Universidade Federal do Rio Grande do Sul

Instituto de Matemática

Programa de Pós-Graduação em Matemática

Cópulas em Processos Estocásticos

por

GUILHERME PUMI

Porto Alegre, 23 de agosto de 2006.

Dissertação submetida por Guilherme Pumi[1] como requisito parcial para obtenção do grau de Mestre em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professora Orientadora:
Dra. Sílvia Regina Costa Lopes

Banca Examinadora:
Dr. Artur Oscar Lopes
Dra. Sara Ianda Correa Carmona
Dra. Sílvia Regina Costa Lopes
Dr. Luiz Koodi Hotta (IMECC/UNICAMP)

Data da Defesa: 23 de agosto de 2006.

[1]Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.

Agradecimentos

Aos meus pais pelo apoio incondicional. Minha mãe, exemplo absoluto de mulher, que me ensinou a enfrentar os problemas de cabeça erguida e sem hesitar sequer um instante e meu pai, exemplo de serenidade, tranquilidade e retidão. Obrigado.

Em especial, agradeço à Sílvia, orientadora e amiga que, com seus conselhos, orientação e apoio, ensinou-me muito mais do que a teoria. Obrigado.

À todos os colegas do Programa de Pós-Graduação em Matemática da UFRGS, em especial à Cleonis Viater Figueira, amiga e exemplo de mulher com a qual tive a oportunidade de estudar, à Rosane Fydryzewski, Edite Tauffer, Adriana Neumann de Oliveira, Leandro, Edson Werle (in memorian). Em especial também aos colegas da estatística matemática, Cléber Bisognin, Marcus Alexandre Nunes e Márcio Valk, pelo apoio e amizade. Obrigado.

Quero agradecer em especial à Tatiane Bagatini, que acompanhou-me desde o início do mestrado até praticamente o final, quando nossas estradas, impelidas por um desses truques que a vida real amiúde teima em impor, se separaram. Obrigado pelo amor, carinho, dedicação, companheirismo, apoio incondicional, pela paciência, pelas brincadeiras, pelo respeito ao meu trabalho e pela postura assumida durante o duro período de preparo desta dissertação. Obrigado também pelo empréstimo do computador que foi de grande valia para esta dissertação. Palavras não podem expressar minha gratidão. Guardo bem os momentos bons que vivemos juntos. Muito obrigado.

À Taiane Shaedler Prass que nesses últimos tempos me fez reconsiderar padrões, redefinir antigos conceitos e definir novos. Obrigado.

À todos os professores da pós, em especial ao Alexandre Tavares Baravieira, Artur Oscar Lopes, Jaime Bruck Ripoll, Luiz Fernando Carvalho da Rocha, Ivan Edgar Pan Peres. Em particular agradeço à professora Sara I. C. Carmona pelo auxílio com o Movimento Browniano e ao professor Luiz Koodi Hotta pelas valiosas sugestões e comentários apresentados para melhorar esta dissertação. Obrigado.

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

À CAPES (Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior) pelo apoio financeiro.

Ao pessoal da CEFAV, casa do estudante que me acolheu desde os meus primeiros dias em Porto alegre. Em especial para a turma do chimarrão que ajudaram a aliviar a tensão das longas horas de estudos, em particular, à Tati, Jeferson e Rosane, Simone, Elizandro e Marta, Edson (in memorian), Daiane, Vanina, Amauri, Bárbara, Paulinha, Roselei, enfim, tantos que se torna difícil citar todos. Obrigado a todos.

Por fim quero agradecer à todas aquelas pessoas que, de uma forma ou de outra, contribuiram para esta fase da minha vida. Obrigado a todos.

Para minha vó Cezira Kostaneski que, por imprudência
ou desatenção de uma outra pessoa,
foi impedida de viver até este momento.
Ela sabe o quão importante foi para todos nós.
Em especial para mim.

... and you will learn to lose everything.
We are temporary arrangements.

Resumo

O presente trabalho discute formalmente vários aspectos da teoria de cópulas tanto no caso bidimensional quanto no caso m-dimensional. São apresentados os principais teoremas da teoria de cópulas, métodos recentes e tradicionais para a construção de cópulas, algumas cópulas associadas a funcionais do Movimento Browniano são calculadas e alguns métodos de estimação baseados em cópulas são apresentados.

Abstract

The present work formally discusses several aspects of the copula's theory in the bidimensional case and in the m-dimensional case. Here we present the main theorems, recent and traditional methods for copulas' construction and copulas associated with some functionals of the Brownian Motion. Some traditional copula-based estimation methods are also presented.

Índice

Introdução

Um dos problemas mais estudados concernentes à probabilidade e à estatística matemática é sem dúvida como lidar com a dependência entre duas ou mais variáveis aleatórias e como caracterizar e modelar tal dependência. Quando trabalhamos com dados reais, geralmente impomos restrições ao tipo de dependência com que estamos tratando, já que uma estrutura de dependência arbitrária é matematicamente intratável. Para abordar este tipo de problema, muitas ferramentas e técnicas foram desenvolvidas com o passar do tempo. As cópulas são uma dessas ferramentas. O desenvovimento da teoria de cópulas nos últimos anos ocorreu especialmente devido ao impulso recebido em campos como estatística, finanças e ciências atuariais, onde muitas aplicações importantes foram descobertas.

Grosseiramente falando, cópulas são funções que ligam uma função distribuição m-dimensional às suas marginais de qualquer dimensão. Sua história está intimamente relacionada ao desenvolvimento da teoria de espaços métricos probabilísticos. A necessidade de se definir um análogo probabilístico à desigualdade triangular deu origem ao estudo da teoria de cópulas. O principal resultado dentro desta teoria, o Teorema de Sklar (veja o Teorema 1.3.1), bem como outros resultados importantes concernentes à teoria de cópulas, foram provados nesta época, que culminou com o lançamento, em 1983, do livro "Probabilistic Metric Spaces" de B. Schweizer e A. Sklar (veja Schweizer e Sklar, 2005 para uma introdução à história do desenvolvimento da teoria de espaços métricos probabilísticos). Após isso, as cópulas permaneceram esquecidas por alguns anos, até serem redescobertas na década de 90. Desde então o interesse no assunto tem crescido rapidamente.

Conforme exposto acima, cópulas são funções que fazem a conexão entre uma função de distribuição e suas marginais de qualquer dimensão. Uma das características mais atraente e importante das cópulas é que estas, sozinhas, possuem todas as informações relevantes a respeito da estrutura de dependência entre as variáveis aleatórias em questão. Isso permite que as cópulas sejam ferramentas flexíveis na modelagem de funções de distribuição e suas marginais. Além disso, cópulas são interessantes também na construção de medidas de dependência livres de escala. Várias aplicações de cópulas foram descobertas nos últimos anos substituindo, inclusive, certos métodos clássicos. Um exemplo clássico é o da fórmula de Black-Scholes, substituída por ferramentas baseadas em cópulas (veja Cherubini et al., 2004).

Esta dissertação tem por objetivo apresentar uma revisão bibliográfica sobre cópulas envolvendo teoria e aplicações. O assunto é exposto de uma maneira formal apresentando provas a quase todos os resultados presentes no texto. Sempre que possível, apresentamos provas diferentes daquelas encontradas na literatura básica do assunto. Nos preocupamos também em apresentar aspectos da teoria que não estão a disposição nos poucos livros sobre o assunto, mas que se encontram em artigos da área, procurando sempre tornar a teoria e as provas mais detalhadas do que nos respectivos artigos.

O trabalho é dividido como segue. O Capítulo 1 trata das propriedades gerais das cópulas. Introduzimos inicialmente os conceitos necessários ao trabalho com cópulas apresentando, a seguir, os principais resultados concernentes às cópulas, incluindo o Teorema de Sklar. Apresentamos também a transformada integral de probabilidade e a representação canônica de cópulas. No final do capítulo tratamos da teoria de cópulas m-dimensionais incluindo os principais teoremas.

O Capítulo 2 trata da construção de cópulas. Apresentamos as construções da Família Ali-Mikhail-Haq, Família Plackett, a construção de Dolati e Úbeda-Flores e a construção da Família Arquimediana de cópulas.

O Capítulo 3 trata da teoria de cópulas em relação à funcionais do Processo de Movimento Browniano. Iniciamos o capítulo com uma breve introdução aos processos de Markov a tempo contínuo e Movimento Browniano. Nas seções seguintes é calculada a cópula associada ao Movimento Browniano e seu máximo e do máximo do Movimento Browniano e o seu tempo de primeira chegada. Muitos dos resultados desse capítulo são inéditos e serão, posteriormente, preparados em forma de artigo para submissão à revista internacional.

O Capítulo 4 trata do cálculo da estrutura de covariância, via Lema de Hoeffding, de famílias de cópulas reparametrizadas convenientemente. Este capítulo deu origem ao trabalho Lopes e Pumi (2006).

O Capítulo 5 apresenta vários métodos de estimação baseados em cópulas como o método da Máxima Verossimilhança, o método IFM (*Inference for the Margins*), o método da Pseudo-Verossimilhança, um método de estimação baseado na medida de associação conhecida por τ de Kendall e um método de estimação em Famílias Arquimedianas.

As conclusões finais referentes a esta dissertação são apresentadas no Capítulo 6.

Capítulo 1

Cópulas

O que são cópulas? A grosso modo podemos dizer que cópulas são funções que ligam funções de distribuições multivariadas a suas marginais. Equivalentemente podemos dizer que cópulas são funções de distribuições cujas marginais são uniformes no intervalo $[0, 1]$. Suponha que dispomos de duas variáveis aleatórias X e Y, com função de distribuição $F(x) = P(X \leq x)$ e $G(y) = P(Y \leq y)$, respectivamente, e função de distribuição conjunta $H(x, y) = P(X \leq x, Y \leq y)$. Dado um ponto $(x, y) \in \mathbb{R}^2$ podemos associar a ele três números: $F(x)$, $G(y)$ e $H(x, y)$. Note que cada um desses números pertence ao intervalo $[0, 1]$. Assim, para cada ponto $(x, y) \in \mathbb{R}^2$, associamos um par $(F(x), G(y)) \in [0, 1] \times [0, 1]$, e a este par ordenado corresponde um número $H(x, y) \in [0, 1]$. Essa correspondência que, para cada par ordenado de valores das funções de distribuição individuais, atribui o valor da distribuição conjunta é de fato uma função (e vamos provar isso) a qual chamamos de cópula.

É claro que nenhuma das idéias acima é uma definição. Na próxima seção vamos introduzir alguns conceitos básicos para a definição de cópulas.

A importância e o recente crescimento do interesse na teoria de cópulas muito se deve ao fato de as cópulas ligarem uma função de distribuição m-dimensional às suas marginais de qualquer ordem de forma que todas as informações relevantes sobre a estrutura de dependência ficam unicamente caracterizadas pela cópula em questão. Este fato permite às cópulas uma flexibilidade muito grande na modelagem de funções de distribuições e suas marginais. Além disso, as cópulas são ferramentas interessantes na construção de medidas de dependência livres de escala, na simulação de famílias de distribuições m-dimensionais, na modelagem de dependência entre variáveis aleatórias, etc.

1.1 Conceitos Preliminares

Neste trabalho, denotaremos por $\mathbb{R}$ a reta real $(-\infty, \infty)$. Um *retângulo* em $\mathbb{R}^2$ é o produto cartesiano R de dois intervalos fechados, i.e., $R = [x_1, x_2] \times [y_1, y_2]$. Os vértices do retângulo R são os pontos (x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2). Também denotaremos por I^2 o quadrado unitário $I \times I = [0,1] \times [0,1]$. Uma função $f : S \subseteq \mathbb{R}^2 \to \mathbb{R}$ será chamada de *2-função real*.

Definição 1.1.1. Sejam S_1 e S_2 dois subconjuntos não-vazios de $\mathbb{R}$, e seja H uma função cujo domínio seja $\mathrm{Dom}(H) = S_1 \times S_2$. Seja $R = [x_1, x_2] \times [y_1, y_2]$ um retângulo cujos vértices estejam contidos em $\mathrm{Dom}(H)$.
Definimos o *H-volume de R* por

$$V_H(R) = H(x_1, y_1) + H(x_2, y_2) - H(x_1, y_2) - H(x_2, y_1). \tag{1.1}$$

Definição 1.1.2. Uma 2-função real H é dita ser 2-*crescente* se $V_H(R) \geq 0$ para todo retângulo R cujos vértices estejam contidos em $\mathrm{Dom}(H)$.

Observação 1.1.1. Alguns autores se referem a funções 2-crescentes por funções *quase-monótonas*.

Conforme mostra os próximos exemplos, uma função H ser 2-crescente não implica nem é implicado por H ser não-decrescente em cada argumento.

Exemplo 1.1.1. Considere a função definida em I^2 por $H(x,y) = (1-x)(1-y)$. É claro que H é não-crescente em cada um dos seus argumentos. Porém, dado o retângulo $R = [x_1, x_2] \times [y_1, y_2] \subseteq I^2$, temos

$$\begin{aligned} V_H(R) &= (1-x_1)(1-y_1) + (1-x_2)(1-y_2) - (1-x_1)(1-y_2) \\ &\quad -(1-x_2)(1-y_1) \\ &= (1-x_1)(y_2-y_1) + (1-x_2)(y_1-y_2) \\ &= (y_2-y_1)(x_2-x_1) \geq 0. \end{aligned}$$

Logo, H é 2-crescente.

Exemplo 1.1.2. Considere a função $H(x,y) = \ln(x+y+1)$ definida em I^2. É claro que H é crescente em cada um de seus argumentos, mas,

$$\begin{aligned} V_H([1/2, 1] \times [1/2, 1]) &= \ln(2) + \ln(3) - \ln(5/2) - \ln(5/2) \\ &= 3\ln(2) + \ln(3) - 2\ln(5) \\ &= \ln\left(\frac{24}{25}\right) < 0. \end{aligned}$$

Logo, H não é 2-crescente.

Exemplo 1.1.3. Considere a função definida em I^2 por $H(x, y) = e^{x+y}$. É claro que H é crescente em cada argumento. Além disso, dado o retângulo $R = [x_1, x_2] \times [y_1, y_2] \subseteq I^2$,

$$\begin{aligned} V_H(R) &= e^{x_1+y_1} + e^{x_2+y_2} - e^{x_1+y_2} - e^{x_2+y_1} \\ &= e^{x_1}(e^{y_1} - e^{y_2}) + e^{x_2}(e^{y_2} - e^{y_1}) \\ &= (e^{x_2} - e^{x_1})(e^{y_2} - e^{y_1}) \geq 0. \end{aligned}$$

Logo, H é 2-crescente.

Sejam S_1 e S_2 subconjuntos não-vazios de $\mathbb{R}$ e suponha que S_1 tenha um elemento minimal a_1 e um elemento maximal b_1 e que S_2 possua um elemento minimal a_2 e um elemento maximal b_2. Dizemos que uma função $H : S_1 \times S_2 \to \mathbb{R}$ é grounded quando $H(x, a_2) = 0$ e $H(a_1, y) = 0$ para todo $(x, y) \in S_1 \times S_2$ e dizemos que a função tem *marginais* $F : S_1 \longrightarrow \mathbb{R}$ e $G : S_2 \longrightarrow \mathbb{R}$ dadas por

$$\begin{aligned} F(x) &= H(x, b_2) \\ G(y) &= H(b_1, y). \end{aligned}$$

Exemplo 1.1.4. Considere a função H com $\mathrm{Dom}(H) = [0, 1] \times [0, \infty)$ definida por

$$H(x, y) = xy \left[e^{-y} + \frac{e^{1-x}}{y+1} \right].$$

Então, H é *grounded* e tem marginais

$$\begin{aligned} F(x) &= H(x, \infty) = \lim_{y\to\infty} x \left[\frac{y}{e^y} + e^{1-x} \frac{y}{y+1} \right] = xe^{1-x}, \\ G(y) &= H(1, y) = y \left[e^{-y} + \frac{1}{y+1} \right]. \end{aligned}$$

No próximo teorema (veja Nelsen, 1999, página 7) provamos que uma 2-função real que seja 2-crescente e *grounded* também é não-decrescente em cada argumento.

Lema 1.1.1. *Sejam S_1 e S_2 dois subconjuntos não-vazios de $\mathbb{R}$ e seja H uma função 2-crescente e* grounded *cujo domínio seja $S_1 \times S_2$. Então, H é não-decrescente em cada argumento.*

Prova: Sejam $x_1, x_2 \in S_1$ e $y_1, y_2 \in S_2$ tais que $x_1 < x_2$ e $y_1 < y_2$. Vamos provar primeiramente que a função,

$$\begin{aligned} f : S_1 &\longrightarrow \mathbb{R} \\ t &\longmapsto H(t, y_2) - H(t, y_1) \end{aligned}$$

é não-decrescente em S_1 e a função,

$$\begin{array}{rcl} g : S_2 & \longrightarrow & \mathbb{R} \\ t & \longmapsto & H(x_2, t) - H(x_1, t) \end{array}$$

é não-decrescente em S_2.
Para isso, note que, como H é 2-crescente, $V_H(R) \geq 0$ qualquer que seja o retângulo R cujos vértices estejam contidos em $S_1 \times S_2$. Sejam $t_1, t_2 \in S_1$ com $t_1 < t_2$. Assim,

$$\begin{array}{rcl} f(t_2) - f(t_1) & = & H(t_2, y_2) - H(t_2, y_1) - [H(t_1, y_2) - H(t_1, y_1)] \\ & = & V_H([t_1, t_2] \times [y_1, y_2]) \geq 0. \end{array}$$

Logo, $t_1 < t_2 \Longrightarrow f(t_1) \leq f(t_2)$, o que prova que $f(\cdot)$ é não-decrescente em S_1. Analogamente prova-se que $g(\cdot)$ é não-decrescente em S_2. Seja a_1 o menor elemento em S_1 e sejam y_1, $y_2 \in S_2$ com $y_1 < y_2$. Dessa forma,

$$f(t) - f(a_1) = H(t, y_2) - H(t, y_1) - [H(a_1, y_2) - H(a_1, y_1)] \geq 0,$$

para todo $t \in S_1$. Como H é *grounded*, temos que $H(a_1, y_1) = 0 = H(a_1, y_2)$. Logo,

$$y_1 < y_2 \Longrightarrow H(t, y_1) \leq H(t, y_2), \quad \text{para todo } t \in S_1.$$

Isto prova que H é não-decrescente no segundo argumento. A prova para o primeiro argumento é análoga. □

1.2 Cópulas

Nesta seção vamos definir formalmente cópulas e provar algumas de suas propriedades mais importantes. Provaremos que as cópulas são funções uniformemente contínuas, Lipschitzianas e que possuem derivadas parciais em quase toda a parte. Vamos mostrar também que o subconjunto do espaço das funções contínuas de $I^2 \to I$, formado por todas as cópulas é um conjunto convexo.

Definição 1.2.1. Uma *cópula bidimensional* (ou *2-cópula* ou simplesmente *cópula*) é uma função $C : I^2 \to I$ *grounded*, com marginais u e v e 2-crescente. Equivalentemente, uma cópula é uma função $C : I^2 \to I$ com as seguintes propriedades:

(i). Para todo u, v em I,

$$C(u, 0) = 0 = C(0, v), \tag{1.2}$$

e

$$C(u, 1) = u \quad \text{e} \quad C(1, v) = v; \tag{1.3}$$

(ii). Para todo retângulo $R = [u_1, v_1] \times [u_2, v_2] \subseteq I^2$,

$$V_C(R) = C(u_1, u_2) + C(v_1, v_2) - C(u_1, v_2) - C(u_2, v_1) \geq 0. \tag{1.4}$$

O próximo teorema estabelece a continuidade uniforme das cópulas. Note primeiramente que, pelo Teorema 1.1.1, uma cópula é não-decrescente em cada um de seus argumentos e, portanto, por (1.3), é limitada.

Teorema 1.2.1. *Seja C uma cópula. Então C é uniformemente contínua em I^2.*

Prova: Sejam u_1, u_2, v_1, $v_2 \in I$. Vamos provar inicialmente que C satisfaz a condição de Lipschitz. Pela desigualdade triangular temos que,

$$\begin{aligned} |C(u_2, v_2) - C(u_1, v_1)| &= |C(u_2, v_2) - C(u_1, v_2) + C(u_1, v_2) - C(u_1, v_1)| \\ &\leq |C(u_2, v_2) - C(u_1, v_2)| + |C(u_1, v_2) - C(u_1, v_1)|. \end{aligned} \tag{1.5}$$

Suponha $u_1 \leq u_2$ e considere o retângulo $R_1 = [u_1, u_2] \times [v_2, 1]$, assim,

$$V_C(R_1) = C(u_1, v_2) - C(u_2, v_2) + u_2 - u_1 \geq 0.$$

Segue do Lema 1.1.1 e da equação acima que,

$$0 \leq C(u_2, v_2) - C(u_1, v_2) \leq u_2 - u_1. \tag{1.6}$$

Suponha $u_2 \leq u_1$ e considere $R_2 = [u_2, u_1] \times [v_2, 1]$ temos,

$$V_C(R_2) = C(u_2, v_2) - C(u_1, v_2) + u_1 - u_2 \geq 0.$$

Segue da equação acima e do Lema 1.1.1 que,

$$0 \geq C(u_2, v_2) - C(u_1, v_2) \geq u_2 - u_1. \tag{1.7}$$

Juntando (1.6) e (1.7) obtemos,

$$|C(u_2, v_2) - C(u_1, v_2)| \leq |u_2 - u_1|. \tag{1.8}$$

Analogamente prova-se que,

$$|C(u_1, v_2) - C(u_1, v_1)| \leq |v_2 - v_1|. \tag{1.9}$$

Substituindo (1.8) e (1.9) em (1.5) obtemos,

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|. \tag{1.10}$$

Denotando por $\| \cdot \|_S$ a norma da soma em $\mathbb{R}^2$, de (1.10) segue que,

$$|C(u_2, v_2) - C(u_1, v_1)| \leq \| (u_2, v_2) - (u_1, v_1) \|_S .$$

O que prova que C é Lipschitziana (com constante de Lipschitz igual a 1) e, portanto, uniformemente contínua. $\square$

Além de uniformemente contínuas, as cópulas também são diferenciáveis em quase toda a parte em relação à medida de Lebesgue, como mostra o próximo teorema.

Teorema 1.2.2. *Seja C uma cópula qualquer. Então, $\frac{\partial C(u,v)}{\partial u}$ e $\frac{\partial C(u,v)}{\partial v}$ existem em quase toda a parte de I, tomam valores em I e são não-decrescentes em quase toda a parte.*

Prova: A existência das derivadas parciais é imediata já que, pelo Lema 1.1.1, uma cópula é não-decrescente em cada argumento e funções não-decrescentes em intervalos compactos são diferenciáveis em quase toda a parte (veja Fernandez, 1976, página 154). Pela equação (1.10), fixado $u \in [0,1]$, se $v \in (0,1)$ temos,

$$0 \leq \lim_{h\to 0} \frac{C(u,v+h) - C(u,v)}{h} \leq \lim_{h\to 0} \frac{|u-u| + |v+h-v|}{h} = 1,$$

em quase toda a parte, ou seja,

$$0 \leq \frac{\partial C(u,v)}{\partial v} \leq 1,$$

para quase todo $v \in (0,1)$. Usando-se limites laterais na desigualdade acima prova-se que a desigualdade vale para quase todo $v \in [0,1]$. Analogamente prova-se a desigualdade para $\frac{\partial C(u,v)}{\partial u}$. Tomando-se $v_1 \leq v_2$ e usando (1.7), o Teorema 1.2.1 garante que a função $u \longmapsto C(u,v_2) - C(u,v_1)$ é não-decrescente. Logo, a função,

$$\frac{\partial \left(C(u,v_2) - C(u,v_1)\right)}{\partial u},$$

está bem definida e é não-negativa em quase toda a parte em I. Pela linearidade da derivada segue que,

$$0 \leq \frac{\partial \left(C(u,v_2) - C(u,v_1)\right)}{\partial u} = \frac{\partial}{\partial u} C(u,v_2) - \frac{\partial}{\partial u} C(u,v_1).$$

Logo,

$$v_1 \leq v_2 \implies \frac{\partial C(u,v_1)}{\partial u} \leq \frac{\partial C(u,v_2)}{\partial u},$$

em quase toda a parte de I. Analogamente prova-se que $\frac{\partial C(u,v)}{\partial v}$ é não-decrescente em quase toda a parte. Uma prova um pouco diferente, porém mais curta, pode ser encontrada em Nelsen (1999), página 11. $\square$

Exemplo 1.2.1. A função $M : I^2 \to I$ dada por $M(u,v) = \min\{u,v\}$ é uma cópula. É evidente que M é *grounded* e possui marginais u e v. Para provar que M é 2-crescente considere o retângulo $R = [u_1,v_1] \times [u_2,v_2] \subseteq I^2$. Teremos 6 casos distintos para analisar, dependendo da ordem entre u_1, v_1, u_2 e v_2.
Se $v_1 \leq u_2$ temos,

$$\begin{aligned} V_M(R) &= \min\{u_1,u_2\} + \min\{v_1,v_2\} - \min\{u_1,v_2\} - \min\{v_1,u_2\} \\ &= u_1 + v_1 - u_1 - v_1 = 0. \end{aligned}$$

Se $v_2 \leq u_1$ temos,

$$V_M(R) = u_2 + v_2 - u_2 - v_2 = 0.$$

Se $u_1 \leq u_2 \leq v_1 \leq v_2$ temos,

$$V_M(R) = u_1 + v_1 - u_1 - v_1 = v_1 - u_2 \geq 0.$$

Se $u_2 \leq u_1 \leq v_2 \leq v_1$ temos,

$$V_M(R) = u_2 + v_2 - u_1 - u_2 = v_2 - u_1 \geq 0.$$

Se $u_2 \leq u_1 \leq v_1 \leq v_2$ temos,

$$V_M(R) = u_2 + v_1 - u_1 - u_2 = v_1 - u_1 \geq 0.$$

Se $u_1 \leq u_2 \leq v_2 \leq v_1$ temos,

$$V_M(R) = u_1 + v_2 - u_1 - u_2 = v_2 - u_2 \geq 0.$$

Logo, $V_M(R) = \max\{0, \min\{v_1, v_2\} - \max\{u_1, u_2\}\} \geq 0$ em qualquer caso, o que prova que M é 2-crescente e, portanto, uma cópula. A cópula M é chamada *limite superior de Fréchet-Hoeffding.*

Exemplo 1.2.2. A função $W : I^2 \to I$ dada por $W(u,v) = \max\{u+v-1, 0\}$ é uma cópula e é chamada de *limite inferior de Fréchet-Hoeffding.* É fácil ver que W é *grounded* e tem marginais u e v. Para provar que W é 2-crescente usa-se procedimento análogo ao do exemplo anterior.

As cópulas M e W apresentadas nos Exemplos 1.2.1 e 1.2.2 são conhecidas como limites inferior e superior de Fréchet-Hoeffding, respectivamente, devido ao teorema a seguir (veja Nelsen, 1999, página 8):

Teorema 1.2.3. *Seja C uma cópula qualquer e sejam W e M os limites inferior e superior de Fréchet-Hoeffding, respectivamente. Então, para todo $(u,v) \in I^2$,*

$$W(u,v) \leq C(u,v) \leq M(u,v). \tag{1.11}$$

Prova: Seja $(u,v) \in I^2$. Em virtude do Lema 1.1.1, C é não-decrescente em ambos os argumentos e assim,

$$C(u,v) \leq C(u,1) = u \quad \text{e} \quad C(u,v) \leq C(1,v) = v,$$

o que estabelece $C(u,v) \leq \min\{u,v\} = M(u,v)$.
Por outro lado, como C é 2-crescente, segue que,

$$\begin{aligned} V_C([u,1] \times [v,1]) &= C(u,v) + C(1,1) - C(u,1) - C(1,v) \\ &= C(u,v) + 1 - u - v \geq 0 \iff \\ C(u,v) &\geq u + v - 1. \end{aligned} \tag{1.12}$$

Além disso, como C é não-decrescente e *grounded*, temos que $C(u,v) \geq 0$ qualquer que seja $(u,v) \in I^2$, desigualdade que combinada a (1.12) estabelece $C(u,v) \geq \max\{u+v-1, 0\} = W(u,v)$. □

Exemplo 1.2.3. Sejam C_1 e C_2 cópulas quaisquer. Considere a função C, definida em I^2, dada por $C(u,v) = C_1(u,v) + C_2(u,v)$. Note que C é *grounded*, 2-crescente e é não-decrescente em ambos argumentos. Porém C falha em ser uma cópula pois tem marginais $2u$ e $2v$.

Exemplo 1.2.4. Sejam C_1 e C_2 cópulas quaisquer e considere $\alpha \in I$ e a função C, definida em I^2, dada por $C(u,v) = \alpha C_1(u,v) + (1-\alpha)C_2(u,v)$. Novamente é fácil ver que C é *grounded*, possui marginais u e v e é 2-crescente. Sendo assim, C é uma cópula, quaisquer que sejam as cópulas C_1 e C_2 e $\alpha \in I$.

O Exemplo 1.2.3 mostra que a soma de duas cópulas em geral não é uma cópula. No próximo teorema apresentamos uma condição necessária e suficiente para que uma combinação linear de cópulas seja uma cópula. Antes provaremos um lema necessário à prova do teorema.

Lema 1.2.1. *Sejam $C_1, C_2, \cdots, C_n$ cópulas quaisquer e $\alpha_1, \alpha_2, \cdots, \alpha_n$ números reais nem todos nulos. Defina $C(u,v) := \sum_{i=1}^{n} \alpha_i C_i(u,v)$. Seja R um retângulo cujos vértices estejam contidos em I^2. Então,*

$$V_C(R) = \sum_{i=1}^{n} \alpha_i V_{C_i}(R). \tag{1.13}$$

Além disso, se $V_C(R) \geq 0$ para todo o retângulo $R \subseteq I^2$, então, $\sum_{i=1}^{n} \alpha_i \geq 0$. Ainda, se $\alpha_i \geq 0$, para todo i, então, $V_C(R) \geq 0$, para todo o retângulo $R \subseteq I^2$.

Prova: Sejam u_1, u_2, v_1, $v_2 \in I$, com $u_1 \leq u_2$ e $v_1 \leq v_2$ e seja o retângulo $R = [u_1, u_2] \times [v_1, v_2] \subseteq I^2$. Então,

$$\begin{aligned}
V_C(R) &= C(u_1,v_1) + C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) \\
&= \sum_{i=1}^{n} \alpha_i C_i(u_1,v_1) + \sum_{i=1}^{n} \alpha_i C_i(u_2,v_2) \\
&\quad - \sum_{i=1}^{n} \alpha_i C_i(u_1,v_2) - \sum_{i=1}^{n} \alpha_i C_i(u_2,v_1) \\
&= \sum_{i=1}^{n} \alpha_i [C_i(u_1,v_1) + C_i(u_2,v_2) - C_i(u_1,v_2) - C_i(u_2,v_1)] \\
&= \sum_{i=1}^{n} \alpha_i V_{C_i}(R).
\end{aligned}$$

Em particular, $V_C(I^2) = \sum_{i=1}^{n} \alpha_i V_{C_i}(I^2) = \sum_{i=1}^{n} \alpha_i \geq 0$. A última afirmação é óbvia, já que, para todo i, C_i é cópula. $\square$

A condição $\sum_{i=1}^{n} \alpha_i \geq 0$ sozinha não é suficiente para garantir que $V_C(R) \geq 0$, para todo retângulo $R \subseteq I^2$ como mostra o próximo exemplo.

Exemplo 1.2.5. Defina $C(u,v) := 4W(u,v) - \Pi(u,v) - 2M(u,v)$ em I^2, onde $\Pi(u,v) = uv$ é a *cópula independência* (ou *cópula produto*). Note que C é uma combinação linear de cópulas e que a soma de seus coeficientes é 1. Mas se u_1, u_2, v_1 e $v_2 \in I$ são tais que $u_1 < u_2 \leq v_1 < v_2$ e $v_1 + v_2 \leq 1$ (note que estas condições implicam $W(u_i, v_j) = 0$, para $i, j = 1, 2$), então, após algumas simplificações,

$$
\begin{aligned}
V_C([u_1, u_2] \times [v_1, v_2]) &= (u_2 - u_1)(v_1 - v_2) - 2u - 2v + 2u + 2v \\
&= (u_2 - u_1)(v_1 - v_2) < 0.
\end{aligned}
$$

Teorema 1.2.4. *Sejam $C_1, C_2, \cdots, C_n$ cópulas e $\alpha_1, \alpha_2, \cdots, \alpha_n$ números reais não-negativos e nem todos nulos. Defina,*

$$C(u,v) := \sum_{i=1}^{n} \alpha_i C_i(u,v).$$

Então, $C(u,v)$ é uma cópula se, e somente se, $\displaystyle\sum_{i=1}^{n} \alpha_i = 1.$

Prova: Note primeiramente que cada C_i é definida em I^2, logo $\mathrm{Dom}(C) = I^2$. Além disso é evidente que C é *grounded*, já que cada C_i o é. Suponha que C é uma cópula. Como $\alpha_1, \alpha_2, \cdots, \alpha_n$ são não-negativos, segue do Lema 1.2.1 que C é 2-crescente. As marginais para que C seja cópula devem ser tais que,

$$v = C(1,v) = \sum_{i=1}^{n} \alpha_i C_i(1,v) = v \sum_{i=1}^{n} \alpha_i \quad \Longleftrightarrow \quad \sum_{i=1}^{n} \alpha_i = 1.$$

Analogamente para a marginal $C(u,1)$.

Reciprocamente, suponha que $\displaystyle\sum_{i=1}^{n} \alpha_i = 1$. Então,

$$C(u,1) = \sum_{i=1}^{n} \alpha_i C_i(u,1) = u \sum_{i=1}^{n} \alpha_i = u.$$

Analogamente, $C(1,v) = v$. Isso mostra que C tem marginais u e v. Pela equação (1.13), como C_i é cópula e $\alpha_i \geq 0$, para todo $i = 1, \cdots, n$, segue que $V_C(R) \geq 0$. Logo, C é cópula. $\square$

Uma conseqüência imediata do Teorema 1.2.4 é que não é qualquer combinação linear de cópulas que é cópula. O Exemplo 1.2.4 mostra que qualquer combinação linear convexa de cópulas é uma cópula. Isto quer dizer que o subconjunto das funções de I^2 em I formadas pelas cópulas é um conjunto convexo.

Caracterizadas as combinações lineares de cópulas, podemos nos perguntar se existem cópulas que podem ser escritas como produto de funções de seus argumentos. Isto é, será que existem cópulas que podem ser fatoradas como $C(u, v) = f(u)g(v)$ para funções f e g convenientes? A cópula do Exemplo 1.2.5 é um exemplo onde a fatoração vale. Será que existem outras? A resposta negativa é dada pela próxima proposição.

Proposição 1.2.1. *Seja C uma cópula e suponha que existem funções f e g definidas em I^2 tais que $C(u, v) = f(u)g(v)$. Então, $C = \Pi$.*

Prova: Como C é cópula,

$$v = C(1, v) = f(1)g(v) \quad \Longrightarrow \quad g(v) = \frac{v}{f(1)} \text{ para todo } v \in I.$$

Analogamente $f(u) = u[g(1)]^{-1}$, $\forall u \in I$. Logo, $C(u, v) = [f(1)g(1)]^{-1}uv$. Segue do Teorema 1.2.4 que, obrigatoriamente, $[f(1)g(1)]^{-1} = 1$ o que completa a prova. $\square$

1.3 Teorema de Sklar

O Teorema de Sklar (veja Teorema 1.3.1) é resultado chave na teoria de cópulas e é responsável pela maioria das aplicações das cópulas à estatística. É ele quem elucida o papel das cópulas na relação entre as funções de distribuição conjunta e suas marginais univariadas.

No que segue, letras maiúsculas, como X e Y, denotarão variáveis aleatórias e letras minúsculas, como x e y, denotarão valores assumidos por variáveis aleatórias. Assumiremos também que as variáveis aleatórias em questão estão definidas num mesmo espaço de probabilidade $(\Omega, \mathcal{A}, P)$, onde Ω é um conjunto não-vazio qualquer, chamado *espaço amostral*, $\mathcal{A}$ é a classe não-vazia de subconjuntos de Ω, também chamada de *classe dos eventos aleatórios* e $P : \mathcal{A} \longrightarrow [0, 1]$ uma função que, a cada evento $A \in \mathcal{A}$, associa uma *probabilidade* $P(A)$.

Definição 1.3.1. Seja Ω um conjunto qualquer. Uma *σ-álgebra* $\mathcal{A}$ é uma classe de subconjuntos não-vazios de Ω que satisfaz as seguintes propriedades:

(i). $\Omega \in \mathcal{A}$;

(ii). Se $A \in \mathcal{A}$, então, $A^{\complement} \in \mathcal{A}$;

(iii). Se $\{A_n\}_{n\in T}$ é uma seqüência enumerável de conjuntos em $\mathcal{A}$, então, $\cup_{n\in T} A_n \in \mathcal{A}$, onde $T \subseteq \mathbb{N}$ é um conjunto de índices.

No que segue vamos assumir que $\mathcal{A}$ na terna $(\Omega, \mathcal{A}, P)$ é a σ-álgebra de Borel. Para o leitor interessado em uma introdução aos espaços de probabilidade, veja Rohatgi (1976). Mais detalhes podem ser encontrados em Billingsley (1995) e Laha e Rohatgi (1979). Vamos também nos referir ao conjunto $\{\omega \in \Omega; X(\omega) \leq x\}$ por $[X \leq x]$.

Definição 1.3.2. Uma *função de distribuição* é uma função $F : \mathbb{R} \to I$ tal que

(i). F é não-decrescente;

(ii). F é contínua à direita;

(iii). $F(-\infty) = 0$ e $F(+\infty) = 1$.

Diremos que F é uma função de distribuição de uma variável aleatória X quando, para todo $x \in \mathbb{R}$, $F(x) = P(X \leq x)$. Note que a função F assim definida é uma função de distribuição também no sentido da Definição 1.3.2. Diremos também que uma variável aleatória X é contínua se sua função de distribuição o for.

Definição 1.3.3. Uma *função de distribuição conjunta* é uma função $H : \mathbb{R}^2 \to I$ tal que

(i). H é não-decrescente e contínua à direita em ambos os argumentos;

(ii). H é 2-crescente;

(iii). $H(x, -\infty) = H(-\infty, y) = 0$ e $H(+\infty, +\infty) = 1$.

Note que H é *grounded*, e como $\mathrm{Dom}(H) = \mathbb{R}^2$, H tem marginais $F(x)$ e $G(y)$ dadas por $F(x) = H(x, +\infty)$ e $G(y) = H(+\infty, y)$. Pelo Lema 1.1.1 segue que F e G são funções de distribuição.

Para duas ou mais variáveis aleatórias, adotaremos uma notação análoga ao caso univariado, assim o vetor de variáveis aleatórias (X, Y) é descrito por uma distribuição conjunta $H(x, y) = P(X \leq x, Y \leq y)$. Pode-se provar que $H(\cdot\,, \cdot)$ assim definida é também uma função de distribuição conjunta no sentido da Definição 1.3.3. Veja Laha e Rohatgi (1979), página 20.

Dadas duas variáveis aleatórias X e Y, definimos a *probabilidade condicional* de X dado Y por

$$P(X \in A | Y \in B) = \frac{P(X \in A, Y \in B)}{P(Y \in B)}\,,$$

para todo A e B borelianos em $\mathbb{R}$, sempre que $P(Y \in B) > 0$.

Exemplo 1.3.1. A função

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt,$$

definida em $\mathbb{R}$ é uma função de distribuição, chamada de *distribuição Gaussiana*. De fato, como $e^{\frac{-t^2}{2}}$ é uma função não-negativa, segue que Φ é não-decrescente. Além disso,

$$\Phi(-\infty) = \lim_{x \to -\infty} \Phi(x) = \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt = 0.$$

Para provar que $\Phi(+\infty) = 1$, basta provar que $\left(\int_{\mathbb{R}} e^{\frac{-t^2}{2}} dt \right)^2 = 2\pi$. Mas,

$$\begin{aligned} \left(\int_{\mathbb{R}} e^{\frac{-t^2}{2}} dt \right)^2 &= \left(\sqrt{2} \int_{\mathbb{R}} e^{-x^2} dx \right) \left(\sqrt{2} \int_{\mathbb{R}} e^{-y^2} dy \right) \\ &= 2 \iint_{\mathbb{R}^2} e^{-x^2-y^2} dxdy. \end{aligned}$$

Usaremos coordenadas polares, $x = r\cos\theta$, $y = r\sin\theta$ e $dxdy = rdrd\theta$. Substituindo acima obtemos,

$$\begin{aligned} \left(\int_{\mathbb{R}} e^{\frac{-t^2}{2}} dx \right)^2 &= 2 \int_0^{2\pi} \int_0^{\infty} re^{-r^2} drd\theta \\ &= 2 \int_0^{2\pi} \frac{1}{2}\, d\theta \\ &= 2\pi. \end{aligned}$$

Exemplo 1.3.2. Seja a função com domínio $\mathbb{R}^2$ dada por

$$H(x,y) = \begin{cases} \dfrac{(1-e^{-y})x}{(x-5)e^{-y}+5}, & \text{se} \quad (x,y) \in [0,5] \times [0,\infty), \\ 1 - e^{-y}, & \text{se} \quad (x,y) \in [5,\infty) \times [0,\infty), \\ 0, & \text{c.c.} \end{cases}$$

É claro que H assim definida é tal que $H(x,-\infty) = 0 = H(-\infty,y)$, ou seja, H é *grounded*. Mostrar que H é 2-crescente é elementar, porém trabalhoso. Além disso, H tem marginais dadas por,

$$\begin{aligned} G(y) = \lim_{x\to\infty} H(x,y) &= 1 - e^{-y}, \quad y \geq 0, \\ F(x) = \lim_{y\to\infty} H(x,y) &= \lim_{y\to\infty} \frac{(1-e^{-y})x}{(x-5)e^{-y}+5} = \frac{x}{5}, \quad \text{se } 0 \leq x \leq 5. \end{aligned}$$

Além disso, se $x > 5$, $F(x) = \lim_{y\to\infty} 1 - e^{-y} = 1$ e 0, se $x < 0$. Resumindo,

$$F(x) = \begin{cases} 0, & \text{se} \quad x < 0, \\ x/5, & \text{se} \quad 0 \leq x \leq 5, \\ 1, & \text{se} \quad x > 5. \end{cases}$$

Das expressões para as marginais fica claro que $H(+\infty, +\infty) = 1$.

Teorema 1.3.1. *(Teorema de Sklar). Sejam X e Y duas variáveis aleatórias com função de distribuição F e G, respectivamente, e distribuição conjunta H. Então, existe uma cópula C tal que,*

$$H(x, y) = C(F(x), G(y)), \tag{1.14}$$

para todo $x, y \in \mathbb{R}$.

Se F e G são contínuas, então C é única; caso contrário, C é unicamente determinada em $Im(F) \times Im(G)$. Reciprocamente, se C é uma cópula e F e G são funções de distribuição, então a função H, definida em (1.14), é uma função de distribuição conjunta com marginais F e G.

Prova: A prova do teorema de Sklar é bastante técnica e pode ser encontrada no original de Sklar (1959). Provas mais simples podem ser encontradas em Schweizer e Sklar (2005), para o caso m-dimensional, e em Nelsen (1999), página 18, para o caso bidimensional. □

O Teorema de Sklar nos fornece uma maneira de expressar uma função de distribuição conjunta em função de sua cópula e de suas marginais. Note que, dada duas funções de distribuição, variando a cópula na equação (1.14) podemos obter uma infinidade de funções de distribuição conjuntas cujas marginais são F e G.

É comum encontrarmos nos livros básicos de estatística um exercício solicitando ao leitor que este construa uma função de distribuição conjunta que não seja normal bivariada e que possua marginais normais. Com o auxílio de uma cópula e do Teorema de Sklar, isto se torna um exercício fácil, como mostra o próximo exemplo.

Exemplo 1.3.3. Seja Φ a função de distribuição Gaussiana do Exemplo 1.3.1 e seja C a cópula do Exemplo 1.2.4, isto é,

$$\begin{aligned} F(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2} dt \\ C(u, v) &= \alpha M(u, v) + (1 - \alpha) W(u, v), \quad \text{com } \alpha \in I. \end{aligned}$$

Pelo Teorema de Sklar, a função,

$$\begin{aligned} H(x, y) &= \alpha \min\left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2} dt, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2} dt \right\} + \\ &\quad + (1 - \alpha) \max\left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2} dt + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2} dt - 1, 0 \right\}, \end{aligned}$$

é uma função de distribuição conjunta que não é normal bivariada e cujas marginais são $\Phi(x)$ e $\Phi(y)$. De fato, para encontrar uma função de distribuição bivariada que não seja Gaussiana e possua marginais normais, basta utilizar a mesma idéia anterior com uma cópula qualquer que não a cópula Gaussiana (veja o Exemplo 1.5.3).

Vamos introduzir agora o conceito de "quase-inversa"ou "inversa generalizada"de funções de distribuição, conceito este que utilizaremos para deduzir uma forma inversa da equação (1.14) que inclua o caso em que as marginais não sejam estritamente crescentes. Utilizaremos a notação compacta

$$\inf_{x\in\mathbb{R}}\{\cdot\} = \inf\{x \in \mathbb{R} | \cdot \}.$$

Analogamente para o supremo, máximo e mínimo.

Definição 1.3.4. Seja $F : \mathbb{R} \to I$ uma função de distribuição. Então, a *quase-inversa* ou *inversa generalizada* de F, $F^{(-1)} : I \to \mathbb{R}$ é definida como

$$F^{(-1)}(t) = \begin{cases} \sup\limits_{x\in\mathbb{R}}\{F(x) = 0\}, & t = 0, \\ \inf\limits_{x\in\mathbb{R}}\{F(x) \geq t\}, & 0 < t \leq 1. \end{cases}$$

A quase-inversa de uma função de distribuição possui as seguintes propriedades.

Lema 1.3.1. *Seja F uma função de distribuição e seja $F^{(-1)}$ sua quase-inversa. Então,*

(i). *se F é contínua (não é contínua), $F^{(-1)}$ é crescente (não-decrescente) e contínua à direita em I;*

(ii). *$F(F^{(-1)}(y)) \geq y$, para todo $y \in (0,1)$;*

(iii). *$F^{(-1)}(F(x)) \leq x$, $x \in \mathbb{R}$;*

(iv). *se F é contínua em $F^{(-1)}(y)$, $y \in (0,1)$, $F(F^{(-1)}(y)) = y$;*

(v). *se $F^{(-1)}$ é contínua em F(x), $x \in \mathbb{R}$ e $F(x) \in (0,1)$, $F^{(-1)}(F(x)) = x$.*

Prova:
(i). Sejam $u, v \in I$ tal que $u < v$ e defina

$$\begin{aligned} x_u &:= \inf_{x\in\mathbb{R}}\{F(x) \geq u\}, \\ x_v &:= \inf_{x\in\mathbb{R}}\{F(x) \geq v\}. \end{aligned}$$

Então,
$\diamond$ se F for contínua, x_u e x_v são tais que $x_u < x_v$, ou seja,

$$u < v \Longrightarrow F^{(-1)}(u) < F^{(-1)}(v);$$

$\diamond$ se F não for contínua, x_u e x_v são tais que $x_u \leq x_v$, ou seja,

$$u < v \Longrightarrow F^{(-1)}(u) \leq F^{(-1)}(v).$$

Seja $a \in (0,1]$. Como F é uma função de distribuição, é contínua à direita, então,

$$\lim_{x \to a^+} F^{(-1)}(x) = \lim_{x \to a^+} \inf_{y \in \mathbb{R}} \{F(y) \geq x\} = \inf_{y \in \mathbb{R}} \{F(y) \geq a\} = F^{(-1)}(a).$$

(ii). Seja $y \in (0,1)$. Então, definindo $F^{(-1)}(y) := x_y := \inf_{x \in \mathbb{R}} \{F(x) \geq y\}$, temos,

$$F \circ F^{(-1)}(y) = F(x_y) \geq y,$$

por definição de x_y.

