

Helder Parra Palaro

*Aplicação de acoplamento no cálculo
do valor em risco*

Campinas - SP

2004

APLICAÇÃO DE ACOPLAMENTO NO CÁLCULO DO VALOR EM RISCO

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Helder Parra Palaro e aprovada pela comissão julgadora.

Campinas, 18 de Março de 2004.

.....
Prof. Dr. Luiz Koodi Hotta
Orientador

Banca Examinadora:

1. Prof. Dr. Luiz Koodi Hotta (Orientador) - IMECC/UNICAMP.
2. Prof. Dr. Pedro Luiz Valls Pereira - IBMEC-SP
3. Profa. Dra. Verónica Andrea González-Lopes - IMECC/UNICAMP.

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para a obtenção do Título de Mestre em Estatística.

Aos meus pais, Geraldo e Miriam

Agradecimentos

Agradeço aos meus pais, Geraldo Palaro e Miriam Parra Palaro, que sempre me incentivaram a estudar e me deram todo o suporte para que eu me concentrasse em meus estudos.

Ao meu orientador, Professor Dr. Luiz Koodi Hotta, que sempre incentivou a troca de idéias entre seus alunos, e sempre esteve disponível para solucionar as dúvidas que surgiram neste trabalho. Sua experiência e objetividade que me guiaram nesta dissertação.

À todos os Professores do Departamento de Estatística do IMECC, em especial aos Professores Drs. Mauro Marques, Ronaldo Dias e Filidor Vilca Labra, pelos cursos ministrados.

À todos os alunos do mestrado em Estatística. Primeiramente aos amigos que compartilharam todos os momentos juntos, desde a admissão até a conclusão : Alexandre Rübesam, Ricardo Takeyama, Nélon Lopes e Tatiana Benaglia. Ao Benilton Carvalho, que muito me ajudou em Latex e Matlab. Ao Anderson Motta, que sempre me ajudou em minhas escolhas acadêmicas. Agradeço também a Paula, Fernanda, Izabella, Rossana, Rogério, Rosemeire, Clécio, Roberta, Edmílson, Samara e Daniel.

Ao IMECC, Instituto de Matemática, Estatística e Computação Científica, pelos excelentes recursos computacionais, de impressão e de estudo disponíveis para os alunos de mestrado

À Fapesp, Fundação de Amparo à Pesquisa do Estado de São Paulo, responsável pelo suporte financeiro.

Resumo

O Valor-em-Risco (VaR) desempenha um papel muito importante no gerenciamento de risco. Existem diversos métodos para estimá-lo, como o de simulação histórica, o método analítico ou da variância-covariância e simulação de Monte-Carlo. Enquanto o primeiro método não assume nenhuma distribuição, os dois últimos necessitam do conhecimento da distribuição multivariada, que no método analítico geralmente é a normal. A teoria de acoplamento aparece como uma ferramenta fundamental na formulação da distribuição multivariada. Ela permite definir esta distribuição a partir das distribuições marginais e da dependência entre as variáveis. Recentemente, a teoria de acoplamentos foi estendida para o caso condicional, nos permitindo o uso dos acoplamentos na modelagem de dependência que pode variar no tempo. A variação temporal no primeiro e segundo momentos condicionais é amplamente discutida na literatura. Deste modo permitir a variação temporal na dependência condicional entre séries temporais é natural. O trabalho apresenta alguns conceitos e funções de acoplamento, bem como algumas propriedades e a aplicação no cálculo do VaR para duas carteiras, uma composta pelas taxas de câmbio Libra esterlina/Dólar e Iene/Dólar, e a outra composta pelos índices de ações Nasdaq e S&P500 dos EUA.

Abstract

Value at Risk (VaR) has a central role in the risk management. There are several approaches for the estimation of VaR, such as the variance-covariance (also known as analytical), the historical simulation and the Monte Carlo approaches. Whereas the first approach does not assume any distribution, the last two approaches demand the joint distribution to be known, which in the analytical approach is frequently the normal distribution. The copula theory is a fundamental tool in the modelling of the (joint) multivariate distribution. It allows the definition of the joint distribution through the marginal distributions and the dependence between the variables. Recently the copula theory has been extended to the conditional case, allowing us the use of copulas to model time-varying dependence, which is time varying. Variation in time of the first and second conditional moments is widely discussed in the literature, and it is natural to allow for the time variation of the conditional dependence seems to be natural. This work presents some concepts of copula functions, and some issues involved in the application of the copulas theory in the estimation of VaR of portfolios. The methodology is applied to two portfolios, the first composed by the Sterling Pound/US dollar and the Japanese Yen/US dollar exchange rates, and the second composed by the Nasdaq and S&P500 US stock indexes.

Sumário

Lista de Tabelas	p. x
Lista de Figuras	p. xii
1 Introdução	p. 1
2 Funções de Acoplamento	p. 5
2.1 Introdução	p. 5
2.2 Caso não-condicional	p. 6
2.2.1 Definição	p. 6
2.2.2 Propriedades	p. 7
2.3 Caso condicional	p. 12
2.3.1 Definição	p. 12
2.3.2 Propriedades	p. 13
3 Medidas de Dependência	p. 15
3.1 Introdução	p. 15
3.2 Dependência em séries financeiras	p. 16
3.3 Dependência perfeita	p. 17
3.4 Correlação Linear	p. 19

3.5	Medidas de concordância	p. 22
3.5.1	Concordância	p. 22
3.5.2	Tau de Kendall	p. 25
3.5.3	Rho de Spearman	p. 26
3.5.4	Características adicionais do tau de Kendall e rho de Spearman	p. 27
3.5.5	Vantagens da medida de concordância sobre a correlação linear	p. 28
3.6	Dependência nas caudas	p. 29
4	Classes de funções de acoplamento	p. 33
4.1	Acoplamentos Elípticos	p. 33
4.1.1	Distribuições elípticas	p. 34
4.1.2	Acoplamentos gaussianos	p. 35
4.1.3	Acoplamentos t	p. 36
4.2	Acoplamentos arquimedianos	p. 37
4.2.1	Definição	p. 37
4.2.2	Propriedades	p. 40
4.2.3	Acoplamentos arquimedianos multivariados	p. 42
4.3	Acoplamento de Plackett	p. 43
4.4	Algumas aplicações encontradas na literatura	p. 45
5	Estimação e Identificação de Funções de Acoplamento	p. 49
5.1	Estimação Paramétrica	p. 49
5.1.1	Método de máxima verossimilhança exata	p. 51

5.1.2	Método IFM (Inference Function for Margins)	p. 51
5.1.3	Método de Máxima Verossimilhança Canônica	p. 53
5.1.4	Estimação de parâmetros através de medidas de de- pendência	p. 54
5.1.5	Caso condicional	p. 54
5.2	Estimação Não-Paramétrica	p. 57
5.2.1	Acoplamentos empíricos	p. 57
5.3	Seleção da função de acoplamento	p. 59
5.3.1	Acoplamentos Arquimedianos	p. 59
5.3.2	Acoplamento Empírico	p. 60
5.3.3	Critério de Informação de Akaike (AIC)	p. 61
6	Estimação das Distribuições Marginais	p. 63
6.1	Fatos estilizados de retornos financeiros	p. 63
6.2	Modelos de volatilidade	p. 64
6.2.1	Modelo GARCH(p,q)	p. 65
7	Geração de vetores pseudo aleatórios	p. 69
7.1	Método geral	p. 69
7.2	Método de rejeição	p. 72
7.3	Acoplamentos Gaussianos	p. 73
7.4	Acoplamentos t-Student	p. 74
7.5	Acoplamentos arquimedianos	p. 74
7.6	Acoplamento de Plackett	p. 76

8	Aplicação	p. 77
8.1	Descrição da amostra	p. 77
8.2	Análise exploratória de dados	p. 78
8.3	Modelagem das distribuições marginais	p. 81
8.4	Modelagem de dependência	p. 91
8.4.1	Análise Exploratória	p. 91
8.4.2	Estimação da função de acoplamento	p. 92
8.4.3	Comparação e Diagnóstico	p. 94
8.5	Estimação do Valor em Risco	p. 99
8.6	Comparação com outros modelos	p. 104
9	Conclusão	p. 113
	Apêndice A – Aplicação 2	p. 115
A.1	Análise exploratória de dados	p. 115
A.2	Modelagem das distribuições marginais	p. 116
A.3	Modelagem da dependência entre as séries	p. 120
A.4	Simulação do Valor em Risco	p. 122
A.5	Comparação com outros modelos	p. 123
	Referências Bibliográficas	p. 125

Lista de Tabelas

4.1	Uma tabela de contingência 2×2	p. 44
8.1	Estatísticas descritivas dos retornos de Libra e Iene.	p. 79
8.2	Parâmetros estimados dos modelos GARCH em (8.1) e (8.2) e erro padrão das estimativas.	p. 84
8.3	Parâmetros estimados das funções de acoplamento (Desvios padrão em parênteses).	p. 93
8.4	Parâmetros estimados das funções de acoplamento por medi- das de dependência.	p. 94
8.5	Distância quadrática entre o acoplamento estimado e o aco- plamento empírico na cauda inferior (quadrado $[0, \delta]^2$). . . .	p. 95
8.6	Proporção de observações (e número de observações), para t $= 750$ a 3502 , em que a perda do portfolio excede o VaR para $\alpha = 0,05$, $\alpha = 0,01$ e $\alpha = 0,005$. * = Número esperado de observações.	p. 101
8.7	Proporção de dados em que a perda do portfolio excede o VaR $100(1 - \alpha)\%$ e P-valor do teste de Christoffersen (entre parênteses) para diversos valores de α - Libra Esterlina e Iene. p.	110
8.8	Proporção de dados em que a perda do portfolio excede o VaR e P-valor do teste de Christoffersen (entre parênteses) para diversos valores de α - Índices Nasdaq e S&P500. . . .	p. 112
A.1	Estatísticas descritivas dos retornos de NASDAQ e S&P500.	p. 116

A.2	Parâmetros estimados dos modelos das distribuições marginais e erro padrão das estimativas.	p. 118
A.3	Correlações condicionais estimadas para diversos ξ	p. 120
A.4	Parâmetros estimados das funções de acoplamento (Desvios padrão em parênteses).	p. 121
A.5	Parâmetros estimados das funções de acoplamento por medidas de dependência.	p. 121
A.6	Distância quadrática entre o acoplamento estimado e o acoplamento empírico na cauda inferior (quadrado $[0, \delta]^2$). . . .	p. 122
A.7	Proporção de dados (e número de observações), para $t = 751$ a 2971 , em que a perda do portfolio excede o VaR para $\alpha = 0,05$, $\alpha = 0,01$ e $\alpha = 0,005$. * = Número esperado de observações.	p. 123
A.8	Proporção de dados em que a perda do portfolio excede o VaR e P-valor do teste de Christoffersen (entre parênteses) para diversos valores de α - Índices Nasdaq e S&P500. . . .	p. 124

Lista de Figuras

2.1	Curvas de nível para os acoplamentos W , Π e M , $n = 2$. . .	p. 10
3.1	Correlações mínimas e máximas atingíveis $X \sim LN(0, 1)$ e $Y \sim LN(0, \sigma^2)$	p. 21
5.1	Parâmetros das 16 regiões para θ_t	p. 57
8.1	Gráficos de Retornos e Retornos Absolutos (Libra e Iene). .	p. 80
8.2	Função de Autocorrelação e Função de Autocorrelação Parcial - retornos de Libra e Iene.	p. 81
8.3	Diagrama de dispersão Retornos da Libra x Retornos do Iene.	p. 82
8.4	Função de Autocorrelação das inovações observadas - $\hat{\epsilon}_{1,t}$ e $\hat{\epsilon}_{2,t}$.	p. 86
8.5	Função de Autocorrelação das inovações observadas ao quadrado - $\hat{\epsilon}_{1,t}^2$ e $\hat{\epsilon}_{2,t}^2$	p. 87
8.6	Função de Autocorrelação das inovações padronizadas - $\hat{\eta}_{1,t}$ e $\hat{\eta}_{2,t}$	p. 88
8.7	Função de Autocorrelação das inovações padronizadas ao quadrado $\hat{\eta}_{1,t}^2$ e $\hat{\eta}_{2,t}^2$	p. 89
8.8	Função de Distribuição empírica - Séries u_1 e u_2 - Modelos GARCH-N e GARCH-t.	p. 90
8.9	Gráficos das inovações ajustadas padronizadas (Libra(vendido) e Iene).	p. 91
8.10	Correlações condicionais para as inovações dos modelos GARCH-N e GARCH-t.	p. 92

8.11	Curvas de nível da função de Acoplamento SJC estimada (linha cheia) e Empírica (linha tracejada). $((u, v) : C(u, v) = k$, onde $k = 0,005 ; 0,010; 0,015; 0,020; \dots$ a partir do canto inferior esquerdo).	p. 96
8.12	Curvas de nível da função de Acoplamento de Plackett estimada (linha cheia) e Empírica (linha tracejada). $((u, v) : C(u, v) = k$, onde $k = 0,005; 0,010; 0,015; 0,020; \dots$ a partir do canto inferior esquerdo).	p. 97
8.13	Curvas de nível da função de Acoplamento t-Student estimada (linha cheia) e Empírica (linha tracejada). $((u, v) : C(u, v) = k$, onde $k = 0,005 ; 0,010; 0,015; 0,020; \dots$ a partir do canto inferior esquerdo).	p. 98
8.14	Parâmetros estimados das funções de acoplamento para simulação.	p. 102
8.15	Parâmetros estimados das marginais GARCH para simulação.	p. 103
8.16	Retornos do portfolio e VaR estimado (acoplamento t-Student e marginais GARCH-t)	p. 105
8.17	VaR estimado (acoplamento SJC e marginais GARCH-t (tracejado) e GARCH-t (pontilhado) no portfolio) com $\alpha = 0,01$.	p. 111
A.1	Gráficos de Retornos e Ret. Absolutos - NASDAQ e S&P500.	p. 117
A.2	Gráficos das inovações ajustadas padronizadas (NASDAQ e S&P500).	p. 119

1 Introdução

O Valor-em-Risco (VaR) desempenha um papel muito importante em sistemas de gerenciamento de risco, sendo umas das mais populares medidas de risco. Apesar da simplicidade desta medida, ela não é facilmente estimada. Existem várias metodologias para a estimação do VaR, como o método de covariância-variância (também chamado de analítico), o método de simulações históricas e o método de Monte Carlo. O método analítico tornou-se bastante difundido com a divulgação de metodologia Riskmetrics, e adota a suposição de que os retornos dos ativos tenham distribuição conjunta normal multivariada. Neste caso, a matriz de covariância é uma medida natural da dependência entre os ativos e a variância é uma boa medida de risco. Na área financeira raramente esta suposição é adequada. É interessante notar que mesmo a normalidade na distribuição dos retornos de cada ativo da carteira, que também pode ser questionada, não é suficiente para que a distribuição conjunta destes retornos seja normal multivariada. O desvio da normalidade pode causar uma estimação inadequada do VaR, de modo que a instituição financeira pode estar correndo mais risco do que o estimado, ou por outro lado, sendo mais conservadora do que o necessário.

Descartando a suposição de normalidade multivariada dos retornos da carteira, a teoria de funções de acoplamento aparece como uma ferramenta muito poderosa na modelagem da distribuição conjunta destes retornos. Segundo esta teoria, qualquer distribuição conjunta n -dimensional pode ser decomposta em suas n distribuições marginais e uma função de acoplamento. Reciprocamente, a função de acoplamento produz uma distribuição multivariada a

partir de suas distribuições marginais e da estrutura de dependência entre as variáveis. As funções de acoplamento têm sido bastante estudadas na literatura estatística, sendo os livros textos de Joe (1997) e Nelsen (1999) citados freqüentemente pela maioria dos pesquisadores. Embora de uso recente em finanças, tem havido um número considerável de aplicações nesta área. Os artigos de Embrechts, McNeil e Strauman (1999b), Bouyé et al. (2003), Embrechts, Lindskog e McNeil (2003) dão uma boa idéia das possíveis aplicações de acoplamento em finanças. Apenas para citar algumas aplicações mais específicas, temos : Embrechts, McNeil e Strauman (1999a), Embrechts, Höing e Juri (2003) utilizam o acoplamento para modelar valores extremos e limites de risco; Cherubini e Luciano (2000) calculam o VaR utilizando a família de função de acoplamento de Arquimedes e a distribuição empírica histórica como estimativa das marginais; Rockinger e Jondeau (2001) utilizam o acoplamento de Plackett e a distribuição GARCH com perturbações t-Student assimétricas generalizada de Hansen (1994) para propor uma nova medida de dependência condicional; Georges (2001) utilizam o acoplamento normal para modelar tempos de exercício e precificação de derivativos; Meneguzzo e Vecchiato (2002) utilizam os acoplamentos para modelagem do risco de derivativos de créditos e Fortin e Kuzmics (2002) utilizam combinações lineares convexas de acoplamentos para modelagem do VaR de um portfolio composto pelos índices FTSE e DAX. Estes artigos, todos recentes, mostram a larga utilização que acoplamento pode ter em finanças. A dissertação de mestrado de Lucas (2003) utiliza a classe de acoplamentos de valores extremos e distribuições marginais empíricas e Generalizada de Pareto (GPD) para a estimação do VaR de um portfolio composto pelos índices Ibovespa e Merval.

Recentemente, a teoria de acoplamentos não-condicionais foi estendida para o caso condicional, permitindo o uso dos acoplamentos na modelagem de dependência que pode variar no tempo. A variação temporal no primeiro e segundo momentos condicionais é amplamente discutida na literatura, e deste modo permitir a variação temporal na dependência condicional entre séries temporais é natural. Patton (2003a) utiliza acoplamentos condicionais para a modelagem da distribuição conjunta de retornos de ativos.

No Capítulo 2, definiremos as funções de acoplamento e algumas propriedades que mostram sua importância, tanto para o caso não-condicional como o caso condicional. Algumas medidas de dependência entre as variáveis serão apresentadas e discutidas no Capítulo 3, assim como o conceito de dependência nas caudas. As funções de acoplamentos podem ser agrupadas em classes e famílias. Algumas famílias importantes de acoplamentos serão apresentadas no Capítulo 4, bem como algumas propriedades. No Capítulo 5 discutiremos algumas técnicas de inferência para os acoplamentos, com alguns métodos de estimação para os casos não-condicional e condicional. Também neste capítulo serão citadas alguns métodos de seleção da função de acoplamento adequada. Alguns modelos para as distribuições marginais são discutidos no Capítulo 6. Técnicas de simulação para funções de acoplamento serão apresentadas no Capítulo 7. No Capítulo 8 iremos modelar as distribuições marginais e conjunta para uma carteira composta de 2 ativos (taxas de câmbio da Libra Esterlina e do Iene, cotados em Dólar Americano). Serão utilizadas 3 funções de acoplamentos para a modelagem da dependência, sendo que o objetivo deste trabalho é a comparação de diferentes funções de acoplamento, e não encontrar a melhor delas. Após a modelagem e estimação de parâmetros, iremos utilizar técnicas de simulação para testar a adequação de cada modelo na estimação do VaR. Iremos comparar os resultados obtidos na estimação do VaR através de acoplamentos com os obtidos por métodos tradicionais como Simulações Históricas, modelos GARCH bivariado e univariado e alisamento exponencial. As conclusões dos resultados obtidos serão apresentadas no Capítulo 9. No Apêndice A serão apresentados resumidamente os resultados obtidos para outra carteira, composta pelos índices norte-americanos Nasdaq e S&P500.

2 *Funções de Acoplamento*

2.1 Introdução

Neste capítulo definiremos as funções de acoplamento e algumas propriedades e teoremas que mostram sua importância, tanto para o caso não-condicional como o caso condicional.

No livro de Nelsen (1999), (Capítulo 2) são definidas e apresentadas as principais propriedades das funções de acoplamento. De acordo com Nelsen (pg. 1), o acoplamento pode ser visto sob dois pontos de vista : "sob um ponto de vista, acoplamento são funções que juntam ou acoplam funções distribuições conjuntas à suas funções distribuições marginais. Alternativamente, acoplamento são funções distribuições multivariadas, cujas marginais unidimensionais são uniformes no intervalo (0,1)".

Antes de definirmos as funções de acoplamento, iremos apresentar algumas notações que serão utilizadas no resto da dissertação. Por notação temos que $\overline{\mathbb{R}}$ é a reta estendida dos reais, ou seja, $\mathbb{R} \cup \{-\infty, +\infty\}$. Se F for uma função distribuição, de uma variável aleatória X iremos utilizar a notação $X \sim F$. Neste caso, chamaremos de $F^{(-1)}$ a inversa generalizada de F , ou seja, $F^{(-1)} = \inf\{x | F(x) \geq u\}, u \in [0, 1]$. Similarmente, para o caso condicional, seja X condicionada em um vetor aleatório $\mathbf{W}$. Utilizaremos a mesma notação que no caso anterior, ou seja, teremos $X/\mathbf{W} \sim F$ onde F é a distribuição condicional de X dado $\mathbf{W}$ e a inversa generalizada de F em relação ao primeiro argumento será também denotada por $F^{(-1)}$.

2.2 Caso não-condicional

2.2.1 Definição

Nesta seção apresentaremos o caso não-condicional, ou seja, em que a dependência e as distribuições marginais, mesmo condicionadas, não variam com o tempo.

Definição 2.1 (*Funções de acoplamento não-condicionais*) : Uma função C , com domínio em $[0, 1]^n$ e contradomínio em $[0, 1]$, é uma função de acoplamento n -dimensional se ela tem as seguintes propriedades:

1. Para todo $\mathbf{x} \in [0, 1]^n$ tal que pelo menos um elemento de $\mathbf{x}$ é 0, então $C(\mathbf{x}) = 0$;
2. $C(1, \dots, 1, x_i, 1, \dots, 1) = x_i$ para todo $i \in \{1, \dots, n\}$, $x_i \in [0, 1]$;
3. para $\forall (a_1, \dots, a_n), (b_1, \dots, b_n) \in [0, 1]^n$ com $a_i \leq b_i$, temos :

$$V_C([a, b]) = \sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{i_1+\dots+i_n} C(x_{1i_1}, \dots, x_{ni_n}) \geq 0. \quad (2.1)$$

onde $x_{j1} = a_j$ e $x_{j2} = b_j$ para todo $j \in \{1, \dots, n\}$. ■

Nesta dissertação muitas vezes chamaremos a função de acoplamento simplesmente de acoplamento.

Se considerarmos as funções de acoplamento em um contexto de variáveis aleatórias, podemos definí-las de uma outra maneira, que podemos provar ser equivalente à definição anterior, no caso contínuo, que é o caso de nosso interesse.

Definição 2.2 Considere o vetor aleatório $\mathbf{X} = (X_1, \dots, X_n)^t$ com função distribuição conjunta H , e onde X_i tem distribuição marginal F_i , e F_i contínua

para $i = 1, \dots, n$. A função de acoplamento C do vetor $\mathbf{X}$ é a função distribuição conjunta do vetor aleatório $\mathbf{U} = (U_1, \dots, U_n)$ onde $U_i \equiv F_i(X_i)$ para $i = 1, \dots, n$. ■

Segue então que:

$$\begin{aligned} H(x_1, \dots, x_n) &= P[X_1 \leq x_1, \dots, X_n \leq x_n], \\ &= P[F_1(X_1) \leq F(x_1), \dots, F_n(X_n) \leq F(x_n)], \\ &= C(F_1(x_1), \dots, F_n(x_n)). \end{aligned} \tag{2.2}$$

onde C é a função de acoplamento do vetor $\mathbf{X}$. Ou seja, podemos interpretar o acoplamento como uma função que liga as distribuições marginais do vetor aleatório em sua distribuição conjunta.

2.2.2 Propriedades

Vamos agora mostrar alguns teoremas e propriedades satisfeitas pelas funções de acoplamento. Maiores detalhes podem ser encontrados em Nelsen (1999), Bouyé et al. (2003) e Patton (2003a).

Teorema 2.1 (Teorema de Sklar) *Seja H uma função distribuição conjunta n -dimensional com marginais $F_1, F_2, \dots, F_n$. Então existe um acoplamento n -dimensional C tal que para qualquer $\mathbf{x} \in \mathbb{R}^n$,*

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \tag{2.3}$$

Se $F_1, F_2, \dots, F_n$ são todos contínuos, então C é único, caso contrário C é unicamente determinado no conjunto $Im(F_1) \times Im(F_2) \times \dots \times Im(F_n)$. Reciprocamente, se C é um acoplamento n -dimensional e $F_1, F_2, \dots, F_n$ são funções de distribuição, então a função H definida por (2.3) é uma função distribuição conjunta n -dimensional. ■

Prova. Pode ser encontrada em Nelsen (1999), (pg 18).

A recíproca do teorema de Sklar é muito importante para a modelagem de densidades multivariadas em finanças. Esta recíproca implica que a partir de duas distribuições univariadas, de qualquer tipo (não necessariamente da mesma família), e de qualquer acoplamento, teremos definido uma distribuição multivariada válida. Este resultado é particularmente útil uma vez que na literatura da área estatística e econômica, nós temos uma vasta seleção de distribuições paramétricas univariadas, mas o conjunto de distribuições multivariadas disponíveis é muito menor.

O corolário do teorema de Sklar apresentado por Nelsen (1999), (pg 19) mostra que o número de distribuições paramétricas multivariadas que podem ser definidas aumenta consideravelmente, uma vez que nós podemos extrair uma função de acoplamento de qualquer distribuição multivariada e usá-la independentemente das distribuições marginais originais.

Corolário 2.1 *Sejam $H, C, F_1, \dots, F_n$ como no Teorema 2.1 e sejam $F_1^{(-1)}, \dots, F_n^{(-1)}$ as inversas generalizadas de $F_1, \dots, F_n$ respectivamente. Então para qualquer $\mathbf{u} \in [0, 1]^n$*

$$C(u_1, \dots, u_n) = H(F_1^{(-1)}(u_1), F_2^{(-1)}(u_2), \dots, F_n^{(-1)}(u_n)). \quad (2.4)$$

■

Exemplo 2.1 *Sejam $n=2$,*

$$H(x, y) = \begin{cases} \frac{(x+1)(e^y-1)}{x+2e^y-1} & \text{se } (x, y) \in [-1, 1] \times [0, +\infty]; \\ (1 - e^{-y}) & \text{se } (x, y) \in (1, +\infty] \times [0, +\infty]; \\ 0 & \text{caso contrário.} \end{cases}$$

Temos então que

$$F_1(x) = H(x, +\infty) = \begin{cases} 0 & \text{se } x < -1; \\ (x+1)/2 & \text{se } x \in [-1, 1]; \\ 1 & \text{se } x > 1, \end{cases}$$

$$F_2(y) = H(+\infty, y) = \begin{cases} 0 & \text{se } y < 0; \\ 1 - e^{-y} & \text{se } y \geq 0. \end{cases}$$

As inversas generalizadas de F_1 e F_2 são dadas por $F_1^{(-1)}(u) = 2u - 1$ e $F_2^{(-1)}(v) = -\ln(1 - v)$ para $u, v \in [0, 1]$. Como $\text{Im}(F_1) = \text{Im}(F_2) = [0, 1]$, utilizando o corolário (2.1) temos que o acoplamento C é dado por :

$$C(u, v) = \frac{uv}{u + v - uv}. \quad (2.5)$$

■

Considere agora as funções M^n, Π^n e W^n definidas em $[0, 1]^n$ como :

$$M^n(\mathbf{u}) = \min(u_1, \dots, u_n),$$

$$\Pi^n = u_1 \cdot u_2 \dots u_n,$$

$$W^n(\mathbf{u}) = \max(\sum_{i=1}^n u_i - n + 1, 0).$$

As funções M^n e Π^n são acoplamentos n -dimensionais para todo $n \geq 2$, enquanto a função W^n não é uma função de acoplamento para qualquer $n \geq 3$, conforme o seguinte exemplo, encontrado em Embrechts, Lindskog e McNeil (2003). Para o caso $n = 2$ chamaremos estas funções simplesmente de W, Π e M .

Exemplo 2.2 (*Exemplo de que W^n não ser acoplamento*)

Considere o n -cubo $[1/2, 1]^n$.

Temos então que $V_{W^n}([1/2, 1]^n) = 1 - n/2 < 0$ para $n \geq 3$. ■

A seguir temos dois teoremas relacionados com estas funções.

Teorema 2.2 (*Inequação de Fréchet-Hoeffding*). Se C é qualquer acoplamento n -dimensional, então para todo $\mathbf{u} \in [0, 1]^n$,

$$W^n(\mathbf{u}) \leq C(\mathbf{u}) \leq M^n(\mathbf{u}). \quad (2.6)$$

■

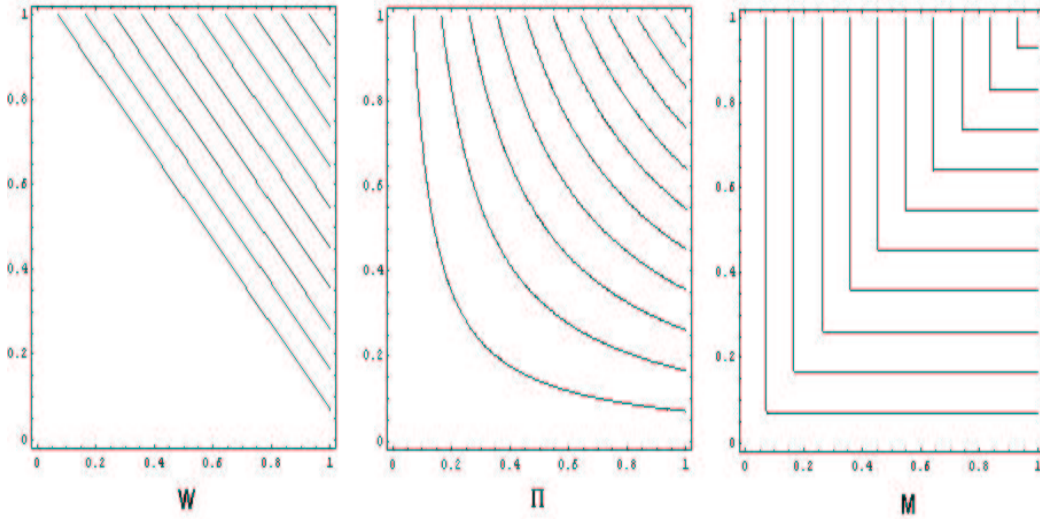


Figura 2.1: Curvas de nível para os acoplamentos W , Π e M , $n = 2$.

Estas funções são conhecidas como limitantes de Fréchet-Hoeffding. Na Figura 2.2.2 temos as curvas de nível para os acoplamentos W , Π e M . Apesar de W^n não ser uma função de acoplamento para $n \geq 3$, temos o seguinte teorema encontrado em Nelsen (1999) (pg 42).

Teorema 2.3 *Para qualquer $n \geq 3$ e qualquer $\mathbf{u} \in [0, 1]^n$, existe um acoplamento n -dimensional C (o qual depende de $\mathbf{u}$) tal que*

$$C(\mathbf{u}) = W^n(\mathbf{u}). \quad (2.7)$$

■

Prova. Em Sklar (1998).

Podemos definir uma ordenação entre as funções de acoplamento, da seguinte maneira :

Definição 2.3 *Sejam C_1 e C_2 acoplamentos, nós dizemos que C_1 é menor do que C_2 , e escrevemos $C_1 \prec C_2$ se $C_1(\mathbf{u}) \leq C_2(\mathbf{u})$, $\forall \mathbf{u} \in [0, 1]^n$. ■*

Temos então para qualquer função de acoplamento C que $W_n \prec C \prec M_n$. Deste resultado temos que :

$$F_L(\mathbf{x}) \leq F(\mathbf{x}) \leq F_U(\mathbf{x}), \forall \mathbf{x} \in \mathfrak{R}^n, \quad (2.8)$$

onde $F_L(\mathbf{x}) = W^n(F_1(x_1), \dots, F_n(x_n))$ e $F_U(\mathbf{x}) = M^n(F_1(x_1), \dots, F_n(x_n))$. Deste modo os limitantes de Fréchet-Hoeffding para os acoplamentos determinam também limitantes para a função distribuição multivariada.

Duas funções de acoplamento nem sempre são comparáveis entre si, uma delas pode ser maior que a outra em uma parte do domínio e menor em outra parte.

Temos famílias de acoplamentos que são totalmente ordenadas. Nós chamamos uma família de acoplamentos $\{C_\theta\}$ totalmente ordenada de *positivamente ordenada* se $C_\alpha \prec C_\beta$ toda vez que $\alpha \prec \beta$. Analogamente chamamos $\{C_\theta\}$ de *negativamente ordenada* se $C_\alpha \succ C_\beta$ toda vez que $\alpha \prec \beta$.

O seguinte teorema em Nelsen (1999), (pg 43) caracteriza a dependência no acoplamento Π^n .

Teorema 2.4 *Para $n \geq 2$, sejam $X_1, \dots, X_n$ variáveis aleatórias contínuas. Então $X_1, X_2, \dots, X_n$ são independentes se, e somente se, a função de acoplamento de $X_1, X_2, \dots, X_n$ é Π^n . ■*

A proposição seguinte mostra uma característica interessante das funções de acoplamento na representação da dependência entre as variáveis. Esta característica é que a estrutura de dependência representada pelo acoplamento é *invariante* sob transformações crescentes e contínuas nas marginais. Vamos utilizar a definição encontrada em Embrechts, McNeil e Strauman (1999b).

Proposição 2.1 *Se $(X_1, X_2, \dots, X_n)^t$ tem acoplamento C e $T_1, \dots, T_n$ são funções contínuas crescentes, então $((T_1(X_1), T_2(X_2), \dots, T_n(X_n))^t$ também tem acoplamento C . ■*

Como simples exemplo da utilidade deste resultado, vamos supor que estejamos modelando os retornos de ativos de diversos tipos. Caso decidamos trabalhar com o logaritmo dos retornos, a função de acoplamento não irá se alterar.

2.3 Caso condicional

2.3.1 Definição

Como notado em Patton (2003a), em economia, e particularmente em análise de séries temporais econômicas, o caso em que as variáveis aleatórias são condicionadas em algumas variáveis pré-determinadas é muito importante. Nesta seção nós iremos mostrar como os resultados obtidos na teoria de funções de acoplamento pode ser estendido para permitir o condicionamento.

