

Universidade Federal do Rio Grande do Sul

Instituto de Matemática

Programa de Pós-Graduação em Matemática

**Análise e Estimação de Medidas de Risco em  
Processos FIEGARCH**

por

**TAIANE SCHAEGLER PRASS**

Porto Alegre, Março de 2008.

Dissertação submetida por Taiane Schaedler Prass<sup>1</sup> como requisito parcial para obtenção do grau de Mestre em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professor Orientador:

Prof<sup>a</sup> Dr<sup>a</sup> Sílvia Regina Costa Lopes

Banca Examinadora:

Prof. Dr. Artur Oscar Lopes (PPGMat-UFRGS)

Prof<sup>a</sup> Dr<sup>a</sup> Chang Chiann (IME-USP)

Prof. Dr. Rafael Rigão Souza (PPGMat-UFRGS)

Prof<sup>a</sup> Dr<sup>a</sup> Sílvia Regina Costa Lopes (PPGMat-UFRGS)

Data da Defesa: 21 de Março de 2008.

---

<sup>1</sup>Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## AGRADECIMENTOS

*Aos meus pais pelo apoio incondicional desde o início de minha jornada como estudante. Minha mãe, exemplo de mulher, que me ensinou a enfrentar os problemas de cabeça erguida e sem hesitar sequer um instante e meu pai, que sempre ensinou que o que importa é quem nós somos e não o que temos. Obrigada.*

*A todos de minha família que acreditaram em mim e sempre me apoiaram, em especial, aos meus avós, Ivo (in memoriam) e Loni, pelo amor, carinho e dedicação. Por me ensinarem tantas coisas. Por serem mais pais do que avós. Obrigada.*

*Em especial, agradeço a Sílvia, orientadora e amiga, pelo voto de confiança, e que sempre disposta a auxiliar, com seus conselhos, orientação e apoio, ensinou-me muito mais do que a teoria. Obrigada.*

*Aos amigos Raquel e Arthur e ao pequeno Henrique, pelo apoio desde a fase inicial até a fase final do mestrado. Pela acolhida na minha chegada, pelas horas de carinho e atenção. Pelo apoio nos momentos difíceis. Pelo chimarrão de manhã e de noite. Pelas jantas e almoços nos dias de correria. Certamente devo muito a vocês. Obrigada.*

*A todos os colegas do Programa de Pós-Graduação em Matemática da UFRGS. Aos colegas e amigos do LCPM, em especial Cléber Bisognin, Marcus Alexandre Nunes, Márcio Valk e Pablo Cristini Guedes que me deram tanto apoio e auxílio, principalmente na reta final da dissertação. Obrigada.*

*A todos os amigos e amigas, em especial ao Carlos Eduardo Souza Ferreira pelas tantas horas de diversão e pelo apoio naqueles momentos difíceis. A minha querida amiga Tatiele Fabiana Sell, que mesmo distante esteve sempre tão presente. Você não imagina como você fez falta nesses dois anos. Ninguém “discute” exercícios como você. A minha amiga querida, um anjo que apareceu na minha vida, Paula Leão, para os amigos, Paulinha. Mais do que professora de Inglês, foi quem me fez sorrir quando eu só sentia vontade de chorar.*

*A todos os professores da pós com quem tive aulas, por todo o conhecimento transmitido. Obrigada.*

*A Rosane, nossa querida secretária do Programa de Pós-Graduação, que conhece meus passos desde meus primeiros dias de UFRGS. Obrigada.*

*Ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pelo apoio financeiro.*

*A ele, que ficou por último nesta lista, não por ser o menos importante, mas simplesmente porque não tenho palavras pra descrever o quanto sou agradecida. Mesmo que eu repetisse tudo que já disse anteriormente, não seria o suficiente. Ao meu noivo, o grande amor da minha vida, ao melhor amigo, companheiro de todas as horas. Que mesmo estando em outro país esteve comigo todos os dias, me apoiando, com suas palavras, cartas, e-mails com seu amor incondicional, com sua atenção, com horas de conversa pelo skype, com aquele sorriso lindo que me faz sentir tão especial. A você Guilherme Pumi por ser tão especial e me ensinar que o amor existe sim, que vale a pena sonhar, que vale a pena esperar. Por ser tão especial e me ensinar a viver e ser eu mesma. Obrigada.*

*Por fim quero agradecer a todas aquelas pessoas que, de uma forma ou de outra, contribuíram para esta fase da minha vida. Obrigada a todos.*

Este trabalho é dedicado a Ivo Osvaldo Prass (*in memoriam*).  
Mais do que meu avô, meu primeiro professor!

...Quando a paixão não dá certo  
não há porque me culpar...  
E recomeço do zero sem reclamar...  
...As pessoas se convencem  
de que a sorte me ajudou  
mas plantei cada semente  
que o meu coração desejou...

# Resumo

Neste trabalho apresentamos os principais resultados encontrados na literatura, relacionados a processos ARCH, GARCH e EGARCH. Descrevemos os processos FIEGARCH e introduzimos resultados relacionados a estacionariedade e a ergodicidade desses processos. Retomamos os resultados clássicos, relacionados aos métodos de identificação/seleção de modelos para séries temporais, os principais métodos de estimação dos parâmetros e apresentamos vários resultados relacionados a previsão. Apresentamos os principais conceitos relacionados às medidas de risco e aquelas utilizadas na literatura. Como aplicação, apresentamos a estimação de medidas de risco, tanto para processos FIEGARCH simulados como para séries temporais reais.

# Abstract

In this work we present the main results found in the literature regarding ARCH, GARCH and EGARCH processes. We describe the FIEGARCH processes and introduce results related to stationarity and ergodicity of these processes, among other interesting results. We review the classical results and methods in model identification/selection for time series, the main methods of parameter estimation and several important results concerning the forecasting. We also present the main concepts related to risk measures and the one considered in the literature. As an application of the whole methodology, we present the estimation of risk measures for both simulated FIEGARCH processes and observed time series.

# Sumário

<b>Introdução</b>	<b>8</b>
<b>1 Conceitos Preliminares</b>	<b>12</b>
1.1 Processos Estocásticos . . . . .	12
1.2 Funções e Séries de Potências . . . . .	21
1.3 Processos com Longa Dependência . . . . .	22
1.4 Processos Ergódicos . . . . .	25
<b>2 Processos Não-Lineares</b>	<b>27</b>
2.1 Processos ARCH . . . . .	27
2.2 Processos GARCH . . . . .	30
2.3 Processos IGARCH e Modelos EWMA . . . . .	38
2.4 Processos EGARCH . . . . .	39
2.5 Processos FIEGARCH . . . . .	45
<b>3 Identificação, Estimação e Previsão</b>	<b>61</b>
3.1 Identificação do Modelo . . . . .	61
3.1.1 Análise da Função de Autocorrelação Amostral . . . . .	61
3.1.2 Teste de Normalidade Jarque-Bera . . . . .	62
3.1.3 Critérios AIC, BIC e HQC . . . . .	62
3.1.4 Seleção do Modelo Baseado nos Erros de Previsão . . . . .	63
3.1.5 Testes de Adequação do Modelo . . . . .	64
3.1.6 Estatísticas de Teste para Detectar Longa Dependência . . . . .	65
3.2 Métodos de Estimação . . . . .	69
3.2.1 Método de Estimação da Máxima Verossimilhança . . . . .	69
3.2.2 Método de Estimação da Pseudo-Máxima Verossimilhança . . . . .	70
3.3 Estimação dos Parâmetros dos Processos Não-Lineares . . . . .	72
3.3.1 Processos ARCH, GARCH, IGARCH e EGARCH . . . . .	72
3.3.2 Modelos EWMA . . . . .	73
3.3.3 Processos FIEGARCH . . . . .	73
3.4 Previsão . . . . .	75
3.4.1 Previsão em Processos Estacionários . . . . .	75
3.4.2 Previsão da Volatilidade em Processos ARCH e GARCH . . . . .	77
3.4.3 Previsão da Volatilidade em Modelos IGARCH e EWMA . . . . .	79
3.4.4 Previsão da Volatilidade em Processos EGARCH e FIEGARCH . . . . .	80

<b>4</b>	<b>Medidas de Risco</b>	<b>84</b>
4.1	Medidas de Risco Coerentes . . . . .	85
4.1.1	Axiomas de Coerência . . . . .	85
4.2	Retornos . . . . .	87
4.2.1	Fatos Estilizados sobre os Retornos . . . . .	89
4.3	Seleção de <i>Portfolios</i> Eficientes . . . . .	91
4.3.1	Linha de Mercado de Capitais - CML . . . . .	93
4.3.2	Modelo de Precificação de Ativos - CAPM . . . . .	95
4.3.3	Interpretação do Coeficiente $\beta$ . . . . .	96
4.3.4	Ações Subavaliadas e Superavaliadas. Coeficiente $\alpha$ da Ação . . . . .	97
4.4	Fatores de Risco e Distribuição de Perdas . . . . .	97
4.4.1	Função de Distribuição Condicional e Não-condicional das Perdas . . . . .	98
4.5	Medidas de Risco . . . . .	100
4.5.1	Medidas de Risco Baseadas na Distribuição de Perdas e Ganhos . . . . .	101
4.5.2	Cálculo do Valor em Risco - $\text{VaR}_p$ . . . . .	102
4.5.3	Cálculo do <i>Expected Shortfall</i> - ES . . . . .	109
4.6	Análise de Cenários - Teste de Estresse . . . . .	110
4.6.1	Descrição do Teste de Estresse . . . . .	111
4.6.2	Avaliação do <i>Portfolio</i> . . . . .	111
4.6.3	Identificação dos Máximos e Mínimos para Fatores de Risco Individuais . . . . .	112
4.6.4	Perda Máxima - <i>MaxLoss</i> . . . . .	112
4.6.5	Alguns Critérios de Admissão para Cenários . . . . .	113
<b>5</b>	<b>Simulações</b>	<b>115</b>
5.1	Geração de Processos FIEGARCH . . . . .	115
5.2	Ajuste de Modelos FIEGARCH . . . . .	118
5.3	Estimação do Parâmetro de Longa Dependência . . . . .	120
5.4	Previsão . . . . .	125
5.5	Estimação da Medida de Risco VaR . . . . .	128
5.6	Exemplo de Análise de um <i>Portfolio</i> . . . . .	132
5.6.1	Características das Séries Temporais Simuladas . . . . .	132
5.6.2	Ajuste e Previsão para as Séries Simuladas . . . . .	138
5.6.3	Cálculo dos Pesos do <i>Portfolio</i> . . . . .	147
5.6.4	Estatísticas Descritivas dos Ativos do <i>Portfolio</i> . . . . .	147
5.6.5	Modelo CAPM para as Séries Temporais do <i>Portfolio</i> . . . . .	148
5.6.6	Cálculo das Medidas de Risco VaR e ES . . . . .	150
5.6.7	Análise de Cenários para as Séries Temporais Simuladas . . . . .	154
5.6.8	Construção dos Cenários Utilizando Dados Históricos . . . . .	155
5.6.9	Perda Máxima sob Domínios de Admissibilidade . . . . .	157
<b>6</b>	<b>Análise de Séries Temporais Reais</b>	<b>158</b>
6.1	Características das Séries Temporais . . . . .	158
6.2	Ajuste dos Modelos FIEGARCH . . . . .	166
6.2.1	Previsão da Volatilidade . . . . .	168
6.3	Estatísticas Descritivas dos Ativos Financeiros . . . . .	169
6.3.1	Cálculo dos Pesos do <i>Portfolio</i> . . . . .	170
6.3.2	Modelo CAPM para as Séries Temporais . . . . .	170
6.4	Cálculo das Medidas de Risco VaR e ES . . . . .	172

6.5	Análise de Cenários (Teste de Estresse) . . . . .	178
6.5.1	Construção dos Cenários Utilizando Dados Históricos . . . . .	178
6.5.2	Perda Máxima sob Domínios de Admissibilidade. . . . .	180
<b>7</b>	<b>Conclusões</b>	<b>181</b>
	<b>Referências</b>	<b>186</b>
	<b>Apêndice A</b>	<b>190</b>
	<b>Apêndice B</b>	<b>193</b>
	<b>Apêndice C</b>	<b>202</b>
	<b>Apêndice D</b>	<b>204</b>
D.1	Método Clássico de Análise da Média-Variância . . . . .	204
D.2	Cálculo dos Pesos do <i>Portfolio</i> Eficiente . . . . .	205
D.3	Cálculo do Coeficiente $\beta$ e do $\alpha$ de um Ativo . . . . .	206
D.4	Cálculo do Coeficiente $\beta$ de um <i>Portfolio</i> . . . . .	208
	<b>Apêndice E</b>	<b>210</b>
E.1	Ajuste dos Modelos EGARCH e FIEGARCH à Série das Perdas do <i>Portfolio</i> para as Séries Temporais Simuladas . . . . .	210
E.2	Ajuste de Modelos EGARCH aos Ativos do <i>Portfolio</i> das Séries Temporais Simuladas . . . . .	213
	<b>Apêndice F</b>	<b>216</b>
F.1	Modelos para as Séries Temporais dos Log-retornos dos Ativos do <i>Portfolio</i>	216
F.2	Modelos para a Série Temporal dos Log-retornos do <i>Portfolio</i> . . . . .	221



# Introdução

Atualmente, uma linha de pesquisa amplamente difundida é aquela que apresenta o estudo de modelos não-lineares para séries temporais, visando aplicação, principalmente, à análise daquelas relacionadas ao mercado financeiro. A análise de séries temporais financeiras é, em geral, análoga àquelas que ocorrem em outras áreas, como meteorologia, oceanografia, etc. Entretanto, séries temporais financeiras apresentam uma importante característica conhecida como *volatilidade*, que pode ser definida de diversas formas, mas não é diretamente observável.

Em termos financeiros, a volatilidade está associada ao risco de um ativo. Como o risco precisa ser mensurado, a volatilidade acaba sendo a medida utilizada. Por volatilidade podemos entender a medida estatística da possibilidade do valor de um ativo cair ou subir muitas vezes, de forma significativa, em um determinado período de tempo. Como medida de risco a volatilidade pode ser calculada de várias formas. A medida mais utilizada, mas não a única, para medir o risco é a variância ou o desvio padrão da rentabilidade histórica de um determinado investimento.

Para modelar séries temporais na presença de grupos (*clusters*) de volatilidade é necessário recorrer a modelos heteroscedásticos condicionais. Estes modelos consideram a variância de um retorno em um dado instante de tempo como uma função que depende de retornos passados e de outras informações disponíveis até aquele momento, obtendo assim, uma *variância condicional*. Essa variância condicional é diferente da variância global (não-condicional) da série observada pois varia com o tempo.

Entre os modelos não-lineares mais conhecidos estão os da família ARCH (*Autoregressive Conditional Heteroskedasticity*) e suas extensões. Tais modelos são utilizados para modelar a variância condicional das séries temporais. Os modelos autoregressivos com heteroscedasticidade condicional, denotados por ARCH( $p$ ), foram introduzidos por Engle (1982), com o objetivo de estimar a variância da inflação. Esses modelos assumem que, dado um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , as variáveis aleatórias são não-correlacionadas, mas a volatilidade (variância condicional) depende dos  $p$  valores anteriores do processo por meio de uma função quadrática.

Sabe-se que, no caso dos modelos lineares, os modelos ARMA podem ser mais parcimoniosos, no sentido de apresentar menor número de parâmetros, do que um modelo autoregressivo (AR) ou um média móvel (MA). Sob esse ponto de vista, Bollerslev (1986) introduz os modelos GARCH( $p, q$ ) (*Generalized ARCH*), que são uma generalização dos modelos ARCH( $p$ ). A generalização dos processos ARCH, através dos processos GARCH, ocorre no sentido que a volatilidade, depende tanto dos quadrados dos  $p$  valores anteriores do processo como dos  $q$  valores das volatilidades passadas.

Apesar dos modelos ARCH e GARCH serem amplamente utilizados para modelar séries de retornos financeiros, tais modelos apresentam uma deficiência. Estes modelos

tratam os retornos de forma simétrica, já que a volatilidade é uma função quadrática dos mesmos. Levando em conta que a volatilidade reage de forma assimétrica aos retornos, tendendo a ser maior para retornos negativos, Nelson (1991) introduz a classe de modelos EGARCH( $p, q$ ) (*Exponential GARCH*). Estes modelos apresentam uma representação da forma exponencial para a volatilidade.

Assim como no caso linear, onde os modelos ARFIMA surgem como uma generalização natural dos modelos ARMA, no caso não-linear os modelos FIGARCH (Baillie *et al.*, 1996) e FIEGARCH surgem como generalizações dos modelos GARCH e EGARCH, respectivamente. Uma das vantagens dos processos FIEGARCH em relação aos processos FIGARCH, que não são apresentados neste trabalho, é a sua estacionariedade. Um estudo mais detalhado sobre os processos ARFIMA e FIEGARCH (entre outros processos não lineares), incluindo a geração de séries temporais e estimação dos parâmetros, pode ser encontrado em Lopes (2008) e em Lopes e Mendes (2006).

O modelo FIEGARCH (*Fractionally Integrated Exponential Generalized Autoregressive Conditional Heteroskedasticity*) foi introduzido por Bollerslev e Mikkelsen (1996). Além da volatilidade variando com o tempo e presença de *clusters* de volatilidade (características já presentes em processos ARCH e GARCH), o modelo leva em conta a longa dependência da volatilidade (assim como no modelo FIGARCH), além da assimetria (a parte exponencial do modelo EGARCH).

Entre as principais aplicações dos modelos da família ARCH, podemos citar o cálculo de medidas de risco para séries temporais financeiras, onde tais modelos são utilizados em combinação com modelos ARMA, para estimar a média e a variância condicional de uma série temporal para um determinado horizonte de tempo.

Atualmente, um dos problemas mais importantes em finanças é avaliar os riscos de uma posição financeira. A questão central em gerenciamento de riscos são as *medidas de risco* que são utilizadas, na prática, para vários propósitos. Entre eles, podemos citar a *determinação do capital em risco*, isto é, mede-se a exposição aos riscos em uma instituição financeira, para determinar a quantidade de capital necessário para honrar possíveis perdas inesperadas no *portfolio*. As medidas de risco também são úteis quando deseja-se limitar os riscos aos quais uma instituição financeira está exposta.

Como ressaltamos anteriormente, a volatilidade pode ser vista como uma medida de risco. Quando definimos a volatilidade como sendo a variância da rentabilidade histórica de um investimento, obtemos uma medida de volatilidade absoluta. Essa medida varia com o período de tempo determinado, de forma que na hora do cálculo, o período escolhido é de vital importância.

Além de medidas de volatilidade absoluta, existem também formas de avaliar a volatilidade de forma relativa, ou seja, em relação à volatilidade do mercado, por exemplo. A medida mais usada é o coeficiente  $\beta$  do modelo CAPM. Esse coeficiente mede a volatilidade de um ativo frente a um índice de mercado.

Os métodos mais utilizados no mercado para calcular medidas de risco baseiam-se na suposição de distribuição normal para os processos de inovação. Um dos problemas apresentados neste método é que essa função de distribuição possui caudas mais leves que as observadas em séries temporais financeiras, subestimando assim as perdas. Uma medida de risco muito utilizada é o VaR (Valor em Risco). Até pouco tempo atrás, os modelos para VaR assumiam que os retornos de ações eram normalmente distribuídos.

Muitas distribuições alternativas vêm sendo propostas na literatura. Mesmo utilizando distribuições de caudas mais pesadas para modelar os retornos financeiros, sabe-se que as medidas de risco ficam subestimadas para probabilidades de ocorrência pequenas (eventos extremos).

Outra ferramenta amplamente utilizada para medir os riscos de um *portfolio* são os denominados *testes de estresse*. Os testes de estresse não baseiam-se em hipóteses sobre a função de distribuição das mudanças de fatores de risco. Por esse motivo não são influenciados por caudas pesadas. Como tais testes não quantificam a probabilidade de ocorrência de cenários individuais, eles são utilizados como ferramenta auxiliar para verificar e complementar as estatísticas do tipo medidas de risco como, por exemplo, o VaR. Mesmo assim, os cenários precisam ser minimamente plausíveis e para isso é necessária uma idéia, ainda que vaga, das probabilidades de ocorrência de cada cenário.

A medida *perda máxima*, introduzida por Studer (1997) pode ser vista como uma maneira sistemática de realizar um teste de estresse. Esta é uma medida de risco que pode ser interpretada como a pior perda possível que pode ocorrer em um *portfolio* em um determinado instante de tempo. Entretanto, em alguns casos, o pior cenário pode não existir, dado que a função para calcular o valor do *portfolio* pode ser não-limitada inferiormente. Como a probabilidade de ocorrência de um cenário no qual o mercado a muito tempo não se encontra é muito baixa, para o cálculo dessa medida restringe-se a atenção à cenários em um certo domínio de admissibilidade, também denominado *região de confiança*.

Esta dissertação tem por objetivo apresentar uma revisão bibliográfica sobre os processos não-lineares, citados anteriormente, sobre os principais métodos de seleção de *portfolios* eficientes e apresentar as principais medidas de risco utilizadas na literatura, envolvendo teoria e aplicações. Procuramos expor os assuntos de maneira formal, apresentamos as provas de quase todos os resultados presentes no texto, muitas delas inexistentes na literatura. Para os resultados que já foram amplamente discutidos, apresentamos uma indicação de onde encontrar tais provas. Além disso, apresentamos também aspectos da teoria que não estão a disposição em livros sobre o assunto, mas que se encontram em artigos da área. Sempre que possível, procuramos apresentar provas mais detalhadas do que as apresentadas nos respectivos artigos, principalmente, para os processos FIEGARCH, cuja literatura é excessiva.

O trabalho é dividido como segue. O Capítulo 1 trata dos principais conceitos necessários à compreensão e ao estudo dos modelos para séries temporais. Apresentamos as idéias básicas sobre processos estocásticos e as principais definições relacionadas a eles, tais como estacionariedade, ergodicidade, funções de autocovariância, autocorrelação e densidade espectral. Bem como, as definições das medidas de assimetria e curtose. Também relembramos algumas definições relacionadas aos processos lineares e teoremas que são estendidos aos processos não-lineares. Além disso, lembramos alguns resultados básicos sobre séries de potências necessários para o desenvolvimento dos próximos capítulos.

No Capítulo 2 apresentamos alguns dos modelos não-lineares mais utilizados na análise de séries temporais que apresentam heteroscedasticidade. Primeiramente, discutimos os modelos ARCH e GARCH. Apresentamos as principais características e os principais resultados relacionados a tais processos. Em seguida, tratamos dos modelos EGARCH. Finalmente, com o objetivo de modelar séries temporais que, além de apresentarem volatilidade, também apresentam característica de longa dependência, definimos os processos

FIEGARCH. Para estes processos provamos diversas propriedades importantes, tais como estacionariedade, ergodicidade, causalidade e inversibilidade. Além disso, provamos que se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo FIEGARCH( $p, d, q$ ) então, o processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, p$ ) e, sob certas condições, o processo estocástico  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, 0$ ) com inovações que, em geral, são não-Gaussianas. Esses resultados não aparecem na literatura. Eles são de grande importância pois auxiliam na seleção dos modelos para as séries temporais.

No Capítulo 3 relembramos os principais métodos, apresentados na literatura, para identificação de um modelo qualquer para uma série temporal. Tratamos da estimação dos parâmetros dos modelos, apresentamos os métodos de estimação de pseudo-máxima verossimilhança e quase-máxima verossimilhança e a relação entre eles. O método de estimação da quase-máxima verossimilhança é o mais utilizado atualmente para a estimação dos parâmetros dos processos não-lineares. As propriedades assintóticas do estimador de quase-máxima verossimilhança para modelos ARCH e GARCH foram analisadas por diversos autores, entre eles, Lee e Hansen (1994), Lumsdaine (1996), Berkes *et al.* (2003), Berkes e Horváth (2003) e Hall e Yao (2003). Resultados para o estimador de quase-máxima verossimilhança para modelos EGARCH podem ser encontrados em Straumann e Mikosch (2003). No caso de modelos FIEGARCH, apesar do estimador de quase-máxima verossimilhança ser amplamente utilizado na prática, resultados assintóticos ainda permanecem um problema em aberto (ver Palma, 2007). Por último apresentamos resultados relacionados à previsão de valores futuros para o processo estocástico a partir do modelo ajustado.

No Capítulo 4 apresentamos um contexto probabilístico para a modelagem de riscos financeiros. Apresentamos o método de média-variância para seleção de *portfolios* eficientes e o modelo de precificação de ativos (CAPM), que é utilizado como ferramenta para análise da correlação dos ativos com o mercado financeiro. Além de definirmos distribuição de lucros e perdas, fatores de risco e funções de risco, apresentamos a definição de medida de risco coerente, bem como as principais medidas de risco utilizadas na literatura.

O Capítulo 5 é destinado à apresentação dos resultados obtidos a partir da análise de séries temporais simuladas. Essas séries temporais foram geradas a partir de processos FIEGARCH( $p, d, q$ ), com parâmetros semelhantes aos dos modelos ajustados às séries temporais reais analisadas no Capítulo 6. Além da estimação dos parâmetros dos modelos e da previsão da volatilidade, calculamos também as medidas de risco VaR e *Expected Shortfall* (ES), para essas séries temporais. Apresentamos, além das análises citadas anteriormente, um exemplo de análise completa de um *portfolio* composto por ativos cujos log-retornos são representados pelas séries temporais simuladas.

O Capítulo 6 trata da estimação e análise de medidas de risco para um *portfolio* de ações. O *portfolio* é formado por ações de quatro empresas brasileiras, sendo elas: Bradesco, Brasil Telecom, Gerdau e Petrobrás. Tais ações são negociadas na Bolsa de Valores de São Paulo (Bovespa). Os valores do *portfolio* do mercado são representados pelo índice Bovespa. Os métodos que utilizamos neste capítulo, para o cálculo das medidas de risco, são aqueles descritos no Capítulo 4. Também utilizamos os processos FIEGARCH, apresentados no Capítulo 2, para obter a variância condicional dos log-retornos dos ativos e calcular as medidas de risco VaR e ES.

As conclusões finais referentes a esta dissertação e as propostas para futuros trabalhos são apresentadas no Capítulo 7.

# Capítulo 1

## Conceitos Preliminares

Neste capítulo apresentamos diversos conceitos necessários à compreensão e ao estudo dos modelos para séries temporais discutidos neste trabalho. Primeiramente, apresentamos as idéias básicas sobre processos estocásticos. Apresentamos a definição do momento de ordem  $r$ , para  $r \in \mathbb{N} - \{0\}$ . Definimos as funções de autocovariância e autocorrelação, que nos fornecem informações a respeito do grau de dependência entre quaisquer duas variáveis aleatórias de um processo. Definimos a função densidade espectral e as medidas de assimetria e curtose. Um estudo mais completo pode ser encontrado em Brockwell e Davis (1991) e Wei (1990). Além disso, lembramos alguns resultados básicos sobre séries de potências necessários para o desenvolvimento dos próximos capítulos. Mais detalhes sobre séries de potências podem ser encontrados em Ahlfors (1979) e Lima (2006).

### 1.1 Processos Estocásticos

Nesta seção apresentamos definições importantes para a análise de processos estocásticos e séries temporais.

**Definição 1.1. (Processo Estocástico).** Um *processo estocástico* é uma família de variáveis aleatórias  $\{X_t\}_{t \in T}$  definidas em um mesmo espaço de probabilidades  $(\Omega, \mathcal{F}, \mathbb{P})$ , onde  $T \neq \emptyset$  é um conjunto de índices arbitrário, podendo ser discreto ou contínuo,  $\Omega$  é o espaço amostral,  $\mathcal{F}$  é a  $\sigma$ -álgebra de eventos aleatórios e  $\mathbb{P}$  é a medida de probabilidade definida em  $\mathcal{F}$ .

Note que, se fixarmos  $t \in T$ ,  $X_t(\cdot)$  é uma função sobre o conjunto  $\Omega$ . Por outro lado, para cada  $\omega \in \Omega$  fixo,  $X_t(\omega)$  é uma função sobre o conjunto  $T$ . Para cada  $\omega \in \Omega$  fixo, o conjunto de pontos  $\{X_t(\omega); t \in T\}$  é denominado uma *trajetória* ou uma *realização* do processo estocástico  $\{X_t\}_{t \in T}$ .

Neste trabalho consideramos os processos onde  $T = \mathbb{Z}$  ou  $T = \{1, \dots, n\}$ , que denotamos, respectivamente, por  $\{X_t\}_{t \in \mathbb{Z}}$  e  $\{X_t\}_{t=1}^n$ . Se  $T = \{1, \dots, n\}$  então o conjunto  $\{X_t\}_{t=1}^n$  é denominado uma *série temporal*. Portanto, uma *série temporal*  $\{X_t\}_{t=1}^n$  é um conjunto de observações de um fenômeno aleatório, medidas em diferentes tempos. Os índices diários da Bolsa de Valores, os registros de temperatura de uma cidade, as medidas dos níveis de água em uma bacia ou rio e a taxa de desemprego mensal são exemplos de séries temporais.

Na Figura 1.1 apresentamos um exemplo de uma série temporal. Esta série temporal é composta por  $n = 1.729$  observações e representa os índices diários da Bolsa de Valores

de São Paulo no período de janeiro de 1995 a dezembro de 2001.

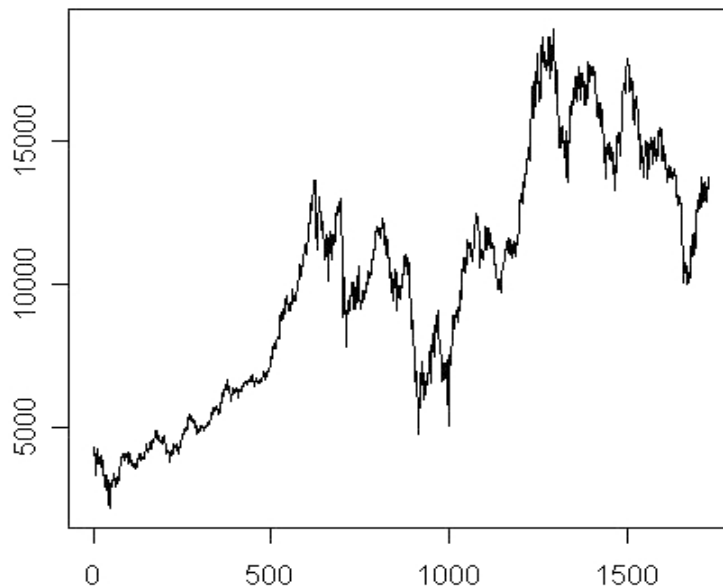


Figura 1.1: Série Temporal dos Índices Diários da Bolsa de Valores de São Paulo no Período de Janeiro de 1995 a Dezembro de 2001.

A seguir definimos a *função de distribuição  $n$ -dimensional* (ou *função de distribuição conjunta*) de um processo estocástico. O processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  estará especificado se conhecermos suas distribuições finito-dimensionais dadas na equação (1.1), para todo  $n \geq 1$ .

**Definição 1.2. (Função de Distribuição Finito-dimensional).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico, conforme a Definição 1.1. As *funções de distribuição finito-dimensionais* de  $\{X_t\}_{t \in \mathbb{Z}}$  são as funções  $\{F_{\mathbf{X}}(\cdot); \mathbf{t} \in \mathcal{T}\}$  definidas por

$$\begin{aligned} F_{\mathbf{X}}(\mathbf{x}) &= F_{X_{t_1}, \dots, X_{t_m}}(\mathbf{x}) \\ &= \mathbb{P}(X_{t_1} \leq x_1, \dots, X_{t_m} \leq x_m), \end{aligned} \quad (1.1)$$

onde  $\mathbf{t} = (t_1, \dots, t_m)' \in \mathcal{T}$ ,  $\mathbf{x} = (x_1, \dots, x_m)' \in \mathbb{R}^m$  e  $\mathcal{T}$  é o conjunto de todos os vetores  $\{\mathbf{t} = (t_1, \dots, t_m)' \in \mathbb{Z}^m : t_1 < \dots < t_m, m \in \mathbb{N} - \{0\}\}$ .

Na prática, as distribuições finito-dimensionais são muito difíceis de serem obtidas. O que se faz então é tentar caracterizar o processo por momentos até determinada ordem. Apresentamos a seguir a definição de *momento de ordem  $r$*  das variáveis aleatórias de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ .

**Definição 1.3. (Momento de Ordem  $r$ ).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Para todo  $r \in \mathbb{N} - \{0\}$ , o *momento de ordem  $r$*  de qualquer variável aleatória  $X_t$  é definido por

$$\mathbb{E}(X_t^r) = \int_{\mathbb{R}} x^r dF_X(x),$$

para todo  $t \in \mathbb{Z}$ , onde  $dF_X(x) = f_X(x)dx$  se a função  $F_X(\cdot)$  for diferenciável em  $x$ . A função  $f_X(\cdot)$  é denominada *função densidade de probabilidade* da variável aleatória  $X_t$ , para todo  $t \in \mathbb{Z}$ .

Neste trabalho analisamos, em geral, os momentos de ordem menor ou igual a 4. O momento de ordem 1, de uma variável aleatória  $X_t$  qualquer, é denominado *esperança matemática* e nos dá uma idéia da evolução média do processo no decorrer do tempo. Os momentos de ordem 1 e 2 são utilizados no cálculo da *variância* da variável aleatória  $X_t$ , para todo  $t \in \mathbb{Z}$ . A variância nos dá uma idéia da dispersão do processo em relação à sua esperança matemática. Enquanto que, os momentos de ordem 3 e 4 são fundamentais à análise das caudas da função densidade de probabilidade.

Quando utilizamos um determinado modelo para descrever algum processo físico, um dos problemas que enfrentamos é como lidar com a dependência entre duas ou mais variáveis aleatórias e como caracterizar e modelar tal dependência. Geralmente impomos restrições ao tipo de dependência com que estamos tratando, já que uma estrutura de dependência arbitrária é matematicamente intratável. Em geral, assumimos que os processos estocásticos são estacionários. As definições a seguir referem-se à estacionariedade de processos estocásticos.

**Definição 1.4. (Estacionariedade Estrita).** Um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é *estritamente estacionário* se as distribuições finito-dimensionais, dadas na expressão (1.1), forem invariantes sob translações de tempo, isto é,

$$\mathbb{P}(X_{t_1+h} \leq x_1, \dots, X_{t_k+h} \leq x_k) = \mathbb{P}(X_{t_1} \leq x_1, \dots, X_{t_k} \leq x_k),$$

para todo inteiro positivo  $k$  e para quaisquer  $t_1, \dots, t_k, h \in \mathbb{Z}$ .

**Definição 1.5. (Função de Autocovariância).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico tal que  $\text{Var}(X_t) < \infty$ , para todo  $t \in \mathbb{Z}$ . A *função de autocovariância* de  $\{X_t\}_{t \in \mathbb{Z}}$ , denotada por  $\gamma_X(\cdot, \cdot)$ , é definida por

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = \mathbb{E}\left[(X_r - \mathbb{E}(X_r))(X_s - \mathbb{E}(X_s))\right],$$

para todo  $r, s \in \mathbb{Z}$ .

**Definição 1.6. (Função de Autocorrelação).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico tal que  $\text{Var}(X_t) < \infty$ , para todo  $t \in \mathbb{Z}$ . A *função de autocorrelação* de  $\{X_t\}_{t \in \mathbb{Z}}$ , denotada por  $\rho_X(\cdot, \cdot)$ , é definida por

$$\rho_X(r, s) = \text{Corr}(X_r, X_s) = \frac{\gamma_X(r, s)}{\sqrt{\text{Var}(X_r)}\sqrt{\text{Var}(X_s)}}, \quad (1.2)$$

para todo  $r, s \in \mathbb{Z}$ .

As funções de autocorrelação e autocovariância fornecem o grau de interdependência entre as variáveis aleatórias. Essas funções são importantes pois são uma das ferramentas utilizadas na identificação do modelo mais adequado para uma série temporal.

**Definição 1.7. (Estacionariedade Fraca).** Um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é *fracamente estacionário* se, e somente se,

- (i)  $\mathbb{E}(|X_t|^2) < \infty$ , para todo  $t \in \mathbb{Z}$ ;
- (ii)  $\mathbb{E}(X_t) = \mu$ , para todo  $t \in \mathbb{Z}$ ;
- (iii)  $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ , para todo  $r, s$  e  $t \in \mathbb{Z}$ .

**Observação 1.1.** Na literatura, *estacionariedade de segunda ordem* e *estacionariedade em covariância* aparecem como sinônimos de *estacionariedade fraca*. Para facilitar a notação, omitimos a palavra fracamente da Definição 1.7 e diremos que o processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é *estacionário*.

É fácil ver que, se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo estocástico estritamente estacionário, as variáveis aleatórias  $X_t$  têm a mesma distribuição, para todo  $t \in \mathbb{Z}$  e ainda, se  $\mathbb{E}(|X_t|^2) < \infty$ , o processo é estacionário. Estacionariedade, entretanto, nem sempre implica estacionariedade estrita (para um exemplo veja Brockwell e Davis, 1991, página 13). Um importante caso em que estacionariedade implica estacionariedade estrita é o do *processo Gaussiano*, cuja definição apresentamos a seguir.

**Definição 1.8. (Processo Gaussiano).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Dizemos que  $\{X_t\}_{t \in \mathbb{Z}}$  é um *processo Gaussiano* se, e somente se, todas as funções de distribuição do processo estocástico são distribuições normais multivariadas.

Note que, se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo Gaussiano estacionário então, para todo  $m \in \mathbb{N}$  e para todos  $h, t_1, \dots, t_m \in \mathbb{Z}$ , os vetores aleatórios  $(X_{t_1}, \dots, X_{t_m})'$  e  $(X_{t_1+h}, \dots, X_{t_m+h})'$  tem a mesma média e a mesma matriz de covariância e, portanto, a mesma distribuição conjunta. Segue que  $\{X_t\}_{t \in \mathbb{Z}}$  é estritamente estacionário.

**Observação 1.2.**

1. Se o processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é estacionário, então

$$\begin{aligned}\gamma_X(r, s) &= \gamma_X(r - s, s - s) \\ &= \gamma_X(r - s, 0), \quad \text{para todo } r, s \in \mathbb{Z}.\end{aligned}$$

Assim podemos, equivalentemente, definir a função de autocovariância para um processo estocástico estacionário por

$$\gamma_X(h) \equiv \gamma_X(h, 0) = \text{Cov}(X_{t+h}, X_t), \quad \text{para todo } t, h \in \mathbb{Z}. \quad (1.3)$$

Além disso,  $\gamma_X(0) = \text{Var}(X_t) \equiv \sigma_X^2$ .

2. As funções de autocovariância e autocorrelação de um processo estocástico estacionário  $\{X_t\}_{t \in \mathbb{Z}}$  apresentam as seguintes propriedades, cujas provas são imediatas

- (i)  $\gamma_X(0) \geq 0$ ;
- (ii)  $|\gamma_X(h)| \leq \gamma_X(0)$ , para todo  $h \in \mathbb{Z}$ ;
- (iii)  $\gamma_X(-h) = \gamma_X(h)$ , para todo  $h \in \mathbb{Z}$ ;
- (iv)  $\gamma_X(\cdot)$  é uma função não-negativa definida, isto é, para todo inteiro  $n > 0$  e para todos os vetores  $\mathbf{a} = (a_1, \dots, a_m)' \in \mathbb{R}^m$  e  $\mathbf{t} = (t_1, \dots, t_m)' \in \mathbb{Z}^m$  tem-se

$$\sum_{i=1}^m \sum_{j=1}^m a_i \gamma_X(t_i - t_j) a_j \geq 0.$$

- (v)  $\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}$ , para todo  $h \in \mathbb{Z}$ ;

- (vi)  $\rho_X(-h) = \rho_X(h)$ , para todo  $h \in \mathbb{Z}$ ;



(vii)  $|\rho_X(h)| \leq \rho_X(0) = 1$ , para todo  $h \in \mathbb{Z}$ .

Assim como a função de autocorrelação, a *função de autocorrelação parcial*, definida a seguir, fornece informações sobre a estrutura de dependência do processo. Ela é importante quando deseja-se investigar a correlação entre as variáveis  $X_t$  e  $X_{t+h}$  após a remoção das dependências lineares das variáveis aleatórias intermediárias  $X_{t+1}, \dots, X_{t+h-1}$ .

**Definição 1.9. (Função de Autocorrelação Parcial).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário, com função de autocovariância  $\gamma_X(\cdot)$  tal que  $\gamma_X(h) \rightarrow 0$ , quando  $h \rightarrow \infty$ . A *função de autocorrelação parcial*, denotada por  $\phi_X(h, j)$ , para  $j = 1, 2, \dots, h$ , e  $h \in \mathbb{N} - \{0\}$ , é o coeficiente  $\phi_X(h, j)$  da equação

$$\mathcal{P}_{\overline{sp}\{X_1, X_2, \dots, X_h\}}(X_{h+1}) = \sum_{j=1}^h \phi_X(h, j) X_{h+1-j},$$

onde  $\mathcal{P}_{\overline{sp}\{X_1, X_2, \dots, X_h\}}(X_{h+1})$  denota a projeção ortogonal de  $X_{h+1}$  no subespaço fechado  $\overline{sp}\{X_1, X_2, \dots, X_h\}$  gerado pelas observações anteriores.

A partir da equação

$$\langle X_{h+1} - \mathcal{P}_{\overline{sp}\{X_1, X_2, \dots, X_h\}}(X_{h+1}), X_j \rangle = 0, \quad \text{para } j = 1, \dots, h,$$

onde  $\langle \cdot, \cdot \rangle$  denota o produto interno sobre o espaço de Hilbert  $L^2(\Omega, \mathcal{F}, \mathbb{P})$ , dado por  $\langle X, Y \rangle = \mathbb{E}(XY)$ , obtemos o sistema de equações

$$\begin{bmatrix} 1 & \rho_X(1) & \rho_X(2) & \cdots & \rho_X(h-1) \\ \rho_X(h-1) & 1 & \rho_X(1) & \cdots & \rho_X(h-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_X(h-1) & \rho_X(h-2) & \rho_X(h-3) & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_X(h, 1) \\ \phi_X(h, 2) \\ \vdots \\ \phi_X(h, h) \end{bmatrix} = \begin{bmatrix} \rho_X(1) \\ \rho_X(2) \\ \vdots \\ \rho_X(h) \end{bmatrix} \quad (1.4)$$

onde  $\rho_X(\cdot)$  é a função de autocorrelação do processo  $\{X_t\}_{t \in \mathbb{Z}}$ . Os coeficientes  $\phi_X(h, j)$  são unicamente determinados pelo sistema (1.4). Para um estudo mais completo, veja Brockwell e Davis (1991).

A seguir, definimos os estimadores, pelo método dos momentos, para o momento de ordem  $r$ , para  $r \in \mathbb{N} - \{0\}$  e para as funções de autocovariância, autocorrelação e autocorrelação parcial de um processo estocástico estacionário  $\{X_t\}_{t \in \mathbb{Z}}$ , observando que

- (i)  $\mathbb{E}(X_t^r) = \mu_r$ , para todo  $t \in \mathbb{Z}$  e todo  $r \in \mathbb{N} - \{0\}$ . Isto é, para cada  $r \in \mathbb{N} - \{0\}$ , o momento de ordem  $r$  é uma constante  $\mu_r$  que não depende de  $t \in \mathbb{Z}$ ;
- (ii)  $\text{Var}(X_t) = \sigma_X^2$ , para todo  $t \in \mathbb{Z}$ ;
- (iii)  $\text{Cov}(X_{t+h}, X_t) = \gamma_X(h)$ , para todo  $t, h \in \mathbb{Z}$ .

Para obter estes estimadores consideramos uma série temporal  $\{X_t\}_{t=1}^n$ , com  $n$  observações, obtida a partir do processo estocástico estacionário  $\{X_t\}_{t \in \mathbb{Z}}$ .

**Definição 1.10. (Momento Amostral de Ordem  $r$ ).** O estimador para o *momento amostral de ordem  $r$*  é definido como

$$\bar{X}^r = \frac{1}{n} \sum_{i=1}^n X_i^r.$$

Pela Definição 1.10 um estimador da média  $\mu_X$  do processo  $\{X_t\}_{t \in \mathbb{Z}}$  é a média amostral dada por

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (1.5)$$

**Definição 1.11. (Função de Autocovariância Amostral).** O estimador para a função de autocovariância  $\gamma_X(\cdot)$  é a *função de autocovariância amostral* definida por

$$\hat{\gamma}_X(h) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-h} (X_{t+h} - \bar{X})(X_t - \bar{X}), & \text{se } 0 \leq h < n \\ \hat{\gamma}_X(-h), & \text{se } -n < h \leq 0, \end{cases} \quad (1.6)$$

onde  $\bar{X}$  é a *média amostral* dada na expressão (1.5).

Pela expressão (1.6) o estimador para a variância  $\sigma_X^2$  do processo  $\{X_t\}_{t \in \mathbb{Z}}$ , dada no ítem 1. da Observação 1.2, é definido por

$$\hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2. \quad (1.7)$$

**Definição 1.12. (Função de Autocorrelação Amostral).** O estimador para a função de autocorrelação é a *função de autocorrelação amostral* definida por

$$\hat{\rho}_X(h) = \frac{\hat{\gamma}_X(h)}{\hat{\gamma}_X(0)}, \quad \text{para todo } -n < h < n, \quad (1.8)$$

onde  $\hat{\gamma}_X(\cdot)$  é dada na equação (1.6).

**Definição 1.13. (Função de Autocorrelação Parcial Amostral).** A *função de autocorrelação parcial amostral*,  $\hat{\phi}_X(h, j)$ , é obtida substituindo-se  $\rho_X(\cdot)$  por  $\hat{\rho}_X(\cdot)$  na expressão (1.4).

Para facilitar a resolução do sistema (1.4) utilizamos o algoritmo de Durbin-Levinson, que é um método recursivo, sugerido por Box *et al.* (1994), dado por

$$\hat{\phi}_X(h+1, h+1) = \frac{\hat{\rho}_X(h+1) - \sum_{j=1}^h \hat{\phi}_X(h, j) \hat{\rho}_X(h+1-j)}{1 - \sum_{j=1}^h \hat{\phi}_X(h, j) \hat{\rho}_X(j)},$$

$$\hat{\phi}_X(h+1, j) = \hat{\phi}_X(h, j) - \hat{\phi}_X(h+1, h+1) \hat{\phi}_X(h, h+1-j),$$

para todo  $j = 1, \dots, h$  e todo  $h \in \{1, \dots, n-1\}$ , onde  $\hat{\phi}_X(1, 1) = \hat{\rho}_X(1)$ .

O algoritmo de Durbin-Levinson também pode ser utilizado para obter o valor teórico da função de autocorrelação parcial,  $\phi_X(\cdot, \cdot)$ , a partir da função de autocorrelação  $\rho_X(\cdot)$ .

Quando queremos informações sobre as caudas da função de distribuição das variáveis aleatórias em questão devemos também considerar as medidas de *assimetria* e *curtose*, que definimos a seguir.

**Definição 1.14. (Medidas de Assimetria e de Curtose).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário tal que  $\mathbb{E}(X_t) = \mu_X$  e  $\text{Var}(X_t) = \sigma_X^2$ . As *medidas de assimetria e curtose* da variável aleatória  $X_t$ , para todo  $t \in \mathbb{Z}$ , são dadas por

$$A_X = \frac{\mathbb{E}[(X_t - \mu_X)^3]}{\sigma_X^3} \quad \text{e} \quad K_X = \frac{\mathbb{E}[(X_t - \mu_X)^4]}{\sigma_X^4},$$

respectivamente, onde  $\sigma_X$  é o desvio padrão do processo  $\{X_t\}_{t \in \mathbb{Z}}$ .

Note que se  $\mathbb{E}(X_t) = \mu_X = 0$ , então as medidas de assimetria e de curtose da variável aleatória  $X_t$  podem ser representadas, respectivamente, por

$$A_X = \frac{\mathbb{E}(X_t^3)}{[\mathbb{E}(X_t^2)]^{3/2}} \quad \text{e} \quad K_X = \frac{\mathbb{E}(X_t^4)}{[\mathbb{E}(X_t^2)]^2}, \quad \text{para todo } t \in \mathbb{Z}.$$

Definimos a seguir os estimadores, obtidos pelo método dos momentos, para as medidas de assimetria e curtose de um processo estocástico estacionário  $\{X_t\}_{t \in \mathbb{Z}}$ .

**Definição 1.15. (Medidas de Assimetria e de Curtose Amostrais).** Seja  $\{X_t\}_{t=1}^n$  uma série temporal obtida de um processo estocástico estacionário  $\{X_t\}_{t \in \mathbb{Z}}$ . As *medidas de assimetria e curtose amostrais* são definidas como

$$\hat{A}_X = \frac{1}{n \hat{\sigma}_X^3} \sum_{j=1}^n (X_j - \bar{X})^3 \quad \text{e} \quad \hat{K}_X = \frac{1}{n \hat{\sigma}_X^4} \sum_{j=1}^n (X_j - \bar{X})^4, \quad (1.9)$$

onde  $\bar{X}$  e  $\hat{\sigma}_X^2$  são, respectivamente, os estimadores de  $\mu_X$  e  $\sigma_X^2$  dados pelas equações (1.5) e (1.7).

As seguintes definições referem-se à análise espectral de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ .

**Definição 1.16. (Função Densidade Espectral).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário com função de autocovariância  $\gamma_X(\cdot)$  absolutamente convergente, isto é,  $\sum_{h \in \mathbb{Z}} |\gamma_X(h)| < \infty$ . A função densidade espectral de  $\{X_t\}_{t \in \mathbb{Z}}$ , denotada por  $f_X(\cdot)$ , é dada por

$$\begin{aligned} f_X(\lambda) &= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_X(h) e^{-i\lambda h} \\ &= \frac{1}{2\pi} \left[ \gamma_X(0) + 2 \sum_{h=1}^{\infty} \gamma_X(h) \cos(\lambda h) \right], \quad \text{para } \lambda \in [-\pi, \pi], \end{aligned}$$

onde  $\gamma_X(\cdot)$  é definida pela expressão (1.3).

O próximo teorema é de grande importância pois caracteriza a função de autocovariância de um processo estacionário  $\{X_t\}_{t \in \mathbb{Z}}$ , que pode ser escrita na forma

$$\gamma_X(h) = \int_{(-\pi, \pi]} e^{ih\lambda} dF_X(\lambda),$$

onde  $F_X(\cdot)$  é uma função de distribuição limitada, com massa concentrada em  $(-\pi, \pi]$ .

**Teorema 1.1. (Herglotz).** Uma função  $\gamma_X : \mathbb{C} \rightarrow \mathbb{Z}$  é não-negativa definida se, e somente se,

$$\gamma_X(h) = \int_{(-\pi, \pi]} e^{ih\lambda} dF_X(\lambda), \quad \text{para todo } h \in \mathbb{Z},$$

onde  $F_X(\cdot)$  é uma função limitada em  $[-\pi, \pi]$ , contínua à direita, não-decrescente, tal que  $F_X(-\pi) = 0$ . A função  $F_X(\cdot)$  é denominada função distribuição espectral de  $\{X_t\}_{t \in \mathbb{Z}}$ . Se

$$F_X(\lambda) = \int_{-\pi}^{\lambda} f_X(\omega) d\omega, \quad -\pi \leq \lambda \leq \pi,$$

então  $f_X(\cdot)$  é denominada função densidade espectral de  $\{X_t\}_{t \in \mathbb{Z}}$ .

**Prova:** Veja Brockwell e Davis (1991), páginas 117 e 118. □

Pelo Teorema de Herglotz e pela Definição 1.16, é possível concluir que a função densidade espectral de um processo estocástico estacionário é a transformada de Fourier da função de autocovariância, quando esta for absolutamente convergente. Enquanto que a função de autocovariância é a transformada de Fourier inversa da função de densidade espectral. Ou seja, a função de densidade espectral e a função de autocovariância formam um par pela transformada de Fourier, com uma sendo unicamente determinada pela outra.

O seguinte teorema apresenta algumas propriedades da função densidade espectral de um processo estocástico estacionário.

**Teorema 1.2.** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário com  $\sum_{h \in \mathbb{Z}} |\gamma_X(h)| < \infty$ . A função densidade espectral  $f_X(\cdot)$  de  $\{X_t\}_{t \in \mathbb{Z}}$  possui as seguintes propriedades:

- (i)  $f_X(\lambda)$  é uma função real contínua;
- (ii)  $f_X(-\lambda) = f_X(\lambda)$ , para todo  $\lambda \in (-\pi, \pi]$ ;
- (iii)  $f_X(\lambda) \geq 0$ , para todo  $\lambda \in (-\pi, \pi]$ ;
- (iv)  $\int_{-\pi}^{\pi} f_X(\lambda) d\lambda < \infty$ .

**Prova:** Para uma prova e para um estudo mais detalhado veja Brockwell e Davis (1991) e Wei (1990). □

Na próxima definição apresentamos um estimador, pelo método dos momentos, para a função densidade espectral  $f_X(\cdot)$ , de um processo estocástico estacionário  $\{X_t\}_{t \in \mathbb{Z}}$ .

**Definição 1.17. (Função Periodograma).** Seja  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir de um processo estocástico estacionário  $\{X_t\}_{t \in \mathbb{Z}}$ . A função periodograma,  $I_n(\cdot)$ , da série temporal  $\{X_t\}_{t=1}^n$  é definida por

$$I_n(\lambda) = \frac{1}{2\pi} \left[ \hat{\gamma}_X(0) + 2 \sum_{h=1}^{n-1} \hat{\gamma}_X(h) \cos(\lambda h) \right], \quad \text{para } \lambda \in [-\pi, \pi]. \quad (1.10)$$

A função periodograma é um estimador não-viciado mas inconsistente para a função densidade espectral  $f_X(\cdot)$ . Definimos, a seguir, um estimador consistente para a função densidade espectral (ver Priestley, 1981).

**Definição 1.18. (Função Periodograma Suavizado).** A função periodograma suavizado para uma série temporal  $\{X_t\}_{t=1}^n$ , denotada por  $f_s(\cdot)$ , é definida por

$$f_s(\lambda) = \frac{1}{2\pi} \sum_{h=-(n-1)}^{n-1} \kappa(h) \hat{\gamma}_X(h) \cos(\lambda h), \quad \text{para todo } \lambda \in (-\pi, \pi],$$

onde  $\kappa(\cdot)$  é uma função de ponderação conhecida como janela espectral.

Apresentamos, a seguir, as definições de processos *ruído branco* e *martingale difference*. Tais processos aparecem com frequência nas definições e demonstrações nos próximos capítulos.

**Definição 1.19. (Processo Ruído Branco).** Seja  $\{Z_t\}_{t \in \mathbb{Z}}$  um processo estocástico. O processo  $\{Z_t\}_{t \in \mathbb{Z}}$  é um *processo ruído branco* com média 0 e variância  $\sigma_Z^2$  se, e somente se,

$$\mathbb{E}(Z_t) = 0 \quad \text{e} \quad \gamma_Z(h) = \begin{cases} \sigma_Z^2, & h = 0 \\ 0, & h \neq 0. \end{cases}$$

**Definição 1.20. (Processo Martingale Difference).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico tal que  $E(|X_t|) < \infty$ , para todo  $t \in \mathbb{Z}$ . Seja  $\{\mathcal{F}_t\}_{t \in \mathbb{Z}}$  uma *filtração*, isto é,  $\{\mathcal{F}_t\}_{t \in \mathbb{Z}}$  é uma seqüência de  $\sigma$ -álgebras tais que, para todo  $t \in \mathbb{Z}$ ,  $\mathcal{F}_{t-1} \subset \mathcal{F}_t$ . Dizemos que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *martingale difference* em relação a filtração  $\{\mathcal{F}_t\}_{t \in \mathbb{Z}}$  se

$$E(X_t | \mathcal{F}_{t-1}) = 0, \quad \text{para todo } t \in \mathbb{Z}.$$

**Proposição 1.1.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo martingale difference em relação à filtração  $\{\mathcal{F}_t\}_{t \in \mathbb{Z}}$ . Então,*

- (i)  $E(X_t) = 0$ , para todo  $t \in \mathbb{Z}$ ;
- (ii) se  $E(X_t^2) < \infty$ , para todo  $t \in \mathbb{Z}$ ,

$$\gamma_X(t, s) = 0, \quad \text{para todo } t, s \in \mathbb{Z} \text{ tais que } t \neq s;$$

- (iii) se  $\text{Var}(X_t) = \mathbb{E}(X_t^2) = \sigma_X^2$  não depende de  $t \in \mathbb{Z}$ ,  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo ruído branco, conforme a Definição 1.19.

**Prova:** Para provar (i), note que, se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *martingale difference* em relação à  $\{\mathcal{F}_t\}_{t \in \mathbb{Z}}$ . Então,

$$\mathbb{E}(X_t) = \mathbb{E}(\mathbb{E}(X_t | \mathcal{F}_{t-1})) = 0.$$

Para (ii), suponha que  $E(X_t^2) < \infty$ , para todo  $t \in \mathbb{Z}$ . Pelo item (i),  $\mathbb{E}(X_t) = 0$ , para todo  $t \in \mathbb{Z}$ . Então,

$$\begin{aligned} \gamma_X(t, s) &= \text{Cov}(X_t, X_s) \\ &= \mathbb{E}(X_t X_s) \\ &= \begin{cases} \mathbb{E}(\mathbb{E}(X_t X_s | \mathcal{F}_{s-1})) = \mathbb{E}(X_t \mathbb{E}(X_s | \mathcal{F}_{s-1})) = 0, & \text{se } t < s \\ \mathbb{E}(\mathbb{E}(X_t X_s | \mathcal{F}_{t-1})) = \mathbb{E}(X_s \mathbb{E}(X_t | \mathcal{F}_{t-1})) = 0, & \text{se } t > s. \end{cases} \end{aligned}$$

Portanto, para todo  $t, s \in \mathbb{Z}$  tais que  $t \neq s$ , segue que

$$\gamma_X(t, s) = 0.$$

Suponha agora que  $\text{Var}(X_t) = \mathbb{E}(X_t^2) = \sigma_X^2$  não depende de  $t \in \mathbb{Z}$ . Por (i),  $\mathbb{E}(X_t) = 0$ , que não depende de  $t \in \mathbb{Z}$ . Por (ii) e pela hipótese, segue que

$$\gamma_X(h, 0) = \gamma_X(h) = \begin{cases} \sigma_X^2, & \text{se } h = 0 \\ 0, & \text{se } h \neq 0. \end{cases}$$

Ou seja,  $\{X_t\}_{t \in \mathbb{Z}}$  é um ruído branco, de acordo com a Definição 1.19. □

## 1.2 Funções e Séries de Potências

Nesta seção apresentamos os principais resultados sobre funções e séries de potências utilizados nas demonstrações nos capítulos que seguem. Para mais detalhes e provas das proposições veja Ahlfors (1979) e Lima (2006).

**Proposição 1.2.** *Sejam  $f(\cdot)$  e  $g(\cdot)$  funções de séries de potência dadas por  $f(x) = \sum_{k=0}^{\infty} f_k x^k$  e  $g(x) = \sum_{k=0}^{\infty} g_k x^k$ , onde  $f_k$  e  $g_k \in \mathbb{R}$ , para todo  $k \geq 0$ . Então,*

(i)  $f(x) = g(x)$  se, e somente se,  $f_k = g_k$ , para todo  $k \geq 0$ .

(ii)  $f(x)g(x) = \sum_{k=0}^{\infty} c_k x^k$ , onde

$$c_n = \sum_{j=0}^k f_j g_{k-j} = \sum_{j=0}^k f_{k-j} g_j = \sum_{j=0}^k g_j f_{k-j}.$$

Para fixar a notação utilizada ao longo do trabalho apresentamos as seguintes definições.

**Definição 1.21.** Sejam  $\{a_n\}_{n \in \mathbb{N}}$  uma seqüência de números reais e  $k \in \mathbb{R}$ . Dizemos que  $a_n \sim n^{-k}$  se, existem constantes reais  $c_1, c_2 > 0$  tais que, para todo  $n \in \mathbb{N}$ ,

$$c_1 < \left| \frac{a_n}{n^{-k}} \right| < c_2.$$

**Definição 1.22.** Sejam  $f : \mathbb{R} \rightarrow \mathbb{R}$  e  $g : \mathbb{R} \rightarrow \mathbb{R}$  duas funções mensuráveis quaisquer. Dizemos que  $f$  é aproximadamente igual a  $g$ , denotando por  $f(x) \sim g(x)$ , quando  $x$  tende ao infinito se

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1.$$

**Definição 1.23.** Sejam  $f : \mathbb{R} \rightarrow \mathbb{R}$  e  $g : \mathbb{R} \rightarrow \mathbb{R}$  duas funções mensuráveis quaisquer. Dizemos que  $f$  é aproximadamente igual a  $g$ , denotando por  $f(x) \sim g(x)$ , quando  $x$  tende a zero se

$$\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = 1.$$

**Definição 1.24.** Sejam  $f : \mathbb{R} \rightarrow \mathbb{R}$  e  $g : \mathbb{R} \rightarrow \mathbb{R}$  duas funções mensuráveis quaisquer. A função  $f$  é da ordem de  $g$  quando  $x$  tende ao infinito, denotada por  $f(x) = o(g(x))$  se, e somente se,

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0.$$

### 1.3 Processos com Longa Dependência

A presença de longa dependência em séries temporais foi primeiramente observada por Hurst (1951) quando analisava a série temporal de níveis do rio Nilo, e mais tarde, por Mandelbrot e Wallis (1968) e McLeod e Hipel (1978). Os primeiros modelos para séries temporais com longa dependência na média foram introduzidos por Granger e Joyeux (1980) e Hosking (1981). Recentemente, os economistas observaram que há evidências de que processos com longa dependência descrevem, de modo satisfatório, dados econômicos e financeiros, tais como taxas de juros e inflação. Discutiremos com mais detalhes tais processos no Capítulo 2. A propriedade de longa dependência para um processo estocástico pode ser definida de diferentes maneiras, podendo ser no domínio do tempo ou no domínio da frequência (equivalentemente). Uma abordagem mais formal para a definição de longa dependência é dada a seguir.

**Definição 1.25. (Longa Dependência).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário com função de autocorrelação  $\rho_X(\cdot)$  e função densidade espectral  $f_X(\cdot)$ . No domínio do tempo, dizemos que  $\{X_t\}_{t \in \mathbb{Z}}$  apresenta *longa dependência* se, para algum número real  $a \in (0, 1)$ ,

$$\rho_X(h) \sim h^{-a}, \quad \text{quando } h \rightarrow \infty.$$

Equivalentemente, no domínio da frequência, dizemos que  $\{X_t\}_{t \in \mathbb{Z}}$  apresenta *longa dependência* se, para algum número real  $b \in (0, 1)$ ,

$$f_X(\lambda) \sim \lambda^{-b}, \quad \text{quando } \lambda \rightarrow 0.$$

Note que o decaimento hiperbólico lento da função de autocorrelação de um processo estocástico estacionário com longa dependência implica que

$$\sum_{h=-m}^m |\rho_X(h)| \rightarrow \infty, \quad \text{quando } m \rightarrow \infty.$$

Apresentamos a seguir algumas definições e resultados referentes a processos com longa dependência na média. Um estudo mais completo sobre tais processos pode ser encontrado em Hosking (1981), Brockwell e Davis (1991), Beran (1994) e Olbermann (2002). Apresentamos, no Capítulo 2, um estudo mais detalhado de processos com longa dependência na volatilidade, contendo resultados equivalentes aos apresentados nesta seção. Sem perda de generalidade consideramos, nas próximas definições, que  $\mathbb{E}(X_t) = 0$ .

**Definição 1.26. (Processos ARMA).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *autoregressivo média móvel*, denotado por  $\text{ARMA}(p, q)$  se, e somente se,

$$\phi(\mathcal{B})X_t = \theta(\mathcal{B})\epsilon_t, \quad \text{para todo } t \in \mathbb{Z}, \quad (1.11)$$

onde  $\{\epsilon_t\}_{t \in \mathbb{Z}}$  é um processo ruído branco, de acordo com a Definição 1.19,  $\phi(\cdot)$  e  $\theta(\cdot)$  são polinômios de grau  $p$  e  $q$ , respectivamente, dados por

$$\phi(z) \equiv - \sum_{i=0}^p \phi_i z^i, \quad \theta(z) \equiv - \sum_{j=0}^q \theta_j z^j, \quad \text{onde } \phi_0 \equiv -1 \equiv \theta_0 \quad (1.12)$$

e  $\mathcal{B}$  é o operador de defasagem (*backward shift operator*) dado por

$$\mathcal{B}^j(X_t) = \mathcal{B}^{j-1}\mathcal{B}(X_t) = \mathcal{B}^{j-1}(X_{t-1}) = X_{t-j}. \quad (1.13)$$

Chamaremos os polinômios  $\phi(\cdot)$  e  $\theta(\cdot)$  de polinômios autoregressivo e média móvel, respectivamente.

Apresentamos, a seguir, a definição de *causalidade*. Causalidade não é uma propriedade apenas do processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  ou do processo estocástico  $\{\epsilon_t\}_{t \in \mathbb{Z}}$ , mas sim de ambos. Utilizamos essa propriedade, e os resultados que seguem, em diversas situações nos próximos capítulos.

**Definição 1.27. (Causalidade).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo ARMA( $p, q$ ), definido pela expressão (1.11). Então, o processo  $\{X_t\}_{t \in \mathbb{Z}}$  é *causal*, ou mais especificamente, é uma função causal das variáveis aleatórias  $\{\epsilon_t\}_{t \in \mathbb{Z}}$  se existe uma seqüência de constantes  $\{\zeta_j\}_{j \in \mathbb{Z}}$  tais que  $\sum_{j=0}^{\infty} |\zeta_j| < \infty$  e

$$X_t = \sum_{j=0}^{\infty} \zeta_j \epsilon_{t-j}, \quad \text{para todo } t \in \mathbb{Z}. \quad (1.14)$$

O teorema que segue apresenta as condições necessárias e suficientes para que um processo ARMA seja causal.

**Teorema 1.3.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo ARMA( $p, q$ ), definido pela expressão (1.11), tal que os polinômios  $\phi(\cdot)$  e  $\theta(\cdot)$  não possuem raízes em comum. Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é causal se, e somente se,  $\phi(z) \neq 0$  para todo  $z \in \mathbb{C}$  tal que  $|z| \leq 1$ . Além disso, os coeficientes  $\{\zeta_j\}_{j \in \mathbb{Z}}$ , dados na expressão (1.14), são unicamente determinados pela expressão*

$$\zeta(z) = \sum_{j=0}^{\infty} \zeta_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1.$$

**Prova:** Veja Brockwell e Davis, 1991, página 85. □

A seguir apresentamos a definição de inversibilidade dos processos ARMA e o teorema de sua caracterização.

**Definição 1.28. (Inversibilidade).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo ARMA( $p, q$ ), definido pela expressão (1.11). Então, o processo  $\{X_t\}_{t \in \mathbb{Z}}$  é *inversível* se existe uma seqüência de constantes  $\{\pi_j\}_{j \in \mathbb{Z}}$  tais que  $\sum_{j=0}^{\infty} |\pi_j| < \infty$  e

$$\epsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad \text{para todo } t \in \mathbb{Z}. \quad (1.15)$$

**Teorema 1.4.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo ARMA( $p, q$ ), definido pela expressão (1.11), tal que os polinômios  $\phi(\cdot)$  e  $\theta(\cdot)$  não possuem raízes em comum. Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é inversível se, e somente se,  $\theta(z) \neq 0$  para todo  $z \in \mathbb{C}$  tal que  $|z| \leq 1$ . Além disso, os coeficientes  $\{\pi_j\}_{j \in \mathbb{Z}}$ , dados na expressão (1.15), são unicamente determinados pela expressão*

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad |z| \leq 1.$$

**Prova:** Veja Brockwell e Davis, 1991, página 87. □



Até agora tratamos dos processos ARMA, que não apresentam longa dependência. A seguir apresentamos a definição dos processos ARFIMA. Estes processos apresentam longa dependência na média.

**Definição 1.29. (Processos ARFIMA).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *autoregressivo fracionariamente integrado de média móvel*, denotado por ARFIMA( $p, d, q$ ), com  $d \in (-0, 5; 0, 5)$ , se e somente se,

$$\phi(\mathcal{B})(1 - \mathcal{B})^d X_t = \theta(\mathcal{B})\epsilon_t, \quad \text{para todo } t \in \mathbb{Z}, \quad (1.16)$$

onde  $\{\epsilon_t\}_{t \in \mathbb{Z}}$  é um processo ruído branco, conforme a Definição 1.19 e  $\phi(\cdot)$  e  $\theta(\cdot)$  são polinômios de grau  $p$  e  $q$ , respectivamente, dados pela expressão (1.12).

O parâmetro  $d$  da expressão (1.16) é o grau de diferenciação fracionária e o operador  $(1 - \mathcal{B})^d$  é definido em termos de sua expansão através de séries de Maclaurin (veja Lopes e Mendes, 2006) por

$$(1 - \mathcal{B})^d = \sum_{k=0}^{\infty} \frac{\Gamma(k - d)}{\Gamma(k + 1)\Gamma(-d)} \mathcal{B}^k. \quad (1.17)$$

Note que, como  $\Gamma(x + 1) = x\Gamma(x)$ , para todo  $x \in \mathbb{R}$ , a expressão (1.17) pode ser reescrita como

$$\begin{aligned} (1 - \mathcal{B})^d &= 1 - d \times \sum_{k=1}^{\infty} \frac{\Gamma(k - d)}{\Gamma(k + 1)\Gamma(1 - d)} \mathcal{B}^k \\ &= 1 - \sum_{k=1}^{\infty} \delta_{d,k} \mathcal{B}^k \\ &= \sum_{k=0}^{\infty} (-\delta_{d,k}) \mathcal{B}^k = \delta_d(\mathcal{B}). \end{aligned} \quad (1.18)$$

Os coeficientes  $\delta_{d,k}$  que aparecem na expressão (1.18) são dados por  $\delta_{d,0} \equiv -1$  e

$$\delta_{d,k} = \frac{d\Gamma(k - d)}{\Gamma(k + 1)\Gamma(1 - d)}, \quad \text{para todo } k \geq 1, \quad (1.19)$$

e são tais que

$$\begin{aligned} \delta_{d,k} &= \frac{d\Gamma(k - d)}{\Gamma(k + 1)\Gamma(1 - d)} = \frac{d(k - d - 1)\Gamma(k - 1 - d)}{k\Gamma(k)\Gamma(1 - d)} \\ &= \left( \frac{k - 1 - d}{k} \right) \delta_{d,k-1}. \end{aligned}$$

O próximo teorema apresenta as características de um processo ARFIMA.

**Teorema 1.5.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo ARFIMA( $p, d, q$ ), definido pela expressão (1.16). Suponha que os polinômios  $\phi(\cdot)$  e  $\theta(\cdot)$  não possuem raízes em comum. Segue que,*

- (i) *se  $\phi(z) \neq 0$  para  $|z| = 1$  então, existe uma única solução estacionária para a equação (1.16) dada por*

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j}, \quad \text{para todo } t \in \mathbb{Z},$$

onde

$$\psi(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j = (1-z)^{-d} \frac{\theta(z)}{\phi(z)}.$$

- (ii) o processo  $\{X_t\}_{t \in \mathbb{Z}}$  é causal se, e somente se,  $\phi(z) \neq 0$  para  $|z| \leq 1$ .
- (iii) o processo  $\{X_t\}_{t \in \mathbb{Z}}$  é inversível se, e somente se,  $\theta(z) \neq 0$  para  $|z| \leq 1$ .
- (iv) se o processo  $\{X_t\}_{t \in \mathbb{Z}}$  é causal e inversível então, sua função de autocorrelação  $\rho_X(\cdot)$  e sua densidade espectral  $f_X(\cdot)$  satisfazem,

$$\rho_X(h) \sim Ch^{2d-1}, \quad \text{quando } h \rightarrow \infty,$$

com  $C \neq 0$  e,

$$f_X(\lambda) = \frac{\sigma_\epsilon^2 |\theta(e^{-i\lambda})|^2}{2\pi |\phi(e^{-i\lambda})|^2} |1 - e^{-i\lambda}|^{-2d} \sim \frac{\sigma_\epsilon^2}{2\pi} \left[ \frac{\theta(1)}{\phi(1)} \right]^2 \lambda^{-2d}, \quad \text{quando } \lambda \rightarrow 0.$$

**Prova:** Veja Brockwell e Davis (1991) e Palma (2007). □

## 1.4 Processos Ergódicos

Nesta seção, apresentamos alguns conceitos básicos da Teoria Ergódica, como a definição de processo ergódico e os principais teoremas relacionados à ergodicidade. Consideramos apenas processos estritamente estacionários, no sentido da Definição 1.4. Os resultados apresentados nesta seção são de grande importância na caracterização dos processos estudados no Capítulo 2.

**Definição 1.30. (Transformação que Preserva Probabilidade).** Seja  $(\Omega, \mathcal{F}, \mathbb{P})$  um espaço de probabilidade e  $\varphi : \Omega \rightarrow \Omega$  uma transformação. A transformação  $\varphi(\cdot)$  preserva a probabilidade  $\mathbb{P}$  se, e somente se,  $\mathbb{P}(\varphi^{-1}(A)) = \mathbb{P}(A)$ , para todo  $A \in \mathcal{F}$ . Dizemos também que  $\mathbb{P}$  é *invariante* para  $\varphi$ .

**Definição 1.31. (Operador *Shift*):** Seja  $\Omega = S^{\mathbb{Z}}$  a  $\sigma$ -álgebra gerada pelos cilindros. A transformação  $\varphi : \Omega \rightarrow \Omega$  definida por  $\varphi(\omega) = (\cdots \omega_{-1} \omega_0 \mid \omega_1 \omega_2 \cdots)$ , para todo  $\omega = (\cdots \omega_{-1} \mid \omega_0 \omega_1 \omega_2 \cdots) \in \Omega$ , é denominada *operador shift*.

Dado um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , o operador *shift* induz uma transformação  $\varphi(\cdot)$  em  $\Omega$ . O processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  será estritamente estacionário se, e somente se, o *shift* preserva a probabilidade  $\mathbb{P}$  (ver Billingsley, 1995).

**Definição 1.32. (Ergodicidade).** Seja  $\varphi : \Omega \rightarrow \Omega$  uma transformação que preserva a probabilidade  $\mathbb{P}$ . Então,  $\mathbb{P}$  é *ergódica* se, e somente se, para todo  $A \in \mathcal{F}$  tal que  $\varphi^{-1}(A) = A$  tem-se  $\mathbb{P}(A) = 0$  ou  $\mathbb{P}(A) = 1$ .

No próximo lema mostramos que todo processo estocástico cujas variáveis aleatórias são independentes e identicamente distribuídas é estritamente estacionário e ergódico.

**Lema 1.1.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico cujas variáveis aleatórias são independentes e identicamente distribuídas. Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é estritamente estacionário e ergódico.*

**Prova:** Veja Stout (1974), página 182. □

O seguinte teorema é de grande importância pois estabelece que transformações mensuráveis preservam estacionariedade estrita e ergodicidade.

**Teorema 1.6.** *Seja  $\{Z_t\}_{t \in \mathbb{Z}}$  um processo estocástico estritamente estacionário e ergódico e  $\phi : \mathbb{R}^\infty \rightarrow \mathbb{R}$  uma transformação mensurável. Seja  $\{X_t\}_{t \in \mathbb{Z}}$  o processo estocástico definido por  $X_t \equiv \phi(Z_t, Z_{t-1}, \dots)$ . Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é estritamente estacionário e ergódico.*

**Prova:** Veja Stout (1974), página 182. □

O próximo teorema é de grande importância pois será utilizado para caracterizar os processos apresentados no Capítulo 2.

**Teorema 1.7.** *Seja  $\{Z_t\}_{t \in \mathbb{Z}}$  uma seqüência de variáveis aleatórias independentes e identicamente distribuídas, com média zero e variância finita. Se  $\sum_{j=0}^{\infty} a_j^2 < \infty$ , então o processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  definido por*

$$X_t = \sum_{j=0}^{\infty} a_j Z_{t-j}, \quad \text{para todo } t \in \mathbb{Z},$$

*é estritamente estacionário e ergódico.*

**Prova:** Aplicação imediata do Lema 1.1 e do Teorema 1.6. □

# Capítulo 2

## Processos Não-Lineares

Neste capítulo apresentamos alguns modelos não-lineares utilizados na análise de séries temporais que apresentam heteroscedasticidade, ou seja, a variância condicional varia com o tempo. Primeiramente, discutimos os modelos ARCH e GARCH. Apresentamos as principais características e os principais resultados relacionados a tais processos. Em seguida, tratamos dos modelos EGARCH. Finalmente, com o objetivo de modelar séries temporais que, além de apresentarem variância condicional variando no tempo, também apresentam característica de longa dependência na volatilidade, definimos os processos FIEGARCH. Para estes processos, apresentados na Seção 2.5, provamos diversas propriedades importantes, tais como estacionariedade, ergodicidade, causalidade e inversibilidade. Além disso, provamos que se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo FIEGARCH( $p, d, q$ ) então, sob certas condições, o processo estocástico  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, 0$ ), com inovações não-Gaussianas.

Séries temporais financeiras apresentam uma importante característica conhecida como *volatilidade*, que pode ser definida de diversas formas, mas não é diretamente observável. Em termos financeiros, a volatilidade pode ser vista como a medida estatística da possibilidade do valor de um ativo cair ou subir muitas vezes, de forma significativa, em um determinado período de tempo. Neste capítulo, dado um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , a volatilidade é a variância condicional desse processo.

Neste capítulo, dado um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , denotamos por  $\mathcal{F}_t$  a  $\sigma$ -álgebra de eventos gerada pelas informações disponíveis até o instante  $t$ , ou seja,  $\mathcal{F}_t \equiv \sigma(\{X_s; s \leq t\})$ .

### 2.1 Processos ARCH

Nesta seção apresentamos os principais resultados relacionados à classe de modelos autorregressivos com heteroscedasticidade condicional, também denominada *família* ARCH (*Autoregressive Conditional Heteroskedasticity*), introduzida por Engle (1982).

**Definição 2.1. (Processos ARCH).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Dizemos que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *auto-regressivo com heteroscedasticidade condicional de ordem*  $p$ , denotado por ARCH( $p$ ), se

$$X_t = \sigma_t Z_t, \tag{2.1}$$

$$\sigma_t^2 = \alpha_0 + \alpha(\mathcal{B})X_t^2, \quad \text{para todo } t \in \mathbb{Z}, \tag{2.2}$$

onde  $\{Z_t\}_{t \in \mathbb{Z}}$  são variáveis aleatórias independentes e identicamente distribuídas com média zero e variância um,  $\mathcal{B}$  é o operador de defasagem definido pela expressão (1.13) e

$$\alpha(\mathcal{B}) \equiv \sum_{i=1}^p \alpha_i \mathcal{B}^i, \quad (2.3)$$

$\alpha_0 > 0$  e  $\alpha_i \geq 0$ , para  $i = 1, 2, \dots, p$ , são constantes reais.

É fácil ver que  $Z_t$  e  $\mathcal{F}_{t-1}$  são independentes, para todo  $t \in \mathbb{Z}$ . Segundo McNeil *et al.* (2005), este resultado segue do fato que um processo ARCH deve ser causal, ou seja, as equações dadas pelas expressões (2.1) e (2.2) devem ter uma solução da forma  $X_t = f(Z_t, Z_{t-1}, \dots)$  para alguma função  $f(\cdot)$  mensurável, de forma que  $Z_t$  é independente dos valores anteriores ao tempo  $t$  do processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ . Este fato é um contraste com os modelos ARMA para os quais as equações podem ter soluções não-causais. Com base nisso, é imediato que, se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo ARCH( $p$ ), então

$$\begin{aligned} \text{Var}(X_t | \mathcal{F}_{t-1}) &= \mathbb{E}(X_t^2 | \mathcal{F}_{t-1}) - (\mathbb{E}(X_t | \mathcal{F}_{t-1}))^2 = \mathbb{E}(\sigma_t^2 Z_t^2 | \mathcal{F}_{t-1}) - (\mathbb{E}(\sigma_t Z_t | \mathcal{F}_{t-1}))^2 \\ &= \sigma_t^2 \mathbb{E}(Z_t^2 | \mathcal{F}_{t-1}) - (\sigma_t \mathbb{E}(Z_t | \mathcal{F}_{t-1}))^2 = \sigma_t^2 \mathbb{E}(Z_t^2) - \sigma_t^2 (\mathbb{E}(Z_t))^2 \\ &= \sigma_t^2, \quad \text{para todo } t \in \mathbb{Z}, \end{aligned}$$

ou seja, a volatilidade é uma função que depende do quadrado dos valores passados do processo.

No teorema a seguir apresentamos uma condição necessária e suficiente para a estacionariedade de um processo ARCH( $p$ ).

**Teorema 2.1.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo ARCH( $p$ ), dado na Definição 2.1. Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é estacionário de segunda ordem se, e somente se,*

$$1 - \alpha(z) \neq 0, \quad \text{para todo } z \in \mathbb{C} \text{ tal que } |z| \leq 1,$$

onde  $\alpha(\cdot)$  é dado pela expressão (2.3).

**Prova:** Veja Engle (1982). □

Apresentamos a seguir um exemplo de um processo ARCH( $p$ ), para  $p = 2$ , simulado através da função *simulate.garch* do programa S-Plus.

**Exemplo 2.1. (Processo ARCH( $\mathbf{p}$ ), com  $\mathbf{p} = 2$ ).** Considere o processo ARCH( $p$ ), dado pelas expressões (2.1) e (2.2), com  $p = 2$ . Para gerar a série temporal  $\{X_t\}_{t=1}^n$ , com  $n = 250$  observações, utilizamos  $\alpha_0 = 0,25$ ,  $\alpha_1 = 0,47$  e  $\alpha_2 = 0,1$ .

A Figura 2.1 apresenta a série temporal simulada  $\{X_t\}_{t=1}^n$ , para  $n = 250$ , e seu desvio padrão condicional (raiz quadrada da volatilidade). A Figura 2.2 apresenta as funções de autocorrelação amostral das séries temporais  $\{X_t^2\}_{t=1}^{250}$  e  $\{\sigma_t^2\}_{t=1}^{250}$ . Note como essas séries temporais apresentam funções de autocorrelação amostral semelhantes. Essa é uma característica dos processos ARCH( $p$ ).

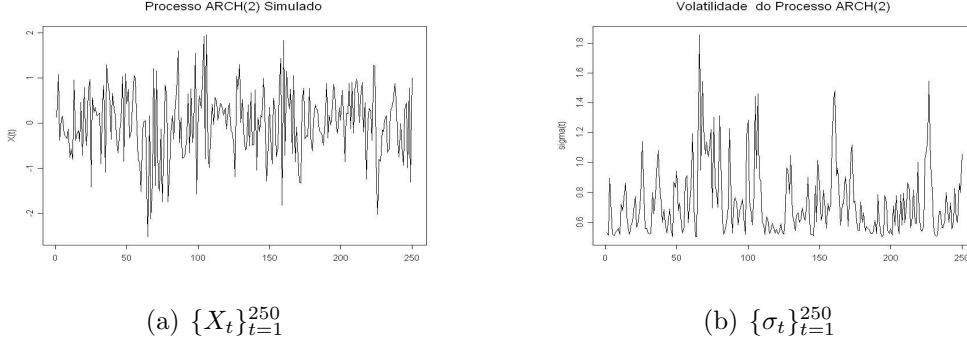


Figura 2.1: Séries Temporais Simuladas do Exemplo 2.1.

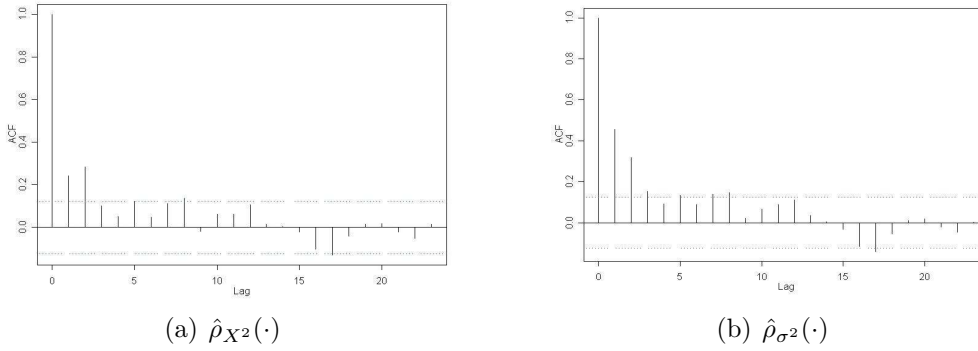


Figura 2.2: Funções de Autocorrelação Amostral das Séries Temporais  $\{X_t^2\}_{t=1}^{250}$  e  $\{\sigma_t^2\}_{t=1}^{250}$ .

Observe que a expressão (2.2) pode ser reescrita como

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \cdots + \alpha_p X_{t-p}^2 + X_t^2 - X_t^2,$$

o que é equivalente a

$$(1 - \alpha(\mathcal{B}))X_t^2 = \alpha_0 + v_t, \quad (2.4)$$

onde  $v_t = X_t^2 - \sigma_t^2$  e  $\alpha(\cdot)$  dado pela expressão (2.3). Desta forma,  $\{X_t^2\}_{t \in \mathbb{Z}}$  pode ser escrito como um processo  $\text{AR}(p)$ . Note que, em geral,  $\{v_t\}_{t \in \mathbb{Z}}$  não será um processo independente e identicamente distribuído, mas sim, de acordo com a Definição 1.20, um processo *martingale difference*. De fato,

$$\begin{aligned} \mathbb{E}(v_t) &= \mathbb{E}(X_t^2 - \sigma_t^2) \\ &= \mathbb{E}(\mathbb{E}(\sigma_t^2 Z_t^2 - \sigma_t^2 | \mathcal{F}_{t-1})) \\ &= \mathbb{E}(\sigma_t^2 \mathbb{E}(Z_t^2 - 1 | \mathcal{F}_{t-1})) = 0 \end{aligned}$$

e

$$\mathbb{E}(v_t | \mathcal{F}_{t-1}) = \mathbb{E}(X_t^2 - \sigma_t^2 | \mathcal{F}_{t-1}) = \mathbb{E}(X_t^2 | \mathcal{F}_{t-1}) - \mathbb{E}(\sigma_t^2 | \mathcal{F}_{t-1}) = \sigma_t^2 - \sigma_t^2 = 0.$$

No modelo clássico  $\text{ARCH}(p)$  de Engle (1982), a variância condicional é uma função linear dos quadrados das inovações passadas, implicando dependência Markoviana de ordem  $p$ . O processo  $\text{GARCH}(p, q)$ , que apresentamos na próxima seção, introduzido por Bollerslev (1986), fornece uma estrutura mais flexível pois expressa a volatilidade do processo como uma função tanto dos valores passados quanto das volatilidades passadas do processo.

## 2.2 Processos GARCH

Nesta seção tratamos dos modelos GARCH (*Generalized ARCH*), introduzidos por Bollerslev (1986). Da mesma forma que os modelos ARMA podem ser mais parcimoniosos, no sentido de apresentar menor número de parâmetros que um modelo AR ou MA, os modelos GARCH podem descrever a volatilidade com menos parâmetros do que um modelo ARCH.

**Definição 2.2. (Processos GARCH).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Dizemos que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *autoregressivo com heteroscedasticidade condicional generalizado*, ou GARCH( $p, q$ ) se,

$$X_t = \sigma_t Z_t, \quad (2.5)$$

$$\sigma_t^2 = \alpha_0 + \alpha(\mathcal{B})X_t^2 + \beta(\mathcal{B})\sigma_t^2, \quad \text{para todo } t \in \mathbb{Z}, \quad (2.6)$$

onde  $\{Z_t\}_{t \in \mathbb{Z}}$  são variáveis aleatórias independentes e identicamente distribuídas com média zero e variância um e  $Z_t$  é independente de  $\mathcal{F}_{t-1}$ , para todo  $t \in \mathbb{Z}$ ,  $\alpha_0 > 0$  é uma constante real,  $\alpha(\cdot)$  é o polinômio definido pela expressão (2.3) e o polinômio  $\beta(\cdot)$  é definido por

$$\beta(\mathcal{B}) \equiv \sum_{j=1}^q \beta_j \mathcal{B}^j, \quad (2.7)$$

com  $\beta_j \geq 0$ , para todo  $j = 1, 2, \dots, q$ , e  $\mathcal{B}$  é o operador de defasagem definido pela expressão (1.13).

Os processos GARCH são uma generalização dos processos ARCH no sentido que a volatilidade  $\sigma_t^2$  é dependente tanto dos quadrados dos valores anteriores do processo como das volatilidades passadas. Note que, se  $q = 0$ , então o modelo GARCH( $p, 0$ ) reduz-se ao modelo ARCH( $p$ ) definido na Seção 2.1.

Na Definição 2.2, exigimos que os coeficientes  $\alpha_i$ , para todo  $i = 1, 2, \dots, p$  e  $\beta_j$ , para todo  $j = 1, 2, \dots, q$ , sejam positivos. Esta é uma condição suficiente, mas não necessária para que  $\sigma_t^2 > 0$ . O leitor interessado pode encontrar um estudo mais detalhado e condições gerais para que  $\sigma_t$ , para todo  $t \in \mathbb{Z}$ , esteja bem definida em Nelson e Cao (1992).

Note que, se todas as raízes do polinômio  $(1 - \beta(\mathcal{B}))$  estão fora do círculo unitário, então o processo GARCH( $p, q$ ), definido pelas expressões (2.5) e (2.6), possui uma representação ARCH( $\infty$ ). De fato, observe que

$$\sigma_t^2 = \alpha_0 + \alpha(\mathcal{B})X_t^2 + \beta(\mathcal{B})\sigma_t^2 \iff (1 - \beta(\mathcal{B}))\sigma_t^2 = \alpha_0 + \alpha(\mathcal{B})X_t^2.$$

Ou seja,

$$\begin{aligned} \sigma_t^2 &= (1 - \beta(\mathcal{B}))^{-1} \alpha_0 + (1 - \beta(\mathcal{B}))^{-1} \alpha(\mathcal{B})X_t^2 \\ &= (1 - \beta(1))^{-1} \alpha_0 + \sum_{j=0}^{\infty} \Phi_j X_{t-j}^2, \end{aligned}$$

onde  $\Phi(\mathcal{B}) \equiv (1 - \beta(\mathcal{B}))^{-1} \alpha(\mathcal{B}) = \sum_{j=0}^{\infty} \Phi_j \mathcal{B}^j$ .

Mostramos agora que, assim como os processos ARCH possuem uma representação AR para o processo  $\{X_t^2\}_{t \in \mathbb{Z}}$ , os processos GARCH possuem uma representação ARMA para o processo  $\{X_t^2\}_{t \in \mathbb{Z}}$ . Note que, somando-se  $X_t^2 - \beta(\mathcal{B})X_t^2$  em ambos os lados da igualdade, a expressão (2.6) pode ser reescrita como

$$\begin{aligned} X_t^2 - \beta(\mathcal{B})X_t^2 + \sigma_t^2 &= \alpha_0 + \alpha(\mathcal{B})X_t^2 + \beta(\mathcal{B})\sigma_t^2 + X_t^2 - \beta(\mathcal{B})X_t^2 \iff \\ (1 - \alpha(\mathcal{B}) - \beta(\mathcal{B}))X_t^2 &= \alpha_0 - (1 - \beta(\mathcal{B}))\sigma_t^2 + (1 - \beta(\mathcal{B}))X_t^2 \iff \\ (1 - \alpha(\mathcal{B}) - \beta(\mathcal{B}))X_t^2 &= \alpha_0 + (1 - \beta(\mathcal{B}))v_t, \end{aligned} \quad (2.8)$$

onde  $v_t = X_t^2 - \sigma_t^2$ , para todo  $t \in \mathbb{Z}$ . Como mostramos na Seção 2.1, o processo  $\{v_t\}_{t \in \mathbb{Z}}$  é um processo *martingale difference*. Logo, se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo GARCH( $p, q$ ), o processo  $\{X_t^2\}_{t \in \mathbb{Z}}$  pode ser representado como um processo ARMA( $\max\{p, q\}, q$ ).

Apresentamos a seguir um exemplo de um processo GARCH( $p, q$ ), para  $p = 1$  e  $q = 1$ , simulado através da função *simulate.garch* do programa S-Plus.

**Exemplo 2.2. (Processo GARCH( $p, q$ ) com  $p = 1 = q$ ).** Considere o processo GARCH( $p, q$ ), dado pelas expressões (2.5) e (2.6), com  $p = 1 = q$ . A série temporal  $\{X_t\}_{t=1}^n$ , com  $n = 250$  observações, foi gerada utilizando-se  $\alpha_0 = 0,1$ ,  $\alpha_1 = 0,27$  e  $\beta_1 = 0,32$ . A Figura 2.3 apresenta a série temporal simulada  $\{X_t\}_{t=1}^n$ , para  $n = 250$ , e seu desvio padrão condicional (raiz quadrada da volatilidade). A Figura 2.4 apresenta as funções de autocorrelação amostral das séries temporais  $\{X_t^2\}_{t=1}^{250}$  e  $\{\sigma_t^2\}_{t=1}^{250}$ .

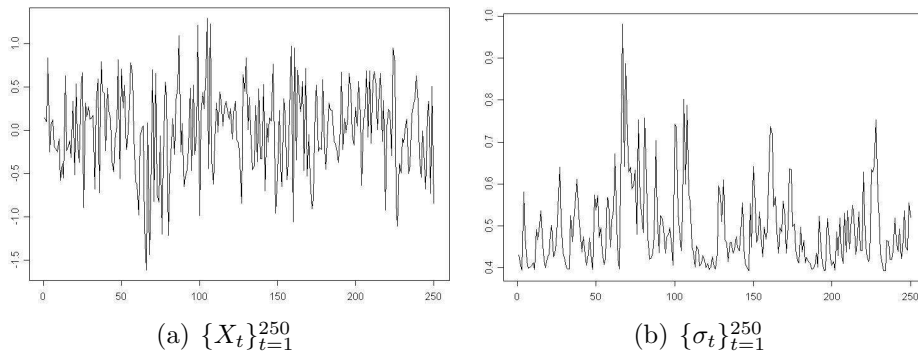


Figura 2.3: Séries Temporais do Exemplo 2.2.

