
Modelos de Volatilidade Estatística

Danilo Kenji Ishizawa

Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Estatística
Departamento de Estatística

Modelos de Volatilidade Estatística

Danilo Kenji Ishizawa

Orientadora: Prof. Dra. Maria Sílvia de Assis Moura

Dissertação apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de São Carlos - PPGEs / UFSCar, como parte dos requisitos necessários para a obtenção do título de Mestre em Estatística.

São Carlos/SP

Outubro - 2008

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

1798mv

Ishizawa, Danilo Kenji.

Modelos de volatilidade estatística / Danilo Kenji Ishizawa.
-- São Carlos : UFSCar, 2008.
60 f.

Dissertação (Mestrado) -- Universidade Federal de São
Carlos, 2008.

1. Análise de séries temporais. 2. Análise de séries
temporais heterocedásticas. I. Título.

CDD: 519.55 (20ª)

*Aos meus pais,
Célia e
Walter.*

Agradecimentos

Agradeço primeiramente a meus pais Célia Fujico Yamaguti Ishizawa e Walter Tikahy Ishizawa por todo o apoio e incentivo que me deram em todas as etapas de minha vida.

Agradeço também o apoio, dedicação e orientação da Prof. Dra. Maria Sílvia de Assis Moura, além de todos os outros professores, funcionários, amigos e companheiros de mestrado que contribuíram nesta caminhada.

Deixo também um agradecimento especial a grande amiga, companheira e namorada Tatiane Silva de Souza que esteve sempre ao meu lado em mais esta jornada.

Resumo

No mercado financeiro costuma-se fazer observações sobre as carteiras sequencialmente ao longo do tempo, caracterizando uma série temporal. Contudo, o maior interesse está em prever o comportamento destas carteiras. Motivado por este fato, foram criados muitos modelos de previsão baseando-se em observações passadas considerando a média e variância constantes no tempo. Porém, nas séries financeiras uma característica muito presente é a chamada volatilidade, que pode ser observada pela variância não constante no tempo. A fim de captar esta característica, desenvolveram-se os modelos da família GARCH, que modelam a variância condicional através de informações passadas. Estes modelos foram muito utilizados e sofreram muitas modificações nas formulações para poderem captar diferentes efeitos, como o efeito de leverage (EGARCH). Assim, deseja-se estimar modelos de volatilidade obedecendo às especificações da família GARCH, verificando quais deles descrevem melhor os dados dentro e fora da amostra.

Palavras Chaves: função de auto-correlação (a.c.f.), função de auto-correlação parcial (p.a.c.f.), modelo ARCH, modelo GARCH, modelo EGARCH.

Abstract

In the financial market usually notices are taken of the shares sequentially over the time in order to characterize them a time series. However, the major interest is to forecast the behavior of these shares. Motivated by this fact, a lot of models were created based on the past information considering constant averages and variance over time. Although, in financial series a feature often presented is called volatility, which can be noticed by the variance to vary in time. In order to catch this characteristic were developed the models of the family GARCH, that model the conditional variance through known information. These models were well used and have passed by many formulation modifications to be able to catch different effects, such as the effect leverage EGARCH. Thus, the goal is to estimate volatility patterns obeying the specifications of the family GARCH verifying which ones of them describe better the data inside and outside the sample.

Keywords: ARCH model, auto correlation function (a.c.f.), EGARCH model, GARCH model, partial auto correlation function (p.a.c.f.).

Sumário

Introdução	x
1 Modelos de Volatilidade Estatística	1
1.1 Modelos ARCH	2
1.1.1 Identificação do Modelo ARCH	5
1.1.2 Estimação	6
1.1.3 Verificação	9
1.1.4 Previsão	9
1.2 Modelos GARCH	9
1.2.1 Identificação do Modelo GARCH	14
1.2.2 Estimação	15
1.2.3 Verificação	17
1.2.4 Previsão	17
1.3 Modelo EGARCH	18
1.3.1 Identificação do Modelo EGARCH	22
1.3.2 Estimação	22
1.3.3 Verificação	25
1.3.4 Previsão	25
2 Estudo das Propriedades dos Estimadores dos Parâmetros via Simulação	27
2.1 Modelo ARCH(1)	28
2.1.1 Convergência	29
2.1.2 Probabilidade de Cobertura	30
2.2 Modelo GARCH(1, 1)	32
2.2.1 Distribuições dos Lags dos Modelos	33

2.2.2	Convergência	34
2.2.3	Probabilidade de Cobertura	36
2.3	Modelo EGARCH(1, 1)	38
2.3.1	Distribuições dos Lags dos Modelos	39
2.3.2	Convergência	40
3	Escolha do Melhor Modelo	43
3.1	CrITÉrio de Akaike	43
3.2	Conteúdo de Informação	44
3.3	Probabilidade de Acerto	45
4	Aplicação	46
5	Considerações Finais, Conclusões e Propostas Futuras	51
A	Apêndice	53
A.1	Apêndice 1	53
A.1.1	Teste de Box-Pierce-Ljung	53
A.1.2	Teste dos Multiplicadores de Lagrange	53
A.2	Apêndice 2	54
A.2.1	Algoritmo Scoring	55
A.2.2	Algoritmo BHHH	56
	Referências Bibliográficas	59

Introdução

No mercado financeiro é muito comum observar o preço de uma determinada ação ou de uma determinada carteira de ações ao longo de um período de tempo. Estas observações podem ser por minuto, por hora, por mês ou por ano. O objetivo do acompanhamento dos preços das ações é coletar informações que permitam uma avaliação dos riscos de investimentos.

Neste trabalho estamos interessados na avaliação e modelagem das séries temporais financeiras ou mais especificamente, das séries temporais dos retornos das ações. O retorno pode ser chamado de *retorno composto continuamente* ou simplesmente *log – retorno*, definido como

$$r_t = \log \left[\frac{P_t}{P_{t-1}} \right] = p_t - p_{t-1},$$

onde P_t é o preço de um ativo (ações) num determinado tempo t e p_t é o logaritmo na base e de P_t . Esta definição é muito utilizada e nesta dissertação iremos nos referir a r_t simplesmente como *retorno*. A justificativa para se trabalhar com os retornos é que eles são livres de escala e apresentam propriedades estatísticas interessantes, como estacionariedade e ergodicidade.

As séries temporais dos retornos apresentam uma alta taxa de mudança na variância condicional para um determinado tempo t , a chamada volatilidade. Com isso, não se pode tratar estas séries pelas técnicas usuais de séries temporais, que supõem variância condicional constante ao longo do tempo.

A volatilidade não é diretamente observada, porém pode ser trabalhada segundo os modelos denominados heteroscedásticos. Seu bom entendimento é fundamental para o mercado financeiro, influenciando na avaliação dos riscos de investimento e no preço das ações. Além disso, a volatilidade indica se o preço de um ativo está variando muito ou pouco, sendo uma medida da incerteza quanto às variações de preço. Épocas nas quais a variabilidade dos retornos está muito alta são aquelas que possibilitam os maiores

lucros ou também as maiores perdas, isto é, são aquelas com maiores riscos. Quando a volatilidade é baixa, o risco é menor.

A volatilidade está diretamente ligada ao preço de um ativo, tendo como uma importante característica, o efeito de alavanca (leverage effect). No caso da presença constante de altas dos preços, este efeito é identificado pela baixa volatilidade nas séries dos retornos. Já no caso da presença constante de baixas dos preços, a série dos retornos apresenta uma alta volatilidade.

Logo, para o mercado financeiro a volatilidade é uma importante medida para a descrição da velocidade de mudança do mercado. Mercados que se movem lentamente são mercados com baixa volatilidade e os que se movem com mais rapidez são mercados com alta volatilidade.

Assim, os retornos financeiros apresentam as seguintes características

- em geral não são autocorrelacionados;
- apresentam a média próxima de zero;
- seus quadrados são autocorrelacionados;
- suas séries apresentam agrupamentos de volatilidade ao longo do tempo;
- apresentam efeito de assimetria;
- apresentam efeito de alavanca;
- suas distribuições apresentam caudas mais pesadas que as da normal.

Os modelos tradicionais de séries temporais assumem seus dados correlacionados com média e variância condicional constante no tempo. Já os modelos ditos heteroscedásticos assumem que as observações não são serialmente correlacionadas, porém são dependentes, com média constante, mas com a variância condicional podendo variar no tempo. Então, um dos primeiros objetivos será de ajustar um modelo auto-regressivo para remover a autocorrelação presente na série dos retornos (se for preciso), para no segundo momento ajustar um modelo heteroscedástico a série dos resíduos do modelo auto-regressivo, retirando a autocorrelação presente na variância.

Para modelar esta variabilidade condicional não constante, Engle (1982) propôs o modelo ARCH (“Autoregressive Conditional Heteroscedasticity”), que possui média zero, não é serialmente correlacionado e tem variância condicional não constante no tempo, dependendo de uma função das informações passadas.

O processo ARCH teve muito sucesso devido a sua grande utilidade para descrever fatos presentes nas variáveis econométricas, como explicar que grandes e pequenos

erros tendem a se agrupar e também que, nos previsores econométricos pode haver a existência de variação na habilidade de prever o futuro, entre um período e outro. No entanto, estes modelos na prática utilizavam muitos parâmetros devido à longa memória encontrada nas séries, além da existência do problema com a variância condicional negativa.

Para solucionar estas dificuldades, Bollerslev (1986) sugeriu uma extensão para os modelos ARCH, o chamado modelo GARCH (“generalized ARCH”). Este modelo é mais parcimonioso, isto é, descreve a volatilidade com menos parâmetros. A idéia foi introduzir na fórmula da variância condicional regressores da variância condicional passada, possibilitando ao modelo ser mais flexível, no sentido de descrever uma memória mais longa e com menos parâmetros (comparada aos modelos ARCH).

Os modelos GARCH foram os modelos econométricos mais amplamente usados para descrever as mudanças na variância condicional dos retornos. Contudo, sua equação da variância condicional leva em conta apenas a magnitude dos retornos, não importando assim sua positividade ou negatividade.

Esta característica de tratamento simétrico dos retornos positivos e negativos pode ser vista na formulação quadrática dos retornos encontrada na equação da variância. E ainda, diante do fato de muitos pesquisadores encontrarem evidências de que os choques dos retornos são negativamente correlacionados com a volatilidade, isto é, a volatilidade tende a aumentar em resposta aos retornos negativos. Assim, os modelos GARCH possuem a fraqueza de não conseguir captar a influência assimétrica dos retornos na modelagem da volatilidade das séries financeiras.

Uma outra limitação destes modelos é detectada no fato de que seus coeficientes devem ser não negativos. Esta imposição garante que a formulação da variância permaneça não negativa com probabilidade igual a um. Entretanto, muitas vezes esta imposição não é satisfeita ou dificulta muito a estimação destes parâmetros.

Para solucionar todas estas fraquezas dos modelos GARCH, Nelson (1991) propôs o modelo EGARCH (“exponencial GARCH”), que utiliza a forma exponencial para a variância e trata de maneira assimétrica, através de uma função g , os retornos positivos e negativos.

Esta dissertação segue a seguinte estrutura. No Capítulo 1, iremos apresentar os modelos de volatilidade estatística da família GARCH e uma de suas extensões (que admitem o efeito de leverage). Também será feito o desenvolvimento de algumas das suas propriedades juntamente com o desenvolvimento das previsões para os modelos propostos. Em cada caso, será mostrada uma maneira de se identificar, estimar, verificar a adequação do modelo e fazer previsões.

No Capítulo 2, iremos desenvolver um estudo de simulação com os modelos da família GARCH visando verificar as propriedades dos estimadores, como a convergência assintótica para o verdadeiro valor dos parâmetros e avaliar a probabilidade de cobertura dos intervalos de confiança dos parâmetros.

Como podem existir muitos modelos adequados a um conjunto de dados, necessitamos de alguma técnica para a sua discriminação. Assim, no Capítulo 3 apresentaremos alguns critérios para a escolha do “melhor” modelo segundo seu ajuste ou previsões. Já o Capítulo 4 apresenta uma aplicação para a série de fechamento de preços do Itaú.

Modelos de Volatilidade Estatística

Neste capítulo iremos estudar modelos de séries temporais denominados heteroscedásticos, isto é, modelos que descrevem a variância condicional dos retornos (volatilidade). Ao contrário dos modelos tradicionais de séries temporais, como o ARIMA que assumem a variância condicional constante no tempo, estes modelos permitem que a variância condicional varie no tempo, formulando-a como uma função que depende de valores dos retornos passados.

Estes modelos heteroscedásticos são denominados modelos de volatilidade. O primeiro a desenvolvê-los foi Engle (1982), que através de valores dos retornos passados, fez a modelagem da volatilidade para a inflação inglesa. O modelo desenvolvido por Engle recebeu o nome de ARCH que significa *Autoregressive Conditional Heteroskedasticity*.

Já Bollerslev (1986) apresentou uma generalização dos modelos ARCH, denominado modelo GARCH (*Generalized ARCH*). A principal diferença do modelo GARCH em relação ao modelo ARCH foi a inclusão de volatilidades anteriores na fórmula da variância condicional, permitindo ao modelo ser mais parcimonioso.

Os modelos GARCH foram muito utilizados para descrever séries financeiras, porém, na literatura econométrica é conhecido que retornos negativos aumentam mais a volatilidade que os retornos positivos. Como os modelos GARCH tratam os retornos positivos e negativos de forma similar, criou-se então a necessidade de suprir estas suas fraquezas.

Para captar a característica do efeito de assimetria dos retornos na volatilidade, Nelson (1991) apresentou o modelo *Exponential GARCH*, denominado EGARCH. Sua idéia foi utilizar a formulação exponencial para a volatilidade, além de uma função $g(\cdot)$ que trata de forma assimétrica os retornos positivos e negativos.

Assim, apresentaremos uma revisão de cada modelo, mostrando suas definições, processo de estimação, adequabilidade e previsões.

1.1 Modelos ARCH

Uma das características muito presente nas séries financeiras são que os retornos não têm variância constante no tempo, formando assim grupos com diferentes graus de volatilidade e média constante. Para modelar estes fatos, Engle (1982), propôs o modelo auto-regressivo com heterocedasticidade condicional (ARCH). Sua idéia básica é que o retorno a_t é não correlacionado serialmente, mas sua volatilidade (variância condicional ao tempo) depende de retornos passados por meio de uma função quadrática dos valores passados da série.

O modelo ARCH(r) é definido como

$$a_t = \sqrt{h_t}e_t, \quad (1.1)$$

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_r a_{t-r}^2, \quad (1.2)$$

onde e_t é uma seqüência de variáveis independentes, identicamente distribuídas com média zero e variância condicional igual a um (denotado por i.i.d. $(0,1)$), $\alpha_0 > 0$, $\alpha_i \geq 0$, $i > 0$.

Usualmente assumiremos que e_t possui uma distribuição normal com média zero e variância um ($e_t \sim N(0,1)$). Podemos assumir também que e_t possui uma distribuição t-Student com v graus de liberdade ($e_t \sim t_v$) ou uma outra distribuição que descreva melhor as caudas pesadas das séries financeiras. Os coeficientes α_i , com $i = 1, \dots, r$, devem satisfazer certas condições, dependendo do tipo de imposição que colocarmos sobre o processo a_t .

Pela própria definição, valores grandes de a_t são seguidos por outros valores grandes da série. Isto pode ser observado pelo fato que h_{t+1} depende exclusivamente do valor do retorno passado a_t ponderado por um coeficiente α_1 (no caso do modelo ARCH(1)).

Para investigar algumas propriedades dos modelos ARCH(r), consideremos o caso $r = 1$ com $e_t \sim N(0,1)$ e F_t sendo toda a informação observada até o tempo t , assim o modelo é definido como

$$a_t = \sqrt{h_t}e_t, \quad (1.3)$$

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2, \quad (1.4)$$

com $\alpha_0 > 0$, $\alpha_1 \geq 0$ e $t = 1, \dots, T$.

A média e variância de a_t condicionada a F_{t-1} dos retornos a_t são dados por

$$\begin{aligned} E(a_t|F_{t-1}) &= E(\sqrt{h_t}e_t|F_{t-1}) = \sqrt{h_t}E(e_t|F_{t-1}) = 0, \\ Var(a_t|F_{t-1}) &= E(a_t^2|F_{t-1}) = E(h_t e_t^2|F_{t-1}) = h_t E(e_t^2|F_{t-1}) = h_t. \end{aligned}$$

A média e variância não condicionais dos retornos são dadas pelas equações

$$\begin{aligned} E(a_t) &= E[E(a_t|F_{t-1})] = E(\sqrt{h_t}E(e_t|F_{t-1})) = 0, \\ Var(a_t) &= E(a_t^2) = E[E(a_t^2|F_{t-1})] \\ &= E(\alpha_0 + \alpha_1 a_{t-1}^2) = \alpha_0 + \alpha_1 E(a_{t-1}^2) \\ &= \alpha_0 + \alpha_1 E[E(a_{t-1}^2|F_{t-2})] = \alpha_0 + \alpha_1 [\alpha_0 + \alpha_1 E(a_{t-2}^2)] \\ &= \alpha_0 + \alpha_0 \alpha_1 + \alpha_1^2 E(a_{t-2}^2) = \dots = \frac{\alpha_0}{1 - \alpha_1}. \end{aligned} \quad (1.5)$$

Para satisfazer a definição de variância e as suposições que foram feitas ao modelo, isto é, $0 < Var(a_t) < \infty$, devemos então tomar $0 \leq \alpha_1 < 1$.

A covariância de a_t e a_{t+k} é dada por

$$\begin{aligned} Cov(a_t, a_{t+k}) &= E(a_t a_{t+k}) = E[E(a_t a_{t+k}|F_{t+k-1})] \\ &= E[a_t E(a_{t+k}|F_{t+k-1})] = E[a_t E(\sqrt{h_{t+k}} e_{t+k}|F_{t+k-1})] \\ &= E[a_t \sqrt{h_{t+k}} E(e_{t+k}|F_{t+k-1})] = 0, \end{aligned}$$

para $k > 0$, pois a_t e $\sqrt{h_{t+k}}$ estão em F_{t+k-1} e $E(e_{t+k}|F_{t+k-1}) = 0$.

Dessa forma a função de autocovariância de a_t é dada por

$$\gamma_{a_t}(k) = 0, \quad k \geq 1,$$

indicando que a_t é uma seqüência de variáveis não correlacionadas (ruído branco), com média zero e variância dada por (1.5).

Na literatura econométrica é conhecido que a função de densidade dos retornos apresentam geralmente caudas mais pesadas que a da normal, de modo que a curtose é maior do que 3. Para calcular a curtose do processo, supondo que a_t siga o modelo (1.3)-(1.4) é necessário calcular o momento de quarta ordem de a_t . Como os e_t são normais i.i.d. (0,1), então temos que

$$\begin{aligned} E(a_t^4|F_{t-1}) &= E(h_t^2 e_t^4|F_{t-1}) = h_t^2 E(e_t^4) \\ &= E(e_t^4)(\alpha_0 + \alpha_1 a_{t-1}^2)^2 = 3(\alpha_0 + \alpha_1 a_{t-1}^2)^2, \end{aligned} \quad (1.6)$$

pois $E(e_t^4) = 3$ na distribuição normal, do que decorre

$$\begin{aligned} E(a_t^4) &= E(3(\alpha_0 + \alpha_1 a_{t-1}^2)^2) = 3E(\alpha_0^2 + 2\alpha_0 \alpha_1 a_{t-1}^2 + \alpha_1^2 a_{t-1}^4) \\ &= 3[\alpha_0^2 + 2\alpha_0 \alpha_1 E(a_{t-1}^2) + \alpha_1^2 E(a_{t-1}^4)]. \end{aligned}$$

Admitindo que o processo seja estacionário de quarta ordem, o momento de quarta ordem pode ser escrito $\mu_4 = E(a_t^4)$ e teremos

$$\begin{aligned}\mu_4 &= 3(\alpha_0^2 + 2\alpha_0\alpha_1\text{Var}(a_{t-1}) + \alpha_1^2\mu_4) \\ &= 3(\alpha_0^2 + 2\alpha_0\alpha_1[\alpha_0/(1 - \alpha_1)] + \alpha_1^2\mu_4) \\ &= 3\alpha_0^2(1 + 2\alpha_1/(1 - \alpha_1)) + 3\alpha_1^2\mu_4.\end{aligned}$$

Ao isolar μ_4 na equação acima, obtemos finalmente a expressão do momento de quarta ordem de a_t dada por

$$\mu_4 = \frac{3\alpha_0^2(1 + 2\alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}. \quad (1.7)$$

Para que os momentos de quarta ordem sejam finitos e positivos precisamos ter $1 - 3\alpha_1^2 > 0$, ou seja, $0 \leq \alpha_1^2 < 1/3$. Portanto, quanto mais restrições impusermos ao processo, mais restrições teremos para os coeficientes do modelo. Isto é verdade para o modelo geral ARCH(r).

