



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

SANTIAGO JIMENEZ RAMOS

Estimação de cópulas dinâmicas sob abordagem Bayesiana

Recife

2025

SANTIAGO JIMENEZ RAMOS

Estimação de cópulas dinâmicas sob abordagem Bayesiana

Dissertação apresentada ao programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Mestre em Estatística

Área de Concentração: Estatística Aplicada

Orientador: Francielle de Lima Medina

Coorientador: Aldo William Medina Garay

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Jimenez Ramos, Santiago.

Estimação de cópulas dinâmicas sob abordagem Bayesiana / Santiago Jimenez Ramos. - Recife, 2025.

76f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Ciências Exatas e da Natureza, Programa de Pós-Graduação em Estatística, 2025.

Orientação: Francielle de Lima Medina.

Coorientação: Aldo William Medina Garay.

Inclui referências.

1. Cópulas; 2. Polinômios; 3. Séries temporais; 4. Inferência Bayesiana. I. Medina, Francielle de Lima. II. Garay, Aldo William Medina. III. Título.

UFPE-Biblioteca Central

SANTIAGO JIMENEZ RAMOS

**ESTIMAÇÃO DE CÓPULAS DINÂMICAS SOB ABORDAGEM
BAYESIANA**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Estatística.

Aprovado em: 27 de fevereiro de 2025.

BANCA EXAMINADORA

Profa. Dra. Francielle de Lima Medina
Presidente (Orientadora)

Prof. Dr. Rodrigo Lambert
Examinador Externo

Prof. Dr. Roberto Ferreira Manghi
Examinador Interno

Dedico este trabalho aos meus pais, Ana Maria Ramos e José Andrés Jiménez, ao meu irmão, Mateo Jiménez, e à minha doce filha, Amelia. Obrigado por serem a base da minha vida e por me mostrarem que, independentemente das circunstâncias, sempre posso contar com vocês. A todos, meu mais sincero agradecimento.

AGRADECIMENTOS

Primeiramente, agradeço a Deus por me permitir chegar até aqui, proporcionando novas experiências, aprendizados, amizades e oportunidades. Sou grato por suas bênçãos, por colocar pessoas maravilhosas em meu caminho, mesmo em lugares desconhecidos, e por fortalecer os laços com aqueles que já faziam parte da minha vida. Sua orientação tem sido minha maior força ao longo desta experiência.

Aos meus pais, Ana María Ramos e José Andrés Jiménez, por me ensinarem o valor da perseverança e por me mostrarem a importância de não desistir diante das adversidades. Ao meu irmão mais velho, Mateo Jiménez, por ser uma inspiração e um grande exemplo para mim.

À minha filha, Amelia, que ilumina até os momentos mais difíceis e me dá a motivação necessária para alcançar os meus objetivos. Obrigado por ser minha maior fonte de inspiração.

Manifesto minha profunda gratidão à minha orientadora, Dra. Francielle de Lima Medina, por me oferecer a oportunidade de trabalhar neste projeto, confiar em minhas capacidades e compartilhar seu inestimável conhecimento e orientação ao longo do desenvolvimento deste trabalho. Da mesma forma, ao Dr. Aldo Medina Garay, meu coorientador, pelo apoio contínuo, experiência e sabedoria. Sem a orientação de ambos, a conclusão deste trabalho não teria sido possível.

Agradeço também a todas as pessoas que me acompanharam durante este processo. Em especial, aos meus amigos colombianos: Ana, sinônimo de paciência e perseverança; Jhon, um exemplo de disciplina; e José e Yess, que representam talento e alegria. Aos meus novos amigos no Brasil: Camila, Thaina, Diego e Luiz, "o Japa", agradeço por tornarem minha adaptação a esta nova etapa de vida muito mais leve. A companhia e o apoio de vocês foram inestimáveis.

Por fim, expresso minha gratidão à Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE), instituição que financiou integralmente meu mestrado, tornando possível a realização deste trabalho.

RESUMO

Este trabalho propõe cópulas dinâmicas para analisar a estrutura de dependência entre duas séries temporais. Supomos que o parâmetro da cópula pode ser caracterizado como uma função $\theta(t)$, que varia ao longo do tempo. Para aproximar essa função, são propostos diferentes polinômios, possibilitando a realização de inferências sobre um vetor de coeficientes, que descreve o comportamento dinâmico da função $\theta(t)$. Propomos uma abordagem Bayesiana e o algoritmo Metropolis-Hasting Adaptativo para estimar os coeficientes. Foram realizados estudos de simulação, além de uma aplicação prática, que indicaram a adequação da metodologia proposta.

Palavras-chave: Cópulas, Polinômios, Séries temporais, Inferência Bayesiana

ABSTRACT

This work proposes the use of dynamic copulas to analyze the dependence structure between two time series. It is assumed that the copula parameter can be characterized as a function $\theta(t)$, which varies over time. To approximate this function, different polynomials are proposed, allowing for inferences about a vector of coefficients that describe the dynamic behavior of $\theta(t)$. A Bayesian approach is proposed, along with the Adaptive Metropolis-Hastings algorithm to estimate the coefficients. Simulation studies were conducted, along with a practical application, which demonstrated the adequacy of the proposed methodology.

Keywords: Copulas, Polynomials, Time Series, Bayesian Inference.

LISTA DE FIGURAS

Figura 1 – Cópulas W (esquerda), Π (centro), M (direita).	26
Figura 2 – Densidade das cópulas Clayton, Gumbel, Frank e Joe com $\tau \approx 0.5$	31
Figura 3 – Contornos das densidades das cópulas Clayton, Gumbel, Frank e Joe com $\tau \approx 0.5$	32
Figura 4 – Função cópula e densidade gaussiana bivariada com parâmetro $\theta = 0.8$	33
Figura 5 – Curvas de nível e densidade da função cópula gaussiana bivariada com parâmetro $\theta = 0.8$	33
Figura 6 – Função cópula t e densidade bivariada com parâmetro $\theta = (2, 0.8)$	35
Figura 7 – Curvas de nível da função cópula e densidade t bivariada com parâmetro $\theta = (2, 0.8)$	35
Figura 8 – Viés, Viés Relativo e EQM para os coeficientes estimados com Cópula Gaussiana.	51
Figura 9 – Viés, Viés Relativo e EQM para os coeficientes estimados com Cópula Clayton.	52
Figura 10 – Viés, Viés Relativo e EQM para os coeficientes estimados com Cópula Gumbel.	53
Figura 11 – Aproximação da função seno com cópula Gaussiana	56
Figura 12 – Retornos de EUR/USD e BRL/USD.	58
Figura 13 – Log-retornos de EUR/USD	59
Figura 14 – Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) dos log-retornos de EUR/USD.	60
Figura 15 – Log-retornos ao quadrado de EUR/USD	60
Figura 16 – Funções de autocorrelação e autocorrelação parcial dos log-retornos ao quadrado de EUR/USD.	61
Figura 17 – Log-retornos de BRL/USD	62
Figura 18 – Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) dos log-retornos de BRL/USD.	63
Figura 19 – Log-retornos ao quadrado de BRL/USD	63
Figura 20 – Funções de autocorrelação e autocorrelação parcial dos log-retornos ao quadrado de BRL/USD.	64
Figura 21 – $\theta(t)$ ajustada para a Cópulas Frank	67
Figura 22 – VaR estimado com cópula Frank	70

LISTA DE TABELAS

Tabela 1 – Fórmulas para $\cos(n\theta)$ e polinômios de Chebyshev $T_n(x)$	18
Tabela 2 – Exemplo de cópulas arquimedianas bivariadas e seus geradores	30
Tabela 3 – Funções de transformação invertíveis para cópulas arquimedeanas	43
Tabela 4 – Métricas e Critérios de informação para a função $\theta(t)$ com cópula Gaussiana	55
Tabela 5 – Estatísticas para a série X_{1t}	59
Tabela 6 – Estatísticas para a série X_{2t}	62
Tabela 7 – Critérios AIC e BIC para diferentes cópulas e graus de polinômios utilizando o método de Bernstein	65
Tabela 8 – Critérios AIC e BIC para diferentes cópulas e graus de polinômios utilizando o método de Chebyshev	66
Tabela 9 – Critérios AIC e BIC para diferentes cópulas e graus de polinômios utilizando o método de Taylor	66
Tabela 10 – VaR utilizando Cópula Frank	69

SUMÁRIO

1	INTRODUÇÃO	12
2	PRELIMINARES	14
2.1	SÉRIES TEMPORAIS	14
2.1.1	Estacionariedade	15
2.2	POLINÔMIOS	16
2.2.1	Polinômios de Bernstein	16
2.2.2	Polinômios de Chebyshev	17
2.2.3	Polinômios de Taylor	19
2.3	CÓPULAS	20
2.3.1	Cópulas	20
2.3.2	Os Limites de Fréchet-Hoeffding e a Cópula Independente	24
2.3.3	Medidas de associação e sua conexão com cópulas	25
2.3.4	Famílias Paramétricas de funções Cópulas	29
2.3.5	Família de Cópulas Arquimedianas	29
2.3.6	Família de Cópulas Elípticas	32
3	CÓPULAS DINÂMICAS	36
3.1	CONCEITOS	36
3.1.1	Função de verossimilhança	39
3.1.2	Fundamentos da Inferência Bayesiana	40
3.1.3	Família de cópulas elípticas	42
3.1.4	Família de cópulas Arquimedeanas	43
4	ESTIMAÇÃO	45
4.1	ALGORITMO METROPOLIS-HASTINGS	45
4.2	ALGORITMO METROPOLIS-HASTINGS ADAPTATIVO	46
4.3	ESTIMATIVA PONTUAL DOS COEFICIENTES	48
5	ESTUDO DE SIMULAÇÃO	50
5.1	CENÁRIO 1: AVALIAÇÃO DO VIÉS E EQM	50
5.2	CENÁRIO 2: AVALIAÇÃO DE MÉTRICAS PARA $\theta(t)$:	53
6	APLICAÇÃO	57
6.1	ANÁLISE DAS SÉRIES	57

6.2	MEDIDA DE RISCO (VAR)	67
7	CONCLUSÕES	71
	REFERÊNCIAS	73

1 INTRODUÇÃO

O estudo da dependência entre fenômenos é um aspecto central em diversas áreas da ciência e aplicações práticas. Em finanças, por exemplo, a relação entre as cotações do dólar e do euro influencia decisões estratégicas; na administração pública, informações sobre enchentes, como volume e duração, são essenciais para ações preventivas e de planejamento. Modelar essas relações não apenas aprimora a compreensão dos fenômenos, mas também possibilita tomadas de decisão mais precisas.

A partir da década de 1950, as cópulas (SKLAR, 1959) surgiram como funções para modelar dependências entre variáveis aleatórias. A principal vantagem das cópulas está na capacidade de modelar a dependência entre variáveis aleatórias de maneira separada da escolha das distribuições marginais, permitindo modelar relações complexas entre variáveis que não seguem distribuições normais multivariadas (NELSEN, 2006). Além disso, as cópulas fornecem uma descrição completa da dependência entre as variáveis, permitindo avaliar o tipo de dependência em qualquer ponto do domínio conjunto, o que é particularmente útil para analisar caudas conjuntas ou eventos extremos. Isso proporciona maior flexibilidade e precisão, especialmente em situações onde os dados apresentam assimetria ou caudas pesadas, características que os modelos baseados na normalidade multivariada não conseguem capturar adequadamente. Essa técnica tem sido amplamente aplicada em áreas como finanças e seguros (DENUIT et al., 2006), saúde pública (DEB; TRIVEDI; ZIMMER, 2014; ZHAO; ZHOU, 2012), e hidrologia (SALVADORI; MICHELE, 2007).

