

UNIVERSIDADE DE SÃO PAULO  
Instituto de Ciências Matemáticas e de Computação  
ISSN 0103-2577

---

**Abordagem Bayesiana com MCMC versus Máxima  
Verossimilhança para Modelos ARCH(p)**

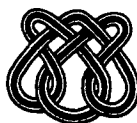
**Marinho Gomes Andrade  
Valéria A. Martins Ferreira  
Sandra C. Oliveira**

N<sup>o</sup> 72

---

NOTAS

Série Estatística



São Carlos – SP  
Ago./2002

**Abordagem Bayesiana com MCMC  
versus Máxima Verossimilhança  
para Modelos ARCH(p)**

Marinho G. Andrade  
Valéria A. Martins Ferreira  
Sandra C. Oliveira

# Abordagem Bayesiana com MCMC versus Máxima Verossimilhança para Modelos ARCH(p)

M. G. Andrade

*Departamento de Ciência da Computação e Estatística, USP-ICMC, São Carlos  
CEP 13560-970, Caixa Postal 668, e-mail: marinho@icmc.sc.usp.br*

V. A. Martins Ferreira

*Departamento de Ciência da Computação e Estatística, USP-ICMC, São Carlos  
CEP 13560-970, Caixa Postal 668, e-mail: va\_ferreira@bol.com.br*

S. C. Oliveira

*Departamento de Ciência da Computação e Estatística, USP-ICMC, São Carlos /  
Departamento de Estatística, DES-UEM, Maringá  
CEP 13560-970, Caixa Postal 668, e-mail: scris@icmc.sc.usp.br*

## Resumo

Neste trabalho consideramos um problema de inferência dos parâmetros dos modelos auto-regressivos com heterocedasticidade condicional, ARCH(p). As estimativas são obtidas usando inferência clássica e inferência Bayesiana via simulação de Monte Carlo em Cadeias de Markov (MCMC). Na abordagem clássica consideramos dados de 500 séries geradas com os modelos ARCH(2) e 1000 séries geradas com os modelos ARCH(5) e ARCH(7). Esses dados são usados para construir a distribuição empírica dos estimadores de máxima verossimilhança dos parâmetros destes modelos. Na análise Bayesiana consideramos uma reparametrização do modelo e adotamos a distribuição a priori normal para os parâmetros transformados. O resumo das distribuições a posteriori é obtido por simulação de Monte Carlo em Cadeias de Markov (MCMC). A comparação entre as duas abordagens é feita para três modelos ARCH(2) sendo: um estacionário, um próximo da região de não estacionariedade e um fora da região de estacionariedade. O efeito da ordem do modelo no problema de inferência é avaliado considerando um modelo ARCH(5) e um modelo ARCH(7). Em geral, as estimativas pontuais de máxima verossimilhança e Bayesiana apresentam-se próximas, mas as estimativas da raiz quadrada do erro quadrático médio (RMSE) e do erro percentual absoluto médio (MAPE) revelam que a abordagem Bayesiana é mais precisa que a abordagem clássica. Em algumas séries o intervalo de confiança para certos parâmetros do modelo apresenta valores negativos, o que viola as restrições impostas aos parâmetros de um modelo ARCH.

**Palavras-chave:** Modelos ARCH, Inferência Bayesiana, Séries Econométricas.

## Abstract

In this work, we consider a inference problem of the parameters for Autoregressive Conditional Heteroskedasticity (ARCH) models. The estimates are obtained by using classical approach and a Bayesian analysis using Markov Chain Monte Carlo methods (MCMC). In the classical approach we consider 500 series gerated with each ARCH(2) model and 1000 series gerated with ARCH(5) and ARCH(7) models. These data are used constructing a empirical distribution for the maximum likelihood estimators of parameters for those models. In the Bayesian analysis we present a reparametrization of model using prior normal densities for the transformed parameters. The posterior densities are obtained by using Monte Carlo Markov Chain methods (MCMC). The comparison between two approaches are illustred considering three ARCH(2) models: one stationary, one close on non-stationary region and other stationary region outside. We also consider ARCH(5) and ARCH(7) models to evaluate the effect of model order. In general, puntual maximum likelihood estimates and Bayesian estimates are close, but the root mean squared error (RMSE) and the mean absolute percentual error (MAPE) reveal that a Bayesian approach is more exact than classical approach. In some series, the confidence interval for some parameters produces negative values, violating the restrictions imposed to parameters of the ARCH model.

**Keywords:** ARCH models, Bayesian inference, Econometric series.

## 1 – INTRODUÇÃO

Uma simples análise gráfica de séries financeiras revela que essas séries apresentam uma alta volatilidade para alguns períodos de tempo. Descrever como essa volatilidade muda com o tempo é fundamental para a avaliação do risco de investimentos e precificação de opções. Além disso, uma inferência econométrica eficiente da média e variância incondicional requer uma especificação correta da variância condicional.

Os modelos mais consagrados para descrever a volatilidade de séries financeiras são os modelos ARCH, propostos por Engle (1982), e os modelos ARCH – generalizados (GARCH) propostos por Bollerslev (1986). Uma ampla revisão das propriedades desses modelos pode ser encontrada em Bollerslev et al (1992).

Nos modelos ARCH é considerado que a volatilidade em dado instante de tempo depende de valores passados da série. A determinação de estimadores de máxima verossimilhança para os parâmetros dos modelos ARCH exige a maximização de uma função não linear. Portanto, estimativas de máxima verossimilhança só podem ser obtidas numericamente. Engle (1982) sugere o uso do método de Newton como um método iterativo de cálculo das estimativas de máxima verossimilhança. Esse procedimento relaxa as restrições impostas aos parâmetros que asseguram estacionariedade na covariância. Por outro lado, calcular estimadores de máxima verossimilhança com restrições envolve várias dificuldades (Geweke, 1986a).

Além disso, procedimentos para identificação, ajuste e diagnóstico de modelos, assim como previsão de valores futuros de séries econométricas, usam propriedades da teoria assintótica. Como os modelos estão muito distantes da linearidade, essa abordagem torna-se analiticamente intratável. Uma alternativa para contornar tais dificuldades é considerar a abordagem Bayesiana desses modelos.

Um dos primeiros trabalhos a propor a abordagem Bayesiana para modelos ARCH foi Geweke (1989). Nesse trabalho foi considerado um caso particular de parametrização que facilitou a utilização de distribuições a priori não informativas. As estimativas dos parâmetros são obtidas usando algoritmo de simulação de Monte Carlo. Mais recentemente uma abordagem Bayesiana foi proposta para modelos ARCH – generalizados (GARCH) em Migon et al (1999) dentro da classe de modelos dinâmicos. Além disso, Nakatsuma (2000), desenvolveu um método de Monte Carlo em Cadeias de Markov (MCMC) para modelos de regressão linear com erros ARMA-GARCH, utilizando distribuições a priori normais para os parâmetros dos mesmos e o algoritmo de simulação Metropolis-Hastings para gerar amostras da distribuição a posteriori de interesse.