Em Patton (2003a) o condicionamento é apresentado apenas para duas variáveis de interesse e a variável de condicionamento com dimensão 1. Iremos apresentar um resultado mais geral para n variáveis.

Como no caso não-condicional, iremos apresentar duas definições de acoplamento condicional, uma matemática e outra no contexto de variáveis aleatórias.

Definição 2.4 (*Função de Acoplamento Condicional*) Uma função C , com domínio em $[0, 1]^n \times \Lambda$ e contradomínio em $[0, 1]$, é uma função de acoplamento condicional n -dimensional se ela tem as seguintes propriedades:

1. Para cada $w \in \Lambda$, e para todo $\mathbf{x} \in [0, 1]^n$ tal que pelo menos um elemento de $\mathbf{x}$ é 0, então $C(\mathbf{x}/\mathbf{w}) = 0$,
2. $C(1, \dots, 1, x_i, 1, \dots, 1/\mathbf{w}) = x_i$ para todo $i \in \{1, \dots, n\}$, $x_i \in [0, 1]$ e para cada $w \in \Lambda$,
3. para $\forall (a_1, \dots, a_n), (b_1, \dots, b_n) \in [0, 1]^n$ com $a_i \leq b_i$, e para cada $\mathbf{w}$

$\in \Lambda$, temos :

$$V_{C/\mathbf{w}}([a, b]) = \sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{i_1+\dots+i_n} C(x_{1i_1}, \dots, x_{ni_n}/\mathbf{w}) \geq 0, \quad (2.9)$$

onde $x_{j1} = a_j$ e $x_{j2} = b_j$ para todo $j \in \{1, \dots, n\}$. ■

Definição 2.5 A função de acoplamento condicional de $\mathbf{X}/\mathbf{W}$, onde $X_i/\mathbf{W} \sim F_i$ e F_i contínua para $i = 1, \dots, n$, é a função de distribuição condicional conjunta do vetor $\mathbf{U} = (U_1, \dots, U_n)$ onde $U_i \equiv F_i(X_i/\mathbf{W})$ para $i=1, \dots, n$. ■

Ou seja, temos que o acoplamento n-dimensional é a função distribuição conjunta de n variáveis aleatórias que condicionalmente são uniformemente distribuídas.

2.3.2 Propriedades

O teorema de Sklar pode ser estendido para o caso condicional, como feito e provado por Patton (2003a) (a prova foi feita apenas para o caso em que $\mathbf{W}$ é unidimensional e $\mathbf{X}$ bidimensional, mas no artigo é citado que a extensão é feita do mesmo modo).

Teorema 2.5 (Teorema de Sklar para distribuições condicionais) Seja F_i a distribuição condicional de $X_i/\mathbf{W}$ para $i = 1, \dots, n$ e seja H a distribuição condicional conjunta de $\mathbf{X}/\mathbf{W}$. Então existe um acoplamento n-dimensional C tal que para qualquer $\mathbf{x} \in \mathbb{R}^n$ e para qualquer $\mathbf{w} \in \Lambda$,

$$H(x_1, \dots, x_n/\mathbf{w}) = C(F_1(x_1/\mathbf{w}), \dots, F_n(x_n/\mathbf{w})/\mathbf{w}). \quad (2.10)$$

Se $F_1, F_2, \dots, F_n$ são todos contínuos, então C é único, caso contrário C é unicamente determinado no conjunto $Im(F_1) \times Im(F_2) \times \dots \times Im(F_n)$. Reciprocamente, se C é um acoplamento n-dimensional e F_i é função de distribuição

condicional de $X_i/\mathbf{W}$ para $i = 1, \dots, n$, então a função H definida por (2.10) é uma função distribuição conjunta condicional. ■

Uma restrição na extensão do teorema de Sklar é que o vetor de condicionamento $\mathbf{W}$ deve ser o mesmo para a cópula e para as distribuições marginais. Da mesma maneira que no caso não-condicional, temos o seguinte corolário :

Corolário 2.2 *Sejam $H, C, F_1, \dots, F_n$ como no Teorema (2.5) e sejam $F_1^{(-1)}, \dots, F_n^{(-1)}$ as inversas generalizadas com respeito ao primeiro argumento de $F_1, \dots, F_n$ respectivamente. Então para qualquer $\mathbf{u} \in [0, 1]^n$ e $\mathbf{w} \in \Lambda$*

$$C(u_1, \dots, u_n/\mathbf{w}) = H(F_1^{(-1)}(u_1/\mathbf{w}), F_2^{(-1)}(u_2/\mathbf{w}), \dots, F_n^{(-1)}(u_n/\mathbf{w})/\mathbf{w}). \quad (2.11)$$

■

3 Medidas de Dependência

3.1 Introdução

Funções de acoplamento são um caminho natural para estudar e caracterizar a dependência entre variáveis aleatórias. Para a identificação da função de acoplamento adequada, devemos ter uma forma apropriada de caracterizar esta dependência, se possível através de um índice, que chamaremos de medida de dependência.

Neste capítulo iremos apresentar diversas medidas de dependências entre as variáveis aleatórias e suas propriedades. O objetivo do capítulo é apresentarmos as medidas de dependência que sejam adequadas para o uso com funções de acoplamento.

Iniciaremos caracterizando a dependência em séries financeiras na Seção 3.2. A seguir, na Seção 3.3, discutiremos o caso extremo de dependência, que é o de dependência perfeita. Após isto apresentaremos o coeficiente de correlação linear. Esta é uma das medidas de associação clássica, que mostraram ser geralmente inadequadas para séries de retornos. Desta forma apresentaremos algumas medidas de associação alternativas, que chamaremos de medidas de concordância. Finalmente, na Seção 3.6, apresentaremos um conceito alternativo de dependência nas caudas.

3.2 Dependência em séries financeiras

Dados de retornos financeiros costumam apresentar os chamados fatos estilizados, que são desvios em relação à distribuição normal. Estes fatos estilizados para as distribuições univariadas de séries financeiras serão discutidos com mais detalhes na Seção 6.1, onde também é dada a definição de retorno financeiro.

Estudos recentes relatam também a existência de desvios da normalidade multivariada, como por exemplo a dependência assimétrica, em que dois ativos exibem maior dependência quando os mercados estão em baixa do que quando estão em alta. Esta dependência assimétrica também não é encontrada na classe de distribuições elípticas, como observado por Longin e Solnik (2001).

Longin e Solnik (2001) utilizam a teoria de valores extremos para estudar a estrutura de dependência em mercados internacionais no caso bivariado. Eles definem a correlação entre os retornos de dois ativos condicionada à uma certa região destes retornos, acima ou abaixo de um certo nível especificado. Ang e Chen (2002) utilizam a notação que apresentaremos a seguir para este tipo de correlação condicional. Seja $(\tilde{x}, \tilde{y})$ o vetor de retornos padronizados de dois ativos, com correlação ρ . A correlação linear condicionada à um certo nível ξ é definida como :

$$\bar{\rho}(\xi) = \begin{cases} \text{corr}(\tilde{x}, \tilde{y} / \tilde{x} > \xi, \tilde{y} > \xi; \rho) & \text{se } \xi \geq 0, \\ \text{corr}(\tilde{x}, \tilde{y} / \tilde{x} < \xi, \tilde{y} < \xi; \rho) & \text{se } \xi \leq 0, \end{cases} \quad (3.1)$$

onde *corr* denota a correlação condicionada à região especificada.

Sob a hipótese de normalidade multivariada com correlação constante, podemos demonstrar que a correlação entre grandes retornos (acima de um certo nível dado) assintoticamente vai para zero conforme este nível cresce. Da mesma maneira a correlação entre retornos negativos abaixo de um nível dado vai para zero conforme este nível (que é negativo) cresce em valor absoluto.

Longin e Solnik (2001) não encontraram este comportamento nos retornos financeiros dos índices dos cinco maiores mercados de ações (EUA, Reino

Unido, França, Alemanha e Japão), onde os estudos foram feitos considerando-se pares de retornos. Empiricamente, a correlação condicionada entre grandes retornos negativos não converge para zero, mas tende a crescer conforme o nível ξ ($\xi \leq 0$), cresce em valor absoluto. Deste modo, apenas considerando este aspecto, hipótese da normalidade multivariada já é rejeitada com evidência altamente significativa. Já a correlação entre grandes retornos positivos tende a decrescer e convergir para zero quando o nível ξ cresce, de modo que a hipótese de normalidade multivariada não é rejeitada. Deste modo, Longin e Solnik (2001) encontraram evidências de que a correlação condicionada entre retornos financeiros cresce em relação ao caso normal em mercados que estão em baixa, mas não em mercados que estão em alta.

No caso de uma posição comprada em diversos ativos, este fato empírico pode acarretar problemas para quem trabalha com gerenciamento de risco de mercado, uma vez que as correlações entre os ativos tendem a aumentar em mercados em baixa, de modo a cancelar possíveis efeitos benéficos que uma baixa correlação entre os ativos traz para o controle de risco de uma carteira. Deste modo, deve ser escolhida uma função de acoplamento que possibilite a modelagem deste fato estilizado.

3.3 Dependência perfeita

Como dito na Subseção 2.2.2, para cada função de acoplamento n -dimensional C , nós temos pela inequação de Fréchet-Hoeffding (Teorema 2.2) as funções limitantes W^n e M^n , onde $W^n(\mathbf{u}) = \max(\sum_{i=1}^n u_i - n + 1, 0)$ é o limitante inferior para as funções de acoplamento e $M^n(\mathbf{u}) = \min(u_1, \dots, u_n)$ o limitante superior. Nestes casos, dizemos que temos um dependência perfeita. A seguir vamos estudar a dependência perfeita para os casos bidimensional e n -dimensionais.

Definição 3.1 *Seja $\mathbf{X} = (X_1, \dots, X_n)$ um vetor aleatório com função de acoplamento M^n . Neste caso, dizemos que as variáveis aleatórias são **comonotônicas**. ■*

Uma definição equivalente encontrada em Embrechts, Höing e Juri (2003) é dada por :

Definição 3.2 *Seja $\mathbf{X} = (X_1, \dots, X_n)$ um vetor aleatório. As variáveis aleatórias $X_1, X_2, \dots, X_n$ são chamadas de comonotônicas se :*

$$(X_1, \dots, X_n) = (f_1(Z), \dots, f_n(Z)) \text{ em lei ,} \quad (3.2)$$

onde $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$ são crescentes e Z é alguma variável aleatória. ■

Esta é precisamente a representação que motiva o uso do conceito de comonotonicidade em aplicações na área financeira. A variável aleatória Z pode ser vista como um fator comum nos diversos riscos existentes.

No caso bidimensional, podemos caracterizar também o limitante inferior, como mostra o teorema a seguir, cuja prova pode ser encontrada em Embrechts, McNeil e Strauman (1999b).

Teorema 3.1 *Seja $(X, Y)^T$ um vetor aleatório com acoplamento W ou M . Então existem duas funções monótonas $\alpha, \beta : \mathbb{R} \rightarrow \mathbb{R}$ e uma variável aleatória Z , tal que (X, Y) tem mesma distribuição que $(\alpha(Z), \beta(Z))$ com α crescente e β decrescente no primeiro caso (W) e ambas α e β crescentes no último caso (M). A recíproca também é verdadeira. ■*

Equivalentemente, W e M são funções de distribuições bivariadas dos vetores aleatórios $(U, 1 - U)^T$ e $(U, U)^T$ respectivamente, onde $U \sim U(0, 1)$.

Definição 3.3 *Se $(X, Y)^T$ é um vetor aleatório com acoplamento M , então X e Y são ditas serem **comonotônicas**. Se o vetor tem cópula W então X e Y são ditas serem **contra-monotônicas** ■*

Sejam F_1 e F_2 as funções distribuições de X e Y respectivamente. Se pelo menos uma das funções F_1 ou F_2 tiver descontinuidade, então o acoplamento não é único, assim W e M são acoplamentos possíveis.

No caso contínuo, temos uma interpretação simples destas definições. Quando as variáveis aleatórias X e Y são comonotônicas, então Y é quase certamente uma função crescente de X . Já quando as variáveis X e Y são contra-monotônicas, então Y é quase certamente uma função decrescente de X .

No caso em que F_1 e F_2 são contínuas uma versão mais forte do resultado pode ser utilizada, como citado em Embrechts, McNeil e Strauman (1999b):

Teorema 3.2 *Seja $(X, Y)^T$ um vetor aleatório, onde X e Y tem distribuições marginais contínua F_1 e F_2 respectivamente e acoplamento W ou M . Então temos :*

$C = W \Leftrightarrow Y = T(X)$ quase certamente, onde $T = F_2^{-1} \circ (1 - F_1)$ (é decrescente),

$C = M \Leftrightarrow Y = T(X)$ quase certamente, onde $T = F_2^{-1} \circ F_1$ (é crescente). ■

3.4 Correlação Linear

Definição 3.4 (Correlação Linear). *Seja $(X, Y)^T$ um vetor aleatório com variâncias finitas não-nulas. O coeficiente de correlação linear para $(X, Y)^T$ é definido como :*

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}, \quad (3.3)$$

onde $Cov(X, Y)$ é a covariância de $(X, Y)^T$ e $Var(X)$ e $Var(Y)$ são variâncias de X e Y respectivamente. ■

No caso de um vetor aleatório n -dimensional $\mathbf{X} = (X_1, \dots, X_n)$, as correlações são dispostas em uma matriz. Definiremos então

$\Sigma = \{Cov(X_i, X_j)\}_{i=1, \dots, n, j=1, \dots, n}$ a matriz de covariâncias do vetor $\mathbf{X}$ e a matriz de correlações $\Omega = \{\rho(X_i, X_j)\}_{i=1, \dots, n, j=1, \dots, n}$.

Como notado em Embrechts, McNeil e Strauman (1999a), a correlação linear é o coração de modelos como o CAPM (Capital Asset Pricing Model) e a Teoria de Precificação por Arbitragem, onde é utilizada como uma medida de dependência entre instrumentos financeiros. No entanto, a correlação linear tem sido utilizada em todas as áreas do gerenciamento de risco, mesmo quando a suposição de riscos com distribuição multivariada normal é completamente inadequada.

Temos diversos problemas no uso da correlação linear, que discutiremos a seguir : existência de momentos, limites atingíveis e em especial a não-invariância sob transformações monótonas.

As variâncias das variáveis aleatórias tem que ser finitas para que a correlação linear seja definida. Isto não é ideal para uma medida de dependência e pode causar problemas quando trabalhamos com distribuições com caudas pesadas. Por exemplo, a covariância e correlação entre os dois componentes de uma distribuição t_ν bivariada não estão definidas para $\nu \leq 2$.

Vamos supor que temos duas variáveis aleatórias X e Y com funções distribuição marginais F_1 e F_2 , respectivamente. Em geral, nem todos os valores de correlação linear entre -1 e +1 são atingíveis para uma dada distribuição conjunta F especificada.

Hoeffding (1940) apresenta a ligação entre a covariância das variáveis aleatórias X e Y , e suas distribuições conjunta e marginais. Temos que ρ deve obedecer

$$\rho_{min} \leq \rho \leq \rho_{max}, \quad (3.4)$$

onde

$$\rho_{min} = \frac{\int \int_D (W(x, y) - F_1(x)F_2(y)) dx dy}{\sqrt{Var(X)}\sqrt{Var(Y)}},$$

e

$$\rho_{max} = \frac{\int \int_D (M(x, y) - F_1(x)F_2(y)) dx dy}{\sqrt{Var(X)}\sqrt{Var(Y)}}.$$

O valor ρ_{min} é atingido no caso contramonotônico (dependência negativa perfeita) e ρ_{max} no caso comonotônico (dependência positiva perfeita). Isto é uma característica inadequada da correlação linear, fora do caso normal variáveis comonôtonicas não tem necessariamente correlação 1 e variáveis contramonôtonicas não tem necessariamente correlação -1. Por exemplo, sejam X e Y variáveis aleatórias com distribuições $LN(0,1)$ (Log-Normal padrão). Então pode ser demonstrado que : $\rho_{min} = \frac{e^{-\sigma}-1}{\sqrt{e-1}\sqrt{e^{\sigma^2}-1}}$ e $\rho_{max} = \frac{e^{\sigma}-1}{\sqrt{e-1}\sqrt{e^{\sigma^2}-1}}$. O gráfico destas funções é apresentado na Figura 3.1.

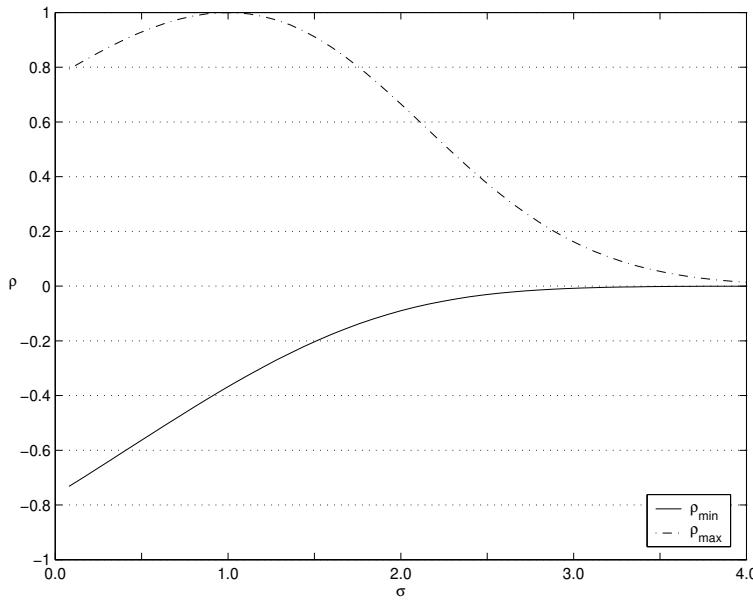


Figura 3.1: Correlações mínimas e máximas atingíveis $X \sim LN(0,1)$ e $Y \sim LN(0, \sigma^2)$.

Por último, a correlação linear possui uma grave deficiência, pois não é invariante sob uma transformação estritamente crescente não-linear $T : \Re \rightarrow \Re$. Deste modo, ela se torna uma medida inadequada para ser utilizada na seleção de funções de acoplamento, uma vez que os acoplamentos são invariantes sob quaisquer transformações estritamente crescentes. Para duas variáveis aleatórias em $\Re$ nós temos em geral

$$\rho(T(X), T(Y)) \neq \rho(X, Y).$$

Por exemplo, se nós considerarmos a distribuição normal bivariada padrão com correlação ρ e a transformação $T(x) = \Phi(x)$ (a função de distribuição da normal padrão), nós temos em Joag-dev (1984):

$$\rho(T(X), T(Y)) = \frac{6}{\pi} \arcsen\left(\frac{\rho}{2}\right). \quad (3.5)$$

3.5 Medidas de concordância

Apresentaremos nesta seção medidas alternativas à correlação linear, que são as chamadas medidas de concordância, sendo duas delas muito conhecidas, o *tau de Kendall* e o *rho de Spearman*.

3.5.1 Concordância

Sejam $(x, y)^T$ e $(\tilde{x}, \tilde{y})^T$ duas observações do vetor $(X, Y)^T$, onde X e Y são variáveis aleatórias contínuas. Então $(x, y)^T$ e $(\tilde{x}, \tilde{y})^T$ são ditos serem **concordantes** se $(x - \tilde{x})(y - \tilde{y}) > 0$ e **discordantes** caso $(x - \tilde{x})(y - \tilde{y}) < 0$. Ou seja, o par será concordante se $x < \tilde{x}$ e $y < \tilde{y}$ ou $x > \tilde{x}$ e $y > \tilde{y}$. O par será discordante se $x < \tilde{x}$ e $y > \tilde{y}$ ou $x > \tilde{x}$ e $y < \tilde{y}$.

Para demonstrarmos a função que os acoplamentos têm nas medidas de concordância, vamos definir uma função de concordância Q , que é a diferença

entre as probabilidades de concordância e discordância entre dois vetores aleatórios (X, Y) e $(\tilde{X}, \tilde{Y})$ de variáveis aleatórias contínuas, com funções distribuição conjuntas H_1 e H_2 , mas com marginais comuns F_1 e F_2 . Então mostraremos pelo próximo teorema, encontrado em Nelsen (1999), (pg 127) que esta função Q depende das distribuições de (X, Y) e $(\tilde{X}, \tilde{Y})$ apenas através de suas funções de acoplamento.

Teorema 3.3 *Sejam (X, Y) e $(\tilde{X}, \tilde{Y})$ vetores aleatórios independentes formados por variáveis aleatórias contínuas, com funções distribuição conjunta H e $\tilde{H}$ respectivamente, e com função marginais comuns F_1 (de X e $\tilde{X}$) e F_2 (de Y e $\tilde{Y}$). Sejam C e $\tilde{C}$ as funções de acoplamento de (X, Y) e $(\tilde{X}, \tilde{Y})$ respectivamente, de modo que $H(x, y) = C(F_1(x), F_2(y))$ e $\tilde{H}(x, y) = \tilde{C}(F_1(x), F_2(y))$. Então definiremos :*

$$Q = P\{(X - \tilde{X})(Y - \tilde{Y}) > 0\} - P\{(X - \tilde{X})(Y - \tilde{Y}) < 0\}.$$

É válido então que :

$$Q = Q(C, \tilde{C}) = 4 \int \int_{[0,1]^2} \tilde{C}(u, v) dC(u, v) - 1. \blacksquare$$

Esta função Q tem outras propriedades interessantes, como pode ser visto por este corolário.

Corolário 3.1 *Sejam C , $\tilde{C}$ e Q como no Teorema (3.3). Então :*

1. *Q é simétrica em seus argumentos, ou seja $Q(C, \tilde{C}) = Q(\tilde{C}, C)$;*
2. *Q é não decrescente em cada argumento, ou seja, se $C \prec C'$ então $Q(C, \tilde{C}) \leq Q(C', \tilde{C})$;*
3. *Seja $\hat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$ o acoplamento de sobrevivência. Então $Q(C, \tilde{C}) = Q(\hat{C}, \hat{\tilde{C}})$. $\blacksquare$*

Exemplo 3.1 A função Q pode facilmente ser avaliada para os acoplamentos W , Π e M . Sabemos que o suporte de M está na diagonal $u = v$ de $[0, 1]^2$. Segue então que :

$$\int \int_{[0,1]^2} g(u, v) dM(u, v) = \int_0^1 g(u, u) du.$$

Então nós temos :

$$\begin{aligned} Q(M, M) &= 4 \int \int_{[0,1]^2} \min(u, v) dM(u, v) - 1 = 4 \int_0^1 u du - 1 = 1, \\ Q(M, \Pi) &= 4 \int \int_{[0,1]^2} uv dM(u, v) - 1 = 4 \int_0^1 u^2 du - 1 = 1/3, \\ Q(M, W) &= 4 \int \int_{[0,1]^2} \max(u + v - 1, 0) dM(u, v) - 1, \\ &= 4 \int_{1/2}^1 (2u - 1) du - 1 = 0. \end{aligned}$$

Similarmente, como o suporte de W é a diagonal secundária $v = 1 - u$, nós temos :

$$\int \int_{[0,1]^2} g(u, v) dW(u, v) = \int_0^1 g(u, 1 - u) du.$$

Então nós temos :

$$\begin{aligned} Q(W, \Pi) &= 4 \int \int_{[0,1]^2} uv dW(u, v) - 1 = 4 \int_0^1 u(1 - u) du - 1 = -1/3, \\ Q(W, W) &= 4 \int \int_{[0,1]^2} \max(u + v - 1, 0) dW(u, v) - 1 = 4 \int_{1/2}^1 0 du - 1 = -1. \end{aligned}$$

Finalmente, como $d\Pi(u, v) = dudv$,

$$Q(\Pi, \Pi) = 4 \int \int_{[0,1]^2} uv d\Pi(u, v) - 1 = 4 \int_{1/2}^1 uv dudv - 1 = 0.$$

Seja C uma função de acoplamento arbitrária. Como Q é a diferença en-

tre duas probabilidades, $Q(C, C) \in [-1, 1]$. Como consequência da parte 2 do corolário 3.1 e os valores de Q neste exemplo, segue que : $Q(C, M) \in [0, 1]$, $Q(C, W) \in [-1, 0]$ e $Q(C, \Pi) \in [-1/3, 1/3]$ ■

Temos algumas propriedades que são importantes para que uma medida possa ser utilizada como uma medida de concordância. A definição seguinte pode ser encontrada em Scarsini (1984).

Definição 3.5 Uma medida nos reais κ de dependência entre duas variáveis aleatórias contínuas X e Y com função de acoplamento C é uma **medida de concordância** se satisfaz as seguintes propriedades :

1. κ é definido para cada para X, Y de variáveis aleatórias contínuas;
2. $-1 \leq \kappa_{X,Y} \leq 1$, $\kappa_{X,Y} = 1$ e $\kappa_{X,-Y} = -1$;
3. $\kappa_{X,Y} = \kappa_{Y,X}$;
4. Se X e Y são independentes, então $\kappa_{X,Y} = \kappa_{\Pi} = 0$;
5. $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$;
6. Se C e $\tilde{C}$ são acoplamentos tais que $C \prec \tilde{C}$, então $\kappa_C \leq \kappa_{\tilde{C}}$;
7. Se $\{(X_n, Y_n)\}$ é uma seqüência de variáveis aleatórias contínuas com acoplamentos C_n e se $\{C_n\}$ converge pontualmente para C , então $\lim_{n \rightarrow \infty} \kappa_{\tilde{C}} = \kappa_C$. ■

3.5.2 Tau de Kendall

Definição 3.6 O tau de Kendall para o vetor aleatório $(X, Y)^T$ é definido como:

$$\tau(X, Y) = P\{(X - \tilde{X})(Y - \tilde{Y}) > 0\} - P\{(X - \tilde{X})(Y - \tilde{Y}) < 0\}, \quad (3.6)$$

onde $(\tilde{X}, \tilde{Y})^T$ e $(X, Y)^T$ são vetores aleatórios i.i.d. ■

Deste modo o tau de Kendall para $(X, Y)^T$ é simplesmente a probabilidade de concordância menos a probabilidade de discordância.

Teorema 3.4 *Seja $(X, Y)^T$ um vetor de variáveis aleatórias contínuas com função de acoplamento C . Então o tau de Kendall para $(X, Y)^T$ é dado por :*

$$\begin{aligned}\tau(X, Y) &= Q(C, C) = 4 \int \int_{[0,1]^2} C(u, v) dC(u, v) - 1, \\ &= 4 E(C(U, V)) - 1. \blacksquare\end{aligned}\quad (3.7)$$

3.5.3 Rho de Spearman

Definição 3.7 *O rho de Spearman para o vetor aleatório $(X, Y)^T$ é definido como:*

$$\rho_S(X, Y) = P\{(X - \tilde{X})(Y - Y') > 0\} - P\{(X - \tilde{X})(Y - Y') < 0\}, \quad (3.8)$$

onde $(X, Y)^T$, $(\tilde{X}, \tilde{Y})^T$ e $(X', Y')^T$ são cópias independentes. $\blacksquare$

Note que $\tilde{X}$ e Y' são independentes. Utilizando o teorema 3.3 e a primeira parte do corolário 3.1 nós temos o seguinte resultado :

Teorema 3.5 *Seja $(X, Y)^T$ um vetor de variáveis aleatórias contínuas com função de acoplamento C . Então o rho de Spearman para $(X, Y)^T$ é dado por*

$$\begin{aligned}\rho_S(X, Y) &= 3Q(C, \Pi) = 12 \int \int_{[0,1]^2} uv dC(u, v) - 3, \\ &= 12 \int \int_{[0,1]^2} C(u, v) dudv - 3. \blacksquare\end{aligned}$$

Neste caso de variáveis aleatórias contínuas, se tivermos X e Y com funções distribuição marginal F_1 e F_2 respectivamente, e fizermos $U = F_1(X)$ e $V =$

$F_2(Y)$, então :

$$\begin{aligned}\rho_S(X, Y) &= 12 \int \int_{[0,1]^2} uv dC(u, v) - 3 = 12E(UV) - 3, \\ &= \frac{E(UV) - 1/4}{1/12} = \frac{Cov(U, V)}{\sqrt{Var(U)}\sqrt{Var(V)}}, \\ &= \rho(F_1(X), F_2(Y)).\end{aligned}$$

Ou seja, o rho de Spearman do vetor aleatório $(X, Y)^T$ pode ser interpretado como o coeficiente de correlação linear de $(F_1(X), F_2(Y))$.

3.5.4 Características adicionais do tau de Kendall e rho de Spearman

A prova do seguinte teorema pode ser encontrada em Nelsen (1999), e relaciona o tau de Kendall e o rho de Spearman com a definição geral de medida de concordância dada.

Teorema 3.6 *Se X e Y são variáveis aleatórias contínuas com função de acoplamento C , então o tau de Kendall e o rho de Spearman satisfazem as propriedades da Definição 3.5 para uma medida de concordância. ■*

O teorema 3.7, cuja prova pode ser encontrada em Embrechts, McNeil e Strauman (1999b) mostra uma propriedade adicional, não citada na Definição 3.5.

Teorema 3.7 *Sejam X e Y variáveis aleatórias contínuas, com função de acoplamento C , e denotemos por κ o tau de Kendall ou o rho de Spearman. Então :*

1. $\kappa(X, Y) = 1 \Leftrightarrow C = M$;
2. $\kappa(X, Y) = -1 \Leftrightarrow C = W$. ■

O tau de Kendall e o rho de Spearman são medidas de dependência entre duas variáveis aleatórias. Para extensão para dimensões maiores, nós simplesmente escreveremos estas medidas pareadas em uma matriz $n \times n$ como fazemos com a correlação linear.

3.5.5 Vantagens da medida de concordância sobre a correlação linear

Seja κ uma medida de concordância para duas variáveis aleatórias contínuas X e Y . Como consequência da Definição 3.5, se Y é uma função quase certamente crescente de X , então $\kappa_{X,Y} = \kappa_M = 1$ e se Y é uma função quase certamente decrescente de X , então $\kappa_{X,Y} = \kappa_W = -1$. Além disto, se α e β são funções quase certamente estritamente crescentes em $Im(X)$ e $Im(Y)$ respectivamente, então $\kappa_{\alpha(X),\beta(Y)} = \kappa_{X,Y}$.

Então a medida de concordância κ resolve alguns problemas que apontamos na seção anterior para o coeficiente de correlação linear ρ :

1. Independentemente das distribuições marginais, a medida κ atinge todos valores entre -1 e 1, o que vimos que não ocorre com ρ . No caso de dependência perfeita positiva, temos $\kappa = 1$. No caso de dependência perfeita negativa, temos $\kappa = -1$.
2. A medida de concordância κ é invariante em relação às transformações crescentes não-lineares, e não somente às transformações lineares, como no caso da medida ρ .
3. Medidas de concordância como o Tau de Kendall e o Rho de Spearman não dependem da existência de variâncias como no caso do coeficiente de correlação linear.

3.6 Dependência nas caudas

O conceito de dependência nas caudas é relacionado com a quantidade de dependência na cauda do quadrante direito superior ou na cauda do quadrante esquerdo inferior de uma distribuição bivariada, como por exemplo a distribuição do vetor aleatório $(X, Y)^T$. Se as distribuições marginais destas variáveis aleatórias X e Y são contínuas, então esta medida de dependência é também uma função de seu acoplamento, e deste modo invariante sob transformações estritamente crescentes.

Definição 3.8 *Seja o vetor $(X, Y)^T$ de variáveis aleatórias contínuas, com distribuições marginais F_1 e F_2 . O coeficiente de dependência na cauda (**superior**) de $(X, Y)^T$ é definido como :*

$$\lim_{\alpha \rightarrow 1-} P\{Y > F_2^{-1}(\alpha) / X > F_1^{-1}(\alpha)\} = \lambda_U,$$

*dado que o limite $\lambda_U \in [0, 1]$ exista. Se $\lambda_U \in (0, 1]$, então X e Y são chamadas de **assintoticamente dependentes na cauda superior**. Se $\lambda_U = 0$, X e Y são chamadas de **assintoticamente independentes na cauda superior**. Caso o limite não exista, a dependência na cauda não é definida. ■*

Uma vez que $P\{Y > F_2^{-1}(\alpha) / X > F_1^{-1}(\alpha)\}$ pode ser reescrita como

$$\frac{1 - P\{X \leq F_1^{-1}(\alpha)\} - P\{Y \leq F_2^{-1}(\alpha)\} + P\{X \leq F_1^{-1}(\alpha), Y \leq F_2^{-1}(\alpha)\}}{1 - P\{X \leq F_1^{-1}(\alpha)\}},$$

uma definição equivalente para variáveis aleatórias contínuas pode ser encontrada em Joe (1997), p.33.

Definição 3.9 *Se C é um acoplamento bivariado tal que*

$$\lim_{\alpha \rightarrow 1-} \frac{(1 - 2\alpha + C(\alpha, \alpha))}{1 - \alpha} = \lambda_U$$

existe, então C tem dependência na cauda superior se $\lambda_U \in (0, 1]$, e é independente na cauda superior se $\lambda_U = 0$. ■

Para funções de acoplamento C sem uma forma fechada simples, os limites das definições 3.8 e 3.9 podem ser complicados de serem calculados. Consideremos então um vetor $(U, V)^T$ de variáveis aleatórias contínuas, com distribuições marginais uniforme padrão, cuja distribuição conjunta é C . Então podemos aplicar a regra de l'Hospital para obtermos :

$$\begin{aligned}\lambda_U &= \lim_{\alpha \rightarrow 1-} \frac{1 - 2\alpha + C(\alpha, \alpha)}{1 - \alpha}, \\ &= - \lim_{\alpha \rightarrow 1-} \left(-2 + \frac{\partial}{\partial s} C(s, t)|_{s=t=\alpha} + \frac{\partial}{\partial t} C(s, t)|_{s=t=\alpha} \right), \\ &= \lim_{\alpha \rightarrow 1-} P\{V > \alpha/U = \alpha\} + P\{U > \alpha/V = \alpha\}.\end{aligned}\tag{3.9}$$

Caso C seja uma função de acoplamento que satisfaz $C(u, v) = C(v, u)$, $\forall u, v$, então a expressão da equação 3.9 para λ_U se simplifica para

$$\lambda_U = 2 \lim_{\alpha \rightarrow 1-} P\{V > \alpha/U = \alpha\}.\tag{3.10}$$

O conceito de dependência na cauda inferior pode ser definido de maneira similar :

Definição 3.10 *Se C é um acoplamento bivariado tal que*

$$\lim_{\alpha \rightarrow 0+} \frac{C(\alpha, \alpha)}{\alpha} = \lambda_L$$

existe, então C tem dependência na cauda inferior se $\lambda_L \in (0, 1]$, e é independente na cauda inferior se $\lambda_L = 0$. ■

Da mesma maneira que anteriormente, no caso de uma função de acoplamento C sem forma fechada simples, consideraremos o vetor aleatório $(U, V)^T$ com distribuição conjunta C . Então teremos :

$$\begin{aligned}\lambda_L &= \lim_{\alpha \rightarrow 0+} \frac{C(\alpha, \alpha)}{\alpha}, \\ &= \lim_{\alpha \rightarrow 0+} (P\{V < \alpha/U = \alpha\} + P\{U < \alpha/V = \alpha\}).\end{aligned}\tag{3.11}$$

No caso de um acoplamento em que $C(u, v) = C(v, u), \forall u, v$, se resume a

$$\lambda_L = \lim_{\alpha \rightarrow 0+} 2P\{V < \alpha/U = \alpha\}. \quad (3.12)$$

A seguir apresentaremos exemplos de cálculo de λ_U pelos dois métodos apresentados.