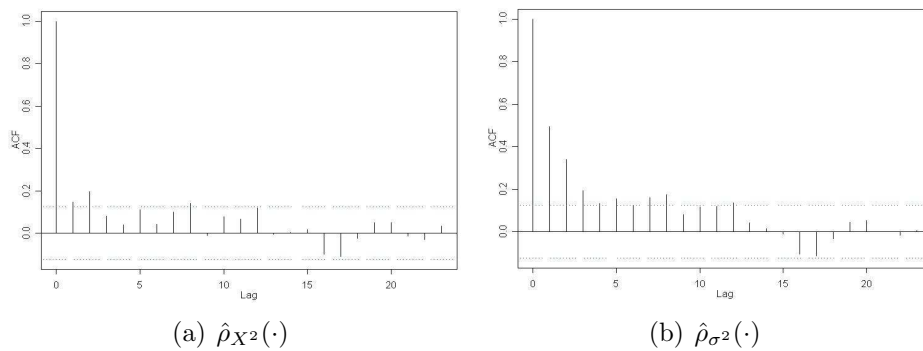


Figura 2.4: Funções de Autocorrelação Amostral das Séries Temporais  $\{X_t^2\}_{t=1}^{250}$  e  $\{\sigma_t^2\}_{t=1}^{250}$ .



No próximo teorema apresentamos uma condição necessária e suficiente para a estacionariedade do processo GARCH( $p, q$ ).

**Teorema 2.2.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo GARCH( $p, q$ ), dado na Definição 2.2. Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é estacionário se, e somente se,*

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1.$$

**Prova:** Veja Bollerslev (1986). □

O próximo lema mostra que todo processo GARCH é um processo *martingale difference*. Utilizamos este resultado na prova da Proposição 2.1.

**Lema 2.1.** *Sejam  $\{X_t\}_{t \in \mathbb{Z}}$  um processo GARCH( $p, q$ ) e  $\mathcal{F}_t = \sigma(\{X_s\}; s \leq t)$ . Então,  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo martingale difference em relação à filtração  $\{\mathcal{F}_t\}_{t \in \mathbb{Z}}$ .*

**Prova:** Note que,

$$\mathbb{E}(X_t) = \mathbb{E}(\mathbb{E}(X_t | \mathcal{F}_{t-1})) = \mathbb{E}(\mathbb{E}(\sigma_t Z_t | \mathcal{F}_{t-1})) = \mathbb{E}(\sigma_t \mathbb{E}(Z_t | \mathcal{F}_{t-1})) = 0.$$

Além disso,

$$\mathbb{E}(X_t | \mathcal{F}_{t-1}) = \mathbb{E}(\sigma_t Z_t | \mathcal{F}_{t-1}) = \sigma_t \mathbb{E}(Z_t | \mathcal{F}_{t-1}) = 0.$$

Segue, da Definição 1.20, que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *martingale difference*. □

Na próxima proposição apresentamos os principais resultados que caracterizam os processos GARCH( $p, q$ ).

**Proposição 2.1.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo GARCH( $p, q$ ), dado na Definição 2.2. Então,*

(i) *se  $\{X_t\}_{t \in \mathbb{Z}}$  é estacionário, sua variância é dada por*

$$\text{Var}(X_t) = \frac{\alpha_0}{1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j}; \quad (2.9)$$

(ii) *se  $\mathbb{E}(X_t^2) < \infty$ , para todo  $t \in \mathbb{Z}$ , a função de autocovariância  $\gamma_X(\cdot, \cdot)$  é tal que  $\gamma_X(t, s) = 0$ , para todo  $t, s \in \mathbb{Z}$  tais que  $t \neq s$ ;*

(iii) *se o polinômio  $(1 - \alpha(\mathcal{B}) - \beta(\mathcal{B}))$ , dado na expressão (2.8), não possui raízes no círculo unitário, se  $\mathbb{E}(X_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$  e não depende de  $t$  e ainda, se  $\mathbb{E}(Z_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ , a curtose da variável aleatória  $X_t$ , para todo  $t \in \mathbb{Z}$ , é dada por*

$$K_X = \frac{\mathbb{E}(Z_t^4)}{\mathbb{E}(Z_t^4) - (\mathbb{E}(Z_t^4) - 1) \sum_{j=0}^{\infty} \Psi_j^2},$$

onde

$$\Psi(\mathcal{B}) \equiv \sum_{j=0}^{\infty} \Psi_j \mathcal{B}^j = \frac{1 - \beta(\mathcal{B})}{1 - \alpha(\mathcal{B}) - \beta(\mathcal{B})}, \text{ com } \sum_{j=1}^{\infty} \Psi_j^2 < \infty. \quad (2.10)$$

**Prova:** Para provar (i) note que o processo  $\{X_t\}_{t \in \mathbb{Z}}$  é estacionário. Logo,  $\mathbb{E}(X_t^2) = \mathbb{E}(X_{t+h}^2)$ , para todo  $h \in \mathbb{Z}$ . Na prova do Lema 2.1, mostramos que  $\mathbb{E}(X_t) = 0$ , para todo  $t \in \mathbb{Z}$ . Além disso,

$$\mathbb{E}(X_t^2) = \mathbb{E}(\sigma_t^2 Z_t^2) = \mathbb{E}(\mathbb{E}(\sigma_t^2 Z_t^2 | \mathcal{F}_{t-1})) = \mathbb{E}(\sigma_t^2 \mathbb{E}(Z_t^2 | \mathcal{F}_{t-1})) = \mathbb{E}(\sigma_t^2), \quad (2.11)$$

para todo  $t \in \mathbb{Z}$ . Por esses resultados e pelas expressões (2.6) e (2.11) segue que

$$\mathbb{E}(X_t^2) = \alpha_0 + \sum_{i=1}^p \alpha_i \mathbb{E}(X_t^2) + \sum_{j=1}^q \beta_j \mathbb{E}(X_t^2).$$

Assim

$$\mathbb{E}(X_t^2) \left( 1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j \right) = \alpha_0.$$

Como  $\{X_t\}_{t \in \mathbb{Z}}$  é estacionário, pelo Teorema 2.2,  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ . Portanto, segue que

$$\begin{aligned} \text{Var}(X_t) &= \mathbb{E}(X_t^2) \\ &= \frac{\alpha_0}{1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j}. \end{aligned}$$

Para provar (ii) note que, pelo Lema 2.1,  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *martingale difference* em relação à  $\{\mathcal{F}_t\}_{t \in \mathbb{Z}}$ . Logo, pela Proposição 1.1,  $\gamma_X(t, s) = 0$  para todo  $t, s \in \mathbb{Z}$  tais que  $t \neq s$ .

Para provar (iii) note que, se todas as raízes do polinômio  $(1 - \alpha(\mathcal{B}) - \beta(\mathcal{B}))$  estão fora do círculo unitário podemos reescrever a expressão (2.8) como

$$\begin{aligned} X_t^2 &= \frac{\alpha_0}{(1 - \alpha(1) - \beta(1))} + \frac{1 - \beta(\mathcal{B})}{(1 - \alpha(\mathcal{B}) - \beta(\mathcal{B}))} v_t \\ &= \frac{\alpha_0}{(1 - \alpha(1) - \beta(1))} + \sum_{j=0}^{\infty} \Psi_j v_{t-j}, \end{aligned} \quad (2.12)$$

onde  $\Psi(\cdot)$  é o polinômio dado pela expressão (2.10). Além disso,

$$\mathbb{E}(X_t^4) = \mathbb{E}(\sigma_t^4 Z_t^4) = \mathbb{E}(\mathbb{E}(\sigma_t^4 Z_t^4 | \mathcal{F}_{t-1})) = \mathbb{E}(\sigma_t^4 \mathbb{E}(Z_t^4 | \mathcal{F}_{t-1})) = \mathbb{E}(\sigma_t^4) \mathbb{E}(Z_t^4). \quad (2.13)$$

Se  $\mathbb{E}(X_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$  e não depende de  $t \in \mathbb{Z}$ , e  $\mathbb{E}(Z_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ , como as variáveis aleatórias  $Z_t$  são identicamente distribuídas, é imediato, pela expressão (2.13), que  $\mathbb{E}(\sigma_t^4) < \infty$  e não depende de  $t \in \mathbb{Z}$ . Portanto,

$$\begin{aligned} \sigma_v^2 &= \mathbb{E}(v_t^2) = \mathbb{E}((X_t^2 - \sigma_t^2)^2) = \mathbb{E}(X_t^4) - 2\mathbb{E}(X_t^2 \sigma_t^2) + \mathbb{E}(\sigma_t^4) \\ &= \mathbb{E}(X_t^4) - 2\mathbb{E}(\mathbb{E}(X_t^2 \sigma_t^2 | \mathcal{F}_{t-1})) + \mathbb{E}(\sigma_t^4) \\ &= \mathbb{E}(X_t^4) - 2\mathbb{E}(\sigma_t^2 \mathbb{E}(X_t^2 | \mathcal{F}_{t-1})) + \mathbb{E}(\sigma_t^4) = \mathbb{E}(X_t^4) - 2\mathbb{E}(\sigma_t^4) + \mathbb{E}(\sigma_t^4) \\ &= \mathbb{E}(X_t^4) - \mathbb{E}(\sigma_t^4) < \infty, \quad \text{para todo } t \in \mathbb{Z}. \end{aligned} \quad (2.14)$$

Pela expressão (2.14) e pelo item (iii) da Proposição 1.1,  $\{v_t\}_{t \in \mathbb{Z}}$  é um processo ruído branco. Portanto, pela expressão (2.12) segue que

$$\text{Var}(X_t^2) = \sigma_v^2 \sum_{j=0}^{\infty} \Psi_j^2, \quad (2.15)$$

onde  $\sigma_v^2 = \text{Var}(v_t)$ , para todo  $t \in \mathbb{Z}$ . Note ainda que, pelas expressões (2.13) e (2.14),

$$\begin{aligned}\sigma_v^2 &= \mathbb{E}(X_t^4) - \mathbb{E}(\sigma_t^4) = \mathbb{E}(\sigma_t^4)\mathbb{E}(Z_t^4) - \mathbb{E}(\sigma_t^4) \\ &= \mathbb{E}(\sigma_t^4)(\mathbb{E}(Z_t^4) - 1), \quad \text{para todo } t \in \mathbb{Z}.\end{aligned}$$

Portanto, substituindo este resultado na expressão (2.15) temos,

$$\text{Var}(X_t^2) = \mathbb{E}(\sigma_t^4)(\mathbb{E}(Z_t^4) - 1) \sum_{j=0}^{\infty} \Psi_j^2. \quad (2.16)$$

Por outro lado,

$$\begin{aligned}\text{Var}(X_t^2) &= \mathbb{E}(X_t^4) - (\mathbb{E}(X_t^2))^2 = \mathbb{E}(\sigma_t^4)\mathbb{E}(Z_t^4) - (\mathbb{E}(\sigma_t^2 Z_t^2))^2 \\ &= \mathbb{E}(\sigma_t^4)\mathbb{E}(Z_t^4) - [\mathbb{E}(\mathbb{E}(\sigma_t^2 Z_t^2 | \mathcal{F}_{t-1}))]^2 \\ &= \mathbb{E}(\sigma_t^4)\mathbb{E}(Z_t^4) - [\mathbb{E}(\sigma_t^2 \mathbb{E}(Z_t^2 | \mathcal{F}_{t-1}))]^2 \\ &= \mathbb{E}(\sigma_t^4)\mathbb{E}(Z_t^4) - (\mathbb{E}(\sigma_t^2))^2.\end{aligned} \quad (2.17)$$

Segue, pelas equações (2.16) e (2.17) que,

$$\mathbb{E}(\sigma_t^4)(\mathbb{E}(Z_t^4) - 1) \sum_{j=0}^{\infty} \Psi_j^2 = \mathbb{E}(\sigma_t^4)\mathbb{E}(Z_t^4) - (\mathbb{E}(\sigma_t^2))^2.$$

Assim, passando os termos que multiplicam  $\mathbb{E}(\sigma_t^4)$  para o mesmo lado da igualdade temos,

$$\mathbb{E}(\sigma_t^4) \left[ \mathbb{E}(Z_t^4) - (\mathbb{E}(Z_t^4) - 1) \sum_{j=0}^{\infty} \Psi_j^2 \right] = (\mathbb{E}(\sigma_t^2))^2.$$

Logo, se  $\mathbb{E}(Z_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ , então

$$\frac{\mathbb{E}(\sigma_t^4)}{(\mathbb{E}(\sigma_t^2))^2} = \frac{1}{\mathbb{E}(Z_t^4) - (\mathbb{E}(Z_t^4) - 1) \sum_{j=0}^{\infty} \Psi_j^2}. \quad (2.18)$$

Assim, pela Definição 1.14 temos que

$$\begin{aligned}K_X &= \frac{\mathbb{E}(X_t^4)}{(\mathbb{E}(X_t^2))^2} = \frac{\mathbb{E}(\sigma_t^4 Z_t^4)}{(\mathbb{E}(\sigma_t^2 Z_t^2))^2} = \frac{\mathbb{E}(\mathbb{E}(\sigma_t^4 Z_t^4 | \mathcal{F}_{t-1}))}{[\mathbb{E}(\mathbb{E}(\sigma_t^2 Z_t^2 | \mathcal{F}_{t-1}))]^2} = \frac{\mathbb{E}(\sigma_t^4 \mathbb{E}(Z_t^4 | \mathcal{F}_{t-1}))}{[\mathbb{E}(\sigma_t^2 \mathbb{E}(Z_t^2 | \mathcal{F}_{t-1}))]^2} \\ &= \frac{\mathbb{E}(Z_t^4)\mathbb{E}(\sigma_t^4)}{(\mathbb{E}(\sigma_t^2))^2}.\end{aligned} \quad (2.19)$$

Pelas expressões (2.18) e (2.19) segue que,

$$K_X = \frac{\mathbb{E}(Z_t^4)}{\mathbb{E}(Z_t^4) - (\mathbb{E}(Z_t^4) - 1) \sum_{j=0}^{\infty} \Psi_j^2},$$

onde  $\Psi(\cdot)$  é dado pela expressão (2.10). Isso completa a prova.  $\square$

No Exemplo 2.3 apresentamos o cálculo da variância, da curtose do processo  $\{X_t\}_{t \in \mathbb{Z}}$  e das funções de autocovariância e autocorrelação do processo  $\{X_t^2\}_{t \in \mathbb{Z}}$ , onde  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo ARCH( $p$ ), com  $p = 1$ . No Exemplo 2.4 apresentamos o cálculo dessas funções quando  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo GARCH( $p, q$ ), com  $p = 1 = q$ .

**Exemplo 2.3. (Propriedades do Processo ARCH( $p$ ), com  $p = 1$ ).** Considere o processo ARCH( $p$ ), dado pelas expressões (2.1) e (2.2), com  $p = 1$ . Apresentamos o cálculo da variância e da curtose do processo estacionário  $\{X_t\}_{t \in \mathbb{Z}}$  e das funções de autocovariância e autocorrelação do processo  $\{X_t^2\}_{t \in \mathbb{Z}}$ .

Note que, pela Proposição 2.1, para todo  $t \in \mathbb{Z}$ ,

$$\mathbb{E}(X_t) = 0.$$

Pelo Teorema 2.1, se  $|\alpha_1| < 1$  então,  $\{X_t\}_{t \in \mathbb{Z}}$  é estacionário. Pelo item (i) da Proposição 2.1

$$\text{Var}(X_t) = \frac{\alpha_0}{1 - \alpha_1}. \quad (2.20)$$

Suponha que  $\{X_t^2\}_{t \in \mathbb{Z}}$  é estacionário e  $\mathbb{E}(Z_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ . Como

$$\Psi(\mathcal{B}) = \frac{1}{1 - \alpha(\mathcal{B})} = \frac{1}{1 - \alpha_1 \mathcal{B}} = \sum_{j=0}^{\infty} \alpha_1^j \mathcal{B}^j,$$

temos

$$\sum_{j=0}^{\infty} \Psi_j^2 = \sum_{j=0}^{\infty} (\alpha_1^j)^2 = \frac{1}{1 - \alpha_1^2}. \quad (2.21)$$

Segue, do item (iii) da Proposição 2.1 e da expressão (2.21), que

$$\begin{aligned} K_X &= \frac{\mathbb{E}(Z_t^4)}{\mathbb{E}(Z_t^4) - (\mathbb{E}(Z_t^4) - 1) \sum_{j=0}^{\infty} \Psi_j^2} \\ &= \frac{\mathbb{E}(Z_t^4)(1 - \alpha_1^2)}{(1 - \alpha_1^2)\mathbb{E}(Z_t^4) - (\mathbb{E}(Z_t^4) - 1)} \\ &= \frac{\mathbb{E}(Z_t^4)(1 - \alpha_1^2)}{1 - \mathbb{E}(Z_t^4)\alpha_1^2}. \end{aligned} \quad (2.22)$$

Particularmente, se  $Z_t \sim \mathcal{N}(0, 1)$  então,  $\mathbb{E}(Z_t^4) = 3$ . Logo, pela expressão (2.22),

$$K_X = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2} > \frac{3(1 - \alpha_1^2)}{1 - \alpha_1^2} = 3.$$

Esse fato mostra que um processo ARCH( $p$ ), com  $p = 1$ , possui caudas mais pesadas do que aquelas de uma distribuição normal.

Além disso, como

$$K_X = \frac{\mathbb{E}(X_t^4)}{(\mathbb{E}(X_t^2))^2}, \quad (2.23)$$

pelas expressões (2.20) e (2.23), temos que

$$\begin{aligned} \mathbb{E}(X_t^4) &= \frac{3(1 - \alpha_1^2)}{(1 - 3\alpha_1^2)} \times \left( \frac{\alpha_0}{1 - \alpha_1} \right)^2 \\ &= \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}. \end{aligned}$$

Com isso conclui-se que os momentos de quarta ordem de  $\{X_t\}_{t \in \mathbb{Z}}$  são positivos e finitos se, e somente se,

$$0 < 1 - 3\alpha_1^2 \iff 0 \leq \alpha_1^2 < \frac{1}{3}.$$

Pela expressão (2.4), o processo  $\{X_t^2\}_{t \in \mathbb{Z}}$  é um processo AR( $p$ ), com  $p = 1$ . Lembramos que, na prova da Proposição 2.1, mostramos que, se  $\mathbb{E}(X_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ , e não depende de  $t$  e  $\mathbb{E}(Z_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ , então, o processo  $\{v_t\}_{t \in \mathbb{Z}}$ , definido na expressão (2.4), é um processo ruído branco. Portanto, se o processo  $\{X_t^2\}_{t \in \mathbb{Z}}$  é estacionário e  $\mathbb{E}(Z_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ , segue que

$$\gamma_{X^2}(h) = \alpha_1^h \gamma_{X^2}(0), \quad \text{para todo } h > 0 \quad \text{e}$$

$$\rho_{X^2}(h) = \alpha_1^h, \quad \text{para todo } h > 0.$$

Para maiores detalhes sobre o cálculo das funções de autocorrelação e autocovariância de processos AR( $p$ ) veja Brockwell e Davis (1991).

**Exemplo 2.4. (Propriedades do Processo GARCH( $p, q$ ) com  $p = 1 = q$ ).** Considere o processo GARCH( $p, q$ ) estacionário, dado pelas expressões (2.5) e (2.6) com  $p = 1 = q$ . Apresentamos o cálculo da variância e da curtose do processo  $\{X_t\}_{t \in \mathbb{Z}}$  e das funções de autocovariância e autocorrelação do processo  $\{X_t^2\}_{t \in \mathbb{Z}}$ .

Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo GARCH( $p, q$ ) com  $p = 1 = q$ . Lembramos que

$$\mathbb{E}(X_t) = 0, \quad \text{para todo } t \in \mathbb{Z}.$$

Pelo Teorema 2.2, se  $\alpha_1 + \beta_1 < 1$  então,  $\{X_t\}_{t \in \mathbb{Z}}$  é estacionário. Segue, pelo item (i) da Proposição 2.1, que

$$\text{Var}(X_t) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

Para calcular a medida de curtose  $K_X$ , para o processo GARCH( $p, q$ ), com  $p = 1 = q$ , precisamos obter os coeficientes do polinômio  $\Psi(\cdot)$ , definido pela expressão (2.10). Para isso, observe que

$$\Psi(\mathcal{B}) = \frac{1 - \beta_1 \mathcal{B}}{1 - (\alpha_1 + \beta_1) \mathcal{B}} = \frac{\alpha_1 \mathcal{B} + (1 - (\alpha_1 + \beta_1) \mathcal{B})}{1 - (\alpha_1 + \beta_1) \mathcal{B}} = \frac{\alpha_1 \mathcal{B}}{1 - (\alpha_1 + \beta_1) \mathcal{B}} + 1. \quad (2.24)$$

Além disso, como  $\alpha_1 + \beta_1 < 1$ , segue que

$$\frac{1}{1 - (\alpha_1 + \beta_1) \mathcal{B}} = \sum_{j=0}^{\infty} (\alpha_1 + \beta_1)^j \mathcal{B}^j. \quad (2.25)$$

Substituindo a expressão (2.25) na expressão (2.24) obtemos

$$\begin{aligned} \Psi(\mathcal{B}) &= \frac{1 - \beta_1 \mathcal{B}}{1 - (\alpha_1 + \beta_1) \mathcal{B}} = 1 + \alpha_1 \mathcal{B} + \alpha_1 (\alpha_1 + \beta_1) \mathcal{B}^2 + \dots + \alpha_1 (\alpha_1 + \beta_1)^{n-1} \mathcal{B}^n + \dots \\ &= 1 + \sum_{j=1}^{\infty} \alpha_1 (\alpha_1 + \beta_1)^{j-1} \mathcal{B}^j. \end{aligned} \quad (2.26)$$

Segue, da expressão (2.26), que  $\Psi_0 = 1$  e  $\Psi_j = \alpha_1 (\alpha_1 + \beta_1)^{j-1}$ , para todo  $j > 1$ . Logo,

$$\sum_{j=0}^{\infty} \Psi_j^2 = 1 + \sum_{j=1}^{\infty} \alpha_1^2 (\alpha_1 + \beta_1)^{2(j-1)} = 1 + \frac{\alpha_1^2}{1 - (\alpha_1 + \beta_1)^2} = \frac{1 - (\alpha_1 + \beta_1)^2 + \alpha_1^2}{1 - (\alpha_1 + \beta_1)^2}. \quad (2.27)$$

Segue, do item (iii) da Proposição 2.1 e da expressão (2.27), que

$$\begin{aligned} K_X &= \frac{\mathbb{E}(Z_t^4)}{\mathbb{E}(Z_t^4) - (\mathbb{E}(Z_t^4) - 1) \sum_{j=0}^{\infty} \Psi_j^2} \\ &= \frac{\mathbb{E}(Z_t^4)(1 - (\alpha_1 + \beta_1)^2)}{1 - (\alpha_1 + \beta_1)^2 - (\mathbb{E}(Z_t^4) - 1)\alpha_1^2}. \end{aligned}$$

Particularmente, se  $Z_t \sim \mathcal{N}(0, 1)$  então,  $\mathbb{E}(Z_t^4) = 3$  e

$$K_X = \frac{3(1 - (\alpha_1 + \beta_1)^2)}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3.$$

Esse fato mostra que um processo GARCH( $p, q$ ), com  $p = 1 = q$ , possui caudas mais pesadas do que aquelas de uma distribuição normal.

Para encontrar a função de autocorrelação do processo estocástico  $\{X_t^2\}_{t \in \mathbb{Z}}$ , lembramos que, pela expressão (2.8), este processo possui uma representação ARMA( $\max\{p, q\}, q$ ). Além disso, para que este processo seja estacionário, o polinômio autoregressivo desta representação não pode ter raízes no círculo unitário. Segue, da expressão (2.12) que

$$\text{Cov}(X_{t+h}^2, X_t^2) = \text{Cov}\left(a + \sum_{i=0}^{\infty} \Psi_i v_{t+h-i}, a + \sum_{j=0}^{\infty} \Psi_j v_{t-j}\right) \quad (2.28)$$

onde  $a = \frac{\alpha_0}{1 - \alpha(1) - \beta(1)}$  e os coeficientes  $\Psi_i$ , para todo  $i \geq 0$ , são definidos pela expressão (2.10).

Como  $\{v_t\}_{t \in \mathbb{Z}}$  é um processo *martingale difference*, segue pela Proposição 1.1, que  $\text{Cov}(v_t, v_s) = 0$ , para todo  $t, s \in \mathbb{Z}$  tais que  $t \neq s$ . Logo, pela expressão (2.28) temos

$$\begin{aligned} \gamma_{X^2}(h) &= \text{Cov}(X_{t+h}^2, X_t^2) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \Psi_i \Psi_j \text{Cov}(v_{t+h-i}, v_{t-j}) \\ &= \sum_{j=0}^{\infty} \Psi_j \Psi_{j+h} \text{Var}(v_t) \\ &= \sigma_v^2 \sum_{j=0}^{\infty} \Psi_j \Psi_{j+h}, \quad \text{para todo } h \geq 0 \end{aligned} \quad (2.29)$$

e

$$\rho_{X^2}(h) = \frac{\gamma_{X^2}(h)}{\gamma_{X^2}(0)} = \frac{\sum_{j=0}^{\infty} \Psi_j \Psi_{j+h}}{\sum_{j=0}^{\infty} \Psi_j^2}, \quad \text{para todo } h \geq 0. \quad (2.30)$$

Pelas expressões (2.26), (2.29) e (2.30) temos que as funções de autocovariância e autocorrelação do processo estocástico  $\{X_t^2\}_{t \in \mathbb{Z}}$ , onde  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo GARCH( $p, q$ ), com  $p = q = 1$ , são dadas, respectivamente, por

$$\begin{aligned}
\gamma_{X^2}(h) &= \sigma_v^2 \Psi_0 \Psi_h + \sigma_v^2 \sum_{j=1}^{\infty} \Psi_j \Psi_{j+h} \\
&= \sigma_v^2 \alpha_1 (\alpha_1 + \beta_1)^{h-1} + \sigma_v^2 \sum_{j=1}^{\infty} \alpha_1 (\alpha_1 + \beta_1)^{j-1} \alpha_1 (\alpha_1 + \beta_1)^{j+h-1} \\
&= \sigma_v^2 \alpha_1 (\alpha_1 + \beta_1)^{h-1} + \sigma_v^2 \sum_{j=1}^{\infty} \alpha_1 (\alpha_1 + \beta_1)^{j-1} \alpha_1 (\alpha_1 + \beta_1)^{j+h-1} \\
&= \sigma_v^2 \alpha_1 (\alpha_1 + \beta_1)^{h-1} + \sigma_v^2 \sum_{j=1}^{\infty} \alpha_1^2 (\alpha_1 + \beta_1)^{2j+h-2} \\
&= \sigma_v^2 \alpha_1 (\alpha_1 + \beta_1)^{h-1} + \sigma_v^2 \alpha_1^2 (\alpha_1 + \beta_1)^{h-2} \sum_{j=1}^{\infty} (\alpha_1 + \beta_1)^{2j} \\
&= \sigma_v^2 \left\{ \alpha_1 (\alpha_1 + \beta_1)^{h-1} + \frac{\alpha_1^2 (\alpha_1 + \beta_1)^h}{1 - (\alpha_1 + \beta_1)^2} \right\}, \text{ para todo } h \geq 0,
\end{aligned}$$

onde  $\sigma_v^2 = \text{Var}(v_t)$ , para todo  $t \in \mathbb{Z}$  e

$$\begin{aligned}
\rho_{X^2}(h) &= \left\{ \alpha_1 (\alpha_1 + \beta_1)^{h-1} + \frac{\alpha_1^2 (\alpha_1 + \beta_1)^h}{1 - (\alpha_1 + \beta_1)^2} \right\} \times \frac{1 - (\alpha_1 + \beta_1)^2}{1 - (\alpha_1 + \beta_1)^2 + \alpha_1^2} \\
&= \frac{\alpha_1 (\alpha_1 + \beta_1)^{h-1} - \alpha_1 (\alpha_1 + \beta_1)^{h+1} + \alpha_1^2 (\alpha_1 + \beta_1)^h}{1 - (\alpha_1 + \beta_1)^2 + \alpha_1^2}, \text{ para todo } h \geq 0.
\end{aligned}$$

## 2.3 Processos IGARCH e Modelos EWMA

Nesta seção apresentamos os processos IGARCH (*Integrate GARCH*) e os modelos EWMA (*Exponentially Weighted Moving Average*). Como veremos a seguir os modelos EWMA são um caso particular dos processos IGARCH. Tais modelos são amplamente utilizados devido à sua fácil implementação.

Os processos IGARCH foram apresentados por Engle e Bollerslev (1986). A introdução de tais processos deve-se ao fato que em muitas aplicações práticas, ao tentar estimar o polinômio dado pela expressão (2.8), encontra-se um polinômio  $1 - \hat{\alpha}(z) - \hat{\beta}(z)$ , com uma raiz estatisticamente indistingüível da unidade. A seguir apresentamos a definição formal destes processos.

**Definição 2.3. (Processos IGARCH).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Dizemos que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo GARCH *integrado*, ou IGARCH( $p, q$ ) se,

$$X_t = \sigma_t Z_t, \quad (2.31)$$

$$\sigma_t^2 = \alpha_0 + \alpha(\mathcal{B})X_t^2 + \beta(\mathcal{B})\sigma_t^2, \text{ para todo } t \in \mathbb{Z}, \quad (2.32)$$

onde  $\{Z_t\}_{t \in \mathbb{Z}}$  são variáveis aleatórias independentes e identicamente distribuídas com média zero e variância um e  $Z_t$  é independente de  $\mathcal{F}_{t-1}$ , para todo  $t \in \mathbb{Z}$ ,  $\alpha_0 \geq 0$  é uma constante real e  $\alpha(\cdot)$  e  $\beta(\cdot)$  são os polinômios definidos, respectivamente, pelas expressões (2.3) e (2.7), com

$$\alpha(1) + \beta(1) = 1.$$

Note que, se o polinômio  $1 - \alpha(\mathcal{B}) - \beta(\mathcal{B})$  possui uma raiz unitária, podemos decompô-lo como

$$1 - \alpha(\mathcal{B}) - \beta(\mathcal{B}) = \phi(\mathcal{B})(1 - \mathcal{B}),$$

com  $\phi(\mathcal{B}) \neq 0$ , para todo  $|\mathcal{B}| \leq 1$ . Assim, reescrevemos a equação (2.8) como

$$(1 - \alpha(\mathcal{B}) - \beta(\mathcal{B}))X_t^2 = \phi(\mathcal{B})(1 - \mathcal{B})X_t^2 = \alpha_0 + (1 - \beta(\mathcal{B}))v_t, \quad (2.33)$$

onde  $v_t = X_t^2 - \sigma_t^2$ , para todo  $t \in \mathbb{Z}$ . Segue que um processo IGARCH( $p, q$ ) possui uma representação ARIMA( $m, 1, q$ ), onde  $m = \max\{p, q\} - 1$ , para o processo  $\{X_t^2\}_{t \in \mathbb{Z}}$ .

Outro modelo amplamente utilizado para modelar a volatilidade é o modelo EWMA, introduzido por Roberts (1959). Como veremos na próxima definição este modelo pode ser visto como um caso particular dos processos IGARCH( $p, q$ ).

**Definição 2.4. (Modelos EWMA).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico e  $\mathcal{F}_t = \sigma(\{X_s\}; s \leq t)$ . Dizemos que a variância condicional de  $\{X_t\}_{t \in \mathbb{Z}}$  segue um modelo *média móvel ponderado exponencialmente*, denotado por EWMA (*Exponentially Weighted Moving Average*), se  $\sigma_t^2 = \text{Var}(X_t | \mathcal{F}_{t-1})$  é dada por

$$\sigma_t^2 = \lambda X_{t-1}^2 + (1 - \lambda)\sigma_{t-1}^2, \quad \text{para todo } t \in \mathbb{Z}, \quad (2.34)$$

onde  $0 < \lambda < 1$  é uma constante real.

Na prática, dada uma série temporal  $\{X_t\}_{t=1}^n$  obtida a partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , o parâmetro  $\lambda$  no modelo EWMA, é escolhido de forma a minimizar a soma dos quadrados dos resíduos de ajuste ou de previsão.

Pelas expressões (2.33) e (2.34), concluímos que o modelo EWMA pode ser visto como um modelo IGARCH( $p, q$ ), com  $p = 1 = q$  e  $\alpha_0 = 0$ .

Os modelos ARCH e GARCH (IGARCH) são amplamente utilizados para modelar séries de retornos financeiros (veja Definição 4.4). Como podemos perceber nas Definições 2.1 e 2.2, tais modelos tratam os retornos de forma simétrica, pois a volatilidade é uma função quadrática dos mesmos. Levando em conta que a volatilidade reage de forma assimétrica aos retornos, tendendo a ser maior para retornos negativos, Nelson (1991) introduz a classe de modelos EGARCH (*Exponential GARCH*), que apresentamos na próxima seção.

## 2.4 Processos EGARCH

Nesta seção definimos os processos EGARCH (*Exponential GARCH*), introduzidos por Nelson (1991) e apresentamos os principais resultados relacionados à estacionariedade e ergodicidade de tais processos.

**Definição 2.5. (Processos EGARCH).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Dizemos que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo GARCH *exponencial*, denotado por EGARCH( $p, q$ ) se,

$$X_t = \sigma_t Z_t, \quad (2.35)$$

$$\ln(\sigma_t^2) = \omega + \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})} g(Z_{t-1}), \quad \text{para todo } t \in \mathbb{Z}, \quad (2.36)$$



onde  $\omega \in \mathbb{R}$ ,  $\{Z_t\}_{t \in \mathbb{Z}}$  são variáveis aleatórias independentes e identicamente distribuídas com média zero e variância um,

$$\alpha(\mathcal{B}) \equiv \sum_{i=0}^p (-\alpha_i) \mathcal{B}^i = 1 - \sum_{i=1}^p \alpha_i \mathcal{B}^i, \quad \beta(\mathcal{B}) \equiv \sum_{j=0}^q (-\beta_j) \mathcal{B}^j = 1 - \sum_{j=1}^q \beta_j \mathcal{B}^j, \quad (2.37)$$

com  $\alpha_0 = -1 = \beta_0$ ,  $\beta(\mathcal{B}) \neq 0$  para todo  $\mathcal{B}$  tal que  $|\mathcal{B}| \leq 1$  e

$$g(Z_t) = \theta Z_t + \gamma[|Z_t| - \mathbb{E}(|Z_t|)], \quad \text{para todo } t \in \mathbb{Z}, \text{ e } \theta, \gamma \in \mathbb{R}. \quad (2.38)$$

Observe que  $g(Z_t)$  pode ser reescrita como

$$g(Z_t) = \begin{cases} (\theta + \gamma)Z_t - \gamma\mathbb{E}(|Z_t|), & Z_t \geq 0 \\ (\theta - \gamma)Z_t - \gamma\mathbb{E}(|Z_t|), & Z_t < 0, \end{cases}$$

para todo  $t \in \mathbb{Z}$ . Essa assimetria permite que a volatilidade responda mais rapidamente a retornos negativos do que a positivos.

O seguinte lema trata da independência entre quaisquer variáveis aleatórias e é utilizado na prova da Proposição 2.2 onde mostramos que o processo estocástico  $\{g(Z_t)\}_{t \in \mathbb{Z}}$ , com  $g(\cdot)$  definida pela expressão (2.38), é um processo ruído branco. Apesar de vários autores afirmarem que as variáveis aleatórias  $g(Z_t)$ , para todo  $t \in \mathbb{Z}$ , são independentes e identicamente distribuídas, nenhum deles apresenta uma prova formal dessa afirmação, nem do fato de  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  ser um processo ruído branco.

**Lema 2.2.** *Sejam  $X_1, \dots, X_n$  variáveis aleatórias independentes. Então, funções mensuráveis de famílias disjuntas das variáveis  $X_i$ 's são também independentes.*

**Prova:** Veja Durrett (1991), página 23. □

**Proposição 2.2.** *Seja  $\{Z_t\}_{t \in \mathbb{Z}}$  um processo estocástico tal que as variáveis aleatórias  $Z_t$ , para todo  $t \in \mathbb{Z}$ , são independentes e identicamente distribuídas, com média zero e variância um. Seja  $g(\cdot)$  a função definida pela expressão (2.38). Então  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  é um processo ruído branco.*

**Prova:** Seja  $g(x) = \theta x + \gamma|x| - c$ , com  $\theta, \gamma, c \in \mathbb{R}$ . É imediato que  $g(\cdot)$  é mensurável em relação à  $\sigma$ -álgebra de Borel  $\mathfrak{B}$ .

Por hipótese, para todo  $t \in \mathbb{Z}$ , as variáveis aleatórias  $Z_t$  possuem a mesma distribuição. Logo,  $\mathbb{E}(|Z_t|)$  é constante, para todo  $t \in \mathbb{Z}$ . Desta forma a função  $g(\cdot)$ , dada na expressão (2.38), pode ser reescrita como

$$g(Z_t) = \theta Z_t + \gamma|Z_t| - c, \quad \text{para todo } t \in \mathbb{Z},$$

com  $c = \gamma\mathbb{E}(|Z_t|)$ . Segue do Lema 2.2 que, para todo  $t \in \mathbb{Z}$ ,  $g(Z_t)$  e  $g(Z_{t+h})$  são independentes sempre que  $h \in \mathbb{Z} - \{0\}$ .

Observe que, como  $Z_t$  e  $Z_{t+h}$  são identicamente distribuídas

$$\mathbb{P}(Z_t \in A) = \mathbb{P}(Z_{t+h} \in A), \quad \text{para todo boreliano } A \in \mathfrak{B}.$$

Em particular, para todo  $A = (-\infty, x] \in \mathfrak{B}$ , com  $x \in \mathbb{R}$ , teremos

$$\begin{aligned}\mathbb{P}(g(Z_t) \leq x) &= \mathbb{P}\left(Z_t \in g^{-1}((-\infty, x])\right) \\ &= \mathbb{P}\left(Z_{t+h} \in g^{-1}((-\infty, x])\right) \\ &= \mathbb{P}(g(Z_{t+h}) \leq x).\end{aligned}$$

Portanto,  $g(Z_t)$  e  $g(Z_{t+h})$  possuem a mesma distribuição, para todo  $t, h \in \mathbb{Z}$ .

Já provamos que  $g(Z_t)$  e  $g(Z_{t+h})$  são independentes e identicamente distribuídas. Logo,

$$\begin{aligned}\mathbb{E}(g(Z_t)) &= \mathbb{E}[\theta Z_t + \gamma(|Z_t| - \mathbb{E}(|Z_t|))] \\ &= \mathbb{E}[\theta Z_t + \gamma|Z_t| - \gamma\mathbb{E}(|Z_t|)] \\ &= \theta\mathbb{E}(Z_t) + \gamma\mathbb{E}(|Z_t|) - \gamma\mathbb{E}(|Z_t|) \\ &= 0,\end{aligned}$$

pois  $\mathbb{E}(Z_t) = 0$ , para todo  $t \in \mathbb{Z}$ . E ainda, para todo  $h \in \mathbb{Z}$ , pela independência das variáveis aleatórias  $\{g(Z_t)\}_{t \in \mathbb{Z}}$ , temos que

$$\gamma_g(h) = \text{Cov}(g(Z_{t+h}), g(Z_t)) = \begin{cases} \sigma_g^2, & \text{se } h = 0 \\ 0, & \text{se } h \neq 0, \end{cases}$$

onde  $\sigma_g^2 = \mathbb{E}([g(Z_t)]^2)$ , para todo  $t \in \mathbb{Z}$ , e não depende de  $t$ .

Para completar a prova precisamos mostrar que  $\mathbb{E}(|g(Z_t)|^2) < \infty$ .

Note que,  $Z_t|Z_t| \leq Z_t^2$ , para todo  $t \in \mathbb{Z}$ . Segue, pelas propriedades da esperança, que

$$\mathbb{E}(Z_t|Z_t|) \leq \mathbb{E}(Z_t^2) = 1, \text{ para todo } t \in \mathbb{Z}. \quad (2.39)$$

Como  $\mathbb{E}(|Z_t|^2) = \mathbb{E}(Z_t^2) = 1$  e  $\mathbb{E}(Z_t) = 0$ , segue que

$$\begin{aligned}\mathbb{E}(|g(Z_t)|^2) &= \mathbb{E}(g^2(Z_t)) \\ &= \mathbb{E}\left[\left(\theta Z_t + \gamma\left[|Z_t| - \mathbb{E}(|Z_t|)\right]\right)^2\right] \\ &= \mathbb{E}\left[\theta^2 Z_t^2 + 2\theta Z_t \gamma\left(|Z_t| - \mathbb{E}(|Z_t|)\right) + \gamma^2\left(|Z_t| - \mathbb{E}(|Z_t|)\right)^2\right] \\ &= \mathbb{E}\left[\theta^2 Z_t^2 + 2\theta Z_t \gamma\left(|Z_t| - \mathbb{E}(|Z_t|)\right) + \gamma^2\left(|Z_t|^2 - 2|Z_t|\mathbb{E}(|Z_t|) + (\mathbb{E}(|Z_t|))^2\right)\right] \\ &= \theta^2\mathbb{E}(Z_t^2) + 2\theta\gamma\mathbb{E}(Z_t|Z_t|) - 2\theta\gamma\mathbb{E}(Z_t)\mathbb{E}(|Z_t|) \\ &\quad + \gamma^2\mathbb{E}(|Z_t|^2) - 2\gamma^2\mathbb{E}(|Z_t|)\mathbb{E}(|Z_t|) + \gamma^2(\mathbb{E}(|Z_t|))^2 \\ &= \theta^2\mathbb{E}(Z_t^2) + 2\theta\gamma\mathbb{E}(Z_t|Z_t|) - 0 + \gamma^2\mathbb{E}(Z_t^2) - 2\gamma^2(\mathbb{E}(|Z_t|))^2 + \gamma^2(\mathbb{E}(|Z_t|))^2 \\ &= \theta^2 + 2\theta\gamma\mathbb{E}(Z_t|Z_t|) + \gamma^2 - c^2 \\ &= \theta^2 + \gamma^2 - c^2 + 2\theta\gamma\mathbb{E}(Z_t|Z_t|),\end{aligned} \quad (2.40)$$

onde  $c = \gamma\mathbb{E}(|Z_t|)$ .

Pelas expressões (2.39) e (2.40) segue que

$$\mathbb{E}(|g(Z_t)|^2) < \infty, \quad \text{para todo } t \in \mathbb{Z}$$

e não depende de  $t$ . Portanto,  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  é um processo ruído branco.  $\square$

Da Proposição 2.2 segue que o processo estocástico  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  é fracamente estacionário. No lema que segue mostramos que este processo é também estritamente estacionário e ergódico.

**Lema 2.3.** *Sejam  $\{Z_t\}_{t \in \mathbb{Z}}$  variáveis aleatórias independentes e identicamente distribuídas. Então, o processo estocástico  $\{g(Z_t)\}_{t \in \mathbb{Z}}$ , definido pela expressão (2.38), é estritamente estacionário e ergódico.*

**Prova:** Note que as variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  são independentes e identicamente distribuídas. Logo o resultado segue do Teorema 1.6. □

O próximo lema trata da convergência da soma de variáveis aleatórias. Este resultado será utilizado nas demonstrações que seguem neste capítulo. No Teorema 2.3, apresentamos resultados relacionados aos processos EGARCH( $p, q$ ). O teorema trata da estacionariedade dos processos  $\{X_t\}_{t \in \mathbb{Z}}$ ,  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  e  $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$ . Este teorema foi primeiramente apresentado por Nelson (1991). Entretanto, apresentamos uma prova mais detalhada do que a apresentada pelo autor. No que segue, denotamos por  $L^0(\Omega, \mathcal{F}, \mathbb{P})$  o conjunto de todas as variáveis aleatórias sobre  $(\Omega, \mathcal{F})$  que são finitas em quase toda parte.

**Lema 2.4.** *Sejam  $\{Y_t\}_{t \geq 0}$  variáveis aleatórias independentes tais que  $\mathbb{E}(Y_t) = 0$ , para todo  $t \geq 0$ . Se  $\sum_{k=0}^{\infty} \text{Var}(Y_k) < \infty$ , então  $\sum_{k=0}^{\infty} Y_k$  converge com probabilidade 1.*

**Prova:** Veja Billingsley (1995), página 289. □

**Teorema 2.3.** *Sejam  $\{X_t\}_{t \in \mathbb{Z}}$  e  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  os processos estocásticos definidos pelas expressões (2.35) e (2.36). Suponha que as constantes  $\gamma$  e  $\theta$ , dadas na expressão (2.38), não são ambas iguais a zero. Então,*

- (i) *os processos estocásticos  $\{X_t\}_{t \in \mathbb{Z}}$ ,  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  e  $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$ , onde  $\omega \in \mathbb{R}$ , são estritamente estacionários e ergódicos;*
- (ii) *o processo estocástico  $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$  é fracamente estacionário.*

**Prova:** Para provar (i) note que, por definição, o polinômio  $\beta(\cdot)$ , dado na expressão (2.37) não possui raízes no círculo unitário. Seja  $\Phi(\mathcal{B}) \equiv (\beta(\mathcal{B}))^{-1} \alpha(\mathcal{B})$ . Segue que  $\Phi(\cdot)$  possui uma representação por série de potências dada por

$$\Phi(\mathcal{B}) = \sum_{k=0}^{\infty} \Phi_k \mathcal{B}^k, \quad \text{com} \quad \sum_{k=0}^{\infty} \Phi_k^2 < \infty. \quad (2.41)$$

Na Proposição 2.2 mostramos que as variáveis aleatórias  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  são independentes e identicamente distribuídas com  $\mathbb{E}(g(Z_t)) = 0$  e  $\text{Var}(g(Z_t)) < \infty$ , para todo  $t \in \mathbb{Z}$ . Aplicando o Lema 2.4 à seqüência de variáveis aleatórias  $\{\Phi_k g(Z_{t-1-k})\}_{k \geq 0}$ , para todo  $t \in \mathbb{Z}$  e  $\Phi_k$ , para todo  $k \geq 0$ , definido pela expressão (2.41), concluímos que as variáveis aleatórias  $\ln(\sigma_t^2) - \omega \in L^0(\Omega, \mathcal{F}, \mathbb{P})$ , para todo  $t \in \mathbb{Z}$ . Segue que  $\sigma_t^2$  e  $X_t \in L^0(\Omega, \mathcal{F}, \mathbb{P})$ , para todo  $t \in \mathbb{Z}$ .

Pelo Lema 2.3, o processo estocástico  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  é estritamente estacionário e ergódico. Segue, da expressão (2.36) e do Teorema 1.6, que  $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$  é estritamente estacionário e ergódico. De forma análoga, o processo estocástico  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  é estritamente

estacionário e ergódico. Pelas expressões (2.35), (2.36) e (2.41) temos

$$\begin{aligned} X_t &= \sigma_t Z_t \\ &= Z_t \exp \left\{ \frac{1}{2} \left( \omega + \sum_{k=0}^{\infty} \Phi_k g(Z_{t-1-k}) \right) \right\}, \quad \text{para todo } t \in \mathbb{Z}, \end{aligned} \quad (2.42)$$

Como as variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  são independentes e identicamente distribuídas, segue do Lema 1.1 e do Teorema 1.6 que o processo  $\{X_t\}_{t \in \mathbb{Z}}$  é estritamente estacionário e ergódico.

Pelo item (i) o processo  $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$  é estritamente estacionário. Logo, para provar (ii) só precisamos mostrar que  $\mathbb{E}((\ln(\sigma_t^2) - \omega)^2) < \infty$ , para todo  $t \in \mathbb{Z}$ . Note que, pela Proposição 2.2,  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  é um processo ruído branco, logo  $\mathbb{E}([g(Z_t)]^2) < \infty$ , para todo  $t \in \mathbb{Z}$ , e não depende de  $t$  e ainda,  $\mathbb{E}(g(Z_{t+h})g(Z_t)) = 0$ , sempre que  $h \neq 0$ . Por isso e pela expressão (2.41) temos que

$$\begin{aligned} \mathbb{E}((\ln(\sigma_t^2) - \omega)^2) &= \mathbb{E} \left( \left[ \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})} g(Z_{t-1}) \right]^2 \right) \\ &= \mathbb{E} \left( \left[ \sum_{k=0}^{\infty} \Phi_k g(Z_{t-1-k}) \right]^2 \right) \\ &= \mathbb{E}([g(Z_t)]^2) \sum_{k=0}^{\infty} \Phi_k^2 < \infty. \end{aligned}$$

Segue que o processo estocástico  $\{\ln(\sigma_t^2) - \omega\}_{t \in \mathbb{Z}}$  é fracamente estacionário.  $\square$

Pelo Teorema 2.3, os processos estocásticos  $\{X_t\}_{t \in \mathbb{Z}}$  e  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  são estritamente estacionários e ergódicos. Isto porém, não implica estacionariedade fraca pois, para alguma distribuição das variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$ , a média ou a variância destes processos pode não ser finita (veja Nelson, 1991).

Apresentamos a seguir, resultados para uma família de distribuições amplamente utilizadas, a distribuição de erro generalizada (GED), da qual fazem parte as funções de distribuição normal, dupla exponencial e uniforme. Para esta família de distribuições os processos  $\{X_t\}_{t \in \mathbb{Z}}$  e  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  possuem momentos de ordem  $r$  finitos, para todo  $r \in \mathbb{N} - \{0\}$ . Na Observação 2.1 da Seção 2.5 mostramos que os resultados apresentados a seguir podem ser estendidos aos processos FIEGARCH.

**Definição 2.6. (Distribuição de Erro Generalizada).** Seja  $X_t$  uma variável aleatória, para cada  $t \in \mathbb{Z}$ . A função de distribuição de  $X_t$  é uma *função de distribuição de erro generalizada*, denotada por GED, se a função densidade de probabilidade da variável aleatória normalizada, é dada por

$$f_X(z) = \frac{v \exp \left\{ -\frac{1}{2} \left| \frac{z}{\eta} \right|^v \right\}}{\eta 2^{1+1/v} \Gamma \left( \frac{1}{v} \right)}, \quad z \in \mathbb{R}, \quad v > 0,$$

onde  $\Gamma(\cdot)$  é a função Gama e

$$\eta \equiv \left[ 2^{(-2/v)} \frac{\Gamma(1/v)}{\Gamma(3/v)} \right]^{1/2}.$$

O parâmetro  $v$  determina o comportamento das caudas da distribuição. Quando  $v = 2$ , a variável aleatória tem distribuição normal padrão. Para  $v < 2$ , a distribuição possui caudas mais leves que a normal (quando  $v = 1$  temos a distribuição exponencial dupla) e para  $v > 2$  a função de distribuição possui caudas mais pesadas que a distribuição normal.

Note que, como as variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  são independentes e identicamente distribuídas, pela expressão (2.42) temos,

$$\begin{aligned}
\mathbb{E}(X_t^r) &= \mathbb{E} \left( Z_t^r \exp \left\{ \frac{r}{2} \left( \omega + \sum_{k=0}^{\infty} \Phi_k g(Z_{t-1-k}) \right) \right\} \right) \\
&= e^{r\omega/2} \mathbb{E} \left( Z_t^r \exp \left\{ \frac{r}{2} \sum_{k=0}^{\infty} \Phi_k g(Z_{t-1-k}) \right\} \right) \\
&= e^{r\omega/2} \mathbb{E}(Z_t^r) \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \left\{ \frac{r}{2} \Phi_k g(Z_{t-1-k}) \right\} \right) \\
&= e^{r\omega/2} \mathbb{E}(Z_t^r) \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \left\{ \frac{r}{2} \Phi_k g(Z_t) \right\} \right), \tag{2.43}
\end{aligned}$$

onde os coeficientes  $\Phi_k$  são dados pela expressão (2.41) e  $g(\cdot)$  é dada por (2.38).

O teorema que segue apresenta resultados que nos permitem avaliar a expressão (2.43) para o caso em que as variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  possuem distribuição GED.

**Teorema 2.4.** *Dados  $r \in \mathbb{N}$  e  $b \in \mathbb{R}$ , sejam  $\{Z_t\}_{t \in \mathbb{Z}}$  variáveis aleatórias independente e identicamente distribuídas, com média zero e variância um, e  $g(\cdot)$  a função definida pela expressão (2.38).*

- (i) *Se, para todo  $t \in \mathbb{Z}$ , a variável aleatória  $Z_t$  tem distribuição GED( $v$ ), com  $v > 1$ , então*

$$\begin{aligned}
\mathbb{E}(Z_t^r \exp\{b g(Z_t)\}) &= 2^{r/v} \eta^r \exp \left\{ \frac{-b \gamma \Gamma(2/v) \lambda 2^{1/v}}{\Gamma(1/v)} \right\} \\
&\times \sum_{k=0}^{\infty} (2^{1/v} \eta b)^k [(\gamma + \theta)^k + (-1)^r (\gamma - \theta)^k] \frac{\Gamma((r+k+1)/v)}{2\Gamma(1/v)\Gamma(k+1)} < \infty.
\end{aligned}$$

- (ii) *Se, para todo  $t \in \mathbb{Z}$ ,  $Z_t$  tem distribuição GED( $v$ ), com  $v < 1$ , ou  $Z_t$  tem distribuição  $t$ -Student com  $\nu > 2$  graus de liberdade, então  $\mathbb{E}(z^r \exp\{b g(Z_t)\}) < \infty$  se, e somente se*

$$b \gamma + |b \theta| \leq 0.$$

- (iii) *Se, para todo  $t \in \mathbb{Z}$ , a variável aleatória  $Z_t$  tem distribuição GED( $v$ ), com  $v = 1$ , então  $\mathbb{E}(z^r \exp\{b g(Z_t)\}) < \infty$  se, e somente*

$$b \gamma + |b \theta| < 2^{1/2}.$$

**Prova:** Veja Nelson (1991). □

Com base nos resultados do Teorema 2.4 enunciamos o Teorema 2.5 que trata da existência dos momentos de ordem  $r$ , para  $r \in \mathbb{N} - 0$ , dos processos EGARCH( $p, q$ ).

**Teorema 2.5.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo EGARCH( $p, q$ ), definido pelas expressões (2.35), (2.36) e (2.38). Suponha que  $\theta$  e  $\gamma$  não são ambos iguais a zero. Se, para todo  $t \in \mathbb{Z}$ , as variáveis aleatórias  $Z_t$  possuem distribuição GED( $v$ ), com  $v > 1$ , então os processos estocásticos  $\{X_t\}_{t \in \mathbb{Z}}$  e  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  são tais que  $\mathbb{E}(X_t^r) < \infty$  e  $\mathbb{E}((\sigma_t^2)^r) < \infty$ , para todo  $t \in \mathbb{Z}$  e todo  $r > 0$ .*

**Prova:** Veja Nelson (1991). □

Na próxima seção tratamos, com detalhes, os processos FIEGARCH( $p, d, q$ ), introduzidos por Bollerslev e Mikkelsen (1996). Ressaltamos que os processos EGARCH( $p, q$ ) são um caso particular dos processos FIEGARCH( $p, d, q$ ), quando  $d = 0$ .

## 2.5 Processos FIEGARCH

Nesta seção apresentamos o processo EGARCH *Fracionariamente Integrado*, denotado por FIEGARCH( $p, d, q$ ), (*Fractionally Integrated Exponential Generalized Autoregressive Conditional Heteroskedasticity*). Esta classe de processos, introduzida por Bollerslev e Mikkelsen (1996), além da volatilidade variando com o tempo e *clusters* de volatilidade (efeitos oriundos dos processos ARCH e GARCH), leva em conta a longa dependência da volatilidade, e a assimetria (oriunda da parte exponencial no processo EGARCH). Além das principais propriedades deste processo, apresentamos resultados relacionados à ordem de convergência dos coeficientes do polinômio que descreve a volatilidade. Estes resultados são de extrema importância para o desenvolvimento das simulações apresentadas no Capítulo 5.

**Definição 2.7. (Processos FIEGARCH).** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Dizemos que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo EGARCH *Fracionariamente Integrado*, e denotamos por FIEGARCH( $p, d, q$ ), se*

$$X_t = \sigma_t Z_t, \quad (2.44)$$

$$\ln(\sigma_t^2) = \omega + \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})(1 - \mathcal{B})^d} g(Z_{t-1}), \quad \text{para todo } t \in \mathbb{Z}, \quad (2.45)$$

onde  $\omega \in \mathbb{R}$ ,  $\{Z_t\}_{t \in \mathbb{Z}}$  são variáveis aleatórias independentes e identicamente distribuídas, com média zero e variância um,  $g(\cdot)$  é a função definida pela expressão (2.38), os polinômios  $\alpha(\cdot)$  e  $\beta(\cdot)$  são dados pela expressão (2.37) e o operador  $(1 - \mathcal{B})^d$  é definido pela expressão (1.17).

No Lema 2.1 mostramos que se  $\{X_t\}_{t \in \mathbb{Z}}$  é processo GARCH( $p, q$ ), então  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *martingale difference*. É imediato esse resultado pode ser estendido aos processos FIEGARCH( $p, d, q$ ) e a demonstração é análoga.

Na literatura encontramos diferentes maneiras de definir um processo FIEGARCH. Na próxima proposição provamos que sob certas restrições, dadas na expressão (2.47), a expressão (2.45), que faz parte da definição de um processo FIEGARCH( $p, d, q$ ), proposta por Bollerslev e Mikkelsen (1996), é equivalente à apresentada na equação (2.46). A expressão (2.46) faz parte da definição do mesmo processo, proposta por Zivot e Wang (2005). Este resultado é de extrema importância para o desenvolvimento do Capítulo 5. Ressaltamos que a prova formal desse resultado não aparece na literatura.

**Proposição 2.3.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ), dado na Definição 2.7. Então, a expressão de  $\ln(\sigma_t^2)$ , dada em (2.45), pode ser reescrita como*

$$\beta(\mathcal{B})(1 - \mathcal{B})^d \ln(\sigma_t^2) = a + \sum_{i=0}^p \left( \psi_i |Z_{t-1-i}| + \gamma_i Z_{t-1-i} \right), \quad (2.46)$$

onde

$$\begin{aligned} a &= (-\gamma)\alpha(1)\mathbb{E}(|Z_t|) \\ \psi_0 &= -\gamma\alpha_0 = -\gamma(-1) = \gamma, \\ \psi_i &= -\gamma\alpha_i, \quad \text{para todo } 1 \leq i \leq p, \\ \gamma_0 &= -\theta\alpha_0 = -\theta(-1) = \theta, \\ \gamma_i &= -\theta\alpha_i, \quad \text{para todo } 1 \leq i \leq p. \end{aligned} \quad (2.47)$$

**Prova:** Note que, pela igualdade (1.18),

$$0 = (1 - 1)^d = 1 - \sum_{k=1}^{\infty} \delta_{d,k},$$

ou seja,

$$\sum_{k=1}^{\infty} \delta_{d,k} = 1. \quad (2.48)$$

Pela igualdade (2.48), segue que

$$\begin{aligned} (1 - \mathcal{B})^d \omega &= \left( 1 - \sum_{k=1}^{\infty} \delta_{d,k} \mathcal{B}^k \right) \omega \\ &= \omega - \sum_{k=1}^{\infty} \delta_{d,k} \omega \\ &= \omega \left( 1 - \sum_{k=1}^{\infty} \delta_{d,k} \right) = 0, \end{aligned}$$

para todo  $\omega \in \mathbb{R}$ .

A expressão (2.45) pode ser reescrita como

$$\begin{aligned} \beta(\mathcal{B})(1 - \mathcal{B})^d \ln(\sigma_t^2) &= \beta(\mathcal{B})(1 - \mathcal{B})^d \omega + \alpha(\mathcal{B})g(Z_{t-1}) \\ &= 0 + \alpha(\mathcal{B})g(Z_{t-1}). \end{aligned} \quad (2.49)$$

Da expressão (2.38) podemos reescrever a expressão (2.49) por

$$\beta(\mathcal{B})(1 - \mathcal{B})^d \ln(\sigma_t^2) = \left( 1 - \sum_{i=1}^p \alpha_i \mathcal{B}^i \right) \left( \theta Z_t + \gamma [|Z_{t-1}| - \mathbb{E}(|Z_{t-1}|)] \right). \quad (2.50)$$

Dado que as variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  são independentes e identicamente distribuídas temos que  $\mathbb{E}(|Z_t|)$  é constante, para todo  $t \in \mathbb{Z}$ . Assim,

$$\begin{aligned} \left( 1 - \sum_{i=1}^p \alpha_i \mathcal{B}^i \right) \left( -\gamma \mathbb{E}(|Z_{t-1}|) \right) &= \left( 1 - \sum_{i=1}^p \alpha_i \right) \left( -\gamma \mathbb{E}(|Z_t|) \right) \\ &= (-\gamma)\alpha(1)\mathbb{E}(|Z_t|) = a. \end{aligned} \quad (2.51)$$

Segue, pela expressão (2.51), que a expressão (2.50) pode ser reescrita como

$$\begin{aligned}\beta(\mathcal{B})(1 - \mathcal{B})^d \ln(\sigma_t^2) &= a + \left(1 - \sum_{i=1}^p \alpha_i \mathcal{B}^i\right) (\theta Z_{t-1} + \gamma |Z_{t-1}|) \\ &= a + \sum_{i=0}^p \left(\psi_i |Z_{t-1-i}| + \gamma_i Z_{t-1-i}\right),\end{aligned}$$

onde  $\psi_i$ , para todo  $0 \leq i \leq p$ , e  $\gamma_i$ , para todo  $0 \leq i \leq p$ , são dados pela expressão (2.47). Ou seja,

$$\beta(\mathcal{B})(1 - \mathcal{B})^d \ln(\sigma_t^2) = a + \sum_{i=0}^p \left(\psi_i |Z_{t-1-i}| + \gamma_i Z_{t-1-i}\right),$$

com  $a$ ,  $\psi_i$  e  $\gamma_i$ , para  $0 \leq i \leq p$ , definidos pela expressão (2.47). Isto mostra que, sob as condições dadas, as expressões (2.45) e (2.46) são equivalentes.  $\square$

É fácil ver que a definição dada por Zivot e Wang (2005) é mais geral que a apresentada por Bollerslev e Mikkelsen (1996) já que os coeficientes  $\psi_j$  e  $\gamma_j$ , para  $j = 0, \dots, p$  não precisam, necessariamente, satisfazer as condições da expressão (2.47).

Seja  $\lambda(\mathcal{B}) \equiv \beta(\mathcal{B})^{-1}(1 - \mathcal{B})^{-d}\alpha(\mathcal{B})$  o polinômio definido na expressão (2.45). Na próxima proposição apresentamos uma fórmula de recorrência para o cálculo dos coeficientes  $\lambda_{d,k}$ , para todo  $k \in \mathbb{N}$ , deste polinômio. Utilizamos esta fórmula nas simulações apresentadas no Capítulo 5. Apesar dos processos FIEGARCH terem sido introduzidos em 1996, não encontramos na literatura um resultado, equivalente ao apresentado na Proposição 2.4, para tais processos.

**Proposição 2.4.** *Sejam*

$$\alpha(\mathcal{B}) = \sum_{i=0}^p (-\alpha_i) \mathcal{B}^i \quad \text{e} \quad \beta(\mathcal{B}) = \sum_{j=0}^q (-\beta_j) \mathcal{B}^j,$$

com  $\alpha_0 = -1 = \beta_0$ . Suponha que  $\beta(\cdot)$  não possui raízes no círculo unitário e que  $\alpha(\cdot)$  e  $\beta(\cdot)$  não possuem raízes em comum. Se

$$\lambda(\mathcal{B}) = \beta(\mathcal{B})^{-1}(1 - \mathcal{B})^{-d}\alpha(\mathcal{B}), \tag{2.52}$$

então, os coeficientes  $\lambda_{d,k}$ , para todo  $k \in \mathbb{N}$ , do polinômio  $\lambda(\mathcal{B}) = \sum_{k=0}^{\infty} \lambda_{d,k} \mathcal{B}^k$ , são dados por

$$\begin{aligned}\lambda_{d,0} &= 1 \\ \lambda_{d,k} &= -\alpha_k - \sum_{i=0}^{k-1} \lambda_i \left( \sum_{j=0}^{k-i} \delta_{d,k-i-j} \beta_j \right), \quad \text{para todo } k \geq 1.\end{aligned} \tag{2.53}$$

Na expressão (2.53), por definição,  $\alpha_k = 0$ , se  $k > p$ ,  $\beta_k = 0$ , se  $k > q$  e os coeficientes  $\delta_{d,k}$ , para todo  $k \in \mathbb{N}$ , são dados na expressão (1.19).



**Prova:** Por definição

$$\alpha(\mathcal{B}) = \beta(\mathcal{B})(1 - \mathcal{B})^d \left( \sum_{k=0}^{\infty} \lambda_{d,k} \mathcal{B}^k \right). \quad (2.54)$$

Note que, pela Proposição 1.2, item (ii), se definirmos  $\beta_k = 0$ , para todo  $k > q$ , e utilizarmos a expressão (1.18), temos

$$\begin{aligned} \beta(\mathcal{B})(1 - \mathcal{B})^d &= \left( \sum_{k=0}^{\infty} -\beta_k \mathcal{B}^k \right) \left( \sum_{k=0}^{\infty} -\delta_{d,k} \mathcal{B}^k \right) \\ &= \sum_{k=0}^{\infty} \left( \sum_{i=0}^k (-\beta_i)(-\delta_{d,k-i}) \right) \mathcal{B}^k \\ &= \sum_{k=0}^{\infty} \left( \sum_{i=0}^k \beta_i \delta_{d,k-i} \right) \mathcal{B}^k \\ &= \sum_{k=0}^{\infty} \left( \sum_{i=0}^k \delta_{d,k-i} \beta_i \right) \mathcal{B}^k. \end{aligned} \quad (2.55)$$

Assim,

$$\begin{aligned} \beta(\mathcal{B})(1 - \mathcal{B})^d \left( \sum_{k=0}^{\infty} \lambda_{d,k} \mathcal{B}^k \right) &= \left( \sum_{k=0}^{\infty} \left( \sum_{j=0}^n \delta_{d,k-j} \beta_j \right) \mathcal{B}^k \right) \left( \sum_{k=0}^{\infty} \lambda_{d,k} \mathcal{B}^k \right) \\ &= \sum_{k=0}^{\infty} \left( \sum_{i=0}^k \lambda_{d,i} \left( \sum_{j=0}^{k-i} \delta_{d,k-i-j} \beta_j \right) \right) \mathcal{B}^k \\ &= \sum_{k=0}^{\infty} \left( \lambda_{d,k} \delta_{d,0} \beta_0 + \sum_{i=0}^{k-1} \lambda_{d,i} \left( \sum_{j=0}^{k-i} \delta_{d,k-i-j} \beta_j \right) \right) \mathcal{B}^k \\ &= \sum_{k=0}^{\infty} \left( \lambda_{d,k} + \sum_{i=0}^{k-1} \lambda_{d,i} \left( \sum_{j=0}^{k-i} \delta_{d,k-i-j} \beta_j \right) \right) \mathcal{B}^k. \end{aligned} \quad (2.56)$$

Portanto, pela Proposição 1.2, item (i), se definirmos  $\alpha_k = 0$ , para todo  $k > p$ , pela expressão (2.56), a igualdade dada pela expressão (2.54) é verdadeira se, e somente se,

$$-\alpha_0 = \lambda_{d,0}, \text{ ou seja, } \lambda_{d,0} = 1$$

e

$$-\alpha_k = \lambda_{d,k} + \sum_{i=0}^{k-1} \lambda_{d,i} \left( \sum_{j=0}^{k-i} \delta_{d,k-i-j} \beta_j \right),$$

para todo  $k \geq 1$ . Assim temos

$$\lambda_{d,k} = -\alpha_k - \sum_{i=0}^{k-1} \lambda_{d,i} \left( \sum_{j=0}^{k-i} \delta_{d,k-i-j} \beta_j \right), \text{ para todo } k \geq 1.$$

Suponha agora que  $\lambda(\mathcal{B}) = \sum_{k=0}^{\infty} \lambda_{d,k} \mathcal{B}^k$  e que os coeficientes  $\lambda_{d,k}$ ,  $k \geq 0$ , sejam definidos pela expressão (2.53). É fácil ver que, substituindo os coeficientes  $\lambda_{d,k}$ , para todo  $k \geq 0$ , na expressão (2.56) temos como resultado  $\sum_{k=0}^{\infty} (-\alpha_k) \mathcal{B}^k = \alpha(\mathcal{B})$ . Isto completa a prova.  $\square$

No próximo lema apresentamos uma aproximação para os coeficientes  $\pi_{d,k}$  do polinômio  $(1 - \mathcal{B})^{-d}$ , quando  $k \rightarrow \infty$ . Este resultado, já conhecido, é utilizado para mostrar que a ordem de convergência ao infinito dos coeficientes  $\lambda_{d,k}$ , dados na expressão (2.53), é dada por  $\lambda_{d,k} = o(k^d)$ . Esse resultado é apresentado no Teorema 2.6. A prova formal dessa afirmação não é encontrada na literatura existente sobre processos FIEGARCH.

**Lema 2.5.** *Seja  $(1 - \mathcal{B})^{-d} = \sum_{i=0}^{\infty} \pi_{d,k} \mathcal{B}^k$ , com  $-0.5 < d < 0.5$ . Então, a ordem de convergência ao infinito para os coeficientes  $\pi_{d,k}$  é dada por*

$$\pi_{d,k} = o(k^d), \quad \text{quando } k \rightarrow \infty.$$

**Prova:** Para provar o lema, mostraremos que

$$\pi_{d,k} \sim \frac{1}{\Gamma(d)k^{1-d}}, \quad \text{quando } k \rightarrow \infty.$$

Note que, deste resultado segue que  $\lim_{k \rightarrow \infty} \frac{\pi_{d,k}}{k^d} = 0$ , ou seja, pela Definição 1.24,  $\pi_{d,k} = o(k^d)$ .

Note que  $(1 - \mathcal{B})^{-d}$  possui a mesma expansão binomial que o polinômio dado na expressão (1.18), substituindo-se  $d$  por  $-d$ . Segue que

$$\pi_{d,k} = \frac{d\Gamma(k+d)}{\Gamma(k+1)\Gamma(1+d)} = \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)} = -\delta_{-d,k}, \quad (2.57)$$

para todo  $k \geq 0$  com  $\delta_{d,k}$  definido pela expressão (1.19).

Pela fórmula de Stirling (veja Karlin e Taylor, 1975, página 36), para  $k$  suficientemente grande,

$$\Gamma(k) \sim \sqrt{2\pi} e^{-k} k^{k+1/2}. \quad (2.58)$$

Além disso,

$$\lim_{k \rightarrow \infty} \left(1 + \frac{d}{k}\right)^k = e^d.$$

Como  $\lim_{k \rightarrow \infty} \left(1 + \frac{d}{k}\right)^{1/2} = 1$ , segue que

$$\lim_{k \rightarrow \infty} \left(\frac{k+d}{k}\right)^{k+1/2} = \lim_{k \rightarrow \infty} \left[ \left(1 + \frac{d}{k}\right)^k \left(1 + \frac{d}{k}\right)^{1/2} \right] = e^d. \quad (2.59)$$

E ainda

$$\frac{(k+d)^d}{k} \sim \frac{1}{k^{1-d}}, \quad \text{quando } k \rightarrow \infty. \quad (2.60)$$

Assim, pela expressão (2.58) temos que, para  $k$  suficientemente grande,

$$\begin{aligned} \pi_{d,k} &= \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)} = \frac{\Gamma(k+d)}{k\Gamma(k)\Gamma(d)} \\ &\sim \frac{1}{k\Gamma(d)} \cdot \frac{\sqrt{2\pi} e^{-(k+d)} (k+d)^{(k+d)+1/2}}{\sqrt{2\pi} e^{-k} k^{k+1/2}} \\ &= \frac{e^{-d}}{\Gamma(d)} \frac{(k+d)^d}{k} \left(\frac{k+d}{k}\right)^{k+1/2}. \end{aligned}$$

Pelas expressões (2.59) e (2.60) segue que, para  $k$  suficientemente grande,

$$\pi_{d,k} \sim \frac{1}{\Gamma(d)k^{1-d}}.$$

Assim,  $\lim_{k \rightarrow \infty} \frac{\pi_{d,k}}{k^d} = 0$ . Logo, pela Definição 1.24,  $\pi_{d,k} = o(k^d)$ .  $\square$

No próximo teorema provamos que a ordem de convergência ao infinito dos coeficientes  $\lambda_{d,k}$ , dados na expressão (2.53), é dada por  $\lambda_{d,k} = o(k^d)$ . Este resultado é utilizado nas simulações apresentadas no Capítulo 5.

**Teorema 2.6.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ), com  $-0.5 < d < 0.5$ , definido pelas expressões (2.44) e (2.45). Seja  $\lambda(\cdot)$  o polinômio definido pela expressão (2.52), com coeficientes  $\lambda_{d,k}$  dados pela expressão (2.53). Então, a ordem de convergência ao infinito para os coeficientes  $\lambda_{d,k}$  é dada por*

$$\lambda_{d,k} = o(k^d), \quad \text{quando } k \rightarrow \infty.$$

**Prova:** Para a prova consideramos os casos  $p = 0 = q$  e  $p \geq 0$  e  $q \geq 0$ . Para mais detalhes veja os casos  $p > 0$  e  $q = 0$ ,  $p \geq 0$  e  $q = 1$  no Apêndice A.

CASO 1:  $p = 0 = q$ .

Note que, se  $p = 0 = q$ , pelas expressões (2.45) e (2.52) temos

$$\ln(\sigma_t^2) = \omega + \frac{1}{(1 - \mathcal{B})^d} g(Z_{t-1})$$

e

$$\lambda(\mathcal{B}) = \frac{1}{(1 - \mathcal{B})^d}.$$

Portanto, para todo  $k \geq 0$ ,  $\lambda_{d,k} = \pi_{d,k}$ , com  $\pi_{d,k}$  dado pela expressão (2.57). Segue, do Lema 2.5 que

$$\lambda_{d,k} = o(k^d), \quad \text{quando } k \rightarrow \infty.$$

CASO 2:  $p \geq 0$  e  $q \geq 0$ .

Neste caso, pela expressão (2.45), temos

$$\ln(\sigma_t^2) = \omega + \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})(1 - \mathcal{B})^d} g(Z_{t-1}).$$

Lembramos que o polinômio  $\beta(\cdot)$  não possui raízes no círculo unitário. Definindo  $f(\mathcal{B}) \equiv \beta(\mathcal{B})^{-1}$ , temos (ver Ahlfors, 1979)

$$\beta(\mathcal{B})^{-1} \equiv f(\mathcal{B}) = \sum_{k=0}^{\infty} f_k \mathcal{B}^k,$$

onde  $f_k = \frac{f^{(k)}(0)}{k!}$ , para todo  $k \geq 0$ . Segue, pelas expressões (2.52) e (2.57), que

$$\begin{aligned} \lambda(\mathcal{B}) &= \beta(\mathcal{B})^{-1} (1 - \mathcal{B})^{-d} \alpha(\mathcal{B}) \\ &= \left( \sum_{k=0}^{\infty} \left( \sum_{j=0}^k \pi_{d,k-j} f_j \right) \mathcal{B}^k \right) \left( \sum_{i=0}^p (-\alpha_i) \mathcal{B}^i \right) \\ &= \sum_{k=0}^{\infty} \left( \sum_{i=0}^{\min\{p,k\}} (-\alpha_i) \left( \sum_{j=0}^{k-i} \pi_{d,k-i-j} f_j \right) \right) \mathcal{B}^k. \end{aligned} \quad (2.61)$$

Da expressão (2.61) segue que, para todo  $k \geq 0$ ,

$$\lambda_{d,k} = \sum_{i=0}^{\min\{p,k\}} (-\alpha_i) \left( \sum_{j=0}^{k-i} \pi_{d,k-i-j} f_j \right).$$

Note que, para  $k > p$ , temos

$$\begin{aligned} \lambda_{d,k} &= \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=0}^{k-i} \pi_{d,k-i-j} f_j \right) \\ &= \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=0}^{k-i} \pi_{d,j} f_{k-i-j} \right). \end{aligned}$$

Além disso, como  $f_k \rightarrow 0$ , quando  $k \rightarrow \infty$ , segue que, dado  $\varepsilon > 0$ , existe  $k_0 > 0$  tal que, para todo  $m > 0$  fixo, e para  $k > k_0$ ,

$$|\pi_{d,j} f_{k-i-j}| < \frac{\varepsilon}{m},$$

para todo  $0 \leq j \leq m$  e para todo  $0 \leq i \leq p$ . Portanto, para  $k$  suficientemente grande,

$$\begin{aligned} \lambda_{d,k} &= \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=0}^m \pi_{d,j} f_{k-i-j} + \sum_{j=m+1}^{k-i} \pi_{d,j} f_{k-i-j} \right) \\ &\sim \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=m+1}^{k-i} \pi_{d,j} f_{k-i-j} \right). \end{aligned}$$

Note ainda que, como  $\pi_{d,k} \rightarrow 0$ , quando  $k \rightarrow \infty$ ,  $m$  pode ser escolhido de forma que

$$\pi_{d,k} \sim \pi_{d,k-i} \sim \pi_{d,j},$$

para todo  $m+1 \leq j \leq k-i$  e para todo  $0 \leq i \leq p$ . Logo,

$$\begin{aligned} \lambda_{d,k} &\sim \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=m+1}^{k-i} \pi_{d,j} f_{k-i-j} \right) \\ &\sim \sum_{i=0}^p (-\alpha_i) \pi_{d,k} \left( \sum_{j=m+1}^{k-i} f_{k-i-j} \right) \\ &= \pi_{d,k} \left( \sum_{i=0}^p (-\alpha_i) \right) \left( \sum_{j=0}^{k-i-(m+1)} f_j \right) \\ &\sim \pi_{d,k} \left( \sum_{i=0}^p (-\alpha_i) \right) \left( \sum_{j=0}^{\infty} f_j \right) = \pi_{d,k} \alpha(1) f(1) \\ &= \pi_{d,k} \frac{\alpha(1)}{\beta(1)}. \end{aligned}$$

Segue, pelo Lema 2.5 que

$$\lambda_{d,k} = o(k^d), \quad \text{quando } k \rightarrow \infty.$$

Isto conclui a prova. □

Na Proposição 2.2 mostramos que  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  é um processo ruído branco. Essa proposição é utilizada para provar o Teorema 2.7 que apresenta as características do processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ , dado pela expressão (2.45). Nesse teorema apresentamos resultados sobre a estacionariedade, ergodicidade, causalidade e inversibilidade desse processo, além da sua função densidade espectral. As afirmações deste teorema já foram apresentadas por muitos autores. Entretanto, a prova do teorema é inexistente na literatura.

**Teorema 2.7.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ), definido pelas expressões (2.44) e (2.45). Então, para o processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ , dado pela expressão (2.45) temos,*

- (i) *se  $d < 0,5$ , o processo é fracamente estacionário;*
- (ii) *se  $d > -0,5$  e  $\alpha(\mathcal{B}) \neq 0$ , para  $|\mathcal{B}| \leq 1$ , o processo é inversível;*
- (iii) *se  $d \in (-0,5; 0,5)$  e  $\alpha(\mathcal{B}) \neq 0$ , para  $|\mathcal{B}| \leq 1$ ,*

$$\rho_{\ln(\sigma^2)}(h) \sim ch^{2d-1}, \quad \text{quando } h \rightarrow \infty,$$

onde  $c \neq 0$  e

$$\begin{aligned} f_{\ln(\sigma^2)}(\lambda) &= \frac{\sigma_g^2 |\alpha(e^{-i\lambda})|^2}{2\pi |\beta(e^{-i\lambda})|^2} |1 - e^{-i\lambda}|^2 \\ &\sim \frac{\sigma_g^2}{2\pi} \left( \frac{\alpha(1)}{\beta(1)} \right)^2 \lambda^{-2d}, \quad \text{quando } \lambda \rightarrow 0; \end{aligned}$$

- (iv) *se  $d < 0,5$ , o processo é estritamente estacionário e ergódico.*

**Prova:** Pela Proposição 2.2 temos que o processo  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  é um processo ruído branco. Segue que, pela expressão (2.45), que  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, p$ ), para  $d \in (-0,5; 0,5)$ . Lembramos que, por definição, todas as raízes do polinômio  $\beta(\cdot)$  estão fora do círculo unitário. Sendo assim, os itens (i), (ii) e (iii) seguem do Teorema 1.5. Para o item (iv) seguimos as idéias da prova do Teorema 2.3. Note que, se  $d < 0,5$ , então

$$\begin{aligned} \lambda(\mathcal{B}) &= \beta(\mathcal{B})^{-1} (1 - \mathcal{B})^{-d} \alpha(\mathcal{B}) \\ &= \sum_{k=0}^{\infty} \lambda_{d,k} \mathcal{B}^k \end{aligned}$$

é tal que  $\sum_{k=0}^{\infty} \lambda_{d,k}^2 < \infty$ . Além disso, pela Proposição 2.2, as variáveis aleatórias  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  são independentes e identicamente distribuídas, com média zero e variância finita. Segue, pelo Teorema 1.7, que o processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$  é estritamente estacionário e ergódico. □

**Corolário 2.1.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ). Se  $d < 0,5$  os processos estocásticos  $\{X_t\}_{t \in \mathbb{Z}}$  e  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  são estritamente estacionários e ergódicos.*

**Prova:** A estacionariedade estrita e a ergodicidade dos processos  $\{X_t\}_{t \in \mathbb{Z}}$  e  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  são consequência imediata do item (iv) do Teorema 2.7 (veja a demonstração do Teorema 2.3). □

**Observação 2.1.** Para os processos EGARCH( $p, q$ ), Nelson (1991) utiliza a seguinte representação para o logaritmo da volatilidade

$$\ln(\sigma_t^2) = \alpha_t + \sum_{j=1}^{\infty} \beta_j g(Z_{t-j}), \quad \text{para todo } t \in \mathbb{Z}, \quad (2.62)$$

onde  $\{\alpha_t\}_{t \in \mathbb{Z}}$  e  $\{\beta_j\}_{j \geq 1}$  são seqüências de números reais. Note que, pelas expressões (2.45) e (2.52), segue que para os processos FIEGARCH( $p, d, q$ ), o logaritmo da volatilidade pode ser representado por

$$\ln(\sigma_t^2) = \omega + \sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{t-1-k}), \quad (2.63)$$

Segue que as expressões (2.62) e (2.63) são equivalentes, basta tomar  $\alpha_t = \omega$ , para todo  $t \in \mathbb{Z}$ , e  $\beta_j = \lambda_{d,j-1}$ , para todo  $j \geq 1$ . Assim, se  $d < 0,5$ , a série  $\sum_{k=0}^{\infty} \lambda_{d,k}^2 < \infty$  e os resultados apresentados no Teorema 2.5 são válidos também para os processos FIEGARCH( $p, d, q$ ), e a demonstração é análoga.

No Teorema 2.7 apresentamos resultados sobre a estacionariedade do processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ , que é não-observável. No Teorema 2.8, a seguir, apresentamos resultados sobre a estacionariedade do processo estocástico  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$ , que é observável. Ressaltamos que esse resultado não é encontrado na literatura. Consideramos apenas os processos FIEGARCH( $p, d, q$ ) em que o processo  $\{Z_t\}_{t \in \mathbb{Z}}$  é tal que  $\mathbb{E}([\ln(Z_t^2)]^2) < \infty$ , para todo  $t \in \mathbb{Z}$ . Note que, em particular, se  $Z_t \sim \mathcal{N}(0, 1)$  essa condição é satisfeita pois

$$\begin{aligned} \mathbb{E}([\ln(Z_t^2)]^2) &= \int_{-\infty}^{\infty} (\ln(z^2))^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\ &= 2 \int_0^{\infty} (\ln(z^2))^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\ &= (\ln(2) + \gamma)^2 + \frac{1}{2} \pi^5 < \infty, \end{aligned}$$

onde  $\gamma \approx 0,577$  é a constante de Euler. Para mais detalhes sobre o cálculo desta integral veja Gradshteyn e Ryzhik (2000), fórmula (4.335, FI II 808), e referências ali contidas.

**Teorema 2.8.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ), dado na Definição 2.7. Se  $d < 0,5$  e  $\mathbb{E}([\ln(Z_t^2)]^2) < \infty$ , para todo  $t \in \mathbb{Z}$ , então o processo  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é estacionário, no sentido da Definição 1.7.*

**Prova:** Note que, pela expressão (2.44) temos

$$\ln(X_t^2) = \ln(\sigma_t^2) + \ln(Z_t^2). \quad (2.64)$$

Pelo Teorema 2.7, para  $d < 0,5$ , o processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$  é estacionário. Por definição, as variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  são independentes e identicamente distribuídas. Segue, pelo Lema 2.2, que as variáveis aleatórias  $\{\ln(Z_t^2)\}_{t \in \mathbb{Z}}$  são independentes. É imediato que essas variáveis são identicamente distribuídas. Sendo assim,  $\mathbb{E}(\ln(Z_t^2))$  e  $\mathbb{E}([\ln(Z_t^2)]^2)$  não dependem de  $t \in \mathbb{Z}$ . Dessa forma, pela expressão (2.64), temos que

$$\mathbb{E}(\ln(X_t^2)) = \mathbb{E}(\ln(\sigma_t^2)) + \mathbb{E}(\ln(Z_t^2)),$$

e não depende de  $t \in \mathbb{Z}$ .

Pelas expressões (2.45) e (2.52),  $\sigma_t$  é uma função  $\mathcal{F}_{t-1}$ -mensurável, onde  $\mathcal{F}_{t-1} \equiv \sigma(\{Z_s\}; s \leq t-1)$ . Segue, pelo Lema 2.2, que  $\sigma_t$  e  $Z_t$  são independentes, para todo  $t \in \mathbb{Z}$ . O mesmo vale para qualquer função mensurável de  $\sigma_t$  e  $Z_t$ . Segue que,

$$\text{Var}(\ln(X_t^2)) = \text{Var}(\ln(\sigma_t^2)) + \text{Var}(\ln(Z_t^2)), \quad (2.65)$$

e não depende de  $t \in \mathbb{Z}$ .

Pela estacionariedade do processo  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$  temos que  $\text{Var}(\ln(\sigma_t^2)) < \infty$ . Por hipótese,  $\mathbb{E}([\ln(Z_t^2)]^2) < \infty$ . Segue que  $\text{Var}(\ln(Z_t^2)) < \infty$ . Logo,

$$\text{Var}(\ln(X_t^2)) < \infty.$$

Assim,  $\mathbb{E}(|\ln(X_t^2)|^2) < \infty$  e não depende de  $t \in \mathbb{Z}$ .

Resta mostrar que  $\gamma_{\ln(X^2)}(h) = \text{Cov}(\ln(X_{t+h}^2), \ln(X_t^2))$ , para todo  $h \in \mathbb{Z}$ , não depende de  $t \in \mathbb{Z}$ . Já consideramos o caso  $h = 0$  na expressão (2.65), tratamos agora do caso  $h \neq 0$ .

Note que, pela expressão (2.64),

$$\begin{aligned} \text{Cov}(\ln(X_{t+h}^2), \ln(X_t^2)) &= \text{Cov}(\ln(\sigma_{t+h}^2) + \ln(Z_{t+h}^2), \ln(\sigma_t^2) + \ln(Z_t^2)) \\ &= \text{Cov}(\ln(\sigma_{t+h}^2), \ln(\sigma_t^2)) + \text{Cov}(\ln(Z_{t+h}^2), \ln(Z_t^2)) \\ &\quad + \text{Cov}(\ln(\sigma_{t+h}^2), \ln(Z_t^2)) + \text{Cov}(\ln(Z_{t+h}^2), \ln(\sigma_t^2)). \end{aligned} \quad (2.66)$$

Dado que o processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$  é estacionário temos

$$\text{Cov}(\ln(\sigma_{t+h}^2), \ln(\sigma_t^2)) = \gamma_{\ln(\sigma^2)}(h), \quad (2.67)$$

isto é, não depende de  $t \in \mathbb{Z}$ . Pela independência das variáveis aleatórias  $\{\ln(Z_t^2)\}_{t \in \mathbb{Z}}$ , segue que,

$$\text{Cov}(\ln(Z_{t+h}^2), \ln(Z_t^2)) = 0, \quad \text{para todo } h \neq 0. \quad (2.68)$$

Pelas expressões (2.45) e (2.52), pela independência das variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  e pelo Lema 2.2 temos

$$\begin{aligned} \text{Cov}(\ln(\sigma_{t+h}^2), \ln(Z_t^2)) &= \text{Cov}\left(\omega + \sum_{i=0}^{\infty} \lambda_{d,i} g(Z_{t+h-1-i}), \ln(Z_t^2)\right) \\ &= \text{Cov}\left(\omega, \ln(Z_t^2)\right) + \sum_{i=0}^{\infty} \lambda_{d,i} \text{Cov}\left(g(Z_{t+h-1-i}), \ln(Z_t^2)\right) \\ &= \begin{cases} \lambda_{d,h-1} \text{Cov}\left(g(Z_t), \ln(Z_t^2)\right), & \text{se } h > 0 \\ 0, & \text{se } h < 0, \end{cases} \end{aligned} \quad (2.69)$$

que não depende de  $t \in \mathbb{Z}$  e

$$\begin{aligned} \text{Cov}(\ln(Z_{t+h}^2), \ln(\sigma_t^2)) &= \text{Cov}\left(\ln(Z_{t+h}^2), \omega + \sum_{i=0}^{\infty} \lambda_{d,i} g(Z_{t-1-i})\right) \\ &= \text{Cov}\left(\ln(Z_{t+h}^2), \omega\right) + \sum_{i=0}^{\infty} \lambda_{d,i} \text{Cov}\left(\ln(Z_{t+h}^2), g(Z_{t-1-i})\right) \\ &= \begin{cases} 0, & \text{se } h > 0 \\ \lambda_{d,-h-1} \text{Cov}\left(\ln(Z_{t+h}^2), g(Z_{t+h})\right), & \text{se } h < 0, \end{cases} \end{aligned} \quad (2.70)$$

que também não depende de  $t \in \mathbb{Z}$ .

Pelas expressões (2.67), (2.68), (2.69) e (2.70), segue que a expressão (2.66) não depende de  $t \in \mathbb{Z}$ . Portanto, o processo estocástico  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é estacionário.  $\square$

Lembramos que as variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  são independentes e identicamente distribuídas. Como  $\ln(\sigma_t^2)$  e  $\ln(Z_t^2)$ , para todo  $t \in \mathbb{Z}$ , são funções mensuráveis dessas variáveis aleatórias, segue do Lema 1.1 e do Teorema 1.6 que o processo estocástico  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é estritamente estacionário e ergódico. No próximo teorema provamos que, se  $\mathbb{E}([\ln(Z_t^2)]^2) < \infty$  e  $\mathbb{E}(Z_t \ln(Z_t^2)) < \infty$ , para todo  $t \in \mathbb{Z}$ , então o processo estocástico  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, 0$ ), com inovações que, em geral, são não-Gaussianas. Este resultado, que não aparece na literatura, é importante pois auxilia na seleção dos modelos para séries temporais. Lopes e Mendes (2006) apresentam um resultado semelhante para os modelos LMSV (*Long Memory Stochastic Volatility Models*), introduzidos por Breidt, Crato e de Lima (1998).

Note que, como a função  $h(z) = z \ln(z^2)$  é uma função ímpar, se a distribuição das variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  for simétrica então,  $\mathbb{E}(Z_t \ln(Z_t^2)) = 0$ , para todo  $t \in \mathbb{Z}$ .

**Teorema 2.9.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ), dado pela Definição 2.7. Se  $d \in (-0, 5; 0, 5)$ ,  $\mathbb{E}(\ln(Z_t^2)^2) < \infty$  e  $\mathbb{E}(Z_t \ln(Z_t^2)) < \infty$ , para todo  $t \in \mathbb{Z}$ , então o processo  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, 0$ ) dado por*

$$\beta(\mathcal{B})(1 - \mathcal{B})^d \ln(X_t^2) = \varepsilon_t, \text{ para todo } t \in \mathbb{Z},$$

onde  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  é um processo ruído branco com variância  $\sigma_\varepsilon^2$ .

**Prova:** Note que, pelas expressões (2.44) e (2.45) temos

$$\ln(X_t^2) = \omega + \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})(1 - \mathcal{B})^d} g(Z_{t-1}) + \ln(Z_t^2), \text{ para todo } t \in \mathbb{Z}. \quad (2.71)$$

Pela expressão (2.48) segue que  $\beta(\mathcal{B})(1 - \mathcal{B})^d \omega = 0$ . Sendo assim, podemos reescrever a expressão (2.71) como

$$\begin{aligned} \beta(\mathcal{B})(1 - \mathcal{B})^d \ln(X_t^2) &= \alpha(\mathcal{B})g(Z_{t-1}) + \beta(\mathcal{B})(1 - \mathcal{B})^d \ln(Z_t^2) \\ &= \varepsilon_t \end{aligned} \quad (2.72)$$

com

$$\varepsilon_t = \alpha(\mathcal{B})g(Z_{t-1}) + \beta(\mathcal{B})(1 - \mathcal{B})^d \ln(Z_t^2), \quad (2.73)$$

para todo  $t \in \mathbb{Z}$ .