(iii). $F^{(-1)}(F(x)) = \inf_{y \in \mathbb{R}} \{F(y) \geq F(x)\} \leq x$, pois pode-se ter $F(y) = F(x)$ com $y < x$, já que F é não-decrescente.

(iv). Se F é contínua em $u = F^{(-1)}(y)$ e $y \in (0,1)$, então,

$$\lim_{x \to u^+} F(x) = \lim_{x \to u^-} F(x) = F(u) = F(F^{(-1)}(y)) = y.$$

(v). Se $F^{(-1)}$ é contínua em $F(x)$, $x \in \mathbb{R}$ e $F(x) \in (0,1)$, então,

$$F^{(-1)}(F(x)) = \lim_{u \to F(x)^+} F^{(-1)}(u) = \lim_{u \to F(x)^+} \inf_{y \in \mathbb{R}} \{F(y) \geq u\} \geq x.$$

Combinando a desigualdade acima com o item (iii) deste lema, segue a igualdade $F^{(-1)}(F(x)) = x$. $\square$

Usando o conceito de quase-inversa obtemos o seguinte corolário do Teorema de Sklar.

Corolário 1.3.1. *Seja H uma função de distribuição conjunta contínua com marginais F e G. Sejam $F^{(-1)}$ e $G^{(-1)}$ as quase-inversas de F e G, respectivamente, e C a cópula associada a H, F e G como no Teorema de Sklar. Então, para todo $(u,v) \in I^2$,*

$$C(u,v) = H\big(F^{(-1)}(u), G^{(-1)}(v)\big). \tag{1.15}$$

Prova: Basta escrever $u = F(x) \Rightarrow x = F^{(-1)}(u)$ e $v = G(y) \Rightarrow y = G^{(-1)}(v)$ na equação (1.14) e o resultado segue. $\square$

Este resultado nos fornece um método para a construção de cópulas a partir de distribuições conjuntas que será usado mais adiante (no Capítulo 3). O próximo exemplo serve para ilustrar o método.

Exemplo 1.3.4. Seja H a função de distribuição do Exemplo 1.3.2:

$$H(x,y) = \begin{cases} \dfrac{(1-e^{-y})x}{(x-5)e^{-y}+5}, & \text{se} \quad (x,y) \in [0,5] \times [0,\infty), \\ 1-e^{-y}, & \text{se} \quad (x,y) \in [5,\infty) \times [0,\infty), \\ 0, & \text{c.c.} \end{cases}$$

cujas marginais são,

$$F(x) = \begin{cases} 0, & \text{se} \quad x < 0, \\ x/5, & \text{se} \quad 0 \le x \le 5, \\ 1, & \text{se} \quad x > 5. \end{cases} \quad \text{e} \quad G(y) = \begin{cases} 1-e^{-y}, & \text{se} \quad y \ge 0, \\ 0, & \text{se} \quad y < 0. \end{cases}$$

As quase-inversas de F e G são dadas por $F^{(-1)}(u) = 5u$ e $G^{(-1)}(v) = -\ln(1-v)$, para todo $u, v \in I^2$. Substituindo isso na equação (1.15) obtemos a cópula,

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

Uma das propriedades que torna as cópulas interessantes é sua invariância por transformações estritamente crescentes, como mostra o próximo teorema.

Teorema 1.3.2. *Sejam X e Y variáveis aleatórias contínuas com cópula $C_{X,Y}$. Se f e g são funções estritamente crescentes em quase toda a parte em Im(X) e Im(Y), respectivamente, então $C_{f(X),g(Y)} = C_{X,Y}$.*

Prova: Sejam F_1, G_1, F_2 e G_2 as funções de distribuição de X, Y, $f(X)$ e $g(Y)$, respectivamente. Como f e g são contínuas e crescentes em quase toda a parte, $F_2(x) = P(f(X) \le x) = P(X \le f^{-1}(x)) = F_1(f^{-1}(x))$. Analogamente, obtém-se $G_2(y) = G_1(g^{-1}(y))$.
Assim, para $x, y \in \mathbb{R}$,

$$\begin{aligned} C_{f(X),g(Y)}\big(F_2(x), G_2(y)\big) &= P\big(f(X) \le x, g(Y) \le y\big) \\ &= P\big(X \le f^{-1}(x), Y \le g^{-1}(y)\big) \\ &= C_{X,Y}\big(F_1(f^{-1}(x)), G_1(g^{-1}(y))\big) \\ &= C_{X,Y}\big(F_2(x), G_2(y)\big), \end{aligned}$$

em quase toda a parte. Como X e Y são contínuas e como f e g são monótonas em quase toda a parte, então $f(X)$ e $g(Y)$ são contínuas também e $\text{Im}(F_2) = \text{Im}(G_2) = I$. Logo, $C_{f(X),g(Y)} = C_{X,Y}$ em quase toda a parte de I^2. $\square$

Para o caso em que pelo menos uma das transformações é estritamente decrescente em quase toda a parte, obtemos resultados em que a cópula de $f(X)$ e $g(Y)$ varia de uma forma bastante simples e previsível. Este é o conteúdo do próximo teorema.

Teorema 1.3.3. *Sejam X e Y variáveis aleatórias contínuas com cópula $C_{X,Y}$. Sejam f e g transformações estritamente monótonas em quase toda a parte em Im(X) e Im(Y), respectivamente.*

(i). *Se f é estritamente decrescente em quase toda a parte e g é estritamente crescente em quase toda a parte, então,*

$$C_{f(X),g(Y)}(u,v) = v - C_{X,Y}(1-u,v);$$

(ii). *Se f é estritamente crescente em quase toda a parte e g é estritamente decrescente em quase toda a parte, então,*

$$C_{f(X),g(Y)}(u,v) = u - C_{X,Y}(u,1-v);$$

(iii). *Se f e g são estritamente decrescentes em quase toda a parte, então,*

$$C_{f(X),g(Y)}(u,v) = u + v - 1 + C_{X,Y}(1-u,1-v).$$

Prova: (i). Sejam f uma função estritamente decrescente e g uma função estritamente crescente (quase toda a parte). Sejam F_1, G_1, F_2 e G_2 as distribuições de X, Y, $f(X)$ e $g(Y)$, respectivamente. Observe que

$$\begin{aligned}
C_{f(X),g(Y)}\big(F_2(x), G_2(y)\big) &= P\big(f(X) \leq x, g(Y) \leq y\big) \\
&= P\big(X > f^{-1}(x), Y \leq g^{-1}(y)\big) \\
&= P\big(X > f^{-1}(x) | Y \leq g^{-1}(y)\big)\, P\big(Y \leq g^{-1}(y)\big) \\
&= \big[1 - P\big(X \leq f^{-1}(x) | Y \leq g^{-1}(y)\big)\big] P\big(Y \leq g^{-1}(y)\big) \\
&= P\left(Y \leq g^{-1}(y)\right) - P\big(X \leq f^{-1}(x), Y \leq g^{-1}(y)\big) \\
&= G_1(g^{-1}(y)) - C_{X,Y}\big(F_1(f^{-1}(x)), G_1(g^{-1}(y))\big)\,.
\end{aligned}$$

Mas,

$$F_1(f^{-1}(x)) = P(X \leq f^{-1}(x)) = P(f(X) \geq x) = 1 - F_2(x),$$

e

$$G_1(g^{-1}(y)) = P(Y \leq g^{-1}(y)) = P\left(g(Y) \leq y\right) = G_2(y).$$

Logo,

$$C_{f(X),g(Y)}\big(F_2(x), G_2(y)\big) = G_2(y) - C_{X,Y}\big(1 - F_2(x), G_2(y)\big), \quad x, y \in \mathbb{R}.$$

Igualdade válida em quase toda a parte de $\mathbb{R}^2$. Como F_2 e G_2 são contínuas, $\mathrm{Im}(F_2) = \mathrm{Im}(G_2) = I$ e, portanto, em quase toda a parte de I^2,

$$C_{f(X),g(Y)}(u, v) = u - C_{X,Y}(u, 1 - v).$$

(ii). A prova do item (ii) é análoga à anterior.

(iii). Suponha agora que f e g são estritamente decrescentes em quase toda a parte. Então,

$$\begin{aligned}
&C_{f(X),g(Y)}\big(F_2(x), G_2(y)\big) = \\
&\quad = P\big(f(X) \leq x, g(Y) \leq y\big) \\
&\quad = P\left(X \geq f^{-1}(x), Y \geq g^{-1}(y)\right) \\
&\quad = P\left(X \geq f^{-1}(x) | Y \geq g^{-1}(y)\right) P\left(Y \geq g^{-1}(y)\right) \\
&\quad = P\big(Y \geq g^{-1}(y)\big) - P\left(X \leq f^{-1}(x), Y \geq g^{-1}(y)\right) \\
&\quad = 1 - G_1(g^{-1}(y)) - P\left(Y \geq g^{-1}(y) | X \leq f^{-1}(x)\right) P\left(X \leq f^{-1}(x)\right) \\
&\quad = 1 - G_1(g^{-1}(y)) - P\big(X \leq f^{-1}(x)\big) + P\big(X \leq f^{-1}(x), Y \leq g^{-1}(y)\big) \\
&\quad = 1 - G_1(g^{-1}(y)) - F_1(f^{-1}(x)) + C_{X,Y}\left(F_1(f^{-1}(x)), G_1(g^{-1}(y))\right). \quad (1.16)
\end{aligned}$$

Mas,

$$F_1(f^{-1}(x)) = P(X \leq f^{-1}(x)) = P(f(X) \geq x) = 1 - F_2(x),$$

e

$$G_1(g^{-1}(y)) = P(Y \leq g^{-1}(y)) = P\left(g(Y) \geq y\right) = 1 - G_2(y).$$

Substituindo na equação (1.16), obtemos,

$$C_{f(X),g(Y)}\big(F_2(x), G_2(y)\big) = F_2(x) + G_2(y) - 1 + C_{X,Y}\big(1 - F_2(x), 1 - G_2(y)\big).$$

Igualdade válida em quase toda a parte de $\mathbb{R}^2$. Como F_2 e G_2 são contínuas, $\mathrm{Im}(F_2) = \mathrm{Im}(G_2) = I$ e, portanto,

$$C_{f(X),g(Y)}(u, v) = u + v - 1 + C_{X,Y}(1 - u, 1 - v),$$

o que completa a prova. □

Note que, basta que f e g sejam contínuas e monótonas, para os Teoremas 1.3.2 e 1.3.3 valerem.

1.4 Transformada Integral de Probabilidade

Vamos trabalhar agora no sentido de caracterizar a cópula independência (Π) e os limites de Fréchet-Hoeffding (cópulas M e W), ligando essas cópulas ao tipo de dependência a que cada uma está associada. Iniciemos com alguns conceitos e resultados necessários, como a conhecida *Transformada Integral de Probabilidade* que, como veremos adiante, também é uma ferramenta essencial na teoria de cópulas.

Teorema 1.4.1. *(Transformada Integral de Probabilidade). Seja X uma variável aleatória qualquer com distribuição F e seja $U \sim \mathcal{U}([0,1])$. Então,*

(i). *se X é contínua, temos $F(X) \sim \mathcal{U}([0,1])$;*

(ii). $F^{(-1)}(U) \sim X$;

(iii). $F^{(-1)}(1-U) \sim X$.

Prova: Suponha, primeiramente, que X seja uma variável aleatória contínua.
(i). Como F é não-decrescente e $F^{(-1)}$ é crescente, segue que,

$$P(F(X) \leq x) = P(X \leq F^{(-1)}(x)) = F(F^{(-1)}(x)) = x,$$

qualquer que seja $x \in \mathbb{R}$.
(ii). Segue do item (i) que,

$$P(F^{(-1)}(U) \leq x) = P(U \leq F(x)) = F(x) = P(X \leq x),$$

para todo $x \in \mathbb{R}$.
(iii). Segue do item (i) que,

$$\begin{aligned} P(F^{(-1)}(1-U) \leq x) &= P(U \geq 1 - F(x)) \\ &= 1 - P(U \leq 1 - F(x)) \\ &= 1 - (1 - F(x)) \\ &= P(X \leq x), \end{aligned}$$

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

Para o caso em que X é uma variável aleatória discreta, vamos provar apenas o item (ii), visto que a construção para o caso (iii) é muito similar. A construção a seguir pode ser encontrada em Rohatgi (1976), página 203. Suponha, então, que F não é contínua. Defina $P(X = x_k) := p_k,\ k \in \mathbb{N}$. Defina h como segue,

$$h(x) = \begin{cases} x_1, & \text{se} \quad 0 \leq x < p_1, \\ x_2, & \text{se} \quad p_1 \leq x < p_1 + p_2, \\ \vdots & \qquad\qquad \vdots \end{cases}$$

Então,

$$P(h(X) = x_1) = P(0 \leq X < p_1) = p_1,$$
$$P(h(X) = x_2) = P(p_1 \leq X < p_1 + p_2) = p_2,$$

e, em geral,

$$P(h(X) = x_k) = p_k, \quad k \in \mathbb{N}.$$

Logo, $h(X)$ é uma variável aleatória discreta com função de distribuição F. $\square$

Iniciaremos caracterizando a estrutura de dependência induzida pela cópula Π, que recebe o nome de *cópula independência* em virtude do próximo teorema.

Teorema 1.4.2. *Sejam X e Y duas variáveis aleatórias contínuas. Então, X e Y são independentes se, e somente se, $C_{X,Y} = \Pi$.*

Prova: Sejam F e G as funções de distribuição de X e Y, respectivamente, H a distribuição conjunta de X e Y e C a cópula associada. Suponha que X e Y são independentes. Isso quer dizer que $H(x,y) = F(x)G(y)$, para todo $(x,y) \in \mathbb{R}^2$. Então, por (1.15),

$$C(u,v) = H\big(F^{(-1)}(u), G^{(-1)}(v)\big) = F\big(F^{(-1)}(u)\big)\, G\big(G^{(-1)}(v)\big) = uv, \quad u,v \in I.$$

Reciprocamente, suponha que X e Y tenham cópula Π. Então, pelo Teorema de Sklar,

$$H(x,y) = \Pi\big(F(x), G(y)\big) = F(x)G(y).$$

Logo, X e Y são independentes. $\square$

No próximo teorema, devido a Wang e Dhaene (1998), tratamos da estrutura de dependência induzida pela cópula M. Nossa prova é ligeiramente diferente da original, principalmente no que se refere a recíproca. Por $\stackrel{d}{=}$ queremos dizer igualdade em distribuição.

Teorema 1.4.3. *Sejam X e Y duas variáveis aleatórias contínuas com função de distribuição F e G, respectivamente, e distribuição conjunta H. Então, X e Y têm cópula M se, e somente se, existem duas funções crescentes f e g e uma variável aleatória Z tais que $(X,Y) \stackrel{d}{=} (f(Z), g(Z))$.*

Prova: Suponha que existem duas funções crescentes f e g e uma variável aleatória Z tais que $(X,Y) \stackrel{d}{=} (f(Z), g(Z))$, então,

$$H(x,y) = P(f(Z) \leq x, g(Z) \leq y) = P(Z \in A, Z \in B),$$

onde A é um intervalo da forma $[0,x]$ ou $[0,x)$ e B é um intervalo da forma $[0,y]$ ou $[0,y)$. Como $A \subseteq B$ ou $B \subseteq A$, temos

$$\begin{aligned} H(x,y) &= P(Z \in A, Z \in B) = \min\{P(Z \in A), P(Z \in B)\} \\ &= \min\{P(X \le x), P(Y \le y)\} \\ &= \min\{F(x), G(y)\}, \end{aligned}$$

para todo $x, y \in \mathbb{R}$. Logo, pelo Teorema de Sklar, $C(u,v) = \min\{u,v\}$. Reciprocamente, assuma que X e Y tenham cópula M. Seja U uma variável aleatória com distribuição $\mathcal{U}([0,1])$ e seja F_U sua função de distribuição. Então,

$$\begin{aligned} P(X \le x, Y \le y) &= \min\{F(x), G(y)\} \\ &= \min\{F_U(F(x)), F_U(G(y))\} \\ &= F_U(\min\{F(x), G(y)\}) \\ &= P(U \le \min\{F(x), G(y)\}) \\ &= P(U \le F(x), U \le G(y)) \\ &= P(F^{(-1)}(U) \le x, G^{(-1)}(U) \le y), \end{aligned}$$

para todo $x, y \in \mathbb{R}$. Logo, $(X,Y) \stackrel{d}{=} (F^{(-1)}(U), G^{(-1)}(U))$. Assim, definindo $Z = U$, $f = F^{(-1)}$ e $g = G^{(-1)}$, segue que $(X,Y) \stackrel{d}{=} (f(Z), g(Z))$ e, pelo Lema 1.3.1, f e g são crescentes, o que completa a prova. $\square$

Usando o Teorema 1.4.3, obtemos um resultado que mostra que duas variáveis aleatórias X e Y têm cópula M se, e somente se, entre elas existe uma dependência perfeita, ou seja, se uma for função crescente da outra. Este é o conteúdo do próximo teorema. O enunciado pode ser encontrado em Embrechts et al. (2002), porém nenhuma prova é dada pelos autores. O mesmo ocorre com o Teorema 1.4.6.

Teorema 1.4.4. *Sejam X e Y duas variáveis aleatórias contínuas com função de distribuição F e G, respectivamente. Então, X e Y têm cópula M se, e somente se, existe uma função crescente h tal que $Y \stackrel{d}{=} h(X)$ quase certamente.*

Prova: Suponha que a cópula associada a X e Y seja M. Então o Teorema 1.4.3 garante a existência de uma variável aleatória Z e duas funções crescentes, f e g, tais que $(X,Y) \stackrel{d}{=} (f(Z), g(Z))$. Defina $h := g \circ f^{-1}$. Então,

$$h(X) = (g \circ f^{-1})(X) \stackrel{d}{=} g(f^{-1}(X)) \stackrel{d}{=} g(Z) \stackrel{d}{=} Y.$$

Além disso, como f e g são crescentes, segue que h é crescente. Reciprocamente, suponha que existe uma função h crescente tal que $Y \stackrel{d}{=} h(X)$.

Tomando $f = I$, $g = h$ e $Z = X$, onde I é a função identidade, temos que existem duas funções crescentes f e g e uma variável aleatória Z tais que $(X, Y) \stackrel{d}{=} (f(Z), g(Z))$. Então, pelo Teorema 1.4.3 a cópula de X e Y é M e isto completa a prova. □

O próximo teorema caracteriza a relação entre X e Y mediante a cópula W. O teorema é devido a Embrechts et al. (2002).

Teorema 1.4.5. *Sejam X e Y duas variáveis aleatórias contínuas com função de distribuição F e G, respectivamente, e distribuição conjunta H. Então, X e Y têm cópula W se, e somente se, existem uma variável aleatória Z, uma função crescente f e uma função decrescente g tais que $(X, Y) \stackrel{d}{=} (f(Z), g(Z))$.*

Prova: Seja U uma variável aleatória com distribuição $\mathcal{U}([0, 1])$ e suponha que X e Y têm cópula W. Note primeiramente que,

$$\begin{aligned} P(U \leq F(x), U \geq 1 - G(y)) &= P\big(U \geq 1 - G(y) | U \leq F(x)\big) P(U \leq F(x)) \\ &= \big[1 - P\big(U \leq 1 - G(y) | U \leq F(x)\big)\big] P(U \leq F(x)) \\ &= P(U \leq F(x)) - P(U \leq F(x), U \leq 1 - G(y)). \end{aligned} \tag{1.17}$$

Então,

$$\begin{aligned} P(X \leq x, Y \leq y) &= \max\{0, F(x) + G(y) - 1\} \\ &= -\min\{0, 1 - F(x) - G(y)\} \\ &= F(x) - (F(x) + \min\{0, 1 - F(x) - G(y)\}) \\ &= F(x) - \min\{F(x), 1 - G(y)\} \\ &= P(U \leq F(x)) - P(U \leq F(x), U \leq 1 - G(y)) \\ &=^{1} P(U \leq F(x), U \geq 1 - G(y)) \\ &= P(U \leq F(x), 1 - U \leq G(y)) \\ &= P(F^{(-1)}(U) \leq x, G^{(-1)}(1 - U) \leq y), \end{aligned}$$

Isso mostra que $(X, Y) \stackrel{d}{=} (F^{(-1)}(U), G^{(-1)}(1 - U))$. Definindo $Z = U$, $f := F^{(-1)}$ e $g := G^{(-1)} \circ (1 - \mathrm{I})$, obtemos $(X, Y) \stackrel{d}{=} (f(Z), g(Z))$. Além disso, pelo Lema 1.3.1, $F^{(-1)}$ e $G^{(-1)}$ são funções crescentes e $h = 1 - \mathrm{I}$ é função decrescente, segue que f é crescente e g é decrescente.

Reciprocamente, se $(X, Y) \stackrel{d}{=} (f(Z), g(Z))$ com f crescente e g decrescente, defina $A := \{Z \in f^{-1}((-\infty, x])\}$ e $B := \{Z \in g^{-1}((-\infty, y])\}$. Note primeiramente que, $f^{-1}((-\infty, x])$ é um intervalo da forma $(-\infty, f^{-1}(x)]$ pois f é uma função contínua e crescente, e como g é uma função contínua e

[1] Por (1.17).

decrescente, $g^{-1}((-\infty, y])$ é um intervalo da forma $[\,g^{-1}(y), \infty)$. Assim,

$$\begin{aligned} P(A) &= P(Z \in f^{-1}((-\infty, x])) = P(Z \leq f^{-1}(x)) \\ &= P(f(Z) \leq x) = P(X \leq x) \\ &= F(x). \end{aligned}$$

Por outro lado, como g é função decrescente,

$$\begin{aligned} P(B) &= P(Z \in g^{-1}((-\infty, y])) = P(Z \geq g^{-1}(y)) \\ &= P(g(Z) \leq y) = P(Y \leq y) \\ &= G(y). \end{aligned}$$

Analogamente mostra-se que $P(A \cap B) = P(X \leq x, Y \leq y)$.
Se $A \cap B \neq \emptyset$ então, pela definição de A e B,

$$1 = P(\mathbb{R}) = P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

e assim,

$$\begin{aligned} P(X \leq x, Y \leq y) = P(A \cap B) &= P(A) + P(B) - 1 \\ &= F(x) + G(y) - 1. \end{aligned}$$

Se $A \cap B = \emptyset$ então,

$$P(X \leq x, Y \leq y) = P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0,$$

De qualquer forma temos $P(X \leq x, Y \leq y) = \max\{F(x) + G(y) - 1, 0\} = W(F(x), G(y))$, para todo $x, y \in \mathbb{R}$. Logo, pelo Teorema de Sklar, W é a cópula de X e Y. $\square$

Usando o Teorema 1.4.6, a seguir obtemos um resultado que caracteriza duas variáveis aleatórias X e Y quando a cópula associada é a cópula W. Isto ocorre se, e somente se, X e Y estão em dependência perfeita negativa, ou seja, uma é função decrescente da outra. Este teorema pode ser encontrado em Embrechts et al. (2002), porém nenhuma prova é dada pelos autores.

Teorema 1.4.6. *Sejam X e Y duas variáveis aleatórias contínuas com função de distribuição F e G, respectivamente. Então, X e Y têm cópula W se, e somente se, existe uma função decrescente h, tal que $Y \stackrel{d}{=} h(X)$ quase certamente.*

Prova: Suponha primeiramente que X e Y têm cópula W. Então, pelo Teorema 1.4.5, existem uma variável aleatória Z, uma função crescente f e uma função decrescente g tais que $(X, Y) \stackrel{d}{=} (f(Z), g(Z))$. Observe que

$Z \stackrel{d}{=} f^{-1}(X)$. Defina $h := g \circ f^{-1}$. Note que h assim definida é uma função decrescente e tal que

$$h(X) = (g \circ f^{-1})(X) = g(f^{-1}(X)) \stackrel{d}{=} g(Z) \stackrel{d}{=} Y.$$

Reciprocamente, suponha que $Y \stackrel{d}{=} h(X)$ para alguma função decrescente h. Então,

$$F(h^{-1}(y)) = P(X \leq h^{-1}(y)) = P(h(X) \geq y) = 1 - P(Y \leq y) = 1 - G(y). \tag{1.18}$$

Por outro lado,

$$\begin{aligned} H(x,y) &= P(X \leq x, h(X) \leq y) \\ &= P(X \leq x, X \geq h^{-1}(y)) \\ &= P(X \leq x) - P(X \leq x, X \leq h^{-1}(y)) \\ &= F(x) - F(\min\{x, h^{-1}(y)\}) \\ &= F(x) - \min\{F(x), F(h^{-1}(y))\} \\ &=^{2} F(x) - \min\{F(x), 1 - G(y)\} \\ &= -\min\{F(x) - F(x), 1 - F(x) - G(y)\} \\ &= \max\{F(x) + G(y) - 1, 0\}, \end{aligned}$$

para todo x, $y \in \mathbb{R}$. Segue pelo Teorema de Sklar que a cópula associada a X e Y é W. □

1.5 Derivadas e Representação Canônica de Uma Cópula

Na introdução deste capítulo, comentamos que, equivalentemente à Definição 1.2.1, cópulas podem também ser vistas como funções de distribuições cujas marginais são $\mathcal{U}([0,1])$. Este fato segue da definição de cópulas e pela transformada integral de probabilidade. Com efeito, dadas duas variáveis aleatórias contínuas X e Y com distribuição F e G, respectivamente, distribuição conjunta H e cópula C, temos,

$$\begin{aligned} C(u,v) &= H(F^{(-1)}(u), G^{(-1)}(v)) \\ &= P(X \leq F^{(-1)}(u), Y \leq G^{(-1)}(v)) \\ &= P(F(X) \leq u, G(X) \leq v) \\ &= P(U_1 \leq u, U_2 \leq v), \end{aligned} \tag{1.19}$$

onde U_1 e U_2 são duas variáveis aleatórias com distribuição $\mathcal{U}([0,1])$.

[2] Por (1.18).

Note que não mencionamos nada a respeito da estrutura de dependência entre U_1 e U_2. Acontece que a estrutura de dependência é caracterizada pela cópula em questão, vide, por exemplo, Teoremas 1.4.2, 1.4.4 e 1.4.6. Esta é uma das qualidades mais importantes das cópulas. Podemos, a princípio, substituir o problema de caracterizar o tipo de dependência entre as variáveis aleatórias, pelo problema, aparentemente mais fácil, de caracterizar sua cópula. Este problema será estudado mais adiante, no capítulo 4.

Usando a definição de probabilidade condicional e a equação (1.19), as seguintes propriedades para duas variáveis aleatórias U_1 e U_2 com distribuição $\mathcal{U}([0,1])$ são de verificação imediata:

(i). $P(U_1 \leq u, U_2 \geq v) = u - C(u,v)$;

(ii). $P(U_1 \geq u, U_2 \leq v) = v - C(u,v)$;

(iii). $P(U_1 \leq u | U_2 \leq v) = \dfrac{C(u,v)}{v}$;

(iv). $P(U_1 \leq u | U_2 \geq v) = \dfrac{v - C(u,v)}{1-v}$.

Note que, para obter resultados para duas variáveis aleatórias X e Y contínuas quaisquer, com distribuição F e G, respectivamente, basta usar a transformada integral de probabilidade e substituir $U_1 \stackrel{d}{=} F(X)$ e $U_2 \stackrel{d}{=} G(Y)$ acima.

Dada uma cópula C, o Teorema 1.2.2 garante a existência das derivadas parciais em quase toda a parte. A próxima proposição nos dá uma relação entre as derivadas parciais de uma cópula e as probabilidades condicionais das variáveis aleatórias em questão.

Proposição 1.5.1. *Sejam U_1 e U_2 duas variáveis aleatórias com distribuição $\mathcal{U}([0,1])$. Então,*

(i). $C_v(u,v) := \dfrac{\partial C(u,v)}{\partial v} = P(U_1 \leq u | U_2 = v)$;

(ii). $C_u(u,v) := \dfrac{\partial C(u,v)}{\partial u} = P(U_2 \leq v | U_1 = u)$,

onde
$$P(U_1 \leq u | U_2 = v) = \lim_{h \to 0^+} P(U_1 \leq u | v \leq U_2 \leq v + h).$$

Prova: (i). Sejam U_1 e U_2 variáveis aleatórias com distribuição $\mathcal{U}([0,1])$. Seja H sua distribuição conjunta com marginais F e G, respectivamente. Note primeiramente que $H(u,v) = C\big(F(u), G(v)\big) = C(u,v)$, onde a segunda igualdade decorre do fato de U_1 e U_2 serem uniformemente distribuídas. Assim,

$$\begin{aligned}
P(U_1 \leq u | U_2 = v) &= \lim_{h \to 0^+} P(U_1 \leq u | v \leq U_2 \leq v + h) \\
&= \lim_{h \to 0^+} \frac{P(U_1 \leq u, v \leq U_2 \leq v + h)}{P(v \leq U_2 \leq v + h))} \\
&= \lim_{h \to 0^+} \frac{H(u, v + h) - H(u, v)}{P(U_2 \leq v + h) - P(U_2 \leq v)} \\
&= \lim_{h \to 0^+} \frac{C(u, v + h) - C(u, v)}{h} \\
&= \frac{\partial C(u, v)}{\partial v},
\end{aligned}$$

onde a derivada existir. A prova do item (ii) é análoga. □

Novamente, se quisermos resultados para duas variáveis aleatórias contínuas X e Y, pela transformada integral de probabilidade, basta tomarmos $U_1 \stackrel{d}{=} F(X)$ e $U_2 \stackrel{d}{=} G(Y)$ na Proposição 1.5.1.

Exemplo 1.5.1. (*Família de Cópulas Ali-Mikhail-Haq*). Para $\theta \in [-1, 1]$, considere a função

$$C_\theta(u, v) = \frac{uv}{1 - \theta(1 - u)(1 - v)}.$$

Esta função é de fato uma cópula e sua construção será feita na Seção 2.1 do Capítulo 2. Sejam X uma variável aleatória com distribuição exponencial com parâmetro $\lambda = 1$, Y uma variável aleatória com distribuição Cauchy padrão e $H(x, y)$ a distribuição conjunta de X e Y. Assim,

$$F(x) = 1 - e^{-x}, \quad x \geq 0, \qquad \text{e} \qquad G(y) = \frac{1}{2} + \frac{\arctan(y)}{\pi}, \quad y \in \mathbb{R}.$$

Suponhamos que X e Y tenham cópula Ali-Mikhail-Haq. Assim,

$$\begin{aligned}
H(x, y) = C_\theta(F(x), G(y)) &= \frac{(1 - e^{-x})(\frac{1}{2} + \frac{\arctan(y)}{\pi})}{1 - \theta e^{-x}(\frac{1}{2} - \frac{\arctan(y)}{\pi})} \\
&= \frac{(e^{-x} - 1)(\pi + 2\arctan(y))}{\theta e^{-x}(\pi - 2\arctan(y)) - 2\pi},
\end{aligned}$$

para todo $(x, y) \in \mathbb{R}^+ \times \mathbb{R}$. As derivadas parciais de C_θ são dadas por

$$\frac{\partial C_\theta(u, v)}{\partial u} = \frac{v[1 - \theta(v - 1)]}{[1 - \theta(1 - u)(1 - v)]^2},$$

e

$$\frac{\partial C_\theta(u, v)}{\partial v} = \frac{u[1 - \theta(u - 1)]}{[1 - \theta(1 - u)(1 - v)]^2}.$$

Pela Proposição 1.5.1 temos,

$$\begin{aligned}
P(Y \leq y|X = x) &= \left.\frac{\partial C_\theta(u,v)}{\partial u}\right|_{(F(x),G(y))} \\
&= \frac{(\pi + 2\arctan(y))[2\pi + \theta(\pi - 2\arctan(y))]}{[2\pi - \theta e^{-x}(\pi - 2\arctan(y))]^2} \\
&= \frac{\pi^2(\theta+2) + 4\arctan(y)[\pi - \theta\arctan(y)]}{[2\pi - \theta e^{-x}(\pi - 2\arctan(y))]^2},
\end{aligned}$$

$$\begin{aligned}
P(X \leq x|Y = y) &= \left.\frac{\partial C_\theta(u,v)}{\partial v}\right|_{(F(x),G(y))} \\
&= \frac{4\pi^2(1 - e^{-x})(1 + \theta e^{-x})}{[2\pi - \theta e^{-x}(\pi - 2\arctan(y))]^2}.
\end{aligned}$$

Exemplo 1.5.2. (*Família de Cópulas Plackett*). Considere a seguinte cópula:

$$C_\theta(u,v) = \frac{[1 + (\theta-1)(u+v)] - \sqrt{[1 + (\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}}{2(\theta-1)}, \tag{1.20}$$

com $\theta > 0, \theta \neq 1$ e $C_1(u,v) = uv$.

Esta é a conhecida família de cópulas Plackett. A construção desta família será feita na Seção 2.2 do Capítulo 2. Para uma aplicação desta família, veja Rockinger e Jondeau (2006). Sejam U_1 e U_2 duas variáveis aleatórias $\mathcal{U}([0,1])$. Então, pela Proposição 1.5.1,

$$P(U_1 \leq u|U_2 = v) = \frac{\partial C_\theta(u,v)}{\partial v} = \frac{1}{2} + \frac{u + v - 1 + \theta(u - v)}{2\sqrt{[1 + (\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}}.$$

Observe que calcular as probabilidades condicionais acima, pelo método descrito na Proposição 1.5.1, é relativamente fácil em relação ao método tradicional, especialmente quando não é possível obter a distribuição conjunta das variáveis aleatórias. É claro que essa relativa facilidade teve um preço: exigiu conhecermos a cópula associada às variáveis aleatórias em questão.

O início do Exemplo 1.5.1 também nos dá uma idéia de como proceder quando queremos encontrar a distribuição conjunta de variáveis aleatórias cujas distribuições são conhecidas e que tenham uma certa estrutura de dependência. Note que, neste caso, a estrutura de dependência entre as variáveis aleatórias é modelada pela cópula associada.

Vamos introduzir agora os conceitos necessários para determinar a *decomposição canônica* de uma função de distribuição conjunta (veja a Proposição 1.5.2). Como cópulas são funções de distribuição, também possuem a noção de densidade.

Definição 1.5.1. Seja C uma cópula qualquer. A *densidade* associada a uma cópula C é a função $c(\cdot,\cdot)$ dada por

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v} = \frac{\partial^2 C(u,v)}{\partial v \partial u}.$$

Observe que o Teorema 1.2.2 garante a existência da densidade de uma cópula em quase toda a parte de I^2. Além disso a densidade é não-negativa. De fato, quaisquer que sejam $u, v \in I$, temos

$$\begin{aligned} \frac{\partial^2 C(u,v)}{\partial u \partial v} &= \frac{\partial}{\partial u}\left[\frac{\partial C(u,v)}{\partial v}\right] \\ &= \lim_{h \to 0^+} \frac{\frac{\partial C(u+h,v)}{\partial v} - \frac{\partial C(u,v)}{\partial v}}{h} \geq 0, \end{aligned}$$

pois, pelo Teorema 1.2.2, $\frac{\partial C(u,v)}{\partial v}$ é não-decrescente e não-negativa.

Exemplo 1.5.3. (*Cópula Gaussiana*). Seja $0 \leq \rho \leq 1$. A *cópula Gaussiana* é dada por (veja Joe, 1997, página 140 e também Cherubini et al., 2004)

$$\begin{aligned} C_\rho(u,v) &= \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)) \\ &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{2\rho xy - x^2 - y^2}{2(1-\rho^2)}\right) dxdy, \end{aligned}$$

onde $\Phi_\rho(\cdot\,,\cdot)$ é a distribuição normal bivariada padrão com coeficiente de correlação ρ e $\Phi^{-1}(\cdot)$ é a inversa (ou função quantil) de uma distribuição $\mathcal{N}(0,1)$. A densidade da cópula Gaussiana é dada por

$$c(u,v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{\xi_u^2 + \xi_v^2}{2} + \frac{2\rho\xi_u\xi_v - \xi_u^2 - \xi_v^2}{2(1-\rho^2)}\right),$$

onde $\xi_u = \Phi^{-1}(u)$ e $\xi_v = \Phi^{-1}(v)$.

Seguimos explorando as conseqüências da existência das derivadas parciais de uma cópula. Assim como as funções de distribuições, as cópulas também possuem a noção de componentes singular e absolutamente contínuas, conforme a próxima definição (ver, Nelsen 1999, página 23).

Definição 1.5.2. Seja C uma cópula. Defina

$$A_C(u,v) = \int_0^u \int_0^v c(x,y)\, dxdy.$$

Então,

$$C(u,v) = A_C(u,v) + S_C(u,v),$$

onde $S_C(u,v) = C(u,v) - A_C(u,v)$. A função $A_C(u,v)$ é chamada *componente absolutamente contínua* da cópula C e $S_C(u,v)$ é chamada *componente singular* de C.

Cada função de distribuição conjunta bivariada H induz uma única medida de probabilidade em $\mathbb{R}^2$ que satisfaz $\Delta_R H = H(u_1, v_1) + H(u_2, v_2) - H(u_1, v_2) - H(u_2, v_1)$ para todo retângulo limitado $R = [u_1, u_2] \times [v_1, v_2]$ em $\mathbb{R}^2$, desde que $\Delta_R H \geq 0$, qualquer que seja R em $\mathbb{R}^2$. Porém, pode-se provar que uma função de distribuição conjunta de duas variáveis aleatórias sempre satisfaz esta condição (veja Rohatgi, 1976, página 107 e Tucker, 1967, página 24). O leitor interessado em mais informações, pode encontrar em Billingsley (1995), páginas 176-178, Teorema 12.5.

Se usarmos o fato de que $V_H((-\infty, x] \times (-\infty, y]) = H(x, y)$ podemos, através de uma extensão padrão dos Borelianos em $\mathbb{R}^2$ (junto com o Teorema de Extensão de Kolmogorov[3]), concluir que H induz uma única medida de probabilidade em $\mathbb{R}^2$ definida pelo H-volume. Como cópulas são funções de distribuição conjunta, a condição de ser 2-crescente faz com que cada cópula induza uma (única) medida de probabilidade em I^2 através de $V_C([0, u] \times [0, v]) = C(u, v)$. Assim a *C-medida* de um conjunto R em I^2 é dada por $\mu_C(R) = V_C(R)$.

Note que, dado um intervalo $S = [u_1, u_2] \subseteq I$, temos

$$
\begin{aligned}
V_C(S \times I) &= V_C([u_1, u_2] \times [0, 1]) \\
&= C(u_1, 0) + C(u_2, 1) - C(u_1, 1) - C(u_2, 0) \\
&= u_2 - u_1 \\
&= \mu([u_1, u_2]),
\end{aligned}
$$

onde μ é a medida de Lebesgue em I. Note também que $V_C(S \times I) = \mu(S) = V_C(I \times S)$.

Definição 1.5.3. Uma cópula é dita ser *absolutamente contínua* se $C \equiv A_C$ em I^2 e é dita ser *singular* se $C \equiv S_C$ em I^2. Neste caso, teremos $c(u, v) = 0$ em quase toda a parte de I^2.

É claro que existem cópulas que não são nem absolutamente contínuas, nem singulares. Nestes casos, dizemos que C tem componentes absolutamente contínua e singular e a C-medida dessas componentes são $A_C(1, 1)$ e $S_C(1, 1)$, respectivamente.

Para provar que uma dada função é uma cópula temos que verificar as 3 exigências da Definição 1.2.1. Certamente a parte mais árdua é provar que $V_C(R) \geq 0$ qualquer que seja o retângulo $R \subseteq I^2$. No caso de cópulas absolutamente contínuas, o resultado é muitas vezes mais simples de se obter, como mostra o próximo lema.

[3] O enunciado preciso e uma prova para o Teorema de Extensão de Kolmogorov podem ser encontrados em Laha e Rohatgi 1979, página 23.

Lema 1.5.1. *Seja $C : I^2 \longrightarrow I$ uma função satisfazendo as condições de contorno (1.2) e (1.3) para cópulas. Se*

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}$$

existe, é não-negativa e tal que C pode ser escrita como

$$C(u,v) = \int_0^u \int_0^v c(u,v)\, dudv, \quad \forall\, u, v \in I,$$

então, C é 2-crescente.

Prova: Temos que provar que $V_C(R) \geq 0$ qualquer que seja o retângulo $R \subseteq I^2$. Para isso, seja $R = [u_1, v_1] \times [u_2, v_2] \subseteq I^2$, então,

$$\begin{aligned}
V_C(R) &= C(u_1,u_2) + C(v_1,v_2) - C(u_1,v_2) - C(v_1,u_2) \\
&= \int_0^{u_1}\int_0^{u_2} c(u,v)\, dudv + \int_0^{v_1}\int_0^{v_2} c(u,v)\, dudv - \int_0^{u_1}\int_0^{v_2} c(u,v)\, dudv \\
&\quad - \int_0^{v_1}\int_0^{u_2} c(u,v)\, dudv \\
&= \int_0^{u_1}\int_{v_2}^{u_2} c(u,v)\, dudv + \int_0^{v_1}\int_{u_2}^{v_2} c(u,v)\, dudv \\
&= -\int_0^{u_1}\int_{u_2}^{v_2} c(u,v)\, dudv + \int_0^{v_1}\int_{u_2}^{v_2} c(u,v)\, dudv \geq 0,
\end{aligned}$$

pois $c(u,v)$ é não-negativa e não-decrescente, $u_1 < v_1$ e pela linearidade da integral. $\square$

Ao contrário das funções de distribuição em geral, as marginais de uma cópula são contínuas, assim não existem pontos individuais em I^2 com C-medida positiva (átomos).

Exemplo 1.5.4. (*Família de Cópulas Gumbel-Bernett*). Considere a família de cópulas dada por

$$C_\theta(u,v) = uv \exp\{-\theta \ln(u) \ln(v)\}, \tag{1.21}$$

onde $\theta \in [0,1]$. Esta é a chamada família Gumbel-Bernett de cópulas. Esta família é composta apenas de cópulas absolutamente contínuas. De fato, como $C_\theta(u,v)$ dada por (1.21) pode ser reescrita como

$$C_\theta(u,v) = vu^{1-\theta \ln(v)},$$

obtemos,

$$\begin{aligned}
\frac{\partial C_\theta(u,v)}{\partial u} &= [1 - \theta \ln(v)] v u^{-\theta \ln(v)}, \\
\frac{\partial^2 C_\theta(u,v)}{\partial v \partial u} &= \{1 - \theta[1 + \ln(v) + \ln(u)(1 - \theta \ln(v))]\}\, u^{-\theta \ln(v)}.
\end{aligned}$$

Assim, a parte absolutamente contínua da cópula $C_\theta(u,v)$ é dada por

$$\begin{aligned}
A_{C_\theta}(u,v) &= \int_0^u\int_0^v c(x,y)dxdy \\
&= \int_0^u\int_0^v \big\{1-\theta\big[1+\ln(y)+\ln(x)(1-\theta\ln(y))\big]\big\}\, x^{-\theta\ln(y)}dxdy \\
&=^4 \int_0^u \big[1-\theta\ln(x)\big]x^{1-\theta\ln y}dy \\
&=^5 v\exp\big\{\big(1-\theta\,\ln(v)\big)\ln(x)\big\}\Big|_0^u \\
&= uv\exp\{-\theta\ln(u)\ln(v)\}.
\end{aligned}$$

Isto prova que C_θ é absolutamente contínua.

Assim como as funções de distribuições, as cópulas também possuem a noção de *suporte*, que é o complemento da união de todos os subconjuntos de I^2 com C-medida nula.

Definição 1.5.4. Seja C uma cópula. O *suporte* da cópula C é o complemento da união de todos os subconjuntos abertos em I^2 que possuem C-medida nula. Se o suporte da cópula C for todo I^2, dizemos que a cópula tem *suporte completo.*

Exemplo 1.5.5. Pelos cálculos do Exemplo 1.2.1, concluímos que o suporte da cópula M é a diagonal principal de I^2, já que todo o retângulo R cujos vértices estão abaixo da diagonal principal é tal que $V_H(R)=0$.

A próxima proposição é uma conseqüência direta do Teorema de Sklar para variáveis aleatórias contínuas e nos dá uma expressão para a densidade conjunta de duas variáveis aleatórias em função de suas densidades marginais e da densidade da cópula associada.

Proposição 1.5.2. *Sejam X e Y duas variáveis aleatórias contínuas com cópula C, densidade de cópula c e função de distribuição F e G, respectivamente. Sejam f, g e h as funções densidades marginais de X e Y e sua função densidade conjunta, respectivamente, e suponha que as mesmas estejam bem definidas. Então, para todo $x,y\in\mathbb{R}$,*

$$h(x,y) = c\big(F(x),G(y)\big)f(x)g(y). \tag{1.22}$$

[4] Aqui usamos a fórmula (Gradshteyn e Ryzhik, 2000, fórmula 2.723, página 234),

$$\int x^n\ln(x)dx = x^{(n+1)}\left[\frac{\ln(x)}{n+1}-\frac{1}{(n+1)^2}\right].$$

[5] Lembramos que $\displaystyle\int a^{\ln(y)}dy = \frac{y^{1+\ln(a)}}{1+\ln(a)}$.

Prova: Sejam F, G e H as funções de distribuição de X, Y e a função de distribuição conjunta de X e Y, respectivamente. Pelo Teorema de Sklar, para quaisquer $x, y \in \mathbb{R}$, temos $H(x,y) = C(F(x), G(y))$. Diferenciando $H(x,y)$ em relação a x e a y temos,

$$\begin{aligned} h(x,y) &= \frac{\partial^2 H(x,y)}{\partial x \partial y} = \frac{\partial^2 C\big(F(x), G(y)\big)}{\partial x \partial y} \\ &= c\big(F(x), G(y)\big) \begin{vmatrix} \dfrac{\partial F(x)}{\partial x} & \dfrac{\partial F(x)}{\partial y} \\ \dfrac{\partial G(y)}{\partial x} & \dfrac{\partial G(y)}{\partial y} \end{vmatrix} \\ &= c\big(F(x), G(y)\big) f(x) g(y), \end{aligned}$$

para todo $x, y \in \mathbb{R}$. □

Da equação (1.22) definimos o que muitos autores chamam de *representação canônica* (veja, por exemplo, Cherubini et al., 2004) e que desempenha um papel fundamental na estimação de máxima verossimilhança baseada em cópulas, como veremos mais adiante, no Capítulo 5.

