

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO  
DEPARTAMENTO DE FÍSICA E MATEMÁTICA

Programa de Pós-graduação em Biometria

**MODELAGEM E PREVISÃO DA  
VOLATILIDADE DOS RETORNOS DO CAFÉ  
ARÁBICA PRODUZIDO NO BRASIL**

Marcela Verônica Alves de Souza

DISSERTAÇÃO DE MESTRADO

RECIFE-PE  
28 de fevereiro de 2005

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO  
DEPARTAMENTO DE FÍSICA E MATEMÁTICA

Marcela Verônica Alves de Souza

**MODELAGEM E PREVISÃO DA VOLATILIDADE DOS RETORNOS  
DO CAFÉ ARÁBICA PRODUZIDO NO BRASIL**

*Trabalho apresentado ao Programa de Pós-graduação em Biometria do DEPARTAMENTO DE FÍSICA E MATEMÁTICA da UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO como requisito parcial para obtenção do grau de Mestre em Mestrado em Biometria.*

Orientador: *Prof. Dr. Gauss Moutinho Cordeiro*

RECIFE-PE  
28 de fevereiro de 2005

*Aos meus irmãos Ricardo, Robertson e Márcia*

## AGRADECIMENTOS

A Deus, pela saúde, paz e discernimento para fazer minhas escolhas e por concluir mais uma etapa da minha vida.

A meus irmãos Ricardo, Robertson e Márcia pelo incentivo e apoio que sempre me deram.

Ao meu orientador, Gauss Moutinho Cordeiro, pela paciência, incentivo e confiança que sempre depositou em mim.

A minha sogra Fleunice e cunhada/irmã Hildenice pela paciência e carinho.

Ao meu marido Hildeberto, presente em todas as horas, pela ajuda, confiança e paciência.

A Oscar Neto, pela eficiência e organização com que trabalha, pela solidariedade e principalmente pela paciência.

Aos professores do Programa de Mestrado em Estatística da UFRPE pela atenção e disponibilidade.

A Eufrásio Lima Neto, pela simpática colaboração e principalmente pela disponibilidade.

A Marco Santos, pela eficiência, organização e presteza.

Aos colegas da minha turma de mestrado, em especial Cleto, Luiz de França, Paulo, Illes, Hebertz, Dirac e Márcio pelas horas divertidas e enriquecedoras de estudo.

*“Viva como se fosse morrer  
amanhã, aprenda como se fosse  
viver para sempre”.*

—MAHATMA GHANDI

## RESUMO

O principal objetivo da presente dissertação é apresentar uma metodologia de modelagem e previsão da volatilidade do café Arábica, com o intuito não só de explicar e prever sua heteroscedasticidade condicional, como de obter o risco de mercado para seus investimentos. Ao longo desta dissertação trabalhamos com a série temporal dos retornos diários do café Arábica de setembro de 1996 a dezembro de 2004 para modelagem e previsão da volatilidade. Analisamos os modelos ARIMA, assim como utilizamos os modelos não-lineares ARCH e sua generalização, GARCH. Contudo, os resultados prevêm alta volatilidade e perda nos investimentos para até nove passos à frente. Estes resultados também revelam o modelo GARCH ajustado aos resíduos de um modelo AR, dentre os modelos considerados, como o modelo de maior poder preditivo, um passo à frente.

**Palavras-chave:** Previsão, Retornos, Volatilidade, Risco, Café.

## ABSTRACT

The main aim of the present dissertation is to present a modeling methodology and forecast of the volatility of the Arabic coffee, with the intent not only to predict and explain its conditional heteroskedasticity, but also to obtain the market risk for its investments. For modeling and prevision of volatility in this dissertation, the temporal series of Arabic coffee returns from September 1996 to December 2004, was used. The ARIMA models were analyzed, together with non-linear ARCH models, and their generalization GARCH. The results predict high volatility and investment loss for up to nine steps ahead. The results also indicate that the GARCH model, adjusted to the residuals of the AR model, presents the largest forecast power one step ahead.

**Keywords:** Forecast, Returns, Volatility, Risk, Coffee.

# SUMÁRIO

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Revisão Bibliográfica</b>	<b>3</b>
2.1	Estacionariedade, Tendência e Sazonalidade . . . . .	3
2.2	Funções de Autocovariância e Autocorrelação . . . . .	5
2.3	Modelo de Alisamento Exponencial . . . . .	9
2.3.1	Modelo de Alisamento Exponencial Simples . . . . .	9
2.3.2	Modelo de Alisamento Exponencial de Holt . . . . .	10
2.3.3	Alisamento Exponencial de Holt-Winters . . . . .	11
2.4	Modelos Auto-Regressivos AR(p) . . . . .	12
2.5	Modelos de Médias Móveis MA(q) . . . . .	13
2.6	Modelos ARMA, ARIMA, SARIMA, ARMAX . . . . .	14
2.7	Modelagem de Box-Jenkins . . . . .	16
2.8	Identificação, Estimação e Checagem . . . . .	16
<b>3</b>	<b>Metodologia</b>	<b>18</b>
3.1	Modelos ARIMA . . . . .	18
3.1.1	Identificação . . . . .	18
3.1.2	Estimação . . . . .	22
3.1.3	Diagnóstico . . . . .	26
3.1.4	Previsão . . . . .	29
3.2	Considerações Gerais Sobre Econometria Financeira . . . . .	31
3.2.1	Retornos . . . . .	31
3.2.2	Distribuição de Retornos . . . . .	32
3.2.3	Assimetria e Curtose . . . . .	33



3.2.4	Caracterização dos Retornos . . . . .	34
3.2.5	Volatilidade . . . . .	34
3.2.6	Modelos Auto-Regressivos com Heteroscedasticidade Condicional (ARCH) . . . . .	36
3.2.7	Modelos GARCH . . . . .	41
3.2.8	Valor em Risco . . . . .	43
3.2.9	VaR Usando Modelos ARIMA e GARCH . . . . .	45
<b>4</b>	<b>Resultados e Discussão</b>	<b>46</b>
4.1	Índices diários . . . . .	46
4.2	Análise Exploratória . . . . .	47
4.3	Modelagem e Previsão . . . . .	48
<b>5</b>	<b>Conclusões</b>	<b>67</b>
<b>6</b>	<b>Referências</b>	<b>68</b>

## LISTA DE FIGURAS

<b>Figura 2.1:</b> Comportamento explosivo. . . . .	4
<b>Figura 2.2:</b> Correlograma de uma série simulada. . . . .	7
<b>Figura 2.3:</b> Correlograma do índice de custo de vida no município de São Paulo; observações mensais de janeiro de 1970 a junho de 1980. . . . .	7
<b>Figura 4.1:</b> Retornos diários da série <i>cafeA</i> . . . . .	47
<b>Figura 4.2:</b> Histograma dos retornos da série <i>cafeA</i> . . . . .	48
<b>Figura 4.3:</b> Gráfico das <i>fac</i> e <i>facp</i> dos retornos diários da série <i>cafeA</i> . . . . .	49
<b>Figura 4.4:</b> Gráfico das <i>fac</i> e <i>facp</i> dos quadrados dos resíduos do modelo (4.1). . . . .	54
<b>Figura 4.5:</b> Estimativa fornecida pelo modelo (4.3) para o desvio padrão condicional dos retornos diários da série <i>cafeA</i> . . . . .	59
<b>Figura 4.6:</b> Estimativa fornecida pelo modelo (4.5) para o desvio padrão condicional dos retornos diários da série <i>cafeA</i> . . . . .	63
<b>Figura 4.7:</b> Representação gráfica do intervalo de confiança para as previsões do modelo $AR(2)-ARCH(5)$ e das previsões da volatilidade. . . . .	65
<b>Figura 4.8:</b> Representação gráfica do intervalo de confiança para as previsões do modelo $AR(2)-GARCH(3,3)$ e das previsões da volatilidade. . . . .	65

## LISTA DE TABELAS

<b>Tabela 3.1:</b> Comportamento de $\text{fac}$ e $\text{facp}$ segundo a ordem do modelo. . . . .	20
<b>Tabela 4.1:</b> $\text{Fac}$ e $\text{facp}$ dos retornos da série $\text{cafeA}$ . . . . .	49
<b>Tabela 4.2:</b> Valores do $AIC$ e $BIC$ de modelos $AR(j)$ , $j = 1, \dots, 10$ , ajustados aos retornos diários da série $\text{cafeA}$ . . . . .	50
<b>Tabela 4.3:</b> Ajustamento de um modelo $AR(7)$ aos retornos diários da série $\text{cafeA}$ . . . . .	50
<b>Tabela 4.4:</b> Ajustamento do modelo (4.1) aos retornos diários da série $\text{cafeA}$ . . . . .	51
<b>Tabela 4.5:</b> $\text{Fac}$ e $\text{facp}$ dos resíduos do modelo (4.1). . . . .	52
<b>Tabela 4.6:</b> $\text{Fac}$ e $\text{facp}$ dos quadrados dos resíduos do modelo (4.1). . . . .	53
<b>Tabela 4.7:</b> Valores do $AIC$ e $BIC$ de modelos $ARCH(j)$ , $j = 1, \dots, 6$ , ajustados aos resíduos do modelo (4.1). . . . .	54
<b>Tabela 4.8:</b> Ajustamento do modelo (4.2) aos retornos diários da série $\text{cafeA}$ . . . . .	55
<b>Tabela 4.9:</b> Ajustamento do modelo $AR(2)$ - $ARCH(5)$ aos retornos diários da série $\text{cafeA}$ . . . . .	55
<b>Tabela 4.10:</b> $\text{Fac}$ e $\text{facp}$ dos resíduos padronizados do modelo (4.3). . . . .	56
<b>Tabela 4.11:</b> $\text{Fac}$ e $\text{facp}$ dos quadrados dos resíduos padronizados do modelo (4.3). . . . .	57
<b>Tabela 4.12:</b> Aplicação do teste ML à seqüência dos quadrados dos resíduos padronizados do modelo (4.3). . . . .	58
<b>Tabela 4.13:</b> Valores do $AIC$ e $BIC$ de modelos $GARCH(j, k)$ , $j, k = 1, \dots, 3$ , ajustados aos resíduos do modelo $AR(2)$ . . . . .	60
<b>Tabela 4.14:</b> Ajustamento do modelo $AR(2)$ - $GARCH(3, 3)$ aos retornos diários da série $\text{cafeA}$ . . . . .	60
<b>Tabela 4.15:</b> $\text{Fac}$ e $\text{facp}$ dos resíduos padronizados do modelo (4.5). . . . .	61
<b>Tabela 4.16:</b> $\text{Fac}$ e $\text{facp}$ dos quadrados dos resíduos padronizados do modelo (4.5). . . . .	62
<b>Tabela 4.17:</b> Aplicação do teste ML à seqüência dos quadrados dos resíduos padronizados do modelo (4.5). . . . .	63
<b>Tabela 4.18:</b> Valores do $AIC$ e $BIC$ dos modelos (4.3) e (4.5), ajustados aos retornos diários da série $\text{cafeA}$ . . . . .	64
<b>Tabela 4.19:</b> Medidas de precisão para previsões 1, 3, 6 e 9 passos à frente, para os retornos diários da série $\text{cafeA}$ . . . . .	64
<b>Tabela 4.20:</b> Valores do $VaR[h]$ , $h = 1, \dots, 9$ dos modelos (4.3) e (4.5), ajustados aos retornos diários da série $\text{cafeA}$ . . . . .	66

# 1 INTRODUÇÃO

Diversos fenômenos podem ser explicados através de dados organizados seqüencialmente no tempo, tais dados são chamados de séries temporais. A intenção de modelar séries temporais vem junto com a necessidade de explicar o comportamento destas séries, descrever e até prever determinado fenômeno.

A abordagem Box-Jenkins vem sendo comumente utilizada para modelagem de séries temporais. Trata-se de um método que consiste em ajustar modelos ARIMA e suas variações. A família de modelos ARIMA é usada para modelar séries que admitem inovações comportando-se conforme um processo ruído branco, com média zero e variância constante. Entretanto, as séries financeiras apresentam heteroscedasticidade condicional (volatilidade), ou seja, modelos não-lineares no que se refere à variância.

Existe uma grande variedade de modelos não-lineares disponíveis na literatura, mas iremos nos concentrar na classe de modelos ARCH, introduzida por Engle (1982). A idéia básica para esses modelos é de que os dados são não correlacionados serialmente, mas a variância condicional evolui no tempo dependendo de dados passados por meio de uma função quadrática. Este tipo de comportamento é observado freqüentemente em dados financeiros.

Um dos objetivos em finanças é a avaliação dos riscos de uma carteira de ativos ou risco de ativos financeiros. Neste trabalho o principal objetivo será calcular a medida de um tipo particular de risco, o chamado risco de mercado. O valor em risco é freqüentemente medido em termos de variações de preços dos ativos. Comumente esta variação tem comportamento condicional, ou seja, a variação de hoje depende da variação de ontem e assim sucessivamente, melhor dizendo, possuem volatilidade.

A variação relativa de preços é chamada de retorno do ativo nos instantes definidos. Quando o interesse é modelar a volatilidade, é preferível trabalhar com retornos, que são livres de escalas, do que com preços, pois os primeiros têm propriedades estatísticas mais interessantes, como estacionariedade e ergodicidade.

Devido à importante influência financeira do mercado do café no Brasil, estamos interessados, então, em modelar seus retornos e determinar o valor em risco de seus investimentos. Sua eco-

nomia é caracterizada por ciclos de prosperidade e declínio em função de variações no consumo. Esta variação é considerada como um fator determinante ou principal elemento para explicar o comportamento da economia do café. Daí a importância de modelar e prever esta variação para os investidores e para o setor produtivo. Enfim, o objetivo da presente dissertação é apresentar uma metodologia de modelagem e previsão da volatilidade do café Arábica, com o intuito não só de explicar e prever sua heteroscedasticidade condicional, como de obter o risco de mercado para seus investimentos.

A Seção 2, revisando a literatura, apresenta conceitos de séries temporais como estacionariedade, função de autocovariância, função de autocorrelação e de autocorrelação parcial, ruído branco, métodos de modelagem e previsão. Nas Seções 3 e 4 é feita uma exposição dos materiais e métodos utilizados, a saber: método de Box e Jenkins para modelos ARIMA e ferramentas da econometria financeira. Na Seção 5 é feita uma análise exploratória dos dados diários da série de preços do café Arábica que compreende o período de setembro de 1996 a dezembro de 2004 e efetuada a modelagem da série. Para isso, usamos a modelagem de Box-Jenkins e os modelos ARCH-GARCH. Por fim, na Seção 6 estão as conclusões obtidas.

Os pacotes computacionais estatísticos utilizados foram o R e o Eviews. O programa R é uma versão gratuita do programa S-PLUS, e pode ser obtido da página <http://www.r-project.org>.

## 2 REVISÃO BIBLIOGRÁFICA

### SÉRIES TEMPORAIS: CONCEITOS IMPORTANTES E MODELOS

#### 2.1 Estacionariedade, Tendência e Sazonalidade

Uma das suposições a respeito de uma série temporal é a de que ela se desenvolve aleatoriamente no tempo ao redor de uma média constante, demonstrando alguma forma de equilíbrio ou estabilidade. Em outras palavras, uma série temporal estacionária (ou fracamente estacionária) ocorre se suas variações cíclicas foram removidas, se com a evolução do tempo sua média se mantém constante (sem tendência), e se o mesmo acontece com a variância.

A maior parte da teoria de séries temporais trata de séries estacionárias e, por essa razão, a análise é feita transformando-se séries não estacionárias em estacionárias, Cordeiro, D.M. (2002).

Considere uma série temporal  $\{y_t, t \in T\}$ , onde a média de  $y_t$  é denotada por

$$\mu_t = E(y_t),$$

a covariância entre  $y_t$  e  $y_{t+h}$  é

$$\gamma(t, t+h) = Cov(y_t, y_{t+h}) = E[(y_t - \mu_t)(y_{t+h} - \mu_{t+h})],$$

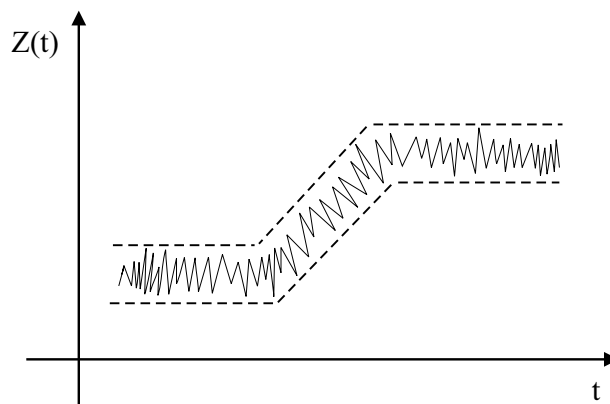
então, sua variância será dada por  $\gamma(t, t)$ . Esta série temporal é formalmente dita ser fracamente estacionária se:

- i)  $E(y_t) = \mu, \forall t \in T$  (a média de  $y_t$  é constante);
- ii)  $Var(y_t) = \sigma^2, \forall t \in T$  (a variância de  $y_t$  é constante);
- iii)  $Cov(y_t, y_{t+h}) = \gamma(h), \forall t, t+h \in T$  (a covariância entre  $y_t$  e  $y_{t+h}$  não depende de  $t$  e sim apenas de  $h$ ).

Se a função de distribuição de probabilidade de  $y_t$  é a mesma para todos os pontos do conjunto índice, então a série é dita ser fortemente ou estritamente estacionária. Uma série temporal

fortemente estacionária que possui seus dois primeiros momentos finitos também é dita fracamente estacionária. A recíproca desta afirmação não pode ser dita como verdadeira, pois só estão garantidos os dois primeiros momentos.

Geralmente as séries econômicas e financeiras apresentam a forma mais comum de não estacionariedade onde a série apresenta uma tendência, ou seja, a série flutua em volta de uma reta com alguma inclinação. Uma série também pode apresentar comportamento explosivo (Figura 2.1), caso em que a serie é estacionária, flutuando ao redor de um nível por certo tempo, depois muda de nível e flutua ao redor de um novo nível e assim por diante, Morettin e Tolo (1985).



**Figura 2.1:** *Comportamento explosivo.*

A classe de modelos ARIMA, que será discutida mais adiante, é capaz de descrever satisfatoriamente séries estacionárias e não estacionárias, mas que não apresentem comportamento explosivo.

Um teste muito usado para verificar a estacionariedade é o teste da raiz unitária. O meio mais fácil de apresentar este teste é considerar o seguinte modelo:

$$Y_t = \rho Y_{t-1} + u_t \quad (2.1)$$

ou na forma alternativa

$$\Delta Y_t = (\rho - 1) Y_{t-1} + u_t = \delta Y_{t-1} + u_t \quad (2.2)$$

em que  $\delta = (\rho - 1)$ ,  $\Delta Y_t = (Y_t - Y_{t-1})$ ,  $\rho$  é o coeficiente estimado para  $Y_{t-1}$  e  $u_t$  segue processo ruído branco.

Se  $\delta = 0$  ou da mesma forma  $\rho = 1$ , pode-se escrever (2.2) como  $\Delta Y_t = (Y_t - Y_{t-1}) = u_t$ , o que nos diz que as primeiras diferenças da série temporal seguem caminho aleatório ou que a série temporal é estacionária, pois, por hipótese,  $u_t$  é puramente aleatório.

Sob a hipótese nula  $\rho = 1$ , a estatística  $\tau$  tem seus valores críticos tabulados por Dickey e Fuller (1979) com base em simulações. Então, para saber se uma série é ou não estacionária, estima-se uma regressão do tipo (2.1), dividindo o coeficiente  $\rho$  estimado por seu desvio padrão para calcular a estatística  $\tau$ . Desta forma, segue-se consultando as tabelas de Dickey e Fuller (DF) para ver se a hipótese nula  $\rho = 1$  foi rejeitada. Se o valor absoluto calculado da estatística  $\tau$  (isto é,  $|\tau|$ ) excede os valores críticos absolutos  $\tau$  de DF, então não rejeita-se a hipótese de que a dada série temporal é estacionária, Gujarati, D.N. (2000).

Se uma seqüência  $\{y_t, t \in T\}$  tem  $E(y_t) = 0 \forall t \in T$ ,  $Var(y_t) = \sigma^2 \forall t (0 < \sigma^2 < \infty)$ , e  $Cov(y_t, y_{t+h}) = 0 \forall t, t+h \in T$  com  $h \neq 0$ , então  $y_t$  é dita ser ruído branco. Quando  $y_t \sim N$ , então trata-se de ruído branco Gaussiano.

Uma série  $\{y_t, t \in T\}$  é um passeio aleatório quando  $y_t = y_{t-1} + \xi_t$ , onde  $\xi_t \sim RB(0, \sigma^2)$ . Se  $\xi_t$  possui distribuição normal diz-se, então, que é um passeio aleatório Gaussiano.