Vamos provar que o processo  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ , dado na expressão (2.72), é um processo ruído branco. Note que, pela independência das variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  e como  $\ln(Z_t^2) < Z_t^2$ , para todo  $t \in \mathbb{Z}$ , segue que  $\mathbb{E}(\ln(Z_t^2)) < \mathbb{E}(Z_t^2) = 1$ , para todo  $t \in \mathbb{Z}$ . Portanto,  $\mathbb{E}(\ln(Z_t^2))$  é constante e finita, para todo  $t \in \mathbb{Z}$ . Da expressão (2.48), temos que

$$\beta(\mathcal{B})(1 - \mathcal{B})^d \mathbb{E}(\ln(Z_t^2)) = 0. \quad (2.74)$$

Logo, pela Proposição 2.2,

$$\alpha(\mathcal{B})\mathbb{E}(g(Z_{t-1})) = 0. \quad (2.75)$$



Assim, pela expressão (2.73) e pelas expressões (2.74) e (2.75) temos

$$\mathbb{E}(\varepsilon_t) = 0, \text{ para todo } t \in \mathbb{Z}.$$

Seja  $\phi(\mathcal{B}) = \beta(\mathcal{B})(1 - \mathcal{B})^d = \sum_{i=0}^{\infty} \phi_i \mathcal{B}^i$ . Note que

$$\begin{aligned} \text{Cov}(\varepsilon_{t+h}, \varepsilon_t) &= \text{Cov}\left(\alpha(\mathcal{B})g(Z_{t+h-1}) + \phi(\mathcal{B}) \ln(Z_{t+h}^2), \alpha(\mathcal{B})g(Z_{t-1}) + \phi(\mathcal{B}) \ln(Z_t^2)\right) \\ &= \text{Cov}\left(\alpha(\mathcal{B})g(Z_{t+h-1}), \alpha(\mathcal{B})g(Z_{t-1})\right) + \text{Cov}\left(\phi(\mathcal{B}) \ln(Z_{t+h}^2), \phi(\mathcal{B}) \ln(Z_t^2)\right) \\ &\quad + \text{Cov}\left(\alpha(\mathcal{B})g(Z_{t+h-1}), \phi(\mathcal{B}) \ln(Z_t^2)\right) + \text{Cov}\left(\phi(\mathcal{B}) \ln(Z_{t+h}^2), \alpha(\mathcal{B})g(Z_{t-1})\right). \end{aligned} \quad (2.76)$$

Pela Proposição 2.2 temos,

$$\begin{aligned} \text{Cov}\left(\alpha(\mathcal{B})g(Z_{t+h-1}), \alpha(\mathcal{B})g(Z_{t-1})\right) &= \text{Cov}\left(\sum_{i=0}^p \alpha_i g(Z_{t+h-1-i}), \sum_{j=0}^p \alpha_j g(Z_{t-1-j})\right) \\ &= \sum_{i=0}^p \sum_{j=0}^p \alpha_i \alpha_j \text{Cov}\left(g(Z_{t+h-1-i}), g(Z_{t-1-j})\right) \\ &= \begin{cases} \text{Var}(g(Z_t)) \sum_{i=|h|}^p \alpha_i \alpha_{i-|h|}, & \text{se } -p \leq h \leq p \\ 0, & \text{se } h < -p, \text{ ou } h > p, \end{cases} \end{aligned} \quad (2.77)$$

que não depende de  $t \in \mathbb{Z}$ . Pela independência das variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$ ,

$$\begin{aligned} \text{Cov}\left(\alpha(\mathcal{B})g(Z_{t+h-1}), \phi(\mathcal{B}) \ln(Z_t^2)\right) &= \text{Cov}\left(\sum_{i=0}^p \alpha_i g(Z_{t+h-1-i}), \sum_{j=0}^{\infty} \phi_j \ln(Z_{t-j}^2)\right) \\ &= \sum_{i=0}^p \sum_{j=0}^{\infty} \alpha_i \phi_j \text{Cov}\left(g(Z_{t+h-1-i}), \ln(Z_{t-j}^2)\right) \\ &= \begin{cases} \text{Cov}(g(Z_t), \ln(Z_t^2)) \sum_{i=0}^p \alpha_i \phi_{i-h+1}, & \text{se } h < 1 \\ \text{Cov}(g(Z_t), \ln(Z_t^2)) \sum_{i=h-1}^p \alpha_i \phi_{i-h+1}, & \text{se } 1 \leq h \leq p+1 \\ 0, & \text{se } h > p+1, \end{cases} \end{aligned} \quad (2.78)$$

que também não depende de  $t \in \mathbb{Z}$ . E, ainda,

$$\begin{aligned}
\text{Cov}\left(\phi(\mathcal{B}) \ln(Z_{t+h}^2), \alpha(\mathcal{B})g(Z_{t-1})\right) &= \text{Cov}\left(\sum_{i=0}^{\infty} \phi_i \ln(Z_{t+h-i}^2), \sum_{j=0}^p \alpha_j g(Z_{t-1-i})\right) \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^p \phi_i \alpha_j \text{Cov}\left(\ln(Z_{t+h-i}^2), g(Z_{t-1-i})\right) \\
&= \begin{cases} 0, & \text{se } h < -(p+1) \\ \text{Cov}\left(\ln(Z_t^2), g(Z_t)\right) \sum_{i=|h|-1}^p \alpha_i \phi_{i+h+1}, & \text{se } -(p+1) \leq h \leq -1 \\ \text{Cov}\left(\ln(Z_t^2), g(Z_t)\right) \sum_{i=0}^p \alpha_i \phi_{i+h+1}, & \text{se } h > -1, \end{cases}
\end{aligned} \tag{2.79}$$

que não depende de  $t \in \mathbb{Z}$ . Além disso, pela independência das variáveis aleatórias  $\{\ln(Z_t^2)\}_{t \in \mathbb{Z}}$  temos

$$\begin{aligned}
\text{Cov}\left(\phi(\mathcal{B}) \ln(Z_{t+h}^2), \phi(\mathcal{B}) \ln(Z_t^2)\right) &= \text{Cov}\left(\sum_{i=0}^{\infty} \phi_i \ln(Z_{t+h-i}^2), \sum_{j=0}^{\infty} \phi_j \ln(Z_{t-j}^2)\right) \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \phi_i \phi_j \text{Cov}\left(\ln(Z_{t+h-i}^2), \ln(Z_{t-j}^2)\right) \\
&= \text{Var}\left(\ln(Z_t^2)\right) \sum_{i=|h|}^{\infty} \phi_i \phi_{i-|h|}, \quad h \in \mathbb{Z},
\end{aligned} \tag{2.80}$$

que não depende de  $t \in \mathbb{Z}$ . Pelas expressões (2.77), (2.80), (2.78) e (2.79), concluímos que a expressão (2.76) não depende de  $t \in \mathbb{Z}$ .

Resta mostrar que  $\mathbb{E}(|\varepsilon_t|^2) < \infty$ . Para isso note que, como  $\mathbb{E}(\varepsilon_t) = 0$ , segue que  $\mathbb{E}(|\varepsilon_t|^2) = \text{Var}(\varepsilon_t) = \gamma_\varepsilon(0)$ . Pela expressão (2.76) temos

$$\begin{aligned}
\gamma_\varepsilon(0) &= \text{Cov}\left(\alpha(\mathcal{B})g(Z_{t-1}), \alpha(\mathcal{B})g(Z_{t-1})\right) + \text{Cov}\left(\phi(\mathcal{B}) \ln(Z_t^2), \phi(\mathcal{B}) \ln(Z_t^2)\right) \\
&\quad + \text{Cov}\left(\alpha(\mathcal{B})g(Z_{t-1}), \phi(\mathcal{B}) \ln(Z_t^2)\right) + \text{Cov}\left(\phi(\mathcal{B}) \ln(Z_t^2), \alpha(\mathcal{B})g(Z_{t-1})\right).
\end{aligned} \tag{2.81}$$

Pelas expressões (2.77), (2.80), (2.78), (2.79) e pela expressão (2.81) concluímos que

$$\gamma_\varepsilon(0) = \text{Var}(g(Z_t)) \sum_{i=0}^p \alpha_i^2 + \text{Var}(\ln(Z_t^2)) \sum_{i=0}^{\infty} \phi_i^2 + 2\text{Cov}(g(Z_t), \ln(Z_t^2)) \sum_{i=0}^p \alpha_i \phi_{i+1}. \tag{2.82}$$

Pela Proposição 2.2  $\text{Var}(g(Z_t)) < \infty$ . Por hipótese  $\mathbb{E}([\ln(Z_t^2)]^2) < \infty$ , logo  $\text{Var}(\ln(Z_t^2)) < \infty$ . Além disso,  $\sum_{i=0}^{\infty} \phi_i^2 < \infty$ , pois a série  $\sum_{i=0}^{\infty} \phi_i \mathcal{B}^i$  é convergente. Precisamos provar que  $\text{Cov}(g(Z_t), \ln(Z_t^2)) < \infty$ .

Note que,

$$\text{Cov}(g(Z_t), \ln(Z_t^2)) = \mathbb{E}(g(Z_t) \ln(Z_t^2)) - \mathbb{E}(g(Z_t))\mathbb{E}(\ln(Z_t^2)). \tag{2.83}$$

Pela Proposição 2.2,  $\mathbb{E}(g(Z_t)) = 0$ . Além disso, pela expressão (2.38), temos,

$$\begin{aligned}\mathbb{E}\left(g(Z_t) \ln(Z_t^2)\right) &= \mathbb{E}\left(\left(\theta Z_t + \gamma\left[|Z_t| - \mathbb{E}(|Z_t|)\right]\right) \ln(Z_t^2)\right) \\ &= \theta \mathbb{E}\left(Z_t \ln(Z_t^2)\right) + \gamma \mathbb{E}\left(|Z_t| \ln(Z_t^2)\right) - \gamma \mathbb{E}(|Z_t|) \mathbb{E}(\ln(Z_t^2)).\end{aligned}\quad (2.84)$$

Note que, como  $\mathbb{E}(Z_t) = 0$ , segue que  $\mathbb{E}(|Z_t|) < \infty$  e ainda,  $\mathbb{E}(\ln(Z_t^2)) < \infty$ . Logo,  $\gamma \mathbb{E}(|Z_t|) \mathbb{E}(\ln(Z_t^2)) = c < \infty$ . Pela expressão (2.84) temos,

$$\mathbb{E}\left(g(Z_t) \ln(Z_t^2)\right) = \theta \mathbb{E}\left(Z_t \ln(Z_t^2)\right) + \gamma \mathbb{E}\left(|Z_t| \ln(Z_t^2)\right) - c. \quad (2.85)$$

Observe que

$$\begin{aligned}\mathbb{E}\left(|Z_t| \ln(Z_t^2)\right) &= \int_{-\infty}^{\infty} |z| \ln(z^2) dF_Z(z) \\ &= 2 \int_{-\infty}^{\infty} |z| \ln(|z|) dF_Z(z) \\ &\leq 2 \int_{-\infty}^{\infty} |z| |z| dF_Z(z) \\ &= 2 \int_{-\infty}^{\infty} z^2 dF_Z(z) \\ &= 2 \mathbb{E}(Z_t^2) = 2 < \infty.\end{aligned}\quad (2.86)$$

Por hipótese,  $\mathbb{E}(Z_t \ln(Z_t^2)) < \infty$ , para todo  $t \in \mathbb{Z}$ . Então, pelas expressões (2.85), e (2.86) segue que

$$\mathbb{E}\left(g(Z_t) \ln(Z_t^2)\right) < \infty. \quad (2.87)$$

Desta forma, pelas expressões (2.87) e (2.83) segue que  $\text{Cov}(g(Z_t), \ln(Z_t^2)) < \infty$ . Portanto,  $\gamma_\varepsilon(0) < \infty$ . Assim,  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  é um processo ruído branco com variância  $\sigma_\varepsilon^2 = \gamma_\varepsilon(0)$  dada pela expressão (2.82). Isto prova que  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, 0$ ).  $\square$

**Observação 2.2.** Segue, do Teorema 2.9, que os resultados sobre estacionariedade fraca, estacionariedade estrita e ergodicidade, apresentados no Teorema 2.7, para o processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ , são também válidos para o processo  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  e a prova é análoga.

Ruiz e Veiga (2007) apresentam a medida de curtose para processos FIEGARCH( $1, d, 0$ ). Apresentamos, nas proposições que seguem, as medidas de curtose e assimetria para o caso geral dos processos FIEGARCH( $p, d, q$ ). Na Proposição 2.5 apresentamos a medida de curtose, enquanto que a Proposição 2.6 apresenta a medida de assimetria para os processos FIEGARCH( $p, d, q$ ) estacionários.

**Proposição 2.5.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ) estacionário, tal que  $\mathbb{E}(X_t^4) < \infty$  e  $\mathbb{E}(Z_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ . A medida de curtose de  $\{X_t\}_{t \in \mathbb{Z}}$  é dada por*

$$K_X = \mathbb{E}(Z_t^4) \frac{\prod_{k=0}^{\infty} \mathbb{E}\left(\exp\{2\lambda_{d,k} g(Z_t)\}\right)}{\left[\prod_{k=0}^{\infty} \mathbb{E}\left(\exp\{\lambda_{d,k} g(Z_t)\}\right)\right]^2}.$$

**Prova:** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ) estacionário, definido pelas expressões (2.44) e (2.45). Seja  $\lambda(\cdot)$  o polinômio definido pela expressão (2.52), com coeficientes  $\lambda_{d,k}$  dados pela expressão (2.53). Note que  $\mathbb{E}(X_t^2) = \mathbb{E}(\sigma_t^2 Z_t^2) = \mathbb{E}(\sigma_t^2)$ , para todo  $t \in \mathbb{Z}$ , pois  $\sigma_t$  e  $Z_t$  são independentes e  $\mathbb{E}(Z_t^2) = 1$ . De forma análoga,  $\mathbb{E}(X_t^4) = \mathbb{E}(\sigma_t^4)\mathbb{E}(Z_t^4)$ .

Pela expressão (2.45) segue que

$$\begin{aligned} \mathbb{E}(\sigma_t^2) &= \mathbb{E} \left( \exp \left\{ \omega + \sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{t-1-k}) \right\} \right) \\ &= e^{\omega} \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \{ \lambda_{d,k} g(Z_t) \} \right), \end{aligned} \quad (2.88)$$

pois as variáveis aleatórias  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  são independentes e identicamente distribuídas. De forma análoga

$$\begin{aligned} \mathbb{E}(\sigma_t^4) &= \mathbb{E}((\sigma_t^2)^2) \\ &= \mathbb{E} \left( \exp \left\{ 2\omega + 2 \sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{t-1-k}) \right\} \right) \\ &= e^{2\omega} \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \{ 2\lambda_{d,k} g(Z_t) \} \right). \end{aligned} \quad (2.89)$$

Pela Definição 1.14, a medida de curtose de  $\{X_t\}_{t \in \mathbb{Z}}$  é dada por

$$K_X = \frac{\mathbb{E}(X_t^4)}{(\mathbb{E}(X_t^2))^2} = \frac{\mathbb{E}(\sigma_t^4)\mathbb{E}(Z_t^4)}{(\mathbb{E}(\sigma_t^2))^2}.$$

Logo, pelas expressões (2.88) e (2.89) segue que

$$\begin{aligned} K_X &= \mathbb{E}(Z_t^4) \frac{e^{2\omega} \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \{ 2\lambda_{d,k} g(Z_t) \} \right)}{\left[ e^{\omega} \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \{ \lambda_{d,k} g(Z_t) \} \right) \right]^2} \\ &= \mathbb{E}(Z_t^4) \frac{\prod_{k=0}^{\infty} \mathbb{E} \left( \exp \{ 2\lambda_{d,k} g(Z_t) \} \right)}{\left[ \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \{ \lambda_{d,k} g(Z_t) \} \right) \right]^2}. \end{aligned}$$

□

**Proposição 2.6.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ) estacionário, tal que  $\mathbb{E}(X_t^3) < \infty$  e  $\mathbb{E}(Z_t^3) < \infty$ , para todo  $t \in \mathbb{Z}$ . A medida de assimetria de  $\{X_t\}_{t \in \mathbb{Z}}$  é dada por*

$$A_X = \mathbb{E}(Z_t^3) \frac{\prod_{k=0}^{\infty} \mathbb{E} \left( \exp \left\{ \frac{3}{2} \lambda_{d,k} g(Z_t) \right\} \right)}{\left[ \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \{ \lambda_{d,k} g(Z_t) \} \right) \right]^{3/2}}.$$

**Prova:** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ) estacionário, definido pelas expressões (2.44) e (2.45). Seja  $\lambda(\cdot)$  o polinômio definido pela expressão (2.52), com coeficientes  $\lambda_{d,k}$  dados pela expressão (2.53). Note que  $\mathbb{E}(X_t^2) = \mathbb{E}(\sigma_t^2 Z_t^2) = \mathbb{E}(\sigma_t^2)$ , para todo  $t \in \mathbb{Z}$ , pois  $\sigma_t$  e  $Z_t$  são independentes e  $\mathbb{E}(Z_t^2) = 1$ . De forma análoga,  $\mathbb{E}(X_t^3) = \mathbb{E}(\sigma_t^3)\mathbb{E}(Z_t^3)$ . Pela expressão (2.45) segue que

$$\begin{aligned} \mathbb{E}(\sigma_t^3) &= \mathbb{E} \left( \exp \left\{ \frac{3}{2}\omega + \frac{3}{2} \sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{t-1-k}) \right\} \right) \\ &= e^{3/2\omega} \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \left\{ \frac{3}{2} \lambda_{d,k} g(Z_t) \right\} \right). \end{aligned} \quad (2.90)$$

Pela Definição 1.14, a medida de assimetria de  $\{X_t\}_{t \in \mathbb{Z}}$  é dada por

$$A_X = \frac{\mathbb{E}(X_t^3)}{(\mathbb{E}(X_t^2))^{3/2}} = \frac{\mathbb{E}(\sigma_t^3)\mathbb{E}(Z_t^3)}{(\mathbb{E}(\sigma_t^2))^{3/2}}.$$

Logo, pelas expressões (2.88) e (2.90) segue que

$$\begin{aligned} A_X &= \mathbb{E}(Z_t^3) \frac{e^{2/3\omega} \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \left\{ \frac{3}{2} \lambda_{d,k} g(Z_t) \right\} \right)}{\left[ e^{\omega} \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \{ \lambda_{d,k} g(Z_t) \} \right) \right]^{3/2}} \\ &= \mathbb{E}(Z_t^3) \frac{\prod_{k=0}^{\infty} \mathbb{E} \left( \exp \left\{ \frac{3}{2} \lambda_{d,k} g(Z_t) \right\} \right)}{\left[ \prod_{k=0}^{\infty} \mathbb{E} \left( \exp \{ \lambda_{d,k} g(Z_t) \} \right) \right]^{3/2}}. \end{aligned}$$

□

# Capítulo 3

## Identificação, Estimação e Previsão

Neste capítulo apresentamos os principais métodos de identificação de um modelo qualquer para séries temporais apresentados na literatura. Além disso, tratamos da estimação dos parâmetros dos modelos e por último apresentamos resultados relacionados à previsão de valores futuros para o processo estocástico a partir do modelo ajustado.

### 3.1 Identificação do Modelo

Uma das fases mais difíceis do processo de análise de uma série temporal  $\{X_t\}_{t=1}^n$ , obtida à partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , é a identificação do modelo a ser ajustado. Existem vários critérios para identificar o melhor modelo. Nesta seção apresentamos os principais critérios utilizados na literatura.

#### 3.1.1 Análise da Função de Autocorrelação Amostral

A análise das funções de autocorrelação e autocorrelação amostral de uma série temporal é importante tanto na seleção do modelo quanto na análise dos resíduos. Tais funções nos dão uma idéia do grau de dependência entre as variáveis aleatórias (veja Capítulo 1).

Para identificar o modelo precisamos verificar se a função de autocorrelação  $\rho_X(\cdot)$ , dada pela expressão (1.2), é nula além de um determinado valor  $h$ . Seja  $\hat{\rho}_X(\cdot)$  a função de autocorrelação amostral, dada pela expressão (1.8). Uma expressão aproximada da variância de  $\hat{\rho}_X(\cdot)$ , para um processo estacionário Gaussiano é dada por (veja Morettin e Tolo, 2004)

$$\text{Var}(\hat{\rho}_X(h)) \sim \frac{1}{n} \sum_{k=-\infty}^{\infty} \left[ [\rho_X(k)]^2 + \rho_X(k+h)\rho_X(k-h) - 4\rho_X(h)\rho_X(k)\rho_X(k-h) + 2[\rho_X(k)]^2[\rho_X(h)]^2 \right], \quad (3.1)$$

para todo  $1 \leq h < n$ . Para um processo em que a função de autocorrelação é nula para  $h > q$ , a expressão (3.1) é dada por

$$\text{Var}(\hat{\rho}_X(h)) \sim \frac{1}{n} \left[ 1 + 2 \sum_{k=1}^q [\rho_X(k)]^2 \right], \quad h > q. \quad (3.2)$$

Substituindo  $\rho_X(\cdot)$  por  $\hat{\rho}_X(\cdot)$  na expressão (3.2), obtemos a estimativa para a variância, dada por

$$\widehat{\text{Var}}(\hat{\rho}_X(h)) \sim \frac{1}{n} \left[ 1 + 2 \sum_{k=1}^q [\hat{\rho}_X(k)]^2 \right], \quad h > q. \quad (3.3)$$

Para  $n$  suficientemente grande e sob a hipótese de que  $\rho_X(h) = 0$ , para  $h > q$ , a distribuição de  $\hat{\rho}_X(h)$  é aproximadamente normal, com média zero e variância dada pela expressão (3.3). Desta forma, um intervalo de confiança aproximado para a função de autocorrelação, é dado por

$$\hat{\rho}_X(h) \pm t_{n-1,\gamma} \widehat{\text{Var}}(\hat{\rho}_X(h)),$$

onde  $t_{n-1,\gamma}$  é o valor da estatística  $t$  de Student com  $n - 1$  graus de liberdade, tal que  $\mathbb{P}(-t_{n-1,\gamma} < t < t_{n-1,\gamma}) = 1 - \gamma$ . Na prática, usa-se  $t_{n-1,\gamma} = 2$ , de modo que podemos considerar  $\rho_X(h)$  significativamente diferente de zero se

$$|\hat{\rho}_X(h)| > 2 \widehat{\text{Var}}(\hat{\rho}_X(h)), \quad h > q.$$

### 3.1.2 Teste de Normalidade Jarque-Bera

Seja  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir de um processo estocástico estacionário  $\{X_t\}_{t \in \mathbb{Z}}$ . Se as variáveis aleatórias são normalmente distribuídas então, para  $n$  grande,

$$\hat{A}_X \sim \mathcal{N}\left(0, \frac{6}{n}\right) \quad \text{e} \quad \hat{K}_X \sim \mathcal{N}\left(3, \frac{24}{n}\right),$$

onde  $\hat{A}_X$  e  $\hat{K}_X$  são as medidas de assimetria e de curtose amostrais de  $\{X_t\}_{t=1}^n$ , dadas pela expressão (1.9). Estes fatos são utilizados para testar a normalidade de uma série temporal.

Podemos utilizar a estatística de teste  $\sqrt{n/6}\hat{A}_X$  para testar a hipótese  $H_0 : A_X = 0$ , onde  $A_X$  é a medida de assimetria, e utilizar a estatística de teste  $\sqrt{n/24}(\hat{K}_X - 3)$  para testar a hipótese  $H_0 : K_X = 3$ , onde  $K_X$  é a medida de curtose. Sob  $H_0$ , ambas as estatísticas terão distribuição normal padrão.

Um teste amplamente utilizado em Econometria é o teste Jarque-Bera (Bera e Jarque, 1981), que combina os dois testes acima, utilizando a estatística de teste

$$JB = \left(\frac{n}{6}\right) \hat{A}_X^2 + \left(\frac{n}{24}\right) (\hat{K}_X - 3)^2,$$

que, sob  $H_0$  : a série temporal é Gaussiana, tem distribuição qui-quadrado com dois graus de liberdade.

### 3.1.3 Critérios AIC, BIC e HQC

Além da análise da função de autocorrelação, existem os métodos baseados em uma função penalizadora. A idéia é escolher o modelo cujo número de parâmetros minimize a quantidade

$$P(T) = \ln(\hat{\sigma}^2) + \frac{TC(n)}{n},$$

onde  $\hat{\sigma}^2$  é uma estimativa da variância residual obtida ao se ajustar o modelo às  $n$  observações da série temporal,  $T$  é o número de parâmetros a serem estimados e  $C(n)$  é uma função do tamanho da série temporal.

Os critérios que apresentamos a seguir são os mais utilizados na literatura. Estes critérios utilizam o estimador de máxima verossimilhança para a variância, mais uma função penalizadora que difere para cada critério. Nas definições que seguem,  $T$  é o número de parâmetros estimados e  $\hat{\sigma}^2$  é o estimador de máxima verossimilhança da variância.

#### A) Critério de Informação de Akaike (AIC)

$$\text{AIC} = \ln(\hat{\sigma}^2) + \frac{2T}{n}.$$

Hurvich e Tsay (1989) propõem uma correção para o AIC, dada por

$$\text{AIC}_c = \text{AIC} + \frac{2(T+1)(T+2)}{n-T+2}.$$

#### B) Critério de Informação Bayesiano (BIC ou SIC)

$$\text{BIC} = \ln(\hat{\sigma}^2) + T \frac{\ln(n)}{n}.$$

#### C) Critério de Hannan e Quinn (HQC)

$$\text{HQC} = \ln(\hat{\sigma}^2) + 2Tc \frac{\ln(\ln(n))}{n},$$

onde  $c > 1$ .

### 3.1.4 Seleção do Modelo Baseado nos Erros de Previsão

Muitas vezes existe mais de um modelo adequado para descrever uma série temporal. Se o objetivo principal for utilizar o modelo para previsões futuras, um critério alternativo para seleção do modelo pode ser baseado nos erros de previsão.

**Definição 3.1. (Erro de Previsão).** Sejam  $X_t$  e  $\hat{X}_t$ , respectivamente, o valor observado da série temporal e o valor previsto pelo modelo no tempo  $t$ . O *erro de previsão*  $e_t$  é dado por

$$e_t = X_t - \hat{X}_t \quad \text{para todo } t \in \mathbb{Z}. \quad (3.4)$$

Nas definições que seguem,  $m$  é o número de valores previstos e  $e_t$  é o erro de previsão no tempo  $t$ , definido pela expressão (3.4).

**Definição 3.2. (Erro Percentual Médio).** O *erro percentual médio*, denotado por MPE, mede o vício da previsão e é dado por

$$\text{MPE} = \frac{1}{m} \sum_{t=1}^m \frac{e_t}{\hat{X}_t},$$



**Definição 3.3. (Erro Quadrático Médio).** O *erro quadrático médio*, denotado por MSE, é dado por

$$\text{MSE} = \frac{1}{m} \sum_{t=1}^m e_t^2. \quad (3.5)$$

**Definição 3.4. (Erro Absoluto Médio).** O *erro absoluto médio*, denotado por MAE, é dado por

$$\text{MAE} = \frac{1}{m} \sum_{t=1}^m |e_t|.$$

**Definição 3.5. (Erro Percentual Absoluto Médio).** O *erro percentual absoluto médio*, denotado por MAPE, é dado por

$$\text{MAPE} = \frac{1}{m} \sum_{t=1}^m \left| \frac{e_t}{X_t} \right|.$$

### 3.1.5 Testes de Adequação do Modelo

Os testes baseiam-se, em geral, nas autocorrelações estimadas dos resíduos. Dentre os principais testes, podemos citar

#### Teste de Autocorrelação Residual.

Suponha que um determinado modelo foi ajustado aos dados. Sejam  $\{\hat{\varepsilon}_t\}_{t=1}^n$  os resíduos estimados. Se o modelo for adequado, as variáveis aleatórias  $\{\hat{\varepsilon}_t\}_{t=1}^n$  deverão ser aproximadamente não-correlacionadas, isto é, se  $\hat{\rho}_\varepsilon(\cdot)$  é a função de autocorrelação amostral dos resíduos  $\hat{\varepsilon}_t$ , devemos ter  $\hat{\rho}_\varepsilon(h) \simeq 0$ , para todo  $h > 1$ .

#### Teste de Box-Pierce-Ljung

Box e Pierce (1970) sugerem um teste para as autocorrelações dos resíduos estimados. Sejam  $\{\hat{\varepsilon}_t\}_{t=1}^n$  os resíduos estimados e  $\hat{\rho}_\varepsilon(\cdot)$  a sua função de autocorrelação amostral. Se o modelo for adequado, a estatística

$$Q(m) = n \sum_{j=1}^m [\hat{\rho}_\varepsilon(j)]^2$$

terá aproximadamente distribuição  $\mathcal{X}^2$  com  $m - T$  graus de liberdade, onde  $T$  é o número de parâmetros estimados.

Ljung e Box (1978) sugerem uma modificação para este teste. Eles sugerem que, se o modelo for adequado, a estatística

$$Q(m) = n(n+2) \sum_{j=1}^m \frac{[\hat{\rho}_\varepsilon(j)]^2}{n-j},$$

terá aproximadamente uma distribuição  $\mathcal{X}^2$  com  $m - T$  graus de liberdade, onde  $T$  é o número de parâmetros estimados. A hipótese é rejeitada para valores grandes de  $Q(m)$ .

## Teste dos Multiplicadores de Lagrange (ML)

Este teste, sugerido por Engle (1982), é utilizado para verificar se a série temporal apresenta heteroscedasticidade e consiste em testar a hipótese nula  $H_0 : \alpha_i = 0$ , para todo  $i = 1, \dots, r$ , na regressão

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2 + u_t, \quad (3.6)$$

para  $t = r + 1, \dots, n$ . A estatística do teste é

$$ML = nR^2 \sim \mathcal{X}^2(r),$$

onde  $R^2$  é o quadrado do coeficiente de correlação múltipla da regressão dada na expressão (3.6).

### 3.1.6 Estatísticas de Teste para Detectar Longa Dependência

Apresentamos, a seguir, alguns dos procedimentos utilizados para testar se uma série temporal apresenta característica de longa dependência. Utilizamos as estatísticas apresentadas para testar a hipótese  $H_0 : d = 0$ , onde  $d$  é o parâmetro de diferenciação fracionária que caracteriza a longa dependência.

No que segue, adotamos a seguinte notação:

- $\{W(t)\}_{t \in [0,1]}$  é o Movimento Browniano padrão.
- $\{W^0(t)\}_{t \in [0,1]}$ , onde  $W^0(t) = W(t) - tW(1)$ , é uma Ponte Browniana.
- o símbolo  $\implies$  denota convergência fraca das variáveis aleatórias.

#### Estatística R/S

A estatística R/S foi introduzida por Hurst (1951) com o nome *rescaled range*, com o propósito de testar a existência de longa dependência em séries temporais. Dada uma série temporal  $\{X_t\}_{t=1}^n$ , a estatística R/S é definida como,

$$Q_n = \frac{1}{S_n} \left[ \max_{1 \leq k \leq n} \sum_{j=1}^k (X_j - \bar{X}) - \min_{1 \leq k \leq n} \sum_{j=1}^k (X_j - \bar{X}) \right], \quad (3.7)$$

onde  $\bar{X}$  é a média amostral, definida pela expressão (1.5) e  $S_n = \sqrt{\hat{\sigma}_X^2}$  é o desvio padrão amostral, com  $\hat{\sigma}_X^2$  definida pela expressão (1.7).

Se as variáveis aleatórias  $X_t$ , para todo  $t \in \mathbb{Z}$ , são independentes e identicamente distribuídas, com distribuição comum  $\mathcal{N}(\mu_X, \sigma_X^2)$ , então  $Q_n/\sqrt{n}$  converge fracamente para uma variável aleatória que está no domínio de atração de uma Ponte Browniana no intervalo  $[0, 1]$ . Isto é,

$$\frac{Q_n}{\sqrt{n}} \implies U_{R/S}$$

onde

$$U_{R/S} = \max_{0 \leq t \leq 1} W^0(t) - \min_{0 \leq t \leq 1} W^0(t). \quad (3.8)$$

## Estatística R/S Modificada

Lo (1991) propõe uma modificação para a estatística R/S, definida pela expressão (3.7). O autor sugere a modificação

$$\tilde{Q}_n = \frac{1}{\hat{\sigma}_n(q)} \left[ \max_{1 \leq k \leq n} \sum_{j=1}^k (X_j - \bar{X}) - \min_{1 \leq k \leq n} \sum_{j=1}^k (X_j - \bar{X}) \right], \quad (3.9)$$

onde

$$\hat{\sigma}_n^2(q) = \hat{\gamma}_X(0) + 2 \sum_{j=1}^q \omega_j(q) \hat{\gamma}_X(j), \quad \omega_j(q) = 1 - \frac{j}{q+1}, \quad q < n, \quad (3.10)$$

e  $\hat{\gamma}_X(\cdot)$  é definida pela expressão (1.6). Os pesos  $\omega_j(q)$  são os mesmos utilizados por Newey e West (1987), que sugerem escolher  $q = 4 \left( \frac{n}{100} \right)^{2/9}$ . Zivot e Wang (2005) utilizam  $q = 4 \left( \frac{n}{100} \right)^{1/4}$ . Neste trabalho, a notação para esta estatística será  $R/S_m$ .

Lo (1991) mostra que, na ausência de longa dependência, a variável aleatória  $\tilde{Q}_n/\sqrt{n}$  também converge fracamente para a variável aleatória  $U_{R/S}$ , dada pela expressão (3.8).

## Estatística KPSS

Introduzido por Kwiatkowski *et al.* (1992), o KPSS é um teste que, assim como o R/S, baseia-se em somas parciais. Ele foi inicialmente utilizado para testar a hipótese nula de estacionariedade contra a hipótese de existência de uma raiz unitária. Lee e Schmidt (1996) utilizam a estatística KPSS para testar a presença de longa dependência em séries temporais estacionárias.

Dada uma série temporal  $\{X_t\}_{t=1}^n$ , a estatística KPSS é dada por

$$T_n = \frac{1}{\hat{\sigma}_n^2(q)n^2} \sum_{k=1}^n \left( \sum_{j=1}^k (X_j - \bar{X}) \right)^2,$$

onde  $\hat{\sigma}_n^2(q)$  é definida pela expressão (3.10).

Sob as hipóteses de estacionariedade e ausência de longa dependência temos,

$$T_n \implies U_{KPSS},$$

onde

$$U_{KPSS} = \int_0^1 (W^0(t))^2 dt.$$

## Estatística V/S

Giraitis *et al.* (2003) propõem uma modificação para a estatística KPSS. A estatística modificada foi denominada *rescaled variance*, ou simplesmente, V/S. Dada uma série temporal  $\{X_t\}_{t=1}^n$ , a estatística V/S é dada por

$$M_n = \frac{1}{\hat{\sigma}_n^2(q)n^2} \left[ \sum_{k=1}^n \left( \sum_{j=1}^k (X_j - \bar{X}) \right)^2 - \frac{1}{n} \left( \sum_{k=1}^n \sum_{j=1}^k (X_j - \bar{X}) \right)^2 \right],$$

onde  $\hat{\sigma}_n^2(q)$  é definida pela expressão (3.10).

Sob as hipóteses de estacionariedade e ausência de longa dependência temos,

$$M_n \implies U_{V/S},$$

onde

$$U_{V/S} = \int_0^1 (W^0(t))^2 dt - \left( \int_0^1 W^0(t) dt \right)^2.$$

Para o cálculo das estatísticas R/S, R/S<sub>m</sub>, KPSS ou V/S, subdivide-se as  $n$  observações da série temporal  $\{X_t\}_{t=1}^n$  em blocos adjacentes e que não se sobrepõem, cada um de tamanho  $[n/B]$ . Obtém-se assim a seqüência  $\{t_i\}_{i=1}^B$ , onde  $t_i = (i-1)[n/B] + 1$ , para todo  $i = 1, \dots, B$ . Para cada ponto da seqüência  $\{t_i\}_{i=1}^B$  define-se a seqüência crescente  $\{[t_i, t_i + l_j]\}_{j=1}^K$ , de  $K$  blocos com origem em  $t_i$ , tal que  $t_i + l_j < n$ . A seqüência  $\{l_j\}_{j=1}^K$  representa o tamanho amostral utilizado para o cálculo da estatística. Giraitis *et al.* (2003) sugere utilizar  $l = 40$  como tamanho mínimo amostral para cada bloco. Nas simulações do Capítulo 5 utilizamos  $l \in \{40, 100\}$ .

O procedimento para identificação do parâmetro  $d$  consiste em construir gráficos (na escala log-log) de R/S, R/S<sub>m</sub>, KPSS ou V/S, contra o tamanho amostral. Para uma série temporal com característica de longa dependência a reta deve ter inclinação  $H > 0,5$ , no caso das estatísticas R/S, R/S<sub>m</sub> e inclinação  $b > 1$ , no caso das estatísticas KPSS e V/S, para grandes amostras. O coeficiente  $H$  é denominado *coeficiente de Hurst* e satisfaz  $H = d + 0,5$ ; o coeficiente  $b$  satisfaz  $b = 2H - 1 = 2d$ .

## Teste GPH

Este método para estimação do parâmetro de longa dependência foi proposto por Geweke e Porter-Hudak (1983). O teste consiste em estimar o parâmetro  $d$  através do método de regressão linear baseado na função periodograma e tester a hipótese  $H_0 : d = 0$  contra  $H_1 : d \neq 0$ .

Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo ARFIMA( $p, d, q$ ), conforme Definição 1.16, com  $\mathbb{E}(X_t) = 0$ , para todo  $t \in \mathbb{Z}$ . Sabemos que a função densidade espectral desse processo (veja item (iv) do Teorema 1.5) é dada por

$$f_X(\lambda) = f_U(\lambda) \left[ 2 \operatorname{sen} \left( \frac{\lambda}{2} \right) \right]^{-2d}, \quad \text{para } \lambda \in (0, \pi], \quad (3.11)$$

onde

$$f_U(\lambda) = \frac{\sigma_Z^2 |\theta(e^{-i\lambda})|^2}{2\pi |\phi(e^{-i\lambda})|^2}, \quad \text{para } \lambda \in (0, \pi],$$

é a função densidade espectral do processo ARMA( $p, q$ ) dado por  $\phi(\mathcal{B})U_t = \theta(\mathcal{B})Z_t$ , para todo  $t \in \mathbb{Z}$ .

Aplicando o logaritmo e somando  $\ln(f_U(0))$  em ambos os lados da equação (3.11), obtemos a expressão

$$\ln(f_X(\lambda)) = \ln(f_U(0)) - d \ln \left( 2 \operatorname{sen} \left( \frac{\lambda}{2} \right) \right)^2 + \ln \left( \frac{f_U(\lambda)}{f_U(0)} \right), \quad \text{para } \lambda \in (0, \pi]. \quad (3.12)$$

Substituindo, na equação (3.12), a frequência  $\lambda$  pelas frequências de Fourier  $\lambda_\ell = \frac{2\pi\ell}{n}$ , para  $\ell = 1, \dots, g(n)$  e somando  $\ln(I_n(\lambda_\ell))$  em ambos os lados da equação (3.12), onde  $I_n(\cdot)$  é a função periodograma dada pela expressão (1.10), obtemos

$$\ln(I_n(\lambda_\ell)) = \ln(f_U(0)) - d \ln\left(2 \operatorname{sen}\left(\frac{\lambda_\ell}{2}\right)\right)^2 + \ln\left(\frac{I_n(\lambda_\ell)}{f_X(\lambda_\ell)}\right) + \ln\left(\frac{f_U(\lambda_\ell)}{f_U(0)}\right), \quad (3.13)$$

para  $\ell = 1, \dots, g(n)$ . Se  $\lambda_\ell$  está próximo de zero o último termo é desprezível se comparado com os outros termos do lado direito da igualdade na expressão (3.13) (ver Geweke e Porter-Hudak, 1983). Desta forma, podemos escrever a expressão (3.13) como uma equação de regressão linear simples

$$y_\ell = a + b x_\ell + \epsilon_\ell, \quad \text{para } \ell = 1, \dots, g(n),$$

onde

$$y_\ell = \ln(I_n(\lambda_\ell)), \quad x_\ell = \ln\left(2 \operatorname{sen}\left(\frac{\lambda_\ell}{2}\right)\right)^2, \quad (3.14)$$

$$a = \ln(f_U(0)), \quad b = -d, \quad \text{e} \quad \epsilon_\ell = \ln\left(\frac{I_n(\lambda_\ell)}{f_X(\lambda_\ell)}\right). \quad (3.15)$$

Geralmente, utiliza-se  $g(n) = n^\alpha$ , para  $\alpha \in (0, 1)$ . O estimador  $\hat{d}_{GPH}$  de  $d$  é obtido pelo método dos mínimos quadrados e é dado por

$$\hat{d}_{GPH} = -\frac{\sum_{\ell=1}^{g(n)} (x_\ell - \bar{x})(y_\ell - \bar{y})}{\sum_{\ell=1}^{g(n)} (x_\ell - \bar{x})^2},$$

onde  $y_\ell$  e  $x_\ell$  são dados pela expressão (3.14) e

$$\bar{x} = \frac{1}{g(n)} \sum_{\ell=1}^{g(n)} x_\ell \quad \text{e} \quad \bar{y} = \frac{1}{g(n)} \sum_{\ell=1}^{g(n)} y_\ell. \quad (3.16)$$

Geweke e Porter-Hudak (1983) mostram que o estimador  $\hat{d}_{GPH}$  é normalmente distribuído para grandes amostras, isto é, se  $g(n) = n^\alpha$ , para  $\alpha \in (0, 1)$ ,

$$\hat{d}_{GPH} \sim \mathcal{N}\left(d, \frac{\pi^2}{6 \sum_{\ell=1}^{g(n)} (x_\ell - \bar{x})^2}\right),$$

com  $x_\ell$  e  $\bar{x}$  dados pelas expressões (3.14) e (3.16). Sob a hipótese nula de não existência de longa dependência, isto é, sob a hipótese de que  $d = 0$ , a estatística de teste

$$T = \hat{d}_{GPH} \times \left(\frac{\pi^2}{6 \sum_{\ell=1}^{g(n)} (x_\ell - \bar{x})^2}\right)^{-1/2},$$

tem distribuição assintótica normal.

## 3.2 Métodos de Estimação

Nesta seção relembramos o método de estimação da máxima verossimilhança e definimos os estimadores de pseudo-máxima verossimilhança e quase-máxima verossimilhança que são os métodos de estimação mais utilizados para estimar os parâmetros de processos estocásticos não-lineares. Dado um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , denotamos por  $\mathcal{F}_t$  a  $\sigma$ -álgebra gerada por  $\{X_s\}_{s \leq t}$ , isto é,  $\mathcal{F}_t \equiv \sigma(X_s; s \leq t)$ .

### 3.2.1 Método de Estimação da Máxima Verossimilhança

**Definição 3.6.** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico. Para todo  $t \in \mathbb{Z}$ , considere  $f_{X_t}(\cdot; \theta)$  a função densidade de probabilidade da variável aleatória  $X_t$ , onde  $\theta$  é um parâmetro desconhecido que pertence ao espaço de parâmetros  $\Theta \subseteq \mathbb{R}$ .

- (i) Para cada variável aleatória  $X_t$  e para cada  $x$  fixo, a *função de verossimilhança* é uma função de  $\theta$  definida por

$$L(\theta; x) \equiv f_{X_t}(x; \theta).$$

- (ii) O *estimador de máxima verossimilhança*  $\hat{\theta}$ , de  $\theta$ , é definido como sendo

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \Theta} \{L(\theta; x)\}.$$

- (iii) Seja  $g(\cdot)$  uma função mensurável de  $\theta$ . Se  $\hat{\theta}$  é o estimador de máxima verossimilhança para  $\theta$ , então,  $\hat{v} = g(\hat{\theta})$  é o estimador de máxima verossimilhança para  $v = g(\theta)$ .

Intuitivamente, estamos escolhendo  $\theta$  que maximiza a probabilidade de ocorrência de  $x$ . Em alguns casos existe mais de um valor que maximiza a função. Em outras situações,  $L(\theta; x)$  pode ser ilimitada. Neste caso, evidentemente, não existe o estimador de máxima verossimilhança para  $\theta$ .

Note que a função  $L(\theta; x)$  é sempre não-negativa, além disso, maximizá-la é equivalente a maximizar a função  $\ln(L(\theta; x))$ . Sendo assim, definimos a *função de log-verossimilhança* por

$$\mathcal{L}(\theta; x) \equiv \ln(L(\theta; x)).$$

Supondo que  $\Theta$  é um conjunto aberto em  $\mathbb{R}$  e que a função de verossimilhança é diferenciável em relação a  $\theta$ , se  $\hat{\theta}$  existir, deve satisfazer a *equação de verossimilhança* dada por

$$\frac{\partial \mathcal{L}(\theta; x)}{\partial \theta} = 0.$$

**Observação 3.1.** De forma equivalente, dada uma série temporal  $\{X_t\}_{t=1}^n$ , obtida a partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , com função densidade de probabilidade conjunta dada por

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n; \boldsymbol{\theta}) = \prod_{i=1}^n f_{X_i}(x_i, \boldsymbol{\theta}), \quad (3.17)$$

onde  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_s)' \in \Theta \subseteq \mathbb{R}^s$ , a função de verossimilhança de  $\boldsymbol{\theta}$  é dada por

$$L(\boldsymbol{\theta}; \mathbf{x}) \equiv f_{X_1, \dots, X_n}(\mathbf{x}; \boldsymbol{\theta}),$$

com  $\mathbf{x} = (x_1, \dots, x_n)'$ . O estimador de máxima verossimilhança de  $\boldsymbol{\theta}$  é dado por

$$\hat{\boldsymbol{\theta}} = \operatorname{argmax}_{\boldsymbol{\theta} \in \Theta} \{L(\boldsymbol{\theta}; \mathbf{x})\}. \quad (3.18)$$

Neste caso, pela equação (3.17), a função log-verossimilhança é dada por

$$\mathcal{L}(\boldsymbol{\theta}; \mathbf{x}) = \sum_{i=1}^n \ln(f_{X_i}(x_i; \boldsymbol{\theta})). \quad (3.19)$$

Pela expressão (3.19) a equação de verossimilhança torna-se

$$\sum_{i=1}^n \frac{\partial \ln(f_{X_i}(x_i; \boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} = 0.$$

Suponha que  $\Theta \subseteq \mathbb{R}^s$  é um conjunto aberto e que o estimador de máxima verossimilhança para  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_s)' \in \Theta \subseteq \mathbb{R}^s$ , dado pela expressão (3.18), existe. Então, deve satisfazer as equações de verossimilhança

$$\frac{\partial \mathcal{L}(\boldsymbol{\theta}; \mathbf{x})}{\partial \theta_i} = 0,$$

para todo  $i = 1, \dots, s$  e para todo  $\mathbf{x}$  tal que  $\mathcal{L}(\boldsymbol{\theta}; \mathbf{x})$  tenha derivadas de primeira ordem em relação à  $\boldsymbol{\theta}$ .

### 3.2.2 Método de Estimação da Pseudo-Máxima Verossimilhança

Seja  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  tal que, para todo  $t \in \mathbb{Z}$ , a variável aleatória  $X_t$  possui função densidade de probabilidade  $g_{X_t}(\cdot; \boldsymbol{\theta})$ . Para obter o estimador de máxima-verossimilhança para  $\boldsymbol{\theta}$ , definido na expressão (3.18), devemos, necessariamente, conhecer a função densidade de probabilidade das variáveis aleatórias  $X_t$ , para todo  $1 \leq t \leq n$ . Em geral, na prática, isso não ocorre. Nessas situações, utilizamos o método de estimação de *pseudo-máxima verossimilhança*, introduzido por Besag (1975). Para maiores detalhes veja Gouriéroux (1997), e referências ali contidas.

Nas definições que seguem, considere  $\mathcal{F}_{t-1} \equiv \sigma(X_s; s \leq t-1)$  e  $f_{X_t|\mathcal{F}_{t-1}}(\cdot; \boldsymbol{\theta})$  não é necessariamente a verdadeira função densidade de probabilidade condicional da variável aleatória  $X_t$ , dado que essa nem sempre é conhecida, mas sim uma função que supomos ser a função densidade de probabilidade condicional da variável aleatória. As funções de distribuição *Gaussiana* e *t-Student* são as mais utilizadas para obter a *função de pseudo-verossimilhança*, que definimos a seguir.

**Definição 3.7.** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico qualquer. Para cada  $t \in \mathbb{Z}$ , assumamos que a variável aleatória  $X_t$  possui função densidade de probabilidade condicional denotada por  $f_{X_t|\mathcal{F}_{t-1}}(\cdot; \boldsymbol{\theta})$ , onde  $\boldsymbol{\theta}$  pertence ao espaço de parâmetros  $\Theta \subseteq \mathbb{R}^s$ .

- (i) Para todo  $t \in \mathbb{Z}$ , dada a variável aleatória  $X_t$ , para cada  $x$  fixo, definimos a função de verossimilhança condicional de  $\boldsymbol{\theta}$ , denominada *função de pseudo-verossimilhança* de  $\boldsymbol{\theta}$ , por

$$\ell_t(\boldsymbol{\theta}; x_t) = f_{X_t|\mathcal{F}_{t-1}}(x_t|x_{t-1}, x_{t-2}, \dots; \boldsymbol{\theta}). \quad (3.20)$$

- (ii) Seja  $\{X_t\}_{t=1}^n$  uma s rie temporal obtida a partir do processo estoc stico  $\{X_t\}_{t \in \mathbb{Z}}$ . A fun o de verossimilhan a condicionada em  $\mathcal{F}_u = \sigma(X_u, \dots, X_1)$ , com  $1 \leq u \leq n$ , tamb m denominada *fun o de pseudo-verossimilhan a* de  $\boldsymbol{\theta}$ ,   definida por

$$\ell(\boldsymbol{\theta}; \mathbf{x}) = \ell(\boldsymbol{\theta}; x_{u+1}, \dots, x_n | x_u, \dots) = \prod_{i=u+1}^n \ell_i(\boldsymbol{\theta}; x_i), \quad (3.21)$$

onde  $\mathbf{x} = (x_{u+1}, \dots, x_n)'$  e  $\ell_i(\boldsymbol{\theta}; x_i)$ , para todo  $u + 1 \leq i \leq n$ ,   dada pela express o (3.20).

- (iii) O *estimador de pseudo-m xima verossimilhan a* de  $\boldsymbol{\theta}$    dado por

$$\hat{\boldsymbol{\theta}}_{PL} = \operatorname{argmax}_{\boldsymbol{\theta} \in \Theta} \{ \ln(\ell(\boldsymbol{\theta}; \mathbf{x})) \}.$$

### Observa o 3.2.

1. Na maioria dos casos, para construir a fun o de pseudo-verossimilhan a, dada na express o (3.21), consideramos  $u = 0$  (veja, por exemplo, Morettin e Toloi, 2004). Neste caso, como n o conhecemos os valores de  $X_t$ , para  $t < 1$ , consideramos  $\mathcal{F}_0 = I_0$ , onde  $I_0$    um conjunto de condi es iniciais a serem definidas adequadamente.
2. A fun o de pseudo-verossimilhan a constru da sob a hip tese de normalidade da vari vel aleat ria  $X_t | \mathcal{F}_{t-1}$    tamb m conhecida como *fun o de quase-verossimilhan a* e o estimador obtido a partir de tal fun o   denominado *estimador de quase-m xima verossimilhan a*.

A partir das express es (3.20) e (3.21), a *fun o de log-pseudo-verossimilhan a*   dada por

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}; \mathbf{x}) &= \ln(\ell(\boldsymbol{\theta}; \mathbf{x})) \\ &= \sum_{t=r+1}^n \ln(\ell_t(\boldsymbol{\theta}; x_t)). \end{aligned}$$

As propriedades do estimador de pseudo-m xima verossimilhan a dependem tanto da distribui o verdadeira como da distribui o utilizada para construir o estimador.

As propriedades assint ticas do estimador de quase-m xima verossimilhan a para modelos ARCH e GARCH foram analisadas por diversos autores, dentre eles, Lee e Hansen (1994), Lumsdaine (1996), Berkes *et al.* (2003), Berkes e Horv th (2003) e Hall e Yao (2003). Sabe-se que, para tais processos (veja McNeil *et al.*, 2005, Degiannakis, 2004, Gouri roux, 1997 e Hamilton, 1994), sob certas condi es de regularidade, o estimador   consistente. Al m disso, este estimador   assintoticamente normal com matriz de vari ncia-covari ncia dada por

$$\operatorname{Var}_{\boldsymbol{\theta}} \left( \sqrt{n}(\hat{\boldsymbol{\theta}}_{PL} - \boldsymbol{\theta}) \right) = J^{-1} I J^{-1},$$

onde as matrizes quadradas  $I$  e  $J$ , de ordem  $s \times s$ , s o dadas por

$$I = \mathbb{E}_0 \left[ \frac{\partial \ln(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \boldsymbol{\theta}} \times \frac{\partial \ln(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \boldsymbol{\theta}'} \right] \quad \text{e} \quad J = \mathbb{E}_0 \left[ -\frac{\partial^2 \ln(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right], \quad (3.22)$$



onde  $\mathbb{E}_0(\cdot)$  indica que a esperança é considerada em relação a função distribuição verdadeira.

Os estimadores consistentes para as matrizes  $I$  e  $J$ , dadas na expressão (3.22), são dados, respectivamente, por

$$\hat{I} = \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial \ln(\ell_t(\hat{\boldsymbol{\theta}}_{PL}; x_t))}{\partial \boldsymbol{\theta}} \cdot \frac{\partial \ln(\ell_t(\hat{\boldsymbol{\theta}}_{PL}; x_t))}{\partial \boldsymbol{\theta}'} \right] \quad (3.23)$$

e

$$\hat{J} = \frac{1}{n} \sum_{t=1}^n \left[ -\frac{\partial^2 \ln(\ell_t(\hat{\boldsymbol{\theta}}_{PL}; x_t))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right]. \quad (3.24)$$

Assim, a variância assintótica é então estimada por

$$\widehat{\text{Var}}_{\boldsymbol{\theta}} \left( \sqrt{n}(\hat{\boldsymbol{\theta}}_{PL} - \boldsymbol{\theta}) \right) = \hat{J}^{-1} \hat{I} \hat{J}^{-1}.$$

Apresentamos, no Apêndice B, dois exemplos onde obtemos o estimador de quase-máxima verossimilhança e analisamos a sua precisão assintótica. No Exemplo B.1 as variáveis aleatórias são independentes e identicamente distribuídas. No Exemplo B.2 as variáveis aleatórias são tais que a média e a variância condicional variam com o tempo. Para finalizar, aplicamos os resultados do Exemplo B.2 para um processo ARCH( $p$ ) (ver Exemplo B.3).

Resultados para o estimador de quase-máxima verossimilhança para modelos EGARCH podem ser encontrados em Straumann e Mikosch (2003). O artigo apresenta uma discussão sobre o estimador de quase-máxima verossimilhança considerando os modelos heteroscedásticos, em que  $X_t = \sigma_t Z_t$ , onde a volatilidade  $\sigma_t$  é uma função não-observável de  $(X_{t-1}, \dots, X_{t-p}, \sigma_{t-1}, \dots, \sigma_{t-q})$  e  $Z_t$  tem função de distribuição normal.

No caso de modelos FIEGARCH, Palma (2007), página 121, quando refere-se ao estimador de quase-máxima verossimilhança, diz que, apesar de ser amplamente utilizado na prática, resultados assintóticos para este estimador ainda permanecem um problema em aberto.

## 3.3 Estimação dos Parâmetros dos Processos Não-Lineares

### 3.3.1 Processos ARCH, GARCH, IGARCH e EGARCH

Para os processos ARCH, GARCH, IGARCH e EGARCH, os estimadores dos seus parâmetros são obtidos pelo método da pseudo-verossimilhança pois, na prática, não conhecemos a verdadeira função de distribuição das variáveis aleatórias. As funções de distribuição mais utilizadas para construir o estimador de pseudo-verossimilhança são a Gaussiana e a  $t$ -Student.

Considere  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ . Seja  $\boldsymbol{\theta}$  o vetor de parâmetros a ser estimado. Sob a hipótese de normalidade das variáveis aleatórias  $X_t | \mathcal{F}_{t-1}$ , para todo  $t \in \mathbb{Z}$ , e observando que  $Z_t = \frac{X_t}{\sigma_t}$ , para todo

$t \in \mathbb{Z}$ , de acordo com a Definição 3.7 o estimador  $\hat{\boldsymbol{\theta}}_{PL}$  de  $\boldsymbol{\theta}$  é obtido maximizando-se a expressão (ver Apêndice B)

$$\ln(\ell(\boldsymbol{\theta}; \mathbf{x})) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^n \left[ \ln(\sigma_t^2) + \frac{X_t^2}{\sigma_t^2} \right]. \quad (3.25)$$

De forma análoga, se assumirmos que as variáveis aleatórias  $X_t | \mathcal{F}_{t-1}$ , para todo  $t \in \mathbb{Z}$ , possuem distribuição  $t$ -Student com  $\nu$  graus de liberdade, a função de pseudo verossimilhança da variável aleatória  $X_t | \mathcal{F}_{t-1}$  normalizada, isto é, da variável aleatória  $\frac{X_t}{\sigma_t \sqrt{\nu/(\nu-2)}}$ , será dada por

$$\ell_t(\boldsymbol{\theta}; x_t) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2) \sqrt{(\nu-2)\pi}} \left( 1 + \frac{X_t^2}{\sigma_t^2(\nu-2)} \right)^{-\frac{\nu+1}{2}}, \quad \text{para todo } t \in \mathbb{Z}. \quad (3.26)$$

Pela expressão (3.26) e pela Definição 3.7 o estimador  $\hat{\boldsymbol{\theta}}_{PL}$  de  $\boldsymbol{\theta}$  é obtido maximizando-se a expressão

$$\ln(\ell(\boldsymbol{\theta}; \mathbf{x})) = - \sum_{t=1}^n \left[ \frac{\nu+1}{2} \ln \left( 1 + \frac{X_t^2}{\sigma_t^2(\nu-2)} \right) \right] + n \ln \left( \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2) \sqrt{(\nu-2)\pi}} \right). \quad (3.27)$$

A maximização das expressões (3.25) e (3.27) pode ser realizada através da utilização de algoritmos de otimização numérica, tais como Newton-Raphson, Gauss-Newton, etc. Para mais detalhes veja Hamilton (1994) e Press *et al.* (1989).

Note ainda que, dado que não conhecemos os valores do processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  quando  $t < 1$ , a função de pseudo-verossimilhança é obtida condicionando-se  $\{X_t\}_{t=1}^n$  sobre  $I_0$ , onde  $I_0$  é um conjunto de condições iniciais. Bollerslev (1986) sugere utilizar  $I_0 = \{\sigma_t^2 = X_t^2 = \hat{\sigma}_X^2, \text{ para } t < 1\}$ , para os processos GARCH.

### 3.3.2 Modelos EWMA

Dada uma série temporal  $\{X_t\}_{t=1}^n$  obtida a partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , o parâmetro  $\lambda$ , no modelo EWMA, dado pela expressão (2.34), é escolhido de forma a minimizar a soma dos quadrados dos resíduos de ajuste ou de previsão.

### 3.3.3 Processos FIEGARCH

Seja  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ . Para ajustar um modelo FIEGARCH( $p, d, q$ ) à essa série temporal, devemos levar em conta que a volatilidade  $\sigma_t$  é definida recursivamente em termos de  $\sigma_{t-1}, \dots, \sigma_{t-q}$  e  $Z_{t-1}, \dots, Z_{t-1-p}$ , onde  $Z_t = \frac{X_t}{\sigma_t}$ , para todo  $t = 1, \dots, n$ .

Se considerarmos a expressão (2.45), o vetor de parâmetros  $\boldsymbol{\theta}$  a ser estimado é dado por

$$\boldsymbol{\theta} = (d; \omega; \theta; \lambda; \alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_q)' \in \mathbb{R}^{p+q+4}. \quad (3.28)$$

Sob a hipótese de normalidade das variáveis aleatórias  $X_t | \mathcal{F}_{t-1}$ , de acordo com a Definição 3.7, o estimador  $\hat{\boldsymbol{\theta}}_{PL}$  de  $\boldsymbol{\theta}$  é obtido maximizando-se a expressão

$$\ln(\ell(\boldsymbol{\theta}; \mathbf{x})) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^n \left[ \ln(\sigma_t^2) + \frac{X_t^2}{\sigma_t^2} \right]. \quad (3.29)$$

Dado que não conhecemos os valores do processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  quando  $t < 1$ , a função de quase-verossimilhança é obtida condicionando-se  $\{X_t\}_{t=1}^n$  sobre  $I_0$ , onde  $I_0$  é um conjunto de condições iniciais. Bollerslev e Mikkelsen (1996) sugerem as seguintes condições iniciais

$$X_t = 0 \quad \text{e} \quad X_t^2 = \sigma_t^2 = \hat{\sigma}_X^2, \quad \text{para todo } t < 1, \quad (3.30)$$

onde  $\hat{\sigma}_X^2$  é a variância amostral da série temporal  $\{X_t\}_{t=1}^n$ . Note que, dadas as condições iniciais na expressão (3.30) temos, pela expressão (2.45),

$$\hat{\sigma}_t^2 = e^{\hat{\omega}}, \quad \text{para todo } t \leq 1,$$

pois,  $Z_t = \frac{X_t}{\sigma_t} = 0$ , sempre que  $t < 1$ . Logo,  $\hat{\omega} = \ln(\hat{\sigma}_X^2)$ .

Por hipótese,  $X_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, \sigma_t^2)$ . Logo,  $\frac{X_t}{\sigma_t} | \mathcal{F}_{t-1} \sim \mathcal{N}(0, 1)$ . Ou seja,  $Z_t \sim \mathcal{N}(0, 1)$ , pois  $Z_t$  é independente de  $\mathcal{F}_{t-1}$ , para todo  $t \in \mathbb{Z}$ . Segue que

$$\mathbb{E} \left( \left| \frac{X_t}{\sigma_t} \right| \right) = \mathbb{E} (|Z_t|) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z| e^{-z^2/2} dz = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} z e^{-z^2/2} dz = \sqrt{\frac{2}{\pi}},$$

e

$$\begin{aligned} \ln(\sigma_t^2) &= \omega + \lambda(\mathcal{B}) \left( \theta \frac{X_{t-1}}{\sigma_{t-1}} + \gamma \left[ \left| \frac{X_{t-1}}{\sigma_{t-1}} \right| - \mathbb{E} \left( \left| \frac{X_{t-1}}{\sigma_{t-1}} \right| \right) \right] \right), \\ &= \omega - \lambda(1) \gamma \sqrt{\frac{2}{\pi}} + \lambda(\mathcal{B}) \left( \theta \frac{X_{t-1}}{\sigma_{t-1}} + \gamma \left| \frac{X_{t-1}}{\sigma_{t-1}} \right| \right), \quad \text{para todo } t > 1, \end{aligned}$$

onde  $\lambda(\cdot)$  é dado pela expressão (2.52). Note que, para calcular o estimador de quase-máxima verossimilhança, será necessário truncar a série  $\lambda(\cdot)$  definida pela expressão (2.52) em algum ponto de truncamento  $m$ . Tal fato tem grande influência na precisão dos cálculos. Bollerslev e Mikkelsen (1996) sugerem utilizar  $m = 1000$  para a estimação do parâmetro  $\theta$  dado pela expressão (3.28).

Pela Proposição 2.3, as expressões (2.45) e (2.46) são equivalentes, dessa forma, se considerarmos a expressão (2.46), temos

$$\theta = (d; a; \phi_1, \dots, \phi_p; \gamma_1, \dots, \gamma_p; \beta_1, \dots, \beta_q)' \in \mathbb{R}^{2p+q+2}.$$

Novamente, de acordo com a Definição 3.7, o estimador  $\hat{\theta}_{PL}$  de  $\theta$  é então obtido maximizando-se a expressão (3.29). Dadas as condições iniciais na expressão (3.30) temos, pela expressão (2.45),

$$\beta(\mathcal{B})(1 - \mathcal{B})^d \ln(\sigma_t^2) = a, \quad \text{para } t = 1. \quad (3.31)$$

Ou seja, considerando as expressões (2.55) e (3.31),

$$\sum_{k=0}^{\infty} \left( \sum_{i=1}^{\min\{k,q\}} (-\delta_{d,k-i}) \beta_i \right) \ln(\sigma_{t-k}^2) = a, \quad \text{para } t = 1. \quad (3.32)$$

Como definimos  $\sigma_t^2 = \hat{\sigma}_X^2$ , para todo  $t < 1$ , segue que a expressão (3.32) pode ser reescrita por

$$\ln(\sigma_t^2) = a + \sum_{k=1}^{\infty} \left( \sum_{i=1}^{\min\{k,q\}} \delta_{d,k-i} \beta_i \right) \ln(\hat{\sigma}_X^2), \quad \text{para } t = 1.$$

Enquanto que, para todo  $t > 1$ , temos

$$\beta(\mathcal{B})(1 - \mathcal{B})^d \ln(\sigma_t^2) = a + \sum_{i=0}^p \left( \psi_i |Z_{t-1-i}| + \gamma_i Z_{t-1-i} \right). \quad (3.33)$$

Ou seja, pelas expressões (2.55) e (3.33), para todo  $t > 1$

$$\ln(\sigma_t^2) = a + \sum_{k=1}^{\infty} \left( \sum_{i=1}^{\min\{k,q\}} \delta_{d,k-i} \beta_i \right) \ln(\sigma_{t-k}^2) + \sum_{i=0}^p \left( \psi_i \left| \frac{X_{t-1-i}}{\sigma_{t-1-i}} \right| + \gamma_i \frac{X_{t-1-i}}{\sigma_{t-1-i}} \right).$$

Note que, neste caso, para calcular o estimador de quase-máxima verossimilhança, será necessário truncar a série  $\beta(\mathcal{B})(1 - \mathcal{B})^d$ . Como dito anteriormente, tal fato tem grande influência na precisão dos cálculos.

Para o leitor interessado, apresentamos, no Apêndice C, o valor aproximado da soma  $S = -\sum_{k=1}^m \delta_{d,k}$ , para diferentes valores de  $m$  e  $d$ . A Tabela C.1 nos mostra que, para valores  $m \geq 4.000$  e  $d \geq 0,25$ , a soma  $S$  é aproximadamente igual a 0,9.

## 3.4 Previsão

Nesta seção tratamos da previsão da volatilidade em modelos ARCH, GARCH, IGARCH, EWMA e FIEGARCH. Devido à analogia com os processos ARMA e ARFIMA apresentamos alguns resultados relacionados também a estes processos. Para um estudo mais detalhado veja Brockwell e Davis (1991).

### 3.4.1 Previsão em Processos Estacionários

Sejam  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário e  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir deste processo. Nosso objetivo é utilizar as observações até o instante  $n$  para prever o comportamento do processo  $\{X_t\}_{t \in \mathbb{Z}}$  para algum instante  $t > n$ .

**Definição 3.8. (Melhor Preditor Linear).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico estacionário. Considere a  $\sigma$ -álgebra  $\mathcal{F}_n = \sigma(\{X_s\}; s \leq n)$ . O *melhor preditor linear* de  $X_{n+h}$ , para  $h \in \mathbb{Z}$  passos à frente, em relação ao erro quadrático médio, é dado por

$$\hat{X}_{n+h} \equiv \mathbb{E}(X_{n+h} | \mathcal{F}_n).$$

Seja  $\{Y_t\}_{t \in \mathbb{Z}}$  um processo ARFIMA( $p, d, q$ ), com  $d \in (-0,5; 0,5)$ , dado na Definição 1.29. Se  $\{Y_t\}_{t \in \mathbb{Z}}$  é causal e inversível então

$$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \quad \text{e} \quad \epsilon_t = \sum_{j=0}^{\infty} \pi_j Y_{t-j}, \quad \text{para todo } t \in \mathbb{Z}, \quad (3.34)$$

onde  $\{\epsilon\}_{t \in \mathbb{Z}}$  é um processo ruído branco,  $\psi(\mathcal{B}) = \frac{\theta(\mathcal{B})}{\phi(\mathcal{B})}(1 - \mathcal{B})^{-d}$  e  $\pi(\mathcal{B}) = \frac{\phi(\mathcal{B})}{\theta(\mathcal{B})}(1 - \mathcal{B})^d$ .

Seja  $n$  um inteiro positivo fixo. Então, pela Definição 3.8 e pelas equações dadas na expressão (3.34), segue que, para todo  $h \in \mathbb{Z}$ , o preditor  $h$  passos à frente  $\hat{Y}_{n+h}$  é dado

por

$$\begin{aligned}\hat{Y}_{n+h} = \mathbb{E}(Y_{n+h}|\mathcal{F}_n) &= -\sum_{j=1}^{\infty} \pi_j \hat{Y}_{n+h-j} \\ &= -\sum_{j=1}^{h-1} \pi_j \hat{Y}_{n+h-j} - \sum_{j=h}^{\infty} \pi_j Y_{n+h-j}.\end{aligned}$$

Isto é,

$$\hat{Y}_{n+h} = -\sum_{j=1}^{h-1} \pi_j \hat{Y}_{n+h-j} - \sum_{j=0}^{\infty} \pi_{j+h} Y_{n-j}. \quad (3.35)$$

A equação (3.35) pode ser calculada recursivamente dado que  $\hat{Y}_t = Y_t$ , se  $t \leq n$ .

De forma equivalente,  $\hat{Y}_{n+h}$ , dado na expressão (3.35), pode ser obtido por

$$\hat{Y}_{n+h} = \sum_{j=0}^{\infty} \psi_j \hat{\epsilon}_{n+h-j} = \sum_{j=h}^{\infty} \psi_j \epsilon_{n+h-j} = \sum_{j=0}^{\infty} \psi_{h+j} \epsilon_{n-j}, \quad (3.36)$$

pois  $\mathbb{E}(\epsilon_{n+h}|\mathcal{F}_n) = \mathbb{E}(\epsilon_{n+h}) = 0$ , para todo  $h > 1$  e  $\mathbb{E}(\epsilon_t|\mathcal{F}_n) = \epsilon_t$ , para todo  $t \leq n$ .

Pelas expressões (3.34) e (3.36), temos que o erro quadrático médio (teórico) de previsão é dado por

$$\begin{aligned}\mathbb{E}((Y_{n+h} - \hat{Y}_{n+h})^2) &= \mathbb{E}\left(\left(\sum_{j=0}^{\infty} \psi_j \epsilon_{n+h-j} - \sum_{j=h}^{\infty} \psi_j \epsilon_{n+h-j}\right)^2\right) \\ &= \mathbb{E}\left(\left(\sum_{j=0}^{h-1} \psi_j \epsilon_{n+h-j}\right)^2\right) \\ &= \sigma_{\epsilon}^2 \sum_{j=0}^{h-1} \psi_j^2,\end{aligned} \quad (3.37)$$

onde  $\sigma_{\epsilon}^2 = \text{Var}(\epsilon_t)$ , para todo  $t \in \mathbb{Z}$ .

Note que, quando  $d = 0$  na equação (1.16), temos que  $\{Y_t\}_{t \in \mathbb{Z}}$  é um processo ARMA( $p, q$ ) e as equações (3.35) e (3.37) continuam válidas. Note que, neste caso

$$\phi(\mathcal{B})Y_t = \theta(\mathcal{B})\epsilon_t \quad \Leftrightarrow \quad Y_t = \sum_{i=1}^p \phi_i Y_{t-i} - \sum_{j=0}^q \theta_j \epsilon_{t-j}, \quad (3.38)$$

para todo  $t \in \mathbb{Z}$ . Neste caso, o preditor  $h$  passos à frente  $\hat{Y}_{n+h}$  pode ser obtido, recursivamente, através da expressão

$$\hat{Y}_{n+h} = \sum_{i=1}^p \phi_i \hat{Y}_{n+h-i} - \sum_{j=0}^q \theta_j \hat{\epsilon}_{n+h-j}, \quad (3.39)$$

onde

$$\begin{aligned}\hat{Y}_{n+h} &= Y_{n+h}, \quad \text{se } h \leq 0; \\ \hat{\epsilon}_{n+h} &= \epsilon_{n+h}, \quad \text{se } h \leq 0; \\ \hat{\epsilon}_{n+h} &= 0, \quad \text{se } h > 0.\end{aligned} \quad (3.40)$$

### 3.4.2 Previsão da Volatilidade em Processos ARCH e GARCH

Nesta seção tratamos da previsão em processos ARCH e GARCH. Note que os processos ARCH( $p$ ) são um caso particular dos processos GARCH( $p, q$ ), quando  $q = 0$ . Portanto, sem perda de generalidade, consideramos o caso mais geral, isto é, tratamos da previsão em processos GARCH.

Seja  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , à qual ajustamos um modelo GARCH( $p, q$ ). A previsão de valores futuros da volatilidade para este modelo, isto é, a previsão de valores  $\sigma_{n+h}^2$ , para  $h \geq 1$ , está fortemente relacionada com a previsão de valores futuros para o processo  $\{X_t^2\}_{t \in \mathbb{Z}}$ , como vemos a seguir.

Dado que o processo GARCH( $p, q$ ) possui uma representação ARMA( $\max\{p, q\}, q$ ), dada pela expressão (2.8), utilizamos métodos de previsão análogos aos utilizados para previsão em processos ARMA.

**Lema 3.1.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo GARCH( $p, q$ ) estacionário. Então, os preditores de  $X_{n+h}$  e  $X_{n+1}^2$ , dado  $\mathcal{F}_n$ , são  $\hat{X}_{n+h} = 0$  e  $\hat{X}_{n+1}^2 = \sigma_{n+1}^2$ , respectivamente.*

**Prova:** Note que, se o processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  é estacionário então

$$\mathbb{E}(X_t^2) = \mathbb{E}(\sigma_t^2) < \infty, \quad \text{para todo } t \in \mathbb{Z}.$$

Além disso, dado que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo *martingale difference*,  $\mathbb{E}(X_{n+h} | \mathcal{F}_n) = 0$ . Pela Definição 3.8, segue que o preditor de  $X_{n+h}$ , dado  $\mathcal{F}_n$ , é  $\hat{X}_{n+h} = 0$ . Ainda, pela Definição 3.8, segue que,

$$\hat{X}_{n+1}^2 \equiv \mathbb{E}(X_{n+1}^2 | \mathcal{F}_n) = \sigma_{n+1}^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{n+1-i}^2 + \sum_{j=1}^q \beta_j \sigma_{n+1-j}^2. \quad (3.41)$$

Note que, se  $\mathbb{E}(X_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ , então esse é o preditor ótimo em termos de menor erro quadrático médio.  $\square$

É fácil ver que, para  $h > 1$  passos à frente dado  $\mathcal{F}_n$ , as predições das variáveis aleatórias  $X_{n+h}^2$  e  $\sigma_{n+h}^2$  coincidem, pois

$$\hat{X}_{n+h}^2 \equiv \mathbb{E}(X_{n+h}^2 | \mathcal{F}_n) = \mathbb{E}(\sigma_{n+h}^2 Z_{n+h}^2 | \mathcal{F}_n) = \mathbb{E}(\sigma_{n+h}^2 | \mathcal{F}_n) \mathbb{E}(Z_{n+h}^2 | \mathcal{F}_n) = \mathbb{E}(\sigma_{n+h}^2 | \mathcal{F}_n) \equiv \hat{\sigma}_{n+h}^2, \quad (3.42)$$

pois  $Z_t$  é independente de  $\sigma_t$  e de  $\mathcal{F}_t$ , para todo  $t \in \mathbb{Z}$  e  $\mathbb{E}(Z_{n+h}^2 | \mathcal{F}_n) = \mathbb{E}(Z_{n+h}^2) = 1$ . Considere agora a representação dada na expressão (2.8). Segue que, como  $v_t = X_t^2 - \sigma_t^2$ , para todo  $t \in \mathbb{Z}$ , então

$$\mathbb{E}(v_{t+h} | \mathcal{F}_n) = \mathbb{E}(X_{n+h}^2 | \mathcal{F}_n) - \mathbb{E}(\sigma_{n+h}^2 | \mathcal{F}_n) = 0, \quad \text{para todo } h \geq 1. \quad (3.43)$$

Como  $\{X_t^2\}_{t \in \mathbb{Z}}$  é um processo ARMA( $\max\{p, q\}, q$ ), de acordo com as expressões (3.38) e (3.39), temos

$$\begin{aligned} (1 - \alpha(\mathcal{B}) - \beta(\mathcal{B}))X_{n+h}^2 &= \alpha_0 + (1 - \beta(\mathcal{B}))v_{n+h} \\ &\Downarrow \\ X_{n+h}^2 &= \alpha_0 + (\alpha(\mathcal{B}) + \beta(\mathcal{B}))X_{n+h}^2 + (1 - \beta(\mathcal{B}))v_{n+h} \\ &\Downarrow \\ \hat{X}_{n+h}^2 &= \alpha_0 + (\alpha(\mathcal{B}) + \beta(\mathcal{B}))\hat{X}_{n+h}^2 + (1 - \beta(\mathcal{B}))\hat{v}_{n+h}, \end{aligned} \quad (3.44)$$

onde

$$\begin{aligned}\hat{X}_{n+h} &= X_{n+h}, & \text{se } h \leq 0; \\ \hat{v}_{n+h} &= v_{n+h}, & \text{se } h \leq 0; \\ \hat{v}_{n+h} &= 0, & \text{se } h > 0.\end{aligned}$$

Por outro lado, dado que

$$\sigma_{n+h}^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{n+h-i}^2 + \sum_{j=1}^q \beta_j \sigma_{n+h-j}^2,$$

pela expressão (3.42) segue que

$$\begin{aligned}\hat{X}_{n+h}^2 &\equiv \mathbb{E}(X_{n+h}^2 | \mathcal{F}_n) = \mathbb{E}(\sigma_{n+h}^2 | \mathcal{F}_n) \\ &= \alpha_0 + \sum_{i=1}^p \alpha_i \mathbb{E}(X_{n+h-i}^2 | \mathcal{F}_n) + \sum_{j=1}^q \beta_j \mathbb{E}(\sigma_{n+h-j}^2 | \mathcal{F}_n) \\ &= \alpha_0 + \sum_{i=1}^p \alpha_i \hat{X}_{n+h-i}^2 + \sum_{j=1}^q \beta_j \hat{\sigma}_{n+h-j}^2,\end{aligned}\tag{3.45}$$

onde  $\hat{X}_t^2 = X_t$ , se  $t \leq n$ , e  $\hat{\sigma}_t^2 = \sigma_t^2$ , se  $t \leq n+1$ .

Note que as expressões (3.44) e (3.45) são equivalentes pois

$$\alpha_0 + (\alpha(\mathcal{B}) + \beta(\mathcal{B}))X_{n+h}^2 + (1 - \beta(\mathcal{B}))v_{n+h} = \alpha_0 + \alpha(\mathcal{B})X_{n+h}^2 + \beta(\mathcal{B})X_{n+h}^2 + X_{n+h}^2 - \sigma_{n+h}^2 - \beta(\mathcal{B})X_{n+h}^2 + \beta(\mathcal{B})\sigma_{n+h}^2$$

↓

$$\begin{aligned}\alpha_0 + (\alpha(\mathcal{B}) + \beta(\mathcal{B}))\hat{X}_{n+h}^2 + (1 - \beta(\mathcal{B}))\hat{v}_{n+h} &= \alpha_0 + \alpha(\mathcal{B})\hat{X}_{n+h}^2 + \beta(\mathcal{B})\hat{X}_{n+h}^2 + \hat{X}_{n+h}^2 \\ &\quad - \hat{\sigma}_{n+h}^2 - \beta(\mathcal{B})\hat{X}_{n+h}^2 + \beta(\mathcal{B})\hat{\sigma}_{n+h}^2 \\ &= \alpha_0 + \alpha(\mathcal{B})\hat{X}_{n+h}^2 + \beta(\mathcal{B})\hat{\sigma}_{n+h}^2,\end{aligned}$$

pois  $\hat{X}_{n+h}^2 = \hat{\sigma}_{n+h}^2$ , para todo  $h > 1$ .

No próximo exemplo apresentamos a fórmula geral para o cálculo das previsões  $h$  passos à frente em um modelo GARCH( $p, q$ ), com  $p = 1 = q$ .

**Exemplo 3.1. (Modelo GARCH( $p, q$ ), com  $p = 1 = q$ ).** Considere  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ . Se ajustarmos um modelo GARCH( $p, q$ ), com  $p = 1 = q$ , à essa série temporal, pela expressão (3.41), a previsão 1 passo à frente é dada por

$$\hat{X}_{n+1}^2 = \mathbb{E}(X_{n+1}^2 | \mathcal{F}_n) = \sigma_{n+1}^2 = \alpha_0 + \alpha_1 X_n^2 + \beta_1 \sigma_n^2.\tag{3.46}$$

Pela expressão (3.45), a previsão para  $h > 1$  passos à frente é dada por

$$\begin{aligned}
\hat{X}_{n+h}^2 = \mathbb{E}(X_{n+h}^2 | \mathcal{F}_n) &= \mathbb{E}(\sigma_{n+h}^2 | \mathcal{F}_n) \\
&= \alpha_0 + \alpha_1 \mathbb{E}(X_{n+h-1}^2 | \mathcal{F}_n) + \beta_1 \mathbb{E}(\sigma_{n+h-1}^2 | \mathcal{F}_n) \\
&= \alpha_0 + (\alpha_1 + \beta_1) \mathbb{E}(X_{n+h-1}^2 | \mathcal{F}_n) \\
&= \alpha_0 + (\alpha_1 + \beta_1) (\alpha_0 + \alpha_1 \mathbb{E}(X_{n+h-2}^2 | \mathcal{F}_n) + \beta_1 \mathbb{E}(\sigma_{n+h-2}^2 | \mathcal{F}_n)) \\
&= \alpha_0 + \alpha_0(\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 \mathbb{E}(X_{n+h-2}^2 | \mathcal{F}_n) \\
&\vdots \\
&= \alpha_0 \sum_{j=0}^{h-2} (\alpha_1 + \beta_1)^j + (\alpha_1 + \beta_1)^{h-1} \mathbb{E}(X_{n+1}^2 | \mathcal{F}_n). \tag{3.47}
\end{aligned}$$

Substituindo a expressão (3.46) na expressão (3.47) temos,

$$\begin{aligned}
\hat{X}_{n+h}^2 &= \alpha_0 \sum_{j=0}^{h-2} (\alpha_1 + \beta_1)^j + (\alpha_1 + \beta_1)^{h-1} (\alpha_0 + \alpha_1 X_n^2 + \beta_1 \sigma_n^2) \\
&= \alpha_0 \sum_{j=0}^{h-1} (\alpha_1 + \beta_1)^j + (\alpha_1 + \beta_1)^{h-1} (\alpha_1 X_n^2 + \beta_1 \sigma_n^2).
\end{aligned}$$

Note que,  $\hat{X}_{n+h}^2 \rightarrow \sigma_X^2$ , quando  $h \rightarrow \infty$ , onde  $\sigma_X^2$  é a variância não-condicional do processo GARCH( $p, q$ ), com  $p = 1 = q$ , dada na expressão (2.9). □

Mais detalhes sobre previsão em modelos GARCH podem ser encontrados em McNeil *et al.* (2005).

### 3.4.3 Previsão da Volatilidade em Modelos IGARCH e EWMA

Neste trabalho tratamos apenas da previsão da volatilidade dos processos IGARCH( $p, q$ ), com  $p = 1 = q$ , dado que a previsão da volatilidade nos modelos EWMA é equivalente a previsão de tais processos.

Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo IGARCH( $p, q$ ), com  $p = 1 = q$ , com variância condicional dada por

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{3.48}$$

com  $\alpha_1 + \beta_1 = 1$ . Assim como nos processos GARCH( $p, q$ ), para  $n$  fixo, um preditor natural de  $X_{n+1}^2$  dado  $\mathcal{F}_n$  é sua média condicional  $\sigma_{n+1}^2$  dada por

$$\hat{X}_{n+1}^2 \equiv \mathbb{E}(X_{n+1}^2 | \mathcal{F}_n) = \sigma_{n+1}^2 = \alpha_0 + \alpha_1 X_n^2 + \beta_1 \sigma_n^2. \tag{3.49}$$

Agora, note que, pela expressão (3.42), para  $h > 1$ ,

$$\hat{X}_{n+h}^2 \equiv \mathbb{E}(X_{n+h}^2 | \mathcal{F}_n) = \mathbb{E}(\sigma_{n+h}^2 | \mathcal{F}_n) \equiv \hat{\sigma}_{n+h}^2. \tag{3.50}$$

Pelas expressões (3.48), (3.49) e (3.50), observe que



$$\begin{aligned}
\hat{\sigma}_{n+2}^2 = \mathbb{E}(\sigma_{n+2}^2 | \mathcal{F}_n) &= \alpha_0 + \alpha_1 \mathbb{E}(X_{n+1}^2 | \mathcal{F}_n) + \beta_1 \mathbb{E}(\sigma_{n+1}^2 | \mathcal{F}_n) \\
&= \alpha_0 + \alpha_1 \sigma_{n+1}^2 + \beta_1 \sigma_{n+1}^2 \\
&= \alpha_0 + \sigma_{n+1}^2, \quad \text{pois } \mathbb{E}(\sigma_{n+1}^2 | \mathcal{F}_n) = \sigma_{n+1}^2; \\
\hat{\sigma}_{n+3}^2 = \mathbb{E}(\sigma_{n+3}^2 | \mathcal{F}_n) &= \alpha_0 + \alpha_1 \mathbb{E}(X_{n+2}^2 | \mathcal{F}_n) + \beta_1 \mathbb{E}(\sigma_{n+2}^2 | \mathcal{F}_n) \\
&= \alpha_0 + (\alpha_1 + \beta_1) \mathbb{E}(\sigma_{n+2}^2 | \mathcal{F}_n) \\
&= \alpha_0 + \hat{\sigma}_{n+2}^2 \\
&= 2\alpha_0 + \sigma_{n+1}^2; \\
&\vdots \\
\hat{\sigma}_{n+h}^2 = \mathbb{E}(\sigma_{n+h}^2 | \mathcal{F}_n) &= (h-1)\alpha_0 + \sigma_{n+1}^2. \tag{3.51}
\end{aligned}$$

Ou seja, pela expressão (3.51), para todo  $h > 1$ , o preditor  $h$  passos à frente para a volatilidade é dado por

$$\hat{\sigma}_{n+h}^2 = (h-1)\alpha_0 + \sigma_{n+1}^2. \tag{3.52}$$

Lembramos que o modelo EWMA é obtido quando  $\alpha_0 = 0$ , na expressão (3.48), e ainda,  $\alpha_1 = \lambda$  e  $\beta_1 = 1 - \lambda$ . Segue, da expressão (3.52) que, para todo  $h > 1$ , o preditor  $h$  passos à frente para a volatilidade é dado por

$$\hat{\sigma}_{n+h}^2 = \sigma_{n+1}^2. \tag{3.53}$$

### 3.4.4 Previsão da Volatilidade em Processos EGARCH e FIEGARCH

Assim como os processos ARCH são um caso particular dos processos GARCH, os processos EGARCH( $p, q$ ) são um caso particular dos processos FIEGARCH( $p, d, q$ ), quando  $d = 0$ . Portanto, nesta seção consideramos apenas o caso mais geral.

Seja  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir de um processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , à qual ajustamos um modelo FIEGARCH( $p, d, q$ ). Assim como nos modelos GARCH, a previsão de valores futuros da volatilidade para o modelo FIEGARCH está fortemente relacionada com a previsão de valores futuros para o processo  $\{X_t^2\}_{t \in \mathbb{Z}}$ .

Dado que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo FIEGARCH( $p, d, q$ ), o processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$ , definido pela expressão (2.45), possui uma representação ARFIMA( $q, d, p$ ). Sendo assim, utilizamos métodos de previsão análogos aos utilizados para previsão em processos ARFIMA.

**Lema 3.2.** *Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo FIEGARCH( $p, d, q$ ) estacionário. Então, os preditores de  $X_{n+h}$  e  $X_{n+1}^2$ , dado  $\mathcal{F}_n$ , são  $\hat{X}_{n+h} = 0$  e  $\hat{X}_{n+1}^2 = \sigma_{n+1}^2$ , respectivamente.*

**Prova:** Dado que um processo FIEGARCH( $p, d, q$ ) é um processo *martingale-difference* (ver Lema 2.1), segue, de imediato que  $\hat{X}_{n+h} = \mathbb{E}(X_{n+h} | \mathcal{F}_n) = 0$ . Lembramos que, por definição,  $\mathbb{E}(X_{n+1}^2 | \mathcal{F}_n) = \sigma_{n+1}^2$ . Segue que o preditor  $X_{n+1}^2$ , dado  $\mathcal{F}_n$ , é  $\sigma_{n+1}^2$ . Pela expressão (2.45),  $\sigma_{n+1}^2$  é dada por

$$\sigma_{n+1}^2 = \exp \left\{ \omega + \sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{n-k}) \right\},$$

onde os coeficientes  $\lambda_{d,k}$ , para todo  $k \in \mathbb{Z}$ , são definidos pela expressão (2.53) e a função  $g(\cdot)$  é definida pela expressão (2.38). Note que, se  $\mathbb{E}(X_t^4) < \infty$ , para todo  $t \in \mathbb{Z}$ , então esse é o preditor ótimo em termos de menor erro quadrático médio.  $\square$

Para a previsão de  $X_{n+h}^2$ , para todo  $h > 1$ , lembramos que, pela expressão (3.42),

$$\hat{X}_{n+h}^2 = \mathbb{E}(X_{n+h}^2 | \mathcal{F}_n) = \mathbb{E}(\sigma_{n+h}^2 | \mathcal{F}_n) = \hat{\sigma}_{n+h}^2.$$

Além disso, o processo estocástico  $\{\ln(\sigma_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, p$ ) pois, pela Proposição 2.2,  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  é um processo ruído branco. Como

$$\begin{aligned} \ln(\sigma_{n+h}^2) &= \omega + \frac{\alpha(\mathcal{B})g(Z_{n+h-1})}{\beta(\mathcal{B})(1-\mathcal{B})^d} \\ &= \omega + \sum_{k=0}^{\infty} \lambda_{d,k}g(Z_{n+h-1-k}), \end{aligned} \quad (3.54)$$

segue, por analogia à expressão (3.36), que

$$\ln(\hat{\sigma}_{n+h}^2) = \omega + \sum_{k=h-1}^{\infty} \lambda_{d,k}g(Z_{n+h-1-k}) = \omega + \sum_{k=0}^{\infty} \lambda_{d,k+h-1}g(Z_{n-k}). \quad (3.55)$$

De fato, basta considerar  $Y_t = \ln(\sigma_t^2)$  e  $\epsilon_t = g(Z_{t-1})$ , para todo  $t \in \mathbb{Z}$ , e o resultado segue. Assim, pela expressão (3.55), o preditor  $h$  passos à frente para a volatilidade é dado por

$$\hat{\sigma}_{n+h}^2 = \exp \left\{ \omega + \sum_{k=0}^{\infty} \lambda_{d,k+h-1}g(Z_{n-k}) \right\}. \quad (3.56)$$

Pelo Teorema 2.6,  $\lambda_{d,k} = o(k^d)$ . Logo,  $\lim_{k \rightarrow \infty} \lambda_{d,k} = 0$ . Segue, pela expressão (3.56) que,

$$\hat{\sigma}_{n+h}^2 \rightarrow e^\omega, \quad \text{quando } h \rightarrow \infty.$$

Note que, como conhecemos apenas os valores de  $X_t$  e  $\sigma_t$ , para  $t = 1, \dots, n$ , a expressão (3.56) será aproximada por

$$\hat{\sigma}_{n+h}^2 = \exp \left\{ \omega + \sum_{k=0}^{n-1} \lambda_{d,k+h-1}g(Z_{n-k}) \right\}. \quad (3.57)$$

Dessa forma, pela expressão (3.57), o erro quadrático médio de previsão, para o logaritmo da variância condicional, é dado por

$$\begin{aligned} \mathbb{E} \left( \left( \ln(\sigma_{n+h}^2) - \ln(\hat{\sigma}_{n+h}^2) \right)^2 \right) &= \mathbb{E} \left( \left( \sum_{k=0}^{\infty} \lambda_{d,k}g(Z_{n+h-1-k}) - \sum_{k=0}^{n-1} \lambda_{d,k+h-1}g(Z_{n-k}) \right)^2 \right) \\ &= \mathbb{E} \left( \left( \sum_{k=0}^{\infty} \lambda_{d,k}g(Z_{n+h-1-k}) - \sum_{k=h-1}^{n+h-2} \lambda_{d,k}g(Z_{n+h-1-k}) \right)^2 \right) \\ &= \mathbb{E} \left( \left( \sum_{k=0}^{h-2} \lambda_{d,k}g(Z_{n+h-1-k}) + \sum_{k=n+h-1}^{\infty} \lambda_{d,k}g(Z_{n+h-1-k}) \right)^2 \right). \end{aligned} \quad (3.58)$$

Note que

$$\begin{aligned}
& \mathbb{E} \left( \left( \sum_{k=0}^{h-2} \lambda_{d,k} g(Z_{n+h-1-k}) + \sum_{k=n+h-1}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) \right)^2 \right) \\
&= \mathbb{E} \left( \left( \sum_{j=0}^{h-2} \lambda_{d,j} g(Z_{n+h-1-j}) \right)^2 + 2 \left( \sum_{j=0}^{h-2} \lambda_{d,j} g(Z_{n+h-1-j}) \right) \left( \sum_{k=n+h-1}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) \right) \right. \\
&\quad \left. + \left( \sum_{k=n+h-1}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) \right)^2 \right). \tag{3.59}
\end{aligned}$$

Como  $\{g(Z_t)\}_{t \in \mathbb{Z}}$  é um processo ruído branco temos

$$\mathbb{E}(g(Z_t)) = 0 \quad \text{e} \quad \mathbb{E}(g(Z_t)g(Z_k)) = 0, \quad \text{para todo } t, k \in \mathbb{Z}, \text{ tais que } t \neq k.$$

Segue, da expressão (3.59) que

$$\mathbb{E} \left( \left( \sum_{k=0}^{h-2} \lambda_{d,k} g(Z_{n+h-1-k}) + \sum_{k=n+h-1}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) \right)^2 \right) = \sigma_g^2 \left( \sum_{k=0}^{h-2} \lambda_{d,k}^2 + \sum_{k=n+h-1}^{\infty} \lambda_{d,k}^2 \right), \tag{3.60}$$

onde  $\sigma_g^2 \equiv \text{Var}(g(Z_t)) = \mathbb{E}([g(Z_t)]^2)$ , para todo  $t \in \mathbb{Z}$ .