Definição 1.5.5. Sejam X e Y duas variáveis aleatórias contínuas com cópula C, densidade de cópula c e função de distribuição F e G, respectivamente. Sejam f, g e h as funções densidades marginais de X e Y e sua função densidade conjunta, respectivamente, e suponha que as mesmas estejam bem definidas. Definimos a *representação canônica* da densidade conjunta de X e Y em relação às suas marginais e à cópula C por

$$h(x,y) = f(x) g(y) c\big(F(x), G(y)\big).$$

Exemplo 1.5.6. Sejam X e Y duas variáveis aleatórias com distribuição logística univariada padrão. Sejam F, f, G e g as funções de distribuição e densidade de X e Y, respectivamente. Isto significa que,

$$X \sim \begin{cases} F(x) &= \dfrac{1}{1+e^{-x}} \\ f(x) &= \dfrac{e^{-x}}{(1+e^{-x})^2} \end{cases}, \qquad Y \sim \begin{cases} G(y) &= \dfrac{1}{1+e^{-y}} \\ g(y) &= \dfrac{e^{-y}}{(1+e^{-y})^2} \end{cases}.$$

Suponha que X e Y tenham cópula $C(u,v) = \dfrac{uv}{u+v-uv} \cdot$ Note que C é um membro da família de cópulas Ali-Mikhail-Haq do Exemplo 1.5.1 com $\theta = 1$. Derivando em relação a u e a v obtemos,

$$c(u,v) = \frac{2uv}{(u+v-uv)^3} \cdot$$

Dessa forma, após algumas simplificações, a densidade conjunta de X e Y é dada, por,

$$\begin{aligned} h(x,y) &= c\big(F(x),G(y)\big)f(x)g(y) \\ &= \left(\frac{2(1+e^{-x})^2(1+e^{-y})^2}{(1+e^{-x}+e^{-y})^3}\right)\left(\frac{e^{-x}}{(1+e^{-x})^2}\right)\left(\frac{e^{-y}}{(1+e^{-y})^2}\right) \\ &= \frac{2e^{-x-y}}{(1+e^{-x}+e^{-y})^3}. \end{aligned}$$

Note que esta é a densidade da distribuição logística bivariada de Gumbel, ou seja, duas variáveis aleatórias com distribuição logística padrão e cópula C têm distribuição conjunta logística bivariada.

1.6 Extensões m-dimensionais

O objetivo desta seção é estender alguns dos resultados das seções precedentes para o caso m-dimensional. A maioria dos teoremas das seções anteriores possuem análogos m-dimensionais, porém nem todos. Para trabalharmos com m-cópulas precisamos estender também algumas notações. Denotaremos por I^m o m-cubo $\underbrace{[0,1]\times\cdots\times[0,1]}_{m\ vezes}$. Um retângulo em $\mathbb{R}^m$ é o produto cartesiano de m intervalos fechados $[u_1,v_1]\times\cdots\times[u_m,v_m]$.

Definição 1.6.1. Seja $R=[u_1,v_1]\times\cdots\times[u_m,v_m]$ um retângulo em $\mathbb{R}^m$. Definimos o *conjunto de vértices* de R como sendo o conjunto

$$A_m(R)=\{\mathbf{w}=(w_1,\cdots,w_m)\in R \text{ tais que } w_k=u_k \text{ ou } w_k=v_k, k=1,\cdots,m\}.$$

Exemplo 1.6.1. Seja $R=[u_1,v_1]\times[u_2,v_2]\times[u_3,v_3]$ um retângulo em $\mathbb{R}^3$. Então,

$$\begin{aligned} A_3(R) = \{&(u_1,u_2,u_3),(u_1,u_2,v_3),(u_1,v_2,u_3),(v_1,u_2,u_3), \\ &(u_1,v_2,v_3),(v_1,u_2,v_3),(v_1,v_2,u_3),(v_1,v_2,v_3)\}. \end{aligned}$$

Note que se R é um retângulo em $\mathbb{R}^m$, então o conjunto $A_m(R)$ tem 2^m elementos.

Definição 1.6.2. Seja $R=[u_1,v_1]\times\cdots\times[u_m,v_m]$ um retângulo em R^m e seja $\mathbf{w}=(w_1,\cdots,w_m)\in A_m(R)$. Definimos a *função co-sinal* de $\mathbf{w}$ como

$$\operatorname{csgn}(\mathbf{w})=\begin{cases} 1, & \text{se } w_i=u_i \text{ para um número par de } i\text{'s (ou nenhum } i), \\ -1, & \text{se } w_i=u_i \text{ para um número ímpar de } i\text{'s.} \end{cases}$$

Exemplo 1.6.2. Seja $R = [u_1, v_1] \times [u_2, v_2] \times [u_3, v_3]$ um retângulo em $\mathbb{R}^3$. Então, (veja o Exemplo 1.6.1)

$$\text{csgn}(\mathbf{w}) = \begin{cases} 1, & \text{se } \mathbf{w} \in \{(u_1, u_2, v_3), (u_1, v_2, u_3), (v_1, u_2, u_3), (v_1, v_2, v_3)\}, \\ -1, & \text{se } \mathbf{w} \in \{(u_1, u_2, u_3), (u_1, v_2, v_3), (v_1, u_2, v_3), (v_1, v_2, u_3)\}. \end{cases}$$

Definição 1.6.3. Seja H uma função de distribuição m-dimensional e $R = [u_1, v_1] \times \cdots \times [u_m, v_m]$ um retângulo em R^m. Definimos o H-*volume* de R por

$$V_H(R) = \sum_{\mathbf{w} \in A_m(R)} \text{csgn}(\mathbf{w}) H(\mathbf{w}). \tag{1.23}$$

Exemplo 1.6.3. Seja $H : \mathbb{R}^3 \longrightarrow \mathbb{R}$. Então, (veja o Exemplo 1.6.2)

$$\begin{aligned} V_H(R) &= \sum_{\mathbf{w} \in A_3(R)} \text{csgn}(\mathbf{w}) H(\mathbf{w}) \\ &= \text{csgn}(u_1, u_2, v_3)\, H(u_1, u_2, v_3) + \text{csgn}(u_1, v_2, u_3)\, H(u_1, v_2, u_3) \\ &\quad + \text{csgn}(v_1, u_2, u_3)\, H(v_1, u_2, u_3) + \text{csgn}(v_1, v_2, v_3)\, H(v_1, v_2, v_3) \\ &\quad + \text{csgn}(u_1, u_2, u_3)\, H(u_1, u_2, u_3) + \text{csgn}(u_1, v_2, v_3)\, H(u_1, v_2, v_3) \\ &\quad + \text{csgn}(v_1, u_2, v_3)\, H(v_1, u_2, v_3) + \text{csgn}(v_1, v_2, u_3)\, H(v_1, v_2, u_3) \\ &= H(u_1, u_2, v_3) + H(u_1, v_2, u_3) + H(v_1, u_2, u_3) + H(v_1, v_2, v_3) \\ &\quad - H(u_1, u_2, u_3) - H(u_1, v_2, v_3) - H(v_1, u_2, v_3) - H(v_1, v_2, u_3), \end{aligned}$$

para todo retângulo $R = [u_1, v_1] \times [u_2, v_2] \times [u_3, v_3] \subseteq \mathbb{R}^3$.

Definição 1.6.4. Uma função $H : S_1 \times \cdots \times S_m \to \mathbb{R}$, onde $S_i \subset \mathbb{R}$, para $i = 1, 2, \cdots, m$, é dita ser m-*crescente* se $V_H(R) \geq 0$, para todo o retângulo m-dimensional $R \subseteq S_1 \times \cdots \times S_m$.

Definição 1.6.5. Seja $H : S_1 \times \cdots \times S_m \to \mathbb{R}$, onde os S_i's são subconjuntos não-vazios de $\mathbb{R}$ que contém um menor elemento a_i, $i = 1, \cdots, m$. Então, dizemos que H é *grounded* se, para todos os pontos $(u_1, \cdots, u_m) \in \text{Dom}(H)$ tais que $u_k = a_k$ para pelo menos um índice k, temos $H(u_1, \cdots, u_m) = 0$.

Assim como no caso bidimensional, uma função H ser m-crescente não implica nem é implicado por H ser não-decrescente em cada um de seus argumentos, como mostram os próximos exemplos.

Exemplo 1.6.4. Seja m um inteiro positivo par. Considere a função $H : I^m \to \mathbb{R}$ dada por

$$H(u_1, \cdots, u_m) = \max\{u_1, \cdots, u_m\}.$$

É claro que H assim definida é não-decrescente em cada um de seus argumentos, mas H não é m-crescente. Para ver isso, considere o retângulo $R = I^m$. Note que $A_m(R)$ neste caso possui um único elemento que não possui alguma coordenada igual a 1, a saber, o vetor $(0, \cdots, 0)$. O conjunto $A_m(R)$ tem 2^m elementos sendo que entre esses elementos, digamos $\mathbf{w_i}$, $i = 1, \cdots, 2^m$, metade deles é tal que $\text{csgn}(\mathbf{w_i}) = 1$ (o vetor $(0, \cdots, 0)$ é um deles, já que m é par) e a outra metade é tal que $\text{csgn}(\mathbf{w_i}) = -1$. Além disso, $H(\mathbf{w_i}) = 1$ para todo $\mathbf{w_i} \in A_m(R) - \{(0, \cdots, 0)\}$ e $H(0, \cdots, 0) = 0$. Logo,

$$\begin{aligned} V_H(R) &= \sum_{\mathbf{w}\in A_m(R)} \text{csgn}(\mathbf{w})H(\mathbf{w}) \\ &= \left[\frac{2^m}{2} - 1 + H(0, \cdots, 0)\right] - \left[\frac{2^m}{2}\right] \\ &= -1. \end{aligned}$$

Portanto, $V_H(R) < 0$.

Exemplo 1.6.5. Considere a função $H : I^m \to I$ dada por

$$H(u_1, \cdots, u_m) = (1 - u_1)(1 - u_2) \cdots (1 - u_m).$$

É claro que a função H assim definida é decrescente em cada um de seus argumentos. Vamos provar por indução (na dimensão do retângulo R) que, dado $R = [u_1, v_1] \times \cdots \times [u_m, v_m] \subseteq I^m$, temos

$$V_H(R) = (-1)^m \prod_{i=1}^{m} (v_i - u_i). \tag{1.24}$$

Para $k = 2$ o resultado foi mostrado no Exemplo 1.1.1. Suponha que o resultado seja válido para $k = m - 1$, isto é, suponha que para todo $R' = [u_1, v_1] \times \cdots \times [u_{m-1}, v_{m-1}] \subseteq I^{m-1}$ valha

$$V_H(R') = \sum_{\mathbf{w}\in A_{m-1}(R')} \text{csgn}(\mathbf{w})H(\mathbf{w}) = (-1)^{m-1} \prod_{i=1}^{m-1} (v_i - u_i). \tag{1.25}$$

Inferiremos daí a validade do resultado para $k = m$. Dado $R' = [u_1, v_1] \times \cdots \times [u_{m-1}, v_{m-1}] \subseteq I^{m-1}$, considere $R = R' \times [u_m, v_m] \subseteq I^m$ e os conjuntos

$$\begin{aligned} A_m^{u_m}(R) = \{&\mathbf{w} = (w_1, \cdots, w_{m-1}, u_m) \in R \text{ tais que } w_i = u_i \text{ ou } w_i = v_i, \\ &i = 1, \cdots, m-1\}, \\ A_m^{v_m}(R) = \{&\mathbf{w} = (w_1, \cdots, w_{m-1}, v_m) \in R \text{ tais que } w_i = u_i \text{ ou } w_i = v_i, \\ &i = 1, \cdots, m-1\}. \end{aligned}$$

Note que $A_m^{u_m}(R)$ é o conjunto dos vértices de R cuja m-ésima coordenada é u_m e $A_m^{v_m}(R)$ é o conjunto dos vértices de R cuja m-ésima coordenada é v_m. Além disso, é claro que $A_m(R) = A_m^{u_m}(R) \cup A_m^{v_m}(R)$ e $A_m^{u_m}(R) \cap A_m^{v_m}(R) = \emptyset$, ou seja, $A_m^{u_m}(R)$ e $A_m^{v_m}(R)$ formam uma partição de $A_m(R)$. Agora,

$$\begin{aligned}
V_H(R) &= \sum_{\mathbf{w}\in A_m(R)} \operatorname{csgn}(\mathbf{w})H(\mathbf{w}) \\
&= \sum_{\mathbf{w}\in A_m^{u_m}(R)} \operatorname{csgn}(\mathbf{w})H(\mathbf{w}) + \sum_{\mathbf{w}\in A_m^{v_m}(R)} \operatorname{csgn}(\mathbf{w})H(\mathbf{w}) \\
&= \sum_{\mathbf{w}\in A_m^{u_m}(R)} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i)(1-u_m) + \sum_{\mathbf{w}\in A_m^{v_m}(R)} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i)(1-v_m) \\
&= \sum_{\mathbf{w}\in A_m^{u_m}(R)} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i) + \sum_{\mathbf{w}\in A_m^{v_m}(R)} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i) \\
&\quad - u_m \sum_{\mathbf{w}\in A_m^{u_m}(R)} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i) - v_m \sum_{\mathbf{w}\in A_m^{v_m}(R)} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i). \quad (1.26)
\end{aligned}$$

Note que, para cada $\mathbf{w} = (w_1, \cdots, w_{m-1}, u_m) \in A_m^{u_m}(R)$ temos

$$\begin{aligned}
\operatorname{csgn}(\mathbf{w})\big|_{\mathbf{w}\in A_m^{u_m}(R)} &= \operatorname{csgn}(w_1, \cdots, w_{m-1}, u_m) \\
&= -\operatorname{csgn}(w_1, \cdots, w_{m-1}) = \operatorname{csgn}(\mathbf{w})\big|_{\mathbf{w}\in A_{m-1}(R')}.
\end{aligned}$$

Por exemplo, o vetor $(u_1, v_2, \cdots, v_{m-1}, u_m) \in A_m^{u_m}(R)$ é tal que

$$\operatorname{csgn}\underbrace{(u_1, v_2, \cdots, v_{m-1}, u_m)}_{\in A_m^{u_m}(R)} = 1 = -\operatorname{csgn}\underbrace{(u_1, v_2, \cdots, v_{m-1})}_{\in A_{m-1}(R')}.$$

Dessa forma temos,

$$\sum_{\mathbf{w}\in A_m^{u_m}(R)} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i) = -\sum_{\mathbf{w}\in A_{m-1}(R')} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i). \qquad (1.27)$$

Por outro lado, para cada $\mathbf{w} = (w_1, \cdots, w_{m-1}, v_m) \in A_m^{v_m}(R)$ temos

$$\begin{aligned}
\operatorname{csgn}(\mathbf{w})\big|_{\mathbf{w}\in A_m^{v_m}(R)} &= \operatorname{csgn}(w_1, \cdots, w_{m-1}, v_m) \\
&= \operatorname{csgn}(w_1, \cdots, w_{m-1}) = \operatorname{csgn}(\mathbf{w})\big|_{\mathbf{w}\in A_{m-1}(R')}.
\end{aligned}$$

Desta forma,

$$\sum_{\mathbf{w}\in A_m^{v_m}(R)} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i) = \sum_{\mathbf{w}\in A_{m-1}(R')} \operatorname{csgn}(\mathbf{w})\prod_{i=1}^{m-1}(1-w_i). \qquad (1.28)$$

Substituindo (1.27) e (1.28) em (1.26), obtemos,

$$\begin{aligned}
V_H(R) &= -\sum_{\mathbf{w}\in A_{m-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{i=1}^{m-1}(1-w_i) + \sum_{\mathbf{w}\in A_{m-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{i=1}^{m-1}(1-w_i) \\
&\quad + u_m \sum_{\mathbf{w}\in A_{m-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{i=1}^{m-1}(1-w_i) - v_m \sum_{\mathbf{w}\in A_{m-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{i=1}^{m-1}(1-w_i) \\
&= (u_m - v_m) \sum_{\mathbf{w}\in A_{m-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{i=1}^{m-1}(1-w_i) \\
&=^{6} (u_m - v_m)(-1)^{m-1} \prod_{i=1}^{m-1}(v_i - u_i) \\
&= (-1)^m \prod_{i=1}^{m}(v_i - u_i).
\end{aligned}$$

Isto prova (1.24). Note que, se $u_i < v_i$, $i = 1, \cdots, n$, por (1.24), o H-volume de um retângulo $R = [u_1, v_1] \times \cdots \times [u_m, v_m]$ é positivo se m é par e é negativo se m é ímpar. Assim, $H(u_1, \cdots, u_m) = \prod_{i=1}^{m}(1-u_i)$ é uma função decrescente em cada um de seus argumentos que é m-crescente se m é par e não é m-crescente se m é ímpar.

Lema 1.6.1. *Seja $H : S_1 \times \cdots \times S_m \to \mathbb{R}$, onde os S_i's são conjuntos não-vazios de $\mathbb{R}$, uma função m-crescente e* grounded. *Então, H é não-decrescente em cada um dos seus argumentos.*

Prova: Sejam $k \in \mathbb{N}$ e a_i o menor elemento em S_i, $i = 1, \cdots, m$. Considere o retângulo $R = [a_1, v_1] \times \cdots \times [a_{k-1}, v_{k-1}] \times [x_k, y_k] \times [a_{k+1}, v_{k+1}] \times \cdots \times [a_m, v_m] \subset S_1 \times \cdots \times S_m$, com $a_i < v_i$, para $i = 1, \cdots, m$ e $a_k < x_k < y_k$. Então, o conjunto $A_m(R)$ contém apenas dois elementos que não possuem coordenadas iguais a a_i, a saber, $\mathbf{w}_1 = (v_1, \cdots, v_{k-1}, x_k, v_{k+1}, \cdots, v_m)$ e $\mathbf{w}_2 = (v_1, \cdots, v_{k-1}, y_k, v_{k+1}, \cdots, v_m)$. Note que como x_k é elemento minimal do intervalo $[x_k, y_k]$, temos $\operatorname{csgn}(\mathbf{w_1}) = -1$ e $\operatorname{csgn}(\mathbf{w_2}) = 1$ e como H é *grounded* e m-crescente, segue que

$$\begin{aligned}
0 \le V_H(R) &= \sum_{\mathbf{w}\in A_m(R)} \operatorname{csgn}(\mathbf{w}) H(\mathbf{w}) \\
&= \operatorname{csgn}(\mathbf{w}_1) H(\mathbf{w}_1) + \operatorname{csgn}(\mathbf{w}_2) H(\mathbf{w}_2) \\
&= H(\mathbf{w}_2) - H(\mathbf{w}_1) \implies H(\mathbf{w}_1) \le H(\mathbf{w}_2).
\end{aligned}$$

Logo, H é não-decrescente em cada um de seus argumentos. □

[6]Por (1.25) (hipótese de indução).

Definição 1.6.6. Seja $H : S_1 \times \cdots \times S_m \to \mathbb{R}$, onde os S_i's são subconjuntos não-vazios de $\mathbb{R}$. Suponha que cada S_i possua um elemento maximal $b_i, i = 1, \cdots, m$. Então dizemos que H tem *marginais*. As marginais unidimensionais de H, $F_i : S_i \to \mathbb{R}$, são dadas por

$$F_i(x) = H(b_1, \cdots, b_{i-1}, x, b_{i+1}, \cdots, b_m).$$

As marginais de dimensões maiores são definidas analogamente, ou seja, fixando-se pontos maximais em H e deixando-se variar os outros pontos.

Lema 1.6.2. *Seja H uma função que é m-crescente,* grounded *e que tenha marginais. Então, cada marginal F_i de H é não-decrescente, e se*

$$(u_1, \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m) \in Dom(H),$$

então,

$$0 \leq H(u_1 \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m) \leq F_k(x),$$

para todo $x \in Dom(F_k)$. Enunciado equivalente vale para as marginais de dimensão maior.

Prova: Pelo Lema 1.6.1, H é não-decrescente em cada um de seus argumentos. Assim, com a mesma notação da Definição 1.6.6, segue que

$$H(u_1, \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m) \leq H(b_1, \cdots, b_{k-1}, x, b_{k+1}, \cdots, b_m) = F_k(x),$$

para todo $x \in \mathrm{Dom}(F_k)$. Por outro lado, como H é *grounded*, pelo Lema 1.6.1, segue que,

$$0 = H(a_1, \cdots, a_{k-1}, x, a_{k+1}, \cdots, a_m) \leq H(u_1, \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m),$$

para todo $x \in \mathrm{Dom}(F_k)$. Usando um processo análogo, prova-se a desigualdade para o caso das marginais de dimensões maiores. □

Proposição 1.6.1. *Seja $H : S_1 \times \cdots \times S_m \to \mathbb{R}$ uma função m-crescente,* grounded *e com marginais. Para qualquer k, $1 \leq k \leq m$, com $x < y$, sejam $(u_1, \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m)$ e $(u_1, \cdots, u_{k-1}, y, u_{k+1}, \cdots, u_m) \in Dom(H)$. Então,*

$$\begin{aligned} &H(u_1,\cdots, u_{k-1}, y, u_{k+1}, \cdots, u_m) - \\ &\qquad - H(u_1, \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m) \leq F_k(y) - F_k(x). \end{aligned} \tag{1.29}$$

Prova: Se $m = 2$, considere o retângulo $R_1 = [x, y] \times [x_2, b_2] \subseteq S_1 \times S_2$. Como H é 2-crescente e *grounded*, temos que,

$$\begin{aligned} &H(x, x_2) + H(y, b_2) - H(x, b_2) - H(y, x_2) \geq 0 \implies \\ &H(x, x_2) - H(y, x_2) + F_1(y) - F_1(x) \geq 0 \implies \\ &\qquad H(y, x_2) - H(x, x_2) \leq F_1(y) - F_1(x). \end{aligned} \tag{1.30}$$

Por outro lado, considerando o retângulo $R_2 = [x_1, b_1] \times [x, y] \subseteq S_1 \times S_2$ temos

$$\begin{aligned} & H(x_1, x) + H(b_1, y) - H(b_1, x) - H(x_1, y) \geq 0 \quad \Longrightarrow \\ & H(x_1, y) - H(x_1, x) \leq F_2(y) - F_2(x). \end{aligned} \tag{1.31}$$

As desigualdades (1.30) e (1.31) juntas estabelecem (1.29).
Considere $m > 2$ e $x, y \in S_k$, $x < y$. Dados $(u_1, \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m)$ e $(u_1, \cdots, u_{k-1}, y, u_{k+1}, \cdots, u_m) \in \text{Dom}(H)$, definimos os $m-1$ retângulos m-dimensionais $R_k(l)$, $l = 1, \cdots, m$, $l \neq k$, como segue,

$$R_k(l) = [x_1, y_1] \times \cdots \times [x_m, y_m],$$

$$\text{com} \quad x_i = \begin{cases} a_i, & \text{se} \quad i \neq l, k, \\ u_l, & \text{se} \quad i = l, \\ x, & \text{se} \quad i = k, \end{cases} \quad \text{e} \quad y_i = \begin{cases} b_i, & \text{se} \quad i \leq l, \, i \neq k, \\ u_i, & \text{se} \quad i > l, \, i \neq k, \\ y, & \text{se} \quad i = k, \end{cases},$$

onde $a_i = \min\{S_i\}$ e $b_i = \max\{S_i\}$, $i = 1, \cdots, m$. Por exemplo, se $k = 5$, temos,

$$R_5(2) = [a_1, b_1] \times [u_2, b_2] \times [a_3, u_3] \times [a_4, u_4] \times [x, y] \times [a_6, u_6] \times \cdots \times [a_m, u_m],$$

$$R_5(3) = [a_1, b_1] \times [a_2, b_2] \times [u_3, b_3] \times [a_4, u_4] \times [x, y] \times [a_6, u_6] \times \cdots \times [a_m, u_m].$$

Como H é m-crescente então, $V_H(R_k(l)) \geq 0$, para cada l, assim,

$$\sum_{\substack{l=1 \\ l \neq k}}^{m} V_H(R_k(l)) \geq 0. \tag{1.32}$$

Como H é *grounded*, cada $V_H(R_k(l))$ se reduz a uma soma de 4 termos que envolvem os 4 vértices de $R_k(l)$ que não contém a_i's. Por exemplo, se k=5, os vértices de $R_5(2)$ que não contém a_i's são $(b_1, u_2, u_3, u_4, x, u_6, \cdots, u_m)$, $(b_1, u_2, u_3, u_4, y, u_6, \cdots, u_m)$, $(b_1, b_2, u_3, u_4, x, u_6, \cdots, u_m)$ e $(b_1, b_2, u_3, u_4, y, u_6, \cdots, u_m)$. Quando estas $m-1$ somas são substituidas em (1.32), a soma é telescópica (veja Exemplo 1.6.6) e resulta

$$\begin{aligned} 0 \quad \leq \quad & H(b_1, \cdots, b_{k-1}, y, b_{k+1} \cdots, b_m) - H(u_1, \cdots, u_{k-1}, y, u_{k+1} \cdots, u_m) \\ & - H(b_1, \cdots, b_{k-1}, x, b_{k+1} \cdots, b_m) + H(u_1, \cdots, u_{k-1}, x, u_{k+1} \cdots, u_m) \\ = \quad & H(u_1, \cdots, u_{k-1}, x, u_{k+1} \cdots, u_m) - H(u_1, \cdots, u_{k-1}, y, u_{k+1} \cdots, u_m) \\ & + F_k(y) - F_k(x). \end{aligned} \tag{1.33}$$

Logo, (1.33) reduz (1.32) a (1.29) e isto completa a prova. □

O próximo exemplo serve para clarear a idéia da prova da Proposição 1.6.1, pois, num primeiro momento, não parece claro que a soma em (1.32) é telescópica. Para esclarecer faremos o caso particular $m = 4$ e $k = 3$.

Exemplo 1.6.6. Seja $m = 4$ e $k = 3$. Vamos construir os retângulos $R_k(l)$ conforme a Proposição 1.6.1. Lembrando da definição dos retângulos $R_k(l)$, temos, para $l = 1$

$$R_3(1) = [u_1, b_1] \times [a_2, u_2] \times [x, y] \times [a_4, u_4].$$

Os vértices de $R_3(1)$ que não possuem a_i's são

$$(u_1, u_2, x, u_4), (u_1, u_2, y, u_4), (b_1, u_2, x, u_4) \text{ e } (b_1, u_2, y, u_4).$$

Assim, como H é *grounded*, $V_H(R_3(1))$, dado em (1.23), se reduz a

$$\begin{aligned} V_H(R_3(1)) = {} & H(u_1, u_2, x, u_4) - H(u_1, u_2, y, u_4) - H(b_1, u_2, x, u_4) \\ & + H(b_1, u_2, y, u_4). \end{aligned} \tag{1.34}$$

Para $l = 2$, temos,

$$R_3(2) = [a_1, b_1] \times [u_2, b_2] \times [x, y] \times [a_4, u_4].$$

Os vértices de $R_3(2)$ que não possuem a_i's são

$$(b_1, u_2, x, u_4), (b_1, u_2, y, u_4), (b_1, b_2, x, u_4) \text{ e } (b_1, b_2, y, u_4).$$

Assim, $V_H(R_3(2))$ se reduz a

$$\begin{aligned} V_H(R_3(2)) = {} & H(b_1, u_2, x, u_4) - H(b_1, u_2, y, u_4) - H(b_1, b_2, x, u_4) \\ & + H(b_1, b_2, y, u_4). \end{aligned} \tag{1.35}$$

Para $l = 4$, temos,

$$R_3(4) = [a_1, b_1] \times [a_2, b_2] \times [x, y] \times [u_4, b_4].$$

Os vértices de $R_3(4)$ que não possuem a_i's são

$$(b_1, b_2, x, u_4), (b_1, b_2, y, u_4), (b_1, b_2, x, b_4) \text{ e } (b_1, b_2, y, b_4).$$

Assim, $V_H(R_3(4))$ se reduz a

$$\begin{aligned} V_H(R_3(4)) = {} & H(b_1, b_2, x, u_4) - H(b_1, b_2, y, u_4) - H(b_1, b_2, x, b_4) \\ & + H(b_1, b_2, y, b_4). \end{aligned} \tag{1.36}$$

Substituindo (1.34), (1.35) e (1.36) na expressão (1.32) obtemos,

$$\begin{aligned} & \sum_{l=1,2,4} V_H(R_3(l)) = \\ & = H(u_1, u_2, x, u_4) - H(u_1, u_2, y, u_4) - H(b_1, u_2, x, u_4) + H(b_1, u_2, y, u_4) \\ & \quad + H(b_1, u_2, x, u_4) - H(b_1, u_2, y, u_4) - H(b_1, b_2, x, u_4) + H(b_1, b_2, y, u_4) \\ & \quad + H(b_1, b_2, x, u_4) - H(b_1, b_2, y, u_4) - H(b_1, b_2, x, b_4) + H(b_1, b_2, y, b_4) \\ & = H(u_1, u_2, x, u_4) - H(u_1, u_2, y, u_4) - H(b_1, b_2, x, b_4) + H(b_1, b_2, y, b_4) \\ & = H(u_1, u_2, x, u_4) - H(u_1, u_2, y, u_4) + F_3(y) - F_3(x). \end{aligned}$$

Como $\sum_{l=1,2,4} V_H(R_3(l)) \geq 0$, segue que

$$H(u_1, u_2, y, u_4) - H(u_1, u_2, x, u_4) \leq F_3(y) - F_3(x).$$

O próximo lema estabelece uma desigualdade fundamental que será útil na prova de várias propriedades de funções m-crescentes, *grounded* e com marginais.

Lema 1.6.3. *Seja $H : S_1 \times \cdots \times S_m \to \mathbb{R}$ uma função m-crescente,* grounded *e com marginais. Sejam $(u_1, \cdots, u_m)$ e $(v_1, \cdots, v_m) \in Dom(H)$. Então,*

$$|H(u_1, \cdots, u_m) - H(v_1, \cdots, v_m)| \leq \sum_{i=1}^{m} |F_i(u_i) - F_i(v_i)|. \tag{1.37}$$

Prova: Sejam $(u_1, \cdots, u_m)$ e $(v_1, \cdots, v_m) \in \mathrm{Dom}(H)$. Então, o Lema 1.6.2 e a Proposição 1.6.1 juntos são equivalentes a

$$\begin{aligned} &|H(u_1, \cdots, u_{k-1}, y, u_{k+1}, \cdots, u_m) - \\ &- \;\; H(u_1, \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m)| \leq |F_k(y) - F_k(x)|, \end{aligned}$$

para todo $(u_1, \cdots, u_{k-1}, y, u_{k+1}, \cdots, u_m)$ e $(u_1, \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m) \in \mathrm{Dom}(H)$. Assim,

$$\begin{aligned} |H(u_1, u_2, u_3, \cdots, u_m) - H(v_1, u_2, u_3, \cdots, u_m)| &\leq |F_1(u_1) - F_1(v_1)|, \\ |H(v_1, u_2, u_3, \cdots, u_m) - H(v_1, v_2, u_3, \cdots, u_m)| &\leq |F_2(u_2) - F_2(v_2)|, \\ &\vdots \\ |H(v_1, \cdots, v_{m-1}, u_m) - H(v_1, \cdots, v_{m-1}, v_m)| &\leq |F_n(u_m) - F_n(v_m)|. \end{aligned}$$

Somando as desigualdades acima obtemos,

$$\begin{aligned} &|H(u_1, u_2, u_3, \cdots, u_m) - H(v_1, u_2, u_3, \cdots, u_m)| + \\ &+ \;\; |H(v_1, u_2, u_3, \cdots, u_m) - H(v_1, v_2, u_3, \cdots, u_m)| + \cdots + \\ &+ \;\; |H(v_1, \cdots, v_{m-1}, u_m) - H(v_1, \cdots, v_{m-1}, v_m)| \leq \sum_{i=1}^{m} |F_i(u_i) - F_i(v_i)|. \end{aligned}$$

Por outro lado, pela desigualdade triangular,

$$\begin{aligned} &|H(u_1, \cdots, u_m) - H(v_1, \cdots, v_m)| = \\ &= \;\; |H(u_1, \cdots, u_m) - H(v_1, u_2, u_3, \cdots, u_m) + H(v_1, u_2, u_3, \cdots, u_m) - \\ &\quad\; - H(v_1, v_2, u_3, \cdots, u_m) + H(v_1, v_2, u_3, \cdots, u_m) + \cdots + \\ &\quad\; - H(v_1, \cdots, v_{n-1}, u_m) + H(v_1, \cdots, v_{n-1}, u_m) - H(v_1, \cdots, v_m)| \\ &\leq \;\; |H(u_1, \cdots, u_m) - H(v_1, u_2, \cdots, u_m)| + \\ &\quad\; + |H(v_1, u_2, u_3, \cdots, u_m) - H(v_1, v_2, u_3, \cdots, u_m)| + \cdots + \\ &\quad\; + |H(v_1, \cdots, v_{m-1}, u_m) - H(v_1, \cdots, v_m)| \\ &\leq \;\; \sum_{i=1}^{m} |F_i(u_i) - F_i(v_i)|. \end{aligned}$$

Logo,

$$|H(u_1, \cdots, u_m) - H(v_1, \cdots, v_m)| \leq \sum_{i=1}^{m} |F_i(u_i) - F_i(v_i)|.$$

Isto completa a prova. $\square$

Definição 1.6.7. Seja $m \in \mathbb{N}$, $m \geq 2$. Uma função H é dita ser uma *função de distribuição m-dimensional* se satisfaz as seguintes condições:

(i). $\mathrm{Dom}(H) = \mathbb{R}^m$;

(ii). H é m-crescente e *grounded*;

(iii). $H(\infty, \cdots, \infty) = 1$.

Diremos que H é a *função de distribuição conjunta* das variáveis aleatórias $X_1, \cdots, X_m$ se, para todo $\mathbf{x} = (x_1, \cdots, x_m) \in \mathbb{R}^m$,

$$H(\mathbf{x}) = P(X_1 \leq x_1, \cdots, X_m \leq x_m).$$

A função H assim definida será também uma função de distribuição m-dimensional conforme a Definição 1.6.7 (veja Tucker, 1967, página 24). Além disso, segue do Lema 1.6.1, que cada marginal de uma distribuição conjunta m-dimensional é uma função de distribuição, o mesmo valendo para as marginais de ordem maior. Em geral vamos denotar por $F_1, \cdots, F_m$ as marginais unidimensionais de uma distribuição conjunta m-dimensional e nos referiremos a elas simplesmente por marginais. As marginais de dimensão maiores serão chamadas de k-marginais, $2 \leq k < m$.

Definição 1.6.8. Seja $m \in \mathbb{N}$, $m \geq 2$. Uma *cópula m-dimensional* (ou, abreviadamente, uma *m-cópula*) é uma função C que satisfaz as seguintes condições:

(i). $\mathrm{Dom}(C) = I^m$;

(ii). C é m-crescente e *grounded*;

(iii). Para todo $k \in \mathbb{N}$, $k \leq m$, as marginais de C, C_k, são tais que

$$C_k(u_k) = C(1, \cdots, 1, u_k, 1, \cdots, 1) = u_k. \tag{1.38}$$

Note que uma m-cópula C é uma função de distribuição no sentido da Definição 1.6.7. Logo, todos os resultados vistos até aqui valem para m-cópulas. Além disso, as k-marginais de uma m-cópula qualquer, são também k-cópulas. O próximo teorema garante a continuidade uniforme de uma m-cópula qualquer. A prova é adaptada de Schweizer e Sklar (2005).

Teorema 1.6.1. *Uma m-cópula C é uniformemente contínua.*

Prova: Dado $\varepsilon > 0$ e $\mathbf{u} = (u_1, \cdots, u_m) \in I^m$, seja $\mathbf{v} = (v_1, \cdots, v_m) \in I^m$ tal que

$$\sum_{i=1}^{m} |u_i - v_i| < \varepsilon.$$

Como C é cópula, por (1.37) e (1.38), temos,

$$|C(\mathbf{u}) - C(\mathbf{v})| \leq \sum_{i=1}^{m} |u_i - v_i| < \varepsilon.$$

Isto prova que C é contínua. Como I^m é um intervalo compacto e $\mathrm{Dom}(C) = I^m$, segue que C é uniformemente contínua em I^m. □

Assim como no caso bidimensional, uma m-cópula também é diferenciável em quase toda a parte em cada um de seus argumentos, conforme mostra o próximo teorema.

Teorema 1.6.2. *Seja C uma m-cópula qualquer. Então, as derivadas parciais* $\frac{\partial C(u_1,\cdots,u_m)}{\partial u_k}$, $k = 1, \cdots, m$ *existem, tomam valores em I e são não-decrescentes quase toda a parte.*

Prova: A existência das derivadas é imediata, já que C é não-decrescente em cada um de seus argumentos e I^m é compacto. Logo, C é diferenciável em quase toda a parte. Fixado $k \in \{1, \cdots, m\}$ e $u_i \in [0, 1]$, $i = 1, \cdots, m$, $i \neq k$ e se $u_k \in (0, 1)$, temos por (1.37) que,

$$\begin{aligned} 0 &\leq \lim_{h \to 0} \left| \frac{C(u_1, \cdots, u_{k-1}, u_k + h, u_{k+1}, \cdots, u_m)}{h} - \right. \\ &\qquad\qquad \left. - \frac{C(u_1, \cdots, u_{k-1}, u_k, u_{k+1}, \cdots, u_m)}{h} \right| \\ &\leq \lim_{h \to 0} \frac{\sum_{\substack{i=1 \\ i \neq k}}^{m} |u_i - u_i| + |u_k + h - u_k|}{h} = 1, \end{aligned}$$

em quase toda a parte, ou seja,

$$0 \leq \frac{\partial C(u_1, \cdots, u_m)}{\partial u_k} \leq 1, \quad \forall k = 1, \cdots, m.$$

Usando-se limites laterais, prova-se a desigualdade para u_k em quase toda a parte de $[0, 1]$. Tomando-se $x \leq y$, o Lema 1.6.1 garante que a função

$$\begin{aligned} (u_1, \cdots, u_{k-1}, u_{k+1}, \cdots, u_m) \quad &\longmapsto \quad C(u_1, \cdots, u_{k-1}, y, u_{k+1}, \cdots, u_m) - \\ &\qquad -C(u_1, \cdots, u_{k-1}, x, u_{k+1}, \cdots, u_m), \end{aligned}$$

é não-decrescente em quase toda a parte. Logo, a função

$$\frac{\partial\big[C(u_1,\cdots,u_{k-1},y,u_{k+1},\cdots,u_m)-C(u_1,\cdots,u_{k-1},x,u_{k+1},\cdots,u_m)\big]}{\partial u_k}\geq 0,$$

está bem definida e é não-negativa quase toda a parte. Pela linearidade da derivada obtemos

$$\frac{\partial C(u_1,\cdots,u_{k-1},y,u_{k+1},\cdots,u_m)}{\partial u_k}\geq\frac{\partial C(u_1,\cdots,u_{k-1},x,u_{k+1},\cdots,u_m)}{\partial u_k},$$

o que prova que $\frac{\partial C(u_1,\cdots,u_m)}{\partial u_k}$ é não-decrescente em quase toda a parte para todo $k=1,\cdots,m$. $\square$

A versão do Teorema de Sklar para m-cópulas é a seguinte.

Teorema 1.6.3. *(Versão m-dimensional do Teorema de Sklar). Sejam X_1, $\cdots,X_m$ variáveis aleatórias com função de distribuição $F_1,\cdots,F_m$, respectivamente, e distribuição conjunta H. Então, existe uma m-cópula C tal que,*

$$H(x_1,\cdots,x_m)=C(F_1(x_1),\cdots,F_m(x_m)). \tag{1.39}$$

Se $F_1,\cdots,F_m$ são todas contínuas, então C é única. Caso contrário, C é unicamente determinada em $Im(F_1)\times\cdots\times Im(F_m)$.

Prova: Veja Schweizer e Sklar (2005) e referências ali contidas. $\square$

Corolário 1.6.1. *Seja $X_1,\cdots,X_m$ variáveis aleatórias com função de distribuição $F_1,\cdots,F_m$ respectivamente, e distribuição conjunta H. Então, para todo $(u_1,\cdots,u_m)\in\mathbb{R}^m$,*

$$C(u_1,\cdots,u_m)=H(F_1^{(-1)}(u_1),\cdots,F_m^{(-1)}(u_m)), \tag{1.40}$$

onde $F_k^{(-1)}$ é a quase-inversa de F_k.

Exemplo 1.6.7. (*Cópulas independência e limites de Fréchet-Hoeffding*).Os análogos m-dimensionais às cópulas Π, M e W são dados, respectivamente, por

$$\begin{aligned}\Pi^m(\mathbf{u}) &= u_1u_2\cdots u_m\,;\\ M^m(\mathbf{u}) &= \min\{u_1,\cdots,u_m\};\\ W^m(\mathbf{u}) &= \max\{u_1+\cdots+u_m-m+1,0\}.\end{aligned}$$

É fácil ver que Π^m é *grounded* e possui marginais univariadas conforme (1.38), isto é, $C_k(u_k)=u_k$, $k=1,\cdots,m$. Note também que dado $R=[u_1,v_1]\times\cdots\times[u_m,v_m]\subseteq I^m$,

$$V_{\Pi^m}(R)=\sum_{\mathbf{w}\in A_m(R)}\text{csgn}(\mathbf{w})\Pi(\mathbf{w})=\sum_{\mathbf{w}\in A_m(R)}\text{csgn}(\mathbf{w})\prod_{i=1}^m w_i={}^7\prod_{i=1}^m(v_i-u_i)\geq 0.$$

[7] A prova deste fato é dada no Lema 2.3.2 na Seção 2.3.

Também é fácil ver que M^m é uma cópula. De fato, é claro que M^m é *grounded* e suas marginais satisfazem (1.38). Além disso, um argumento similar ao do Exemplo 1.2.1 prova que, dado $R = [u_1, v_1] \times \cdots \times [u_m, v_m] \subseteq I^m$,

$$V_{M^m}(R) = \max\{0, \min\{v_1, \cdots, v_m\} - \max\{u_1, \cdots, u_m\}\} \geq 0.$$

Neste ponto começam a aparecer as diferenças entre 2-cópulas e m-cópulas. A função W^m não é copula qualquer que seja o inteiro $m > 2$. Para ver isso, seja $m > 2$ e considere o retângulo $R = [\frac{1}{2}, 1] \times \cdots \times [\frac{1}{2}, 1] \subseteq I^m$. Defina o conjunto

$$B_1 {=} \{\mathbf{w} \in A_m(R) |\text{ no máximo } m-2 \text{ coordenadas de } \mathbf{w} \text{ são iguais a } 1\}.$$

Note que, dado $\mathbf{w}' \in B_1$, $W^m(\mathbf{w}') = 0$. Defina também o conjunto

$$B_2 {=} \{\mathbf{w} \in A_m(R) |\, m-1 \text{ coordenadas de } \mathbf{w} \text{ são iguais a } 1\}.$$

Note que o conjunto B_2 possui m elementos e se $\mathbf{w}'' \in B_2$, então,

$$W^n(\mathbf{w}'') = \max\left\{m - 1 + \frac{1}{2} - m + 1, 0\right\} = \frac{1}{2},$$

e $\operatorname{csgn}(\mathbf{w}'') = -1$. Além disso, $A_m(R) = B_1 \cup B_2 \cup \{(1, \cdots, 1)\}$. Logo,

$$\begin{aligned} V_{W^m}(R) &= \sum_{\mathbf{w} \in A_m(R)} \operatorname{csgn}(\mathbf{w}) W^m(\mathbf{w}) = \\ &= \sum_{\mathbf{w} \in B_1} \operatorname{csgn}(\mathbf{w}) W^m(\mathbf{w}) + \sum_{\mathbf{w} \in B_2} \operatorname{csgn}(\mathbf{w}) W^m(\mathbf{w}) + W^m(1, \cdots, 1) \\ &= \sum_{\mathbf{w} \in B_2} \operatorname{csgn}(\mathbf{w}) W^m(\mathbf{w}) + 1 \\ &= 1 - \frac{m}{2} < 0, \end{aligned}$$

para todo $m > 2$.

O próximo teorema é o análogo multidimensional dos limites de Fréchet-Hoeffding. Ainda que W^m não seja uma m-cópula para $m > 2$, continua sendo uma cota inferior para o conjunto das m-cópulas.

Teorema 1.6.4. *Seja C uma m-cópula qualquer, $m > 2$. Então, para todo $\mathbf{u} \in I^m$,*

$$W^m(\mathbf{u}) \leq C(\mathbf{u}) \leq M^m(\mathbf{u}).$$

Prova: Seja $(u_1, \cdots, u_m) \in I^m$. O Lema 1.6.2 e a condição (1.38) implicam

$$C(u_1, \cdots, u_m) \leq u_k, \quad \forall\, k = 1, \cdots, m.$$

Logo,

$$C(u_1, \cdots, u_m) \leq \min\{u_1, \cdots, u_m\}.$$

Por outro lado, pelo Lema 1.6.3,

$$\begin{aligned}
|C(u_1, \cdots, u_m) - C(1, \cdots, 1)| &\leq \sum_{i=1}^{m} |C_i(u_i) - C_i(1)| \implies \\
1 - C(u_1, \cdots, u_m) &\leq \sum_{i=1}^{m} (1 - C_i(u_i)) \implies \\
1 - C(u_1, \cdots, u_m) &\leq m - \sum_{i=1}^{m} u_i.
\end{aligned}$$

Logo,

$$C(u_1, \cdots, u_m) \geq \sum_{i=1}^{m} u_i - m + 1, \tag{1.41}$$

qualquer que seja $(u_1, \cdots, u_m) \in I^m$. Além disso, como C é *grounded* e não-decrescente em cada argumento, segue que

$$C(u_1, \cdots, u_m) \geq 0,$$

desigualdade que combinada a (1.41) estabelece

$$C(u_1, \cdots, u_m) \geq \max\{0, u_1 + \cdots + u_m - m + 1\},$$

o que completa a prova. □

Assim como no caso bidimensional, as m-cópulas Π^m e M^m também possuem interpretações em relação às características de dependência entre as variáveis aleatórias envolvidas. Isto é o que mostra os próximos teoremas.

