

**INFERÊNCIA BAYESIANA PARA OS
MODELOS DA CLASSE ARCH
COM POTÊNCIA ASSIMÉTRICA**

WASHINGTON SANTOS DA SILVA

2006

WASHINGTON SANTOS DA SILVA

Inferência Bayesiana para os Modelos da Classe ARCH com Potência Assimétrica

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Doutorado em Estatística e Experimentação Agropecuária, área de concentração em Teoria Matemática e Métodos Estatísticos, para obtenção do título de “Doutor”.

Orientadora: Prof.^a Dr.^a Thelma Sáfyadi

Lavras, Minas Gerais

2006

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Silva, Washington Santos da

Inferência bayesiana para os modelos da classe ARCH com potência
assimétrica / Washington Santos da Silva. -- Lavras : UFLA, 2005.
82 p. : il.

Orientador: Thelma Sáfyadi.

Tese (Doutorado) - UFLA.

Bibliografia.

1. APARCH. 2. Inferência Bayesiana. 3. Metrópolis-Hastings. 4. Ibo-
vespa. I. Universidade Federal de Lavras. II. Título.

CDD-519.542

WASHINGTON SANTOS DA SILVA

Inferência Bayesiana para os Modelos da Classe ARCH com Potência Assimétrica

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Doutorado em Estatística e Experimentação Agropecuária, área de concentração em Teoria Matemática e Métodos Estatísticos, para obtenção do título de “Doutor”.

APROVADA em 24 de Abril de 2006

Prof. Dr. Júlio Sílvio Bueno Filho	UFLA
Prof. Dr. Mário Javier Ferrua Vivanco	UFLA
Prof. Dr. Luiz Gonzaga de Castro Júnior	UFLA
Prof. Dr. Renato Ribeiro de Lima	UFLA
Prof. Dr. Wagner Moura Lamounier	UFMG

Prof.^a Dr.^a Thelma Sáfadi
UFLA
(Orientadora)

Lavras, Minas Gerais
2006

*PARA ANDRÉA, CAIO E
MELL*

*DAS COISAS BOAS EM MI-
NHA VIDA VOÇÊS SÃO, CER-
TAMENTE, A MELHOR*

Agradecimentos

À minha mulher Andréa, pelo amor, pela grande mulher e mãe em que ela se tornou e por me tornar a cada dia, pacientemente, um homem muito melhor. Aos nossos filhos Caio e Mell, que nos fazem ver todos os dias que vale a pena viver.

À minha mãe Emília e à minha sogra Mara pelas orações, pelo apoio constante e fundamental. Ao meu sogro e grande amigo Rozendo, que nos deixou recentemente e nos faz muita falta. Ao meu pai Amado, *em memória*.

À minha orientadora, Prof.^a Thelma, por sua amizade, profissionalismo e ensinamentos. Aprendi muito com sua postura profissional e humana ao mesmo tempo. Muito Obrigado por me orientar, e defender, durante o doutorado.

Aos meus professores do doutorado, Lucas, Mário e Júlio, pelos ensinamentos. Principalmente ao Professor Mário, pela grande amizade.

Aos meus colegas de curso e grandes amigos Luiz Alberto e Marcelo Cirillo, pelo companheirismo e pela força.

Ao Prof. Aladir pela amizade (e pelas caronas). Aos meus amigos da cidade de Formiga: Robson, Eugênio e Daniel.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa de estudos.

E a Deus, pelo trabalho concluído.

RESUMO

SILVA, Washington Santos. **Inferência Bayesiana para os Modelos da Classe ARCH com Potência Assimétrica**. Lavras: UFLA, 2006. 82p. (Tese - Doutorado em Estatística e Experimentação Agropecuária)^{*}

Este trabalho desenvolve uma análise bayesiana do modelo ARCH com potência assimétrica. A análise envolve a estimação de parâmetros, a predição da variância condicional e a seleção de modelos. Os procedimentos de inferência bayesiana são implementados usando-se o algoritmo de Metropolis-Hastings. Aplica-se a metodologia proposta a dados simulados e reais.

^{*}Orientadora: Thelma Sáfadi - UFLA

ABSTRACT

SILVA, Washington Santos. **Bayesian Inference for the Asymmetric Power ARCH Class of Models**. Lavras: UFLA, 2006. 82p. (Thesis - Phd Program in Statistics and Agricultural Experimentation)

This work proposed a bayesian analysis of the asymmetric power ARCH model. The analysis involve parameter estimation, conditional variance prediction and model selection. The procedures of the bayesian approach are implemented using the Metropolis-Hastings algorithm. The methodology is applied to simulated an real series.

*Supervisor: Thelma Sáfadi - UFLA

SUMÁRIO

Resumo	i
Abstract	ii
Lista de Figuras	vi
Lista de Tabelas	viii
1 Introdução	1
2 Revisão da Literatura	3
2.1 Teoria econômica e métodos bayesianos	3
2.2 Inferência bayesiana para modelos ARCH	5
2.3 Modelo APARCH	11
3 Referencial Teórico	12
3.1 Modelo GARCH com Potência Assimétrica	12
3.1.1 Modelo ARCH	12
3.1.2 Modelo GARCH	15

3.1.3	Modelo APARCH	17
3.2	Inferência Bayesiana	20
3.2.1	Conceitos básicos	21
3.2.2	Estimação pontual e intervalar	24
3.2.3	Seleção e diagnóstico preditivo de modelos	25
3.2.3.1	Fator de Bayes	25
3.2.3.2	Pseudo-fator de Bayes	27
3.2.4	Simplicidade e unicidade da inferência bayesiana	30
3.3	Métodos de Monte Carlo via Cadeias de Markov	31
3.3.1	Métodos de simulação Monte Carlo via Cadeias de Markov	31
3.3.2	O algoritmo de Metropolis-Hastings	35
3.3.3	Procedimentos para verificação da convergência da Cadeia de Markov	38
4	Material e Métodos	40
4.1	Dados	40
4.1.1	Dados simulados	40
4.1.2	Ibovespa	42
4.2	Metodologia	44
5	Resultados e Discussão	50

5.1	Série Simulada	50
5.2	Série de retornos do Ibovespa	54
6	Conclusões	59
	Referências	61
	Apêndice A - Figuras - Dados simulados	67
	Apêndice B - Figuras - retornos do Ibovespa	71
	Apêndice C - Programa R para o modelo APARCH	75

LISTA DE FIGURAS

1	Série, funções de autocorrelação e autocorrelação parcial amostrais - dados simulados	41
2	Série, funções de autocorrelação e autocorrelação parcial amostrais - Ibovespa	43
3	Esperanças a posteriori de σ_t^δ	53
4	Esperanças a posteriori de σ_t^δ - Ibovespa	57
5	Log-densidades preditivas um-passo-à-frente para 30 observações - Ibovespa.	57
6	Traço e histograma das amostras das marginais a posteriori dos parâmetros - dados simulados.	68
7	Traço e histograma das amostras das marginais a posteriori dos parâmetros - dados simulados.	69
8	Evolução dos quantis das cadeias dos parâmetros - dados simulados.	70
9	Traço e histograma das amostras das marginais a posteriori dos parâmetros - Ibovespa.	72
10	Traço e histograma das amostras das marginais a posteriori dos parâmetros - Ibovespa.	73

11	Evolução dos quantis das cadeias dos parâmetros - Ibo- vespa.	74
----	--	----

LISTA DE TABELAS

1	Ibovespa - Estatísticas	42
2	Resultados da estimação - modelo APARCH(1,1) gaussiano - Série simulada	51
3	Resultados da estimação - modelo APARCH(1,1) gaussiano - retornos do Ibovespa	54

1 INTRODUÇÃO

Modelar variâncias que mudam no tempo é a motivação básica da classe de modelos ARCH, introduzida no artigo seminal de Engle (1982). ARCH é a sigla para *Autoregressive Conditional Heteroskedasticity*. Especificamente, Robert F. Engle desenvolveu um modelo de séries temporais (processo estocástico) para a heterocedasticidade constatada em séries temporais econômicas. Esta estratégia, não apenas melhora a eficiência dos estimadores usuais, mas também fornece uma predição da variância de cada termo do erro. Tal previsão é de grande interesse em macroeconomia e, em particular, em Finanças.

A partir do trabalho seminal de Engle (1982), diversos modelos buscaram generalizar o modelo ARCH. O modelo GARCH (de *Generalized Autoregressive Conditional Heteroskedasticity*) proposto Bollerslev (1986) foi uma das generalizações mais importantes. A generalização de um modelo ARCH para um modelo GARCH é análoga, a uma generalização de um modelo autorregressivo (AR) para um modelo autorregressivo com médias-móveis (ARMA), o que gera ganhos substanciais em termos de parcimônia e qualidade do ajuste.

Apesar de a literatura clássica sobre a classe de modelos ARCH ser imensa, a literatura bayesiana, relativamente, está em sua infância, apesar dos avanços recentes. Geweke (1989) e Dellaportas et al. (2005) abordaram modelos ARCH. Geweke (1994), Bawens e Lubrano

(1998), Muller & Pole (1998) e Dellaportas et al. (2000) forneceram estratégias bayesianas para os modelos GARCH e EGARCH, além de poucos outros trabalhos.

A contribuição deste trabalho é desenvolver uma abordagem bayesiana para o modelo ARCH com potência assimétrica (APARCH, de *asymmetric power ARCH*), para o qual não há tratamento na literatura bayesiana, usando Métodos de Monte Carlo via Cadeias de Markov, especificamente, usando-se o algoritmo de Metropolis-Hastings. Tal modelo generaliza significativamente o modelo ARCH original na medida em que possui, como sub-casos, sete outros modelos propostos na literatura. Os procedimentos sugeridos envolvem a estimação de parâmetros, a previsão da variância condicional e a seleção de modelos via a densidade preditiva. Testa-se o algoritmo em dados simulados e, em seguida, aplicam-se os procedimentos desenvolvidos à uma série de retornos do Ibovespa.

Além desta introdução o restante do trabalho está organizado da seguinte forma. Na Seção 2, o referencial teórico, revisa-se a literatura sobre as abordagens bayesianas propostas para modelos da classe ARCH. Na Seção 3, metodologia, relatam-se os procedimentos a serem seguidos. Na Seção 4 constam os resultados e discussão dos mesmos. E finalmente, na Seção 5 colocam-se os comentários finais.

2 REVISÃO DA LITERATURA

Nesta seção, em primeiro lugar faz-se uma breve análise do histórico da relação entre as teorias econômica e bayesiana. Em seguida, revisa-se a literatura econométrica e estatística sobre a aplicação da metodologia bayesiana a modelos da classe ARCH. E, por fim, faz-se uma resenha da literatura sobre as aplicações do modelo APARCH.

2.1 Teoria econômica e métodos bayesianos

A teoria econômica e a teoria da decisão bayesiana possuem interconexões antigas. O trabalho clássico de Keynes (1921) sobre a teoria subjetiva lógica da probabilidade e o trabalho fundamental de Morgenstern & Neumann (1944), trabalho seminal sobre a moderna teoria da utilidade esperada, são exemplos das profundas relações entre ambas as áreas. A teoria econômica da utilidade constitui um fundamento básico das principais formulações axiomáticas da teoria da decisão bayesiana. Por sua vez, a aprendizagem bayesiana é uma metodologia há muito incorporada à teoria econômica.

Para ilustrar a relação entre as duas áreas, Zellner (1985) cita a conhecida amizade entre Milton Friedman e Leonard Savage, que escreveram a quatro mãos dois artigos influentes sobre a teoria econô-

mica da Escolha sob Incerteza, Friedman & Savage (1948) e Friedman e Savage (1952). Segundo Zellner (1985), estes artigos evidenciaram o papel central da teoria econômica da utilidade em estatística pela ênfase no papel central da maximização da utilidade esperada ou, de forma equivalente, da minimização da perda esperada no processo de escolha sob incerteza na determinação de, por exemplo, estimativas, previsões e modelos. O trabalho conjunto de Friedman e Savage ajudou a esclarecer, para economistas e estatísticos, a conexão íntima entre a teoria econômica da utilidade e a teoria estatística. Assim, haveria considerável suporte para a utilização da inferência bayesiana e dos procedimentos de decisão bayesianos em economia, os quais foram aplicados na análise de muitos problemas econométricos, obtendo-se resultados muito satisfatórios. Boyer e Kihlstrom (1984) é uma coletânea que ilustra muito bem a aplicação dos métodos bayesianos à teoria econômica.

No que concerne especificamente à Econometria, segundo Zellner (1985), a Econometria bayesiana teve início na década de 1960 com os trabalhos de Jacques Dréze, Tom Rothemberg, Walter Fischer, Albert Ando, Gordon Kaufman e do próprio Arnold Zellner, entre poucos outros. Data destes primórdios o trabalho clássico de Zellner (1971). A partir destes primórdios, o desenvolvimento e aplicação da metodologia bayesiana a problemas econométricos é crescente e contínuo. Em praticamente todas as linhas de pesquisa econômica, a consulta a revistas da área revela algum trabalho bayesiano (Zellner (1997)). Além dos pioneiros citados anteriormente, atualmente, economistas como Dale Poirier, John Geweke, Luc Bawens, Michel Lubrano, Jean-François Richard, Gary Koop, Herman Van Dijk, entre outros, têm contribuições

importantes à econometria e à estatística bayesianas.

Uma tendência recente e importante é o surgimento de livros de Econometria bayesiana em nível de graduação, como Koop (2003), Poirier (1995) e Lancaster (2004). Tal fato certamente terá um impacto significativo na disseminação da metodologia bayesiana em Econometria, na medida em que as novas gerações serão expostas à metodologia bayesiana já na graduação. Além desses, publicaram-se livros mais avançados importantes, como Bawens et al. (1999), cujo foco principal é a abordagem bayesiana de modelos econométricos envolvendo séries temporais, no qual se tratam tópicos importantes e controversos, como a abordagem bayesiana ao problema da raiz unitária. Geweke (2005) aborda diversos tópicos importantes, como as técnicas modernas para seleção bayesiana de modelos. O livro de Rossi, Allenby e McCulloch (2006) explora a abordagem bayesiana de modelos que representam o estado da arte na área de microeconomia e marketing.

2.2 Inferência bayesiana para modelos ARCH

O artigo seminal da literatura sobre inferência bayesiana aplicada a modelos ARCH foi escrito por Geweke (1989). Neste artigo, além de desenvolver uma estratégia de estimação via amostragem por importância com a função de importância sendo uma distribuição t-Student, o autor cria um algoritmo baseado em simulação para a obtenção das densidades preditivas exatas m -passos à frente. A estimação utiliza-se de uma reparametrização do modelo ARCH original. A distribuição *a priori*, é uma variante da distribuição *a priori* não-informativa de Jef-

freys. Aplica-se a metodologia desenvolvida para obter o ajuste diário em uma carteira de ativos.

O foco do trabalho de Geweke (1994) não foram os modelos ARCH, mas métodos para a obtenção de verossimilhanças marginais, fundamentais para a seleção bayesiana de modelos. Entretanto, como uma das ilustrações das técnicas apresentadas, Geweke (1994) estimou um modelo GARCH gaussiano com distribuições a priori não-informativas utilizando o algoritmo de Metropolis-Hastings independente, tendo utilizado como distribuição candidata uma distribuição t-Student Multivariada. Apesar de não ser um trabalho publicado, é muito citado na literatura.

Kleinbergen & Dijk (1993) desenvolveram uma abordagem bayesiana ao problema da verificação das condições de não-estacionariedade da variância condicional em modelos GARCH e à análise de raiz unitária para o modelo AR(1) com erros GARCH com distribuição condicional t-Student. Interpretando a não-estacionariedade em termos de persistência infinita de choques, os autores obtiveram resultados diferentes dos normalmente reportados na literatura. Os modelos GARCH e Autorregressivos com erros GARCH analisados, com distribuições a priori não-informativas, foram estimados como em Geweke (1989), isto é, via amostragem por importância com a função de importância sendo uma t-Student multivariada.

Muller & Pole (1998) propuseram um novo esquema híbrido de simulação de Monte Carlo via Cadeias de Markov para a estimação dos parâmetros de modelos GARCH multivariados (modelos *VECH*). Em primeiro lugar, Muller & Pole (1998) formulam os modelos GARCH

multivariados como modelos dinâmicos bayesianos, isto é, como modelos em espaço de estados. O algoritmo de estimação proposto é uma variação do algoritmo de Metropolis-Hastings independente. Apesar de muito geral, esta abordagem é ainda pouco operacional, fato confirmado por Migon & Mazucheli (1999). Trata-se de um algoritmo relativamente complexo.