Substituindo a expressão (3.60) na expressão (3.58), segue que o erro quadrático médio de previsão, para o logaritmo da volatilidade, é dado por

$$\mathbb{E} \left( \left( \ln(\sigma_{n+h}^2) - \ln(\hat{\sigma}_{n+h}^2) \right)^2 \right) = \sigma_g^2 \left( \sum_{k=0}^{h-2} \lambda_{d,k}^2 + \sum_{k=n+h-1}^{\infty} \lambda_{d,k}^2 \right), \tag{3.61}$$

onde  $\sigma_g^2 = \mathbb{E}([g(Z_t)]^2)$ , para todo  $t \in \mathbb{Z}$ . Além disso, dado que  $Z_{t+h}$  é independente de  $X_t$ , para todo  $t \in \mathbb{Z}$  e para todo  $h > 0$ , segue que  $\hat{Z}_{n+h}^2 \equiv \mathbb{E}(Z_{n+h}^2 | \mathcal{F}_n) = \mathbb{E}(Z_{n+h}^2) = 1$ . Logo,

$$\ln(\hat{X}_{n+h}^2) = \ln(\hat{\sigma}_{n+h}^2) \iff \hat{X}_{n+h}^2 = \hat{\sigma}_{n+h}^2, \tag{3.62}$$

Segue, pelas expressões (3.57) e (3.62), que o erro quadrático médio de previsão para  $\ln(X_{n+h}^2)$  é dado por

$$\begin{aligned}
& \mathbb{E} \left( \left( \ln(X_{n+h}^2) - \ln(\hat{X}_{n+h}^2) \right)^2 \right) = \mathbb{E} \left( \left( \ln(\sigma_{n+h}^2) + \ln(Z_{n+h}^2) - \ln(\hat{\sigma}_{n+h}^2) \right)^2 \right) \\
&= \mathbb{E} \left( \left( \sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) + \ln(Z_{n+h}^2) - \sum_{k=0}^{n-1} \lambda_{d,k+h-1} g(Z_{n-k}) \right)^2 \right) \\
&= \mathbb{E} \left( \left( \sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) + \ln(Z_{n+h}^2) - \sum_{k=h-1}^{n+h-2} \lambda_{d,k} g(Z_{n+h-1-k}) \right)^2 \right) \\
&= \mathbb{E} \left( \left( \sum_{k=0}^{h-2} \lambda_{d,k} g(Z_{n+h-1-k}) + \sum_{k=n+h-1}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) + \ln(Z_{n+h}^2) \right)^2 \right). \tag{3.63}
\end{aligned}$$

Note que

$$\begin{aligned}
& \mathbb{E} \left( \left( \sum_{k=0}^{h-2} \lambda_{d,k} g(Z_{n+h-1-k}) + \sum_{k=n+h-1}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) + \ln(Z_{n+h}^2) \right)^2 \right) \\
&= \mathbb{E} \left( \left[ \sum_{k=0}^{h-2} \lambda_{d,k} g(Z_{n+h-1-k}) + \sum_{k=n+h-1}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) \right]^2 \right) + \mathbb{E} \left( \left( \ln(Z_{n+h}^2) \right)^2 \right) \\
&+ 2 \mathbb{E} \left( \left[ \sum_{k=0}^{h-2} \lambda_{d,k} g(Z_{n+h-1-k}) + \sum_{k=n+h-1}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) \right] \ln(Z_{n+h}^2) \right). \quad (3.64)
\end{aligned}$$

Como as variáveis aleatórias  $\{Z_t\}_{t \in \mathbb{Z}}$  são independentes,  $g(Z_t)$  e  $\ln(Z_{t+k})$  são independentes, para todo  $t \in \mathbb{Z}$  e todo  $k \neq 0$ . Segue que  $\mathbb{E}(g(Z_t) \ln(Z_{t+k})) = 0$ . Assim o terceiro termo do lado direito da igualdade (3.64) é igual a zero. Logo, pela expressão (3.61), a expressão (3.64) pode ser reescrita como

$$\begin{aligned}
& \mathbb{E} \left( \left( \sum_{k=0}^{h-2} \lambda_{d,k} g(Z_{n+h-1-k}) + \sum_{k=n+h-1}^{\infty} \lambda_{d,k} g(Z_{n+h-1-k}) + \ln(Z_{n+h}^2) \right)^2 \right) \\
&= \sigma_g^2 \left( \sum_{k=0}^{h-2} \lambda_{d,k}^2 + \sum_{k=n+h-1}^{\infty} \lambda_{d,k}^2 \right) + \mathbb{E} \left( \left( \ln(Z_{n+h}^2) \right)^2 \right), \quad (3.65)
\end{aligned}$$

onde  $\sigma_g^2 = \mathbb{E}([g(Z_t)]^2)$ , para todo  $t \in \mathbb{Z}$ . Pela expressão (3.65), o erro quadrático médio de previsão para  $\ln(X_{n+h}^2)$ , dado na expressão (3.63), é dado por

$$\mathbb{E} \left( \left( \ln(X_{n+h}^2) - \ln(\hat{X}_{n+h}^2) \right)^2 \right) = \sigma_g^2 \left( \sum_{k=0}^{h-2} \lambda_{d,k}^2 + \sum_{k=n+h-1}^{\infty} \lambda_{d,k}^2 \right) + \mathbb{E} \left( \left( \ln(Z_{n+h}^2) \right)^2 \right),$$

onde  $\sigma_g^2 = \mathbb{E}([g(Z_t)]^2)$ , para todo  $t \in \mathbb{Z}$ .

# Capítulo 4

## Medidas de Risco

Neste capítulo apresentamos um contexto probabilístico para a modelagem de riscos financeiros. Apresentamos o método da média-variância para seleção de *portfolios* eficientes e o modelo de precificação de ativos (CAPM), utilizado como ferramenta para análise da correlação dos ativos com o mercado financeiro. Além de definirmos distribuição de lucros e perdas, fatores de risco e funções de risco, apresentamos a definição de medida de risco coerente, bem como as principais medidas de risco utilizadas na literatura.

Em termos financeiros, *risco* é a possibilidade de que um investimento tenha um retorno diferente do esperado, incluindo a possibilidade de perda de uma parte ou até mesmo de todo o investimento original. Um ativo é dito ser *livre de risco* se é não correlacionado com o mercado, para o qual o retorno esperado é sempre igual ao retorno observado. Ou seja, um ativo arriscado é aquele que tem maior chance de perda, enquanto que um ativo mais seguro é aquele em que a probabilidade de perda é menor.

Um *portfolio* de aplicações financeiras é uma coleção de investimentos mantida por uma instituição ou indivíduo. Neste trabalho, sempre que nos referimos a *portfolio* estamos considerando uma carteira de ações.

A questão central em gerenciamento de riscos são as *medidas de risco*, que definimos neste capítulo. As medidas de risco são utilizadas para vários propósitos, entre eles, podemos citar a *determinação do capital em risco*, isto é, mede-se a exposição aos riscos em uma instituição financeira, para determinar a quantidade de capital necessário para honrar possíveis perdas inesperadas no *portfolio*. As medidas de risco também são úteis quando deseja-se limitar os riscos aos quais a instituição financeira está exposta.

Um conceito freqüentemente associado ao risco é a volatilidade. A volatilidade pode ser calculada de diferentes maneiras. No Capítulo 2, definimos a volatilidade como sendo a variância condicional. Entretanto, essa não é a única definição existente. A variância é uma medida de volatilidade absoluta. Existem maneiras de avaliar a volatilidade de forma relativa, ou seja, em relação à volatilidade do mercado, por exemplo. A medida mais utilizada é o coeficiente  $\beta$  do modelo CAPM, que apresentamos na Subseção 4.3.2. Esse coeficiente mede a volatilidade de um ativo frente a volatilidade do mercado. Um ativo mais volátil, portanto, é aquele de varia de forma mais significativa em relação às flutuações do mercado.

## 4.1 Medidas de Risco Coerentes

Nesta seção apresentamos as propriedades que uma medida de risco adequada deve ter. Tais propriedades foram primeiramente propostas por Artzner *et al.* (1999), para aplicações em gerenciamento de riscos financeiros. Baseados no ponto de vista financeiro os autores especificam axiomas que as denominadas “medidas de risco coerentes” devem satisfazer.

### 4.1.1 Axiomas de Coerência

A fim de introduzir os axiomas de coerência, definimos formalmente medidas de risco. Sejam  $(\Omega, \mathcal{F}, \mathbb{P})$  um espaço de probabilidade,  $h$  um horizonte de tempo e  $L^0(\Omega, \mathcal{F}, \mathbb{P})$  o conjunto de todas as variáveis aleatórias sobre  $(\Omega, \mathcal{F})$  que são finitas em quase toda parte.

Riscos financeiros são representados por um conjunto  $\mathcal{M} \subset L^0(\Omega, \mathcal{F}, \mathbb{P})$  de variáveis aleatórias que interpretamos como perdas do *portfolio* sobre o horizonte de tempo  $h$ . Assumimos que  $\mathcal{M}$  é um *cone convexo*, isto é, se  $L_1 \in \mathcal{M}$  e  $L_2 \in \mathcal{M}$ , então  $L_1 + L_2 \in \mathcal{M}$  e  $cL_1 \in \mathcal{M}$ , para todo  $c > 0$ .

**Definição 4.1. (Medida de Risco).** Dado um cone convexo  $\mathcal{M}$  de variáveis aleatórias, tal que  $\mathcal{M} \subset L^0(\Omega, \mathcal{F}, \mathbb{P})$ , uma *medida de risco* com domínio  $\mathcal{M}$  é uma função real  $\varrho : \mathcal{M} \rightarrow \mathbb{R}$ .

Em termos econômicos, interpretamos  $\varrho(L)$  como a quantidade de capital que deve ser adicionada ao *portfolio* com perda  $L$  para que este *portfolio* torne-se aceitável dada alguma restrição.

A seguir, apresentamos um exemplo de uma medida de risco. Na Seção 4.5 definimos formalmente essa medida como sendo o *valor em risco* (VaR) do *portfolio*.

**Exemplo 4.1. (Medida de Risco).** Considere um *portfolio*  $\mathcal{P}$  formado por  $N$  ativos financeiros. Seja  $V_t$  o valor deste *portfolio* em um determinado instante  $t$ . Suponha que um investidor possui o valor necessário para comprar esse *portfolio*. Entretanto, ele deseja reservar uma quantidade de capital adicional  $R$ . Por hipótese, a quantidade  $R$  é destinada a compensar possíveis perdas que possam ocorrer em um horizonte de tempo  $h$  (isto é, esse valor não será utilizado para a compra dos ativos no instante  $t$ ). O objetivo do investidor é reservar um quantidade de capital  $R$  tal que, o valor  $V_{t+h}$  (valor do *portfolio* no instante  $t + h$ ) mais o valor  $R$  seja negativo com uma probabilidade menor ou igual a  $1-p$ . Essa condição pode ser escrita como

$$\mathbb{P}(V_{t+h} + R < 0) \leq 1 - p. \quad (4.1)$$

O capital necessário, no instante  $t$ , é a soma do valor do *portfolio* mais a reserva, isto é,

$$\varrho = V_t + R. \quad (4.2)$$

Pela expressão (4.2),  $R$  pode ser escrito como  $R = \varrho - V_t$ . Segue que,  $V_{t+h} + R = V_{t+h} + \varrho - V_t$ . Além disso,

$$V_{t+h} + \varrho - V_t < 0 \Leftrightarrow -(V_{t+h} - V_t) > \varrho.$$

Substituindo esses resultados na expressão (4.1) obtemos

$$\mathbb{P}(-(V_{t+h} - V_t) > \varrho) \leq 1 - p. \quad (4.3)$$

Como veremos na Seção 4.4, o valor  $-(V_{t+h} - V_t)$  representa a perda  $L_{t+h}$  do *portfólio* no horizonte  $h$ . Observe que, por definição,  $\varrho$  é o p-quantil da função de distribuição da variável aleatória  $L_{t+h}$ . A medida  $\varrho$  é uma função que depende do instante  $t$ , do horizonte  $h$  e da probabilidade de perda  $p$ . Esta medida é definida, na Seção 4.5, como o *valor em risco* ( $\text{VaR}_{p,t+h}$ ) do *portfólio*.

Ressaltamos que, neste trabalho, as variáveis aleatórias  $L \in \mathcal{M}$  representam perdas no *portfólio*. Esta abordagem é a mesma de McNeil *et al.* (2005). Em Artzner *et al.* (1999), as variáveis aleatórias  $L \in \mathcal{M}$  são interpretadas como valores futuros de uma posição atual.

Introduzimos agora os axiomas de coerência para uma medida de risco qualquer.

**Axioma 1 (Invariância por Translação).** Para qualquer  $L \in \mathcal{M}$  e para todo  $c \in \mathbb{R}$ , uma medida de risco *invariante por translação* deve satisfazer

$$\varrho(L + c) = \varrho(L) + c.$$

**Axioma 2 (Subaditividade).** Para quaisquer  $L_1, L_2 \in \mathcal{M}$ , uma medida de risco *subaditiva* deve satisfazer

$$\varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2).$$

**Axioma 3 (Homogeneidade Positiva).** Para qualquer  $L_1 \in \mathcal{M}$  e para todo  $c > 0$ , uma medida de risco *homogênea positiva* deve satisfazer

$$\varrho(cL) = c\varrho(L).$$

**Axioma 4 (Monotonicidade).** Para quaisquer  $L_1, L_2 \in \mathcal{M}$  tais que  $L_1 \leq L_2$  quase certamente, uma medida de risco *monótona* deve satisfazer

$$\varrho(L_1) \leq \varrho(L_2).$$

**Definição 4.2. (Medida de Risco Coerente).** Seja  $\varrho(\cdot)$  uma medida de risco cujo domínio é um cone convexo  $\mathcal{M}$ . Então,  $\varrho(\cdot)$  é uma *medida de risco coerente* (sobre  $\mathcal{M}$ ) se satisfazer os Axiomas 1 – 4.

Apresentamos, a seguir, um exemplo retirado de Frey e McNeil (2002). Este exemplo mostra que a medida de risco  $\varrho$  apresentada no Exemplo 4.1 não é coerente.

**Exemplo 4.2. (Medida de Risco Coerente).** Considere um *portfólio* formado por títulos (ou apólices), para os quais existe possibilidade de *default*. Em termos financeiros, considera-se *default* qualquer falta de pagamento, pagamento impontual, ou repactuações unilaterais. Assuma que os *defaults* de diferentes títulos são independentes e que a probabilidade de *default* é idêntica para todos eles e igual a 2%. O valor nominal dos títulos é 100; esta quantidade é paga em  $T = t + h$  se, e somente se, não houver *default*. O preço atual (tempo  $t$ ) dos títulos é igual a 95. Sendo assim, a perda para o título  $i$  é dada por (veja Definição 4.8)

$$L_{i,t+h} = -(100(1 - Y_i) - 95) = 100Y_i - 5, \quad \text{para cada } i = 1, \dots, 50,$$

onde  $Y_i$  é, por definição, igual a 1, se ocorre *default* e igual a 0, caso contrário. Sendo assim,  $\{L_{i,t+h}\}_{i=1}^{50}$  é uma seqüência de variáveis aleatórias independentes e identicamente distribuídas tais que

$$\mathbb{P}(L_{i,t+h} = -5) = 0,98 \quad \text{e} \quad \mathbb{P}(L_{i,t+h} = 95) = 0,02.$$

Comparamos dois *portfolios*, ambos com valor atual (tempo  $t$ ) igual a 9500 e formados por 100 títulos. O *portfolio* A é fixo e constituído apenas pelos títulos 1. O *portfolio* B é diversificável e é constituído por duas unidades de cada título. Considere  $p=95\%$  e a medida  $\varrho$ , definida no Exemplo 4.1. Lembramos que,  $\varrho$  é o  $p$ -quantil da função de distribuição de  $L_{t+h}$ , isto é,

$$\varrho(L_{t+h}) = \inf\{x \in \mathbb{R}, \mathbb{P}(L_{1,t+h} > x) \leq 1 - p\} = \inf\{x \in \mathbb{R}, \mathbb{P}(L_{1,t+h} \leq x) \geq p\}.$$

Para o *portfolio* A a perda é dada por  $L_{A,t+h} = 100L_{1,t+h}$ . Temos que

$$\mathbb{P}(L_{1,t+h} \leq -5) = \mathbb{P}(L_{1,t+h} = -5) = 0,98 \implies \mathbb{P}(100L_{1,t+h} = -500) = 0,98$$

Logo,

$$\varrho(L_{A,t+h}) = 100\varrho(L_{1,t+h}) = -500.$$

Isso significa que, mesmo com a retirada de 500, o *portfolio* ainda é aceitável, ao nível de confiança de 95%.

Para o *portfolio* B a perda é dada por

$$L_{B,t+h} = 2 \sum_{i=1}^{50} L_{i,t+h} = 2 \sum_{i=1}^{50} (100Y_i - 5) = 200 \sum_{i=1}^{50} Y_i - 500.$$

Segue que  $\varrho(L_{B,t+h}) = 200q_p(\sum_{i=1}^{50} Y_i) - 500$ , onde  $q_p(\cdot)$  é a função quantil. A soma  $M \equiv \sum_{i=1}^{50} Y_i$  tem função de distribuição binomial com probabilidade de sucesso igual a 0,02. Segue que  $q_{0,95}(M) = 3$  (veja Frey e McNeil, 2002). Segue que

$$\varrho(L_{B,t+h}) = 100.$$

Neste caso, o banco precisaria um capital igual a 100 para satisfazer as exigências necessárias ao nível de 95% de confiança.

Segundo Frey e McNeil (2002), uma instituição financeira diria que o *portfolio* B apresenta menos risco que o *portfolio* A. Os autores apresentam um lema que prova que essa afirmação é verdadeira quando utiliza-se uma medida de risco coerente. Isso ocorre pois, para uma medida de risco coerente  $\varrho(\cdot)$ , a desigualdade

$$\varrho\left(\sum_{i=1}^{50} 2L_{i,t+h}\right) \leq \sum_{i=1}^{50} \varrho(L_{i,t+h}) = 100\varrho(L_{1,t+h}) = \varrho(100L_{1,t+h})$$

é sempre satisfeita.

## 4.2 Retornos

Dado que, em um *portfolio*, o risco é freqüentemente medido em termos de variações de preços, apresentamos nesta seção as principais definições relacionadas a variação de preços de ativos financeiros.

**Definição 4.3. (Retorno Líquido Simples).** Seja  $P_t$  o preço de um ativo no instante  $t$ . A variação de preços do ativo entre os instantes  $t - 1$  e  $t$  é dada por  $\Delta P_t = P_t - P_{t-1}$



e a variação relativa de preços ou *retorno líquido simples* deste ativo entre os mesmos instantes é definido por

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

Note que,  $R_t = \frac{P_t}{P_{t-1}} - 1$ . Chamamos  $1 + R_t = \frac{P_t}{P_{t-1}}$  de *retorno bruto simples*. Em geral, expressamos  $R_t$  em termos de porcentagem, relativamente ao período (um dia, um mês, um ano, etc). A quantidade  $R_t$  é também conhecida como *taxa de retorno*.

**Definição 4.4. (Log-retorno).** Definimos o *retorno composto continuamente* ou simplesmente *log-retorno* como

$$r_t = \ln(1 + R_t) = \ln(P_t) - \ln(P_{t-1}). \quad (4.4)$$

**Definição 4.5. (Retorno Simples de Período h).** O *retorno simples de período h*, entre os instantes  $t - k$  e  $t$  é dado por

$$R_t[h] = \frac{P_t - P_{t-h}}{P_{t-h}} = \frac{P_t}{P_{t-h}} - 1. \quad (4.5)$$

A partir da Definição 4.5 obtemos o *log-retorno de período h* que é então definido por

$$r_t[h] = \ln(1 + R_t[h]) = \ln\left(\frac{P_t}{P_{t-h}}\right) = \sum_{j=0}^{h-1} r_{t-j}, \quad (4.6)$$

onde  $r_t$  é dado pela expressão (4.4) e o log-retorno de período 1 é  $r_t$ .

**Definição 4.6. (Retorno Anualizado).** O *retorno anualizado* é definido como

$$R_t[h]\text{anualizado} = \left[ \prod_{j=0}^{h-1} (1 + R_{t-j}) \right]^{1/h} - 1. \quad (4.7)$$

É fácil ver que a expressão (4.7) pode ser aproximada por  $(1/h) \sum_{j=0}^{h-1} R_{t-j}$ , utilizando a expansão de Taylor até primeira ordem. A expressão apresentada na equação (4.6) é denominada *agregação temporal* dos retornos. No caso de um *portfolio* podemos ter também uma agregação denominada “cross-section”.

Considere um *portfolio*  $\mathcal{P} = \{A_1, \dots, A_N\}$  formado por  $N$  ativos com pesos  $\mathbf{a} = (a_1, \dots, a_N)'$ , isto é, para todo  $i = 1, \dots, N$ ,  $a_i$  representa a participação porcentual do ativo  $A_i$  em relação ao total do *portfolio*, de forma que  $\sum_{i=1}^N a_i = 1$ . Sejam  $P_{i,t}$  e  $c_i$ , respectivamente, o preço e o número de unidades do ativo  $A_i$ , para todo  $i = 1, \dots, N$ , no instante  $t$ . É fácil ver que se  $P_t$  é o valor do *portfolio* no instante  $t$ , devemos ter

$$c_i P_{i,t} = a_i P_t, \quad \text{para todo } i = 1, \dots, N. \quad (4.8)$$

Logo, pela expressão (4.8), segue que

$$a_i = \frac{c_i P_{i,t}}{P_t}, \quad \text{para todo } i = 1, \dots, N \text{ e para todo } t \in \mathbb{Z}. \quad (4.9)$$

Sob essas condições, o valor do *portfolio* no instante  $t$  é dado por

$$P_t = \sum_{i=1}^N c_i P_{i,t}, \quad \text{para todo } t \in \mathbb{Z}. \quad (4.10)$$

Pelas expressões (4.9) e (4.10), se  $P_0$  é o valor inicial do *portfolio*, após um período de tempo, o retorno simples do *portfolio* é dado por

$$R_{\mathcal{P}} = \frac{P_1 - P_0}{P_0} = \frac{1}{P_0} \sum_{i=1}^N c_i (P_{i,1} - P_{i,0}) = \sum_{i=1}^N c_i \frac{P_{i,0}}{P_0} \left( \frac{P_{i,1} - P_{i,0}}{P_{i,0}} \right) = \sum_{i=1}^N a_i R_i, \quad (4.11)$$

onde denotamos por  $R_i$  o retorno simples do ativo  $A_i$ , para  $i = 1, \dots, N$ , após um período de tempo. Seja  $r_i$  o log-retorno do ativo  $A_i$ , para todo  $i = 1, \dots, N$ , após um período de tempo. Segue, da expressão (4.11) que o log-retorno do *portfolio* após um período de tempo é dado por

$$r_{\mathcal{P}} = \ln \left( \frac{P_1}{P_0} \right) = \ln(1 + R_{\mathcal{P}}) = \ln \left( 1 + \sum_{i=1}^N a_i R_i \right). \quad (4.12)$$

Mas,  $\sum_{i=1}^N a_i = 1$  e  $1 + R_i = e^{r_i}$ , para todo  $i = 1, \dots, N$ . Segue, pela expressão (4.12), que

$$r_{\mathcal{P}} = \ln \left( \sum_{i=1}^N a_i + \sum_{i=1}^N a_i R_i \right) = \ln \left( \sum_{i=1}^N a_i e^{r_i} \right).$$

A próxima definição resume os tipos de agregações possíveis.

**Definição 4.7. (Agregação).** Considere um *portfolio*  $\mathcal{P} = \{A_1, \dots, A_N\}$ , formado por  $N$  ativos com pesos  $\mathbf{a} = (a_1, \dots, a_N)'$ , tais que  $\sum_{i=1}^N a_i = 1$ . Sejam  $R_{i,t}$  o retorno simples no instante  $t$  e  $r_{i,t}$  o log-retorno no instante  $t$ , do ativo  $A_i$ , para todo  $i = 1, \dots, N$ .

(i) *Agregação temporal* para os ativos  $A_i$ , para todo  $i = 1, \dots, N$ , é dada por

$$R_{i,t}[h] = \prod_{j=0}^{h-1} (1 + R_{i,t-j}) - 1 \quad \text{e} \quad r_{i,t}[h] = \sum_{j=0}^{h-1} r_{i,t-j}, \quad (4.13)$$

para todo  $h \geq 1$ , onde  $h$  representa o período.

(ii) *Agregação "cross-section"* para o *portfolio*, no tempo  $t$  é dada por

$$R_{\mathcal{P},t} = \sum_{i=1}^N a_i R_{i,t} \quad \text{e} \quad r_{\mathcal{P},t} = \ln \left( \sum_{i=1}^N a_i e^{r_{i,t}} \right), \quad (4.14)$$

onde  $R_{i,t}$  e  $r_{i,t}$  são, respectivamente, o retorno simples e o log-retorno do ativo  $A_i$  no instante  $t$ , para todo  $i = 1, \dots, N$ .

Na prática utiliza-se a aproximação  $r_{\mathcal{P},t} \approx \sum_{i=1}^N a_i r_{i,t}$ .

## 4.2.1 Fatos Estilizados sobre os Retornos

Séries temporais financeiras apresentam características que são comuns a outras séries temporais, tais como tendência, sazonalidade, observações influentes (ou atípicas), heteroscedasticidade condicional e não-linearidade. De modo geral, a *não-linearidade* das séries é caracterizada pelo fato de séries financeiras responderem de maneira diferente a grandes ou pequenos choques, ou ainda, a choques negativos ou positivos.

No entanto, séries de retornos financeiros raramente apresentam tendência ou sazonalidades (com exceção eventualmente de retornos intra-diários) e possuem ainda algumas características que não são comuns a outras séries temporais tais como:

1. Em geral, são não-autocorrelacionados.
2. Os quadrados dos retornos são autocorrelacionados, apresentando uma correlação para  $h = 1$  pequena e depois uma queda lenta para os demais valores de  $h$ .
3. Séries de retornos apresentam agrupamento de volatilidades ao longo do tempo.
4. A função de distribuição (não-condicional) dos retornos apresenta caudas mais pesadas do que uma função de distribuição normal. Além disso, a distribuição é em geral leptocúrtica, isto é, é mais alta e alongada, logo possui caudas pesadas, que a distribuição normal, embora aproximadamente simétrica.
5. Algumas séries de retornos são não-lineares, no sentido discutido no primeiro parágrafo desta subseção.

O exemplo a seguir apresenta o gráfico de uma série temporal financeira e o gráfico dos retornos, onde podemos observar algumas das características aqui citadas.

**Exemplo 4.3. (Fatos Estilizados dos Retornos).** Considere a série temporal composta por  $n = 1.729$  observações que representa os índices diários da Bolsa de Valores de São Paulo no período de janeiro de 1995 a dezembro de 2001, apresentada na Figura 1.1 (no Capítulo 1). Neste exemplo apresentamos a sua correspondente série de retornos no período citado.

Na Figura 4.1 percebemos que o histograma da série temporal dos log-retornos tem o formato de uma curva leptocúrtica, apresentando aparente simetria. Observe que a série de retornos, apresentada na Figura 4.2, apresenta os fatos estilizados citados anteriormente, como aparente estacionariedade, média em torno de zero e agrupamentos de volatilidades.

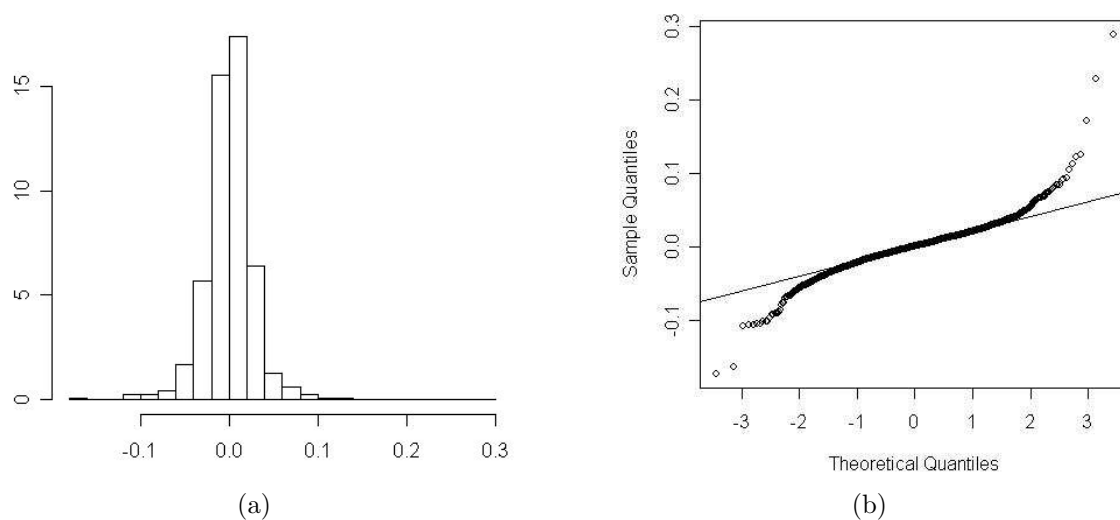


Figura 4.1: (a) Histograma e (b)  $Q \times Q$  Plot dos Log-retornos Diários do IBovespa

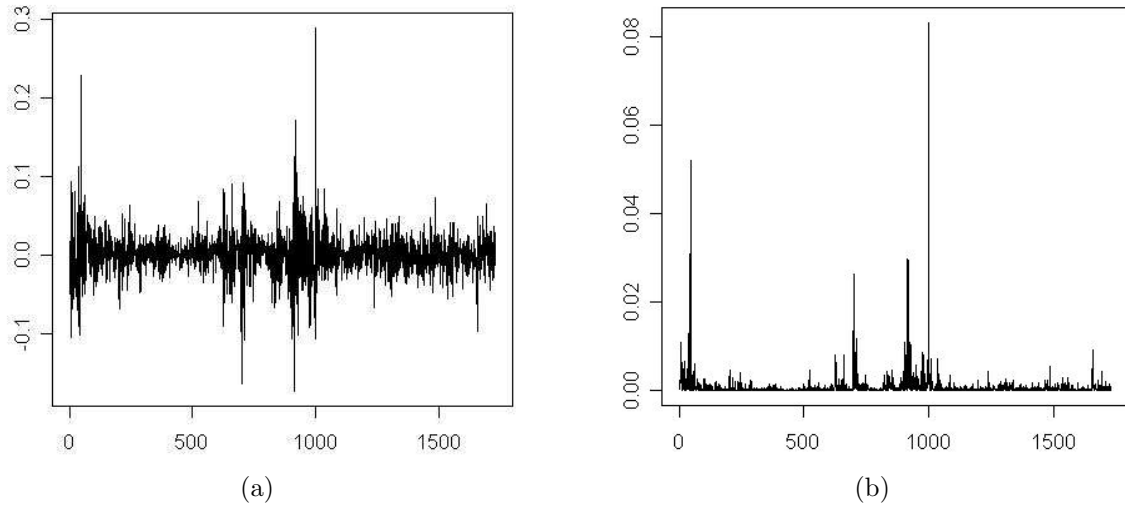


Figura 4.2: Série dos (a) Log-retornos Diários e (b) Quadrados dos Log-retornos Diários do IBovespa.

Na Figura 4.3 observamos que, enquanto a série de retornos é pouco correlacionada, a série do quadrado dos retornos apresenta valores altos da função de autocorrelação para vários valores de  $h$ .

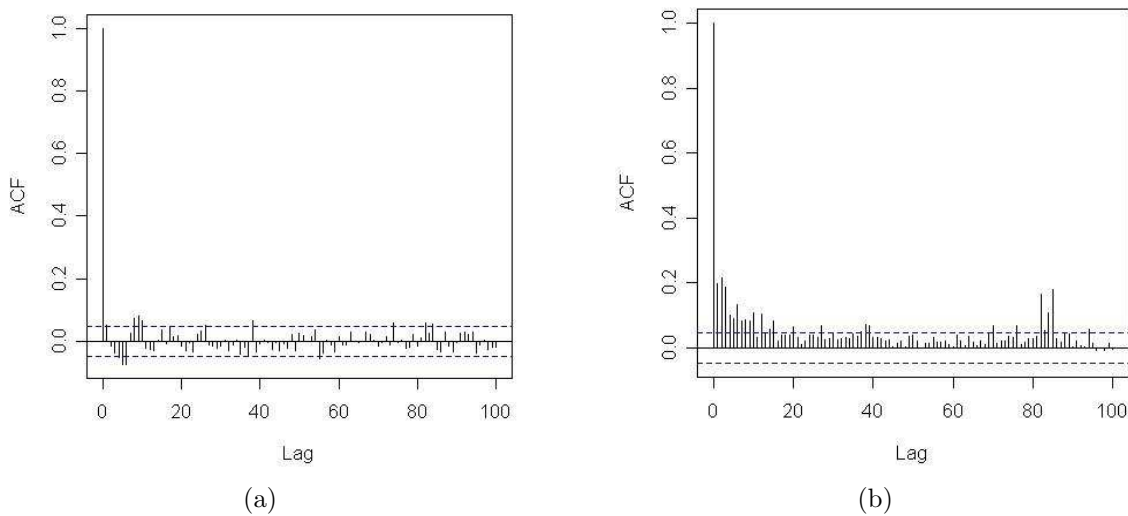


Figura 4.3: Função de Autocorrelação Amostral (a) dos Log-retornos Diários e (b) do Quadrado dos Log-retornos Diários do IBovespa.

### 4.3 Seleção de *Portfolios* Eficientes

Nesta seção tratamos da seleção de *portfolios* eficientes baseado no método da análise de média-variância (*M-V analysis*), introduzido por Markowitz (1952). Dizemos que um *portfolio* é *eficiente* se ele contém uma combinação de ativos que proporciona o máximo retorno para uma dada classe de risco, ou de forma equivalente, o mínimo risco para uma dada classe de retornos.

A maioria dos investidores espera que um *portfolio* eficiente satisfaça ambas as condições, máximo retorno e mínima variância. Segundo Markowitz (1952), o *portfolio* com

máximo valor esperado não é, necessariamente, o de mínima variância. Existe uma proporção tal que o investidor pode aumentar os retornos esperados aumentando a variância ou reduzir a variância diminuindo os retornos esperados. Em seu trabalho, o autor apresenta um estudo sobre a seleção de *portfolios* eficientes baseado no valor esperado e na variância dos retornos (*E-V rule*). De acordo com os conceitos desenvolvidos, o investidor pode reduzir o risco total de seus investimentos através da combinação de ativos que não apresentam correlação perfeita e positiva. E a redução dos riscos será tão maior quanto menor for esta correlação.

Considere um *portfolio*  $\mathcal{P} = \{A_1, \dots, A_N\}$  composto por  $N$  ativos com pesos  $\mathbf{a} = (a_1, \dots, a_N)'$ . Seja  $\mathbf{R} = (R_1, \dots, R_N)'$  o vetor formado pelo retorno  $R_i$  do ativo  $A_i$ , para todo  $i = 1, \dots, N$ . Segue que o retorno do *portfolio* para um período de tempo é dado pela expressão (4.11). Assim, o valor esperado do retorno do *portfolio* é dado por

$$\mathbb{E}(R_{\mathcal{P}}) = \sum_{i=1}^N a_i \mathbb{E}(R_i). \quad (4.15)$$

E a variância do retorno do *portfolio* é dada por

$$\text{Var}(R_{\mathcal{P}}) = \sum_{i=1}^N \sum_{k=1}^N a_j a_k \text{Cov}(R_i, R_k). \quad (4.16)$$

Denotamos por  $R_M$  o retorno do mercado e  $a_{iM}$  é a proporção do ativo  $A_i$  no *portfolio* do mercado, para  $i = 1, \dots, N$ . Substituindo-se  $R_{\mathcal{P}}$  por  $R_M$  e  $a_j$  por  $a_{iM}$ , para  $i = 1, \dots, N$ , na expressão (4.16) (veja Fabozzi *et al.*, 2006) obtemos a variância do retorno do mercado dada por

$$\text{Var}(R_M) = \sum_{i=1}^N \sum_{k=1}^N a_{iM} a_{kM} \text{Cov}(R_i, R_k). \quad (4.17)$$

Observando que a covariância do retorno  $R_i$ , para todo  $i = 1, \dots, N$ , com o retorno do *portfolio* do mercado é dada por

$$\text{Cov}(R_i, R_M) = \sum_{k=1}^N a_{kM} \text{Cov}(R_i, R_k), \quad (4.18)$$

e substituindo a expressão (4.18) na expressão (4.17), obtemos

$$\begin{aligned} \text{Var}(R_M) &= \sum_{i=1}^N a_{iM} \left( \sum_{k=1}^N a_{kM} \text{Cov}(R_i, R_k) \right) \\ &= \sum_{i=1}^N a_{iM} \text{Cov}(R_i, R_M). \end{aligned}$$

Sharpe (1964) define o valor  $\text{Cov}(R_i, R_M)$  como *risco sistemático* do ativo  $A_i$ , para todo  $i = 1, \dots, N$ . O risco sistemático é aquele que resulta das condições econômicas e do mercado e que não pode ser eliminado através da diversificação. Assim, segundo Sharpe (1964) o *risco total* de um *portfolio* pode ser subdividido em dois tipos:

- **risco diversificável:** afeta um, ou um número pequeno de ativos. Em geral, ele independe do mercado e pode ser reduzido ou eliminado através da diversificação,

com o aumento da quantidade de diferentes ativos. Um exemplo de risco diversificável é um problema de produção ou logística. Isso pode afetar a ação de uma empresa, mas o efeito disso será muito pequeno sobre os papéis de outras empresas;

- **risco sistemático:** é a parcela do risco que não pode ser eliminada pela diversificação e está relacionada com o comportamento do mercado como um todo. Como exemplos de risco sistemático podemos citar: aumento nas taxas de juros, crises políticas, mudanças de cenários internacionais, entre outros.

Sendo assim, o objeto de estudo passa a ser o risco sistemático, que é a componente do risco relacionada com o mercado. Neste caso, a medida de risco para uma ação passa a ser a covariância dos retornos desse ativo com os retornos do mercado. Com base nestes conceitos e naqueles desenvolvidos por Markowitz (1952), os trabalhos de Sharpe (1964) e de Lintner (1965) apresentam o modelo de precificação de ativos, denotado por CAPM (*Capital Asset Pricing Model*), que definimos na Seção 4.3.2. Este é um dos modelos mais utilizados atualmente para a avaliação de ativos financeiros.

### 4.3.1 Linha de Mercado de Capitais - CML

Sharpe (1964) e Lintner (1965) demonstram que o conjunto de *portfolios* eficientes disponíveis para o investidor que emprega a análise de média-variância, na ausência de um ativo livre de risco, é inferior ao disponível quando existe ativo desse tipo. Apresentamos, nesta subseção esta formulação. As idéias básicas do método clássico de análise de média-variância são apresentadas no Apêndice D, Seção D.1.

Assuma que existe um ativo livre de risco, denotado por  $F$ , onde o investidor pode emprestar e tomar emprestado a uma única taxa, denotada por  $R_F$ . O investidor precisa escolher uma combinação de  $N$  ativos de risco  $A_1, \dots, A_N$  mais o ativo livre de risco. Neste caso o *portfolio* será dado por  $\mathcal{P} = \{A_1, \dots, A_N, F\}$ . Seja  $\mathbf{a} = (a_1, \dots, a_N)'$ , o vetor de pesos dos ativos  $A_i$ , para  $i = 1, \dots, N$ . Estes pesos não devem somar 1, para que a percentagem restante possa ser investida no ativo livre de risco. Se considerarmos ambos os casos, compra e venda do ativo livre de risco (ver Fabozzi *et al.*, 2006), a proporção  $1 - \sum_{i=1}^N a_i$  pode ser tanto positiva quanto negativa. O retorno esperado  $\mathbb{E}(R_{\mathcal{P}})$  e a variância do retorno do *portfolio*  $\text{Var}(R_{\mathcal{P}})$  são dados, respectivamente, pelas expressões (4.15) e (4.16). Como  $\mathbb{E}(R_F) = R_F$ , pois o ativo é livre de risco, segue que

$$\mathbb{E}(R_{\mathcal{P}}) = \sum_{i=1}^N a_i \mathbb{E}(R_i) + \left(1 - \sum_{i=1}^N a_i\right) R_F \quad \text{e} \quad \text{Var}(R_{\mathcal{P}}) = \sum_{i=1}^N \sum_{k=1}^N a_i a_k \text{Cov}(R_i, R_k), \quad (4.19)$$

pois o ativo livre de risco tem variância zero e é não-correlacionado com os ativos de risco.

O objetivo do investidor é, agora, escolher as alocações (pesos), de forma a atingir um determinado nível de retorno esperado  $\mu_0$ . O vetor de pesos é estimado resolvendo-se o problema de otimização

$$\min_{\mathbf{a}} \left\{ \sum_{i=1}^N \sum_{k=1}^N a_i a_k \text{Cov}(R_i, R_k) \right\},$$

sujeito à condição

$$\mu_0 = \sum_{i=1}^N a_i \mathbb{E}(R_i) + \left(1 - \sum_{i=1}^N a_i\right) R_F.$$

Segundo Fabozzi *et al.* (2006) os pesos ótimos para o *portfolio* são dados por

$$\mathbf{a} = C \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_F \mathbf{1}), \quad (4.20)$$

onde  $\boldsymbol{\Sigma}$  é a matriz de variância-covariância dos retornos,  $\boldsymbol{\mu} = (\mathbb{E}(R_1), \dots, \mathbb{E}(R_N))'$ ,  $\mathbf{1}$  é o vetor coluna com  $N$  entradas iguais a 1 e  $C$  é dado pela expressão

$$C = \frac{\mu_0 - R_F}{(\boldsymbol{\mu} - R_F \mathbf{1})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - R_F \mathbf{1})}. \quad (4.21)$$

A expressão (4.20) mostra que todos os pesos para ativos de risco para um *portfolio* de mínima variância são proporcionais ao vetor  $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_F \mathbf{1})$ , com constante de proporção  $C$ , dada pela expressão (4.21). Portanto, quando consideramos um ativo livre de risco, todos os *portfolios* de mínima variância são combinações lineares de um ativo livre de risco com um *portfolio* de risco. Esta combinação é representada por uma reta denominada *Linha de Mercado de Capitais*, ou simplesmente, CML. O *portfolio* consistindo apenas dos ativos de risco é denominado *portfolio tangente*. O *portfolio* tangente é freqüentemente denominado *portfolio do mercado*, ou simplesmente, *mercado*.

Para o *portfolio* do mercado o vetor de pesos  $\mathbf{a}_M$  é dado por

$$\mathbf{a}_M = \frac{1}{\mathbf{1}' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - R_F \mathbf{1})} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - R_F \mathbf{1}). \quad (4.22)$$

Pode-se mostrar que  $\mathbf{a}_M$  pode ser obtido pela otimização da expressão

$$\max_{\mathbf{a}} \left\{ \frac{\mathbf{a}' \boldsymbol{\mu} - R_F}{\sqrt{\mathbf{a}' \boldsymbol{\Sigma} \mathbf{a}}} \right\},$$

com a condição  $\mathbf{a} \mathbf{1} = 1$ . Para um estudo mais detalhado veja Fabozzi *et al.* (2006).

### Como Obter a Linha de Mercado de Capitais (CML)

Suponha que todos os investidores possam criar um *portfolio* eficiente dado por  $\mathcal{P} = \{A_1, \dots, A_N, F\}$  com pesos  $a_F$ , para o ativo livre de risco, e  $a_M = \sum_{i=1}^N a_i$ , para os ativos com risco. Então,  $a_F + a_M = 1$ . Segue, da expressão (4.19), que o retorno esperado do *portfolio* é dado por

$$\mathbb{E}(R_{\mathcal{P}}) = a_F R_F + a_M \mathbb{E}(R_M).$$

Desde que  $a_F = 1 - a_M$ , temos

$$\mathbb{E}(R_{\mathcal{P}}) = (1 - a_M) R_F + a_M \mathbb{E}(R_M) = R_F + a_M [\mathbb{E}(R_M) - R_F]. \quad (4.23)$$

Como o retorno do ativo livre de risco e o retorno do *portfolio* de mercado são não-correlacionados e a variância do ativo livre de riscos é zero, a variância do retorno do *portfolio* é dada por

$$\text{Var}(R_{\mathcal{P}}) = a_M^2 \text{Var}(R_M) \quad \implies \quad a_M = \sqrt{\frac{\text{Var}(R_{\mathcal{P}})}{\text{Var}(R_M)}}. \quad (4.24)$$

Substituindo a expressão (4.24) na expressão (4.23), obtemos a expressão para a *linha de mercado de capitais* (CML), dada por

$$\mathbb{E}(R_{\mathcal{P}}) = R_F + \left[ \frac{\mathbb{E}(R_M) - R_F}{\sqrt{\text{Var}(R_M)}} \right] \sqrt{\text{Var}(R_{\mathcal{P}})}. \quad (4.25)$$

O valor

$$\frac{\mathbb{E}(R_M) - R_F}{\sqrt{\text{Var}(R_M)}}$$

é conhecido como *prêmio do risco*. O prêmio de risco corresponde, basicamente, a uma previsão dos rendimentos que os investidores irão exigir no futuro para aplicarem o seu capital no mercado de ações. Como o retorno do *portfolio* de mercado é incerto, há um prêmio para o investidor por continuar investindo nele, ao invés de investir em ativos sem risco.

Apresentamos no Apêndice D, Seção D.2, um exemplo do cálculo dos pesos de um *portfolio* eficiente para dois ativos de risco.

### 4.3.2 Modelo de Precificação de Ativos - CAPM

O modelo CAPM, desenvolvido por Sharpe (1964) e Lintner (1965), baseia-se nas seguintes hipóteses (veja Fabozzi *et al.*, 2006):

- inexistência de custos de transação (mercado perfeito);
- ausência de informação confidencial, o que faz com que não haja ativos subavaliados ou superavaliados no mercado;
- a decisão do investidor baseia-se unicamente no retorno esperado e no risco;
- os investidores estimam o risco em função da variabilidade das taxas de retorno estimadas;
- existência de um ativo livre de risco, denotado por  $F$ , onde os investidores podem emprestar e tomar emprestado a uma única taxa, denotada por  $R_F$ , visando obter alocações ótimas;
- os investidores ajustam a decisão de alocação às preferências de risco, decidindo dessa forma, quando investirão em ativos livres de risco ou ativos arriscados.

Suponha que o mercado esteja em equilíbrio. Lembramos que a situação é representada pela linha de capital de mercado (CML), dada pela expressão (4.25). Sob essas condições, o retorno esperado de um *portfolio* é uma função linear do valor esperado do *portfolio* do mercado e ativos individuais não aparecem sobre a CML. Segundo Fabozzi *et al.* (2006), pode-se demonstrar que, para ativos individuais temos a seguinte relação

$$\mathbb{E}(R_i) = R_F + \left[ \frac{\mathbb{E}(R_M) - R_F}{\text{Var}(R_M)} \right] \text{Cov}(R_i, R_M), \quad \text{para todo } i = 1, \dots, N. \quad (4.26)$$

A equação (4.26) é denominada *Linha de Mercado de Títulos (Security Market Line-SML)*. Em situações de equilíbrio, o retorno esperado de ativos individuais deve estar sobre a linha SML e não sobre a linha CML.

A versão empírica da equação (4.26) é a regressão linear, denominada *linha característica*, dada por

$$R_{i,t} - R_{F,t} = \alpha_i + \beta_i (R_{M,t} - R_{F,t}) + \varepsilon_{i,t}, \quad \text{para todo } i = 1, \dots, N \text{ e } t \in 1, \dots, n, \quad (4.27)$$



onde  $n$  é o tamanho amostral,  $\varepsilon_{i,t}$  é o erro (representa o risco não-sistemático),  $\hat{\beta}_i$  é o estimador da expressão

$$\beta_i \equiv \frac{\text{Cov}(R_i, R_M)}{\text{Var}(R_M)}, \quad \text{para todo } i = 1, \dots, N, \quad (4.28)$$

que representa o risco sistemático e  $\hat{\alpha}_i$  é o estimador do  $\alpha$  da ação, definido por,

$$\alpha_i = \bar{R}_i - \mathbb{E}(R_i), \quad \text{para todo } i = 1, \dots, N. \quad (4.29)$$

Substituindo-se  $\beta_i$ , dado pela expressão (4.28), na expressão (4.26) obtemos o modelo CAPM, dado por

$$\mathbb{E}(R_i) = R_F + \beta_i(\mathbb{E}(R_M) - R_F), \quad \text{para todo } i = 1, \dots, N. \quad (4.30)$$

A razão (4.28) também pode ser estimada através da covariância e da variância amostrais das séries de retorno do *portfolio* do mercado e do ativo.

**Observação 4.1.** O coeficiente  $\beta$  do *portfolio*  $\mathcal{P} = \{A_1, \dots, A_N\}$ , denotado por  $\beta_{\mathcal{P}}$ , é dado pela expressão

$$\beta_{\mathcal{P}} = \sum_{i=1}^N a_i \beta_i, \quad (4.31)$$

onde  $\beta_i$ , para  $i = 1, \dots, N$ , é o coeficiente  $\beta$  do ativo  $A_i$ , definido pela expressão (4.28).

### 4.3.3 Interpretação do Coeficiente $\beta$

O coeficiente angular  $\beta_i$ , dado na expressão (4.30), pode ser visto como uma medida de volatilidade das taxas de retorno de um ativo qualquer com relação às taxas de retorno do mercado como um todo. Esse pressuposto parte do princípio de que todos os ativos tendem a ter os seus preços alterados com maior ou menor proporção às alterações do mercado como um todo. O coeficiente angular  $\beta_i$ , para todo  $i = 1, \dots, N$ , dado na expressão (4.30), pode ser interpretado como segue:

- $\beta_{R_M} = 1$ . De fato, a covariância do mercado é igual a variância do mercado, ou seja, uma carteira tem  $\beta$  igual a 1 em relação a ela mesma.
- $\beta_i = 1$ . O ativo  $A_i$  pode ser considerado *neutro*. Por exemplo: à medida que os retornos do mercado como um todo sobem 5% (índice Bovespa), os retornos do ativo  $A_i$  tendem a subir 5%. À medida que os retornos do mercado como um todo (índice Bovespa) caem 2%, os retornos do ativo  $A_i$  tendem a cair 2%.
- $\beta_i > 1$ . O ativo  $A_i$  pode ser visto como *agressivo*. Por exemplo: uma ação com  $\beta_i = 1,15$  significa que se o mercado (índice Bovespa) como um todo apresentar uma queda de 10%, o ativo  $A_i$  deverá sofrer uma baixa de 11,50% no seu preço.
- $\beta_i < 1$ . O ativo  $A_i$  pode ser classificado como *defensivo*. Por exemplo: uma ação com  $\beta_i = 0,5$  significa que se o mercado (índice Bovespa) cair 6%, o retorno do ativo  $A_i$  deverá sofrer uma baixa de somente 3%.

Lembramos que o retorno esperado depende apenas do risco sistemático. Como ativos com valores de  $\beta$  maiores têm riscos sistemáticos mais altos, têm também retornos esperados maiores. Dessa forma, conhecendo-se as características de risco (seu respectivo  $\beta_i$ ) de uma ação, é possível estimar-se o preço justo (ou valor intrínseco), tendo-se a indicação se o título é, ou não, uma boa opção de compra.

### 4.3.4 Ações Subavaliadas e Superavaliadas. Coeficiente $\alpha$ da Ação

Os investidores sempre têm o interesse em saber se no mercado existem ações cujo valor intrínseco não coincide com o seu preço de mercado. Tais ações podem estar *subavaliadas*, isto é, com preços inferiores aos seus respectivos valores, significando que seu retorno esperado verdadeiro é maior do que aquele fixado pelo mercado em geral. Ou podem estar *superavaliadas*, isto é, cotadas a preços superiores aos seus respectivos valores intrínsecos.

A diferença entre o retorno esperado pelo investidor  $\bar{R}_i$  (média dos valores históricos do ativo  $A_i$ ) e o retorno esperado  $\mathbb{E}(R_i)$  calculado pelo modelo CAPM é chamado  $\alpha$  de uma ação dado pela expressão (4.29).

O coeficiente  $\alpha_i$  é obtido através da regressão linear dada pela expressão (4.27) e pode ser interpretado como segue:

- $\alpha_i > 0$ . O ativo  $A_i$  está subavaliado, isto é, o preço de suas ações é menor do que seu valor intrínseco. É conveniente comprá-lo comparativamente a outras ações com um mesmo nível de risco em situação de equilíbrio.
- $\alpha_i < 0$ . O ativo  $A_i$  está superavaliado na visão do investidor, portanto é conveniente vendê-lo.

No Apêndice D, Seção D.3, apresentamos um exemplo do cálculo dos coeficientes  $\alpha$  e  $\beta$  dos ativos de um *portfolio*. Na Seção D.4 apresentamos um exemplo do cálculo do  $\beta$  do *portfolio*.

## 4.4 Fatores de Risco e Distribuição de Perdas

Dado um *portfolio* fixo  $\mathcal{P} = \{A_1, \dots, A_N\}$ , denotamos por  $V_t$  o *valor do portfolio* no tempo  $t \in T$ , onde  $T$  é um conjunto de índices. Em geral, neste trabalho, consideramos  $T = \mathbb{N}$ . Assumimos que  $V_t$ , para todo  $t \in \mathbb{N}$ , são variáveis aleatórias observáveis.

**Definição 4.8. (Distribuição de Perdas).** Considere um *portfolio*  $\mathcal{P} = \{A_1, \dots, A_N\}$  fixo. Seja  $V_t$  o seu valor no tempo  $t$ . Dado um horizonte  $h$ , a *perda do portfolio*  $\mathcal{P}$  sobre este período é dada por

$$L_{t+h} \equiv -(V_{t+h} - V_t). \quad (4.32)$$

A distribuição da variável aleatória  $L_{t+h}$  é denominada *distribuição de perdas*.

**Observação 4.2.** Em gerenciamento de riscos, o principal interesse é analisar a probabilidade de grandes perdas, ou seja, analisar a cauda direita da distribuição de perdas. Na literatura também encontramos a expressão *distribuição de perdas e ganhos*. A distribuição de perdas e ganhos é a distribuição das variáveis aleatórias  $-L_{t+h}$ . Em muitos casos os autores utilizam a mesma notação para a distribuição de perdas e para a distribuição de perdas e ganhos.

Para todo  $t \in \mathbb{N}$ , a variável aleatória  $V_t$  é uma função de um vetor aleatório,  $m$ -dimensional,  $\mathbf{Z}_t = (Z_{1,t}, \dots, Z_{m,t})'$ , de *fatores de risco*, isto é,

$$V_t = f(t, \mathbf{Z}_t), \quad \text{para todo } t \in \mathbb{N}, \quad (4.33)$$

para alguma função mensurável  $f : \mathbb{R} \times \mathbb{R}^m \longrightarrow \mathbb{R}$ . A função  $f(\cdot)$ , dada pela expressão (4.33), é denominada *função dos riscos*.

Assumimos que os fatores de risco são observáveis. Portanto,  $\mathbf{Z}_t$  é observável no tempo  $t$ . A escolha dos fatores de risco e da função  $f(\cdot)$  depende tanto do *portfolio* quanto do nível de precisão desejados. Na prática, os fatores de risco mais utilizados são o logaritmo dos preços dos ativos financeiros, retornos e o logaritmo da taxa de câmbio.

Em alguns casos, é conveniente definir a série de mudança de fatores de risco  $\{\mathbf{X}_t\}_{t \in \mathbb{N}}$  da maneira como segue.

**Definição 4.9. (Mudanças de Fatores de Risco).** Seja  $\mathbf{Z}_t = (Z_{1,t}, \dots, Z_{m,t})'$  um vetor aleatório  $m$ -dimensional de fatores de risco. As variáveis aleatórias  $\mathbf{X}_t$ , para todo  $t \in \mathbb{N}$ , definidas por

$$\mathbf{X}_t \equiv \mathbf{Z}_t - \mathbf{Z}_{t-1}, \quad \text{para todo } t \in \mathbb{N} \quad (4.34)$$

são denominadas *mudanças de fatores de risco*.

Por exemplo, se  $Z_{i,t} = \ln(P_{i,t})$ , onde  $P_{i,t}$  é o preço do ativo  $A_i$  no tempo  $t$ , para cada  $i = 1, \dots, m$ , então  $\mathbf{X}_t = (X_{1,t}, \dots, X_{m,t})'$  é o vetor dos log-retornos.

Dessa forma, pelas expressões (4.33) e (4.34), a expressão (4.32), para  $h = 1$ , pode ser reescrita como

$$L_{t+1} = -(f(t+1, \mathbf{Z}_t + \mathbf{X}_{t+1}) - f(t, \mathbf{Z}_t)). \quad (4.35)$$

Como  $\mathbf{Z}_t$  é observável no tempo  $t$ , a distribuição das perdas é determinada pela distribuição da mudança de fator de risco  $\mathbf{X}_{t+1}$ .

#### 4.4.1 Função de Distribuição Condicional e Não-condicional das Perdas

Em gerenciamento de riscos, muitas vezes precisamos decidir se estamos interessados na distribuição condicional ou incondicional das perdas. Como já ressaltamos anteriormente, a distribuição das variáveis aleatórias  $L_t$  é determinada pela distribuição das variáveis aleatórias  $\mathbf{X}_t$ , para todo  $t \in \mathbb{N}$ .

Suponha que o processo  $\{\mathbf{X}_t\}_{t \in \mathbb{N}}$ , de mudanças de fatores de risco, é um processo estritamente estacionário. Segue que a função de distribuição  $F_{\mathbf{X}}(\cdot)$ , da variável aleatória  $\mathbf{X}_t$ , para todo  $t \in \mathbb{N}$ , é invariante sob translações de tempo. Ressaltamos que a maioria dos modelos para séries temporais utilizados para modelar mudanças de fatores de risco satisfaz esta propriedade.

Seja  $\mathcal{F}_t$  a  $\sigma$ -álgebra gerada pelas informações disponíveis até o tempo  $t$ , isto é,  $\mathcal{F}_t \equiv \sigma(\{X_s\}, s \leq t)$ . Denotemos por  $F_{\mathbf{X}_{t+1}|\mathcal{F}_t}(\cdot)$  a distribuição condicional de  $\mathbf{X}_{t+1}$  dada a informação  $\mathcal{F}_t$ . Na maioria dos casos,  $F_{\mathbf{X}_{t+1}|\mathcal{F}_t}(\cdot)$  não coincide com  $F_{\mathbf{X}}(\cdot)$ . Isto só acontecerá quando as variáveis aleatórias do processo  $\{\mathbf{X}_t\}_{t \in \mathbb{N}}$  forem independentes e identicamente distribuídas.

Ambas as distribuições são de grande importância e são utilizadas de acordo com a medida de risco que se deseja calcular.

Para entender melhor as definições apresentadas nas últimas seções consideremos o seguinte exemplo onde apresentamos o vetor de fatores de risco, a função dos riscos e calculamos o valor do *portfolio*.

**Exemplo 4.4. (Cálculo das Perdas do *Portfolio*).** Considere um *portfolio* fixo  $\mathcal{P} = \{A_1, \dots, A_N\}$ , com  $N$  posições (ativos financeiros). Seja  $c_i$  o número de ações do ativo  $A_i$  no tempo  $t$ . Denotemos por  $P_{i,t}$  o preço por ação do ativo  $i$  no instante  $t$ . Segue que, para  $i = 1, \dots, N$ ,  $\{P_{i,t}\}_{t \in \mathbb{N}}$  é um processo estocástico que representa o preço do ativo financeiro do ativo  $A_i$ .

Se considerarmos o logaritmo do preço das ações do ativo como fator de risco, temos

$$\mathbf{Z}_t = (Z_{1,t}, \dots, Z_{N,t})' \quad \text{onde} \quad Z_{i,t} = \ln(P_{i,t}), \quad \text{para todo } i = 1, \dots, N.$$

Assim a mudança de fator de risco, do ativo  $A_i$ , definida pela expressão (4.34), é dada por

$$X_{i,t+1} = \ln(P_{i,t+1}) - \ln(P_{i,t}), \quad \text{para todo } t \in \mathbb{N}.$$

Note que, de acordo com a Definição 4.4, as variáveis  $X_{i,t}$  são os log-retornos do processo  $\{P_{i,t}\}_{t \in \mathbb{N}}$ .

O preço  $V_t$  do *portfolio*, no tempo  $t$ , é dado por

$$V_t = \sum_{i=1}^N c_i P_{i,t} = \sum_{i=1}^N c_i \exp\{Z_{i,t}\}. \quad (4.36)$$

Neste caso, a função  $f(\cdot)$ , dada pela expressão (4.33), é definida como

$$f(t, \mathbf{Z}_t) = \sum_{i=1}^N c_i \exp\{Z_{i,t}\}, \quad \text{para todo } t \in \mathbb{N}.$$

Assim, a perda no *portfolio* no instante  $t + 1$  é dada por

$$\begin{aligned} L_{t+1} &= -(V_{t+1} - V_t) \\ &= -\left( \sum_{i=1}^N c_i \exp\{Z_{i,t+1}\} - \sum_{i=1}^N c_i \exp\{Z_{i,t}\} \right) \\ &= -\sum_{i=1}^N c_i \left( \exp\{Z_{i,t+1}\} - \exp\{Z_{i,t}\} \right) \\ &= -\sum_{i=1}^N c_i \exp\{Z_{i,t}\} \left( \exp\{Z_{i,t+1} - Z_{i,t}\} - 1 \right) \\ &= -\sum_{i=1}^N c_i \exp\{Z_{i,t}\} \left( \exp\{X_{i,t+1}\} - 1 \right) \\ &= -V_t \sum_{i=1}^N a_i \left( \exp\{X_{i,t+1}\} - 1 \right), \end{aligned} \quad (4.37)$$

onde

$$a_i = \frac{c_i \exp\{Z_{i,t}\}}{V_t} = \frac{c_i P_{i,t}}{V_t}, \quad \text{para todo } i = 1, \dots, N, \quad (4.38)$$

é a proporção do valor do *portfolio* investido no ativo  $A_i$ , no tempo  $t$ . Ou ainda,

$$L_{t+1} = -V_t \sum_{i=1}^N a_i R_{i,t+1}, \quad (4.39)$$

onde  $R_{i,t+1}$  é o retorno do ativo  $A_i$ , no tempo  $t + 1$ . Note que a expressão (4.37) satisfaz a expressão (4.35). Desta forma, se conhecermos a distribuição das variáveis aleatórias  $X_{i,t}$ , para todo  $t \in \mathbb{N}$  e para todo  $i = 1, \dots, N$  (ou seja, a distribuição dos log-retornos), conheceremos a distribuição das perdas do *portfolio*  $\mathcal{P}$ .

**Observação 4.3.** Sob as mesmas hipóteses do Exemplo 4.4 pode-se mostrar que, para um horizonte  $h$  qualquer,

$$L_{t+h} \equiv -V_t \sum_{i=1}^N a_i R_{i,t}[h] \approx -V_t \sum_{i=1}^N a_i r_{i,t}[h]. \quad (4.40)$$

Para isso basta notar que  $X_{i,t} = r_{i,t}$ , para todo  $i = 1, \dots, N$ , onde  $r_{i,t}$  é o log-retorno do ativo  $A_i$  no tempo  $t$ . Logo,

$$\begin{aligned} Z_{i,t+h} - Z_{i,t} &= (Z_{i,t+h} - Z_{i,t+h-1}) + (Z_{i,t+h-1} - Z_{i,t+h-2}) + \dots + (Z_{i,t+1} - Z_{i,t}) \\ &= X_{i,t+h} + \dots + X_{i,t+1} \\ &= X_t[h]. \end{aligned}$$

Além disso, pelas expressões (4.5) e (4.6)

$$\exp \{X_{i,t}[h]\} - 1 = R_{i,t}[h].$$

Utilizando estes fatos chegamos ao resultado desejado.

## 4.5 Medidas de Risco

Segundo McNeil *et al.* (2005), existem várias abordagens para o cálculo de medidas de risco, mas de modo geral, podemos classificá-las em quatro diferentes categorias. São elas:

- *Abordagem do valor nominal.* É a abordagem mais antiga para medir os riscos de um *portfolio*. Nesta abordagem, o risco de um *portfolio* é definido como a soma dos valores nominais dos componentes individuais do *portfolio*.
- *Medidas de risco fator-sensitivas.* São aquelas que fornecem a variação do valor do *portfolio* para uma dada variação, pré-determinada, em um dos fatores de risco.
- *Medidas de risco baseadas em distribuição de perdas.* As medidas de risco mais modernas são estatísticas que descrevem a distribuição condicional e não-condicional das perdas do *portfolio* sobre um horizonte de tempo pré-determinado. Alguns exemplos são a variância, o Valor-em-Risco (VaR) e o Valor-em-Risco Condicional (CVaR ou *Expected Shortfall*).
- *Medidas de risco baseadas em cenários.* Nesta abordagem, para calcular as medidas de risco em um *portfolio*, consideramos um número de possíveis mudanças de fatores de risco (cenários), tais como 10% de variação nas taxas de câmbio ou queda de 20% no preço das ações. O risco do *portfolio* é então medido como a perda máxima do *portfolio* sobre todos os possíveis cenários.

As duas últimas abordagens são consideradas neste trabalho e descritas, a seguir, nesta seção.

### 4.5.1 Medidas de Risco Baseadas na Distribuição de Perdas e Ganhos

Uma das medidas de risco mais utilizadas é a *variância* da distribuição de perdas e ganhos. Este fato deve-se ao grande impacto que a teoria de *portfolios* de Markowitz, que utiliza variância como uma medida de risco, apresenta no estudo teórico e prático da área de finanças. Entretanto, como uma medida de risco, a variância apresenta dois problemas. O primeiro deles é que necessitamos assumir que a função de distribuição de perdas possui segundo momento finito. Além disso, essa medida não faz distinção entre desvios positivos e negativos da média. Segundo McNeil *et al.* (2005), a variância é uma boa medida de risco somente para distribuições que são (aproximadamente) simétricas, tais como a normal ou a *t*-Student (com variância finita).

Outra abordagem amplamente utilizada atualmente é a análise dos quantis da distribuição de perdas. Considere um *portfolio*  $\mathcal{P} = \{A_1, \dots, A_N\}$  de ativos de risco e um horizonte  $h$  fixos. Seja  $F_L(\ell) = \mathbb{P}(L \leq \ell)$  a função de distribuição de perdas correspondente. Nosso objetivo é definir uma estatística baseada em  $F_L(\cdot)$  que seja capaz de medir os riscos do *portfolio* sobre o período  $h$ . Um candidato natural é a perda máxima possível dada por

$$\inf \{ \ell \in \mathbb{R} : F_L(\ell) = 1 \}.$$

Entretanto, em muitos modelos o suporte de  $F_L(\cdot)$  é não-limitado. Logo, a perda máxima é infinito. A idéia então, é repassar “perda máxima” por “perda máxima que não é ultrapassada com uma certa probabilidade”, com esta probabilidade denominada *nível de confiança*.

**Definição 4.10. (Valor em Risco).** Seja  $\mathcal{P} = \{A_1, \dots, A_N\}$  um *portfolio* fixo. Dado um nível de confiança  $p \in (0, 1)$ , o *valor em risco* do *portfolio*, denotado por  $\text{VaR}_{p, \mathcal{P}}$ , é definido como

$$\begin{aligned} \text{VaR}_{p, \mathcal{P}} &\equiv \inf \{ \ell \in \mathbb{R} : \mathbb{P}(L \geq \ell) \leq 1 - p \} \\ &= \inf \{ \ell \in \mathbb{R} : F_L(\ell) \geq p \}. \end{aligned} \tag{4.41}$$

Em termos probabilísticos,  $\text{VaR}_p$  é o  $p$ -quantil da função de distribuição de perdas. Na prática, os valores freqüentemente utilizados para  $p$  são 0,95 ou 0,99 e para  $h$  são 1 ou 10 dias.

Assim como na análise da medida de risco pela variância, o  $\text{VaR}_p$  apresenta certas desvantagens. Artzner *et al.* (1999) mostram que o  $\text{VaR}_p$  não é uma medida de risco coerente pois não satisfaz o axioma da subaditividade. Isto é, o  $\text{VaR}_p$  de *portfolios* somados pode não ser limitado pela soma do  $\text{VaR}_p$  dos *portfolios* individuais. Isso contradiz a idéia de que os riscos podem ser diminuídos através da diversificação, ou seja, através da compra ou venda de ativos.

Apresentamos, a seguir, uma medida de risco que é coerente, no sentido da Definição 4.2 (veja também o Exemplo 4.2).

**Definição 4.11. (Expected Shortfall).** Considere uma perda  $L$  com função de distribuição  $F_L(\cdot)$ , tal que  $\mathbb{E}(|L|) < \infty$ . O *expected shortfall*, denotado por  $\text{ES}_p$ , ao nível de confiança  $p \in (0, 1)$ , é definido como

$$\text{ES}_p \equiv \frac{1}{1 - p} \int_p^1 q_u(F_L) du,$$

onde  $q_u(\cdot)$  é a função quantil definida por  $q_u(F_L) \equiv \inf\{\ell \in \mathbb{R} : F_L(\ell) \geq u\}$ .

É imediato ver que as medidas de risco *Expected Shortfall* e VaR estão relacionados através da expressão

$$\text{ES}_p \equiv \frac{1}{1-p} \int_p^1 \text{VaR}_u du.$$

Mostra-se (veja McNeil *et al.*, 2005) que se  $L$  for integrável com função de distribuição  $F_L(\cdot)$  contínua então

$$\text{ES}_p = \mathbb{E}(L|L \geq \text{VaR}_p).$$

**Observação 4.4.** Na literatura, encontramos variações para a medida de risco ES, dada na Definição 4.11, tais como *tail conditional expectation* (TCE), *worst conditional expectation* (WCE) e *conditional VaR* (CVaR). Apesar de apresentarem definições um pouco diferentes, quando a função de distribuição das perdas for contínua todas essas medidas coincidem.

Na prática, para calcular o VaR e o ES, devemos estimar a função de distribuição das perdas  $F_L(\cdot)$ . É imediato que diferentes formas de estimar  $F_L(\cdot)$  fornecerão diferentes valores para essas medidas. Como ressaltamos na Seção 4.4, a função de distribuição da variável aleatória  $L_t$  fica determinada se conhecermos a distribuição do vetor aleatório  $\mathbf{X}_t$  de mudança de fatores de risco, no instante  $t$ . No Exemplo 4.4 consideramos o caso em que os fatores de risco são os log-retornos. Vimos, pela expressão (4.39) que

$$L_{t+1} = -V_t \sum_{i=1}^N a_i R_{i,t+1},$$

onde  $R_{i,t}$  é o retorno do ativo  $A_i$ , no tempo  $t + 1$ . Na prática  $R_t$  e  $r_t$  são valores muito próximos entre si. Assim, podemos assumir que

$$L_{t+1} = -V_t \sum_{i=1}^N a_i r_{i,t+1}.$$

Segue que

$$\mathbb{E}(L_{t+1}) = -V_t \mathbf{a}' \boldsymbol{\mu} \quad \text{e} \quad \text{Var}(L_{t+1}) = V_t^2 \mathbf{a}' \boldsymbol{\Sigma} \mathbf{a}, \quad (4.42)$$

onde  $\mathbf{a} = (a_1, \dots, a_N)'$  é o vetor de pesos,  $\boldsymbol{\mu} = (\mathbb{E}(r_{1,t+1}), \dots, \mathbb{E}(r_{N,t+1}))'$  e  $\boldsymbol{\Sigma}$  é a matriz de variâncias-covariâncias. Como  $V_t$  é um valor conhecido no tempo  $t$  e portanto, constante em relação a  $t + 1$ , podemos calcular o  $\text{VaR}_p$  considerando  $V_t = 1$ . Como os retornos do *portfolio* são, em geral, dados em porcentagem, o  $\text{VaR}_p$  obtido fixando-se  $V_t = 1$  representa também uma porcentagem. Multiplicando o valor obtido por  $V_t$  obtemos o risco do *portfolio* em termos de unidades monetárias.

Na próxima subseção apresentamos algumas das abordagens mais utilizadas para o cálculo do  $\text{VaR}_p$ .

## 4.5.2 Cálculo do Valor em Risco - $\text{VaR}_p$

Nesta subseção apresentamos as principais abordagens, existentes na literatura, para o cálculo do  $\text{VaR}_p$ . Para isto, consideramos tanto a função de distribuição condicional como

a não-condicional dos fatores de mudança de risco. Entre as abordagens existentes para o cálculo desta medida podemos citar:

**(i) VaR<sub>p</sub> Empírico**

Esta é a abordagem não-paramétrica para o cálculo do VaR<sub>p</sub>. O VaR<sub>p</sub> empírico é o  $p$ -quantil da função de distribuição empírica dos dados, dada por

$$F_L(\ell) = \frac{1}{n} \sum_{t=1}^n I_{\{L_t \leq \ell\}}, \quad \text{para todo } \ell \in \mathbb{R},$$

onde  $I_{\{B\}}$  é a função indicadora do conjunto  $B$ . Segue que

$$\text{VaR}_p = \inf \left\{ L_t \in \{L_t\}_{t=1}^n : F_L(L_t) \geq p \right\}.$$

Sob essa abordagem, o VaR<sub>p</sub> de  $h$  períodos é igual ao VaR<sub>p</sub> de 1 período.

**(ii) VaR<sub>p</sub> Normal ou Método da Variância-Covariância**

Considere a função de distribuição não-condicional do vetor de mudança de fatores de risco  $\mathbf{X}_{t+1}$  e assumamos que  $\mathbf{X}_{t+1} = (r_{1,t+1}, \dots, r_{N,t+1})$  tem distribuição normal multivariada com média  $\boldsymbol{\mu} = (\mathbb{E}(r_{1,t+1}), \dots, \mathbb{E}(r_{N,t+1}))'$  e matriz de variância-covariância  $\boldsymbol{\Sigma}$  constantes. Lembre que, por simplicidade, estamos fixando  $V_t = 1$ , de forma que o VaR<sub>p</sub> calculado será o VaR<sub>p</sub> para cada unidade monetária. Segue, pela expressão (4.42), que

$$L_{t+1} \sim \mathcal{N}(-\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}).$$

Os parâmetros  $\boldsymbol{\mu}$  e  $\boldsymbol{\Sigma}$  são estimados através da média, variância e covariância amostrais. Como o VaR<sub>p</sub> é, por definição, o  $p$ -quantil da função de distribuição das perdas, ele é dado por

$$\text{VaR}_p = \bar{L} + \Phi^{-1}(p)\sqrt{\hat{\sigma}_L^2}, \tag{4.43}$$

onde

$$\bar{L} = -\sum_{i=1}^N a_i \bar{r}_i, \quad \hat{\sigma}_L^2 = \sum_{i=1}^N \sum_{j=1}^N a_i a_j \text{Cov}(r_i, r_j), \tag{4.44}$$

$\bar{r}_i = \sum_{t=1}^n r_{i,t}$  é a média amostral do ativo  $A_i$ , para todo  $i = 1, \dots, N$ ,  $\text{Cov}(r_i, r_j)$  é a covariância amostral dos log-retornos  $r_i$  e  $r_j$ , dada por

$$\text{Cov}(r_i, r_j) = \frac{1}{n} \sum_{t=1}^n (r_{i,t} - \bar{r}_i)(r_{j,t} - \bar{r}_j)$$

e  $\Phi(\cdot)$  é a função de distribuição normal padrão. Assim, o valor em risco do *portfolio* no tempo  $t + 1$  será dado por

$$\text{Risco} = V_t \times \text{VaR}_p.$$

Para o cálculo do VaR<sub>p</sub> de  $h$  períodos note que, pela expressão (4.40), tomando  $V_t = 1$ ,

$$\begin{aligned} L_{t+h} &= -\sum_{i=1}^N a_i r_{i,t}[h] \\ &= -\sum_{i=1}^N a_i (r_{i,t+1} + \dots + r_{i,t+h}) \\ &= -\sum_{k=1}^h \sum_{i=1}^N a_i r_{i,t+k}. \end{aligned}$$



Assim, os estimadores amostrais para a média e a variância são, respectivamente,

$$-h \times \sum_{i=1}^N a_i \times \bar{r}_i = -h\bar{L} \quad \text{e} \quad h \sum_{i=1}^N \sum_{j=1}^N a_i a_j \text{Cov}(r_i, r_j) = h\hat{\sigma}_L^2,$$

com  $\bar{L}$  e  $\hat{\sigma}_L^2$  dadas pela expressão (4.44). De onde segue que

$$\text{VaR}_p[h] = -h\bar{L} + \sqrt{h}\Phi^{-1}(p)\sqrt{\hat{\sigma}_L^2}.$$

Podemos também considerar a função de distribuição condicional do vetor aleatório  $\mathbf{X}_{t+1} = (r_{t+1,1}, \dots, r_{t+1,N})'$ . Por exemplo, se assumirmos que  $\mathbf{X}_{t+1}|\mathcal{F}_t \sim \mathcal{N}(\boldsymbol{\mu}_{t+1}, \boldsymbol{\Sigma}_{t+1})$ . Segue que

$$L_{t+1} \sim \mathcal{N}(-V_t \mathbf{a}' \boldsymbol{\mu}_{t+1}, V_t^2 \mathbf{a}' \boldsymbol{\Sigma}_{t+1} \mathbf{a}).$$

Fixando  $V_t = 1$ , obtemos

$$\text{VaR}_{p,t+1} = -\mathbf{a}' \boldsymbol{\mu}_{t+1} + \Phi^{-1}(p) \times \sqrt{\mathbf{a}' \boldsymbol{\Sigma}_{t+1} \mathbf{a}}.$$

E o valor em risco do *portfolio* no tempo  $t + 1$  é então dado por

$$\text{Risco} = V_t \times \left( -\mathbf{a}' \boldsymbol{\mu}_{t+1} + \Phi^{-1}(p) \times \sqrt{\mathbf{a}' \boldsymbol{\Sigma}_{t+1} \mathbf{a}} \right).$$

Neste caso, ajustamos um determinado modelo aos dados e utilizamos os valores previstos através do modelo para estimar a média e a matriz de variâncias-covariâncias.

As abordagens que apresentamos a seguir são as mais utilizadas no caso em que considera-se a função de distribuição condicional.

### (iii) Abordagem RiskMetrics

Considere, primeiramente, o caso de um único ativo. A metodologia RiskMetrics foi desenvolvida por J.P. Morgan para o cálculo do  $\text{VaR}_p$ . Em sua forma mais simples a metodologia assume que

$$r_t | \mathcal{F}_{t-1} \sim \mathcal{N}(\mu_t, \sigma_t^2),$$

onde a média e a variância condicionais são tais que

$$\mu_t = 0 \quad \text{e} \quad \sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2, \quad 0 < \lambda < 1, \quad (4.45)$$

isto é,  $\{\sigma_t^2\}_{t \in \mathbb{Z}}$  segue um modelo EWMA, dado na Definição 2.4. Segue que  $L_{t+1} = -r_{t+1}$ . Logo, o  $\text{VaR}_p$  do ativo no tempo  $t + 1$ , é dado por

$$\text{VaR}_{p,t+1} = \Phi^{-1}(p) \times \sigma_{t+1},$$

onde  $\Phi(\cdot)$  é a função de distribuição normal padrão e  $\sigma_{t+1}^2$  é obtida através do modelo dado na expressão (4.45). Multiplicando pelo preço do ativo no tempo  $t$ , obtemos o risco do ativo no tempo  $t + 1$ , isto é,

$$\text{Risco} = V_t \times \text{VaR}_{p,t+1}.$$

Para o cálculo do  $\text{VaR}_p$  para um horizonte  $h$ , note que pela expressão (4.40), tomando  $V_t = 1$ ,

$$L_{t+h} = -r_t[h].$$

Pela expressão (4.6), o log-retorno de  $h$  períodos é dado por

$$r_t[h] = r_{t+1} + \cdots + r_{t+h}.$$

Como um processo EWMA é um caso particular de um IGARCH( $p, q$ ), com  $p = 1 = q$ , e este é um processo *martingale difference* (veja Proposição 2.1), temos

$$\sigma_t^2[h] \equiv \text{Var}(r_t[h]|\mathcal{F}_t) = \sum_{j=1}^h \text{Var}(r_{t+j}|\mathcal{F}_t), \quad (4.46)$$

pois  $\text{Cov}(r_t, r_s) = 0$ , para todo  $t, s$  tais que  $t \neq s$ . O valor  $\text{Var}(r_{t+j}|\mathcal{F}_t)$ , para  $j = 1, \dots, h$ , na expressão (4.46) é obtido através do modelo dado pela equação (4.45). Pela expressão (3.53), segue que

$$\text{Var}(r_{t+h}|\mathcal{F}_t) = \mathbb{E}(r_{t+h}^2|\mathcal{F}_t) = \mathbb{E}(\sigma_{t+h}^2|\mathcal{F}_t) = \sigma_{t+1}^2.$$

Então,  $\sigma_t^2[h] = h\sigma_{t+1}^2$ . Como o método RiskMetrics assume que  $r_t[h]|\mathcal{F}_t \sim \mathcal{N}(0, \sigma_t^2[h])$ , segue que o  $\text{VaR}_p$  para o horizonte  $h$  é dado por

$$\text{VaR}_{p,t}[h] = \Phi^{-1}(p)\sqrt{h}\sigma_{t+1},$$

onde  $\Phi(\cdot)$  é a função de distribuição normal padrão e  $\sigma_{t+1}^2$  é dada pela expressão (4.45).

Assim, o valor em risco do *portfolio* para um horizonte  $h$  será dado por

$$\text{Risco} = V_t \times \Phi^{-1}(p) \times \sqrt{h} \times \sigma_{t+1}.$$

Esta abordagem é amplamente utilizada devido à fácil implementação do modelo EWMA.

É fácil ver que, se  $\mu_t = 0$ , então  $\text{VaR}_{p,t}[h] = \sqrt{h}\text{VaR}_{p,t+1}$ . Ressaltamos que quando  $\mu_t \neq 0$  essa igualdade não é mais válida. Suponha que  $\mu_t = \mu$ , para todo  $t \in \mathbb{N}$ , segue que  $\text{VaR}_{p,t+1} = \mu + \Phi^{-1}(p) \times \sigma_{t+1}$ , enquanto que

$$\text{VaR}_{p,t}[h] = h \times \mu + \sqrt{h} \times \Phi^{-1}(p) \times \sigma_{t+1} \neq \sqrt{h} \times \text{VaR}_{p,t+1},$$

para todo  $h > 1$  (veja Morettin, 2006). Para o caso em que o *portfolio* é formado por  $N$  ativos  $\mathcal{P} = \{A_1, \dots, A_N\}$ , o  $\text{VaR}_p$ , no instante  $t + 1$  do *portfolio* é dado por

$$\text{VaR}_{p,\mathcal{P},t+1} = \sqrt{\sum_{i=1}^N (\text{VaR}_{p,i,t+1})^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \rho_{ij} \text{VaR}_{p,i,t+1} \text{VaR}_{p,j,t+1}}, \quad (4.47)$$

onde  $\text{VaR}_{p,i,t+1}$  é o valor em risco do ativo  $A_i$ , no instante  $t + 1$ ,

$$\rho_{ij} = \text{Corr}(r_{i,t+1}, r_{j,t+1}) = \frac{\text{Cov}(r_{i,t+1}, r_{j,t+1})}{\sqrt{\text{Var}(r_{i,t+1})}\sqrt{\text{Var}(r_{j,t+1})}}$$

e  $\text{Cov}(r_{i,t+1}, r_{j,t+1}) = \gamma_{ij,t+1}$  é estimada através da equação

$$\gamma_{ij,t} = \lambda\gamma_{ij,t-1} + (1 - \lambda)r_{i,t-1}r_{j,t-1}, \quad \text{para } 0 < \lambda < 1.$$

Para mais detalhes veja Zangari (1996).

#### (iv) Abordagem Econométrica

Uma série de retornos financeiros é, em geral, não-correlacionada mas apresenta correlação no quadrado dos retornos. Se isto ocorre a volatilidade pode ser modelada através dos modelos apresentados no Capítulo 2. Em alguns casos as séries temporais dos retornos apresentam autocorrelação significativa entre as variáveis aleatórias. Quando isso ocorre ajustamos primeiramente um modelo linear, por exemplo um modelo ARMA, e em seguida um modelo heteroscedástico.

Considere o caso de um único ativo financeiro. Suponha que o modelo escolhido seja um ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ). Então, por hipótese,  $r_t|\mathcal{F}_{t-1}$  tem função de distribuição  $F_{r_t|\mathcal{F}_{t-1}}(\cdot)$ , com  $\text{Var}(r_t|\mathcal{F}_{t-1}) = \sigma_t^2$ , para todo  $t > 0$ , e satisfaz

$$r_t = \phi_0 + \sum_{l=1}^{p_1} \phi_l r_{t-l} + X_t - \sum_{k=1}^{q_1} \theta_k X_{t-k} \quad (4.48)$$

$$X_t = \sigma_t Z_t \quad (4.49)$$

$$\ln(\sigma_t^2) = \omega + \frac{1 - \sum_{i=1}^{p_2} \alpha_i \mathcal{B}^i}{(1 - \sum_{j=1}^{q_2} \beta_j \mathcal{B}^j)(1 - \mathcal{B})^d} g(Z_{t-1}), \quad (4.50)$$

com  $1 - \sum_{j=1}^{q_2} \beta_j \mathcal{B}^j \neq 0$ , para todo  $|\mathcal{B}| \leq 1$ , e  $g(\cdot)$  definida pela expressão (2.38). Se denotarmos

$$\mu_t = \phi_0 + \sum_{l=1}^{p_1} \phi_l r_{t-l} - \sum_{k=1}^{q_1} \theta_k X_{t-k}, \quad (4.51)$$

teremos  $r_t = \mu_t + \sigma_t Z_t$ . Além disso,

$$\mathbb{E}(r_t|\mathcal{F}_{t-1}) = \mathbb{E}(\mu_t + \sigma_t Z_t|\mathcal{F}_{t-1}) = \mathbb{E}(\mu_t|\mathcal{F}_{t-1}) + \mathbb{E}(\sigma_t Z_t|\mathcal{F}_{t-1}) = \mu_t, \quad (4.52)$$

pois  $\sigma_t$  e  $\mu_t$  são  $\mathcal{F}_{t-1}$ -mensurável e  $\mathbb{E}(Z_t|\mathcal{F}_{t-1}) = 0$ .

Supondo que  $Z_t \sim \mathcal{N}(0, 1)$ , segue que  $r_t|\mathcal{F}_{t-1} \sim \mathcal{N}(\mu_t, \sigma_t^2)$ . De fato,  $\mu_t$  e  $\sigma_t$  são  $\mathcal{F}_{t-1}$ -mensuráveis. Logo,

$$F_{r_t|\mathcal{F}_{t-1}}(z) = \mathbb{P}\left(\frac{r_t - \mu_t}{\sigma_t} \leq \frac{z - \mu_t}{\sigma_t} \middle| \mathcal{F}_{t-1}\right) = \mathbb{P}\left(Z_t \leq \frac{z - \mu_t}{\sigma_t} \middle| \mathcal{F}_{t-1}\right) = \mathbb{P}\left(Z_t \leq \frac{z - \mu_t}{\sigma_t}\right)$$

e o resultado segue.

Lembramos que  $L_{t+1} = -V_t r_{t+1}$ , onde  $V_t$  é o valor do ativo no tempo  $t$ . Tomando  $V_t = 1$ , segue que

$$\text{VaR}_{p,t+1} = -\mu_{t+1} + \Phi^{-1}(p)\sigma_{t+1},$$

com  $\mu_{t+1}$  e  $\sigma_{t+1}$  estimadas pelas equações (4.48), (4.49) e (4.50). Logo, o risco do *portfolio* no tempo  $t + 1$  é dado por

$$\text{Risco} = V_t \times (-\mu_{t+1} + \Phi^{-1}(p)\sigma_{t+1}).$$

Para o cálculo do  $\text{VaR}_p$  para um horizonte  $h$ , note que, pela expressão (4.40),  $L_{t+h} = -V_t r_t[h]$ . Portanto, precisamos estimar a média e a variância condicional do retorno  $h$  dias à frente,  $r_t[h]$ , dada a informação até o instante  $t$ ,  $\mathcal{F}_t$ . A proposição que segue trata da estimação da média e da variância do retorno de período  $h$ .

**Proposição 4.1.** *Seja  $\{r_t\}_{t \in \mathbb{Z}}$  um processo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), definido pelas expressões (4.48), (4.49) e (4.50). Segue que*

(i) A previsão da média do retorno de período  $h$  é dada por

$$\hat{r}_t[h] = \sum_{j=1}^h \hat{r}_{t+j}, \quad (4.53)$$

onde  $\hat{r}_{t+j}$  é a previsão  $j$  passos à frente para o modelo descrito pelas equações (4.48).

(ii) O erro de previsão é dado por

$$e_n[h] = \sum_{j=0}^{h-1} \left[ \left( \sum_{i=0}^j \psi_i \right) X_{t+h-j} \right],$$

onde  $\psi_i$ , para todo  $i > 0$  são os coeficientes da representação MA( $\infty$ ) do processo estocástico  $\{r_t\}_{t \in \mathbb{Z}}$ , definidos pela expressão (1.14).

(iii) A previsão da volatilidade do retorno de período  $h$  é dada por

$$\hat{\sigma}_t^2[h] = \sum_{j=0}^{h-1} \left[ \left( \sum_{i=0}^j \psi_i \right)^2 \hat{\sigma}_{t+h-j}^2 \right], \quad (4.54)$$

onde  $\psi_i$ , para todo  $i > 0$ , são os coeficientes da representação MA( $\infty$ ) do processo estocástico  $\{r_t\}_{t \in \mathbb{Z}}$ , definidos pela expressão (1.14) e  $\hat{\sigma}_{t+j}^2$  é a previsão  $j$  passos à frente para o modelo descrito pelas equações (4.49) e (4.50).

**Prova:** Para provar (i), note que

$$r_t[h] = \sum_{j=1}^h r_{t+j} \implies \hat{r}_t[h] = \mathbb{E}(r_t[h] | \mathcal{F}_{t-1}) = \sum_{j=1}^h \mathbb{E}(r_{t+j} | \mathcal{F}_{t-1}) = \sum_{j=1}^h \hat{r}_{t+j}, \quad (4.55)$$

onde  $\hat{r}_{t+j}$  é a previsão  $j$  passos à frente para o modelo descrito pelas equações (4.48). Pela expressão (3.39),  $\hat{r}_{n+j}$  é dado por

$$\hat{r}_{n+j} = \phi_0 + \sum_{i=1}^{p_1} \phi_i \hat{r}_{n+j-i} - \sum_{j=0}^{q_1} \theta_j \hat{X}_{n+j-j},$$

com  $\hat{r}_{n+j} = r_{n+j}$ , se  $j \leq 0$ ,  $\hat{X}_{n+j} = X_{n+j}$ , se  $j \leq 0$  e  $\hat{X}_{n+j} = 0$ , se  $j > 0$ , para todo  $1 \leq j \leq h$ . Ou ainda, pela expressão (3.36),

$$\hat{r}_{n+j} = \sum_{i=0}^{\infty} \psi_i \hat{X}_{n+j-i} = \sum_{i=j}^{\infty} \psi_i X_{n+j-i}, \quad (4.56)$$

onde  $\psi_i$ , para todo  $i > 0$ , são os coeficientes da representação MA( $\infty$ ) do processo estocástico  $\{r_t\}_{t \in \mathbb{Z}}$ , definidos pela expressão (1.14).

Para provar (ii) note que, pela expressão (4.56), o erro de previsão, para cada  $1 \leq j \leq h$ , é dado por

$$\begin{aligned} e_n(j) = r_{n+j} - \hat{r}_{n+j} &= \sum_{i=0}^{\infty} \psi_i X_{n+j-i} - \sum_{i=j}^{\infty} \psi_i X_{n+j-i} \\ &= \sum_{i=0}^{j-1} \psi_i X_{n+j-i}. \end{aligned} \quad (4.57)$$

Segue, da expressão (4.57), que o erro de previsão para o retorno de período  $h$  é dado por

$$\begin{aligned}
e_n[h] &= \sum_{j=1}^h e_n(j) = \psi_0 X_{t+1} + (\psi_0 X_{t+2} + \psi_1 X_{t+1}) + \cdots + (\psi_0 X_{t+h} + \cdots + \psi_{h-1} X_{t+1}) \\
&= \psi_0 X_{t+h} + (\psi_0 + \psi_1) X_{t+h-1} + \cdots + (\psi_0 + \cdots + \psi_{h-1}) X_{t+1} \\
&= \sum_{j=0}^{h-1} \left[ \left( \sum_{i=0}^j \psi_i \right) X_{t+h-j} \right].
\end{aligned} \tag{4.58}$$

Para (iii) note que,

$$\sigma_t^2[h] = \text{Var}(r_t[h]|\mathcal{F}_t) = \text{Var}(r_t[h] - \mu_t[h]|\mathcal{F}_t). \tag{4.59}$$

Como  $\hat{\mu}_t[h] = \hat{r}_t[h]$ , com  $\hat{r}_t[h]$  dado pela expressão (4.55), segue que

$$\hat{\sigma}_t^2[h] = \text{Var}(r_t[h] - \hat{r}_t[h]|\mathcal{F}_t) = \text{Var}(e_n[h]|\mathcal{F}_t), \tag{4.60}$$

com  $e_n[h]$  dado pela expressão (4.58). Suponha que  $0 < j < k$ . Como  $\mathbb{E}(X_{t+j}|\mathcal{F}_t) = \mathbb{E}(\sigma_{t+j} Z_{t+j}|\mathcal{F}_t) = \mathbb{E}(\sigma_{t+j}|\mathcal{F}_t)\mathbb{E}(Z_{t+j}) = 0$ , segue que

$$\begin{aligned}
\text{Cov}(X_{t+j}, X_{t+k}|\mathcal{F}_t) &= \mathbb{E}(X_{t+j} X_{t+k}|\mathcal{F}_t) - \mathbb{E}(X_{t+j}|\mathcal{F}_t)\mathbb{E}(X_{t+k}|\mathcal{F}_t) \\
&= \mathbb{E}(\mathbb{E}(X_{t+j} X_{t+k-1}|\mathcal{F}_t)|\mathcal{F}_{t+j}) - 0 \\
&= \mathbb{E}(\mathbb{E}(X_{t+j} X_{t+k}|\mathcal{F}_{t+k-1})|\mathcal{F}_t) \\
&= \mathbb{E}(X_{t+j} \mathbb{E}(X_{t+k}|\mathcal{F}_{t+k-1})|\mathcal{F}_t) \\
&= \mathbb{E}(X_{t+j} 0|\mathcal{F}_t) \\
&= 0.
\end{aligned} \tag{4.61}$$

Assim, pelas expressões (4.58), (4.60) e (4.61) segue que

$$\hat{\sigma}_t^2[h] = \sum_{j=0}^{h-1} \left[ \left( \sum_{i=0}^j \psi_i \right)^2 \hat{\sigma}_{t+h-j}^2 \right], \tag{4.62}$$

onde  $\psi_i$ , para todo  $i > 0$ , são os coeficientes da representação  $\text{MA}(\infty)$  do processo estocástico  $\{r_t\}_{t \in \mathbb{Z}}$ , definidos pela expressão (1.14) e  $\hat{\sigma}_{t+j}^2$  é a previsão  $j$  passos à frente para o modelo descrito pelas equações (4.49) e (4.50).