Teorema 1.6.5. *Seja um inteiro $m \geq 2$. Sejam $X_1, \cdots, X_m$ variáveis aleatórias contínuas. Então, $X_1, \cdots, X_m$ são independentes se, e somente se, a m-cópula associada a $X_1, \cdots, X_m$ é Π^m.*

Prova: Suponha que $X_1, \cdots, X_m$ sejam independentes. Seja H a sua distribuição conjunta com marginais $F_1, \cdots, F_m$, respectivamente. Então,

$$H(x_1, \cdots, x_m) = F_1(x_1) \cdots F_m(x_m). \tag{1.42}$$

Substituindo $x_i = F^{(-1)}(u_i)$, para todo $(u_1, \cdots, u_m) \in I^m$, em (1.42), por (1.40), obtemos,

$$C(u_1, \cdots, u_m) = u_1 \cdots u_m = \Pi^m(u_1, \cdots, u_m).$$

Reciprocamente, suponha que a m-cópula associada a $X_1, \cdots, X_m$ seja Π^m. Então, por (1.39),

$$H(x_1, \cdots, x_m) = \Pi^m\big(F_1(x_1), \cdots, F_m(x_m)\big) = F_1(x_1) \cdots F_m(x_m),$$

o que mostra que $X_1, \cdots, X_m$ são independentes e completa a prova. □

Teorema 1.6.6. *Seja m um inteiro $m \geq 2$. Sejam $X_1, \cdots, X_m$ variáveis aleatórias contínuas. Então, a m-cópula associada a $X_1, \cdots, X_m$ é M^m se, e somente se, as variáveis aleatórias $X_1, \cdots, X_m$ são quase certamente uma função crescente de alguma delas.*

Prova: Veja Schweizer e Sklar (2005) e referências ali contidas. □

Teorema 1.6.7. *Sejam $X_1, \cdots, X_m$ variáveis aleatórias contínuas com m-cópula associada C. Sejam f_i :$Im(X_i) \longrightarrow \mathbb{R}$ funções crescentes, para todo $i = 1, \cdots, m$. Considere as variáveis aleatórias $f_1(X_1), \cdots, f_m(X_m)$. Então a m-cópula associada a $f_1(X_1), \cdots, f_m(X_m)$ é C. Isto é, C é invariante mediante transformações estritamente crescentes.*

Prova: Veja Schweizer e Sklar (2005) e referências ali contidas. Veja também as referências contidas em Anjos et al. (2004). □

Se $1 \leq k < m$ e C é uma m-cópula, então as marginais k-dimensionais de C são as funções obtidas igualando-se $m - k$ dos argumentos a 1. Assim, uma m-cópula possui $\binom{m}{k}$ k-marginais e, para $k \geq 2$, estas marginais são k-cópulas também. Porém a recíproca não vale em geral. Dadas $\binom{m}{k}$ k-cópulas nem sempre elas são as marginais k-dimensionais de alguma m-cópula. Se isto ocorrer, então as $\binom{m}{k}$ k-cópulas são ditas ser *compatíveis*. O problema de determinar condições para que um conjunto de k-cópulas sejam as marginais de alguma m-cópulas é chamado de *problema de compatibilidade*. O problema de compatibilidade é muito importante, porém difícil de ser abordado. Apesar de vários avanços, o problema ainda está longe de estar completamente resolvido.

Se C_1, C_2 e C_3 são 2-cópulas quaisquer, então, C_1, C_2 e C_3 são compatíveis se, e somente se, existe uma 3-cópula tal que suas marginais bidimensionais são, em alguma ordem, C_1, C_2 e C_3. Não é nosso objetivo aqui estudar este problema de forma aprofundada. Ao leitor interessado, referenciamos Joe (1997) e Nelsen (1999). Outras referências podem ser encontradas também em Schweizer e Sklar (2005).

Capítulo 2

Construindo Cópulas

Neste capítulo apresentaremos algumas técnicas para a construção de cópulas. Construir cópulas não é um problema fácil. No caso m-dimensional, existem ainda muitas limitações técnicas além do problema de compatibilidade, o que torna a construção de cópulas m-dimensionais uma tarefa bastante difícil.

Existem muitos métodos para se construir cópulas, mas nenhum deles é geral o bastante a ponto de permitir que qualquer cópula possa ser construída através dele. Aqui apresentamos algumas construções particularmente interessantes, como a das Famílias Ali-Mikhail-Haq, Plackett e Arquimediana. Além dessas apresentamos também a interessante construção de Dolati e Úbeda-Flores. As famílias de cópulas Ali-Mikhail-Haq, Plackett e Arquimediana são muito utilizadas em aplicações devido a suas propriedades interessantes e ao grande número de cópulas importantes que fazem parte destas famílias. Para uma aplicação da Família Plackett, veja Rockinger e Jondeau (2006). Para mais detalhes a respeito da Família Ali-Mikhail-Haq, veja Ali et al. (1978). Para mais detalhes a respeito da Família Arquimediana de cópulas, veja Nelsen (1999), página 89. Veja também Anjos et al. (2004).

2.1 Família Ali-Mikhail-Haq de Cópulas

Sejam X e Y duas variáveis aleatórias com função de distribuição F e marginais F_1 e F_2, respetivamente. Quando X e Y denotam tempo de vida de objetos ou seres, é natural falar a respeito de diferenças de sobrevivência. Isto é, a razão entre a probabilidade de sobrevivência além do tempo x e de falha até o tempo x. Matematicamente isto quer dizer estudar o comportamento da razão

$$\frac{P(X > x)}{P(X \leq x)} = \frac{1 - F_1(x)}{F_1(x)}\,.$$

Analogamente, para o caso bivariado podemos nos perguntar como se comporta a razão $\dfrac{1 - F(x, y)}{F(x, y)}$.

Exemplo 2.1.1. Suponha que X e Y tenham função de distribuição logística bivariada de Gumbel (veja Exemplo 1.5.4), isto é

$$F(x,y) = \frac{1}{1+e^{-x}+e^{-y}}, \qquad \forall\, x,y \in \mathbb{R}.$$

Temos

$$\frac{1-F(x,y)}{F(x,y)} = \frac{1-(1+e^{-x}+e^{-y})^{-1}}{(1+e^{-x}+e^{-y})^{-1}} = e^{-x}+e^{-y}.$$

Do Exemplo 1.5.6, sabemos que as marginais de X e Y são dadas por

$$F_1(x) = \frac{1}{1+e^{-x}} \qquad \text{e} \qquad F_2(y) = \frac{1}{1+e^{-y}}\cdot$$

Assim,

$$\frac{1-F_1(x)}{F_1(x)} = e^{-x} \qquad \text{e} \qquad \frac{1-F_2(y)}{F_2(y)} = e^{-y}.$$

Logo, F satisfaz a equação funcional

$$\frac{1-F(x,y)}{F(x,y)} = \frac{1-F_1(x)}{F_1(x)} + \frac{1-F_2(y)}{F_2(y)}\cdot \tag{2.1}$$

Exemplo 2.1.2. Suponha que X e Y sejam variáveis aleatórias independentes com função de distribuição conjunta F e marginais F_1 e F_2, respectivamente. Assim $F(x,y) = F_1(x)F_2(y)$, para todo $x,y \in \mathbb{R}$. Note que podemos escrever

$$F_1(x) = \left[1+\frac{1-F_1(x)}{F_1(x)}\right]^{-1},$$

e analogamente para F_2. Dessa forma obtemos,

$$\begin{aligned}
\frac{1-F(x,y)}{F(x,y)} &= \frac{1}{F_1(x)F_2(y)} - 1 \\
&= \left[1+\frac{1-F_1(x)}{F_1(x)}\right]\left[1+\frac{1-F_2(y)}{F_2(y)}\right] - 1 \\
&= \frac{1-F_1(x)}{F_1(x)} + \frac{1-F_2(y)}{F_2(y)} + \frac{[1-F_1(x)][1-F_2(y)]}{F_1(x)F_2(y)}\cdot \qquad (2.2)
\end{aligned}$$

Exemplo 2.1.3. Seja X uma variável aleatória com distribuição $\mathcal{E}(\lambda_1)$ e Y uma variável aleatória com distribuição $\mathcal{E}(\lambda_2)$, $\lambda_1, \lambda_2 > 0$. Sejam F_1 e F_2 as funções de distribuição de X e Y, respectivamente. Assim, $F_1(x) = 1-e^{-\lambda_1 x}$ e $F_2(y) = 1-e^{-\lambda_2 y}$. Suponha que a distribuição conjunta de X e Y seja dada por

$$F(x,y) = \frac{2(1-e^{-\lambda_1 x})(1-e^{-\lambda_2 y})}{2-e^{-(\lambda_1 x+\lambda_2 y)}}\cdot$$

Nestas condições, vamos mostrar que F_1, F_2 e F satisfazem a equação funcional

$$\frac{1-F(x,y)}{F(x,y)} = \frac{1-F_1(x)}{F_1(x)} + \frac{1-F_2(y)}{F_2(y)} + \frac{1}{2}\left(\frac{[1-F_1(x)][1-F_2(y)]}{F_1(x)F_2(y)}\right). \quad (2.3)$$

Por um lado temos,

$$\begin{aligned} \frac{1-F(x,y)}{F(x,y)} &= \frac{2-e^{-(\lambda_1 x+\lambda_2 y)}-2(1-e^{-\lambda_1 x})(1-e^{-\lambda_2 y})}{2(1-e^{-\lambda_1 x})(1-e^{-\lambda_2 y})} \\ &= \frac{2e^{-\lambda_1 x}+2e^{-\lambda_2 y}-3e^{-(\lambda_1 x+\lambda_2 y)}}{2(1-e^{-\lambda_1 x})(1-e^{-\lambda_2 y})}. \end{aligned} \quad (2.4)$$

Por outro lado,

$$\begin{aligned} \frac{1-F_1(x)}{F_1(x)} + \frac{1-F_2(y)}{F_2(y)} + \frac{1}{2}\left(\frac{[1-F_1(x)][1-F_2(y)]}{F_1(x)F_2(y)}\right) &= \\ = \frac{e^{-\lambda_1 x}}{1-e^{-\lambda_1 x}} + \frac{e^{-\lambda_2 y}}{1-e^{-\lambda_2 y}} + \frac{1}{2}\left(\frac{e^{-(\lambda_1 x+\lambda_2 y)}}{(1-e^{-\lambda_1 x})(1-e^{-\lambda_2 y})}\right) \\ = \frac{2e^{-\lambda_1 x}(1-e^{-\lambda_2 y})+2e^{-\lambda_2 y}(1-e^{-\lambda_1 x})+e^{-(\lambda_1 x+\lambda_2 y)}}{2(1-e^{-\lambda_1 x})(1-e^{-\lambda_2 y})} \\ = \frac{2e^{-\lambda_1 x}+2e^{-\lambda_2 y}-3e^{-(\lambda_1 x+\lambda_2 y)}}{2(1-e^{-\lambda_1 x})(1-e^{-\lambda_2 y})}. \end{aligned} \quad (2.5)$$

De (2.4) e (2.5) concluímos (2.3).

A construção apresentada por Ali et al. (1978) é baseada nas soluções de uma equação diferencial parcial sob certas condições de contorno que apresentaremos a seguir. Antes disso, algumas definições e resultados necessários à construção são apresentados.

Sejam então X e Y duas variáveis aleatórias contínuas com função de distribuição conjunta $F(x,y)$ e distribuições marginais $F_1(x)$ e $F_2(y)$, respectivamente. Defina a *função razão de diferenças de sobrevivência* bidimensional H por

$$H(x,y) = \frac{1-F(x,y)}{F(x,y)},$$

e defina as *funções razões marginais de diferenças de sobrevivência* por

$$H_1(x) = \frac{1-F_1(x)}{F_1(x)} \quad \text{e} \quad H_2(y) = \frac{1-F_2(y)}{F_2(y)}. \quad (2.6)$$

Note que podemos escrever

$$F(x,y) = \frac{1}{1+H(x,y)}.$$

Observe também que como F_1 e F_2 são contínuas e não-decrescentes, H_1 e H_2 são não-crescentes e

$$\lim_{x \to -\infty} H_1(x) = \lim_{x \to -\infty} \frac{1 - F_1(x)}{F_1(x)} = \infty,$$

e

$$\lim_{x \to \infty} H_1(x) = \lim_{x \to \infty} \frac{1 - F_1(x)}{F_1(x)} = 0.$$

Analogamente para H_2. Assim H_1 e H_2 são funções que decrescem de ∞ para 0 no domínio de X e Y. Deste fato segue que $H(x, y)$ deve satisfazer as seguintes propriedades:

(i). $H(x, \infty) = \lim_{y \to \infty} H(x, y) = \lim_{y \to \infty} \frac{1 - F(x, y)}{F(x, y)} = \frac{1 - F_1(x)}{F_1(x)} = H_1(x)$.
Analogamente, $H(\infty, y) = H_2(y)$;

(ii). $H(x, -\infty) = \lim_{y \to -\infty} H(x, y) = \lim_{y \to -\infty} \frac{1 - F(x, y)}{F(x, y)} = \infty$.
Analogamente, $H(-\infty, y) = \infty$;

(iii). Como $F(x, y)$ é função de distribuição, então, dado o retângulo $[x_1, y_1] \times [x_2, y_2]$ contido no domínio de (X, Y), temos

$$F(x_1, x_2) + F(y_1, y_2) - F(x_1, y_2) - F(x_2, y_1) \geq 0,$$

o que é equivalente a

$$\frac{1}{1 + H(x_1, x_2)} + \frac{1}{1 + H(y_1, y_2)} - \frac{1}{1 + H(x_1, y_2)} - \frac{1}{1 + H(x_2, y_1)} \geq 0.$$

Ali et al. (1978) estudaram as soluções da equação diferencial

$$\frac{\partial^2 H(x, y)}{\partial H_1(x) \partial H_2(y)} = \alpha, \tag{2.7}$$

onde α é um parâmetro real e $H(\cdot\,, \cdot)$ é uma função de distribuição a ser determinada e $H_1(\cdot)$ e $H_2(\cdot)$ são dadas em (2.6).

O objetivo é obter soluções de (2.7) que satisfaçam as condições de contorno (i) a (iii) acima para obter $H(\cdot\,, \cdot)$ em função de $F(\cdot\,, \cdot)$ e então aplicar o Teorema de Sklar. Dessa forma, obtém-se uma cópula ou família de cópulas que satisfazem a equação diferencial parcial (2.7) com condições de contorno (i) a (iii) com as mudanças óbvias necessárias.

Pode-se mostrar que a solução da equação diferencial (2.7) é

$$H(x, y) = S(H_1(x)) + R(H_2(y)) + \alpha H_1(x) H_2(y),$$

onde $S(\cdot)$ e $R(\cdot)$ são funções diferenciáveis arbitrárias definidas na imagem de H_1 e H_2, respectivamente. Note que

$$\begin{aligned} H_2(y) &= \lim_{x\to\infty} H(x,y) \\ &= \lim_{x\to\infty} S(H_1(x)) + R(H_2(y)) + \alpha H_1(x)H_2(y) \\ &= S(0) + R(H_2(y)). \end{aligned}$$

Logo,

$$R(H_2(y)) = H_2(y) - S(0).$$

Analogamente,

$$S(H_1(x)) = H_1(x) - R(0).$$

Com isso, a solução de (2.7) se torna

$$H(x,y) = H_1(x) + H_2(y) + \alpha H_1(x)H_2(y) - S(0) - R(0). \tag{2.8}$$

Como

$$\begin{aligned} 0 &= \lim_{x,\, y\to\infty} H(x,y) \\ &= \lim_{x,\, y\to\infty} H_1(x) + H_2(y) + \alpha H_1(x)H_2(y) - S(0) - R(0) \\ &= -S(0) - R(0), \end{aligned}$$

segue que $S(0) + R(0) = 0$. Logo,

$$H(x,y) = H_1(x) + H_2(y) + \alpha H_1(x)H_2(y).$$

Aplicando a reparametrização $\alpha = 1 - \theta$ e substituindo as expressões de H, H_1 e H_2 em relação à F, F_1 e F_2 em (2.8), obtemos

$$\frac{1-F(x,y)}{F(x,y)} = \frac{1-F_1(x)}{F_1(x)} + \frac{1-F_2(y)}{F_2(y)} + (1-\theta)\frac{(1-F_1(x))(1-F_2(y))}{F_1(x)F_2(y)}, \tag{2.9}$$

para alguma constante $\theta \in \mathbb{R}$. Note que (2.1), (2.2) e (2.3) satisfazem (2.9) com $\theta = 1$, 0 e $1/2$, respectivamente. Aplicando as transformações $u = F_1(x)$ e $v = F_2(y)$ e o Teorema de Sklar em (2.9) obtemos,

$$\frac{1-C_\theta(u,v)}{C_\theta(u,v)} = \frac{1-u}{u} + \frac{1-v}{v} + (1-\theta)\left[\frac{(1-u)(1-v)}{uv}\right].$$

Após alguma álgebra obtemos,

$$\begin{aligned} uv(1-C_\theta(u,v)) &= v(1-u)C_\theta(u,v) + u(1-v)C_\theta(u,v) \\ &\quad + (1-\theta)(1-u)(1-v)C_\theta(u,v), \iff \\ uv - uvC_\theta(u,v) &= C_\theta(u,v)[v - uv + u - uv + (1-u)(1-v) \\ &\quad - \theta(1-u)(1-v)] \iff \\ C_\theta(u,v) &= \frac{uv}{1-\theta(1-u)(1-v)}. \end{aligned} \tag{2.10}$$

Resta saber se $C_\theta(u,v)$ definido em (2.10) é de fato uma cópula. É claro que $C_\theta(u,v)$ satisfaz a equação diferencial (2.7) e às condições de contorno (i) à (iii). É imediato também que C_θ é *grounded* e tem marginais u e v. Para que $C_\theta(u,v) \geq 0$, para todo $u, v \in I$, é preciso que $1 - \theta(1-u)(1-v) > 0$. Mas,

$$0 < 1 - \theta(1-u)(1-v) \leq 1 - \theta.$$

Logo, $\theta < 1$. Se $\theta = 1$ temos

$$C_1(u,v) = \frac{uv}{u + v - uv} \geq 0, \quad \forall\, u, v \in I.$$

Logo, $C_\theta(u,v) \geq 0$, para todo $u, v \in I \iff \theta \leq 1$. Para que C_θ seja 2-crescente é suficiente que (veja Lema 1.5.1) C seja absolutamente contínua e que

$$\frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} \geq 0.$$

Pela regra do quociente temos,

$$\frac{\partial C_\theta(u,v)}{\partial u} = \underbrace{\frac{v}{1 - \theta(1-u)(1-v)}}_{A(u,v)} - \underbrace{\frac{uv\theta(1-v)}{[1 - \theta(1-u)(1-v)]^2}}_{B(u,v)}.$$

Novamente a regra do quociente nos dá

$$\frac{\partial A(u,v)}{\partial v} = \frac{1}{1 - \theta(1-u)(1-v)} - \frac{\theta v(1-u)}{[1 - \theta(1-u)(1-v)]^2},$$

e

$$\frac{\partial B(u,v)}{\partial v} = \frac{\theta u(1-v) - uv\theta}{[1 - \theta(1-u)(1-v)]^2} - \frac{2\theta^2 uv(1-u)(1-v)}{[1 - \theta(1-u)(1-v)]^3}.$$

Logo,

$$\begin{aligned}
\frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} &= \frac{1}{[1 - \theta(1-u)(1-v)]} - \frac{\theta v(1-u)}{[1 - \theta(1-u)(1-v)]^2} \\
&\quad - \frac{\theta u(1-v) - uv\theta}{[1 - \theta(1-u)(1-v)]^2} + \frac{2\theta^2 uv(1-u)(1-v)}{[1 - \theta(1-u)(1-v)]^3} \\
&= \frac{1}{[1 - \theta(1-u)(1-v)]} + \frac{3\theta uv - \theta(u+v)}{[1 - \theta(1-u)(1-v)]^2} \\
&\quad + \frac{2\theta^2 uv(1-u)(1-v)}{[1 - \theta(1-u)(1-v)]^3} \\
&= \frac{1 - \theta + 2\theta uv}{[1 - \theta(1-u)(1-v)]^2} + \frac{2\theta^2 uv(1-u)(1-v)}{[1 - \theta(1-u)(1-v)]^3} \\
&= \frac{1-\theta}{u^2v^2} C_\theta^2(u,v) + \frac{2\theta}{uv} + \frac{2\theta^2(1-u)(1-v)C_\theta^3(u,v)}{u^2v^2} \\
&= \frac{C_\theta^2(u,v)}{u^2v^2} [1 - \theta + 2\theta uv + 2\theta^2(1-u)(1-v)C_\theta(u,v)]. \quad (2.11)
\end{aligned}$$

Mas, da expressão (2.10), temos que

$$\begin{aligned} 2\theta uv &= 2\theta[1-\theta(1-u)(1-v)]C_\theta(u,v) \\ &= [\,2\theta - 2\theta^2(1-u)(1-v)]C_\theta(u,v). \end{aligned} \tag{2.12}$$

Substituindo (2.12) em (2.11), obtemos,

$$\frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} = \frac{C_\theta(u,v)}{u^2v^2}\,[1-\theta+2\theta C_\theta(u,v)]. \tag{2.13}$$

Para que $0 \le C_\theta(u,v) \le 1$ e $\frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} \ge 0$ é necessário e suficiente que $1-\theta+2\theta C_\theta(u,v) \ge 0$, mas,

$$0 \le 1-\theta+2\theta C_\theta(u,v) \le 1-\theta+2\theta = 1+\theta \iff \theta \ge -1.$$

Além disso, um cálculo extenso, porém elementar, nos mostra que

$$\int_0^u\int_0^v \frac{\partial^2 C_\theta(x,y)}{\partial x \partial y}\, dxdy = C_\theta(u,v).$$

Dessa forma, a função definida por

$$C_\theta(u,v) = \frac{uv}{1-\theta(1-u)(1-v)}\,,$$

é uma cópula para todo $\theta \in [-1,1]$, satisfaz a equação diferencial (2.7) e as condições de contorno (i) à (iii). A esta família de cópulas chamamos *Família Ali-Mikhail-Haq* (veja Ali et al., 1978).

2.2 Família Plackett de Cópulas

Nosso objetivo nesta seção é apresentar a construção da família Plackett de cópulas (veja Nelsen, 1999, página 79). Sejam X e Y duas variáveis aleatórias contínuas com distribuição conjunta H e marginais F e G, respectivamente. Considere a razão

$$\theta = \frac{H(x,y)[1-F(x)-G(y)+H(x,y)]}{[F(x)-H(x,y)][G(y)-H(x,y)]} \tag{2.14}$$

$$= \frac{H(x,y)-F(x)H(x,y)-G(y)H(x,y)+H^2(x,y)}{F(x)G(y)-F(x)H(x,y)-G(y)H(x,y)+H^2(x,y)}\,. \tag{2.15}$$

Podemos notar que a diferença entre o numerador e o denominador no lado direito da expressão (2.15) é $H(x,y)-F(x)G(y)$. Assim podemos interpretar θ como uma maneira de se "medir"o quão próximo (ou distante) está a dependência entre X e Y em relação à independência. É imediato da expressão (2.15) que se X e Y são independentes, então $\theta = 1$. Suponha agora que X

e Y possuam uma relação de dependência perfeita positiva, isto é, suponha que exista uma função crescente f tal que $Y \stackrel{d}{=} f(X)$. Denotando por F a função de distribuição de X, temos

$$P(X \leq x, Y \leq y) = P(X \leq x, X \leq f^{-1}(y)) = \min\left\{F(x), F(f^{-1}(y))\right\}. \tag{2.16}$$

Escrevendo $A := -F(x)H(x,y) - G(y)H(x,y) + H^2(x,y)$, e como

$$0 \leq F(x), F(f^{-1}(y)) \leq 1 \Longrightarrow F(x)F(f^{-1}(y)) \leq \min\left\{F(x), F(f^{-1}(y))\right\}, \tag{2.17}$$

de (2.15) segue que,

$$\theta = \frac{\min\{F(x), F(f^{-1}(y))\} + A}{F(x)F(f^{-1}(y)) + A} > 1.$$

Por outro lado, se X e Y apresentam uma dependência perfeita negativa, isto é, se existe uma função decrescente g tal que $Y \stackrel{d}{=} g(X)$, então,

$$\begin{aligned} P(Y \leq y) &= P(X \geq g^{-1}(y)) \\ &= 1 - P(X \leq g^{-1}(y)) \\ &= 1 - F(g^{-1}(y)). \end{aligned} \tag{2.18}$$

Por outro lado,

$$\begin{aligned} P(X \leq x, Y \leq y) &= P(X \leq x, X \geq g^{-1}(y)) \\ &= P(X \geq g^{-1}(y) | X \leq x) P(X \leq x) \\ &= \left[1 - P(X \leq g^{-1}(y) | X \leq x)\right] P(X \leq x) \\ &= P(X \leq x) - P(X \leq x, X \leq g^{-1}(y)) \\ &= F(x) - \min\left\{F(x), F(g^{-1}(y))\right\}. \end{aligned} \tag{2.19}$$

Desta forma, substituindo (2.18) e (2.19) em (2.15), e pelo análogo de (2.17) para $F(g^{-1}(y))$ obtemos,

$$\theta = \frac{F(x) - \min\{F(x), F(g^{-1}(y))\} + A}{F(x) - F(x)F(g^{-1}(y)) + A} < 1.$$

É claro também que $\theta \geq 0$. Assim, temos $\theta = 1$ em caso de independência, $\theta > 1$ em caso de dependência positiva perfeita e $\theta < 1$ em caso de dependência negativa perfeita.

Note que no caso de independência θ não é uma função do ponto (x, y) utilizado em (2.14), já que $\theta = 1$. Porém no caso de dependência perfeita positiva e negativa, θ depende das funções f e g envolvidas e possivelmente até do ponto (x, y) escolhido. De fato, é de se esperar que, para a maioria das funções de distribuição, θ será função do ponto (x, y) utilizado para calcular (2.14). A questão que Plackett tentou responder foi a seguinte: Existe alguma função de distribuição para a qual θ é constante qualquer que seja o ponto $(x, y) \in \mathbb{R}^2$?

A resposta é afirmativa, conforme veremos adiante, e as distribuições que satisfazem esta propriedade são os membros da família de distribuições que originou a Família Plackett de cópulas.

Usando as transformações $u = F(x)$ e $v = G(y)$ e o Teorema de Sklar, podemos reescrever (2.14) como

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

onde C_θ é a cópula de X e Y. Se $\theta = 1$, resolvendo a igualdade acima obtemos, como esperado, $C_1 = \Pi$. Se $\theta \neq 1$ temos,

$$\begin{aligned} uv\theta - \theta(u + v)C_\theta(u, v) + \theta C_\theta^2(u, v) &= C_\theta(u, v) - uC_\theta(u, v) - vC_\theta(u, v) + C_\theta^2(u, v) \\ \iff (1 - \theta)C_\theta^2(u, v) &+ [(\theta - 1)(u + v) + 1]C_\theta(u, v) - uv\theta = 0. \end{aligned}$$

Resolvendo a equação quadrática acima obtemos,

$$C_\theta(u, v) = \frac{(\theta - 1)(u + v) + 1 \pm \sqrt{[1 + (\theta - 1)(u + v)]^2 - 4uv\theta(\theta - 1)}}{2(\theta - 1)}. \tag{2.20}$$

Para que C_θ em (2.20) seja *grounded*, precisamos que

$$C_\theta(u, 0) = \frac{[1 + (\theta - 1)u] \pm [1 + (\theta - 1)u]}{2(\theta - 1)} = 0, \ \ \forall\, u \in I.$$

Isto acontece somente para a raiz com o sinal "$-$"precedendo o radical. Analogamente para $C_\theta(0, v)$. Para que C_θ tenha marginais u e v, precisamos que

$$C_\theta(u, 1) = \frac{[\theta + (\theta - 1)u] \pm [\theta - (\theta - 1)u]}{2(\theta - 1)} = u, \ \ \forall\, u \in I,$$

que também só vai ocorrer para a raiz com o sinal "$-$"precedendo o radical. Analogamente para $C_\theta(1, v)$. Assim, se $\theta \neq 1$, C_θ definida em (2.20) satisfaz as condições de contorno somente se o sinal precedendo o radical é "$-$". Para provar que

$$C_\theta(u, v) = \frac{(\theta - 1)(u + v) + 1 - \sqrt{[1 + (\theta - 1)(u + v)]^2 - 4uv\theta(\theta - 1)}}{2(\theta - 1)},$$

é 2-crescente, basta provar (veja o Lema 1.5.1) que C é absolutamente contínua e que

$$\frac{\partial^2 C_\theta(u, v)}{\partial u \partial v} \geq 0.$$

Pelo Exemplo 1.5.2, temos

$$\frac{\partial C_\theta(u,v)}{\partial v} = \frac{1}{2} + \frac{u+v-1+\theta(u-v)}{2\sqrt{[1+(\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}}.$$

Denotando por $f(u) := u+v-1+\theta(u-v)$ temos,

$$\frac{\partial f(u)}{\partial u} = 1+\theta.$$

Denotando por $g(u) := 2\sqrt{[1+(\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}$, pela regra da cadeia temos

$$\frac{\partial g(u)}{\partial u} = \frac{2(\theta-1)[1+(\theta-1)(u+v)] - 4v\theta(\theta-1)}{\sqrt{[1+(\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}}.$$

Dessa forma, pela regra do quociente temos

$$\begin{aligned}
\frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} &= \frac{\partial}{\partial u}\left(\frac{1}{2} + \frac{f(u)}{g(u)}\right) \\
&= \frac{f(u)g'(u) - f'(u)g(u)}{g^2(u)}.
\end{aligned}$$

Denotando

$$B := \sqrt{[1+(\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}$$

e

$$D := 2(\theta-1)[1+(\theta-1)(u+v)] - 4uv\theta(\theta-1),$$

obtemos,

$$\begin{aligned}
\frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} &= \frac{2(1+\theta)\sqrt{B}}{4B} - \frac{1}{4B}\left[\frac{D(u+v+\theta(u-v)-1)}{\sqrt{B}}\right] \\
&= \frac{2(1+\theta)B - D(u+v+\theta(u-v)-1)}{4B^{\frac{3}{2}}} \\
&= \frac{\theta[(\theta-1)(u+v) - 2uv(\theta-1)+1]}{B^{\frac{3}{2}}} \\
&= \frac{\theta[(\theta-1)(u+v-2uv)+1]}{B^{\frac{3}{2}}}.
\end{aligned}$$

Note que, como $0 \le u, v \le 1$, então, $0 \le u+v-2uv \le 1$ e como $\theta \ge 0$ segue que,

$$\theta[(\theta-1)(u+v-2uv)+1] \ge 0.$$

Como $B \ge 0$, finalmente segue que

$$\frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} \ge 0, \quad \forall\, u, v \in I.$$

É elementar, porém bastante trabalhoso (veja Nelsen, 1999), mostrar que

$$\int_0^v \int_0^u \frac{\partial^2 C_\theta(x,y)}{\partial x \partial y}\, dxdy = C_\theta(u,v),$$

ou seja, C_θ é absolutamente contínua e, pelo Lema 1.5.1, isso mostra que C_θ é 2-crescente. Desta forma obtemos a família de cópulas Plackett, que é dada por

$$C_\theta(u,v) = \begin{cases} \frac{(\theta-1)(u+v)+1-\sqrt{[1+(\theta-1)(u+v)]^2-4uv\theta(\theta-1)}}{2(\theta-1)}, & \text{se} \quad \theta > 0, \theta \neq 1, \\ uv, & \text{se} \quad \theta = 1. \end{cases} \tag{2.21}$$

Esta família contém como casos particulares as cópulas M e W (famílias com essa propriedade são chamadas de *famílias comprensivas*). Para ver isso, escreva C_θ como,

$$C_\theta(u,v) = \frac{1}{2(\theta-1)} + \frac{u+v}{2} - \frac{1}{2}\sqrt{\left[\frac{1}{\theta} - 1 + u + v\right]^2 - \frac{4uv\theta}{\theta - 1}}\,\cdot \tag{2.22}$$

Tomando o limite para $\theta \to \infty$ em (2.22), obtemos,

$$\begin{aligned} C_\infty(u,v) &= \lim_{\theta\to\infty} \frac{1}{2(\theta-1)} + \frac{u+v}{2} - \frac{1}{2}\sqrt{\left[\frac{1}{\theta} - 1 + u + v\right]^2 - \frac{4uv\theta}{\theta - 1}} \\ &= \frac{u+v}{2} - \frac{1}{2}\sqrt{(u+v)^2 - 4uv} \\ &= \frac{u+v}{2} - \frac{1}{2}\sqrt{(u-v)^2} \\ &= \frac{u+v}{2} - \frac{|u-v|}{2} \\ &= \min\{u,v\}, \end{aligned}$$

para todo $u, v \in I$. Logo, $C_\infty = M$. Por outro lado, escrevendo

$$C_\theta(u,v) = \frac{1}{2(\theta-1)} + \frac{u+v}{2} - \frac{\sqrt{[1+(\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}}{2(\theta-1)},$$

e tomando o limite para $\theta \to 0^+$ obtemos,

$$\begin{aligned} C_0(u,v) &= \lim_{\theta\to 0^+} \frac{1}{2(\theta-1)} + \frac{u+v}{2} - \frac{\sqrt{[1+(\theta-1)(u+v)]^2 - 4uv\theta(\theta-1)}}{2(\theta-1)} \\ &= \frac{u+v-1}{2} + \frac{\sqrt{(u+v-1)^2}}{2} \\ &= \frac{u+v-1}{2} + \frac{|u+v-1|}{2} \\ &= \max\{u+v-1, 0\}, \end{aligned}$$

para todo $u, v \in I$. Logo, $C_0 = W$.

2.3 Construção de Dolati e Úbeda-Flores

A construção apresentada em Dolati e Úbeda-Flores (2005) é um pouco diferente das demais existentes na literatura e das apresentadas até aqui. A grande diferença é que a construção de Dolati e Úbeda-Flores é livre do problema de compatibilidade entre as marginais. A idéia é construir uma cópula m-dimensional a partir de marginais bivariadas quaisquer dadas (sem nos preocuparmos se são ou não compatíveis) e é baseada em uma função que definiremos adiante.

Note primeiramente que dada uma m-cópula qualquer, temos $\binom{m}{2}$ marginais bivariadas. A seguir definimos a função que será o objeto de estudo desta seção.

Definição 2.3.1. Seja $\{C_{ij} : 1 \leq i < j \leq m\}$ um conjunto de $\binom{m}{2}$ cópulas bivariadas. Definimos a *função Dolati e Úbeda-Flores* $C : I^m \to I$ por

$$C(u_1, \cdots, u_m) = \sum_{1 \leq i < j \leq m} C_{ij}(u_i, u_j) \prod_{\substack{t=1 \\ t \neq i,j}}^{m} u_t - \frac{(m-2)(m+1)}{2} \prod_{t=1}^{m} u_t. \tag{2.23}$$

A função de Dolati e Úbeda-Flores nem sempre é uma cópula. O resultado principal desta seção será apresentar uma condição necessária e suficiente para que a função (2.23) seja uma m-cópula. Antes de apresentarmos tal resultado, vamos precisar de alguns lemas. O primeiro lema mostra que a função de Dolati e Úbeda-Flores é *grounded*, tem marginais univariadas u_i, $i = 1, 2, \cdots, m$ e marginais bivariadas $C_{ij}(u_i, u_j)$, quaisquer que sejam as $\binom{m}{2}$ cópulas bivariadas dadas.

Lema 2.3.1. *Seja $\{C_{ij} : 1 \leq i < j \leq m\}$ um conjunto de $\binom{m}{2}$ cópulas bivariadas e seja C a função de Dolati e Úbeda-Flores dada por (2.23). Então,*

(i). $C(u_1, \cdots, u_{k-1}, 0, u_{k+1}, \cdots, u_m) = 0$, *quaisquer que sejam* $u_1, \cdots, u_{k-1}$, $u_{k+1}, \cdots, u_m \in I$ *e* $k \in \{1, 2, \cdots, m\}$;

(ii). $C(1, \cdots, 1, u_k, 1, \cdots, 1) = u_k$, *qualquer que seja* $u_k \in I$, $k \in \{1, 2, \cdots, m\}$;

(iii). $C(1, \cdots, 1, u_k, 1, \cdots, 1, u_l, 1, \cdots, 1) = C_{kl}(u_k, u_l)$ *quaisquer que sejam* $u_k, u_l \in I$, $k, l \in \{1, 2, \cdots, m\}$ *e* $k < l$.

Prova: (i). Sejam $u_1, \cdots, u_{k-1}, u_{k+1}, \cdots, u_m \in I$, $k \in \{1, 2, \cdots, m\}$. Considere o vetor

$$\mathbf{u} = (u_1, \cdots, u_{k-1}, 0, u_{k+1}, \cdots, u_m).$$

Então,

$$\begin{aligned} C(\mathbf{u}) &= \sum_{1\le i<j\le m} C_{ij}(u_i,u_j)\prod_{\substack{t=1\\ t\neq i,j}}^{m} u_t - \frac{(m-2)(m+1)}{2}\prod_{t=1}^{m} u_t \\ &= \sum_{\substack{i=1\\ i<j}}^{m}\sum_{\substack{j=1\\ i,j\neq k}}^{m} C_{ij}(u_i,u_j)\prod_{\substack{t=1\\ t\neq i,j}}^{m} u_t + \sum_{j=k+1}^{m} C_{kj}(u_k,u_j)\prod_{\substack{t=1\\ t\neq i,k}}^{m} u_t \\ &\quad + \sum_{i=1}^{k-1} C_{ik}(u_i,u_k)\prod_{\substack{t=1\\ t\neq k,j}}^{m} u_t - \frac{(m-2)(m+1)}{2}\prod_{t=1}^{m} u_t. \end{aligned}$$

Note que, como C_{ij} é cópula para todo $i,j=1,2,\cdots,m$ e $u_k=0$, segue que $C_{kj}(u_k,u_j)=0=C_{ik}(u_i,u_k)$. Como $i,j\neq k$, $\prod\limits_{\substack{t=1\\ t\neq i,j}}^{m} u_t$ sempre conterá o termo $u_k=0$ para todos os índices da soma dupla, logo,

$$\sum_{\substack{i=1\\ i<j}}^{m}\sum_{\substack{j=1\\ i,j\neq k}}^{m} C_{ij}(u_i,u_j)\prod_{\substack{t=1\\ t\neq i,j}}^{m} u_t = 0.$$

Como $\prod\limits_{t=1}^{m} u_t = 0$, segue que $C(\mathbf{u})=0$.

(ii). Seja $u_k\in I$, $k=1,2,\cdots,m$ e considere o vetor $\mathbf{u}=(1,\cdots,1,u_k,1,\cdots,1)\in I^m$. Precisamos calcular

$$C(\mathbf{u}) = \sum_{\substack{i=1\\ i<j}}^{m}\sum_{j=1}^{m} C_{ij}(u_i,u_j)\prod_{\substack{t=1\\ t\neq i,j}}^{m} u_t + \frac{(m-2)(m+1)}{2}\prod_{t=1}^{m} u_t.$$

Note que, para qualquer combinação de $i,j\in\{1,\cdots,m\}$ com $i<j$, temos

$$C_{ij}(u_i,u_j)\prod_{\substack{t=1\\ t\neq i,j}}^{m} u_t = \begin{cases} C_{ik}(u_i,u_k)\prod\limits_{\substack{t=1\\ t\neq i,k}}^{m} u_t = u_k, & \text{se} \quad j=k, \\ C_{kj}(u_k,u_j)\prod\limits_{\substack{t=1\\ t\neq k,j}}^{m} u_t = u_k, & \text{se} \quad i=k, \\ \underbrace{C_{ij}(u_i,u_j)}_{=\underbrace{C_{ij}(1,1)}_{=1}}\underbrace{\prod\limits_{\substack{t=1\\ t\neq i,j}}^{m} u_t}_{=u_k} = u_k, & \text{se} \quad i,j\neq k. \end{cases}$$

Assim,

$$C_{ij}(u_i, u_j) \prod_{\substack{t=1 \\ t \neq i,j}}^{m} u_t = u_k,$$

para todo $i, j \in \{1, \cdots, m\}$, com $i < j$. Logo,

$$\begin{aligned}
\sum_{\substack{i=1 \\ i<j}}^{m} \sum_{j=1}^{m} C_{ij}(u_i, u_j) \prod_{\substack{t=1 \\ t \neq i,j}}^{m} u_t &= u_k \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} 1 \\
&= u_k \left[\sum_{i=1}^{m-1} (m-i) \right] \\
&= u_k \left[m(m-1) - \sum_{i=1}^{m-1} i \right] \\
&= u_k \left[m(m-1) - \frac{m(m-1)}{2} \right] \\
&= \frac{m(m-1)}{2} \; u_k.
\end{aligned}$$

Como

$$\frac{(m-2)(m+1)}{2} \prod_{t=1}^{m} u_t = \frac{(m-2)(m+1)}{2} \; u_k,$$

segue que,

$$\begin{aligned}
C(\mathbf{u}) &= \frac{m(m-1)}{2} \; u_k - \frac{(m-2)(m+1)}{2} \; u_k \\
&= \left(\frac{m^2 - m - m^2 + m + 2}{2} \right) u_k \\
&= u_k.
\end{aligned}$$

(iii). Sejam u_k, $u_l \in I$, para $k < l$. Considere o vetor

$$\mathbf{u} = (1, \cdots, 1, u_k, 1, \cdots, 1, u_l, 1, \cdots, 1).$$

Temos que calcular

$$\begin{aligned}
C(\mathbf{u}) &= \sum_{1 \leq i < j \leq m} C_{ij}(u_i, u_j) \prod_{\substack{t=1 \\ t \neq i,j}}^{m} u_t - \frac{(m-2)(m+1)}{2} \prod_{t=1}^{m} u_t \\
&= \sum_{\substack{i=1 \\ i<j}}^{m} \sum_{j=1}^{m} C_{ij}(u_i, u_j) \prod_{\substack{t=1 \\ t \neq i,j}}^{m} u_t - \frac{(m-2)(m+1)}{2} \prod_{t=1}^{m} u_t. \qquad (2.24)
\end{aligned}$$

Note primeiramente que,

$$\begin{aligned}
\sum_{\substack{i=1 \\ i<j}}^{m}\sum_{j=1}^{m} C_{ij}(u_i,u_j)\prod_{\substack{t=1 \\ t\neq i,j}}^{m} u_t &= \sum_{i=1}^{m-1}\sum_{j=i+1}^{m} C_{ij}(u_i,u_j)\prod_{\substack{t=1 \\ t\neq i,j}}^{m} u_t \\
&= \sum_{\substack{i=1 \\ i\neq k,l}}^{m-1}\sum_{j=i+1}^{m} C_{ij}(u_i,u_j)\prod_{\substack{t=1 \\ t\neq i,j}}^{m} u_t + \sum_{j=k+1}^{m} C_{kj}(u_k,u_j)\prod_{\substack{t=1 \\ t\neq k,j}}^{m} u_t \\
&\quad + \sum_{j=l+1}^{m} C_{lj}(u_l,1)\prod_{\substack{t=1 \\ t\neq l,j}}^{m} u_t \\
&= \sum_{\substack{i=1 \\ i\neq k,l}}^{m-1}\sum_{j=i+1}^{m} C_{ij}(u_i,u_j)\prod_{\substack{t=1 \\ t\neq i,j}}^{m} u_t + \sum_{\substack{j=k+1 \\ j\neq l}}^{m} u_k\prod_{\substack{t=1 \\ t\neq k,j}}^{m} u_t \\
&\quad + C_{kl}(u_k,u_l) + \sum_{j=l+1}^{m} u_l\prod_{\substack{t=1 \\ t\neq l,j}}^{m} u_t \\
&= \sum_{\substack{i=1 \\ i\neq k,l}}^{m-1}\sum_{j=i+1}^{m} u_k u_l + \sum_{\substack{j=k+1 \\ j\neq l}}^{m} u_k u_l + C_{kl}(u_k,u_l) + \sum_{j=l+1}^{m} u_k u_l \\
&= C_{kl}(u_k,u_l) + \left(\sum_{\substack{i=1 \\ i\neq k,l}}^{m-1}\sum_{j=i+1}^{m} 1 + \sum_{\substack{j=k+1 \\ j\neq l}}^{m} 1 + \sum_{j=l+1}^{m} 1\right) u_k u_l. \qquad (2.25)
\end{aligned}$$

Note que,

$$\sum_{\substack{j=k+1 \\ j\neq l}}^{m} 1 = m-k-1 \quad \text{e} \quad \sum_{j=l+1}^{m} 1 = m-l, \qquad (2.26)$$

e que

$$\begin{aligned}
\sum_{\substack{i=1 \\ i\neq k,l}}^{m-1}\sum_{j=i+1}^{m} 1 &= \sum_{\substack{i=1 \\ i\neq k,l}}^{m-1}(m-(i+1)+1) = \sum_{\substack{i=1 \\ i\neq k,l}}^{m-1}(m-i) \\
&= m(m-3) - \left[\sum_{i=1}^{m-1} i - (k+l)\right] \\
&= m(m-3) - \left[\frac{(m-1+1)(m-1-1+1)}{2} - k - l\right] \\
&= m(m-3) + k + l - \frac{m(m-1)}{2}\cdot \qquad (2.27)
\end{aligned}$$

Substituindo (2.27) e (2.26) em (2.25), obtemos,

$$\begin{aligned}
C(\mathbf{u}) &= C_{kl}(u_k,u_l) + u_k u_l\left(m(m-3)+k+l-\frac{m(m-1)}{2}+m-k-1+m-l\right)\\
&= C_{kl}(u_k,u_l) + u_k u_l\left(\frac{2m^2-6m-m^2+m}{2}+2m-1\right)\\
&= C_{kl}(u_k,u_l) + u_k u_l\left(\frac{m^2-m-2}{2}\right)\\
&= C_{kl}(u_k,u_l) + \frac{(m-2)(m+1)}{2}\, u_k u_l. \qquad (2.28)
\end{aligned}$$

Substituindo (2.28) em (2.24), obtemos,

$$\begin{aligned}
C(\mathbf{u}) &= C_{kl}(u_k,u_l) + \frac{(m-2)(m+1)}{2}\, u_k u_l - \frac{(m-2)(m+1)}{2}\prod_{t=1}^{m} u_t\\
&= C_{kl}(u_k,u_l) + \frac{(m-2)(m+1)}{2}\, u_k u_l - \frac{(m-2)(m+1)}{2}\, u_k u_l\\
&= C_{kl}(u_k,u_l).
\end{aligned}$$

Isto completa a prova. □

Nosso próximo passo é provar um lema técnico necessário à prova do Teorema 2.3.1, que é o resultado principal desta seção. Seja m um inteiro positivo e $R=[u_1,v_1]\times\cdots\times[u_m,v_m]\subseteq I^m$. Lembre da Seção 1.6 que o conjunto dos vértices de R é dado por

$$A_m(R)=\{\mathbf{w}=(w_1,\cdots,w_m)\in R \,|\, w_k=u_k \text{ ou } w_k=v_k,\ k=1,\cdots,m\}. \qquad (2.29)$$

Lembre também da Seção 1.6 que a função csgn($\mathbf{w}$) é definida por

$$\text{csgn}(\mathbf{w})=\begin{cases} 1, & \text{se } \mathbf{w} \text{ possui um número par de } u_i\text{'s, ou nenhum,}\\ -1, & \text{se } \mathbf{w} \text{ possui um número ímpar de } u_i\text{'s.}\end{cases}$$

Lema 2.3.2. *Sejam m um inteiro positivo e $R=[u_1,v_1]\times\cdots\times[u_m,v_m]$ um retângulo em I^m. Então,*

$$\sum_{\boldsymbol{w}\in A_m(R)} \text{csgn}(\boldsymbol{w})\prod_{t=1}^{m} w_t=\prod_{t=1}^{m}(v_t-u_t). \qquad (2.30)$$

Prova: A prova será feita por indução em m, a dimensão do vetor $\mathbf{w}$. Seja $R=[u_1,v_1]\times\cdots\times[u_m,v_m]$ um retângulo em I^m. Suponha que $m=2$. Então,

$$A_2(R)=\{(u_1,u_2),(u_1,v_2),(v_1,u_2),(v_1,v_2)\}. \qquad (2.31)$$

Assim,

$$\begin{aligned}\sum_{\mathbf{w}\,\in A_2(R)} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{2} w_t &= u_1u_2 + v_1v_2 - u_1v_2 - u_2v_1 \\ &= (v_1-u_1)(v_2-u_2) \\ &= \prod_{t=1}^{2}(v_t-u_t).\end{aligned}$$

Suponha que o resultado seja válido para $m = k-1$, para algum k inteiro positivo. Assim, para todo $R' = [u_1, v_1] \times \cdots \times [u_{k-1}, v_{k-1}] \subseteq I^{k-1}$,

$$\sum_{\mathbf{w}\in A_{k-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k-1} w_t = \prod_{t=1}^{k-1}(v_t-u_t). \tag{2.32}$$

Agora dado o retângulo $R' = [u_1, v_1] \times \cdots \times [u_{k-1}, v_{k-1}] \subseteq I^{k-1}$ considere $R = R' \times [u_k, v_k] \in I^k$. Defina os conjuntos

$$\begin{aligned} A_k^{v_k}(R) &= \{\mathbf{w} = (w_1, \cdots, w_{k-1}, v_k) \in R \,|\, w_j = u_j \text{ ou } w_j = v_j, \\ &\quad\; j = 1, 2, \cdots, k-1\}, \\ A_k^{u_k}(R) &= \{\mathbf{w} = (w_1, \cdots, w_{k-1}, u_k) \in R \,|\, w_j = u_j \text{ ou } w_j = v_j, \\ &\quad\; j = 1, 2, \cdots, k-1\}. \end{aligned}$$

Note que, $A_k(R) = A_k^{u_k}(R) \cup A_k^{v_k}(R)$ e que $A_k^{u_k}(R) \cap A_k^{v_k}(R) = \emptyset$. Note também que,

$$\sum_{\mathbf{w}\in A_k(R)} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k} w_t = \sum_{\mathbf{w}\in A_k^{v_k}(R)} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k-1} w_t v_k + \sum_{\mathbf{w}\in A_k^{u_k}(R)} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k-1} w_t u_k.$$

Observe que, da definição dos conjuntos $A_k^{v_k}(R)$ e $A_k^{u_k}(R)$ e da definição da função $\operatorname{csgn}(\cdot)$, segue que,[1]

$$\sum_{\mathbf{w}\in A_k^{v_k}(R)} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k-1} w_t v_k = v_k \sum_{\mathbf{w}\in A_{k-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k-1} w_t,$$

e

$$\sum_{\mathbf{w}\in A_k^{u_k}(R)} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k-1} w_t u_k = -u_k \sum_{\mathbf{w}\in A_{k-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k-1} w_t.$$

[1] Veja também o Exemplo 1.6.5.