No entanto, em muitos casos, as relações entre variáveis não permanecem estáticas ao longo do tempo. Fenômenos como flutuações de mercado ou impactos climáticos exigem modelos capazes de capturar dependências dinâmicas. As cópulas variantes no tempo foram desenvolvidas para atender a essa necessidade, permitindo que o parâmetro de dependência, $\theta(t)$, varie ao longo do tempo. Patton (2002) realizou uma revisão detalhada das técnicas de cópulas aplicadas à análise de séries temporais, enquanto Zhang e Guegan (2008) propuseram o uso de modelos GARCH combinados com cópulas variantes no tempo. Essas abordagens representam avanços importantes, mas há lacunas na literatura, especialmente em métodos semiparamétricos para cópulas dinâmicas.

No contexto de análise de risco financeiro, Ausin e Lopes (2010) introduziram uma abordagem totalmente Bayesiana para modelar dependências dinâmicas, incluindo a estimação de

medidas como o Value at Risk (VaR) e o Conditional Value at Risk (CVaR). Embora robusta, sua metodologia é predominantemente paramétrica, o que pode limitar sua flexibilidade em cenários mais complexos.

Para superar essas limitações, Abegaz, Gijbels e Veraverbeke (2012) propuseram uma abordagem semiparamétrica baseada na aproximação da dependência condicional por kernel. Inspirados nesse trabalho, este estudo desenvolve uma metodologia para modelar cópulas variantes no tempo, utilizando polinômios de Taylor, Bernstein e Chebyshev para aproximar a função de dependência $\theta(t)$, estimando os coeficientes sob enfoque Bayesiano e propondo o algoritmo de Metropolis-Hastings Adaptativo. A metodologia proposta é validada por meio de estudos de simulação e aplicações práticas, com destaque para o cálculo de medidas de risco financeiro, como o VaR.

Este texto está organizado da seguinte forma: no Capítulo 1 apresentamos a introdução e o contexto do trabalho. No Capítulo 2 revisamos os conceitos básicos de séries temporais, polinômios e cópulas. No Capítulo 3 apresentamos o conceito de cópulas variantes no tempo e discutimos a formulação da metodologia proposta. No Capítulo 4 detalhamos os métodos de estimação, incluindo o algoritmo de Metropolis-Hastings e sua versão adaptativa, além do uso de técnicas de Monte Carlo para aproximar distribuições a posteriori; no Capítulo 5 apresentamos o estudo de simulação e as aplicações práticas; e, finalmente, no Capítulo 6 apresentamos as considerações finais e sugestões para trabalhos futuros.

2 PRELIMINARES

Neste capítulo, apresentamos alguns conceitos fundamentais, que serão utilizados ao longo do texto.

2.1 SÉRIES TEMPORAIS

Nesta seção, apresentamos os conceitos fundamentais de séries temporais, com base na definição de processo estocástico e sua interpretação como sequência de observações ao longo do tempo. De acordo com Morettin e Tolo, (2018), um processo estocástico é definido como segue:

Definição 2.1.1. *Seja T um conjunto não vazio. Um processo estocástico é uma família $X = \{X(t), t \in T\}$, tal que, para cada $t \in T$, $X(t)$ é uma variável aleatória.*

Nessas condições, um processo estocástico é uma coleção de variáveis aleatórias (v.a.), definidas em um mesmo espaço de probabilidade $(\Omega, \mathcal{A}, \mathcal{P})$. O conjunto T , é normalmente escolhido como o conjunto dos inteiros $\mathbb{Z} = \{0, 1, 2, \dots\}$ ou o conjunto dos números reais \mathbb{R} . Para cada $t \in T$, $X(t)$ será uma variável aleatória real.

Como $X(t)$ representa uma variável aleatória para cada $t \in T$, o processo $X(t)$ pode ser visto como uma função de dois argumentos, $X(t, \omega)$, em que $t \in T$ e $\omega \in \Omega$. Essa notação ilustra que para cada valor de t , temos uma variável aleatória $X(t, \omega)$.

Para um valor fixo de $\omega_0 \in \Omega$, obtemos uma função em relação a t , que caracteriza uma trajetória específica do processo estocástico. Esta trajetória também é conhecida como uma série temporal e representa uma sequência de observações de $X(t)$ ao longo do tempo. Uma série temporal, portanto, é uma realização de um processo estocástico.

Em nossas considerações futuras, teremos para estudar uma série temporal $X(t)$ (uma realização de um processo estocástico), que denotaremos simplesmente como X_t . As observações serão feitas em instantes discretos e igualmente espaçados no tempo, denotados por X_1, X_2, \dots, X_T .

Para uma série temporal, definem-se as seguintes funções:

Definição 2.1.2. *A função média de X_t o simplesmente média é definida por*

$$\mu(t) = \mathbb{E}(X_t) = \int_{-\infty}^{\infty} x f_t(x) dx,$$

em que $\mathbb{E}(\cdot)$ denota o valor esperado de X_t , e $f_t(x)$ é a PDF de $X(t)$.

Definição 2.1.3. A função de autocovariância (facv) de X_t é

$$\gamma(t_1, t_2) = \mathbb{E}(X_{t_1}X_{t_2}) - \mathbb{E}(X_{t_1})\mathbb{E}(X_{t_2}),$$

$t_1, t_2 \in T$, respectivamente.

Em particular, se $t_1 = t_2 = t$, temos a variância do processo X_t :

$$\gamma(t, t) = \text{Var}(X_t) = V(t) = \mathbb{E}(X_t^2) - \mathbb{E}(X_t)^2.$$

2.1.1 Estacionariedade

Seja $X = \{X(t), t \in T\}$ um processo estocástico. Para quaisquer instantes $t_1, \dots, t_n \in T$, a distribuição conjunta das variáveis aleatórias $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ é descrita pela função de distribuição acumulada (CDF) conjunta:

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = \mathbb{P}(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n).$$

Esta notação será utilizada para descrever propriedades de estacionariedade.

Definição 2.1.4. Um processo estocástico $X = \{X_t, t \in T\}$ diz-se estritamente estacionário se todas as distribuições conjuntas permanecem as mesmas sob translações no tempo, ou seja,

$$F(x_1, \dots, x_n; t_1 + \tau, \dots, t_n + \tau) = F(x_1, \dots, x_n; t_1, \dots, t_n)$$

para quaisquer t_1, \dots, t_n, τ de T .

A função $F(\cdot)$ representa a função de distribuição acumulada conjunta das variáveis aleatórias $X_{t_1}, X_{t_2}, \dots, X_{t_n}$. Esta condição implica que a distribuição conjunta das variáveis é invariante sob translações no tempo, caracterizando assim um processo estritamente estacionário.

Isso significa, em particular, que todas as distribuições unidimensionais são invariantes sob translação no tempo, logo a média $\mu(t)$ e a variância $V(t)$ são constantes, isto é,

$$\mu(t) = \mu, \quad V(t) = \sigma^2, \quad \forall t \in T.$$

Definição 2.1.5. Um processo estocástico $X = \{X_t, t \in T\}$ é dito fracamente estacionário ou estacionário de segunda ordem se e somente se:

- i) $\mathbb{E}(X_t) = \mu(t) = \mu$, constante para todo $t \in T$.
- ii) $\mathbb{E}(X_t^2) < \infty$, para todo $t \in T$.
- iii) $\gamma(t_1, t_2) = \text{Cov}(X_{t_1}, X_{t_2}) = \gamma(h)$, $h(t_1, t_2) = |t_1 - t_2|$.

Em processos estacionários de segunda ordem, a função de autocovariância depende apenas da diferença de tempo entre duas observações, chamada de *defasagem temporal* ou “lag”. Denotamos essa defasagem por

$$h = |t_1 - t_2|,$$

simplificando assim a notação de $\gamma(t_1, t_2)$ para $\gamma(h)$.

Definição 2.1.6. A função de autocorrelação (ACF) de um processo estacionário é definida como:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)},$$

em que $\gamma(h)$ é a autocovariância no lag h e $\gamma(0)$ é a variância do processo.

2.2 POLINÔMIOS

Nesta seção, estão apresentados alguns polinômios que desempenham um papel fundamental em matemática, física, estatística e diversas outras áreas. Estes polinômios são frequentemente utilizados para aproximar funções complexas e desempenham um papel crucial em diversas aplicações práticas. Abordaremos três tipos de polinômios e discutiremos algumas de suas características e aplicações.

2.2.1 Polinômios de Bernstein

As bases de Bernstein de grau n , definidas no intervalo $[0, 1]$, são dadas por:

$$B_i^n(x) = \binom{n}{i} (1-x)^{n-i} x^i, \quad i = 0, 1, \dots, n, \quad x \in [0, 1], \quad (2.1)$$

em que

$$\binom{n}{i} = \frac{n!}{(n-i)!i!}.$$

Além disso, satisfazem a propriedade:

$$\sum_{i=0}^n B_{i,n}(x) = 1, \quad \forall x \in [0, 1]. \quad (2.2)$$

Definição 2.2.1. Se $f(x)$ é uma função definida no intervalo $[0, 1]$, então cada polinômio da forma

$$\mathcal{B}_n(x) = \sum_{i=0}^n f\left(\frac{i}{n}\right) B_i^n(x)$$

pode ser chamado de polinômio de Bernstein da função $f(x)$.

O seguinte teorema garante a aproximação de uma função $f(x)$ por polinômios de Bernstein.

Teorema 2.2.2 (Teorema de Bernstein). Se $f(x)$ é contínua em $[0, 1]$, então,

$$\lim_{n \rightarrow \infty} \mathcal{B}_n(x) = f(x).$$

Demonstração. Ver Natanson, (1961). □

Os polinômios de Bernstein são empregados em teoria da aproximação e geometria computacional. Esses polinômios foram aplicados para fornecer uma prova construtiva do teorema de aproximação de Weierstrass (BERNSTEIN, 1912). Com o avanço da computação digital, os polinômios de Bernstein ganharam popularidade devido à sua estabilidade numérica e à versatilidade em aplicações geométricas, permitindo a construção intuitiva e eficiente de formas geométricas complexas (FAROUKI, 2012).

2.2.2 Polinômios de Chebyshev

Existem vários tipos de polinômios de Chebyshev. Em particular, introduziremos os polinômios de primeiro tipo. Alguns livros e muitos artigos utilizam a expressão *polinômio de Chebyshev* para se referir exclusivamente ao polinômio de Chebyshev $T_n(x)$, de primeiro tipo.

Definição 2.2.3. Conforme definido em Natanson, (1961), um polinômio da forma

$$T_n(x) = \cos(n \arccos(x))$$

é chamado de polinômio de Chebyshev.

Note que, se $x \in [-1, 1]$, então, para $\theta = \arccos(x)$, temos $\theta \in [0, \pi]$. Dessa forma, podemos representar os polinômios de Chebyshev utilizando identidades trigonométricas. Do Teorema de De Moivre (MOIVRE, 1730), sabemos que $\cos(n\theta)$ pode ser expresso como um polinômio de grau n em $\cos(\theta)$.

A Tabela 1 apresenta as fórmulas elementares para $\cos(n\theta)$ e os polinômios de Chebyshev correspondentes.

Tabela 1 – Fórmulas para $\cos(n\theta)$ e polinômios de Chebyshev $T_n(x)$.

n	$\cos(n\theta)$	$T_n(x)$
0	1	1
1	$\cos(\theta)$	x
2	$2\cos^2(\theta) - 1$	$2x^2 - 1$
3	$4\cos^3(\theta) - 3\cos(\theta)$	$4x^3 - 3x$
4	$8\cos^4(\theta) - 8\cos^2(\theta) + 1$	$8x^4 - 8x^2 + 1$

Fonte: Elaborada pelo autor (2024)

Na prática, não é conveniente nem eficiente calcular cada polinômio $T_n(x)$ diretamente a partir da definição. Em vez disso, podemos utilizar uma relação de recorrência para gerar os polinômios de Chebyshev de forma mais rápida. Combinando a identidade trigonométrica

$$\cos(n\theta) = 2\cos(\theta)\cos((n-1)\theta) - \cos((n-2)\theta),$$

e aplicando a Definição 2.2.3, obtemos a seguinte relação de recorrência:

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x), \quad n = 2, 3, \dots \quad (2.3)$$

Com as condições iniciais

$$T_0(x) = 1 \quad \text{e} \quad T_1(x) = x,$$

podemos gerar recursivamente todos os polinômios $\{T_n(x)\}$ de maneira computacionalmente rápida e prática.