A curtose de a_t será dada então por

$$K = \frac{\mu_4}{[\text{Var}(a_t)]^2} = 3 \frac{\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)} \frac{(1 - \alpha_1)^2}{\alpha_0^2} = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2} > 3. \quad (1.8)$$

Como a curtose do modelo ARCH é maior que 3, suas caudas serão mais pesadas que a da normal e este fato é uma das características dos ativos financeiros. Logo este modelo é apropriado para descrever as séries financeiras.

Porém, ao se avaliar a função da variância condicional, percebemos que os retornos negativos e positivos são tratados de forma similar. Contudo, é conhecido empiricamente que a volatilidade reage de forma diferente aos retornos positivos e negativos, sendo mais influenciada pelo último.

Um outro fato que deve ser destacado, em relação as fraquezas do modelo, é que devido aos termos quadráticos presentes na função da variância condicional, alguns retornos grandes e isolados podem conduzir a super previsões.

Utilizando (1.3) e (1.4) e calculando $v_t = a_t^2 - h_t$, temos que

$$v_t = a_t^2 - (\alpha_0 + \alpha_1 a_{t-1}^2) = a_t^2 - h_t = h_t(e_t^2 - 1),$$

ou seja,

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + v_t, \quad (1.9)$$

no qual

$$v_t = h_t(e_t^2 - 1) = h_t(A - 1), \quad (1.10)$$

onde A é uma v.a. com distribuição $\chi^2(1)$, o que mostra que temos um modelo AR(1) para a_t^2 , mas com erros não gaussianos.

Calculando a esperança e variância, condicionada as informações passadas, para a variável v_t temos que

$$\begin{aligned} E(v_t|F_{t-1}) &= E(h_t(e_t^2 - 1)|F_{t-1}) \\ &= h_t[E(e_t^2|F_{t-1}) - 1] = 0, \\ Var(v_t|F_{t-1}) &= E(v_t^2|F_{t-1}) = E[h_t^2(e_t^2 - 1)^2|F_{t-1}] \\ &= h_t^2[E(e_t^4|F_{t-1}) - 2E(e_t^2|F_{t-1}) + 1] = 2h_t^2. \end{aligned}$$

Portanto $\{v_t\}$ é uma seqüência de v.a. com média zero, não correlacionadas, mas com variância condicional não constante.

Assim, pelo fato de que o processo ARCH(1) pode ser escrito como um processo AR(1), dado pela equação (1.9), temos que a f.a.c. (função de autocorrelação) de a_t^2 é caracterizada por um decaimento exponencial, e sua expressão é dada por

$$\rho_{a_t^2}(k) = \alpha_1^k, \quad k > 0.$$

Fazendo os mesmos passos para o modelo ARCH(r) temos

$$a_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i a_{t-i}^2 + v_t, \quad (1.11)$$

onde os v_t não seguem mais a distribuição normal, como no caso $r = 1$, tendo assim uma distribuição com caudas mais pesadas. Ou seja, temos um modelo AR(r) para a_t^2 , com inovações não gaussianas.

1.1.1 Identificação do Modelo ARCH

Como os modelos ARCH consideram que a série é não correlacionada serialmente, temos que o primeiro passo no ajuste da série é tentar retirar toda a autocorrelação da série, se esta existir, ajustando um modelo ARMA(p, q). No caso da série ser autocorrelacionada, temos a seguinte expressão

$$\phi(B)X_t = \theta_0 + \theta(B)a_t,$$

onde B é o operador retroativo (isto é, leva a variável no tempo t para o tempo imediatamente anterior $t-1$), $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$, $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$,

X_t é a série original dos retornos e $a_t \sim \text{ARCH}(r)$ é o resíduo do modelo. No que segue, quando nos referirmos a a_t , estaremos supondo ou que a série é não correlacionada, ou então ela é o resíduo da aplicação de um modelo ARMA à série original (X_t). No caso da série correlacionada X_t não ser estacionária, podemos aplicar a ela algumas diferenças $(1 - B)^d$, para assim atingir a estacionaridade. Com isso, estaremos ajustando um modelo ARIMA(p, d, q), onde d é o número de diferenças. Logo a expressão para este modelo é dada por

$$\phi(B)\nabla^d X_t = \theta_0 + \theta(B)a_t,$$

onde $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$, $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$, $\nabla = (1 - B)$, X_t é a série original dos retornos e $a_t \sim \text{ARCH}(r)$ é o resíduo do modelo.

Para verificar se a série a_t^2 apresenta heteroscedasticidade condicional, podem ser utilizado dois testes:

- (i) Teste de Box-Pierce-Ljung para a_t^2 (A.1.1);
- (ii) Teste de multiplicadores de Lagrange - ML (A.1.2).

Como a_t^2 é um estimador não viesado de h_t (condicionado aos retornos passados) e dada a forma (1.2) de modelarmos a volatilidade, notamos que o valor atual do quadrado do retorno depende de quadrados de retornos passados como em um modelo auto-regressivo. Com isso podemos utilizar f.a.c.p. (função de autocorrelação parcial) de a_t^2 para encontrar a ordem do modelo ARCH(r), onde o número (finito) de “lags” seqüenciais antes do corte (lags significativos) será a ordem “ r ” do modelo. Porém este método não é muito confiável. Assim recomenda-se ajustar alguns modelos, tendo uma idéia de sua ordem dada pela f.a.c.p., escolhendo o melhor por algum critério de informação, como por exemplo o AIC (critério de Akaike).

Vale observar que os “lags” de ordem r , seriam a influência das informações passadas (F_{t-r}) no modelo, dado um determinado tempo t .

1.1.2 Estimação

Como em qualquer técnica estatística é necessário a estimação de seus parâmetros para que o modelo fique completamente identificado, dado a escolha da ordem r a ser utilizado para o modelo ARCH. Logo as estimativas podem ser obtidas pelo método de máxima verossimilhança condicional, onde a função de verossimilhança é dada por

$$\begin{aligned} L(a_1, \dots, a_T | \alpha) &= f(a_T | F_{T-1}) f(a_{T-1} | F_{T-2}) \dots f(a_{r+1} | F_r) f(a_1, \dots, a_r | \alpha). \\ &= \prod_{t=r+1}^T f(a_t | F_{t-1}) f(a_1, \dots, a_r | \alpha). \end{aligned} \quad (1.12)$$

Os erros e_t do modelo ARCH podem seguir qualquer distribuição com média zero e variância unitária, segundo a definição do modelo. Vamos apresentar agora o desenvolvimento da verossimilhança para o caso das distribuições Normal e t-Student padrões.

Distribuição Normal

Supondo que $e_t \sim N(0, 1)$ temos que a função de densidade condicional de a_t dado as informações passadas segue uma distribuição normal com média zero e variância h_t dada por

$$f(a_t|F_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} \exp\left\{\frac{-a_t^2}{2h_t}\right\}.$$

Utilizando a verossimilhança dada pela equação (1.12) e assumindo que para T grande, $f(a_1, \dots, a_r|\alpha)$ pode ser desprezado. Então a verossimilhança condicionada aos “ r ” primeiros retornos é dada por

$$L(a_{r+1}, \dots, a_T|\alpha, a_1, \dots, a_r) = \prod_{r+1}^T (\sqrt{2\pi h_t})^{-1} \exp\left\{\frac{-a_t^2}{2h_t}\right\}, \quad (1.13)$$

onde a volatilidade h_t é obtida recursivamente.

Logo temos que maximizar a função de verossimilhança condicional para encontrar os estimadores de máxima verossimilhança para os parâmetros.

No artigo original Engle (1982) é utilizada a log-verossimilhança média com erros seguindo uma distribuição normal padrão para estimar os parâmetros, que tem a finalidade de facilitar os resultados assintóticos, dada por

$$l(\alpha) = \frac{1}{T} \sum_{t=1}^T l_t(\alpha),$$

onde $l_t(\alpha)$ é a log-verossimilhança para o tempo t , que tem a seguinte forma

$$l_t(\alpha) \propto -\frac{1}{2} \ln(h_t) - \frac{a_t^2}{2h_t}.$$

No caso particular de um modelo ARCH(1), temos

$$L(a_2, \dots, a_T|\alpha_0, \alpha_1, a_1) = f(a_T|a_{T-1})f(a_{T-1}|a_{T-2})\dots f(a_2|a_1),$$

em que

$$(a_t|a_{t-1}) \sim N(0, h_t),$$

e $h_t = \alpha_0 + \alpha_1 a_{t-1}^2$. Segue que

$$L(a_2, \dots, a_T | \alpha_0, \alpha_1, a_1) = (2\pi)^{-T/2} \prod_{t=2}^T (\alpha_0 + \alpha_1 a_{t-1}^2)^{-1/2} \exp \left\{ \frac{-a_t^2}{2(\alpha_0 + \alpha_1 a_{t-1}^2)} \right\}.$$

A log-verossimilhança fica

$$l(a_2, \dots, a_T | \alpha_0, \alpha_1, a_1) \propto -\frac{1}{2} \sum_{t=2}^T \ln(\alpha_0 + \alpha_1 a_{t-1}^2) - \frac{1}{2} \sum_{t=2}^T \left(\frac{a_t^2}{\alpha_0 + \alpha_1 a_{t-1}^2} \right). \quad (1.14)$$

Distribuição t-Student

Supondo que $e_t \sim t_v$ (t de Student com v graus de liberdade) temos que a função de densidade condicional de a_t dado as informações passadas segue também uma distribuição t de Student dada por

$$f(a_t | F_{t-1}) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2) \sqrt{(v-2)\pi h_t}} \left(1 + \frac{a_t^2}{(v-2)h_t} \right)^{-(v+1)/2}.$$

Utilizando a verossimilhança dada pela equação (1.12) e assumindo que para T grande, $f(a_1, \dots, a_r | \alpha)$ pode ser desprezado. Então a verossimilhança condicionada aos “ r ” primeiros retornos é dada por

$$L(a_{r+1}, \dots, a_T | \alpha, a_1, \dots, a_r) = \prod_{t=r+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2) \sqrt{(v-2)\pi h_t}} \left(1 + \frac{a_t^2}{(v-2)h_t} \right)^{-(v+1)/2}. \quad (1.15)$$

onde a volatilidade h_t é obtida recursivamente.

Logo temos que maximizar a função de verossimilhança condicional para encontrar os estimadores de máxima verossimilhança para os parâmetros. Neste caso, podemos estimar os parâmetros fixando um determinado grau de liberdade ou podemos também estimar os graus de liberdades juntamente com os parâmetros.

Em ambas as situações de erros normais ou t de Student, algum procedimento não linear de otimização numérica terá que ser utilizado. No artigo de Engle (1982) é utilizado o algoritmo scoring para estimar os parâmetros, este método iterativo é descrito em (A.2.1).

1.1.3 Verificação

Para um modelo ARCH(r), com e_t normal ou t -Student, os resíduos padronizados

$$\tilde{a}_t = \frac{a_t}{\sqrt{h_t}}$$

são v.a. i.i.d. com distribuição normal padrão ou t -Student. Logo, uma maneira de verificar se o modelo é adequado, é calcular a estatística Q de Ljung-Box (descrita em A.1.1) para a seqüência \tilde{a}_t . Além disso, podemos calcular os coeficientes de assimetria e curtose estimados e fazer um gráfico *quantil* \times *quantil* para avaliar a suposição de normalidade.

Para verificar se ainda existe heteroscedasticidade condicional nos resíduos, pode-se aplicar o teste ML para a seqüência \tilde{a}_t^2 .

1.1.4 Previsão

As previsões para a volatilidade, utilizando o modelo ARCH(r) dado pelas equações (1.1)-(1.2) são obtidas recursivamente. Assim,

$$\hat{h}_t(1) = \alpha_0 + \alpha_1 a_t^2 + \dots + \alpha_r a_{t-r+1}^2, \quad (1.16)$$

é a previsão de h_{t+1} (um passo à frente), com origem fixada no instante t . As previsões l passos à frente, com origem em t , são dadas por

$$\hat{h}_t(l) = \alpha_0 + \sum_{i=1}^r \alpha_i \hat{h}_t(l-i), \quad (1.17)$$

em que $\hat{h}(l-i) = a_{t+l-i}^2$, se $l-i \leq 0$.

1.2 Modelos GARCH

O processo ARCH(r) introduzido por Engle (1982) descreveu o comportamento de um grande número de fenômenos econômicos. Porém, em muitos trabalhos, a ordem destes modelos apresentou-se muito alta devido a longa memória encontrada, além da existência do problema com a variância condicional negativa. Conseqüentemente, na prática, a estimação dos modelos ARCH não é simples.

Bollerslev (1986) sugeriu uma extensão para os modelos ARCH, o chamado modelo GARCH (“generalized ARCH”). Este modelo é mais parcimonioso, isto é, descreve a volatilidade com menos parâmetros. A idéia foi introduzir na fórmula da variância condicional regressores da variância condicional passada, possibilitando ao modelo ser

mais flexível, no sentido de descrever uma memória mais longa e com menos parâmetros comparado aos modelos ARCH.

O modelo GARCH(r, s) é definido por

$$a_t = \sqrt{h_t} e_t \quad (1.18)$$

$$h_t = \alpha_0 + \sum_{i=1}^r \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}, \quad (1.19)$$

com e_t i.i.d. $(0,1)$, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $\sum_{i=1}^q (\alpha_i + \beta_i) < 1$, $q = \max(r, s)$.

Os coeficientes positivos dão uma condição suficiente para que $h_t > 0$, já a desigualdade $\sum_{i=1}^q (\alpha_i + \beta_i) < 1$, com $q = \max(r, s)$, satisfaz a condição necessária. Para condições gerais, veja Nelson e Cao (1992).

Como no caso dos modelos ARCH, usualmente supomos que os e_t são normais ou seguem uma distribuição t de Student, ou ainda uma outra distribuição que descreva melhor as caudas pesadas das séries financeiras.

Para a investigação de algumas propriedades dos modelos GARCH, iremos utilizar o modelo de ordem $r = 1$ e $s = 1$ com $e_t \sim N(0, 1)$, muito utilizado na prática, para o qual é expresso como

$$\begin{aligned} a_t &= \sqrt{h_t} e_t \\ h_t &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1}, \end{aligned} \quad (1.20)$$

com $0 \leq \alpha_1, \beta_1 < 1$, $\alpha_1 + \beta_1 < 1$.

Reescrevendo a equação (1.20) temos

$$\begin{aligned} h_t &= \frac{\alpha_0}{1 - \beta_1} + \frac{\alpha_1 a_{t-1}^2}{1 - \beta_1 L} \\ &= \frac{\alpha_0}{1 - \beta_1} + \sum_{i=1}^{\infty} \alpha_1 a_{t-1}^2 (\beta_1 L)^{i-1} \\ &= \frac{\alpha_0}{1 - \beta_1} + \sum_{i=1}^{\infty} \alpha_1 \beta_1^{i-1} a_{t-i}^2 \\ &= \delta_0 + \sum_{i=1}^{\infty} \delta_i a_{t-i}^2, \end{aligned} \quad (1.21)$$

onde $\delta_0 = \alpha_0 / (1 - \beta_1)$ e $\delta_i = \alpha_1 \beta_1^{i-1}$. Assim a equação (1.21) juntamente com (1.18), descrevem um modelo ARCH(∞). Portanto o modelo GARCH(1,1) pode ser visto como um modelo ARCH de ordem infinita, sendo assim mais parcimonioso.

Calculando a média e variância condicionais dos retornos chegamos que

$$\begin{aligned} E(a_t|F_{t-1}) &= E(\sqrt{h_t}e_t|F_{t-1}) = \sqrt{h_t}E(e_t|F_{t-1}) = 0, \\ Var(a_t|F_{t-1}) &= E(a_t^2|F_{t-1}) = E(h_t e_t^2|F_{t-1}) = h_t E(e_t^2|F_{t-1}) = h_t E(e_t^2) = h_t. \end{aligned}$$

A média e variância dos retornos são dadas pelas equações

$$\begin{aligned} E(a_t) &= E[E(a_t|F_{t-1})] = 0, \\ Var(a_t) &= E(a_t^2) = E[E(a_t^2|F_{t-1})] \\ &= E(h_t) = E(\alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1}) \\ &= \alpha_0 + \alpha_1 E(a_{t-1}^2) + \beta_1 E(h_{t-1}) \\ &= \alpha_0 + \alpha_1 E[E(h_{t-1} e_{t-1}^2|F_{t-2})] + \beta_1 E(\alpha_0 + \alpha_1 a_{t-2}^2 + \beta_1 h_{t-2}) \\ &= \alpha_0 + \alpha_1 [\alpha_0 + \alpha_1 E(a_{t-2}^2) + \beta_1 E(h_{t-2})] + \beta_1 [\alpha_0 + \alpha_1 E(a_{t-2}^2) + \beta_1 E(h_{t-2})] \\ &= \alpha_0 + \alpha_0(\alpha_1 + \beta_1) + \alpha_1^2 E(a_{t-2}^2) + \alpha_1 \beta_1 E(h_{t-2}) + \alpha_1 \beta_1 E(a_{t-2}^2) + \beta_1^2 E(h_{t-2}) \\ &= \alpha_0 + \alpha_0(\alpha_1 + \beta_1) + \alpha_0(\alpha_1 + \beta_1)^2 + \dots = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}. \end{aligned} \quad (1.22)$$

Para poder atender a definição de variância e as suposições feitas ao modelo, isto é $0 < Var(a_t) < \infty$, devemos então tomar $0 \leq \alpha_1 + \beta_1 < 1$.

No caso geral do modelo GARCH(r, s) a variância será dada por

$$Var(a_t) = \frac{\alpha_0}{1 - (\sum_{i=1}^q \alpha_i + \beta_i)}$$

onde $q = \max(r, s)$. Para atender a definição de variância e as suposições feitas ao modelo, temos que $0 \leq \sum_{i=1}^q \alpha_i + \beta_i < 1$.

Calculando a covariância de a_t e a_{t+k} temos que

$$\begin{aligned} Cov(a_t, a_{t+k}) &= E(a_t a_{t+k}) = E[E(a_t a_{t+k}|F_{t+k-1})] \\ &= E[a_t E(a_{t+k}|F_{t+k-1})] = E[a_t E(\sqrt{h_{t+k}} e_{t+k}|F_{t+k-1})] \\ &= E[a_t \sqrt{h_{t+k}} E(e_{t+k}|F_{t+k-1})] = 0, \end{aligned}$$

para $k > 0$, pois a_t e $\sqrt{h_{t+k}}$ estão em F_{t+k-1} e $E(e_{t+k}|F_{t+k-1}) = 0$.

Dessa forma a função de autocovariância de a_t será

$$\gamma_{a_t}(k) = 0, \quad k \geq 1,$$

indicando que a_t é uma seqüência de variáveis não correlacionadas (ruído branco), com média zero e variância dada por (1.22).

