Modelos GARCH com Distribuições Não Gaussianas

Rogério Oliveira Ribeiro

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU
DE
MESTRE EM ESTATÍSTICA

Área de Concentração : **Estatística**
Orientadora : **Mariane Streibel**

São Paulo, fevereiro de 2000

Modelos GARCH com Distribuições Não Gaussianas

Este exemplar corresponde à redação final da dissertação de mestrado devidamente corrigida e defendida por Rogério Oliveira Ribeiro e aprovada pela comissão julgadora.

São Paulo, 07 de abril de 2000

Banca Examinadora

- Prof.^a. Dra. Mariane Streibel (Orientadora) – IME – USP
- Prof. Dr. Luiz Koodi Hotta – IMECC-UNICAMP
- Prof. Dr. Luiz Valls Pereira – IME - USP

Agradecimentos

Agradeço aos meus familiares pelo constante apoio, em especial à minha mãe e ao meu irmão.

Agradeço à Prof. Mariane Striebel pela orientação.

Agradeço ao Prof. Pedro Valls pelas instrutivas conversas de corredor.

Resumo

A ênfase do trabalho está na modelagem de série financeiras. O presente trabalho investiga, em modelos GARCH, como a suposição de distribuição dos erros padronizados pode afetar a previsão por intervalo da série em estudo. Simulações foram realizadas com o intuito de comparar os modelos GARCH com erros padronizados seguindo distribuição normal, t-student e normal contaminada. Uma série real foi estudada utilizando três diferentes suposições sobre a distribuição dos erros padronizados. A conclusão geral é que existe considerável benefício na previsão por intervalo quando são utilizadas distribuições com excesso de curtose quando comparadas à distribuição normal.

Abstract

The emphasis of this dissertation is financial time series models. This dissertation investigates, in GARCH models, how the assumption about standard error distribution affects the time series prediction by interval. Simulations were carried out in order to compare GARCH models with standard errors following normal distribution, t-student distribution and contaminated normal distribution. A real time series was analyzed assuming three different distribution for standard errors. The overall conclusion is that there can be considerable benefits in predictions by interval using distributions with excess of curtose with respect to gaussian distribution.

Sumário

0.1	Introdução	4
1	MODELOS ARCH, GARCH E ALGUMAS VARIAÇÕES	6
1.1	Introdução	6
1.2	Modelo ARCH	7
1.2.1	Estimativas de Máxima Verossimilhança Condicionais	9
1.3	Modelo GARCH	10
1.3.1	Estimativas de Máxima Verossimilhança Condicionais	12
1.4	Modelos IGARCH, EGARCH e ARCH - M	14
1.4.1	IGARCH	14
1.4.2	EGARCH	14
1.4.3	ARCH-M	15
2	MODELO GARCH COM DISTRIBUIÇÃO T-STUDENT	16
2.1	Introdução	16
2.2	GARCH com distribuição T- Student	17
3	MODELOS ARCH E GARCH COM MUDANÇAS DE REGIME	19
3.1	Introdução	19
3.2	Mistura de Distribuições IID	21
3.2.1	Estimativas de Máxima Verossimilhança e Algoritmo EM	22

3.3	Modelos com Mudanças de Regimes Markovianas	23
3.3.1	Probabilidade Filtrada	24
3.3.2	Probabilidade Suavizada	26
3.3.3	Estimativas de Máxima Verossimilhança Condicionais e Algoritmo EM	26
3.4	Modelos Heterocedásticos com Mudanças de Regimes Markovianas	28
3.4.1	Modelos SWARCH	29
3.4.2	SWGARCH proposto por Dueker	33
3.5	Função de verossimilhança	35
4	MODELO GARCH COM NORMAL CONTAMINADA	36
4.1	Introdução	36
4.2	Modelo GARCH com Normal Contaminada (NC)	37
4.2.1	Distribuição Normal Contaminada	37
4.2.2	Modelo GARCH com Normal Contaminada	38
4.3	Testes de Hipóteses	40
5	SIMULAÇÃO COM ALGUNS MODELOS GARCH	44
5.1	Introdução	44
5.2	Gerando os dados	45
5.3	Estimando os Parâmetros e Analisando os Resultados	49
5.3.1	Estimando as séries com erros normais	50
5.3.2	Estimando com erros t-student	52
5.3.3	Estimando com erros normais contaminados (normal cont I)	55
5.3.4	Estimando com erros normais contaminados (normal cont II)	58
5.4	Comentários finais	61
6	APLICAÇÃO EM UMA SÉRIE REAL	62
6.1	Introdução	62
6.2	Dados	62
6.3	Análise dos Dados	65

6.3.1	Modelo GARCH	66
6.3.2	Modelo GARCH-T	70
6.3.3	Modelo GARCH-NC	74
6.4	Comparação da Volatilidades Estimadas	78
7	Comentários Finais	80
A	Tabelas de Decis	82
A.1	Erros seguindo distribuição normal	82
A.2	Erros seguindo distribuição t - student	84
A.3	Erros seguindo distribuição normal contaminada I	86
A.4	Erros seguindo distribuição normal contaminada II	88
	Bibliografia	90

0.1 Introdução

Até o final dos anos 70 perdurou a idéia de que as séries financeiras pudessem ser modeladas através de processos lineares. Porém, quando algumas transformações eram aplicadas, a presença de correlação serial significativa indicava que a suposição de independência não era verdadeira. Em virtude deste fato, surge a preocupação em desenvolver modelos não lineares que expliquem de forma mais apropriada o comportamento de séries financeiras.

No início dos anos 80 a idéia do uso de modelos não lineares, para explicar o comportamento de séries financeiras, começa a ganhar força. A grande maioria dos modelos não lineares pode ser descrito por:

$$y_t = g(y_{t-1}; y_{t-2}; \dots) + h(y_{t-1}; y_{t-2}; \dots) \cdot v_t$$

onde $g(y_{t-1}; y_{t-2}; \dots)$ representa a esperança de y_t condicionada à informação passada;
 $h(y_{t-1}; y_{t-2}; \dots)$ representa a variância de y_t condicionada à informação passada;
 $E(v_t) = 0$ e $Var(v_t) = 1$.

Modelos nos quais $g(\bullet)$ é uma função não linear são chamados não lineares na média e, analogamente, modelos em que $h(\bullet)$ é uma função não linear são chamados modelos não lineares na variância.

Um exemplo de modelos não lineares na variância é o modelo ARCH proposto por Engle em 1982 onde $h(\bullet) = \sqrt{\delta_0 + \sum_{i=1}^p \varepsilon_{t-i}^2 \delta_i}$ com $\varepsilon_t = h(\bullet)v_t$.

Neste trabalho, iremos apresentar alguns modelos não lineares na variância e considerar, basicamente, dois objetivos:

1. Fazer um resumo dos diversos modelos não lineares na variância pertencentes a classe dos modelos GARCH;
2. Apresentar uma nova variação do modelo GARCH com suposição dos erros padronizados seguindo a distribuição proveniente de uma mistura de gaussianas e compará-la com outros já existentes no que se refere a previsão por intervalo da série em estudo.

Para contemplar o primeiro objetivo apresentamos o modelo ARCH proposto por Engle (1982) no capítulo 1. Ainda no capítulo 1, iremos mostrar a generalização do modelo ARCH para o modelo GARCH proposto por Bollerslev em 1986 e, para finalizar este capítulo, fizemos uma breve apresentação dos modelo IGARCH proposto por Engle e Bollerslev (1986), EGARCH proposto por Nelson (1991) e ARCH - M proposto por Engle, Lilian e Robins (1987).

Preocupado com o fato das séries financeiras, em geral, apresentarem excesso de curtose quando comparadas à distribuição normal, Bollerslev (1987) propôs o modelo GARCH - T e este é apresentado no capítulo 2.

Apesar dos modelos SWARCH e SWGARCH não serem parte integrante das simulações e resultados finais, em virtude do primeiro objetivo, houve por bem apresentá-los no capítulo 3. Neste capítulo consideramos as propostas apresentadas por Cai (1994), Hamilton e Susmel (1994) e Dueker (1997).

Passamos a contemplar o segundo objetivo a partir do capítulo 4, onde apresentamos uma proposta de modelo GARCH com os erros provenientes da mistura de distribuições normais (normal contaminada). Iremos nos referir a este modelo com GARCH - NC.

No capítulo 5 apresentamos o resultado de algumas simulações englobando os modelos GARCH, GARCH - T e GARCH - NC. Ainda neste capítulo fizemos a comparação dos três modelos no que se refere à previsão por intervalo da série em estudo.

No capítulo 6, aplicamos os três modelos utilizados no capítulo 5 aos retornos diários da ação ITAÚ PN e comparamos os resultados obtidos.

No final do trabalho, breves conclusões são apresentadas.

Capítulo 1

MODELOS ARCH, GARCH E ALGUMAS VARIAÇÕES

1.1 Introdução

Neste capítulo, iremos apresentar uma breve revisão sobre o modelo ARCH introduzidos por Engle (1982) e o modelo GARCH apresentado por Bollerslev (1986).

Embora este trabalho esteja mais fortemente relacionado com os modelos acima mencionados também iremos apresentar, brevemente, algumas variações dos modelos GARCH tais como o IGARCH apresentado por Engle e Bollerslev (1986), o modelo EGARCH introduzido por Nelson (1991) e o ARCH-M apresentado por Engle, Lilien e Robins (1987).

Uma boa descrição dos modelos acima mencionados também podem ser encontrada em Hamilton (1994) e Palm (1996).

Nas seções que seguem iremos descrever os modelos anteriormente citados.

1.2 Modelo ARCH

Seja $\{y_t\}$ uma série temporal qualquer.

O modelo ARCH(p) é dado por:

$$y_t = x_t' \beta + \varepsilon_t \quad (1.1)$$

$$\varepsilon_t = \sqrt{h_t} v_t \quad (1.2)$$

$$h_t = \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \delta_2 \varepsilon_{t-2}^2 + \dots + \delta_p \varepsilon_{t-p}^2 \quad (1.3)$$

onde x_t' é o vetor de variáveis explicativas podendo conter lag's da série temporal $\{y_t\}$;

β é um vetor de parâmetros, relacionados ao nível da série temporal $\{y_t\}$, a serem estimados;

δ_i , com $i = 0, 1, \dots, p$, são os parâmetros, relacionados à variância da série temporal $\{y_t\}$, a serem estimados;

v_t é seqüência de variáveis aleatórias iid com $E(v_t) = 0$ e $E(v_t^2) = 1$.

Se somarmos ε_t^2 em 1.3 de ambos os lados temos que:

$$\varepsilon_t^2 + h_t = \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \delta_2 \varepsilon_{t-2}^2 + \dots + \delta_p \varepsilon_{t-p}^2 + \varepsilon_t^2 \quad (1.4)$$

$$\varepsilon_t^2 = \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \delta_2 \varepsilon_{t-2}^2 + \dots + \delta_p \varepsilon_{t-p}^2 + z_t \quad (1.5)$$

$$\text{onde } z_t = \varepsilon_t^2 - h_t.$$

Lembremos que h_t é a previsão para ε_t^2 baseados nos lag's de ε_t^2 . Logo, temos que z_t é o erro de previsão associado e é ruído branco.

Desta forma temos que a esperança do erro quadrático de previsão de $\{y_t\}$ é dado por:

$$E\left(\varepsilon_t^2 \mid \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots\right) = \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \delta_2 \varepsilon_{t-2}^2 + \dots + \delta_p \varepsilon_{t-p}^2 \quad (1.6)$$

Para assegurar que 1.5 é não negativa e 1.6 é positiva é suficiente que $\delta_i > 0$, com $i = 0, 1, \dots, p$, e $z_t > -\delta_0$ para qualquer t .

Como podemos notar, através de 1.5, ε_t^2 segue um processo AR(p) e portanto, para que seja um processo estacionário é necessário que as raízes da equação $1 - \delta_1 x - \delta_2 x^2 - \dots - \delta_p x^p = 0$ estejam fora do círculo unitário.

Ter as raízes da equação característica, acima mencionada, fora do círculo unitário e $\delta_i > 0$, para $i = 1, \dots, p$ é equivalente a dizer que:

$$\sum_{j=1}^p \delta_j < 1 \quad (1.7)$$

Quando estas condições estão satisfeitas temos que a variância incondicional de ε_t é dada por:

$$\sigma^2 = E\left(\varepsilon_t^2\right) = \frac{\delta_0}{1 - \delta_1 - \delta_2 - \dots - \delta_p} \quad (1.8)$$

1.2.1 Estimativas de Máxima Verossimilhança Condicionais

Suponhamos que estejamos interessados em estimar os parâmetros $\psi = (\beta; \delta)$ onde $\delta = (\delta_0; \delta_1; \dots; \delta_p)$ baseado em y_t com $t = 1, \dots, T$.

Seja $\mathfrak{S}_t$ o vetor de todas as observações até a data t , ou seja:

$$\mathfrak{S}_t = (y_t; y_{t-1}; \dots; y_1; x'_t; x'_{t-1}; \dots x'_1)'$$

Para que seja possível escrever a verossimilhança é necessário que se faça uma suposição sobre a distribuição de v_t , portanto, suponhamos também que $v_t \sim N(0; 1)$.

Desta forma, podemos escrever a função densidade condicional de y_t dado o vetor $\mathfrak{S}_{t-1}$ e x'_t da forma que segue:

$$f(y_t | \mathfrak{S}_{t-1}; x'_t) = \frac{1}{\sqrt{2\pi h_t}} \exp\left(\frac{-\varepsilon_t^2}{2h_t}\right) \quad (1.9)$$

onde h_t é dado pela equação 1.3 e ε_t da equação 1.1.

O Log da função de verossimilhança, condicionada nas primeiras p observações é dado por:

$$\zeta(\psi) = \sum_{j=p+1}^T \ln(f(y_j | \mathfrak{S}_{j-1}; x'_j)) = -\frac{T-p}{2} \ln(2\pi) - \frac{1}{2} \sum_{j=p+1}^T \ln(h_j) - \frac{1}{2} \sum_{j=p+1}^T \frac{\varepsilon_j^2}{h_j} \quad (1.10)$$

As estimativas dos parâmetros são encontradas através de maximização numérica da função de verossimilhança.

1.3 Modelo GARCH

O modelo proposto por Bollerslev (1986), é uma generalização do modelo ARCH anteriormente apresentado. Imaginemos que no processo dado pelas equações 1.1, 1.2 e 1.3 a variância condicional dependesse de um número infinito de lag's de ε_t^2 , ou seja, o novo processo seria dado por:

$$y_t = x_t' \beta + \varepsilon_t \quad (1.11)$$

$$\varepsilon_t = \sqrt{h_t} v_t \quad (1.12)$$

$$h_t = \delta_0 + \pi(L) \varepsilon_t^2 \quad (1.13)$$

$$\pi(L) = \sum_{j=1}^{\infty} \pi_j L^j \quad (1.14)$$

É bastante razoável pensar em $\pi(L)$ como uma razão de dois polinômios de ordem finita, ou seja, que $\pi(L)$ pode ser escrito da forma:

$$\pi(L) = \frac{\alpha(L)}{1 - \phi(L)} = \frac{\sum_{j=1}^p \alpha_j L^j}{1 - \sum_{j=1}^q \phi_j L^j} \quad (1.15)$$

onde supomos que as raízes da equação $1 - \phi(L) = 0$ estão fora do círculo unitário.

Desta forma, podemos multiplicar a equação 1.13 por $1 - \phi(L)$ de ambos os lados e iremos obter que:

$$(1 - \phi(L)) h_t = (1 - \phi(L)) \delta_0 + (1 - \phi(L)) \pi(L) \varepsilon_t^2 \quad (1.16)$$

$$h_t = \gamma + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \phi_j h_{t-j} \quad (1.17)$$

$$\text{onde } \gamma = \left(1 - \sum_{j=1}^q \phi_j\right) \delta_0$$

É fácil notar que o processo ε_t^2 segue um ARMA(d,q) onde $d = \max(p; q)$. Se somarmos ε_t^2 em ambos os lados da equação 1.17 podemos reescreve-la da forma abaixo:

$$h_t + \varepsilon_t^2 = \gamma - \sum_{j=1}^q \phi_j (\varepsilon_{t-j}^2 - h_{t-j}) + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \phi_j \varepsilon_{t-j}^2 + \varepsilon_t^2 \quad (1.18)$$

que implica em:

$$\varepsilon_t^2 = \gamma - \sum_{j=1}^q \phi_j z_{t-j} + \sum_{j=1}^d (\phi_j + \alpha_j) \varepsilon_{t-j}^2 + z_t \quad (1.19)$$

onde $z_t = \varepsilon_t^2 - h_t$, $d = \max(p; q)$, $\phi_j = 0$ para $j > q$ e $\alpha_j = 0$ para $j > p$.

Lembremos que h_t é a previsão para ε_t^2 baseados nos lag's de ε_t^2 . Logo, temos que z_t é o erro de previsão associado e é ruído branco.

Para que ε_t^2 seja não negativo é necessário que $\gamma > 0$, $\phi_j \geq 0$ e $\alpha_j \geq 0$ para $j = 1, \dots, d$.

Para que o processo ε_t^2 seja um processo estacionário é necessário que as raízes da equação $1 - (\alpha_1 + \phi_1)x - (\alpha_2 + \phi_2)x^2 - \dots - (\alpha_p + \phi_p)x^p = 0$ estejam fora do círculo unitário.

