_t representa um vetor de variáveis exógenas, $\boldsymbol{\beta}$ é um vetor de parâmetros desconhecidos e Ω_{t-1} representa o conjunto de informações disponíveis até o instante $t-1$. Vamos considerar ainda que o ruído branco z_t é um processo *i.i.d.* $N(0, \tau^{-1})$, sendo $\tau = \frac{1}{\sigma^2} > 0$.

Para que o processo $AR(p)$ seja estacionário, todas as raízes do polinômio $1 - \sum_{i=1}^p \beta_i l^i$ (l : operador translação) devem estar fora do círculo de raio unitário.

2. Função de Verossimilhança

Consideremos uma trajetória observada $\mathbf{Y} = \{y_t, t = 1, 2, \dots, T\}$ do processo y_t . Então, a função de verossimilhança de y_t , $t = p+1, \dots, T$, condicionada às p primeiras observações (assumindo que elas são conhecidas) pode ser escrita como:

$$L(\boldsymbol{\beta}, \tau) \propto \prod_{t=p+1}^T \tau^{1/2} \exp\left\{-\frac{\tau}{2}(y_t - \mathbf{x}_t \boldsymbol{\beta})^2\right\} \quad (4)$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ e $\tau > 0$.

Podemos ainda escrever a expressão (4) na forma matricial (Oliveira, 1998):

$$L(\boldsymbol{\beta}, \tau) \propto \tau^{\frac{1}{2}(T-p)} \exp\left\{-\frac{\tau}{2}[(\mathbf{x} - \mathbf{H}\boldsymbol{\beta})'(\mathbf{x} - \mathbf{H}\boldsymbol{\beta})]\right\} \quad (5)$$

onde $\mathbf{x} = (y_{p+1}, \dots, y_T)'$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$, $\tau > 0$ e

$$\mathbf{H}' = \begin{pmatrix} y_p & y_{p+1} & y_{p+2} & \cdots & y_{T-1} \\ y_{p-1} & y_p & y_{p+1} & \cdots & y_{T-2} \\ y_{p-2} & y_{p-1} & y_p & \cdots & y_{T-3} \\ \vdots & \vdots & \vdots & & \vdots \\ y_1 & y_2 & y_3 & \cdots & y_{T-p} \end{pmatrix} \quad (6)$$

Mas,

$$(\mathbf{x} - \mathbf{H}\boldsymbol{\beta})'(\mathbf{x} - \mathbf{H}\boldsymbol{\beta}) = (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + (\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\beta}})'(\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\beta}}) \quad (7)$$

onde $\hat{\boldsymbol{\beta}} = (\mathbf{H}'\mathbf{H})^{-1}(\mathbf{H}'\mathbf{x})$ e $\boldsymbol{\Sigma}^{-1} = \mathbf{H}'\mathbf{H}$. Assim, a função de verossimilhança condicionada às p primeiras observações é dada por (Box & Jenkins, 1976; Andrade & Barreto, 2000):

$$L(\boldsymbol{\beta}, \tau) \propto \tau^{\frac{1}{2}(T-p)} \exp\left\{-\frac{\tau}{2}\left[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + (\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\beta}})'(\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\beta}})\right]\right\} \quad (8)$$

O logaritmo natural da função de verossimilhança (8) é dado por:

$$l(\boldsymbol{\beta}, \tau) \propto \frac{1}{2}(T-p)\ln(\tau) - \frac{\tau}{2}\left[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + (\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\beta}})'(\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\beta}})\right] \quad (9)$$

Diferenciando a função (9) em relação a τ e em relação a $\boldsymbol{\beta}$ obtemos:

$$\frac{\partial l(\boldsymbol{\beta}, \tau)}{\partial \tau} = \frac{(T-p)}{2\tau} - \frac{1}{2} [(\mathbf{x} - \mathbf{H}\boldsymbol{\beta})'(\mathbf{x} - \mathbf{H}\boldsymbol{\beta})] \quad (10)$$

$$\frac{\partial l(\boldsymbol{\beta}, \tau)}{\partial \boldsymbol{\beta}} = \tau \mathbf{H}'(\mathbf{x} - \mathbf{H}\boldsymbol{\beta}) \quad (11)$$

Igualando estas expressões a zero e resolvendo as equações resultantes, obtemos os estimadores de máxima verossimilhança para os parâmetros τ e $\boldsymbol{\beta}$ de um modelo AR(p).

Assim, os estimadores $\hat{\tau}$ e $\hat{\boldsymbol{\beta}}$ são obtidos diretamente das expressões (10) e (11):

$$\hat{\tau}^{-1} = \frac{1}{(T-p)} (\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\beta}})'(\mathbf{x} - \mathbf{H}\hat{\boldsymbol{\beta}}) \quad (12)$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{H}'\mathbf{H})^{-1}(\mathbf{H}'\mathbf{x}) \quad (13)$$

onde $\mathbf{x} = (y_{p+1}, \dots, y_T)'$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$, $\tau > 0$ e \mathbf{H} é a matriz definida em (6).

APÊNDICE A.2

Método de Simulação Bootstrap

1. O Método Bootstrap

O Bootstrap é um método computacional para avaliar precisão de estimativas e testes, usando o esforço computacional, sem necessidade de muitas suposições ou desenvolvimentos analíticos complicados. O método foi proposto por Efron (1979) e tem ajudado na solução dos mais diversos problemas estatísticos.

A idéia principal do método Bootstrap é tratar os dados como se fosse a população e retirar amostras (com reposição) dos dados como se fosse uma amostragem da população. O método consiste em repetir este procedimento um grande número de vezes (digamos, R), calculando para cada amostra obtida a quantidade de interesse. Então, com os R valores das quantidades de interesse, estima-se suas distribuições desconhecidas.

Se comparado a outras técnicas, o método Bootstrap teve sua difusão um tanto tardia, devido a sua dependência do uso intensivo de cálculos computacionais. Os progressos da informática experimentados nas últimas décadas do século XX possibilitaram a popularização do uso do computador e incrementaram o surgimento e acesso a softwares matemáticos e estatísticos. Conseqüentemente, as aplicações de métodos Bootstrap às mais diferentes áreas da estatística se intensificaram.

Ao usuário de inferência estatística, compete o entendimento de que a amostra obtida em seu experimento é apenas uma dentre várias ou infinitas possibilidades. A estimativa $\hat{\theta}$ de θ obtida em sua amostra, poderia ser diferente em alguma outra amostra obtida da mesma população, através do mesmo processo de amostragem. Realizando o experimento uma única vez, o pesquisador tem em mãos apenas uma destas possíveis estimativas $\hat{\theta}$. Para serem obtidas inferências para θ pode ser utilizada a normalidade assintótica dos estimadores de

Máxima Verossimilhança. A utilização de propriedades assintóticas é direcionada pelo tamanho da amostra, que deve ser suficientemente grande.

Pode ser usado o método Bootstrap paramétrico e/ou o método Bootstrap não-paramétrico na construção de intervalos de confiança e testes de hipóteses. No Bootstrap paramétrico, as estimativas de Máxima Verossimilhança são obtidas por meio do modelo ajustado, isto é, os dados são gerados do modelo ajustado com os valores dos parâmetros fixados nas estimativas de Máxima Verossimilhança obtidas da amostra original; e no Bootstrap não-paramétrico, as estimativas de Máxima Verossimilhança são baseadas em R reamostras com reposição obtidas da amostra original. O Bootstrap não-paramétrico é mais robusto contra suposições distribucionais, ao passo que o Bootstrap paramétrico é mais eficiente quando as suposições paramétricas são verdadeiras. Maiores detalhes sobre a técnica Bootstrap podem ser encontrados em Wehrens et al (2000).

2. Intervalo de Confiança via Método Bootstrap

Em inferência estatística temos interesse na quantificação do erro cometido ao se estimar um parâmetro de interesse $\hat{\theta}$ através de θ . Uma estratégia usual para a busca de medidas de incerteza, que expressem este erro, é a estimação do erro padrão de θ . Entretanto, métodos analíticos para a obtenção destas medidas nem sempre são disponíveis, ou constituem processos altamente complexos, enquanto métodos assintóticos, nos quais a construção de intervalos de confiança é baseada, dependem de aproximações nem sempre alcançadas. Neste contexto, o método Bootstrap constitui uma eficiente alternativa, fornecendo estimativas do erro padrão de $\hat{\theta}$ livres de complexidades algébricas e possibilitando a obtenção de intervalos de confiança sem necessidade de pressupostos sobre a distribuição do estimador.

Desta forma, o método Bootstrap é utilizado para a obtenção de estimativas intervalares empíricas para os estimadores dos parâmetros de interesse, através da reamostragem do conjunto de dados originais.

Seja μ o parâmetro de interesse. Para cada amostra, devemos calcular a estimativa de Máxima Verossimilhança para μ , e temos no final de R reamostragens, $\hat{\mu}_1^* < \dots < \hat{\mu}_R^*$ valores das estimativas de Máxima Verossimilhança ordenadas. Utilizamos, então,

$\hat{\mu}_{1(R+1)\left(\frac{\alpha}{2}\right)}^*$ e $\hat{\mu}_{2(R+1)\left(1-\frac{\alpha}{2}\right)}^*$ como sendo os limites inferiores e superiores do intervalo $100(1-\alpha)\%$ de confiança para μ . Em geral, o número de reamostragens R é fixado em 999.

Desta forma, através de $\hat{\mu}_{1(R+1)\left(\frac{\alpha}{2}\right)}^*$ e $\hat{\mu}_{2(R+1)\left(1-\frac{\alpha}{2}\right)}^*$ podemos obter intervalos de confiança percentil Bootstrap $100(1-\alpha)\%$ para o parâmetro de interesse.

Segundo Efron (1979), intervalos Bootstrap também são aproximados, entretanto, oferecem melhores aproximações do que os intervalos de confiança padrão.

APÊNDICE A.3

Inferência Bayesiana

1. Introdução

A abordagem Bayesiana é adequada quando ocorrem situações nas quais os dados avaliados são incompatíveis com modelos simples e, outros modelos mais realistas (conseqüentemente mais complexos) são necessários.

Como na inferência clássica, a inferência Bayesiana utiliza observações y , cujo valor é inicialmente incerto e descrito através de uma distribuição de probabilidades $f(y|\theta)$. A quantidade θ serve como indexador da família de distribuições das observações, representando características de interesse que se deseja conhecer para ter uma descrição completa do processo. Se o pesquisador tem informações sobre esta quantidade, então, é possível que esse conhecimento prévio seja incorporado à análise. O método clássico não admite estas informações por não serem observáveis, mas o Bayesiano as incorpora à análise através de uma densidade $\pi(\theta)$, mesmo que estas informações não sejam muito precisas.