Exemplo 3.2 *Considere o acoplamento bivariado da família de Gumbel, dado por*

$$C_\theta(u, v) = \exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}),$$

para $\theta \geq 1$. Então temos

$$\frac{1 - 2\alpha + C(\alpha, \alpha)}{1 - \alpha} = \frac{1 - 2\alpha + \alpha^{2^{1/\theta}}}{1 - \alpha},$$

e assim

$$\lambda_U = \lim_{\alpha \rightarrow 1-} \frac{(1 - 2\alpha + C(\alpha, \alpha))}{1 - \alpha} = 2 - 2^{1/\theta}.$$

Deste modo teremos dependência na cauda superior para $\theta > 1$. ■

Exemplo 3.3 *Seja $(X, Y)^T$ um vetor aleatório com distribuição padrão normal bivariada, com coeficiente de correlação linear ρ , que denotaremos por $\mathcal{N}_\rho$. Ou seja, o acoplamento C satisfaz*

$$C(u, v) = \mathcal{N}_\rho(\Phi^{-1}(u), \Phi^{-1}(v)), \quad (3.13)$$

onde Φ denota a função distribuição da normal padrão univariada.

Como o acoplamento C satisfaz a condição $C(u, v) = C(v, u), \forall u, v$, temos que :

$$\lambda_U = 2 \lim_{\alpha \rightarrow 1-} P\{V > \alpha/U = \alpha\}.$$

E como Φ é uma função distribuição com $\Phi(x) = 1$ somente para x infinito, temos :

$$\begin{aligned} \lim_{\alpha \rightarrow 1-} P\{V > \alpha/U = \alpha\} &= \lim_{\alpha \rightarrow 1-} P\{\Phi^{-1}(V) > x/\Phi^{-1}(U) = x\}, \\ &= P\{Y > x/X = x\}, \end{aligned}$$

onde $x = \Phi^{-1}(\alpha)$.

Usando o fato que $Y/X = x \sim \mathcal{N}(\rho x, 1 - \rho^2)$, nós obtemos :

$$\begin{aligned} \lambda_u &= 2 \lim_{x \rightarrow \infty} \bar{\Phi}((x - \rho x)/\sqrt{1 - \rho^2}), \\ &= 2 \lim_{x \rightarrow \infty} \bar{\Phi}(x((1 - \rho)/\sqrt{1 + \rho})) = 0 \text{ para } |\rho| \neq 1. \end{aligned}$$

De onde segue que o acoplamento C não tem dependência na cauda superior para $\rho < 1$. ■

4 *Classes de funções de acoplamento*

Neste capítulo serão apresentadas algumas classes de famílias de acoplamentos que têm sido utilizadas na análise de séries de finanças. Também serão apresentadas algumas relações com as medidas de dependência apresentadas no capítulo anterior. Estas relações poderão ser importantes na escolha da função de acoplamento adequada. Na ultima seção são apresentadas algumas aplicações dos acoplamentos encontradas na literatura.

4.1 Acoplamentos Elípticos

A classe dos acoplamentos elípticos é uma rica fonte de distribuições multivariadas que compartilham muitas das propriedades da distribuição normal multivariada e possibilita a modelagem de formas de dependência não-normal. Os acoplamentos elípticos são simplesmente os acoplamentos das distribuições elípticas, obtidos pelo corolário do Teorema de Sklar (Corolário 2.1). Gerar dados de distribuições elípticas é simples, e devido ao teorema de Sklar, temos que também é simples gerar dados de acoplamentos elípticos, como será visto no Capítulo 7.

Nesta seção definiremos as distribuições elípticas e algumas propriedades, e então consideraremos a família de acoplamentos gaussiana e t-Student.

As distribuições elípticas tem este nome porque suas curvas de contornos

são elipses.

4.1.1 Distribuições elípticas

Definição 4.1 *Seja $\mathbf{X}$ um vetor aleatório n -dimensional. Seja $\mu \in \mathbb{R}^n$ e Σ alguma matriz $n \times n$ simétrica, não-negativa definida. Se para algum μ a função característica $\varphi_{\mathbf{X}-\mu}(\mathbf{t})$ de $\mathbf{X} - \mu$ é função da forma quadrática $\mathbf{t}^T \Sigma \mathbf{t}$, ou seja, $\varphi_{\mathbf{X}-\mu}(\mathbf{t}) = \phi(\mathbf{t}^T \Sigma \mathbf{t})$, nós dizemos que $\mathbf{X}$ tem **distribuição elíptica** com parâmetros μ , Σ e ϕ e denotamos por $\mathbf{X} \sim E_n(\mu, \Sigma, \phi)$. ■*

Quando $n = 1$ a classe de distribuições elípticas coincide com a classe de distribuições simétricas unidimensionais. A função ϕ da Definição 4.1 é chamada de função geradora característica.

Teorema 4.1 *Teremos que $\mathbf{X} \sim E_n(\mu, \Sigma, \phi)$ com $\text{posto}(\Sigma) = k$ se, e somente se, existem uma variável aleatória $R \geq 0$ independente de $\mathbf{U}$, que é um vetor aleatório k -dimensional com distribuição uniforme na hiper-esfera unitária $\{\mathbf{z} \in \mathbb{R}^k | \mathbf{z}^T \mathbf{z} = 1\}$, e uma matriz $n \times k$ A , com $AA^T = \Sigma$, tal que :*

$$\mathbf{X} =_d \mu + RA\mathbf{U}. \quad \blacksquare$$

O sinal $=_d$ denota igualdade em distribuição. Para a prova do Teorema 4.1 consultar Fang, Kotz e Ng (1987).

Exemplo 4.1 *Seja $\mathbf{X} \sim \mathcal{N}_n(\mathbf{0}, \mathbf{I}_n)$. Como os componentes $X_i \sim \mathcal{N}(0, 1)$ $i = 1, \dots, n$ são independentes e a função característica de X_i é $\exp(-t_i^2/2)$, então a função característica de $\mathbf{X}$ é :*

$$\exp\left\{-\frac{1}{2}(t_1^2 + \dots + t_n^2)\right\} = \exp\left\{-\frac{1}{2}\mathbf{t}^T \mathbf{t}\right\}.$$

Assim pela Definição 4.1 segue que $\mathbf{X} \sim E_n(\mathbf{0}, \mathbf{I}_n, \phi)$, onde $\phi(u) = \exp(-u/2)$. Podemos mostrar que, aplicando o Teorema 4.1, teremos $\mathbf{X} =_d R\mathbf{U}$, onde $R \sim \sqrt{\chi_n^2}$. ■

Exemplo 4.2 Dizemos que $\mathbf{X}$ tem distribuição *t-Student* n -dimensional com média μ e matriz de covariância $\frac{\nu}{\nu-2}\Sigma$ (caso $\nu \geq 2$), se $\mathbf{X}$ tem a seguinte representação estocástica :

$$\mathbf{X} =_d \mu + \frac{\sqrt{\nu}}{\sqrt{S}} \mathbf{Z}, \quad (4.1)$$

onde $\mu \in \mathbb{R}^n$, $S \sim \chi_n^2$ e $\mathbf{Z} \sim \mathcal{N}_n(\mathbf{0}, \Sigma)$, e onde S e $\mathbf{Z}$ são independentes. ■

4.1.2 Acoplamentos gaussianos

Definição 4.2 Seja ρ uma matriz simétrica, positiva definida com $\text{diag}(\rho) = 1$ e seja Φ_ρ a distribuição normal multivariada padronizada, com matriz de correlação ρ . Utilizando o Corolário 2.1, definimos então o **acoplamento gaussiano multivariado** como :

$$C_\rho^{Ga}(u_1, u_2, \dots, u_n) = \Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_n)). \quad (4.2)$$

■

onde Φ^{-1} é a inversa da função distribuição normal padrão univariada.

Para o caso bivariado, onde R_{12} é o coeficiente de correlação linear da distribuição normal bivariada correspondente, temos :

$$C_\rho^{Ga}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1 - R_{12}^2)^{1/2}} \exp \left\{ -\frac{s^2 - 2R_{12}st + t^2}{2(1 - R_{12}^2)} \right\} ds dt.$$

No Exemplo 3.3 vimos que o acoplamento Gaussiano não tem dependência na cauda superior ($\lambda_U = 0$). Como as distribuições elípticas são simétricas radialmente, os coeficientes de dependência nas caudas superior e inferior são iguais. Assim os acoplamentos Gaussianos não tem dependência na cauda inferior ($\lambda_L = 0$).

Aplicando o acoplamento gaussiano em distribuições marginais normais univariadas (ou seja $H(x_1, \dots, x_n) = C_\rho^{Ga}(\Phi(x_1), \dots, \Phi(x_n))$), teremos como distribuição conjunta H a normal multivariada. Caso as distribuições marginais não sejam normais não teremos mais a normal multivariada como distribuição conjunta.

4.1.3 Acoplamentos t

Definição 4.3 *Seja ρ uma matriz simétrica, positiva definida com $\text{diag}(\rho) = 1$ e seja $T_{\rho, \nu}$ a distribuição t -student (como definida por (4.1)), com $\mu = 0$, ν graus de liberdade e matriz de correlação ρ (ou seja $\rho_{ij} = \Sigma_{ij} / \sqrt{\Sigma_{ii}\Sigma_{jj}}$ para $i, j \in 1, \dots, n$). Então o **acoplamento t -Student multivariado**, é definido como*

$$C_{\rho, \nu}^t(\mathbf{u}) = T_{\rho, \nu}(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), \dots, t_\nu^{-1}(u_n)),$$

onde t_ν^{-1} é a inversa da função distribuição t_ν padrão univariada. ■

Para o caso bivariado, onde R_{12} é o coeficiente de correlação linear da distribuição t_ν bivariada correspondente, temos :

$$C_{\rho, \nu}^t(u, v) = \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi(1 - R_{12}^2)^{1/2}} \left\{ 1 + \frac{s^2 - 2R_{12}st + t^2}{\nu(1 - R_{12}^2)} \right\}^{-(\nu+2)/2} ds dt.$$

Embrechts, Lindskog e McNeil (2003) mostram que no caso bivariado os coeficientes de dependência nas caudas são dados por :

$$\lambda_U = \lambda_L = 2 \left(1 - t_{\nu+1} \left(\sqrt{\nu+1} \sqrt{1 - R_{12}} / \sqrt{1 + R_{12}} \right) \right).$$

Deste modo a dependência nas caudas aumenta quando a correlação R_{12} aumenta e diminui quando o número de graus de liberdade ν aumenta. No limite, temos que $\lambda_U \rightarrow 0$ quando $\nu \rightarrow \infty$ e $R_{12} < 1$.

Aplicando o acoplamento t em distribuições marginais t univariadas com mesmo grau de liberdade ν (ou seja $H(x_1, \dots, x_n) = C_{\rho, \nu}^t(t_\nu(x_1), \dots, t_\nu(x_n))$),

teremos como distribuição conjunta H a distribuição t multivariada. Caso as distribuições marginais não sejam t ou tenham graus de liberdades diferentes, não teremos mais a distribuição t multivariada como distribuição conjunta.

4.2 Acoplamentos arquimedianos

Vimos na seção anterior que é fácil gerar dados de acoplamentos elípticos. Apesar disto, esta classe de acoplamentos possui algumas desvantagens, como não possuírem fórmulas fechadas. Em especial, no caso de séries financeiras, os acoplamentos elípticos estão restritos a terem simetria radial, o que não é adequado uma vez que os retornos financeiros tendem a ter uma dependência maior em grandes perdas do que em grandes ganhos, como visto na Seção 3.2.

Nesta seção discutiremos a importante classe dos acoplamentos Arquimedianos. Esta classe de famílias de acoplamentos tem uma grande importância por diversos motivos. Muitas famílias paramétricas de acoplamentos são arquimedianas, e a classe de acoplamentos arquimedianos permite uma grande variedade de estruturas de dependência, incluindo as existentes em séries financeiras. Além disto, os acoplamentos arquimedianos podem ser construídos com grande facilidade. Outras propriedades interessantes serão discutidas nesta seção.

Vamos apresentar a definição e propriedades da classe de acoplamentos arquimedianos para o caso bivariado. Na última subseção estenderemos o resultado para o caso multivariado.

4.2.1 Definição

Iniciaremos com uma definição geral de acoplamentos arquimedianos, que pode ser encontrada em Nelsen (1999), (pg 90).

Definição 4.4 *Seja φ uma função contínua, estritamente decrescente, de $[0, 1]$ para $[0, \infty]$, tal que $\varphi(1) = 0$. A pseudo-inversa de φ é uma função*

$\varphi^{[-1]} : [0, \infty] \rightarrow [0, 1]$, dada por :

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & \text{se } 0 \leq t \leq \varphi(0); \\ 0 & \text{se } \varphi(0) \leq t \leq \infty. \end{cases}$$

■

Note que $\varphi^{[-1]}(t)$ é contínua e não-crescente em $[0, \infty]$ e estritamente decrescente em $[0, \varphi(0)]$. Além disto, $\varphi^{[-1]}(\varphi(u)) = u$ em $[0, 1]$, e

$$\varphi(\varphi^{[-1]}(t)) = \begin{cases} t & \text{se } 0 \leq t \leq \varphi(0); \\ \varphi(0) & \text{se } \varphi(0) \leq t \leq \infty. \end{cases}$$

Finalmente, se $\varphi(0) = \infty$, então $\varphi^{[-1]} = \varphi^{-1}$.

Teorema 4.2 *Seja φ uma função contínua, estritamente decrescente, de $[0, 1]$ para $[0, \infty]$, tal que $\varphi(1) = 0$ e seja $\varphi^{[-1]}$ a pseudo-inversa de φ . Seja C uma função de $[0, 1]^2$ para $[0, 1]$, dada por :*

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)). \quad (4.3)$$

Então C é uma função de acoplamento se, e somente se, φ é convexa. ■

A prova pode ser encontrada em Nelsen (1999), (pg 91).

Funções de acoplamento da forma (4.3) são chamadas de **funções de acoplamentos arquimedianas**. A função φ é chamada de **geradora do acoplamento**. Se $\varphi(0) = \infty$, nós dizemos que φ é uma **geradora estrita**, o que implica que $\varphi^{[-1]} = \varphi^{-1}$.

Mostraremos a seguir alguns exemplos de acoplamentos arquimedianos. Para mais famílias de acoplamentos, consultar Nelsen (1999), (pg 94).

Exemplo 4.3 *Seja $\varphi(t) = -\ln t$ para $t \in [0, 1]$. Como $\varphi(0) = \infty$, φ é estrita, e então temos que $\varphi^{[-1]}(t) = \varphi^{-1}(t) = \exp(-t)$. Utilizando (4.3), temos que $C(u, v) = uv = \Pi$. Ou seja, Π é um acoplamento arquimadiano.* ■

Exemplo 4.4 Seja $\varphi(t) = 1-t$ para $t \in [0, 1]$. Então $\varphi^{[-1]}(t) = \max(1-t, 0)$ e $C(u, v) = \max(u+v-1, 0) = W(u, v)$. Ou seja, W também é um acoplamento arquimadiano. ■

Exemplo 4.5 Seja $\varphi(t) = (-\ln t)^\theta$, onde $\theta \geq 1$.

Claramente $\varphi(t)$ é contínua e $\varphi(1) = 0$. Temos também que $\varphi'(t) = -\theta(-\ln t)^{\theta-1} \frac{1}{t}$, então φ é um função estritamente decrescente de $[0, 1]$ para $[0, \infty]$. $\varphi''(t) \geq 0$ em $[0, 1]$, então φ é convexa. Além disto, $\varphi(0) = \infty$, então temos que φ é uma função geradora estrita.

Aplicando (4.3) temos que :

$$C_\theta(u, v) = \exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}).$$

Esta família é conhecida como a família de Gumbel, muito utilizada em teoria de valores extremos. Temos além disto que $C_1 = \Pi$ e $\lim_{\theta \rightarrow \infty} C_\theta = M$. ■

Exemplo 4.6 Seja $\varphi(t) = (t^{-\theta} - 1)/\theta$, onde $\theta \in [-1, \infty) \setminus \{0\}$. Esta geradora nos dá a família de acoplamentos de Clayton

$$C_\theta(u, v) = \max([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0). \quad (4.4)$$

Temos então que $C_{-1} = W$, $\lim_{\theta \rightarrow 0} C_\theta = \Pi$ e $\lim_{\theta \rightarrow \infty} C_\theta = M$. A família de Clayton tem dependência na cauda inferior para $\theta > 0$. ■

Exemplo 4.7 A família de acoplamentos de Frank é gerada por $\varphi = -\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$. Temos então

$$C_\theta(u, v) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right). \quad (4.5)$$

Temos então que $\lim_{\theta \rightarrow -\infty} C_\theta = W$, $C_0 = \Pi$, $\lim_{\theta \rightarrow \infty} C_\theta = M$. ■

4.2.2 Propriedades

O seguinte resultado, cuja prova pode ser encontrada em Embrechts, Lindskog e McNeil (2003) irá nos permitir estender os resultados para o caso multivariado.

Teorema 4.3 *Seja C um acoplamento arquimediato com gerador φ . Então :*

1. C é simétrico, isto é, $C(u, v) = C(v, u)$, $\forall u, v \in [0, 1]$;
2. C é associativo, isto é, $C(C(u, v), w) = C(u, C(v, w))$, $\forall u, v, w \in [0, 1]$.

■

Esta propriedade não é compartilhada pelas funções de acoplamento em geral.

Os resultados seguintes facilitarão o cálculo das medidas de dependência estudadas na seção (3.5) na classe de acoplamentos arquimedianos. As provas podem ser encontradas também em Embrechts, Lindskog e McNeil (2003).

Teorema 4.4 *Sejam X e Y variáveis aleatórias com acoplamento arquimediato C , gerado por φ . O tau de Kendall para X e Y é dado por :*

$$\tau_C = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt. \quad (4.6)$$

■

Calcular o tau de Kendall pela equação (3.7) em geral pode ser complicado. Desta maneira, nos acoplamentos arquimedianos esta tarefa se torna mais simples. Vamos exemplificar o uso do resultado.

Exemplo 4.8 Para a família de acoplamentos de Gumbel, sabemos que $\varphi(t) = (-\ln t^\theta)$, para $\theta \geq 1$. Então temos $\varphi(t)/\varphi'(t) = (t \ln t)/\theta$. Utilizando o Teorema 4.4, temos :

$$\tau_\theta = 1 + 4 \int_0^1 \frac{t \ln t}{\theta} dt = 1 - 1/\theta. \blacksquare$$

Para acoplamentos arquimedianos, a dependência nas caudas pode ser expressada em termos dos geradores.

Teorema 4.5 Seja φ uma função geradora estrita, tal que φ^{-1} pertence à classe de transformadas de Laplace de variáveis aleatórias estritamente positivas. Se $\varphi^{-1'}(0)$ é finita, então

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$$

não tem dependência na cauda superior. Se C tem dependência na cauda superior, então $\varphi^{-1'}(0) = -\infty$ e o coeficiente de dependência na cauda superior é dado por :

$$\lambda_U = 2 - 2 \lim_{s \rightarrow 0+} [\varphi^{-1'}(2s)/\varphi^{-1'}(s)]. \blacksquare$$

A prova do teorema pode ser encontrada em Joe (1997), (pg 103). Podemos aplicar este resultado para a família de Gumbel, e compararmos o resultado com o obtido no Exemplo 3.2.

Exemplo 4.9 Seja a família de acoplamentos Gumbel com função geradora $\varphi(t) = -(\ln t)^\theta$.

Então temos $\varphi^{-1}(s) = \exp(-s^{1/\theta})$ e $\varphi^{-1'}(s) = -s^{1/\theta-1} \exp(-s^{1/\theta})/\theta$.

Utilizando o Teorema 4.5 temos:

$$\begin{aligned} \lambda_U &= 2 - 2 \lim_{s \rightarrow 0+} [\varphi^{-1'}(2s)/\varphi^{-1'}(s)], \\ &= 2 - 2^{1/\theta} \lim_{s \rightarrow 0+} \frac{\exp(-(2s)^{1/\theta})}{\exp(-s^{1/\theta})} = 2 - 2^{1/\theta}. \end{aligned}$$

Deste modo o resultado coincide com o obtido anteriormente. $\blacksquare$

4.2.3 Acoplamentos arquimedianos multivariados

Nesta subseção iremos estender os acoplamentos arquimedianos para o caso multivariado. No exemplo 4.3 vimos que para $\varphi(t) = -\ln t$ temos que $C(u, v) = \exp(-[(-\ln u) + (-\ln v)]) = uv = \Pi$. A extensão para o acoplamento Π^n , onde $\mathbf{u} = (u_1, u_2, \dots, u_n)$ pode ser escrita como $\Pi^n(\mathbf{u}) = \exp(-[(-\ln u_1) + \dots + (-\ln u_n)]) = u_1 \dots u_n$, o que leva naturalmente à seguinte generalização de (4.3) :

$$C^n(\mathbf{u}) = \varphi^{[-1]}(\varphi(u_1) + \varphi(u_2) + \dots + \varphi(u_n)). \quad (4.7)$$

As funções $C^n(\mathbf{u})$ como definidas em (4.7) podem ser obtidas iterativamente através da composição de funções, pois para $n \geq 3$ temos que $C^n(u_1, \dots, u_n) = C(C^{n-1}(u_1, \dots, u_{n-1}, u_n))$, uma vez que os acoplamentos arquimedianos são simétricos e associativos (Teorema 4.3).

Porém considere o caso em que $\varphi(t) = 1 - t$, que gera W^n . Mas sabemos pelo Exemplo 2.2 que W^n não é uma função de acoplamento para $n \geq 3$. Assim temos que investigar quais as condições necessárias para que a função $C^n(\mathbf{u})$ como definida em (4.7) seja realmente uma função de acoplamento.

Definição 4.5 *Uma função $g(t)$ é **completamente monótona** no intervalo $[0, 1]$ se tem derivadas de todas as ordens que alternam o sinal, ou seja, se satisfaz*

$$(-1)^k \frac{d^k}{dt^k} g(t) \geq 0$$

para todo $t \in (0, 1)$ e $k = 0, 1, 2, \dots$ ■

Como consequência, se $g(t)$ é completamente monótona em $[0, \infty)$ e $g(c) = 0$ para algum $c > 0$ finito, então $g(t) = 0$ para todo $t \in [0, +\infty)$ (Widder (1941)). Deste modo, se a pseudo-inversa $\varphi^{[-1]}$ de um gerador arquimediano φ é completamente monótona, temos que $\varphi^{[-1]} > 0$ para todo $t \in [0, \infty)$ e $\varphi^{[-1]} = \varphi^{-1}$.

O seguinte teorema de Kimberling (1974) dá condições necessárias e suficientes para que a função em (4.7) seja realmente uma função de acoplamento.

Teorema 4.6 *Seja φ uma função contínua estritamente decrescente de $[0, 1]$ para $[0, \infty)$ tal que $\varphi(0) = \infty$ e $\varphi(1) = 0$, e seja φ^{-1} a inversa de φ . Se C^n é a função de $[0, 1]^n$ para $[0, 1]$ dada por (4.7), então C^n é um acoplamento n -dimensional para todo $n \geq 2$ se, e somente se, φ^{-1} é completamente monótona em $[0, \infty)$. ■*

Exemplo 4.10 *Seja $\varphi_\theta(t) = t^{-\theta} - 1$ para $\theta > 0$, o que gera uma subfamília da família Clayton bivariada, como no Exemplo 4.6, das quais as funções geradoras desta subfamília são estritas. Então $\varphi_\theta^{-1}(t) = (1+t)^{-1/\theta}$, que podemos mostrar ser completamente monótona em $[0, \infty)$. Desta maneira, podemos generalizar o família de Clayton de acoplamentos bivariados para uma família de acoplamentos n -variados para $\theta > 0$ e $n \geq 2$:*

$$C_\theta^n(\mathbf{u}) = (u_1^{-\theta} + u_2^{-\theta} + \dots + u_n^{-\theta} - n + 1)^{-1/\theta}. \quad \blacksquare$$

Este último resultado mostra que um acoplamento arquimediato gerado pela extensão de um acoplamento bivariado utilizando (4.7) deve ter dependência positiva. A prova está em Nelsen (1999), p. 122.

Corolário 4.1 *Se a inversa φ^{-1} de uma função geradora estrita φ de um acoplamento arquimediato C é completamente monótona, então $C \succ \Pi$, isto é $C(u, v) \geq uv \forall u, v \in [0, 1]$ ■*

4.3 Acoplamento de Plackett

Uma medida de "associação" ou "dependência" em uma tabela de contingência 2×2 é a chamada razão do produto cruzado, ou em inglês *odds ratio*, que denotaremos por θ . Por exemplo, seja a Tabela 4.3, onde denotamos as categorias de cada variável como "baixa" e "alta":

Variável Linha	Variável Coluna			
		Baixa	Alta	
	Baixa	a	b	a+b
	Alta	c	d	c+d
		a+c	b+d	n

Tabela 4.1: Uma tabela de contingência 2×2 .

Se a contagem observada das categorias mostradas na tabela foram a , b , c , d , então temos $\theta = (ad)/(bc)$. Caso $\theta = 1$, temos que cada entrada "observada" (como a) é igual ao seu "valor esperado" sob independência (neste caso $(a+b)(a+c)/n$), onde $n = a+b+c+d$. Caso $\theta > 1$, as observações estão concentradas nas células "baixa-baixa" e "alta-alta". Caso $0 < \theta < 1$, as observações estão concentradas nas outras duas células da tabela.

Plackett (1965) teve a idéia de associar esta tabela com distribuições marginais contínuas. Sejam X e Y variáveis aleatórias contínuas com funções distribuições marginais conjuntas F_1 e F_2 e função distribuição conjunta H . Seja (x, y) qualquer par de números e as categorias "baixa" e "alta" sejam respectivamente " $X \leq x$ " e " $X > x$ " para a variável X , e respectivamente " $Y \leq y$ " e " $Y > y$ " para a variável Y . Então trocando os números a , b , c , d por suas probabilidades de ocorrência, temos :

$$\theta(x, y) = \frac{H(x, y)[1 - F_1(x) - F_2(y) + H(x, y)]}{[F_1(x) - H(x, y)][F_2(y) - H(x, y)]}. \quad (4.8)$$

Para a maioria das distribuições $\theta(x, y)$ não é independente de (x, y) . Porém podemos encontrar distribuições conjuntas que não dependem de (x, y) . Seja $u = F_1(x)$ e $v = F_2(y)$. Utilizando o Teorema de Sklar (2.1) temos que :

$$\theta = \frac{C(u, v)[1 - u - v + C(u, v)]}{[u - C(u, v)][v - C(u, v)]}, \quad (4.9)$$

onde C é o acoplamento entre as variáveis X e Y .

Resolvendo para C , temos : $C_\theta(u, v) =$

$$\begin{cases} \frac{1}{2(\theta-1)} \left[1 + (\theta-1)(u+v) - \sqrt{[1 + (\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)} \right] & \text{for } \theta \neq 1, \\ uv, & \text{for } \theta = 1 \end{cases}$$

que é definida para $\theta > 0$, e satisfaz as condições que definem uma função de acoplamento. Assim $H(x, y) = C_\theta(F(x), G(y))$ é a função distribuição conjunta entre x e y . Chamaremos C de acoplamento de Plackett.

Para esta família de acoplamentos temos que $\lim_{\theta \rightarrow 0} C_\theta = W$, $\lim_{\theta \rightarrow \infty} C_\theta = M$.

Para esta função de acoplamento temos $\lambda_U = \lambda_L = 0$.

4.4 Algumas aplicações encontradas na literatura

Nesta seção iremos apresentar algumas aplicações encontradas na literatura sobre acoplamentos.

Cherubini e Luciano (2000) utilizam as famílias de Gumbel, Clayton e Frank com o objetivo de calcular o Valor em Risco (VaR) para índices de mercados norte-americano (S&P100) e inglês (FTSE100).

Romano (2002) utiliza os acoplamentos Gaussiano, t-Student, e os acoplamentos arquimedianos de Gumbel, Frank e Clayton para modelar os retornos de duas ações italianas (Olivetti e TIM).

Meneguzzo e Vecchiato (2002) utilizam os acoplamentos elípticos Gaussianos e t-Student e os acoplamentos arquimedianos de Clayton e Frank para modelagem do risco de derivativos de créditos.

Bouyé et al. (2003) utiliza os acoplamentos Gaussianos e t-Student para modelar o VaR de um portfolio composto por cinco ativos.

Nelsen (1999) (pg 64), cita que qualquer combinação linear convexa de

acoplamentos também é uma função de acoplamento. Por exemplo, caso C_1 e C_2 sejam funções de acoplamento, então a função de acoplamento definida por

$$C(u, v) = \lambda C_1(u, v) + (1 - \lambda)C_2(u, v), \quad (4.10)$$

onde $(u, v) \in [0, 1]^2$ e $\lambda \in (0, 1)$ é uma função de acoplamento. Fortin e Kuzmics (2002) utilizam combinações lineares convexas de acoplamentos elípticos (Gaussiana e t-Student) e acoplamentos arquimedianos (Clayton, Gumbel e Joe) para modelagem do VaR de um portfólio composto pelos índices FTSE e DAX, de modo que o acoplamento elíptico seja responsável por modelar os valores centrais da distribuição e os acoplamentos arquimedianos pelo comportamento dos retornos nas caudas.

Patton (2003a) propõe o acoplamento simetrizado de Joe-Clayton para a modelagem dos retornos das taxas de câmbio Iene-Dólar americano e Marco Alemão-Dólar americano. A razão do interesse neste acoplamento é que não impõe dependência simétrica na distribuição dos retornos (apesar disto, este acoplamento tem a dependência simétrica como um caso particular). O acoplamento de Joe-Clayton, que é arquimediato, é dado por :

$$C_{JC}(u, v | \tau_U, \tau_L) = 1 - \left(\{ [1 - (1 - u)^\kappa]^{-\gamma} + [1 - (1 - v)^\kappa]^{-\gamma} - 1 \}^{-1/\gamma} \right)^{1/\kappa} \quad (4.11)$$

onde

$$\begin{aligned} \kappa &= 1/\log_2(2 - \tau_U), \\ \gamma &= -1/\log_2(\tau_L), \\ \tau_U &\in (0, 1), \tau_L \in (0, 1). \end{aligned}$$

Este acoplamento tem dois parâmetros, τ_U e τ_L , que são exatamente os coeficientes de dependência nas caudas superior e inferior, como nas Definições 3.9 e 3.10. Este acoplamento se aproxima do limitante superior de Fréchet ($M(u, v) = \min(u, v)$) quando $\tau_U \rightarrow 1$ e $\tau_L \rightarrow 1$. O acoplamento de Joe-Clayton ainda possui uma certa assimetria quando $\tau_U = \tau_L$, o que é um inconveniente. Deste modo, é proposto o acoplamento simetrizado de Joe-Clayton,

dado por :

$$C_{SJC}(u, v|\tau_U, \tau_L) = \quad (4.12)$$

$$0,5 C_{JC}(u, v|\tau_U, \tau_L) + 0,5 C_{JC}(1 - u, 1 - v|\tau_L, \tau_U) + u + v - 1.$$

Este acoplamento é simétrico quando $\tau_U = \tau_L$.

Rockinger e Jondeau (2001) utilizam esta família de acoplamentos e uma medida de dependência similar para verificar se a dependência linear é constante no tempo, utilizando séries das bolsas européias e os índices S&P500 e Nikkei. Uma dificuldade neste tipo de acoplamento é sua generalização para mais do que duas variáveis.

5 *Estimação e Identificação de Funções de Acoplamento*

Em situações práticas, temos uma amostra de um vetor aleatório n -dimensional, digamos $\{(x_{1,t}, \dots, x_{n,t})\}_{t=1}^T$. Iremos agora discutir métodos de estimação não-paramétricos e paramétricos para as funções de acoplamento e para as distribuições marginais. Como argumentado por Patton (2003a), o principal problema com a abordagem não-paramétrica é a falta de precisão na estimação quando a dimensão da distribuição de interesse é moderadamente grande (por exemplo quatro) ou quando nós consideramos distribuições multivariadas condicionadas em um vetor. Por outro lado, no caso paramétrico temos o problema de especificação incorreta do modelo.

Além dos métodos que serão apresentados, existem outros métodos de estimação, como o método de momentos, que não serão discutidos.

Após a estimação, iremos discutir como selecionar entre as diversas famílias de acoplamentos.

5.1 *Estimação Paramétrica*

Um parâmetro de uma função de distribuição é um funcional que associa a cada função de distribuição um número real (ou um vetor de números reais).

Ou seja, se F é uma função de distribuição, $F \in \mathfrak{N}$ (onde $\mathfrak{N}$ é o espaço de funções distribuições consideradas) e $G : \mathfrak{F} \rightarrow \mathfrak{R}^p$ um funcional, temos que $\theta = G(F)$ é um *parâmetro* da função de distribuição ($\theta \in \mathfrak{R}^p$). Desta maneira, na modelagem paramétrica assumimos que F é completamente conhecida com exceção do parâmetro θ , que deve ser estimado. Ou seja, temos que $F \in \{F_\theta, \theta \in \Theta\}$, onde Θ é chamado de espaço paramétrico.