Em uma série temporal existem flutuações irregulares as quais são possíveis de se observar através de decomposição da série. As duas principais componentes não observáveis de uma série temporal são tendência e variação sazonal. Sazonalidade são flutuações periódicas (diária, semanal, anual, etc). É uma flutuação fácil de ser medida e extraída dos dados obtendo uma série dessazonalizada. As tendências são efeitos longos (persistentes, continuados) na média, de aumento ou decréscimo. A dificuldade é definir os prazos muito longos, Morettin e Tolo (1985).

## 2.2 Funções de Autocovariância e Autocorrelação

Em uma série temporal a função de autocovariância é uma ferramenta importante, pois avalia se existe dependência (relação) entre as observações da série, veja, Cordeiro, D.M. (2002). Seja  $\{y_t, t \in T\}$  uma seqüência tal que  $Var(y_t) = \sigma^2 < \infty, \forall t \in T$ , então sua função de



autocovariância é

$$\gamma(t, t+h) = Cov(y_t, y_{t+h}) = E[(y_t - \mu_t)(y_{t+h} - \mu_{t+h})].$$

Quando se trata de um processo estacionário, Brockwe e Davis (1991), então:

- i)  $\gamma_0 \geq 0$   $Var(y_t) = \sigma_y^2$ ;
- ii)  $|\gamma_h| \leq \gamma_0 \forall h$ ;
- iii) a função  $\gamma(\cdot)$  é par, isto é,  $\gamma_h = \gamma_{-h} \quad \forall h$ .

A autocovariância depende da unidade de medida de  $y$ , portanto, é conveniente normalizar as autocovariâncias dividindo-as pelo produto dos respectivos desvios padrão, obtendo assim a função de autocorrelação (fac).

A fac de um processo estacionário, caso em que a variância é constante, é da forma

$$\rho_h = \frac{Cov(y_t, y_{t+h})}{\sigma_y^2} = \frac{\gamma_h}{\gamma_0}.$$

A função de autocorrelação de modo geral é dada por

$$\rho_h = \frac{Cov(y_t, y_{t+h})}{\sigma_t \sigma_{t+h}} = \frac{\gamma_h}{\sigma_t \sigma_{t+h}},$$

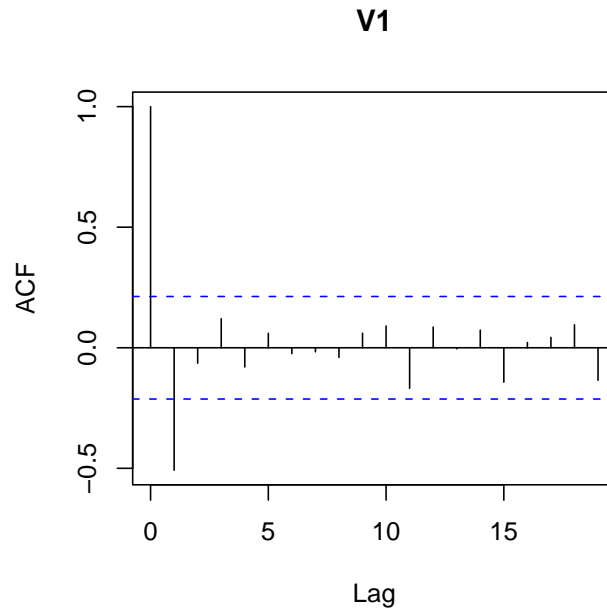
e tem as seguintes características básicas:

- i)  $\rho_0 = 1$ ;
- ii)  $|\rho_h| \leq 1 \quad \forall h$ ;
- iii) a função  $\rho(\cdot)$  é par, isto é,  $\rho_h = \rho_{-h} \quad \forall h$ .

Os valores da autocorrelação calculados com base numa amostra  $\hat{\rho}_h$  é chamada de autocorrelação amostral. O gráfico dos valores de  $\hat{\rho}_h$  versus os valores não-negativos de  $h$  é chamado de correlograma ou função de autocorrelação amostral.

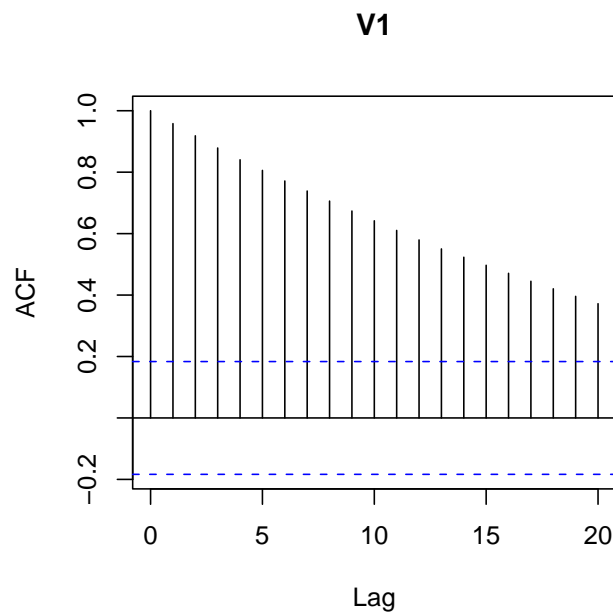
Obtém-se, então, uma ferramenta útil para uma análise preliminar dos dados observados de uma série, pois o correlograma permite verificar se o processo é estacionário ou não, salvo alguns casos. Este gráfico também oferece sugestões de modelos para processos estacionários.

O correlograma de um processo estacionário decai bruscamente para “zero”, ou melhor dizendo, para um intervalo onde admiti-se  $\rho_h$  nula. Um exemplo é visto na Figura 2.2 que apresenta o correlograma de uma série estacionária.



**Figura 2.2:** Correlograma de uma série simulada.

Caso o processo não seja estacionário o correlograma apresenta um decaimento exponencial (lento) ou ondas senóides amortecidas. Um exemplo é visto na Figura 2.3 que apresenta o correlograma de uma série não estacionária.



**Figura 2.3:** Correlograma do índice de custo de vida no município de São Paulo; observações mensais de Janeiro de 1970 a Junho de 1980.

A significância estatística de qualquer  $\hat{\rho}_k$  pode ser julgada por seu erro-padrão. Bartlett (1946),

mostrou que se uma série temporal for puramente aleatória, ou seja, se exibir ruído branco, os coeficientes de autocorrelação amostral são, aproximadamente, distribuídos normalmente com média zero e variância  $1/n$ , logo desvio-padrão  $1/\sqrt{n}$ , em que  $n$  é o tamanho da amostra. Então, seguindo as propriedades da distribuição normal padrão, o intervalo de confiança de 95% para qualquer  $\hat{\rho}_k$  será  $\pm 1,96 (1/\sqrt{n})$ . Assim, se um  $\rho_k$  estimado se situar no intervalo  $(-1,96 (1/\sqrt{n}), +1,96 (1/\sqrt{n}))$ , não rejeitamos a hipótese de que o verdadeiro  $\rho_k$  seja zero. Mas se ele se encontrar fora desse intervalo de confiança, então podemos rejeitar a hipótese de que o verdadeiro  $\rho_k$  é zero, Gujarati, D.N. (2000).

A autocorrelação parcial pode ser considerada como uma correlação entre duas variáveis quaisquer  $y_{t-k}$  e  $y_t$ , separadas por  $k$  períodos com os efeitos das variáveis intermediárias  $y_{t-k+1}$ ,  $y_{t-k+2} \dots, y_{t-1}$  eliminadas, tendo-se portanto

$$Corr (y_{t-k}, y_t | y_{t-k+1}, y_{t-k+2} \dots, y_{t-1}),$$

que é a influência direta de  $y_{t-k}$  em  $y_t$ .

A função de autocorrelação parcial (facp) é outra ferramenta que também contribui na escolha de um modelo adequado, e é denotada por

$$\phi_{kk} = \frac{|P_k^*|}{|P_k|},$$

onde  $P_k$  é a matriz de autocorrelações e  $P_k^*$  é a matriz  $P_k$  com a última coluna substituída pelo vetor de autocorrelações. Obtém-se, assim,

$$\begin{aligned} \phi_{11} &= \rho_1, \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}, \\ \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}. \end{aligned}$$

## 2.3 Modelo de Alisamento Exponencial

O principal interesse em análise de séries temporais é o de previsão de valores futuros com base em observações passadas. A observação mais recente, é de se esperar, tem mais informação sobre o futuro que observações de um passado distante. O método chamado alisamento exponencial é baseado numa possível solução para este problema. A grande popularidade deste método é devido à facilidade nos cálculos, à necessidade pequena de armazenamento de informação, ao uso de cálculos recursivos, à eficiência computacional e à sua razoável precisão.

**2.3.1 Modelo de Alisamento Exponencial Simples** O modelo de alisamento exponencial simples se adequa a situações onde não existe tendência ou sazonalidade. Este método se baseia na suposição de que mudanças futuras na média são imprevisíveis. Dessa forma, o algoritmo estima o nível atual da série (média) e assim prevê valores futuros.

Com base nas observações  $y_t, y_{t-1}, y_{t-2}, y_{t-3}, \dots$  existem duas possibilidades para se estimar o nível da série:

- i) usar uma média aritmética de todos os valores e assumir que todas as observações possuem o mesmo peso. Neste caso, é desconsiderado que as observações mais recentes contenham informações mais úteis;
- ii) admitir que o último valor da série como estimativa do nível. Neste caso, desprezando toda contribuição que as observações não tão recentes são capazes de fornecer.

O modelo de alisamento exponencial simples consiste em uma junção das duas possibilidades acima descritas. Trata-se de uma média ponderada de todas as observações, onde o peso dado a cada observação cresce exponencialmente à medida que o tempo cresce. Desta forma as observações mais recentes recebem pesos maiores que observações mais antigas. Portanto, o nível estimado da série no instante  $t$ , para  $0 \leq \alpha \leq 1$ , é dado por:

$$N_t = \alpha y_t + \alpha (\alpha - 1)^1 y_{t-1} + \alpha (\alpha - 1)^2 y_{t-2} + \alpha (\alpha - 1)^3 y_{t-3} + \dots \quad (2.3)$$

É importante observar que  $\alpha + \alpha (\alpha - 1)^1 + \alpha (\alpha - 1)^2 + \alpha (\alpha - 1)^3 + \dots = 1$

A escolha da constante de alisamento  $\alpha$  é o problema central agora. Tal constante pode ser escolhida de 3 (três) formas diferentes:

- Por experiência passada (o usuário pode já ter trabalhado com séries semelhantes antes);
- Por análise gráfica: Ao se traçar um gráfico do comportamento da série ao longo do tempo é razoável pensar em  $\alpha$  perto de zero, quando a série evolui de forma irregular, ou  $\alpha$  próximo de 1, quando sua evolução se dá de forma suave;
- Ou, por fim, usando a minimização dos erros quadrados de previsão,

$$S_\alpha = \sum_{t=2}^n e_t^2,$$

onde

$$e_t = y_t - N_{t-1}, \text{ e } N_{t-1} = \hat{y}_{t-1}(1), \quad t = 3, 4, 5, \dots, n$$

e  $\hat{y}_{t-1}(1)$  é a previsão da série para  $y_t$  no tempo  $t - 1$ . É então razoável concluir que  $N_t = N_{t-1} + \alpha e_t$ , assim se  $e_t = 0$ , conclui-se que a previsão foi exata e não há necessidade de correção. Já, se o valor da série foi subestimado, se faz necessário a adição de uma quantidade para correção deste valor na obtenção do próximo valor da série. É possível então observar que este algoritmo tem uma estrutura que permite corrigir, caso necessário, o erro em uma próxima previsão.

**2.3.2 Modelo de Alisamento Exponencial de Holt** Este procedimento difere do alisamento exponencial simples por levar em conta a tendência, porém ainda neste se faz necessária a suposição de que não existe sazonalidade. Neste método, além de estimar o nível atual da série, também se faz necessária a estimativa da inclinação corrente da série (tendência). O algoritmo de tendência linear de Holt usa como base para tais estimativas as seguintes fórmulas:

$$N_t = \alpha y_t + (1 - \alpha) (N_{t-1} + T_{t-1}), \quad 0 \leq \alpha \leq 1,$$

$$T_t = \beta (N_t - N_{t-1}) + (1 - \beta) T_{t-1}, \quad 0 \leq \beta \leq 1,$$

onde  $N_t$  é a estimativa do nível,  $T_t$  é a estimativa da tendência, no instante  $t$ , e  $\alpha$  e  $\beta$  são constantes de suavização. A previsão do nível da série no instante  $t + h$  é:

$$\hat{y}_t(h) = N_t + hT_t, \quad t = 4, 5, \dots, n.$$

Admita que a previsão foi exata ( $e_t = 0$ ) então mantém-se a previsão da série. Já se o valor da série foi subestimado, se faz necessário a adição de uma quantidade para correção deste valor na obtenção do próximo valor da série. Contudo, se  $e_t < 0$ , é necessário uma correção diminuindo as estimativas do nível e da tendência em valor proporcional ao erro. A correção no nível é controlada por  $\alpha$  enquanto que a correção na tendência é controlada por  $\alpha$  e  $\beta$ .

**2.3.3 Alisamento Exponencial de Holt-Winters** O modelo de alisamento exponencial de Holt-Winters objetiva permitir a acomodação de sazonalidade ao algoritmo de tendência linear de Holt. Para estimar o fator de sazonalidade  $F_t$  existem dois tipos de procedimentos que serão escolhidos de acordo com as características da série considerada.

**Modelo Sazonal Multiplicativo** Se a sazonalidade for multiplicativa, as fórmulas de recorrência são:

$$\begin{aligned} N_t &= \alpha \frac{y_t}{F_{t-s}} + (1 - \alpha) (N_{t-1} + T_{t-1}), \quad 0 \leq \alpha \leq 1, \\ T_t &= \beta (N_t - N_{t-1}) + (1 - \beta) T_{t-1}, \quad 0 \leq \beta \leq 1, \\ F_t &= \gamma \frac{y_t}{N_t} + (1 - \gamma) F_{t-s}, \quad 0 \leq \gamma \leq 1 \end{aligned}$$

onde  $s$  é fator sazonal e  $\alpha$ ,  $\beta$  e  $\gamma$  são constantes de alisamento. Para tanto, as previsões são:

$$\begin{aligned} \hat{y}_t(h) &= (N_t + hT_t) F_{t+h-s}, \quad h = 1, 2, \dots, s \\ \hat{y}_t(h) &= (N_t + hT_t) F_{t+h-2s}, \quad h = s, \dots, 2s \\ &\vdots \end{aligned}$$

aqui, a forma de correção dos erros é

$$\begin{aligned} N_t &= N_{t-1} + T_{t-1} + \alpha \frac{e_t}{F_{t-s}}; \\ T_t &= T_{t-1} + \frac{\alpha\beta e_t}{F_{t-s}}; \\ F_t &= F_{t-s} + \frac{\gamma(1-\alpha)e_t}{N_t}. \end{aligned}$$

**Modelo Sazonal Aditivo** O método multiplicativo pode ser adaptado para lidar com situações onde o fator sazonal é aditivo. Neste caso, as fórmulas de recorrência são:

$$\begin{aligned} N_t &= \alpha (y_t - F_{t-s}) + (1 - \alpha) (N_{t-1} + T_{t-1}), \quad 0 \leq \alpha \leq 1, \\ T_t &= \beta (N_t - N_{t-1}) + (1 - \beta) T_{t-1}, \quad 0 \leq \beta \leq 1, \\ F_t &= \gamma (y_t - N_t) + (1 - \gamma) F_{t-s}, \quad 0 \leq \gamma \leq 1 \end{aligned}$$

E, neste caso, as previsões são:

$$\begin{aligned}\hat{y}_t(h) &= N_t + hT_t + F_{t+h-2}, \quad h = 1, 2, \dots, s \\ \hat{y}_t(h) &= N_t + hT_t + F_{t+h-2s}, \quad h = s, \dots, 2s \\ &\vdots\end{aligned}$$

Por fim, o mecanismo de correção dos erros é dado por

$$\begin{aligned}N_t &= N_{t-1} + T_{t-1} + \alpha e_t; \\ T_t &= T_{t-1} + \alpha \beta e_t; \\ F_t &= F_{t-s} + \gamma(1 - \alpha) e_t.\end{aligned}$$

## 2.4 Modelos Auto-Regressivos AR(p)

Suponha  $\{y_t, t \in T\}$  uma série temporal sem movimentos sazonais. Portanto,  $\{y_t\}$  pode ser escrita como função linear da variável imediatamente anterior mais um ruído branco, ou seja,

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t, \quad (2.4)$$

onde  $u_t \sim RB(0, \sigma^2)$ ,  $\phi_1, \dots, \phi_p$  e  $c$  são parâmetros desconhecidos, e ainda  $\phi_1, \dots, \phi_p$  são parâmetros auto-regressivos. É razoável também que (2.4) seja vista como,

$$c + u_t = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) y_t \quad (2.5)$$

onde

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

ou ainda

$$\phi(B) y_t = c + u_t,$$

com

$$B y_t = y_{t-1}.$$

Um processo AR(p) é estacionário quando todas as raízes da equação  $1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0$ , no plano complexo, encontram-se fora do círculo unitário. Já com relação à invertibilidade, não existe restrições, todo processo AR(p) é invertível.

- Média

$$E(y_t) = E(c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t)$$

$$c = (1 - \phi_1 - \phi_2 - \dots - \phi_p) \mu$$

$$\mu = \frac{c}{1 - \phi_1 - \phi_2 - \dots - \phi_p};$$

- Autocovariância.

A autocovariância de ordem  $j$  é dada por

$$E[(y_t - \mu)(y_{t-j} - \mu)] = \phi_1 E[(y_{t-1} - \mu)(y_{t-j} - \mu)] + \dots + \phi_p E[(y_{t-p} - \mu)(y_{t-j} - \mu)] + E[u_t(y_{t-j} - \mu)],$$

assim,

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}, \quad j = 1, 2, \dots \quad (2.6)$$

- Autocorrelação.

Dividindo (2.6) por  $\gamma_0$

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p}, \quad j = 1, 2, \dots$$

## 2.5 Modelos de Médias Móveis MA(q)

Suponha  $\{y_t, t \in T\}$  uma série temporal estacionária. Portanto,  $\{y_t\}$  pode ser escrita como função linear do valor atual de um processo ruído branco mais o valor deste mesmo processo no instante imediatamente anterior multiplicado por uma constante, Gujarati, Damodar N. (2000), ou seja,

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \dots + \theta_q u_{t-q} \quad (2.7)$$

ou

$$y_t = \mu + \theta(B) u_t,$$

para

$$\theta(B) = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q),$$

onde  $u_t \sim RB(0, \sigma^2)$ ,  $\theta_1, \dots, \theta_q$  e  $\mu$  são parâmetros desconhecidos, e ainda  $\theta_1, \dots, \theta_q$  são chamados de parâmetros de médias móveis.



- Média

$$E(y_t) = E(\mu + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q})$$

$$E(y_t) = \mu;$$

- Variância

$$\text{var}(y_t) = \text{var}(\mu + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q})$$

$$\text{var}(y_t) = (1 + \theta_1^2 + \cdots + \theta_q^2) \sigma^2;$$

- Autocovariância

$$\gamma_j = \begin{cases} (\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_q\theta_{q-j}) \sigma^2, & \text{para } j = 1, 2, \dots, q, \\ 0, & \forall j > q \end{cases}$$

Em um processo MA( $q$ ) não existem restrições com relação a estacionariedade, uma vez que o processo é sempre estacionário. Todavia, a condição de invertibilidade é satisfeita se todas as raízes  $\theta(B) = 0$  estiverem fora do círculo unitário.

## 2.6 Modelos ARMA, ARIMA, SARIMA, ARMAX

Um processo ARMA( $p, q$ ) é definido como sendo um modelo que possui termos auto-regressivos e termos de médias móveis, simultaneamente, Brockwell e Davis (1996), isto é

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q}$$

ou

$$\phi(B) y_t = c + \theta(B) u_t$$

onde  $u_t \sim RB(0, \sigma^2)$ ,  $\phi_1, \dots, \phi_p$  são parâmetros auto-regressivos e  $\theta_1, \dots, \theta_q$  são parâmetros de médias móveis. Ainda aqui o polinômio auto-regressivo  $\phi(B)$  continua atendendo às condições de estacionariedade, assim como o polinômio de médias móveis  $\theta(B)$  continua atendendo às condições de invertibilidade.

Existem processos cujo polinômio AR não atende as condições de estacionariedade, contudo sua primeira diferença atende, e neste caso é possível modelar  $\Delta^1 y_t$  através de uma estrutura ARMA( $p, q$ ). Uma série pode se tornar estacionária através de  $d$  diferenças sucessivas. O

processo original então será integrado de ordem  $d$ , ou seja, um processo ARIMA( $p, d, q$ ) definido como

$$\phi(B) \left[ (1 - B)^d y_t - \mu \right] = \theta(B) u_t,$$

onde  $d$  indica a ordem de diferenciação, Box e Jenkins (1976).