Para calcular a curtose do processo a_t , supondo que siga o modelo GARCH(1,1), é necessário calcular o momento de quarta ordem. Como os e_t são normais i.i.d., então temos que

$$\begin{aligned} E(a_t^4 | F_{t-1}) &= E(h_t^2 e_t^4 | F_{t-1}) = h_t^2 E(e_t^4 | F_{t-1}) \\ &= E(e_t^4) (\alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1})^2 = 3(\alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1})^2, \end{aligned}$$

pois $E(e_t^4) = 3$ na distribuição normal, do que decorre

$$\begin{aligned} E(a_t^4) &= 3E[(\alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1})^2] \\ &= 3E(\alpha_0^2 + \alpha_1^2 a_{t-1}^4 + \beta_1^2 h_{t-1}^2 + 2\alpha_0 \alpha_1 a_{t-1}^2 + 2\alpha_0 \beta_1 h_{t-1} + 2\alpha_1 \beta_1 a_{t-1}^2 h_{t-1}) \\ &= 3[\alpha_0^2 + \alpha_1^2 E(a_{t-1}^4) + \beta_1^2 E(h_{t-1}^2) + 2\alpha_0 \alpha_1 E(a_{t-1}^2) + 2\alpha_0 \beta_1 E(h_{t-1}) + 2\alpha_1 \beta_1 E(a_{t-1}^2 h_{t-1})]. \end{aligned}$$

Admitindo que o processo seja estacionário de quarta ordem, o momento de quarta ordem pode ser escrito $\mu_4 = E(a_t^4)$ e teremos

$$\begin{aligned} \mu_4 &= 3[\alpha_0^2 + \alpha_1^2 \mu_4 + \beta_1^2 E(h_{t-1}^2) + 2\alpha_0 \alpha_1 E(a_{t-1}^2) + 2\alpha_0 \beta_1 E(a_{t-1}^2) + 2\alpha_1 \beta_1 E(h_{t-1}^2)] \\ &= 3 \left[\alpha_0^2 + \alpha_1^2 \mu_4 + \frac{2\alpha_0^2(\alpha_1 + \beta_1)}{1 - (\alpha_1 + \beta_1)} + (2\alpha_1 \beta_1 + \beta_1^2) E(h_{t-1}^2) \right] \\ &= 3 \left[\alpha_0^2 + \alpha_1^2 \mu_4 + \frac{2\alpha_0^2(\alpha_1 + \beta_1)}{1 - (\alpha_1 + \beta_1)} \right] + 3(2\alpha_1 \beta_1 + \beta_1^2) E(h_{t-1}^2) \end{aligned}$$

Considere $A = \left[\alpha_0^2 + \alpha_1^2 \mu_4 + \frac{2\alpha_0^2(\alpha_1 + \beta_1)}{1 - (\alpha_1 + \beta_1)} \right]$,
então

$$\begin{aligned} \mu_4 &= 3A + 3(2\alpha_1 \beta_1 + \beta_1^2) E(h_{t-1}^2) \\ &= 3A + 3(2\alpha_1 \beta_1 + \beta_1^2) [A + (2\alpha_1 \beta_1 + \beta_1^2) E(h_{t-2}^2)] \\ &= 3A + 3A(2\alpha_1 \beta_1 + \beta_1^2) + 3A(2\alpha_1 \beta_1 + \beta_1^2)^2 [A + (2\alpha_1 \beta_1 + \beta_1^2) E(h_{t-3}^2)] \\ &= 3A + 3A(2\alpha_1 \beta_1 + \beta_1^2) + 3A(2\alpha_1 \beta_1 + \beta_1^2)^2 + 3A(2\alpha_1 \beta_1 + \beta_1^2)^3 + \dots \\ &= \frac{3A}{1 - (2\alpha_1 \beta_1 + \beta_1^2)} \end{aligned}$$

Substituindo A por sua equação e isolando μ_4 na equação acima, obtemos finalmente a expressão do momento de quarta ordem de a_t dada por

$$\mu_4 = \frac{3\alpha_0^2(1 + \alpha_1 + \beta_1)}{[1 - (\alpha_1 + \beta_1)](1 - \beta_1^2 - 2\alpha_1 \beta_1 - 3\alpha_1^2)}. \quad (1.23)$$

Para que os momentos de quarta ordem sejam finitos e positivos, devemos ter $(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2) > 0$, ou seja, $[(\alpha_1 + \beta_1)^2 + 2\alpha_1^2] < 1$.

Assim, a curtose do modelo (1.20) é obtida por

$$\begin{aligned} K &= \frac{\mu_4}{[Var(a_t)]^2} \\ &= \frac{3\alpha_0^2(1 + \alpha_1 + \beta_1)}{[1 - (\alpha_1 + \beta_1)](1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)} \frac{[1 - (\alpha_1 + \beta_1)]^2}{\alpha_0^2} \\ &= \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3, \end{aligned}$$

dado que o denominador seja positivo, mostrando novamente que as caudas de a_t serão mais pesadas do que as da normal.

Para os modelos GARCH temos então as mesmas vantagens e desvantagens dos modelos ARCH. Volatilidades altas são precedidas de retornos ou volatilidades grandes, as caudas da distribuição de a_t são mais pesadas que a da distribuição normal e retornos positivos e negativos são tratados de mesma forma, tendo em vista que os retornos ao quadrado entram na fórmula da volatilidade.

Utilizando (1.18) e calculando $a_t^2 - h_t$, temos que

$$v_t = a_t^2 - h_t = h_t(e_t^2 - 1) = h_t(A - 1), \quad (1.24)$$

de modo que, substituindo (1.19) obtemos

$$\begin{aligned} a_t^2 &= \alpha_0 + \sum_{i=1}^r \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j} + \sum_{j=1}^s \beta_j a_{t-j}^2 - \sum_{j=1}^s \beta_j a_{t-j}^2 \\ &= \alpha_0 + \sum_{i=1}^q (\alpha_i + \beta_i) a_{t-i}^2 - \sum_{j=1}^s \beta_j v_{t-j} + v_t. \end{aligned} \quad (1.25)$$

onde $q = \max(r, s)$ e a_t^2 segue um processo ARMA com a distribuição dos erros v_t não mais normal, dado que A é uma v.a. com distribuição $\chi^2(1)$.

Calculando a esperança e variância, condicionada as informações passadas, para a variável v_t temos que

$$\begin{aligned} E(v_t | F_{t-1}) &= E(h_t(e_t^2 - 1) | F_{t-1}) \\ &= h_t[E(e_t^2 | F_{t-1}) - 1] = 0, \\ Var(v_t | F_{t-1}) &= E(v_t^2 | F_{t-1}) = E[h_t^2(e_t^2 - 1)^2 | F_{t-1}] \\ &= h_t^2[E(e_t^4 | F_{t-1}) - 2E(e_t^2 | F_{t-1}) + 1] = 2h_t^2. \end{aligned}$$

Assim, temos um modelo ARMA(r, s) para a_t^2 , porém a seqüência de v.a. $\{v_t\}$ não é mais um processo i.i.d., isto é, sua variância condicional não é mais constante

para todo t , sendo influenciada pelos retornos e volatilidades passadas provenientes da fórmula (1.19).

1.2.1 Identificação do Modelo GARCH

Assim como os modelos ARCH, os modelos GARCH, consideram que a série é não correlacionada serialmente, então temos que o primeiro passo é tentar retirar toda a autocorrelação da série, se esta existir, ajustando um modelo ARMA. No caso da série ser autocorrelacionada, temos a seguinte expressão

$$\phi(B)X_t = \theta_0 + \theta(B)a_t,$$

onde $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$, $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ e $a_t \sim \text{GARCH}(r, s)$. Logo, quando nos referirmos a a_t , estaremos supondo ou que a série é não correlacionada, ou então ela é o resíduo da aplicação de um modelo ARMA à série original (X_t). No caso da série correlacionada X_t não ser estacionária, podemos aplicar a ela algumas diferenças $(1 - B)^d$, para assim atingir a estacionaridade. Com isso, estaremos ajustando um modelo ARIMA(p, d, q), onde d é o número de diferenças. Logo a expressão para este modelo é dada por

$$\phi(B)\nabla^d X_t = \theta_0 + \theta(B)a_t,$$

onde $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$, $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$, $\nabla = (1 - B)$, X_t é a série original dos retornos e $a_t \sim \text{GARCH}(r, s)$ é o resíduo do modelo.

Para verificarmos se a série apresenta heteroscedasticidade condicional, podemos utilizar dois testes para série a_t^2 :

- (i) Teste de Box-Pierce-Ljung (A.1.1);
- (ii) Teste de multiplicadores de Lagrange - ML (A.1.2).

Novamente a f.a.c. pode ser utilizado para verificar se a série a_t^2 segue um processo GARCH, seu comportamento é marcado por um decaimento exponencial.

Para identificarmos a ordem do modelo GARCH(r, s), utilizaremos o número de lags significantes da f.a.c. para estimarmos o valor de “ s ” e o número de lags significantes da f.a.c.p. para estimarmos o valor de “ r ”, assim como no caso dos modelos ARMA. No entanto este método de identificação não é muito confiável, servindo apenas para dar uma idéia da ordem do modelo a ser ajustado. Recomenda-se, assim como no caso dos modelos ARCH, que se ajuste alguns modelos, com base na ordem dada pelas f.a.c. e f.a.c.p., usando algum critério de informação, como no caso do AIC para discriminar o melhor.

Normalmente os modelos usados para descrever as séries reais são de ordem baixa, como (1,1), (1,2), (2,1) ou (2,2).

1.2.2 Estimação

Para a estimação dos parâmetros do modelo GARCH(r, s), dados pelas equações (1.18)-(1.19), utilizaremos o método de máxima verossimilhança condicional, onde a função de verossimilhança é dada por

$$\begin{aligned} L(a_1, \dots, a_T | \alpha, \beta) &= f(a_T | F_{T-1}) f(a_{T-1} | F_{T-2}) \dots f(a_{r+1} | F_r) f(a_1, \dots, a_r | \alpha, \beta) \\ &= \prod_{t=r+1}^T f(a_t | F_{t-1}) f(a_{t-1} | F_{t-2}) f(a_1, \dots, a_r | \alpha, \beta), \end{aligned} \quad (1.26)$$

Os erros e_t do modelo GARCH podem seguir qualquer distribuição com média zero e variância um, segundo a definição do modelo. Apresentaremos agora o desenvolvimento da verossimilhança para o caso das distribuições Normal e t-Student padrões.

Distribuição Normal

Supondo que $e_t \sim N(0, 1)$ temos que a função de densidade condicional de a_t dado as informações passadas segue uma distribuição normal com média zero e variância h_t dada por

$$f(a_t | F_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} \exp \left\{ \frac{-a_t^2}{2h_t} \right\}.$$

Utilizando a verossimilhança dada pela equação (1.26) e assumindo que para T grande, $f(a_1, \dots, a_r | \alpha, \beta)$ pode ser desprezado. Então a verossimilhança condicionada aos "r" primeiros retornos é dada por

$$L(a_{r+1}, \dots, a_T | \alpha, \beta, a_1, \dots, a_r) = \prod_{t=r+1}^T (\sqrt{2\pi h_t})^{-1} \exp \left\{ \frac{-a_t^2}{2h_t} \right\} \quad (1.27)$$

onde a volatilidade h_t é obtida recursivamente.

Assim, o logaritmo da verossimilhança condicionado às primeiras "r" observações é dada por

$$l(a_{r+1}, \dots, a_T | \alpha, \beta, a_1, \dots, a_r) \propto -\frac{1}{2} \sum_{t=r+1}^T \ln(h_t) - \frac{1}{2} \sum_{t=r+1}^T \frac{a_t^2}{h_t}. \quad (1.28)$$

Para encontrar os estimadores de máxima verossimilhança para os parâmetros, temos então que maximizar a função de verossimilhança condicional.

Assim como foi feito no artigo de Engle (1982), Bollerslev (1986), para facilitar os resultados assintóticos, utiliza a log-verossimilhança média com erros seguindo uma distribuição normal padrão para estimar os parâmetros do modelo dada por

$$l(\alpha, \beta) = \frac{1}{T} \sum_{t=1}^T l_t(\alpha, \beta),$$

onde $l_t(\alpha, \beta)$ é a log-verossimilhança para o tempo t , que tem a seguinte forma

$$l_t(\alpha, \beta) \propto -\frac{1}{2} \ln(h_t) - \frac{a_t^2}{2h_t}.$$

No caso particular de um modelo GARCH(1,1), temos

$$L(a_2, \dots, a_T | \alpha_0, \alpha_1, \beta_1, a_1) = f(a_T | a_{T-1}) f(a_{T-1} | a_{T-2}) \dots f(a_2 | a_1),$$

em que

$$(a_t | a_{t-1}) \sim N(0, h_t),$$

e $h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1}$. Segue que

$$L(a_2, \dots, a_T | \alpha_0, \alpha_1, \beta_1, a_1) = (2\pi)^{-T/2} \prod_{t=2}^T h_t^{-1/2} \exp \left\{ \frac{-a_t^2}{2h_t} \right\}.$$

Então a log-verossimilhança será

$$l(a_2, \dots, a_T | \alpha_0, \alpha_1, \beta_1, a_1) \propto -\frac{1}{2} \sum_{t=2}^T \ln(h_t) - \frac{1}{2} \sum_{t=2}^T \left(\frac{a_t^2}{h_t} \right). \quad (1.29)$$

Distribuição t-Student

Supondo que $e_t \sim t_v$ (t de Student com v graus de liberdade) temos que a função de densidade condicional de a_t dado as informações passadas segue também uma distribuição t de Student dada por

$$f(a_t | F_{t-1}) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2) \sqrt{(v-2)\pi h_t}} \left(1 + \frac{a_t^2}{(v-2)h_t} \right)^{-(v+1)/2}.$$

Utilizando a verossimilhança dada pela equação (1.26) e assumindo que para T grande, $f(a_1, \dots, a_r | \alpha, \beta)$ pode ser desprezado. Então a verossimilhança condicionada

aos “ r ” primeiros retornos é dada por

$$L(a_{r+1}, \dots, a_T | \alpha, \beta, a_1, \dots, a_r) = \prod_{r+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi h_t}} \left(1 + \frac{a_t^2}{(v-2)h_t}\right)^{-(v+1)/2}. \quad (1.30)$$

onde a volatilidade h_t é obtida recursivamente.

Logo temos que maximizar a função de verossimilhança condicional para encontrar os estimadores de máxima verossimilhança para os parâmetros. Neste caso, podemos estimar os parâmetros fixando um determinado grau de liberdade ou podemos também estimar os graus de liberdades juntamente com os parâmetros.

Para encontrar as estimativas dos parâmetros para ambas as distribuições, será necessário a utilização de algum método numérico de maximização. Em Bollerslev (1986) é utilizado o processo iterativo BHHH para obter as estimativas de máxima verossimilhança. Este método está descrito em (A.2.2).

1.2.3 Verificação

Para um modelo GARCH(r, s), com e_t normal ou t -Student, os resíduos padronizados

$$\tilde{a}_t = \frac{a_t}{\sqrt{h_t}}$$

são v.a. i.i.d. com distribuição normal padrão ou t -Student. Logo uma maneira de verificar se o modelo é adequado, como no caso do modelo ARCH, é calcular a estatística Q de Ljung-Box para a seqüência \tilde{a}_t . Além disso, podemos calcular os coeficientes de assimetria e curtose estimados e fazer um gráfico *quantil* \times *quantil* para avaliar a suposição de normalidade.

Para se verificar se ainda existe heteroscedasticidade condicional nos resíduos, pode-se aplicar o teste ML para a seqüência \tilde{a}_t^2 .

1.2.4 Previsão

Previsões da volatilidade, usando um modelo GARCH, podem ser calculadas de forma similar ao do modelo ARMA. As previsões, com origem t , considerando um modelo GARCH(1,1) da forma (1.20), são dadas por

$$\hat{h}_t(1) = \alpha_0 + \alpha_1 a_t^2 + \beta_1 h_t,$$

e para $l > 1$,

$$\begin{aligned}\hat{h}_t(l) &= \alpha_0 + \alpha_1 \hat{a}_t^2(l-1) + \beta_1 \hat{h}_t(l-1), \\ &= \alpha_0 + \alpha_1 \hat{h}_t(l-1) \hat{e}_t^2(l-1) + \beta_1 \hat{h}_t(l-1),\end{aligned}$$

pois $a_t = \sqrt{\hat{h}_t} e_t$. Substituindo $\hat{e}_t^2(l-1)$ por $E(e_{t+l-1}^2) = 1$, temos que

$$\hat{h}_t(l) = \alpha_0 + (\alpha_1 + \beta_1) \hat{h}_t(l-1), \quad l > 1. \quad (1.31)$$

1.3 Modelo EGARCH

Os modelos GARCH foram os modelos econométricos mais amplamente usados para descrever as mudanças na variância condicional dos retornos. Nestes modelos, a forma da variância condicional leva em conta apenas a magnitude dos retornos, não importando assim sua positividade ou negatividade.

Assim uma de suas características é o tratamento de forma simétrica dos retornos positivos e negativos, encontrada na formulação da variância. Isto pode ser visto, devido à forma quadrática dos retornos encontrada na equação da variância.

Porém muitos pesquisadores encontraram evidências de que os choques dos retornos são negativamente correlacionados com a volatilidade, isto é, a volatilidade tende a aumentar em resposta aos retornos negativos.

Com isso, os modelos GARCH possuem a fraqueza de não conseguir captar a influência assimétrica dos retornos, na modelagem da volatilidade das séries financeiras.

Contudo, existe uma outra limitação destes modelos devido aos seus coeficientes que não devem ser negativos. Esta imposição feita sobre os coeficientes, garante que a formulação da variância permaneça não negativa com probabilidade um. Entretanto, muitas vezes esta imposição não é satisfeita ou dificulta muito a estimação destes parâmetros.

Para solucionar todas estas fraquezas, Nelson (1991) propôs o modelo denominado “exponential GARCH” (EGARCH), que trata de forma assimétrica os retornos positivos e negativo através da função

$$g(e_t) = \theta e_t + \gamma [|e_t| - E(|e_t|)], \quad (1.32)$$

onde θ e γ são constantes reais. Ambos, e_t e $|e_t| - E(|e_t|)$ são sequências com distribuição contínua i.i.d. com média zero. Então $E[g(e_t)] = 0$. A assimetria de $g(e_t)$ pode ser vista facilmente reescrevendo a função como

$$g(e_t) = \begin{cases} (\theta + \gamma)e_t - \gamma E(|e_t|) & \text{se } e_t \geq 0, \\ (\theta - \gamma)e_t - \gamma E(|e_t|) & \text{se } e_t < 0. \end{cases}$$

Para e_t seguindo uma distribuição normal padrão temos que

$$\begin{aligned}
 E(|e_t|) &= \int_{-\infty}^{+\infty} |e_t| f(e_t) de_t \\
 &= \int_{-\infty}^{+\infty} |e_t| \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{e_t^2}{2}\right\} de_t \\
 &= 2 \int_0^{+\infty} e_t \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{e_t^2}{2}\right\} de_t \\
 &= -2 \int_0^{-\infty} \frac{1}{\sqrt{2\pi}} \exp\{u\} du \\
 &= \sqrt{2/\pi}.
 \end{aligned}$$

Para e_t seguindo uma distribuição t-Student padrão com v graus de liberdade, cuja a função de densidade é dado por

$$f(e_t) = \frac{\Gamma[(v+1)/2]}{\sqrt{\pi(v-2)}\Gamma(v/2)} \left(1 + \frac{e_t^2}{v-2}\right)^{-(v+1)/2}$$

temos que

$$\begin{aligned}
 E(|e_t|) &= \int_{-\infty}^{+\infty} |e_t| f(e_t) de_t \\
 &= \int_{-\infty}^{+\infty} |e_t| \frac{\Gamma[(v+1)/2]}{\sqrt{\pi(v-2)}\Gamma(v/2)} \left(1 + \frac{e_t^2}{v-2}\right)^{-(v+1)/2} de_t \\
 &= 2 \frac{\Gamma[(v+1)/2]}{\sqrt{\pi(v-2)}\Gamma(v/2)} \int_0^{+\infty} e_t \left(1 + \frac{e_t^2}{v-2}\right)^{-(v+1)/2} de_t \\
 &= 2 \frac{\Gamma[(v+1)/2]}{\sqrt{\pi(v-2)}\Gamma(v/2)} \int_0^{+\infty} \frac{r-2}{2} u^{-(v+1)/2} du \\
 &= \frac{2\sqrt{v-2} \Gamma((v+1)/2)}{(v-1)\Gamma(v/2)\sqrt{\pi}}.
 \end{aligned}$$

Um modelo EGARCH(r, s) é escrito como

$$a_t = \sqrt{h_t} e_t, \quad (1.33)$$

$$\ln(h_t) = \alpha_0 + \frac{1 + \beta_1 B + \dots + \beta_s B^s}{1 - \alpha_1 B - \dots - \alpha_r B^r} g(e_{t-1}), \quad (1.34)$$

onde α_0 é uma constante, B é o operador retroativo, tal que, $Bg(e_t) = g(e_{t-1})$ e $\beta(B) = 1 + \beta_1 B + \dots + \beta_s B^s$ e $\alpha(B) = 1 - \alpha_1 B - \dots - \alpha_r B^r$ são polinômios com raízes fora do círculo unitário e sem fatores comuns. Para os zeros dos polinômios

$\alpha(B)$ e $\beta(B)$ estarem fora do círculo unitário, significa que os valores absolutos devem ser menores que um. Agora a equação (1.34) usa a usual parametrização ARMA para descrever a evolução da variância condicional de a_t . Baseado nesta observação, algumas propriedades dos modelos EGARCH podem ser obtidas de maneira similar aos modelos GARCH.