Ter as raízes da equação acima mencionada, fora do círculo unitário e $\gamma > 0$, $\phi_j \geq 0$ e $\alpha_j \geq 0$ com $j = 1, \dots, d$ é equivalente a dizer que:

$$\sum_{j=1}^d (\alpha_j + \phi_j) < 1 \quad (1.20)$$

Se as condições de estacionaridade estiverem satisfeitas, a esperança incondicional de ε_t^2 é dada por:

$$\sigma^2 = E(\varepsilon_t^2) = \frac{\gamma}{1 - \sum_{j=1}^d (\alpha_j + \phi_j)} \quad (1.21)$$

Finalmente, podemos rescrever o modelo GARCH(p;q), dado pelas equações 1.11, 1.12 e 1.17 da forma que segue:

$$\begin{aligned} y_t &= x_t' \beta + \varepsilon_t \\ \varepsilon_t &= \sqrt{h_t} v_t \\ h_t &= \gamma + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \phi_j h_{t-j} \end{aligned}$$

1.3.1 Estimativas de Máxima Verossimilhança Condicionais

Suponhamos que estejamos interessados em estimar os parâmetros do vetor $\psi = (\beta; \alpha; \phi; \gamma)$ onde $\alpha = (\alpha_1; \alpha_2; \dots; \alpha_p)$ e $\phi = (\phi_1; \phi_2; \dots; \phi_q)$ baseado em y_t com $t = 1, \dots, T$.

Seja $\mathfrak{S}_t$ o vetor de todas as observações até a data t , ou seja:

$$\mathfrak{S}_t = (y_t; y_{t-1}; \dots; y_1; x_t'; x_{t-1}'; \dots x_1')'$$

Para que seja possível escrever a verossimilhança é necessário que se faça uma suposição sobre a distribuição de v_t , portanto, suponhamos também que $v_t \sim N(0; 1)$.

Desta forma, podemos escrever a função densidade condicional de y_t dado o vetor $\mathfrak{S}_{t-1}$ e x_t' da forma que segue:

$$f(y_t | \mathfrak{S}_{t-1}; x_t') = \frac{1}{\sqrt{2\pi h_t}} \exp\left(\frac{-\varepsilon_t^2}{2h_t}\right) \quad (1.22)$$

onde h_t é dado pela equação 1.17 e ε_t é obtido a partir da equação 1.11.

O log da função de verossimilhança, condicionada nas primeiras p observações é dado por:

$$\zeta(\psi) = \sum_{t=p+1}^T \ln(f(y_t | \mathfrak{S}_{t-1}; x'_t)) = -\frac{T-p}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=p+1}^T \ln(h_t) - \frac{1}{2} \sum_{t=p+1}^T \frac{\varepsilon_t^2}{h_t} \quad (1.23)$$

Para calcular o valor h_{q+1} são necessários os valores h_j , $j = 1, \dots, q$. Uma alternativa é utilizar a variância incondicional dada pela equação 1.21. Bollerslev (1986) sugere que se faça $h_j = \widehat{\sigma}^2$, $j = 1, \dots, q$ onde $\widehat{\sigma}^2$ é dado abaixo.

$$\widehat{\sigma}^2 = \frac{1}{T} \sum_{j=1}^T (y_j - x'_j \beta)^2 \quad (1.24)$$

As estimativas dos parâmetros são encontradas através de maximização numérica da função de verossimilhança.

1.4 Modelos IGARCH, EGARCH e ARCH - M

Apenas com o intuito de complementar a apresentação deste capítulo, iremos fazer uma breve menção sobre algumas variações existentes dos modelos ARCH e GARCH a seguir.

1.4.1 IGARCH

O modelo IGARCH é assim chamado por Engle e Bollerslev (1986) quando tem-se o modelo GARCH dado pelas equações 1.11, 1.12 e 1.17 com:

$$\sum_{j=1}^p (\alpha_j + \phi_j) = 1 \quad (1.25)$$

Desta forma tem-se que a variância incondicional de y_j não é definida e portanto ε_t^2 não satisfaz a definição de estacionaridade.

1.4.2 EGARCH

Uma outra variação dos modelos GARCH é a proposta realizada por Nelson (1991) onde o mesmo é dado pelas equações 1.11, 1.12, porém a variância condicional é modelada da forma que segue:

$$\ln(h_t) = \gamma + \sum_{j=1}^{\infty} \pi_j \{|\varepsilon_{t-j}| - E|\varepsilon_{t-j}| + \theta \varepsilon_{t-j}\} \quad (1.26)$$

Uma clara característica deste modelo é o fato que não existe a necessidade de restrições nos parâmetros para que h_t seja positivo, pois mesmo que $\ln(h_t)$ seja negativo temos que $h_t > 0$.

Outra característica interessante do modelo EGARCH (ou exponencial GARCH) é que o mesmo permite que surpresas negativas ($v_{t-j} < 0$) afetem a variância condicional de forma diferente das surpresas positivas ($v_{t-j} > 0$) quando $\theta \neq 0$. Por razão de facilidade de raciocínio, suponhamos que $\pi_j > 0$ para qualquer j .

Se tivermos $-1 < \theta < 0$ ($0 < \theta < 1$) teremos que surpresas negativas (positivas) irão provocar um aumento na variância condicional maior que surpresas positivas (negativas). Por outro lado, se tivermos que $\theta < -1$ ($\theta > 1$) teremos que surpresas negativas (positivas) aumentam a variância condicional enquanto surpresas positivas (negativas) reduzem a mesma.

Acredita-se que, em séries financeiras, uma surpresa negativa ($v_t < 0$) provoque um aumento da variância condicional maior que uma surpresa positiva ($v_t > 0$), portanto espera-se que $\theta < 0$.

Assim como no modelo GARCH, é razoável pensar em $\pi(L)$ como a razão de dois polinômios finitos, e portanto, podemos rescrever 1.26 como:

$$\ln(h_t) = \gamma + \sum_{j=1}^p \phi_j \ln(h_{t-j}) + \sum_{j=1}^q \alpha_j \{|v_{t-j}| - E|v_{t-j}| + \theta v_{t-j}\} \quad (1.27)$$

Nelson (1991) mostra que quando $\sum_{j=1}^{\infty} \pi_j^2 < \infty$ tem-se que $\ln(h_t)$, h_t e ε_j são estritamente estacionários.

1.4.3 ARCH-M

A teoria financeira sugere que maiores retornos (ganhos) está associada à exposição de maiores riscos (maior variância). Sob a luz desta idéia Engle, Lilien e Robins (1987) sugeriram o modelo ARCH-M que é dado pelas equações que seguem abaixo:

$$y_t = x_t' \beta + \varsigma h_t + \varepsilon_t \quad (1.28)$$

$$\varepsilon_t = \sqrt{h_t} v_t \quad (1.29)$$

$$h_t = \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \delta_2 \varepsilon_{t-2}^2 + \dots + \delta_p \varepsilon_{t-p}^2 \quad (1.30)$$

Desta forma, o modelo sugere que quando houver um aumento do risco (variância condicional) existirá um aumento na esperança condicional de ganho (nível da série) captado, no modelo, pelo parâmetro ς .

Capítulo 2

MODELO GARCH COM DISTRIBUIÇÃO T-STUDENT

2.1 Introdução

No modelo GARCH apresentado no capítulo anterior, a distribuição condicional de $\{y_t\}$ foi descrita como sendo a distribuição gaussiana, porém, algumas séries financeiras possuem distribuições com curtose excessiva quando comparadas à distribuição normal.

Com o intuito de permitir que a distribuição condicional de $\{y_t\}$ possua caudas pesadas Bollerslev (1987) propõe o uso da distribuição T-Student no modelo GARCH.

Um resultado conhecido da distribuição T- Student é que a mesma se aproxima da distribuição gaussiana quando o grau de liberdade da mesma vai para infinito. Devido a este resultado o modelo GARCH com distribuição T- Student, doravante GARCH – T, proposto por Bollerslev (1987) pode ser entendido como uma generalização do modelo GARCH proposto por ele mesmo em 1986.

A possibilidade de entender o modelo GARCH – T como uma generalização do modelo GARCH usual deve-se ao fato de ser possível testar se os graus de liberdade da distribuição utilizada é alto o suficiente para que se possa utilizar a distribuição normal ao invés da T – Student.

Neste capítulo, iremos fazer uma breve apresentação do modelo GARCH – T proposto por Bollerslev (1987).

2.2 GARCH com distribuição T- Student

O modelo GARCH – T é dado por:

$$y_t = x_t' \beta + \varepsilon_t \quad (2.1)$$

$$\varepsilon_t = \sqrt{h_t} \sqrt{\frac{n-2}{n}} v_t \quad (2.2)$$

$$h_t = \gamma + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \phi_j h_{t-j} \quad (2.3)$$

onde v_t tem distribuição T – Student com n graus de liberdade.

Observe que, definido desta forma temos que:

$$E(\varepsilon_t) = 0; \quad (2.4)$$

$$Var(\varepsilon_t) = h_t \frac{n-2}{n} Var(v_t) = h_t \quad (2.5)$$

Podemos escrever a função densidade condicionada de y_t dado o vetor $\mathfrak{S}_{t-1}$ (passado de y_t e x_t') e x_t' é da forma que segue:

$$f(y_t | \mathfrak{S}_{t-1}; x_t') = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right) \sqrt{(n-2)\pi h_t}} \left(1 + \frac{\varepsilon_t^2}{(n-2)h_t}\right)^{-\frac{(n+1)}{2}} \quad (2.6)$$

onde h_t é dado pela equação 2.3 e ε_t pela equação 2.1;

$\Gamma(\bullet)$ é a função Gama.

O logaritmo da função de verossimilhança, condicionada nas primeiras p observações é dado por:

$$\zeta(\psi) = \sum_{t=p+1}^T \ln(f(y_t | \mathcal{S}_{t-1}; x'_t)) \quad (2.7)$$

$$\begin{aligned} \zeta(\psi) = & -\frac{(T-p)}{2} \ln(\pi) - \frac{(T-p)}{2} \ln(n-2) + (T-p) \left(\ln \left(\Gamma \left(\frac{n+1}{2} \right) \right) - \ln \left(\Gamma \left(\frac{n}{2} \right) \right) \right) \\ & - \frac{1}{2} \sum_{t=p+1}^T \ln(h_t) - \frac{(n+1)}{2} \sum_{t=p+1}^T \ln \left(1 + \frac{\varepsilon_t^2}{(n-2)h_t} \right) \end{aligned} \quad (2.8)$$

As estimativas de máxima verossimilhança condicionada do vetor $\psi = (\beta; \alpha; \phi; \gamma; n)$ onde $\alpha = (\alpha_1; \alpha_2; \dots; \alpha_p)$ e $\phi = (\phi_1; \phi_2; \dots; \phi_q)$ baseado em y_t com $t = 1, \dots, T$ são obtidas através de maximização numérica da função de verossimilhança condicionada.

Conforme mencionado anteriormente, é possível construir um teste que permita avaliar se $\frac{1}{n} = 0$. Bollerslev (1987) menciona que o teste de razão de verossimilhança possui distribuição de probabilidades mais concentrada em torno da origem que a χ_1^2 , e portanto, utilizar a distribuição χ_1^2 leva a um teste conservador em favor da hipótese nula $(H_0 : \frac{1}{n} = 0)$. No mesmo artigo, é sugerido que, com amostras de tamanho razoável, para o nível de 5% o valor apropriado para o teste de razão de verossimilhança é aproximadamente 2.7. Maiores detalhes podem ser encontrados em Bollerslev (1987).

Capítulo 3

MODELOS ARCH E GARCH COM MUDANÇAS DE REGIME

3.1 Introdução

Durante a história, foi possível observar que alguns eventos fazem com que o comportamento de algumas séries mude significativamente de forma repentina. Estes eventos podem ser de natureza diversa como guerras, crises globais da economia ou mudanças de política econômica governamental.

Em um passado recente pudemos observar a crise da Ásia em 1997, a moratória russa em 1998 e a liberação do câmbio no Brasil em 1999.

A data em que estes eventos ocorrem não é exatamente o que se possa classificar de previsível e tal informação não está implícita na série temporal em estudo. Em função disto, é razoável em pensar em um modelo que permita mudanças de regime, ou seja, que o processo gerador da série temporal seja diferente em instantes diferentes.

Neste capítulo iremos apresentar alguns modelos que permitam a mudança de regimes. Iremos iniciar com a “Mistura de Distribuições I.I.D.” e em seguida, considerar modelos nos quais as mudanças de regime são regidas por uma cadeia de Markov. Ambos os modelos podem ser encontrados em Hamilton (1990) e Hamilton (1994).

A proposta de modelos combinando o modelo ARCH com as mudança de regimes surgiu com Cai (1994) e Hamilton e Susmel (1994). Um pouco mais tarde Dueker (1997) propõe um modelo combinando mudanças de regime com o modelo GARCH.

Um boa apresentação dos modelos acima mencionados também é apresentada por Rabi J.(1998).

3.2 Mistura de Distribuições IID

O modelo de mistura de distribuições é dado por:

$$y_t = \mu_{s_t} + \varepsilon_t \quad (3.1)$$

onde s_t é uma sequência de variáveis aleatórias i.i.d. que indica qual regime está em vigor na data t e assume valores de 1 até N .

ε_t é uma sequência de variáveis aleatórias i.i.d. com $E(\varepsilon_t) = 0$ e $Var(\varepsilon_t) = \sigma_{s_t}^2$.

μ_{s_t} é a esperança do processo y_t quando ao regime s_t está vigente.

Temos ainda que $P(s_t = j) = \pi_j$, $j = 1, \dots, N$ e $\sum_{j=1}^N \pi_j = 1$.

Com o processo descrito da forma acima temos que a função densidade de y_t é dada por:

$$f(y_t) = \sum_{j=1}^N \pi_j f(y_t | s_t = j) \quad (3.2)$$

Como s_t é uma variável não observável nós não temos certeza sobre o regime que esta em vigor na data t , porém é possível calcular a probabilidade, condicionada ao valor de y_t , de estar no regime j da forma que segue:

$$P(s_t = j | y_t) = \frac{P(s_t = j; y_t)}{f(y_t)} = \frac{f(y_t | s_t = j) \pi_j}{f(y_t)} \quad (3.3)$$

Como y_i é independente de y_j para $i \neq j$ temos que o log da função de verossimilhança é dada por:

$$\ell(\psi) = \sum_{t=1}^T \ln f(y_t) \quad (3.4)$$

onde $\psi = (\mu_1; \mu_2; \dots; \mu_N; \sigma_1^2; \sigma_2^2; \dots; \sigma_N^2; \pi_1; \pi_2; \dots; \pi_N)$

3.2.1 Estimativas de Máxima Verossimilhança e Algoritmo EM

Para encontrarmos as estimativas de máxima verossimilhança é necessário maximizar 3.4 sujeito a restrição de $\sum_{j=1}^N \pi_j = 1$. Desta forma, construímos o Lagrangeano que segue:

$$L(\psi) = \ell(\psi) + \lambda \left(1 - \sum_{j=1}^k \pi_j \right) \quad (3.5)$$

Com a suposição de que ε_t segue a distribuição gaussiana temos que

$$f(y_t | s_t = j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp \left(-\frac{(y_t - \mu_j)^2}{2\sigma_j^2} \right) \quad (3.6)$$

Hamilton (1994) mostra que os estimadores de máxima verossimilhança são obtidos resolvendo o sistema de equações não lineares abaixo:

$$\hat{\mu}_j = \frac{\sum_{t=1}^T y_t P(s_t = j | y_t; \hat{\psi})}{\sum_{t=1}^T P(s_t = j | y_t; \hat{\psi})} \quad (3.7)$$

$$\hat{\sigma}_j^2 = \frac{\sum_{t=1}^T (y_t - \hat{\mu}_j)^2 P(s_t = j | y_t; \hat{\psi})}{\sum_{t=1}^T P(s_t = j | y_t; \hat{\psi})} \quad (3.8)$$

$$\hat{\pi}_j = \frac{\sum_{t=1}^T P(s_t = j | y_t; \hat{\psi})}{T} \quad (3.9)$$

A solução é obtida iterativamente e consiste das seguintes etapas:

1. Atribuir estimativas iniciais para ψ ;
2. Com base nestes valores calcular 3.6, 3.2 e 3.3;
3. Recalcular 3.7, 3.8 e 3.9;
4. Repetir os passos 2 e 3 até que os valores não se alterem significativamente sob algum critério previamente estipulado.

O procedimento descrito é um caso especial do algoritmo EM (Expectation Maximization).

3.3 Modelos com Mudanças de Regimes Markovianas

Um modelo que permite mudanças de regimes regido por uma cadeia de Markov pode ser expresso como segue. Suponhamos que $s_t \in \{1; 2; \dots; k\}$ é a variável que indica qual regime a série temporal y_t está na data t . Desta forma podemos escrever o modelo de regressão com mudança de regime por:

$$y_t = x_t' \beta_{s_t} + \varepsilon_t \quad (3.10)$$

onde x_t' é o vetor de variáveis explicativas podendo conter valores defasados de y_t ;

β_{s_t} é um vetor de parâmetros, relacionados ao nível da série temporal y_t , quando o regime s_t está em vigor, $s_t = 1, \dots, k$;

ε_t é ruído branco com $E(\varepsilon_t) = 0$ e $Var(\varepsilon_t) = \sigma_{s_t}^2$.

Além disso, suponhamos também que s_t segue uma Cadeia de Markov de primeira ordem com matriz de transição P .

