UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA
DEPARTAMENTO DE ESTATÍSTICA

Função de acoplamento t-Student assimétrica: modelagem de dependência assimétrica

Erick Andrade Busato
Orientador: Prof. Dr. Luiz Koodi Hotta

Dissertação apresentada junto ao Departamento de Estatística do Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas, para a obtenção do Título de Mestre em Estatística.

Campinas - SP
2008

Função de Acoplamento t-Student Assimétrica: Modelagem de Dependência Assimétrica

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Erick Andrade Busato e aprovada pela comissão julgadora.

Campinas, 03 de Dezembro de 2008.

Prof. Dr. Luiz Koodi Hotta
Orientador

Banca Examinadora:
1. Prof. Dr. Luiz Koodi Hotta (Orientador) – IMECC - UNICAMP.
2. Prof. Dr. Pedro Luiz Valls Pereira – EESP - FGV.
3. Prof. Dr. Mauricio Enrique Zevallos Herencia – IMECC - UNICAMP.

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

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP
Bibliotecária: Maria Júlia Milani Rodrigues – CRB8a 2116

	Busato, Erick Andrade
B96f	Função de acoplamento t-Student assimétrica: modelagem de dependência assimétrica / Erick Andrade Busato -- Campinas, [S.P. :s.n.], 2008.
	Orientador : Luiz Koodi Hotta Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.
	1. Copulas. 2. Risco (Economia) - Modelos matemáticos. 3. Distribuição (Probabilidades). 4. Análise multivariada. I. Hotta, Luiz Koodi. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Título em inglês: Skewed t-Student copula function: skewed dependence modelling

Palavras-chave em inglês (Keywords): 1. Copula. 2. Risk (Economics) – Mathematical models. 3. Distribution (Probability). 4. Multivariate analysis.

Área de concentração: Teoria de Copulas

Titulação: Mestre em Estatística

Banca examinadora:

Prof. Dr. Luiz Koodi Hotta (IMECC – UNICAMP)
Prof. Dr. Pedro Luiz Valls Pereira (EESP – FGV)
Prof. Dr. Mauricio Enrique Zevallos Herencia (IMECC – UNICAMP)

Data da defesa: 03/12/2008

Programa de pós-graduação: Mestrado em Estatística

Dissertação de Mestrado defendida em 03 de dezembro de 2008 e aprovada

Pela Banca Examinadora composta pelos Profs. Drs.

Prof(a). Dr(a). LUIZ KOODI HOTTA

Prof(a). Dr(a). PEDRO LUIZ VALLS PEREIRA

Prof(a). Dr(a). MAURICIO ENRIQUE ZEVALLOS HERENCIA

Agradecimentos

Agradeço aos meus pais Sérgio Busato e Vera Mazzoni, e aos meus irmãos Jonathan Busato e Caetano Ranieri pelos incentivos à continuidade dos estudos, aos amigos Henrique Leme Felizatti e Daniel Koji Ito, que contribuíram diretamente para o desenvolvimento desta dissertação.

Aos familiares e amigos: Gabriel, Cláudia, Maria Luiza, Fernando, Tatiana, Marques, Rita, Marcus, Rodolfo, Paulo, Célia, Patrick, Marcos, Carlos e Sara, dentre outros, que me auxiliaram de muitas formas e me deram apoio e suporte. À CAPES pelo auxílio financeiro nos primeiros meses de mestrado.

Especialmente ao estimado orientador Prof. Dr. Luiz Koodi Hotta por todo o auxílio e empenho dedicados a este trabalho.

Resumo

A família de distribuições t-Student Assimétrica, construída a partir da mistura em média e variância da distribuição normal multivariada com a distribuição Inversa Gama possui propriedades desejáveis de flexibilidade para as mais diversas formas de assimetria. Essas propriedades são exploradas na construção de funções de acoplamento que possuem dependência assimétrica. Neste trabalho são estudadas as características e propriedades da distribuição t-Student Assimétrica e a construção da respectiva função de acoplamento, fazendo-se uma apresentação de diferentes estruturas de dependência que pode originar, incluindo assimetrias da dependência nas caudas. São apresentados métodos de estimação de parâmetros das funções de acoplamento, com aplicações até a terceira dimensão da cópula. Essa função de acoplamento é utilizada para compor um modelo ARMA-GARCH-Cópula com marginais de distribuição t-Student Assimétrica, que será ajustado para os log-retornos de preços do Petróleo e da Gasolina, e log-retornos do Índice de Óleo AMEX, buscando o melhor ajuste, principalmente, para a dependência nas caudas das distribuições de preços. Esse modelo será comparado, através de medidas de Valor em Risco e AIC, além de outras medidas de bondade de ajuste, com o modelo de Função de Acoplamento t-Student Simétrico.

Palavras-Chave: Função de Acoplamento, Assimetria da Dependência, t-Student Assimétrica, Petróleo, Valor em Risco.

Abstract

The Skewed t-Student distribution family, constructed upon the multivariate normal mixture distribution, known as mean-variance mixture, composed with the Inverse-Gamma distribution, has many desirable flexibility properties for many distribution asymmetry structures. These properties are explored by constructing copula functions with asymmetric dependence. In this work the properties and characteristics of the Skewed t-Student distribution and the construction of a respective copula function are studied, presenting different dependence structures that the copula function generates, including tail dependence asymmetry. Parameter estimation methods are presented for the copula, with applications up to the 3^{rd} dimension. This copula function is used to compose an ARMA-GARCH-Copula model with Skewed t-Student marginal distribution that is adjusted to log-returns of Petroleum and Gasoline prices and log-returns of the AMEX Oil Index, emphasizing the return's tail distribution. The model will be compared, by the means of the VaR (Value at Risk) and Akaike's Information Criterion, along with other Goodness-of-fit measures, with models based on the Symmetric t-Student Copula.

Keywords: Copula Function, Skewed Dependence, Skewed t-Student, Petroleum, Value at Risk.

Sumário

Lista de Tabelas

Lista de Figuras

Capítulo 1

Introdução

O conhecimento da distribuição conjunta de probabilidade é importante em vários tipos de aplicações estatísticas, pois contém informações sobre cada componente marginal, além da estrutura de dependência entre as variáveis.

Pode-se ter uma idéia da estrutura através, por exemplo, da observação de medidas de dependência e gráficos comparativos conjuntos e marginais, que dão condições ao analista de selecionar um modelo probabilístico completo que possa ajustar a maior parte das características observadas no conjunto de observações.

Na maioria das vezes, as distribuições marginais, quando bem selecionadas, podem ser ajustadas de forma consistente. Entretanto, quando se procura uma distribuição conjunta que modele adequadamente os eventos marginais, e também a estrutura de dependência entre eles, não se obtém, em geral, resultados adequados.

O conceito de funções de acoplamento (cópulas), proposto por Sklar (1959) mas denominado e expandido por Joe (1997) e Nelsen (2006), faz a ligação entre distribuições marginais, modelando a estrutura de dependência entre as variáveis, permitindo que fenômenos ajustados por quaisquer distribuições marginais sejam conjuntamente modelados pela estrutura de dependência mais adequada.

A facilidade de construção das funções de acoplamento (através da transformada integral de probabilidade e do teorema de Sklar) permitiu o surgimento de diversas cópulas. Cada uma delas destina-se à modelagem de distintos perfis de dependência (veja Joe, 1997, Cap. 2), e cabe ao pesquisador estudar a dependência entre os fenômenos para selecionar uma função de acoplamento que se ajuste adequadamente à dependência entre todos os eventos (mesmo entre os eventos de menor probabilidade de ocorrência).

Uma área que vem fazendo uso cada vez maior das funções de acoplamento como forma de modelar a dependência entre fenômenos é a área de gestão de ativos, principalmente no controle do risco de mercado. Nesse caso, os gestores buscam acoplamentos que se ajustem bem às dependências entre fenômenos críticos, ou seja, que se ajustem bem à dependência presente nas caudas das distribuições (por exemplo, quando ocorrem eventos críticos, grandes perdas em um ativo, e os impactos que isso causa em um outro ativo, ver Caillault & Guégan, 2003).

O objetivo dos técnicos em gestão de risco durante uma modelagem de dependência através de cópulas é encontrar uma função de acoplamento que se ajuste bem à dependência nas caudas das distribuições, pois só assim poderão modelar o verdadeiro comportamento do mercado durante crises e calcular o risco financeiro (por exemplo como uma perda limítrofe de acordo com uma probabilidade e nível de significância).

Diversas funções de acoplamento podem modelar a dependência nas caudas. É o caso da Cópula t-Student (com dependência simétrica, possui o mesmo nível de dependência nas caudas inferior e superior), da cópula arquimediana Clayton (com dependência apenas na cauda inferior) e da Gumbel (apenas cauda superior); para maiores detalhes, veja Joe (1997), Cap. 5. Embora sejam cópulas de grande utilidade, mostram-se limitadas quando existem graus diversos de assimetria na dependência, e/ou ao se generalizar a função para dimensões das séries ou ativos maiores do que 2.

Na análise de risco é de interesse obter uma cópula que possa ser calibrada de forma a modelar, além de casos de dependência simétrica, diferentes perfis ou proporções de assimetria na dependência nas caudas, por exemplo, dependência apenas na cauda inferior, apenas na cauda superior, ou qualquer outra combinação assimétrica de dependência. Não apenas isso. Busca-se encontrar uma cópula flexível o bastante que ajuste corretamente a dependência entre dois fenômenos, tanto na região central dos eventos possíveis, aquela com maior massa de probabilidade, quanto em suas caudas, quando ocorrer qualquer assimetria na dependência.

Para suprir a necessidade de tal função de acoplamento, alguns métodos foram propostos na literatura. Patton (2004) utiliza uma mistura da cópula Clayton (que possui dependência na cauda inferior) com a cópula Clayton de sobrevivência, gerando uma cópula com características assimétricas. Stefanova (2007) utiliza uma mistura das cópulas Gaussiana (que não possui dependência nas caudas), Gumbel (com dependência na cauda superior) e Gumbel sobrevivência.

Outra alternativa é buscar distribuições conjuntas de probabilidade que possuam elevado nível de flexibilidade, e utilizar tal distribuição na obtenção de uma função de acoplamento. Nadaraja & Dey (2005) apresentam uma revisão de diversas formulações conhecidas de distribuições t-Student Assimétricas. Demarta & McNeil (2005) propõem uma outra distribuição t-Student Assimétrica a partir da adição de um termo de distribuição Inversa Gama à distribuição t-Student, e propõem seu uso na construção de uma função de acoplamento flexível. Essa mistura é chamada de mistura em média e variância da distribuição Normal e a distribuição originada é utilizada por Hu e Kercheval (2005) para estudar a estrutura de dependência entre moratórias em uma carteira de *'swaps'* de crédito, e por Sun *et al* (2008) para analisar a dependência assimétrica e não-linear no mercado de ações alemão.

O objetivo deste trabalho é estudar, a partir da distribuição t-Student Assimétrica proposta por Demarta e McNeil, que aparece também em McNeil, Frey e Embrechts (2005, Cap 3, p. 75), a capacidade dessa distribuição em modelar dependência nas caudas inferior e superior, atestando a sua flexibilidade no ajuste de dependências simétricas ou assimétricas. A Cópula construída comporá um modelo ARMA-GARCH-Cópula

(Filtragem e Estrutura de Dependência) que será aplicado a preços *'Spot'* de ativos físicos de energia da NYMEX (New York Mercantile Exchange): o Petróleo WTI (Óleo Cru *'West Texas Intermediate'*) e a Gasolina Convencional do Porto de Nova Iorque, e também o Índice de Óleo AMEX ("American Exchange Oil Index"). Modelos para esses ativos foram criados que descrevem o comportamento de seus preços e retornos (D. Pilipovic, 2007, Cap. 6), e trabalhos descrevendo a dependência entre os mesmos (p.ex. Biglova et al (2007) utiliza modelos de componentes principais para a dependência, filtrando as séries de preços através de modelos ARMA-GARCH). Grégoire et al (2008) utiliza funções de acoplamento do tipo Clayton e Frank, entre outras, para relacionar o preço 'spot' do Petróleo com o Gás Natural (não tratado nesta dissertação), filtrando também as séries através de modelos ARMA-GARCH.

A abordagem proposta aqui se inicia com a apresentação da metodologia geral da análise das séries e composição do modelo ARMA-GARCH-Cópulas, que tem por base a filtragem das séries de dados com posterior estimação de distribuições marginais e dependência através de cópula. A metodologia será apresentada no **Capítulo 2**.

Os resultados teóricos para a análise serão apresentados no **Capítulo 3**. Nesse capítulo, a distribuição t-Student Assimétrica de Demarta & McNeil será revista. Serão discutidas as propriedades da distribuição e os métodos de estimação de seus parâmetros. No mesmo capítulo será desenvolvida a função de acoplamento t-Student Assimétrica, mostrando-se a flexibilidade da distribuição em modelar os diferentes valores de dependência nas caudas inferior e superior que essa distribuição pode assumir.

No **Capítulo 4** realiza-se o estudo das séries de preços do Óleo Cru WTI, Índice de Óleo AMEX e Gasolina Convencional, propondo-se modelos ARMA-GARCH, cujos resíduos (inovações) serão ajustados por funções de acoplamento. Será apresentada a descrição das séries em estudo, suas características empíricas marginais e de dependência e toda a modelagem das distribuições marginais e de acoplamento das inovações, realizando-se também uma comparação com a cópula t-Student através de medidas de bondade de ajuste (descritas no apêndice D).

As conclusões serão apresentadas no **Capítulo 5**, seguida pelo **Apêndice**, apresentando, dentre outras, as definições e propriedades de funções de acoplamento, as propriedades da dependência, conceitos de simetria e assimetria, dependência nas caudas, ferramentas para medição de dependência e suas propriedades e alguns testes de bondade de ajuste para cópulas.

Capítulo 2

Metodologia

O objetivo deste trabalho é utilizar a função de acoplamento construída a partir da distribuição t-Student Assimétrica, proposta feita em Demarta & McNeil (2005), para compor um modelo ARMA-GARCH-Cópula e ajustar um conjunto de dados de preços de ativos energéticos, com o objetivo de se verificar a flexibilidade da cópula no ajuste de dados reais, principalmente no que diz respeito aos eventos caudais.

Desenvolver uma função de acoplamento a partir da distribuição conjunta t-Student Assimétrica significa utilizar o Teorema de Sklar (1959) e seu Corolário que, de uma forma geral, traz a relação entre uma função de distribuição conjunta com uma função de acoplamento que é única para variáveis aleatórias contínuas, distinguindo a distribuição conjunta em partes, que são as distribuições marginais e a relação de dependência entre elas (cópula).

O fato de que já existe uma grande gama de funções de acoplamento para atender a diversos perfis de dependência mostra que a utilização de uma nova opção de função de acoplamento deve ser feita apenas se trouxer ganhos substanciais quanto à flexibilidade da função em caracterizar todo, ou grande parte do perfil de dependência entre as variáveis.

No arcabouço teórico serão apresentadas as versões marginal e conjunta da distribuição t-Student Assimétrica, bem como algumas de suas propriedades, estimação dos parâmetros das distribuições e simulação de variáveis aleatórias. Em seguida será apresentada a função de acoplamento t-Student Assimétrica, construída a partir do Teorema de Sklar, e também a estimação dos parâmetros da cópula através do método de máxima verossimilhança.

A existência da flexibilidade na modelagem da dependência será verificada através de ajuste da função de acoplamento em dados reais do setor de energia. Outros exemplos da flexibilidade da função serão mostrados através da simulação da função de acoplamento.

O ajuste do modelo ARMA-GARCH-Cópula para um conjunto de dados se inicia com a filtragem através de modelos ARMA-GARCH das séries históricas, para que se obtenham séries livres de efeitos não aleatórios. Após a passagem do filtro ARMA-GARCH, as séries filtradas terão os seus efeitos marginais modelados através da função t-Student Assimétrica univariada e da Normal Assimétrica univariada.

Realiza-se então o ajuste da função de acoplamento. Uma abordagem é utilizar uma estimativa da distribuição marginal através da função distribuição empírica, calculando a inversa da função distribuição empírica acumulada de forma a obter marginais de distribuição Uniforme(0,1). A esses percentís são ajustadas as funções de acoplamento. Esse método é conhecido como Método de Máxima Verossimilhança Canônica para Cópulas. Outra alternativa é utilizar as funções distribuição acumuladas marginais paramétricas ajustadas, sendo esse método chamado de IFM (*"Inference Function for Marginals"*). Na aplicação serão utilizadas as distribuições marginais t-Student Assimétrica (que possui propriedades de caudas pesadas e assimetria) e/ou Normal Assimétrica.

Espera-se que a distribuição marginal t-Student Assimétrica possa suprir a necessidade de uma distribuição marginal que assuma valores extremos. Essa capacidade será medida nas estatísticas de VaR (Valor em Risco, veja Apêndice D).

O conjunto de dados resultante da aplicação da distribuição acumulada possuirá distribuição multivariada Uniforme (0,1), segundo a Transformada Integral de Probabilidade. Sobre esses dados dar-se-á a estimação dos parâmetros da função de acoplamento, através de máxima verossimilhança. Estatísticas baseadas no Valor em Risco, e outros testes de bondade de ajuste, são então utilizados para verificar a qualidade do ajuste da função, e se a mesma atendeu às características de dependência nas caudas.

Tanto a teoria quanto a prática serão apresentadas no caso bivariado ou 3-variado, sendo que, com exceção das medidas de dependência nas caudas (definidas para o caso bivariado), podem ser ampliadas para maiores dimensões, implicando em aumento considerável no número de parâmetros a serem estimados e no tempo de processamento, principalmente no algoritmo de máxima verossimilhança para estimação dos parâmetros da função de acoplamento t-Student Assimétrica. Entretanto é possível ter uma estimação aproximada a um custo computacional bem mais baixo, como será apresentado na seção 3.3.

Apenas por praticidade serão utilizados em todo o texto números decimais separados por ponto (como em 0.05), já que as saídas numéricas do software estatístico utilizado (R-2.7.1) são nesse formato.

Capítulo 3

Modelagem

O uso da distribuição Normal Multivariada para modelar distribuições conjuntas (e portanto a dependência) entre observações de um conjunto de variáveis aleatórias só é adequado em situações simétricas (simetria radial vide apêndice B), sem características de dependência nas caudas ou caudas pesadas. Além disso o modelo Normal considera apenas dependência linear entre as variáveis.

No mercado financeiro sabe-se que a dependência entre os diversos fatores de risco, ativos, moedas, etc. não atende à maioria das premissas para o uso da distribuição normal multivariada, motivando o surgimento de outros modelos de dependência, por exemplo, utilizando-se a distribuição t-Student multivariada em substituição à distribuição Normal. Essa abordagem atende à característica de caudas pesadas geralmente observada nas séries financeiras. Entretanto essa distribuição também é simétrica, e não caracteriza corretamente séries financeiras que possuam assimetria.

Busca-se então distribuições que sejam flexíveis na caracterização tanto da característica de caudas pesadas, quanto da assimetria.

Tomando-se como base o Teorema de Sklar, distribuições univariadas podem ser utilizadas para modelar as características marginais de um conjunto de variáveis aleatórias, modelando a dependência através de funções de acoplamento.

Neste capítulo serão tratadas as distribuições univariadas assimétricas Normal e t-Student, para modelagem de distribuições marginais (Seção 3.1). Será apresentada também a distribuição t-Student Assimétrica, que será a base para a construção de uma função de acoplamento elíptica flexível (Seção 3.2).

Na literatura existem diversas formulações de distribuições assimétricas baseadas na distribuição t-Student e na distribuição Normal. Decidiu-se dar ênfase à formulação da distribuição t-Student Assimétrica definida em Demarta & McNeil (2005), já que a mesma pode se ajustar à ocorrência de caudas pesadas e dependência nas caudas. Como alternativa para a modelagem marginal será utilizada a formulação de Azzalini (1985) da distribuição Normal Assimétrica.

3.1. Distribuições Assimétricas

Distribuições assimétricas é um termo geral que designa toda a família de distribuições que se afastam do conceito de simetria, por exemplo, no caso univariado é qualquer distribuição cuja probabilidade de ocorrência de eventos abaixo da média (se a mesma existir) seja diferente da probabilidade de ocorrência de eventos acima da média. Adicionalmente pode ocorrer que a média não seja coincidente com a mediana ou moda da distribuição. Nesse caso encontram-se diversas distribuições univariadas que podem ter essas características, como a Gama, F-Snedecor, Qui-Quadrada, entre outras.

O interesse, nesta dissertação, são as distribuições multivariadas assimétricas, onde aqui serão tratadas as distribuições formadas a partir de modificações nas distribuições simétricas Normal e t-Student. São distribuições mais flexíveis, que podem ser ajustadas para diferentes perfis de distribuições marginais, inclusive a normal e t-student simétricas, e generalizadas facilmente para dimensões maiores. São de interesse também as versões univariadas dessas distribuições, que podem ser utilizadas para modelagem das marginais.

Uma versão da distribuição multivariada Normal Assimétrica (Skew-Normal ou SN) é apresentada por Azzalini (1985), construída através de uma das famílias de distribuições que generaliza a distribuição normal, por incluir essa distribuição, permitindo assimetria diferente de zero (aqui será mostrado seu caso univariado). Quanto à distribuição t-Student Assimétrica multivariada, diversas formas já foram estudadas, sendo que algumas delas podem ser consultadas em Arellano-Valle & Genton (2005) ou Nadarajah & Dey (2005). Neste texto será utilizada a construção de Demarta & McNeil, apresentada na seção 3.1.2.

3.1.1 Distribuição Normal Assimétrica Univariada

Segundo Azzalini (1985) e Azzalini & Capitaneo (2003) uma construção possível de distribuições assimétricas é realizada a partir de uma forma "padrão", através das distribuições simétricas, pela multiplicação da função densidade pela função acumulada. A versão **univariada** da Distribuição Normal Assimétrica definida dessa forma é:

Definição 3.1: (Distribuição Normal Assimétrica) Seja $\phi(.)$ a função densidade da distribuição Normal, com função distribuição acumulada $\Phi(.)$. Então, X possui Distribuição Normal Assimétrica se sua densidade é dada por:

$$f_X(x) = 2\phi(x)\Phi(\alpha x), \qquad [3\text{-}1]$$

para algum parâmetro $\alpha \in \Re$. Adicionando-se os parâmetros de posição e escala, através da transformação $x \to \dfrac{x-\mu}{\sigma}$, têm-se a função densidade de probabilidade:

$$f_X(x) = \frac{2}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \int_{-\infty}^{\alpha\left(\frac{x-\mu}{\sigma}\right)} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \, . \qquad [3\text{-}2]$$

Algumas propriedades da Distribuição Normal Assimétrica são:

1. Quando $\alpha = 0$, a assimetria se anula e obtém-se a Distribuição Normal;
2. Conforme α aumenta (em valor absoluto), a assimetria da distribuição aumenta;

3. Quando $\alpha \to \infty$, a densidade converge à densidade da distribuição "*folded-normal*" (Se X é variável aleatória de distribuição Normal, $Y = |X|$ possui distribuição "*folded-normal*");
4. Se o sinal de α for alterado, a assimetria é refletida no lado oposto do eixo vertical.

Na figura 3.1 são apresentadas curvas da distribuição normal assimétrica para diferentes valores do parâmetro de assimetria.

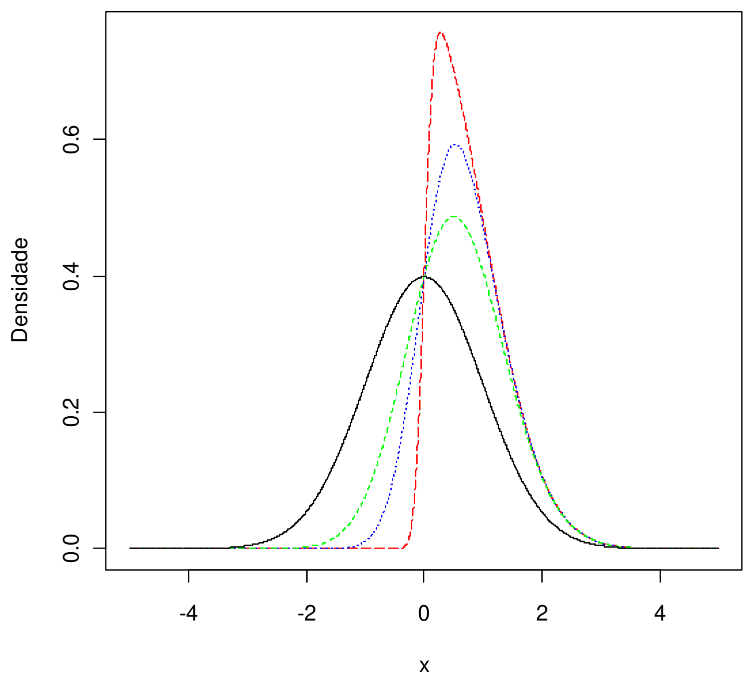

Figura 3. 1. Densidade da Distribuição Normal Assimétrica, para $\alpha = 0$ (———), $\alpha = 1$ (- - - - -), $\alpha = 2$ (·······) e $\alpha = 8$ (− − − −).

Essa distribuição pode ser utilizada quando os dados apresentam formato semelhante à distribuição normal, entretanto, com assimetria entre -1 e 1 e baixo nível de excesso de curtose (até aproximadamente 0.9, vide figura 3.2).

Se $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$, a esperança, variância, assimetria e curtose dessa distribuição são dadas por:

$$E(X) = \mu + \sigma\delta\sqrt{\frac{2}{\pi}}\,, \qquad [3\text{-}3]$$

$$Var(X) = \sigma^2\left(1 - \frac{2\delta^2}{\pi}\right), \qquad [3\text{-}4]$$

$$Assim(X) = \frac{(4-\pi)\left(\delta\sqrt{2/\pi}\right)^3}{2\left(1 - \frac{2\delta^2}{\pi}\right)^{3/2}}\,, \qquad [3\text{-}5]$$

$$Exc.Curtose(X) = 2(\pi - 3)\frac{\left(\delta\sqrt{2/\pi}\right)^4}{\left(1 - \frac{2\delta^2}{\pi}\right)^2} - 3\,. \qquad [3\text{-}6]$$

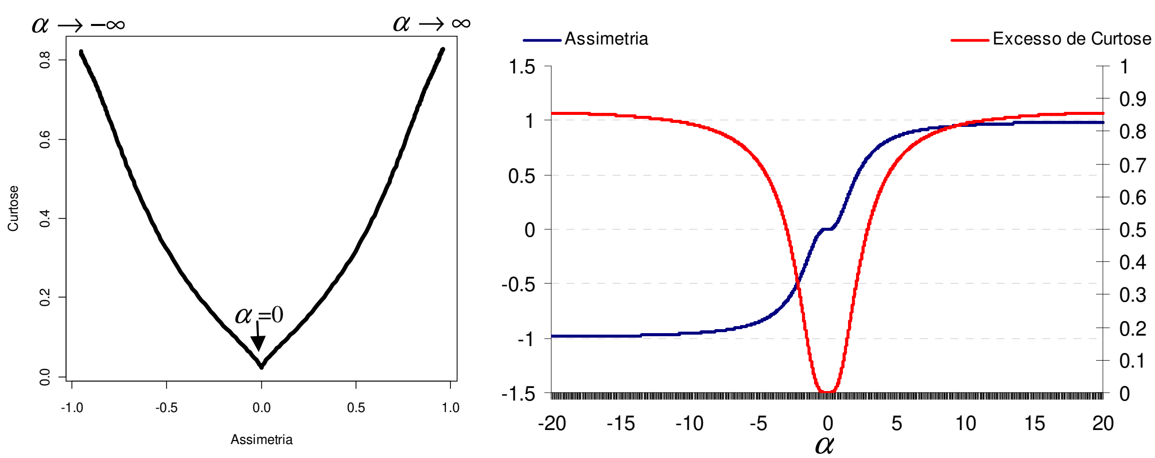

Figura 3. 2. Gráfico de Assimetrias e Curtoses da distribuição Normal Assimétrica para parâmetros de assimetria α entre -20 e 20.

O cálculo dos limites das equações [3-5] e [3-6] mostra que quando $\alpha \to -\infty$ a assimetria converge para -0.9953 e o excesso de curtose para 0.8692 .Quando $\alpha \to \infty$ a assimetria converge para 0.9953 e o excesso de curtose também para 0.8692. Esse movimento pode ser observado na figura 3.2, onde se observa também que a assimetria e o excesso de curtose são unicamente determinados por um parâmetro de assimetria, e que se for fixado o excesso de curtose, o valor absoluto da assimetria também estará determinado, sendo sua direção determinada pelo sinal do parâmetro de assimetria. Essa distribuição não é apropriada para ajuste quando o excesso de curtose é maior do que 0.8692.

3.1.2. Distribuição t-Student Assimétrica

Uma distribuição assimétrica, que tem por base a distribuição t-Student, foi proposta por Demarta & McNeil. O fato de que essa distribuição é flexível, já que se ajusta a diferentes valores do 1º ao 4º momentos centrais, torna-a uma boa alternativa para utilização na modelagem das distribuições marginais, quando da estimação de uma função de acoplamento na caracterização da estrutura de dependência de um conjunto de dados.

Sendo diferente da Distribuição Normal Assimétrica de Azzalini (1985), e das Distribuições t-Student Assimétricas apresentadas por Azzalini & Capitaneo (2003) ou selecionadas por Nadarajah & Dey (2005), a Distribuição t-Student Assimétrica apresentada por Demarta & McNeil (2005), possui uma formulação distinta. Essa formulação é um caso particular da *família das distribuições de misturas em variância da distribuição normal*, que por si é um caso particular da *família de distribuição de misturas em média e variância*. A família de distribuições de misturas em variância da distribuição normal inclui a distribuição t-Student simétrica, sendo que a deformação da simetria da distribuição t-Student é realizada através da adição de um termo estocástico.

Definição 3.2: (Distribuições de Mistura em Variância da Distribuição Normal). Seja Z e W variáveis aleatórias tal que $Z \sim N(0,\sigma^2)$ e W é independente de Z. Define-se por Distribuição de Mistura em Variância da Distribuição Normal a distribuição cuja construção é $X =_d \mu + Z\sqrt{W}$. (A notação $=_d$ significa igualdade em distribuição).

Definição 3.3: (Distribuições de Mistura em Média e Variância da Distribuição Normal). Seja uma função g e um conjunto de parâmetros $\mu_i, \gamma_i \in \Re$, $i=1,2,...d$. Sejam $\mathbf{Z} \sim N_d(0,\Sigma)$ Normal d-dimensional e W univariada independente de **Z,** variáveis aleatórias. Define-se por distribuição de Mistura em Média e Variância da Distribuição Normal a distribuição X construída por: $\mathbf{X} =_d \boldsymbol{\mu} + \boldsymbol{\gamma}' g(W) + \mathbf{Z}\sqrt{W}$, onde $\boldsymbol{\mu}$ e $\boldsymbol{\gamma}$ são vetores *d*-dimensionais.

A distribuição t-Student Assimétrica é construída a partir da formulação de Mistura em Média e Variância da Distribuição Normal.

Definição 3.4: (Distribuição t-Student Assimétrica). Suponha que $\mathbf{X} = (X_1, X_2, ..., X_d)$ seja um vetor aleatório cuja distribuição pertença à família das distribuições de mistura em média e variância. Se *g(W) = W*, onde *W ~ Ig(ν/2, ν/2)* a distribuição Inversa Gamma univariada cujos parâmetros remetem à distribuição Qui-Quadrada, então **X** possui distribuição t-Student Assimétrica *d*-dimensional, cuja função densidade é dada por:

$$f_{\mathbf{X}}(\mathbf{x}) = c\frac{K_{\frac{\nu+d}{2}}\left(\sqrt{\left(\nu+(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}\right)\exp\left((\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}\right)}{\left(\sqrt{\left(\nu+(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}\right)^{-\frac{\nu+d}{2}}\left(1+\frac{(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{\nu}\right)^{\frac{\nu+d}{2}}}, \qquad [3\text{-}7]$$

onde K_λ é a função de Bessel Modificada de Segundo Tipo (veja Abramowitz & Stegun (1965), Capítulos 9 e 10), $\mathbf{x} = (x_1, x_2, ..., x_d)$, $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_d)$, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, ..., \gamma_d)$ e $\Sigma = \begin{pmatrix} \sigma_1^{\ 2} & \cdots & \sigma_{1,d} \\ \vdots & \ddots & \vdots \\ \sigma_{d,1} & \cdots & \sigma_d^{\ 2} \end{pmatrix}$. A constante de normalização é dada por:

$$c = \frac{2^{\frac{2-(\nu+d)}{2}}}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}. \qquad [3\text{-}8]$$

A notação para um vetor aleatório **X** de distribuição t-Student Assimétrica é:

$$\mathbf{X} \sim \mathrm{st}_d(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma}).$$

Distribuição Marginal Univariada

Seja $\mu_i, \gamma_i \in \Re$, $\sigma_i^2 > 0$ parâmetros de posição, assimetria e escala, respectivamente e $\nu > 0$ o número de graus de liberdade. Uma variável aleatória *X* de distribuição t-Student Assimétrica possui função densidade de probabilidade:

$$f_i(x_i) = \frac{2^{\frac{1-v}{2}} K_{\frac{v+1}{2}}\left(\sqrt{\left(v + \frac{(x_i-\mu_i)^2}{\sigma_i^{\,2}}\right)(\gamma^2)}\right)\exp((x_i-\mu_i)\gamma_i)}{\Gamma\left(\frac{v}{2}\right)(\pi v)^{\frac{1}{2}}\sigma_i\left(\sqrt{\left(v + \frac{(x_i-\mu_i)^2}{\sigma_i^{\,2}}\right)(\gamma_i^{\,2})}\right)^{-\frac{v+1}{2}}\left(1 + \frac{(x_i-\mu_i)^2}{v_i\sigma_i^{\,2}}\right)^{\frac{v+1}{2}}} . \qquad [3\text{-}9]$$

A distribuição t-Student Assimétrica Univariada é denotada por $X_i \sim \mathrm{st}_1(v,\mu_i,\sigma_i^2,\gamma_i)$, sendo a i-ésima distribuição marginal univariada da distribuição t-Student Assimétrica multivariada apresentada na definição 3.4.

Essa distribuição pode ser considerada uma generalização da distribuição t-Student Simétrica pois retorna-se a ela quando $\gamma_i \to 0$. Nota-se ainda que a presença da função de Bessel na formulação torna-a menos tratável analiticamente.

Sejam as variáveis aleatórias Z e W independentes e com distribuições Normal$(0,\sigma^2)$ e Inversa-Gamma, respectivamente. A esperança, variância, assimetria e curtose da distribuição t-Student Assimétrica $\mathrm{st}_1(v,\mu_i,\sigma^2,\gamma)$ univariada são obtidas da definição da distribuição t-Student Assimétrica ($X =_d \mu + \gamma W + Z\sqrt{W}$) e dadas por:

$$E(X) = \frac{v}{v-2}\gamma + \mu , \qquad [3\text{-}10]$$

$$Var(X) = \frac{v}{v-2}\sigma^2 + \frac{2v^2}{(v-2)^2(v-4)}\gamma^2 , \qquad [3\text{-}11]$$

$$Assim(X) = \frac{\gamma^3 E(W^3) + 3\gamma.E(W^2) - 3E(X).Var(X) - [E(X)]^3}{\left(\sqrt{Var(X)}\right)^3} , \qquad [3\text{-}12]$$

$$Exc.Curtose(X) = \frac{\gamma^4 E(W^4) + 6\gamma^2.E(W^3) + E(Z^4).E(W^2)}{\left(\sqrt{Var(X)}\right)^4} - 3 , \qquad [3\text{-}13]$$

onde,

$$E(W^2) = \left(\frac{v}{v-2}\right)^2 + \frac{v^2}{2(v-2)^2(v-4)}, \qquad [3\text{-}14]$$

$$E(W^3) = \frac{k^3[3(k-2)(k-3)+(k-2)^2(k-3)+4]}{(k-1)^3(k-2)^2(k-3)}, \text{ onde } k = \frac{v}{2}, \qquad [3\text{-}15]$$

$$E(W^4) = E(W)[4E(W^3)+6E(W^2)E(W)-(3E(W))^4] + \left[\frac{30k-66}{(k-3)(k-4)}+3\right]\left[\sqrt{Var(W)}\right]^4, \text{ onde } k = \frac{v}{2}, \qquad [3\text{-}16]$$

$$E(Z^4) = 3\sigma^2. \qquad [3\text{-}17]$$

É importante notar que a existência do 2º momento central da distribuição t-Student Assimétrica depende de graus de liberdade maiores do que 4, a existência do 3º momento central depende de graus de liberdade maiores do que 6 e a existência do 4º momento central depende de graus de liberdade maiores do que 8. A figura 3.3 mostra exemplos da densidade da t-Student Assimétrica para alguns parâmetros de assimetria e 8.5 g.l.

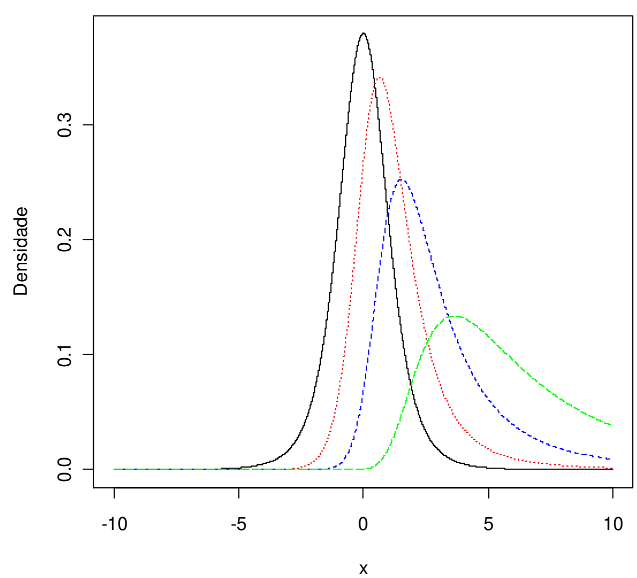

Figura 3. 3. Densidade de probabilidade da distribuição t-Student Assimétrica padronizada com 5 graus de liberdade e parâmetros de assimetria $\gamma = 0$ (——), $\gamma = 0.8$ (- - - -), $\gamma = 2$ (········) e $\gamma = 5$ (- - - - -).

Através da figura 3.3 nota-se que parâmetros de assimetria positivos causam a ocorrência de assimetria em direção ao eixo positivo da distribuição, o que ocorre de forma análoga com parâmetros de assimetria negativos. Através da fórmula para a Assimetria (3º momento central padronizado) obtém-se a seguinte propriedade:

Proposição 3.1: Suponha X e Y variáveis aleatórias tais que $X \sim \mathrm{st}_1(\nu,\mu,\sigma^2,\gamma)$ e $Y \sim \mathrm{st}_1(\nu,\mu,\sigma^2,-\gamma)$. Então, Assimetria($X$)=-Assimetria($Y$).

Prova: Diretamente através da fórmula [3-12], substituindo-se γ por $-\gamma$ nos 4 termos da equação.■

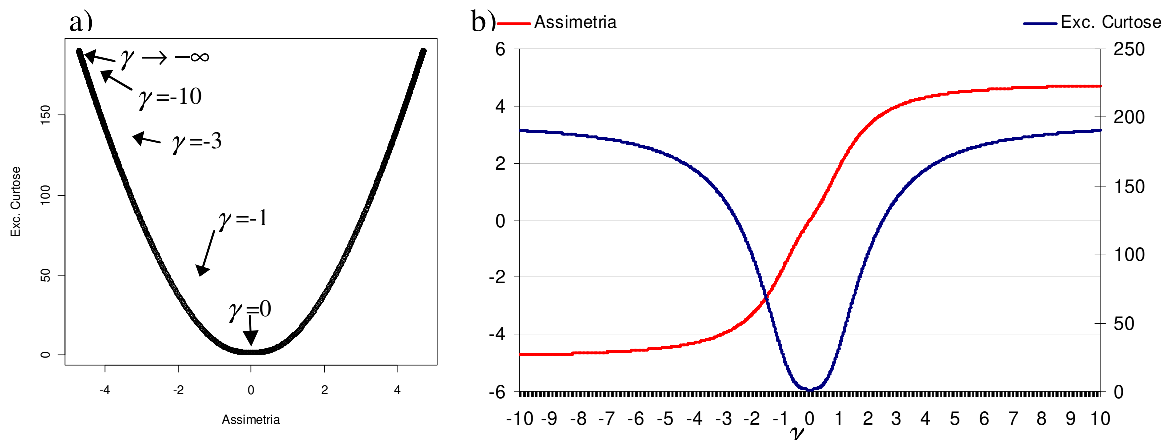

Figura 3. 4. Gráfico de Assimetria e Exc. Curtose da distribuição t-Student Assimétrica para parâmetros de assimetria γ entre -10 e 10, calculadas para 8.5 graus de liberdade.

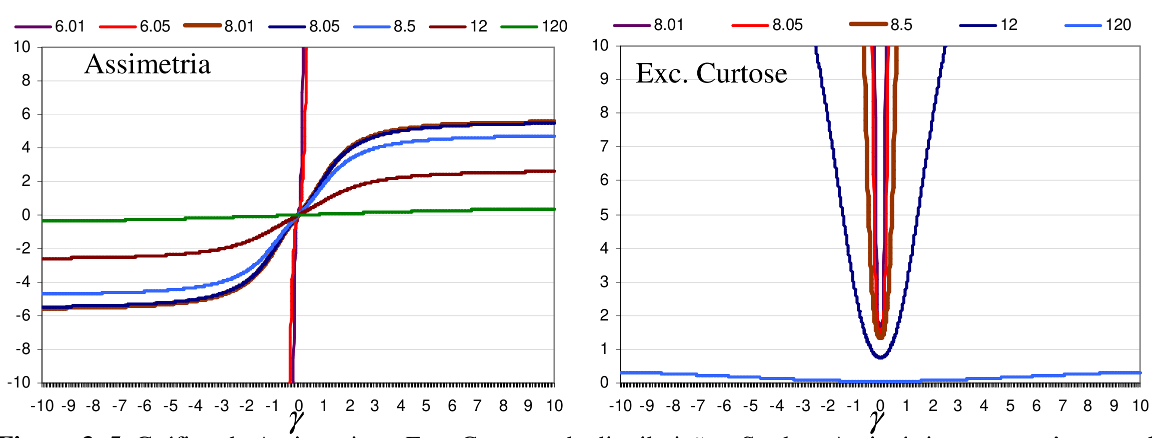

Figura 3. 5. Gráfico de Assimetrias e Exc. Curtoses da distribuição t-Student Assimétrica para parâmetros de assimetria γ entre -10 e 10, calculadas para diversos graus de liberdade.

