

# Abordagem Bayesiana para modelos estocásticos com heterocedasticidade para os retornos IBovespa

**Sandra C. Oliveira**

Universidade Estadual Paulista – UNESP/Campus de Tupã – Tupã-SP  
sandra@tupa.unesp.br

**Marinho G. Andrade**

Universidade de São Paulo – USP/ICMC – São Carlos-SP  
marinho@icmc.usp.br

## Resumo

Neste trabalho comparamos as estimativas Bayesianas obtidas para os parâmetros de processos da família ARCH considerando distribuição normal e  $t$  de *student* para a distribuição condicional da série de retornos. Adotamos ainda distribuição a priori não-informativa e consideramos uma reparametrização dos modelos estudados para mapear o espaço dos parâmetros no espaço real. Os sumários a posteriori foram obtidos por meio dos métodos de simulação de Monte Carlo em Cadeias de Markov (MCMC). Com a reparametrização adotada, reduzimos a taxa de rejeição do algoritmo de simulação MCMC, acelerando o processo de convergência deste. A metodologia foi avaliada considerando uma série de retornos IBovespa. Para selecionar o modelo que apresentou melhor ajuste aos dados, utilizamos o Critério da Densidade Preditiva Ordenada.

**Palavras-chave:** processos da família ARCH, abordagem Bayesiana, métodos MCMC, retornos financeiros.

## 1. Introdução

Existe uma ampla variedade de modelos para a estimação da volatilidade de séries de retorno de ativos financeiros, e os mais difundidos na literatura são os modelos auto-regressivos com heterocedasticidade (ARCH), propostos por Engle (1982), e sua extensão, os modelos ARCH Generalizados (GARCH), propostos por Bollerslev (1986). Tais modelos caracterizam uma dependência não-linear entre os retornos, função da dependência serial da variância condicional. Uma ampla revisão das propriedades destes modelos pode ser encontrada em Bollerslev (2008).

Como é considerado que a volatilidade em um dado instante de tempo depende dos valores passados da série, a determinação de estimadores de Máxima Verossimilhança (MV) dos parâmetros de modelos da família ARCH requer a maximização de uma função não-linear. Portanto, as estimativas só podem ser obtidas numericamente. Engle (1982) sugere o uso do método de Newton como um método iterativo para o cálculo das estimativas de MV. Este procedimento

relaxa as restrições impostas aos parâmetros, que asseguram estacionariedade na covariância. Por outro lado, a determinação de estimadores de MV com tais restrições envolve algumas dificuldades.

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

Neste contexto, o objetivo deste trabalho é comparar as estimativas Bayesianas para parâmetros de modelos  $AR(p)$ - $ARCH(q)$ ,  $ARCH(q)$  e  $GARCH(q,r)$  considerando distribuições normal e  $t$  de student para a distribuição condicional da série. Adotamos ainda distribuições a priori não-informativas e consideramos, em cada caso, uma reparametrização dos modelos estudados para mapear o espaço dos parâmetros no espaço real. Os sumários a posteriori foram obtidos por meio dos métodos de simulação de Monte Carlo em Cadeias de Markov (MCMC).

## 2. Modelos estocásticos com heterocedasticidade para séries temporais

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

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

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

$$y_t | \Omega_{t-1} \sim P(\mathbf{x}_t \boldsymbol{\beta}, h_t) \quad (2)$$

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

$$z_t = y_t - \mathbf{x}_t \boldsymbol{\beta} \quad (4)$$

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

$$y_t = \sum_{i=1}^p \beta_i y_{t-i} + z_t \quad (5)$$

Seja  $z_t$  um processo que satisfaz o modelo  $z_t = h_t^{1/2} \varepsilon_t$  em que  $\{\varepsilon_t, t \geq 0\}$  é uma seqüência de variáveis aleatórias independentes e identicamente distribuídas com média zero e variância um, independente de  $\mathbf{x}_t$ . Na prática, é comumente suposto que  $\varepsilon_t \sim N(0,1)$  ou  $\varepsilon_t \sim t_\nu$  (distribuição  $t$  de student com  $\nu$  graus de liberdade). Uma interpretação para o modelo definido em (1)-(4) é que os distúrbios na regressão linear seguem um processo ARCH( $q$ ).

