

PAULO HENRIQUE SOTO COSTA

SÉRIES DE RETORNOS DE AÇÕES BRASILEIRAS:
VOLATILIDADE E VALOR EM RISCO

Tese apresentada ao Departamento de Engenharia Industrial da PUC / Rio como parte dos requisitos para obtenção do título de Doutor em Engenharia de Produção.

Orientador: Tara Keshar Nanda Baidya

Departamento de Engenharia Industrial

Pontifícia Universidade Católica do Rio de Janeiro

Rio de Janeiro, 25 de abril de 2001.

Meus agradecimentos

a Tara Baidya, meu orientador, pelo apoio e amizade.

à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES, pelo auxílio financeiro recebido.

RESUMO

O objetivo principal do trabalho é o estudo dos resultados obtidos com a aplicação de diferentes modelos para estimar a volatilidade das ações brasileiras.

Foram analisadas as séries de retornos diários de seis ações, num período de 1200 dias de pregão.

Inicialmente, as séries foram estudadas quanto a suas propriedades estatísticas: estacionariedade, distribuição incondicional e independência. Concluiu-se que as séries são estacionárias na média, mas não houve conclusão quanto à variância, nesta análise inicial. A distribuição dos retornos não é normal, por apresentar leptocurtose. Os retornos mostraram dependência no tempo, linear e, principalmente, não linear.

Modelada a dependência linear, foram aplicados dez modelos diferentes para tentar capturar a dependência não linear através da modelagem da volatilidade: os modelos foram avaliados, dentro e fora da amostra, pelos seus resíduos e pelos erros de previsão. Os resultados indicaram que os modelos menos elaborados tendem a representar pior o processo gerador dos dados, mas que os modelos pouco parcimoniosos são de difícil estimação e seus resultados não correspondem ao que seria esperado em função de sua sofisticação.

As volatilidades estimadas pelos dez modelos foram utilizadas para prever valor em risco (VaR), usando-se dois processos para determinar os quantis das distribuições dos resíduos: distribuição empírica e teoria de valores extremos. Os resultados indicaram que os modelos menos elaborados prevêem melhor o VaR. Isto se deve à não estacionariedade das séries na variância, que fica evidente ao longo do trabalho.

ABSTRACT

This thesis aims to study the results of applying different models to estimate Brazilian stock volatilities. The models are applied to six series of daily returns, and each series has 1200 days.

We studied first the series' main statistical features: Stationarity, unconditional distribution and independence. We concluded that the series are mean stationary, but there was no conclusion on variance stationarity, in this first analysis. Return distribution is not normal, because of the high kurtosis. Returns showed time dependence, linear and, mainly, not linear.

We modeled the linear dependence, and then applied ten different volatility models, in order to try to capture the non linear dependence. We evaluated the different models, in sample and out of sample, by analyzing their residuals and their forecast errors. The results showed that the less sophisticated models tend to give a worst representation of the data generating process; they also showed that the less parsimonious models are difficult to estimate, and their results are not as good as we could expect from their sophistication.

We used the ten models' volatility forecasts to estimate value-at-risk (VaR) and two methods to estimate the residual distribution quantiles: empirical distribution and extreme value theory. The results showed that the less sophisticated models give better VaR estimates. This is a consequence of the variance non stationarity, that became apparent along the thesis.

SUMÁRIO

LISTA DE ILUSTRAÇÕES.....	VI
LISTA DE TABELAS.....	VII
CAPÍTULO 1: INTRODUÇÃO.....	1
1.1 – Objetivos do trabalho.....	1
1.2 – Revisão da literatura.....	3
1.3 – Estrutura do trabalho.....	7
CAPÍTULO 2: PROPRIEDADES ESTATÍSTICAS.....	11
2.1 – Introdução.....	11
2.2 – As séries.....	13
2.3 – Estacionariedade.....	17
2.3.1 – Introdução.....	17
2.3.2 – Teste ADF.....	19
2.3.3 – Modelagem GARCH.....	20
2.3.4 – Média condicional.....	22
2.3.4.1 – Teste de bicorrelação	
2.3.4.2 – Regressão localmente ponderada	
2.3.4.3 – Estacionariedade da média: conclusão	
2.3.5 – Estacionariedade: conclusão.....	29
2.4 – Distribuição incondicional.....	31
2.5 – Independência.....	39
2.5.1 – Introdução	39
2.5.2 – Dependência linear.....	40
2.5.2.1 – Autocorrelação	
2.5.2.2 – Modelagem ARMA	
2.5.3 – Dependência não linear.....	45
2.5.3.1 – Resíduos quadrados	
2.5.3.2 – Teste BDS	
2.5.4 – Independência: conclusão.....	50
2.6 – Conclusão do capítulo.....	50
2.7 – Apêndice do capítulo 2.....	52
2.7.1 – Comentário sobre a persistência e a estacionariedade dos modelos GARCH e IGARCH.....	52
2.7.2 – Detalhamento do uso da regressão localmente ponderada...	53
2.7.3 - Detalhamento do teste de Harvey, Leybourne e Newbold...	57

CAPÍTULO 3: MODELOS DE VOLATILIDADE.....	59
3.1 – Introdução.....	59
3.2 – Os modelos.....	61
3.2.1 – Janela de tempo.....	61
3.2.2 – Amortecimento exponencial do retorno quadrado.....	62
3.2.3 - Amortecimento exponencial do retorno absoluto.....	63
3.2.4 – GARCH.....	64
3.2.5 – EGARCH.....	65
3.2.6 – Hentschel.....	66
3.2.7 – Gray completo.....	68
3.2.8 – Gray simplificado.....	70
3.2.9 – Volatilidade de Parkinson.....	70
3.2.10 – Taylor.....	72
3.3 – Resultados da estimação	75
3.3.1 – Janela de tempo.....	75
3.3.2 – Amortecimento exponencial do retorno quadrado.....	75
3.3.3 - Amortecimento exponencial do retorno absoluto.....	76
3.3.4 – GARCH.....	77
3.3.5 – EGARCH.....	78
3.3.6 – Hentschel.....	78
3.3.7 – Gray completo.....	80
3.3.8 – Gray simplificado.....	82
3.3.9 – Volatilidade de Parkinson.....	83
3.3.10 – Taylor.....	84
3.4 – Avaliação dos modelos.....	86
3.4.1 – Introdução.....	86
3.4.2 – Análise dos resíduos.....	86
3.4.3 – Análise das previsões de volatilidade.....	94
3.5 – Conclusão do capítulo.....	103
3.6 – Apêndice do capítulo 3.....	106
3.6.1 – Modelagem da assimetria pelo EGARCH.....	106
3.6.2 – Casos particulares mais importantes do modelo de Hentschel.....	108
3.6.3 – Detalhamento do modelo de Gray.....	109
3.6.4 – Soluções inadequadas do modelo de Gray.....	112
3.6.5 – Procedimento de estimação do modelo de Taylor.....	114

CAPÍTULO 4: VALOR EM RISCO.....	116
4.1 – Introdução.....	116
4.2 – VaR a partir dos quantis empíricos.....	118
4.3 – VaR a partir dos quantis estimados pela TVE.....	126
4.4 – Conclusão do capítulo.....	130
4.5 – Apêndice do capítulo 4.....	132
4.5.1 – Conceituação do VaR.....	132
4.5.1.1 – Cálculo do VaR	
4.5.1.2 – Verificação das estimativas de VaR	
4.5.2 – TVE: resumo teórico.....	136
4.5.3 – TVE: critérios de aplicação do método.....	140
4.5.4 – Resultados da estimação dos quantis de 5% por TVE.....	143
4.5.5 – Resultados do VaR de 99%.....	149
 CAPÍTULO 5: CONCLUSÃO.....	 154
5.1 – Introdução.....	154
5.2 – Propriedades estatísticas.....	155
5.3 – Modelos de volatilidade.....	157
5.4 – Valor em risco.....	158
5.5 – Possíveis extensões.....	159
 BIBLIOGRAFIA.....	 161

LISTA DE ILUSTRAÇÕES

FIGURA 1	14
FIGURA 2	15
FIGURA 3	15
FIGURA 4	36
FIGURA 5	37
FIGURA 6	42
FIGURA 7	44
FIGURA 8	46
FIGURA 9	55
FIGURA 10.....	62
FIGURA 11.....	85
FIGURA 12.....	107
FIGURA 13.....	110
FIGURA 14.....	111
FIGURA 15.....	113
FIGURA 16.....	114
FIGURA 17.....	124
FIGURA 18.....	124
FIGURA 19.....	125
FIGURA 20.....	125
FIGURA 21.....	133
FIGURA 22.....	140
FIGURA 23.....	141

LISTA DE TABELAS

TABELA 1	16
TABELA 2	20
TABELA 3	21
TABELA 4	25
TABELA 5	26
TABELA 6	28
TABELA 7	35
TABELA 8	36
TABELA 9	38
TABELA 10	41
TABELA 11	43
TABELA 12	43
TABELA 13	45
TABELA 14	49
TABELA 15	75
TABELA 16	75
TABELA 17	77
TABELA 18	78
TABELA 19	79
TABELA 20	81
TABELA 21	81
TABELA 22	82
TABELA 23	84
TABELA 24	85
TABELA 25	89
TABELA 26	90
TABELA 27	92
TABELA 28	93
TABELA 29	97
TABELA 30	98
TABELA 31	99
TABELA 32	100
TABELA 33	119
TABELA 34	120
TABELA 35	121
TABELA 35	122
TABELA 37	127
TABELA 38	127
TABELA 39	128
TABELA 40	143

TABELA 41.....	144
TABELA 42	144
TABELA 43	145
TABELA 44	145
TABELA 45	146
TABELA 46	146
TABELA 47	147
TABELA 48	147
TABELA 49	148
TABELA 50	148
TABELA 51	149
TABELA 52	150
TABELA 53	151
TABELA 54	151
TABELA 55	152
TABELA 56	153

CAPÍTULO 1 – INTRODUÇÃO

1.1 – OBJETIVOS DO TRABALHO

Existe já há alguns anos, no Brasil e na maioria dos países, uma tendência a se usar cada vez mais a ferramenta ‘Valor em Risco’. Isto decorre da exigência de bancos centrais e outros órgãos reguladores no sentido que as instituições financeiras mantenham reservas proporcionais ao VaR de suas posições em ativos financeiros, e decorre também do próprio interesse das instituições em usar o VaR como ferramenta de controle e gerenciamento do risco de suas operações.

A pergunta inicial que motivou este trabalho é: ‘como se deve proceder para calcular – **bem** – o VaR de ações brasileiras?’.

O cálculo do VaR, em si, é trivial mas, para que ele seja bem feito, é indispensável a escolha de um modelo de volatilidade adequado. Daí surge uma segunda pergunta, fundamental e relevante: ‘como se comportam os modelos de variância condicional, quando aplicados a séries de retornos de ações brasileiras?’.

Responder a segunda pergunta tornou-se então o objetivo central deste trabalho.

Evidentemente, este é um trabalho empírico. Não estão sendo criados novos modelos: está-se procurando analisar os resultados que podem ser obtidos com o uso dos modelos existentes.

A abrangência do trabalho é a sua característica mais importante. Esta abrangência diz respeito ao número e à diversidade de modelos, às séries com e sem observações aberrantes, e à análise de resultados por resíduos, previsões e VaR.

Foram selecionados para análise dez modelos de volatilidade: cinco de uso relativamente corrente e cinco mais elaborados, que nunca foram aplicados a séries brasileiras e não estão em ‘pacotes’ de computador.

Estes modelos foram aplicados a séries de retornos de seis ações brasileiras. Não houve a preocupação em formar ou otimizar carteiras de ações e/ou outros ativos financeiros, porque isto não contribuiria com o objetivo central do trabalho, que é analisar *os modelos*, através dos resultados que podem ser obtidos com seu uso nas séries.

Num estudo deste tipo é comum dar algum tratamento às observações aberrantes, para que elas não venham distorcer as análises. Visando quantificar a influência destas observações nos resultados dos modelos de volatilidade, foram criadas séries ‘modificadas’ para os retornos das seis ações, sem alguns valores de retornos correspondentes a dias atípicos.

Assim, passou-se de seis para doze séries de retornos que, servindo de dados aos dez modelos de volatilidade, geraram 120 séries-modelos para análise.

Estas combinações de modelos e séries foram analisadas quanto aos resíduos (4 critérios) e erros de previsão (4 critérios); posteriormente tiveram seu VaR estimado por dois processos e com dois valores de probabilidade.

Ao lado da abrangência, houve a preocupação de se procurar bons métodos para desenvolver as análises necessárias, alguns deles pouco ou nunca usados em séries financeiras brasileiras, como a regressão localmente ponderada, o teste robusto de Harvey, Leybourne e Newbold (1997) para erros de previsão, e o teste BDS aplicado a resíduos GARCH com as correções de Fernandes e Preumont (1996).

Esta preocupação com o uso de bons métodos motivou também o capítulo 2, onde são estudadas as propriedades estatísticas das séries, para nortear a aplicação dos modelos. Neste capítulo ficou caracterizada uma dependência linear nos retornos, que fez com que modelos de volatilidade e VaR tivessem como dados não os retornos, mas os resíduos de modelos lineares aplicados aos retornos.

O uso da Teoria de Valores Extremos na determinação do VaR também é fruto desta procura dos melhores métodos.

Entre os objetivos de um trabalho empírico está também a observação de características e padrões nos dados. Algumas constatações interessantes surgiram do manuseio das séries e modelos; entre elas pode-se citar:

- A verificação da não estacionariedade das séries de retornos e das dificuldades que ela causa na modelagem e no VaR.
- A verificação da irrelevância das observações aberrantes na qualidade dos resultados dos modelos e do VaR.
- A verificação da assimetria da volatilidade de nossas ações com relação ao sinal do choque de retorno.
- A verificação de que o fator de decaimento do amortecimento exponencial recomendado pelo documento técnico do RiskMetrics (1995) é muito alto para nossas ações.

1.2 – REVISÃO DA LITERATURA

O estudo das propriedades estatísticas das séries financeiras inicia-se com Bachelier (1900), em um trabalho notável que implicava em que os retornos tivessem distribuição normal independente e idêntica. Alguns outros trabalhos se seguiram, entre os quais pode-se citar o de Working (1934) que estudou séries de preços de futuros de *commodities*, e o de Kendall (1953) que estudou séries de preços de ações e de produtos agrícolas. No entanto, foi com Fama (1965) que se iniciou a publicação de um grande número de trabalhos sobre a estatística das séries financeiras; ele estudou em detalhe as séries de retornos das 30 ações do *Dow Jones Industrial Average Index*, do mercado de ações norte-americano, num longo e fundamental artigo onde conclui:

- que as séries têm distribuições com caudas pesadas incompatíveis com as da normal.
- que as distribuições das séries podem ser representadas por distribuições estáveis paretianas com expoente característico menor que dois.

- que a dependência entre os retornos, medida pelos coeficientes de autocorrelação e pelo número e comprimento dos *runs* é muito pequena para ter significado econômico, e que não é possível obter ganhos continuados usando a técnica de “filtro” proposta por Alexander (1961).
- que o processo estocástico dos preços se aproxima do modelo de “passeio aleatório”, não sendo portanto possível usar preços passados para prever preços futuros.

Nos muitos trabalhos que se seguiram, embora a constatação da não normalidade dos retornos seja amplamente aceita, o mesmo não acontece com as distribuições estáveis paretianas.

É interessante observar que o uso destas duas distribuições para os retornos têm uma justificativa econômica, pois o retorno (logarítmico) em um dado período de tempo é a soma dos retornos verificados em períodos menores, ou mesmo transação a transação, e tanto a normal como as distribuições estáveis paretianas são preservadas quando se somam variáveis aleatórias. Ao contrário, supor distribuição t (por exemplo) para os retornos diários, significa também supor que os retornos horários ou semanais têm distribuições diferentes da t .

A explicação mais aceita atualmente para as caudas pesadas das distribuições de retornos é a de mistura de distribuições; muitos autores, como Praetz (1972) e Clark (1973) argumentam que esta mistura é de distribuições normais. A idéia da mistura de distribuições deu origem a muitos modelos não lineares para as séries de retornos e os mais difundidos nesta linha são os da família GARCH, introduzida por Engel (1982) e generalizada por Bollerslev (1986), onde fica caracterizada uma dependência não linear entre os retornos, função da dependência serial da variância condicional.

O modelo do passeio aleatório também foi muito estudado, dando origem a grande número de artigos, entre os quais podem-se destacar Granger e Morgenstern (1970), Cunningham (1973), e Cooper (1982), entre muitos outros voltados para retornos de ações. De maneira geral os autores concordam com a validade deste modelo (ou de variações dele), pelo menos no que diz respeito à imprevisibilidade dos retornos. Taylor (1986), por exemplo, testa a hipótese de passeio aleatório e a previsibilidade de 40

séries de preços, concluindo que os preços de ações seguem o passeio aleatório, porém encontra alguma previsibilidade em preços de futuros de *commodities*.

Mais recentemente, aceita a impossibilidade de obter lucro usando séries de retornos para prever retornos, a ênfase dos trabalhos com séries temporais financeiras deslocou-se para a modelagem da variância: os retornos esperados são previstos através de outros tipos de modelos (análise de fatores, por exemplo) e apenas o risco vem das séries temporais de retornos.

Nesta linha, os modelos GARCH foram expandidos e generalizados em várias direções: por exemplo, o GARCH exponencial [de Nelson (1991)], o de limiar [Zakoian (1991)], a generalização de Hentschel (1995) que engloba vários modelos em uma só especificação, e os modelos com parâmetros variáveis no tempo de acordo com processos de Markov [Hamilton e Susmel (1994) e Gray (1996)]. Outras técnicas também foram empregadas, como por exemplo os modelos de volatilidade estocástica [ver Hsieh (1993), que emprega modelo baseado na volatilidade de Parkinson (1980), ou Kim, Shephard e Chib (1998) que comparam volatilidade estocástica e modelos da família GARCH] e também os modelos utilizando distribuição de valores extremos [ver por exemplo Embrechts, Resnick, e Samorodnitsky (1998) ou McNeil (1998)]

Encontram-se também, nos periódicos especializados, diversos trabalhos sobre a modelagem de séries temporais financeiras em países emergentes; entre eles, pode-se citar Bekaert e Harvey (1997), Duarte e Mendes (1997), Campbell (1995) e Susmel (1998).

Não são raros em nossos periódicos artigos onde se estudam séries temporais de retornos de ativos financeiros brasileiros. Alguns artigos, como este trabalho, comparam modelos de variância condicional e, dentre eles destacam-se:

- Herencia, Hotta e Pereira (1998), que comparam modelos GARCH e de volatilidade estocástica; a ênfase é na comparação do GARCH (1,1) com o AR(1) VE, que é um modelo onde a variância segue um processo linear autorregressivo de primeira ordem. Os autores fazem comparações de previsões de volatilidade, qualitativamente e através do 'critério de retorno observado'. Eles usam como

dados os retornos de Telebrás e também empregam séries de retornos simulados, onde se tem a vantagem de conhecer o processo estocástico gerador dos dados.