A figura 3.4 mostra a assimetria e excesso de curtose da distribuição t-Student assimétrica a) compara a variação conjunta da assimetria e excesso de curtose conforme varia-se o parâmetro de assimetria; b) mostra a variação distinta desses momentos conforme varia se o parâmetro. Note que assimetria é infinita para graus de liberdade menores ou iguais a 6 e excesso de curtose é infinito para graus de liberdade menores ou iguais a 8. É possível que exista assimetria menor do que infinito quando excesso de curtose é "igual" a infinito. Por possuir 1º a 4º momentos bem definidos quando o número de graus de liberdade é maior do que 8, a estimação dos parâmetros da distribuição através do método de momentos pode ser definida, para $\theta = \{\mu, \sigma^2, \gamma, v\}$ pelas equações:

$$T_1 = \frac{v}{v-2}\gamma + \mu - \bar{x}\,, \tag{3-18}$$

$$T_2 = \frac{v}{v-2}\sigma^2 + \frac{2v^2}{(v-2)^2(v-4)}\gamma^2 - S_x^{\ 2}\,, \tag{3-19}$$

$$T_3 = \frac{\gamma^3 E(W^3) + 3\gamma.E(W^2) - 3E(X).Var(X) - [E(X)]^3}{\left(\sqrt{Var(X)}\right)^3} - Assimetria(x)\,, \tag{3-20}$$

$$T_4 = \frac{\gamma^4 E(W^4) + 6\gamma^2.E(W^3) + E(Z^4).E(W^2)}{\left(\sqrt{Var(X)}\right)^4} - 3 - Exc.Curtose(x)\,, \tag{3-21}$$

onde $\bar{x}$, $S_x^{\ 2}$, $Assimetria(x)$ e $Exc.Curtose(x)$ são estatísticas amostrais de média variância, assimetria e excesso de curtose de um vetor de observações *x*. As estimativas são obtidas minimizando-se a soma ponderada do quadrado das equações nos parâmetros em θ. O uso desse método não gerou resultados satisfatórios pois não se observou convergência adequada na minimização, para diversas simulações geradas. Entretanto, quando considera-se um caso padrão, onde $\mu = 0$ e $\sigma^2 = 1$ as equações [3-18] e [3-19] podem ser igualadas à zero e resolvidas analiticamente. Nesse caso, um método de momentos para a distribuição t-Student Assimétrica padrão pode ser realizado através de:

1. Resolver a equação de 2º grau:

$$\left(\frac{\sum(x-\overline{x})^2}{n-1}-1\right).\hat{v}^2+\left(4-2\overline{x}-6\frac{\sum(x-\overline{x})^2}{n-1}\right).v+\left(4\overline{x}^2+8\frac{\sum(x-\overline{x})^2}{n-1}\right)=0, \quad [3\text{-}22]$$

2. $\hat{\gamma}=\frac{(\hat{v}-2)}{\hat{v}}\overline{x}$. [3-23]

Mas, se os dados possuírem média 0 e variância 1, ou seja, dados padronizados, uma solução trivial para os parâmetros será encontrada ($v=4$ e $\gamma=0$), impossibilitando a utilização desses momentos, tornando necessária a utilização de momentos de ordens superiores, por exemplo, 3º e 4º momentos para estimação dos parâmetros. Importante notar que existem duas estimativas para o grau de liberdade, advindas da solução da equação de segundo grau. Escolhe-se o grau de liberdade cujo valor seja maior do que 4, para garantir a existência do segundo momento.

Se ambos forem maiores do que 4, pode-se utilizar uma média dentre os dois, ou a estimativa mais próxima de uma estimativa de máxima verossimilhança. No caso que ambos são menores do que 4, deve-se utilizar o método de máxima verossimilhança para determinar o real valor da estimativa. Se a estimativa por máxima verossimilhança também for menor do que 4, então o modelo ajustado não possui variância finita, e pode-se utilizar a estimativa pelo método de momentos mais próxima da máxima verossimilhança. Esta escolha, no entanto, leva a perda de uma das principais propriedades do Método de Momentos, comparado com o Método de Máxima Verossimilhança, que é a facilidade de estimação.

As estimativas pelo método de máxima verossimilhança podem ser obtidas numericamente, maximizando nos parâmetros a soma do logaritmo da função [3-9] aplicada ao conjunto de dados. Na prática esse método gera estimativas mais precisas para os parâmetros, como será visto nos exemplos da seção 3.3.

Distribuição t-Student Assimétrica Bivariada

Seja $\gamma_i, \gamma_j \in \Re$, *i,j*=1,...,d, parâmetros de assimetria, v o número de graus de liberdade e $\Sigma_{i,j} = \begin{pmatrix} \sigma_i^2 & \sigma_{i,j} \\ \sigma_{j,i} & \sigma_j^2 \end{pmatrix}$ uma partição da matriz Σ da definição 3.4. Um par de variáveis aleatórias (X_i, X_j) possui distribuição conjunta t-Student Assimétrica, cuja função densidade de probabilidade conjunta é definida pela equação [3-7] fixando-se a dimensão em 2.

Um caso particular de interesse, que será utilizado para construção de uma função de acoplamento é definido quando fixa-se $\mu_i = \mu_j = 0$ e $\sigma_i^2 = \sigma_j^2 = 1$. Nesse caso uma matriz de correlações será necessária ao modelo, e não uma matriz de variâncias-covariâncias completa. Isso tornará a função de acoplamento que será construída mais simples de ser interpretada. A função bivariada nesse caso é dada por:

$$f_{X_1,X_2}(x_1,x_2) = c\frac{K_{\frac{v+2}{2}}\left(\sqrt{\left(v+\frac{x_1{}^2-2\rho x_1x_2+x_2{}^2}{1-\rho^2}\right)\left(\frac{\gamma_1^2-2\rho\gamma_1\gamma_2+\gamma_2^2}{1-\rho^2}\right)}\right).\exp\left(\frac{\gamma_1(x_1-\rho x_2)+\gamma_2(x_2-\rho x_1)}{1-\rho^2}\right)}{\left(\sqrt{\left(v+\frac{x_1{}^2-2\rho x_1x_2+x_2{}^2}{1-\rho^2}\right)\left(\frac{\gamma_1^2-2\rho\gamma_1\gamma_2+\gamma_2^2}{1-\rho^2}\right)}\right)^{-\frac{v+2}{2}}\left(1+\frac{x_1{}^2-2\rho x_1x_2+x_2{}^2}{v.(1-\rho^2)}\right)^{\frac{v+2}{2}}}, \quad [3\text{-}24]$$

$$\text{onde } c = \frac{2^{\frac{-v}{2}}}{\Gamma\left(\frac{v}{2}\right)(\pi v)\sqrt{1-\rho^2}}. \quad [3\text{-}25]$$

Denota-se essa distribuição por $\mathbf{X} \sim \mathrm{st}_2(v,\mathbf{0},\rho,\boldsymbol{\gamma})$, onde $\mathbf{X} = (X_1, X_2)$, $\mathbf{0} = (0,0)$, $\rho \in [-1,1]$ e $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$.

Quando $\gamma_1, \gamma_2 \to 0$ essa distribuição converge para a distribuição t-Student Simétrica bivariada. As distribuições marginais obtidas a partir da distribuição conjunta t-Student Assimétrica são também t-Student Assimétrica univariada ($X_i \sim \mathrm{st}_1(v,0,1,\gamma_i)$).

Para a estimação dos parâmetros dessa distribuição podem ser utilizados métodos baseados no algoritmo EM (ver por exemplo Hu, 2005), métodos de momentos e máxima verossimilhança. A estimação dos parâmetros da distribuição bivariada [3-24] pelo método de momentos é realizada de forma semelhante à estimação dos parâmetros da distribuição univariada padrão. O parâmetro de correlação que compõe a distribuição pode ser estimado a partir da covariância amostral. Os métodos de momentos tanto para o caso univariado padrão quanto para o caso bivariado padrão são obtidos através da solução das equações da média e variância da distribuição, que no caso bivariado é:

$$\begin{pmatrix} E(X_1) \\ E(X_2) \end{pmatrix} = \left(\frac{v}{v-2}\right)\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \qquad [3\text{-}26]$$

$$\begin{pmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_1, X_2) & Var(X_2) \end{pmatrix} = \left(\frac{v}{v-2}\right)\begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_2^2 \end{pmatrix} + \frac{2v^2}{(v-2)^2(v-4)}\begin{pmatrix} \gamma_1^2 & \gamma_1\gamma_2 \\ \gamma_2\gamma_1 & \gamma_2^2 \end{pmatrix}. \qquad [3\text{-}27]$$

Para o caso bivariado padrão ($\mathrm{st}_2(v,\mathbf{0},\rho,\boldsymbol{\gamma})$), como existem 5 equações para estimação de 4 parâmetros é necessária a aplicação do método de momentos generalizado, atribuindo-se, por exemplo, pesos iguais a 0.5 às equação [3-30] e [3-31], e pesos elevados às outras equações. Aproximadamente isso é equivalente a realizar uma minimização da soma dos quadrados dessas funções no parâmetro *v*, igualando as outras equações à zero. As equações [3-28] e [3-29] são construídas diretamente a partir da equação [3-26] e as equações [3-30] a [3-32] a partir da equação [3-27].

$$T1 = \frac{(\hat{v}-2)}{\hat{v}}\bar{x} - \hat{\gamma}_1; \qquad [3\text{-}28]$$

$$T2 = \frac{(\hat{v}-2)}{\hat{v}}\bar{y} - \hat{\gamma}_2; \qquad [3\text{-}29]$$

$$T3 = \left(\frac{\sum(x-\bar{x})^2}{n-1} - 1\right).\hat{v}^2 + \left(4 - 2\bar{x} - 6\frac{\sum(x-\bar{x})^2}{n-1}\right).\hat{v} + \left(4\bar{x}^2 + 8\frac{\sum(x-\bar{x})^2}{n-1}\right); \qquad [3\text{-}30]$$

$$T4 = \left(\frac{\sum(y-\bar{y})^2}{n-1} - 1\right).\hat{v}^2 + \left(4 - 2\bar{y} - 6\frac{\sum(y-\bar{y})^2}{n-1}\right).\hat{v} + \left(4\bar{y}^2 + 8\frac{\sum(y-\bar{y})^2}{n-1}\right); \qquad [3\text{-}31]$$

$$T5 = \left(\frac{\Sigma(x-\bar{x})(y-\bar{y})}{n-1} - \frac{2\bar{x}.\bar{y}}{\hat{v}-4}\right)\left(\frac{\hat{v}-2}{\hat{v}}\right) - \hat{\rho}; \qquad [3\text{-}32]$$

No caso bivariado completo, $\mathrm{st}_2(v,\boldsymbol{\mu},\Sigma,\boldsymbol{\gamma})$, o 1º e 2º momentos não são suficientes para a estimação de todos os parâmetros. Nesse caso serão necessários momentos de ordens superiores, como a assimetria e excesso de curtose para a estimação.

As figuras 3.6 e 3.7 exibem gráficos de contorno da densidade da distribuição t-Student Assimétrica bidimensional, bem como o contorno da função acumulada dessa distribuição com os parâmetros $v = 6$, $\rho = 0.7$ e assimetria $(\gamma_1, \gamma_2) = (0.8\ ,\ 0.8)$ e $(\gamma_1, \gamma_2) = (-0.8\ ,\ 0.8)$. Veja na figura 3.6 que, embora apresente assimetria marginal, a igualdade dos parâmetros de assimetria mantém a simetria radial, já que ambas as marginais terão o mesmo comportamento e perfil assimétrico. Já na figura 3.7, a presença de valores diferentes de assimetria causa tanto assimetria marginal quanto assimetria radial, já que as marginais possuem comportamentos assimétricos inversos.

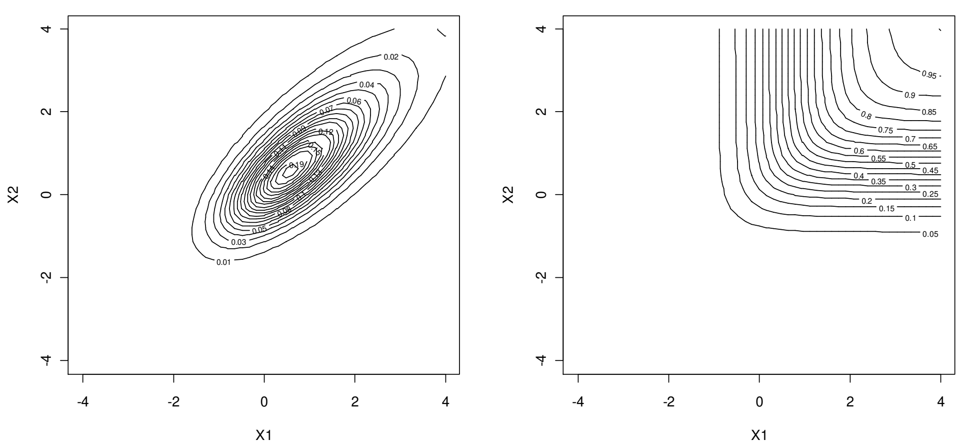

Figura 3. 6. Contornos da Distribuição t-Student Assimétrica – Função Densidade e Função Acumulada com parâmetros $v = 6$, $\rho = 0.7$ e assimetria $(\gamma_1, \gamma_2) = (0.8\ ,\ 0.8)$

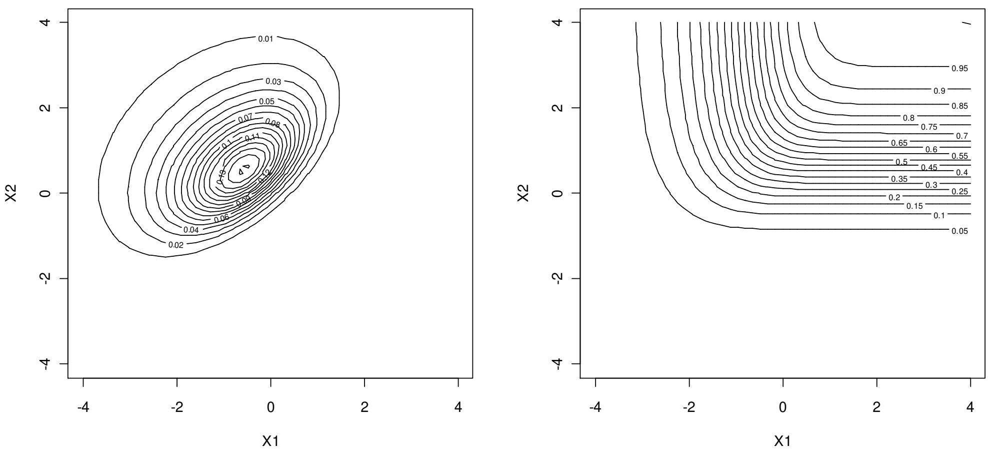

Figura 3. 7. Contornos da Distribuição t-Student Assimétrica – Função Densidade e Função Acumulada com parâmetros $\nu = 6$, $\rho = 0.7$ e assimetria $(\gamma_1, \gamma_2) =$ (-0.8 , 0.8)

O método de máxima verossimilhança para a versão bidimensional da distribuição t-Student Assimétrica é análogo ao caso univariado. Realiza-se através da maximização numérica (nos parâmetros) da soma do logaritmo da função [3-24] aplicada ao conjunto de dados. Além disso, o caso particular bivariado é extensível para dimensões superiores, utilizando-se parâmetros de posição nulos e parâmetros de escala iguais a 1 na equação [3-9]. Mais detalhes sobre a estimação de parâmetros através da máxima verossimilhança para dimensões maiores do que 2 serão dados na seção 3.3.

3.2. Função de Acoplamento

As funções de acoplamento são distribuições conjuntas de probabilidade de marginais Uniformes (0,1) que, através da *Transformada Integral de Probabilidade* (Veja Apêndice A) juntam (acoplam) diferentes distribuições marginais, possibilitando a caracterização da dependência entre eventos aleatórios, mesmo que sejam eventos de diferentes descontinuidades.

Será tratada neste capítulo a família de funções de acoplamento gerada pela distribuição t-Student Assimétrica, sendo uma alternativa para a caracterização de dependências assimétricas entre conjuntos de dados. A construção da função de acoplamento tem por base o Corolário do Teorema de Sklar (Apêndice A-2).

Função de Acoplamento t-Student Assimétrica

A função de acoplamento é definida a partir do Corolário do Teorema de Sklar, utilizando a equação da distribuição t-Student Assimétrica (equação [3-7]), e será chamada de Cópula t-Student Assimétrica Multivariada, e é definida por:

Definição 3.6. Seja v o parâmetro de graus de liberdade, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, ..., \gamma_d)$ um vetor com d parâmetros de assimetria e $\mathbf{P}$ uma matriz de correlações de ordem dxd. Define-se Função de Acoplamento t-Student Assimétrica Multivariada a distribuição:

$$\begin{aligned} &C_{ST}(u_1, u_2, ..., u_d; v, \mathbf{P}, \boldsymbol{\gamma}) = \\ &ST_d(ST_1^{-1}(u_1; v, 0, 1, \gamma_1), ST_1^{-1}(u_2; v, 0, 1, \gamma_2), ..., ST_1^{-1}(u_d; v, 0, 1, \gamma_d); v, \mathbf{0}, \mathbf{P}, \boldsymbol{\gamma}) \end{aligned}, \qquad [3\text{-}33]$$

onde $ST_1 = \int_{-\infty}^{x} st_1(t; v, 0, 1, \gamma_i) dt$ é a função distribuição acumulada da t-Student Assimétrica univariada e ST_1^{-1} a sua inversa. ST_d é a função distribuição acumulada da t-Student Assimétrica d-dimensional.

3.2.1. Dependência

A estrutura distribucional da t-Student Assimétrica em d dimensões é flexível no sentido de que se ajusta a diferentes perfis de assimetria na distribuição elíptica em qualquer de suas dimensões. Quando construída sua função de acoplamento, as diferentes estruturas de assimetria da distribuição d-dimensional são convertidas em diferentes estruturas de assimetria na dependência entre variáveis. Isso sugere a possibilidade de se utilizar a mesma função de acoplamento para ajustar diferentes níveis de assimetria da distribuição d-dimensional. No caso de dependências nas caudas pode-se ter $\lambda_{\sup}(i, j) \neq \lambda_{\inf}(i, j)$, $\lambda_{\sup}(i, j) \neq \lambda_{\sup}(i', j')$, $\lambda_{\inf}(i, j) \neq \lambda_{\sup}(i', j')$, para $(i, j) \neq (i', j')$, com

$i, j, i', j' = 1,...d$ dimensões, onde $\lambda_{\sup}$ é o coeficiente de dependência nas caudas superior e $\lambda_{\inf}$ na cauda inferior (veja Seção 3.2.2 e apêndice B).

As figuras 3.8 a 3.11 exemplificam diferentes estruturas de dependência originadas por diferentes parâmetros de assimetria e correlação da função de acoplamento, através dos contornos da função densidade e da função acumulada da distribuição C_{ST} bidimensional.

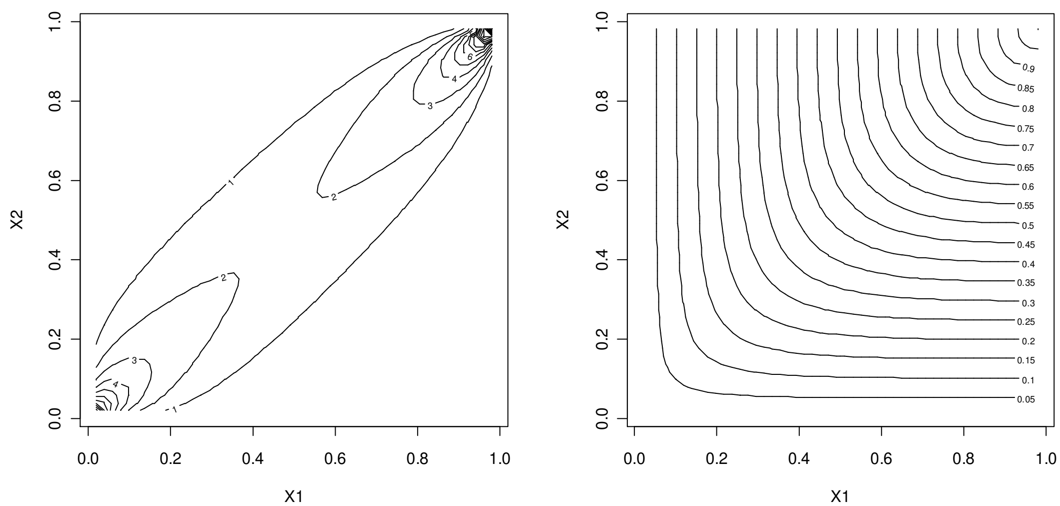

Figura 3. 8. Funções densidade e distribuição da cópula t-Student Assimétrica com parâmetros $\nu = 7$, $\rho = 0.8$ e $(\gamma_1, \gamma_2) = (0.8 , 0.8)$.

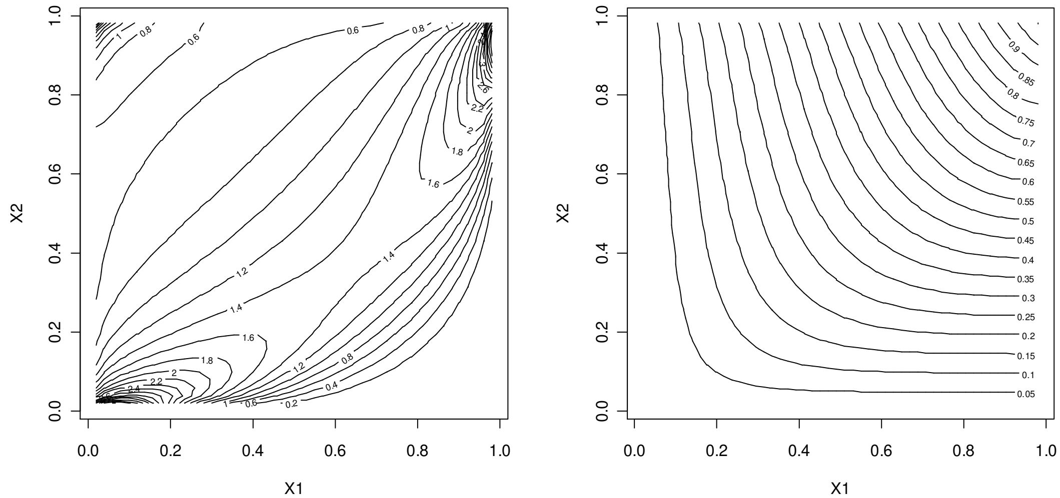

Figura 3. 9. Funções densidade e distribuição da cópula t-Student Assimétrica com parâmetros $v = 7$, $\rho = 0.8$ e $(\gamma_1, \gamma_2) =$(-0.8 , 0.8).

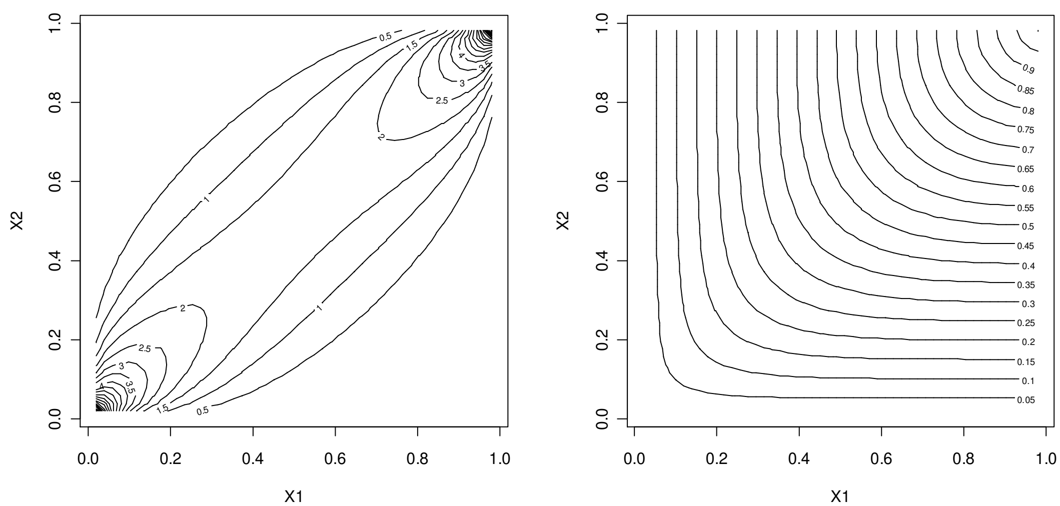

Figura 3. 10. Funções densidade e distribuição da cópula t-Student Assimétrica com parâmetros $v = 120$, $\rho = 0.8$ e $(\gamma_1, \gamma_2) =$(2 , 2).

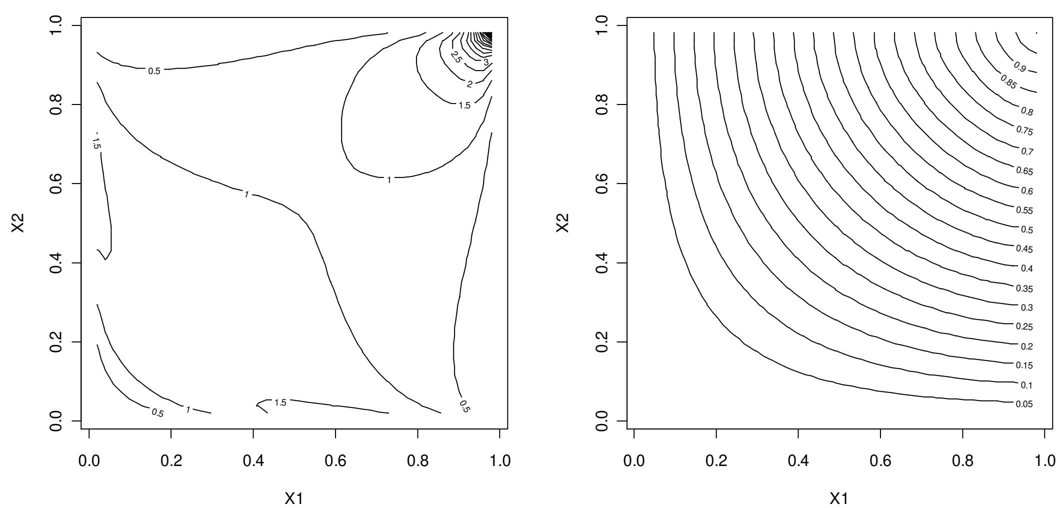

Figura 3. 11. Funções densidade e distribuição da cópula t-Student Assimétrica com parâmetros $v = 7$, $\rho = -0.8$ e $(\gamma_1, \gamma_2) = (2\ ,\ 2)$.

Em todas as figuras veja que a densidade localizada nos cantos superior direito e inferior esquerdo indicam dependências entre valores positivo-positivo e negativo-negativo. A densidade concentrada nos cantos superior esquerdo e inferior direito indica dependência entre valores positivo-negativo e negativo-positivo.

A figura 3.8 indica maior dependência na cauda superior do que na cauda inferior, considerando para essa constatação que existe maior concentração da densidade no canto superior direito do que no inferior esquerdo. 3.9 exibe dependência além dos cantos superior direito e inferior esquerdo, também no superior esquerdo.

Na figura 3.10 e 3.11, parâmetros mais extremos foram selecionados. Em 3.10 foram utilizados 120 graus de liberdade, o que, mesmo com valores significativos de assimetria, manteve simetria na densidade. 3.11 exibe dependência no canto superior direito, mesmo com correlação negativa entre as variáveis.

3.2.2. Dependência nas Caudas

No estudo da dependência nas caudas (por exemplo, para a criação de um modelo de risco), deseja-se um modelo que caracterize de maneira coerente os eventos extremos de um conjunto de dados (por exemplo, ativos financeiros interdependentes).

O fato de que muitas vezes as características da dependência em valores extremos são distintas, por exemplo, nas perdas e nos ganhos extremos de um par de ativos financeiros, sugere que distribuições conjuntas assimétricas devem ser utilizadas nesses casos.

A distribuição t-Student Assimétrica mostra-se uma opção adequada para muitos casos de dependência nas caudas, já que apresenta flexibilidade na caracterização da dependência em valores extremos nas caudas superior e inferior. A definição de dependência nas caudas é dada por:

Definição 3.7: 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 (inferior) de $(X,Y)^T$ é definido como: $\lim_{\alpha\to 1^-} P\{Y > F_2^{-1}(\alpha) / X > F_1^{-1}(\alpha)\} = \lambda_{\sup}$, $(\lim_{\alpha\to 0^+} P\{Y \le F_2^{-1}(\alpha) / X \le F_1^{-1}(\alpha)\} = \lambda_{\inf})$, $\lambda_{\sup} \in [0,1]$ ($\lambda_{\inf} \in [0,1]$), dado que o limite exista. Se $\lambda_{\sup} \in (0,1]$ ($\lambda_{\inf} \in (0,1]$), então X e Y são chamadas de assintoticamente dependentes na cauda superior (inferior). Se $\lambda_{\sup} = 0$ ($\lambda_{\inf} = 0$) X e Y são chamadas de assintoticamente independentes na cauda superior (inferior). Caso o limite não exista, a dependência na cauda não é definida.

Analiticamente, a partir da formulação teórica da dependência nas caudas e da distribuição univariada e bivariada t-Student Assimétrica é possível obter uma equação da dependência nas caudas da cópula t-Student Assimétrica. Esse resultado pode ser obtido a partir do mesmo argumento utilizado por Embrechts et al. (2002) para a distribuição t-Student Simétrica. Entretanto o cálculo da distribuição condicional necessária para o resultado final será mostrado em parte, dada a dificuldade de se trabalhar analiticamente com função Bessel modificada de segundo tipo.

O resultado é obtido inicialmente aplicando a Regra de L'Hôpital às expressões [B-14] e [B-15] do Apêndice B.3.5. Se C é a cópula de X e Y, onde $F_1 = ST(.)$ e $F_2 = ST(.)$, então obtém-se:

$$\lambda_{\inf} = \lim_{u \to 0^+} \frac{dC(u,u)}{du} = \lim_{u \to 0^+} P(U_2 \le u \,|\, U_1 = u) + \lim_{u \to 0^+} P(U_1 \le u \,|\, U_2 = u)\,, \qquad [3\text{-}34]$$

$$\begin{aligned} \lambda_{\sup} &= \lim_{u \to 1^-} \frac{\dfrac{d(1-2u+C(u,u))}{du}}{\dfrac{d(1-u)}{du}} = -\frac{2 + \lim_{u \to 1^-} \dfrac{dC(u,u)}{du}}{-1} = \\ &= 2 - \lim_{u \to 1^-} P(U_2 \le u \,|\, U_1 = u) - \lim_{u \to 1^-} P(U_1 \le u \,|\, U_2 = u), \end{aligned} \qquad [3\text{-}35]$$

onde (U_1, U_2) é um vetor aleatório cuja distribuição é C, e a segunda igualdade segue de uma propriedade da derivação de cópulas (Nelsen, 2006, Cap. 2). Suponha que se defina $Y_1 = ST^{-1}(U_1; v, 0, 1, \gamma_1)$ e $Y_2 = ST^{-1}(U_2; v, 0, 1, \gamma_2)$, de forma que $(Y1, Y2) \sim ST_2(v, \mathbf{0}, \mathbf{P}, \boldsymbol{\gamma})$. Daí,

$$\begin{aligned} &\lim_{u \to 1^-} P(U_2 \le u \,|\, U_1 = u) = \\ &\lim_{u \to 1^-} P(ST^{-1}(U_2; v, 0, 1, \gamma_2) \le ST^{-1}(u; v, 0, 1, \gamma_2) \,|\, ST^{-1}(U_1; v, 0, 1, \gamma_1) = ST^{-1}(u; v, 0, 1, \gamma_1)) \end{aligned}, \qquad [3\text{-}36]$$

$$\begin{aligned} &\lim_{u \to 0^+} P(U_1 \le u \,|\, U_2 = u) = \\ &\lim_{u \to 0^+} P(ST^{-1}(U_1; v, 0, 1, \gamma_1) \le ST^{-1}(u; v, 0, 1, \gamma_1) \,|\, ST^{-1}(U_2; v, 0, 1, \gamma_2) = ST^{-1}(u; v, 0, 1, \gamma_2)) \end{aligned}. \qquad [3\text{-}37]$$

Então, como ST^{-1} é a inversa da acumulada da t-Student Assimétrica, é, portanto, monótona. Se $y = ST^{-1}(u; v, 0, 1, \gamma_2)$, e $u \to 0^+$, implica em $y \to -\infty^+$. Análogo para o caso $u \to 1^-$, em que $y \to \infty^-$. Dessa forma,

$$\lambda_{\inf} = \lim_{y \to -\infty^+} P(Y_2 \le y \,|\, Y_1 = y) + \lim_{y \to -\infty^+} P(Y_1 \le y \,|\, Y_2 = y)\text{, e} \tag{3-38}$$

$$\lambda_{\sup} = 2 - \lim_{y \to \infty^-} P(Y_2 \le y \,|\, Y_1 = y) - \lim_{y \to \infty^-} P(Y_1 \le y \,|\, Y_2 = y)\,. \tag{3-39}$$

onde,

$$\lim_{y \to \alpha} P(Y_2 \le y \,|\, Y_1 = y) = \lim_{y \to \alpha} \int_{-\infty}^{y} f_{Y_2/Y_1=y}(x / Y_1 = y)dx\text{, e} \tag{3-40}$$

$$\lim_{y \to \alpha} P(Y_1 \le y \,|\, Y_2 = y) = \lim_{y \to \alpha} \int_{-\infty}^{y} f_{Y_1/Y_2=y}(x / Y_2 = y)dx\,. \tag{3-41}$$

Depende-se então do cálculo das densidades condicionais de $Y_1 / Y_2 = y$ e $Y_2 / Y_1 = y$. Apresenta-se a seguir o cálculo da distribuição condicional:

$$f_{Y_1/Y_2=y}(x / Y_2 = y) = \frac{f(Y_1 = x, Y_2 = y)}{f(Y_2 = y)} = \tag{3-42}$$

$$= \frac{c_1}{c_2} \frac{\dfrac{K_{\frac{v+2}{2}}\left(\sqrt{\left(v + \dfrac{x^2 - 2\rho xy + y^2}{1-\rho^2}\right)\left(\dfrac{\gamma_1^2 - 2\rho\gamma_1\gamma_2 + \gamma_2^2}{1-\rho^2}\right)}\right)}{K_{\frac{v+1}{2}}\left(\sqrt{\left(v + y^2\right)\gamma_2^2}\right)} \cdot \dfrac{\exp\left(\dfrac{\gamma_1(x-\rho y) + \gamma_2(y - \rho x)}{1-\rho^2}\right)}{\exp(y\gamma_2)}}{\dfrac{\left(\sqrt{\left(v + \dfrac{x^2 - 2\rho xy + y^2}{1-\rho^2}\right)\left(\dfrac{\gamma_1^2 - 2\rho\gamma_1\gamma_2 + \gamma_2^2}{1-\rho^2}\right)}\right)^{-\frac{v+2}{2}}}{\left(\sqrt{\left(v + y^2\right)\gamma_2^2}\right)^{-\frac{v+1}{2}}} \dfrac{\left(1 + \dfrac{x^2 - 2\rho xy + y^2}{v.(1-\rho^2)}\right)^{\frac{v+2}{2}}}{\left(1 + \dfrac{y^2}{v}\right)^{\frac{v+1}{2}}}} = \tag{3-43}$$

$$= \frac{c_1.\Omega_1}{c_2.\Omega_2} \frac{\exp\left(\dfrac{\gamma_1(x-\rho y) + \gamma_2(y-\rho x)}{1-\rho^2} - y\gamma_2\right)}{\left(1 + \dfrac{x^2 - 2\rho xy + y^2}{v.(1-\rho^2)}\right)^{\frac{v+2}{2}} \left(1 + \dfrac{y^2}{v}\right)^{-\frac{v+2}{2}} \left(1 + \dfrac{y^2}{v}\right)^{\frac{1}{2}}} = \tag{3-44}$$

$$
= \frac{c_1.\Omega_1}{c_2.\Omega_2} \frac{\exp\left(\dfrac{(x-\rho y)(\gamma_1 - \rho\gamma_2)}{1-\rho^2}\right)}{\left(\dfrac{(v+y^2).(1-\rho^2)+(x-\rho y)^2}{v.(1-\rho^2)}\right)^{\frac{v+2}{2}} \left(\dfrac{v}{v+y^2}\right)^{\frac{v+2}{2}} \left(1+\dfrac{y^2}{v}\right)^{\frac{1}{2}}} = \qquad [3\text{-}45]
$$

$$
= \frac{c_1.\Omega_1}{c_2.\Omega_2} \frac{\exp\left(\dfrac{(x-\rho y)}{\sqrt{1-\rho^2}}\left(\dfrac{v+1}{v+y^2}\right)^{\frac{1}{2}} \dfrac{(\gamma_1-\rho\gamma_2)}{\sqrt{1-\rho^2}}\left(\dfrac{v+y^2}{v+1}\right)^{\frac{1}{2}}\right)}{\left(1+\dfrac{(x-\rho y)^2}{(1-\rho^2)(v+y^2)}\right)^{\frac{v+2}{2}}\left(1+\dfrac{y^2}{v}\right)^{\frac{1}{2}}}, \qquad [3\text{-}46]
$$

onde,

$$
\Omega_1 = \frac{K_{\frac{v+2}{2}}\left(\sqrt{\left(v+\dfrac{x^2-2\rho xy+y^2}{1-\rho^2}\right)\left(\dfrac{\gamma_1^2-2\rho\gamma_1\gamma_2+\gamma_2^2}{1-\rho^2}\right)}\right)}{K_{\frac{v+1}{2}}\left(\sqrt{(v+y^2)\gamma_2^2}\right)}, \text{ e} \qquad [3\text{-}47]
$$

$$
\Omega_2 = \frac{\left(\sqrt{\left(v+\dfrac{x^2-2\rho xy+y^2}{1-\rho^2}\right)\left(\dfrac{\gamma_1^2-2\rho\gamma_1\gamma_2+\gamma_2^2}{1-\rho^2}\right)}\right)^{-\frac{v+2}{2}}}{\left(\sqrt{(v+y^2)\gamma_2^2}\right)^{-\frac{v+1}{2}}}. \qquad [3\text{-}48]
$$

O termo Ω_1 é difícil de ser tratado analiticamente, sendo necessário isolar dentro dele os termos relacionados às variáveis *x* e *y*, de forma que se obtenha uma distribuição de probabilidade final conhecida, ou que possa ser integrada e ter seu limite calculado analiticamente. c_1 e c_2 são constantes referentes a $f(Y_1 = x | Y_2 = y)$ e $f(Y_2 = y)$. O resultado analítico para $y_1 = y_2 = 0$ podem ser encontrados em Demarta & McNeil (2005).

Embora não tenha sido possível calcular o valor teórico de dependência nas caudas exatamente, pelo fato da dificuldade de encontrar os limites [3-38] e [3-39], eles podem ser obtidos através de simulação, diretamente pela definição de dependência nas caudas. Para a obtenção dos resultados, dado um conjunto de parâmetros, são simulados 1 milhão de observações e calculado o número de ultrapassagens conjuntas de limiares percentís inferiores (entre 0.001 e 0.0001) e superiores (entre 0.999 a 0.9999). A figura 3.12 exibe a dependência na cauda inferior da cópula t-Student Assimétrica para simulações realizadas com $v = 5$, $\rho = 0.7$, estando os parâmetros de assimetria (γ_1, γ_2) nos eixos.

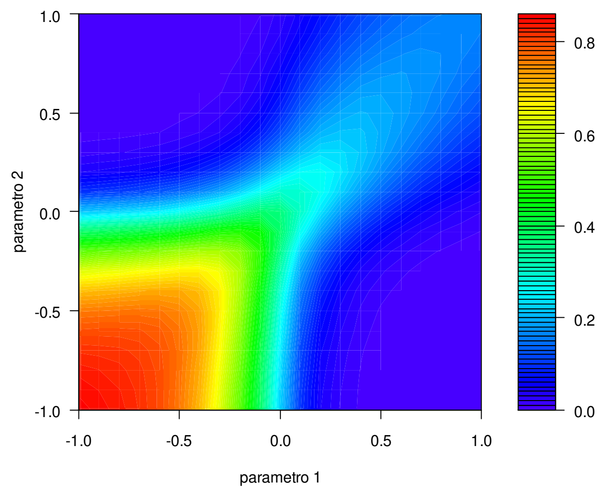

Figura 3. 12. Dependência na cauda inferior estimada através de simulação para função de acoplamento t-Student Assimétrica com $v = 5$, $\rho = 0.7$, e parâmetros de assimetria (γ_1, γ_2) nos eixos. A cor distingue dependências nas caudas baixa ou nula (azul) até alta (acima de 0.8, em vermelho).

Observa-se que, para um par de parâmetros de assimetria em (γ_1, γ_2)=(-0.5,-0.5) a dependência na cauda inferior possui valor de aproximadamente 0.75. Para 5 graus de liberdade a figura 3.13 mostra a variação causada na dependência nas caudas inferior e superior de acordo com diversos valores do parâmetro de correlação e parâmetros de assimetria.

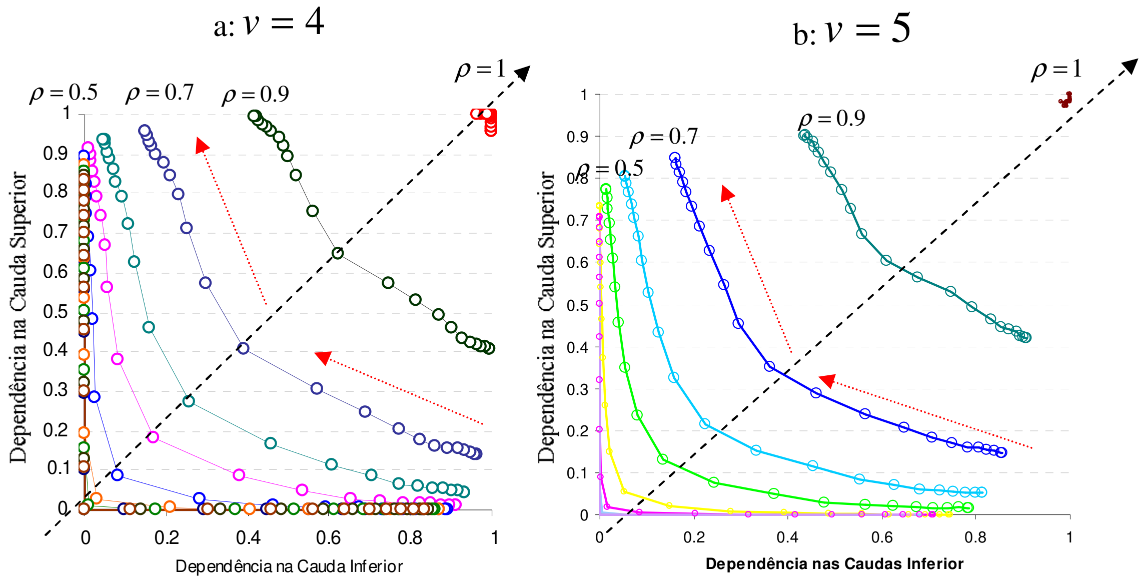

Figura 3. 13. Variação da dependência nas caudas inferior e superior por variações na correlação (seta preta tracejada), e parâmetros de assimetria (setas vermelhas pontilhadas) entre -1 e 1 em passos de 0.1 com $\gamma_1 = \gamma_2$, a: $v = 4$ e b: $v = 5$.