Para que o modelo (1)-(4) seja plausível ( $h_t > 0$  para todo  $t$ ), devemos ter  $\alpha_0 > 0$  e  $\alpha_j \geq 0, j=1, \dots, q$ . Além disso, o processo  $y_t$  tem variância finita e, portanto, tem covariância estacionária se, e somente se, todas as raízes dos polinômios  $1 - \sum_{i=1}^p \beta_i l^i$  e  $1 - \sum_{j=1}^q \alpha_j l^j$  estiverem fora do círculo de raio unitário (ENGLE, 1982: Teorema 2). Quando estas condições são satisfeitas,

pode ser mostrado que a variância incondicional de  $y_t$  é dada por: 
$$\frac{\alpha_0}{\left(1 - \sum_{j=1}^q \alpha_j\right)} \left/ \left(1 - \sum_{i=1}^p \beta_i \rho_i\right)\right.,$$

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

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

Seja  $\mathbf{Y} = \{y_t, t=1, 2, \dots, T\}$  uma trajetória observada do processo  $y_t$  e que  $\mathbf{x}_t$  envolva somente os “ $p$ ” valores passados de  $y_t$ . Considerando normalidade para  $\varepsilon_t$ , temos que a função de verossimilhança de  $y_t, t=p+q+1, \dots, T$ , condicionada às  $p+q$  primeiras observações (assumindo que elas são conhecidas) é definida como:

$$L(\mathbf{Y} | \boldsymbol{\alpha}, \boldsymbol{\beta}) = (2\pi)^{-\frac{(T-p-q)}{2}} \prod_{t=p+q+1}^T \left(\frac{1}{h_t}\right)^{1/2} \exp\left\{-\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t}\right\} \quad (6)$$

sendo  $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$  e  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ . Assumindo que  $\varepsilon_t$  tem uma distribuição  $t$  de student

padronizada, tal que  $t_\nu \sim t$  de student com  $\nu$  graus de liberdade, isto é,  $\varepsilon_t = \frac{t_\nu}{\sqrt{\nu/(\nu-2)}}$ , temos que a

função de verossimilhança de  $y_t, t=q+1, \dots, T$ , condicionada às  $p+q$  primeiras observações, é dada por:

$$L(\mathbf{Y} | \boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{t=p+q+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}} \left(\frac{1}{h_t}\right)^{1/2} \left(1 + \frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t(v-2)}\right)^{-\frac{(v+1)}{2}} \quad (7)$$

sendo  $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$ ,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$  e para algum  $v$ .

## 2.2. Modelos ARCH( $q$ )

Se no modelo de regressão proposto por Engle (1982), definido pelas expressões (1)-(4), considerarmos  $\boldsymbol{\beta} = \mathbf{0}$ , então, temos uma nova estrutura, resumida como:

$$y_t = z_t \quad (8)$$

$$z_t | \Omega_{t-1} \sim P(0, h_t) \quad (9)$$

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j z_{t-j}^2 \quad (10)$$

em que  $z_t$  representa uma série de retornos,  $P(\cdot)$  é uma distribuição paramétrica, e  $\Omega_{t-1}$  representa o conjunto de informações disponíveis até o instante  $t-1$ . Uma interpretação para o modelo definido em (8)-(10) é que os retornos na regressão linear seguem um processo ARCH( $q$ ).

Da mesma forma, para que o modelo (8)-(10) seja plausível ( $h_t > 0$  para todo  $t$ ), devemos ter  $\alpha_0 > 0$  e  $\alpha_j \geq 0$  para  $j = 1, \dots, q$ . Além disso, o processo  $z_t$  tem variância finita e, portanto, tem

covariância estacionária se, e somente se, todas as raízes do polinômio  $1 - \sum_{j=1}^q \alpha_j l^j$  estiverem fora

do círculo de raio unitário. Assim, pode ser mostrado que a variância incondicional de  $z_t$  é dada por

$\alpha_0 / \left(1 - \sum_{j=1}^q \alpha_j\right)$ , cuja condição para que o processo tenha covariância estacionária é  $\sum_{j=1}^q \alpha_j < 1$ .