Dadas as informações y obtidas sobre um modelo paramétrico $f(y|\theta)$, o procedimento de inferência Bayesiana é baseado na forma familiar do Teorema de Bayes, dada por (Box & Tiao, 1973):

$$p(\theta|y) = \frac{l(\theta|x)\pi(\theta)}{\int l(\theta|x)\pi(\theta)d\theta} \quad (1)$$

em que $p(\theta|y)$ é chamada de distribuição a posteriori, que resulta da combinação da função de verossimilhança $l(\theta|x)$ com a distribuição a priori $\pi(\theta)$. Portanto, a distribuição a posteriori representa a distribuição de probabilidade de θ após observar o valor y , que passa a fazer parte do conjunto de informação disponível. Se o interesse é específico em algum

componente de θ , digamos $\theta_i, (i = 1, \dots, k)$, a distribuição marginal correspondente pode ser facilmente obtida integrando a distribuição conjunta $p(\theta | y)$, isto é,

$$p(\theta_i | y) = \int l(\theta | x) \pi(\theta) d\theta_{-i} \quad (2)$$

na qual o termo $-i$ implica a integração de todos os componentes exceto θ_{-i} . Se inferências sumárias na forma de esperanças a posteriori são requeridas, elas são obtidas por:

$$E(\psi(\theta) | y) = \int \psi(\theta) \pi(\theta | y) d\theta$$

para uma escolha $\psi(\theta)$ de interesse.

Nas últimas décadas, alguns métodos numéricos têm sido desenvolvidos para resolver as questões mencionadas acima. Dentre os métodos numéricos baseados em amostragem destacam-se os métodos de Monte Carlo em Cadeias de Markov (MCMC). Estes são mais simples para implementação e não apresentam restrições quanto ao número de parâmetros a serem estimados.

2. Métodos de Simulação de Monte Carlo em Cadeias de Markov (MCMC)

O método de simulação de Monte Carlo em Cadeias de Markov (MCMC) é uma forma de integração de Monte Carlo. A idéia é simular uma cadeia de Markov irreduzível aperiódica cuja distribuição estacionária é a distribuição de interesse $p(\theta | y)$, ou seja, a distribuição a posteriori. Existem dois métodos para gerar cadeia de Markov com distribuição estacionária especificada. O Metropolis-Hastings (Chib & Greenberg, 1995), que tem sido usado por muitos anos em física estatística; e o Gibbs Sampler, que foi trazido para a literatura estatística por Gelfand & Smith (1990).

É importante observar que o uso destes algoritmos, em geral, é necessário se a geração não iterativa da distribuição da qual se deseja obter uma amostra for muito complicada ou custosa (Gamerman, 1996).

O Algoritmo Metropolis-Hastings

Quando as distribuições condicionais a posteriori não são facilmente identificadas como possuidoras de uma forma padrão (normal, gama, etc.), que impossibilita a geração direta a partir destas distribuições, usa-se o algoritmo Metropolis-Hastings. Esta técnica

requer um núcleo, isto é, uma densidade de transição $q(\theta, \theta^*)$, que não necessariamente tem probabilidade de equilíbrio π , mas que represente uma regra de passagem que defina uma cadeia. Considere também a probabilidade de aceitação $p(\theta^{(r-1)}, \theta^*)$ definida abaixo. O algoritmo segue os seguintes passos:

Passo 1: Atribua um valor inicial $\theta = \theta^{(0)}$ e inicie o contador de iterações em $r = 1$.

Passo 2: Mova a cadeia para um novo valor de θ^* gerado da densidade $q(\theta^{(r-1)}, \theta^*)$.

Passo 3: Gere u da distribuição uniforme $(0,1)$.

Passo 4: Aceite o valor gerado θ^* se

$$u \leq p(\theta^{(r-1)}, \theta^*) = \min\left(1, \frac{\pi(\theta^*)q(\theta^*, \theta^{(r-1)})}{\pi(\theta^{(r-1)})q(\theta^{(r-1)}, \theta^*)}\right)$$

Caso contrário, fique com $\theta^{(r)} = \theta^{(r-1)}$.

Passo 5: Incremente r e volte ao passo 2.

Assim, para r suficientemente grande, $\theta^{(r)}$ é uma amostra da distribuição a posteriori $\pi(\theta)$. Para o caso vetorial $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$, teremos uma densidade de transição dada por $q(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ e uma probabilidade de aceitação dada por $p(\boldsymbol{\theta}^{(r-1)}, \boldsymbol{\theta}^*)$, e devemos proceder da mesma forma.

3. Diagnósticos de Convergência

Os métodos MCMC são uma ótima ferramenta para a resolução de muitos problemas práticos, porém, alguns problemas surgem com o uso dos mesmos. Dentre eles podemos destacar: o número necessário de iterações para se obter convergência, a possibilidade das iterações iniciais da amostra serem influenciadas pelos valores iniciais dos parâmetros e, ainda, de as seqüências de valores apresentarem correlações entre os parâmetros. Não existe uma técnica geral para resolver estas questões, mas, existem métodos de verificação de convergência baseados nas propriedades da cadeia de Markov, que indicam a convergência ou não, da amostra selecionada para a distribuição marginal.

Uma avaliação da convergência pode ser feita preliminarmente analisando os gráficos ou as medidas descritivas obtidas a partir dos valores simulados para os parâmetros de interesse. Os gráficos mais freqüentes são os gráficos da quantidade de interesse estimada ao longo das iterações e o gráfico da estimativa da distribuição marginal a posteriori deste

parâmetro. As medidas descritivas são a média, o desvio-padrão e os quantis. Uma outra avaliação da convergência pode ser feita usando algumas técnicas de diagnóstico. As técnicas mais populares são descritas por Geweke (1992), Gelman & Rubin (1992), Raftery & Lewis (1992) e Heidelberger & Welch (1983). Cowles & Carlin (1996) comparam alguns métodos concluindo que, apesar deles detectarem comumente os problemas na convergência, estas técnicas também podem falhar no seu propósito, não sendo possível afirmar qual delas é mais eficiente. Por isso, é recomendável que o uso destes diagnósticos seja combinado com a análise gráfica e com as medidas descritivas.

Neste trabalho, além da análise gráfica e das medidas descritivas, usamos o diagnóstico de Geweke para avaliar a convergência.

Critério de Diagnóstico de Geweke

O diagnóstico de Geweke (1992) foi desenvolvido para indicar a convergência da média a posteriori de uma função da quantidade amostrada, $g(\theta)$, em uma única cadeia, isto é, os valores $g(\theta^{(i)})$ são calculados a cada iteração da simulação, formando uma série temporal. A partir da cadeia, a variância assintótica $S_g(0)$ da medida de $g(\theta)$ pode ser estimada. De uma forma resumida, o diagnóstico consiste nos seguintes passos:

Passo 1: Divida a cadeia gerada de tamanho N em duas seqüências, em que a primeira possui as N_a primeiras iterações e a segunda, as N_b últimas iterações.

Passo 2: Calcule as médias m_a e m_b e as variâncias assintóticas $\hat{S}_a(0)$ e $\hat{S}_b(0)$, respectivamente. As variâncias são determinadas pela estimação da densidade espectral.

Passo 3: Se a cadeia completa for estacionária, as médias m_a e m_b destas seqüências estarão próximas.

Pode-se mostrar que, se as razões $\frac{N_a}{N}$ e $\frac{N_b}{N}$ são fixas e $N \rightarrow \infty$,

$$\frac{m_a - m_b}{\sqrt{\frac{\hat{S}_a(0)}{N_a} + \frac{\hat{S}_b(0)}{N_b}}} \sim N(0,1).$$

Assim, um teste pode ser construído e, se a diferença padronizada entre as médias for muito grande, existe indicação de não convergência. Geweke (1992) sugere que as médias sejam construídas após algumas iterações iniciais terem sido descartadas e que $N_a = 0.1N$ e $N_b = 0.5N$.

APÊNDICE A.4

Gráficos da Seção 5.2

Apresentamos neste apêndice os gráficos obtidos através da abordagem Bayesiana para modelos ARCH(q), referentes ao capítulo 5 (seção 5.2). As figuras representam os trezentos últimos valores gerados para os parâmetros dos modelos ARCH(q) ajustados às séries IBovespa ($q = 3$) e Telebrás ($q = 7$) considerando distribuição a priori informativa e distribuição a priori não-informativa. Estes gráficos mostram a convergência gráfica dos parâmetros obtida através do algoritmo de simulação Metropolis-Hastings.

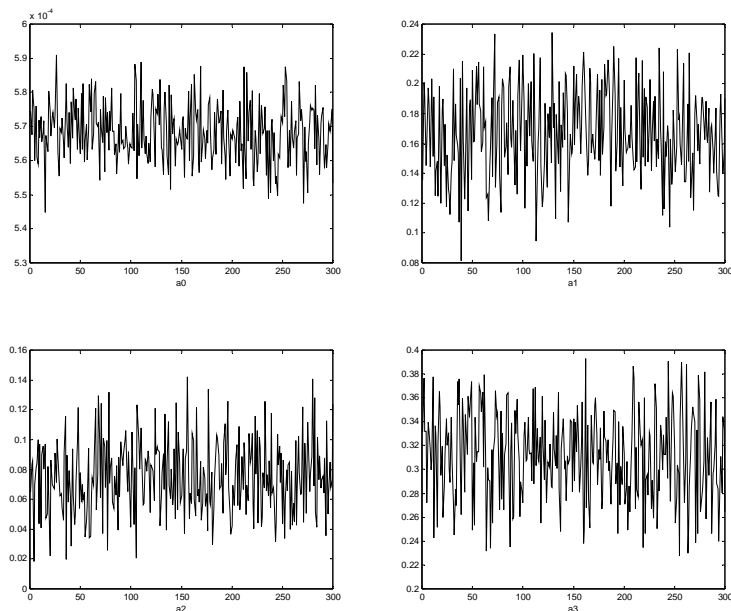


Figura 1: Valores gerados para os parâmetros do modelo ARCH(3) ajustado à série IBovespa considerando distribuição a priori informativa.