Supondo um processo ARMA que além de ser não-estacionário apresenta um comportamento sazonal de período  $s$ . Supondo ainda que pode ser feita uma transformação linear na série  $\{y_t\}$  resultando em uma nova série  $\{z_t\}$  não sazonal. Esta série pode ser ARIMA, mas com a presença de relações entre as observações separadas em  $s$  períodos, Newbold e Bos (1994). Seja  $y_t$  a série de interesse observada com período de sazonalidade  $s$ . Sejam ainda

$$\Phi(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{sP}$$

o operador auto-regressivo sazonal, estacionário, de ordem  $P$ ,

$$\Theta(B^s) = 1 + \Theta_1 B^s + \dots + \Theta_Q B^{sQ}$$

o operador de médias móveis sazonal, invertível, de ordem  $Q$  e  $\Delta_s^D = (1 - B^{sQ})$ , onde  $D$  é o indicador de diferenças sazonais. O modelo SARIMA  $(p, d, q) \times (P, D, Q)$  é dado por

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) (1 - \Phi_1 B^s - \dots - \Phi_P B^{sP})$$

$$\left[ (1 - B)^d (1 - B^s)^D - \mu \right] y_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) (1 + \Theta_1 B^s + \dots + \Theta_Q B^{sQ}) u_t$$

ou

$$\phi(B) \Phi(B^s) \left[ (1 - B)^d (1 - B^s)^D - \mu \right] y_t = \theta(B) \Theta(B^s) u_t,$$

onde  $u_t \sim RB(0, \sigma^2)$ .

Box e Jenkins (1970), usaram o modelo SARIMA  $(0, 1, 1) \times (0, 1, 1)$  para modelar o logaritmo do número mensal de passageiros em companhias aéreas. Mais tarde este modelo se mostrou útil para modelar outras séries e foi chamado de modelo “airline”. Esse modelo é dado por

$$(1 - B)^d (1 - B^s)^D y_t = \mu + (1 + \theta_1 B) (1 + \Theta_1^s B^s) u_t$$

Já o modelo ARMAX (“autoregressive-moving average with explanatory variables”) é uma estrutura de modelagem que possui considerável generalidade para se analisar impactos de

regressores dinamicamente. Aqui, a atenção será dada ao caso de uma variável dependente  $y$ ; e uma única variável independente ou explicativa  $x$ . O modelo é definido como

$$y_t = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k x_{tk} + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q},$$

onde  $u_t \sim RB(0, \sigma^2)$ ,  $\beta_i, \phi_j$  e  $\theta_r$ , com  $i = 0, \dots, k$ ,  $j = 1, \dots, p$ ,  $r = 1, \dots, q$ , são parâmetros fixados. Notamos que no caso especial onde os  $\beta$ 's são todos iguais a zero, obtém-se o modelo ARMA( $p, q$ ) para a variável dependente  $y_t$ . Quando os  $\theta$ 's são todos iguais a zero, os outros parâmetros do modelo são facilmente estimados por mínimos quadrados ordinários. Contudo, quando os  $\theta$ 's são não-nulos, a estimação dos parâmetros requer o uso de um algoritmo de otimização não-linear. É importante lembrar que mais de uma variável explanatória pode ser incluída no modelo, contanto que a estrutura de defasagem para cada variável seja finita e que o número total de parâmetros a serem estimados não exceda o número de observações.

## 2.7 Modelagem de Box-Jenkins

Este método consiste em ajustar modelos ARIMA (para mais detalhes veja Seção 3) e suas variações. A estratégia para construção do modelo é baseada em um ciclo iterativo e que utiliza os próprios dados para a escolha da estrutura deste modelo. Inicialmente, é proposta uma classe de modelos e procede-se à identificação de um modelo particular com base em critérios específicos. Em seguida, os parâmetros são estimados, e os resíduos do modelo ajustado são então avaliados. Caso o modelo selecionado não seja adequado, o ciclo é reiniciado. Em geral, os modelos mais convenientes são parcimoniosos, ou seja, contêm poucos parâmetros e mesmo assim geram previsões precisas.

## 2.8 Identificação, Estimação e Checagem

É importante, quando a intenção é identificar que modelo será mais adequado, observar o gráfico das autocorrelações parciais amostrais. Neste caso, o modelo seja AR( $p$ ), deve apresentar um corte na ordem  $p$ . Já em processos MA( $q$ ), o corte deve ocorrer na ordem  $q$  no gráfico das autocorrelações amostrais. Contudo, sendo o processo misto, a identificação por este procedimento

torna-se muito difícil. Neste caso, três critérios de seleção de modelos são úteis, são eles:

$$\begin{aligned}AIC &= -2 \log \hat{L} + 2(p + q), \\AIC_C &= -2 \log \hat{L} + \frac{2(p + q)T}{T - p - q - 1}, \\BIC &= -2 \log \hat{L} + (p + q) \log T,\end{aligned}$$

onde  $T$  é o número de observações e  $\hat{L}$  é a verossimilhança maximizada. O melhor modelo é aquele que apresentar o menor  $AIC$ ,  $AIC_C$  e/ou o menor BIC. O  $AIC$  (“Akaike Information Criterion”) e o  $AIC_C$  (“Akaike Information Criterion Corrected”) não são consistentes, a dimensão do modelo pode ser super estimada, ao passo que o  $BIC$  (“Bayesian Information Criterion”) é consistente e mais parcimonioso. O  $BIC$  por impor uma penalidade mais pesada à inclusão de parâmetros, escolhe modelos cuja dimensão não ultrapassa a dos modelos selecionados pelo  $AIC$  quando há mais de oito observações.

A estimação, como no item anterior, é feita com base nas observações e para tanto podem ser usados diferentes métodos, são eles:

- Método de máxima verossimilhança,
- Método de mínimos quadrados exatos,
- Método de mínimos quadrados condicional.

Por último, a etapa de verificação, busca decidir se o modelo encontrado fornece uma representação estatisticamente adequada dos dados, ou melhor dizendo, fazer a checagem de diagnóstico. Uma das formas de checagem de diagnóstico é a análise dos resíduos, já que estes fornecem uma estimativa dos erros, que são desconhecidos. Caso o modelo represente bem os dados da série, os resíduos devem comportar-se como um processo ruído branco, logo não deve haver autocorrelação e a variância deve ser constante.

Os procedimentos de previsão utilizados na prática variam muito, podendo ser simples e intuitivos Morettin (1981). Um modelo não conduz, necessariamente, a um procedimento (ou fórmula) de previsão. Será necessário especificar, além do modelo, uma função perda como erro quadrático médio (EQM), para chegar ao procedimento Morettin e Toloí (1985).

### 3 METODOLOGIA

#### 3.1 Modelos ARIMA

**3.1.1 Identificação** Nesta seção será tratada a primeira fase do método de Box e Jenkins para o particular modelo ARIMA. Nesta fase o interesse maior é de reduzir o número de possíveis modelos do que escolher o “melhor” modelo. Uma das formas para este estudo consiste na representação gráfica, que se dá por meio da plotagem dos dados ao longo do tempo. Dessa forma, é possível identificar com facilidade os efeitos da tendência e da sazonalidade. No entanto, esta escolha é feita principalmente com base nas autocorrelações e autocorrelações parciais estimadas, que se espera representem adequadamente as respectivas quantidades teóricas que são desconhecidas.

A função de autocorrelação  $\rho_j$  (fac) é estimada por

$$r_j = \frac{c_j}{c_0}, \quad j = 0, 1, \dots, N-1, \quad (3.1)$$

onde  $c_j$  é a estimativa da função de autocovariância  $\gamma_j$  (facv),

$$c_j = \frac{1}{N} \sum_{t=1}^{N-j} [(Y_t - \bar{Y})(Y_{t+j} - \bar{Y})], \quad j = 0, 1, \dots, N-1 \quad (3.2)$$

sendo  $\bar{Y} = \frac{1}{N} \sum_{t=1}^N Y_t$  a média amostral.

Uma expressão aproximada para a variância de  $r_j$ , para um processo estacionário normal, é dada por

$$\text{Var}(r_j) \cong \frac{1}{N} \sum_{v=-\infty}^{\infty} [\rho_v^2 + \rho_{v+j}\rho_{v-j} - 4\rho_j\rho_v\rho_{v-j} + 2\rho_v^2\rho_j^2]. \quad (3.3)$$

Num processo em que para  $v > q$  as autocorrelações são nulas, todos os termos do lado direito de (3.3) anulam-se para  $j > q$ , exceto o primeiro, ficando então,

$$\text{Var}(r_j) \cong \frac{1}{N} \left[ 1 + 2 \sum_{v=1}^q \rho_v^2 \right], \quad j > q. \quad (3.4)$$

Como a autocorrelação  $\rho_v$  é desconhecida, será usada sua estimativa  $r_v$ , obtendo-se

$$\hat{\sigma}^2(r_j) \cong \frac{1}{N} \left[ 1 + 2 \sum_{v=1}^q r_v^2 \right], \quad j > q. \quad (3.5)$$

Logo, sob a hipótese  $\rho_j = 0$ , para  $j > q$  e  $N$  suficientemente grande, a distribuição de  $r_j$  é, aproximadamente normal, com média igual a zero e variância dada por (3.5). Com isso, é possível então construir um intervalo de confiança aproximado para as autocorrelações.

$$r_j \pm t_\gamma \cdot \hat{\sigma}(r_j), \quad (3.6)$$

onde  $t_\gamma$  é o valor da estatística  $t$  de Student com  $N - 1$  graus de liberdade. Na prática, usa-se  $t_\gamma = 2$  e, então, pode-se considerar  $\rho_v$  como sendo significativamente diferente de zero se

$$|r_j| > 2\hat{\sigma}(r_j), \quad j > q \quad (3.7)$$

Quenouille (1949), mostra que as facp estimadas de ordens  $p + 1, p + 2, \dots$  são, sob a hipótese de que o processo seja  $AR(p)$ , aproximadamente independentemente distribuídas, com

$$Var(\hat{\phi}_{jj}) \cong \frac{1}{N}, \quad j > p \quad (3.8)$$

e, se o número de observações  $N$  for suficientemente grande,  $\hat{\phi}_{jj}$  tem distribuição aproximadamente normal com média zero e variância dada como em (3.8). Portanto pode-se considerar  $\hat{\phi}_{jj}$  significativamente diferente de zero se

$$\left| \hat{\phi}_{jj} \right| > \frac{2}{\sqrt{N}}, \quad j > p. \quad (3.9)$$

O procedimento de identificação dos modelos  $ARIMA(p, d, q)$  tem como objetivo central determinar os valores de  $p, d, q$ , e consiste necessariamente de três partes:

- I – Observar a necessidade de estabilizar a variância por meio de uma transformação. Veja, com detalhes, Morettin e Tolo (2004).
- II – Tornar a série obtida no item I estacionária, tomando diferenças até que o processo  $\Delta^d Y_t$  seja reduzido a um  $ARMA(p, q)$ . As diferenças são necessárias até que a fac amostral de  $W_t = \Delta^d Y_t$  decresça rapidamente para zero. Então, um teste para verificar a existência de raízes unitárias poderá ser de grande utilidade.
- III – Analisar o comportamento das autocorrelações e autocorrelações parciais estimadas com o objetivo de identificar o modelo  $ARMA(p, q)$  resultante. Estes comportamentos devem imitar os comportamentos das respectivas quantidades teóricas. Um resumo das propriedades desses modelos está na Tabela 3.1. Para maiores detalhes, veja Morettin e Tolo (2004).

Aqui, serão ajustados modelos ARMA de baixa ordem como,  $(1, 1)$ ,  $(1, 2)$ ,  $(2, 1)$  e utilizados critérios que permitam escolher o modelo mais adequado. Isso porque a *fac* e *facp* não são muito úteis para identificar modelos ARMA, dada a forma complicada destas funções. A Tabela 3.1 contém informação sobre o comportamento típico da *fac* e *facp* segundo a possível ordem do modelo.

Ordem	$(1, d, 0)$	$(0, d, 1)$
Comportamento de $\rho_k$	Decai exponencialmente	Somente $\rho_k \neq 0$
Comportamento de $\phi_{kk}$	Somente $\phi_{kk} \neq 0$	Decai exponencialmente
Ordem	$(2, d, 0)$	$(0, d, 2)$
Comportamento de $\rho_k$	Denominada por misturas de exponenciais ou ondas senóides amortecidas	Somente $\rho_k \neq 0$
Comportamento de $\phi_{kk}$	Somente $\phi_{kk} \neq 0$	Denominada por misturas de exponenciais ou ondas senóides amortecidas
Ordem	$(1, d, 1)$	
Comportamento de $\rho_k$	Decai exponencialmente após o lag 1	
Comportamento de $\phi_{kk}$	Denominada decaimento exponencial após o lag 1	

**Tabela 3.1:** *Comportamento da fac e facp segundo a ordem do modelo.*

No estágio citado no item II do procedimento, o maior problema é evitar um excesso de diferenças. McLeod (1983), faz alguns comentários:

1. Um número excessivo de diferenças resulta em um valor negativo da autocorrelação de ordem 1 da série diferenciada.
2. O monitoramento da variância é muito importante para escolher o valor apropriado de  $d$ , pois quando existe um excesso de diferenças a variância da série transformada aumenta e o contrário acontece quando a série é corretamente diferenciada.

Na prática,  $d = 0, 1$ , ou  $2$  e é suficiente inspecionar as primeiras 15 ou 20 autocorrelações da série e de suas diferenças.

Na metodologia já mencionada para identificação de modelos ARIMA um dos maiores obstáculos é a identificação do melhor modelo. Nesta seção, serão abordadas algumas das mais usadas formas alternativas de identificação apresentadas na literatura.

Apresentam-se aqui, métodos baseados em uma função penalizadora. Seja a quantidade,

$$P(k, l) = \ln \hat{\sigma}_{k,l}^2 + (k + l) \frac{C(N)}{N} \quad (3.10)$$

onde  $\hat{\sigma}_{k,1}^2$  é uma estimativa da variância dos resíduos quando ajustado um modelo ARMA( $k, l$ ),  $N$  é o número de observações da série, e  $C(N)$  é uma função do tamanho da série. O objetivo é encontrar valores de  $k$  e  $l$  que minimizem o valor de  $P(k, l)$ .

Quando o número de parâmetros aumenta, o termo  $(k + l) \frac{C(N)}{N}$ , chamado penalizador, também aumenta, enquanto a variância residual  $\hat{\sigma}_{k,l}$  diminui.

a) Critério de Informação Akaike (*AIC*).

O *AIC* “Akaike Information Criterion” talvez tenha sido um dos primeiros critérios, proposto por Akaike (1973, 1974). A proposta é então escolher  $k$  e  $l$  que minimizem o critério

$$AIC(k, d, l) = N \ln \hat{\sigma}_u^2 + \frac{N}{N-d} 2(k + l + l + \partial_{d0}) + N \ln 2\pi + N, \quad (3.11)$$

onde

$$\partial_{d0} = \begin{cases} 1, & d = 0 \\ 0, & d \neq 0 \end{cases}$$

e  $\hat{\sigma}_u^2$  é o estimador de máxima verossimilhança de  $\sigma_u^2$ .

Levado-se em conta a comparação de modelos com  $N$  fixado e que geralmente a série identificada apropriadamente diferenciada, tem-se

$$AIC(k, l) = N \ln \hat{\sigma}_a^2 + 2(k + l + 2) \quad (3.12)$$

e, então, para determinar os valores de  $p$  e  $q$ , admite-se limites superiores  $K$  e  $L$  para  $k$  e  $l$  respectivamente. Com isso calculando-se o valor de (3.12) para todas as possíveis combinações  $(k, l)$ . Os limites  $K$  e  $L$  geralmente são funções do tamanho da série  $N$ , como por exemplo,  $K = L = \ln N$ . Para o caso dos modelos AR( $p$ ), este critério é dado da seguinte forma

$$AIC(k) = N \ln \hat{\sigma}_k^2 + 2k, \quad K \leq k \quad (3.13)$$

Uma versão modificada que inclui uma correção de viés incorporando uma penalidade ainda maior para modelos mais parametrizados é o  $AIC_C$  dado por

$$AIC_C(k) = N \ln \hat{\sigma}_k^2 + 2k + \frac{2(k+1)(k+2)}{N-k+2}, \quad K \leq k \quad (3.14)$$



Hurvich e Tsai (1989), utilizando simulações, mostram que quando  $N$  é pequeno ou quando  $K$  é uma fração “relativamente grande” de  $N$ , esta correção torna-se muito útil.

b) Critério de Informação Bayesiano (*BIC*).

O *BIC* “Bayesian Information Criterion” provoca um desencorajamento à inclusão de parâmetros nos modelos, procurando alcançar o maior grau possível de parcimônia. Akaike (1977), Rissanen (1978) e Schwarz (1978), propõem minimizar

$$BIC(k, l) = \ln \hat{\sigma}_{k,l}^2 + (k + l) \frac{\ln N}{N}, \quad (3.15)$$

onde  $\hat{\sigma}_{k,l}^2$  é a estimativa de máxima verossimilhança da variância residual do modelo ARMA( $k, l$ ).

Hannan (1980, 1982) mostra a forte consistência dos valores de  $p$  e  $q$  encontrados quando minimizada a expressão (3.15).

**3.1.2 Estimação** Uma vez que o modelo tenha sido selecionado de acordo com os procedimentos da seção anterior, o próximo passo é estimar os parâmetros desconhecidos do modelo.

Seja um modelo ARIMA( $p, d, q$ ) e um vetor  $\xi(\phi, \theta, \sigma_u^2)$ , em que  $\phi = (\phi_1, \dots, \phi_p)$  e  $\theta = (\theta_1, \dots, \theta_q)$  onde serão colocados os  $p + q + 1$  parâmetros do modelo. E considere também  $\eta(\phi, \theta)$ , com  $\phi$  e  $\theta$  como já visto. Supondo aqui, que quando  $d > 0$   $\mu_w = 0$ , em caso contrário, o número de parâmetros será  $p + q + 2$ , pois  $\mu_w$  será incluído como um parâmetro do modelo.

Considere-se as  $N$  observações  $Y_1, \dots, Y_N$ , e a função de máxima verossimilhança  $L(\xi|Y_1, \dots, Y_N)$  considerada como função de  $\xi$ . Os valores que maximizam  $L$  ou  $l = \log L$  serão os estimadores de máxima verossimilhança (EMV) de  $\xi$ .

Trabalhando com a suposição de que o processo  $a_t$  tem distribuição normal com média zero e variância  $\sigma_u^2$  para cada  $t$ , então os EMV serão, aproximadamente, estimadores de mínimos quadrados (EMQ).

Alcançada a estacionariedade, após as  $d$  diferenças necessárias, obtém-se  $n = N - d$  observações  $W_1, \dots, W_n$  onde  $W_t = \Delta^d Y_t$ . O modelo ARMA( $p, q$ ) obtido é estacionário e invertível, dado

por

$$u_t = \tilde{W}_t - \phi_1 \tilde{W}_{t-1} - \cdots - \phi_p \tilde{W}_{t-p} + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q}, \quad (3.16)$$

onde  $\tilde{W}_t = W_t - \mu_w$ .

a) Método dos Momentos.

A idéia, neste método, é substituir os momentos teóricos pelos seus correspondentes momentos amostrais, nas equações em que estes se relacionam com os parâmetros do modelo. Os estimadores dos parâmetros obtidos desta forma serão utilizados, na maioria das vezes, como estimativas iniciais em procedimentos iterativos de estimação não-linear.

b) Método de Máxima Verossimilhança.

Este método necessita de valores iniciais para os  $\tilde{W}'s$  e para os  $u's$  da expressão (3.16).

Para tanto, são conhecidos dois procedimentos, sendo um condicional e outro não-condicional.

i) Procedimento condicional.

A função de densidade conjunta de  $u_1, \dots, u_n$ , sob a suposição de  $u_t \sim N(0, \sigma_u^2)$ , é dada por

$$f(u_1, \dots, u_n) = (2\pi)^{-n/2} (\sigma_u)^{-n} \exp \left\{ - \sum_{t=1}^n \frac{u_t^2}{2\sigma_u^2} \right\}. \quad (3.17)$$

Serão denotados de  $W_t^*$  e  $u_t^*$  os  $p$  valores supostos de  $W_t$  e os  $q$  valores também supostos de  $u_t$  usados para calcular  $u_1, \dots, u_n$ .

