**Estimação indireta de modelos
R-GARCH**

Jhames Matos Sampaio

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
DOUTOR EM CIÊNCIAS

Programa: Estatística
Orientador: Prof. Dr. Pedro Alberto Morettin

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da
CAPES/CNPq

São Paulo, 2012

Estimação indireta de modelos R-GARCH

Este exemplar corresponde à redação
final da tese devidamente corrigida
e defendida por Jhames Matos Sampaio
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Prof. Dr. Pedro Alberto Morettin (Orientador) - IME-USP.
- Profa. Dra. Clélia Maria Castro Toloí - IME-USP.
- Profa. Dra. Chang Chung Yu Dorea - UnB.
- Profa. Dra. Beatriz Vaz Mendes - UFRJ.
- Prof. Dr. Pedro Luiz Valls Pereira - FGV-SP.

Agradecimentos

Este trabalho é fruto de um projeto de muitos anos de dedicação e creio que todo doutorando concordaria comigo que este é o capítulo mais fácil de uma tese. Agradecer às pessoas importantes que possibilitaram a realização desta tese é também extremamente gratificante e nostálgico, e sendo a lista de nomes consideravelmente grande agradeço de forma sincera a qualquer um que, mesmo não citado, saiba da contribuição para com meu trabalho e minha vida.

Em primeiro lugar, agradeço à minha família que é a razão da minha vida e mantenedora das minhas perspectivas: meus pais Maria José e José Antônio e meu irmão Jheyse.

Agradeço ao professor Pedro Morettin, pela confiança em mim depositada e orientação indispensável no decorrer do doutorado assim como as conversas informais e produtivas.

Ao professor Giorgio Calzolari, pela simpatia e ajuda na solução de alguns problemas.

À professora Sílvia Ferrari, pelas ótimas aulas e por representar um exemplo de profissional.

À professora Airlane Alencar que, muito prestativa e dedicada, foi minha porta de entrada a uma nova visão da ciência.

Aos grandes e inesquecíveis amigos: Luís, Marcus, Guto, Elmar e Rafael, pelos momentos de lazer, descontração e por estarem presentes nos melhores e piores momentos.

Aos meus colegas do grupo de séries temporais: Hugo, Luz, Michele e Fran pela amizade, simpatia e ajuda.

Agradeço também aos membros da comissão examinadora, por aceitar o convite em fazer parte da banca e por todas as sugestões, correções e comentários que melhoraram o trabalho.

Emfim, a todos os colegas, professores e funcionários do Instituto de Matemática e Estatística (IME) que me acolheram.

Resumo

Estimação indireta de modelos R-GARCH

Processos lineares não capturam a estrutura dos dados em finanças. Há uma variedade muito grande de modelos não lineares disponíveis na literatura. A classe de modelos ARCH (Autoregressive Conditional Heterokedastic) foi introduzida por Engle (1982) com o objetivo de estimar a variância da inflação. A idéia nesta classe é que os retornos sejam não correlacionados serialmente, mas a volatilidade (variância condicional) dependa de retornos passados. A classe de modelos GARCH (Generalized Autoregressive Conditional Heterokedastic) sugerida por Bollerslev (1986, 1987, 1988) pode ser usada para descrever a volatilidade com menos parâmetros que um modelo ARCH. Modelos da classe GARCH são processos estocásticos não lineares, suas distribuições tem cauda pesada com variância condicional dependente do tempo e modelam agrupamentos de volatilidade.

Apesar da razoável descrição, a forma como os modelos acima foram construídos apresentaram algumas limitações no que se refere ao peso das caudas em suas distribuições não condicionais. Muitos estudos em dados financeiros apontam para caudas com peso considerável. Modelos R-GARCH (Randomized Generalized Autoregressive Conditional Heterokedastic) foram propostos por Nowicka (1998) e incluem os modelos ARCH e GARCH possibilitando o uso de inovações estáveis além da conhecida distribuição normal. Estas permitem captar melhor a propriedade de cauda pesada. Como a função de autocovariância não existe para tais processos introduz-se novas medidas de dependência. Neste trabalho apresentamos a inferência indireta como método de estimação da classe R-GARCH. Apresentamos também a classe R_t -GARCH permitindo distribuições t de Student para os resíduos do modelo e a classe R_s -GARCH que estende a classe R-GARCH, permitindo distribuições α -estáveis para os resíduos. Do mesmo modo, propomos o método de inferência indireta para estimar os parâmetros dos modelos R_t -GARCH e R_s -GARCH. Duas aplicações empíricas são realizadas para as três classes de modelos em questão utilizando dados intradiários do índice de ações Ibovespa e das ações da Telemar.

Palavras chave: Distribuições estáveis, R-GARCH, R_t -GARCH, R_s -GARCH, inferência indireta, séries temporais, finanças.

Abstract

Indirect inference of R-GARCH models

Linear processes do not capture the structure of financial data. There is a large variety of nonlinear models available in literature. The class of ARCH models (Autoregressive Conditional Heterokedastic) was introduced by Engle (1982) in order to estimate inflation's variance. The idea is that, in this class, returns are serially uncorrelated, but the volatility (conditional variance) depends on past returns. The class of GARCH models (Generalized Autoregressive Conditional Heterokedastic) suggested by Bollerslev (1986, 1987, 1988) can be used to describe the volatility with less parameters than ARCH-type models. GARCH-type models are nonlinear stochastic processes, their distribution are heavy-tailed with time-dependent conditional variance model and they model clustering of volatility.

Despite the reasonable description, the way that GARCH models are built imposes limits on the heaviness of the tails of their unconditional distribution. Many studies in financial data point to considerable heaviness of the tails. The class of Randomized Generalized Autoregressive Conditional Heterokedastic (R-GARCH) were proposed by Nowicka (1998) and include the ARCH and GARCH models allowing the use of stable innovations in place of normal distribution. This distribution allows to capture the heaviness tail property. As the autocovariance function does not exist for these processes a new measure of dependence was introduced. Estimation methods and empirical analysis of R-GARCH class, as well as their measures of dependence are not available in literature and are the focus of this work. This paper presents the indirect inference method to estimate the R-GARCH class. We also present the class R_t -GARCH allowing the Student t distributions for the residuals of the model and the class R_s -GARCH which extends the class R-GARCH, allowing distributions α -stable to waste. Similarly, we propose the method of indirect inference to estimate the parameters of the models R_t -GARCH and R_s -GARCH. Two empirical applications are held for three classes of models in question using intraday data of the Bovespa stock index and the shares of Telemar.

Key words: Stable distributions, R-GARCH, R_t -GARCH, R_s -GARCH, indirect inference, time series, finance.

Sumário

Agradecimentos	i
Resumo	iii
Abstract	v
Lista de Abreviaturas	xi
Lista de Notações	xiii
Lista de Figuras	xv
Lista de Tabelas	xix
Introdução	xxi
1 Distribuições estáveis	1
1.1 O teorema limite central	1
1.1.1 Fenômenos com variância infinita	2
1.2 Estabilidade	3
1.2.1 Definições e caracterização das constantes normalizadoras	4
1.2.2 Função Característica	9
1.2.3 Significado dos parâmetros	10
1.2.4 Distribuições simétricas e estabilidade estrita	12
1.2.5 Momentos e outras propriedades	13
1.2.6 Combinação linear e padronização	14
1.2.7 Função densidade e função distribuição	14

1.2.8	Simulação de uma variável aleatória α -estável	16
1.3	Estimação e inferência	17
1.3.1	Métodos baseados em quantis	17
1.3.2	Métodos baseados na função característica	18
1.3.3	Máxima verossimilhança	20
1.3.4	Métodos baseados em simulação	21
1.4	Testes para o comportamento estável	21
1.5	Processos de Lévy α -estáveis	23
1.5.1	Simulação de um processo de Lévy α -estável	25
2	Processos ARMA e GARCH	27
2.1	Processos ARMA	27
2.2	Processos GARCH	28
3	Estimação indireta	31
3.1	Propriedades assintóticas	33
3.2	Trabalhando com restrições	35
3.3	Inferência indireta para distribuições e processos α -estáveis	35
4	Processos R-GARCH	37
4.1	Processos R-GARCH com inovações estáveis	39
4.2	Estrutura de dependência	43
4.3	Inferência indireta para o modelo R-GARCH(1,1,0)	60
4.3.1	O modelo auxiliar	60
4.3.2	Simulações	61
4.3.3	Aplicação empírica (TLM)	69
4.3.4	Aplicação empírica (IBV)	70
5	Processos R_t-GARCH	73
5.1	Considerações e exemplos	73
5.2	Inferência indireta para o modelo R_t -GARCH(1,1,0)	74
5.2.1	Simulações	75
5.2.2	Aplicação empírica (TLM)	75
5.2.3	Aplicação empírica (IBV)	76

6	Processos R_s-GARCH	79
6.1	Processos R_s -GARCH com inovações estáveis	80
6.2	Estrutura de dependência	84
6.3	Inferência indireta para o modelo R_s -GARCH(1,1,0)	95
6.3.1	Simulações	95
6.3.2	Aplicação empírica (TLM)	96
6.3.3	Aplicação empírica (IBV)	97
7	Conclusões e trabalhos futuros	99
	Referências Bibliográficas	101

Lista de Abreviaturas

i.i.d.	Independentes e identicamente distribuídas
MC	Monte Carlo
q.c.	Quase certamente
RB	Ruído branco
v.a.	Variável aleatória

Lista de Notações

$\mathcal{P}$	Probabilidade
$f(x; \theta)$	Função densidade de probabilidade
$F(x; \theta)$	Função de distribuição
$\Phi_X(t)$	Função característica da v.a. X
$\mathbf{E}$	Valor esperado
Var	Variância
Cov	Covariância
$I_k(\xi_1, \xi_2)$	Função de autocovariação
$S_\alpha(\sigma, \beta, \mu)$	Distribuição α -estável com parâmetros α , β , σ e μ
$\text{S}\alpha\text{S}$	Distribuição simétrica α -estável
Z_α	Processo de Lévy α -estável
i	Número imaginário: $i = \sqrt{-1}$
$\mathcal{F}$	Sigma álgebra
$\mathcal{L}$	Função de verossimilhança
$\mathcal{N}$	Distribuição normal
$\sim$	Distribuído segundo
$\stackrel{d}{=}$	Igualdade em distribuição
$\stackrel{d}{\Rightarrow}$	Convergência em distribuição
$\mathbb{N}$	Conjunto dos números naturais
$\mathbb{R}$	Conjunto dos números reais
$\mathbb{Z}$	Conjunto dos números inteiros

Lista de Figuras

1.1	Histograma e logretornos de IBV15m	2
1.2	Variograma convergente e não convergente	3
1.3	Variograma dos logretornos de IBV15m	3
1.4	Índice de estabilidade e Parâmetro de assimetria	10
1.5	Parâmetro de escala e Parâmetro de posição	10
1.6	Função distribuição com diferentes valores para os parâmetros.	16
1.7	Processos de Lévy α -estáveis com diferentes valores para α	26
4.1	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.9$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$	49
4.2	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$	50
4.3	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.5$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$	51
4.4	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.3$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$	52
4.5	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.1$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$	53
4.6	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.1$ e $\phi_1 = 0.6$	54
4.7	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.01$ e $\phi_1 = 0.6$	55
4.8	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.001$ e $\phi_1 = 0.6$	56
4.9	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.3$ e $\phi_1 = 0.5$	57

4.10	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.3$ e $\phi_1 = 0.4$	58
4.11	Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.3$ e $\phi_1 = 0.2$	59
4.12	Histograma e densidade ajustada das estimativas do parâmetro α com 100 réplicas e diferentes valores para T.	62
4.13	Histograma e densidade ajustada das estimativas do parâmetro θ_1 com 100 réplicas e diferentes valores para T.	63
4.14	Histograma e densidade ajustada das estimativas do parâmetro ϕ_1 com 100 réplicas e diferentes valores para T.	64
4.15	Histograma e densidade ajustada das estimativas do parâmetro α com 500 réplicas e diferentes valores para T.	66
4.16	Histograma e densidade ajustada das estimativas do parâmetro θ_1 com 500 réplicas e diferentes valores para T.	67
4.17	Histograma e densidade ajustada das estimativas do parâmetro ϕ_1 com 500 réplicas e diferentes valores para T.	68
4.18	Logretornos e variograma de TLM15m	69
4.19	QxQ plot dos resíduos do modelo R-GARCH(1,1,0) ajustado a TLM15m versus quantis da Normal Padrão.	70
4.20	Logretornos e variograma de IBV15m	70
4.21	QxQ plot dos resíduos do modelo R-GARCH(1,1,0) ajustado a IBV15m versus quantis da Normal Padrão.	71
5.1	Retornos e volatilidade de um processo R_t -GARCH(1,1,0) para $\alpha = 1.8$, $\theta_1 = 0.001$ e $\phi_1 = 0.7$	74
5.2	Retornos e volatilidade de um processo R_t -GARCH(1,1,0) para $\alpha = 1.7$, $\theta_1 = 0.0001$ e $\phi_1 = 0.8$	74
5.3	QxQ plot dos resíduos do modelo R_t -GARCH(1,1,0) ajustado a TLM15m versus quantis de um t de Student com quatro graus de liberdade.	76
5.4	QxQ plot dos resíduos do modelo R_t -GARCH(1,1,0) ajustado a IBV15m versus quantis de um t de Student com três graus de liberdade.	77
6.1	Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.9$, $\theta_1 = 0.0000001$, $\phi_1 = 0.8$, $\lambda = 1.8$ e $\sigma = 0.5$	89
6.2	Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.0000001$, $\phi_1 = 0.8$, $\lambda = 1.8$ e $\sigma = 0.5$	90
6.3	Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.0000001$, $\phi_1 = 0.5$, $\lambda = 1.8$ e $\sigma = 0.5$	91

6.4	Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.0001$, $\phi_1 = 0.8$, $\lambda = 1.8$ e $\sigma = 0.5$	92
6.5	Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.0000001$, $\phi_1 = 0.8$, $\lambda = 1.3$ e $\sigma = 0.5$	93
6.6	Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariação teórica para $\alpha = 1.7$, $\theta_1 = 0.0000001$, $\phi_1 = 0.8$, $\lambda = 1.8$ e $\sigma = 2.0$	94
6.7	Densidades das distribuições normal padrão, t de Student e estável.	96
6.8	QxQ plot dos resíduos do modelo R_s -GARCH(1,1,0) ajustado a TLM15m versus quantis de uma $S_{1.8}(0.5, 1, 0)$	97
6.9	QxQ plot dos resíduos do modelo R_s -GARCH(1,1,0) ajustado a IBV15m versus quantis de uma estável $S_{1.8}(0.5, 0, 0)$	98

Lista de Tabelas

4.1	Média e erro padrão MC para a estimação de um R-GARCH(1,1,0) com 100 réplicas T observações.	61
4.2	Média e erro padrão MC para a estimação de um R-GARCH(1,1,0) com 500 réplicas e T observações.	65
4.3	Média e erro padrão MC para a estimação de um R-GARCH(1,1,0) com 100 réplicas e 100000 observações.	65
4.4	Média e erro padrão MC ao ajustarmos um R-GARCH(1,1,0) aos logretornos de TLM15m.	69
4.5	Erro quadrático médio ao ajustarmos um modelo GARCH(0,1) e R-GARCH(1,1,0) aos logretornos de TLM15m	69
4.6	Média e erro padrão MC ao ajustarmos um R-GARCH(1,1,0) aos logretornos de IBV15m.	71
4.7	Erro quadrático médio ao ajustarmos um modelo GARCH(0,1) e R-GARCH(1,1,0) aos logretornos de IBV15m	71
5.1	Média e erro padrão MC para a estimação de um R_t -GARCH(1,1,0) com 100 réplicas e 100000 observações.	75
5.2	Média e erro padrão MC ao ajustarmos um R_t -GARCH(1,1,0) aos logretornos de TLM15m.	75
5.3	Erro quadrático médio ao ajustarmos os modelos GARCH(0,1), R-GARCH(1,1,0) e R_t -GARCH(1,1,0) aos logretornos de TLM15m	76
5.4	Média e erro padrão MC ao ajustarmos um R_t -GARCH(1,1,0) aos logretornos de IBV15m.	77
5.5	Erro quadrático médio ao ajustarmos os modelos GARCH(0,1), R-GARCH(1,1,0) e R_t -GARCH(1,1,0) aos logretornos de IBV15m	77
6.1	Média e erro padrão MC para a estimação de um R_s -GARCH(1,1,0) com 100 réplicas e 100000 observações.	96
6.2	Média e erro padrão MC ao ajustarmos um R_s -GARCH(1,1,0) aos logretornos de TLM15m.	97
6.3	Erro quadrático médio ao ajustarmos os modelos GARCH(0,1), R-GARCH(1,1,0), R_t -GARCH(1,1,0) e R_s -GARCH(1,1,0) aos logretornos de TLM15m	97

- 6.4 Média e erro padrão MC ao ajustarmos um R_s -GARCH(1,1,0) aos logretornos de IBV15m. 98
- 6.5 Erro quadrático médio ao ajustarmos os modelo GARCH(0,1), R-GARCH(1,1,0), R_t -GARCH(1,1,0) e R_s -GARCH(1,1,0) aos logretornos de IBV15m 98

Introdução

Neste capítulo iremos apresentar a relação entre os sucessivos capítulos presentes no trabalho e os resultados obtidos visando antecipar as principais idéias que induziram a obtenção e evolução dos mesmos. O link entre os capítulos se concentra no uso de metodologias computacionais intensivas para a estimação de modelos estatísticos baseados nas distribuições α -estáveis.

Os capítulos foram escritos de maneira que possam ser lidos de forma autônoma; algumas repetições de conceitos são, por vezes, inevitáveis para o melhor acompanhamento. Em particular, os conceitos fundamentais relativos às distribuições α -estáveis são apresentados no Capítulo 1. Este capítulo é introdutório, de modo que não há resultados originais em seu corpo. Começamos o capítulo motivando o natural surgimento das distribuições estáveis em função do Teorema Limite Central e fenômenos com variância infinita. Em seguida, apresentamos as definições e principais propriedades dessa classe de distribuições. Terminamos o Capítulo 1 discutindo a simulação, inferência e testes para as distribuições estáveis. No Capítulo 2, apenas relembramos a definição e alguns resultados dos modelos ARCH e GARCH que são motivadores e compartilham algumas propriedades com a classe de modelos R-GARCH que será apresentada no Capítulo 4.

As distribuições estáveis são excelentes para a descrição de alguns fenômenos por permitirem caudas pesadas, assimetria, curtose elevada e generalizar a classe de distribuições Gaussianas. Apesar de todos os prós, o fato de não possuir formas fechadas para a densidade tem dificultado o seu emprego. Não teremos assim forma fechada para modelos que envolvam as distribuições estáveis, pois também não teremos forma fechada para a função de verossimilhança. Felizmente, com o crescente avanço tecnológico e computacional, metodologias computacionais intensivas para estimação de modelos estatísticos estão sendo viabilizadas e tem apresentado excelentes resultados. Seguindo essa perspectiva, apresentamos no Capítulo 3, o método de inferência indireta de Gouriéroux et. al. (1993) e principais resultados. O método de inferência indireta é aplicável em situações onde não possuímos forma fechada para a função de verossimilhança do modelo original, mas podemos simular as distribuições facilmente. Utilizamos então, um modelo auxiliar que possua função de verossimilhança com forma fechada e obtemos as estimativas para os parâmetros do modelo original via simulações e um processo iterativo que visa aproximar as distribuições ou processos simulados com seus respectivos correspondentes aos parâmetros estimados com o modelo auxiliar. Sendo as distribuições α -estáveis e os processos descritos na tese de fácil simulação iremos propor um modelo auxiliar para os mesmos e investigar o emprego da estimação por inferência indireta.

No Capítulo 4 definimos a classe de modelos R-GARCH (Randomized GARCH) que generaliza a classe GARCH ao tornarmos a volatilidade também aleatória, substituindo o termo constante por um fator com distribuição positiva. Em particular, uma teoria muito promissora surge ao escolhermos as ditribuições α -estáveis para este fator. Essa classe foi proposta por Nowicka (1998) e compartilha

alguns resultados importantes da classe de modelos GARCH. Os processos R-GARCH com inovações α -estáveis são estacionários, possuem distribuição α -estável simétrica assim como seus retornos agregados. Como as distribuições estáveis não possuem segundo momento, exceto no caso normal, a função de autocovariância não existe para tais processos e se faz necessário uma nova medida de dependência. Nowicka (1998) sugere a função de codiferença como alternativa e mostra que para um caso particular dos processos R-GARCH (os processos R-GARCH(1,1,0)), ela é limitada com decaimento exponencial. Visando possibilitar a estimação de uma função que avalie a dependência de processos com distribuição α -estável, nós utilizamos a função de autocovariância de Kokoszka e Taqqu (1994) que generaliza a função codiferença. Mostramos alguns resultados relevantes no que concerne à dependência e generalizamos o resultado de Nowicka (1998), ao demonstrarmos que para os processos R-GARCH(1,1,0) a função de autocovariância também é limitada com decaimento exponencial.

Ainda no Capítulo 4 apresentamos um dos principais resultados obtidos na tese. Apresentamos o método de inferência indireta como meio para estimar os parâmetros de um processo R-GARCH(1,1,0) e, para tal, propomos o modelo GARCH(1,1) com distribuições t de Student como modelo auxiliar. As propriedades dos estimadores são verificados por meio de um estudo detalhado de simulações. Seguindo o capítulo, duas aplicações empíricas são realizadas para a classes de modelo R-GARCH(1,1,0) utilizando dados intradiários do índice de ações Ibovespa e das ações da Telemar.

Os processos R-GARCH estimados nas aplicações empíricas apresentam desempenho muito bom, no entanto a hipótese de normalidade para os resíduos se mostra bastante inadequado quando fazemos um QxQ plot. Com a intenção de melhorar essa questão residual, propomos, no Capítulo 5, a classe R_t -GARCH que permite distribuições t de Student para os resíduos do modelo. Do mesmo modo, propomos a inferência indireta como método de estimação para os seus parâmetros utilizando o mesmo modelo auxiliar. Estudamos o caso particular R_t -GARCH(1,1,0) e avaliamos as propriedades dos estimadores por meio de um estudo detalhado de simulações. Com a idéia de comparar o desempenho dos processos R-GARCH(1,1,0) e R_t -GARCH(1,1,0), também realizamos duas aplicações empíricas utilizando dados intradiários do índice de ações Ibovespa e das ações da Telemar.

Os processos R_t -GARCH, apresentam um melhor desempenho que os processos R-GARCH assim como melhor adequação em relação a hipótese sobre os resíduos. Infelizmente a falta de uma teoria para a questão de autodependência impossibilita essa importante análise. Pensando em solucionar esta questão, propomos no Capítulo 6, a classe de modelos R_s -GARCH atribuindo também aos resíduos a hipótese de distribuições α -estáveis. Como primeiro ganho temos o fato de que os processos R_s -GARCH generalizam os processos R-GARCH. No desenrolar deste capítulo provamos, para os processos R_s -GARCH, resultados semelhantes aos existentes para a classe de modelos R-GARCH onde todos eles generalizam os resultados em questão. Por exemplo, utilizamos novamente a função de autocovariância de Kokoszka e Taqqu (1994) e demonstramos que para os processos R_s -GARCH(1,1,0) a função de autocovariância também é limitada com decaimento exponencial. Novamente propomos a inferência indireta como método de estimação para os seus parâmetros utilizando o mesmo modelo auxiliar. Estudamos o caso particular R_s -GARCH(1,1,0) e avaliamos as propriedades dos estimadores por meio de um estudo detalhado de simulações. Com a idéia de comparar o desempenho dos processos R-GARCH(1,1,0), R_t -GARCH(1,1,0) e R_s -GARCH(1,1,0) também realizamos duas aplicações empíricas utilizando dados intradiários do índice de ações Ibovespa e das ações da Telemar.

Capítulo 1

Distribuições estáveis

Distribuições e processos gaussianos são há muito tempo estudados e sua utilidade em modelos estocásticos está bem estabelecida. Análises de finanças e seguros frequentemente indicam a presença de caudas pesadas de modo que os processos estáveis se mostram como um recurso apropriado para modelos probabilísticos de cauda pesada. Neste capítulo apresentaremos a teoria inicial sobre distribuições estáveis e como referências básicas citamos Breiman (1968), Feller (1966) e Samorodnitsky e Taqqu (1994).

1.1 O teorema limite central

O teorema limite central é um dos pilares na inferência estatística. Na sua formulação mais simples fica estabelecido que, dado uma sequência de n variáveis aleatórias independentes e identicamente distribuídas, seu somatório converge, à medida que n cresce, para uma distribuição normal. Este teorema é de suma importância por duas razões:

- Muitas estatísticas envolvem somas de variáveis aleatórias que se referem aos indivíduos na amostra.
- Vários fenômenos de interesse estatístico podem ser pensados como agregações de contribuições de fatores menores.

Como resultado, a distribuição normal tem sido amplamente empregada na inferência estatística e modelagem estatística. Abaixo formalizamos uma das versões mais gerais do teorema limite central que mesmo sendo de conhecimento geral é importante lembrá-lo para efeitos de comparações com resultados apresentados nos próximos parágrafos.

Teorema 1.1. *Dada uma sequência $\{X_i\}$, $i = 1, \dots, n$ de variáveis aleatórias i.i.d., onde $0 < \mathbf{E}X_1^2 < \infty$, existem sequências reais $\{c_n\}$, $\{d_n\}$ com $c_n > 0$ tais que:*

$$\frac{S_n - d_n}{c_n} \xrightarrow{d} Z \sim \mathcal{N}(0, 1),$$

onde $S_n = X_1 + \dots + X_n$.

1.1.1 Fenômenos com variância infinita

Em muitos casos empíricos as hipóteses teóricas acima caem por terra e a suposição de que as variáveis aleatórias são gaussianas passam a constituir um equívoco. É possível observar, na prática, muitos casos em que os resíduos mostram comportamentos de caudas muito mais pesadas do que o esperado se considerássemos que os mesmos possuem distribuição normal. Isso mostra que uma de nossas hipóteses, somas de variáveis aleatórias i.i.d. ou variância finita, deve estar errada nesses casos. Seguindo este raciocínio, vamos explorar a segunda hipótese.

Como um primeiro exemplo, considere os logretornos de 15 minutos do índice de ações Ibovespa (IBV15m) no período de 06/04/1998 até 13/08/2003, totalizando 37960 observações. Como podemos

Figura 1.1: Histograma e logretornos de IBV15m

observar na Figura 1.1, o seu histograma apresenta caudas bem mais pesadas que o esperado para um modelo Gaussiano. Este poderia ser um sintoma de que a variância dos componentes é não finita.

Vamos introduzir uma ferramenta descritiva bastante útil para observar a finitude da variância de um vetor aleatório. O *variograma* de uma sequência aleatória i.i.d. é definido como o vetor $\mathbf{V}$ cuja k -ésima entrada é dada por

$$V_k = \sum_{i=1}^k \frac{(X_i - \bar{X}_k)^2}{k-1}, \quad k \geq 2$$

onde $\bar{X}_k$ denota a média amostral das primeiras k observações. Desse modo V_k é a variância amostral dos primeiros k itens de um conjunto de dados.

Se o teorema limite central valer, a variância amostral é um estimador consistente para a variância populacional, logo o variograma, dado que o tamanho da amostra é suficientemente grande, deveria manter um comportamento convergente. Quando isso não ocorre, como há “saltos” causados por observações pertencentes à cauda, o comportamento foge da convergência mostrando um indício de não finitude. Um exemplo de padrão convergente e não convergente para o variograma é mostrado na Figura 1.2; o variograma dos logretornos de IBV15m é mostrado na Figura 1.3.

Figura 1.2: Variograma convergente e não convergente

Figura 1.3: Variograma dos logretornos de IBV15m

1.2 Estabilidade

De posse das idéias apresentadas na seção anterior, é natural a indagação sobre o que deveríamos fazer quando os componentes da variância apresentarem indícios de não finitude. Se voltarmos ao Teorema 1.1 algumas perguntas seriam imediatas, tais quais as apresentadas abaixo:

- Qual seria o comportamento da distribuição do fator $\frac{S_n - d_n}{c_n}$ se relaxássemos a hipótese de finitude da variância?
- Haveria convergência?
- Caso haja a convergência, para qual classe de distribuições convergiria?

Se pensarmos em $\frac{S_n - d_n}{c_n} \stackrel{d}{=} X$, para $c_n > 0$ e d_n real, com variância não necessariamente finita, uma nova classe de distribuições começa a surgir. Começamos a imaginar um tipo de variável aleatória X que, nas condições acima, tenha a seguinte representação para $n \geq 2$:

$$X_1 + \cdots + X_n \stackrel{d}{=} c_n X + d_n. \quad (1.1)$$

De fato, a essa classe de distribuições iremos chamar *distribuições estáveis*. Um dos fatos mais interessantes que constituem a primeira versão de um teorema limite central generalizado (sem hipótese de variância finita) é que as distribuições estáveis são o único caso possível para suas leis limites. O resultado foi provado pela primeira vez em Lévy (1924).