Neste trabalho apresentamos uma abordagem Bayesiana para a inferência dos parâmetros de processos ARCH, utilizando técnicas de simulação MCMC. Um dos objetivos deste trabalho é comparar as estimativas Bayesianas com as estimativas de máxima verossimilhança. Esta comparação é feita usando as distribuições dos estimadores de máxima verossimilhança, obtidas empiricamente através da simulação de 500 séries com os modelos ARCH(2) e de 1000 séries com os modelos ARCH(5) e ARCH(7) tendo, cada série, 300 observações. Para cada série gerada, as estimativas de máxima verossimilhança dos parâmetros dos modelos são calculadas numericamente, constituindo uma amostra do estimador de máxima verossimilhança, que permite estimar os erros quadráticos médios e os erros absolutos percentuais médios com relação aos verdadeiros parâmetros do modelo. Esses erros, juntamente com os intervalos de confiança dados pelos percentis 2.5% e 97.5% dessa amostra de estimativas, são comparados aos erros obtidos com a estimativa Bayesiana e aos intervalos com 95% de credibilidade das distribuições a posteriori geradas com o algoritmo Metropolis-Hastings.

Na abordagem Bayesiana apresentada neste trabalho, as distribuições a priori para os parâmetros refletem as restrições gerais que os parâmetros de qualquer modelo ARCH(p) devem satisfazer para assegurar volatilidade finita, ou seja, denotando por  $\alpha_j, j = 1, \dots, p$  os parâmetros do modelo, as distribuições a priori escolhidas devem satisfazer  $\alpha_0 > 0$  e  $0 \leq \alpha_j < 1, j = 1, \dots, p$ . As restrições nos parâmetros que asseguram a estacionariedade na

covariância do processo não são incorporadas a priori, uma vez que um dos objetivos da análise Bayesiana deste trabalho é verificar essa propriedade a posteriori.

## 2 – MODELO ARCH(p)

O modelo linear ARCH(p) proposto por Engle (1982) pode ser resumido como:

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

$$z_t | \Omega_{t-1} \sim P(0, h_t) \quad (2)$$

$$h_t = \alpha_0 + \sum_{j=1}^p \alpha_j z_{t-j}^2 \quad (3)$$

onde:

$P(\cdot)$ : é uma distribuição paramétrica, usualmente a Normal ou  $t$ -Student

$\mathbf{x}_t$ : representa um vetor de variáveis exógenas

$\boldsymbol{\beta}$ : é um vetor de parâmetros desconhecidos

$\Omega_{t-1}$ : representa o conjunto de informações disponíveis até o instante  $t-1$ .

Neste trabalho, vamos considerar  $\boldsymbol{\beta} = 0$  e o processo  $z_t$  satisfazendo o modelo

$$z_t = h_t^{1/2} \varepsilon_t \quad (4)$$

$$h_t = \alpha_0 + \sum_{j=1}^p \alpha_j z_{t-j}^2 \quad (5)$$

onde  $\{\varepsilon_t, t \geq 0\}$  é um ruído branco *i.i.d.*  $N(0,1)$ .

A estimativa de máxima verossimilhança para os parâmetros  $\alpha_j$  parte da hipótese da distribuição de  $z_t$ , ou similarmente de  $\varepsilon_t$ . Para que o modelo (4)-(5) seja plausível ( $h_t > 0$  para todo  $t$ ) devemos ter  $\alpha_0 > 0$  e  $\alpha_j \geq 0$  para  $j = 1, \dots, p$ . Além disso, para que o processo  $z_t$  tenha variância finita, a condição necessária e suficiente é que todas as raízes do polinômio

$1 - \sum_{j=1}^p \alpha_j q^j$  estejam fora do círculo de raio unitário. Quando esta condição é satisfeita, a

variância incondicional de  $z_t$  é dada por  $\alpha_0 / \left(1 - \sum_{j=1}^p \alpha_j\right)$ . A condição necessária para covariância estacionária é que  $\alpha_0 > 0$  e  $\sum_{j=1}^p \alpha_j < 1$ . Neste trabalho essas restrições serão relaxadas, para evitar as dificuldades de maximização da função de verossimilhança sujeita a restrições (Geweke, 1986a).

### 3 – ABORDAGEM CLÁSSICA PARA MODELOS ARCH(p)

Dada uma trajetória observada  $\mathbf{Z} = \{z_t, t = 1, 2, \dots, T\}$  do processo  $z_t$ , a função de verossimilhança de  $z_t$ ,  $t = p+1, \dots, T$  é dada por:

$$L(\boldsymbol{\alpha} | \mathbf{Z}) \propto \prod_{t=p+1}^T \left(\frac{1}{h_t}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t}\right\} \quad (6)$$

onde  $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_p)$ . Denotamos a média do log da função de verossimilhança por:

$$l(\boldsymbol{\alpha} | \mathbf{Z}) = \frac{1}{T} \sum_{t=p+1}^T l_t(\boldsymbol{\alpha} | \mathbf{Z}) \quad (7)$$

onde  $l_t(\boldsymbol{\alpha} | \mathbf{Z})$  indica o log da verossimilhança da  $t$ -ésima observação, ou seja,

$$l_t(\boldsymbol{\alpha} | z_t) = -\frac{1}{2} \log h_t - \frac{z_t^2}{2h_t} \quad (8)$$

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

$$\frac{\partial l}{\partial \boldsymbol{\alpha}} = \frac{1}{T} \sum_{t=p+1}^T \frac{\partial l_t}{\partial \boldsymbol{\alpha}} \quad (9)$$

onde  $\frac{\partial l_t}{\partial \boldsymbol{\alpha}}$  é o vetor  $1 \times (p+1)$  cujas componentes são:

$$\frac{\partial l_t}{\partial \alpha_j} = \frac{1}{2h_t} \left( \frac{\partial h_t}{\partial \alpha_j} \right) \left( \frac{z_t^2}{h_t} - 1 \right) \quad j = 0, 1, \dots, p \quad (10)$$

Da equação (5) temos:

$$\frac{\partial h_t}{\partial \alpha_j} = \begin{cases} 1 & j = 0 \\ z_{t,j}^2 & j = 1, 2, \dots, p, \quad t \geq p+1 \end{cases} \quad (11)$$