O modelo EGARCH difere do GARCH por vários motivos. Primeiro, ele usa o logaritmo da variância condicional para relaxar necessidade dos coeficientes do modelo serem positivos. Segundo, o uso de $g(e_t)$ capacita o modelo a responder de forma assimétrica aos valores positivos e negativos de a_t .

Além disso, o modelo EGARCH original utiliza α_t no lugar de α_0 , permitindo ao modelo variar no tempo. A forma utilizada no artigo de Nelson (1991) foi $\alpha_t = \alpha + \ln(1 + N_t\delta)$, com N_t sendo o número de dias entre $t - 1$ e t e os parâmetros α e δ como constantes. Este formato permitiu ao modelo captar a maior variação após os dias sem negociação. Outras propriedades adicionais estão em Nelson (1991).

Podemos escrever a equação(1.34) como

$$\ln(h_t) = \alpha_* + \sum_{i=0}^s \beta_i g(e_{t-i-1}) + \sum_{j=1}^r \alpha_j \ln(h_{t-j}), \quad (1.35)$$

onde $\alpha_* = \alpha_0(1 - \alpha_1 - \dots - \alpha_r)$ e $\beta_0 = 1$.

Substituindo em (1.35), $g(e_t)$ por sua equação e supondo normalidade de e_t , temos que uma forma alternativa para o modelo EGARCH(r, s) dada por

$$\begin{aligned} \ln(h_t) &= \alpha_* + \sum_{i=0}^s \beta_i (\theta e_{t-i-1} + \gamma_i [|e_{t-i-1}| - E(|e_{t-i-1}|)]) + \sum_{j=1}^r \alpha_j \ln(h_{t-j}) \\ &= \omega + \sum_{i=0}^s \beta_i^* \left(\frac{\gamma_i^* a_{t-i-1} + |a_{t-i-1}|}{\sqrt{h_{t-i-1}}} \right) + \sum_{j=1}^r \alpha_j \ln(h_{t-j}), \end{aligned}$$

onde $\omega = \alpha_* - \sum_{i=0}^s \beta_i \gamma_i \sqrt{2/\pi}$, $\gamma_i^* = \theta/\gamma_i$ e $\beta_i^* = \beta_i \gamma_i$.

O programa S-Plus utiliza uma forma parecida para o modelo EGARCH(r, s), onde a forma da variância é dada por

$$\ln(h_t) = \alpha_0 + \sum_{i=1}^s \alpha_i \frac{\gamma_i a_{t-i} + |a_{t-i}|}{\sqrt{h_{t-i}}} + \sum_{j=1}^r \beta_j \ln(h_{t-j}). \quad (1.36)$$

Aqui a_{t-i} positivo contribui com $\alpha_i(1 + \gamma_i)|e_{t-i}|$ para o log da volatilidade, contudo a_{t-i} negativo contribui com $\alpha_i(1 - \gamma_i)|e_{t-i}|$, onde $e_{t-i} = a_{t-i}/\sqrt{h_{t-i}}$. Os parâmetros γ_i então significam os efeitos de alavanca dos a_{t-i} . Agora, esperamos que γ_i sejam negativos nas aplicações reais.

Como no caso dos modelos GARCH, podemos obter a média e variância condicionais do modelo EGARCH, dados por

$$\begin{aligned} E(a_t|F_{t-1}) &= E(\sqrt{h_t}e_t|F_{t-1}) = \sqrt{h_t}E(e_t|F_{t-1}) = 0, \\ Var(a_t|F_{t-1}) &= E(a_t^2|F_{t-1}) = E(h_t e_t^2|F_{t-1}) = h_t E(e_t^2|F_{t-1}) = h_t E(e_t^2) = h_t. \end{aligned}$$

Com isso, temos que a média dos retornos é igual a zero ($E(a_t) = 0$). Logo, calculando a covariância de a_t e a_{t+k} temos que

$$\begin{aligned} Cov(a_t, a_{t+k}) &= E(a_t a_{t+k}) = E[E(a_t a_{t+k}|F_{t+k-1})] \\ &= E[a_t E(a_{t+k}|F_{t+k-1})] = E[a_t E(\sqrt{h_{t+k}} e_{t+k}|F_{t+k-1})] \\ &= E[a_t \sqrt{h_{t+k}} E(e_{t+k}|F_{t+k-1})] = 0, \end{aligned}$$

para $k > 0$, pois a_t e $\sqrt{h_{t+k}}$ estão em F_{t+k-1} e $E(e_{t+k}|F_{t+k-1}) = 0$.

Dessa forma a função de autocovariância de a_t será

$$\gamma_{a_t}(k) = 0, \quad k \geq 1,$$

indicando que a_t é uma seqüência de variáveis não correlacionadas (ruído branco), com média zero e variância condicional dada por h_t .

Para entender melhor o modelo EGARCH, vamos considerar o modelo simples de ordem (1,0). Portanto o modelo EGARCH(1,0) será dado pelas equações

$$(1 - \alpha B) \ln(h_t) = (1 - \alpha) \alpha_0 + g(e_{t-1}), \quad (1.37)$$

onde o e_t é uma normal padrão i.i.d. e o subscrito de α_1 foi omitido. Neste caso, $E(|e_t|) = \sqrt{2/\pi}$ e o modelo para $\ln(h_t)$ é determinado por

$$(1 - \alpha B) \ln(h_t) = \begin{cases} \alpha_* + (\gamma + \theta) e_{t-1} & \text{se } e_{t-1} \geq 0, \\ \alpha_* + (\gamma - \theta) (-e_{t-1}) & \text{se } e_{t-1} < 0, \end{cases} \quad (1.38)$$

onde $\alpha_* = (1 - \alpha) \alpha_0 - \gamma \sqrt{2/\pi}$.

Como desejamos que os retornos negativos tenham maior impacto sobre a volatilidade, devemos então fazer $\gamma < 0$. Com esta assimetria pode-se permitir que o modelo responda mais rapidamente a retornos negativos do que a retornos positivos, fato este conhecido como “efeito de alavancagem”.

Assim temos que

$$h_t = \begin{cases} \frac{\exp\{\alpha_*\}}{\exp\{\gamma\sqrt{2/\pi}\}} \exp\{(\gamma + \theta)e_{t-1}\} h_{t-1}^\alpha & \text{se } e_{t-1} \geq 0, \\ \frac{\exp\{\alpha_*\}}{\exp\{\gamma\sqrt{2/\pi}\}} \exp\{(\gamma - \theta)e_{t-1}\} h_{t-1}^\alpha & \text{se } e_{t-1} < 0. \end{cases}$$

Portanto a variação h_t é maior em resposta a retornos negativos, considerando $e_t = a_t/h_t$ e $\gamma < 0$.

1.3.1 Identificação do Modelo EGARCH

Os modelos EGARCH, consideram que a série é não correlacionada serialmente, então temos que o primeiro passo é tentar retirar toda a autocorrelação da série, se esta existir, ajustando um modelo ARMA. No caso da série ser autocorrelacionada, temos a seguinte expressão

$$\phi(B)X_t = \theta_0 + \theta(B)a_t,$$

onde $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$, $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ e $a_t \sim \text{EGARCH}(r, s)$. Logo, quando nos referirmos a a_t , estaremos supondo ou que a série é não correlacionada, ou então ela é o resíduo da aplicação de um modelo ARMA à série original (X_t). No caso da série correlacionada X_t não ser estacionária, podemos aplicar a ela algumas diferenças $(1 - B)^d$, para assim atingir a estacionaridade. Com isso, estaremos ajustando um modelo ARIMA(p, d, q), onde d é o número de diferenças. Logo a expressão para este modelo é dada por

$$\phi(B)\nabla^d X_t = \theta_0 + \theta(B)a_t,$$

onde $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$, $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$, $\nabla = (1 - B)$, X_t é a série original dos retornos e $a_t \sim \text{EGARCH}(r, s)$ é o resíduo do modelo.

Para verificarmos se a série apresenta heteroscedasticidade condicional, podemos utilizar dois testes para série a_t^2 :

- (i) Teste de Box-Pierce-Ljung A.1.1;
- (ii) Teste de multiplicadores de Lagrange (ML) A.1.2.

A f.a.c. de a_t^2 pode ser utilizado para verificar se a série segue um processo EGARCH, onde seu comportamento é marcado por um decaimento exponencial.

Para se ter uma idéia da ordem do modelo EGARCH(r, s), podemos utilizar a mesma técnica apresentada para os modelos GARCH(r, s), onde observa-se os números de lags significativos da f.a.c. e f.a.c.p. e ajusta-se modelos baseados nestes valores. Assim critérios de informação, como o AIC, serão necessários para discriminação dos melhores modelos.

1.3.2 Estimação

A estimação dos parâmetros do modelo EGARCH(r, s) dados pelas equações (1.33)-(1.34) são obtidas pelo método de máxima verossimilhança condicional, onde a função

de verossimilhança é dada por

$$\begin{aligned} L(a_1, \dots, a_T | \alpha, \beta, \gamma) &= f(a_T | F_{T-1}), \dots, f(a_{r+1} | F_r) f(a_1, \dots, a_r | \alpha, \beta, \gamma) \\ &= \prod_{t=r+1}^T f(a_t | F_{t-1}) f(a_1, \dots, a_r | \alpha, \beta, \gamma), \end{aligned} \quad (1.39)$$

Os erros e_t do modelo EGARCH podem seguir qualquer distribuição com média zero e variância um, segundo a definição do modelo. Apresentaremos agora o desenvolvimento da verossimilhança para o caso das distribuições Normal e t-Student padrões.

Distribuição Normal

Supondo que $e_t \sim N(0, 1)$ temos que a função de densidade condicional de a_t dado as informações passadas segue uma distribuição normal com média zero e variância h_t dada por

$$f(a_t | F_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} \exp \left\{ \frac{-a_t^2}{2h_t} \right\}.$$

Utilizando a verossimilhança dada pela equação (1.39) e assumindo que para T grande, $f(a_1, \dots, a_r | \alpha, \beta, \gamma)$ pode ser desprezado. Então a verossimilhança condicionada aos "r" primeiros retornos é dada por

$$L(a_{r+1}, \dots, a_T | \alpha, \beta, \gamma, a_1, \dots, a_r) = \prod_{t=r+1}^T (\sqrt{2\pi h_t})^{-1} \exp \left\{ \frac{-a_t^2}{2h_t} \right\} \quad (1.40)$$

onde a volatilidade h_t é obtida recursivamente.

Assim, o logaritmo da verossimilhança condicionado às primeiras "r" observações é dada por

$$l(a_{r+1}, \dots, a_T | \alpha, \beta, \gamma, a_1, \dots, a_r) \propto -\frac{1}{2} \sum_{t=r+1}^T \ln(h_t) - \frac{1}{2} \sum_{t=r+1}^T \frac{a_t^2}{h_t}. \quad (1.41)$$

Para encontrar os estimadores de máxima verossimilhança para os parâmetros, temos então que maximizar a função de verossimilhança condicional.

Como no caso de Engle (1982) e Bollerslev (1986), será utilizado a log-verossimilhança média com erros normais (padrão), que facilitam a obtenção de resultados assintóticos, para estimar os parâmetros do modelo (θ) dada por

$$l(\theta) = \frac{1}{T} \sum_{t=1}^T l_t(\theta),$$

onde $l_t(\theta)$ é a log-verossimilhança para o tempo t , que tem a seguinte forma

$$l_t(\theta) \propto -\frac{1}{2} \ln(h_t) - \frac{a_t^2}{2h_t}.$$

No caso particular de um modelo EGARCH(1,0), temos que a verossimilhança condicional é dada por

$$L(a_2, \dots, a_T | \alpha_0, \alpha_1, a_1) = f(a_T | a_{T-1}) f(a_{T-1} | a_{T-2}) \dots f(a_2 | a_1). \quad (1.42)$$

Supondo que

$$(a_t | a_{t-1}) \sim N(0, h_t),$$

e tomando

$$h_t = \exp\{\alpha_0 + \alpha_1 \ln(h_{t-1}) + g(e_{t-1})\}$$

temos que (1.42) é dada então por

$$L(a_2, \dots, a_T | \alpha_0, \alpha_1, a_1) = \prod_{t=2}^T (2\pi h_t)^{-1/2} \exp\left\{\frac{-a_t^2}{2h_t}\right\}.$$

Então a log-verossimilhança será

$$l(a_2, \dots, a_T | \alpha_0, \alpha_1, a_1) \propto -\frac{1}{2} \sum_{t=2}^T \ln(h_t) - \frac{1}{2} \sum_{t=2}^T \left(\frac{a_t^2}{h_t}\right). \quad (1.43)$$

Distribuição t-Student

Supondo que $e_t \sim t_v$ (t de Student com v graus de liberdade) temos que a função de densidade condicional de a_t dado as informações passadas segue também uma distribuição t de Student dada por

$$f(a_t | F_{t-1}) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2) \sqrt{(v-2)\pi h_t}} \left(1 + \frac{a_t^2}{(v-2)h_t}\right)^{-(v+1)/2}.$$

Utilizando a verossimilhança dada pela equação (1.39) e assumindo que para T grande, $f(a_1, \dots, a_r | \alpha, \beta, \gamma)$ pode ser desprezado. Então a verossimilhança condicionada aos “ r ” primeiros retornos é dada por

$$L(a_{r+1}, \dots, a_T | \alpha, \beta, \gamma, a_1, \dots, a_r) = \prod_{t=r+1}^T \frac{\Gamma((v+1)/2)}{\Gamma(v/2) \sqrt{(v-2)\pi h_t}} \left(1 + \frac{a_t^2}{(v-2)h_t}\right)^{-(v+1)/2}. \quad (1.44)$$

onde a volatilidade h_t é obtida recursivamente.

Logo temos que maximizar a função de verossimilhança condicional para encontrar os estimadores de máxima verossimilhança para os parâmetros. Neste caso, podemos estimar os parâmetros fixando um determinado grau de liberdade ou podemos também estimar os graus de liberdades juntamente com os parâmetros.

Para encontrar as estimativas dos parâmetros, podemos utilizar o processo iterativo BHHH para obter as estimativas de máxima verossimilhança, como em Bollerslev (1986). Este método está descrito em (A.2.2).

1.3.3 Verificação

Para um modelo EGARCH(r, s), com e_t normal ou t -Student, os resíduos padronizados

$$\tilde{a}_t = \frac{a_t}{\sqrt{h_t}}$$

são v.a. i.i.d. com distribuição normal padrão ou t -Student. Logo, uma maneira de verificar se o modelo é adequado é calcular a estatística Q de Ljung-Box para a seqüência \tilde{a}_t .

Para se verificar se ainda existe heteroscedasticidade condicional nos resíduos, pode-se aplicar o teste ML para a seqüência \tilde{a}_t^2 .

1.3.4 Previsão

Para ilustrar a previsão do modelo EGARCH, utilizaremos o modelo EGARCH(1,1), com parâmetros conhecidos e os e_t sejam de uma normal padrão. Para tal modelo, temos

$$\begin{aligned} \ln(h_t) &= (1 - \alpha_1)\alpha_0 + \alpha_1 \ln(h_{t-1}) + g(e_{t-1}), \\ g(e_{t-1}) &= \theta e_{t-1} + \gamma(|e_{t-1}| - \sqrt{2/\pi}). \end{aligned}$$

Tomando a exponencial, o modelo fica

$$\begin{aligned} h_t &= h_{t-1}^{\alpha_1} \exp\{(1 - \alpha_1)\alpha_0\} \exp\{g(e_{t-1})\}, \\ g(e_{t-1}) &= \theta e_{t-1} + \gamma(|e_{t-1}| - \sqrt{2/\pi}). \end{aligned} \tag{1.45}$$

Seja t a previsão original. Para uma previsão de 1 - *Passo* a frente, temos

$$h_{t+1} = h_t^{\alpha_1} \exp\{(1 - \alpha_1)\alpha_0\} \exp\{g(e_t)\},$$

onde todas as quantidades do lado direito são conhecidas. Então, a previsão de 1 – *Passo* a frente para a volatilidade, na origem t é simplesmente $\hat{h}_t(1) = h_{t+1}$ dada anteriormente. Para o 2 – *Passos* a frente da previsão, a eq. (1.45) dada por

$$h_{t+2} = h_{t+1}^{\alpha_1} \exp\{(1 - \alpha_1)\alpha_0\} \exp\{g(e_{t+1})\}.$$

Tomando a esperança condicional do tempo t , temos

$$h_t(2) = h_t^{\alpha_1} \exp\{(1 - \alpha_1)\alpha_0\} E_t[\exp\{g(e_{t+1})\}],$$

onde E_t denota a esperança condicional dada para o tempo original t . A esperança a priori pode ser obtida da seguinte forma

$$\begin{aligned} E\{\exp[g(e_t)]\} &= \int_{-\infty}^{\infty} \exp[\theta e_t + \gamma(|e_t| - \sqrt{2/\pi})] f(e_t) de_t \\ &= \exp(-\gamma\sqrt{2/\pi}) \left[\int_0^{\infty} \exp\{(\theta + \gamma)e_t\} \frac{1}{\sqrt{2\pi}} \exp\{-e_t^2/2\} de_t \right. \\ &\quad \left. + \int_{-\infty}^0 \exp\{(\theta + \gamma)e_t\} \frac{1}{\sqrt{2\pi}} \exp\{-e_t^2/2\} de_t \right] \\ &= \exp(-\gamma\sqrt{2/\pi}) \left[\exp\{(\theta + \gamma)^2/2\} \Phi(\theta + \gamma) + \exp\{(\theta - \gamma)^2/2\} \Phi(\gamma - \theta) \right], \end{aligned}$$

onde $f(e_t)$ e $\Phi(x)$ são a função de densidade e função densidade acumulada da distribuição normal padrão, respectivamente. Portanto, a previsão da volatilidade para 2 – *Passos* a frente é

$$\begin{aligned} h_t(2) &= \hat{h}_t^{\alpha_1}(1) \exp[(1 - \alpha_1)\alpha_0 - \gamma\sqrt{2/\pi}] \\ &\quad \times \{ \exp[(\theta + \gamma)^2/2] \Phi(\theta + \gamma) + \exp[(\theta - \gamma)^2/2] \Phi(\gamma - \theta) \}. \end{aligned}$$

Repetindo o processo de previsão, obtemos recursivamente a fórmula para previsão de j – *Passos* a frente:

$$\begin{aligned} h_t(j) &= \hat{h}_t^{\alpha_1}(j-1) \exp(\omega) \\ &\quad \times \{ \exp[(\theta + \gamma)^2/2] \Phi(\theta + \gamma) + \exp[(\theta - \gamma)^2/2] \Phi(\gamma - \theta) \}, \end{aligned}$$

onde $\omega = (1 - \alpha_1)\alpha_0 - \gamma\sqrt{2/\pi}$.