Teorema 1.2 (Lévy). *Dada uma sequência $\{X_i\}$, $i = 1, \dots, n$ de variáveis aleatórias i.i.d. e sequências reais $\{c_n\}$ e $\{d_n\}$ com $c_n > 0$, a condição necessária e suficiente para que $F(x)$ seja a distribuição limite de*

$$\frac{S_n - d_n}{c_n},$$

é que $F(x)$ seja estável.

1.2.1 Definições e caracterização das constantes normalizadoras

Nas seções passadas construímos a motivação para a importância das distribuições estáveis. Daqui em diante apresentaremos as principais definições e propriedades dessa classe promissora. Tomaremos como base a visão ilustrativa de Samorodnitsky e Taqqu (1994).

Definição 1.1. *Dizemos que uma variável aleatória X é estável se, para $n \geq 2$, existem um número positivo real c_n e um número real d_n tais que:*

$$X_1 + \cdots + X_n \stackrel{d}{=} c_n X + d_n,$$

onde $X_1, \dots, X_n$ são cópias independentes de X .

Essencialmente, isto significa que uma variável aleatória é estável se puder ser quebrada em peças idênticas com constantes normalizadoras.

Exemplo 1.1. *Seja Z uma v.a. com distribuição degenerada, isto é, $P(Z = c) = 1$ para alguma constante c , e $X_1, X_2, \dots$ cópias independentes de Z . Então, as funções características satisfazem:*

$$\Phi_Z(t) = \mathbf{E}(e^{itc}) = e^{itc},$$

$$\Phi_{S_n}(t) = e^{itnc} \implies \Phi_{\frac{S_n}{n}}(t) = \Phi_{S_n}(t/n) = e^{itc}.$$

Portanto, $\frac{S_n}{n} \stackrel{d}{=} Z$ ou ainda $S_n \stackrel{d}{=} nZ$.

Exemplo 1.2. *Seja Z uma v.a. com distribuição $\mathcal{N}(0, 1)$ e $X_1, X_2, \dots$ cópias independentes de Z . Sabemos que $\Phi_Z(t) = e^{-t^2/2}$ de forma que*

$$\Phi_{\frac{S_n}{\sqrt{n}}}(t) = \Phi_{S_n}(t/\sqrt{n}) = \left[e^{-\frac{(t/\sqrt{n})^2}{2}} \right]^n = e^{-t^2/2}.$$

Portanto $\frac{S_n}{\sqrt{n}} \stackrel{d}{=} Z$ ou ainda $S_n \stackrel{d}{=} \sqrt{n}Z$. Se considerarmos $Z \sim \mathcal{N}(\mu, \sigma^2)$, pode-se realizar cálculos análogos para obter $S_n \stackrel{d}{=} \sqrt{n}Z + (n\mu - \sqrt{n}\mu)$.

Se $d_n = 0$ na Definição 1.1, dizemos que a distribuição é estritamente estável. Apresentaremos adiante outras duas definições equivalentes à Definição 1.1.

Definição 1.2. Dizemos que X tem distribuição estável se dados $a > 0$, $b > 0$ e X_1, X_2 cópias independentes de X , existem $c > 0$ e $d \in \mathbb{R}$ tais que

$$aX_1 + bX_2 \stackrel{d}{=} cX + d. \quad (1.2)$$

Se $d = 0$, dizemos que X tem distribuição estritamente estável.

Definição 1.3. Dada uma variável aleatória X com função distribuição F , definimos o domínio de atração de F como sendo o conjunto das distribuições G tais que se $X_i \stackrel{d}{=} G$, para $i \geq 1$, existem constantes normalizadoras c_n e d_n tais que

$$\frac{S_n - d_n}{c_n} \xrightarrow{d} X.$$

Definição 1.4. Dizemos que uma variável aleatória X tem distribuição estável se seu domínio de atração é não vazio.

Pode-se mostrar que as Definições 1.1, 1.2 e 1.4 são todas equivalentes. A demonstração desta propriedade é bastante técnica e pode ser encontrada em Samorodnitsky e Taqqu (1994), motivo pelo qual iremos omiti-la.

Uma outra questão que surge naturalmente é a possibilidade de caracterizarmos as constantes normalizadoras c_n e d_n . De fato, iremos demonstrar no teorema seguinte que isso é possível e teremos uma forma fechada para as constantes normalizadoras.

Teorema 1.3. Seja X uma v.a. com distribuição F estável, não degenerada, e sejam $X_1, \dots, X_n$ cópias independentes de X . Então,

- (i) existe $0 < \alpha \leq 2$ único tal que $c_n = n^{1/\alpha}$, isto é, $S_n = X_1 + \dots + X_n \stackrel{d}{=} n^{1/\alpha}X + d_n$,
- (ii) se $\alpha \neq 1$, existe um $b \in \mathbb{R}$ tal que $d_n = -b(n^{1/\alpha} - n)$ e b é tal que $X - b$ possui distribuição estritamente estável.

Dem. Vamos dividir a demonstração do item (i) em algumas etapas:

- (i1) Tome $Y \stackrel{d}{=} X$, X e Y independentes.

Sejam $X_1, \dots, X_n$ cópias independentes de X , e $Y_1, \dots, Y_n$ cópias independentes de Y . Segue assim que

$$X_1 + \dots + X_n \stackrel{d}{=} c_n X + d_n$$

e

$$Y_1 + \dots + Y_n \stackrel{d}{=} c_n Y + d_n.$$

Temos então, pela independência, que

$$(X_1 - Y_1) + \cdots + (X_n - Y_n) \stackrel{d}{=} c_n(X - Y).$$

Como $(X - Y)$ tem distribuição simétrica e mesmo coeficiente c_n podemos simplificar o teorema para o caso simétrico, isto é, X é tal que

$$S_n = X_1 + \cdots + X_n \stackrel{d}{=} c_n X.$$

(i2) Seja $X \stackrel{d}{=} F$, F estável e simétrica. Vamos mostrar que $\{c_n\}$ satisfaz:

$$\begin{aligned} c_{m+n}X &\stackrel{d}{=} c_m X_1 + c_n X_2, \\ c_{mn} &= c_m c_n. \end{aligned}$$

Considere $S_{m+n} = X_1 + \cdots + X_n + X_{n+1} + \cdots + X_{m+n}$, e defina

$$S_n = X_1 + \cdots + X_n$$

e

$$S_{m+n} - S_n = X_{n+1} + \cdots + X_{m+n}.$$

Temos então, pela independência

$$\begin{aligned} S_n &\stackrel{d}{=} c_n X \stackrel{d}{=} c_n X_2, \\ S_{m+n} - S_n &\stackrel{d}{=} c_m X \stackrel{d}{=} c_m X_1. \end{aligned}$$

Portanto

$$S_{m+n} \stackrel{d}{=} c_m X_1 + c_n X_2 \stackrel{d}{=} c_{m+n} X.$$

Resumindo,

$$c_{m+n}X \stackrel{d}{=} c_m X_1 + c_n X_2.$$

Se $m = n$,

$$c_{2n}X \stackrel{d}{=} c_n X_1 + c_n X_2 = c_n(X_1 + X_2) = c_n c_2 X.$$

Logo $c_{2n} = c_2 c_n$.

Mostra-se facilmente, por indução, que $c_{mn} = c_m c_n$.

(i3) Agora um candidato que se torna natural é $c_n = n^{1/\alpha}$. Mais que isso, $0 < \alpha \leq 2$ é único.

Dado $k \in \mathbb{N}$, seja $c_k = k^{1/\alpha}$, onde $\alpha = \frac{\log k}{\log c_k}$.

Se $n = k^l$, $c_n = c_{k^l} = c_k^l = (k^{1/\alpha})^l = (k^l)^{1/\alpha} = n^{1/\alpha}$.

Se $v = m^i$ e $v \neq k^l$, $\forall l$, então $c_m = m^{1/\beta}$ e $c_v = v^{1/\beta}$.

Devemos mostrar que $\alpha = \beta$.

Seja j tal que $k^j < m^i < k^{j+1}$. Então,

$$k^{j/\beta} < (m^i)^{1/\beta} < k^{1/\beta} k^{j/\beta}$$

e

$$c_{kj}^{\alpha/\beta} < c_v < k^{1/\beta} c_{kj}^{\alpha/\beta}. \quad (1.3)$$

Para isto, observe que $k^{j/\beta} = [(k^j)^{1/\alpha}]^{\alpha/\beta}$.

Segue da equação (1.3) que

$$1 < k^{1/\beta} \frac{c_{kj}}{c_v} c_{kj}^{(\alpha/\beta-1)}.$$

Como $c_{kj} \rightarrow \infty$, segue que $\alpha \geq \beta$. Invertendo os papéis de m e k , obtemos $\alpha \leq \beta$ e, portanto, $\alpha = \beta$.

(i4) Vamos mostrar agora que $\alpha > 0$.

$\alpha = 0$ implica claramente uma indeterminação. Suponha $\alpha < 0$, dessa forma como $c_m > 0$, teremos a igualdade abaixo

$$\mathcal{P}\{|X| > z\} = \mathcal{P}\{|c_m X| > c_m z\} = \mathcal{P}\{|S_m| > c_m z\}. \quad (1.4)$$

Mas o evento $\{|S_m| > c_m z\}$ contém o evento abaixo:

$$\left\{ \left(\max_{1 \leq j \leq m} |X_j| > c_m z, S_m - \max_{1 \leq j \leq m} |X_j| \geq 0 \right) \cup \left(\max_{1 \leq j \leq m} |X_j| < -c_m z, S_m - \max_{1 \leq j \leq m} |X_j| \leq 0 \right) \right\}.$$

Logo, pela equação (1.4), temos

$$\mathcal{P}\{|X| > z\} \geq \frac{1}{2} \mathcal{P} \left\{ \max_{1 \leq j \leq m} |X_j| > c_m z \right\} \geq \frac{1}{2} \mathcal{P}\{|X| > c_m z\}.$$

Como $\alpha < 0$, $c_m = m^{1/\alpha} \rightarrow 0$, quando $m \rightarrow \infty$.

Assim, temos que $\mathcal{P}\{|X| > z\} \geq \frac{1}{2}$, $\forall z > 0$, o que é uma contradição.

(i5) Resta mostrar apenas que $\alpha \leq 2$.

Suponha que $\alpha > 2$,

$$\begin{aligned} \mathcal{P}\{|S_n| > c_n z\} &\leq \frac{1}{2} \mathcal{P} \left\{ \max_{1 \leq j \leq n} |X_j| > c_n z \right\} \\ &= \frac{1}{2} (1 - \mathcal{P}^n\{|X| \leq c_n z\}) \\ &\geq \frac{1}{2} \left(1 - e^{-n(1 - \mathcal{P}\{|X| \leq c_n z\})} \right). \end{aligned}$$

Onde a última desigualdade segue do resultado conhecido,

$$(1 - x^n) \geq 1 - e^{-n(1-x)}, \quad 0 < x < 1.$$

Podemos escolher um $z \in \mathbb{R}$ tal que $\mathcal{P}\{|S_n| > c_n z\} < 1/4, \forall n$. Segue assim que $n(1 - \mathcal{P}\{|X| \leq c_n z\})$ permanece limitado, ou seja, $\exists M < \infty$ tal que $n(1 - \mathcal{P}\{|X| \leq c_n z\}) \leq M$.

Em particular, se $n = m^\alpha$ e $z = 1$, então $m^\alpha(1 - \mathcal{P}\{|X| \leq c_n\}) \leq M < \infty$ e, portanto, $\mathbf{E}X^2 < \infty$, o que é uma contradição (ver lema 2 em Feller (1966)).

(ii)

$$\begin{cases} S_n \stackrel{d}{=} c_n X_1 + d_n, \\ S_{2n} - S_n \stackrel{d}{=} c_n X_2 + d_n, \\ \vdots \\ S_{mn} - S_{(m-1)n} \stackrel{d}{=} c_n X_m + d_n. \end{cases}$$

Segue por independência que

$$S_{mn} \stackrel{d}{=} c_n(X_1 + \cdots + X_m) + md_n \stackrel{d}{=} c_n c_m X + c_n d_m + md_n.$$

Analogamente, obtemos $S_{mn} \stackrel{d}{=} c_m c_n X + c_m d_n + nd_m$, de forma que $(c_n - n)d_m = (c_m - m)d_n$.

Sendo assim, basta tomarmos $d_n = -b(n^{1/\alpha} - n)$.

Como X é estável, é fácil mostrar que $X - b$ é estritamente estável. De fato, sabendo que $c_n = n^{1/\alpha}$ e $d_n = -b(n^{1/\alpha} - n)$ para algum $b \in \mathbb{R}$, então

$$c_n X + d_n \stackrel{d}{=} X_1 + \cdots + X_n$$

ou

$$\begin{aligned} n^{1/\alpha} X - n^{1/\alpha} b + nb &\stackrel{d}{=} X_1 + \cdots + X_n \stackrel{d}{=} X_1 + \cdots + X_n - nb + nb \\ \implies n^{1/\alpha}(X - b) + nb &\stackrel{d}{=} (X_1 - b) + \cdots + (X_n - b) + nb \\ \implies n^{1/\alpha}(X - b) &\stackrel{d}{=} (X_1 - b) + \cdots + (X_n - b). \end{aligned}$$

■

Exemplo 1.3. Para ilustrar um caso em que as constantes normalizadoras c_n e d_n serão conhecidas é razoável decidirmos pelo caso em que temos uma distribuição normal. Retornemos então ao Exemplo 1.2. Como vimos, se considerarmos uma variável aleatória $X \sim \mathcal{N}(\mu, \sigma^2)$, então S_n terá forma conhecida de modo que $c_n = n^{1/2}$ e $d_n = -\mu(\sqrt{n} - n)$.

1.2.2 Função Característica

A maneira mais simples de caracterizarmos as distribuições estáveis é por meio de suas funções características, cuja expressão genérica é dada no teorema abaixo. É interessante notar que, como o teorema funciona nas duas direções, ele nos dá mais uma alternativa para caracterizar as distribuições estáveis. Daqui em diante iremos nos referir a uma variável aleatória com distribuição estável simplesmente chamando-a de estável.

Teorema 1.4 (Lévy-Khintchine). *A função característica de uma variável aleatória estável tem a forma*

$$\ln \Phi_X(t) = \begin{cases} it\mu - \sigma^\alpha |t|^\alpha \left[1 - i\beta \operatorname{sinal}(t) \tan\left(\frac{\pi\alpha}{2}\right) \right], & \text{se } \alpha \neq 1, \\ it\mu - \sigma |t| [1 + i\beta \operatorname{sinal}(t) \ln(t)], & \text{se } \alpha = 1, \end{cases} \quad (1.5)$$

onde $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\sigma > 0$ e $\mu \in \mathbb{R}$.

A função $\operatorname{sinal}(t)$ é definida por

$$\operatorname{sinal}(t) = \begin{cases} 1 & \text{se } t > 0, \\ 0 & \text{se } t = 0, \\ -1 & \text{se } t < 0. \end{cases}$$

O teorema acima nos permite visualizar uma variável aleatória estável por meio de seus parâmetros na função característica. Desse modo, podemos nos referir a uma v.a. com distribuição estável apenas chamando-a *v.a. α -estável* e introduzimos uma notação bastante simples com informação preciosa sobre seus parâmetros. Como veremos adiante, cada um desses parâmetros carrega consigo um significado prático. Sendo assim, de agora em diante, simbolizaremos uma v.a. α -estável com o uso da notação $X \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu)$.

Exemplo 1.4 (Distribuição Gaussiana). *Quando $\alpha = 2$, a função característica se reduz a*

$$\Phi(t) = \exp\{i\mu t - \sigma^2 t^2\},$$

que coincide com a função característica de uma normal de média μ e variância $2\sigma^2$.

Exemplo 1.5 (Distribuição de Cauchy). *Quando fazemos $\alpha = 1$ e $\beta = 0$, a função característica se reduz a*

$$\Phi(t) = \exp\{i\mu t - \sigma |t|\},$$

que coincide com a função característica da distribuição de Cauchy.

Exemplo 1.6 (Distribuição de Lévy). *Quando fazemos $\alpha = 1/2$ e $\beta = \pm 1$, a função característica se reduz a*

$$\Phi(t) = \exp\{i\mu t - \sqrt{\sigma |t|} [1 - i \operatorname{sinal}(t)]\},$$

que coincide com a função característica da distribuição de Lévy com densidade

$$f(x; \mu, \sigma) = \sqrt{\frac{\sigma/2\pi}{(x - \mu)^3}} e^{-\frac{\sigma}{2(x - \mu)}}, \quad \forall x > \mu.$$

1.2.3 Significado dos parâmetros

Cada parâmetro na função característica tem um significado prático. O parâmetro $\alpha \in (0, 2]$ é conhecido como “Índice de estabilidade” e $\beta \in [-1, 1]$ como “Parâmetro de assimetria”. A Figura 1.4 abaixo justifica os nomes. Em particular, como visto no Exemplo 1.4, se $\alpha = 2$ então a v.a. estável tem distribuição $\mathcal{N}(\mu, 2\sigma^2)$.

Figura 1.4: Índice de estabilidade e Parâmetro de assimetria

O parâmetro $\sigma \in (0, +\infty)$ é conhecido como “Parâmetro de escala” e $\mu \in \mathbb{R}$ como “Parâmetro de posição”. A Figura 1.5 abaixo justifica os nomes. Veremos que se $1 < \alpha < 2$ então a média da distribuição é μ , caso contrário a distribuição não possui primeiro momento.

Figura 1.5: Parâmetro de escala e Parâmetro de posição

Após a motivação dada pelas figuras acima, segue nas propriedades abaixo a justificativa teórica para os termos utilizados para identificar os parâmetros.

Propriedade 1.1. *Seja $X \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu)$ e seja a uma constante real. Então $X + a \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu + a)$.*

A demonstração da Propriedade 1.1 segue trivialmente da função característica. A razão para nos referir ao parâmetro μ como parâmetro de posição fica agora bastante clara.

Propriedade 1.2. *Seja $X \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu)$ e a uma constante real não nula. Então*

$$\begin{aligned} aX &\stackrel{d}{=} S_\alpha(|a|\sigma, \text{ sinal}(a)\beta, a\mu), & \text{se } \alpha \neq 1, \\ aX &\stackrel{d}{=} S_1\left(|a|\sigma, \text{ sinal}(a)\beta, a\mu - \frac{2}{\pi}a(\ln|a|)\sigma\beta\right), & \text{se } \alpha = 1. \end{aligned}$$

Dem. Pela função característica, temos para $\alpha \neq 1$

$$\begin{aligned} \ln[\mathbf{E}(\exp(itaX))] &= i\mu(ta) - |ta|^\alpha \sigma^\alpha \left(1 - i\beta \text{ sinal}(at) \tan\left(\frac{\pi\alpha}{2}\right)\right) \\ &= i(\mu a)t - (\sigma|a|)^\alpha |t|^\alpha \left(1 - i\beta \text{ sinal}(a) \text{ sinal}(t) \tan\left(\frac{\pi\alpha}{2}\right)\right). \end{aligned}$$

A demonstração para $\alpha = 1$ é análoga. ■

O parâmetro σ é então chamado parâmetro de escala. Observe que quando $\alpha = 1$, a multiplicação por uma constante afeta a posição de um modo não linear. Então, o termo “parâmetro de escala” para σ passa a não fazer sentido quando $\alpha = 1$ e $\beta = 0$. Quando $\mu = 0$ temos a propriedade abaixo.

Propriedade 1.3. *Para todo $0 < \alpha < 2$,*

$$X \stackrel{d}{=} S_\alpha(\sigma, \beta, 0) \iff -X \stackrel{d}{=} S_\alpha(\sigma, -\beta, 0).$$

A demonstração da Propriedade 1.3 segue trivialmente da função característica.

Dizemos que X tem distribuição simétrica se $X \stackrel{d}{=} -X$ e temos dessa maneira que, quando $\beta = 0$ a distribuição é simétrica. Quando $\beta > 0$ a distribuição é assimétrica à direita enquanto se $\beta < 0$ é assimétrica à esquerda. O caso $\beta = \pm 1$ corresponde à assimetria perfeita à direita ou esquerda de acordo com o sinal.

Pelo resultado seguinte iremos identificar α como o parâmetro de cauda: à medida que decresce, a cauda tende a ficar mais pesada seguindo o comportamento como uma potência de α .

Propriedade 1.4. *Seja $X \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu)$ com $0 < \alpha < 2$. Então*

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \lambda^\alpha \mathcal{P}\{X > \lambda\} &= C_\alpha \frac{1 + \beta}{2} \sigma^\alpha, \\ \lim_{\lambda \rightarrow \infty} \lambda^\alpha \mathcal{P}\{X < -\lambda\} &= C_\alpha \frac{1 - \beta}{2} \sigma^\alpha, \end{aligned}$$

onde

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \text{sen}(x) dx \right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha) \cos(\pi\alpha/2)}, & \text{se } \alpha \neq 1, \\ \frac{2}{\pi}, & \text{se } \alpha = 1. \end{cases}$$

1.2.4 Distribuições simétricas e estabilidade estrita

Aqui veremos algumas propriedades úteis das distribuições simétricas e outras de distribuições com estabilidade estrita. Como ilustração faremos a demonstração de apenas algumas propriedades.

Propriedade 1.5. *Seja $X \stackrel{d}{=} F$, então F é simétrica se, e somente se, $\Phi_X(t)$ é real.*

Dem. Temos que

$$\begin{aligned}\Phi_X(t) &= \int \cos(tx)dF(x) + i \int \sin(tx)dF(x), \\ \Phi_{-X}(t) &= \int \cos(tx)dF(x) - i \int \sin(tx)dF(x),\end{aligned}$$

então

$$\Phi_X(t) + \Phi_{-X}(t) = 2 \int \cos(tx)dF(x). \quad (1.6)$$

Como $X \stackrel{d}{=} -X$, temos $\Phi_X(t) = \Phi_{-X}(t)$. Segue da equação (1.6) que,

$$\Phi_X(t) = \int \cos(tx)dF(x).$$

Portanto $\Phi_X(t) \in \mathbb{R}$.

Por outro lado $\Phi_X(t) \in \mathbb{R} \ \forall t$ se, e somente se, $\Phi_X(t) = \int \cos(tx)dF(x)$.

Da equação (1.6) segue que $\Phi_X(t) = \Phi_{-X}(t)$ ou ainda $X \stackrel{d}{=} -X$. ■

Propriedade 1.6. *Se U e V são independentes e simétricas então $U - V$ e $U + V$ também são simétricas.*

Propriedade 1.7. *Suponha $X_1, \dots, X_n$ cópias independentes de X . Se X é simétrica, temos que $S_n = X_1 + \dots + X_n$ é simétrica.*

Propriedade 1.8. *Se X é simétrica e $\mathbf{E}X$ existe, então $\mathbf{E}X = 0$.*

Propriedade 1.9. *Se X é estável e independente de Y com $Y \stackrel{d}{=} X$, então $X - Y$ é estável e simétrica.*

Propriedade 1.10. *$X \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu)$ é simétrica se, e somente se, $\beta = 0$ e $\mu = 0$. É simétrica em relação a μ se, e somente se, $\beta = 0$.*

Dem. Pela Propriedade 1.5, basta mostrar que a função característica é real. Este caso ocorre somente se $\beta = 0$ e $\mu = 0$. Da Propriedade 1.1 segue a segunda afirmação. ■

É interessante notar que se X é simétrica α -estável (SoS) a função característica toma uma forma bastante simples

$$\mathbf{E}(\exp\{i\xi X\}) = e^{-\sigma^\alpha |\xi|^\alpha}. \quad (1.7)$$

Essa forma simples de visualizar as distribuições simétricas α -estáveis será bastante útil no decorrer das demonstrações de alguns resultados posteriores.

Propriedade 1.11. *Seja $X \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu)$.*

(i) *Se $\alpha \neq 1$, então X é estritamente estável se, e somente se, $\mu = 0$.*

(ii) *Se $\alpha = 1$, então X é estritamente estável se, e somente se, $\beta = 0$.*

1.2.5 Momentos e outras propriedades

Nós mostramos que os quatro parâmetros das distribuições estáveis estão diretamente ligados à posição, escala, assimetria e forma. Poderíamos nos perguntar se existe uma relação próxima entre eles e os momentos teóricos. O segundo momento, exceto no caso normal, não existe para uma distribuição estável e a média existe apenas quando $\alpha > 1$. A propriedade a seguir mostra que momentos de ordem maior que α existem somente quando $\alpha < 2$.

Propriedade 1.12 (Momentos). *Seja $X \sim S_\alpha(\sigma, \beta, \mu)$ com $0 < \alpha < 2$. Então $\mathbf{E}(|X|^p) < \infty$ se, e somente se, $0 < p < \alpha$.*

Propriedade 1.13 (Média). *Quando $1 < \alpha \leq 2$, a média é igual ao parâmetro de posição μ .*

A transformada de Laplace de uma variável aleatória com distribuição estável será muito útil na demonstração de resultados subsequentes. Por essa razão, formalizamos essa importante propriedade abaixo.

Propriedade 1.14 (Transformada de Laplace). *A transformada de Laplace $\mathbf{E}(e^{-\gamma Z})$, $\gamma \geq 0$, de uma variável aleatória $Z \sim S_\alpha(\sigma, 1, 0)$, $0 < \alpha \leq 2$, é dada por*

$$\mathbf{E}(e^{-\gamma Z}) = \exp \left\{ -\frac{\sigma^\alpha}{\cos \frac{\pi\alpha}{2}} \gamma^\alpha \right\}, \quad \text{se } \alpha \neq 1, \quad (1.8)$$

e

$$\mathbf{E}(e^{-\gamma Z}) = \exp \left\{ \sigma \frac{2}{\pi} \gamma \ln \gamma \right\}, \quad \text{se } \alpha = 1. \quad (1.9)$$

Propriedade 1.15. *Sejam $Y \sim \mathcal{N}(0, 1)$ e $A \sim S_{\alpha/2}(\sigma, 1, 0)$ com $0 < \alpha < 2$ e suponha Y e A independentes. Então*

$$Z = A^{1/2}Y \sim S_\alpha \left(\frac{\sqrt{\sigma}}{\sqrt{2} (\cos \frac{\pi\alpha}{4})^{1/\alpha}}, 0, 0 \right). \quad (1.10)$$

Dem. Usando a expressão da função característica de uma v.a. normal padrão, a transformada de Laplace da variável A e o fato de que Y e A são independentes, nós temos

$$\begin{aligned} \mathbf{E}(\exp\{i\xi Z\}) &= \mathbf{E}(\exp\{i\xi A^{1/2}Y\}) \\ &= \mathbf{E}(\mathbf{E}(\exp\{i(\xi A^{1/2})Y\} | A)) \\ &= \mathbf{E} \left(\exp \left\{ -\frac{\xi^2 A}{2} \right\} \right) \\ &= \exp \left\{ -|\xi|^\alpha \left(\frac{\sigma}{2} \right)^{\alpha/2} \frac{1}{\cos \frac{\pi\alpha}{4}} \right\}, \end{aligned}$$

para qualquer real ξ . Segue de (1.7) que Z tem distribuição dada por (1.10). ■

1.2.6 Combinação linear e padronização

Propriedade 1.16 (Combinação linear). *Sejam X_1 e X_2 variáveis aleatórias independentes com $X_i \stackrel{d}{=} S_\alpha(\sigma_i, \beta_i, \mu_i)$, $i = 1, 2$. Então $X_1 + X_2 \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu)$, onde*

$$\sigma = (\sigma_1^\alpha + \sigma_2^\alpha)^{1/\alpha}, \quad \beta = \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}, \quad \mu = \mu_1 + \mu_2.$$

Dem. Primeiramente, vamos verificar para $\alpha \neq 1$. Por independência

$$\begin{aligned} & \ln[\mathbf{E}(\exp(it(X_1 + X_2)))] \\ &= \ln[\mathbf{E}(\exp(itX_1))] + \ln[\mathbf{E}(\exp(itX_2))] \\ &= it(\mu_1 + \mu_2) - (\sigma_1^\alpha + \sigma_2^\alpha) |t|^\alpha + i|t|^\alpha \operatorname{sign}(t) \tan\left(\frac{\pi\alpha}{2}\right) (\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha) \\ &= it(\mu_1 + \mu_2) - (\sigma_1^\alpha + \sigma_2^\alpha) |t|^\alpha \left[1 - i \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha} \operatorname{sign}(t) \tan\left(\frac{\pi\alpha}{2}\right) \right]. \end{aligned}$$

A prova para $\alpha = 1$ é análoga. ■

Propriedade 1.17 (Padronização). *Seja $X \sim S_\alpha(1, \beta, 0)$. Então*

$$Z = \begin{cases} \sigma X + \mu, & \text{se } \alpha \neq 1, \\ \sigma X + \frac{2}{\pi} \beta \sigma \log \sigma + \mu, & \text{se } \alpha = 1, \end{cases} \quad (1.11)$$

tem distribuição $S_\alpha(\sigma, \beta, \mu)$.