O objetivo da estimação paramétrica é estimarmos o parâmetro θ baseado em uma amostra de F .

No nosso caso, consideraremos C_θ uma função de acoplamento n -dimensional que pertence à família $\{C_\theta, \theta \in \Theta\}$, onde C_θ é completamente conhecida à exceção do parâmetro θ (que pode ser vetorial). Sejam $X_i, i = 1, \dots, n$ variáveis aleatórias com distribuição conjunta H e distribuições marginais $F_1, \dots, F_n$ respectivamente. Supondo que cada distribuição marginal dependa somente do parâmetro ϑ_i , teremos pelo teorema de Sklar que

$$H(x_1, \dots, x_n) = C(F_1(x_1; \vartheta_1), \dots, F_n(x_n; \vartheta_n); \theta). \quad (5.1)$$

Ou seja, para conhecermos completamente a distribuição conjunta H devemos estimar o vetor de parâmetros $\vartheta = (\vartheta_1, \dots, \vartheta_n, \theta)$.

Derivando (5.1) para todas as variáveis, temos a expressão da função densidade h :

$$h(x_1, x_2, \dots, x_n) = c(F_1(x_1), F_2(x_2), \dots, F_n(x_n)) \prod_{i=1}^n f_i(x_i), \quad (5.2)$$

onde f_i é a função densidade da função distribuição marginal F_i e c é a densidade da função de acoplamento, dada por :

$$c(u_1, \dots, u_n) = \frac{\partial C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n}. \quad (5.3)$$

Vamos a seguir discutir alguns métodos para estimação do vetor ϑ .

5.1.1 Método de máxima verossimilhança exata

Suponha que temos uma amostra de tamanho T de um vetor aleatório n -dimensional, $\{(x_{1,t}, \dots, x_{n,t})\}_{t=1}^T$. Seja, como denotado anteriormente, $\vartheta = (\vartheta_1, \dots, \vartheta_n, \theta)$ o vetor de parâmetros a ser estimado, onde ϑ_i é o vetor de parâmetros da distribuição marginal F_i para $i = 1, \dots, n$ e θ é o vetor de parâmetros da função de acoplamento C .

A função de log-verossimilhança é definida então por :

$$l(\vartheta) = \sum_{t=1}^T \ln c(F_1(x_{1,t}; \vartheta_1), \dots, F_n(x_{n,t}; \vartheta_n); \theta) + \sum_{t=1}^T \sum_{i=1}^n \ln f_i(x_{i,t}; \vartheta_i). \quad (5.4)$$

Deste modo o estimador de máxima verossimilhança $\hat{\vartheta}$ do vetor de parâmetros ϑ é aquele que maximiza (5.4), ou seja, que satisfaz :

$$\hat{\vartheta} = \arg \max_{\vartheta} l(\vartheta).$$

A estimação utilizando este método pode ser complicada no caso de distribuições com dimensão grande porque requer a estimação conjunta dos parâmetros das distribuições marginais e dos parâmetros da estrutura de dependência. Deste modo, este estimador não é muito utilizado na literatura sobre acoplamento, ver, por exemplo, Patton (2003a) e Bouyé et al. (2003).

5.1.2 Método IFM (Inference Function for Margins)

Neste método a estimação dos parâmetros é realizada em dois estágios. Este procedimento é computacionalmente mais simples que estimar todos os parâmetros simultaneamente como no método anterior. Uma otimização numérica com muitos parâmetros consome muito mais tempo do que diversas pequenas otimizações numéricas, cada uma com um número menor de parâmetros.

No primeiro estágio estimamos os parâmetros das distribuições marginais

F_i através do método de máxima verossimilhança, ou seja, obtendo as estimativas $\hat{\vartheta}_i$ através de :

$$\hat{\vartheta}_i = \arg \max_{\vartheta_i} \sum_{t=1}^T \ln f_i(x_{i,t}; \vartheta_i).$$

No segundo estágio utilizamos o método de máxima verossimilhança para estimarmos os parâmetros do acoplamento, onde a verossimilhança é condicionada às estimativas das marginais, através de :

$$\hat{\theta} = \arg \max_{\theta} \sum_{t=1}^T \ln c \left(F_1(x_{1,t}; \hat{\vartheta}_1), \dots, F_n(x_{n,t}; \hat{\vartheta}_n); \theta \right).$$

Em geral temos que o estimador para θ pelo método de máxima verossimilhança exata e pelo método IFM diferem. A comparação de eficiência pode se basear na comparação da matriz de variância-covariância dos estimadores de máxima verossimilhança e IFM. Outro método de comparação se baseia em simulações de Monte Carlo para comparação os dois estimadores (através de erro médio quadrático por exemplo). Em uma tese não publicada, Xu (1996) realiza diversas comparações, obtendo eficiências relativas (razão entre os erros médios quadráticos dos estimadores IFM e de máxima verossimilhança) próximas de um.

Sob certas condições de regularidade nas log-verossimilhanças, nas funções de escore e nas matrizes hessianas, Patton (2003b) mostra que o estimador de máxima verossimilhança em dois estágios (método IFM) é assintoticamente consistente e Normal.

O IFM é o método citado no livro de Joe (1997) e utilizado em Patton (2003a) e Rockinger e Jondeau (2001). Nestes trabalhos não houve problemas de convergência nas estimativas de máxima verossimilhança.

5.1.3 Método de Máxima Verossimilhança Canônica

Este método surge da idéia de utilizar a distribuição empírica para estimar as distribuições marginais, em uma abordagem não-paramétrica. Após isto, na estimação do parâmetro θ da função de acoplamento é utilizada uma abordagem paramétrica. A diferença em relação ao método anterior (IFM) é que nenhuma suposição é feita sobre a forma paramétrica das distribuições marginais.

Antes de apresentarmos o método, vamos definir a função de distribuição empírica.

Definição 5.1 *Seja $Y_1, \dots, Y_n$ uma amostra aleatória da variável Y , que tem distribuição F . A função de distribuição empírica de Y é dada por :*

$$F_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(Y_i \in (-\infty, y])}. \quad (5.5)$$

■

O método tem dois estágios. No primeiro estágio, transformamos o conjunto de dados $(x_{1,t}, \dots, x_{n,t})$, $t = 1, \dots, T$ em variáveis uniformes $(\hat{u}_{1,t}, \dots, \hat{u}_{n,t})$ utilizando as distribuições empíricas. Pela teoria de probabilidades sabemos que $U = F_X^{-1}(X) \sim U(0, 1)$. Na versão amostral temos $u_t = F_n^{-1}(x_t)$, $t = 1, \dots, T$ e F_n é a distribuição empírica de X . Assim aplicamos este método para as n variáveis consideradas.

No segundo estágio, estimamos o parâmetro de dependência θ da seguinte maneira:

$$\hat{\theta} = \arg \max_{\theta} \sum_{t=1}^T \ln c \left((\hat{u}_{1,t}, \dots, \hat{u}_{n,t}; \hat{\vartheta}_n); \theta \right).$$

Este método também é consistente sob condições de regularidade.

Romano (2002) utiliza o método de máxima verossimilhança canônica na

estimação de uma função de acoplamento bivariada entre dois ativos do mercado italiano.

5.1.4 Estimação de parâmetros através de medidas de dependência

Este método é utilizado apenas para estimação de parâmetros de funções de acoplamento bivariadas uniparamétricas, quando existe uma relação entre seu parâmetro e sua medida de dependência. No caso de acoplamentos arquimedianos, temos pelo Teorema 4.4 a relação entre o tau de Kendall e o parâmetro θ da função de acoplamento.

Por exemplo, na família de Acoplamentos de Gumbel, temos que $\tau_\theta = 1 - 1/\theta$, o que implica que $\theta = 1/(1 - \tau_\theta)$.

Podemos estimar então o Tau de Kendall através de :

$$\hat{\tau} = \left(\frac{T}{2} \right)^{-1} \sum_{i < j} \text{sign} [(x_{1,i} - x_{1,j})(x_{2,i} - x_{2,j})]. \quad (5.6)$$

Utilizando $\hat{\tau}$ podemos estimar θ por $\hat{\theta} = 1/(1 - \hat{\tau})$.

Como no caso de séries financeiras o tamanho da amostra T geralmente é grande, a perda causada pela discretização da estimativa é bem pequena.

5.1.5 Caso condicional

Utilizando a versão condicional do Teorema de Sklar (Teorema 2.5), podemos escrever a função de verossimilhança para o caso condicional. Seja F_i a distribuição condicional de $X_i/\mathbf{W}$ para $i = 1, \dots, n$ e seja H a distribuição condicional conjunta (absolutamente contínua) de $\mathbf{X}/\mathbf{W}$, onde $\mathbf{X} = (X_1, \dots, X_n)$ tem função de acoplamento condicional C . Então temos :

$$H(x_1, \dots, x_n/\mathbf{w}) = C(F_1(x_1/\mathbf{w}), \dots, F_n(x_n/\mathbf{w})).$$

Deste modo, uma expressão equivalente à (5.2) é dada por :

$$c(F_1(x_1/\mathbf{w}), \dots, F_n(x_n/\mathbf{w})/\mathbf{w}) \prod_{i=1}^n f_i(x_i/\mathbf{w}), \quad (5.7)$$

onde $f_i(x_i/\mathbf{w})$ é a densidade condicional de $X_i/\mathbf{W} = \mathbf{w}$ e

$$c(u_1, \dots, u_n/\mathbf{w}) = \frac{\partial C(u_1, \dots, u_n/\mathbf{w})}{\partial u_1 \dots \partial u_n}. \quad (5.8)$$

Para o caso de uma amostra de tamanho T de um vetor aleatório n -dimensional, $\{(x_{1,t}, \dots, x_{n,t})\}_{t=1}^T$, temos a expressão equivalente à (5.4) :

$$\begin{aligned} l(\vartheta) = \sum_{t=1}^T \ln c(F_1(x_{1,t}; \vartheta_1/\mathbf{w}_t), \dots, F_n(x_{n,t}; \vartheta_n/\mathbf{w}_t); \theta/\mathbf{w}_t) \\ + \sum_{t=1}^T \sum_{i=1}^n \ln f_i(x_{i,t}; \vartheta_i/\mathbf{w}_t). \end{aligned} \quad (5.9)$$

Assim podemos utilizar os métodos citados anteriormente, como máxima verossimilhança exata, IFM ou máxima verossimilhança canônica para estimarmos os parâmetros do modelo.

No caso do método IFM podemos obter as estimativas $\hat{\vartheta}_i$ através de

$$\hat{\vartheta}_i = \arg \max_{\vartheta_i} \sum_{t=1}^T \ln f_i(x_{i,t}; \vartheta_i/\mathbf{w}_t), \quad (5.10)$$

e então estimamos o parâmetro de dependência θ por

$$\hat{\theta} = \arg \max_{\theta} \sum_{t=1}^T \ln c(F_1(x_{1,t}; \hat{\vartheta}_1/\mathbf{w}_t), \dots, F_n(x_{n,t}; \hat{\vartheta}_n/\mathbf{w}_t); \theta/\mathbf{w}_t). \quad (5.11)$$

O método IFM também pode ser pensado da maneira a seguir. No primeiro

passo, fazemos as transformações :

$$u_{i,t} = F_t(x_{i,t}; \hat{\vartheta}_i) / \mathbf{w}_t, \quad (5.12)$$

para $t = 1, \dots, T$ e para $i = 1, \dots, n$ e onde $\hat{\vartheta}_i$ é calculado por (5.10).

Diebold, Gunther e Tay (1998) mostram que se o modelo especificado está correto, as séries transformadas $\{u_{1,t}\}$, $\{u_{2,t}\}$, ..., $\{u_{n,t}\}$ serão amostras i.i.d. de distribuições Uniforme(0,1). Ou seja, as transformações em (5.12) eliminam a dependência temporal em cada série, embora a dependência entre uma série e outra continue existindo (que é justamente a dependência que desejamos modelar através da função de acoplamento).

No segundo estágio, a Equação 5.11 se torna :

$$\hat{\theta} = \arg \max_{\theta} \sum_{t=1}^T \ln c(u_{1,t}, \dots, u_{n,t}; \theta). \quad (5.13)$$

Exemplo 5.1 *Rockinger e Jondeau (2001) modelam o parâmetro de dependência θ_t no caso bivariado de acordo com a posição do vetor de retornos passados (uniformizados) (u_1, u_2) em $[0, 1]^2$. então para o tempo t que o vetor de condicionamento é $\mathbf{w}_t = (u_{1,t-1}, u_{2,t-1})$. Eles dividem a região $[0, 1]^2$ em 16 subregiões, onde o parâmetro θ_t será constante em cada uma destas subregiões. Mais precisamente, o modelo básico considerado é :*

$$\ln(\theta_t) = \sum_{j=1}^{16} d_j \mathbf{I}[(u_{1,t-1}, u_{2,t-1}) \in \mathcal{A}_j], \quad (5.14)$$

onde $\mathcal{A}_j$ é a j -ésima subregião de $[0, 1]^2$. Para cada subregião $\mathcal{A}_j$ temos um parâmetro d_j associado, como visto na Figura 5.1.

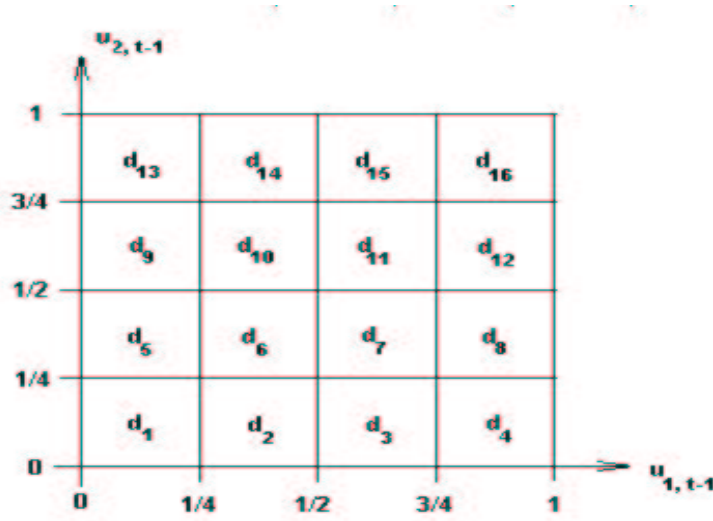


Figura 5.1: Parâmetros das 16 regiões para θ_t .

A escolha de 16 subregiões é puramente arbitrária. Esta parametrização nos permite testar diversas hipóteses sobre o impacto dos retornos passados na dependência subsequente. ■

5.2 Estimação Não-Paramétrica

5.2.1 Acoplamentos empíricos

Definição 5.2 Seja uma amostra $\{(x_{1,t}, \dots, x_{n,t})\}_{t=1}^T = \mathbf{x}_t$. O **acoplamento empírico** é dado por :

$$\hat{C} \left(\frac{t_1}{T}, \frac{t_2}{T}, \dots, \frac{t_n}{T} \right) = \frac{1}{T} \sum_{t=1}^T \mathbf{1}_{[x_{1,t} \leq x_{1(t_1)}, x_{2,t} \leq x_{2(t_2)}, \dots, x_{n,t} \leq x_{n(t_n)}]}, \quad (5.15)$$

onde $x_i(t_j)$, $i = 1, \dots, n$, $j = 1, \dots, n$ são estatísticas de ordem e $1 \leq t_1, \dots, t_n \leq T$. ■

Ou seja, o acoplamento empírico é a proporção de elementos da amostra

que satisfazem $x_{1,t} \leq x_{1,(t_1)}, x_{2,t} \leq x_{2,(t_2)}, \dots, x_{n,t} \leq x_{n,(t_n)}$.

Definição 5.3 *Seja uma amostra como na definição anterior. A **freqüência de acoplamento empírico** é dada por*

$$\hat{c}\left(\frac{t_1}{T}, \frac{t_2}{T}, \dots, \frac{t_n}{T}\right) = \begin{cases} \frac{1}{T} & \text{se } (x_{1,(t_1)}, x_{2,(t_2)}, \dots, x_{n,(t_n)}) \text{ pertence à amostra} \\ 0 & \text{caso contrário.} \end{cases} \quad (5.16)$$

■

As relações entre $\hat{C}$ e $\hat{c}$ são dadas por :

$$\hat{C}\left(\frac{t_1}{T}, \frac{t_2}{T}, \dots, \frac{t_n}{T}\right) = \sum_{i_1=1}^{t_1} \dots \sum_{i_n=1}^{t_n} \hat{c}\left(\frac{i_1}{T}, \frac{i_2}{T}, \dots, \frac{i_n}{T}\right) \quad (5.17)$$

e

$$\hat{c}\left(\frac{t_1}{T}, \frac{t_2}{T}, \dots, \frac{t_n}{T}\right) = \sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{i_1+\dots+i_n} \hat{C}\left(\frac{t_1-i_1+1}{T}, \frac{t_2-i_2+1}{T}, \dots, \frac{t_n-i_n+1}{T}\right) \quad (5.18)$$

Os acoplamentos empíricos podem ser utilizados para estimarmos medidas de dependência. Vamos considerar o caso bivariado (n=2).

Sejam $\hat{C}$ e $\hat{c}$, respectivamente o acoplamento empírico e a freqüência de acoplamento empírico de uma amostra $\{(x_{1,t}, x_{2,t})\}_{t=1}^T$. Então temos as seguintes estimativas para o rho de Spearman e o tau de Kendall, respectivamente

$$\hat{\rho}_s = \frac{12}{T^2 - 1} \sum_{t_1=1}^T \sum_{t_2=1}^T \left(\hat{C}\left(\frac{t_1}{T}, \frac{t_2}{T}\right) - \frac{t_1 t_2}{T^2} \right) \quad (5.19)$$

$$\hat{\tau} = \frac{2T}{T-1} \sum_{t_1=2}^T \sum_{t_2=2}^T \sum_{p=1}^{t_1-1} \sum_{q=1}^{t_2-1} \left[\hat{c}\left(\frac{t_1}{T}, \frac{t_2}{T}\right) \hat{c}\left(\frac{p}{T}, \frac{q}{T}\right) - \hat{c}\left(\frac{t_1}{T}, \frac{q}{T}\right) \hat{c}\left(\frac{p}{T}, \frac{t_2}{T}\right) \right] \quad (5.20)$$

Nelsen (1999), (pg 177) mostra que estas estimativas são equivalentes às expressões para as estimativas de ρ_s e τ encontradas usualmente na literatura.

5.3 Seleção da função de acoplamento

Nesta seção iremos discutir como selecionar a melhor família de acoplamentos a partir de uma determinada amostra.

5.3.1 Acoplamentos Arquimedianos

No caso da classe de Acoplamentos Arquimedianos, a abordagem seguinte, proposta por Genest e Rivest (1993), permite a seleção da função de acoplamento que melhor se ajusta aos dados reais. Esta abordagem é baseada no seguinte teorema, cuja prova pode ser encontrada em Nelsen (1999), p. 101.

Teorema 5.1 *Seja C um acoplamento Arquimadiano gerado por φ e seja*

$$K_c(t) = V_C(\{(u, v) \in [0, 1]^2 \mid C(u, v) \leq t\}), \quad (5.21)$$

(onde V_C é calculado como no item (3) da Definição 2.1). Então para qualquer $t \in [0, 1]$,

$$K_C(t) = t - \frac{\varphi(t)}{\varphi'(t+)}.$$

■

$K_c(t)$ pode ser interpretada como a medida abaixo da curva de nível $C(u, v) = t$. Uma estimação não-paramétrica de (5.21) é dada por:

$$\hat{K}_C(t) = \frac{1}{T} \sum_{i=1}^T \mathbf{1}_{[\vartheta_i \leq t]}, \quad (5.23)$$

onde $\vartheta_i = \frac{1}{T-1} \sum_{j=1}^T \mathbf{1}_{[x_{1,j} \leq x_{1,i}; x_{2,j} \leq x_{2,i}; \dots; x_{n,j} \leq x_{n,i}]}$, para $i=1, \dots, T$.

Neste ponto, nós podemos escolher diversas representações paramétricas dos acoplamentos através de algumas funções geradoras φ_i . Podemos então comparar a estimativa não-paramétrica $\hat{K}$ e as funções K teóricas de diversas famílias arquimedianas, e escolhermos a que melhor se ajusta à estimativa $\hat{K}$. (a estimação do parâmetro θ do acoplamento arquimediano foi discutida na Seção 5.1). Esta comparação pode ser feita através de um gráfico Quantil-Quantil.

Uma outra sugestão é escolher o acoplamento que minimiza a distância baseada na norma L^2 entre (5.22) e (5.23). Esta distância é dada por :

$$d_2(\hat{K}, K) = \int_0^1 \left[K(z) - \hat{K}(z) \right]^2 dz. \quad (5.24)$$

5.3.2 Acoplamento Empírico

Seja $\{C_k\}_{1 \leq k \leq K}$ o conjunto de acoplamentos disponíveis. A escolha das famílias de acoplamentos a serem consideradas está relacionada com a estrutura de dependência dos dados a serem modelados e também com os objetivos da modelagem. Pelo método de seleção através do acoplamento empírico, utilizado por Romano (2002), escolheremos a função de acoplamento C_k que minimiza a seguinte distância entre C_k e o acoplamento empírico definido em (5.15) :

$$\bar{d}_n(\hat{C}, C_k) = \left[\sum_{t_1=1}^T \dots \sum_{t_n=1}^T \left(\hat{C}\left(\frac{t_1}{T}, \dots, \frac{t_n}{T}\right) - C_k\left(\frac{t_1}{T}, \dots, \frac{t_n}{T}\right) \right)^2 \right]^{1/2}. \quad (5.25)$$

Esta distância também pode ser utilizada para estimar o vetor de parâmetros $\vartheta \in \Theta$ de um acoplamento dado $C(\mathbf{u}, \vartheta)$ da seguinte maneira :

$$\hat{\vartheta} = \arg \min_{\vartheta \in \Theta} \left[\sum_{\mathbf{u}} \left(\hat{C}(\mathbf{u}) - C(\mathbf{u}; \vartheta) \right)^2 \right]^{1/2}. \quad (5.26)$$

5.3.3 Critério de Informação de Akaike (AIC)

Suponha que um modelo estatístico de Q parâmetros seja ajustado aos dados. Para assegurar a qualidade do ajuste do modelo, Akaike (1973) introduziu um critério de informação. O critério é chamado na literatura de AIC (*Akaike's information criterion*), e é definido como :

$$AIC(Q) = -2\ln(\text{máxima verossimilhança}) + 2Q. \quad (5.27)$$

Podemos utilizar este critério para a escolha do melhor acoplamento, ou seja, escolher a função de acoplamento estimada que nos dá o mínimo AIC, onde a função de verossimilhança é dada por (5.4) ou (5.9).

Smith (2003) utiliza este critério para seleção da função de acoplamento.

6 *Estimação das Distribuições Marginais*

Seja uma série de preços de um ativo $\{P_t\}_{t=1}^T$. Então seja $X_t = \ln(P_t) - \ln(P_{t-1})$, para $t = 0, \dots, n$. Esta série $\{X_t\}_{t=1}^T$ é a chamada série de log-retornos, ou retornos compostos continuamente, que chamaremos nesta dissertação apenas de retornos. No nosso caso, temos 2 ou mais ativos dos quais queremos estimar as distribuições marginais.

Iremos discutir brevemente os fatos estilizados normalmente existentes em dados de retornos financeiros, e em seguida iremos apresentar os modelos da família GARCH.

A utilização da distribuição empírica para as marginais, que é um método não-paramétrico, já foi discutida na Subseção 5.1.3.

6.1 Fatos estilizados de retornos financeiros

Diversos fatos estilizados sobre os retornos financeiros são conhecidos, entre eles temos :

1. Retornos - Os retornos de séries financeiras apresentam, em geral, uma distribuição simétrica com média próxima de zero.
2. Autocorrelação - as séries dos retornos são, via de regra, não correlacionadas. Entretanto, as séries dos quadrados dos retornos possuem uma

pequena, mas significativa, autocorrelação. Este fato sugere que há não linearidade no processo gerador dos retornos, e irá nos motivar a modelar o quadrado dos retornos na Seção .

3. Caudas pesadas - Os retornos destas séries apresentam distribuições com caudas, em geral, mais pesadas do que a da distribuição normal, assemelhando-se mais a distribuição t de Student.
4. Agrupamentos de volatilidade - Os retornos apresentam períodos de alta e outros de baixa volatilidade, ou seja, períodos de alta volatilidade tendem a continuar com alta volatilidade, o mesmo ocorrendo para os períodos de baixa volatilidade.

Desejamos que os modelos especificados sejam capazes de modelar estes fatos observados empiricamente.

6.2 Modelos de volatilidade

Os dois modelos paramétricos mais utilizados para modelar a volatilidade são os modelos da família GARCH (ver Bollerslev, Engle e Nelson (1994)) e o modelo de volatilidade condicional (ver Ghysels, Harvey e Renault (1996) e Shephard (1996)). Nesta dissertação iremos considerar apenas os modelos da família GARCH.

A série de retornos X_t é modelada como um nível μ_t mais uma perturbação ϵ_t . É considerada uma relação aditiva entre os componentes, e que a perturbação é o produto de dois processos independentes. De maneira formal, temos :

$$\begin{aligned} X_t &= \mu_t + \epsilon_t, \\ \epsilon_t &= \sigma_t \eta_t. \end{aligned} \tag{6.1}$$

As suposições são de que :

- as perturbações ϵ_t e o nível μ_t são processos independentes;
- os processos $\{\eta_t\}$ e $\{\sigma_t\}$ são independentes;
- $\{\eta_t\}$ é um processo ruído branco com média zero e variância 1.

Notamos que a definição do modelo implica que $E[\epsilon_t] = 0$ e $Cov(\epsilon_t, \epsilon_{t-k}) = 0$ para $k > 0$. O fato de $\{\epsilon_t\}$ ser não-correlacionado está de acordo com o fato estilizado (2).

No caso mais simples temos $\mu_t = c$ (constante).

A modelagem para o componente ϵ_t será discutida na subseção a seguir.

6.2.1 Modelo GARCH(p,q)

GARCH é a sigla para a expressão em inglês *Generalized Autoregressive Conditional Heteroskedasticity*. O modelo é dado por :

$$\begin{aligned}\epsilon_t &= \sigma_t \eta_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2,\end{aligned}\tag{6.2}$$

onde $\{\eta_t\}$ é um processo ruído branco com média zero e variância 1. Em geral supõe-se que $\eta_t \sim \mathcal{N}(0, 1)$ ou $\eta_t \sim t_\nu$ padronizada. Bollerslev (1986) mostra que impondo as restrições $\alpha_0 > 0$, $\alpha_i \geq 0$ para $i = 1, \dots, p$, $\beta_j \geq 0$ para $j = 1, \dots, q$, $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$ é suficiente para que a variância condicional seja positiva. Na prática os modelos GARCH são estimados sem estas restrições. Nelson e Cao (1992) relaxam estas restrições e dão condições suficientes para que a volatilidade condicional nos modelos GARCH(1,q) e GARCH(2,q) seja positiva.

Para este modelo, podemos provar que a variância não-condicional é dada

por :

$$Var(\epsilon_t) = \frac{\alpha_0}{1 - \left(\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j \right)}.$$

Esta variância é finita se $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$. Podemos provar também que a curtose é maior do que 3, que é a curtose do caso normal. Isto reproduz o fato estilizado (3). Também é possível mostrar que o modelo reproduz o fato estilizado (4), que é a existência de aglomerados de volatilidade.

Por último, após algumas contas, temos também que :

$$\epsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\epsilon_{t-1}^2 + \dots + (\alpha_p + \beta_p)\epsilon_{t-p}^2 + \omega_t - \beta_1\omega_{t-1} - \dots - \beta_q\omega_{t-q} \quad (6.3)$$

onde $\omega_t = \epsilon_t^2 - \sigma_t^2$, que pode ser provado ser um processo ruído branco.

Assim temos que $\epsilon_t^2 \sim ARMA(p, q)$. No caso do modelo GARCH(1,1), a autocorrelação deste processo ARMA(1,1) tem decaimento segundo o valor $\alpha_1 + \beta_1$, que é chamado de *persistência* do processo.

Exemplo 6.1 *Para exemplificar a metodologia de estimação das funções de acoplamento no caso dos modelos de volatilidade, vamos considerar o caso bivariado*

$\mathbf{X} = (X_1, X_2)$, onde temos uma amostra $\{(X_{1,t}, X_{2,t})\}_{t=1}^T$.

O condicionamento se dá nas realizações passadas das variáveis, ou seja $\mathbf{W}_{t-1} = \mathcal{F}_{t-1} = (X_{1,t-1}, X_{2,t-1}, X_{1,t-2}, X_{2,t-2}, \dots)$. Seja C_θ a função de acoplamento condicional.

Considerando no exemplo para cada marginal modelos GARCH(1,1) com níveis μ_t constantes, temos para X_1 :

$$\begin{aligned} X_{1,t} &= c_1 + \epsilon_{1,t}, \\ \epsilon_{1,t} &= \sigma_{1,t}\eta_{1,t}, \\ \sigma_{1,t}^2 &= \alpha_1 + \beta_1\epsilon_{1,t-1}^2 + \gamma_1\sigma_{1,t-1}^2, \end{aligned}$$

e para X_2 :

$$\begin{aligned} X_{2,t} &= c_2 + \epsilon_{2,t}, \\ \epsilon_{2,t} &= \sigma_{2,t} \eta_{2,t}, \\ \sigma_{2,t}^2 &= \alpha_2 + \beta_2 \epsilon_{2,t-1}^2 + \gamma_2 \sigma_{2,t-1}^2, \end{aligned}$$

onde $\{\eta_{1,t}\}$ e $\{\eta_{2,t}\}$ são processos ruído branco com distribuição normal padrão.

Podemos então estimar os parâmetros do modelo GARCH para as marginais, ou seja, os parâmetros $\alpha_{1,0}, \alpha_{1,1}, \beta_{1,1}, \alpha_{2,0}, \alpha_{2,1}, \beta_{2,1}$ e então estimarmos o parâmetro θ da função de acoplamento usando o método IFM ou método de máxima verossimilhança canônica. ■

Este é apenas um simples exemplo de modelagem das distribuições marginais. Modelos mais complexos podem ser utilizados. Um exemplo seria um modelo GARCH multivariado em que a distribuição da variável X_i depende das realizações passadas das "outras variáveis", como por exemplo o modelo BEKK proposto por Engle e Kroner (1995), que será utilizado na Seção 8.6.

7 *Geração de vetores pseudo aleatórios*

Uma das utilizações mais importantes das funções de acoplamento é na geração de vetores pseudo aleatórios, que pode ser utilizada para compreendermos e testarmos modelos de sistemas do mundo real, como por exemplo modelos de Valor em Risco. A simulação também pode ser utilizada para estudos de propriedades estatísticas de estimadores, como a comparação entre diversos métodos, estudos de robustês ou da concordância de valores assintóticos com valores para pequenas amostras.

Iremos apresentar um método geral de geração de vetores pseudo aleatórios e após isto iremos discutir alguns algoritmos específicos para os acoplamentos Gaussianos, t-Student, arquimedianos e de Plackett .

7.1 Método geral

Iremos iniciar com o método bivariado de simulação e após isto iremos estender o algoritmo para o caso n -dimensional.

Sabemos dos resultados básicos de probabilidade que um algoritmo simples para gerar uma observação x de uma variável aleatória univariada X , com função distribuição F , é dado por :

1. Gere uma observação u de uma variável aleatória U , cuja distribuição é

Uniforme(0,1);

2. Faça $x = F^{(-1)}(u)$, onde $F^{(-1)}$ é a inversa generalizada de F .

Agora iremos apresentar um algoritmo pra gerar observações (x, y) de um vetor aleatório (X, Y) , cuja distribuição conjunta é H e as distribuições marginais de X e Y são F_1 e F_2 , respectivamente. Neste algoritmo, utilizaremos as funções de acoplamento como ferramenta. Devido ao teorema de Sklar, basta gerarmos um par de observações (u, v) do vetor aleatório (U, V) , onde U e V têm distribuição Uniforme(0,1) e (U, V) tem distribuição conjunta C . Um procedimento para a geração deste par (u, v) é o método de distribuição condicional. Para este método, nós precisaremos da distribuição condicional de V dado $U = u$, que denotaremos por $c_u(v)$:

$$\begin{aligned} c_u(v) &= P(V \leq v/U = u) = F_{V/U}(v/u) \\ &= \lim_{\Delta u \rightarrow 0} \frac{C(u + \Delta u, v) - C(u, v)}{\Delta u} = \frac{\partial}{\partial u} C(u, v). \end{aligned} \quad (7.1)$$

Desta maneira temos o seguinte algoritmo para geração do par desejado (x, y) :

1. Gere duas variáveis Uniforme(0,1) independentes, u e t ;
2. Considere $v = c_u^{(-1)}(t)$, onde $c_u^{(-1)}$ é a inversa generalizada de c_u ;
3. Como $c_u(t) = F_{V/U}(t/u)$, e t é uma observação de uma variável aleatória uniforme padrão, então $v = c_u^{(-1)}(t)$ é uma observação da variável aleatória $V/U = u$. Temos então que o par (u, v) é uma observação do vetor aleatório (U, V) que tem distribuição conjunta C ;
4. Podemos transformar então o vetor uniforme (u, v) no vetor desejado (x, y) , utilizando as inversas $x = F_1^{(-1)}(u)$ e $y = F_2^{(-1)}(v)$.

Este algoritmo pode ser generalizado caso desejemos gerar uma observação $(x_1, \dots, x_n)$ do vetor $(X_1, \dots, X_n)$, cuja distribuição conjunta é H , cuja função

de acoplamento é C e as distribuições marginais de $X_1, \dots, X_n$ são $F_1, \dots, F_n$ respectivamente.