A partir de (3.16) e (3.17), será obtida a função de verossimilhança, condicional à escolha de  $W_t$  e  $u_t$ . Tem-se,

$$L(\xi | W, W^*, u^*) = (2\pi)^{-n/2} (\sigma_u)^{-n} \cdot \exp \left\{ - \frac{1}{2\sigma_u^2} \sum_{t=1}^n \left( \tilde{W}_t - \phi_1 \tilde{W}_{t-1} - \cdots - \phi_p \tilde{W}_{t-p} + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q} \right)^2 \right\}.$$

Usando um asterisco para denotar  $l$  e  $S$  condicionais a  $W = (W_1, \dots, W_n)$ ,  $W^* = (W_1^*, \dots, W_p^*)$ ,  $u^* = (u_1^*, \dots, u_q^*)$  e, em seguida, aplicando o logaritmo, obtém-se

$$l_*(\xi | W, W^*, u^*) \propto -n \log(\sigma_u) - \frac{S_*(\eta)}{2\sigma_u^2}, \quad (3.18)$$

onde  $S_*(\eta) = S(\eta | W, W^*, u^*) = \sum_{t=1}^n u_t^2(\eta | W, W^*, u^*)$  é denominada soma dos quadrados (SQ) condicional.

Existem duas opções para a escolha de  $W^*$  e  $u^*$  :

- a) O primeiro é admitir que seus elementos são iguais às suas esperanças  $E(u_t) = 0$ , e se  $E(W_t) \neq 0$  substitui-se todos os elementos de  $W^*$  por  $\bar{W}$ .
- b) A segunda é indicada quando alguma raiz de  $\phi(B) = 0$  estiver próxima do círculo unitário. O indicado então, é calcular  $u_{p+1}, \dots, u_{p+n}$  a partir de (3.16) e assumindo os valores anteriores de  $a_t$  iguais a zero. E, assim estarão sendo usados os valores observados de  $w_t$ , obtendo-se

$$u_{p+1} = \tilde{W}_{p+1} - \phi_1 \tilde{W}_p - \dots - \phi_p \tilde{W}_1 + \theta_1 u_p + \dots + \theta_q u_{p-q+1}$$

etc.

- ii) Procedimento não-condicional.

Dada a função de verossimilhança não-condicional

$$l(\xi) \cong -n \log(\sigma_u) - \frac{S(\eta)}{2\sigma_u^2}, \quad (3.19)$$

onde

$$S(\eta) = S(\phi, \theta) = \sum_{t=-\infty}^n [u_t(\eta, W)]^2 \quad (3.20)$$

chamada soma de quadrados não-condicional, em que

$$[u_t(\eta, W)] = E(u_t | \eta, W). \quad (3.21)$$

Esta demonstração pode ser encontrada em Box, Jenkins e Reinsel (1994).

Então, como boa aproximação para EMV são os EMQ obtidos após a minimização de (3.20). Dado então, um  $\eta$  deseja-se calcular a SQ e, para tanto, deve-se então calcular as esperanças condicionais (3.21) por meio de (3.16). Porém um processo chamado “backforecasting” será necessário para inicializar o processo, obtendo-se assim, valores para  $[W_{-j}]$  e  $[u_{-j}]$ ,  $j = 1, 2, \dots$ , em fim, valores antes do início da série.

Suponha um modelo ARIMA usual

$$\phi(B) W_t = \theta(B) u_t. \quad (3.22)$$

Então, pode-se demonstrar, usando a função geradora de autocovariância, que a estrutura probabilística de  $W_1, \dots, W_n$  é igualmente explicada pelo modelo (3.22) ou pelo modelo

$$\phi(F) W_t = \theta(F) e_t, \quad (3.23)$$

onde  $F$  é o operador de transição para o futuro e  $e_t$  é um ruído branco com a mesma variância que  $u_t$ ; (3.23) é chamada forma “backward” do processo e fornece uma representação estacionária e invertível na qual  $W_t$  é expressa somente em termos de valores futuros de  $W_t$  e de  $e_t$ .

Desta maneira, o valor  $W_j$  tem a mesma relação probabilística com  $W_1, \dots, W_n$  que  $W_{n+j+1}$  tem com  $W_n, W_{n-1}, \dots, W_1$ , ou seja, fazer previsão antes que a série se inicie é equivalente a prever a série reversa.

c) Função de Verossimilhança Exata.

Ansley e Newbold (1980), mostram, via simulação de Monte Carlo, que a estimação por máxima verossimilhança exata é geralmente mais confiável em amostras de tamanho moderado. Para o modelo ARMA( $p, q$ ) a função de verossimilhança é bem complicada, Newbold (1974). Portanto, apresentamos aqui o desenvolvimento da função de verossimilhança exata para um processo AR(1). Considere  $W_t$  uma série estacionária e gerada por

$$\tilde{W}_t = \phi \tilde{W}_{t-1} + u_t, \quad |\phi| < 1 \quad (3.24)$$

onde  $\tilde{W}_t = W_t - \mu$  e  $u_t \sim N(0, \sigma_u^2)$  são independentes.

Pode-se, também, escrever a expressão (3.23) da forma

$$\tilde{W}_t = \sum_{j=0}^{\infty} \phi^j u_{t-j} \quad (3.25)$$

em que  $\tilde{W}_t \sim N\left(0, \sigma_u^2 (1 - \phi^2)^{-1}\right)$ .

Com as duas últimas expressões obtém-se

$$\begin{aligned} \tilde{W}_1 &= \sum_{j=0}^{\infty} \phi^j u_{1-j} = v_1, \\ \tilde{W}_t &= \phi \tilde{W}_{t-1} + u_t, \quad t = 2, \dots, n, \end{aligned} \quad (3.26)$$

de onde  $(v_1, u_2, \dots, u_n) \sim N(0, \Sigma)$  e  $\Sigma = \text{diag}\left((1 - \phi^2)^{-1} \sigma_u^2, \sigma_u^2, \dots, \sigma_u^2\right)$ .

Logo, a função conjunta de  $(v_1, u_2, \dots, u_n)$  e dada por

$$\begin{aligned} f(v_1, u_2, \dots, u_n) &= \left[ \frac{(1 - \phi^2)}{2\pi\sigma_u^2} \right]^{1/2} \exp \left[ -\frac{\tilde{W}_1^2 (1 - \phi^2)}{2\sigma_u^2} \right] \cdot \left[ \frac{1}{2\pi\sigma_u^2} \right]^{(n-1)/2} \\ &\quad \exp \left[ -\frac{1}{2\sigma_u^2} \sum_{t=2}^n \left( \tilde{W}_t - \phi \tilde{W}_{t-1} \right)^2 \right]. \end{aligned}$$

E, conseqüentemente, o logaritmo da função de verossimilhança exata de

$$\left(\tilde{W}_1, \dots, \tilde{W}_n\right)$$

é

$$l(W, \xi) = -\frac{n}{2} \ln 2\pi + \frac{1}{2} \ln (1 - \phi^2) - \frac{n}{2} \ln \sigma_u^2 - \frac{S(\eta)}{2\sigma_u^2}, \quad (3.27)$$

onde

$$S(\eta) = (W_1 - \mu)^2 (1 - \phi^2) + \sum_{t=2}^n [(W_t - \mu) - \phi(W_{t-1} - \mu)]^2 \quad (3.28)$$

Os estimadores de “máxima verossimilhança exata” são obtidos com a derivação de (3.27) com respeito a  $\mu$ ,  $\phi$  e  $\sigma_u^2$  e fazendo as equações obtidas iguais a zero.

Cordeiro e Klein (1994) mostram o viés do estimador de máxima verossimilhança para os modelos AR(1), AR(2), MA(1), ARMA(1, 1).

**3.1.3 Diagnóstico** Esta etapa busca decidir se o modelo estimado é estatisticamente adequado aos dados observados. Para tanto, o estudo dos erros contidos neste modelo (resíduos) é uma importante ferramenta, pois fornecem uma estimativa natural para os erros verdadeiros que são desconhecidos.

Suponha que o modelo ajustado seja

$$\phi(B)W_t = \theta(B)u_t$$

em que,

$$W_t = \Delta^d Y_t,$$

Os “erros verdadeiros”  $u_t = \theta^{-1}(B)\phi(B)W_t$  serão ruído branco, se este modelo for verdadeiro.

Para se verificar a adequação de um modelo se faz uso de testes de diagnóstico que, geralmente, são baseados nas autocorrelações estimadas dos resíduos.

Considere-se agora o teste de autocorrelação residual, os resíduos estimados

$$\hat{u}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)W_t \quad (3.29)$$

são obtidos a partir das estimativas de  $\theta$  e  $\phi$ . Se o modelo for adequado, os  $\hat{u}_t$  deverão ser aproximadamente não-correlacionados, ou seja, as autocorrelações, aqui chamadas  $\hat{r}_k$ , deverão ser iguais a zero. Em particular, sob a suposição de que o modelo é adequado, deve-se ter, aproximadamente,

$$\hat{r}_k \sim N\left(0, \frac{1}{n}\right). \quad (3.30)$$

As autocorrelações são calculadas da seguinte forma

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{u}_t \hat{u}_{t-k}}{\sum_{t=1}^n \hat{u}_t^2}. \quad (3.31)$$

Box, Jenkins e Reinsel (1994), mostram, para um modelo AR(1), que

$$\begin{aligned} Var[\hat{r}_k] &\cong \frac{1}{n} [1 - \phi^{2(k-1)} (1 - \phi^2)] \\ Cov[\hat{r}_i, \hat{r}_j] &\cong \frac{1}{n} \{ \delta_{ij} - \phi^{i+j-2} (1 - \phi^2) \}, \end{aligned} \quad (3.32)$$

onde  $\delta_{ij}$  é o delta de Kronecker. Então, para  $k$  grande ou moderado, os  $\hat{r}_k$  são não-correlacionados e a aproximação de sua variância é de  $1/n$ .

De qualquer modo, a comparação de  $\hat{r}_k$  com os limites  $\pm 2/\sqrt{n}$  sugere uma indicação geral de que  $u_t$  não se comporta segundo ruído branco, lembrando que estes limites subestimarão a significância de alguma discrepância quando o valor de  $k$  for pequeno.

Apresenta-se agora o teste proposto por Box e Pierce (1970), este teste não avalia se o comportamento das autocorrelações corresponde a um ruído branco, mas detecta se esses valores são muito altos. Porém, uma variação foi sugerida por Ljung e Box (1978).

A estatística dada por

$$Q(K) = n(n+2) \sum_{j=1}^K \frac{\hat{r}_j^2}{(n-j)} \quad (3.33)$$

terá, aproximadamente, uma distribuição  $\chi^2$  com  $K - p - q$  graus de liberdade, caso o modelo seja apropriado. Na prática basta usar as 10 ou 15 primeiras  $\hat{r}_k$ , a hipótese de ruído branco será rejeitada para valores grandes de  $K$ .

Já o teste da autocorrelação cruzada é baseado na correlação entre valores passados da série e o valor presente do ruído branco.

A função de correlação cruzada (fcc) deve ser investigada quando,  $u_t$  e  $Y_{t-k}$  são não-correlacionadas, ou seja,  $Cov\{a_t, Y_{t-k}\} = \gamma_{uz}(k) = 0$ ,  $k \geq 1$ . Tem-se,

$$s_k = \frac{\sum u_t (Y_{t-k} - \bar{Y})}{\left[\sum u_t^2 \sum (Y_t - \bar{Y})^2\right]^{1/2}}, \quad k = 1, 2, 3, \dots \quad (3.34)$$

Se o modelo que está sendo testado é um  $AR(p)$  e para um dado  $k_0$ ,  $s_k$  tem valor “alto”, isso pode significar que o modelo é inadequado. Neste caso, um novo termo auto-regressivo deve ser incluído no modelo. Como não conhecemos os verdadeiros  $u_t$ , consideramos os resíduos estimados  $\hat{u}_t$  e substituímos  $s_k$  por

$$\hat{s}_k = \frac{\sum_{t=k+1}^n \hat{u}_t (Y_{t-k} - \bar{Y})}{\left[\sum_{t=1}^n \hat{u}_t^2 \sum_{t=1}^n (Y_t - \bar{Y})^2\right]^{1/2}}, \quad k = 1, 2, \dots \quad (3.35)$$

Já a verificação das autocorrelações  $\hat{r}_k$  dá informação sobre novos termos de médias móveis a serem incluídos no modelo. Se a estimativa de uma autocorrelação  $|\hat{r}_f| > 2/\sqrt{n}$  então, um termo  $\theta_f u_{t-f}$  deve ser incluído no modelo.

Pode-se demonstrar que se  $Y_t$  for estacionário, com  $\rho_k$ , então

$$E(s_k) \rightarrow 0 \quad (3.36)$$

$$Var(s_k) \rightarrow \frac{1}{n-k} \cong \frac{1}{n}, \quad k > 0 \quad (3.37)$$

$$Cov\{s_k, s_{k+1}\} \rightarrow \frac{1}{n-k} \rho_l \cong \frac{1}{n} \rho_l, \quad k > 0 \quad (3.38)$$

As relações (3.36) e (3.37) mostram que  $\gamma_{uz}(k)$  é significativamente diferente de zero se  $|s_k| > 2/\sqrt{n}$ . Mas (3.36) e (3.37) não são válidas quando são usados os resíduos estimados  $\hat{u}_t$ .

Contudo, Hokstad (1983), mostra que  $1/n$  é um limite superior para  $Var(\hat{s}_k)$  quando  $Z_t \sim AR(p)$ . Portanto, o critério de julgar  $s_k$  significativamente quando  $|s_k| > 2/\sqrt{n}$  é razoável, exceto para  $k$  pequeno. Observe que para  $k$  pequeno o mesmo problema ocorre para o teste da fac dos resíduos.

Assim, podemos utilizar esses resultados para construir um modelo ARMA:

- (i) começando com um  $AR(p)$  de baixa ordem, podemos incluir novos termos auto-regressivos, analisando a fcc  $\hat{s}_k$ ;

- (ii) quando  $s_k$  não se apresentar mais significativa, a fac  $\hat{r}_k$  pode indicar termos de médias móveis a serem incluídos;
- (iii) se termos de médias móveis são incluídos num estágio anterior de identificação, a interpretação de valores grandes para  $|s_k|$  não é tão óbvia.

Usa-se os resíduos para modificar o modelo, suponha que os resíduos  $b_t$  do modelo ajustado

$$\phi_0(B) \Delta^{d_0} Y_t = \theta_0(B) b_t, \quad (3.39)$$

não sejam aleatórios. Usando o método de identificação da seção anterior, pode-se descrever os resíduos através do modelo

$$\bar{\phi}(B) \Delta^{\bar{d}} b_t = \bar{\theta}(B) u_t, \quad (3.40)$$

Substituindo (3.39) em (3.40) temos um novo modelo

$$\phi_0(B) \bar{\phi}(B) \Delta^{d_0} \Delta^{\bar{d}} Y_t = \theta_0(B) \bar{\theta}(B) u_t, \quad (3.41)$$

cujos resíduos são aleatórios, e que deverá ser ajustado aos dados. O ciclo de identificação, estimação e verificação deve ser continuado, até que um modelo satisfatório seja encontrado.

**3.1.4 Previsão** Esta etapa visa utilizar o modelo identificado e estimado para fazer previsões. O interesse agora é prever um valor  $Y_{t+h}$ ,  $h \geq 1$ , supondo que se tem observações  $\dots, Y_{t-2}, Y_{t-1}, Y_t$ , até o instante  $t$ , que é chamado origem das previsões.

É possível expressar a previsão  $\hat{Y}_t(h)$  da seguinte forma:

$$\begin{aligned} E(Y_{t+h}) &= \phi_1 E(Y_{t+h-1}) + \dots + \phi_{p+d} E(Y_{t+h-p-d}) \\ &\quad - \theta_1 E(u_{t+h-1}) - \dots - \theta_q E(u_{t+h-q}) + E(u_{t+h}), \end{aligned} \quad (3.42)$$

para  $h \geq 1$ . Aqui, devemos utilizar os seguintes fatos

$$\begin{aligned} E(Y_{t+h}) &= \hat{Y}_t(h), & h > 0 \\ E(Y_{t+h}) &= Y_{t+h}, & h \leq 0 \\ E(u_{t+h}) &= 0, & h > 0 \\ E(u_{t+h}) &= u_{t+h}, & h \leq 0 \end{aligned}$$

observações:



- i) os termos de médias móveis desaparecem para  $h > q$ ;
- ii) para calcular  $\hat{Y}_t(h)$  precisa-se de  $\hat{Y}_t(h-1), \hat{Y}_t(h-2), \dots$  que são calculados recursivamente;
- iii) existe uma certa aproximação quando utilizamos esse procedimento, pois, na prática, só se conhece um número finito de dados passados. Portanto, na realidade, a previsão é  $E(Y_{t+h} | Y_t, \dots, Y_1)$ , que é diferente da previsão ótima  $E(Y_{t+h} | Y_t, Y_{t-1} \dots)$ . Entretanto as duas fórmulas fornecem resultados semelhantes para um valor grande de  $t$ . Essa aproximação é introduzida quando atribuímos valores iniciais para calcular os valores da seqüência  $\{u_t\}$ ;
- iv) As previsões para um  $AR(p)$  são exatas, uma vez que pode ser demonstrado que, para esse modelo,

$$E(Y_{t+h} | Y_t, Y_{t-1} \dots) = E(Y_{t+k} | Y_t, \dots, Y_{t+1-p})$$

- v) As inovações  $u_t$  são obtidas recursivamente.

Para fazer previsões de valores futuros é necessário avaliar a precisão das previsões futuras oriundas do método selecionado. Considere uma série temporal com  $T$  observações e retire desta as  $n$  últimas observações, então gere previsões  $n$  passos à frente a partir do modelo estimado com  $T - n$  observações e obtenha os erros de previsão,

$$e_j = y_j - \hat{y}_j$$

onde  $\hat{y}_j$  denota a previsão de  $y_j$ .

A capacidade preditiva do modelo estimado pode ser avaliada utilizando as medidas de precisão a seguir, onde  $k$  é o número de passos à frente ( $k = T - n$ ):

- i) Erro quadrático médio (EQM):

$$EQM = \frac{1}{n} \sum_{j=k+1}^T e_j^2$$

- ii) Erro absoluto médio (EAM):

$$EAM = \frac{1}{n} \sum_{j=k+1}^T |e_j|$$

- iii) Erro médio percentual (EMP):

$$EMP = \left( \frac{1}{n} \sum_{j=k+1}^T \frac{e_j}{y_j} \right) \times 100$$

## 3.2 Considerações Gerais Sobre Econometria Financeira

**3.2.1 Retornos** A modelagem de preços de ativos é um dos objetivos de finanças, contudo, por estarem presos a alguns fatores como escala e não estacionariedade, será preferível trabalhar com retornos. Numa situação em que não houve dividendos pagos em um determinado período, podemos então considerar  $P_t$  o preço de um ativo num instante  $t$ . O “retorno líquido simples” de um ativo será dado pela variação de preços entre os instantes  $t - 1$  e  $t$  ( $\Delta P_t = P_t - P_{t-1}$ ) dividido pelo preço no instante  $t - 1$  ( $P_{t-1}$ ), também chamado de variação relativa dos preços. Ou seja,

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (3.43)$$

ou

$$R_t = \frac{P_t}{P_{t-1}} - 1$$

Assim, o que chamamos de “retorno bruto simples” será dado por

$$1 + R_t = \frac{P_t}{P_{t-1}}.$$

A “taxa de retorno” será o  $R_t$  expresso em termos percentuais, relativo ao período (um dia, um mês, etc). O logaritmo (na base  $e$ ) de  $P_t$  é definido como  $p_t = \log P_t$ , que será usado para determinar o “retorno composto continuamente”, “log retorno” ou ainda, simplesmente “retorno” dado por

$$r_t = \log \frac{P_t}{P_{t-1}} = \log(1 + R_t) = p_t - p_{t-1}. \quad (3.44)$$

Observe que em (3.44) pode-se verificar que  $R_t = e^{r_t} - 1$ . Note também que, para  $\mu$  pequeno  $\log(1 + \mu) \approx \mu$ , indicando que os retornos simples  $R_t$  e os log retornos  $r_t$  terão geralmente valores aproximados. O retorno simples de período  $k$  será dado por

$$R_t[k] = \frac{P_t - P_{t-k}}{P_{t-k}}. \quad (3.45)$$

Logo,

$$R_t[k] = \frac{P_t}{P_{t-k}} - 1, \quad (3.46)$$

sendo que seus  $k$  períodos estão compreendidos entre os instantes  $t - k$  e  $t$ , e pode ser chamado de retorno multiperíodos. Então, o log retorno de período  $k$  será expresso por

$$r_t[k] = \log \frac{P_t}{P_{t-k}} = \log(1 + R_t[k]) \quad (3.47)$$

No caso de haver pagamento de dividendos ( $D_t$ ) no período, então a expressão dos retornos será dada por

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1 \quad (3.48)$$

e

$$r_t = \log(1 + R_t) = \log(P_t + D_t) - \log P_{t-1} \quad (3.49)$$

**3.2.2 Distribuição de Retornos** Seja a série de retornos  $\{r_t, t = 1, \dots, T\}$ , onde os instantes de tempo são igualmente espaçados. Como visto na Seção 2, esta série pode ser vista como parte de um processo estocástico  $\{r_t, t \in Y\}$ , onde  $Y = \{\dots, -2, -1, 0, 1, 2, \dots\}$ . O processo, então, estará completamente especificado quando conhecermos as distribuições de dimensões finitas

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = P(r(t_1) \leq x_1, \dots, r(t_n) \leq x_n), \quad (3.50)$$

para quaisquer  $t_1, \dots, t_n$  e qualquer  $n \geq 1$ .