1.2.7 Função densidade e função distribuição

Como mostramos anteriormente, existem funções densidade com forma fechada para as distribuições estáveis em alguns casos especiais (Normal, Cauchy e Lévy). Infelizmente não há outros casos onde a função densidade possua forma fechada conhecida. Mesmo não estando disponível, a função densidade tem propriedades analíticas muito importantes.

Propriedade 1.18. *Cada distribuição estável tem função densidade de probabilidade contínua e infinitamente diferenciável.*

Propriedade 1.19. *O suporte das distribuições estáveis é a reta real quando $|\beta| \neq 1$ ou $\alpha \geq 1$, $[\mu, +\infty)$ quando $\beta = 1$ e $\alpha < 1$ e $(-\infty, \mu]$ quando $\beta = -1$ e $\alpha < 1$.*

Propriedade 1.20. *Distribuições estáveis são unimodais. A moda coincide com a média para as distribuições estáveis simétricas com $1 < \alpha \leq 2$. Em outros casos ela não tem forma fechada e necessita ser aproximada por métodos numéricos.*

O problema da não existência de uma forma fechada para a função densidade de probabilidade tem sido uma das principais barreiras ao emprego das distribuições estáveis. Felizmente, com a disponibilidade computacional atual, essa questão pode ser superada por meio de métodos de aproximações numéricas.

A primeira idéia que se poderia explorar seria inverter a função característica de modo a obter a densidade.

$$\begin{aligned} f(x; \alpha, \beta) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \Phi(t; \alpha, \beta) dt \\ &= \frac{1}{\pi} \Re \left[\int_0^{\infty} e^{-itx} \Phi(t; \alpha, \beta) dt \right]. \end{aligned} \quad (1.12)$$

Em princípio, a expressão acima pode ser calculada rapidamente por meio de uma transformada de Fourier, ver Mittnik, S., Doganoglu, T. & Chenyao, D. (1999) para mais detalhes, mas produz apenas um conjunto de abscissas com as densidades associadas. O comportamento suave da função densidade precisa então ser reproduzido por meio de uma função interpoladora. Abaixo apresentamos duas maneiras alternativas para o cômputo da função densidade.

O comportamento assintótico, quando $x \rightarrow \infty$, da densidade pode ser concebido a partir das expansões de Bergström obtidas em Bergström (1952):

$$\begin{aligned} f_2(x; \alpha, \beta) &= \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{\Gamma(k\alpha + 1)}{\Gamma(k + 1)} \operatorname{sen} \left[\frac{k\pi\alpha}{2} K(\alpha) \right] (-1)^{k-1} x^{-k\alpha-1}, \\ f_2(x; \alpha, \beta) &= \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{\Gamma(k/\alpha + 1)}{\Gamma(k + 1)} \operatorname{sen} \left[\frac{k\pi\alpha}{2\alpha} K(\alpha) \right] (-x)^{-k\alpha-1}, \end{aligned} \quad (1.13)$$

onde

$$K(\alpha) = \alpha + \beta \min(\alpha, 2 - \alpha);$$

a forma é uma expansão assintótica quando $x \rightarrow +\infty$ para $\alpha \in (1, 2)$ e uma série absolutamente convergente $\forall x > 0$ para $\alpha \in (0, 1)$, a última é uma expansão assintótica quando $x \rightarrow 0^+$ para $\alpha \in (0, 1)$ e uma série absolutamente convergente $\forall x > 0$ para $\alpha \in (1, 2)$. Note que as expressões acima não estão definidas para valores negativos de x . Esse problema pode ser facilmente superado ao explorar a propriedade

$$f(x; \alpha; \beta) = f(-x; \alpha; -\beta).$$

Note também que as expressões acima não se aplicam quando $\alpha = 1$. Nesse caso, nós temos para $\beta > 0$,

$$f_2(x; \alpha, \beta) = \frac{1}{\pi} \sum_{k=1}^{\infty} (-1)^{k-1} b_k k x^{k-1},$$

onde

$$b_k = \frac{1}{\Gamma(k + 1)} \int_0^{\infty} e^{-\beta u \ln u} u^{k-1} \operatorname{sen} \left[(1 + \beta) u \frac{\pi}{2} \right] du.$$

Uma alternativa rápida para o caso de distribuições estáveis simétricas é também apresentada em McCulloch (1998b). Outros resultados expandem a equação (1.13) em uma expressão mais manejável

e fácil de computar, reparametrizando a função característica. Os detalhes podem ser encontrados em Nolan (1997).

Resultados numéricos também estão disponíveis para a função distribuição. Nas Figuras 1.4 e 1.5 nós antecipamos alguns exemplos de função densidade variando alguns parâmetros das distribuições estáveis; na Figura 1.6 abaixo disponibilizamos alguns exemplos para a função distribuição.

Figura 1.6: Função distribuição com diferentes valores para os parâmetros.

1.2.8 Simulação de uma variável aleatória α -estável

Como vimos, existe uma dificuldade considerável para trabalharmos com a função densidade e a função distribuição por não termos uma forma fechada para as mesmas. Felizmente, simular uma variável aleatória com distribuição estável é bastante simples. Para tal considere $\zeta = \arctan(\beta \tan(\frac{\pi\alpha}{2})/\alpha)$, $\eta = [1 + \beta^2 \tan^2(\frac{\pi\alpha}{2})]^{1/(2\alpha)}$, $W \sim \exp(1)$ e $U \sim U(-\pi/2, \pi/2)$. Por Weron (1995), sabemos que

$$X = \begin{cases} \eta \frac{\sin \alpha(\zeta + U)}{(\cos U)^{1/\alpha}} \left[\frac{\cos(U - \alpha(U + \zeta))}{W} \right]^{\frac{1-\alpha}{\alpha}} & \text{se } \alpha \neq 1, \\ \frac{2}{\pi} [(\pi/2 + \beta U) \tan U] - \beta \ln \frac{\pi/2 W \cos U}{\pi/2 + \beta U} & \text{se } \alpha = 1, \end{cases} \quad (1.14)$$

tem distribuição $S_\alpha(1, \beta, 0)$. Podemos obter uma v.a. com distribuição $S_\alpha(\sigma, \beta, \mu)$ facilmente ao aplicarmos a Propriedade 1.17.

1.3 Estimação e inferência

Apesar do seu imediato apelo teórico e da inadequação da hipótese Gaussiana para uma grande variedade de situações práticas, as distribuições estáveis tem encontrado uma insistente dificuldade na sua difusão em estatística aplicada devido à questão intrigante de sua estimação, muito por ausência de procedimentos de estimação considerados usuais e a inexistência de forma fechada para sua densidade. Nesta seção iremos discutir uma breve revisão dos principais métodos de estimação existentes na literatura.

1.3.1 Métodos baseados em quantis

A primeira idéia explorada para a estimação dos parâmetros de uma distribuição estável, em particular sobre a influência de α sobre as caudas, aparece nos métodos baseados em quantis. As primeiras contribuições nesta direção foram focadas em distribuições α estáveis simétricas, S α S. Fama & Roll (1968) mostra que o parâmetro de posição μ pode ser estimado por uma média truncada de modo não viesado, isto é, a média da porção “central” da amostra. Eles propõem, para situações práticas, o uso da média truncada em 50% quando é sabido que α está no intervalo de 0 a 2. Em um artigo subsequente, Fama & Roll (1971), os autores mostram como estimar σ e α . No que concerne ao parâmetro de escala, eles observaram que o quantil 0.72 de uma distribuição $S_\alpha(1, 0, 0)$ é muito próximo (± 0.003) de 0.827 independentemente de α , então propuseram

$$\hat{\sigma} = \frac{\hat{q}_{.72} - \hat{q}_{.28}}{2 \cdot 0.827}.$$

Como este estimador é uma combinação linear de estatísticas de ordem, o mesmo tem distribuição assintótica normal. Um estudo Monte Carlo apontou que o viés assintótico é menor que 0.4%. Um procedimento similar é aplicado para a estimação de α . Como este parâmetro determina o comportamento da cauda da distribuição, pode-se usar

$$\hat{z}_f = \frac{\hat{q}_f - \hat{q}_{1-f}}{2\hat{\sigma}}$$

e então procurar por um valor $\hat{\alpha}$ cujo quantil teórico seja equivalente a $\hat{z}_f$. A escolha de f é uma questão complicada. Como o método de estimação lida com o comportamento de cauda, pode-se escolher f suficientemente grande; por outro lado, valores muito altos de f tendem a aumentar a dispersão da amostra. Um estudo Monte Carlo revelou que valores de 0.95 a 0.97 são robustos contra a variação do “verdadeiro” α .

Uma extensão mais refinada dos métodos de estimação baseados em quantis foi proposto por McCulloch (1986). Vamos considerar a distribuição $S_\alpha(\sigma, \beta, \mu)$. As quantidades

$$\nu_\alpha = \frac{q_{.95} - q_{.05}}{q_{.75} - q_{.25}}$$

e

$$\nu_\beta = \frac{q_{.95} + q_{0.05} - 2q_{.5}}{q_{.95} - q_{.05}}$$

não dependem de σ e μ e são tabuladas pelo autor como uma função de α e β . A estimação dos dois parâmetros é então baseada ao confrontar as quantias amostrais análogas às expressões com suas

quantias teóricas. Uma vez obtidas as estimativas de α e β , podemos seguir para a estimação do parâmetro σ . A quantidade

$$\nu_\sigma = \frac{q_{.75} - q_{.25}}{\sigma}$$

foi tabulada como função de α e β , então

$$\hat{\sigma} = \frac{\hat{q}_{.75} - \hat{q}_{.25}}{\nu_\sigma(\hat{\alpha}, \hat{\beta})}.$$

Finalmente, o parâmetro de posição μ pode ser estimado como segue. Primeiro consideremos a transformação

$$\zeta = \begin{cases} \mu + \beta\sigma \tan \frac{\pi\alpha}{2}, & \text{se } \alpha \neq 1, \\ \mu, & \text{se } \alpha = 1. \end{cases}$$

A quantidade

$$\nu_\zeta = \frac{\zeta - q_{.5}}{\sigma}$$

foi tabulada como função de α e β , então $\hat{\zeta} = \hat{q}_{.5} + \hat{\sigma}\nu_\zeta(\hat{\alpha}, \hat{\beta})$ e

$$\hat{\mu} = \hat{\zeta} - \hat{\beta}\hat{\sigma} \tan \frac{\pi\hat{\alpha}}{2}.$$

O estimadores acima são consistentes mas não eficientes. Mais ainda, esse tipo de aproximação não pode ser diretamente estendida para modelos lineares e séries temporais e portanto requer um método de estimação em dois passos.

1.3.2 Métodos baseados na função característica

A idéia de estimar os parâmetros de uma distribuição estável por meio da função característica empírica começa com Press (1972) e Paulson et. al. (1975). O fundamento por trás deste método é minimizar numericamente, com respeito aos parâmetros, um critério apropriado de distância entre a função característica empírica e a função característica teórica. Dada uma amostra $\mathbf{y}$ de n unidades, a função característica empírica é

$$\hat{\phi} = \frac{1}{n} \sum_{j=1}^n e^{ity_j} = \frac{1}{n} \sum_{j=1}^n [\cos(ty_j) - i \sin(ty_j)]. \quad (1.15)$$

O método consiste em minimizar

$$\int_{-\infty}^{+\infty} |\hat{\phi}(t) - \phi(t)|^2 w(t) dt \quad (1.16)$$

com respeito aos parâmetros da distribuição estável, onde $w(t)$ é uma função peso apropriada. Os autores sugerem, por razões computacionais, $w(t) = e^{-t^2}$; e a integral em (1.16) é calculada por quadratura numérica. Simulações implementadas pelo autor mostraram sucesso do método de estimação somente para valores de σ e μ próximos de 1 e 0 respectivamente. Desse modo, eles propuseram uma espécie de procedimento em dois passos no qual estimativas da posição e da escala são obtidos e então a estimação é realizada com os valores amostrais padronizados.

Um procedimento diferente foi proposto por Koutrouvelis (1980). O autor notou que o logaritmo da função característica pode ser expresso, para $\alpha \neq 1$, como:

$$\ln \phi(t) = -|\sigma t|^\alpha + i \left[\mu t + |\sigma t|^\alpha \text{sinal}(t) \beta \tan \frac{\pi \alpha}{2} \right]. \quad (1.17)$$

A parte real da expressão acima é portanto

$$\Re[\ln \phi(t)] = -\sigma |t|^\alpha,$$

então

$$\ln \{-\Re[\ln \phi(t)]\} = \alpha \ln |t| + \alpha \ln \sigma. \quad (1.18)$$

Similarmente, a parte imaginária de (1.17) pode ser expressa como

$$\Im[\ln \phi(t)] = \mu t + |\sigma t|^\alpha \text{sinal}(t) \beta \tan \frac{\pi \alpha}{2}. \quad (1.19)$$

O método proposto é portanto:

- (i) Estimar σ e μ por método dos quantis e dados padronizados.
- (ii) Calcular a função característica empírica (1.15) e as análogas amostrais de (1.18) e (1.19) por meio de

$$\Re[\ln \hat{\phi}(t)] = \ln |\hat{\phi}(t)| \quad \text{e} \quad \Im[\ln \hat{\phi}(t)] = \arctan \left\{ \frac{\Re[\hat{\phi}(t)]}{\Im[\hat{\phi}(t)]} \right\}.$$

Os pontos nos quais a função característica é calculada são determinados com base em uma tabela proposta pelo autor.

- (iii) Fazer a regressão de $\ln \{-\Re[\ln \hat{\phi}(t)]\}$ em uma constante e $\ln |t|$, ver (1.18); dividir a constante por $\hat{\alpha}$ é a base para melhorar $\hat{\sigma}$.
- (iv) Fazer uma regressão da parte imaginária da função característica empírica em uma constante e t , ver (1.19), para obter uma estimativa de β e um melhoramento de μ .

Mais tarde, este método foi aprimorado em Koutrouvelis (1981) por meio de uma regressão GLS iterada onde as covariâncias dos erros da regressão fazem o papel de peso.

Uma técnica ainda mais refinada explorando a função característica foi recentemente proposta por Kogon & Williams (1998). Os autores utilizaram idéia parecida no que diz respeito à regressão, no entanto realizaram uma reparametrização da função característica. Apesar do uso de uma diferente parametrização, a diferença entre este método e o proposto por Koutrouvelis (1980) reside na escolha dos pontos onde a função característica é calculada. Ao invés de olhar uma tabela, os autores usam um intervalo fixo independente do parâmetro de escala e do parâmetro de locação. Este intervalo é determinado ao observar que a função característica amostral é determinística e igual à sua teórica correspondente em $t = 0$, então o intervalo ótimo proposto é $t \in [0.1, 1.0]$. Um exercício de simulação mostra que o número ótimo de pontos igualmente espaçados em que o intervalo deve ser dividido é 10.

Após estudo detalhado de simulações, observa-se que este estimador tem performance, em geral, muito boa, e tende a ter desempenho superior ao método de quantis de McCulloch (1986). Esta performance, entretanto, é pior que o método GLS de Koutrouvelis (1981), especialmente quando α é pequeno. No entanto, esta perda de acurácia é compensada por uma formulação simples dado que não requer o uso de uma tabela e nem procedimento iterativo.

1.3.3 Máxima verossimilhança

O potencial do método de máxima verossimilhança foi estudado primeiramente por DuMochel (1973a). O autor mostra que o estimador segue a teoria assintótica padrão. Entretanto, o autor cita que a função de verossimilhança tem comportamento estranho já que

$$\lim_{\alpha \rightarrow 0, x \rightarrow \mu} \mathcal{L}(x; \alpha, \mu) = +\infty.$$

Portanto, ela apresenta picos sempre que os valores observados se aproximam do parâmetro de posição μ . Quando μ é conhecido, isto obviamente não acontece, dado que $\mathcal{P}(X = \mu) = 0$. O problema pode ser superado se restringirmos a admissibilidade de valores para α em $\epsilon < \alpha \leq 2$, com ϵ arbitrariamente pequeno. Como a função de verossimilhança tem uma descontinuidade em $\alpha = 1$, também exclui-se este valor de parâmetro. Iremos denotar o espaço amostral restrito por Θ , e C será um conjunto aberto arbitrário de Θ .

A prova da normalidade assintótica e consistência do estimador de máxima verossimilhança é baseado nas condições de LeCam (1952):

(i) A densidade $f(x; \theta)$ é contínua $\forall \theta \in \Theta$ e tem primeira e segunda derivadas parciais contínuas $\forall \theta_0 \in C^c$, onde C^c denota o complemento de C .

(ii) A densidade conjunta da amostra é tal que

$$\mathbf{E}_{\theta} \left[\sup_{\theta \in C^c} \ln \frac{\prod_{i=1}^n f(x_i; \theta)}{\prod_{i=1}^n f(x_i; \theta_0)} \right] < \infty$$

para todo $\theta_0 \in C^c$.

(iii) Sejam $g(\theta)$ e $H(\theta)$, respectivamente, o gradiente e a Hessiana da função verossimilhança. Então existe uma função $m(x)$ tal que

$$\int_{-\infty}^{+\infty} m(x) f(x; \theta) dx \leq M < +\infty$$

e $H(\theta)$ é limitada em valor absoluto por $m(x)$.

(iv) $\forall \theta \in C^c$, $\mathbf{E}[g(\theta)] = 0$ e $I(\theta) = -\mathbf{E}[H(\theta)]$.

(v) $\forall \theta \in \Theta$ e $\forall \theta_0 \in C^c$,

$$\int_{-\infty}^{+\infty} |f(x; \theta) - f(x; \theta_0)| dx > 0.$$

(vi) A matriz de informação é não singular $\forall \theta \in C^c$.

A prova de que as condições acima valem é dada em DuMochel (1973a), permitindo estabelecer o seguinte teorema.

Teorema 1.5 (Estimador de máxima verossimilhança). *O estimador de máxima verossimilhança é consistente e assintoticamente normal com matriz de covariâncias $I^{-1}(\theta)$ contanto que $\theta \in \Theta$.*

Na prática, a aproximação normal serve apenas para certos valores de α e β : quando os dois parâmetros estão próximos dos seus limites, a distribuição assintótica dos estimadores de máxima verossimilhança tende a ficar degenerada para um tamanho de amostra fixo. Entretanto, este resultado é bastante promissor e, em princípio, deveria influenciar na preferência da máxima verossimilhança em relação aos outros métodos descritos nas seções anteriores. Infelizmente, este método carrega uma séria dificuldade de emprego devido à ausência de forma fechada para a densidade, então o cálculo da verossimilhança requer a inversão numérica da função característica para cada diferente vetor de parâmetros e pode consumir bastante tempo.

A primeira proposta para superar esta dificuldade computacional foi apresentada por DuMochel (1971): ele propõe dividir os dados em intervalos de comprimento fixo e então calcular a transformada inversa rápida de Fourier para obter a densidade de grupos “centrais” de observações e a expansão assintótica (1.13) para as observações “extremas”. A perda de informação associada com essa técnica de agrupamento é analisada em DuMochel (1975).

Dois métodos para o cálculo prático do estimador de máxima verossimilhança foram recentemente propostos: o primeiro, ver Nolan (2002), explora a representação em séries de Nolan (1997); o segundo, explorado inicialmente por Mittnik et. al. (1999), usa uma interpolação linear da transformada rápida de Fourier da função característica (Mittnik, S., Doganoglu, T. & Chenyao, D. (1999)). Os dois métodos apresentam uma significativa redução do erro médio quadrático em relação ao método de quantis de McCulloch (1986).

1.3.4 Métodos baseados em simulação

Recentemente métodos baseados em simulação foram propostos por Gouriéroux et. al. (1993) e Gallant e Tauchen (1996) para situações onde é difícil encontrar formas fechadas para a função de verossimilhança mas é fácil simular os processos e ou distribuições em questão. Método de estimação indireta é o nome que usaremos para nos referir a este princípio. Recentemente, Lombardi & Calzolari (2008) aplicaram este método para estimar os parâmetros da distribuição estável. Sendo este método de suma importância para o trabalho dedicaremos um capítulo detalhado sobre o mesmo. Mais à frente usaremos o método para estimar parâmetros de modelos para volatilidade.

1.4 Testes para o comportamento estável

Em princípio, por ter quatro parâmetros ao invés de dois, seria razoável não ter dúvidas de que as distribuições estáveis se ajustam melhor aos dados que a normal. Entretanto, em muitos casos, dados empíricos não estão muito longe da normalidade e passa a ser fundamental uma maneira de testar tais distribuições.¹

¹Uma fonte muito útil sobre testes está disponível em Nardelli (1997) e inspirou esta seção.

O problema da escolha entre a distribuição normal e distribuições estáveis mais gerais foi primeiro discutida em DuMouchel (1983). Como comentamos nos parágrafos anteriores, o estimador de máxima verossimilhança segue a teoria assintótica padrão para valores dentro dos limites impostos pelo espaço paramétrico. Por outro lado, quando α se aproxima de 2, o estimador se torna super eficiente, isto é seu erro padrão tende a zero para uma amostra de tamanho fixo: isto é obviamente o maior problema na estrutura de teste. Entretanto o autor mostra que

$$\lim_{n \rightarrow \infty} \mathcal{P}(\hat{\alpha} = 2) = 1.$$

Um exercício de simulação, na qual observações distribuídas segundo a normal foram ajustadas a um modelo estável, considera a propriedade do simples teste construído sob a regra “rejeitar quando $\hat{\alpha} < 2$ ”. Aconteceu que com uma amostra de tamanho 1000, o teste incorre em erro tipo I em 16% das vezes. Mais adiante, o autor conjecturou que

$$\mathcal{P}(\hat{\alpha} = 2) \approx \frac{k}{\ln n},$$

onde k é um termo constante, de modo que precisaríamos de alguns milhões de observações antes de obter um teste com 5% de significância.

Por causa desta dificuldade, os procedimentos de teste tem sido focados em técnicas visuais. Além do já mencionado variograma, muitos autores tem desenvolvido alternativas para testar dados: Nolan (2002), por exemplo, analisa o uso de QxQ e PxP plots.

Alguns autores (Fama & Roll (1971)) tentaram explorar a propriedade de estabilidade: se um conjunto de dados é estável, agregações deveriam também ser estáveis e exibir o mesmo índice de estabilidade. Este método conferiu rejeição da hipótese de estabilidade para diversos fenômenos, pois notou-se que à medida que os níveis de agregação se moviam α se aproximava de 2. Estas metodologias para testes não podem ser consideradas apropriadas por duas razões: primeiro, como mostra Diebold (1986), o que rejeitamos é o fato das variáveis aleatórias não serem i.i.d.'s, elas poderiam ser estáveis com diferentes parâmetros; segundo, como mostrado em Fielitz & Smith (1971), se os dados forem gerados por misturas de distribuições estáveis com diferentes parâmetros de escala e posição, o valor estimado de α tende a crescer com o nível de agregação.

Na maioria dos casos, entretanto, trabalhos práticos focaram no comportamento de cauda da distribuição: é comum se valer de gráficos da distribuição empírica em escala log-log que são usadas para justificar o emprego de um modelo estável ao observar o decaimento da potência da cauda (linear em escala log-log). Além do fato de não ser formal e não lidar com a parte “central” da distribuição, que poderia muito bem estar longe de ser estável, o método tende ao erro e deveria ser evitado pois o decaimento da potência de cauda ocorre assintoticamente e é difícil decidir quando x é grande o suficiente para justificá-lo.

Um teste amplamente empregado que lida com o comportamento de cauda é o chamado gráfico de Hill (Hill (1975)). A idéia é calcular recursivamente

$$h_k = \left[\frac{1}{k} \sum_{i=n-k}^n \ln x_{(i)} - \ln x_{(n-k)} \right]^{-1}$$

para $k = 1, \dots, n$ e ver como a quantidade se comporta. Se a cauda tiver decaimento de potência esta quantidade deveria se aproximar do expoente. Este método é bem simples e tem encontrado

forte apelo. Como vários estudos empíricos tem encontrado $\hat{\alpha} > 2$, isto tem sido interpretado como evidência contra a estabilidade. Esta linha de raciocínio foi fortemente criticada por McCulloch(1997). Ele mostra que amostras com distribuição estável onde α é menor que 1,65 também pode gerar $\hat{\alpha} > 2$.

No mesmo artigo, os valores críticos da estatística LR sobre a hipótese nula $\alpha = 2$ são tabulados, tornando possível o uso da razão de verossimilhança para discernir entre modelos normal e estáveis.

1.5 Processos de Lévy α -estáveis

Nesta seção, apresentaremos a definição e algumas importantes propriedades dos processos de Lévy α -estáveis. Inicialmente, vamos apresentar o conceito de uma variável aleatória com distribuição infinitamente divisível.

Definição 1.5. *Uma variável aleatória X tem distribuição infinitamente divisível se para todo $n \in \mathbb{N}$ existem variáveis aleatórias $X_1^{(n)}, X_2^{(n)}, \dots, X_n^{(n)}$ i.i.d. tais que*

$$X \stackrel{d}{=} X_1^{(n)} + \dots + X_n^{(n)}.$$

Daqui em diante diremos que X é infinitamente divisível se X tiver distribuição infinitamente divisível.

Definição 1.6. *Um processo estocástico $X = \{X(t) : t \geq 0\}$ com valores em $\mathbb{R}^d$ é dito um processo de Lévy se :*

- (i) $X(0) = 0$ q.c.
- (ii) X tem incrementos independentes.
- (iii) X tem incrementos estacionários, isto é, $X(t+h) - X(t) \stackrel{d}{=} X(h), \forall t > 0$.
- (iv) X é estocasticamente contínuo, isto é, $\lim_{t \rightarrow t_0} \mathcal{P}\{|X(t) - X(t_0)| > \epsilon\} = 0$.
- (v) Para quase todo ω , $X(t, \omega)$ é uma trajetória contínua à direita com limite à esquerda.

Observações:

- (1) Pode-se provar que, se X satisfaz os itens (i), (ii), (iii) e (iv), então existe uma versão $\bar{X}$ satisfazendo (v).
- (2) As condições (i), (ii) e (iii) equivalem a um processo estacionário com incrementos independentes.

Proposição 1.1. *Se X é um processo de Lévy, então X possui distribuição infinitamente divisível.*

Dem. Para $t > 0$, tomando $t_0 = 0$ e $t_k = t$, temos

$$\begin{aligned} X(t) &\stackrel{d}{=} (X(t_1) - X(t_0)) + (X(t_2) - X(t_1)) + \cdots + (X(t_k) - X(t_{k-1})) \\ &\stackrel{d}{=} X(t_k) - X(t_0) \stackrel{d}{=} X(t) - X(0). \end{aligned}$$

Fazendo $t_j = \frac{j}{k}t$, $j = 0, \dots, k$ segue que os termos

$$X(t_j) - X(t_{j-1}) \stackrel{d}{=} X(t_1)$$

são independentes. Portanto, $X(t)$ tem distribuição infinitamente divisível. ■

Definição 1.7. Dizemos que X é um processo de Lévy se X for estacionário com incrementos independentes e $X(1)$ possuir distribuição infinitamente divisível.

Pode-se demonstrar que as Definições 1.6 e 1.7 são equivalentes.

Definição 1.8. Dizemos que $Z_\alpha = \{Z_\alpha(t) : t \geq 0\}$ é um processo (movimento) de Lévy α -estável se:

- (i) $\mathcal{P}\{Z_\alpha(0) = 0\} = 1$ ou $Z_\alpha(0) = 0$ q.c.
- (ii) Z_α possui incrementos independentes.
- (iii) $Z_\alpha(t+h) - Z_\alpha(t) \stackrel{d}{=} S_\alpha(\sigma h^{1/\alpha}, \beta, 0)$.

Definição 1.9. Um processo Z_α iniciado em 0, com incrementos independentes e estacionários, é dito Lévy α -estável se

$$Z_\alpha(1) \stackrel{d}{=} S_\alpha(\sigma, \beta, 0).$$

Proposição 1.2. A Definição 1.8 é equivalente à Definição 1.9.