Isto completa a prova. □

A Proposição 4.1 nos fornece as estimativas da média e da variância condicional do retorno de período  $h$ . Sendo assim, sob a hipótese de normalidade,

$$L_{t+h}|\mathcal{F}_t \sim \mathcal{N}(-V_t \hat{r}_t[k], V_t^2 \hat{\sigma}_t^2[h]),$$

com  $\hat{r}_t[k]$  e  $\hat{\sigma}_t^2[h]$  dados, respectivamente, pelas expressões (4.53) e (4.62). Tomando  $V_t = 1$  segue que

$$\text{VaR}_{p,t}[h] = -\hat{r}_t[h] + \Phi^{-1}(p)\hat{\sigma}_t[h]$$

e o risco do *portfolio* para o horizonte  $h$  é dado por

$$\text{Risco} = V_t \times \text{VaR}_{p,t}[h].$$

### Observação 4.5.

1. A abordagem econométrica pode ser estendida para o caso de  $N$  ativos financeiros. Existem vários trabalhos onde modelos multivariados ARMA-GARCH são utilizados para obter as estimativas da média e da matriz de variância-covariância do vetor das mudanças de fatores de risco. Para mais detalhes veja McNeil *et al.* (2005) e referências ali contidas. A abordagem multivariada não será considerada neste trabalho.
2. Note que  $\gamma_{ij,t} = \rho_{ij,t}\sigma_{i,t}\sigma_{j,t}$ , onde  $\gamma_{ij,t}$  e  $\rho_{ij,t}$  são, respectivamente, a covariância e a correlação dos log-retornos, para todo  $i, j$  e  $t \in \mathbb{Z}$ . Partindo desta igualdade vários autores sugerem utilizar modelos ARMA-GARCH multivariados (ou univariados) para obter estimativas para  $\sigma_{i,t}^2$  e modelar a dependência entre as variáveis aleatórias através de cópulas. Para um exemplo veja Palaro e Hotta (2006).
3. Uma alternativa para o cálculo das medidas de risco é, ao invés de modelar as séries temporais dos retornos dos ativos, é modelar diretamente a série temporal de retornos (ou log-retornos) do *portfolio* como um todo (veja Palaro e Hotta, 2006). Reduz-se assim a abordagem multivariada para um caso univariado.

### 4.5.3 Cálculo do *Expected Shortfall* - ES

Apresentamos nesta subseção resultados relacionados ao cálculo do *Expected Shortfall*.

**Proposição 4.2.** *Seja  $L$  a variável aleatória que representa as perdas do portfolio. Se  $L$  possui função de distribuição normal com média  $\mu$  e variância  $\sigma^2$  então, para todo  $p \in (0, 1)$ ,*

$$\text{ES}_p = \mu + \sigma \frac{\phi(\Phi^{-1}(p))}{1-p},$$

onde  $\phi(\cdot)$  e  $\Phi(\cdot)$  são a função densidade de probabilidade e a função de distribuição normal padrão, respectivamente.

**Prova:** Por definição,  $\mathbb{P}(L \geq \text{VaR}_p) = 1 - p$  e  $\text{ES}_p = \mathbb{E}(L|L \geq \text{VaR}_p)$ . Por hipótese  $L \sim \mathcal{N}(\mu, \sigma^2)$ . Logo,  $\text{VaR}_p = \mu + \sigma\Phi^{-1}(p)$ . Segue que

$$\text{ES}_p = \mathbb{E}(L|L \geq \text{VaR}_p) = \frac{1}{1-p} \int_{\mu+\sigma\Phi^{-1}(p)}^{\infty} \ell dF_L(\ell) = \frac{1}{1-p} \int_{\mu+\sigma\Phi^{-1}(p)}^{\infty} \ell \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\ell-\mu)^2}{2\sigma^2}} d\ell. \quad (4.63)$$

Tomando  $z = \frac{\ell-\mu}{\sigma}$  na expressão (4.63), obtemos

$$\begin{aligned} \text{ES}_p &= \frac{1}{1-p} \int_{\Phi^{-1}(p)}^{\infty} (\mu + \sigma z) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2}} \sigma dz = \frac{1}{1-p} \int_{\Phi^{-1}(p)}^{\infty} \mu \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2}} \sigma dz \\ &\quad + \frac{1}{1-p} \int_{\Phi^{-1}(p)}^{\infty} \sigma z \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2}} \sigma dz. \end{aligned} \quad (4.64)$$

Note que a primeira integral do lado direito da igualdade (4.64) é, por definição, igual a  $1 - \mathbb{P}(Z \leq \Phi^{-1}(p)) = 1 - p$ , onde  $Z$  é uma variável aleatória com função de distribuição normal padrão. Segue que

$$\text{ES}_p = \mu + \frac{\sigma}{1-p} \int_{\Phi^{-1}(p)}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz. \quad (4.65)$$

Pela expressão (4.65), utilizando a mudança de variável  $u = z^2/2$ , segue que

$$\begin{aligned}
 \text{ES}_p &= \mu + \frac{\sigma}{1-p} \int_{\Phi^{-1}(p)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u} du \\
 &= \mu - \frac{\sigma}{1-p} \left[ \frac{1}{\sqrt{2\pi}} e^{-u} \right]_{\Phi^{-1}(p)}^{\infty} \\
 &= \mu - \frac{\sigma}{1-p} \left[ \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \right]_{\Phi^{-1}(p)}^{\infty}.
 \end{aligned} \tag{4.66}$$

Como  $\lim_{z \rightarrow \infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} = 0$ , pela expressão (4.66), segue que

$$\text{ES}_p = \mu + \sigma \frac{\phi(\Phi^{-1}(p))}{1-p}.$$

□

O resultado da Proposição 4.2 é válido também para a função de distribuição condicional e a demonstração é análoga.

**Proposição 4.3.** *A função Expected Shortfall (ES) é uma medida de risco coerente.*

**Prova:** Veja McNeil *et al.* (2005), página 243.

□

Como consequência da Proposição 4.3, segue que, para um *portfolio*  $\mathcal{P} = \{A_1, \dots, A_N\}$  formado por  $N$  ativos, com pesos  $\mathbf{a} = (a_1, \dots, a_N)'$ , a desigualdade

$$\text{ES}_{\mathcal{P}, t+1} \leq \sum_{i=1}^N a_i \text{ES}_{p, i, t+1},$$

é sempre válida.

## 4.6 Análise de Cenários - Teste de Estresse

A maioria das abordagens apresentadas na literatura para o cálculo do  $\text{VaR}_p$  assume que as mudanças nos fatores de risco são normalmente distribuídas. Entretanto, na maioria dos casos, mudanças em séries temporais financeiras apresentam caudas mais pesadas que as da função de distribuição normal.

Como veremos a seguir, testes de estresse não baseiam-se em hipóteses sobre a função de distribuição das mudanças de fatores de risco. Por esse motivo não são influenciados por caudas pesadas. Como tais testes não quantificam a probabilidade de ocorrência de cenários individuais, eles são utilizados como ferramenta auxiliar para verificar e complementar as estatísticas do tipo medidas de risco tais como o  $\text{VaR}_p$ . Mesmo assim, os cenários precisam ser minimamente plausíveis e para isso é necessária uma idéia, ainda que vaga, das probabilidades de ocorrência de cada cenário.

### 4.6.1 Descrição do Teste de Estresse

O teste de estresse baseia-se na idéia de que o valor do *portfolio* depende dos fatores de risco (veja Seção 4.4). Sejam  $\mathbf{Z} = (Z_1, \dots, Z_m)'$  o vetor dos  $m$  fatores de risco que influenciam no valor do *portfolio* e  $f(\cdot)$  a função que determina o valor deste *portfolio*. Então, dizemos que o vetor  $\mathbf{Z}$  descreve a situação do mercado e que  $f(\mathbf{Z})$  é o valor do *portfolio* para esta situação (ou cenário). No que segue, nos referimos a  $\mathbf{Z}_{AM}$  como a *situação atual do mercado* e conseqüentemente  $f(\mathbf{Z}_{AM})$  representa o *valor atual do portfolio*.

Como ressaltamos na Seção 4.4, a escolha dos fatores de risco e da função  $f(\cdot)$  dependem do *portfolio*. Nem todos os *portfolios* são influenciados pelos mesmos fatores de risco e diferentes *portfolios* podem ter seus valores calculados por diferentes funções e estas, em geral, não são funções explícitas dos fatores de risco.

O teste de estresse nos diz o que aconteceria se uma dada situação de mercado  $\mathbf{Z}$  repentinamente ocorresse. Para o teste de estresse,  $k$  diferentes cenários  $\mathbf{Z}_1, \dots, \mathbf{Z}_k$  são selecionados de acordo com algum critério específico. Calcula-se, então os valores do *portfolio*  $f(\mathbf{Z}_1), \dots, f(\mathbf{Z}_k)$  sobre esses cenários. Comparando com o valor atual do *portfolio*  $f(\mathbf{Z}_{AM})$ , podemos determinar as perdas que ocorreriam se o mercado passasse repentinamente da situação  $\mathbf{Z}_{AM}$  para alguma das  $k$  situações  $\mathbf{Z}_1, \dots, \mathbf{Z}_k$ .

### 4.6.2 Avaliação do *Portfolio*

No teste de estresse, o primeiro passo após a escolha dos cenários é determinar o valor do *portfolio* sobre cada cenário. Dado que o mercado encontra-se na situação  $\mathbf{Z}_{AM}$ , com o valor do *portfolio* dado por  $f(\mathbf{Z}_{AM})$ , existem duas abordagens para determinar o valor do *portfolio* sobre um outro cenário  $\mathbf{Z}_K$ , para  $K = 1, \dots, k$ . Descrevemos, a seguir, estas duas abordagens:

**Reavaliação Completa:** consiste em calcular diretamente o valor do *portfolio* utilizando os novos valores  $Z_i$ , para  $i = 1, \dots, m$ , dos fatores de risco para o cenário  $\mathbf{Z}_K$ , para  $K = 1, \dots, k$ . Teremos então que o valor do *portfolio* sobre o cenário  $\mathbf{Z}_K$  é  $f(\mathbf{Z}_K)$ , para todo  $K = 1, \dots, k$ .

**Aproximação Linear:** sob essa abordagem, primeiramente estimamos os valores  $\delta_i$ , dados pela expressão (4.67), para  $i = 1, \dots, m$ , que indicam o quanto o valor do *portfolio* é afetado pela variação de cada fator de risco. Para obter esses valores, selecionamos variações “típicas”  $\Delta_1, \dots, \Delta_m$  para cada fator de risco. Então, o valor  $\delta_i$  é dado por

$$\delta_i = \frac{f(Z_1, \dots, Z_i, \dots, Z_m) - f(Z_1, \dots, Z_i + \Delta_i, \dots, Z_m)}{\Delta_i}, \quad (4.67)$$

para  $i = 1, \dots, m$ . É fácil ver que  $\delta_i$  representa a inclinação média da função  $f(\cdot)$  através da distância  $\Delta_i$ . Portanto, se  $f(\cdot)$  não é uma função linear para o  $i$ -ésimo fator de risco o valor de  $\delta_i$  dependerá da escolha de  $\Delta_i$ .

Assim, tendo os valores de  $\delta_i$ , para todo  $i = 1, \dots, m$ , o segundo passo é obter a estimativa para o valor do *portfolio*. Uma aproximação  $\hat{f}(\cdot)$  para o valor do *portfolio* é dada por

$$\hat{f}(Z_1, \dots, Z_m) = f(Z_{AM,1}, \dots, Z_{AM,m}) - \sum_{i=1}^m (Z_i - Z_{AM,i})\delta_i,$$



onde  $\mathbf{Z}_{AM} = (Z_{AM,1}, \dots, Z_{AM,m})'$  representa o vetor de  $m$  fatores de risco para a situação atual do mercado e  $\delta_j$ , para todo  $i = 1, \dots, m$ , é dado pela expressão (4.67).

Em geral, a reavaliação completa é mais precisa e mais eficiente do que a aproximação linear. Em nosso caso, a reavaliação do *portfolio* é simples e envolve apenas o logaritmo dos preços dos ativos. Lembramos que um *portfolio* é um conjunto qualquer de investimentos. Logo, em alguns casos, a reavaliação do *portfolios* pode exigir vários cálculos. A aproximação linear mostra-se mais eficiente, na prática, somente quando a reavaliação completa do *portfolio* exige muitos cálculos, quando o número de cenários é muito maior do que o número de fatores de risco ou quando os valores  $\delta_i$  forem utilizados posteriormente para outro propósito.

### 4.6.3 Identificação dos Máximos e Mínimos para Fatores de Risco Individuais

Sejam  $\mathcal{P}$  um *portfolio* fixo e  $\mathbf{Z} = (Z_1, \dots, Z_m)'$  o vetor dos  $m$  fatores de risco que influenciam no valor deste *portfolio*. O método tradicional de realizar um teste de estresse consiste em construir os cenários baseando-se em dados históricos dos fatores de risco. Define-se o *período de observação histórico* como o período no qual a série temporal é considerada, por exemplo, 1 ou 10 anos. O período de observação histórico é então sobreposto por janelas de igual duração, por exemplo, 1 ou 10 dias. Para cada janela temporal determinamos as mudanças de fatores de risco  $\Delta Z_i$ , para cada  $i = 1, \dots, m$ . O máximo ou mínimo das mudanças de fatores de risco entre todas as janelas temporais é então fixado como a mudança de fator de risco  $\Delta Z_i$ , para cada  $i = 1, \dots, m$ .

A mudança dos fatores de risco é geralmente definida como a variação entre o primeiro e o último dia da janela temporal e é denominada *Start to End (StE)*. Note que, se assumimos que os fatores de risco são os logaritmos dos preços dos ativos, então as mudanças de fatores de risco são os log-retornos de período igual a duração da janela. Alternativamente, a mudança dos fatores de risco pode ser definida como o máximo das variações entre dois pontos quaisquer da janela temporal. Essa variação é denominada *drawdown (DD)*.

É fácil ver que, se escolhermos a variação mínima dentro de cada janela, então  $\Delta Z_i$ , para cada  $i = 1, \dots, m$ , representa a redução máxima dos fatores de risco. Se o máximo é selecionado, então estamos considerando o aumento máximo nos fatores de risco. Podemos também considerar  $\Delta Z_i$ , para cada  $i = 1, \dots, m$ , como máximo para o valor absoluto dessas variações. Neste caso, obtemos a variação máxima dos fatores de risco não levando em consideração se a mudança é positiva ou negativa.

### 4.6.4 Perda Máxima - *MaxLoss*

A *perda máxima*, denotada por *MaxLoss (maximum loss)* foi introduzido por Studer (1997) e pode ser vista como uma maneira sistemática de realizar um teste de estresse. Esta é uma medida de risco que pode ser interpretada como a pior perda possível que pode ocorrer. Para muitos *portfolios* o pior cenário pode não existir, dado que a função para calcular o valor do *portfolio* pode ser não-limitada inferiormente. Como a probabilidade de ocorrência de um cenário no qual o mercado a muito tempo não se encontra é muito baixa, restringe-se a atenção à cenários em um certo domínio de admissibilidade, também denominado *região de confiança*. Isto é, um certo conjunto de cenários com alta

probabilidade de ocorrência (veja Studer, 1997). Por exemplo, se assumirmos que os log-retornos possuem uma função de distribuição elíptica, como a  $t$ -Student ou a normal, então o domínio de admissibilidade é um elipsóide (veja Studer, 1997). A seguir apresentamos a definição formal de perda máxima.

**Definição 4.12. (Perda Máxima).** Dado um domínio de admissibilidade  $A$ , a *perda máxima* de um *portfolio* contido em  $A$  é dado por

$$MaxLoss_A(f) \equiv f(\mathbf{Z}_{AM}) - \min_{\mathbf{Z} \in A} \{f(\mathbf{Z})\},$$

onde  $f(\cdot)$  é a função que determina o preço do *portfolio* e  $\mathbf{Z}_{AM} = (Z_{AM,1}, \dots, Z_{AM,m})'$  representa o vetor de  $m$  fatores de risco para a situação atual do mercado.

Em contraste com o  $VaR_p$ , a perda máxima é uma medida de risco coerente (veja Studer, 1997). Portanto, a propriedade de subaditividade é sempre válida. Além disso, o  $MaxLoss$  informa, além da dimensão da perda, o cenário em que a pior perda ocorre.

**Observação 4.6.** Note que para calcular a perda máxima deve ser escolhida uma região de confiança fechada  $A$ , com uma dada probabilidade  $p$  de ocorrência. Dessa forma, uma maneira equivalente de definir a perda máxima é

$$MaxLoss_A(f) = \max \left\{ f(\mathbf{Z}_{AM}) - f(\mathbf{Z}) : \mathbf{Z} \in A \text{ e } \mathbb{P}(A) = p \right\}.$$

Observamos que a expressão  $f(\mathbf{Z}_{AM}) - f(\mathbf{Z})$  pode ser comparada à expressão (4.35), isto é, esta expressão representa a perda  $L$  (ou  $-L$ , se  $\mathbf{Z}$  é medido em um instante de tempo anterior a  $AM$ ) do *portfolio*. Dizemos que um *portfolio* é *linear* se a expressão (4.35) é uma função linear em relação a cada uma das mudanças de fatores de risco. O teorema que segue fornece a expressão da medida  $MaxLoss$  para um *portfolio* linear.

**Teorema 4.1.** *Sejam  $\mathcal{P}$  um portfolio linear e  $f(\cdot)$  a função que determina o valor do portfolio, então  $f(\mathbf{X}) = \mathbf{a}'\mathbf{X}$ , onde  $\mathbf{a} \in \mathbb{R}^m$  é um vetor de constantes reais e  $\mathbf{X} \in \mathbb{R}^m$  é o vetor de mudanças de fatores de risco. Segue que, dado um nível de confiança  $p$ , a perda máxima do portfolio é dada por*

$$MaxLoss = -\sqrt{c_p} \sqrt{\mathbf{a}'\Sigma\mathbf{a}}, \quad (4.68)$$

onde  $\Sigma$  é a matriz de variância-covariância das mudanças de fatores de risco e  $c_p$  é o  $p$ -quantil da função de distribuição  $\chi^2$  com  $m$  graus de liberdade. O pior cenário é dado por

$$\mathbf{Z}^* = -\frac{\sqrt{c_p}}{\sqrt{\mathbf{a}'\Sigma\mathbf{a}}} \Sigma\mathbf{a}. \quad (4.69)$$

**Prova:** Veja Studer (1997), Teorema 3.15. □

Na próxima subseção apresentamos alguns critérios de admissibilidade para um *portfolio* qualquer.

#### 4.6.5 Alguns Critérios de Admissão para Cenários

Apresentamos nesta subseção alguns dos critérios mais utilizados na literatura para seleção de cenários admissíveis.

## Critérios de Admissão que Ignoram a Correlação

Sejam  $\mathcal{P}$  um *portfolio* fixo e  $\mathbf{Z} = (Z_1, \dots, Z_m)' \in \mathbb{R}^m$  o vetor de fatores de risco que influenciam no valor deste *portfolio*. Sejam  $\sigma_i$  o desvio padrão de  $Z_i$ , para todo  $i = 1, \dots, m$  e  $\mathbf{Z}_{AM} = (Z_{AM,1}, \dots, Z_{AM,m})'$  o vetor de  $m$  fatores de risco para a situação atual do mercado. Para cada inteiro positivo  $k$  podemos definir os seguintes domínio de admissibilidade:

- **“Cubóide”**. Este domínio admite todos os cenários  $\mathbf{Z} = (Z_1, \dots, Z_m)'$  que satisfazem

$$Z_{AM,i}(1 - k\sigma_i) \leq Z_i \leq Z_{AM,i}(1 + k\sigma_i),$$

para cada fator de risco  $Z_i$ . Esta condição admite somente cenários que estão situados dentro de um cubóide  $m$ -dimensional com arestas  $2k\sigma_i Z_{AM,i}$  e centro  $\mathbf{Z}_{AM}$ . Note que, quanto maior o valor de  $k$ , mais cenários são admitidos, e portanto, maior será o *MaxLoss*.

- **“Cubóide em escala logarítmica”**. Esta condição admite todos os cenários  $\mathbf{Z} = (Z_1, \dots, Z_m)'$  que satisfazem

$$Z_{AM,i}e^{-k\sigma_i} \leq Z_i \leq Z_{AM,i}e^{k\sigma_i},$$

para cada fator de risco  $Z_i$ . Esta condição é útil para *portfolios* de ações pois, em geral, assume-se que os fatores de risco possuem função de distribuição log-normal.

## Critérios de Admissão que Consideram a Correlação

Assuma que o vetor de mudanças de fatores de risco  $\mathbf{X}$  possui matriz de variância-covariância dada por

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_m^2 \end{pmatrix},$$

onde  $\sigma_i^2 = \mathbb{E}((\Delta Z_i - \mu_i)^2)$  e  $\sigma_{ij} = \sigma_{ji} = \mathbb{E}((\Delta Z_i - \mu_i)(\Delta Z_j - \mu_j))$  e  $\mu_i = \mathbb{E}(\Delta Z_i)$ , para todo  $i, j = 1, \dots, m$  e  $\mathbf{X} \equiv \Delta \mathbf{Z} = (\Delta Z_1, \dots, \Delta Z_m)'$ .

Neste caso, admite-se todos os cenários  $\mathbf{Z} = (Z_1, \dots, Z_m)'$  que satisfazem

$$(\mathbf{Z}_{AM} - \mathbf{Z})' \Sigma^{-1} (\mathbf{Z}_{AM} - \mathbf{Z}) \leq k^2.$$

# Capítulo 5

## Simulações

Neste capítulo apresentamos resultados obtidos a partir da análise de séries temporais simuladas. Tais séries temporais são amostras de processos FIEGARCH( $p, d, q$ ), com parâmetros semelhantes aos dos modelos ajustados às séries temporais reais analisadas no Capítulo 6. Além da estimação dos parâmetros dos modelos e da previsão da volatilidade, calculamos também as medidas de risco VaR e ES (Definições 4.10 e 4.11, respectivamente), para essas séries temporais. Apresentamos também um exemplo de análise completa de um *portfolio*, com o mesmo número de ativos daquele proposto no Capítulo 6, utilizando séries temporais simuladas.

### 5.1 Geração de Processos FIEGARCH

Nesta seção apresentamos resultados relacionados à geração dos processos FIEGARCH. No que segue, as séries temporais representadas pelos modelos denotados por  $M_i$ , para  $i \in \{1, \dots, 4\}$ , representam os log-retornos dos preços das ações dos ativos financeiros e as séries temporais representadas pelo modelo denotado por  $M_5$  representam os log-retornos dos índices do mercado financeiro.

Os seguintes critérios foram utilizados para gerar as séries temporais:

- Os parâmetros dos modelos gerados são semelhantes àqueles obtidos na análise das séries temporais reais apresentadas no Capítulo 6. Para gerar as séries temporais utilizamos a representação dada na Definição 2.7. Isto é, consideramos a representação da volatilidade, proposta por Bollerslev e Mikkelsen (1996), dada por

$$\ln(\sigma_{i,t}^2) = \omega + \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})(1 - \mathcal{B})^d} g(Z_{i,t-1}), \quad \text{para cada } i = 1, \dots, 5, \quad (5.1)$$

onde o índice  $i$  refere-se ao modelo  $M_i$ ,  $i \in \{1, \dots, 5\}$ , considerado. Em todos os casos, assumimos que  $Z_{i,t} \sim \mathcal{N}(0, 1)$ . Os parâmetros da expressão (5.1), para cada modelo, são apresentados na Tabela 5.1.

- Os coeficientes do polinômio  $\lambda(\cdot)$ , definido pela expressão (2.52), são apresentados na Proposição 2.4 e dados pela expressão (2.53).
- Para cada modelo foram realizadas 1.000 replicações, com dois diferentes tamanhos amostrais, sendo eles,  $n = 2.000$  e  $n = 5.000$ . Utilizamos o tamanho amostral  $n = 2.000$  por ser aproximadamente aquele das séries temporais reais analisadas neste

trabalho. O valor  $n = 5.000$  foi utilizado para verificar as propriedades assintóticas dos estimadores considerados.

- Para truncar o polinômio  $\lambda(\cdot)$  utilizamos  $m = 10.000$  nos casos em que  $d > 0,4$  e  $m = 50.000$  para o caso em que  $d < 0,4$ . Estes valores foram escolhidos considerando-se os resultados apresentados na Tabela C.1 e no Teorema 2.6.

Tabela 5.1: Valores dos Parâmetros da Expressão (5.1) Utilizados na Simulação dos Modelos FIEGARCH( $p, d, q$ ).

	$\omega$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\alpha_1$	$\theta$	$\gamma$	$d$
M1	0,00	0,45	-	-	-	-	-0,14	0,38	0,45
M2	0,00	0,90	-	-	-	0,80	0,04	0,38	0,45
M3	0,00	0,22	0,18	0,47	-0,45	-	-0,04	0,40	0,26
M4	0,00	0,58	-	-	-	-	-0,11	0,33	0,42
M5	0,00	0,71	-	-	-	-	-0,17	0,28	0,34

**Observação 5.1.** Para todos os modelos simulados consideramos  $\omega = 0$  pois, pelas expressões (2.44) e (2.45),

$$X_t = Z_t \exp \left\{ \omega + \sum_{k=0}^{\infty} \lambda_{d,k} g(Z_{t-1-k}) \right\}, \quad \text{para todo } t \in \mathbb{Z},$$

com  $\lambda_{d,k}$ , para todo  $k \geq 0$ , dado pela expressão (2.53). Segue que, para gerar séries temporais com  $\omega \neq 0$ , basta multiplicar a série temporal gerada pelo método utilizado neste trabalho (isto é, considerando  $\omega = 0$ ) por  $e^\omega$ .

Tabela 5.2: Valores dos Coeficientes  $\lambda_{d,k}$ , do Polinômio  $\lambda(\cdot)$ , Utilizados para Gerar os Modelos  $M_i$ , para  $i \in \{1, \dots, 5\}$ .

$k \setminus d$	$\lambda_{d,k}$				
	0,45 (M1)	0,45 (M2)	0,26 (M3)	0,42 (M4)	0,34 (M5)
0	1,0000	1,0000	1,0000	1,0000	1,0000
10	0,2722	0,3130	-0,0905	0,3302	0,3806
100	0,0736	0,0164	0,0165	0,0786	0,0640
1.000	0,0207	0,0228	0,0030	0,0205	0,0138
5.000	0,0085	0,0094	0,0009	0,0081	0,0048
10.000	0,0058	0,0064	0,0005	0,0054	0,0030
25.000	-	-	0,0003	-	0,0016
50.000	-	-	0,0002	-	0,0010

**Nota:** Os valores marcados com o sinal (-) não foram utilizados na simulação.

A Tabela 5.2 apresenta os valores dos coeficientes  $\lambda_{d,k}$ , para  $k = 0, 10, 100, 1.000, 5.000, 10.000, 20.000$  e  $50.000$  para cada um dos modelos simulados. Note que, para os

primeiros valores de  $k$ , o decaimento dos coeficientes é mais rápido do que para os valores de  $k \geq 5.000$ .

Lembramos que, pelo Teorema 2.5,  $\lambda_{d,k} = o(k^d)$ . Na Tabela 5.3 apresentamos os valores do quociente  $\frac{\lambda_{d,k}}{k^d}$  para aqueles valores de  $\lambda_{d,k}$  considerados na Tabela 5.2. Note que, para  $k = 10.000$  e  $50.000$  este quociente é próximo de zero.

Tabela 5.3: Valores do Quociente  $\frac{\lambda_{d,k}}{k^d}$ , para os Valores  $\lambda_{d,k}$  Apresentados na Tabela 5.2.

$k \setminus d$	0, 45	0, 45	0, 26	0, 42	0, 34
0	1,0000	1,0000	1,0000	1,0000	1,0000
10	0,1031	0,1149	0,0588	0,1358	0,1920
100	0,0093	0,0105	0,0050	0,0114	0,0134
1.000	0,0009	0,0010	0,0005	0,0011	0,0013
5.000	0,0002	0,0002	0,0001	0,0002	0,0003
10.000	0,0001	0,0001	0,0000	0,0001	0,0001
25.000	-	-	0,0000	-	0,0001
50.000	-	-	0,0000	-	0,0000

**Nota:** Os valores marcados com o sinal (-) não foram utilizados na simulação.

A Figura 5.1 apresenta os gráficos dos coeficientes do polinômio  $\lambda(\cdot)$ , definidos pela expressão (2.53), para os modelos M1 e M3 simulados. Apesar de não apresentarmos aqui os gráficos destes coeficientes para os demais modelos, observamos que, para todos os modelos simulados, o valor dos coeficientes apresenta um decaimento rápido a zero. Além disso, para valores de  $k > 5.000$  estes coeficientes assumem valores próximos de zero.

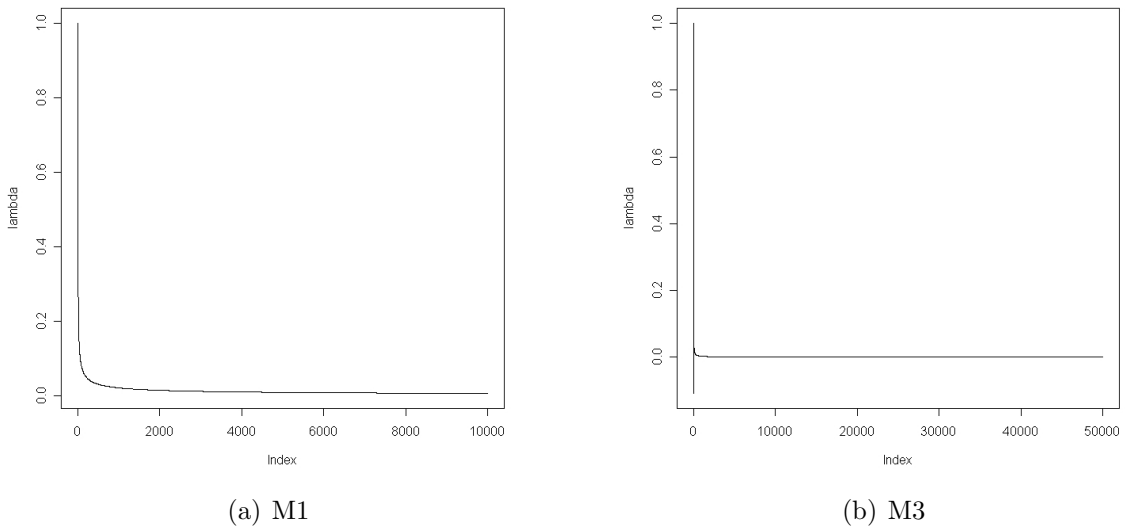


Figura 5.1: Gráfico dos Coeficientes do Polinômio  $\lambda(\cdot)$ , Definido pela Expressão (2.52), para as Modelos Simulados.

## 5.2 Ajuste de Modelos FIEGARCH

Nesta seção apresentamos os resultados dos ajustes de modelos FIEGARCH( $p, d, q$ ) às séries temporais simuladas. Os resultados aqui apresentados nos dão uma idéia da qualidade do ajuste dos modelos às séries temporais reais analisadas no Capítulo 6. Ressaltamos ainda, que a qualidade do ajuste do modelo está diretamente relacionada às propriedades assintóticas do estimador. Como mencionamos na Subseção 3.2.2, resultados sobre as propriedades assintóticas do estimador de quase-máxima verossimilhança, para os processos FIEGARCH, ainda permanecem um problema em aberto.

Para o ajuste dos modelos FIEGARCH( $p, d, q$ ) utilizamos a função *fgarch* do programa S-Plus. O método utilizado para a estimação dos parâmetros é o método da quase-verossimilhança, descrito na Subseção 3.2.2. A rotina utilizada considera a expressão (2.46) para definir o logaritmo da volatilidade. Para cada série temporal, utilizamos  $n - 10$  valores, onde  $n$  é o tamanho amostral, para o ajuste dos modelos e reservamos 10 valores (portanto, um horizonte de 10 dias) para estimar os erros de previsão (veja Seção 5.4). Lembramos que, na Proposição 2.3 mostramos que, sob certas condições (veja expressão (2.47)), a expressão utilizada para gerar as séries temporais e a utilizada pelo S-Plus são equivalentes, sendo a última a mais geral.

Nas Tabelas 5.4 - 5.6 apresentamos os valores dos parâmetros utilizados na simulação dos processos (equivalentes aos apresentados na Tabela 5.1), e seus respectivos estimadores. Os critérios MAE e MSE são, respectivamente, o erro absoluto médio e o erro quadrático médio de estimação, definidos na Subseção 3.1.4. Os valores apresentados nas tabelas referem-se à média dos valores estimados para as 1.000 replicações de cada modelo.

Comparando os valores utilizados na simulação com aqueles ajustados às séries temporais percebemos que, com exceção de M2, para os demais modelos o valor estimado para o parâmetro  $d$  está muito próximo do valor simulado, o que é confirmado pelo baixo valor do erro quadrático médio. Quanto aos demais parâmetros, a precisão dos seus ajustes varia de modelo para modelo. Observamos valores altos do erro quadrático médio. Entretanto, para os modelos M1, M4 e M5, esse valor é menor ou igual a 0,0773.

Tabela 5.4: Valor Estimado para os Parâmetros do Modelo M1.

Modelo	Parâmetro	Estimador	MAE	MSE
M1 $n = 2.000$	$a = -0,30$	-0,1778	0,1222	0,0155
	$\beta_1 = 0,45$	0,7249	0,2749	0,0773
	$\psi_0 = 0,38$	0,2232	0,1568	0,0254
	$\gamma_0 = -0,14$	-0,0870	0,0530	0,0031
	$d = 0,45$	0,4329	0,0483	0,0038
M1 $n = 5.000$	$a = -0,30$	-0,1779	0,1221	0,0151
	$\beta_1 = 0,45$	0,7195	0,2695	0,0733
	$\psi_0 = 0,38$	0,2235	0,1565	0,0248
	$\gamma_0 = -0,14$	-0,0856	0,0544	0,0031
	$d = 0,45$	0,4423	0,0292	0,0014

Tabela 5.5: Valor Estimado para os Parâmetros dos Modelos M2, M3 e M4.

Modelo	Parâmetro	Estimador	MAE	MSE
M2 $n = 2.000$	$a = -0,06$	-0,3053	0,2453	0,0627
	$\beta_1 = 0,90$	-0,0140	0,9148	0,8816
	$\psi_0 = 0,38$	0,0003	0,3797	0,1469
	$\psi_1 = -0,30$	0,3813	0,6813	0,4685
	$\gamma_0 = 0,04$	0,0003	0,0429	0,0026
	$\gamma_1 = -0,03$	0,0415	0,0718	0,0061
	$d = 0,45$	0,5968	0,1570	0,0299
M2 $n = 5.000$	$a = -0,06$	-0,3066	0,2466	0,0617
	$\beta_1 = 0,90$	-0,0398	0,9398	0,8990
	$\psi_0 = 0,38$	0,0038	0,3762	0,1427
	$\psi_1 = -0,30$	0,3800	0,6800	0,4639
	$\gamma_0 = 0,04$	0,0012	0,0392	0,0019
	$\gamma_1 = -0,03$	0,0387	0,0687	0,0051
	$d = 0,45$	0,6157	0,1659	0,0293
M3 $n = 2.000$	$a = -0,32$	-0,1137	0,2063	0,0438
	$\beta_1 = 0,22$	1,6573	1,4373	2,2583
	$\beta_2 = 0,18$	-1,4983	1,6851	3,2459
	$\beta_3 = 0,47$	1,0480	0,6636	0,5639
	$\beta_4 = -0,45$	-0,4350	0,1907	0,0618
	$\psi_0 = 0,40$	0,1429	0,2571	0,0681
	$\gamma_0 = -0,04$	-0,0146	0,0258	0,0008
	$d = 0,26$	0,2497	0,1284	0,0240
M3 $n = 5.000$	$a = -0,32$	-0,1198	0,2002	0,0408
	$\beta_1 = 0,22$	1,5648	1,3448	1,9310
	$\beta_2 = 0,18$	-1,3433	1,5237	2,5138
	$\beta_3 = 0,47$	0,9398	0,4781	0,2707
	$\beta_4 = -0,45$	-0,4047	0,1261	0,0260
	$\psi_0 = 0,40$	0,1508	0,2492	0,0633
	$\gamma_0 = -0,04$	-0,0158	0,0243	0,0007
	$d = 0,26$	0,2838	0,0870	0,0113
M4 $n = 2.000$	$a = -0,26$	-0,1658	0,0942	0,0093
	$\beta_1 = 0,58$	0,7606	0,1806	0,0340
	$\psi_0 = 0,33$	0,2078	0,1222	0,0157
	$\gamma_0 = -0,11$	-0,0737	0,0365	0,0016
	$d = 0,42$	0,4190	0,0449	0,0033
M4 $n = 5.000$	$a = -0,26$	-0,1671	0,0929	0,0088
	$\beta_1 = 0,58$	0,7548	0,1748	0,0311
	$\psi_0 = 0,33$	0,2097	0,1203	0,0148
	$\gamma_0 = -0,11$	-0,0723	0,0377	0,0015
	$d = 0,42$	0,4319	0,0307	0,0015



Tabela 5.6: Valor Estimado para os Parâmetros do Modelo M5.

Modelo	Parâmetro	Estimador	MAE	MSE
M5 $n = 2.000$	$a = -0,22$	-0,1551	0,0649	0,0046
	$\beta_1 = 0,71$	0,7926	0,0826	0,0076
	$\psi_0 = 0,28$	0,1942	0,0858	0,0080
	$\gamma_0 = -0,17$	-0,1204	0,0496	0,0027
	$d = 0,34$	0,3883	0,0575	0,0049
M5 $n = 5.000$	$a = -0,22$	-0,1561	0,0639	0,0042
	$\beta_1 = 0,71$	0,7921	0,0821	0,0070
	$\psi_0 = 0,28$	0,1957	0,0843	0,0074
	$\gamma_0 = -0,17$	-0,1199	0,0501	0,0026
	$d = 0,34$	0,3895	0,0506	0,0033

### 5.3 Estimação do Parâmetro de Longa Dependência

Nesta seção apresentamos resultados relacionados à estimação do parâmetro  $d$  de longa dependência, dos processos FIEGARCH( $p, d, q$ ), para as séries temporais simuladas. Para estimar o parâmetro de longa dependência utilizamos os métodos descritos na Subseção 3.1.6. Para os métodos R/S, R/S <sub>$m$</sub> , KPSS e V/S, utilizamos  $l = 40$  e  $l = 100$  como número mínimo de observações para cada bloco e para o estimador GPH utilizamos valores de  $\alpha \in \{0, 5; 0, 6; 0, 7; 0, 8; 0, 9\}$ .

No Teorema 2.9 mostramos que, se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo FIEGARCH( $p, d, q$ ), tal que  $\mathbb{E}([\ln(Z_t^2)]^2) < \infty$  e  $\mathbb{E}(Z_t \ln(Z_t^2)) < \infty$ , para todo  $t \in \mathbb{Z}$ , então o processo  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, 0$ ) com inovações que, em geral, são não-Gaussianas. Partindo deste resultado, para estimar o parâmetro  $d$ , utilizamos ambas as séries temporais,  $\{X_{i,t}^2\}_{t=1}^n$  e  $\{\ln(X_{i,t}^2)\}_{t=1}^n$ , para cada  $i = 1, \dots, 5$ , onde  $i$  refere-se ao modelo utilizado para simulação e  $n \in \{2.000, 5.000\}$ . Para cada um dos cinco modelos consideramos as 1.000 replicações realizadas. Para cada uma das séries temporais, estimamos o valor do parâmetro  $d$  através dos diferentes métodos citados e em seguida, consideramos a média aritmética destes valores.

Nas Tabelas 5.7 e 5.8 apresentamos as médias dos valores estimados para o parâmetro  $d$  para as séries temporais  $\{X_{i,t}^2\}_{t=1}^{2.000}$ , para cada  $i = 1, \dots, 5$ , onde  $i$  refere-se ao modelo utilizado para simulação, e ainda, o erro absoluto médio (MAE) e o erro quadrático médio (MSE). Nas Tabelas 5.9 e 5.10 apresentamos esses valores para as séries temporais  $\{\ln(X_{i,t}^2)\}_{t=1}^{2.000}$ , para cada  $i = 1, \dots, 5$ .

Observamos que, em todos dos casos, os resultados encontrados para as séries temporais  $\{X_{i,t}^2\}_{t=1}^n$  e  $\{\ln(X_{i,t}^2)\}_{t=1}^n$ , para cada  $i = 1, \dots, 5$ , foram semelhantes. Em geral, os valores encontrados pelos diferentes estimadores, foram todos muito próximos. Entretanto, os resultados obtidos por esses métodos foram muito ruins se comparados com os valores obtidos pelo método da quase-verossimilhança (veja Tabelas 5.4 - 5.6). Apesar dos altos valores do vício, o Método GPH, com  $\alpha = 0, 8$  foi o que, na maioria dos casos, apresentou os melhores resultados (em relação ao erro quadrático médio). Além disso, o vício apresentado por esse estimador para os processos FIEGARCH foi, em geral, menor que aquele encontrado por Hurvich *et al.* (2005), para os processos LMSV. O erro quadrático médio apresentado pelos estimadores, para o tamanho amostral  $n = 5.000$  foi menor do que o apresentado para  $n = 2.000$ , como era esperado.

Tabela 5.7: Valor Médio Estimado do Parâmetro  $d$ , para as Séries Temporais  $\{X_{i,t}^2\}_{t=1}^n$ , para  $i = 1, 2, 3$ , onde  $i$  Refere-se ao Modelo  $Mi$ .

Método	R/S		R/S Modificado		KPSS		V/S		GPH				
	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$\alpha = 0, 5$	$\alpha = 0, 6$	$\alpha = 0, 7$	$\alpha = 0, 8$	$\alpha = 0, 9$
M1, $n = 2.000, d = 0, 45$													
$\hat{d}$	0,2506	0,2665	0,2272	0,2496	0,2173	0,2542	0,2111	0,2480	0,2705	0,2670	0,3015	0,2992	0,1599
MAE	0,1994	0,1838	0,2228	0,2005	0,2330	0,2120	0,2389	0,2071	0,1881	0,1843	0,1504	0,1517	0,2901
MSE	0,0416	0,0388	0,0513	0,0448	0,0655	0,0646	0,0643	0,0574	0,0464	0,0401	0,0263	0,0256	0,0856
M1, $n = 5.000, d = 0, 45$													
$\hat{d}$	0,2514	0,2703	0,2267	0,2540	0,2179	0,2520	0,2137	0,2456	0,2788	0,2618	0,2848	0,3173	0,1799
MAE	0,1986	0,1797	0,2233	0,1960	0,2321	0,2003	0,2363	0,2048	0,1753	0,1884	0,1656	0,1329	0,2701
MSE	0,0401	0,0341	0,0504	0,0401	0,0594	0,0514	0,0596	0,0501	0,0388	0,0396	0,0296	0,0191	0,0737
M2, $n = 2.000, d = 0, 45$													
$\hat{d}$	0,2710	0,2998	0,2498	0,2836	0,2533	0,2860	0,2410	0,2844	0,3462	0,3296	0,2820	0,2448	0,1310
MAE	0,1790	0,1508	0,2002	0,1665	0,1984	0,1910	0,2091	0,1789	0,1266	0,1268	0,1684	0,2054	0,3190
MSE	0,0338	0,0281	0,0417	0,0328	0,0503	0,0551	0,0511	0,0466	0,0240	0,0216	0,0319	0,0444	0,1028
M2, $n = 5.000, d = 0, 45$													
$\hat{d}$	0,2750	0,3074	0,2529	0,2918	0,2638	0,3009	0,2514	0,2978	0,3410	0,3445	0,3025	0,2591	0,1539
MAE	0,1750	0,1426	0,1971	0,1582	0,1862	0,1549	0,1986	0,1543	0,1219	0,1092	0,1478	0,1909	0,2961
MSE	0,0313	0,0225	0,0395	0,0269	0,0406	0,0347	0,0437	0,0319	0,0213	0,0153	0,0239	0,0375	0,0882
M3, $n = 2.000, d = 0, 26$													
$\hat{d}$	0,1553	0,1443	0,1402	0,1356	0,0921	0,0999	0,0992	0,1014	0,1068	0,1156	0,1949	0,2039	0,1104
MAE	0,1049	0,1190	0,1198	0,1267	0,1716	0,1837	0,1623	0,1721	0,1617	0,1463	0,0711	0,0601	0,1496
MSE	0,0126	0,0181	0,0158	0,0200	0,0378	0,0497	0,0326	0,0415	0,0365	0,0264	0,0070	0,0049	0,0232
M3, $n = 5.000, d = 0, 26$													
$\hat{d}$	0,1516	0,1432	0,1362	0,1353	0,0919	0,1026	0,0973	0,1047	0,1107	0,1077	0,1527	0,2156	0,1269
MAE	0,1084	0,1168	0,1238	0,1247	0,1684	0,1627	0,1627	0,1578	0,1523	0,1524	0,1073	0,0461	0,1331
MSE	0,0123	0,0153	0,0159	0,0171	0,0340	0,0357	0,0304	0,0320	0,0306	0,0264	0,0129	0,0027	0,0181

Tabela 5.8: Valor Médio Estimado do Parâmetro  $d$ , para as Séries Temporais,  $\{X_{i,t}^2\}_{t=1}^n$ , para  $i = 4, 5$ , onde  $i$  Refere-se ao Modelo Mi.

Método	R/S		R/S Modificado		KPSS		V/S		GPH				
	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$\alpha = 0,5$	$\alpha = 0,6$	$\alpha = 0,7$	$\alpha = 0,8$	$\alpha = 0,9$
M4, $n = 2.000, d = 0,42$													
$\hat{d}$	0,2531	0,2644	0,2294	0,2472	0,2190	0,2411	0,2177	0,2391	0,2515	0,2691	0,3221	0,2933	0,1538
MAE	0,0325	0,0572	0,0390	0,0547	0,0865	0,1308	0,0723	0,1055	0,0944	0,0680	0,0720	0,0487	0,2664
MSE	0,0017	0,0051	0,0024	0,0048	0,0120	0,0263	0,0085	0,0174	0,0145	0,0075	0,0083	0,0041	0,0725
M4, $n = 5.000, d = 0,42$													
$\hat{d}$	0,2556	0,2666	0,2306	0,2500	0,2173	0,2420	0,2173	0,2393	0,2670	0,2658	0,3046	0,3205	0,1747
MAE	0,0208	0,0348	0,0316	0,0337	0,0688	0,0901	0,0597	0,0750	0,0751	0,0508	0,0519	0,0612	0,2453
MSE	0,0007	0,0019	0,0014	0,0018	0,0076	0,0126	0,0057	0,0089	0,0089	0,0041	0,0043	0,0052	0,0610
M5, $n = 2.000, d = 0,34$													
$\hat{d}$	0,2493	0,2459	0,2239	0,2277	0,2028	0,2092	0,2083	0,2118	0,2150	0,2602	0,3352	0,3039	0,1737
MAE	0,0327	0,0549	0,0425	0,0574	0,0886	0,1234	0,0736	0,1002	0,1005	0,0664	0,0827	0,0559	0,1669
MSE	0,0016	0,0047	0,0027	0,0052	0,0129	0,0246	0,0089	0,0165	0,0161	0,0071	0,0107	0,0053	0,0295
M5, $n = 5.000, d = 0,34$													
$\hat{d}$	0,2471	0,2424	0,2209	0,2256	0,2010	0,2081	0,2075	0,2107	0,2154	0,2320	0,3037	0,3248	0,1878
MAE	0,0221	0,0365	0,0398	0,0435	0,0745	0,0927	0,0633	0,0785	0,0833	0,0547	0,0526	0,0659	0,1523
MSE	0,0008	0,0021	0,0021	0,0028	0,0087	0,0136	0,0063	0,0098	0,0113	0,0047	0,0044	0,0061	0,0241

Tabela 5.9: Valor Médio Estimado do Parâmetro  $d$ , para as Séries Temporais,  $\{\ln(X_{i,t}^2)\}_{t=1}^n$ , para  $i = 1, 2, 3$ , onde  $i$  Refere-se ao Modelo  $M_i$ .

Método	R/S		R/S Modificado		KPSS		V/S		GPH				
	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$\alpha = 0,5$	$\alpha = 0,6$	$\alpha = 0,7$	$\alpha = 0,8$	$\alpha = 0,9$
M1, $n = 2.000, d = 0,45$													
$\hat{d}$	0,2701	0,2861	0,2451	0,2690	0,2367	0,2765	0,2299	0,2688	0,3047	0,2987	0,3215	0,2938	0,1526
MAE	0,1799	0,1648	0,2049	0,1814	0,2152	0,1995	0,2203	0,1925	0,1560	0,1519	0,1287	0,1562	0,2974
MSE	0,0345	0,0329	0,0439	0,0382	0,0585	0,0598	0,0570	0,0525	0,0340	0,0281	0,0189	0,0256	0,0891
M1, $n = 5.000, d = 0,45$													
$\hat{d}$	0,2698	0,2890	0,2439	0,2728	0,2384	0,2755	0,2325	0,2678	0,3147	0,2955	0,3082	0,3143	0,1711
MAE	0,1802	0,1610	0,2061	0,1772	0,2117	0,1815	0,2175	0,1840	0,1391	0,1545	0,1418	0,1357	0,2789
MSE	0,0332	0,0281	0,0431	0,0334	0,0512	0,0443	0,0517	0,0429	0,0264	0,0269	0,0214	0,0190	0,0781
M2, $n = 2.000, d = 0,45$													
$\hat{d}$	0,2811	0,3110	0,2585	0,2944	0,2724	0,3104	0,2570	0,3062	0,3707	0,3397	0,2825	0,2332	0,1228
MAE	0,1689	0,1415	0,1915	0,1564	0,1823	0,1881	0,1939	0,1677	0,1136	0,1145	0,1675	0,2168	0,3272
MSE	0,0307	0,0261	0,0386	0,0303	0,0455	0,0562	0,0466	0,0439	0,0196	0,0175	0,0306	0,0482	0,1076
M2, $n = 5.000, d = 0,45$													
$\hat{d}$	0,2841	0,3154	0,2604	0,2995	0,2755	0,3138	0,2619	0,3099	0,3621	0,3580	0,3050	0,2495	0,1438
MAE	0,1659	0,1346	0,1896	0,1505	0,1748	0,1481	0,1881	0,1455	0,1022	0,0940	0,1450	0,2005	0,3062
MSE	0,0284	0,0207	0,0367	0,0249	0,0372	0,0328	0,0403	0,0296	0,0159	0,0117	0,0224	0,0407	0,0940
M3, $n = 2.000, d = 0,26$													
$\hat{d}$	0,1553	0,1443	0,1402	0,1356	0,0921	0,0999	0,0992	0,1014	0,1068	0,1156	0,1949	0,2039	0,1056
MAE	0,1049	0,1190	0,1198	0,1267	0,1716	0,1837	0,1623	0,1721	0,1617	0,1463	0,0711	0,0601	0,1544
MSE	0,0126	0,0181	0,0158	0,0200	0,0378	0,0497	0,0326	0,0415	0,0365	0,0264	0,0070	0,0049	0,0245
M3, $n = 5.000, d = 0,26$													
$\hat{d}$	0,1516	0,1432	0,1362	0,1353	0,0919	0,1026	0,0973	0,1047	0,1107	0,1077	0,1527	0,2156	0,1216
MAE	0,1084	0,1168	0,1238	0,1247	0,1684	0,1627	0,1627	0,1578	0,1523	0,1524	0,1073	0,0461	0,1384
MSE	0,0123	0,0153	0,0159	0,0171	0,0340	0,0357	0,0304	0,0320	0,0306	0,0264	0,0129	0,0027	0,0194

Tabela 5.10: Valor Médio Estimado do Parâmetro  $d$ , para as Séries Temporais,  $\{\ln(X_{i,t}^2)\}_{t=1}^n$ , para  $i = 4, 5$ , onde  $i$  Refere-se ao Modelo  $M_i$ .

Método	R/S		R/S Modificado		KPSS		V/S		GPH				
	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$l = 40$	$l = 100$	$\alpha = 0, 5$	$\alpha = 0, 6$	$\alpha = 0, 7$	$\alpha = 0, 8$	$\alpha = 0, 9$
M4, $n = 2.000, d = 0, 42$													
$\hat{d}$	0,2699	0,2805	0,2448	0,2633	0,2378	0,2640	0,2350	0,2589	0,2867	0,2986	0,3374	0,2842	0,1458
MAE	0,0357	0,0638	0,0346	0,0581	0,0899	0,1370	0,0732	0,1122	0,0912	0,0683	0,0803	0,0357	0,2742
MSE	0,0020	0,0062	0,0019	0,0052	0,0125	0,0287	0,0085	0,0194	0,0135	0,0072	0,0085	0,0020	0,0759
M4, $n = 5.000, d = 0, 42$													
$\hat{d}$	0,2721	0,2814	0,2460	0,2652	0,2348	0,2599	0,2341	0,2567	0,2970	0,2938	0,3252	0,3129	0,1635
MAE	0,0244	0,0415	0,0235	0,0362	0,0672	0,0944	0,0569	0,0792	0,0756	0,0521	0,0661	0,0531	0,2565
MSE	0,0009	0,0026	0,0009	0,0020	0,0071	0,0138	0,0051	0,0096	0,0089	0,0043	0,0056	0,0034	0,0661
M5, $n = 2.000, d = 0, 34$													
$\hat{d}$	0,2726	0,2681	0,2458	0,2504	0,2262	0,2346	0,2319	0,2361	0,2575	0,3029	0,3588	0,2964	0,1654
MAE	0,0364	0,0586	0,0344	0,0558	0,0872	0,1295	0,0710	0,1037	0,0870	0,0686	0,0995	0,0422	0,1746
MSE	0,0020	0,0054	0,0018	0,0050	0,0123	0,0262	0,0083	0,0173	0,0122	0,0070	0,0122	0,0026	0,0311
M5, $n = 5.000, d = 0, 34$													
$\hat{d}$	0,2702	0,2623	0,2426	0,2461	0,2233	0,2310	0,2295	0,2325	0,2521	0,2720	0,3393	0,3209	0,1784
MAE	0,0230	0,0362	0,0243	0,0363	0,0687	0,0914	0,0573	0,0772	0,0708	0,0447	0,0797	0,0610	0,1616
MSE	0,0008	0,0020	0,0009	0,0020	0,0077	0,0134	0,0054	0,0095	0,0080	0,0032	0,0076	0,0043	0,0264

## 5.4 Previsão

Nesta seção apresentamos resultados relacionados à previsão da volatilidade para um dado horizonte  $h$ , a partir dos modelos FIEGARCH( $p, d, q$ ) ajustados às séries temporais simuladas. Tais resultados são importantes pois nos fornecem uma idéia dos erros de previsão que encontramos na análise das séries temporais reais. Além disso, os valores previstos são utilizados para o cálculo da medida de risco VaR, que apresentamos ainda neste capítulo. Lembramos que, para o ajuste dos modelos, utilizamos os primeiros  $n - 10$  valores das séries temporais, onde  $n$  é o tamanho amostral, e reservamos os 10 últimos valores para previsão.

Os métodos de previsão para processos FIEGARCH( $p, d, q$ ) são apresentados na Seção 3.4. Nas Tabelas 5.11 - 5.15, apresentamos a média dos valores  $X_{i,t+h}^2$  e  $\sigma_{i,t+h}$  simulados (portanto, conhecidos), para  $h = 1, \dots, 10$ ,  $t = n - 10$ ,  $n \in \{2.000, 5.000\}$  e  $i = 1, \dots, 5$ , juntamente com a média dos valores previstos  $\hat{X}_{i,t+h}^2$  e  $\hat{\sigma}_{i,t+h}$ , o erro absoluto médio (MAE) e o erro quadrático médio (MSE). Observamos que, para  $\sigma_t$ , o erro quadrático médio de previsão variou entre 0,0037 e 0,3227. Enquanto que, para  $X_t^2$ , o erro quadrático médio de previsão variou entre 2,0725 e 12,1379. Sendo, em ambos os casos, maior para o tamanho amostral  $n = 5.000$ .

Tabela 5.11: Média, para as 1.000 Replicações, dos Valores Previstos  $\hat{X}_{1,t+h}^2$  e  $\hat{\sigma}_{1,t+h}$ , para  $h = 1, \dots, 10$ .

$h$	$\hat{\sigma}_{1,t+h}$	$\sigma_{1,t+h}$	MAE	MSE	$\hat{X}_{1,t+h}^2$	$X_{1,t+h}^2$	MAE	MSE
$n = 2.000$								
1	1,0797	1,0682	0,0923	0,0196	1,3442	1,3462	1,3096	6,4641
2	1,0765	1,0698	0,0668	0,0100	1,4255	1,3281	1,2998	6,5364
3	1,0722	1,0753	0,1314	0,0350	1,4290	1,3040	1,3298	6,6864
4	1,0679	1,0842	0,1717	0,0590	1,2863	1,2810	1,3309	7,4938
5	1,0641	1,0884	0,1931	0,0832	1,2250	1,2612	1,2948	5,4407
6	1,0609	1,0780	0,1958	0,0782	1,2711	1,2448	1,2939	7,1591
7	1,0582	1,0638	0,1994	0,0748	1,2631	1,2314	1,3265	6,8337
8	1,0561	1,0645	0,2093	0,0848	1,3821	1,2205	1,3622	9,8998
9	1,0543	1,0664	0,2220	0,1042	1,3176	1,2115	1,2817	5,3681
10	1,0528	1,0697	0,2256	0,1040	1,2616	1,2041	1,2818	5,9562
$n = 5.000$								
1	1,0680	1,0851	0,4082	0,2983	1,3020	1,3623	1,5534	7,1325
2	1,0658	1,0875	0,4035	0,2888	1,2911	1,3614	1,5619	8,4032
3	1,0620	1,0869	0,3923	0,2645	1,2718	1,3244	1,4454	6,3784
4	1,0581	1,0807	0,3809	0,2523	1,2518	1,3737	1,4988	8,1271
5	1,0545	1,0773	0,3626	0,2327	1,2339	1,2271	1,3409	5,6984
6	1,0515	1,0794	0,3524	0,2241	1,2187	1,3217	1,4644	8,4122
7	1,0490	1,0766	0,3477	0,2138	1,2061	1,3278	1,4328	7,4953
8	1,0470	1,0698	0,3386	0,2147	1,1958	1,2262	1,3499	6,6178
9	1,0453	1,0689	0,3322	0,2034	1,1873	1,2526	1,3638	6,0584
10	1,0439	1,0629	0,3275	0,1933	1,1803	1,3677	1,4547	7,5470

Tabela 5.12: Média, para as 1.000 Replicações, dos Valores Previstos  $\hat{X}_{2,t+h}^2$  e  $\hat{\sigma}_{2,t+h}$ , para  $h = 1, \dots, 10$ .

$h$	$\hat{\sigma}_{2,t+h}$	$\sigma_{2,t+h}$	MAE	MSE	$\hat{X}_{2,t+h}^2$	$X_{2,t+h}^2$	MAE	MSE
$n = 2.000$								
1	1,0639	1,0621	0,0420	0,0037	1,2451	1,2083	1,1642	3,4973
2	1,0556	1,0557	0,1050	0,0215	1,2103	1,2753	1,2172	5,4859
3	1,0529	1,0553	0,1210	0,0267	1,1982	1,2802	1,2110	4,1119
4	1,0507	1,0592	0,1284	0,0314	1,1895	1,3534	1,2941	6,3247
5	1,0492	1,0654	0,1344	0,0347	1,1835	1,2831	1,2528	6,2395
6	1,0480	1,0709	0,1459	0,0412	1,1788	1,2307	1,2162	7,6093
7	1,0470	1,0665	0,1481	0,0438	1,1748	1,3429	1,2934	6,0690
8	1,0461	1,0621	0,1492	0,0419	1,1715	1,2735	1,2475	6,0593
9	1,0453	1,0663	0,1568	0,0470	1,1685	1,2809	1,2546	5,2953
10	1,0447	1,0669	0,1584	0,0471	1,1659	1,2694	1,2591	4,7061
$n = 5.000$								
1	1,0602	1,0504	0,3429	0,2034	1,2492	1,2016	1,3914	6,2332
2	1,0485	1,0409	0,3271	0,1856	1,2053	1,3984	1,4738	7,1674
3	1,0465	1,0409	0,3207	0,1801	1,1945	1,1774	1,3377	5,6101
4	1,0449	1,0534	0,3297	0,1856	1,1870	1,2650	1,3555	5,3730
5	1,0436	1,0453	0,3246	0,1815	1,1816	1,2594	1,3464	5,7636
6	1,0426	1,0485	0,3215	0,1806	1,1772	1,2964	1,3819	6,6531
7	1,0418	1,0505	0,3213	0,1762	1,1736	1,2597	1,3554	10,0613
8	1,0411	1,0533	0,3208	0,1759	1,1705	1,2317	1,2787	4,9245
9	1,0404	1,0520	0,3191	0,1751	1,1678	1,0795	1,2196	3,4894
10	1,0399	1,0557	0,3151	0,1681	1,1654	1,2293	1,3138	4,8063

Tabela 5.13: Média, para as 1.000 Replicações, dos Valores Previstos  $\hat{X}_{3,t+h}^2$  e  $\hat{\sigma}_{3,t+h}$ , para  $h = 1, \dots, 10$ .

$h$	$\hat{\sigma}_{3,t+h}$	$\sigma_{3,t+h}$	MAE	MSE	$\hat{X}_{3,t+h}^2$	$X_{3,t+h}^2$	MAE	MSE
$n = 2.000$								
1	1,0271	1,0186	0,0977	0,0163	1,0936	1,1053	1,0790	3,0490
2	1,0235	1,0173	0,0839	0,0123	1,0835	1,1119	1,0578	2,3588
3	1,0202	1,0171	0,1263	0,0275	1,0700	1,1006	1,0796	3,1225
4	1,0192	1,0203	0,1368	0,0324	1,0606	1,1309	1,1351	3,8089
5	1,0159	1,0197	0,1364	0,0319	1,0474	1,0595	1,0542	2,5950
6	1,0117	1,0198	0,1600	0,0431	1,0350	1,0701	1,0705	2,8101
7	1,0112	1,0200	0,1596	0,0424	1,0320	1,0578	1,0767	3,2305
8	1,0125	1,0183	0,1555	0,0417	1,0337	1,0639	1,0457	2,5841
9	1,0111	1,0174	0,1619	0,0474	1,0301	1,0781	1,0819	2,9045
10	1,0102	1,0179	0,1587	0,0428	1,0280	1,0261	1,0648	3,3794
$n = 5.000$								
1	1,0305	1,0307	0,2159	0,0776	1,0990	1,1323	1,1717	3,4025
2	1,0299	1,0170	0,2089	0,0741	1,0942	1,0546	1,0953	2,8701
3	1,0265	1,0226	0,2013	0,0682	1,0788	0,9851	1,0299	2,4174
4	1,0231	1,0276	0,1953	0,0641	1,0650	1,0691	1,0579	2,6751
5	1,0198	1,0107	0,1787	0,0527	1,0528	1,0206	1,0038	2,0725
6	1,0168	1,0211	0,1694	0,0500	1,0422	0,9999	1,0094	2,3678
7	1,0150	1,0192	0,1615	0,0425	1,0365	1,1499	1,1421	3,7239
8	1,0138	1,0058	0,1539	0,0412	1,0331	1,0456	1,0573	4,3785
9	1,0130	1,0241	0,1636	0,0512	1,0308	0,9883	0,9933	2,1910
10	1,0124	1,0144	0,1606	0,0441	1,0294	1,0335	1,0728	2,9688

Tabela 5.14: Média, para as 1.000 Replicações, dos Valores Previstos  $\hat{X}_{4,t+h}^2$  e  $\hat{\sigma}_{4,t+h}$ , para  $h = 1, \dots, 10$ .

$h$	$\hat{\sigma}_{4,t+h}$	$\sigma_{4,t+h}$	MAE	MSE	$\hat{X}_{4,t+h}^2$	$X_{4,t+h}^2$	MAE	MSE
$n = 2.000$								
1	1,0667	1,0524	0,0833	0,0150	1,2882	1,2947	1,3198	11,2799
2	1,0642	1,0555	0,0552	0,0067	1,2762	1,1426	1,1718	3,5852
3	1,0604	1,0529	0,1122	0,0272	1,2574	1,2231	1,2583	11,6102
4	1,0564	1,0468	0,1405	0,0398	1,2381	1,3325	1,3305	6,8781
5	1,0528	1,0439	0,1614	0,0582	1,2204	1,1350	1,2092	4,0434
6	1,0495	1,0494	0,1851	0,0744	1,2051	1,3484	1,2643	5,2165
7	1,0467	1,0468	0,1955	0,0774	1,1922	1,2982	1,2685	5,9811
8	1,0444	1,0556	0,2042	0,0817	1,1813	1,2555	1,2388	5,4190
9	1,0423	1,0559	0,2074	0,0804	1,1721	1,2646	1,2439	5,2147
10	1,0406	1,0565	0,2137	0,0830	1,1644	1,2346	1,1912	4,1274
$n = 5.000$								
1	1,0734	1,0644	0,3730	0,2492	1,3053	1,2768	1,4835	6,7601
2	1,0716	1,0580	0,3664	0,2382	1,2966	1,2673	1,4857	8,1326
3	1,0681	1,0570	0,3619	0,2306	1,2788	1,1967	1,3901	5,8127
4	1,0642	1,0554	0,3599	0,2267	1,2593	1,2731	1,4193	6,8665
5	1,0606	1,0533	0,3532	0,2149	1,2410	1,2355	1,3265	7,9746
6	1,0574	1,0562	0,3424	0,2069	1,2250	1,2417	1,3266	6,1027
7	1,0546	1,0598	0,3347	0,2007	1,2113	1,2213	1,3641	7,3400
8	1,0522	1,0606	0,3229	0,1820	1,1997	1,3829	1,4504	7,4792
9	1,0503	1,0558	0,3221	0,1753	1,1900	1,2920	1,3471	6,3358
10	1,0486	1,0636	0,3300	0,1898	1,1818	1,4040	1,5179	9,1731

Tabela 5.15: Média, para as 1.000 Replicações, dos Valores Previstos  $\hat{X}_{5,t+h}^2$  e  $\hat{\sigma}_{5,t+h}$ , para  $h = 1, \dots, 10$ .

$h$	$\hat{\sigma}_{5,t+h}$	$\sigma_{5,t+h}$	MAE	MSE	$\hat{X}_{5,t+h}^2$	$X_{5,t+h}^2$	MAE	MSE
$n = 2.000$								
1	1,0918	1,0753	0,0893	0,0183	1,3874	1,3684	1,3568	6,1264
2	1,0880	1,0794	0,0658	0,0106	1,3679	1,3079	1,3009	5,1825
3	1,0823	1,0747	0,1218	0,0363	1,3388	1,2941	1,3289	8,4456
4	1,0764	1,0730	0,1688	0,0585	1,3086	1,3191	1,3436	7,1734
5	1,0708	1,0690	0,1932	0,0750	1,2808	1,1912	1,2287	5,1002
6	1,0659	1,0716	0,2169	0,0961	1,2564	1,2296	1,3034	6,8245
7	1,0615	1,0660	0,2265	0,0964	1,2355	1,1693	1,1770	4,4440
8	1,0578	1,0600	0,2330	0,1027	1,2179	1,3049	1,3416	5,7768
9	1,0547	1,0535	0,2396	0,1071	1,2031	1,3334	1,3313	5,8234
10	1,0519	1,0566	0,2426	0,1086	1,1906	1,1696	1,2551	4,9864
$n = 5.000$								
1	1,0483	1,0679	0,4185	0,3227	1,2583	1,2706	1,5127	7,8570
2	1,0457	1,0655	0,4030	0,3069	1,2435	1,2973	1,5094	9,7420
3	1,0421	1,0686	0,3986	0,2928	1,2238	1,3077	1,4925	9,6828
4	1,0385	1,0656	0,3900	0,2774	1,2036	1,3548	1,4738	6,9389
5	1,0352	1,0649	0,3808	0,2625	1,1852	1,4764	1,5597	8,8766
6	1,0323	1,0716	0,3743	0,2541	1,1691	1,2696	1,4020	12,1379
7	1,0299	1,0830	0,3764	0,2628	1,1554	1,3011	1,3784	8,9583
8	1,0278	1,0803	0,3745	0,2523	1,1439	1,4781	1,5727	12,1276
9	1,0261	1,0830	0,3682	0,2441	1,1342	1,4159	1,5161	11,4614
10	1,0247	1,0871	0,3669	0,2476	1,1261	1,3463	1,4573	11,8576



## 5.5 Estimação da Medida de Risco VaR

Nesta seção apresentamos os resultados obtidos na estimação das medidas de risco VaR para as séries temporais simuladas. Para o cálculo do VaR utilizamos os primeiros  $n - 10$  valores das séries temporais simuladas, onde  $n$  é o tamanho amostral, para a estimação da média e da variância, tanto condicional como não-condicional. O valor real de  $\text{VaR}_{p,i,t+1}$  é dado por

$$\text{VaR}_{p,i,t+1} = \Phi^{-1}(p) \times \sigma_{i,t+1}, \quad \text{para cada } i = 1, \dots, 5, \quad (5.2)$$

onde  $t = n - 10$ ,  $\Phi(\cdot)$  é a função de distribuição normal padrão e  $\sigma_{i,t+1}^2$  é o valor da variância condicional gerado através do modelo  $M_i$ , para cada  $i = 1, \dots, 5$ .

Para o cálculo do VaR, para cada uma das 1.000 replicações dos modelos simulados, utilizamos as abordagens descritas na Seção 4.5.2. Para a abordagem econométrica, utilizamos modelos EGARCH( $p, q$ ), com  $p = 1 = q$  e os modelos FIEGARCH( $p, d, q$ ) ajustados às séries temporais, apresentados na Seção 5.2.

Tabela 5.16: Médias dos Valores Estimados da Medida de Risco VaR, sob Diferentes Abordagens, com  $p \in \{0,95; 0,99\}$  para as Séries Temporais Geradas pelos Modelos M1 e M2.

Abordagem	VaR <sub>0,95</sub>	$\widehat{\text{VaR}}_{0,95}$	MAE	MSE	VaR <sub>0,99</sub>	$\widehat{\text{VaR}}_{0,99}$	MAE	MSE
M1, $n = 2.000$								
Empírico	1,6726	1,7996	0,2891	0,1378	2,3656	2,8355	0,5773	0,5257
Normal	1,6726	1,8458	0,3069	0,1534	2,3656	2,6108	0,4332	0,3051
EWMA	1,6726	1,7159	0,3493	0,2386	2,3656	2,4370	0,4900	0,4750
EGARCH	1,6726	1,7707	0,2694	0,1722	2,3656	2,5044	0,3810	0,3445
FIEGARCH	1,6726	1,7759	0,2851	0,1936	2,3656	2,5117	0,4033	0,3872
M1, $n = 5.000$								
Empírico	1,6673	1,7967	0,2883	0,1287	2,3581	2,8577	0,5858	0,5206
Normal	1,6673	1,8433	0,3063	0,1453	2,3581	2,6072	0,4326	0,2895
EWMA	1,6673	1,6912	0,3474	0,2270	2,3581	2,4029	0,4890	0,4548
EGARCH	1,6673	1,7533	0,2649	0,1521	2,3581	2,4797	0,3746	0,3043
FIEGARCH	1,6673	1,7567	0,2758	0,1621	2,3581	2,4845	0,3901	0,3243
M2, $n = 2.000$								
Empírico	1,6765	1,7773	0,2480	0,1014	2,3710	2,7020	0,4439	0,3285
Normal	1,6765	1,8004	0,2548	0,1077	2,3710	2,5462	0,3602	0,2152
EWMA	1,6765	1,7287	0,2702	0,1353	1,6765	2,4425	0,2146	0,2694
EGARCH	1,6765	1,7494	0,2146	0,0925	2,3710	2,4743	0,3036	0,1850
FIEGARCH	1,6765	1,7499	0,2209	0,1035	2,3710	2,4749	0,3125	0,2071
M2, $n = 5.000$								
Empírico	1,6691	1,7689	0,2573	0,1094	2,3607	2,7229	0,4743	0,3625
Normal	1,6691	1,7938	0,2664	0,1170	2,3607	2,5369	0,3765	0,2337
EWMA	1,6691	1,7192	0,2745	0,1352	2,3607	2,4291	0,3884	0,2697
EGARCH	1,6691	1,7342	0,2172	0,0902	2,3607	2,4527	0,3072	0,1805
FIEGARCH	1,6691	1,7439	0,2382	0,1114	2,3607	2,4664	0,3369	0,2228

Tabela 5.17: Médias dos Valores Estimados da Medida de Risco VaR, sob Diferentes Abordagens, com  $p \in \{0,95; 0,99\}$  para as Séries Temporais Geradas pelos Modelos M3, M4 e M5.

Abordagem	VaR <sub>0,95</sub>	$\widehat{\text{VaR}}_{0,95}$	MAE	MSE	VaR <sub>0,99</sub>	$\widehat{\text{VaR}}_{0,99}$	MAE	MSE
M3, $n = 2.000$								
Empírico	1,6517	1,6852	0,1453	0,0326	2,3360	2,5002	0,2566	0,0965
Normal	1,6517	1,7028	0,1474	0,0330	2,3360	2,4086	0,2077	0,0654
EWMA	1,6517	1,6864	0,1999	0,0683	2,3360	2,3858	0,2805	0,1349
EGARCH	1,6517	1,6886	0,1468	0,0414	2,3360	2,3882	0,2076	0,0828
FIEGARCH	1,6517	1,6895	0,1571	0,0466	2,3360	2,3894	0,2222	0,0932
M3, $n = 5.000$								
Empírico	1,6594	1,6968	0,1437	0,0326	2,3469	2,5208	0,2572	0,0967
Normal	1,6594	1,7129	0,1483	0,0339	2,3469	2,4224	0,2093	0,0674
EWMA	1,6594	1,6805	0,1882	0,0598	2,3469	2,3778	0,2654	0,1190
EGARCH	1,6594	1,6976	0,1401	0,0359	2,3469	2,4009	0,1982	0,0717
FIEGARCH	1,6594	1,6950	0,1446	0,0382	2,3469	2,3973	0,2045	0,0765
M4, $n = 2.000$								
Empírico	1,6637	1,7643	0,2677	0,1194	2,3530	2,7593	0,5227	0,4399
Normal	1,6637	1,8064	0,2843	0,1330	2,3530	2,5549	0,4016	0,2642
EWMA	1,6637	1,6987	0,3153	0,1958	2,3530	2,4115	0,4438	0,3912
EGARCH	1,6637	1,7478	0,2426	0,1327	2,3530	2,4719	0,3432	0,2654
FIEGARCH	1,6637	1,7546	0,2633	0,1559	2,3530	2,4816	0,3724	0,3117
M4, $n = 5.000$								
Empírico	1,6719	1,7967	0,2775	0,1197	2,3645	2,8391	0,5613	0,4714
Normal	1,6719	1,8383	0,2942	0,1340	2,3645	2,5998	0,4154	0,2667
EWMA	1,6719	1,7078	0,3234	0,1901	2,3645	2,4250	0,4561	0,3805
EGARCH	1,6719	1,7610	0,2491	0,1314	2,3645	2,4906	0,3524	0,2629
FIEGARCH	1,6719	1,7656	0,2683	0,1512	2,3645	2,4971	0,3795	0,3025
M5, $n = 2.000$								
Empírico	1,6763	1,8038	0,3036	0,1473	2,3708	2,9105	0,6522	0,6476
Normal	1,6763	1,8672	0,3308	0,1708	2,3708	2,6411	0,4665	0,3384
EWMA	1,6763	1,7402	0,3740	0,2926	2,3708	2,4790	0,5258	0,5869
EGARCH	1,6763	1,7863	0,2710	0,1780	2,3708	2,5264	0,3833	0,3560
FIEGARCH	1,6763	1,7959	0,2940	0,2088	2,3708	2,5400	0,4159	0,4176
M5, $n = 5.000$								
Empírico	1,6484	1,8001	0,2909	0,1237	2,3314	2,9232	0,6562	0,5871
Normal	1,6484	1,8593	0,3205	0,1464	2,3314	2,6298	0,4529	0,2917
EWMA	1,6484	1,6606	0,3823	0,2636	2,3314	2,3665	0,5370	0,5277
EGARCH	1,6484	1,7200	0,2529	0,1363	2,3314	2,4326	0,3577	0,2726
FIEGARCH	1,6484	1,7242	0,2742	0,1557	2,3314	2,4386	0,3878	0,3114

Nas Tabelas 5.16 e 5.17 apresentamos os valores reais, obtidos a partir da expressão 5.2, e os valores estimados da medida de risco VaR, para  $p \in \{0,95; 0,99\}$ . Os valores apresentados nas tabelas correspondem à média dos valores estimados para as 1.000 replicações e  $n \in \{2.000, 5.000\}$ . Os critérios MAE e MSE representam, respectivamente, o erro absoluto médio e o erro quadrático médio.

Observamos que, em média, os valores estimados, sob as diferentes abordagens, para a medida de risco VaR, são muito próximos. Em todos os casos, o valor estimado foi maior que o valor real, tanto para  $p = 0,95$  como para  $p = 0,99$ . Observa-se pouca diferença entre os valores do erro quadrático médio quando o tamanho amostral varia de 2.000 para  $n = 5.000$ . Na maioria dos casos, os valores estimados pela abordagem Empírica foram os que apresentaram o maior erro quadrático médio. Entretanto, observou-se pouca diferença entre os valores obtidos pelas abordagens Econométrica e Normal. Além disso, essas duas abordagens apresentaram, em geral, melhores resultados do que aqueles da abordagem RiskMetrics. Devemos levar em consideração que os modelos FIEGARCH ajustados às séries temporais, apresentam, em geral, parâmetros diferentes daqueles utilizados nas simulações. Este fato certamente influenciou nos resultados obtidos.

Tabela 5.18: Médias dos Valores Estimados da Medida de Risco VaR, com  $p \in \{0,95; 0,99\}$  para as Séries Temporais Geradas pelos Modelos M1 e M2.

Abordagem	$-r_{i,t+1}$	$\widehat{\text{VaR}}_{0,95,i,t+1}$	MAE	MSE	$\widehat{\text{VaR}}_{0,99,i,t+1}$	MAE	MSE
M1, $n = 2.000$							
Empírico	0,0649	1,7996	1,8155	4,5099	2,8355	2,7895	9,4227
Normal	0,0649	1,8458	1,8555	4,6781	2,6108	2,5698	8,1359
EWMA	0,0649	1,7159	1,7184	4,4996	2,4370	2,4370	6,9760
EGARCH	0,0649	1,7707	1,7707	3,5713	2,5044	2,5044	7,1437
FIEGARCH	0,0649	1,7759	1,7759	3,6421	2,5117	2,5117	7,2853
M1, $n = 5.000$							
Empírico	-0,0185	1,7967	1,8891	4,7085	2,8577	2,8906	9,8172
Normal	-0,0185	1,8433	1,9311	4,8815	2,6072	2,6468	8,3910
EWMA	-0,0185	1,6912	1,8273	4,6719	2,4029	2,4709	8,0643
EGARCH	-0,0185	1,7533	1,8700	4,8073	2,4797	2,5391	8,2666
FIEGARCH	-0,0185	1,7567	1,8775	4,8570	2,4845	2,5461	8,3692
M2, $n = 2.000$							
Empírico	-0,0015	1,7773	1,8275	4,4530	2,7020	2,7106	8,7487
Normal	-0,0015	1,8004	1,8481	4,5403	2,5462	2,5554	7,8898
EWMA	-0,0015	1,7287	1,7723	4,5726	2,4425	2,4425	6,6325
EGARCH	-0,0015	1,7494	1,7494	3,3268	2,4743	2,4743	6,6547
FIEGARCH	-0,0015	1,7499	1,7499	3,3688	2,4749	2,4749	6,7386
M2, $n = 5.000$							
Empírico	0,0424	1,7689	1,7918	4,2751	2,7229	2,6927	8,6155
Normal	0,0424	1,7938	1,8143	4,3669	2,5369	2,5132	7,6183
EWMA	0,0424	1,7192	1,7721	4,3423	2,4291	2,4236	7,5702
EGARCH	0,0424	1,7342	1,7769	4,3238	2,4527	2,4388	7,5427
FIEGARCH	0,0424	1,7439	1,7902	4,4094	2,4664	2,4560	7,7190

**Nota:** O erro médio absoluto (MAE) e erro quadrático médio (MSE) foram medidos em relação ao log-retorno observado.

Tabela 5.19: Médias dos Valores Estimados da Medida de Risco VaR, com  $p \in \{0,95; 0,99\}$  para as Séries Temporais Geradas pelos Modelos M3, M4 e M5.

Abordagem	$-r_{i,t+1}$	$\widehat{\text{VaR}}_{0,95,i,t+1}$	MAE	MSE	$\widehat{\text{VaR}}_{0,99,i,t+1}$	MAE	MSE
M3, $n = 2.000$							
Empírico	-0,0380	1,6852	1,7752	4,0809	2,5002	2,5521	7,5819
Normal	-0,0380	1,7028	1,7908	4,1427	2,4086	2,4627	7,1071
EWMA	-0,0380	1,6864	1,7726	4,1591	2,3858	2,3858	5,8684
EGARCH	-0,0380	1,6886	1,6886	2,9299	2,3882	2,3882	5,8607
FIEGARCH	-0,0380	1,6895	1,6895	2,9587	2,3894	2,3894	5,9182
M3, $n = 5.000$							
Empírico	0,0353	1,6968	1,7282	3,8915	2,5208	2,4967	7,3138
Normal	0,0353	1,7129	1,7431	3,9452	2,4224	2,4011	6,8317
EWMA	0,0353	1,6805	1,7272	3,9260	2,3778	2,3659	6,7917
EGARCH	0,0353	1,6976	1,7412	3,9708	2,4009	2,3858	6,8787
FIEGARCH	0,0353	1,6950	1,7417	3,9892	2,3973	2,3834	6,9152
M4, $n = 2.000$							
Empírico	-0,0369	1,7643	1,8644	4,6457	2,7593	2,8161	9,4031
Normal	-0,0369	1,8064	1,9018	4,8021	2,5549	2,6143	8,2306
EWMA	-0,0369	1,6987	1,7900	4,7666	2,4115	2,4115	6,6759
EGARCH	-0,0369	1,7478	1,7478	3,4045	2,4719	2,4719	6,8100
FIEGARCH	-0,0369	1,7546	1,7546	3,4854	2,4816	2,4816	6,9718
M4, $n = 5.000$							
Empírico	-0,0103	1,7967	1,8723	4,6180	2,8391	2,8570	9,5968
Normal	-0,0103	1,8383	1,9102	4,7760	2,5998	2,6229	8,2499
EWMA	-0,0103	1,7078	1,8303	4,7099	2,4250	2,4814	8,1433
EGARCH	-0,0103	1,7610	1,8610	4,8067	2,4906	2,5357	8,2991
FIEGARCH	-0,0103	1,7656	1,8690	4,8756	2,4971	2,5459	8,4360
M5, $n = 2.000$							
Empírico	-0,0589	1,8038	1,9263	4,9339	2,9105	2,9897	10,4852
Normal	-0,0589	1,8672	1,9838	5,1815	2,6411	2,7255	8,8529
EWMA	-0,0589	1,7402	1,8625	5,2978	2,4790	2,4790	7,3238
EGARCH	-0,0589	1,7863	1,7863	3,6491	2,5264	2,5264	7,2993
FIEGARCH	-0,0589	1,7959	1,7959	3,7537	2,5400	2,5400	7,5085
M5, $n = 5.000$							
Empírico	0,0525	1,8001	1,8199	4,3492	2,9232	2,8808	9,6276
Normal	0,0525	1,8593	1,8726	4,5625	2,6298	2,5939	7,9827
EWMA	0,0525	1,6606	1,7403	4,3588	2,3665	2,3758	7,6434
EGARCH	0,0525	1,7200	1,7736	4,3689	2,4326	2,4200	7,6058
FIEGARCH	0,0525	1,7242	1,7820	4,4497	2,4386	2,4297	7,7618

**Nota:** O erro médio absoluto (MAE) e erro quadrático médio (MSE) foram medidos em relação ao log-retorno observado.

Na prática, a volatilidade é não observável. Portanto, não é possível obter o verdadeiro valor da medida de risco VaR, para calcular os valores do erro absoluto médio (MAE) e do erro quadrático médio (MSE). O que fazemos então é comparar os valores estimados para essa medida de risco com o valor observado do log-retorno.

Lembramos que o  $\text{VaR}_{p,i,t+1}$ , para cada  $i = 1, \dots, 5$ , representa a perda máxima que pode ocorrer, dada uma certa probabilidade, no instante  $t + 1$ . Enquanto que o valor  $-r_{i,t+1}$ , para cada  $i = 1, \dots, 5$ , representa a perda sofrida pelo ativo  $A_i$  no instante  $t + 1$ .

As Tabelas 5.18 e 5.19 apresentam o valor médio  $-r_{i,t+1}$  (que é conhecido), para cada  $i = 1, \dots, 5$ , e os valores médios estimados do  $\text{VaR}_{p,i,t+1}$ , para  $t = n - 10$ , onde  $n$  é o tamanho amostral e  $p \in \{0,95; ,99\}$ .

Observamos que, em média, os valores estimados, sob as diferentes abordagens, para a medida de risco VaR, são muito próximos. Na maioria dos casos, os valores estimados pela abordagem RiskMetrics (utilizando modelos EWMA) foram os que apresentaram o menor erro quadrático médio. Além disso, os valores estimados por essa abordagem foram mais próximos dos valores reais do que os valores estimados pelas abordagens Empírica e Normal. Observou-se pouca diferença entre os valores obtidos pelas abordagens Econométrica e RiskMetrics. Assim como no caso anterior, devemos levar em consideração que os modelos FIEGARCH ajustados às séries temporais, apresentam, em geral, parâmetros diferentes daqueles utilizados nas simulações. Este fato certamente influenciou nos resultados obtidos.

## 5.6 Exemplo de Análise de um *Portfolio*

Nas seções anteriores tratamos do ajuste e previsão para a volatilidade e para a medida de risco VaR para todas as séries temporais simuladas. Nesta seção apresentamos o exemplo de uma análise completa de um *portfolio*, desde o cálculo dos pesos dos ativos até a estimação das medidas de risco. Como seria muito difícil realizar tal análise para todas as 1.000 replicações de cada modelo, selecionamos apenas uma série temporal, cuja estimação dos parâmetros apresentou um erro quadrático médio pequeno, para compor o *portfolio*.

Por hipótese, os processos simulados através dos modelos M1-M4 representam log-retornos dos preços dos ativos e os processos simulados através do modelo M5 representam os log-retornos do mercado financeiro. Denotamos, respectivamente, por  $A_i$  e  $\{X_{i,t}\}_{t=1}^n$  o ativo financeiro correspondente e a série temporal simulada através do modelo  $M_i$  para  $i = 1, \dots, 4$  e  $n = 2.000$ . Lembramos que a notação  $A_M$  se refere ao mercado financeiro, assim  $\{X_{5,t}\}_{t=1}^n \equiv \{X_{M,t}\}_{t=1}^n$  são os log-retornos dos índices do mercado e  $n = 2.000$ . Para todos os cálculos apresentados consideramos apenas os primeiros 1.990 valores das séries temporais e reservamos os 10 últimos valores para todas as comparações aqui apresentadas.

### 5.6.1 Características das Séries Temporais Simuladas

Nesta seção apresentamos a análise das características das séries temporais dos log-retornos dos preços das ações dos ativos e dos log-retornos dos índices do mercado. Tal análise é importante na seleção de modelos para essas séries temporais. Para verificar se as séries temporais apresentam característica de longa dependência e selecionar o modelo

mais adequado para cada uma delas, utilizamos os métodos descritos na Seção 3.1 (análise da função de autocorrelação, análise de resíduos, etc. ).

A Figura 5.2 apresenta os valores  $\{X_{i,t}\}_{t=1}^{1.990}$  das séries temporais simuladas e os valores  $\{X_{i,t}^2\}_{t=1}^{1.990}$ , para cada  $i = 1, \dots, 5$ . Note que as cinco séries temporais apresentam as características comuns às séries temporais de retornos financeiros como, aparente estacionariedade, média em torno de zero e agrupamentos de volatilidades.

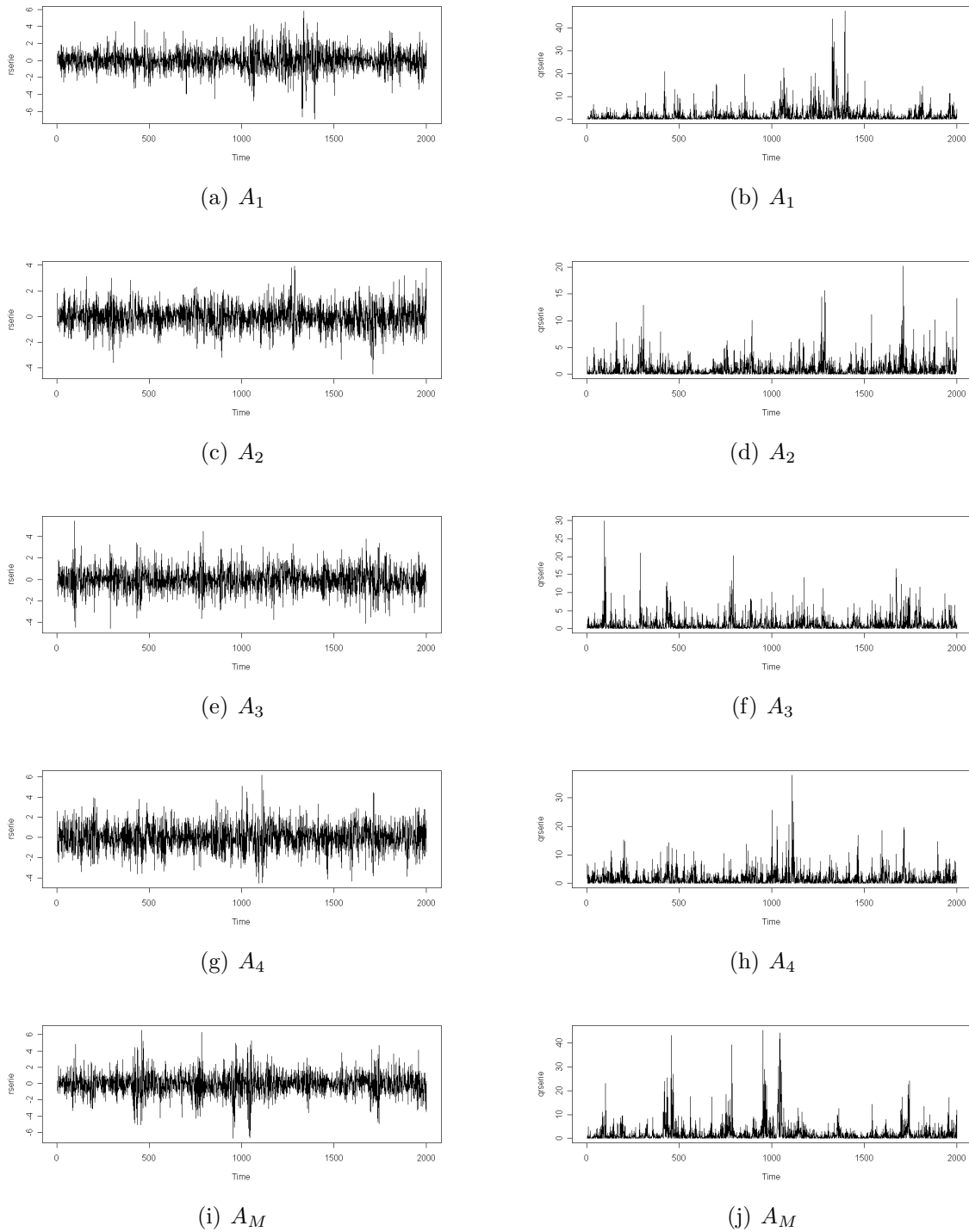


Figura 5.2: Séries Temporais  $\{X_{i,t}\}_{t=1}^{1.990}$  dos Log-retornos dos Ativos e  $\{X_{i,t}^2\}_{t=1}^{1.990}$  (Simuladas).

A Figura 5.3 apresenta o histograma, o Q×Q plot e a função de autocorrelação amostral das cinco séries temporais simuladas. Observamos que, para todas as séries temporais simuladas, a função de distribuição é aproximadamente simétrica. Além disso, pelos gráficos das funções de autocorrelação amostral percebemos que as variáveis aleatórias são não-correlacionadas já que  $\hat{\rho}_X(h) \simeq 0$ , para  $h > 1$ , para as cinco séries temporais simuladas.