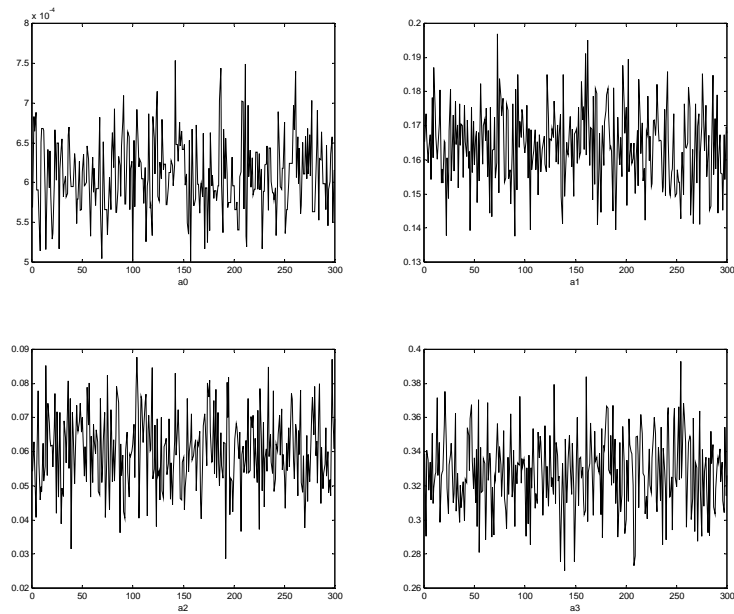


Figura 2: Valores gerados para os parâmetros do modelo ARCH(3) ajustado à série IBovespa considerando distribuição a priori não-informativa.

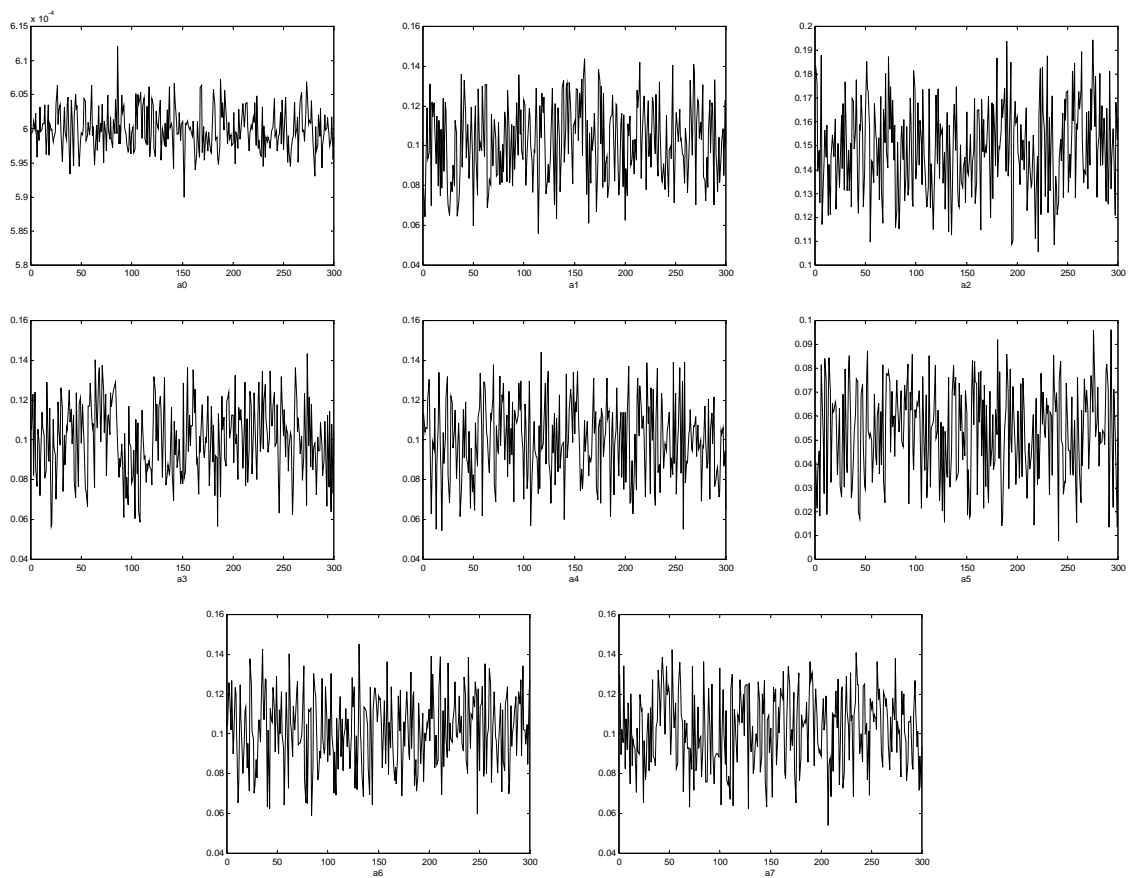


Figura 3: Valores gerados para os parâmetros do modelo ARCH(7) ajustado à série Telebrás considerando distribuição a priori informativa.

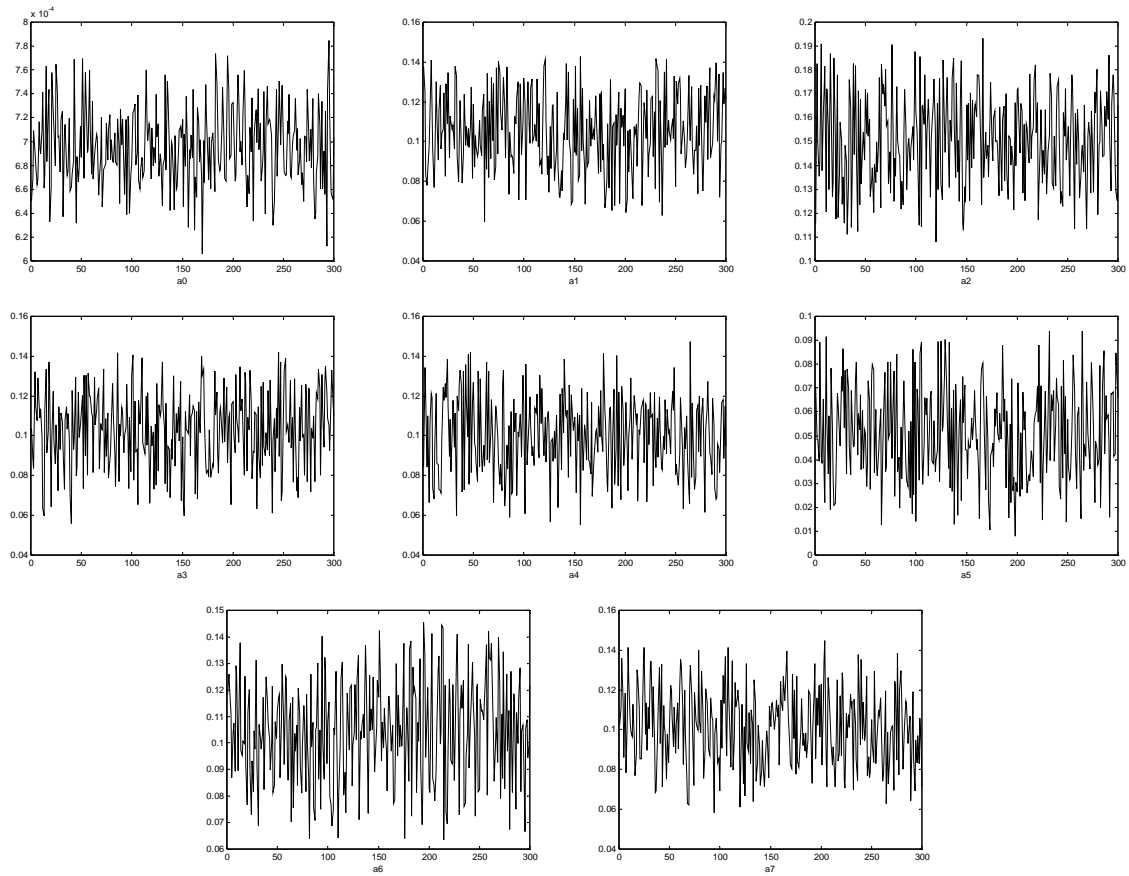


Figura 4: Valores gerados para os parâmetros do modelo ARCH(7) ajustado à série Telebrás considerando distribuição a priori não-informativa.

APÊNDICE A.5

Gráficos da Seção 5.3

Nesta seção apresentamos os gráficos obtidos através da abordagem Bayesiana para modelos $AR(p)$ - $ARCH(q)$, referentes ao capítulo 5 (seção 5.3). As figuras representam os trezentos últimos valores gerados para os parâmetros dos modelos $AR(p)$ - $ARCH(q)$ ajustados às séries IBovespa ($p = 6$ e $q = 3$) e Telebrás ($p = 2$ e $q = 6$) considerando distribuição a priori informativa e distribuição a priori não-informativa. Estes gráficos mostram a convergência gráfica dos parâmetros obtida através do algoritmo de simulação Metropolis-Hastings.

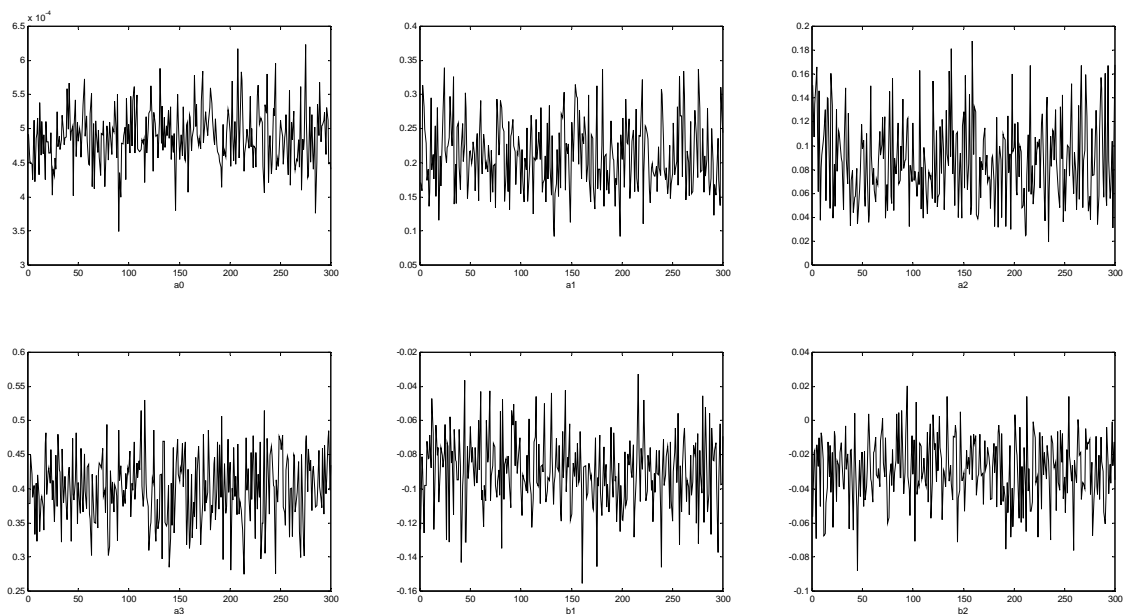


Figura 1: Valores gerados para os parâmetros do modelo $AR(6)$ - $ARCH(3)$ ajustado à série IBovespa considerando distribuição a priori informativa.