Estudo das Propriedades dos Estimadores dos Parâmetros via Simulação

Neste capítulo apresentaremos um estudo de simulação das propriedades dos estimadores dos parâmetros para os modelos apresentados até agora, como a convergência assintótica para o verdadeiro valor e a probabilidade de cobertura dos parâmetros.

Como o verdadeiro valor dos parâmetros de uma série é desconhecido, optou-se então por fazer um estudo de simulação, onde gera-se séries a partir de um modelo pré-determinado (coeficientes fixos e conhecidos) e então calcula-se o valor das estimativas para os parâmetros. Com isso, verificamos a existência ou não da convergência assintótica para o valor pré-determinado dos coeficientes, além de avaliarmos a probabilidade de cobertura dos coeficientes.

A probabilidade de cobertura será definida como o número de vezes que o intervalo de confiança, construído a partir das simulações, conteve o valor pré-determinado do parâmetro.

Para construirmos o intervalo de confiança para os coeficientes, usaremos a Informação observada de Fisher, definida por

$$I(\theta) = - \sum_t \frac{1}{T} \left[\frac{\partial^2 l_t(\theta)}{\partial \theta^2} \Big|_{\theta=\hat{\theta}} \right],$$

onde $l(\theta)$ é o log da verossimilhança.

Considerando que assintoticamente a distribuição dos estimadores é dado por

$$\sqrt{T}(\hat{\theta} - \theta) \sim N(0, I^{-1}(\theta)),$$

logo o intervalo de confiança para θ será dado por $\hat{\theta} \pm z_{\alpha/2} I^{-1/2}(\theta)$, onde $z_{\alpha/2}$ é o quantil $(1 - \alpha/2)$ da normal e $(1 - \alpha)$ é nível de confiança.

2.1 Modelo ARCH(1)

Nesta seção iremos analisar a convergência e a probabilidade de cobertura para os parâmetros do modelo ARCH(1), através de um estudo de simulação.

O modelo ARCH(1) é dado por

$$\begin{aligned} a_t &= \sqrt{h_t} e_t \\ h_t &= \alpha_0 + \alpha_1 a_{t-1}^2, \end{aligned}$$

com $0 \leq \alpha_1 < 1$ e $e_t \sim N(0, 1)$.

Para simular algumas séries que seguem este processo, precisamos fixar os valores dos parâmetros. Então, iremos atribuir dois diferentes valores para os coeficientes do modelo ARCH(1) apresentados na Tabela 2.1, no qual o primeiro processo terá uma menor influência do valor passado em relação ao segundo.

Modelo	α_0	α_1
1	1,0	0,5
2	1,0	0,9

Tabela 2.1: Parâmetros dos Modelos

Contudo, para estimar a variância de cada parâmetro dos modelos, utilizaremos a Informação observada de Fisher que será dado por

$$I(\alpha_0) = \frac{1}{T} \sum_{t=2}^T \frac{1}{2h_t^2},$$

$$I(\alpha_1) = \frac{1}{T} \sum_{t=2}^T \frac{1}{2h_t^2} (a_{t-1}^2)^2.$$

2.1.1 Convergência

A partir dos coeficientes especificados na Tabela 2.1, simulou-se uma série de tamanho amostral $n = 20500$ com ponto de início $a_0 = 0$, para cada modelo.

Fazendo um *burn in* de 500 observações, para tirar a influência do ponto inicial, estimamos o modelo ARCH(1) com os primeiros 500 dados (isto é, a_{501} até a_{1000}). Repetindo este processo, aumentando o tamanho da amostra de 100 em 100 até o fim da amostra total (isto é, de a_{501} a a_{1100} , de a_{501} a a_{1200} , ..., de a_{501} a a_{20500}), estimamos então os parâmetros do modelo. Diante disso teremos 200 séries com tamanhos variando de 500 até 20500, observando dessa forma se existe a convergência assintótica.

Modelo 1

As Figuras 2.1e 2.2 mostram a existência da convergência assintótica dos parâmetros devido a estabilização das estimativas em torno dos valores centrais que são os valores estabelecidos para os coeficientes do modelo e nas periferias estão os limites inferior e superior das estimativas (com 95% de confiança).

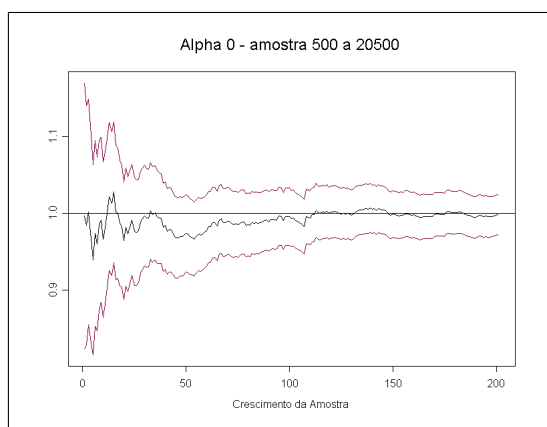


Figura 2.1: Convergência-Alpha 0

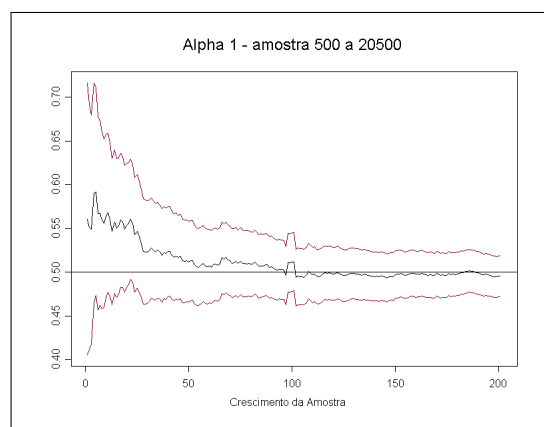


Figura 2.2: Convergência-Alpha 1

Modelo 2

As Figuras 2.3 e 2.4 mostram a existência da convergência assintótica dos parâmetros para o valor central correspondente ao valor do parametro e nas periferias estão os limites inferior e superior das estimativas.

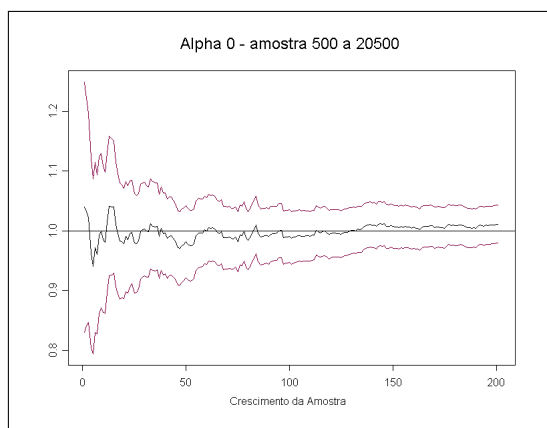


Figura 2.3: Convergência-Alpha 0

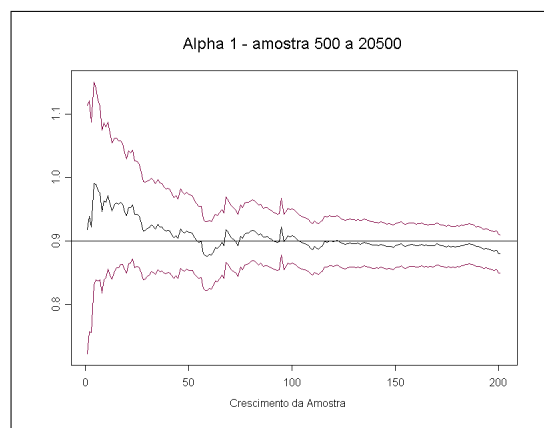


Figura 2.4: Convergência-Alpha 1

2.1.2 Probabilidade de Cobertura

Para estudarmos a probabilidade de cobertura, simulamos 100 séries de tamanho $n = 10500$ com ponto de início $a_0 = 0$. Fazendo um *burn in* de 500 observações em cada série, para tirar a influência do ponto inicial, estimou-se o modelo ARCH(1) com as primeiras 500 observações (isto é, a_{501} até a_{1000}) para cada série. Repetindo este processo, aumentando o tamanho das amostras de 100 em 100 até o final da amostra de cada série (isto é, de a_{501} a a_{1100} , de a_{501} a a_{1200} , ..., de a_{501} a a_{10500}), ficaremos assim com 100 grupos de cada uma das 100 séries com tamanhos que variam de 500 a 10500.

Tomando o grupo das séries com o mesmo tamanho amostral, construímos os intervalos de confiança dos coeficientes para cada série, avaliando assim a probabilidade de cobertura, que é determinada pela porcentagem de intervalos que contém o verdadeiro valor do parâmetro.

Deseja-se verificar que, ao aumentar o tamanho da amostra (comprimento da série), a probabilidade de cobertura aumentará e se estabilizará em torno da confiança estabelecida. Para isto foi utilizado os 100 grupos de 100 séries com tamanho variando de 500 a 10500.

Modelo 1

A Figura 2.5 mostra o resultado da probabilidade de cobertura para o modelo ARCH(1), juntamente com o comprimento médio dos intervalos de confiança, para cada grupo de séries.

Podemos notar que a probabilidade de cobertura não está se mantendo próximo dos 95% para α_0 e α_1 como era o esperado. A probabilidade de cobertura parece estar sendo subestimada pelo modelo.

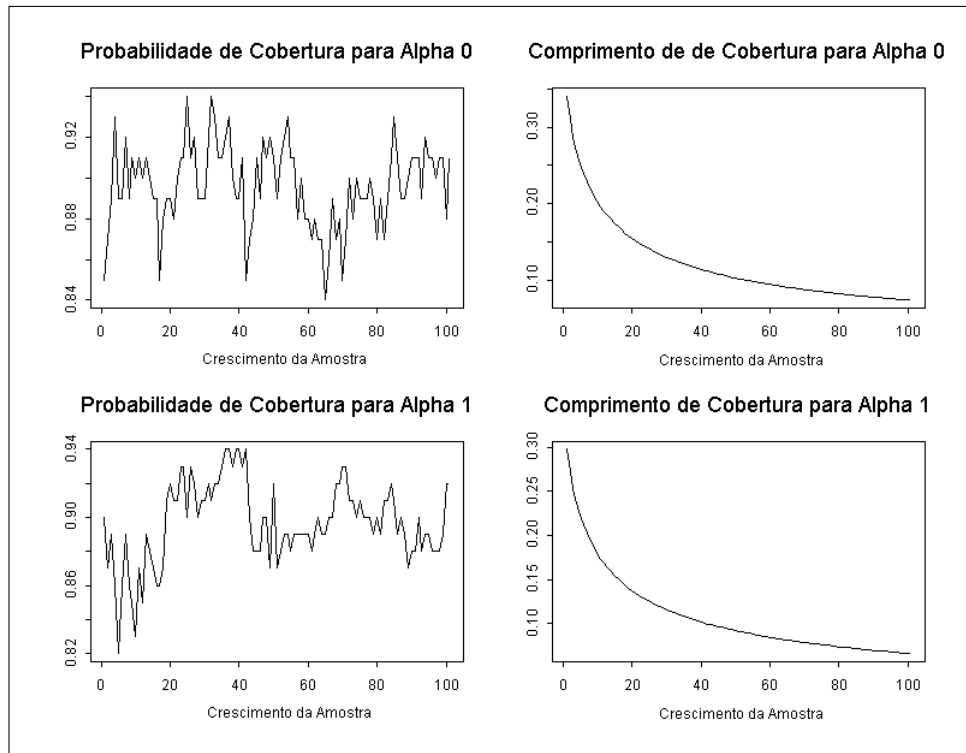


Figura 2.5: Probabilidade de Cobertura

Modelo 2

A Figura 2.6 mostra o resultado da probabilidade de cobertura para o modelo ARCH(1), juntamente com o comprimento médio dos intervalos de confiança, para cada grupo de séries.

Podemos notar que a probabilidade de cobertura para α_0 está em torno dos 95%, como esperado. Porém, ao aumentarmos o tamanho da amostra esta probabilidade de cobertura fica abaixo do esperado e isso pode ser explicado pelo comprimento do intervalo que teve uma certa diminuição. Já a probabilidade de cobertura para α_1 está se mantendo aleatoriamente em torno dos 95% como esperado.

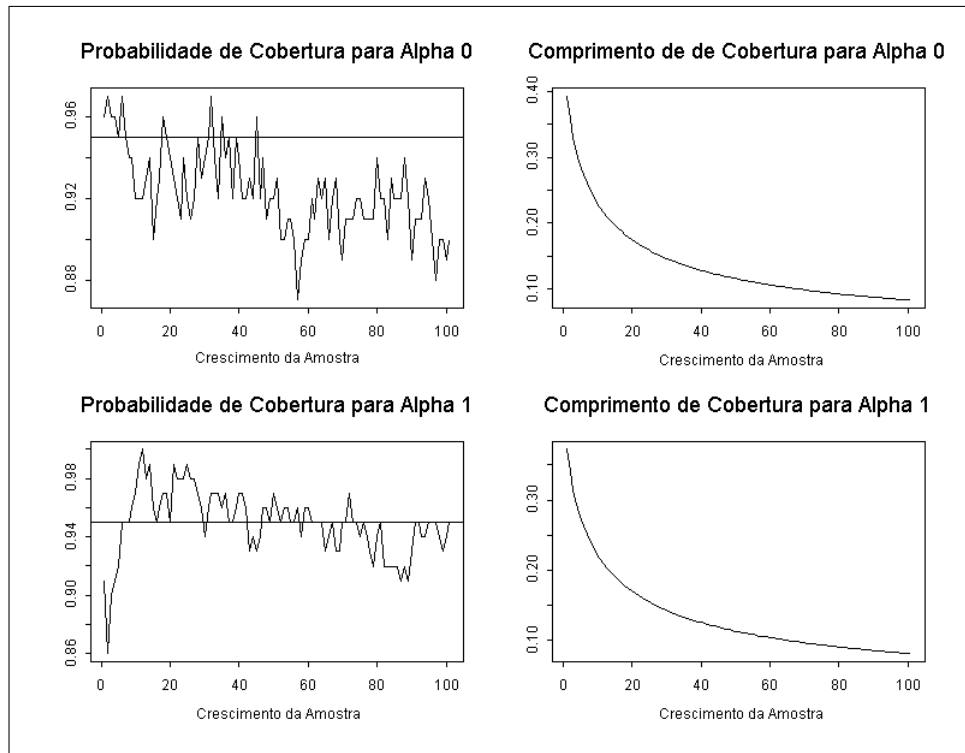


Figura 2.6: Probabilidade de Cobertura

2.2 Modelo GARCH(1,1)

Nesta seção iremos analisar a convergência e a probabilidade de cobertura para os parâmetros do modelo GARCH(1,1), através de um estudo de simulação.

O modelo GARCH(1,1) é dado por

$$a_t = \sqrt{h_t}e_t$$

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1},$$

com $0 \leq \alpha_1, \beta_1 < 1$, $\alpha_1 + \beta_1 < 1$ e $e_t \sim N(0, 1)$.

Para simular algumas séries que seguem este processo, precisamos fixar os valores dos parâmetros. Então, iremos atribuir dois diferentes valores para os coeficientes do modelo GARCH(1,1) apresentados na Tabela 2.2, no qual o primeiro processo terá uma persistência menor em relação ao segundo.

Modelo	α_0	α_1	β_1
1	1,0	0,7	0,2
2	1,0	0,2	0,7

Tabela 2.2: Parâmetros dos Modelos

Contudo, para estimar a variância de cada parâmetro dos modelos, utilizaremos a Informação observada de Fisher que será dado por

$$I(\alpha_0) = \frac{1}{T} \sum_{t=2}^T \frac{1}{2h_t^2} \left(\frac{\partial h_t}{\partial \alpha_0} \right)^2,$$

$$I(\alpha_1) = \frac{1}{T} \sum_{t=2}^T \frac{1}{2h_t^2} \left(\frac{\partial h_t}{\partial \alpha_1} \right)^2,$$

$$I(\beta_1) = \frac{1}{T} \sum_{t=2}^T \frac{1}{2h_t^2} \left(\frac{\partial h_t}{\partial \beta_1} \right)^2.$$

Para o modelo GARCH(1, 1) temos que os termos recursivos são dados por

$$\frac{\partial h_t}{\partial \alpha_0} = 1 + \beta_1 \frac{\partial h_{t-1}}{\partial \alpha_0},$$

$$\frac{\partial h_t}{\partial \alpha_1} = a_{t-1}^2 + \beta_1 \frac{\partial h_{t-1}}{\partial \alpha_1},$$

$$\frac{\partial h_t}{\partial \beta_1} = h_{t-1} + \beta_1 \frac{\partial h_{t-1}}{\partial \beta_1}.$$

2.2.1 Distribuições dos Lags dos Modelos

As Figuras 2.7 e 2.8 mostram as distribuições dos “lags” para os modelos GARCH(1, 1) definido com os parâmetros pela Tabela 2.2. Pode-se observar que a influência dos coeficientes não é muito longa no Modelo 1, visto que a partir do “lag 4” ela é praticamente nula, já no Modelo 2 a influência é muito maior, sendo considerada nula somente a partir do “lag 14”.

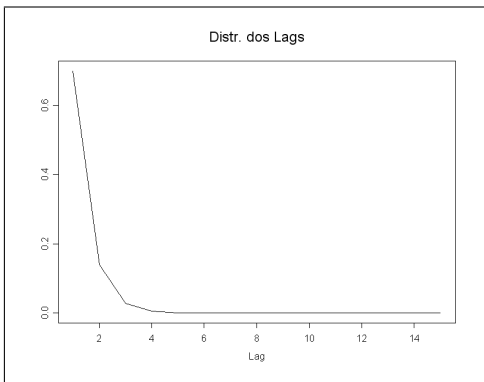


Figura 2.7: Lags Modelo 1

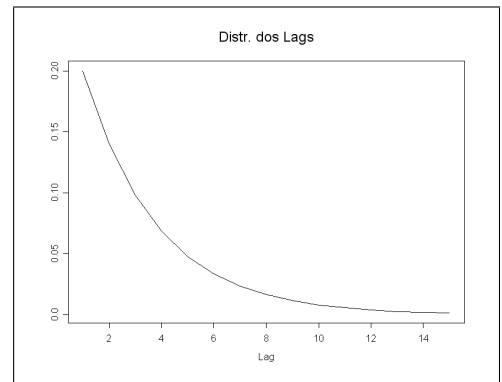


Figura 2.8: Lags Modelo 2

2.2.2 Convergência

A partir dos coeficientes especificados na Tabela 2.2, simularemos uma série de tamanho amostral $n = 20500$ com ponto de início $a_0 = 0$, para cada modelo. Fazendo um *burn in* de 500 observações, para tirar a influência do ponto inicial, iremos estimar o modelo GARCH(1, 1) com os primeiros 500 dados (isto é, a_{501} até a_{1000}). Repetindo este processo, aumentando o tamanho da amostra de 100 em 100 até o fim da amostra total (isto é, de a_{501} a a_{1100} , de a_{501} a a_{1200} , ..., de a_{501} a a_{20500}), iremos então estimar os parâmetros do modelo. Diante disso teremos 200 séries com tamanhos variando de 500 até 20500, observando dessa forma se existe a convergência assintótica.

Modelo 1

As Figuras 2.9, 2.10 e 2.11 mostram a existência da convergência assintótica dos parâmetros onde, no centro estão o valor verdadeiro e a estimativa do parâmetro e nas periferias estão os limites inferior e superior para as estimativas dos coeficientes.