Bawens e Lubrano (1998) estimaram o modelo GARCH via uma variação do algoritmo amostrador de Gibbs conhecida como amostrador Griddy-Gibbs; este algoritmo foi proposto na literatura por Ritter & Tanner (1992). O amostrador de Gibbs é eficiente e elegante quando se conhecem as distribuições condicionais completas dos parâmetros; no caso dos modelos GARCH estas distribuições não são conhecidas. A idéia do algoritmo Griddy-Gibbs é avaliar o kernel das distribuições condicionais em um *griddy* de pontos, possibilitando, então, o cálculo da função de distribuição correspondente via uma regra de integração determinística, como a regra de Simpson iterativa. Em seguida, pode-se gerar uma amostra de cada parâmetro pela inversão da distribuição em um valor aleatório amostrado uniformemente no intervalo $[0, 1]$. Para analisar a eficiência do algoritmo, os autores implementaram um pequeno estudo piloto de simulação comparando o desempenho do amostrador Griddy-Gibbs com a amostragem por importância e o algoritmo de Metropolis-Hastings. Os critérios utilizados no estudo foram a acurácia na estimação pontual, a qualidade das densidades marginais e o tempo de execução. As distribuições a priori foram não-informativas. Os resultados reportados mostram que em termos de estimação pontual, os algoritmos têm desempenho muito parecido. Em termos da qualidade do gráfico das densidades margi-

nais, o algoritmo Griddy-Gibbs conseguiu obter gráficos acurados em poucas iterações, enquanto a amostragem por importância conseguiu densidades de boa qualidade somente após muitas iterações e o algoritmo de Metropolis-Hastings não produziu densidades tão acuradas quanto as duas técnicas anteriores. Em termos de tempo de execução, o algoritmo Griddy-Gibbs foi aproximadamente 10 a 15 vezes mais lento que as outras duas técnicas. Apesar do tempo de execução mais lento, no algoritmo Griddy-Gibbs não há a necessidade do ajuste, muitas vezes complexo, de uma função por importância, no caso da amostragem por importância, e da distribuição candidata no caso do algoritmo Metropolis-Hastings. Com base nestes resultados, Bawens e Lubrano (1998) concluem que o algoritmo Griddy-Gibbs constitui uma alternativa relevante. Além deste estudo piloto, os autores aplicaram a metodologia desenvolvida para estimar um modelo GARCH assimétrico com as inovações seguindo uma distribuição t-Student para uma série do índice da Bolsa de Valores de Bruxelas.

Migon & Mazucheli (1999) é o único artigo publicado na literatura econométrica brasileira sobre o tema. Como Muller & Pole (1998), mas para o caso univariado, Migon & Mazucheli (1999) formulam o modelo GARCH gaussiano como um modelo dinâmico bayesiano e apresentam algumas técnicas numéricas para a estimação dos modelos considerados. Os algoritmos numéricos considerados são a quadratura gaussiana com base fixa que usa toda a informação disponível e a aproximação de Laplace na forma exponencial. O modelo GARCH gaussiano analisado é estimado para séries de retornos do Cbond, da taxa de câmbio Real/Dólar e da Telebrás. Migon & Mazucheli (1999) também fazem um estudo de desempenho da acurácia preditiva da abordagem bayesiana

dinâmica com estrutura GARCH em relação à abordagem bayesiana dinâmica sem estrutura GARCH, as conclusões são relativamente ambíguas, com ligeira vantagem para os modelos dinâmicos com estrutura GARCH.

Nakatsuma (2000) desenvolveu uma metodologia para a estimação de modelos ARMA com erros GARCH gaussianos e utilizando distribuições a priori normais (provavelmente motivado por um argumento assintótico frequentista, ou seja, os estimadores pontuais bayesianos teriam distribuição assintoticamente gaussiana), a qual é uma combinação do algoritmo baseado no amostrador de Gibbs proposto por Chib & Greenberg (1994) para a estimação bayesiana de modelos de regressão com erros ARMA, com o algoritmo do tipo Metropolis proposto por Muller & Pole (1998). Uma característica importante do trabalho é que o autor ilustra como é fácil testar condições complexas de não-estacionariedade dos modelos GARCH.

O artigo de Dellaportas et al. (2000) refere-se a inferência bayesiana completa pelo fato de abordar a estimação de parâmetros, a seleção de modelos e a predição do processo de volatilidade de modelos GARCH e EGARCH gaussianos e com as inovações seguindo uma distribuição de erros generalizada. Os modelos são estimados via o algoritmo de Metropolis-Hastings independente com a distribuição candidata sendo uma distribuição normal multivariada. O desenvolvimento de um algoritmo para a implementação do procedimento denominado na literatura de *Bayesian model averaging* é uma contribuição inovadora do artigo. As técnicas desenvolvidas são aplicadas a dados simulados e a uma série de dados do índice da Bolsa de Valores de Atenas.

Goldman & Tsurumi (2005) aplicam a metodologia de estimação proposta por Nakatsuma (2000) para estimar modelos ARMA-GARCH gaussianos; a inovação do trabalho é a aplicação desta modelagem para o teste da teoria da paridade do poder de compra das taxas de câmbio, impondo dupla-truncagem nas variáveis, isto é, a taxa de câmbio pode assumir valores apenas em um intervalo fechado pré-especificado, como é o caso da regime cambial nos países membros da União Européia.

Kaufmann & Fruhwirth-Schnatter (2002) propuseram uma análise bayesiana do modelo ARCH gaussiano com mudança de regime markoviana proposto por Hamilton & Susmel (1994). A estratégia de estimação de Monte Carlo via Cadeias de Markov é uma combinação de amostradores multi-movimentos (ver Chib (1996) e Carter & Kohn (1994)) com o algoritmo proposto por Nakatsuma (2000). O amostrador de Gibbs é utilizado para simular as probabilidades de transição do processo markoviano latente, uma vez que as distribuições condicionais completas são conhecidas, para os demais parâmetros, cujas densidades condicionais completas são desconhecidas, utiliza-se o algoritmo de Metropolis-Hastings. Kaufmann & Fruhwirth-Schnatter (2002), utilizando a prioris informativas, implementam a seleção bayesiana de modelos via fatores de bayes.

Dellaportas et al. (2005) apresentam uma abordagem bayesiana ao que denominam modelo ARCH não-observado, ou seja, as componentes do modelo ARCH são observadas com erro. O algoritmo MCMC proposto procura simplificar ao máximo o procedimento de amostragem dos parâmetros do modelo, por exemplo, os autores incluem variáveis latentes de forma que os parâmetros tenham densidades condicionais completas com formas conhecidas. O algoritmo de estimação é com-

posto somente de passos Gibbs, com a utilização do amostrador de variáveis auxiliares para amostrar as variáveis latentes.

2.3 Modelo APARCH

Desde o artigo de Ding et al. (1993) o qual introduziu o modelo APARCH, diversos trabalhos têm utilizado esta especificação, e algumas extensões foram propostas. A Seguir encontram-se comentários sobre alguns destes estudos.

Tse & Tsui (1997) utilizaram o processo APARCH para analisar a volatilidade de taxas de câmbio, concluindo que esta especificação modelava adequadamente as taxas de câmbio analisadas.

Giot & Laurent (2003) analisaram o desempenho do modelos da classe ARCH na estimação do Valor em Risco (VaR) para posições compradas e vendidas e reportaram que as estimativas de VaR do modelo APARCH forneceram previsões mais acuradas que a metodologia tradicional, dentro e fora da amostra.

Rengifo & Rombouts (2004) propuseram um modelo de seleção de portfolios dinâmicos que maximiza os retornos esperados sujeitos a uma restrição do Valor em Risco, com o objetivo de se fazerem recomendações futuras de investimento. Os autores utilizaram o modelo APARCH e o modelo GARCH para estimar o Valor em Risco, relatando que o modelo APARCH apresentou desempenho preditivo superior ao modelo GARCH.

3 REFERENCIAL TEÓRICO

Na seção 3.1 descreve-se o modelo GARCH com potência assimétrica, entretanto, para enfatizar a evolução do modelo ARCH original, apresentam-se antes os modelos ARCH e GARCH. Os procedimentos de inferência bayesiana são descritos na seção 3.2. Enquanto na seção 3.3, apresenta-se um resumo da teoria dos métodos de simulação de Monte Carlo via Cadeias de Markov e uma abordagem mais detalhada do algoritmo de Metropolis-Hastings

3.1 Modelo GARCH com Potência Assimétrica

Inicialmente apresenta-se o modelo ARCH original de Engle (1982). Em seguida, a generalização proposta por Bollerslev (1986) e, por último, a generalização construída por Ding et al. (1993).

3.1.1 Modelo ARCH

Um processo ARCH pode ser definido em vários contextos. O modelo ARCH(p) proposto no artigo seminal de Engle (1982) pode ser

sumarizado por:

$$y_t = x_t' \xi + \varepsilon_t, \quad (3.1a)$$

$$\varepsilon_t | \Psi_{t-1} \sim N(0, \sigma_t^2), \quad (3.1b)$$

$$\sigma_t^2 \equiv E[\varepsilon_t^2 | \Psi_{t-1}] = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 \equiv \omega + \alpha(L) \varepsilon_t^2, \quad (3.1c)$$

$$\omega > 0, \alpha_i \geq 0, \forall i = 1, \dots, p, \quad (3.1d)$$

em que $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_p L^p$, sendo L o operador de defasagens.

As restrições representadas por (3.1d) asseguram a positividade da variância. x_t é um vetor $k \times 1$ de variáveis exógenas, o qual pode incluir valores defasados da variável dependente $\{y_t\}$, e $\{\xi\}$ é um vetor $k \times 1$ de parâmetros. O modelo ARCH, representado pelas equações (3.1b) e (3.1c), caracteriza a distribuição das inovações ε_t condicional aos valores realizados do conjunto de informação $\{\Psi_{t-1} = y_{t-1}, x_{t-1}, y_{t-2}, x_{t-2}, \dots\}$. Note-se que como $\{\varepsilon_{t-i} = y_{t-i} - x_{t-i}' \xi\}$, $\{\sigma_t^2\}$ é uma função dos elementos de $\{\Psi_{t-1}\}$.

Conforme Bollerslev et al. (1994), a característica marcante do modelo ARCH não é o fato de a variância condicional $\{\sigma_t^2\}$ ser uma função do conjunto de informação $\{\Psi_{t-1}\}$, mas sim a forma funcional especificada. Episódios de volatilidade são geralmente caracterizados pelo agrupamento de grandes choques na variável dependente. O processo da variância condicional (3.1c) é formulado de forma a reproduzir este fenômeno. No modelo de regressão, um grande choque é representado por um grande desvio de $\{y_t\}$ em relação à sua média condicional $x_t' \xi$ ou, de forma equivalente, um grande valor positivo ou negativo de $\{\varepsilon_t\}$.

No modelo de regressão ARCH, a variância da inovação atual, condicional aos valores realizados das inovações defasadas $\{\varepsilon_{t-i}, i = 1, \dots, p\}$, é uma função crescente da magnitude das inovações defasadas, seja qual for o sinal. Portanto, em concordância com a regularidade empírica denominada agrupamento da volatilidade, grandes erros de qualquer sinal tendem a ser seguidos por grandes erros de qualquer sinal. De forma similar, pequenos erros de qualquer sinal tendem a ser seguidos por pequenos erros de qualquer sinal. A ordem do lag p determina o intervalo de tempo no qual um choque persiste em condicionar a variância dos erros subsequentes. Quanto maior o valor de p , maior a duração dos episódios de volatilidade.

Uma representação muito conveniente de um processo ARCH é

$$\varepsilon_t = z_t \sqrt{\sigma_t^2}, \quad (3.2)$$

em que $z_t \sim IID N(0, 1)$ e $\{\sigma_t^2\}$ é dado por (3.1c). Posto que $\{\sigma_t^2\}$ é uma função de $\{\Psi_{t-1}\}$ e, portanto, fixado quando condicionado à Ψ_{t-1} , pode-se verificar que $\{\varepsilon_t\}$ como dado em (3.2) será condicionalmente normal com $E(\varepsilon_t | \Psi_{t-1}) = \sqrt{\sigma_t^2} E(z_t | \Psi_{t-1}) = 0$ e $Var(\varepsilon_t | \Psi_{t-1}) = \sigma_t^2 Var(z_t | \Psi_{t-1}) = \sigma_t^2$. Portanto, o processo especificado por (3.2) é idêntico ao processo ARCH (3.1c). A equação geradora (3.2) demonstra que um processo ARCH reescala um processo subjacente de inovação gaussiana $\{z_t\}$ pela multiplicação deste pelo desvio padrão condicional, o qual é uma função do conjunto de informação $\{\Psi_{t-1}\}$.

Conforme Bollerslev (1986), note-se a semelhança entre o pro-

cesso (3.1c) e um modelo AR(p):

$$y_t = \phi_0 + \sum_{i=1}^p \phi_i y_{t-i} + \epsilon_t \equiv \phi_0 + \phi(L)y_t + \epsilon_t.$$

De fato, definindo $\{\nu_t \equiv \varepsilon_t^2 - \sigma_t^2\}$ e $\nu_t \sim N(0, 1)$, pode-se reescrever o modelo ARCH(p) em 3.1c como,

$$\varepsilon_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \nu_t, \quad (3.3)$$

e tal modelo corresponde diretamente a um modelo AR(p) para os erros quadráticos, $\{\varepsilon_t^2\}$. Este processo é fracamente estacionário se, e somente se, a soma dos parâmetros autorregressivos positivos for menor que um, conforme Engle (1982).

3.1.2 Modelo GARCH

Nas primeiras aplicações empíricas do modelo ARCH à relação entre o nível e a volatilidade da taxa de inflação, Engle (1982) verificou que se requeria um grande valor para o lag p no processo da variância condicional. Isto implicava na necessidade de estimar um grande número de parâmetros sujeitos a restrições de desigualdade.

Para contornar este problema, Bollerslev (1986) propôs o modelo ARCH generalizado, ou modelo GARCH(p, q). Seja $\{\varepsilon_t\}$ um processo

estocástico em tempo discreto com valores reais, sendo definido por,

$$y_t = x_t' \xi + \varepsilon_t, \quad (3.4a)$$

$$\varepsilon_t | \Psi_{t-1} \sim D(0, \sigma_t^2), \quad (3.4b)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 = \omega + \alpha(L) \varepsilon_t^2 + \beta(L) \sigma_t^2, \quad (3.4c)$$

no qual as restrições de desigualdade,

$$\omega > 0; \quad \alpha_i \geq 0, \forall i = 1, \dots, p; \quad \beta_i \geq 0, \forall j = 1, \dots, q, \quad (3.5)$$

são impostas para assegurar a positividade estrita de $\{\sigma_t^2\}$, $D(\cdot)$ é uma distribuição paramétrica e $\alpha(L)$ e $\beta(L)$ são polinômios de ordem finita no operador de defasagens L . Para $q = 0$, o processo reduz-se a um processo ARCH(p), e para $p = q = 0$, ε_t é simplesmente um processo ruído branco. No processo ARCH(p) a variância condicional é especificada como uma função somente das variâncias amostrais passadas, enquanto o processo GARCH(p, q) permite, além destas, a inclusão de variâncias condicionais defasadas ((BOLLERSLEV, 1986)).

A formulação do processo GARCH(p, q) pode ser compreendida a partir de uma analogia direta com os modelos da classe AR(p), no qual a solução parcimoniosa é incluir termos MA(q), dando origem a um modelo ARMA(p, q). Para verificar como um modelo GARCH fornece uma flexibilidade similar à de um modelo ARMA, considere um modelo GARCH(1, 1)

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \quad (3.6)$$

definido como em (3.3), $\{\nu_t \equiv \varepsilon_t^2 - \sigma_t^2\}$ como o erro de predição da

variância condicional, com a propriedade $\nu_t \sim N(0, 1)$; pode-se resolver (3.6) em termos de $\{\varepsilon_t^2\}$ e $\{\nu_t\}$ defasados e atuais para obter,

$$\varepsilon_t^2 = \omega + (\alpha + \beta)\varepsilon_{t-1}^2 - \beta\nu_{t-1} + \nu_t, \quad (3.7)$$

que é um processo ARMA(1, 1) para $\{\varepsilon_t^2\}$. Conforme Bollerslev (1986), através do mesmo procedimento, pode-se mostrar que um processo GARCH (p, q) corresponde a um processo ARMA($\max(p, q), q$) para $\{\varepsilon_t^2\}$, assim,

$$\varepsilon_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \varepsilon_{t-j}^2 - \sum_{j=1}^q \beta_j \nu_{t-j} + \nu_t, \quad (3.8)$$

ou

$$\phi(L)(1 - L)\varepsilon_t^2 = \omega + [1 - \beta(L)]\nu_t, \quad (3.9)$$

em que $\phi(L) \equiv [1 - \alpha(L) - \beta(L)](1 - L)^{-1}$ tem ordem $(m - 1)$. Conforme Bollerslev et al. (1994), esta analogia com os modelos da classe ARMA permite o uso das técnicas usuais na identificação das ordens p e q .