A figura indica que essa função de acoplamento permite grande gama de combinações de dependências nas caudas inferior e superior, chegando muito próximo da cobertura total das possibilidades de combinações com 4 graus de liberdade (até aproximadamente 0.999 para parâmetros de assimetria se aproximando de 2. A figura 3.14 exibe a variação das dependências nas caudas com parâmetro de assimetria γ_1 entre -1 e 0.8 em passos de .1 e $\gamma_2 = \gamma_1 + 0.2$

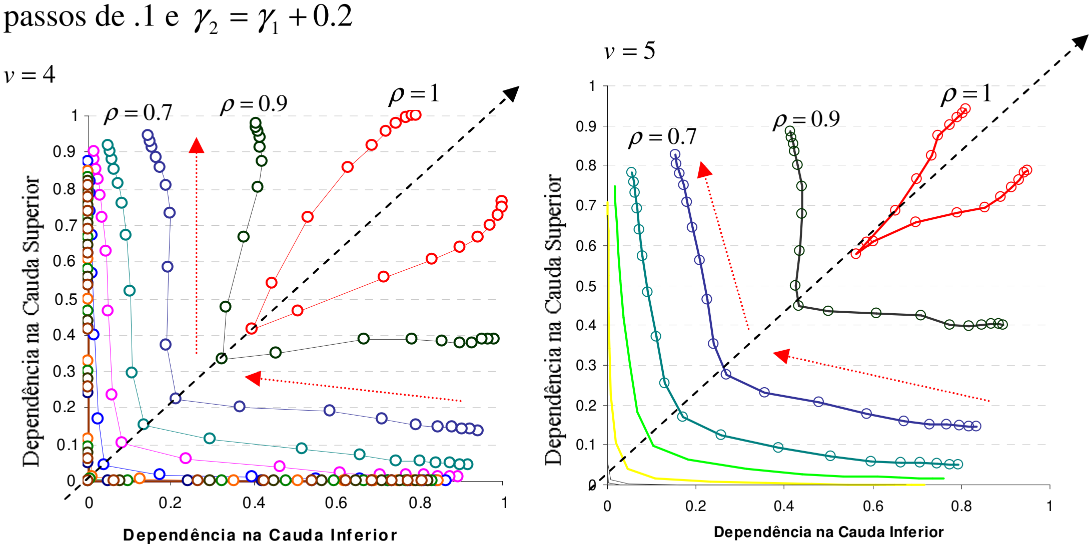

Figura 3. 14. Variação da dependência nas caudas inferior e superior por variações na correlação (seta preta tracejada), e parâmetros de assimetria (setas vermelhas pontilhadas) γ_1 entre -1 e 0.8 em passos de .1, com $\gamma_2 = \gamma_1 + 0.2$, $v = 4$ e $v = 5$

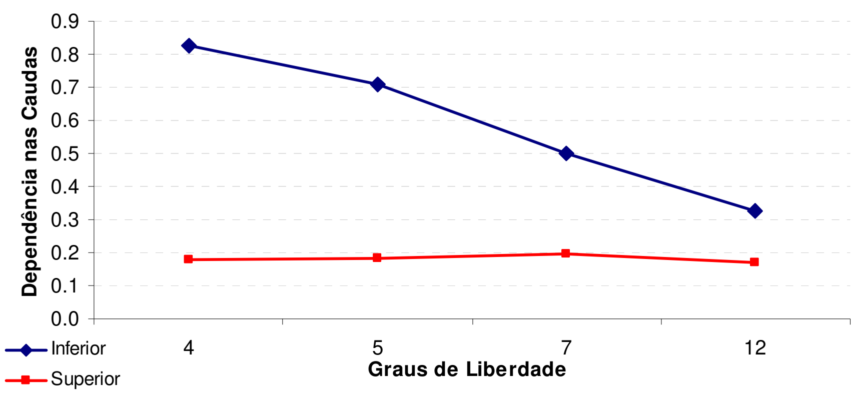

Figura 3. 15. Variação da dependência nas caudas inferior e superior através de variações no grau de liberdade, com $\rho = 0.7$ e $\gamma_1 = \gamma_2 = -0.4$.

A figura 3.15 mostra a redução na dependência nas caudas inferior e superior conforme aumenta-se o número de graus de liberdade, com correlação e parâmetros de assimetria fixados em $\rho = 0.7$ e $\gamma_1 = \gamma_2 = -0.4$. Note que o grau de liberdade é um parâmetro que influencia diretamente na dependência nas caudas, sendo determinante para sua existência. Quando $v \to \infty$ obtém-se uma cópula normal assimétrica e, portanto, a dependência nas caudas reduz-se para zero, para quaisquer parâmetros de assimetria e quaisquer parâmetros de correlação diferentes de 1.

Quando a correlação é negativa (veja exemplo da distribuição na figura 3.11 e da dependência na cauda inferior segundo diferentes parâmetros de assimetria na figura 3.16) a dependência nas caudas concentra-se em apenas uma direção, de acordo com o sinal dos parâmetros de assimetria. Nesse caso, mesmo que o parâmetro de correlação seja igual a -1 a dependência nas caudas decai para zero quando $v \to \infty$.

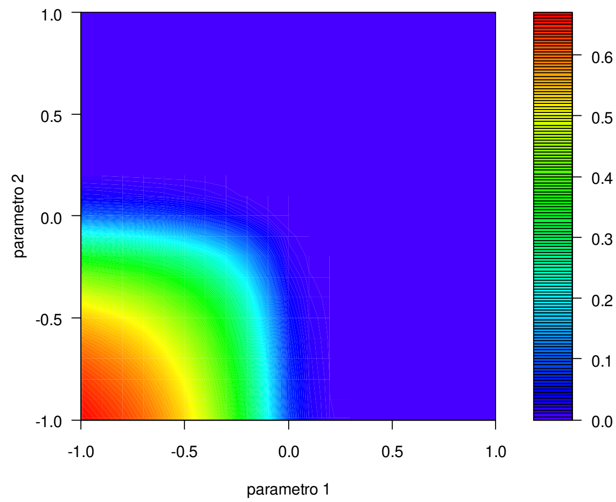

Figura 3. 16. Dependência na cauda inferior estimada através de simulação para função de acoplamento t-Student Assimétrica com $\nu = 5$, $\rho = -0{,}7$, e parâmetros de assimetria (γ_1, γ_2) nos eixos. A cor distingue dependências nas caudas baixa ou nula (azul) até alta (acima de 0,6, em vermelho).

Uma propriedade da dependência nas caudas da função de acoplamento t-Student Assimétrica é a de que a assimetria na cauda inferior é equivalente à assimetria na cauda superior quando os parâmetros de assimetria possuem sinais inversos, ou seja,

Proposição 3.2: Seja $C_{ST}(u_1, u_2, ..., u_d; \nu, \mathbf{P}, \boldsymbol{\gamma})$ uma função de acoplamento t-Student Assimétrica de d dimensões, onde **P** é uma matriz de correlações e $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, ..., \gamma_d)$ os parâmetros de assimetria. Se $\lambda_{\inf}^{i,j}(\nu, \rho_{i,j}, \gamma_i, \gamma_j)$ é o coeficiente de dependência nas caudas inferior para a marginal bidimensional (i,j) desse acoplamento e $\lambda_{\sup}^{i,j}(\nu, \rho_{i,j}, \gamma_i, \gamma_j)$ o coeficiente de dependência nas caudas superior para a marginal bidimensional (i,j) do acoplamento, então,

$$\lambda_{\inf}^{i,j}(\nu, \rho_{i,j}, \gamma_i, \gamma_j) = \lambda_{\sup}(\nu, \rho, -\gamma_i, -\gamma_j). \quad [3\text{-}49]$$

Foram calculadas as dependências nas caudas para outros valores de parâmetros além dos já citados e as principais conclusões são:

- Se os parâmetros de assimetria para um par (*i*,*j*), (γ_i, γ_j), i,j=1,...,d, $i \neq j$ são de sinais contrários, ou seja, um deles de valor negativo e o outro de valor positivo, o par possuirá dependências nas caudas não nulas para correlações positivas e grau de liberdade pequeno (4 a 6, aproximadamente), decaindo rapidamente para zero conforme se aumenta o grau de liberdade;
- Conforme o valor absoluto dos parâmetros de assimetria se aproxima de zero a dependência nas caudas se aproxima do caso da função de acoplamento t-Student Simétrica;
- Com correlação igual a 1, parâmetros de assimetria de valores iguais mostram sempre dependências nas caudas inferior e superior iguais a 1, mesmo para altos graus de liberdade;
- Diferentemente do que foi exposto em Sun *et al* (2008) o grau de liberdade é um parâmetro importante para a determinação da dependência nas caudas em uma função de acoplamento;
- Aparecem nas figuras 3.12, 3.13, 3.14 e 3.16 dependências nas caudas relativas à parâmetros de assimetria entre -1 e 1. Porém esses parâmetros podem assumir valores maiores. Com $v = 5$ e $\rho = 0.7$ obtém-se $\lambda_{\inf} = 0.087$ e $\lambda_{\sup} = 0.953$ com $\gamma_1 = \gamma_2 = 2$ e obtém-se $\lambda_{\inf} = 0.138$ e $\lambda_{\sup} = 0.9712$ com $\gamma_1 = \gamma_2 = 4$, mostrando a possibilidade de obtenção de níveis ainda mais extremos dos que os que aparecem nos gráficos, cobrindo todo o retângulo $(0,1)^2$.

3.2.3. Simulação da Função de Acoplamento

A simulação de uma variável aleatória distribuída de acordo com uma função de acoplamento será realizada através de uma adaptação do método de aceitação-rejeição (Asmussen & Glynn, 2007, Cap. 2). Esse método utiliza a função densidade da função de acoplamento.

Algoritmo de Aceitação-Rejeição para Cópulas

Seja $c(u_1, u_2)$ a função densidade de uma função de acoplamento C. A geração de um vetor aleatório de distribuição C é feita através de :

1. Gera-se 3 valores u_1, u_2 e w de distribuição Uniforme(0,1) independentes.
2. Se $w \leq c(u_1, u_2)/k$, onde $c(u,v) \leq k, \forall u, v \in [0,1]$, aceita-se os valores u_1 e u_2. Caso contrário, retorna-se a 1.
3. O vetor (u_1, u_2) é uma observação gerada da função de acoplamento C.
4. Se $F_X(x)$ e $G_Y(y)$ são funções acumuladas de distribuições de probabilidade quaisquer, então $(F_X^{-1}(u_1), G_Y^{-1}(u_2))$ possui estrutura de dependência determinada pela função de acoplamento C e marginais f e g.

Veja que é necessária a existência da função densidade da função de acoplamento para a operacionalização desse método, e que o mesmo é facilmente expandido para dimensões maiores. Especificamente para a função de acoplamento t-Student Assimétrica a função densidade é apresentada na proposição 3.2.

Proposição 3.2. Seja $C_{ST}(u_1, u_2, ..., u_d; v, \mathbf{P}, \boldsymbol{\gamma})$ como definida na equação [3-33] a densidade da função de acoplamento t-Student Assimétrica, denotada por $c_{ST}(u_1, u_2, ..., u_d; v, \mathbf{P}, \boldsymbol{\gamma})$ é dada por:

$$c_{ST}(u_1, u_2, ..., u_d; v, \mathbf{P}, \boldsymbol{\gamma}) = \frac{st_2(ST_1^{-1}(u_1; v, 0, 1, \gamma_1), ..., ST_1^{-1}(u_d; v, 0, 1, \gamma_d); v, \mathbf{P}, \boldsymbol{\gamma})}{st_1(ST_1^{-1}(u_1; v, 0, 1, \gamma_1); v, 0, 1, \gamma_1) ... st_1(ST_1^{-1}(u_d; v, 0, 1, \gamma_d); v, 0, 1, \gamma_d)} . \quad [3\text{-}50]$$

Prova: A função densidade de uma função de acoplamento C é dada por:

$$c(u_1, u_2, ..., u_d) = \frac{\partial^d C(u_1, u_2, ..., u_d)}{\partial u_1 \partial u_2 ... \partial u_d} \quad [3\text{-}51]$$

Então, pela regra da cadeia e teorema da função inversa (Stewart, 2002),

$$c_{ST}(u_1,u_2,...,u_d;v,\mathbf{P},\boldsymbol{\gamma}) = \frac{\partial^d C_{ST}(u_1,u_2,...,u_d;v,\mathbf{P},\boldsymbol{\gamma})}{\partial u_1 \partial u_2 ... \partial u_d} =$$
$$= \frac{\partial^d ST_d(ST_1^{-1}(u_1;v,0,1,\gamma_1), ST_1^{-1}(u_2;v,0,1,\gamma_2),..., ST_1^{-1}(u_d;v,0,1,\gamma_d);v,\mathbf{0},\mathbf{P},\boldsymbol{\gamma})}{\partial u_1 \partial u_2 ... \partial u_d} = \quad [3\text{-}52]$$

$$= \frac{st_2(ST_1^{-1}(u_1;v,0,1,\gamma_1),..., ST_1^{-1}(u_d;v,0,1,\gamma_d);v,\mathbf{P},\boldsymbol{\gamma})}{st_1(ST_1^{-1}(u_1;v,0,1,\gamma_1);v,0,1,\gamma_1)...st_1(ST_1^{-1}(u_d;v,0,1,\gamma_d);v,0,1,\gamma_d)}. \quad [3\text{-}53]$$

■

A figura 3.17 mostra uma simulação de 1000 valores da função de acoplamento t-Student assimétrica com $v=7$, $\rho=0.8$ e $(\gamma_1,\gamma_2)=(0.8\ ,\ 0.8)$ e $v=7$, $\rho=0.8$ e $(\gamma_1,\gamma_2)=(\text{-}0.8\ ,\ 0.8)$.

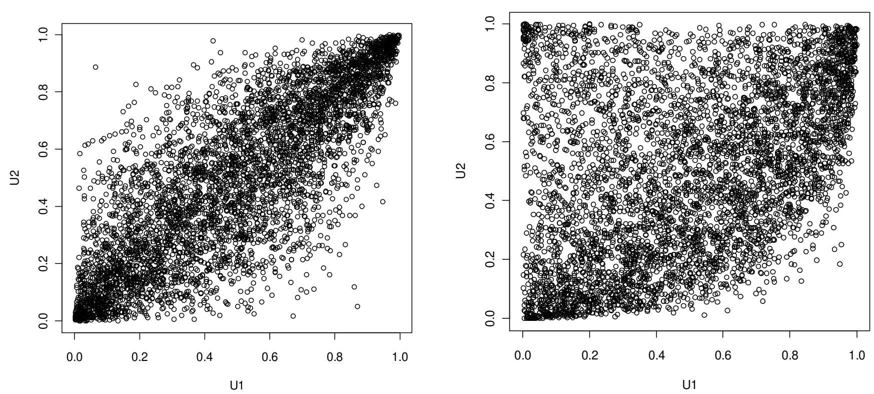

Figura 3. 17. Pontos simulados da Função de Acoplamento t-Student Assimétrica com parâmetros $v=7$, $\rho=0.8$ e $(\gamma_1,\gamma_2)=(0.8\ ,\ 0.8)$ e $v=7$, $\rho=0.8$ e $(\gamma_1,\gamma_2)=(\text{-}0.8\ ,0.8)$

Pode-se observar que, aparentemente, o método de aceitação-rejeição produz simulações condizentes com a densidade teórica, apresentadas nas figuras 3.8 e 3.9.

3.3. Estimação da Função de Acoplamento

A estimação paramétrica dessa função de acoplamento seguirá os métodos IFM ("Inference Function for Margins" – Inferência nas funções marginais) que realiza a estimação por máxima verossimilhança das marginais, seguindo distribuições paramétricas pré-determinadas, maximizando, posteriormente, a verossimilhança baseada na densidade da função de acoplamento, e o método de Máxima Verossimilhança Canônica, que utiliza a distribuição empírica nas marginais, obtendo a máxima verossimilhança da densidade da função de acoplamento.

Método IFM

Sejam $(X_1, X_2, ..., X_d)$ um vetor aleatório d-dimensional de distribuição conjunta $C_{ST}(F_{X_1}(x_2;\vartheta_1), F_{X_2}(x_2;\vartheta_2), ..., F_{X_d}(x_d;\vartheta_d); v, \mathbf{P}, \boldsymbol{\gamma})$, onde $F_{X_1}, F_{X_2}, ..., F_{X_d}$ são as distribuições marginais de $X_1, X_2, ..., X_d$ respectivamente.

Primeiro estágio: Realiza-se a estimação dos parâmetros das marginais F_{X_i}, pelo método de máxima verossimilhança:

$$\hat{\vartheta}_i = \arg\max_{\vartheta} \sum_{t=1}^{T} \ln f_{X_i}(x_{i,t};\vartheta_i)\text{, i=1,...,d.} \qquad [3\text{-}54]$$

No segundo estágio, estima-se o parâmetro $\boldsymbol{\theta} = (v, \mathbf{P}, \boldsymbol{\gamma})$ utilizando o método de máxima verossimilhança condicionado ao que foi obtido no primeiro estágio. Então,

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \sum_{t=1}^{T} \ln c_{st}(F_1(x_{1,t};\hat{\vartheta}_1), ..., F_n(x_{n,t};\hat{\vartheta}_n); \boldsymbol{\theta}). \qquad [3\text{-}55]$$

Método de Máxima Verossimilhança Canônica

Primeiro Estágio: Primeiramente usa-se a distribuição empírica para estimar as distribuições marginais (abordagem não-paramétrica, ou seja, nenhuma suposição é feita sobre as distribuições marginais).

Definição 3.7: Seja $Y_1, ..., Y_n$ uma amostra aleatória da variável Y, que tem distribuição F. A função distribuição empírica de Y é dada por

$$F_n(y) = \tfrac{1}{n} \sum_{i=1}^{n} \mathbf{1}\{Y_i \in (-\infty, y]\}. \qquad [3\text{-}56]$$

Utilizando a distribuição empírica, transforma-se o conjunto de dados $(x_{1,i}, ..., x_{n,i})$, $i = 1,2,...d$, em variáveis aleatórias uniformes $(\hat{u}_{1,i}, ..., \hat{u}_{n,i})$. Pela teoria de probabilidades,

$U_i = F_{X_i}^{-1}(X) \sim U(0,1)$, se F_{X_i} é contínua, então $\hat{u}_i = F_n^{-1}(x_i)$, $i = 1,2,...d$, em que F_n é a distribuição empírica de X.

Segundo Estágio: Estima-se o parâmetro $\boldsymbol{\theta} = (\nu, \rho, \gamma_1, \gamma_2)$ da seguinte forma:

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \sum_{t=1}^{T} \ln c_{st}(\hat{u}_{1,t},...,\hat{u}_{n,t};\boldsymbol{\theta}). \qquad [3\text{-}57]$$

Para simulações diretas da função de acoplamento bivariada obtêm-se vetores aleatórios bivariados cujas marginais são Uniformes (0,1). Nesse caso, a estimação é realizada apenas através do segundo estágio dos métodos IFM ou Máxima Verossimilhança Canônica, que nesse caso são equivalentes.

Além do exposto, o objetivo desta seção é apresentar um estudo das propriedades dos estimadores dos parâmetros da distribuição t-Student Assimétrica e da função de acoplamento t-Student Assimétrica bivariada e 3-variada. Inicialmente serão discutidos brevemente os algoritmos utilizados na estimação de parâmetros. Esse estudo é necessário pois se observaram diferentes resultados na otimização dos parâmetros através de variações, por exemplo, de valores iniciais, o que indicaria uma falta de regularidade das funções de verossimilhança empregadas. Outro fato de grande importância é o tempo de execução dos algoritmos, que é de aproximadamente 10 minutos para a otimização de uma função de acoplamento 3-variada, com precisão média (a precisão do algoritmo será discutida na seção 3.3.1 e 3.3.6).

Este estudo é organizado em: a) Comentários sobre os algoritmos auxiliares da verossimilhança e da função de verossimilhança (seção 3.3.1), b) Seleção de valores Iniciais (seção 3.3.2) c) Estimação de parâmetros das distribuições marginais (seção 3.3.3), e d) Estimação de parâmetros das funções de acoplamento Bivariada e Trivariada (seções 3.3.4 e 3.3.5). Toda a programação foi realizada através do software estatístico R-2.7.1.

3.3.1. Comentários sobre Algoritmos

As ferramentas computacionais utilizadas podem ser distribuídas em 3 classes principais: algoritmos de otimização, funções auxiliares e funções de verossimilhança. Na classe de otimização se encontram os métodos computacionais que, dada uma determinada função objetivo, encontram o ponto máximo ou mínimo da função. Os algoritmos de otimização testados foram o método Simplex (Nelder & Mead, 1965, que é a busca pelo valor ótimo baseado em um elemento geométrico que possui n+1 vértices para n dimensões), o método quasi-Newton (Byrd *et al.*, 1995, através da aproximação da função por um polinômio e ajuste da primeira e segunda derivadas para encontrar o ponto de gradiente nulo), e o método de gradientes conjugados (Fletcher & Reeves, 1964, através do gradiente em um ponto determina-se a direção de maior crescimento da função). O método de Nelder & Mead (1965) foi o que mostrou os melhores resultados dentre os métodos implementados no software estatístico R para as otimizações de parâmetros nas simulações realizadas, e seus resultados serão apresentados neste texto.

Dentro da classe das funções auxiliares estão os algoritmos que calculam as distribuições de probabilidade, funções densidade e acumulada da distribuição t-Student Assimétrica e sua função de acoplamento, as funções de geração de variáveis aleatórias e outras que fazem cálculo da função acumulada e sua inversa para a distribuição t-Student Assimétrica, dado um conjunto de parâmetros. A necessidade da programação de cada uma dessas funções adveio da ausência de pacotes que incluam essa distribuição. Foram programadas também as funções de verossimilhança, construídas a partir das funções auxiliares.

Dentre as funções auxiliares, a função que calcula a inversa da função acumulada é a que contribui mais drasticamente para o aumento no tempo de execução de uma otimização. Dado um conjunto de parâmetros da distribuição t-Student Assimétrica, essa função constrói uma matriz de duas colunas (quantís da distribuição e percentís da função acumulada). Essa matriz possui número de linhas relativo ao número de casas decimais que se deseja considerar para os quantís, e também relativo ao valor inicial e final considerado para os quantís da distribuição. Por exemplo, quantís entre -120 e 120 com passos de 0.002 implicam em uma matriz de 120000 linhas e duas colunas (x e $P(X \leq x)$, onde X é distribuída pela distribuição t-Student Assimétrica). Essa seleção garante que a soma cumulativa da probabilidade em -120 seja próxima de zero e em 120, próxima de 1.

Em uma otimização o cálculo dessa matriz deve ser realizado um número suficiente de vezes até que se encontre o parâmetro mais adequado para o conjunto de dados. Além disso, se um conjunto de dados possui β percentís, o algoritmo deve encontrar os β quantís correspondentes. O tamanho do passo (p.ex. 0.002) para a construção do vetor de quantís relacionado ao vetor de probabilidades acumuladas foi considerado a **precisão** desse algoritmo. Maior precisão implica em um vetor maior, e maior tempo de processamento. Maiores detalhes sobre precisão serão dados na seção 3.3.6.

3.3.2. Valores Iniciais

A seleção de valores iniciais para os algoritmos de otimização é de fundamental importância para que se possa encontrar a estimativa de máxima verossimilhança para as funções de acoplamento t-Student Assimétricas. Esse fato é causado pela falta de regularidade (suavidade) da função de verossimilhança do acoplamento em alguns dos parâmetros. Para contornar esse problema, no caso bivariado pode-usar um método de "grid", otimizando a máxima verossimilhança nos parâmetros de assimetria e correlação para vários graus de liberdade, distantes em algum nível um do outro (por exemplo, de 4 a 20 de 2 em 2).

Seleciona-se as 3 estimativas de maior verossimilhança e refina-se os graus de liberdade a partir do maior intervalo composto por essas 3 estimativas. Esse passo é repetido até que se obtenha um grau de liberdade com número de casas decimais desejado. Junto com ele vêm sub-estimativas dos outros parâmetros do acoplamento. Todos esses valores são então utilizados como valor inicial para a otimização completa da cópula bivariada.

Existe aí a possibilidade de que se encontre um bom valor inicial para o grau de liberdade e não para os parâmetros de assimetria (exclui-se do problema o parâmetro de correlação pelo bom comportamento deste na verossimilhança). Poderia ser necessário utilizar um método de "grid" também para esses parâmetros. Entretanto o custo computacional seria demasiado e essa alternativa foi descartada, e empiricamente, observou-se pelas simulações que a seleção de um bom valor inicial para o grau de liberdade já traz estimativas do parâmetro de assimetria válidas para serem utilizadas como valor inicial na estimação completa.

No caso trivariado a função de verossimilhança completa aumenta de 4 parâmetros (2 de assimetria, correlação e graus de liberdade) a serem estimados para 7 (adição de 1 parâmetro de assimetria e 2 de correlação). Conforme a ordem da função de acoplamento é elevada aumentam-se sempre mais 1 parâmetro de assimetria e mais (ordem - 1) parâmetros de correlação (na 4ª ordem aumentam-se mais 3 parâmetros de correlação). Pensou-se na possibilidade de usar, nas estimativas de parâmetros por máxima verossimilhança da função de acoplamento de ordem 3 ou mais, os parâmetros de correlação fixados nas correlações estimadas pelo caso bivariado. Isso significa organizar um algoritmo com a seguinte seqüência de eventos:

1) Estimar parâmetros da função de acoplamento bivariada para cada par (no caso 3-variado, são 3 pares, no caso 4-variado, são 6 pares, etc.).
2) Estimar parâmetros da função de acoplamento de ordem superior, mantendo fixos os parâmetros de correlação nos estimados em (1), com valores iniciais definidos pelas outras estimativas de parâmetros de (1).
3) Utilizar os parâmetros estimados em (2) e mais os parâmetros de correlação estimados em (1) como valores iniciais para uma estimação completa do acoplamento de ordem superior.

O segundo passo sugere que a correlação estimada pela cópula bivariada é próximo do parâmetro estimado pela função de ordens superiores. Para ilustrar esse fato a figura 3.18 mostra o resultado da estimação bivariada da correlação em 170 simulações de 2000 valores da função de acoplamento t-Student Assimétrica 3-variada contra os parâmetros estimados da cópula 3-variada.

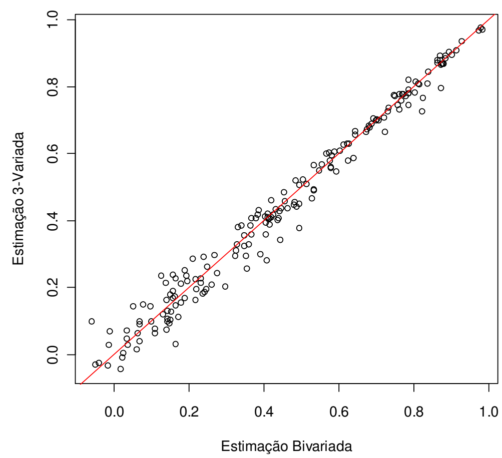

Figura 3. 18. Parâmetros de correlação verdadeiros de 170 simulações da função de acoplamento t-Student Assimétrica 3-variada contra parâmetros estimados pela cópula bivariada

Nota-se que as estimativas dos parâmetros estimados pelo acoplamento 3-variado são centradas nos estimados pelo acoplamento bivariado, embora que, para correlações baixas exista uma dispersão maior da estimativa contra os valores verdadeiros.

O uso do algoritmo descrito acima até o passo 2 já vem gerando resultados adequados na estimação de parâmetros da cópula 3-variada. Existe a possibilidade de se estudar modelos com dimensões superiores, até que se chegue no limite computacional de estimação.

3.3.3. Estimação da Distribuição t-Student Assimétrica Univariada

A estimação de parâmetros da distribuição t-Student Assimétrica univariada, que será utilizada como distribuição marginal das funções de acoplamento, como foi visto na seção 3.1.2, é realizada através dos métodos de momentos ou máxima verossimilhança. Os métodos tornam-se de difícil tratamento analítico quando são necessárias as estimações, além dos parâmetros de graus de liberdade e de assimetria, dos parâmetros de posição e escala.

Exemplo 3.1 a) Foram simuladas 1000 amostras com 1000 observações de uma distribuição t-Student assimétrica de parâmetros v=5.76 e γ = 0.68 e os parâmetros estimados pelo método de momentos e também pelo método de máxima verossimilhança (Tabela 3.1). Veja que, embora a variabilidade das estimativas seja elevada, principalmente nas estimativas pelo método de momentos, a média é próxima do valor verdadeiro, tanto nos graus de liberdade, quanto no parâmetro de assimetria.

Tabela 3. 1. Média e dispersão dos parâmetros estimados pelo método de momentos e pelo método de máxima verossimilhança para mil simulações distintas de 1000 valores da distribuição t-Student Assimétrica de parâmetros v=5.76 e γ = .68.

	Graus Liberdade		Assimetria	
	M. Momentos	M. Verossim.	M. Momentos	M. Verossim.
Verdadeiro	5.76	5.76	0.68	0.68
Média	6.0698	5.8099	0.6965	0.6824
Desvio Padrão	0.5874	0.4769	0.0377	0.0300
Erro Quad. Médio	0.4407	0.2297	0.0017	0.0009

As estimativas pelo método de momentos são menos precisas e mais dispersas do que pelo método de máxima verossimilhança, quando são estimados apenas os graus de liberdade e parâmetro de assimetria. O Exemplo 3.1 b) mostra a estimação por máxima verossimilhança dos parâmetros graus de liberdade, assimetria, posição e escala. Não são apresentadas as estimativas dos 4 parâmetros pelo método de momentos pois, através das equações [3-18] a [3-21] o método divergiu em todas as amostras realizadas.

A estimação de parâmetros de posição e escala pela máxima verossimilhança para a distribuição t-Student Assimétrica univariada, embora seja mais complexa no sentido de envolver a função de Bessel, apresentou otimizações razoáveis dos parâmetros de posição, escala, assimetria e graus de liberdade, como será apresentado.

Exemplo 3.1 b) Foram simulados 1000 amostras de 500 observações ,1000 amostras de 1000 observações, 1000 amostras de 3000 observações e mais 1000 amostras com 10000 simulações, todos com 5.76 graus de liberdade, parâmetro de assimetria no valor de 0.68, posição em 0.75 e escala em 1.75. Na tabela 3.2 observam-se os parâmetros estimados. Para cada um desses conjuntos foram estimadas as matrizes hessianas que, através da raiz de sua inversa dão uma estimativa do desvio padrão para a estimativa dos parâmetros. Serão comparados os desvios padrão da matriz hessiana com o amostral (Tabela 3.3).

Tabela 3. 2. Parâmetros estimados, desvio padrão, erro quadrático médio das estimativas por máxima verossimilhança dos parâmetros da distribuição t-Student Assimétrica univariada, com 5.76 graus de liberdade, e parâmetro de assimetria no valor de 0.68, posição em 0.75 e escala em 1.75 para 1000 replicações de 500, 1000, 3000 e 10000 observações.

	Graus de Liberdade				Assimetria			
	500	1000	3000	10000	500	1000	3000	10000
Verdadeiro	5.76	5.76	5.76	5.76	0.68	0.68	0.68	0.68
Média	5.9939	5.9852	5.7833	5.7615	0.7173	0.7143	0.6839	0.6807
Mediana	5.8402	5.8639	5.7309	5.7706	0.6929	0.6980	0.6756	0.6822
Desvio Padrão	1.3634	1.0351	0.6726	0.4078	0.2091	0.1731	0.1315	0.0769
Erro Quad. Médio	1.9118	1.1210	0.4525	0.1661	0.0451	0.0311	0.0173	0.0059

	Posição				Escala			
	500	1000	3000	10000	500	1000	3000	10000
Verdadeiro	0.75	0.75	0.75	0.75	1.75	1.75	1.75	1.75
Média	0.7033	0.7127	0.7474	0.7505	1.7396	1.7440	1.7474	1.7493
Mediana	0.7187	0.7225	0.7576	0.7507	1.7371	1.7443	1.7500	1.7507
Desvio Padrão	0.2414	0.1960	0.1494	0.0869	0.0848	0.0598	0.0362	0.0220
Erro Quad. Médio	0.0604	0.0398	0.0223	0.0075	0.0073	0.0036	0.0013	0.0005

Para uma amostra relativamente pequena (500 observações) observou-se que existe um afastamento da média das estimativas do grau de liberdade e parâmetro de assimetria com relação aos parâmetros utilizados nas simulações. Entretanto, no caso do parâmetro de graus de liberdade a mediana mostra um afastamento menos acentuado, o que é uma indicativa de que o procedimento de otimização para pequeno número de observações pode gerar estimativas muito distantes do valor verdadeiro. De fato, como se pode observar na figura 3.19 e 3.20, que mostram a dispersão das estimativas do grau de liberdade para 500, 1000, 3000 e 10000 observações, as estimativas podem assumir valores muito acima do verdadeiro para poucas informações, reduzindo a variabilidade na estimação conforme se aumenta o número de observações. Observa-se também que com um número maior de observações simuladas obtêm-se estimativas mais precisas e de menor variabilidade para todos os parâmetros.

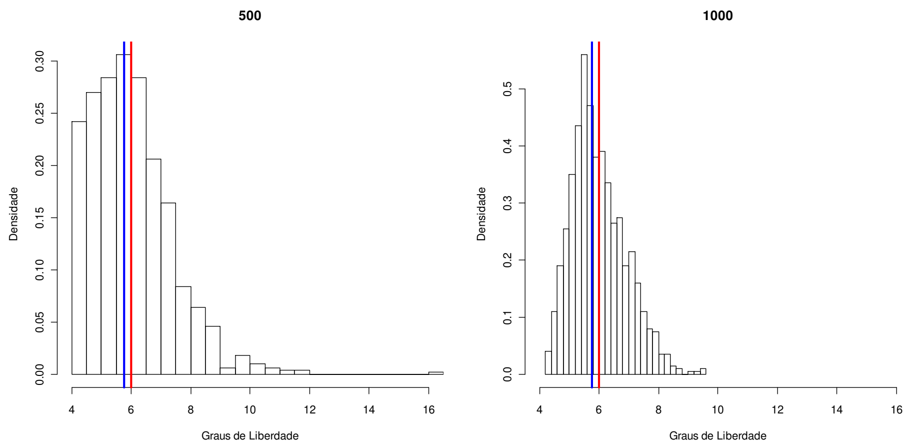

Figura 3. 19. Graus de liberdade estimados em 1000 simulações de 3000 observações da distribuição t-Student Assimétrica com 5.76 graus de liberdade (linha azul), parâmetro de assimetria no valor de 0.68, posição em 0.75 e escala em 1.75. A linha vermelha é a média das estimativas.

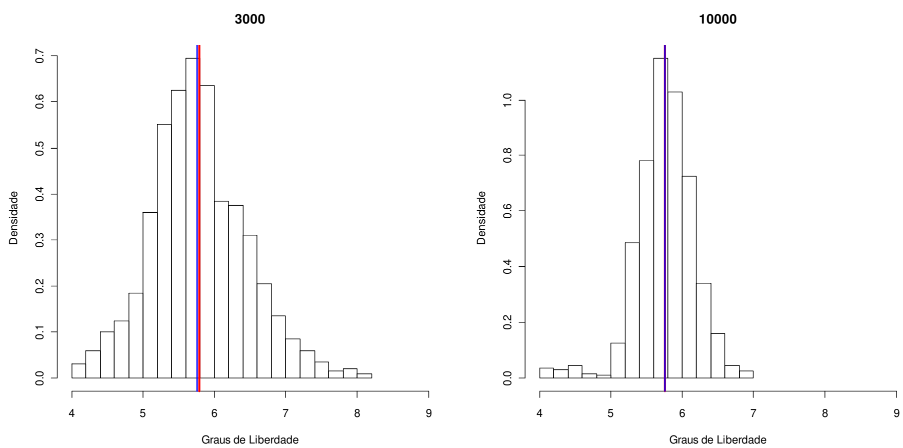

Figura 3. 20. Histograma dos graus de liberdade estimados em 1000 simulações de 3000 e 10000 observações da distribuição t-Student Assimétrica com 5.76 graus de liberdade (linha azul), parâmetro de assimetria no valor de 0.68, posição em 0.75 e escala em 1.75. A linha vermelha é a média das estimativas.

A tabela 3.3 mostra o desvio padrão médio estimado através da matriz hessiana, juntamente com medidas de variabilidade do desvio padrão estimado para as simulações realizadas com 500, 1000, 3000 e 10000 observações.

Tabela 3. 3. Desvio padrão médio estimado através do hessiano das estimações de parâmetros de 1000 simulações de 500, 1000, 3000 e 10000 observações da distribuição t-Student Assimétrica univariada, com 5.76 graus de liberdade, e parâmetro de assimetria no valor de 0.68, posição em 0.75 e escala em 1.75.

	Graus de Liberdade				Assimetria			
	500	1000	3000	10000	500	1000	3000	10000
Amostral	1.3634	1.0351	0.6726	0.4078	0.2091	0.1731	0.1315	0.0769
Média	1.3303	1.0014	0.5798	0.3195	0.1926	0.1736	0.1159	0.0662
Mediana	1.1953	0.9492	0.5738	0.3203	0.2157	0.1961	0.1197	0.0671
Desvio Padrão	0.6973	0.3304	0.1325	0.0456	0.1539	0.0943	0.0361	0.0108

	Posição				Escala			
	500	1000	3000	10000	500	1000	3000	10000
Amostral	0.2414	0.1960	0.1494	0.0869	0.0848	0.0598	0.0362	0.0220
Média	0.2268	0.2016	0.1317	0.0756	0.0856	0.0608	0.0353	0.0194
Mediana	0.2640	0.2285	0.1385	0.0772	0.0853	0.0607	0.0353	0.0194
Desvio Padrão	0.1658	0.1002	0.0428	0.0139	0.0069	0.0027	0.0010	0.0004

Observe a redução da média e mediana do desvio padrão quando aumenta-se o número de observações e que os desvios padrão das estimações com 500 observações são de 3 a 5 vezes maiores do que com 3000 ou 10000 observações. A aproximação normal, conjuntamente com o Hessiano são utilizados para construir intervalos de confiança para as estimativas. Através de simulações será analisado se essa aproximação é válida. Será verificado, para um determinado nível, o percentual de intervalos de confiança construídos que cobrem o valor verdadeiro comparando-se com o percentual nominal (ou seja, se foi construído um intervalo de confiança de 95%, espera-se que em aproximadamente 95% das simulações o valor verdadeiro esteja dentro desse intervalo). A tabela 3.4 mostra os percentuais observados para intervalos de confiança de níveis 90%, 95%, 97.5% e 99%.

Tabela 3. 4. Percentual em que o intervalo de confiança de níveis 90%, 95%, 97.5% e 99% abrangeu o verdadeiro parâmetro nas estimativas realizadas em 1000 simulações de 500, 1000, 3000 e 10000 observações da distribuição t-Student Assimétrica univariada, com 5.76 graus de liberdade, e parâmetro de assimetria no valor de 0.68, posição em 0.75 e escala em 1.75.

	Graus de Liberdade				Assimetria			
	500	1000	3000	10000	500	1000	3000	10000
90.00%	93.24%	94.09%	91.66%	87.56%	98.40%	98.43%	93.30%	90.54%
95.00%	96.44%	96.06%	96.05%	93.63%	98.93%	99.21%	96.93%	95.07%
97.50%	99.29%	98.03%	97.48%	97.23%	99.47%	100.00%	98.13%	97.64%
99.00%	99.47%	98.82%	98.79%	98.56%	99.64%	100.00%	99.45%	99.08%

	Posição				Escala			
	500	1000	3000	10000	500	1000	3000	10000
90.00%	99.47%	100.00%	92.65%	90.44%	93.06%	91.08%	91.77%	89.93%
95.00%	99.47%	100.00%	97.59%	95.89%	96.62%	95.01%	96.27%	95.07%
97.50%	99.47%	100.00%	99.23%	98.25%	98.75%	97.64%	97.91%	97.33%
99.00%	99.47%	100.00%	100.00%	98.87%	99.29%	98.82%	99.23%	99.18%

A tabela 3.4 indica que os intervalos de confiança foram sobreestimados para amostras com 500 e 1000 observações. Possivelmente a matriz hessiana ficou subestimada por conta do baixo número de observações nas simulações. Para um número maior de observações nota-se que a existência do valor verdadeiro dos parâmetros dentro dos intervalos de confiança das estimativas, de acordo com os níveis estabelecidos.

3.3.4. Estimação da Distribuição t-Student Assimétrica de 2 e 3 Dimensões

Exemplo 3.2 a) Foram simulados 3000 valores de uma distribuição t-Student Assimétrica de parâmetros $\nu = 8.35$, $\rho = 0.77$ e assimetria $(\gamma_1, \gamma_2) = (0.67\ ,\ -0.32)$. A figura 3.21 mostra a dispersão dos pontos simulados, juntamente com os contornos da função densidade teórica com esses parâmetros.

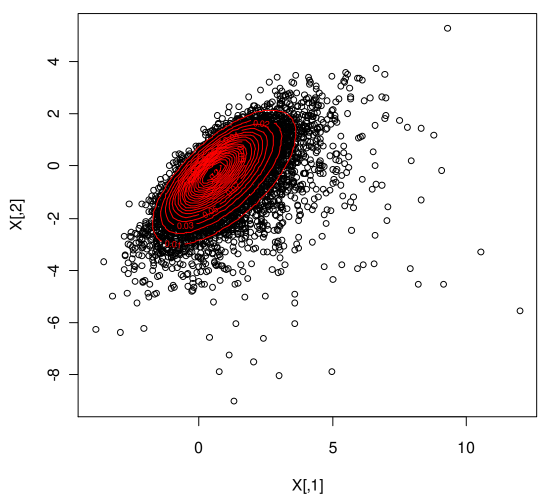

Figura 3. 21. Dispersão de pontos simulados e contornos da Distribuição t-Student Assimétrica com parâmetros $\nu = 8.35$, $\rho = 0.77$ e assimetria $(\gamma_1, \gamma_2) = (0.67\ ,\ -0.32)$.

Como feito no caso univariado, serão estimados os parâmetros através dos dados simulados pelos métodos de momentos e máxima verossimilhança, para 1000 replicações da simulação acima. Os parâmetros estimados por esses métodos são exibidos na tabela 3.5.

Tabela 3. 5. Parâmetros estimados pelo método de momentos generalizado e pelo método de máxima verossimilhança para 1000 amostras aleatórias distintas de 3000 valores da distribuição t-Student Assimétrica bivariada de parâmetros $\nu = 8.35$, $\rho = 0.77$ e assimetria $(\gamma_1, \gamma_2) = (0.67\ , -0.32)$.

	Graus de Liberdade		Assimetria 1		Assimetria2		Correlação	
	M.V.	M.M.G.	M.V.	M.M.G.	M.V.	M.M.G.	M.V.	M.M.G.
Verdadeiro	8.35	8.35	0.67	0.67	-0.32	-0.32	0.77	0.77
Média	8.3899	8.4319	0.6712	0.6696	-0.3180	-0.3194	0.7697	0.7721
Desvio Padrão	0.6840	0.9721	0.0314	0.0351	0.0286	0.0283	0.0167	0.0595
Erro Quad. Médio	0.4690	0.9507	0.0010	0.0012	8.23E-04	8.00E-04	2.79E-04	3.54E-03

M.V.: Máxima Verossimilhança
M.M.G.: Método de Momentos Generalizado

Novamente, maior variabilidade na estimação dos parâmetros é observada nas estimativas pelo método de momentos principalmente na estimativa do grau de liberdade. Mesmo assim, ao cabo de 1000 simulações a média da estimação conjunta dos 4 parâmetros da distribuição foi próxima entre os dois métodos de estimação. Importante ressaltar que o uso do método de momentos através das funções do 1º e 2º momentos só é válido se os dados não forem padronizados, caso contrário esse método retorna valores triviais (4 graus de liberdade, parâmetros de assimetria e correlação nulos), já que estão sendo desprezados os parâmetros de posição e escala. A correlação entre as estimativas pelo método de momentos e as estimativas por máxima verossimilhança neste caso são de 0.85, para os graus de liberdade.