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

$$L(\mathbf{Z} | \boldsymbol{\alpha}) = (2\pi)^{-\frac{(T-q)}{2}} \prod_{t=q+1}^T \left(\frac{1}{h_t}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t}\right\} \quad (11)$$

sendo  $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$ . Assumindo que  $\varepsilon_t$  tem uma distribuição  $t$  de student padronizada, temos que a função de verossimilhança de  $z_t$ ,  $t = q+1, \dots, T$ , condicionada às  $q$  primeiras observações, é dada por (MORETTIN, 2008):

$$L(\mathbf{Z} | \mathbf{a}) = \prod_{t=q+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}} \left(\frac{1}{h_t}\right)^{1/2} \left(1 + \frac{z_t^2}{h_t(v-2)}\right)^{-\frac{(v+1)}{2}} \quad (12)$$

sendo  $\mathbf{a} = (\alpha_0, \alpha_1, \dots, \alpha_q)'$  e para algum  $v$ .

Assim, as funções de verossimilhança descritas em (6), (7), (11) e (12) podem ser maximizadas com respeito aos respectivos parâmetros desconhecidos.

### 2.3. Modelos GARCH( $q, r$ )

O modelo de regressão proposto por Engle (1982), e generalizado por Bollerslev (1986), com média nula e expresso como uma combinação linear de variáveis exógenas, tem uma estrutura que pode ser resumida como:

$$y_t = z_t \quad (13)$$

$$z_t | \Omega_{t-1} \sim P(\mathbf{0}, h_t) \quad (14)$$

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j z_{t-j}^2 + \sum_{i=1}^r \lambda_i h_{t-i} \quad (15)$$

em que  $z_t$  representa uma série de retornos,  $P(\cdot)$  é uma distribuição paramétrica e  $\Omega_{t-1}$  representa o conjunto de informações disponíveis até o instante  $t-1$ . Quando  $r=0$ , o processo se reduz a um ARCH( $q$ ).

Seja  $z_t$  um processo que satisfaz o modelo  $z_t = h_t^{1/2} \varepsilon_t$  tal que  $\{\varepsilon_t, t \geq 0\}$  é uma seqüência de variáveis aleatórias independentes e identicamente distribuídas com média zero e variância um. Desta forma, uma interpretação para o modelo definido em (13)-(15) é que os distúrbios na regressão linear seguem um processo GARCH( $q, r$ ).

Para que o modelo (13)-(15) seja plausível ( $h_t > 0$  para todo  $t$ ), devemos ter  $q > 0$ ,  $r \geq 0$ ,  $\alpha_0 > 0$ ,  $\alpha_j \geq 0, j = 1, 2, \dots, q$ ,  $\lambda_i \geq 0, i = 1, 2, \dots, r$ . Assim, pode ser mostrado que a

variância incondicional de  $z_t$  é dada por  $\alpha_0 / \left(1 - \sum_{j=1}^q \alpha_j - \sum_{i=1}^r \lambda_i\right)$ , cuja condição para que o

processo tenha covariância estacionária é  $\sum_{j=1}^q \alpha_j - \sum_{i=1}^r \lambda_i < 1$ .

Seja  $\mathbf{Z} = \{z_t, t = 1, 2, \dots, T\}$  uma trajetória observada do processo  $z_t$  e  $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \dots, \alpha_q, \lambda_1, \lambda_2, \dots, \lambda_r)'$ . Considerando normalidade para  $\varepsilon_t$ , temos que a função de verossimilhança de  $z_t, t = q+r+1, \dots, T$ , condicionada às  $q+r$  primeiras observações (assumindo

que elas são conhecidas) é dada por:

$$L(\mathbf{Z} | \boldsymbol{\theta}) = (2\pi)^{-\frac{(T-q-r)}{2}} \prod_{t=q+r+1}^T \left( \frac{1}{h_t} \right)^{1/2} \exp \left\{ -\frac{z_t^2}{2h_t} \right\} \quad (16)$$

Assumindo que  $\varepsilon_t$  tem uma distribuição  $t$  de student padronizada, isto é,  $\varepsilon_t = \frac{t_v}{\sqrt{v/(v-2)}}$  tal que  $t_v \sim t$  de student com  $v$  graus de liberdade, temos que a função de verossimilhança de  $z_t$ ,  $t = q+r+1, \dots, T$ , condicionada às  $q+r$  primeiras observações é dada por (MORETTIN, 2008):

$$L(\mathbf{Z} | \boldsymbol{\theta}) = \prod_{t=q+r+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)}\pi} \left( \frac{1}{h_t} \right)^{1/2} \left( 1 + \frac{z_t^2}{h_t(v-2)} \right)^{-\frac{(v+1)}{2}} \quad (17)$$

$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \dots, \alpha_q, \lambda_1, \lambda_2, \dots, \lambda_r)'$  e para algum  $v$ . Assim, as funções de verossimilhança descritas em (16) e (17) podem ser maximizadas com respeito aos respectivos parâmetros desconhecidos.