Em relação ao produto interno

$$\langle f, g \rangle = \int_{-1}^1 w(x)f(x)g(x)dx,$$

considerando o intervalo $[-1, 1]$, os polinômios de Chebyshev de primeiro tipo são ortogonais em relação à respectiva função de peso

$$w(x) = \frac{1}{\sqrt{1-x^2}}.$$

No caso específico de polinômios do primeiro tipo, a expansão é obtida como

$$f(x) \approx \sum_{i=0}^{\infty} c_i T_i(x),$$

em que

$$c_i = \frac{2}{\pi} \int_{-1}^1 (1-x^2)^{-\frac{1}{2}} f(x) T_i(x) dx.$$

2.2.3 Polinômios de Taylor

Os polinômios de Taylor desempenham um papel fundamental na aproximação de funções contínuas e diferenciáveis, pois permitem aproximar uma função ao redor de um ponto específico, por meio de uma série polinomial que incorpora informações sobre as derivadas da função nesse ponto. Mais formalmente, consideremos o conjunto das funções que são n -deriváveis e têm derivadas contínuas até a n -ésima ordem, denotado por

$$C^n = \{f : I \rightarrow \mathbb{R} : f^{(n)}(x) \text{ é contínua}\}.$$

Os polinômios de Taylor utilizam as derivadas de f para construir uma aproximação local para $f \in C^n$ ao redor de um ponto x_0 . Mais precisamente, quando temos uma função $f \in C^n$ e um polinômio $P_n(x)$ de grau n que satisfaça $P_k(x_0) = f^{(k)}(x_0)$, em que $f^{(k)}(x_0)$ denota a k -ésima derivada avaliada no ponto x_0 , para $k = 0, 1, 2, \dots, n$. Podemos afirmar que o polinômio de Taylor de ordem n de $f(x)$ em torno de x_0 terá a seguinte forma:

$$P_n(x) = f(x_0) + \frac{f^{(1)}(x_0)}{1!}(x - x_0) + \frac{f^{(2)}(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n,$$

e este é chamado de polinômio de Taylor de ordem n de $f(x)$ em torno de x_0 , pois fornece uma aproximação de $f(x)$ ao redor de x_0 , com erro que depende da ordem n e da natureza da função f .

Teorema 2.2.4 (Teorema de Taylor). *Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função $n + 1$ vezes diferenciável em torno do ponto $x_0 \in \mathbb{R}$, então, $\forall x \in \mathbb{R}, \exists \xi \in (a, x)$, tal que a função $f(x)$ pode ser aproximada por meio da série de Taylor:*

$$f(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{f^{(2)}(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + R_n(x),$$

em que

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$$

é definido como o resto de Lagrange para algum c entre x e x_0 .

Demonstração. Veja Bartle e Sherbert (1992) para mais detalhes. □

Note que

$$f(x) = P_n(x) + R_n(x),$$

Além disso, observe que se $|f^{(n+1)}(x)| \leq M, \forall x \in I$ e existe $r > 0$ tal que $I \subseteq (x_0 - r, x_0 + r)$, então

$$\begin{aligned} |R_n(x)| &= \left| f^{(n+1)}(c) \frac{(x - x_0)^{n+1}}{(n+1)!} \right| \\ &\leq \frac{M}{(n+1)!} |x - x_0|^{n+1} \\ &\leq M \frac{r^{n+1}}{(n+1)!}. \end{aligned}$$

Como

$$\lim_{n \rightarrow \infty} \frac{r^{n+1}}{(n+1)!} = 0,$$

segue-se que

$$\lim_{n \rightarrow \infty} |R_n(x)| = 0, \quad \forall x \in I.$$

Em outras palavras, para uma função com suporte compacto e derivadas contínuas de todas as ordens, o polinômio de Taylor $P_n(x)$ converge para $f(x)$ à medida que aumentamos a ordem n . É importante notar que podemos reescrever o polinômio de Taylor de ordem n na forma

$$P_n(x) = a_0 + a_1x + \dots + a_nx^n,$$

em que

$$a_i = \frac{f^{(i)}(x)}{i!}, \quad i = 0, 1, \dots, n.$$

2.3 CÓPULAS

Nesta seção, apresentamos os conceitos fundamentais das funções cópulas e as principais famílias de cópulas paramétricas. Em seguida, discutimos as medidas de dependência associadas a cada função cópula.

2.3.1 Cópulas

Sejam X_1, X_2, \dots, X_n variáveis aleatórias definidas em um mesmo espaço de probabilidade $(\Omega, \mathcal{A}, \mathcal{P})$ e considere a função de distribuição conjunta

$$F(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n). \quad (2.4)$$

Suponha que o vetor aleatório (X_1, \dots, X_n) seja transformado componente por componente para que cada variável aleatória no vetor resultante tenha uma distribuição marginal uniforme no intervalo $(0, 1)$. Se X_1, \dots, X_n têm distribuições marginais contínuas $F_{X_1}(x_1), \dots, F_{X_n}(x_n)$, isto pode ser feito utilizando a transformação:

$$T_n : \mathbb{R}^n \rightarrow (0, 1)^n, \quad (x_1, \dots, x_n) \mapsto (F_{X_1}(x_1), \dots, F_{X_n}(x_n)).$$

A função distribuição conjunta $C(\cdot)$ de $F_{X_1}(X_1), \dots, F_{X_n}(X_n)$, é então chamada de *função cópula* do vetor aleatório (X_1, \dots, X_n) , ou equivalentemente, cópula associada a função de distribuição conjunta $F(\cdot)$. Segue que,

$$C(u_1, \dots, u_n) = F(F_{X_1}^{-1}(u_1), \dots, F_{X_n}^{-1}(u_n)),$$

em que $u_i = F_{X_i}(x_i), i = 1, 2, \dots, n$. Para facilitar a leitura, apresentamos algumas definições e conceitos para o caso bivariado, no entanto, é importante ressaltar que os resultados e conclusões apresentados podem ser generalizados para o contexto multivariado.

Sejam X e Y variáveis aleatórias com funções de distribuições marginais $F_X(x)$ e $F_Y(y)$, respectivamente. Então a função de distribuição conjunta é dada por

$$F_{XY}(x, y) = P(X \leq x, Y \leq y).$$

Em geral, as funções de distribuições marginais podem ser obtidas utilizando a função de distribuição conjunta $F_{XY}(x, y)$ e as seguintes relações:

$$F_X(x) = \lim_{y \rightarrow \infty} F_{XY}(x, y), \quad F_Y(y) = \lim_{x \rightarrow \infty} F_{XY}(x, y).$$

Esta abordagem baseia-se na função de distribuição conjunta para determinar as funções de distribuição marginal de cada variável aleatória. A densidade conjunta $f_{XY}(x, y)$ é definida por

$$f_{XY}(x, y) = \frac{\partial^2}{\partial x \partial y} \{F_{XY}(x, y)\}.$$

As variáveis aleatórias X e Y são independentes se e somente se,

$$F_{XY}(x, y) = F_X(x)F_Y(y),$$

para todos os valores de $x, y \in \mathbb{R}$.

O Teorema de Sklar (SKLAR, 1959) é um marco fundamental na teoria das cópulas, pois estabelece a existência e unicidade da função cópula $C(\cdot)$. Este teorema é amplamente reconhecido como um dos resultados mais importantes nesta área, fornecendo uma estrutura teórica fundamental para a modelagem e análise de dependência estatística.

Teorema 2.3.1. *Seja $F_{XY}(x, y)$ uma função de distribuição conjunta com marginais $F_X(x)$ e $F_Y(y)$. Então existe uma função cópula $C(\cdot)$ tal que para todo $x, y \in \mathbb{R}$,*

$$F_{XY}(x, y) = C(F_X(x), F_Y(y)). \quad (2.5)$$

Se $F_X(x)$ e $F_Y(y)$ são contínuas, então a função cópula $C(\cdot)$ é única, caso contrário, $C(\cdot)$ está definida unicamente em $Im(F_X(x)) \times Im(F_Y(y))$ em que $Im(F_X(x))$, $Im(F_Y(y))$ representam os conjuntos de valores assumidos pelas funções de distribuição F_X y F_Y , respectivamente. Reciprocamente, se $C(\cdot)$ é uma função cópula e $F_X(x)$ e $F_Y(y)$ são funções de distribuições univariadas, então $F_{XY}(x, y)$ definida em (2.5) é a função de distribuição conjunta com marginais $F_X(x)$ e $F_Y(y)$.

O seguinte corolário afirma que é possível representar a dependência entre variáveis aleatórias por meio de uma função cópula que é independente das distribuições marginais específicas. Em outras palavras, a função cópula captura a estrutura de dependência, permitindo modelá-la separadamente das características das distribuições marginais.

Corolário 2.3.2. *Sejam $F_{XY}(x, y)$, $F_X(x)$ e $F_Y(y)$ como no Teorema de Sklar, e sejam $F_X^{-1}(x)$ e $F_Y^{-1}(y)$ as inversas generalizadas de $F_X(x)$ e $F_Y(y)$, respectivamente. Então, para qualquer $u_1 = F_X(x)$ e $u_2 = F_Y(y)$, existe uma cópula bivariada $C(\cdot)$ tal que*

$$C(u_1, u_2) = F_{XY}(F_X^{-1}(u_1), F_Y^{-1}(u_2)). \quad (2.6)$$

O Teorema 2.3.1 estabelece uma relação entre a cópula associada a uma função de distribuição bivariada $F_{XY}(x, y)$ e as distribuições marginais $F_X(x)$ e $F_Y(y)$, indicando que a estrutura de dependência entre as variáveis aleatórias X e Y é unicamente determinada pela função de distribuição bivariada $F_{XY}(x, y)$. Assim, o teorema mostra como a cópula está relacionada às distribuições marginais e à estrutura de dependência entre as variáveis aleatórias. Para o caso contínuo, a densidade de cópula é definida por

$$c(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} C(u_1, u_2).$$

Seja $F_{XY}(x, y)$ a função de distribuição conjunta com marginais $F_X(x)$ e $F_Y(y)$, conforme definida em (2.5). Utilizando a regra da cadeia, a densidade conjunta de X e Y pode ser obtida como

$$\begin{aligned} f_{XY}(x, y) &= \frac{\partial^2}{\partial x \partial y} \{F_{XY}(x, y)\} \\ &= \frac{\partial^2 C(F_X(x), F_Y(y))}{\partial F_X \partial F_Y} \cdot \frac{\partial F_X(x)}{\partial x} \cdot \frac{\partial F_Y(y)}{\partial y} \\ &= c(F_X(x), F_Y(y)) f_X(x) f_Y(y). \end{aligned}$$

O resultado acima mostra que sempre é possível obter a função de densidade conjunta de duas variáveis aleatórias, utilizando suas funções de densidades marginais e a função de densidade da função cópula associada. Isso é importante porque fornece uma maneira de expressar a função de densidade conjunta em termos das densidades marginais e da densidade de função cópula.

Proposição 2.3.3. *Sejam X e Y variáveis aleatórias contínuas com funções de distribuição marginais F_X , F_Y , densidades marginais f_X , f_Y , e função de densidade conjunta f_{XY} . Seja $C(u, v)$ a cópula associada ao par (X, Y) , com função de densidade $c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}$. Então, as seguintes afirmações são equivalentes:*

- (i) X e Y são independentes;
- (ii) A cópula associada é a cópula de independência: $C(u, v) = uv$;
- (iii) A densidade da cópula satisfaz $c(u, v) = 1$ para todo $(u, v) \in [0, 1]^2$;
- (iv) A função de densidade conjunta é escrita como $f_{XY}(x, y) = f_X(x)f_Y(y)$.

Demonstração.