1. Gere n variáveis Uniforme(0,1) independentes, $(v_1, \dots, v_n)$;
2. Considere $u_1 = v_1$;
3. Seja

$$\begin{aligned} C_{m \setminus 1, \dots, m-1}(u_1, \dots, u_m) &= P(U_m \leq u_m \mid (U_1, \dots, U_{m-1}) = (u_1, \dots, u_{m-1})) \\ &= \frac{\partial_{(u_1, \dots, u_{m-1})}^{m-1} C(u_1, \dots, u_m, 1, \dots, 1)}{\partial_{(u_1, \dots, u_{m-1})}^{m-1} C(u_1, \dots, u_{m-1}, 1, \dots, 1)}. \end{aligned} \quad (7.2)$$

Então calcule recursivamente $u_m = C_{m \setminus 1, \dots, m-1}^{-1}(u_1, \dots, u_{m-1}, v_m)$ para $m = 2, \dots, n$. Deste modo teremos que o vetor $(u_1, \dots, u_n)$ gerado é uma observação do vetor $(U_1, \dots, U_n)$, que tem distribuição conjunta C ;

4. Transforme o vetor uniforme $(u_1, \dots, u_n)$ no vetor desejado $(x_1, \dots, x_n)$, utilizando as inversas $x_i = F_i^{(-1)}(u_i)$ para $i = 1, \dots, n$.

Exemplo 7.1 *Seja a função de acoplamento do Exemplo 2.1. Temos então*

$$C(u, v) = \frac{uv}{u + v - uv}, \quad (7.3)$$

e deste modo $c_u(v) = \frac{\partial}{\partial u} C(u, v) = \left(\frac{v}{u+v-uv}\right)^2$ e $c_u^{-1}(t) = \frac{u\sqrt{t}}{1-(1-u)\sqrt{t}}$. As marginais F_1 e F_2 têm inversas $F_1^{-1}(u) = 2u - 1$ e $F_2^{-1}(v) = -\ln(1 - v)$ respectivamente.

Temos então o algoritmo de simulação :

1. Gere duas variáveis u e t uniforme(0,1) independentes;
2. Calcule $v = \frac{u\sqrt{t}}{1-(1-u)\sqrt{t}}$;
3. Faça $x = 2u - 1$ e $y = -\ln(1 - v)$;

4. O par desejado (x,y) é uma observação do vetor (X,Y) que tem distribuição conjunta H . ■

O algoritmo é computacionalmente intensivo para valores grandes de n . De fato, pode ser muito complicado o cálculo da distribuição condicional em (7.2). Para certas funções de acoplamento existem algoritmos mais simples, que serão apresentados nas próximas seções.

7.2 Método de rejeição

Podem existir dificuldades na implementação do algoritmo da seção anterior caso a inversa da derivada da função de acoplamento $c_u^{-1}(v)$ seja difícil de ser calculada. Neste caso, o algoritmo a seguir, baseado no método da rejeição (consulte Gamerman (1996) (pg. 16-19)), pode ser utilizado para gerar observações da função de acoplamento C . Apresentaremos o algoritmo apenas para o caso bivariado, mas a extensão para maiores dimensões é semelhante. Seja c a função de densidade do acoplamento C , como definido em (5.3).

1. Gere três variáveis u , v e w Uniforme(0,1) independentes;
2. Aceitar os valores gerados se $w \leq c(u, v)$. Caso contrário retornar ao passo 1;
3. O vetor (u, v) é uma observação gerada da função de acoplamento C .

Prova: Queremos provar que $f(u, v/\{w \leq c(u, v)\}) = c(u, v)$. Pelo Teorema de Bayes, temos que :

$$f(u, v/\{w \leq c(u, v)\}) = \frac{P(w \leq c(u, v)/(u, v))f_{U,V}(u, v)}{\int_0^1 \int_0^1 P(w \leq c(u, v)/(u, v))f_{U,V}(u, v)dvdu}$$

Como w é independente de (u,v) e Uniforme(0,1), temos que $P(w \leq c(u, v)/(u, v)) = c(u, v)$. Como u e v são variáveis aleatórias uniforme padrão independentes,

então na região $[0, 1]^2$ temos que $f_{U,V}(u, v) = 1$ (antes da aceitação/rejeição). Deste modo, teremos que :

$$f(u, v/\{w \leq c(u, v)\}) = \frac{c(u, v)}{\int_0^1 \int_0^1 c(u, v) dv du} = c(u, v). \quad \square$$

7.3 Acoplamentos Gaussianos

Considere a função de acoplamento gaussiano, denotada por C_R^{Ga} , como definida na Subseção 4.1.2. Temos que R tem dimensão $n \times n$ positiva definida. Neste caso, existe uma matriz A também $n \times n$ tal que $R = AA^T$. Também é assumido que as variáveis aleatórias $Z_1, \dots, Z_n$ são normais padrão independentes. Seja $\mathbf{Z} = (Z_1, \dots, Z_n)$. Então temos que

$$\mu + AZ \sim \mathcal{N}_n(\mu, R).$$

A matriz A pode ser encontrada através da decomposição de Cholesky da matriz R . Temos então o seguinte algoritmo :

1. Encontre a decomposição de Cholesky A da matriz R ;
2. Simule n variáveis normal padrão independentes $\mathbf{z} = (z_1, \dots, z_n)^T$;
3. Faça $\mathbf{x} = A\mathbf{z}$;
4. Determine os componentes $u_i = \phi(x_i)$, onde ϕ é a função distribuição acumulada da Normal padrão, para $i = 1, \dots, n$;
5. O vetor $(u_1, \dots, u_n)^T$ é uma observação do vetor $(U_1, \dots, U_n)$ que tem distribuição C_R^{Ga} ;
6. Transforme o vetor uniforme $(u_1, \dots, u_n)$ no vetor desejado $(x_1, \dots, x_n)$, utilizando as inversas $x_i = F_i^{(-1)}(u_i)$ para $i = 1, \dots, n$.

7.4 Acoplamentos *t*-Student

A Equação (4.1) nos dá uma maneira simples de gerar observações de um acoplamento *t*, denotado por $C_{\rho,\nu}^t$, como definido na Subseção 4.1.3. O algoritmo é apresentado a seguir :

1. Encontre a decomposição de Cholesky A da matriz R ;
2. Simule n variáveis normal padrão independentes $\mathbf{z} = (z_1, \dots, z_n)^T$;
3. Simule uma variável s , de uma distribuição χ_ν^2 , independente de $\mathbf{z}$;
4. Faça $\mathbf{y} = A\mathbf{z}$;
5. Faça $\mathbf{x} = \frac{\sqrt{\nu}}{\sqrt{s}}\mathbf{y}$;
6. Determine os componentes $u_i = t_\nu(x_i)$, para $i = 1, \dots, n$;
7. O vetor $(u_1, \dots, u_n)^T$ é uma observação do vetor $(U_1, \dots, U_n)$ que tem distribuição $C_{\rho,\nu}^t$;
8. Transforme o vetor uniforme $(u_1, \dots, u_n)$ no vetor desejado $(x_1, \dots, x_n)$, utilizando as inversas $x_i = F_i^{(-1)}(u_i)$ para $i = 1, \dots, n$.

7.5 Acoplamentos arquimedianos

No caso de acoplamentos arquimedianos, temos dois resultados que irão nos auxiliar na construção de um algoritmo de simulação. O primeiro deles é um corolário do Teorema 5.1.

Corolário 7.1 *Seja $(U, V)^T$ um vetor aleatório com distribuição conjunta C , onde C é um acoplamento arquimadiano gerado por φ . Então a função K_C dada por (5.22) é a função de distribuição da variável aleatória $C(U, V)$. ■*

A prova do teorema a seguir pode ser encontrada em Nelsen (1999), (pg 104).

Teorema 7.1 *Sob as hipóteses do Corolário 7.1, a função de distribuição conjunta $H(s, t)$ das variáveis aleatórias $S = \varphi(U)/[\varphi(U) + \varphi(V)]$ e $T = C(U, V)$ é dada por $H(s, t) = sK_C(t)$ para todo $(s, t) \in [0, 1]^2$. Então S e T são independentes, e S é uniformemente distribuída em $[0, 1]$. ■*

Temos então o seguinte algoritmo para simulação :

1. Gere duas variáveis Uniforme(0,1) independentes, s e q ;
2. Faça $t = K_C^{-1}(q)$, onde K_C é a função de distribuição da variável aleatória $C(U, V)$. Deste modo temos que t é uma observação da variável aleatória $T = C(U, V)$;
3. Faça $u = \varphi^{[-1]}(s\varphi(t))$ e $v = \varphi^{[-1]}((1-s)\varphi(t))$;
4. Temos que o par (u, v) é uma observação de um vetor aleatório (U, V) com distribuição conjunta C .

Prova: Como $T = C(U, V)$, e C é um acoplamento arquimadiano com função geradora φ , então $T = \varphi^{[-1]}(\varphi(U) + \varphi(V))$. Deste modo, temos:

$$\varphi(T) = \varphi(U) + \varphi(V). \quad (7.4)$$

Da definição de $S = \varphi(U)/[\varphi(U) + \varphi(V)]$ e de (7.4), temos que $S = \varphi(U)/\varphi(T)$, o que implica que $U = \varphi^{[-1]}(S\varphi(T))$, que é a transformação realizada no passo 3 do algoritmo. De maneira semelhante para V , da definição de S temos que :

$$S = \varphi(U)/[\varphi(U) + \varphi(V)] = 1 - \varphi(V)/[\varphi(U) + \varphi(V)]. \quad (7.5)$$

Assim utilizando novamente (7.4) temos que $S = 1 - \varphi(V)/\varphi(T)$, o que implica que $1 - S = \varphi(V)/\varphi(T)$, e desta maneira $V = \varphi^{[-1]}((1 - S)\varphi(T))$. □

Note que as variáveis s e t correspondem às variáveis aleatórias S e T do Teorema 7.1.

7.6 Acoplamento de Plackett

Um algoritmo para simular dados distribuídos de acordo com o acoplamento de Plackett é encontrado em Rockinger e Jondeau (2001), baseado em Johnson (1987). Consiste em simular uma observação de uma das marginais e então simular a outra marginal utilizando o acoplamento. O algoritmo é apresentado a seguir :

1. Gere duas variáveis Uniforme(0,1) independentes, u e t ;
2. Defina $a = t(1 - t)$ e $b = \sqrt{\theta} \sqrt{\theta + 4au(1 - u)(1 - \theta)^2}$;
3. Calcule $v = [2a(u\theta^2 + 1 - u) + \theta(1 - 2a) - (1 - 2t)b] / [2\theta + a(\theta - 1)^2]$;
4. u e v serão distribuídos de acordo com uma distribuição de Plackett.

8 *Aplicação*

Neste capítulo nós aplicaremos a teoria de funções de acoplamentos condicionais na modelagem da distribuição bivariada condicional das taxas diárias de câmbio de fechamento Libra Esterlina - Dólar Americano e Iene - Dólar Americano, durante o período de 02 de Janeiro de 1990 a 05 de Dezembro de 2003. Após a estimação de alguns acoplamentos e de modelos para as distribuições marginais, iremos utilizar técnicas de simulação para verificar se a estimação do Valor em Risco (VaR) obtida pelos modelos estimados é adequada. No final do Capítulo apresentamos um pequeno resumo sobre a análise de outra carteira, composta pelos índices americanos Nasdaq e S&P500, cuja análise mais detalhada está no Apêndice A.

8.1 Descrição da amostra

Como dito anteriormente, o conjunto de dados usado para análise compreende as taxas diárias de câmbio Libra Esterlina - Dólar Americano e Iene - Dólar Americano, durante o período de 01 de Janeiro de 1990 a 05 de Dezembro de 2003, num total de 3503 observações. As taxas utilizadas correspondem ao câmbio comercial oficial do FED (Banco Central norte-americano), taxa de fechamento diário de venda. É o referencial cambial mais comum do mercado financeiro, embora as pessoas físicas não comprem este tipo de moeda. A taxa chamada de venda se refere a quanto o banco está cobrando para vender a moeda.

Por notação, denotaremos por $P_{1,t}$ a taxa de câmbio Libra Esterlina - Dólar Americano no dia t , onde $t = 0, \dots, 3502$. De maneira análoga, denotaremos por $P_{2,t}$ a taxa de câmbio Iene - Dólar Americano no dia t .

Não temos problema de sincronização dos dados, com dados faltantes em apenas um dos ativos, pois ambos os ativos são negociados no mesmo mercado.

Como dito no início do Capítulo 6, iremos trabalhar com os retornos financeiros destas taxas, fazendo as transformações :

$$\begin{aligned} X_{1,t} &= \ln(P_{1,t}) - \ln(P_{1,t-1}), \\ X_{2,t} &= \ln(P_{2,t}) - \ln(P_{2,t-1}), \end{aligned}$$

para $t = 1, \dots, 3502$.

Temos então as duas séries de retornos financeiros que iremos estudar e modelar nas próximas seções deste capítulo.

8.2 Análise exploratória de dados

Nesta seção iremos inicialmente fazer uma análise exploratória dos retornos univariados das séries consideradas (Libra), ajustando então modelos que foram considerados no Capítulo 6.

Temos a seguir algumas estatísticas descritivas para as séries consideradas.

Observamos que ambas as séries possuem assimetria negativa, devido à cauda maior na esquerda (perdas). A distribuição dos retornos do Iene possui uma curtose maior do que a dos retornos da Libra, sendo ambas maiores do que o caso normal. Isto indica a existência do fato estilizado conhecido como caudas pesadas.

Temos na Figura 8.1 os gráficos da série de retornos e de retornos absolutos. Podemos notar no gráfico de retornos absolutos a existência dos chamados agrupamentos de volatilidade, com regiões de alta e baixa volatilidade, também um fato estilizado citado no Capítulo 6. A região de alta volatilidade do Iene

Estatística	Libra	Iene
Média	0,00002	-0,00009
Média (anualizada)	0,502%	-2,174%
Desvio Padrão	0,00571	0,00705
Mínimo	-0,03286	-0,05630
Mediana	0,00013	0,00000
Máximo	0,02889	0,03240
Excesso de Curtose	2,59090	4,15646
Assimetria	-0,24432	-0,54509

Tabela 8.1: Estatísticas descritivas dos retornos de Libra e Iene.

em 1998 corresponde ao período da crise da Rússia e da quebra do *hedge fund* americano LTCM (*Long Term Capital Management*), em Agosto de 1998.

Na figura 8.2 vemos as autocorrelações e correlações parciais do retornos de Libra e Iene. Para os retornos de Libra são estatisticamente significantes as autocorrelações de ordens 1, 4 e 9 (estatística t vale 3,36, 2,61 e 2,40 respectivamente, com p-valores de 0,0008; 0,0090 e 0,0163 respectivamente). Para os retornos de Iene temos estatisticamente significativa a autocorrelação de ordem 3 ($t = -2,71$, p-valor de 0,0067). Iremos modelar o nível dos retornos da Libra através de um processo AR, como veremos na subseção seguinte.

A Figura 8.3 é o diagrama de dispersão das séries de retornos e das séries de retornos uniformizadas por suas respectivas distribuições empíricas. Notamos a evidência de associação negativa entre os ativos, tanto pelo gráfico dos retornos como pela séries uniformizadas, onde vemos uma maior concentração de pontos na diagonal secundária. Porém iremos considerar uma posição vendida em Libra Esterlina, de modo que a função de acoplamento irá modelar a distribuição conjunta entre X_1 e X_2 onde

$$\begin{aligned} X_{1,t} &= -(\ln(P_{1,t}) - \ln(P_{1,t-1})), \\ X_{2,t} &= \ln(P_{2,t}) - \ln(P_{2,t-1}), \end{aligned}$$

para $t = 1, \dots, 3502$. Neste caso a associação entre as variáveis será positiva.

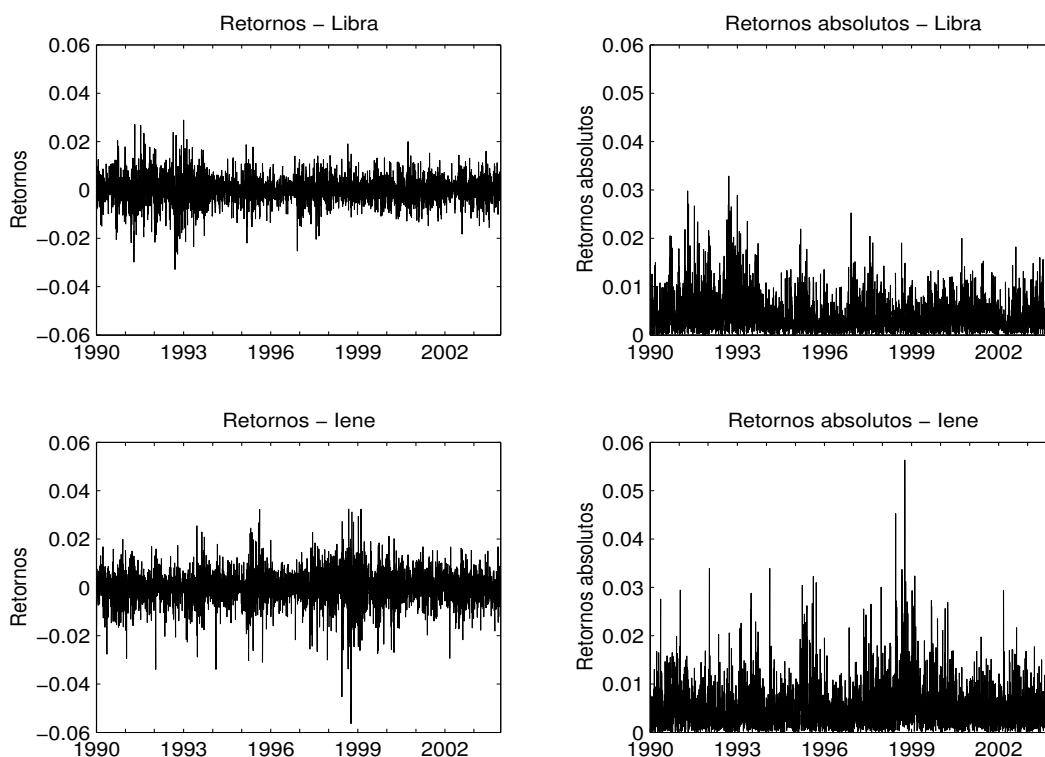


Figura 8.1: Gráficos de Retornos e Retornos Absolutos (Libra e Iene).

Caso tivéssemos uma posição comprada nos dois ativos, poderíamos modelar a dependência entre eles utilizando um acoplamento de associação negativa ou através da rotação de um acoplamento de associação positiva. É importante lembrar que no caso especial em que temos uma carteira composta pelas taxas -Libra/Dólar e Iene/Dólar na mesma proporção em termos de valores, estamos simplesmente modelando a taxa de câmbio Iene/Libra. Mas iremos considerar em nosso estudo, na Seção 8.5, uma outra proporção para os ativos na carteira ($2/3$ e $1/3$).

A função de acoplamento não será aplicada diretamente nesta série de retornos bivariada, mas sim nas séries transformadas que estimaremos na próxima seção.

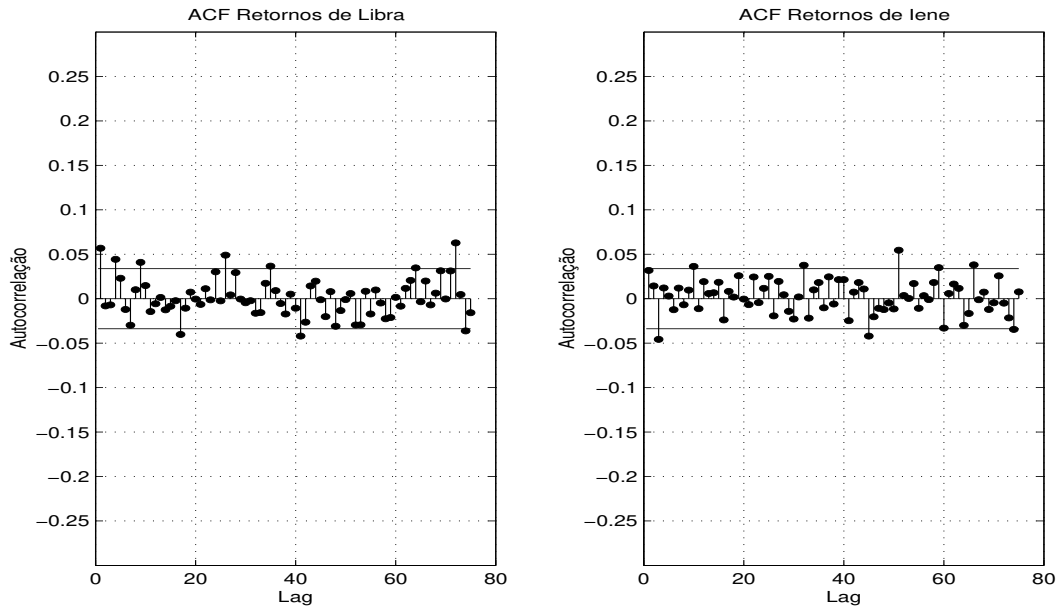


Figura 8.2: Função de Autocorrelação e Função de Autocorrelação Parcial - retornos de Libra e Iene.

8.3 Modelagem das distribuições marginais

Ajustamos como modelos tentativos iniciais processos AR(1) com erros GARCH(1,1) para ambas as séries X_1 e X_2 . Utilizamos inovações com distribuição normal, t-Student e distribuição empírica. O método de estimação utilizada é o de máxima verossimilhança, sem restrições, através da função de maximização do MATLAB. No caso da distribuição empírica, o método utilizado é o da máxima quase-verossimilhança, em que os parâmetros são estimados como se as inovações fossem normais (e portanto os mesmos do caso normal).

Para a série X_1 (retornos da posição vendida em Libra) o coeficiente AR(1) é estatisticamente significativamente ao nível de significância de 5 % (p-valor de 0,0430) para o modelo GARCH-N. Ao incluirmos coeficientes MA(1) e AR(2) estes não são significantes ao nível de significância de 5% (p-valores de 0,6028 e 0,7123 respectivamente). Já para a série X_2 (retornos de Iene) nem o coeficiente

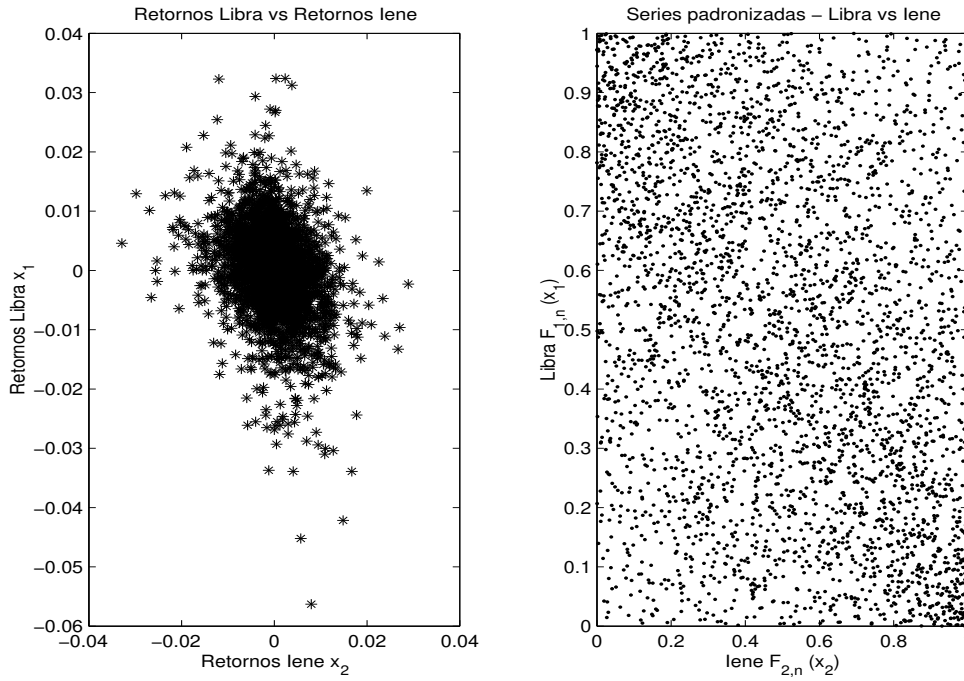


Figura 8.3: Diagrama de dispersão Retornos da Libra x Retornos do Iene.

AR(1) é estatisticamente significativa a 5% de nível de significância (p-valor de 0,1903), de modo que iremos considerar apenas o modelo GARCH(1,1). Os modelos então são dados por :

$$\begin{cases} X_{1,t} = \mu_1 + \phi_1 X_{1,t-1} + \epsilon_{1,t} ; \\ \epsilon_{1,t} = \sigma_{1,t} \eta_{1,t} ; \\ \sigma_{1,t}^2 = \alpha_1 + \beta_1 \epsilon_{1,t-1}^2 + \gamma_1 \sigma_{1,t-1}^2, \end{cases} \quad (8.1)$$

$$\begin{cases} X_{2,t} = \mu_2 + \epsilon_{2,t} ; \\ \epsilon_{2,t} = \sigma_{2,t} \eta_{2,t} ; \\ \sigma_{2,t}^2 = \alpha_2 + \beta_2 \epsilon_{2,t-1}^2 + \gamma_2 \sigma_{2,t-1}^2, \end{cases} \quad (8.2)$$

onde $\{\eta_{1,t}\}$ e $\{\eta_{2,t}\}$ são processos ruído branco.

Temos então :

$$\begin{aligned}\epsilon_{1,t} &= x_{1,t} - \mu_1 - \phi_1 x_{1,t-1}, \\ \epsilon_{2,t} &= x_{2,t} - \mu_2\end{aligned}$$

Para os modelos GARCH-Normal (GARCH-N) temos

$$\eta_{i,t} = \frac{\epsilon_{i,t}}{\sigma_{i,t}} / \mathcal{F}_{i,t-1} \sim \mathcal{N}(0, 1). \quad (8.3)$$

Para os modelos GARCH-t temos :

$$\eta_{i,t} = \frac{\epsilon_{i,t}}{\sigma_{i,t}} / \mathcal{F}_{i,t-1} \sim t_{\nu_i}. \quad (8.4)$$

Para os modelos GARCH-Empírica (GARCH-E) temos :

$$\eta_{i,t} = \frac{\epsilon_{i,t}}{\sigma_{i,t}} / \mathcal{F}_{i,t-1} \sim F_{\eta_i}. \quad (8.5)$$

para $i = 1, 2$, onde $t = 1, \dots, T$ e onde F_{η_1} e F_{η_2} são as distribuições empíricas das séries $\{\hat{\eta}_{1,t}\}$ e $\{\hat{\eta}_{2,t}\}$ (inovações observadas).

As estimativas de máxima verossimilhança obtidas através do programa são dadas pela Tabela 8.2. Podemos ver que a série de retornos da posição vendida em Libra tem persistência estimada de 0,9866 e 0,9751 ($\beta_1 + \gamma_1$) (nos modelos GARCH-N e GARCH-t respectivamente) e a série de retornos de Iene tem uma persistência bem próxima, de 0,9842 e 0,9757 ($\beta_2 + \gamma_2$).

Pelo critério de AIC o modelo GARCH-t é o melhor para ambas as séries.

Temos então uma análise dos resultados obtidos nos modelos ajustados. Para isto estudaremos tanto o comportamento das inovações observadas ($\hat{\epsilon}_{1,t}$ e $\hat{\epsilon}_{2,t}$) como o das inovações padronizadas ($\hat{\eta}_{1,t}$ e $\hat{\eta}_{2,t}$), dadas por:

$$\hat{\eta}_{1,t} = \frac{\hat{\epsilon}_{1,t}}{\hat{\sigma}_{1,t}} = \frac{x_{1,t} - \hat{\mu}_1 - \hat{\phi}_1 x_{1,t-1}}{\sqrt{\hat{\alpha}_1 + \hat{\beta}_1 \hat{\epsilon}_{1,t-1}^2 + \hat{\gamma}_1 \hat{\sigma}_{1,t-1}^2}},$$

	Parâmetro	GARCH-N	GARCH-t
Libra (vendido)	μ_1	$-9,95 \times 10^{-5}$ ($8,41 \times 10^{-5}$)	$-1,48 \times 10^{-4}$ ($7,72 \times 10^{-5}$)
	ϕ_1	0,0349 (0,0177)	0,0164 (0,0160)
	α_1	$4,19 \times 10^{-7}$ ($1,15 \times 10^{-7}$)	$1,80 \times 10^{-7}$ ($7,96 \times 10^{-8}$)
	β_1	0,0475 (0,0069)	0,0346 (0,0058)
	γ_1	0,9391 (0,0093)	0,9405 (0,0102)
	ν_1		5,7566 (0,4136)
	AIC_1	$-2,672 \times 10^{-4}$	$-2,688 \times 10^{-4}$
Iene	μ_2	$-5,40 \times 10^{-5}$ ($1,08 \times 10^{-4}$)	$1,07 \times 10^{-4}$ ($9,87 \times 10^{-5}$)
	α_2	$7,65 \times 10^{-7}$ ($2,12 \times 10^{-7}$)	$3,55 \times 10^{-7}$ ($1,24 \times 10^{-7}$)
	β_2	0,0383 (0,0064)	0,0212 (0,0039)
	γ_2	0,9459 (0,0095)	0,9545 (0,0086)
	ν_2		5,6850 (0,5313)
	AIC_2	$-2,512 \times 10^{-4}$	$-2,537 \times 10^{-4}$

Tabela 8.2: Parâmetros estimados dos modelos GARCH em (8.1) e (8.2) e erro padrão das estimativas.

$$\hat{\eta}_{2,t} = \frac{\hat{\epsilon}_{2,t}}{\hat{\sigma}_{2,t}} = \frac{x_{2,t} - \hat{\mu}_2}{\sqrt{\hat{\alpha}_2 + \hat{\beta}_2 \hat{\epsilon}_{2,t-1}^2 + \hat{\gamma}_2 \hat{\sigma}_{2,t-1}^2}},$$

para $t = 1, \dots, 3502$.

Caso o modelo esteja bem especificado, esperamos que as séries $\hat{\epsilon}_{1,t}$ e $\hat{\epsilon}_{2,t}$ não apresentem autocorrelações significantes indicando que o modelo AR(1) modelou adequadamente o nível dos retornos. Esperamos que as séries $\hat{\epsilon}_{1,t}^2$ e $\hat{\epsilon}_{2,t}^2$ apresentem autocorrelações significantes de modo que temos evidência da necessidade do modelo GARCH. Caso o modelo seja adequado, esperamos também que as séries $\hat{\eta}_{1,t}$ e $\hat{\eta}_{2,t}$ não apresentem autocorrelações significantes. Finalmente esperamos que o quadrado destas séries não sejam autocorrelacionados, o que indica que o modelo explicou adequadamente a variância condicional dos retornos.

Pela Figura 8.4 vemos que não existem autocorrelações altamente significantes nas séries $\hat{\epsilon}_{1,t}$ e $\hat{\epsilon}_{2,t}$. Temos como exceção na série $\hat{\epsilon}_{1,t}$ a autocorrelação

de ordem 4, cuja estimativa vale 0,04, e cuja estatística t do teste de correlação nula valem 2,60 (p-valor de 0,0093). Apesar desta autocorrelação significativa, o modelo AR(1) diminui a autocorrelação existente originalmente nos retornos, e modelos de ordens superiores não eliminaram esta autocorrelação significativa, de modo que iremos manter este modelo. Na série $\hat{\epsilon}_{2,t}$ temos a autocorrelação de ordem 3, cuja estimativa vale -0,05 e a estatística t vale -2,70 (p-valor de 0,0069). Os valores citados são para os resíduos dos modelos GARCH-N, sendo que os valores obtidos para os modelos GARCH-t são bem próximos. Já na Figura 8.5 vemos que existe evidência de autocorrelações significantes nas séries $\hat{\epsilon}_{1,t}^2$ e $\hat{\epsilon}_{2,t}^2$ indicando a necessidade do modelo GARCH para a volatilidade condicional.

Pelas Figura 8.6 vemos que não existem correlações altamente significantes para ambas as séries $\hat{\eta}_{1,t}$ e $\hat{\eta}_{2,t}$ nos dois modelos. Temos como exceção na série $\hat{\eta}_{1,t}$ a autocorrelação de ordem 26, cuja estimativa vale 0,05, e cuja estatística t do teste de correlação nula valem 2,86 (p-valor de 0,0042). Na série $\hat{\eta}_{2,t}$ temos como significativa a autocorrelação de ordem 10, cuja estimativa vale -0,05 e a estatística t vale 2,71 (p-valor de 0,0067). O teste de Ljung-Box para autocorrelações até ordem 10 não rejeita a hipótese de autocorrelações nulas a 5% de significância. Os p-valores para as séries $\hat{\eta}_{1,t}$ e $\hat{\eta}_{2,t}$ são de 0,3200 e 0,0871 no modelo GARCH-N e de 0,1853 e 0,0833 no modelo GARCH-t.

Pelas Figura 8.7 vemos que não existem autocorrelações altamente significantes para as séries $\hat{\eta}_{1,t}^2$ e $\hat{\eta}_{2,t}^2$, de modo que os modelos explicam adequadamente a heteroscedasticidade da série original. Temos como exceção a autocorrelação de ordem 25 da série $\hat{\eta}_{1,t}^2$, cuja estatística t do teste de autocorrelação nula vale 3,17 (p-valor de 0,0015). O teste de Ljung-Box para autocorrelações até ordem 10 não rejeita a hipótese de autocorrelações nulas. Os p-valores para as séries $\hat{\eta}_{1,t}^2$ e $\hat{\eta}_{2,t}^2$ são de 0,9589 e 0,7592 no modelo GARCH-N e de 0,9679 e 0,6362 no modelo GARCH-t.

Utilizaremos as transformações descritas em (5.12), ou seja, $u_{1,t} = F_{1,t}(x_{1,t}; \hat{\vartheta}_1)/\mathbf{w}_t$ e $u_{2,t} = F_{2,t}(x_{2,t}; \hat{\vartheta}_2)/\mathbf{w}_t$ onde os coeficientes estimados dos vetores $\vartheta_1 = (\mu_1, \phi_1, \alpha_1, \beta_1, \gamma_1, \nu_1)$ e $\vartheta_2 = (\mu_2, \alpha_2, \beta_2, \gamma_2, \nu_2)$ estão na Tabela 8.2 (ν_1 e ν_2 existem apenas para o modelo GARCH-t).