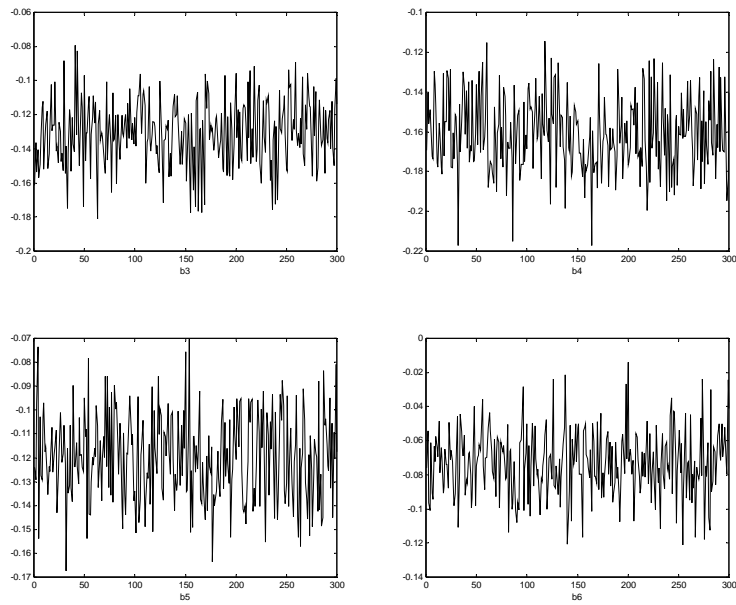


Figura 1 (cont.): Valores gerados para os parâmetros do modelo AR(6)-ARCH(3) ajustado à série IBovespa considerando distribuição a priori informativa.

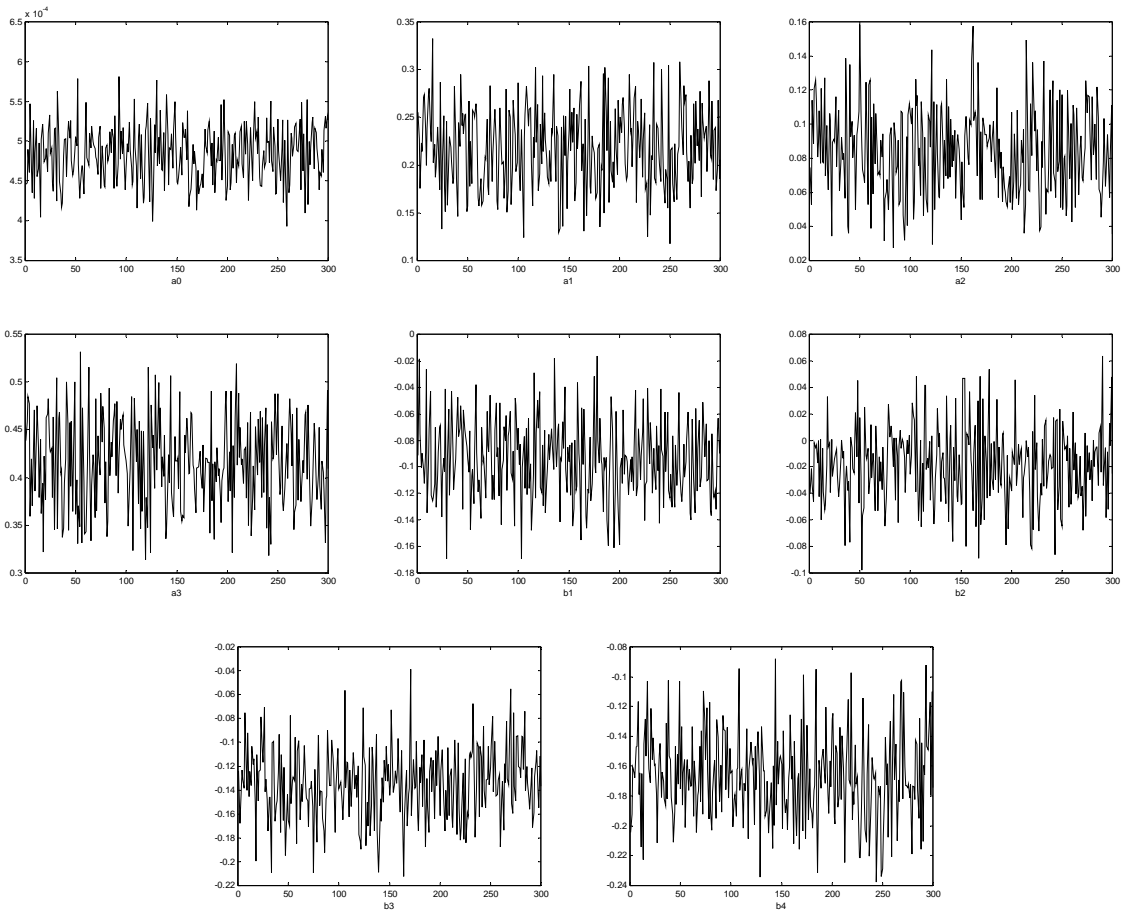


Figura 2: Valores gerados para os parâmetros do modelo AR(6)-ARCH(3) ajustado à série IBovespa considerando distribuição a priori não-informativa.

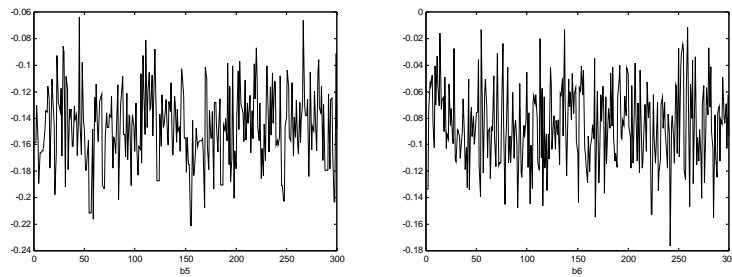


Figura 2 (cont.): Valores gerados para os parâmetros do modelo AR(6)-ARCH(3) ajustado à série IBovespa considerando distribuição a priori não-informativa.

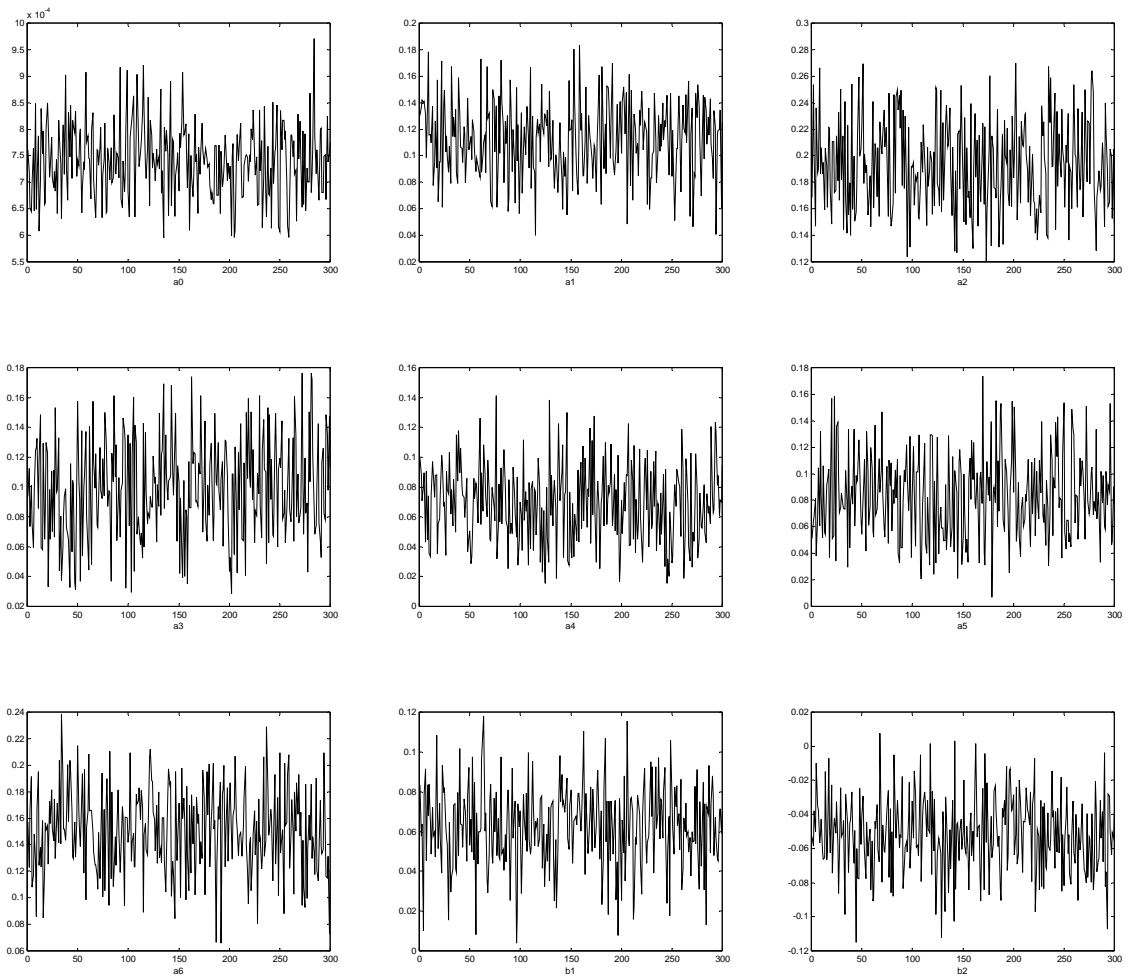


Figura 3: Valores gerados para os parâmetros do modelo AR(2)-ARCH(6) ajustado à série Telebrás considerando distribuição a priori informativa.