Contudo, conhecer essas distribuições não é fácil, ou impossível, e por isso na prática o processo é caracterizado por seus momentos até uma dada ordem. Por exemplo, a média, dada por

$$E(r_t) = \int_{-\infty}^{\infty} r dF(r; t) \quad (3.51)$$

e a função de autocovariância

$$\gamma(t_1, t_2) = E(r_{t_1}, r_{t_2}) - E(r_{t_1})E(r_{t_2}), \quad t_1, t_2 \in Z. \quad (3.52)$$

Como, também, se faz uso de algumas suposições como a estacionariedade, a ergodicidade ou a normalidade do processo. Apesar de os log-retornos serem estacionários (quando os preços não o são), a suposição de normalidade não é válida para estes.

Já para o caso das distribuições

$$F(r_{1,1}, \dots, r_{N,1}; \dots; r_{1,T}, \dots, r_{N,T}),$$

onde  $N$  é o número de ativos,  $r_{it}$  são os log-retornos e  $T$  o número de instantes de tempo, o estudo dessas distribuições é muito geral e ainda se faz necessário o uso de restrições. Como por exemplo, a suposição de independência e a de invariância temporal (mesma distribuição para todo instante de tempo).

Um processo é ergótico se em uma única trajetória do processo pudermos estabelecer características como a média, a função de autocovariância, etc., por exemplo, se a média amostral convergir em probabilidade para o verdadeiro valor da média do processo. Outra suposição muitas vezes usada é a de que a distribuição dos retornos se comportaria segundo uma distribuição estável.

**3.2.3 Assimetria e Curtose** Dada uma variável aleatória qualquer  $X$  onde  $E(X) = \mu$  e  $Var(X) = \sigma^2$ , então sua assimetria é expressa por

$$A(X) = E\left(\frac{(X - \mu)^3}{\sigma^3}\right) \quad (3.53)$$

e sua curtose definida por

$$K(X) = E\left(\frac{(X - \mu)^4}{\sigma^4}\right). \quad (3.54)$$

Em uma distribuição normal a assimetria é igual a zero  $A = 0$ , e sua curtose igual a 3,  $K = 3$ , então, se define como “excesso de curtose” a quantidade  $K(X) - 3$ . Quanto mais pesadas as caudas das distribuições maior o seu excesso de curtose, podendo mesmo ser infinito.

Então, com uma amostra  $X_1, \dots, X_T$ ,  $\hat{\mu} = \bar{X}$  como estimativa da média e  $\hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=1}^T (X_t - \hat{\mu})^2$  como estimativa da variância, se tem a assimetria e a curtose definidas como

$$\hat{A}(X) = \frac{1}{(T-1)\hat{\sigma}^3} \sum_{t=1}^T (X_t - \hat{\mu})^3 \quad (3.55)$$

e

$$\hat{K}(X) = \frac{1}{(T-1)\hat{\sigma}^4} \sum_{t=1}^T (X_t - \hat{\mu})^4 \quad (3.56)$$

respectivamente.

Se a condição de que os log-retornos  $r_t$  são independentes, têm mesma distribuição e são normais, é aceita, então os retornos brutos serão log-normais. Ou seja, se  $r_t \sim N(\mu, \sigma^2)$ , então, sabendo que  $r_t = \log(1 + R_t)$ , tem-se que  $1 + R_t$  será log-normal, com média

$$E(R_t) = e^{\mu + \sigma^2/2} - 1 \quad (3.57)$$

e variância

$$Var(R_t) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1). \quad (3.58)$$

Ao se fazer uma análise amostral dos retornos  $r_t$ , é possível observar que sua distribuição é aproximadamente simétrica, porém com excesso de curtose.

**3.2.4 Caracterização dos Retornos** É importante ressaltar algumas questões básicas em séries econômicas e financeiras. Como dito anteriormente, nestas, existem características em comum com outras séries temporais. Algumas destas características já foram citadas na seção anterior, como tendência e sazonalidade. Outra característica em comum é a não-linearidade. Uma série econômica ou financeira é não-linear quando responde de maneira diferente a choques negativos ou positivos. Por exemplo, um choque negativo pode causar maior volatilidade que choques positivos.

No entanto retornos financeiros raramente apresentam tendência ou sazonalidade. Segue-se algumas considerações a respeito de séries de retornos:

- 1) geralmente não são auto-correlacionados;
- 2) existe auto-correlação nos quadrados dos retornos, correlação de lag um, pequena, e depois uma queda lenta;
- 3) apresentam agrupamento de volatilidade ao longo do tempo;
- 4) apresentam caldas mais pesadas que a normal em sua distribuição (incondicional);
- 5) tem distribuição geralmente leptocúrtica, embora aproximadamente simétrica.

**3.2.5 Volatilidade** Considerada uma das mais importantes ferramentas em séries econômicas e financeiras, a volatilidade é de certa forma uma medida da velocidade do mercado. Teoricamente, a volatilidade é a variação de preço de uma mercadoria referente a um desvio padrão, expresso em porcentagem, em um intervalo de tempo. Para se falar em volatilidade é sempre bom explicitar o tipo de volatilidade a que se esta referindo, o período considerado e sua base (se diária, mensal, anual, etc.).

- a) volatilidade implícita.

É a volatilidade que o mercado, através dos preços das opções, implica para o ativo em questão. A volatilidade implícita se refere exclusivamente a opções de um ativo. Uma maneira de determinar seu valor é equacionar um preço de mercado observado com o preço modelado de uma opção, fórmula de Black-Scholes (1973). Para tanto se supõe normalidade dos preços e volatilidade constante.

b) volatilidade estatística.

É obtida a partir da modelagem direta da volatilidade da série de retornos usando alguma família de modelos como ARCH.

c) volatilidade histórica.

É calculada usando séries históricas de certo ativo, como por exemplo, por meio de uma média móvel de uma função dos últimos  $k$  retornos. Para tanto se pode considerar os quadrados dos retornos ou os valores absolutos dos retornos nesta média móvel. Em geral, a volatilidade para cada instante  $t$  se dá por

$$v_t = \left[ \frac{1}{k} \sum_{j=0}^{k-1} |r_{t-j}|^p \right]^{\frac{1}{p}},$$

onde  $p > 0$ . Os casos mais comuns são  $p = 2$  e  $p = 1$ .

A volatilidade é dita persistente se o retorno atual tem efeito grande na volatilidade esperada  $k$  períodos a frente, que é dada por

$$h_{t+k|t} = E_t [(r_{t+k} - \mu_{t+k})^2] \quad (3.59)$$

Seja

$$\theta_{t+k|t} = \frac{\partial h_{t+k|t}}{\partial r_t^2} \quad (3.60)$$

a “persistência futura”, que para muitos modelos de volatilidade decresce geometricamente. Já a persistência acumulada, que é uma medida relacionada à persistência futura, mede o impacto de um choque na variância média do retorno do ativo sobre o período  $[t, t + k]$ , dada por

$$\varphi_{t+k|t} = \frac{\partial \left( \frac{1}{k} (h_{t+k|t} + h_{t+k-1|t} + \dots + h_{t+1|t}) \right)}{\partial r_t^2}. \quad (3.61)$$

Quando se fala de reversão à média, o que se quer dizer é que existe um nível normal de volatilidade ao qual esta retorna após um período de alta volatilidade. Usando (3.60), é possível

afirmar que a persistência futura tende a zero em probabilidade, quando  $k \rightarrow \infty$ , para todo  $t$ . O que significa que previsões de volatilidade de longo prazo deveriam convergir para esse nível normal de volatilidade.

Os choques negativos e positivos afetam os retornos de ações de maneira assimétrica. Quando o preço de uma ação cai, aumenta a volatilidade dos retornos, isso se dá devido ao efeito alavancagem (mede o efeito potencial de cada observação sobre a estimativa dos parâmetros). Também existe a assimetria provocada pela chamada “aversão ao risco”, que ocorre quando a volatilidade aumenta, e então ocorre uma redução na demanda da ação o que faz baixar o preço e provocar mais volatilidade.

Veremos que os modelos ARCH e GARCH supõem que a variância condicional é afetada simetricamente por inovações negativas ou positivas.

A volatilidade dos retornos de um ativo em geral é afetada por outros fatores ou preços de mercado. É possível incluir valores de variáveis exógenas em modelos de volatilidade.

Como já visto, retornos têm estimativas de curtose significativamente maiores que três, o que indica a não-normalidade destes. Modelos como ARCH e GARCH absorvem estas características.

### 3.2.6 Modelos Auto-Regressivos com Heteroscedasticidade Condicional (ARCH)

Robert F. Engle (1982), introduziu os modelos ARCH, definidos por

$$X_t = \sqrt{h_t} \xi_t, \quad (3.62)$$

com

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2, \quad (3.63)$$

onde  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  e  $i > 0$ , com o objetivo de estimar a variância da inflação. Como visto, o retorno  $X_t$  é não-correlacionado, porém existe correlação entre a volatilidade e retornos passados por meio de uma função quadrática. No entanto, para aplicar tais modelos é necessário fazer algumas suposições, tais como:

- $\xi_t \sim N(0, 1)$ ,  $\xi_t \sim t_v$  ou uma distribuição que melhor descreva o excesso de curtose das séries financeiras.

- Os  $\xi_t$  devem ser independentes e identicamente distribuídos.

### I. Modelo ARCH(1)

Para efeito de estudo de algumas propriedades é útil considerar o caso em que  $r = 1$ , ou seja, o modelo ARCH é dado por

$$X_t = \sqrt{h_t} \xi_t, \quad (3.64)$$

onde

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2, \quad (3.65)$$

com  $\alpha_0 > 0, \alpha_1 \geq 0$ .

Neste caso,

$$\text{i) } E(X_t) = E\{E(X_t | F_{t-1})\} = 0;$$

$$\text{ii) } Var(X_t) = E(X_t^2) = E\{E(X_t^2 | F_{t-1})\} = E(\alpha_0 + \alpha_1 X_{t-1}^2), \text{ considerando o processo estacionário de segunda ordem, então } E(X_t^2) = E(X_{t-1}^2) = Var(X_t), \forall t. \text{ Logo,}$$

$$Var(X_t) = \frac{\alpha_0}{1 - \alpha_1}, \quad (3.66)$$

como  $Var(X_t) > 0$  então, deve-se ter  $0 \leq \alpha_1 < 1$ ;

$$\text{iii) } Cov(X_t, X_{t+k}) = E(X_t X_{t+k}) = E\{E(X_t X_{t+k} | F_{t+k-1})\} = E\{X_t E(X_{t+k} | F_{t+k-1})\} = E\left\{X_t E\left(\sqrt{h_{t+k}} \xi_{t+k} | F_{t+k-1}\right)\right\} = 0, \text{ para } k > 0, \text{ pois } X \text{ está em } F_{t+k-1} \text{ e } E(\xi_{t+k} | F_{t+k-1}) = 0. \text{ Logo } \gamma_X(k) = 0, k \geq 1;$$

são respectivamente média, variância e covariância incondicional da série. Portanto  $X_t$  é uma seqüência de variáveis não-correlacionadas, ou seja, ruído branco com  $E(X_t) = 0$  e  $Var(X_t) = \frac{\alpha_0}{1 - \alpha_1}$ . Ainda é necessário o cálculo do momento de quarta ordem para determinar a curtose que sabemos ser maior que três. Sejam os  $\xi_t$  normais (para efeito de cálculo), tem-se que

$$E(X_t^4 | F_{t-1}) = 3(\alpha_0 + \alpha_1 X_{t-1}^2)^2 = E(h_t^2 \xi_t^4 | F_{t-1})$$

lembrando que a curtose de uma variável com distribuição normal é igual a três, então  $E(\xi_t^4) = 3$ ; logo,

$$E(X_t^4) = 3E(\alpha_0 + \alpha_1 X_{t-1}^2)^2 = 3E(\alpha_0^2 + 2\alpha_1 \alpha_0 X_{t-1}^2 + \alpha_1^2 X_{t-1}^4) \quad (3.67)$$



O momento de quarta ordem, já que se admite um processo estacionário, será escrito como

$$\begin{aligned}
\mu_4 &= E(X_t^4) = 3(\alpha_0^2 + 2\alpha_1\alpha_0\text{Var}(X_t) + \alpha_1^2\mu_4) \\
&= 3\left(\alpha_0^2 + 2\alpha_1\alpha_0\left(\frac{\alpha_0}{1-\alpha_1}\right) + \alpha_1^2\mu_4\right) \\
&= 3\alpha_0^2\left(1 + \frac{2\alpha_1}{1-\alpha_1}\right) + 3\alpha_1^2\mu_4 \\
&= \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)} \tag{3.68}
\end{aligned}$$

No entanto,  $(1 - 3\alpha_1^2)$  deve ser maior que zero e, conseqüentemente,  $0 \leq \alpha_1^2 < 1/3$ , para que seja aceita a suposição de que momentos de quarta ordem são finitos e positivos. Isso leva a crer que ao se aumentar o número de restrições impostas ao processo, maior também será o número de restrições aos parâmetros do modelo.

No que se refere à curtose de  $X_t$  será, por fim, dada como

$$K = \frac{\mu_4}{[\text{Var}(X_t)]^2} = 3\frac{\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}\frac{(1-\alpha_1)^2}{\alpha_0^2} = 3\frac{(1-\alpha_1^2)}{(1-3\alpha_1^2)} \tag{3.69}$$

de onde se percebe então, que a curtose do processo de retornos é maior que três, indicando caudas mais pesadas que as da normal.

A volatilidade comporta-se de forma diferente a retornos positivos e negativos; isso é fácil perceber na prática. Porém, como o cálculo da volatilidade é feito através dos quadrados dos retornos, esta diferença não é percebida pelo modelo ARCH.

Pode-se achar  $X_t^2 - h_t$  usando (3.64) e (3.65)

$$\begin{aligned}
X_t^2 - (\alpha_0 + \alpha_1 X_{t-1}^2) &= h_t(\xi_t^2 - 1) \\
X_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + \nu_t \tag{3.70}
\end{aligned}$$

pois

$$\nu_t = h_t(\xi_t^2 - 1) = h_t(X - 1) \tag{3.71}$$

sendo  $X$  uma variável aleatória com distribuição  $\chi_1^2$ .

É possível observar que  $\{\nu_t\}$  é uma seqüência de variáveis aleatórias não-correlacionadas, de média zero, porém a variância não é constante. Nota-se, também que se tem um modelo AR(1) para  $X_t^2$ , mas os erros não são normalmente distribuídos.

A função de autocorrelação de  $X_t^2$  por (3.70) é escrita como

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

## II. Modelo ARCH( $r$ )

Em um modelo ARCH( $r$ )

$$X_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 + \nu_t \quad (3.72)$$

onde os  $\nu_t$  se comportam como no caso  $r = 1$  e ainda, é possível mostrar que os retornos também formam ruído branco com variância igual a

$$\text{Var}(X_t) = \frac{\alpha_0}{1 - \sum_{i=1}^r \alpha_i}.$$

Na etapa de identificação do modelo, considera-se que na construção de modelos ARCH, um procedimento importante é remover a correlação serial (caso exista) e, para tanto, é indicado tentar ajustar modelos ARIMA. No caso de existir correlação serial, tem-se

$$\varphi(B) X_t = \theta_0 + \theta(B) \xi_t,$$

onde  $\xi_t \sim ARCH(r)$ .

Sabemos então que  $X_t$  é não-correlacionada ou é o resíduo da aplicação de um modelo ARIMA à série original.

Outro interesse seria a verificação da presença de heteroscedasticidade condicional, para tanto se pode contar com dois testes:

- Teste de Box-Pierce-Ljung para  $X_t^2$ ;
- Teste de Multiplicadores de Lagrange (ML); Engle (1982).

Queremos testar a hipótese  $H_0 : \alpha_i = 0, \forall i = 1, \dots, r$ , na regressão

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

para  $t = r + 1, \dots, T$ .

Desde que  $X_t^2$  é um estimador não-viesado de  $h_t$  e que o valor atual do quadrado do retorno depende de valores passados dos quadrados dos retornos, nota-se então, que isto caracteriza um comportamento similar a de um modelo auto-regressivo. Desta forma, a função de auto correlação parcial de  $X_t^2$  pode ser usada para encontrar a ordem  $r$  do modelo ARCH( $r$ ).

Os estimadores dos parâmetros do modelo são obtidos pelo método de máxima verossimilhança condicional. Supondo normalidade dos resíduos  $\xi_t$  pode-se escrever a função de verossimilhança da forma

$$L(x_1, \dots, x_T | \alpha) = f(x_T | F_{T-1}) f(x_{T-1} | F_{T-2}) \cdots f(x_{r+1} | F_r) f(x_1, \dots, x_r | \alpha),$$

Engle (1982), mostra que para  $T$  grande  $f(x_1, \dots, x_r | \alpha)$  pode ser desprezado. Logo a função de verossimilhança a ser maximizada será

$$L(x_1, \dots, x_T | \alpha) = \prod_{t=r+1}^T (\sigma_t \sqrt{2\pi})^{-1} \exp \left\{ \frac{-x_t^2}{2\sigma_t^2} \right\} \quad (3.73)$$

onde  $\sigma_t^2 = h_t$  é a volatilidade obtida recursivamente.

Lembrando que em um modelo ARCH( $r$ ) os resíduos padronizados dados por

$$\tilde{X}_t = \frac{X_t}{\sqrt{\hat{h}_{t-1}(1)}} \quad (3.74)$$

são variáveis aleatórias independentes e identicamente distribuídas, com distribuição normal padrão ou  $t$ -student. Sendo assim, a estatística  $Q$  de Ljung-Box vista em (3.33) é uma maneira de verificar a adequação do modelo, a hipótese de normalidade será rejeitada para valores altos de  $Q$ .

A heteroscedasticidade condicional também deve ser verificada e, para tanto, pode-se aplicar o teste dos Multiplicadores de Lagrange para a seqüência  $\tilde{X}_t^2$ .

A volatilidade modelada utilizando o modelo ARCH( $r$ ) tem suas previsões obtidas recursivamente. Assim,

$$\hat{h}_t(1) = \alpha_0 + \alpha_1 X_t^2 + \cdots + \alpha_r X_{t-r}^2 \quad (3.75)$$

é a previsão de  $h_{t+1}$ , com origem fixada no instante em  $t$ . As previsões  $\ell$  passos à frente, com origem em  $t$ , são dadas por

$$\hat{h}_t(\ell) = \alpha_0 + \sum_{i=1}^r \alpha_i \hat{h}_t(\ell - i) \quad (3.76)$$

onde  $\hat{h}_t(\ell - i) = X_{t+\ell-i}^2$  se  $\ell - i \leq 0$ . K

**3.2.7 Modelos GARCH** Bollerslev (1986, 1987, 1988) sugere um modelo mais parcimonioso para descrever a volatilidade, o chamado modelo GARCH (“generalized ARCH”), uma generalização do modelo ARCH. Definido por

$$X_t = \sqrt{h_t} \xi_t \quad (3.77)$$

$$h_t = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j} \quad (3.78)$$

onde os  $\xi_t$  são independentes e identicamente distribuídos, i.i.d.  $(0, 1)$ ,  $\xi_t \sim N(0, 1)$ ,  $\xi_t \sim t_v$  ou uma distribuição de erro generalizada,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$ ,  $\sum_{i=1}^q (\alpha_i + \beta_j) < 1$ ,  $q = \max(r, s)$ .

A condição de positividade dos coeficientes é suficiente para que  $h_t > 0$  mas não é necessária.

Chamando  $\nu_t = X_t^2 - h_t$ , e substituindo em (3.38) chega-se

$$X_t^2 = \alpha_0 + \sum_{i=1}^q (\alpha_i + \beta_i) X_{t-i}^2 + \nu_t + \sum_{j=1}^s \beta_j \nu_{t-j}, \quad (3.79)$$

tem-se então, um modelo ARMA( $q, s$ ) para  $X_t^2$ , porém  $\nu_t$  geralmente não é um processo i.i.d.

Em longo prazo, a volatilidade convergirá para a média

$$E(X_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^q (\alpha_i + \beta_i)}.$$

I. Modelo GARCH(1, 1)

Ainda aqui se preservam as vantagens e desvantagens do modelo ARCH, onde volatilidades altas são precedidas de retornos ou volatilidade grandes, e observam-se grupos de volatilidade.

A volatilidade do modelo GARCH(1, 1) é dada por

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 h_{t-1} \quad (3.80)$$

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

A curtose para este modelo pode ser facilmente obtida por

$$K = \frac{E(X_t^4)}{[E(X_t^2)]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3, \quad (3.81)$$

o que mostra, considerando que o denominador seja positivo, que se  $X_t$  segue um modelo GARCH suas caudas serão mais longas que as da normal.