Portanto,

$$\frac{\partial l_t}{\partial \alpha} = \frac{1}{2h_t} v_t \left( \frac{z_t^2}{h_t} - 1 \right) \quad (12)$$

onde  $v_t = (1, z_{t-1}^2, \dots, z_{t-p}^2)$  e o gradiente de  $l(\alpha | \mathbf{Z})$  pode ser calculado substituindo (12) em (9), resultando em:

$$\frac{\partial l}{\partial \alpha} = \frac{1}{T} \sum_{t=p+1}^T \frac{1}{2h_t} v_t \left( \frac{z_t^2}{h_t} - 1 \right) \quad (13)$$

Denotando por:

$$\begin{aligned} \tilde{v}_t &= (1, z_{t-1}^2, \dots, z_{t-p}^2) / h_t \quad ; \quad t = p+1, \dots, T \\ \tilde{v}' &= (\tilde{v}'_{p+1}, \dots, \tilde{v}'_T) \\ f_t &= \left( \frac{z_t^2}{h_t} - 1 \right) \quad ; \quad t = p+1, \dots, T \\ f' &= (f_{p+1}, \dots, f_T) \end{aligned} \quad (14)$$

podemos escrever  $\frac{\partial l}{\partial \alpha}$  na forma matricial como:

$$\frac{\partial l}{\partial \alpha} = \frac{1}{2T} \tilde{v}' f \quad (15)$$

Denotando a matriz Hessiana por  $H = \left\{ \frac{\partial^2 l}{\partial \alpha_i \partial \alpha_j}, \quad i, j = 0, 1, \dots, p \right\}$  temos:

$$\frac{\partial^2 l}{\partial \alpha_i \partial \alpha_j} = \frac{1}{T} \sum_{t=p+1}^T \frac{\partial^2 l_t}{\partial \alpha_i \partial \alpha_j} \quad (16)$$

onde:

$$\frac{\partial^2 l_t}{\partial \alpha_i \partial \alpha_j} = \frac{\partial}{\partial \alpha_i} \left[ \frac{1}{2h_t} \frac{\partial h_t}{\partial \alpha_j} \left( \frac{z_t^2}{h_t} - 1 \right) \right] \quad (17)$$

$$\frac{\partial^2 l_t}{\partial \alpha_i \partial \alpha_j} = -\frac{1}{2h_t^2} \left( \frac{\partial h_t}{\partial \alpha_i} \right) \left( \frac{\partial h_t}{\partial \alpha_j} \right) \left( \frac{z_t^2}{h_t} - 1 \right) + \frac{1}{2h_t} \frac{\partial^2 h_t}{\partial \alpha_i \partial \alpha_j} \left( \frac{z_t^2}{h_t} - 1 \right) + \frac{1}{2h_t} \left( \frac{\partial h_t}{\partial \alpha_j} \right) \left( -\frac{z_t^2}{h_t^2} \frac{\partial h_t}{\partial \alpha_i} \right) \quad (18)$$



Da equação (5) temos que  $\frac{\partial^2 h_t}{\partial \alpha_i \partial \alpha_j} = 0$ . Assim, (18) pode ser simplificada para:

$$\frac{\partial^2 l_t}{\partial \alpha_i \partial \alpha_j} = -\frac{1}{2h_t^2} \left( \frac{\partial h_t}{\partial \alpha_i} \right) \left( \frac{\partial h_t}{\partial \alpha_j} \right) \left( 2 \frac{z_t^2}{h_t} - 1 \right) \quad (19)$$

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

$$I_{ij} = -E \left[ \frac{1}{T} \sum_{t=p+1}^T \frac{\partial^2 l_t}{\partial \alpha_i \partial \alpha_j} \right] \quad i, j = 0, 1, \dots, p \quad (20)$$

Usando o fato que  $E \left( 2 \frac{z_t^2}{h_t} - 1 \right) = 2 \frac{E(z_t^2)}{h_t} - 1 = 1$ , os elementos da matriz de informação  $\{\hat{I}_{ij}, i, j = 0, 1, \dots, p\}$  podem ser consistentemente estimados por:

$$\hat{I}_{ij} = \frac{1}{T} \sum_{t=p+1}^T \frac{1}{2h_t^2} \left\{ \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \alpha_j} \right\}; i, j = 0, 1, \dots, p \quad (21)$$

Usando a definição  $\tilde{v}_t = (1, z_{t,1}^2, \dots, z_{t,p}^2) / h_t, t = p+1, \dots, T$  e  $\tilde{v}' = (\tilde{v}'_{p+1}, \dots, \tilde{v}'_T)$ , podemos escrever a matriz de informação por:

$$\hat{I} = \frac{1}{2T} \sum_{t=p+1}^T \tilde{v}_t' \tilde{v}_t = \frac{1}{2T} \tilde{v}' \tilde{v} \quad (22)$$

Um método iterativo para calcular os estimadores de máxima verossimilhança para os parâmetros  $\alpha_j, j = 0, 1, \dots, p$ , proposto em Engle (1982) consiste em iteragir a seguinte equação recursiva:

$$\alpha^{(k+1)} = \alpha^{(k)} + [\hat{I}^{(k)}]^{-1} \left[ \frac{\partial l^{(k)}}{\partial \alpha} \right] \quad (23)$$

onde  $[\hat{I}^{(k)}]$  é a matriz de informação em (22) e  $[\partial l^{(k)} / \partial \alpha]$  o gradiente em (15), ambos avaliados com a estimativa  $\alpha^{(k)}$ . Substituindo (15) e (22) em (23) temos:

$$\alpha^{(k+1)} = \alpha^{(k)} + [\tilde{v}' \tilde{v}]^{-1} \tilde{v}' f \quad (24)$$

Sob as condições em Engle (1982), os estimadores de máxima verossimilhança têm distribuição assintótica normal com limite  $\sqrt{T}(\hat{\alpha} - \alpha) \xrightarrow{D} N(0, I^{-1})$ .