(i) \Rightarrow (ii): Se X e Y são independentes, então:

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) = P(X \leq x) \cdot P(Y \leq y) = F_X(x)F_Y(y).$$

Logo, pela definição de cópula, temos:

$$C(u, v) = uv,$$

que é a cópula de independência.

(ii) \Rightarrow (iii): Se $C(u, v) = uv$, então:

$$\frac{\partial^2 C(u, v)}{\partial u \partial v} = \frac{\partial^2 (uv)}{\partial u \partial v} = 1.$$

Ou seja, $c(u, v) = 1$ em $[0, 1]^2$.

(iii) \Rightarrow (iv): A densidade conjunta pode ser escrita como:

$$f_{XY}(x, y) = c(F_X(x), F_Y(y)) \cdot f_X(x) \cdot f_Y(y).$$

Se $c(F_X(x), F_Y(y)) = 1$, então:

$$f_{XY}(x, y) = f_X(x) f_Y(y),$$

o que mostra que a função de densidade conjunta pode ser escrita como o produto das funções de densidade marginais.

(iv) \Rightarrow (i): Se $f_{XY}(x, y) = f_X(x) f_Y(y)$, então:

$$F_{XY}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(s, t) dt ds = \left(\int_{-\infty}^x f_X(s) ds \right) \left(\int_{-\infty}^y f_Y(t) dt \right) = F_X(x) F_Y(y),$$

mostrando que X e Y são independentes.

□

2.3.2 Os Limites de Fréchet-Hoeffding e a Cópula Independente

Para qualquer função cópula $C : [0, 1]^2 \rightarrow [0, 1]$ e $(u_1, u_2) \in [0, 1]^2$, existem limites inferiores e superiores que toda cópula deve satisfazer. Estes limites são conhecidos como os limites de Fréchet-Hoeffding e estabelecem o intervalo dentro do qual qualquer cópula deve estar:

$$W(u_1, u_2) \leq C(u_1, u_2) \leq M(u_1, u_2),$$

em que:

- A função $W(u_1, u_2) = \max\{u_1 + u_2 - 1, 0\}$ é chamada de limite inferior de Fréchet-Hoeffding. Ela representa o caso de dependência negativa máxima entre as variáveis.
- A função $M(u_1, u_2) = \min\{u_1, u_2\}$ é chamada de limite superior de Fréchet-Hoeffding. Ela representa o caso de dependência positiva máxima entre as variáveis.

Esses limites são essenciais em diversas áreas, especialmente na análise de riscos extremos, pois permitem modelar a dependência entre variáveis em eventos de cauda direita (associada ao limite superior $M(u_1, u_2)$) e cauda esquerda (associada ao limite inferior $W(u_1, u_2)$). Em contextos de risco, essas caudas representam eventos extremos em que as variáveis atingem valores altos ou baixos simultaneamente, fornecendo informações críticas para a gestão de riscos (RÜSCHENDORF, 2012).

Além dos limites de Fréchet-Hoeffding, outra função cópula importante é a função cópula produto ou cópula independente, definida como:

$$\Pi(u_1, u_2) = u_1 \cdot u_2.$$

A função cópula independente representa a ausência de qualquer dependência entre as variáveis aleatórias X e Y . É utilizada como uma referência para avaliar se a dependência entre variáveis aleatórias é mais forte ou mais fraca do que a independência completa. Em particular, ao comparar a dependência observada com a independência, é possível identificar se existe uma dependência significativa entre as variáveis, especialmente em eventos extremos, em que a dependência pode amplificar o impacto dos riscos (EMBRECHTS; LINDSKOG; MCNEIL, 2001).

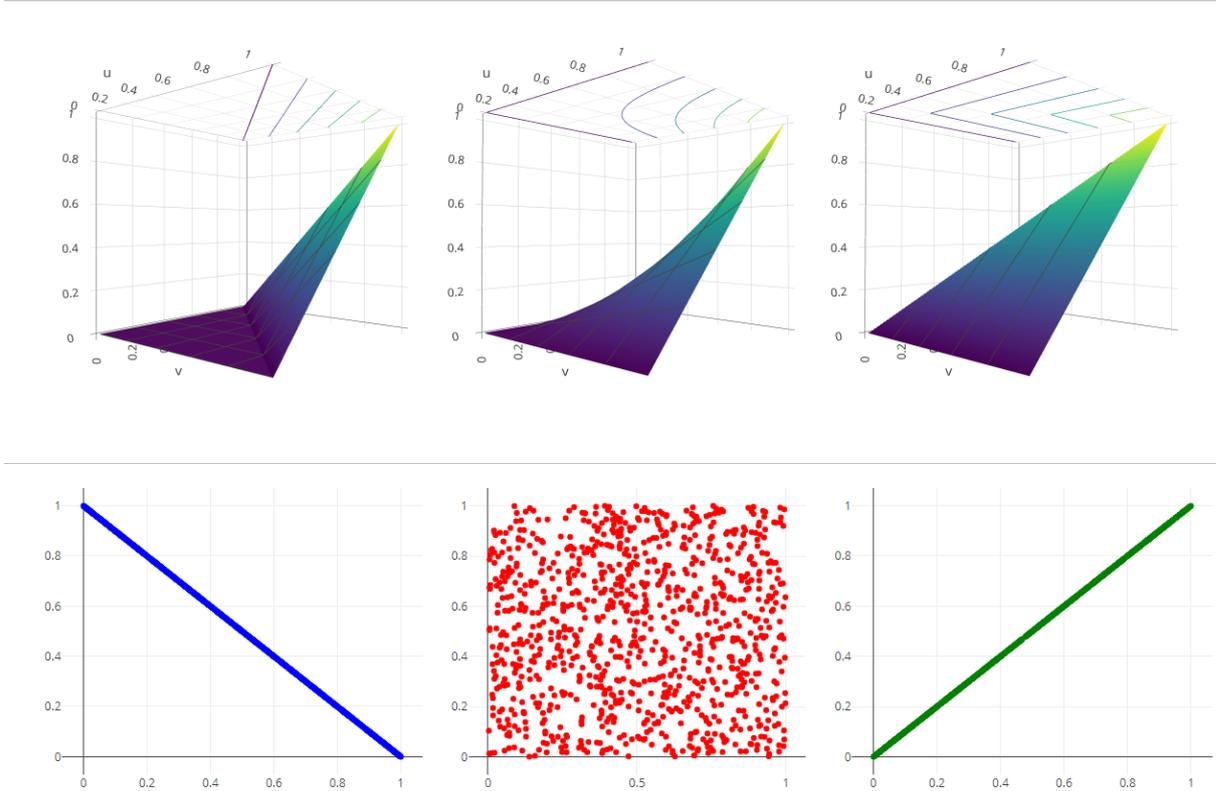
Assim, o conceito de cópulas independentes e os limites de Fréchet-Hoeffding fornecem uma base para entender a estrutura de dependência entre variáveis aleatórias, especialmente em contextos em que os eventos extremos desempenham um papel importante na avaliação e na gestão de riscos.

Na Figura 1, a primeira linha apresenta os gráficos das superfícies das cópulas W , Π e M , com suas respectivas curvas de nível. As curvas de nível conectam pontos com o mesmo valor da função, facilitando a visualização da estrutura de dependência. Na segunda linha, são mostrados os gráficos de dispersão correspondentes a W , Π e M , ilustrando diferentes padrões de dependência entre as variáveis.

2.3.3 Medidas de associação e sua conexão com cópulas

As medidas de associação para gerenciamento de riscos financeiros estão vastamente descritas na literatura, como pode ser visto em (EMBRECHTS; KLÜPPELBERG; MIKOSCH, 2013). Para duas variáveis aleatórias X e Y , (EMBRECHTS; LINDSKOG; MCNEIL, 2001), listaram quatro

Figura 1 – Cópulas W (esquerda), Π (centro), M (direita).



Fonte: Elaborado pelo autor (2024)

propriedades desejáveis de uma medida de associação $\delta(X, Y)$:

- I) $\delta(X, Y) = \delta(Y, X)$;
- II) $-1 \leq \delta(X, Y) \leq 1$;
- III) $\delta(X, Y) = 1$ se X e Y são cromotônicas e $\delta(X, Y) = -1$ se X e Y são contramonotônicas.
- IV) Se T for uma transformação monótona,

$$\delta(T(X), Y) = \begin{cases} \delta(X, Y), & \text{se } T \text{ crescente,} \\ -\delta(X, Y), & \text{se } T \text{ decrescente.} \end{cases}$$

Essas quatro propriedades são fundamentais para garantir que a medida de associação seja consistente, simétrica e fácil de interpretar. Elas asseguram que a medida capture adequadamente a relação entre as variáveis aleatórias, mantendo a mesma interpretação sob transformações monótonas. Além disso, o fato de a medida pertencer ao intervalo $[-1, 1]$ facilita a comparação direta da dependência entre diferentes conjuntos de dados (EMBRECHTS;

LINDSKOG; MCNEIL, 2001), (EMBRECHTS; HÖING; JURI, 2003), (EMBRECHTS; FREY; MCNEIL, 2011).

Dizemos que X e Y são comonotônicos se Y (ou X) for uma função estritamente crescente de X (ou Y) em quase toda parte e são contramonotônicos se a função for estritamente decrescente.

(EMBRECHTS; KLÜPPELBERG; MIKOSCH, 2013) argumentaram que a correlação linear de Pearson satisfaz apenas as duas primeiras propriedades e demonstraram que as medidas de correlação de Rho de Spearman e o Tau de Kendall satisfazem todas as quatro propriedades.

Embora as medidas apresentadas a seguir como o τ de Kendall e o ρ de Spearman sejam tecnicamente medidas de associação, elas podem ser reinterpretadas como medidas de dependência por meio da teoria das cópulas, conforme demonstrado por Nelsen (NELSEN, 2006). Esta seção discute essa transição de maneira gradual.

a) **Correlação linear de Pearson**

Sejam X e Y duas variáveis aleatórias definidas no mesmo espaço de probabilidade, o coeficiente de correlação de Pearson de X e Y é definido da seguinte forma:

$$\rho_P(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

em que $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ é a covariância entre X e Y , e σ_X, σ_Y denotam os desvios padrões de X e Y , respectivamente. Notamos que:

- i) $-1 \leq \rho_P \leq 1$;
- ii) Se X e Y são independentes, então $\rho_P = 0$;
- iii) $\rho_P = \pm 1$ se, e somente se, $Y = aX + b$ quase certamente, com $a \neq 0$, ou seja, X e Y estão linearmente relacionados com probabilidade 1.

Como vimos anteriormente, o coeficiente de correlação de Pearson satisfaz apenas as propriedades I) e II). Além disso, ρ_P requer que as variâncias de X e Y sejam finitas. Note que $\rho_P = 0$ não implica independência entre X e Y , a menos que (X, Y) siga uma distribuição normal bivariada. Observe também que $\rho_P = \pm 1$ ocorre apenas quando existe uma relação linear determinística entre as variáveis. Finalmente, ρ_P não é invariante sob transformações não lineares, o que limita seu uso em situações com relações não lineares entre X e Y .

b) τ de Kendall e ρ de Spearman

Suponha que as variáveis aleatórias X e Y representem os retornos de títulos financeiros. Em situações típicas, é comum que tanto X quanto Y apresentem variações semelhantes, ou seja, ambos assumem valores altos ou baixos simultaneamente. No entanto, em circunstâncias menos frequentes, é possível que X assume um valor alto enquanto Y assume um valor baixo, ou vice-versa, ao mesmo tempo. Os conceitos de concordância e discordância são utilizados para medir esses tipos de associação.

Suponha que a v.a. (X, Y) tenha distribuição F , contínua. Sejam, $(X_1, Y_1), (X_2, Y_2)$ dois pares independentes obtidos de F . Dizemos que estes pares são concordantes se $X_1 > X_2$ sempre que $Y_1 > Y_2$, $X_1 < X_2$ sempre que $Y_1 < Y_2$. Os pares são discordantes se $X_1 > X_2$ sempre que $Y_1 < Y_2$, $X_1 < X_2$ sempre que $Y_1 > Y_2$.