Exemplo 3.2 b) Para exemplificar o uso da distribuição t-Student Assimétrica em maiores dimensões foram realizadas 1000 replicações de amostras aleatórias de tamanho 3000 da distribuição t-Student Assimétrica de 3 dimensões. As médias e dispersão das estimativas por máxima verossimilhança dos parâmetros pelo algoritmo trivariado e pelo algoritmo bivariado figuram nas tabelas 3.6 e 3.7.

Tabela 3. 6. Parâmetros estimados pelo método de máxima verossimilhança para 1000 replicações de 3000 amostras da distribuição t-Student Assimétrica 3-variada de parâmetros $\nu = 6.23$, $\rho_{1,2} = 0.71$, $\rho_{1,3} = 0.33$, $\rho_{2,3} = 0.23$ e assimetria $(\gamma_1, \gamma_2, \gamma_3) = (0.78\ , -0.67\ , -0.43)$.

	g.l.	Assim. 1	Assim. 2	Assim. 3	Corr. 1,2	Corr. 1,3	Corr. 2,3
Verdadeiro	6.23	0.78	-0.67	-0.43	0.71	0.33	0.23
Média	6.283	0.807	-0.702	-0.431	0.734	0.328	0.230
Desvio Padrão	0.474	0.029	0.029	0.028	0.028	0.035	0.032
Erro Quad. Médio	2.31E-01	2.85E-03	3.84E-03	7.94E-04	1.97E-03	1.24E-03	1.05E-03

Tabela 3. 7. Parâmetros estimados pelo método de máxima verossimilhança para 1000 simulações distintas de 3000 valores da distribuição t-Student Assimétrica 3-variada de parâmetros $\nu = 6.23$, $\rho_{1,2} = 0.71$, $\rho_{1,3} = 0.33$, $\rho_{2,3} = 0.23$ e assimetria $(\gamma_1, \gamma_2, \gamma_3) =$ (0.78 , -0.67 , -0.43). Estimação para marginais bidimensionais.

	Marginal Bidimensional 1				Marginal Bidimensional 2			
	g.l.	Assim. 1	Assim. 2	Corr. 1,2	g.l.	Assim. 1	Assim. 3	Corr. 1,3
Verdadeiro	6.23	0.78	-0.67	0.71	6.23	0.78	-0.43	0.33
Média	6.289	0.858	-0.764	0.757	6.285	0.782	-0.431	0.329
Desvio Padrão	0.488	0.024	0.027	0.035	0.451	0.032	0.028	0.036
Erro Quad. Médio	3.27E-01	6.60E-03	9.61E-03	3.44E-03	2.06E-01	1.03E-03	8.01E-04	1.32E-03

	Marginal Bidimensional 3			
	g.l.	Assim. 2	Assim. 3	Corr. 2,3
Verdadeiro	6.23	-0.67	-0.43	0.23
Média	6.290	-0.670	-0.430	0.230
Desvio Padrão	0.533	0.033	0.029	0.034
Erro Quad. Médio	2.87E-01	1.06E-03	8.31E-04	1.13E-03

Dessas tabelas é interessante notar que as estimativas dos parâmetros das marginais bivariadas são próximas das estimativas dos parâmetros obtidas de forma completa (3-variada). Esse fato será utilizado no contexto de cópulas, principalmente no caso dos parâmetros de correlação, para permitir o aumento do número de dimensão sem aumentar demasiadamente o custo computacional. A correlação entre as estimativas do grau de liberdade no caso 3-variado e as estimativas no caso bivariado são em torno de 0.75.

O objetivo de se estudar a estimação de parâmetros da distribuição t-Student Assimétrica de dimensões maiores do que 1 é dar uma base para a criação de métodos computacionais para a estimação de parâmetros da função de acoplamento baseada nessa distribuição, além de mostrar praticamente propriedades da distribuição, como a hereditariedade dos parâmetros da distribuição 3-variada para suas marginais bivariadas e univariadas.

3.3.5. Estimação de Parâmetros do Acoplamento Bivariado

Exemplo 3.3 a) A tabela 3.8 mostra a média e dispersão das estimativas dos parâmetros de 200 simulações de 3000 observações da função de acoplamento t-Student Assimétrica com parâmetros $\nu = 6.85$, $\rho = 0.81$ e $(\gamma_1, \gamma_2) = (-0.78, -0.67)$.

Tabela 3. 8. Parâmetros estimados por máxima verossimilhança de 200 simulações distintas de 3000 valores da função de acoplamento t-Student Assimétrica bivariada de parâmetros $\nu = 6.85$, $\rho = 0.81$ e $(\gamma_1, \gamma_2) = (-0.78, -0.67)$.

	g.l.	Assim. 1	Assim. 2	Correlação
Verdadeiro	6.85	-0.78	-0.67	0.81
Média	6.978	-0.726	-0.628	0.809
Desvio Padrão	0.115	0.071	0.080	0.011
Erro Quad. Médio	2.98E-02	7.93E-03	8.14E-03	1.29E-04

Veja que na tabela ainda não aparecem menções a respeito de resultados distintos dos métodos IFM ou Máxima Verossimilhança Canônica. Isso é devido ao fato de que as simulações obtidas pelo método de aceitação-rejeição já são percentís de uma distribuição marginal, ou seja, são quantís da distribuição Uniforme (0,1), que são diretamente aplicados ao algoritmo de otimização de parâmetros da função de acoplamento. Isso significa que o exemplo de otimização exposto é relativo a simulações da própria distribuição (conjunta de marginais uniformes) que é a função de acoplamento. Observe que os parâmetros de assimetria foram estimados em média abaixo do valor verdadeiro, e que a correlação ficou próxima do valor real.

O próximo exemplo utiliza simulações da função de acoplamento, porém com a transformação das marginais uniformes para as distribuições marginais Normal-Assimétricas e t-Student Assimétricas, com posterior estimação dos parâmetros pelos métodos IFM e Máxima Verossimilhança Canônica. Em princípio espera-se que, se a estimação dos parâmetros das distribuições marginais, no caso do método IFM, estiver correta, obter, após a transformada integral de probabilidade, percentís próximos aos simulados, dessa forma obter-se-ia, para um número suficiente de simulações, estimativas dos parâmetros condizentes com os parâmetros utilizados para a simulação.

Exemplo 3.3 b): Foram realizadas 100 amostras de 3000 observações da função de acoplamento t-Student Assimétrica com parâmetros $\nu = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$, sendo as marginais Uniforme(0,1) transformadas na distribuição Normal-Assimétrica com parâmetros $\mu = 0.43$, $\sigma = 0.81$ e $\alpha = 2.21$, e na distribuição t-Student Assimétrica com parâmetros $\nu = 5.35$, $\gamma = 0.86$. As tabelas 3.9 e 3.10 exibem estimações dos parâmetros das marginais e da função de acoplamento pelo método IFM e da função de acoplamento pelo método de Máxima Verossimilhança Canônica.

Tabela 3. 9. Parâmetros estimados pelo método IFM para 100 simulações distintas de 3000 valores da função de acoplamento t-Student Assimétrica bivariada de parâmetros $v = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$ com marginal 1: Normal-Assimétrica com parâmetros $\mu = 0.43$, $\sigma = 0.81$ e $\alpha = 2.21$, e marginal 2: t-Student Assimétrica com parâmetros $v = 5.35$, $\gamma = 0.86$.

	Marginal Normal Assimet.			Marg. T-Student Assimet.	
	Posição	Escala	Assimetria	G.l.	Assimetria
Verdadeiro	0.43	0.81	2.21	5.35	0.86
Média	0.4221	0.8152	2.4013	5.4988	0.8739
Desvio Padrão	0.0135	0.0156	0.1289	0.2627	0.0218
Erro Quad. Médio	0.0002	0.0003	0.0529	0.0898	0.0007

	Função de Acoplamento t-Student Assimétrica			
	Assimetria 1	Assimetria 2	Correlação	Graus Lib.
Verdadeiro	-0.48	-0.72	0.65	7.21
Média	-0.4178	-0.6609	0.6524	7.6335
Desvio Padrão	0.0965	0.1253	0.0177	0.9890
Erro Quad. Médio	0.0130	0.0189	0.0003	1.1378

Tabela 3. 10. Parâmetros estimados pelo método de máxima verossimilhança canônica para 100 simulações distintas de 3000 valores da função de acoplamento t-Student Assimétrica bivariada de parâmetros $v = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$ com marginal 1: Normal-Assimétrica com parâmetros $\mu = 0.43$, $\sigma = 0.81$ e $\alpha = 2.21$, e marginal 2: t-Student Assimétrica com parâmetros $v = 5.35$, $\gamma = 0.86$.

	Função de Acoplamento t-Student Assimétrica			
	Assimetria 1	Assimetria 2	Correlação	Graus Lib.
Verdadeiro	-0.48	-0.72	0.65	7.21
Média	-0.4059	-0.6856	0.6588	8.6681
Desvio Padrão	0.1268	0.1680	0.0171	1.6097
Erro Quad. Médio	0.0212	0.0288	0.0004	4.6654

Novamente observa-se a subestimação dos parâmetros pelo método de máxima verossimilhança canônica, indicando que a seleção do método para estimação influencia nos resultados obtidos. Em geral o grau de liberdade foi estimado com maior precisão e menor variabilidade pelo método IFM.

Foram estimadas as matrizes hessianas com o intuito de calcular o desvio padrão das estimativas dos parâmetros do acoplamento baseado na distribuição t-Student Assimétrica. A tabela 3.11 mostra os desvios padrão estimados através da matriz hessiana para os métodos IFM e Máxima Verossimilhança Canônica. Aparecem também na tabela os desvios padrão calculados para as 100 estimações de parâmetros.

Tabela 3. 11. Desvio padrão médio estimado através do hessiano das estimações de parâmetros de 100 simulações de 3000 observações da distribuição t-Student Assimétrica bivariada de parâmetros $\nu = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$.

	Assimetria 1	Assimetria 2	Correlação	Graus Lib.
Amostral	0.0965	0.1253	0.0177	0.9890
Médio	0.0068	0.0075	0.0101	0.0092
Mediano	0.0063	0.0066	0.0100	0.0086
Desvio Padrão	0.0021	0.0035	0.0005	0.0028

Observe que o desvio padrão médio calculado através do hessiano obtido automaticamente pelo algoritmo de otimização para cada simulação é muito inferior do que o desvio padrão das estimativas. Isso sugere que a função de verossimilhança não é uma função suave para pequenas variações (deltas) dos parâmetros. Isso significa que, para chegar em uma estimativa que maximize a função de verossimilhança devem-se considerar passos entre estimativas de tamanhos suficientes, de forma a evitar a ocorrência de otimizações em pontos máximos locais (e buscar o máximo global).

Em termos de estimativa do desvio padrão descobriu-se que os algoritmos de otimização da função de verossimilhança utilizavam, no método numérico de cálculo das segundas derivadas da função, definido na função [3-58] o mesmo *h* para todos os parâmetros, no valor de 0.001.

$$\mathrm{H}_{i,j}(\theta) = \begin{cases} \dfrac{L(\theta_i + 2h_i, \cdots; x) - 2L(\theta; x) - L(\theta_i - 2h_i, \cdots; x)}{4h_i^2}, & \text{se} \quad \mathrm{i = j} \\ \\ \dfrac{L(\theta_i + h_i, \theta_j + h_j, \cdots; x) - L(\theta_i + h_i, \theta_j - h_j, \cdots; x) - L(\theta_i - h_i, \theta_j + h_j, \cdots; x) + L(\theta_i - h_i, \theta_j - h_j, \cdots; x)}{4h_i h_j} & \text{se} \quad \mathrm{i \neq j} \end{cases} \quad .[3\text{-}58]$$

Esse fato manteve a estimativa do desvio padrão muito longe da amostral. Além disso a chance de que um intervalo de confiança para os parâmetros construído com esses desvios padrão não inclua o parâmetro real tornou-se muito alta. Decidiu-se então por uma análise para definir o melhor *h* para cada parâmetro. Para o cálculo final da matriz hessiana considerou-se a função de verossimilhança calculada em um ponto θ através da média da função avaliada nos pontos $\theta, \theta-\varepsilon$ e $\theta+\varepsilon$, reduzindo o impacto da falta de suavidade da função de verossimilhança.

A análise do *h* utilizou gráficos de funções de verossimilhança para diversas simulações, em cada um dos parâmetros, e em escalas (deltas) diferentes. As figura 3.22 a 3.24 mostram a função de verossimilhança quando é variado um dos parâmetros, mantendo os demais fixos, para uma das simulações da função de acoplamento do exemplo 3.5. Observe que quando a variação é realizada com passos (*h*) pequenos os resultados da função de verossimilhança tornam-se irregulares. Dessa forma o desvio padrão calculado através da segunda derivada (hessiano) torna-se pequeno e irregular.

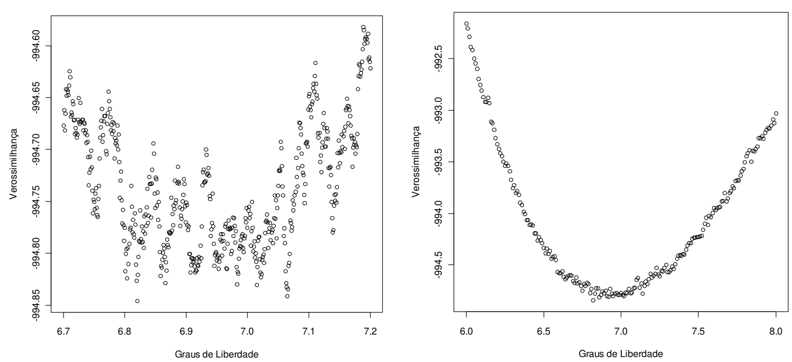

Figura 3. 22. Função de verossimilhança calculadas em uma simulação da função de acoplamento com $v = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$. Com variação de a) graus de liberdade em passos de .001 e b) graus de liberdade em passos de .01.

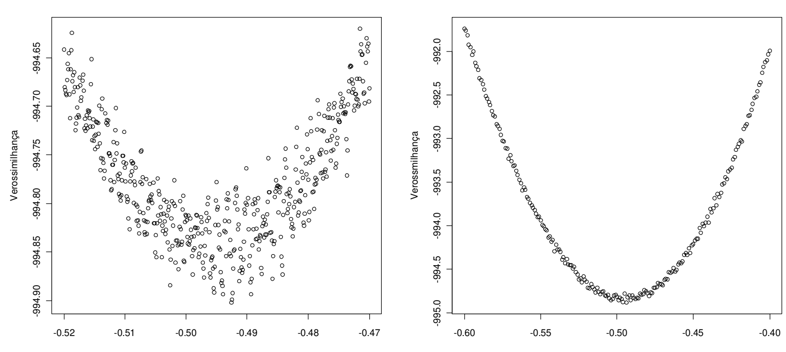

Figura 3. 23. Função de verossimilhança calculadas em uma simulação da função de acoplamento com $v = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$. Com variação de a) Parâmetro de Assimetria 1 (-0.48) em passos de .001 e b) Parâmetro de Assimetria 1 (-0.48) em passos de .01

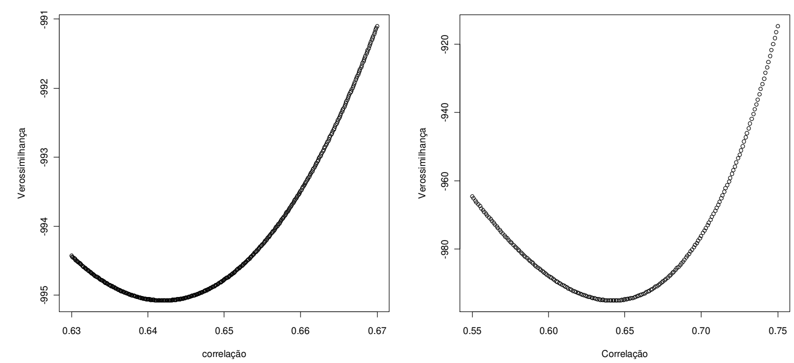

Figura 3. 24. Função de verossimilhança calculadas em uma simulação da função de acoplamento com $v = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$. Com variação de a) graus de liberdade em passos de .0001 e b) graus de liberdade em passos de .001.

Com h maiores, passando de 0.001 para 0.01 ou 0.1, no caso dos graus de liberdade, obteve-se uma função de verossimilhança mais regular, cuja segunda derivada produzirá resultados mais adequados de dispersão para os parâmetros de assimetria e graus de liberdade. Para a correlação a função de verossimilhança é regular e suave, facilitando a maximização da verossimilhança nesse parâmetro e cálculo da segunda derivada. Por esse motivo, como visto na seção 3.3.2, os parâmetros de correlação estimados através da função de acoplamento bivariada para cada par de um conjunto de dados multivariado podem ser fixados na função de verossimilhança multidimensional, o que reduz o número de parâmetros a serem estimados em maiores dimensões. A tabela 3.12 mostra os h's e ε selecionados durante a análise.

Tabela 3. 12. Diferenciais utilizados no cálculo numérico da matriz hessiana para cada parâmetro da função de acoplamento t-Student Assimétrica

	Assimetria 1	Assimetria 2	Correlação	Graus Lib.
h	0.01	0.01	0.001	0.1
ε	0.00001	0.00001	0.00001	0.00001

O mesmo valor de ε mostrou-se suficiente para o cálculo do hessiano em todos os parâmetros. A tabela 3.13 mostra estimativas do desvio padrão através do hessiano calculado com esses diferenciais.

Tabela 3. 13. Desvio padrão médio estimado através do hessiano das estimações de parâmetros de 100 simulações de 3000 observações da distribuição t-Student Assimétrica bivariada de parâmetros $\nu = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$

	Assimetria 1	Assimetria 2	Correlação	Graus Lib.
Amostral	0.0965	0.1253	0.0177	0.9890
Médio	0.0613	0.0792	0.0100	0.5468
Mediano	0.0486	0.0648	0.0083	0.4600
Desvio Padrão	0.0342	0.0306	0.0055	0.1375

Note agora que o desvio padrão calculado pela matriz hessiana encontra-se mais próximo do desvio padrão amostral das estimativas dos parâmetros comparado com o caso h=0.001, mas encontra-se ainda longe do desvio padrão amostral, podendo ainda gerar intervalos de confiança que não incluam o parâmetro verdadeiro, principalmente nos graus de liberdade.

Foram utilizados esses desvios padrão para a determinação de intervalos de confiança de 90%, 95%, 97.5% e 99% das estimativas dos parâmetros. Com esses intervalos de confiança, como feito para o caso marginal, foi verificada a proporção do número de séries em que o parâmetro verdadeiro se encontra dentro dos intervalos de confiança. Os resultados são apresentados na tabela 3.14

Tabela 3. 14. Percentual em que o intervalo de confiança de níveis 90%, 95%, 97.5% e 99% abrangeu o verdadeiro parâmetro nas estimativas realizadas em 100 simulações de 3000 observações da função de acoplamento t-Student Assimétrica com $\nu = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$.

	90%	95%	97.50%	99%
Assimetria 1	88.00%	95.00%	98.00%	100.00%
Assimetria 2	86.00%	97.00%	99.00%	100.00%
Correlação	84.00%	95.00%	95.00%	98.00%
Graus Lib.	89.00%	91.00%	92.00%	95.00%

Veja que os intervalos de confiança não cobriram o verdadeiro parâmetros em muitas das simulações, principalmente o grau de liberdade. É possível que sejam necessários h's ainda maiores para melhorar o cálculo das segundas derivadas.

Foram realizados ajustes de modelos polinomiais quadráticos à função de verossimilhança calculada em um *"grid"* em torno da estimativa de máxima verossimilhança. O ajuste mostrou-se adequado e o Hessiano da função quadrática ajustada ficou próximo ao Hessiano estimado pelo método proposto. Esse resultado auxilia na validação do método proposto.

3.3.6. Estimação de Parâmetros da Função de Acoplamento 3-variada

Nesta seção serão apresentados exemplos semelhantes aos existentes para a função de acoplamento bivariada, mostrando a possibilidade de ajuste de funções de acoplamento de maiores dimensões.

Exemplo 3.4 a) A tabela 3.15 mostra a média e dispersão das estimativas dos parâmetros de 100 replicações de simulações de 3000 observações da função de acoplamento t-Student Assimétrica 3-variada com parâmetros $v = 6.79$, $\rho_{1,2} = 0.588$, $\rho_{1,3} = -.705$, $\rho_{2,3} = 0.074$ e $(\gamma_1, \gamma_2, \gamma_3) = (-0.54, 0.82, -0.05)$.

Tabela 3. 15. Parâmetros estimados por máxima verossimilhança de 100 simulações de 3000 valores da função de acoplamento t-Student Assimétrica 3-variada de parâmetros $v = 6.79$, $\rho_{1,2} = 0.588$, $\rho_{1,3} = -.705$, $\rho_{2,3} = 0.074$ e $(\gamma_1, \gamma_2, \gamma_3) = (-0.54, 0.82, -0.05)$.

	Assimet. 1	Assimet. 2	Assimet. 3	Correl. 1,2	Correl. 1,3	Correl. 2,3	Graus Lib.
Verdadeiro	-0.5400	0.8200	-0.0500	0.5880	-0.7050	0.0780	6.7900
Média	-0.5669	0.7739	-0.0655	0.5814	-0.7089	0.0782	6.9720
Desvio Padrão	0.0551	0.0302	0.0378	0.0150	0.0073	0.0164	0.3421
Erro Quad. Médio	3.72E-03	3.03E-03	1.66E-03	2.67E-04	6.76E-05	2.65E-04	1.49E-01

Note a proximidade das estimativas dos parâmetros de correlação aos parâmetros originais. O próximo exemplo utiliza as mesmas simulações obtidas aqui, porém com a transformação das marginais uniformes para as distribuições marginais Normal-Assimétricas e t-Student Assimétricas, com posterior estimação dos parâmetros pelos métodos IFM e Máxima Verossimilhança Canônica.

Exemplo 3.4 b) Foram realizadas 100 rodadas de simulações de 3000 valores da função de acoplamento t-Student Assimétrica 3-variada com parâmetros $v = 6.79$, $\rho_{1,2} = 0.588$, $\rho_{1,3} = -.705$, $\rho_{2,3} = 0.074$ e $(\gamma_1, \gamma_2, \gamma_3) = (-0.54, 0.82, -0.05)$, sendo a primeira marginal

Uniforme(0,1) transformada na distribuição Normal-Assimétrica com parâmetros $\mu = 0.54$, $\sigma = 0.72$ e $\alpha = 3.35$. A 2º e 3ª marginais na distribuição t-Student Assimétrica com parâmetros $v = 6.78$, $\gamma = -0.78$ e $v = 16.78$, $\gamma = 0.078$, respectivamente. As tabelas 3.16 e 3.17 exibem estimações dos parâmetros das marginais e da função de acoplamento pelo método IFM e estimações dos parâmetros da função de acoplamento pelo método de Máxima Verossimilhança Canônica.

Tabela 3. 16. Parâmetros estimados pelo método IFM para 100 simulações distintas de 3000 valores da função de acoplamento t-Student Assimétrica 3-variada com parâmetros $v = 6.79$, $\rho_{1,2} = 0.588$, $\rho_{1,3} = -.705$, $\rho_{2,3} = 0.074$ e $(\gamma_1, \gamma_2, \gamma_3) = (-0.54, 0.82, -0.05)$, sendo a primeira marginal Uniforme(0,1) transformada na distribuição Normal-Assimétrica com parâmetros $\mu = 0.54$, $\sigma = 0.72$ e $\alpha = 3.35$. A 2º e 3ª marginais na distribuição t-Student Assimétrica com parâmetros $v = 6.78$, $\gamma = -0.78$ e $v = 16.78$, $\gamma = 0.078$, respectivamente.

	Marg. 1 Normal Assimétrica			Marg 2. t-Student Assim.		Marg 2. t-Student Assim.	
	Posição	Escala	Assimetria	Graus Lib.	Assimetria	Graus Lib.	Assimetria
Verdadeiro	0.5400	0.7200	3.3500	6.7800	-0.7800	16.7800	0.0780
Média	0.5400	0.7185	3.4066	6.7916	-0.7865	19.7339	0.0797
Desvio Padrão	0.0119	0.0127	0.2050	0.3224	0.0191	3.8876	0.0158
Erro Quad. Médio	1.40E-04	1.63E-04	4.48E-02	1.03E-01	4.04E-04	2.37E+01	2.49E-04

Acoplamento	Assimet. 1	Assimet. 2	Assimet. 3	Correl. 1,2	Correl. 1,3	Correl. 2,3	Graus Lib.
Verdadeiro	-0.5400	0.8200	-0.0500	0.5880	-0.7050	0.0780	6.7900
Média	-0.5604	0.7821	-0.0759	0.5817	-0.7076	0.0797	7.1049
Desvio Padrão	0.0527	0.0365	0.0386	0.0171	0.0089	0.0228	0.3195
Erro Quad. Médio	3.17E-03	2.75E-03	2.14E-03	3.30E-04	8.59E-05	5.16E-04	2.00E-01

Comparativamente ao obtido no exemplo 3.4 a), os parâmetros foram estimados mais afastados dos parâmetros originais, na maioria deles, com erro quadrático médio maior. Entretanto exibe-se ainda um ajuste adequado, cujos parâmetros originais estão dentro do intervalo de confiança de 95% dos parâmetros estimados. A tabela 3.17 exibe os parâmetros estimados pelo método de máxima verossimilhança canônica. Novamente observa-se uma perda da precisão do estimador e aumento do erro quadrático médio comparativamente aos casos anteriores.

Tabela 3. 17. Parâmetros estimados pelo método de máxima verossimilhança canônica para 100 simulações distintas de 3000 valores da função de acoplamento t-Student Assimétrica 3-variada com parâmetros $v = 6.79$, $\rho_{1,2} = 0.588$, $\rho_{1,3} = -.705$, $\rho_{2,3} = 0.074$ e $(\gamma_1, \gamma_2, \gamma_3) = (-0.54, 0.82, -0.05)$,

	Assimet. 1	Assimet. 2	Assimet. 3	Correl. 1,2	Correl. 1,3	Correl. 2,3	Graus Lib.
Verdadeiro	-0.5400	0.8200	-0.0500	0.5880	-0.7050	0.0780	6.7900
Média	-0.5729	0.7709	-0.0659	0.5820	-0.7078	0.0779	7.0967
Desvio Padrão	0.0572	0.0335	0.0467	0.0174	0.0100	0.0205	0.3498
Erro Quad. Médio	4.33E-03	3.52E-03	2.41E-03	3.35E-04	1.07E-04	4.18E-04	2.15E-01

A tabela 3.18 mostra o desvio padrão calculado pelo hessiano das estimativas dos parâmetros do exemplo 3.4 (b).

Tabela 3. 18. Desvio padrão médio estimado através do hessiano das estimações de parâmetros de 100 simulações de 3000 observações da distribuição t-Student Assimétrica 3-variada de parâmetros $v = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$

	Assimetria 1	Assimetria 2	Assimetria 2	Correlação 1,2	Correlação 1,3	Correlação 2,3	Graus Lib.
Amostral	0.0551	0.0302	0.0378	0.0150	0.0073	0.0164	0.3421
Médio	0.0054	0.0082	0.0051	0.0116	0.0075	0.0153	0.0070
Mediano	0.0047	0.0072	0.0045	0.0116	0.0075	0.0153	0.0064
Desvio Padrão	0.0024	0.0033	0.0020	0.0005	0.0002	0.0004	0.0026

Novamente observam-se desvios padrão muito baixos quando calculados pela matriz hessiana. Por esse motivo foram utilizados os seguintes diferenciais para o cálculo do hessiano. Nota-se também que o desvio padrão amostral do caso 3-variado é menor do que no bivariado.

Tabela 3. 19. Diferenciais utilizados no cálculo numérico da matriz hessiana para cada parâmetro da função de acoplamento t-Student Assimétrica

	Assimetria 1	Assimetria 2	Assimetria 2	Correlação 1,2	Correlação 1,3	Correlação 2,3	Graus Lib.
h	0.01	0.01	0.01	0.001	0.001	0.001	0.1
ε	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001

Os resultados do desvio padrão calculado pelo hessiano para essa cópula são exibidos na tabela 3.20.

Tabela 3. 20. Desvio padrão médio estimado através do hessiano das estimações de parâmetros de 100 simulações de 3000 observações da distribuição t-Student Assimétrica bivariada de parâmetros $\nu = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$

	Assimetria 1	Assimetria 2	Assimetria 2	Correlação 1,2	Correlação 1,3	Correlação 2,3	Graus Lib.
Amostral	0.0551	0.0302	0.0378	0.0150	0.0073	0.0164	0.3421
Médio	0.0381	0.0522	0.0296	0.0147	0.0083	0.0168	0.2805
Mediano	0.0356	0.0490	0.0290	0.0144	0.0082	0.0168	0.2533
Desvio Padrão	0.0096	0.0145	0.0038	0.0017	0.0004	0.0006	0.0810

Foram utilizados esses hessianos para a determinação de intervalos de confiança de 90%, 95%, 97.5% e 99% das estimativas dos parâmetros. Com esses intervalos de confiança, como feito para o caso marginal, foi verificada a proporção do número de séries em que o parâmetro verdadeiro se encontra dentro dos intervalos de confiança. Os resultados são apresentados na tabela 3.21

Tabela 3. 21. Percentual em que o intervalo de confiança de níveis 90%, 95%, 97.5% e 99% abrangeu o verdadeiro parâmetro nas estimativas realizadas em 100 simulações de 3000 observações da função de acoplamento t-Student Assimétrica com $\nu = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$.

	90%	95%	97.50%	99%
Assimetria 1	88.00%	94.00%	98.00%	98.00%
Assimetria 2	92.00%	98.00%	100.00%	100.00%
Assimetria 3	92.00%	94.00%	94.00%	94.00%
Correlação 1,2	78.00%	84.00%	96.00%	100.00%
Correlação 1,3	81.00%	89.00%	93.00%	98.00%
Correlação 1,3	71.00%	78.00%	86.00%	94.00%
Graus Lib.	82.00%	92.00%	94.00%	98.00%

Observa-se que o parâmetro verdadeiro ficou dentro dos intervalos de confiança em menos séries do que no caso bivariado, com exceção dos graus de liberdade, ficando também menos próximo do percentual teórico.

3.3.7. Precisão do Algoritmo

Como mencionado anteriormente, a precisão do algoritmo, caracterizada aqui como o número de casas decimais do calculo da inversa da função acumulada da t-Student Assimétrica pode melhorar ou reduzir a qualidade do resultado final das estimativas dos parâmetros da função de acoplamento t-Student Assimétrica. Nesta seção serão apresentados ajustes de funções de acoplamento para diversas precisões, verificando possíveis melhorias com o aumento da precisão e também o inverso, procurando a melhor opção em termos de custo/benefício computacional.

Os resultados serão apresentados para a função de acoplamento bivariada, tendo extensão natural para ordens maiores do acoplamento. A tabela 3.22 mostra o tempo de estimação da função de acoplamento para precisão entre 0.0001 a 0.001, em um computador de núcleo duplo com 3Ghz reais de velocidade de processamento, barramento de 1333Mhz e 8GB de Memória RAM, juntamente com estimativas dos parâmetros para cada precisão selecionada.

Tabela 3. 22. Velocidade até a obtenção de estimativas de parâmetros para 1 simulação da função de acoplamento t-Student Assimétrica bivariada com $\nu = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$, com precisão de valores 0.01, 0.005, 0.004, 0.003, 0.002, 0.001, 0.0007, 0.0005, 0.0002 e 0.0001.

	Tempo	Estimativas dos Parâmetros			
Precisão	segundos	Assimetria 1	Assimetria 2	Correlação	Graus Lib.
0.01	22.16	-0.533	-0.792	0.654	7.254
0.005	42.65	-0.537	-0.784	0.652	7.237
0.004	48.81	-0.540	-0.796	0.651	7.282
0.003	63.55	-0.522	-0.773	0.653	7.292
0.002	95.54	-0.545	-0.807	0.652	7.288
0.001	190.44	-0.539	-0.791	0.652	7.303
0.0007	277.63	-0.537	-0.795	0.652	7.285
0.0005	372.30	-0.541	-0.792	0.652	7.317
0.0002	1084.25	-0.542	-0.795	0.652	7.353
0.0001	> 30 Mins	-	-	-	-
Parâmetros Verdadeiros		-0.48	-0.72	0.65	7.21

A tabela mostra que o tempo até a finalização da estimação cresce de maneira quase exponencial conforme se aumenta a precisão do algoritmo. Observando os resultados das estimativas pode-se supor que o uso de uma precisão baixa não gera redução da qualidade das estimativas dos parâmetros. Contudo, nesse caso os desvios padrão calculados através do hessiano com os diferenciais *h*'s selecionados (seção 3.3.4), ficam ainda mais subestimados (Tabela 3.23).

Tabela 3. 23. Desvios padrão calculados pelo hessiano para precisão 0.01 e 0.001 de 1 simulação da função de acoplamento t-Student Assimétrica bivariada com $\nu = 7.21$, $\rho = 0.65$ e $(\gamma_1, \gamma_2) = (-0.48, -0.72)$

Precisão	Assimetria 1	Assimetria 2	Correlação	Graus Lib.
0.01	0.0320	0.0504	0.0107	0.4876
0.001	0.0939	0.1451	0.0169	0.6555

Mesmo assim o aumento da precisão não gera resultados de melhoria das estimativas, apenas aumento do custo computacional e a possibilidade de melhor cálculo de desvios padrão.

Para garantir o melhor ajuste possível para os dados reais será utilizada a precisão 0.002 para as estimativas de acoplamentos bivariados (para encontrar os valores iniciais dos parâmetros para uso no algoritmo de estimação por máxima verossimilhança) e 0.001 para as estimativas da função de acoplamento completa. A execução da estimação de parâmetros de uma função de acoplamento 3-variada completa dura em torno de 10 minutos para essa precisão.

As estimativas de parâmetros do acoplamento para cálculo do VaR (Valor em Risco, apêndice D) para as séries de dados reais será realizada com precisão de 0.01 para redução do custo computacional (já que nesse caso são executadas mais de 50 estimações de parâmetros da função de acoplamento 3-variada).

Capítulo 4

Aplicação – Petróleo e Derivados

O Petróleo é um elemento energético bruto de produção global em larga escala. Possuindo uma diversa gama de derivados, cujas produções de maior volume são os óleos combustíveis e gasolina, o também chamado Óleo Cru possui um mercado amplo, pois seus derivados são utilizados como combustível em motores, em aquecedores, na produção de plásticos e lubrificantes, entre outros.

As negociações desse ativo físico são realizadas através das Bolsas Mercantis e de Futuros, que categorizam o Óleo Cru através do local de sua produção (p. ex. Petróleo do tipo Brent, 'West Texas', Dubai-Oman e OPEC (acrônimo de língua inglesa de Organização dos Países Exportadores de Petróleo), dentre outros), e possuem diferentes cotações, preços que são relacionados com os custos de extração, transporte e processamento do produto nas refinarias.

De um barril de petróleo (que possui aproximadamente 159 litros de óleo cru), extrai-se, em média, 45% de seu conteúdo em gasolina e outros 20% em óleos combustíveis (essencialmente os óleos para aquecimento residencial e o óleo Diesel) (*'Texas Oil and Gas Association'*, 1995). Esses percentuais, que são uma média obtida das refinarias de petróleo dos EUA, variam anualmente de acordo com o mix (proporções) de saída desejado pela indústria, que tem por base as necessidades dos consumidores/mercado.

Os preços de gasolina ou óleos combustíveis são, por sua vez, vinculados em algum grau ao preço do óleo cru, além de outros fatores, como variações do consumo e *mix*, e podem possuir alterações distintas de acordo com a demanda do mercado ou especulação por parte dos agentes do mercado ou mesmo dos agentes que vendem os produtos refinados aos consumidores finais, indicando muitas vezes assimetria entre a dependência de preços de petróleo vs. produtos refinados (Borenstein *et al*, 1992).

O fato de já ter havido evidências para supor que exista assimetria na dependência entre os preços do Óleo Cru e de seus derivados dá valia a uma análise de funções de acoplamento que possa de alguma forma mensurar estocasticamente essa dependência assimétrica e propiciar melhores condições de gerenciamento de risco, por exemplo, no caso da necessidade de se utilizar um modelo que vincule preços de petróleo com os preços de seus refinados. Neste trabalho serão estudadas séries históricas diárias de preços de elementos energéticos negociados na Bolsa Mercantil de Nova York (NYMEX – *"New York Mercantile Exchange"*):

1. Histórico de preços de curto prazo do Petróleo do tipo WTI (*"West Texas Intermediate"*);
2. Histórico de preços de curto prazo da Gasolina Convencional com entrega no porto de Nova York (*"New York Harbor Conventional Gasoline Regular"*);
3. Histórico de valores do Índice de Óleo AMEX (*"American Exchange Oil Index"*).

Espera-se com este trabalho detectar a existência de assimetria na dependência entre esses ativos energéticos através de funções de acoplamento.

Log-retornos (aproximadamente a variação percentual do preço com base em uma data inicial) de ativos, em particular dos ativos energéticos possuem muitos dos chamados fatos estilizados de séries financeiras, como persistência da série, excesso de curtose, conglomerados de volatilidade e dependência entre retornos.

O objetivo inicial deste capítulo é obter uma modelagem das marginais para as séries de preços de curto prazo mencionadas, de forma que se tenha como resultados resíduos que possam ser comparados a fim de se obter a dependência real entre os ativos, desvinculada de dependências seriais. Para obter tais séries, as séries originais serão filtradas por um modelo ARMA-GARCH, com posterior ajuste de função de acoplamento aos resíduos desse ajuste.

4.1. Características das Séries

4.1.1. Óleo Cru WTI

O Óleo Cru WTI é explorado e extraído no oeste do estado do Texas nos EUA, e negociado dentro da NYMEX. Sua cotação é usada como *"benchmark"* (padrão) dos preços de petróleo nos Estados Unidos, ou seja, mesmo o petróleo produzido em outras partes daquele país tende a ser cotado por esse preço.

A série de preços do Petróleo WTI, obtida através do Departamento de Energia do governo dos Estados Unidos da América (DOE-USA), inicia-se na data de 02 de Janeiro de 1986, com informações diárias colhidas até a data de 01 de Julho de 2008, excluindo-se finais de semana, feriados e outras datas em que negociações não foram realizadas em pelo menos uma das 3 séries em estudo.

A figura 4.1 mostra a série de preços curto prazo (*"spot"*) do Óleo Cru WTI, entre 02/01/1986 a 01/07/2008. O comportamento da série mostra que esse produto sofre influência de eventos geopolíticos, como a ocorrência de conflitos, aumentos ou reduções da produção como estratégia de alavancagem dos lucros e a depreciação de reservatórios, principalmente dos campos localizados no oriente médio.

Após os choques do petróleo que ocorreram na década de 70, uma ampliação da extração de petróleo, principalmente nos campos sauditas, mantiveram os preços desse ativo estáveis até 1990. Entre 1990 e 1991 com o advento da primeira guerra dos Estados Unidos contra o Iraque (Guerra do Golfo), os campos de produção do Kuwait foram desabilitados (Parra, 2004, Cap. 15). Fatores Geopolíticos fizeram com que o preço praticamente dobrasse de valor, retornando à estabilidade nos anos posteriores à guerra, até 1998.

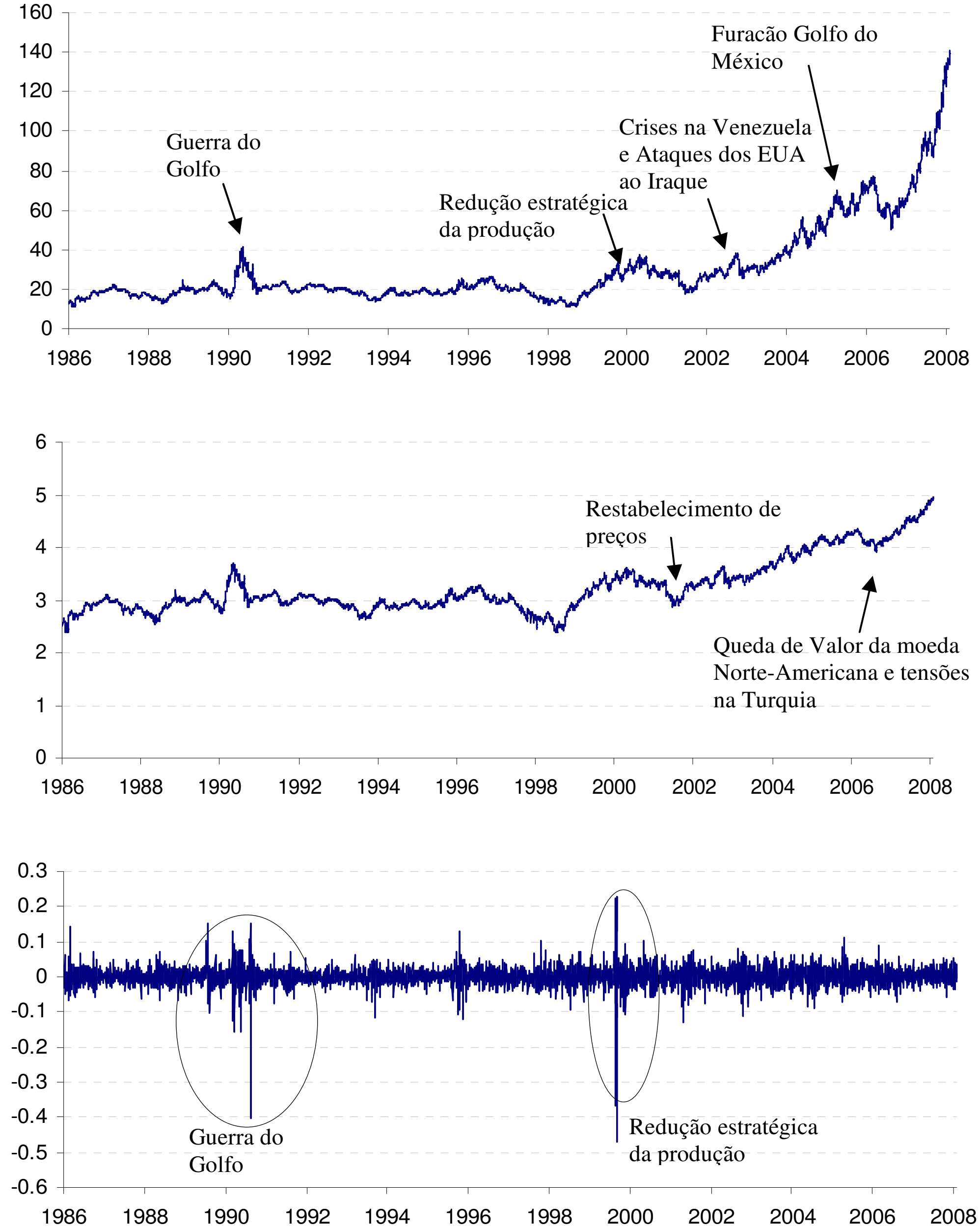

Figura 4. 1. Preços de curto prazo do Petróleo WTI, cotação NYMEX a) em Dólares por Barril, b) log(preços) em Dólares por Barril, c) log-retornos de preços

A partir de 1998, o aumento da demanda principalmente pelos mercados asiáticos induziu os países da OPEP a elevarem a produção, o que reduziu o preço do petróleo mundialmente. Entretanto, entre os anos de 1999 e 2000 a OPEP reduziu, estrategicamente, a produção de óleo, mesmo com alta demanda, o que aumentou o preço do minério, seguido de um restabelecimento de seu preço em patamares menos elevados até meados de 2001. Após 2001, diversos choques geopolíticos fizeram com que o preço crescesse e chegasse ao patamar que hoje se encontra. Pode-se citar: 2002-03: crise na Venezuela e ataques dos EUA e aliados ao Iraque; 2004: nova redução da produção da OPEP; 2005: Furacão afeta campos do Golfo do México; 2006: possibilidade de novos conflitos entre Irã e Iraque aliados com a alta demanda por óleo pelos EUA; 2007: tensões na Turquia e queda do valor da moeda Norte-americana (Williams, 2007).