#### 4 – ABORDAGEM BAYESIANA PARA MODELOS ARCH(p)

Considerando uma trajetória observada dos retornos  $\mathbf{Z} = (z_1, \dots, z_T)$ , a abordagem Bayesiana do problema de inferência dos parâmetros do modelo ARCH(p) inicia-se combinando a função de verossimilhança dessa trajetória  $L(\boldsymbol{\alpha} | \mathbf{Z})$  com a distribuição a priori para os parâmetros  $\pi_0(\boldsymbol{\alpha})$ , resultando na distribuição a posteriori dada por:

$$\pi(\boldsymbol{\alpha} | \mathbf{Z}) \propto L(\boldsymbol{\alpha} | \mathbf{Z})\pi_0(\boldsymbol{\alpha}) \quad (25)$$

Nos modelos ARCH(p), os parâmetros estão restritos às condições  $\alpha_0 > 0, \alpha_j \geq 0, j = 1, \dots, p$  e  $\sum_{j=1}^p \alpha_j < 1$  para assegurar estacionariedade na covariância, uma vez que a restrição  $\sum_{j=1}^p \alpha_j < 1$  exige que todos os  $\alpha_j \in (0,1)$ . Na abordagem Bayesiana, essas restrições nos permite assegurar faixas de variação para esses parâmetros, por exemplo, podemos assegurar que existem intervalos  $[a_j, b_j]$  com  $0 < a_j < b_j < 1$ , tais que  $a_j \leq \alpha_j \leq b_j, j = 1, 2, \dots, p$ . Podemos considerar também que  $a_0 \leq \alpha_0 \leq b_0$  com  $a_0 \geq 0$ . Esta análise nos leva à escolha de uma transformação que mapeia os intervalos  $[a_j, b_j]$  em um domínio  $(-\infty, +\infty)$ . Esta reparametrização pode apresentar vantagens quando se pretende usar algoritmos MCMC. Assim, optamos por uma reparametrização dada por:

$$\phi_j = \log\left(\frac{\alpha_j - a_j}{b_j - \alpha_j}\right), \quad j = 0, 1, 2, \dots, p \quad (26)$$

Com esta reparametrização, a função de verossimilhança é escrita da mesma forma que em (6), mas com a volatilidade  $h_t$  reescrita com os parâmetros  $\alpha_j$ , devidamente transformados em  $\phi_j$  por (26), ou seja:

$$L(\boldsymbol{\phi} | \mathbf{Z}) \propto \prod_{t=p+1}^T \left(\frac{1}{h_t(\boldsymbol{\phi})}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t(\boldsymbol{\phi})}\right\} \quad (27)$$

onde  $\boldsymbol{\phi} = (\phi_0, \phi_1, \dots, \phi_p)$ . Denotando por  $\Sigma(\boldsymbol{\phi})$  uma matriz diagonal dada por  $\Sigma(\boldsymbol{\phi}) = \text{diag}(h_{p+1}(\boldsymbol{\phi}), \dots, h_T(\boldsymbol{\phi}))$ , podemos reescrever (27) como:

$$L(\varphi | \mathbf{Z}) \propto [\det(\Sigma(\varphi))]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \mathbf{Z}' [\Sigma^{-1}(\varphi)] \mathbf{Z}\right\} \quad (28)$$

Assumindo que os  $\phi_j, j = 0, 1, 2, \dots, p$ , são independentes com distribuições a priori normal, com média zero e variância conhecida  $\sigma_j^2$  e, denotando por  $\pi_j(\phi_j) = Normal(0, \sigma_j^2), j = 0, 1, 2, \dots, p$ , a distribuição conjunta a priori para  $\varphi = (\phi_0, \phi_1, \dots, \phi_p)$  é dada por:

$$\Pi(\varphi) = \prod_{j=0}^p \pi_j(\phi_j) \quad (29)$$

Com (28) e (29) podemos escrever a distribuição conjunta a posteriori para  $\varphi$  como:

$$\pi(\varphi | \mathbf{Z}) \propto \left( \prod_{i=p+1}^T \left( \frac{1}{h_i(\varphi)} \right)^{1/2} \exp\left\{-\frac{z_i^2}{2h_i(\varphi)}\right\} \right) \left( \prod_{j=0}^p \pi_j(\phi_j) \right) \quad (30)$$

As distribuições condicionais a posteriori para  $\phi_j, j = 0, 1, \dots, p$ , obtidas de (30), são dadas por:

$$\Pi_j(\phi_j | \varphi_{-j}, \mathbf{Z}) \propto L(\varphi | \mathbf{Z}) \pi_j(\phi_j) \quad (31)$$

onde  $\varphi_{-j}$  é um vetor com os parâmetros do modelo exceto o parâmetro  $\phi_j$ , ou seja,  $\varphi_{-j} = (\phi_0, \phi_1, \dots, \phi_{j-1}, \phi_{j+1}, \dots, \phi_p)$ .

As distribuições condicionais dadas em (31) são utilizadas no algoritmo de simulação MCMC, Metropolis-Hastings (Chib & Greenberg, 1995), para gerar uma amostra da distribuição conjunta a posteriori (30).

### Algoritmo Metropolis-Hastings:

- 1) Forneça valores iniciais arbitrários  $\varphi^{(0)} = (\phi_j^{(0)}), j = 0, 1, \dots, p$  e faça  $i=1$ .
- 2) Gere um novo valor  $\beta$  da densidade  $\pi_j(\phi_j)$ , fazendo:

$$\beta = (\phi_0^{(i-1)}, \phi_1^{(i-1)}, \dots, \phi_{j-1}^{(i-1)}, \beta, \phi_{j+1}^{(i-1)}, \dots, \phi_p^{(i-1)})$$

- 3) Calcule a probabilidade de aceitação do novo valor gerado  $\beta$ :

$$\lambda(\phi_j^{(i-1)}, \beta) = \begin{cases} \min\{1, L(\beta)/L(\varphi^{(i-1)})\} & \text{se } L(\varphi^{(i-1)}) \neq 0 \\ 1, & \text{caso contrário} \end{cases}$$

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

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

5) Repetir 2 a 5 para  $j = 0, \dots, p$ .

6) Faça  $i \leftarrow i + 1$  e volte ao passo 2) até a convergência.

Após a convergência os parâmetros originais  $\alpha_j, j = 0, 1, \dots, p$  podem ser recuperados pela transformação inversa:

$$\alpha_j = \frac{b_j e^{\phi_j} + a_j}{1 + e^{\phi_j}}$$

**Nota:** O uso de distribuições a priori informativas como núcleo do algoritmo Metropolis-Hastings é apropriado sempre que tais distribuições forem adequadamente escolhidas e o domínio das mesmas envolverem o domínio da distribuição a posteriori. Isso é possível no caso dos modelos ARCH(p), escolhendo adequadamente os parâmetros dos intervalos  $[a_j, b_j], j = 0, 1, \dots, p$ . A escolha de  $a_j = 0$  e  $b_j = 1$  para  $j = 0, 1, \dots, p$ , pode levar a uma alta taxa de rejeição no algoritmo Metropolis-Hastings.

O problema de inferência Bayesiana pode ser expresso como a avaliação da esperança de uma função de interesse  $g(\alpha)$  sob a densidade a posteriori  $\pi(\alpha | \mathbf{Z})$ :

$$E[g(\alpha)] = \int_{\alpha} g(\alpha) \pi(\alpha | \mathbf{Z}) d\alpha \quad (32)$$

A solução exata da integral múltipla em (32) não pode ser calculada analiticamente no presente caso. No entanto, sempre que  $E[g(\alpha)]$  existe, uma aproximação para (32) pode ser obtida por simulação de Monte Carlo (Geweke, 1986b).