Em uma série real a identificação da ordem de um modelo GARCH é geralmente difícil, o que recomenda-se é que usando modelos de baixa ordem (no máximo 2, para  $r$  e  $s$ ) seja feita a identificação do modelo usando algum critério como *AIC*, *BIC* ou valores de assimetria e curtose.

Aqui também os estimadores dos parâmetros do modelo são obtidos pelo método de máxima verossimilhança condicional. Supondo normalidade dos resíduos  $\xi_t$  pode-se escrever a log-verossimilhança, condicional às primeiras  $r$  observações, da forma

$$l(x_{r+1}, \dots, x_T | \alpha, \beta, x_1, \dots, x_r) \propto -\frac{1}{2} \sum_{t=r+1}^T \ln(h_t) - \frac{1}{2} \sum_{t=r+1}^T \frac{x_t^2}{h_t} \quad (3.82)$$

Bollerslev (1986) utiliza em (3.82),  $h_t = \hat{\sigma}^2$ ,  $t = 1, \dots, s$ , com  $\hat{\sigma}^2 = \sum_{t=1}^T \frac{X_t^2}{T}$ .

As estimativas dos parâmetros são obtidas por meio de métodos numéricos de maximização.

As previsões da volatilidade, usando um modelo GARCH, podem ser calculadas de forma similar àquelas de modelo ARMA. As previsões, com origem  $t$ , considerando um modelo GARCH(1, 1) da forma (3.80), são dadas por

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

e para  $\ell > 1$ ,

$$\begin{aligned} \hat{h}_t(\ell) &= \alpha_0 + \alpha_1 \hat{X}_t^2(\ell - 1) + \beta_1 \hat{h}_t(\ell - 1), \\ &= \alpha_0 + \alpha_1 \hat{h}_t(\ell - 1) \hat{\xi}_t^2(\ell - 1) + \beta_1 \hat{h}_t(\ell - 1), \end{aligned}$$

pois  $X_t = \sqrt{h_t} \xi_t$ . Substituindo  $\hat{\xi}_t^2(\ell - 1)$  por  $E(\hat{\xi}_{t+\ell-1}^2) = 1$ , temos que

$$\hat{h}_t(\ell) = \alpha_0 + (\alpha_1 + \beta_1) \hat{h}_t(\ell - 1), \quad \ell > 1 \quad (3.83)$$

em muitas situações praticas podemos obter, por exemplo no GARCH(1, 1),  $\alpha_1 + \beta_1$  próximos de um. Se a soma desses parâmetros for um, teremos o modelo IGARCH (“integrated GARCH”).

Neste caso teremos

$$X_t = \sqrt{h_t} \xi_t$$

$$h_t = \alpha_0 + \beta_1 h_{t-1} + (1 - \beta_1) X_{t-1}^2,$$

com  $0 < \beta_1 < 1$  Morettin, P.A. e Toloi, C.M.C., (2004).

**3.2.8 Valor em Risco** O chamado risco de mercado ou valor em risco VaR é um tipo particular de risco, cujo cálculo envolve a volatilidade de um ativo financeiro ou de uma carteira de instrumentos financeiros. Nesta seção já foram discutidos vários aspectos relacionados com a volatilidade, aqui, se tratará somente com volatilidade estatística e serão usados os modelos ARCH e GARCH. Essa abordagem é denominada às vezes de abordagem econométrica.

As empresas em geral estão expostas a três classes de risco: operacional, estratégico e financeiro. Os riscos financeiros estão ligados às variações de variáveis financeiras (como juros e taxas de câmbio) que podem implicar em perdas financeiras. Os riscos financeiros ainda podem ser de vários tipos, como operacional, de crédito, de liquidez, legal e de mercado. O interesse aqui será somente nos riscos financeiros de mercado, que estão relacionados às perdas potenciais associadas ao comportamento do mercado. Portanto, o VaR será uma medida de risco financeiro de mercado e que dá uma medida do grau de incerteza sobre retornos líquidos futuros.

Informalmente, o VaR é uma medida da variação potencial máxima do valor de um ativo (ou carteira de ativos), sobre um período pré-fixado, com dada probabilidade. Ou seja, quanto se pode perder, com probabilidade  $p$ , sobre um horizonte  $h$  fixado. Do ponto de vista de uma empresa, o VaR é uma medida de perda associada a um evento extremo, sob condições normais de mercado.

Para que se possa definir formalmente o VaR é necessário definir os dois tipos de posições financeiras em uso.

Uma posição financeira comprada (ou “long”) significa possuir determinado ativo (ou carteira de ativos). Uma posição financeira vendida (ou “short”) envolve vender um ativo que não se possui. Esta operação é realizada alugando-se o ativo. Em data futura, o vendedor é obrigado a comprar exatamente o mesmo número de cotas ou ações alugadas (e não o valor em moeda),

para pagar o débito. Como o pagamento é em cotas ou ações, o vendedor ganha com a queda do preço do ativo.

Suponha que no instante  $t$  estejamos interessados em calcular o risco de uma posição financeira para o horizonte  $h > 0$ . Seja

$$\Delta P(h) = P(t+h) - P(t)$$

a variação do valor do ativo entre os dois instantes. A quantidade  $\Delta P(h)$  representa o lucro ou a perda da posição sobre o horizonte  $h$ . Chamemos de  $F_h(\cdot)$  a função de distribuição acumulada (f.d.a.) de  $\Delta P(h)$ .

**Definição 3.2.1** *Defini-se o VaR de uma posição comprada sobre o horizonte  $h$  com probabilidade  $p$ , por meio de*

$$p = P(\Delta P(h) \leq VaR) = F_h(VaR) \quad (3.84)$$

Observa-se que o VaR depende de  $p$  e de  $h$ , e que além disso, o valor em risco aumenta com  $p$  diminuindo ou com  $h$  aumentando.

Algumas observações são necessárias aqui:

- i) O VaR em (3.2.1) é dado em unidades monetárias (u.m.), por exemplo, reais. Lembrando que os retornos simples,  $R_t$ , são dados em porcentagem e que os log-retornos  $r_t$  são aproximadamente iguais a  $R_t$ , logo podemos supor que os  $r_t$  medem, aproximadamente, variações percentuais. Assim sendo, usaremos log-retornos no que segue.
- ii) A definição (3.2.1) mostra que o VaR é o  $p$ -quantil da distribuição  $F_h(\cdot)$ . Na prática, teremos que estimar este quantil usando, por exemplo, a distribuição empírica dos retornos.
- iii) O VaR calculado em (3.2.1) tem valor negativo, pois quem tem uma posição comprada sofre uma perda se  $\Delta P(h) < 0$ .
- iv) A quantia em u.m. no cálculo do VaR é obtida multiplicando o valor da posição financeira pelo VaR do log-retorno. A posição financeira em u.m. é usualmente o valor do ativo marcado pelo mercado (“mark-to-market”).

No caso de uma posição vendida, há perda se  $\Delta P(h) > 0$ , ou seja, o preço do ativo aumenta. Neste caso o VaR é definido por

$$p = P(\Delta P(h) \geq VaR) = 1 - F_h(VaR) \quad (3.85)$$

que tipicamente é positivo para  $p$  pequeno. O sinal positivo aqui indica perda. As definições (3.84) e (3.85) implicam que o VaR é calculado usando a cauda esquerda da distribuição  $F_h(\cdot)$ , para uma posição comprada e usando a cauda direita, para uma posição vendida. Também, a definição (3.85) aplica-se a uma posição vendida se usarmos a distribuição de  $-\Delta P(h)$ . Portanto, basta considerar o cálculo do VaR para uma dessas posições.

**3.2.9 VaR Usando Modelos ARIMA e GARCH** Lembremos que uma série de retornos é, em geral, não-correlacionada, mas dependente. Se este for o caso, a volatilidade é modelada por um dos modelos heteroscedásticos considerados nesta seção. Mas vimos também que algumas séries de retornos ainda exibem a presença de autocorrelação, havendo a necessidade de eliminá-la por meio do ajuste inicial de um modelo linear, por exemplo da família ARMA.

A estratégia é, portanto, modelar a média da série de retornos  $r_t$  por meio de um modelo ARMA e depois modelar os resíduos  $u_t$  deste modelo por um membro da família ARCH. Por exemplo, se escolhermos um modelo GARCH( $r, s$ ) para usar, teremos o modelo ARMA( $p, q$ )-GARCH( $r, s$ ):

$$r_t = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} + u_t - \sum_{j=1}^q \theta_j u_{t-j}, \quad (3.86)$$

$$u_t = \sigma_t \xi_t, \quad (3.87)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i u_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2. \quad (3.88)$$

Supondo  $\xi_t \sim N(0, 1)$ , resulta que

$$r_{t+1} | F_t \sim N(\hat{r}_t(1), \hat{\sigma}_t^2(1)),$$

onde  $\hat{r}_t(1)$  e  $\hat{\sigma}_t^2(1)$  são as previsões a um passo da média e variância usando (3.86) e (3.88), respectivamente. Neste caso, o VaR será dado por

$$VaR = \hat{r}_t(1) - Z_p \hat{\sigma}_t(1)$$

onde  $Z_p$  é o  $p$ -quantil da distribuição normal padrão. Para maiores detalhes sobre esta seção ver Morettin, P.A. e Toloi, C.M.C. (2004).



## **4 RESULTADOS E DISCUSSÃO**

### **ANÁLISE EXPLORATÓRIA, MODELAGEM E PREVISÃO**

Nesta seção serão feitas algumas considerações gerais a respeito da série de retornos do café Arábica. Será realizada também uma análise exploratória, modelagem e previsão da mesma. Para efeito de cálculo, quando não informado o nível de significância, deve ser considerado 5%.

#### **4.1 Índices diários**

O CEPEA (Centro de Estudos Avançados em Economia Aplicada) levanta diariamente os preços da saca de 60kg dos cafés Arábica e Conillon no mercado físico. Desde setembro de 1996, uma parceria entre o CEPEA e a BM&F (Bolsa de Mercadorias & Futuro) deu início ao levantamento dos preços do Arábica, com o objetivo de oferecer a esse setor um indicador de alta confiabilidade. Os indicadores CEPEA se tornaram um mecanismo de apoio para os participantes da cadeia cafeeira, constituindo-se num elemento referencial para análises de mercado, pois registram fielmente os movimentos de preços no mercado físico e, indiretamente, refletem as oscilações do mercado futuro. No caso do café Arábica, a comparação entre físico e futuro é favorecida pela similaridade do padrão do produto cotado pelo CEPEA e do negociado pela BM&F. Mensalmente, o CEPEA elabora informativos que trazem uma síntese dos mercados, acompanhada de tabelas de preços e gráficos das cotações físicas e futuras (Bolsa de Nova York, de Londres e BM&F). Estas análises conjunturais, para o café Arábica, são divulgadas no último dia útil de cada mês.

O Indicador do Café Arábica fornecido pelo CEPEA representa uma média ponderada dos preços do grão tipo 6, nas principais praças de comercialização, sendo o fator de ponderação proporcional à participação de cada região na produção deste café. Para isto, agentes atuantes no cerrado e sul de Minas Gerais, nas regiões Mogiana e Paulista de São Paulo e noroeste do Paraná são diariamente consultados.

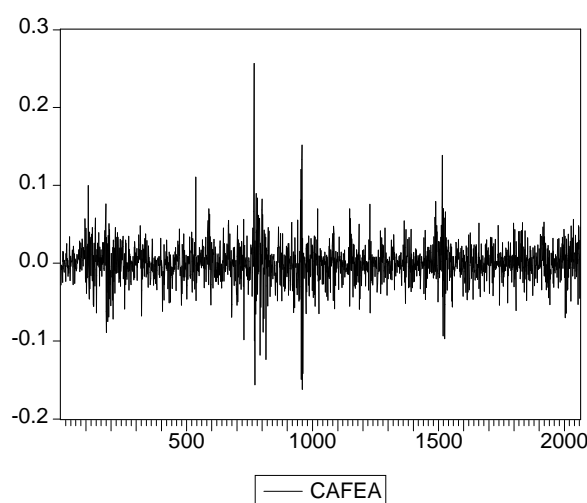
Após a coleta, os preços são depurados mediante tratamento estatístico e acrescidos dos custos de frete da região de origem para a praça de São Paulo. Assim, o Indicador se refere ao valor médio, à vista, do café posto na praça de São Paulo ([www.cepea.esalq.usp.br](http://www.cepea.esalq.usp.br)). O valor de

venda da saca de 60kg é cotado diariamente às 16:30h, ao todo são 2072 observações diárias que cobrem o período de 2 de setembro de 1996 a 23 de dezembro 2004. A série que compreende estas observações até 17 de setembro de 2004, omitindo os nove últimos registros, será denotada por *cafeA*. As nove observações do final da série, serão usadas para verificação das previsões.

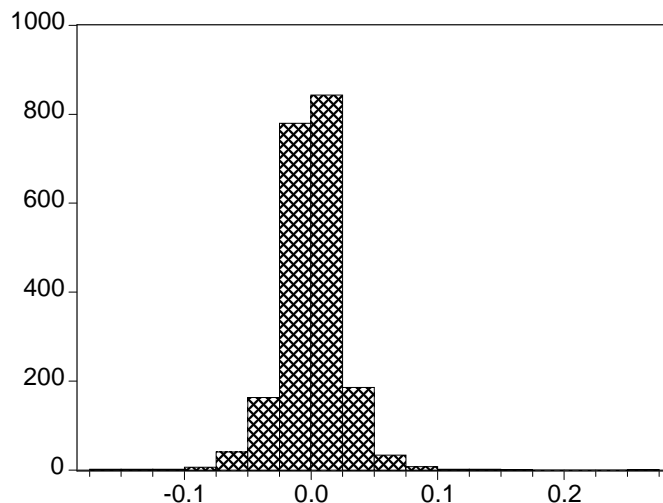
## 4.2 Análise Exploratória

Num primeiro passo obtivemos os retornos a partir da série dos preços da saca de 60kg do café arábica cotados diariamente pelo CEPEA. Ou seja, calculamos os retornos diários da série já denotada como *cafeA*.

A Figura 4.1 apresenta as observações da série ao longo do tempo. O gráfico temporal de uma série permite visualizar seu comportamento, sendo muitas vezes possível identificar, por exemplo, a existência de sazonalidade, tendência, estacionariedade, agrupamentos de volatilidade. Pode-se perceber nesta figura que a série de retornos diários do café apresenta média ao redor de zero, aparente estacionariedade e agrupamentos de volatilidade. A estacionariedade da série pode ser confirmada a partir do teste da raiz unitária (ou teste de Dickey-Fuller) cujo resultado para estatística  $\tau$  é  $-11,758$  e o  $p$ -valor é menor que  $0,01$ . A Figura 4.2 revela o histograma dos retornos, no qual se constata o excesso de curtose e pode se perceber a presença de valores afastados da parte central da distribuição (caudas longas).



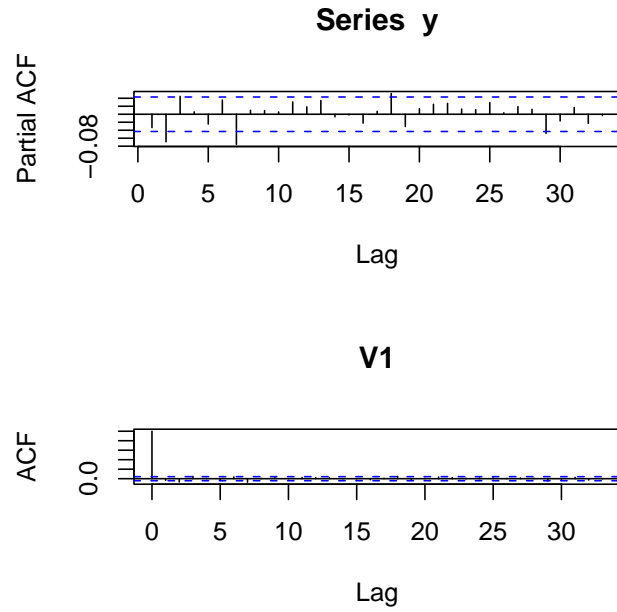
**Figura 4.1:** *Retornos diários da série cafeA.*



**Figura 4.2:** *Histograma dos retornos da série cafeA.*

### 4.3 Modelagem e Previsão

A Tabela 4.1 e a Figura 4.3 apresentam as funções de autocorrelações e autocorrelações parciais amostrais dos retornos da série *cafeA*. A análise dessas funções indica a existência de uma dependência linear entre as observações e que um modelo apropriado é um  $AR(7)$ , pois é o maior valor que a função assume e está fora do intervalo de confiança. Entretanto, verificando os valores de alguns critérios de ajustamento para modelos  $AR(j)$ ,  $j = 1, \dots, 10$  apresentados na Tabela 4.2, verificamos que o modelo escolhido pelo critério *AIC* comprova a adequação de um modelo  $AR(7)$  enquanto que o *BIC* escolhe o modelo  $AR(2)$ . Como visto anteriormente, o *BIC* impõe uma penalidade mais pesada que o *AIC*, por esta razão escolhe modelos mais parcimoniosos. Apesar disto, neste trabalho consideramos os dois modelos, o escolhido pelo *AIC* e o escolhido pelo *BIC*.



**Figura 4.3:** Gráfico das *fac* e *facp* dos retornos diários da série *cafeA*.

Amostra: 1 2063  
 Observações: 2063

Autocorrelações	Autocorrelações Parciais	FAC	FACP	Q	p-valor	
		1	-0,033	-0,033	2,3022	0,129
*	*	2	-0,068	-0,069	11,748	0,003
		3	0,049	0,045	16,759	0,001
		4	0,007	0,006	16,870	0,002
		5	-0,031	-0,025	18,873	0,002
		6	0,038	0,035	21,909	0,001
*	*	7	-0,073	-0,076	32,940	0,000
		8	0,007	0,010	33,041	0,000
		9	0,021	0,009	33,987	0,000
		10	-0,002	0,006	33,993	0,000
		11	0,025	0,031	35,337	0,000
		12	0,024	0,018	36,487	0,000
		13	0,024	0,034	37,637	0,000
		14	-0,003	-0,006	37,656	0,001
		15	-0,003	-0,002	37,674	0,001
		16	-0,022	-0,022	38,673	0,001
		17	0,010	0,007	38,875	0,002
		18	0,049	0,052	43,827	0,001
		19	-0,036	-0,031	46,574	0,000
		20	0,006	0,014	46,639	0,001
		21	0,036	0,025	49,344	0,000
		22	0,022	0,027	50,376	0,001
		23	0,009	0,013	50,542	0,001
		24	0,016	0,011	51,086	0,001

**Tabela 4.1:** *Fac* e *facp* dos retornos da série *cafeA*.

Modelo	AIC	BIC
AR(1)	-4,535872	-4,530410
<b>AR(2)</b>	-4,539182	<b>-4,530985</b>
AR(3)	-4,539836	-4,528903
AR(4)	-4,538416	-4,524745
AR(5)	-4,537604	-4,521192
AR(6)	-4,537422	-4,518266
<b>AR(7)</b>	<b>-4,542208</b>	-4,520307
AR(8)	-4,540953	-4,516305
AR(9)	-4,539573	-4,512175
AR(10)	-4,538163	-4,508014

**Tabela 4.2:** Valores do AIC e BIC de modelos  $AR(j)$ ,  $j = 1, \dots, 10$ , ajustados aos retornos diários da série *cafeA*.

A Tabela 4.3 apresenta o ajustamento do modelo  $AR(7)$ , de onde se verifica que vários coeficientes não são significativos, a um nível de 5%. Eliminando-os, obtemos o modelo

$$Y_t = -0,070195Y_{t-2} - 0,075242Y_{t-7} + X_t \quad (4.1)$$

com resultados apresentados na Tabela 4.4.

Variável Dependente: *cafeA*  
Método: Mínimos Quadrados  
Amostra(ajustada): 8 2063  
Observações: 2056 após o ajuste

Variável	Coefficiente	D.P.	t	p-valor
Constante	0,000437	0,000492	0,888503	0,3744
AR(1)	-0,029481	0,022044	-1,337383	0,1812
<b>AR(2)</b>	<b>-0,068245</b>	<b>0,022059</b>	<b>-3,093784</b>	<b>0,0020</b>
AR(3)	0,042312	0,022102	1,914419	0,0557
AR(4)	0,010249	0,022121	0,463342	0,6432
AR(5)	-0,028796	0,022111	-1,302332	0,1929
AR(6)	0,032571	0,022071	1,475736	0,1402
<b>AR(7)</b>	<b>-0,076088</b>	<b>0,022086</b>	<b>-3,445125</b>	<b>0,0006</b>
Log Verossimilhança	4677,390	Critério de Inf. Akaike		-4,542208
Estat.Durbin-Watson	1,996401	Critério de Schwarz		-4,520307

**Tabela 4.3:** Ajustamento de um modelo  $AR(7)$  aos retornos diários da série *cafeA*.

Variável Dependente: cafeA  
Método: Mínimos Quadrados  
Amostra(ajustada): 8 2063  
Observações: 2056 após o ajuste

Variável	Coefficiente	D.P.	t	p-valor
AR(2)	-0,070195	0,021992	-3,191851	0,0014
AR(7)	-0,075242	0,022015	-3,417749	0,0006
Log Verossimilhança	4671,605	Critério de Inf. Akaike		-4,542418
Estat.Durbin-Watson	2,055975	Critério de Schwarz		-4,536942

**Tabela 4.4:** Ajustamento do modelo (4.1) aos retornos diários da série *cafeA*.

A análise de resíduos do modelo (4.1) fornece  $Q(17) = 17,380$  com  $p$ -valor igual a 0,297 apresentado na Tabela 4.5, indicando que o modelo (4.1) eliminou a correlação serial da série de retornos diários.