Iremos denotar a probabilidade de ir do regime i para o regime j por p_{ij} , ou seja, $P(s_t = j | s_{t-1} = i) = p_{ij}$. Observe também, que $\sum_{j=1}^k p_{ij} = 1$. Desta forma temos que P é dada por:

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1k} \\ p_{21} & p_{22} & \dots & p_{2k} \\ \dots & \dots & \dots & \dots \\ p_{k1} & p_{k2} & \dots & p_{kk} \end{pmatrix} \quad (3.11)$$

Vale a pena mencionar que, por hipótese, temos :

$$P(s_t | s_{t-1}; s_{t-2}; s_{t-3}; \dots) = P(s_t | s_{t-1}) \quad (3.12)$$

Com o processo descrito desta forma, temos que a função densidade de y_t dado o vetor $\mathfrak{S}_{t-1}$ e x'_t é da forma que segue:

$$f(y_t|\mathfrak{S}_{t-1}; x'_t; \psi) = \sum_{j=1}^N P(s_t = j|\mathfrak{S}_{t-1}; \psi) f(y_t|s_t = j; \mathfrak{S}_{t-1}; x'_t; \psi) \quad (3.13)$$

onde $\psi = (\psi_1; \psi_2);$

$$\psi_1 = (\beta_1; \beta_2; \dots; \beta_k; \sigma_1^2; \sigma_2^2; \dots; \sigma_k^2);$$

$$\psi_2 = (p_{11}; p_{12}; \dots; p_{1k}; p_{21}; \dots; p_{2k}; \dots; p_{k1}; \dots; p_{kk}).$$

Da mesma forma que na seção anterior, temos que s_t é uma variável não observável e portanto nós não temos certeza sobre o regime que está em vigor na data t .

A diferença entre o modelo de misturas de distribuições iid e o de mudanças de regimes markovianas é que no primeiro, em virtude da ausência de estrutura autoregressiva no processo estocástico e independência de s_i e s_m para $i \neq m$, a probabilidade de $s_t = j$ foi condicionada apenas à observação y_t enquanto no segundo, devido à presença de estrutura autoregressiva, incorporada no modelo através de valores defasados de y_t , e ausência de independência de s_t e s_{t-1} , a probabilidade de $s_t = j$ deve ser condicionada ao vetor $\mathfrak{S}_{t-1}$.

Antes de mencionarmos o processo de estimação dos parâmetros do modelo iremos definir duas quantidades que serão úteis durante o processo de estimação.

Chamaremos de probabilidade filtrada a probabilidade dada por $P(s_t = j|\mathfrak{S}_t; \psi)$ e chamaremos de probabilidade suavizada a probabilidade dada por $P(s_t = j|\mathfrak{S}_T; \psi)$.

3.3.1 Probabilidade Filtrada

Com a finalidade de apresentar um algoritmo para o cálculo das probabilidades filtradas, iremos introduzir notação na forma matricial. Desta forma, seja:

η_t o vetor $(N \times 1)$, cujo j -ésimo elemento é dado por $f(y_t|s_t = j; \mathfrak{S}_{t-1}; x'_t; \psi);$

$\xi_{t|t}$ o vetor $(N \times 1)$, cujo j -ésimo elemento é dado por $P(s_t = j|\mathfrak{S}_t; \psi).$

Observe que:

$$P(s_t = j | \mathfrak{S}_t; \psi) = P(s_t = j | y_t; x'_t; \mathfrak{S}_{t-1}; \psi) = \frac{f(y_t; s_t = j | \mathfrak{S}_{t-1}; x'_t; \psi)}{f(y_t | \mathfrak{S}_{t-1}; x'_t; \psi)} \quad (3.14)$$

mas:

$$f(y_t; s_t = j | \mathfrak{S}_{t-1}; x'_t; \psi) = f(y_t | s_t = j; \mathfrak{S}_{t-1}; x'_t; \psi) P(s_t = j | \mathfrak{S}_{t-1}; x'_t; \psi) \quad (3.15)$$

porém $P(s_t = j | \mathfrak{S}_{t-1}; x'_t; \psi) = P(s_t = j | \mathfrak{S}_{t-1}; \psi)$ pois x'_t não contém informação sobre s_t .

Desta forma temos que:

$$P(s_t = j | \mathfrak{S}_t; \psi) = \frac{f(y_t | s_t = j; \mathfrak{S}_{t-1}; x'_t; \psi) P(s_t = j | \mathfrak{S}_{t-1}; \psi)}{f(y_t | \mathfrak{S}_{t-1}; x'_t; \psi)} \quad (3.16)$$

Em notação matricial temos que:

$$\xi_{t|t} = \frac{(\xi_{t|t-1} \odot \eta_t)}{1' (\xi_{t|t-1} \odot \eta_t)} \quad (3.17)$$

onde $\odot$ significa produto elemento por elemento;

$1'$ é um vetor linha de uns.

Além de 3.17, podemos mostrar:

$$\xi_{t+1|t} = P \xi_{t|t} \quad (3.18)$$

onde P é a matriz de transição.

Para o cálculo das probabilidades filtradas é necessário, além do conhecimento de ψ , o vetor $\xi_{1|0}$ que pode ser considerado como sendo composto pelas probabilidades invariantes da cadeia de Markov.

3.3.2 Probabilidade Suavizada

Conforme já mencionado, a probabilidade suavizada é dada por $P(s_t = j | \mathfrak{S}_T; \psi)$. A probabilidade suavizada pode ser calculada utilizando o algoritmo desenvolvido por Kim (1994). Em notação matricial o algoritmo pode ser escrito da forma:

$$\xi_{t|T} = \xi_{t|t} \odot \left\{ P' \left[\xi_{t+1|T} (\div) \xi_{t+1|t} \right] \right\} \quad (3.19)$$

onde $(\div)$ significa a divisão elemento por elemento.

Este algoritmo é inicializado através do cálculo da probabilidade filtrada $(\xi_{T|T})$, utilizando as equações 3.17 e 3.18, e então calcula-se para os valores $t = T - 1, T - 2, \dots, 1$.

A dedução do algoritmo de cálculo pode ser encontrada em Hamilton (1994) ou Rabi J. (1998).

3.3.3 Estimativas de Máxima Verossimilhança Condicionais e Algoritmo EM

Hamilton (1990) demonstra que, para um processo autoregressivo de ordem m , as estimativas de máxima verossimilhança das probabilidades de transição são dadas por

$$\hat{p}_{ji} = \frac{\sum_{t=m+1}^T P(s_t = i, s_{t-1} = j | \mathfrak{S}_T; \hat{\psi})}{\sum_{t=m+1}^T P(s_{t-1} = j | \mathfrak{S}_T; \hat{\psi})} \quad (3.20)$$

enquanto que as demais são obtidas resolvendo

$$\sum_{t=m+1}^T \sum_{s_t=1}^k \dots \sum_{s_{t-m}=1}^k \frac{\partial \ln f(y_t | s_t, \dots, s_{t-m}, y_{t-1}, \dots, y_{t-m}; \psi)}{\partial \psi_1} P(s_t, \dots, s_{t-m} | \mathfrak{S}_T; \psi) = 0 \quad (3.21)$$

Percebamos que são as probabilidades suavizadas que aparecem nas equações que proporcionam o cálculo das estimativas de máxima verossimilhança, e por essa razão existe a necessidade de desenvolver um algoritmo para o cálculo das mesmas.

Particularmente, no modelo dado por 3.10, temos que as equações dadas por 3.21 podem ser escritas da forma:

$$\sum_{t=m+1}^T (y_t - x_t' \hat{\beta}_{s_t}) x_t P(s_t = j | \mathfrak{S}_T; \hat{\psi}) = 0, \text{ para } j = 1, 2, \dots, k \quad (3.22)$$

$$\hat{\sigma}_j^2 = \frac{\sum_{t=m+1}^T (y_t - x_t' \hat{\beta}_{s_t})^2 P(s_t = j | \mathfrak{S}_T; \hat{\psi})}{\sum_{t=m+1}^T P(s_t = j | \mathfrak{S}_T; \hat{\psi})} \quad (3.23)$$

Podemos observar que as equações 3.22 são as mesmas dos estimadores de mínimos quadrados a menos da ponderação pela probabilidade de estar no regime j . Neste caso, podemos dizer que $\hat{\beta}_j$ pode ser estimado através dos estimadores de mínimos quadrados da regressão de $\tilde{y}_t(j)$ sobre $\tilde{x}_t(j)$, ou seja:

$$\hat{\beta}_j = \left[\sum_{t=m+1}^T \tilde{x}_t(j) \tilde{x}_t'(j) \right]^{-1} \left[\sum_{t=m+1}^T \tilde{x}_t(j) \tilde{y}_t(j) \right], \text{ para } j = 1, 2, \dots, k \quad (3.24)$$

onde $\tilde{y}_t(j) = y_t \sqrt{P(s_t = j | \mathfrak{S}_T; \hat{\psi})}$

$$\tilde{x}_t(j) = x_t \sqrt{P(s_t = j | \mathfrak{S}_T; \hat{\psi})}$$

Antes de mencionarmos como deve-se proceder com o algoritmo EM, iremos mencionar como calcular o valor $P(s_t = i, s_{t-1} = j | \mathfrak{S}_T; \psi)$ que aparece no numerador da equação 3.20.

Hamilton (1994) demonstra que:

$$P(s_t = i, s_{t-1} = j | \mathfrak{S}_T; \psi) = P(s_t = i | \mathfrak{S}_T; \psi) \frac{p_{ji} P(s_{t-1} = j | \mathfrak{S}_t; \psi)}{P(s_t = i | \mathfrak{S}_t; \psi)} \quad (3.25)$$

Desta forma, para a obtenção das estimativas de máxima verossimilhança, o algoritmo EM é calculado da forma que segue:

1. Atribuir estimativas iniciais para ψ ;
2. Com base nestes valores calcular as probabilidades filtradas utilizando 3.17 e 3.18
3. Calcular as probabilidades suavizadas com 3.19
4. Calcular as estimativas de máxima verossimilhança utilizando 3.23, 3.24 e 3.20
5. Repetir os passos 2 e 4 até que os valores não se alterem significativamente sob algum critério previamente estipulado.

3.4 Modelos Heterocedásticos com Mudanças de Regimes Markovianas

Nos modelos apresentados anteriormente a variância, dado que $s_t = j$, era constante. Nesta seção iremos combinar os modelos com mudanças de regimes markovianas com o modelo ARCH (conhecidos como modelos SWARCH) e o modelo GARCH (conhecidos como modelos SWGARCH). Conforme já mencionado, estes modelos foram propostos por Cai (1994), Hamilton e Susmel (1994) e Dueker (1997).

3.4.1 Modelos SWARCH

SWARCH proposto por Cai

Em sua proposta Cai (1994) propõe o uso do modelo que segue:

$$y_t = c_0 + c_1 s_t + z_t \quad (3.26)$$

$$z_t = b_1 z_{t-1} + \dots + b_k z_{t-k} + \varepsilon_t \quad (3.27)$$

$$\varepsilon_t = \sqrt{h_t} \nu_t, \quad \nu_t \text{ iid } \sim N(0; 1) \quad (3.28)$$

$$h_t = \gamma(s_t) + \sum_{i=1}^g \alpha_i \varepsilon_{t-i}^2 \quad \alpha_i \geq 0 \quad (3.29)$$

onde $\gamma(s_t) = \gamma_0 + \gamma_1 s_t, \quad \gamma_0 \geq 0, \gamma_1 \geq 0$

s_t assume os valores 0 e 1 e segue um processo de Markov de primeira ordem com as probabilidades de transição dadas por:

$$p(s_t = 0 | s_{t-1} = 0) = p$$

$$p(s_t = 1 | s_{t-1} = 0) = 1 - p$$

$$p(s_t = 1 | s_{t-1} = 1) = q$$

$$p(s_t = 0 | s_{t-1} = 1) = 1 - q$$

Observe que no modelo acima, a mudança de regimes afeta o nível através de $c_1 s_t$ e a variância através $\gamma(s_t)$. Neste modelo, o valor esperado de ε_{t+1}^2 condicionado em $y_t, y_{t-1}, \dots, y_1$ é dado por:

$$\begin{aligned} & E(\varepsilon_{t+1}^2 | y_t, y_{t-1}, \dots, y_1) \\ &= \sum_{s_{t+1}} \sum_{s_t} \dots \sum_{s_{t-m+1}} E(\varepsilon_{t+1}^2 | y_t, y_{t-1}, \dots, y_1, s_{t+1}, s_t, \dots, s_{t-m+1}) \\ & \quad \cdot P(s_{t+1}, s_t, \dots, s_{t-m+1} | y_t, y_{t-1}, \dots, y_1) \\ &= \sum_{s_{t+1}} \sum_{s_t} \dots \sum_{s_{t-m+1}} \left(\gamma_0 + \gamma_1 s_{t+1} + \sum_{i=1}^g \alpha_i \varepsilon_{t+1-i}^2 \right) \cdot P(s_{t+1} | s_t) \\ & \quad \cdot P(s_t, s_{t-1}, \dots, s_{t-m+1} | y_t, y_{t-1}, \dots, y_1) \end{aligned} \quad (3.30)$$

A extensão de previsão da variância condicional para j períodos a frente pode ser feita da forma que segue por:

$$E\left(\varepsilon_{t+j}^2 | y_t, y_{t-1}, \dots, y_1\right) = \sum_{s_{t+j}} \sum_{s_{t+j-1}} \dots \sum_{s_{t-m+1}} E\left(\varepsilon_{t+j}^2 | y_t, y_{t-1}, \dots, y_1, s_{t+j}, s_{t+j-1}, \dots, s_{t-m+1}\right) \cdot P\left(s_{t+j}, s_{t+j-1}, \dots, s_{t-m+1} | y_t, y_{t-1}, \dots, y_1\right) \quad (3.31)$$

De onde temos que:

$$E\left(\varepsilon_{t+j}^2 | y_t, y_{t-1}, \dots, y_1, s_{t+j}, s_{t+j-1}, \dots, s_{t-m+1}\right) = \gamma_0 + \gamma_1 s_{t+1} + \sum_{i=1}^g \alpha_i \varepsilon_{t+1-i}^2 \quad \text{para } j = 1 \quad (3.32)$$

$$= \gamma_0 + \gamma_1 s_{t+j} + \sum_{i=1}^g \alpha_i h_{t+j-i|t} \quad \text{para } j > 1 \quad (3.33)$$

onde $h_{\tau|t} = \varepsilon_{\tau}^2$ para $\tau \leq t$ e $h_{\tau|t} = E\left(\varepsilon_{\tau}^2 | y_t, y_{t-1}, \dots, y_1, s_{t+\tau}, s_{t+\tau-1}, \dots, s_{t-m+1}\right)$ para $\tau > t$

Também temos que:

$$P\left(s_{t+j}, s_{t+j-1}, \dots, s_{t-m+1} | y_t, y_{t-1}, \dots, y_1\right) = P\left(s_{t+j} | s_{t+j-1}\right) \cdot \dots \cdot P\left(s_{t+1} | s_t\right) \cdot P\left(s_t, s_{t-1}, \dots, s_{t-m+1} | y_t, y_{t-1}, \dots, y_1\right) \quad (3.34)$$

Embora Cai(1994) não comente como deve ser escolhido m , é fácil notar que $m = k + g$ pois h_t é função de $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-g}$ e temos de 3.26 e 3.27 que:

$$\begin{aligned} \varepsilon_{t-g} &= z_{t-g} - b_1 z_{t-g-1} - \dots - b_k z_{t-g-k} \\ &= (y_{t-g} - c_0 + c_1 s_{t-g}) - b_1 (y_{t-g-1} - c_0 + c_1 s_{t-g-1}) - \dots - b_k (y_{t-g-k} - c_0 + c_1 s_{t-g-k}) \end{aligned} \quad (3.35)$$

Portanto, h_t depende dos estados $s_t, s_{t-1}, \dots, s_{t-g-k}$.

SWARCH proposto por Hamilton e Susmel

Hamilton e Susmel (1994) propõe o modelo com vários regimes que segue:

$$y_t = \mu + \beta y_{t-1} + \varepsilon_t \quad (3.36)$$

$$\varepsilon_t = \sqrt{g_{s_t}} \tilde{\varepsilon}_t \quad (3.37)$$

$$\tilde{\varepsilon}_t = \sqrt{h_t} v_t \quad (3.38)$$

$$h_t = \gamma + \alpha_1 \tilde{\varepsilon}_{t-1}^2 + \alpha_2 \tilde{\varepsilon}_{t-2}^2 + \dots + \alpha_q \tilde{\varepsilon}_{t-q}^2 + \xi d_{t-1} \tilde{\varepsilon}_{t-1}^2 \quad (3.39)$$

onde $d_{t-1} = 1$ se $\tilde{\varepsilon}_{t-1} < 0$ e $d_{t-1} = 0$ se $\tilde{\varepsilon}_{t-1} > 0$;

$E(v_t) = 0$ e $Var(v_t) = 1$;

s_t assume valores de 1 até k ;

g_{s_t} assume os valores $g_1, g_2, \dots, g_k$.

Observe que o parâmetro $\xi \neq 0$ faz com que a variância condicional h_t seja afetada de forma distinta para valores positivos e negativos de $\tilde{\varepsilon}_t$. A previsão de ε_{t+j}^2 condicionada no conjunto de informações até a data t é dada por:

$$\begin{aligned} E(\varepsilon_{t+j}^2 | s_t, \dots, s_{t-q+1}; \tilde{\varepsilon}_t, \dots, \tilde{\varepsilon}_{t-q+1}) &= E(g_{s_{t+j}} \tilde{\varepsilon}_{t+j}^2 | s_t, \dots, s_{t-q+1}; \tilde{\varepsilon}_t, \dots, \tilde{\varepsilon}_{t-q+1}) \\ &= E(g_{s_{t+j}} | s_t, \dots, s_{t-q+1}) \cdot E(\tilde{\varepsilon}_{t+j}^2 | \tilde{\varepsilon}_t, \dots, \tilde{\varepsilon}_{t-q+1}) \end{aligned} \quad (3.40)$$

A última igualdade segue do fato que s_t é independente de $\tilde{\varepsilon}_\tau$ e v_τ são independentes para todo t e τ . Como s_t segue uma cadeia de Markov temos que o primeiro termo de 3.40 é dado por:

$$E(g_{s_{t+j}} | s_t, s_{t-1}, \dots, s_{t-q+1}) = \sum_{j=1}^k g_j P(s_{t+j} = j | s_t) \quad (3.41)$$

As probabilidades de transição m -períodos a frente pode ser calculada multiplicando a matriz de transição P por ela mesma j vezes. Desta forma, podemos escrever que:

$$E(g_{s_t+j}|s_t = i) = g' P^j e_i$$

onde $g' = (g_1; g_2 \dots; g_k);$

e_i é a i -ésima coluna da matriz identidade.