- Morais e Portugal (1999) que, na mesma linha do artigo acima, comparam modelos GARCH e de volatilidade estocástica. Os dados aqui são a série de retornos do IBOVESPA e os modelos GARCH são extensões do modelo original de Bollerslev (1986): o EGARCH e o GJR-GARCH, que são aqui abordados no capítulo 3 (EGARCH) e no apêndice do capítulo 3 (GJR, como caso particular do modelo de Hentschel).
- Ziegelman e Pereira (1997), que comparam 7 modelos de volatilidade estocástica: um sem deformação temporal (DT), três com DT em tempo discreto, e três com DT em tempo contínuo. Os dados são também os retornos do IBOVESPA e a média condicional é modelada por um AR(1) antes da aplicação dos modelos de volatilidade.
- Issler (1999), que compara modelos da família ARCH: o GARCH, o EGARCH, o TARCH e o SWARCH. O TARCH também está aqui citado no apêndice do capítulo 3, como caso particular do modelo de Hentschel. O SWARCH é um modelo de parâmetros variáveis segundo um processo de Markov, na mesma linha do modelo de Gray (aqui estudado); no entanto, ele é uma extensão do ARCH e não do GARCH, como é o modelo de Gray. O autor usa quatro séries diferentes como dados (dólar, cacau, C-bond e Telebrás) e compara o ajuste e as previsões de volatilidade dos modelos.
- Pereira, Hotta, Souza e Almeida (1999) que, num trabalho abrangente, comparam modelos da família GARCH (o GARCH comum e o EGARCH), volatilidade estocástica (AR[1] VE) e o SWARCH. Eles usam três critérios de comparação de previsões de volatilidade, e analisam a autocorrelação dos resíduos e dos resíduos quadrados. Os dados usados são séries de retornos do dólar, C-bond e Telebrás, e os autores usam variáveis *dummies* para modelar o efeito de observações extremas. A média condicional é modelada por um AR(p), estimado simultaneamente com os modelos de volatilidade.

Pode-se notar, pelo conteúdo dos artigos acima citados, que a comparação empírica de modelos de volatilidade aplicados a dados brasileiros é um tema atual e em desenvolvimento. O presente trabalho contribui para o desenvolvimento deste tema:

- ampliando o leque de modelos analisados, que inclui também modelos mais simples (como o de janela de tempo), o modelo de Hentschel, que engloba a maioria dos modelos GARCH em uma só especificação, o modelo de Gray, que é um *SW-GARCH*, e dois modelos de volatilidade estocástica diferentes do AR(1) VE;
- aplicando os modelos a séries com e sem observações extremas;
- apresentando uma comparação sistemática de resíduos e erros de previsão, por diferentes critérios;
- indo além da modelagem da volatilidade, ao empregar os modelos para determinação de valor em risco, permitindo analisar os resultados da modelagem sob um outro ponto de vista.

1.3 – ESTRUTURA DO TRABALHO

O capítulo dois –PROPRIEDADES ESTATÍSTICAS – inicia-se detalhando como foram definidas as séries usadas no trabalho, passando-se em seguida para o estudo de sua estacionariedade onde, além da aplicação do teste ADF¹, usual para processos lineares, faz-se uma análise supondo que o processo gerador dos dados é um GARCH (1,1). A seguir, buscando maior generalidade no estudo da estacionariedade, são feitas análises da média condicional, sem suposições a respeito do processo gerador dos dados: estas análises consistem em testar o coeficiente de bicorneleração das séries, e na sua modelagem através do método de regressão localmente ponderada.

¹ Nesta seção introdutória não são indicadas as referências para os métodos citados; elas aparecem nos próximos capítulos, quando eles são efetivamente empregados.

A análise da estacionariedade conclui que, se o processo gerador dos dados for linear, ele é estacionário; se ele for GARCH (1,1) ele tem persistência extremamente alta, não se descartando a possibilidade de que ele seja um GARCH *integrado*, sem variância incondicional finita. Os testes com a média condicional, de maneira geral, indicam que as séries são estacionárias na média, não sendo possível concluir quanto à variância.

Ao contrário da análise da estacionariedade, a da distribuição incondicional é bastante simples: as caudas pesadas das distribuições dos retornos levam à evidente rejeição da normalidade incondicional. Esta rejeição permanece mesmo quando são retirados da amostra dias atípicos, e também quando se usa teste robusto que não leva em conta o coeficiente de curtose que, por envolver a quarta potência do desvio com relação à média, é muito sensível à presença de observações aberrantes.

Passando-se à independência, os testes usuais de autocorrelação indicam alguma dependência linear, enquanto que os de autocorrelação dos quadrados e o teste BDS indicam forte dependência não linear entre os retornos. A dependência não linear é um fato estilizado das séries de retornos e é justamente o que vai-se tentar modelar no capítulo seguinte. Para que a dependência linear existente não interfira na modelagem da variância, são aplicados modelos ARMA (p,q) aos retornos, e os resíduos destes modelos, supostamente livres de dependência linear, é que serão os dados para os modelos de volatilidade.

O capítulo três –MODELOS DE VOLATILIDADE – inicia-se com a descrição de cada um dos modelos, seu equacionamento e suas características mais importantes. Em seguida os parâmetros destes modelos são estimados: alguns com o auxílio de rotinas prontas em programas como o EViews, enquanto que para outros foi necessário programar algumas passagens do processo de estimação. Esta estimação mostrou-se simples e com bons resultados para a maior parte dos modelos mas, para os menos parcimoniosos, houve problemas de vários tipos, inclusive falta de convergência dos algoritmos de estimação, em algumas séries.

Estimados os modelos, passou-se à análise dos resultados obtidos, para avaliar o desempenho dos modelos.

O primeiro critério de avaliação foi o estudo dos resíduos. Ele parte do princípio que, se o modelo conseguiu capturar a dinâmica do processo gerador dos dados, os resíduos não podem ter estrutura, e devem ser independentes no tempo, com distribuição idêntica. Como estes resíduos podem ser interpretados como *retornos* re-escalados (divididos pelo desvio padrão condicional), sua distribuição deveria ser normal, se se supõe verdadeira a tese de que os retornos são (incondicionalmente) misturas de distribuições normais. Aliás, todos os modelos estimados por máxima verossimilhança supuseram distribuição normal para os resíduos.

No entanto, todas as séries de resíduos apresentaram distribuições com caudas pesadas que levaram à rejeição da normalidade; um dos critérios de análise dos resíduos foi o coeficiente de curtose de sua distribuição: quanto menor (mais próximo de três), melhor.

Mesmo rejeitada a normalidade, um bom modelo não pode ter resíduos com estrutura: testou-se a existência de autocorrelação nos resíduos e em seus quadrados (que se mostrou significativa em alguns casos), e se sua distribuição é independente e idêntica – teste BDS – (o que foi rejeitado em vários casos). A conclusão aqui é que, em geral, há alguma estrutura nos resíduos dos modelos.

Como outro critério de avaliação dos modelos, procurou-se verificar seus erros de previsão de volatilidade: nesta verificação as previsões foram comparadas com *representantes* da volatilidade. Um dos critérios consistiu em comparar a variância condicional prevista com o quadrado do retorno observado; o outro consistiu em comparar o desvio padrão condicional previsto com o valor absoluto do retorno observado. Como não há uma referência absoluta para servir de padrão, que permita dizer se os erros de previsão são altos ou baixos, esta análise permitiu apenas que se comparassem os erros dos diferentes modelos nas diferentes séries.

A avaliação dos modelos, neste capítulo 3, indicou uma tendência de piores resultados, em termos de resíduos e erros de previsão, nos modelos menos elaborados e nas séries correspondentes aos papéis menos líquidos. Esta tendência não é muito pronunciada, e não há perfeita coerência entre os diferentes critérios de análise: alguns modelos são

bons para algumas séries em determinadas análises mas não para outras séries e/ou outros critérios de avaliação.

No capítulo quatro –VALOR EM RISCO – os modelos de volatilidade são usados para prever o VaR de aplicações nas ações das séries; para determinar o VaR, além da previsão de volatilidade, é necessário conhecer o correspondente quantil da distribuição dos resíduos do modelo. Como houve rejeição de normalidade para os resíduos, não se podem usar quantis da distribuição normal para cálculo do VaR. Então, estes quantis foram aqui determinados por dois processos: primeiramente, eles foram tirados diretamente das distribuições empíricas dos resíduos e, no segundo processo, usou-se a teoria de valores extremos para estimar a cauda esquerda da distribuição dos resíduos; este segundo processo não se mostrou significativamente melhor que o primeiro.

A maneira mais simples de avaliar previsões de valor em risco é através da contagem do número de vezes em que o retorno observado é pior (mais negativo) que o VaR. Este método foi empregado aqui, e os resultados indicam que os modelos mais simples, que tiveram propriedades piores que os outros nas avaliações do capítulo três, mostraram agora um desempenho, em média, superior aos demais. Este melhor desempenho é mais acentuado nas séries menos líquidas e praticamente não ocorre em Telebrás (a série mais líquida).

Esta aparente contradição é explicada pela não estacionariedade das séries de retornos, na variância: os modelos que tiveram os melhores resultados de contagem do número de ultrapassagens do VaR, além de serem os mais simples, são também modelos não estacionários².

Os testes do capítulo 2 não são conclusivos quanto à estacionariedade (ou não estacionariedade) das séries, na variância. Porém, ao longo do trabalho vão surgindo evidências que levam à conclusão que as séries não são estacionárias.

No capítulo cinco estão as conclusões do trabalho.

² O modelo de janela de tempo simplesmente mede a variância amostral nos últimos k dias, e os dois modelos de amortecimento exponencial aqui usados são, na verdade, modelos ARMA (1,1), no quadrado dos retornos ou em seu valor absoluto, com o coeficiente do termo autorregressivo igual a um e, portanto, têm raiz unitária. O modelo de amortecimento exponencial do retorno quadrado também pode ser visto como um IGARCH (1,1).

CAPÍTULO 2 – PROPRIEDADES ESTATÍSTICAS

2.1 – INTRODUÇÃO

A caracterização das propriedades estatísticas das séries de retornos das ações é importante para permitir a correta aplicação de modelos aos dados, que permitam inferir a respeito das características dos retornos, sobretudo no que diz respeito a média e variância, que vão determinar o retorno esperado e a previsão da volatilidade para os próximos períodos. Esta determinação é fundamental para a decisão de investimento não só na ação como também em seus derivativos.

Este capítulo analisa estacionariedade, distribuição e independência em seis séries de retornos de ações: Telebrás, Petrobrás, Eletrobrás, Bradesco, Vale do Rio Doce e Brahma, todas preferenciais. A escolha foi feita entre as mais líquidas do mercado, de setores diferentes da economia.

Todas as séries têm 1200 observações de retornos diários, num período de aproximadamente cinco anos, de julho de 1994 a maio de 1999. Este período foi escolhido porque se inicia logo após o ‘plano Real’, dispensando o uso de deflatores, e por ser suficientemente longo, como são em geral os períodos de dados para estudos deste tipo.

As séries temporais são analisadas quanto a estacionariedade. Inicialmente ela é estudada supondo os dados gerados por processo linear (teste ADF) ou por processo GARCH, e se conclui pela estacionariedade, embora no segundo processo a variância talvez não seja sempre finita. A seguir é estudada a estacionariedade da média sem premissas a respeito do processo gerador dos dados, através do coeficiente de bicorrelação (que indica que as séries estudadas são estacionárias) e através de regressão localmente ponderada (que indica estacionariedade em quatro séries e é inconclusiva para as outras duas). Pelo conjunto dos testes, conclui-se que as séries são

estacionárias na média, mas a análise quanto à estacionariedade na variância é inconclusiva³.

A análise da distribuição incondicional dos retornos é bastante simples, pois está centrada na constatação da evidente não normalidade dos retornos diários, que apresentam leptocurtose. A forma da função de distribuição das séries mais líquidas se aproxima à da *t* de Student. Uma possível explicação para a não normalidade incondicional seria distribuição condicional não independente e idêntica dos retornos diários, e isto motiva a modelagem do retorno condicional, desenvolvida no capítulo 3.

A independência entre os retornos consecutivos é estudada inicialmente através da função de autocorrelação, que indica existência de dependência (linear) nas séries. Ela é modelada através de um ARMA(1,1) mas os resíduos continuam apresentando correlação, o que indica uma possível dependência também não linear, o que é confirmado pelo fato de os quadrados dos resíduos serem fortemente correlacionados.

A seguir, é aplicado o teste BDS aos retornos e aos resíduos dos modelos ARMA. Como há rejeição da hipótese de os retornos e resíduos serem independentes e identicamente distribuídos, conclui-se formalmente que há dependência não linear decorrente de estrutura na variância condicional.

A maior parte das análises foi feita com o auxílio do programa EVIEWS 3.0 ; os testes BDS empregaram programa específico do Departamento de Economia da Universidade de Houston, EUA, (copyright 1987, 1988, por W. D. Dechert); as análises de biconexão e a regressão localmente ponderada foram feitas no programa EXCEL, em sub-rotinas escritas pelo autor , que se coloca à disposição para fornecer maiores informações sobre suas características.

Na seção 2.2 , a seguir, são apresentadas as séries e suas características mais gerais; na seção 2.3 estuda-se a estacionariedade; na seção 2.4 é analisada a distribuição incondicional, e em 2.5 a independência dos retornos, concluindo-se na seção 2.6. No fim do capítulo há um apêndice, seção 2.7 .

³ Como será visto mais adiante neste capítulo, será estudada a estacionariedade de segunda ordem das séries de retornos. A expressão “a série é estacionária na média (variância)” é usada com o sentido de “o

2.2 – AS SÉRIES

As séries de retornos foram montadas a partir dos preços de fechamento diários das ações, em Reais, corrigidos para dividendos e outros eventos, obtidos no banco de dados Economatica. O primeiro dia é 4/7/94 (logo após a criação do Real) e o último é 13/5/99.

Os trabalhos recentes e importantes de autores brasileiros sobre modelagem de volatilidade usam este mesmo período para os dados. Como exemplo, pode-se citar Moraes e Portugal (1999) que usam 1060 observações diárias, de 4/7/94 a 9/10/98; Pereira e outros (1999) que usam 989 dias no mesmo período; Issler (1999) que também usa 4 anos de dados diários logo após o ‘plano Real’.

O período coberto pelos dados contém algumas ocasiões de crise nos mercados financeiros: em março/95 houve a crise do México, em outubro/97 a da Ásia, em setembro/98 a da Rússia e em janeiro/99 a desvalorização do Real; entre maio/95 e julho/97 há um longo período de estabilidade.

A partir das séries de preços, os retornos foram calculados através do logaritmo neperiano do quociente entre dois preços consecutivos. Na seção 2.4, neste capítulo, estão apresentados os histogramas e algumas estatísticas descritivas de todas as séries de retornos. A seguir estão apresentados, apenas para ilustração do processo de obtenção dos retornos, três gráficos relativos à ação da Telebrás.

No gráfico da figura 1 estão os preços de fechamento das ações da Telebrás, nos 1200 dias do período estudado. Neste gráfico, pode-se observar que houve uma tendência de crescimento do preço no período considerado, e esta tendência está ilustrada por uma linha correspondente a uma função exponencial.

processo estocástico gerador da série de retornos é tal que o valor esperado (ou a variância) do retorno é constante”.



FIGURA 1: Preços das ações da Telebrás

Neste gráfico, como em todos os outros apresentados neste trabalho, os dias estão numerados de 1 a 1200; as datas correspondentes são:

<u>dia</u>	<u>data</u>
1	04/07/94
201	27/04/95
401	15/02/96
601	04/12/96
801	26/09/97
1001	21/07/98
1200	13/05/99

Na figura 2 estão os retornos diários da ação da Telebrás.

Os retornos diários de Telebrás têm média 0,00131 (positiva, coerente com a tendência de crescimento do preço) e desvio padrão 0,0367. Para uniformização e facilidade de comparação entre as séries elas foram padronizadas, ou seja, de cada observação foi subtraída a média, e o resultado foi dividido pelo desvio padrão, resultando séries de média zero e variância um. Todas as análises deste trabalho se referem às séries de retornos padronizados das ações.

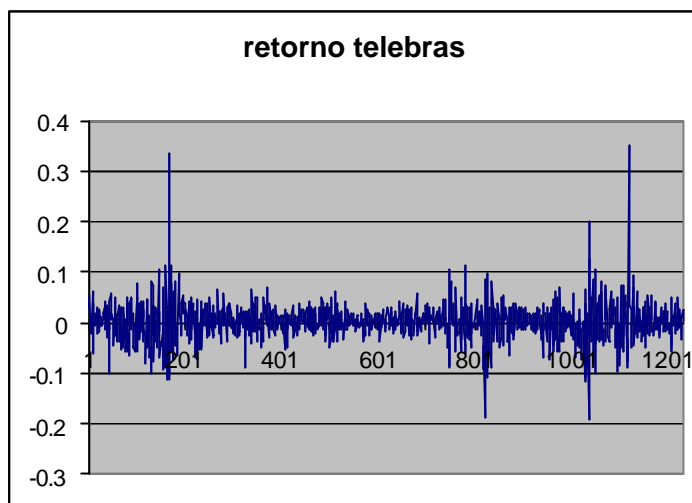


FIGURA 2: Retornos das ações da Telebrás

Na figura 3 estão os retornos padronizados das ações da Telebrás.

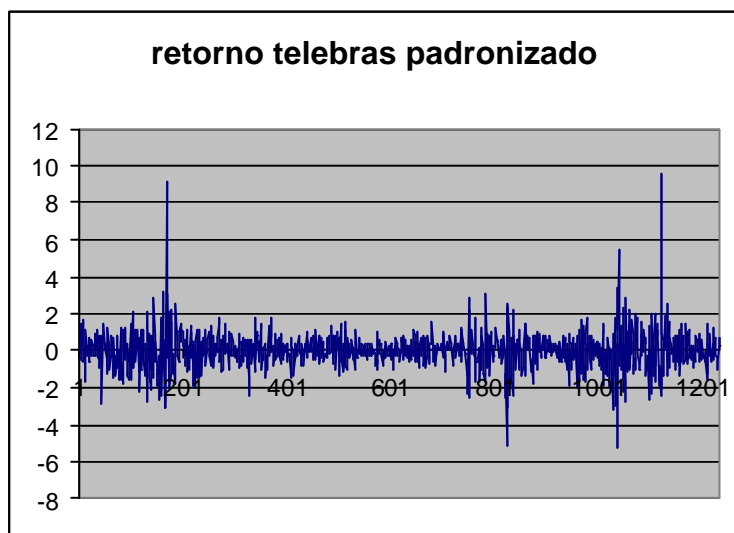


FIGURA 3: Retornos padronizados das ações da Telebrás

Observando a série dos retornos padronizados da Telebrás, nota-se que existem algumas observações aberrantes, com valores da ordem de até dez desvios padrão. Estes valores extremos de retornos ocorreram em dias atípicos. Em função disto foi formada uma nova série de retornos para cada ação, da qual foram retirados cinco valores de retorno, correspondentes aos dias 10/3/95 (crise do México), 27/10/97 (crise

da Ásia), 10/9/98 e 15/9/98 (crise da Rússia), e 15/1/99 (crise do Real) e foram acrescentados mais cinco dias no final, passando o último dia a ser 20/5/99.

Esta nova série foi chamada de *série modificada* de retornos. A série modificada de retornos diários Telebrás tem média 0,00085 e desvio padrão 0,0325. As séries modificadas também foram padronizadas, resultando duas séries de retornos padronizados para cada ação: dados originais e dados modificados (sem os cinco dias de crise).

As séries de retornos padronizados receberam siglas, conforme a tabela 1.

AÇÃO	DADOS ORIGINAIS	DADOS MODIFICADOS
TELEBRÁS	OTELB	MTELB
PETROBRÁS	OPETR	MPETR
ELETROBRÁS	OELET	MELET
BRADESCO	OBBDC	MBBDC
VALE DO RIO DOCE	OVALE	MVALE
BRAHMA	OBRHA	MBRHA

TABELA 1: Siglas das séries de retornos padronizados

Estas são as 12 séries que terão suas propriedades estudadas a seguir. É interessante observar que a retirada de valores aberrantes de séries de retorno é procedimento adotado por muitos autores, mas contestado por outros.

A modificação das séries foi feita aqui apenas com o objetivo de verificar sua influência nos resultados, e não implica em julgamento da conveniência ou da correção deste tipo de modificação. Telebrás, por exemplo, está representada por duas séries de retornos (OTELB e MTELB), que serão analisadas como se fossem séries de papéis diferentes.