Logo,

$$
\begin{aligned}
\sum_{\mathbf{w}\in A_k(R)} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k} w_t = &\ v_k \sum_{\mathbf{w}\in A_{k-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k-1} w_t - u_k \sum_{\mathbf{w}\in A_{k-1}(R')} \operatorname{csgn}(\mathbf{w}) \prod_{t=1}^{k-1} w_t \\
=^{2} &\ v_k \left(\prod_{t=1}^{k-1} (v_t - u_t) \right) - u_k \left(\prod_{t=1}^{k-1} (v_t - u_t) \right) \\
= &\ \prod_{t=1}^{k} (v_t - u_t).
\end{aligned}
$$

O que completa a prova. □

O próximo teorema é o resultado principal desta seção. Nele é apresentado uma condição necessária e suficiente para que a equação (2.23) defina uma cópula m-dimensional. O teorema é enunciado em Dolati e Úbeda-Flores (2005), mas a prova dada pelos autores deixa várias lacunas para o leitor que formalizamos ao longo desta seção.

Teorema 2.3.1. *Seja $\{C_{ij} : 1 \leq i < j \leq m\}$ um conjunto de $\binom{m}{2}$ cópulas bivariadas e seja C a função de Dolati e Úbeda-Flores dada por (2.23). Então, C é uma m-cópula cujas marginais bivariadas são C_{ij} se, e somente se,*

$$
\sum_{\substack{i=1 \\ i<j}}^{m} \sum_{j=1}^{m} \frac{V_{C_{ij}}([u_i, v_i] \times [u_j, v_j])}{(v_i - u_i)(v_j - u_j)} \geq \frac{(m-2)(m+1)}{2}, \tag{2.33}
$$

quaisquer que sejam $u_k, v_k \in I$, $k = 1, 2, \cdots, m$, tais que $u_k < v_k$.

Prova: Pelo Lema 2.3.1 basta mostrar que, dado um retângulo $R \subseteq I^m$,

$$
V_C(R) \geq 0 \iff \sum_{\substack{i=1 \\ i<j}}^{m} \sum_{j=1}^{m} \frac{V_{C_{ij}}([u_i, v_i] \times [u_j, v_j])}{(v_i - u_i)(v_j - u_j)} \geq \frac{(m-2)(m+1)}{2},
$$

quaisquer que sejam $u_k, v_k \in I$, $k = 1, 2, \cdots, m$, tais que $u_k < v_k$.
Primeiramente note que, dado $R = [u_1, v_1] \times \cdots \times [u_m, v_m] \subseteq I^m$ e $A_m(R)$ dado por (2.29), temos

$$
\begin{aligned}
V_C(R) &= \sum_{\mathbf{w}\in A_m(R)} \operatorname{csgn}(\mathbf{w}) C(\mathbf{w}) \\
&= \sum_{\mathbf{w}\in A_m(R)} \operatorname{csgn}(\mathbf{w}) \left[\sum_{\substack{i=1 \\ i<j}}^{m} \sum_{j=1}^{m} C_{ij}(w_i, w_j) \prod_{\substack{t=1 \\ t\neq i,j}}^{m} w_t - \frac{(m-2)(m+1)}{2} \prod_{t=1}^{m} w_t \right]
\end{aligned}
$$

[2]Pela hipótese de indução (2.32).

$$= \sum_{\substack{i=1 \\ i<j}}^{m}\sum_{j=1}^{m}\sum_{\mathbf{w}\in A_m(R)} \operatorname{csgn}(\mathbf{w})\, C_{ij}(w_i,w_j)\prod_{\substack{t=1\\ t\neq i,j}}^{m} w_t -$$

$$-\frac{(m-2)(m+1)}{2}\left(\sum_{\mathbf{w}\in A_m(R)} \operatorname{csgn}(\mathbf{w})\prod_{t=1}^{m} w_t\right). \qquad (2.34)$$

Para cada $i,j \in \{1,2,\cdots,m\}$, defina,

$$A_2(R_{ij}) = \{\mathbf{w} = (w_i,w_j)\,|\, w_k = u_k \text{ ou } w_k = v_k, k = i,j\}.$$

Note que, para cada $i,j \in \{1,2,\cdots,m\}$, com $i<j$, $A_2(R_{ij})$ é o conjunto dos vértices do retângulo $R_{ij} = [u_i,v_i]\times[u_j,v_j]$. Vamos denotar por $B := A_m(R) - A_2(R_{ij})$ o conjunto dos vértices contidos no retângulo $\prod_{\substack{k=1\\ k\neq i,j}}^{m}[u_k,v_k]$, onde o produto deve ser interpretado como produto cartesiano. Formalmente, o conjunto B pode ser definido como

$$\begin{aligned} B = \{&\mathbf{w} = (w_1,\cdots,w_{i-1},w_{i+1},\cdots,w_{j-1},w_{j+1},\cdots,w_m)\,|\, w_k = u_k \\ &\text{ou } \ w_k = v_k, k = 1,\cdots,i-1,i+1,\cdots,j-1,j+1,\cdots,m\}. \end{aligned}$$

Observe que $A_m(R) = A_2(R_{ij}) \cup B$ e $A_2(R_{ij}) \cap B = \emptyset$, $i,j \in \{1,2,\cdots,m\}$, com $i<j$. Defina,

$$\underset{A_2(R_{ij})}{\operatorname{csgn}}(\mathbf{w}) = \operatorname{csgn}(\mathbf{w})\big|_{\mathbf{w}\in A_2(R_{ij})} \ ,$$

e

$$\underset{B}{\operatorname{csgn}}(\mathbf{w}) = \operatorname{csgn}(\mathbf{w})\big|_{\mathbf{w}\in B} \ .$$

É claro que,

$$\operatorname{csgn}(\mathbf{w}) = \underset{A_2(R_{ij})}{\operatorname{csgn}}(\mathbf{w})\ \underset{B}{\operatorname{csgn}}(\mathbf{w}).$$

Então, pelo Lema 2.3.2, podemos reescrever (3.8) como

$$\begin{aligned} V_C(R) &= \sum_{\substack{i=1 \\ i<j}}^{m}\sum_{j=1}^{m}\left\{\sum_{\mathbf{w}\in A_m(R)}\left[\underset{B}{\operatorname{csgn}}(\mathbf{w})\prod_{\substack{t=1\\ t\neq i,j}}^{m} w_t\right]\left[\underset{A_2(R_{ij})}{\operatorname{csgn}}(\mathbf{w})C_{ij}(w_i,w_j)\right]\right\} \\ &\quad -\frac{(m-2)(m+1)}{2}\prod_{t=1}^{m}(v_t-u_t) \\ &= \sum_{\substack{i=1 \\ i<j}}^{m}\sum_{j=1}^{m} V_{C_{ij}}(R_{ij})\prod_{\substack{t=1\\ t\neq i,j}}^{m}(v_t-u_t) - \frac{(m-2)(m+1)}{2}\prod_{t=1}^{m}(v_t-u_t). \end{aligned}$$

Logo,

$$\begin{aligned}V_C(R) &= \left[\sum_{\substack{i=1\\i<j}}^{m}\sum_{j=1}^{m}\frac{V_{C_{ij}}(R_{ij})}{(v_i-u_i)(v_j-u_j)}\right]\prod_{t=1}^{m}(v_t-u_t)-\frac{(m-2)(m+1)}{2}\prod_{t=1}^{m}(v_t-u_t)\\ &= \left[\sum_{\substack{i=1\\i<j}}^{m}\sum_{j=1}^{m}\frac{V_{C_{ij}}(R_{ij})}{(v_i-u_i)(v_j-u_j)}-\frac{(m-2)(m+1)}{2}\right]\prod_{t=1}^{m}(v_t-u_t).\end{aligned}$$

Como $\prod_{t=1}^{m}(v_t-u_t)\geq 0$ e $(v_i-u_i)(v_j-u_j)\geq 0$, segue que,

$$V_C(R)\geq 0 \iff \sum_{\substack{i=1\\i<j}}^{m}\sum_{j=1}^{m}\frac{V_{C_{ij}}\left([u_i,v_i]\times[u_j,v_j]\right)}{(v_i-u_i)(v_j-u_j)}\geq\frac{(m-2)(m+1)}{2},$$

quaisquer que sejam $u_k, v_k \in I$, $k = 1, 2, \cdots, m$, tais que $u_k < v_k$, o que prova o teorema. $\square$

2.4 Família Arquimediana de Cópulas

Nesta seção apresentamos a construção da conhecida Família Arquimediana de Cópulas. Antes de definirmos as cópulas que pertencem a esta família, precisamos de alguns conceitos e resultados que apresentamos a seguir.

Definição 2.4.1. Seja $\varphi : I \to \mathbb{R}^+$ uma função contínua, decrescente e tal que $\varphi(1) = 0$. A *pseudo-inversa* de φ é a função

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

Note que $\varphi^{[-1]}$ é contínua, não-decrescente em $\mathbb{R}^+$, decrescente em $[0, \varphi(0)]$, $\varphi^{[-1]}(\varphi(u)) = u$ em I e

$$\begin{aligned}\varphi(\varphi^{[-1]}(t)) &= \begin{cases} t, & \text{se} \quad 0 \leq t \leq \varphi(0),\\ \varphi(0), & \text{se} \quad t \geq \varphi(0),\end{cases}\\ &= \min\{t, \varphi(0)\}.\end{aligned}$$

Além disso, se $\lim\limits_{t\to 0^+}\varphi(t) = \infty$, então $\varphi^{[-1]} = \varphi^{-1}$. Dada uma função $\varphi : I \to \mathbb{R}^+$ contínua, decrescente e tal que $\varphi(1) = 0$, defina a função $C : I^2 \to I$ por

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)). \tag{2.35}$$

Estamos interessados em estabelecer condições para que a função definida em (2.35) seja uma cópula. A resposta é dada no próximo teorema, cuja prova é adaptada de Nelsen (1999), página 91.

Teorema 2.4.1. *Seja $C : I^2 \to I$ dada por (2.35) para alguma função $\varphi : I \to \mathbb{R}^+$ contínua, decrescente e tal que $\varphi(1) = 0$. Então, C é uma cópula se, e somente se, φ é convexa.*

Prova: Vamos iniciar provando que C satisfaz as condições de contorno (1.2) e (1.3) de cópulas. Com efeito, dado $u \in I$, temos

$$C(u,0) = \varphi^{[-1]}(\varphi(u) + \varphi(0)) = \varphi^{-1}(\varphi(0)) = 0$$

e

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

Por simetria, $C(0,v) = 0$ e $C(1,v) = v$, qualquer que seja $v \in I$. Resta provar que C é 2-crescente se, e somente se, φ é convexa. A prova deste fato será feita em duas etapas. Primeiramente vamos provar que C é 2-crescente se, e somente se,

$$C(u_2,v) - C(u_1,v) \leq u_2 - u_1, \tag{2.36}$$

quaisquer que sejam $u_1, u_2, v \in I$, com $u_1 \leq u_2$. Note que (2.36) é equivalente a $V_C([u_1,u_2] \times [v,1])$ que é positivo sempre que C for 2-crescente. Reciprocamente, suponha que (2.36) valha. Sejam $v_1, v_2 \in I$ tais que $v_1 \leq v_2$. Então,

$$C(0,v_1) \leq v_1 \leq v_2 \leq C(1,v_2).$$

Como φ e $\varphi^{[-1]}$ são contínuas, C é contínua. Assim, existe $t \in I$ tal que $C(t,v_2) = v_1$ e isto implica que $\varphi(v_2) + \varphi(t) = \varphi(v_1)$. Então,

$$\begin{aligned}
C(u_2,v_1) - C(u_1,v_1) = & \\
= \ & \varphi^{[-1]}(\varphi(u_2) + \varphi(v_1)) - \varphi^{[-1]}(\varphi(u_1) + \varphi(v_1)) \\
= \ & \varphi^{[-1]}(\varphi(u_2) + \varphi(v_2) + \varphi(t)) - \varphi^{[-1]}(\varphi(u_1) + \varphi(v_2) + \varphi(t)) \\
= \ & C(C(u_2,v_2),t) - C(C(u_1,v_2),t) \\
\leq^3 \ & C(u_2,v_2) - C(u_1,v_2).
\end{aligned}$$

Logo, $V_C([u_1,u_2] \times [v_1,v_2]) \geq 0$ e C é 2-crescente. Vamos provar agora que (2.36) vale se, e somente se, φ é convexa. Para isso, note que φ é convexa se, e somente se, $\varphi^{[-1]}$ é convexa. Note também que (2.36) é equivalente a

$$u_1 + \varphi^{[-1]}(\varphi(u_2) + \varphi(v)) \leq u_2 + \varphi^{[-1]}(\varphi(u_1) + \varphi(v)),$$

com $u_1 \leq u_2$. Escreva $a = \varphi(u_1)$, $b = \varphi(u_2)$ e $c = \varphi(v)$. Então, $a \geq b$, $c \geq 0$ e

$$\varphi^{[-1]}(a) + \varphi^{[-1]}(b+c) \leq \varphi^{[-1]}(b) + \varphi^{[-1]}(a+c). \tag{2.37}$$

[3] Por (2.36).

Se (2.36) vale então, (2.37) vale e, para $0 \leq s < t \in \mathbb{R}^+$, escrevendo $a = \frac{s+t}{2}$, $b = s$ e $c = \frac{t-s}{2}$ e substituindo em (2.37), obtemos

$$\varphi^{[-1]}\left(\frac{s+t}{2}\right) \leq \frac{\varphi^{[-1]}(s) + \varphi^{[-1]}(t)}{2}. \tag{2.38}$$

Como $\varphi^{[-1]}$ é contínua, (2.38) prova que $\varphi^{[-1]}$ é convexa. Reciprocamente, suponha que $\varphi^{[-1]}$ seja convexa. Dados $a, b, c \in I$ tais que $a \geq b$ e $c \leq 0$, defina $t = \frac{a-b}{a-b+c}$. Assim, $a = (1-t)b + t(a+c)$ e $b + c = tb + (1-t)(a+c)$, e assim

$$\varphi^{[-1]}(a) \leq (1-t)\varphi^{[-1]}(b) + t\varphi^{[-1]}(a+c)$$

e

$$\varphi^{[-1]}(b+c) \leq t\varphi^{[-1]}(b) + (1-t)\varphi^{[-1]}(a+c).$$

Somando tudo obtemos,

$$\varphi^{[-1]}(a) + \varphi^{[-1]}(b+c) \leq \varphi^{[-1]}(b) + \varphi^{[-1]}(a+c),$$

que é equivalente a (2.37) e isto completa a prova. □

O resultado acima pode ser utilizado como um método indireto de construção de cópulas. Para isso, basta escolher uma função $\varphi : I \to \mathbb{R}^+$ contínua, convexa, decrescente e tal que $\varphi(1) = 0$ para que a função definida em (2.35) seja uma cópula. As cópulas obtidas desta maneira recebem o nome de cópulas *Arquimedianas*.

Definição 2.4.2. Dizemos que uma cópula C é *Arquimediana* se pode ser escrita como

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

quaisquer que sejam $u, v \in I$, para alguma função $\varphi : I \to \mathbb{R}^+$ contínua, convexa, decrescente e tal que $\varphi(1) = 0$. A função φ é dita ser o *gerador* da cópula C.

Uma explicação para o porquê do termo “Arquimediana” pode ser encontrada em Nelsen (1999) página 98. Para mais detalhes veja Nelsen (1999), capítulo 4.

Exemplo 2.4.1. Tomando-se $\varphi(t) = -\ln(t)$ obtemos a cópula independência $\Pi(u, v) = uv$.

Exemplo 2.4.2. Tomando-se $\varphi(t) = \ln\left(\frac{1-\theta(1-t)}{t}\right)$, obtemos a família Ali-Mikhail-Haq de cópulas, dada por

$$C(u, v) = \frac{uv}{1 - \theta(1-u)(1-v)}.$$

Exemplo 2.4.3. Tomando-se $\varphi(t) = \ln(1 - \theta \ln(t))$ obtemos a família de cópulas Gumbel, dada por

$$C(u, v) = uv \exp\{-\theta \ln u \ln v\}.$$

Capítulo 3

Cópulas e Movimento Browniano

Neste capítulo apresentaremos alguns resultados envolvendo cópulas e o Movimento Browniano Padrão. A idéia básica é calcularmos alguns funcionais do Movimento Browniano e usá-los para obter as respectivas cópulas associadas. Iniciamos com uma breve introdução aos processos Markovianos a tempo contínuo.

3.1 Processos Markovianos a Tempo Contínuo

Nesta seção revisaremos rapidamente alguns conceitos de processos de Markov a tempo contínuo. Iniciamos com sua definição.

Definição 3.1.1. (*Processo de Markov a Tempo Contínuo*). Um processo estocástico $\{X_t\}_{t\geq 0}$ é dito ser um *Processo de Markov a tempo contínuo* se para $k \geq 1$, $0 \leq t_1 \leq \cdots \leq t_k \leq u$ e B boreliano em $\mathbb{R}$,

$$P(X_u \in B | X_{t_1}, \cdots, X_{t_k}) = P(X_u \in B | X_{t_k}). \tag{3.1}$$

A condição (3.1) é chamada de *propriedade de Markov*. Sejam B_1, B_2 borelianos em $\mathbb{R}$ e $s < t < u$. De (3.1) temos que,

$$\begin{aligned} P(X_s \in B_1, X_u \in B_2 | X_t) &= P(X_u \in B_2 | X_t, X_s \in B_1) P(X_s \in B_1 | X_t) \\ &= P(X_u \in B_2 | X_t) P(X_s \in B_1 | X_t). \end{aligned} \tag{3.2}$$

De (3.2) podemos concluir que a propriedade de Markov (3.1) informalmente estabelece que, dado o presente, o futuro e o passado de um processo Markoviano são independentes um do outro.

Definição 3.1.2. Um processo de Markov é dito ser *homogêneo* (ou ter *incrementos independentes*) se, para quaisquer $0 \leq s \leq t$ e B boreliano em $\mathbb{R}$,

$$P(X_t \in B | X_s = x) = P(X_{t-s} \in B | X_0 = x). \tag{3.3}$$

Intuitivamente a condição (3.3) significa que o comportamento do processo no tempo $t+s$ dado que o processo iniciou em s é probabilisticamente o mesmo se olharmos para o processo no tempo t dado que o processo iniciou em 0.

Definição 3.1.3. (*Tempo de Parada*). Uma variável aleatória $\tau : \Omega \to [0, \infty]$ é chamada de *tempo de parada* se para quaisquer $\omega_1, \omega_2 \in \Omega$ tais que $\omega_1(s) = \omega_2(s)$, para todo $0 \leq s \leq t$ e $\tau(\omega_1) \leq t$, implica $\tau(\omega_1) = \tau(\omega_2)$.

Considere $\{X_t\}_{t \geq 0}$ um processo Markoviano homogêneo com espaço de estados arbitrário $\{S, \mathcal{S}\}$, onde S é o conjunto de valores do processo e $\mathcal{S}$ é a σ-álgebra gerada por S. A propriedade de Markov e a homogeneidade do processo implicam que, para quaisquer x, $x_j \in S$, $j = 1, \cdots, m$, $B_i \in \mathcal{S}$, $i = 1, \cdots, n$, $0 \leq s_1 < s_2 < \cdots < s_m < s$ e $0 < t_1 < \cdots < t_n$,

$$\begin{aligned} P(X_{t_1+s} \in B_1, \cdots, X_{t_n+s} \in B_n | X_{s_1} = x_1, \cdots, X_{s_m} = x_m, X_s = x) = \\ = P(X_{t_1} \in B_1, \cdots, X_{t_n} \in B_n | X_0 = x). \end{aligned} \tag{3.4}$$

Parece intuitivamente razoável esperar que (3.4) seja válida quando trocamos s por um tempo de parada τ qualquer. Isto porém não é verdadeiro em geral[1]. Processos Markovianos que possuem esta propriedade são ditos ter a *propriedade forte de Markov.* Muitas vezes um processo Markoviano que possua a propriedade forte de Markov é dito ser um *Processo Forte de Markov.* Formalmente,

Definição 3.1.4. Um processo Markoviano $\{X_t\}_{t \geq 0}$ com espaço de estados $\{S, \mathcal{S}\}$ é dito ter a *propriedade forte de Markov* se, para quaisquer x, $x_j \in S$, $j = 1, \cdots, m$, $B_i \in \mathcal{S}$, $i = 1, \cdots, n$, $0 \leq s_1 < s_2 < \cdots < s_m < \tau$ e $0 < t_1 < \cdots < t_n$,

$$\begin{aligned} P(X_{t_1+\tau} \in B_1, \cdots, X_{t_n+\tau} \in B_n | X_{s_1} = x_1, \cdots, X_{s_m} = x_m, X_\tau = x) = \\ = P(X_{t_1} \in B_1, \cdots, X_{t_n} \in B_n | X_0 = x), \end{aligned}$$

qualquer que seja o tempo de parada τ.

Reconhecer que um processo Markoviano possui a propriedade forte de Markov pela definição não é uma tarefa fácil. Existem várias caracterizações para Processos Fortes de Markov disponíveis na literatura. Referimos ao leitor interessado o capítulo 14 (seção 4) de Karlin e Taylor (1981) e a seção 8.6 de Todorovic (1992).

[1]Para o caso em que $\{X_t\}_{t \in T}$ é um processo de Markov a tempo discreto sempre vale a propriedade forte de Markov. Para mais detalhes veja Karlin e Taylor (1975).

3.2 Movimento Browniano

Segundo Lipster e Shiryayev (1977), página 30, “Na classe dos processos com incrementos independentes e estacionários, o Movimento Browniano é a peça chave”. Sem dúvida, um dos processos mais importantes e estudados é o Movimento Browniano. Nesta seção faremos uma breve revisão dos principais conceitos envolvendo Movimento Browniano e resultados necessários para as seções que seguem. Iniciamos com sua definição.

Definição 3.2.1. (*Movimento Browniano*). O *Movimento Browniano* ou *Processo de Wiener* é um processo estocástico $\{X_t\}_{t\geq 0}$ com as seguintes propriedades:

(i). Os incrementos $X_{t+h} - X_t$ têm distribuição normal com média zero e variância $\sigma^2 h$, onde $\sigma > 0$ é um parâmetro fixo, isto é, $X_{t+h} - X_t \sim N(0, \sigma^2 h)$;

(ii). Para cada par de intervalos de tempos $[t_1, t_2]$ e $[t_3, t_4]$, tais que, $t_1 < t_2 \leq t_3 < t_4$, os incrementos $X_{t_4} - X_{t_3}$ e $X_{t_2} - X_{t_1}$ são variáveis aleatórias independentes com distribuição dada no item (i);

(iii). $X_0 = 0$ e X_t é contínua em $t = 0$ com probabilidade 1.

Pode-se provar a existência de um processo satisfazendo as condições da Definição 3.2.1. A construção é bastante técnica e não será feita aqui. O leitor interessado poderá encontrar a construção do Movimento Browniano no capítulo 2 (seção 9) de Billingsley (1968) e na seção 37 de Billingsley (1995). Karatzas e Shreve (1988) apresentam, nas seções 2.2 e 2.3, duas construções distintas para o Movimento Browniano. Veja também o capítulo 1 de Revuz e Yor (1994).

Dizemos que X_t é um *Movimento Browniano Padrão* se for um Movimento Browniano com $\sigma^2 = 1$. Neste caso,

$$P(X_{s+t} - X_s \leq x) = P(X_t \leq x) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{x} e^{-\frac{u^2}{2t}} du = \Phi\left(\frac{x}{\sqrt{t}}\right), \quad (3.5)$$

para todo $s, t > 0$, onde $\Phi(\cdot)$ é a função de distribuição de uma variável aleatória $\mathcal{N}(0, 1)$.

Um resultado clássico em processos estocásticos é o fato de que as trajetórias de um Movimento Browniano são funções contínuas mas não-diferenciáveis em quase toda a parte. O leitor interessado pode encontrar este e outros aspectos das trajetórias do Movimento Browniano no capítulo 7 de Karlin e Taylor (1975), no capítulo 15 de Karlin e Taylor (1981), na seção 3.4 de Todorovic (1992), e na seção 37 de Billingsley (1995). Veja também o capítulo 1 de Revuz e Yor (1994).

Não é difícil provar que o Movimento Browniano é um processo Gaussiano, ou seja, qualquer vetor finito-dimensional $(X_{t_1}, \cdots, X_{t_n})$ de um Movimento Browniano possui distribuição conjunta normal multivariada. Uma prova deste fato pode ser encontrada em Todorovic (1992), página 98. Uma outra prova mais elegante (via martingales) pode ser encontrada em Karlin e Taylor (1975), página 376.

Devido aos incrementos independentes, é imediato que o Movimento Browniano satisfaz a propriedade de Markov, sendo, portanto, um processo Markoviano. Além disso, o Movimento Browniano possui a propriedade forte de Markov. Este importante fato é demasiado profundo para ser provado aqui, por isso referenciamos ao leitor interessado a seção 2.6 Karatzas e Shreve (1988). Assim, o Movimento Browniano é um processo estocástico a tempo contínuo, homogêneo, Gaussiano, começando em 0 e que possui a Propriedade Forte de Markov.

3.3 Cópula do Movimento Browniano e seu Máximo

Nesta seção nosso objetivo central será determinar a cópula associada ao vetor $(X_t, \max\limits_{0 \leq s \leq t} X_s)$. Esta distribuição conjunta é importante no estudo de processos que podem ser modelados pelo Movimento Browniano quando se deseja analisar a relação entre o processo num dado instante de tempo t e o que ocorreu em termos do máximo do processo até este tempo t.

Para fixar idéias, vamos pensar que desejamos investir em um portfólio de ações e que, de alguma forma, sabemos que estas ações atingem um certo patamar máximo apenas uma vez ao dia. Vamos supor que este portfólio seja bem modelado por um Movimento Browniano (embora isto normalmente não ocorra). Se desejamos maximizar o ganho, seria interessante ter alguma informação a respeito do instante em que, com probabilidade alta, este patamar seja alcançado. Neste caso, conhecer o comportamento do vetor $(X_t, \max\limits_{0 \leq s \leq t} X_s)$ pode ser de grande valia, já que ele nos fornece, em termos de máximo, informações sobre o que ocorreu até um certo instante t em relação ao que está ocorrendo exatamente no tempo t. Conhecer a cópula associada também será de grande importância, já que toda a estrutura de dependência do vetor em questão fica unicamente caracterizada por ela.

Antes de determinarmos a distribuição conjunta de X_t e $\max\limits_{0 \leq s \leq t} X_s$, precisamos determinar as distribuições de X_t e de $\max\limits_{0 \leq s \leq t} X_s$. Para a variável aleatória $\max\limits_{0 \leq s \leq t} X_s$, aplicando o princípio da reflexão (veja Karlin e Taylor, 1975, página 345) obtemos,

$$\begin{aligned} G_t(y) = P(\max_{0\leq s\leq t} X_s \leq y) &= 1 - P(\max_{0\leq s\leq t} X_s \geq y) \\ &= 1 - 2P(X_t \geq y) \\ &= 1 - 2\left[1 - \Phi\left(\frac{y}{\sqrt{t}}\right)\right] \\ &= 2\Phi\left(\frac{y}{\sqrt{t}}\right) - 1. \end{aligned} \tag{3.6}$$

Note que, se $x \leq y$,

$$\begin{aligned} P(X_t \leq x, \max_{0\leq s\leq t} X_s \leq y) &= P(X_t \leq x) - P(X_t \leq x, \max_{0\leq s\leq t} X_s \geq y) \\ &=^2 P(X_t \leq x) - P(X_t \geq 2y - x) \\ &= \Phi\left(\frac{x}{\sqrt{t}}\right) - \Phi\left(\frac{x-2y}{\sqrt{t}}\right). \end{aligned}$$

Se $x > y$,

$$P(X_t \leq x, \max_{0\leq s\leq t} X_s \leq y) = P(\max_{0\leq s\leq t} X_s \leq y) = 2\Phi\left(\frac{y}{\sqrt{t}}\right) - 1.$$

Logo, a distribuição conjunta H de X_t e $\max\limits_{0\leq s\leq t} X_s$ é dada por,

$$H_t(x,y) = \begin{cases} \Phi\left(\frac{x}{\sqrt{t}}\right) - \Phi\left(\frac{x-2y}{\sqrt{t}}\right), & \text{se} \quad x \leq y, \\ 2\Phi\left(\frac{y}{\sqrt{t}}\right) - 1, & \text{se} \quad x > y. \end{cases} \tag{3.7}$$

Sendo F_t a função de distribuição de X_t, as quase-inversas de F_t e G_t são,

$$F_t^{(-1)}(u) = \sqrt{t}\,\Phi^{-1}(u), \quad \text{e} \quad G_t^{(-1)}(v) = \sqrt{t}\,\Phi^{-1}\left(\frac{v+1}{2}\right).$$

Pelo corolário do Teorema de Sklar, se $u \leq \frac{v+1}{2}$ (expressão obtida ao se escrever $x = F_t^{(-1)}(u)$ e $y = G_t^{(-1)}(v)$ e determinar quando se tem $x \leq y$, ou seja, $\sqrt{t}\,\Phi^{-1}(u) \leq \sqrt{t}\,\Phi^{-1}\left(\frac{v+1}{2}\right)$),

$$\begin{aligned} C(u,v) &= H\left(\sqrt{t}\,\Phi^{-1}(u), \sqrt{t}\,\Phi^{-1}\left(\frac{v+1}{2}\right)\right) \\ &= \Phi\left(\frac{\sqrt{t}\,\Phi^{-1}(u)}{\sqrt{t}}\right) - \Phi\left(\frac{\sqrt{t}\,\Phi^{-1}(u) - 2\sqrt{t}\,\Phi^{-1}\left(\frac{v+1}{2}\right)}{\sqrt{t}}\right) \\ &= u - \Phi\left[\Phi^{-1}(u) - 2\Phi^{-1}\left(\frac{v+1}{2}\right)\right]. \end{aligned}$$

[2] Pelo princípio da reflexão.

Se $u > \frac{v+1}{2}$,

$$\begin{aligned} C(u,v) &= H\left(\sqrt{t}\,\Phi^{-1}(u), \sqrt{t}\,\Phi^{-1}\left(\frac{v+1}{2}\right)\right) \\ &= 2\Phi\left(\frac{\sqrt{t}\,\Phi^{-1}(\frac{v+1}{2})}{\sqrt{t}}\right) - 1 \\ &= 2\left(\frac{v+1}{2}\right) - 1 = v. \end{aligned}$$

Logo, a cópula associada ao Movimento Browniano e seu máximo é dada por

$$C(u,v) = \begin{cases} u - \Phi\left[\Phi^{-1}(u) - 2\,\Phi^{-1}\left(\dfrac{v+1}{2}\right)\right], & \text{se} \quad u \leq \dfrac{v+1}{2}, \\ v, & \text{se} \quad u > \dfrac{v+1}{2}\cdot \end{cases} \tag{3.8}$$

Como o Movimento Browniano possui incrementos independentes, é simétrico, Gaussiano e homogêneo, a cópula $C(\cdot\,,\cdot)$ associada à X_t e $\max\limits_{0\leq s\leq t} X_s$ é independente de t, já que t, nessas condições, se comporta como um parâmetro de escala e cópulas são invariantes por mudanças de escala.

3.4 Cópula Associada ao Máximo de um Movimento Browniano e ao Tempo de Primeira Chegada

Nesta seção desejamos determinar a cópula associada ao máximo de um Movimento Browniano e seu tempo de primeira chegada. Seja $\{X_t\}_{t\geq 0}$ um Movimento Browniano Padrão. Defina,

$$\mathcal{T}_x(\omega) = \inf_{t>0}\{X_t(\omega) = x\}. \tag{3.9}$$

A variável aleatória $\mathcal{T}$ pode ser interpretada como o tempo de primeira chegada a x. É claro que,

$$[\mathcal{T}_x \leq t] = [\max_{0\leq s\leq t} X_s \geq x]. \tag{3.10}$$

Além disso, pela simetria do Movimento Browniano, segue que $\mathcal{T}_x$ e $\mathcal{T}_{-x}$ têm a mesma distribuição. De fato, a variável aleatória $\mathcal{T}_x$, definida em (3.9), é um tempo de parada, como mostra o próximo lema.

Lema 3.4.1. *A variável aleatória $\mathcal{T}_x$, definida em (3.9), é um tempo de parada.*

Prova: Suponha que para $x \in S$, $t > 0$ e $\omega_1 \in \Omega$ tenhamos $\mathcal{T}_x(\omega_1) < t$. Isto quer dizer que existe $t_1 < t$ tal que $X_{t_1}(\omega_1) = x$. Suponha que para $\omega_2 \in \Omega$ tenhamos $X_s(\omega_2) = X_s(\omega_1)$, $0 \leq s < t$, então, $X_{t_1}(\omega_2) = X_{t_1}(\omega_1)$. Isto mostra que $\mathcal{T}_x(\omega_1) \geq \mathcal{T}_x(\omega_2)$. Por simetria, segue que $\mathcal{T}_x(\omega_1) = \mathcal{T}_x(\omega_2)$. □

O lema a seguir será utilizado nas provas das próximas proposições.

Lema 3.4.2. *Seja X uma variável aleatória com função de distribuição F. Seja $B \subseteq \mathbb{R}$ um boreliano. Então,*

$$P(X \in B) = \int_B dF,$$

onde a integral é de Lebesgue-Stieltjes.

Prova: Seja μ_F a medida de Lebesgue-Stieltjes sobre $\mathbb{R}$ determinada por F. Considere o conjunto B da forma $B = \bigcup_{i=1}^{n}(a_i, b_i]$, onde $b_i \leq a_{i+1}$ e $n < \infty$. Então,

$$\begin{aligned} \mu_F(B) &= \sum_{i=1}^{n}[F(b_i) - F(a_i)] \\ &= \sum_{i=1}^{n} P(a_i < X \leq b_i) \\ &= P(\cup_{i=1}^{n}[a_i < X \leq b_i]) := PX^{-1}(B). \end{aligned} \tag{3.11}$$

A igualdade (3.11) mostra que PX^{-1} é uma medida sobre o conjunto

$$A = \{\mathcal{B} \subseteq \mathbb{R} | \mathcal{B} = \cup_{i=1}^{n}(a_i, b_i],\ b_i \leq a_{i+1}, i = 1, \cdots, n\},$$

o conjunto de todos os conjuntos da forma de B. Como A gera a σ-álgebra de Borel então, pelo teorema de extensão de Kolmogorov, μ_F e PX^{-1} podem ser unicamente estendidos até (e assim são iguais sobre) a σ-álgebra de Borel em $\mathbb{R}$. Isto é,

$$P(X \in B) = PX^{-1}(B) = \int_B d\mu_F = \int_B dF,$$

qualquer que seja B boreliano em $\mathbb{R}$. □

A próxima proposição nos dá a distribuição de $\mathcal{T}_x$ e pode ser encontrada em Todorovic (1992), página 65.

Proposição 3.4.1. *Para todo $x \in \mathbb{R}$ e $t \geq 0$,*

$$P(\mathcal{T}_x \leq t) = 2P(X_t \geq x) = 2 - 2\Phi\left(\frac{x}{\sqrt{t}}\right). \tag{3.12}$$

Prova: Suponha inicialmente que $x > 0$. De (3.9) podemos deduzir que, se $t \leq 0$,

$$[X_t \geq x] \Longrightarrow [\mathcal{T}_x \leq t] \Longrightarrow [X_t \geq x] \subseteq [\mathcal{T}_x \leq t]. \tag{3.13}$$

Pela simetria do Movimento Browniano e de (3.10), para todo $0 \leq s \leq t$ e $x \in \mathbb{R}$,

$$P(X_t \leq x | X_s = x) = P(X_t \geq x | X_s = x),$$

segue que,

$$P(X_t \leq x | \mathcal{T}_x \leq t) = P(X_t \geq x | \mathcal{T}_x \leq t), \tag{3.14}$$

pois no instante $\mathcal{T}_x$ o processo estava em x. Levando em conta (3.13), obtemos,

$$P(X_t \leq x | \mathcal{T}_x \leq t) = \frac{P(X_t \geq x)}{P(\mathcal{T}_x \leq t)}. \tag{3.15}$$

De (3.14), (3.15) e da simetria do Movimento Browniano resulta,

$$\begin{aligned} P(X_t \geq x) &= P(X_t \leq x | \mathcal{T}_x \leq t) P(\mathcal{T}_x \leq t) \\ &= P(X_t \geq x, \mathcal{T}_x \leq t) \\ &= P(\mathcal{T}_x \leq t) - P(X_t \leq x, \mathcal{T}_x \leq t) \\ &= P(\mathcal{T}_x \leq t) - P(X_t \geq x, \mathcal{T}_x \leq t) \\ &= P(\mathcal{T}_x \leq t) - P(X_t \geq x). \end{aligned}$$

Logo,

$$P(\mathcal{T}_x \leq t) = 2P(X_t \geq x) = 2 - 2\Phi\left(\frac{x}{\sqrt{t}}\right).$$

Se $x < 0$ a proposição segue de que $\mathcal{T}_x$ e $\mathcal{T}_{-x}$ têm a mesma distribuição. $\square$

A próxima proposição (veja Todorovic, 1992) nos fornece a distribuição conjunta das variáveis aleatórias $\mathcal{T}_a$ e $\max_s X_s$ para o caso $a \leq x$.

Proposição 3.4.2. *Para todo $t > 0$ e $a \leq x$,*

$$P(\max_{0 \leq s \leq t} X_s \leq x, \mathcal{T}_a \leq y) = \frac{a}{\pi} \int_0^y \int_0^{x-a} \frac{1}{\sqrt{z^3(t-z)}} e^{-\frac{1}{2}\left[\frac{u^2}{t-z} + \frac{a^2}{z}\right]} \, du \, dz. \tag{3.16}$$

Prova: Seja $t > 0$ e $a > 0$. Pelo Lema 3.4.1, $\mathcal{T}_a$ é um tempo de parada. Como o processo $\{X_t\}_{t \geq 0}$ satisfaz a propriedade forte de Markov então, dado que $\mathcal{T}_a = z$, onde $0 < z \leq t$, para todo $0 \leq a \leq x$ temos,

$$\begin{aligned} P(\max_{0 \leq s \leq t} X_s \leq x | \mathcal{T}_a = z) &= P(\max_{z \leq s \leq t} X_s \leq x | \mathcal{T}_a = z) \\ &=^3 P(a + \max_{0 \leq s \leq t-z} X_s \leq x) \\ &= P(\max_{0 \leq s \leq t-z} X_s \leq x - a) \\ &= 2\Phi\left(\frac{x-a}{\sqrt{t-z}}\right) - 1. \end{aligned}$$

[3] Pela homogeneidade do Movimento Browniano.

Observe que, como $a \leq x$ e $z < t$,

$$
\begin{aligned}
2\Phi\left(\frac{x-a}{\sqrt{t-z}}\right) - 1 &= \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sqrt{t-z}}} e^{-\frac{v^2}{2}}\, dv - 1 \\
&= \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sqrt{t-z}}} e^{-\frac{v^2}{2}}\, dv - \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{v^2}{2}}\, dv \\
&= \sqrt{\frac{2}{\pi}} \int_{0}^{\frac{x-a}{\sqrt{t-z}}} e^{\frac{-v^2}{2}}\, dv \\
&=^{4} \sqrt{\frac{2}{\pi(t-z)}} \int_{0}^{x-a} e^{-\frac{u^2}{2(t-z)}}\, du.
\end{aligned}
$$

Logo,

$$
P(\max_{0\leq s\leq t} X_s \leq x | \mathcal{T}_a = z) = \sqrt{\frac{2}{\pi(t-z)}} \int_{0}^{x-a} e^{-\frac{u^2}{2(t-z)}}\, du. \tag{3.17}
$$

Condicionando no tempo de parada $\mathcal{T}_a$, temos,

$$
P(\max_{0\leq s\leq t} X_s \leq x, \mathcal{T}_a \leq y) = \int_0^y P(\max_{0\leq s\leq t} X_s \leq x | \mathcal{T}_a = z)\, dP(\mathcal{T}_a \leq z). \tag{3.18}
$$

Pela expressão (3.12) temos,

$$
\begin{aligned}
dP(\mathcal{T}_a \leq z) &= \frac{\partial}{\partial z}\left[2 - 2\Phi\left(\frac{a}{\sqrt{z}}\right)\right] dz \\
&= -2\,\frac{\partial}{\partial z}\left(\frac{1}{\sqrt{2\pi z}} \int_{-\infty}^{a} e^{-\frac{u^2}{2z}}\, du\right) dz \\
&= -\sqrt{\frac{2}{\pi}} \left[\int_{-\infty}^{a} \frac{\partial}{\partial z}\left(\frac{e^{-\frac{u^2}{2z}}}{\sqrt{z}}\right) du\right] dz \\
&= -\sqrt{\frac{2}{\pi}} \left[\int_{-\infty}^{a} \frac{1}{z}\left(\frac{u^2 e^{-\frac{u^2}{2z}}}{2z^{\frac{3}{2}}} - \frac{e^{-\frac{u^2}{2z}}}{2z^{\frac{1}{2}}}\right) du\right] dz.
\end{aligned}
$$

Assim,

$$
dP(\mathcal{T}_a \leq z) = -\sqrt{\frac{2}{\pi}} \left[\int_{-\infty}^{a} \frac{u^2 e^{-\frac{u^2}{2z}}}{2z^{\frac{5}{2}}}\, du - \int_{-\infty}^{a} \frac{e^{-\frac{u^2}{2z}}}{2z^{\frac{3}{2}}}\, du\right] dz. \tag{3.19}
$$

Integração por partes com $f = u$, $dg = ue^{-\frac{u^2}{2z}}\, du$ e $g = -ze^{-\frac{u^2}{2z}}$, resulta,

$$
\begin{aligned}
\int_{-\infty}^{a} \frac{u^2 e^{-\frac{u^2}{2z}}}{2z^{\frac{5}{2}}}\, du &= \frac{1}{2z^{\frac{5}{2}}}\left[-uze^{-\frac{u^2}{2z}}\Big|_{-\infty}^{a} + z\int_{-\infty}^{a} e^{-\frac{u^2}{2z}} du\right] \\
&= -\frac{ae^{-\frac{u^2}{2z}}}{2z^{\frac{3}{2}}} + \int_{-\infty}^{a} \frac{e^{-\frac{u^2}{2z}}}{2\,z^{\frac{3}{2}}}\, du\,.
\end{aligned} \tag{3.20}
$$

[4] Fazendo a mudança de variáveis $u = v\sqrt{t-z}$.

Substituindo (3.20) em (3.19), obtemos,

$$\begin{aligned} dP(\mathcal{T}_a \leq z) &= -\sqrt{\frac{2}{\pi}}\left[-\frac{ae^{-\frac{u^2}{2z}}}{2z^{\frac{3}{2}}}\right]dz \\ &= \frac{a}{\sqrt{2\pi}}\left[\frac{e^{-\frac{u^2}{2z}}}{z^{\frac{3}{2}}}\right]dz. \end{aligned} \tag{3.21}$$

Substituindo (3.21) e (3.17) em (3.18), obtemos,

$$\begin{aligned} P(\max_{0\leq s\leq t} X_s \leq x, \mathcal{T}_a \leq z) &= \int_0^y \left[\sqrt{\frac{2}{\pi(t-z)}}\int_0^{x-a} e^{-\frac{u^2}{2(t-z)}}\,du\right]\frac{a}{\sqrt{2\pi}}\left[\frac{e^{-\frac{u^2}{2z}}}{z^{\frac{3}{2}}}\right]dz \\ &= \frac{a}{\pi}\int_0^y\int_0^{x-a}\frac{1}{\sqrt{z^3(t-z)}}\,e^{-\frac{1}{2}\left(\frac{u^2}{t-z}+\frac{a^2}{z}\right)}\,du\,dz. \end{aligned}$$

Isto completa a prova. □

A proposição a seguir estende o resultado da Proposição 3.4.2 e completa a descrição da distribuição conjunta de $\mathcal{T}_a$ e $\max\limits_{0\leq s\leq t} X_s$.