### 3. Inferência Bayesiana

Considerando uma trajetória observada de retornos  $\mathbf{Y} = \{y_t, t = 1, 2, \dots, T\}$ , a abordagem Bayesiana para a inferência dos parâmetros dos processos AR( $p$ )-ARCH( $q$ ), ARCH( $q$ ) e GARCH( $q, r$ ) parte da combinação das funções de verossimilhança definidas para esta trajetória com uma distribuição a priori para os parâmetros, através da regra de Bayes:

$$\pi(\boldsymbol{\bullet} | \mathbf{Y}) \propto L(\mathbf{Y} | \boldsymbol{\bullet})\pi(\boldsymbol{\bullet}) \quad (18)$$

A expressão  $\pi(\boldsymbol{\bullet} | \mathbf{Y})$  é chamada de distribuição a posteriori do(s) parâmetro(s) de interesse e diz como esta(s) variável(eis) aleatória(s) está(ão) distribuída(s) após os dados terem sido observados.

Quando se tem pouco ou nenhum conhecimento prévio sobre a distribuição dos parâmetros de um modelo, pode-se adotar uma distribuição a priori não-informativa (BOX & TIAO, 1973). Neste trabalho propomos distribuição a priori não-informativa para os parâmetros de processos AR( $p$ )-ARCH( $q$ ) e ARCH( $q$ ) baseada na proposta de Geweke (1989), definida como:

$$\pi(\boldsymbol{\phi}) \propto \begin{cases} \left( \frac{1 - \sum_{j=1}^q \alpha_j}{\alpha_0} \right) & , \alpha_0 > 0, \alpha_j \geq 0, j = 1, \dots, q \\ 0 & , \text{caso contrário} \end{cases} \quad (19)$$

Condicional a  $\alpha_j = 0, j = 1, \dots, q$ , esta é a distribuição a priori invariante de Jeffreys para o modelo de regressão linear normal. Consideramos ainda uma reparametrização definida por (OLIVEIRA, 2005):

$$\phi_j = \log\left(\frac{\alpha_j - a_j}{b_j - \alpha_j}\right) \quad j = 0, 1, 2, \dots, q \quad (20)$$

Os valores de  $a_j$  e  $b_j$  podem ser escolhidos com base em alguma informação obtida por meio de estudos anteriores sobre a série analisada. Esta reparametrização conduz à escolha de uma transformação que mapeia os intervalos  $(-\infty, +\infty)$  no domínio  $(a_j, b_j)$  e vice-versa. Assim, pode-se reduzir a taxa de rejeição do algoritmo de simulação MCMC, acelerando o processo de convergência deste. Usando a reparametrização definida em (20), tem-se:

$$\pi(\boldsymbol{\phi}) \propto \begin{cases} \left(\frac{1}{k(\boldsymbol{\phi})}\right)^{\frac{1}{2}}, & -\infty < \phi_j < \infty, \quad j = 0, 1, \dots, q \\ 0 & , \text{ caso contrário} \end{cases} \quad (21)$$

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

De forma similar, uma distribuição a priori para os parâmetros de processos GARCH( $q, r$ ), baseada na proposta de Geweke (1989) e na reparametrização  $\phi_j = \log\left(\frac{\theta_j - a_j}{b_j - \theta_j}\right)$ , sendo  $\theta_j = \alpha_j$  ou  $\lambda_j, j = 0, 1, \dots, q + r$ , é dada por:

$$\pi(\boldsymbol{\phi}) \propto \begin{cases} \left(\frac{1}{k(\boldsymbol{\phi})}\right)^{\frac{1}{2}}, & -\infty < \phi_j < \infty, \quad j = 0, 1, \dots, q + r \\ 0 & , \text{ caso contrário} \end{cases} \quad (22)$$

onde  $\boldsymbol{\phi} = (\phi_0, \phi_1, \dots, \phi_{q+r})'$  e  $k(\boldsymbol{\phi}) = \alpha_0 / \left(1 - \sum_{j=1}^q \alpha_j + \sum_{i=1}^r \lambda_j\right)$ .

Assim, escrevendo a volatilidade  $h_t$  em função dos parâmetros dos modelos, devidamente transformados em  $\phi_j$ , obtivemos as distribuições conjuntas a posteriori para  $\boldsymbol{\phi}$ , as quais estão descritas na Tabela 1 a seguir.