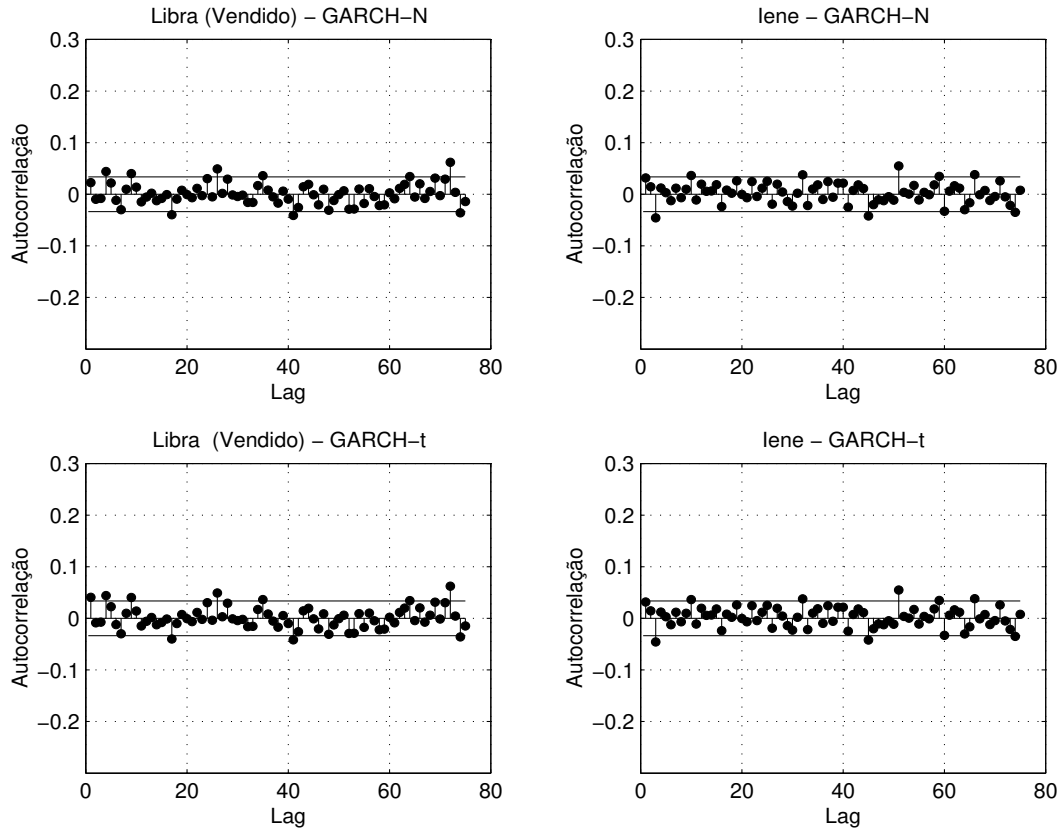


Figura 8.4: Função de Autocorrelação das inovações observadas - $\hat{\epsilon}_{1,t}$ e $\hat{\epsilon}_{2,t}$.

É interessante notar que é equivalente transformarmos as séries $\{x_{1,t}\}$ e $\{x_{2,t}\}$ utilizando-se as distribuições $F_{1,t}$ e $F_{2,t}$ ou transformar as séries de inovações $\{\eta_{1,t}\}$ e $\{\eta_{2,t}\}$ utilizando-se as distribuições consideradas em 8.3, 8.4 e 8.5. Por exemplo, os resultados obtidos serão os mesmos no modelo GARCH-N utilizando-se a transformação $u_{1,t} = F_{1,t}(x_{1,t})$ onde $F_{1,t}$ é a função distribuição acumulada de uma $\mathcal{N}(\mu_1 + \phi_1 x_{1,t-1}, \sigma_{1,t}^2)$ ou utilizando-se a transformação $u_{1,t} = \Phi(\eta_{1,t})$ onde Φ é a função distribuição acumulada de uma Normal padrão.

Após feitas as transformações, vamos analisar as séries $\{u_{1,t}\}$ e $\{u_{2,t}\}$.

No caso teórico em que os modelos para distribuições marginais fossem

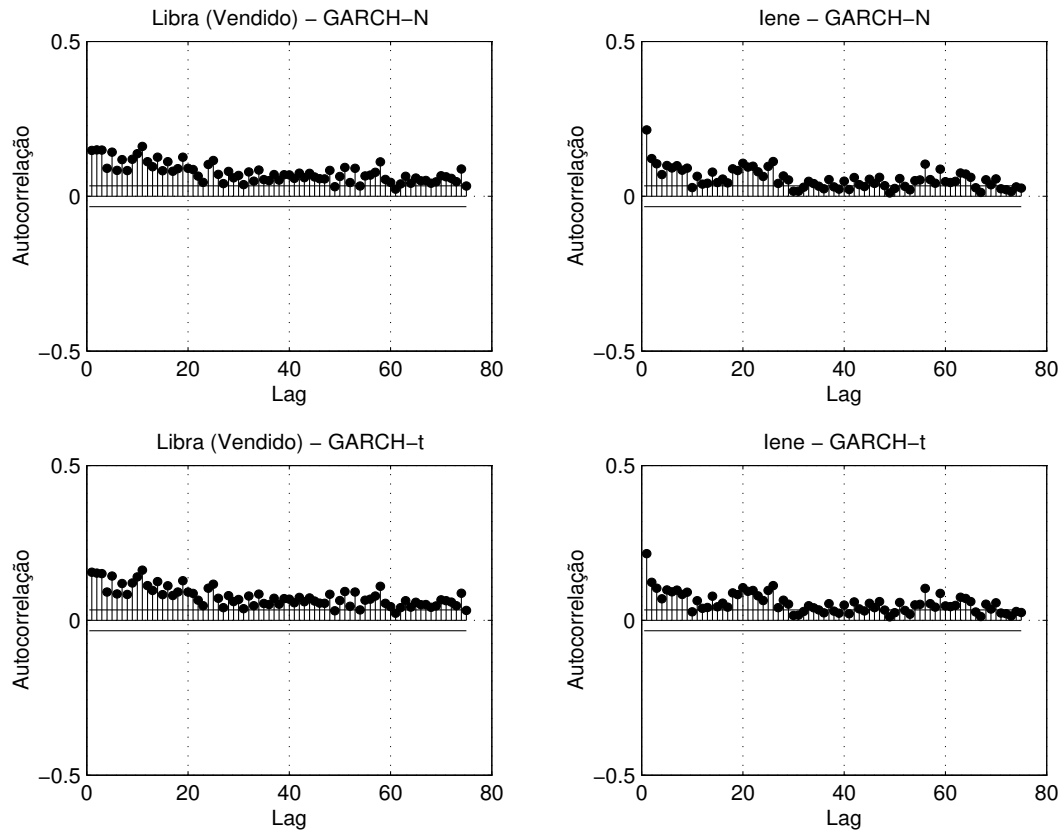


Figura 8.5: Função de Autocorrelação das inovações observadas ao quadrado - $\hat{\epsilon}_{1,t}^2$ e $\hat{\epsilon}_{2,t}^2$.

especificados e estimados sem erro, a distribuição teórica de ambas as séries transformadas $\{u_{1,t}\}$ e $\{u_{2,t}\}$ seria Uniforme(0,1) i.i.d., como dito na Subseção 5.1.5. Na Figura 8.8, podemos observar que a distribuição empírica das duas séries transformadas se aproximam da distribuição Uniforme(0,1). Para ambas séries o resultado obtido com o modelo GARCH-t é melhor. O modelo GARCH-E não foi exibido pois sua distribuição empírica irá sempre coincidir com a Uniforme[0,1] por construção.

Assim, visto que as séries transformadas $\{u_{1,t}\}$ e $\{u_{2,t}\}$ são aproximadamente duas amostras Uniforme (0,1) i.i.d. para os modelos considerados, a

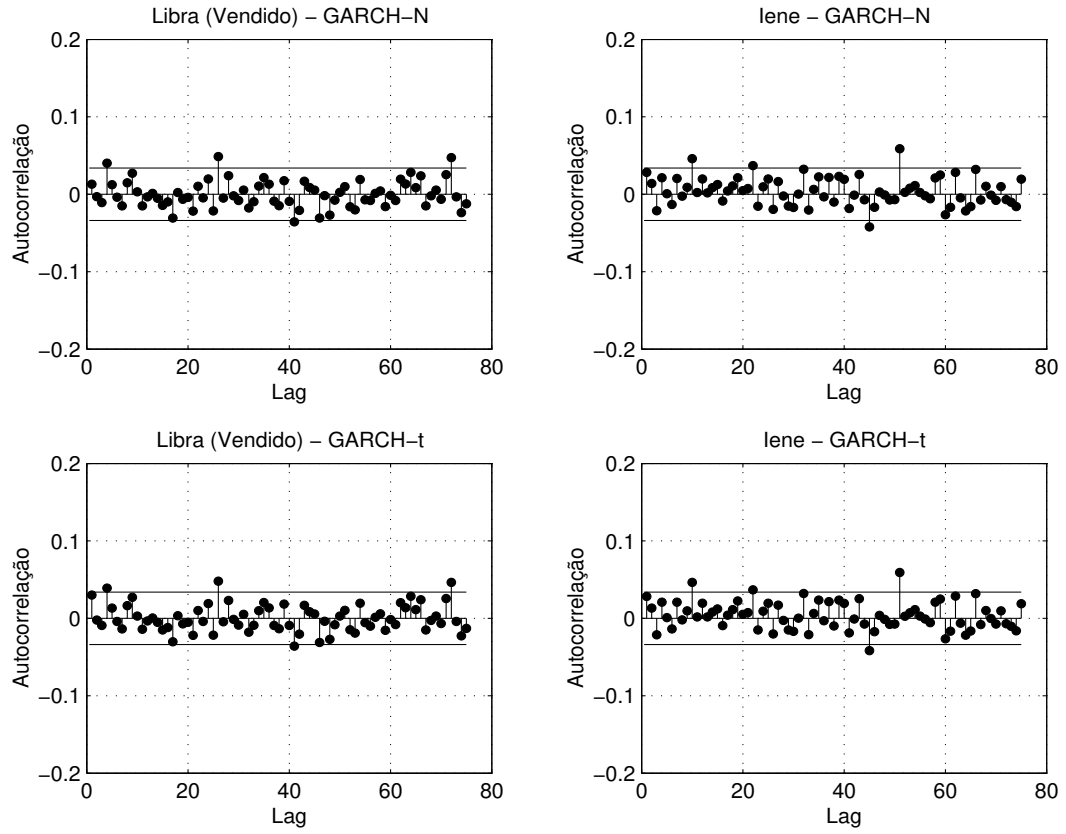


Figura 8.6: Função de Autocorrelação das inovações padronizadas - $\hat{\eta}_{1,t}$ e $\hat{\eta}_{2,t}$.

dependência temporal nas séries originais $\{x_{1,t}\}$ e $\{x_{2,t}\}$ foi satisfatoriamente explicada pelos modelos para as distribuições marginais dos retornos e removida com as transformações em (5.12). A dependência entre as duas séries ainda existe, e será modelada a seguir, através de função de acoplamento.

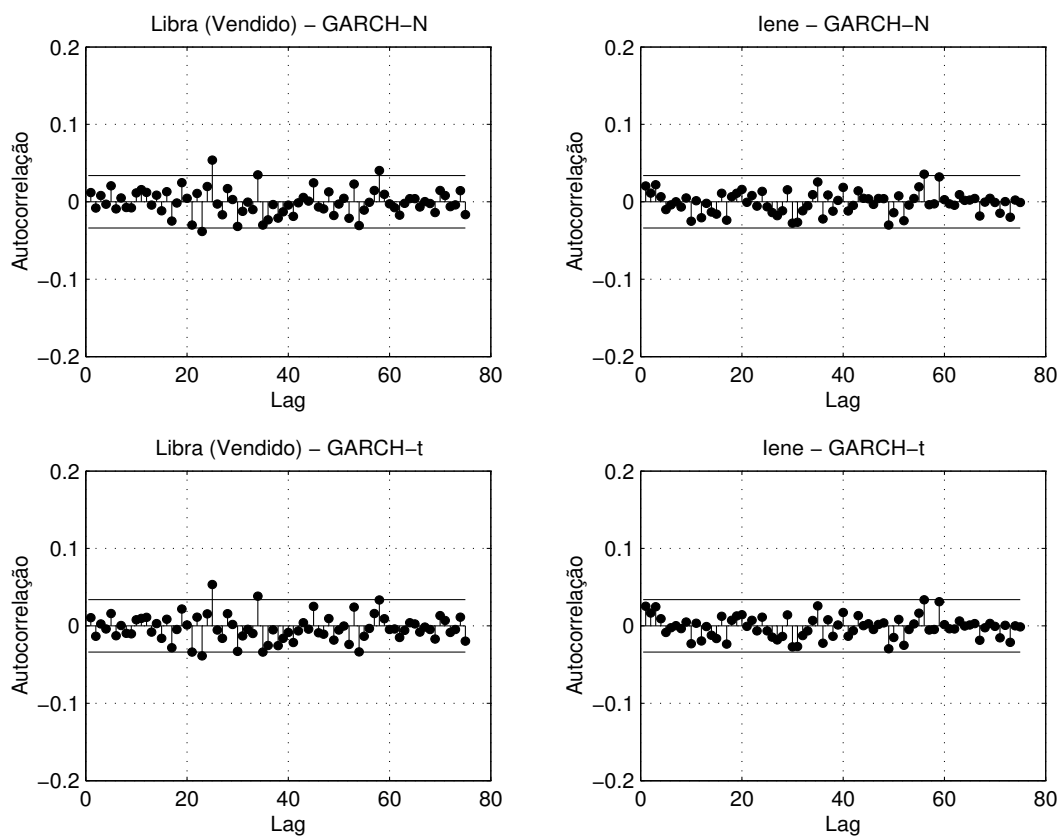


Figura 8.7: Função de Autocorrelação das inovações padronizadas ao quadrado $\hat{\eta}_{1,t}^2$ e $\hat{\eta}_{2,t}^2$.

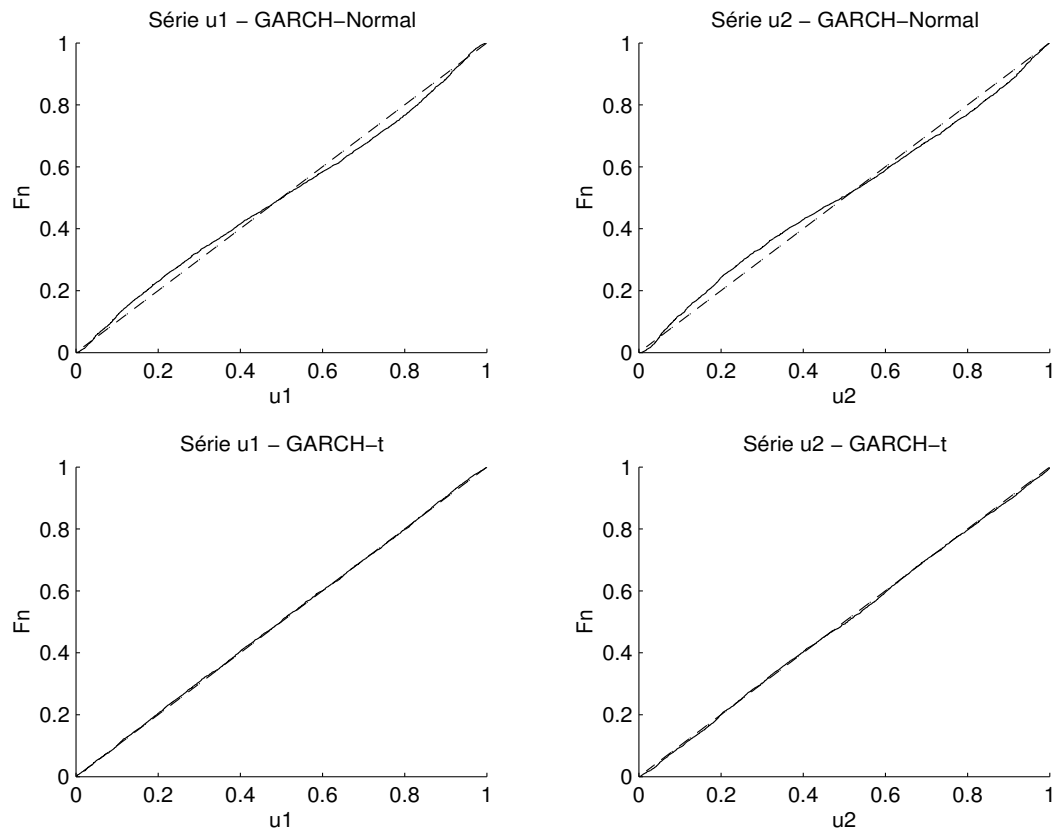


Figura 8.8: Função de Distribuição empírica - Séries u_1 e u_2 - Modelos GARCH-N e GARCH-t.

8.4 Modelagem de dependência

8.4.1 Análise Exploratória

Na Figura 8.9 observamos os gráficos bivariados das inovações ajustadas padronizadas para ambos os modelos.

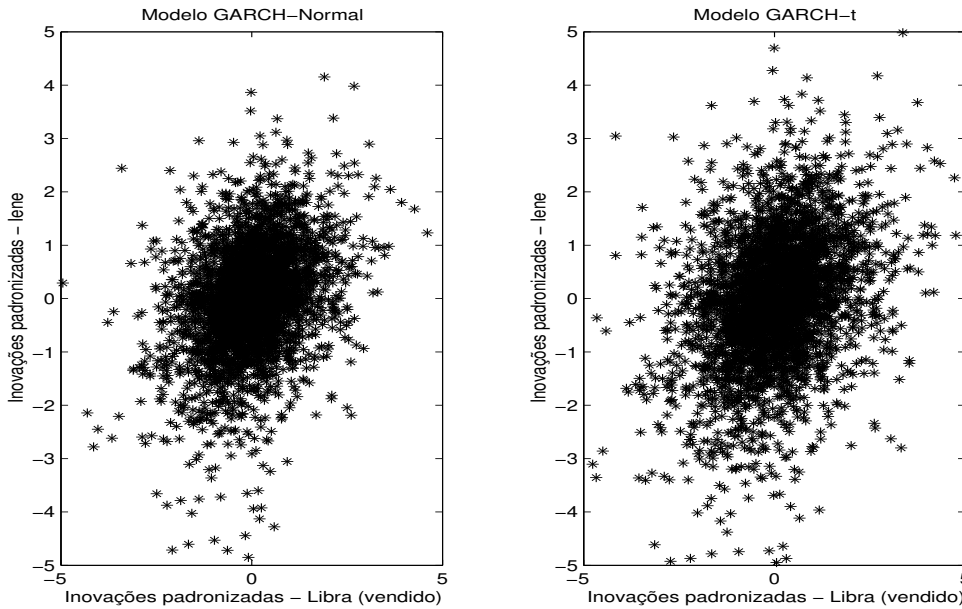


Figura 8.9: Gráficos das inovações ajustadas padronizadas (Libra(vendido) e Iene).

Para verificar o fato de dependência assimétrica, utilizaremos a correlação condicional, como definida em (3.1). Calculamos a correlação condicional para diversos valores de ξ entre -3 e 3 para as inovações dos modelos GARCH-N e GARCH-t. Obtivemos então a Figura 8.10.

Pela Figura 8.10 temos grande evidência de que a correlação entre os retornos não cai para zero no caso das perdas, mas cai mais rapidamente a zero no caso dos ganhos. Isto é uma evidência de existência de dependência assimétrica. Assim temos evidências de um desvio da distribuição normal bivariada (e da classe das distribuições elípticas).

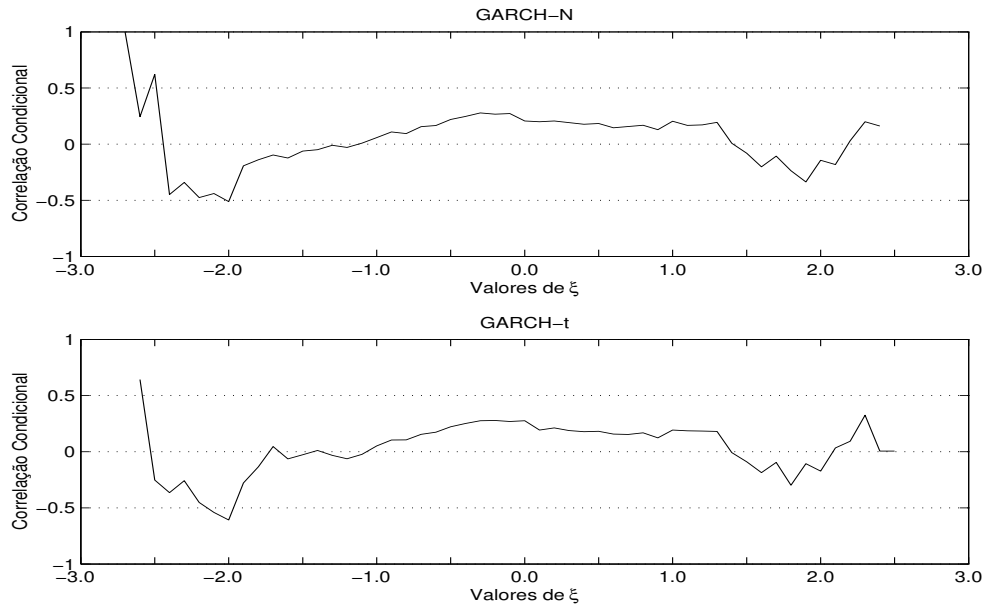


Figura 8.10: Correlações condicionais para as inovações dos modelos GARCH-N e GARCH-t.

8.4.2 Estimação da função de acoplamento

Iremos modelar a dependência através de três funções de acoplamento que tem sido utilizada em finanças : simetrizada de Joe-Clayton, t-Student e Plackett, apresentadas respectivamente na Seção 4.4, na Subseção 4.1.3 e na Seção 4.3. Depois iremos comparar os acoplamentos ajustados com o acoplamento empírico. Escolhemos estas três funções de acoplamento por serem bem diferentes, a SJC é baseada no acoplamento arquimediano de Joe Clayton, a de Plackett baseada na medida de *Odds Ratio* e a t-Student é elíptica.

Temos então a estimativa por máxima verossimilhança (método IFM) dos parâmetros dos acoplamentos dados na Tabela 8.3.

Podemos ver que na função de acoplamento simetrizada de Joe Clayton, o parâmetro τ_L , que pode justamente ser interpretado como o coeficiente de dependência na cauda inferior, tem valor estimado próximo de 0.40, indicando dependência na cauda inferior. Já o parâmetro τ_U tem valor estimado 0 nos

Acoplamento	Parâmetro	GARCH-N	GARCH-t	GARCH-E
SJC	τ_L	0,4038 (0,0126)	0,3951 (0,0150)	0,4006 (0,0050)
	τ_U	0,0000 (—)	0,0000 (—)	0,0000 (—)
	(AIC)	$-5,218 \times 10^4$	$-5,256 \times 10^4$	
Plackett	θ	3,3575 (0,1713)	2,9125 (0,1427)	2,9031 (0,1421)
	(AIC)	$-5,229 \times 10^4$	$-5,267 \times 10^4$	
t-Student	ν	14.5705 (2,4744)	8.4567 (1,0607)	8.9535 (1,1834)
	R_{12}	0,3428 (0,0158)	0,3350 (0,0157)	0,3360 (0,0157)
	(AIC)	$-5,228 \times 10^4$	$-5,270 \times 10^4$	

Tabela 8.3: Parâmetros estimados das funções de acoplamento (Desvios padrão em parênteses).

três modelos considerados, indicando a não existência de dependência na cauda superior. Este fato está de acordo com o fato estilizado de dependência assimétrica, com maior correlação nas perdas do que nos ganhos dos retornos dos ativos. O fato de $\tau_U = 0$ implica que o estimador de máxima verossimilhança não tem necessariamente distribuição assintótica normal, de modo que não foi exibido seu desvio padrão. Na estimação do desvio padrão do parâmetro τ_L assumiu-se que $\tau_U = 0$. O parâmetro estimado ν do acoplamento t-Student nos dá evidência da existência moderada de caudas pesadas. Temos para estes parâmetros estimados que $\lambda_U = \lambda_L$ vale 0,0142, 0,0566 e 0,0505 (para os três modelos considerados). Pelo critério de AIC temos que o acoplamento t-Student com marginais GARCH-t proporcionou o melhor ajuste.

Também iremos estimar os parâmetros por medidas de dependência para os acoplamentos de Plackett e t-Student. No caso do acoplamento de Plackett,

temos a seguinte relação entre o rho de Spearman e o parâmetro θ :

$$\rho_S = \frac{\theta + 1}{\theta - 1} - \frac{2\theta}{(\theta - 1)^2} \ln \theta. \quad (8.6)$$

Estimamos o rho de Spearman amostralmente e usamos (8.6) para obtermos a estimativa do parâmetro θ . No caso do acoplamento t-Student, utilizamos o método sugerido em Mashal. e Zeevi (2002). Estimamos o Tau de Kendall amostral utilizando o estimador $\hat{\tau}_k$ do Tau de Kendall em (5.6). Então estimamos o parâmetro R_{12} do acoplamento t-Student por

$$\hat{R}_{12} = \text{sen} \left(\frac{\pi}{2} \hat{\tau}_k \right).$$

Após isto, estimamos o parâmetro ν por máxima verossimilhança. Temos na Tabela 8.4 as medidas de dependência e parâmetros das funções de acoplamento estimados

Acoplamento	Medida de dependência estimada	Parâmetro estimado
Plackett	$\rho_S = 0,3205$	$\theta = 2,6998$
t-Student	$\tau_k = 0,2198$	$R_{12} = 0,3384$
		$\nu = 11,8845$

Tabela 8.4: Parâmetros estimados das funções de acoplamento por medidas de dependência.

8.4.3 Comparação e Diagnóstico

Para verificar o ajuste dos acoplamentos estimados podemos também calcular a distância quadrática em relação ao acoplamento empírico, como definido em (5.25). Como estamos interessados principalmente no ajuste na cauda inferior para efeitos de estimação do VaR, iremos calcular esta distância nas regiões $[0; 0, 2]^2$, $[0; 0, 1]^2$ e $[0; 0, 05]^2$. Iremos também fazer o gráfico das curvas de níveis dos acoplamentos empíricos e estimados na região $[0; 0, 2]^2$.

Podemos ver pela Tabela 8.5 que o acoplamento t-Student com marginais

Modelo	$\delta = 0,20$	$\delta = 0,10$	$\delta = 0,05$
SJC + GARCH-N	3,2104	1,3729	0,3201
SJC + GARCH-t	1,2302	0,5242	0,0764
SJC + GARCH-E	3,4554	1,4675	0,3507
Plackett + GARCH-N	4,9878	1,5740	0,3200
Plackett + GARCH-t	6,6791	1,8840	0,3689
Plackett + GARCH-E	6,7289	1,8929	0,3702
Plackett (usando ρ_S)	7,6065	2,0472	0,3940
t-Student + GARCH-N	1,3423	0,4492	0,0881
t-Student + GARCH-t	0,9795	0,3216	0,1407
t-Student + GARCH-E	1,0249	0,3327	0,1329
t-Student (usando τ_k)	1,2736	0,4105	0,0998

Tabela 8.5: Distância quadrática entre o acoplamento estimado e o acoplamento empírico na cauda inferior (quadrado $[0, \delta]^2$).

GARCH-t apresentou o melhor ajuste para as regiões $[0; 0, 1]^2$ e $[0; 0, 2]^2$, mas para a região $[0; 0, 05]^2$ o melhor foi o acoplamento Simetrizado de Joe Clayton com marginais GARCH-t. Provavelmente o ajuste melhor dado pelo acoplamento SJC em relação ao acoplamento t-Student na região de perdas extremas $[0; 0, 05]^2$ deve-se a maior flexibilidade desta função de acoplamento na modelagem da cauda da distribuição, pois o acoplamento SJC permite dependência assimétrica. O acoplamento de Plackett apresentou o pior ajuste em todos os casos. Este acoplamento também não permite dependência assimétrica entre as variáveis. Nas Figuras 8.11, 8.12 e 8.13 podemos comparar os acoplamentos estimados com o acoplamento empírico. As curvas de nível correspondem ao conjunto de pontos $(u, v) : C(u, v) = k$. No nosso caso as curvas traçadas correspondem, partindo do canto inferior esquerdo, aos valores de $k = 0,005 ; 0,010; 0,015; 0,020; \dots$

Dentre todos os modelos, considerando esta distância em relação ao acoplamento empírico na cauda inferior, os melhores modelos são SJC e t-Student com marginais GARCH-t. Poderíamos selecionar apenas estes modelos e utilizá-los na estimação do VaR, mas iremos prosseguir também com os outros

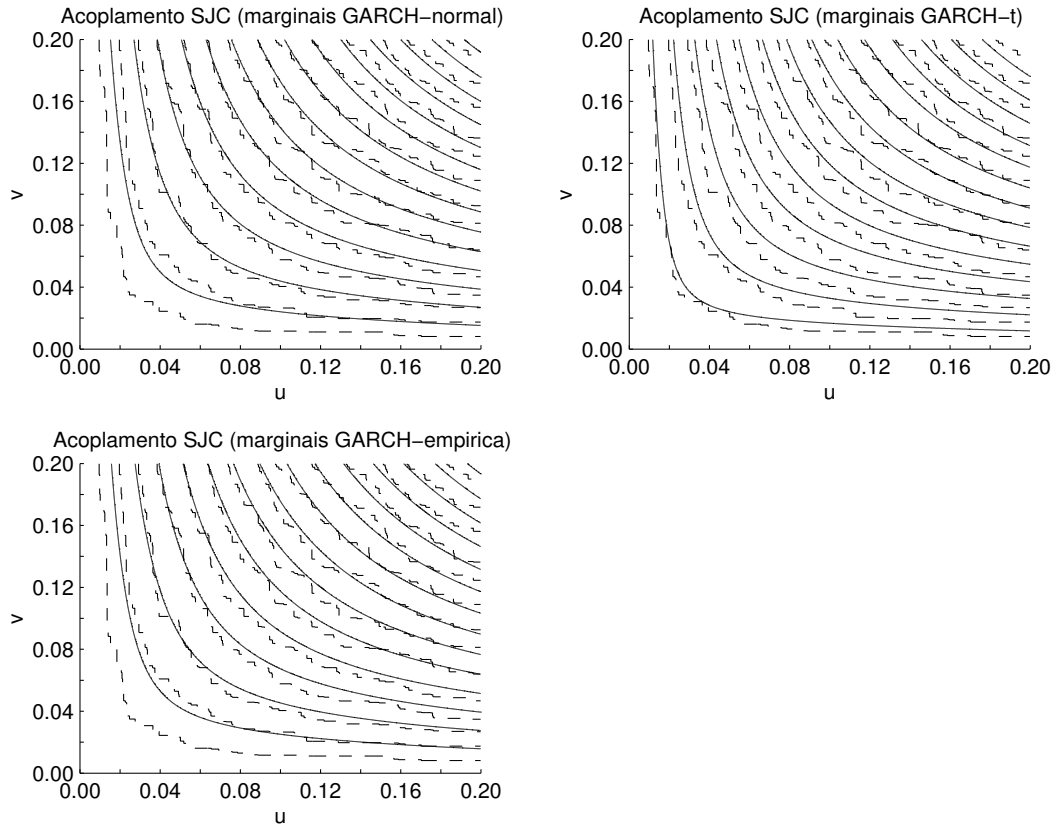


Figura 8.11: Curvas de nível da função de Acoplamento SJC estimada (linha cheia) e Empírica (linha tracejada). $((u, v) : C(u, v) = k$, onde $k = 0,005 ; 0,010; 0,015; 0,020; \dots$ a partir do canto inferior esquerdo).

modelos apenas para efeitos de comparação.

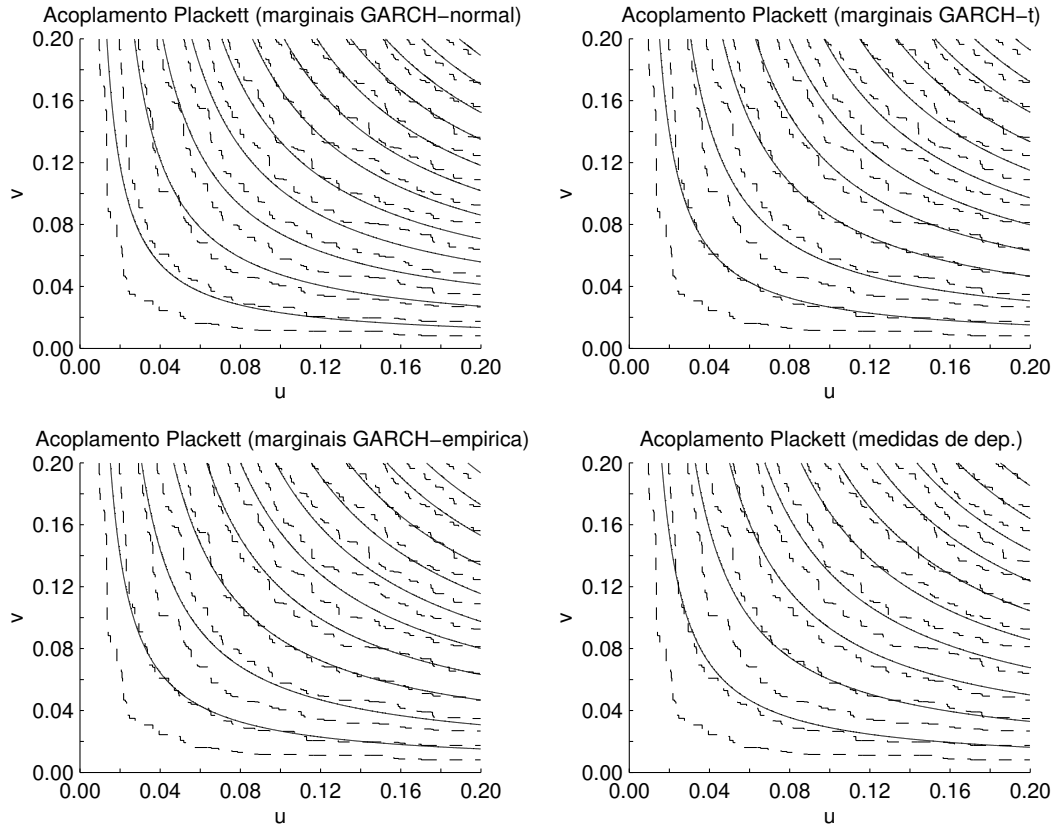


Figura 8.12: Curvas de nível da função de Acoplamento de Plackett estimada (linha cheia) e Empírica (linha tracejada). $((u, v) : C(u, v) = k$, onde $k = 0,005; 0,010; 0,015; 0,020; \dots$ a partir do canto inferior esquerdo).