2.3 – ESTACIONARIEDADE

2.3.1 – INTRODUÇÃO

Um processo estocástico representado por uma seqüência de variáveis aleatórias $\{... Y_1, Y_2, \dots, Y_n, \dots\}$ é estritamente estacionário se, para todos os inteiros i, j e para todos os inteiros positivos k , a distribuição multivariada de $(Y_i, Y_{i+1}, \dots, Y_{i+k})$ é idêntica à de $(Y_j, Y_{j+1}, \dots, Y_{j+k})$. Na prática só é possível testar algumas das conseqüências da premissa de que um processo é estritamente estacionário; em particular, se as distribuições multivariadas são idênticas, então $E(Y_t)$ é igual a alguma constante m e $cov(Y_t, Y_{t+\tau})$ é igual a uma constante I_t que depende apenas de t . Um processo que atenda a estas duas condições é dito *estacionário de segunda ordem*, que será aqui abreviado por *estacionário*. Algumas características dos processos estacionários são:

- Reversão à média: a série flutua em torno de uma média constante de longo prazo. Mais especificamente, a previsão do valor da variável representada pela série, feita em t para o instante $t+t$, dada por $E_t \{Y_{t+t}\}$, tende para a média m quando t é grande.
- O processo pode ser condicionalmente heteroscedástico mas sua variância incondicional é finita e constante no tempo. A previsão da variância condicional feita em t para o instante $t+t$, dada por $E_t \{Y_{t+t} - E_t \{Y_{t+t}\}\}^2$ tende para $var(Y_t)$ quando t é grande.
- A covariância I_t decresce com t , tendendo a zero quando t é grande.

A questão da estacionariedade ou não do processo estocástico gerador dos dados tornou-se extremamente importante nos anos 70, com o desenvolvimento da modelagem linear das séries temporais e a popularização dos modelos ARMA, a partir do trabalho de Box e Jenkins (1976). Se a série não for estacionária é preciso fazer

transformações (tomar logaritmos, fazer transformação de Box e Cox [1964], fazer diferenças) de modo a *estabilizar* a série, ou seja, torná-la estacionária pois os parâmetros estimados para séries não estacionárias não têm boas propriedades estatísticas.

Dentro do contexto de séries representadas por modelos lineares, existe um caso particular de não estacionariedade, que é aquela gerada pela presença de raiz unitária. O teste ADF, descrito e aplicado a seguir, na seção 2.3.2, tem como hipótese nula a presença de raiz unitária no modelo linear que porventura represente o processo gerador dos dados.

Séries financeiras, e especialmente séries de retornos de ações, dificilmente são bem representadas por processos lineares. Uma vez admitido que o processo não é linear, a questão da estacionariedade diminui de importância pois processos não lineares não estacionários podem ter parâmetros com boas propriedades (pelo menos as assintóticas) e são às vezes empregados para representar séries financeiras.

Taylor (1986, pag. 63) diz que “Pode ser razoável esperar que um modelo estacionário aproxime adequadamente o processo verdadeiro por um período relativamente curto, por exemplo um ano. Para períodos maiores, entretanto, pode-se esperar alguma não estacionariedade. Não é fácil determinar a aplicabilidade de ajustar modelos estacionários a séries que cobrem vários anos” .

É interessante notar que a estacionariedade não é dos dados em si, mas do processo estocástico gerador dos dados: uma série de preços de uma ação, por exemplo, sempre tem uma média e uma variância amostrais, que não necessariamente coincidem com a média e variância do processo estocástico gerador dos dados, que pode até nem ter média e variância.

Os modelos da família GARCH são muito usados para representar séries de retornos. A formulação do GARCH(p,q) é:

$$y_t = \mathbf{s}_t \mathbf{e}_t \text{ com } \mathbf{e}_t \sim N(0,1)$$

$$\mathbf{s}_t^2 = \mathbf{w} + \sum_{i=1}^q \mathbf{a}_i y_{t-i}^2 + \sum_{i=1}^p \mathbf{b}_i \mathbf{s}_{t-i}^2$$

Bollerslev (1986) demonstra, em seu *teorema 1*, que este processo é estacionário (de segunda ordem) com $E(y_t) = 0$, $\text{var}(y_t) = \omega / (1 - \sum a_i - \sum b_i)$ e $\text{cov}(y_t, y_s) = 0$ se e somente se $\sum a_i + \sum b_i < 1$. Na seção 2.3.3 é estimado modelo GARCH(1,1) para todas as séries objetivando verificar se a condição de estacionariedade é atendida.

Se a série for melhor representada por outro modelo não linear, a verificação da estacionariedade depende, evidentemente, do modelo e não são comuns artigos onde se estabeleçam formalmente as condições de estacionariedade para modelos fora da família GARCH. Na seção 2.3.4 estuda-se a estacionariedade da média (se $E(Y_t)$ é constante) sem fazer suposições a respeito do processo gerador dos dados, usando processos indiretos que analisam a média condicional.

2.3.2 – TESTE ADF

O primeiro teste de estacionariedade aqui usado é o de Dickey e Fuller (1979) aumentado (ADF), descrito a seguir, e que será aplicado às séries de retornos, visando indicar, caso elas possam ser representadas por modelo linear, se o modelo é estacionário.

Um processo linear auto-regressivo de ordem p , $AR(p)$, pode ser escrito como

$$A(L)y_t = \mathbf{m} + \mathbf{e}_t$$

Onde $A(L)$ é um polinômio de grau p no operador L , de *lag*:

$$A(L) = 1 - \mathbf{a}_1 L - \mathbf{a}_2 L^2 - \dots - \mathbf{a}_p L^p$$

O processo é estacionário se $A(L)$ tem todas suas raízes fora do círculo unitário; se houver raiz dentro do círculo unitário, o processo é explosivo e evidentemente não estacionário. No entanto, se houver raiz sobre o círculo, ele é não estacionário mas não é explosivo

O teste ADF tem como hipótese nula a presença de raiz unitária no processo AR(p) , e o teste original (não *aumentado*) DF estuda o AR(1): ele determina se $\alpha = 1$ no modelo

$y_t = \alpha y_{t-1} + e_t$, enquanto o teste aumentado verifica se $\alpha_1 = 1$ no modelo

$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + e_t$, e a não rejeição desta hipótese indica a presença de raiz unitária no AR(p).

Nesta aplicação foi usado o valor crítico de MacKinnon (1991) para 1200 observações, e foi usado $p=4$. Para um nível de significância de 5% , o valor crítico é 2,86 (em valor absoluto).

Na tabela 2 estão os valores da estatística do teste obtidos para as doze séries, que indicam a rejeição da hipótese nula em todas elas.

	ORIGINAL	MODIFICADA
TELB	-17,73	-17,18
PETR	-17,35	-16,27
ELET	-18,60	-17,94
BBDC	-17,03	-16,48
VALE	-16,87	-17,56
BRHA	-16,51	-16,29

TABELA 2: Teste de Dickey-Fuller aumentado

A conclusão desta seção é que, se as séries puderem ser representadas por processo estocástico linear, este processo é estacionário.

2.3.3 - MODELAGEM GARCH

Como já foi dito anteriormente, a verificação da estacionariedade de uma série é sempre feita dentro de hipóteses que limitam a generalidade do problema: aqui será suposto que as séries podem ser representadas por processo GARCH(1,1), que é um modelo bastante usado para representar séries de retornos.

Mais adiante neste trabalho serão abordados em detalhe modelos da família GARCH e eles serão usados para estimar volatilidades. No momento, o que está sendo feito é apenas uma verificação de estacionariedade, caso as séries possam ser representadas pelo GARCH(1,1). Este processo é dado por:

$$y_t = \mathbf{s}_t \mathbf{e}_t \text{ com } \mathbf{e}_t \sim N(0,1)$$

$$\mathbf{s}_t^2 = \mathbf{w} + \mathbf{a} y_{t-1}^2 + \mathbf{b} \mathbf{s}_{t-1}^2$$

Ele é estacionário se $(\mathbf{a} + \mathbf{b}) < 1$. Para verificar esta condição, foi ajustado um modelo GARCH(1,1) a cada uma das séries; os coeficientes \mathbf{a} e \mathbf{b} foram sempre significantes, num nível de 5% , e os resultados estão na tabela 3:

	a	b	a+b
otelb	0,159	0,816	0,975
mtelb	0,126	0,850	0,977
opetr	0,076	0,915	0,990
mpetr	0,066	0,926	0,992
oelet	0,106	0,881	0,987
melet	0,096	0,893	0,988
obbdc	0,420	0,497	0,917
mbbdc	0,361	0,560	0,921
ovale	0,225	0,732	0,957
mvale	0,207	0,750	0,957
obrha	0,070	0,911	0,981
mbrha	0,071	0,910	0,981

TABELA 3: Parâmetros do GARCH(1,1)

Embora as somas dos coeficientes sejam elevadas, indicando alto grau de persistência, os valores de $(\mathbf{a} + \mathbf{b})$ são todos menores que um, ou seja, todas as séries podem ser modeladas por GARCH(1,1) estacionário. Note-se, entretanto, que para algumas séries o valor de $\mathbf{a} + \mathbf{b}$ está bastante próximo da unidade, sugerindo que o modelo apropriado poderia ser o IGARCH(1,1), que é o modelo GARCH(1,1) com $\mathbf{a} + \mathbf{b} = 1$. A aplicação de teste formal para esta hipótese não é trivial e está fora do escopo deste trabalho.

No apêndice deste capítulo, seção 2.7.1 , está comentada a questão da persistência e da estacionariedade dos modelos GARCH e IGARCH.

Em vista dos resultados acima, a conclusão desta seção é que, se as séries puderem ser bem representadas por processo GARCH(1,1), este processo é estacionário mas não se descarta a hipótese de que, em algumas séries, a variância possa não existir pois o modelo mais adequado seria o IGARCH (1,1).

2.3.4 – MÉDIA CONDICIONAL

Uma outra maneira de abordar a estacionariedade é através da média condicional:

$$\mathbf{m}_t = E\mathbf{G}_t | y_{t-1}, y_{t-2}, \dots, \mathbf{r}.$$

Se a média condicional não depende do tempo ($\mathbf{m}_t = \text{constante}$) então a média incondicional é igual à condicional e não depende do tempo, logo há *estacionariedade da média*. Neste trabalho, como a média amostral de todas as séries é zero, se a média condicional for constante ela será também nula, portanto será testado $\mathbf{m}_t = \text{zero}$. Note-se que média condicional variável não significa que a média incondicional também seja variável: por exemplo um AR(1) estacionário do tipo $y_t = \mathbf{a} y_{t-1} + \mathbf{e}_t$ tem $\mathbf{m}_t = \mathbf{a} y_{t-1}$ e média (incondicional) zero. Desta forma, a rejeição da hipótese $\mathbf{m}_t = \text{zero}$, não necessariamente implica na não estacionariedade, mas a não rejeição implica na estacionariedade.

Os testes de média condicional são importantes, mais que para verificar estacionariedade, para verificar previsibilidade dos retornos: se a média condicional depende do tempo, a modelagem desta dependência pode levar a modelos que consigam prever retornos a partir de retornos passados.

2.3.4.1 – Teste de Bicorrelação

Hsieh (1989) propõe um teste para a hipótese de média condicional igual a zero, que usa o fato de que, se a média condicional de y_t é zero, então os coeficientes de bicorrelação dados por

$$r_{i,j} = E(y_t \cdot y_{t-i} \cdot y_{t-j}) / \text{var}(y_t)^{3/2}$$

são zero para $i, j \geq 1$. O que é testado de fato no teste de Hsieh é a hipótese nula “o coeficiente de bicorrelação é zero”, contra a hipótese alternativa de que ele é diferente de zero. O coeficiente de bicorrelação amostral é

$$r_{i,j} = \frac{1}{n-i} \sum_{t=i}^n y_t y_{t-i} y_{t-j} / \left(\frac{1}{n} \sum_{t=1}^n y_t^2 \right)^{3/2}, \quad i \geq j;$$

se a hipótese nula é verdadeira e satisfeitas algumas condições de regularidade, $\sqrt{n} \cdot r_{i,j}$ tem distribuição assintótica normal, com média zero e variância que pode ser consistentemente estimada por

$$\hat{s}_{\sqrt{n} \cdot r_{i,j}}^2 = \frac{1}{n-i} \sum_{t=i}^n y_t^2 y_{t-i}^2 y_{t-j}^2 / \left(\frac{1}{n} \sum_{t=1}^n y_t^2 \right)^3, \quad i \geq j$$

O teste então consiste em calcular a estatística $z = \sqrt{n} \cdot r_{i,j} / \hat{s}_{\sqrt{n} \cdot r_{i,j}}$ que, sob a hipótese nula, tem distribuição assintótica normal padronizada. Nas tabelas 4 e 5 estão, para quinze pares i, j , os valores dos coeficientes de bicorrelação amostrais para as seis séries originais e modificadas; em itálico estão as estatísticas z do teste de significância de cada coeficiente isoladamente e a estatística do teste de significância dos quinze coeficientes em conjunto (soma dos quadrados dos quinze valores de z).

Para um nível de 5%, o valor crítico da estatística para cada coeficiente isoladamente é igual a 1,96 e para os quinze coeficientes em conjunto é 25,0 e, nas tabelas, os valores significantes estão marcados com um asterisco.

Na primeira tabela observa-se que para nenhuma série houve significância dos 15 coeficientes em conjunto, OBBDC apresenta três coeficientes significantes e OVALE apresenta um, podendo-se dizer que o teste de maneira geral não rejeita a hipótese de coeficiente de bicorrelação zero para as séries originais.

Na segunda tabela estão os mesmos valores para as séries modificadas; embora os números tenham mudado um pouco, continua a não haver significância conjunta dos 15 coeficientes (embora MBBDC tenha se aproximado do valor crítico), e poucos coeficientes são individualmente significantes.

Concluindo, de maneira geral os coeficientes de bicorrelação não são significativamente diferentes de zero, o que é consistente com a média condicional nula e com a estacionariedade da média incondicional, para todas as séries.

i,j		otelb	opetr	oelet	obbdc	ovale	obrha
1.1	coeficiente teste	0,28 1,43	0,20 0,99	0,33 1,80	0,67 (*) 2,47	-0,27 -1,02	0,22 1,11
1.2	coeficiente teste	0,14 1,52	0,09 0,73	0,07 0,55	0,16 0,87	0,07 0,86	-0,02 -0,27
2.2	coeficiente teste	0,32 1,24	0,18 0,99	0,32 1,10	0,58 (*) 2,04	0,29(*) 2,15	0,33 1,48
1.3	coeficiente teste	-0,03 -0,25	0,02 0,19	-0,03 -0,20	-0,16 -1,41	-0,12 -1,47	-0,16 -1,10
2.3	coeficiente teste	-0,07 -0,54	-0,09 -0,72	-0,26 -1,61	-0,04 -0,28	-0,13 -1,62	-0,01 -0,06
3.3	coeficiente teste	0,12 0,56	0,12 0,67	0,16 1,00	0,08 0,40	0,24 0,88	0,32 1,31
1.4	coeficiente teste	0,05 0,62	0,02 0,17	-0,01 -0,08	0,03 0,29	0,01 0,11	-0,12 -1,23
2.4	coeficiente teste	-0,01 -0,08	-0,01 -0,08	-0,05 -0,34	-0,08 -0,91	-0,06 -0,90	-0,01 -0,13
3.4	coeficiente teste	0,04 0,44	0,01 0,07	0,08 0,81	0,06 0,68	0,08 1,30	-0,05 -0,34
4.4	coeficiente teste	-0,03 -0,14	-0,06 -0,40	-0,02 -0,11	-0,05 -0,55	-0,03 -0,23	0,21 1,41
1.5	coeficiente teste	-0,03 -0,42	0,01 0,14	-0,10 -1,26	-0,07 -0,83	0,04 0,71	-0,10 -1,39
2.5	coeficiente teste	0,08 0,85	0,08 0,58	-0,05 -0,34	0,24 (*) 2,44	0,01 0,15	0,06 0,60
3.5	coeficiente teste	0,02 0,22	0,14 1,46	0,02 0,20	0,00 0,02	-0,03 -0,33	0,08 1,03
4.5	coeficiente teste	-0,11 -1,41	0,07 0,65	-0,04 -0,50	-0,01 -0,11	-0,01 -0,13	0,07 1,02
5.5	coeficiente teste	-0,14 -0,81	-0,04 -0,22	-0,08 -0,49	-0,18 -1,21	0,09 0,94	0,02 0,15
teste 15 coefic.		10,72	6,65	11,42	23,07	16,07	14,47

TABELA 4: Coeficientes de bicorrelação: séries originais

i,j		mtelb	mpetr	melet	mbbdc	mvale	mbrha
1.1	coeficiente	-0,06	-0,06	0,07	0,35 (*)	-0,02	0,32
	teste	-0,59	-0,47	0,55	2,33	-0,23	1,18
1.2	coeficiente	0,08	0,03	-0,08	0,02	0,01	-0,02
	teste	1,14	0,39	-0,94	0,28	0,18	-0,09
2.2	coeficiente	0,05	-0,05	-0,05	0,24	0,16	0,37
	teste	0,43	-0,39	-0,30	1,60	1,72	1,77
1.3	coeficiente	-0,02	0,03	-0,11	-0,04	0,03	-0,22
	teste	-0,39	0,34	-1,46	-0,46	0,62	-1,66
2.3	coeficiente	0,00	-0,06	-0,08	0,00	-0,04	0,14
	teste	0,07	-0,86	-0,96	0,03	-0,63	1,06
3.3	coeficiente	-0,05	0,07	-0,10	0,13	0,02	0,20
	teste	-0,59	0,57	-0,80	0,93	0,21	1,43
1.4	coeficiente	0,03	-0,08	-0,13 (*)	-0,09	-0,02	-0,07
	teste	0,44	-1,29	-2,04	-1,16	-0,30	-0,71
2.4	coeficiente	0,02	0,06	0,08	0,05	-0,05	0,03
	teste	0,25	0,78	0,68	0,72	-0,76	0,39
3.4	coeficiente	0,02	0,05	0,11	0,04	0,04	0,05
	teste	0,27	0,60	1,59	0,58	0,65	0,64
4.4	coeficiente	0,01	0,04	0,04	-0,04	-0,01	0,18
	teste	0,13	0,35	0,34	-0,35	-0,11	1,45
1.5	coeficiente	0,02	-0,01	-0,12	-0,19 (*)	0,06	-0,12
	teste	0,29	-0,16	-1,33	-2,77	1,12	-1,45
2.5	coeficiente	0,10	0,04	0,16	0,11	0,02	0,07
	teste	1,42	0,57	1,61	1,55	0,35	0,80
3.5	coeficiente	0,07	0,08	0,15	0,03	-0,12 (*)	0,04
	teste	0,91	1,21	1,59	0,46	-2,22	0,49
4.5	coeficiente	-0,10	-0,01	-0,04	-0,05	0,02	0,06
	teste	-1,55	-0,13	-0,49	-0,91	0,40	0,85
5.5	coeficiente	-0,11	-0,11	-0,08	-0,13	0,05	-0,01
	teste	-1,09	-1,07	-0,54	-1,29	0,54	-0,06
teste 15 coefic.		9,21	7,41	19,67	24,27	11,76	17,30

TABELA 5: Coeficientes de bicorrelação: séries modificadas

2.3.4.2 – Regressão Localmente Ponderada

Segundo Pemberton e Tong (1981) existem modelos não lineares com coeficientes de bicorrelação nulos e médias condicionais diferentes de zero. Por isto será feita uma outra abordagem à questão da média condicional, através de sua estimação direta por

método não paramétrico. Se y_t é gerado por um modelo do tipo $y_t = g(y_{t-1}, y_{t-2}, \dots) + e_t$, onde e_t é IID e a função g é suficientemente bem comportada, Stone (1977) mostra que métodos de regressão não paramétricos podem ser usados para estimar a função g consistentemente.