Deste modo, usando o algoritmo de simulação Monte Carlo em Cadeias de Markov (MCMC), podemos obter amostras representativas das distribuições a posteriori descritas na Tabela 10, e recuperar os parâmetros estimados por meio de transformações inversas.

**Tabela 1:** Distribuições conjuntas a posteriori.

<b>Modelo AR(p)-ARCH(q)</b>	
Distribuição normal para $\varepsilon_t$ (M1)	$\pi(\boldsymbol{\varphi}   \mathbf{Y}) \propto \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{1/2} \times \prod_{t=p+q+1}^T \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \exp\left\{-\frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{2h_t(\boldsymbol{\varphi})}\right\}$
Distribuição $t$ de student para $\varepsilon_t$ (M2)	$\pi(\boldsymbol{\varphi}   \mathbf{Y}) \propto \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{1/2} \times \prod_{t=p+q+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}} \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \left(1 + \frac{(y_t - \mathbf{x}_t \boldsymbol{\beta})^2}{h_t(\boldsymbol{\varphi})(v-2)}\right)^{-\frac{(v+1)}{2}}$
<b>Modelo ARCH(q)</b>	
Distribuição normal para $\varepsilon_t$ (M3)	$\pi(\boldsymbol{\varphi}   \mathbf{Z}) \propto \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{1/2} \times \prod_{t=q+1}^T \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t(\boldsymbol{\varphi})}\right\}$
Distribuição $t$ de student para $\varepsilon_t$ (M4)	$\pi(\boldsymbol{\varphi}   \mathbf{Z}) \propto \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{1/2} \times \prod_{t=q+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}} \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \left(1 + \frac{z_t^2}{h_t(\boldsymbol{\varphi})(v-2)}\right)^{-\frac{(v+1)}{2}}$
<b>Modelo GARCH(q,r)</b>	
Distribuição normal para $\varepsilon_t$ (M5)	$\pi(\boldsymbol{\varphi}   \mathbf{Z}) \propto \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{1/2} \times (2\pi)^{-\frac{(T-q-r)}{2}} \prod_{t=q+r+1}^T \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \exp\left\{-\frac{z_t^2}{2h_t(\boldsymbol{\varphi})}\right\}$
Distribuição $t$ de student para $\varepsilon_t$ (M6)	$\pi(\boldsymbol{\varphi}   \mathbf{Z}) \propto \left(\frac{1}{k(\boldsymbol{\varphi})}\right)^{1/2} \times \prod_{t=q+r+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}} \left(\frac{1}{h_t(\boldsymbol{\varphi})}\right)^{1/2} \left(1 + \frac{z_t^2}{h_t(\boldsymbol{\varphi})(v-2)}\right)^{-\frac{(v+1)}{2}}$

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

Quando as distribuições condicionais a posteriori não são facilmente identificadas como possuidoras de uma forma padrão (normal, gama, etc.), que impossibilita a geração direta a partir destas distribuições, usamos o algoritmo Metropolis-Hastings. Esta técnica requer um núcleo, isto é, uma densidade de transição  $q(\gamma, \gamma')$ , que não necessariamente tem probabilidade de equilíbrio  $\pi$ , mas que represente uma regra de passagem que defina uma cadeia. Consideremos também a probabilidade de aceitação  $p(\gamma^{(l-1)}, \gamma')$  definida abaixo. O algoritmo segue os seguintes passos:

**Passo 1:** Atribuir um valor inicial  $\gamma = \gamma^{(0)}$  e iniciar o contador de iterações em  $l = 1$ .

**Passo 2:** Mover a cadeia para um novo valor de  $\gamma'$  gerado da densidade  $q(\gamma^{(l-1)}, \gamma')$ .

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



**Passo 4:** Aceitar o valor gerado  $\gamma'$  se

$$u \leq p(\gamma^{(l-1)}, \gamma') = \min\left(1, \frac{\pi(\gamma')q(\gamma', \gamma^{(l-1)})}{\pi(\gamma^{(l-1)})q(\gamma^{(l-1)}, \gamma')}\right)$$

caso contrário, ficar com  $\gamma^{(l)} = \gamma^{(l-1)}$ .