3.1.3 Modelo APARCH

Ding et al. (1993) efetuaram um estudo de Monte-Carlo cuja principal conclusão é que não há razão óbvia de porque se deve assumir que o processo da variância condicional seja uma função linear dos resíduos quadráticos defasados, como no modelo GARCH(p,q) de Bollerslev, ou que o processo do desvio-padrão condicional seja uma função linear dos resíduos absolutos defasados como no modelo de Taylor (1986). Ou seja, no modelo de Taylor, modela-se o desvio-padrão condicional (σ_t) , ao invés da variância condicional. Isto posto, Ding

et al. (1993) propuseram um modelo que generaliza substancialmente os modelos da classe GARCH. Denominou-se tal modelo de modelo GARCH com potência assimétrica (*Asymmetric power ARCH*), que será denominado modelo APARCH(p,q), o qual pode ser representado por,

$$\varepsilon_t = z_t \sqrt{\sigma_t}, \quad z_t \sim N(0, 1) \quad (3.10)$$

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta, \quad (3.11)$$

$$\omega > 0, \quad (3.12)$$

$$\delta \geq 0, \quad (3.13)$$

$$\alpha_i \geq 0, \quad (3.14)$$

$$\beta_j \geq 0 \quad \text{e} \quad (3.15)$$

$$-1 < \gamma_i < 1. \quad (3.16)$$

O parâmetro ω é a constante da equação do processo do desvio-padrão condicional, os parâmetros α_i representam os termos ARCH e os parâmetros β_i , os termos GARCH. O modelo concilia a flexibilidade de um expoente variante com o coeficiente de assimetria. Desta forma, este modelo é capaz de detectar impactos assimétricos de choques sobre a volatilidade. Especificamente, se $\gamma > 0$, verifica-se o efeito alavancagem, isto é, choques negativos tem um impacto maior sobre a volatilidade da série que choques positivos.

O modelo impõe uma transformação Box-Cox para o processo do desvio-padrão condicional e para os resíduos absolutos assimétricos. A forma funcional do processo do desvio-padrão condicional é familiar para economistas, trata-se da função de produção com elasticidade

de substituição constante (CES). Conforme Ding et al. (1993), esta especificação permite linearizar modelos que, de outra forma, seriam não lineares

Caso se assuma que a distribuição condicional dos dados seja normal, Ding et al. (1993) mostraram que a condição de estacionariedade fraca para o modelo APARCH(p,q) é dada por,

$$\frac{1}{\sqrt{2\pi}} \sum_{i=1}^p \alpha_i \{(1+\gamma_i)^\delta (1-\gamma_i)^\delta\} 2^{\frac{\delta-1}{2}} \Gamma\left(\frac{\delta+1}{2}\right) + \sum_{j=1}^q \beta_j < 1 \quad (3.17)$$

A abordagem baseada na verossimilhança é o procedimento mais utilizado na estimação de modelos ARCH, sendo que o processo de construção da função de verossimilhança é o mesmo para vários modelos da classe ARCH. É conveniente colocar os parâmetros a serem estimados em um vetor θ de dimensão $(a \times 1)$,

$$\theta \equiv (\alpha, \beta, \gamma, \delta).$$

Isto posto, a função de log verossimilhança condicional às primeiras m observações é,

$$L_T(\theta) = \sum_{t=1}^T \ln(f(y_t | \Psi_{t-1}; \theta)) = \sum_{t=1}^T l_t(\cdot), \quad (3.18)$$

em que $l_t(\cdot)$ é a contribuição individual da verossimilhança, especificamente,

$$L_T(\theta) = -\left(\frac{T}{2}\right) \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln(\sigma_t^2) - \frac{1}{2} \sum_{t=1}^T \left(\frac{\varepsilon_t^2}{\sigma_t^2}\right). \quad (3.19)$$

Para um dado valor inicial do vetor de parâmetros θ , a sequência de variâncias condicionais pode ser calculada pela equação (3.11), sendo então utilizada para calcular a função de verossimilhança condicional (3.19).

Para enfatizar a generalidade do modelo APARCH(p,q), note-se que tal modelo incorpora sete modelos da classe GARCH propostos na literatura:

- O modelo ARCH de Engle (1982), quando $\delta = 2$, $\gamma_i = 0$ e $\beta_j = 0$.
- O modelo GARCH de Bollerslev (1986), quando $\delta = 2$ e $\gamma_i = 0$.
- O modelo GARCH de Taylor (1986), quando $\delta = 1$ e $\gamma_i = 0$.
- O modelo GJR de Glosten et. al. (1993), se $\delta = 2$.
- O modelo TARCH de Zakoian (1994), quando $\delta = 1$.
- O modelo NARCH de Bera & Higgins (1992), quando $\gamma_i = 0$ e $\beta_j = 0$
- O modelo Log-ARCH de Geweke (1992), quando $\delta \rightarrow 0$.

3.2 Inferência Bayesiana

Esta seção baseia-se, sobretudo, em Zellner (1971), Koop (2003) e Gelman et al. (2003)

3.2.1 Conceitos básicos

Os procedimentos de inferência bayesiana são baseados em simples propriedades da teoria da probabilidade. A simplicidade e unidade teóricas são duas das vantagens da abordagem bayesiana. Todas as coisas que um pesquisador deseja saber, tais como as estimativas dos parâmetros de um modelo, avaliar o desempenho de diferentes modelos e obter previsões, envolvem as mesmas regras de probabilidade. Assim, os métodos bayesianos são extremamente gerais e podem ser utilizados em quaisquer circunstâncias em que um pesquisador estiver interessado em utilizar dados para aprender sobre algum fenômeno.

Para mostrar a simplicidade da abordagem bayesiana, considere-se duas variáveis aleatórias contínuas X e Y . Propriedades simples da teoria das probabilidades implicam que,

$$P(X \cap Y) = P(X|Y)P(Y), \quad (3.20)$$

sendo $P(X \cap Y)$ a probabilidade conjunta da ocorrência de X e Y , $P(X|Y)$ a probabilidade condicional de ocorrência de X , dada a ocorrência de Y , e $P(Y)$, a probabilidade marginal de ocorrência de Y . De forma similar, pode-se reverter os papéis de X e Y e encontrar uma outra expressão para a probabilidade conjunta de X e Y ,

$$P(X \cap Y) = P(Y|X)P(X). \quad (3.21)$$

Igualando as duas expressões e arranjando os termos, obtém-se a versão mais simples do teorema de Bayes, que é a peça fundamental dos procedimentos bayesianos de inferência:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}. \quad (3.22)$$

Considere-se um modelo de regressão clássico, no qual o interesse frequentemente se centra sobre os coeficientes do modelo, e o pesquisador está interessado em estimar estes coeficientes, ou seja, os coeficientes são os parâmetros sob estudo. Seja Y um vetor ou matriz de dados e Θ um vetor ou matriz que contém os parâmetros do modelo que busca explicar Y . O pesquisador está interessado em aprender sobre Θ , baseado nos dados Y . Os procedimentos bayesianos utilizam o teorema de Bayes para aprender sobre os parâmetros, dada a informação contida nos dados.

Assim, substituindo Y por θ e X por Y , obtém-se,

$$P(\theta|Y) = \frac{P(Y|\theta)P(\theta)}{P(Y)}. \quad (3.23)$$

Do Teorema da Probabilidade Total, tem-se que,

$$P(Y \cap \theta) = P(Y|\theta)P(\theta) \rightarrow P(Y) = \int P(\theta)P(Y|\theta) d\theta. \quad (3.24)$$

Portanto, pode-se reescrever o teorema de Bayes como,

$$P(\theta|Y) = \frac{P(Y|\theta)P(\theta)}{\int P(\theta)P(Y|\theta) d\theta}. \quad (3.25)$$

Uma vez que $\int P(\theta)P(Y|\theta) d\theta$ não é uma função de θ , pode-se escrever,

$$P(\theta|Y) \propto P(Y|\theta)P(\theta) \quad (3.26)$$

$P(\theta|Y)$ é de fundamental interesse em inferência bayesiana. Esta

densidade responde diretamente à questão “Dadas as observações, o que se sabe sobre θ ”? Denomina-se a densidade a posteriori para o vetor de parâmetros θ , dada a informação amostral Y . O termo densidade a posteriori denota que a incerteza sobre os parâmetros, na lógica da inferência bayesiana, deve sempre ser representada por uma distribuição de probabilidades. Especificamente a incerteza sobre os parâmetros, após observarem-se os dados, é modelada pela densidade a posteriori. Portanto, a inferência bayesiana implica em um processo de aprendizagem sobre algo desconhecido (parâmetros de um modelo), dado o que é conhecido (os dados), e considerando a densidade condicional do desconhecido, dado o que é conhecido, como sendo a melhor forma de sumarizar o que foi aprendido. O teorema de Bayes pode ser entendido como uma regra de atualização na qual os dados permitem atualizar o conhecimento a priori sobre θ ; o resultado do processo de atualização é a posteriori, que combina a informação amostral e a informação não-amostral.

$P(\theta)$ é denominada distribuição a priori dos parâmetros, esta densidade sumariza o conhecimento em relação a θ , antes de obterem-se os dados. A densidade a priori é um aspecto controverso da teoria bayesiana. Existem distribuições a priori informativas e não-informativas em relação aos parâmetros. $P(Y|\theta)$ é a função de verossimilhança, que não é uma densidade. O princípio da verossimilhança postula que toda a informação relevante sobre um fenômeno, após obterem-se os dados, está contida na função de verossimilhança.

3.2.2 Estimação pontual e intervalar

Do exposto anteriormente, deve estar claro que os procedimentos bayesianos de estimação, quando implementados, fornecem a distribuição de probabilidade a posteriori completa para o vetor de parâmetros θ . Caso se deseje, pode-se resumir esta distribuição usando um pequeno número de medidas de posição como estimativas pontuais. No contexto da teoria da decisão pode-se ter uma função perda $L(\theta, \hat{\theta})$, sendo $\hat{\theta} = \hat{\theta}(y)$ uma estimativa pontual que depende das observações $y = (y_1, \dots, y_n)$. Num contexto de decisão, um princípio bayesiano geralmente utilizado para se obter estimativas pontuais, é encontrar o valor de $\hat{\theta}$ que minimize a esperança da função perda, ou seja,

$$\arg \min_{\hat{\theta}} E(L(\theta, \hat{\theta})) = \arg \min_{\hat{\theta}} \int_{\Theta} L(\theta, \hat{\theta}) p(\theta|y) d\theta. \quad (3.27)$$

Tal expressão assume, implicitamente, que $E(L(\theta, \hat{\theta}))$ é finita e que exista um mínimo, lembrando que, sob funções perda quadráticas, a média da distribuição a posteriori é o estimador pontual ótimo, e que, sob funções perda erro absoluto, a mediana da distribuição a posteriori é o estimador pontual ótimo.

Seja θ o vetor de parâmetros de interesse, $\theta \in \mathbb{R}^p$. Seja $\omega = g(\theta)$ um vetor de dimensão m de funções de θ , o qual é definido em uma região do espaço paramétrico Θ , com $m \leq p$. Seja C uma região em Θ , tal que, $C \subseteq \Theta$.

Definição 1 *O conjunto $C \subseteq \Theta$ é um conjunto com $100(1 - \alpha)\%$ de*

credibilidade em relação à $P(\omega|y)$ se:

$$P(\omega \in C|y) = \int_C P(\omega|y)d\omega = 1 - \alpha. \quad (3.28)$$

Por exemplo, suponha que $\omega = g(\theta) = \theta_j$, seja um coeficiente único. Assim, um intervalo com 95% de credibilidade para θ_j é qualquer intervalo $[a, b]$, tal que,

$$P(a \leq \theta_j \leq b|y) = \int_a^b P(\theta_j|y)d\theta = 0,95. \quad (3.29)$$

Como para um dado nível de credibilidade, existem inúmeros intervalos de credibilidade, é prática comum reportar o intervalo com menor área, neste caso denomina-se o intervalo de credibilidade de intervalo com densidade a posteriori máxima (*highest posterior density interval*). Formalmente,

Definição 2 *Um intervalo com densidade a posteriori máxima para ω é um intervalo com $100(1 - \alpha)\%$ de credibilidade para ω com a propriedade de que tal intervalo tem menor comprimento que qualquer outro intervalo com $100(1 - \alpha)\%$ de credibilidade para ω .*

3.2.3 Seleção e diagnóstico preditivo de modelos

3.2.3.1 Fator de Bayes

Em geral, um pesquisador analisa diversos modelos completos simultaneamente, torna-se então natural avaliar o desempenho relativo destes modelos. Formalmente, um modelo M_i completo é caracterizado por uma função de verossimilhança e uma densidade a priori.

Suponha-se que existam m modelos diferentes, M_i para $i = 1, \dots, m$, que tentam explicar y . Os modelos M_i dependem de parâmetros θ_i . Portanto, a densidade a posteriori para os parâmetros calculados sob M_i pode ser expressa como:

$$P(\theta^i|y, M_i) = \frac{P(y|\theta^i, M_i)P(\theta^i|M_i)}{P(y|M_i)}. \quad (3.30)$$

A lógica da inferência bayesiana sugere que seja utilizado o teorema de Bayes para estimar uma probabilidade sobre o que não se conhece (isto é, qual modelo é mais consistente com os dados), condicional ao que se conhece (os dados). Tal probabilidade é denominada probabilidade a posteriori do modelo (*posterior model probability*), a qual pode ser utilizada para calcular o grau de suporte que os dados fornecem para o modelo M_i . Utilizando o teorema de Bayes tem-se que:

$$P(M_i|y) = \frac{P(y|M_i)P(M_i)}{P(y)}, \quad (3.31)$$

sendo $P(M_i)$ a probabilidade a priori do modelo (*prior model probability*). Uma vez que não envolve os dados, a probabilidade a priori do modelo quantifica o quão provável se considera o modelo M_i consistente com os dados antes de se observarem os dados. $P(y|M_i)$ é denominada *verossimilhança marginal* e é calculada utilizando (3.31) e algumas manipulações. Especificamente, integrand ambos os lados de (3.31) em relação à θ_i , sabendo-se que $\int P(\theta_i|y, M_i) d\theta_i = 1$ e rearranjando os termos, obtém-se,

$$P(y|M_i) = \int P(y|\theta_i, M_i)P(\theta_i|M_i) d\theta_i, \quad (3.32)$$

note-se que a verossimilhança marginal depende somente da densidade a priori e da função de verossimilhança. Diversos métodos têm sido sugeridos na literatura para calcular (3.32).

Dado que o cálculo direto do denominador de (3.31) é complexo na maioria das situações, tornou-se prática comum comparar dois modelos i e j , utilizando a razão de chances a posteriori (*posterior odds ratio*), que é simplesmente a razão das probabilidades a posteriori, isto é,

$$PO_{ij} = \frac{P(M_i|y)}{P(M_j|y)} = \frac{P(y|M_i)P(M_i)}{P(y|M_j)P(M_j)}. \quad (3.33)$$

Em diversas situações, o pesquisador pode estar interessado em avaliar a consistência dos modelos com os dados, fornecendo a mesma probabilidade a priori para cada modelo. Neste caso, a razão de chances a posteriori entre os modelos, torna-se simplesmente a razão das verossimilhanças marginais e recebe a denominação de *Fator de Bayes*, que é definido por,

$$BF_{ij} = \frac{P(y|M_i)}{P(y|M_j)}, \quad (3.34)$$

selecionando-se o modelo que apresenta o maior fator de Bayes.

3.2.3.2 Pseudo-fator de Bayes

A metodologia bayesiana usual para a comparação do desempenho relativo de modelos é a razão de chances a posteriori, que é a probabilidade relativa de dois modelos completamente especificados, ou o fator de Bayes. Entretanto, muitas vezes, como é o caso deste trabalho, o pesquisador deseja utilizar distribuições a priori não-informativas ou impróprias. Neste caso, a utilização da razão de chances a pos-

teriori ou do fator de Bayes apresenta sérios problemas se as priors não-informativas são utilizadas para parâmetros que não são comuns a todos os modelos (ver Koop (2003)). Contudo, a abordagem proposta por Gelfand et al. (1992), denominada *pseudo-fator de Bayes*, é adequada para lidar com estas situações. Este critério foi utilizado por Dellaportas et al. (2003) para selecionar modelos GARCH multivariados.