O segundo termo de 3.40 pode ser calculado como:

$$E(\tilde{\varepsilon}_{t+j}^2 | \tilde{\varepsilon}_t, \dots, \tilde{\varepsilon}_{t-q+1}) = \gamma + \alpha_1 \tilde{\varepsilon}_t^2 + \dots + \alpha_q \tilde{\varepsilon}_{t-q+1}^2 + \xi d_t \tilde{\varepsilon}_t^2 \quad \text{para } j = 1 \quad (3.42)$$

$$= \gamma + (\alpha_1 + \xi/2) \tilde{h}_{t+j-1|t} + \alpha_2 \tilde{h}_{t+j-2|t} + \dots + \alpha_q \tilde{h}_{t+j-q|t} \quad \text{para } j > 1 \quad (3.43)$$

onde $\tilde{h}_{\tau|t} = \tilde{\varepsilon}_t^2$ para $\tau \leq t$ e $\tilde{h}_{\tau|t} = E(\tilde{\varepsilon}_\tau^2 | \tilde{\varepsilon}_t^2, \tilde{\varepsilon}_{t-1}^2, \dots)$ para $\tau > t$

Como $\tilde{\varepsilon}_t = \frac{\varepsilon_t}{\sqrt{g_{s_t}}}$ temos que:

$$E(\tilde{\varepsilon}_{t+j}^2 | \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}, \dots, \tilde{\varepsilon}_{t-q+1}) = \tilde{h}_{t+j|t}(s_t, s_{t-1}, \dots, s_{t-q+1}; \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}) \quad (3.44)$$

Logo temos que a previsão é dada por:

$$E(\varepsilon_{t+j}^2 | s_t, s_{t-1}, \dots, s_{t-q+1}; \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}, \dots, \tilde{\varepsilon}_{t-q+1}) = (g' P^j e_{s_t}) \cdot \tilde{h}_{t+j|t}(s_t, s_{t-1}, \dots, s_{t-q+1}; \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}) \quad (3.45)$$

Como os valores de $s_t, s_{t-1}, \dots, s_{t-q+1}$ são desconhecidos temos que:

$$\begin{aligned}
 E\left(\varepsilon_{t+j}^2 | \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}, \dots, \tilde{\varepsilon}_{t-q+1}\right) &= \sum_{s_t} \sum_{s_{t-1}} \dots \sum_{s_{t-q+1}} \left(g' P^j e_{s_t}\right) \\
 &\quad \tilde{h}_{t+j|t}(s_t, s_{t-1}, \dots, s_{t-q+1}; \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}) \cdot \\
 &\quad P(s_t, s_{t-1}, \dots, s_{t-q+1} | y_t, y_{t-1}, \dots)
 \end{aligned} \tag{3.46}$$

3.4.2 SWGARCH proposto por Dueker

Os modelos propostos por Cai(1994) e Hamilton e Susmel(1994) são modelos ARCH com mudanças de regime seguindo uma cadeias de Markov. A extensão destes modelos para o GARCH com mudanças de regime seguindo uma cadeia de Markov (SWGARCH) não é trivial devido ao fato que a variância condicional ficaria dependente de todos os estados desde o instante inicial.

A dependência da variância condicional de todos os estados desde o instante inicial torna o processo de estimação praticamente inviável devido a grande quantidade de probabilidades necessárias. Por exemplo, imaginemos que temos que analisar uma série que contenha 1000 observações e queremos ajustar um modelo GARCH com mudanças de regimes markovianas com apenas dois estados. Desta forma seria necessário o cálculo de 2^{1000} probabilidades para que se possa obter a variância condicional.

Em 1997, Dueker propõe uma extensão destes modelos de forma que seja utilizado o modelo GARCH com mudanças de regime seguindo uma cadeia de Markov. Em tal proposta Dueker faz uso de um procedimento proposto por Kim (1994).

Kim (1994) afirma que, se h_t segue um processo autoregressivo de ordem k é possível considerar h_t como uma função de $k + 1$ regimes mais recentes sem perda significativa de precisão no cálculo do log da função de verossimilhança. Desta forma, temos que no SWGARCH(1,1), onde $k = 1$, h_t fica sendo uma função apenas de s_t e s_{t-1} .

Para exemplificar, iremos apresentar a seguir, como ficam os modelos propostos por Cai

(1994) e Hamilton e Susmel (1994).

O modelo proposto por Cai (1994) pode ser adaptado para SWGARCH da forma que segue:

$$h_t^{(i,j)} = \gamma(s_t = i) + \alpha \left(\varepsilon_{t-1}^{(j)} \right)^2 + \phi h_{t-1}^{(j)} \quad (3.47)$$

$$\text{onde} \quad h_{t-1}^{(j)} = \sum_i P(s_{t-2} = i | s_{t-1} = j; \mathfrak{S}_t) h_{t-1}^{(j,i)}$$

A variância condicional é dada por:

$$\widehat{h}_t = \sum_i P(s_t = i | \mathfrak{S}_t) h_t^{(i)} \quad (3.48)$$

O modelo proposto por Hamilton e Susmel (1994), ao contrário do que ocorre com o modelo proposto por Cai (1994), a variância condicional depende de s_t apenas através de s_{t-1} . Isto faz como que o modelo adaptado para SWGARCH seja um pouco mais simples.

O modelo adaptado é descrito da forma que segue:

$$h_t^{(j)} = \gamma + \frac{\alpha_1}{g_j} \left(\varepsilon_{t-1}^{(j)} \right)^2 + \phi \widehat{h}_{t-1} \quad (3.49)$$

Novamente temos que a variância condicional é dada pela equação 3.48.

3.5 Função de verossimilhança

O log da função de verossimilhança condicionada, associada aos modelos SWARCH e SWGARCH, é dada por:

$$\ell(\psi) = \sum_t \ln f(y_t | \mathfrak{F}_{t-1}; \psi) \quad (3.50)$$

Para obter a função de verossimilhança é preciso calcular $f(y_t | s_t, \dots, s_{t-m}, y_{t-1}, \dots, y_{t-m}; \psi)$ cujo cálculo passa a ser descrito a seguir:

1. Suponha $P(s_t, \dots, s_{t-m} | \mathfrak{F}_t; \psi)$ seja conhecida;
2. Multiplica-se a quantidade apresentada no passo 1 por $f(y_{t+1} | s_{t+1}, \dots, s_{t-m+1}; \mathfrak{F}_t; \psi)$ e por $p_{s_t | s_{t+1}}$ obtendo desta forma $P(s_{t+1}, \dots, s_{t-m+1}, s_{t-m}, y_{t+1} | \mathfrak{F}_t; \psi)$;
3. Sabe-se que $f(y_{t+1} | \mathfrak{F}_t; \psi) = \sum_{s_{t+1}=1}^k \sum_{s_t=1}^k \dots \sum_{s_{t-m}=1}^k P(s_{t+1}, \dots, s_{t-m}, y_{t+1} | \mathfrak{F}_t; \psi)$

Observe que para obter a probabilidade filtrada $P(s_{t+1}, \dots, s_{t-m+1} | \mathfrak{F}_{t+1}; \psi)$ basta calcular:

$$P(s_{t+1}, \dots, s_{t-m+1} | \mathfrak{F}_{t+1}; \psi) = \frac{\sum_{s_{t-m}=1}^k P(s_{t+1}, \dots, s_{t-m+1}, s_{t-m}, y_{t+1} | \mathfrak{F}_t; \psi)}{f(y_{t+1} | \mathfrak{F}_t; \psi)} \quad (3.51)$$

O algoritmo acima descrito deve-se iniciar com $P(s_0, \dots, s_{-m} | \mathfrak{F}_0; \psi)$ utilizando as probabilidades invariantes do processo de Markov.

As probabilidades suavizadas continuam sendo dadas através da equação 3.19.

Capítulo 4

MODELO GARCH COM NORMAL CONTAMINADA

4.1 Introdução

No capítulo I, nós mencionamos os modelos ARCH e GARCH propostos Engle (1982) e Bollerslev (1986) respectivamente.

No capítulo II, foi apresentado o modelo GARCH - T proposto por Bollerslev (1987) em função do fato de que algumas séries financeiras possuem distribuições com excesso de curtose quando comparadas à distribuição normal.

No capítulo III, foi apresentado o modelo de mistura de distribuições e o modelo de mudanças de regimes markovianas.

Neste capítulo, iremos apresentar modelos similares ao mencionado no capítulo I, porém, da mesma forma que o modelo proposto por Bollerslev (1987), este modelo permite a existência de uma distribuição com excesso de curtose quando comparada à distribuição gaussiana. Para construir o modelo mencionado iremos nos valer de alguns conceitos similares aos utilizados nos modelos de mistura de distribuições e mudanças de regimes markovianas apresentados no capítulo III.

Doravante iremos nos referir à mistura de distribuições gaussianas independentes como

normal contaminada (NC). Um resumo de misturas de distribuições iid's pode ser encontrado em Everit e Hand (1981).

4.2 Modelo GARCH com Normal Contaminada (NC)

Antes de apresentarmos o modelo ARCH com distribuição normal contaminada, iremos fazer algumas breves observações sobre algumas propriedades da distribuição normal contaminada.

4.2.1 Distribuição Normal Contaminada

Seja x uma variável aleatória que segue uma distribuição normal contaminada com parâmetros μ_i, σ_i^2, π_i com $i = 1, \dots, k$ com $\sum_{i=1}^k \pi_i = 1$. Denotaremos que:

$$x \sim NC(\mu_1; \mu_2; \dots; \mu_k; \sigma_1^2; \sigma_2^2; \dots; \sigma_k^2; \pi_1; \pi_2; \dots; \pi_k).$$

Cuja função densidade é dada por:

$$f(y_t) = \sum_{j=1}^k \frac{\pi_j}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(y_t - \mu_j)^2}{2\sigma_j^2}\right)$$

É fácil notar que os momentos da normal contaminada são dados por:

$$E(x^i) = \sum_{j=1}^k \pi_j E_j(x^i) \quad (4.1)$$

onde $E_j(x^i)$ é o i -ésimo momento da j -ésima distribuição normal que está sendo misturada.

A função geradora de momentos da normal contaminada é dada por:

$$M_x(t) = \sum_{j=1}^k \pi_j \exp\left(\mu_j t + \frac{\sigma_j^2 t^2}{2}\right) \quad (4.2)$$

Através da função geradora de momentos da normal contaminada podemos notar que:

Se $x \sim NC(\mu_1; \mu_2; \dots; \mu_k; \sigma_1^2; \sigma_2^2; \dots; \sigma_k^2; \pi_1; \pi_2; \dots; \pi_k)$ e $y = \frac{a+x}{b}$ com $a, b \neq 0$ então $y \sim NC(\frac{a+\mu_1}{b}; \frac{a+\mu_2}{b}; \dots; \frac{a+\mu_k}{b}; \frac{\sigma_1^2}{b^2}; \frac{\sigma_2^2}{b^2}; \dots; \frac{\sigma_k^2}{b^2}; \pi_1; \pi_2; \dots; \pi_k)$

4.2.2 Modelo GARCH com Normal Contaminada

O modelo GARCH com normal contaminada (GARCH - NC) é dado por:

$$y_t = x_t' \beta + \varepsilon_t \quad (4.3)$$

$$\varepsilon_t = \sqrt{h_t} v_t \quad (4.4)$$

$$h_t = z_t' \delta \quad (4.5)$$

onde $t = 1, \dots, T$

x_t' é o vetor $(1 \times s)$ de variáveis explicativas podendo conter lag's da série temporal $\{y_t\}$;

β é um vetor de s parâmetros, relacionados ao nível da série temporal $\{y_t\}$, a serem estimados;

δ' , é o vetor $(\gamma; \alpha_1; \dots; \alpha_p; \phi_1; \dots; \phi_q)'$ de parâmetros, relacionados à variância da série temporal $\{y_t\}$, a serem estimados;

z_t' é o vetor $(1; \varepsilon_{t-1}^2; \varepsilon_{t-2}^2; \dots; \varepsilon_{t-p}^2; h_{t-1}; h_{t-2}; \dots; h_{t-q})$;

v_t é sequência de variáveis aleatórias iid com distribuição

$NC(0; \dots; 0; \sigma_1^2; \dots; \sigma_k^2; \pi_1; \dots, \pi_k)$ com a restrição $\sum_{j=1}^k \pi_j \sigma_j^2 = 1$.

A restrição $\sum_{j=1}^k \pi_j \sigma_j^2 = 1$ se faz necessária para que $E(v_t^2) = 1$ e consequentemente $E(\varepsilon_t^2 | \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots) = h_t$.

Observe que, dado $\mathfrak{F}_{t-1}$, ε_t segue a distribuição normal contaminada de parâmetros $(0; \dots; 0; h_t \sigma_1^2; \dots; h_t \sigma_k^2; \pi_1; \dots; \pi_k)$. Observe também que a distribuição de y_t , dado $\mathfrak{F}_{t-1}$ e x_t , é normal contaminada com parâmetros $(x_t' \beta; \dots; x_t' \beta; h_t \sigma_1^2; \dots; h_t \sigma_k^2; \pi_1; \dots; \pi_k)$.

Com o processo descrito da forma acima temos que a função densidade de y_t , dado $\mathfrak{S}_{t-1}$ e x_t , é dada por:

$$f(y_t|\mathfrak{S}_{t-1}; x_t) = \sum_{j=1}^k \pi_j f(y_t|s_t = j; \mathfrak{S}_{t-1}; x_t) \quad (4.6)$$

$$\text{onde} \quad f(y_t|s_t = j; \mathfrak{S}_{t-1}; x_t) = \frac{1}{\sqrt{2\pi h_t \sigma_j^2}} \exp\left(\frac{-(y_t - x_t' \beta)^2}{2 h_t \sigma_j^2}\right)$$

Como s_t é uma variável não observável nós não temos certeza sobre o regime que está em vigor na data t , porém é possível calcular a probabilidade, condicionada ao valor de y_t , $\mathfrak{S}_{t-1}$ e x_t , de estar no regime j da forma que segue:

$$P(s_t = j|y_t; \mathfrak{S}_{t-1}; x_t) = \frac{P(s_t = j; y_t|\mathfrak{S}_{t-1}; x_t)}{f(y_t|\mathfrak{S}_{t-1}; x_t)} = \frac{f(y_t|s_t = j; \mathfrak{S}_{t-1}; x_t) \pi_j}{f(y_t|\mathfrak{S}_{t-1}; x_t)} \quad (4.7)$$

Temos que o ln da função de verossimilhança é dado por:

$$\zeta(\psi) = \sum_t \ln f(y_t|\mathfrak{S}_{t-1}; x_t) \quad (4.8)$$

$$\text{onde} \quad \psi = (\beta; \delta; \sigma_1^2; \sigma_2^2; \dots; \sigma_k^2; \pi_1; \pi_2; \dots; \pi_k)$$

4.3 Testes de Hipóteses

Quando estimamos o modelo GARCH - NC, surgem três perguntas, são elas:

1. Como testar se os parâmetros associados ao nível da série (β) são diferentes de zero ?
2. Como testar se os parâmetros associados à variância da série (δ) são diferentes de zero ?
3. Como testar se os erros padronizados não são provenientes da distribuição normal ?

Para as três perguntas acima podemos usar a estatística de razão de verossimilhança. O teste de Razão de Verossimilhança consiste em calcular a estatística $-2\log \frac{\zeta(\hat{\psi}_0)}{\zeta(\hat{\psi}_a)}$ onde $\hat{\psi}_0$ são as estimativas com as restrições da hipótese nula e $\hat{\psi}_a$ são as estimativas dos parâmetros sem restrições. Infelizmente, nestes casos, esta estatística não possui distribuição conhecida e existe a necessidade de encontrar esta distribuição de forma empírica. No artigo em que Cai (1994) apresenta o SWARCH, é proposto que a distribuição empírica, da estatística de interesse, seja encontrada da forma que segue:

1. Estimar os parâmetros sob a hipótese nula ;
2. Gerar uma grande quantidade de séries sob a hipótese nula , com mesmo tamanho da série que está sendo estudada, com os parâmetros estimados no passo 1;
3. Estimar os parâmetros das séries geradas sob a hipóteses nula e alternativa;
4. Calcular os percentis da estatística de razão de verossimilhança baseado nas verossimilhanças encontradas no passo 3;
5. Verificar se a estatística de razão de verossimilhança é maior que o percentil de confiança estipulado previamente.

Na questão 3, além da estatística de razão de verossimilhança, iremos apresentar uma outra estatística, cuja proposta passa a ser justificada abaixo.

Devido à construção da função densidade da distribuição $NC(0; \dots; 0; \sigma_1^2; \dots; \sigma_k^2; \pi_1; \dots, \pi_k)$, podemos notar que a mesma restringe-se à distribuição normal se $\pi_j = 1$ para algum $j = 1, \dots, k$ ou $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$ (note que, em virtude da restrição $\sum_{j=1}^k \pi_j \sigma_j^2 = 1$, $\sigma^2 = 1$). Tal fato pode ser comprovado, observando os momentos da distribuição normal e normal contaminada.