As medidas de concordância são invariantes sob transformações estritamente crescentes de X e Y , ou seja, não se alteram quando as variáveis são substituídas por funções que preservam a ordem relativa de seus valores. Essa propriedade garante que tais medidas sejam robustas e consistentes, independentemente da escala em que as variáveis são observadas. Como resultado, as medidas de concordância podem ser expressas como funções de cópula entre X e Y . A correlação linear de Pearson, por outro lado, não é invariante sob transformações crescentes, e portanto, não é uma medida de concordância. Duas medidas de concordância comuns são o *Rho de Spearman* e o *Tau de Kendall*.

τ de Kendall A estatística τ de Kendall para a distribuição F é definida como:

$$\tau = P\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - P\{(X_1 - X_2)(Y_1 - Y_2) < 0\}. \quad (2.7)$$

ρ de Spearman

Suponha que (X, Y) tenha distribuição conjunta F_{XY} , com marginais F_X e F_Y , respectivamente. O coeficiente ρ_S de Spearman também é baseado em concordância e discordância entre variáveis aleatórias. Seja $\{(X_i, Y_i)\}_{i=1}^3$ uma amostra de tamanho três de pares aleatórios independentes e identicamente distribuídos com a mesma distribuição de (X, Y) . Então,

$$\rho_S = 3 \left(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0] \right). \quad (2.8)$$

Essas medidas podem ser expressas em termos de cópulas; de fato, (NELSEN, 2006) prova os seguintes resultados.

Teorema 2.3.4. *Sejam X e Y variáveis aleatórias contínuas com função cópula $C(\cdot)$. Então o τ de Kendall para X e Y é definido por*

$$\tau = 4 \int_0^1 \int_0^1 C(u_1, u_2) c(u_1, u_2) du_1 du_2 - 1. \quad (2.9)$$

Teorema 2.3.5. *Sejam X e Y variáveis aleatórias contínuas com função cópula $C(\cdot)$. Então o ρ de Spearman para X e Y é definido por*

$$\begin{aligned} \rho_S &= 12 \int_0^1 \int_0^1 u_1 u_2 c(u_1, u_2) - 3 \\ &= 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3 \\ &= 12 \int_0^1 \int_0^1 [C(u_1, u_2) - u_1 u_2] du_1 du_2. \end{aligned}$$

Lembrando que $c(\cdot)$ é a função de densidade da função cópula $C(\cdot)$. Os teoremas 2.3.4 e 2.3.5 estabelecem que o τ de Kendall e o ρ de Spearman podem ser expressos em termos da função cópula, fornecendo uma maneira de calcular esta medida diretamente a partir da estrutura de dependência representada pela função cópula. Esses teoremas são importantes pois estabelecem uma conexão entre medidas de dependência comumente utilizadas e a teoria das funções cópulas, possibilitando compreender e calcular a dependência entre variáveis aleatórias de maneira mais direta.

2.3.4 Famílias Paramétricas de funções Cópulas

Nesta seção apresentamos algumas funções cópulas, que são frequentemente utilizadas na prática. Cada uma delas depende de um ou mais parâmetros, que denotaremos pelo vetor θ . Como antes, nos restringimos ao caso bivariado.

2.3.5 Família de Cópulas Arquimedianas

A família de funções cópulas arquimedianas é amplamente utilizada em diversas aplicações, devido à flexibilidade de suas propriedades de dependência, que podem ser ajustadas por meio de diferentes escolhas da função geradora, também chamada de função arquimediana. As cópulas arquimedianas são definidas a partir de uma função geradora φ , com as seguintes propriedades:

Definição 2.3.6. *Uma **função geradora** (ou simplesmente geradora) é uma função $\varphi : [0, \infty) \rightarrow [0, 1]$ que satisfaz:*

1. $\varphi(0) = 1$ e $\lim_{t \rightarrow \infty} \varphi(t) = 0$;
2. $\varphi(t)$ é contínua;
3. $\varphi(t)$ é decrescente em $[0, \infty)$ e estritamente decrescente em $[0, \inf\{t > 0 : \varphi(t) = 0\})$, com $\inf(\emptyset) = 0$.

Definição 2.3.7. *Cópulas Arquimedianas bivariadas são definidas por*

$$C_\varphi(u_1, u_2) = \varphi(\varphi^{-1}(u_1) + \varphi^{-1}(u_2)), \quad (2.10)$$

em que φ é uma função geradora e φ^{-1} sua inversa generalizada.

A utilidade das cópulas arquimedianas se deve à flexibilidade em ajustar diferentes estruturas de dependência por meio da escolha da função geradora $\varphi(\cdot)$. Além disso, possuem relativa simplicidade na estimação dos parâmetros. Diversas cópulas arquimedianas estão bem documentadas na literatura, e permitem representar padrões variados de associação entre variáveis. Alguns exemplos incluem as funções cópulas de Ali-Mikhail-Haq (AMH), Clayton, Frank, Gumbel-Hougaard e Joe, frequentemente utilizadas devido à sua capacidade de capturar diferentes padrões de dependência com uma forma funcional relativamente simples. O parâmetro $\theta \in \Theta$, apresentado na Tabela 2, indica o tipo de dependência entre as variáveis aleatórias consideradas.

Tabela 2 – Exemplo de cópulas arquimedianas bivariadas e seus geradores

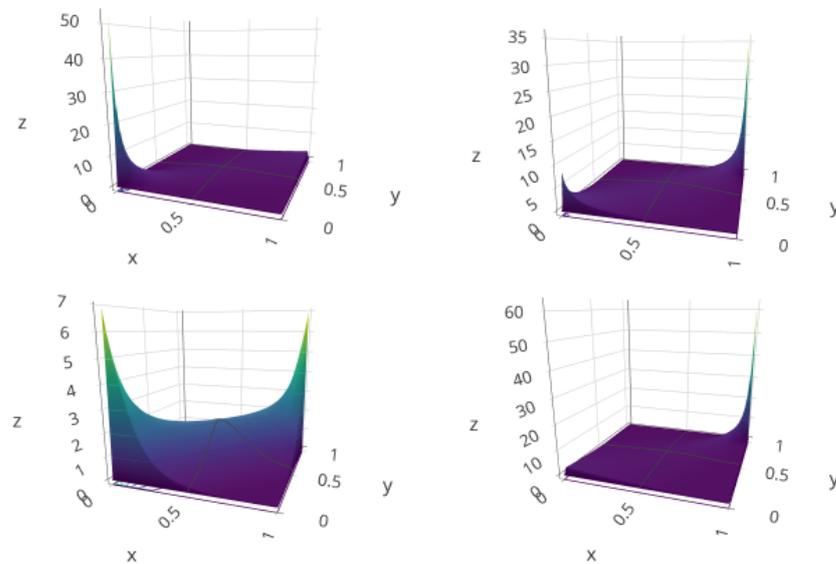
Cópula	$C(u_1, u_2)$	Θ	$\varphi(t)$
AMH	$\frac{u_1 u_2}{1 - \theta(1-u_1)(1-u_2)}$	$[-1, 1]$	$\log \left[\frac{1 - \theta(1-t)}{t} \right]$
Clayton	$(u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}$	$(0, \infty)$	$\frac{1}{\theta}(t^{-\theta} - 1)$
Frank	$-\frac{1}{\theta} \log \left[1 + \frac{(\exp(-\theta u_1) - 1)(\exp(-\theta u_2) - 1)}{\exp(-\theta) - 1} \right]$	$(0, \infty)$	$-\log \left(\frac{\exp(-\theta t) - 1}{\exp(-\theta) - 1} \right)$
Gumbel-Hougaard	$\exp \left[-((-\log(u_1))^\theta + (-\log(u_2))^\theta)^{\frac{1}{\theta}} \right]$	$[1, \infty)$	$(-\log(t))^\theta$
Joe	$1 - [(1 - u_1)^\theta + (1 - u_2)^\theta - (1 - u_1)^\theta(1 - u_2)^\theta]^{\frac{1}{\theta}}$	$[1, \infty)$	$-\log(1 - (1 - t)^\theta)$

Fonte: Elaborada pelo autor (2024)

Na Figura 2, são apresentados os gráficos das funções de densidade para as cópulas Clayton, Gumbel, Frank e Joe. Todas as funções cópulas foram parametrizadas de forma que o τ de Kendall seja próximo de $\frac{1}{2}$, utilizando a Equação (2.9). Apesar do mesmo valor aproximado de τ de Kendall, os gráficos destacam como diferentes estruturas de dependência podem ser capturadas por cada cópula, evidenciando padrões distintos de associação. Por fim, na Figura 3 são apresentados os contornos correspondentes às funções densidades das cópulas.

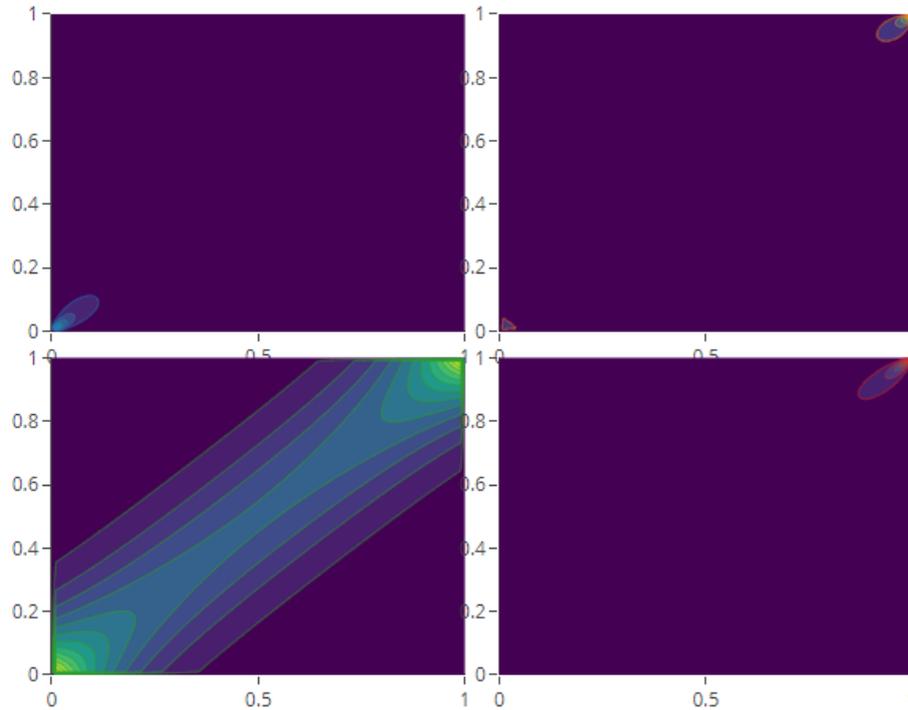
Neste exemplo, é possível observar que, embora as cópulas tenham o mesmo valor de τ de Kendall, elas representam a dependência de maneira distinta. Nota-se que, nas cópulas Clayton, Gumbel e Joe, há uma dependência positiva assimétrica, enquanto a cópula Frank exibe uma dependência positiva e simétrica, com valores maiores nos extremos.

Figura 2 – Densidade das cópulas Clayton, Gumbel, Frank e Joe com $\tau \approx 0.5$.



Fonte: Elaborado pelo autor (2024)

Figura 3 – Contornos das densidades das cópulas Clayton, Gumbel, Frank e Joe com $\tau \approx 0.5$.



Fonte: Elaborado pelo autor (2024)

2.3.6 Família de Cópulas Elípticas

Ao contrário das cópulas arquimedianas, geralmente caracterizadas por expressões analíticas claras que facilitam sua implementação prática, as cópulas elípticas são definidas de forma implícita e estão diretamente relacionadas às distribuições elípticas, como a normal multivariada e a *t* de Student (MAI; SCHERER, 2017).