O método da regressão localmente ponderada (RLP) foi desenvolvido por Cleveland e Devlin (1988) e foi usado em séries financeiras por Diebold e Nason (1990), Meese e Rose (1991) e Hsieh (1993). Hsieh (1991) usa simulação para mostrar que a RLP pode detectar todos os modelos de dinâmica não linear mais citados na literatura de séries temporais.

O método foi aqui empregado usando apenas um regressor (um *lag* de y_t), ou seja, foi suposto que o retorno é uma função do tipo $y_t = a_{t-1} + b_{t-1} y_{t-1} + e_t$. O método consiste em estimar os coeficientes da regressão, para cada valor de y_{t-1} , usando como amostra apenas os r vizinhos mais próximos (em valor) de y_{t-1} e empregando alguma função de ponderação que dê mais peso às observações mais próximas (em valor) de y_{t-1} .

No apêndice deste capítulo, seção 2.7.2, está detalhado e exemplificado o uso deste método.

A RLP foi usada aqui de forma semelhante ao que foi feito em Hsieh (1993), no que diz respeito à escolha das condições em que são feitas as regressões: o número de vizinhos mais próximos foi arbitrado em 50, 100 e 200 observações. Foi usado o sistema de ponderação com pesos tricúbicos, proposto em Cleveland e Devlin (1988) e foi feita a previsão dos retornos no terço final das séries (400 valores), usando como amostra os 800 valores iniciais.

Foi então determinado o erro quadrático médio destas previsões, e o mesmo foi feito para as previsões baseadas na hipótese de média condicional nula (previsão de y_t igual a zero), cujo erro quadrático médio foi tomado como referência e finalmente foram comparados os erros dos dois tipos de previsão.

Na tabela 6, nas colunas de “quocientes” estão os valores dos quocientes entre os erros quadráticos médios obtidos pelos dois processos (erro da RLP / erro da hipótese de

referência); os valores maiores que um indicam, portanto, que as previsões do método de RLP foram piores que as de referência.

vizinhos:	50		100		200	
	quocientes	estatísticas	quocientes	estatísticas	quocientes	estatísticas
otelb	0,97	1,07	0,96 (*)	2,12	0,96 (*)	3,76
mtelb	1,02	-0,86	1,00	0,24	1,00	0,31
opetr	1,10	-2,82	1,06	-2,95	1,03	-2,54
mpetr	1,03	-0,51	1,01	-0,19	0,99	0,22
oelet	0,94 (*)	1,67	0,95	1,64	0,96	1,52
melet	1,01	-0,29	1,00	-0,06	1,00	0,07
obbdc	0,96	0,35	0,96	0,44	0,95	0,68
mbbdc	0,95	1,33	0,95	1,79	0,95 (*)	1,99
ovale	1,03	-1,60	1,01	-1,41	1,01	-2,28
mvale	1,07	-2,78	1,03	-2,68	1,04	-3,58
obrha	1,00	-0,01	0,99	0,42	0,98	0,78
mbrha	0,98	0,67	0,97	1,38	0,96 (*)	2,03
média	1,00		0,99		0,99	

TABELA 6: Regressão Localmente Ponderada - comparação de erros de previsão

Observando-se os resultados nota-se que, na média, os dois métodos de previsão têm erros equivalentes mas que em algumas séries a RLP consegue erros menores, indicando que, para elas, talvez a média condicional não seja nula (especialmente para BBDC, que também apresentou coeficientes de bicorrelação elevados).

Para verificar se as reduções de erro obtidas pela RLP são significantes, foi empregado o teste de Harvey, Leybourne e Newbold (1997), descrito na seção 2.7.3 , no apêndice deste capítulo.

Na tabela 6, nas colunas de “estatísticas”, estão os valores da estatística deste teste. O valor crítico (em valor absoluto) para um nível de 5% é 1,65 , ou seja, rejeita-se a hipótese de igualdade dos erros em favor da hipótese alternativa (erros da RLP menores que os de referência) se a estatística for maior que este valor crítico. As séries onde os erros de previsão da RLP são significativamente menores que os do modelo de referência estão com quocientes marcados com um asterisco.

Na tabela 6 observa-se que a significância da redução de erro para TELB original desaparece para a série modificada, sugerindo que a redução está mais ligada aos elevados valores dos retornos (e dos erros quadrados de previsão) nos dias atípicos do que à existência de estrutura nos retornos diários, supostamente modelada pela RLP.

Para BBDC e BRHA há redução de erro nas séries originais e nas modificadas, mais acentuada nas modificadas e significativa apenas na RLP de 200 vizinhos.

2.3.4.3 - Estacionariedade da média: conclusão

Como BBDC apresentou alguns coeficientes de bicorrelação individualmente significantes e, nos quinze coeficientes tomados em conjunto, a estatística de teste ficou próxima do valor crítico, há indicação de que a média condicional talvez não seja nula para esta série, já que a RLP também indica alguma estrutura.

Para as demais séries, o conjunto dos dois testes indica média condicional nula e conseqüentemente estacionariedade na média.

Observe-se que, como já foi dito, a média condicional não nula não implica em que necessariamente a média da série não seja estacionária e, portanto, BBDC pode ser estacionária na média, mesmo que se conclua que sua média condicional varia.

2.3.5 – ESTACIONARIEDADE: CONCLUSÃO

As séries aqui analisadas são de observações que oscilam em torno de uma média (zero). O desvio das observações com relação à média varia no tempo, crescendo em algumas ocasiões e diminuindo em outras, e não aparenta um comportamento explosivo.

O teste ADF indica estacionariedade, se o processo gerador dos dados for linear. Se o processo for GARCH(1,1) ele também é estacionário, embora em alguns casos se aproxime do IGARCH e, neste caso, ele não tem variância mas tem a média incondicional constante. Os testes da média condicional, de maneira geral, indicam também estacionariedade na média.

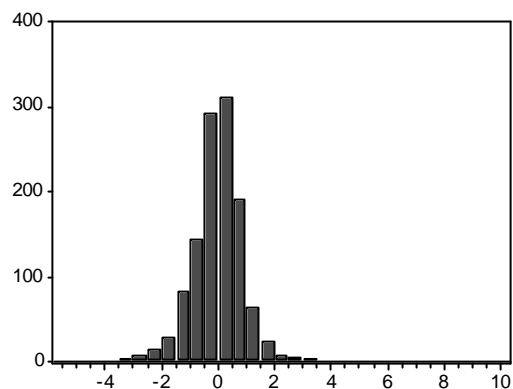
Este conjunto de informações leva à conclusão que as séries são bem modeladas por modelos de média incondicional constante. Quanto à variância, no caso de modelagem não linear, não é possível concluir se as séries são melhor representadas por modelos de variância incondicional constante ou variável.

2.4 – DISTRIBUIÇÃO INCONDICIONAL

Nas figuras a seguir estão apresentados os histogramas e algumas estatísticas descritivas das séries padronizadas.

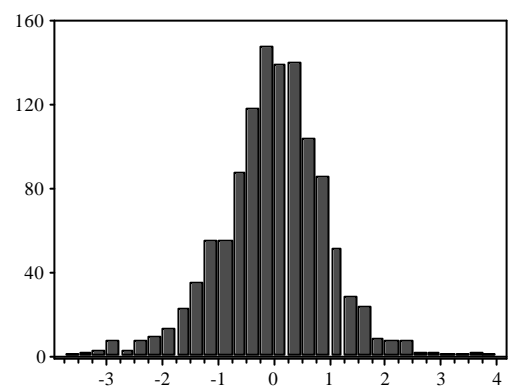
OTELB

MÉDIA	0.000
MEDIANA	0.023
MÁXIMO	9.608
MÍNIMO	-5.299
DESV.PADRÃO	1.000
ASSIMETRIA	1.122
CURTOSE	17.487



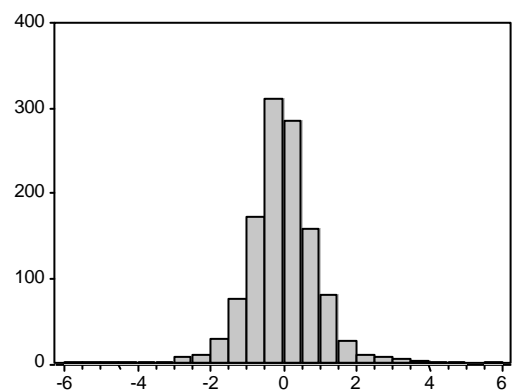
MTELB

MÉDIA	0.000
MEDIANA	0.035
MÁXIMO	3.844
MÍNIMO	-3.594
DESV.PADRÃO	1.000
ASSIMETRIA	-0.167
CURTOSE	4.411



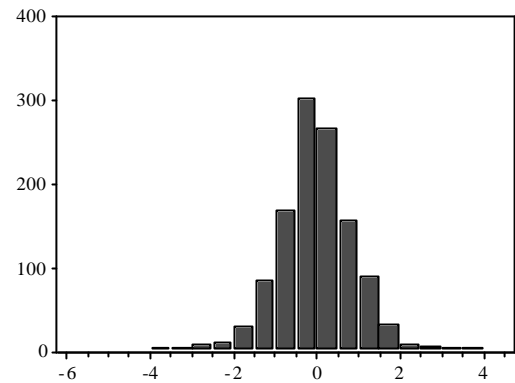
OPETR

MÉDIA	0.000
MEDIANA	-0.026
MÁXIMO	5.535
MÍNIMO	-5.596
DESV.PADRÃO	1.000
ASSIMETRIA	-0.054
CURTOSE	7.778

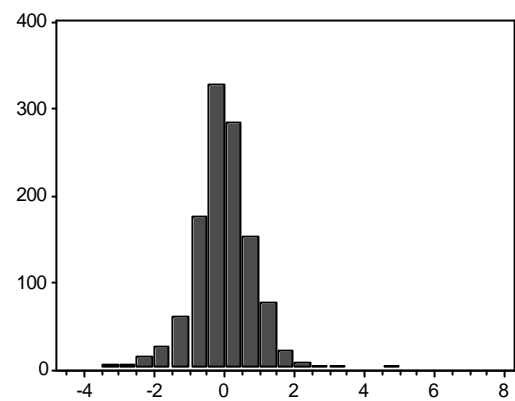


MPETR

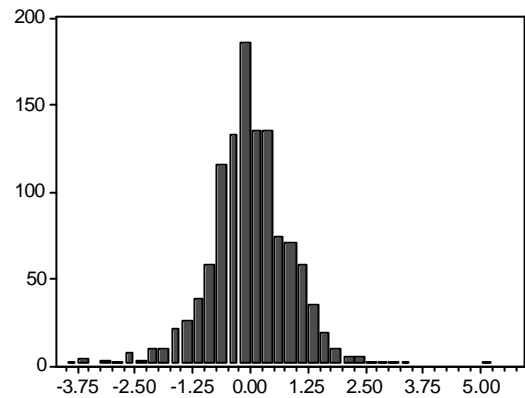
MÉDIA	0.000
MEDIANA	-0.022
MÁXIMO	4.100
MÍNIMO	-5.909
DESV.PADRÃO	1.000
ASSIMETRIA	-0.135
CURTOSE	6.208

**OELET**

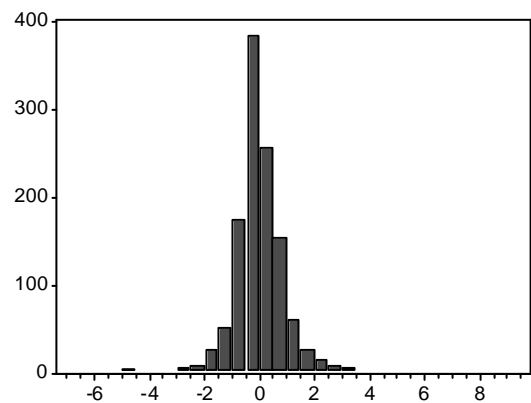
MÉDIA	0.000
MEDIANA	-0.019
MÁXIMO	7.556
MÍNIMO	-4.456
DESV.PADRÃO	1.000
ASSIMETRIA	0.786
CURTOSE	10.566

**MELET**

MÉDIA	0.000
MEDIANA	-0.010
MÁXIMO	5.565
MÍNIMO	-3.834
DESV.PADRÃO	1.000
ASSIMETRIA	0.212
CURTOSE	6.725

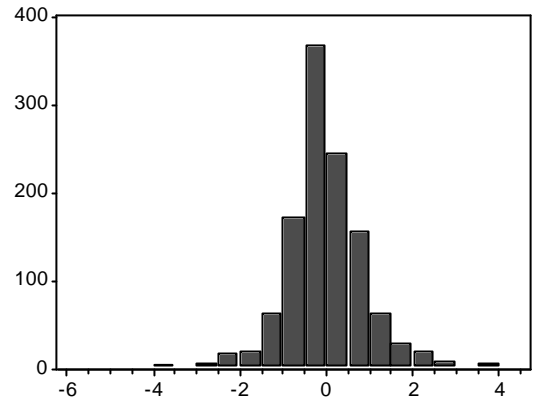
**OBBDC**

MÉDIA	0.000
MEDIANA	-0.038
MÁXIMO	9.132
MÍNIMO	-6.949
DESV.PADRÃO	1.000
ASSIMETRIA	0.211
CURTOSE	13.545

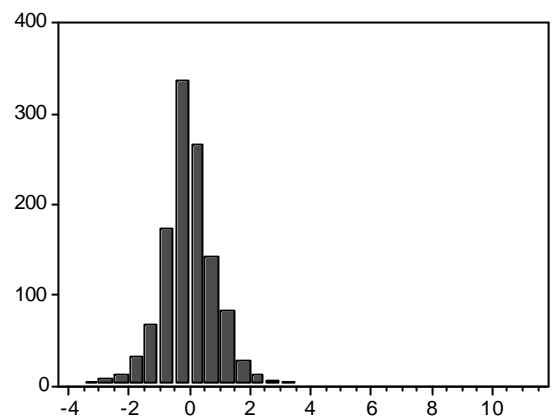


MBBDC

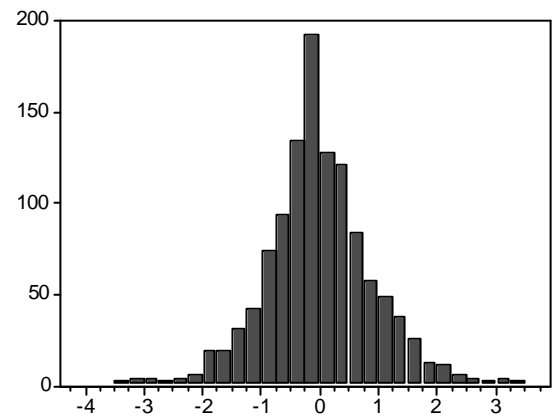
MÉDIA	0.000
MEDIANA	-0.035
MÁXIMO	4.425
MÍNIMO	-5.636
DESV.PADRÃO	1.000
ASSIMETRIA	-0.112
CURTOSE	7.175

**OVALE**

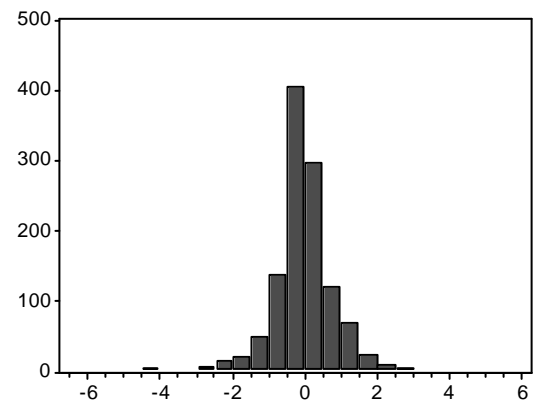
MÉDIA	0.000
MEDIANA	-0.031
MÁXIMO	11.117
MÍNIMO	-3.896
DESV.PADRÃO	1.000
ASSIMETRIA	1.373
CURTOSE	17.685

**MVALE**

MÉDIA	0.000
MEDIANA	-0.016
MÁXIMO	3.685
MÍNIMO	-4.243
DESV.PADRÃO	1.000
ASSIMETRIA	0.008
CURTOSE	4.596

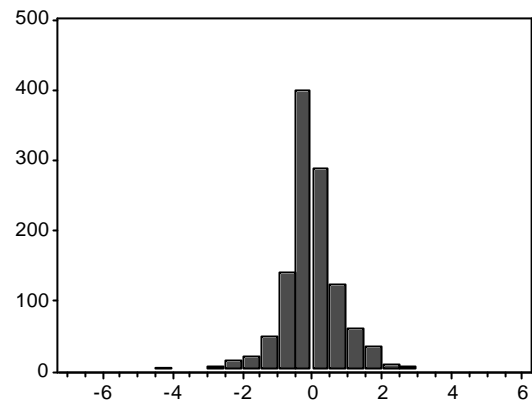
**OBRHA**

MÉDIA	0.000
MEDIANA	-0.051
MÁXIMO	5.813
MÍNIMO	-6.456
DESV.PADRÃO	1.000
ASSIMETRIA	-0.051
CURTOSE	10.845



MBRHA

MÉDIA	0.000
MEDIANA	-0.048
MÁXIMO	5.640
MÍNIMO	-6.695
DESV.PADRÃO	1.000
ASSIMETRIA	-0.344
CURTOSE	10.153



Observando os dados acima pode-se inicialmente notar a grande influência das observações aberrantes na forma da distribuição: as séries modificadas sempre têm coeficiente de curtose menor que as séries originais e, em geral, têm o coeficiente de assimetria mais próximo de zero.

Mesmo após a retirada dos dias atípicos, as séries modificadas têm as caudas bem mais pesadas que as da distribuição normal, como sempre acontece com as séries de retornos de ações. Ao contrário do observado por Talyor (1986) não se nota predominância de valores positivos para o coeficiente de assimetria.

O modelo de Bachelier (1900) e Osborne (1959) assume que o preço das ações varia em decorrência do somatório de um número grande de mudanças que ocorrem transação a transação. Se estas mudanças são independentes, o teorema do limite central leva à conclusão que as variações diárias (ou em outro intervalo de tempo maior) têm distribuição normal, se o número de transações for suficientemente elevado.

Para testar normalidade das distribuições foi usado inicialmente o teste de Jarque-Bera, cuja hipótese nula é coeficiente de assimetria nulo e de curtose igual a três, simultaneamente.

A estatística do teste é

$$S = \frac{b}{n} g^2 + \frac{b}{n} \phi - 3c$$

onde A é o coeficiente de assimetria dado por

$$A = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^3 / \hat{s}^3$$

e C é o coeficiente de curtose dado por

$$C = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^4 / \hat{s}^4 .$$

Nas duas expressões, \bar{y} e \hat{s} são a média e o desvio padrão amostrais.

Se a distribuição é normal, a estatística de teste tem distribuição qui-quadrado com dois graus de liberdade e o valor crítico, para um nível de 5% , é igual a 5,99. Na tabela 7 estão os valores da estatística para as doze séries:

	ORIGINAL	MODIFICADA
TELB	10766	106
PETR	1145	520
ELET	2993	705
BBDC	5581	877
VALE	11181	128
BRHA	3085	2588

TABELA 7: Teste de Jarque-Bera

Um teste simples para normalidade, robusto à presença de observações aberrantes que aumentam excessivamente a curtose, é o teste da *amplitude padronizada*, onde a amplitude da amostra (valor máximo menos valor mínimo) é padronizada pelo desvio padrão, resultando a estatística de teste

$$S = (\max y_i - \min y_i) / \hat{s} .$$

A hipótese nula de normalidade é rejeitada se o valor de S excede os valores críticos obtidos em Pearson e Hartley (1966); o valor crítico para nível de 5% é 2,772 . Na tabela 8 estão os valores da estatística para as doze séries:

	ORIGINAL	MODIFICADA
TELB	14,09	7,44
PETR	11,13	10,01
ELET	12,01	9,40
BBDC	16,08	10,06
VALE	15,01	8,11
BRHA	12,27	12,33

TABELA 8: Teste de amplitude padronizada

Os resultados dos dois testes acima indicam fortemente a rejeição da hipótese de normalidade.