Temos que os momentos centrados da distribuição normal, com média zero, são dados pela equação que segue :

$$E_{Normal}(X^i) = 1 \cdot 3 \cdot \dots \cdot (i-1) \cdot \sigma^i \quad (4.9)$$

Por conseqüência, temos que os momentos da distribuição $NC(0; \dots; 0; \sigma_1^2; \dots; \sigma_k^2; \pi_1; \dots, \pi_k)$ são:

$$E_{NC}(X^i) = \sum_{m=1}^k \pi_m E_m(x^i) \quad (4.10)$$

onde $E_j(x^i)$ é o i -ésimo momento da j -ésima distribuição normal que está sendo misturada.

Se $\pi_j = 1$ para algum $j = 1, \dots, k$ temos que:

$$E_{NC}(X^i) = \sum_{m=1}^k \pi_m E_m(x^i) = \pi_j E_j(x^i) = 1 \cdot 3 \cdot \dots \cdot (i-1) \cdot \sigma_j^i \quad (4.11)$$

Por outro lado, se $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$ temos que:

$$E_{NC}(X^i) = \sum_{m=1}^k \pi_m E_m(x^i) = 1 \cdot 3 \cdot \dots \cdot (i-1) \cdot \sigma^i \cdot \sum_{m=1}^k \pi_m = 1 \cdot 3 \cdot \dots \cdot (i-1) \cdot \sigma^i \quad (4.12)$$

Portanto, quando as condições $\pi_j = 1$ para algum $j = 1, \dots, k$ ou $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$ estão atendidas temos que a $NC(0; \dots; 0; \sigma_1^2; \dots; \sigma_k^2; \pi_1; \dots, \pi_k)$ se resume a distribuição

gaussiana e não existe a necessidade de utilizar a forma geral. Podemos notar também que as restrições acima são equivalentes à equação que segue:

$$\left(\prod_{m=1}^k (1 - \pi_m) \right) \left(\sum_{m=1}^k (\sigma_m^2 - 1)^2 \right) = 0 \quad (4.13)$$

Podemos então, basear nosso teste na estatística abaixo:

$$U = \left(\prod_{m=1}^k (1 - \widehat{\pi}_m) \right) \left(\sum_{m=1}^k (\widehat{\sigma}_m^2 - 1)^2 \right) \quad (4.14)$$

Assim como a razão de verossimilhança, esta estatística possui distribuição ignorada e devemos encontrar a distribuição desta, de forma empírica. A distribuição empírica, da estatística U , deve ser encontrada da forma que segue:

1. Estimar os parâmetros sob a hipótese nula (no caso usando o modelo GARCH usual);
2. Gerar uma grande quantidade de séries sob a hipótese nula (no caso usando o modelo GARCH usual), com mesmo tamanho da série que está sendo estudada, com os parâmetros estimados no passo 1;
3. Estimar os parâmetros das séries geradas sob a hipótese alternativa (no caso o modelo GARCH - NC);
4. Calcular os percentis da estatística U baseado nas verossimilhanças encontradas no passo 3;
5. Verificar se a estatística U é maior que o percentil de confiança estipulado previamente.

Devemos notar que, embora este procedimento demande bastante trabalho, no caso 3, o uso da estatística U demanda menos trabalho do que o uso da estatística de razão de verossimilhança. Tal fato ocorre porque, caso queiramos utilizar a estatística de razão de verossimilhança, devemos além de executar os passos acima mencionados, estimar os parâmetros das séries geradas sob a hipótese nula.

Os testes de hipótese apresentados nesta seção possuem o objetivo de colaborar na identificação do modelo adequado para a série em estudo. Vale a pena lembrar que a metodologia de Box - Jenkins, aplicada aos resíduos quadrados, e os critérios de AKAIKE e Schwarz também podem ser usados na identificação do modelo apropriado. Em vários casos é possível que o uso da metodologia Box - Jenkins e os critérios de AKAIKE e Schwarz tragam indicações suficientes sobre o modelo sem que para isso seja necessário encontrar a distribuição empírica das estatísticas anteriormente mencionadas.

Capítulo 5

SIMULAÇÃO COM ALGUNS MODELOS GARCH

5.1 Introdução

Ao longo do capítulo I até o capítulo III nós apresentamos diversos modelos de variância condicional não constante desde o ARCH proposto por Engle em 1982 até o GARCH com mudanças markovianas de regime proposto por Dueker em 1997. No capítulo IV fizemos a proposta de modelo GARCH com os erros padronizados seguindo uma distribuição normal contaminada.

Neste capítulo iremos fazer, através de simulação, uma comparação entre os modelos GARCH com suposição de distribuição dos erros normal, t-student e normal contaminada.

5.2 Gerando os dados

Para realizar as simulações foram geradas 4000 séries com amostras de tamanho 1101, sendo que os erros destas foram gerados seguindo as distribuições abaixo:

- 1000 - Normal com média 0 e variância 1;
- 1000 - t-student com 5 graus de liberdade;
- 1000 - Normal contaminada com $\pi_1 = 0.876869$, $\sigma_1^2 = 0.470053$, $\sigma_2^2 = 4.773966$, e $\mu_1 = \mu_2 = 0$;
- 1000 - Normal contaminada com $\pi_1 = 0.688773$, $\sigma_1^2 = 0.47168$, $\sigma_2^2 = 2.169216$ e $\mu_1 = \mu_2 = 0$;

Para gerar todas as séries o procedimento adotado foi:

1. Gerar número aleatório ω entre 0 e 1 (sub rotina do pacote SAS);
2. Calcular o valor x da distribuição tal que $P(X < x) = \omega$.

Embora não seja de suma importância, foram utilizados os mesmos ω para gerar os quatro grupos de séries.

Os parâmetros das distribuições normais contaminadas foram escolhidos de duas formas a saber:

1. Os parâmetros da primeira normal contaminada, que passa a ser denominada normal cont I, foram escolhidos de forma que os quatro primeiros momentos sejam similares aos momentos da distribuição t-student, padronizada, com 5 graus de liberdade. Com os parâmetros escolhidos desta forma, temos que:

Tabela 5.2.1 - Momemtos da t - student com 5 g.l. e normal contaminada

Momentos	t - student	Normal Contaminada
$E(x)$	0	0
$E(x^2)$	1	1
$E(x^3)$	0	0
$E(x^4)$	9.00	9.00

2. Os parâmetros da segunda normal contaminada, que passa a ser denominada normal cont II, foram escolhidos de forma que a função abaixo fosse minimizada.

$$h(p; \sigma_1^2; \sigma_2^2) = \sum_j \left(f_{NC}(j; p; \sigma_1^2; \sigma_2^2) - f_t(j; 5) \right)^2, j = -10; -9.99; \dots; 9.99; 10.$$

onde $f_{NC}(j; p; \sigma_1^2; \sigma_2^2)$ é a função densidade da distribuição normal contaminada de parâmetros $NC(0; 0; \sigma_1^2; \sigma_2^2; \pi_1; 1 - \pi_1)$ no ponto j ;

$f_t(j; 5)$ é a função densidade da distribuição t-student padronizada (função densidade da variável $a = \sqrt{\frac{n-2}{n}}b$ onde b segue a distribuição t-student de n graus de liberdade - neste caso $n = 5$) no ponto j ;

Segue abaixo o gráfico contendo a função densidade de probabilidade das quatro distribuições acima mencionadas.

Gráfico 5.2.1 - Funções Densidade de Probabilidade

Embora o gráfico acima apresente as funções densidade das distribuições t-student padronizada e da segunda normal contaminada praticamente iguais, as mesmas diferem acima de 2 desvios padrões como pode ser observado no "zoom" dado no gráfico que segue.

Gráfico 5.2.2 - Comparação t-student, padronizada, de 5 graus de liberdade e segunda normal contaminada

Como pode ser observado, a função densidade da distribuição t-student assume valores maiores que a função densidade da distribuição segunda normal contaminada para valores tais que $|x| > 4$.

Após gerar os erros da forma acima mencionada, as séries finais foram geradas segundo o modelo que segue:

$$y_t = -0.07 \cdot y_{t-1} + \varepsilon_t \quad (5.1)$$

$$\varepsilon_t = \sqrt{h_t} \nu_t \quad (5.2)$$

$$h_{t-1} = 0.0001 + 0.1 \cdot \varepsilon_{t-1}^2 + 0.8 \cdot h_{t-1} \quad (5.3)$$

onde ν_t segue uma das quatro distribuições: t-student, padronizada, de 5 graus, normal, normal contaminada I ou normal contaminada II.

5.3 Estimando os Parâmetros e Analisando os Resultados

Após gerar as 4000 séries, os primeiros 100 valores de cada uma delas foram descartados e os 1000 valores seguintes foram utilizados para estimar modelos GARCH(1,1) com suposições de distribuição normal, t-student e normal contaminada. Por último foi realizada a previsão por intervalo do último valor gerado.

Os modelos foram estimados utilizando maximização numérica da função de verossimilhança.

Na estimação do modelo GARCH - T, os graus de liberdade foram limitados em 102. Tal restrição foi necessária em virtude do computador não conseguir calcular a função gama, existente na função densidade da distribuição t - student, quando os graus de liberdade superam o valor 102. Encontrar graus de liberdade estimados com um alto valor é bastante comum no caso em que séries foram geradas com erros gaussianos.

Vale a pena lembrar que esta limitação não é algo muito ruim, pois a distribuição t-student com 102 graus de liberdade é aproximadamente normal.

Um ponto de interesse deste trabalho são as previsões por intervalo para os valores futuros da série. É evidente que para realizar as previsões por intervalo é necessário obter estimativas dos percentis da distribuição, portanto, temos três casos, são eles:

1. Estimando modelo GARCH : Percentis são funções da distribuição normal com média zero e variância 1 e portanto, as estimativas são sempre as mesmas;
2. Estimando o modelo GARCH - T: Percentis são uma função da estimativa dos graus de liberdade;
3. Estimando o modelo GARCH - NC: Percentis são uma função da estimativas dos parâmetros π , σ_1^2 e σ_2^2 ;

É de se esperar que estimativas ruins dos percentis conduzam à intervalos de confiança que não refletem a confiança desejada.

Neste trabalho queremos fazer estimativas por intervalos de 90%, 95% e 98% de confiança.

5.3.1 Estimando as séries com erros normais

Após estimar os parâmetros das séries geradas com erros seguindo a distribuição normal construímos a tabela que segue abaixo:

Tabela 5.3.1.1 - Resultados das estimativas para séries com erros normais

Estimado com Suposição		β_1	γ	α_1	ϕ_1	GL	π	σ_1^2	σ_2^2	U
Valor Real		-0.07	0.0001	0.1	0.8	∞ (102)	*	*	*	0.00
Normal	Média	-0.0694	0.00012	0.101	0.780	-	-	-	-	-
	Mediana	-0.0686	0.00011	0.100	0.792	-	-	-	-	-
	EQM ^{1/2}	0.0345	0.00006	0.028	0.083	-	-	-	-	-
T - Student	Média	-0.0694	0.00012	0.101	0.781	88.37	-	-	-	-
	Mediana	-0.0684	0.00011	0.100	0.791	102.00	-	-	-	-
	EQM ^{1/2}	0.0346	0.00006	0.029	0.083	22.25	-	-	-	-
Normal Contaminada	Média	-0.0693	0.00012	0.100	0.781	-	0.4122	0.8421	1.1971	0.0162
	Mediana	-0.0687	0.00010	0.100	0.792	-	0.1227	0.9970	1.0111	0.0001
	EQM ^{1/2}	0.0345	0.00006	0.029	0.083	-	0.3904	0.2690	0.7380	0.0348

Os números em vermelho indicam o cálculo do desvio padrão * Podemos ter três situações ($P = 1, \sigma_z^2 = 1$); ($P = 0, \sigma_z^2 = 1$); ($\sigma_z^2 = \sigma_z^2 = 1$)

Como pode ser observado na tabela acima, os EQM's das estimativas dos parâmetros β_1 , γ , α_1 e ϕ_1 não são muito diferentes independentemente do modo de estimação utilizado. Em todos os três casos analisados as estimativas foram, em média e mediana, bastante boas.

Quando as estimativas foram feitas supondo o modelo GARCH - T, como era esperado, tivemos um grande número de séries (58.0%) atingindo o valor máximo permitido para os graus de liberdade. O valor mínimo encontrado foi de 10.86.

No modelo GARCH - NC, a avaliação das estimativas de π , σ_1^2 e σ_2^2 é relativamente difícil, em virtude de termos três configurações diferentes da normal contaminada que levam à distribuição normal. Devido a este fato, optamos por calcular a estatística U . Em 49.7% das séries, a estatística U encontrada foi igual à zero. Um outro ponto interessante é que 99.1 % das vezes a estatística U forneceu um valor inferior à .14 (o valor mínimo encontrado quando a estimativa do modelo GARCH - NC foi aplicada às séries geradas com erros provenientes das distribuições t-student e normal contaminada).

É importante ressaltar que, os dois modelos (GARCH - T e GARCH -NC) sugerem (GARCH - T através dos graus de liberdade estimados e o GARCH - NC através da estatística U) o modelo GARCH tradicional para as séries em questão.

Foram realizadas também, as estimativas dos percentis necessárias para construção dos intervalos com 90% (percentil 95), 95% (percentil 97.5) e 98% (percentil 99) de confiança. Os resultados destas estimativas é apresentado na tabela abaixo:

Tabela 5.3.1.2 - Resultados das estimativas dos percentis para as séries com erros normais

Intervalo de		90%	95%	98%
Real		1.645	1.960	2.325
Normal	Média	1.645	1.960	2.325
	Mediana	1.645	1.960	2.325
T - Student	Média	1.643	1.965	2.345
	Mediana	1.644	1.964	2.340
Normal	Média	1.645	1.965	2.339
Contaminada	Mediana	1.645	1.960	2.326

Podemos observar, que os percentis foram adequadamente estimados quando as suposições sobre a distribuição não eram corretas, ou seja, os modelos GARCH - T e GARCH - NC estimaram de forma adequada os percentis de interesse.

No apêndice A pode ser encontrada uma tabela de decis de cada uma das estimativas (parâmetros e percentis) acima.

Por último, fizemos a previsão por intervalo do último valor gerado com 90, 95 e 98 % de confiança. A eficácia da previsão por intervalo foi analisada através da contagem de vezes em que o intervalo não contém o valor previsto, e os resultados são os que seguem:

Tabela 5.3.1.3 - Resultados da contagem de pontos fora

NC	90%	95%	98%
Valor Esperado	100	50	20
Normal	101	49	25
T - Student	100	54	22
Normal Cont.	100	55	22

Vale a pena observar que as quantidades de pontos fora, sempre estiveram próximas aos valores esperados. Testando as hipóteses de $p = 10\%$, 5% e 2% obtemos a tabela de p-values abaixo.

Tabela 5.3.1.4 - P-values dos testes $p = 10\%$, 5% e 2%

NC	90%	95%	98%
Normal	91.6%	94.2%	90.8%
T - Student	100.0%	61.2%	73.5%
Normal Cont.	100.0%	51.4%	73.5%

Claramente, as hipóteses nulas não devem ser rejeitadas, pois os p-values encontrados sempre tiveram valores altos.

5.3.2 Estimando com erros t-student

Após estimar os parâmetros das séries geradas com erros seguindo a distribuição t - student com 5 graus de liberdade, construímos a tabela que segue:

Tabela 5.3.2.1 - Resultados das estimativas para séries com erros t-student com 5 graus de liberdade

Estimado com Suposição		β_1	γ	α_1	ϕ_1	GL	π	σ_1^2	σ_2^2	U
Valor Real		-0.07	0.0001	0.1	0.8	5				
Normal	Média	-0.0691	0.00012	0.105	0.768	-	-	-	-	-
	Médiana	-0.0682	0.00011	0.098	0.790	-	-	-	-	-
	EQM ^{1/2}	0.0377	0.00008	0.049	0.115	-	-	-	-	-
T - Student	Média	-0.0693	0.00012	0.101	0.781	5.28	-	-	-	-
	Médiana	-0.0672	0.00010	0.098	0.793	5.14	-	-	-	-
	EQM ^{1/2}	0.0336	0.00007	0.033	0.092	0.97	-	-	-	-
Normal Contaminada	Média	-0.0695	0.00012	0.102	0.778	-	0.8084	0.5814	3.6321	0.7664
	Médiana	-0.0681	0.00010	0.098	0.794	-	0.8299	0.5834	2.9685	0.5684
	EQM ^{1/2}	0.0340	0.00007	0.035	0.099	-	0.1113	0.0970	3.0370	1.3256

Assim como ocorreu com os dados gerados com distribuição normal, os erros das estimativas dos parâmetros β_1 , γ e α_1 não são muito diferentes independentemente da suposição

de distribuição feita. A excessão é feita ao parâmetro ϕ_1 que teve o erro quadrático médio ligeiramente maior quando os séries foram estimadas com o modelo GARCH. Em geral, as estimativas foram, em média e mediana, bastante razoáveis.

Observamos que as estimativas dos graus de liberdade também foram boas, 78.6% dos casos entre 4 e 6. Vale a pena lembrar que quando estimamos o modelo GARCH - T utilizando as séries com erro normais, 58% das vezes tivemos os graus de liberdade estimado em 102.

No modelo GARCH - NC, a estatística U apresentou valores maiores do que quando as séries foram geradas com a distribuição gaussiana, o que é uma indicação de que o modelo GARCH usual não é uma boa aproximação para as séries em questão.