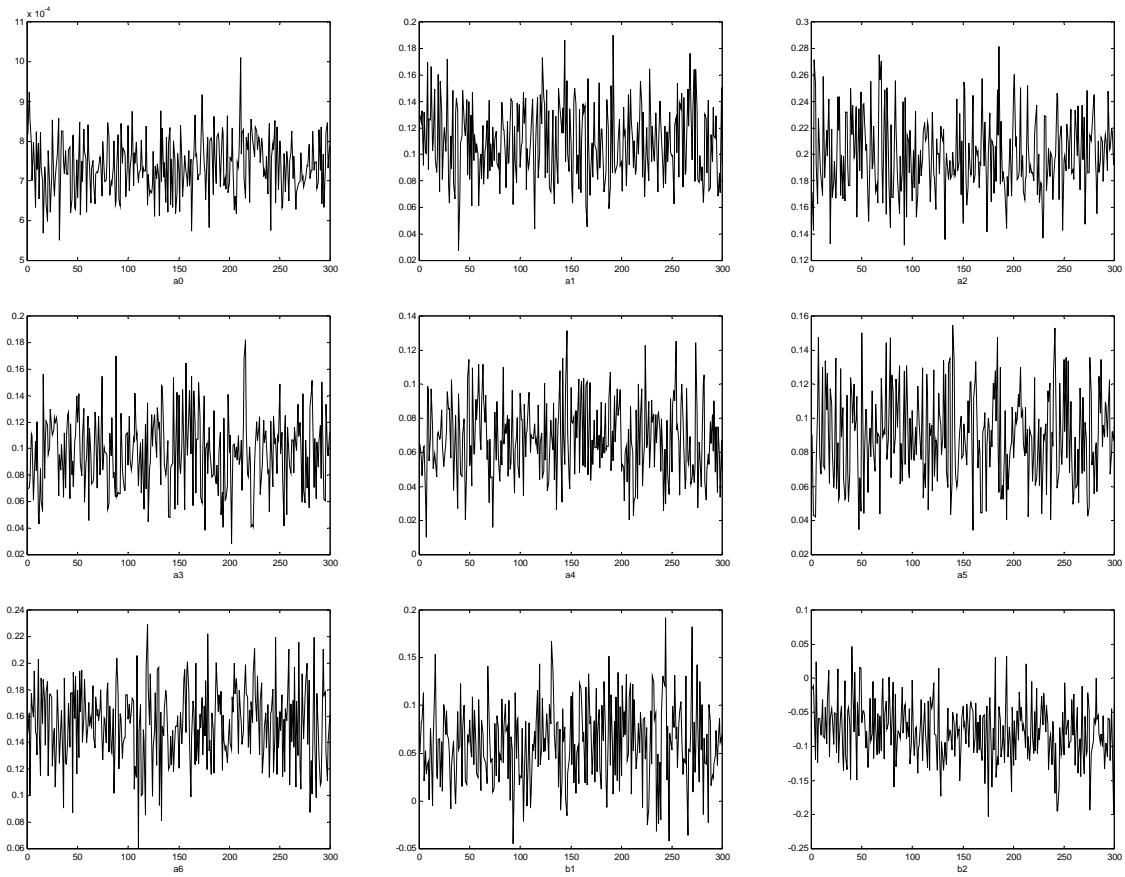


Figura 4: Valores gerados para os parâmetros do modelo AR(2)-ARCH(6) ajustado à série Telebrás considerando distribuição a priori não-informativa.

APÊNDICE A.6

Gráficos do Capítulo 6

Apresentamos neste apêndice os gráficos obtidos no estudo comparativo entre as abordagens de Máxima Verossimilhança e Bayesiana, considerando distribuição a priori informativa, para modelos ARCH(q), referentes ao capítulo 6. As figuras mostram a distribuição empírica (histograma) das estimativas de Máxima Verossimilhança e das estimativas Bayesianas dos parâmetros do processo ARCH(2) – modelo M2 e do processo ARCH(2) – modelo M3.

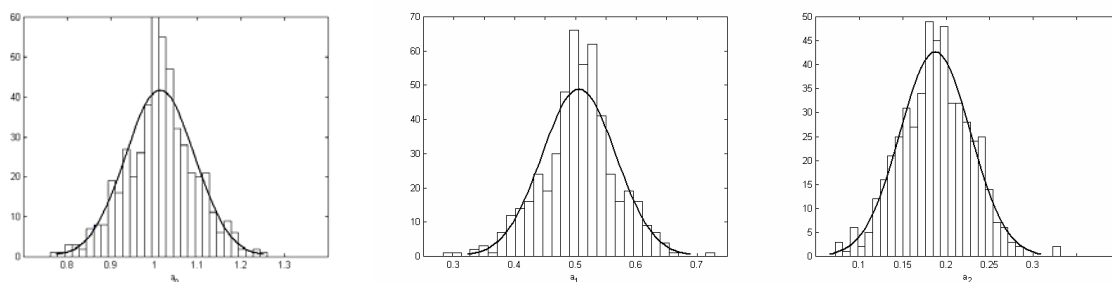


Figura 1: Histograma das Estimativas de Máxima Verossimilhança - Modelo M2.

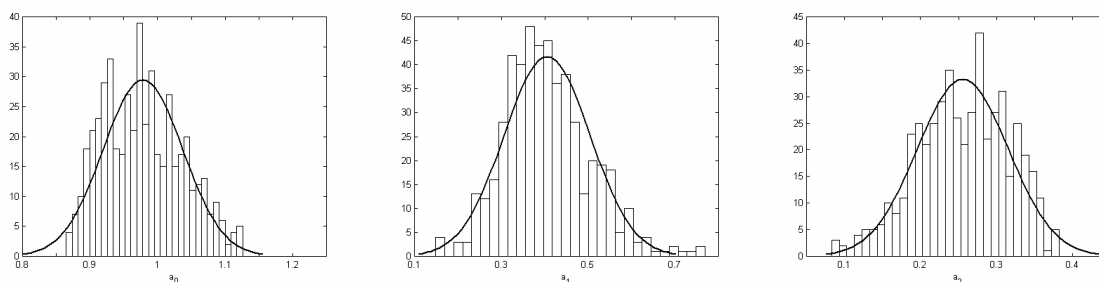


Figura 2: Histograma das Estimativas Bayesianas - Modelo M2.

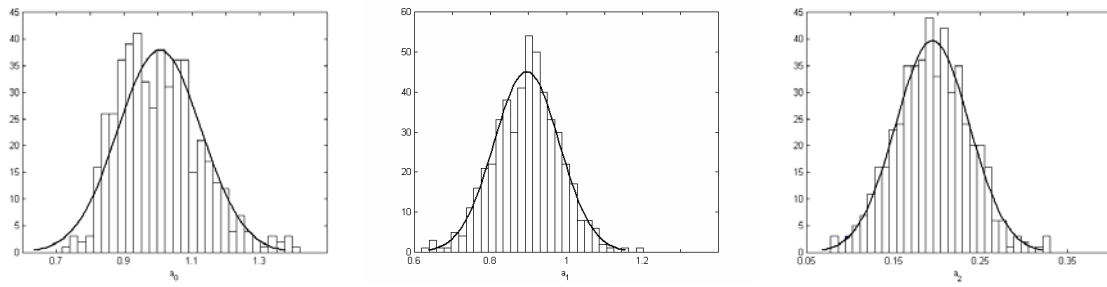


Figura 3: Histograma das Estimativas de Máxima Verossimilhança - Modelo M3.

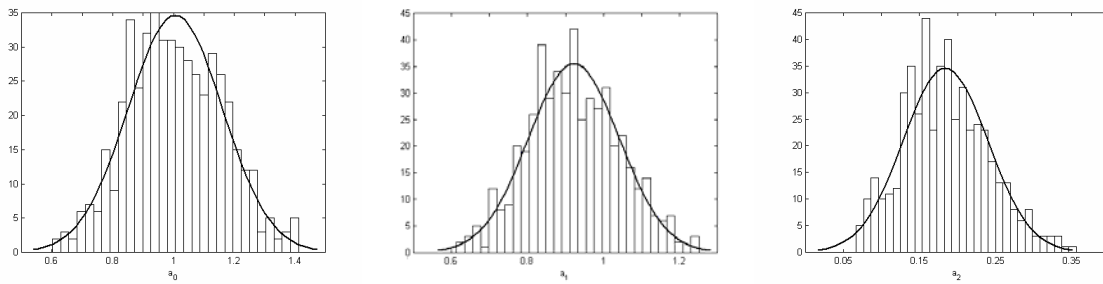
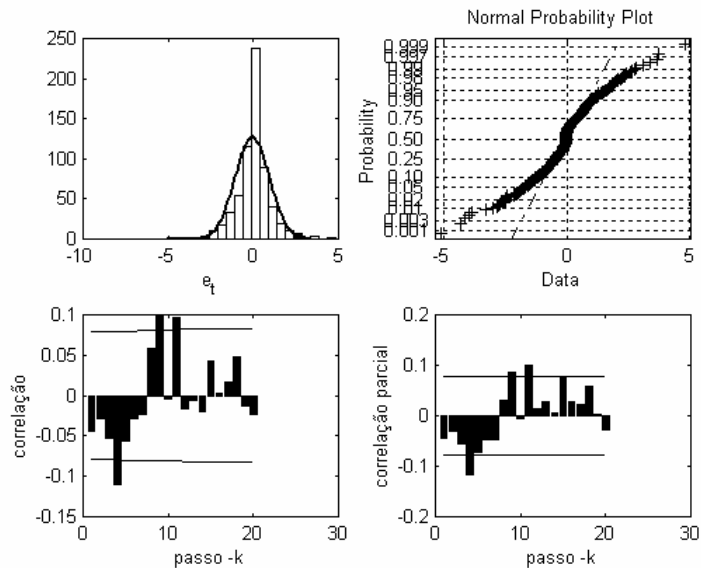


Figura 4: Histograma das Estimativas Bayesianas - Modelo M3.