Definição 2.3.8. *Seja (X, Y) um vetor aleatório bivariado com distribuição elíptica $F_R(x, y)$, matriz de correlação R , e funções de distribuição marginais F_X e F_Y . A cópula elíptica associada a (X, Y) , obtida a partir do Teorema de Sklar, é definida por:*

$$C_R(u_1, u_2) = F_R(F_X^{-1}(u_1), F_Y^{-1}(u_2)), \quad (2.11)$$

em que F_X^{-1} e F_Y^{-1} são as funções quantil de X e Y , respectivamente.

Na Definição 2.3.8, a matriz R é uma matriz de correlação 2×2 , em que os elementos da diagonal principal são iguais a 1 e os elementos fora da diagonal são iguais a θ . Este parâmetro θ caracteriza a dependência linear entre as variáveis da cópula. Assim, as cópulas

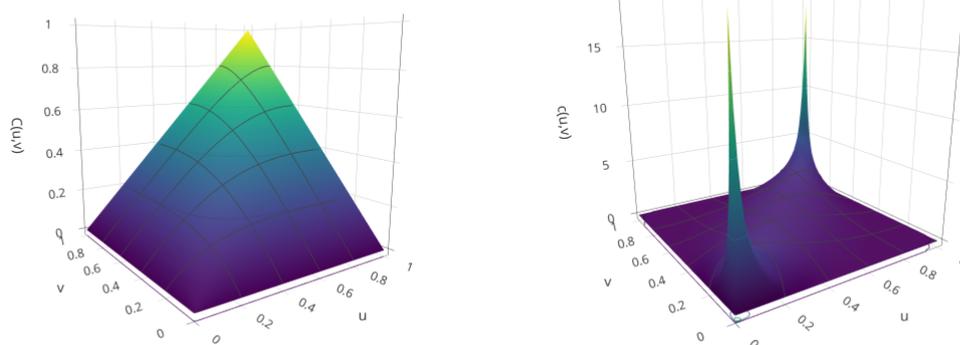
elípticas permitem descrever estruturas de dependência mais complexas em comparação a outros tipos de cópulas (HYRŠ; SCHWARZ, 2015).

O primeiro exemplo de função cópula elíptica é a cópula Gaussiana bivariada:

$$C_N(u_1, u_2) = \Phi_R(\Phi^{-1}(u_1), \Phi^{-1}(u_2)),$$

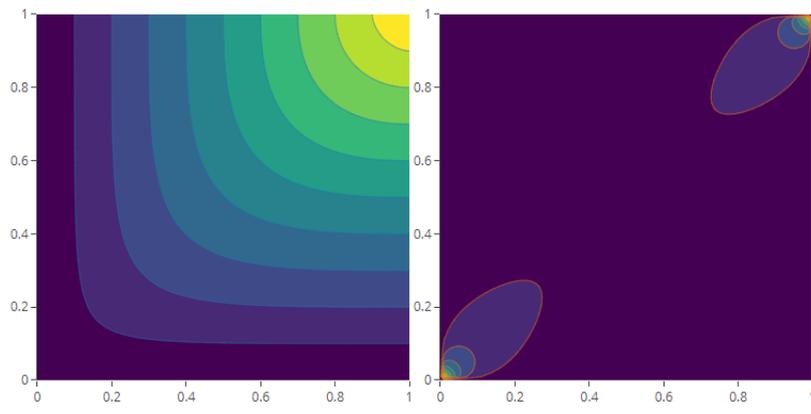
em que $\Phi_R(\cdot, \cdot)$ denota a função de distribuição acumulada (CDF) de uma distribuição normal padrão bivariada com coeficiente de correlação θ , e $\Phi^{-1}(\cdot)$ denota a função quantil de uma distribuição normal padrão univariada. Nas Figuras 4, 5 temos os gráficos para a cópula Gaussiana, com o coeficiente de dependência $\theta = \rho_P = 0.8$. Notamos uma dependência positiva e simétrica.

Figura 4 – Função cópula e densidade gaussiana bivariada com parâmetro $\theta = 0.8$.



Fonte: Elaborado pelo autor (2024)

Figura 5 – Curvas de nível e densidade da função cópula gaussiana bivariada com parâmetro $\theta = 0.8$.



Fonte: Elaborado pelo autor (2024)

O segundo exemplo é a cópula t bivariada, associada à família da distribuição t de Student:

$$C_t(u_1, u_2) = t_{\Sigma, \nu} \left(t_{\nu}^{-1}(u_1), t_{\nu}^{-1}(u_2) \right), \quad (2.12)$$

em que:

1. $t_{\Sigma, \nu}$ é a função de distribuição acumulada (CDF) de uma distribuição t de Student bivariada com:
 - $\nu > 2$: Graus de liberdade, garantindo momentos finitos.
 - Σ : Matriz de correlação 2×2 , com elementos na diagonal principal iguais a 1 e elementos na diagonal secundária a ρ_P .
2. t_{ν}^{-1} denota a função quantil (CDF inversa) de uma distribuição t de Student univariada com ν graus de liberdade.

Assim, os parâmetros da cópula t são especificados por:

$$\theta = (\nu, \rho_P),$$

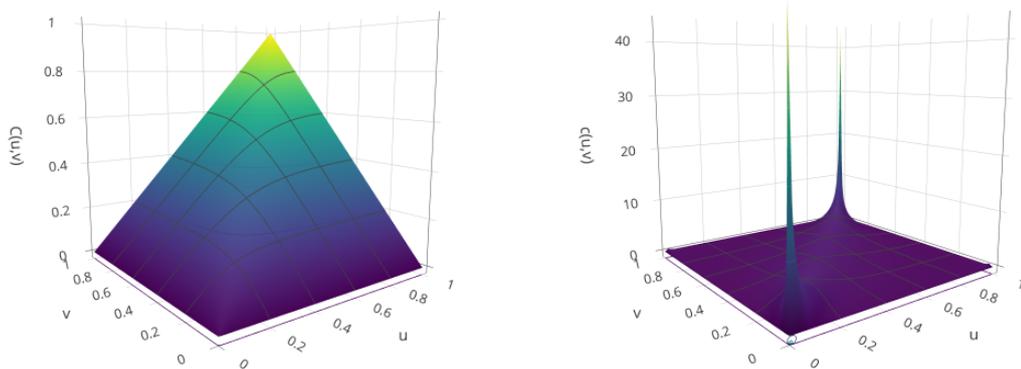
em que:

- $\nu > 2$: são os graus de liberdade.
- $\rho_P \in (-1, 1)$: Representa a correlação linear entre as variáveis.

As figuras, 6, 7, mostram a cópula t , com $\rho_P(0.8)$ e número de graus de liberdade $\nu = 2$. Obtemos gráficos semelhantes ao caso anterior (cópula Gaussiana bivariada com $\theta = 0.8$) mas com maior presença de valores extremos nos limites inferior e superior.

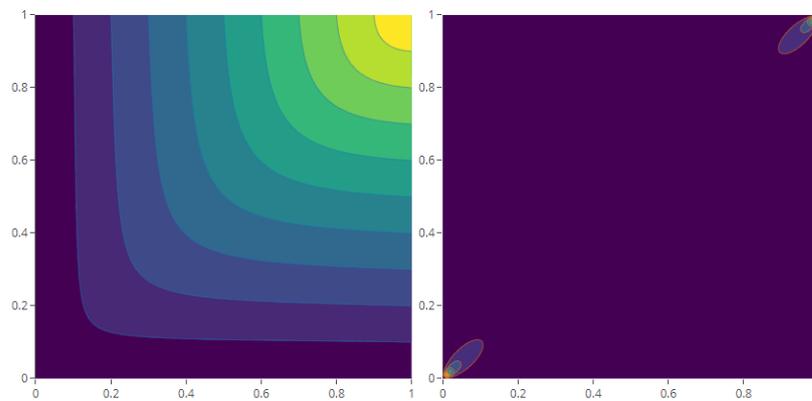
Em geral, as cópulas elípticas são amplamente utilizadas na estatística multivariada e na análise de dados. Sua capacidade de descrever uma ampla variedade de estruturas de dependência as torna em funções importantes na avaliação de riscos, finanças e em diversas outras áreas (MAI; SCHERER, 2014).

Figura 6 – Função cópula t e densidade bivariada com parâmetro $\theta = (2, 0.8)$.



Fonte: Elaborado pelo autor (2024)

Figura 7 – Curvas de nível da função cópula e densidade t bivariada com parâmetro $\theta = (2, 0.8)$.



Fonte: Elaborado pelo autor (2024)

3 CÓPULAS DINÂMICAS

Cópulas dinâmicas são amplamente utilizadas para analisar a dependência entre ativos financeiros Cherubini et al. (2011), Braun, (2011), Jammazi et al, (2015) . Estas cópulas permitem modelar de forma mais precisa as relações entre variáveis ao capturar a variação temporal das dependências, algo muito importante para entender e prever comportamentos em mercados financeiros voláteis. Como discutido no capítulo anterior, em muitos casos, as cópulas tradicionais utilizam um único parâmetro, como o coeficiente de correlação de Pearson no caso da cópula gaussiana, para caracterizar a função de dependência ao longo de todo o período de avaliação. No entanto, ao permitir que esse parâmetro varie dinamicamente com o tempo, as cópulas dinâmicas oferecem uma abordagem mais adaptativa e sensível às mudanças temporais, o que é fundamental para modelar com precisão as relações de dependência em séries temporais (PATTON, 2006a). Com essa abordagem, o parâmetro θ utilizado na função de dependência pode ser funcional, ou seja, pode variar de acordo com o tempo, permitindo uma representação mais precisa e flexível da evolução das relações entre os dados ao longo do tempo.

Neste contexto, esta pesquisa tem como objetivo propor que o parâmetro da cópula pode variar ao longo do tempo, sendo aproximado por funções não paramétricas, como polinômios de Taylor, Bernstein e Chebyshev. Consideramos que o parâmetro de uma cópula dinâmica é uma função dependente do tempo, e essa função será aproximada utilizando polinômios. Essa abordagem nos permitirá capturar com maior precisão as relações de dependência em séries temporais, proporcionando uma modelagem mais robusta e flexível (KRUPSKII; JOE, 2020).

3.1 CONCEITOS

Considere uma série temporal bivariada $\{\mathbf{X}_t\}_{t \in T}$, em que $\mathbf{X}_t = (X_{1t}, X_{2t})$, definida em um espaço de probabilidade (Ω, \mathcal{F}, P) , em que Ω é o espaço amostral, \mathcal{F} é uma σ -álgebra definida em Ω , e P é uma medida de probabilidade sobre (Ω, \mathcal{F}) ¹.

A σ -álgebra \mathcal{F}_{t-1} é gerada pelas variáveis aleatórias X_{1s} e X_{2s} para $1 \leq s \leq t-1$, definida por:

$$\mathcal{F}_{t-1} = \sigma(X_{1s}, X_{2s} | 1 \leq s \leq t-1).$$

¹ Para simplificar a notação, adotamos as abreviações $\mathbf{X}_t = (X_{1t}, X_{2t})$ e $\mathbf{x}_t = (x_{1t}, x_{2t})$.

Esta σ -álgebra contém toda a informação disponível até o instante $t - 1$, permitindo modelar as distribuições marginais condicionais $X_{1t}|\mathcal{F}_{t-1}$ e $X_{2t}|\mathcal{F}_{t-1}$ (PATTON, 2013). Note que, quando temos a série temporal bivariada \mathbf{X}_t , a amostra $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_T$ não é composta por pares independentes, pois existe correlação serial dentro de cada série e correlação cruzada entre as séries, de modo que os pares $\mathbf{X}_r, \mathbf{X}_s$ não são independentes para $r \neq s$.

Para contornar este problema, um procedimento comum consiste em ajustar modelos ARMA-GARCH individualmente a cada série e obter os resíduos padronizados (MORETTIN, 2017). Suponha que α denote o vetor de parâmetros dos modelos marginais ARMA-GARCH. A estimativa obtida por máxima verossimilhança será denotada por $\hat{\alpha}$. Com base nisso, os resíduos padronizados são definidos como:

$$\epsilon_{it} = \frac{X_{it} - \hat{\mu}_{it}}{\hat{\sigma}_{it}}, \quad i = 1, 2, \quad t = 1, 2, \dots, T,$$

em que:

- $\hat{\mu}_{it}$ é a média condicional estimada com base em \mathcal{F}_{t-1} por um modelo ARMA ajustado com os parâmetros $\hat{\alpha}$;
- $\hat{\sigma}_{it}$ é o desvio padrão condicional estimado por um modelo GARCH, também com base em $\hat{\alpha}$.