As estimativas dos percentis também foram feitas e os resultados são os que seguem:

Tabela 5.3.2.2 - Resultados das estimativas dos percentis para as séries com erros t - student de 5 graus de liberdade

		90%	95%	98%
	Real	1.561	1.991	2.606
Normal	Média	1.645	1.960	2.325
	Mediana	1.645	1.960	2.325
T - Student	Média	1.562	1.989	2.597
	Mediana	1.566	1.993	2.600
Normal	Média	1.550	1.989	2.669
Contaminada	Mediana	1.553	1.997	2.682

Podemos observar, que os percentis foram adequadamente estimados quando os modelos utilizados eram os GARCH - T e GARCH - NC. Quando utilizamos o modelo GARCH usual os percentis utilizados são os mesmos da distribuição normal.

Podemos notar também que, quando usamos o modelo GARCH usual, o percentil 95 (utilizado para fazer intervalos de 90%) é super-estimado e os demais são sub-estimados. Tal situação levará à construção de intervalos de confiança que não possuirão 90%, 95% e 98% de confiança.

No apêndice A pode ser encontrada uma tabela de decis de cada uma das estimativas acima.

Por último, fizemos a previsão por intervalo do último valor gerado com 90, 95 e 98 % de confiança. A eficácia da previsão por intervalo foi analisada através da contagem de vezes em que o intervalo não contém o valor previsto, e os resultados são os que seguem:

Tabela 5.3.2.3 - Resultados da contagem de pontos fora

Nível de Conf.	90%	95%	98%
Valor Esperado	100	50	20
Normal	92	57	34
T - Student	104	56	21
Normal Cont.	104	56	21

Vale a pena observar que as quantidades de pontos fora, tendem a estar acima do valor esperado quando o intervalo foi construído com 98% de confiança quando o modelo GARCH usual foi utilizado. Quando os modelos GARCH - T e GARCH - NC foram utilizados, a contagem de pontos fora ficou próxima dos valores esperados. Testando as hipóteses de $p = 10\%$, 5% e 2% obtemos a tabela de p-values abaixo.

Tabela 5.3.2.4 - P-values dos testes $p = 10\%$, 5% e 2%

Nível de Conf.	90%	95%	98%
Normal	46.0%	34.5%	0.3%
T - Student	67.3%	42.5%	91.0%
Normal Cont.	67.3%	42.5%	91.0%

Como podemos observar, a hipótese $p = 2\%$ quando, o modelo GARCH usual foi utilizado, deve ser rejeitada e as demais hipóteses não foram rejeitadas. A não rejeição das hipóteses, no caso em que o modelo GARCH usual foi utilizado, para $p = 10\%$ e 5% indicam que os percentis correspondentes, embora mal estimados, não distanciam dos verdadeiros percentis o suficiente para comprometer a previsão por intervalo quando os erros são provenientes da distribuição t-student com 5 graus de liberdade.

5.3.3 Estimando com erros normais contaminados (normal cont I)

Após estimar os parâmetros das séries geradas com erros seguindo a normal cont I construímos a tabela que segue abaixo:

Tabela 5.3.3.1 - Resultados das estimativas para séries com erros com distribuição normal contaminada I

Estimado com Suposição		β_1	γ	α_1	ϕ_1	GL	π	σ_1^2	σ_2^2	U
Valor Real		-0.07	0.0001	0.1	0.8		0.8769	0.5	4.8	1.57
Normal	Média	-0.0691	0.00013	0.111	0.759	-	-	-	-	-
	Mediana	-0.0696	0.00011	0.102	0.787	-	-	-	-	-
	EQM ^{1/2}	0.0413	0.00010	0.065	0.139	-	-	-	-	-
T - Student	Média	-0.0694	0.00012	0.107	0.787	3.50	-	-	-	-
	Mediana	-0.0684	0.00011	0.104	0.795	3.45	-	-	-	-
	EQM ^{1/2}	0.0318	0.00005	0.036	0.073	0.39	-	-	-	-
Normal Contaminada	Média	-0.0696	0.00011	0.102	0.790	-	0.8763	0.4718	4.9842	1.6754
	Mediana	-0.0687	0.00010	0.099	0.798	-	0.8778	0.4709	4.7429	1.5326
	EQM ^{1/2}	0.0312	0.00005	0.032	0.066	-	0.0297	0.0540	1.2749	0.7026

Podemos observar que os EQM's das estimativas dos parâmetros β_1 , γ , α_1 e ϕ_1 são maiores quando a suposição de normalidade dos erros foi feita e que, aparentemente, não existem grandes diferenças, entre os EQM's encontrados destas estimativas, quando os modelos GARCH - T e GARCH - NC foram utilizados.

Observamos também que, quando o modelo GARCH tradicional foi utilizado, as estimativas do parâmetro α_1 tiveram uma leve tendência a valores altos (5.7% foram maiores que 0.2 enquanto apenas 1.2% e 0.7% das estimativas foram maiores que 0.2 quando os modelos GARCH - T e GARCH - NC foram utilizados respectivamente).

Uma outra observação que pode ser feita é que, quando o modelo GARCH tradicional foi utilizado, as estimativas do parâmetro ϕ_1 tiveram uma leve tendência a valores baixos (8% foram menores que 0.6 enquanto apenas 1.5% e 1.2% das estimativas foram inferiores à 0.2 quando os modelos GARCH - T e GARCH - NC, respectivamente, foram utilizados).

Notamos que, quando estimamos o modelo GARCH - NC, as estimativas dos parâmetros π , σ_1^2 e σ_2^2 assumiram valores bastante razoáveis em média e mediana.

Por último, vemos que a estatística U apresentou valores maiores do que quando as séries foram geradas com a distribuição gaussiana, o que é uma indicação de que o modelo GARCH usual não é uma boa aproximação para as séries em questão.

As estimativas dos percentis também foram feitas e os resultados são os que seguem:

Tabela 5.3.3.2 - Resultados das estimativas dos percentis para as séries com erros gerados pela distribuição normal cont I

		90%	95%	98%
	Real	1.402	1.944	3.053
Normal	Média	1.645	1.960	2.325
	Mediana	1.645	1.960	2.325
T - Student	Média	1.442	1.911	2.646
	Mediana	1.448	1.919	2.654
Normal	Média	1.402	1.938	3.026
Contaminada	Mediana	1.410	1.943	3.029

Podemos observar, que os percentis foram adequadamente estimados quando o modelo utilizado é o GARCH - NC. O modelo GARCH - T apresentou tendência em sub-estimar, em média e mediana, o percentil 99 (utilizado para fazer intervalo de confiança 98%).

Quando utilizamos o modelo GARCH usual os percentis estimados são os mesmos da distribuição normal.

Podemos notar também que, quando usamos o modelo GARCH usual, o percentil 95 (utilizado para fazer intervalos de 90%) é super-estimado enquanto o percentil 99% (utilizado para fazer intervalo de confiança de 98%) é sub-estimado. Tal situação levará à construção de intervalos de confiança que não possuirão 90% e 98% de confiança.

No apêndice A pode ser encontrada uma tabela de decis de cada uma das estimativas acima.

Por último, fizemos a previsão por intervalo do último valor gerado com 90, 95 e 98 % de confiança. A eficácia da previsão por intervalo foi analisada através da contagem de vezes em que o intervalo não contém o valor previsto, e os resultados são os que seguem:

Tabela 5.3.3.3 - Resultados da contagem de pontos fora

NC	90%	95%	98%
Valor Esperado	100	50	20
Normal	75	55	39
T - Student	96	56	29
Normal Cont.	103	57	21

Vale a pena observar que as quantidades de pontos fora, tendem a estar acima do valor esperado quando o intervalo foi construído com 98% de confiança e o modelo GARCH usual foi utilizado. Observamos também, que a contagem de pontos fora está bem abaixo do esperado quando o intervalo foi construído com 90% de confiança e o modelo GARCH usual foi utilizado.

O modelo GARCH - T, como era de se esperar (devido à tabela 5.3.3.2), apresentou uma quantidade de pontos fora acima do esperado quando intervalos de 98% foram construídos. Já com o modelo GARCH - NC, tivemos contagens de pontos fora próximas aos valores esperados.

Testando as hipóteses de $p = 10\%$, 5% e 2% obtemos a tabela de p-values abaixo.

Tabela 5.3.3.4 - P-values dos testes $p = 10\%$, 5% e 2%

Nível de Conf.	90%	95%	98%
Normal	1.2%	51.4%	0.0%
T - Student	75.2%	42.5%	5.3%
Normal Cont.	75.2%	24.4%	91.0%

Como podemos observar, a hipótese $p = 10\%$ e $p = 2\%$ deve ser rejeitada quando o modelo GARCH usual foi utilizado. A hipótese de $p = 5\%$ não foi rejeitada.

Vale a pena notar que a hipótese $p = 2\%$, quando o modelo GARCH - T foi utilizado, apresenta um valor baixo para o p-value. Embora tal p-value não nos leve a rejeição da hipótese (ao nível de 5%), o baixo valor do p-value é uma indicação de que devemos observar este resultado com cautela.

5.3.4 Estimando com erros normais contaminados (normal cont II)

Após estimar os parâmetros das séries geradas com erros seguindo a normal cont II construímos a tabela que segue abaixo:

Tabela 5.3.4.1 - Resultados das estimativas das séries com erros normais

Estimado com Suposição		β_1	γ	α_1	ϕ_1	GL	π	σ_1^2	σ_2^2	U
Valor Real		-0.07	0.0001	0.1	0.8					
Normal	Média	-0.0692	0.00012	0.104	0.774	-	-	-	-	-
	Mediana	-0.0685	0.00011	0.100	0.790	-	-	-	-	-
	EQM ^{1/2}	0.0366	0.00008	0.039	0.101	-	-	-	-	-
T - Student	Média	-0.0693	0.00012	0.104	0.780	5.15	-	-	-	-
	Mediana	-0.0673	0.00011	0.102	0.792	5.04	-	-	-	-
	EQM ^{1/2}	0.0334	0.00007	0.034	0.092	0.80	-	-	-	-
Normal Contaminada	Média	-0.0693	0.00012	0.102	0.781	-	0.6862	0.4747	2.3473	0.4084
	Mediana	-0.0677	0.00010	0.100	0.793	-	0.6871	0.4701	2.1536	0.3467
	EQM ^{1/2}	0.0333	0.00007	0.033	0.090	-	0.0969	0.0815	1.0444	0.3086

A exemplo do que ocorreu com os dados gerados com distribuição t - student, os erros das estimativas dos parâmetros β_1 , γ e α_1 não são muito diferentes independentemente da suposição de distribuição feita. A exceção é feita ao parâmetro ϕ_1 que teve o erro quadrático médio um pouco maior quando as séries foram estimadas com o modelo GARCH. Em geral, as estimativas foram, em média e mediana, bastante razoáveis.

Notamos que, quando estimamos o modelo GARCH - NC, as estimativas dos parâmetros π , σ_1^2 e σ_2^2 assumiram valores bastante razoáveis em média e mediana.

Por último, vemos que a estatística U apresentou valores maiores do que quando as séries foram geradas com a distribuição gaussiana, o que é uma indicação de que o modelo GARCH usual não é uma boa aproximação para as séries em questão.

As estimativas dos percentis também foram feitas e os resultados são os que seguem:

Tabela 5.3.4.2 - Resultados das estimativas dos percentis para as séries com erros gerados pela distribuição normal cont II

		90%	95%	98%
	Real	1.600	2.092	2.727
Normal	Média	1.645	1.960	2.325
	Mediana	1.645	1.960	2.325
T - Student	Média	1.559	1.988	2.601
	Mediana	1.562	1.992	2.605
Normal Contaminada	Média	1.599	2.077	2.710
	Mediana	1.603	2.086	2.707

Podemos observar, que os percentis foram adequadamente estimados quando o modelo utilizado é GARCH - NC. O modelo GARCH - T apresentou pequena tendência em sub-estimar, em média e mediana, o percentil 99 (utilizado para fazer intervalo de confiança 98%).

Quando utilizamos o modelo GARCH usual os percentis estimados são os mesmos da distribuição normal.

Podemos notar também que, quando usamos o modelo GARCH usual, os percentis 97.5 e 99 (utilizados para fazer intervalos com 95% e 98% de confiança respectivamente) são sub-estimados. Tal situação levará à construção de intervalos de confiança que não possuirão 95% e 98% de confiança.

No apêndice A pode ser encontrada uma tabela de decis de cada uma das estimativas acima.

Por último, fizemos a previsão por intervalo do último valor gerado com 90, 95 e 98 % de confiança. A eficácia da previsão por intervalo foi analisada através da contagem de vezes em que o intervalo não contém o valor previsto, e os resultados são os que seguem:

Tabela 5.3.4.3 - Resultados da contagem de pontos fora

NC	90%	95%	98%
Valor Esperado	100	50	20
Normal	95	64	40
T - Student	105	61	26
Normal Cont.	101	57	22

Vale a pena observar que as quantidades de pontos fora tendem a estar acima do valor esperado quando os intervalos foram construídos com 95% e 98% de confiança e o modelo GARCH usual foi utilizado. Observamos também, que a contagem de pontos fora está abaixo do esperado quando o intervalo foi construído com 90% de confiança e o modelo GARCH usual foi utilizado.

Quando o modelo GARCH - T foi utilizado e os intervalos de confiança construídos com 98% de confiança, observamos leve tendência de saída de pontos acima do esperado. A surpresa ficou para os intervalos construídos com 95% de confiança. Nestes, houve tendência de saídas de pontos acima do esperado.

Quando o modelo GARCH - NC foi utilizado tivemos as contagens de pontos fora próximas dos valores esperados.

Testando as hipóteses de $p = 10\%$, 5% e 2% obtemos a tabela de p-values abaixo.

Tabela 5.3.4.4 - P-values dos testes $p = 10\%$, 5% e 2%

Nível de Conf.	90%	95%	98%
Normal	67.3%	5.0%	0.0%
T - Student	59.8%	12.7%	21.3%
Normal Cont.	91.6%	34.5%	73.5%

Como podemos observar, a hipótese $p = 5\%$ e $p = 2\%$ deve ser rejeitada, ao nível de 5%, quando o modelo GARCH usual foi utilizado. As demais hipóteses não devem ser rejeitadas.

5.4 Comentários finais

Neste estudo de simulação percebemos que, em séries grandes (usamos 1000 pontos por série), as estimativas dos parâmetros da equação da variância são pouco afetadas pela suposição de distribuições dos erros padronizados. Em geral, as estimativas estiveram próximas dos valores reais mesmo quando foi realizada uma suposição de distribuição dos erros padronizados errônea.

Notamos também que, o modelo GARCH - NC é mais robusto que os demais no que se refere às estimativas dos percentis da distribuição dos erros e, por consequência, faz previsão por intervalo, da série em estudo, com confiança mais próxima da desejada.

O modelo GARCH- T é mais robusto, no que se refere às estimativas dos percentis da distribuição dos erros, que o modelo GARCH usual, porém, perde um pouco da precisão nas estimativas de percentis mais próximos dos extremos (percentil 99 por exemplo) quando a suposição de distribuição dos erros não é verdadeira. Vale a pena lembrar que, a distribuição t-student converge para a distribuição gaussiana quando o número de graus de liberdade vai para infinito e, portanto, neste caso, a suposição de distribuição dos erros converge para a verdadeira distribuição dos erros e as estimativas dos percentis tendem a ter boa precisão.

É importante ressaltar que o "preço" pago para ter maior precisão nas estimativas por intervalo, da série em estudo, é o aumento do número de parâmetros do modelo. O modelo GARCH - NC utilizado possui um parâmetro a mais que o modelo GARCH - T que, por sua vez, possui um parâmetro a mais que o modelo GARCH usual.

Capítulo 6

APLICAÇÃO EM UMA SÉRIE REAL

6.1 Introdução

No capítulo anterior fizemos uma simulação para comparar a eficácia da previsão por intervalo, da série estudada, dos modelos GARCH, GARCH - T e GARCH - NC. Neste capítulo iremos estimar esses três modelos em uma série real e analisar os resíduos para verificar se a suposição sobre a distribuição dos erros é razoável.

6.2 Dados

A série que será analisada neste capítulo é a ação PN do banco Itáu cujo código é ITAU4 na bolsa de valores de São Paulo. O período utilizado para análise compreende de 01 de agosto de 1994 até 14 de setembro de 1999 totalizando 1264 dados.

O gráfico do preço unitário da ação é apresentado abaixo:

Gráfico 6.2.1 - Ação ITAÚ PN

Os modelos são aplicados na variável retorno que é definido da forma que segue:

$$R_t = \ln \left(\frac{P_t}{P_{t-1}} \right) \quad (6.1)$$

onde P_t é o preço da ação na data t .

Observe que temos 1263 dados da variável retorno. O gráfico dos retornos é apresentado abaixo:

Gráfico 6.2.2 - Retornos do ITAÚ PN

6.3 Análise dos Dados

Os modelos GARCH, GARCH - T e GARCH - NC foram aplicados à série de retornos definida na seção anterior. O modelo aplicado é da forma que segue abaixo:

$$y_t = \sum_{j=1}^d \beta y_{t-j} + \varepsilon_t \quad (6.2)$$

$$\varepsilon_t = \sqrt{h_t} v_t \quad (6.3)$$

$$h_t = \gamma + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \phi_j h_{t-j} \quad (6.4)$$

onde v_t segue uma das três distribuições: t-student padronizada, normal ou normal contaminada.

Os três modelos de interesse foram aplicados com $d = 0, 1, 2$; $p = 1, 2, 3$ e $q = 0, 1, 2, 3$ e os critério de Akaike (AIC) e Schwarz (BIC) (ver Harvey 1993) foram utilizados para selecionar o número adequado dos parâmetros d, p e q .

Por último foi realizada a análise de resíduos para averiguar se o modelo escolhido é adequado.