No contexto atual, fatores geopolíticos continuam a afetar os preços. Um dos fatos relevantes é a alta demanda pelo líquido nos países asiáticos, principalmente na China, que já vinha reportando deficiências energéticas, inicialmente de energia elétrica, a mais de 4 anos. Esse fato, aliado com o aumento da demanda pelo minério em todo o mundo, num cenário de estagnação da produção, dada a ausência parcial de descobertas de novos poços, mantém os preços elevados (Whipple, 2008). Como a série finda em julho de 2008 não são observados os impactos do rompimento da bolha imobiliária em seus preços. Devido ao fato de que a série possui um comportamento de aumento de preços após 2002, selecionou-se o início desse ano como um período de corte. Os modelos ARMA-GARCH-Cópula serão estimados para as séries, então, nesses 2 períodos: Período 1: **Pré Choques 2002,** e Período 2: **Pós Choques 2002.**

Filtragem do Nível e Volatilidade

O modelo ARMA-GARCH-Cópulas é um modelo que caracteriza três aspectos de um conjunto de séries temporais: o nível de cada série (Modelo ARMA), sua volatilidade (modelo GARCH) e a dependência entre as séries (Função de Acoplamento). Serão estimados os modelos ARMA-GARCH para os log-retornos de preços de Petróleo WTI nos períodos Pré e Pós Choques de 2002.

A figura 4.1 (c) mostra os log-retornos dos preços de petróleo WTI, em todo o período da série. Note a existência de períodos de maior volatilidade, os chamados conglomerados de volatilidade, que ocorrem nos momentos de choques dos preços do petróleo (ocorrência de influências geopolíticas no preço).

A Autocorrelação dos log-retornos do Óleo Cru (figura 4.2) indica leve correlação entre as defasagens de ordens 2 e 3 da série no período 1 e defasagem 1 no período 2. A análise dos resíduos do ajuste dos modelos ARMA para o período 1 (pré choques de 2002, veja figura 4.3) indica que a maior parte da correlação pode ser removida pelo modelo ARMA (2,0). Foi ajustado um modelo ARMA (3,0) à série mas o mesmo também não conseguiu filtrar suficientemente a autocorrelação na defasagem 3. A inclusão de um termo MA de ordem 3 resolveria esse problema mas dado não existir justificativas econômicas para um choque com defasagem 3, o valor da autocorrelação estimado não ser muito alto e por questões de parcimônia foi adotado o ARMA (2,0). Para o período 2 foi ajustado o modelo ARMA (1,0).

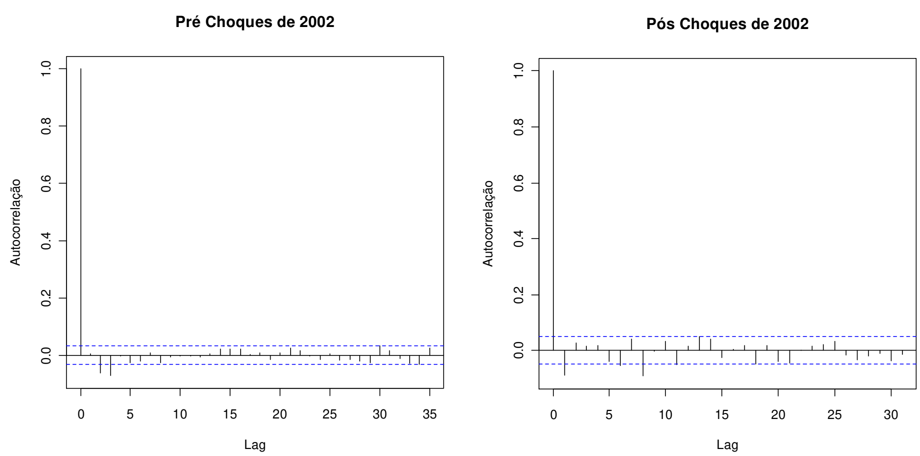

Figura 4. 2. Autocorrelação da série de log-retornos do Óleo Cru WTI (NYMEX) períodos pré e pós choques de 2002.

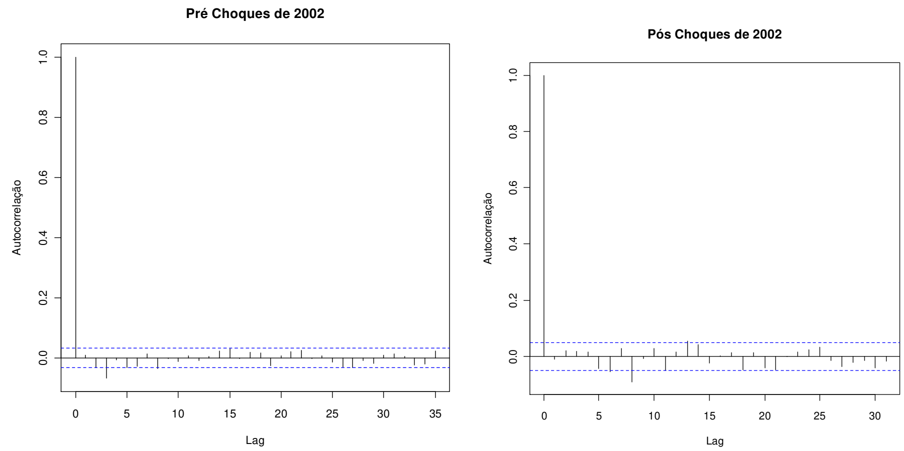

Figura 4. 3. Autocorrelação dos resíduos do ajuste dos modelos ARMA para o Óleo Cru WTI (NYMEX) períodos pré e pós choques de 2002.

Não existem razões para supor que a correlação serial nos resíduos (figura 4.3) observada na defasagem de ordem 8 do período pós choques de 2002 seja mais do que uma correlação espúria (metade da 2ª semana). Daí foi ajustado apenas uma constante nesse período.

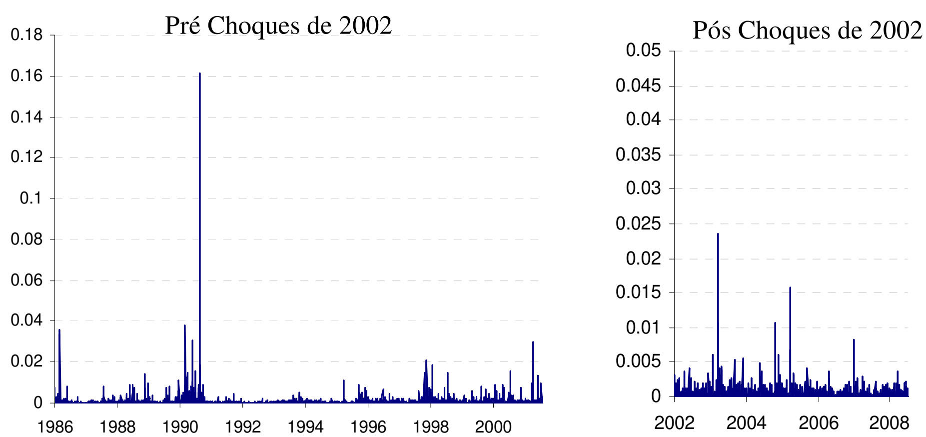

Figura 4. 4. Resíduos quadráticos do modelo ARMA (2,0) (Período 1) e ARMA (1,0) (Período 2), utilizado nos log-retornos do Óleo Cru WTI

A figura 4.4 mostra os retornos quadráticos dos resíduos de um modelo ARMA (2,0) (Período 1) e ARMA (1,0) (Período 2). Nesse gráfico, os conglomerados de volatilidade ficam mais evidentes. As variações da volatilidade da série serão também capturadas pelo modelo, no componente GARCH (Bollerslev, 1986). Esse modelo propõe que a variabilidade de uma série possui correlação temporal. Assim, justifica-se a utilização do mesmo quando existe autocorrelação no quadrado da série de retornos, depois de filtrado o nível da série, pois são indicativos da variabilidade em cada período.

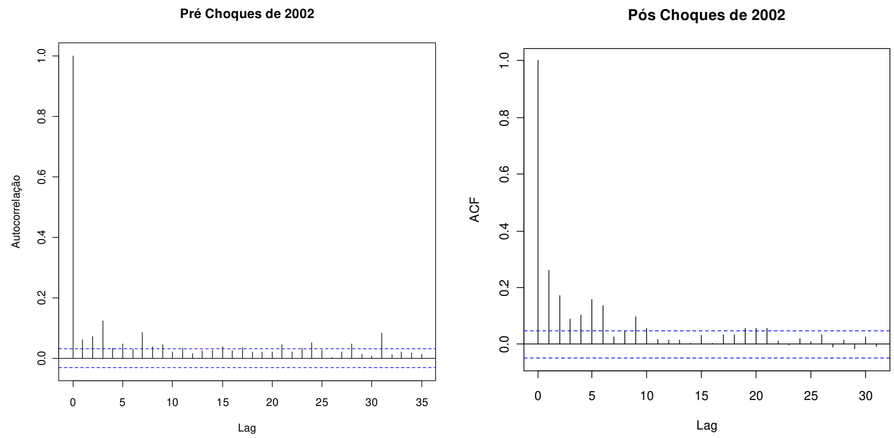

Figura 4. 5. Autocorrelação dos resíduos quadráticos do modelo ARMA (2,0) (Período 1) e ARMA(1,0) (Período 2), utilizado nos log- retornos do Óleo Cru WTI

Vê-se através da figura 4.5 que existe correlação temporal evidente principalmente nas defasagens de ordens 1 a 3 da série de resíduos ao quadrado no período 1 e até a sexta defasagem no período 2. Isso indica a necessidade de um modelo para a volatilidade da série. Nesta aplicação será utilizado o modelo GARCH com inovações t-Student.

Verificou-se que o modelo de maior parcimônia foi o GARCH (1,1), já que modelos inferiores como (1,0) e (0,1) foram insuficientes para estabilizar a variância da série. Modelos com maior número de parâmetros resultaram em coeficientes não significativos (nível $\alpha = 5\%$). Dessa forma, o modelo que filtra de forma aceitável a série de preços do Óleo Cru WTI é o modelo ARMA(2,0)-GARCH(1,1) para o período 1 e ARMA(1,0)-GARCH(1,1) para o período 2. A figura 4.7 apresenta a série temporal filtrada.

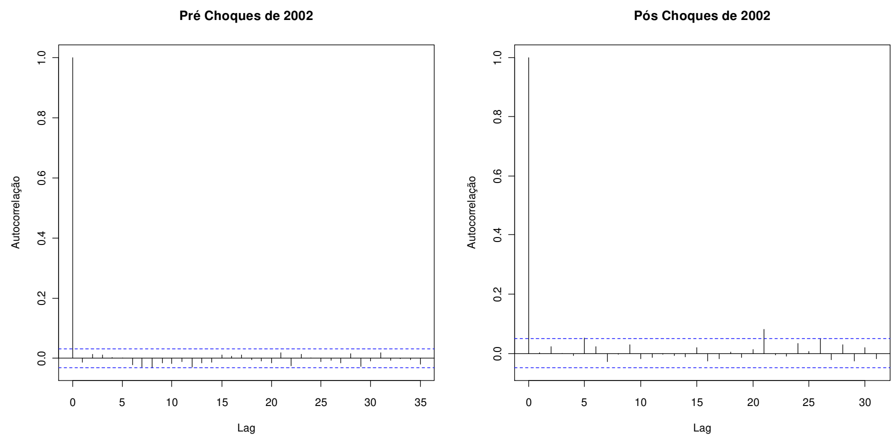

Figura 4. 6. Autocorrelação dos resíduos quadráticos do modelo ARMA (2,0)-GARCH(1,1) (Período 1) e ARMA(1,0)-GARCH(1,1) (Período 2), utilizado nos log- retornos do Óleo Cru WTI

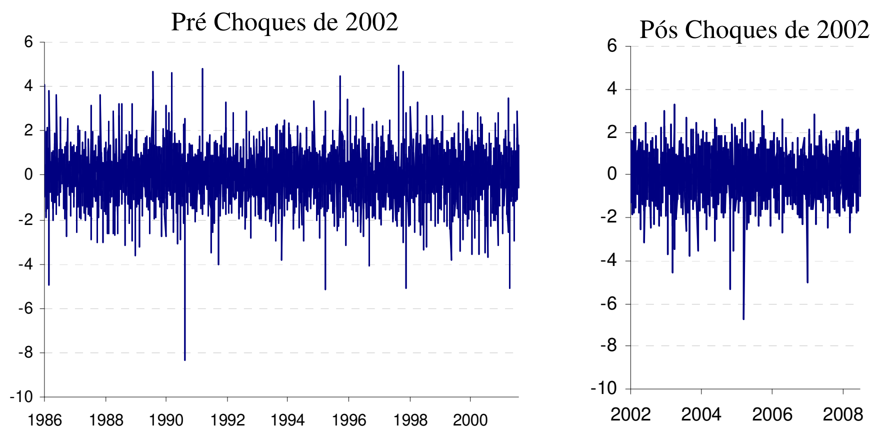

Figura 4. 7. Log-Retornos do Óleo Cru WTI filtrados por modelo ARMA(2,0)-GARCH(1,1), (Período 1) e ARMA(1,0)-GARCH(1,1) (Período 2).

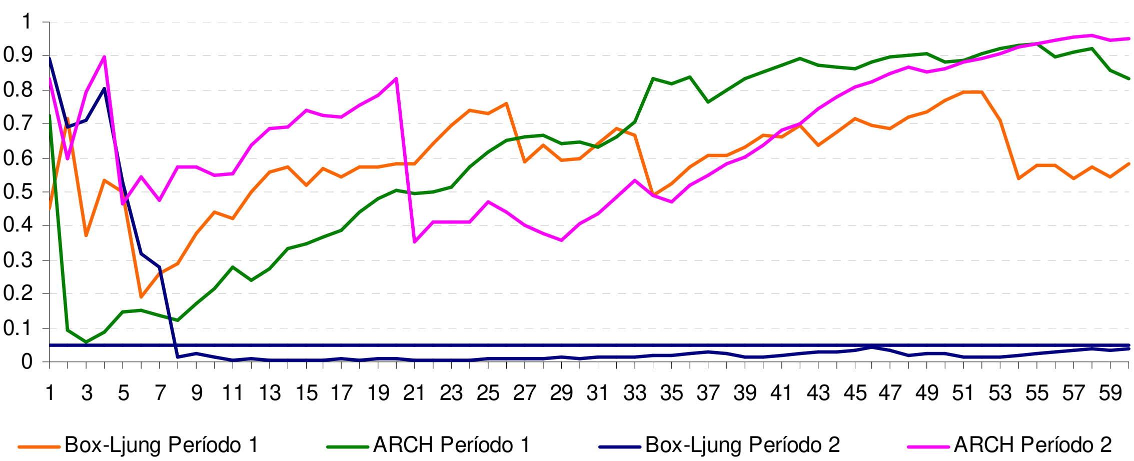

Figura 4. 8. P-valores dos Testes de Multiplicadores de Lagrange (ARCH LM) e de Autocorrelação (Box-Ljung) para o nível das séries filtradas de log-retornos de preços do Petróleo WTI.

A figura 4.8 mostra os testes de Box-Ljung para autocorrelação dos resíduos dos modelos ARMA e os testes de Multiplicadores de Lagrange para os resíduos dos modelos ARMA-GARCH dos log-retornos de preços de Petróleo WTI nos períodos 1 e 2. Note que após a 8ª defasagem, o teste de Box-Ljung passa mostrar evidência da falta de ajuste do modelo ARMA para o período 2. Entretanto foram utilizados outros modelos ARMA de ordens superiores e não surtiram efeitos expressivos em termos de bondade de ajuste pelo teste de Box-Ljung, além da baixa significância estatística dos parâmetros de modelos de ordens superiores (considerando nível $\alpha = 5\%$). Observe também que os testes ARCH LM não dão evidências da necessidade de novos ajustes de modelos para a volatilidade.

A tabela 4.1 mostra os parâmetros ajustados dos modelos ARMA-GARCH com perturbações modeladas pela distribuição t-Student padronizada. Nas séries seguintes também serão utilizadas a distribuição t-Student padronizada no modelo GARCH(1,1), para os períodos Pré e Pós Choques de 2002 da série de log retornos do Óleo Cru WTI.

Tabela 4.1. Parâmetros ajustados dos modelos ARMA(2,0)-GARCH(1,1), (Período 1) e ARMA(1,0)-GARCH(1,1) (Período 2) para o log-retornos do Óleo Cru WTI.

Óleo Cru WTI Período Pré Choques 2002

	Parâmetro	Erro Padrão	t-Valor	P(T>t)
Média	0.0003	0.0003	1.1377	0.2552
AR 1	-0.0068	0.0153	-0.4432	0.6576
AR 2	-0.0324	0.0157	-2.0586	0.0395
Média Vol	0.0000	0.0000	4.1022	0.0000
GARCH(P,.)	0.0716	0.0107	6.6674	0.0000
GARCH(.,Q)	0.9194	0.0108	85.3330	0.0000
g.l. t-Student	4.7221	0.3613	13.0711	0.0000
Persistência:	0.9910			

Óleo Cru WTI Período Pós Choques 2002

	Parâmetro	Erro Padrão	t-Valor	P(T>t)
Média	0.0020	0.0005	3.8299	0.0001
AR 1	-0.0753	0.0254	-2.9670	0.0030
Média Vol	0.0000	0.0000	2.4510	0.0142
GARCH(P,.)	0.0471	0.0137	3.4255	0.0006
GARCH(.,Q)	0.8886	0.0348	25.4986	0.0000
g.l. t-Student	8.4265	1.5135	5.5677	0.0000
Persistência	0.9328			

4.1.2. Gasolina Convencional

A série de preços da Gasolina Convencional possui observações colhidas pelo Departamento de Energia dos Estados Unidos da América, com início no dia 03 de junho de 1986 e fim em 01 de julho de 2008.

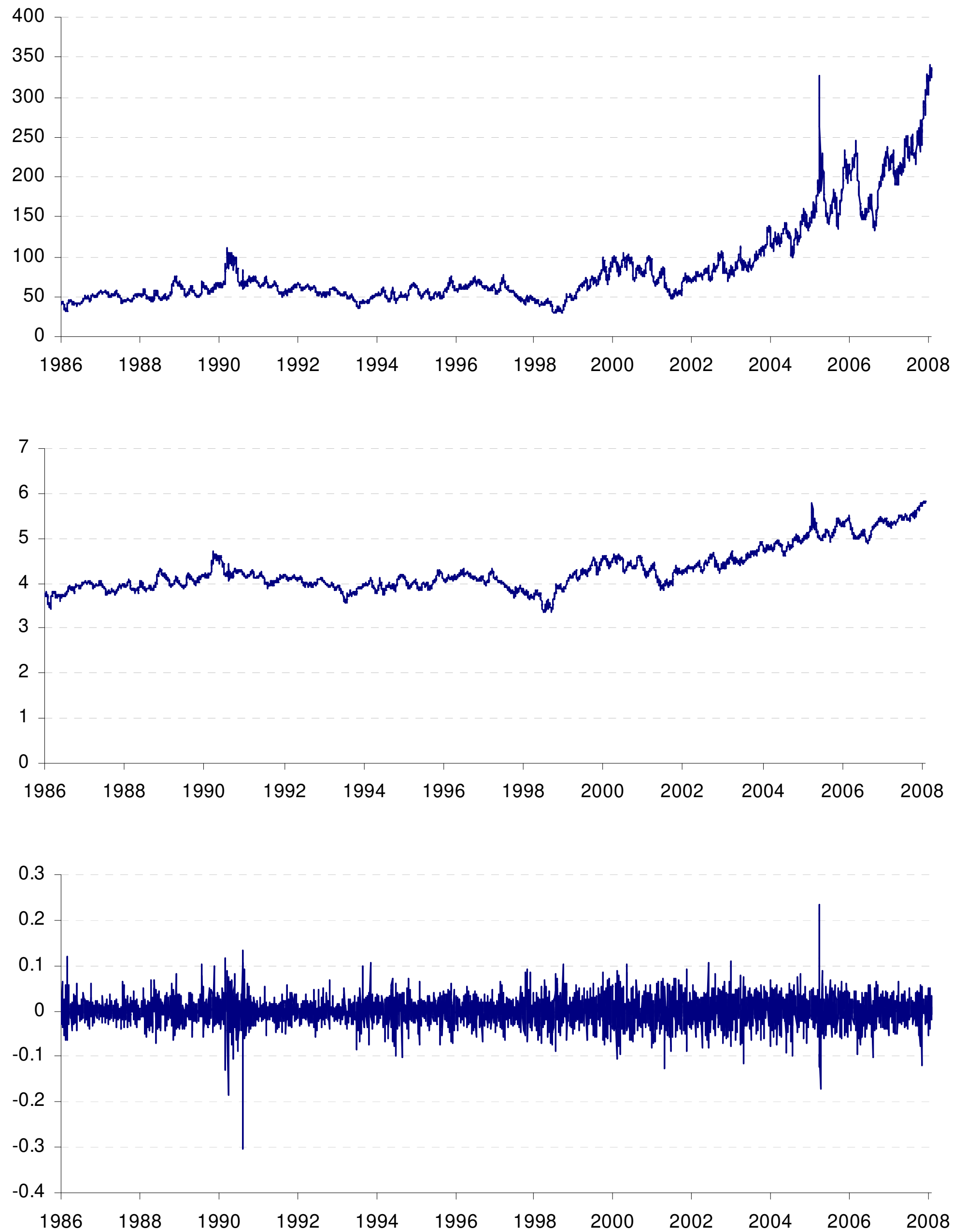

Figura 4. 9. Gasolina Convencional com entrega no porto de Nova York a) preços em centavos de dólar, b) log(preços) em centavos de dólar, c) log-retornos de preços.

A figura 4.9 ilustra o comportamento da série de preços da Gasolina Convencional, medido em US¢/Galão (Centavos de Dólar por Galão; 1 Barril = 42 Galões), onde pode-se observar que a série possui comportamento semelhante à série de preços do Óleo Cru WTI, ou seja, sofre também influências dos fatores geopolíticos assim como o petróleo.

Entretanto, o preço em todo o período é mais volátil do que o petróleo, por sofrer ainda variações de acordo com a demanda dos consumidores finais nas datas de maior consumo, como feriados, com conseqüente variação do *mix* de produção das indústrias (Brown & Virmani, 2007).

Da mesma forma como foi filtrada a série de preços do Óleo Cru WTI, foi ajustado um modelo ARMA-GARCH na série de preços da Gasolina Convencional. Os resultados da seleção e estimação do modelo podem ser observados na tabela 4.2, sendo que o modelo de maior parcimônia foi o ARMA(1,0)-GARCH(1,1), para o Período Pré Choques 2002 e ARMA(0,0)-GARCH(1,1), para o Período Pós Choques de 2002. Gráficos comparativos da autocorrelação pré e pós ajustes dos modelos são apresentados no apêndice C.

Tabela 4. 2. Parâmetros ajustados dos modelos ARMA(1,0)-GARCH(1,1), (Período 1) e ARMA(0,0)-GARCH(1,1) (Período 2) para o log-retornos da Gasolina Convencional.

Gasolina Convencional Período Pré Choques 2002

	Parâmetro	Erro Padrão	t-Valor	P(T>t)
Média	0.0000	0.0005	0.0613	0.9512
AR 1	0.0660	0.0163	4.0444	0.0001
Média Vol	0.0000	0.0000	3.6171	0.0003
GARCH(P,.)	0.0822	0.0122	6.7595	0.0000
GARCH(.,Q)	0.9015	0.0144	62.6887	0.0000
g.l. t-Student	6.8700	0.7296	9.4166	0.0000
Persistência	0.9837			

Gasolina Convencional Período Pós Choques 2002

	Parâmetro	Erro Padrão	t-Valor	P(T>t)
Média	0.0019	0.0007	2.7392	0.0062
Média Vol	0.0001	0.0000	2.8826	0.0039
GARCH(P,.)	0.0721	0.0190	3.7898	0.0002
GARCH(.,Q)	0.8029	0.0537	14.9647	0.0000
g.l. t-Student	9.0276	1.8368	4.9149	0.0000
Persistência	0.8751			

Na figura 4.10 são exibidos os p-valores dos testes de Box-Ljung e Multiplicadores de Lagrange dos modelos ARMA-GARCH selecionados para os log-retornos de preços da gasolina convencional. Embora sejam observadas medidas significativas de falta de bondade de ajuste, dando um indicativo de que os modelos foram sub-ajustados aos log-retornos de preços, outras configurações de modelos ARMA-GARCH foram utilizadas, que resultaram em ajustes infimamente superiores. Optou-se por utilizar os modelos indicados.

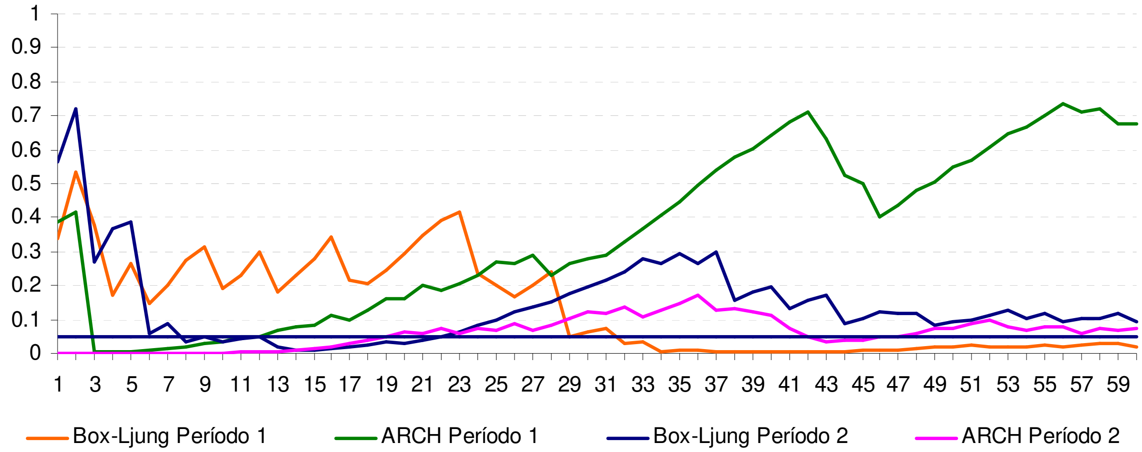

Figura 4. 10. P-valores dos Testes de Multiplicadores de Lagrange (ARCH LM) e de Autocorrelação (Box-Ljung) para o nível das séries filtradas de log-retornos de preços da Gasolina Convencional.

4.1.3. Índice de Óleo AMEX

O índice de Óleo AMEX é um índice da bolsa de valores AMEX (American Exchange), recém adquirida pela NYMEX, que agrega informações de uma variada gama de empresas de exploração de Petróleo e derivados, nas suas diversas fases de produção. É constituído de preços ponderados, medindo a performance da indústria do Petróleo através de mudanças de preços dos ativos que o compõe.

A série utilizada se inicia no dia 27 de agosto de 1984 e finda em 01 de julho de 2008 e foi obtida através do site da WEB Yahoo Finance. A figura 4.11 mostra os valores diários do índice, excluindo observações de finais de semana e feriados e outras de forma que fique exatamente pareado com as outras séries em análise.

Observam-se variações conforme os efeitos e impactos geopolíticos, com comportamento semelhante ao comportamento da série de petróleo WTI, em termos do crescimento no período final (entre 2002 e 2008). Nota-se também que o impacto de alguns efeitos geopolíticos (como por exemplo a Guerra do Golfo em 1991) não é tão grave quanto o mesmo na série do Óleo Cru. Notam-se conglomerados de volatilidade nos log-retornos da série de preços.

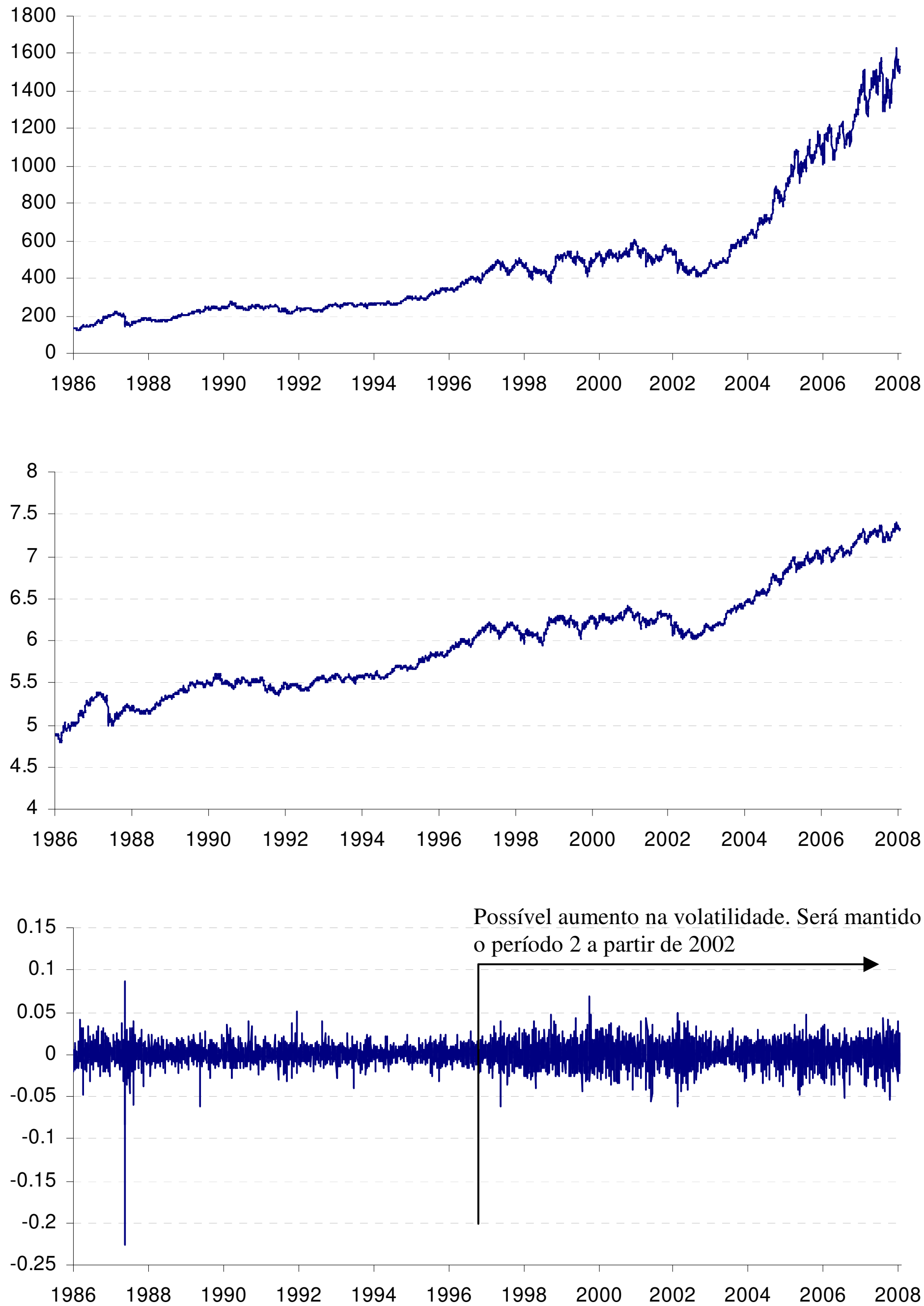

Figura 4. 11. Índice de Óleo AMEX a) valores diários, b) log(valores), c) log-retornos.

O possível aumento na volatilidade após meados de 1996 não será considerado como ponto inicial de um novo período, mantendo-se o início do novo período em 2002. Para filtragem da série utilizou-se também o modelo ARMA(2,0)-GARCH(1,1) no período

1 e ARMA(0,0)-GARCH(1,1)-t no período 2. A tabela 4.3 mostra os resultados do ajuste dos modelos, e a figura 4.12 o p-valor dos testes de bondade de ajuste dos modelos ARMA-GARCH para os períodos 1 e 2. Novamente os gráficos de autocorrelação pré e pós ajustes dos modelos são apresentados no apêndice C.

Tabela 4. 3. Parâmetros ajustados dos modelos ARMA(2,0)-GARCH(1,1), (Período 1) e ARMA(0,0)-GARCH(1,1) (Período 2) para o log-retornos do Índice de Óleo AMEX.

Índice Óleo AMEX Período Pré Choques 2002

	Parâmetro	Erro Padrão	t-Valor	P(T>t)
Média	0.0003	0.0001	2.0815	0.0374
AR 1	0.0534	0.0160	3.3395	0.0008
AR 2	-0.0365	0.0159	-2.2960	0.0217
Média Vol	0.0000	0.0000	2.7058	0.0068
GARCH(P,.)	0.0451	0.0073	6.1993	0.0000
GARCH(.,Q)	0.9495	0.0080	118.3367	0.0000
g.l. t-Student	6.1521	0.5672	10.8472	0.0000
Persistência	0.9946248			

Índice Óleo AMEX Período Pós Choques 2002

	Parâmetro	Erro Padrão	t-Valor	P(T>t)
Média	0.0011	0.0003	3.5839	0.0003
Média Vol	0.0000	0.0000	1.9564	0.0504
GARCH(P,.)	0.0692	0.0151	4.5962	0.0000
GARCH(.,Q)	0.9140	0.0209	43.7259	0.0000
g.l. t-Student	15.4453	5.4015	2.8594	0.0042
Persistência	0.9832068			

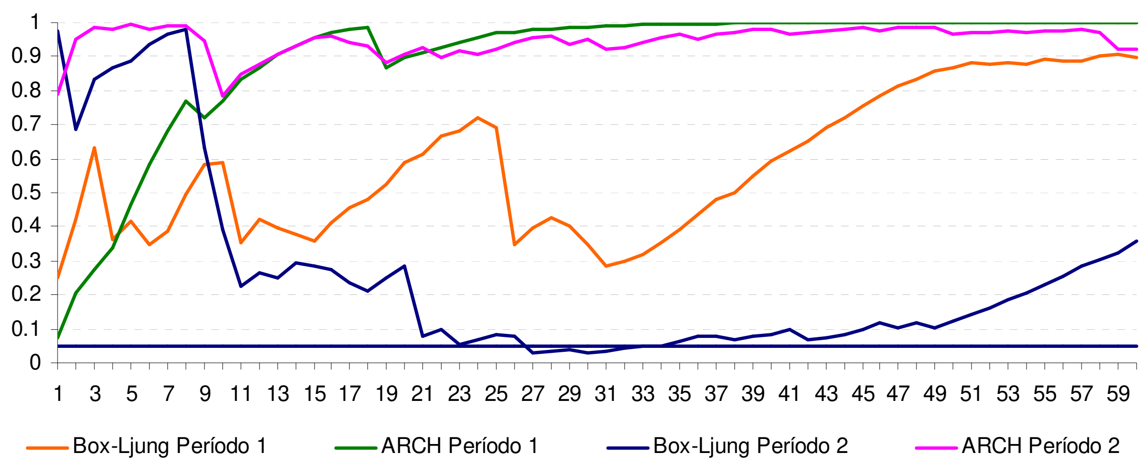

Figura 4. 12. P-valores dos Testes de Multiplicadores de Lagrange (ARCH LM) e de Autocorrelação (Box-Ljung) para o nível das séries filtradas de log-retornos de preços da Gasolina Convencional.

Observam-se no 2º período p-valores significativamente baixos no teste de Box-Ljung para o modelo ARMA, indicando que poderiam ser ajustados modelos de ordem superior. Entretanto o ajuste desses modelos não resultou em modificações no teste de bondade de ajuste Box-Ljung, continuando-se com o ARMA(0,0) (modelo trivial apenas média estática).

Em todas as séries foram também ajustados modelos EGARCH, mas em nenhum caso esse modelo mostrou performance superior ao GARCH (1,1).

4.2. Características Empíricas

Observar as características empíricas, tanto marginais quanto de dependência, pode auxiliar na seleção/desenvolvimento de modelos paramétricos para os conjuntos de dados.

4.2.1. Características Empíricas Marginais

A partir das séries filtradas dos log-retornos das três séries em estudo podem-se observar as distribuições empíricas de cada série, a fim de se conhecer suas características distribucionais, como sua assimetria e o peso das caudas. Essas informações são importantes para desenvolver um modelo estocástico marginal que posteriormente pode ser utilizado na estimação das funções de acoplamento. A figura 4.13 mostra o gráfico de probabilidade normal da série filtrada de log-retornos do Petróleo WTI.

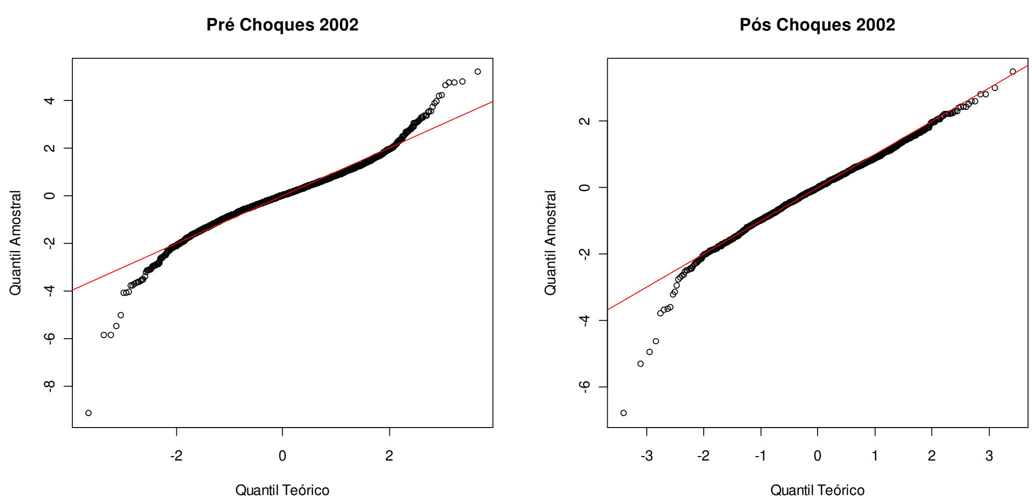

Figura 4. 13. Graficos de comparação de quantís da série filtrada dos log-retornos dos preços do Petróleo WTI nos períodos pré e pós choques 2002, e a distribuição Normal (linha vermelha)

Nota-se que a distribuição da série filtrada no período 1 é aproximadamente simétrica, possuindo caudas mais pesadas do que a distribuição normal. No período 2 observa-se caudas mais pesadas do que a distribuição Normal apenas em valores negativos, uma evidência de assimetria.

Da forma como se comporta, pode-se sugerir o uso de uma distribuição t-Student assimétrica para ajustar a distribuição marginal, que será ajustada na seção 4.3. Seguem, nas figuras 4.14 e 4.15, ilustração do gráfico quantil-quantil das séries filtradas dos log-retornos dos preços da Gasolina Convencional e Índice de Óleo AMEX.

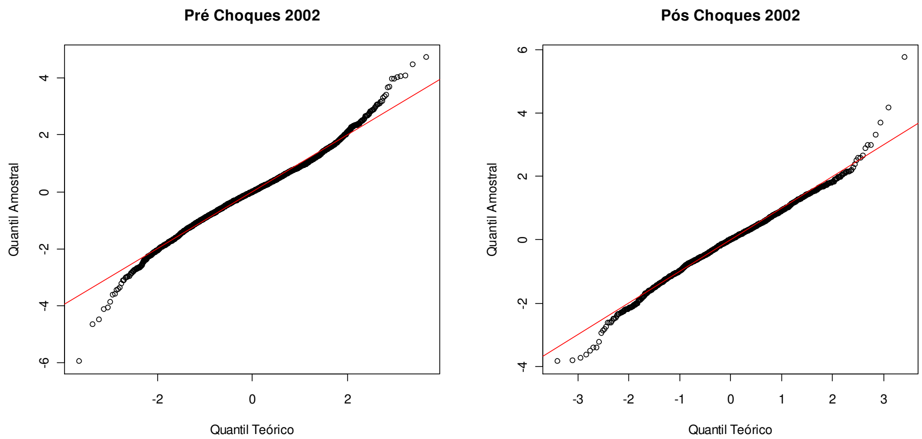

Figura 4. 14. Graficos de comparação de quantís da série filtrada dos log-retornos dos preços da Gasolina Convencional nos períodos pré e pós choques 2002, e a distribuição Normal (linha vermelha)

Observe nas figuras que a distribuição das séries da Gasolina e Índice Óleo AMEX são menos assimétricas do que o Óleo Cru WTI, mas possuem também caudas mais pesadas do que a distribuição Normal, sendo as caudas do 2º período do índice óleo AMEX próximas do peso das caudas da distribuição Normal.

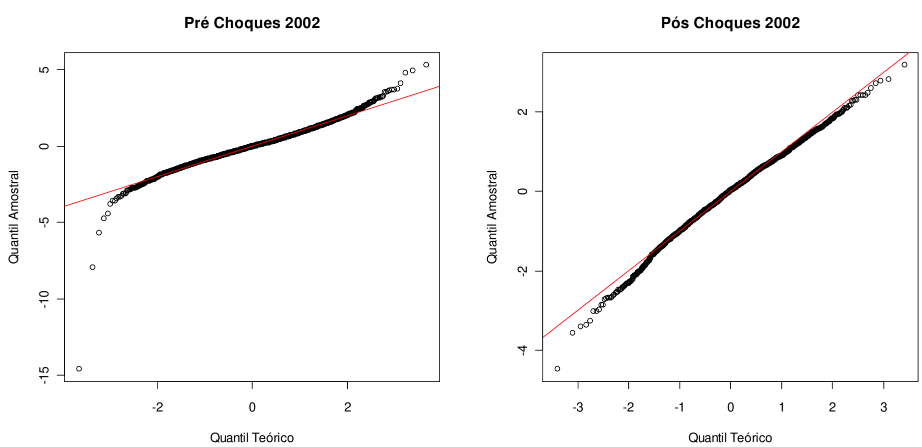

Figura 4. 15. Graficos de comparação de quantís da série filtrada dos log-retornos dos preços do Índice de Óleo AMEX nos períodos pré e pós choques 2002, e a distribuição Normal (linha vermelha)

A Tabela 4.4 exibe as estatísticas de Assimetria (3º Momento Central padronizado) e Excesso de Curtose (4º Momento Central relativo à distribuição Normal) das séries em estudo. Observa-se que a série filtrada de log-retornos do Índice de Óleo AMEX no período pré choques de 2002 é a série que possui maior assimetria e caudas pesadas, seguida pela série de preços do Petróleo WTI. Na Gasolina a assimetria é menos marcante, sendo os impactos geopolíticos amenizados pelo controle do *mix* de produção das refinarias.

Tabela 4. 4. Estatísticas de Assimetria e Curtose (3º e 4º Momentos Centrais) das séries filtradas dos log-retornos dos preços do Petróleo WTI, Gasolina Convencional e valores do Índice de Óleo AMEX, comparativamente à Distribuição Normal.