Seja  $\{\alpha^{(i)}, i = 1, \dots, M\}$  uma seqüência i.i.d de um vetor aleatório  $(p+1)$  dimensional com distribuição conjunta  $\pi(\alpha | \mathbf{Z})$ . Numericamente aproximamos  $\bar{g} \equiv E[g(\alpha)]$  por:

$$\bar{g}_M = \frac{1}{M} \sum_{i=1}^M g(\alpha^{(i)}) \quad (33)$$

Sob condições regulares, para a aplicação do Teorema Central do Limite (Geweke, 1986b) temos que  $\bar{g}_M \rightarrow E[g(\alpha)]$ , e pode ser mostrado que  $M^{1/2}(\bar{g}_M - \bar{g}) \rightarrow N(0, \sigma^2)$  com  $\sigma^2 = E[g(\alpha) - \bar{g}]^2$ . Se  $\hat{\sigma}_M^2 = \frac{1}{M-1} \sum_{i=1}^M (g(\alpha^{(i)}) - \bar{g}_M)^2$ , então,  $M\hat{\sigma}_M^2 \rightarrow \sigma^2$ . Referimo-nos a  $\hat{\sigma}_M$  como o erro padrão numérico na avaliação de  $\bar{g}$ . Então, a precisão numérica pode ser assegurada sob a suposição  $\bar{g}_M \sim N(\bar{g}, \hat{\sigma}_M^2)$ , quando  $M$  é grande. Um resumo das principais estatísticas dos parâmetros  $\alpha_j, j = 0, \dots, p$  podem ser obtidas pelo método de Monte Carlo, onde  $M$  representa o tamanho da cadeia gerada com o algoritmo Metropolis-Hastings.

## 5 – COMPARAÇÃO ENTRE AS ABORDAGENS

A comparação entre as abordagens é feita calculando a raiz quadrada do erro quadrático médio (RMSE) e o erro absoluto percentual médio (MAPE) entre os valores verdadeiros dos parâmetros e os valores estimados com cada uma das técnicas. É evidente que esse tipo de comparação só pode ser aplicado para situações hipotéticas, onde os parâmetros dos modelos são conhecidos, como é o caso deste estudo. Mesmo não sendo realista, esta situação hipotética ressalta a potencialidade de cada técnica quanto à precisão na inferência dos parâmetros, quando não está presente o erro devido ao desconhecimento do verdadeiro valor dos parâmetros. Assim, os erros de estimação utilizados para comparar as duas abordagens são:

$$RMSE_j = \sqrt{K^{-1} \sum_{i=1}^K (\theta_j - \hat{\theta}_j^{(i)})^2} \quad (34)$$

$$MAPE_j = \frac{100}{K} \sum_{i=1}^K \left| \frac{\theta_j - \hat{\theta}_j^{(i)}}{\theta_j} \right| \quad (35)$$

onde  $j = 1, 2, \dots, p$ .

Na análise destes erros foram utilizados, para cada modelo, valores estimados com dados das séries simuladas.

## 6 – APLICAÇÕES

### 6.1 – Modelos ARCH(2)

Para compararmos as abordagens clássica e Bayesiana foram geradas 500 séries de tamanho 300, segundo um modelo ARCH(2), e as estimativas de máxima verossimilhança e Bayesiana foram obtidas a partir dos algoritmos propostos neste trabalho. Estas estimativas foram utilizadas na construção da distribuição empírica dos estimadores. Então, foram calculados a raiz quadrada do erro quadrático médio (RMSE) e o erro absoluto percentual médio (MAPE) das estimativas, com relação aos parâmetros verdadeiros. Os três modelos ARCH(2) simulados estão representados pela Tabela 1. Note que o modelo M3 viola a restrição de covariância estacionária  $\alpha_1 + \alpha_2 < 1$ . Este método foi adotado para verificarmos o comportamento das duas técnicas nesta situação.

**Tabela 1: Modelos ARCH(2) simulados**

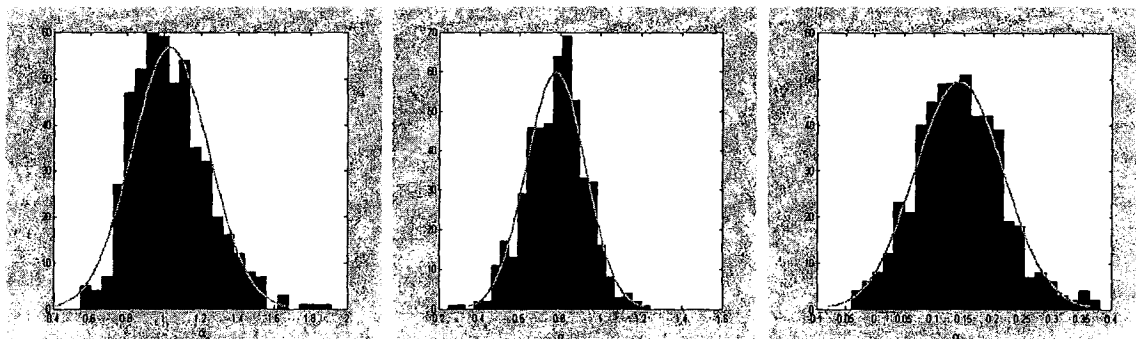
Modelos	$\alpha_0$	$\alpha_1$	$\alpha_2$
M1	1.00	0.80	0.15
M2	1.00	0.50	0.20
M3	1.00	0.90	0.20

Na abordagem Bayesiana consideramos os seguintes hiperparâmetros:

$$a_0 = 0.5, b_0 = 1.0; \quad a_1 = 0.0, b_1 = 1.0; \quad a_2 = 0.0, b_2 = 0.5 \quad \text{e} \quad \sigma_i^2 = 1.0, i = 1, 2, 3$$

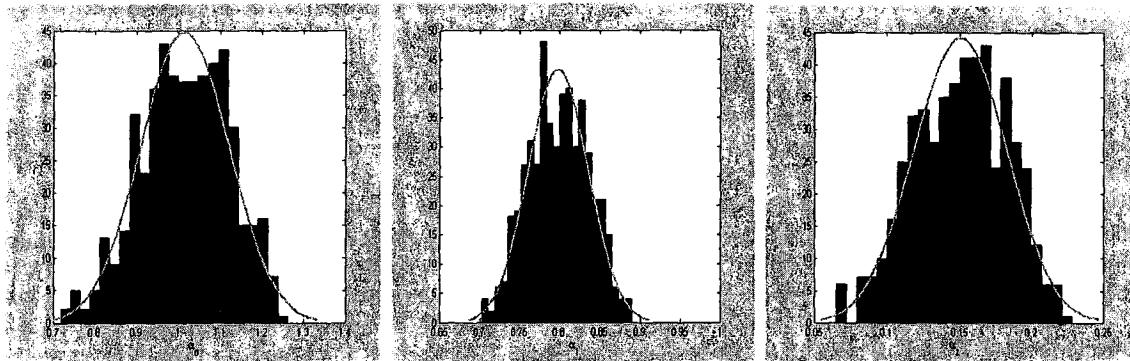
Na implementação do algoritmo Metropolis-Hastings, para cada parâmetro foi simulada uma cadeia com 5000 iterações, descartado 50% e tomado valores de 5 em 5, totalizando uma amostra de tamanho 500. A convergência dos parâmetros foi verificada pelo critério de Geweke (1992).