O pseudo-fator de Bayes é apropriado para este trabalho uma vez que, em geral, notadamente em Econometria, os pesquisadores estão interessados em obter predições. Isto é, dadas as observações, y , o pesquisador deseja prever algum dado não-observado y^* . A lógica bayesiana postula que se deve sumarizar a incerteza sobre o desconhecido (y^*) via uma distribuição condicional. Portanto, os procedimentos bayesianos de previsão baseiam-se na *densidade preditiva* $P(y^*|y)$ (ou, caso existam vários modelos, $P(y^*|y, M_i)$). Como se pode obter uma distribuição marginal a partir da distribuição conjunta associada e aplicando o Teorema da Probabilidade Total a distribuições condicionais, a densidade preditiva pode ser escrita como,

$$P(y^*|y) = \int P(y^*, \theta|y) d\theta, \quad (3.35)$$

$$= \int P(y^*|y, \theta)P(\theta|y) d\theta. \quad (3.36)$$

A equação (3.36), indica que a densidade preditiva pode ser vista como uma média ponderada das densidades preditivas condicionais, $P(y^*|y, \theta)$, com a densidade a posteriori conjunta para θ funcionando como uma função de ponderação. No contexto de séries temporais, a densidade preditiva (3.35) pode ser utilizada para implementar predi-

ções probabilísticas sobre y_{T+1}^* dado y_t .

Para motivar o pseudo-fator de Bayes, suponha-se que \mathbf{y} seja um vetor contendo T observações, θ um vetor n -dimensional dos parâmetros de interesse e y_{T+1} ; um observável desconhecido. Assim, a distribuição de y_{T+1} é a distribuição preditiva a posteriori, sendo dada por,

$$P(y_{T+1}|y) = \int P(y_{T+1}|y, \theta)P(\theta|y)d\theta. \quad (3.37)$$

Deve-se checar as distribuições preditivas $P(y_{T+1}|y)$ contra o verdadeiro y_{T+1} na medida em que, se o modelo é adequado, y_{T+1} pode ser considerada como uma realização de $P(y_{T+1}|y)$.

Para avaliar a acurácia preditiva dos modelos sob consideração, deve-se utilizar as estimativas das densidade preditivas $\hat{P}(y_{T+1}|y)$. Se os métodos MCMC produzem observações da posteriori $P(\theta|y)$, pode-se utilizar o resultado destas simulações, θ_s , ($s = 1, \dots, S$), para efetuar a comparação dos modelos. A estimativa da densidade preditiva um passo-à-frente será dada por,

$$\hat{P}(y_{T+1}|y) = \frac{\sum_{s=1}^S P(y_{T+1}|y, \theta_s)}{S}, \quad (3.38)$$

para estimar-se K densidades preditivas um passo-à-frente, $\hat{P}(y_{T+1}|y)$, utilizam-se os $t = T, T+1, \dots, T+K-1$ dados. Gelfand et al. (1992) propuseram utilizar,

$$\prod_{t=T}^{T+K-1} \hat{P}(y_{t+1}|y), \quad (3.39)$$

como base para a comparação de modelos. Seleciona-se o modelo com maior $\prod_{t=T}^{T+K-1} \hat{P}(y_{t+1}|y)$. Especificamente, para comparar dois mode-

los M_i e M_j , Gelfand et al. (1992) propuseram a seguinte quantidade,

$$D = \log \left(\frac{\prod_{t=T}^{T-1} \hat{P}(y_{t+1} = y_{T+1} | \Psi_t, \theta, M_i)}{\prod_{t=T}^{T-1} \hat{P}(y_{t+1} = y_{t+1} | \Psi_t, \theta, M_2)} \right), \quad (3.40)$$

sendo y_{t+1} uma realização do processo estocástico no tempo $t + 1$; Ψ_t , o conjunto de informação disponível até o tempo t ; e θ , o vetor de médias a posteriori da modelo especificado. Seleciona-se o modelo M_i se $D > 0$ (ou M_j se $D < 0$). Gelfand et al. (1992) denominaram $\exp(D)$ de pseudo-fator de Bayes.

3.2.4 Simplicidade e unicidade da inferência bayesiana

Todos os conceitos teóricos relevantes da inferência bayesiana foram expostos nestas poucas páginas, ou seja, os conceitos para aprender sobre os parâmetros, sobre como avaliar modelos e para fazer previsões. Note-se a unidade e simplicidade teórica da abordagem bayesiana. Conforme Koop (2003), aceitando-se que a incerteza sobre as grandezas desconhecidas (θ , M_i e y^*) seja representada por variáveis aleatórias, o restante da abordagem bayesiana não apresenta controvérsias.

Trata-se simplesmente de aplicar regras simples de probabilidade para fazer inferências. Assim, diante de um novo modelo, basta se lembrar de que a inferência bayesiana exige o estabelecimento da verossimilhança e de uma densidade a priori, informativa ou não informativa. Estas são, então, utilizadas para encontrar a densidade a posteriori dos parâmetros, que é base para todos os procedimentos de inferência sobre os parâmetros do modelo. Caso se considerem vários modelos, o pesquisador pode diagnosticá-los, melhorá-los e/ou compará-los utilizando

as probabilidades a posteriori dos modelos, as razões de chances a posteriori ou o fator de Bayes, no caso de a prioris informativas. No caso de a prioris não-informativas, pode-se utilizar o pseudo-fator de Bayes. As previsões são feitas a partir da densidade preditiva, $p(y^*, y)$. Estas poucas equações são utilizadas para implementar inferência estatística em *qualquer* aplicação.

3.3 Métodos de Monte Carlo via Cadeias de Markov

A teoria dos Métodos de Monte Carlo via Cadeias de Markov (MCMC) é extremamente técnica. Procurou-se nesta seção, apresentar os principais fatos que fundamentam a metodologia. Apresenta-se, de forma mais detalhada, o algoritmo de Metropolis-Hastings. Por fim, descrevem-se os procedimentos utilizados para verificar a convergência do algoritmo de Metropolis-Hastings. Esta seção é basicamente uma resenha do trabalhos de Chib & Greenberg (1995) principalmente, além de Bawens et al. (1999), Geweke (1996) e Geweke (1997)

3.3.1 Métodos de simulação Monte Carlo via Cadeias de Markov

Suponha que um pesquisador deseje utilizar a média a posteriori como um estimador pontual ou esteja interessado em qualquer outra função de θ . Todas estas características associadas às distribuições a

posteriori possuem a forma,

$$E[g(\theta)|y] = \int_{\Theta} g(\theta)P(\theta|y) d\theta \quad (3.41)$$

sendo $g(\cdot)$ uma função de interesse. Com a exceção de poucos casos, estas integrais não podem ser avaliadas analiticamente. Neste caso o pesquisador deve recorrer a técnicas numéricas. As mais utilizadas são as técnicas de Monte Carlo via cadeias de Markov, cuja teoria básica é apresentada nesta seção.

Segundo Chib & Greenberg (1995), a abordagem tradicional da teoria das cadeias de Markov sobre um espaço de estados contínuo inicia-se com um núcleo de transição $P(x, A)$ para $x \in \mathbb{R}^d$ e $A \in \mathbb{B}$, sendo \mathbb{B} uma σ -álgebra de Borel sobre \mathbb{R}^d . O núcleo de transição é uma função de distribuição condicional que representa a probabilidade da cadeia de Markov movimentar-se de x para outro ponto no conjunto A . Como se trata de uma função de distribuição, tem-se que $P(x, \mathbb{R}^d) = 1$, permitindo que a cadeia possa fazer uma transição de x para x , ou seja, $P(x|x)$ não é necessariamente zero.

Conforme Chib & Greenberg (1995), uma questão central na teoria das cadeias de Markov é determinar condições sob as quais existe uma distribuição invariante π^* e condições sob as quais as iterações do núcleo de transição convergem para a distribuição invariante. Uma distribuição invariante deve satisfazer,

$$\pi^*(dy) = \int_{\mathbb{R}^d} P(x, dy)\pi(x) dx, \quad (3.42)$$

sendo π a densidade com relação à medida de Lebesgue de π^* (Assim, $\pi^*(dy) = \pi(y)dy$). A n -ésima iteração da cadeia de Markov é dada

por,

$$P^n(x, A) = \int_{\mathbb{R}^p} P^{n-1}(x, dy)P(y, A), \quad (3.43)$$

sendo $P^1(x, dy) = P(x, dy)$. Em seguida apresentam-se as condições sob as quais se pode mostrar que a n -ésima iteração da cadeia converge para a distribuição invariante, na medida em que $n \rightarrow \infty$.

Os métodos *MCMC* revertem a teoria: a distribuição invariante é conhecida (a menos de uma constante multiplicativa) - a densidade invariante é $\pi(\cdot)$, a densidade objetivo que se deseja amostrar - mas o núcleo de transição é desconhecido. Para gerar amostras de $\pi(\cdot)$, a metodologia busca encontrar um núcleo de transição $P(x, dy)$ cuja n -ésima iteração converge para $\pi(\cdot)$, para n grande. O processo pode ser iniciado em um ponto arbitrário x , e a partir deste ponto pode-se iterar a cadeia um grande número de vezes. Após muitas iterações da cadeia de Markov, a distribuição das amostras geradas por simulação é, aproximadamente, a distribuição objetivo.

Isto posto, o problema é encontrar um núcleo de transição $P(x, dy)$ adequado. A estratégia para encontrar tal núcleo pode ser descrita como a seguir. Suponha-se que o núcleo de transição, para alguma função $p(x, y)$, possa ser representado por,

$$P(x, dy) = p(x, y)(dy) + r(x)\delta_x d(y), \quad (3.44)$$

sendo $p(x, x) = 0$, $\delta_x d(y) = 1$ se $x \in dy$ e 0 caso contrário, e $r(x) = 1 - \int_{\mathbb{R}^d} p(x, y)dy$ é a probabilidade de que a cadeia permaneça em x na próxima iteração. Conforme Chib & Greenberg (1995), da possibilidade que $r(x) \neq 0$ torna-se claro que a integral de $p(x, y)$ em relação a y não é necessariamente 1.

Caso a função $p(x, y)$ em (3.44) satisfizer a condição de reversibilidade dada por,

$$\pi(x)p(x, y) = \pi(y)p(y, x), \quad (3.45)$$

então π é a distribuição invariante de $P(x, \cdot)$. Seguindo Chib & Greenberg (1995), pode-se mostrar isto avaliando o lado direito da equação (3.42),

$$\begin{aligned} \int P(x, A)\pi(x) dx &= \int \left[\int_A p(x, y) dy \right] \pi(x) dx + \int r(x)\delta_x(A)\pi(x) dx, \\ &= \int_A \left[\int p(x, y)\pi(x) dx \right] dy + \int_A r(x)\pi(x) dx, \\ &= \int_A \left[\int p(y, x)\pi(y) dx \right] dy + \int_A r(x)\pi(x) dx, \\ &= \int_A (1 - r(y))\pi(y) dy + \int_A r(x)\pi(x) dx, \\ &= \int_A \pi(y) dy \end{aligned} \quad (3.46)$$

Interpretando informalmente, o lado esquerdo da condição de reversibilidade é a probabilidade incondicional de a cadeia de Markov mover-se de x para y , sendo x gerado a partir de $\pi(\cdot)$; e o lado direito é probabilidade incondicional de movimento da cadeia de y para x , sendo y também gerado de $\pi(\cdot)$. Portanto, a condição de reversibilidade implica que estas probabilidades são iguais, e o resultado mostrado anteriormente mostra que $\pi^*(\cdot)$ é a distribuição invariante de $P(\cdot, \cdot)$. Conforme Tierney (1994), este resultado mostra que a reversibilidade é uma condição suficiente que deve ser satisfeita pelo núcleo de transição $p(x, y)$.

3.3.2 O algoritmo de Metropolis-Hastings

O algoritmo de Metropolis-Hastings é um poderoso método MCMC que pode ser utilizado para amostrar uma distribuição intratável $\pi^*(\cdot)$. Este algoritmo foi proposto por Metropolis et al. (1953) e generalizado por Hastings (1970). É um algoritmo em que, uma vez que a convergência tenha sido obtida, fornece amostras da densidade a posteriori de interesse. Relaciona-se à amostragem por importância, na medida em que utiliza uma aproximação da posteriori para simular amostras aleatórias. Entretanto, há um mecanismo de rejeição (ao invés de uma função de ponderação) para decidir se uma amostra pertence ou não à densidade posteriori de interesse. Pode-se descrever o funcionamento do algoritmo como a seguir.

Suponha-se que possam ser geradas amostras candidatas de uma determinada densidade. Como tais amostras são geradas de cadeias de Markov, a densidade depende do estado atual do processo. Em conformidade com tal fato, denota-se por $q(x, y)$ uma densidade geradora de candidatos, ou densidade instrumental, sendo $\int q(x, y)dy = 1$. Se o processo (cadeia de Markov) está em um determinado ponto x , pode-se gerar uma amostra y da densidade candidata $q(x, y)$. Caso $q(x, y)$ satisfaça a condição de reversibilidade dada em (3.45), para todo x e y , pode-se terminar a busca pela densidade invariante. Entretanto, conforme Chib & Greenberg (1995), muito provavelmente isto não ocorrerá. Como exemplo, pode ocorrer que, a cadeia de Markov se movimente com muita frequência de x para y , e muito raramente de y para x , ou seja,

$$\pi(x)q(x, y) > \pi(y)q(y, x), \quad (3.47)$$

uma maneira de evitar esta possibilidade é reduzindo o número de movimentos de x para y pela introdução de uma probabilidade $\alpha(x, y)$ de o movimento ocorrer. Se o movimento não é realizado, a cadeia de Markov retorna para x , considerando-o um valor da densidade objetivo. Assim, a transição da cadeia de x para y ($y \neq x$) é feita de acordo com,

$$P_{MH}(x, y) = q(x, y)\alpha(x, y) \quad x \neq y, \quad (3.48)$$

em que $\alpha(x, y)$ deve ser calculada. A equação (3.47) implica que movimentos de y para x não acontecem frequentemente. Portanto, deve-se definir $\alpha(y, x)$ tão grande quanto possível, e como $\alpha(y, x)$ é uma probabilidade, seu limite superior é igual a 1. Isto implica que a probabilidade de movimento $\alpha(x, y)$ é determinada pela condição de que $P_{MH}(x, y)$ satisfaça à condição de reversibilidade; assim, tem-se que,

$$\begin{aligned} \pi(x)q(x, y)\alpha(x, y) &= \pi(y)q(y, x)\alpha(y, x), \\ &= \pi(y)q(y, x), \end{aligned} \quad (3.49)$$

portanto,

$$\alpha(x, y) = \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)} \quad (3.50)$$

Assim, conforme Chib & Greenberg (1995), vê-se que as probabilidades $\alpha(x, y)$ e $\alpha(y, x)$ asseguram que P_{MH} satisfaça a condição de reversibilidade. Pode-se concluir então que, para P_{MH} ser reversível, a probabilidade de movimento da cadeia de Markov deve calculada por,

$$\alpha(x, y) = \begin{cases} \min \left\{ \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}, 1 \right\} & \text{se } \pi(x)q(x, y), \\ 1 & \text{caso contrário,} \end{cases} \quad (3.51)$$

Para completar a definição do núcleo de transição da cadeia de Metropolis-Hastings, deve-se considerar a probabilidade, que pode ser positiva, de a cadeia permanecer no ponto x . Pela definição dada anteriormente,

$$r(x) = \int_{\mathbb{R}^d} P(x, dy)\pi(x)dx \quad (3.52)$$

portanto, o núcleo de transição da cadeia de Metropolis-Hastings, denotado por $P_{MH}(x, dy)$, será dado por,

$$P_{MH}(x, dy) = q(x, y)\alpha(x, y)dy + \left[1 - \int_{\mathbb{R}^d} q(x, y)\alpha(x, y)dy\right]\delta_x(dy)$$

que é um caso particular de (3.44). Conforme Chib & Greenberg (1995), uma vez que $P_{MH}(x, dy)$ é reversível por construção, segue-se do argumento da igualdade em (3.46), segundo o qual o núcleo de transição da cadeia de Markov gerada pelo algoritmo tem distribuição invariante dada por $\pi(x)$.

O algoritmo de Metropolis-Hastings fica determinado após ser definida a distribuição geradora de candidatos, duas escolhas comuns são as distribuições normal e t-Student multivariadas. No caso de distribuições candidatas simétricas, como as duas citadas, tem-se um importante caso especial, isto é, $q(x, y) = q(y, x)$ e a probabilidade de movimento reduz-se a $\pi(y)/\pi(x)$; portanto, se $\pi(y) \geq \pi(x)$, a cadeia move-se para y , caso contrário, a cadeia se move para x com probabilidade $\pi(y)/\pi(x)$. Um sumário do algoritmo de Metropolis-Hastings pode ser dado por,

1. Escolhe-se um valor inicial θ^0 para a cadeia de Markov.
2. Gera-se um candidato, θ^* , da densidade geradora de candidatos,

$q(\theta^{s-1}; \theta)$.