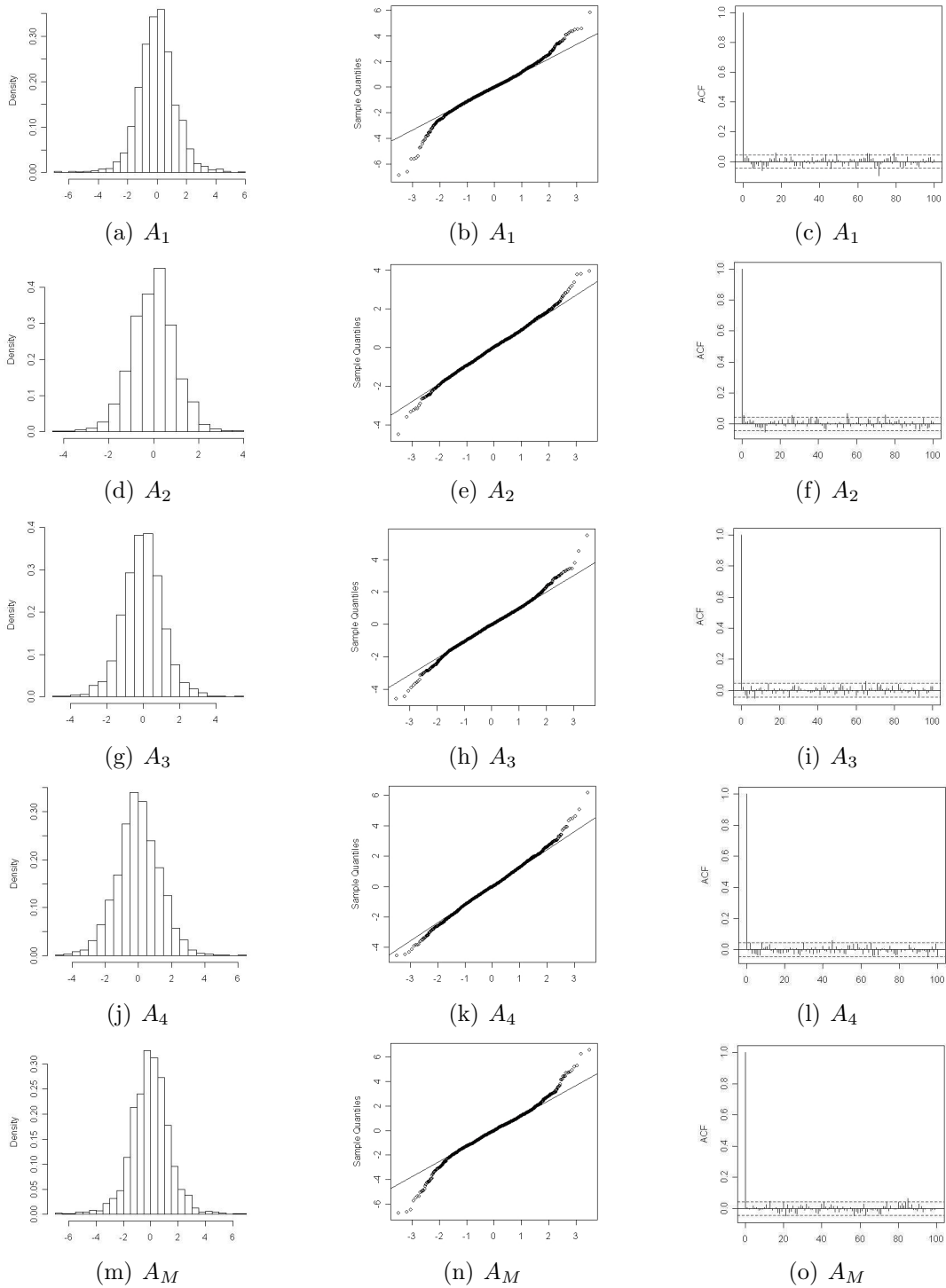


Figura 5.3: Histograma, Q×Q Plot e Função de Autocorrelação Amostral das Séries Temporais dos Log-retornos dos Ativos (Simuladas).

Pela Figura 5.4 podemos perceber que os valores da função de autocorrelação amostral de ambas as séries temporais,  $\{X_{i,t}^2\}_{t=1}^{1.990}$  e  $\{\ln(X_{i,t}^2)\}_{t=1}^{1.990}$ , para cada  $i = 1, \dots, 5$ , são significativos para vários valores de  $h > 1$ , o que indica a presença de longa dependência.

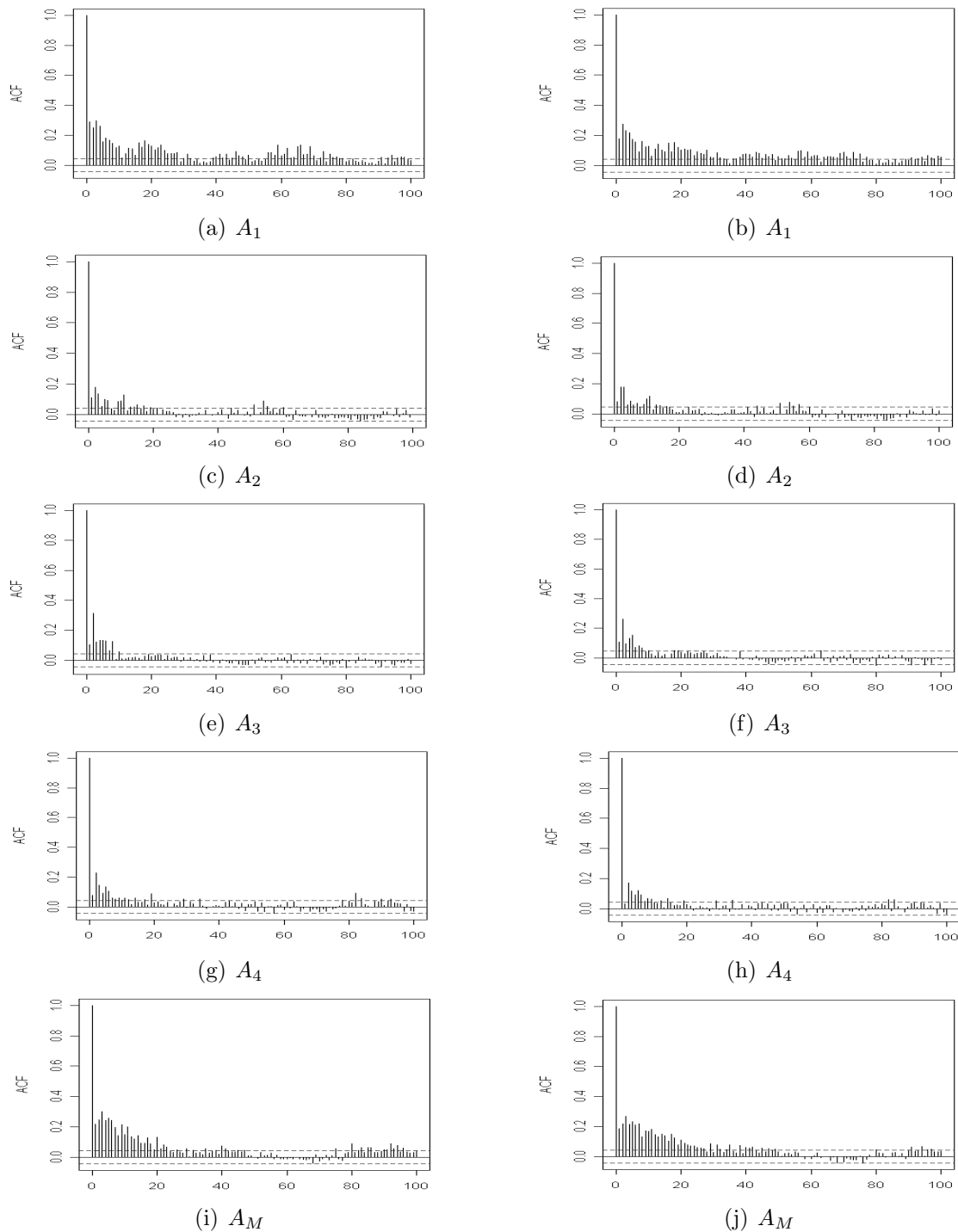
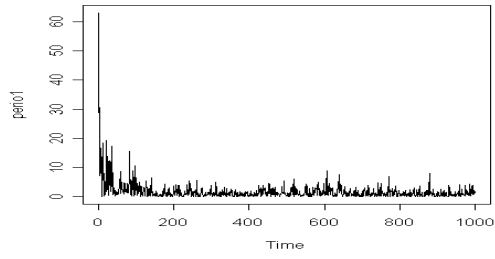
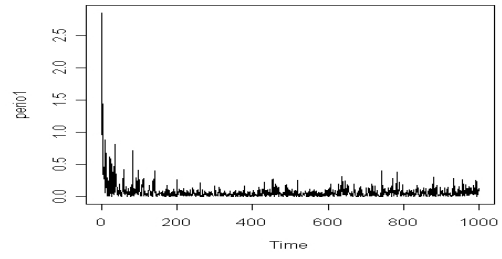


Figura 5.4: Função de Autocorrelação Amostral das Séries Temporais  $\{X_{i,t}^2\}_{t=1}^{1.990}$  e  $\{\ln(X_{i,t}^2)\}_{t=1}^{1.990}$  (Simuladas).

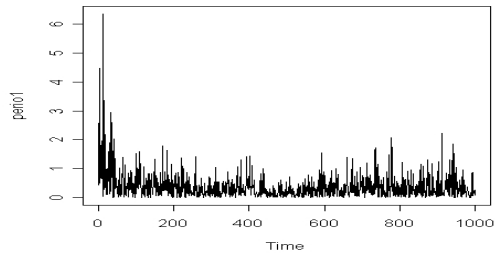




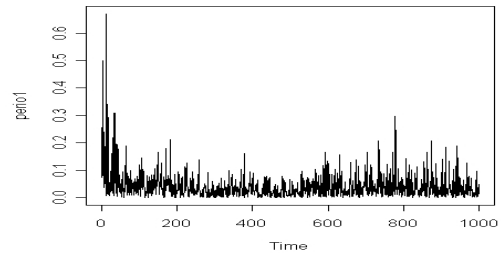
(a)  $A_1$



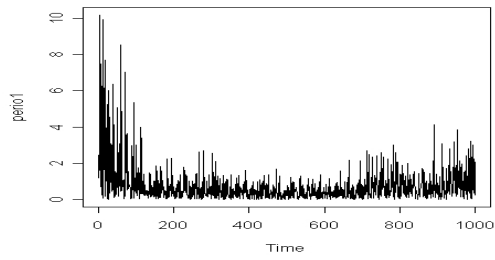
(b)  $A_1$



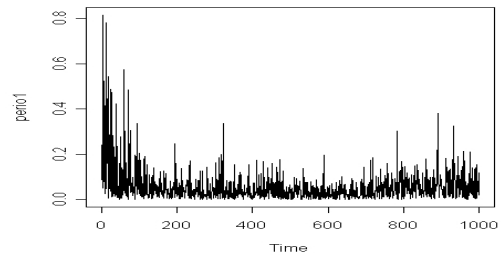
(c)  $A_2$



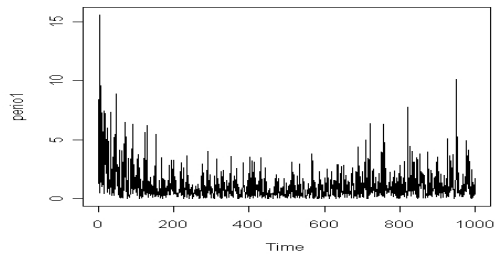
(d)  $A_2$



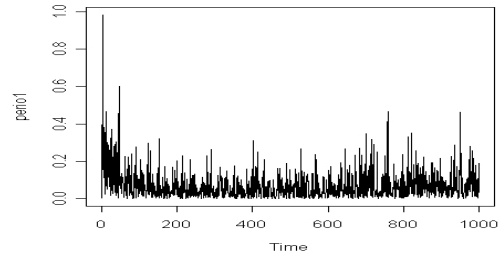
(e)  $A_3$



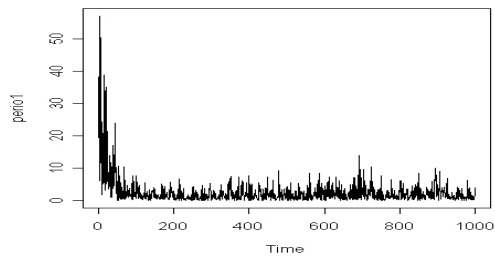
(f)  $A_3$



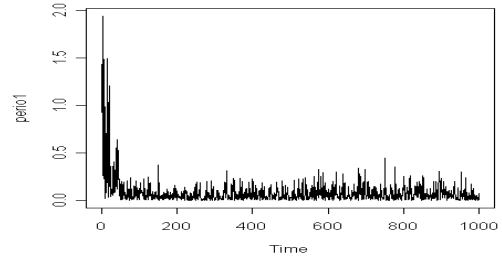
(g)  $A_4$



(h)  $A_4$



(i)  $A_M$



(j)  $A_M$

Figura 5.5: Função Periodograma das Séries Temporais  $\{X_{i,t}^2\}_{t=1}^{1.990}$  e  $\{\ln(X_{i,t}^2)\}_{t=1}^{1.990}$  (Simuladas).

Na Figura 5.5 apresentamos a função periodograma das séries temporais analisadas. Note que, em todos os casos, tal função apresenta um pico próximo de zero, o que é outro indicativo da existência de longa dependência. Esse fato é confirmado pelos valores estimados do parâmetro  $d$  para essas séries temporais. Esses valores foram estimados através dos diferentes métodos apresentados na Subseção 3.1.6 e são apresentados, juntamente com os seus respectivos erros quadráticos médios (valor apresentado entre parênteses), na Tabela 5.20. Observamos que os valores do parâmetro  $d$  obtidos para as séries temporais  $\{X_{i,t}^2\}_{t=1}^{1.990}$  e  $\{\ln(X_{i,t}^2)\}_{t=1}^{1.990}$ , para cada  $i = 1, \dots, 5$  são muito próximos. Esse fato indica que não é necessário analisar as duas séries temporais pois ambas as análises apresentam resultados semelhantes. Entretanto, esses resultados são ruins, se comparados com os obtidos pelo estimador de quase-máxima verossimilhança.

Tabela 5.20: Valor Estimado do Parâmetro  $d$  para as Séries Temporais dos Valores Quadráticos e do Logaritmo dos Valores Quadráticos dos Log-retornos do Índice do Mercado e dos Preços dos Ativos.

$d$	0,45	0,45	0,26	0,42	0,34
Método da Quase-verossimilhança					
	$A_1$	$A_2$	$A_3$	$A_4$	$A_M$
	0,4730 (0,0005)	0,4370 (0,0002)	0,2660 (0,0000)	0,3870 (0,0011)	0,3350 (0,0000)
$\{X_{i,t}^2\}_{t=1}^{2.000}$					
Método \ Ativo	$A_1$	$A_2$	$A_3$	$A_4$	$A_M$
R/S $l = 100$	0,2628 (0,0350)	0,3017 (0,0220)	0,1409 (0,0142)	0,2653 (0,0239)	0,2461 (0,0088)
R/S <sub><math>m</math></sub> $l = 100$	0,2452 (0,0419)	0,2859 (0,0269)	0,1323 (0,0163)	0,2482 (0,0295)	0,2283 (0,0125)
KPSS $l = 100$	0,2321 (0,0475)	0,2825 (0,0281)	0,1018 (0,0250)	0,2507 (0,0287)	0,2076 (0,0175)
VS $l = 100$	0,2296 (0,0486)	0,2830 (0,0279)	0,1006 (0,0254)	0,2412 (0,0320)	0,2118 (0,0164)
GPH $\alpha = 0,5$	0,2595 (0,0363)	0,3448 (0,0111)	0,1096 (0,0226)	0,2567 (0,0267)	0,2227 (0,0138)
$\{\ln(X_{i,t}^2)\}_{t=1}^{2.000}$					
Método \ Ativo	$A_1$	$A_2$	$A_3$	$A_4$	$A_M$
R/S $l = 100$	0,2815 (0,0284)	0,3097 (0,0197)	0,1391 (0,0146)	0,2796 (0,0197)	0,2667 (0,0711)
R/S <sub><math>m</math></sub> $l = 100$	0,2644 (0,0344)	0,2933 (0,0246)	0,1307 (0,0167)	0,2624 (0,0248)	0,2493 (0,0622)
KPSS $l = 100$	0,2528 (0,0389)	0,2969 (0,0234)	0,1026 (0,0248)	0,2671 (0,0234)	0,2283 (0,0521)
VS $l = 100$	0,2479 (0,0408)	0,2952 (0,0240)	0,1032 (0,0246)	0,2571 (0,0265)	0,2330 (0,0543)
GPH $\alpha = 0,5$	0,2951 (0,0240)	0,3638 (0,0074)	0,1149 (0,0211)	0,2878 (0,0175)	0,2593 (0,0672)

**Nota:** Os valores entre parênteses representam o erro quadrático médio (MSE).

## 5.6.2 Ajuste e Previsão para as Séries Simuladas

Nesta seção apresentamos os resultados relacionados ao ajuste de processos FIEGARCH e à previsão da volatilidade para as séries temporais apresentadas na Subseção 5.6.1. Estes resultados são necessários para o cálculo das medidas de risco que apresentamos na Subseção 5.6.6. Primeiramente tentamos ajustar modelos com o mesmo grau de  $p$  e  $q$  dos modelos simulados. Quando necessário, reajustamos um novo modelo. Utilizamos, na seleção dos modelos, os critérios apresentados na Seção 3.1. Observamos que, em alguns casos, o modelo mais adequado para a série temporal difere em relação aos valores  $p$  e  $q$  que foram simulados. Consideramos apenas os parâmetros significativos dos modelos, como sugerido por Ruey Tsay em suas notas de aula (*Lecture Note of Bus 41202, Spring 2007: More Volatility Models*).

Tabela 5.21: Modelo FIEGARCH( $p, d, q$ ), com  $p = 0, q = 1$  Ajustado aos Log-retornos dos Preços do Ativo  $A_1$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,201	0,032	-6,362	0,000
$\beta_1$	0,666	0,124	5,391	0,000
$\psi_0$	0,253	0,040	6,323	0,000
$\gamma_0$	-0,082	0,019	-4,297	0,000
$d$	0,473	0,101	4,691	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 6229,694		Estatística	p-valor	
BIC = 6257,673		0,1395	0,9326	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
11,21	0,51	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
18,71	0,10	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
19,20	0,08	1,76	0,14	

Na Tabela 5.21 apresentamos o resultado do ajuste de um modelo FIEGARCH( $p, d, q$ ), com  $p = 0$  e  $q = 1$  à série temporal simulada  $\{X_{1,t}\}_{t=1}^{1,990}$ , correspondente aos log-retornos dos preços do ativo  $A_1$ . Percebemos que, ao nível de 5% de significância, todos os coeficientes ajustados são significativos. O  $d$  estimado é igual a 0,473 com desvio padrão igual a 0,101 (lembramos que o valor deste parâmetro, utilizado na simulação, é  $d = 0,45$ ). O teste de normalidade de Jarque-Bera forneceu  $JB = 0,1395$ , com p-valor igual a 0,9326, o que indica que os resíduos são normalmente distribuídos. O teste de Ljung-Box, para os resíduos padronizados forneceu  $Q(12) = 11,21$ , com p-valor igual a 0,51, o que indica que os resíduos são não-correlacionados, enquanto que o mesmo teste para os quadrados dos resíduos padronizados forneceu  $Q(12) = 18,71$ , com p-valor igual a 0,10, o que indica que os resíduos não apresentam heteroscedasticidade. O resultado é confirmado pelo teste dos Multiplicadores de Lagrange que fornece  $ML = 19,20$ , com p-valor igual a 0,08, o que indica que, ao nível de 5% de significância, os resíduos não apresentam heteroscedasti-

cidade condicional. Sob essas condições, o modelo final para a variância condicional da série temporal  $\{X_{1,t}\}_{t=1}^{1.990}$ , correspondente aos log-retornos do ativo  $A_1$  é dado por

$$(1 - 0,666\mathcal{B})(1 - \mathcal{B})^{0,473} \ln(\sigma_t^2) = -0,201 + 0,253|Z_{t-1}| - 0,082Z_{t-1}.$$

Tabela 5.22: Modelo FIEGARCH( $p, d, q$ ), com  $p = 1, q = 1$  Ajustado aos Log-retornos dos Preços do Ativo  $A_2$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,316	0,041	-7,598	0,000
$\beta_1$	0,348	0,168	2,068	0,019
$\psi_0$	-0,003	0,054	-0,048	0,481
$\psi_1$	0,397	0,068	5,802	0,000
$\gamma_0$	0,020	0,034	0,722	0,235
$\gamma_1$	-0,009	0,034	-0,265	0,396
$d$	0,437	0,091	4,820	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 5330,272		Estatística	p-valor	
BIC = 5369,443		0,0382	0,9811	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
13,33	0,35	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
17,40	0,14	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
17,57	0,13	1,61	0,19	

Na Tabela 5.22 apresentamos o resultado do ajuste de um modelo FIEGARCH( $p, d, q$ ), com  $p = 1 = q$ , à série temporal simulada  $\{X_{2,t}\}_{t=1}^{1.990}$ , correspondente aos log-retornos do ativo  $A_2$ . Percebemos que, ao nível de 5% de significância, os coeficientes  $\psi_0, \gamma_0$  e  $\gamma_1$  são não-significativos, com p-valores iguais a 0,481, 0,235 e 0,395, respectivamente. Enquanto que todos os outros coeficientes ajustados são significativos. Sendo assim, um modelo aproximado para o logaritmo da volatilidade é dado por

$$(1 - 0,348\mathcal{B})(1 - \mathcal{B})^{0,437} \ln(\sigma_t^2) = -0,316 + 0,397|Z_{t-2}|.$$

O  $d$  estimado é igual a 0,437 com desvio padrão igual a 0,091 (para a simulação utilizamos  $d = 0,45$ ). O teste de normalidade de Jarque-Bera forneceu  $JB = 0,0382$ , com p-valor igual a 0,9811, o que indica que os resíduos são normalmente distribuídos. O teste de Ljung-Box, para os resíduos padronizados forneceu  $Q(12) = 13,33$ , com p-valor igual a 0,35, o que indica que os resíduos são não-correlacionados, enquanto que o mesmo teste para os quadrados dos resíduos padronizados forneceu  $Q(12) = 17,40$ , com p-valor igual a 0,14, o que indica que os resíduos não apresentam heteroscedasticidade. O resultado é confirmado pelo teste dos Multiplicadores de Lagrange que fornece  $ML = 17,57$ , com p-valor igual a 0,13, o que indica que, ao nível de 5% de significância, os resíduos não apresentam heteroscedasticidade condicional.

Tabela 5.23: Modelo FIEGARCH( $p, d, q$ ), com  $p = 0, q = 4$  Ajustado aos Log-retornos dos Preços do Ativo  $A_3$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,210	0,028	-7,570	0,000
$\beta_1$	1,022	0,092	11,124	0,000
$\beta_2$	-0,687	0,100	-6,871	0,000
$\beta_3$	0,890	0,107	8,338	0,000
$\beta_4$	-0,511	0,079	-6,448	0,000
$\psi_0$	0,269	0,036	7,379	0,000
$\gamma_0$	-0,019	0,015	-1,279	0,101
$d$	0,266	0,104	2,554	0,005
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 5843,276		Estatística	p-valor	
BIC = 5888,043		2,748	0,253	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
12,05	0,44	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
34,600	0,001	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
33,420	0,001	3,090	0,019	

Na Tabela 5.23 apresentamos o resultado do ajuste de um modelo FIEGARCH( $p, d, q$ ), com  $p = 0$  e  $q = 4$ , à série temporal simulada  $\{X_{3,t}\}_{t=1}^{1.990}$ , correspondente aos log-retornos do ativo  $A_3$ . Percebemos que o teste de Ljung-Box, para os quadrados dos resíduos padronizados forneceu  $Q(12) = 34,610$ , com p-valor igual a 0,001, o que indica que os resíduos ainda apresentam heteroscedasticidade. O resultado é confirmado pelo teste dos Multiplicadores de Lagrange que fornece  $ML = 33,420$ , com p-valor igual a 0,001, o que indica que, ao nível de 5% de significância, os resíduos apresentam heteroscedasticidade condicional. Ajustamos novos modelos, com valores de  $p$  e  $q$  entre 0 e 4. Os modelos para os quais os resíduos não apresentam heteroscedasticidade são apresentados a seguir.

Nas Tabelas 5.24 e 5.25 apresentamos dois modelos alternativos para a série temporal  $\{X_{3,t}\}_{t=1}^{1.990}$ . Ambos os modelos solucionam o problema da heteroscedasticidade condicional no quadrado dos resíduos padronizados. Note porém que o parâmetro  $d$  do modelo apresentado na Tabela 5.24 apresenta um erro quadrático de 0,043 em relação ao parâmetro simulado ( $d = 0,26$ ). Por sua vez, o modelo apresentado na Tabela 5.25 apresenta vários parâmetros a mais que o modelo simulado. Considerando apenas os parâmetros significativos, obtemos os seguintes modelos, para o logaritmo da volatilidade,

$$(1 + 0,369\mathcal{B})(1 - \mathcal{B})^{0,466} \ln(\sigma_t^2) = -0,546 + 0,138|Z_{t-1}| + 0,556|Z_{t-2}| + 0,075Z_{t-1}$$

e

$$\beta(\mathcal{B})(1 - \mathcal{B})^{0,264} \ln(\sigma_t^2) = 0,110|Z_{t-1}| + 0,441|Z_{t-2}| - 0,447|Z_{t-3}| - 0,103Z_{t-2},$$

onde  $\beta(\mathcal{B}) = 1 - 0,667\mathcal{B} - 0,257\mathcal{B}^3 + 0,406\mathcal{B}^4$ .

Tabela 5.24: Modelo FIEGARCH( $p, d, q$ ), com  $p = 1, q = 1$  Ajustado aos Log-retornos dos Preços do Ativo  $A_3$  e Resultados da Análise de Resíduos.

Parâmetro	Coeficiente	Desvio Padrão	Estatística t	p-valor
$a$	-0,546	0,060	-9,113	0,000
$\beta_1$	-0,369	0,110	-3,355	0,000
$\psi_0$	0,138	0,052	2,643	0,004
$\psi_1$	0,556	0,056	9,971	0,000
$\gamma_0$	0,075	0,031	2,391	0,008
$\gamma_1$	-0,042	0,028	-1,475	0,070
$d$	0,466	0,053	8,819	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 5823,263		Estatística	p-valor	
BIC = 5862,43		2,991	0,224	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
14,30	0,28	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
18,75	0,09	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
19,93	0,07	1,83	0,13	

Tabela 5.25: Modelo FIEGARCH( $p, d, q$ ), com  $p = 3, q = 4$  Ajustado aos Log-retornos dos Preços do Ativo  $A_3$  e Resultados da Análise de Resíduos.

Parâmetro	Coeficiente	Desvio Padrão	Estatística t	p-valor
$a$	-0,179	0,117	-1,529	0,063
$\beta_1$	0,667	0,160	4,183	0,000
$\beta_2$	0,212	0,145	1,460	0,072
$\beta_3$	0,257	0,101	2,556	0,005
$\beta_4$	-0,406	0,095	-4,271	0,000
$\psi_0$	0,110	0,056	1,983	0,024
$\psi_1$	0,441	0,079	5,587	0,000
$\psi_2$	-0,447	0,151	-2,961	0,002
$\psi_3$	0,121	0,109	1,113	0,133
$\gamma_0$	0,049	0,032	1,564	0,059
$\gamma_1$	-0,103	0,038	-2,684	0,004
$\gamma_2$	0,074	0,044	1,679	0,047
$\gamma_3$	-0,014	0,032	-0,430	0,334
$d$	0,264	0,141	1,867	0,031
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 5802,667		Estatística	p-valor	
BIC = 5881,010		0,6941	0,7068	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
13,440	0,338	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
1,3590	0,9999	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
1,3540	0,9999	0,1232	1	

Na Tabela 5.26 apresentamos o resultado do ajuste de um modelo FIEGARCH( $p, d, q$ ), com  $p = 0$  e  $q = 1$ , à série temporal simulada  $\{X_{4,t}\}_{t=1}^{1.990}$ , correspondente aos log-retornos do ativo  $A_4$ . Percebemos que o teste de Ljung-Box para os quadrados dos resíduos padronizados forneceu  $Q(12) = 27,800$ , com p-valor igual a 0,006, o que indica que os resíduos apresentam heteroscedasticidade. O resultado é confirmado pelo teste dos Multiplicadores de Lagrange que fornece  $ML = 26,490$ , com p-valor igual a 0,009, o que indica que, ao nível de 5% de significância, os resíduos apresentam heteroscedasticidade condicional. Apresentamos, na Tabela 5.27 um modelo alternativo para essa série temporal. Note que, para este modelo o quadrado dos resíduos padronizados não apresentam heteroscedasticidade condicional. Além disso, ao nível de 5% de significância, todos os parâmetros são significativos para o modelo. Sendo assim, obtemos

$$(1 - 1,364\mathcal{B} + 0,552\mathcal{B}^2)(1 - \mathcal{B})^{0,379} \ln(\sigma_t^2) = -0,103 + 0,133|Z_{t-1}| - 0,033Z_{t-1}.$$

Tabela 5.26: Modelo FIEGARCH( $p, d, q$ ), com  $p = 0, q = 1$  Ajustado aos Log-retornos dos Preços do Ativo  $A_4$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,145	0,026	-5,510	0,000
$\beta_1$	0,767	0,126	6,080	0,000
$\psi_0$	0,187	0,034	5,545	0,000
$\gamma_0$	-0,050	0,016	-3,103	0,000
$d$	0,387	0,156	2,476	0,007
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 6508,603		Estatística	p-valor	
BIC = 6536,582		0,2117	0,8996	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
18,45	0,10	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
27,800	0,006	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
26,490	0,009	2,441	0,048	

Na Tabela 5.28 apresentamos o resultado do ajuste de um modelo FIEGARCH( $p, d, q$ ), com  $p = 0$  e  $q = 1$ , à série temporal simulada  $\{X_{5,t}\}_{t=1}^{1.990}$ , correspondente aos log-retornos do índice do mercado. O teste de Ljung-Box para os quadrados dos resíduos padronizados forneceu  $Q(12) = 30,440$ , com p-valor igual a 0,002, o que indica que os resíduos apresentam heteroscedasticidade. O resultado é confirmado pelo teste dos Multiplicadores de Lagrange que forneceu  $ML = 31,010$ , com p-valor igual a 0,002, o que indica que, ao nível de 5% de significância, os resíduos apresentam heteroscedasticidade condicional. Apresentamos, na Tabela 5.29 um modelo alternativo para essa série temporal. Note que, para este modelo, o quadrado dos resíduos padronizados não apresentam heteroscedasticidade

condicional. Além disso, ao nível de 5% de significância, todos os parâmetros são significativos para este modelo. No entanto, o valor estimado para o parâmetro  $d$  foi 0,264 com desvio padrão de 0,090, enquanto que o valor simulado deste parâmetro foi  $d = 0,34$ . Sob essas condições obtemos o modelo

$$(1 - 0,679\mathcal{B} - 0,734\mathcal{B}^2 + 0,524\mathcal{B}^3)(1 - \mathcal{B})^{0,264} \ln(\sigma_t^2) = -0,135 + 0,173|Z_{t-1}| - 0,084Z_{t-1}.$$

Tabela 5.27: Modelo FIEGARCH( $p, d, q$ ), com  $p = 0, q = 2$  Ajustado aos Log-retornos dos Preços do Ativo  $A_4$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,103	0,020	-5,232	0,000
$\beta_1$	1,364	0,104	13,111	0,000
$\beta_1$	-0,552	0,090	-6,108	0,000
$\psi_0$	0,133	0,026	5,151	0,000
$\gamma_0$	-0,033	0,011	-3,076	0,001
$d$	0,379	0,101	3,752	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 6493,184		Estatística	p-valor	
BIC = 6526,759		0,7238	0,6963	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
17,31	0,14	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
18	0,12	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
17,23	0,14	1,58	0,20	

Tabela 5.28: Modelo FIEGARCH( $p, d, q$ ), com  $p = 0, q = 1$  Ajustado aos Log-retornos dos Preços do Ativo  $A_5$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,153	0,021	-7,185	0,000
$\beta_1$	0,839	0,071	11,849	0,000
$\psi_0$	0,194	0,028	6,993	0,000
$\gamma_0$	-0,091	0,018	-5,088	0,000
$d$	0,335	0,125	2,677	0,004
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 6535,781		Estatística	p-valor	
BIC = 6563,760		0,6423	0,7253	
Teste de Ljung-Box test para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
1,8170	0,9996	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
30,440	0,002	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
31,010	0,002	2,864	0,026	



Tabela 5.29: Modelo FIEGARCH( $p, d, q$ ), com  $p = 0, q = 3$  Ajustado aos Log-retornos dos Preços do Ativo  $A_5$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,135	0,022	-6,041	0,000
$\beta_1$	0,679	0,097	7,035	0,000
$\beta_1$	0,734	0,116	6,303	0,000
$\beta_1$	-0,542	0,080	-6,433	0,000
$\psi_0$	0,173	0,029	5,884	0,000
$\gamma_0$	-0,084	0,016	-5,201	0,000
$d$	0,264	0,089	2,960	0,002
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 6524,74		Estatística	p-valor	
BIC = 6563,9		0,916	0,633	
Teste de Ljung-Box test para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
1,7190	0,9997	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
19,62	0,07	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
20,81	0,05	1,91	0,11	

Após selecionar o modelo mais adequado para cada série temporal simulada, obtivemos a previsão da volatilidade para  $h = 1, \dots, 10$  dias à frente, como descrito no Capítulo 3. Para obter as previsões, apenas os parâmetros significativos dos modelos foram utilizados. No Capítulo 3, mostramos que  $\hat{X}_{i,t+h}^2 = \hat{\sigma}_{i,t+h}^2$ , para todo  $h > 1$ , para cada  $i = 1, \dots, 5$ . Utilizamos estes resultados para verificar a precisão da previsão a partir dos modelos ajustados.

Nas Tabelas 5.30 - 5.34 apresentamos os valores simulados de  $X_{i,t}^2$  e os valores previstos  $\hat{X}_{i,t}^2$ , para cada  $t = 1.991, \dots, 2.000$  e  $i = 1, \dots, 5$ , a partir dos modelos iniciais (modelos simulados), que denotamos por MI, e dos modelos finais ajustados a cada série temporal, que denotamos por MF, correspondente aos log-retornos dos preços dos ativos  $A_i$ ,  $i \in \{1, \dots, 5\}$ , respectivamente.

Na Tabela 5.32, MF<sub>1</sub> refere-se ao modelo final ajustado à série temporal dos log-retornos do ativo  $A_3$ , apresentado na Tabela 5.24. Enquanto que, MF<sub>2</sub> refere-se ao modelo apresentado na Tabela 5.25.

Observamos que, em todos os casos, o erro de previsão, definido pela expressão (3.4), para os modelos iniciais (MI) e finais (MF) assume valores muito próximos.

Tabela 5.30: Previsões para o Valor Quadrático dos Log-retornos do Ativo  $A_1$ .

$t$	$X_{1,t}^2$	$\hat{X}_{1,t}^2$ (MI)	$X_{1,t}^2 - \hat{X}_{1,t}^2$	$\hat{X}_{1,t}^2$ (MF)	$X_{1,t}^2 - \hat{X}_{1,t}^2$
1.991	0,2878	2,2128	-1,9251	2,1841	-1,8964
1.992	0,0698	2,1269	-2,0570	2,1133	-2,0435
1.993	0,9926	2,0933	-1,1006	2,0710	-1,0783
1.994	0,0906	2,0700	-1,9794	2,0393	-1,9487
1.995	0,0832	2,0491	-1,9659	2,0129	-1,9297
1.996	0,4110	2,0292	-1,6181	1,9896	-1,5786
1.997	0,1045	2,0100	-1,9055	1,9685	-1,8640
1.998	4,3324	1,9918	2,3406	1,9490	2,3834
1.999	0,9721	1,9747	-1,0025	1,9308	-0,9587
2.000	0,4923	1,9585	-1,4662	1,9137	-1,4214

**Nota:** MI Refere-se à Previsão a Partir do Modelo Utilizado na Simulação e MF Refere-se ao Modelo Final Ajustado à Série Temporal.

Tabela 5.31: Previsões para o Valor Quadrático dos Log-retornos do Ativo  $A_2$ .

$t$	$X_{1,t}^2$	$\hat{X}_{1,t}^2$ (MI)	$X_{2,t}^2 - \hat{X}_{2,t}^2$	$\hat{X}_{2,t}^2$ (MF)	$X_{2,t}^2 - \hat{X}_{2,t}^2$
1.991	2,4301	1,1379	1,2921	1,1310	1,2990
1.992	0,1291	0,9237	-0,7946	1,1168	-0,9877
1.993	0,0061	0,8519	-0,8458	1,1123	-1,1062
1.994	1,4623	0,8150	0,6473	1,1086	0,3537
1.995	0,0858	0,7951	-0,7093	1,1049	-1,0191
1.996	0,6675	0,7855	-0,1180	1,1012	-0,4336
1.997	4,5980	0,7825	3,8155	1,0975	3,5005
1.998	2,3776	0,7841	1,5935	1,0940	1,2836
1.999	14,1703	0,7891	13,3812	1,0907	13,0797
2.000	2,7279	0,7965	1,9314	1,0875	1,6404

**Nota:** MI Refere-se à Previsão a Partir do Modelo Utilizado na Simulação e MF Refere-se ao Modelo Final Ajustado à Série Temporal.

Tabela 5.32: Previsões para o Valor Quadrático dos Log-retornos do Ativo  $A_3$ .

$t$	$X_{3,t}^2$	$\hat{X}_{3,t}^2$ (MI)	$X_{3,t}^2 - \hat{X}_{3,t}^2$	$\hat{X}_{3,t}^2$ (MF <sub>1</sub> )	$X_{3,t}^2 - \hat{X}_{3,t}^2$	$\hat{X}_{3,t}^2$ (MF <sub>2</sub> )	$X_{3,t}^2 - \hat{X}_{3,t}^2$
1.991	1,6588	1,5482	0,1107	1,1160	0,5429	3,1177	-1,4589
1.992	1,7114	1,1410	0,5704	0,8957	0,8157	1,9901	-0,2787
1.993	0,6831	1,2915	-0,6084	1,0871	-0,4040	1,3839	-0,7008
1.994	0,0142	1,2132	-1,1989	1,0392	-1,0249	1,5341	-1,5198
1.995	2,0267	1,0414	0,9853	1,0677	0,9590	1,5735	0,4532
1.996	0,0123	1,2083	-1,1959	1,0639	-1,0516	1,7848	-1,7724
1.997	0,2104	1,1134	-0,9030	1,0696	-0,8592	2,3345	-2,1241
1.998	0,5101	1,0736	-0,5635	1,0707	-0,5606	2,7241	-2,2140
1.999	1,4233	1,2039	0,2194	1,0727	0,3506	3,1078	-1,6845
2.000	0,0725	1,1026	-1,0301	1,0739	-1,0014	3,4713	-3,3988

**Nota:** MI Refere-se à Previsão a Partir do Modelo Utilizado na Simulação e MF Refere-se ao Modelo Final Ajustado à Série Temporal.

Tabela 5.33: Previsões para o Valor Quadrático dos Log-retornos do Ativo  $A_4$ .

$t$	$X_{1,t}^2$	$\hat{X}_{1,t}^2$ (MI)	$X_{4,t}^2 - \hat{X}_{4,t}^2$	$\hat{X}_{4,t}^2$ (MF)	$X_{4,t}^2 - \hat{X}_{4,t}^2$
1991	1,5009	1,6488	-0,1480	1,3755	0,1254
1992	$7,91 \times 10^{-6}$	1,5849	-1,5849	1,3791	-1,3791
1993	3,3353	1,5734	1,7619	1,4742	1,8611
1994	7,4029	1,5745	5,8284	1,5048	5,8981
1995	0,9531	1,5780	-0,6249	1,4580	-0,5049
1996	0,5829	1,5811	-0,9981	1,4220	-0,8390
1997	0,0364	1,5830	-1,5466	1,4308	-1,3945
1998	2,7934	1,5839	1,2096	1,4510	1,3424
1999	2,0500	1,5838	0,4662	1,4506	0,5994
2000	2,2839	1,5832	0,7007	1,4368	0,8471

**Nota:** MI Refere-se à Previsão a Partir do Modelo Utilizado na Simulação e MF Refere-se ao Modelo Final Ajustado à Série Temporal.

Tabela 5.34: Previsões para o Valor Quadrático dos Log-retornos dos Índices do Mercado  $A_M$ .

$t$	$X_{1,t}^2$	$\hat{X}_{1,t}^2$ (MI)	$X_{5,t}^2 - \hat{X}_{5,t}^2$	$\hat{X}_{5,t}^2$ (MF)	$X_{5,t}^2 - \hat{X}_{5,t}^2$
1.991	1,3625	1,6616	-0,2991	2,2107	-0,8482
1.992	0,3527	1,6138	-1,2610	1,7772	-1,4244
1.993	4,7542	1,6297	3,1245	1,6220	3,1323
1.994	0,4706	1,6627	-1,1922	1,4332	-0,9626
1.995	1,6892	1,6980	-0,0088	1,3903	0,2989
1.996	9,7724	1,7297	8,0427	1,3087	8,4637
1.997	0,0032	1,7561	-1,7529	1,3139	-1,3108
1.998	11,7870	1,7768	10,0102	1,2816	10,5054
1.999	3,2052	1,7923	1,4129	1,3054	1,8998
2.000	0,0717	1,8033	-1,7315	1,2948	-1,2231

**Nota:** MI Refere-se à Previsão a Partir do Modelo Utilizado na Simulação e MF Refere-se ao Modelo Final Ajustado à Série Temporal.

Os critérios de seleção relacionadas ao erro de previsão, definidos na Subseção 3.1.4, para as séries temporais correspondentes aos log-retornos dos preços dos ativos  $A_i$ ,  $i \in \{1, \dots, 5\}$ , são apresentadas na Tabela 5.35. Observamos que os erros de previsão fornecidos pelos modelos finais, isto é, pelos modelos ajustados às séries temporais, são muito próximos dos erros de previsão fornecidos pelos modelos iniciais, isto é, aqueles que foram simulados. Note que, para o ativo  $A_3$ , obtivemos duas sugestões de modelos. De acordo com os valores apresentados na Tabela 5.35, o modelo que melhor se ajusta à série temporal simulada é o modelo  $MF_1$ , apresentado na Tabela 5.24. Note que, para o ativo  $A_4$ ,  $X_{4,1.992}^2 = 7,91 \times 10^{-6}$ . Devido a esse fato, obtivemos altos valores para os critérios de seleção MPE e MAPE para esse ativo.

Tabela 5.35: Critérios de Seleção para o Erro de Previsão.

Ativo	Modelo	MPE	MSE	MAE	MAPE
$A_1$	MI	-10,8376	3,1814	1,7361	10,9456
	MF	-10,6646	3,1043	1,7103	10,7746
$A_2$	MI	-14,8631	20,3838	2,5129	15,6884
	MF	-19,7271	19,2908	2,4703	20,4475
$A_3$	MI	-20,0516	0,6789	0,7386	20,2597
	MF <sub>1</sub>	-17,5213	0,6378	0,7570	17,8260
	MF <sub>2</sub>	-31,4646	3,2157	1,5605	31,5094
$A_4$	MI	-20041,5544	4,5559	1,4869	20042,0110
	MF	-17438,5109	4,5951	1,4791	17439,0274
$A_M$	MI	-58,0739	18,5819	2,8836	58,6279
	MF	-43,4111	20,2396	3,0069	44,0483

**Nota:** MI Refere-se à Previsão a Partir do Modelo Utilizado na Simulação e MF Refere-se ao Modelo Final Ajustado à Série Temporal.

### 5.6.3 Cálculo dos Pesos do *Portfolio*

Nesta seção apresentamos o cálculo do *portfolio* eficiente utilizando o método da análise de média-variância descrito na Seção 4.3. O *portfolio* é representado por  $\mathcal{P} = \{A_1, A_2, A_3, A_4\}$  e o vetor de pesos é representado por  $\mathbf{a} = (a_1, a_2, a_3, a_4)'$ . Assumimos que o log-retorno médio do ativo livre de risco no período considerado é igual a 0,02. Sob essas condições, utilizando a expressão (4.22), obtemos

$$(a_1, a_2, a_3, a_4) = (0, 16; 0, 36; 0, 39; 0, 09), \quad (5.3)$$

isto é, 16% do valor total a ser investido será aplicado no ativo  $A_1$ , 36% no ativo  $A_2$ , 39% no ativo  $A_3$  e 9% no ativo  $A_4$ .

**Observação 5.2.** Note que, no *portfolio* aqui considerado não estamos incluindo o ativo livre de risco. Por essa razão, não utilizamos a expressão (4.20) para o cálculo dos pesos do *portfolio*. Observe que, se fixarmos um retorno esperado  $\mu_0 = 1\%$  e utilizarmos a expressão (4.20) para o cálculo dos pesos do *portfolio*, encontramos

$$(a_1, a_2, a_3, a_4) = (0, 03; 0, 08; 0, 09; 0, 02).$$

Note que,  $a_1 + a_2 + a_3 + a_4 = 0, 22$ , o que significa que 22% do valor total a ser investido deve ser destinado aos ativos de risco (ações) e 78% deve ser reservado para investir no ativo livre de risco (poderia ser, por exemplo, a caderneta de poupança). Para diferentes valores de  $\mu_0$ , obtemos diferentes pesos. Observamos ainda que, para este caso, quanto menor o valor de  $\mu_0$ , maior a proporção dos ativos de risco em relação ao total do *portfolio*.

### 5.6.4 Estatísticas Descritivas dos Ativos do *Portfolio*

Na Tabela 5.36 apresentamos as estatísticas descritivas das séries temporais dos log-retornos dos ativos do *portfolio*. Observamos que o ativo que apresentou a melhor desempenho, quantificado pelo valor da mediana dos retornos diários igual a 0,012, foi o ativo

$A_2$  (veja Tabela 5.36). Observa-se também que a desempenho do mercado foi inferior ao desempenho dos quatro ativos considerados. O ativo que apresenta maior variabilidade é o  $A_4$ , com variância igual a 1,663. Essa variância é menor do que a variância do mercado, que é igual a 1,970. O ativo  $A_2$  é o menos volátil, com variância igual a 0,921. Entre todos os ativos, o que apresentou o menor retorno no período foi o ativo  $A_1$  (-6,889), e este foi menor do que o menor retorno apresentado pelo mercado no mesmo período (-6,731). O ativo que apresentou o maior retorno no período foi o ativo  $A_4$  (6,163) e foi inferior ao maior retorno do mercado (6,560) no período. Percebemos também que, para todas as séries temporais de log-retornos, a distribuição dos dados é aproximadamente simétrica, já que, em todos os casos, a medida de assimetria assume valores próximos de zero. Entretanto, todas as séries temporais podem ser consideradas de caudas pesadas, sendo a série temporal dos log-retornos do ativo  $A_5$  a que apresenta a maior medida de curtose (5,081).

Tabela 5.36: Estatísticas Descritivas para as Séries Simuladas.

Ativo	Média	Variância	Mediana	Máximo	Mínimo	Assimetria	Curtose
$A_1$	-0,012	1,651	-0,011	5,826	-6,889	-0,148	5,064
$A_2$	-0,022	0,921	0,012	3,947	-4,497	-0,027	3,684
$A_3$	-0,040	1,215	-0,029	5,471	-4,582	0,016	4,094
$A_4$	0,002	1,663	-0,018	6,163	-4,536	0,059	3,640
$A_M$	-0,042	1,970	-0,027	6,560	-6,731	-0,150	5,081

### 5.6.5 Modelo CAPM para as Séries Temporais do *Portfolio*

Nesta subseção apresentamos o cálculo do coeficiente  $\beta$  para os ativos do *portfolio*  $\mathcal{P}$ , cujos log-retornos são representados pelas séries temporais simuladas. Lembramos que, por hipótese, a série temporal simulada pelo modelo M5 representa o mercado financeiro.

Na prática, a correlação entre os ativos e a correlação dos ativos com o mercado financeiro (e portanto, o valor do coeficiente  $\beta$ ) pode sofrer grandes alterações no decorrer dos anos. Sendo assim, consideramos apenas as 1.000 últimas observações (aproximadamente, 4 anos) para o cálculo do coeficiente  $\beta$ . Na Tabela 5.37 apresentamos a correlação amostral entre as séries temporais simuladas. Note que, de acordo com os resultados obtidos, os log-retornos dos ativos são pouco correlacionados, tanto entre si como com o mercado. Isso difere do que ocorre na prática, como vemos nas séries temporais reais apresentadas no Capítulo 6.

Tabela 5.37: Correlação Amostral dos Log-retornos Simulados.

	$A_1$	$A_2$	$A_3$	$A_4$	$A_M$
$A_1$	1,0000	0,0017	-0,0015	-0,0388	0,0053
$A_2$	0,0017	1,0000	-0,0093	0,0361	0,0584
$A_3$	-0,0015	-0,0093	1,0000	-0,0197	0,0880
$A_4$	-0,0388	0,0361	-0,0197	1,0000	0,0239
$A_M$	0,0053	0,0584	0,0880	0,0239	1,0000

Assumimos que o log-retorno do ativo livre de risco no período considerado foi igual a 0,02. Para cada  $i = 1, \dots, 4$ , o parâmetro  $\beta_i$  foi estimado através da expressão (4.28). Segue que

$$\beta_1 = 0,006, \quad \beta_2 = 0,043, \quad \beta_3 = 0,070 \quad \text{e} \quad \beta_4 = 0,023.$$

Note que, para todos os ativos,  $\beta_i < 1$ , o que significa que os ativos são defensivos. Isto significa que se houverem quedas no índice do mercado, os retornos dos ativos sofrerão quedas menores (em termos de porcentagem) do que a queda do índice do mercado. De forma análoga, se o índice do mercado sofrer um determinado aumento, o aumento sofrido pelos retornos também será menor. Pelas expressões (4.31) e (5.3) segue que o coeficiente  $\beta$  do *portfolio* é dado por

$$\beta_{\mathcal{P}} = 0,16 \times 0,006 + 0,36 \times 0,043 + 0,39 \times 0,070 + 0,09 \times 0,023 = 0,046.$$

Obtivemos  $\beta_{\mathcal{P}} = 0,046$ , o que indica que, se o mercado sofre uma queda de 10%, o valor do *portfolio* sofrerá uma queda de apenas 0,46%. Note que, sob o modelo CAPM, o retorno esperado do ativo  $A_i$  é dado por

$$\mathbb{E}(R_i) = R_F + \beta_i(\mathbb{E}(R_M) - R_F), \quad \text{para cada } i = 1, \dots, 4.$$

O valor  $\mathbb{E}(R_M)$  é estimado através da média amostral dos retornos do mercado. Os valores da média amostral para as séries temporais simuladas foram obtidos da Tabela 5.36. Segue que

$$\mathbb{E}(R_i) = 0,02 + \beta_i(-0,042 - 0,02), \quad \text{para cada } i = 1, \dots, 4.$$

Assim, o retorno esperado para o período de investimento, segundo o modelo CAPM, é dado por

$$\mathbb{E}(R_1) = 0,0196, \quad \mathbb{E}(R_2) = 0,0173, \quad \mathbb{E}(R_3) = 0,0157 \quad \text{e} \quad \mathbb{E}(R_4) = 0,0186.$$

Lembramos que, para um período  $h = 10$  dias o retorno dos investimentos é dado por

$$\ln(P_{i,1.990+h}) - \ln(P_{i,1.990}) = X_{i,1.990}[10] = X_{i,1.991} + \dots + X_{i,2.000} \quad \text{para cada } i = 1, \dots, 4.$$

A Tabela 5.38 apresenta os valores (observados) dos retornos de período 1 e 10 dias, que são dados, respectivamente, por  $X_{i,1.991}$  e  $X_{i,1.990}[10]$ , para  $i = 1, \dots, 4$ , a média amostral dos retornos observados e os retornos estimados pelo modelo CAPM.

Tabela 5.38: Retornos dos Ativos.

Ativo	$R_i$	$R_i[10]$	$\mathbb{E}(R_i)$	$\bar{R}_i$
$A_1$	-0,5364	-0,9071	0,0196	-0,013
$A_2$	1,5589	9,3960	0,0173	-0,017
$A_3$	1,2880	0,7426	0,0157	-0,040
$A_4$	-1,2251	1,3954	0,0186	0,003

**Nota:**  $R_i \equiv X_{i,1.991}$  e  $R_i[10] \equiv X_{i,1.990}[10]$ .

Comparando os valores obtidos pelo modelo CAPM com os valores observados, para o período  $h = 1$  dia, notamos que, para os ativos  $A_1$  e  $A_4$  os retornos observados foram

menores que os retornos estimados, enquanto que, para os ativos  $A_2$  e  $A_3$ , os resultados obtidos através do modelo CAPM foram inferiores ao observado. Enquanto que, para um horizonte  $h = 10$  dias, o retorno previsto pelo modelo CAPM só foi superior ao retorno observado no caso do ativo  $A_1$ . Tais resultados indicam que, para o *portfolio* analisado, o modelo CAPM não apresenta um padrão em relação às estimativas para o horizonte  $h = 1$ . Entretanto, para um período maior, o modelo tende a estimar retornos menores do que os observados.

O coeficiente  $\alpha$  da ação, que indica se ela está subavaliada ou superavaliada, é definido pela expressão (4.29). Para este *portfolio* temos:

$$\alpha_1 = -0,0326; \quad \alpha_2 = -0,0340; \quad \alpha_3 = -0,0551 \quad \text{e} \quad \alpha_4 = -0,0154.$$

Na visão do investidor, as ações estão superavaliadas pelo mercado (pois  $\alpha$  é negativo) o que deverá proporcionar uma rentabilidade inferior àquela calculada pelo modelo de equilíbrio (CAPM). Portanto, as ações devem ser vendidas.

### 5.6.6 Cálculo das Medidas de Risco VaR e ES

Nesta subseção calculamos as medidas de risco baseadas na distribuição das perdas do *portfolio*. Como sugerido por Palaro e Hotta (2006), utilizamos a série temporal de log-retornos do *portfolio*, definida pela expressão (4.14). No caso univariado, isto é, utilizando os log-retornos do *portfolio*, o cálculo do VaR e do ES foi realizado considerando-se as diferentes abordagens discutidas na Seção 4.5.2. Apresentamos também o cálculo do VaR e do ES, considerando tanto a distribuição condicional, como a distribuição não-condicional do vetor de mudanças de fatores de risco. Para isso utilizamos, respectivamente, a abordagem RiskMetrics e o método da variância-covariância. Por fim, apresentamos a análise das medidas de risco para cada ativo do *portfolio*. Lembramos que a medida ES é uma medida de risco coerente, no sentido da Definição 4.2. Além disso, por definição,  $\text{VaR} \leq \text{ES}$ . Sendo assim, obtendo o ES de cada ativo obtivemos uma cota superior para o VaR do *portfolio*.

Para esta subseção fixamos que:

- os fatores de risco para este *portfolio* são os logaritmos dos preços das ações. Sendo assim, o vetor de fatores de risco é dado por  $\mathbf{Z}_t = (Z_{1,t}, Z_{2,t}, Z_{3,t}, Z_{4,t})'$ , onde  $Z_{i,t}$  é o logaritmo do preço do ativo  $A_i$  no instante  $t$ , para cada  $i = 1, \dots, 4$ . Como consequência, o vetor de mudanças de fatores de risco  $\mathbf{X}_t$ , definido pela expressão (4.34), será dado por

$$\mathbf{X}_t = (X_{1,t}, X_{2,t}, X_{3,t}, X_{4,t})',$$

onde  $X_{i,t}$ , para  $i = 1, \dots, 4$ , representa o log-retorno do ativo  $A_i$  no instante  $t$ ;

- a perda do *portfolio* no instante  $t$  é representada pela variável aleatória  $L_t$ , definida pela expressão (4.32). Dado que, na prática, os retornos simples e os log-retornos são valores muito próximos entre si e pela expressão (4.39), apresentada no Exemplo 4.4, assumimos que a perda do *portfolio* no tempo  $t$  é dada por

$$L_t = -V_{t-1} \sum_{i=1}^4 a_i X_{i,t}, \quad (5.4)$$

onde  $V_{t-1}$  é o valor do *portfolio* no instante  $t - 1$  e  $\mathbf{a} = (a_1, a_2, a_3, a_4)'$  é o vetor de pesos dado pela expressão (5.3);

- para os cálculos realizados, utilizamos os primeiros 1.990 valores das séries temporais para o ajuste dos modelos e reservamos os 10 últimos valores para comparar com os resultados obtidos;
- a função de distribuição considerada para o cálculo das medidas de risco, para todas as abordagens, foi a normal multivariada (ou univariada).

Apresentamos a seguir os resultados da estimação do VaR e do ES para a série temporal das perdas do *portfolio*.

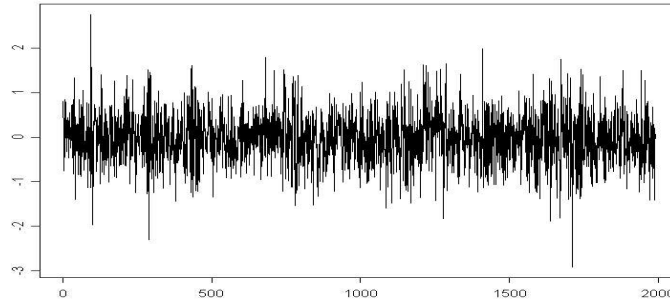


Figura 5.6: Log-retornos do *Portfolio* das Séries Temporais Simuladas.

A série temporal  $\{r_{\mathcal{P},t}\}_{t=1}^{1.990}$ , onde  $r_{\mathcal{P},t}$ , para  $t = 1, \dots, 1.990$ , é dado pela expressão (4.14) com pesos dados pela expressão (5.3), é apresentada na Figura 5.6 e representa os log-retornos do *portfolio*. Da expressão (5.4), temos que a perda do *portfolio*, no instante  $t$ , é dada por

$$L_t = -V_{t-1}r_{\mathcal{P},t}. \quad (5.5)$$

Para os cálculos realizados fixamos  $V_{t-1} = 1$ . Na Tabela 5.39 apresentamos os valores estimados do VaR e do ES, para  $h = 1$  e  $p = 95\%$  e  $99\%$ . O VaR Empírico é o VaR estimado através da distribuição empírica dos log-retornos, o VaR Normal foi estimado através do método da variância-covariância. Esses dois métodos consideram a distribuição não condicional dos dados. O VaR EWMA (estimado através da abordagem RiskMetrics), VaR EGARCH e VaR FIEGARCH (estimados através da abordagem Econométrica) consideram a distribuição condicional dos dados.

O retorno real do *portfolio* para  $h = 1$  é dado por  $r_{\mathcal{P},1.991} = 0,8674$ . Isto significa que a perda do *portfolio*, em  $t = 1.991$ , é dada por  $L_{1.991} = -0,8674$ . Uma perda negativa significa que o investidor ganhou dinheiro neste horizonte de tempo.

Lembramos que o  $\text{VaR}_p$  é a medida tal que  $\mathbb{P}(L_{t+1} \geq \text{VaR}_p) \leq 1-p$ , o que é equivalente a dizer que  $\mathbb{P}(-L_{t+1} \geq -\text{VaR}_p) \geq p$ . Essa última expressão nos diz que a probabilidade do retorno do *portfolio* no tempo  $t + 1$  ser maior do que  $-\text{VaR}$  é maior ou igual a  $p\%$ .

A Figura 5.7 mostra a função densidade de probabilidade das perdas do *portfolio* e os valores do VaR e do ES empíricos para  $p = 99\%$ . Na Tabela 5.39 observamos que, ao nível de confiança de  $99\%$ , o VaR Empírico, apresentado em negrito, foi a medida mais próxima do retorno do *portfolio* para  $t = 1.991$  (que foi igual a  $0,8674$ ), pois obtivemos que  $\mathbb{P}(L_{1.991} \leq 1,3783) \geq 99\%$ , isto significa  $\mathbb{P}(r_{\mathcal{P},1.991} \geq -1,3783) \geq 99\%$ . O mesmo ocorre ao nível de  $95\%$  de confiança. Observamos entretanto que todas as medidas forneceram estimativas semelhantes, sendo que o VaR EWMA foi o que, em ambos



os casos afastou-se mais do valor observado. Para detalhes sobre o ajuste dos modelos EGARCH e FIEGARCH à série temporal de retornos do *portfolio* veja o Apêndice E.

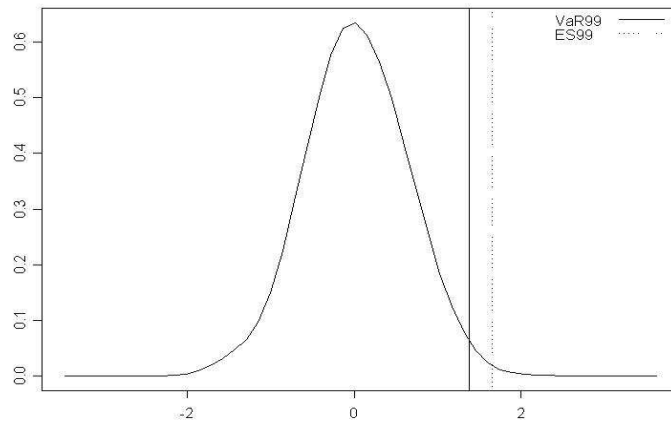


Figura 5.7: Função Densidade das Perdas do *Portfolio* e os Valores do VaR e do ES com  $p = 99\%$ .

Tabela 5.39: Valores Estimados do VaR e do ES para a Série Temporal de Perdas do *Portfolio*. Caso Univariado.

VaR \ P	99%	95%	ES \ P	99%	95%
Empírico	<b>1,3783</b>	<b>0,9958</b>	Empírico	1,6495	1,2395
Normal	1,4040	1,0002	Normal	2,2590	1,7541
EWMA	1,4737	1,0500	EWMA	2,3705	1,8409
EGARCH	1,4442	1,0212	EGARCH	2,3400	1,8110
FIEGARCH	1,4306	1,0115	FIEGARCH	2,3179	1,7939

Os resultados que seguem são referentes à estimativa das medidas de risco considerando-se a distribuição multivariada do vetor de mudanças de fatores de risco.

Na Tabela 5.40 apresentamos os resultados para o cálculo do VaR considerando-se a distribuição multivariada, não-condicional (VaR Normal) e a distribuição condicional (VaR EWMA), do vetor aleatório  $\mathbf{X}_t$ . Observamos que os resultados obtidos no caso do VaR Normal foram os mesmos obtidos quando consideramos a função de distribuição dos retornos do *portfolio*. Para o caso do VaR EWMA os valores encontrados (e apresentados em negrito na Tabela 5.40) foram inferiores aos obtidos pelo método anterior, porém, mais próximos da perda real.

Tabela 5.40: Valores Estimados do VaR e do ES para a Série Temporal de Perdas do *Portfolio*. Caso multivariado.

VaR \ P	99%	95%	ES \ P	99%	95%
Normal	1,4040	1,0002	Normal	2,2590	1,7541
EWMA	<b>1,3624</b>	<b>0,9714</b>	EWMA	2,1903	1,7014

Apresentamos, a seguir, o cálculo das medidas de risco para cada ativo do *portfolio*.

Nas Tabelas 5.41 e 5.42 apresentamos o valor do VaR e do ES para cada um dos ativos simulados. Observamos que, para este *portfolio*, as desigualdades

$$\text{VaR}_{p,p,t+1} \leq \sum_{i=1}^4 a_i \text{VaR}_{p,i,t+1} \quad \text{e} \quad \text{ES}_{p,p,t+1} \leq \sum_{i=1}^4 a_i \text{ES}_{p,i,t+1}$$

estão satisfeitas, tanto ao nível de 95% quanto ao nível de 99% de confiança. Essas desigualdades são válidas tanto considerando a função de distribuição (condicional ou não-condicional) das perdas do *portfolio* quanto a função de distribuição (condicional ou não-condicional) do vetor de mudança de fatores de risco.

As perdas reais dos ativos  $A_i$ , para cada  $i = 1, \dots, 4$ , no instante  $t = 1.991$ , são dadas por 0,5364, -1,5589, -1,2879, e 1,2251, respectivamente. Os valores que mais se aproximaram das perdas reais aparecem em negrito na Tabela 5.41.

Tabela 5.41: Cálculo do VaR para os Ativos  $A_i$ ,  $i \in \{1, \dots, 4\}$ , do *Portfolio*.

p = 99%					
VaR	$A_1$	$A_2$	$A_3$	$A_4$	$\sum_{i=1}^4 a_i \text{VaR}_i$
Empírico	3,3692	2,4256	2,8469	3,1563	2,8066
Normal	3,0018	<b>2,2554</b>	2,6049	2,9979	2,5780
EWMA	<b>2,7816</b>	2,4500	2,4690	<b>2,7372</b>	2,5363
EGARCH	3,3345	2,4579	<b>2,3468</b>	3,1487	2,6170
FIEGARCH	3,4380	2,4740	2,4576	3,4589	2,7105
p = 95%					
VaR	$A_1$	$A_2$	$A_3$	$A_4$	$\sum_{i=1}^4 a_i \text{VaR}_i$
Empírico	1,9773	<b>1,5770</b>	1,8013	2,1185	1,7773
Normal	2,1261	1,6012	1,8537	2,1191	1,8302
EWMA	<b>1,9746</b>	1,7353	1,7597	<b>1,9387</b>	1,8014
EGARCH	2,3576	1,7379	<b>1,6593</b>	2,2263	1,8503
FIEGARCH	2,4309	1,7493	1,7376	2,4456	1,9165

Observamos que, para o *portfolio* considerado, o VaR obtido a partir do ajuste de um processo FIEGARCH tende a superestimar as perdas reais mais do que o VaR obtido através das outras abordagens.

Tabela 5.42: Cálculo do ES para os Ativos  $A_i$ ,  $i \in \{1, \dots, 4\}$ , do *Portfolio*.

p = 99%					
ES	$A_1$	$A_2$	$A_3$	$A_4$	$\sum_{i=1}^4 a_i \text{ES}_i$
Empírico	4,5836	2,8716	3,3786	3,6665	3,4148
Normal	4,8559	3,6404	4,1955	4,8587	4,1610
EWMA	4,4901	3,9632	3,9706	4,4279	4,0922
EGARCH	5,4025	3,9823	3,8023	5,1015	4,2401
FIEGARCH	5,5704	4,0085	3,9818	5,6042	4,3916
p = 95%					
ES	$A_1$	$A_2$	$A_3$	$A_4$	$\sum_{i=1}^4 a_i \text{ES}_i$
Empírico	2,8756	2,0619	2,4568	2,7300	2,4062
Normal	3,7610	2,8225	3,2562	3,7598	3,2261
EWMA	3,4811	3,0696	3,0838	3,4295	3,1734
EGARCH	4,1812	3,0821	2,9427	3,9483	3,2815
FIEGARCH	4,3111	3,1023	3,0817	4,3373	3,3988

### 5.6.7 Análise de Cenários para as Séries Temporais Simuladas

Para realizar o teste de estresse utilizamos duas abordagens diferentes. A primeira delas consiste na escolha dos cenários baseados nos dados históricos, este é o denominado *teste de estresse tradicional*. A segunda abordagem consiste na escolha dos cenários sob um certo domínio de admissibilidade. Obtemos então a medida denominada *perda máxima*, ou *MaxLoss*.

Primeiramente estimamos os valores  $\delta_i$ , para cada  $i = 1, \dots, 4$ , definidos pela expressão (4.67). Para isso, assumimos que:

- a variação do fator de risco é dada por  $\Delta_i = 0,05$ , para cada  $i = 1, \dots, 4$ ;
- a situação atual do mercado é  $\mathbf{Z}_{1.990} = (Z_{1,1.990}, Z_{2,1.990}, Z_{3,1.990}, Z_{4,1.990})'$ , onde  $Z_{i,1.990}$  é o logaritmo do preço do ativo  $A_i$ , para cada  $i = 1, \dots, 4$ , no instante  $t = 1.990$ . O correspondente vetor de mudanças de fatores de risco, utilizado no cálculo do retorno do *portfolio*, é

$$\mathbf{X}_{1.990} = (X_{1,1.990}, X_{2,1.990}, X_{3,1.990}, X_{4,1.990})' = (0,7408; 0,4206; -0,7995; -0,0833)'$$

Lembramos que os valores  $\delta_i$ , para  $i = 1, \dots, 4$ , indicam o quanto o valor do *portfolio* é afetado pela variação de cada fator de risco. Obtivemos  $\delta_i = -a_i$ , para cada  $i = 1, \dots, 4$ , onde  $a_i$  é o peso do ativo  $A_i$ . Isso indica que o fator de risco que mais influencia no preço do *portfolio* é o logaritmo do preço do ativo que possui a maior participação porcentual, em relação ao total do *portfolio*.

A Figura 5.8 apresenta o valor dos retornos do *portfolio* mantendo os dois últimos valores dos fatores de risco fixos e variando os dois primeiros. É fácil ver que, o valor do *portfolio* é uma função linear dos fatores de mudança de risco.

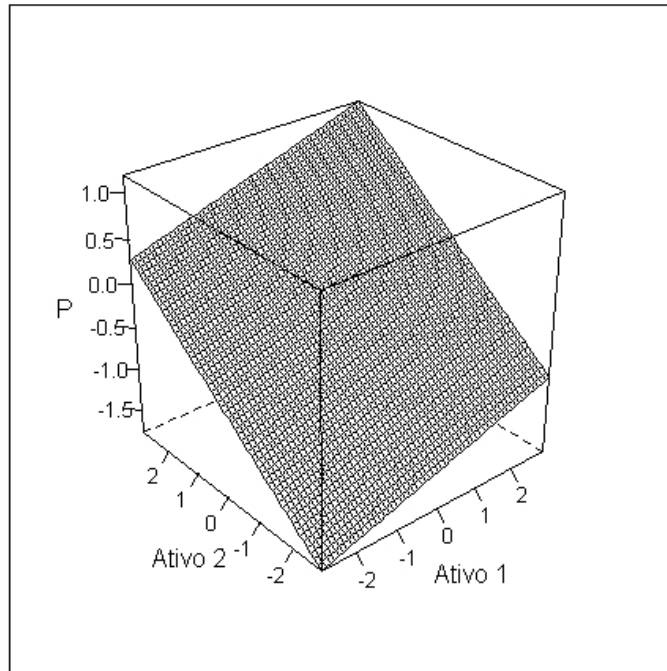


Figura 5.8: Log-retornos do *Portfolio* para o Valor dos Log-retornos dos Ativos  $A_3$  e  $A_4$  Fixos.

### 5.6.8 Construção dos Cenários Utilizando Dados Históricos

Para a construção de cenários baseados em dados históricos, o método mais utilizado na prática, e apresentado nesta seção, é determinar a variação máxima de cada fator de risco e, em seguida, combinar tudo em um cenário. Os procedimentos utilizados nesta subseção para determinar a variação máxima dos fatores de risco (*Start to End* e *drawdown*) são descritos na Subseção 4.6.3. Estes procedimentos são denotados por StE e DD, respectivamente, de acordo com a notação definida na Subseção 4.6.3.

Lembramos que, como os fatores de risco são o logaritmo dos preços, a variação dos fatores de risco, para um período  $h$ , são os log-retornos de período  $h$  (veja Definição 4.5). Apresentamos, na Tabela 5.43, as variações máximas e mínimas dos fatores de risco para cada um dos ativos e para o mercado financeiro, durante o período histórico, considerando janelas de 1 e de 10 dias.

Os cenários considerados são apresentados na Tabela 5.44. Por definição,  $X_{i,Sj}$  representa o valor do log-retorno do ativo  $A_i$  sob o cenário  $Sj$ , para cada  $i = 1, \dots, 4$  e  $j = 1, \dots, 12$ . Os cenários são definidos por

- Cenários S1, S2 e S3: são os cenários otimistas, formados pelos log-retornos máximos dos ativos, obtidos pelo método StE e DD, com janelas de 1 e 10 dias.
- Cenários S4, S5 e S6: são os cenários pessimistas, formados pelos retornos mínimos dos ativos, obtidos pelo método StE e DD, com janelas de 1 e 10 dias.
- Cenários S7, S8 e S9: são os cenários otimistas, construídos considerando-se os log-retornos máximos (obtido das observações históricas) do mercado e os coeficientes

$\beta_i$ , para cada  $i = 1, \dots, 4$ , fornecidos pelo modelo CAPM.

- Cenários S10, S11 e S12: são os cenários pessimistas, construídos considerando-se os log-retornos mínimos do mercado (obtido das observações históricas) e os coeficientes  $\beta_i$ , para cada  $i = 1, \dots, 4$ , fornecidos pelo modelo CAPM.

Tabela 5.43: Variação Máxima e Mínima dos Fatores de Risco.

	1 dia (StE)		10 dias (StE)		10 dias (DD)	
Ativo	Máx	Mín	Máx	Mín	Máx	Mín
$A_1$	5,826	-6,889	8,618	-20,611	17,709	-23,839
$A_2$	3,947	-4,497	10,960	-7,411	14,107	-9,607
$A_3$	5,471	-4,582	9,538	-9,923	11,441	-11,781
$A_4$	6,163	-4,536	12,884	-11,996	12,884	-14,443
$A_M$	6,560	-6,731	11,375	-27,224	13,022	-29,261

Tabela 5.44: Cenários Selecionados.

	$X_{1,S_j}$	$X_{2,S_j}$	$X_{3,S_j}$	$X_{4,S_j}$
S1	5,826	3,947	5,471	6,163
S2	8,618	10,960	9,538	12,884
S3	17,709	14,107	11,441	12,884
S4	-6,889	-4,497	-4,582	-4,536
S5	-20,611	-7,411	-9,923	-11,996
S6	-23,839	-9,607	-11,781	-14,443
S7	0,039	0,282	0,459	0,151
S8	0,068	0,489	0,796	0,262
S9	0,078	0,560	0,912	0,299
S10	-0,040	-0,290	-0,471	-0,155
S11	-0,163	-1,171	-1,906	-0,626
S12	-0,176	-1,258	-2,048	-0,673

Os retornos do *portfolio* sob cada cenário são apresentados na Tabela 5.45. Os retornos reais (obtidos a partir da série temporal observada) do *portfolio* para um 1 e 10 dias foram, respectivamente, 0,867 e 3,653. Observamos que os retornos reais (que são conhecidos) foram superiores aos previstos pelos cenários pessimistas (ver Tabela 5.45). Os cenários otimistas calculados a partir dos retornos máximos dos ativos superestimaram os retornos, enquanto que, os cenários otimistas que consideram o valor do coeficiente  $\beta_i$ , para cada  $i = 1, \dots, 4$ , fornecidos pelo modelo CAPM, forneceram estimativas inferiores aos retornos observados.

Tabela 5.45: Retornos do *Portfolio* sob os Diferentes Cenários.

Cenário	$r_{\mathcal{P},S_j}$	Cenário	$r_{\mathcal{P},S_j}$	Cenário	$r_{\mathcal{P},S_j}$	Cenário	$r_{\mathcal{P},S_j}$
S1	5,041	S4	-4,916	S7	0,300	S10	-0,308
S2	10,204	S5	-10,915	S8	0,521	S11	-1,247
S3	13,534	S6	-13,167	S9	0,597	S12	-1,340

### 5.6.9 Perda Máxima sob Domínios de Admissibilidade

Nesta subseção apresentamos a análise de cenários escolhidos sob um domínio de admissibilidade. A perda calculada sob esses cenários é denominada perda máxima, ou *MaxLoss*. Pela expressão (5.5), segue que o *portfolio* é linear, no sentido discutido na Subseção 4.6.4. Portanto, segue do Teorema 4.1, que a um nível de confiança  $p\%$ , a perda máxima do *portfolio* é dada pela expressão (4.68) e o cenário onde isso ocorre é dado pela expressão (4.69).

Na Tabela 5.46 apresentamos a perda máxima (*MaxLoss*) do *portfolio* calculada para diferentes valores de  $p$ , e os respectivos cenários para os quais essa perda ocorre. Por definição  $X_{i,ML}$  representa o log-retorno do ativo  $A_i$  sob o cenário de perda máxima (*MaxLoss*), para cada  $i = 1, \dots, 4$ .

Tabela 5.46: Perda Máxima para Diferentes Valores de  $p$  e seu Respetivo Cenário.

p	<i>MaxLoss</i>	Cenário			
		$X_{1,ML}$	$X_{2,ML}$	$X_{3,ML}$	$X_{4,ML}$
0,50	-1,0856	-0,7793	-0,9940	-1,4462	-0,4343
0,55	-1,1378	-0,8167	-1,0418	-1,5157	-0,4551
0,65	-1,2482	-0,8960	-1,1429	-1,6628	-0,4993
0,75	-1,3751	-0,9870	-1,2590	-1,8317	-0,5501
0,85	-1,5389	-1,1046	-1,4090	-2,0500	-0,6156
0,95	-1,8252	-1,3101	-1,6712	-2,4313	-0,7301
0,99	-2,1591	-1,5498	-1,9769	-2,8761	-0,8637

Observando as Tabelas 5.45 e 5.46 podemos perceber que as perdas que ocorrem sob os cenários pessimistas construídos sob dados históricos são maiores do que aquelas que ocorrem quando consideramos cenários sob um certo domínio de admissibilidade.

# Capítulo 6

## Análise de Séries Temporais Reais

Neste capítulo apresentamos a estimação e análise de medidas de risco para um *portfolio* real de ações. Este *portfolio* é formado por ações de quatro empresas brasileiras (também denominadas ativos do *portfolio*). Denotamos esses ativos por  $A_i$ ,  $i = 1, \dots, 4$ , de forma que:

$A_1$ : Representa as ações do Bradesco.

$A_2$ : Representa as ações da Brasil Telecom.

$A_3$ : Representa as ações da Gerdau.

$A_4$ : Representa as ações da Petrobrás.

Tais ações são negociadas na bolsa de valores de São Paulo (Bovespa). Utilizamos a notação  $A_M$  (ou, equivalentemente,  $A_5$ ) para denotar o mercado financeiro. Os valores do *portfolio* do mercado são representados pelo índice Bovespa (para obter maiores detalhes sobre este índice, veja o site <http://www.bovespa.com.br/Principal.asp>. Veja também <http://www.bovespa.com.br/Pdf/Indices/IBovespa.pdf>).

Os métodos utilizados neste capítulo, para o cálculo das medidas de risco, são descritos no Capítulo 4. Utilizamos os processos FIEGARCH, apresentados no Capítulo 2, para obter a variância condicional dos log-retornos dos ativos e calcular as medidas de risco VaR e ES. Também apresentamos o teste de estresse e o cálculo da perda máxima (*MaxLoss*), discutidos na Seção 4.6.

### 6.1 Características das Séries Temporais

Nesta seção apresentamos a análise das características das séries temporais dos log-retornos dos preços das ações dos ativos e dos log-retornos do índice de mercado. Tal análise é indispensável para a seleção de modelos para essas séries temporais. Para verificar se as séries temporais apresentam característica de longa dependência e selecionar o modelo mais adequado para cada uma delas, utilizamos os critérios descritos na Seção 3.1.

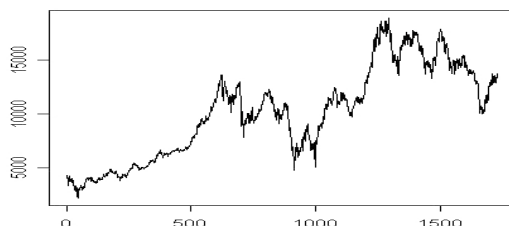


Figura 6.1: Índice Diário da Bolsa de Valores de São Paulo no Período de Janeiro de 1995 a Dezembro de 2001.

A Figura 6.1 apresenta a série temporal composta por  $n = 1.729$  observações que representa o índice diário da Bolsa de Valores de São Paulo no período de janeiro de 1995 a dezembro de 2001. A Figura 6.2 apresenta as séries temporais compostas por  $n = 1.729$  observações que representam, respectivamente, os preços das ações do Bradesco, Brasil Telecom, Gerdau e Petrobrás, no período de janeiro de 1995 a dezembro de 2001. Nas Figuras 6.1 e 6.2 observa-se que tanto o índice do mercado como os preços dos ativos sofreram uma forte queda em seus valores no período em torno de  $t = 1.000$  (15 de janeiro de 1999). Esse período corresponde à janeiro de 1999 e foi marcado pela desvalorização do real.

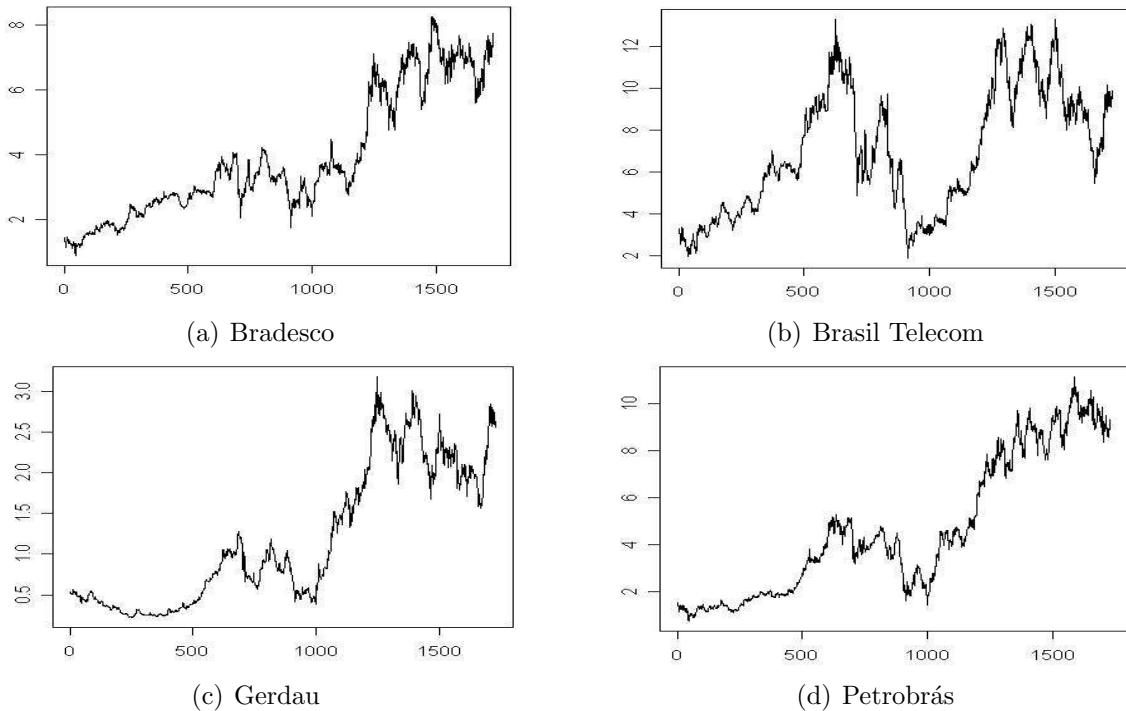


Figura 6.2: Séries Temporais dos Preços das Ações no Período de Janeiro de 1995 a Dezembro de 2001.

A Figura 6.3 apresenta as séries temporais dos log-retornos do IBovespa e dos preços das ações do Bradesco, Brasil Telecom, Gerdau e Petrobrás e os valores quadráticos das mesmas. Observamos, nestas séries temporais, as características comuns àquelas de retornos financeiros, tais como média aproximadamente em torno de zero e agrupamentos de volatilidade.

A Figura 6.4 apresenta o histograma, o  $Q \times Q$  plot e a função de autocorrelação amostral da série temporal dos log-retornos do IBovespa e dos log-retornos dos preços das ações do Bradesco, Brasil Telecom, Gerdau e Petrobrás. Observamos que, para todas as séries temporais, a distribuição dos dados é, aparentemente, leptocúrtica (possui caudas mais pesadas do que a distribuição normal). Percebemos que, tanto para os log-retornos do IBovespa, como para os log-retornos dos ativos, as variáveis aleatórias apresentam valores significativos da função de autocorrelação amostral  $\hat{\rho}_X(h)$ , para alguns valores de  $h > 1$ . Esse fato sugere o ajuste de um modelo do tipo ARMA. Para verificar se as séries temporais apresentam heteroscedasticidade condicional, observamos a função de autocorrelação amostral dos valores quadráticos dos log-retornos. Esta também nos fornece informações sobre a presença da característica de longa dependência.



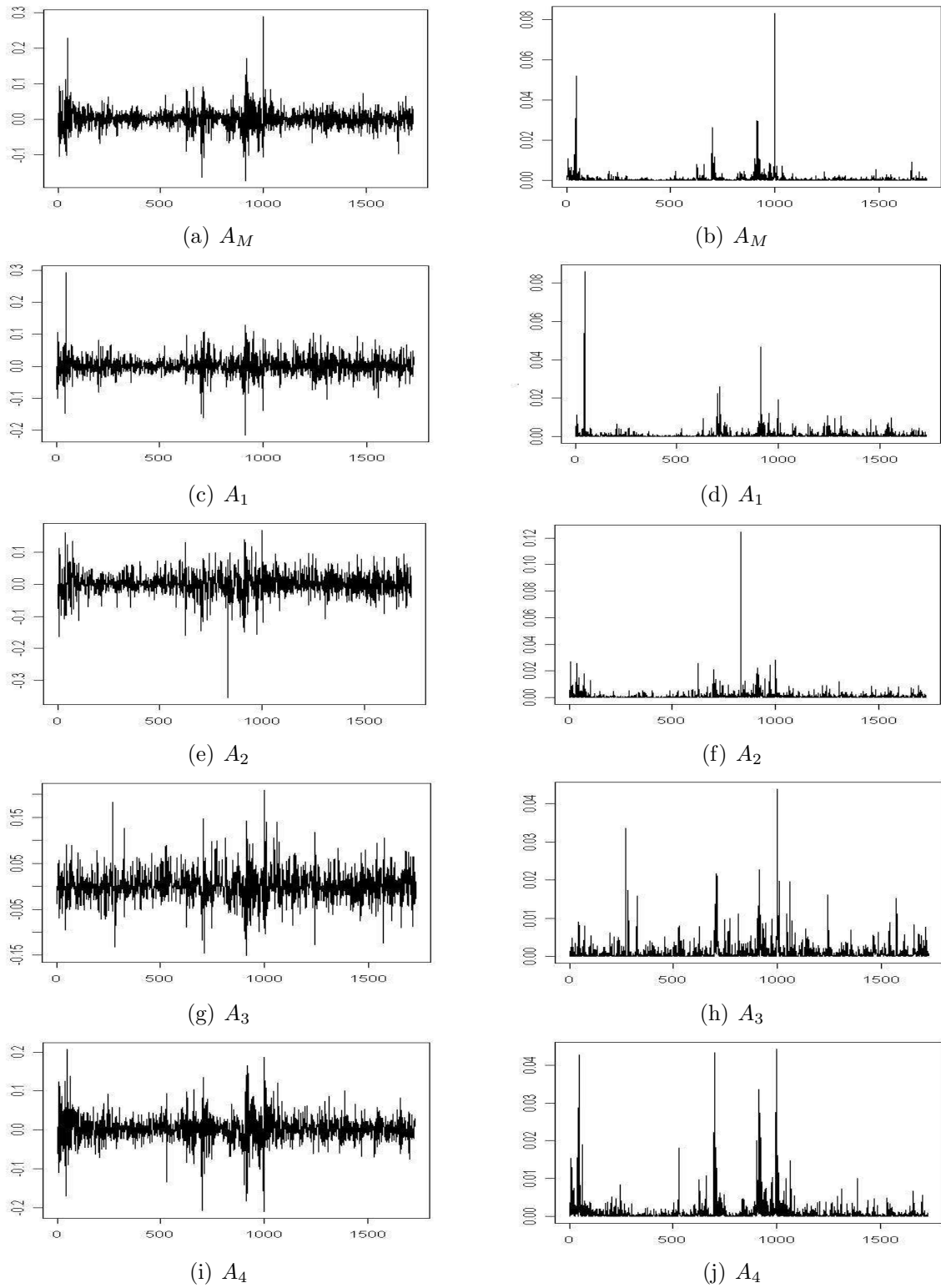


Figura 6.3: (a) Séries Temporais dos Log-retornos  $\{r_{i,t}\}_{t=1}^{1,728}$  e dos Valores Quadráticos dos Log-retornos  $\{r_{i,t}^2\}_{t=1}^{1,728}$ .

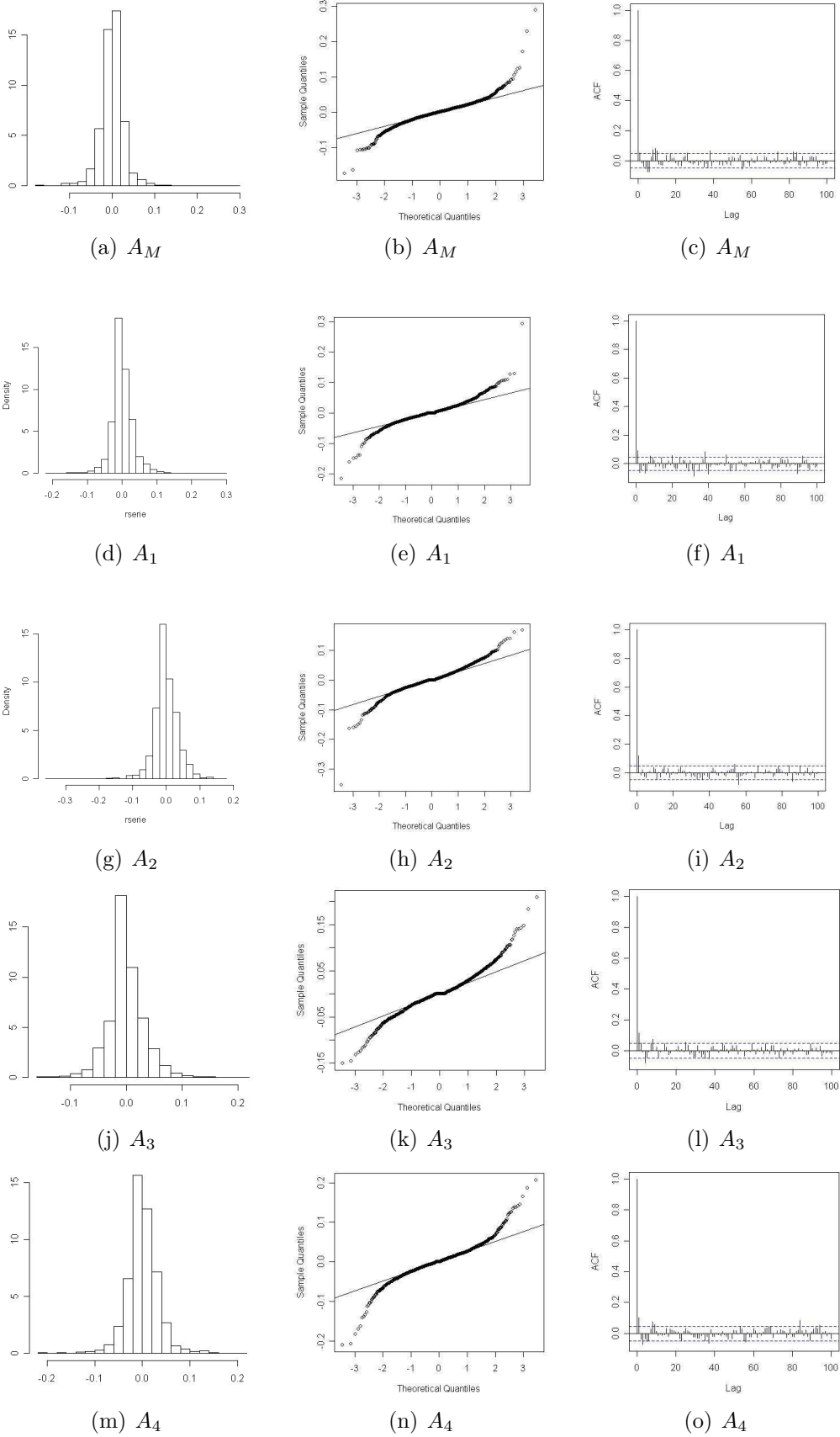


Figura 6.4: Histograma, Q×Q Plot e Função de Autocorrelação Amostral da Série Temporal dos Log-retornos.

No Teorema 2.9, do Capítulo 2, provamos que, sempre que  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo FIEGARCH( $p, d, q$ ), tal que  $\mathbb{E}([\ln(Z_t^2)]^2) < \infty$  e  $\mathbb{E}(Z_t \ln(Z_t^2)) < \infty$ , para todo  $t \in \mathbb{Z}$ , o processo  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, 0$ ) com inovações que, em geral, são não-Gaussianas. Com base neste resultado, analisamos a função de autocorrelação amostral das séries temporais dos valores quadráticos dos log-retornos e as do logaritmo dos valores quadráticos dos log-retornos, tanto para o mercado financeiro como para os quatro ativos, a fim de verificar a presença de heteroscedasticidade condicional e da característica de longa dependência.

A Figura 6.5 apresenta a função de autocorrelação amostral das séries temporais dos valores quadráticos dos log-retornos  $\{r_{M,t}^2\}_{t=1}^{1728}$  e do logaritmo dos valores quadráticos dos log-retornos  $\{\ln(r_{M,t}^2)\}_{t=1}^{1729}$  do índice do mercado (IBovespa). Observamos que tais funções apresentam valores significativos para grandes valores de  $h$ , o que indica a existência tanto de heteroscedasticidade quanto de longa dependência. A presença de longa dependência é confirmada pelas funções periodograma dessas séries temporais, apresentadas na Figura 6.6. As mesmas características são observadas nas funções de autocorrelação amostral e periodograma dos ativos, apresentadas, respectivamente, nas Figuras 6.7 e 6.8.