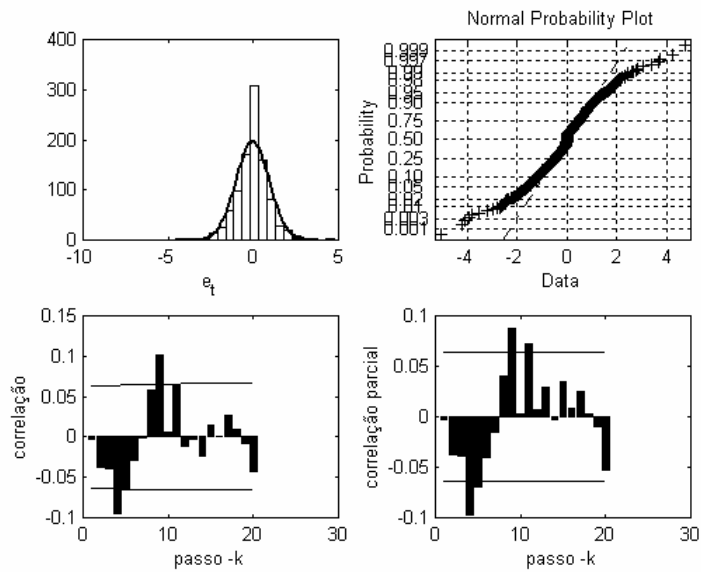
APÊNDICE A.7

Análise de Resíduos e Volatilidade Estimada

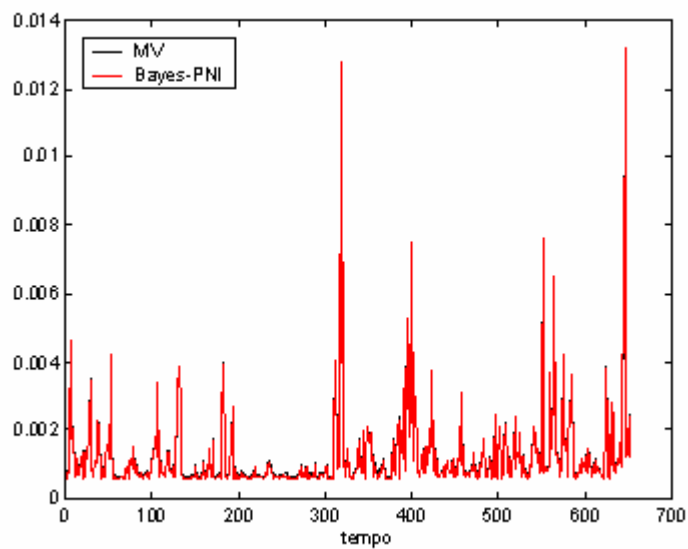
Nesta seção apresentamos os gráficos obtidos através das abordagens de Máxima Verossimilhança e Bayesiana para modelos $AR(p)$ - $ARCH(q)$ e $ARCH(q)$, referentes ao capítulo 5 (seção 5.2 e seção 5.3). As figuras a seguir representam a análise de resíduos e as volatilidades estimadas obtidas pelas estimativas Bayesianas, com distribuição a priori não-informativa, e pelas estimativas de Máxima Verossimilhança dos parâmetros dos modelos $ARCH(q)$ ajustados às séries IBovespa ($q = 3$) e Telebrás ($q = 7$), respectivamente.



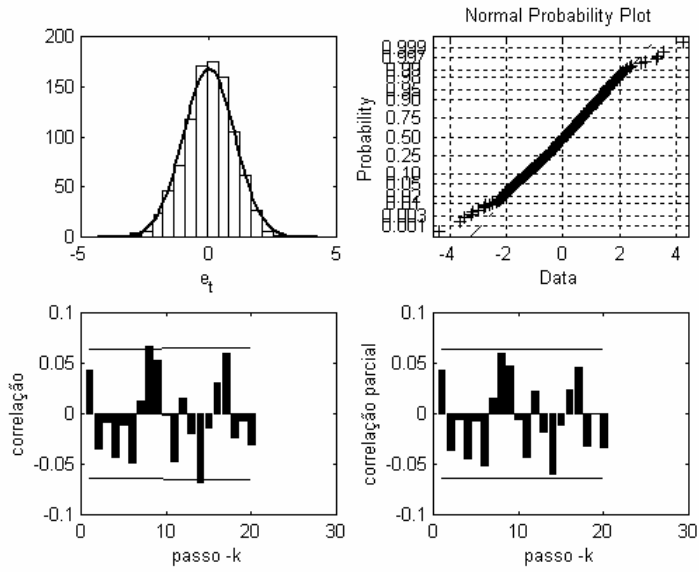
**Figura 1 – Análise de resíduos – Série IBovespa
Abordagem de Máxima Verossimilhança.**



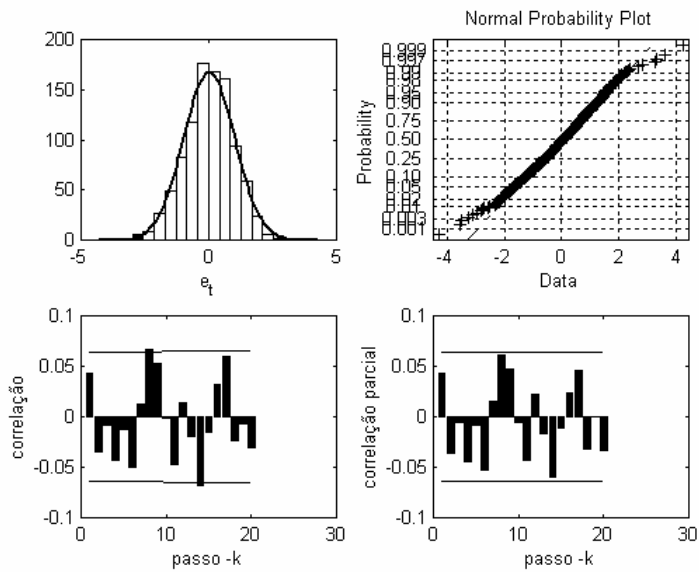
**Figura 2 – Análise de resíduos – Série IBovespa
Abordagem Bayesiana com distribuição a priori não-informativa.**



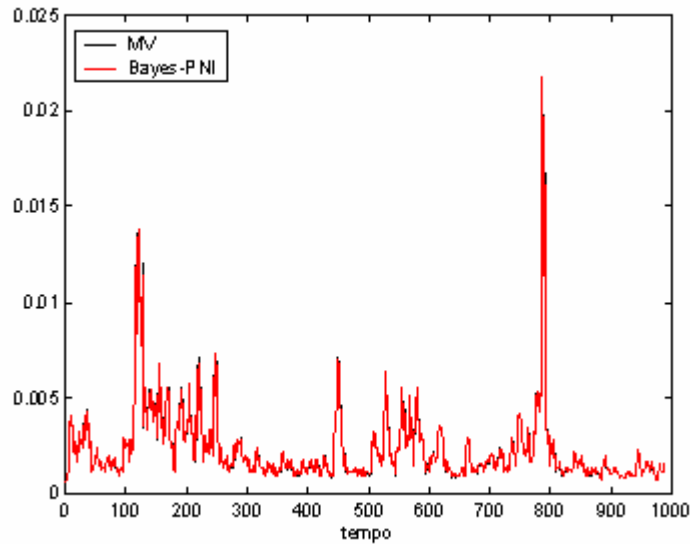
**Figura 3 – Volatilidade Estimada
Modelo ARCH(3) ajustado à série IBovespa.**



**Figura 4 – Análise de resíduos – Série Telebrás
Abordagem de Máxima Verossimilhança.**

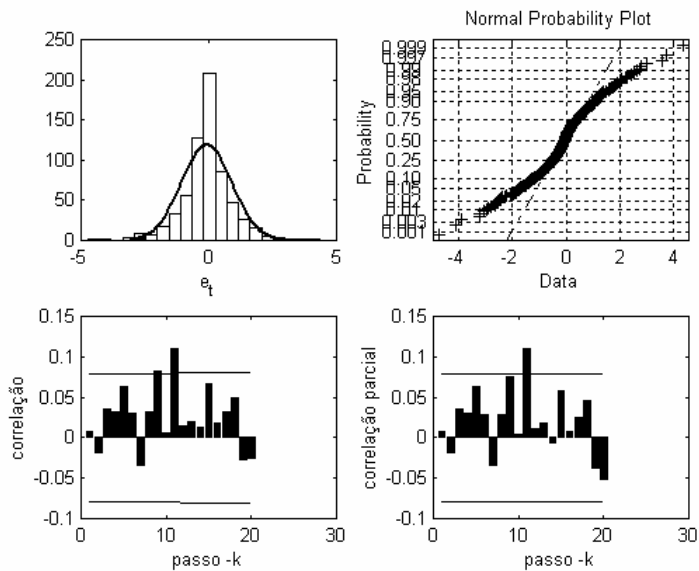


**Figura 5 – Análise de resíduos – Série Telebrás
Abordagem Bayesiana com distribuição a priori não-informativa.**

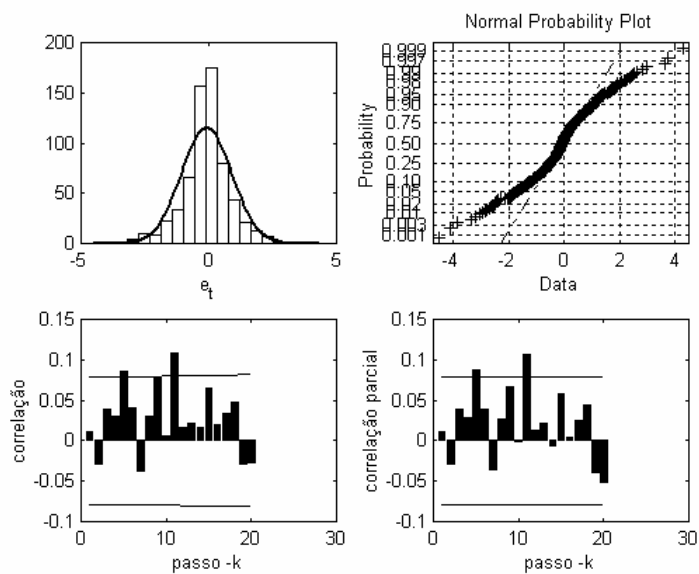


**Figura 6 – Volatilidade Estimada
Modelo ARCH(7) ajustado à série Telebrás.**

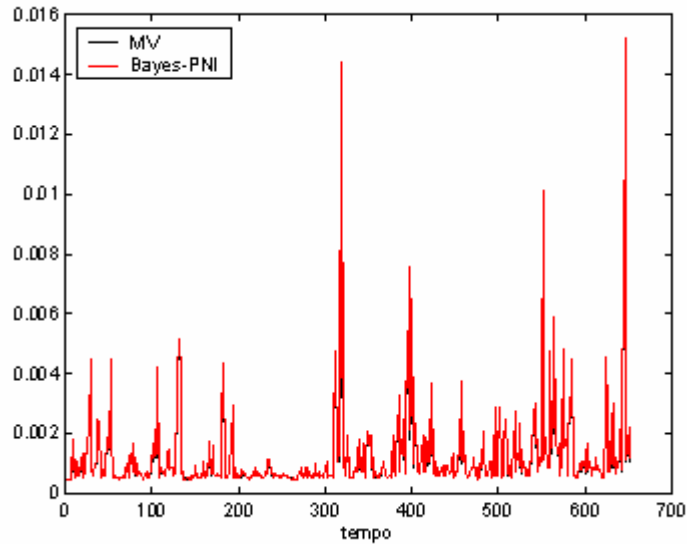
As figuras a seguir mostram a análise de resíduos e as volatilidades estimadas obtidas pelas estimativas Bayesianas, com distribuição a priori não-informativa, e pelas estimativas de Máxima Verossimilhança dos parâmetros dos modelos $AR(p)$ - $ARCH(q)$ ajustados às séries IBovespa ($p = 6$ e $q = 3$) e Telebrás ($p = 2$ e $q = 6$), respectivamente.