Para verificar se os resíduos do modelo (4.1) apresentam heteroscedasticidade condicional examinamos as  $fac$  e  $facp$  dos quadrados dos resíduos, que estão apresentadas na Tabela 4.6. A  $fac$  indica a existência de heteroscedasticidade e a Figura 4.4, gráfico da  $facp$  dos quadrados dos resíduos, sugere um modelo ARCH(6). Para confirmar este fato, é verificado os valores de alguns critérios de seleção de modelos ARCH( $j$ ),  $j = 1, \dots, 6$ , apresentados na Tabela 4.7. A adequação de um modelo ARCH(6) ainda é comprovada considerando o critério  $AIC$ . Contudo, o modelo ARCH(6) apresenta o coeficiente  $\alpha_6$  não significativo, portanto recaímos no modelo ARCH(5) indicado pelo  $BIC$ , dado por

$$\begin{aligned}
Y_t &= \phi_2 Y_{t-2} + \phi_7 Y_{t-7} + X_t, \quad X_t = \sqrt{h_t} \xi_t \\
h_t &= \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2 + \alpha_3 X_{t-3}^2 + \alpha_4 X_{t-4}^2 + \alpha_5 X_{t-5}^2.
\end{aligned} \tag{4.2}$$

Amostra: 8 2063  
 Observações: 2056

Autocorrelações		Autocorrelações Parciais		FAC	FACP	Q	p-valor	
				1	-0,029	-0,029	1,7719	
				2	0,001	0,000	1,7737	
				3	0,046	0,046	6,1143	0,013
				4	0,011	0,014	6,3630	0,042
				5	-0,032	-0,031	8,4362	0,038
				6	0,039	0,035	11,519	0,021
				7	0,000	0,001	11,519	0,042
				8	0,007	0,010	11,630	0,071
				9	0,012	0,010	11,910	0,104
				10	0,005	0,004	11,964	0,153
				11	0,033	0,035	14,251	0,114
				12	0,018	0,018	14,919	0,135
				13	0,028	0,029	16,597	0,120
				14	-0,006	-0,008	16,670	0,162
				15	0,002	-0,002	16,676	0,215
				16	-0,016	-0,017	17,215	0,245
				<b>17</b>	<b>0,009</b>	<b>0,006</b>	<b>17,380</b>	<b>0,297</b>
				18	0,052	0,054	23,050	0,112
				19	-0,032	-0,031	25,138	0,092
				20	0,014	0,011	25,539	0,111
				21	0,036	0,030	28,178	0,080
				22	0,020	0,024	29,049	0,087
				23	0,011	0,014	29,311	0,107
				24	0,021	0,010	30,196	0,114

**Tabela 4.5:** *Fac e facp dos resíduos do modelo (4.1).*

Amostra: 8 2063  
 Observações: 2056

Autocorrelações		Autocorrelações Parciais		FAC	FACP	Q	p-valor	
**		**		1	0,242	0,242	120,94	
**		*		2	0,204	0,155	206,81	
**		**		3	0,303	0,244	396,16	0,000
*				4	0,163	0,037	450,92	0,000
*				5	0,142	0,033	492,77	0,000
*				6	0,087	-0,042	508,41	0,000
*				7	0,099	0,027	528,71	0,000
*				8	0,077	0,006	540,88	0,000
				9	0,051	0,007	546,35	0,000
				10	0,037	-0,014	549,24	0,000
				11	0,032	-0,002	551,35	0,000
				12	0,035	0,009	553,86	0,000
				13	0,047	0,033	558,48	0,000
				14	0,026	0,001	559,89	0,000
				15	0,013	-0,012	560,26	0,000
				16	0,026	0,001	561,62	0,000
				17	0,039	0,025	564,72	0,000
				18	0,012	-0,007	565,01	0,000
				19	0,013	-0,001	565,37	0,000
				20	0,010	-0,012	565,60	0,000
				21	0,045	0,042	569,75	0,000
				22	0,027	0,010	571,29	0,000
				23	0,030	0,019	573,20	0,000
*				24	0,081	0,056	586,95	0,000

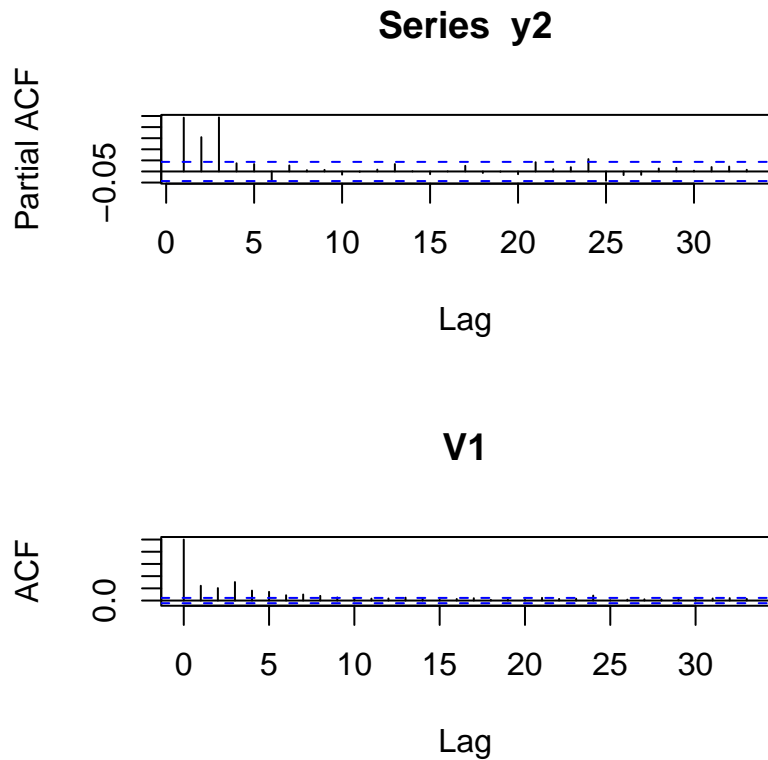
**Tabela 4.6:** *Fac e facp dos quadrados dos resíduos do modelo (4.1).*

Assumindo  $\xi \sim N(0, 1)$ , temos na Tabela 4.8 os resultados do ajustamento do modelo (4.2), que revelam a não significância do parâmetro  $\phi_7$  a um nível de 7%. Retirando-o do modelo, recaímos no modelo AR(2) escolhido anteriormente pelo *BIC* na Tabela 4.2. Procedemos então a um novo processo de estimação apresentado na Tabela 4.9 que fornece como modelo final

$$\begin{aligned}
 Y_t &= -0,036553Y_{t-2} + X_t, \quad X_t = \sqrt{h_t}\xi_t \\
 h_t &= 0,00023 + 0,1897X_{t-1}^2 + 0,0383X_{t-2}^2 + 0,0610X_{t-3}^2 + \\
 &\quad + 0,1708X_{t-4}^2 + 0,1626X_{t-5}^2
 \end{aligned} \tag{4.3}$$

com todos os coeficientes significativamente diferentes de zero, a um nível de 7%.





**Figura 4.4:** Gráfico das *fac* e *facp* dos quadrados dos resíduos do modelo (4.1).

Modelo	AIC	BIC
ARCH(1)	-4,673593	-4,662642
ARCH(2)	-4,713070	-4,699382
ARCH(3)	-4,736185	-4,719759
ARCH(4)	-4,768950	-4,749786
<b>ARCH(5)</b>	-4,788580	<b>-4,766679</b>
<b>ARCH(6)</b>	<b>-4,788806</b>	-4,764168

**Tabela 4.7:** Valores do *AIC* e *BIC* de modelos  $ARCH(j)$ ,  $j = 1, \dots, 6$ , ajustados aos resíduos do modelo (4.1).

A Tabela 4.10 apresenta as *fac* e *facp* dos resíduos padronizados, juntamente com as estatísticas de Ljung-Box, com  $p$ -valores maiores 0,05, indicando a adequação do modelo para modelar a dependência linear entre os retornos sucessivos a um nível de 5%.

A Tabela 4.11 apresenta as *fac* e *facp* dos quadrados dos resíduos padronizados; analisado as estatísticas de Ljung-Box podemos concluir que não existe heteroscedasticidade condicional nos resíduos do modelo. Para confirmar este resultado aplicamos o teste de Multiplicadores

de Lagrange (teste ML). Os resultados estão na Tabela 4.12 de onde extraímos os valores  $T = 15,55312$  com  $P = 0,412356$ , confirmando nossa hipótese.

Variável Dependente: cafeA  
Método: MV - ARCH  
Amostra(ajustada): 8 2063  
Observações: 2056 após o ajuste

	Coeficiente	D.P.	z	p-valor
AR(2)	-0,038965	0,019949	-1,953217	0,0508
AR(7)	-0,021946	0,015982	-1,373192	0,1697
Equação da Variância				
Contante	0,000235	1,14E-05	20,68862	0,0000
ARCH(1)	0,190120	0,016198	11,73720	0,0000
ARCH(2)	0,038386	0,013964	2,748896	0,0060
ARCH(3)	0,059323	0,017396	3,410159	0,0006
ARCH(4)	0,169035	0,024576	6,878021	0,0000
ARCH(5)	0,162110	0,026023	6,229573	0,0000
Log Verossimilhança	4930,660	Critério de Inf. Akaike		-4,788580
Estat.Durbin-Watson	2,061411	Critério de Schwarz		-4,766679

**Tabela 4.8:** Ajustamento do modelo (4.2) aos retornos diários da série cafeA.

Variável Dependente: cafeA  
Método: MV - ARCH  
Amostra(ajustada): 3 2063  
Observações: 2061 após o ajuste

	Coeficiente	D.P.	z	p-valor
AR(2)	-0,036555	0,019816	-1,844717	0,0651
Equação da Variância				
Constante	0,000234	1,14E-05	20,48643	0,0000
ARCH(1)	0,189647	0,015576	12,17562	0,0000
ARCH(2)	0,038254	0,013874	2,757223	0,0058
ARCH(3)	0,060968	0,017339	3,516193	0,0004
ARCH(4)	0,170793	0,024279	7,034500	0,0000
ARCH(5)	0,162562	0,026127	6,221896	0,0000
Log Verossimilhança	4944,322	Critério de Inf. Akaike		-4,791191
Estat.Durbin-Watson	2,063663	Critério de Schwarz		-4,772066

**Tabela 4.9:** Ajustamento do modelo AR(2)-ARCH(5) aos retornos diários da série cafeA.

Amostra: 3 2063  
 Observações: 2061

Autocorrelações	Autocorrelações Parciais	FAC	FACP	Q	p-valor	
		1	0,029	0,029	1,7848	
		2	-0,015	-0,015	2,2251	0,136
		3	0,062	0,063	10,071	0,007
		4	0,021	0,017	10,958	0,012
		5	0,000	0,001	10,958	0,027
		6	0,028	0,024	12,534	0,028
		7	-0,035	-0,039	15,100	0,019
		8	-0,001	0,002	15,101	0,035
		9	0,024	0,020	16,290	0,038
		10	0,004	0,007	16,331	0,060
		11	0,021	0,023	17,210	0,070
		12	0,016	0,011	17,717	0,088
		13	0,014	0,014	18,114	0,112
		14	-0,002	-0,007	18,126	0,153
		15	0,000	-0,003	18,127	0,201
		16	0,006	0,005	18,196	0,252
		17	-0,010	-0,011	18,406	0,301
		18	0,045	0,047	22,618	0,162
		19	-0,015	-0,019	23,089	0,187
		20	0,013	0,017	23,424	0,219
		21	0,018	0,011	24,129	0,237
		22	0,008	0,006	24,277	0,280
		23	-0,012	-0,012	24,574	0,318
		24	0,033	0,028	26,785	0,265

**Tabela 4.10:** *Fac e facp dos resíduos padronizados do modelo (4.3).*

Amostra: 3 2063  
 Observações: 2061

Autocorrelações		Autocorrelações Parciais		FAC	FACP	Q	p-valor	
				1	0,019	0,019	0,7791	
				2	-0,019	-0,019	1,5141	0,219
				3	0,010	0,010	1,7101	0,425
				4	-0,010	-0,011	1,9275	0,588
				5	-0,001	0,000	1,9281	0,749
				6	-0,029	-0,030	3,7132	0,591
				7	0,031	0,032	5,6962	0,458
				8	0,015	0,012	6,1608	0,521
				9	0,008	0,009	6,2865	0,615
				10	0,000	-0,001	6,2869	0,711
				11	0,003	0,004	6,3029	0,789
				12	0,012	0,011	6,5837	0,832
				13	0,061	0,063	14,262	0,284
				14	-0,011	-0,014	14,524	0,338
				15	-0,016	-0,014	15,077	0,373
				16	-0,001	-0,002	15,078	0,446
				17	0,009	0,011	15,258	0,506
				18	-0,018	-0,018	15,911	0,530
				19	0,004	0,007	15,938	0,597
				20	0,021	0,015	16,843	0,601
				21	0,012	0,011	17,162	0,642
				22	0,023	0,023	18,283	0,631
				23	0,031	0,032	20,306	0,564
*			*	24	0,090	0,089	37,351	0,030

**Tabela 4.11:** *Fac e facp dos quadrados dos resíduos padronizados do modelo (4.3).*

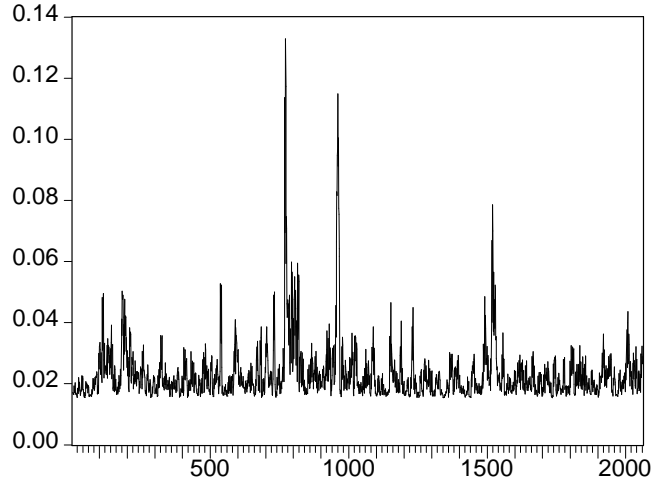
Teste ARCH :				
F-Estatística	1,036646	Probabilidade	0,413265	
R-quadrado	15,55312	Probabilidade	0,412356	
Variável Dependente: STD_RESID <sup>2</sup>				
Método: Mínimos Quadrados				
Amostra(ajustada): 18 2063				
Observações: 2046 após o ajuste				
Variável	Coefficiente	Desvio Padr.	t-Estatística	Prob.
Constante	0,935420	0,096926	9,650840	0,0000
STD_RESID <sup>2</sup> (-1)	0,019941	0,022195	0,898444	0,3691
STD_RESID <sup>2</sup> (-2)	-0,019594	0,022199	-0,882649	0,3775
STD_RESID <sup>2</sup> (-3)	0,011412	0,022158	0,515045	0,6066
STD_RESID <sup>2</sup> (-4)	-0,013086	0,022158	-0,590567	0,5549
STD_RESID <sup>2</sup> (-5)	0,000222	0,022160	0,010029	0,9920
STD_RESID <sup>2</sup> (-6)	-0,032297	0,022160	-1,457465	0,1451
STD_RESID <sup>2</sup> (-7)	0,034593	0,022185	1,559292	0,1191
STD_RESID <sup>2</sup> (-8)	0,012021	0,022198	0,541536	0,5882
STD_RESID <sup>2</sup> (-9)	0,009358	0,022267	0,420247	0,6743
STD_RESID <sup>2</sup> (-10)	-0,001666	0,022259	-0,074835	0,9404
STD_RESID <sup>2</sup> (-11)	0,004426	0,022257	0,198879	0,8424
STD_RESID <sup>2</sup> (-12)	0,009014	0,022255	0,405030	0,6855
STD_RESID <sup>2</sup> (-13)	0,063142	0,022254	2,837277	0,0046
STD_RESID <sup>2</sup> (-14)	-0,013570	0,022294	-0,608687	0,5428
STD_RESID <sup>2</sup> (-15)	-0,013886	0,022290	-0,622972	0,5334
Log Verossimilhança	-4485,719	Critério de Inf. Akaike	4,400508	
Estat.Durbin-Watson	1,999868	Critério de Schwarz	4,444485	

**Tabela 4.12:** Aplicação do teste ML à seqüência dos quadrados dos resíduos padronizados do modelo (4.3).

A Figura 4.5 apresenta a estimativa do desvio padrão condicional ( $\sqrt{h_t}$ ). Analisando esta figura observamos que os seis maiores picos (em ordem cronológica) na volatilidade estimada, correspondem a:

- 1°)  $t \approx 109$  – fevereiro de 1997;
- 2°)  $t \approx 179$  – maio de 1997;
- 3°)  $t \approx 536$  – novembro de 1998;
- 4°)  $t \approx 768$  – outubro de 1999;
- 5°)  $t \approx 958$  – julho de 2000;
- 6°)  $t \approx 1516$  – outubro de 2002.

Todos correspondem a oscilações na bolsa de New York, segundo as informações do CEPEA.



**Figura 4.5:** *Estimativa fornecida pelo modelo (4.3) para o desvio padrão condicional dos retornos diários da série cafeA.*

Vamos agora analisar novamente os retornos diários da série *cafeA*. A idéia é verificar a adequação de um modelo GARCH, talvez mais parcimonioso.

A identificação da ordem de um modelo GARCH é geralmente difícil. Na prática é recomendado que se use modelos de baixa ordem e se escolha o modelo com base em critérios de seleção como o *AIC* e *BIC*. Em nosso caso, a ordem do modelo é encontrada com a verificação dos valores destes critérios em modelos  $\text{GARCH}(j, k)$ ,  $j, k = 1, \dots, 3$  apresentados na Tabela 4.13. Assim, o modelo escolhido com base em ambos os critérios é  $\text{GARCH}(3, 3)$  dado por

$$Y_t = \phi_2 Y_{t-2} + X_t, \quad X_t = \sqrt{h_t} \xi_t$$

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2 + \alpha_3 X_{t-3}^2 + \beta_1 h_{t-1} + \beta_2 h_{t-2} + \beta_3 h_{t-3}. \quad (4.4)$$

Supondo  $\xi \sim N(0, 1)$  a Tabela 4.14 apresenta o ajustamento do modelo (4.4) onde comprovamos que todos os parâmetros são significantes. Então, o modelo proposto é

$$Y_t = -0,0398 Y_{t-2} + X_t, \quad X_t = \sqrt{h_t} \xi_t$$

$$h_t = 0,000006 + 0,187 X_{t-1}^2 - 0,145 X_{t-2}^2 + 0,197 X_{t-3}^2 + 0,838 h_{t-1} -$$

$$-0,771 h_{t-2} + 0,594 h_{t-3} \quad (4.5)$$

Modelo	AIC	BIC
GARCH(1,1)	-4,778188	-4,767260
GARCH (1,2)	-4,784693	-4,771033
GARCH (1,3)	-4,790748	-4,774355
GARCH (2,1)	-4,782709	-4,769049
GARCH (2,2)	-4,776106	-4,759713
GARCH (2,3)	-4,785139	-4,766014
GARCH(3,1)	-4,789185	-4,772793
GARCH (3,2)	-4,789866	-4,770741
<b>GARCH (3,3)</b>	<b>-4,806061</b>	<b>-4,784204</b>

**Tabela 4.13:** Valores do AIC e BIC de modelos GARCH( $j, k$ ),  $j, k = 1, \dots, 3$ , ajustados aos resíduos do modelo AR(2).