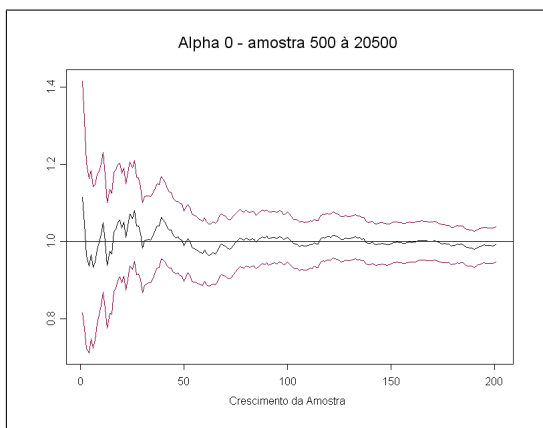


Figura 2.9: Convergência-Alpha 0

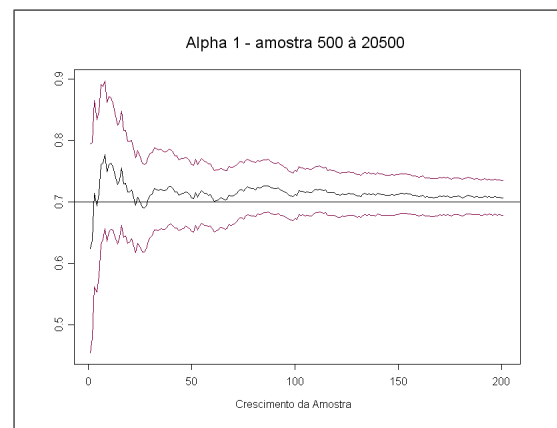


Figura 2.10: Convergência-Alpha 1

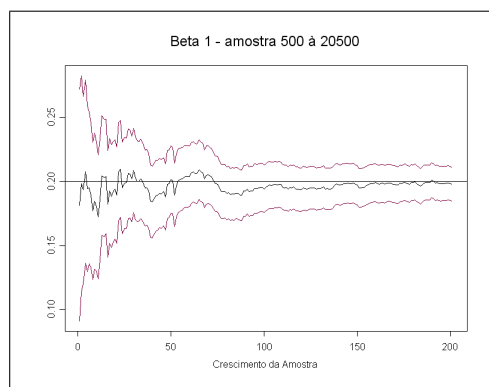
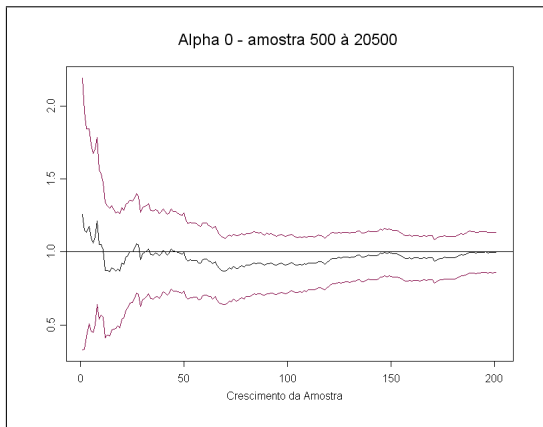
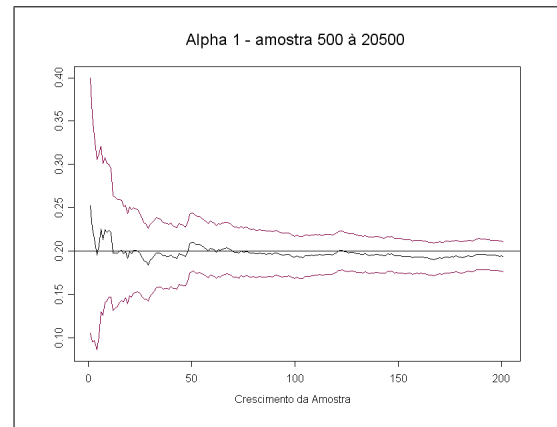
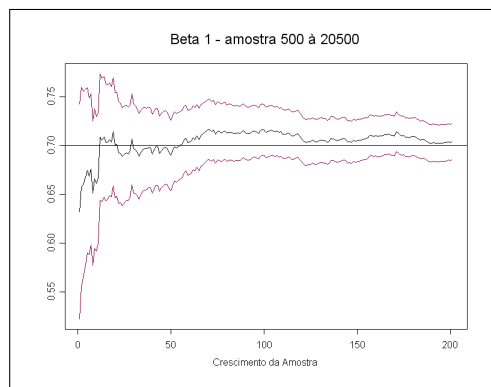


Figura 2.11: Convergência-Beta 1

Modelo 2

As Figuras 2.12, 2.13 e 2.14 mostram a existência da convergência assintótica dos parâmetros onde, no centro estão o valor verdadeiro e a estimativa do parâmetro e nas periferias estão os limites inferior e superior para as estimativas dos coeficientes.

**Figura 2.12:** Convergência-Alpha 0**Figura 2.13:** Convergência-Alpha 1**Figura 2.14:** Convergência-Beta 1

2.2.3 Probabilidade de Cobertura

Para estudarmos a probabilidade de cobertura, simularemos 100 séries de tamanho $n = 10500$ com ponto de início $a_0 = 0$. Fazendo um *burn in* de 500 observações em cada série, para tirar a influência do ponto inicial, iremos estimar o modelo GARCH(1,1) com as primeiras 500 observações (isto é, a_{501} até a_{1000}) para cada série. Repetindo este processo, aumentando os tamanhos das amostras de 100 em 100 até o final da amostra de cada série (isto é, de a_{501} a a_{1100} , de a_{501} a a_{1200} , ..., de a_{501} a a_{10500}), ficaremos assim com 100 grupos de cada uma das 100 séries com tamanho que variam de 500 a 10500.

Portanto, tomando o grupo de séries com o mesmo tamanho amostral, construiremos os intervalos de confiança dos coeficientes para cada série, avaliando assim a probabilidade de cobertura, que é definida como a porcentagem de intervalos que contém o verdadeiro valor do parâmetro.

Deseja-se verificar que ao aumentar o tamanho da amostra (comprimento da série), a probabilidade de cobertura aumentará e se estabilizará em torno da confiança estabelecida. Para isto foi utilizado os 100 grupos de 100 séries com tamanho variando de 500 a 10500.

Modelo 1

A Figura 2.15 mostra o resultado da probabilidade de cobertura para o modelo GARCH(1,1), juntamente com o comprimento médio dos intervalos de confiança, para cada grupo de séries.

Podemos notar que a probabilidade de cobertura não está se mantendo próximo dos 95% para α_0 e β_1 como era o esperado. Porém para α_0 a probabilidade de cobertura parece estar convergindo para os 95% conforme aumentamos a amostra. Já no caso de β_1 , além da probabilidade estar aumentando conforme aumentamos o tamanho da amostra, percebemos também que o comprimento do intervalo é bem menor que a dos demais, podendo assim ser um fato a explicar esta probabilidade baixa de cobertura.

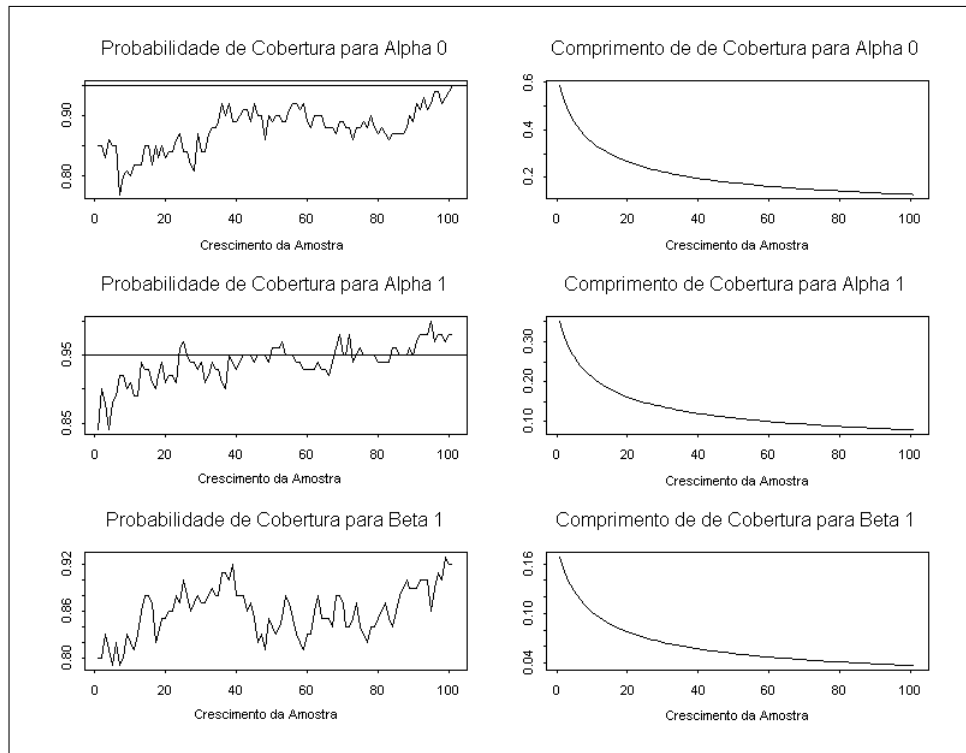


Figura 2.15: Probabilidade de Cobertura

Modelo 2

A Figura 2.16 mostra o resultado da probabilidade de cobertura para o modelo GARCH(1,1), juntamente com o comprimento médio dos intervalos de confiança, para cada grupo de séries.

Podemos notar que a probabilidade de cobertura não está se mantendo próximo dos 95% para β_1 como era o esperado. Porém para o comprimento do intervalo é bem pequeno (em relação ao valor do parâmetro), podendo assim ser um fato a explicar esta probabilidade baixa de cobertura em relação aos outros.

Já a probabilidade de cobertura para α_1 está muito acima dos 95%. Este fato pode ser explicado também pelo comprimento do intervalo, que é grande em relação ao valor do parâmetro.

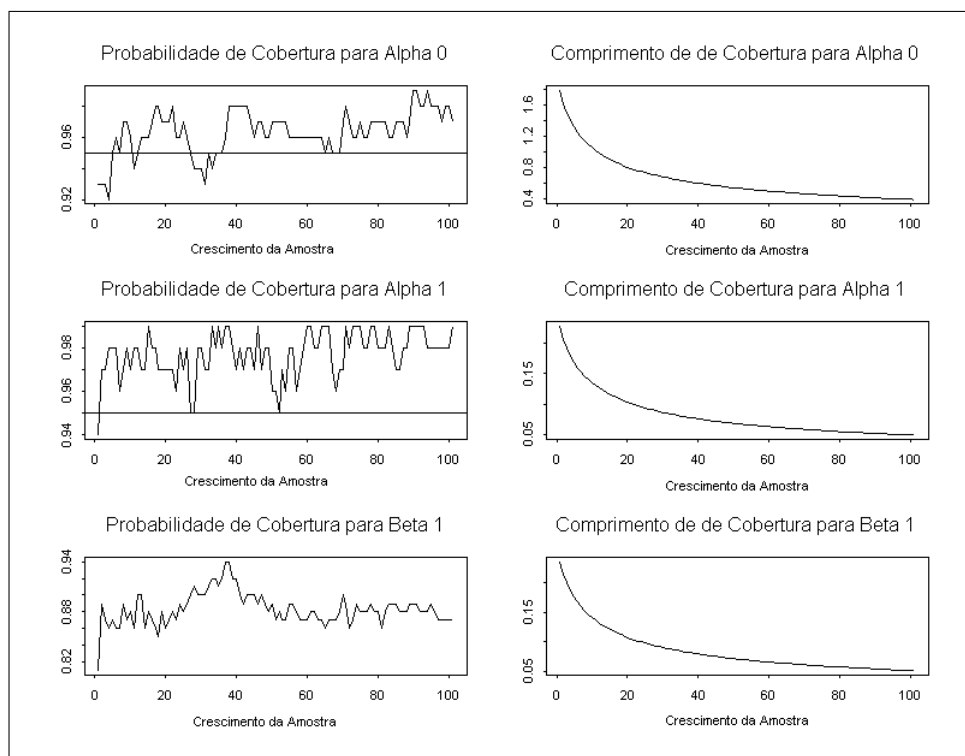


Figura 2.16: Probabilidade de Cobertura

2.3 Modelo EGARCH(1, 1)

Nesta seção iremos analisar a convergência e a probabilidade de cobertura para os parâmetros do modelo EGARCH(1, 1), através de um estudo de simulação.

O modelo EGARCH(1, 1) especificado no S-Plus é dado por

$$a_t = \sqrt{h_t} e_t$$

$$\ln(h_t) = \alpha_0 + \alpha_1 \frac{(|a_{t-1}| + \gamma a_{t-1})}{\sigma_{t-1}} + \beta_1 \ln(h_{t-1})$$

com $\sigma_t = \sqrt{\exp(h_t)}$ e $e_t \sim N(0, 1)$.

Para simularmos algumas séries que seguem este processo, precisamos fixar os valores dos parâmetros. Então, atribuímos dois valores diferentes para cada um dos coeficientes do modelo EGARCH(1, 1) que são apresentados na Tabela 2.3.

Modelo	α_0	α_1	β_1	γ
1	0,1	0,8	0,3	-0,2
2	0,1	0,3	0,8	-0,2

Tabela 2.3: Parâmetros dos Modelos

Contudo, para estimar a variância de cada parâmetro dos modelos, utilizaremos a Informação observada de Fisher que neste caso será dado por

$$\begin{aligned} I(\alpha_0) &= \frac{1}{T} \sum_{t=2}^T \frac{1}{2h_t^2} \left(\frac{\partial h_t}{\partial \alpha_1} \right)^2, \\ I(\alpha_1) &= \frac{1}{T} \sum_{t=2}^T \frac{1}{2h_t^2} \left(\frac{\partial h_t}{\partial \alpha_0} \right)^2, \\ I(\beta_1) &= \frac{1}{T} \sum_{t=2}^T \frac{1}{2h_t^2} \left(\frac{\partial h_t}{\partial \beta_1} \right)^2, \\ I(\gamma) &= \frac{1}{T} \sum_{t=2}^T \frac{1}{2h_t^2} \left(\frac{\partial h_t}{\partial \gamma} \right)^2. \end{aligned}$$

Para o modelo EGARCH(1, 1) (forma do S-Plus) temos que os termos recursivos são dados por

$$\frac{\partial h_t}{\partial \alpha_0} = h_t \left[1 + \frac{1}{h_{t-1}} \left(\beta_1 - \frac{\alpha_1 |a_{t-1}| + \gamma a_{t-1}}{\sqrt{h_{t-1}}} \right) \frac{\partial h_{t-1}}{\partial \alpha_0} \right],$$

$$\frac{\partial h_t}{\partial \alpha_1} = h_t \left[\frac{|a_{t-1}| + \gamma a_{t-1}}{\sqrt{h_{t-1}}} + \frac{1}{h_{t-1}} \left(\beta_1 - \frac{\alpha_1 |a_{t-1}| + \gamma a_{t-1}}{\sqrt{h_{t-1}}} \right) \frac{\partial h_{t-1}}{\partial \alpha_1} \right],$$

$$\frac{\partial h_t}{\partial \gamma} = h_t \left[\frac{\alpha_1 a_{t-1}}{\sqrt{h_{t-1}}} + \frac{1}{h_{t-1}} \left(\beta_1 - \frac{\alpha_1 |a_{t-1}| + \gamma a_{t-1}}{\sqrt{h_{t-1}}} \right) \frac{\partial h_{t-1}}{\partial \gamma} \right],$$

$$\frac{\partial h_t}{\partial \beta_1} = h_t \left[\ln(h_{t-1}) + \frac{1}{h_{t-1}} \left(\beta_1 - \frac{\alpha_1 |a_{t-1}| + \gamma a_{t-1}}{\sqrt{h_{t-1}}} \right) \frac{\partial h_{t-1}}{\partial \beta_1} \right].$$

2.3.1 Distribuições dos Lags dos Modelos

As Figuras 2.17 e 2.18 mostram as distribuições dos lags para os modelos EGARCH(1, 1) definido com os parâmetros pela Tabela 2.3. Pode-se observar que a influência dos coeficientes não é muito longa no Modelo 1, visto que a partir do “lag” 6 ela é praticamente nula, já no Modelo 2 a influência é muito maior, sendo considerada nula somente a partir do “lag” 15.

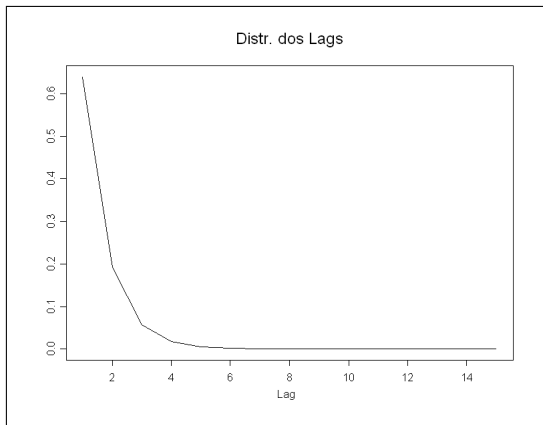


Figura 2.17: Lags Modelo 1

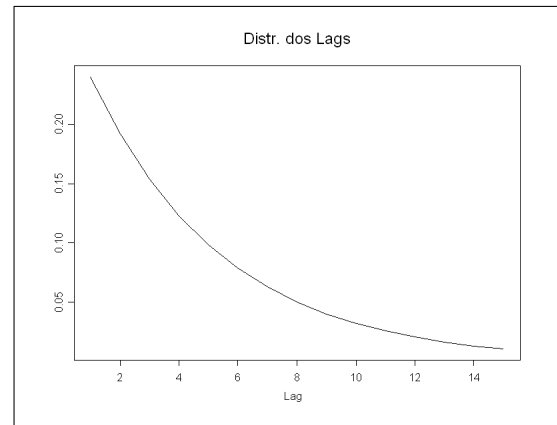


Figura 2.18: Lags Modelo 2

2.3.2 Convergência

A partir dos coeficientes especificados na Tabela 2.3, simularemos uma série de tamanho amostral $n = 20500$ com ponto de início $a_0 = 0$, para cada modelo. Fazendo um *burn in* de 500 observações, para tirar a influência do ponto inicial, iremos estimar o modelo EGARCH(1, 1) com os primeiros 500 dados (isto é, a_{501} até a_{1000}). Repetindo este processo, aumentando o tamanho da amostra de 100 em 100 até o fim da amostra total (isto é, de a_{501} a a_{1100} , de a_{501} a a_{1200} , ..., de a_{501} a a_{20500}), iremos então estimar os parâmetros do modelo. Diante disso teremos 200 séries com tamanhos variando de 500 até 20500, observando dessa forma se existe a convergência assintótica.

Modelo 1

As Figuras 2.19, 2.20, 2.21 e 2.22 mostram que não existe a convergência assintótica dos parâmetros, visto que suas estimativas não se estabilizam e também os intervalos de confiança não contém os “verdadeiros valores” dos coeficientes, em alguns momentos.

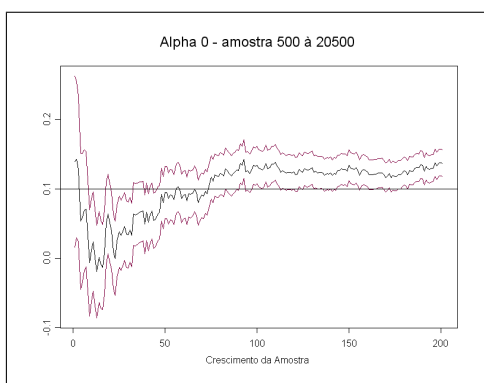


Figura 2.19: Alpha-0

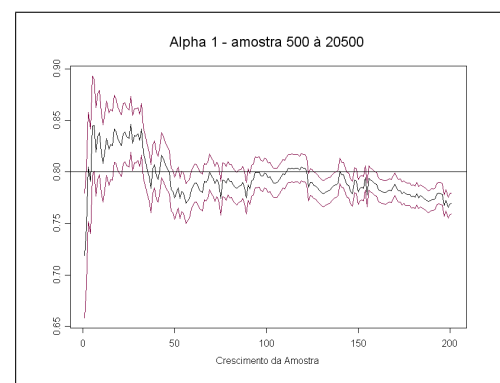


Figura 2.20: Alpha-1

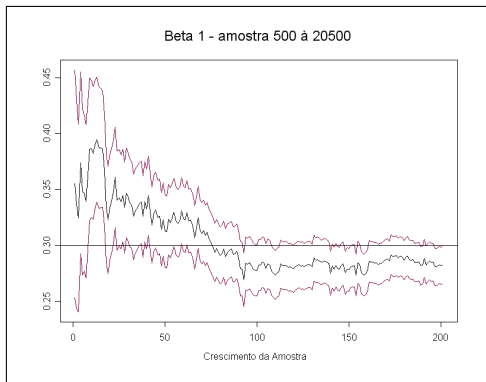


Figura 2.21: Beta-1

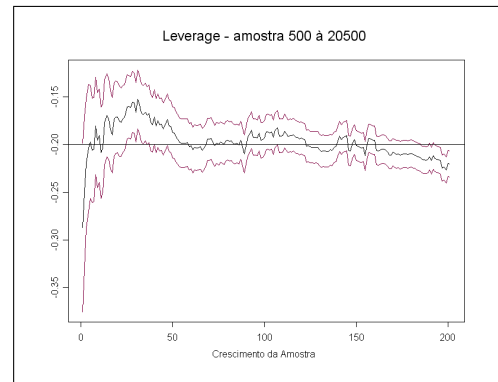


Figura 2.22: Leverage

Modelo 2

As Figuras 2.23, 2.24, 2.25 e 2.26 mostram a existência da convergência assintótica dos parâmetros α_0 , γ e β_1 (tendo neste um pequeno vício) onde, no centro estão o valor verdadeiro e a estimativa do parâmetro e nas periferias estão os limites inferior e superior. Porém para α_1 não conseguiu-se obter a convergência, visto que o valor estimado não se estabilizou e também seu intervalo de confiança não conteve o “verdadeiro valor” do coeficiente, em alguns momentos.