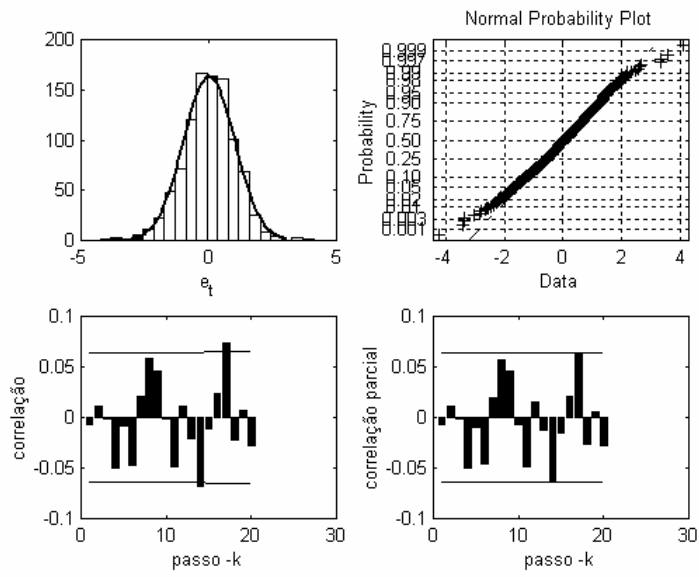
**Figura 7 – Análise de resíduos – Série IBovespa
Abordagem de Máxima Verossimilhança.**



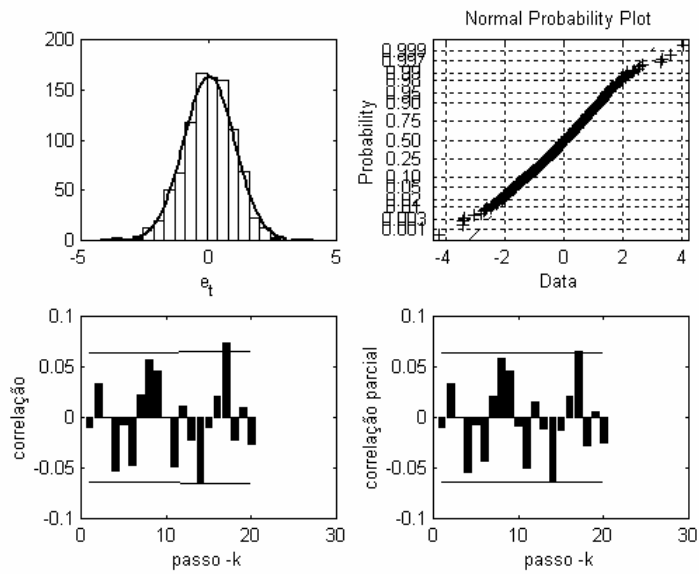
**Figura 8 – Análise de resíduos - Série IBOvespa
Abordagem Bayesiana com distribuição a priori não-informativa.**



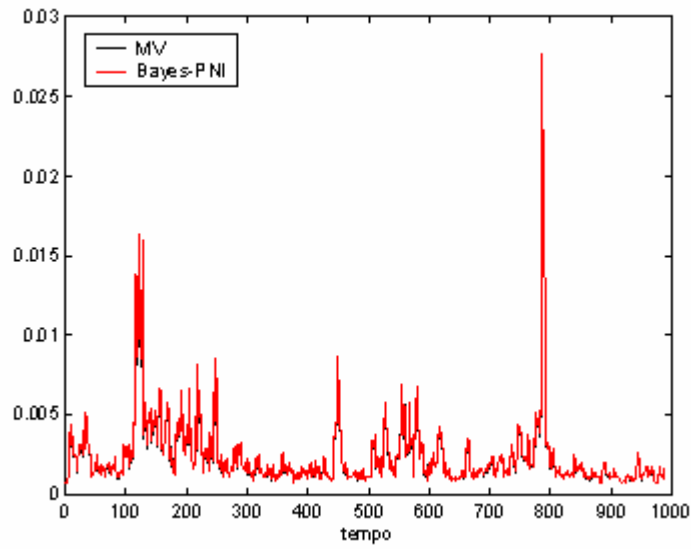
**Figura 9 – Volatilidade Estimada
Modelo AR(6)-ARCH(3) ajustado à série IBOvespa.**



**Figura 10 – Análise de resíduos - Série Telebrás
Abordagem de Máxima Verossimilhança.**



**Figura 11 – Análise de resíduos - Série Telebrás
Abordagem Bayesiana com distribuição a priori não-informativa.**



**Figura 12 – Volatilidade Estimada
Modelo AR(2)-ARCH(6) ajustado à série Telebrás.**

BIBLIOGRAFIA

AKAIKE, H. *A New Look at the Statistical Identification Model*. IEEE Trans. on Automatic Control, Vol. 19, pp. 716-723, 1974.

ANDRADE, M.G.; BARRETO, G.A. *Estimação e Análise de Robustez na Abordagem Bayesiana de Modelos AR(p): Estudo de Caso em Previsão de Vazão*. Relatório Técnico No. 97, ICMC-USP, 2000.

BARRETO, G. A.; OLIVEIRA, S. C.; ANDRADE, M. G. *Estimação Bayesiana de Parâmetros de Modelos Arch(p) via Simulação de Monte Carlo em Cadeias de Markov*. XV Congresso Brasileiro de Automática, Gramado-RS, 2004.

BERA, A.; ZUO, X. *Specification Test for a Linear Regression Model with ARCH Process*. Journal of Statistical Planning and Inference, Vol. 50, pp. 283-308, 1996.

BERNDT, E. K.; HALL, B.H.; HALL, R.E.; HAUSMAN, J.A. *Estimation Inference in Nonlinear Structural Models*. Annals of Economic and Social Measurement, N° 4. pp.653-665, 1974.

BOLLERSLEV, T. *Generalized Autoregressive Conditional Heteroskedasticity*. Journal of Econometrics, Vol. 31, pp. 307-327, 1986.

BOLLERSLEV, T.; CHOU, R.Y.; KRONER, K.F. *ARCH Modeling in Finance: a Review of the Theory and Empirical Evidence*. Journal of Econometrics, Vol. 52, pp. 5-59, 1992.

BORKOVEC, M. *Extremal Behavior of Autoregressive Process with ARCH(1) Errors*. Munich University Technology, Germany, 1991.

BOX, G. E.; JENKINS, G. M. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, 1976.

BOX, G. E.; JENKINS, G. M.; REINSEL, G. *Time Series Analysis: Forecasting and Control*. Third Edition. Englewood Cliffs: Prentice Hall, 1994.

BOX, G. E.; TIAO, G. C. *Bayesian Inference in Statistical Analysis*. Addison-Wesley, New York, 1973.

CARLIN, B.; CHIB, S. *Bayesian Model Choice via Markov Chain Monte Carlo Methods*. Journal of the Royal Statistical Society, B, Vol. 57, No. 3, pp. 473-484, 1995.

CARLIN, B.; LOUIS, T. *Bayes and Empirical Bayes Methods for Data Analysis*. 2nd ed., Vol. 47, Texts in Statistical Science Series, London: Chapman and Hall, 2000.

- CHIB, S.; GREENBERG, E. *Understanding the Metropolis-Hastings Algorithm*. The American Statistician, Vol. 49, No. 4, pp. 327-335, 1995.
- CHIB, S.; GREENBERG, E. *Bayes Inference in Regression Models with ARMA(p,q) Errors*. Journal of Econometrics, Vol. 64, pp. 183-206, 1994.
- CLARK, W. *A Subordinated Stochastic Process Model with Finite Variance for Speculative Price*. Econometrica, Vol. 41, pp. 135-155, 1973.
- COSTA, P. H. S.; BAIDYA, T.K.N. *Propriedades Estatísticas das Séries de Retornos das Principais Ações Brasileiras*. Pesquisa Operacional, Vol. 21, No. 1, pp. 61-87, 2001.
- COWLES, M.K.; CARLIN, B.P. *Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Study*. J. Amer. Statist. Assoc. Vol. 91, pp. 883-904, 1996.
- CROWDER, M. J. *Maximum Likelihood Estimation for Dependent Observations*. Journal of the Royal Statistical Society, Series B, Vol. 38, pp. 45-53, 1976.
- DUARTE, A.; HEIL, T.; PINHEIRO, M. *Previsão de Volatilidade de Ativos e Índices Brasileiros*. Resenha da BMF. No. 112, pp. 15-27, 1996.
- DUARTE, A.; MENDES, B. *Robust Estimation of Systematic Risk in Emerging Stock Markets*. Emerging Markets Quarterly, Vol. 1, No.4, pp. 85-95, 1997.
- DUTTA, H. *Large Sample Tests for a Regression Model with Autoregressive Conditional Heteroscedastic Errors*. Comm. Stat. Theory and Methods, Vol. 28, pp. 105-117, 1999.
- EFRON, B. *Bootstrap Methods: Another Look at the Jackknife*. The Annals of Statistics, Vol. 49, No. 4, pp. 2-26, 1979.
- ENGLE, R. *Financial econometrics: A New Discipline with New Methods*. Journal of Econometrics, Vol. 100, pp. 53-56, 2000.
- ENGLE, R. *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation*. Econometrica, Vol. 50, pp. 987-1007, 1982.
- ENGLE, R.; BOLLERSLEV, T. *Modelling the Persistence of Conditional Variances*. Econometric Reviews, Vol. 5, pp. 1-50, 1986.
- ENGLE, R.; LILIEN, D.; ROBINS, R. *Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model*. Econometrica, Vol. 55, No. 2, pp. 391-407, 1987.
- FERREIRA, V. A. M. *Uso de MCMC na Abordagem Bayesiana de Modelos ARCH e GARCH*. Dissertação de Mestrado, ICMC-USP, 2001.
- FRANCO, G.; KOOPMAN, S. SOUZA, R. *Bootstrap Tests when Parameters of Nonstationary Time Series Models lie on the Boundary of the Parameter Space*. Manuscript (UFMG, Minas Gerais, Brazil).
- GAMERMAN, D. *Simulação Estocástica via Cadeias de Markov*. 12º Simpósio Brasileiro de Probabilidade e Estatística, UFRJ, Rio de Janeiro – RJ, 1996.