3. Calcula-se a probabilidade de aceitação,

$$\alpha(\theta^{s-1}; \theta^*) = \min \left\{ \frac{P(\theta = \theta^* | y) q(\theta^*; \theta = \theta^{s-1})}{P(\theta = \theta^{s-1} | y) q(\theta^{s-1}; \theta = \theta^*)}, 1 \right\} \quad (3.53)$$

4. Faz-se,

$$\theta^{s+1} = \begin{cases} \theta^* & \text{com probabilidade } \alpha(\theta^{s-1}; \theta^*), \\ \theta^{s-1} & \text{com probabilidade } 1 - \alpha(\theta^{s-1}; \theta^*) \end{cases} \quad (3.54)$$

5. Repetem-se os passos 1, 2 e 3 S vezes.

3.3.3 Procedimentos para verificação da convergência da Cadeia de Markov

Além de métodos gráficos, como a monitoração dos traços e dos quantis acumulados das cadeias, será utilizado o procedimento sugerido por Geweke (1992), baseado em métodos de séries temporais, para verificar a convergência das cadeias de Markov geradas pelo algoritmo de Metropolis-Hastings. O procedimento baseia-se em um teste de igualdade das médias da primeira e última parte da cadeia de Markov (usualmente, os primeiros $m_1 = 10\%$ e os últimos $m_2 = 50\%$). Se as amostras são retiradas da distribuição estacionária da cadeia, as duas médias tendem a ser iguais, e sob a hipótese nula de convergência, a estatística de teste possui uma distribuição assintótica normal padronizada. Portanto, trata-se de um teste Z-score padrão: a diferença entre as duas médias amostrais dividida pelo erro-padrão estimado. O erro-padrão é estimado a partir da densidade espectral em zero, levando-se

em consideração portanto, qualquer, autocorrelação presente.

Especificamente, a estatística de teste para o diagnóstico de convergência é dada por,

$$(\bar{\theta}_1 - \bar{\theta}_2)/(\hat{S}_1^\theta(0)/m_1 + \hat{S}_2^\theta(0)/m_2)^{1/2} \rightarrow N(0, 1) \quad (3.55)$$

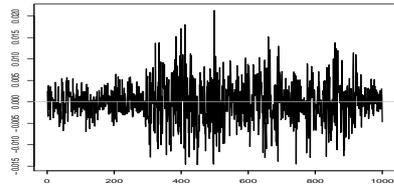
sendo $\hat{S}_1^\theta(\cdot)$ a estimativa da densidade espectral para m_i iterações da cadeia.

4 MATERIAL E MÉTODOS

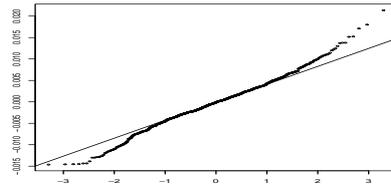
4.1 Dados

4.1.1 Dados simulados

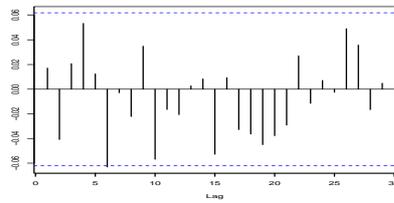
Para testar os algoritmos utilizados para estimar o modelo ARCH com potência assimétrica gaussiano, implementou-se um estudo via simulação. Os detalhes da implementação da série simulada são dados na seção 4.2. A Figura 1 mostra a série de dados simulada. Por inspeção visual, notam-se claramente algumas das regularidades empíricas de séries financeiras, ou seja, agrupamento de períodos de baixa e alta volatilidades, retorno médio muito próximo de zero e dependência desprezível no nível da série (em consonância com a hipótese da eficiência informacional dos mercados), a leptocurtose da distribuição dos dados e a dependência significativa dos retornos quadráticos e absolutos.



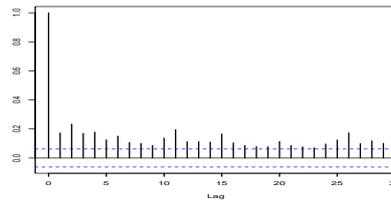
(a) Série simulada



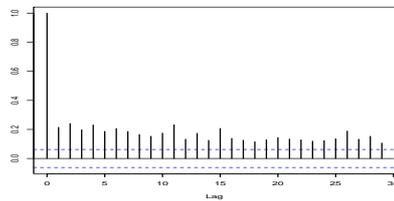
(b) qqplot normal



(c) fac da serie simulada



(d) fac da série quadrática



(e) fac do módulo da série

FIGURA 1: Série, funções de autocorrelação e autocorrelação parcial amostrais - dados simulados

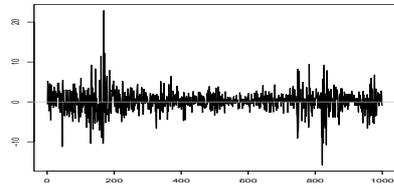
4.1.2 Ibovespa

A abordagem bayesiana proposta para o modelo APARCH será ilustrada utilizando T=1000 retornos diários do índice da Bolsa de Valores do Estado de São Paulo (Ibovespa), no período 1994-1997. Os retornos foram computados da forma usual, ou seja, $y_t = \ln(I_t/I_{t-1})$, para $t = 1, \dots, T$. A Tabela 1 mostra algumas estatísticas básicas da série sob análise. Note-se as características básicas das séries financeiras, como retorno médio muito próximo de zero, leptocurtose e alguma assimetria na distribuição dos retornos.

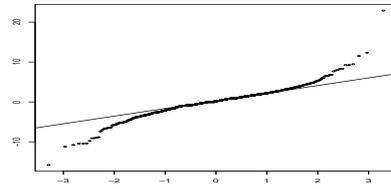
TABELA 1: Ibovespa - Estatísticas

Estatística	
Retorno médio (%)	0,0001
Retorno mediano (%)	0,0002
Curtose	7,6
Assimetria	0,12
Mínimo (%)	-0,16
Máximo (%)	0,23

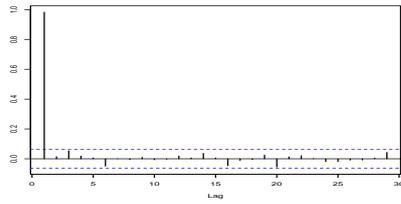
A Figura 2 ilustra a série temporal dos retornos, o gráfico quantil-quantil normal, a função de autocorrelação parcial do nível dos retornos e as funções de autocorrelação dos retornos quadráticos e absolutos. Novamente, vêem-se as regularidades empíricas típicas de séries econômicas e financeiras, dependência muito fraca e com padrão não evidente no nível da série e dependência significativa nos retornos quadráticos e absolutos, além da leptocurtose da distribuição dos retornos mostrada pelo gráfico quantil-quantil normal.



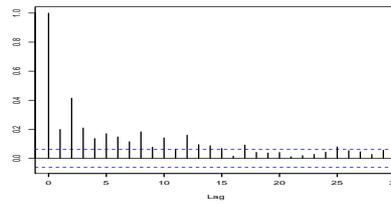
(a) Retornos



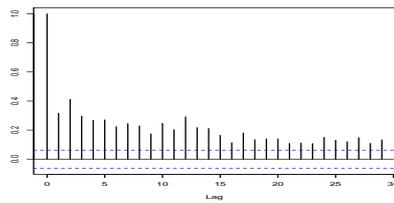
(b) qqplot normal



(c) facp dos retornos do Ibovespa



(d) fac dos retornos quadráticos do Ibovespa



(e) fac dos retornos absolutos

FIGURA 2: Série, funções de autocorrelação e autocorrelação parcial amostrais - Ibovespa

4.2 Metodologia

Todos os programas utilizados para implementar a abordagem bayesiana para os modelos considerados foram elaborados na linguagem R, (R Development Core Team (2005)). Os estimadores de máxima verossimilhança foram obtidos com o pacote `fSeries` (Wuertz (2006)) e com programas da linguagem `Ox` criados por Laurent & Peters (2002). O apêndice C, contem o programa R utilizado para implementar os procedimentos bayesianos para o modelo APARCH considerado.

A série simulada $(y_1, y_2, \dots, y_{1000})$ de um modelo APARCH(1,1) gaussiano foi obtida usando a seguinte especificação:

$$\begin{aligned} \sigma_t^{1,2} &= 0,000014 + 0,083(|y_{t-1}| - 0,373y_{t-1})^{1,2} + 0,92\sigma_{t-1}^{1,2}, \\ y_t &\sim N(0, \sigma_t^\delta), \end{aligned} \quad (4.1)$$

e, como usual na literatura, considerou-se σ_0^δ uma constante conhecida.

A função de verossimilhança para o modelo APARCH(1,1) gaussiano pode ser escrita como,

$$L(\omega, \alpha_1, \gamma_1, \beta_1, \delta|y) = \prod_{i=1}^T \sigma_t^{-\delta/2} \exp\left(-\frac{(y_i)^2}{2\sigma_t^\delta}\right) \quad (4.2)$$

sendo $y = (y_1, \dots, y_T)$

Como informação a priori, utilizaram-se as seguintes a prioris não-

informativas,

$$\begin{aligned}\omega &> 0, \\ \delta &\geq 0, \\ \alpha_1 &\geq 0, \\ \beta_1 &\geq 0, \\ -1 &< \gamma_1 < 1,\end{aligned}$$

que representam as restrições necessárias para assegurar a positividade da variância condicional.

Para estimar o modelo APARCH(1,1) gaussiano propôs-se o algoritmo de Metropolis-Hastings com uma distribuição t-Student multivariada como distribuição candidata e com atualização simultânea de todos os parâmetros. O objetivo é gerar uma sequência simultânea de cadeias de Markov para cada parâmetro, de forma que a distribuição limite das cadeias convirja para a distribuição a posteriori conjunta dos parâmetros. A escolha do esquema de atualização simultânea e da distribuição t-Student multivariada como distribuição geradora de candidatos deu-se por experimentação. O esquema de atualização em blocos e a distribuição normal multivariada como distribuição candidata foram considerados, entretanto, produziram resultados inferiores.

Os procedimentos bayesianos foram implementados como se segue. Seja $\theta = (\omega, \alpha_1, \gamma_1, \beta_1, \delta)$ o vetor de parâmetros e $P(\theta|y)$ a distribuição a posteriori dos parâmetros. Isto posto, o núcleo da distribuição a

posteriori conjunta será formado por,

$$\begin{aligned}
P(\theta|y) \propto & L(\omega, \alpha_1, \gamma_1, \beta_1, \delta|y) \times I_{(\omega>0)} \times I_{(\alpha_1 \geq 0)} \times I_{(\beta_1 \geq 0)} \\
& \times I_{(-1 < \gamma_1 < 1)} \times I_{(\delta \geq 0)}
\end{aligned} \tag{4.3}$$

sendo $I_{(\cdot)}$ a função indicadora.

Em primeiro lugar estimou-se a moda ($\tilde{\theta}$) do núcleo da densidade logarítmica a posteriori, para isto utilizou-se um algoritmo do tipo Newton implementado na função `nlminb` da linguagem R. Como sub-produto da estimação da moda, obteve-se uma aproximação numérica do hessiano H , sendo que o inverso do hessiano H^{-1} foi utilizado para obter-se uma aproximação da matriz de variâncias e covariâncias dos parâmetros. .

Construiu-se a distribuição candidata da seguinte forma: seja $K(\cdot; \mu, V, \nu)$ o núcleo de uma distribuição t-Student multivariada com vetor de parâmetros de locação μ , matriz de escala V e com ν graus de liberdade. Isto posto, considerou-se,

$$\mu = \tilde{\theta}, \tag{4.4}$$

$$V = -(c)H^{-1}, \tag{4.5}$$

$$\nu = 6 \tag{4.6}$$

sendo c uma constante utilizada para ajustar a razão de aceitação do algoritmo de Metropolis-Hastings e V uma aproximação da matriz de variâncias e covariâncias dos parâmetros da distribuição a posteriori conjunta. Assim, o núcleo de transição do algoritmo de Metropolis-

Hastings fica dado por,

$$\alpha(\theta^{s-1}; \theta^*) = \min \left\{ \frac{P(\theta = \theta^* | y) / K(\theta^*; \mu, V, \nu)}{P(\theta = \theta^{s-1} | y) / K(\theta^{s-1}; \mu, V, \nu)}, 1 \right\} \quad (4.7)$$

portanto, com probabilidade $\alpha(\theta^{s-1}; \theta^*)$, $\theta = \theta^*$ e com probabilidade $1 - \alpha(\theta^{s-1}; \theta^*)$, $\theta = \theta^{s-1}$.

As amostras resultantes do algoritmo foram obtidas da seguinte forma. Seguindo Dellaportas et al. (2000), em primeiro lugar, obteve-se uma grande amostra da cadeia (400.000 iterações) e uma parte desta cadeia foi descartada (200.000 iterações) após inspeção visual do traço e dos quantis acumulados da cadeia de cada parâmetro. Analisou-se, em seguida, a função de autocorrelação amostral da cadeia de cada parâmetro. Isto posto, escolheu-se o intervalo (*thin*=20) com o qual seriam tomadas amostras, de forma a se obter uma amostra aproximadamente não-correlacionada da cadeia de cada parâmetro, totalizando-se, portanto, uma amostra de 10000 candidatos aceitos. E, por fim, testou-se a convergência de cada cadeia implementando o teste Z-score de Geweke (1992) e usando métodos gráficos de verificação da convergência das cadeias.

Computados os resultados a posteriori dos parâmetros, o cálculo das esperanças a posteriori das variâncias condicionais é relativamente simples, uma vez que σ_t^δ é função somente dos parâmetros e dos dados. Para o modelo APARCH(1,1), tem-se que,

$$\begin{aligned} E(\sigma_t^\delta | y_t) &= E(\omega | y_t) + E(\alpha_1 | y_t)(|y_{t-1}| - E(\gamma_1 | y_t)y_{t-1})^{E(\delta | y_t)} \\ &\quad + E(\beta_1 | y_t)\sigma_{t-1}^{E(\delta | y_t)}, \end{aligned} \quad (4.8)$$

Ding et al. (1993) demonstraram que se a distribuição de y_t é condicionalmente normal, então a condição para a estacionariedade fraca do modelo APARCH(1,1), isto é, a condição para a existência de $E(\sigma_t^\delta)$ e de $E(|\sigma_t^\delta|)$, é dada por,

$$f(\alpha_1, \gamma_1, \beta_1, \delta) = \frac{1}{\sqrt{2\pi}} \alpha_1 \{(1 + \gamma_1)^\delta + (1 - \gamma_1)^\delta\} 2^{\frac{\delta-1}{2}} \Gamma\left(\frac{\delta+1}{2}\right) + \beta_1,$$

$$f(\alpha_1, \gamma_1, \beta_1, \delta) < 1 \quad (4.9)$$

se esta condição é satisfeita, a esperança incondicional de σ_t^δ (a variância incondicional do processo APARCH(1,1)) será,

$$E(\sigma_t^\delta) = \frac{\omega}{1 - \left(\frac{1}{\sqrt{2\pi}} \alpha_1 \{(1 + \gamma_1)^\delta + (1 - \gamma_1)^\delta\} 2^{\frac{\delta-1}{2}} \Gamma\left(\frac{\delta+1}{2}\right) + \beta_1\right)} \quad (4.10)$$

Para verificar a condição (4.11) de que o processo APARCH(1,1) gaussiano possui ou não variância incondicional finita na abordagem bayesiana, simplesmente estima-se a probabilidade a posteriori dada por,

$$P(f(\alpha_1, \gamma_1, \beta_1, \delta) < 1) \approx \frac{1}{S} \sum_{s=1}^S I\left(\frac{1}{\sqrt{2\pi}} \alpha_1^s \{(1 + \gamma_1^s)^\delta (1 - \gamma_1^s)^\delta\} 2^{\frac{\delta^s-1}{2}} \times \Gamma\left(\frac{\delta^s+1}{2}\right) + \beta_1^s\right),$$

sendo $I(\cdot)$ a função indicadora.