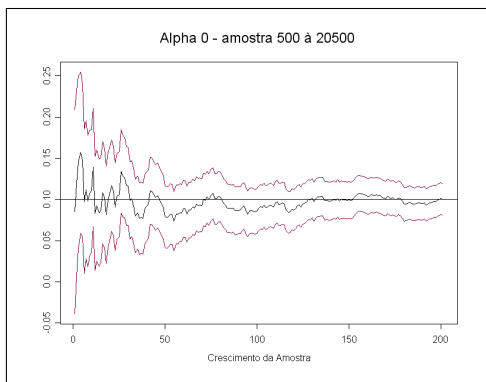


Figura 2.23: Alpha-0

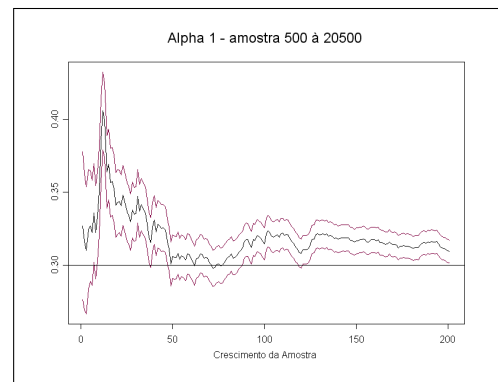


Figura 2.24: Alpha-1

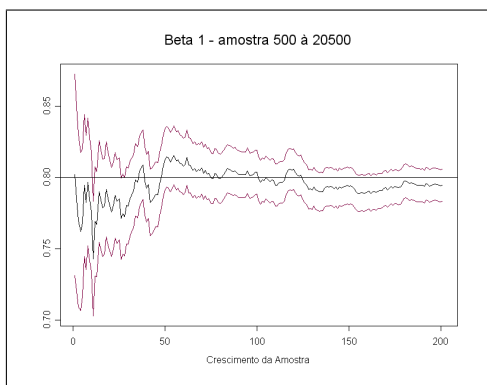


Figura 2.25: Beta-1

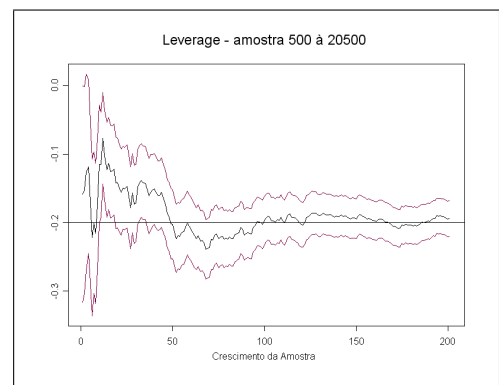


Figura 2.26: Leverage

Escolha do Melhor Modelo

Muitas vezes, após ajustarmos alguns modelos, podemos encontrar muitos que são adequados ao conjunto de dados em estudo. Diante disso, é necessário a utilização de alguma técnica de discriminação para a escolha do “melhor”.

No intuito de ajudar a procurar um modelo que melhor se ajuste aos dados e além disso tenha uma ótima previsão, iremos apresentar algumas técnicas utilizadas para estas escolhas.

3.1 Critério de Akaike

Em muitas aplicações podemos encontrar vários modelos que são adequados aos dados. Uma maneira de discriminar estes modelos é verificar qual deles possui melhor qualidade de ajuste. Para isto, pode-se utilizar a maior verossimilhança.

No entanto, espera-se que modelos com um maior número de coeficientes, tenham melhor qualidade de ajuste. Assim, é necessário que exista uma técnica de avaliação de modelos que pondere a qualidade do ajuste e o número de parâmetros utilizados.

Uma técnica muito conhecida para esta avaliação é o critério de informação de Akaike, definida como

$$AIC = -2 \log \text{verossimilhança maximizada} + 2m,$$

onde m é o número de coeficientes do modelo ajustado. Para dados normalmente distribuídos e utilizando as estimativas de máxima verossimilhança dos coeficientes, temos que

$$AIC = n \log(\hat{\sigma}_e^2) + 2m,$$

onde $\hat{\sigma}_e^2 = 1/n \sum_{i=1}^n \hat{e}_i^2$ e \hat{e}_t é o resíduo do modelo.

Existem muitos outros critérios de informação que são apenas variações do *AIC*. Um muito utilizado é o critério de informação bayesiano ou critério de Schwarz, definido por

$$BIC = -2 \log \text{verossimilhança maximizada} + 2m + m \log n,$$

onde n é o tamanho da amostra. Podemos observar que o *BIC* penaliza mais os modelos com maior número de parâmetros que o *AIC*.

3.2 Conteúdo de Informação

Após a estimação dos modelos de volatilidade desejamos avaliar, dentre os adequados, qual contém mais informação tanto para a previsão dentro da amostra quanto para a previsão do dia seguinte, isto é, previsão um passo à frente.

Desta forma, consideremos $\sigma_{t,+1}$ como a volatilidade realizada ao longo do próximo dia de transação da ação definida como

$$\sigma_{t,+1} = |R_{t+1}|$$

onde $|R_{t+1}|$ é o retorno do ativo-objeto entre t e $t + 1$ em módulo. Multiplicando esta volatilidade por $\sqrt{252}$, iremos obter a estimativa anual da volatilidade. Logo, como proposto por Portugal e Gabe (2002), para avaliar o conteúdo de informação da previsão um passo à frente, isto é, a habilidade das variáveis explicativas em prever a volatilidade entre t e $t + 1$, é realizada a seguinte regressão

$$\sigma_{t,+1} = \beta_0 + \beta_1 \hat{\sigma}_t + u_t \tag{3.1}$$

onde $\hat{\sigma}_t$ é a previsão da volatilidade um passo à frente do modelo escolhido e $u_t \sim IIN(0, 1)$. O R^2 (coeficiente de determinação) obtido na equação (3.1), pode ser utilizado para inferir o poder de explicação da variável independente sobre a volatilidade um passo à frente, assim como $\beta_1 \neq 0$ significa que o previsor tem alguma informação sobre a volatilidade do dia seguinte.

Esta regressão pode também ser usada com a finalidade de verificar o conteúdo de informação na previsão dentro da amostra, substituindo respectivamente os valores de $\hat{\sigma}_t$ e $\sigma_{t,+1}$ pelos valores do modelo ajustado e retorno, ambos no tempo t .

Como muitas vezes não se tem apenas um único modelo para a volatilidade, pode-se utilizar uma regressão múltipla para verificar qual dos modelos possui maior informação, analisando a eficiência relativa dos previsores, através da seguinte equação

$$\sigma_{t,+1} = \beta_0 + \beta_1 \hat{\sigma}_t^{V1} + \beta_2 \hat{\sigma}_t^{V2} + u_t \quad (3.2)$$

onde se a estimativa da volatilidade de $\hat{\sigma}_t^{V1}$ for relativamente mais eficiente que a de $\hat{\sigma}_t^{V2}$, o seu coeficiente β_1 deve ser estatisticamente significativo e $\beta_2 = 0$.

Novamente podemos utilizar esta regressão para verificar o conteúdo de informação na previsão dentro da amostra de forma similar ao mostrado para uma única variável explicativa.

3.3 Probabilidade de Acerto

Nesta seção, iremos apresentar a probabilidade de acerto do modelo. Esta técnica tem a finalidade, assim como o conteúdo de informação, de verificar se o modelo proposto consegue tanto se ajustar aos dados em estudo como fazer previsões a curto prazo.

Para definirmos a probabilidade de acerto, temos utilizar o seguinte coeficiente

$$s = \text{volatilidade estimada} / \text{volatilidade verdadeira}.$$

Assim, se o modelo estiver perfeitamente ajustado a amostra, esperamos que o valor de s seja igual a um. No entanto, sabemos que sempre existirá uma pequena variação do valor real para o ajustado, necessitando assim construir um intervalo onde este valor possa ser considerado bom.

Logo, a probabilidade de acerto será dada como a porcentagem dos s contidos no intervalo (0,95;1,05) que é puramente intuitivo. Os valores contidos acima deste de 1,05 são consideradas superestimções e os valores contidos abaixo de 0,95 são considerados subestimções.

Aplicação

Utilizando, o retorno da série dos preços de fechamento, das ações do ITAÚ de 02/01/1996 a 29/09/2006, iremos ajustar um modelo de volatilidade estatística que melhor descreve comportamento dos dados.

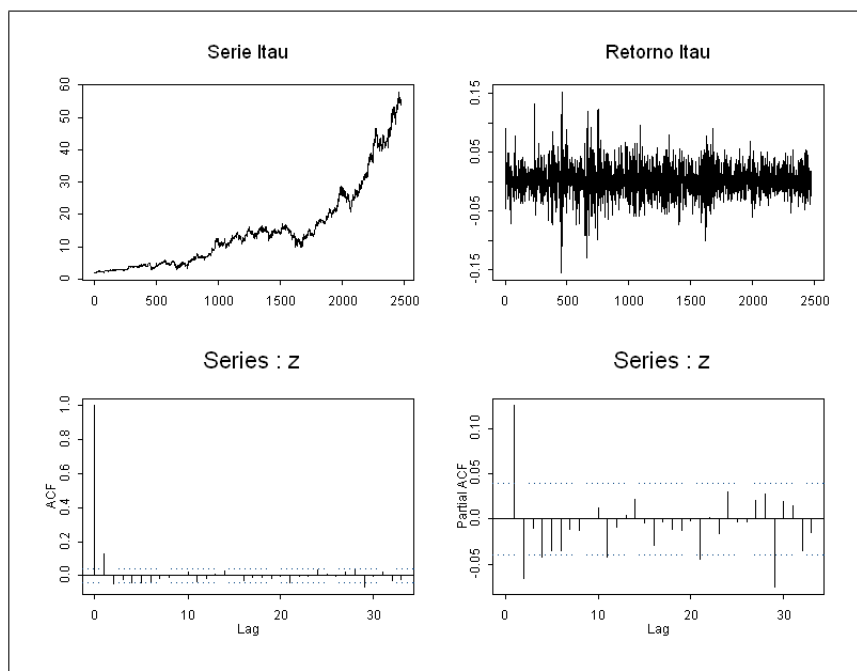


Figura 4.1: Série de Preço do Itaú

A Figura 4.1 mostra a série de fechamento das ações do ITAÚ, seus retornos, função de autocorrelação (f.a.c.) e função de autocorrelação parcial (f.a.c.p.). Como podemos

perceber pela f.a.c. e pela f.a.c.p., existe uma autocorrelação na série de retornos. Este fato é comprovado pelo teste de Ljung-Box ($p - \text{valor} < 0,05$), indicando que é necessário o ajuste de um modelo ARMA.

Os testes para a presença de raiz unitária e memória longa foram rejeitados, já o teste para estacionaridade da série foi aceito. Assim, como as suposições feitas ao modelo ARMA foram satisfeitas, selecionamos através do critério de informação (AIC) o modelo ARMA(1,7).

Pela Figura 4.2 podemos observar os resíduos do modelo ARMA(1,7), onde a f.a.c. dos resíduos indica que o modelo retirou toda a autocorrelação da série dos retornos. No entanto, a f.a.c. para o quadrado dos resíduos mostra a presença de autocorrelação na variância, sendo comprovada pelos testes de Ljung-Box e de efeito ARCH ($p - \text{valor} < 0,05$). Logo os modelos da família GARCH serão utilizados para remover esta autocorrelação presente na variância.

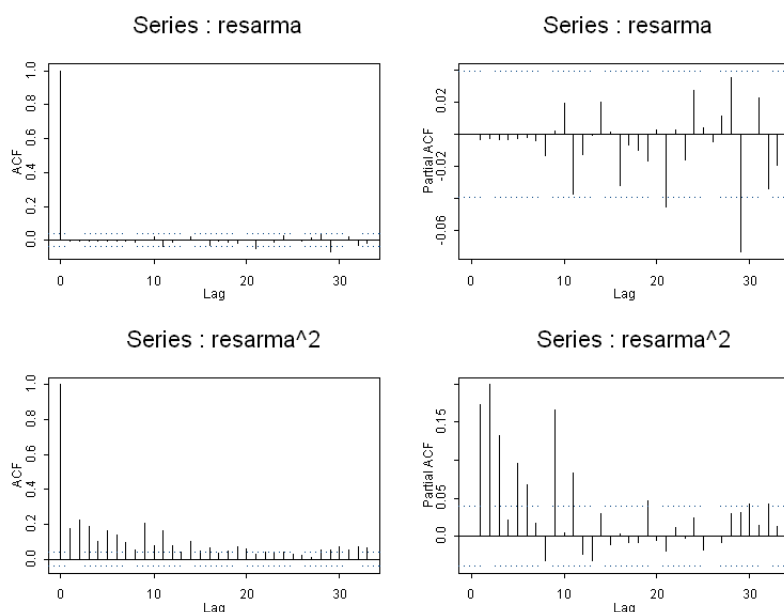


Figura 4.2: Resíduos ARMA(1,7)

Utilizando o critério de Akaike (menor AIC), além da convergência e adequação do modelo (teste de Ljung-Box para os resíduos padronizados), selecionou-se os modelos apresentados na Tabela 4.1. São apresentados também, na Tabela 4.1, a probabilidade de acerto e probabilidade de cobertura das previsões um passo à frente e o erro quadrático médio dos modelos ajustados.

Modelos	P. Acerto	P. Cobertura	E.Q.M	AIC	Meia-Vida
ARMA(1,7)-GARCH(1,1)	0,019	0,9198	0,0000009	-11481,64	11,7885
ARMA(1,7)-GARCH(3,2)	0,0206	0,9091	0,0000014	-11477,79	7,0211

Tabela 4.1: Modelos Selecionados

As Figuras 4.3 e 4.4 mostram o a.c.f. e o p.a.c.f. para os resíduos e os resíduos padronizados dos modelos ARMA(1,7)-GARCH(1,1) e ARMA(1,7)-GARCH(3,2) respectivamente. Podemos notar pelos gráficos que ambos os modelos retiraram a autocorrelação presente na variância, visto que não temos “lags” significativos. Este fato é comprovado ao aplicarmos o teste Ljung-Box para o quadrado dos resíduos padronizados, com nível de significância de 1%.

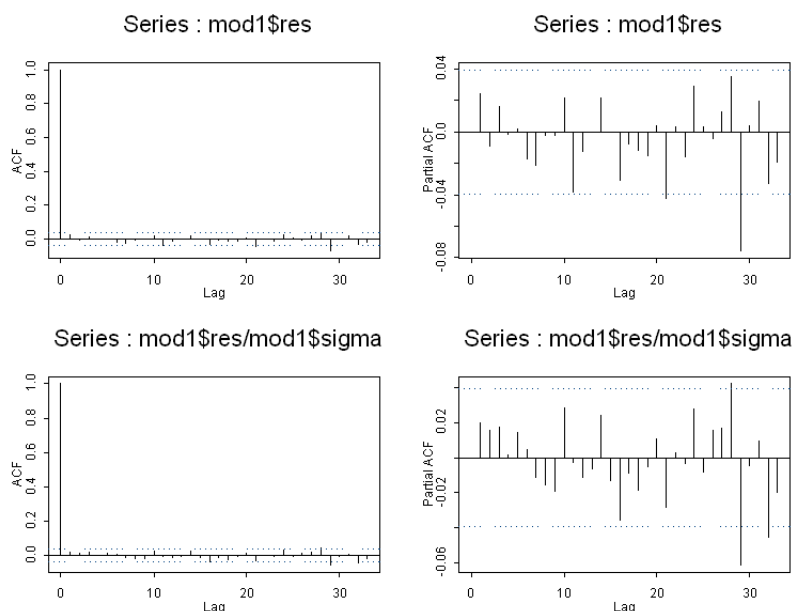


Figura 4.3: Resíduos Modelo 1

A Figura 4.5 mostra o intervalo de cobertura dos modelos, para dentro e fora da amostra.

Na Figura 4.6 podemos observar as regressões de conteúdo de informação, para dentro e fora da amostra, para os modelos. A Tabela 4.2 mostra os R^2 das regressões de conteúdo de informação.

Podemos notar que ambos os modelos apresentam a probabilidade de cobertura, probabilidade de acerto, E.Q.M. e AIC muito próximos. Porém, fazendo a regressão múltipla para os modelos, temos que o modelo ARMA(1,7)-GARCH(3,2) é significativo ao nível de significância de 1%.