Os gráficos da Figura 1 apresentam a distribuição empírica dos estimadores de máxima verossimilhança dos parâmetros do modelo M1. Este resultado é semelhante aos obtidos para os modelos M2 e M3, que serão omitidos aqui.



**Figura 1 – Distribuição empírica dos estimadores do modelo M1**

A Figura 2 apresenta os histogramas construídos com as amostras geradas por MCMC da distribuição a posteriori dos parâmetros  $\alpha_0, \alpha_1$  e  $\alpha_2$ .



**Figura 2 – Distribuição a posteriori dos parâmetros do modelo M1**

Os erros de estimativas apresentados na Tabela 2 a seguir, para cada um dos modelos da Tabela 1, são usados para compararmos as duas abordagens. Estas avaliações mostram que a abordagem Bayesiana apresenta estimativas mais precisas do que as estimativas de máxima verossimilhança. Este comportamento não muda conforme os parâmetros do modelo saem da região de estacionariedade.

**Tabela 2: Comparação para os modelos ARCH(2) simulados**

Modelo M1	$\hat{\alpha}$	RMSE	MAPE(%)	IC(95%)	
$\alpha_0$	Bayes	1.0178	0.1068	8.78	[0.8082,1.2024]
	MLE	1.0365	0.2115	16.42	[0.7081,1.5019]
$\alpha_1$	Bayes	0.7989	0.0375	3.88	[0.7277,0.8661]
	MLE	0.7830	0.1456	14.26	[0.4894,1.0649]
$\alpha_2$	Bayes	0.1509	0.0324	17.56	[0.0855,0.2079]
	MLE	0.1435	0.0741	39.28	[0.0024,0.3026]
Modelo M2	$\hat{\alpha}$	RMSE	MAPE(%)	IC(95%)	
$\alpha_0$	Bayes	1.0098	0.1028	8.49	[0.8099,1.1982]
	MLE	1.0261	0.1665	13.55	[0.7418,1.4299]
$\alpha_1$	Bayes	0.5025	0.0920	14.69	[0.3232,0.6683]
	MLE	0.4836	0.1180	18.73	[0.2615,0.7058]
$\alpha_2$	Bayes	0.2202	0.0554	22.25	[0.1171,0.3279]
	MLE	0.1935	0.0827	33.05	[0.1851,0.3726]

Modelo M3	$\hat{\alpha}$	RMSE	MAPE(%)	IC(95%)	
$\alpha_0$	Bayes	1.0041	0.1056	8.62	[0.7971,1.1890]
	MLE	1.0315	0.2108	16.64	[0.6622,1.5060]
$\alpha_1$	Bayes	0.9240	0.1069	9.75	[0.7309,1.1164]
	MLE	0.8750	0.1543	13.62	[0.5599,1.1625]
$\alpha_2$	Bayes	0.2194	0.0539	21.75	[0.1208,0.3205]
	MLE	0.1949	0.0746	29.99	[0.0429,0.3401]

### 6.2 – Modelo ARCH(5) e modelo ARCH(7)

Para verificarmos o efeito da ordem do modelo na precisão das estimativas, consideramos dois modelos de ordem elevada: um modelo ARCH(5) e um ARCH(7), com os parâmetros dados pela Tabela 3. Na abordagem Bayesiana consideramos as distribuições a priori dos parâmetros definidas nas faixas  $[a_i, b_i]$ , dadas pela Tabela 4. Foi considerado ainda, em ambos os modelos,  $\sigma_i^2 = 1.0$ ,  $i = 1, \dots, p$ , como variância para as distribuições a priori.

**Tabela 3: Modelos simulados**

Modelos	$\alpha_0$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$
ARCH(5)	1.00	0.50	0.20	0.15	0.08	0.05	-	-
ARCH(7)	1.00	0.50	0.20	0.10	0.05	0.03	0.06	0.02

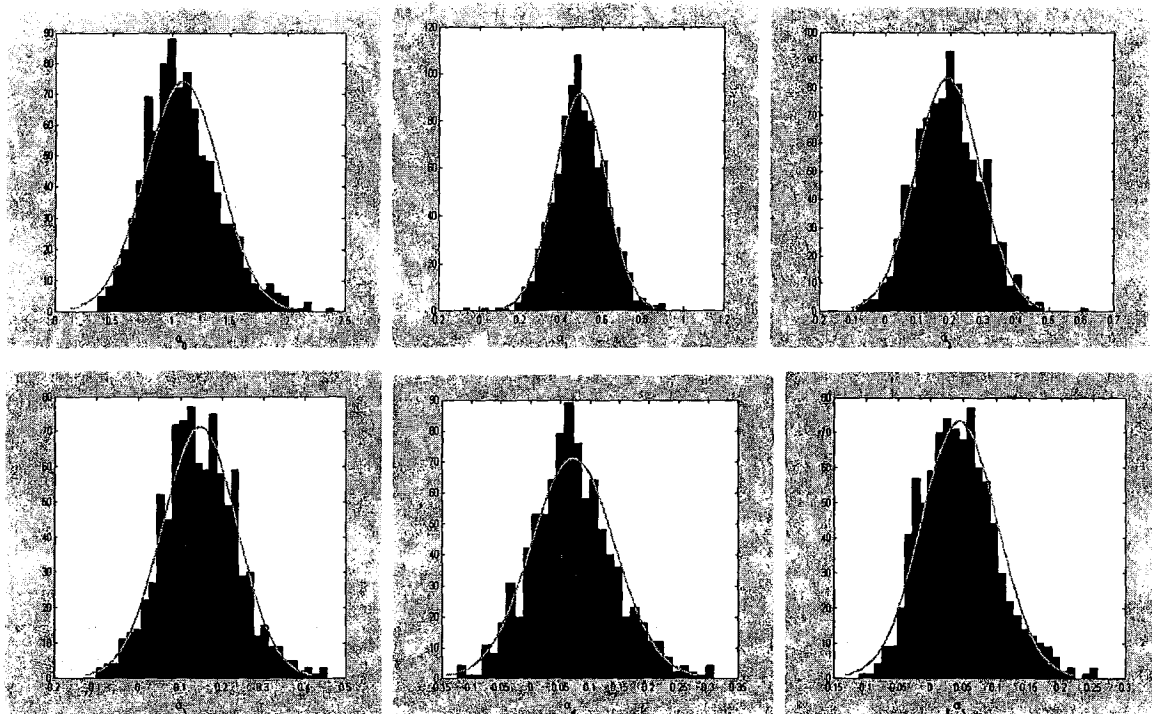
**Tabela 4: Distribuições a priori dos modelos ARCH(5) e ARCH(7) simulados**