**Passo 5:** Incrementar  $l$  e voltar ao passo 2.

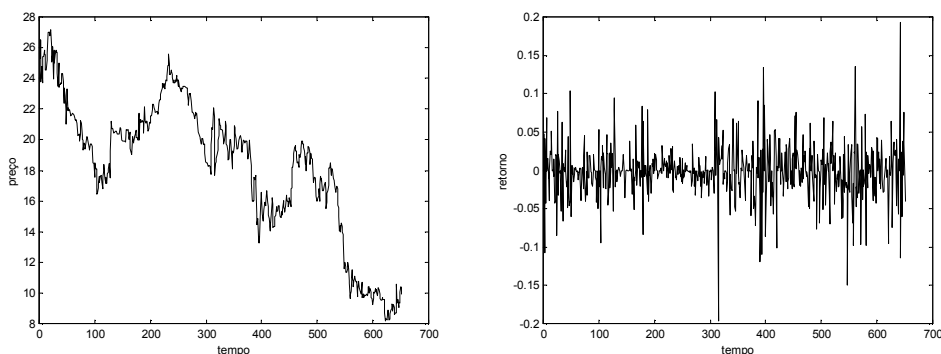
Assim, para  $l$  suficientemente grande,  $\gamma^{(l)}$  é uma amostra da distribuição a posteriori  $\pi(\gamma)$ . Para o caso vetorial  $\gamma = [\gamma_0 \ \gamma_1 \ \dots \ \gamma_q]$ , teremos uma densidade de transição dada por  $q(\gamma, \gamma')$  e uma probabilidade de aceitação dada por  $p(\gamma^{(l-1)}, \gamma')$ , e devemos proceder da mesma forma.

Existem vários critérios que podem ser usados para a seleção de modelos, sob um contexto Bayesiano. Neste trabalho usamos o critério baseado na densidade preditiva ordenada (CPO), construído a partir da distribuição de  $y_{T+m}$ , condicionada aos dados  $\mathbf{Y}$ , e aos parâmetros de cada modelo considerado. O CPO consiste em escolher o modelo  $l$  que apresentar o *maior* valor de  $\hat{c}(l)$  (quantidade obtida por meio da estimativa de Monte Carlo da densidade preditiva). Maiores detalhes podem ser encontrados em Carlin & Chib (1995).

#### 4. Aplicações

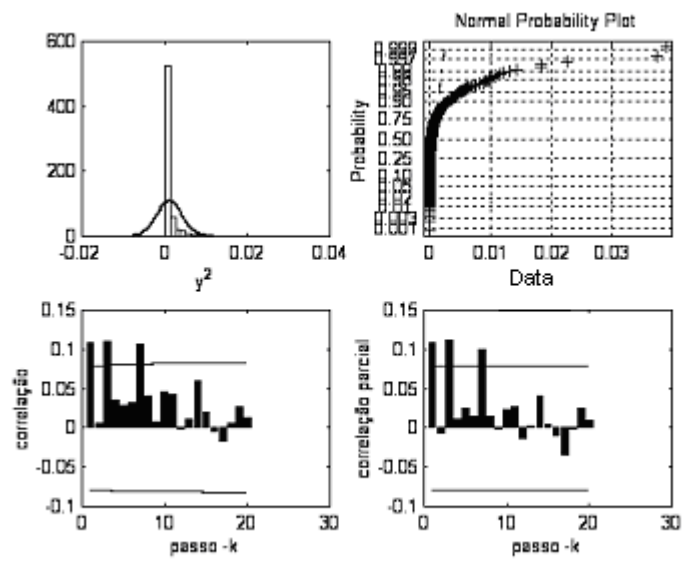
Apresentamos a seguir os resultados obtidos na implementação dos modelos propostos para a inferência dos parâmetros de processos AR( $p$ )-ARCH( $q$ ), ARCH( $q$ ) e GARCH( $q,r$ ), considerando distribuição normal e  $t$  de student para  $\varepsilon_t$ .

A série histórica analisada apresenta informações sobre índices de fechamento da Bovespa (IBovespa) registrados no período de 02/01/1996 a 01/02/1999, totalizando 986 informações. Seja  $p_t$  o preço de fechamento do índice Bovespa em um dia de negócio. Denotando  $P_t = \ln p_t$ , temos o retorno dado por  $y_t = \ln(p_t/p_{t-1}) = P_t - P_{t-1}$ . Os índices da Bovespa e os respectivos retornos estão ilustrados na Figura 1 a seguir.