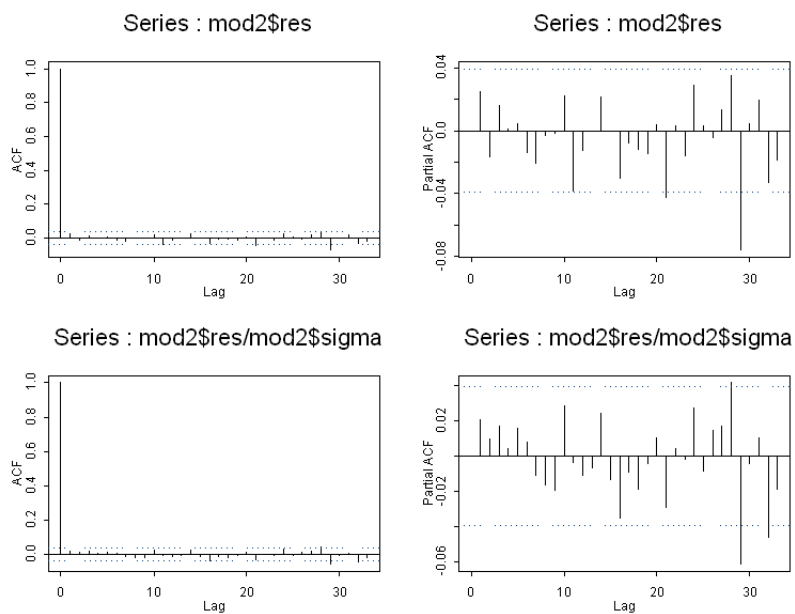


Figura 4.4: Resíduos Modelo 2

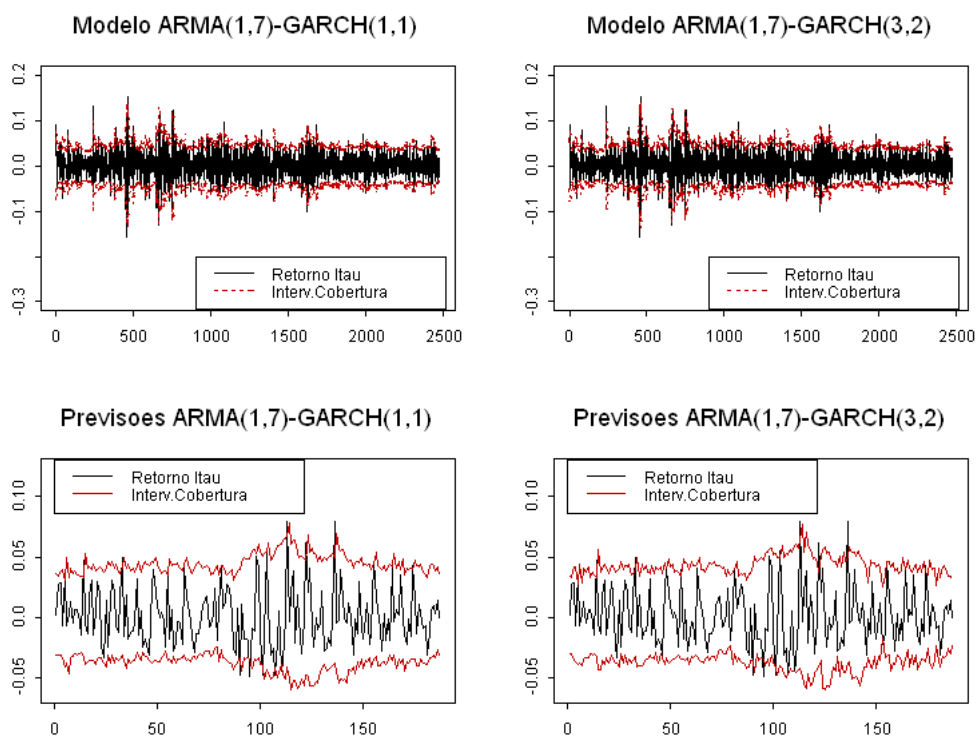


Figura 4.5: Intervalo de Cobertura

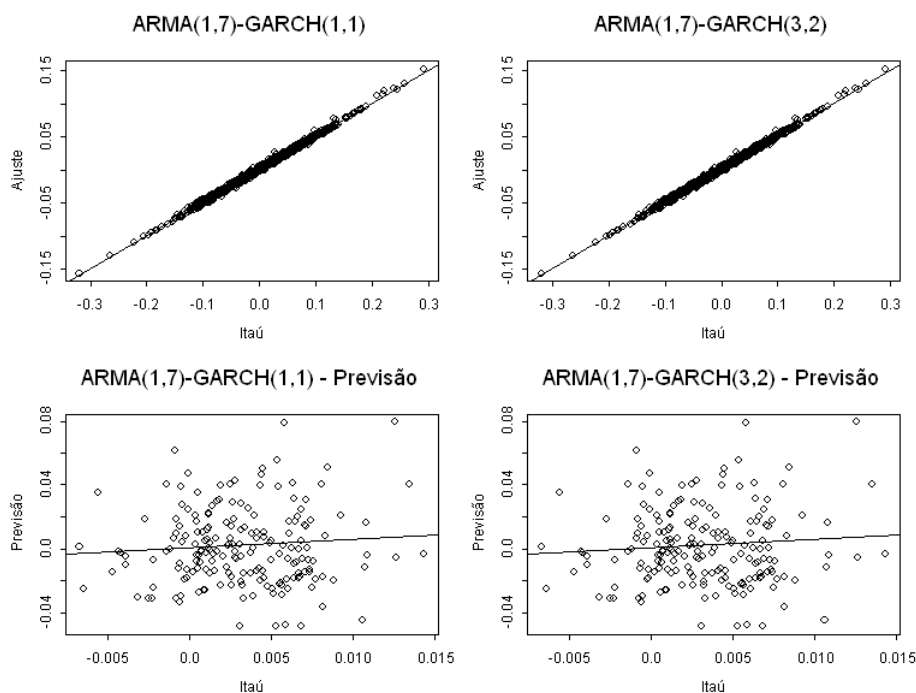


Figura 4.6: Regressões de Conteúdo de Informação

Modelos	Dentro	Previsões
ARMA(1,7)-GARCH(1,1)	0,99553	0,0014
ARMA(1,7)-GARCH(3,2)	0,99558	0,0004

Tabela 4.2: R^2 da Regressão Conteúdo de Informação

Portanto o “melhor” modelo é o ARMA(1,7)-GARCH(3,2). Contudo, os testes de memória longa e estacionaridade para o quadrado dos retornos são aceitos apenas ao nível de significância de 1% indicando que um modelo fracionário poderia ser ajustado. Também vale notar que um modelo que captasse o efeito de “leverage” dos retornos financeiros seria interessante para explicar a maior influência de retornos negativos na volatilidade.

Considerações Finais, Conclusões e Propostas Futuras

Nesta dissertação estudamos os modelos desenvolvidos inicialmente por Engle (1982), generalizados por Bollerslev (1986) e uma de suas extensões desenvolvida por Nelson (1991). Para cada modelo da família GARCH desenvolvemos algumas de suas propriedades, mostramos como identificar a presença da volatilidade, sua estimação pelo método de máxima verossimilhança (para o caso dos erros seguindo uma distribuição Normal ou t-Student), a verificação da adequação dos modelos aos dados e as suas previsões.

Também foi feito um estudo de simulação visando verificar as propriedades assintóticas dos parâmetros do modelos, como a convergência para o valor verdadeiro e a probabilidade de cobertura.

Infelizmente, estas propriedades não foram verificadas para a maioria dos modelos. Devido a este problema, foi estudado os algoritmos de estimação apresentados nos apêndices. Este estudo foi importante para conhecer como cada um destes algoritmos funcionam computacionalmente, além de nos oferecer como proposta futura a utilização de outros métodos e algoritmos de estimação para um estudo comparativo entre eles.

No Capítulo 3, foram apresentados alguns critérios para escolher qual seria o “melhor modelo” entre todos. Foi verificado na aplicação pouco poder de discriminação destes critérios. Assim, também podemos deixar como proposta futura o desenvolvimento de um critério que discrimine melhor os modelos.

A aplicação foi importante para verificarmos se os modelos são aplicáveis realmente. Contudo apenas ajustar o modelo não traz muita informação aos usuários do mercado financeiro. Assim, deixo também a possibilidade de, no futuro, estudar o valor em risco (VaR) que é uma medida muito utilizada no mercado financeiro.

Em linhas gerais, esta dissertação procura servir como material de referência para as pessoas que desejam utilizar estes modelos de heteroscedasticidade, podendo encontrar desde suas definições e propriedades até os algoritmos de estimação.

Apêndice

A.1 Apêndice 1

A.1.1 Teste de Box-Pierce-Ljung

Box e Pierce(1970) sugeriram um teste para as autocorrelações dos resíduos estimados, que apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muitos altos. Uma modificação deste teste foi proposta por Ljung e Box (1978), que é apresentada a seguir.

Se o modelo for apropriado, a estatística

$$Q = n(n+2) \sum_{i=1}^K \frac{\hat{r}_i^2}{(n-i)} \quad (\text{A.1})$$

terá uma distribuição χ_r^2 com $r = K - p - q$ graus de liberdade. A hipótese de ruído branco para os resíduos é rejeitada para valores grandes de Q . Em geral basta utilizar as 15 ou 20 primeiras \hat{r}_i .

A.1.2 Teste dos Multiplicadores de Lagrange

Teste de multiplicadores de Lagrange (ML); veja Engle(1982). Queremos testar $H_0 : \alpha_i = 0$, para todo $i = 1, \dots, r$, na regressão

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_r a_{t-r}^2 + u_t,$$

para $t = r + 1, \dots, T$. A estatística do teste é $S = TR^2$, que tem distribuição assintótica $\chi^2(r)$ sob H_0 . Aqui, R^2 é o quadrado do coeficiente de correlação múltipla da regressão acima. Um teste assintoticamente equivalente, que pode ter propriedades melhores para amostras pequenas, consiste em utilizar a estatística

$$F = \frac{(SQR_0 - SQR_1)/r}{SQR_1/(T - 2r - 1)} \sim F(r, T - 2r - 1), \quad (\text{A.2})$$

na qual $SQR_0 = \sum_{t=r+1}^T (a_t^2 - \bar{a})^2$ e $SQR_1 = \sum_{t=r+1}^T \hat{u}_t^2$, com \bar{a} denotando a média amostral dos a_t^2 e \hat{u}_t os resíduos de mínimos quadrados da regressão acima. Se o valor de F for significativo, dizemos que há heteroscedasticidade condicional na série.

A.2 Apêndice 2

Para a estimação dos parâmetros dos modelos da família "GARCH", via maximização da função de verossimilhança condicional, será necessário a utilização de um algoritmo iterativo. Antes da apresentação de dois destes algoritmos, iremos desenvolver alguns conceitos.

Supondo que o processo da família "GARCH" tenha distribuição condicional normal, temos que a t -ésima log-verossimilhança condicional será dada por

$$l_t(\theta) \propto -\frac{1}{2} \log h_t + \frac{1}{2} \frac{a_t^2}{h_t}, \quad (\text{A.3})$$

onde θ é o vetor dos parâmetros do modelo da família "GARCH".

Considerando o tamanho amostral T , temos que a média da log-verossimilhança condicional será dada por $l(\theta) = \sum_{t=1}^T l_t(\theta)/T$.

A primeira derivada da equação (A.3) em relação ao vetor paramétrico é dada por

$$\frac{\partial l_t(\theta)}{\partial \theta} = \frac{1}{2h_t} \frac{\partial h_t}{\partial \theta} \left(\frac{a_t^2}{h_t} - 1 \right), \quad (\text{A.4})$$

onde $\partial h_t / \partial \theta$ é o vetor $W' = (1, a_{t-1}^2, a_{t-2}^2, \dots, a_{t-r}^2)$ no caso do modelo ARCH(r) ou será o termo recursivo no caso do modelo GARCH(r, s) dado por

$$\frac{\partial h_t}{\partial \theta} = W + \sum_{i=1}^s \beta_i \frac{\partial h_{t-i}}{\partial \theta}.$$

A segunda derivada será dada por

$$\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'} = -\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta'} \left(\frac{a_t^2}{h_t} \right) + \frac{\partial}{\partial \theta'} \left[\frac{1}{2h_t} \frac{\partial h_t}{\partial \theta} \right] \left(\frac{a_t^2}{h_t} - 1 \right). \quad (\text{A.5})$$

Logo podemos calcular a matriz de informação de Fisher, que é obtida através da esperança da segunda derivada negativa da log-verossimilhança condicional, ou seja, no caso da log-verossimilhança média condicional temos

$$\widehat{I}(\theta) = E \left[-\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'} \right]. \quad (\text{A.6})$$

Calculando a esperança dos últimos fatores da primeira e segunda parte da expressão (A.5) temos que

$$\begin{aligned} E \left[\frac{a_t^2}{h_t} \right] &= E \left[E \left[\frac{a_t^2}{h_t} / F_{t-1} \right] \right] = E \left[\frac{1}{h_t} E \left[a_t^2 / F_{t-1} \right] \right] = 1; \\ E \left[\frac{a_t^2}{h_t} - 1 \right] &= E \left[E \left[\frac{a_t^2}{h_t} - 1 / F_{t-1} \right] \right] = E \left[\frac{1}{h_t} E \left[a_t^2 / F_{t-1} \right] - 1 \right] = 0. \end{aligned}$$

Logo a matriz de informação de Fisher é consistentemente estimada pela expressão

$$\widehat{I}(\theta) = \frac{1}{2T} \sum_{t=1}^T \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta'} \right]. \quad (\text{A.7})$$

A.2.1 Algoritmo Scoring

Para a estimação dos parâmetros do modelo ARCH(r), via maximização da função de verossimilhança condicional, será necessário a utilização de um algoritmo iterativo. Em Engle (1982) foi utilizado o algoritmo scoring. As iterações deste algoritmo são calculadas baseando-se na seguinte expressão

$$\theta^{i+1} = \theta^i + [I^i(\theta)]^{-1} \frac{1}{T} \sum_t \frac{\partial l_t^i(\theta)}{\partial \theta} \quad (\text{A.8})$$

onde $\theta' = (\alpha_0, \alpha_1, \dots, \alpha_r)$ é o vetor dos parâmetros, T é o tamanho amostral, $l_t(\theta)$ é a t -ésima função de verossimilhança condicional e $I(\theta)$ é a matriz de informação de fisher.

Para o modelo ARCH(r) temos que a fórmula da volatilidade é dada por

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_r a_{t-r}^2,$$

podemos escrever esta função da seguinte forma

$$h_t = z_t' \alpha, \quad (\text{A.9})$$

onde $z_t' = (1, a_{t-1}^2, \dots, a_{t-p}^2)$ e $\alpha' = (\alpha_0, \alpha_1, \dots, \alpha_p)$.

Utilizando a expressão (A.9) na equações (A.4) e (A.7), temos as seguintes expressões

$$\begin{aligned} \frac{\partial l_t(\alpha)}{\partial \alpha} &= \frac{1}{2h_t} z_t \left(\frac{a_t^2}{h_t} - 1 \right), \\ \hat{I}(\alpha) &= \frac{1}{2T} \sum_{t=1}^T \left[\frac{z_t' z_t}{h_t^2} \right]. \end{aligned}$$

Assim podemos reescrever o algoritmo scoring (A.8) da seguinte maneira

$$\alpha^{i+1} = \alpha^i + [\tilde{z}' \tilde{z}]^{-1} \tilde{z}' f^i,$$

onde

$$\begin{aligned} \tilde{z}_t &= (1, a_{t-1}^2, \dots, a_{t-p}^2) / h_t^i, \\ \tilde{z}' &= (\tilde{z}'_1, \dots, \tilde{z}'_T), \\ f_t^i &= (a_t^2 - h_t^i) / h_t^i, \\ f^{i'} &= (f_1^i, \dots, f_T^i). \end{aligned}$$

Sob as condições do teorema de Crowder's (1976) para martinagales, os estimadores de máxima verossimilhança são assintoticamente normalmente distribuídos, com distribuição limite dada por

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

A.2.2 Algoritmo BHHH

Para a estimação dos parâmetros do modelo GARCH, via maximização da função de verossimilhança condicional, será necessário a utilização de um algoritmo iterativo. Em Bollerslev (1986) foi utilizado o algoritmo iterativo BHHH, desenvolvido por Berndt, Hall, Hall e Hausman (1974). As iterações deste algoritmo são calculadas baseando-se na seguinte expressão

$$\theta^{i+1} = \theta^i + \lambda^i \left(\sum_t \frac{\partial^2 l_t^i(\theta)}{\partial \theta \partial \theta'} \right)^{-1} \left(\sum_t \frac{\partial l_t^i(\theta)}{\partial \theta} \right) \quad (\text{A.10})$$

onde $l_t(\theta)$ é a t -ésima log-verossimilhança condicional, θ é o vetor paramétrico e λ^i é uma variável que determina o tamanho do salto na i -ésima iteração e esta variável é escolhida de forma a maximizar a função de verossimilhança.

Para a escolher a variável λ^i , Berndt (1974) introduziu o critério descrito a seguir.

Seja $V(\theta)$ uma função contínua e diferenciável. Temos por objetivo encontrar θ que maxime $V(\theta)$.

Considere o gradiente de $V(\theta)$ com relação a θ , $g = [\partial V(\theta)]/(\partial \theta)$, e a matriz Hessiana Q obtida da inversa da matriz de segundas derivadas de $V(\theta)$.

Considere também a função $\gamma(\theta, \lambda)$,

$$\gamma(\theta, \lambda) = \frac{V(\theta + \lambda d) - V(\theta)}{\lambda d' g}, \quad (\text{A.11})$$

onde $d = Q \cdot g$.

Tomando uma constante δ dentro do intervalo $(0, 1/2)$, utilizaremos o seguinte critério para a escolha de λ .

Se $\gamma(\theta, 1) \geq \delta$, faça $\lambda = 1$. Senão, escolha um valor de λ que satisfaça a condição

$$\delta \leq \gamma(\theta, \lambda) \leq 1 - \delta. \quad (\text{A.12})$$

Sendo $V(\theta)$ uma função contínua e diferenciável, existe um valor de λ que satisfaz o critério (A.12).

Para demonstrar este fato, calculemos o seguinte limite

$$\begin{aligned} \lim_{\lambda \rightarrow 0} \gamma(\theta, \lambda) &= \lim_{\lambda \rightarrow 0} \frac{V(\theta + \lambda d) - V(\theta)}{\lambda d' g} \\ &= \frac{1}{d' g} \lim_{\lambda \rightarrow 0} \frac{V(\theta + \lambda d) - V(\theta)}{\lambda} \\ &= \frac{1}{d' g} d' \frac{\partial V(\theta)}{\partial \theta} = \frac{1}{d' g} d' g = 1. \end{aligned}$$

Assim, pelo fato da função $V(\theta)$ ser contínua, $\gamma(\theta, 1) < \delta$, $\lim_{\lambda \rightarrow 0} \gamma(\theta, \lambda) = 1$ e pelo teorema do valor intermediário, temos que existe um λ que satisfaça o critério (A.12).

Obedecido estas condições, verificamos a convergência no processo de maximização da função $V(\theta)$. O teorema de convergência para esse caso é provado com detalhes em Goldstein (1967). Apresentamos a seguir alguns aspectos do mesmo.

Assumindo que $V(\theta)$, é duplamente diferenciável em um conjunto compacto com contorno definido e a seqüência $\theta^1, \theta^2, \dots$, onde

$$\theta^{i+1} = \theta^i + \lambda^i d^i, \quad (\text{A.13})$$

$$d^i = Q^i g^i, \quad (\text{A.14})$$

e Q^i satisfazendo $(d'g)/(d'd) > \alpha$, sendo α um valor constante positivo menos que um e λ^i satisfazendo (A.12), então $\lim_{i \rightarrow \infty} g^i = 0$.

O processo de estimação para os outros modelos da família “GARCH” segue o mesmo procedimento apresentado com algumas pequenas variações, tendo em vista que são simplificações ou complementações do modelo GARCH.

Referências Bibliográficas

- [1] Baillie, R. T. [1996] *Long memory processes and fractional integration in econometrics*, Journal of Econometrics, 73, 5-59.
- [2] Berndt, E. K.; Hall, B.; Hall, R. and Hausman, J. A. [1974] *Estimation and Inference in Nonlinear Structural Models*, Annals of Economic and Social Measurement, 3, 653-665.
- [3] Bollerslev, T. [1986] *Generalize Autoregressive Conditional Heteroskedasticity*, Journal of Econometrics, 31, 307-327.
- [4] Bollerslev, T. and H. O. Mikkelsen [1996] *Modeling and pricing long-memory in stock market volatility*, Journal of Econometrics.
- [5] Box, G. E. P. and Pierce, D. A. [1970] *Distribution of the Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models*, Journal of the American Statistical Association, 65, 1509-1526.
- [6] Crowder, M. J. [1976] *Maximum Likelihood Estimation for Dependent Observations*, Journal of the Royal Statistical Society, Series B, 38, 45-53.
- [7] Engle, R. F. [1982] *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation*, Econometrica, 50, 987-1008.
- [8] Goldstein, A. [1967] *Constructive Real Analysis*, Harper & Row, New York.
- [9] Ljung, G. M. and Box, G. E. P. [1978] *On a Measure of Lack of Fit in Time Series Models*, Biometrika, 65, 297-303.
- [10] Morais, I. A. C. and Portugal, M. S. [1999] *Modelagem e Previsão de Volatilidade Determinística e Estocástica para a Série do IBOVESPA*, Est. Econ., 3, 303-341.

- [11] Morettin, P. A. [2006] *Econometria Financeira. Um Curso em Séries Temporais Financeiras*, Departamento de Estatística do Instituto de Matemática e Estatística da Universidade de São Paulo.
- [12] Nelson, D. B. [1991] *Conditional heteroskedasticity in asset returns: A new approach*, *Econometrica*, 59, 347-370.
- [13] Nelson, D. B. and Cao, C.Q. [1992] *Inequality Constraints in the Univariate GARCH model*, *Journal of Business and Economic Statistics*, 10, 229-235.
- [14] Portugal, M. S. and Gabe, J. [2002] *Volatilidade implícita versus volatilidade estatística: Uma avaliação para o mercado brasileiro a partir dos dados de opções e ações da TELEMAR S.A.*, *Econometrica*, 59, 347-370.