Uma vez que fornecer previsões é um dos principais papéis de modelos de séries temporais em geral, notadamente dos modelos de heterocedasticidade condicional, propôs-se o seguinte algoritmo para a obtenção das densidades preditivas um-passo-à-frente para o modelo

APARCH(1,1) gaussiano, lembrando que a densidade preditiva de y_{T+1} é dada por,

$$P(y_{T+1}|y) = \int_{\Theta} P(y_{T+1}|y, \theta)P(\theta|y) d\theta \quad (4.11)$$

tem-se,

1. Implementa-se o algoritmo de Metropolis-Hastings e obtém-se uma amostra θ^s ($s = 1, \dots, S$) da distribuição conjunta a posteriori $P(\theta|y)$
2. Geram-se amostras de tamanho S das variâncias condicionais σ_{t+1}^δ . Para isto, utilizam-se as amostras de θ^s ($s = 1, \dots, S$)
3. Para cada $\sigma_{t+1}^\delta(s)$, calcula-se $\hat{P}(y_{T+1}|\theta^s, y)$.
4. Isto posto, estima-se a densidade preditiva por,

$$\hat{P}(y_{T+1}|y) = \frac{1}{S} \sum_{s=1}^S \hat{P}(y_{T+1}|\theta^s, y) \quad (4.12)$$

Para calcular o Pseudo-Fator de Bayes, tanto para os dados simulados quanto para a série de retornos do Ibovespa, ambas com $T = 1000$, utilizaram-se as 970 primeiras observações para estimar os momentos a posteriori, em seguida aplicou-se o algoritmo descrito para obter as log-densidades preditivas um-passo-à-frente para as 30 observações restantes. A Figura 6 ilustra as log-densidades preditivas um-passo-à-frente avaliadas nas 30 observações fora da amostra geradas pelo modelo APARCH(1,1).

5 RESULTADOS E DISCUSSÃO

5.1 Série Simulada

Os resultados obtidos para a série simulada encontram-se na Tabela 2. As estimativas dos erros de Monte Carlo são bastante razoáveis. O valor verdadeiro de cada parâmetro está contido no intervalo com 95% de credibilidade com máxima densidade a posteriori (HPD). As estimativas pontuais das médias e medianas das distribuições marginais a posteriori são adequadas. Além disso, a razão de aceitação total foi de aproximadamente 50%, valor sustentado na literatura como sendo adequado (Gelman et al. (2003)). Como referência para a análise do algoritmo, reportam-se também as estimativas de máxima verossimilhança.

As Figuras 6 e 7, do apêndice A, ilustram o traço da cadeia de Markov de cada parâmetro e os respectivos histogramas das amostras das densidades marginais a posteriori. Por inspeção visual, os traços de todos os parâmetros não revelam nenhum sinal evidente de não-convergência.

Observando-se os histogramas das marginais a posteriori, vê-se os desvios da normalidade devidos às assimetrias das densidades marginais a posteriori. Note-se a acentuada assimetria à direita do parâmetro ω , causada pela restrição de positividade modelada na priori. A

TABELA 2: Resultados da estimação - modelo APARCH(1,1) gaussiano - Série simulada

			sumário da posteriori		
	θ	$\hat{\theta}$	média	$Q_{50\%}$	z -score
ω	0,000014	0.006	0,00356	0,00282	0,60125
$d.p$			0,00305		
e_{MC}			$3,30 \times 10^{-5}$		
$HPD_{95\%}$			$[6,81 \times 10^{-7}; 0,00094]$		
α_1	0,083	0.094	0,09466	0,09386	0,20305
$d.p$			0,03176		
e_{MC}			$3,05 \times 10^{-4}$		
$HPD_{95\%}$			$[0,0288; 0,1573]$		
γ_1	0,373	0,491	0,41426	0,414486	0,14971
$d.p$			0,16701627		
e_{MC}			$1,92 \times 10^{-3}$		
$HPD_{95\%}$			$[0,082; 0,7500]$		
β_1	0,92	0,908	0,90243	0,902893	0,26152
$d.p$			0,02728		
e_{MC}			$2,17 \times 10^{-4}$		
$HPD_{95\%}$			$[0,849; 0,9592]$		
δ	1,2	1,110	1,28674	1,28721	0,48563
$d.p$			0,22551		
e_{MC}			$2,19 \times 10^{-3}$		
$HPD_{95\%}$			$[0,8355; 1,7325]$		
razão			0,49		

Notas: θ representa o verdadeiro valor do parâmetro. $\hat{\theta}$ denota a estimativa de máxima verossimilhança dos respectivos parâmetros. $d.p$ é o desvio-padrão da posteriori. e_{MC} é a estimativa do erro de Monte Carlo. $HPD_{95\%}$ é o intervalo com 95% de credibilidade de máxima densidade a posteriori. z -score denota o p-valor do teste de Geweke. $Q_{50\%}$ é a mediana da posteriori. $razão$ denota a razão de aceitação do algoritmo de Metropolis-Hastings.

distribuição marginal a posteriori de α_1 também apresenta assimetria à direita, enquanto a distribuição marginal a posteriori de γ_1 mostra assimetria à esquerda. A distribuição marginal a posteriori de δ é aproximadamente simétrica, apresentando uma pequena assimetria à esquerda.

A assimetria das marginais a posteriori dos parâmetros do modelo APARCH(1,1) gaussiano é a principal causa da diferença entre as estimativas de máxima verossimilhança e os momentos a posteriori. Como sustentado por Bawens e Lubrano (1998), pode-se concluir que são três os efeitos da assimetria: em primeiro lugar o estimador de máxima verossimilhança pode ser maior (no caso de assimetria a direita) ou menor (no caso de assimetria a esquerda) que a média a posteriori, e o erro-padrão do estimador de máxima verossimilhança tende a ser maior que o desvio-padrão a posteriori, uma vez que o cálculo do estimador de máxima verossimilhança não leva as restrições em consideração.

Uma terceira conclusão que se deve destacar é que a assimetria noticiada das marginais a posteriori implica em dúvidas sobre a adequabilidade de densidades a priori normais, as quais foram utilizadas por Nakatsuma (2000) para estimar o modelo GARCH. Uma possível justificativa para se utilizarem densidades a priori gaussianas é um argumento assintótico frequentista na medida em que, teoricamente, os estimadores de máxima verossimilhança são assintoticamente normais. Entretanto, as evidências reportadas neste trabalho, e em conformidade com os resultados de Bawens e Lubrano (1998) e Dellaportas et al. (2000), mostram que a utilização de densidades a priori normais pode ser inconsistente com as marginais a posteriori obtidas na literatura.

A Figura 3 mostra os resultados de $E(\sigma_t^\delta | y_t)$ para os dados simulados. Comparando a Figura 5 com a série simulada representada na Figura 1, vê-se que o modelo APARCH(1,1) modelou adequadamente o processo das volatilidades da série simulada.

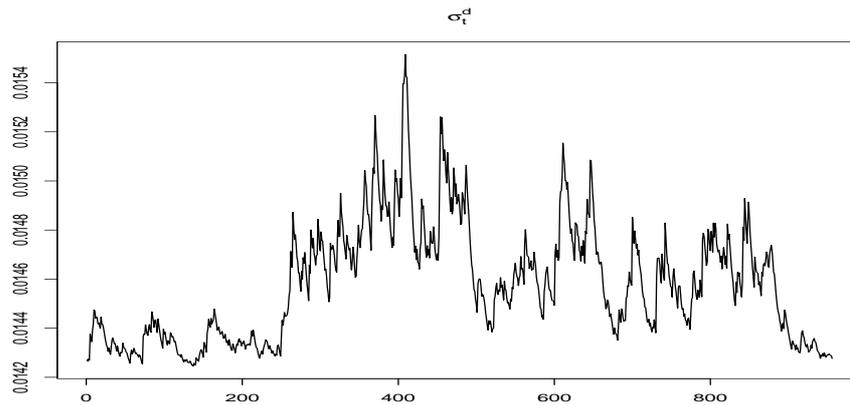


Figura 3: Esperanças a posteriori de σ_t^δ

A Figura 8, do apêndice A, ilustra a evolução dos quantis acumulados das cadeias de todos os parâmetros. Observando as figuras, percebe-se que todos os quantis estabilizaram-se, sem qualquer sinal aparente de não-convergência, fato constatado pelo teste $z - score$ de Geweke reportado na Tabela 2, na qual se pode observar que os p-valores do teste para as cadeias de cada parâmetro não acusam nenhuma evidência de não convergência.

5.2 Série de retornos do Ibovespa

A Tabela 3 contém os resultados da estimação do modelo APARCH (1,1) gaussiano para os retornos diários do Ibovespa.

TABELA 3: Resultados da estimação - modelo APARCH(1,1) gaussiano - retornos do Ibovespa

	$\hat{\theta}$	sumário da posteriori		z-score
		média	$Q_{50\%}$	
ω	3,756	$1,54 \times 10^{-5}$	$1,4 \times 10^{-5}$	0,72
d.p		$9,2 \times 10^{-6}$		
e_{mc}		$8,39 \times 10^{-8}$		
HPD _{95%}		$[1,34 \times 10^{-7}; 3,2 \times 10^{-5}]$		
α_1	0,160	0,036	0,031	0,41
d.p		0,02		
e_{mc}		$2,9 \times 10^{-4}$		
HPD _{95%}		$[3,2 \times 10^{-4}; 0,07]$		
γ_1	0,489	0,52	0,51	0,30
d.p		0,139		
e_{mc}		$1,3 \times 10^{-3}$		
HPD _{95%}		[0,24;0,79]		
β_1	0,795	0,81	0,88	0,22
d.p		0,034		
e_{mc}		$3,4 \times 10^{-4}$		
HPD _{95%}		[0,74;0,87]		
δ	1,393	1,06	1,08	0,47
d.p		0,19		
e_{mc}		$2,21 \times 10^{-3}$		
HPD _{95%}		[0,68;1,40]		
razão		0,45		
$P[E(\sigma_t^\delta) < \infty)$		0,9846		

Notas: θ representa o verdadeiro valor do parâmetro. $\hat{\theta}$ denota a estimativa de máxima verossimilhança dos respectivos parâmetros. *d.p* é o desvio-padrão da posteriori. e_{MC} é a estimativa do erro de Monte Carlo. *HPD*_{95%} é o intervalo com 95% de credibilidade de máxima densidade a posteriori. *z-score* denota o p-valor do teste de Geweke. $Q_{50\%}$ é a mediana da posteriori. *razão* denota a razão de aceitação do algoritmo de Metropolis-Hastings.

As estimativas dos erros de Monte Carlo estão em níveis adequados e a razão de aceitação do algoritmo de Metropolis-Hastings foi de aproximadamente 45%, nível adequado para a dimensão do vetor de parâmetros. A média a posteriori do parâmetro γ_1 foi de 0.52, resultado que sugere a presença do efeito alavancagem no índice Ibovespa, isto é, choques negativos sobre o índice possuem impacto maior sobre a volatilidade do índice que choques positivos.

Observando a Tabela 3, vê-se que para o modelo APARCH(1,1) ajustado para os retornos do índice Ibovespa, a estimativa foi $P(f(\alpha_1, \gamma_1, \beta_1, \delta) < 1) = 0.9846$, indicando que o processo APARCH(1,1) possui variância incondicional finita, o que é consistente com a evidência empírica.

As Figuras 9 e 10 do apêndice B ilustram o traço da cadeia de Markov de cada parâmetro e os respectivos histogramas das amostras das densidades marginais a posteriori dos parâmetros do modelo APARCH(1,1) estimado para os retornos do Ibovespa. Confirmando os resultados do estudo de simulação reportados na seção anterior, as densidades marginais a posteriori dos parâmetros apresentam acentuado grau de assimetria e, portanto, evidentes desvios da hipótese de normalidade. A marginal a posteriori do parâmetro ω apresenta assimetria à direita, o mesmo acontecendo com a marginal a posteriori do parâmetro α_1 . Apesar de não tão acentuada, a marginal a posteriori do parâmetro γ é assimétrica à esquerda. A marginal a posteriori do parâmetro β_1 apresenta uma leve assimetria à esquerda e é leptocúrtica. Enquanto a marginal a posteriori do parâmetro δ é claramente assimétrica à esquerda. A assimetria pronunciada das marginais a posteriori deve explicar a diferença entre as estimativas de máxima verossimi-

lhança e as médias e medianas a posteriori.

A inspeção visual dos traços das cadeias de Markov de cada parâmetro não evidenciam nenhum sinal de não-convergência, fato confirmado pelos p-valores do teste Z-score de Geweke reportados na Tabela 3. A Figura 11, apêndice B, ilustra a evolução dos quantis acumulados das cadeias de todos os parâmetros do modelo APARCH(1,1) ajustado. Analisando os gráficos, percebe-se que todos os quantis estabilizaram-se, sem qualquer sinal aparente de não-convergência.

A Figura 4 ilustra os cálculos das esperanças a posteriori de σ_t^δ ($E(\sigma_t^\delta | y_t)$) fornecidas pelo modelo APARCH(1,1) ajustado para o índice Ibovespa. Comparando a Figura 11 com a série temporal do índice representada na Figura 2, vê-se que o modelo APARCH(1,1) modelou adequadamente o processo de volatilidade da série simulada. Analisando a Figura 4, identificam-se os períodos de alta volatilidade no início da série, em 1994 especificamente, período no qual a volatilidade acentuou-se após implantação do plano Real e a crise do México; em seguida o período é de relativa calma, com o câmbio fixo artificial praticado pela política monetária. No final da série temporal, a crise da Ásia, em 1997, provocou o aumento da volatilidade do índice observado.

Para se obterem as densidades preditivas, utilizou-se o algoritmo descrito na seção anterior e aplicado à série simulada. Para ilustrar o algoritmo com dados reais, fez-se o mesmo procedimento para os retornos diários do Ibovespa, ou seja, com as 970 primeiras observações da série de retornos do Ibovespa estimaram-se os momentos a posteriori e, em seguida, aplicou-se o algoritmo para obter as log-densidades preditivas

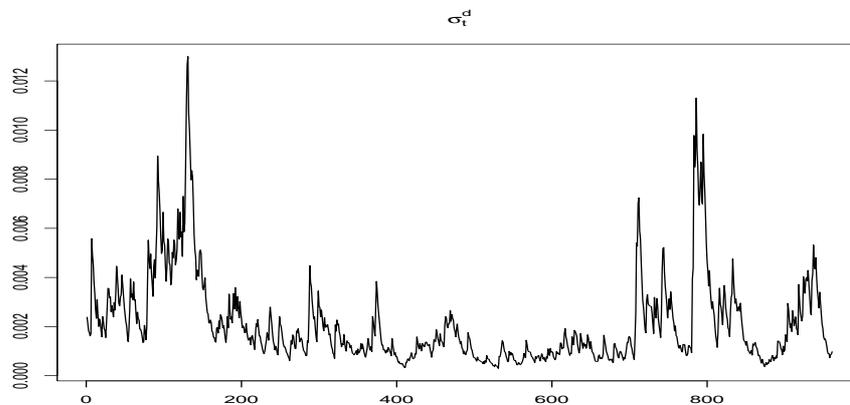


Figura 4: Esperanças a posteriori de σ_t^δ - Ibovespa

um-passo-à-frente para as 30 observações restantes. A Figura 12 ilustra as ordenadas das log-densidades preditivas um-passo-à-frente avaliadas nas 30 observações fora da amostra geradas pelo modelo APARCH(1,1) gaussiano.

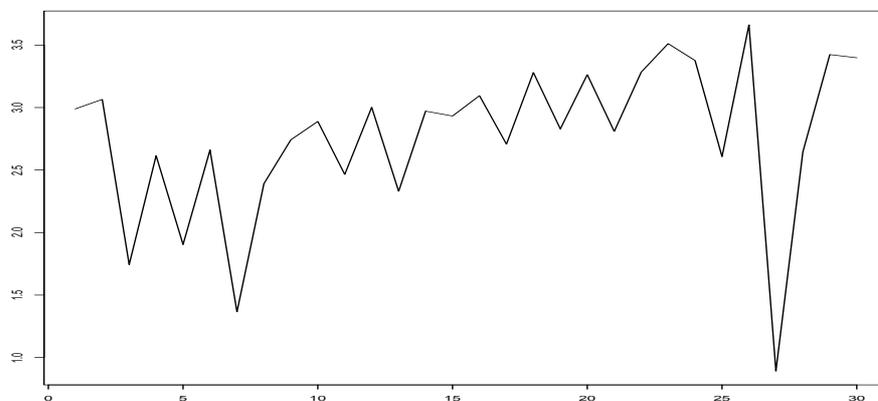


Figura 5: Log-densidades preditivas um-passo-à-frente para 30 observações - Ibovespa.