Os parâmetros $\hat{\alpha}$ representam os coeficientes estimados dos modelos ARMA-GARCH ajustados individualmente para as séries X_{1t} e X_{2t} . Após o ajuste dos modelos marginais e a obtenção dos resíduos padronizados, as distribuições acumuladas marginais estimadas são aproximadas utilizando funções ajustadas com base nesses resíduos. Assim, temos:

$$\hat{F}_{1t}(x_{1t}) \approx \int_{-\infty}^{x_{1t}} \hat{f}_{1t}(z) dz, \quad \hat{F}_{2t}(x_{2t}) \approx \int_{-\infty}^{x_{2t}} \hat{f}_{2t}(z) dz,$$

em que $\hat{f}_{1t}(x_{1t})$ e $\hat{f}_{2t}(x_{2t})$ representam as funções de densidade ajustadas aos resíduos padronizados ϵ_{it} . Essas funções podem ser obtidas empiricamente ou a partir de uma distribuição teórica apropriada, como a normal padrão ou outra que melhor se ajuste ao comportamento observado. Sejam

$$\hat{u}_1 = \hat{F}_{1t}(x_{1t}), \quad \hat{u}_2 = \hat{F}_{2t}(x_{2t}),$$

as pseudo-observações associadas às variáveis X_{1t} e X_{2t} , transformadas para o intervalo $[0, 1]$, então, a função cópula condicional paramétrica $C(\hat{u}_1, \hat{u}_2|\theta(t))$ modela a dependência entre as

variáveis X_{1t} e X_{2t} , em que o parâmetro da cópula $\theta(t)$ é uma função do tempo t . A função de distribuição acumulada conjunta de $\mathbf{X}_t|\mathcal{F}_{t-1}$ pode ser expressa como:

$$\begin{aligned} F(x_{1t}, x_{2t}|\mathcal{F}_{t-1}) &= C_t(F_{1t}(x_{1t}), F_{2t}(x_{2t})) \\ &= C(F_{1t}(x_{1t}), F_{2t}(x_{2t})|\theta(t)), \end{aligned}$$

em que $C_t(\cdot, \cdot)$ é a função cópula condicionada em \mathcal{F}_{t-1} . Por outro lado, a função de densidade de $\mathbf{X}_t|\mathcal{F}_{t-1}$ é dada por:

$$\begin{aligned} f(x_{1t}, x_{2t}|\mathcal{F}_{t-1}) &= c_t(F_{1t}(x_{1t}), F_{2t}(x_{2t}))f_{1t}(x_{1t})f_{2t}(x_{2t}) \\ &= c(F_{1t}(x_{1t}), F_{2t}(x_{2t})|\theta(t))f_{1t}(x_{1t})f_{2t}(x_{2t}), \end{aligned} \quad (3.0)$$

em que $c_t(\cdot, \cdot)$ é a função de densidade da cópula condicionada a \mathcal{F}_{t-1} ; $F_{it}(\cdot)$ e $f_{it}(\cdot)$ são, respectivamente, as funções de distribuição acumulada (f.d.a.) condicionais e as funções de densidade de probabilidade (f.d.p.) condicionais de $X_{it}|\mathcal{F}_{t-1}$ para $i = 1, 2$; $C(\cdot, \cdot)$ é a função cópula e $c(\cdot, \cdot)$ é a respectiva função de densidade da cópula no instante t para $\mathbf{X}_t|\mathcal{F}_{t-1}$.

Seja $\theta(t)$ uma função definida sobre um domínio \mathcal{D} , em que \mathcal{D} representa o conjunto de valores possíveis para $\theta(t)$. Em vez de interpretarmos $\theta(t)$ como um parâmetro fixo pertencente a um espaço paramétrico, consideramos $\theta(t)$ como uma função que varia ao longo do tempo e que descreve a dependência dinâmica entre as variáveis de interesse. O domínio \mathcal{D} pode ser restrito de acordo com a cópula utilizada; por exemplo:

- Para uma cópula Gaussiana, $\mathcal{D} = [-1, 1]$,
- Para a cópula de Clayton, $\mathcal{D} = (0, \infty)$.

Embora $\theta(t)$ seja uma função, é comum que o domínio \mathcal{D} não corresponda ao conjunto de todos os números reais \mathbb{R} . Para facilitar a estimação e a aplicação de técnicas numéricas, utilizamos uma função de transformação invertível $\psi : \mathcal{D} \rightarrow \mathbb{R}$, cujas propriedades são baseadas nos princípios gerais de análise funcional e otimização descritos por Lang, (2012) e Boyd e Vandenberghe, (2004):

1. **Invertibilidade:** Existe uma função inversa $\psi^{-1} : \mathbb{R} \rightarrow \mathcal{D}$ tal que:

$$\psi^{-1}\{\psi(\theta(t))\} = \theta(t), \quad \forall \theta(t) \in \mathcal{D}.$$

Isso garante que cada valor transformado em \mathbb{R} possa ser revertido ao seu valor original no domínio \mathcal{D} .

2. **Suavidade:** As funções ψ e ψ^{-1} devem ser continuamente diferenciáveis. Essa propriedade é essencial para que a transformação e sua inversa possam ser aplicadas em métodos de otimização e estimação que envolvem derivadas contínuas, além de facilitar a implementação numérica.
3. **Escalabilidade:** A transformação ψ deve mapear o domínio \mathcal{D} para um intervalo adequado em \mathbb{R} . Isso assegura que os valores de $\psi(\theta(t))$ estejam dentro de uma escala apropriada.

A transformação $\psi : \mathcal{D} \rightarrow \mathbb{R}$, descrita anteriormente, permite que o parâmetro $\theta(t)$ seja estimado em um espaço mais amplo, mesmo que originalmente seja restrito a um subconjunto dos números reais. Definimos a transformação $\psi\{\theta(t)\} = h(v)$, em que v é uma normalização de t : $v = \frac{t-1}{T}$, $t = 1, 2, \dots, T$.

A função $h(v)$, para $v \in [0, 1]$, pode ser aproximada da seguinte forma:

$$h(v) = \sum_{i=0}^n a_i \varphi_i(v),$$

em que $\varphi_i(v)$ pode ser algum tipo de polinômio base e a_i com $i = 0, 1, \dots, n$ são os coeficientes de aproximação que determinam a forma exata de $h(v)$. Seja $\mathbf{a} = (a_0, a_1, \dots, a_n)$ o vetor dos coeficientes desconhecidos. Esses coeficientes determinam a forma da função $\theta(t)$ por meio da aproximação polinomial definida por $h(v)$, em que v é uma reescalagem do tempo t . Portanto, o desenvolvimento da metodologia para determinar o valor de \mathbf{a} é descrito nas seções a seguir.

3.1.1 Função de verossimilhança

Seja $\mathbf{X}_t = (X_{1t}, X_{2t})$ uma série temporal bivariada com $t = 1, 2, \dots, T$, e seja \mathbf{x}_t a realização observada de \mathbf{X}_t . Suponha que $\mathbf{a} = (a_0, a_1, \dots, a_n)$ denota o vetor de parâmetros do modelo. , então $\forall t = 1, 2, \dots, T$, a função de verossimilhança é dada por

$$L(\mathbf{a} | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) = f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T | \mathbf{a}), \quad (3.1)$$

em que f é a função de densidade conjunta de $\mathbf{X}_1, \dots, \mathbf{X}_T$. A função de verossimilhança apresentada em (3.1) pode ser reformulada com base na decomposição da densidade conjunta

em distribuições condicionais, como segue

$$\begin{aligned} L(\mathbf{a}) &= f(\mathbf{x}_T|\mathbf{x}_1, \dots, \mathbf{x}_{T-1}) \cdot f(\mathbf{x}_{T-1}|\mathbf{x}_1, \dots, \mathbf{x}_{T-2}) \cdots f(\mathbf{x}_2|\mathbf{x}_1) \cdot f(\mathbf{x}_1) \\ &= f(\mathbf{x}_T|\mathcal{F}_{T-1}) \cdot f(\mathbf{x}_{T-1}|\mathcal{F}_{T-2}) \cdots f(\mathbf{x}_2|\mathcal{F}_1) \cdot f(\mathbf{x}_1) \\ &= f(\mathbf{x}_1) \prod_{t=2}^T f(\mathbf{x}_t), \end{aligned}$$

e pelo resultado 3.0 temos que

$$L(\mathbf{a}) = f(\mathbf{x}_1) \prod_{t=2}^T c(F_{1t}(x_{1t}), F_{2t}(x_{2t})|\theta(t)) f_{1t}(x_{1t}) f_{2t}(x_{2t}), \quad (3.2)$$

em que $f(\mathbf{x}_1)$ é a função de densidade marginal de $\mathbf{x}_1 = (x_{1t}, x_{2t})$. Note que, $f_{1t}(x_{1t})$, $f_{2t}(x_{2t})$ e $f(\mathbf{x}_1)$ são funções que não dependem de \mathbf{a} e portanto podemos simplificar (3.2) por

$$L(\mathbf{a}) \propto \prod_{t=2}^T c(F_{1t}(x_{1t}), F_{2t}(x_{2t})|\theta(t)).$$

3.1.2 Fundamentos da Inferência Bayesiana

A inferência bayesiana é um método estatístico que utiliza distribuições de probabilidade para representar a incerteza sobre parâmetros desconhecidos. Antes de observar os dados, essa incerteza é descrita por uma distribuição a priori. Depois da observação, essa distribuição é atualizada por meio da regra de Bayes, resultando em uma distribuição a posteriori, que reflete o que se sabe sobre os parâmetros após incorporar a informação dos dados.

Esse método não se limita a testes de hipótese. Ele é utilizado para estimar parâmetros, construir intervalos de credibilidade, fazer previsões e tomar decisões com base em distribuições completas.

De acordo com Hoff (2009), a regra de Bayes oferece uma maneira sistemática de atualizar a informação sobre parâmetros após a observação de novos dados. Resultados anteriores de Cox (1946, 2001) e Savage (1954, 1972) mostram que, sob certas condições, essa regra é a forma mais consistente de atualizar probabilidades.

Neste trabalho, utilizamos inferência bayesiana para estimar os coeficientes \mathbf{a} que determinam o comportamento da função $\theta(t)$, que descreve a estrutura dinâmica da série temporal bivariada.

Distribuição a Priori

Para modelar a incerteza inicial sobre os coeficientes \mathbf{a} , assumimos uma distribuição *a priori* fracamente informativa:

$$\mathbf{a} \sim N_d(\mathbf{0}, \sigma^2 \mathbf{I}),$$

em que em que \mathbf{a} é o tamanho do vetor, \mathbf{I} é uma matriz identidade de dimensão $d \times d$, e σ^2 é um escalar positivo conhecido que controla a variabilidade inicial assumida para os coeficientes. Essa escolha reflete uma crença inicial de que os coeficientes estão centrados em zero, com variação limitada, antes de observar os dados. Além disso, prioris normais multivariadas são amplamente utilizadas em análises bayesianas quando há pouca informação prévia disponível, como sugerido por Hoff (2009).