A título de ilustração, estão apresentados na figura 4 os “PxP plots” para TELB e VALE, originais, tomando como referência a normal.

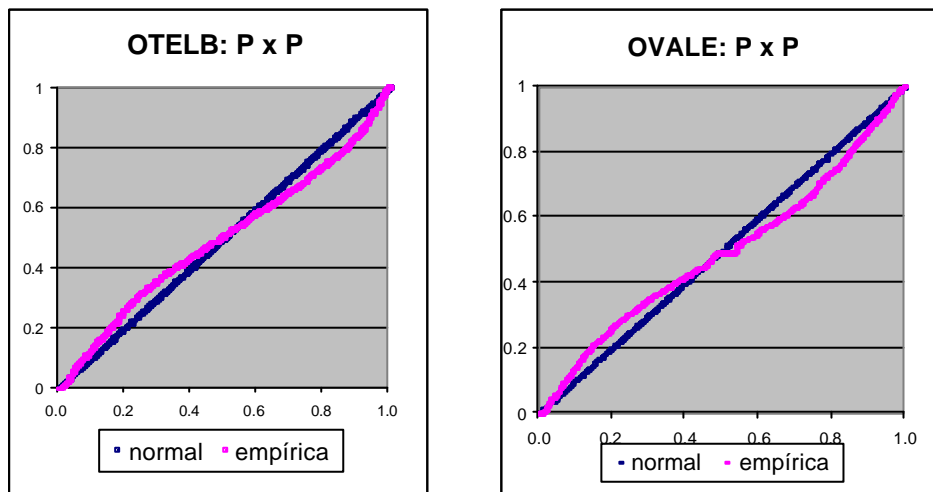


FIGURA 4: “PxP plot” – comparação com a normal

A rejeição de normalidade é um fato que ocorre em praticamente todos os estudos sobre séries de retornos de ações, e muitos autores sugerem que cada mudança de preço seja normal mas que não necessariamente as variâncias destas normais sejam iguais. Logo, para um período de tempo t qualquer (um dia, por exemplo) o retorno será uma variável aleatória X_t que tem distribuição condicional $X_t | w_t \sim N[m, f(w_t)]$, onde m é constante e $f(w_t)$ é a variância condicional.

Desta forma, a série de observações de X_t vem de uma mistura de distribuições normais. A quantidade w_t pode ser o volume negociado (Clark[1973]), a quantidade de informação que chega ao mercado (Tauchen e Pitts [1983]) ou outra variável qualquer.

De qualquer maneira, a variância condicional $f(w_t)$ terá alguma distribuição e, conseqüentemente, X_t terá uma distribuição incondicional. Se, conforme Praetz (1972), a distribuição da variância condicional for a gama invertida, a distribuição incondicional dos retornos será uma t de Student.

Os “PxP plots” foram feitos novamente para as mesmas duas séries, agora tomando como referência a distribuição “t” com quatro graus de liberdade e estão mostrados na figura 5. O número de graus de liberdade foi o que mais aproximou a “t” da distribuição empírica.

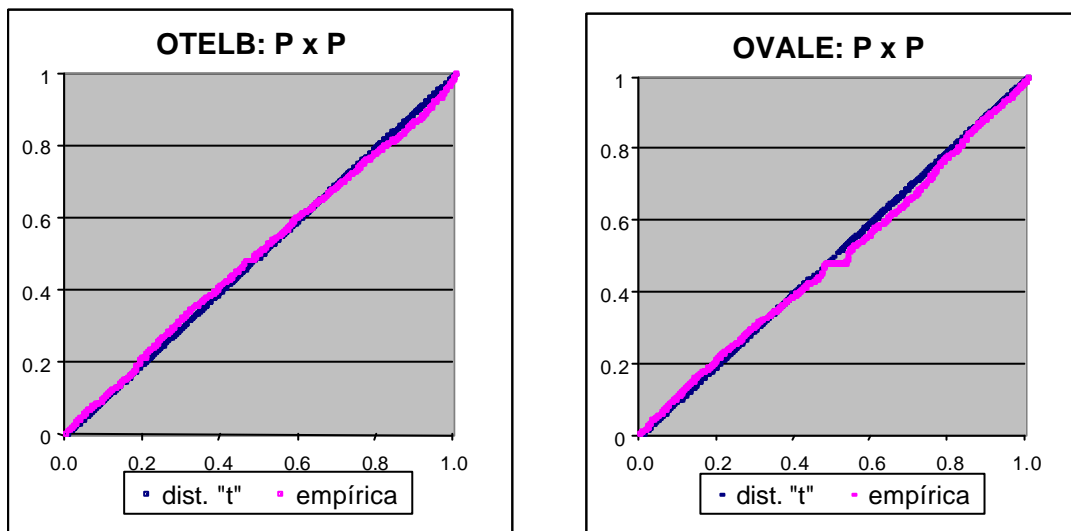


FIGURA 5: “PxP plot” – comparação com a t de Student

O teste de Kolmogorov-Smirnov, embora tenha baixa potência, é útil para verificar a distribuição de um conjunto de dados. Ele é baseado no máximo da diferença entre a função de distribuição empírica dos dados, $F_n(x)$, e a hipotética, $F_0(x)$ (no caso, a “t” com 4 graus de liberdade). Formalmente, este máximo é

$$D_n = \sup_y |F_n(y) - F_0(y)|$$

e, se for válida a hipótese nula de igualdade entre as funções de distribuição,

$$\lim_{n \rightarrow \infty} \Pr \left\{ \sup_{0 \leq t \leq 1} |F_n(t) - F(t)| \leq \epsilon \right\} = 1 - \sum_{i=1}^{\infty} 2^{-i} \exp(-2i^2 \epsilon^2).$$

O valor crítico de D_n para rejeição da hipótese nula, com 1200 observações e um nível de 5% , é igual a 0,0393 . Na tabela 9 estão os valores da estatística para as doze séries; os valores marcados com um asterisco correspondem à rejeição da hipótese de distribuição “t₄”, e esta rejeição ocorre justamente nas seis séries correspondentes aos três papéis menos líquidos.

	ORIGINAL	MODIFICADA
TELB	0.034	0.028
PETR	0.025	0.025
ELET	0.037	0.028
BBDC	0.064 (*)	0.064 (*)
VALE	0.047 (*)	0.041 (*)
BRHA	0.079 (*)	0.073 (*)

TABELA 9: Teste de Kolmogorov-Smirnov para distribuição “t₄”

Observando os “PxP plots” das séries, na figura 5, nota-se que o que afasta a curva empírica de OVALE da curva teórica é a grande quantidade de retornos nulos, e isto ocorre para os três papéis menos líquidos. Por exemplo, enquanto TELB tem 27 retornos nulos, VALE tem 54 , nas 1200 observações.

O retorno nulo ocorre quando o mesmo preço de fechamento se repete em dias consecutivos; note-se que na amostra todos os papéis foram negociados todos os dias (não houve volume igual a zero).

Embora o trabalho de Praetz (1972) forneça uma justificativa teórica para uma distribuição “t” nos retornos, não há evidência de que a distribuição da variância seja gama invertida, e é preferível interpretar os resultados dos testes acima apenas como uma coincidência de formato das funções de distribuição.

A hipótese de normalidade condicional, sem suposição de distribuição gama invertida para a variância, é uma hipótese com mais consistência teórica e será usada mais adiante na estimação dos modelos de variância condicional.

Como conclusão desta seção, pode-se dizer que a distribuição incondicional dos retornos certamente não é normal e o que mais afasta as distribuições empíricas da normal são suas caudas pesadas.

O teste de amplitude padronizada, mais robusto a observações aberrantes, também rejeita a normalidade, por ampla margem, o que indica que a rejeição da normalidade pelo teste de Jarque-Bera não se deve ao fato de o coeficiente de curtose depender da quarta potência dos retornos.

A distribuição empírica se aproxima, em formato, da “t” com quatro graus de liberdade para os papéis de maior liquidez e, para os menos líquidos, dificilmente será possível o bom ajuste de uma distribuição paramétrica, por causa do excesso de retornos nulos.

2.5- INDEPENDÊNCIA

2.5.1 - INTRODUÇÃO

A questão da existência ou não de dependência em séries de retornos é de grande importância e resultou em inúmeras publicações, nem todas chegando às mesmas conclusões. Como o tema está relacionado com a eficiência dos mercados e com a possibilidade de obtenção de ganhos continuados, explorando possíveis padrões de comportamento dos preços, este foi o primeiro enfoque abordado pelos autores.

Atualmente, a maior parte dos autores concorda que é muito difícil prever preços a partir de preços passados, mas a verificação da existência de dependência entre retornos consecutivos e a modelagem desta dependência são importantes para a quantificação (e gerência) do risco.

No campo teórico, uma explicação geralmente aceita para a não normalidade das séries de retornos é o fato de não serem eles independentes e identicamente distribuídos (IID) condicionalmente. Desta forma, os retornos teriam distribuição normal apenas condicionalmente à informação disponível.

A verificação da existência de dependência temporal entre os retornos abre a perspectiva da modelagem desta dependência, e do teste dos resíduos dos modelos para verificação de sua normalidade. Esta verificação é importante porque a normalidade dos resíduos dos modelos significa também normalidade condicional dos retornos, geralmente aceita em Finanças.

Neste capítulo, a dependência temporal entre as distribuições de retornos é inicialmente estudada através do teste da existência de dependência linear, que é aquela que ocorre quando o retorno é representado por uma distribuição que é combinação linear de retornos passados. Verifica-se que, embora exista alguma dependência linear nas séries estudadas, ela não explica a contento o comportamento dos retornos, e são feitos testes que caracterizam a existência de dependência não linear.

2.5.2 - DEPENDÊNCIA LINEAR

2.5.2.1 - Autocorrelações

Na tabela 10 estão os coeficientes de autocorrelação para as séries originais. O tipo de função de autocorrelação é o mesmo para as séries originais e modificadas. Entretanto as observações aberrantes parecem introduzir nas séries um *ruído* que atenua as características das funções de autocorrelação, que são mais acentuadas nas séries modificadas. Por isto, toda a análise de dependência será feita apenas para as séries modificadas; as séries originais apresentam as mesmas características, porém de forma menos nítida.

Lag	mtelb	mpetr	melet	mbbdc	mvale	mbrha
1	0.030	0.160	0.160	0.171	0.044	0.147
2	-0.074	-0.066	-0.069	-0.082	-0.074	-0.025
3	-0.001	-0.034	-0.047	-0.005	-0.055	-0.038
4	-0.040	-0.037	-0.097	-0.023	-0.031	-0.038
5	-0.049	-0.014	-0.085	-0.057	-0.050	-0.034
6	-0.022	-0.005	-0.036	-0.045	-0.011	-0.083
7	0.019	0.020	0.035	0.000	-0.013	-0.020
8	0.065	0.022	0.017	0.068	0.009	0.022
9	0.056	0.035	0.077	0.098	0.052	0.033
10	0.040	0.006	0.064	-0.012	0.021	0.013
11	-0.048	-0.018	-0.016	-0.109	-0.044	0.012
12	0.004	0.034	-0.019	0.004	-0.003	0.078
13	0.027	0.022	-0.003	0.007	0.061	0.036
14	-0.022	-0.022	-0.039	-0.051	0.011	-0.034
15	-0.003	0.031	-0.023	-0.072	-0.050	-0.063

TABELA 10: Função de autocorrelação: séries modificadas

A significância das autocorrelações é testada por meio da estatística de Ljung-Box (1978):

$$Q = n(n+2) \sum_{m=1}^k \frac{r_m^2}{n-m},$$

onde r_m é o coeficiente de autocorrelação de ordem m e são considerados k coeficientes em conjunto. A hipótese nula é que não há autocorrelação até a ordem k e, neste caso, Q tem distribuição assintótica qui-quadrado com k graus de liberdade. Os valores numéricos das estatísticas não estão apresentados.

Observando a tabela 10, pode-se notar que a autocorrelação de *lag* um é positiva para todas as séries. Do *lag* 2 ao *lag* 6 todas as autocorrelações são negativas, para todas as séries. Tomadas em conjunto, as autocorrelações dos m primeiros *lags* (com m variando de 1 a 15 são significantes para todas as séries (exceto MTELB e MVALE para $m=1$), o que implica na existência de dependência, pelo menos linear.

Nos gráficos da figura 6 pode-se visualizar, para duas das séries, o padrão de senóide amortecida das funções de autocorrelação, característico de modelos lineares

autorregressivos; este padrão se repete para todas as séries. As funções de autocorrelação parcial, que não são mostradas aqui, também apresentam padrão de senóide amortecida, característico de modelos de médias móveis.

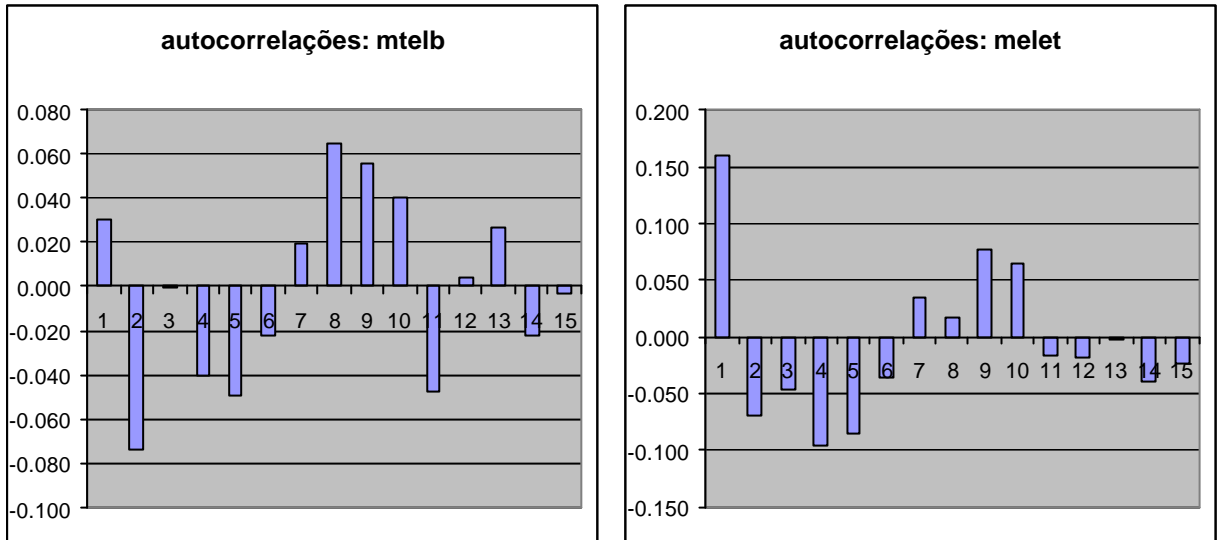


FIGURA 6: Gráfico das funções de autocorrelação para duas das séries

2.5.2.2 - Modelagem ARMA

As funções de autocorrelação sugerem que um modelo ARMA possa representar as séries em estudo; foram estimados modelos ARMA(p,q) para as seis séries modificadas. Embora o formato das funções de autocorrelação possam sugerir valores de p e q maiores, o critério de informação de Schwarz indicou, para a maior parte das séries $p=q=1$. O modelo considerado para todas as séries foi

$$y_t = a + by_{t-1} + ce_{t-1} + e_t$$

onde os valores obtidos para as constantes a , b e c estão na tabela 11:

Em *itálico* está o erro padrão das estimativas; o coeficiente a não é significativo, enquanto que os coeficientes b e c são significantes, para a maioria das séries.

	a	b	c
mtelb	0.000 0.030	-0.676 0.179	0.724 0.168
mpetr	0.000 0.032	-0.277 0.136	0.454 0.126
melet	0.000 0.032	-0.312 0.130	0.491 0.119
mbbdc	0.001 0.032	-0.363 0.107	0.558 0.095
mvale	0.001 0.030	-0.351 0.373	0.409 0.363
mbrha	0.000 0.031	-0.535 0.107	0.655 0.095

TABELA 11: coeficientes do modelo linear

Se existe apenas dependência linear e toda ela foi capturada pelo modelo acima, os resíduos devem ser descorrelacionados.

A tabela 12 mostra os coeficientes de autocorrelação dos resíduos:

	mtelb	mpetr	melet	mbbdc	mvale	mbrha
1	-0.012	-0.002	-0.001	-0.004	-0.013	0.040
2	-0.043	-0.016	-0.012	-0.007	-0.050	0.035
3	-0.020	-0.040	-0.050	-0.026	-0.055	-0.070
4	-0.024	-0.027	-0.075	-0.001	-0.022	-0.013
5	-0.057	-0.010	-0.069	-0.054	-0.051	-0.040
6	-0.013	-0.007	-0.036	-0.029	-0.008	-0.069
7	0.012	0.021	0.040	-0.006	-0.013	-0.020
8	0.067	0.013	0.001	0.056	0.010	0.021
9	0.049	0.031	0.069	0.083	0.051	0.028
10	0.041	0.006	0.055	-0.003	0.020	0.013
11	-0.049	-0.021	-0.018	-0.108	-0.047	0.006
12	0.006	0.031	-0.016	0.023	0.000	0.074
13	0.025	0.023	0.005	0.001	0.057	0.029
14	-0.020	-0.032	-0.041	-0.036	0.010	-0.027
15	-0.004	0.041	-0.010	-0.066	-0.049	-0.061

**TABELA 12: Função de autocorrelação: resíduos ARMA(1,1)
Séries modificadas**

Pode-se ver que os valores das autocorrelações para os primeiros *lags* são em geral negativos e que, em valor absoluto, elas são bem menores que para os retornos.

Entretanto as m primeiras autocorrelações, embora de pequena magnitude, são conjuntamente significantes para m entre 4 e 10, em todas as séries, exceto MPETR.

Os valores das autocorrelações dos resíduos dos modelos ARMA(1,1) aplicado a duas séries estão plotados (barras mais claras) nos gráficos da figura 7. As barras mais escuras são as autocorrelações dos retornos e pode-se ver que o modelo influenciou apenas os valores dos primeiros *lags*, ou seja, a modelo ARMA foi eficaz e reduziu as autocorrelações apenas nos primeiros *lags* mas não conseguiu representar a dependência que produz as autocorrelações nos *lags* mais altos. Isto acontece também para as autocorrelações parciais e nas demais séries.