Modelos	$\alpha_0$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$
$a_i$	0.50	0.00	0.00	0.00	0.00	0.00	-	-
$b_i$	1.50	1.00	0.50	0.30	0.20	0.10	-	-
$a_i$	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$b_i$	1.50	1.00	0.50	0.20	0.10	0.06	0.12	0.04

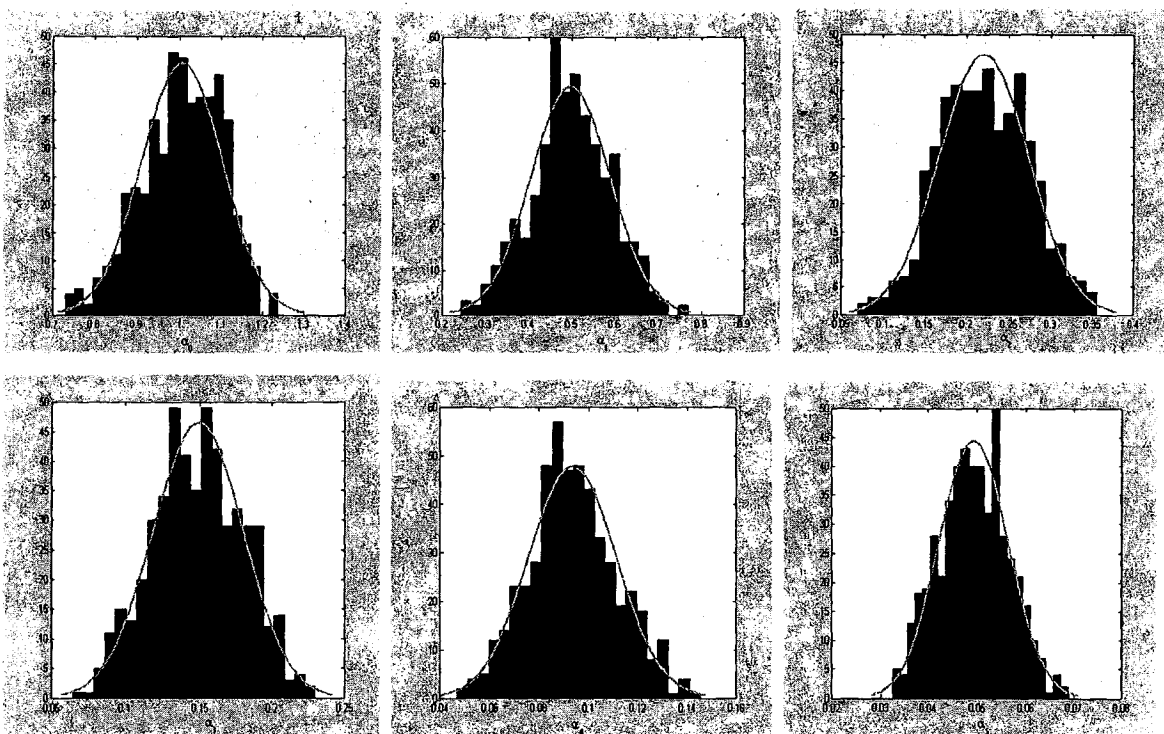
No cálculo das estimativas de máxima verossimilhança do modelo ARCH(5) foram geradas 1000 séries, mas o método iterativo de cálculo das estimativas só convergiu para 901 destas séries. Na abordagem Bayesiana foi implementado o algoritmo Metropolis-Hastings, através de procedimento análogo ao do modelo ARCH(2).



Os histogramas representando as distribuições empíricas dos estimadores de máxima verossimilhança e as distribuições a posteriori dos parâmetros do modelo ARCH(5), são apresentados nas Figuras 3 e 4, respectivamente.



**Figura 3 – Distribuição empírica dos estimadores do modelo ARCH(5)**



**Figura 4 – Distribuição a posteriori dos parâmetros do modelo ARCH(5)**

A Tabela 5 a seguir mostra os erros RMSE e MAPE, juntamente com os intervalos de 95% de confiança dos estimadores clássicos e os intervalos de 95% de credibilidade dos estimadores Bayesianos.

**Tabela 5: Comparação para os modelos ARCH(5) simulados**

Modelo M1	$\hat{\alpha}$	RMSE	MAPE(%)	IC(95%)	
$\alpha_0$	Bayes	1.0076	0.0976	7.94	[0.7978,1.1734]
	MLE	1.0959	0.3328	25.33	[0.5461,1.8299]
$\alpha_1$	Bayes	0.4981	0.0924	14.72	[0.3144,0.6650]
	MLE	0.4905	0.1263	19.86	[0.2458,0.7347]
$\alpha_2$	Bayes	0.2208	0.0571	23.12	[0.1157,0.3273]
	MLE	0.1928	0.0985	39.51	[0.0103,0.3934]
$\alpha_3$	Bayes	0.1496	0.0307	17.79	[0.0908,0.2059]
	MLE	0.1483	0.0898	47.63	[-0.0263,0.3424]
$\alpha_4$	Bayes	0.0937	0.0225	22.53	[0.0599,0.1302]
	MLE	0.0719	0.0711	70.04	[-0.0556,0.2215]
$\alpha_5$	Bayes	0.0494	0.0071	11.59	[0.0361,0.0633]
	MLE	0.0445	0.0586	93.59	[-0.0533,0.1764]

Os resultados mostram que as estimativas Bayesianas apresentam menores erros quadráticos médios e percentuais do que as estimativas de máxima verossimilhança. Além disso, os intervalos Bayesianos com 95% de credibilidade são menores do que os intervalos clássicos com 95% de confiança.

Para os parâmetros  $\alpha_3, \alpha_4$  e  $\alpha_5$ , os intervalos clássicos apresentam limites inferiores negativos, o que nos leva erroneamente à suposição de que estes parâmetros não são significativamente diferentes de zero.