Dem. Primeiro vamos considerar o processo Z_α da Definição 1.8. Pela condição (iii), temos incrementos estacionários. Logo, tomando $t = 1$ e $h = 0$, temos

$$Z_\alpha(1) \stackrel{d}{=} S_\alpha(\sigma, \beta, 0).$$

Agora consideremos a Definição 1.9. As condições (i) e (ii) já estão satisfeitas, então basta mostrar a condição (iii). Nós temos que

$$Z_\alpha(0) = 0$$

e

$$Z_\alpha(1) \stackrel{d}{=} Z_\alpha(1) - Z_\alpha(0) \stackrel{d}{=} S_\alpha(\sigma, \beta, 0).$$

Pela Propriedade 1.16 temos, para U e V independentes, que se

$$U \stackrel{d}{=} S_\alpha(\sigma_U, \beta_U, 0) \quad \text{e} \quad V \stackrel{d}{=} S_\alpha(\sigma_V, \beta_V, 0)$$

então

$$U + V \stackrel{d}{=} S_\alpha \left((\sigma_U^\alpha + \sigma_V^\alpha)^{1/\alpha}, \frac{\beta_U \sigma_U^\alpha + \beta_V \sigma_V^\alpha}{\sigma_U^\alpha + \sigma_V^\alpha}, 0 \right).$$

Dessa forma $\forall n \in \mathbb{N}$,

$$Z_\alpha(n) \stackrel{d}{=} Z_\alpha(n) - Z_\alpha(n-1) + Z_\alpha(n-1) - Z_\alpha(n-2) + \cdots + Z_\alpha(1) - Z_\alpha(0).$$

Como o processo Z_α tem, por hipótese, incrementos independentes e estacionários, então temos para cada $0 \leq i \leq n$ que $Z_\alpha(i) - Z_\alpha(i-1) \stackrel{d}{=} S_\alpha(\sigma, \beta, 0)$. Portanto, $Z_\alpha(n) \stackrel{d}{=} S_\alpha(\sigma n^{1/\alpha}, \beta, 0)$. Também temos,

$$\begin{aligned} Z_\alpha(1) &= Z_\alpha\left(\frac{1}{m}m\right) \stackrel{d}{=} S_\alpha\left(\frac{\sigma_1}{m}m^{1/\alpha}, \beta, 0\right) \\ \implies \frac{\sigma_1}{m} &= \sigma \left(\frac{1}{m}\right)^{1/\alpha} \\ \implies Z_\alpha\left(\frac{1}{m}\right) &\stackrel{d}{=} S_\alpha\left(\sigma \left(\frac{1}{m}\right)^{1/\alpha}, \beta, 0\right). \end{aligned}$$

Usamos argumentos semelhantes para obter o resultado para os racionais e reais. ■

Definição 1.10. Um processo estocástico X é dito *autosimilar com expoente de similaridade H* se existe $H > 0$ tal que para qualquer escalar $c > 0$,

$$(X_{ct})_{t \geq 0} \stackrel{d}{=} (c^H X_t)_{t \geq 0}.$$

Proposição 1.3. Todo processo Lévy α -estável é autosimilar com expoente de similaridade $H = 1/\alpha$.

A demonstração segue imediatamente da Propriedade 1.2.

Exemplo 1.7. O processo de Lévy α -estável mais simples que podemos citar é o famoso movimento Browniano $B = \{B(t) : t \geq 0\}$. Sabemos que B é um processo estocástico com incrementos independentes e estacionários. Sabemos também que

$$B(1) \stackrel{d}{=} \mathcal{N}(0, \sigma^2) \stackrel{d}{=} S_2(\sigma/\sqrt{2}, 0, 0).$$

Portanto, pela Definição 1.9, B é um processo de Lévy 2-estável e pela Proposição 1.3 é autosimilar, com expoente de similaridade $H = 1/2$ ou ainda

$$B(t) \stackrel{d}{=} t^{1/2} B(1), t > 0.$$

1.5.1 Simulação de um processo de Lévy α -estável

Para simularmos a trajetória discretizada $(Z(t_1), \dots, Z(t_n))$ de um processo de Lévy α -estável o procedimento também é muito simples. Basta considerar os tempos igualmente espaçados $t_i = t_{i-1} + h$, $i = 2, \dots, n$. Devemos ter $\Delta Z_i = Z(t_i) - Z(t_{i-1}) \stackrel{d}{=} Z_h \sim S_\alpha(\sigma h^{1/\alpha}, \beta, 0)$.

Simulamos então n v.a.'s independentes $U_i \sim U(-\pi/2, \pi/2)$ e n v.a.'s independentes $W_i \sim \exp(1)$. Usando a equação (1.14) teremos n v.a.'s com distribuição $S_\alpha(1, \beta, 0)$ segundo

$$\Delta X_i = \begin{cases} \eta \frac{\sin \alpha(\zeta + U_i)}{(\cos U_i)^{1/\alpha}} \left[\frac{\cos(U - \alpha(U_i + \zeta))}{W_i} \right]^{\frac{1-\alpha}{\alpha}}, & \text{se } \alpha \neq 1, \\ \frac{2}{\pi} [(\pi/2 + \beta U_i) \tan U_i] - \beta \ln \frac{\pi/2 + \beta U_i}{\pi/2 - \beta U_i}, & \text{se } \alpha = 1. \end{cases}$$

Desse modo os saltos do processo, que tem distribuição $S_\alpha(\sigma h^{1/\alpha}, \beta, 0)$, podem ser simulados segundo a transformação estabelecida pela Propriedade 1.17 e obtemos

$$\Delta Z_i = \begin{cases} \sigma h^{1/\alpha} \Delta X_i, & \text{se } \alpha \neq 1, \\ \sigma h^{1/\alpha} \Delta X_i + \frac{2}{\pi} \beta \sigma \log \sigma, & \text{se } \alpha = 1. \end{cases}$$

A trajetória discretizada será dada por $Z(t_i) = \sum_{k=1}^i \Delta Z_k$.

Alguns exemplos de trajetórias podem ser observados na Figura 1.7 (abaixo seguem respectivamente $\alpha = 0.5$, $\alpha = 1.0$, $\alpha = 1.5$ e $\alpha = 2.0$). À medida que aumentamos o índice de estabilidade α a trajetória apresenta menos saltos. Em particular, para $\alpha = 2$ temos o movimento browniano.

Figura 1.7: Processos de Lévy α -estáveis com diferentes valores para α .

Capítulo 2

Processos ARMA e GARCH

Processos ARMA (Autoregressivo e de Médias Móveis) pertencem a classe de processos lineares. Na literatura clássica, as inovações são Gaussianas ou não Gaussianas com variância finita, de modo que a probabilidade de tomarem valores relativamente grandes é demasiadamente pequena. A análise e principais resultados desses processos podem ser encontrados em Brockwell e Davis (1991) e Morettin (2011).

Apesar dos processos lineares serem apropriados para diversos fenômenos na vida real, eles não capturam algumas características de séries temporais financeiras e outros processos que apresentem caudas pesadas. Processos ARCH e GARCH foram desenvolvidos com o propósito de modelar o comportamento de ações, índices de ações, retornos, moedas etc.

2.1 Processos ARMA

Modelos ARMA com variância finita e diversas aplicações são apresentados de forma sistemática em Brockwell e Davis (1991) e Morettin (2011). São exploradas propriedades do modelo, métodos de estimação, média, funções de autocovariância e autocorrelação e previsões no tempo. Abaixo definimos o modelo e comentamos algumas propriedades.

Definição 2.1. *Um processo estacionário $\{X_t\}$ é dito um processo ARMA(p, q), $p, q \in \mathbb{N}$ se para todo t inteiro*

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}, \quad (2.1)$$

onde as inovações ϵ_t são um ruído branco $RB(0, \sigma^2)$.

A equação (2.1) pode ser simbolicamente escrita numa forma mais compacta

$$\Phi(B)X_t = \Theta(B)\epsilon_t,$$

onde $\Phi(z)$ e $\Theta(z)$ são os polinômios

$$\begin{aligned} \Phi(z) &= 1 - \phi_1 z - \cdots - \phi_p z^p, \\ \Theta(z) &= 1 + \theta_1 z + \cdots + \theta_q z^q, \end{aligned}$$

e B é o operador retroativo definido por

$$B^j X_t = X_{t-j}, \quad j \in \mathbb{Z}.$$

Um processo ARMA(p,q) é dito causal se existe uma sequência de constantes $\{c_j\}$, $c_0 = 1$, tais que $\sum_{j=0}^{\infty} c_j < \infty$ e

$$X_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j}, \quad \text{q.c. } t \in \mathbb{Z}.$$

O processo causal ARMA(p,q) é uma média móvel infinita e pode ser denotado por MA(∞).

Considerando as inovações com distribuição normal padrão, suponha que os polinômios $\Phi(z)$ e $\Theta(z)$ não tem raízes em comum. Então $\{X_t\}$ é causal se, e somente se $\Theta(z)$ não tem raízes no disco unitário $\{z : |z| \leq 1\}$. Mais ainda, $\{X_t\}$ é estacionário, Gaussiano com média zero e função de autocovariância

$$\gamma_\tau = \sigma^2 \sum_{j=0}^{\infty} c_j c_{j+|\tau|},$$

que caracteriza a estrutura de dependência.

A distribuição de X_t condicionada a $\mathcal{F}_{t-1} = \sigma\{\epsilon_s : s \leq t-1, s \in \mathbb{Z}\}$ é Gaussiana

$$X_t | \mathcal{F}_{t-1} \sim N \left(\sum_{j=1}^{\infty} c_j \epsilon_{t-j}, \sigma^2 \right).$$

Isto significa que a variância condicional é constante enquanto a média condicional depende do passado e constitui a principal diferença deste para os modelos do tipo GARCH.

2.2 Processos GARCH

A classe de processos ARCH foi introduzida em Engle (1982) para permitir que a variância condicional de uma série temporal dependesse da informação passada.

Como o objetivo é modelar a volatilidade dos retornos de séries temporais financeiras iremos considerar

$$r_t = \ln(P_t) - \ln(P_{t-1}),$$

onde P_t é o preço do ativo no tempo t . Para $\mathcal{F}_{t-1} = \sigma\{\epsilon_s : s \leq t-1, s \in \mathbb{Z}\}$ supomos

$$\mu_t = \mathbf{E}(r_t | \mathcal{F}_{t-1}), \quad h_t = \text{Var}(r_t | \mathcal{F}_{t-1})$$

e definimos o processo ARCH abaixo.

Definição 2.2. Um processo ARCH(q), $q \in \mathbb{N}$ é definido por

$$r_t = \sqrt{h_t} \epsilon_t, \quad t \in \mathbb{Z}, \tag{2.2}$$

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j r_{t-j}^2, \tag{2.3}$$

onde as inovações ϵ_t são i.i.d. com média zero e $\alpha_0 > 0$, $\alpha_q > 0$ e $\alpha_j \geq 0$, $j = 1, \dots, q-1$.

Uma generalização dos modelos ARCH foi sugerida por Bollerslev (1986), Bollerslev (1987) e Bollerslev (1988), o chamado modelo GARCH (generalized ARCH). Um modelo ARMA pode ser mais parcimonioso, no sentido de apresentar menos parâmetros do que um modelo AR ou MA puro. Do mesmo modo, um modelo GARCH pode ser usado para descrever a volatilidade com menos parâmetros do que um modelo ARCH.

Definição 2.3. Um processo $GARCH(p,q)$, $p, q \in \mathbb{N}$ é definido por

$$r_t = \sqrt{h_t} \epsilon_t, \quad t \in \mathbb{Z}, \quad (2.4)$$

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \quad (2.5)$$

onde

$$\begin{aligned} \alpha_0 &> 0, \alpha_p > 0, \alpha_i \geq 0, i = 1, \dots, p-1, \\ \beta_q &> 0, \beta_j \geq 0, j = 1, \dots, q-1, \\ \sum_{k=1}^m (\alpha_k + \beta_k) &< 1, \quad m = \max(p, q), \end{aligned}$$

as inovações ϵ_t são v.a's i.i.d. com média zero e variância 1.

Usualmente supõe-se que as inovações ϵ_t são normais ou seguem uma distribuição t de Student, ou ainda, uma distribuição de erro generalizada.

Processos GARCH pertencem a classe de processos não lineares e modelam agrupamentos de volatilidade. Se tomarmos as inovações Gaussianas, a sua distribuição condicional também será Gaussiana com variância condicional variando no tempo, isto é

$$X_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, h_t).$$

A identificação da ordem do modelo GARCH a ser ajustado costuma ser um pouco complicada. Recomenda-se o uso de modelos de ordem baixa, como (1,1), (1,2), (2,1) ou (2,2) para então decidir o modelo mais adequado por meio de critérios como AIC, BIC, valores de assimetria e curtose, logverossimilhança e de alguma função perda.

Os estimadores dos parâmetros do modelo apresentado em (2.4)-(2.5) podem ser obtidos pelo método de máxima verossimilhança condicional. Supondo a normalidade dos ϵ_t 's, temos que a logverossimilhança, condicional às primeiras m observações, é dada por

$$\ln \mathcal{L}(r_{m+1}, \dots, r_T | \alpha, \beta, r_1, \dots, r_m) \propto -\frac{1}{2} \sum_{t=m+1}^T \ln(h_t) - \frac{1}{2} \sum_{t=m+1}^T \frac{r_t^2}{h_t}.$$

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

Processos ARMA e GARCH com inovações apresentando caudas pesadas tem sido de grande interesse nas pesquisas recentes. Vários casos empíricos apontam para essa abordagem como, por

exemplo, sinais telefônicos, áudio, turbulência aérea e dados financeiros. O caso mais imediato o qual podemos considerar seria um processo ARMA com inovações α -estáveis simétricas. Claramente algumas dificuldades, principalmente na estimação, são impostas a esses modelos por não termos formas fechadas para a função densidade e função de distribuição. Também não teremos disponível a importante função de autocovariância empregada nos processos ARMA e GARCH com inovações Gaussianas. Apesar das dificuldades, o caso ARMA caso já está bem estudado, as definições e principais propriedades podem ser encontradas, por exemplo, em Brockwell e Davis (1991) e Samorodnitsky e Taqqu (1994), a estimação dos parâmetros em Bhansali (1993), Calder e Davis (1996), Mikosch et. al. (1995) e Lombardi & Calzolari (2008) e a análise de dependência em Kokoszka e Taqqu (1994), Nowicka (1997a) e Nowicka e Weron (1995).

Processos GARCH com inovações estáveis também tem sido bastantes explorados, no entanto sua análise se demonstrou bastante difícil, onde propriedades que costumam ser bastante comuns em outros modelos são desconhecidas. Apenas alguns aspectos acerca das simulações são discutidos em Panorska et. al. (1995). Várias formulações foram propostas e algumas são discutidas em De Vries (1991) e Nowicka (1997b).

Nós iremos discutir com detalhes nos capítulos seguintes os processos R-GARCH, que generalizam os processos GARCH permitindo caudas mais pesadas no modelo e iremos propor um método para estimação dos seus parâmetros. A estrutura de dependência também será discutida ao apresentarmos uma alternativa à função de autocovariância.

Capítulo 3

Estimação indireta

A inferência indireta, proposta em Gouriéroux et. al. (1993), é um método inferencial aplicável em situações em que a estimação do modelo de interesse é muito difícil de ser obtida de modo direto. Situações onde a obtenção do estimador de máxima verossimilhança do modelo de interesse é muito difícil mas a simulação desses modelos é fácil, tem o método de estimação indireta como uma forte alternativa. Tomamos Gouriéroux e Monfort (1996) como referência básica para este capítulo.

O método de estimação indireta trabalha sob uma base muito simples. Vamos considerar dois modelos: o *modelo de interesse* e o *modelo auxiliar*. Suponha que nós temos T observações $\mathbf{y}$ e um modelo de interesse cuja função de verossimilhança, $\mathcal{L}^*(\mathbf{y}; \theta)$, é difícil de manejar e maximizar; o modelo pode também depender de uma matriz explicativa X . O estimador de máxima verossimilhança de $\theta \in \Theta$, dado por

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \sum_{t=1}^T \ln \mathcal{L}^*(\theta; y_t), \quad (3.1)$$

está portanto indisponível. Tomemos agora um modelo alternativo, dependendo de um vetor de parâmetros $\zeta \in \Upsilon$, o qual iremos chamar modelo auxiliar. Suponha que o modelo auxiliar, diferentemente do modelo de interesse, tem função de verossimilhança, $\tilde{\mathcal{L}}(\mathbf{y}; \theta)$, fácil de manejar e maximizar. Claramente o estimador

$$\hat{\zeta} = \arg \max_{\zeta \in \Upsilon} \sum_{t=1}^T \ln \tilde{\mathcal{L}}(\zeta; y_t) \quad (3.2)$$

não é necessariamente consistente. A idéia é explorar as simulações do modelo de interesse (modelo original) para corrigir o viés do estimador $\hat{\zeta}$. Como primeiro passo calculamos o estimador de máxima verossimilhança de ζ , que denotaremos por $\hat{\zeta}$. Em seguida simulamos S vetores de tamanho T do modelo original com base num vetor de parâmetros $\hat{\theta}^{(0)}$ que denominaremos “valor inicial”. Vamos denotar cada um desses vetores por $\mathbf{y}^s(\hat{\theta}^{(0)})$, $s = 1, \dots, S$. Desse modo, estimamos os parâmetros do modelo auxiliar com os valores simulados do modelo original,

$$\hat{\zeta}_S \left(\hat{\theta}^{(0)} \right) = \arg \max_{\zeta \in \Upsilon} \sum_{s=1}^S \sum_{t=1}^T \ln \tilde{\mathcal{L}} \left[\zeta; y_t^s \left(\hat{\theta}^{(0)} \right) \right]. \quad (3.3)$$

A idéia é melhorar o valor inicial $\hat{\theta}^{(0)}$ numericamente ao minimizar a distância

$$\left[\hat{\zeta} - \hat{\zeta}_S \right]' \Omega \left[\hat{\zeta} - \hat{\zeta}_S \right],$$

onde Ω é uma matriz simétrica não negativa definida e que define a métrica. Para um estimador $\hat{\theta}^{(p)}$, o procedimento gera $\hat{\theta}^{(p+1)}$; repetimos o procedimento até que $\hat{\theta}^{(p)}$ convirja. O estimador é dado pelo limite numérico

$$\hat{\theta} = \lim_{p \rightarrow \infty} \hat{\theta}^{(p)}.$$

Um método alternativo e bastante similar, introduzido por Gallant e Tauchen (1996), considera a função escore do modelo auxiliar

$$\sum_{t=1}^T \frac{\partial \ln \tilde{\mathcal{L}}(\zeta; y_t)}{\partial \zeta},$$

que é claramente zero para o estimador de quasi-verossimilhança de β . A idéia é fazer o escore computado nas observações simuladas o mais próximo possível de zero. No caso,

$$\min_{\theta} \left\{ \sum_{s=1}^S \sum_{t=1}^T \frac{\partial \ln \tilde{\mathcal{L}}[\zeta; y_t^s(\theta)]}{\partial \zeta} \right\}' \Delta \left\{ \sum_{s=1}^S \sum_{t=1}^T \frac{\partial \ln \tilde{\mathcal{L}}[\zeta; y_t^s(\theta)]}{\partial \zeta} \right\} \quad (3.4)$$

onde Δ é uma matriz simétrica não negativa definida. Como no caso anterior, o estimador é obtido numericamente e o estimador é dado por

$$\check{\theta} = \lim_{p \rightarrow \infty} \check{\theta}^{(p)}.$$

Esse método é bastante útil quando as expressões analíticas para o gradiente do modelo auxiliar estão disponíveis, o que nos permite evitar a otimização numérica para o cálculo de $\hat{\zeta}_S$.

Propriedade 3.1. *Se $\dim \zeta = \dim \theta$ e T é suficientemente grande:*

- (i) $\hat{\theta}$ não depende de Ω .
- (ii) $\check{\theta}$ não depende de Δ .
- (iii) $\hat{\theta} = \check{\theta}$.

Desse modo as duas técnicas são equivalentes e poderemos escolher qual delas é melhor aplicável nos problemas práticos a serem analisados.

3.1 Propriedades assintóticas

Introduzimos nesta seção um novo conceito que é de grande utilidade na análise do comportamento assintótico de estimadores obtidos por inferência direta.

Consideremos o comportamento assintótico da logverossimilhança do modelo auxiliar:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \ln \tilde{\mathcal{L}}(y_t; \theta) = \mathbf{E}_{\theta} \left[\ln \tilde{\mathcal{L}}(\zeta; y_t) \right].$$

A solução do problema de maximização para o caso assintótico é

$$b(\theta) = \max_{\zeta \in \Upsilon} \mathbf{E}_{\theta} \left[\ln \tilde{\mathcal{L}}(\zeta; y_t) \right]. \quad (3.5)$$

A função $b(\theta)$ é chamada *função de ligação* e mapeia o espaço paramétrico do modelo original no espaço paramétrico do modelo auxiliar. O estimador θ , obtido por inferência indireta, é portanto baseado no cálculo da função de ligação sobre o “verdadeiro” ótimo θ^* .

Consideremos as condições de regularidade abaixo que serão necessárias para obter algumas propriedades importantes dos estimadores de inferência indireta.

- CR1.** Os processos $\{y_t\}$ e $\{x_t\}$ são estacionários, e $\{x_t\}$ é independente do ruído branco $\{\epsilon_t\}$.
- CR2.** A função de verossimilhança do modelo auxiliar $\tilde{\mathcal{L}}$ tende quase certamente, quando $T \rightarrow \infty$, para um limite não estocástico.
- CR3.** O limite da função de verossimilhança é contínuo com respeito a ζ e tem um máximo único.
- CR4.** A função de ligação é uma aplicação bijetora de Θ em B e sua primeira derivada com respeito a θ é de posto coluna completo.

Propriedade 3.2. *Se as condições de regularidade acima forem satisfeitas, o estimador de inferência indireta $\hat{\theta}$ é consistente para S fixo e $T \rightarrow \infty$.*

Adicionando algumas outras condições de regularidade sobre o comportamento da função de verossimilhança do modelo auxiliar obtemos normalidade assintótica. As condições adicionais seguem abaixo:

- CR5.** A matriz Hessiana da função de verossimilhança do modelo auxiliar converge para um limite não estocástico J_0 .
- CR6.** O gradiente da função de verossimilhança do modelo auxiliar converge em distribuição para a lei Gaussiana.
- CR7.** A covariância assintótica entre os gradientes de duas unidades s_1 e s_2 de amostras simuladas é constante.

Propriedade 3.3. *Se as condições de regularidade **CR1-CR7** valem, então o estimador de inferência indireta $\hat{\theta}$ é assintoticamente normal para S fixo e $T \rightarrow \infty$:*

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}[0, W(S, \Omega)]. \quad (3.6)$$

A matriz de covariâncias W é definida como segue abaixo. Primeiro vamos considerar duas matrizes:

$$J_0 = \lim_{T \rightarrow \infty} -\frac{\partial^2 \ln \tilde{\mathcal{L}}[b(\theta); \mathbf{y}]}{\partial \beta \partial \beta'},$$

$$I_0 = \lim_{T \rightarrow \infty} \text{Var} \left\{ \sqrt{T} \frac{\partial \ln \tilde{\mathcal{L}}[b(\theta); \mathbf{y}]}{\partial \beta} \right\}.$$

A escolha ótima para a métrica é

$$\Omega^* = J_0 I_0^{-1} J_0$$

e a matriz de covariâncias assintótica é

$$W(S, \Omega) = \left(1 + \frac{1}{S}\right) \left[\frac{\partial b'(\theta)}{\partial \theta} \Omega^* \frac{\partial b(\theta)}{\partial \theta'} \right]^{-1}.$$

No que concerne à versão (3.4), pode-se demonstrar que Δ e $J_0 \Delta J_0$ são equivalentes, de tal modo que a escolha ótima para a métrica é

$$\Delta^* = I_0^{-1}$$

com matriz de covariâncias assintótica

$$W(S, \Delta) = \left(1 + \frac{1}{S}\right) \left[\frac{\partial b'(\theta)}{\partial \theta} I_0^{-1} \frac{\partial b(\theta)}{\partial \theta'} \right]^{-1}.$$

Infelizmente, a expressão acima não pode ser calculada diretamente a menos que se explicita a função de ligação e sua derivada. Em geral, essa é uma tarefa bastante difícil. Entretanto, podemos estimá-la consistentemente por meio da seguinte expressão que foi obtida no apêndice 2 de Gouriéroux et. al. (1993):

$$W(S, \Delta) = \left(1 + \frac{1}{S}\right) \left[\frac{\partial^2 \ln \tilde{\mathcal{L}}}{\partial \theta \partial \zeta'} \hat{I}_0^{-1} \frac{\partial^2 \ln \tilde{\mathcal{L}}}{\partial \theta' \partial \zeta} \right]^{-1},$$

onde

$$\hat{I}_0 = \hat{H}_0 + \sum_{k=1}^K \left(1 - \frac{k}{K+1}\right) (\hat{H}_k + \hat{H}_k'),$$

com

$$\hat{H}_k = \frac{1}{T} \sum_{t=k+1}^T \left[\frac{\partial \ln \tilde{\mathcal{L}}(\zeta; y_{t-k})}{\partial \zeta} \Big|_{\zeta=\hat{\zeta}} \times \frac{\partial \ln \tilde{\mathcal{L}}(\zeta; y_t)}{\partial \zeta'} \Big|_{\zeta=\hat{\zeta}} \right].$$

O valor de K depende da estrutura de autocorrelação do gradiente do modelo auxiliar e está bastante relacionado à condição **CR7**.

3.2 Trabalhando com restrições

A inferência indireta, como vimos, requer que os parâmetros do modelo auxiliar não tenham restrições de modo que seus estimadores tenham distribuição assintoticamente normal e matriz de covariâncias de posto completo sob as condições de regularidade. Em muitas situações essa hipótese não é realística e seria de suma importância embutir restrições que garantam a boa definição da função de verossimilhança. Felizmente, a inferência indireta foi estendida nesse sentido em Calzolari et. al. (2001).

Para tal, basta otimizar a função de Lagrange

$$\mathcal{Q}(\zeta) = \tilde{\mathcal{L}}(\zeta; \mathbf{y}) + \lambda h'(\zeta), \quad (3.7)$$

onde $h(\zeta)$ é o vetor de funções sumarizando as restrições e λ é o vetor de multiplicadores de Lagrange. A função de ligação é então obtida por uma maximização restrita da função de verossimilhança do modelo auxiliar. Mostra-se, também em Calzolari et. al. (2001), que sob três condições que imitam as condições de regularidade do caso sem restrições, o estimador de inferência indireta tem distribuição assintótica normal. Mais ainda, também no caso com restrições, o método de Gouriéroux et. al. (1993) e Gallant e Tauchen (1996) são equivalentes.

3.3 Inferência indireta para distribuições e processos α -estáveis

A grande questão no métodos de inferência indireta é a relação que existe entre a dificuldade de se obter formas fechadas para a função de verossimilhança e facilidade de simulação das distribuições e ou processos em voga. O primeiro passo é ser capaz de especificar um modelo auxiliar adequado. Como adiantamos anteriormente, este método foi bem empregado em Lombardi & Calzolari (2008) para obter estimadores dos parâmetros de uma distribuição estável assim como um caso específico de processo ARMA com inovações α -estáveis.

No capítulo seguinte apresentaremos o modelo R-GARCH proposto por Nowicka (1998) que é uma generalização dos processos GARCH ao aleatorizar a volatilidade h_t incluindo distribuições estáveis em sua formulação. Neste capítulo propomos a inferência indireta como meio para estimar os parâmetros dos processos R-GARCH.

Nos Capítulos 5 e 6 propomos outros dois modelos, R_t -GARCH e R_s -GARCH, onde também empregamos a inferência indireta como meio para a estimação dos parâmetros dos modelos em questão.

Capítulo 4

Processos R-GARCH

Os processos GARCH, principalmente com inovações Gaussianas, estão definidos de uma maneira tal que impõe limites ao peso das caudas de sua distribuição não condicional. Para dados financeiros sabemos que as caudas costumam ser bastante pesadas de modo que, para séries temporais financeiras, esses modelos não se mostram perfeitamente adequados. Inovações com distribuição t de Student constituem uma alternativa que permite caudas mais pesadas para modelos GARCH se compararmos com as inovações Gaussianas, no entanto perdemos a desejável estabilidade. A estabilidade é desejável pois a classe de distribuições estáveis possui domínio de atração que possibilita aproximações muito boas. Como a classe de distribuições estáveis (que inclui a distribuição normal como caso particular) permite caudas bastante pesadas e possui a desejável estabilidade é natural pensar num modelo de volatilidade que as inclui.

Os processos R-GARCH (Randomized-GARCH) são propostos em Nowicka (1998). Como essa classe de modelos é de suma importância neste trabalho, iremos apresentar a definição e principais resultados neste capítulo. Ainda neste capítulo iremos propor a inferência indireta para estimar os parâmetros do modelo assim como generalizar alguns resultados envolvendo a relação de dependência.

Definição 4.1. Um modelo R-GARCH(r, p, q), $r, p, q \in \mathbb{N}$ é definido por

$$r_t = \sqrt{h_t} \epsilon_t, \quad t = 0, \pm 1, \pm 2, \dots, \quad (4.1)$$

$$h_t = \sum_{i=1}^r \theta_i \eta_{t-i} + \sum_{j=1}^p \phi_j h_{t-j} + \sum_{k=1}^q \psi_k r_{t-k}^2, \quad (4.2)$$

onde

$$r \geq 1, \theta_r > 0, \theta_i \geq 0, i = 1, \dots, r-1, \quad (4.3)$$

$$p \geq 0, \phi_p > 0, \phi_j \geq 0, j = 1, \dots, p-1, \quad (4.4)$$

$$q \geq 0, \psi_q > 0, \psi_k \geq 0, k = 1, \dots, q-1, \quad (4.5)$$

as inovações ϵ_t são v.a.'s i.i.d. normais padrão ($\epsilon_t \sim \mathcal{N}(0, 1)$), as inovações η_t são v.a.'s i.i.d. positivas e $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes.