Em todos os três modelos, os critérios AIC e BIC indicam para o modelo que segue:

$$y_t = \beta_1 y_{t-1} + \varepsilon_t \quad (6.5)$$

$$\varepsilon_t = \sqrt{h_t} v_t \quad (6.6)$$

$$h_t = \gamma + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \phi_2 h_{t-2} \quad (6.7)$$

Nas seções que seguem iremos apresentar os resultados encontrados com cada um dos modelos.

6.3.1 Modelo GARCH

Quando aplicamos o modelo GARCH obtivemos as estimativas apresentadas na tabela abaixo:

Tabela 6.3.1.1 - Estimativas do Modelo GARCH

Parâmetros	$\hat{\beta}_1$	$\hat{\gamma}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\phi}_2$
Valores	.1960	.000096	.2744	.0710	.5573

Tabela 6.3.1.2 - Log da Verossimilhança e critérios de AIC e BIC

Log Verossimilhança.	AIC	BIC
2849.82	-5689.63	-5663.92

Na tabela que segue, apresentamos os testes Ljung e Box para os resíduos padronizados (v_t) e resíduos padronizados ao quadrado (v_t^2).

Tabela 6.3.1.3 - Testes de Ljung e Box para os resíduos padronizados e resíduos padronizados ao quadrado

Lag	Estatística Qui-Quadrado ($\hat{v}_t$)	p-value	Estatística Qui-Quadrado ($\hat{v}_t^2$)	p-value
12	12.09	.438	5.77	.927
24	21.06	.635	17.24	.938

Como podemos notar, em ambos os casos ($\hat{v}_t$ e $\hat{v}_t^2$) não rejeitamos a hipótese nula de ausência de auto-correlação nos resíduos padronizados ($\hat{v}_t$) e resíduos padronizados ao quadrado ($\hat{v}_t^2$).

Quando estimamos o modelo GARCH estamos supondo que os erros padronizados seguem a distribuição normal com média zero e variância 1. Uma forma de checar esta suposição é através da construção do gráfico Q x Q plot.

Gráfico 6.3.1.1 - QxQ Plot dos resíduos padronizados

No gráfico acima, temos a indicação que os resíduos padronizados não seguem a distribuição gaussiana. Podemos perceber, através dos extremos, a indicação da existência de caudas pesadas. Com o objetivo de testar se os resíduos padronizados seguem a distribuição gaussiana, foi feito a tabela que segue:

Tabela 6.3.1.4 - Valores esperados e observados dos resíduos

Intervalos	Valores Esperados	Valores Observados
$-\infty < \rightarrow -3$	1.7	8
$-3 < \rightarrow -2$	27.0	25
$-2 < \rightarrow -1$	171.6	120
$-1 < \rightarrow 0$	431.1	445
$0 < \rightarrow +1$	431.1	481
$+1 < \rightarrow +2$	171.6	146
$+2 < \rightarrow +3$	27.0	32
$+3 < \rightarrow +\infty$	1.7	6

Fazendo um teste de hipótese para verificar a aderência, temos que a estatística qui-quadrado é igual a 60.72 e o p-value igual a .000, portanto, rejeitamos a hipótese de que os resíduos padronizados sigam a distribuição normal padrão.

Utilizando as estimativas apresentadas na tabela 6.3.1.1, construímos intervalos de confiança para a série de retornos com 90%, 95% e 98% de confiança. O gráfico com intervalos com 98% de confiança é apresentado a seguir:

Gráfico 6.3.1.2 - Intervalo de 98% de confiança

Novamente, fizemos a contagem de pontos que incidiram fora dos intervalos construídos, testamos as hipóteses $p = 10\%$, 5% e 2% e os resultados são os que seguem:

Tabela 6.3.1.5 - Contagem de pontos fora do intervalo

Intervalos	Percentil	Pontos Fora	Pontos Fora Esperado	p-value
90%	1.645	130	126	.742
95%	1.960	73	63	.204
98%	2.325	44	25	.000

Podemos confirmar a existência de caudas pesadas através do excessivo número de pontos

fora do intervalo de confiança construído com 98% de confiança. Vale a pena ressaltar que a hipótese $p = 2\%$ deve ser rejeitada.

Devemos ressaltar que, em virtude de termos construído intervalos de confiança para os mesmos pontos que foram utilizados nas estimativas dos parâmetros, era de se esperar resultados melhores que os apresentados na tabela 6.3.1.3. A pobreza dos resultados deve-se a falsa suposição de distribuição normal dos erros padronizados.

6.3.2 Modelo GARCH-T

Quando aplicamos o modelo GARCH-T obtivemos as estimativas apresentadas na tabela abaixo:

Tabela 6.3.2.1 - Estimativas do Modelo GARCH-T

Parâmetros	$\hat{gl}$	$\hat{\beta}_1$	$\hat{\gamma}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\phi}_2$
Valores	4.24	.1454	.000087	.2899	.0866	.5762

Tabela 6.3.2.2 - Log da Verossimilhança e critérios de AIC e BIC

Log Verossimilhança.	AIC	BIC
2894.29	-5776.68	-5745.74

Na tabela que segue, apresentamos os testes Ljung e Box para os resíduos padronizados (v_t) e resíduos padronizados ao quadrado (v_t^2).

Tabela 6.3.2.3 - Testes de Ljung e Box para os resíduos padronizados e resíduos padronizados ao quadrado

Lag	Estatística Qui-Quadrado ($\hat{v}_t$)	p-value	Estatística Qui-Quadrado ($\hat{v}_t^2$)	p-value
12	15.99	.192	6.30	.900
24	24.49	.434	16.96	.850

Como podemos notar, em ambos os casos ($\hat{v}_t$ e $\hat{v}_t^2$) não rejeitamos a hipótese nula de ausência de auto-correlação nos resíduos padronizados ($\hat{v}_t$) e resíduos padronizados ao quadrado ($\hat{v}_t^2$).

Quando estimamos o modelo GARCH - T supomos que os erros padronizados seguem a distribuição t - student padronizada. Uma forma de checar esta suposição é através da construção do gráfico Q x Q plot onde os graus de liberdade da distribuição t - student é dado por $\hat{g}l$.

Gráfico 6.3.2.1 - Q x Q Plot dos resíduos padronizados

Ao contrário do que ocorreu com os resíduos gerados pelo modelo GARCH, o gráfico acima indica que os resíduos padronizados, resultantes da aplicação do modelo GARCH - T, seguem a distribuição t-student, padronizada, com 4.24 graus de liberdade.

Com o objetivo de testar se os resíduos padronizados seguem a distribuição t - student com 4.24 graus de liberdade, foi feita a tabela que segue:

Tabela 6.3.2.4 - Valores esperados e observados dos resíduos

Intervalos	Valores Esperados	Valores Observados
$-\infty < \rightarrow -3$	8.1	7
$-3 < \rightarrow -2$	22.3	21
$-2 < \rightarrow -1$	118.5	116
$-1 < \rightarrow 0$	482.5	461
$0 < \rightarrow +1$	482.5	486
$+1 < \rightarrow +2$	118.5	139
$+2 < \rightarrow +3$	22.3	27
$+3 < \rightarrow +\infty$	8.1	6

Fazendo um teste de hipótese para verificar a aderência, temos que a estatística qui-quadrado é igual a 6.34 e o p-value igual a .500 portanto, não rejeitamos a hipótese de que os resíduos padronizados sigam a distribuição t-student com 4.24 graus de liberdade.

Utilizando as estimativas apresentadas na tabela 6.3.2.1, construímos intervalos de confiança para a série de retornos com 90%, 95% e 98% de confiança. O gráfico com intervalos com 98% de confiança é apresentado a seguir:

Gráfico 6.3.2.2 - Intervalo de 98% de confiança

Novamente, fizemos a contagem de pontos que incidiram fora dos intervalos construídos, testamos as hipóteses $p = 10\%$, 5% e 2% e os resultados são os que seguem:

Tabela 6.3.2.5 - Contagem de pontos fora do intervalo

Intervalos	Percentil	Pontos Fora	Pontos Fora Esperado	p-value
90%	1.525	143	126	.122
95%	1.974	64	63	.949
98%	2.640	22	25	.614

Como pode ser visto na tabela acima, as hipóteses $p = 10\%$, 5% e 2% não são rejeitadas devido aos altos valores dos p-values.

6.3.3 Modelo GARCH-NC

Quando aplicamos o modelo GARCH-NC obtivemos as estimativas apresentadas na tabela abaixo:

Tabela 6.3.3.1 - Estimativas do Modelo GARCH-NC

Parâmetros	$\hat{\pi}$	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\hat{\beta}_1$	$\hat{\gamma}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\phi}_2$
Valores	.363	.164	1.476	.1355	.000098	.2872	.0789	.5397

Tabela 6.3.2.2 - Log da Verossimilhança e critérios de AIC e BIC

Log Verossimilhança.	AIC	BIC
2900.11	-5786.22	-5750.24

Na tabela que segue, apresentamos os testes Ljung e Box para os resíduos padronizados (v_t) e resíduos padronizados ao quadrado (v_t^2).

Tabela 6.3.3.3 - Testes de Ljung e Box para os resíduos padronizados e resíduos padronizados ao quadrado

Lag	Estatística Qui-Quadrado ($\hat{v}_t$)	p-value	Estatística Qui-Quadrado ($\hat{v}_t^2$)	p-value
12	17.78	.123	6.91	.864
24	26.26	.345	17.40	.831

Como podemos notar, em ambos os casos ($\hat{v}_t$ e $\hat{v}_t^2$) não rejeitamos a hipótese nula de ausência de auto-correlação nos resíduos padronizados ($\hat{v}_t$) e resíduos padronizados ao quadrado ($\hat{v}_t^2$).

Quando estimamos o modelo GARCH - NC supomos que os erros padronizados seguem a distribuição normal contaminada. Uma forma de checar esta suposição é através da construção do gráfico Q x Q plot onde os parâmetros da normal contaminada são dados por $\hat{\pi}$, $\hat{\sigma}_1^2$ e $\hat{\sigma}_2^2$.

Gráfico 6.3.3.1 - Q x Q Plot dos resíduos padronizados.

O gráfico acima fornece indicação de que os resíduos provenientes da aplicação do modelo GARCH - NC seguem a distribuição normal contaminada com parâmetros dados pela tabela 6.3.3.1.

Com o objetivo de testar se os resíduos padronizados seguem a distribuição normal contaminada com $\hat{\pi} = .363$, $\hat{\sigma}_1^2 = .164$ e $\hat{\sigma}_2^2 = 1.476$, foi feita a tabela de contingência que segue:

Tabela 6.3.3.4 - Valores esperados e observados dos resíduos

Intervalos	Valores Esperados	Valores Observados
$-\infty <-> -3$	5.4	8
$-3 <-> -2$	34.7	23
$-2 <-> -1$	128.1	120
$-1 <-> 0$	463.3	453
$0 <-> +1$	463.3	475
$+1 <-> +2$	128.1	145
$+2 <-> +3$	34.7	33
$+3 <-> +\infty$	5.4	6

Fazendo um teste de hipótese para verificar a aderência, temos que a estatística qui-quadrado é igual a 8.61 e o p-value igual a .282 portanto, não rejeitamos a hipótese de que os resíduos padronizados sigam a distribuição normal contaminada com $\hat{\pi} = .363$, $\hat{\sigma}_1^2 = .164$ e $\hat{\sigma}_2^2 = 1.476$.

Utilizando as estimativas apresentadas na tabela 6.3.3.1, construímos intervalos de confiança para a série de retornos com 90%, 95% e 98% de confiança. O gráfico com intervalos com 98% de confiança é apresentado a seguir:

Gráfico 6.3.3.2 - Intervalo de 98% de confiança

Novamente, fizemos a contagem de pontos que incidiram fora dos intervalos construídos, testamos as hipóteses $p = 10\%$, 5% e 2% e os resultados são os que seguem:

Tabela 6.3.3.5 - Contagem de pontos fora do intervalo

Intervalos	Percentil	Pontos Fora	Pontos Fora Esperado	p-value
90%	1.720	116	126	.373
95%	2.138	56	63	.401
98%	2.615	26	25	.920

Como pode ser visto na tabela acima, as hipóteses $p = 10\%$, 5% e 2% não são rejeitadas devido aos altos valores dos p-values.

6.4 Comparação da Volatilidades Estimadas

Com o intuito ilustrativo, iremos apresentar o gráfico contendo as volatilidades estimadas pelos três modelos já apresentados a seguir:

Gráfico 6.4.1 - Comparação das volatilidades estimadas pelos modelos GARCH, GARCH-T e GARCH-NC

Como podemos observar, as volatilidades estimadas não são muito diferentes independentemente do modelo utilizado. Este fato já era esperado em virtude dos valores apresentados nas tabelas 6.3.1.1, 6.3.2.1 e 6.3.3.1 estarem todos muito próximos.

A suposição sobre a distribuição dos erros padronizados se revelou importante quando o interesse é a previsão por intervalo, porém não se apresentou importante no que se refere às estimativas de volatilidades. A diferença está nas estimativas dos percentis de interesse.

É importante notar que os critérios de AIC e BIC sinalizam que o modelo GARCH-NC é o mais adequado para a série estudada como pode ser visto na tabela abaixo.

Tabela 6.4.1 - Critérios de AIC e BIC

Modelos	AIC	BIC
GARCH	-5689.63	-5663.92
GARCH - T	-5776.68	-5745.74
GARCH - NC	-5786.22	-5750.24

Devemos ressaltar que, não podemos afirmar que com amostras menores (neste trabalho a amostra possui 1263 dados) a suposição sobre a distribuição sobre os erros padronizados não se revelará importante nas estimativas das volatilidades.

Capítulo 7

Comentários Finais

Analisando os resultados apresentados ao longo do capítulo 6, percebemos que os três modelos estimaram valores similares para os parâmetros que forma a equação da variância (h_t). Este resultado não é surpreendente pois as simulações apresentadas no capítulo V já sinalizavam neste sentido.

Nas simulações realizadas no capítulo V tivemos a indicação de que o modelo GARCH - NC estima os percentis com precisão ligeiramente superior ao GARCH - T, porém no caso de séries reais, as quais desconhecemos o processo gerador, não podemos avaliar as estimativas dos percentis como fizemos no capítulo V. Em virtude deste fato, devemos nos restringir à contagem de pontos que incidiram fora dos intervalos de confiança.

Os testes de hipótese, aplicados ao longo do capítulo 6, não fornecem indicações de que os intervalos de confiança construídos com os modelos GARCH - T e GARCH NC possuam confiança diferente da desejada. Em virtude disto, podemos afirmar que os modelos em questão se mostraram adequados no que se refere à construção de intervalos de confiança para a série estudada. Vale a pena lembrar que os critérios de AIC e BIC sinalizam que o modelo GARCH-NC é o mais adequado para a série ITAÚ PN.

O modelo GARCH não se mostrou robusto no que se refere a construção de intervalos de confiança.

Em virtude de todos os resultados apresentados ao longo dos capítulos V e VI, podemos

sugerir que, quando houver interesse em previsões por intervalo os modelos GARCH - T e GARCH - NC fornecem resultados mais precisos.

É importante ressaltar que os estudos realizados ao longo deste trabalho envolveram amostras grandes ($N = 1000$) e portanto, as conclusões se restringem à análises que envolvam amostras desta magnitude ou maior. Não podemos afirmar que em amostras menores que as utilizadas, a robustez dos modelos GARCH - NC e GARCH - T serão similares nem mesmo que os parâmetros da equação da variância (h_t) seriam similares nos três modelos estudados. É possível que em amostras menores a suposição de distribuição dos resíduos padronizados influencie mais fortemente as estimativas dos parâmetros da equação da variância.