- GELFAND, A. E.; SMITH, A. F. *Sampling – Based Approaches to Calculating Marginal Densities*. Journal of the Statistical Association, Vol. 5, No. 410, pp. 398-409, 1990.
- GELMAN, A.; CARLIN, J.; STERN, H.; RUBIN, D. *Bayesian Data Analysis*. Chapman and Hall, London, 1995.
- GELMAN, A.; RUBIN, D. *Inference from Iterative Simulation using Multiple Sequences*. Statistical Science, Vol. 7, No. 4, pp. 457-511, 1992.
- GEWEKE, J. *Evaluating the Accuracy of sampling-based approaches to calculating posterior moments (with discussion)*. Bayesian Statistics, Vol. 4, Oxford University Press, Oxford, 169-193, 1992.
- GEWEKE, J. *Exact Predictive Densities for Linear Models with ARCH Disturbances*. Journal of Econometrics, Vol. 40, pp. 63-86, 1989.
- GEWEKE, J. *Bayesian Inference in Econometric Models using Monte Carlo Integration*. Manuscript (Duke University, Durham, NC), 1986c.
- GEWEKE, J. *Exact Inference in the Inequality Constrained Normal Linear Regression Model*. Journal of Applied Econometrics, Vol. 1, pp. 127-141, 1986a.
- GIAKOUMATOS, S.; DELLAPORTAS, P.; POLITIS, D. *Bayesian Analysis of the Unobserved ARCH Model using Auxiliary Variable Sampling*. Technical Report, University of Athens, 1998.
- GIACOMINI, R.; GOTTSCHLING, A.; HAEFKE, C.; WHITE, H. *Hypernormal Densities*. Department of Economics, University of California, US, 2002.
- GÖING, A. *Estimation in Financial Models*. Department of Mathematics, Zurich, Switzerland, 1996.
- GONÇALVES, S.; KILIAN, L. *Bootstrapping Autoregression with Conditional Heteroskedasticity of Unknown Form*. Journal of Econometrics, forthcoming, 2003.
- GONÇALVES, S.; WHITE, H. *Maximum Likelihood and the Bootstrap for Nonlinear Dynamic Models*. Journal of Econometrics, forthcoming, 2003.
- HA, J.; LEE, S. *Coefficient Constancy Test in AR-ARCH Models*. Statistics and Probability Letters, Vol. 57, No. 1, pp. 65-77, 2002.
- HAMILTON, J. *Time Series Analysis*. Princeton University Press, 1994.
- HAMILTON, J. D.; SUSMEL, R. *Autoregressive Conditional Heteroskedasticity and Changes in Regime*. Journal of Econometrics, Vol. 64, pp. 307-333, 1994.
- HARVEY, A C. *Time Series Models*, 2th ed., Harvester Wheatsheaf, London, 1997.
- HEIDELBERGER, P.; WELCH, P.W. *Simulation Run Length Control in the Presence of an Initial Transient*. Operations Research, Vol. 31, pp.1109–1144, 1983.

- HEIMANN, G.; KREISS, J. *Bootstrapping General First Order Autoregression*. Statistics and Probability Letters, Vol. 30, pp. 87-98, 1996.
- ISSLER, J. V. *Estimating and Forecasting the Volatility of Brazilian Finance Series using ARCH Models*. The Brazilian Review of Econometrics, Vol. 19, No. 1, pp. 5-56, 1999.
- KARANASOS, M.; KIM, J. *Moments of the ARMA-EGARCH Model*. University of York, UK, 2001.
- KARLIS, D.; XEKALAKI, E. *Autoregressive Conditional Heteroscedasticity (ARCH) Models: A review*. Technical report, University of Athenas, 2003.
- KOOP, G. *Bayesian Semi-nonparametric ARCH Models*. Review of Econometrics and Statistics, Vol. 76, No. 1, pp. 176-181, 1994.
- LEE, J. H. H. *A Survey of ARCH and Related Models*. Department of Economics & Finance. Faculty of Business. Melbourne-Australia, 1992.
- LEITE, J. G.; SINGER, J. M. *Métodos Assintóticos em Estatística – Fundamentos e Aplicações*. 9º Simpósio Brasileiro de Probabilidade e Estatística, IME-USP, São Paulo – SP, 1990.
- LUENBERGER, D. G. *Linear and Nonlinear Programming*. Addison-Wesley, Reading, Mass, 1984.
- MAECKER, G. *Efficient Estimation in AR Models with ARCH Errors*. Institut für Mathematische Stochastik, Germany, 1991.
- MENDES, B.; DUARTE, A. *Robust Estimation of ARCH Models*. Revista de Econometria, Vol. 19, No.1, pp. 139-180, 1999.
- MIGON, H. S.; MAZUCHELI, J. *Modelos GARCH Bayesianos: Métodos Aproximados e Aplicações*. The Brazilian Review of Econometrics, Vol. 19, No. 1, pp. 111-138, 1999.
- MORETTIN, P. A.; TOLOI, C. M. *Análise de Séries Temporais*. Editora Edgard Blücher Ltda, São Paulo, 2004.
- NAKATSUMA, T. *Bayesian Analysis of ARMA-GARCH Models: A Markov Chain Sampling Approach*. Journal of Econometrics, Vol. 95, pp. 57-69, 2000.
- NAKATSUMA, T.; TSURUMI, H. *ARMA-GARCH Models: Bayes Estimation versus MLE, and Bayes Non-Stationarity Test*. Department of Economics, New Brunswick, USA, 1996.
- NELSON, D. B. *Conditional Heteroskedasticity in Asset Returns: A new approach*. Econometrica, Vol. 45, pp. 347-370, 1991.
- OLIVEIRA, J. R. *Uso do Amostrador de Gibbs e Metropolis-Hastings em Análise Bayesiana de Modelos AR(p)*. Dissertação de Mestrado, ICMC-USP, 1998.
- OLIVEIRA, S. C.; ANDRADE, M. G. *Modelos Estocásticos com Heterocedasticidade para Séries Temporais em Finanças – Tese em Andamento*. IX Simpósio de Teses e Dissertações, São Carlos-SP, 2004.

OLIVEIRA, S. C.; ANDRADE, M. G. *Estimation of Autoregressive Models with ARCH Errors using Maximum Likelihood Approach*. 16o Simpósio Nacional de Probabilidade e Estatística, Caxambu-MG, 2004.

OLIVEIRA, S. C.; ANDRADE, M. G. *Uma reparametrização do modelo MA(q) com enfoque Bayesiano*. 9o Escola de Séries Temporais e Econometria, Belo Horizonte-MG, 2001.

OLIVEIRA, S. C.; ANDRADE, M. G.; SILVA, R. G. *A Comparison of Bayesian and Bootstrapped Maximum Likelihood Approaches for ARCH Models: Evidences from Brazilian Financial Time Series*. First Brazilian Conference on Stochastic Modeling in Insurance and Finance, Ubatuba-SP, p. 250-253, 2003.

OLIVEIRA, S. C.; FERREIRA, V. A. M.; ANDRADE, M. G.; BARRETO, G. A. *Abordagem Bayesiana com MCMC vs MV em Modelos ARCH(p)*. 10o Escola de Séries Temporais e Econometria, São Pedro-SP, 2003.

PANTULA, S. *Estimation of Autoregressive Models with ARCH Errors*. Indian Journal of Statistics, Vol. 50, pp. 119-138, 1988.

PELEGRIN, C. B.; ORTEGA, E. R. *Estimación de Modelos de Volatilidad Estocástica Y Modelos de Componentes Inobservados Condicionalmente Heterocedásticos*. Memória de Doctor, Universidad Carlos III de Madrid, Departamento de Estadística, 2004.

PEREIRA, P.; HOTTA, L.; SOUZA, L.; ALMEIDA, N. *Alternative Models to Extract Asset Volatility: A Comparative Study*. The Brazilian Review of Econometrics, Vol. 19, No. 1, pp. 57-109, 1999.

PRAETZ, P. *The Distribution of Share Price Changes*. Journal of Business. Vol. 35, pp. 49-55, 1972.

POLASEK, W. *MCMC Methods for Periodic AR-ARCH Models*. Institute of Statistics and Econometrics, University of Basel, Switzerland, 2001.

POLASEK, W.; KOZUMI, H. *The VAR-VARCH Model: A Bayesian Approach*. Institute of Statistics and Econometrics, University of Basel, Switzerland, 2000.

RAFTERY, A. E.; LEWIS, S. M. *One Long Run with Diagnostics: Implementation Strategies for Markov Chain Monte Carlo*. Statistical Science, Vol. 7, 493-497, 1992.

SARKAR, N. *ARCH Model with Box-Cox Transformed Dependent Variable*. Statistics and Probability Letters, Vol. 50, pp. 365-374, 2000.

SCHWARZ, G. *Estimating the Dimension of a Model*. The Annals of Statistics, Vol. 6, No. 2, pp. 461-464, 1978.

SILVEIRA, M. A.; BARCINSKI, A.; ALMEIDA, B. C.; GARCIA, M. G. *Estimação da Volatilidade do Retorno de Ações Brasileiras: Um Método Alternativo à Família GARCH*. Departamento de Controle de Risco do Banco Icatu, Internet, 2001.

SIMS, C. *Bayesian Skepticism on Unit Root Econometrics*. Journal of Economic Dynamics and Control, Vol. 12, pp. 463-474, 1988.

SIMS, C.; UHLIG, H. *Understanding Unit Rooters: Na helicopter Toor*. *Econometrica*, Vol. 59, No. 6, pp. 1591-1599, 1991.

SOBRINHO, R. F. *Extração da Volatilidade do IBovespa*. Resenha da BMF. No. 144, pp. 17-39, 2004.

SOUZA, R. C.; CAMARGO, M. E. *Análise e Previsão de Séries Temporais: Os Modelos ARIMA*. Santa Maria-RS, 1996.

TSAY, R. S. *Financial Time Series Analysis*. 10^o Escola de Séries Temporais e Econometria, São Pedro-SP, 2003.

XEKALAKI, E.; KARLIS, D. *Choosing Initial Values for the EM Algorithm for Finite mixtures*. *Computational Statistics and Data Analysis*. Vol. 41, pp. 577-590, 2003.

WEHRENS, R.; PUTTER, H.; BUYDENS, L. *The Bootstrap: a Tutorial*. *Chemometrics and Intelligent Laboratory Systems*, Vol. 54, pp. 35-52, 2000.

WEISS, A. *ARMA Models with ARCH Errors*. *Journal of Time Series Analysis*, Vol. 5, N^o 2, pp. 129-143, 1984.