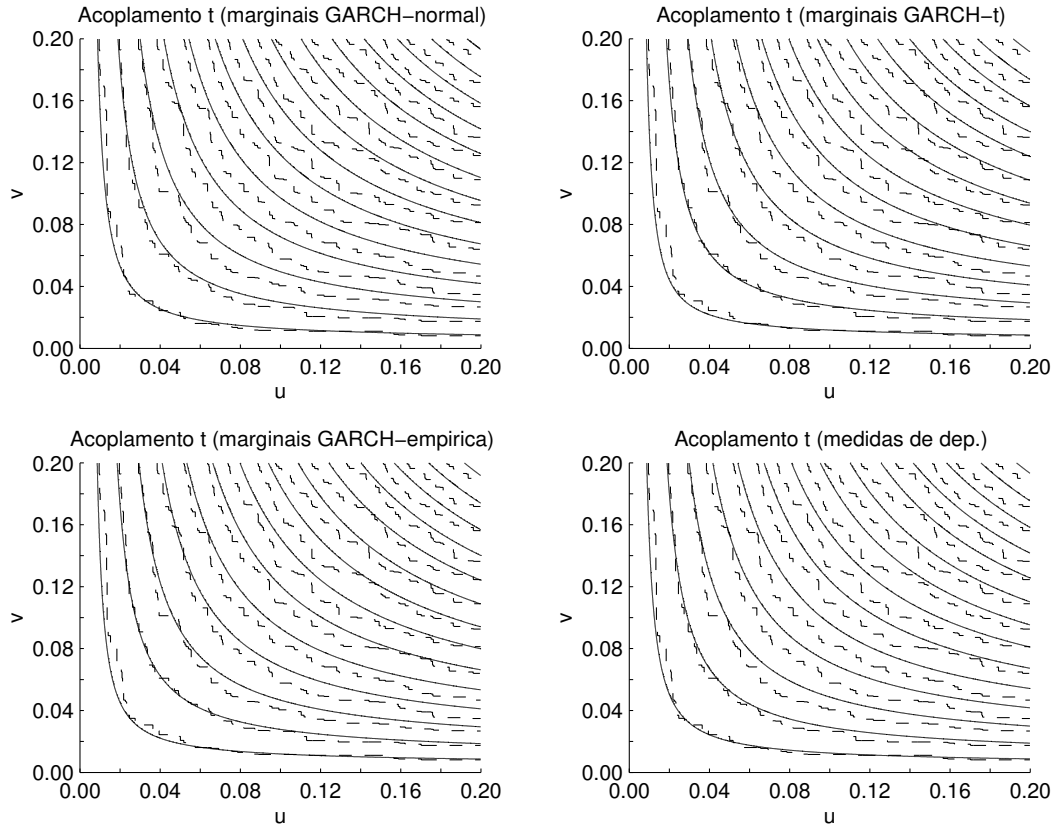


Figura 8.13: Curvas de nível da função de Acoplamento t-Student estimada (linha cheia) e Empírica (linha tracejada). $((u, v) : C(u, v) = k$, onde $k = 0,005 ; 0,010 ; 0,015 ; 0,020 ; \dots$ a partir do canto inferior esquerdo).

8.5 Estimação do Valor em Risco

Como definido no início deste capítulo, temos as séries de retornos de Libra $\{X_{1,t}\}$ (posição vendida) e Iene $\{X_{2,t}\}$, cujas distribuições marginais no tempo t foram denotadas por $F_{1,t}$ e $F_{2,t}$ e cuja distribuição conjunta no tempo t pode ser denotada por H_t . Como exercício de simulação iremos considerar agora que temos uma carteira composta destes dois ativos, cuja série de retornos será denotada por $\{X_{p,t}\}$, com distribuição $F_{p,t}$ no tempo t . Nesta carteira consideraremos uma proporção de 2:1 para a posição vendida em Libra, de modo que temos aproximadamente $X_{p,t} = (2/3)X_{1,t} + (1/3)X_{2,t}$. Esta escolha foi feita arbitrariamente, qualquer outra proporção poderia ser escolhida.

O Valor em Risco (VaR) da carteira no tempo t , com nível de confiança $1 - \alpha$, onde $\alpha \in (0, 1)$, é definido como :

$$VaR(\alpha)_t = \inf_s \{s : F_{p,t}(s) \geq \alpha\}.$$

Sabemos que $F_{p,t}$ poderia ser teoricamente calculada a partir da distribuição conjunta dos retornos H_t , que por sua vez pode ser obtida conhecendo-se a função de acoplamento C e as distribuições marginais. Nesta seção iremos simular dados da distribuição conjunta H_t para cada tempo t fixo, obtendo assim uma estimativa de $VaR(\alpha)_t$ para cada tempo t , definido por $\widehat{VaR(\alpha)}_t$. Esta simulação é necessária pois é bastante complicado obter uma forma fechada (analítica) da distribuição conjunta dos retornos a partir da função de acoplamento e das distribuições marginais. Após isto, iremos calcular a proporção de valores $X_{p,t}$ na amostra que são menores que $\widehat{VaR(\alpha)}_t$. Caso a função de acoplamento e as distribuições marginais sejam adequadas, esta proporção deve ser próxima de α . Para o acoplamento t-Student o algoritmo utilizado é o algoritmo específico da Seção 7.4. Para o acoplamento de Plackett o algoritmo utilizado é o da Seção 7.6. Já para o acoplamento SJC é utilizado o algoritmo geral proposto na Seção 7.1.

Como no tempo t conhecemos os dados para tempos iguais ou anteriores a t , os parâmetros das distribuições marginais e da função de acoplamento

deveriam ser estimados a cada novo tempo t para utilizar toda a informação disponível, o que acarretaria um grande custo computacional para o exercício de simulação. Devido a isto, iremos estimar o modelo apenas a cada 50 observações, começando a simulação com dados de $t = 1$ a $t = 750$, que é um tamanho adequado de amostra para permitir uma estimação adequada dos modelos GARCH. Desta forma, em $t = 800$ reestimamos os parâmetros dos modelos das distribuições marginais e do acoplamento, utilizando dados de $t = 50$ a $t = 800$, e assim por diante para $t = 850, 900, \dots, 3450$ e 3500 . Apesar dos parâmetros serem estimados a cada 50 observações, as distribuições marginais $F_{1,t}$ e $F_{2,t}$ (e deste modo a distribuição conjunta) são atualizadas dia a dia. Para cada tempo t foram simuladas 5000 observações. Em um computador Athlon XP 950 Mhz com 128 MB de RAM, utilizando o Matlab 12.1, o tempo para estimação de um modelo (por exemplo SJC + GARCH-t), que deve ser feito a cada 50 observações, é de 27,78 segundos e o tempo da atualização de 1 dia para estimação do VaR é de 7,62 segundos.

Para verificar a estabilidade dos parâmetros dos acoplamentos no tempo, temos na Figura 8.14 os parâmetros estimados pelos métodos IFM e medida de dependência a cada passo. Pelos parâmetros τ_L , θ e R_{12} , respectivamente dos acoplamentos SJC, Plackett e t-Student, vemos que a dependência entre os ativos é alta no início da amostra, vai caindo com o passar do tempo, até que se estabiliza entre 1999 e 2002, quando então volta a crescer.

Em Aguilar e West (2000), é comentado que o Iene não apresenta uma grande dependência na volatilidade em relação às outras moedas consideradas (Marco Alemão, Libra Esterlina, Franco Francês, Dólar Canadense e Pesseta Espanhola), especialmente após 1997. Esta dependência na volatilidade está de acordo com os resultados de dependência nos retornos encontrados na Figura 8.14, em que a dependência cai após 1997.

Na Figura 8.15 temos os parâmetros estimados para as marginais GARCH. Para os parâmetros dos modelos GARCH, verificamos uma razoável alteração no tempo, sendo que os parâmetros do modelo GARCH-t são menos suscetíveis à variações do que os parâmetros do modelo GARCH-N, provavelmente devido ao fato de que o modelo GARCH-t dá menos peso aos valores nas caudas da

distribuição. Isto é notado em especial para os parâmetros α_1 e β_1 do modelo GARCH entre 1999 e 2001.

Verificamos então a proporção de dados em que a perda do portfolio excede o VaR, isto é, os tempos t em que $x_{p,t} < \widehat{VaR}(\alpha)_t$, para $t = 751, \dots, 3502$ para os três acoplamentos utilizados e para $\alpha = 0,005$, $\alpha = 0,01$ e $\alpha = 0,05$. Os resultados estão na Tabela 8.6.

	Método IFM			Med. dep.
Copula	GARCH-N	GARCH-t	GARCH-E	
$\alpha = 0,05$ (137)*				
SJC	0,0472 (130)	0,0469 (129)	0,0440 (121)	
Plackett	0,0450 (124)	0,0534 (147)	0,0461 (127)	0,0483 (133)
t-Student	0,0472 (130)	0,0509 (140)	0,0483 (133)	0,0465 (128)
$\alpha = 0,01$ (27)*				
SJC	0,0229 (63)	0,0142 (39)	0,0134 (37)	
Plackett	0,0193 (53)	0,0156 (43)	0,0131 (36)	0,0203 (56)
t-Student	0,0182 (50)	0,0116 (32)	0,0094 (26)	0,0196 (54)
$\alpha = 0,005$ (14)*				
SJC	0,0178 (49)	0,0051 (14)	0,0073 (20)	
Plackett	0,0124 (34)	0,0073 (20)	0,0065 (18)	0,0134 (37)
t-Student	0,0084 (23)	0,0054 (15)	0,0058 (18)	0,0098 (27)

Tabela 8.6: Proporção de observações (e número de observações), para $t = 750$ a 3502, em que a perda do portfolio excede o VaR para $\alpha = 0,05$, $\alpha = 0,01$ e $\alpha = 0,005$. * = Número esperado de observações.

O método de simulação indica que o acoplamento de Plackett subestima o VaR, para qualquer modelo para as distribuições marginais. Para $\alpha = 0,05$ e $\alpha = 0,01$ o melhor modelo é o acoplamento t-Student com marginais GARCH-t e GARCH-E respectivamente, seguido pelo acoplamento SJC + GARCH-t. Já para $\alpha = 0,005$ a situação entre estes dois modelos se inverte, sendo o melhor modelo acoplamento SJC + GARCH-t, seguido pelo modelo acoplamento t-Student + GARCH-t. Estes resultados coincidem com os obtidos nos ajustes do acoplamentos (Tabela 8.5), em que o melhor ajustado foi obtido pelo

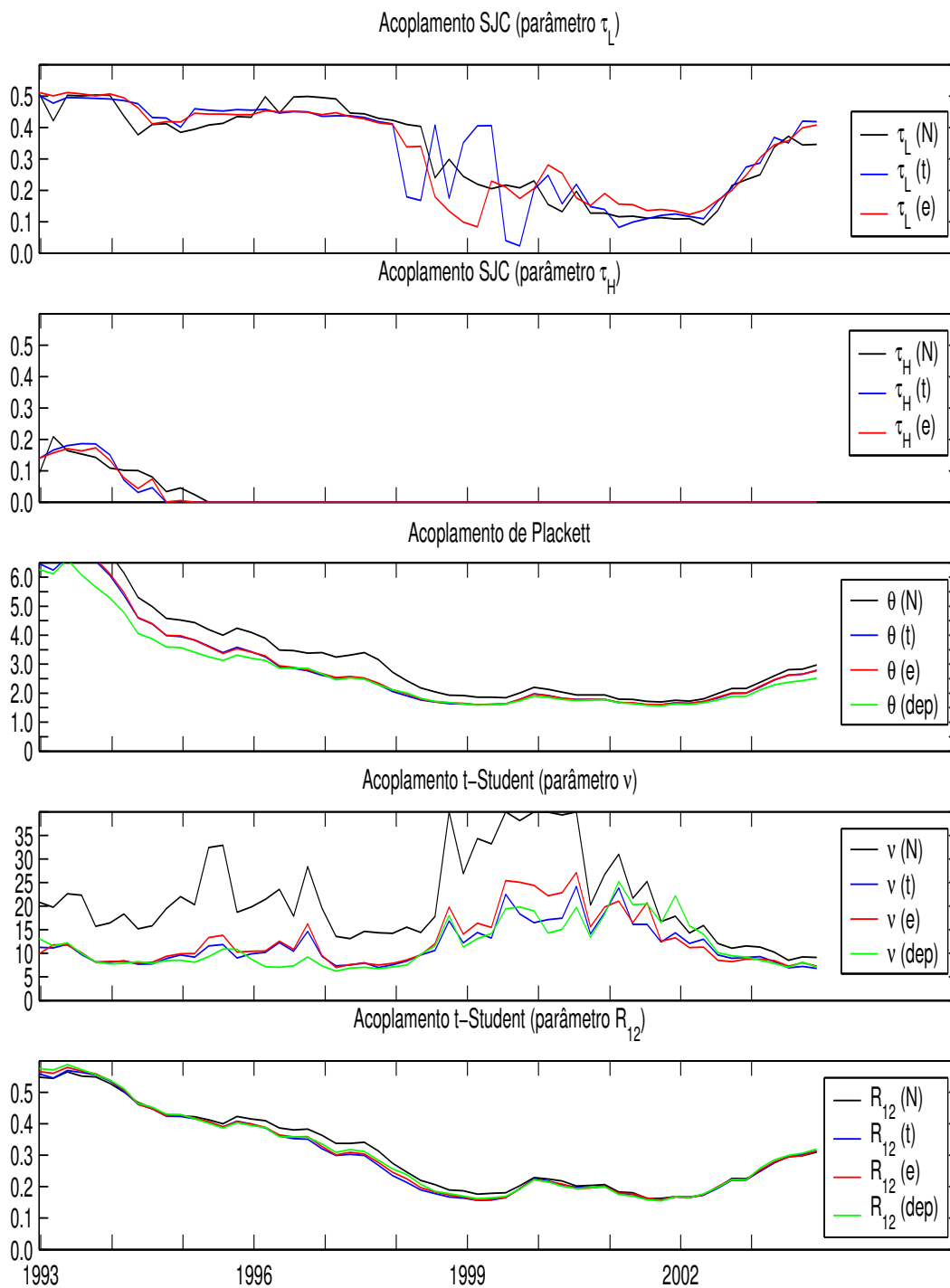


Figura 8.14: Parâmetros estimados das funções de acoplamento para simulação.

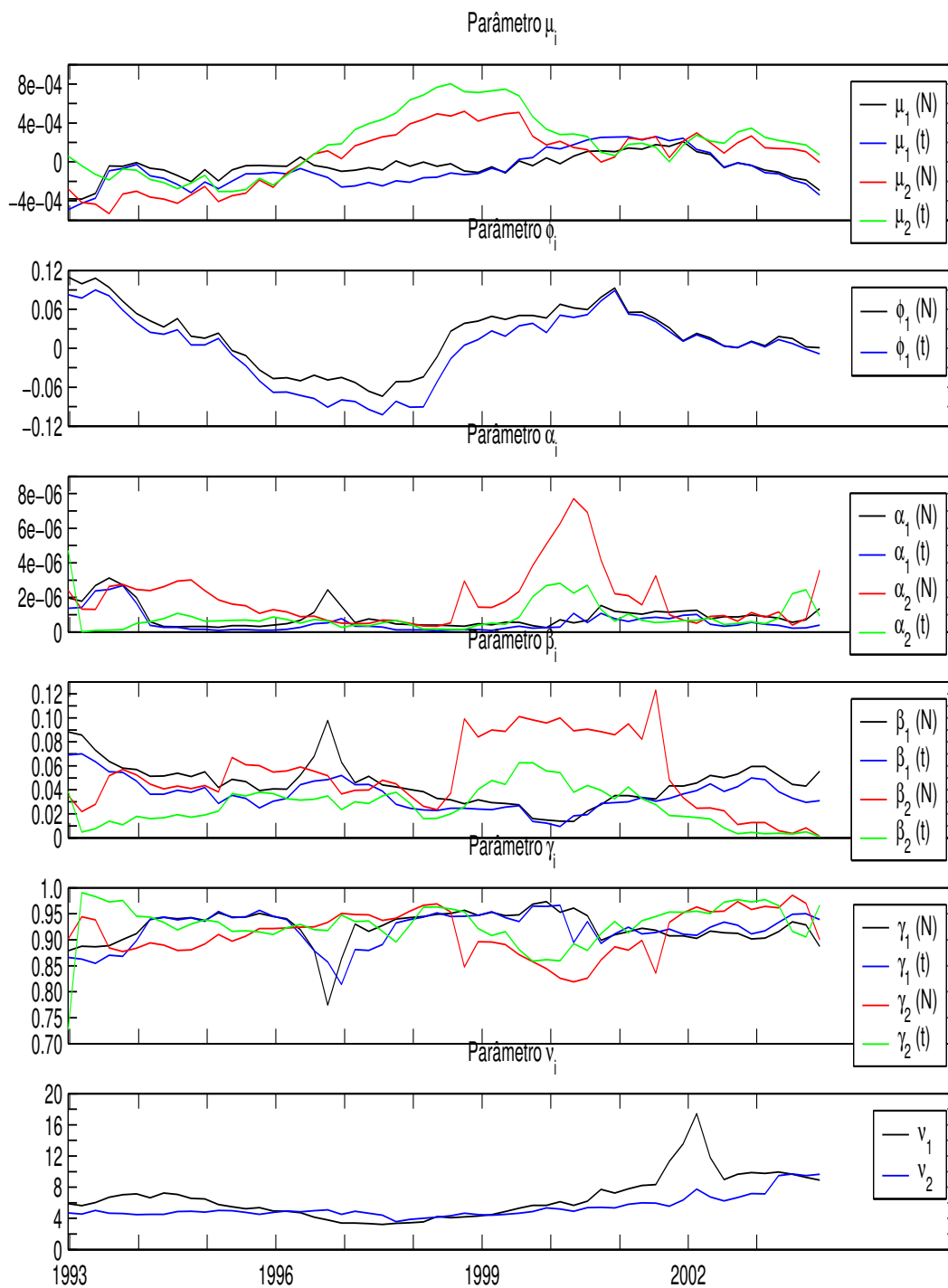


Figura 8.15: Parâmetros estimados das marginais GARCH para simulação.

acoplamento t-Student $\alpha = 0,05$ e $\alpha = 0,01$ e pelo acoplamento SJC para $\alpha = 0,005$. Isto era esperado uma vez que esperamos em geral que o acoplamento estimado com a menor distância em relação ao acoplamento empírico, especialmente nas caudas, produza a melhor estimação do VaR.

Como exemplo temos na Figura 8.16 o gráfico dos retornos do portfolio e do VaR estimado com o acoplamento t-Student e marginais GARCH-t, para $\alpha = 0,05$, $\alpha = 0,01$ e $\alpha = 0,005$. Na próxima seção iremos também avaliar estatisticamente a qualidade da estimativa do VaR através do teste de Christoffersen (1998).

8.6 Comparação com outros modelos

Nesta seção iremos comparar o resultado obtido pelos melhores modelos utilizando-se funções de acoplamentos (acoplamentos SJC e t-Student + GARCH-t) com diversos outros modelos univariados e bivariados (na estimação do VaR). No caso univariado, aplicado diretamente no retorno do portfolio, iremos utilizar os modelos GARCH-N e GARCH-t e os métodos EWMA e de estimação por simulações históricas. No caso bivariado iremos considerar os modelos GARCH (modelos BEKK e DCC) e o método EWMA.

Podemos modelar diretamente a distribuição dos retornos do portfolio, no nosso caso aproximado por $X_{p,t} = (2/3)X_{1,t} + (1/3)X_{2,t}$. Por exemplo, podemos aplicar diretamente o modelo GARCH na série de retornos do portfolio. Selecionando o melhor modelo pelo critério de AIC, temos que o melhor modelo é um GARCH(1,1) dado por :

$$\begin{cases} x_{p,t} = \mu_t + \sigma_t \epsilon_t ; \\ \sigma_t^2 = \alpha + \beta \epsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 . \end{cases} \quad (8.7)$$

Iremos considerar para as inovações as distribuições Normal e t-Student.

O método EWMA (Exponential Weighted Moving Average) é um dos mais preferidos pelos utilizadores da norma Riskmetrics graças a sua praticidade. Seja $\sigma_{p,t}^2$ a variância do retorno do portfolio no tempo t . A variância estimada

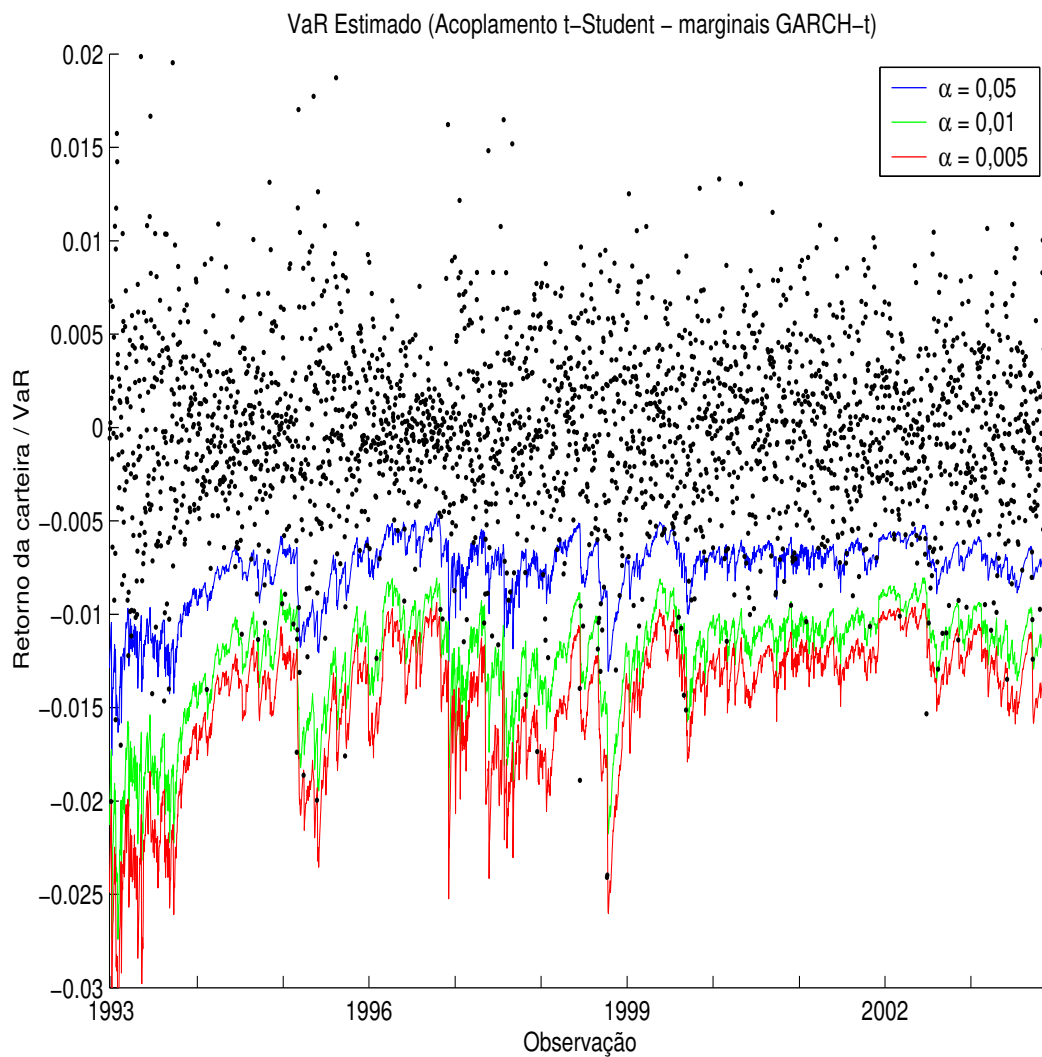


Figura 8.16: Retornos do portfolio e VaR estimado (acoplamento t-Student e marginais GARCH-t)

é dada por

$$\sigma_{p,t/t-1}^2 = (1 - \lambda)x_{p,t-1}^2 + \lambda\sigma_{p,t-1/t-2}^2.$$

Considerando a distribuição normal, temos : $x_{p,t} \sim \mathcal{N}(0, \sigma_{p,t}^2)$. O parâmetro λ ótimo é reestimado a cada tempo t , obtendo-se o λ que minimiza o erro $\sqrt{\sum_{i=t_0}^{t-1} (\sigma_{(i+1)/i}^2 - x_{i+1}^2)^2}$, onde $\sigma_{(i+1)/i}^2$ denota o valor alisado para variância do tempo $i + 1$ considerando-se a informação até o tempo i e x_{i+1}^2 denota o retorno ao quadrado realmente ocorrido no tempo $i + 1$. Na estimação do λ foi considerado $t_0 = 50$.

O último modelo considerado é o de Simulações Históricas. Seja um tempo t_0 fixado. A estimativa do VaR para o tempo $t_0 + 1$ com $1 - \alpha$ de confiança é dada pela αt_0 -ésima estatística de ordem da amostra de retornos do portfolio entre os tempos $1, 2, \dots, t_0$.

A primeira classe de modelos GARCH bivariados utilizada é a de BEKK (veja Engle e Kroner (1995)). Sejam $x_t = \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix}$, $\epsilon_t = \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}$, $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$, $C = \begin{pmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{pmatrix}$ e seja $\Sigma_t = \begin{pmatrix} \sigma_{1,t}^2 & \sigma_{12,t} \\ \sigma_{12,t} & \sigma_{2,t}^2 \end{pmatrix} = \text{Cov}[\epsilon_t | \mathfrak{S}_{t-1}]$. Temos então que

$$\begin{cases} x_t = \mu + \epsilon_t ; \\ \Sigma_t = C'C + A'\epsilon_{t-1}\epsilon_{t-1}'A + B'\Sigma_{t-1}B. \end{cases} \quad (8.8)$$

e assumimos que o vetor de inovações ϵ_t tem distribuição condicional bivariada $\mathcal{N}(0, \Sigma_t)$. O programa utilizado para a estimação foi o SAS 8.2, proc VARMAX.

A segunda classe de modelos GARCH bivariados utilizada é a DCC, proposta por Engle (2002). Sejam x_i os retornos com média zero, $i = 1, \dots, n$. As correlações e variâncias condicionais são definidas por

$$\sigma_{i,j,t} = E[x_{i,t}x_{j,t} | \mathfrak{S}_{t-1}] / \sqrt{E[x_{i,t}^2 | \mathfrak{S}_{t-1}] E[x_{j,t}^2 | \mathfrak{S}_{t-1}]}. \text{ Seja } \sigma_{i,t}^2 = E[x_{i,t}^2 | \mathfrak{S}_{t-1}]$$

e $\eta_{i,t} = x_{i,t}/\sigma_{i,t}$. Então a correlação pode ser escrita como $\sigma_{i,j,t} = E[\eta_{i,t} \eta_{j,t} | \mathfrak{S}_{t-1}]$. Engle (2002) sugere estimarmos os processos GARCH

$$q_{i,j,t} = \bar{\sigma}_{i,j} + \alpha(\eta_{i,t-1} \eta_{j,t-1} - \bar{\sigma}_{i,j}) + \beta(q_{i,j,t-1} - \bar{\sigma}_{i,j}) \quad (8.9)$$

para $i,j = 1, \dots, n$ e obtermos $\sigma_{i,j,t} = q_{i,j,t}/\sqrt{q_{i,i,t} q_{j,j,t}}$. Podemos interpretar $\bar{\sigma}_{i,j}$ como a correlação não-condicional entre as inovações dos modelos GARCH univariados $\eta_{i,t}$ e $\eta_{j,t}$. Deste modo, as correlações e variâncias são modeladas individualmente como processos GARCH com parâmetros comuns α e β , e com esperanças não-condicionais distintas $\bar{\sigma}_{i,j}$.

Podemos utilizar o método de alisamento exponencial também para o caso bivariado. Neste caso as variâncias e a covariância estimada são dadas por

$$\begin{cases} \sigma_{1,t/t-1}^2 = (1 - \lambda_1)x_{1,t-1}^2 + \lambda_1\sigma_{1,t-1/t-2}^2; \\ \sigma_{2,t/t-1}^2 = (1 - \lambda_2)x_{2,t-1}^2 + \lambda_2\sigma_{2,t-1/t-2}^2; \\ \sigma_{12,t/t-1} = (1 - \lambda_{12})x_{1,t-1}x_{2,t-1} + \lambda_{12}\sigma_{12,t-1/t-2}. \end{cases} \quad (8.10)$$

onde os parâmetros ótimos λ_1 , λ_2 e λ_{12} são obtidos a cada tempo t minimizando-se as quantidades $\sqrt{\sum_{i=t_0}^{t-1} (\sigma_{1,(i+1)/i}^2 - x_{1,i+1}^2)^2}$, $\sqrt{\sum_{i=t_0}^{t-1} (\sigma_{2,(i+1)/i}^2 - x_{2,i+1}^2)^2}$ e $\sqrt{\sum_{i=t_0}^{t-1} (\sigma_{12,(i+1)/i} - x_{1,i+1}x_{2,i+1})^2}$ respectivamente. Considerando a distribuição normal bivariada, nós temos então :

$$\begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{1,t}^2 & \sigma_{12,t} \\ \sigma_{12,t} & \sigma_{2,t}^2 \end{pmatrix} \right].$$

Consideramos também um alisamento exponencial para a média dos retornos, mas os resultados obtidos foram muito próximos, então mantivemos o modelo com vetor de médias igual a zero.

Como no caso das funções de acoplamento, iremos começar a simulação com dados de $t = 1$ a $t = 750$. No caso dos modelos GARCH-bivariado e univariado iremos manter os parâmetros das distribuições marginais e do acoplamento fixos a cada 50 retornos. Apesar dos parâmetros serem estimados a cada 50 observações, as distribuições bivariada ou univariada são atualizadas

dia a dia. Já nos modelos EWMA e de Simulações Históricas, os valores são atualizados dia a dia a partir de $t = 750$.

Estimamos então o VaR para todos os modelos e comparamos os resultados obtidos com o melhor modelo que utilizamos função de acoplamento (acoplamento SJC com marginais GARCH-t). Uma medida descritiva é dada pela proporção de valores em que a perda do portfolio superar o VaR estimado. Porém esta medida não é suficiente, uma vez que desejaríamos que a probabilidade de falha (perda superando o VaR) fosse constante ao longo do tempo. Poderíamos utilizar, por exemplo, o teste de Kupiec, descrito em Jorion (1997). O problema principal deste é que sua natureza é não condicional, enquanto que os intervalos de confiança são estimados condicionalmente. Christoffersen (1998) desenvolve um método de avaliação condicional para intervalos de precisão. Iremos descrever brevemente as hipóteses deste teste.

Seja $\{r_t\}$ uma série temporal e seja $\{(L_{t/t-1}(\alpha), U_{t/t-1}(\alpha))\}_{1 \leq t \leq T}$ a correspondente sequência de intervalos de previsão fora da amostra, onde $L_{t/t-1}$ e $U_{t/t-1}$ são os limites inferior e superior do intervalo de precisão ex-ante para o período t feito no período $t-1$ ao nível de confiança $1-\alpha$.

$$I_t = \begin{cases} 1 & \text{se } r_t \in (L_{t/t-1}(\alpha), U_{t/t-1}(\alpha)); \\ 0 & \text{se } r_t \notin (L_{t/t-1}(\alpha), U_{t/t-1}(\alpha)). \end{cases}$$

Diz-se que a sequência de intervalos de precisão, $\{(L_{t/t-1}(\alpha), U_{t/t-1}(\alpha))\}_{1 \leq t \leq T}$, é eficiente com respeito ao conjunto de informação Ψ_{t-1} , se $E(I_t/\Psi_{t-1}) = 1-\alpha$, para todo o t . Christoffersen (1998) mostra que testar $E(I_t/\Psi_{t-1}) = 1-\alpha = E(I_t/I_{t-1}, \dots, I_1) = 1-\alpha$ para todo t , é equivalente a testar que a hipótese nula de que a sequência $\{I_t\}_{1 \leq t \leq T}$ é idêntica e independentemente distribuída como Bernoulli, com parâmetro $1-\alpha$, isto é $\{I_t\} \sim b(1-\alpha)$ i.i.d, $\forall t$. Para maiores detalhes sobre a estatística do teste e sua distribuição, consulte Christoffersen (1998).

Na Tabela 8.7 temos a proporção de falhas e, entre parênteses, o p-valor do teste de Christoffersen, considerando os diversos modelos ajustados e três valores de α . Podemos ver que para $\alpha = 0,05$ todos os modelos apresentaram proporções razoavelmente próximas de 0,05 exceto os modelos GARCH

Bivariado (BEKK) e de Simulações Históricas. O melhor resultado foi obtido pelo modelo GARCH bivariado (DCC). Já para $\alpha = 0,01$ os modelos que se saíram razoavelmente bem foram o acoplamento t-Student + GARCH-t e o GARCH-t aplicado diretamente no portfólio. Para $\alpha = 0,005$ os melhores modelos foram as funções de acoplamento (t-Student e SJC) com marginais GARCH-t e o GARCH-t no portfólio. Não rejeitamos a hipótese nula do teste de Christoffersen apenas para o acoplamento SJC e para $\alpha = 0,005$. Para o acoplamento t-Student e $\alpha = 0,01$ temos uma evidência moderada contra a hipótese nula.

De uma maneira resumida, vemos que as funções de acoplamento e o modelo GARCH-t aplicado diretamente no portfólio foram os melhores modelos para a estimação do VaR. É importante lembrar, que como observado em Fortin e Kuzmics (2002), em certos problemas como na otimização de portfólios é necessário estudar a estrutura de dependência entre os ativos, de modo que modelos aplicados diretamente no portfólio não são adequados.

Como exemplo temos na Figura 8.17 temos o VaR estimado para $\alpha = 0,05$, $\alpha = 0,01$ e $\alpha = 0,005$ utilizando-se o modelo Acoplamento t-Student + GARCH-t e o modelo GARCH-t aplicado diretamente no portfólio. Os resultados obtidos são bastante próximos.