Para verificar o desempenho preditivo relativo do modelo APARCH (1,1) estimou-se um modelo GARCH(1,1) gaussiano para os dados de retorno do Ibovespa e efetuou-se o mesmo procedimento descrito anteriormente para o modelo APARCH(1,1). Utilizando o pseudo-Fator de Bayes e fazendo $M_i = \text{APARCH}(1,1)$ e $M_j = \text{GARCH}(1,1)$, obteve-se,

$$D = \left(\frac{\sum_{t=970}^{1000} \log \hat{P}(Y_{t+1} = y_{t+1} | \Psi_t, \theta, M_i)}{\sum_{t=970}^{1000} \log \hat{P}(Y_{t+1} = y_{t+1} | \Psi_t, \theta, M_j)} \right) = 77,00306 \quad (5.1)$$

$D = 77,00306 > 0$ implica que, para os dados e o período analisados, o modelo APARCH(1,1) possui desempenho, sobretudo preditivo, superior a um modelo GARCH(1,1), que por sinal é um sub-modelo encaixado no modelo APARCH(1,1). Neste caso, a complexidade adicional do modelo parece ter sido compensada por um desempenho preditivo superior.

6 CONCLUSÕES

A contribuição deste trabalho foi desenvolver uma abordagem bayesiana do modelo GARCH com potência assimétrica usando algoritmos de Monte Carlo via Cadeias de Markov. O modelo generaliza consideravelmente a classe de modelos ARCH, na medida em que possui sete modelos como sub-casos. Desenvolveram-se procedimentos para estimar os parâmetros, obter previsões da volatilidade e selecionar modelos via estimativas da densidade preditiva.

A abordagem bayesiana proposta teve um desempenho adequado. Para a série simulada com $T = 1000$ observações, obtiveram-se boas estimativas dos parâmetros, sendo que algumas densidades marginais a posteriori mostraram-se assimétricas.

Para a série de retornos do Ibovespa, as densidades marginais a posteriori também mostraram-se assimétricas. O pseudo-fator de Bayes, mostrou que a complexidade adicional do modelo APARCH resultou em um ganho em termos de acurácia preditiva em relação a um modelo GARCH(1,1). A condição de estacionariedade não foi imposta a priori, foi estimada pela probabilidade a posteriori da condição, mostrando que o modelo APARCH(1,1) gaussiano é fracamente estacionário.

Dada a assimetria reportada para algumas marginais a posteriori e, considerando que o Teorema de Bayes constitui um processo de atualização e aprendizagem, a utilização de densidades a priori assimétricas

como a distribuição gama entre outras, em trabalhos futuros poderia resultar em ganhos nos procedimentos de inferência.

Pode-se pensar em diversas extensões do presente trabalho, tais como criar uma estratégia bayesiana para o modelo ARCH com potência assimétrica fracionalmente integrado e testar o desempenho de outras técnicas numéricas como amostragem por importância e o amostrador Griddy-Gibbs.

Entretanto, com a enorme variedade de modelos com variâncias tempo-variantes, penso que o desenvolvimento de uma abordagem bayesiana que envolva a modelagem explícita da incerteza em relação aos modelos seria uma grande contribuição. Refiro-me à metodologia denominada “*Bayesian Model Averaging*”, a partir da qual desenvolveu-se uma nova técnica para simulação de distribuições a posteriori denominada *Composição de Modelos com Métodos de Monte Carlo via Cadeias de Markov* (MC^3) (Markov Chain Monte Carlo Model Composition). Em outras palavras, nesta abordagem bayesiana, devem ser obtidos resultados para todos os modelos em consideração e tomadas “médias” destes resultados. Os pesos no processo de “tomar médias” são dados pelas probabilidades a posteriori dos modelos.

REFERÊNCIAS

- BAWENS, L.; LUBRANO, M. Bayesian inference on garch models using the gibbs sampler. **The Econometrics Journal**, Oxford, v. 1, p. 23–46, 1998.
- BAWENS, L.; LUBRANO, M.; RICHARD, J.-F. **Bayesian inference in dynamic econometric models**. Oxford: Oxford University Press, 1999, 376p.
- BERA, A. K.; HIGGINS, M. L. A class of nonlinear ARCH models. **International Economic Review**, Philadelphia, v. 33, n. 1, p. 137–158, Feb. 1992.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, Lousanne, v. 31, n. 3, p. 307–327, Apr. 1986.
- BOLLERSLEV, T.; ENGLE, R. F.; NELSON, D. B. ARCH Models. In: ENGLE, R. F.; MACFADDEN, D. L. (Ed.). **Handbook of Econometrics**. Amsterdan: North-Holland, 1994. v. 4, cap. 49, p. 2959–3038.
- BOYER, M.; KIHSTROM, R. (Ed.). **Bayesian models in economic theory**. Amsterdan: North-Holland, 1984, 317p
- CARTER, C. K.; KOHN, R. On gibbs sampling for state space models. **Biometrika**, London, v. 81, n. 3, p. 541–553, Sept. 1994.
- CHIB, S. Calculating posterior distributions and modal estimates in markov mixture models. **Journal of Econometrics**, Lousanne, v. 75, n. 1, p. 79–97, Nov. 1996.

- CHIB, S.; GREENBERG, E. Bayesian inference in regression models with arma(p,q) errors. **Journal of Econometrics**, Lousanne, v. 49, n. 1/2 p. 183–206, Sept./Oct. 1994.
- CHIB, S.; GREENBERG, E. Understanding the metropolis-hastings algorithm. **The American Statistician**, Washington, v. 49, n. 4, 1995.
- DELLAPORTAS, P.; GIAKOUMATOS, S. G.; POLITIS, D. N. Bayesian analysis of the unobserved arch model. **Statistics and Computing**, London, v. 15, n.2, p. 103–111, Apr. 2005.
- DELLAPORTAS, P.; POLITIS, D. N.; VRONTOS, I. D. Full bayesian inference for garch and egarch models. **Journal of Business and Economic Statistics**, Washington, v. 18, n. 2, p. 187–198, 2000.
- DELLAPORTAS, P.; POLITIS, D. N.; VRONTOS, I. D. Inference for some multivariate arch and garch models. **Journal of Forecasting**, Chichester, v. 22, p. 427–446, 2003.
- DING, Z.; ENGLE, R. F.; GRANGER, C. W. J. A long memory property of stock market returns and a new model. **Journal of Empirical Finance**, Amsterdam, v. 1, p. 83–106, 1993.
- ENGLE, R. F. Autoregressive Conditional Heteroskedasticity with Estimates of the United Kingdom Inflation. **Econometrica**, Avon, v. 50, n. 4, p. 987–1008, 1982.
- FRIEDMAN, M.; SAVAGE, L. J. The utility analysis of choices involving risk. **Journal of Political Economy**, Chicago, v. 56, n. 4, 1948.
- FRIEDMAN, M.; SAVAGE, L. J. The expected-utility hypothesis and the measurability of utility. **Journal of Political Economy**, Chicago, v. 60, n. 3, 1952.
- GELFAND, A. E.; DEY, D. K.; CHANG, H. Model determination using predictive distributions with implementation via sampling-based methods. In: BERNARDO, J. M. et al. (Ed.). **Bayesian statistics**. Oxford: Oxford University Press, 1992. p. 169–193.

- CARLIN, J. B.; GELMAN, A.; RUBIN, D. B.; STERN, H. S. **Bayesian data analysis**. 2 ed. Oxford, England: Chapman Hall CRC Press, 2003, 696p.
- GEWEKE, J. Exact predictive densities for linear models with arch disturbances. **Journal of Econometrics**, Lousanne, v. 40, n. 1, p. 63–86, Jan. 1989.
- GEWEKE, J. Evaluating the accuracy of sampling based approaches to the calculation of posterior moments. In: BERNARDO, J. M. et al. (Ed.). **Bayesian statistics**. Oxford: Oxford University Press, 1992. p. 169–193.
- GEWEKE, J. Bayesian comparison of econometric models. **Working Paper 532, Federal Reserve Bank of Minneapolis**, 1994.
- GEWEKE, J. Monte carlo simulation and numerical integration. In: AMMAN, H. M.; KENDRICK, D. A.; RUST, J. (Ed.). **Handbook of Computational Economics**. Amsterdam: North-Holland, 1996. v. 1, cap. 15, p. 731–800.
- GEWEKE, J. Posterior simulators in econometrics. In: KREPS, D.; WALLIS, K. (Ed.). **Advances in Economics and Econometrics: Theory and Applications**. Cambridge: Cambridge University Press, 1997. III, cap. 15, p. 731–800.
- GEWEKE, J. **Contemporary bayesian econometrics and statistics**. New York: John Wiley and Sons, 2005. 300p.
- GIOT, P.; LAURENT, S. Value-at-risk for long and short trading positions. **Journal of Applied Econometrics**, Chichester, v. 18, n. 6, p. 641–663, Nov./Dec. 2003.
- GLOSTEN, L. R.; JAGANNATHAN, R.; RUNKLE, D. On the relation between the expected value and volatility of the nominal excess return on stocks. **Journal of Finance**, New York, v. 48, n. 5, p. 1779–1801, Dec. 1993.
- GOLDMAN, E.; TSURUMI, H. Bayesian analysis of a doubly truncated arma-garch model. **Studies in Nonlinear Dynamics and Econometrics**, Berkeley, v. 9, n. 2, p. 1–36, May. 2005.

- HAMILTON, J. D.; SUSMEL, R. Autoregressive conditional heteroscedasticity and changes in regime. **Journal of Econometrics**, Lousanne, v. 64, n. 1/2, p. 307–333, Sept./Oct. 1994.
- HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. **Biometrika**, London, v. 57, n. 1, p. 97–109, Mar. 1970.
- KAUFMANN, S.; FRUHWIRTH-SCHNATTER, S. Bayesian analysis of switching arch models. **Journal of Time Series Analysis**, Clevedon, v. 23, n. 4, p. 425–458, July. 2002.
- KEYNES, J. M. **A Treatise on Probability**. London: John Wiley and Sons, 1921. 480p.
- KLEINBERGEN, F.; DIJK, H. K. V. Non-stationarity in garch models: A bayesian analysis. **Journal of Applied Econometrics**, Chichester, v. 8, n. 1, p. 41–61, 1993.
- KOOP, G. **Bayesian Econometrics**. Chichester: John Wiley and Sons, 2003. 374p
- LANCASTER, T. **An Introduction to modern bayesian econometrics**. New York: Blackwell Publishers, 2004. 401p.
- LAURENT, S.; PETERS, J. P. GARCH 2.3: An Ox package for estimating and forecasting various ARCH Models. **Journal of Economic Surveys**, Clevedon, v. 16, n. 3, p. 447–485, July. 2002.
- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equations of state calculations by fast computing machines. **Journal of Chemical Physics**, New York, v. 21, p. 1087–1091, 1953.
- MIGON, H. S.; MAZUCHELLI, J. Modelos garch bayesianos: métodos aproximados e aplicações. **The Brazilian Review Of Econometrics**, Rio de Janeiro, v. 19, n. 1, p. 111–138, 1999.
- MORGENSTERN, O.; NEUMANN, J. V. **Theory of games and economic behavior**. Princeton: Princeton University Press, 1944. 648p.

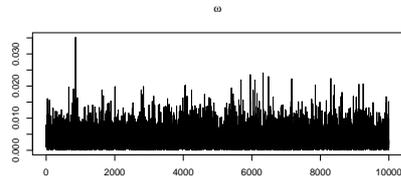
- MULLER, P.; POLE, A. Monte carlo posterior integration in garch models. **Sankhya**, Calcutta, v. 60, n 1, p. 127–144, 1998.
- NAKATSUMA, T. Bayesian analysis of arma-garch models: A markov chain sampling approach. **Journal of Econometrics**, Lousanne, v. 95, n. 1, p. 57–69, Mar. 2000.
- POIRIER, D. J. **Intermediate statistics and econometrics: a comparative approach**. Princeton: MIT Press, 1995. 731p.
- R Development Core Team. **R: A language and environment for statistical computing**. Vienna, Austria, 2005.
- RENGIFO, E.; ROMBOUTS, J. V. K. Dynamic optimal portfolio selection in a var framework. **CORE Discussion Paper**, Louvain, n. 57, p. 447–485, July. 2004.
- RITTER, C.; TANNER, M. A. Facilitating the gibbs sampler: the gibbs stopper and the griddy-gibbs sampler. **Journal of the American Statistical Association**, Washington, v. 87, n. 4/9, p. 861–868, Sept. 1992.
- ROSSI, P.; ALLENBY, G.; MCCULLOCH, R. **Bayesian statistics and marketing**. New York: John Wiley and Sons, 2006. 368p.
- TAYLOR, S. **Modelling financial time series**. New York: John Wiley and Sons, 1986. 284p.
- TIERNEY, L. Markov chains for exploring posterior distributions. **Annals of Statistics**, Hayward, v. 22, n. 4, p. 1701–1762, Dec. 1994.
- TSE, Y. K.; TSUI, A. K. C. Conditional volatility in foreign exchange rates: evidence form the malaysian ringgit and Singapore dollar. **Pacific-Basin Financial Journal**, v. 5, p. 345–356, 1997.
- WUERTZ, D. **fSeries: Rmetrics - The Dynamical Process Behind Markets**. 2006. R package version 221.10065.
- ZAKOIAN, J. M. Threshold heteroskedasticity models. **Journal of Economic Dynamics and Control**, Amsterdam, v. 18, n. 5, p. 931–955, 1994.

ZELLNER, A. **An introduction to bayesian inference in econometrics**. New York: John Wiley and Sons, 1971. 448p.

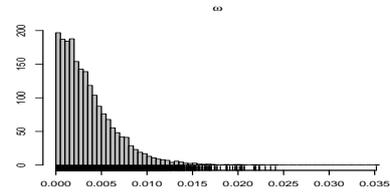
ZELLNER, A. Bayesian econometrics. **Econometrica**, Avon, v. 53, n. 2, p. 253–269, Mar./Apr. 1985.

ZELLNER, A. The current state of bayesian econometrics. In: ZELLNER, A. (Ed.). **Basic issues in econometrics**. Chicago: Chicago University Press, 1997. p. 306–321.

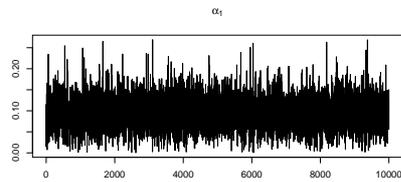
**APÊNDICE A – FIGURAS - DADOS
SIMULADOS**



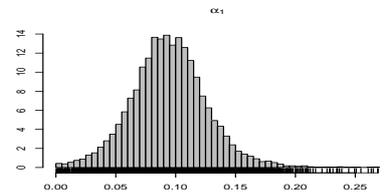
(a) Traço da cadeia de ω



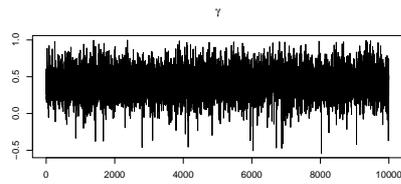
(b) Marginal a posteriori de ω



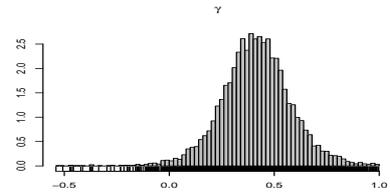
(c) Traço da cadeia de α_1



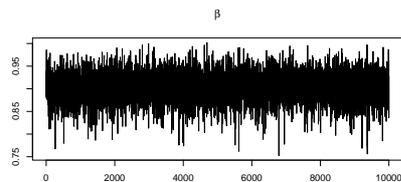
(d) Marginal a posteriori de α_1



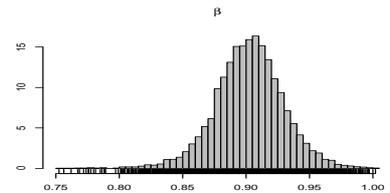
(e) Traço da cadeia de γ_1



(f) Marginal a posteriori de γ_1

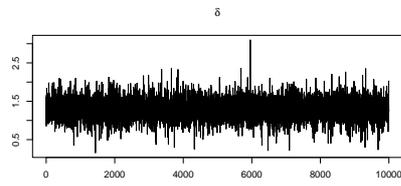


(g) Traço da cadeia de β_1

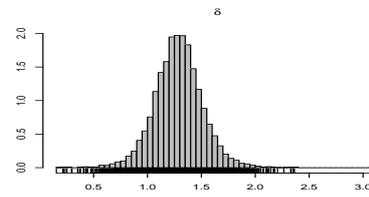


(h) Marginal a posteriori de β_1

FIGURA 6: Traço e histograma das amostras das marginais a posteriori dos parâmetros - dados simulados.