**Figura 1: Preço e Retorno - Série IBOvespa.**

Os gráficos da Figura 2 mostram ainda o comportamento da série dos quadrados dos retornos do IBovespa,  $y_t^2$ . Podemos observar que esta série é correlacionada. Tal comportamento é tipicamente associado aos modelos da família ARCH.



**Figura 2: Histograma, Gráfico Normal Probabilístico, Autocorrelação e Autocorrelação parcial de  $y_t^2$  - Série IBovespa.**

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

Assim, os modelos ajustados à série IBovespa, de acordo com o critério da CPO, foram: ARCH(3), AR(3)-ARCH(6) e GARCH(1,1), respectivamente.

As Tabelas 2, 3 e 4 mostram as estimativas Bayesianas para os modelos ajustados, sendo M1 e M2 os processos ARCH, M3 e M4 os processos AR-ARCH, e M5 e M6 os processos GARCH, com distribuição normal e  $t$  de *student* para  $\varepsilon_t$ , respectivamente.

**Tabela 2: Estimativas Bayesianas dos modelos ARCH – M1 e M2.**

Parâmetro	M1			M2		
	Média	Desvio Padrão	Intervalo Credibilidade (95%)	Média	Desvio Padrão	Intervalo Credibilidade (95%)
$\alpha_0$	0,00061	0,00005	0,00052;0,00070	0,00070	0,00003	0,00060;0,00071
$\alpha_1$	0,16350	0,01113	0,14187;0,18514	0,21672	0,00591	0,20524;0,22858
$\alpha_2$	0,06065	0,01139	0,03885;0,08257	0,05228	0,00205	0,04818;0,05624
$\alpha_3$	0,33162	0,02134	0,28888;0,37152	0,34961	0,00926	0,33156;0,36665

**Tabela 3:** Estimativas Bayesianas dos modelos AR-ARCH – M3 e M4.

Parâmetro	M3			M4		
	Média	Desvio Padrão	Intervalo Credibilidade (95%)	Média	Desvio Padrão	Intervalo Credibilidade (95%)
$\alpha_0$	0,00048	0,00003	0,00042;0,00056	0,00050	0,00003	0,00044;0,00056
$\alpha_1$	0,21579	0,04142	0,13543;0,29508	0,18757	0,00567	0,17646;0,19868
$\alpha_2$	0,08565	0,02706	0,03649;0,14026	0,06340	0,00331	0,05691;0,06988
$\alpha_3$	0,41482	0,04587	0,32397;0,49956	0,38771	0,00900	0,37007;0,40535
$\beta_1$	-0,09571	0,02879	-0,15259;-0,03821	-0,10020	0,00577	-0,11151;-0,08889
$\beta_2$	-0,01679	0,02853	-0,07421;0,03983	-0,02365	0,00542	-0,03427;-0,01303
$\beta_3$	-0,13483	0,02876	-0,19079;-0,07799	-0,12111	0,00531	-0,13152;-0,11070
$\beta_4$	-0,16456	0,02804	-0,21835;-0,10972	-0,15437	0,00522	-0,16460;-0,14414
$\beta_5$	-0,14604	0,02958	-0,20397;-0,08949	-0,13210	0,00689	-0,14560;-0,11860
$\beta_6$	-0,08765	0,02993	-0,14678;-0,02814	-0,07353	0,00695	-0,08715;-0,05991

**Tabela 4:** Estimativas Bayesianas dos modelos GARCH – M5e M6.

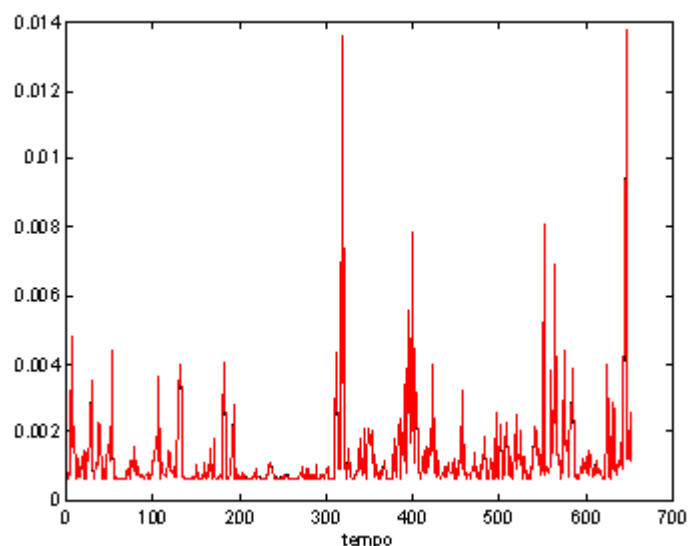
Parâmetro	M5			M6		
	Média	Desvio Padrão	Intervalo Credibilidade (95%)	Média	Desvio Padrão	Intervalo Credibilidade (95%)
$\alpha_0$	0,00007	1,5e-009	0,000069;0,000071	0,00007	1,4e-009	0,000069;0,000071
$\alpha_1$	0,15010	6,7e-004	0,14870;0,15140	0,15035	6,5e-004	0,14920;0,15100
$\beta_1$	0,75050	1,3e-003	0,74790;0,75290	0,75027	1,2e-003	0,74850;0,75250