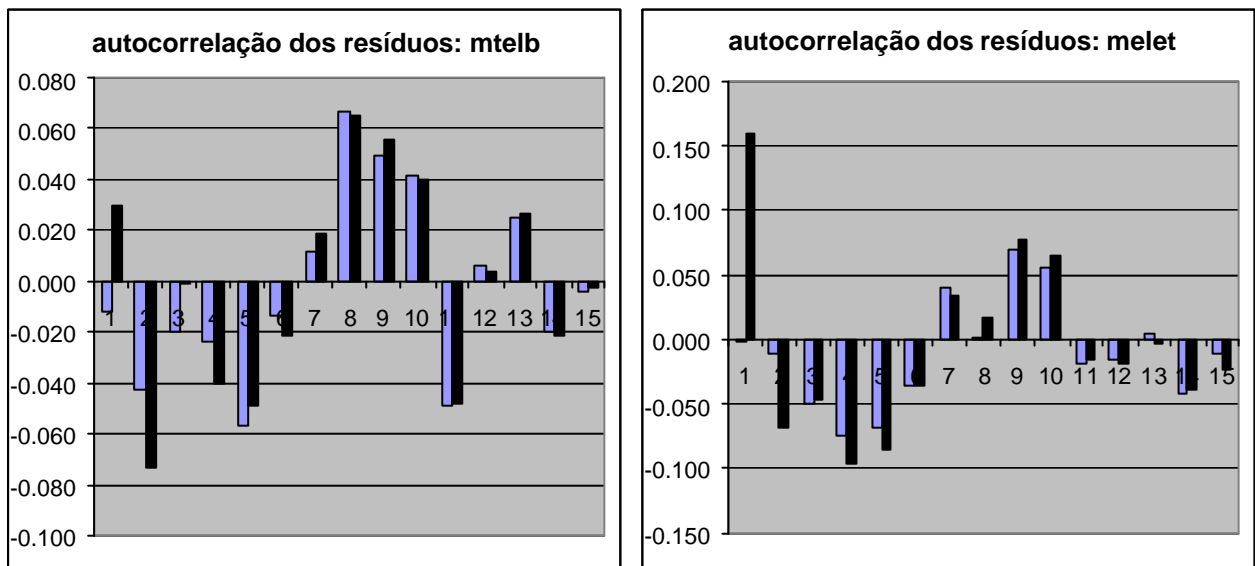


FIGURA 7: Gráfico das funções de autocorrelação dos resíduos ARMA(1,1) para duas das séries

Observe-se que a existência da dependência linear e a significância estatística dos parâmetros dos modelos ARMA não significa necessariamente que haja ineficiência (fraca) no mercado de ações.

Na seção 2.3.4.2 foi visto que o erro quadrático médio da previsão de retorno da regressão localmente ponderada, que é um modelo bem mais geral que o ARMA, não é, em geral, significativamente menor que o erro da previsão ingênua (retorno igual a zero) ou seja, há eficiência fraca.

Na mesma linha, Costa e Baidya (1999) mostram que, no período deste trabalho, não foi possível prever retornos do IBOVESPA a partir de retornos passados.

2.5.3 - DEPENDÊNCIA NÃO LINEAR

2.5.3.1 - Resíduos quadrados

Se os resíduos são independentes então os quadrados dos resíduos também o são. Na tabela 13 estão os coeficientes de autocorrelação dos quadrados dos resíduos do modelo ARMA(1,1):

	mtelb	mpetr	melet	mbbdc	mvale	mbrha
1	0.271	0.233	0.244	0.281	0.206	0.380
2	0.309	0.257	0.284	0.173	0.223	0.186
3	0.230	0.191	0.195	0.187	0.163	0.117
4	0.243	0.194	0.179	0.106	0.243	0.091
5	0.269	0.165	0.269	0.079	0.132	0.067
6	0.206	0.138	0.237	0.135	0.176	0.063
7	0.264	0.207	0.346	0.131	0.118	0.112
8	0.155	0.160	0.137	0.046	0.085	0.099
9	0.212	0.163	0.228	0.135	0.129	0.055
10	0.152	0.134	0.132	0.084	0.109	0.058
11	0.186	0.157	0.124	0.103	0.126	0.056
12	0.231	0.115	0.186	0.084	0.099	0.038
13	0.169	0.146	0.206	0.142	0.137	0.017
14	0.195	0.096	0.192	0.108	0.049	0.047
15	0.138	0.115	0.084	0.055	0.093	0.012

TABELA 13: Função de autocorrelação: resíduos ARMA(1,1) ao quadrado
Séries modificadas

Nota-se que a autocorrelações dos quadrados são bem maiores que as dos resíduos; elas são todas positivas e há uma tendência de declínio com o aumento do valor do *lag*.

Tomadas em conjunto para os m primeiros *lags* (com m variando de 1 a 15), elas são todas significativas. Para que se tenha noção de ordem de grandeza, estão plotados no gráfico da figura 8 as autocorrelações dos resíduos e dos quadrados dos resíduos para a série MPETR; as barras mais escuras correspondem aos resíduos.

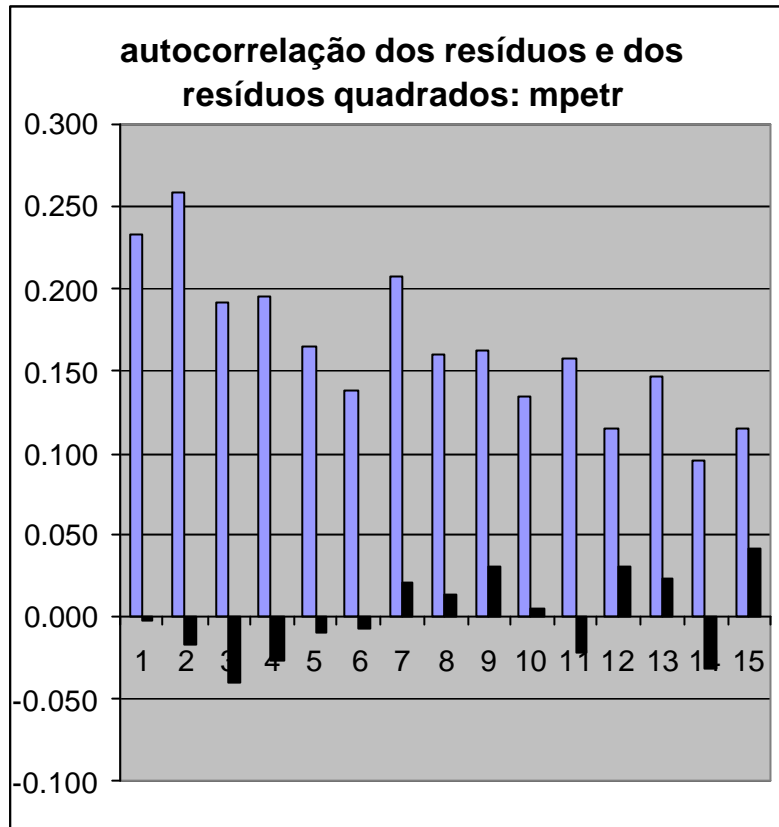


FIGURA 8: Comparação de autocorrelações

As elevadas autocorrelações dos resíduos ao quadrado indicam a existência de dependência temporal não linear entre os resíduos e, conseqüentemente, entre os retornos.

2.5.3.2 - Teste BDS

Este teste foi desenvolvido por Brock, Dechert e Scheinkman (1987) e é usado para testar se as variáveis aleatórias que compõem uma série são IID. Ele pode detectar várias situações em que as variáveis não são IID, tais como não estacionariedade, não linearidade e caos determinístico.

Neste capítulo, ele complementa as conclusões de não linearidade da dependência, tiradas das funções de autocorrelação, no item anterior; mais adiante ele vai ser usado na análise de resíduos de modelos de variância.

A idéia central do teste é analisar os dados sob uma ótica m -dimensional, onde m é suficientemente alto para que se possa distinguir um processo IID de um onde haja dependência não linear. Desta forma, as observações y_t são organizadas em m -histórias

$$y_t^m = \mathbf{I}(y_t, y_{t-1}, \dots, y_{t-m+1}) \mathbf{C},$$

e determina-se a fração de pares de m -histórias que estão perto umas das outras. Isto é feito estimando a distância entre cada par de m -histórias e verificando se ela é menor que um número d arbitrário; esta distância é dada pela máxima norma e não pela distância euclidiana usual.

Assim, para cada par y_t^m e y_s^m , verifica-se se é satisfeita a condição

$$\max_{i=0, \dots, m-1} |y_{t-i} - y_{s-i}| < d.$$

A fração do número total de pares que satisfazem a esta condição é

$$C_{m,n}(d) = \frac{\sum_{s=1}^n \sum_{t=1}^n I_{s,t}}{[n(n-1)g/2]},$$

onde $I_{s,t}$ é um indicador que tem valor um quando a condição é atendida e zero caso contrário, e n é o tamanho da amostra. A *integral de correlação* é definida como o limite desta fração quando a amostra aumenta:

$$C_m(d) = \lim_{n \rightarrow \infty} C_{m,n}(d).$$

Brock, Dechert e Scheinkman (1987) mostraram que se $\mathbf{I}(y_t, \mathbf{C})$ é IID, então

$$C_m(d) = [C_1(d)]^m;$$

Brock e Baeck (1991) mostraram que, sendo válida a hipótese nula de independência, a estatística

$$S = n^{1/2} [C_m \text{bg} C_1 \text{bg}^m] / V^{1/2}$$

tem distribuição assintótica normal padronizada.

Scheinkman e LeBaron (1989) e Hsieh (1991) fornecem expressões explícitas para cálculo do estimador da variância assintótica da expressão entre colchetes, V , que é dada por uma função não trivial da integral de correlação.

Conforme Hsieh (1993), o valor recomendado de d é entre meio e dois desvios padrão, e a precisão da distribuição assintótica se deteriora para altas dimensões, especialmente para m igual a dez ou mais. Na aplicação a seguir, foram arbitrados seis valores para d , entre meio e três desvios padrão, e cinco valores para o número de dimensões, de 2 a 6, resultando 30 testes para cada série. Os testes foram aplicados às séries de retornos (séries modificadas) e aos resíduos dos modelos ARMA.

Os resultados estão apresentados na tabela 14; todos os valores da estatística do teste são significantes para um nível de 5%, que tem valor crítico igual a 1,96, e isto implica na rejeição da hipótese nula, ou seja, as séries não são IID. Esta constatação vem reforçar a conclusão de que existe dependência não linear nas séries, evidenciada anteriormente pelas autocorrelações dos lags mais altos dos retornos e resíduos ARMA e, principalmente, pela autocorrelação de seus quadrados.

Observe-se que os valores da estatística do teste praticamente não se alteram para retornos e resíduos de um mesmo papel, indicando que a dependência linear capturada pelo modelo ARMA é menos importante que a dependência não linear que passa para os resíduos.

Foi visto no capítulo de estacionariedade que o teste de bicorrelação indicou coeficiente de bicorrelação nulo para os retornos. A partir daí pode-se concluir, conforme Cromwell, Labys e Terraza (1994), que a dependência (não linear) existente é do tipo multiplicativo, ou seja, é causada por variância condicional não constante e não por dependência não linear na média condicional.

δ	m	mtelb		mpetr		melet		mbbdc		mvale		mbrha	
		res	ret	res	ret	res	ret	res	ret	res	ret	res	ret
3.0	2	7.4	6.4	5.9	7.5	7.3	6.3	10.1	10.4	5.0	5.3	10.9	11.3
3.0	3	10.4	9.8	9.1	9.9	10.2	9.7	11.7	11.6	7.0	7.3	11.7	11.1
3.0	4	11.7	11.2	10.3	11.0	11.3	10.9	12.2	12.3	8.4	8.2	11.5	11.1
3.0	5	12.5	12.0	11.0	11.7	11.6	11.4	12.2	12.3	9.6	9.6	11.2	10.9
3.0	6	13.3	12.8	11.4	12.0	12.2	11.8	12.0	12.2	10.1	10.3	10.9	10.6
2.5	2	9.1	8.4	7.9	8.3	11.0	9.6	10.6	10.2	6.2	6.3	10.4	10.2
2.5	3	12.1	11.5	11.3	11.1	13.7	12.7	12.6	12.1	8.5	8.7	11.4	10.7
2.5	4	13.6	13.0	12.7	12.4	14.7	14.0	13.4	13.0	9.8	9.7	11.5	11.0
2.5	5	14.3	13.7	13.4	13.2	15.1	14.6	13.6	13.2	10.9	11.0	11.5	11.2
2.5	6	15.0	14.5	13.8	13.6	15.6	15.2	13.8	13.4	11.6	11.7	11.5	11.2
2.0	2	10.6	10.2	9.0	8.4	12.3	11.6	10.3	10.4	7.2	7.4	9.6	10.0
2.0	3	13.5	13.1	12.4	11.4	14.8	14.4	12.6	12.4	9.8	9.9	11.1	11.1
2.0	4	15.2	14.8	14.2	13.0	16.0	15.8	13.7	13.4	11.3	11.3	11.7	11.8
2.0	5	16.3	15.9	15.2	13.9	16.9	16.9	14.2	13.7	12.6	12.7	12.1	12.2
2.0	6	17.4	17.0	16.1	14.5	17.9	17.8	14.6	14.0	13.6	13.6	12.5	12.5
1.5	2	11.2	10.9	9.5	8.9	12.4	12.5	10.2	10.1	8.4	8.8	9.9	9.6
1.5	3	14.2	14.0	12.9	12.2	14.9	15.2	12.9	12.6	11.1	11.4	11.2	10.9
1.5	4	16.3	16.1	15.3	14.7	16.7	17.0	14.2	13.7	13.2	13.4	11.8	11.5
1.5	5	18.3	18.1	16.9	16.4	18.3	18.7	15.0	14.5	15.3	15.5	12.4	12.0
1.5	6	20.4	20.2	18.4	17.9	19.8	20.3	15.9	15.3	17.0	17.2	12.8	12.5
1.0	2	11.4	11.1	10.7	9.9	11.8	12.8	10.2	9.8	9.6	9.8	10.1	9.8
1.0	3	14.9	14.6	14.4	13.6	14.6	15.6	13.1	12.6	12.6	12.6	11.9	11.5
1.0	4	17.7	17.4	18.3	17.3	17.6	18.8	15.1	14.4	15.9	15.8	13.0	12.7
1.0	5	21.1	20.8	21.3	20.0	20.7	22.3	16.5	15.8	19.3	19.3	14.1	13.9
1.0	6	25.2	24.8	25.0	23.2	24.2	26.1	18.0	17.4	22.9	22.9	15.2	15.3
0.5	2	10.7	10.7	11.7	10.8	11.1	12.5	9.3	8.7	10.7	10.9	11.2	10.1
0.5	3	15.0	15.1	16.1	15.3	14.3	15.5	12.3	11.8	14.4	14.5	14.6	12.8
0.5	4	18.6	19.2	22.0	20.9	18.2	19.7	14.8	14.5	19.5	19.4	17.5	15.2
0.5	5	23.9	25.2	27.8	25.9	23.0	25.2	17.1	17.3	26.1	26.2	21.6	18.1
0.5	6	31.3	33.8	35.6	33.8	28.1	31.7	20.1	20.9	35.4	35.5	27.6	21.9

TABELA 14: Teste BDS para resíduos e retornos - séries modificadas

2.5.4 – INDEPENDÊNCIA: CONCLUSÃO

As autocorrelações dos retornos indicam a existência de dependência linear, modelada pelo ARMA(1,1), que reduziu bastante a autocorrelação nos *lags* mais baixos.

Entretanto, passou para os resíduos autocorrelação nos *lags* mais altos, e os quadrados dos resíduos se mostraram altamente correlacionados, indicando dependência não linear nos retornos.

O teste BDS confirma esta dependência e, aplicado aos retornos e aos resíduos do ARMA(1,1), sua estatística tem resultados numéricos semelhantes, indicando que a dependência linear modelada é menos importante que a não linear.

Do teste de bicorrelação pode-se concluir também que a dependência não linear é multiplicativa, causada por estrutura na variância condicional das séries.

2.6 - CONCLUSÃO DO CAPÍTULO

As seis ações aqui estudadas representam, em termos de volume, cerca de 60% do mercado acionário brasileiro no período em consideração, e as suas características gerais mostraram-se bastante homogêneas, podendo-se considerar que estas são, de modo geral, as características do grupo de ações mais líquidas de nosso mercado.

Quanto à estacionariedade, o teste ADF indicou a inexistência de raiz unitária, raiz esta característica de modelos lineares não estacionários, mas como existe dependência não linear entre os retornos, a indicação de estacionariedade do teste ADF não é válida. As séries puderam ser modeladas por GARCH(1,1) estacionário, embora de alta persistência. Embora o teste de bicorrelação indique estacionariedade da média, a regressão localmente ponderada tem resultados inconclusivos quanto a esta estacionariedade em algumas séries.

O conjunto das análises, porém, leva à conclusão de que o processo estocástico gerador das séries pode ser considerado estacionário na média. Quanto à variância, não há uma conclusão: não foram feitos testes de estacionariedade fora das premissas de linearidade (que não se mostrou verdadeira) ou de modelo GARCH (1,1), cujos coeficientes sugerem que o processo gerador de algumas séries pode ser o IGARCH (1,1), sem variância finita.

Como ocorre na maioria dos estudos deste tipo, as distribuições dos retornos apresentaram caudas pesadas incompatíveis com a normal. As séries de retornos de papéis mais líquidos têm função de distribuição com formato análogo ao da *t* de Student.

As autocorrelações dos retornos indicam a existência de dependência linear, modelada pelo ARMA(1,1), que reduziu bastante a autocorrelação nos *lags* mais baixos.

Entretanto, passou para os resíduos autocorrelação nos *lags* mais altos, indicando dependência também não linear nos retornos. Os quadrados dos resíduos se mostraram altamente correlacionados, o que também é indício de dependência não linear, e o teste BDS rejeita a hipótese de retornos IID. A conclusão é que existe dependência linear e, principalmente, não linear multiplicativa.

No próximo capítulo parte-se da informação aqui levantada para tentar modelar a dependência não linear existente entre os retornos.

2.7 – APÊNDICE DO CAPÍTULO 2

2.7.1 – COMENTÁRIO SOBRE A PERSISTÊNCIA E A ESTACIONARIEDADE DOS MODELOS GARCH e IGARCH

A persistência está associada com o tempo durante o qual se fazem sentir os efeitos de uma dada observação. Por exemplo, um choque nos retornos (um valor elevado) aumenta a variância do período seguinte e, voltando os retornos a seu nível usual, a variância tende a voltar a seu nível usual, tão mais rapidamente quanto *menor* for a persistência. A previsão da variância j períodos adiante, feita pelo GARCH(1,1) estacionário é

$$E_t \sigma_{t+j}^2 = b + a g \sigma_t^2 - \frac{w}{1-a-b} \left[1 - \frac{w}{1-a-b} \right];$$

note-se que a fração representa a variância incondicional, s^2 , logo a previsão é

$$E_t \sigma_{t+j}^2 = b + a g \sigma_t^2 - s^2 h s^2.$$

Nesta expressão fica clara a reversão à média, ou seja, quando j é grande, a variância tende a seu valor incondicional, e pode-se perceber que, quanto maior a soma $a+b$, maior a persistência e mais duradouro efeito de um choque.

As séries de retornos, modeladas pelo GARCH, normalmente apresentam elevada persistência, o que levou Engle e Bollerslev (1986) a chamar de GARCH *integrado*, IGARCH, o GARCH(1,1) com $a+b = 1$. Este modelo apresenta propriedades interessantes: Nelson (1990) mostra que, satisfeitas algumas condições, s_t^2 tem distribuição limite estritamente estacionária e ergódica e, conseqüentemente, o retorno dado por $y_t = s_t e_t$ é estritamente estacionário por que é o produto de dois processos estritamente estacionários. Então, curiosamente, o processo IGARCH não é estacionário (de segunda ordem) porque não atende à condição do teorema 1 de

Bollerslev, mas é estritamente estacionário. Isto significa que a distribuição dos retornos tem média zero, mas suas caudas tão espessas que a distribuição não tem variância ou momentos de ordem mais elevada. Outros autores já propuseram distribuições sem variância para os retornos e o mais famoso talvez seja Fama (1965) que sugeriu que os retornos pudessem ser representados por distribuições estáveis paretianas, com expoente característico menor que dois.

É também interessante notar que, ao contrário dos processos lineares com raiz unitária (*integrados*), os estimadores de máxima verossimilhança dos parâmetros do IGARCH são bem comportados, com distribuição assintótica normal, conforme mostra Lumsdaine (1990).

A previsão da variância j períodos adiante, feita pelo IGARCH(1,1) é

$$E_t \sigma_{t+j}^2 = s_t^2 + jw .$$

Pode-se notar que a persistência é tal que o efeito de um choque é permanente (não decresce com j), e que, para j grande, a previsão vai para infinito.