Proposição 3.4.3. *Para todo $t > 0$ e $a \geq x$,*

$$\begin{aligned} P(\max_{0\leq s\leq t} X_s \leq x, \mathcal{T}_a \leq y) &= P(\mathcal{T}_a \leq y) - P(\max_{0\leq s\leq t} X_s \geq a) \\ &+ P(\max_{0\leq s\leq t} X_s < x)\int_0^x P(\max_{0<s\leq y-t} X_s \geq a-z)\,dP(X_t \leq z) \\ &- P(\max_{0\leq s\leq t} X_s < a)\int_0^a P(\max_{0<s\leq y-t} X_s \geq a-z)\,dP(X_t \leq z). \end{aligned} \tag{3.22}$$

Prova: Suponha que $0 \leq x < a < \infty$ e $t < y < \infty$.

$$P(\max_{0\leq s\leq t} X_s \leq x, \mathcal{T}_a \leq y) = P(\mathcal{T}_a \leq y) - P(\max_{0\leq s\leq t} X_s > x, \mathcal{T}_a \leq y).$$

Mas,

$$\left[\max_{0\leq s\leq t} X_s > x, \mathcal{T}_a \leq y\right] = \left[\max_{0\leq s\leq t} X_s \geq a\right] \cup \left[\max_{0\leq s\leq t} X_s > x, t \leq \mathcal{T}_a < y\right].$$

Assim,

$$\begin{aligned} P(\max_{0\leq s\leq t} X_s \leq x, \mathcal{T}_a \leq y) &= P(\mathcal{T}_a \leq y) - P(\max_{0\leq s\leq t} X_s \geq a) \\ &\quad - P(\max_{0\leq s\leq t} X_s > x, t < \mathcal{T}_a \leq y). \end{aligned}$$

Mas,

$$\left[\max_{0\leq s\leq t} X_s > x, t < \mathcal{T}_a \leq y\right] = \left[x < \max_{0\leq s\leq t} X_s < a, \max_{t<s\leq y} X_s \geq a\right].$$

Assim,

$$
\begin{aligned}
P(\max_{0\le s\le t} X_s > x, t < \mathcal{T}_a \le y) &= P(x < \max_{0\le s\le t} X_s < a, \max_{t<s\le y} X_s \ge a) \\
&= P(\max_{0\le s\le t} X_s < a, \max_{t<s\le y} X_s \ge a) - P(\max_{0\le s\le t} X_s < x, \max_{t<s\le y} X_s \ge a).
\end{aligned}
$$

Logo,

$$
\begin{aligned}
P(\max_{0\le s\le t} X_s \le x, \mathcal{T}_a \le y) &= P(\mathcal{T}_a \le y) - P(\max_{0\le s\le t} X_s \ge a) \\
&\quad + P(\max_{0\le s\le t} X_s < x, \max_{t<s\le y} X_s \ge a) \\
&\quad - P(\max_{0\le s\le t} X_s < a, \max_{t<s\le y} X_s \ge a). \qquad (3.23)
\end{aligned}
$$

Primeiramente note que, como $X_0 = 0$ e como queremos que $\max_{0\le s\le t} X_s < x$, então,

$$
\begin{aligned}
&P(\max_{0\le s\le t} X_s < x, \max_{t<s\le y} X_s \ge a) = \\
&= P(\max_{t<s\le y} X_s \ge a | \max_{0\le s\le t} X_s < x) P(\max_{0\le s\le t} X_s < x) \\
&= P(\max_{0\le s\le t} X_s < x) \int_0^x P(\max_{t<s\le y} X_s \ge a | \max_{0\le s\le t} X_s < x, X_t = z) dP(X_t \le z).
\end{aligned}
$$

Mas, pela propriedade de Markov,

$$
P(\max_{t<s\le y} X_s \ge a | \max_{0\le s\le t} X_s < x, X_t = z) = P(\max_{t<s\le y} X_s \ge a | X_t = z).
$$

Assim,

$$
\begin{aligned}
&P(\max_{0\le s\le t} X_s < x, \max_{t<s\le y} X_s \ge a) = \\
&\quad = P(\max_{0\le s\le t} X_s < x) \int_0^x P(\max_{t<s\le y} X_s \ge a | X_t = z)\, dP(X_t \le z) \\
&\quad =^5 P(\max_{0\le s\le t} X_s < x) \int_0^x P(\max_{0<s\le y-t} X_s \ge a - z | X_0 = 0)\, dP(X_t \le z) \\
&\quad = P(\max_{0\le s\le t} X_s < x) \int_0^x P(\max_{0<s\le y-t} X_s \ge a - z)\, dP(X_t \le z). \qquad (3.24)
\end{aligned}
$$

Por outro lado,

$$
\begin{aligned}
&P(\max_{0\le s\le t} X_s < a, \max_{t<s\le y} X_s \ge a) = \\
&= P(\max_{t<s\le y} X_s \ge a | \max_{0\le s\le t} X_s < a) P(\max_{0\le s\le t} X_s < a) \\
&= P(\max_{0\le s\le t} X_s < a) \int_0^a P(\max_{t<s\le y} X_s \ge a | \max_{0\le s\le t} X_s < a, X_t = z)\, dP(X_t \le z).
\end{aligned}
$$

[5] Pela homogeneidade do Movimento Browniano.

Novamente, pela propriedade de Markov e pela homogeneidade do Movimento Browniano,

$$\begin{aligned} P(\max_{t<s\leq y} X_s \geq a | \max_{0\leq s\leq t} X_s < a, X_t = z) &= P(\max_{t<s\leq y} X_s \geq a | X_t = z) \\ &= P(\max_{0<s\leq y-t} X_s \geq a - z | X_0 = 0) \\ &= P(\max_{0<s\leq y-t} X_s \geq a - z). \end{aligned}$$

Logo,

$$\begin{aligned} &P(\max_{0\leq s\leq t} X_s < a, \max_{t<s\leq y} X_s \geq a) = \\ &\quad = P(\max_{0\leq s\leq t} X_s < a) \int_0^a P(\max_{0<s\leq y-t} X_s \geq a - z) \, dP(X_t \leq z). \end{aligned} \tag{3.25}$$

Substituindo (3.24) e (3.25) em (3.23), obtemos,

$$\begin{aligned} P(\max_{0\leq s\leq t} X_s \leq x, \mathcal{T}_a \leq y) &= P(\mathcal{T}_a \leq y) - P(\max_{0\leq s\leq t} X_s \geq a) \\ &+ P(\max_{0\leq s\leq t} X_s < x) \int_0^x P(\max_{0<s\leq y-t} X_s \geq a - z) \, dP(X_t \leq z) \\ &- P(\max_{0\leq s\leq t} X_s < a) \int_0^a P(\max_{0<s\leq y-t} X_s \geq a - z) \, dP(X_t \leq z). \end{aligned}$$

Isto completa a prova. □

Corolário 3.4.1. *Para todo* $0 < t < y$ *e* $a \geq x$,

$$P(\max_{0\leq s\leq t} X_s \leq x, \mathcal{T}_a \leq y) = \Lambda_{t,y}(x) - \Lambda_{t,y}(a) + \sqrt{\frac{2}{\pi}} \int_{\frac{a}{\sqrt{y}}}^{\frac{a}{\sqrt{t}}} e^{-\frac{z^2}{2}} \, dz, \tag{3.26}$$

onde,

$$\Lambda_{t,y}(k) = \frac{1}{\pi^{\frac{3}{2}}} \sqrt{\frac{2}{y-t}} \int_0^{\frac{k}{\sqrt{t}}} \int_0^k \int_{-\infty}^{z-a} e^{-\frac{1}{2}\left[\frac{v^2}{y-t} + u^2 + z^2\right]} dv \, dz \, du.$$

Prova: Note primeiramente que,

$$\begin{aligned} 2\Phi\left(\frac{x}{\sqrt{t}}\right) - 1 &= \frac{2}{\sqrt{2\pi t}} \int_{-\infty}^x e^{\frac{u^2}{2t}} du - \frac{2}{\sqrt{2\pi t}} \int_{-\infty}^0 e^{-\frac{u^2}{2t}} du \\ &= \sqrt{\frac{2}{\pi t}} \int_0^x e^{-\frac{u^2}{2t}} du. \end{aligned} \tag{3.27}$$

Além disso, temos que,

$$2\Phi\left(\frac{a}{\sqrt{t}}\right) - 2\Phi\left(\frac{a}{\sqrt{y}}\right) = \sqrt{\frac{2}{\pi}} \int_{\frac{a}{\sqrt{y}}}^{\frac{a}{\sqrt{t}}} e^{-\frac{u^2}{2}} du, \tag{3.28}$$

e

$$1-\Phi\left(\frac{a-z}{\sqrt{y-t}}\right)=\Phi\left(\frac{z-a}{\sqrt{y-t}}\right)=\frac{1}{\sqrt{2\pi(y-t)}}\int_{-\infty}^{z-a}e^{-\frac{u^2}{2(y-t)}}du. \quad (3.29)$$

Como $\{X_t\}_{t\geq 0}$ é um Movimento Browniano Padrão, segue que,

$$dP(X_t\leq z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz. \quad (3.30)$$

Utilizando (3.6), (3.27), (3.28), (3.29) e (3.30) obtemos,

$$\begin{aligned}
P(\max_{0\leq s\leq t}X_s<x)\int_0^x P(\max_{0\leq s\leq y-t}X_s\geq a-z)\,dP(X_t\leq z) &= \\
=\left(2\Phi\left(\frac{x}{\sqrt{t}}\right)-1\right)\int_0^x\left[2-2\Phi\left(\frac{a-z}{\sqrt{y-t}}\right)\right]dP(X_t\leq z) & \\
=\left(\sqrt{\frac{2}{\pi}}\int_0^{\frac{x}{\sqrt{t}}}e^{-\frac{u^2}{2}}\,du\right)\int_0^x\left[\frac{2}{\sqrt{2\pi(y-t)}}\int_{-\infty}^{z-a}e^{-\frac{v^2}{2(y-t)}}\,dv\right]\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\,dz & \\
=\frac{1}{\pi^{\frac{3}{2}}}\sqrt{\frac{2}{y-t}}\int_0^{\frac{x}{\sqrt{t}}}e^{-\frac{u^2}{2}}\,du\int_0^x\int_{-\infty}^{z-a}e^{-\frac{1}{2}\left[\frac{v^2}{y-t}+z^2\right]}\,dv\,dz & \\
=\frac{1}{\pi^{\frac{3}{2}}}\sqrt{\frac{2}{y-t}}\int_0^{\frac{x}{\sqrt{t}}}\int_0^x\int_{-\infty}^{z-a}e^{-\frac{1}{2}\left[\frac{v^2}{y-t}+u^2+z^2\right]}\,dv\,dz\,du. &
\end{aligned}$$

Definindo,

$$\Lambda_{t,y}(k)=\frac{1}{\pi^{\frac{3}{2}}}\sqrt{\frac{2}{y-t}}\int_0^{\frac{k}{\sqrt{t}}}\int_0^k\int_{-\infty}^{z-a}e^{-\frac{1}{2}\left[\frac{v^2}{y-t}+u^2+z^2\right]}\,dv\,dz\,du,$$

obtemos,

$$P(\max_{0\leq s\leq t}X_s<x)\int_0^x P(\max_{0\leq s\leq y-t}X_s\geq a-z)\,dP(X_t\leq z)=\Lambda_{t,y}(x). \quad (3.31)$$

Analogamente,

$$P(\max_{0\leq s\leq t}X_s<a)\int_0^a P(\max_{0\leq s\leq y-t}X_s\geq a-z)\,dP(X_t\leq z)=\Lambda_{t,y}(a). \quad (3.32)$$

Com (3.31) e (3.32) podemos reescrever (3.22) como

$$P(\max_{0\leq s\leq t}X_s\leq x,\mathcal{T}_a\leq y)=P(\mathcal{T}_a\leq y)-P(\max_{0\leq s\leq t}X_s\geq a)+\Lambda_{t,y}(x)-\Lambda_{t,y}(a).$$

Por (3.6) e (3.12), temos

$$\begin{aligned}
P(\max_{0\leq s\leq t}X_s\leq x,\mathcal{T}_a\leq y) &= 2-2\Phi\left(\frac{a}{\sqrt{t}}\right)-2+2\Phi\left(\frac{a}{\sqrt{y}}\right)+\Lambda_{t,y}(x)-\Lambda_{t,y}(a) \\
&= 2\Phi\left(\frac{a}{\sqrt{y}}\right)-2\Phi\left(\frac{a}{\sqrt{t}}\right)+\Lambda_{t,y}(x)-\Lambda_{t,y}(a). \quad (3.33)
\end{aligned}$$

Agora substituindo (3.28) em (3.33), (3.26) segue. $\square$

A expressão para $\Lambda_{t,y}(\cdot)$ pode ser escrita de uma forma um pouco mais compacta. Este é o conteúdo do próximo lema.

Lema 3.4.3. *Para todo a, t e $y \in \mathbb{R}^+$ fixos,*

$$\Lambda_{t,y}(k) = \left(2\Phi\left(\frac{k}{\sqrt{t}}\right) - 1\right)\sqrt{\frac{2}{\pi}}\int_0^k e^{-\frac{z^2}{2}}\,\Phi\left(\frac{z-a}{\sqrt{y-t}}\right)dz, \tag{3.34}$$

onde $\Phi(\cdot)$ é a função de distribuição de uma variável aleatória $Z \sim \mathcal{N}(0,1)$.

Prova: No Corolário 3.4.1, definimos

$$\Lambda_{t,y}(k) = \frac{1}{\pi^{\frac{3}{2}}}\sqrt{\frac{2}{y-t}}\int_0^{\frac{k}{\sqrt{t}}}\int_0^k\int_{-\infty}^{z-a} e^{-\frac{1}{2}\left[\frac{v^2}{y-t}+u^2+z^2\right]}dv\,dz\,du. \tag{3.35}$$

Note que,

$$\int_{-\infty}^{z-a} e^{-\frac{1}{2}\left[\frac{v^2}{y-t}+u^2+z^2\right]}dv = e^{-\frac{1}{2}\left[u^2+z^2\right]}\int_{-\infty}^{z-a} e^{-\frac{v^2}{2(y-t)}}dv. \tag{3.36}$$

Lembrando que $\Phi(\cdot)$ é a função de distribuição de uma variável aleatória $Z \sim \mathcal{N}(0,1)$, podemos escrever (3.36) como

$$\int_{-\infty}^{z-a} e^{-\frac{1}{2}\left[\frac{v^2}{y-t}+u^2+z^2\right]}dv = \sqrt{2\pi(y-t)}\,\Phi\left(\frac{z-a}{\sqrt{y-t}}\right)e^{-\frac{1}{2}\left[\frac{u^2}{t}+z^2\right]}. \tag{3.37}$$

Substituindo (3.37) em (3.35) obtemos

$$\begin{aligned}
\Lambda_{t,y}(k) &= \frac{1}{\pi^{\frac{3}{2}}}\sqrt{\frac{2}{y-t}}\sqrt{2\pi(y-t)}\int_0^{\frac{k}{\sqrt{t}}}\int_0^k e^{-\frac{1}{2}\left[u^2+z^2\right]}\Phi\left(\frac{z-a}{\sqrt{y-t}}\right)dz\,du \\
&= \frac{2}{\pi}\int_0^{\frac{k}{\sqrt{t}}} e^{-\frac{u^2}{2}}\left[\int_0^k e^{-\frac{z^2}{2}}\,\Phi\left(\frac{z-a}{\sqrt{y-t}}\right)dz\right]du \\
&= \frac{2}{\pi}\int_0^k e^{-\frac{z^2}{2}}\Phi\left(\frac{z-a}{\sqrt{y-t}}\right)dz\int_0^{\frac{k}{\sqrt{t}}} e^{-\frac{u^2}{2}}\,du. \qquad (3.38)
\end{aligned}$$

Mas,

$$\begin{aligned}
\int_0^{\frac{k}{\sqrt{t}}} e^{-\frac{u^2}{2}}du &= \sqrt{2\pi}\int_0^{\frac{k}{\sqrt{t}}}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{u^2}{2}}\,du \\
&= \sqrt{2\pi}\left[\int_{-\infty}^{\frac{k}{\sqrt{t}}}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{u^2}{2}}\,du - \frac{1}{2}\right] \\
&= \sqrt{2\pi}\left[\Phi\left(\frac{k}{\sqrt{t}}\right) - \frac{1}{2}\right]. \qquad (3.39)
\end{aligned}$$

Substituindo (3.39) em (3.38), obtemos

$$\begin{aligned}\Lambda_{t,y}(k) &= \frac{2}{\pi}\sqrt{2\pi}\left[\Phi\left(\frac{k}{\sqrt{t}}\right)-\frac{1}{2}\right]\int_0^k e^{-\frac{z^2}{2}}\,\Phi\Big(\frac{z-a}{\sqrt{y-t}}\Big)\,dz\\ &= \left(2\Phi\left(\frac{k}{\sqrt{t}}\right)-1\right)\sqrt{\frac{2}{\pi}}\int_0^k e^{-\frac{z^2}{2}}\,\Phi\Big(\frac{z-a}{\sqrt{y-t}}\Big)\,dz.\end{aligned}$$

Isto completa a prova. $\square$

Combinando o resultado do Lema 3.4.3 com o Corolário 3.4.1, obtemos imediatamente o próximo resultado.

Corolário 3.4.2. *Para todo $t > 0$ e $a \geq x$,*

$$P(\max_{0\leq s\leq t} X_s \leq x, \mathcal{T}_a \leq y) = \Lambda_{t,y}(x) - \Lambda_{t,y}(a) + \sqrt{\frac{2}{\pi}}\int_{\frac{a}{\sqrt{y}}}^{\frac{a}{\sqrt{t}}} e^{-\frac{z^2}{2}}\,dz, \qquad (3.40)$$

onde

$$\Lambda_{t,y}(k) = \left(2\Phi\left(\frac{k}{\sqrt{t}}\right)-1\right)\sqrt{\frac{2}{\pi}}\int_0^k e^{-\frac{z^2}{2}}\,\Phi\Big(\frac{z-a}{\sqrt{y-t}}\Big)\,dz.$$

Da expressão (3.35) notamos que a distribuição conjunta de $\max\limits_{0\leq s\leq t} X_s$ e $\mathcal{T}_a$ é uma função a quatro argumentos, x, a, y e t. A distribuição conjunta de $\max\limits_{0\leq s\leq t} X_s$ e $\mathcal{T}_a$ é tomada para cada a e $t \in \mathbb{R}^+$ fixos, isto é,

$$\begin{aligned} H_{t,a} \ &: \ \mathbb{R}^+\times\mathbb{R}^+ \longrightarrow I\\ &\quad (x,y) \longmapsto H_{t,a}(x,y),\end{aligned}$$

com $H_{t,a}(x,y) = P(\max\limits_{0\leq s\leq t} X_s \leq x, \mathcal{T}_a \leq y)$.

Pela Proposição 3.4.2 e o Corolário 3.4.2, a distribuição conjunta $H_{t,a}(\cdot\,,\cdot)$ é dada por

$$H_{t,a}(x,y) = \begin{cases} \Lambda_{t,y}(x) - \Lambda_{t,y}(a) + \sqrt{\dfrac{2}{\pi}}\displaystyle\int_{\frac{a}{\sqrt{y}}}^{\frac{a}{\sqrt{t}}} e^{-\frac{z^2}{2}}\,dz, & \text{se} \quad \begin{array}{l} a > x,\\ y > 0;\end{array}\\[2ex] \dfrac{a}{\pi}\displaystyle\int_0^y\int_0^{x-a} \frac{1}{\sqrt{z^3(t-z)}}\, e^{-\frac{1}{2}\left[\frac{u^2}{t-z}+\frac{a^2}{z}\right]}\,du\,dz, & \text{se} \quad \begin{array}{l} a \leq x,\\ y > t;\end{array}\\[2ex] 0, & \text{se} \quad \begin{array}{l} a \leq x,\\ y \leq t,\end{array}\end{cases} \qquad (3.41)$$

onde

$$\Lambda_{t,y}(k) = \left(2\Phi\left(\frac{k}{\sqrt{t}}\right)-1\right)\sqrt{\frac{2}{\pi}}\int_0^k e^{-\frac{z^2}{2}}\,\Phi\Big(\frac{z-a}{\sqrt{y-t}}\Big)\,dz.$$

Fixados $a, t \in \mathbb{R}^+$, queremos determinar a cópula $C_{t,a}(u, v)$ associada ao vetor $(\max\limits_{0\leq s\leq t} X_s, \mathcal{T}_a)$. Denotando por F_t e G_a as marginais de $H_{t,a}(\cdot\,,\cdot)$, de (3.6) e (3.12) temos

$$F_t(x) = P(\max_{0\leq s\leq t} X_s \leq x) = 2\Phi\left(\frac{x}{\sqrt{t}}\right) - 1, \tag{3.42}$$

e

$$G_a(y) = P(\mathcal{T}_a \leq y) = 2 - 2\Phi\left(\frac{a}{\sqrt{y}}\right), \tag{3.43}$$

quaisquer que sejam $x, y \in \mathbb{R}^+$. As inversas de $F_t(\cdot)$ e $G_a(\cdot)$ são dadas por

$$F_t^{-1}(u) = \sqrt{t}\,\Phi^{-1}\left(\frac{u+1}{2}\right) \text{ e } G_a^{-1}(v) = \left(\frac{a}{\Phi^{-1}\left(\frac{2-v}{2}\right)}\right)^2, \quad u, v \in I.$$

Usando as transformações $x = F_t^{-1}(u)$ e $y = G_a^{-1}(v)$, com $u, v \in I$, pelo Corolário 1.3.1, $C_{t,a}(u, v)$ é dada por

$$C_{t,a}(u, v) = H_{t,a}(F_t^{-1}(u), G_a^{-1}(v)), \quad \forall\, u, v \in I,$$

que deve ser analisado em cada caso de (3.41). Isto é feito na próxima proposição.

Proposição 3.4.4. *Seja $\{X_t\}_{t\geq 0}$ um Movimento Browniano Padrão. Então, fixados a e $t \in \mathbb{R}^+$, a cópula associada às variáveis aleatórias $\max\limits_{0\leq s\leq t} X_s$ e $\mathcal{T}_a$ (conforme definida em (3.9)) é dada por*

$$C_{t,a}(u, v) = \begin{cases} \Lambda_{t,\xi_v}(\zeta_u) - \Lambda_{t,\xi_v}(a) + \sqrt{\dfrac{2}{\pi}} \displaystyle\int_{\Phi^{-1}\left(\frac{2-v}{2}\right)}^{\frac{a}{\sqrt{t}}} e^{-\frac{z^2}{2}}\, dz, & \mathit{se} \quad u < F_t(a); \\[2ex] \dfrac{a}{\pi} \displaystyle\int_0^{\xi_v}\int_0^{\zeta_u - a} \frac{1}{\sqrt{z^3(t-z)}}\, e^{-\frac{1}{2}\left[\frac{w^2}{t-z} + \frac{a^2}{z}\right]}\, dw\, dz, & \mathit{se} \quad \begin{array}{l} u \geq F_t(a), \\ v > G_a(t); \end{array} \\[2ex] 0, & \mathit{se} \quad \begin{array}{l} u \geq F_t(a), \\ v \leq G_a(t), \end{array} \end{cases}$$

onde $F_t(\cdot)$ e $G_a(\cdot)$ são dadas em (3.42) e (3.43), respectivamente,

$$\zeta_u = \sqrt{t}\;\Phi^{-1}\left(\frac{u+1}{2}\right) \quad e \quad \xi_v = \left(\frac{a}{\Phi^{-1}\left(\frac{2-v}{2}\right)}\right)^2,$$

e

$$\Lambda_{t,y}(k) = \left(2\Phi\left(\frac{k}{\sqrt{t}}\right) - 1\right)\sqrt{\frac{2}{\pi}} \int_0^k e^{-\frac{z^2}{2}}\, \Phi\left(\frac{z-a}{\sqrt{y-t}}\right) dz.$$

Prova: Pelo Corolário 1.3.1, a cópula $C_{t,a}(\cdot,\cdot)$ é dada por

$$C_{t,a}(u,v) = H_{t,a}(F_t^{-1}(u), G_a^{-1}(v)),$$

onde

$$H_{t,a}(x,y) = P(\max_{0\le s\le t} X_s \le x, \mathcal{T}_a \le y),$$

é dada por (3.41). Defina

$$\zeta_u := \sqrt{t}\,\Phi^{-1}\left(\frac{u+1}{2}\right) \quad \text{e} \quad \xi_v := \left(\frac{a}{\Phi^{-1}(\frac{2-v}{2})}\right)^2, \quad u,v \in I.$$

Note que ζ_u e ξ_v nada mais são do que as inversas das funções de distribuição $F_t(\cdot)$ de $\max\limits_{0\le s\le t} X_s$ e $G_a(\cdot)$ de $\mathcal{T}_a$ dadas em (3.42) e (3.43), respectivamente. Tendo em vista as transformações $x = F_t^{-1}(u)$ e $y = G_a^{-1}(v)$ e (3.41), a condição $a > x$ é equivalente a $u < F_t(a)$. Assim se $u < F_t(a)$, temos

$$\begin{aligned} C_{t,a}(u,v) &= \Lambda_{t,G_a^{-1}(v)}(F_t^{-1}(u)) - \Lambda_{t,G_a^{-1}(v)}(a) + \sqrt{\frac{2}{\pi}}\int_{\frac{a}{\sqrt{G_a^{-1}(v)}}}^{\frac{a}{\sqrt{t}}} e^{-\frac{z^2}{2}}\,dz \\ &= \Lambda_{t,\xi_v}(\zeta_u) - \Lambda_{t,\xi_v}(a) + \sqrt{\frac{2}{\pi}}\int_{\Phi^{-1}\left(\frac{2-v}{2}\right)}^{\frac{a}{\sqrt{t}}} e^{-\frac{z^2}{2}}\,dz, \end{aligned}$$

onde

$$\Lambda_{t,y}(k) = \left(2\Phi\left(\frac{k}{\sqrt{t}}\right) - 1\right)\sqrt{\frac{2}{\pi}}\int_0^k e^{-\frac{z^2}{2}}\,\Phi\left(\frac{z-a}{\sqrt{y-t}}\right)dz.$$

As condições $a \le x$ e $t < y$, mediante as transformações $x = F_t^{-1}(u)$ e $y = G_a^{-1}(v)$, são equivalentes a $u \ge F_t(a)$ e $v > G_a(t)$ e nessas condições temos

$$\begin{aligned} C_{t,a}(u,v) &= \frac{a}{\pi}\int_0^{G_a^{-1}(v)}\int_0^{F_t^{-1}(u)-a} \frac{1}{\sqrt{z^3(t-z)}}\, e^{-\frac{1}{2}\left[\frac{w^2}{t-z}+\frac{a^2}{z}\right]}\,dw\,dz \\ &= \frac{a}{\pi}\int_0^{\xi_v}\int_0^{\zeta_u-a} \frac{1}{\sqrt{z^3(t-z)}}\, e^{-\frac{1}{2}\left[\frac{w^2}{t-z}+\frac{a^2}{z}\right]}\,dw\,dz. \end{aligned}$$

As condições $a \le x$ e $t \ge y$, mediante as mesmas transformações $x = F_t^{-1}(u)$ e $y = G_a^{-1}(v)$, são equivalentes a $u \ge F_t(a)$ e $v \le G_a(t)$. Nessas condições temos

$$C_{t,a}(u,v) = 0,$$

e isto completa a prova. □

Ainda que a cópula associada ao vetor $(\max_{0\le s\le t} X_s, \mathcal{T}_a)$ dada na Proposição 3.4.4 pareça tão ou mais complicada que a distribuição conjunta dada em (3.41), a vantagem está no fato de a cópula conter, sozinha, todas as informações relevantes sobre a estrutura de dependência entre as variáveis aleatórias em questão. A importância do estudo do vetor $\max_{0\le s\le t} X_s$ está na potencialidade de aplicações, que não serão feitas aqui.

Capítulo 4

Famílias de Cópulas Modificadas

O objetivo principal deste capítulo é apresentar reparametrizações para algumas famílias conhecidas de cópulas de forma a obter formas mais simples para essas porém, com um parâmetro adicional. Estas reparametrizações visam a flexibilização das famílias de cópulas em questão ao mesmo tempo que, devido às formas mais simples assumidas pelas cópulas, o cálculo da estrutura de dependência induzida pelas mesmas seja facilitado. Inicialmente são apresentados alguns resultados necessários ao desenvolvimento do trabalho, como o Lema de Hoeffding e sua versão para cópulas. Este capítulo é baseado em Lopes e Pumi (2006).

4.1 Lema de Hoeffding

O foco desta seção será determinar uma fórmula para a covariância entre variáveis aleatórias baseada em cópulas. A covariância entre duas variáveis aleatórias X e Y é definida por

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y), \tag{4.1}$$

sempre que as esperanças envolvidas existirem. Uma fórmula alternativa envolvendo a distribuição conjunta de X e Y e suas marginais apareceu pela primeira vez em Hoeffding (1941) e ficou conhecido como Lema de Hoeffding. A prova apresentada é baseada em Lehmann (1966).

Lema 4.1.1. *(Lema de Hoeffding). Sejam X e Y duas variáveis aleatórias com distribuição conjunta H e marginais F e G, respectivamente, e suponha que $\mathbb{E}(X)$, $\mathbb{E}(Y)$ e $\mathbb{E}(XY)$ existam. Então,*

$$\mathrm{Cov}(X,Y) = \iint_D H(x,y)\,dxdy - \iint_D F(x)\,G(y)\,dxdy, \tag{4.2}$$

onde D denota o domínio de (X,Y).

Prova: Sejam (X_1, Y_1) e (X_2, Y_2) dois pares de variáveis aleatórias independentes e tais que $(X_1, Y_1) \stackrel{d}{=} (X, Y)$ e $(X_2, Y_2) \stackrel{d}{=} (X, Y)$. Seja D o domínio de (X, Y) Note que,

$$\begin{aligned}
\mathbb{E}\left[(X_1 - X_2)(Y_1 - Y_2)\right] &= \mathbb{E}\left[(X_1Y_1) - (X_1Y_2) - (X_2Y_1) + (X_2Y_2)\right] \\
&= \mathbb{E}(X_1Y_1) + \mathbb{E}(X_2Y_2) - \mathbb{E}(X_1)\,\mathbb{E}(Y_2) - \mathbb{E}(X_2)\,\mathbb{E}(Y_1) \\
&= 2\left[\mathbb{E}(X_1Y_1) - \mathbb{E}(X_1)\mathbb{E}(Y_1)\right] \\
&= 2\left[\mathbb{E}(X, Y) - \mathbb{E}(X)\mathbb{E}(Y)\right]. \qquad (4.3)
\end{aligned}$$

Mas,

$$\begin{aligned}
&\mathbb{E}\big[(X_1 - X_2)(Y_1 - Y_2)\big] = \\
&= \mathbb{E}\left[\iint_D \big[\,\mathrm{I}(X_1 \geq u) - \mathrm{I}(X_2 \geq u)\big]\big[\,\mathrm{I}(Y_1 \geq v) - \mathrm{I}(Y_2 \geq v)\big] dudv\right].
\end{aligned}$$

Como as esperanças em questão existem, temos

$$\begin{aligned}
&= \iint_D \mathbb{E}\left([\,\mathrm{I}(X_1 \geq u) - \mathrm{I}(X_2 \geq u)][\,\mathrm{I}(Y_1 \geq v) - \mathrm{I}(Y_2 \geq v)]\right) dudv \\
&= 2\iint_D H(x, y) - F(x)\,G(y)\,dxdy. \qquad (4.4)
\end{aligned}$$

Comparando (4.3) com (4.4), obtemos o resultado desejado. $\square$

Usando o Teorema de Sklar no Lema de Hoeffding é fácil converter (4.2) em outra expressão em termos da cópula e das marginais de X e Y.

Corolário 4.1.1. *(Versão do Lema de Hoeffding em Termos de Cópulas). Sejam X e Y duas variáveis aleatórias contínuas com cópula associada C. Então,*

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \iint_{I^2} \frac{C(u, v)}{F'\big(F^{(-1)}(u)\big)G'\big(G^{(-1)}(v)\big)}\, dudv \\
&\quad - \iint_{I^2} \frac{uv}{F'\big(F^{(-1)}(u)\big)G'\big(G^{(-1)}(v)\big)}\, dudv. \qquad (4.5)
\end{aligned}$$

Prova: Fazendo a mudança de variáveis $u = F(x)$ e $v = G(y)$, cujo Jacobiano é

$$\begin{vmatrix} \dfrac{\partial F(x)}{\partial x} & \dfrac{\partial F(x)}{\partial y} \\[2ex] \dfrac{\partial G(y)}{\partial x} & \dfrac{\partial G(y)}{\partial y} \end{vmatrix} = \begin{vmatrix} F'(x) & 0 \\ 0 & G'(y) \end{vmatrix} = F'(x)G'(y),$$

em (4.2), obtemos

$$
\begin{aligned}
\operatorname{Cov}(X,Y) &= \iint_{I^2} \frac{H\big(F^{-1}(u),G^{-1}(v)\big)}{F'\big(F^{(-1)}(u)\big)G'\big(G^{(-1)}(v)\big)}\,dudv \\
&\quad - \iint_{I^2} \frac{uv}{F'\big(F^{(-1)}(u)\big)G'\big(G^{(-1)}(v)\big)}\,dudv \\
&=^{1} \iint_{I^2} \frac{C(u,v)}{F'\big(F^{(-1)}(u)\big)G'\big(G^{(-1)}(v)\big)}\,dudv \\
&\quad - \iint_{I^2} \frac{uv}{F'\big(F^{(-1)}(u)\big)G'\big(G^{(-1)}(v)\big)}\,dudv,
\end{aligned}
$$

o que completa a prova. □

Observação 4.1.1. Note que se X e Y têm distribuição $\mathcal{U}([0,1])$, então $F'(x)=1$, para todo x e $G'(y)=1$, para todo o y. Assim, a equação (4.5) é equivalente a

$$\operatorname{Cov}(X,Y) = \iint_{I^2} C(u,v)\,dudv - \frac{1}{4}\cdot \tag{4.6}$$

4.2 Família Ali-Mikhail-Haq Modificada

Considere a família de cópulas Ali-Mikhail-Haq, dada por,

$$C_\theta(u,v) = \frac{uv}{1-\theta(1-u)(1-v)}\,, \tag{4.7}$$

para $\theta \in [-1,1)$. Esta família de cópulas é um membro da Família Arquimediana de cópulas e seu gerador é dado por $\varphi(t) = \ln\left(\frac{1-\theta(1-t)}{t}\right)$. Na Seção 2.1 apresentamos a construção desta família em detalhes. Note que esta família de cópulas possui, como caso particular, $C_0 = \Pi$.

Na próxima proposição, determinamos a estrutura de covariância associada a duas variáveis aleatórias com distribuição exponencial e cópula associada pertencente à Família Ali-Mikhail-Haq através da versão para cópulas do Lema de Hoeffding.

Proposição 4.2.1. *Sejam X e Y variáveis aleatórias com distribuição $\mathcal{E}(\lambda_1)$ e $\mathcal{E}(\lambda_2)$, respectivamente. Suponha que a cópula associada a X e Y seja um membro da Família Ali-Mikhail-Haq de cópulas, dada por (4.7), com parâmetro $\theta \in [-1,1)$. Então,*

$$\operatorname{Cov}(X,Y) = \frac{\operatorname{dilog}(1-\theta)}{\lambda_1\lambda_2}\,, \tag{4.8}$$

[1] Pelo Corolário do Teorema de Sklar.

onde dilog$(\cdot)$ *é a função dilogaritmo dada por* $\mathrm{dilog}(x) = \displaystyle\int_1^x \frac{\ln(t)}{1-t}\, dt.$

Prova: Seja X e Y variáveis aleatórias com distribuição $\mathcal{E}(\lambda_1)$ e $\mathcal{E}(\lambda_2)$, respectivamente. Desta forma, temos

$$F(x) = 1 - e^{-\lambda_1 x}, \qquad F'(x) = \lambda_1 e^{-\lambda_1 x}, \qquad F^{-1}(u) = -\frac{\ln(1-u)}{\lambda_1},$$

e

$$G(y) = 1 - e^{-\lambda_2 y}, \qquad G'(y) = \lambda_2 e^{-\lambda_2 y}, \qquad G^{-1}(v) = -\frac{\ln(1-v)}{\lambda_2}.$$

Assim,

$$\begin{aligned} F'\big(F^{-1}(u)\big) &= \lambda_1 e^{-\lambda_1\left(-\frac{\ln(1-u)}{\lambda_1}\right)} \\ &= \lambda_1(1-u). \end{aligned}$$

Analogamente, $G'\big(G^{-1}(v)\big) = \lambda_2(1-v)$. Assim, (4.5) se torna

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \iint_{I^2} \frac{1}{\lambda_1\lambda_2(1-u)(1-v)} \left[\frac{uv}{1-\theta(1-u)(1-v)} - uv\right] dudv \\ &= \iint_{I^2} \frac{\theta(1-u)(1-v)}{1-\theta(1-u)(1-v)} \cdot \frac{1}{\lambda_1\lambda_2(1-u)(1-v)}\, dudv \\ &= \frac{\theta}{\lambda_1\lambda_2} \int_0^1 \left[\int_0^1 \frac{1}{1-\theta(1-u)(1-v)}\, du\right] dv. \end{aligned} \tag{4.9}$$

Fazendo a mudança de variáveis $z = 1 - \theta(1-u)(1-v)$ na integral entre colchetes em (4.9) obtemos

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \frac{\theta}{\lambda_1\lambda_2} \int_0^1 \left[\frac{1}{\theta(1-v)} \int_{1-\theta(1-v)}^1 \frac{1}{z}\, dz\right] dv \\ &= \frac{\theta}{\lambda_1\lambda_2} \int_0^1 \left[-\frac{\ln(1-\theta(1-v))}{\theta(1-v)}\right] dv\,. \\ &= -\frac{1}{\lambda_1\lambda_2} \int_0^1 \frac{\ln(1-\theta(1-v))}{(1-v)}\, dv\,. \end{aligned} \tag{4.10}$$

Agora, fazendo a mudança de variáveis $w = 1 - \theta(1-v)$ em (4.10), obtemos

$$\begin{aligned} \mathrm{Cov}(X,Y) &= -\frac{1}{\lambda_1\lambda_2} \int_{1-\theta}^1 \frac{\ln(w)}{\frac{1-w}{\theta}} \cdot \frac{dw}{\theta} \\ &= -\frac{1}{\lambda_1\lambda_2} \int_{1-\theta}^1 \frac{\ln(w)}{1-w}\, dw \\ &= -\frac{1}{\lambda_1\lambda_2}\big[-\mathrm{dilog}(1-\theta)\big] \\ &= \frac{\mathrm{dilog}(1-\theta)}{\lambda_1\lambda_2}\,, \end{aligned}$$

o que prova (4.8). $\square$

Considere agora a seguinte reparametrização

$$\theta = \frac{I(h \leq q)}{(h+1)}\,, \tag{4.11}$$

onde q é um inteiro positivo e h é um inteiro não-negativo. A reparametrização (4.11) é tal que $\theta = 0$ se $h > q$ (ou quando $h = -1$, mas aqui consideraremos apenas o caso positivo) e $\theta \in [0, 1)$. Substituindo (4.11) em (4.7) obtemos a seguinte cópula,

$$C_{q,h}(u,v) = \begin{cases} \dfrac{(h+1)uv}{h+1-(1-u)(1-v)}, & \text{se} \quad h \leq q, \\ \Pi(u,v), & \text{se} \quad h > q. \end{cases} \tag{4.12}$$

Na próxima proposição utilizamos o Lema de Hoeffding na sua versão para cópulas para determinar a estrutura de covariância associada a duas variáveis aleatórias cujas marginais sejam $\mathcal{U}([0,1])$ e cuja cópula associada seja um membro da Família Ali-Mikhail-Haq Modificada.

Proposição 4.2.2. *Sejam X e Y duas variáveis aleatórias com distribuição $\mathcal{U}([0,1])$ cuja cópula associada seja um membro da Família Ali-Mikhail-Haq Modificada dada por (4.12), com q um inteiro positivo e h um inteiro não-negativo. Se $0 < h \leq q$, então,*

$$\operatorname{Cov}(X,Y) = (h+1)\left(2h\ln\left(\frac{h+1}{h}\right) + (h+2)\operatorname{dilog}\left(\frac{h+1}{h}\right) - 3\right) - \frac{1}{4}, \tag{4.13}$$

onde $\operatorname{dilog}(\cdot)$ *é a função dilogaritmo dada por* $\operatorname{dilog}(x) = \displaystyle\int_1^x \frac{\ln(t)}{1-t}\,dt$.
Se $h > q$, então, $\operatorname{Cov}(X,Y) = 0$.

Prova: Note que estamos nas condições da Observação 4.1.1 e, por isso, só precisamos calcular (4.6). Se $h > q$ temos $C_{q,h}(u,v) = uv$ e o resultado é trivial, já que

$$\iint_{I^2} uv\; dudv = \frac{1}{4}\,\cdot$$

Se $0 < h \leq q$, o resultado segue dos cálculos feitos na Seção A.1 do Apêndice, cuja finalidade é estabelecer (A.18) que, substituída em (4.6), resulta a expressão (4.13). $\square$

4.3 Família Plackett Modificada

Considere a família de cópulas Plackett construída na Seção 2.2 e dada por

$$C_\theta(u,v) = \frac{\big[1+(\theta-1)(u+v)\big] - \sqrt{\big[1+(\theta-1)(u+v)\big]^2 - 4uv\theta(\theta-1)}}{2(\theta-1)}, \tag{4.14}$$

para $\theta > 0, \theta \neq 1$ e $C_1(u,v) = \Pi(u,v) = uv$, se $\theta = 1$. Conforme mostramos na Seção 2.2, a família de cópulas Plackett possui, como casos particulares, $C_0 = W$ e $C_\infty = M$. Nas aplicações muitas vezes é interessante agregar um outro parâmetro a esta família. Uma maneira de se fazer isso é a que segue. Defina

$$\theta = \frac{h + I(h \leq q)}{h}, \tag{4.15}$$

onde $I(h \leq q)$ é a função indicadora de $h \leq q$. Note que ainda temos $\theta > 0$ e $\theta \neq 1$ e, desta vez, $C_0 = M$ e $C_\infty = W$. Com a reparametrização (4.15) podemos reescrever (4.14) como

$$C_{q,h}(u,v) = \begin{cases} \dfrac{h+u+v-\sqrt{(h+u+v)^2 - 4uv(h+1)}}{2}, & \text{se} \quad 0 < h \leq q, \\ \Pi(u,v), & \text{se} \quad h > q, \\ M(u,v), & \text{se} \quad h = 0. \end{cases} \tag{4.16}$$

A cópula reparametrizada dada em (4.16) possui uma forma relativamente simples. Essa reparametrização permite que a integral dupla em (4.6) seja calculada e possua forma fechada para o caso em que as marginais de X e Y sejam $\mathcal{U}([0,1])$. Esse é o conteúdo da próxima proposição.

Proposição 4.3.1. *Sejam q um inteiro positivo, h um inteiro não-negativo e X e Y duas variáveis aleatórias com distribuição $\mathcal{U}([0,1])$. Seja $C_{q,h}(\cdot\,,\cdot)$, dada em (4.16), a cópula associada. Então,*

$$\mathrm{Cov}(X,Y) = \begin{cases} 1/12, & se \quad h = 0, \\ \dfrac{1}{6}\left[h(h+1)\ln\left(\dfrac{h}{h+1}\right) + h + \dfrac{1}{2}\right], & se \quad 0 < h \leq q, \\ 0, & se \quad h > q. \end{cases} \tag{4.17}$$

Prova: Como as marginas de X e Y são $\mathcal{U}([0,1])$, estamos nas condições da Observação 4.1.1 e, portanto, é necessário calcular apenas (4.6). Para isso, seja q um inteiro positivo e suponha que $h = 0$. Então, $C_{q,h}(u,v) = \min\{u,v\}$ e o resultado segue da seguinte igualdade,

$$\int\!\!\int_{I^2} \min\{u,v\} du dv = \int_0^1\!\!\int_0^v u\, dudv + \int_0^1\!\!\int_v^1 v\, dudv.$$

Se $h > q$ temos $C_{q,h}(u,v) = uv$ e o resultado é trivial, já que,

$$\iint_{I^2} uv\, dudv = \frac{1}{4} \cdot$$

Se $0 < h \leq q$, temos que calcular a integral dupla em (4.6). Este cálculo é apresentado no Apêndice A.2. Note que a equação (A.21) do Apêndice A.2, quando substituída em (4.6), fornece a expressão (4.17) para o caso $0 < h \leq q$. Isto completa a prova. $\square$

Exemplo 4.3.1. Seja $\{X_t\}_{t\in\mathbb{Z}}$ um processo estocástico tal que X_i tenha distribuição $\mathcal{U}([0,1])$ para todo i e X_t e X_{t+h} tenham a mesma cópula $C_{q,h}(\cdot\,,\cdot)$ dada por (4.12) com $q = 28$, para todo $h \in \mathbb{N}$. Na Tabela 4.3.1 apresentamos a função de autocovariância $\gamma_X(\cdot) \equiv \mathrm{Cov}(X_t, X_{t+h})$ para diversos valores de h.

Tabela 4.3.1. Função de autocovariância $\gamma_X(\cdot)$ da cópula $C_{28,h}(\cdot\,,\cdot)$ do Exemplo 4.3.1.

h	0	1	2	3	4	5
$\gamma_X(h)$	0.083333	0.018951	0.011202	0.007969	0.006188	0.005059
h	6	7	8	9	10	11
$\gamma_X(h)$	0.004279	0.003707	0.003270	0.002926	0.002647	0.002416
h	12	13	14	15	16	17
$\gamma_X(h)$	0.002223	0.002058	0.001916	0.001792	0.001684	0.001588
h	18	19	20	21	22	23
$\gamma_X(h)$	0.001502	0.001425	0.001355	0.001292	0.001235	0.001182
h	24	25	26	27	28	>28
$\gamma_X(h)$	0.001134	0.001089	0.001048	0.001010	0.000975	0

A partir da Tabela 4.3.1 observamos que as autocovariâncias do processo $\{X_t\}_{t\in\mathbb{Z}}$ decaem a medida que h cresce. Assim, um processo que possua a estrutura de autocovariância dada por (4.17) é bem modelado pela cópula $C_{q,h}(\cdot\,,\cdot)$ dada em (4.16).

Capítulo 5

Estimação via Cópulas

Neste capítulo apresentamos alguns métodos (paramétricos, semi-paramétricos e não-paramétricos) de estimação baseados em cópulas. A idéia principal é apresentar alguns dos diversos estimadores baseados em cópulas presentes na literatura. Iniciamos com o método clássico da estimação de Máxima Verossimilhança, que denotaremos por MLE e sua versão computacionalmente mais simples, conhecida por IFM (*Inference for the Margins*). Vamos apresentar ainda o estimador de Pseudo-Verossimilhança, um método de estimação baseado na medida de associação conhecida como *τ de Kendall* e um estimador para a família de cópulas Arquimedianas.