	Assimetria		Excesso de Curtose	
	Pré Choques	Pós Choques	Pré Choques	Pós Choques
Óleo Cru WTI	-0.3173	-0.5215	4.4039	2.3757
Gasolina Conv.	-0.0286	-0.0720	1.5618	1.3587
Índice AMEX	-0.7669	-0.3214	12.7809	0.4399

Para as séries pode-se sugerir o uso das distribuições t-Student Simétrica ou assimétrica.

4.2.2. Perfil Empírico de Dependência

A dependência entre as séries será avaliada a fim de se obter as características gerais das interações entre os ativos energéticos. Para isso serão utilizadas estatísticas, as chamadas medidas de dependência ou concordância (veja Apêndice B).

A caracterização da dependência se dará através da estimação das medidas de concordância *"Tau de Kendall", "Rho de Spearman"*, que são medidas gerais para a obtenção do perfil relacional entre as séries, e da medida de dependência correlação linear, cujo uso nem sempre é adequado na caracterização de séries financeiras, no caso, de preços e índices de ativos energéticos, pois nem sempre existe uma estrutura de dependência linear entre elas. Outra estatística de interesse é o chamado Coeficiente de Dependência nas Caudas entre as séries. Inicialmente serão apresentada análise exploratória através de gráficos de dispersão entre as séries filtradas.

Dispersão entre Petróleo WTI e Gasolina Convencional

Apesar do preço da gasolina se movimentar com o mesmo perfil do preço do petróleo nos Estados Unidos de um modo geral, causado pelo modelo de comercialização do produto naquele país, existem fatores que causam descolamento entre o perfil das séries. Como exemplo pode-se citar que existe aumento sistemático do preço da gasolina quando é elevado o preço do petróleo, o que não acontece automaticamente quando o preço do óleo é reduzido, assim como a variação na demanda e eventuais especulações de preço do mercado (Brown & Virmani, 2007).

A figura 4.14 mostra o gráfico de dispersão de pontos das séries filtradas de log-retornos dos preços do Óleo Cru WTI e da Gasolina Convencional (porto de Nova York), períodos pré e pós 2002.

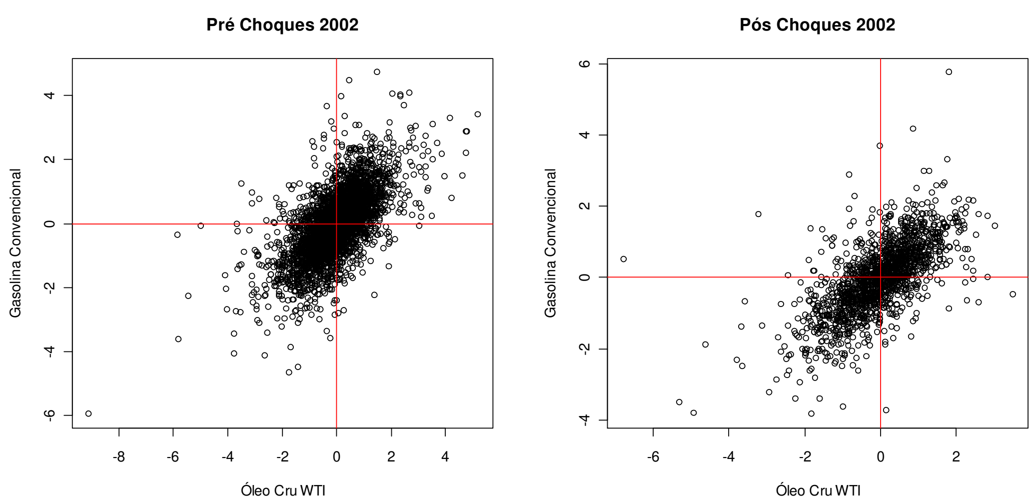

Figura 4. 16. Dispersão entre as séries filtradas de log-retornos dos preços do Petróleo WTI e da Gasolina Convencional, períodos pré e pós choques 2002.

Como se pode observar, os pontos exibem um perfil ligeiramente assimétrico na cauda das perdas. Note o perfil elíptico das interações entre as séries e a assimetria na cauda principalmente no eixo Petróleo WTI. A função de acoplamento deverá capturar esse perfil elíptico com assimetria nas caudas. Pode-se sugerir o uso de uma função de acoplamento elíptica-assimétrica como a cópula t-Student Assimétrica.

Dispersão entre Petróleo WTI e Índice Óleo AMEX

Da mesma forma que a gasolina, o Índice do Óleo AMEX possui comportamento semelhante ao preço do Óleo Cru WTI. Esse fato também pode ser atribuído à presença dos fatores geopolíticos nessa série de preços. A dispersão entre as séries filtradas de log-retornos de preços do Petróleo WTI e do índice pode ser observada na figura 4.15.

A interação entre as séries filtradas do Petróleo WTI e do Índice AMEX exibem um perfil praticamente independente no período 1, mas exibem um perfil elíptico-assimétrico no período 2. A função de acoplamento entre essas séries também pode ser a cópula t-Student Assimétrica, pois as mesma pode atender até mesmo os casos simétrico e de independência.

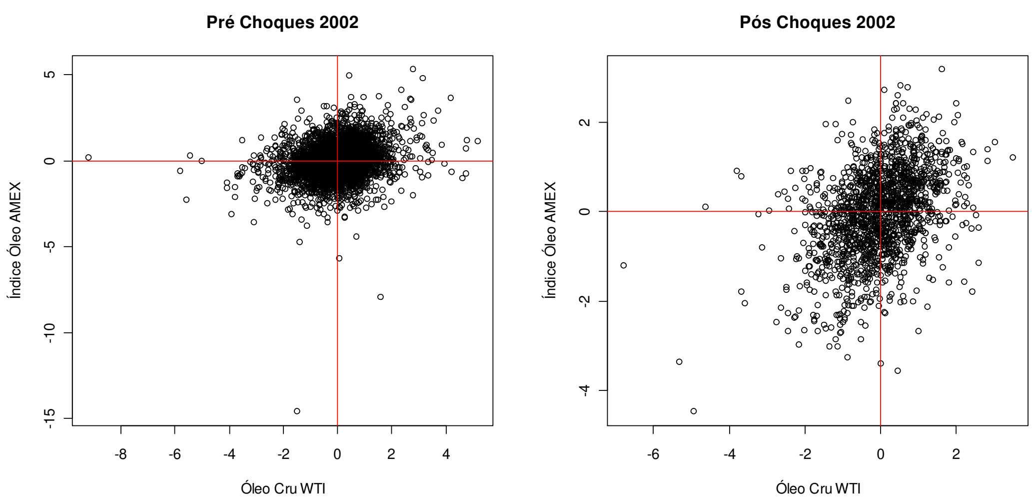

Figura 4. 17. Dispersão entre as séries filtradas de log-retornos dos preços do Petróleo WTI e Índice de Óleo AMEX, Períodos Pré e Pós Choques de 2002.

Dispersão entre Gasolina Convencional e Índice de Óleo AMEX

A figura 4.18 mostra a dispersão entre a gasolina convencional e o Índice de Óleo AMEX.

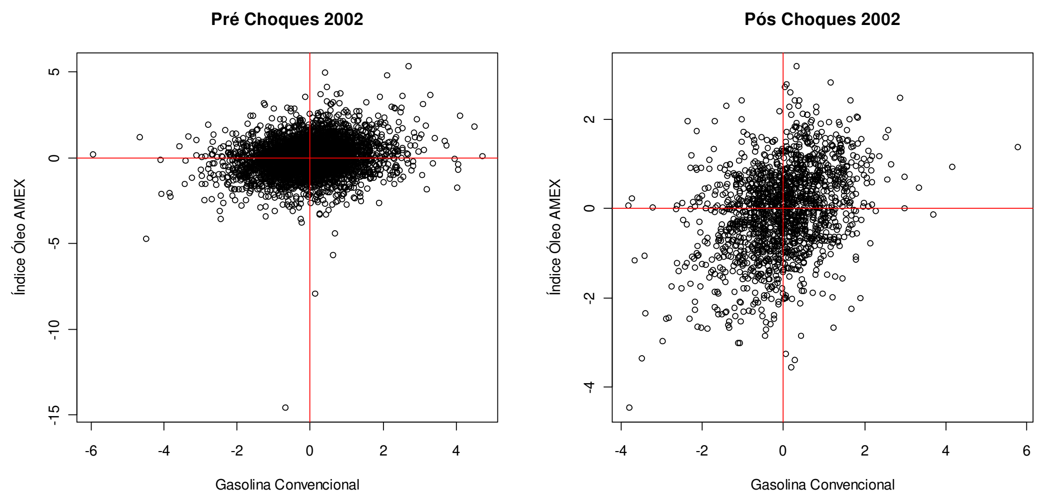

Figura 4. 18. Dispersão entre as séries filtradas de log-retornos dos preços da Gasolina Convencional e Índice de Óleo AMEX, Períodos Pré e Pós Choques de 2002.

Como se pode observar a interação entre a gasolina convencional e o Índice de Óleo AMEX no primeiro período é praticamente simétrica nos eixos horizontal e vertical, tendo o índice AMEX a menor dispersão. Adicionalmente existe indicações de independência entre as séries no período 1. No período 2 nota-se leve assimetria na dispersão entre as séries, onde pode-se usar uma função de acoplamento elíptico-assimétrica para modelagem da dependência.

Medidas de Dependência entre as Séries

O coeficiente de dependência nas caudas é calculado através da estimação da cópula empírica entre os pares de ativos energéticos (vide Apêndice A).

A tabela 4.5 mostra as estatísticas de *"Tau de Kendall"*, *"Rho de Spearman"*, Correlação Linear e Coeficiente de Dependência nas Caudas.

Tabela 4. 5. Medidas amostrais de dependência e concordância: Correlação Linear, *"Tau de Kendall", "Rho de Spearman"*, e Dependência nas Caudas Inferior e Superior calculada utilizando cópula empírica, para os pares de ativos energéticos nos períodos Pré e Pós Choques 2002.

		Dep. Caudas Inferior	Dep. Caudas Superior	Correlação Linear	Tau de Kendall	Rho de Spearman
Pré Choques 2002	Óleo Cru vs. Gasolina	0.1892	0.2973	0.6494	0.4783	0.6578
	Óleo Cru vs. Índice AMEX	0.0476	0.0952	0.2249	0.1400	0.2073
	Gasolina vs. Índice AMEX	0.0270	0.0811	0.1942	0.1210	0.1790
Pós Choques 2002	Óleo Cru vs. Gasolina	0.3673	0.2653	0.6711	0.5249	0.7063
	Óleo Cru vs. Índice AMEX	0.1892	0.0811	0.4114	0.2832	0.4144
	Gasolina vs. Índice AMEX	0.1304	0.0435	0.3411	0.2225	0.3279

Com exceção da dependência nas caudas superior, todas as medidas indicam a existência de maior dependência no segundo período. Note que a maior dependência observada é entre a gasolina e o óleo cru, e que a menor é entre a gasolina e o índice de óleo AMEX. Veja que ocorre inversão da dependência nas caudas nos períodos 1 e 2. No período 1 observava-se maior dependência entre os ganhos, enquanto que no período 2 a maior dependência é entre as perdas.

Observe que a medida de dependência nas caudas é um bom indicador da existência de assimetria entre as séries, indicando assimetria em todos os pares de séries em estudo, principalmente no 2º período.

Existência de dependência nas caudas não nula é indicativo de que, por exemplo, uma cópula Gaussiana Assimétrica não é adequada, já que possui dependência nas caudas nula. A Cópula T-Student Assimétrica é boa alternativa para estimação da dependência entre os ativos energéticos.

4.3. Modelagem das Séries Filtradas

Como o interesse são por modelos ARMA-GARCH com estrutura de dependência definida por função de acoplamento, a partir das séries filtradas devem-se ajustar modelos marginais (para a estrutura assimétrica marginal) e uma função de acoplamento (estrutura de dependência entre as séries).

Serão ajustadas e comparadas as distribuições marginais Normal Assimétrica (Seção 3.1.1) e t-Student Assimétrica (Seção 3.1.2) que, juntamente com o ajuste das funções de acoplamento t-Student (caso comum) e t-Student Assimétrica, modelarão a distribuição conjunta das séries em estudo. O ajuste das funções de acoplamento serão realizados através do método IFM (que depende do ajuste das distribuições marginais). Os resultados desses ajustes são apresentados nas seções 4.3.1 e 4.3.2. Em seguida apresentam-se métricas de bondade de ajuste (seção 4.3.3) baseadas na Função de Acoplamento Empírica e na Transformada de Kendall (Cn e Cnk, veja Genest *et al*, 2007), e também um comparativo do ajuste caudal da dependência, através de estatísticas baseadas no VaR (Valor em Risco, Morgan, 1976). O Apêndice D dá mais detalhes sobre os testes e estatísticas utilizados.

4.3.1. Ajustes de Distribuições Marginais

Para as séries filtradas de preços de Petróleo WTI, Gasolina (Porto de Nova York) e valores do Índice do Óleo AMEX, em estudo, serão ajustadas as distribuições marginais Normal Assimétrica e t-Student Assimétrica. O objetivo é não somente obter o ajuste das distribuições marginais para posterior uso na modelagem das distribuições conjuntas e do processo de preços dos ativos, mas também verificar a eficácia do ajuste de distribuições assimétricas a essas séries financeiras e observar possíveis diferenças entre períodos pré e pós os choques geopolíticos de 2002.

Inicialmente serão ajustadas as distribuições assimétricas marginais para a série filtrada de preços do Petróleo WTI no período pré choques. A figura 4.17 mostra um gráfico comparativo de quantil-quantil para o ajuste da distribuição Normal Assimétrica e t-Student Assimétrica.

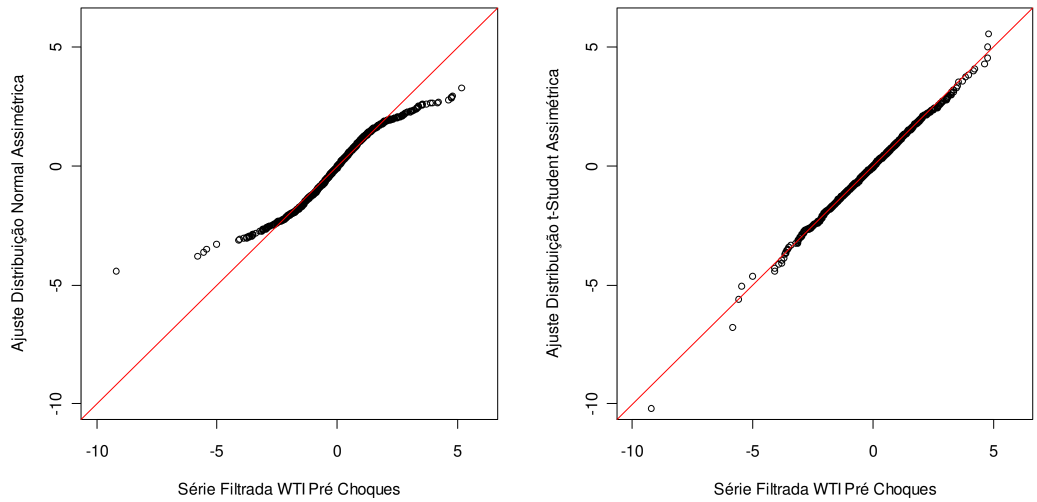

Figura 4. 19. Gráficos comparativos de probabilidades para os ajustes das distribuições Normal Assimétrica e t-Student Assimétrica à série filtrada de preços de Petróleo WTI Pré Choques Geopolíticos de 2002.

Note que o ajuste da distribuição Normal Assimétrica não se mostra adequado na região caudal da série de dados, enquanto que a distribuição t-Student Assimétrica aparentemente atendeu à estrutura distribucional da série de dados. As figuras de 4.18 a 4.22 mostram gráficos semelhantes para as séries filtradas de Petróleo WTI Pós Choques de 2002, Índice de Óleo AMEX Pré e Pós choques e Gasolina Pré e Pós Choques.

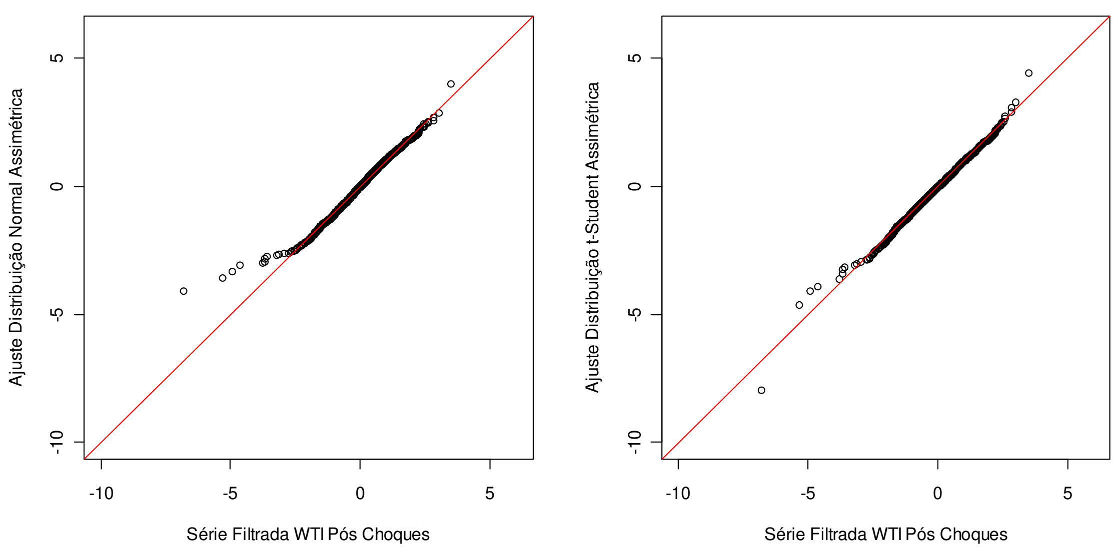

Figura 4. 20. Gráficos comparativos de probabilidades para os ajustes das distribuições Normal Assimétrica e t-Student Assimétrica à série filtrada de preços de Petróleo WTI Pós Choques Geopolíticos de 2002.

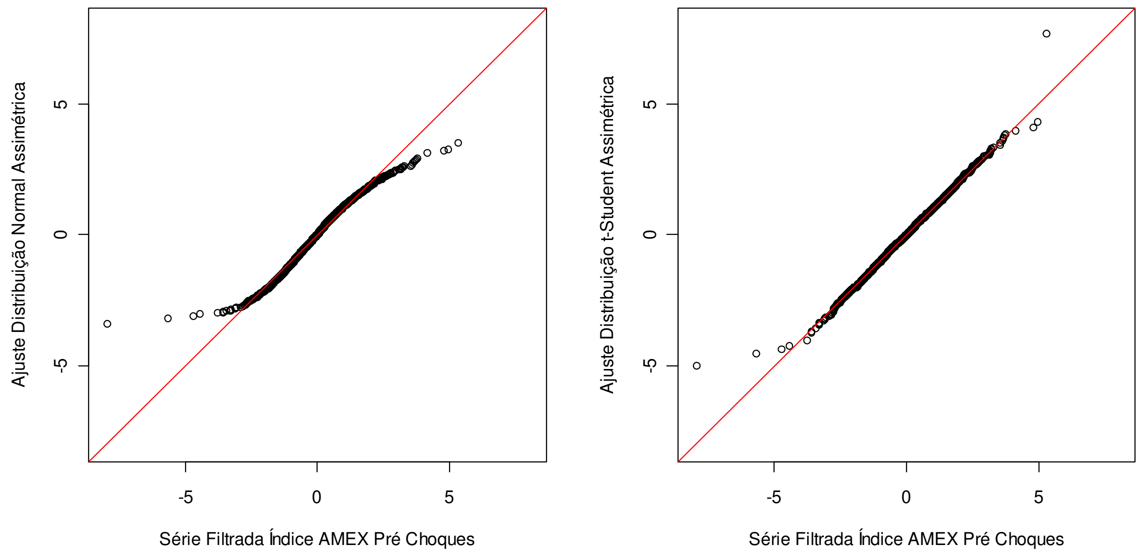

Figura 4. 21. Gráficos comparativos de probabilidades para os ajustes das distribuições Normal Assimétrica e t-Student Assimétrica à série filtrada de valores do Índice de óleo AMEX Pré Choques Geopolíticos de 2002.

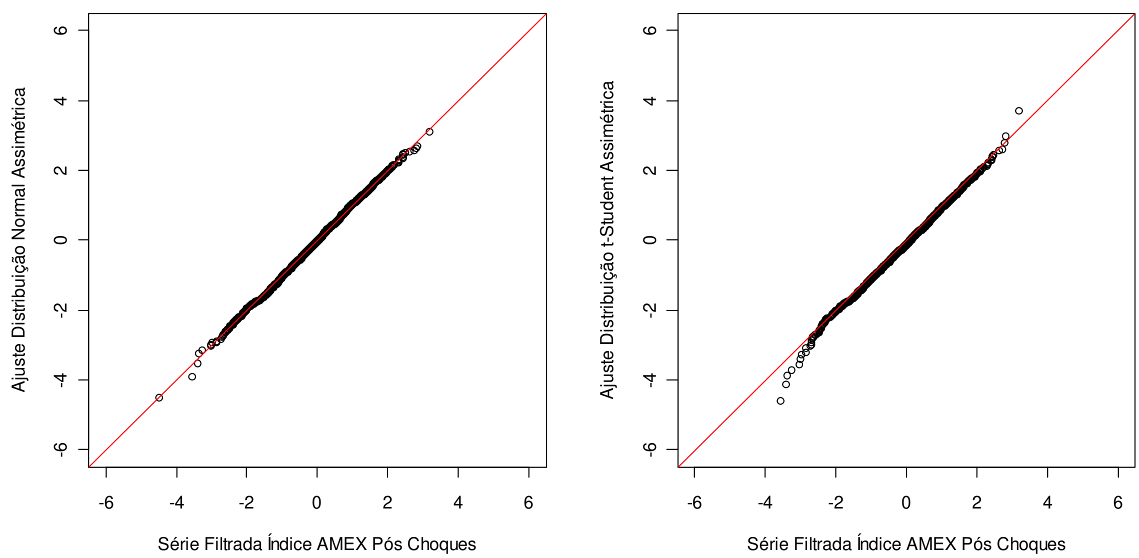

Figura 4. 22. Gráficos comparativos de probabilidades para os ajustes das distribuições Normal Assimétrica e t-Student Assimétrica à série filtrada de valores do Índice de óleo AMEX Pós Choques Geopolíticos de 2002.

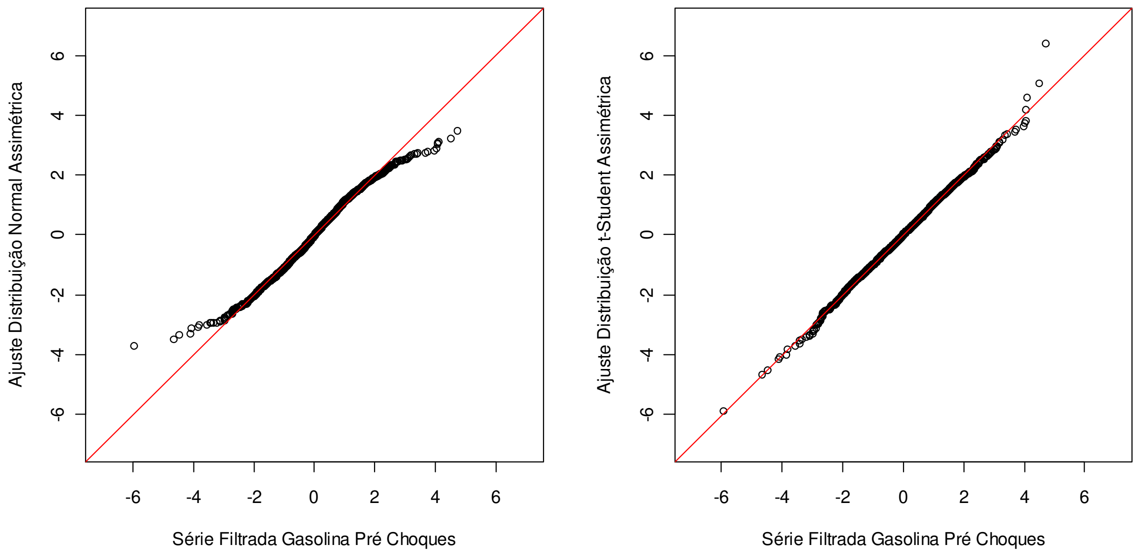

Figura 4. 23. Gráficos comparativos de probabilidades para os ajustes das distribuições Normal Assimétrica e t-Student Assimétrica à série filtrada de preços de Gasolina Pré Choques Geopolíticos de 2002.

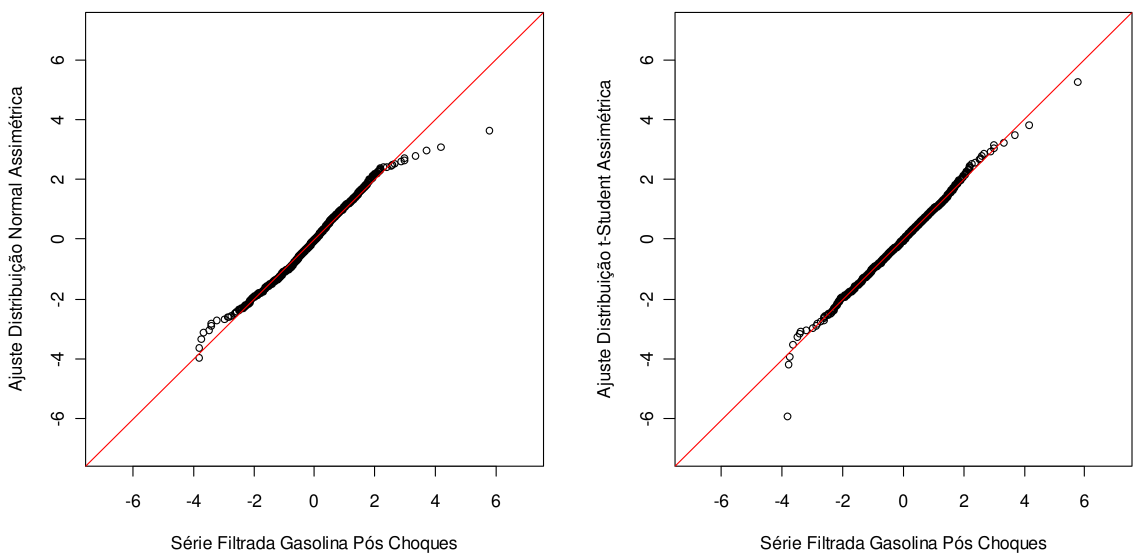

Figura 4. 24. Gráficos comparativos de probabilidades para os ajustes das distribuições Normal Assimétrica e t-Student Assimétrica à série filtrada de preços de Gasolina Pós Choques Geopolíticos de 2002.

Visualmente, através dos gráficos comparativos de probabilidade, nota-se um melhor ajuste da distribuição t-Student Assimétrica do que a distribuição Normal Assimétrica. Esses resultados são corroborados numericamente através do Critério de

Informação de Akaike (AIC, vide Akaike (1973)) e de medidas de bondade de ajuste dos modelos completos ARMA-GARCH-Cópula (Seção 4.3.3). Em todos os casos poucos são os valores extremos que não são ajustados pela distribuição t-Student Assimétrica, indicando ser desnecessário para essas séries a utilização de distribuições com caudas mais pesadas, por exemplo, baseadas na Teoria de Valores Extremos.

A tabela 4.6 mostra os parâmetros ajustados das distribuições marginais Normal e t-Student Assimétricas para a série filtrada de preços, juntamente com seus desvios padrão. Nota-se que o segundo período, pós choques de 2002, possui graus de liberdade em geral maiores do que no primeiro.

Tabela 4. 6. Parâmetros das distribuições Normal e t-Student Assimétricas ajustados para as séries marginais filtradas de preços.

		WTI Pré		WTI Pós		Óleo AMEX Pré		Óleo AMEX Pós		Gasolina Pré		Gasolina Pós	
		Param.	DP	Param.	DP	Param.	DP	Param.	DP	Param.	DP	Param.	DP
t-Student Assimétrica	Graus Liberdade	4.763	0.360	9.554	2.027	6.158	0.562	22.473	4.693	6.752	0.724	9.159	1.857
	Assimetria	-0.055	0.027	-0.301	0.127	0.025	0.035	-0.988	0.025	-0.020	0.041	-0.149	0.094
	Posição	0.060	0.031	0.302	0.128	-0.030	0.040	0.950	0.028	0.024	0.046	0.147	0.101
	Escala	0.760	0.015	0.873	0.023	0.821	0.015	0.903	0.024	0.839	0.016	0.880	0.025
Normal Assim.	Posição	0.666	0.045	0.822	0.062	0.680	0.046	0.842	0.067	0.517	0.087	0.608	0.104
	Escala	1.457	0.022	1.742	0.037	1.503	0.021	1.757	0.042	1.263	0.038	1.399	0.048
	Assimetria	-1.004	0.065	-1.399	0.112	-0.977	0.063	-1.460	0.127	-0.708	0.114	-0.897	0.145

O Critério de Informação de Akaike, que pondera a função de verossimilhança e o número de parâmetros para sugerir, dentre os modelos ajustados, qual o de melhor ajuste, é exibido na tabela 4.7. Aqui o melhor dentre os modelos é selecionado através das medidas de menor valor.

Tabela 4. 7. AIC (Akaike's Information Criterion) para o ajuste das distribuições marginais Normal Assimétrica e t-Student Assimétrica às séries filtradas de preços de Petróleo Bruto WTI, Índice de Óleo AMEX e Gasolina Convencional.

	WTI Período Pré Choques	WTI Período Pós Choques	HO2 Período Pré Choques	HO2 Período Pós Choques	Gasol. Período Pré Choques	Gasol. Período Pós Choques
Normal Assimétrica	-10500.98	-4394.23	-10653.04	-4428.32	-10747.32	-4400.02
t-Student Assimétrica	-10867.24	-4431.51	-10888.22	-4435.19	-10894.20	-4436.85

Observe que a função t-Student Assimétrica, mesmo possuindo um parâmetro a mais do que a distribuição Normal Assimétrica é a que se ajustou de forma mais adequada às séries em estudo.

4.3.2. Ajustes de Funções de Acoplamento

Nesta sessão são apresentados os resultados dos ajustes das funções de acoplamento t-Student Assimétrica e t-Student Simétrica para os preços filtrados de ativos energéticos, através do método IFM, testes de bondade de ajuste das funções de acoplamento e resultados do cálculo do VaR da simulação dos modelos ARMA-GARCH-Cópula. A função de acoplamento t-Student Simétrica é utilizada para efeito de comparação dos resultados com a contraparte assimétrica. A tabela 4.8 exibe os resultados dos ajustes dessas funções de acoplamento, juntamente com o desvio padrão das estimativas de seus parâmetros. Serão apresentados os resultados dos ajustes bivariados e 3-variado, com o objetivo de comparar as estimativas nesses casos para as cópulas mencionadas.

Tabela 4. 8. Ajustes das funções de acoplamento t-Student Assimétricas bivariadas para os pares Óleo Cru WTI e Gasolina Convencional, Óleo Cru WTI e Índice Óleo AMEX, e Gasolina Convencional e Índice Óleo AMEX nos períodos Pré e Pós choques 2002. Acompanham respectivos desvios padrão.

		Assimetria 1	Assimetria 2	Correlação	Graus Lib.
Período Pré Choques 2002	Óleo Cru WTI e Gasolina Conv.	-0.1300	0.0024	0.6770	6.6455
	Desvio Padrão	0.03549	0.03374	0.00855	0.29863
	Óleo Cru WTI e Índice Óleo AMEX	0.3966	0.1066	0.2172	13.2103
	Desvio Padrão	0.08002	0.11317	0.01783	0.77583
	Gasolina Conv. e Índice Óleo AMEX	0.4142	-0.0053	0.1938	14.8435
	Desvio Padrão	0.07786	0.08304	0.01719	0.75503
Período Pós Choques 2002	Óleo Cru WTI e Gasolina Conv.	-0.1819	-0.0440	0.7247	5.8969
	Desvio Padrão	0.0460	0.0426	0.0118	0.4673
	Óleo Cru WTI e Índice Óleo AMEX	-0.1402	-0.1885	0.4103	14.7533
	Desvio Padrão	0.0678	0.0839	0.0211	0.7335
	Gasolina Conv. e Índice Óleo AMEX	-0.3218	0.1287	0.3440	16.1096
	Desvio Padrão	0.0783	0.1014	0.0230	1.5976

Observam-se distinções principalmente na direção da assimetria entre os períodos pré e pós choques de 2002. As funções de acoplamento ajustadas para o par Óleo Cru WTI e Gasolina Convencional nesses períodos indicam maior dependência entre aumentos de preços dos ativos no período pré choques e maior dependência entre as reduções de preços no período pós choques. Observa-se também aumento na correlação do período 1 para o período 2.

Para comparação foi ajustada a função de acoplamento t-Student Simétrica que, com graus de liberdade elevados se aproxima da função de acoplamento Gaussiana. O resultado do ajuste é mostrado na tabela 4.9.

Tabela 4. 9. Ajustes das funções de acoplamento t-Student Simétrica bivariadas para os pares Óleo Cru WTI e Gasolina Convencional, Óleo Cru WTI e Índice Óleo AMEX, e Gasolina Convencional e Índice Óleo AMEX nos períodos Pré e Pós choques 2002. Acompanham respectivos desvios padrão.

	Período Pré Choques		Período Pós Choques	
	Correlação	Graus Lib.	Correlação	Graus Lib.
Óleo Cru WTI e Gasolina Conv.	0.6744	6.6825	0.7205	5.3196
Desvio Padrão	0.0086	0.7702	0.0120	0.7150
Óleo Cru WTI e Índice Óleo AMEX	0.2221	14.9415	0.4195	20.6629
Desvio Padrão	0.0160	3.9608	0.0202	9.2220
Gasolina Conv. e Índice Óleo AMEX	0.1906	15.1610	0.3422	24.2839
Desvio Padrão	0.0162	4.1739	0.0221	13.8550

Observa-se aumento expressivo dos graus de liberdade das cópulas que ajustam o Índice Óleo AMEX. Esse fato implica em redução na dependência nas caudas entre os ativos físicos e o índice AMEX. As estimativas de correlação e seus respectivos desvios padrão possuem variação muito pequena da t-Student Assimétrica para a t-Student Simétrica, mostrando robustez em relação à esse erro de especificação (uso da t-Student Simétrica no lugar da t-Student Assimétrica). As estimativas do grau de liberdade, principalmente para o primeiro período, também são próximas, entretanto os desvios padrão são mais afastados.

As funções de acoplamento apresentadas são bivariadas. As tabelas 4.10 e 4.11 mostram o ajuste das funções de acoplamento 3-variadas t-Student Assimétrica e t-Student Simétrica.

Tabela 4. 10. Ajustes das funções de acoplamento t-Student Assimétricas 3-variada para as séries filtradas de Óleo Cru WTI, Gasolina Convencional e Índice Óleo AMEX nos períodos Pré e Pós choques 2002. Acompanham respectivos desvios padrão.

	Período Pré Choques 2002		Período Pós Choques 2002	
	Parâmetros	Desvio Padrão	Parâmetros	Desvio Padrão
Assimetria 1	0.1995	0.0522	-0.3366	0.0635
Assimetria 2	0.4255	0.0604	-0.1347	0.0805
Assimetria 3	0.0851	0.0641	-0.0950	0.1105
Correlação 1,2	0.6722	0.0087	0.7152	0.0120
Correlação 1,3	0.2296	0.0161	0.4036	0.0223
Correlação 2,3	0.1977	0.0172	0.3329	0.0230
Graus Lib.	10.6197	0.5744	10.9160	0.8438

1: Óleo Cru WTI 2: Gasolina Convencional 3: Índice de Óleo AMEX

Tabela 4. 11. Ajustes das funções de acoplamento t-Student Simétrica 3-variada para as séries filtradas de Óleo Cru WTI, Gasolina Convencional e Índice Óleo AMEX nos períodos Pré e Pós choques 2002. Acompanham respectivos desvios padrão.

	Período Pré Choques 2002	Período Pós Choques 2002
Correlação 1,2	0.6738	0.7176
Desvio Padrão	0.0081	0.0113
Correlação 1,3	0.2211	0.4169
Desvio Padrão	0.0160	0.0210
Correlação 2,3	0.1905	0.3450
Desvio Padrão	0.0162	0.0225
Graus Lib.	11.4087	10.2755
Desvio Padrão	1.3265	1.4238

1: Óleo Cru WTI 2: Gasolina Convencional
3: Índice de Óleo AMEX

Observe que os parâmetros de correlação e graus de liberdade das cópulas t-Assimétrica 3-variada e t-Student simétrica 3-variada são semelhantes, estando nos intervalos de confiança das estimativas.

Comparando-se com as cópulas bivariadas observam-se grandes diferenças nas estimativas dos parâmetros (exceto as correlações). Enquanto que na bivariada obtiveram-se parâmetros de assimetria referentes ao Óleo Cru WTI no valor de -0.13 e 0.39, nas cópulas relacionadas com a Gasolina e o Índice AMEX no período 1, respectivamente, obteve-se aproximadamente 0.20 no mesmo período na cópula 3-variada. Isso indica que a cópula 3-variada considerou um parâmetro próximo da média dos parâmetros bivariados. O mesmo ocorre com os outros parâmetros. O ajuste do acoplamento 3-variado t-Student

Simétrico não possuiu graus de liberdade tão grandes quanto os observados no caso bivariado. Isso pode ser devido ao alto desvio padrão das estimativas bivariadas, o que não trazem garantias firmes de que o grau de liberdade estimado foi bem ajustado. Outro fato importante de ser mencionado é que a função de acoplamento 3 dimensional restringe o grau de liberdade a um valor único (no caso bivariado existe um grau de liberdade para cada par). Se os graus de liberdade dos casos bivariados são muito diferentes do caso 3-variado é possível que se esteja "engessando" a estrutura de dependência através das médias dos graus de liberdade.

Esses resultados indicam que ao se ajustar a cópula t-Student Assimétrica pode-se estimar a matriz de correlação através dos modelos bivariados. Isso diminui de forma expressiva o número de parâmetros da função de acoplamento a ser estimada. Dessa forma, a cada nova variável aumenta-se apenas 1 parâmetro adicional de assimetria, viabilizando a utilização da cópula para dimensões superiores.

Bondades de Ajuste

Foram realizados testes de bondade de ajuste das funções de acoplamento 3-variadas. Os p-valores dos testes da cópula empírica e da transformada de Kendall são mostrados na tabela 4.12, para os períodos pré e pós choques de 2002. Esses testes são baseados apenas nas funções de acoplamento ajustadas, desconsiderando o ajuste das distribuições marginais e modelos ARMA-GARCH.

Tabela 4. 12. P-Valores dos testes de bondade de ajuste baseados na Cópula Empírica e na Transformada Kendall das funções de acoplamento t-Student Assimétrica e t-Student Simétrica nos períodos pré e pós choques de 2002.

Período 1	t-Student Assimétrica	t-Student Simétrica
Cópula Empírica	0.565	0.578
Transformada Kendall	0.965	0.298

Período 2	t-Student Assimétrica	t-Student Simétrica
Cópula Empírica	0.109	0.181
Transformada Kendall	0.026	0.036

Observe que o teste gerou p-valores que não rejeitam a hipótese nula (de modelos de funções de acoplamentos adequados) no período pré choques de 2002. Nesse período observa-se no entanto um melhor ajuste do acoplamento t-Student Assimétrico. Em contrapartida, no período pós choques de 2002 observou-se ajustes menos adequados. O teste baseado na Transformada Kendall rejeita a hipótese de adequação das funções de acoplamento com nível α de 0.05. O teste baseado na Cópula Empírica não rejeita a hipótese de adequação, entretanto qualifica a cópula t-Student Simétrica como melhor ajustada (se comparados os resultados dos testes das cópulas Simétrica e Assimétrica).

As estatísticas baseadas no VaR (ver metodologia no apêndice) são mostradas nas tabelas 4.13 (período pré choques 2002) e 4.14 (período pós choques) para duas carteiras construídas a partir dos 2 atívos físicos mais o Índice Óleo AMEX. A primeira carteira é constituída igualitariamente pelos 3 itens. A segunda é constituída de 50% de Óleo Cru WTI, 20% de Gasolina Convencional e 30% do Índice de Óleo AMEX. Esse teste considera o ajuste completo do modelo ARMA-GARCH-Cópulas, e, se o ajuste caudal foi adequado pelas funções de acoplamento, os resultados apresentados devem ser próximos do VaR de nível α (ou 1-α no caso de níveis maiores ou iguais a .9). Mais detalhes sobre a geração das estatísticas pode ser verificado no apêndice D.

Tabela 4. 13. Estatísticas do VaR (Valor em Risco) para simulações em passos de 50 dias (janela de 1000 observações) dos modelo ARMA-GARCH-Cópula t-Student Assimétrica e ARMA-GARCH-Cópula t-Student Simétrica para o período 1. Número de observações para cálculo do VaR é de 2841.

	t-Student Assimétrica		t-Student Simétrica		t-Student Assimétrica		t-Student Simétrica	
Período 1	Carteira 1	Ocorrências	Carteira 1	Ocorrências	Carteira 2	Ocorrências	Carteira 2	Ocorrências
VaR 0.01%	0.04%	1	0.04%	1	0.00%	0	0.04%	1
VaR 0.1%	0.21%	6	0.07%	2	0.32%	9	0.21%	6
VaR 0.5%	0.77%	22	0.67%	19	0.81%	23	0.74%	21
VaR 1%	1.41%	40	1.16%	33	1.51%	43	1.23%	35
VaR 5%	6.58%	187	6.23%	177	6.58%	187	6.62%	188
VaR 10%	11.33%	322	11.02%	313	11.23%	319	10.84%	308
VaR 90%	11.69%	332	11.30%	321	11.76%	334	11.47%	326
VaR 95%	5.77%	164	5.70%	162	5.74%	163	5.49%	156
VaR 99%	1.51%	43	1.37%	39	1.48%	42	1.51%	43
VaR 99.5%	0.88%	25	0.81%	23	0.77%	22	0.81%	23
VaR 99.9%	0.14%	4	0.11%	3	0.14%	4	0.21%	6
VaR 99.99%	0.00%	0	0.04%	1	0.04%	1	0.11%	3

Observa-se, tanto na carteira 1 quanto na carteira 2 proximidade entre as estatísticas da t-Student Assimétrica e t-Student Simétrica. Na carteira 2 foi subestimado o VaR na cauda superior da t-Student Simétrica, causado exatamente pelo seu comportamento simétrico. A tabela 4.14 mostra as estatísticas para o período pós choques de 2002.

Tabela 4. 14. VaR (Valor em Risco) para simulações em passos de 50 dias (janela de 1000 observações) dos modelo ARMA-GARCH-Cópula t-Student Assimétrica e ARMA-GARCH-Cópula t-Student Simétrica para o período 1. Número de observações para cálculo do VaR é de 556.

	t-Student Assimétrica		t-Student Simétrica		t-Student Assimétrica		t-Student Simétrica	
Período 2	Carteira 1	Ocorrências	Carteira 1	Ocorrências	Carteira 2	Ocorrências	Carteira 2	Ocorrências
VaR 0.1%	0.18%	1	0.18%	1	0.18%	1	0.18%	1
VaR 0.5%	0.71%	4	0.71%	4	0.53%	3	0.71%	4
VaR 1%	1.06%	6	1.06%	6	0.71%	4	0.88%	5
VaR 5%	5.83%	33	6.01%	34	6.54%	37	6.71%	38
VaR 10%	14.31%	81	14.13%	80	14.31%	81	14.13%	80
VaR 90%	9.54%	54	9.54%	54	9.36%	53	9.36%	53
VaR 95%	4.95%	28	4.95%	28	4.77%	27	4.77%	27
VaR 99%	0.71%	4	0.53%	3	0.88%	5	0.71%	4
VaR 99.5%	0.18%	1	0.00%	0	0.18%	1	0.18%	1

Nesse caso observa-se grande semelhança entre as carteiras e também entre as funções de acoplamento t-Student simétrica e assimétrica. Observe que ocorreram o mesmo número de ultrapassagens do VaR 90% na cópula t-Student Assimétrica e t-Student Simétrica, em ambas as carteiras, o que aconteceu também com outros níveis de VaR.