Variável Dependente: cafeA  
Método: MV - ARCH  
Amostra(ajustada): 3 2063  
Observações: 2061 após o ajuste

	Coefficiente	D.P.	z	p-valor
AR(2)	-0,039805	0,020926	-1,902148	0,0572
Equação da Variância				
C	6,28E-05	8,76E-06	7,161414	0,0000
ARCH(1)	0,186803	0,010882	17,16686	0,0000
ARCH(2)	-0,145232	0,013603	-10,67642	0,0000
ARCH(3)	0,196597	0,016870	11,65333	0,0000
GARCH(1)	0,837843	0,034631	24,19365	0,0000
GARCH(2)	-0,770798	0,037591	-20,50500	0,0000
GARCH(3)	0,593625	0,028106	21,12070	0,0000
Log Verossimilhança	4960,646	Critério de Inf. Akaike		-4,806061
Estat.Durbin-Watson	2,063571	Critério de Schwarz		-4,784204

**Tabela 4.14:** Ajustamento do modelo AR(2)–GARCH(3,3) aos retornos diários da série cafeA.

Para verificar a adequação do modelo aplicamos o teste de Ljung-Box aos resíduos padronizados apresentados na Tabela 4.15 de onde obtemos  $Q(17) = 17,991$  com  $p = 0,324$  indicando que o modelo é adequado para modelar a dependência linear entre os retornos. A Tabela 4.16 expõe o quadrado dos resíduos padronizados onde observamos  $Q(23) = 14,013$  com valor de  $p$  igual a 0,901. Além disso, a Tabela 4.17 contendo os resultados da aplicação do teste ML, fornece os valores  $T = 9,721295$  com  $p = 0,836895$  o que confirma que o modelo (4.5) também é adequado para ajustar os retornos diários da série cafeA.

Amostra: 3 2063  
 Observações: 2061

Autocorrelações		Autocorrelações Parciais		FAC	FACP	Q	p-valor	
				1	0,029	0,029	1,7889	
				2	-0,009	-0,010	1,9520	0,162
				3	0,062	0,063	9,9428	0,007
				4	0,024	0,021	11,179	0,011
				5	-0,004	-0,004	11,206	0,024
				6	0,030	0,027	13,049	0,023
				7	-0,031	-0,035	14,989	0,020
				8	-0,001	0,002	14,990	0,036
				9	0,024	0,020	16,171	0,040
				10	0,007	0,008	16,268	0,061
				11	0,022	0,024	17,284	0,068
				12	0,011	0,006	17,530	0,093
				13	0,009	0,009	17,697	0,125
				14	-0,004	-0,008	17,727	0,168
				15	-0,005	-0,007	17,772	0,217
				16	0,006	0,006	17,858	0,270
				<b>17</b>	<b>-0,008</b>	<b>-0,009</b>	<b>17,991</b>	<b>0,324</b>
				18	0,041	0,044	21,516	0,204
				19	-0,014	-0,017	21,906	0,236
				20	0,024	0,026	23,058	0,235
				21	0,016	0,009	23,595	0,261
				22	0,015	0,012	24,042	0,291
				23	-0,018	-0,020	24,745	0,309
				24	0,031	0,026	26,701	0,269

**Tabela 4.15:** *Fac e facp dos resíduos padronizados do modelo (4.5).*



Amostra: 3 2063  
 Observações: 2061

Autocorrelações		Autocorrelações Parciais		FAC	FACP	Q	p-valor	
				1	0,030	0,030	1,8669	
				2	-0,015	-0,016	2,3073	0,129
				3	0,007	0,008	2,4199	0,298
				4	0,002	0,001	2,4252	0,489
				5	0,032	0,032	4,5565	0,336
				6	-0,010	-0,012	4,7675	0,445
				7	0,014	0,015	5,1488	0,525
				8	-0,010	-0,012	5,3494	0,617
				9	-0,019	-0,018	6,1125	0,635
				10	-0,006	-0,007	6,1899	0,721
				11	0,002	0,003	6,2000	0,798
				12	-0,018	-0,019	6,8395	0,812
				13	0,028	0,031	8,5155	0,744
				14	-0,006	-0,008	8,5882	0,803
				15	-0,018	-0,016	9,2448	0,815
				16	-0,027	-0,026	10,735	0,771
				17	-0,010	-0,007	10,925	0,814
				18	-0,024	-0,027	12,112	0,793
				19	-0,019	-0,016	12,860	0,800
				20	-0,008	-0,008	12,988	0,839
				21	-0,013	-0,012	13,368	0,861
				22	0,015	0,017	13,849	0,876
				23	0,009	0,010	14,013	0,901
				24	0,054	0,053	19,997	0,642

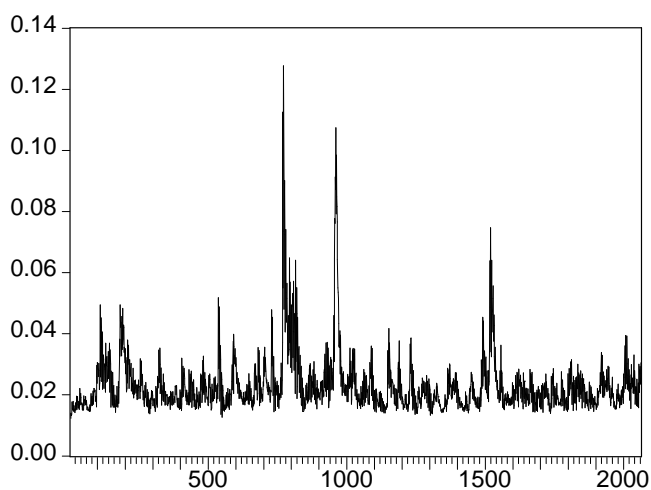
**Tabela 4.16:** *Fac e facp dos quadrados dos resíduos padronizados do modelo (4.5).*

A estimativa do desvio padrão condicional para o modelo (4.5),  $\sqrt{h_t}$ , encontra-se na Figura 4.6 da qual vemos que também aqui houve influência, das crises citadas para o modelo (4.3). Contudo, a estimativa da volatilidade fornecida pelo modelo AR(2)–GARCH(3, 3) é um pouco mais suave que a fornecida pelo modelo AR(2)–ARCH(5).

Teste ARCH :

F-Estatística	0,646088	Probabilidade	0,838223
R-quadrado	9,721295	Probabilidade	0,836895
Variável Dependente: STD_RESID <sup>2</sup>			
Método: Mínimos Quadrados			
Amostra(ajustada): 18 2063			
Observações: 2046 após o ajuste			
Variável	Coefficiente	D.P.	t p-valor
Constante	0,990467	0,095799	10,33897 0,0000
STD_RESID <sup>2</sup> (-1)	0,031411	0,022194	1,415301 0,1571
STD_RESID <sup>2</sup> (-2)	-0,016580	0,022209	-0,746536 0,4554
STD_RESID <sup>2</sup> (-3)	0,008678	0,022201	0,390895 0,6959
STD_RESID <sup>2</sup> (-4)	4,31E-05	0,022198	0,001944 0,9984
STD_RESID <sup>2</sup> (-5)	0,033404	0,022198	1,504806 0,1325
STD_RESID <sup>2</sup> (-6)	-0,014099	0,022210	-0,634815 0,5256
STD_RESID <sup>2</sup> (-7)	0,016129	0,022229	0,725590 0,4682
STD_RESID <sup>2</sup> (-8)	-0,012280	0,022231	-0,552402 0,5807
STD_RESID <sup>2</sup> (-9)	-0,017660	0,022362	-0,789745 0,4298
STD_RESID <sup>2</sup> (-10)	-0,006966	0,022365	-0,311463 0,7555
STD_RESID <sup>2</sup> (-11)	0,004003	0,022346	0,179142 0,8578
STD_RESID <sup>2</sup> (-12)	-0,020158	0,022345	-0,902135 0,3671
STD_RESID <sup>2</sup> (-13)	0,031218	0,022349	1,396851 0,1626
STD_RESID <sup>2</sup> (-14)	-0,007371	0,022356	-0,329724 0,7416
STD_RESID <sup>2</sup> (-15)	-0,015927	0,022343	-0,712847 0,4760
Log Verossimilhança	-4439,780	Critério de Inf. Akaike	4,355601
Estat.Durbin-Watson	2,000707	Critério de Schwarz	4,399578

**Tabela 4.17:** Aplicação do teste ML à seqüência dos quadrados dos resíduos padronizados do modelo (4.5).



**Figura 4.6:** Estimativa fornecida pelo modelo (4.5) para o desvio padrão condicional dos retornos diários da série *cafeA*.

Com o objetivo de comparar os modelos (4.3) e (4.5), apresentamos os valores de alguns critérios de ajustamento de cada um dos modelos na tabela 4.18. Analisando os resultados desta tabela podemos dizer que o modelo  $AR(2) - GARCH(3,3)$ , ajusta melhor a série pois tem menor  $AIC$  e  $BIC$ .

Modelo	AIC	BIC
AR(2)-ARCH(5,0)	-4,791191	-4,772066
<b>AR(2)-GARCH(3,3)</b>	<b>-4,806061</b>	<b>-4,784204</b>

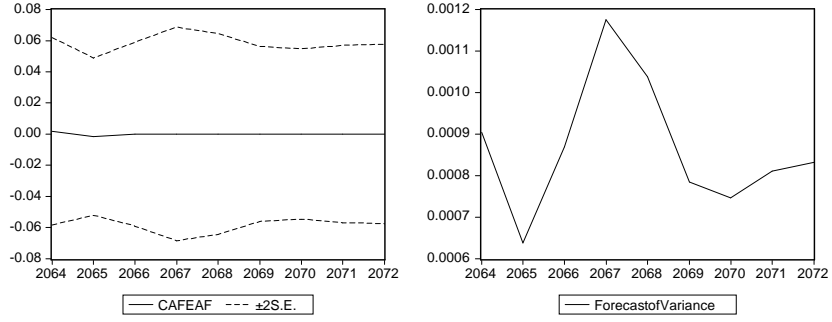
**Tabela 4.18:** Valores do  $AIC$  e  $BIC$  dos modelos (4.3) e (4.5), ajustados aos retornos diários da série *cafeA*.

Na Tabela 4.19 encontra-se o erro quadrático médio e o erro absoluto médio para previsões 1; 3; 6 e 9 passos à frente. Todos para as previsões obtidas para a volatilidade dos retornos diários da série *cafeA* pelos modelos (4.3) e (4.5). Verificamos novamente que o modelo (4.3) foi superior ao (4.5) pois para o número de passos à frente maior ou igual a três, os EQM e EMA para o modelo (4.3) foram menores que para o (4.5); se o interesse recair sobre a previsão um passo à frente o modelo (4.5) é o mais apropriado.

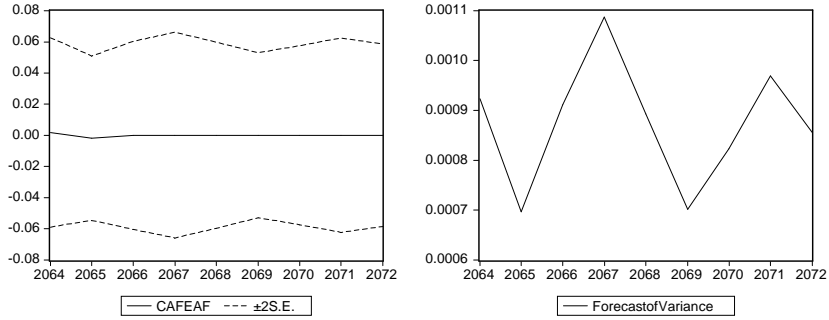
AR(2)-ARCH(5)	Passos à frente			
Medidas de precisão	1	3	6	9
EQM	0,002255	<b>0,014290</b>	<b>0,015961</b>	<b>0,014924</b>
EMA	0,002255	<b>0,010705</b>	<b>0,014072</b>	<b>0,013473</b>
AR(2)-GARCH(3,3)	Passos à frente			
Medidas de precisão	1	3	6	9
EQM	<b>0,002108</b>	0,014310	0,015972	0,014932
EMA	<b>0,002108</b>	0,010710	0,014077	0,013476

**Tabela 4.19:** Medidas de precisão para previsões 1, 3, 6 e 9 passos à frente, para os retornos diários da série *cafeA*.

A Figura 4.7 mostra a representação gráfica do intervalo de confiança para as previsões nove passos à frente do modelo (4.3) como também o comportamento das previsões da volatilidade. Percebemos então, que este modelo prevê volatilidade alta e por conta disto um largo intervalo de confiança, porém com estreitamento para previsões mais distantes. Já a Figura 4.8 contém a mesma representação gráfica para o caso do modelo (4.5). Nesta figura podemos observar um comportamento mais suave para a volatilidade e um intervalo de confiança mais estreito no todo.



**Figura 4.7:** Representação gráfica do intervalo de confiança para as previsões do modelo  $AR(2)-ARCH(5)$  e das previsões da volatilidade.



**Figura 4.8:** Representação gráfica do intervalo de confiança para as previsões do modelo  $AR(2)-GARCH(3,3)$  e das previsões da volatilidade.

O valor em risco, portanto, deve ser calculado para ambos os modelos, visto que os dois modelos são apropriados, dependendo da distância das previsões. Lembrando que

$$VaR = \hat{r}_t(1) - Z_p \hat{\sigma}_t(1),$$

para a previsão um passo à frente, então no modelo (4.3) temos

$$r_t = -0,036553r_{t-2} + u_t, \quad u_t = h_t \xi_t$$

$$\sigma_t^2 = 0,00023 + 0,1897u_{t-1}^2 + 0,0383u_{t-2}^2 + 0,0610u_{t-3}^2 + 0,1708u_{t-4}^2 + 0,1626u_{t-5}^2.$$

Dos dados temos  $r_{2063} = 0,0465$  e  $r_{2062} = -0,0451$ , logo a previsão da série e da volatilidade um passo à frente são  $\hat{r}_{2063}(1) = 0,001649$  e  $\hat{\sigma}_{2063}^2(1) = 0,000904$ , respectivamente.

Então,  $VaR = 0,001649 - 1,96 \times 0,030066 = -0,057280$  com 95% de confiança (quantil 0,05 da distribuição normal padrão).

A Tabela 4.20 contém os *Valores em Risco* para o horizonte  $h = 1, \dots, 9$  com base nos modelos (4.3) e (4.5). Os *VaR's* calculados, independente do modelo tido como base, mostram valores negativos que indicam perda, no caso de uma posição comprada. Estes valores mostram ainda que estas perdas podem variar de 4,376% a 6,716%.

	ARCH(5)	GARCH(3,3)
VaR[1]	-0,057280	-0,04833
VaR[2]	-0,051193	-0,04541
VaR[3]	-0,057906	-0,04992
VaR[4]	-0,067160	-0,05436
VaR[5]	-0,063165	-0,04934
VaR[6]	-0,054959	-0,04376
VaR[7]	-0,053596	-0,04738
VaR[8]	-0,055852	-0,05139
VaR[9]	-0,056585	-0,04832

**Tabela 4.20:** *Valores do  $VaR[h]$ ,  $h = 1, \dots, 9$  dos modelos (4.3) e (4.5), ajustados aos retornos diários da série *cafeA*.*

## 5 CONCLUSÕES

As principais conclusões da análise apresentada ao longo da seção anterior são as seguintes:

1. O modelo que melhor ajustou os dados e portanto explica melhor sua variância condicional foi o modelo GARCH(3,3).
2. No que diz respeito à capacidade preditiva, mais que três passos à frente, o modelo ARCH(5) se mostrou mais eficaz, do que o modelo GARCH(3,3).
3. Se o interesse recair na geração de previsão um passo a frente, então o modelo que apresenta melhor capacidade preditiva é o GARCH(3,3).
4. No caso de uma posição comprada os valores em risco indicam perda, ou seja, indicam baixa no valor da saca de 60kg com relação ao ponto de origem das previsões.
5. A maior baixa acontece para um horizonte de quatro passos à frente, com uma possível perda de 6,716% do valor investido.

## 6 REFERÊNCIAS

- 1 Ansley, C.F. and Newbold, P. (1980). *Finite amostra properties of estimators for autoregressive moving average models*. Journal of Econometrics, **13**, 159–183.
- 2 Akaike , H. (1973). *Maximum likelihood identification of gaussian autoregressive moving average models*. Biometrika, **60**, 255–265.
- 3 Akaike, H. (1974). *A new look at the statistical model identification*. IEEE Transactions on Automatic Control, AC-19, 716–723.
- 4 Akaike H. (1977). *On Entropy Maximization Principle*. In Applications of Statistics (P.R. Krishnaiah, ed.), 27–41. Amsterdam: North-Holland.
- 5 Bartlett, M. S., *On the theoretical specification of sampling properties of autocorrelated time series*. Journal of the Royal Statistical Society, série B, vol. 27, 1946, pp.27–41.
- 6 Black, F. and Scholes, M. (1973). *The pricing of options and corporate liabilities*. Journal of Political Economy, 81, 635–654.
- 7 Bollerslev, T. (1986). *Generalized autoregressive conditional heteroscedasticity*. Journal of Econometrics, 31, 307–327.
- 8 Bollerslev, T. (1987). *A conditionally heterokedastic time series model for speculative process ande rates of return*. Review of Economics and Statistics, **69**, 542–547.
- 9 Bollerslev, T. (1988). *On the correlation structure for the generalized autoregressive conditional heteroscedasticity*. Journal of Times Series Analysis, **9**, 121–132.
- 10 Box, G.E.P. and Jenkins, G.M. (1970). *Time Series Analysis, Forecasting and Control*. San Francisco: Holden Day.
- 11 Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis, Forecasting and Control*. San Francisco: Holden Day. Second Edition.
- 12 Box, G.E.P., Jenkins, G.M. and Reinsel, G., (1994). *Time Series Analysis: Forecasting and Control*. Third Edition. Englewood Cliffs: Prentice Hall.

- 13 Box, G.E.P. and Pierce, D.A. (1970). *Distribution of residual autocorrelations in autoregressive-integrated moving average time series models*. Journal of the American Statistical Association, **64**, 1509–1526.
- 14 Brockwell, P.J. and Davis, R.A. (1991). *Times Series: Theory and Methods*. Second Edition. New York: Springer.
- 15 Brockwell, P.J. and Davis, R.A. (1996). *Introduction to Time Series and Forecasting*, New York: Springer-Verlag.
- 16 Centro de Estudos Avançados em Economia Aplicada (CEPEA). *Indicadores Diários do Café Arábica*. Disponível em: <[www.cepea.esalq.usp.br](http://www.cepea.esalq.usp.br)>. Acesso em: 27/09/2004.
- 17 Cordeiro, D.M. (2002), *Séries Temporais, Análise Quantitativa, Teoria e Aplicações*. Recife, EDUPE.
- 18 Cordeiro, G.M. and Klein, R. (1994), *Bias Correction of Maximum Likelihood Estimates for ARMA Models*. Probability and Statistics Letters, 169–76.
- 19 Dickey, D.A. and Fuller, W.A. (1979), *Distribution of the estimators for autoregressive time series with a unit root*. Journal of the American Statistical Association, vol. 74, pp. 427–431.
- 20 Engle R.F. (1982). *Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation*. Econometrica, **50**, 987–1008.
- 21 Gujarati, D.N., (2000). *Econometria Básica*, 3<sup>a</sup> ed. São Paulo, Makron Books.
- 22 Hannan E.J. (1980). *The estimation of the order of an ARMA process*. Annals of Statistics, **8**, 1071–1081.
- 23 Hannan E.J. (1982). *Testing for autocorrelation and Akaike's criterion*. In *Essays in Statistical Science, special volume 19A of Journal of Applied Probability*. The Applied Probability Trust, Sheffied, 403–412.
- 24 Hurvich, C.M. and Tsai, C.L. (1989). *Regression and time series model selection in small samples*. Biometrika, **76**, 297–307.



- 25 Hokstad P. (1983). *A method for diagnostic checking of time series models*. Journal of Time series Analysis, **4**, 177–184.
- 26 Ljung, G. and Box, G.E.P (1978). *On a measure of lack of fit in times series models*. Biometrika, **65**, 297–303.
- 27 McLeod G. (1983). *Box-Jenkins in Practice*. Lancaster: Gwilym Jenkins and Partners Ltd.
- 28 Morettin, P.A., (1981). *Modelo para Previsões de Série Temporais*, volumes I e II. Rio de Janeiro IMPA – CNPq.
- 29 Morettin, P.A. e Toloi, C.M.C., (1985). *Previsão de Séries Temporais*. Atual Editora, São Paulo.
- 30 Morettin, P.A. e Toloi, C.M.C., (2004). *Análise de Séries Temporais*. São Paulo, Blücher.
- 31 Newbold, P. and Bos, T. (1994). *Introductory Bussiness & Economic Forecasting*. 2<sup>a</sup> ed. Cincinnati: South-Wertern Publishing.
- 32 Quenouille, M.H. (1949). *Approximate test of correlation in time series*. Journal of the Royal Statistical Society, Series B, **11**, 68–84.
- 33 Rissanem, J. (1978). *Modelling by shortest data description*. Automatica, **14**, 465–471.
- 34 Schwarz, G. (1978). *Estimating the dimension of a model*. The Annals of Statistics, **6**, 461–464.