Os erros na Tabela 5 e os erros na Tabela 2 mostram que, para o modelo de ordem maior, ARCH(5), as estimativas de máxima verossimilhança são menos precisas do que para o modelo ARCH(2).

No cálculo das estimativas de máxima verossimilhança do modelo ARCH(7) foram geradas 1000 séries, mas o método numérico apresentou convergência para 655 das mesmas. A comparação entre as abordagens clássica e Bayesiana para o modelo ARCH(7) é apresentada na Tabela 6 a seguir.

**Tabela 6: Comparação para os modelos ARCH(7) simulados**

Modelo M1	$\hat{\alpha}$	RMSE	MAPE(%)	IC(95%)	
$\alpha_0$	Bayes	1.0227	0.0968	7.96	[0.8298,1.1911]
	MLE	1.1274	0.3903	28.60	[0.5461,1.8299]
$\alpha_1$	Bayes	0.4973	0.0629	10.29	[0.3712,0.6082]
	MLE	0.4849	0.1275	20.22	[0.2458,0.7347]
$\alpha_2$	Bayes	0.2215	0.0544	21.71	[0.1165,0.3153]
	MLE	0.1963	0.0971	39.05	[0.0103,0.3934]
$\alpha_3$	Bayes	0.0995	0.0179	14.54	[0.0650,0.1337]
	MLE	0.0949	0.0773	61.57	[-0.0263,0.3424]
$\alpha_4$	Bayes	0.0492	0.0063	10.33	[0.0378,0.0618]
	MLE	0.0455	0.0730	118.11	[-0.0556,0.2215]
$\alpha_5$	Bayes	0.0298	0.0028	7.46	[0.0249,0.0335]
	MLE	0.0285	0.0628	167.45	[-0.0533,0.1764]
$\alpha_6$	Bayes	0.0596	0.0092	12.45	[0.0428,0.0776]
	MLE	0.0535	0.0639	86.83	[-0.0533,0.1764]
$\alpha_7$	Bayes	0.0198	0.0016	6.59	[0.0170,0.0234]
	MLE	0.0215	0.0557	217.58	[-0.0533,0.1764]

## CONCLUSÃO

Este trabalho faz uma comparação entre as abordagens clássica e Bayesiana para estimação dos parâmetros de modelos ARCH(p), usando séries simuladas a partir de modelos conhecidos e estimando os parâmetros destes modelos com as duas técnicas propostas: a inferência clássica, que propõe métodos quasi-Newton para calcular as estimativas de máxima verossimilhança, e a inferência Bayesiana, junto com algoritmo de simulação MCMC.

A comparação apresentada neste trabalho nos leva à conclusão de que a abordagem Bayesiana é muito mais precisa no que se refere ao cálculo das estimativas pontuais e por intervalo, ou seja, os intervalos com 95% de credibilidade obtidos com a inferência Bayesiana têm amplitude inferior aos intervalos com 95% de confiança obtidos com a inferência clássica. Além disso, as estimativas pontuais tornam-se mais imprecisas à medida que a ordem do modelo cresce, o que não ocorre com a inferência Bayesiana.

Assim, podemos concluir que, apesar do maior esforço computacional exigido pelos algoritmos de simulação MCMC em relação aos métodos quasi-Newton usados no cálculo da inferência clássica, a precisão das estimativas recomenda a abordagem Bayesiana para a inferência dos parâmetros dos modelos ARCH(p), para  $p > 2$ .

## REFERÊNCIAS

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

Bollerslev, T.; Chou, R.Y.; Kroner, K.F. (1992). *ARCH Modeling in Finance: a review of the theory and empirical evidence*, Journal of Econometrics, Vol. 52, pp. 5-59.

Chib, S.; Greenberg, E. (1995). *Understanding the Metropolis-Hastings Algorithm*, The American Statistician, Vol. 49, No. 4, pp. 327-335.

Engle, R.F. (1982). *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation*, Econometrica, Vol. 50, pp. 987-1007.

Geweke, J. (1986a). *Exact inference in the inequality constrained normal linear regression model*, Journal of Applied Econometrics, Vol. 1, pp. 127-141.

Geweke, J. (1986b). *Bayesian inference in econometric models using Monte Carlo integration*, Manuscript (Duke University, Durham, NC).

Geweke, J. (1989). *Exact Predictive densities for linear models with ARCH disturbances*, Journal of Econometrics, Vol. 40, pp. 63-86.

Migon, H. S.; Mazucheli, J. (1999). *Modelos GARCH Bayesianos: Métodos Aproximados e Aplicações*, Revista de Econometria, Vol. 19, No.1, pp. 111-138.

Nakatsuma, T. (2000). *Bayesian analysis of ARMA-GARCH models: A Markov chain sampling approach*. Journal of Econometrics, Vol. 95, pp. 57-69.

# NOTAS DO ICMC

## SÉRIE ESTATÍSTICA

- 071/2002 ANDRADE, MARINHO G.; FERREIRA, VALÉRIA A. MARTINS; OLIVEIRA, SANDRA C. – Abordagem bayesiana com MCMC versus máxima verossimilhança para modelos ARCH(p)
- 070/2002 LEÃO JR., D.; FRAGOS, M. D.; VAL, J. B. R. DO; ANDRADE, M. G. – Separable Hansdorff measurable Radon spaces
- 069/2002 AOKI, R.; ACHCAR, J.A.; BOLFARINR, H.; SINGER, J.M. – Bayesian analysis of paired data in mixture of normal distributions.
- 068/2001 LOIBEL, S.; VAL, J.B.R.; ANDRADE, M.G. – Comparação entre as abordagens Bayesiana e clássica do problema de inferência com modelo de crescimento populacional de Richards.
- 067/2001 ACHCAR, J.A; BARBOSA, V.F. – Análise bayesiana para dados de sobrevivência com funções de risco em forma de U.
- 066/2001 MAZUCHELI, J.; LOUZADA-NETO, F.; ACHCAR, J.A. – A survival model with a surviving fraction.
- 065/2001 MAZUCHELI, J.; LOUZADA-NETO, F.; ACHCAR, J.A. – Infant mortality model for lifetime data.
- 064/2001 SOUZA, C.N.; ACHCAR, J.A.; MAZUCHELI, J. – Uso de métodos MCMC para análise bayesiana de dados de sobrevivência na presença de covariáveis.
- 063/2001 ACHCAR J.A.; JUNQUEIRA, J.J.G. – Extra-biomial variability: a bayesian approach.
- 062/2000 WRUCK, E.; ACHCAR J.A.; MAZUCHELI, J. – Classification and discrimination for populations with mixture of multivariate normal distributions.