2.7.2 – DETALHAMENTO DO USO DA REGRESSÃO LOCALMENTE PONDERADA

A RLP foi feita usando como amostra os retornos dos dias $t = 1, 2, \dots, 800$ e foram previstos retornos para os dias $t = 801, 802, \dots, 1200$. O método da RLP parte da ordenação da amostra segundo os valores crescentes do retorno: a amostra ordenada dos retornos dos primeiros 800 dias é que serve de base para a determinação dos parâmetros que serão depois usados nas previsões. Para cada um dos 800 retornos, y_t , será determinado um par de parâmetros a_t e b_t ; estes parâmetros são determinados por meio de uma regressão *local*, usando como dados apenas os r vizinhos mais próximos de y_t na amostra ordenada. Neste trabalho foram empregados três valores de r (50, 100

e 200), e a seguir é descrito o procedimento para $r = 50$, já que os outros casos são análogos.

Sem perda de generalidade, pode-se supor que estejam sendo estimados, por exemplo, os coeficientes da regressão para $t = 247$. Não ser usados nesta estimação os 50 valores de retorno mais próximos de y_{247} na amostra dos retornos ordenados; cada um destes 50 retornos ocorreu num determinado dia k e no dia seguinte ocorreu um retorno y_{k+1} .

Usando como dados os 51 pares (y_{k+1}, y_k) , correspondentes a y_{247} e seus 50 vizinhos mais próximos, são estimados os parâmetros da regressão $y_{k+1} = a_{247} + b_{247}y_k + e_{k+1}$ que relaciona linearmente o retorno de um dia com o retorno observado na véspera, na vizinhança do valor observado no dia 247.

Mas a regressão, além de local, é também *ponderada*. Os retornos mais próximos de y_{247} na amostra ordenada, pesam mais na regressão que os mais afastados. Então a estimação de a_{247} e b_{247} é feita por mínimos quadrados ponderados, usando-se um vetor de pesos *tricúbicos* definido a seguir.

O peso do retorno do dia k é dado por $w_k = \frac{|y_k - y_{247}|^3}{\max_k |y_k - y_{247}|^3}$.

Note-se que k não varia de 1 a 51; ele varia no conjunto dos índices t correspondentes aos 50 dias em que o retorno foi mais próximo de y_{247} . No denominador da fração entre parênteses na expressão do peso está a maior diferença de valor entre y_{247} e cada um de seus 50 vizinhos mais próximos, e no numerador está a diferença entre y_{247} e o retorno do dia k . A fração tem seu valor entre zero e um e, conseqüentemente, os pesos também variam entre zero (quando y_k corresponde ao vizinho mais afastado de y_{247} , entre os 50 mais próximos) e um (quando y_k é igual a y_{247}). Na figura 9 estão os pesos usados na regressão do retorno do dia 247 da série OTELB.

O número na abcissa indica o ordenamento dos vizinhos de y_{247} e não o dia t . O número 1 corresponde ao vizinho mais afastado para menos e o número 51 ao mais afastado para mais, entre os 50 vizinhos mais próximos; o número 26 corresponde ao próprio y_{247} .

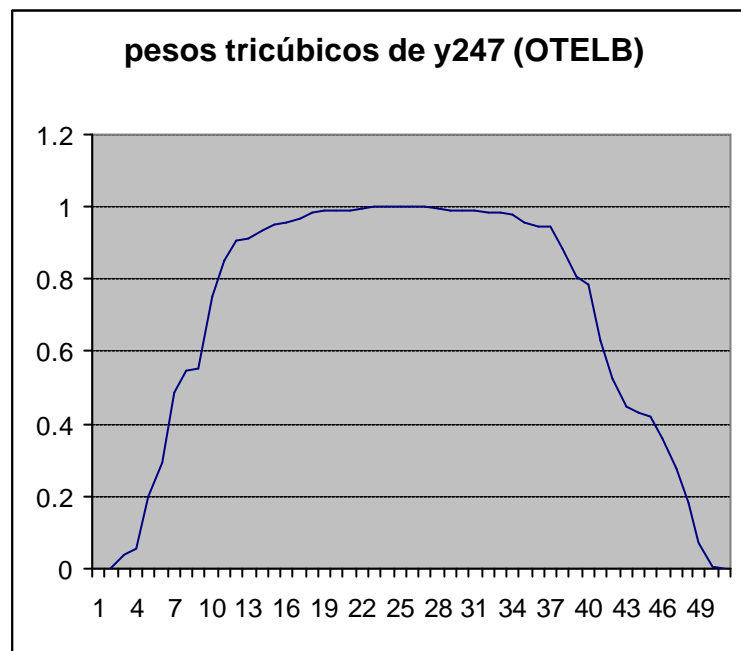


FIGURA 9: Exemplo de pesos tricúbicos

O procedimento descrito acima é repetido para cada um dos 800 retornos iniciais: encontrar os 50 vizinhos mais próximos de cada retorno, determinar os pesos tricúbicos e determinar os parâmetros da regressão. Desta forma é criado um “banco de dados” onde a cada valor de retorno está associado um par de parâmetros de RLP.

O próximo passo é a previsão dos retornos para cada um dos 400 dias finais da amostra: por exemplo, para prever o retorno do dia 917 parte-se do retorno do dia 916 e procura-se no “banco de dados” o valor de retorno mais próximo de y_{916} . A este valor de retorno (por hipótese y_{658}) corresponde um par de parâmetros a_{658} e b_{658} e a previsão de retorno será

$$\hat{y}_{917} = a_{658} + b_{658} y_{916} .$$

O erro desta previsão é simplesmente

$$\mathbf{e}_{RLP,917} = y_{917} - \hat{y}_{917} .$$

A idéia central do método é comparar os erros de previsão da RLP com os erros da previsão baseada na hipótese de que a média condicional é zero. Formalmente esta hipótese se traduz em

$$\mathbf{m}_t = E \mathbf{G}_t | y_{t-1}, y_{t-2}, \dots \neq E \mathbf{G}_t | y_{t-1} \neq 0 ,$$

e a previsão baseada nesta hipótese é

$$\tilde{y}_t = E \mathbf{G}_t | y_{t-1} \neq 0 ;$$

seu erro é

$$\mathbf{e}_{m=0,t} = y_t - 0 = y_t .$$

A comparação é feita através do erro quadrático médio de cada uma das duas previsões: para a RLP ele é dado por

$$EQM_{RLP} = \sum_{t=801}^{1200} \mathbf{e}_{RLP,t}^2 / 400$$

e para a hipótese de média condicional zero ele é

$$EQM_{m=0} = \sum_{t=801}^{1200} \mathbf{e}_{m=0,t}^2 / 400 .$$

Se $EQM_{RLP} \geq EQM_{m=0}$, as previsões feitas pelo método da RLP não são melhores que as baseadas na hipótese de média condicional zero, podendo-se rejeitar a hipótese de média condicional diferente de zero e, finalmente, concluir pela estacionariedade da média do processo gerador dos dados, independentemente do modelo que seja assumido para este processo.

Já se $EQM_{RLP} < EQM_{m=0}$ não se pode rejeitar a hipótese de média condicional zero e, portanto, não se pode concluir a respeito da estacionariedade.

Na tabela 6 , seção 2.3.4.2 , nas colunas de “quocientes” estão os valores dos quocientes entre os erros quadráticos médios obtidos pelos dois processos

$$\frac{EQM_{RLP}}{EQM_{m=0}}.$$

2.7.3 – DETALHAMENTO DO TESTE DE HARVEY, LEYBOURNE E NEWBOLD

Harvey, Leybourne e Newbold (1997) propõem uma variante do teste de Morgan (1939), Granger e Newbold (1977) , MGN daqui em diante, para testar igualdade de erros de previsão um passo à frente. Esta variante, ao contrário do teste original, é robusta à não normalidade dos erros, ou seja, ela apresenta bons resultados ainda que, como acontece nas séries de retornos, os erros de previsão apresentem caudas mais pesadas que as da normal.

O teste MGN original pode ser resumido como segue: supondo que estão sendo comparadas previsões um passo adiante feitas a partir de modelos diferentes, que os erros de previsão não são autocorrelacionados e têm média zero, então a igualdade dos erros quadráticos médios das previsões é equivalente à igualdade das variâncias dos erros. Neste caso, se a distribuição dos erros for normal bivariada, o teste não tendencioso, uniformemente mais poderoso para a hipótese nula de igualdade de variâncias é o teste usual para correlação zero entre

$$u_{1,t} = e_{1,t} - e_{2,t} \quad \text{e} \quad u_{2,t} = e_{1,t} + e_{2,t},$$

onde $e_{1,t}$ e $e_{2,t}$ são os erros das duas previsões, e a estatística de teste é

$$S = \left[\frac{1 - r^2}{n - 1} \right]^{1/2} \cdot r$$

onde

$$r = \left[\sum u_{1,t}^2 \sum u_{2,t}^2 \right]^{-1/2} \cdot \sum u_{1,t} u_{2,t}$$

e S tem distribuição t de Student com $n-1$ graus de liberdade.

No teste modificado, é conveniente entrar num contexto de regressão: considerando-se

$$u_{2,t} = \mathbf{b} u_{1,t} + \mathbf{e}_t ,$$

a estatística S é idêntica à estatística usual para testar $\mathbf{b} = 0$ mas, como não será suposta normalidade dos erros de previsão, calcula-se $\hat{\mathbf{b}}$ por mínimos quadrados e emprega-se, para testar $\mathbf{b} = 0$, a estatística modificada

$$S^* = \frac{\sum u_{1,t}^2 \hat{\mathbf{e}}_t^2 / (n-1)}{\sum u_{1,t}^2} \cdot \hat{\mathbf{b}} ,$$

cuja distribuição sob a hipótese nula não é exata mas se aproxima de uma t de Student com $n-1$ graus de liberdade. Os autores do teste modificado apresentam resultados de simulações que mostram uma boa aproximação, para amostras grandes ($n=512$).

Na tabela 6, nas colunas de “estatísticas”, estão os valores de S^* calculados para o teste MGN modificado; o valor crítico (em valor absoluto) para um nível de 5% é 1,65, ou seja, rejeita-se H_0 (igualdade dos erros) em favor da hipótese alternativa (erros da RLP menores que os de referência) se S^* for maior que este valor crítico.

CAPÍTULO 3 – MODELOS DE VOLATILIDADE

3.1 - INTRODUÇÃO

Este é o capítulo central deste trabalho, onde são descritos, estimados e avaliados 10 modelos de volatilidade, com diferentes graus de elaboração. Procurou-se uma certa concisão na descrição dos modelos e foram indicadas as referências para um maior detalhamento quanto à sua formulação.

A estimação dos modelos foi feita por diferentes processos, de acordo com as características de cada um.

O modelo mais simples, de janela de tempo, como não tem parâmetros não necessita estimação.

Para os modelos de amortecimento exponencial, foi feita uma rotina usando o otimizador 'Solver', do programa EXCEL, para determinar fatores de decaimento ótimos, que minimizam somas de erros quadrados.

Os modelos da família GARCH foram estimados por (quase) máxima verossimilhança, com o auxílio do programa EVIEWS 3.0 que usa uma variante do algoritmo de Berndt, Hall, Hall e Hausman (1974) para chegar iterativamente aos parâmetros; para três destes cinco modelos foi necessário programar a função de verossimilhança.

Os dois modelos de volatilidade estocástica foram estimados por métodos indiretos recomendados por seus autores, e que envolveram também o programa EVIEWS e a rotina Solver.

O objetivo deste trabalho está na modelagem da volatilidade e posterior previsão de valor em risco. A presença de dependência linear nos retornos, constatada no capítulo 2, tem que ser levada em consideração quando se estimam os modelos de volatilidade.

Nos modelos da família GARCH, isto pode ser feito supondo um modelo linear para os retornos, um modelo da família GARCH para os resíduos do modelo linear, e estimando por máxima verossimilhança os parâmetros de ambos os modelos simultaneamente.

Este procedimento, entretanto, não pode ser usado nos outros modelos de volatilidade: é necessário eliminar a dependência linear previamente, aplicando um modelo linear aos retornos e, numa segunda etapa, aplicar o modelo de volatilidade aos resíduos do modelo linear.

Evidentemente, o segundo método é menos eficiente que o outro, mas foi empregado aqui para todos os modelos de volatilidade⁴, para que houvesse uniformidade de critérios na modelagem. Então, para cada uma das 12 séries de retornos padronizados, foi estimado ‘dentro da amostra’ um modelo linear (diferente do ARMA[1,1] do capítulo 2) e os resíduos deste modelo, supostamente livres de dependência linear, são os dados, dentro e fora da amostra, para todos os modelos de volatilidade.

Assim, os modelos de volatilidade foram comparados com o mesmo conjunto de séries servindo como dados para todos eles. Note-se que, numa aplicação real, os modelos da família GARCH levariam a vantagem de poderem ser aplicados diretamente aos retornos, sem a necessidade de eliminação prévia da dependência linear.

Estes resíduos dos modelos lineares serão denominados neste capítulo por x_t , diferentemente dos retornos padronizados, que foram denominados y_t no capítulo 2.

Por questão de simplicidade, os resíduos x_t continuarão sendo chamados de *retornos* neste capítulo, enquanto que os resíduos dos modelos de volatilidade é que serão chamados de *resíduos*.

Desta forma, para todos os modelos de volatilidade, o modelo do *retorno* é da forma $x_t = \hat{\mathbf{S}}_t \mathbf{e}_t$, onde \mathbf{e}_t é independente e identicamente distribuído, com média zero e variância um; neste modelo, toda a dependência em x_t (não linear, multiplicativa) vem da estrutura na variância condicional.

⁴ Exceto o modelo de volatilidade de Parkinson, que não estima as volatilidades a partir dos retornos.

Foi mantida a distinção entre séries originais e modificadas, mas estas só tiveram observações aberrantes removidas ‘dentro da amostra’, para a estimação dos modelos. Fora da amostra não houve retirada de observações aberrantes.

Os modelos foram avaliados dentro e fora da amostra (capacidade explicativa e preditiva), através de resíduos e erros de previsão; como não há um valor ‘observado’ de volatilidade, foram utilizados dois métodos diferentes de comparação entre previsto e ‘real’.

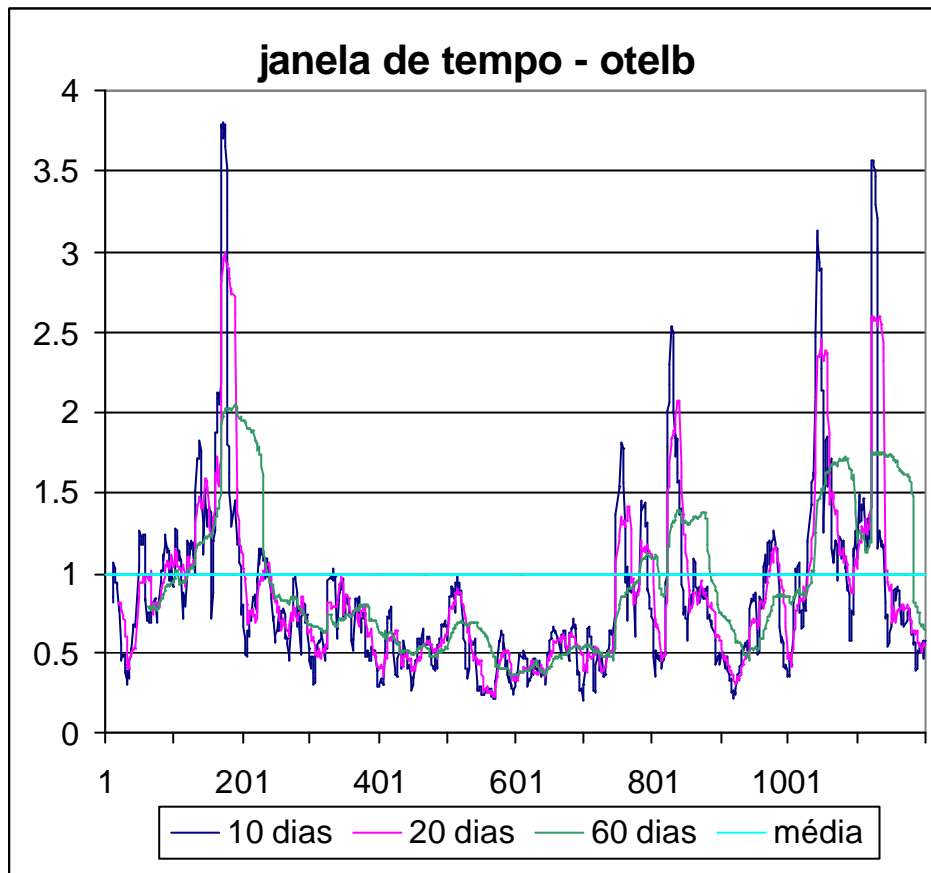
Na seção 3.2, a seguir estão apresentados os dez modelos de volatilidade; na seção 3.3 estão os resultados da estimação dos modelos, enquanto que na seção 3.4 estão os resultados de sua avaliação. A conclusão do capítulo está na seção 3.5 , seguindo-se um apêndice ao capítulo na seção 3.6 .

3.2 – OS MODELOS

3.2.1- JANELA DE TEMPO

Este é o modelo mais simples e intuitivo: a volatilidade é definida como sendo o desvio padrão amostral observado nos últimos k dias. Foram tentados três valores para k (10, 20 e 60 dias de pregão) e concluiu-se que 20 dias (quatro semanas) é o valor que apresenta melhores resultados, simplesmente observando os gráficos com as previsões de volatilidade: as previsões de 10 dias são extremamente “nervosas” e as de 60 dias refletem muito lentamente as oscilações da volatilidade. A figura 10 mostra estes gráficos para uma das séries.

FIGURA 10: Previsões de volatilidade para três janelas de tempo



3.2.2 – AMORTECIMENTO EXPONENCIAL DO RETORNO QUADRADO

No método anterior, a cada dia uma nova observação é incluída na amostra, e a mais distante no tempo é excluída; isto pode ocasionar saltos no valor da volatilidade quando entra ou sai da amostra um retorno elevado (positivo ou negativo). Na figura 10 é fácil observar estes saltos, quando é usada a janela de 60 dias. O método de amortecimento exponencial computa o desvio padrão amostral dando mais peso às observações mais recentes, evitando saltos na volatilidade quando um retorno elevado sai da amostra, e ao mesmo tempo privilegia a informação mais atual. Sua formulação é:

$$\hat{s}_t^2 = (1-q)\sigma^2 + q\hat{s}_{t-1}^2$$

onde: \hat{s}_t = volatilidade estimada para o dia t

x_{t-1} = retorno do dia $t-1$

q = parâmetro, $0 < q < 1$

Através de substituições sucessivas das previsões passadas de volatilidade, chega-se a:

$$\hat{s}_t^2 = \omega - q \hat{s}_{t-1}^2 + q x_{t-2}^2 + q^2 x_{t-3}^2 + \dots$$

Nesta segunda formulação é fácil perceber que a volatilidade ao quadrado é a média ponderada dos retornos ao quadrado passados e que os pesos decrescem numa progressão geométrica, e também que a soma dos pesos tende para a unidade. Também é interessante notar que o uso dos pesos decrescentes resolve o problema da escolha da janela de tempo a considerar: usam-se todos os dados disponíveis mas, de fato, a partir de um certo tempo o peso torna-se tão pequeno que o dado não influencia mais a estimativa da volatilidade.

O valor do parâmetro q foi escolhido, para cada série, de maneira a minimizar a soma dos quadrados das diferenças entre o quadrado do retorno e o quadrado da volatilidade prevista para cada período; seu valor variou entre 0,736 e 0,948.

Os resultados completos da estimação dos parâmetros deste e dos outros modelos estão na seção 3.3 (resultados da estimação).