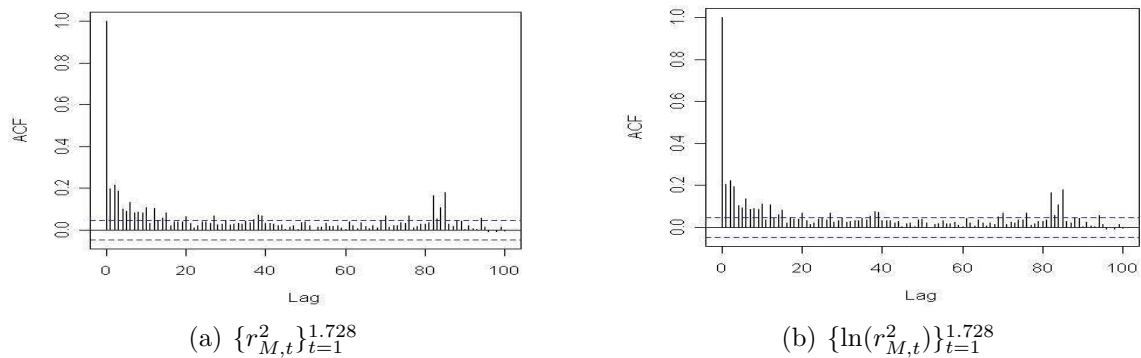


Figura 6.5: Função de Autocorrelação Amostral das Séries Temporais Reais.

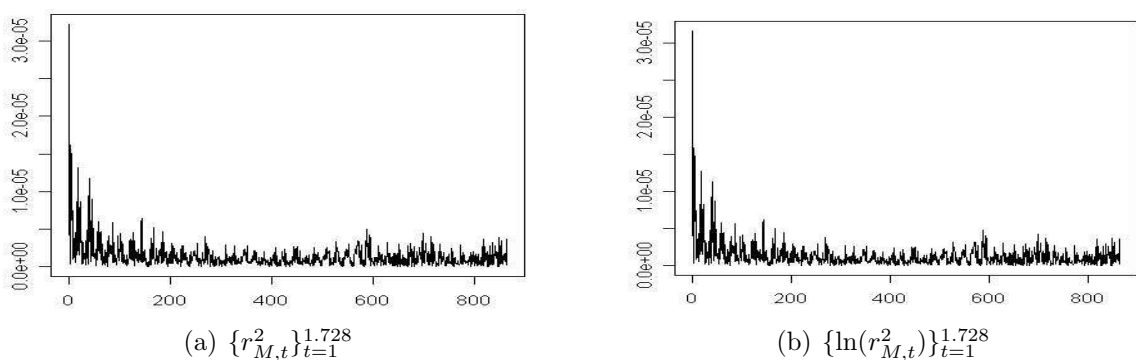


Figura 6.6: Função Periodograma das Série Temporais Reais.

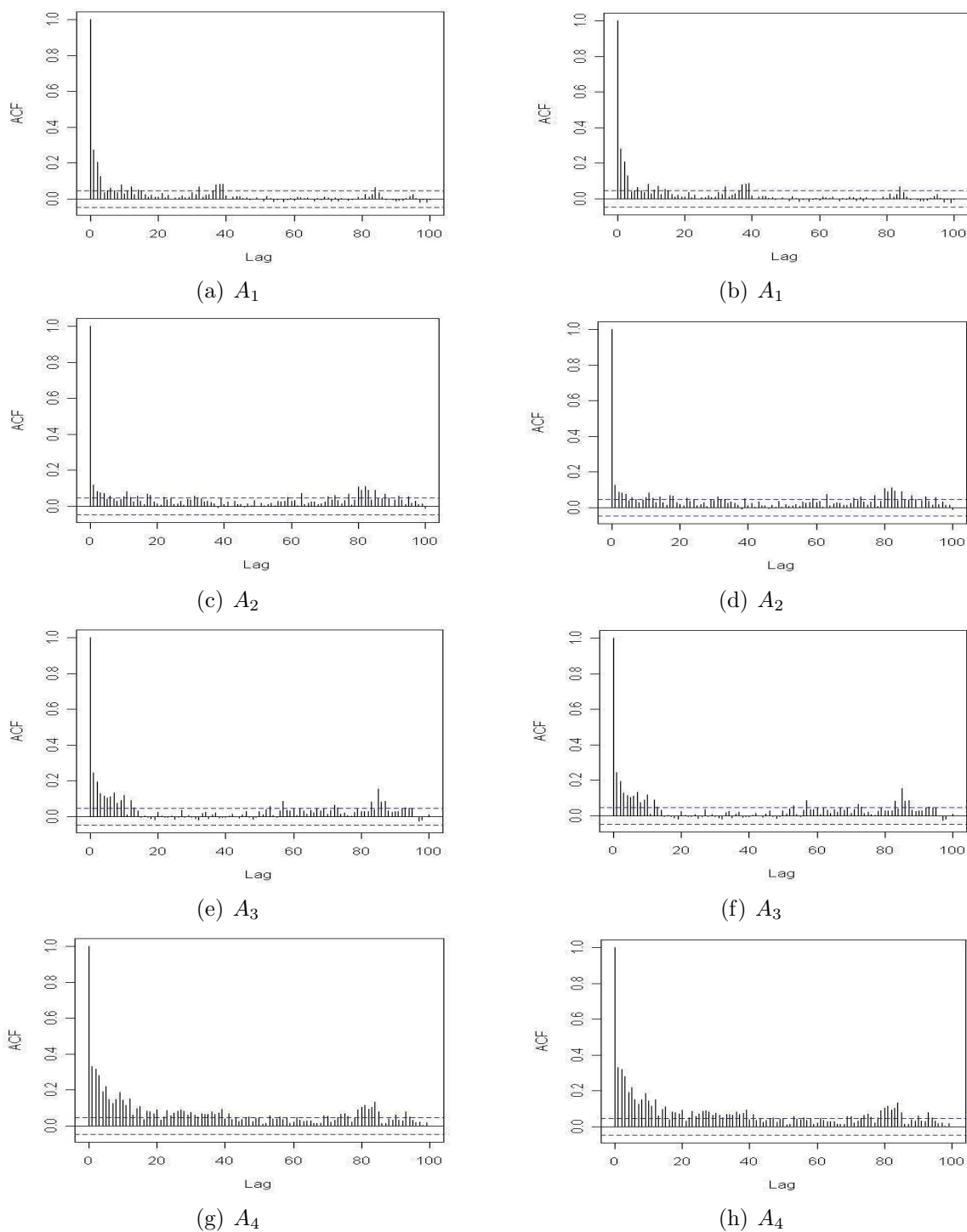
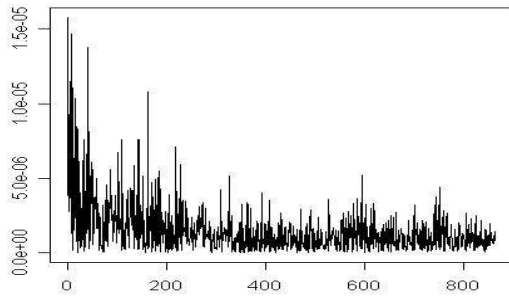
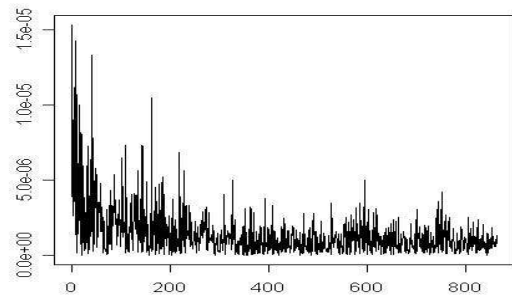


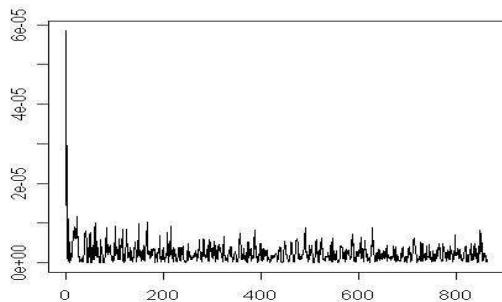
Figura 6.7: Função de Autocorrelação Amostral das Séries Temporais  $\{r_{i,t}^2\}_{t=1}^{1.728}$  e  $\{\ln(r_{i,t}^2)\}_{t=1}^{1.728}$ .



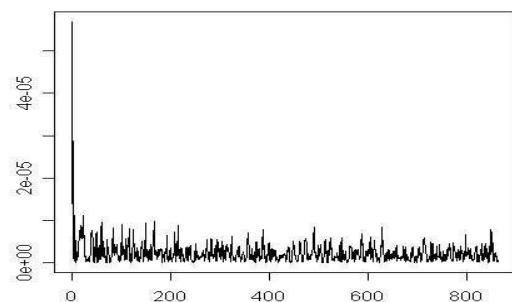
(a)  $A_1$



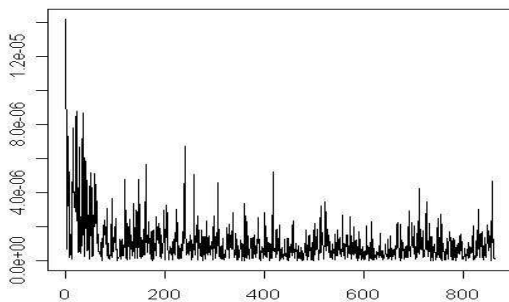
(b)  $A_1$



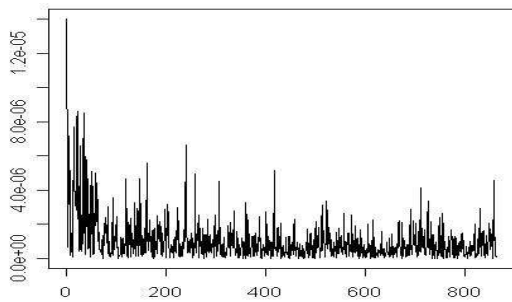
(c)  $A_2$



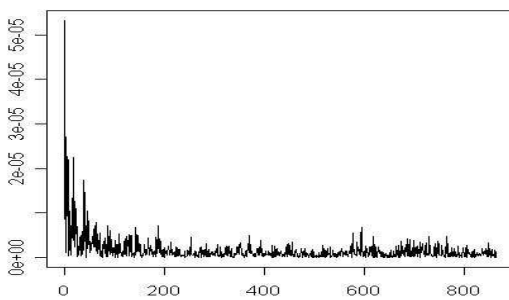
(d)  $A_2$



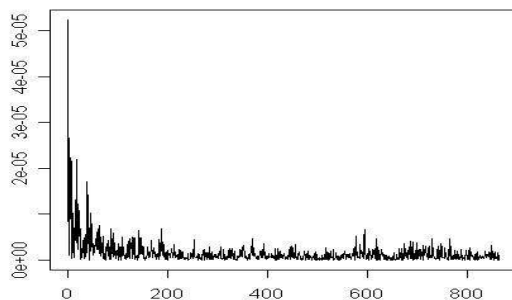
(e)  $A_3$



(f)  $A_3$



(g)  $A_4$



(h)  $A_4$

Figura 6.8: Função Periodograma das Séries Temporais  $\{r_{i,t}^2\}_{t=1}^{1.728}$  e  $\{\ln(r_{i,t}^2)\}_{t=1}^{1.728}$ .

Além da análise das funções de autocorrelação amostral e periodograma de cada série temporal, apresentamos também a estimação do parâmetro  $d$  de longa dependência. Para estimar o parâmetro de longa dependência utilizamos os procedimentos descritos na Subseção 3.1.6. Para os métodos R/S, R/S<sub>*m*</sub>, KPSS e V/S, utilizamos  $l = 40$  e  $l = 100$  como número mínimo de observações para cada bloco e para o estimador GPH utilizamos  $\alpha \in \{0,5; 0,6; 0,7; 0,8; 0,9\}$ .

A Tabela 6.1 apresenta o valor estimado do parâmetro  $d$  para as séries temporais dos valores quadráticos dos log-retornos do IBovespa e dos quatro ativos. A Tabela 6.2 apresenta o valor estimado do parâmetro  $d$  para o logaritmo dessas séries temporais. Observamos que os valores obtidos, utilizando-se o valor quadrático dos log-retornos e o logaritmo do valor quadrático dos log-retornos, foram muito próximos. Em ambos os casos, para o ativo  $A_3$  (Gerdau), o valor do parâmetro  $d$  estimado pelos métodos KPSS e V/S, com  $l = 40$ , foi negativo. Os demais valores estimados pertencem ao intervalo  $(0; 0,5)$  o que indica a presença de longa dependência.

Tabela 6.1: Valores dos Estimadores do  $d$  para a Série Temporal dos Valores Quadráticos dos Log-retornos do IBovespa e para as Séries Temporais dos Valores Quadráticos dos Log-retornos dos Preços dos Ativos.

Método		$\hat{d}$				
		IBovespa	Bradesco	Brasil Telecom	Gerdau	Petrobrás
R/S	$l = 40$	0,2731	0,1929	0,2295	0,1600	0,3039
	$l = 100$	0,3198	0,2708	0,3080	0,2493	0,4212
R/S <sub><i>m</i></sub>	$l = 40$	0,2364	0,1723	0,2133	0,1423	0,2620
	$l = 100$	0,2910	0,2463	0,2952	0,2264	0,3807
KPSS	$l = 40$	0,3587	0,2442	0,3037	-0,0313	0,3986
	$l = 100$	0,4883	0,4582	0,4100	0,0753	0,4440
V/S	$l = 40$	0,2833	0,1443	0,2687	-0,0131	0,2983
	$l = 100$	0,4502	0,3730	0,3345	0,0380	0,4274
GPH	$\alpha = 0,5$	0,2206	0,2522	0,2183	0,0916	0,2700
	$\alpha = 0,6$	0,2176	0,1757	0,1979	0,1963	0,2908
	$\alpha = 0,7$	0,2232	0,1456	0,0733	0,2420	0,2995
	$\alpha = 0,8$	0,2202	0,2207	0,0628	0,2041	0,3107
	$\alpha = 0,9$	0,1885	0,2046	0,0632	0,2130	0,2751

Tabela 6.2: Valores dos Estimadores do  $d$  para a Série Temporal do Logaritmo dos Valores Quadráticos dos Log-retornos do IBovespa e para as Séries Temporais dos Logaritmos dos Valores Quadráticos dos Log-retornos dos Preços dos Ativos.

Método		$\hat{d}$				
		IBovespa	Bradesco	Brasil Telecom	Gerdau	Petrobrás
R/S	$n = 40$	0,2762	0,1955	0,2335	0,1606	0,3054
	$n = 100$	0,3247	0,2754	0,3141	0,2504	0,4229
R/S <sub>m</sub>	$n = 40$	0,2391	0,1745	0,2168	0,1428	0,2634
	$n = 100$	0,2952	0,2502	0,3006	0,2275	0,3823
KPSS	$n = 40$	0,3604	0,2465	0,3041	-0,0313	0,3994
	$n = 100$	0,4900	0,4605	0,4137	0,0764	0,4459
V/S	$n = 40$	0,2850	0,1463	0,2691	-0,0134	0,2991
	$n = 100$	0,4518	0,3757	0,3382	0,0383	0,4290
GPH	$\alpha = 0,5$	0,2242	0,2557	0,2249	0,0928	0,2717
	$\alpha = 0,6$	0,2215	0,1797	0,2024	0,1974	0,2925
	$\alpha = 0,7$	0,2266	0,1480	0,0798	0,2432	0,3010
	$\alpha = 0,8$	0,2250	0,2241	0,0698	0,2053	0,3111
	$\alpha = 0,9$	0,1921	0,2086	0,0700	0,2137	0,2756

## 6.2 Ajuste dos Modelos FIEGARCh

Nesta seção apresentamos resultados relacionados à seleção e ao ajuste dos modelos FIEGARCh( $p, d, q$ ) às séries temporais dos log-retornos do IBovespa e dos ativos financeiros. Utilizamos tais modelos para a previsão da volatilidade e para os cálculos das medidas de risco, apresentados na Seção 6.4.

A análise das funções de autocorrelações amostrais, apresentada na Seção 6.1, sugere o ajuste de modelos ARMA( $p_1, q_1$ )-FIEGARCh( $p_2, d, q_2$ ), tanto para o IBovespa como para os ativos do *portfolio*. Enquanto os modelos ARMA( $p_1, q_1$ ) modelam a correlação entre as variáveis aleatórias, os modelos FIEGARCh( $p_2, d, q_2$ ) incorporam a existência da heteroscedasticidade e da característica de longa dependência.

Para a estimação dos parâmetros dos modelos utilizamos a função *fgarch* do S-Plus. Inicialmente, utilizamos  $p_1, q_1 \in \{0, 1, 2, 3\}$  e  $p_2, q_2 \in \{0, 1\}$ . O valor do parâmetro  $d$  é estimado pelo método da quase-verossimilhança. Selecionamos aqueles modelos que apresentaram os menores valores para os critérios AIC, BIC e/ou o maior valor da log-verossimilhança. Esses valores são apresentados, em negrito, na Tabela 6.3. Em seguida, realizamos a análise dos resíduos destes modelos. Observamos que, para a série temporal dos log-retornos dos preços das ações da Gerdau, o valor quadrático dos resíduos dos modelos selecionados ainda apresentavam heteroscedasticidade. O problema foi solucionado com o ajuste de um modelo com valor  $p_2 = 4$ .

Tabela 6.3: Valores dos Critérios AIC, BIC e Log-verossimilhança dos Primeiros Modelos Selecionados para as Séries Temporais dos Log-retornos do IBovespa e dos Preços das Ações do Bradesco, Brasil Telecom, Gerdau e Petrobrás.

	ARMA		FIEGARCH			Critério de Seleção		
	$p_1$	$q_1$	$p_2$	$d$	$q_2$	AIC	BIC	Log-verossimilhança
IBovespa	0	1	0	0,359	1	-8263,249	<b>-8225,066</b>	4138,625
	1	2	0	0,350	1	<b>-8269,095</b>	-8220,003	4143,548
	2	1	1	0,374	1	-8267,582	-8207,580	<b>4144,791</b>
Bradesco	0	1	0	0,451	1	-7700,331	<b>-7662,148</b>	3857,165
	2	2	1	0,447	1	-7708,651	-7643,194	<b>3866,325</b>
	2	2	0	0,453	1	<b>-7710,530</b>	-7655,983	3865,265
Brasil Telecom	0	1	0	0,584	0	-6921,569	<b>-6888,841</b>	3466,785
	0	1	1	0,451	1	<b>-6935,136</b>	-6886,043	3476,568
	2	2	1	0,430	1	-6931,100	-6865,644	<b>3477,550</b>
Gerdau	1	0	0	0,151	1	-7091,686	<b>-7053,503</b>	3552,843
	2	2	1	0,127	1	<b>-7103,078</b>	-7037,622	<b>3563,539</b>
Petrobrás	0	1	1	0,567	0	-7483,429	<b>-7439,792</b>	3749,715
	0	3	1	0,454	1	<b>-7490,259</b>	-7430,257	3756,130
	2	2	1	0,450	1	-7489,195	-7423,739	<b>3756,598</b>

Os modelos finais selecionados, após a análise dos resíduos, foram:

- Para a série temporal dos log-retornos  $\{r_{M,t}\}_{t=1}^{1.728}$  do IBovespa, optamos pelo modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 0 = p_2, q_1 = 1 = q_2$ , dado por

$$\begin{aligned}
 r_{M,t} &= X_{M,t} - 0,078X_{M,t-1} \\
 X_{M,t} &= \sigma_{M,t}Z_{M,t} \\
 (1 - 0,706\mathcal{B})(1 - \mathcal{B})^{0,339} \ln(\sigma_{M,t}^2) &= -0,346 + 0,275|Z_{M,t-1}| - 0,166Z_{M,t-1}.
 \end{aligned}$$

A análise de resíduos completa é apresentada na Tabela F.5 do Apêndice F.

- Para a série temporal dos log-retornos  $\{r_{1,t}\}_{t=1}^{1.728}$  dos preços das ações do Bradesco, optamos pelo modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 0 = p_2, q_1 = 1 = q_2$ , dado por

$$\begin{aligned}
 r_{1,t} &= X_{1,t} - 0,129X_{1,t-1} \\
 X_{1,t} &= \sigma_{1,t}Z_{1,t} \\
 (1 - 0,453\mathcal{B})(1 - \mathcal{B})^{0,446} \ln(\sigma_{1,t}^2) &= -0,374 + 0,381|Z_{1,t-1}| - 0,135Z_{1,t-1}.
 \end{aligned}$$

A análise de resíduos completa é apresentada na Tabela F.6 do Apêndice F.

- Para a série temporal dos log-retornos  $\{r_{2,t}\}_{t=1}^{1.728}$  dos preços das ações da Brasil Telecom, optamos pelo modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 0,$



$q_1 = 1, p_2 = 1 = q_2$ , dado por

$$\begin{aligned} r_{2,t} &= 0,002 + X_{2,t} - 0,103X_{2,t-1} \\ X_{2,t} &= \sigma_{2,t}Z_{2,t} \\ (1 - 0,905\mathcal{B})(1 - \mathcal{B})^{0,447} \ln(\sigma_{2,t}^2) &= -0,053 + 0,382|Z_{2,t-1}| - 0,331|Z_{2,t-2}| \\ &\quad + 0,044Z_{2,t-1} - 0,066Z_{2,t-2}. \end{aligned}$$

A análise de resíduos completa é apresentada na Tabela F.7 do Apêndice F.

- Para a série temporal dos log-retornos  $\{r_{3,t}\}_{t=1}^{1.728}$  dos preços das ações da Gerdau, optamos pelo modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 1, q_1 = 0 = q_2$  e  $q_2 = 4$ , dado por

$$\begin{aligned} r_{1,t} &= -0,1409r_{3,t-1} + X_{3,t} \\ X_{3,t} &= \sigma_{3,t}Z_{3,t} \\ \beta(\mathcal{B})(1 - \mathcal{B})^{0,256} \ln(\sigma_{3,t}^2) &= -0,769 + 0,395|Z_{3,t-1}| - 0,046Z_{3,t-1}, \end{aligned}$$

com  $\beta(\mathcal{B}) = 1 - 0,216\mathcal{B} - 0,184\mathcal{B}^2 - 0,470\mathcal{B}^3 + 0,450\mathcal{B}^4$ .

A análise de resíduos completa é apresentada na Tabela F.8 do Apêndice F.

- Para a série temporal dos log-retornos  $\{r_{4,t}\}_{t=1}^{1.728}$  dos preços das ações da Petrobrás optamos pelo modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 0 = p_2, q_1 = 1 = q_2$ , dado por

$$\begin{aligned} r_{4,t} &= 0,001 + X_{4,t} - 0,119X_{4,t-1} \\ X_{4,t} &= \sigma_{4,t}Z_{4,t} \\ (1 - 0,575\mathcal{B})(1 - \mathcal{B})^{0,416} \ln(\sigma_{4,t}^2) &= -0,347 + 0,326|Z_{4,t-1}| - 0,110Z_{4,t-1}. \end{aligned}$$

A análise de resíduos completa é apresentada na Tabela F.9 do Apêndice F.

## 6.2.1 Previsão da Volatilidade

Nesta subseção apresentamos a previsão da média e da variância condicional para  $h = 1$  até  $h = 10$  passos à frente, para as séries temporais analisadas. Estes resultados são utilizados para o cálculo das medidas de risco do *portfolio*, que apresentamos na Seção 6.4. Resultados teóricos sobre a previsão para os modelos ARMA e FIEGARCH são apresentados na Seção 3.4.

Lembramos que, se  $\{r_t\}_{t \in \mathbb{Z}}$  é um processo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), definido pelas expressões (4.48), (4.49) e (4.50), então, pela expressão (3.39), a previsão da média  $\hat{r}_{n+h}$ ,  $h$  passos à frente é dada por

$$\hat{r}_{n+h} = \phi_0 + \sum_{i=1}^{p_1} \phi_i \hat{r}_{n+h-i} - \sum_{j=0}^{q_1} \theta_j \hat{X}_{n+h-j},$$

com  $\hat{r}_{n+h} = r_{n+h}$ , se  $h \leq 0$ ;  $\hat{X}_{n+h} = X_{n+h}$ , se  $h \leq 0$  e  $\hat{X}_{n+h} = 0$ , se  $h > 0$ , pois  $\hat{X}_{n+h} = \mathbb{E}(X_{n+h} | \mathcal{F}_n) = 0$ . A previsão do logaritmo da variância condicional  $h$  passos à frente é dada pela expressão (3.55). A previsão da média e da volatilidade, para as séries temporais analisadas, são apresentadas na Tabela 6.4.

Tabela 6.4: Previsão da Média e da Variância Condicional dos Log-retornos dos IBovespa ( $A_M$ ) e das Ações do Bradesco ( $A_1$ ), Brasil Telecom ( $A_2$ ), Gerdau ( $A_3$ ) e Petrobrás ( $A_4$ ) para os Horizontes  $h = 1, \dots, 10$ .

$h$	$A_M$		$A_1$		$A_2$		$A_3$		$A_4$	
	$\hat{r}_{M,n+h}$	$\hat{\sigma}_{M,n+h}$	$\hat{r}_{1,n+h}$	$\hat{\sigma}_{1,n+h}$	$\hat{r}_{2,n+h}$	$\hat{\sigma}_{2,n+h}$	$\hat{r}_{3,n+h}$	$\hat{\sigma}_{3,n+h}$	$\hat{r}_{4,n+h}$	$\hat{\sigma}_{4,n+h}$
1	-0,0012	0,0179	0,0031	0,0215	-0,0030	0,0319	-0,0041	0,0256	0,0019	0,0176
2	0,0000	0,0182	0,0000	0,0221	0,0018	0,0219	0,0000	0,0248	0,0014	0,0179
3	0,0000	0,0182	0,0000	0,0226	0,0018	0,0194	0,0000	0,0195	0,0014	0,0185
4	0,0000	0,0182	0,0000	0,0230	0,0018	0,0181	0,0000	0,0293	0,0014	0,0190
5	0,0000	0,0182	0,0000	0,0233	0,0018	0,0175	0,0000	0,0235	0,0014	0,0194
6	0,0000	0,0182	0,0000	0,0236	0,0018	0,0171	0,0000	0,0219	0,0014	0,0197
7	0,0000	0,0182	0,0000	0,0238	0,0018	0,0169	0,0000	0,0280	0,0014	0,0200
8	0,0000	0,0182	0,0000	0,0239	0,0018	0,0168	0,0000	0,0219	0,0014	0,0202
9	0,0000	0,0182	0,0000	0,0241	0,0018	0,0169	0,0000	0,0233	0,0014	0,0204
10	0,0000	0,0182	0,0000	0,0242	0,0018	0,0170	0,0000	0,0261	0,0014	0,0206

Observamos que, para  $h > 1$ , o valor de  $\sigma_{M,n+h}$  é constante, o que não ocorre para os ativos  $A_i$ , para cada  $i = 1, \dots, 4$ . Aparentemente, para os ativos  $A_1$  e  $A_4$  estes valores estão crescendo, enquanto que para o ativo  $A_2$ , os valores decrescem. Para o ativo  $A_3$  o comportamento parece ser aleatório.

### 6.3 Estatísticas Descritivas dos Ativos Financeiros

Na Tabela 6.5 apresentamos as estatísticas descritivas das séries temporais analisadas. Observamos que todos os ativos apresentam desempenho semelhante, quantificado pelo valor da mediana dos retornos diários igual a zero em todos os casos. Entretanto, esse valor está abaixo daquele apresentado pelo mercado (0,0014) no mesmo período. O ativo que apresentou maior variabilidade foi o  $A_2$ , com variância igual a 0,0013, sendo esse valor maior do que a variância do mercado, que é igual a 0,0008. O ativo  $A_1$  é o menos volátil, com variância igual a 0,0009. Entre todos os ativos, o que apresentou o menor retorno no período foi o ativo  $A_2$  (-0,3530), e este foi menor do que o menor retorno apresentado pelo mercado no mesmo período (-0,1721). O ativo que apresentou o maior retorno no período foi o  $A_1$  (0,2933) e foi superior ao maior retorno do mercado (0,2883) no período. Percebemos também que, para todas as séries temporais de log-retornos, a função de distribuição dos dados é aproximadamente simétrica, dado que a medida de assimetria assume valores pequenos. Entretanto, todas as séries temporais podem ser consideradas como de caudas pesadas. A série temporal dos log-retornos do ativo  $A_1$  apresenta o maior valor da medida de curtose (11,9990) e esse valor é menor do que o da medida de curtose do mercado (15,5581).

Tabela 6.5: Estatísticas Descritivas para as Séries Temporais dos Log-retornos do Índice do Mercado ( $A_M$ ) e dos Preços das Ações do Bradesco ( $A_1$ ), Brasil Telecom ( $A_2$ ), Gerdau ( $A_3$ ) e Petrobrás ( $A_4$ ).

Ativo	$A_1$	$A_2$	$A_3$	$A_4$	$A_M$
Média	0,0010	0,0006	0,0009	0,0010	0,0007
Variância	0,0009	0,0013	0,0011	0,0011	0,0008
Mediana	0,0000	0,0000	0,0000	0,0000	0,0014
Máximo	0,2933	0,1676	0,2092	0,2068	0,2883
Mínimo	-0,2162	-0,3530	-0,1507	-0,2103	-0,1721
Assimetria	0,2390	-0,5773	0,3675	-0,0738	0,6920
Curtose	11,9990	10,6346	6,5431	9,1426	15,5581

### 6.3.1 Cálculo dos Pesos do *Portfolio*

Nesta seção apresentamos o cálculo do *portfolio* eficiente utilizando o método descrito na Seção 4.3. O *portfolio* é representado por  $\mathcal{P} = \{A_1, A_2, A_3, A_4\}$ , onde  $A_1$  representa as ações do Bradesco,  $A_2$  as ações da Brasil Telecom,  $A_3$  as ações da Gerdau e  $A_4$  as ações da Petrobrás. O vetor de pesos é representado por  $\mathbf{a} = (a_1, a_2, a_3, a_4)'$ . Utilizamos a taxa SELIC, cujo log-retorno médio no período é igual a  $R_F = 0,02$ , como ativo livre de risco. Sob essas condições, utilizando a expressão (4.22), obtemos

$$(a_1, a_2, a_3, a_4) = (0,3381; 0,1813; 0,3087; 0,1719), \quad (6.1)$$

isto é, 33,81% do valor total a ser investido será aplicado no ativo  $A_1$ ; 18,13% no ativo  $A_2$ ; 30,87% no ativo  $A_3$  e 17,19% no ativo  $A_4$ .

### 6.3.2 Modelo CAPM para as Séries Temporais

Nesta subseção apresentamos o cálculo do coeficiente  $\beta$  para os ativos do *portfolio*. Como citado na subseção anterior, estamos considerando, como ativo livre de risco, a taxa SELIC. Na Tabela 6.7 apresentamos a correlação amostral entre as séries temporais analisadas. Observamos que os ativos possuem correlação significativa, tanto entre si, como com o mercado financeiro.

Os métodos de estimação para o coeficiente angular  $\beta$  do modelo CAPM são descritos na Subseção 4.3.2. Os valores estimados do parâmetro  $\beta_i$ , para cada  $i = 1, \dots, 4$  são dados por

$$\beta_1 = 0,9525, \quad \beta_2 = 1,0508, \quad \beta_3 = 1,1267 \quad \text{e} \quad \beta_4 = 1,0153.$$

Note que, para o ativo  $A_1$ ,  $\beta_1 < 1$ , o que significa que esse ativo é defensivo. No entanto,  $\beta_i > 1$ , para  $i = 2, 3$ , o que significa que esses ativos são agressivos, isto é, se houverem quedas no índice do mercado, os retornos dos ativos sofrerão quedas maiores (em termos de porcentagem) do que a queda do índice do mercado. De forma análoga, se o índice do

mercado sofrer um determinado crescimento, o aumento sofrido pelos retornos será maior. Pelas expressões (4.31) e (6.1) segue que o  $\beta_{\mathcal{P}}$  do *portfolio* é dado por

$$\beta_{\mathcal{P}} = 0,3381 \times 0,9525 + 0,1813 \times 1,0508 + 0,3087 \times 1,1267 + 0,1719 \times 1,0153 = 1,0349.$$

Obtivemos  $\beta_{\mathcal{P}} = 1,0349$ , o que indica que, se o índice do mercado sofrer um aumento de 10%, o valor do *portfolio* sofrerá um aumento de 10,35%. Note que, sob o modelo CAPM, o retorno esperado do ativo  $A_i$  é dado por

$$\mathbb{E}(R_i) = R_F + \beta_i(\mathbb{E}(R_M) - R_F), \quad \text{para cada } i = 1, \dots, 4.$$

O valor  $\mathbb{E}(R_M)$  é estimado através da média amostral dos retornos do mercado. Os valores da média amostral das séries temporais foram obtidos da Tabela 6.5. Lembramos que o log-retorno médio do ativo livre de risco para o período é igual a 0,02. Segue que

$$\mathbb{E}(R_i) = 0,02 + \beta_i(0,0007 - 0,02), \quad \text{para cada } i = 1, \dots, 4.$$

Assim, o retorno esperado para o período de investimento, segundo o modelo CAPM, é dado por

$$\mathbb{E}(R_1) = 0,0016, \quad \mathbb{E}(R_2) = -0,0003, \quad \mathbb{E}(R_3) = -0,0017 \quad \text{e} \quad \mathbb{E}(R_4) = 0,0004.$$

A Tabela 6.6 apresenta os valores (observados) dos retornos de períodos 1 e 10 dias, que são dados, respectivamente, por  $X_{i,1.991}$  e  $X_{i,1.990}$ [10], para cada  $i = 1, \dots, 4$ , a média amostral dos retornos observados e os retornos estimados pelo modelo CAPM. Comparando com

Tabela 6.6: Retornos dos Ativos.

Ativo	$R_i$	$R_i[10]$	$\mathbb{E}(R_i)$	$\bar{R}_i$
$A_1$	-0,0026	-0,0156	0,0016	-0,013
$A_2$	0,0301	-0,0343	-0,0003	-0,017
$A_3$	0,0680	0,2456	-0,0017	-0,040
$A_4$	0,0021	-0,0551	0,0004	0,003

**Nota:**  $R_i \equiv X_{i,1.991}$  e  $R_i[10] \equiv X_{i,1.990}[10]$ .

os valores obtidos pelo modelo CAPM notamos que, para o ativo  $A_1$  (Bradesco) o log-retorno observado foi menor que o retorno estimado, enquanto que, para os outros ativos os resultados obtidos através do modelo CAPM foram inferiores aos observados. Para um período  $h = 10$  dias o retorno previsto pelo modelo CAPM só foi inferior aos log-retornos observados no caso do ativo  $A_3$ .

O coeficiente  $\alpha$  das ações, que indica se elas estão subavaliadas ou superavaliadas, é definido pela expressão (4.29). Para este *portfolio* temos:

$$\alpha_1 = -0,0006, \quad \alpha_2 = 0,0009, \quad \alpha_3 = 0,0026 \quad \text{e} \quad \alpha_4 = 0,0006.$$

Na visão do investidor, as ações do ativo  $A_1$  estão superavaliadas pelo mercado ( $\alpha < 0$ ) o que deverá proporcionar uma rentabilidade inferior àquela calculada pelo modelo de equilíbrio (CAPM). Portanto, devem ser vendidas. Enquanto que, para os outros ativos, o valor positivo do coeficiente  $\alpha$  indica que, as ações estão subavaliadas, isto é, seu preço é menor do que seu valor intrínseco. Sendo assim, é conveniente comprá-las.

Tabela 6.7: Matriz de Correlação Amostral dos Log-retornos.

	IBovespa	Bradesco	Brasil Telecom	Gerdau	Petrobrás
IBovespa	1,0000	0,7024	0,7817	0,7379	0,8373
Bradesco	0,7024	1,0000	0,6887	0,4517	0,5982
Brasil Telecom	0,7817	0,6887	1,0000	0,5895	0,6435
Gerdau	0,7379	0,4517	0,5895	1,0000	0,5927
Petrobrás	0,8373	0,5982	0,6435	0,5927	1,0000

## 6.4 Cálculo das Medidas de Risco VaR e ES

Nesta seção calculamos as medidas de risco baseadas na distribuição das perdas do *portfolio*. Como sugerido por Palaro e Hotta (2006), utilizamos a série de log-retornos do *portfolio*, definida pela expressão (4.14). O cálculo das medidas de risco VaR e ES foi realizado utilizando-se as diferentes abordagens discutidas na Seção 4.5.2. Apresentamos também o cálculo das medidas de risco VaR e ES, considerando tanto a distribuição condicional, como a distribuição não-condicional do vetor de mudanças de fatores de risco. Para isso utilizamos, respectivamente, a abordagem RiskMetrics (modelo EWMA) e o método da variância-covariância. Por fim, apresentamos a análise das medidas de risco para cada ativo do *portfolio*. Lembramos que a medida ES é uma medida de risco coerente, no sentido da Definição 4.2. Além disso, por definição,  $\text{VaR} \leq \text{ES}$ . Sendo assim, obtendo o ES de cada ativo obtivemos uma cota superior para o VaR do *portfolio*.

Para os cálculos realizados fixamos:

- os fatores de risco para este *portfolio* são os logaritmos dos preços das ações. Sendo assim, o vetor de fatores de risco é dado por

$$\mathbf{Z}_t = (Z_{1,t}, Z_{2,t}, Z_{3,t}, Z_{4,t})' = (\ln(P_{1,t}), \ln(P_{2,t}), \ln(P_{3,t}), \ln(P_{4,t}))',$$

onde  $P_{i,t}$  é o preço da ação do ativo  $A_i$ , no instante  $t$ , para  $i = 1, 2, 3, 4$ . Como consequência, o vetor de mudanças de fatores de risco  $\mathbf{X}_t$ , definido pela expressão (4.34), é dado por

$$\mathbf{X}_t = (X_{1,t}, X_{2,t}, X_{3,t}, X_{4,t})' = (r_{1,t}, r_{2,t}, r_{3,t}, r_{4,t})',$$

onde  $r_{i,t}$ , para  $i = 1, 2, 3$  e  $4$ , é o log-retorno do ativo  $A_i$ , no instante  $t$ ;

- a perda do *portfolio* no instante  $t$  é representada pela variável aleatória  $L_t$ , definida pela expressão (4.32). Na prática, os retornos simples e os log-retornos são valores muito próximos e pela expressão (4.39), apresentada no Exemplo 4.4, assumimos que a perda do *portfolio*, no tempo  $t$ , é dada por

$$L_t = -V_{t-1} \sum_{i=1}^4 a_i r_{i,t} = -V_{t-1} r_{\mathcal{P},t}, \quad (6.2)$$

onde  $V_{t-1}$  é o valor do *portfolio* no instante  $t - 1$  e  $\mathbf{a} = (a_1, a_2, a_3, a_4)'$  é o vetor de pesos dado pela expressão (6.1),  $r_{i,t}$  é o log-retorno do ativo  $A_i$  no tempo  $t$  e  $r_{\mathcal{P},t}$  é o log retorno do *portfolio* no tempo  $t$ .

- a função de distribuição considerada para o cálculo das medidas de risco, para todas as abordagens foi a normal multivariada (ou univariada).

Note que, pela expressão (4.38), se  $c_i$  é o número de ações do ativo  $A_i$ , então

$$c_i = \frac{a_i V_0}{P_{i,0}}, \quad \text{para cada } i = 1, \dots, 4,$$

onde  $V_0$  é o capital inicial investido no *portfolio*,  $P_{i,0}$  é o preço unitário no instante inicial das ações do ativo  $A_i$ , para cada  $i = 1, \dots, 4$  e  $\mathbf{a} = (a_1, a_2, a_3, a_4)'$  são os pesos do *portfolio*, dados pela expressão (6.1). Sendo assim, pela expressão (4.36), o valor  $V_t$  do *portfolio*, no instante  $t$ , é dado por

$$V_t = V_0 \left( \frac{a_1}{P_{1,0}} P_{1,t} + \frac{a_2}{P_{2,0}} P_{2,t} + \frac{a_3}{P_{3,0}} P_{3,t} + \frac{a_4}{P_{4,0}} P_{4,t} \right), \quad t = 1, \dots, n,$$

onde  $n$  é o tamanho amostral,  $P_{i,t}$  é o preço unitário das ações do ativo  $A_i$ , para cada  $i = 1, \dots, 4$ , no instante  $t$ . Como exemplo, assumimos que o instante inicial é igual ao primeiro dia de observação das séries temporais. A Figura 6.9 apresenta a série temporal dos valores do *portfolio*  $\{V_t\}_{t=1}^n$  e a dos log-retornos do *portfolio* para o período de janeiro de 1995 à dezembro de 2001. Comparando a série temporal dos valores do *portfolio* com a série temporal do índice de mercado, apresentada na Figura 6.1, observamos que ambas apresentam um comportamento muito semelhante.

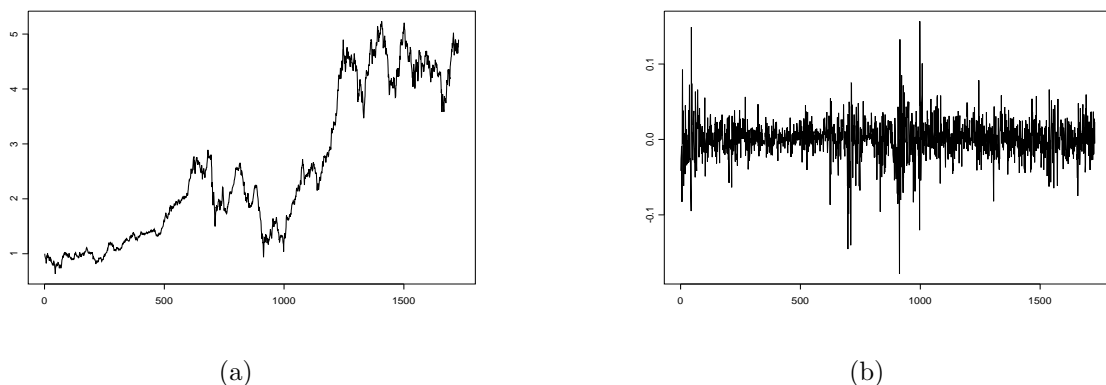


Figura 6.9: Séries Temporais: (a) Valores do *Portfolio*  $\mathcal{P}$ ; (b) Log-retornos do *Portfolio* no Período de Janeiro de 1995 à Dezembro de 2001.

A série temporal das perdas do *portfolio* e sua função densidade de probabilidade são apresentadas na Figura 6.10. Observamos que a maior perda ocorre em  $t = 698$  com  $L_t = 0,3729$  (ou seja, o maior ganho, que é igual a aproximadamente R\$ 0,37 por cada real investido) e a menor perda ocorre em  $t = 1.244$ , com  $L_t = -0,3402$  (ou seja, aproximadamente R\$ 0,34 por cada real investido). A maior perda corresponde a mudança no valor do *portfolio*, ocorrida entre 24/10/1997 (sexta-feira) e 27/10/1997 (segunda-feira). Nesse período o valor do índice Bovespa passou de 11.545,20 para 9.816,80 pontos, o que representa uma queda de 14,97%. Nessa data também ocorreram grandes quedas nos índices Dow Jones ( 7,18%) e S&P 500 (6,87%). Este período é marcado pela forte crise na Ásia. O maior ganho no *portfolio* corresponde a variação ocorrida no período de 13/01/2001 à 14/01/2001. O IBovespa apresenta uma alta de 2,08% neste período.

Enquanto que, para o mercado mundial temos que em 14/01/2000 o valor do índice Dow Jones foi de 11.722,98 pontos. Esse valor só foi superado em 03/10/2006, quando o índice atingiu 11.727,34 pontos.

Essa análise nos fornece uma idéia do que aconteceria se um investidor mantivesse seu *portfolio* constante durante o período considerado. Na prática, o investidor compra e vende ações a fim de obter melhores rendimentos. Sendo assim, a análise das medidas de risco é feita considerando-se ou a série temporal dos log-retornos do *portfolio* ou (equivalentemente) as séries temporais dos log-retornos dos ativos do *portfolio*.

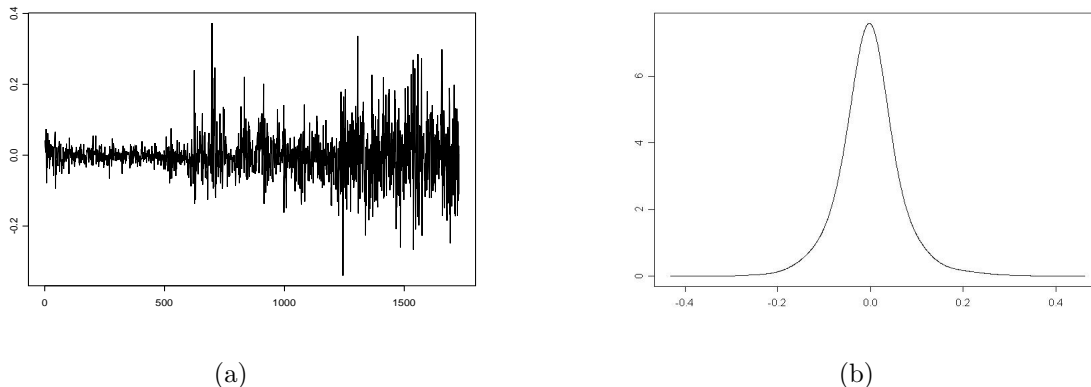


Figura 6.10: (a) Série Temporal das Perdas do *Portfolio* no Período de Janeiro de 1995 à Dezembro de 2001; (b) Função Densidade de Probabilidade das Perdas do *Portfolio*.

Apresentamos a seguir os resultados da estimação das medidas de risco VaR e ES considerando a série temporal dos log-retornos do *portfolio*. Para calcular o  $\text{VaR}_{p,n+1}$  do *portfolio*, onde  $n = 1.729$  é o tamanho amostral e  $p$  é o nível de confiança, lembramos que

$$\text{VaR}_{p,n+1} = \inf\{\ell \in \mathbb{R} : \mathbb{P}(L_{n+1} \geq \ell) \leq 1 - p\},$$

onde  $L_{n+1}$  é dada pela expressão (6.2). Pela expressão (4.42)

$$\mathbb{E}(L_{n+1}) = -V_n \mathbf{a}' \boldsymbol{\mu} \quad \text{e} \quad \text{Var}(L_{t+n}) = V_n^2 \mathbf{a}' \boldsymbol{\Sigma} \mathbf{a},$$

onde  $\mathbf{a} = (a_1, a_2, a_3, a_4)'$  é o vetor de pesos, dado pela expressão (6.1),  $\boldsymbol{\mu}$  é o vetor das médias dos ativos, dado por

$$\boldsymbol{\mu} = (\mathbb{E}(r_{1,n+1}), \mathbb{E}(r_{2,n+1}), \mathbb{E}(r_{3,n+1}), \mathbb{E}(r_{4,n+1}))',$$

$\boldsymbol{\Sigma}$  é a matriz de variâncias - covariâncias e  $V_n$  é o valor do *portfolio*, que é conhecido no instante  $n$  (neste caso,  $n$  é o tamanho amostral). Assumimos que, tanto a função de distribuição condicional como a não-condicional das perdas é a função de distribuição normal e utilizamos as abordagens descritas na Subseção 4.5.2 para o cálculo das medidas de risco.

Para o cálculo do VaR e do ES sob a abordagem econométrica, observamos que a função de autocorrelação amostral da série temporal dos log-retornos do *portfolio* apresenta valores significativos para  $h > 1$ . Para modelar essa dependência utilizamos modelos  $\text{ARMA}(p_1, q_1)$ . A função de autocorrelação amostral dos valores quadráticos dessa série

temporal indica heteroscedasticidade e longa dependência, o que é confirmado pelos valores da função periodograma, apresentada na Figura 6.12. Apresentamos os valores das funções de autocorrelação amostral destas duas séries temporais na Figura 6.11. O modelo  $ARMA(p_1, q_1)$ -EGARCH( $p_2, q_2$ ) selecionado, após a análise de resíduos, foi o que apresentou o menor número de parâmetros e é apresentado na Tabela F.10 do Apêndice F. O critério de seleção do modelo  $ARMA(p_1, q_1)$ -FIEGARCH( $p_2, q_2$ ) para a série temporal dos log-retornos do *portfolio* é apresentado na Seção F.2 do Apêndice F.

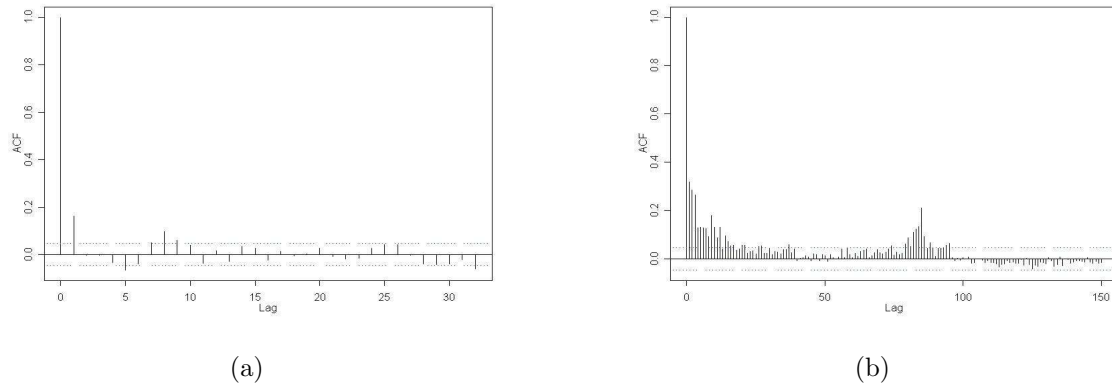


Figura 6.11: Função de Autocorrelação Amostral da Série Temporal: (a)  $\{r_{p,t}\}_{t=1}^{1.728}$  dos Log-Retornos do *Portfolio*; (b)  $\{r_{p,t}^2\}_{t=1}^{1.728}$  dos Valores Quadráticos dos Log-retornos do *Portfolio*.

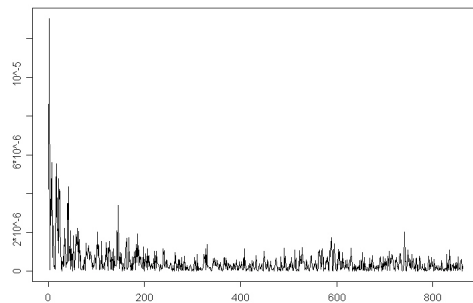


Figura 6.12: Função Periodograma da Série Temporal  $\{r_{p,t}^2\}_{t=1}^{1.728}$ .

Os resultados da estimação das medidas de risco VaR e ES, utilizando a série temporal dos log-retornos do *portfolio*, são apresentados na Tabela 6.8. Lembramos que, para o cálculo das medidas de risco, utilizamos a série temporal  $\{-r_{p,t}\}_{t=1}^{1.728}$ . Sendo assim, para calcular o valor em risco, em reais, do *portfolio* no instante  $n + 1$ , devemos multiplicar os valores apresentados na Tabela 6.8 pelo valor do *portfolio* no instante  $n$ ,  $V_n$ . Lembramos que os valores reais observados para os log-retornos  $r_{i,17.29}$ , para cada  $i = 1, \dots, 4$ , das séries temporais analisadas são, respectivamente, -0,0026, 0,0301, 0,0680 e 0,0021. Portanto, o log-retorno do *portfolio*, em  $t = n + 1$  é igual a 0,0259 e, conseqüentemente, a perda do *portfolio* é igual a  $L_{n+1} = -V_n \times r_{p,n+1} = -V_n \times 0,0259$ .

A Figura 6.13 mostra a função densidade de probabilidade da série temporal  $\{-r_{p,t}\}_{t=1}^{1.728}$ , juntamente com o valor observado  $-r_{p,n+1}$  e os estimados  $\text{VaR}_{p,n+1}$ , para  $p = 99\%$ . Ob-



servamos que o valor estimado, utilizando-se os processos FIEGARCH, foi o que mais se aproximou do valor observado.

Os resultados na Tabela 6.9 são referentes à estimativa das medidas de risco considerando a função de distribuição multivariada do vetor de mudanças de fatores de risco  $\mathbf{X}_t = (r_{1,t}, r_{2,t}, r_{3,t}, r_{4,t})'$ . Neste caso, apresentamos o cálculo do VaR considerando a função de distribuição multivariada, não-condicional (abordagem VaR Normal, expressão (4.43)) e a função de distribuição multivariada condicional (abordagem VaR EWMA, expressão (4.47)), do vetor aleatório  $\mathbf{X}_t$ . Comparando os resultados apresentados nas Tabelas 6.8 e 6.9, observamos que os valores obtidos, para o VaR Normal, em ambos os casos foram os mesmos. Entretanto, para a abordagem VaR EWMA os valores encontrados considerando a função de distribuição univariada dos log-retornos do *portfolio* foram inferiores aos obtidos utilizando-se a função de distribuição multivariada dos fatores de mudança de risco (log-retornos dos ativos).

Tabela 6.8: Valores Estimados para as Medidas de Risco VaR e ES para os Log-retornos do *Portfolio*, ao Nível de Confiança  $p = 95\%$  e  $99\%$  para o Horizonte  $h = 1$  Dia. Caso Univariado.

Abordagem	p	VaR <sub>p,n+1</sub>	ES <sub>p,n+1</sub>
Empírico	0,95	0,0369	0,0588
	0,99	0,0703	0,0966
Normal	0,95	0,0398	0,0712
	0,99	0,0566	0,0923
EWMA	0,95	0,0321	0,0583
	0,99	0,0461	0,0759
EGARCH	0,95	0,0353	0,0625
	0,99	0,0499	0,0808
FIEGARCH	0,95	0,0247	0,0437
	0,99	0,0349	0,0564

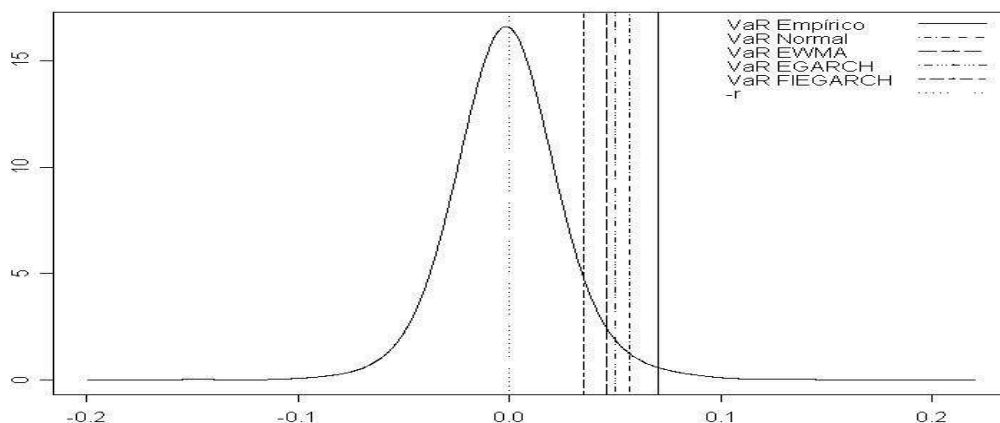


Figura 6.13: Função Densidade de Probabilidade da Série Temporal  $\{-r_{p,t}\}_{t=1}^{1.728}$  com Valor Observado de  $-r_{p,n+1}$  e Valores Estimados para o  $\text{VaR}_{p,p,n+1}$ , ao Nível de Confiança  $p = 99\%$ .

Tabela 6.9: Valores Estimados para as Medidas de Risco VaR e ES para a Série Temporal dos Log-retornos do *Portfolio* ao Nível de Confiança  $p = 95\%$  e  $99\%$  para o Horizonte  $h = 1$ . Caso Multivariado.

Abordagem	p	VaR <sub>p,n+1</sub>	ES <sub>p,n+1</sub>
Normal	0,95	0,0398	0,0712
	0,99	0,0566	0,0923
EWMA	0,95	0,2131	0,3788
	0,99	0,3018	0,4898

Os resultados que seguem referem-se ao cálculo das medidas de risco para os ativos individuais do *portfolio*. Para o cálculo das medidas de risco, sob a abordagem econométrica, os modelo ARMA( $p_1, q_1$ )-EGARCH( $p_2, q_2$ ) selecionados, após a análise de resíduos, foram os que apresentaram o menor número de parâmetros. Estes modelos são apresentados na Seção F.2 do Apêndice F. O processo de seleção dos modelos ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ) para as séries temporais dos log-retornos dos ativos do *portfolio* é apresentado na Seção 6.2 deste capítulo.

Tabela 6.10: Valores Estimados do VaR para as Séries Temporais dos Log-retornos dos Ativos do *Portfolio*, ao Nível de Confiança  $p = 95\%$  e  $99\%$  para o Horizonte  $h = 1$ .

Abordagem	p	VaR <sub>p,1,n+1</sub>	VaR <sub>p,2,n+1</sub>	VaR <sub>p,3,n+1</sub>	VaR <sub>p,4,n+1</sub>	$\sum_{i=1}^4 a_i \text{VaR}_{p,i,n+1}$
Empírico	0,95	0,0427	0,0508	0,0494	0,0483	0,0472
	0,99	0,0785	0,1023	0,0870	0,0905	0,0875
Normal	0,95	0,0487	0,0582	0,0535	0,0539	0,0528
	0,99	0,0693	0,0825	0,0761	0,0767	0,0750
EWMA	0,95	0,2648	0,3056	0,3024	0,2511	0,2814
	0,99	0,3751	0,4327	0,4281	0,3557	0,3985
EGARCH	0,95	0,0370	0,0633	0,0565	0,0342	0,0473
	0,99	0,0536	0,0875	0,0783	0,0489	0,0666
FIEGARCH	0,95	0,0385	0,0495	0,0380	0,0308	0,0390
	0,99	0,0531	0,0712	0,0555	0,0428	0,0554

Comparando os valores apresentados nas Tabelas 6.10 e 6.11 com aqueles apresentados na Tabela 6.8, observamos que, para este *portfolio*, ambas as desigualdades

$$\text{VaR}_{p,\mathcal{P},n+1} \leq \sum_{i=1}^4 a_i \text{VaR}_{p,i,n+1} \quad \text{e} \quad \text{ES}_{p,\mathcal{P},n+1} \leq \sum_{i=1}^4 a_i \text{ES}_{p,i,n+1},$$

são válidas. Observamos que o VaR calculado pela abordagem econométrica, utilizando os processos FIEGARCH foi o que mais se aproximou dos valores reais observados  $-r_{i,1.729}$ , para cada  $i = 1, \dots, 4$ , que são dados, respectivamente, por 0,0026, -0,0301, -0,0680 e -0,0021 (segue que  $-r_{\mathcal{P},1.729} = -0,0259$ ). Note que as perdas foram superestimadas (se estivéssemos analisando a cauda direita da função de distribuição, as perdas seriam subestimadas). Esse fato, amplamente discutido na literatura, já era esperado. Isso ocorre porque utilizamos a função de distribuição normal para estimar as medidas de risco.

Tabela 6.11: Valores Estimados do ES para as Séries Temporais dos Log-retornos dos Ativos do *Portfolio*, ao Nível de Confiança  $p = 95\%$  e  $99\%$  para o Horizonte  $h = 1$ .

Abordagem	p	$ES_{p,1,n+1}$	$ES_{p,2,n+1}$	$ES_{p,3,n+1}$	$ES_{p,4,n+1}$	$\sum_{i=1}^4 a_i ES_{p,i,n+1}$
Empírico	0,95	0,0672	0,0731	0,0623	0,0679	0,0669
	0,99	0,1133	0,1163	0,1049	0,1202	0,1124
Normal	0,95	0,0871	0,1037	0,0956	0,0964	0,0943
	0,99	0,1128	0,1341	0,1238	0,1249	0,1221
EWMA	0,95	0,4707	0,5428	0,5370	0,4464	0,5001
	0,99	0,6086	0,7017	0,6942	0,5771	0,6465
EGARCH	0,95	0,0680	0,1085	0,0973	0,0616	0,0833
	0,99	0,0887	0,1388	0,1246	0,0799	0,1074
FIEGARCH	0,95	0,0658	0,0901	0,0706	0,0532	0,0695
	0,99	0,0841	0,1172	0,0924	0,0682	0,0900

## 6.5 Análise de Cenários (Teste de Estresse)

Para realizar o teste de estresse, utilizamos duas abordagens diferentes, sendo que, a primeira delas consiste na escolha dos cenários baseado nos dados históricos, este é o denominado *teste de estresse tradicional*, a segunda abordagem consiste na escolha dos cenários sob um certo domínio de admissibilidade, obtendo então a medida denominada *perda máxima*, ou *MaxLoss*.

Para realizar o teste de estresse assumimos que os fatores de risco são os logaritmos dos preços dos ativos do *portfolio*. Para o cálculo dos valores  $\delta_i$ , para  $i = 1, \dots, 4$ , definidos pela expressão (4.67), (estes valores indicam o quanto o valor do *portfolio* é afetado pela variação de cada fator de risco) consideramos o valor atual dos preços dos ativos, isto é,  $P_{i,n}$ , para cada  $i = 1, \dots, 4$  e  $\Delta_i = 0,05$ , para cada  $i = 1, \dots, 4$  e  $n = 1.729$ . Obtivemos  $\delta_i = -a_i$ , para cada  $i = 1, \dots, 4$ , onde  $a_i$  é o peso do ativo  $A_i$ . Isso indica que o fator de risco que mais influência no preço do *portfolio* é o logaritmo do preço do ativo que possui a maior participação porcentual, em relação ao total do *portfolio*.

### 6.5.1 Construção dos Cenários Utilizando Dados Históricos

Para a construção de cenários baseados em dados históricos determinamos a variação máxima de cada fator de risco e, em seguida, combinamos tudo em um único cenário. Os procedimentos utilizados nesta subseção para determinar a variação máxima dos fatores de risco são descritos na Subseção 4.6.3. A notação utilizada nas tabelas que seguem é aquela definida na Subseção 4.6.3.

Lembramos que, como os fatores de risco são o logaritmo dos preços, a variação dos fatores de risco, para um período  $h$ , são os log-retornos de período  $h$  (veja Definição 4.5). Na Tabela 6.12 apresentamos as variações máximas e mínimas dos fatores de risco (ou seja, os log-retornos máximos e mínimos) para cada um dos ativos e para o mercado financeiro, durante o período histórico, considerando janelas de 1 e de 10 dias. Os cenários

considerados são apresentados na Tabela 6.13. Por definição,  $X_{i,S_j}$  representa o valor do log-retorno do ativo  $A_i$  sob o cenário  $S_j$ , para cada  $i = 1, \dots, 4$  e  $j = 1, \dots, 12$ . Os cenários são definidos por

- Cenários S1, S2 e S3: são os cenários otimistas, formados pelos log-retornos máximos dos ativos, obtidos pelo método StE e DD, com janelas de 1 e 10 dias.
- Cenários S4, S5 e S6: são os cenários pessimistas, formados pelos retornos mínimos dos ativos, obtidos pelo método StE e DD, com janelas de 1 e 10 dias.
- Cenários S7, S8 e S9: são os cenários otimistas, construídos considerando-se os log-retornos máximos (obtido das observações históricas) do mercado e os coeficientes  $\beta_i$ , para cada  $i = 1, \dots, 4$ , fornecidos pelo modelo CAPM.
- Cenários S10, S11 e S12: são os cenários pessimistas, construídos considerando-se os log-retornos mínimos do mercado (obtido das observações históricas) e os coeficientes  $\beta_i$ , para cada  $i = 1, \dots, 4$ , fornecidos pelo modelo CAPM.

Tabela 6.12: Variações Máximas e Mínimas dos Log-retornos dos Ativos do *Portfolio* e dos Log-retornos do Mercado.

Ativo	1 dia (StE)		10 dias (StE)		10 dias (DD)	
	Máx	Mín	Máx	Mín	Máx	Mín
$A_M$	0,2883	-0,1721	0,2360	-0,2853	0,3848	-0,3713
$A_1$	0,2933	-0,2162	0,3010	-0,4761	0,3669	-0,4761
$A_2$	0,1676	-0,3530	0,3223	-0,5626	0,4037	-0,5626
$A_3$	0,2094	-0,1508	0,3390	-0,3111	0,4558	-0,3189
$A_4$	0,2068	-0,2103	0,3495	-0,2970	0,4242	-0,4937

Tabela 6.13: Cenários para o Teste de Estresse.

	$r_{1,S_j}$	$r_{2,S_j}$	$r_{3,S_j}$	$r_{4,S_j}$
S1	0,2933	0,1676	0,2094	0,2068
S2	0,3010	0,3223	0,3390	0,3495
S3	0,3669	0,4037	0,4558	0,4242
S4	-0,2162	-0,3530	-0,1508	-0,2103
S5	-0,4761	-0,5626	-0,3111	-0,2970
S6	-0,3713	-0,5626	-0,3189	-0,4937
S7	0,2746	0,3030	0,3249	0,2927
S8	0,2248	0,2480	0,2659	0,2396
S9	0,3665	0,4043	0,4335	0,3907
S10	-0,1639	-0,1808	-0,1939	-0,1747
S11	-0,2717	-0,2998	-0,3214	-0,2896
S12	-0,3537	-0,3902	-0,4184	-0,3770

Os log-retornos do *portfolio* sob cada cenário são apresentados na Tabela 6.14. Os log-retornos reais observados do *portfolio* para 1 e 10 dias foram, respectivamente, 0,0259 e 0,0548. Observamos que, em ambos os casos, os log-retornos reais foram superiores aos previstos pelos cenários pessimistas. Entretanto, todos os cenários otimistas superestimaram os log-retornos reais.

Tabela 6.14: Valores dos Log-retornos do *Portfolio* sob os Diferentes Cenários.

Cenário	$r_{\mathcal{P},S_j}$	Cenário	$r_{\mathcal{P},S_j}$	Cenário	$r_{\mathcal{P},S_j}$	Cenário	$r_{\mathcal{P},S_j}$
S1	0,2297	S4	-0,2198	S7	0,2984	S10	-0,1781
S2	0,3249	S5	-0,4100	S8	0,2442	S11	-0,2952
S3	0,4108	S6	-0,4463	S9	0,3982	S12	-0,3843

### 6.5.2 Perda Máxima sob Domínios de Admissibilidade.

Nesta subseção apresentamos a análise de cenários escolhidos sob um domínio de admissibilidade. A perda calculada sob esses cenários é denominada perda máxima, ou *MaxLoss*. Pela expressão (5.5), segue que o *portfolio* é linear, no sentido discutido na Subseção 4.6.4. Portanto, segue do Teorema 4.1, que a um nível de confiança  $p\%$ , a perda máxima do *portfolio* é dada pela expressão (4.68) e o cenário onde isso ocorre é dado pela expressão (4.69).

Na Tabela 6.15 apresentamos a perda máxima do *portfolio* calculada para diferentes valores de  $p$  e os respectivos cenários para os quais essa perda ocorre. Por definição  $r_{i,ML}$ , para cada  $i = 1, \dots, 4$  representa o log-retorno do ativo  $A_i$  sob o cenário de perda máxima. Observando as Tabelas 6.14 e 6.15, percebemos que as perdas que ocorrem sob os cenários pessimistas construídos sob dados históricos são maiores que aquelas que ocorrem quando consideramos cenários sob um certo domínio de admissibilidade. Para todos os valores de  $p$ , a perda máxima estimada foi maior que a perda real observada no instante  $n + 1$ , que foi igual a  $-r_{\mathcal{P}} = -0,0259$ .

Tabela 6.15: Perda Máxima do *Portfolio* para Diferentes Valores de  $p$  e seu Respetivo Cenário.

p	<i>MaxLoss</i>	Cenário			
		$r_{1,ML}$	$r_{2,ML}$	$r_{3,ML}$	$r_{4,ML}$
0,50	-0,0453	-0,0451	-0,0460	-0,0453	-0,0449
0,55	-0,0475	-0,0473	-0,0482	-0,0475	-0,0471
0,65	-0,0521	-0,0519	-0,0529	-0,0521	-0,0517
0,75	-0,0574	-0,0572	-0,0583	-0,0574	-0,0569
0,85	-0,0642	-0,0640	-0,0652	-0,0642	-0,0637
0,95	-0,0762	-0,0759	-0,0773	-0,0762	-0,0756
0,99	-0,0901	-0,0897	-0,0915	-0,0901	-0,0894

# Capítulo 7

## Conclusões

No Capítulo 1 desta dissertação tratamos dos principais conceitos necessários à compreensão e ao estudo dos modelos para séries temporais. Além de apresentar as idéias básicas sobre processos estocásticos, apresentamos as principais definições relacionadas a eles, tais como estacionariedade, ergodicidade, funções de autocovariância, autocorrelação e densidade espectral. Também apresentamos as definições das medidas de assimetria e curtose. Definimos uma série temporal como sendo uma realização, ou parte de uma realização, de um processo estocástico e apresentamos as versões amostrais das funções citadas anteriormente. Também apresentamos algumas definições envolvendo os processos lineares e relembramos alguns dos principais teoremas relacionados a eles. Além disso, retomamos alguns resultados básicos sobre séries de potências que foram utilizados na demonstração de resultados nos capítulos seguintes.

No Capítulo 2 apresentamos os principais processos não-lineares utilizados na literatura para modelar séries temporais que apresentam heteroscedasticidade. Apresentamos condições necessárias e suficientes para a estacionariedade e ergodicidade destes processos. Mostramos que os processos ARCH possuem uma representação AR para o valor quadrático do processo, que os processos GARCH possuem uma representação ARMA para tais valores e que os processos FIEGARCH possuem uma representação ARFIMA para o logaritmo do quadrado da volatilidade. Também mostramos que, se  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo FIEGARCH( $p, d, q$ ) e se  $\mathbb{E}([\ln(Z_t^2)]^2) < \infty$  e  $\mathbb{E}(Z_t \ln(Z_t^2)) < \infty$ , para todo  $t \in \mathbb{Z}$ , então o processo estocástico  $\{\ln(X_t^2)\}_{t \in \mathbb{Z}}$  é um processo ARFIMA( $q, d, 0$ ), com inovações que, em geral, são não-Gaussianas.

No Capítulo 3 relembramos os principais resultados relacionados à análise da função de autocorrelação amostral, o teste de normalidade Jarque-Bera e os principais critérios de seleção de modelos, presentes na literatura, que utilizam o estimador de máxima verossimilhança para a variância mais uma função penalizadora. Também apresentamos alguns dos testes de adequação do modelo, que baseiam-se na função de autocorrelação amostral dos resíduos e as principais estatísticas de teste para detectar longa dependência, encontradas na literatura. Tratamos da estimação dos parâmetros dos modelos, apresentamos os métodos de estimação de pseudo-máxima verossimilhança e ressaltamos que a quase-máxima verossimilhança é um caso particular desse método. E além disso, apresentamos resultados relacionados à previsão de valores futuros para o processo estocástico a partir de um modelo ajustado. Estabelecemos uma relação entre a estimação para modelos lineares e não-lineares.

No Capítulo 4 tratamos da análise dos riscos para séries temporais financeiras. Apre-

sentamos o método de média-variância para seleção de *portfolios* eficientes e o modelo de precificação de ativos (CAPM), que utilizamos como ferramenta para análise da correlação dos ativos com o mercado financeiro. Definimos distribuição de lucros e perdas, fatores de risco e funções de risco. Também apresentamos a definição de medida de risco coerente e as principais medidas de risco utilizadas na literatura.

No Capítulo 5 apresentamos os resultados que obtivemos a partir da análise de séries temporais simuladas. Tais séries temporais foram geradas a partir de processos FIEGARCH, com parâmetros semelhantes aos dos modelos ajustados às séries temporais reais analisadas no Capítulo 6. Utilizamos, para estimação dos parâmetros dos modelos, a função *fgarch* do programa S-Plus. Comparando os valores utilizados na simulação com aqueles ajustados às séries temporais percebemos que, com exceção do modelo M2, para os demais modelos o valor estimado para o parâmetro  $d$  está muito próximo do valor simulado. Essa proximidade foi confirmada pelo valor do erro quadrático médio. Quanto aos demais coeficientes, a precisão dos seus ajustes varia de modelo para modelo, mas em geral, o erro quadrático médio é menor do que 0,0773.

Apresentamos também a estimação do parâmetro  $d$  utilizando as séries temporais  $\{X_{i,t}^2\}_{t=1}^n$  e  $\{\ln(X_{i,t}^2)\}_{t=1}^n$ . Para os métodos R/S, R/S <sub>$m$</sub> , KPSS e V/S, utilizamos  $n = 40$  e  $n = 100$  como número mínimo de observações para cada bloco e para o estimador GPH utilizamos  $g(n) = n^\alpha$  (onde  $n$  é o tamanho amostral) regressores na expressão (3.14), com  $\alpha \in \{0,5; 0,6; 0,7; 0,8; 0,9\}$ . Observamos que os valores encontrados pelos diferentes estimadores para as séries temporais  $\{X_{i,t}^2\}_{t=1}^n$  e  $\{\ln(X_{i,t}^2)\}_{t=1}^n$ , foram todos muito próximos. Entretanto, os valores estimados foram distantes dos valores simulados, em todos os casos. Apesar dos altos valores do vício, o Método GPH, com  $\alpha = 0,8$  foi o que, na maioria dos casos, apresentou os melhores resultados (em relação ao erro quadrático médio). Além disso, o vício apresentado por esse estimador para os processos FIEGARCH foi, em geral, menor que aquele encontrado por Hurvich *et al.* (2005), para os processos LMSV. O erro quadrático médio apresentado pelos estimadores, foi menor quando o tamanho amostral variou de 2.000 para 5.000, como era esperado.

Em relação à previsão dos valores futuros do processo FIEGARCH e da sua volatilidade, observamos que, para o tamanho amostral  $n = 5000$ , o erro quadrático médio foi maior do que o apresentado para  $n = 2000$ . Para o cálculo do VaR utilizamos os valores previstos pelos modelos ajustados às séries temporais. Observamos que os valores obtidos através dos diferentes métodos utilizados foram muito próximos, de forma que não podemos afirmar que um dos métodos é superior aos demais. Ressaltamos, entretanto, que os resultados obtidos utilizando-se os modelos FIEGARCH são fortemente influenciados pela qualidade do ajuste dos modelos. Como as propriedades assintóticas do estimador de quase-máxima verossimilhança para tais processos não são conhecidas, não podemos afirmar que os resultados obtidos no ajuste dos modelos foram satisfatórios.

A análise do *portfolio* formado pelas séries temporais simuladas nos forneceu uma idéia dos resultados que encontraríamos para as séries reais analisadas. Para verificar se as séries temporais apresentam característica de longa dependência e selecionar o modelo mais adequado para cada uma delas, analisamos a função de autocorrelação amostral de ambas as séries temporais, a dos valores quadráticos e a do logaritmo dos valores quadráticos de  $X_t$ , para cada  $t = 1, \dots, n$ , onde  $n$  é o tamanho amostral. Observamos que ambas as séries temporais fornecem resultados semelhantes. Em relação ao ajuste dos modelos FIEGARCH, primeiramente tentamos ajustar modelos com o mesmo grau de  $p$  e  $q$  dos modelos simulados e, quando necessário, reajustamos um novo modelo. Observamos

que, em alguns casos, o modelo mais adequado para a série temporal difere em relação aos valores  $p$  e  $q$  que foram simulados.

Em relação ao modelo CAPM observamos que, para os ativos  $A_1$  e  $A_4$  os retornos observados foram menores que os retornos estimados pelo modelo, enquanto que, para os outros dois ativos os resultados obtidos através do modelo CAPM foram inferiores ao observado para um período  $h = 1$  dia. Para um período  $h = 10$  dias observamos que o retorno previsto pelo modelo CAPM só foi superior ao retorno observado no caso do ativo  $A_1$ .

Para o cálculo das medidas de risco VaR e ES consideramos a função de distribuição dos log-retornos do *portfolio* (caso univariado) e a função de distribuição multivariada dos fatores de risco (log-retornos dos ativos). Consideramos tanto a função de distribuição condicional como a não-condicional. Observamos que, para este *portfolio*, as desigualdades

$$\text{VaR}_{p,\mathcal{P}} \leq \sum_{i=1}^4 a_i \text{VaR}_{p,i} \quad \text{e} \quad \text{ES}_{p,\mathcal{P}} \leq \sum_{i=1}^4 a_i \text{ES}_{p,i} \quad (7.1)$$

são satisfeitas, tanto ao nível de 95% quanto de 99% de confiança. Não foi possível definir qual dos métodos de estimação é o mais adequado, mas observamos que todos eles apresentam resultados semelhantes entre si. Ressaltamos que as perdas estimadas são sempre maiores do que as perdas reais, o que fornece segurança aos agentes ou investidores financeiros.

Para o teste de estresse consideramos cenários baseados nos dados históricos, calculamos os retornos máximos e mínimos para janelas de duração  $h = 1$  e 10 dias e também utilizamos o coeficiente  $\beta$  do modelo CAPM. Observamos que os retornos reais foram superiores aos previstos pelos cenários pessimistas. Os cenários otimistas calculados a partir dos retornos máximos dos ativos superestimaram os retornos, enquanto que, os cenários otimistas que consideram o valor do  $\beta_i$ , para cada  $i = 1, \dots, 4$ , fornecidos pelo modelo CAPM, forneceram estimativas inferiores aos retornos observados. Sendo assim, os cenários pessimistas e os que consideram o coeficiente  $\beta$  podem ser utilizados como uma cota inferior para os log-retornos. Ou, em termos de perdas, podem ser vistos como uma cota superior, dando uma idéia da pior perda que poderia ocorrer no *portfolio* no período de investimento.

Calculamos a medida de risco perda máxima e percebemos que as perdas que ocorrem sob os cenários pessimistas construídos sob dados históricos são maiores que aquelas que ocorrem quando consideramos cenários sob um certo domínio de admissibilidade. Isto é, enquanto que a perda máxima nos fornece uma estimativa mais próxima da realidade, o método tradicional nos fornece uma estimativa para os eventos que possuem menor probabilidade de ocorrência.

No Capítulo 6 tratamos da estimação e análise de medidas de risco para um *portfolio* formado por ações de quatro empresas brasileiras, sendo elas: Bradesco, Brasil Telecom, Gerdau e Petrobrás. Tais ações são negociadas na Bolsa de Valores de São Paulo (Bovespa). Os valores do *portfolio* do mercado foram representados pelo índice Bovespa. Apresentamos as mesmas análises realizadas para as séries temporais simuladas. Observamos que tanto o índice do mercado como os preços dos ativos sofreram uma forte queda em seus valores no período em torno de  $t = 1.000$ , aproximadamente 15 de janeiro de 1999. Este período foi marcado pela desvalorização do real.



Observamos que, para todas as séries temporais, a distribuição dos dados é, aparentemente, leptocúrtica (possui caudas mais pesadas do que a distribuição normal). Pelo gráfico da função de autocorrelação amostral  $\hat{\rho}_X(h)$  percebemos que, tanto para os log-retornos do IBovespa, como para os log-retornos dos ativos, a autocorrelação das variáveis aleatórias apresentam valores significativos para alguns valores de  $h > 1$ . Esse fato sugeriu a necessidade do ajuste de um modelo do tipo ARMA. Observamos que a função de autocorrelação amostral dos valores quadráticos dessas séries temporais também apresenta valores significativos para grandes valores de  $h$ , o que indicou existência tanto de heteroscedasticidade quanto da característica de longa dependência. A presença de longa dependência foi confirmada pela função periodograma dessas séries temporais. Apresentamos também a estimação do parâmetro  $d$  e observamos que os valores obtidos utilizando-se o valor quadrático dos log-retornos e o logaritmo do valor quadrático dos log-retornos foram muito próximos. Em ambos os casos, para o ativo  $A_3$  (Gerdau), o valor do parâmetro  $d$  estimado pelos métodos KPSS e V/S, com  $n = 40$ , foi negativo. Os demais valores estimados pertencem ao intervalo  $(0; 0,5)$ , o que indica a presença de longa dependência.

Ajustamos modelos ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ) às séries temporais dos log-retornos dos preços dos ativos e aos log-retornos dos índices do mercado e utilizamos os valores previstos por estes modelos para o cálculo das medidas de risco VaR e ES.

Para o ativo  $A_1$ , o coeficiente  $\beta$  do modelo CAPM foi tal que  $\beta_1 < 1$ , o que significa que o ativo  $A_1$  (Bradesco) é defensivo. No entanto, enquanto que  $\beta_i > 1$ , para  $i = 2, 3, 4$ , o que significa que os ativos  $A_2$  (Brasil Telecom),  $A_3$  (Gerdau) e  $A_4$  (Petrobrás), são agressivos, isto é, se houverem quedas no índice do mercado, os retornos dos ativos sofrerão quedas maiores (em termos de porcentagem) do que a queda do índice do mercado. De forma análoga, se o índice do mercado sofrer um determinado aumento, o aumento sofrido pelos retornos desses ativos será maior. Os valores reais observados para os log-retornos  $r_{i,1729}$ , para cada  $i = 1, \dots, 4$ , das séries temporais analisadas foram, respectivamente, -0,0026, 0,0301, 0,0680 e 0,0021. Comparando com os valores obtidos pelo modelo CAPM notamos que, para o ativo  $A_1$  (Bradesco) o log-retorno observado foi menor que o retorno estimado, enquanto que, para os demais ativos os resultados obtidos através do modelo CAPM foram inferiores aos observados. Para um período  $h = 10$  dias os log-retornos de período  $h$ , dados por  $r_{i,t}[h]$ , para  $i = 1, \dots, 4$ , são iguais a -0,0156, -0,0343, 0,2456 e -0,0551, respectivamente. Observamos que, neste caso, o retorno previsto pelo modelo CAPM só foi inferior aos log-retornos observados no caso do ativo  $A_3$ . Com base nos resultados, pode-se dizer que o modelo CAPM não descreve de forma satisfatória o comportamento dos retornos esperados do *portfolio*, já que não apresenta um padrão nas estimativas apresentadas.

Calculamos o valor do *portfolio* assumindo que o instante inicial de investimento era igual ao primeiro dia de observação das séries temporais. Em seguida calculamos as perdas do *portfolio* e observamos que a maior perda (ou seja, o maior ganho, que é igual a aproximadamente R\$ 0,37 por cada real investido) corresponde à mudança no valor do *portfolio*, ocorrida entre 24/10/1997 e 27/10/1997. Nesse período o valor do índice Bovespa passou de 11.545,20 para 9.816,80 pontos, o que representa uma queda de 14,97%. Nessa data também ocorreram grandes quedas nos índices Dow Jones (7,18%) e S&P 500 (6,87%). Este período é marcado pela forte crise na Ásia. O maior ganho no *portfolio* corresponde a variação ocorrida no período de 13/01/2001 à 14/01/2001. O IBovespa apresenta uma alta de 2,08% neste período. Enquanto que, para o mercado

mundial temos que em 14/01/2000 o valor do índice Dow Jones foi de 11.722,98 pontos. Esse valor só foi superado em 03/10/2006, quando o índice atingiu 11.727,34 pontos.

Calculamos as medidas de risco VaR e ES utilizando os log-retornos do *portfolio* e observamos que o VaR estimado utilizando-se os processos FIEGARCH foi o que mais se aproximou do valor real da perda observada. Observamos também que os valores encontrados considerando a distribuição univariada dos log-retornos do *portfolio* foram inferiores aos obtidos utilizando-se a distribuição multivariada dos fatores de mudança de risco (log-retornos dos ativos). Além disso, as desigualdades dadas pela expressão (7.1) são também válidas para as séries temporais reais analisadas. Em relação à análise de cada ativo do *portfolio*, observamos que o VaR calculado pela abordagem econométrica, utilizando os processos FIEGARCH, foi o que mais se aproximou dos valores reais observados.

No teste de estresse, observamos que os log-retornos reais foram superiores aos previstos pelos cenários pessimistas. Entretanto, todos os cenários otimistas superestimaram os log-retornos reais. Assim como para as séries temporais simuladas, neste caso, também acontece que as perdas que ocorrem nos cenários pessimistas construídos a partir dos dados históricos são maiores do que aquelas estimadas quando consideramos cenários com um certo domínio de admissibilidade.

Neste trabalho procuramos apresentar os principais resultados relacionados aos processos não-lineares, principalmente, os processos FIEGARCH, cuja literatura é escassa. Procuramos tratar dos principais métodos de seleção de *portfolios* eficientes e apresentar as principais medidas de risco utilizadas na literatura, envolvendo teoria e aplicações. Procuramos expor os assuntos de maneira formal. Sempre que possível apresentamos as provas dos resultados no texto. Em especial, apresentamos provas completas para os resultados que não aparecem na literatura. Nos defrontamos com vários aspectos da teoria de processos FIEGARCH pouco estudados. Para alguns aspectos tentamos dar atenção especial nesta dissertação, como é o caso dos resultados relacionados à ergodicidade e à estacionariedade, e à previsão da volatilidade e dos valores futuros do processo. Esses resultados são apresentados nos Capítulos 2 e 3.

Ainda em aberto estão as propriedades assintóticas do estimador de quase-máxima verossimilhança. É fato que a análise de uma série temporal depende diretamente da qualidade do ajuste de um modelo para os dados. Como não temos conhecimento sobre as propriedades do estimador de quase-máxima verossimilhança não podemos garantir, de forma teórica, que estamos obtendo bons resultados. Assim, esta é uma linha de pesquisa que, mais do que importante, é necessária.

Relacionado à análise de riscos, temos como proposta, realizar os testes de estresse incluindo, além do logaritmo dos preços, outros fatores de risco. Para isso, é necessário identificar os fatores que afetam o preço das ações de cada ativo. Isso requer uma ampla pesquisa e análise das características das empresas envolvidas.

# Referências Bibliográficas

Ahlfors, L.V. (1979). *Complex Analysis: An Introduction to the Theory of Analytic Functions of One Variable*. New York: Mcgraw-Hill. 3ª ed.

Artzner, P., F. Delbaen, J. Eber, e D. Heath, (1999). “Coherent Measures of Risk”. *Math. Finance*, Vol. **9**, 203 – 228.

Baillie, R., T. Bollerslev e H. Mikkelsen (1996). “Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity”. *Journal of Econometrics*, Vol. **74**, 3 – 30.

Bera, A.K. e C.M. Jarque (1981). “Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals: Monte Carlo evidence”. *Economics Letters*, Vol. **7**(4), 313 – 318.

Beran, J. (1994). *Statistics for Long-Memory Processes*. New York: Chapman & Hall.

Berkes, I. e L. Horváth (2003). “The Rate of Consistency of the Quasi-Maximum Likelihood Estimator”. *Statistics and Probability Letters*, Vol. **61**, 133 – 143.

Berkes, I., L. Horváth, e P. Kokoszka (2003). “GARCH Processes: Structure and Estimation”. *Bernoulli*, Vol. **9**, 201 – 228.

Besag, J. (1975). “Statistical Analysis of Non-lattice Data”. *The Statistician*, Vol. **24**, 179 – 195.

Billingsley, P. (1995). *Probability and Measure*. New York: John Wiley.

Box, G.E.P., G.M. Jenkins e G.C. Reinsel (1994). *Time Series Analysis: Forecasting and Control*. New Jersey: Prentice Hall.

Bollerslev, T. (1986). “Generalized Autoregressive Conditional Heteroskedasticity”. *Journal of Econometrics*, Vol. **31**, 307 – 327.

Bollerslev, T. e H.O. Mikkelsen (1996). “Modeling and Pricing Long Memory in Stock Market Volatility”. *Journal of Econometrics*, Vol. **73**, 151 – 184.

Box, G.E.P e D.A. Pierce (1970). “Distribution of Autocorrelations in Autoregressive Moving Average Models”. *Journal of the American Statistical Association*, Vol. **65**, 1509 – 1526.

Breidt, F., N. Crato e P.J.F. de Lima (1998) “On The Detection and Estimation of Long Memory in Stochastic Volatility”. *Journal of Econometrics*, Vol. **83**, 325 – 348.

Brockwell, P.J. e R.A. Davis (1991). *Time Series: Theory and Methods*. New York: Springer-Verlag.

- Degiannakis, S. e E. Xekalaki (2004). “Autoregressive Conditional Heteroscedasticity (ARCH) Models: A Review”. *Quality Technology and Quantitative Management*, Vol. **1**(2), 271 – 324.
- Durrett, R. (1991). *Probability: Theory and Examples*. Pacific Grove: Wadsworth & Brooks/Cole.
- Engle, R.F. (1982). “Autoregressive Conditional Heteroskedasticity with Estimates of Variance of U.K. Inflation”. *Econometrica*, Vol. **50**, 987 – 1008.
- Engle, R.F. e T. Bollerslev (1986). “Modelling the persistence of conditional variances”. *Econometric Reviews*, Vol. **5**, 1 – 50.
- Fabozzi, F.J., S.M. Focardi e P.N. Kolm (2006). *Financial Modeling of the Equity Market. From CAPM to Cointegration*. New York: John Wiley.
- Frey, R. e A. McNeil (2002). “VaR and expected shortfall in portfolios of dependent credit risk: Conceptual and practical insights”. *Journal of Banking and Finance*, Vol. **27**(6), 1317 – 1334.
- Geweke, J. e S. Porter-Hudak (1983). “The Estimation and Application of Long Memory Time Series Models”. *Journal of Times Series Analysis*, Vol. **4**, 221 – 238.
- Giraitis, L., P. Kokoszka, R. Leipus e G. Teyssière, (2003). “Rescaled Variance and Related Tests for Long Memory in Volatility and Levels”. *Journal of Econometrics*, Vol. **112**, 265 – 294.
- Granger, C.W.J. e R. Joyeux, (1980). “An Introduction to Long Memory Time Series Models and Fractional Differencing”. *Journal of Time Series Analysis*, Vol. **1**, 15 – 29.
- Gradshteyn, I.S. e I.M. Ryzhik, (2000). *Table of Integrals, Series, and Products*. New York: Academic Press.
- Gouriéroux, C. (1997). *ARCH Models and Financial Applications*. New York: Springer-Verlag.
- Hall, P. e Q. Yao (2003). “Inference in ARCH and GARCH Models with Heavy-tailed Errors.” *Econometrica*, Vol. **71**, 285 – 317.
- Hamilton, J.D. (1994). *Time Series Analysis*. New Jersey: Princeton University Press.
- Hosking, J. (1981). “Fractional Differencing”. *Biometrika*, Vol. **68**, 165 – 167.
- Hurst, H.R. (1951). “Long-term Storage in Reservoirs”. *Transactions of the American Society of Civil Engineers*, Vol. **116**, 770 – 799.
- Hurvich, C.M. e C.L. Tsay (1989). “Regression and Time Series Model Selection in Small Samples”. *Biometrika*, Vol. **76**, 297 – 307.
- Hurvich, C.M, E. Moulines e P. Soulier (2005). “Estimating long memory in volatility”. *Econometrica*, Vol. **73**(4), 1283 – 1328.
- Karlin, S. e H.M. Taylor (1975). *A First Course in Stochastic Processes*. New York: Academic Press. 2<sup>a</sup> ed.
- Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, e Y. Shin (1992). “Testing the Null Hy-

- pothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?” *Journal of Econometrics*, Vol. **54**, 159 – 178.
- Lee, S. e B. Hansen (1994). “Asymptotic Properties of the Maximum Likelihood Estimator and Test of the Stability on the GARCH and IGARCH Models.” *Econometric Theory*, Vol. **10**, 29 – 52.
- Lee, D. e P. Schmidt (1996). “On the Power of the KPSS Test of Stationarity Against Fractionally-integrated Alternatives”. *Journal of Econometrics*, Vol. **73**(1), 285 – 302.
- Lintner, J. (1965). “The Valuation of Risk Assets and the Selection of Risk Investments in Stock Portfolios and Capital Budgets”. *Review of Economics and Statistics*, Vol. **47**, 13 – 37.
- Lima, E.L. (2006). *Curso de Análise*. Rio de Janeiro: IMPA/CNPq, Vol. **1**. 12º ed.
- Ljung, G. e G.E.P. Box (1978). “On a Measure of Lack of Fit in Time Series Models”. *Biometrika*, Vol. **65**, 297 – 303.
- Lo, A.W. (1991). “Long Term Memory in Stochastic Market Prices”. *Econometrica*, Vol. **59**, 1279 – 1313.
- Lopes, S.R.C (2008). “Long-range Dependence in Mean and Volatility: Models, Estimation and Forecasting”. Capítulo de Livro Especial Pelos 10 Anos da Escola Brasileira de Probabilidade. M.A. Vares e V. Sidoravicius (eds.). Birkhäuser. No Prelo.
- Lopes, S.R.C e B.V.M. Mendes (2006). “Bandwidth Selection in Classical and Robust Estimation of Long Memory”. *International Journal of Statistics and Systems*, Vol. **1**(2), 177 – 200.
- Lumsdaine, R. (1996). “Asymptotic Properties of the Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models.” *Econometrica*, Vol. **64**, 575 – 596.
- Mandelbrot, B.B. e J.R. Wallis (1969). “Computer Experiments with Fractional Gaussian Noises. Part One, Averages and Variances”. *Water Resources Research*, Vol. **5**(1), 228 – 241.
- Markowitz, H. (1952). “Portfolio Selection”. *Journal of Finance*, Vol. **7**(1), 77 – 91.
- McLeod, A.I. e K.W. Hipel (1978). “Preservation of the Rescaled Adjusted Range. A Reassessment of the Hurst Phenomenon”. *Water Resources Research*, Vol. **14**, 491 – 508.
- McNeil, A.J., R. Frey e P. Embrechts (2005). *Quantitative Risk Management*. New Jersey: Princeton University Press.
- Morettin, P.A. (2006). *Um Curso em Séries Temporais Financeiras*. Minicurso do 17º Sinape. São Paulo: ABE.
- Morettin, P.A. e C.M.C. Tolo (2004). *Análise de Séries Temporais*. São Paulo: Edgard Blücher.
- Nelson, D.B. (1991). “Conditional Heteroskedasticity in Asset Returns: A New Approach”. *Econometrica*, Vol. **59**, 347 – 370.
- Nelson, D.B. e C.Q. Cao (1992). “Inequality Constraints in the Univariate GARCH

- Model”. *Journal of Business and Economic Statistics*, Vol. **10**, 229 – 235.
- Newey, W.K. e K.D. West, (1987). “A Simple Positive Semidefinite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. *Econometrica*, Vol. **55**, 703 – 708.
- Olbermann, B.P. (2002). “Estimação em Classes de Processos Estocásticos com Decaimento Hiperbólico da Função de Autocorrelação”. Tese de Doutorado no Programa de Pós-Graduação em Matemática da UFRGS, Porto Alegre. O endereço URL é: <http://mat.ufrgs.br/~slopes/artigos/teseBarbara.pdf>.
- Palaro, H.P. e L.K. Hotta, (2006). “Using Conditional Copula to Estimate Value at Risk”. *Journal of Data Science*, Vol. **4**, 93 – 115.
- Palma, W. (2007). *Long-Memory Time Series: Theory and Methods*. New Jersey: John Wiley.
- Press, W.H., B.P. Flannery, S.A. Teukolsky e W.T. Vetterling (1989). *Numerical Recipes: The Art of Scientific Computing*. Cambridge: Cambridge University Press.
- Priestley, M.B. (1981). *Spectral Analysis in Time Series*. New York: Academic Press.
- Roberts, S.W. (1959). “Control Charts Based on Geometric Moving Averages”. *Technometrics*, Vol. **1**, 239 – 250.
- Ruiz E. e H. Veiga (2007). “Modelling long-memory volatilities with leverage effect: A-LMSV versus FIEGARCH”. *Computational Statistics and Data Analysis*. Forthcoming.
- Sharpe, W.F. (1964). “Capital Asset Prices: a Theory of Market Equilibrium Under Conditions of Risk”. *Journal of Finance*. Vol. **19**, 425 – 442.
- Stout, W.F. (1974). *Almost Sure Convergence*. New York: Academic Press.
- Straumann, D. e T. Mikosch (2006). “Quasi-Maximum-Likelihood Estimation in Conditionally Heteroscedastic Time Series: A Stochastic Recurrence Equations Approach”. *The Annals of Statistics*, Vol. **34**(5), 2449 – 2495.
- Studer G. (1997). “Maximum Loss for Measurement of Market Risk”. Tese de Doutorado no Programa de Pós-Graduação em Matemática do Swiss Federal Institute of Technology. Zurich.
- Wei, W. (1990). *Time Series Analysis. Univariate and Multivariate Methods*. New York: Addison Wesley.
- Zangari, P. “RiskMetrics Technical Document”. O endereço URL é <http://www.riskmetrics.com/rmcovv.html>.
- Zivot, E. e J. Wang (2005). *Modeling Financial Time Series with S-PLUS*. New York: Springer-Verlag. 2ª ed.