Se compararmos a formulação da variância não condicional nos modelos GARCH e R-GARCH apresentados respectivamente nas equações (2.5) e (4.2), podemos notar que o modelo R-GARCH é obtido ao tornar o termo α_0 aleatório. Isso explica a adoção do termo Randomized-GARCH. De fato, tomando os η_t 's finitos, positivos constantes, isto é, $\eta_t = \eta$, $0 < \eta < \infty$ o processo R-GARCH(r,p,q) se reduz ao processo GARCH(p,q) onde o termo constante $\alpha_0 = \eta \cdot \sum_{j=1}^r \theta_j$. Se além disso tomarmos $p = 0$ o processo se reduz ao processo ARCH(q). Identificamos dessa maneira os processos ARCH e GARCH como casos particulares do processo R-GARCH.

A equação (4.2) pode ser escrita simbolicamente na forma compacta

$$h_t = \Theta(B)\eta_t + \Phi(B)h_t + \Psi(B)r_t^2,$$

onde Θ , Φ e Ψ são

$$\begin{aligned}\Theta(B) &= \theta_1 B + \theta_2 B^2 + \cdots + \theta_r B^r, \\ \Phi(B) &= \phi_1 B + \phi_2 B^2 + \cdots + \phi_p B^p, \\ \Psi(B) &= \psi_1 B + \psi_2 B^2 + \cdots + \psi_q B^q\end{aligned}$$

e B é o operador retroativo definido por

$$B^i X_t = X_{t-i}, \quad i = 0, \pm 1, \pm 2, \dots$$

Vamos agora explorar a distribuição condicional dos processos R-GARCH. Para tal, consideremos novamente $\mathcal{F}_{t-1} = \sigma\{\epsilon_s, \eta_s : s \leq t-1, s \in \mathbb{Z}\}$.

Propriedade 4.1. *A distribuição de r_t condicionada a $\mathcal{F}_{t-1}$ é Gaussiana*

$$r_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, h_t).$$

Dem. A distribuição condicional de r_t segue imediatamente da distribuição de ϵ_t e do fato de que h_t é $\mathcal{F}_{t-1}$ -mensurável. ■

Daí segue que a esperança condicional $\mathbf{E}(r_t | \mathcal{F}_{t-1})$ é constante e igual a zero. Desse modo

$$\text{Var}(r_t | \mathcal{F}_{t-1}) = \mathbf{E}(r_t^2 | \mathcal{F}_{t-1}) = h_t.$$

Observe que nos processos R-GARCH a variância condicional h_t depende do passado como função linear das inovações passadas $\eta_{t-1}, \dots, \eta_{t-r}$, das variâncias condicionais $h_{t-1}, \dots, h_{t-p}$ e também das observações passadas dos retornos ao quadrado $r_{t-1}^2, \dots, r_{t-q}^2$.

Se o processo $\{r_t\}$ for estacionário com segundo momento finito, então sua estrutura de dependência poderá ser descrita pela função de autocovariância e, além disso, os r_t 's serão não correlacionados. Apesar deste fato, em geral, não serão independentes.

Proposição 4.1. *Suponha que as inovações η_t possuam esperança finita, $\mathbf{E}\eta_t = \mu$, $0 < \mu < \infty$. Se o processo R-GARCH(r,p,q) é estacionário com segundo momento finito, então $\mathbf{E}(r_t) = 0$, $\text{Var}(r_t) =$*

$\mu\Theta(1)(1-\Phi(1)-\Psi(1))^{-1}$ e $\text{Cov}(r_t, r_k) = 0$ para $t \neq k$. Uma condição necessária para que o processo seja estacionário e o segundo momento seja finito é que

$$\Phi(1) + \Psi(1) = \sum_{j=1}^p \phi_j + \sum_{j=1}^q \psi_j < 1. \quad (4.6)$$

Dem. Primeiramente podemos notar, pela Propriedade 4.1, que a média de r_t é zero

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

Desse modo, a variância pode ser obtida como a esperança não condicional de r_t^2 .

$$\begin{aligned} \mathbf{E}r_t^2 &= \mathbf{E}(h_t \epsilon_t^2) = \mathbf{E}h_t \\ &= \sum_{j=1}^r \theta_j \mathbf{E}\eta_{t-j} + \sum_{j=1}^p \phi_j \mathbf{E}r_{t-j}^2 + \sum_{j=1}^q \psi_j \mathbf{E}r_{t-j}^2 \\ &= \mu \sum_{j=1}^r \theta_j + \sum_{j=1}^p \phi_j \mathbf{E}r_{t-j}^2 + \sum_{j=1}^q \psi_j \mathbf{E}r_{t-j}^2. \end{aligned} \quad (4.7)$$

Como sabemos, uma condição necessária para que haja estacionariedade é que a variância não condicional seja constante $\mathbf{E}r_t^2 = \mathbf{E}r_{t-j}^2$, $j \geq 1$. Aplicando esse fato na equação (4.7) obtemos a variância

$$\text{Var}(r_t) = \mathbf{E}r_t^2 = \frac{\mu \sum_{j=1}^r \theta_j}{1 - \left(\sum_{j=1}^p \phi_j + \sum_{j=1}^q \psi_j \right)} = \frac{\mu\Theta(1)}{1 - \Phi(1) - \Psi(1)}, \quad (4.8)$$

que deve ser finita e positiva. Obtemos assim a condição (4.6)

$$\Phi(1) + \Psi(1) = \sum_{j=1}^p \phi_j + \sum_{j=1}^q \psi_j < 1.$$

Pela simetria dos ϵ_t 's, segue imediatamente que

$$\text{Cov}(r_t, r_k) = 0 \quad \text{para } t \neq k.$$

■

4.1 Processos R-GARCH com inovações estáveis

Se considerarmos os η_t 's estáveis, por um lado podemos perder a propriedade de segundo momento finito, mas é possível obter alguns resultados muito interessantes sobre distribuições não condicionais dos processos R-GARCH.

Vamos considerar o caso R-GARCH(r,p,0) onde as inovações η_t 's são positivas, estritamente estáveis e totalmente assimétricas à direita com distribuição

$$\eta_t \sim S_{\alpha/2} \left(2 \left(\cos \frac{\pi\alpha}{4} \right)^{2/\alpha}, 1, 0 \right), \quad (4.9)$$

onde $0 < \alpha < 2$. O índice de estabilidade menor que um para η_t implica na não existência do primeiro momento da mesma. A escolha do parâmetro de escala fornece resultados numa forma compacta.

Mais ainda, supomos que os polinômios $\Theta(B)$ e $1 - \Phi(B)$ não tem raízes comuns e o polinômio $1 - \Phi(B)$ não tem raízes no disco unitário $\{z : |z| \leq 1\}$. Essa hipótese garante algumas propriedades importantes e similares aos processos GARCH e serão sumarizadas nas propriedades que se seguem.

Proposição 4.2. *Se os polinômios $\Theta(B)$ e $1 - \Phi(B)$ não tem raízes em comum, então a variância condicional h_t para R-GARCH($r, p, 0$) com inovações dadas por (4.9) pode ser representada na forma*

$$h_t = \sum_{j=1}^{\infty} \delta_j \eta_{t-j}, \quad q.c. \quad (4.10)$$

com δ_j 's positivos¹ satisfazendo $\delta_j < Q^j$ eventualmente², $Q > 1$, se e somente se o polinômio $1 - \Phi(B)$ não tem raízes no disco unitário fechado $\{z : |z| \leq 1\}$. A sequência h_t é então estacionária, estritamente estável com índice de estabilidade $\alpha/2$ e

$$h_t \sim S_{\alpha/2} \left(2 \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right)^{2/\alpha} \left(\cos \frac{\pi\alpha}{4} \right)^{2/\alpha}, 1, 0 \right). \quad (4.11)$$

Os δ_j 's são os coeficientes na série expandida de $\frac{\Theta(B)}{1-\Phi(B)}$, $|B| < 1$.

Dem. Essa proposição pode ser provada de maneira análoga ao Teorema 7.12.2 em Samorodnitsky e Taqqu (1994) mas os coeficientes δ_j 's são determinados pelo seguinte sistema de equações

$$\begin{cases} \delta_1 = \theta_1 \\ \delta_2 - \delta_1 \phi_1 = \theta_2 \\ \delta_3 - \delta_2 \phi_1 - \delta_1 \phi_2 = \theta_3 \\ \vdots \\ \delta_r - \delta_{r-1} \phi_1 - \delta_{r-2} \phi_2 - \cdots - \delta_1 \phi_{r-1} = \theta_r \\ \delta_s - \delta_{s-1} \phi_1 - \delta_{s-2} \phi_2 - \cdots - \delta_1 \phi_{s-1} = 0, \quad s > r, \end{cases} \quad (4.12)$$

com o entendimento que $\phi_i = 0$ se $i > p$. Segue da equação (4.12) e da hipótese (4.3) que os δ_j 's são reais, não negativos e ao menos um deles é positivo. Para que $\delta_1 > 0$ supomos $\theta_1 > 0$. A distribuição de h_t segue imediatamente de (4.9) e (4.10). ■

Teorema 4.1. *A distribuição não condicional de r_t no processo R-GARCH($r, p, 0$) com inovações estáveis dadas em (4.9) é simétrica α -estável*

$$r_t \sim S_{\alpha} \left(\left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right)^{1/\alpha}, 0, 0 \right). \quad (4.13)$$

¹Mais precisamente, os δ_j 's são não negativos com pelo menos um deles positivo.

²" $a_j < b_j$ eventualmente" significa que existe um j_0 tal que $a_j < b_j$ para todo $j > j_0$.

Dem. Como ϵ_t e h_t são independentes por definição, podemos utilizar a Propriedade 1.15 tomando $Y = \epsilon_t$ e $A = h_t$. O parâmetro de escala de h_t é

$$\sigma = 2 \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right)^{2/\alpha} \left(\cos \frac{\pi\alpha}{4} \right)^{2/\alpha}$$

e, portanto, a distribuição não condicional de r_t é dado por (4.13). ■

Segue do teorema acima que, escolhendo o índice de estabilidade apropriado para η_t , a distribuição não condicional de r_t pode ser qualquer distribuição estável, com exceção da normal.

Poderíamos estar interessados não somente na distribuição de retornos mas também na distribuição de retornos agregados. Por exemplo, se o modelo R-GARCH(r,p,0) com η_t 's dadas em (4.9) descreve retornos diários, então a fórmula (4.13) determina a distribuição não condicional desses retornos diários. Se estamos interessados nos retornos semanais, mensais, etc, devemos encontrar a distribuição da soma $\sum_{k=0}^{m-1} r_{t-k}$ onde m denota o número de dias no intervalo dado. A distribuição dos retornos agregados segue no teorema abaixo.

Teorema 4.2. *Considere um processo R-GARCH(r,p,0) com inovações η_t dadas em (4.9). A distribuição não condicional da soma $\sum_{k=0}^{m-1} r_{t-k}$, $m \geq 1$, é α -estável simétrica segundo*

$$\sum_{k=0}^{m-1} r_{t-k} \sim S_{\alpha} \left(\left[\sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right)^{\alpha/2} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right)^{\alpha/2} \right]^{1/\alpha}, 0, 0 \right). \quad (4.14)$$

Dem. Como $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes, a função característica de $\sum_{k=0}^{m-1} r_{t-k}$ pode ser expressa como a transformada de Laplace de $\sum_{k=0}^{m-1} h_{t-k}$. De fato,

$$\begin{aligned} \mathbf{E} \left(\exp \left\{ i\xi \sum_{k=0}^{m-1} r_{t-k} \right\} \right) &= \mathbf{E} \left(\exp \left\{ i\xi \sum_{k=0}^{m-1} \sqrt{h_{t-k}} \epsilon_{t-k} \right\} \right) \\ &= \mathbf{E} \left(\mathbf{E} \left[\exp \left\{ i\xi \sum_{k=0}^{m-1} \sqrt{h_{t-k}} \epsilon_{t-k} \right\} \middle| \eta_{t-1}, \eta_{t-2}, \dots \right] \right) \\ &= \mathbf{E} \left(\exp \left\{ -\frac{\xi^2 \sum_{k=0}^{m-1} h_{t-k}}{2} \right\} \right). \end{aligned} \quad (4.15)$$

Para encontrarmos a transformada de Laplace de $\sum_{k=0}^{m-1} h_{t-k}$ usamos a Proposição 4.2 e escrevemos

$$\begin{aligned} \sum_{k=0}^{m-1} h_{t-k} &= \sum_{k=0}^{m-1} \left(\sum_{j=1}^{\infty} \delta_j \eta_{t-k+j} \right) \\ &= \sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right) \eta_{t-j} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right) \eta_{t-j}, \end{aligned}$$

e, portanto, $\sum_{k=0}^{m-1} h_{t-k}$ é positivo, estritamente estável com índice de estabilidade $\alpha/2$, e

$$\sum_{k=0}^{m-1} h_{t-k} \sim S_{\alpha/2} \left(2 \left[\sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right)^{\alpha/2} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right)^{\alpha/2} \right]^{2/\alpha} \left(\cos \frac{\pi\alpha}{4} \right)^{2/\alpha}, 1, 0 \right).$$

Agora que sabemos a distribuição de $\sum_{k=0}^{m-1} h_{t-k}$, aplicamos a Propriedade 1.14 em (4.15) para a sua transformada de Laplace e obtemos

$$\mathbf{E} \left(\exp \left\{ i\xi \sum_{k=0}^{m-1} r_{t-k} \right\} \right) = \exp \left\{ -|\xi|^\alpha \left[\sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right)^{\alpha/2} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right)^{\alpha/2} \right] \right\}.$$

Segue de (1.7) que $\sum_{k=0}^{m-1} r_{t-k}$ é simétrica α -estável com distribuição

$$\sum_{k=0}^{m-1} r_{t-k} \sim S_\alpha \left(\left[\sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right)^{\alpha/2} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right)^{\alpha/2} \right]^{1/\alpha}, 0, 0 \right).$$

■

Após tantas propriedades importantes seria desejável que o processo R-GARCH(r,p,0) com inovações dada em (4.9) fosse estacionário. De fato, isso se confirma no corolário abaixo.

Corolário 4.1. *O processo R-GARCH(r,p,0) com inovações η_t dada por (4.9) é simétrico α -estável e estacionário, $0 < \alpha < 2$.*

Dem. Utilizando argumento similar ao da demonstração do Teorema 4.2 mostra-se que todas as combinações lineares finitas de r_t 's são simétricas α -estáveis com mesmo índice de estabilidade α . Um resultado conhecido de processos estáveis é que um processo é simétrico estável se, e somente se todas as suas combinações lineares finitas são simétricas estáveis. Segue, dessa maneira, que r_t é um processo estocástico simétrico α -estável e todas as suas distribuições finito dimensionais são simétricas α -estáveis. Mais que isso, o processo r_t é estacionário. De fato, como $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes, a função característica do vetor $(r_{t_1+k}, \dots, r_{t_d+k})$ pode ser calculada como segue abaixo

$$\begin{aligned} \mathbf{E} \left(\exp \left\{ i \sum_{j=1}^d \xi_j r_{t_j+k} \right\} \right) &= \mathbf{E} \left(\exp \left\{ i \sum_{j=1}^d \xi_j \sqrt{h_{t_j+k}} \epsilon_{t_j+k} \right\} \right) \\ &= \mathbf{E} \left(\mathbf{E} \left[\exp \left\{ i \sum_{j=1}^d \xi_j \sqrt{h_{t_j+k}} \epsilon_{t_j+k} \right\} \middle| \eta_{m-1}, \eta_{m-2}, \dots \right] \right) \\ &= \mathbf{E} \left(\exp \left\{ -\frac{\sum_{j=1}^d \xi_j^2 h_{t_j+k}}{2} \right\} \right), \end{aligned}$$

onde $m = \max(t_1 + k, \dots, t_d + k)$.

Pela Proposição 4.2 sabemos que $\{h_t\}$ é estacionário, logo

$$\mathbf{E} \left(\exp \left\{ -\frac{\sum_{j=1}^d \xi_j^2 h_{t_j+k}}{2} \right\} \right) = \mathbf{E} \left(\exp \left\{ -\frac{\sum_{j=1}^d \xi_j^2 h_{t_j}}{2} \right\} \right),$$

e então

$$\mathbf{E} \left(\exp \left\{ i \sum_{j=1}^d \xi_j r_{t_j+k} \right\} \right) = \mathbf{E} \left(\exp \left\{ i \sum_{j=1}^d \xi_j r_{t_j} \right\} \right).$$

Concluimos que para todo $d \geq 1$ e $t_1, \dots, t_d, k \in \mathbb{Z}$

$$(r_{t_1+k}, \dots, r_{t_d+k}) \stackrel{d}{=} (r_{t_1}, \dots, r_{t_d}),$$

e portanto o processo $\{r_t\}$ é estacionário. ■

4.2 Estrutura de dependência

Como destacamos, uma das principais diferenças entre processos com inovações estáveis e processos com inovações Gaussianas, dado que a variância não existe no primeiro caso, é a indisponibilidade da função de autocovariância. Um novo conceito para a análise de dependência é apresentado em Kokoszka e Taqqu (1994), a função de autocovariância definida abaixo:

$$I_k(\xi_1, \xi_2) = -\ln \mathbf{E} \left[e^{i(\xi_1 X_t + \xi_2 X_{t-k})} \right] + \ln \mathbf{E} \left[e^{i\xi_1 X_t} \right] + \ln \mathbf{E} \left[e^{i\xi_2 X_{t-k}} \right]. \quad (4.16)$$

Como a função característica sempre existe, a função de autocovariância não requer condições de momento para o processo original X_t . No caso Gaussiano a expressão se reduz a

$$I_k(\xi_1, \xi_2) = \xi_1 \xi_2 \text{Cov}(X_t, X_{t-k}),$$

de modo que a função é proporcional à autocovariância. Para o caso em que a variância é infinita, a função ainda guarda significado prático, onde podemos avaliar a autodependência de dois processos independentes com mesmos parâmetros. Outro resultado importante é o fato da função de autocovariância decrescer exponencialmente.

Teorema 4.3. *Considere dois processos R -GARCH(r, p, q) independentes, $\{X_t\}$ e $\{Y_t\}$, com os mesmos parâmetros de distribuição. Se $\{X_t\}$ tem mais autocovariância que $\{Y_t\}$, digamos, $I_k^{(x)}(s, -s) \geq I_k^{(y)}(s, -s)$ para todo k , então o processo $\{X_t\}$ é menos autodependente que $\{Y_t\}$.*

Dem. Primeiramente considere

$$\mu_k = -\ln \mathbf{E} \left[e^{is(X_t - X_{t-k})} \right] \quad \text{e} \quad \nu_k = -\ln \mathbf{E} \left[e^{is(Y_t - Y_{t-k})} \right]. \quad (4.17)$$

$I_k^{(x)}(s, -s) \geq I_k^{(y)}(s, -s)$ se, e somente se

$$\begin{aligned}
& - \ln \mathbf{E} \left[e^{is(X_t - X_{t-k})} \right] + \ln \mathbf{E} \left[e^{isX_t} \right] + \ln \mathbf{E} \left[e^{-isX_{t-k}} \right] \geq \\
& - \ln \mathbf{E} \left[e^{is(Y_t - Y_{t-k})} \right] + \ln \mathbf{E} \left[e^{isY_t} \right] + \ln \mathbf{E} \left[e^{-isY_{t-k}} \right]
\end{aligned}$$

e portanto de (4.17) segue que

$$\mu_k \geq \nu_k$$

ou

$$\mu_k^{-1} \nu_k \leq 1. \quad (4.18)$$

Observe que $\frac{X_t - X_{t-k}}{\mu_k}$ e $\frac{Y_t - Y_{t-k}}{\nu_k}$ tem a mesma distribuição, desse modo para todo $c > 0$ temos

$$\begin{aligned}
\mathcal{P}(|X_t - X_{t-k}| > c) &= \mathcal{P}\left(\left|\frac{X_t - X_{t-k}}{\mu_k}\right| > \frac{c}{\mu_k}\right) \\
&= \mathcal{P}\left(\left|\frac{Y_t - Y_{t-k}}{\nu_k}\right| > \frac{c}{\mu_k}\right) \\
&= \mathcal{P}(|Y_t - Y_{t-k}| > c\mu_k^{-1}\nu_k) \\
&\stackrel{(4.18)}{\geq} \mathcal{P}(|Y_t - Y_{t-k}| > c).
\end{aligned}$$

A desigualdade acima significa que Y_k e Y_{t-k} tem menor probabilidade de serem diferentes do que X_t e X_{t-k} . Concluimos assim que o processo $\{Y_t\}$ é mais autodependente que o processo $\{X_t\}$. ■

Após vários comentários relevantes no que se refere à estrutura de dependência em processos com variância infinita, vamos voltar aos processos R-GARCH com inovações dadas em (4.9). Consideremos o processo R-GARCH(r,p,0) que, como vimos no Corolário 4.1, é estacionário, simétrico α -estável com $0 < \alpha < 2$. A estrutura de dependência não pode ser caracterizada pela função de autocovariância.

Utilizamos $I_k(\xi_1, \xi_2)$, no lugar da versão mais simples muitas vezes chamada codiferença $I_k(-1, 1)$ e abordada em Nowicka (1998), pensando em possibilitar uma estimação futura visto que artigos recentes (como Rosadi & Deistler (2011)) sugerem que a estimação do mesmo só produz resultados coerentes se considerarmos ξ_1 e ξ_2 pequenos. É comum o uso de $I_k(s, -s)$ ou sua versão normalizada

$$\rho_k = \frac{I_k(s, -s)}{I_0(s, -s)},$$

pois como podemos observar, sendo $I_k(s, -s) = -s^2 \text{Cov}(X_t, X_{t-k})$ no caso gaussiano, sua versão normalizada coincide com a função de autocorrelação.

Iremos considerar a função de autocovariação (4.16) e mostrar no teorema abaixo que a mesma possui forma fechada, é negativa e limitada por uma exponencial decrescente.

Teorema 4.4. *Seja Q como na Proposição 4.2. Então para um modelo R-GARCH(r,p,0) com inovações η_t dadas por (4.9) e qualquer $0 < \alpha < 2$, existe uma constante M dependendo de α , Q , ξ_1 e ξ_2 tal que*

$$0 \leq \limsup_{k \rightarrow \infty} Q^{k\frac{\alpha}{2}} |I_k(\xi_1, \xi_2)| \leq M$$

ou

$$-M \leq \limsup_{k \rightarrow \infty} Q^{k\frac{\alpha}{2}} I_k(\xi_1, \xi_2) \leq 0.$$

Dem. Como $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes e o processo R-GARCH(r,p,0) é estacionário,

$$\begin{aligned} \mathbf{E}[\exp\{i(\xi_1 r_t + \xi_2 r_{t-k})\}] &= \mathbf{E}[\exp\{i(\xi_1 r_k + \xi_2 r_0)\}] \\ &= \mathbf{E}\left(\mathbf{E}\left[\exp\left\{i(\xi_1 \sqrt{h_k} \epsilon_k + \xi_2 \sqrt{h_0} \epsilon_0)\right\}\right] \middle| \eta_{k-1}, \eta_{k-2}, \dots\right) \\ &= \mathbf{E}\left[\exp\left\{-\frac{\xi_1^2 h_k + \xi_2^2 h_0}{2}\right\}\right]. \end{aligned} \quad (4.19)$$

Como h é dado por (4.10), para todo k temos

$$\begin{aligned} \xi_1^2 h_k + \xi_2^2 h_0 &= \xi_1^2 \sum_{j=1}^{\infty} \delta_j \eta_{k-j} + \xi_2^2 \sum_{j=1}^{\infty} \delta_j \eta_{-j} \\ &= \xi_1^2 [\delta_1 \eta_{k-1} + \dots + \delta_k \eta_0] + \xi_1^2 [\delta_{k+1} \eta_{-1} + \delta_{k+2} \eta_{-2} + \dots] + \xi_2^2 [\delta_1 \eta_{-1} + \delta_2 \eta_{-2} + \dots] \\ &= \xi_1^2 \sum_{j=1}^k \delta_j \eta_{k-j} + \sum_{j=1}^{\infty} (\xi_1^2 \delta_{k+j} + \xi_2^2 \delta_j) \eta_{-j} \end{aligned}$$

e então, de (4.11), a distribuição de $\xi_1^2 h_k + \xi_2^2 h_0$ é positiva, estritamente estável com índice de estabilidade $\alpha/2$ dada abaixo:

$$\xi_1^2 h_k + \xi_2^2 h_0 \sim S_{\alpha/2} \left(2 \left[\sum_{j=1}^k |\xi_1|^\alpha \delta_j^{\alpha/2} + \sum_{j=1}^{\infty} (\xi_1^2 \delta_{k+j} + \xi_2^2 \delta_j)^{\alpha/2} \right]^{2/\alpha} \left(\cos \frac{\pi \alpha}{4} \right)^{2/\alpha}, 1, 0 \right).$$

Aplicando a fórmula para a transformada de Laplace de $\xi_1^2 h_k + \xi_2^2 h_0$ em (4.19) segue que

$$\mathbf{E}[\exp i(\xi_1 r_k + \xi_2 r_0)] = \exp \left\{ - \left[\sum_{j=1}^k |\xi_1|^\alpha \delta_j^{\alpha/2} + \sum_{j=1}^{\infty} (\xi_1^2 \delta_{k+j} + \xi_2^2 \delta_j)^{\alpha/2} \right] \right\}. \quad (4.20)$$

Com cálculos similares aos cálculos acima chegamos a

$$\mathbf{E}[\exp(i \xi_1 r_k)] = \exp \left\{ - |\xi_1|^\alpha \sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right\}, \quad (4.21)$$

$$\mathbf{E}[\exp(i \xi_2 r_0)] = \exp \left\{ - |\xi_2|^\alpha \sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right\}. \quad (4.22)$$

De (4.20), (4.21) e (4.22) segue que

$$I_k(\xi_1, \xi_2) = - \sum_{j=1}^{\infty} |\xi_1|^\alpha \delta_{k+j}^{\alpha/2} - \sum_{j=1}^{\infty} |\xi_2|^\alpha \delta_j^{\alpha/2} + \sum_{j=1}^{\infty} (\xi_1^2 \delta_{k+j} + \xi_2^2 \delta_j)^{\alpha/2}. \quad (4.23)$$

Para continuidade da demonstração lembremos um resultado importante.

Lembrete. Para $0 \leq \gamma \leq 1$ e $r, s \in \mathbb{R}$ temos

$$|r + s|^\gamma \leq |r|^\gamma + |s|^\gamma \quad (4.24)$$

e

$$||r + s|^\gamma - |r|^\gamma - |s|^\gamma| \leq 2|r|^\gamma. \quad (4.25)$$

Suponha que $k \geq 1$ é tal que $0 < \delta_j < Q^j$ para cada $j > k$. Então, após aplicarmos (4.25) em cada termo de (4.23) (com $r = \xi_1^2 \delta_{k+j}$, $s = \xi_2^2 \delta_j$ e $\gamma = \alpha/2$) temos

$$\begin{aligned} Q^{k\frac{\alpha}{2}} |I_k(\xi_1, \xi_2)| &\leq 2Q^{k\frac{\alpha}{2}} \sum_{j=1}^{\infty} |\xi_1|^\alpha \delta_{k+j}^{\alpha/2} < 2Q^{k\frac{\alpha}{2}} \sum_{j=1}^{\infty} |\xi_1|^\alpha Q^{-(k+j)\frac{\alpha}{2}} \\ &= 2|\xi_1|^\alpha \sum_{j=1}^{\infty} Q^{-j\frac{\alpha}{2}} = 2|\xi_1|^\alpha \frac{Q^{-\alpha/2}}{1 - Q^{-\alpha/2}} \\ &= |\xi_1|^\alpha K. \end{aligned} \quad (4.26)$$

É fácil perceber que se trocarmos r e s de lugar nas equações (4.24) e (4.25) a última igualdade em (4.26) terá ξ_2 no lugar de ξ_1 . Também é fácil observar que se $|\xi_1| \leq 1$ ou $|\xi_2| \leq 1$,

$$Q^{k\frac{\alpha}{2}} |I_k(\xi_1, \xi_2)| \leq K.$$

Considere $M = |\xi_1|^\alpha K$ se $|\xi_1| > 1$ e $|\xi_2| > 1$ e $M = K$ caso contrário. De (4.24) sabemos que $I_k(\xi_1, \xi_2) \leq 0$, logo

$$0 \leq \limsup_{k \rightarrow \infty} Q^{k\frac{\alpha}{2}} |I_k(\xi_1, \xi_2)| \leq M$$

ou

$$-M \leq \limsup_{k \rightarrow \infty} Q^{k\frac{\alpha}{2}} I_k(\xi_1, \xi_2) \leq 0. \quad \blacksquare$$

O Teorema 4.4 mostra que $I_k(\xi_1, \xi_2)$ tende a zero exponencialmente ou mais rápido para processos R-GARCH(r,p,0) com inovações estritamente estáveis como em (4.9).

Observação: Se $p = 0$ então

$$h_k = \sum_{j=1}^r \theta_j \eta_{k-j} = \sum_{j=1}^{\infty} \delta_j \eta_{k-j},$$

onde $\delta_j = \theta_j$ para $1 \leq j \leq r$ e $\delta_j = 0$ para $j > r$, e r_k e r_0 são independentes para $k \geq r$.