3.2.3 – AMORTECIMENTO EXPONENCIAL DOS RETORNOS ABSOLUTOS

O modelo supõe que os retornos seguem um processo produto $X_t = V_t U_t$, onde U_t tem distribuição independente normal padronizada e V_t é o processo do desvio padrão condicional, independente de U_t .

Tomando o valor absoluto vem $|X_t| = V_t |U_t|$;

Em termos de valor esperado: $E[|X_t|] = E[V_t]E[|U_t|]$;

Assim, chamando de M_t o processo do valor absoluto do retorno, a previsão do desvio padrão condicional é dada por

$$\hat{s}_t = \hat{v}_t = \hat{m}_t / \sqrt{2/p} .$$

A constante no denominador é o valor esperado do valor absoluto de uma distribuição normal padronizada.

Desta forma, obtendo-se a previsão do valor absoluto do retorno, obtém-se também a previsão do desvio padrão condicional. Este mesmo raciocínio será empregado na seção 3.2.10, para a formulação do modelo de Taylor.

Uma boa previsão do retorno absoluto (Taylor [1986]) é obtida por médias móveis com amortecimento exponencial: $\hat{m}_t = q \hat{m}_{t-1} + (1-q) X_{t-1}$, com parâmetro q , $0 < q < 1$, resultando um modelo análogo ao anterior.

Aqui também o parâmetro q de cada série é estimado de modo a minimizar erro quadrado, resultando em valores entre 0,858 e 0,932.

2.4 – GARCH (p,q)

Este modelo, desenvolvido por Bollerslev (1986) a partir do trabalho pioneiro de Engle (1982) é amplamente empregado em Finanças, especialmente na modelagem de séries de retornos. Ele também assume que os retornos são da forma $X_t = V_t U_t$, onde U_t tem distribuição independente normal padronizada (na verdade, basta que a distribuição tenha média zero e variância um, IID) mas agora o desvio padrão condicional é função determinística dos retornos passados e é dado por

$$\hat{s}_t^2 = w + \sum_{i=1}^q a_i x_{t-i}^2 + \sum_{j=1}^p b_j \hat{s}_{t-j}^2 , \quad \text{com } w > 0 ; a_i, b_i \geq 0$$

Os parâmetros w , a_i e b_j não podem ser negativos, na formulação original de Bollerslev, mas foram estimados aqui sem restrição. Quando ocorreu algum coeficiente negativo, as previsões de volatilidade da série foram inspecionadas, para verificar se havia valores negativos de volatilidade. Os valores de p e q são determinados por critérios de informação (Schwarz, Akaike).

Das 12 séries estudadas, 10 foram melhor representadas pelo GARCH (1,1) e duas pelo GARCH (2,1).

3.2.5 – EGARCH (p,q)

Uma importante variante dos modelos GARCH é o EGARCH (“E” de exponencial), proposto por Nelson (1991). Nele, a variância condicional é dada por:

$$\log \sigma_t^2 = w + \sum_{i=1}^q a_i \left| \frac{x_{t-i}}{\hat{\sigma}_{t-i}} \right| + g_i \frac{x_{t-i}}{\hat{\sigma}_{t-i}} + \sum_{j=1}^p b_j \log \sigma_{t-j}^2$$

Os parâmetros w , a_i e b_j , análogos aos do modelo original, agora podem assumir qualquer valor. A dependência entre a variância condicional e os retornos passados é exponencial e não quadrática; e surgem os parâmetros g_i para modelar a assimetria da variância.

Alguns autores constataram que a variância aumenta mais depois de choques retornos negativos do que positivos, ou seja, para um mesmo valor absoluto de retorno em um dado dia t , a volatilidade em $t+1$, se o retorno em t for negativo, tende a ser maior do que seria se ele fosse positivo. Este é o efeito modelado por valores negativos para os parâmetros g_i , conforme está discutido no apêndice 3.6.1.

A ordem dos modelos EGARCH foi (1,1) para onze séries e (2,1) para uma.

As doze séries aqui analisadas apresentaram valores negativos para a constante de assimetria, sendo onze significativamente diferentes de zero, num nível de 5% ; os valores estimados ficaram entre -0,041 e -0,110.

3.2.6 – HENTSCHEL: GARCH(1,1) GENERALIZADO

Hentschel (1995) desenvolveu um modelo que engloba a maior parte dos modelos da família GARCH. Ele vai ser aqui apresentado, por simplicidade, como generalização do GARCH (1,1) , que é o mais usual em Finanças. Sua formulação da volatilidade é:

$$\frac{\hat{s}_t^I - 1}{I} = w + a\hat{s}_{t-1}^I f^x \mathbf{b}_{t-1} \mathbf{g} b \frac{\hat{s}_{t-1}^I - 1}{I} \quad ; \quad I \geq 0$$

$$\mathbf{e}_t = x_t / \hat{s}_t$$

$$f \mathbf{b}_{t-1} \mathbf{g} |e_t - b| - c \mathbf{b}_{t-1} - b \mathbf{g}$$

Os parâmetros w , a , e b são os equivalentes aos usuais do GARCH (1,1). Não foram impostas restrições sobre estes parâmetros, para garantir a não negatividade da volatilidade. Foi verificado que a volatilidade é sempre positiva, para os valores ótimos estimados dos parâmetros.

A assimetria agora tem dois parâmetros b e c .

Os choques de retorno vão influenciar a volatilidade um período adiante através da função $f(\mathbf{e}_{t-1})$, elevada a uma potência x .

Analisando inicialmente o efeito dos parâmetros de assimetria sobre esta função, pode-se ver que, se $b=0$, o parâmetro c tem a mesma função do \mathbf{g} do EGARCH (com o sinal trocado): ele determina o grau de rotação sofrido pela curva de impacto (vide apêndice 3.6.1).

O parâmetro b desloca o vértice das curvas de impacto, fazendo que um retorno padronizado (dividido pela volatilidade) igual a b tenha impacto zero na volatilidade do período seguinte. Valores de retorno padronizado maiores ou menores que b influenciam a volatilidade, mudando o valor da função f de maneira assimétrica com relação ao ponto $e_t = b$, maneira esta determinada pelo valor de c .

A combinação destes dois parâmetros permite que se modele, por exemplo, uma resposta aproximadamente simétrica para choques pequenos e altamente assimétrica para choques muito grandes. Isto é verdade porque o deslocamento é a fonte dominante de assimetria para choques pequenos, enquanto que a rotação é mais importante para choques grandes.

A influência dos choques de retorno sobre f passa para a equação da volatilidade elevada à potência α . Se este expoente é maior que a unidade, os ramos da curva de impacto deixam de ser retas e se transformam em curvas, convexas com relação ao eixo horizontal (com a segunda derivada positiva); se ele estiver entre zero e um, estas curvas são côncavas. Desta forma, o parâmetro α vai determinar de que forma (elevados a que potência) os choques de retorno, já potencialmente deslocados e girados pela assimetria, irão se refletir na equação da volatilidade.

No GARCH comum, este parâmetro vale 2, ou seja, os choques de retorno entram elevados ao quadrado, na equação da volatilidade (que também está elevada ao quadrado). Já no EGARCH este parâmetro vale 1: os choques de retorno influenciam linearmente o logaritmo da volatilidade.

Da discussão acima, a respeito do parâmetro α , fica claro que sua influência é conjugada com a do parâmetro β , que é o expoente da volatilidade na equação. O parâmetro β pode ser interpretado como o parâmetro de forma de uma transformação de Box-Cox (1964) da volatilidade σ_t .

No GARCH $\beta = 2$, e no EGARCH resulta $\beta = 0$, pois $\lim_{\beta \rightarrow 0} [\sigma_t^\beta - 1/\beta] = \log \sigma_t$.

É importante notar que, conforme ressalta Hentschel (1994) é crucial que $f(\mathbf{e}_t)$ não esteja elevado à potência l , para permitir que o EGARCH seja um caso particular deste modelo.

Os casos particulares mais importantes deste modelo estão mostrados no anexo 3.6.2 .

O modelo de Hentschel foi estimado aqui, como no seu artigo, com uma pequena aproximação para contornar o problema da não diferenciabilidade do valor absoluto dos choques. A aproximação é

$$|\mathbf{e}_t| \cong \sqrt{a^2 + \mathbf{e}_t^2}, \quad \text{com } a=0,001$$

Na estimação do modelo, os parâmetros usuais do GARCH (w , \mathbf{a} e \mathbf{b}) mostraram-se mais estáveis que os de assimetria e os expoentes, que variaram muito de uma série para outra, não havendo indicação de que um único modelo da família possa representar todas as séries, ou a maioria delas.

Houve, em várias séries, problemas de não convergência do algoritmo de estimação, conforme descrito na seção 3.3.6 .

3.2.7 – GRAY: GARCH(1,1) COM TROCA ALEATÓRIA DE PARÂMETROS (COMPLETO)

Este modelo também será aqui apresentado como generalização do GARCH (1,1), mas pode ser usado com quaisquer valores de p e q , e também com outros modelos da família GARCH; a versão a seguir é baseada no trabalho de Gray (1996). A idéia é que a volatilidade segue dois regimes, dependendo de qual dos dois possíveis estados da natureza, $S_t=1$ ou $S_t=2$, ocorre.

Para $S_t=1$, a variância condicional é $\hat{\mathbf{s}}_{t,1}^2 = \mathbf{w}_1 + \mathbf{a}_1 x_{t-1}^2 + \mathbf{b}_1 \hat{\mathbf{s}}_{t-1}^2$

e, para $S_t=2$, vem $\hat{\mathbf{s}}_{t,2}^2 = \mathbf{w}_2 + \mathbf{a}_2 x_{t-1}^2 + \mathbf{b}_2 \hat{\mathbf{s}}_{t-1}^2$.

O estado da natureza não é uma variável observável e é aqui parametrizado como um processo de Markov de primeira ordem, com matriz de probabilidades de transição constante:

$$\begin{aligned} \Pr[S_t = 1 | S_{t-1} = 1] &= P, \\ \Pr[S_t = 2 | S_{t-1} = 1] &= \mathbf{b} - P\mathbf{g} \\ \Pr[S_t = 2 | S_{t-1} = 2] &= Q, \\ \Pr[S_t = 1 | S_{t-1} = 2] &= \mathbf{b} - Q\mathbf{g}; \quad 0 \leq P, Q \leq 1. \end{aligned}$$

Os parâmetros w_i , a_i , e b_i são os equivalentes aos usuais do GARCH (1,1). Aqui também não foram impostas restrições sobre estes parâmetros, para garantir a não negatividade da volatilidade. Foi verificado que a volatilidade é sempre positiva, para os valores ótimos estimados dos parâmetros.

Os parâmetros do modelo são, portanto, os seis coeficientes GARCH (três para cada regime) e as duas probabilidades P e Q.

Os modelos GARCH com mudança de regime apresentam uma característica que torna problemática a sua estimação: a dependência que a variância condicional apresenta com relação a toda a história dos dados. Para evitar este problema, Cai (1994) e Hamilton e Susmel (1994) desenvolvem modelos de variância condicional com troca de regime, nos quais cada regime é caracterizado por um processo ARCH, e não GARCH.

Gray resolve o problema utilizando nas equações da variância condicional ao estado (e à informação disponível), como termo GARCH, a variância passada condicional à informação, mas não ao estado. Este procedimento está detalhado no anexo 3.6.3 .

A estimação deste modelo, cujos resultados estão em 3.3.7 , foi problemática para algumas séries e não houve convergência para as duas séries de Vale do Rio Doce. Em três séries, a solução ótima foi inadequada, conforme mostrado no anexo 3.6.4 .

3.2.8 – GRAY: GARCH (1,1) COM TROCA ALEATÓRIA DE PARÂMETROS SIMPLIFICADO

O modelo de Gray apresentado na seção anterior apresentou alguns problemas na obtenção de soluções ótimas. Visando obter melhores soluções, o modelo foi, aqui neste trabalho, simplificado de maneira a se tornar mais parcimonioso: ele continua, é claro, com dois regimes de volatilidade condicional mas o segundo regime é o primeiro multiplicado por uma constante. Sua formulação é:

Para $S_t=1$, a variância condicional é $\hat{\mathbf{s}}_{t,1}^2 = \mathbf{w} + \mathbf{a} x_{t-1}^2 + \mathbf{b} \hat{\mathbf{s}}_{t-1}^2$

e, para $S_t=2$, vem $\hat{\mathbf{s}}_{t,2}^2 = k \hat{\mathbf{s}}_{t,1}^2$.

Desta forma, os três parâmetros do GARCH da segunda equação do modelo original de Gray são substituídos pela constante k , e o modelo passa a ter 6 parâmetros, contra 8 do modelo completo.

Este tipo de modelo GARCH com troca de regime, onde um regime é proporcional ao outro, é comum na literatura: por exemplo, Almeida e Pereira (1999) empregam um modelo deste tipo aos retornos diários do índice da Bolsa de Valores de São Paulo, comparando os resultados obtidos com os de modelos GARCH tradicionais.

Os resultados obtidos, em termos de convergência foram melhores, pois o modelo convergiu (com soluções ótimas adequadas) para as duas séries de Vale do Rio Doce, onde o modelo completo não convergira. Porém as três séries que apresentaram soluções inadequadas continuaram com os mesmos problemas.

3.2.9 – VOLATILIDADE DE PARKINSON

Este modelo é baseado em trabalho de Hsieh (1993), que descreve a distribuição condicional dos retornos como sendo composta de uma parte previsível, modelada

parametricamente por este modelo de volatilidade autorregressiva, e uma parte imprevisível, modelada não parametricamente por uma distribuição empírica.

Aqui a volatilidade não é determinística e, novamente, um processo produto é suposto para os retornos, $X_t = V_t U_t$. Supõe-se, também, que o processo estocástico de $\{V_t\}$ seja um AR(p) nos logaritmos das volatilidades diárias, v_t . Desta forma, o modelo é especificado por:

$$x_t = v_t u_t$$

$$\log v_t = a + \sum_{i=1}^p b_i \log v_{t-i} + h_t$$

com h_t IID. A medida da volatilidade diária, como em Kupiec (1990), é o estimador de desvio padrão por extensão, de Parkinson (1980):

$$v_t = 0,361 \times 1440 / M \sqrt{\log \max_t / \min_t}$$

onde \max_t e \min_t são o preço máximo e mínimo da ação no dia t , e M é o número de minutos em um dia de pregão. O desvio padrão de Parkinson é uma medida *ex post* da volatilidade e, nesta aplicação, interessa o valor *ex ante*, ou seja, a previsão da

volatilidade condicional, dada por $\log v_t = \hat{a} + \sum_{i=1}^p \hat{b}_i \log v_{t-i}$, onde \hat{a} e \hat{b}_i são os parâmetros estimados da regressão do desvio padrão de Parkinson nos seus *lags*.

O desvio padrão por extensão, de Parkinson, envolve, além dos preços máximo e mínimo diários, uma constante 0,361 que multiplica o inverso da fração do dia em que o mercado funciona (já que 1440 é o número de minutos em um dia). Esta constante pode não ser a mesma para todos os mercados, ou mesmo para todos os papéis de um mesmo mercado. O valor desta constante, por resultar em um fator de escala que multiplica todas as volatilidades *ex post*, vai também resultar em um fator de escala para a volatilidade condicional *ex ante*. Consequentemente, quando se determina o resíduo do modelo (ou retorno re-escalado) $\hat{u}_t = x_t / \hat{S}_t$, o valor daquela constante também vai resultar em um fator de escala para o resíduo. Entretanto, a forma da distribuição de \hat{u}_t e, por conseqüência, os quantis da distribuição condicional do

retorno (que é que Hsieh busca em seu artigo, para quantificar risco) não dependem do valor que é assumido para a constante.

Mais adiante neste trabalho, os modelos de volatilidade vão ser empregados para determinar valor em risco e, como no artigo de Hsieh, os resultados obtidos independem do valor daquela constante. Entretanto, ainda dentro deste capítulo vão ser feitas análises de capacidade preditiva dos modelos de volatilidade e, numa comparação previsto *versus* real, o valor da constante passa a importar. Desta forma, para permitir comparação de erros de previsão, os valores de volatilidade *ex ante* obtidos por este modelo foram multiplicados, em cada série, por uma constante determinada de forma a minimizar a soma, para todos os dias dentro da amostra, dos quadrados das diferenças entre a previsão de volatilidade e o retorno absoluto da série.

Outra particularidade deste modelo é que os dados para ele foram os retornos *verdadeiros*, e não foi aplicado modelo ARMA para eliminar dependência linear entre retornos nem foi feita a sua padronização (para obter séries de média zero e variância um). Foram usados estes dados porque os valores de volatilidade *ex ante* envolvem os preços máximo e mínimo de cada dia, e os únicos retornos consistentes com estes preços são os verdadeiros, calculados diretamente a partir dos preços de fechamento diários.

A estimação deste modelo é muito simples e envolve apenas a determinação dos coeficientes dos modelos AR(p). O valor de p para as diversas séries variou de 2 até 5 .

3.2.10 – TAYLOR: VOLATILIDADE LOGNORMAL AUTORREGRESSIVA

Este modelo é proposto por Taylor (1986) e, como no anterior, a volatilidade não é determinística e, novamente, é suposto um processo produto para os retornos,

$X_t = V_t U_t$. A distribuição de $\log(V_t)$ é suposta Normal ($\boldsymbol{m} \boldsymbol{d}^2$); esta distribuição é conveniente por permitir a obtenção de resultados matemáticos simples, e foi considerada satisfatória e melhor que distribuições alternativas em modelagens

empíricas de Clark (1973) e Tauchen e Pitts (1983). Os parâmetros m e d não podem ser estimados por máxima verossimilhança, porque a densidade multivariada de X_1, X_2, \dots, X_n não pode ser avaliada e, assim, será usado o método aproximado de estimação descrito a seguir.

Taylor (1986) observa, partir de dados empíricos, que as autocorrelações do processo $\log V_t$ podem ser representadas por uma função da forma $r_{t, \log(V)} = f^t$, onde f é uma constante e t é a defasagem. O modelo mais simples que apresenta autocorrelações deste tipo é o AR (1): $\log V_t = f \log V_{t-1} + h_t$.

Taylor mostra que as autocorrelações do processo gaussiano $\log V_t$ são aproximadamente proporcionais às do processo não gaussiano V_t , e que estas são proporcionais às do processo dos valores absolutos dos retornos, $|M_t|$, que é um processo observável. Assim, pode-se concluir que as autocorrelações de $|M_t|$ são aproximadamente proporcionais às de $\log V_t$, assumidas da forma f^t , ou seja, $r_{t, M} \cong k r_{t, \log(V)} = k f^t$.

Este é o ponto central do modelo de Taylor: se o processo de V_t é lognormal autorregressivo, então $r_{t, M} = k f^t$ (aproximadamente).

Esta função de autocorrelação do retorno absoluto é compatível com um processo ARMA(1,1) para $|M_t|$ e, conhecidos k e f , é trivial determinar os parâmetros deste ARMA(1,1).

Determinados estes parâmetros, é possível utilizar este processo para prever \hat{m}_t e, utilizando a relação $\hat{s}_t = \hat{m}_t / \sqrt{2/p}$, já vista em 3.2.3, chega-se finalmente à previsão da volatilidade.

Taylor propõe que k e f sejam estimados de maneira a minimizar a soma dos quadrados das diferenças entre as correlações teóricas dos retornos absolutos e suas correlações amostrais, para as j primeiras defasagens.

Formalmente, o processo é determinar k e f para minimizar a função

$$F(k, f) = \sum_{t=1}^j c_{t,m} - k f^t h, \text{ com } 0 < k, f < 1 .$$

E estimação deste modelo é muito simples e não apresentou problemas.

No anexo 3.6.5 está explicado o procedimento para estimação dos parâmetros e previsão de volatilidades.