De acordo com os valores obtidos pelo CPO pelo ajuste dos modelos propostos, concluímos que o modelo que melhor se adequou aos dados da série IBovespa foi o processo GARCH(1,1) com distribuição  $t$  de student para  $\varepsilon_t$ , ou seja, **o modelo M6**, conforme Tabela 5 a seguir.

**Tabela 5: Critério de seleção de modelos.**

Modelo	CPO
M1	1,163e+018
M2	1,190e+018
M3	1,155e+019
M4	1,264e+019
M5	1,173e+019
<b>M6</b>	<b>1,859e+019</b>

A Figura 3 mostra a volatilidade estimada a partir das estimativas obtidas pela abordagem Bayesiana com distribuição a priori não informativa, por meio do processo GARCH(1,1) com distribuição  $t$  de student para  $\varepsilon_t$ . Podemos observar que, no período estudado, houve alguns fatos que provocaram alterações (“picos”) no comportamento da volatilidade, os quais o modelo ajustado identificou: Moratória da Rússia, em Agosto de 1998; Crise na Ásia, em Outubro de 1998; e a Desvalorização do Real, em Janeiro de 1999.



**Figura 3: Volatilidade Estimada - Série IBovespa: Modelo M6 .**

## 5. Considerações Finais

Apesar de ser uma representação bastante simplificada do processo gerador de dados dos retornos condicionais, a distribuição normal é amplamente utilizada na estimação de modelos de volatilidade. Porém, em algumas situações é mais apropriado assumir que  $\varepsilon_t$  tem distribuição t de student, como pôde ser observado neste caso.

Os resultados deste trabalho mostraram que, de um modo geral, a abordagem Bayesiana proposta fornece estimativas satisfatórias. Além disso, consideramos uma reparametrização dos modelos estudados e adotamos distribuição a priori não-informativa, conseguindo uma convergência mais rápida do processo de inferência de parâmetros de modelos da família ARCH com métodos MCMC.

## Referências

BOLLERSLEV, T. Glossary to ARCH (GARCH). Research paper, School of Economics and Management, University of Aarhus, 2008. Eletronic copy available at: <http://ssrn.com/abstract=1263250>.

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

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

CARLIN, B.; CHIB, S. Bayesian Model Choice via Markov Chain Monte Carlo Methods. *Journal of the Royal Statistical Society*, Vol. 57, pp. 473-484, 1995.

CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, Vol. 49, N° 4, pp. 327-335, 1995.

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

GELFAND, A. E.; SMITH, A. F. Sampling – Based Approaches to Calculating Marginal Densities. *Journal of the Statistical Association*, Vol. 5, N° 410, pp. 398-409, 1990.

GEWEKE, J. Exact Predictive Densities for Linear Models with ARCH Disturbances. *Journal of Econometrics*, Vol. 40, pp. 63-86, 1989.

GEWEKE, J. Evaluating the Accuracy of sampling-based approaches to calculating posterior moments (with discussion). *Bayesian Statistics*, Vol. 4, Oxford University Press, Oxford, pp. 169-193, 1992.

MORETTIN, P. A. *Econometria Financeira: Um curso em Séries Temporais Financeiras*. Edgard Blücher Ltda, São Paulo, 519 p., 2008.

OLIVEIRA, S. C. *Modelos Estocásticos com Heterocedasticidade para Séries Temporais em Finanças*. Tese de Doutorado, ICMC-USP, 2005.

SÁFADI, T., ANDRADE, M. G. (2007). *Abordagem Bayesiana de Modelos de Séries Temporais*. Minicurso, Universidade Federal do Rio Grande do Sul, 12º Escola de Séries Temporais e Econometria.