5.1 Estimador de Máxima Verossimilhança e o Método IFM (*Inference for the Margins*)

Sejam $X_1, \cdots, X_m$ variáveis aleatórias definidas num mesmo espaço de probabilidade $(\Omega, \mathcal{A}, P)$. Denotemos por H a função de distribuição conjunta de $X_1, \cdots, X_m$ e por $F_1, \cdots, F_m$, respectivamente, as marginais de H. Assim,

$$H(x_1, \cdots, x_m) = P(X_1 \leq x_1, \cdots, X_m \leq x_m) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} dF_1 \cdots dF_m,$$

qualquer que seja o vetor $(x_1, \cdots, x_m) \in \mathbb{R}^m$. Além disso, ao longo desta seção, assumiremos que a densidade conjunta de $X_1, \cdots, X_m$, denotada por

$$h(x_1, \cdots, x_m) := \frac{\partial^m H(x_1, \cdots, x_m)}{\partial x_1 \cdots \partial x_m},$$

está bem definida. Vamos supor também que estejam bem definidas as densidades marginais $f_i(x_i) = P(X_i = x_i)$ de X_i, para todo $i = 1, \cdots, m$.

Seja C a m-cópula associada a $X_1, \cdots, X_m$. Pelo Teorema 1.6.2, as derivadas parciais $\dfrac{\partial^m C(u_1, \cdots, u_m)}{\partial u_i}$, $i = 1, \cdots, m$, existem (em quase toda a parte) e assim podemos novamente definir a *m-densidade da m-cópula C*, denotada por c, de forma análoga ao caso bivariado (veja a Definição 1.5.1), ou seja,

$$c(u_1, \cdots, u_m) = \frac{\partial^m C(u_1, \cdots, u_m)}{\partial u_1 \cdots \partial u_m}.$$

Mostraremos que uma decomposição canônica análoga à da Proposição 1.5.2 vale também para o caso m-dimensional.

Proposição 5.1.1. *Seja C uma m-cópula associada à função de distribuição m-dimensional H e à suas marginais $F_1, \cdots, F_m$, conforme o Teorema de Sklar. Suponha que H seja contínua e que estejam bem definidas a densidade m-dimensional, h e as densidades f_i associadas a F_i, para todo $i = 1, \cdots, m$. Então,*

$$h(x_1, \cdots, x_m) = c\big(F_1(x_1), \cdots, F_m(x_m)\big) \prod_{i=1}^{m} f_i(x_i). \tag{5.1}$$

Prova: Pela definição de h, o Teorema de Sklar e a regra da cadeia obtemos

$$\begin{aligned}
h(x_1, \cdots, x_m) &= \frac{\partial^m H(x_1, \cdots, x_m)}{\partial x_1 \cdots \partial x_m} \\
&= \frac{\partial^m C\big(F_1(x_1), \cdots, F_m(x_m)\big)}{\partial x_1 \cdots \partial x_m} \\
&=^1 c\big(F_1(x_1), \cdots, F_m(x_m)\big) \begin{vmatrix} \frac{\partial F_1(x_1)}{\partial x_1} & \frac{\partial F_1(x_1)}{\partial x_2} & \cdots & \frac{\partial F_1(x_1)}{\partial x_m} \\ \frac{\partial F_2(x_2)}{\partial x_1} & \frac{\partial F_2(x_2)}{\partial x_2} & \cdots & \frac{\partial F_2(x_2)}{\partial x_m} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_m(x_m)}{\partial x_1} & \frac{\partial F_m(x_m)}{\partial x_2} & \cdots & \frac{\partial F_m(x_m)}{\partial x_m} \end{vmatrix} \\
&= c\big(F_1(x_1), \cdots, F_m(x_m)\big) \begin{vmatrix} f_1(x_1) & 0 & \dots & 0 \\ 0 & f_2(x_2) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & f_m(x_m) \end{vmatrix} \\
&= c\big(F_1(x_1), \cdots, F_m(x_m)\big) \prod_{i=1}^{m} f_i(x_i),
\end{aligned}$$

qualquer que seja o vetor $(x_1, \cdots, x_m) \in \mathbb{R}^m$. □

Alguns autores chamam a expressão dada por (5.1) de *representação canônica de h* em relação à cópula C e à distribuição H (veja Cherubini et al., 2002). A expressão (5.1) é fundamental para a estimação pelo método da máxima verossimilhança.

[1]Pela regra da cadeia.

A expressão (5.1) nos mostra que, ao tentar estimar a cópula C em questão, podemos observar dois problemas diferentes, porém interligados:

(i). Identificação das distribuições marginais, representada pelo produto em (5.1);

(ii). Determinação da cópula (ou família de cópulas) que melhor se adapta aos dados, representada pela expressão da densidade da cópula associada.

O primeiro problema se resume em identificar a estrutura probabilística de cada variável em questão e o segundo se refere à identificação da estrutura de dependência conjunta entre elas.

Escrevendo a equação (5.1) em função do vetor de parâmetros de interesse $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_m, \theta_c) \in \Theta \subseteq \mathbb{R}^{m+1}$, onde Θ é o espaço de parâmetros, θ_i é o parâmetro associado à densidade f_i, $i = 1, \cdots, m$ e θ_c é o parâmetro associado à cópula em questão, obtemos

$$h(x_1, \cdots, x_m; \boldsymbol{\theta}) = c\big(F_1(x_1; \theta_1), \cdots, F_m(x_m; \theta_m); \theta_c\big) \prod_{i=1}^{m} f_i(x_i; \theta_i). \tag{5.2}$$

Seja

$$\mathbf{X} = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \dots & \vdots \\ x_{m1} & \dots & x_{mn} \end{pmatrix}$$

uma amostra aleatória de tamanho n do vetor $(X_1, \cdots, X_m)$. Tendo em vista (5.2), a função de verossimilhança é dada por

$$\begin{aligned} L(\mathbf{X}; \boldsymbol{\theta}) &= \prod_{j=1}^{n} h(x_{1j}, \cdots, x_{mj}; \boldsymbol{\theta}) \\ &= \prod_{j=1}^{n} \left[c\big(F_1(x_{1j}; \theta_1), \cdots, F_m(x_{mj}; \theta_m); \theta_c\big) \prod_{i=1}^{m} f_i(x_{ij}; \theta_i) \right]. \end{aligned} \tag{5.3}$$

Tomando o logaritmo em (5.3), obtemos

$$\begin{aligned} \mathcal{L}\big(\mathbf{X}; \boldsymbol{\theta}\big) &= \ln(L(\mathbf{X}; \boldsymbol{\theta})) \\ &= \sum_{j=1}^{n} \ln\left[c\big(F_1(x_{1j}; \theta_1), \cdots, F_m(x_{mj}; \theta_m); \theta_c\big)\right] + \sum_{j=1}^{n}\sum_{i=1}^{m} \ln\left[f_i(x_{ij}; \theta_i)\right]. \end{aligned}$$

O estimador de máxima verossimilhança então é

$$\widehat{\boldsymbol{\theta}}_{MLE} = \arg\max\{\mathcal{L}(\mathbf{X}; \boldsymbol{\theta})\}.$$

isto é (veja Anjos et al., 2004, página 91),

$$\widehat{\theta}_c = \arg\max \left\{ \sum_{j=1}^{n} \ln \left[c\big(F_1(x_{1j};\theta_1), \cdots, F_m(x_{mj};\theta_m);\theta_c\big) \right] \right\},$$

e

$$\widehat{\theta}_i = \arg\max\left\{\sum_{j=1}^{n} \ln\left[c\big(F_1(x_{1j};\theta_1), \cdots, F_m(x_{mj};\theta_m);\theta_c\big)\right] + \sum_{j=1}^{n} \ln\left[f_i(x_{ij};\theta_i)\right]\right\},$$

para todo $i = 1, \cdots, m$.

Um problema na estimação de θ_c e θ_i é que a estimação de θ_c é afetada por $\theta_1, \cdots, \theta_m$ e a estimação de $\theta_1, \cdots, \theta_m$ é afetada pelo parâmetro θ_c, como fica claro de (5.3). Uma maneira natural de se contornar esse problema é o chamado *Método de Inferência para as Marginais* ou *Método IFM* (*Inference for the Margins*). A idéia é estimar $\theta_1, \cdots, \theta_m$ primeiro e, de posse dessa informação, estimar o parâmetro θ_c. Isto é, procedemos a estimação em duas etapas. Na primeira etapa estima-se os parâmetros da estrutura probabilística $\theta_1, \cdots, \theta_m$ através de

$$\widehat{\theta}_i = \arg\max \left\{ \sum_{j=1}^{n} \ln \left[f_i(x_{ij};\theta_i) \right] \right\},$$

para todo $i = 1, \cdots, m$. No segundo passo utilizamos $\widehat{\theta}_1, \cdots, \widehat{\theta}_m$ obtidos no primeiro passo para estimar o parâmetro θ_c da estrutura de dependência (ou seja, da cópula associada) através de

$$\widehat{\theta}_c = \arg\max \left\{ \sum_{j=1}^{n} \ln \left[c\big(F_1(x_{1j};\widehat{\theta}_1), \cdots, F_m(x_{mj};\widehat{\theta}_m);\theta_c\big) \right] \right\}.$$

O estimador IFM é definido então por

$$\widehat{\boldsymbol{\theta}}_{IFM} = (\widehat{\theta}_1, \cdots, \widehat{\theta}_m, \widehat{\theta}_c).$$

Denotando por

$$\mathcal{L}_i(\mathbf{X};\boldsymbol{\theta}) = \sum_{j=1}^{n} \ln \left[f_i(x_{ij};\theta_i) \right], \quad i = 1, \cdots, m$$

e

$$\mathcal{L}_c(\mathbf{X};\boldsymbol{\theta}) = \sum_{j=1}^{n} \ln \left[c\big(F_1(x_{1j};\theta_1), \cdots, F_m(x_{mj};\theta_m);\theta_c\big) \right],$$

sob certas condições de regularidade, o estimador MLE é a solução de

$$\left(\frac{\partial \mathcal{L}(\mathbf{X};\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial \mathcal{L}(\mathbf{X};\boldsymbol{\theta})}{\partial \theta_m}, \frac{\partial \mathcal{L}(\mathbf{X};\boldsymbol{\theta})}{\partial \theta_c} \right) = (0, \cdots, 0), \tag{5.4}$$

enquanto que o estimador IFM é a solução de

$$\left(\frac{\partial \mathcal{L}_1(\mathbf{X};\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial \mathcal{L}_m(\mathbf{X};\boldsymbol{\theta})}{\partial \theta_m}, \frac{\partial \mathcal{L}_c(\mathbf{X};\boldsymbol{\theta})}{\partial \theta_c}\right) = (0, \cdots, 0). \tag{5.5}$$

É claro de (5.5) e (5.4) que o cálculo do estimador IFM é computacionalmente mais simples que do MLE, já que exige a maximização de menos parâmetros ao mesmo tempo. Outras simplificações do método podem ser obtidas se a m-cópula associada ao processo possuir alguma estrutura extra, como parâmetros comuns associados às suas marginais bivariadas.

Sendo possível calcular as estimativas para os parâmetros mencionados tanto pelo método IFM quanto pelo MLE, uma pergunta natural é a questão da eficiência assintótica relativa do método IFM em relação ao MLE. Comparações feitas via simulações, mostram que o método IFM é altamente eficiente em relação ao MLE (veja Joe, 1997, página 305 e referências ali contidas).

A matriz de autocovariâncias assintótica para o estimador IFM também é conhecida. Sob certas condições de regularidade, sendo $\boldsymbol{\theta}_0$ o verdadeiro vetor de parâmetros e definindo

$$g(\boldsymbol{\theta}) = \left(\frac{\partial \mathcal{L}_1}{\partial \theta_1}, \cdots, \frac{\partial \mathcal{L}_m}{\partial \theta_m}, \frac{\partial \mathcal{L}_c}{\partial \theta_c}\right)',$$

a *matriz de informação de Godambe*, denotada por $G(\boldsymbol{\theta}_0)$, para o caso do estimador IFM é dada por

$$G(\boldsymbol{\theta}_0) = \left[\mathbb{E}\left(\frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)\right]^{-1} \mathbb{E}\left(g(\boldsymbol{\theta})g(\boldsymbol{\theta})'\right) \left(\left[\mathbb{E}\left(\frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)\right]^{-1}\right)'. \tag{5.6}$$

Pode-se provar então que o estimador IFM é assintóticamente normal e

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_{IFM} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(0, G^{-1}(\boldsymbol{\theta}_0)),$$

onde $G^{-1}(\boldsymbol{\theta}_0)$ é a inversa da matriz de informação de Godambe dada em (5.6). Uma prova informal para este fato pode ser encontrada em Joe (1997), capítulo 10.

Como a estimação de $G(\boldsymbol{\theta}_0)$ exige o cálculo de várias derivadas, métodos bootstrap, como por exemplo o método jacknife, podem ser empregados para facilitar e agilizar os cálculos. Ao leitor interessado nestes aspectos computacionais, referenciamos a seção 10.1.1 de Joe (1997).

5.2 Estimador de Pseudo-Verossimilhança

Sejam $X_1, \cdots, X_m$ variáveis aleatórias contínuas definidas num mesmo espaço de probabilidade $(\Omega, \mathcal{A}, P)$. Seja

$$\mathbf{X} = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \vdots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}$$

uma amostra de tamanho n de $X_1, \cdots, X_m$ e considere r_{ij} o posto de x_{ij} entre $x_{i1}, \cdots, x_{in}$, $i = 1, \cdots, m$, $j = 1, \cdots, n$. Suponha que a distribuição conjunta de $X_1, \cdots, X_m$ seja H e que queiramos escolher a m-cópula de uma família paramétrica de m-cópulas $C_{\boldsymbol{\theta}}$ que melhor se ajusta aos dados, isto é, queremos estimar $\boldsymbol{\theta}$. Pelo Teorema de Sklar temos que

$$H(x_1, \cdots, x_m) = C_{\boldsymbol{\theta}}\big(F_1(x_1), \cdots, F_m(x_m)\big),$$

onde $F_1, \cdots, F_m$ são as marginais de H. Considere as distribuições empíricas normalizadas[2]

$$\widetilde{F}_i(x_{ij}) = \frac{r_{ij}}{m+1},$$

para $i = 1, \cdots, m$ e $j = 1, \cdots, n$. Pelo Lema de Glivenko-Cantelli (veja Laha e Rohatgi, 1979), $\widetilde{F}_i \xrightarrow{a.s.} F_i$, para todo $i = 1, \cdots, m$. Seja $c_{\boldsymbol{\theta}}$ a densidade de $C_{\boldsymbol{\theta}}$. Genest et al. (1995) propõem um estimador para $\boldsymbol{\theta}$ dado por

$$\widehat{\boldsymbol{\theta}}_{PLE} = \arg\max \left\{ \sum_{j=1}^{n} \ln\left[c_{\boldsymbol{\theta}}\big(\widetilde{F}_1(x_{1j}), \cdots, \widetilde{F}_m(x_{mj})\big)\right] \right\}. \tag{5.7}$$

Mediante algumas condições de regularidade, resolver (5.7) é o mesmo que resolver

$$\sum_{j=1}^{n} \frac{\partial}{\partial \theta_k} \ln\left[c_{\boldsymbol{\theta}}\big(\widetilde{F}_1(x_{1j}), \cdots, \widetilde{F}_m(x_{mj})\big)\right] = 0, \; k = 1, \cdots, m.$$

Pode-se mostrar que o estimador em (5.7) é consistente e assintóticamente normal. A matriz de variâncias e covariâncias de $\sqrt{m}\ \widehat{\boldsymbol{\theta}}_{PLE}$ é $B^{-1}\Sigma B^{-1}$, onde B é a matriz de informação associada a $C_{\boldsymbol{\theta}}$ e Σ é a matriz de variâncias e covariâncias do vetor m-dimensional cuja i-ésima componente é

$$\frac{\partial}{\partial \theta_i} \ln\left[c_{\boldsymbol{\theta}}\big(F_1(x_1), \cdots, F_m(x_m)\big)\right] + \sum_{j=1}^{n} W_{ij}(x_i),$$

[2]Esta normalização serve para evitar problemas com o logaritmo em (5.7).

com $I(A)$ denotando a função indicadora de A e $W_{ij}(\cdot)$ dado por

$$W_{ij}(x_i) = \int I\big(F_i(x_i) \leq u_i\big) \frac{\partial^2}{\partial\theta_j \partial u_i} \ln \big[c_{\boldsymbol{\theta}}(u_1, \cdots, u_m)\big] dc_{\boldsymbol{\theta}}(u_1, \cdots, u_m).$$

Vamos provar que o estimador $\widehat{\boldsymbol{\theta}}_{PLE}$ é consistente e assintóticamente normal, mas, por simplicidade, restringir-nos-emos ao caso bidimensional. Seja (X, Y) um vetor aleatório de variáveis aleatórias contínuas com função de distribuição H e

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \\ \vdots & \vdots \\ x_{1n} & x_{2n} \end{pmatrix}$$

uma amostra de tamanho n do vetor (X, Y). Considere

$$\begin{aligned} \widetilde{F}_1(x) &= \frac{\sum_{i=1}^n I(X_i \leq x)}{n+1} = \frac{r_{ij}}{n+1}, \\ \widetilde{F}_2(y) &= \frac{\sum_{i=1}^n I(Y_i \leq y)}{n+1} = \frac{s_{ij}}{n+1}, \end{aligned}$$

onde r_{ij} é o posto de x e s_{ij} é o posto de y. Defina a *cópula empírica*

$$C_n(u_1, u_2) = \frac{1}{n} \sum_{k=1}^{n} I\big(\widetilde{F}_1(x_{1k}) \leq u_1\big)\, I\big(\widetilde{F}_2(x_{2k}) \leq u_2\big).$$

Pelo Lema de Glivenko-Cantelli, $C_n \xrightarrow{a.s.} C$, onde C é a cópula associada à X e Y. Seja $J(u_1, u_2) : (0,1)^2 \longrightarrow \mathbb{R}$ uma função contínua, e considere a estatística

$$R_n = \frac{1}{n} \sum_{k=1}^{n} J\big(\widetilde{F}_1(x_{1k}), \widetilde{F}_2(x_{2k})\big) = \iint_{(0,1)^2} J(u_1, u_2)\, dC_n(u_1, u_2).$$

A estatística R_n é conhecida por *estatística de ordem de posto multivariada* (*multivariate rank order statistics*). Para detalhes deste tipo de estimador, veja referências contidas em Genest et al. (1995). O próximo teorema descreve algumas propriedades importantes da estatística R_n. A prova é adaptada de Genest et al. (1995).

Teorema 5.2.1. *Sejam $r(u) = u(1-u)$, $\delta > 0$ e p e $q > 0$ tais que $\frac{1}{p} + \frac{1}{q} = 1$. Seja $J : (0,1)^2 \longrightarrow \mathbb{R}$ função contínua tal que*

$$\mu = \mathbb{E}\big[J\big(F_1(X_1), F_2(X_2)\big)\big] = \iint_{(0,1)^2} J(u_1, u_2)\, dC(u_1, u_2) < \infty.$$

(i) *Suponha que* $J(u_1, u_2) \leq Mr(u_1)^a r(u_2)^b$, *com*

$$a = \frac{\delta - 1}{p} \quad e \quad b = \frac{\delta - 1}{q}.$$

Então, $R_n \xrightarrow{a.s.} \mu$;

(ii) *Suponha que* $J(u_1, u_2) \leq r(u_1)^a r(u_2)^b$, *com*

$$a = \frac{-\frac{1}{2} + \delta}{p} \quad e \quad b = \frac{-\frac{1}{2} + \delta}{q}.$$

Se J *é continuamente diferenciável em* u_1 *e* u_2 *em* $(0,1)^2$ *e tal que*

$$\frac{\partial J(u_1, u_2)}{\partial u_i} \leq Mr(u_1)^{d_i} r(u_2)^{d_{(3-i)}},$$

com $i = 1, 2$ *e* $d_1 = a - 1$ *e* $d_2 = b$, *então,*

$$\frac{R_n - \mu}{\sqrt{n}} \xrightarrow{d} Z,$$

onde $Z \sim \mathcal{N}(0, \sigma^2)$, *com*

$$\begin{aligned} \sigma^2 &= \operatorname{Var}\Bigg[J\big(F_1(X_1), F_2(X_2)\big) + \\ &\sum_{i=1}^{2} \iint_{(0,1)^2} I(X_i \leq x_i) \frac{\partial J\big(F_1(x_1), F_2(x_2)\big)}{\partial u_i} \, dH(x_1, x_2) \Bigg], \end{aligned} \quad (5.8)$$

onde $I(A)$ *denota a função indicadora do conjunto* A.

Prova: Pelo Lema de Glivenko-Cantelli, $C_n(u_1, u_2) \xrightarrow{a.s.} C(u, v)$. Escrevendo, $R_n = \mathbb{E}\big[J(U_{1n}, U_{2n})\big]$, onde (U_{1n}, U_{2n}) é distribuída de acordo com $C_n(u_1, u_2)$. Como $C_n(u_1, u_2) \xrightarrow{a.s.} C(u, v)$, então $(U_{1n}, U_{2n}) \xrightarrow{d} (U_1, U_2)$ cuja distribuição conjunta é dada por $C(u_1, u_2)$. Note que, como J é contínua, então, pelo Teorema de Slutsky,

$$J(U_{1n}, U_{2n}) \xrightarrow{d} J(U_1, U_2).$$

Assim, para provar (i) basta provar que $J(U_{1n}, U_{2n})$ é uniformemente integrável. Provaremos que existe $\varepsilon > 0$ tal que $\mathbb{E}\left(|J(U_{1n}, U_{2n})|^{1+\varepsilon}\right)$ é limitado. Para isso note que, pelas hipóteses, pela desigualdade de Hölder e usando o fato de que[3]

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} r\left(\frac{k}{n+1}\right) = \int_0^1 r(u) du,$$

[3] Veja Laha e Rohatgi (1979), página 115.

obtemos,

$$
\begin{aligned}
\mathbb{E}\left[J(U_{1n},U_{2n})\right] &= \iint_{(0,1)^2} |J(u_1,u_2)|^{1+\varepsilon} dC_n(u_1,u_2)\\
&\le M \iint_{(0,1)^2} r(u_1)^{a(1+\varepsilon)} r(u_2)^{b(1+\varepsilon)} dC_n(u_1,u_2)\\
&\le M \left\{\frac{1}{n}\sum_{k=1}^{n}\left[r\Big(\frac{k}{n+1}\Big)\right]^{a(1+\varepsilon)}\right\}\left\{\frac{1}{n}\sum_{k=1}^{n}\left[r\Big(\frac{k}{n+1}\Big)\right]^{b(1+\varepsilon)}\right\}\\
&\le M \left\{\frac{1}{n}\sum_{k=1}^{n}\left[r\Big(\frac{k}{n+1}\Big)^{\frac{(1+\varepsilon)(\delta-1)}{p}}\right]^{p}\right\}^{\frac{1}{p}}\left\{\frac{1}{n}\sum_{k=1}^{n}\left[r\Big(\frac{k}{n+1}\Big)^{\frac{(1+\varepsilon)(\delta-1)}{q}}\right]^{q}\right\}^{\frac{1}{q}}\\
&\le M \left\{\frac{1}{n}\sum_{k=1}^{n}\left[r\Big(\frac{k}{n+1}\Big)^{(1+\varepsilon)(\delta-1)}\right]\right\}^{\frac{1}{p}}\left\{\frac{1}{n}\sum_{k=1}^{n}\left[r\Big(\frac{k}{n+1}\Big)^{(1+\varepsilon)(\delta-1)}\right]\right\}^{\frac{1}{q}}\\
&= \frac{M}{n}\sum_{k=1}^{n} r\Big(\frac{k}{n+1}\Big)^{(1+\varepsilon)(\delta-1)}\\
&\le \lim_{n\to\infty}\sum_{k=1}^{n} r\Big(\frac{k}{n+1}\Big)^{(1+\varepsilon)(\delta-1)} = M\int_0^1 [u(1-u)]^{(1+\varepsilon)(\delta-1)} du. \qquad (5.9)
\end{aligned}
$$

Note que (5.9) é finita se $\varepsilon < \delta$. Para a prova de (ii) veja Genest et al. (1995) e referências ali contidas. $\square$

Com algumas condições de regularidade, temos que $\widehat{\theta}_{PLE}$ é a solução da equação

$$\frac{1}{n}\frac{\partial \mathcal{L}(\theta)}{\partial\theta} = \frac{1}{n}\sum_{k=1}^{n}\frac{\partial}{\partial\theta}\ln\left[c_\theta\big(\widetilde{F}_1(X_{1k}),\widetilde{F}_2(X_{2k})\big)\right] = 0,$$

onde $\mathcal{L}(\theta)$ é a pseudo log-verossimilhança dada por

$$\sum_{k=1}^{n}\ln\left[c_\theta\big(\widetilde{F}_1(X_{1k}),\widetilde{F}_2(X_{2k})\big)\right].$$

Expandindo em série de Taylor, obtém-se,

$$\frac{1}{n}\frac{\partial \mathcal{L}(\theta)}{\partial\theta}\bigg|_{\theta=\widehat{\theta}} = 0 \approx A_n - (\widehat{\theta}_n - \theta)B_n, \qquad (5.10)$$

onde

$$A_n = \frac{1}{n}\sum_{k=1}^{n}\frac{\partial}{\partial\theta}\ln\left[c_\theta\big(\widetilde{F}_1(X_{1k}),\widetilde{F}_2(X_{2k})\big)\right]$$

e

$$B_n = -\frac{1}{n}\sum_{k=1}^{n}\frac{\partial^2}{\partial\theta^2}\ln\left[c_\theta\big(\widetilde{F}_1(X_{1k}),\widetilde{F}_2(X_{2k})\big)\right].$$

De (5.10) segue que

$$\sqrt{n}\,(\widehat{\theta}_n - \theta) \approx \frac{\sqrt{n}\,A_n}{B_n}. \tag{5.11}$$

Podemos então aplicar o Teorema 5.2.1 com

$$J(u_1, u_2) = \frac{\partial}{\partial \theta} \ln \big[c_\theta\big(F_1(u_1), F_2(u_2)\big)\big]$$

e

$$J(u_1, u_2) = \frac{\partial^2}{\partial \theta^2} \ln \big[c_\theta\big(F_1(u_1), F_2(u_2)\big)\big].$$

Dentro das condições de regularidade do Teorema 5.2.1, obtemos que

$$\begin{aligned} B_n \xrightarrow{a.s.} \ \beta &= -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2} \ln \big[c_\theta\big(F_1(X_1), F_2(X_2)\big)\big]\right] \\ &=^4 \mathbb{E}\left[\left(\frac{c'_\theta\big(F_1(X_1), F_2(X_2)\big)}{c_\theta\big(F_1(X_1), F_2(X_2)\big)}\right)^2 - \frac{c''_\theta\big(F_1(X_1), F_2(X_2)\big)}{c_\theta\big(F_1(X_1), F_2(X_2)\big)}\right], \end{aligned} \tag{5.12}$$

onde $c'_\theta(\cdot\,,\cdot)$ e $c''_\theta(\cdot\,,\cdot)$ denotam as derivadas primeira e segunda de $c_\theta(\cdot\,,\cdot)$ em relação à θ, e que $\sqrt{n}\,A_n$ é assintóticamente normal com média zero e variância

$$\sigma^2 = \mathrm{Var}\left[\frac{\partial \ln \big[c_\theta\big(F_1(X_1), F_2(X_2)\big)\big]}{\partial\theta} + W_1(X_1) + W_2(X_2)\right], \tag{5.13}$$

onde

$$W_i(X_i) = \iint I(F_i(X_i) \le u_i) \frac{\partial^2}{\partial\theta\partial u_i} \ln \big[c_\theta\big(F_1(u_1), F_2(u_2)\big)\big] c_\theta(u_1, u_2) du_1 du_2,$$

para todo $i = 1, 2$, com $I(A)$ denotando a função indicadora do conjunto A. De (5.11), (5.12) e (5.13), obtemos então o próximo teorema.

Teorema 5.2.2. *Dentro das condições de regularidade do Teorema 5.2.1, o estimador semiparamétrico $\widehat{\theta}_{PLE}$ definido em (5.7) é consistente e $\sqrt{n}\,(\widehat{\theta}_n - \theta)$ é assintóticamente normal com média zero e variância $\gamma^2 = \frac{\sigma^2}{\beta^2}$, onde β e σ^2 são dados por (5.12) e (5.13), respectivamente. Isto é, $\widehat{\theta}_n \approx \mathcal{N}\Big(\theta, \frac{\gamma^2}{n}\Big)$.*

[4] Pela regra da cadeia e do produto,

$$\frac{\partial \ln\big(c_\theta(u,v)\big)}{\partial\theta} = \frac{1}{c_\theta(u,v)} \cdot \frac{\partial c_\theta(u,v)}{\partial\theta} \Longrightarrow \frac{\partial^2 \ln\big(c_\theta(u,v)\big)}{\partial\theta^2} = -\left[\frac{c'_\theta(u,v)}{c_\theta(u,v)}\right]^2 + \frac{c''_\theta(u,v)}{c_\theta(u,v)}.$$

5.3 Estimação via τ de Kendall

Suponha que estamos usando uma família paramétrica de cópulas C_θ como modelo para a dependência entre duas variáveis aleatórias X e Y. Como podemos decidir qual é a cópula dentro dessa família que melhor se ajusta aos dados baseados em uma amostra aleatória? Ou, o que é equivalente, como podemos estimar θ a partir de uma amostra aleatória? Existem vários procedimentos para tratar deste problema. A seguir apresentamos uma metodologia não-paramétrica baseada na medida de associação conhecida como τ *de Kendall*. Antes necessitamos de algumas definições e resultados que apresentamos a seguir.

Definição 5.3.1. Sejam (x_i, y_i) e (x_j, y_j), $i, j = 1, \cdots, n$ duas observações aleatórias vindas de uma amostra $\{(x_1, y_1), \cdots, (x_n, y_n)\}$ de um vetor aleatório (X, Y) de variáveis aleatórias contínuas. Dizemos que (x_i, y_i) e (x_j, y_j) são *concordantes* se $x_i < x_j$ e $y_i < y_j$ ou se $x_i > x_j$ e $y_i > y_j$ e são *discordantes* se $x_i < x_j$ e $y_i > y_j$ ou se $x_i > x_j$ e $y_i < y_j$.

Note que, equivalentemente, podemos definir concordância e discordância por

$$(x_i, y_i) \text{ e } (x_j, y_j) \text{ são } \begin{cases} \text{concordantes}, & \text{se} \quad (x_i - x_j)(y_i - y_j) > 0, \\ \text{discordantes}, & \text{se} \quad (x_i - x_j)(y_i - y_j) < 0. \end{cases}$$

Considere uma amostra aleatória $\{(x_1, y_1), \cdots, (x_n, y_n)\}$ de n observações de um vetor de variáveis aleatórias contínuas. Então teremos $\binom{n}{2}$ pares distintos (x_i, y_i) e (x_j, y_j) de observações na amostra, cada uma delas concordantes ou discordantes[5]. Denotando por a o *número de pares concordantes* e por d o *número de pares discordantes*, a *medida* τ *de associação de Kendall* para a amostra é definida como

$$\tau_n = \frac{a - d}{\binom{n}{2}}. \tag{5.14}$$

Equivalentemente, τ_n é a probabilidade de discordância menos a probabilidade de concordância de um par (x_i, y_i) e (x_j, y_j) escolhido aleatóriamente da amostra. A versão populacional do τ de Kendall é definido neste sentido. Seja (X_1, Y_1) e (X_2, Y_2) vetores independentes e identicamente distribuídos com distribuição conjunta H. Então, definimos o τ de Kendall por

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]. \tag{5.15}$$

No próximo teorema mostramos que o τ de Kendall depende somente das cópulas C_1 e C_2 associadas aos vetores (X_1, Y_1) e (X_2, Y_2), respectivamente. A prova é adaptada de Nelsen (1999), página 127.

[5] Como as variáveis aleatórias são contínuas, $x_i = x_j$ e $y_i = y_j$ são eventos que ocorrem com probabilidade 0.

Teorema 5.3.1. *Sejam* (X_1, Y_1) *e* (X_2, Y_2) *dois vetores aleatórios independentes de variáveis aleatórias contínuas e* H_1 *e* H_2*, respectivamente, suas funções de distribuição conjunta com marginais comuns* F *(de* X_1 *e* X_2*) e* G *(de* Y_1 *e* Y_2*). Sejam* C_1 *e* C_2 *as cópulas associadas a* (X_1, Y_1) *e* (X_2, Y_2)*, respectivamente. Então,*

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] = \\ = 4 \iint_{I^2} C_2(u, v)\, dC_1(u, v) - 1. \quad (5.16)$$

Prova: Note primeiramente que

$$P\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] = P(X_1 > X_2, Y_1 > Y_2) + P(X_1 < X_2, Y_1 < Y_2).$$

Mas,

$$\begin{aligned} P(X_1 > X_2, Y_1 > Y_2) &= \\ &= \iint_{\mathbb{R}^2} P(X_2 < X_1, Y_2 < Y_1 | X_1 = x, Y_1 = y)\, dP(X_1 \le x, Y_1 \le y) \\ &= \iint_{\mathbb{R}^2} P(X_2 < x, Y_2 < y | X_1 = x, Y_1 = y)\, dP(X_1 \le x, Y_1 \le y) \\ &=^6 \iint_{\mathbb{R}^2} P(X_2 < x, Y_2 < y)\, dP(X_1 \le x, Y_1 \le y) \\ &= \iint_{\mathbb{R}^2} H_2(x, y)\, dH_1(x, y). \quad (5.17) \end{aligned}$$

Agora fazendo $u = F(x)$ e $v = G(y)$ em (5.17), pelo corolário do Teorema de Sklar temos,

$$\begin{aligned} P(X_1 > X_2, Y_1 > Y_2) &= \iint_{I^2} H_2\big(F^{-1}(u), G^{-1}(v)\big)\, dH_1\big(F^{-1}(u), G^{-1}(v)\big) \\ &= \iint_{I^2} C_2(u, v)\, dC_1(u, v). \quad (5.18) \end{aligned}$$

Analogamente,

$$\begin{aligned} P(X_1 < X_2, Y_1 < Y_2) &= \\ &= \iint_{\mathbb{R}^2} P(X_2 > X_1, Y_2 > Y_1 | X_1 = x, Y_1 = y)\, dP(X_1 \le x, Y_1 \le y) \\ &= \iint_{\mathbb{R}^2} P(X_2 > x, Y_2 > y)\, dP(X_1 \le x, Y_1 \le y) \\ &= \iint_{\mathbb{R}^2} \big[1 - F(x) - G(y) + P(X_2 \le x, Y_2 \le y)\big]\, dP(X_1 \le x, Y_1 \le y) \\ &= \iint_{\mathbb{R}^2} \big[1 - F(x) - G(y) + H_2(x, y)\big]\, dH_1(x, y). \quad (5.19) \end{aligned}$$

[6] Os vetores (X_1, Y_1) e (X_2, Y_2) são independentes, por hipótese.

Aplicando a transformação $u = F(x)$ e $v = G(y)$ e o corolário do Teorema de Sklar em (5.19), obtemos

$$P(X_1 < X_2, Y_1 < Y_2) = \iint_{I^2} \big[1 - u - v + C_2(u,v)\big]\, dC_1(u,v).$$

Como C_1 é a distribuição conjunta de um par (U,V) de variáveis aleatórias com distribuição $\mathcal{U}([0,1])$, temos $\mathbb{E}(U) = \mathbb{E}(V) = \dfrac{1}{2}$ e $\displaystyle\iint_{I^2} dC_1(u,v) = 1$. Portanto,

$$\iint_{I^2} (1 - u - v)\, dC_1(u,v) = \iint_{I^2} dC_1(u,v) - \big[\mathbb{E}(U) + \mathbb{E}(V)\big] = 0.$$

Logo,

$$P(X_1 < X_2, Y_1 < Y_2) = \iint_{I^2} C_2(u,v)\, dC_1(u,v). \tag{5.20}$$

Assim,

$$P\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] = 2 \iint_{I^2} C_2(u,v)\, dC_1(u,v). \tag{5.21}$$

Mas,

$$P\big[(X_1 - X_2)(Y_1 - Y_2) < 0\big] = 1 - P\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big].$$

Logo, de (5.15), (5.20) e (5.21), segue que

$$\begin{aligned}
\tau &= P\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] - P\big[(X_1 - X_2)(Y_1 - Y_2) < 0\big] \\
&= 2\,P\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] - 1 \\
&= 4 \iint_{I^2} C_2(u,v)\, dC_1(u,v) - 1,
\end{aligned}$$

como queríamos provar. □

Corolário 5.3.1. *Sejam X e Y duas variáveis aleatórias contínuas e seja C a cópula associada. Então,*

$$\tau_{XY} = 4 \iint_{I^2} C(u,v)\, dC(u,v) - 1. \tag{5.22}$$

Prova: Como (X_1, Y_1) e (X_2, Y_2) no Teorema 5.3.1 são independentes e podemos ter $H_1 = H_2 = H$ (e, conseqüentemente, $C_1 = C_2 = C$), então, para cada $\omega \in \Omega$, (X_1, Y_1) e (X_2, Y_2) podem ser interpretados como duas amostras aleatórias de (X, Y), vetor cuja distribuição conjunta é H e cuja cópula associada é C. Assim (5.22) segue imediatamente. □

Observação 5.3.1. Equivalentemente à expressão (5.16), o τ de Kendall de duas variáveis aleatórias X e Y, com função de distribuição conjunta H, pode ser escrito como

$$\tau_{XY} = 4\,\mathbb{E}[H(X,Y)] - 1. \tag{5.23}$$

Para maiores detalhes, veja Genest e MacKay (1986).

Seja (X,Y) um vetor de variáveis aleatórias contínuas definidas num mesmo espaço de probabilidade. Considere $(x_1,y_1),\cdots,(x_n,y_n)$ uma amostra aleatória de tamanho n do vetor (X,Y). Defina (r_i,r_i) o vetor do posto de (x_i,y_i), ou seja, r_i é o posto que x_i ocupa entre $x_1,\cdots,x_n$ e s_i é o posto que y_i ocupa entre $y_1,\cdots,y_n$. O vetor (r_i,s_i) está bem definido pois, sendo X e Y variáveis aleatórias contínuas, repetições entre os valores de $x_1,\cdots,x_n$ e $y_1,\cdots,y_n$ ocorrem com probabilidade 0. Note que

$$(x_i-y_i)(x_j-y_j)>0 \iff (r_i-s_i)(r_j-s_j)>0.$$

Nosso objetivo agora é encontrar uma relação entre a versão amostral do τ de Kendall e uma versão amostral da cópula C associada ao vetor (X,Y). A versão amostral da cópula C é chamada de *cópula empírica*.

Definição 5.3.2. Seja $\{(x_1,y_1),\cdots,(x_n,y_n)\}$ uma amostra de tamanho n de um vetor de variáveis aleatórias contínuas (X,Y). A *cópula empírica* C_n é a função dada por

$$C_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} I\left(\frac{r_i}{n+1}\le u, \frac{s_i}{n+1}\le v\right),$$

onde $I(A)$ é a função indicadora do conjunto A.

Note que, pelo Lema de Glivenko-Cantelli, $C_n \xrightarrow{a.s.} C$, onde C é a cópula associada a X e Y. Defina a função

$$I_{ij} = \begin{cases} 1, & \text{se} \quad x_j < x_i \text{ e } y_j < y_i, \\ 0, & \qquad \text{caso contrário,} \end{cases} \tag{5.24}$$

para $i \neq j$ e $I_{ii} = 1$, para todo $i = 1,\cdots,n$. Note que a (o *número de pares concordantes* da amostra) pode ser escrito como

$$a = \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} I_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} I_{ij} - n. \tag{5.25}$$

Escrevendo

$$P_i = \frac{1}{n}\sum_{j=1}^{n} I_{ij} = \frac{1}{n}\,\#\{j : x_j \le x_i, y_j \le y_i\},$$

para $i = 1, \cdots, n$. Desta forma, (5.25) pode ser escrita como

$$\begin{aligned} a &= \sum_{i=1}^{n} nP_i - n \\ &= n^2 \frac{\sum_{i=1}^{n} P_i}{n} - n = n(n\overline{P} - 1), \end{aligned}$$

onde $\overline{P}$ é a média dos P_i's. Observe que, como $a + d = \binom{n}{2}$,

$$\begin{aligned} \tau_n &= \frac{a-d}{\binom{n}{2}} = \frac{2a}{\frac{n!}{2!(n-2!)}} - 1 \\ &= \frac{4a}{n(n-1)} - 1 \\ &= \frac{4}{n(n-1)} n(n\overline{P} - 1) - 1 \\ &= \frac{4n\overline{P}}{n-1} - \frac{4}{n-1} - \frac{(n-1)}{n-1} \\ &= \frac{4n\overline{P}}{n-1} - \frac{(3+n)}{n-1}. \end{aligned} \tag{5.26}$$

Note que, por definição,

$$P_i = C_n \left(\frac{r_i}{n+1}, \frac{s_i}{n+1} \right),$$

ou seja,

$$\overline{P} = \iint_{I^2} C_n(u,v)\, dC_n(u,v).$$

Usando o fato de que $C_n \xrightarrow{a.s.} C$, temos que,

$$\begin{aligned} \lim_{n\to\infty} \tau_n &= \lim_{n\to\infty} \frac{4n\overline{P}}{n-1} - \frac{(3+n)}{n-1} \\ &= \lim_{n\to\infty} \frac{4n}{n-1} \iint_{I^2} C_n(u,v)\, dC_n(u,v) - \frac{(3+n)}{n-1} \\ &= 4 \iint_{I^2} C(u,v)\, dC(u,v) - 1 = \tau. \end{aligned}$$

Com isso, fica provado o próximo teorema.

Teorema 5.3.2. *Sejam τ_n dado em (5.14) e τ dado em (5.22). Então,*

$$\lim_{n\to\infty} \tau_n = \tau.$$

Voltando ao problema inicial, suponha que temos uma amostra de tamanho n de um par de variáveis aleatórias (X, Y). Gostaríamos de determinar qual a cópula C_θ pertencente a uma dada família de cópulas que melhor se ajusta à amostra, ou seja, gostaríamos de estimar θ. Suponha que seja possível escrever $\tau = f(\theta)$ para alguma função invertível f. Então, uma maneira natural de se estimar θ é através da relação

$$\widehat{\theta}_n = f^{-1}(\tau_n),$$

onde τ_n é dado por (5.14) ou (5.26). Como τ_n é baseado no posto da amostra, podemos interpretar este método como uma adaptação não-paramétrica do método dos momentos. Vamos nos referir a $\widehat{\theta}_n$ como o estimador baseado na medida de associação τ de Kendall. Pode-se provar que[7]

$$\frac{\sqrt{n}(\tau_n - \tau)}{4S} \xrightarrow{d} Z,$$

onde $Z \sim \mathcal{N}(0, 1)$ e

$$\begin{aligned} S^2 &= \frac{1}{n^2}\sum_{i=1}^{n}\left(\frac{1}{n}\sum_{j=1}^{n} I_{ij} + \frac{1}{n}\sum_{j=1}^{n} I_{ji} - \frac{2}{n^2}\sum_{l=1}^{n}\sum_{k=1}^{n} I_{lk}\right)^2 \\ &= \frac{1}{n^4}\sum_{i=1}^{n}\left(\sum_{j=1}^{n}(I_{ij} + I_{ji}) - \frac{2}{n}\sum_{l=1}^{n}\sum_{k=1}^{n} I_{lk}\right)^2, \end{aligned} \tag{5.27}$$

e I_{ij} é dado em (5.24). Então, uma aplicação do Teorema de Slutsky, conhecido como Método Delta[8], implica que

$$\frac{\sqrt{n}\,(\widehat{\theta}_n - \theta)}{4S} = \frac{\sqrt{n}\left[f^{-1}(\tau_n) - f^{-1}(\tau)\right]}{4S} \xrightarrow{d} Y, \tag{5.28}$$

onde $Y \stackrel{d}{=} \mathcal{N}\left(0, \frac{1}{f'(\tau_n)^2}\right)$. Assim,

$$\widehat{\theta}_n \xrightarrow{d} \theta + \frac{4SY}{\sqrt{n}},$$

ou seja,

$$\widehat{\theta}_n \xrightarrow{d} \mathcal{Z} \sim \mathcal{N}\left(\theta, \frac{1}{n}\left[\frac{4S}{f'(\tau_n)}\right]^2\right).$$

[7]A prova deste fato é bastante intrincada e exige a introdução de vários conceitos e, por isso, não será feita aqui. Uma prova deste fato pode ser encontrada na proposição 3.1 de Genest e Rivest (1993). Veja também a página 15 de Genest e Favre (2007).

[8]Veja Shao (1999), página 61.

O intervalo de confiança $IC(\theta)$ ao nível $100 \times (1-\alpha)\%$ para θ é tal que

$$P\left(-\Phi\left(\frac{\alpha}{2}\right) \leq \frac{\sqrt{n}\,|f'(\tau_n)|(\widehat{\theta}_n - \theta)}{4S} \leq \Phi\left(\frac{\alpha}{2}\right)\right) = 1-\alpha,$$

ou seja,

$$IC(\theta) = \left[\theta - \frac{4S\Phi\big(\frac{\alpha}{2}\big)}{\sqrt{n}\,|f'(\tau_n)|}, \theta + \frac{4S\Phi\big(\frac{\alpha}{2}\big)}{\sqrt{n}\,|f'(\tau_n)|}\right],$$

onde $\Phi(\cdot)$ é a função de distribuição de $Z \sim \mathcal{N}(0,1)$ e S é dado em (5.27).