No caso do processo R-GARCH(1,1,0) com inovações dadas por (4.9) é possível encontrar uma fórmula explícita para a função de autocovariância e determinar seu comportamento assintótico exato.

Teorema 4.5. *No caso do processo R-GARCH(1,1,0) com inovações η_t dadas por (4.9) e com $\theta_1 > 0$ e $0 < \phi_1 < 1$,*

$$\lim_{k \rightarrow \infty} \phi_1^{-k \frac{\alpha}{2}} I_k(\xi_1, \xi_2) = -|\xi_1|^\alpha \frac{\theta_1^{\alpha/2}}{1 - \phi_1^{\alpha/2}},$$

i.e. a função de autocovariância é assintoticamente proporcional a $\phi_1^{k \frac{\alpha}{2}}$ e portanto tende a zero exponencialmente.

Dem. Nesse caso, a variancia condicional h_k depende de h_{k-1} e η_{k-1} , isto é

$$h_k = \theta_1 \eta_{k-1} + \phi_1 h_{k-1} = \Theta(B) \eta_k + \Phi(B) h_k,$$

onde $\Theta(z) = \theta_1 z$ e $\Phi(z) = \phi_1 z$. Como $0 < \phi_1 < 1$, o polinômio $1 - \Phi(z)$ não possui raízes no disco unitário fechado $\{z : |z| \leq 1\}$. Pela Proposição 4.2, os coeficientes δ_j 's são os coeficientes na expansão da série $\frac{\theta_1 z}{1 - \phi_1 z}$, $|z| < 1$. Como

$$\frac{\theta_1 z}{1 - \phi_1 z} = \theta_1 z \sum_{j=0}^{\infty} \phi_1^j z^j = \sum_{j=1}^{\infty} \theta_1 \phi_1^{j-1} z^j,$$

os δ_j 's são dados por

$$\delta_j = \theta_1 \phi_1^{j-1}, \quad j = 1, 2, \dots$$

Pelo Corolário 4.1, $\{r_t\}$ é um processo estacionário SaS. De (4.23) temos que para $k \geq 1$

$$\begin{aligned} I_k(\xi_1, \xi_2) &= \sum_{j=1}^{\infty} \left[(\xi_1^2 \delta_{k+j} + \xi_2^2 \delta_j)^{\alpha/2} - |\xi_1|^\alpha \delta_{k+j}^{\alpha/2} - |\xi_2|^\alpha \delta_j^{\alpha/2} \right] \\ &= \sum_{j=1}^{\infty} \left[\left(\xi_1^2 \theta_1 \phi_1^{k+j-1} + \xi_2^2 \theta_1 \phi_1^{j-1} \right)^{\alpha/2} - |\xi_1|^\alpha \left(\theta_1 \phi_1^{k+j-1} \right)^{\alpha/2} - |\xi_2|^\alpha \left(\theta_1 \phi_1^{j-1} \right)^{\alpha/2} \right] \\ &= -\theta_1^{\alpha/2} \sum_{j=1}^{\infty} \left[\phi_1^{(j-1) \frac{\alpha}{2}} \left(|\xi_1|^\alpha \phi_1^{k \frac{\alpha}{2}} + |\xi_2|^\alpha - \left(\xi_1^2 \phi_1^k + \xi_2^2 \right)^{\alpha/2} \right) \right] \\ &= -\theta_1^{\alpha/2} \left[|\xi_1|^\alpha \phi_1^{k \frac{\alpha}{2}} + |\xi_2|^\alpha - \left(\xi_1^2 \phi_1^k + \xi_2^2 \right)^{\alpha/2} \right] \sum_{j=1}^{\infty} \phi_1^{(j-1) \frac{\alpha}{2}} \\ &= -\frac{\theta_1^{\alpha/2}}{1 - \phi_1^{\alpha/2}} \left[|\xi_1|^\alpha \phi_1^{k \frac{\alpha}{2}} + |\xi_2|^\alpha - \left(\xi_1^2 \phi_1^k + \xi_2^2 \right)^{\alpha/2} \right]. \end{aligned}$$

Portanto,

$$\begin{aligned} \lim_{k \rightarrow \infty} \phi_1^{-k \frac{\alpha}{2}} I_k(\xi_1, \xi_2) &= -\frac{\theta_1^{\alpha/2}}{1 - \phi_1^{\alpha/2}} \lim_{k \rightarrow \infty} \left[|\xi_1|^\alpha + |\xi_2|^\alpha \phi_1^{-k \frac{\alpha}{2}} - \left(\xi_1^2 + \xi_2^2 \phi_1^{-k} \right)^{\alpha/2} \right] \\ &= -\frac{\theta_1^{\alpha/2}}{1 - \phi_1^{\alpha/2}} \lim_{k \rightarrow \infty} \left[|\xi_1|^\alpha + \frac{|\xi_2|^\alpha - \left(\phi_1^k \xi_1^2 + \xi_2^2 \right)^{\alpha/2}}{\phi_1^{k \frac{\alpha}{2}}} \right], \end{aligned}$$

e para completar a demonstração basta notar que

$$\lim_{k \rightarrow \infty} \left[\frac{|\xi_2|^\alpha - (\phi^k \xi_1^2 + \xi_2^2)^{\alpha/2}}{\phi_1^{k \frac{\alpha}{2}}} \right] = 0.$$

■

Para ilustrarmos um processo R-GARCH(1,1,0) simulamos o mesmo com diferentes valores para os parâmetros α , θ_1 e ϕ_1 . Nas Figuras 4.1 a 4.11, seguem o gráfico dos retornos, da volatilidade e a função de autocovariação teórica dos diversos processos simulados. Lembrando que a forma fechada para a função de autocovariação do processo R-GARCH(1,1,0) foi encontrada na demonstração do teorema anterior e é dada por

$$I_k(\xi_1, \xi_2) = -\frac{\theta_1^{\alpha/2}}{1 - \phi_1^{\alpha/2}} \left[|\xi_1|^\alpha \phi_1^{k \frac{\alpha}{2}} + |\xi_2|^\alpha - (\xi_1^2 \phi_1^k + \xi_2^2)^{\alpha/2} \right]. \quad (4.27)$$

Como se pode observar nas figuras, os processos R-GARCH podem variar muito de acordo com seus parâmetros. O parâmetro α tem influência muito grande nas caudas e consequentemente na variação dos retornos e da volatilidade. O parâmetro θ_1 também tem forte influência na variação dos retornos e da volatilidade. O parâmetro ϕ_1 tem pequena influência na variação dos retornos e da volatilidade, porém tem forte influência sobre o decaimento exponencial da função de autocovariação; à medida que diminui o decaimento é mais rápido.

Figura 4.1: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.9$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$.

Figura 4.2: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$.

Figura 4.3: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.5$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$.

Figura 4.4: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.3$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$.

Figura 4.5: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.1$, $\theta_1 = 0.3$ e $\phi_1 = 0.6$.

Figura 4.6: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.1$ e $\phi_1 = 0.6$.

Figura 4.7: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.01$ e $\phi_1 = 0.6$.

Figura 4.8: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.001$ e $\phi_1 = 0.6$.

Figura 4.9: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.3$ e $\phi_1 = 0.5$.

Figura 4.10: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.3$ e $\phi_1 = 0.4$.

Figura 4.11: Retornos e volatilidade de um processo R-GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.3$ e $\phi_1 = 0.2$.

4.3 Inferência indireta para o modelo R-GARCH(1,1,0)

Uma vez que especificarmos um modelo auxiliar adequado, a estimação indireta dos parâmetros de um processo R-GARCH(1,1,0) pode ser implementada tomando como base a simulação de variáveis aleatórias α -estáveis dadas em (1.14). Nesse sentido, iremos propor na seção seguinte um modelo auxiliar a partir do qual obtemos as estimativas. Para averiguar as propriedades e a eficiência do método faremos uma análise de simulações variando a quantidade de réplicas e observações. Em seguida, apresentaremos a análise de duas aplicações empíricas aos dados de 15 minutos da ação Telemar e do índice de ações Ibovespa e comparamos com um modelo popular na literatura.

4.3.1 O modelo auxiliar

A idéia inicial era utilizar um modelo conhecido e próximo a um processo R-GARCH(1,1,0). Como já adiantamos, precisamos de um modelo cuja estimação dos parâmetros envolvidos possam ser obtidos por máxima verossimilhança. Um requisito interessante para que possamos economizar tempo computacional é que a convergência para os estimadores, no modelo auxiliar, seja rápida já que iremos usar um método baseado em simulações. Pensando dessa maneira, um candidato interessante para ser o modelo auxiliar seria um modelo GARCH(1,1). Como o modelo R-GARCH admite caudas bem mais pesadas que um modelo GARCH padrão, decidi utilizar como modelo auxiliar um GARCH(1,1) com inovações t de Student. Nesse caso, a logverossimilhança condicional às primeiras m observações é dada por

$$\ln \mathcal{L}(r_{m+1}, \dots, r_t | \alpha, \beta, r_1, \dots, r_m) \propto \sum_{t=m+1}^T \left[\ln \left(\Gamma \left(\frac{\nu+1}{2} \right) \right) - \ln \left(\Gamma \left(\frac{\nu}{2} \right) \right) - \frac{1}{2} \ln((\nu-2)h_t) \right],$$

onde $\Gamma(\cdot)$ é a função gamma e $\nu > 2$ é o número de graus de liberdade da distribuição t de Student.

O método de inferência indireta proposto por Gouriéroux et. al. (1993) foi então implementado (via Matlab) para encontrar os estimadores do modelo R-GARCH(1,1,0). Relembremos os dois modelos, auxiliar e original, de modo que seus respectivos espaços paramétricos fiquem claros daqui em diante.

- (Modelo auxiliar) Processo GARCH(1,1) com inovações ϵ_t distribuídas segundo t de Student:

$$\begin{aligned} r_t &= \sqrt{h_t} \epsilon_t, \quad t \in \mathbb{Z}, \\ h_t &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1}, \end{aligned}$$

onde $\alpha_0 > 0$ e $(\alpha_1 + \beta_1) < 1$.

- (Modelo original) Processo R-GARCH(1,1,0):

$$\begin{aligned} r_t &= \sqrt{h_t} \epsilon_t, \quad t = 0, \pm 1, \pm 2, \dots, \\ h_t &= \theta_1 \eta_{t-1} + \phi_1 h_{t-1}, \end{aligned}$$

onde as inovações ϵ_t são v.a's i.i.d. normais padrão, as inovações η_t são v.a's i.i.d. distribuídas segundo (4.9) e $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes.

Dessa maneira, temos os parâmetros do modelo auxiliar simbolizados por $\zeta = (\alpha_0, \alpha_1, \beta_1)$ assim como os parâmetros do modelo original simbolizados por $\theta = (\alpha, \theta_1, \phi_1)$. Ao tratarmos os dados previamente, os testes indicaram três ou quatro graus de liberdade para a distribuição t de Student do modelo auxiliar.

Como $\dim \zeta = \dim \theta$, motivados pela Propriedade 3.1, escolhemos a matriz simétrica não negativa definida Ω como sendo a matriz identidade. O parâmetro α_0 do modelo auxiliar está diretamente ligado aos parâmetros α e θ_1 de modo que a função de ligação é claramente não bijetora. Apesar disso, como veremos adiante, os estimadores de inferência indireta tem distribuição consideravelmente comportada e muitas vezes aparentam normalidade.

4.3.2 Simulações

Nesta seção conduzimos algumas simulações no intuito de averiguar as propriedades dos estimadores de inferência indireta dos parâmetros α , θ_1 e ϕ_1 do modelo R-GARCH(1,1,0). Os programas e rotinas foram implementados no Matlab e processados em um Pentium Core 2 Duo 2.26GHz com 4Gb de memória RAM.

Simulações com 100 réplicas

As primeiras simulações foram baseadas em 100 réplicas independentes, variando as observações em 25000, 50000 e 100000. A idéia ao variarmos T (números de observações) é averiguar o que ocorre com a distribuição dos estimadores visto que a função de ligação não é bijetora e assim não podemos utilizar a Propriedade 3.3 para obter a distribuição limite dos parâmetros. Os primeiros resultados foram obtidos ao simularmos um processo R-GARCH(1,1,0) com parâmetros $\alpha = 1.6$, $\theta_1 = 0.5$ e $\phi_1 = 0.4$. Seguem na Tabela 4.1 a média e erro padrão Monte Carlo estimados onde $\nu = 4$ para o modelo auxiliar.

	$\alpha = 1.6$	$\theta_1 = 0.5$	$\phi_1 = 0.4$
T = 25000	1.5944 0.0214	0.5080 0.0310	0.3816 0.0204
T = 50000	1.6030 0.0140	0.4884 0.0212	0.4122 0.0133
T = 100000	1.6059 0.0124	0.5031 0.0192	0.4124 0.0163

Tabela 4.1: Média e erro padrão MC para a estimação de um R-GARCH(1,1,0) com 100 réplicas T observações.

Como podemos observar, as estimativas obtidas para os parâmetros do modelo via simulações estão bem próximas do seu valor real com erro padrão consideravelmente baixo. Mais à frente iremos obter outras estimativas com diferentes valores para α , θ_1 e ϕ_1 .

De posse das estimativas dos parâmetros avaliamos as respectivas distribuições dos estimadores de inferência indireta para α , θ_1 e ϕ_1 à medida que aumentamos T. Nas figuras seguintes temos o histograma e a densidade ajustada das estimativas dos parâmetros indicados na Tabela 4.1.

Figura 4.12: Histograma e densidade ajustada das estimativas do parâmetro α com 100 réplicas e diferentes valores para T .

Figura 4.13: Histograma e densidade ajustada das estimativas do parâmetro θ_1 com 100 réplicas e diferentes valores para T .

Figura 4.14: Histograma e densidade ajustada das estimativas do parâmetro ϕ_1 com 100 réplicas e diferentes valores para T .

Simulações com 500 réplicas

Após observarmos, nas Figuras 4.12, 4.13 e 4.14, os histogramas dos estimadores de inferência indireta para α , θ_1 e ϕ_1 com 100 réplicas, concluímos que estes possuem distribuição relativamente comportada. Ainda pensando em observar o comportamento dessas estimativas, obtivemos novas estimativas para os parâmetros agora tomando 500 réplicas independentes. Novamente os resultados foram obtidos com os mesmos dados simulados do processo R-GARCH(1,1,0) com parâmetros $\alpha = 1.6$, $\theta_1 = 0.5$ e $\phi_1 = 0.4$ e seguem na Tabela 4.2 a média e erro padrão Monte Carlo dos valores estimados onde $\nu = 4$ para o modelo auxiliar.

	$\alpha = 1.6$	$\theta_1 = 0.5$	$\phi_1 = 0.4$
T = 25000	1.5967	0.5069	0.3802
	0.0138	0.0208	0.0149
T = 50000	1.6015	0.4854	0.4125
	0.0171	0.0244	0.0171
T = 100000	1.6055	0.5009	0.4111
	0.0129	0.0184	0.0151

Tabela 4.2: Média e erro padrão MC para a estimação de um R-GARCH(1,1,0) com 500 réplicas e T observações.

Como podemos observar, houve uma melhora na estimação dos parâmetros que ficaram ainda mais próximos dos valores simulados. Ainda interessados em avaliar a distribuição dos estimadores à medida que aumentamos T, apresentamos nas Figuras 4.15, 4.16 e 4.17 o histograma e a densidade ajustada das estimativas dos parâmetros indicados na Tabela 4.2. A distribuição das estimativas com 500 réplicas se mostra ainda mais comportada e aparentemente normal. Como forma de averiguar o potencial do método, obtemos mais algumas estimativas variando α , θ_1 e ϕ_1 .

Apesar do comportamento da distribuição das estimativas ser consideravelmente melhor se tomarmos 500 réplicas como base, as estimativas para os outros valores foram baseadas em 100 réplicas e 100000 observações por uma questão de economia computacional já que as mesmas se aproximam muito bem do valor real e tem pequena diferença se tomarmos 500 réplicas como base. Podemos ver os valores escolhidos e suas respectivas estimativas na Tabela 4.3 seguinte.

	$\alpha = 1.9$	$\theta_1 = 0.1$	$\phi_1 = 0.8$
Média	1.8996	0.1030	0.8060
Erro Padrão	0.0077	0.0094	0.0122
	$\alpha = 1.8$	$\theta_1 = 0.4$	$\phi_1 = 0.3$
Média	1.8165	0.4069	0.3088
Erro Padrão	0.0119	0.0100	0.0179
	$\alpha = 1.7$	$\theta_1 = 0.3$	$\phi_1 = 0.6$
Média	1.7030	0.2942	0.6080
Erro Padrão	0.0229	0.0150	0.0149

Tabela 4.3: Média e erro padrão MC para a estimação de um R-GARCH(1,1,0) com 100 réplicas e 100000 observações.

Figura 4.15: Histograma e densidade ajustada das estimativas do parâmetro α com 500 réplicas e diferentes valores para T .

Figura 4.16: Histograma e densidade ajustada das estimativas do parâmetro θ_1 com 500 réplicas e diferentes valores para T .

Figura 4.17: Histograma e densidade ajustada das estimativas do parâmetro ϕ_1 com 500 réplicas e diferentes valores para T .

4.3.3 Aplicação empírica (TLM)

Com o intuito de avaliar a performance e realizar uma análise residual do modelo R-GARCH(1,1,0) com dados reais, iremos ajustá-lo aos dados intradiários de 15 minutos da Telemar (TLM15m). Considerando os logretornos de TLM15m no período de 02/01/2002 até 14/01/2005, totalizando 21428 observações, seguem na Figura 4.18 abaixo o seu gráfico juntamente com o seu variograma.

Figura 4.18: Logretornos e variograma de TLM15m

Podemos observar em seu variograma uma indicação de que o processo possui variância infinita de modo que o modelo R-GARCH(1,1,0) aparece como um candidato adequado. Com base em 100 réplicas independentes e $S = 4$, ajustamos então o modelo R-GARCH(1,1,0) para os logretornos de TLM15m, tomando $\nu = 4$ no modelo auxiliar, e apresentamos na Tabela 4.4 abaixo os valores da média e erro padrão Monte Carlo dos parâmetros α , θ_1 e ϕ_1 .

	α	θ_1	ϕ_1
Média	1.5879	7.9040×10^{-7}	0.7767
Erro Padrão	0.0072	2.0057×10^{-8}	0.0026

Tabela 4.4: Média e erro padrão MC ao ajustarmos um R-GARCH(1,1,0) aos logretornos de TLM15m.

Para efeitos de comparação ajustamos aos dados um modelo GARCH(0,1) (lembrando que o modelo R-GARCH(1,1,0) é uma espécie de GARCH(0,1) com o termo constante aleatorizado) e comparamos o erro quadrático médio obtido pelos dois modelos. Os valores seguem na Tabela 4.5 abaixo:

EQM para GARCH(0,1)	EQM para R-GARCH(1,1,0)
1.9150×10^{-5}	2.2942×10^{-7}

Tabela 4.5: Erro quadrático médio ao ajustarmos um modelo GARCH(0,1) e R-GARCH(1,1,0) aos logretornos de TLM15m

Como podemos observar, o erro quadrático médio do modelo R-GARCH(1,1,0) ajustado é muito menor que o erro quadrático médio do modelo GARCH(0,1) ajustado. Apesar da considerável melhora

no que diz respeito ao erro quadrático médio, ao observarmos os resíduos do modelo R-GARCH(1,1,0) ajustado podemos observar de imediato que a hipótese de normalidade para os mesmos não parece adequada. Na Figura 4.19 abaixo temos um gráfico dos quantis teóricos de uma distribuição normal versus os quantis dos resíduos do modelo que confirmam nossa suspeita. Com o intuito de solucionar esta questão residual e melhorar o modelo R-GARCH(1,1,0), após uma série de observações, iremos propor, nos capítulos seguintes, algumas alterações na distribuição dos resíduos.

Figura 4.19: QxQ plot dos resíduos do modelo R-GARCH(1,1,0) ajustado a TLM15m versus quantis da Normal Padrão.

4.3.4 Aplicação empírica (IBV)

Para reforçar a análise da performance e, novamente, observar o comportamento residual do modelo R-GARCH(1,1,0) com dados reais, iremos ajustá-lo, agora, aos dados intradiários de 15 minutos do índice Ibovespa (IBV15m) no período de 06/04/1998 até 01/06/1998, totalizando 37960 observações. Seguem na Figura 4.20 abaixo o seu gráfico juntamente com o seu variograma.

Figura 4.20: Logretornos e variograma de IBV15m

Como no caso de TLM15m, o variograma indica que o processo possui variância infinita de modo que o modelo R-GARCH(1,1,0) aparece como um candidato adequado. Com base em 100 réplicas

independentes e $S = 5$, ajustamos então o modelo R-GARCH(1,1,0) para os logretornos de IBV15m, tomando $\nu = 3$ no modelo auxiliar, e apresentamos na Tabela 4.6 abaixo os valores da média e erro padrão Monte Carlo dos parâmetros α , θ_1 e ϕ_1 .

	α	θ_1	ϕ_1
Média	1.2867	2.7461×10^{-7}	0.7121
Erro Padrão	0.0179	4.7875×10^{-8}	0.0022

Tabela 4.6: Média e erro padrão MC ao ajustarmos um R-GARCH(1,1,0) aos logretornos de IBV15m.

Novamente ajustamos aos dados um modelo GARCH(0,1) e comparamos o erro quadrático médio obtido pelos dois modelos. Os valores seguem na Tabela 4.7 abaixo:

EQM para GARCH(0,1)	EQM para R-GARCH(1,1,0)
1.6116×10^{-5}	3.5512×10^{-6}

Tabela 4.7: Erro quadrático médio ao ajustarmos um modelo GARCH(0,1) e R-GARCH(1,1,0) aos logretornos de IBV15m

Observamos mais uma vez que o erro quadrático médio do modelo R-GARCH(1,1,0) ajustado é muito menor que o erro quadrático médio do modelo GARCH(0,1) ajustado, todavia ao observarmos os resíduos do modelo R-GARCH(1,1,0) ajustado podemos observar de imediato que a hipótese de normalidade para os mesmos não parece adequada. Na Figura 4.21 abaixo temos um gráfico dos quantis teóricos de uma distribuição normal versus os quantis dos resíduos do modelo que confirmam nossa suspeita.

Figura 4.21: QxQ plot dos resíduos do modelo R-GARCH(1,1,0) ajustado a IBV15m versus quantis da Normal Padrão.

Dado que nos dois casos empíricos analisados a hipótese de normalidade dos ϵ_t 's parece inadequada, iremos seguir a idéia e propor nos capítulos seguintes o modelo R-GARCH substituindo a hipótese de normalidade dos resíduos pela hipótese de que os mesmos possuem distribuição t de Student e a esse modelo daremos o nome de R_t -GARCH.

Ainda na intenção criar um modelo melhor e mais geral, propomos no Capítulo 6 que os resíduos

possuam distribuição α -estável sob algumas condições e a este modelo, que generaliza a classe de processos R-GARCH, daremos o nome de R_s -GARCH.

Capítulo 5

Processos R_t -GARCH

No capítulo anterior estimamos, com sucesso, os parâmetros de um processo R-GARCH(1,1,0) assim como aplicações empíricas com dados reais. Como podemos observar, há uma melhora significativa em termos de erro quadrático médio se compararmos com os modelos GARCH usuais. Apesar da boa performance, o modelo, assim como na maioria dos modelos existentes, possui uma incoerência de proposta no que diz respeito aos resíduos e, como observamos nos gráficos QxQ plot, a hipótese de normalidade destes não parece adequada. Com a intenção de atribuir uma hipótese que melhor se adapte aos resíduos, neste capítulo iremos explorar o potencial da distribuição t de Student como hipótese para os resíduos. Para averiguar as propriedades e a eficiência do método faremos uma análise de simulações e novamente apresentaremos a análise empírica dos dados de 15 minutos da ação Telemar e do índice de ações Ibovespa.

5.1 Considerações e exemplos

Apesar de não obtermos resultados teóricos decidimos averiguar a capacidade de estimação do modelo assim como refazer as aplicações do capítulo anterior sobre esta hipótese. A essa classe de processos chamaremos R_t -GARCH(r,p,q) com definição similar à Definição 4.1 apenas atribuindo distribuição t de Student aos ϵ_t 's. O processo o qual analisaremos, e será o modelo original na estimação indireta, é o equivalente a um processo R-GARCH(1,1,0), o processo R_t -GARCH(1,1,0) com inovações η_t 's dadas por (4.9):

$$\begin{aligned} r_t &= \sqrt{h_t} \epsilon_t, \quad t = 0, \pm 1, \pm 2, \dots, \\ h_t &= \theta_1 \eta_{t-1} + \phi_1 h_{t-1}. \end{aligned}$$

Para ilustrar esta classe simulamos algumas trajetórias com diferentes valores para α , θ_1 e ϕ_1 . Nas figuras seguintes seguem os gráficos dos retornos e da volatilidade das trajetórias simuladas.

Figura 5.1: Retornos e volatilidade de um processo R_t -GARCH(1,1,0) para $\alpha = 1.8$, $\theta_1 = 0.001$ e $\phi_1 = 0.7$.

Figura 5.2: Retornos e volatilidade de um processo R_t -GARCH(1,1,0) para $\alpha = 1.7$, $\theta_1 = 0.0001$ e $\phi_1 = 0.8$.

5.2 Inferência indireta para o modelo R_t -GARCH(1,1,0)

Novamente iremos empregar o método de estimação indireta de Gouriéroux et. al. (1993) para encontrar estimativas dos parâmetros do modelo original R_t -GARCH(1,1,0). O modelo auxiliar utilizado é o mesmo modelo auxiliar que empregamos na estimação dos parâmetros do modelo R-GARCH e o relembramos abaixo.

$$\begin{aligned} r_t &= \sqrt{h_t} \epsilon_t, \quad \epsilon_t \sim t(\nu), \quad t \in \mathbb{Z}, \\ h_t &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1}, \end{aligned}$$

Os parâmetros do modelo auxiliar serão também simbolizados por $\zeta = (\alpha_0, \alpha_1, \beta_1)$ assim como os parâmetros do modelo original simbolizados por $\theta = (\alpha, \theta_1, \phi_1)$. Ao tratarmos os dados previamente, os testes indicaram três ou quatro graus de liberdade para a distribuição t de Student do modelo

auxiliar.

Como $\dim \zeta = \dim \theta$, motivados pela Propriedade 3.1, escolhemos a matriz simétrica não negativa definida Ω como sendo a matriz identidade. O parâmetro α_0 do modelo auxiliar está diretamente ligado aos parâmetros α e θ_1 de modo que a função de ligação é claramente não bijetora. Apesar disso, como veremos nas simulações, os estimadores de inferência indireta são muito próximos do seu valor real.

5.2.1 Simulações

Nesta seção conduzimos, novamente, simulações no intuito de averiguar a estimação dos parâmetros pelo método de inferência indireta. Com a experiência que temos do caso R-GARCH, decidimos utilizar 100 réplicas com 100000 observações para obter as estimativas visto que o comportamento dos estimadores é bem similar. Escolhemos alguns valores para os parâmetros, α , θ_1 e ϕ_1 e apresentamos na Tabela 5.1 a média e erro padrão Monte Carlo estimados onde $\nu = 4$ no modelo auxiliar.

	$\alpha = 1.9$	$\theta_1 = 0.001$	$\phi_1 = 0.3$
Média	1.8834	0.0011	0.3398
Erro Padrão	0.0047	1.2875×10^{-4}	0.0199
	$\alpha = 1.8$	$\theta_1 = 0.001$	$\phi_1 = 0.5$
Média	1.8062	0.001703	0.5175
Erro Padrão	0.0088	1.4108×10^{-4}	0.0126
	$\alpha = 1.7$	$\theta_1 = 0.0001$	$\phi_1 = 0.7$
Média	1.6902	0.0001153	0.7040
Erro Padrão	0.0094	5.9274×10^{-5}	0.0069

Tabela 5.1: Média e erro padrão MC para a estimação de um R_t -GARCH(1,1,0) com 100 réplicas e 100000 observações.

5.2.2 Aplicação empírica (TLM)

Com o intuito de avaliar a performance, realizar uma análise residual do modelo R_t -GARCH(1,1,0) com dados reais e comparar esta performance com a dos processos R-GARCH(1,1,0), novamente ajustamos esse modelo aos dados intradiários de 15 minutos da Telemar (TLM15m) com distribuição t de Student e quatro graus de liberdade com base em 100 réplicas independentes e $S = 5$. Apresentamos na Tabela 5.2 os valores da média e erro padrão Monte Carlo dos parâmetros α , θ_1 e ϕ_1 onde $\nu = 4$ no modelo auxiliar.

	α	θ_1	ϕ_1
Média	1.5879	7.9040×10^{-7}	0.7767
Erro Padrão	0.0072	2.0057×10^{-8}	0.0026

Tabela 5.2: Média e erro padrão MC ao ajustarmos um R_t -GARCH(1,1,0) aos logretornos de TLM15m.