Modelo	$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,005$
Acoplamento SJC + GARCH-t	0,0469 (0,0097)	0,0142 (0,0001)	0,0051 (0,1689)
Acoplamento t-Student + GARCH-t	0,0509 (0,0033)	0,0116 (0,0149)	0,0055 (0,0083)
GARCH Bivariado (BEKK)	0,0679 (0,0001)	0,0265 (0,0000)	0,0189 (0,0000)
GARCH Bivariado (DCC)	0,0494 (0,0017)	0,0203 (0,0000)	0,0116 (0,0000)
EWMA (Bivariado)	0,0461 (0,0022)	0,0174 (0,0000)	0,0131 (0,0000)
GARCH-N (Portfolio)	0,0443 (0,0005)	0,0171 (0,0001)	0,0109 (0,0000)
GARCH-t (Portfolio)	0,0472 (0,0015)	0,0116 (0,0013)	0,0054 (0,0083)
EWMA (Portfolio)	0,0548 (0,0200)	0,0196 (0,0000)	0,0142 (0,0000)
Simulações Históricas (Portfolio)	0,0345 (0,0000)	0,0044 (0,0000)	0,0025 (0,0053)

Tabela 8.7: Proporção de dados em que a perda do portfolio excede o VaR $100(1-\alpha)\%$ e P-valor do teste de Christoffersen (entre parênteses) para diversos valores de α - Libra Esterlina e Iene.

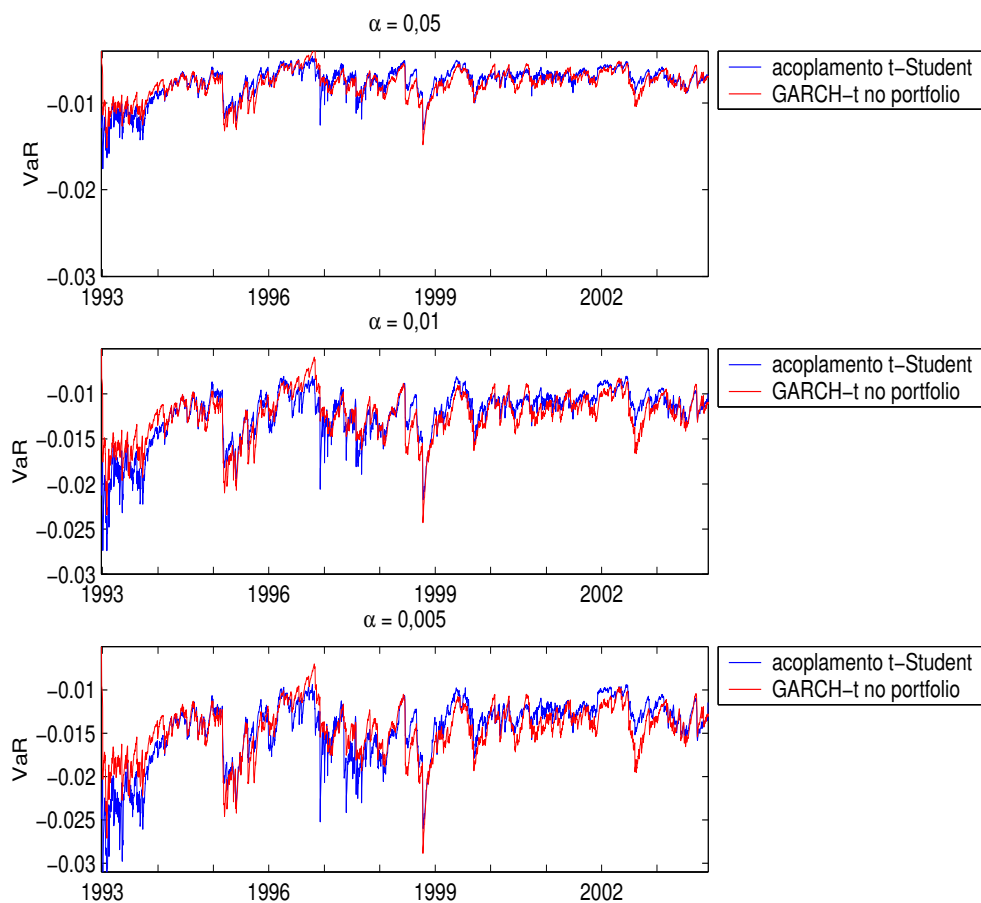


Figura 8.17: VaR estimado (acoplamento SJC e marginais GARCH-t (tracejado) e GARCH-t (pontilhado) no portfolio) com $\alpha = 0,01$.

No Apêndice A iremos repetir os passos da análise para as séries de Libra e Iene para um novo conjunto de dados, que consiste dos valores dos índices NASDAQ e S&P500, durante o período de 02 de Janeiro de 1992 a 01 de Outubro de 2003, num total de 2972 observações. Os parâmetros estimados para as funções de acoplamento estão na Tabela A.3. Os resultados obtidos estão resumidos na Tabela 8.8. Podemos ver que para $\alpha = 0,05$ o modelo EWMA aplicado diretamente no portfolio apresentou os melhores resultados, seguido pela função de acoplamento SJC com marginais GARCH-E. Já para $\alpha = 0,01$ e $\alpha = 0,005$ o melhor modelo foi a função de acoplamento SJC. No caso da função de acoplamento não rejeitamos a hipótese nula do teste de Christoffersen para nenhum dos valores de α considerado.

Modelo	$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,005$
Acoplamento SJC + GARCH-E	0,0558 (0,1342)	0,0104 (0,7251)	0,0041 (0,9063)
GARCH Bivariado (BEKK)	0,0819 (0,0000)	0,0338 (0,0000)	0,0248 (0,0000)
GARCH Bivariado (DCC)	0,0432 (0,0649)	0,0140 (0,1549)	0,0113 (0,0012)
EWMA (Bivariado)	0,0387 (0,0368)	0,0144 (0,7836)	0,0104 (0,1430)
GARCH-N (Portfolio)	0,0666 (0,0026)	0,0207 (0,0000)	0,0144 (0,0000)
GARCH-t (Portfolio)	0,0693 (0,0003)	0,0131 (0,2620)	0,0104 (0,0004)
EWMA (Portfolio)	0,0527 (0,2056)	0,0135 (0,7548)	0,0099 (0,0126)
Simulações Históricas (Portfolio)	0,1220 (0,0000)	0,0293 (0,0000)	0,0144 (0,0000)

Tabela 8.8: Proporção de dados em que a perda do portfolio excede o VaR e P-valor do teste de Christoffersen (entre parênteses) para diversos valores de α - Índices Nasdaq e S&P500.

9 *Conclusão*

Esta dissertação de mestrado mostrou como a teoria de acoplamentos condicionais pode ser bastante poderosa na estimação do Valor em Risco. Foram analisados dois conjunto de dados, o primeiro composto pelas taxas de câmbio Libra/Dólar e Iene/Dólar, e o segundo pelos índices de ações Nasdaq e S&P500. Nós utilizamos diversas funções de acoplamento e distribuições marginais para as inovações dos modelos GARCH. O objetivo do trabalho foi comparar diferentes funções de acoplamento. Além dos métodos tradicionais de estimação, nós utilizamos medidas de dependência como Rho de Spearman e Tau de Kendall. Diversas outras técnicas para estimação do VaR foram utilizadas com o objetivo de comparação, como o modelo GARCH-bivariado, alisamento exponencial (EWMA), Simulações Históricas e GARCH aplicado diretamente no portfolio.

No caso do portfolio composto pelas taxas de câmbio Libra/Dólar e Iene/Dólar, os acoplamentos simetrizado de Joe Clayton e t-Student com marginais GARCH-t produziram o melhor resultado e as estimativas mais confiáveis para o VaR. Entre os modelos de comparação, o único que apresentou performance semelhante foi o modelo GARCH-t aplicado diretamente no portfolio. No caso do portfólio composto pelos índices NASDAQ e S&P500, os melhores resultados foram obtidos pelo acoplamento simetrizado de Joe Clayton com marginais GARCH-E.

A pesquisa pode ser estendida para permitir este tipo de aplicação para maiores dimensões.

APÊNDICE A – Aplicação 2

Iremos agora repetir os passos da análise para as séries de Libra e Iene para um novo conjunto de dados, que consiste dos valores dos índices NASDAQ e S&P500, durante o período de 02 de Janeiro de 1992 a 01 de Outubro de 2003, num total de 2972 observações. O índice NASDAQ Composite compreende uma cesta de mais de 5000 ações negociadas na bolsa eletrônica NASDAQ, que corresponde por cerca de 54% do volume negociado nas bolsas norte-americanas. O índice S&P500 é composto por 500 ações negociadas na NYSE (*New York Stock Exchange*), AMEX (*American Stock Exchange*) e NASDAQ (*National Market System*).

A notação será a mesma do conjunto de dados anterior.

A.1 Análise exploratória de dados

Ambas as séries possuem pouca assimetria. A distribuição dos retornos do NASDAQ possui uma curtose maior do que a dos retornos do S&P500, sendo ambas maiores do que o caso normal (excesso de curtose positivo). Isto indica a existência do fato estilizado conhecido como caudas pesadas.

Temos na Figura A.1 os gráficos da série de retornos e de retornos absolutos. Podemos notar no gráfico de retornos absolutos a existência dos chamados agrupamentos de volatilidade, com regiões de alta e baixa volatilidade, também um fato estilizado citado no Capítulo 6.

Estatística	NASDAQ	S&P500
Média	0,00038	0,00030
Média (anualizada)	10,141%	7,857%
Desvio Padrão	0,01694	0,01076
Mínimo	-0,10168	-0,07113
Mediana	0,00122	0,00028
Máximo	0,13255	0,05574
Excesso de Curtose	4,91481	3,78088
Assimetria	0,01490	-0,10267

Tabela A.1: Estatísticas descritivas dos retornos de NASDAQ e S&P500.

Ambas séries de retornos dos índices não apresentam autocorrelações ou autocorrelações parciais altamente significantes.

A.2 Modelagem das distribuições marginais

Ajustamos como modelos tentativos iniciais processos com erros GARCH(1,1) e média constante para ambas as séries. Utilizamos inovações com distribuição normal, t-Student e distribuição empírica. Outros modelos foram tentados, como AR(1), MA(1) e ARMA(1,1), todos com erros GARCH(1,1), mas os resíduos obtidos não foram melhores, e os parâmetros adicionados não foram significantes. Assim o modelo é dado por:

$$\begin{aligned}
 X_{i,t} &= \mu_i + \phi_i X_{i,t-1} + \epsilon_{i,t}, \\
 \epsilon_{i,t} &= \sigma_{i,t} \eta_{i,t}, \\
 \sigma_{i,t}^2 &= \alpha_i + \beta_i \epsilon_{i,t-1}^2 + \gamma_i \sigma_{i,t-1}^2,
 \end{aligned} \tag{A.1}$$

onde $\{\eta_{i,t}\}$ são processos ruído branco e $i=1, 2$.

Consideramos as três distribuições para as inovações consideradas anteriormente : Normal, t-Student e empírica.

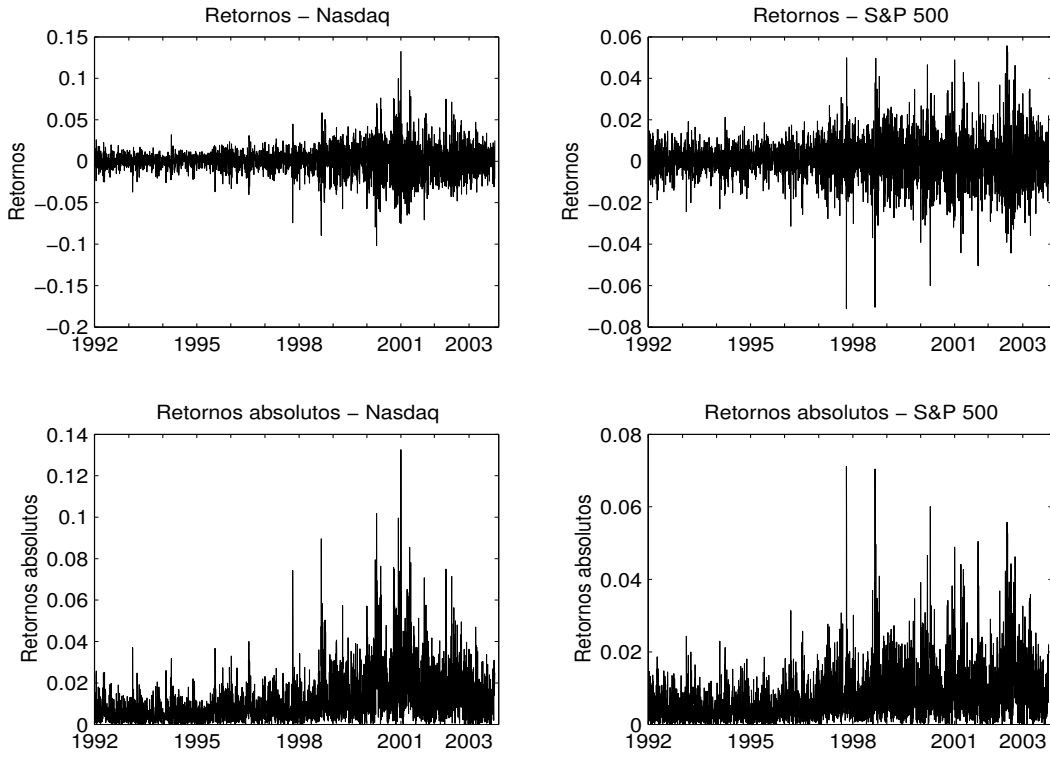


Figura A.1: Gráficos de Retornos e Ret. Absolutos - NASDAQ e S&P500.

As estimativas dos parâmetros dos modelos pelo método de máxima verossimilhança são dadas na Tabela A.2.

Vamos analisar as séries $\hat{\epsilon}_{1,t}$, $\hat{\epsilon}_{2,t}$, $\hat{\eta}_{1,t}$ e $\hat{\eta}_{2,t}$ como no conjunto de dados anterior. Os correlogramas não serão apresentados aqui por questão de espaço. Os resultados obtidos mostram que não existem autocorrelações significantes nas séries $\hat{\epsilon}_{1,t}$ e $\hat{\epsilon}_{2,t}$. Porém nas séries $\hat{\epsilon}_{1,t}^2$ e $\hat{\epsilon}_{2,t}^2$ existem diversas autocorrelações significantes mostrando a necessidade do modelo GARCH.

Já para as séries $\hat{\eta}_{1,t}$, $\hat{\eta}_{2,t}$, $\hat{\eta}_{1,t}^2$ e $\hat{\eta}_{2,t}^2$ não existem autocorrelações significantes ao nível de significância de 5%, exceto : na série $\hat{\eta}_{1,t}$ a estatística t do teste de autocorrelação nula de ordem 12 vale 2,87 (p-valor de 0,0041) e de ordem 13 vale 3,17 (p-valor de 0,0015) (modelo GARCH-N). E na série $\hat{\eta}_{1,t}^2$ a estatística t do teste de autocorrelação de ordem 33 vale 2,94 (p-valor de

Parâmetro	GARCH-N	GARCH-t
μ_1	$8,25 \times 10^{-4}$ ($1,85 \times 10^{-4}$)	$9,43 \times 10^{-4}$ ($1,79 \times 10^{-4}$)
ϕ_1	0,0846 (0,0199)	0,0884 (0,0188)
α_1	$1,44 \times 10^{-6}$ ($4,06 \times 10^{-7}$)	$6,13 \times 10^{-7}$ ($2,69 \times 10^{-7}$)
β_1	0,1022 (0,0145)	0,0652 (0,0111)
γ_1	0,8958 (0,0140)	0,9180 (0,0134)
ν_1		9,6859 (0,0773)
AIC_1	$-1,738 \times 10^4$	$-1,745 \times 10^4$
μ_2	$5,18 \times 10^{-4}$ ($1,42 \times 10^{-4}$)	$5,66 \times 10^{-4}$ ($1,35 \times 10^{-4}$)
ϕ_2	0,0156 (0,0194)	-0,0021 (0,0188)
α_2	$5,76 \times 10^{-6}$ ($1,87 \times 10^{-7}$)	$2,55 \times 10^{-7}$ ($1,20 \times 10^{-7}$)
β_2	0,0701 (0,0097)	0,0430 (0,0069)
γ_2	0,9273 (0,0097)	0,9397 (0,0093)
ν_2		7,2459 (0,1516)
AIC_2	$-1,939 \times 10^4$	$-1,950 \times 10^4$

Tabela A.2: Parâmetros estimados dos modelos das distribuições marginais e erro padrão das estimativas.

0,0033), de modo que os modelos explicam adequadamente a heteroscedasticidade da série original.

Observamos os gráficos bivariados das inovações ajustadas padronizadas para ambos os modelos na A.2. Interessante notar a existência de valores extremos na cauda das perdas.

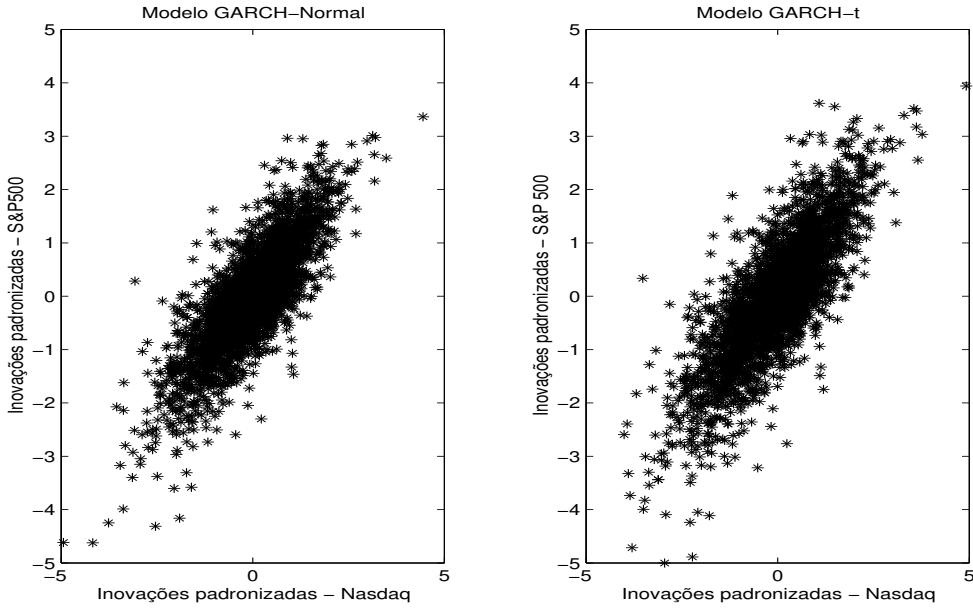


Figura A.2: Gráficos das inovações ajustadas padronizadas (NASDAQ e S&P500).

Para verificar o fato de dependência assimétrica, utilizaremos a correlação condicional, como definida em (3.1), para $\xi \in \{-3, -2, -1, 1, 2, 3\}$.

Apesar de existirem poucos retornos nos valores extremos, temos grande evidência de que a correlação entre os retornos aumenta tanto para as grandes perdas como os grandes ganhos. Assim temos evidências de um desvio da distribuição normal bivariada, em que estes valores deveriam tender a zero. Além disto, notamos que as correlações para as grandes perdas são levemente maiores do que para os grandes ganhos, indicando um grau leve de dependência assimétrica.

ξ	GARCH-N	GARCH-t
Correlação comum	0.8205	0.8193
-3	0,8925 (9)	0,9220 (10)
-2.5	0,8740 (20)	0,8777 (22)
-2	0,8122 (41)	0,8452 (36)
-1	0,7026 (292)	0,7189 (290)
1	0,5559 (258)	0,5268 (259)
2	0,7944 (26)	0,7876 (25)
2.5	0,6404 (8)	0,8233 (8)
3	— (2)	— (2)

Tabela A.3: Correlações condicionais estimadas para diversos ξ .

A.3 Modelagem da dependência entre as séries

Iremos modelar a dependência através das três funções de acoplamento utilizadas anteriormente.

Temos então a estimativa por máxima verossimilhança (método IFM) dos parâmetros dos acoplamentos na Tabela A.3

Temos na Tabela A.3 as estimativas de medidas de dependência e os parâmetros das funções de acoplamento estimados através destas medidas.

Para verificar o ajuste dos acoplamentos estimados podemos calcular a distância quadrática em relação ao acoplamento empírico, como dito em (5.26). Como estamos interessados principalmente no ajuste na cauda inferior para efeitos de estimação do VaR, iremos calcular esta distância nas regiões $[0; 0, 2]^2$, $[0; 0, 1]^2$ e $[0; 0, 05]^2$

Podemos ver pela Tabela A.6 que o acoplamento t-Student foi o que melhor se ajustou aos dados, e em especial na cauda inferior (entre 0 e 0,05), que mais nos interessa para a estimativa do VaR. Os resultados obtidos para o acoplamento simetrizado de Joe Clayton também foram bastante razoáveis,

Acoplamento	Parâmetro	GARCH-N	GARCH-t	GARCH-E
SJC	τ_L	0,6706 (0,0092)	0,6887 (0,0080)	0,7124 (0,0072)
	τ_U	0,5937 (0,0151)	0,5858 (0,0152)	0,5286 (0,0188)
	(AIC)	-3,9918x10 ⁴	-4,0208x10 ⁴	
Plackett	θ	21,6216 (0,9248)	20,0555 (0,8699)	19,9550 (0,8630)
	(AIC)	-3,9939x10 ⁴	-4,0047x10 ⁴	
t-Student	ν	9,5305 (0,6648)	6,9129 (0,5210)	7,5432 (0,5865)
	R_{12}	0,8248 (0,0052)	0,8174 (0,0063)	0,8175 (0,0056)
	(AIC)	-4,0117x10 ⁴	-4,0291x10 ⁴	

Tabela A.4: Parâmetros estimados das funções de acoplamento (Desvios padrão em parênteses).

embora piores para a região $[0; 0,05]^2$. O acoplamento de Plackett apresentou o pior ajuste. Provavelmente o bom ajuste dado pelo acoplamento t-Student, que é elíptico, deve-se ao fato de que a dependência assimétrica encontrada nos dados não é tão alta, pois os coeficientes de dependência nas caudas τ_L e τ_U são próximos.

Acoplamento	Medida de dependência estimada	Parâmetro estimado
Plackett	$\rho_S = 0,8020$	$\theta = 24,6302$
t-Student	$\tau_k = 0,6200$	$R_{12} = 0,8270$
		$\nu = 6,0826$

Tabela A.5: Parâmetros estimados das funções de acoplamento por medidas de dependência.

Modelo	$\delta = 0,20$	$\delta = 0,10$	$\delta = 0,05$
SJC + GARCH-N	4,8896	0,7971	0,4328
SJC + GARCH-t	3,9492	0,8381	0,5181
SJC + GARCH-E	2,8348	1,0524	0,6273
Plackett + GARCH-N	6,7868	2,1017	0,5102
Plackett + GARCH-t	7,4747	2,2761	0,5530
Plackett + GARCH-E	7,5159	2,2880	0,5559
Plackett (usando ρ_S)	5,8743	1,8040	0,4358
t-Student + GARCH-N	3,6668	0,6846	0,2483
t-Student + GARCH-t	3,8616	0,7011	0,2553
t-Student + GARCH-E	3,8996	0,7102	0,2467
t-Student (usando τ_k)	3,3552	0,6440	0,3022

Tabela A.6: Distância quadrática entre o acoplamento estimado e o acoplamento empírico na cauda inferior (quadrado $[0, \delta]^2$).

A.4 Simulação do Valor em Risco

Iremos considerar agora que temos uma carteira composta pelos dois índices Nasdaq Composite e S&P500 em proporções iguais dos dois índices, de modo que temos aproximadamente $X_{p,t} = 0,5X_{1,t} + 0,5X_{2,t}$.

Os parâmetros das distribuições marginais e da função de acoplamento irão ser estimados com dados de $t = 1$ a $t = 750$, mantendo os parâmetros das distribuições marginais e do acoplamento fixos a cada 50 retornos.

Verificamos então a proporção de dados em que a perda do portfolio excede o VaR, isto é, os tempos t em que $x_{p,t} < VaR(\alpha)_t$, para $t = 751, \dots, 2971$ para os três acoplamentos utilizados e para $\alpha = 0,005$, $\alpha = 0,01$ e $\alpha = 0,05$. Os resultados estão na Tabela A.7. Os melhores resultados são obtidos para o acoplamento SJC com marginais GARCH-E.

	Método IFM			Med. dep.
Copula	GARCH-N	GARCH-t	GARCH-E	
$\alpha = 0,05$ (111)*				
SJC	0,0675 (150)	0,0698 (155)	0,0558 (124)	
Plackett	0,0684 (152)	0,0707 (157)	0,0576 (128)	0,0675 (150)
t-Student	0,0716 (159)	0,0761 (169)	0,0626 (139)	0,0720 (160)
$\alpha = 0,01$ (22)*				
SJC	0,0185 (31)	0,0149 (33)	0,0104 (23)	
Plackett	0,0225 (50)	0,0176 (39)	0,0131 (29)	0,0216 (48)
t-Student	0,0243 (54)	0,0189 (42)	0,0126 (28)	0,0248 (55)
$\alpha = 0,005$ (11)*				
SJC	0,0126 (28)	0,0090 (20)	0,0041 (9)	
Plackett	0,0162 (36)	0,0122 (27)	0,0077 (17)	0,0162 (36)
t-Student	0,0171 (38)	0,0126 (28)	0,0077 (17)	0,0162 (36)

Tabela A.7: Proporção de dados (e número de observações), para $t = 751$ a 2971 , em que a perda do portfolio excede o VaR para $\alpha = 0,05$, $\alpha = 0,01$ e $\alpha = 0,005$. * = Número esperado de observações.

A.5 Comparação com outros modelos

Iremos comparar os resultados obtidos pelo melhor modelo utilizando-se funções de acoplamentos (acoplamento SJC + GARCH-E) com diversos outros modelos bivariados e univariados.

Na Tabela A.8 temos a proporção de falhas e, entre parênteses, o p-valor do teste de Christoffersen, considerando os diversos modelos ajustados e três valores de α . Podemos ver que para $\alpha = 0,05$ o modelo EWMA aplicado diretamente no portfolio apresentou os melhores resultados, seguido pela função de acoplamento SJC. Já para $\alpha = 0,01$ e $\alpha = 0,005$ o melhor modelo foi a função de acoplamento SJC. No caso da função de acoplamento não rejeitamos a hipótese nula do teste de Christoffersen para nenhum dos valores de α considerado.

Vemos que para $\alpha = 0,01$ ou $\alpha = 0,005$ a função de acoplamento apresen-

tou um resultado bastante superior aos outros modelos considerados.

Modelo	$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,005$
Acoplamento SJC + GARCH-E	0,0558 (0,1342)	0,0104 (0,7251)	0,0041 (0,9063)
GARCH Bivariado (BEKK)	0,0819 (0,0000)	0,0338 (0,0000)	0,0248 (0,0000)
GARCH Bivariado (DCC)	0,0432 (0,0649)	0,0140 (0,1549)	0,0113 (0,0012)
EWMA (Bivariado)	0,0387 (0,0368)	0,0144 (0,7836)	0,0104 (0,1430)
GARCH-N (Portfolio)	0,0666 (0,0026)	0,0207 (0,0000)	0,0144 (0,0000)
GARCH-t (Portfolio)	0,0693 (0,0003)	0,0131 (0,2620)	0,0104 (0,0004)
EWMA (Portfolio)	0,0527 (0,2056)	0,0135 (0,7548)	0,0099 (0,0126)
Simulações Históricas (Portfolio)	0,1220 (0,0000)	0,0293 (0,0000)	0,0144 (0,0000)

Tabela A.8: Proporção de dados em que a perda do portfolio excede o VaR e P-valor do teste de Christoffersen (entre parênteses) para diversos valores de α - Índices Nasdaq e S&P500.

Referências Bibliográficas

AGUILAR, O.; WEST, M. Bayesian dynamic factor models and portfolio allocation. *Journal of Business and Economic Statistics*, v. 18, p. 338–357, 2000.

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. *Em Second International Symposium on Information Theory, editado por Em B.N. Petrov e F. Csake*, Budapest : Akademiai Kiado, p. 267–281, 1973.

ANG, A.; CHEN, J. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, v. 63(3), p. 443–494, 2002.

BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, v. 31, p. 307–327, 1986.

BOLLERSLEV, T.; ENGLE, R.; NELSON, D. Arch models. *Handbook of Statistics*, Amsterdã : North Holland, p. 2959–3038, 1994.

BOUYÉ, E. et al. Copulas for finance, a reading guide and some applications. *Documento de trabalho*, City University, 2003.

CHERUBINI, U.; LUCIANO, E. Value at risk trade-off and capital allocation with copulas. *Economic Notes*, v. 30, n. 2, 2000.

CHRISTOFFERSEN, P. Evaluating interval forecasts. *International Economic Review*, v. 39, p. 841–862, 1998.

DIEBOLD, F. X.; GUNTHER, T.; TAY, A. S. Evaluating density forecasts with applications to financial risk management. *International Economic Review*, v. 39, p. 863–883, 1998.

EMBRECHTS, P.; HÖING, A.; JURI, A. Using copulae to bound the value-at-risk for functions of dependent risks. *Finance and Stochastics*, v. 7, n. 2, p. 145–167, 2003.

EMBRECHTS, P.; LINDSKOG, F.; MCNEIL, A. Modelling dependence with copulas and applications to risk management. *Em Handbook of Heavy Tailed Distributions in Finance*, ed. S. Rachev, Nova Iorque : Elsevier/North Holland, v. 8, p. 329–384, 2003.

EMBRECHTS, P.; MCNEIL, A.; STRAUMAN, D. Correlation : Pitfalls and alternatives. *RISK*, p. 69–71, 1999.

EMBRECHTS, P.; MCNEIL, A.; STRAUMAN, D. Correlation and dependence in risk management: Properties and pitfalls. *Em Risk management: value at risk and beyond*, editado por Dempster M, Cambridge : Cambridge University Press, 1999.

ENGLE, R. Dynamic conditional correlation - a simple class of multivariate garch. *Journal of Business and Economics Statistics*, v. 17 (5), p. 425–446, 2002.

ENGLE, R.; KRONER, K. F. Multivariate simultaneous generalized arch. *Econometric Theory*, v. 50, p. 987–1007, 1995.

FANG, K.-T.; KOTZ, S.; NG, K.-W. Symmetric multivariate and related distributions. Londres : Chapman & Hall, 1987.

FORTIN, I.; KUZMICS, C. Tail-dependence in stock-return pairs. *International Journal of Intelligent Systems in Accounting, Finance & Management*, v. 11, p. 89–107, 2002.

GAMERMAN, D. *Simulação Estocástica via Cadeias de Markov*. 12º SINAPE: Associação Brasileira de Estatística, 1996.

GENEST, C.; RIVEST, L. Statistical inference procedures for bivariate archimedean copulas. *Journal of American Statistical Association*, v. 88, p. 1034–1043, 1993.

GEORGES, P. et al. Multivariate survival modeling: a unified approach with copulas. *Documento de trabalho*, Credit Lyonnais, 2001.

- GHYSELS, E.; HARVEY, A.; RENAULT, T. Stochastic volatility. *em Statistical Methods in Finance*, Amsterdã : North Holland, p. 119–191, 1996.
- HANSEN, B. Autoregressive conditional density estimation. *International Economic Review*, v. 35, p. 705–730, 1994.
- HOEFFDING, W. Massstabinvariante korrelationstheorie. *Schriften des Mathematischen Seminars und des Instituts für Angewandte Mathematik der Universität Berlin*, v. 5, p. 181–223, 1940.
- JOAG-DEV, K. Measures of dependence. *Handbook of Statistics*, Nova Iorque : North-Holland/Elsevier, v. 4, p. 79–88, 1984.
- JOE, H. *Multivariate Models and Dependence Concepts*. Londres: Chapman And Hall, 1997.
- JOHNSON, M. E. *Multivariate Statistical Simulation*. Nova Iorque: John Wiley and Sons, 1987.
- JORION, P. *Value at Risk*. Nova Iorque: McGraw-Hill, 1997.
- KIMBERLING, H. C. A probabilistic interpretation of complete monotonicity. *Aequationes Mathematicae*, v. 10, p. 152–164, 1974.
- LONGIN, F.; SOLNIK, B. Extreme correlation of international equity markets. *Journal of Finance*, v. 56(2), p. 649–676, 2001.
- LUCAS, E. C. Cálculo do var utilizando acoplamentos e teoria de valores extremos. *Dissertação de Mestrado - IMECC - Unicamp*, 2003.
- MASHAL., R.; ZEEVI, A. Beyond correlation : Extreme co-movements between financial assets. *Documento de trabalho*, Columbia University, 2002.
- MENEGUZZO, D.; VECCHIATO, W. Copula sensitivity in collateralized debt obligations and basket defaults swaps pricing and risk monitoring. *Documento de trabalho*, Veneto Banca, 2002.
- NELSEN, R. B. *Introdution to Copulas*. Nova Iorque: Springer Verlag, 1999.
- NELSON, D.; CAO, C. Inequality constraints in the univariate garch model. *Journal of Business and Economic Statistics*, v. 10, p. 229–235, 1992.

- PATTON, A. Modelling asymmetric exchange rate dependence. *Documento de trabalho*, University of California, San Diego, 2003a.
- PATTON, A. Estimation of multivariate models for time series of possibly different lengths. *Documento de trabalho*, University of California, San Diego, 2003b.
- PLACKETT, R. L. A class of bivariate distributions. *Journal of American Statistical Association*, v. 60, p. 516–522, 1965.
- ROCKINGER, M.; JONDEAU, E. Conditional dependency of financial series: An application of copulas. *Documento de trabalho*, HEC, Paris, 2001.
- ROMANO, C. Calibrating and simulating copula functions : an application to the italian stock market. *Documento de trabalho*, Capitalia Banking Group, Roma, 2002.
- SCARSINI, M. On measures of concordance. *Stochastica*, v. 8, p. 201–218, 1984.
- SHEPHARD, N. Statistical aspects of arch and stochastic volatility. *em Time Series Models in Econometrics, Finance and Other Fields, editado por D. R. Cox, D. V. Hinkley & O. E. Barndorff-Nielsen*, Chapman Hall, Londres, p. 1–67, 1996.
- SKLAR, A. *Personal communications to Roger B. Nelsen*, 1998.
- SMITH, M. Modelling sample selection using archimedean copulas. *Econometrics Journal*, v. 6, p. 99–123, 2003.
- WIDDER, D. V. *The Laplace Transform*. Princeton: Princeton University Press, 1941.
- XU, J. J. Statistical modelling and inference for multivariate and longitudinal discrete response data. *Tese de Doutorado*, Departamento de Estatística, University of British Columbia, 1996.