4.3.3. Comparação dependência nas Caudas

Nesta seção serão verificadas as dependências nas caudas das cópulas ajustadas (teóricas) e do acoplamento empírico. A tabela 4.15 mostra as dependências nas caudas inferior e superior estimada através da cópula empírica e calculada através de simulação da cópula t-Student Assimétrica e teórica da t-Student Simétrica. Foram utilizados os parâmetros estimados pelos acoplamentos 3-variados.

Tabela 4. 15. Dependência nas caudas Inferior e Superior das funções de acoplamento empírica, t-Student Assimétrica e t-Student Simétrica ajustadas.

		Inferior Empírica	Inferior t-Assimétrica	Inferior t-Simétrica	Superior Empírica	Superior t-Assimétrica	Superior t-Simétrica
Pré Choques 2002	Óleo Cru vs. Gasolina	0.1892	0.1762	0.1451	0.2973	0.3101	0.1451
	Óleo Cru vs. Índice AMEX	0.0476	0.0321	0.0152	0.0952	0.0666	0.0152
	Gasolina vs. Índice AMEX	0.0270	0.0167	0.0128	0.0811	0.0780	0.0128
Pós Choques 2002	Óleo Cru vs. Gasolina	0.3673	0.3231	0.1999	0.2653	0.2191	0.1999
	Óleo Cru vs. Índice AMEX	0.1892	0.1292	0.0537	0.0811	0.0590	0.0537
	Gasolina vs. Índice AMEX	0.1304	0.0909	0.0384	0.0435	0.0555	0.0384

Observa-se aqui que a dependência nas caudas do ajuste da cópula t-Student Assimétrica ficou próximo da dependência nas caudas empírica. A cópula t-Student Simétrica subestimou a dependência nas caudas na maioria das vezes.

4.3.4. Ganhos em Relação à Função de Acoplamento t-Student Simétrica

Não é simples estabelecer na prática os ganhos que a função de acoplamento t-Student Assimétrica têm em relação à t-Student Simétrica. Observou-se que a distribuição assimétrica conseguiu um ajuste da dependência nas caudas mais próximo do observado empiricamente. Entretanto as estatísticas baseadas no VaR mostram apenas uma leve subestimação do VaR pela t-Student Simétrica, causada pela necessidade de ajuste assimétrico. Observou-se também, no teste de bondade de ajustes que provavelmente a função de acoplamento t-Student Simétrica se ajustou melhor. Entretanto esse teste dá o mesmo peso à distribuição em toda sua extensão, desprezando, então, uma maior dependência nas caudas.

Análises e aplicações em outros conjuntos de dados podem ser úteis para estabelecer de forma mais confiável o grau de melhoria pelo uso da função de acoplamento t-Student Assimétrica em detrimento da Simétrica.

Além disso, a distribuição t-Student Simétrica é mais simples de ser implementada computacionalmente, gerando resultados mais rapidamente, mesmo para dimensões superiores.

Capítulo 5

Conclusões

Procurou-se mostrar nesta dissertação que a distribuição t-Student Assimétrica é uma boa alternativa, tanto para a modelagem da distribuição marginal quanto para a constituição de uma função de acoplamento que seja flexível o suficiente para caracterizar a estrutura de dependência principal e caudal de um conjunto de dados.

Mostrou-se que o acoplamento t-Student Assimétrico pode descrever uma grande variedade de dependências nas caudas superior e inferior, desde situações simétricas até assimétricas de dependência. A função de acoplamento t-Student Assimétrica é uma boa alternativa para modelagem de eventos extremos em dados reais. Verificou-se que essa distribuição ajusta-se adequadamente à dependência nas caudas, considerando-se como comparação a dependência nas caudas empírica. Observou-se também através das estimações que essa cópula também pode compor um modelo ARMA-GARCH-Cópulas que caracterize um conjunto de dados reais.

Procurou-se mostrar que é possível fazer a estimação conjunta dos parâmetros da função de acoplamento t-Student Assimétrica pela máxima verossimilhança através de métodos computacionais de otimização construídos a partir da função densidade do acoplamento. Os resultados dos parâmetros estimados ficaram em geral próximos dos parâmetros simulados, o que dá maior garantia de que a estimação dos parâmetros para dados reais caracterize corretamente a dependência entre eles. Nesse contexto mostrou-se também que alterações na precisão não modificam a estimativa dos parâmetros drasticamente, mas tem influencia no cálculo do desvio padrão das estimativas. Adicionalmente, é possível aumentar a dimensão da função de acoplamento, com aumento do custo computacional restrito a um parâmetro por dimensão, já que os parâmetros de correlação podem ser ajustados com fidelidade pela função de acoplamento bivariada.

Apêndices

Apêndice A

Funções de Acoplamento

Segundo Nelsen (2006, Cap 1), acoplamentos podem ser vistos de duas formas: "sob um ponto de vista, acoplamentos são funções que juntam ou "acoplam" funções distribuições conjuntas às suas funções distribuições marginais. Alternativamente, acoplamento são funções distribuições multivariadas, cujas marginais unidimensionais são uniformes no intervalo (0,1)". Atualmente, o estudo de funções de acoplamento é de grande interesse, principalmente por suas aplicações em finanças e atuária, como, por exemplo, a análise de risco de mercado, o cálculo do risco de um portfólio de seguros, o apreçamento de derivativos, entre outros.

A Transformada Integral de Probabilidade

Para um bom entendimento das funções de acoplamento, é necessário definir uma ferramenta de extrema importância, que é a transformada integral de probabilidade:

Seja X uma variável aleatória com função de distribuição $F(x) = P(X \leq x), x \in \Re$, sendo isso denotado por $X \sim F$. Suponha que $F(x)$ é contínua. Então, para $u \in (0,1)$, existe um valor mínimo único $x(u)$, tal que $F(x(u)) = u$. Formalmente,

$$x(u) = F^{-1}(u) = \inf\{x / F(x) \geq u\}, \qquad \text{[A-1]}$$

o qual define a função distribuição inversa. Então $F(x) \le u \Leftrightarrow x \le F^{-1}(u)$. Uma vez que $F(x)$ é não-decrescente e contínua, então sua inversa $F^{-1}(u)$ também é não-decrescente e contínua sobre $u \in (0,1)$. Portanto,

$$P(F(X) \le u) = P(X \le F^{-1}(u)) = F(F^{-1}(u)) = u$$

e

$$P(F(X) \le u) = \begin{cases} 1 & \text{se}\, u \ge 1 \\ 0 & \text{se}\, u < 0 \end{cases} \quad . \qquad \text{[A-2]}$$

Portanto,

$$P(F(X) \le u) = \begin{cases} 1 & \text{se}\, u \ge 1 \\ u & \text{se}\, 0 \le u \le 1 \\ 0 & \text{se}\, u < 0 \end{cases}$$

Conseqüentemente, $F(X)$ tem distribuição uniforme em $[0,1]$, i.e., $F(X) \sim Uniforme(0,1)$. A transformação $U = F(X)$ é chamada transformada integral de probabilidade.

Lema A.1: Para quaisquer variáveis aleatórias contínuas $X \sim F$ e $U \sim Uniforme(0,1)$, X e $F^{-1}(U)$ tem a mesma função distribuição.

Realmente, uma vez que $F^{-1}(u)$ é não decrescente tem-se que:

$$P(F^{-1}(u) \le x) = P(U \le F(x)) = F(x).$$

Proposição A.1: Seja $G(x)$ uma função distribuição qualquer e $U \sim Uniforme(0,1)$. Então existe uma função $H(x)$ tal que $P(H(U) \le x) = G(x), x \in \Re$.

Prova: Se $G(x)$ é contínua, tomamos $h(U) = G^{-1}(U)$. Então

$$P(h(U) \le x) = P(U \le G(x)) = G(x).$$

Se $G(x)$ tem saltos de tamanho p_k em y_k, então tomamos $G(y)$ para $x \in (p_1 + p_2 + \cdots + p_{k-1}, p_1 + p_2 + \cdots + p_k], k = 1,2,...$ (assumindo $p_0 = 0$). Uma vez que $U \sim U(0,1)$, $P(h(U) = y_k) = P(p_1 + \cdots + p_{k-1} < U \leq p_1 + \cdots + p_k) = p_k$.

Portanto $h(U)$ é uma variável aleatória discreta com função distribuição $G(x)$.

Então, se Y tem densidade $g(y)$ e função distribuição $G(y)$, então $Y = G^{-1}(F(x))$ transforma $f(.)$ em $g(.)$.

A.1. Distribuições Multivariadas

A dependência entre as variáveis aleatórias de valores reais $X_1, X_2,..., X_n$ é completamente descrita pela função de distribuição conjunta

$$H(x_1, x_2,..., x_n) = P(X_1 \leq x_1,\ X_2 \leq x_2,...,\ X_n \leq x_n). \qquad \text{[A-3]}$$

Se $H(x_1, x_2,..., x_n)$ é conhecida, podemos obter a densidade conjunta (se a mesma existir) de acordo com a formula:

$$h(x_1, x_2,..., x_n) = \frac{\partial^n H(x_1, x_2,..., x_n)}{\partial x_1 \partial x_2 ... \partial x_n}, \qquad \text{[A-4]}$$

para todo $(x_1, x_2,..., x_n) \in \Re^n$ onde $H(x_1, x_2,..., x_n)$ tem derivada parcial mista.

Para qualquer subconjunto $\{i_1, i_2,..., i_j\} \in \{1,2,..., n\}$ a distribuição marginal de $(X_{i_1}, X_{i_2},..., X_{i_j})$ é dada por

$$H_{X_{i_1}, X_{i_2},..., X_{i_j}}(x_{i_1}, x_{i_2},..., x_{i_j}) = \lim_{x_i \to \infty} H(x_1, x_2,..., x_n), i \neq i_1, i_2,..., i_j. \qquad \text{[A-5]}$$

A.2. Definição das Funções de Acoplamento

Suponha que se transforme o vetor aleatório $(X_1, X_2,..., X_n)$ componente a componente para que ele tenha distribuições marginais $Uniforme(0,1)$. Se $X_1, X_2,..., X_n$ têm distribuições marginais contínuas $F_{X_1}(x_1),..., F_{X_n}(x_n)$, então essa transformação pode ser obtida usando-se a transformada integral de probabilidade:

$$T_n : \Re^n \mapsto \Re^n, \quad (x_1,..., x_n) \mapsto (F_{X_1}(x_1),..., F_{X_n}(x_n)). \qquad [A-6]$$

A função distribuição conjunta $C(.)$ de $(F_{X_1}(x_1),..., F_{X_n}(x_n))$ é então chamada de cópula (ou função de acoplamento, ou, muitas vezes, apenas acoplamento) do vetor aleatório $(X_1, X_2,..., X_n)$, ou equivalentemente, cópula associada com a distribuição conjunta $H(.)$.

Segue que:

$$\begin{aligned} &H(x_1, x_2,..., x_n) = \\ &= P(F_{X_1}(X_1) \le F_{X_1}(x_1),\ F_{X_2}(X_2) \le F_{X_2}(x_2),...,\ F_{X_n}(X_n) \le F_{X_n}(x_n)) \\ &= C(F_{X_1}(x_1),\ F_{X_2}(x_2),..., F_{X_n}(x_n). \end{aligned} \qquad [A-7]$$

Definição A.1: Uma função de acoplamento é definida como a função de distribuição conjunta

$C(u_1, u_2,..., u_n) = P(U_1 \le u_1,\ U_2 \le u_2,...,\ U_n \le u_n),\ 0 \le \mathrm{u_i} \le 1$,

sendo $U_i \sim Uniforme(0,1), i = 1,2,..., n$ ou alternativamente

Definição A.2: Uma função de acoplamento é qualquer função $C : [0,1]^n \to [0,1]$ que possui as seguintes propriedades:

i) $C(u_1,..., u_n)$ é não-decrescente em cada componente u_i

ii) $C(1,...,1, u_i, 1,...,1) = u_i \in [0,1]$ $\forall$ $i = 1,2,..., n$

iii) para $\forall$ $(a_1,..., a_n), (b_1,..., b_n) \in [0,1]^n$ com $a_i \le b_i$ tem-se que

$\sum_{i_1=1}^{2}\dots\sum_{i_n=1}^{2}(-1)^{i_1+i_2+\dots+i_n} C(x_{1i_1},\dots,x_{ni_n}) \geq 0$, onde $x_{j1} = a_j$ e $x_{j2} = b_j$, $j = 1,2,\dots,n$.

Sendo essas definições equivalentes.

A.2.1. Propriedades

Apresentaremos agora algumas propriedades e teoremas que as funções de acoplamento satisfazem. Maiores detalhes podem ser obtidos em Nelsen (2006) e Bouyé et al. (2000).

Teorema A.1 (Sklar): Seja $H(.)$ uma função de distribuição conjunta com marginais $F_{X_1},\dots,F_{X_n}$. Então existe uma função de acoplamento n-dimensional $C(.)$ tal que:

$$H(x_1, x_2,\dots,x_n) = C(F_{X_1}(x_1), F_{X_2}(x_2),\dots, F_{X_n}(x_n)). \qquad \text{[A-8]}$$

Reciprocamente, se $C(.)$ é uma função de acoplamento n-dimensional e $F_{X_1},\dots,F_{X_n}$ são funções de distribuição, então $H(.)$, definida por [A-8] é uma distribuição conjunta n-dimensional.

Note que a recíproca desse teorema implica que podemos acoplar quaisquer tipos de distribuições univariadas com qualquer função de acoplamento, assim definindo uma distribuição multivariada válida.

A prova desse teorema pode ser encontrada em Sklar (1959). O corolário a seguir mostra que podemos obter funções de acoplamento a partir de qualquer distribuição multivariada e usá-la independentemente das marginais originais.

Corolário A.1: Seja $H(.)$ uma função de distribuição conjunta com marginais $F_{X_1},\dots,F_{X_n}$ e sejam $F^{-1}{}_{X_1},\dots,F^{-1}{}_{X_n}$ as inversas de $F_{X_1},\dots,F_{X_n}$ respectivamente. Então, para qualquer $u_i \in [0,1]$, existe uma função de acoplamento n-dimensional $C(.)$ tal que:

$$C(u_1, u_2,..., u_n) = H(F^{-1}{}_{X_1}(u_1),..., F^{-1}{}_{X_n}(u_n)) \qquad \text{[A-9]}$$

Então, podemos estudar o fenômeno de dependência entre as variáveis sem fixar as distribuições marginais.

Teorema A.2: Seja $(X_1, X_2,..., X_n)$ um vetor de variáveis aleatórias contínuas com função de acoplamento $C(.)$. Se $\alpha_1(.),...,\alpha_n(.)$ são funções estritamente crescentes em $\mathrm{Im}(X_1),...,\mathrm{Im}(X_n)$, respectivamente, então $(\alpha_1(X_1),...,\alpha_n(X_n))$ também possui acoplamento $C(.)$.

Prova: Sejam $F_{X_1},\cdots,F_{X_n}$ funções distribuição de $X_1,\cdots,X_n$ e sejam $G_1,\cdots,G_n$ as funções distribuição de $\alpha_1(X_1),\cdots,\alpha_1(X_n)$ respectivamente . Seja $(X_1,\cdots,X_n)$ um vetor aleatório com cópula $C(.)$ e seja $C_\alpha(.)$ a cópula de $(\alpha_1(X_1),\cdots,\alpha_1(X_n))$. Uma vez que $\alpha_k(.)$ é estritamente crescente para cada $k = 1,\cdots,n$,

$$G_k(x) = P(\alpha_k(X_k) \le x) = P(X_k \le \alpha_k^{-1}(x)) = F_k(\alpha_k^{-1}(x)).$$

para todo x em $\Re$, portanto

$$\begin{aligned} &C_\alpha(G_1(x_1),..., G_n(x_n)) = P(\alpha_1(X_1) \le x_1,..., \alpha_n(X_n) \le x_n) = \\ &= P(X_1 \le \alpha_1^{-1}(x_1),..., X_n \le \alpha_{n1}^{-1}(x_n)) = C(F_1(\alpha_1^{-1}(x_1)),..., F_n(\alpha_n^{-1}(x_n))) = \\ &= C(G_1(x_1),..., G_n(x_n)). \end{aligned} \qquad \text{[A-10]}$$

Uma vez que $X_1,..., X_n$ são contínuas, $\mathrm{Im}(G_1) = ... = \mathrm{Im}(G_n) = [0,1]$, i.e., $C_\alpha(.) = C(.)$ em $[0,1]^n$.

■

Então, as transformações não lineares estritamente crescentes alteram a distribuição conjunta $H(.)$ e suas marginais, mas a forma analítica da função de acoplamento permanece a mesma.

Teorema A.3: $(X_1, X_2, ..., X_n)$ um vetor de variáveis aleatórias contínuas com função de acoplamento $C(.)$. Então $X_1, X_2, ..., X_n$ são independentes se e somente se

$$C(\mathbf{u}) = \prod_{i=1}^{n} u_i = \prod(\mathbf{u}) \qquad \text{[A-11]}$$

$\prod$ é chamada de cópula de independência (ou independente).

A.2.1. Limites de Fréchet-Hoeffding para Cópulas

Definição 2.3 (Acoplamentos W e M): Sejam M e W funções com domínio em $[0,1]^n$, tais que:

$$M^n(\mathbf{u}) = \min(u_1, u_2, ..., u_n)\,,$$

$$W^n(\mathbf{u}) = \max(u_1 + u_2 + ... + u_n - n + 1 \quad , \quad 0)\,. \qquad \text{[A-12]}$$

Então M é uma função de acoplamento n-dimensional para todo $n \geq 2$, ao passo que W não é uma função de acoplamento para todo $n \geq 3$.

Teorema A.4 (Limites de Fréchet-Hoefding): Se $C(\mathbf{u})$ é um acoplamento qualquer, então, para todo $\mathbf{u} \in [0,1]^n$,

$$W^n(\mathbf{u}) \leq C(\mathbf{u}) \leq M^n(\mathbf{u})\,. \qquad \text{[A-13]}$$

Portanto, o limite inferior de Fréchet-Hoefding $W^2(u,v) = \max(u + v - 1, 0)$ é menor do que qualquer cópula bidimensional, e qualquer cópula $C(\mathbf{u})$ é menor que o limite superior de Fréchet-Hoefding $M^n(\mathbf{u})$.

Prova (caso bidimensional):

1ª parte: Seja $(u,v) \in [0,1]^2$, temos por definição que $C(u,v) \le C(u,1) = u$ e $C(u,v) \le C(1,v) = v$. Diretamente, isso implica que $C(u,v) \le \min(u,v)$. .

2ª parte: Utilizando a definição de cópula (Definição 2.2, item iii), com $u_2 = v_2 = 1$, temos que $C(1,1) - C(u,1) - C(1,v) + C(u,v) \ge 0$. Então $1 - u - v + C(u,v) \ge 0$ e $C(u,v) \ge u + v - 1$. Como $C(u,v) \ge 0$, $C(u,v) \ge \max(u + v - 1\ ,0)$.

■

Apêndice B

Dependência

O conceito de dependência entre variáveis aleatórias está relacionado com a probabilidade conjunta entre os eventos possíveis descritos pelas distribuições marginais. Como Jogdeo (1982) aponta,

> "As relações de dependência entre variáveis aleatórias é um dos temas mais estudados em probabilidade e estatística. A natureza da dependência pode tomar uma variedade de formas e, a não ser que alguma suposição específica seja feita sobre a dependência, nenhum modelo estatístico pode ser criado."

É importante então, para que se possa fazer uma boa modelagem, que se utilize funções que caracterizem corretamente a dependência entre os eventos em estudo. Joe (1997, Cap 2, p. 19-33), apresenta diversos tipos de dependência que são encontradas nas diversas funções conjuntas e/ou de acoplamento. A dependência entre variáveis aleatórias é totalmente definida por sua função de distribuição conjunta, porém, esta não informa especificamente como é essa dependência, nem a mensura. Para isso são desenvolvidas medidas de dependência, que são instrumentos que quantificam essa associação. Cabe ao analista/estatístico selecionar corretamente a função mais adequada ao seu problema.

Serão abordados inicialmente alguns dos diferentes tipos de dependência, mais úteis nas aplicações em gerenciamento de risco, como simetria e assimetria, dependência em quadrantes positivo e negativo, dependência com crescimento estocástico, dependência crescente nas caudas e dependência positiva total.

B.1. Propriedades da Dependência

Para criar um modelo estocástico para algum determinado conjunto de eventos e variáveis aleatórias é necessário que se façam suposições sobre o comportamento aleatório dos mesmos. Essas suposições podem estar baseadas em alguns conceitos e estruturas de dependência para caracterizar o comportamento.

Por estrutura de dependência pode-se entender as características intrínsecas da dependência entre eventos aleatórios de um conjunto de variáveis aleatórias, podendo ser, por exemplo, simétrica ou assimétrica, alongada ou não nas caudas. Entende-se por conceito as características e propriedades que uma determinada estrutura de dependência possui.

Serão apresentadas aqui definições dos conceitos e estruturas de dependência para o caso bivariado, sendo o caso multivariado uma extensão natural dos mesmos conceitos e estruturas.

B.1.1. Dependência em Quadrante Positiva

Este conceito de dependência distingue dependência de independência, quando duas variáveis aleatórias assumem valores em um mesmo quadrante, e é definida por:

Definição B.1: Seja $X = (X_1, X_2)$ um vetor aleatório bivariado com distribuição acumulada (f.d.a.) F. X ou F é dependente em quadrante positivo se:

$$P(X_1 > a_1; X_2 > a_2) \geq P(X_1 > a_1)P(X_2 > a_2) \forall a_1, a_2 \in \Re, \qquad \text{[B-1]}$$

ou equivalentemente:

$$P(X_1 \le a_1; X_2 \le a_2) \ge P(X_1 \le a_1)P(X_2 \le a_2) \forall a_1, a_2 \in \Re . \qquad \text{[B-2]}$$

A razão por que [B-1] e [B-2] são um conceito de dependência positiva é que é mais provável que X_1 e X_2 assumam valores grandes conjuntamente ou pequenos conjuntamente comparados com X'_1 e X'_2, onde $X_1 =_d X_1'$ e $X_2 =_d X_2'$, sendo X'_1 e X'_2 variáveis aleatórias independentes. Da mesma forma, X ou F são dependentes em quadrante negativo se as inequações em [B-1] e [B-2] forem reversas.

Nota-se que esse conceito é geral, pois apenas distingue um par de variáveis aleatórias dependentes de um par independente.

B.1.2. Dependência de Crescimento Estocástico

Este conceito diz respeito ao crescimento estocástico da dependência entre duas variáveis aleatórias conforme o valor de uma das variáveis aumenta.

Definição B.2: Seja $X = (X_1, X_2)$ um vetor aleatório bivariado com distribuição acumulada (f.d.a.) $F \in \mathcal{F}(F_1, F_2)$. X_2 é estocasticamente crescente em X_1, ou a distribuição condicional $F_{2/1}$ é estocasticamente crescente se:

$$P(X_2 > x_2 \mid X_1 = x_1) = 1 - F_{2/1}(x_2 \mid x_1) \uparrow x_1 \forall x_2 . \qquad \text{[B-3]}$$

Se forem invertidos os papéis de X_1 e X_2 em [B-3] então obtém-se X_1 estocasticamente crescente em X_2, ou $F_{1/2}$ estocasticamente crescente. Entende-se por esse conceito que X_2 é mais provável de assumir valores maiores conforme X_1 aumenta. Revertendo-se a direção de monotonicidade de $\uparrow$ para $\downarrow$, resulta-se na condição estocasticamente decrescente.

Note que, com este conceito, aproxima-se o conceito de dependência nas caudas, que possui semelhanças com o conceito de dependência com crescimento estocástico. O próximo conceito aproxima-se ainda mais da dependência nas caudas.

B.1.3. Dependência crescente na cauda direita e decrescente na cauda esquerda

Este conceito é construído a partir do conceito anterior, e envolve não só valores crescentes fixos de uma das variáveis aleatórias, mas toda a cauda além de um limiar crescente em uma das variáveis aleatórias.

Definição B.3: Seja $X = (X_1, X_2)$ um vetor aleatório bivariado com distribuição acumulada (f.d.a.) $F \in \mathcal{F}(F_1, F_2)$. A dependência de X_2 é crescente na cauda direita de X_1 se:

$$P(X_2 > x_2 \mid X_1 > x_1) = 1 - F_{2/1}(x_2 \mid x_1) \uparrow x_1 \forall x_2 \ . \qquad \text{[B-4]}$$

Similarmente, a dependência de X_2 é decrescente na cauda esquerda de X_1 se:

$$P(X_2 \leq x_2 \mid X_1 \leq x_1) = 1 - F_{2/1}(x_2 \mid x_1) \downarrow x_1 \forall x_2 . \qquad \text{[B-5]}$$

Como anteriormente, em [B-4], é mais provável que ocorram valores maiores de X_2 conforme X_1 encontra-se em valores mais elevados na sua distribuição, e em [B-5], é mais provável que ocorram valores menores de X_2 conforme X_1 assume valores menores em sua distribuição.

Veja que os conceitos de dependência B.1.2 e B.1.3 também são conceitos de dependência positiva. Em B.1.3, se for invertida a ordem da monotonicidade de $\uparrow$ para $\downarrow$, em [B-4] e de $\downarrow$ para $\uparrow$ em [B-5], obtém-se conceitos de dependência negativa.

B.1.4. Associação entre Variáveis Aleatórias

Este conceito exprime a existência de uma associação entre variáveis aleatórias quando existe uma covariância positiva entre elas.

Definição B.4: Seja $\mathbf{X} = (X_1, X_2)$ um vetor aleatório bivariado. X possui associação se a inequação:

$$E[g_1(\mathbf{X})g_2(\mathbf{X})] \geq E[g_1(\mathbf{X})]E[g_2(\mathbf{X})] \qquad \text{[B-6]}$$

é válida para toda função real g_1 e g_2 crescentes em cada um de seus componentes, e de forma que cada esperança em [B-6] exista.

B.1.5. Positividade Total de segunda ordem

Definição B.5: Uma função não negativa b em A^2, onde $A \subset \Re$, é totalmente positiva de segunda ordem se, para todo $x_1 < y_1$, $x_2 < y_2$, onde $x_i, y_i \in A$,

$$b(x_1, x_2)b(y_1, y_2) \geq b(x_1, y_2)b(x_2, y_1)\,. \qquad \text{[B-7]}$$

Suponha que se esteja trabalhando com uma função acumulada *F* de densidade *f*. As noções de dependência positiva que podem ser expressas através da positividade total de segunda ordem são as seguintes: i *f*, ii *F*, iii $\overline{F} = 1 - F$ possuem positividade total de segunda ordem. Ou seja, em (i) a condição de dependência positiva significa que para $x_1 < y_1$, $x_2 < y_2$, $f(x_1, x_2)f(y_1, y_2) \geq f(x_1, y_2)f(x_2, y_1)$ significa que é mais provável que sejam observados pares cujos componentes possuam valores alto-alto ou baixo-baixo do que valores alto-baixo ou baixo-alto.

Nota-se, então uma correlação alta entre a positividade total de segunda ordem e o conceito de concordância que será apresentado na seção B.3.

B.1.6. Relações entre os Conceitos de Dependência

Como pode-se notar, cada um dos conceitos apresentados são inter-relacionados, sendo que o conceito mais forte é o da dependência com positividade total de segunda ordem, pois sua existência implica que todas as outras propriedades/conceitos de dependência são válidos. As relações entre os conceitos de dependência são apresentadas no teorema B.1.

Teorema B.1:

(a) Positividade total de segunda ordem na densidade f $\Rightarrow$ Dependência com crescimento estocástico $\Rightarrow$ Dependência crescente na cauda direita e decrescente na cauda esquerda;

(b) Dependência crescente na cauda direita e decrescente na cauda esquerda $\Rightarrow$ Associação entre as variáveis aleatórias $\Rightarrow$ Dependência em Quadrante Positiva;

(c) Positividade total de segunda ordem na densidade f $\Rightarrow$ Positividade total de segunda ordem na distribuição acumulada F e sobrevivência $\overline{F} = 1 - F$.

(d) Positividade total de segunda ordem na densidade f $\Rightarrow$ Dependência decrescente na cauda esquerda e Positividade total de segunda ordem na função de sobrevivência $\overline{F} = 1 - F$ $\Rightarrow$ Dependência crescente na cauda direita.

A prova desse importante teorema pode ser consultada em Joe (1997, Cap 2, p. 26-27). Como pode-se observar, a positividade total de segunda ordem na densidade implica em positividade total de segunda ordem na distribuição acumulada. Como se sabe da definição de Função de Acoplamento (veja Nelsen, 2006, Cap 2.), as cópulas satisfazem a propriedade de positividade total de segunda ordem, o que as transformam em um ferramental poderoso para o estudo da dependência entre variáveis aleatórias.

O teorema B.2 mostra a flexibilidade das propriedades/conceitos de dependência com relação a transformações estritamente crescentes, às quais as Cópulas também atendem.

Teorema B.2: Os conceitos de dependência em Quadrante Positiva, Dependência com Crescimento Estocástico, Dependência Crescente nas Caudas e Positividade Total de segunda ordem são invariantes a transformações estritamente crescentes dos componentes do vetor aleatório.

Isso quer dizer que, por exemplo, se (X_1, X_2) possuem dependência em quadrante positiva, então $(a_1(X_1), a_2(X_2))$ também possuem, para transformações estritamente crescentes a_1 e a_2. A prova desse teorema é trivial e será omitida.

B.2. Estruturas de Dependência

Neste texto serão estudadas as estruturas de dependência simétricas e assimétricas por sua importância em aplicações financeiras e econômicas.

A simetria na dependência ocorre na maioria das distribuições elípticas multivariadas, como Normal e t-Student multivariadas. O fato de que, muitas vezes observa-se maior ocorrência de pares concordantes em eventos negativos (por exemplo, no estudo de contágios financeiros, existe uma dependência maior na cauda das perdas), a utilização das distribuições acima mencionadas pode ocasionar vícios consideráveis quando utilizadas em modelos probabilísticos.

A busca por distribuições multivariadas que possuam dependência assimétrica é importante pois garante que casos como esses sejam modelados de forma não viesada.

B.2.1. Dependência Perfeita Positiva e Negativa

Nesta seção será apresentada a definição de dependência perfeita, que pode-se entender como a mais forte dependência existente entre variáveis aleatórias. A dependência perfeita é caracterizada pela comonotonicidade entre as variáveis aleatórias.

Definição B.6: (Dependência Perfeita): As variáveis aleatórias X e Y são ditas positivamente dependentes perfeitamente se são comonotônicas (por exemplo, $X = Y$ quase certamente) e são negativamente dependentes perfeitamente se forem contramonotônicas (por exemplo, $X = -Y$ quase certamente).

Definição B.7: (Subconjunto Comonotônico): Um subconjunto $A \subseteq \Re^n$ é dito comonotônico se, para qualquer $(x_1,..., x_n)$ e $(y_1,..., y_n)$ em *A*, é válido que $x_i \leq y_i$ ou $y_i \leq x_i$, $\forall i = 1,2,..., n$.

Definição B.8: (Suporte Comonotônico): Seja $\mathbf{X} = (X_1,..., X_n)$ um vetor aleatório n-dimensional. Qualquer subconjunto $A \subseteq \Re^n$ será chamado de suporte de **X** se $P(\mathbf{X} \in A) = 1$ for verdadeira. O suporte A é comonotônico se, para qualquer $\mathbf{x}$ e $\mathbf{y} \in A$, ou $\mathbf{x} < \mathbf{y}$ ou $\mathbf{y} < \mathbf{x}$ é válido.

Definição B.9 (Vetor Aleatório Comonotônico): Um vetor aleatório $\mathbf{X} = (X_1,..., X_n)$ é dito comonotônico se seu suporte é comonotônico.

Pelas definições pode-se concluir que a comonotonicidade é uma estrutura de dependência muito forte. De fato, se $(x_1,..., x_n)$ e $(y_1,..., y_n)$ são elementos do suporte comonotônico de **X**, então são ordenados elemento à elemento. Comonotonicidade de um vetor **X** implica que, quanto maior for o valor de um componente X_j, maior será o valor de qualquer outro componente X_k. Comonotonicidade significa que nenhum X_j neutraliza qualquer outra componente X_k em qualquer forma ou grau.

O teorema B.3 mostra outras caracterizações equivalentes da comonotonicidade em vetores aleatórios.

Teorema B.3: O vetor aleatório $\mathbf{X} = (X_1,..., X_n)$ é comonotônico se e somente se uma das seguintes condições equivalentes é satisfeita:

(i) $(X_1,..., X_n)$ tem suporte comonotônico

(ii) Para todo $(x_1,..., x_n)$: $H(x_1,..., x_n) = \min\{F_{X_1}(x_1),..., F_{X_n}(x_n)\}$

(iii) $(X_1,..., X_n) =_d (F_{X_1}^{-1}(U),..., F_{X_n}^{-1}(U))$ para $U \sim U(0,1)$

(iv) existem funções não-decrescentes $\alpha_i : \Re \to \Re, i = 1,...,n$ e uma variável aleatória Z tal que $(X_1,..., X_n) =_d (\alpha_1(Z),...,\alpha_n(Z))$, onde $=_d$ indica igualdade em distribuição.

Prova: [(i) $\Rightarrow$ (ii)] Assuma que **X** possui suporte comonotônico **B**. Seja $(x_1,..., x_n) \in \Re^n$ e seja A_j definida por:

$$A_j = \{\mathbf{y} \in \mathbf{B} \mid y_j \leq x_j\}, \; j = 1,2,...,n .$$

Como **B** é comonotônico, existe um i de forma que $A_i = \bigcap_{j=1}^{n} A_j$. Assim,

$H(x_1,..., x_n) = P[\mathbf{X} \in \bigcap_{j=1}^{n} A_j] = P[\mathbf{X} \in A_i] = F_{X_i}(x_i)$. Por outro lado, já que $A_i \subset A_j$, $j = 1,2,...,n$, tem-se que $F_{X_i}(x_i) \leq F_{X_j}(x_j)$. Daí, $F_{X_i}(x_i) = \min\{F_{X_1}(x_1),..., F_{X_n}(x_n)\}$.

[(ii) $\Rightarrow$ (iii)] Assuma que $F_{X_i}(x_i) = \min\{F_{X_1}(x_1),..., F_{X_n}(x_n)\}$ para todo $\mathbf{x} = (x_1,..., x_n)$. Então,

$$P[F_{X_1}^{-1}(U) \leq x_1,..., F_{X_n}^{-1}(U) \leq x_n] = P[U \leq F_{X_1}(x_1),..., U \leq F_{X_n}(x_n)] =$$
$$P[U \leq \min_{j=1,...,n}\{F_{X_j}(x_j)\}] = \min_{j=1,...,n}\{F_{x_j}(x_j)\}]$$

[(iii) $\Rightarrow$ (iv)] Conclusão direta.

[(iv) $\Rightarrow$ (i)] Assuma que se pode encontrar uma variável aleatória Z com suporte **B** e funções não decrescentes a_i, $i = 1,2,...,n$ de forma que $\mathbf{X} =_d (a_1(Z), a_2(Z),..., a_n(Z))$.

O conjunto de possíveis eventos de **X** é $\{(a_1(Z), a_2(Z),..., a_n(Z)) \mid Z \in B\}$, que é comonotônico. Isso implica que **X** de fato é comonotônica.

■

B.2.2.Dependência Simétrica

No contexto da dependência, podem-se definir as estruturas de dependência simétrica e assimétrica através de características da estrutura conjunta das variáveis aleatórias em estudo, ou através de sua função de acoplamento. A definição B.10 introduz o conceito de simetria radial bivariada.

Definição B.10: $(X,Y) \sim H = C(F,G)$ são ditas radialmente simétricas em torno de (μ_x, μ_y) se a distribuição conjunta de $(X-\mu_x, Y-\mu_y)$ é a mesma que a distribuição de $(\mu_x - X, \mu_y - Y)$.

Isso significa que, se (X,Y) são radialmente simétricas, então:

$$\begin{aligned} &P[X-\mu_x<0, Y-\mu_y<0] = P[X<\mu_x, Y<\mu_y] = \\ &P[-X>-\mu_x, -Y>-\mu_y] = P[\mu_x - X>0, \mu_y - Y>0] \end{aligned} . \qquad \text{[B-8]}$$

Ou seja, a massa de probabilidade além de um ponto $A=(a,b)$ deve ser a mesma a mesma massa de probabilidade até seu ponto simétrico $A^{'}=(-a,-b)$

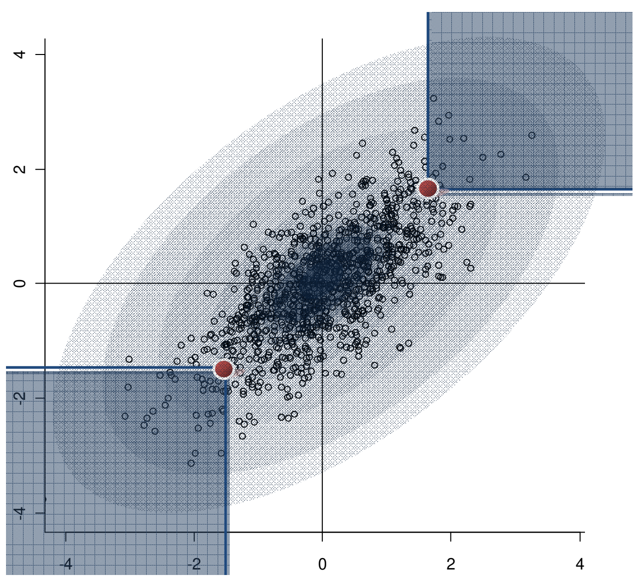

Figura B. 1. Uma distribuição conjunta simétrica padronizada possui a mesma massa de probabilidade em cada quadrante alternado.

A distribuição conjunta de duas variáveis aleatórias é radialmente simétrica se e somente se a estrutura de dependência entre elas é radialmente simétrica. Este conceito é apresentado mais formalmente no teorema B.4, que relaciona a simetria radial da distribuição conjunta com a de uma Cópula a ela referente.

Teorema B.4: Seja $(X,Y) \sim H = C(F,G)$, e seja *X e Y* simétricas em torno de μ_x e μ_y, respectivamente. Então (*X,Y*) são radialmente simétricas em torno de (μ_x, μ_y) se e somente se sua cópula, C, é radialmente simétrica.

A prova desse teorema pode ser encontrada em Nelsen (2006, Cap. 2.7).

Isso significa que se temos duas distribuições simétricas e desejamos construir um modelo que possua estrutura de dependência simétrica, basta selecionar uma função de acoplamento que seja radialmente simétrica.

B.2.3. Dependência Assimétrica

No mercado financeiro muitas vezes não é válida a premissa de simetria na dependência (acontece que em mercados muito voláteis as perdas costumam ser observadas em diversos ativos ao mesmo tempo, o que, excluindo-se alguns fatos isolados e fundamentalmente explicados, não ocorrem com os ganhos). Essa característica da dependência, de ser maior em uma das caudas e menor em uma outra, recebe o nome de dependência assimétrica, e é definida de forma inversa à dependência simétrica.

Definição B.11: Uma estrutura de dependência é assimétrica se a função de acoplamento que a caracteriza não atender ao conceito de simetria do teorema B.4.

A figura B.2 exemplifica estruturas assimétricas de dependência caracterizadas por diversas funções de acoplamento, onde se pode citar a Gumbel, t-Student assimétrica, Joe e Clayton.

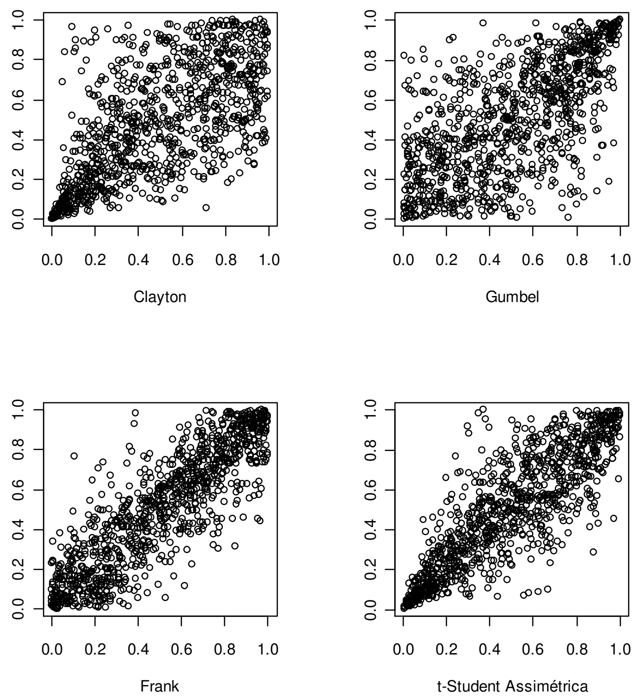

Figura B. 2. Estruturas de dependência assimétricas Gumbel, Clayton, Frank e t-Student Assimétrica

A dependência assimétrica que mais figura no mercado financeiro é a dependência caracterizada pelas cópulas Clayton, Gumbel ou t-Student assimétrica, que mostram dependência mais forte em uma das direções (geralmente perdas) do que na outra.

B.3. Medidas de Dependência e Concordância

As medidas de dependência e concordância são ferramentas que podem ser utilizadas para quantificar a dependência entre variáveis aleatórias, de forma que se possa conhecer o perfil da dependência e selecionar uma correta função conjunta ou de acoplamento que atenda adequadamente ao perfil de dependência dos objetos em estudo.

Em geral as medidas de concordância são mais fortes e exprimem a dependência entre variáveis aleatórias com melhor qualidade, pois atendem a um número maior de propriedades desejáveis que essas medidas devem possuir.

B.3.1. Medidas de Dependência

A definição B.12 exibe propriedades desejáveis para medidas de dependência.

Definição B.12: Seja $T:\Re \to \Re$ estritamente monótona no intervalo de variação de X. Seja $\delta(.,.)$ uma medida de dependência, então:

Simetria $(\delta(X,Y)=\delta(Y,X))$;

$-1 \le \delta(X,Y) \le 1$

$\delta(X,Y)=1 \Leftrightarrow X,Y$ comonotônicas

$\delta(X,Y)=-1 \Leftrightarrow X,Y$ contramonotônica

$\delta(T(X),Y)=\delta(X,Y)$, com T crescente

$\delta(T(X),Y)=-\delta(X,Y)$, com T decrescente

$\delta(X,Y)=0 \Leftrightarrow X,Y$ Independentes.

Embora o Coeficiente de Correlação Linear de Pearson falhe em algumas das propriedades acima, ele é uma medida de dependência.