# Apêndice A

No Teorema 2.6 apresentamos o cálculo da ordem de convergência ao infinito dos coeficientes  $\lambda_{d,k}$ , dados na expressão (2.53), para o caso mais simples em que  $p = 0 = q$  e o caso mais geral, quando  $p \geq 0$  e  $q \geq 0$ . Neste apêndice apresentamos o cálculo para os casos intermediários, ou seja, quando  $p > 0$  e  $q = 0$  e quando  $p \geq 0$  e  $q = 1$ . Estes resultados auxiliam no entendimento do caso mais geral.

CASO:  $p > 0$  e  $q = 0$ .

Neste caso, pela expressão (2.45), temos que

$$\ln(\sigma_t^2) = \omega + \frac{\alpha(\mathcal{B})}{(1 - \mathcal{B})^d} g(Z_{t-1}),$$

onde  $\alpha(\mathcal{B}) = \sum_{i=0}^p (-\alpha_i) \mathcal{B}^i$ . Logo, pelas expressões (2.52) e (2.57) segue que

$$\begin{aligned} \lambda(\mathcal{B}) &= (1 - \mathcal{B})^{-d} \alpha(\mathcal{B}) \\ &= \left( \sum_{i=0}^{\infty} \pi_{d,i} \mathcal{B}^i \right) \left( \sum_{i=0}^p (-\alpha_i) \mathcal{B}^i \right) \\ &= \sum_{k=0}^{\infty} \left( \sum_{i=0}^{\min\{p,k\}} (-\pi_{d,k-i} \alpha_i) \right) \mathcal{B}^k. \end{aligned}$$

Assim, para todo  $k \geq 0$ ,

$$\lambda_{d,k} = \sum_{i=0}^{\min\{p,k\}} (-\pi_{d,k-i}) \alpha_i.$$

Note que, para  $k > p$ ,

$$\lambda_{d,k} = \sum_{i=0}^p (-\pi_{d,k-i}) \alpha_i.$$

Além disso, como  $\pi_{d,k} \rightarrow 0$ , quando  $k \rightarrow \infty$ , para  $k$  suficientemente grande,  $\pi_{d,k} \sim \pi_{d,k-i}$ , para todo  $0 \leq i \leq p$ . Segue que, para  $k$  suficientemente grande,

$$\lambda_{d,k} \sim \pi_{d,k} \sum_{i=0}^p (-\alpha_i).$$

Pelo Lema 2.5 segue que

$$\lambda_{d,k} = o(k^d), \quad \text{quando } k \rightarrow \infty.$$

CASO:  $p \geq 0$  e  $q = 1$ .

Neste caso, pela expressão (2.45), temos que

$$\ln(\sigma_t^2) = \omega + \frac{\alpha(\mathcal{B})}{(1 - \beta_1 \mathcal{B})(1 - \mathcal{B})^d} g(Z_{t-1}).$$

Lembramos que, como o polinômio  $\beta(\mathcal{B}) = (1 - \beta_1 \mathcal{B})$  não possui raízes no círculo unitário, segue que  $|\beta_1| < 1$  e assim,

$$(1 - \beta_1 \mathcal{B})^{-1} = \sum_{k=0}^{\infty} \beta_1^k \mathcal{B}^k.$$

Portanto, pelas expressões (2.52) e (2.57) segue que

$$\begin{aligned} \lambda(\mathcal{B}) &= \beta(\mathcal{B})^{-1} (1 - \mathcal{B})^{-d} \alpha(\mathcal{B}) \\ &= \left( \sum_{i=0}^{\infty} \left( \sum_{j=0}^i \pi_{d,i-j} \beta_1^j \right) \mathcal{B}^i \right) \left( \sum_{i=0}^p (-\alpha_i) \mathcal{B}^i \right) \\ &= \sum_{k=0}^{\infty} \left( \sum_{i=0}^{\min\{p,k\}} (-\alpha_i) \left( \sum_{j=0}^{k-i} \pi_{d,k-i-j} \beta_1^j \right) \right) \mathcal{B}^k. \end{aligned}$$

Segue que, para todo  $k \geq 0$ ,

$$\lambda_{d,k} = \sum_{i=0}^{\min\{p,k\}} (-\alpha_i) \left( \sum_{j=0}^{k-i} \pi_{d,k-i-j} \beta_1^j \right)$$

Note que, para  $k > p$  temos,

$$\begin{aligned} \lambda_{d,k} &= \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=0}^{k-i} \pi_{d,k-i-j} \beta_1^j \right) \\ &= \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=0}^{k-i} \pi_{d,j} \beta_1^{k-i-j} \right). \end{aligned}$$

Além disso, como  $\beta_1^k \rightarrow 0$ , quando  $k \rightarrow \infty$ , segue que, dado  $\varepsilon > 0$ , existe  $k_0 > 0$  tal que, para todo  $m > 0$  fixo, e para  $k > k_0$ ,

$$|\pi_{d,j} \beta_1^{k-i-j}| < \frac{\varepsilon}{m},$$

para todo  $0 \leq j \leq m$  e para todo  $0 \leq i \leq p$ . Portanto, para  $k$  suficientemente grande,

$$\begin{aligned} \lambda_{d,k} &= \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=0}^m \pi_{d,j} \beta_1^{k-i-j} + \sum_{j=m+1}^{k-i} \pi_{d,j} \beta_1^{k-i-j} \right) \\ &\sim \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=m+1}^{k-i} \pi_{d,j} \beta_1^{k-i-j} \right). \end{aligned}$$

Note ainda que, como  $\pi_{d,k} \rightarrow 0$ , quando  $k \rightarrow \infty$ ,  $m$  pode ser escolhido de forma que

$$\pi_{d,k} \sim \pi_{d,k-i} \sim \pi_{d,j},$$



para todo  $m + 1 \leq j \leq k - i$  e para todo  $0 \leq i \leq p$ . Logo,

$$\begin{aligned}
\lambda_{d,k} &\sim \sum_{i=0}^p (-\alpha_i) \left( \sum_{j=m+1}^{k-i} \pi_{d,j} \beta_1^{k-i-j} \right) \\
&\sim \sum_{i=0}^p (-\alpha_i) \pi_{d,k} \left( \sum_{j=m+1}^{k-i} \beta_1^{k-i-j} \right) \\
&= \pi_{d,k} \left( \sum_{i=0}^p (-\alpha_i) \right) \left( \sum_{j=0}^{k-i-(m+1)} \beta_1^j \right) \\
&\sim \pi_{d,k} \left( \sum_{i=0}^p (-\alpha_i) \right) \left( \sum_{j=0}^{\infty} \beta_1^j \right).
\end{aligned}$$

Segue que,

$$\lambda_{d,k} \sim \pi_{d,k} \frac{\sum_{i=0}^p (-\alpha_i)}{1 - \beta_1} = \pi_{d,k} \frac{\alpha(1)}{\beta(1)}.$$

Segue, pelo Lema 2.5 que

$$\lambda_{d,k} = o(k^d), \quad \text{quando } k \rightarrow \infty.$$

# Apêndice B

Neste apêndice apresentamos três exemplos onde obtemos o estimador de quase-máxima verossimilhança e analisamos a sua precisão assintótica, a fim de auxiliar na compreensão dos resultados apresentados na Seção 3.2. No primeiro exemplo as variáveis aleatórias são independentes e identicamente distribuídas. No segundo exemplo as variáveis aleatórias são tais que a média e a variância condicional variam com o tempo. Para finalizar, aplicamos os resultados do segundo exemplo para um processo ARCH( $p$ ).

**Exemplo B.1. (Variáveis Aleatórias Independentes e Identicamente Distribuídas).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico constituído por variáveis aleatórias independentes e identicamente distribuídas, com média  $\mu_X$  e variância  $\sigma_X^2$ , desconhecidas. Seja  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir deste processo. Denotamos por  $\mathbf{X} = (X_1, \dots, X_n)'$ . Neste caso, temos que

$$\boldsymbol{\theta} = (\mu_X, \sigma_X^2) \in \mathbb{R} \times \mathbb{R}^+.$$

Sob a hipótese de normalidade, obtemos

$$\ell_t(\boldsymbol{\theta}; X_t) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left\{-\frac{1}{2\sigma_X^2}(X_t - \mu_X)^2\right\}, \quad \text{para todo } t = 1, \dots, n. \quad (\text{B.1})$$

A partir da expressão (B.1), a função de pseudo-verossimilhança é dada por

$$\begin{aligned} \ell(\boldsymbol{\theta}; \mathbf{X}) &= \prod_{t=1}^n \ell_t(\boldsymbol{\theta}; X_t) \\ &= \prod_{t=1}^n \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left\{-\frac{1}{2\sigma_X^2}(X_t - \mu_X)^2\right\}. \end{aligned} \quad (\text{B.2})$$

Segue, da expressão (B.2), que

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}; \mathbf{X}) &= \log(\ell(\boldsymbol{\theta}; \mathbf{X})) \\ &= -\frac{n}{2} \log(\sigma_X^2) - \frac{n}{2} \log(2\pi) - \frac{1}{2\sigma_X^2} \sum_{t=1}^n (X_t - \mu_X)^2. \end{aligned}$$

As derivadas parciais de  $\mathcal{L}(\boldsymbol{\theta}; \mathbf{X})$ , com respeito a  $\mu_X$  e  $\sigma_X^2$ , são dadas por

$$\frac{\partial \ell(\boldsymbol{\theta}; \mathbf{X})}{\partial \mu_X} = \frac{1}{\sigma_X^2} \sum_{t=1}^n (X_t - \mu_X) \quad \text{e} \quad \frac{\partial \ell(\boldsymbol{\theta}; \mathbf{X})}{\partial \sigma_X^2} = -\frac{n}{2\sigma_X^2} + \frac{1}{2\sigma_X^4} \sum_{t=1}^n (X_t - \mu_X)^2. \quad (\text{B.3})$$

Os estimadores de pseudo-máxima verossimilhança de  $\mu_X$  e  $\sigma_X^2$  são as soluções das equações

$$\frac{\partial \ell(\boldsymbol{\theta}; \mathbf{X})}{\partial \mu_X} = 0 \quad \text{e} \quad \frac{\partial \ell(\boldsymbol{\theta}; \mathbf{X})}{\partial \sigma_X^2} = 0$$

e são dados por

$$\hat{\mu}_X = \bar{X} = \frac{1}{n} \sum_{t=1}^n X_t \quad \text{e} \quad \hat{\sigma}_X^2 = \frac{1}{n} \sum_{t=1}^n (X_t - \bar{X})^2. \quad (\text{B.4})$$

Portanto,  $\hat{\boldsymbol{\theta}}_{PL} = (\hat{\mu}_X, \hat{\sigma}_X^2)$ , onde  $\hat{\mu}_X$  e  $\hat{\sigma}_X^2$  são dados na expressão (B.4). Note que os estimadores obtidos são a média e a variância amostrais da série temporal  $\{X_t\}_{t=1}^n$ . A precisão assintótica do estimador pode ser deduzida das expressões das matrizes  $I$  e  $J$  dadas na expressão (3.22). Note que, pela expressão (3.22),

$$I = \mathbb{E}_0 \left( \begin{array}{cc} \left( \frac{\partial \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \mu_X} \right)^2 & \frac{\partial \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \mu_X} \frac{\partial \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \sigma_X^2} \\ \frac{\partial \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \sigma_X^2} \frac{\partial \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \mu_X} & \left( \frac{\partial \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \sigma_X^2} \right)^2 \end{array} \right).$$

Logo, pelas equações dadas na expressão (B.3), temos

$$I = \mathbb{E}_0 \left( \begin{array}{cc} \frac{1}{\sigma_X^2} \left( \frac{X_t - \mu_X}{\sigma_X} \right)^2 & \left( \frac{X_t - \mu_X}{\sigma_X^2} \right) \left( -\frac{1}{2\sigma_X^2} + \frac{(X_t - \mu_X)^2}{2\sigma^4} \right) \\ \left( \frac{X_t - \mu_X}{\sigma_X^2} \right) \left( -\frac{1}{2\sigma_X^2} + \frac{(X_t - \mu_X)^2}{2\sigma^4} \right) & \left( -\frac{1}{2\sigma_X^2} + \frac{(X_t - \mu_X)^2}{2\sigma^4} \right)^2 \end{array} \right),$$

ou seja,

$$I = \left( \begin{array}{cc} \frac{1}{\sigma_X^2} & \frac{1}{2\sigma_X^3} \mathbb{E}_0 \left( \frac{X_t - \mu_X}{\sigma_X} \right)^3 \\ \frac{1}{2\sigma_X^3} \mathbb{E}_0 \left( \frac{X_t - \mu_X}{\sigma_X} \right)^3 & \frac{1}{4\sigma_X^4} \left( \mathbb{E}_0 \left( \frac{X_t - \mu_X}{\sigma_X} \right)^4 - 1 \right) \end{array} \right).$$

Das equações (3.22) e (B.3), segue que

$$J = -\mathbb{E}_0 \left( \begin{array}{cc} \frac{\partial^2 \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \mu_X^2} & \frac{\partial^2 \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \mu_X \partial \sigma_X^2} \\ \frac{\partial^2 \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \sigma_X^2 \partial \mu_X} & \frac{\partial^2 \log(\ell_t(\boldsymbol{\theta}; X_t))}{(\partial \sigma_X^2)^2} \end{array} \right) = \mathbb{E}_0 \left( \begin{array}{cc} \frac{1}{\sigma_X^2} & 0 \\ 0 & \frac{1}{2\sigma_X^4} \end{array} \right).$$

Ou seja,

$$J = \left( \begin{array}{cc} \frac{1}{\sigma_X^2} & 0 \\ 0 & \frac{1}{2\sigma_X^4} \end{array} \right).$$

Assim,

$$\begin{aligned} \text{Var}_{\boldsymbol{\theta}} \left( \sqrt{n}(\hat{\boldsymbol{\theta}}_{PL} - \boldsymbol{\theta}) \right) &= J^{-1} I J^{-1} \\ &= \left( \begin{array}{cc} \sigma_X^2 & \mathbb{E}_0(X_t - \mu_X)^3 \\ \mathbb{E}_0(X_t - \mu_X)^3 & \sigma_X^4 \left( \mathbb{E}_0 \left( \frac{X_t - \mu_X}{\sigma_X} \right)^4 - 1 \right) \end{array} \right). \end{aligned}$$

Note que se a função de distribuição verdadeira das variáveis aleatórias  $X_t$ , para todo  $t \in \mathbb{Z}$ , for a normal, segue que

$$\mathbb{E}_0(X_t - \mu_X)^3 = 0 \quad \text{e} \quad \mathbb{E}_0\left(\frac{X_t - \mu_X}{\sigma_X}\right)^4 = 3 \quad (\text{B.5})$$

e então,

$$J^{-1}IJ^{-1} = \begin{pmatrix} \sigma_X^2 & 0 \\ 0 & 2\sigma_X^4 \end{pmatrix}.$$

Logo,  $\hat{\mu}_X$  e  $\hat{\sigma}_X^2$  são não-correlacionados,  $\text{Var}_\theta(\sqrt{n}(\hat{\mu}_X - \mu_X)) = \sigma_X^2$  e  $\text{Var}_\theta(\sqrt{n}(\hat{\sigma}_X^2 - \sigma_X^2)) = 2\sigma_X^4$ .

Na maioria dos casos, as afirmações da equação (B.5) não são satisfeitas. Ou seja, os estimadores  $\hat{\mu}_X$  e  $\hat{\sigma}_X^2$  são correlacionados quando a função de distribuição verdadeira exibe assimetria. E ainda, a variabilidade de  $\hat{\sigma}_X^2$  pode ser maior ou menor do que  $\frac{2\sigma_X^2}{n}$ , dependendo do valor da curtose. Em particular, quanto mais pesadas as caudas da função distribuição verdadeira, maior a variabilidade.

No exemplo anterior tratamos do caso em que as variáveis aleatórias são independentes e identicamente distribuídas. No Exemplo B.2 a seguir apresentamos o caso em que a média e a variância condicional das variáveis aleatórias variam com o tempo.

**Exemplo B.2. (Processos Estocásticos com Heteroscedasticidade).** Seja  $\{X_t\}_{t \in \mathbb{Z}}$  um processo estocástico tal que

$$\begin{cases} \mathbb{E}(X_t | \mathcal{F}_{t-1}) = \mu_t(\boldsymbol{\theta}), \text{ para todo } t \in \mathbb{Z} \\ \text{Var}(X_t | \mathcal{F}_{t-1}) = h_t(\boldsymbol{\theta}), \text{ para todo } t \in \mathbb{Z}, \end{cases}$$

onde  $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^s$  é o vetor de parâmetros a ser estimado.

Seja  $\{X_{-m}, \dots, X_0, \dots, X_n\}$  parte de uma amostra deste processo. Condicionando  $X_1, \dots, X_n$  à  $\mathcal{F}_0 = \sigma(X_0, \dots, X_{-m})$ , sob hipótese de normalidade de  $X_t | \mathcal{F}_{t-1}$ , para todo  $t \in \mathbb{Z}$ , temos

$$\ell_t(\boldsymbol{\theta}; X_t) = \frac{1}{\sqrt{2\pi h_t(\boldsymbol{\theta})}} \exp\left\{-\frac{1}{2h_t(\boldsymbol{\theta})}(X_t - \mu_t(\boldsymbol{\theta}))^2\right\}.$$

A função de log-pseudo-verossimilhança é dada por

$$\log(\ell(\boldsymbol{\theta}; \mathbf{X})) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n \left[ \log(h_t(\boldsymbol{\theta})) + \frac{(X_t - \mu_t(\boldsymbol{\theta}))^2}{h_t(\boldsymbol{\theta})} \right].$$

### Condições de Primeira Ordem

A derivada de primeira ordem da função  $\log(\ell(\boldsymbol{\theta}; \mathbf{X}))$ , com respeito ao parâmetro  $\boldsymbol{\theta}$ , é dada por

$$\frac{\partial \log(\ell(\boldsymbol{\theta}; \mathbf{X}))}{\partial \boldsymbol{\theta}} = -\frac{1}{2} \sum_{t=1}^n \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \frac{1}{2} \sum_{t=1}^n \frac{(X_t - \mu_t(\boldsymbol{\theta}))^2}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \sum_{t=1}^n \frac{X_t - \mu_t(\boldsymbol{\theta})}{h_t(\boldsymbol{\theta})} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}. \quad (\text{B.6})$$

**Observação B.1.** Suponha que o vetor de parâmetros  $\boldsymbol{\theta}$  possa ser particionado da forma

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} \in \mathbb{R}^s,$$

onde  $\boldsymbol{\alpha}$  aparece apenas na média condicional e  $\boldsymbol{\beta}$  aparece apenas na variância condicional. Neste caso, temos

$$\begin{cases} \frac{\partial \log(\ell(\boldsymbol{\theta}; \mathbf{X}))}{\partial \boldsymbol{\alpha}} = \sum_{t=1}^n \frac{X_t - \mu_t(\boldsymbol{\alpha})}{h_t(\boldsymbol{\beta})} \frac{\partial \mu_t(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}, \\ \frac{\partial \log(\ell(\boldsymbol{\theta}; \mathbf{X}))}{\partial \boldsymbol{\beta}} = -\frac{1}{2} \sum_{t=1}^n \frac{1}{h_t(\boldsymbol{\beta})} \frac{\partial h_t(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} + \frac{1}{2} \sum_{t=1}^n \frac{(X_t - \mu_t(\boldsymbol{\alpha}))^2}{h_t(\boldsymbol{\beta})} \frac{\partial h_t(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}. \end{cases} \quad (\text{B.7})$$

### Matriz de Variância-Covariância Assintótica

Note que, pela expressão (B.6), a derivada de segunda ordem da função  $\log(\ell(\boldsymbol{\theta}; \mathbf{X}))$ , com respeito ao parâmetro  $\boldsymbol{\theta}$ , é dada por

$$\begin{aligned} \frac{\partial^2 \log(\ell(\boldsymbol{\theta}; \mathbf{X}))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} &= \frac{1}{2} \sum_{t=1}^n \frac{1}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} - \frac{1}{2} \sum_{t=1}^n \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \\ &\quad - \sum_{t=1}^n \frac{(X_t - \mu_t(\boldsymbol{\theta}))^2}{h_t^3(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{1}{2} \sum_{t=1}^n \frac{(X_t - \mu_t(\boldsymbol{\theta}))^2}{h_t^2(\boldsymbol{\theta})} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \\ &\quad - \sum_{t=1}^n \frac{(X_t - \mu_t(\boldsymbol{\theta}))}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \sum_{t=1}^n \frac{(X_t - \mu_t(\boldsymbol{\theta}))}{h_t(\boldsymbol{\theta})} \frac{\partial^2 \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \\ &\quad - \sum_{t=1}^n \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} - \sum_{t=1}^n \frac{(X_t - \mu_t(\boldsymbol{\theta}))}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \end{aligned}$$

Para obter a expressão da matrix  $J$  defina

$$u_t = \frac{X_t - \mu_t(\boldsymbol{\theta})}{h_t^{1/2}(\boldsymbol{\theta})}.$$

Note que

$$\mathbb{E}(u_t | \mathcal{F}_{t-1}) = 0 \quad \text{e} \quad \mathbb{E}(u_t^2 | \mathcal{F}_{t-1}) = 1.$$

Logo,

$$\begin{aligned} J &= \mathbb{E}_0 \left[ -\frac{\partial^2 \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] \\ &= \mathbb{E}_0 \left[ -\frac{1}{2} \frac{1}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{1}{2} \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} + \frac{1}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \mathbb{E}(u_t^2 | \mathcal{F}_{t-1}) \right. \\ &\quad - \frac{1}{2} \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial^2 h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \mathbb{E}(u_t^2 | \mathcal{F}_{t-1}) + \frac{1}{h_t^{3/2}(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbb{E}(u_t^2 | \mathcal{F}_{t-1}) \\ &\quad \left. - \frac{1}{h_t^{1/2}(\boldsymbol{\theta})} \frac{\partial^2 \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \mathbb{E}(u_t | \mathcal{F}_{t-1}) + \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{1}{h_t^{3/2}(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \mathbb{E}(u_t | \mathcal{F}_{t-1}) \right]. \end{aligned} \quad (\text{B.8})$$

Segue de (B.8) que

$$J = \mathbb{E}_0 \left[ \frac{1}{2} \frac{1}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right]. \quad (\text{B.9})$$

Para obter a expressão da matriz  $I$  note que

$$\begin{aligned} I &= \mathbb{E}_0 \left[ \frac{\partial \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \boldsymbol{\theta}} \frac{\partial \log(\ell_t(\boldsymbol{\theta}; X_t))}{\partial \boldsymbol{\theta}'} \right] \\ &= \mathbb{E}_0 \left[ \left( -\frac{1}{2} \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \frac{1}{2} \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} u_t^2 + \frac{1}{h_t^{1/2}(\boldsymbol{\theta})} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} u_t \right) \right. \\ &\quad \times \left. \left( -\frac{1}{2} \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{1}{2} \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} u_t^2 + \frac{1}{h_t^{1/2}(\boldsymbol{\theta})} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} u_t \right) \right] \\ &= \mathbb{E}_0 \left[ \frac{1}{4} \frac{1}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} - \frac{1}{2} \frac{1}{h_t^{3/2}(\boldsymbol{\theta})} \left( \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \mathbb{E}(u_t | \mathcal{F}_{t-1}) \right. \\ &\quad + \left( -\frac{1}{2} \frac{1}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right) \mathbb{E}(u_t^2 | \mathcal{F}_{t-1}) \\ &\quad + \frac{1}{2} \frac{1}{h_t^{3/2}(\boldsymbol{\theta})} \left( \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \mathbb{E}(u_t^3 | \mathcal{F}_{t-1}) \\ &\quad \left. + \frac{1}{4} \frac{1}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \mathbb{E}(u_t^4 | \mathcal{F}_{t-1}) \right]. \quad (\text{B.10}) \end{aligned}$$

Segue de (B.10) que

$$\begin{aligned} I &= \mathbb{E}_0 \left[ \frac{1}{4} \frac{1}{h_t^2(\boldsymbol{\theta})} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} (K_t(\boldsymbol{\theta}) - 1) + \frac{1}{h_t(\boldsymbol{\theta})} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right. \\ &\quad \left. + \frac{1}{2} \frac{1}{h_t^{3/2}(\boldsymbol{\theta})} \left( \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{\partial h_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) M_t^3(\boldsymbol{\theta}) \right], \quad (\text{B.11}) \end{aligned}$$

onde

$$K_t(\boldsymbol{\theta}) = \frac{1}{h_t(\boldsymbol{\theta})^2} \mathbb{E} \left[ (X_t - \mu_t(\boldsymbol{\theta}))^4 | \mathcal{F}_{t-1} \right] = \mathbb{E}(u_t^4 | \mathcal{F}_{t-1}),$$

é a curtose condicional da variável aleatória  $X_t$ , para todo  $t \in \mathbb{Z}$ , e

$$M_t^3(\boldsymbol{\theta}) = \mathbb{E} \left[ (X_t - \mu_t(\boldsymbol{\theta}))^3 | \mathcal{F}_{t-1} \right] = h_t^{3/2} \mathbb{E}(u_t^3 | \mathcal{F}_{t-1}),$$

é o momento condicional de terceira ordem da variável aleatória  $X_t$ , para todo  $t \in \mathbb{Z}$ .

As fórmulas dadas nas expressões (B.9) e (B.11) são extensões daquelas encontradas no Exemplo B.1. É fácil ver que as expressões das matrizes  $I$  e  $J$  coincidem quando a distribuição condicional verdadeira é a distribuição normal pois, neste caso,  $K_t(\boldsymbol{\theta}) = 3$  e  $M_t^3(\boldsymbol{\theta}) = 0$ .

**Observação B.2.** Note que, se o vetor de parâmetros  $\boldsymbol{\theta}$  puder ser particionado na forma

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} \in \mathbb{R}^s$$

onde  $\alpha$  aparece apenas na média condicional e  $\beta$  aparece apenas na variância condicional, então as matrizes  $I$  e  $J$  serão dadas por

$$I = \begin{bmatrix} \mathbb{E}_0 \left( \frac{1}{h_t(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial \mu_t(\alpha)}{\partial \alpha'} \right) & \mathbb{E}_0 \left( \frac{1}{2h_t^3(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial h_t(\beta)}{\partial \beta'} M_{3t}(\theta) \right) \\ \mathbb{E}_0 \left( \frac{1}{2h_t^3(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial \mu_t(\alpha)}{\partial \alpha'} M_{3t}(\theta) \right) & \mathbb{E}_0 \left( \frac{1}{4h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} (K_t(\theta) - 1) \right) \end{bmatrix} \quad (\text{B.12})$$

e

$$J = \begin{bmatrix} \mathbb{E}_0 \left( \frac{1}{h_t(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial \mu_t(\alpha)}{\partial \alpha'} \right) & 0 \\ 0 & \mathbb{E}_0 \left( \frac{1}{2h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} \right) \end{bmatrix}. \quad (\text{B.13})$$

Da expressão (B.13) segue que

$$J^{-1} = \begin{bmatrix} \left( \mathbb{E}_0 \left( \frac{1}{h_t(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial \mu_t(\alpha)}{\partial \alpha'} \right) \right)^{-1} & 0 \\ 0 & \left( \mathbb{E}_0 \left( \frac{1}{2h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} \right) \right)^{-1} \end{bmatrix}. \quad (\text{B.14})$$

Das expressões (B.12) e (B.14) segue que

$$J^{-1} I J^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad (\text{B.15})$$

onde

$$A_{11} = \left( \mathbb{E}_0 \left( \frac{1}{h_t(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial \mu_t(\alpha)}{\partial \alpha'} \right) \right)^{-1},$$

$$A_{12} = \left( \mathbb{E}_0 \left( \frac{1}{h_t(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial \mu_t(\alpha)}{\partial \alpha'} \right) \right)^{-1} \times \mathbb{E}_0 \left( \frac{1}{2h_t^3(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial h_t(\beta)}{\partial \beta'} M_{3t}(\theta) \right) \\ \times \left( \mathbb{E}_0 \left( \frac{1}{2h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} \right) \right)^{-1},$$

$$A_{21} = \left( \mathbb{E}_0 \left( \frac{1}{2h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} \right) \right)^{-1} \times \mathbb{E}_0 \left( \frac{1}{2h_t^3(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial h_t(\beta)}{\partial \beta'} M_{3t}(\theta) \right) \\ \times \left( \mathbb{E}_0 \left( \frac{1}{h_t(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial \mu_t(\alpha)}{\partial \alpha'} \right) \right)^{-1}$$

e

$$A_{22} = \left( \mathbb{E}_0 \left( \frac{1}{2h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} \right) \right)^{-1} \times \mathbb{E}_0 \left( \frac{1}{4h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} (K_t(\theta) - 1) \right) \\ \times \left( \mathbb{E}_0 \left( \frac{1}{2h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} \right) \right)^{-1}.$$

Da expressão (B.15) segue que a variância assintótica do estimador  $\hat{\alpha}_{PL}$  é dada por

$$\text{Var}_{\theta} [\sqrt{n}(\hat{\alpha}_{PL} - \alpha)] = \left( \mathbb{E}_0 \left( \frac{1}{h_t(\beta)} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial \mu_t(\alpha)}{\partial \alpha'} \right) \right)^{-1}$$

e a variância assintótica do estimador  $\hat{\beta}_{PL}$  é dada por

$$\begin{aligned} \text{Var}_{\theta} [\sqrt{n}(\hat{\beta}_{PL} - \beta)] &= \left( \mathbb{E}_0 \left( \frac{1}{2h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} \right) \right)^{-1} \\ &\times \mathbb{E}_0 \left( \frac{1}{4h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} (K_t(\theta) - 1) \right) \times \left( \mathbb{E}_0 \left( \frac{1}{2h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} \right) \right)^{-1}. \end{aligned} \quad (\text{B.16})$$

No Exemplo B.3 a seguir aplicamos os resultados obtidos no Exemplo B.2 para um processo ARCH( $p$ ), com  $p = 1$ . Como veremos, a generalização deste exemplo para um  $p$  qualquer não é difícil, apenas envolve um número maior de cálculo.

**Exemplo B.3. (Processos ARCH( $p$ ), com  $p = 1$ ).** Sejam  $\{X_t\}_{t \in \mathbb{Z}}$  um processo ARCH( $p$ ), definido pelas expressões (2.1) e (2.2) e  $\{X_t\}_{t=1}^n$  uma série temporal obtida a partir deste processo. Temos que,

$$\mathbb{E}(X_t | \mathcal{F}_{t-1}) = 0 \quad \text{e} \quad \text{Var}(X_t | \mathcal{F}_{t-1}) = \sigma_t^2, \quad \text{para todo } t \in \mathbb{Z}.$$

Neste caso,  $\theta = (\alpha_0, \dots, \alpha_p)' = \beta$  e  $\alpha = 0$ , pois não existem parâmetros para a média condicional. Supondo que as variáveis aleatórias  $X_t | \mathcal{F}_{t-1}$  possuem distribuição normal temos, pela Observação B.2, que

$$I = \left[ \mathbb{E}_0 \left( \frac{1}{4h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} (K_t(\theta) - 1) \right) \right] \quad (\text{B.17})$$

e

$$J = \left[ \mathbb{E}_0 \left( \frac{1}{2h_t^2(\beta)} \frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} \right) \right], \quad (\text{B.18})$$

onde  $h_t(\beta) = \sigma_t^2$ , para todo  $t \in \mathbb{Z}$ , e  $\beta = (\alpha_0, \dots, \alpha_p)'$ . Além disso,

$$\frac{\partial h_t(\beta)}{\partial \beta} \frac{\partial h_t(\beta)}{\partial \beta'} = \begin{pmatrix} \frac{\partial \sigma_t^2}{\partial \alpha_0} \\ \vdots \\ \frac{\partial \sigma_t^2}{\partial \alpha_p} \end{pmatrix} \begin{pmatrix} \frac{\partial \sigma_t^2}{\partial \alpha_0} & \frac{\partial \sigma_t^2}{\partial \alpha_1} & \dots & \frac{\partial \sigma_t^2}{\partial \alpha_p} \end{pmatrix}. \quad (\text{B.19})$$

$$= \begin{pmatrix} 1 \\ X_{t-1}^2 \\ \vdots \\ X_{t-p}^2 \end{pmatrix} \begin{pmatrix} 1 & X_{t-1}^2 & \dots & X_{t-p}^2 \end{pmatrix} \quad (\text{B.20})$$

$$= \begin{pmatrix} 1 & X_{t-1}^2 & X_{t-2}^2 & X_{t-3}^2 & \dots & X_{t-p}^2 \\ X_{t-1}^2 & (X_{t-1}^2)^2 & X_{t-1}^2 X_{t-2}^2 & X_{t-1}^2 X_{t-3}^2 & \dots & X_{t-1}^2 X_{t-p}^2 \\ X_{t-2}^2 & X_{t-2}^2 X_{t-2}^2 & (X_{t-2}^2)^2 & X_{t-2}^2 X_{t-3}^2 & \dots & X_{t-2}^2 X_{t-p}^2 \\ X_{t-3}^2 & X_{t-3}^2 X_{t-1}^2 & X_{t-3}^2 X_{t-2}^2 & (X_{t-3}^2)^2 & \dots & X_{t-3}^2 X_{t-p}^2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{t-p}^2 & X_{t-p}^2 X_{t-1}^2 & X_{t-p}^2 X_{t-2}^2 & X_{t-p}^2 X_{t-3}^2 & \dots & (X_{t-p}^2)^2 \end{pmatrix}.$$



Note ainda que,

$$K_t(\boldsymbol{\theta}) = \frac{1}{(\sigma_t^2)^2} \mathbb{E}(X_t^4 | \mathcal{F}_{t-1}) = \mathbb{E}(Z_t^4). \quad (\text{B.21})$$

Apresentamos os cálculos das matrizes  $I$  e  $J$  para o caso em que  $p = 1$ . Para os demais casos o raciocínio é análogo.

Se  $p = 1$  então, pela expressão (B.21), segue que

$$\frac{\partial h_t(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \frac{\partial h_t(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} = \begin{pmatrix} 1 & X_{t-1}^2 \\ X_{t-1}^2 & (X_{t-1}^2)^2 \end{pmatrix} = \begin{pmatrix} 1 & X_{t-1}^2 \\ X_{t-1}^2 & X_{t-1}^4 \end{pmatrix}. \quad (\text{B.22})$$

Das expressões (2.2) e (B.22) segue que

$$\begin{aligned} \frac{1}{h_t^2(\boldsymbol{\beta})} \frac{\partial h_t(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \frac{\partial h_t(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} &= \frac{1}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \begin{pmatrix} 1 & X_{t-1}^2 \\ X_{t-1}^2 & X_{t-1}^4 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} & \frac{X_{t-1}^2}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \\ \frac{X_{t-1}^2}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} & \frac{X_{t-1}^4}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \end{pmatrix}. \end{aligned} \quad (\text{B.23})$$

Segue, da expressão (B.17), (B.21) e (B.23), que a matriz  $I$  é dada por

$$I = \frac{\mathbb{E}(Z_t^4) - 1}{4} \begin{bmatrix} \mathbb{E}_0 \left( \frac{1}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \right) & \mathbb{E}_0 \left( \frac{X_{t-1}^2}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \right) \\ \mathbb{E}_0 \left( \frac{X_{t-1}^2}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \right) & \mathbb{E}_0 \left( \frac{X_{t-1}^4}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \right) \end{bmatrix}. \quad (\text{B.24})$$

Pelas expressões (B.18) e (B.23), a matriz  $J$  é dada por

$$J = \frac{1}{2} \begin{bmatrix} \mathbb{E}_0 \left( \frac{1}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \right) & \mathbb{E}_0 \left( \frac{X_{t-1}^2}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \right) \\ \mathbb{E}_0 \left( \frac{X_{t-1}^2}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \right) & \mathbb{E}_0 \left( \frac{X_{t-1}^4}{(\alpha_0 + \alpha_1 X_{t-1}^2)^2} \right) \end{bmatrix}. \quad (\text{B.25})$$

Lembramos que, em geral, não conhecemos a verdadeira função de distribuição das variáveis aleatórias  $X_t$ , para todo  $t \in \mathbb{Z}$ . Sendo assim, de acordo com as expressões (3.24) e (3.23), as matrizes  $I$  e  $J$ , dadas pelas expressões (B.24) e (B.25), são estimadas por

$$\begin{aligned} \hat{I} &= \frac{1}{n} \sum_{t=1}^n \left[ \frac{1}{4h_t^2(\hat{\boldsymbol{\beta}}_{PL})} \frac{\partial h_t(\hat{\boldsymbol{\beta}}_{PL})}{\partial \boldsymbol{\beta}} \frac{\partial h_t(\hat{\boldsymbol{\beta}}_{PL})}{\partial \boldsymbol{\beta}'} (K_t(\hat{\boldsymbol{\beta}}_{PL}) - 1) \right] \\ &= \frac{(\bar{Z}^4 - 1)}{4n} \sum_{t=1}^n \begin{bmatrix} \frac{1}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} & \frac{X_{t-1}^2}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} \\ \frac{X_{t-1}^2}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} & \frac{X_{t-1}^4}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} \end{bmatrix} \end{aligned} \quad (\text{B.26})$$

e

$$\begin{aligned}
\hat{J} &= \frac{1}{n} \sum_{t=1}^n \left[ \frac{1}{2h_t^2(\hat{\beta}_{PL})} \frac{\partial h_t(\hat{\beta}_{PL})}{\partial \beta} \frac{\partial h_t(\hat{\beta}_{PL})}{\partial \beta'} \right] \\
&= \frac{1}{2n} \sum_{t=1}^n \left[ \begin{array}{cc} \frac{1}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} & \frac{X_{t-1}^2}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} \\ \frac{X_{t-1}^2}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} & \frac{X_{t-1}^4}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} \end{array} \right], \tag{B.27}
\end{aligned}$$

onde  $\bar{Z}^4$  é o quarto momento amostral da variável aleatória  $Z_t$ , para todo  $t \in \mathbb{Z}$ , de acordo com a Definição 1.10.

Segue, da expressão (B.27), que

$$\hat{J}^{-1} = 2n \left[ \begin{array}{cc} \sum_{t=1}^n \left( \frac{1}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} \right) & \sum_{t=1}^n \left( \frac{X_{t-1}^2}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} \right) \\ \sum_{t=1}^n \left( \frac{X_{t-1}^2}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} \right) & \sum_{t=1}^n \left( \frac{X_{t-1}^4}{(\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2)^2} \right) \end{array} \right]^{-1}. \tag{B.28}$$

Assim, a variância assintótica de  $\beta_{PL} = (\hat{\alpha}_0, \hat{\alpha}_1)'$  pode ser estimada através da expressão

$$\widehat{\text{Var}}_{\beta} \left( \sqrt{n}(\hat{\beta}_{PL} - \beta) \right) = \hat{J}^{-1} \hat{I} \hat{J}^{-1},$$

com  $\hat{J}^{-1}$  e  $\hat{I}$  dadas pelas expressões (B.28) e (B.26).

# Apêndice C

Neste apêndice apresentamos o valor aproximado da soma

$$S = - \sum_{k=1}^m \delta_{d,k}, \quad (\text{C.1})$$

para diferentes valores de  $m$  e  $d$ , onde os coeficientes  $\delta_{d,k}$ , para todo  $k \in \mathbb{Z}$ , são dados pela expressão (1.19). Este resultado é de grande importância para as simulações apresentadas no Capítulo 5. De fato, pela Proposição 2.4 podemos perceber que a ordem de convergência dos coeficientes da série  $\sum_{k=0}^{\infty} \lambda_{d,k} \mathcal{B}^k$  está diretamente relacionada com a convergência da soma  $1 + S$ , com  $S$  dada pela expressão (C.1).

Na Figura C.1 apresentamos o gráfico da soma  $S$ , para  $1.000 \leq m \leq 15.000$  e  $0,025 \leq d \leq 0,5$ . Os valores para a soma  $-S$ , com  $S$  dada pela expressão (C.1) são apresentados na Tabela C.1.

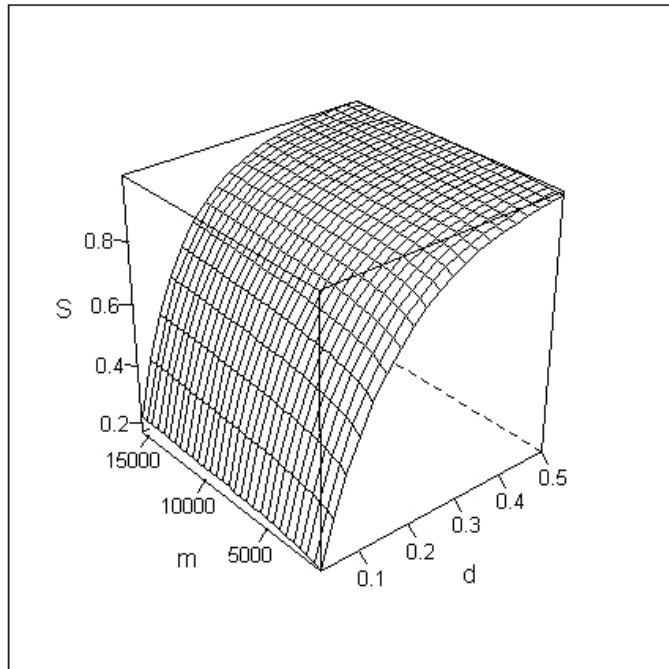


Figura C.1: Valores da expressão (C.1) quando  $1.000 \leq m \leq 15.000$  e  $0,025 \leq d \leq 0,5$ .

Tabela C.1: Valores da soma  $-S$ , com  $S$  dada pela expressão (C.1), para  $1.000 \leq m \leq 15.000$  e  $0,025 \leq d \leq 0,5$ .

$m \setminus d$	0,025	0,05	0,075	0,1	0,125	0,15	0,175	0,2	0,225	0,25	0,275	0,3	0,325	0,35	0,375	0,4	0,425	0,45	0,475	0,5
1000	0,171	0,314	0,432	0,531	0,613	0,681	0,738	0,784	0,823	0,855	0,881	0,903	0,921	0,936	0,948	0,958	0,966	0,972	0,978	0,982
1500	0,179	0,327	0,449	0,550	0,632	0,700	0,755	0,801	0,838	0,869	0,894	0,914	0,931	0,944	0,955	0,964	0,971	0,977	0,982	0,985
2000	0,185	0,337	0,461	0,562	0,645	0,713	0,767	0,812	0,848	0,878	0,902	0,921	0,937	0,950	0,960	0,968	0,974	0,980	0,984	0,987
2500	0,190	0,344	0,470	0,572	0,655	0,722	0,776	0,820	0,856	0,885	0,908	0,926	0,941	0,953	0,963	0,971	0,977	0,982	0,986	0,989
3000	0,194	0,350	0,477	0,580	0,663	0,730	0,783	0,827	0,862	0,890	0,912	0,930	0,945	0,956	0,965	0,973	0,979	0,983	0,987	0,990
3500	0,197	0,355	0,483	0,586	0,669	0,736	0,789	0,832	0,866	0,894	0,916	0,933	0,947	0,958	0,967	0,974	0,980	0,984	0,988	0,990
4000	0,199	0,360	0,488	0,592	0,675	0,741	0,794	0,836	0,870	0,897	0,919	0,936	0,950	0,960	0,969	0,976	0,981	0,985	0,988	0,991
4500	0,202	0,363	0,493	0,596	0,679	0,745	0,798	0,840	0,874	0,900	0,921	0,938	0,951	0,962	0,970	0,977	0,982	0,986	0,989	0,992
5000	0,204	0,367	0,497	0,601	0,684	0,749	0,802	0,844	0,877	0,903	0,924	0,940	0,953	0,963	0,971	0,978	0,983	0,987	0,990	0,992
5500	0,206	0,370	0,500	0,605	0,687	0,753	0,805	0,847	0,879	0,905	0,926	0,942	0,955	0,965	0,972	0,979	0,983	0,987	0,990	0,992
6000	0,207	0,372	0,504	0,608	0,691	0,756	0,808	0,849	0,882	0,907	0,927	0,943	0,956	0,966	0,973	0,979	0,984	0,988	0,991	0,993
6500	0,209	0,375	0,507	0,611	0,694	0,759	0,811	0,852	0,884	0,909	0,929	0,945	0,957	0,967	0,974	0,980	0,985	0,988	0,991	0,993
7000	0,210	0,377	0,509	0,614	0,697	0,762	0,813	0,854	0,886	0,911	0,930	0,946	0,958	0,967	0,975	0,981	0,985	0,988	0,991	0,993
7500	0,212	0,379	0,512	0,617	0,699	0,764	0,816	0,856	0,887	0,912	0,932	0,947	0,959	0,968	0,975	0,981	0,985	0,989	0,991	0,993
8000	0,213	0,381	0,514	0,619	0,702	0,767	0,818	0,858	0,889	0,914	0,933	0,948	0,960	0,969	0,976	0,982	0,986	0,989	0,992	0,994
8500	0,214	0,383	0,516	0,621	0,704	0,769	0,819	0,859	0,891	0,915	0,934	0,949	0,961	0,970	0,977	0,982	0,986	0,989	0,992	0,994
9000	0,215	0,385	0,519	0,624	0,706	0,771	0,821	0,861	0,892	0,916	0,935	0,950	0,961	0,970	0,977	0,982	0,987	0,990	0,992	0,994
9500	0,216	0,387	0,521	0,626	0,708	0,772	0,823	0,862	0,893	0,917	0,936	0,951	0,962	0,971	0,978	0,983	0,987	0,990	0,992	0,994
10000	0,217	0,388	0,522	0,627	0,710	0,774	0,825	0,864	0,895	0,918	0,937	0,951	0,963	0,971	0,978	0,983	0,987	0,990	0,993	0,994
10500	0,218	0,390	0,524	0,629	0,712	0,776	0,826	0,865	0,896	0,919	0,938	0,952	0,963	0,972	0,978	0,983	0,987	0,990	0,993	0,994
11000	0,219	0,391	0,526	0,631	0,713	0,777	0,827	0,866	0,897	0,920	0,939	0,953	0,964	0,972	0,979	0,984	0,988	0,991	0,993	0,995
11500	0,220	0,393	0,527	0,633	0,715	0,779	0,829	0,868	0,898	0,921	0,939	0,953	0,964	0,973	0,979	0,984	0,988	0,991	0,993	0,995
12000	0,221	0,394	0,529	0,634	0,716	0,780	0,830	0,869	0,899	0,922	0,940	0,954	0,965	0,973	0,979	0,984	0,988	0,991	0,993	0,995
12500	0,222	0,395	0,530	0,636	0,718	0,782	0,831	0,870	0,900	0,923	0,941	0,955	0,965	0,973	0,980	0,985	0,988	0,991	0,993	0,995
13000	0,223	0,396	0,532	0,637	0,719	0,783	0,832	0,871	0,901	0,924	0,941	0,955	0,966	0,974	0,980	0,985	0,988	0,991	0,993	0,995
13500	0,223	0,397	0,533	0,638	0,720	0,784	0,834	0,872	0,901	0,924	0,942	0,956	0,966	0,974	0,980	0,985	0,989	0,991	0,994	0,995
14000	0,224	0,398	0,534	0,640	0,722	0,785	0,835	0,873	0,902	0,925	0,943	0,956	0,966	0,974	0,981	0,985	0,989	0,992	0,994	0,995
14500	0,225	0,400	0,535	0,641	0,723	0,786	0,836	0,874	0,903	0,926	0,943	0,957	0,967	0,975	0,981	0,985	0,989	0,992	0,994	0,995
15000	0,225	0,401	0,537	0,642	0,724	0,788	0,837	0,874	0,904	0,926	0,944	0,957	0,967	0,975	0,981	0,986	0,989	0,992	0,994	0,995
15500	0,226	0,402	0,538	0,643	0,725	0,789	0,838	0,875	0,904	0,927	0,944	0,957	0,968	0,975	0,981	0,986	0,989	0,992	0,994	0,995

# Apêndice D

Neste apêndice apresentamos o método clássico da análise de média-variância introduzido por Markowitz (1952) e alguns exemplos relacionados à seleção de *portfolios* e análise de correlação dos ativos com o mercado. Temos, por objetivo, auxiliar o leitor na compreensão dos conceitos apresentados no Capítulo 4, referentes à seleção de *portfolios* eficientes e interpretação dos coeficientes do modelo CAPM. Além disso, apresentamos exemplos do cálculo dos pesos dos ativos para um *portfolio*, do parâmetro  $\beta$  do modelo CAPM e do  $\alpha$  de um ativo. Estes exemplos auxiliam no entendimento dos conceitos apresentados no Capítulo 4, relacionados ao modelo CAPM.

## D.1 Método Clássico de Análise da Média-Variância

Apresentamos nesta seção os principais resultados relacionados à análise da média-variância introduzida por Markowitz (1952).

Considere um *portfolio*  $\mathcal{P} = \{A_1, \dots, A_N\}$  composto por  $N$  ativos com pesos  $\mathbf{a} = (a_1, \dots, a_N)'$ , isto é, para todo  $i = 1, \dots, N$ ,  $a_i$  representa a participação percentual do ativo  $A_i$  em relação ao total do *portfolio*, de forma que  $\sum_{i=1}^N a_i = 1$ . Seja  $\mathbf{R} = (R_1, \dots, R_N)'$  o vetor dos retornos dos ativos, onde  $R_i$  é o retorno do ativo  $A_i$ , para todo  $i = 1, \dots, N$  (veja Definição 4.3). Seja  $\boldsymbol{\mu} = (\mathbb{E}(R_1), \dots, \mathbb{E}(R_N))'$  o vetor dos valores esperados dos retornos e  $\boldsymbol{\Sigma}$  a matriz de variâncias-covariâncias de ordem  $N \times N$ . Sob essas hipóteses o retorno do *portfolio*, dado por  $R_{\mathcal{P}} = \mathbf{a}'\mathbf{R}$ , é uma variável aleatória com valor esperado e variância dados por

$$\begin{aligned}\mathbb{E}(R_{\mathcal{P}}) &= \mathbf{a}'\boldsymbol{\mu} \\ \text{Var}(R_{\mathcal{P}}) &= \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}.\end{aligned}$$

Assume-se que o vetor  $\boldsymbol{\mu}$  e a matriz  $\boldsymbol{\Sigma}$  sejam conhecidos. Na prática isso não acontece e eles devem ser estimados.

Ao escolher os pesos do *portfolio* o investidor deve escolher entre todos os possíveis pares de média-variância disponíveis. Para calcular os pesos para cada possível par o investidor fixa um retorno esperado  $\mu_0$ . Segundo Markowitz (1952), o vetor  $\mathbf{a}$  é dado por

$$\min_{\mathbf{a}} \{\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}\}$$

sujeito às condições

$$\begin{aligned}\mu_0 &= \mathbf{a}'\boldsymbol{\mu} \\ \mathbf{a}'\mathbf{1} &= 1, \quad \mathbf{1} = (1, 1, \dots, 1)'\end{aligned}$$

A solução deste problema é dada por

$$\mathbf{a} = \mathbf{g} + \mathbf{h}\mu_0,$$

onde

$$\mathbf{g} = \frac{1}{ac - b^2} \boldsymbol{\Sigma}^{-1}(c\mathbf{1} - b\boldsymbol{\mu}), \quad \mathbf{h} = \frac{1}{ac - b^2} \boldsymbol{\Sigma}^{-1}(a\boldsymbol{\mu} - b\mathbf{1}),$$

$$a = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}, \quad b = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \quad \text{e} \quad c = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}.$$

Para mais detalhes veja Fabozzi *et al.* (2006).

## D.2 Cálculo dos Pesos do *Portfolio* Eficiente

Apresentamos, nesta seção, o exemplo do cálculo dos pesos dos ativos de um *portfolio* eficiente. Este exemplo ilustra os conceitos desenvolvidos na Seção 4.3.

Dados dois ativos de risco  $A_1$  e  $A_2$ , desejamos saber quais os seus respectivos pesos  $a_1$  e  $a_2$  de forma que o *portfolio* seja considerado eficiente. Note que, estamos construindo um *portfolio* sem o ativo livre de risco. Ele será portanto considerado um *portfolio* tangente (ou *portfolio* do mercado). Dessa forma para calcular os pesos ótimos utilizamos a expressão (4.22), isto é,

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \frac{1}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_F\mathbf{1})} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_F\mathbf{1}), \quad (\text{D.1})$$

onde

$$\boldsymbol{\Sigma} = \begin{bmatrix} \text{Var}(R_1) & \text{Cov}(R_1, R_2) \\ \text{Cov}(R_1, R_2) & \text{Var}(R_2) \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mathbb{E}(R_1) \\ \mathbb{E}(R_2) \end{pmatrix}, \quad \mathbf{1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad (\text{D.2})$$

$R_F$  representa o retorno do ativo livre de risco e  $R_1$  e  $R_2$  são os retornos dos ativos  $A_1$  e  $A_2$ , respectivamente.

Invertendo a matriz  $\boldsymbol{\Sigma}$  e substituindo na expressão (D.1) obtemos

$$a_1 = \frac{(\mathbb{E}(R_1) - R_F)\text{Var}(R_2) - (\mathbb{E}(R_2) - R_F)\text{Cov}(R_1, R_2)}{(\mathbb{E}(R_1) - R_F)\text{Var}(R_2) + (\mathbb{E}(R_2) - R_F)\text{Var}(R_1) - (\mathbb{E}(R_1) + \mathbb{E}(R_2) - 2R_F)\text{Cov}(R_1, R_2)} \quad (\text{D.3})$$

e  $a_2 = 1 - a_1$ . Para estimar a expressão (D.3) assumimos que

$$\mathbb{E}(R_1) = \bar{R}_1, \quad \mathbb{E}(R_2) = \bar{R}_2 \quad \text{e} \quad \text{Cov}(R_1, R_2) = \frac{1}{n} \sum_{t=1}^n (R_{1,t} - \bar{R}_1)(R_{2,t} - \bar{R}_2).$$

No exemplo que segue, apresentamos uma situação hipotética e calculamos os pesos  $a_1$  e  $a_2$  dados pela expressão (D.3).

**Exemplo D.1. (Pesos do *Portfolio* Eficiente).** Considere a seguinte situação ilustrativa:

O *portfolio* é dado por  $\mathcal{P} = \{A, B\}$  dois ativos de risco, com retornos médios e desvio padrão apresentados na Tabela D.1 e  $\text{Cov}(R_A, R_B) = 0,5$ .

Tabela D.1: Média e Desvio Padrão dos Ativos do *Portfolio* no Período Considerado.

Ação	Retorno Esperado (Média Histórica)	Risco (Desvio Padrão)
A	14%	6%
B	8%	3%

Suponha que o retorno do título de renda fixa sem risco (Título do Tesouro Nacional) para o período de investimento é  $R_F = 5\%$ . Pela expressão (D.3) temos

$$a_1 = \frac{(14 - 5) \times 9 - (8 - 5) \times 0,5}{(14 - 5) \times 9 + (8 - 5) \times 36 - (14 + 8 - 2 \times 5) \times 0,5} = 0,42308. \quad (\text{D.4})$$

Portanto,  $a_1 = 42,31\%$  e  $a_2 = 57,69\%$ .

Pela expressão (4.15), segue que o retorno médio esperado deste *portfolio* é

$$\bar{R}_P = 0,4231 \times 14\% + 0,5769 \times 8\% = 10,66\%.$$

### D.3 Cálculo do Coeficiente $\beta$ e do $\alpha$ de um Ativo

Apresentamos, a seguir, o exemplo do cálculo do coeficiente  $\beta$  dos ativos de um *portfolio*. O coeficiente  $\beta$  é um indicador que mede como reage o preço de uma ação às oscilações do índice de mercado (Bovespa, por exemplo). Este exemplo tem por objetivo esclarecer os conceitos desenvolvidos na Seção 4.3.2.

**Exemplo D.2. (Cálculo do Coeficiente  $\beta$  e do  $\alpha$  de Ações).** Na Seção 4.3.2 tratamos do cálculo do coeficiente  $\beta$  de um ativo. Para fixar os resultados apresentados, suponha que os retornos da ação A e os retornos do mercado (índice Bovespa), para os últimos 6 meses são aqueles dados na Tabela D.2.

Tabela D.2: Retornos Mensais do Preço das Ações da Empresa A e do Índice do Mercado.

Mês	Retorno do Mercado (Índice Bovespa) (%)	Retorno da Ação da Empresa A (%)
1	27	30
2	-3	-10
3	12	15
4	-3	-5
5	12	5
6	27	25

Supondo que a taxa de aplicação sem risco seja de  $6,5\%$  (Título do Tesouro Nacional), vamos calcular o  $\beta$  desta ação e em seguida comparar com a taxa de retorno requerida pelo mercado, isto é, aquele dado pelo modelo CAPM.

Sob as hipóteses apresentadas na Tabela D.2 seguem os resultados apresentados na Tabela D.3. A terceira e a sexta coluna da Tabela D.3 apresentam, respectivamente, as diferenças entre o retorno observado e o retorno médio do mercado e da empresa. A quarta e a sétima coluna apresentam os valores quadráticos dessas diferenças. Na última linha da tabela observamos a soma dos valores citados.

Tabela D.3: Tabela dos Retornos do Mercado e da Ação da Empresa A, Retornos Médios e Valores Quadráticos dos Retornos Médios.

Mês ( $t$ )	Retorno do Mercado $R_{M,t}$ (IBovespa)	$R_{M,t} - \bar{R}_M$	$(R_{M,t} - \bar{R}_M)^2$	Retorno da Ação $R_{A,t}$ (Empresa A)	$R_{A,t} - \bar{R}_A$	$(R_{A,t} - \bar{R}_A)^2$
1	27	15	225	30	20	300
2	-3	-15	225	-10	-20	300
3	12	0	0	15	5	0
4	-3	-15	225	-5	-15	225
5	12	0	0	5	-5	0
6	27	15	225	25	15	225
TOTAL	72	0	900	60	0	1.050

Segue, pelos resultados apresentados na Tabela D.3, que

$$\text{Var}(R_M) = \frac{1}{6} \sum_{t=1}^6 (R_{M,t} - \bar{R}_M)^2 = 150 \quad \text{e}$$

$$\text{Cov}(R_A, R_M) = \frac{1}{6} \sum_{t=1}^6 (R_{A,t} - \bar{R}_A)(R_{M,t} - \bar{R}_M) = 175.$$

Logo, pela expressão (4.28), o coeficiente  $\beta$  da ação é dado por

$$\beta_A = \frac{\text{Cov}(R_A, R_M)}{\text{Var}(R_M)} = \frac{175}{150} = 1,167. \quad (\text{D.5})$$

Observe que  $\beta_A > 1$ . Isso significa que o ativo é agressivo, isto é, os retornos e suas perdas tendem a ser maiores do que as do mercado.

Lembramos que a taxa de retorno requerida pelo mercado para esse ativo é dada pelo modelo CAPM. Segue, pela expressão (4.30), que

$$\mathbb{E}(R_A) = R_F + \beta_A [\mathbb{E}(R_M) - R_F].$$

Substituindo-se o valor de  $\beta_A$ , dado na expressão (D.5), o retorno esperado do mercado  $\mathbb{E}(R_M) = \bar{R}_M$  e a taxa livre de risco  $R_F = 6,5\%$ , temos

$$\mathbb{E}(R_A) = 0,065 + 1,167 [0,12 - 0,065] = 0,1291 = 12,91\%.$$



Isso significa que o retorno esperado para esse ativo, em situação de equilíbrio, é igual à 12,91%. O retorno médio da ação é  $\bar{R}_A = \frac{1}{6} \sum_{t=1}^6 R_{A,t} = 10\%$ . Pela expressão (4.29), segue que o  $\alpha$  da ação é dado por

$$\alpha_A = \bar{R}_A - \mathbb{E}(R_A) = -2,91\%.$$

Na visão do investidor, a ação está superavaliada pelo mercado ( $\alpha$  negativo) e deverá proporcionar, no período de investimento, uma rentabilidade inferior àquela calculada pelo modelo de equilíbrio (CAPM). Portanto, deve ser vendida.

## D.4 Cálculo do Coeficiente $\beta$ de um *Portfolio*

Na seção anterior tratamos do cálculo do coeficiente  $\beta$  para um único ativo. No exemplo que segue apresentamos o cálculo do coeficiente  $\beta$  para um *portfolio* de ações. Esse valor nos fornece uma idéia do comportamento dos retornos do *portfolio* em relação aos retornos do mercado.

**Exemplo D.3. (Cálculo do Coeficiente  $\beta$  de um *Portfolio*).** Suponha que um investidor dispõe das informações, apresentadas na Tabela D.4, sobre as ações que compõem o seu *portfolio*  $\mathcal{P} = \{A, B, C\}$ .

Tabela D.4: Coeficiente  $\beta$  e Pesos dos Ativos A,B e C no *Portfolio*  $\mathcal{P}$ .

Ação	$\beta$	Porcentagem da Ação no <i>Portfolio</i> (%)
A	1,10	20
B	0,80	50
C	1,00	30

Suponha ainda que a melhor estimativa para o retorno e o risco do *portfolio* do mercado  $M$  (IBOVESPA, por exemplo) para o período do investimento é  $\mathbb{E}(R_M) = 13\%$  e a taxa livre de risco é de 5%. Sob essas condições vamos calcular o retorno esperado do *portfolio* em situação de equilíbrio e o coeficiente  $\beta$  deste *portfolio*.

Lembramos que, pela expressão (4.15), o retorno esperado do *portfolio* é dado por

$$\mathbb{E}(R_{\mathcal{P}}) = a_A \mathbb{E}(R_A) + a_B \mathbb{E}(R_B) + a_C \mathbb{E}(R_C), \quad (\text{D.6})$$

onde  $a_j$  e  $\mathbb{E}(R_j)$ , para  $j = A, B, C$  representam, respectivamente, o percentual da ação no *portfolio* e o retorno da ação. Na situação de equilíbrio, o retorno esperado de cada ação, segundo o modelo CAPM, é dado pela expressão (4.30). Segue que

$$\begin{aligned} \mathbb{E}(R_A) &= 5 + 1,10(13 - 5) = 13,80\%, \\ \mathbb{E}(R_B) &= 5 + 0,80(13 - 5) = 11,40\%, \\ \mathbb{E}(R_C) &= 5 + 1,00(13 - 5) = 13,00\%. \end{aligned}$$

Pela expressão (D.6), segue que o valor esperado do retorno do *portfolio*, para o período de investimento, é dado por

$$\mathbb{E}(R_{\mathcal{P}}) = 0,20 \times 13,80\% + 0,50 \times 11,40\% + 0,30 \times 13,00\% = 12,36\%.$$

Pela expressão (4.31), coeficiente  $\beta$  (medida do risco sistemático) do *portfolio* é dado por

$$\begin{aligned}\beta_{\mathcal{P}} &= a_A\beta_A + a_B\beta_B + a_C\beta_C \\ &= 0,20 \times 1,10 + 0,50 \times 0,80 + 0,30 \times 1,00 \\ &= 0,92.\end{aligned}$$

Observamos que  $\beta_{\mathcal{P}} < 1$ . Isso indica que, caso ocorra uma queda no índice do mercado, as perdas sofridas pelo *portfolio* são menores do que as do mercado como um todo.

# Apêndice E

Neste apêndice apresentamos resultados auxiliares ao cálculo das medidas de risco VaR e ES do *portfolio* das séries simuladas, apresentado no Capítulo 5.

## E.1 Ajuste dos Modelos EGARCH e FIEGARCH à Série das Perdas do *Portfolio* para as Séries Temporais Simuladas

Nesta seção apresentamos o ajuste dos modelos EGARCH e FIEGARCH à série das perdas do *portfolio*, isto é, à série temporal  $\{-r_{\mathcal{P},t}\}_{t=1}^{1.990}$ , onde  $r_{\mathcal{P},t}$  é o log-retorno do *portfolio* no instante  $t$ . Estes modelos foram utilizados para estimar  $\sigma_{1.991}$ , para o cálculo das medidas de risco apresentadas no Capítulo 5.

Na Figura E.1 apresentamos a série de perdas do *portfolio* e sua função de autocorrelação amostral. Note que, para  $h > 1$  os valores da função de autocorrelação amostral são não-significativos. Entretanto, na Figura E.2 observamos que a função de autocorrelação amostral é significativa para vários valores de  $h > 1$ .

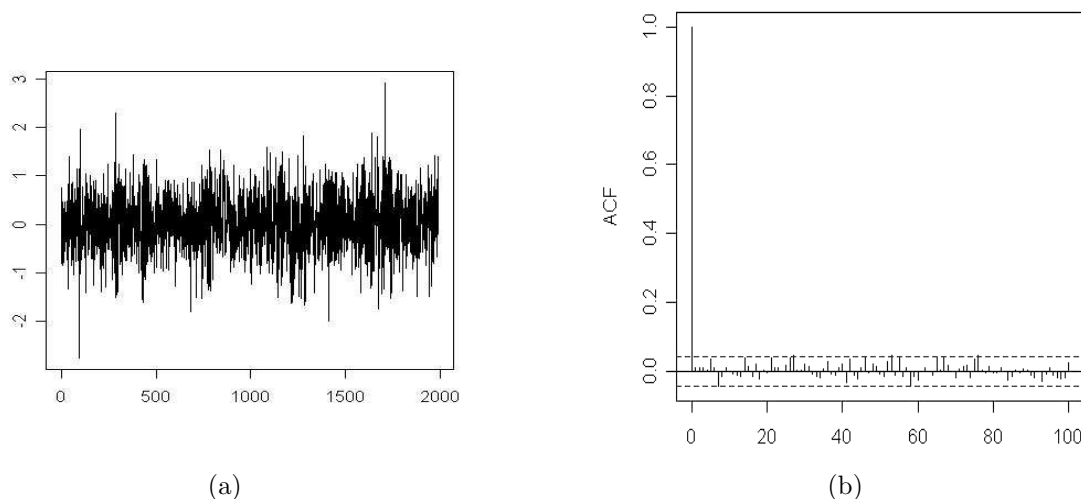


Figura E.1: (a) Série Temporal das Perdas do *Portfolio*. (b) Função de Autocorrelação Amostral.

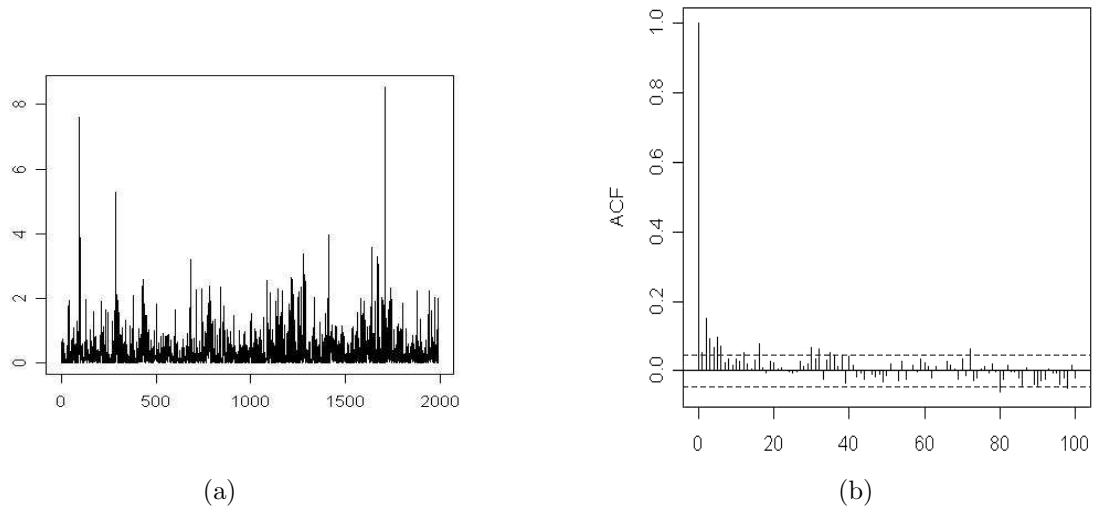


Figura E.2: (a) Série dos Quadrados das Perdas do *Portfolio*. (b) Função de Autocorrelação Amostral.

Na Figura E.3 apresentamos o histograma e o Q×Q plot das perdas do *portfolio*. Pelo Q×Q plot podemos concluir que as perdas são normalmente distribuídas.

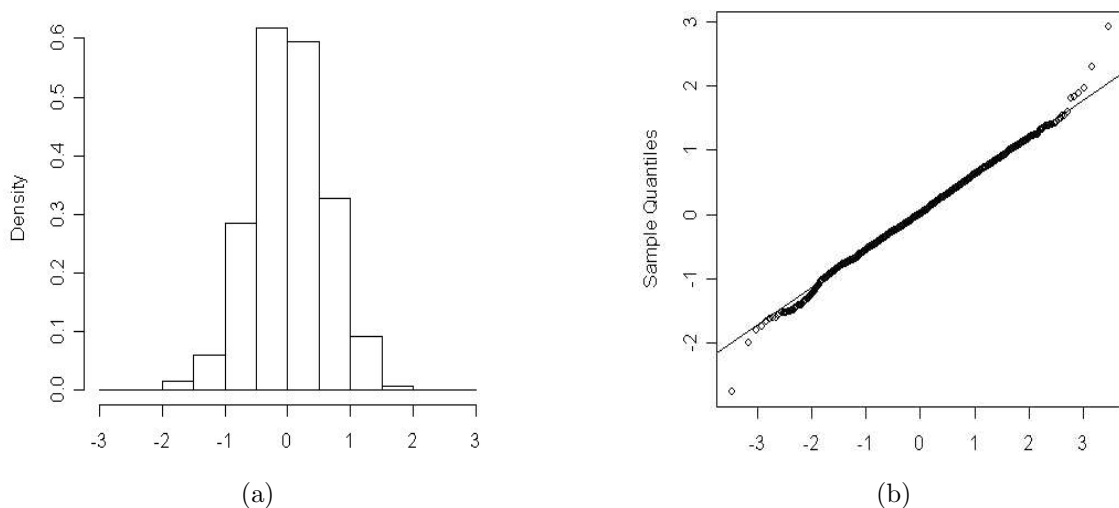


Figura E.3: (a) Histograma, (b) Q×Q plot das Perdas do *Portfolio*.

A função periodograma do quadrado das perdas do *portfolio* é apresentada na Figura E.4. Como podemos perceber, o comportamento desta função não indica a presença de longa dependência na série temporal.

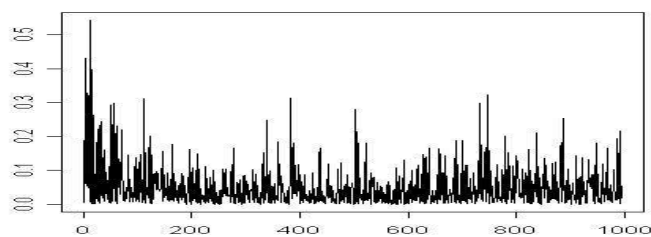


Figura E.4: Função Periodograma da Série Temporal  $\{r_{P,t}^2\}_{t=1}^{1.990}$ .

Os modelos EGARCH e FIEGARCH ajustados à série temporal são apresentados, respectivamente, nas Tabelas E.1 e E.2. Ressaltamos que o S-Plus utiliza a notação apresentada na expressão (E.1) para definir o logaritmo da volatilidade para modelos EGARCH

$$\beta(\mathcal{B}) \ln(\sigma_t^2) = a + \sum_{i=0}^p \psi_i (|Z_{t-1-i}| + \gamma_i Z_{t-1-i}), \quad (\text{E.1})$$

com  $\beta(\cdot)$  definido pela expressão (2.37). Note que esta expressão é equivalente à expressão apresentada na equação (2.36) sob as seguintes condições  $\psi_i = -\alpha_i \times \gamma$  e  $\psi_i \times \gamma_i = -\alpha_i \times \theta$ , onde  $\alpha_i$  é definido pela expressão (2.37), para todo  $i = 1, \dots, p$  e  $\gamma$  e  $\theta$  são definidos na expressão (2.38).

Optamos pelos modelos mais parcimoniosos, isto é, com o menor número de parâmetros.

Tabela E.1: Modelo EGARCH( $p, q$ ), com  $p = 0, q = 1$  Ajustado à Série Temporal das Perdas do *Portfolio* e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,232	0,056	-4,105	0,000
$\beta_1$	0,907	0,033	27,783	0,000
$\psi_0$	0,163	0,033	4,877	0,000
$\gamma_0$	0,130	0,097	1,344	0,090
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 3510,251		Estatística	p-valor	
BIC = 3532,635		5,109	0,078	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
8,484	0,746	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
11,60	0,48	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
11,91	0,45	1,09	0,48	

Note que, em ambos os casos, tanto os resíduos como os quadrados dos resíduos dos modelos são não-correlacionados. Além disso, pelo teste de Jarque-Bera, os resíduos são normalmente distribuídos.

Tabela E.2: Modelo FIEGARCH( $p, d, q$ ), com  $p = 1$ ,  $q = 0$  Ajustado à Série Temporal das Perdas do *Portfolio* e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,282	0,065	-4,332	0,000
$\psi_0$	0,086	0,055	1,582	0,057
$\psi_1$	0,235	0,051	4,589	0,000
$\gamma_0$	0,005	0,032	0,151	0,440
$\gamma_1$	0,044	0,030	1,492	0,068
$d$	0,432	0,075	5,730	0,000

Critério de Seleção		Teste de Normalidade Jarque-Bera	
AIC = 3511,354		Estatística	p-valor
BIC = 3544,929		3,768	0,152

Teste de Ljung-Box para os resíduos padronizados			
Estatística	p-valor	$\chi^2$ (graus de liberdade)	
9,067	0,697	12	

Teste de Ljung-Box para o quadrado dos resíduos padronizados			
Estatística	p-valor	$\chi^2$ (graus de liberdade)	
11,63	0,48	12	

Teste dos Multiplicadores de Lagrange			
ML	p-valor	Estatística F	p-valor
11,31	0,50	1,03	0,53

## E.2 Ajuste de Modelos EGARCH aos Ativos do *Portfolio* das Séries Temporais Simuladas

Nesta seção apresentamos os modelos EGARCH ajustados às séries temporais simuladas apresentadas no Capítulo 5. Para a seleção do modelo ajustamos, à cada série temporal, processos EGARCH( $p, q$ ), com  $p, q \in \{0, 1, 2, 3\}$ . Como critério de seleção utilizamos o teste de Ljung-Box para os resíduos (e seus quadrados) dos modelos ajustados. Os modelos selecionados são apresentados nas Tabelas E.3 - E.6.

Tabela E.3: Modelo EGARCH( $p, q$ ), com  $p = 0, q = 2$  Ajustado à Série Temporal dos Log-retornos do Ativo  $A_1$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,171	0,035	-4,943	0,000
$\beta_1$	1,245	0,144	8,649	0,000
$\beta_2$	-0,294	0,138	-2,132	0,017
$\psi_0$	0,232	0,046	5,046	0,000
$\gamma_0$	0,283	0,075	3,748	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 6239,634		Estatística	p-valor	
BIC = 6267,613		0,2274	0,8925	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
11,4800	0,4885	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
16,660	0,163	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
17,340	0,137	1,591	0,194	

Tabela E.4: Modelo EGARCH( $p, q$ ), com  $p = 0, q = 3$  Ajustado à Série Temporal dos Log-retornos do Ativo  $A_2$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,112	0,020	-5,963	0,000
$\beta_1$	1,958	0,110	17,876	0,000
$\beta_2$	-1,606	0,187	-8,604	0,000
$\beta_3$	0,605	0,091	6,672	0,000
$\psi_0$	0,142	0,024	5,881	0,000
$\gamma_0$	-0,004	0,070	-0,053	0,479
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 5338,327		Estatística	p-valor	
BIC = 5371,903		0,5857	0,7462	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
13,730	0,318	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
20,900	0,052	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
21,230	0,047	1,951	0,104	

Tabela E.5: Modelo EGARCH( $p, q$ ), com  $p = 2, q = 1$  Ajustado à Série Temporal dos Log-retornos do Ativo  $A_3$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,192	0,039	-4,915	0,000
$\beta_1$	0,898	0,033	27,523	0,000
$\psi_0$	0,131	0,054	2,422	0,008
$\psi_1$	0,391	0,069	5,705	0,000
$\psi_2$	-0,270	0,063	-4,273	0,000
$\gamma_0$	-0,518	0,319	-1,623	0,052
$\gamma_1$	0,245	0,106	2,321	0,010
$\gamma_2$	0,136	0,113	1,201	0,115
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 5824,682		Estatística	p-valor	
BIC = 5869,449		2,905	0,234	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
12,35	0,42	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
18,04	0,11	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
19,16	0,08	1,76	0,14	

Tabela E.6: Modelo EGARCH( $p, q$ ), com  $p = 0, q = 2$  Ajustado à Série Temporal dos Log-retornos do Ativo  $A_4$  e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$a$	-0,098	0,024	-4,043	0,000
$\beta_1$	1,459	0,117	12,479	0,000
$\beta_2$	-0,522	0,111	-4,720	0,000
$\psi_0$	0,157	0,034	4,672	0,000
$\gamma_0$	0,236	0,080	2,952	0,002
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = 6497,908		Estatística	p-valor	
BIC = 6525,887		0,3314	0,8473	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
18,19	0,11	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
20,21	0,06	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
19,050	0,087	1,749	0,147	



# Apêndice F

## F.1 Modelos para as Séries Temporais dos Log-retornos dos Ativos do *Portfolio*

Nesta seção apresentamos os modelos  $ARMA(p_1, q_1)$ -EGARCH( $p_2, q_2$ ) e também os modelos  $ARMA(p_1, q_1)$ -FIEGARCH( $p_2, d, q_2$ ) selecionados, após a análise dos resíduos, para as séries temporais dos log-retornos dos ativos do *portfolio*. Esses modelos são utilizados para a previsão da média e da variância condicional, utilizadas para o cálculo das medidas de risco apresentadas na Seção 6.4. Apresentamos também o modelo  $ARMA(p_1, q_1)$ -FIEGARCH( $p_2, d, q_2$ ) ajustado aos log-retornos do IBovespa (veja Seção 6.2 para mais detalhes).

Ressaltamos que, para a série temporal dos log-retornos dos preços das ações da Gerdau não foi possível ajustar um modelo  $ARMA(p_1, q_1)$ -EGARCH( $p_2, d, q_2$ ) de forma a eliminar a correlação entre as variáveis aleatórias da série temporal dos valores quadráticos dos resíduos padronizados. Por essa razão, optamos por um modelo com o mesmo número de parâmetros dos modelos ajustados para as outras séries temporais.

Tabela F.1: Modelo  $ARMA(p_1, q_1)$ -EGARCH( $p_2, d, q_2$ ), com  $p_1 = 1, q_1 = 0, p_2 = 0, q_2 = 1$  Ajustado aos Log-retornos dos Preços das Ações do Ativo  $A_1$  (Bradesco) e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\phi_1$	0,110	0,024	4,531	0,002
$a$	-0,926	0,107	-8,650	0,000
$\beta_1$	0,907	0,014	67,167	0,000
$\psi_0$	0,331	0,026	12,714	0,000
$\gamma_0$	0,383	0,056	6,776	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -7684,604		Estatística	p-valor	
BIC = -7657,330		110,50	0,00	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
13,28	0,35	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
20,05	0,07	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
20,88	0,05	1,92	0,11	

Tabela F.2: Modelo ARMA( $p_1, q_1$ )-EGARCH( $p_2, d, q_2$ ), com  $p_1 = 1, q_1 = 0, p_2 = 1, q_2 = 1$  Ajustado aos Log-retornos dos Preços das Ações do Ativo  $A_2$  (Brasil Telecom) e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\phi_1$	0,114	0,025	4,656	0,000
$a$	-0,591	0,090	-6,587	0,000
$\beta_1$	0,939	0,011	82,223	0,000
$\psi_0$	0,238	0,021	11,353	0,000
$\gamma_0$	0,084	0,040	2,119	0,017
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = - 6900,266		Estatística	p-valor	
BIC = -6872,992		2393,00	0,00	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
8,51	0,74	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
6,18	0,91	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
6,925	0,863	0,632	0,902	

Tabela F.3: Modelo ARMA( $p_1, q_1$ )-EGARCH( $p_2, d, q_2$ ), com  $p_1 = 1, q_1 = 0, p_2 = 1, q_2 = 1$  Ajustado aos Log-retornos dos Preços das Ações do Ativo  $A_3$  (Gerdau) e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\phi_1$	0,135	0,027	5,002	0,000
$a$	-1,904	0,250	-7,610	0,000
$\beta_1$	0,763	0,034	22,587	0,000
$\psi_0$	0,354	0,032	11,145	0,000
$\gamma_0$	0,146	0,055	2,664	0,003
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -7092,381		Estatística	p-valor	
BIC = -7065,107		496,50	0,00	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
17,23	0,14	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
32,450	0,001	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
33,75	0,00	3,13	0,02	

Tabela F.4: Modelo ARMA( $p_1, q_1$ )-EGARCH( $p_2, d, q_2$ ), com  $p_1 = 1, q_1 = 0, p_2 = 0, q_2 = 1$  Ajustado aos Log-retornos dos Preços das Ações do Ativo  $A_4$  (Petrobrás) e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\phi_1$	0,131	0,023	5,675	0,000
$a$	-0,714	0,072	-9,907	0,000
$\beta_1$	0,931	0,009	100,983	0,000
$\psi_0$	0,285	0,016	17,471	0,000
$\gamma_0$	0,402	0,034	11,723	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -7471,269		Estatística	p-valor	
BIC = -7443,995		194,60	0,00	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
13,19	0,36	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
18,71	0,10	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
19,05	0,09	1,75	0,15	

Tabela F.5: Modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 0, q_1 = 1, p_2 = 0, q_2 = 1$  Ajustado aos Log-retornos do IBovespa e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\theta_1$	0,078	0,027	2,950	0,002
$a$	-0,347	0,050	-6,923	0,000
$\beta_1$	0,706	0,077	9,140	0,000
$\psi_0$	0,275	0,024	11,407	0,000
$\gamma_0$	-0,166	0,017	-9,517	0,000
$d$	0,339	0,065	5,242	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -8264,639		Estatística	p-valor	
BIC = -8231,911		90,81	0,00	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
13,92	0,31	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
11,18	0,51	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
10,92	0,54	0,99	0,56	

Tabela F.6: Modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 0, q_1 = 1, p_2 = 0, q_2 = 1$  Ajustado aos Log-retornos dos Preços das Ações do Ativo  $A_1$  (Bradesco) e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\theta_1$	0,129	0,025	5,218	0,000
$a$	-0,374	0,046	-8,086	0,000
$\beta_1$	0,453	0,085	5,315	0,000
$\psi_0$	0,381	0,039	9,726	0,000
$\gamma_0$	-0,135	0,020	-6,611	0,000
$d$	0,446	0,049	9,201	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -7702,270		Estatística	p-valor	
BIC = -7669,542		121,10	0,00	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
12,96	0,37	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
16,33	0,18	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
17,80	0,12	1,64	0,18	

Tabela F.7: Modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 0, q_1 = 1, p_2 = 1, q_2 = 1$  Ajustado aos Log-retornos dos Preços das Ações do Ativo  $A_2$  (Brasil Telecom) e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\phi_0$	0,002	0,001	2,362	0,009
$\theta_1$	0,103	0,029	3,574	0,000
$a$	-0,053	0,021	-2,568	0,005
$\beta_1$	0,905	0,043	20,923	0,000
$\psi_0$	0,382	0,032	11,782	0,000
$\psi_1$	-0,331	0,039	-8,514	0,000
$\gamma_0$	0,044	0,019	2,300	0,010
$\gamma_1$	-0,066	0,021	-3,171	0,000
$d$	0,447	0,065	6,934	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -6935,480		Estatística	p-valor	
BIC = -6886,388		1244,00	0,00	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
8,14	0,77	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
4,39	0,98	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
5,04	0,96	0,46	0,99	

Tabela F.8: Modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 1, q_1 = 0, p_2 = 0, q_2 = 4$  Ajustado aos Log-retornos dos Preços das Ações do Ativo  $A_3$  (Gerdau) e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\phi_1$	0,140	0,026	5,381	0,000
$a$	-0,769	0,200	-3,845	0,000
$\beta_1$	0,216	0,062	3,453	0,000
$\beta_2$	0,184	0,049	3,729	0,000
$\beta_3$	0,470	0,041	11,429	0,000
$\beta_4$	-0,450	0,049	-9,154	0,000
$\psi_0$	0,395	0,034	11,784	0,000
$\gamma_0$	-0,046	0,019	-2,439	0,007
$d$	0,256	0,056	4,542	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -7109,976		Estatística	p-valor	
BIC = -7060,884		431,50	0,00	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
18,11	0,11	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
17,42	0,13	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
17,07	0,15	1,57	0,20	

Tabela F.9: Modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 0, q_1 = 1, p_2 = 0, q_2 = 1$  Ajustado aos Log-retornos dos Preços das Ações do Ativo  $A_4$  (Petrobrás) e Resultados da Análise de Resíduos.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\phi_0$	0,001	0,001	2,108	0,018
$\theta_1$	0,119	0,025	4,788	0,000
$a$	-0,347	0,040	-8,680	0,000
$\beta_1$	0,575	0,097	5,928	0,000
$\psi_0$	0,326	0,023	14,348	0,000
$\gamma_0$	-0,110	0,016	-6,958	0,000
$d$	0,416	0,059	7,019	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -7474,958		Estatística	p-valor	
BIC = -7436,775		168,30	0,00	
Teste de Ljung-Box para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
15,77	0,20	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
18,42	0,10	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
19,56	0,08	1,80	0,13	

## F.2 Modelos para a Série Temporal dos Log-retornos do *Portfolio*

Nesta seção apresentamos o modelo ARMA( $p_1, q_1$ )-EGARCH( $p_2, q_2$ ) selecionado, após a análise dos resíduos, para a série dos log-retornos do *portfolio*. Apresentamos também o processo de seleção de modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ) para a série temporal dos log-retornos do *portfolio* analisado e os valores previstos para a média e volatilidade, utilizados no cálculo das medidas de risco apresentado na Seção 6.4.

O modelo ARMA( $p_1, q_1$ )-EGARCH( $p_2, q_2$ ), com  $p_1 = 1, q_1 = 0, p_2 = 0$  e  $q_2 = 1$ , selecionado para a série dos log-retornos do *portfolio* é dado por

$$\begin{aligned} r_t &= \phi_0 + \phi_1 r_{t-1} + X_t \\ X_t &= \sigma_t Z_t \\ (1 - \beta_1 \mathcal{B}) \ln(\sigma_t^2) &= a + \psi_0 (|Z_{t-1}| + \gamma_0 Z_{t-1}), \end{aligned}$$

com  $\phi_0, \phi_1, \beta_1, \psi_0$  e  $\gamma_0$  dados na Tabela F.10.

Tabela F.10: Modelo ARMA( $p_1, q_1$ )-EGARCH( $p_2, q_2$ ), com  $p_1 = 1, q_1 = 0, p_2 = 0, q_2 = 1$  Ajustado aos Log-retornos do *Portfolio*.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\phi_0$	-0,001	0,001	-1,816	0,035
$\phi_1$	0,172	0,025	6,775	0,000
$a$	-1,007	0,113	-8,909	0,000
$\beta_1$	0,899	0,013	67,028	0,000
$\psi_0$	0,295	0,028	10,485	0,000
$\gamma_0$	0,448	0,073	6,127	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -8449,632		Estatística	p-valor	
BIC = -8416,904		287,00	0,00	
Teste de Ljung-Box test para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
13,00	0,37	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
5,22	0,95	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
5,459	0,941	0,498	0,976	

O primeiro passo para a seleção do modelo mais adequado, considerando a longa-dependência, foi o ajuste de modelos ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1$  e  $q_1$  variando de 0 à 3 e  $p_2$  e  $q_2$  iguais a 0 e 1. Selecionamos os modelos para os quais o valor estimado do parâmetro  $d$  foi menor do que 0,5. Em seguida escolhemos os que apresentavam o menor valor dos critérios AIC e BIC e\ou o maior valor da log-verossimilhança. Esses valores são apresentados, em negrito, na Tabela F.11.

Tabela F.11: Valores dos Critérios AIC, BIC e Log-verossimilhança dos Primeiros Modelos Seleccionados para a Série Temporal dos Log-retornos do *Portfolio*  $\mathcal{P} = \{A_1, A_2, A_3, A_4\}$ , onde  $A_1 = \text{Bradesco}$ ,  $A_2 = \text{Brasil Telecom}$ ,  $A_3 = \text{Gerdau}$  e  $A_4 = \text{Petrobrás}$ .

ARMA		FIEGARCH			Critério de Seleção		
$p_1$	$q_1$	$p_2$	d	$q_2$	AIC	BIC	Log-Verossimilhança
1	0	0	0.2294	1	-8448.2644	<b>-8410.0813</b>	4231.1322
2	1	0	0.2367	1	<b>-8456.7829</b>	-8407.6904	4237.3914
2	1	1	0.2835	1	-8456.5281	-8396.5262	<b>4239.2640</b>

Após a análise dos resíduos verificamos que os três modelos são adequados para descrever a série temporal dos log-retornos do *portfolio*. Sendo assim, optamos pelo mais parcimonioso, que é apresentado na Tabela F.12 e é dado por

$$\begin{aligned}
 r_{\mathcal{P},t} &= -0,001 - 0,173r_{\mathcal{P},t-1} + X_{\mathcal{P},t} \\
 X_{\mathcal{P},t} &= \sigma_{\mathcal{P},t}Z_{\mathcal{P},t} \\
 (1 - 0,754\mathcal{B})(1 - \mathcal{B})^{0,233} \ln(\sigma_{\mathcal{P},t}^2) &= -0,498 + 0,285|Z_{\mathcal{P},t-1}| + 0,127Z_{\mathcal{P},t-1}.
 \end{aligned}$$

Tabela F.12: Modelo ARMA( $p_1, q_1$ )-FIEGARCH( $p_2, d, q_2$ ), com  $p_1 = 1, q_1 = 0, p_2 = 0, q_2 = 1$  Ajustado aos Log-retornos do *Portfolio*.

Parâmetro	Estimador	Desvio Padrão	Estatística t	p-valor
$\phi_0$	-0,001	0,001	-1,721	0,043
$\phi_1$	0,173	0,026	6,731	0,000
$a$	-0,498	0,093	-5,381	0,000
$\beta_1$	0,754	0,060	12,552	0,000
$\psi_0$	0,285	0,029	9,706	0,000
$\gamma_0$	0,127	0,017	7,627	0,000
$d$	0,233	0,056	4,176	0,000
Critério de Seleção		Teste de Normalidade Jarque-Bera		
AIC = -8448,410		Estatística	p-valor	
BIC = -8410,227		413,70	0,00	
Teste de Ljung-Box test para os resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
13,49	0,33	12		
Teste de Ljung-Box para o quadrado dos resíduos padronizados				
Estatística	p-valor	$\chi^2$ (graus de liberdade)		
3,57	0,99	12		
Teste dos Multiplicadores de Lagrange				
ML	p-valor	Estatística F	p-valor	
3,726	0,988	0,340	0,999	

As previsões para a média e para a raiz quadrada da volatilidade da série temporal dos log-retornos do *portfolio*, para o horizonte  $h = 1, \dots, 10$ , são apresentadas na Tabela F.13.

Tabela F.13: Previsão da Média e da Raiz Quadrada da Volatilidade dos Log-retornos do *Portfolio* para os Horizontes  $h = 1, \dots, 10$ .

$h$	$\hat{r}_{\mathcal{P},t+h}$	$\hat{\sigma}_{\mathcal{P},t+h}$	$h$	$\hat{r}_{\mathcal{P},t+h}$	$\hat{\sigma}_{\mathcal{P},t+h}$
1	0,0001	0,0149	6	-0,0010	0,0161
2	-0,0008	0,0152	7	-0,0010	0,0162
3	-0,0010	0,0155	8	-0,0010	0,0164
4	-0,0010	0,0157	9	-0,0010	0,0165
5	-0,0010	0,0159	10	-0,0010	0,0166