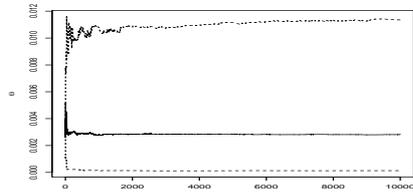


(a) Traço da cadeia de δ

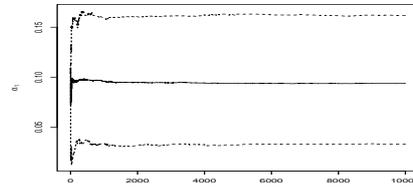


(b) Marginal a posteriori de δ

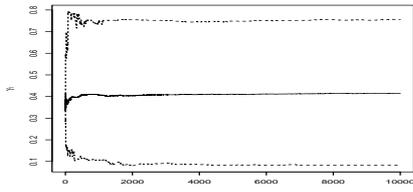
FIGURA 7: Traço e histograma das amostras das marginais a posteriori dos parâmetros - dados simulados.



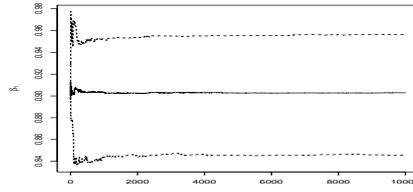
(a) quantis acumulados da cadeia de ω



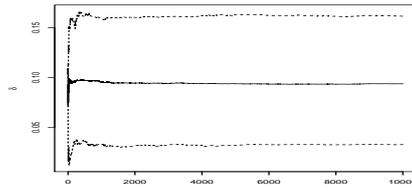
(b) Marginal a posteriori de α_1



(c) quantis acumulados da cadeia de γ



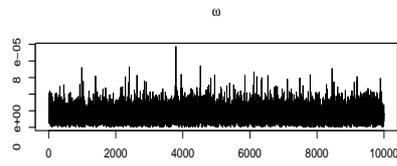
(d) quantis acumulados da cadeia de β_1



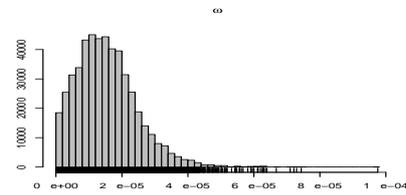
(e) quantis acumulados da cadeia de δ

FIGURA 8: Evolução dos quantis das cadeias dos parâmetros - dados simulados.

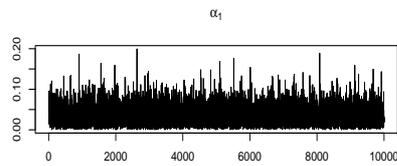
**APÊNDICE B - FIGURAS - RETORNOS DO
IBOVESPA**



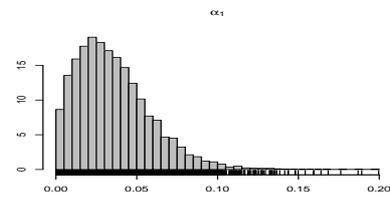
(a) Traço da cadeia de ω



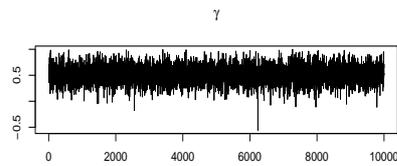
(b) Marginal a posteriori de ω



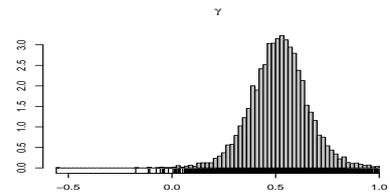
(c) Traço da cadeia de α_1



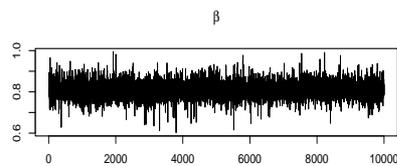
(d) Marginal a posteriori de α_1



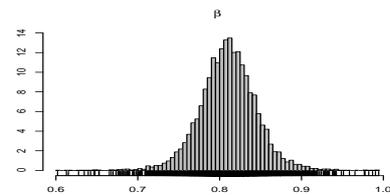
(e) Traço da cadeia de γ_1



(f) Marginal a posteriori de γ_1

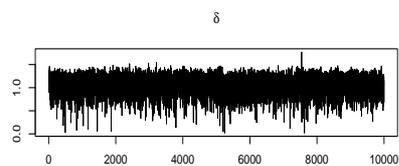


(g) Traço da cadeia de β_1

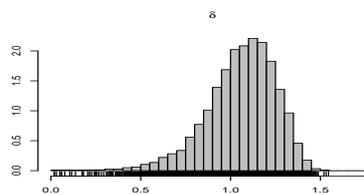


(h) Marginal a posteriori de β_1

FIGURA 9: Traço e histograma das amostras das marginais a posteriori dos parâmetros - Ibovespa.

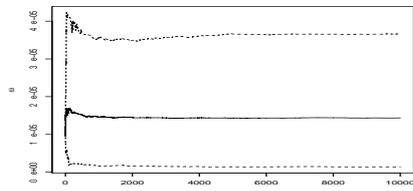


(a) Traço da cadeia de δ

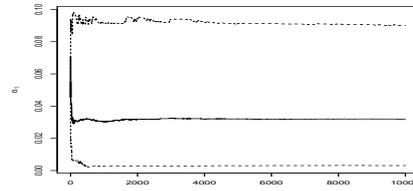


(b) Marginal a posteriori de δ

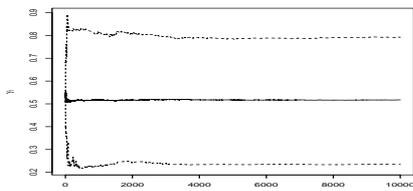
FIGURA 10: Traço e histograma das amostras das marginais a posteriori dos parâmetros - Ibovespa.



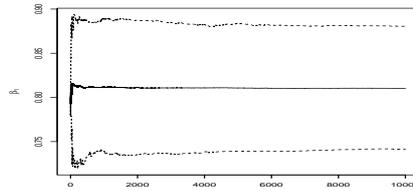
(a) quantis acumulados da cadeia de ω



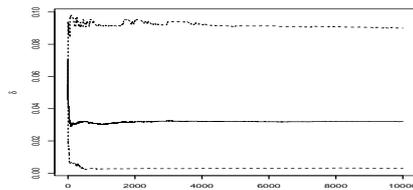
(b) Marginal a posteriori de α_1



(c) quantis acumulados da cadeia de γ



(d) quantis acumulados da cadeia de β_1



(e) quantis acumulados da cadeia de δ

FIGURA 11: Evolução dos quantis das cadeias dos parâmetros - Ibovespa.

APÊNDICE C – PROGRAMA R PARA O MODELO APARCH

```
## COMANDOS PARA DEPURAÇÃO DO PROGRAMA

options(warn=2)

## LIBRARIES UTILIZADAS

library(MASS, keep.source = FALSE)
library(time)

## LENDO OS DADOS

rbov <- scan("ribovespa.txt")

#####

curtose <- function (x, na.rm = FALSE) {
  if (na.rm)
    x <- x[!is.na(x)]
  sum((x-mean(x))^4)/(length(x)*var(x)^2) - 3
}

cat("Retorno médio = ", mean(y), "\n")
cat("Excesso de curtose =", curtose(y), "\n")
cat("Sumario = ", summary(y), "\n")

#####

op <- par (mfrow = c(3, 2), cex = 0.7)

# Gráfico da Série de Log-Retornos:
plot(y, xlab = " ", ylab = " ", col = "black",
main = " ", type="l")
abline (h = 0, col = "grey")

# Histograma:
hist(y, col = "steelblue3", probability = TRUE,
breaks = "FD", border = "white", main="")
abline (h = 0)
lines(density(y, kernel="e"), col="black")

# Grafico Quantil - Quantil:
```

```

qqnorm(y, xlab = " ", ylab = " ", col = "black",
main = " ", cex = 0.5)
qqline(y)

# Autocorrelação Parcial:
acf(y^2, ylab=" ", main=" ")

# Autocorrelação das Volatilidades:
acf(abs(y), ylab=" ", main=" ")

par(op)

#####
# DISTRIBUICOES CANDIDATAS
#####

## T - STUDENT MULTIVARIADA

## GERADOR DE NUMEROS PSEUDO-ALEATORIOS

rmvst <- function(nu,mu,root)
{
nvec=t(root)%*%rnorm(length(mu))
return(nvec/sqrt(rchisq(1,nu)/nu) + mu)
}

## CALCULA A LOG-DENSIDADE

lmvst= function(x,nu,mu,rooti) {
z=as.vector(t(rooti)%*%(x-mu))
return(-((length(x)+nu)/2)*log(nu+z%*%z)
+sum(log(diag(rooti))))
}

## NORMAL MULTIVARIADA

## GERADOR DE NUMEROS PSEUDO-ALEATORIOS

"rmultnorm" <- function(n, mu, vmat, tol = 1e-10)
{
p <- ncol(vmat)
if(length(mu)!=p)
stop(paste("vetor mu tem dimensao errada:",length(mu)))
vs <- La.svd(vmat)
vsqrt <- t(t(vs$vt) %*% (t(vs$u) * sqrt(vs$d)))
ans <- matrix(rnorm(n * p), nrow = n) %*% vsqrt
ans <- sweep(ans, 2, mu, "+")
dimnames(ans) <- list(NULL, dimnames(vmat)[[2]])
ans
}

## CALCULA A LOG-DENSIDADE

lnMv= function(x,mu,rooti) {
z=as.vector(t(rooti)%*%(x-mu))

```

```

return(-.5*(z%*/z) + sum(log(diag(rooti))))
}

#####
#####

METROPOLIS-HASTINGS

## VALORES INICIAIS PARA AS CADEIAS DE MARKOV DE

##ALFAO, ALFA1 EBETA1

a0ini <- 1.266052e-05

a1ini <- 0.02268390

g1ini <- 0.5170186

b1ini <- 0.809478

dini <- 1.128363

## POSTERIORI COM A PRIORI NAO-INFORMATIVA

garch_posteriori <- function(a0, a1, g1, b1, d, y)
{
h <- double(length(y))
h[1] <- var(y)
ht <- double(length(y))
ht[1] <- var(y)
for (i in 2:length(y)) { h[i] <- a0 + a1*((abs(y[i-1]) -
g1*y[i-1])^d) + b1*(h[i-1])
ht[i] <- h[i]^d }
return(-0.5*(sum(log(ht))) - 0.5*(sum((y^2)/ht)))
}

m <- 50000
razao <- 0
canda0p <- numeric(1)
canda1p <- numeric(1)
candg1p <- numeric(1)
candb1p <- numeric(1)
canddp <- numeric(1)
veca0 <- vector("numeric", m)
veca1 <- vector("numeric", m)
vecg1 <- vector("numeric", m)
vecb1 <- vector("numeric", m)
vecd <- vector("numeric", m)

## LAÇO METROPOLIS-HASTINGS

time <- getTime() for (j in 1:m)
{

ve <- c(8.1e-11,0.0006,0.01350941,
0.0007603806, 0.03831023)

```

```

W <- diag(ve)

S <- (1.2)*(W)

beta <- c(1.266052e-05, 2.268390e-02,
          5.170186e-01, 8.094780e-01,
          1.128363)

C <- rmvst(nu=6,mu=beta,root=chol(S))

canda0 <- C[1,1]

canda1 <- C[2,1]

candg1 <- C[3,1]

candb1 <- C[4,1]

candd <- C[5,1]

if(canda0 > 0)
  if(canda1 >= 0)
    if(abs(candg1) < 1)
      if(candb1 >= 0)
        if(candd >= 0)
          {
            canda0p <- canda0
            canda1p <- canda1
            candg1p <- candg1
            candb1p <- candb1
            canddp <- candd
          }

# CALCULO DA LOG-PROBABILIDADE DE ACEITACAO

lprob <- garch_posteriori(canda0p,canda1p,candg1p,candb1p,
                        canddp,y)
-lMvst(x=c(canda0p,canda1p,candg1p,candb1p,canddp),nu=6,
       mu=beta,root=chol(S))
- (garch_posteriori(a0ini,a1ini,g1ini,
                   b1ini,dini,y) -
  lMvst(x=c(a0ini,a1ini,g1ini,b1ini,dini),
        nu=6,mu=beta,root=chol(S)))

if ( lprob > log(runif(1)))
  {
    a0ini <- canda0p
    a1ini <- canda1p
    g1ini <- candg1p
    b1ini <- candb1p
    dini <- canddp
    razao <- (razao*(j-1)+1)/j
  }
veca0[j] <- a0ini
veca1[j] <- a1ini

```

```

        vecg1[j] <- giini
        vecb1[j] <- blini
        vecd[j] <- dini

razao <- razao*((j-1)/j)

cat("iteracao da cadeia= ", j, fill=TRUE)

cat("Taxa de aceitacao = ", razao, fill=TRUE, "\n")

    }
timeReport(time)

##### ANALISE DOS RESULTADOS #####

## VERIFICAÇÃO DA ESTACIONARIEDADE FRACA

sum((1/sqrt(2*pi))*(g*((1+b)^(v)+(1-b)^(v))*2^
((v+1)/2))*gamma((v+1)/2)+k < 1)/10000

## MEDIA E DESVIOS-PADRAO A POSTERIORI
#####

## a0
a0p <- veca0[-c(1:25000)] # BURN-IN 25.000 ITERACOES
z <- a0p[seq(1, 25000, by=5)]
malfa0p <- mean(z)
cat("Media da posteriori de alfa-0 = ", malfa0p, "\n")
cat("DP da posteriori de alfa-0 = ", sd(z), "\n")

## a1
a1p <- veca1[-c(1:25000)]
g <- a1p[seq(1, 25000, by=5)]
malfa1p <- mean(g)
cat("Media da posteriori de alfa-1 = ", malfa1p, "\n")
cat("DP da posteriori de alfa-1 = ", sd(g), "\n")

## g1
g1p <- vecg1[-c(1:25000)]
b <- g1p[seq(1, 25000, by=5)]
malfg1p <- mean(b)
cat("Media da posteriori de gama = ", malfg1p, "\n")
cat("DP da posteriori de gama = ", sd(b), "\n")

## b1
b1p <- vecb1[-c(1:25000)]
k <- b1p[seq(1, 25000, by=5)]
mbetap <- mean(k)
cat("Media da posteriori de beta-1 = ", mbetap, "\n")
cat("DP da posteriori de beta-1 = ", sd(k), "\n")

## d
d1p <- vecd[-c(1:25000)]
v <- d1p[seq(1, 25000, by=5)]
mdp <- mean(v)

```

```

cat("Media da posteriori de delta = ", mdp, "\n")
cat("DP da posteriori de delta = ", sd(v), "\n")

dat <- cbind(z,g,b,k,v)

## DENSIDADES MARGINAIS A POSTERIORI E
## ANALISE DA CONVERGENCIA
##BURN-IN = 25000 ITERAÇÕES

hist(z, col = "gray", ylab=" ", xlab=" ",
main=expression(omega), prob=T)
rug(z)

hist(g, col = "gray", ylab=" ", xlab=" ",
main=expression(alpha[1]),
prob=T)
rug(g)

hist(b, col = "gray", ylab=" ", xlab=" ",
main=expression(gama), prob=T)
rug(b)

hist(k, col = "gray", ylab=" ", xlab=" ",
main=expression(beta), prob=T)
rug(k)

hist(v, col = "gray", ylab=" ", xlab=" ",
main=expression(delta) , prob=T)
rug(v)

## TRAÇO DA CADEIA DE ALFA-0

plot(z, ylab=" ", xlab=" ",
main=expression(omega), type = "l")

## TRAÇO DA CADEIA DE ALFA-1

plot(g, ylab=" ", xlab=" ",
main=expression(alpha[1]),type = "l")

## TRAÇO DA CADEIA DE GAMA

plot(b, ylab=" ", xlab=" ",
main=expression(gama),type = "l")

## TRAÇO DA CADEIA DE BETA-1

plot(k, ylab=" ", xlab=" ",
main=expression(beta),type = "l")

## TRAÇO DA CADEIA DE DELTA

plot(v, ylab=" ", xlab=" ",
main=expression(delta), type = "l")

#####

```

```

#####

## VOLATILIDADE ESTIMADA

volatest <- function(a0, a1, g1, b1, d, y)
{
  h <- double(length(y))
  h[1] <- var(y)
  ht <- double(length(y))
  ht[1] <- var(y)

  for (i in 2:length(y)) {
    h[i] <- a0 + a1*(abs(y[i-1]) - g1*y[i-1])^d +
    b1*(h[i-1])
    ht[i] <- h[i]^d }
  return(ht)
}

x <- volatest(malfa0p, malfa1p, malfg1p, mbetap, mdp, y)

plot(x[40:1000], xlab=" ", ylab=" ",
      main=expression(sigma[t]^d),
      type="l", col="black")

#####
## FIM
#####

```