Nas aplicações em finanças, na maioria das vezes não são válidas suposições como normalidade ou linearidade da dependência, o que requer que ferramentas mais poderosas sejam utilizadas para mensurar a dependência, para viabilizar a seleção de um modelo válido. Nesse sentido são utilizadas medidas de concordância, que auxiliam na seleção dos modelos mais adequados.

B.3.2. Medidas de Concordância

A concordância entre variáveis aleatórias é a característica apresentada por uma variável que assume valores mais altos quando uma segunda variável assume valores altos. A concordância é definida com mais rigor na definição B.13.

Definição B.13. (Concordância): Sejam $(x, y)^T$ e $(x', y')^T$ duas observações do vetor $(X, Y)^T$, onde X e Y são variáveis aleatórias contínuas. Então $(x, y)^T$ e $(x', y')^T$ são ditos serem concordantes se $(x - x')(y - y') > 0$ e discordantes caso $(x - x')(y - y') < 0$. Ou seja, o par será concordante se $x < x'$ e $y < y'$ ou $x > x'$ e $y > y'$. O par será discordante se $x < x'$ e $y > y'$ ou $x > x'$ e $y < y'$.

Definição B.14: 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 par 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 C' são acoplamentos tais que $C \prec C'$, então $\kappa_C \leq \kappa_{C'}$;

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 $\lim_{n \to \infty} \kappa_{C'} = \kappa_C$.

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

B.3.3. O Tau de Kendall

Definição B.15: O tau de Kendall para o vetor aleatório $(X,Y)^T$ é definido como:

$\tau(X,Y) = P\{(X-X')(Y-Y')>0\} - P\{(X-X')(Y-Y')<0\}$, onde $(X,Y)^T$ e $(X',Y')^T$ são vetores aleatórios i.i.d.

Teorema B.6: Sejam (X,Y) e (X',Y') vetores aleatórios independentes formados por variáveis aleatórias contínuas, com funções de distribuição conjunta H e H' respectivamente, e com funções marginais F_1 (de X e X') e F_2 (de Y e Y'). Sejam C e C' as funções de acoplamento de (X,Y) e (X',Y') respectivamente, de modo que $H(x,y) = C(F_1(x), F_2(y))$ e $H'(x,y) = C'(F_1(x), F_2(y))$, $U = F_1(X)$ e $V = F_2(Y)$ Seja Q a medida de concordância definida por:

$$Q(X,Y) = P\{(X-X')(Y-Y')>0\} - P\{(X-X')(Y-Y')<0\}, \qquad \text{[B-9]}$$

então,

$$Q(X,Y) = Q(C,C') = 4\iint_{[0,1]^2} C'(u,v)\,dC(u,v) - 1 = 4E[C(U,V)] - 1\,. \qquad \text{[B-10]}$$

A prova desse teorema pode ser encontrada em Nelsen (2006, Cap. 5, p. 159).

B.3.4. O Rho de Spearman

Definição B.16: O Rho de Spearman para o vetor aleatório $(X,Y)^T$ é definido como:

$\rho_S(X,Y) = 3(P\{(X-X')(Y-Y'')>0\} - P\{(X-X')(Y-Y'')<0\})$, onde $(X,Y)^T, (X',Y')^T$ e $(X'',Y'')^T$ são cópias independentes.

Se forem utilizados os pares (*X,Y*) e (*X′,Y″*), a partir do teorema 2.6, considerando que a Cópula de (*X,Y*) é C e a Cópula de (*X′,Y″*) é a cópula independente $\prod$, pelo fato de que (*X′ e Y″*) são independentes, segue diretamente:

Sejam (X,Y), (X',Y') e (X'',Y'') vetores aleatórios independentes formados por variáveis aleatórias contínuas, todas com funções de distribuição conjunta H, funções marginais F_1 e F_2 e Cópula C. Então,

$$\rho_S(X,Y) = 3Q(C,\prod) = 12\iint_{[0,1]^2} uv dC(u,v) - 3 = 12\iint_{[0,1]^2} C(u,v) du dv - 3, \qquad \text{[B-11]}$$

onde $\prod$ é a cópula independente.

B.3.5. 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 $(X,Y)^T$.

Definição B.17: 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 (inferior) de $(X,Y)^T$ é definido como:

$\lim_{\alpha\to 1^-} P\{Y > F_2^{-1}(\alpha) / X > F_1^{-1}(\alpha)\} = \lambda_{\sup}$, ($\lim_{\alpha\to 0^+} P\{Y \le F_2^{-1}(\alpha) / X \le F_1^{-1}(\alpha)\} = \lambda_{\inf}$), $\lambda_{\sup} \in [0,1]$ ($\lambda_{\inf} \in [0,1]$), dado que o limite exista. Se $\lambda_{\sup} \in (0,1]$ ($\lambda_{\inf} \in (0,1]$), então X e Y são chamadas de assintoticamente dependentes na cauda superior (inferior). Se $\lambda_{\sup} = 0$ ($\lambda_{\inf} = 0$) X e Y são chamadas de assintoticamente independentes na cauda superior (inferior). Caso o limite não exista, a dependência na cauda não é definida.

pode-se reescrever $P\{Y > F_2^{-1}(\alpha) / X > F_1^{-1}(\alpha)\}$ como:

$$\frac{1 - P\{X \le F_1^{-1}(\alpha)\} - P\{Y \le F_2^{-1}(\alpha)\} + P\{X \le F_1^{-1}(\alpha), Y \le F_2^{-1}(\alpha)\}}{1 - P\{X \le F_1^{-1}(\alpha)\}}, \qquad \text{[B-12]}$$

e $P\{Y \le F_2^{-1}(\alpha) / X \le F_1^{-1}(\alpha)\}$ como:

$$\frac{P\{X \le F_1^{-1}(\alpha), Y \le F_2^{-1}(\alpha)\}}{P\{X \le F_1^{-1}(\alpha)\}}, \qquad \text{[B-13]}$$

daí, supondo que X e Y possuam distribuição conjunta $H(X,Y) = C(F_1(X), F_2(Y)) = C(U,V)$, onde C é uma função de acoplamento, aplicando as marginais F_1 e F_2 nos elementos dentro das probabilidade nas equações [B-12] e [B-13] (já que $F_1(X) = U$ e $F_2(Y) = V$) e aplicando os limites descritos na definição B.17 obtém-se definições equivalentes para cópulas:

(1) Se C é o acoplamento bivariado de *X* e *Y* tal que o limite

$$\lim_{\alpha \to 1^-} \frac{1 - 2\alpha - C(\alpha,\alpha)}{1-\alpha} = \lambda_{\sup} \qquad \text{[B-14]}$$

exista, então C tem dependência na cauda superior se $\lambda_{\sup} \in (0,1]$, e é independente na cauda superior se $\lambda_{\sup} = 0$.

(2) Se C é o acoplamento bivariado de *X* e *Y* tal que o limite

$$\lim_{\alpha \to 0^+} \frac{C(\alpha,\alpha)}{\alpha} = \lambda_{\inf} \qquad \text{[B-15]}$$

exista, então C tem dependência na cauda inferior se $\lambda_{\inf} \in (0,1]$, e é independente na cauda inferior se $\lambda_{\inf} = 0$.

Pelo fato de que a aplicação de distribuições conjuntas, sua seleção e estimação sobre dados reais, e mesmo a visualização do perfil de dependência que uma determinada distribuição acarreta, não são fáceis ou diretas, deve-se procurar uma distribuição conjunta cujas marginais atendam ao perfil distribucional de cada variável aleatória, enquanto que a dependência seja totalmente capturada pela conjunta.

Nesse sentido, busca-se solução para que se possam capturar os efeitos da variabilidade marginal separadamente ao perfil de dependência, que pode ser feito pelo uso de funções de acoplamento, sendo o perfil da dependência mais facilmente capturado pelas medidas de concordância, para possibilitar uma melhor seleção da função de acoplamento.

B.3.6. Outras Medidas

Outras medidas de dependência ou concordância presentes na literatura são Latif e Morettin (2007) ou Latif (2008) apresentando medidas de dependência locais para séries temporais, Goodman e Johnson (2003) gerando medidas de dependência multivariadas baseadas em séries ortonormais e distância de Kullback-Leibler, Schmid e Schmidt (2006), mostram extensões multivariadas do rho de Spearman dentre outras técnicas não paramétricas, e Taylor (2006) que revê medidas de concordância apresentadas em Scarsini (1984) para pares ordenados de variáveis aleatórias contínuas. Anjos e Kolev (2005) representam cópulas bivariadas através de medidas locais de dependência.

Apêndice C

Modelos ARMA-GARCH

Suponha uma série de retornos X_t. Essa série pode ser modelada através de um nível μ_t mais uma perturbação ε_t. Podemos considerar que existe uma relação aditiva entre esses componentes, e que a perturbação é o produto de dois processos independentes. Dessa forma, pode-se considerar que:

$$X_t = \mu_t + \varepsilon_t \ , \tag{C-1}$$

$$\varepsilon_t = \sigma_t \eta_t \, . \tag{C-2}$$

Sob as suposições 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 ruído branco com média zero e variância 1, independente do passado de $\{\varepsilon_t\}$.

No caso mais simples, $\mu_t = c$ (constante). Entretanto, o nível da série será modelado através de modelos ARMA(p,q).

C.1. Modelo GARCH(p,q)

O modelo GARCH (Generalized Autoregressive Conditional Heteroskedasticity) é dado por (Bollerslev 1986):

$$\varepsilon_t = \sigma_t \eta_t ,$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \ . \qquad \text{[C-3]}$$

Onde $\{\eta_t\}$ é um processo ruído branco com média zero e variância 1. Em geral, supõe-se que $\eta_t \sim N(0,1)$ ou $\eta_t \sim t_v$ (t-Student padronizada). As restrições que garantem que a variância condicional seja positiva são:

$\alpha_0 > 0$, $\alpha_i \geq 0$ para $i = 1,..., p$, $\beta_j \geq 0$ para $j = 1,..., q$, e $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_j) < 1$.

Para esse modelo, a variância não condicional é dada por:

$$Var(\varepsilon_t) = \frac{\alpha_0}{1 - (\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_j)} . \qquad \text{[C-4]}$$

Que é finita quando $\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_j < 1$. Têm-se também que:

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + ... + (\alpha_p + \beta_p)\varepsilon_{t-p}^2 + \omega_t - \beta_1 \omega_{t-1} - ... - \beta_q \omega_{t-q} \qquad \text{[C-5]}$$

Onde $\omega_t = \varepsilon_t - \sigma_t$ é um ruído branco. Dessa forma têm-se que $\varepsilon_t^2 \sim ARMA(p,q)$. No caso do modelo GARCH(1,1), a autocorrelação desse processo ARMA(1,1) tem decaimento segundo o valor $\alpha_1 + \beta_1$, que é chamado de persistência da série temporal.

C.2. Correlogramas das Séries Energéticas

C.2.1. Gasolina Convencional

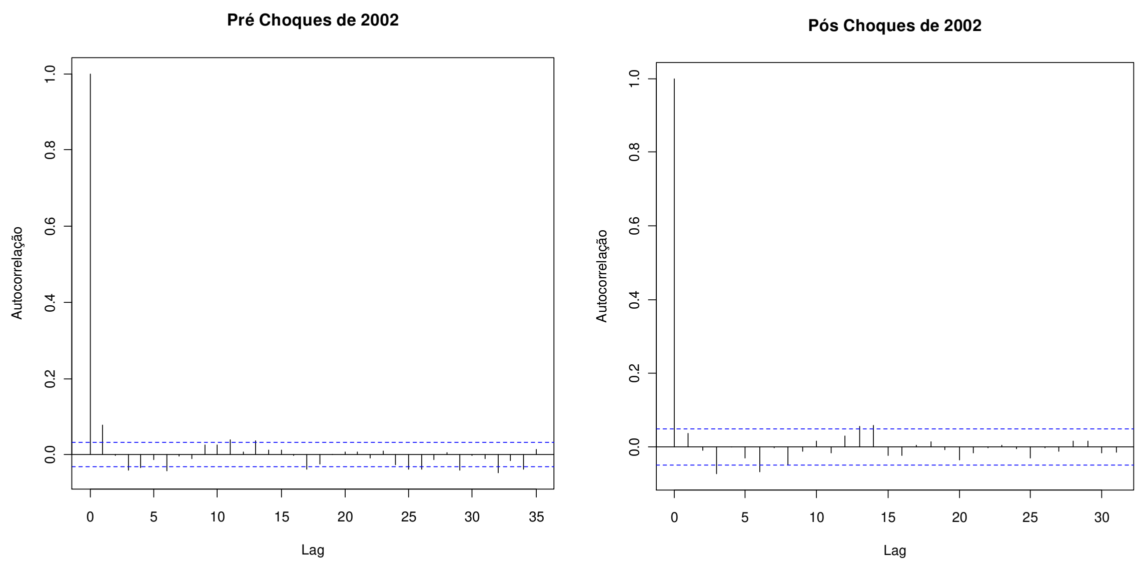

Figura C. 1. Autocorrelação da série de log-retornos da Gasolina Convencional períodos pré e pós choques de 2002.

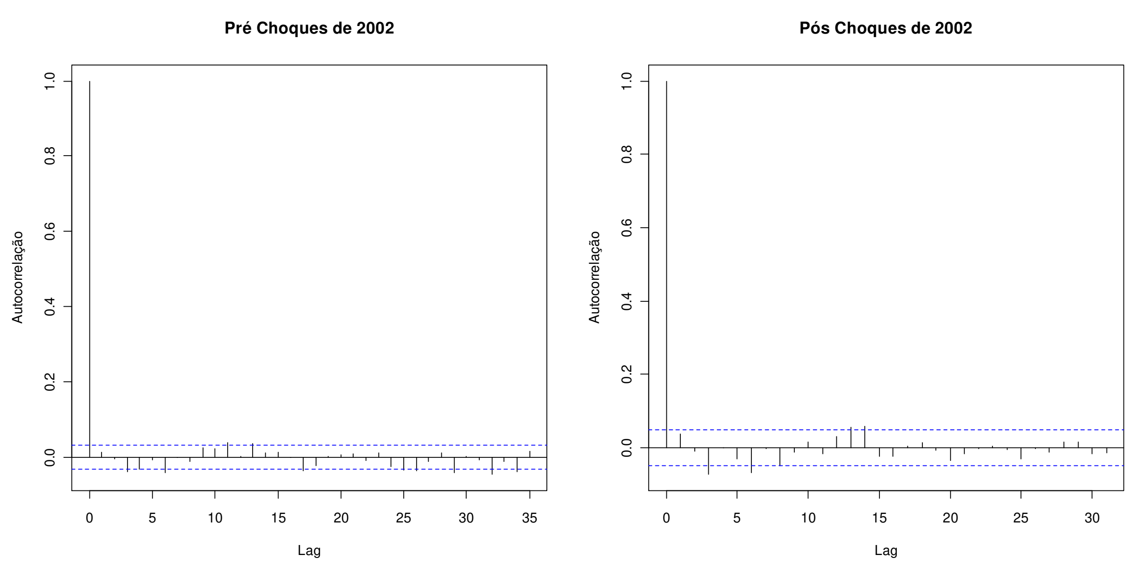

Figura C. 2. Autocorrelação dos resíduos do ajuste dos modelos ARMA(1,0) período 1 e ARMA(0,0) período 2 para a Gasolina Convencional.

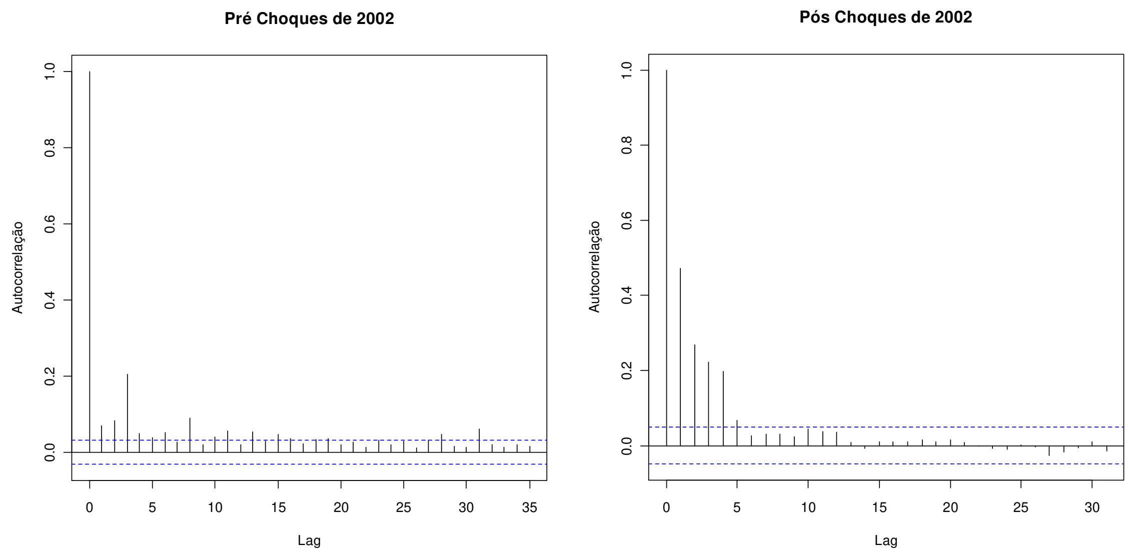

Figura C. 3. Autocorrelação dos resíduos quadráticos do modelo ARMA (1,0) (Período 1) e ARMA(0,0) (Período 2), utilizado nos log- retornos da Gasolina Convencional.

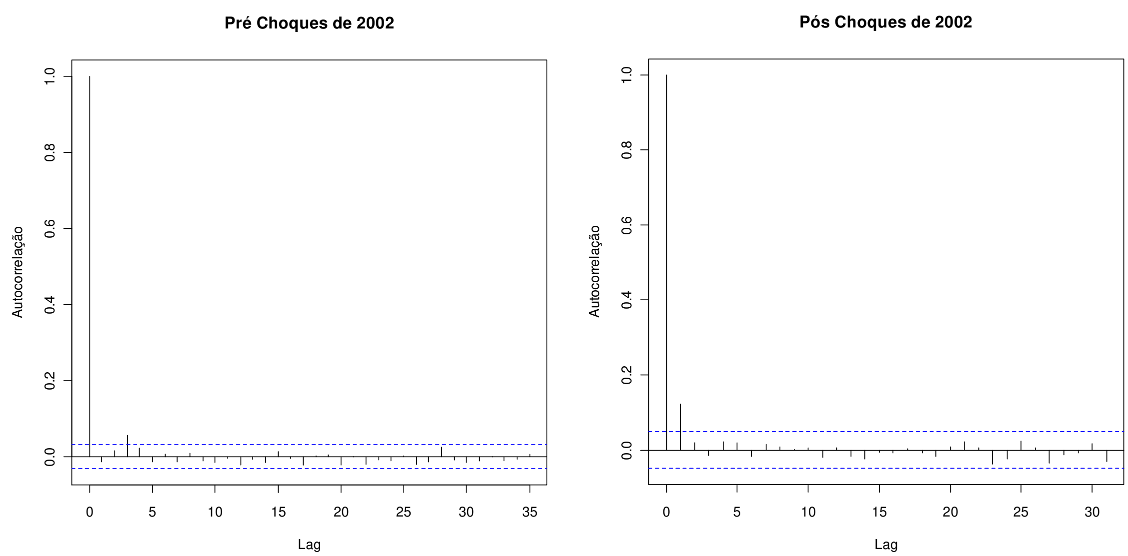

Figura C. 4. Autocorrelação dos resíduos quadráticos do modelo ARMA (1,0)-GARCH(1,1) (Período 1) e ARMA(0,0)-GARCH(1,1) (Período 2), utilizado nos log- retornos da Gasolina Convencional

C.2.2. Índice de Óleo AMEX

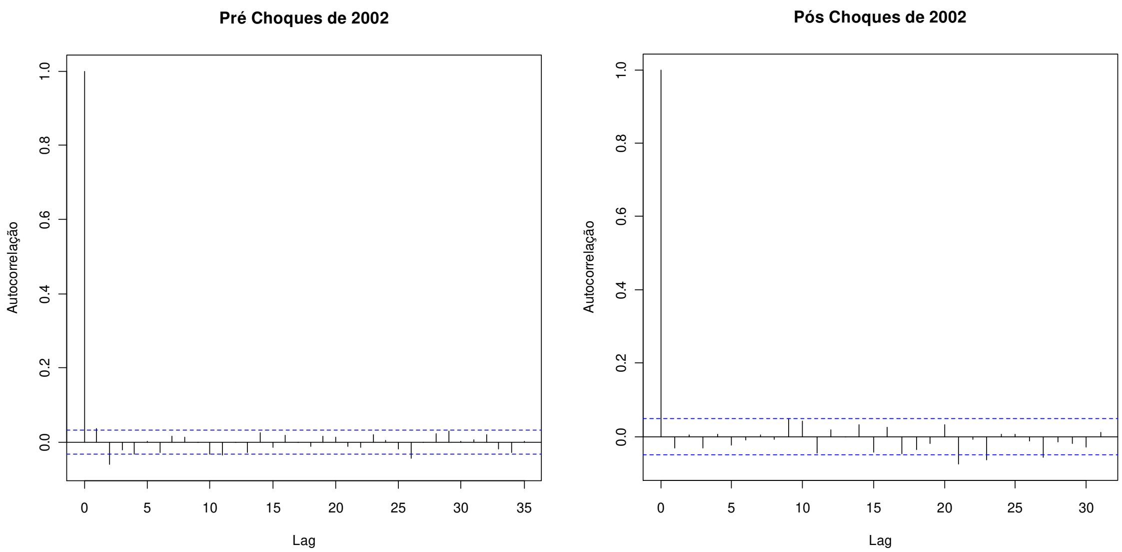

Figura C. 5. Autocorrelação da série de log-retornos do Índice de Óleo AMEX para os períodos pré e pós choques de 2002.

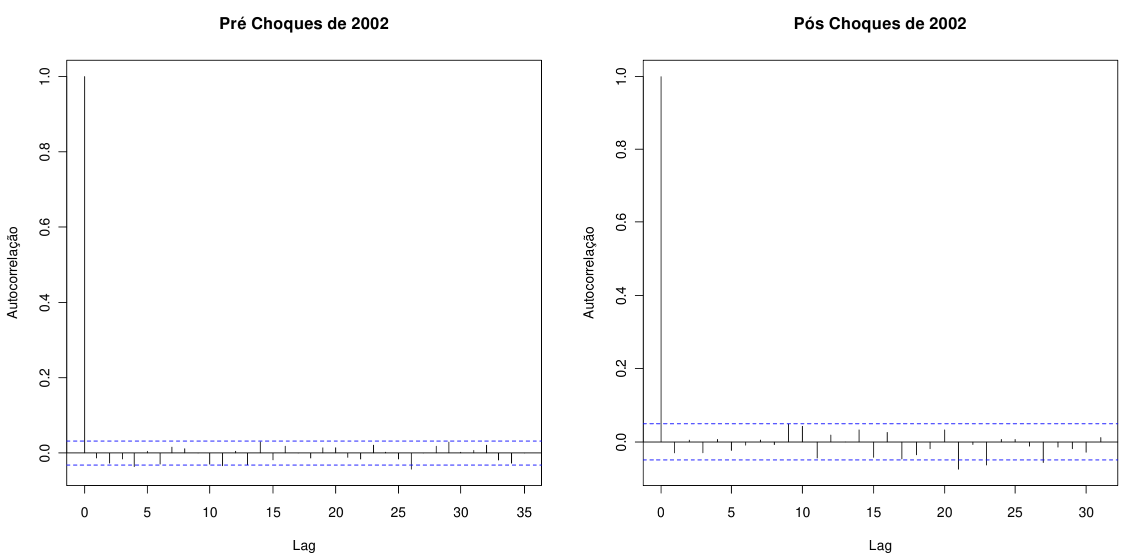

Figura C. 6. Autocorrelação dos resíduos do ajuste dos modelos ARMA(2,0) período 1 e ARMA(0,0) período 2 para o Índice de Óleo AMEX.

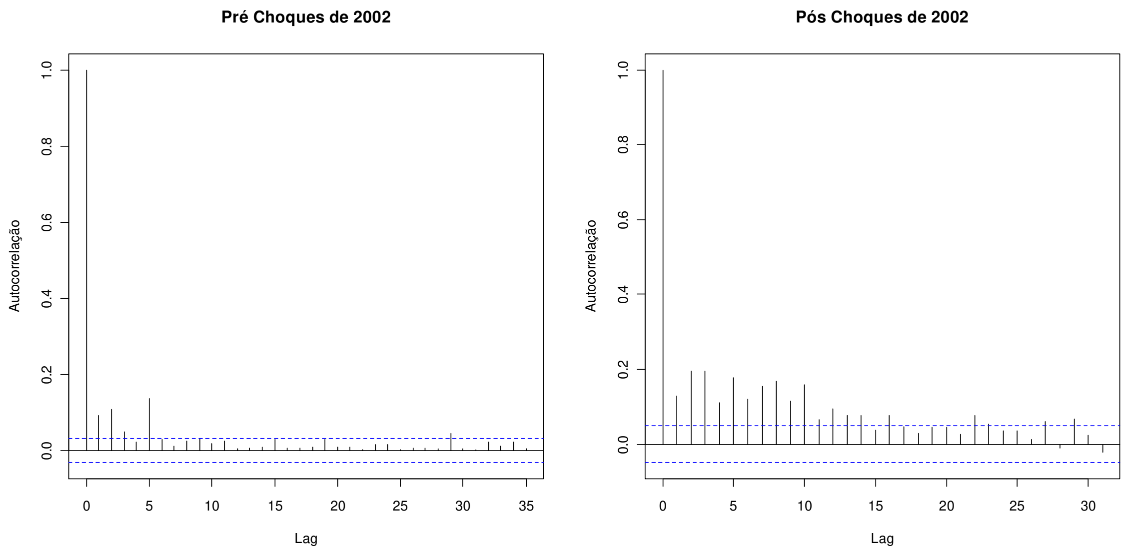

Figura C. 7. Autocorrelação dos resíduos quadráticos do modelo ARMA (2,0) (Período 1) e ARMA(0,0) (Período 2), utilizado nos log- retornos do Índice de Óleo AMEX.

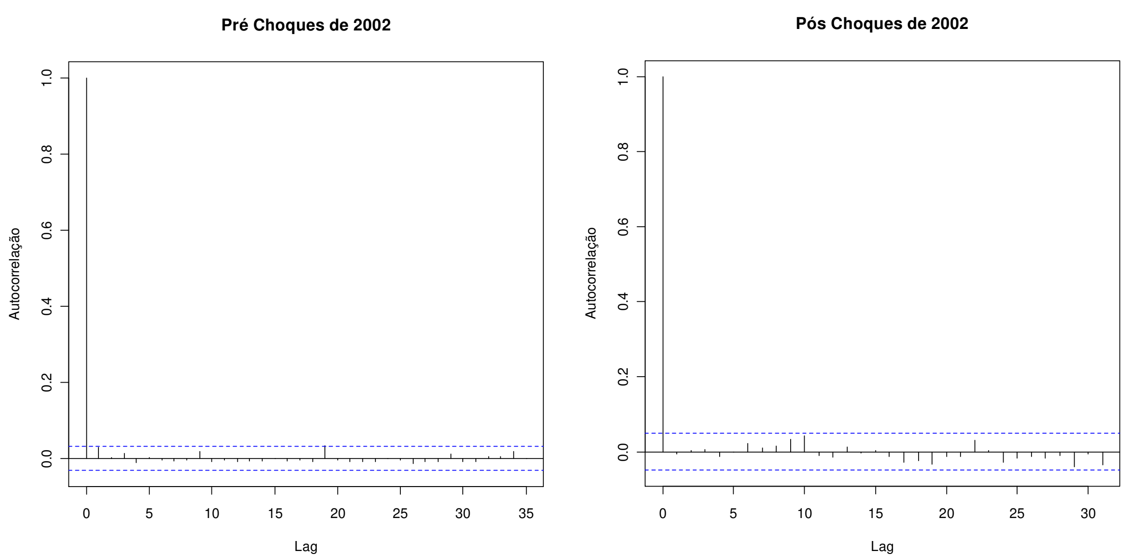

Figura C. 8. Autocorrelação dos resíduos quadráticos do modelo ARMA (2,0)-GARCH(1,1) (Período 1) e ARMA(0,0)-GARCH(1,1) (Período 2), utilizado nos log- retornos do Índice de Óleo AMEX.

Apêndice D

D – Bondade de Ajuste

Nesta seção serão apresentados brevemente os testes de bondade de ajuste utilizados para comparar as funções de acoplamento. Os testes são baseados em: a) na chamada Cópula Empírica (Deheuvels, 1979), b) na Transformada de Kendall (Barbe *et al,*1996), e c) no VaR, Valor em Risco (veja por exemplo Morgan, 1996).

D.1. Teste baseado na Cópula Empírica

O teste da cópula empírica aplica uma estatística de "distancia" entre a cópula ajustada e a empírica. Mais especificamente, o teste considera um acoplamento C desconhecido mas pertencente à classe $C_0 = \{C_\theta : \theta \in \Theta\}$, onde Θ é um espaço aberto dos número Reais de dimensão p. A hipótese nula é definida por $H_0 : C \in C_0$. A estatística do teste da cópula empírica é baseada no processo $\ddot{C}_n = \sqrt{n}(C_n - C_{\theta_n})$, onde C_n é a cópula empírica, e é dada por:

$$S_n = \int_{[0,1]^d} \ddot{C}_n(\mathbf{u})^2 dC_n(\mathbf{u}). \qquad [4\text{-}1]$$

Valores altos dessa estatística resultam na rejeição da hipótese nula. A distribuição assintótica dessa estatística não pode ser tabulada na prática, sendo que um algoritmo de bootstrap paramétrico será utilizado para cálculo do p-valor. Esse algoritmo foi validado por Genest e Rémillard (2008) e está descrito em Genest *et* al (2007).

D.2. Teste baseado na Transformada de Kendall

Este teste utiliza a transformada integral de probabilidade $\mathbf{X} \mapsto V = H(\mathbf{X}) = C(U_1, U_2, ..., U_n)$ para gerar um teste de $H_0 : C \in C_0$, onde $U_i = F_i(X_i), i = 1,...,d$, e a distribuição conjunta de $\mathbf{U} = (U_1,...,U_n)$ é C. Essa transformação é chamada de Transformada de Kendall.

Se $V_i = C_n(U_i)$, então seja $K_n(v) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(V_i \leq v)$, $v \in [0,1]$. A hipótese nula que será testada é a de que $H_0 : K \in K_0 = \{K_\theta : \theta \in \Theta\}$. Então $\ddot{K}_n = \sqrt{n}(K_n - K_{\theta_n})$, e a estatística do teste é dada por:

$$S_n = \int_{[0,1]^d} \ddot{K}_n(v)^2 dK_{\theta_m}(v). \tag{4-1}$$

Novamente, valores altos dessa estatística resultam na rejeição da hipótese nula. Um procedimento paramétrico, descrito em Genest *et al* (2007) será utilizado para o cálculo do p-valor.

D.3. Estatística Baseada no Valor em Risco

O Valor em Risco é uma medida padrão usada por analistas para quantificar o risco de mercado de um ativo ou portfólio. Quando o retorno de cada componente de um portfólio é pequeno, o retorno total do portfólio é aproximadamente dado por $r_t \cong \sum_{j=1}^{n} x_{j,t} r_{j,t}$, onde $x_{j,t}$, j=1,...,n são os pesos do j-ésimo ativo no tempo t. O VaR (Valor em Risco de nível $100\alpha\%$ para um momento t da série em estudo é definido por $P[r_t \leq VaR100\alpha\%_t \mid r_{j,t-1}, j = 1,....,n; t = 1,2,...] = \alpha$.

Depois de estimado um modelo completo para um conjunto de ativos, o $\text{VaR}100\alpha\%_t$ pode ser estimado por simulação, se for possível amostrar da distribuição conjunta das inovações, o que é facilmente realizado através de uma função de acoplamento. Maiores detalhes sobre a metodologia podem ser obtidos em (Hotta *et al, 2008*).

A estatística VaR empregada neste texto é calculada através do seguinte algoritmo:

1. Seleciona-se uma janela de estimação, que é um período inicial para a estimação. O VaR nesse período não será estimado, por exemplo 1000 observações.
2. Seleciona-se o período da estimação, ou seja, considera-se que os modelos ARMA-GARCH-Cópulas não se alteram para um determinado número de dias. O VaR calculado será comparado com as observações desse período (por exemplo 50 dias), que é imediatamente subseqüente à janela.
3. Estima-se o modelo ARMA-GARCH-Cópula para a janela e executa-se uma simulação de carteira a partir desse modelo.
4. Calculam-se limiares positivos e negativos da carteira que são comparado com a carteira real no período de avaliação (50 dias).
5. Verifica-se o número dessas observações que ultrapassaram o VaR.
6. A janela move-se em 50 (período) observações e retorna-se para 2, mantendo fixo o período.

Se o modelo ARMA-GARCH-Cópula foi bem ajustado, espera-se que o percentual, em todo o período, de observações que ultrapassam o VaR de nível α seja próximo de α.

Referências Bibliográficas

Abramowitz, M., Stegun, I., A., *"Handbook of Mathematical Functions*", Dover Publications, Nova Iorque, 1965

Akaike, H, "*Information theory and an extension of the maximum likelihood principle*" Em Second International Symposium on Information Theory, Budapeste: Akademiai Kiado, 267-281, 1973

Anjos, U., U., Kolev, N., V., "*Representation of bivariate copulas via local measure of dependence"* Technical report, IME-USP, 2005

Arellano-Valle, R., Genton, M., G., "*On Fundamental Skew Distributions*", Journal of Multivariate Analysis, 96, 93-116, 2005

Asmussen, S., Glynn, P. W., "*Stochastic Simulation: Algorithms and Analysis*", Springer, Nova Iorque, 2007

Azzalini, A., "*A class of distribution which includes the normal ones*", Scandinavian Journal of Statistics, 12, 171-178, 1985

Azzalini, A., Capitanio, A., "*Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution*", Journal of Royal Statistic Society, série B, volume 65, 367-389, 2003

Bamberger, R, "*The strategic petroleum reserve: history, perspectives, and issues*", CRS Relatório para o Congresso dos EUA, 2008

Barbe, P., Genest, C., Ghoudi ,K., Rémillard, B., "*On kendall's process*", Journal of Multivariate Analysis 58, 197-229, 1996

Bicalho, R., G., Iootty, M., Bomtempo, J., V., Almeida, E., F., Queiroz, H., "*Economia de Energia Fundamentos Econômicos, Evolução Histórica e Organização Industrial*", Editora Campus, 2003

Biglova, A., Kanamura, T., Rachev, S., Stovanov, S., "*Modeling, Risk Assessment and Portfolio Optimization of Energy Futures*", University of Karlsruhe, Alemanha, *J-POWER, Japan e finAnalytica Inc*, 2007

Bollerslev, T., "*Generalized autoregressive conditional heteroskedasticity*", Journal of Econometrics, 31, 307-327, 1986

Borenstein, S., Cameron, A., C., Gilbert, R., "*Do gasoline prices respond asymmetrically to crude oil price changes?*", Working Paper No. 4138, National Bureau of Economic Research, Cambridge, Agosto 1992

Bouyé, E., **Durrleman, V., Nikeghbali, A., Riboulet, G., Roncalli, T.,** "*Copulas for finance: a reading guide and some applications*", Groupe de Recherche Opérationnelle Crédit Lyonnais, Paris, 2000

Brown, S., P., A., Virmani, R., "*What's driving gasoline prices?*", Insights from the Federal Reserve Bank of Dallas, Vol. 2 No. 7, Outubro de 2007

Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C., "*A limited memory algorithm for bound constrained optimization.*" SIAM J. Scientific Computing, 16, 1190–1208, 1995

Caillault, C. e Guégan, D. "*Empirical estimation of tail dependence using copulas: application to asian markets*." Note de Recherche *IDHE-MORA* n. 5, 2003

Chapentier, A., "*Tail distribution and dependence measures*", ENSAE/CREST, França, 2003

Deheuvels, P., "*La fonction de dépendance empirique es sés propriétés – um test non paramétrique d'indépendence*." Academic Royale de Belgique – Bulletin de la classe dês sciences – 5º. Série, 65:274:292, 1979

Demarta, S., McNeil, A., J., "*The t Copula and related copulas*", International Statistical Review, Vol. 73, Número 1, 111-129, 2005

Embrechts, P., McNeil, A. & Straumann, D, "*Correlation and dependency in risk management: properties and pitfalls*", em Risk Management*:* Value at Risk and Beyond, Cambridge University Press, 176–223, 2002

Engdahl, W., *"Iraq and the problem of peak oil"*, Current Concerns Journal, edição de 26 de janeiro de 2004

Ferrarese, C., "*A comparative analysis of correlation skew modeling techniques for CDO index tranches*", Master Thesis, King's College London, MPRA Paper No. 1668, 2006

Fletcher, R. and Reeves, C. M., "*Function minimization by conjugate gradients.",* Computer Journal **7**, 148–154, 1964

Frahm, G., Junker, M., Schmidt, R., "*Estimating the tail-dependence coefficient: properties and pitfalls"*, Insurance: Mathematics and Economics 37, 80-100, 2005

Genest, C., **Rivest, L.**, "*On the multivariate probability integral transformation*", Statistics & Probability Letters 53, 391-399, 2001

Genest, C., Rémillard, B., Quessy, J., "*Goodness-of-fit procedures for copula models based of the probability integral transformation*", Scandinavian Journal of Statistics, vol 33, 337-366, 2006

Genest, C., Rémillard, B., Beaudoin, D., "*Goodness-of-fit tests for copulas: a review and a power study*" Insurance: Mathematics and Economics, doi:10.1016/ j.insmatheco.2007.10.005 , 2007

Genest, C., Rémillard, B., "*Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models*", Annales de l'Institut Henri Poincaré – Probabilités et Statistiques, Volume 44, Number 6, 1096-1127, 2008

Goodman, I., N., Johnson, D., H., "*New multivariate dependence measures and applications to neural ensembles",*ECE Department, Rice University, Houston, TX 2003

Grégoire, V., Genest, C., Gendron, M., "*Using copulas to model price dependence in energy markets", Energy Risk*, energyrisk.com, Março de 2008

Hotta, L., K., Lucas, E., C., Palaro, H., P., "*Estimation of VaR using copula and extreme value theory*", Multinational Finance Journal 12 (3/4), 205-218, 2008

Hu, W., Kercheval, A., N., "*The skewed t distribution for portfolio credit risk*", Florida State University e Bell Trading, 2005

Hu, W., "*Calibration of multivariate generalized hyperbolic distributions using the EM algorithm, with applications in risk management, portfolio optimization and portfolio credit risk", PHd Dissertation, Florida State University, College of Arts and Sciences*, 2005

Jacinto, C., M., C., Andrade, D., F., Barbeta, P., A., de Freitas Filho, P., J., Zunino, N., A., M., "*Using copulas in risk analysis*" Proceedings of the 2006 Winter Simulation Conference, 2006

Joe, H., "*Multivariate Models and Dependence Concepts*" Londres, Chapman & Hall, 1997

Jogdeo, K., "*Concepts of Dependence*", Encyclopedia of Statistical Sciences, Vol. 2, Wiley, Nova Iorque, pp. 324-334. 1982

Kole, E., Koedijk, K., Verbeek, M., "*Testing copulas to model financial dependence*", Dept. of Financial Management, RSM Erasmus University, Rotterdam, Holanda, 2005

Latif, S., A., "*Medidas de Dependência Local para Séries Temporais*". Tese de Doutorado, IME-USP, Brasil. 2008

Latif, S., A., Morettin, P., A., *"A Local Dependence Measure for Time Series."* Third Brazilian Conference on Statistical Modelling in Insurance and Finance, 2007, Maresias.

Proceedings of the Third Brazilian Cnference on Statistical Modelling in Insurance and Finance,. 304-307 2007

Lindskog, F., "*Modelling dependence with copulas*", dissertação de mestrado, Swiss Federal Institute of Technology, Zurich, 2000

Lopez, J., A., "*Methods for evaluating Value-at-Risk estimates*", FRBNY Economic Policy Review, Outubro 1998

Majoras, D., Swindle, O., Leary, T., Harbour, P., Leibowitz, J. "*Gasoline price changes: the dynamic of supply, demand, and competition*", Federal Trade Comission*,* EUA, 2005

Malevergne, Y., Sornette, D., "*Investigating extreme dependence: concepts and tools*", Université de Nice-Sophia Antipolis and University of California, 2002

McNeil, A., J., Frey, R., Embrechts, P., "*Quantitative Risk Management: Concepts, Techniques and Tools*", Princeton University Press, Reino Unido, 2005

Morgan, J., P., "*RiskMetrics Technical Document*", 4ª Edição, J. P. Morgan*,* Nova Iorque 1996

Nadarajah, S., Dey, D., K., "*Multitude of multivariate t-distributions*", Statistics, Vol 39, No. 2, 149-181, Abril 2005

Nelder, J. A. and Mead, R., "*A simplex algorithm for function minimization*" Computer Journal **7**, 308–313, 1965

Nelsen, R., "*An introduction to Copulas.*" 2ª Edição*,* Springer, Nova Iorque, 2006

Palaro, H. P., "*Aplicação de acoplamento no cálculo do valor em risco*", dissertação de mestrado, IMECC-UNICAMP, 2004.

Parra, F., "*Oil Politics - A modern History of Petroleum*", I.B. Tauris, Nova Iorque, 2004

Patton, A., L., "*On the out-of-sample importance of skewness and asymmetric dependence for asset allocation*", Journal of Financial Econometrics, Vol. 2, No. 1, pp. 130-168, 2004

Pilipovic, D., "*Energy Risk : Valuing and Managing Energy Derivatives*", McGraw-Hill, 2ª Edição, 2007

Romano, C., "*Applying Copula Function to Risk Management*", Risk Management Function, Banca di Roma, Viale *U.*, 2001

Scarsini, M., "*On measures of concordance*", Stochastica, VIII, 201-218, 1984

Schimid, F., Schimidt, R., *"Multivariate extensions of spearman's rho and related statistics"*, Statistics & Probability Letters 77, 407-416, 2007

Schmidt, R., Stadtmüller, U., "*Nonparametric estimation of tail dependence*", London School of Economics and University of Ulm, 2003

Sklar, A., "*Fonctions de répartition à n dimensions et leurs marges*" Publications dew l'Institut de Statistique de L'Université de Paris, 8:229-231, (1959).

Stefanova, D., "*Dependence modeling of joint extremes via copulas: a dynamic portfolio allocation perspective*", Canada Research Chair in Risk Management, HEC, Montreal, 2007

Stewart, J., "*Calculus*", 5ª Edição, Thomson Brooks/Cole, 2002

Sun, W., Rachev, S., Stoyanov, S., V., Fabozzi F., J., "*Multivariate skewed student's t copula in the analysis of nonlinear and asymmetric dependence in the German equity market*", Studies in Nonlinear Synamics & Econometrics, Vol. 12, No. 2, Article 3., 2008

Taylor, M., D., "*Multivariate measures of concordance*", The Institute of Statistical Mathematics, Tokyo, 2006

Texas Oil and Gas Association, "*What a Barrel of Crude Oil Makes*", http://www.txoga.org/articles/308/1/WHAT-A-BARREL-OF-CRUDE-OIL-MAKES, 1995

Williams, J., L., "*Oil Price History and Analisys*", WRTG Economics, http://www.wtrg.com/prices.htm, 2007

Whipple, T., "*Peak oil review*", Association for the Study of Peak Oil and Gás, EUA, Edição 14 de Julho de 2008