Para efeitos de comparação seguem na Tabela 5.3 abaixo o erro quadrático médio para os modelos GARCH(0,1), R-GARCH(1,1,0) e R_t -GARCH(1,1,0):

EQM para GARCH(0,1)	1.9150×10^{-5}
EQM para R-GARCH(1,1,0)	2.2942×10^{-7}
EQM para R_t-GARCH(1,1,0)	7.4659×10^{-8}

Tabela 5.3: Erro quadrático médio ao ajustarmos os modelos GARCH(0,1), R-GARCH(1,1,0) e R_t -GARCH(1,1,0) aos logretornos de TLM15m

Como podemos observar, o erro quadrático médio do modelo R_t -GARCH(1,1,0) ajustado é menor que o erro quadrático médio dos modelos GARCH(0,1) e R-GARCH(1,1,0) ajustados. Com o objetivo de avaliar a melhora do modelo em comparação com a hipótese de normalidade assumida sobre os resíduos para os modelos R-GARCH(1,1,0), ilustramos abaixo o QxQ plot dos resíduos do modelo R_t -GARCH(1,1,0) ajustado aos logretornos de TLM15m contra os quantis de uma distribuição t de Student com quatro graus de liberdade. Podemos ver que a hipótese parece bem mais adequada do que a hipótese de normalidade dos modelos R-GARCH(1,1,0).

Figura 5.3: QxQ plot dos resíduos do modelo R_t -GARCH(1,1,0) ajustado a TLM15m versus quantis de um t de Student com quatro graus de liberdade.

5.2.3 Aplicação empírica (IBV)

Mais uma vez, estamos interessados em avaliar a performance, realizar uma análise residual do modelo R_t -GARCH(1,1,0) com dados reais e comparar esta performance com a dos processos R-GARCH(1,1,0). Nessa direção, seguindo a mesma idéia aplicada no exemplo anterior, ajustamos o modelo R_t -GARCH(1,1,0) aos logretornos intradiários de 15 minutos do índice de ações brasileiro Ibovespa (IBV15m) com base em 100 réplicas independentes e $S = 4$ vetores. Neste caso, a hipótese é de que os resíduos possuem distribuição t de Student com três graus de liberdade. Apresentamos na Tabela 5.4 abaixo os valores da média e erro padrão Monte Carlo dos parâmetros α , θ_1 e ϕ_1 estimados, onde $\nu = 3$ no modelo auxiliar, e também a comparação entre os erros quadrático médio para os respectivos modelos GARCH(0,1), R-GARCH(1,1,0) e R_t -GARCH(1,1,0) na Tabela 5.5.

	α	θ_1	ϕ_1
Média	1.4994	1.4562×10^{-7}	0.6765
Erro Padrão	0.0058	1.7135×10^{-8}	0.0053

Tabela 5.4: Média e erro padrão MC ao ajustarmos um R_t -GARCH(1,1,0) aos logretornos de IBV15m.

EQM para GARCH(0,1)	2.4778×10^{-5}
EQM para R-GARCH(1,1,0)	3.5512×10^{-6}
EQM para R_t-GARCH(1,1,0)	6.7555×10^{-7}

Tabela 5.5: Erro quadrático médio ao ajustarmos os modelo GARCH(0,1), R-GARCH(1,1,0) e R_t -GARCH(1,1,0) aos logretornos de IBV15m

Como podemos observar, o erro quadrático médio do modelo R_t -GARCH(1,1,0) ajustado é menor que o erro quadrático médio dos modelos GARCH(0,1) e R-GARCH(1,1,0) ajustados. Novamente com o objetivo de avaliar a melhora do modelo no que diz respeito à hipótese sobre os resíduos, abaixo ilustramos o QxQ plot dos resíduos do modelo R_t -GARCH(1,1,0) ajustado a TLM15m contra os quantis de uma distribuição t de Student com três graus de liberdade. Podemos ver que a hipótese parece bem mais adequada do que a hipótese de normalidade dos modelos R-GARCH(1,1,0).

Figura 5.4: QxQ plot dos resíduos do modelo R_t -GARCH(1,1,0) ajustado a IBV15m versus quantis de um t de Student com três graus de liberdade.

Capítulo 6

Processos R_s -GARCH

Como já destacamos em várias ocasiões, muitos dados reais como dados financeiros, sinais de áudio, etc, apresentam caudas pesadas e uma distribuição que capte essa propriedade das caudas passa a ser de suma importância. Os processos R-GARCH aparecem como alternativa nesse sentido, entretanto, outra característica muito presente, principalmente em dados financeiros, é o fato desses dados reais possuírem distribuição leptocúrtica.

Utilizamos no capítulo anterior a distribuição t de Student como alternativa para melhorar a questão residual dos processos R-GARCH, mas como sabemos, essa classe de distribuições não tem distribuição leptocúrtica. Dessa maneira, propomos neste capítulo a hipótese de estabilidade também para os resíduos de modo que teremos a possibilidade de escolher a forma da nossa distribuição, sendo ela com cauda mais ou menos pesada e mais ou menos leptocúrtica. A esta classe daremos o nome R_s -GARCH e neste capítulo iremos propor a inferência indireta para estimar os parâmetros do modelo assim como generalizar alguns resultados visto que esta classe claramente generaliza a classe de modelos R-GARCH.

Definição 6.1. Um modelo R_s -GARCH(r, p, q), $r, p, q \in \mathbb{N}$ é definido por

$$r_t = \sqrt[r]{h_t} \epsilon_t, \quad 1 < \lambda \leq 2, \quad t = 0, \pm 1, \pm 2, \dots, \quad (6.1)$$

$$h_t = \sum_{i=1}^r \theta_i \eta_{t-i} + \sum_{j=1}^p \phi_j h_{t-j} + \sum_{k=1}^q \psi_k r_{t-k}^2, \quad (6.2)$$

onde

$$r \geq 1, \theta_r > 0, \theta_i \geq 0, i = 1, \dots, r-1, \quad (6.3)$$

$$p \geq 0, \phi_p > 0, \phi_j \geq 0, j = 1, \dots, p-1, \quad (6.4)$$

$$q \geq 0, \psi_q > 0, \psi_k \geq 0, k = 1, \dots, q-1, \quad (6.5)$$

as inovações ϵ_t são v.a's i.i.d. com distribuição $S_\lambda(\sigma, 0, 0)$, as inovações η_t são v.a's i.i.d. positivas e $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes.

A primeira observação pertinente é a de que os processos R_s -GARCH generalizam os processos R-GARCH com distribuição normal para os resíduos. Para tal, basta tomarmos $\lambda = 2$ e $\sigma = 1/\sqrt{2}$.

Como esperamos, a equação (6.2) também pode ser escrita simbolicamente na forma compacta

$$h_t = \Theta(B)\eta_t + \Phi(B)h_t + \Psi(B)r_t^2,$$

onde Θ , Φ e Ψ são

$$\begin{aligned}\Theta(B) &= \theta_1 B + \theta_2 B^2 + \cdots + \theta_r B^r, \\ \Phi(B) &= \phi_1 B + \phi_2 B^2 + \cdots + \phi_p B^p, \\ \Psi(B) &= \psi_1 B + \psi_2 B^2 + \cdots + \psi_q B^q\end{aligned}$$

e B é o operador retroativo definido por

$$B^i X_t = X_{t-i}, \quad i = 0, \pm 1, \pm 2, \dots$$

Vamos agora explorar a distribuição condicional dos processos R_s -GARCH. Para tal, consideremos novamente $\mathcal{F}_{t-1} = \sigma\{\epsilon_s, \eta_s : s \leq t-1, s \in \mathbb{Z}\}$.

Propriedade 6.1. *A distribuição de r_t condicionada a $\mathcal{F}_{t-1}$ é α -estável*

$$r_t | \mathcal{F}_{t-1} \sim S_\lambda(\sigma h_t, 0, 0).$$

Dem. A distribuição condicional de r_t segue imediatamente da distribuição de ϵ_t e do fato de que h_t é $\mathcal{F}_{t-1}$ -mensurável. ■

Daí temos que a esperança condicional $\mathbf{E}(r_t | \mathcal{F}_{t-1})$ é constante e igual a zero e para o caso em que $\lambda = 2$ segue que $\text{Var}(r_t | \mathcal{F}_{t-1}) = \mathbf{E}(r_t^2 | \mathcal{F}_{t-1}) = 2\sigma^2 h_t$. Para outros valores de λ sabemos que o segundo momento não existe para distribuições estáveis.

Como nos modelos R-GARCH, a variância condicional h_t dos modelos R_s -GARCH depende do passado como função linear das inovações passadas $\eta_{t-1}, \dots, \eta_{t-r}$, das variâncias condicionais $h_{t-1}, \dots, h_{t-p}$ e também das observações passadas dos retornos ao quadrado $r_{t-1}^2, \dots, r_{t-q}^2$.

6.1 Processos R_s -GARCH com inovações estáveis

Aqui também iremos considerar os η_t 's estáveis e demonstrar alguns resultados semelhantes aos dos modelos R-GARCH.

Vamos considerar o caso R_s -GARCH(r,p,0) onde as inovações η_t 's são estritamente estáveis e totalmente assimétricas à direita com distribuição

$$\eta_t \sim S_{\alpha/2} \left(2 \left(\cos \frac{\pi\alpha}{4} \right)^{2/\alpha}, 1, 0 \right), \quad (6.6)$$

onde $0 < \alpha < 2$. Supomos também que os polinômios $\Theta(B)$ e $1 - \Phi(B)$ não tem raízes comuns e o polinômio $1 - \Phi(B)$ não tem raízes no disco unitário $\{z : |z| \leq 1\}$. Essa hipótese garante algumas propriedades importantes e similares aos processos R-GARCH e serão sumarizadas nas propriedades que se seguem.

Proposição 6.1. *Se os polinômios $\Theta(B)$ e $1 - \Phi(B)$ não tem raízes em comum, então a variância condicional h_t para R_s -GARCH($r, p, 0$) com inovações dadas por (6.6) pode ser representada na forma*

$$h_t = \sum_{j=1}^{\infty} \delta_j \eta_{t-j}, \quad q.c. \quad (6.7)$$

com δ_j 's positivos¹ satisfazendo $\delta_j < Q^j$ eventualmente², $Q > 1$, se e somente se o polinômio $1 - \Phi(B)$ não tem raízes no disco unitário fechado $\{z : |z| \leq 1\}$. A sequência h_t é então estacionária, estritamente estável com índice de estabilidade $\alpha/2$ e

$$h_t \sim S_{\alpha/2} \left(2 \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right)^{2/\alpha} \left(\cos \frac{\pi\alpha}{4} \right)^{2/\alpha}, 1, 0 \right). \quad (6.8)$$

Os δ_j 's são os coeficientes na série expandida de $\frac{\Theta(B)}{1-\Phi(B)}$, $|B| < 1$.

A demonstração da Proposição 6.1 é rigorosamente análoga à da Proposição 4.2.

Teorema 6.1. *A distribuição não condicional de r_t no processo R_s -GARCH($r, p, 0$) com inovações estáveis dadas em (6.6) é simétrica estável*

$$r_t \sim S_{\frac{\lambda\alpha}{2}} \left(2^{\frac{1}{\lambda}} \sigma \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right)^{\frac{2}{\lambda\alpha}}, 0, 0 \right). \quad (6.9)$$

Dem.

$$\begin{aligned} \mathbf{E}(\exp \{i\xi r_t\}) &= \mathbf{E} \left(\exp \left\{ i\xi \sqrt[\lambda]{h_t} \epsilon_t \right\} \right) \\ &= \mathbf{E} \left[\mathbf{E} \left(\exp \left\{ i\xi \sqrt[\lambda]{h_t} \epsilon_t \right\} \middle| \eta_{k-1}, \eta_{k-2}, \dots \right) \right] \\ &= \mathbf{E} \left(\exp \left\{ -\sigma^\lambda |\xi|^\lambda h_t \right\} \right). \end{aligned} \quad (6.10)$$

Sabemos que $h_t \sim S_{\alpha_h}(\sigma_h, 1, 0)$ onde $\alpha_h = \alpha/2$ e $\sigma_h = 2 \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right)^{2/\alpha} \left(\cos \frac{\pi\alpha}{4} \right)^{2/\alpha}$.

¹Mais precisamente, os δ_j 's são não negativos com pelo menos um deles positivo.

²" $a_j < b_j$ eventualmente" significa que existe um j_0 tal que $a_j < b_j$ para todo $j > j_0$.

Pela Propriedade 1.14 e equação (6.10), a Transformada de Laplace de h_t com $\gamma = \sigma^\lambda |\xi|^\lambda$ é

$$\begin{aligned}
 \mathbf{E} \left(e^{-\gamma h_t} \right) &= \mathbf{E} (\exp \{ i \xi r_t \}) \\
 &= \exp \left\{ -\frac{\sigma_h^{\alpha_h}}{\cos \frac{\pi \alpha_h}{2}} \gamma^{\alpha_h} \right\} \\
 &= \exp \left\{ -\frac{\left[2 \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right)^{2/\alpha} \left(\cos \frac{\pi \alpha}{4} \right)^{2/\alpha} \right]^{\alpha/2} (\sigma^\lambda |\xi|^\lambda)^{\alpha/2}}{\cos \frac{\pi \alpha}{4}} \right\} \\
 &= \exp \left\{ -2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda \alpha}{2}} \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right) |\xi|^{\frac{\lambda \alpha}{2}} \right\} \\
 &= \exp \left\{ -\left(2^{\frac{1}{\lambda}} \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right)^{\frac{2}{\lambda \alpha}} \right)^{\frac{\lambda \alpha}{2}} \sigma^{\frac{\lambda \alpha}{2}} |\xi|^{\frac{\lambda \alpha}{2}} \right\}.
 \end{aligned}$$

Desse modo,

$$r_t \sim S_{\frac{\lambda \alpha}{2}} \left(2^{\frac{1}{\lambda}} \sigma \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right)^{\frac{2}{\lambda \alpha}}, 0, 0 \right).$$

■

Segue do teorema acima que, escolhendo o índice de estabilidade apropriado para η_t , a distribuição não condicional de r_t pode ser qualquer distribuição estável, com exceção da normal.

Novamente, poderíamos estar interessados não somente na distribuição de retornos mas também na distribuição de retornos agregados. A distribuição dos retornos agregados segue no teorema abaixo.

Teorema 6.2. *Considere um processo R_s -GARCH($r, p, 0$) com inovações η_t dadas em (6.6). A distribuição não condicional da soma $\sum_{k=0}^{m-1} r_{t-k}$, $m \geq 1$, é α -estável simétrica segundo*

$$\sum_{k=0}^{m-1} r_{t-k} \sim S_{\frac{\lambda \alpha}{2}} \left(2^{\frac{1}{\lambda}} \sigma \left[\sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right)^{\alpha/2} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right)^{\alpha/2} \right]^{\frac{2}{\lambda \alpha}}, 0, 0 \right). \quad (6.11)$$

Dem. Como $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes, a função característica de $\sum_{k=0}^{m-1} r_{t-k}$ pode ser expressa como a transformada de Laplace de $\sum_{k=0}^{m-1} h_{t-k}$. De fato,

$$\begin{aligned}
\mathbf{E} \left(\exp \left\{ i\xi \sum_{k=0}^{m-1} r_{t-k} \right\} \right) &= \mathbf{E} \left(\exp \left\{ i\xi \sum_{k=0}^{m-1} \sqrt[\lambda]{h_{t-k}} \epsilon_{t-k} \right\} \right) \\
&= \mathbf{E} \left(\mathbf{E} \left[\exp \left\{ i\xi \sum_{k=0}^{m-1} \sqrt[\lambda]{h_{t-k}} \epsilon_{t-k} \right\} \middle| \eta_{t-1}, \eta_{t-2}, \dots \right] \right) \\
&= \mathbf{E} \left(\exp \left\{ -\sigma^\lambda |\xi|^\lambda \sum_{k=0}^{m-1} h_{t-k} \right\} \right). \tag{6.12}
\end{aligned}$$

Para encontrarmos a transformada de Laplace de $\sum_{k=0}^{m-1} h_{t-k}$ encontramos sua distribuição de forma análoga ao que foi feito no Teorema 4.2:

$$\sum_{k=0}^{m-1} h_{t-k} \sim S_{\alpha/2} \left(2 \left[\sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right)^{\alpha/2} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right)^{\alpha/2} \right]^{2/\alpha} \left(\cos \frac{\pi\alpha}{4} \right)^{2/\alpha}, 1, 0 \right).$$

De posse da distribuição de $\sum_{k=0}^{m-1} h_{t-k}$, aplicamos a Propriedade 1.14 e a equação (6.12) de modo que a sua transformada de Laplace com $\gamma = \sigma^\lambda |\xi|^\lambda$ é

$$\begin{aligned}
&\mathbf{E} \left(\exp \left\{ -\gamma \sum_{k=0}^{m-1} h_{t-k} \right\} \right) \\
&= \mathbf{E} \left(\exp \left\{ i\xi \sum_{k=0}^{m-1} r_{t-k} \right\} \right) \\
&= \exp \left\{ -2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \left[\sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right)^{\alpha/2} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right)^{\alpha/2} \right] |\xi|^{\frac{\lambda\alpha}{2}} \right\} \\
&= \exp \left\{ - \left(2^{\frac{1}{\lambda}} \left[\sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right)^{\alpha/2} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right)^{\alpha/2} \right]^{\frac{2}{\lambda\alpha}} \right)^{\frac{\lambda\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} |\xi|^{\frac{\lambda\alpha}{2}} \right\}.
\end{aligned}$$

Desse modo,

$$\sum_{k=0}^{m-1} r_{t-k} \sim S_{\frac{\lambda\alpha}{2}} \left(2^{\frac{1}{\lambda}} \sigma \left[\sum_{j=1}^{m-1} \left(\sum_{i=1}^j \delta_i \right)^{\alpha/2} + \sum_{j=m}^{\infty} \left(\sum_{i=j-m+1}^j \delta_i \right)^{\alpha/2} \right]^{\frac{2}{\lambda\alpha}}, 0, 0 \right).$$

■

Após tantas propriedades importantes seria desejável que o processo R_s -GARCH($r, p, 0$) com inovações dada em (6.6) fosse estacionário. De fato, isso se confirma no corolário abaixo.

Corolário 6.1. *O processo R_s -GARCH($p, q, 0$) com inovações η_t dada por (6.6) é simétrico α -estável e estacionário, $0 < \alpha < 2$ e $1 < \lambda \leq 2$.*

Dem. Utilizando argumento similar ao da demonstração do Teorema 6.2 mostra-se que todas as combinações lineares finitas de r_t 's são simétricas α -estáveis com mesmo índice de estabilidade $\lambda\alpha/2$. Um resultado conhecido de processos estáveis é que um processo é simétrico estável se, e somente se todas as suas combinações lineares finitas são simétricas estáveis. Segue, dessa maneira, que r_t é um processo estocástico simétrico α -estável e todas as suas distribuições finito dimensionais são simétricas α -estáveis. Mais que isso, o processo r_t é estacionário. De fato, como $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes, a função característica do vetor $(r_{t_1+k}, \dots, r_{t_d+k})$ pode ser calculada como segue abaixo

$$\begin{aligned} \mathbf{E} \left(\exp \left\{ i \sum_{j=1}^d \xi_j r_{t_j+k} \right\} \right) &= \mathbf{E} \left(\exp \left\{ i \sum_{j=1}^d \xi_j \sqrt{\lambda h_{t_j+k}} \epsilon_{t_j+k} \right\} \right) \\ &= \mathbf{E} \left(\mathbf{E} \left[\exp \left\{ i \sum_{j=1}^d \xi_j \sqrt{\lambda h_{t_j+k}} \epsilon_{t_j+k} \right\} \middle| \eta_{m-1}, \eta_{m-2}, \dots \right] \right) \\ &= \mathbf{E} \left(\exp \left\{ -\sigma^\lambda \sum_{j=1}^d |\xi_j|^\lambda h_{t_j+k} \right\} \right), \end{aligned}$$

onde $m = \max(t_1 + k, \dots, t_d + k)$.

Pela Proposição 6.1 sabemos que $\{h_t\}$ é estacionário, logo

$$\mathbf{E} \left(\exp \left\{ -\sigma^\lambda \sum_{j=1}^d |\xi_j|^\lambda h_{t_j+k} \right\} \right) = \mathbf{E} \left(\exp \left\{ -\sigma^\lambda \sum_{j=1}^d |\xi_j|^\lambda h_{t_j} \right\} \right),$$

e então

$$\mathbf{E} \left(\exp \left\{ i \sum_{j=1}^d \xi_j r_{t_j+k} \right\} \right) = \mathbf{E} \left(\exp \left\{ i \sum_{j=1}^d \xi_j r_{t_j} \right\} \right).$$

Concluimos que para todo $d \geq 1$ e $t_1, \dots, t_d, k \in \mathbb{Z}$

$$(r_{t_1+k}, \dots, r_{t_d+k}) \stackrel{d}{=} (r_{t_1}, \dots, r_{t_d}),$$

e portanto o processo $\{r_t\}$ é estacionário. ■

6.2 Estrutura de dependência

Iremos considerar novamente a função de autocovariância (4.16) e demonstrar teoremas similares aos teoremas apresentados para os processos R-GARCH.

Teorema 6.3. *Considere dois processos R -GAR $CH(r, p, q)$ independentes, $\{X_t\}$ e $\{Y_t\}$, com os mesmos parâmetros de distribuição. Se $\{X_t\}$ tem mais autocovariância que $\{Y_t\}$, digamos, $I_k^{(x)}(s, -s) \geq I_k^{(y)}(s, -s)$ para todo k , então o processo $\{X_t\}$ é menos autodependente que $\{Y_t\}$.*

A demonstração do Teorema 6.3 é análoga à demonstração do Teorema 4.3.

Teorema 6.4. *Seja Q como na Proposição 6.1. Então para um modelo R_s -GAR $CH(r, p, 0)$ com inovações η_t dadas por (6.6) e qualquer $0 < \alpha < 2$, existe uma constante M dependendo de α , Q , ξ_1 , ξ_2 , λ e σ tal que*

$$0 \leq \limsup_{k \rightarrow \infty} Q^{k\frac{\alpha}{2}} |I_k(\xi_1, \xi_2)| \leq M$$

ou

$$-M \leq \limsup_{k \rightarrow \infty} Q^{k\frac{\alpha}{2}} I_k(\xi_1, \xi_2) \leq 0.$$

Dem. Como $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes e o processo R_s -GAR $CH(r, p, 0)$ é estacionário,

$$\begin{aligned} \mathbf{E}[\exp\{i(\xi_1 r_t + \xi_2 r_{t-k})\}] &= \mathbf{E}[\exp\{i(\xi_1 r_k + \xi_2 r_0)\}] \\ &= \mathbf{E}\left(\mathbf{E}\left[\exp\left\{i(\xi_1 \sqrt[k]{h_k} \epsilon_k + \xi_2 \sqrt[k]{h_0} \epsilon_0)\right\}\right] \middle| \eta_{k-1}, \eta_{k-2}, \dots\right) \\ &= \mathbf{E}\left[\exp\left\{-\sigma^\lambda \left(|\xi_1|^\lambda h_k + |\xi_2|^\lambda h_0\right)\right\}\right]. \end{aligned} \quad (6.13)$$

Como h é dado por (4.10), para todo k temos

$$\begin{aligned} &|\xi_1|^\lambda h_k + |\xi_2|^\lambda h_0 \\ &= |\xi_1|^\lambda \sum_{j=1}^{\infty} \delta_j \eta_{k-j} + |\xi_2|^\lambda \sum_{j=1}^{\infty} \delta_j \eta_{-j} \\ &= |\xi_1|^\lambda [\delta_1 \eta_{k-1} + \dots + \delta_k \eta_0] + |\xi_1|^\lambda [\delta_{k+1} \eta_{-1} + \delta_{k+2} \eta_{-2} + \dots] + |\xi_2|^\lambda [\delta_1 \eta_{-1} + \delta_2 \eta_{-2} + \dots] \\ &= |\xi_1|^\lambda \sum_{j=1}^k \delta_j \eta_{k-j} + \sum_{j=1}^{\infty} \left(|\xi_1|^\lambda \delta_{k+j} + |\xi_2|^\lambda \delta_j\right) \eta_{-j} \end{aligned}$$

e então, de (4.11), a distribuição de $|\xi_1|^\lambda h_k + |\xi_2|^\lambda h_0$ é positiva, estritamente estável com índice de estabilidade $\alpha/2$ dada abaixo:

$$|\xi_1|^\lambda h_k + |\xi_2|^\lambda h_0 \sim S_{\alpha/2} \left(2 \left[\sum_{j=1}^k |\xi_1|^{\frac{\lambda\alpha}{2}} \delta_j^{\alpha/2} + \sum_{j=1}^{\infty} (|\xi_1|^\lambda \delta_{k+j} + |\xi_2|^\lambda \delta_j)^{\alpha/2} \right]^{2/\alpha} \left(\cos \frac{\pi\alpha}{4} \right)^{2/\alpha}, 1, 0 \right).$$

Pela Propriedade 1.14 e equação (6.13), a Transformada de Laplace de $|\xi_1|^\lambda h_k + |\xi_2|^\lambda h_0$ é

$$\mathbf{E}[\exp i(\xi_1 r_k + \xi_2 r_0)] = \exp \left\{ -2^{\frac{\alpha}{2}} \left[\sum_{j=1}^k |\xi_1|^{\frac{\lambda\alpha}{2}} \delta_j^{\alpha/2} + \sum_{j=1}^{\infty} (|\xi_1|^\lambda \delta_{k+j} + |\xi_2|^\lambda \delta_j)^{\alpha/2} \right] \sigma^{\frac{\lambda\alpha}{2}} \right\}. \quad (6.14)$$

Com cálculos similares aos cálculos acima chegamos a

$$\mathbf{E}[\exp(i\xi_1 r_k)] = \exp \left\{ -2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right) |\xi_1|^{\frac{\lambda\alpha}{2}} \right\}, \quad (6.15)$$

$$\mathbf{E}[\exp(i\xi_2 r_0)] = \exp \left\{ -2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right) |\xi_2|^{\frac{\lambda\alpha}{2}} \right\}. \quad (6.16)$$

De (6.14), (6.15) e (6.16) segue que

$$I_k(\xi_1, \xi_2) = -2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right) |\xi_1|^{\frac{\lambda\alpha}{2}} - 2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \left(\sum_{j=1}^{\infty} \delta_j^{\alpha/2} \right) |\xi_2|^{\frac{\lambda\alpha}{2}} + 2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \sum_{j=1}^{\infty} (|\xi_1|^{\lambda} \delta_{k+j} + |\xi_2|^{\lambda} \delta_j)^{\alpha/2}. \quad (6.17)$$

Para continuidade da demonstração relembremos um resultado importante.

Lembrete. Para $0 \leq \gamma \leq 1$ e $r, s \in \mathbb{R}$ temos

$$|r + s|^{\gamma} \leq |r|^{\gamma} + |s|^{\gamma} \quad (6.18)$$

e

$$||r + s|^{\gamma} - |r|^{\gamma} - |s|^{\gamma}| \leq 2|r|^{\gamma}. \quad (6.19)$$

Suponha que $k \geq 1$ é tal que $0 < \delta_j < Q^j$ para cada $j > k$. Então, após aplicarmos (6.19) em cada termo de (6.17) (com $r = |\xi_1|^{\lambda} \delta_{k+j}$, $s = |\xi_2|^{\lambda} \delta_j$ e $\gamma = \alpha/2$) temos

$$\begin{aligned} Q^{k\frac{\alpha}{2}} |I_k(\xi_1, \xi_2)| &\leq 2 \cdot 2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \sum_{j=1}^{\infty} |\xi_1|^{\frac{\lambda\alpha}{2}} \delta_{k+j}^{\alpha/2} \\ &< 2^{\frac{\alpha}{2}+1} \sigma^{\frac{\lambda\alpha}{2}} Q^{k\frac{\alpha}{2}} \sum_{j=1}^{\infty} |\xi_1|^{\frac{\lambda\alpha}{2}} Q^{-(k+j)\frac{\alpha}{2}} \\ &= 2^{\frac{\alpha}{2}+1} \sigma^{\frac{\lambda\alpha}{2}} |\xi_1|^{\frac{\lambda\alpha}{2}} \sum_{j=1}^{\infty} Q^{-j\frac{\alpha}{2}} \\ &= 2^{\frac{\alpha}{2}+1} \sigma^{\frac{\lambda\alpha}{2}} |\xi_1|^{\frac{\lambda\alpha}{2}} \frac{Q^{-\alpha/2}}{1 - Q^{-\alpha/2}} \\ &= |\xi_1|^{\frac{\lambda\alpha}{2}} K. \end{aligned} \quad (6.20)$$

É fácil perceber que se trocarmos r e s de lugar nas equações (6.18) e (6.19) a última igualdade em (6.20) terá ξ_2 no lugar de ξ_1 . Também é fácil observar que se $|\xi_1| \leq 1$ ou $|\xi_2| \leq 1$,

$$Q^{k\frac{\alpha}{2}} |I_k(\xi_1, \xi_2)| \leq K.$$

Considere $M = |\xi_1|^{\frac{\lambda\alpha}{2}} K$ se $|\xi_1| > 1$ e $|\xi_2| > 1$ e $M = K$ caso contrário. De (6.18) sabemos que $I_k(\xi_1, \xi_2) \leq 0$, logo

$$0 \leq \limsup_{k \rightarrow \infty} Q^{k\frac{\alpha}{2}} |I_k(\xi_1, \xi_2)| \leq M$$

ou

$$-M \leq \limsup_{k \rightarrow \infty} Q^{k\frac{\alpha}{2}} I_k(\xi_1, \xi_2) \leq 0.$$

■

O Teorema 6.4 mostra que $I_k(\xi_1, \xi_2)$ tende a zero exponencialmente ou mais rápido para processos $R_s\text{-GARCH}(r, p, 0)$ com inovações estritamente estáveis como em (6.6).