Apêndice A

Tabelas de Decis

A.1 Erros seguindo distribuição normal

Tabela A1.1 - Erros Normais / Suposição de distribuição normal dos erros

	β_1	γ	α_1	ϕ_1
Real	-0,07	0,0001	0,1	0,8
Mínimo	-0,1712	0,00001	0,021	0,031
1º Decil	-0,1133	0,00006	0,065	0,693
2º Decil	-0,0975	0,00007	0,076	0,732
3º Decil	-0,0873	0,00009	0,084	0,757
4º Decil	-0,0773	0,00009	0,092	0,778
5º Decil	-0,0686	0,00011	0,100	0,792
6º Decil	-0,0608	0,00012	0,106	0,806
7º Decil	-0,0526	0,00013	0,116	0,823
8º Decil	-0,0410	0,00015	0,125	0,841
9º Decil	-0,0239	0,00018	0,137	0,862
Máximo	0,0366	0,00072	0,194	0,972

As estimativas dos percentis são sempre iguais aos percentis da

distribuição normal (0;1) - (2.325, 1.960 e 1.645 para 98%, 95% e 90% respectivamente)

Tabela A1.2 -. Erros Normais / Suposição de distribuição t-student dos erros

	β_1	γ	α_1	ϕ_1	GL	98%	95%	90%
Real	0,07	0,0001	0,1	0,8	∞ (102)	2,325	1,960	1,645
Mínimo	-0,1717	0,00001	0,021	0,034	10,858	2,340	1,964	1,624
1º Decil	-0,1137	0,00006	0,064	0,692	41,701	2,340	1,964	1,641
2º Decil	-0,0976	0,00007	0,076	0,734	82,992	2,340	1,964	1,643
3º Decil	-0,0875	0,00009	0,083	0,757	83,014	2,340	1,964	1,643
4º Decil	-0,0771	0,00009	0,093	0,778	83,597	2,340	1,964	1,643
5º Decil	-0,0684	0,00011	0,100	0,791	102,000	2,340	1,964	1,644
6º Decil	-0,0604	0,00012	0,106	0,806	102,000	2,343	1,965	1,644
7º Decil	-0,0526	0,00013	0,115	0,823	102,000	2,343	1,965	1,644
8º Decil	-0,0416	0,00015	0,125	0,842	102,000	2,343	1,965	1,644
9º Decil	-0,0236	0,00018	0,137	0,863	102,000	2,360	1,970	1,644
Máximo	0,0367	0,00072	0,196	0,973	102,000	2,460	1,991	1,644

Tabela A1.3 - Erros Normais / Suposição de distribuição normal contaminada dos erros

	β_1	γ	α_1	ϕ_1	π	σ_1^2	σ_2^2	U	% 98	%95	%90
Real	0,07	0,0001	0,1	0,8	*	*	*	0,00	2,325	1,960	1,645
Mínimo	-0,1712	0,00001	0,021	0,031	0,0124	0,0010	1,0000	0,0000	2,315	1,943	1,618
1º Decil	-0,1128	0,00006	0,064	0,692	0,0829	0,3514	1,0000	0,0000	2,326	1,960	1,641
2º Decil	-0,0978	0,00007	0,076	0,734	0,0853	0,6962	1,0000	0,0000	2,326	1,960	1,645
3º Decil	-0,0876	0,00009	0,084	0,757	0,0862	0,8987	1,0002	0,0000	2,326	1,960	1,645
4º Decil	-0,0772	0,00009	0,092	0,778	0,0936	0,9536	1,0040	0,0000	2,326	1,960	1,645
5º Decil	-0,0687	0,00010	0,100	0,792	0,1227	0,9970	1,0111	0,0001	2,326	1,960	1,645
6º Decil	-0,0605	0,00012	0,106	0,807	0,3781	0,9994	1,0285	0,0009	2,327	1,960	1,645
7º Decil	-0,0524	0,00013	0,115	0,823	0,9129	0,9997	1,0562	0,0181	2,340	1,962	1,645
8º Decil	-0,0408	0,00015	0,125	0,842	0,9144	0,9999	1,1084	0,0328	2,353	1,970	1,645
9º Decil	-0,0238	0,00018	0,137	0,863	0,9167	1,0000	1,4046	0,0520	2,370	1,981	1,652
Máximo	0,0366	0,00072	0,195	0,973	0,9983	1,0000	15,3878	0,3605	2,482	2,042	1,674

* Podemos ter três situações ($P = 1, \sigma_1^2 = 1$); ($P = 0, \sigma_2^2 = 1$); ($\sigma_1^2 = \sigma_2^2 = 1$)

A.2 Erros seguindo distribuição t - student

Tabela A2.1 - Erros t - student / Suposição de distribuição normal dos erros

	β_1	γ	α_1	ϕ_1
Real	-0,07	0,0001	0,1	0,8
Mínimo	-0,1815	0,00001	0,019	0,000
1º Decil	-0,1164	0,00005	0,055	0,647
2º Decil	-0,1003	0,00007	0,069	0,708
3º Decil	-0,0896	0,00008	0,079	0,741
4º Decil	-0,0787	0,00009	0,089	0,768
5º Decil	-0,0682	0,00011	0,098	0,790
6º Decil	-0,0593	0,00012	0,108	0,808
7º Decil	-0,0504	0,00013	0,122	0,827
8º Decil	-0,0365	0,00016	0,137	0,848
9º Decil	-0,0215	0,00021	0,160	0,879
Máximo	0,0752	0,00066	0,652	0,973

As estimativas dos percentis são sempre iguais aos percentis da distribuição normal (0;1) - (2.325, 1.960 e 1.645 para 98%, 95% e 90% respectivamente)

Tabela A2.2 -. Erros t - student / Suposição de distribuição t-student dos erros

	β_1	γ	α_1	ϕ_1	GL	98%	95%	90%
Real	-0,07	0,0001	0,1	0,8	5	2,606	1,991	1,561
Mínimo	-0,1758	0,00001	0,020	0,000	3,167	2,470	1,875	1,398
1º Decil	-0,1117	0,00006	0,060	0,689	4,241	2,551	1,974	1,525
2º Decil	-0,0988	0,00007	0,073	0,733	4,536	2,569	1,983	1,541
3º Decil	-0,0869	0,00008	0,082	0,761	4,754	2,580	1,987	1,551
4º Decil	-0,0764	0,00009	0,091	0,778	4,907	2,592	1,990	1,557
5º Decil	-0,0672	0,00010	0,098	0,793	5,140	2,600	1,993	1,566
6º Decil	-0,0606	0,00011	0,108	0,808	5,342	2,611	1,995	1,572
7º Decil	-0,0525	0,00013	0,117	0,824	5,620	2,617	1,996	1,579
8º Decil	-0,0428	0,00015	0,129	0,844	5,907	2,627	1,997	1,585
9º Decil	-0,0255	0,00018	0,144	0,868	6,435	2,640	1,998	1,594
Máximo	0,0358	0,00071	0,225	0,964	10,140	2,659	1,999	1,622

Tabela A2.3 - Erros t - student / Suposição de distribuição normal contaminada dos erros

	β_1	γ	α_1	ϕ_1	π	σ_1^2	σ_2^2	U	% 98	%95	%90
Real	0,07	0,0001	0,1	0,8	*	*	*	0,00	2,606	1,991	1,561
Mínimo	-0,1829	0,00001	0,021	0,000	0,3476	0,2524	1,3743	0,1414	1,852	1,525	1,269
1º Decil	-0,1130	0,00006	0,060	0,679	0,6551	0,4554	1,9831	0,2764	2,519	1,871	1,483
2º Decil	-0,0993	0,00007	0,072	0,730	0,7193	0,5031	2,2244	0,3483	2,591	1,920	1,510
3º Decil	-0,0867	0,00008	0,081	0,757	0,7642	0,5329	2,4818	0,4237	2,627	1,949	1,527
4º Decil	-0,0767	0,00009	0,090	0,777	0,7989	0,5591	2,7288	0,4908	2,657	1,976	1,541
5º Decil	-0,0681	0,00010	0,098	0,794	0,8299	0,5834	2,9685	0,5684	2,682	1,997	1,553
6º Decil	-0,0608	0,00011	0,107	0,812	0,8526	0,6110	3,2643	0,6598	2,705	2,021	1,566
7º Decil	-0,0533	0,00013	0,119	0,827	0,8790	0,6342	3,6748	0,7705	2,734	2,044	1,578
8º Decil	-0,0424	0,00015	0,131	0,844	0,9023	0,6601	4,2369	0,9343	2,767	2,065	1,594
9º Decil	-0,0260	0,00019	0,146	0,868	0,9365	0,6983	5,6833	1,3354	2,815	2,092	1,615
Máximo	0,0345	0,00065	0,316	0,960	0,9945	0,8773	66,0216	27,1086	3,069	2,219	1,698

A.3 Erros seguindo distribuição normal contaminada I

Tabela A3.1 - Erros normal contaminada I / Suposição de distribuição normal dos erros

	β_1	γ	α_1	ϕ_1
Real	-0,07	0,0001	0,1	0,8
Mínimo	-0,2452	0,00001	0,015	0,000
1º Decil	-0,1215	0,00005	0,050	0,618
2º Decil	-0,1030	0,00006	0,065	0,698
3º Decil	-0,0909	0,00008	0,077	0,733
4º Decil	-0,0798	0,00009	0,089	0,767
5º Decil	-0,0696	0,00011	0,102	0,787
6º Decil	-0,0584	0,00012	0,114	0,807
7º Decil	-0,0489	0,00014	0,128	0,826
8º Decil	-0,0335	0,00017	0,147	0,853
9º Decil	-0,0175	0,00023	0,179	0,885
Máximo	0,0656	0,00078	0,788	0,972

As estimativas dos percentis são sempre iguais aos percentis da distribuição normal (0;1) - (2.325, 1.960 e 1.645 para 98%, 95% e 90% respectivamente)

Tabela A3.2 -. Erros normal contaminada I / Suposição de distribuição t-student dos erros

	β_1	γ	α_1	ϕ_1	GL	98%	95%	90%
Real	-0,07	0,0001	0,1	0,8	5	3,053	1,944	1,402
Mínimo	-0,1736	0,00002	0,017	0,000	2,444	2,336	1,549	1,103
1º Decil	-0,1097	0,00006	0,064	0,709	3,032	2,624	1,846	1,367
2º Decil	-0,0971	0,00008	0,077	0,742	3,165	2,638	1,875	1,398
3º Decil	-0,0871	0,00009	0,087	0,766	3,277	2,646	1,895	1,420
4º Decil	-0,0764	0,00010	0,096	0,781	3,374	2,651	1,909	1,436
5º Decil	-0,0684	0,00011	0,104	0,795	3,454	2,654	1,919	1,448
6º Decil	-0,0615	0,00012	0,112	0,808	3,556	2,656	1,931	1,462
7º Decil	-0,0530	0,00013	0,123	0,824	3,669	2,657	1,941	1,476
8º Decil	-0,0433	0,00015	0,135	0,840	3,795	2,658	1,951	1,489
9º Decil	-0,0292	0,00018	0,151	0,865	4,023	2,658	1,964	1,509
Máximo	0,0335	0,00056	0,250	0,943	5,147	2,659	1,993	1,566

Tabela A3.3 - Erros normal contaminada I / Suposição de distribuição normal contaminada dos erros

	β_1	γ	α_1	ϕ_1	π	σ_1^2	σ_2^2	U	% 98	%95	%90
Real	0,07	0,0001	0,1	0,8	*	*	*	0,00	3,053	1,944	1,402
Mínimo	-0,1814	0,00002	0,022	0,380	0,7681	0,2894	2,8984	0,6843	2,367	1,360	1,028
1º Decil	-0,1100	0,00006	0,065	0,712	0,8383	0,4042	3,8468	1,1141	2,883	1,831	1,342
2º Decil	-0,0972	0,00007	0,076	0,746	0,8525	0,4251	4,1378	1,2405	2,940	1,871	1,370
3º Decil	-0,0852	0,00008	0,084	0,767	0,8612	0,4428	4,3354	1,3438	2,976	1,895	1,386
4º Decil	-0,0762	0,00009	0,091	0,783	0,8699	0,4573	4,5454	1,4428	3,002	1,922	1,398
5º Decil	-0,0687	0,00010	0,099	0,798	0,8778	0,4709	4,7429	1,5326	3,029	1,943	1,410
6º Decil	-0,0608	0,00011	0,108	0,811	0,8849	0,4836	4,9801	1,6447	3,056	1,969	1,421
7º Decil	-0,0534	0,00012	0,116	0,823	0,8923	0,5004	5,2360	1,7537	3,083	1,992	1,433
8º Decil	-0,0441	0,00014	0,128	0,840	0,9009	0,5167	5,5256	1,9147	3,116	2,020	1,446
9º Decil	-0,0302	0,00016	0,145	0,860	0,9119	0,5407	6,0906	2,1894	3,161	2,056	1,464
Máximo	0,0248	0,00042	0,246	0,944	0,9640	0,6663	14,6325	7,9958	3,385	2,185	1,532

A.4 Erros seguindo distribuição normal contaminada

II

Tabela A4.1 - Erros normal contaminada II/ Suposição de distribuição normal dos erros

	β_1	γ	α_1	ϕ_1
Real	-0,07	0,0001	0,1	0,8
Mínimo	-0,1791	0,00001	0,020	0,000
1º Decil	-0,1156	0,00006	0,059	0,671
2º Decil	-0,0999	0,00007	0,073	0,718
3º Decil	-0,0882	0,00008	0,083	0,748
4º Decil	-0,0786	0,00009	0,092	0,772
5º Decil	-0,0685	0,00011	0,100	0,790
6º Decil	-0,0601	0,00012	0,110	0,806
7º Decil	-0,0499	0,00014	0,120	0,825
8º Decil	-0,0385	0,00016	0,133	0,844
9º Decil	-0,0222	0,00020	0,150	0,873
Máximo	0,0451	0,00066	0,430	0,973

As estimativas dos percentis são sempre iguais aos percentis da distribuição normal (0;1) - (2.325, 1.960 e 1.645 para 98%, 95% e 90% respectivamente)

Tabela A4.2 -. Erros normal contaminada II/ Suposição de distribuição t-student dos erros

	β_1	γ	α_1	ϕ_1	GL	98%	95%	90%
Real	-0,07	0,0001	0,1	0,8	5	2,727	2,092	1,600
Mínimo	-0,1747	0,00001	0,021	0,000	3,101	1,600	1,600	1,384
1º Decil	-0,1115	0,00006	0,063	0,688	4,216	2,561	1,973	1,523
2º Decil	-0,0981	0,00007	0,076	0,732	4,488	2,577	1,982	1,539
3º Decil	-0,0866	0,00009	0,085	0,759	4,698	2,586	1,986	1,549
4º Decil	-0,0762	0,00010	0,094	0,776	4,873	2,596	1,989	1,556
5º Decil	-0,0673	0,00011	0,102	0,792	5,039	2,605	1,992	1,562
6º Decil	-0,0606	0,00012	0,112	0,809	5,227	2,612	1,994	1,568
7º Decil	-0,0523	0,00014	0,121	0,824	5,467	2,620	1,996	1,575
8º Decil	-0,0434	0,00015	0,132	0,844	5,703	2,629	1,997	1,581
9º Decil	-0,0258	0,00019	0,147	0,868	6,141	2,641	1,998	1,589
Máximo	0,0346	0,00076	0,232	0,969	9,981	2,658	1,999	1,621

Tabela A4.3 - Erros normal contaminada II/ Suposição de distribuição normal contaminada dos erros

	β_1	γ	α_1	ϕ_1	π	σ_1^2	σ_2^2	U	% 98	%95	%90
Real	0,07	0,0001	0,1	0,8	*	*	*	0,00	2,727	2,092	1,600
Mínimo	-0,1786	0,00001	0,022	0,000	0,3940	0,2441	1,4510	0,1439	2,214	1,650	1,339
1º Decil	-0,1112	0,00006	0,062	0,688	0,5646	0,3716	1,7859	0,2364	2,629	2,012	1,555
2º Decil	-0,0978	0,00007	0,074	0,733	0,6077	0,4069	1,8996	0,2638	2,659	2,047	1,573
3º Decil	-0,0857	0,00008	0,083	0,757	0,6354	0,4326	1,9765	0,2881	2,678	2,063	1,584
4º Decil	-0,0762	0,00009	0,092	0,777	0,6617	0,4531	2,0574	0,3140	2,691	2,075	1,594
5º Decil	-0,0677	0,00010	0,100	0,793	0,6871	0,4701	2,1536	0,3467	2,707	2,086	1,603
6º Decil	-0,0606	0,00012	0,108	0,810	0,7095	0,4902	2,2389	0,3778	2,723	2,097	1,612
7º Decil	-0,0531	0,00013	0,118	0,825	0,7309	0,5121	2,3461	0,4098	2,746	2,108	1,622
8º Decil	-0,0423	0,00015	0,129	0,845	0,7576	0,5355	2,4974	0,4581	2,767	2,120	1,631
9º Decil	-0,0273	0,00019	0,147	0,869	0,8055	0,5767	2,7740	0,5492	2,794	2,135	1,646
Máximo	0,0329	0,00065	0,245	0,965	0,9836	0,8021	14,9586	5,3970	3,007	2,229	1,701

Referências Bibliográficas

- [1] Bollerslev, T. (1986), *Generalized Autoregressive Conditional Heteroskedasticity*, Journal of Econometrics, 31, 307-327.
- [2] Bollerslev, T. (1987), *A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return*, Review of Economics and Statistics, 69, 542-547.
- [3] Cai, J. (1994), *A Markov Model os Switching-Regime ARCH*, Journal of Business and Economic Statistics, 12, 309-316.
- [4] Dueker, M.J. (1997), *Markov Regime in GARCH Processes and Mean-Reverting Stock Market Volatility*, Journal of Business and Economic Statistics, 15, 26-34.
- [5] Engle, R. F. (1982), *Autoregressive Conditional Heterokedasticity with Estimates of the Variance of U.K. Inflation*, Econometrica, 50, 987-1008.
- [6] Engle, R. F. e Bollerslev, T. (1986), *Modeling the Persistence of Conditional Variances*, Econometric Reviews, 5, 1-50.
- [7] Engle, R. F., Lilien, D. M. e Robins, R. P. (1987), *Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model*, Econometrica 55, 391-407.
- [8] Everit, B. S. e Hand, D. J. (1981), *Finite Mixture Distributions*, Cambridge : Cambridge University Press.
- [9] Hamilton, J. D. (1990), *Analisis of Time Series Subject to Changes in Regime*, Journal of Econometrics, 45, 39-70.

- [10] Hamilton, J. D. (1994), *Time Series Analysis*, Princeton : Princeton University Press.
- [11] Hamilton, J.D. e Susmel R. (1994), *Autoregressive Conditional Heteroskedasticity and Changes in Regime*, Journal of Econometrics, 64, 307-333.
- [12] Harvey A. C. (1993), *Times Series Models* , London : Harvester Wheatsheaf
- [13] Kim, C.J. (1994), *Dynamic Linear Models with Markov Switching*, Journal of Econometrics, 45, 39-70.
- [14] Nelson, N. B. (1991). *Conditional Heteroskedasticity in Asset Returns: A New Approach*, Econometrica 59, 347-370.
- [15] Palm, F. C. (1996). *GARCH Models of Volatility*, Maddala, G. S. e Rao, C. R., eds., Handbook of Statistics, 14, 209-240.
- [16] Rabi J.,L. A. (1998). *Modelos de Mudanças Markovianas de Regimes Aplicados a Séries Temporais Financeiras*, Dissertação de Mestrado, IME-USP, São Paulo.