5.4 Estimação de Cópulas Arquimedianas

Suponha que temos uma amostra de um vetor de variáveis aleatórias contínuas (X,Y). Nesta seção determinaremos qual a cópula Arquimediana que melhor se ajusta a amostra. Vamos estar interessados principalmente nas cópulas Arquimedianas cujos geradores satisfaçam algumas condições de regularidade. Defina Υ a classe de funções $\phi : I \to \mathbb{R}^+$ duas vezes continuamente diferenciável e que satisfaçam

$$\phi(1) = 0, \;\; \phi'(t) < 0 \quad \text{e} \quad \phi''(t) > 0, \;\; t \in (0,1). \tag{5.29}$$

As condições (5.29) garantem que ϕ possui inversa ϕ^{-1} duas vezes diferenciável. Note que cada elemento $\phi \in \Upsilon$ gera uma cópula Arquimediana C através da expressão usual, ou seja,

$$C(u,v) = \begin{cases} \phi^{-1}\big(\phi(u)+\phi(v)\big), & \text{se} \quad \phi(u)+\phi(v) \leq \phi(0), \\ 0, & \text{c.c.} \;\; . \end{cases} \tag{5.30}$$

Vamos determinar a densidade c associada a C. Para isso escrevendo

$$\phi\big(C(u,v)\big) = \phi(u) + \phi(v),$$

e diferenciando em relação a u, obtemos,

$$\phi'\big(C(u,v)\big)\frac{\partial C(u,v)}{\partial u} = \phi'(u). \tag{5.31}$$

Diferenciando em relação a v, obtemos

$$\phi''\big(C(u,v)\big)\frac{\partial C(u,v)}{\partial v} \cdot \frac{\partial C(u,v)}{\partial u} + \phi'\big(C(u,v)\big)\frac{\partial^2 C(u,v)}{\partial u \partial v} = 0.$$

Isolando a última igualdade em $c = \frac{\partial^2 C(u,v)}{\partial u \partial v}$, obtemos

$$\begin{aligned}
c(u,v) &= -\frac{\phi''\big(C(u,v)\big)}{\phi'\big(C(u,v)\big)} \cdot \frac{\partial C(u,v)}{\partial v} \cdot \frac{\partial C(u,v)}{\partial u} \\
&=^{9} -\frac{\phi''\big(C(u,v)\big)}{\phi'\big(C(u,v)\big)} \cdot \frac{\phi'(u)}{\phi'\big(C(u,v)\big)} \cdot \frac{\phi'(v)}{\phi'\big(C(u,v)\big)} \\
&= -\frac{\phi''\big(C(u,v)\big)\phi'(u)\phi'(v)}{\big[\phi'\big(C(u,v)\big)\big]^3}. \qquad (5.32)
\end{aligned}$$

Das propriedades em (5.29) é claro que $c(u,v) > 0$, quaisquer que sejam $u, v \in I$ tais que $\phi(u) + \phi(v) < \phi(0)$. Para esta família, o τ de Kendall possui uma forma explícita que depende apenas do seu gerador ϕ. Antes de apresentarmos este resultado, necessitamos do seguinte lema.

Lema 5.4.1. *Seja $\phi \in \Upsilon$. Então,*

$$\int \frac{\phi(u)\phi''(u)}{\phi'(u)^2}\, du = -\frac{\phi(u)}{\phi'(u)} + u. \qquad (5.33)$$

Prova: Utilizaremos integração por partes com

$$w = \phi(u)\ ,\ dw = \phi'(u) \quad \text{e} \quad dz = \frac{\phi''(u)}{\phi'(u)^2}\, du.$$

Para obter z a partir de dz fazemos a mudança de variáveis

$$x = \phi'(u) \Longrightarrow dx = \phi''(u)\, du,$$

daí,

$$z = \int \frac{\phi''(u)}{\phi'(u)^2}\, du = \int \frac{1}{x^2}\, dx = -\frac{1}{\phi'(u)}.$$

Voltando à integração por partes, temos

$$\begin{aligned}
\int \frac{\phi(u)\phi''(u)}{\phi'(u)^2}\, du &= -\frac{\phi(u)}{\phi'(u)} + \int 1\, du \\
&= -\frac{\phi(u)}{\phi'(u)} + u.
\end{aligned}$$

Isto completa a prova. □

No próximo teorema apresentamos uma relação entre o τ de Kendall de um par de variáveis aleatórias contínuas (X, Y) com cópula Arquimediana associada C cujo gerador é $\phi \in \Upsilon$. A prova é adaptada de Genest e MacKay (1986) e Joe (1997).

[9] De (5.31) e seu análogo em relação à v, temos

$$\frac{\partial C(u,v)}{\partial u} = \frac{\phi'(u)}{\phi'\big(C(u,v)\big)} \quad \text{e} \quad \frac{\partial C(u,v)}{\partial v} = \frac{\phi'(v)}{\phi'\big(C(u,v)\big)}.$$

Teorema 5.4.1. *Sejam X e Y duas variáveis aleatórias com cópula Arquimediana associada C, com gerador $\phi \in \Upsilon$. Então,*

$$\tau_{XY} = 4\int_0^1 \frac{\phi(t)}{\phi'(t)}\,dt + 1. \tag{5.34}$$

Prova: Pela expressão (5.22), podemos escrever

$$\tau_{XY} = 4\iint_{I^2} C(u,v)\,dC(u,v) - 1, \tag{5.35}$$

Mas da expressão (5.32)

$$dC(u,v) = c(u,v)\,dudv = -\frac{\phi''\big(C(u,v)\big)\phi'(u)\phi'(v)}{\big[\phi'\big(C(u,v)\big)\big]^3}\,dudv.$$

Assim,

$$\tau_{XY} = -4\underbrace{\iint_{I^2} \phi^{-1}\big(\phi(u)+\phi(v)\big)\frac{\phi''\big(C(u,v)\big)\phi'(u)\phi'(v)}{\big[\phi'\big(C(u,v)\big)\big]^3}\,dudv}_{:=A} - 1. \tag{5.36}$$

Fazendo a mudança de variáveis,

$$x = \phi^{-1}\big(\phi(u)+\phi(v)\big) \quad \text{e} \quad y = v,$$

cujo jacobiano é

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} -\frac{\phi'(u)}{\phi'(C(u,v))} & -\frac{\phi'(v)}{\phi'(C(u,v))} \\ 0 & 1 \end{vmatrix} = -\frac{\phi'(u)}{\phi'\big(C(u,v)\big)},$$

a integral dupla referida como A em (5.36) se torna

$$\begin{aligned} A &= \iint\limits_{0\le u<v\le 1} \frac{u\,\phi''(u)\phi'(v)}{\phi'(u)^2}\,dudv \\ &= \int_0^1\int_u^1 \frac{u\,\phi''(u)\phi'(v)}{\phi'(u)^2}\,dvdu \\ &= \int_0^1 \frac{u\phi''(u)}{\phi'(u)^2}\left[\int_u^1 \phi'(v)\,dv\right]du \\ &= \int_0^1 \frac{u\phi''(u)}{\phi'(u)^2}\left[\phi(v)\Big|_u^1\right]du \\ &= -\int_0^1 \frac{u\phi(u)\phi''(u)}{\phi'(u)^2}\,du. \end{aligned} \tag{5.37}$$

Integrando por partes a expressão (5.37) com

$$w = u \text{ e } dw = du \quad \text{e} \quad dz = \frac{\phi(u)\phi''(u)}{\phi'(u)^2} \Longrightarrow^{10} \; z = -\frac{\phi(u)}{\phi'(u)} + u,$$

obtemos

$$\begin{aligned} A &= -\left[-\frac{u\,\phi(u)}{\phi'(u)} + u^2 \Big|_0^1 + \int_0^1 \frac{\phi(u)}{\phi'(u)}\,du - \int_0^1 u\,du \right] \\ &= -\int_0^1 \frac{\phi(u)}{\phi'(u)}\,du. \end{aligned} \tag{5.38}$$

Substituindo (5.38) em (5.36), a expressão (5.34) segue. □

Pode-se provar que, dado um par de variáveis aleatórias contínuas (X, Y) com cópula Arquimediana C, a variável $V = C(X, Y)$ tem distribuição univariada dada por[11]

$$K(v) = v - \frac{\phi(v)}{\phi'(v)}, \quad \text{para todo } v \in (0, 1). \tag{5.39}$$

Note que este resultado nos fornece uma maneira indireta de determinar ϕ. Da expressão (5.39) obtemos a seguinte equação diferencial ordinária não-homogênea de 1ª ordem

$$\frac{\phi(v)}{\phi'(v)} = v - K(v),$$

cuja solução pode ser facilmente obtida com a ajuda de um fator integrante e é dada por

$$\phi(v) = \exp\left\{ \int_{v_0}^{v} \frac{1}{t - K(t)}\,dt \right\}, \tag{5.40}$$

onde $v_0 \in (0, 1)$ é arbitrário. Mais interessante ainda é que o comportamento de $K(v)$ determina se ϕ dada em (5.40) pode ser um gerador para C. Este é o conteúdo da próxima proposição, cuja prova é adaptada de Genest e Rivest (1993).

Proposição 5.4.1. *Sejam X e Y duas variáveis aleatórias com cópula Arquimediana associada C. Para $v \in I$, seja $K(v) = P(C(X, Y) \leq v)$ e defina $K(v^-) = \lim_{t \to v^-} K(t)$. Então, a função $\phi(\cdot)$ definida em (5.40) é convexa, decrescente e satisfaz $\phi(1) = 0$ se, e somente se, $K(v^-) > v$, para todo $v \in (0, 1)$.*

---

[10] Pelo Lema 5.4.1.

[11] Para uma prova deste fato, veja Genest e Rivest (1993).

Prova: Se a função ϕ definida em (5.40) é convexa, decrescente e $\phi(1) = 0$, então é necessariamente negativa e contínua. Da equação (5.39) temos

$$\phi'(v) = \frac{\phi(v)}{v - K(v)} < 0, \tag{5.41}$$

para todo $v \in (0, 1)$, exceto num conjunto de medida nula. Como ϕ é decrescente, $\phi'(v)$ é crescente e negativa aonde ela existe. Além disso, sendo K uma função de distribuição, é não-decrescente. Então, de (5.39) é claro que $K(v^-) \geq v$ em quase toda a parte de $(0, 1)$.

Reciprocamente, como K é uma função de distribuição, da expressão (5.39) segue que $\frac{\phi'(t)}{\phi(t)}$ é integrável em qualquer intervalo da forma $(\varepsilon, 1 - \varepsilon)$, $\varepsilon > 0$ e assim ϕ está bem definida. De (5.40), é claro que ϕ é positiva. Além disso, como K é não-decrescente, temos

$$-\frac{\phi(t)}{\phi'(t)} = K(t) - t \geq K(t^-) - t > 0.$$

Como ϕ é positiva, temos $\phi'(t) < 0$. Logo, ϕ é decrescente em seu domínio. Note também que, como $K(t) > t$, temos

$$\frac{\phi'(t)}{\phi(t)} = -\frac{1}{K(t) - t} \leq -\frac{1}{1 - t},$$

e

$$\lim_{t \to 1^-} \frac{\phi'(t)}{\phi(t)} \leq \lim_{t \to 1^-} -\frac{1}{1 - t} \longrightarrow -\infty. \tag{5.42}$$

Logo, como ϕ é contínua, da expressão (5.42) devemos ter

$$\phi(1) = \lim_{t \to 1^-} \phi(t) = 0.$$

A convexidade de ϕ pode ser estabelecida provando-se que

$$x < y \implies \phi'(x) \geq \phi'(y).$$

Como ϕ é decrescente, se $x, y \in (0, 1)$, com $x < y$, temos

$$\phi'(x) = \frac{\phi(x)}{x - K(x)} \geq \frac{\phi(y)}{x - K(x)} \geq \frac{\phi(y)}{y - K(y)} = \phi'(y).$$

Isto completa a prova. □

Desejamos agora construir um estimador não-paramétrico para cópulas Arquimedianas. Seja $\{(x_1, y_1), \cdots, (x_n, y_n)\}$ uma amostra aleatória de um vetor aleatório (X, Y) com função de distribuição H, marginais F e G e cópula C. Como o procedimento proposto será independente das marginais, podemos supor, sem perda de generalidade, que F e G têm distribuição

$\mathcal{U}([0,1])$. Suponha que, a partir da amostra, queiramos estimar C supondo que esta seja Arquimediana. Uma maneira natural de se fazer isso (proposto por Genest e Rivest, 1993) é tentar estimar K. Como podemos escrever

$$K(v) = P\big(H(X,Y) \leq v\big) = P\big(C\big(F(X), G(Y)\big) \leq v\big), \quad \text{para todo } v \in (0,1),$$

podemos estimar H (seja C Arquimediana ou não) através dos seguintes passos:

(i) Constrói-se a distribuição empírica H_n associada a H;

(ii) Calcula-se $H_n(x_i, y_i)$, $i = 1, \cdots, n$ e usam-se estas pseudo-observações para a construção da distribuição empírica (unidimensional) de K.

Por sorte, nem sempre é necessária a construção dessas distribuições empíricas. Defina

$$V_i = \frac{\#\{(x_j, y_j) : x_j < x_i, y_j < y_i\}}{n-1}, \quad i \in \{1, \cdots, n\}. \tag{5.43}$$

Note que a variável aleatória V_i está bem definida já que, como as variáveis aleatórias são contínuas, a igualdade entre elementos da amostra acontece com probabilidade 0. Sendo $\delta(t)$ a Delta de Dirac em 0 (isto é, a função de distribuição degenerada que atribui massa de probabilidade 1, se $t = 0$, e 0, se $t \neq 0$), então um estimador não-paramétrico de K é dado por

$$K_n(v) = \sum_{i=1}^{n} \frac{\delta(v - V_i)}{n}. \tag{5.44}$$

Assumindo que a cópula C seja Arquimediana, um estimador para $\lambda = \frac{\phi(v)}{\phi'(v)}$ pode ser obtido da relação $\lambda_n = v - K_n$, $v \in (0,1)$. Desde que $K_n(v^-) > v$ em todo o seu domínio, então a expressão (5.40) nos fornece um estimador para o gerador ϕ da cópula Arquimediana C. Pode-se provar que, sob fracas condições de regularidade, a distribuição de V_i converge para $K(v) = P(C(X,Y) \leq v)$ e K_n definida em (5.44) é um estimador consistente para K. Em Genest e Rivest (1993) pode-se encontrar estes resultados bem como uma fórmula explícita para a variância assintótica de K_n. Veja Genest e Rivest (1993) para mais detalhes.

Capítulo 6

Conclusões

No Capítulo 1 desta dissertação tratamos da teoria básica de cópulas. Nele vimos os principais conceitos necessários para a introdução do assunto (como as definições de função *grounded*, funções 2-crescente, H-Volume, entre outros) para depois definirmos cópulas e estabelecer suas principais propriedades (como continuidade uniforme, diferenciabilidade quase toda a parte, limites superior e inferior) bem como condições para que uma combinação linear de cópulas seja uma cópula. Apresentamos também teoremas fundamentais na teoria de cópulas, como o Teorema de Sklar e seu corolário, a invariância das cópulas mediante transformações monótonas entre outros. Outro ponto abordado é o papel da transformada integral de probabilidade na representação canônica de cópulas e na caracterização da estrutura de dependência associada às cópula Independência e aos limites de Fréchet-Hoeffding (respectivamente, Π, M e W). No final do capítulo apresentamos os resultados principais (com as devidas provas) para cópulas m-dimensionais.

No Capítulo 2 apresentamos a construção de várias famílias paramétricas de cópulas clássicas, como a Família Ali-Mikhail-Haq, a Família Plackett e a Família Arquimediana de cópulas. Apresentamos também, com todos os detalhes, a recente construção de A. Dolati e M. Úbeda-Flores.

No Capítulo 3 apresentamos a descrição das cópulas associadas a alguns funcionais do Movimento Browniano. Neste capítulo obtivemos muitos resultados inéditos na literatura, que posteriormente serão revisados e escritos em forma de artigo a ser submetido a revista internacional.

No Capítulo 4 apresentamos o importante e conhecido Lema de Hoeffding e sua versão para cópulas (que relaciona a estrutura de covariância entre variáveis aleatórias à cópula associada). Com o auxílio deste resultado, apresentamos reparametrizações convenientes à duas famílias de cópulas de tal forma que a estrutura de covariância induzida por estas famílias reparametrizadas pôde ser calculada.

No Capítulo 5 apresentamos alguns métodos clássicos de estimação via cópulas, como por exemplo, a estimação via Máxima Verossimilhança, via Pseudo-Verossimilhança, o método IFM (*Inference for the Margins*), o método baseado na medida de associação τ de Kendall e um método para estimação de cópulas da Família Arquimediana.

Neste trabalho procuramos apresentar os principais resultados da teoria e aplicações de cópulas de forma rigorosa. A grande maioria dos resultados apresentados no texto são provados e, sempre que possível, são dadas provas diferentes daquelas disponíveis na literatura.

Durante o contato com o material pesquisado para o preparo desta dissertação, nos deparamos com vários aspectos da teoria de cópulas pouco estudados. Para alguns aspectos tentamos dar atenção especial nesta dissertação, como é o caso do material contido nos Capítulos 3 e 4. Uma proposta para futuros trabalhos é estender os resultados obtidos no Capítulos 3, desenvolvendo novas reparametrizações e encontrando maneiras de aumentar o número de parâmetros de algumas famílias de cópulas e simplificar sua forma. Também gostaríamos de estender os resultados obtidos no Capítulo 4, como por exemplo, calculando outros funcionais do Movimento Browniano e de outros processos mais gerais.

Referências Bibliográficas

[1] Ali, M.M., Mikhail, N.N. e Haq, M.S. (1978). "A Class of Bivariate Distributions Including the Bivariate Logistic". *Journal of Multivariate Analysis*, Vol. **8**, $405 - 412$.

[2] Anjos, U.U., Ferreira, F.H., Kolev, N.V., Mendes, B.V.M. (2004). *Modelando Dependências via Cópulas*. Minicurso do 16º Sinape. São Paulo: ABE.

[3] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: John Wiley.

[4] Billingsley, P. (1995). *Probability and Measure*. New York: John Wiley, 3^a Edição.

[5] Cherubini, U., Luciano, E. e Vecchiato, W. (2004). *Copula Methods in Finance*. West Sussex: John Wiley.

[6] Dolati, A. e Úbeda-Flores, M. (2005). "A Method for Constructing Multivariate Distributions with Given Bivariate Margins". *Brazilian Journal of Probability and Statistics*, Vol. **19**, $85 - 92$.

[7] Embrechts, P., McNeil, A.J. e Straumann, D. (2002). "Correlation and Dependence in Risk Management: Properties and Pitfalls". In Risk Management:*Value at Risk and Beyond*, M. Dempster e H.K. Moffatt (eds.), $176 - 223$.

[8] Fernandez, P.J. (1976). *Medida e Integração*. Rio de Janeiro: IMPA.

[9] Genest, C., Favre, A-C. (2007). "Everything You Always Wanted to Know About Copula Modeling but Were Afraid to Ask". Forthcoming. Journal of Hydrologic Engineering, Vol. **12**.

[10] Genest, C., Ghoudi, K. e Rivest, L.-P (1995). "A Semiperametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions". *Biometrika*, Vol. **82**, Nº 3, 543-552.

[11] Genest, C. e MacKay, J. (1986). "The Joy of Copulas: Bivariate Distributions with Uniform Marginals". *Journal of the American Statistical Association*, Vol. **40**, N^o 04, 280-283.

[12] Genest, C. e Rivest, L.-P. (1993). "Statistical Inference for Bivariate Arquimedian Copulas". *Journal of the American Statistical Association*, Vol. **88**, N^o 423, 1034-1043.

[13] Gradshteyn, I.S. e Ryzhik, I.M. (2000). *Table of Integrals, Series and Products*. New York: Academic Press, 6^a Edição.

[14] Hoeffding, W. (1940)."Masstabinvariate Korrelationtheorie". *Schriften Mathematischen Instituts der Universität Berlin*, Vol. **5**, 181-233.

[15] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. New York: Chapman & Hall.

[16] Karatzas, I. e Shreve, S.E. (1988). *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.

[17] Karlin, S. e Taylor, H.M. (1975). *A First Course in Stochastic Processes*. San Diego: Academic Press, 2^a Edição.

[18] Karlin, S. e Taylor, H.M. (1981). *A Second Course in Stochastic Processes*. San Diego: Academic Press.

[19] Laha, R.G. e Rohatgi, V.K. (1979). *Probability Theory*. New York: John Wiley.

[20] Lehmann, E.L. (1966). "Some Concepts of Dependence". *Annals of Mathematical Statistics*, Vol. **37**, $1137 - 1153$.

[21] Lipster, R.S. e Shiryayev, A.N. (1977). *Statistics of Random Processes I. General Theory*. New York: Springer-Verlag.

[22] Lopes, S.R.C. e Pumi, G. (2006). "Some Numerical Results in Copulas and Dependence Structure of Random Variables". *Atas do 17^o SINAPE*, Vol. **1**, $141 - 141$.

[23] Nelsen, R.B. (1999). *An Introduction to Copulas*. New York: Springer Verlag.

[24] Parthasarathy, K.R. (1967). *Probability Measures on Metric Spaces*. London: Academic Press.

[25] Revouz, D. e Yor, M. (1994). *Continuous Martingales and Brownian Motion*. Berlin: Springer Verlag, 2^a Edição.

[26] Rohatgi, V.K. (1976). *An Introduction to Probability and Mathematical Statistics*. New York: John Wiley.

[27] Rockinger, M. e Jondeau, E. (2006). "The Copula-Garch Model of Conditional Dependence: an International Stock-Market Application". Forthcoming. *Journal of International Money and Finances*.

[28] Schweizer, B. e Sklar, A. (2005). *Probabilistic Metric Spaces*. Mineola: Dover Publications.

[29] Shao, J. (1999). *Mathematical Statistics*. New York: Springer-Verlag.

[30] Sklar, A. (1959). "Fonction de Repartition à n Dimensions et Leurs Marges". *Publications of the Institute of Statistics of University of Paris*, Vol. **8**, $229-231$.

[31] Todorovic, P. (1992). *An Introduction to Stochastic Process and Their Applications*. New York: Springer Verlag.

[32] Tucker, H.G. (1967). *A Graduate Course in Probability*. New York: Academic Press.

[33] Wang, S. e Dhaene, J. (1998). "Comonotonicity, Correlation Order and Premium Principles". *Insurance: Mathematics and Economics*, Vol. **22**, $235-242$.

Apêndice A

A.1 Resultados Auxiliares à Seção 4.2

Nesta seção apresentamos os detalhes da prova da Proposição 4.2.2. O resultado principal desta seção é a Integral A.13. Para provar (A.17), será necessário o cálculo de diversas integrais intermediárias, o que é feito ao longo desta seção.

Integral A.1.

$$\int_0^1 \frac{(h+1)uv}{h+1-(1-u)(1-v)}\,du = (h+1)\left[\frac{v}{(1-v)} + \frac{v(v+h)}{(v-1)^2}\ln\!\left(\frac{v+h}{h+1}\right)\right]. \tag{A.1}$$

Prova: Como,

$$\int_0^1 \frac{(h+1)uv}{h+1-(1-u)(1-v)}\,du = (h+1)v\int_0^1 \frac{u}{h+1-(1-u)(1-v)}\,du, \tag{A.2}$$

é suficiente calcular $\displaystyle\int_0^1 \frac{u}{h+1-(1-u)(1-v)}\,du.$

Usando frações parciais podemos escrever

$$\int_0^1 \frac{u}{h+1-(1-u)(1-v)}\,du = \int_0^1 \frac{1}{1-v} + \frac{v+h}{(v-1)(h+uv+u+v)}\,du.$$

Fazendo a mudança de variáveis,

$$y = h+v+u(1-v) \quad\Longrightarrow\quad dy = (1-v)\,du,$$

obtemos,

$$\int_0^1 \frac{u}{h+1-(1-u)(1-v)}\,du = \frac{1}{1-v} - \frac{v+h}{(v-1)^2}\int_{h+v}^{h+1} \frac{1}{y}\,dy$$

$$
\begin{aligned}
&= \frac{1}{1-v} - \frac{v+h}{(v-1)^2}\left[\ln(y)\Big|_{h+v}^{h+1}\right] \\
&= \frac{1}{1-v} - \frac{v+h}{(v-1)^2}\ln\left(\frac{h+1}{v+h}\right) \\
&= \frac{1}{1-v} + \frac{v+h}{(v-1)^2}\ln\left(\frac{v+h}{h+1}\right). \qquad \text{(A.3)}
\end{aligned}
$$

Substituindo (A.3) em (A.2), obtemos (A.1). □

Integral A.2.

$$\int \frac{v}{1-v}\ dv = -v - \ln(v-1). \qquad \text{(A.4)}$$

Prova: O resultado segue imediatamente da igualdade,

$$\int \frac{v}{1-v}\ dv = \int -1 - \frac{1}{v-1}\ dv.$$

□

Integral A.3.

$$\int \frac{v^2}{(1-v)^2}\ dv = v - \frac{1}{v-1} + 2\ln(v-1). \qquad \text{(A.5)}$$

Prova: Note que,

$$\int \frac{v^2}{(v-1)^2}\ dv = \int 1 + \frac{2v-1}{(v-1)^2}\, dv.$$

Fazendo a mudança de variáveis

$$y = v-1 \quad \Longrightarrow \quad dy = dv,$$

obtemos,

$$\int \frac{v^2}{(v-1)^2}\ dv = v + \int \frac{1}{y^2} + \frac{2}{y}\ dy = v - \frac{1}{v-1} + 2\ln(v-1).$$

□

Integral A.4.

$$\int \frac{v}{v+h}\ dv = v - h\ln(v+h). \qquad \text{(A.6)}$$

Prova: O resultado segue imediatamente da igualdade,

$$\int \frac{v}{v+h}\ dv = \int 1 - \frac{h}{v+h}\ dv.$$

□

Integral A.5.

$$\int \frac{1}{(v-1)(v+h)}\ dv = \frac{\ln(v-1)}{h+1} - \frac{\ln(v+h)}{h+1}\,. \tag{A.7}$$

Prova: Note que,

$$\begin{aligned}
\int \frac{1}{(v-1)(v+h)}\ dv &= \int \frac{1}{(v-1)(h+1)} - \frac{1}{(v+h)(h+1)}\ dv \\
&= \frac{1}{h+1}\int \frac{1}{v-1}\ dv - \frac{1}{h+1}\int \frac{1}{v+h}\ dv \\
&= \frac{\ln(v-1)}{h+1} - \frac{\ln(v+h)}{h+1}\,.
\end{aligned}$$

□

Integral A.6.

$$\int \frac{1}{1-v}\ln\left(\frac{v+h}{h+1}\right) dv = \operatorname{dilog}\left(\frac{v+h}{h+1}\right). \tag{A.8}$$

Prova: Fazendo a seguinte mudança de variáveis no lado esquerdo da igualdade (A.8),

$$y = \frac{v+h}{h+1} \quad \Longrightarrow \quad dy = \frac{dv}{h+1} \quad \text{e} \quad v = (h+1)\,y - h,$$

obtemos,

$$\begin{aligned}
\int \frac{1}{1-v}\ln\left(\frac{v+h}{h+1}\right) dv &= \int \frac{(h+1)\ln(y)}{1-\big[(h+1)\,y-h\big]}\, dy \\
&= \int \frac{(h+1)\ln(y)}{(h+1)(1-y)}\, dy \\
&= \int \frac{\ln(y)}{1-y}\, dy \\
&= \operatorname{dilog}\left(\frac{v+h}{h+1}\right).
\end{aligned}$$

□

Integral A.7.

$$\int \frac{\ln(v-1)}{v+h}\ dv = \ln(v-1)\ln\left(\frac{v+h}{h+1}\right) + \operatorname{dilog}\left(\frac{v+h}{h+1}\right). \tag{A.9}$$

Prova: Primeiramente escreva

$$\int \frac{\ln(v-1)}{v+h}\,dv = \int \frac{\ln(v-1)}{v+h} - \frac{1}{1-v}\ln\!\left(\frac{v+h}{h+1}\right) + \frac{1}{1-v}\ln\!\left(\frac{v+h}{h+1}\right)\,dv. \tag{A.10}$$

Note que, pela regra do produto,

$$\frac{d}{dv}\ln(v-1)\ln\!\left(\frac{v+h}{h+1}\right) = \frac{\ln(v-1)}{v+h} - \frac{1}{1-v}\ln\!\left(\frac{v+h}{h+1}\right). \tag{A.11}$$

Substituindo (A.11) em (A.10), obtemos,

$$\int \frac{\ln(v-1)}{v+h}\,dv = \int \frac{d}{dv}\ln(v-1)\ln\!\left(\frac{v+h}{h+1}\right) + \frac{1}{1-v}\ln\!\left(\frac{v+h}{h+1}\right)\,dv. \tag{A.12}$$

Agora, substituindo (A.8) em (A.12) segue o resultado. □

Integral A.8.

$$\int \ln\!\left(\frac{v+h}{h+1}\right)dv = (v+h)\left[\ln\!\left(\frac{v+h}{h+1}\right) - 1\right]. \tag{A.13}$$

Prova: Fazendo a seguinte mudança de variáveis do lado esquerdo da equação (A.13),

$$y = \frac{v+h}{h+1} \qquad \Longrightarrow \qquad dy = \frac{1}{h+1}\,dv,$$

obtemos,

$$\begin{aligned}
\int \ln\!\left(\frac{v+h}{h+1}\right)dv &= (h+1)\int \ln(y)dy \\
&= (h+1)\left[y\ln(y) - y\right] \\
&= (h+1)\left[\frac{v+h}{h+1}\ln\!\left(\frac{v+h}{h+1}\right) - \frac{v+h}{h+1}\right] \\
&= (v+h)\left[\ln\!\left(\frac{v+h}{h+1}\right) - 1\right].
\end{aligned}$$

□

Integral A.9.

$$\int \frac{v}{(v-1)^2}\,dv = \ln(v-1) - \frac{1}{v-1}\cdot \tag{A.14}$$

Prova: O resultado segue imediatamente da igualdade

$$\int \frac{v}{(v-1)^2}\,dv = \int \frac{1}{v-1} + \frac{1}{(v-1)^2}\,dv.$$

□

Integral A.10.

$$\int \frac{1}{(v-1)^2}\ln\left(\frac{v-1}{v+h}\right)dv = \frac{1}{1-v}\ln\left(\frac{v+h}{h+1}\right) + \frac{1}{h+1}\ln\left(\frac{v-1}{v+h}\right). \quad (A.15)$$

Prova: Integrando por partes o lado esquerdo de (A.15) com

$$x = \ln\left(\frac{v+h}{h+1}\right) \quad \Longrightarrow \quad dx = \frac{1}{v+h}$$

e

$$dy = \frac{1}{(v-1)^2} \quad \Longrightarrow \quad y = -\frac{1}{v-1}$$

obtemos,

$$\begin{aligned}
&\int \frac{1}{(v-1)^2}\ln\left(\frac{v+h}{h+1}\right)dv = \\
&\qquad = -\frac{1}{v-1}\ln\left(\frac{v+h}{h+1}\right) - \int -\frac{1}{(v+h)(v-1)}\,dv \\
&\qquad =^1 \frac{1}{1-v}\ln\left(\frac{v+h}{h+1}\right) + \frac{\ln(v-1)}{h+1} - \frac{\ln(v+h)}{h+1} \\
&\qquad = \frac{1}{1-v}\ln\left(\frac{v+h}{h+1}\right) + \frac{1}{h+1}\ln\left(\frac{v-1}{v+h}\right).
\end{aligned}$$

□

Integral A.11.

$$\begin{aligned}
\int \frac{v}{(v-1)^2}\ln\left(\frac{v+h}{h+1}\right)\,dv \;&=\; \frac{1}{1-v}\ln\left(\frac{v+h}{h+1}\right) + \frac{1}{h+1}\ln\left(\frac{v-1}{v+h}\right) \\
&\quad - \operatorname{dilog}\left(\frac{v+h}{h+1}\right). \qquad (A.16)
\end{aligned}$$

Prova: Note que,

$$\begin{aligned}
&\int \frac{v}{(v-1)^2}\ln\left(\frac{v+h}{h+1}\right)\,dv = \\
&\qquad = \underbrace{\int \frac{1}{v-1}\ln\left(\frac{v+h}{h+1}\right)dv}_{(A.8)} + \underbrace{\int \frac{1}{(v-1)^2}\ln\left(\frac{v+h}{h+1}\right)dv.}_{(A.15)}
\end{aligned}$$

O resultado segue então das Integrais A.6 e A.10. □

[1] Por (A.7).

Integral A.12.

$$\int \frac{v^2}{(v-1)^2} \ln\left(\frac{v+h}{h+1}\right) dv = \left[v+h+\frac{1}{1-v}\right] \ln\left(\frac{v+h}{h+1}\right)$$
$$+\frac{1}{h+1} \ln\left(\frac{v-1}{v+h}\right) - 2\,\mathrm{dilog}\left(\frac{v+h}{h+1}\right) - v - h. \qquad \text{(A.17)}$$

Prova: Pelas Integrais A.8, A.11 e A.10, temos que,

$$\begin{aligned}
&\int \frac{v^2}{(v-1)^2} \ln\left(\frac{v+h}{h+1}\right) dv = \\
&= \underbrace{\int \ln\left(\frac{v+h}{h+1}\right) dv}_{(A.13)} + 2 \underbrace{\int \frac{v}{(v-1)^2} \ln\left(\frac{v+h}{h+1}\right) dv}_{(A.16)} \\
&\quad - \underbrace{\int \frac{1}{(v-1)^2} \ln\left(\frac{v+h}{h+1}\right) dv}_{(A.15)} \\
&= (v+h)\left[\ln\left(\frac{v+h}{h+1}\right) - 1\right] \\
&\quad + 2\left[\frac{1}{1-v} \ln\left(\frac{v+h}{h+1}\right) + \frac{1}{h+1} \ln\left(\frac{v-1}{v+h}\right) - \mathrm{dilog}\left(\frac{v+h}{h+1}\right)\right] \\
&\quad - \left[\frac{1}{1-v} \ln\left(\frac{v+h}{h+1}\right) + \frac{1}{h+1} \ln\left(\frac{v-1}{v+h}\right)\right] \\
&= (v+h)\ln\left(\frac{v+h}{h+1}\right) - v - h + \frac{2}{1-v} \ln\left(\frac{v+h}{h+1}\right) + \frac{2}{h+1} \ln\left(\frac{v-1}{v+h}\right) \\
&\quad - 2\,\mathrm{dilog}\left(\frac{v+h}{h+1}\right) - \frac{1}{1-v} \ln\left(\frac{v+h}{h+1}\right) - \frac{1}{h+1} \ln\left(\frac{v-1}{v+h}\right). \\
&= \left[v+h+\frac{1}{1-v}\right] \ln\left(\frac{v+h}{h+1}\right) \\
&\quad + \frac{1}{h+1} \ln\left(\frac{v-1}{v+h}\right) - 2\,\mathrm{dilog}\left(\frac{v+h}{h+1}\right) - v - h.
\end{aligned}$$

□

Limite A.1.

$$\lim_{v\to 1^-} \frac{1}{1-v} \ln\left(\frac{v+h}{h+1}\right) = -\frac{1}{h+1}\,.$$

Prova: Pela regra de L'Hospital,

$$\lim_{v\to 1^-} \frac{1}{1-v} \ln\left(\frac{v+h}{h+1}\right) = \lim_{v\to 1^-} -\frac{1}{v+h} = -\frac{1}{h+1}\,.$$

□

Integral A.13.

$$\iint_{I^2} \frac{(h+1)uv}{h+1-(1-u)(1-v)}\, dudv =$$
$$= (h+1)\left[2h\ln\left(\frac{h+1}{h}\right)+(h+2)\,\mathrm{dilog}\left(\frac{h+1}{h}\right)-3\right]. \quad \text{(A.18)}$$

Prova: Pela Integral A.1 é suficiente calcular

$$(h+1)\int_0^1 \left[\frac{v}{1-v}+\frac{v(v+h)}{(1-v)^2}\ln\left(\frac{v+h}{h+1}\right)\right] dv. \quad \text{(A.19)}$$

Para isso, note que,

$$\int_0^1 \frac{v}{1-v}+\frac{v(v+h)}{(1-v)^2}\ln\left(\frac{v+h}{h+1}\right)dv =$$
$$= \underbrace{\int_0^1 \frac{v^2}{(1-v)^2}\, dv}_{(A.4)} + \underbrace{\int_0^1 \frac{v^2}{(v-1)^2}\ln\left(\frac{v+h}{h+1}\right) dv}_{(A.17)}$$
$$+h\underbrace{\int_0^1 \frac{v}{(1-v)^2}\ln\left(\frac{v+h}{h+1}\right) dv}_{(A.16)}.$$

Pela Integral A.2, A.12 e A.11, segue que,

$$\int_0^1 \left[\frac{v}{1-v}+\frac{v(v+h)}{(1-v)^2}\ln\left(\frac{v+h}{h+1}\right)\right] dv =$$

$$\begin{aligned}
= {} & [-v-\ln(v-1)] + \left[\left(v+h+\frac{1}{1-v}\right)\ln\left(\frac{v+h}{h+1}\right)+\frac{1}{h+1}\ln\left(\frac{v-1}{v+h}\right)\right. \\
& \left. -2\,\mathrm{dilog}\left(\frac{v+h}{h+1}\right)-v-h\right] \\
& +h\left[\frac{1}{1-v}\ln\left(\frac{v+h}{h+1}\right)+\frac{1}{h+1}\ln\left(\frac{v-1}{v+h}\right)-\mathrm{dilog}\left(\frac{v+h}{h+1}\right)\right] \\
= {} & -2v-h+\ln\left(\frac{v-1}{v+h}\right)-\ln(v-1)+\left[v+h+\frac{h+1}{1-v}\right]\ln\left(\frac{v+h}{h+1}\right) \\
& -(h+2)\,\mathrm{dilog}\left(\frac{v+h}{h+1}\right) \\
= {} & -2v-h-\ln(v+h)+\ln(v-1)-\ln(v-1)-(h+2)\,\mathrm{dilog}\left(\frac{v+h}{h+1}\right) \\
& +\left[v+h+\frac{h+1}{1-v}\right]\ln\left(\frac{v+h}{h+1}\right)
\end{aligned}$$

$$
\begin{aligned}
&= -2v - h - \ln(v+h) + \left[v + h + \frac{h+1}{1-v}\right]\ln\left(\frac{v+h}{h+1}\right) \\
&\quad - (h+2)\operatorname{dilog}\left(\frac{v+h}{h+1}\right)\Bigg|_0^1 \\
&=^2 -2 - h - \ln(h+1) - 1 \\
&\quad - \left[-h - \ln(h) + (2h+1)\ln\left(\frac{h}{h+1}\right) - (h+2)\operatorname{dilog}\left(\frac{h}{h+1}\right)\right] \\
&= -3 - h - \ln(h+1) + h + \ln(h) - (2h+1)\ln\left(\frac{h}{h+1}\right) \\
&\quad -(h+2)\operatorname{dilog}\left(\frac{h}{h+1}\right) \\
&= -3 - (h+2)\operatorname{dilog}\left(\frac{h}{h+1}\right) + \ln\left(\frac{h}{h+1}\right) - (2h+1)\ln\left(\frac{h}{h+1}\right) \\
&= -3 - (h+2)\operatorname{dilog}\left(\frac{h}{h+1}\right) + 2h\ln\left(\frac{h+1}{h}\right). \qquad \text{(A.20)}
\end{aligned}
$$

Substituindo (A.20) em (A.19), (A.18) segue. □

A.2 Resultados Auxiliares à Seção 4.3

O objetivo principal desta seção é estabelecer a igualdade

$$
\begin{aligned}
&\iint_{I^2} \frac{u+v+h-\sqrt{(u+v+h)^2-4uv(h+1)}}{2}\,dudv = \\
&\qquad = \frac{1}{6}\left[h(h+1)\ln\left(\frac{h}{h+1}\right) + h + 2\right]. \qquad \text{(A.21)}
\end{aligned}
$$

O primeiro passo é calcular a integral (A.21) em relação à u. Para isto vamos usar as fórmulas 2.261 e 2.262-1 de Gradshteyn e Ryzhik (2000) (página 91), que podem ser resumidas no seguinte resultado[3]:

$$
\begin{aligned}
\int \sqrt{cx^2+bx+a}\;dx &= \frac{2cx+b}{4c}\sqrt{cx^2+bx+a} \\
&\quad + \frac{4ac-b^2}{8c^{\frac{3}{2}}}\ln\left(2\sqrt{c(cx^2+bx+a)} + 2cx + b\right). \qquad \text{(A.22)}
\end{aligned}
$$

[2]Pela Integral A.1.

[3]O resultado segue da simples substituição da fórmula 2.261 em 2.262-1 mas, atenção! Na fórmula 2.262-1 é feita referência a Δ, que os autores definem como $\Delta = 4ac - b^2$, o que é o contrário do usual.

Note que,

$$\int_0^1 \frac{u+v+h-\sqrt{(u+v+h)^2-4uv(h+1)}}{2}\,du =$$
$$= \frac{1}{2}\int_0^1 (u+v+h)\,du - \frac{1}{2}\int_0^1 \sqrt{(u+v+h)^2-4uv(h+1)}\,du. \quad \text{(A.23)}$$

Neste caso, vamos primeiramente calcular,

$$A := \int_0^1 \sqrt{(u+v+h)^2-4uv(h+1)}\,du. \quad \text{(A.24)}$$

Mas,

$$\begin{aligned}(u+v+h)^2-4uv(h+1) &= u^2+2u(v+h)+(v+h)^2-4uv(h+1)\\ &= u^2+u\,[\,2(v+h)-4v(h+1)]+(v+h)^2\\ &= u^2+u\,[\,2(h-v-2vh)]+(v+h)^2. \quad \text{(A.25)}\end{aligned}$$

Logo,

$$A = \int_0^1 \sqrt{u^2+u\left[\,2(h-v-2vh)\right]+(v+h)^2}.$$

Assim, substituindo, $a=(v+h)^2$, $b=2(h-v-2vh)$ e $c=1$ em (A.22), obtemos,

$$\begin{aligned}A &= \left(\frac{u+h-v-2vh}{2}\right)\sqrt{(u+v+h)^2-4uv(h+1)}\ \Bigg|_0^1\\ &\quad+\left(\frac{(v+h)^2-(h-v-2vh)^2}{2}\right)\times\\ &\quad\times\left[\ln\!\Big(\sqrt{(u+v+h)^2-4uv(h+1)}+u+h-v-2vh\Big)+\ln(2)\right]\Bigg|_0^1\\ &= \left(\frac{1+h-v-2vh}{2}\right)\sqrt{(1+v+h)^2-4v(h+1)}\\ &\quad-\frac{(h-v-2vh)\sqrt{(v+h)^2}}{2}+\left(\frac{(v+h)^2-(1+h-v-2vh)^2}{2}\right)\times\\ &\quad\times\ln\!\Big(\sqrt{(1+v+h)^2-4v(h+1)}+1+h-v-2vh\Big)\\ &\quad+\left(\frac{(h-v-2vh)^2-(v+h)^2}{2}\right)\ln\!\Big(\sqrt{(v+h)^2}+h-v-2vh\Big). \quad \text{(A.26)}\end{aligned}$$

Na expressão (A.26) calcularemos, primeiramente, a parte sem o logaritmo, isto é, vamos simplificar primeiro a expressão

$$B := \left(\frac{1+h-v-2vh}{2}\right)\sqrt{(1+v+h)^2-4v(h+1)} - \frac{(h-v-2vh)\sqrt{(v+h)^2}}{2}.$$

Note que,

$$\begin{aligned}(1+v+h)^2-4v(h+1) &= (1+h)^2+2v(1+h)+v^2-4hv-4v\\ &= 1+h^2+2h+2v+2hv+v^2-4hv-4v\\ &= h^2+v^2+2h-2v-2hv+1\\ &= v^2-2(1+h)v+(1+h)^2\\ &= (1+h-v)^2. \end{aligned}\tag{A.27}$$

Substituindo (A.27) em B, obtemos,

$$B=\left(\frac{1+h-v-2vh}{2}\right)(1+h-v)-\left(\frac{h-v-2vh}{2}\right)(v+h).\tag{A.28}$$

Mas,

$$\begin{aligned}&(1+h-v-2vh)(1+h-v)=\\ &\quad= 1+h-v+h+h^2-vh-v-vh+v^2-2vh-2vh^2+2v^2h\\ &\quad= v^2+h^2+1+2h-4hv-2v-2vh^2+2v^2h,\end{aligned}\tag{A.29}$$

e

$$\begin{aligned}(h-v-2vh)(v+h) &= vh+h^2-v^2-vh-2v^2h-2vh^2\\ &= h^2-v^2-2v^2h-2vh^2.\end{aligned}\tag{A.30}$$

Substituindo (A.29) e (A.30) em (A.28), obtemos,

$$\begin{aligned}B &= \frac{v^2+h^2+1+2h-4vh-2v-2vh^2+2v^2h-h^2+v^2+2v^2h+2vh^2}{2}\\ &= \frac{2v^2+2h-2v-4vh+4v^2h+1}{2}\\ &= v^2+h-v-2vh+2v^2h+\frac{1}{2}\cdot\end{aligned}\tag{A.31}$$

A parte da expressão (A.26) que contém o logaritmo pode ser escrita como

$$\begin{aligned}C := &-\left[\frac{(h-v-2vh)^2-(v+h)^2}{2}\right]\Big[\ln\big((h+1-v)+1+h-v-2vh\big)\\ &-\ln\big((v+h)+h-v-2vh\big)\Big]\\ = &-\left[\frac{(h-v-2vh)^2-(v+h)^2}{2}\right]\Big[\ln(2h-2v+2-2vh)-\ln\big(2h(1-v)\big)\Big]\\ = &\left[\frac{(h-v-2vh)^2-(v+h)^2}{2}\right]\Big[\ln\big(2h(1-v)\big)-\ln\big(2(h+1)(1-v)\big)\Big]\\ = &\left[\frac{(h-v-2vh)^2-(v+h)^2}{2}\right]\ln\left(\frac{h}{h+1}\right).\end{aligned}$$

Mas,

$$\begin{aligned}
(h-v-2vh)^2-(v+h)^2 &= (h-v)^2+4v^2h^2-4vh(h-v)-v^2-h^2-2vh \\
&= -4vh+4v^2h^2-4vh^2+4v^2h \\
&= 4vh(vh+v-h-1) \\
&= 4vh(h+1)(v-1).
\end{aligned}$$

Logo,

$$C = 2vh(h+1)(v-1)\ln\left(\frac{h}{h+1}\right). \tag{A.32}$$

Substituindo (A.31) e (A.32) na expressão para (A.26), obtemos,

$$A = B+C = v^2+h-v-2vh+2v^2h+\frac{1}{2}+2vh(h+1)(v-1)\ln\left(\frac{h}{h+1}\right). \tag{A.33}$$

Agora estamos em condições de calcular (A.23), que pode ser escrita como

$$\begin{aligned}
&\int_0^1 \frac{u+v+h-\sqrt{(u+v+h)^2-4uv(h+1)}}{2}\,du = \\
&= \int_0^1 \frac{u+v+h}{2}\,du-\frac{1}{2}A \\
&= \frac{1}{4}+\frac{v+h}{2} \\
&\quad -\frac{1}{2}\left[v^2+h-v-2vh+2v^2h+\frac{1}{2}+2vh(h+1)(v-1)\ln\left(\frac{h}{h+1}\right)\right] \\
&= v-\frac{v^2}{2}+vh-v^2h+vh(h+1)(1-v)\ln\left(\frac{h}{h+1}\right).
\end{aligned}$$

Voltando a equação (A.21), agora só precisamos calcular

$$\begin{aligned}
&\int_0^1 v-\frac{v^2}{2}+vh-v^2h+vh(h+1)(1-v)\ln\left(\frac{h}{h+1}\right)\,dv = \\
&= \frac{v^2}{2}-\frac{v^3}{6}+\frac{hv^2}{2}-\frac{hv^3}{3}+h(h+1)\ln\left(\frac{h}{h+1}\right)\left[\frac{v^2}{2}-\frac{v^3}{3}\right]\Bigg|_0^1 \\
&= \frac{h}{6}+\frac{1}{3}+\frac{1}{6}h(h+1)\ln\left(\frac{h}{h+1}\right) \\
&= \frac{1}{6}\left[h(h+1)\ln\left(\frac{h}{h+1}\right)+h+2\right].
\end{aligned}$$

Isto estabelece a igualdade (A.21). □