No caso do processo $R_s\text{-GARCH}(1, 1, 0)$ com inovações dadas por (6.6) é possível encontrar uma fórmula explícita para a função de autocovariância e determinar seu comportamento assintótico exato.

Teorema 6.5. *No caso do processo $R_s\text{-GARCH}(1, 1, 0)$ com inovações η_t dadas por (6.6) e com $\theta_1 > 0$ e $0 < \phi_1 < 1$,*

$$\lim_{k \rightarrow \infty} \phi_1^{-k\frac{\alpha}{2}} I_k(\xi_1, \xi_2) = -\sigma^{\frac{\lambda\alpha}{2}} |\xi_1|^{\frac{\lambda\alpha}{2}} \frac{(2\theta_1)^{\alpha/2}}{1 - \phi_1^{\alpha/2}},$$

i.e. a função de autocovariância é assintoticamente proporcional a $\phi_1^{k\frac{\alpha}{2}}$ e portanto tende a zero exponencialmente.

Dem. Nesse caso, a variancia condicional h_k depende de h_{k-1} e η_{k-1} , isto é

$$h_k = \theta_1 \eta_{k-1} + \phi_1 h_{k-1} = \Theta(B) \eta_k + \Phi(B) h_k,$$

onde $\Theta(z) = \theta_1 z$ e $\Phi(z) = \phi_1 z$. Como $0 < \phi_1 < 1$, o polinômio $1 - \Phi(z)$ não possui raízes no disco unitário fechado $\{z : |z| \leq 1\}$. Pela Proposição 6.1, os coeficientes δ_j 's são os coeficientes na expansão da série $\frac{\theta_1 z}{1 - \phi_1 z}$, $|z| < 1$. Como

$$\frac{\theta_1 z}{1 - \phi_1 z} = \theta_1 z \sum_{j=0}^{\infty} \phi_1^j z^j = \sum_{j=1}^{\infty} \theta_1 \phi_1^{j-1} z^j,$$

os δ_j 's são dados por

$$\delta_j = \theta_1 \phi_1^{j-1}, \quad j = 1, 2, \dots$$

Pelo Corolário 6.1, $\{r_t\}$ é um processo estacionário SaS. De (6.17) temos que para $k \geq 1$

$$\begin{aligned}
& I_k(\xi_1, \xi_2) \\
&= 2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \sum_{j=1}^{\infty} \left[(|\xi_1|^\lambda \delta_{k+j} + |\xi_2|^\lambda \delta_j)^{\alpha/2} - |\xi_1|^{\frac{\lambda\alpha}{2}} \delta_{k+j}^{\alpha/2} - |\xi_2|^{\frac{\lambda\alpha}{2}} \delta_j^{\alpha/2} \right] \\
&= 2^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \sum_{j=1}^{\infty} \left[\left(|\xi_1|^\lambda \theta_1 \phi_1^{k+j-1} + |\xi_2|^\lambda \theta_1 \phi_1^{j-1} \right)^{\alpha/2} - |\xi_1|^{\frac{\lambda\alpha}{2}} \left(\theta_1 \phi_1^{k+j-1} \right)^{\alpha/2} - |\xi_2|^{\frac{\lambda\alpha}{2}} \left(\theta_1 \phi_1^{j-1} \right)^{\alpha/2} \right] \\
&= -(2\theta_1)^{\frac{\alpha}{2}} \sigma^{\frac{\lambda\alpha}{2}} \left[|\xi_1|^{\frac{\lambda\alpha}{2}} \phi_1^{k\frac{\alpha}{2}} + |\xi_2|^{\frac{\lambda\alpha}{2}} - \left(|\xi_1|^\lambda \phi_1^k + |\xi_2|^\lambda \right)^{\alpha/2} \right] \sum_{j=1}^{\infty} \phi_1^{(j-1)\frac{\alpha}{2}} \\
&= -\sigma^{\frac{\lambda\alpha}{2}} \frac{(2\theta_1)^{\frac{\alpha}{2}}}{1 - \phi_1^{\alpha/2}} \left[|\xi_1|^{\frac{\lambda\alpha}{2}} \phi_1^{k\frac{\alpha}{2}} + |\xi_2|^{\frac{\lambda\alpha}{2}} - \left(|\xi_1|^\lambda \phi_1^k + |\xi_2|^\lambda \right)^{\alpha/2} \right].
\end{aligned}$$

Portanto,

$$\begin{aligned}
\lim_{k \rightarrow \infty} \phi_1^{-k\frac{\alpha}{2}} I_k(\xi_1, \xi_2) &= -\sigma^{\frac{\lambda\alpha}{2}} \frac{(2\theta_1)^{\frac{\alpha}{2}}}{1 - \phi_1^{\alpha/2}} \lim_{k \rightarrow \infty} \left[|\xi_1|^{\frac{\lambda\alpha}{2}} + |\xi_2|^{\frac{\lambda\alpha}{2}} \phi_1^{-k\frac{\alpha}{2}} - \left(|\xi_1|^\lambda + |\xi_2|^\lambda \phi_1^{-k} \right)^{\alpha/2} \right] \\
&= -\sigma^{\frac{\lambda\alpha}{2}} \frac{(2\theta_1)^{\frac{\alpha}{2}}}{1 - \phi_1^{\alpha/2}} \lim_{k \rightarrow \infty} \left[|\xi_1|^{\frac{\lambda\alpha}{2}} + \frac{|\xi_2|^{\frac{\lambda\alpha}{2}} - (\phi_1^k |\xi_1|^\lambda + |\xi_2|^\lambda)^{\alpha/2}}{\phi_1^{k\frac{\alpha}{2}}} \right].
\end{aligned}$$

e para completar a demonstração basta notar que

$$\lim_{k \rightarrow \infty} \left[|\xi_1|^{\frac{\lambda\alpha}{2}} + \frac{|\xi_2|^{\frac{\lambda\alpha}{2}} - (\phi_1^k |\xi_1|^\lambda + |\xi_2|^\lambda)^{\alpha/2}}{\phi_1^{k\frac{\alpha}{2}}} \right] = 0.$$

■

Para ilustrarmos um processo R_s -GARCH(1,1,0) simulamos o mesmo com diferentes valores para os parâmetros α , θ_1 e ϕ_1 . Nas figuras seguintes seguem o gráfico dos retornos, da volatilidade e a função de autocovariação teórica dos diversos processos simulados. Lembrando que a forma fechada para a função de autocovariação do processo R_s -GARCH(1,1,0) foi encontrada na demonstração do teorema anterior e é dada por

$$I_k(\xi_1, \xi_2) = -\sigma^{\frac{\lambda\alpha}{2}} \frac{(2\theta_1)^{\frac{\alpha}{2}}}{1 - \phi_1^{\alpha/2}} \left[|\xi_1|^{\frac{\lambda\alpha}{2}} \phi_1^{k\frac{\alpha}{2}} + |\xi_2|^{\frac{\lambda\alpha}{2}} - \left(|\xi_1|^\lambda \phi_1^k + |\xi_2|^\lambda \right)^{\alpha/2} \right]. \quad (6.21)$$

Figura 6.1: Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.9$, $\theta_1 = 0.0000001$, $\phi_1 = 0.8$, $\lambda = 1.8$ e $\sigma = 0.5$.

Figura 6.2: Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.0000001$, $\phi_1 = 0.8$, $\lambda = 1.8$ e $\sigma = 0.5$.

Figura 6.3: Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.0000001$, $\phi_1 = 0.5$, $\lambda = 1.8$ e $\sigma = 0.5$.

Figura 6.4: Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.0001$, $\phi_1 = 0.8$, $\lambda = 1.8$ e $\sigma = 0.5$.

Figura 6.5: Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.0000001$, $\phi_1 = 0.8$, $\lambda = 1.3$ e $\sigma = 0.5$.

Figura 6.6: Retornos e volatilidade de um processo R_s -GARCH(1,1,0) e função de autocovariância teórica para $\alpha = 1.7$, $\theta_1 = 0.0000001$, $\phi_1 = 0.8$, $\lambda = 1.8$ e $\sigma = 2.0$.

6.3 Inferência indireta para o modelo R_s -GARCH(1,1,0)

Novamente iremos empregar o método de estimação indireta de Gouriéroux et. al. (1993) para encontrar estimativas dos parâmetros do modelo original R_s -GARCH(1,1,0). O modelo auxiliar utilizado será o mesmo modelo que empregamos na estimação dos parâmetros dos modelos R -GARCH(1,1,0) e R_t -GARCH(1,1,0) e, para facilitar o prosseguimento, relembramos abaixo os modelos original e auxiliar em questão.

- (Modelo auxiliar) Processo GARCH(1,1) com inovações ϵ_t distribuídas segundo t de Student:

$$\begin{aligned} r_t &= \sqrt{h_t} \epsilon_t, \quad t \in \mathbb{Z}, \\ h_t &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1}, \end{aligned}$$

onde $\alpha_0 > 0$ e $(\alpha_1 + \beta_1) < 1$.

- (Modelo original) Processo R_s -GARCH(1,1,0):

$$\begin{aligned} r_t &= \sqrt[{\lambda}]{h_t} \epsilon_t, \quad t = 0, \pm 1, \pm 2, \dots, \\ h_t &= \theta_1 \eta_{t-1} + \phi_1 h_{t-1}, \end{aligned}$$

onde as inovações ϵ_t são v.a's i.i.d. com distribuição $S_\lambda(\sigma, 0, 0)$, as inovações η_t são v.a's i.i.d. distribuídas segundo (6.6) e $\{\epsilon_t\}$ e $\{\eta_t\}$ são independentes.

Para os modelos R_s -GARCH temos dois parâmetros adicionais, os parâmetros λ e σ dos ϵ_t 's com distribuição estável. Estes não serão estimados e sim escolhidos convenientemente ao compararmos as distribuições utilizadas nos modelos R -GARCH e R_t -GARCH. Dessa maneira, os parâmetros do modelo auxiliar serão também simbolizados por $\zeta = (\alpha_0, \alpha_1, \beta_1)$ assim como os parâmetros do modelo original simbolizados por $\theta = (\alpha, \theta_1, \phi_1)$.

Como $\dim \zeta = \dim \theta$, motivados pela Propriedade 3.1, escolhemos a matriz simétrica não negativa definida Ω como sendo a matriz identidade. O parâmetro α_0 do modelo auxiliar está diretamente ligado aos parâmetros α e θ_1 de modo que a função de ligação é claramente não bijetora. Apesar disso, como veremos nas simulações, os estimadores de inferência indireta são muito próximos do seu valor real.

6.3.1 Simulações

Nesta seção, conduzimos novamente simulações no intuito de averiguar a estimação dos parâmetros pelo método de inferência indireta. Como já adiantado, os parâmetros λ e σ serão escolhidos. A idéia é escolher valores de modo que a distribuição dos retornos seja leptocúrtica. Esses dois parâmetros possuem uma gama de possibilidades infinita então, para assegurar essa característica da distribuição no que diz respeito à sua forma e ao mesmo tempo permitir a estimação dos dados via estimação indireta, escolhemos como primeiro caso $\lambda = 1.8$ e $\sigma = 0.5$. Escolhemos esses valores graficamente ao comparar as distribuições estáveis com as distribuições normal padrão e t de Student utilizadas nos modelos R -GARCH(1,1,0) e R_t -GARCH(1,1,0) respectivamente. Segue na figura abaixo as densidades das respectivas distribuições.

Figura 6.7: Densidades das distribuições normal padrão, t de Student e estável.

Como podemos observar, as distribuições estáveis permitem caudas pesadas e leptocúrticas. Com a experiência que temos do caso R-GARCH, decidimos utilizar 100 réplicas com 100000 observações para obter as estimativas visto que o comportamento dos estimadores é bem similar. Escolhemos alguns valores para os parâmetros, α , θ_1 e ϕ_1 e apresentamos na Tabela 6.1 a média e erro padrão Monte Carlo estimados, onde $\nu = 4$ no modelo auxiliar.

	$\alpha = 1.5$	$\theta_1 = 0.0000001$	$\phi_1 = 0.7$
Média	1.5003	1.3718×10^{-7}	0.7586
Erro Padrão	0.0214	3.7875×10^{-8}	0.0233
	$\alpha = 1.6$	$\theta_1 = 0.0000001$	$\phi_1 = 0.6$
Média	1.5985	1.6802×10^{-7}	0.6245
Erro Padrão	0.0291	2.6712×10^{-8}	0.0134
	$\alpha = 1.7$	$\theta_1 = 0.0000001$	$\phi_1 = 0.8$
Média	1.7091	1.4225×10^{-7}	0.8472
Erro Padrão	0.0116	4.86×10^{-8}	0.0251

Tabela 6.1: Média e erro padrão MC para a estimação de um R_s -GARCH(1,1,0) com 100 réplicas e 100000 observações.

6.3.2 Aplicação empírica (TLM)

Nesta seção repetimos a aplicação empírica realizada nos Capítulos 4 e 5 com o intuito de avaliar a performance, realizar uma análise residual da classe de modelos R_s -GARCH(1,1,0) com dados reais e comparar a performance dessa classe com a dos modelos R-GARCH(1,1,0) e R_t -GARCH(1,1,0). Ajustamos então esse modelo aos dados intradiários de 15 minutos da Telemar (TLM15m) com distribuição estável $S_{1.8}(0.5,0,0)$ tomando como base 100 réplicas independentes e $S = 5$. Apresentamos na Tabela 6.2 abaixo os valores da média e erro padrão Monte Carlo dos parâmetros α , θ_1 e ϕ_1 ajustados, onde $\nu = 4$ no modelo auxiliar.

	α	θ_1	ϕ_1
Média	1.5849	8.6121×10^{-7}	0.8243
Erro Padrão	0.0033	4.8663×10^{-8}	0.0082

Tabela 6.2: Média e erro padrão MC ao ajustarmos um R_s -GARCH(1,1,0) aos logretornos de TLM15m.

Para efeitos de comparação seguem na Tabela 6.3 abaixo o erro quadrático médio para os modelos GARCH(0,1), R-GARCH(1,1,0), R_t -GARCH(1,1,0) e R_s -GARCH(1,1,0):

EQM para GARCH(0,1)	1.9150×10^{-5}
EQM para R-GARCH(1,1,0)	2.2942×10^{-7}
EQM para R_t-GARCH(1,1,0)	7.4659×10^{-8}
EQM para R_s-GARCH(1,1,0)	7.1438×10^{-8}

Tabela 6.3: Erro quadrático médio ao ajustarmos os modelos GARCH(0,1), R-GARCH(1,1,0), R_t -GARCH(1,1,0) e R_s -GARCH(1,1,0) aos logretornos de TLM15m

Como podemos observar, o erro quadrático médio do modelo R_s -GARCH(1,1,0) ajustado é menor que o erro quadrático médio dos modelos GARCH(0,1) e R-GARCH(1,1,0) e muito próximo ao do modelo R_t -GARCH(1,1,0) ajustados. Com o objetivo de avaliar a melhora do modelo em comparação com a hipótese de normalidade assumida sobre os resíduos para os modelos R-GARCH(1,1,0), ilustramos abaixo o QxQ plot dos resíduos do modelo R_s -GARCH(1,1,0) ajustado aos logretornos de TLM15m contra os quantis de uma distribuição $S_{1.8}(0.5, 1, 0)$. Podemos ver que a hipótese parece mais adequada do que a hipótese de normalidade dos modelos R-GARCH(1,1,0).

Figura 6.8: QxQ plot dos resíduos do modelo R_s -GARCH(1,1,0) ajustado a TLM15m versus quantis de uma $S_{1.8}(0.5, 1, 0)$.

6.3.3 Aplicação empírica (IBV)

Mais uma vez, estamos interessados em avaliar a performance, realizar uma análise residual da classe de modelos R_s -GARCH(1,1,0) com dados reais e comparar esta performance com a dos processos

R -GARCH(1,1,0) e R_t -GARCH(1,1,0). Nessa direção, seguindo a mesma idéia aplicada no exemplo anterior, ajustamos o modelo R_s -GARCH(1,1,0) aos logretornos intradiários de 15 minutos do índice de ações brasileiro Ibovespa (IBV15m) com base em 100 réplicas independentes e $S = 4$ vetores. Neste caso, a hipótese é de que os resíduos possuem distribuição $S_{1.8}(0.5, 0, 0)$. Apresentamos na Tabela 6.4 abaixo os valores da média e erro padrão Monte Carlo dos parâmetros α , θ_1 e ϕ_1 estimados, onde $\nu = 3$ no modelo auxiliar, e a comparação entre os erros quadrático médio para os respectivos modelos GARCH(0,1), R -GARCH(1,1,0), R_t -GARCH(1,1,0) e R_s -GARCH(1,1,0) na Tabela 6.5.

	α	θ_1	ϕ_1
Média	1.5031	1.5140×10^{-6}	0.7578
Erro Padrão	0.0040	2.3607×10^{-7}	0.0036

Tabela 6.4: Média e erro padrão MC ao ajustarmos um R_s -GARCH(1,1,0) aos logretornos de IBV15m.

EQM para GARCH(0,1)	2.4778×10^{-5}
EQM para R-GARCH(1,1,0)	3.5512×10^{-6}
EQM para R_t-GARCH(1,1,0)	6.7555×10^{-7}
EQM para R_s-GARCH(1,1,0)	6.2983×10^{-7}

Tabela 6.5: Erro quadrático médio ao ajustarmos os modelo GARCH(0,1), R -GARCH(1,1,0), R_t -GARCH(1,1,0) e R_s -GARCH(1,1,0) aos logretornos de IBV15m

Como no caso anterior, observarmos que o erro quadrático médio do modelo R_s -GARCH(1,1,0) ajustado é menor que o erro quadrático médio dos modelos GARCH(0,1) e R -GARCH(1,1,0) e muito próximo ao do modelo R_t -GARCH(1,1,0) ajustados. Novamente com o objetivo de avaliar a melhora do modelo no que diz respeito à hipótese sobre os resíduos, abaixo ilustramos o QxQ plot dos resíduos do modelo R_s -GARCH(1,1,0) ajustado a TLM15m contra os quantis de uma distribuição estável $S_{1.8}(0.5, 0, 0)$. Podemos ver que a hipótese parece mais adequada do que a hipótese de normalidade dos modelos R -GARCH(1,1,0).

Figura 6.9: QxQ plot dos resíduos do modelo R_s -GARCH(1,1,0) ajustado a IBV15m versus quantis de uma estável $S_{1.8}(0.5, 0, 0)$.

Capítulo 7

Conclusões e trabalhos futuros

Sugerimos o método de inferência indireta para a estimação dos parâmetros dos processos R-GARCH, R_t -GARCH e R_s -GARCH. Em particular, aplicamos o método aos processos R-GARCH(1,1,0), R_t -GARCH(1,1,0) e R_s -GARCH(1,1,0) utilizando o modelo GARCH(1,1) com inovações t de Student como modelo auxiliar. Como podemos constatar, no estudo de simulações, a estimação tem desempenho muito bom para os três casos. Duas aplicações empíricas também foram realizadas para os três casos de onde pudemos concluir que o desempenho dessas classes de processos é superior à classe de processos GARCH com respeito ao erro quadrático médio residual. Ao compararmos uma a uma, observamos um desempenho melhor das classes R_t -GARCH(1,1,0) e R_s -GARCH(1,1,0) em relação à classe R-GARCH(1,1,0) no que diz respeito ao erro quadrático médio residual e também sob a hipótese dos resíduos.

Em termos de adequação dos resíduos, observamos que a hipótese de distribuição t de Student nos processos R_t -GARCH(1,1,0) apresentam gráficos QxQ plot mais próximo do esperado, entretanto os processos R_s -GARCH(1,1,0) também apresentaram um bom comportamento além de uma teoria nova e bem fundamentada. É importante lembrar também que existe uma gama infinita de possibilidades a serem exploradas para os resíduos nos processos R_s -GARCH(1,1,0) visto que os parâmetros λ e σ podem ser escolhidos arbitrariamente dentro de seus intervalos.

Este método é bastante promissor, especialmente pela sua facilidade de extensão para vários outros modelos. Em trabalhos futuros iremos explorar modelos de ordem maior, como por exemplo R-GARCH(1,1,1), R_t -GARCH(1,1,1) e R_s -GARCH(1,1,1). Nesses casos acrescentamos o termo r_t^2 que pode ter papel importante na forma como descrevemos a volatilidade.

Outro tópico importante a ser explorado é a estimação da função de autocovariância que fica mais próxima de ser alcançada visto que obtivemos formas fechadas para a sua versão teórica nos casos R-GARCH(1,1,0) e R_s -GARCH(1,1,0).

Referências Bibliográficas

- Bahnsali, R. J. (1993), *Estimation of the Impulse Response Coefficients of a Linear Processes with infinite variance*. Journal of Multivariate Analysis 45(2), 274-290.
- Bergström, H. (1952), *On some expansions of stable distributions*. Arkiv für Mathematik 2, 375-378.
- Bollerslev, T. (1986), *Generalized autoregressive conditional heteroskedasticity*. Journal of Econometrics, 31, 307-327.
- Bollerslev, T. (1987), *A conditionally heteroskedastic time series model for speculative process and rates of return*. Review of Economics and Statistics, 69, 542-547.
- Bollerslev, T. (1988), *On the correlation structure for the generalized autoregressive conditional heteroskedasticity*. Journal of Time Series Analysis, 9, 121-132.
- Breiman, L. (1968), *Probability*. Addison-Wesley, Reading, MA.
- Brockwell, P. J. e Davis, R. A. (1991), *Time Series: Theory and Methods*., Springer-Verlag, New York, second edition.
- Calder, M. e Davis, R. A. (1996), *Inference for Linear Processes with Stable Noise*. Preprint.
- Calzolari, G., Fiorentini, G. & Sentana, E. (2001), *Constrained indirect inference estimation*. LSE Discussion Paper 384, London School of Economics.
- Chipman, H. A., George, E. I. and McCulloch, R. E. (1998b), *Making Sense of a Forest of Trees*. Proceedings of the 30th Symposium on the Interface, S. Weisberg, Ed., Interface Foundation of North America. pp 84-92.
- De Vries, C. (1991), *On the relation between GARCH and stable processes*. Journal of Econometrics 48, 313-324.
- Diebold, F. (1986), *Temporal aggregation of ARCH processes and the distribution of asset returns*. Special Studies Paper 200, Board of Governors of the Federal Reserve System.
- DuMochel, W. (1971), *Stable Distributions in Statistical Inference, PhD Thesis*. Yale University, New Haven.
- DuMochel, W. (1973a), *On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution*. Annals of Statistics 1, 948-957.

- DuMochel, W. (1975), *Stable distributions in statistical inference: 2. Information from stably distributed samples*. Journal of the American Statistical Association 70, 386-393.
- DuMochel, W. (1983), *Estimating the stable index α in order to measure tail thickness: A critique*. Annals of Statistics 11, 1019-1031.
- Engle, R. F. (1982), *Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation*. Econometrica, 50, 987-1008.
- Fama, E. & Roll, R. (1968), *Some properties of symmetric stable distributions*. Journal of the American Statistical Association 63, 817-836.
- Fama, E. & Roll, R. (1971), *Parameter estimates for symmetric stable distributions*. Journal of the American Statistical Association 66, 331-338.
- Feller, W. (1966), *An Introduction to Probability Theory and Its Applications*, Volume II. John Wiley & Sons.
- Fielitz, B. & Smith, E. (1971), *Asymmetric stable distributions of stock price changes*. Journal of American Statistical Association 67, 813-814.
- Gallant, A. & Tauchen, G. (1996), *Which moments to match?*. Econometric Theory 12, 657-681.
- Gouriéroux, C., Monfort, A. & Renault, E. (1993), *Indirect inference*. Journal of Applied Econometrics 8, 85-118.
- Gouriéroux, C. & Monfort, A. (1996), *Simulated-Based Econometric Methods*. Oxford University Press, Oxford.
- Hill, B. (1975), *A simple general approach to inference about the tail of a distribution*. Annals of Statistics 3, 1163-1174.
- Kogon, S. & Williams, D. (1998), *Characteristic function based estimation of stable distribution parameters*, in R. Adler, R. Feldman & M. Taqqu, eds, 'A practical guide to heavy tails', Birkhäuser, Berlin.
- Kokoszka, P. S. & Taqqu, M. S. (1994), *Infinite Variance Stable ARMA Processes*. Journal of Time Series Analysis 15, 203-221.
- Koutrouvelis, I. (1980), *Regression-type estimation of the parameters of stable laws*. Journal of the American Statistical Association 75, 919-928.
- Koutrouvelis, I. (1981), *An iterative procedure for the estimation of the parameters of stable laws*. Communications in Statistics - Simulation and Computation 10, 17-28.
- LeCam, L. (1952), *On Some Asymptotic Properties of Maximum-Likelihood Estimates and Related Bayes Estimates*, PhD thesis. University of California, Berkeley.
- Lévy, P. (1924), *Théorie des erreurs de la loi de Gauss et les lois exceptionnelles*. Bulletin de la Société de France 52, 49-85.

- Lombardi, M. & Calzolari, G. (2008), *Indirect estimation of α -stable distributions and processes*. Econometrics Journal 11, 1-15.
- McCulloch, J. (1986), *Simple consistent estimators of stable distributions parameters*. Communications in Statistics - Simulation and Computation 15, 1109-1136.
- McCulloch, J. (1997), *Measuring tail thickness to estimate the stable index α : A critique* Journal of Business & Economic Statistics 15, 74-81.
- McCulloch, J. (1998b), *Numerical approximation of the symmetric stable distribution and density*, in R. Adler, R. Feldman & M. Taqqu, eds, 'A practical guide to heavy tails', Birkhäuser, Berlin.
- Mikosch, T., Gadrich, T., Kluppelberg, C. e Adler, R. J. (1995), *Parameter estimation of ARMA models with infinite variance innovations*. The Annals of Statistics 23(1), 305-326.
- Mittnik, S., Doganoglu, T. & Chenyao, D. (1999), *Computing the probability density function of stable Paretian distribution*. Mathematical and Computer Modelling 29, pp. 235-240.
- Mittnik, S., Rachev, S., Doganoglu, T. & Chenyao, D. (1999), *Maximum likelihood estimation of stable Paretian models*. Mathematical and Computer Modelling 29, 275-293.
- Morettin, P. A. (2011), *Econometria Financeira 2 Ed.*. Editora Edgard Blucher.
- Nardelli, S., (1997), *Statistical tests for stable distributions*. Master's Thesis, European University Institute, Fiesole.
- Nolan, J. (1997), *Numerical computation of stable densities and distribution functions*. Communications in Statistics - Stochastic Models 13, pp 759-774.
- Nolan, J. (2002), *Maximum likelihood estimation and diagnostics for stable distributions*. American University, Washington.
- Nowicka, J. (1997a), *Asymptotic behavior of the covariation and the codifference for ARMA models with stable innovations*. Stochastic Models 13(4), 673-686.
- Nowicka, J. (1997b), *Parameter estimation of ARCH-type models with normal and stable innovations*. Proceedings of 16th International Conference on Multivariate Statistical Analysis. Łódź, 13-21.
- Nowicka, J. (1998) *Analysis of measures of dependence for time series with α -stable innovations*. PhD Dissertation, Wrocław University of Technology.
- Nowicka, J e Weron, A. (1995), *Numerical Approximation of Dependence Structure for Symmetric Stable AR(2) Processes*. Preprint.
- Panorska, A. K., Mittnik, S. e Rachev, S. T. (1995), *Stable GARCH Models for Financial Time Series*. Appl. Math. Lett. 8(5), 33-37.
- Paulson, A., Holcomb, E. & Leitch, R. (1975), *The estimation of the parameters of the stable laws*. Biometrika 62, 163-170.
- Press, S. (1972), *Estimation of univariate and multivariate stable distributions*. Journal of the American Statistical Association 67, 842-846.

- Rosadi, D. & Deistler, M. (2011), *Estimating the codifference function of linear time series models with infinite variance*. *Metrika* Volume 73, 3, 395-429.
- Samorodnitsky, G. e Taquq, M.S. (1994), *Stable Non Gaussian Random Process: Stochastic models with infinite variance*. Chapman & Hall/CRC.
- Weron, A. and Weron, R. (1995), *Lectures of Lévy α -stable Variables and Processes*. *Lectures Notes in Physics* 457 Springer-Verlag, pp. 379-392.