Distribuição a posteriori

A distribuição *a posteriori* para \mathbf{a} é aproximada por:

$$\pi(\mathbf{a} | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \propto L(\mathbf{a}) \cdot \exp\left(-\frac{1}{2\sigma^2} \mathbf{a}^\top \mathbf{a}\right),$$

substituindo $L(\mathbf{a})$ por sua expressão na equação (3.2), obtemos:

$$\pi(\mathbf{a} | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \propto \exp\left(-\frac{1}{2\sigma^2} \mathbf{a}^\top \mathbf{a}\right) \prod_{t=2}^T c\left(F_{1t}(x_{1t}), F_{2t}(x_{2t}) \mid \psi^{-1}\left(h\left(\frac{t-1}{T}\right)\right)\right). \quad (3.4)$$

Propriedade Assintótica da Posteriori

Uma propriedade fundamental da inferência Bayesiana é o *comportamento assintótico da distribuição posteriori*. De acordo com o Teorema de Bernstein-von Mises, sob certas condições, a distribuição posteriori tende, assintoticamente, a uma distribuição normal (gaussiana), independentemente da escolha da priori inicial, desde que o número de observações aumente indefinidamente (FREEDMAN, 1965).

Especificamente, para amostras suficientemente grandes e priors bem comportadas, a posteriori converge para uma distribuição normal centrada no valor verdadeiro do parâmetro, com variância decrescente proporcional ao tamanho da amostra. Este resultado é essencial para garantir:

1. Consistência Bayesiana: A posteriori converge em probabilidade para o valor verdadeiro do parâmetro $\theta(t)$ à medida que o número de observações $T \rightarrow \infty$.
2. Redução da influência da priori: à medida que o número de observações aumenta, a distribuição a posteriori passa a depender quase exclusivamente da verossimilhança, tornando as estimativas cada vez menos sensíveis à escolha da distribuição a priori.

No presente trabalho, esta propriedade assintótica é particularmente relevante, pois trabalhamos com séries temporais bivariadas, onde as observações são acumuladas ao longo do tempo. A formulação Bayesiana, combinada com a modelagem por cópulas dinâmicas, permite que o parâmetro $\theta(t)$ seja estimado de forma confiável à medida que a quantidade de dados aumenta, assegurando resultados robustos e consistentes.

Note que, em algumas situações, no entanto, não é possível obter amostras diretamente da distribuição a posteriori, sendo necessário recorrer a técnicas de amostragem como Monte Carlo (METROPOLIS; ULAM, 1949), Gibbs (GEMAN; GEMAN, 1984), Metropolis-Hastings (METROPOLIS et al., 1953; HASTINGS, 1970), entre outras.

Como a Equação (3.4) depende da função cópula, a seguir apresentaremos a formulação da distribuição a posteriori do vetor \mathbf{a} , considerando alguns tipos de cópulas dinâmicas.

3.1.3 Família de cópulas elípticas

A cópula dinâmica Gaussiana tem seu domínio em $\mathcal{D} = (-1, 1)$. Neste contexto, podemos utilizamos a função $\psi : \mathcal{D} \rightarrow \mathbb{R}$ tal que:

$$\psi\{\theta(t)\} = \ln \left\{ \frac{1 + \theta(t)}{1 - \theta(t)} \right\}, \quad \text{com inversa} \quad \psi^{-1}(v) = \frac{\exp(v) - 1}{\exp(v) + 1} \quad \text{e} \quad v = \frac{t - 1}{T}.$$

Como $\psi(\theta(t)) = h(v)$, temos que:

$$\theta(t) = \psi^{-1}\{h(v)\} = \frac{\exp\{h(v)\} - 1}{\exp\{h(v)\} + 1}, \quad \text{em que} \quad h(v) = \sum_{i=0}^n a_i \varphi_i(v),$$

com $\varphi_i(v)$ sendo funções base, como polinômios de Taylor, Bernstein ou Chebyshev. Assim, a distribuição *a posteriori* dos coeficientes \mathbf{a} , conforme a Equação (3.4), é dada por:

$$\pi(\mathbf{a} | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \propto \exp \left(-\frac{1}{2\sigma^2} \mathbf{a}^\top \mathbf{a} \right) \prod_{t=2}^T \frac{1}{\sqrt{1 - \theta(t)^2}} \exp \left(-\frac{\delta_{1t}^2 \theta(t)^2 - 2\delta_{1t} \delta_{2t} \theta(t) + \delta_{2t}^2 \theta(t)^2}{2(1 - \theta(t)^2)} \right),$$

em que $\delta_{it} = \Phi^{-1}(F_{it}(x_{it}))$, $i = 1, 2$ e $\mathbf{a} = (a_0, a_1, \dots, a_n)$

Para o caso da cópula dinâmica t , o domínio da função $\theta(t)$ é $\mathcal{D} = (0, \infty) \times (-1, 1)$. Neste caso, assumimos que os graus de liberdade ν são conhecidos, pois já foram ajustados durante o processo de ajuste das funções marginais como visto no capítulo anterior. Assim, restringimos o domínio da função $\theta(t)$ ao intervalo $(-1, 1)$, ou seja: $\mathcal{D} = (-1, 1)$, mantendo a mesma restrição de correlação válida para o caso estático. Dado que o domínio da função $\theta(t)$ para a cópula t coincide com o da cópula Gaussiana, segue que a distribuição *a posteriori* dos coeficientes \mathbf{a} , de acordo com a Equação (3.4), é dada por:

$$\begin{aligned} \pi(\mathbf{a}|\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \propto \exp\left(-\frac{1}{2\sigma^2}\mathbf{a}^\top\mathbf{a}\right) \prod_{t=2}^T \frac{\Gamma\left(\frac{\nu+2}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}{\left(\Gamma\left(\frac{\nu+1}{2}\right)\right)^2} \cdot \theta(t)^{-\frac{1}{2}} \\ \times \left(1 + \frac{\delta_{1t}^2 - 2\theta(t)\delta_{1t}\delta_{2t} + \delta_{2t}^2}{\nu(1-\theta(t)^2)}\right)^{-\frac{\nu+2}{2}} \\ \times \left(1 + \frac{\delta_{1t}^2}{\nu}\right)^{-\frac{\nu+2}{2}} \left(1 + \frac{\delta_{2t}^2}{\nu}\right)^{-\frac{\nu+2}{2}}, \end{aligned}$$

em que $\delta_{it} = t_\nu^{-1}(F_{it}(x_{it}))$, $i = 1, 2$ e $\mathbf{a} = (a_0, a_1, \dots, a_n)$.

3.1.4 Família de cópulas Arquimedeanas

De maneira geral, conforme a Equação (3.4), todas as distribuições *a posteriori* possuem a mesma estrutura, diferindo apenas na função densidade da cópula correspondente. Para a família das cópulas arquimedeanas, apresentamos a função de transformação ψ , suas respectivas inversas e os domínios associados para cada cópula. A Tabela (3) resume estas funções:

Tabela 3 – Funções de transformação invertíveis para cópulas arquimedeanas

Cópula	\mathcal{D}	$\psi(\theta(t))$	$\psi^{-1}(v)$
AMH	$(-1, 1)$	$\ln\left(\frac{1+\theta(t)}{1-\theta(t)}\right)$	$\frac{\exp(v)-1}{\exp(v)+1}$
Clayton	$(0, \infty)$	$\ln(\theta(t))$	$\exp(v)$
Frank			
Gumbel-Hougaard	$(1, \infty)$	$\ln(\theta(t) - 1)$	$\exp(v) + 1$
Joe			
Gumbel-Barnett	$(0, 1)$	$\ln\left(\frac{\theta(t)}{1-\theta(t)}\right)$	$\frac{\exp(v)}{\exp(v)+1}$

Fonte: Elaborado pelo autor (2024).

Lembrando que $\psi : \Theta \rightarrow \mathbb{R}$ e que o valor de $\theta(t)$ é determinado por:

$$\theta(t) = \psi^{-1}\{h(v)\}, \quad \text{com } v = \frac{t-1}{T} \quad \text{e} \quad h(v) = \sum_{i=0}^n a_i \varphi_i(v),$$

em que $\varphi_i(v)$ pode ser, algum tipo de polinômio base e a_i são os coeficiente de aproximação. Neste trabalho, $\varphi_i(v)$ pode ser Polinômio de Taylor, Bernstein ou Chebyshev. Mais precisamente, $\theta(t)$ é aproximado por

- a) Taylor: $\theta(t) = \psi^{-1}\left(\sum_{i=0}^n a_i \left(\frac{t-1}{T}\right)^i\right), t = 1, 2, \dots, T$ em que $\mathbf{a} = (a_0, a_1, \dots, a_n)^\top$.
- b) Bernstein: $\theta(t) = \psi^{-1}\left(\sum_{i=0}^n a_i B_i^n\left(\frac{t-1}{T}\right)\right), t = 1, 2, \dots, T$ em que, $B_i^n(\cdot)$ é uma base de Bernstein como definido na Equação (2.1) e $\mathbf{a} = (a_0, a_1, \dots, a_n)^\top$.
- c) Chebyshev: $\theta(t) = \psi^{-1}\left(\sum_{i=0}^n a_i T_n\left(\frac{t-1}{T}\right)\right), t = 1, 2, \dots, T$ em que, $T_n(\cdot)$ é uma base de Chebyshev como definida na Equação (2.3) e $\mathbf{a} = (a_0, a_1, \dots, a_n)^\top$.

4 ESTIMAÇÃO

Nesta seção, descrevemos o processo de estimação do vetor de coeficientes $\mathbf{a} = (a_0, a_1, \dots, a_n)$, que permite aproximar a função $\theta(t)$ para cada função cópula dinâmica proposta. Seja $p(\mathbf{a})$ a distribuição *a priori* do vetor \mathbf{a} , conforme apresentado na Seção 3. Assumimos:

$$\mathbf{a} \sim N_d(\mathbf{0}, \sigma^2 \mathbf{I}),$$

em que $d = n + 1$, \mathbf{I} é a matriz identidade de tamanho $d \times d$, e $\sigma^2 > 0$ é o parâmetro de escala que controla a variabilidade inicial.

Por meio de $\pi(\mathbf{a})$ e da função de verossimilhança $L(\mathbf{a})$, conforme descrito no Capítulo 3, em inferência Bayesiana o objetivo é obter a forma da distribuição *a posteriori* $\pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T)$, dada pela Equação (3.4).

Note que a forma da distribuição *a posteriori* pode não ser conhecida e às vezes não é possível obter amostras desta distribuição. Por esta razão, neste trabalho utilizamos métodos de Monte Carlo via Cadeias de Markov (MCMC) (RAFTERY; LEWIS, 1996). Especificamente, adotamos o algoritmo Metropolis-Hastings Adaptativo (METROPOLIS; ULAM, 1949; HASTINGS, 1970; HAARIO; SAKSMAN; TAMMINEN, 2001), detalhado a seguir.

4.1 ALGORITMO METROPOLIS-HASTINGS

O algoritmo Metropolis-Hastings é amplamente utilizado para gerar amostras de uma distribuição de probabilidade $p(\mathbf{a})$, cuja forma analítica é conhecida apenas por uma constante de proporcionalidade. No contexto deste trabalho, $\pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T)$ é a distribuição *a posteriori* do vetor de coeficientes \mathbf{a} .

O algoritmo Metropolis-Hastings constrói uma cadeia de Markov $\{\mathbf{a}_i\}_{i \geq 0}$, cuja distribuição estacionária é $p(\mathbf{a})$.

Considerações Importantes

- **Taxa de Aceitação:** O desempenho do algoritmo depende da taxa de aceitação, que deve ser moderada para evitar dois problemas: se for muito alta, as amostras ficam altamente correlacionadas e a exploração do espaço é ineficiente e se for muito baixa, muitas propostas são rejeitadas e a convergência se torna lenta. Valores de referência

para a taxa de aceitação são 0.44 em distribuições univariadas e 0.234 em distribuições multivariadas (GELMAN; GILKS; ROBERTS, 1997).

- **Distribuição Proposta:** A escolha de $q(\mathbf{a}^*|\mathbf{a}_i)$ desempenha um papel fundamental no desempenho do algoritmo. É importante notar que a escolha da distribuição proposta pode ser um desafio (BROOKS et al., 2011).

Algoritmo 1: Algoritmo de Metropolis-Hastings

Inicializar \mathbf{a}_0 .

Definir a distribuição proposta $q(\cdot | \cdot)$.

for $i = 1, 2, \dots, N$ **do**

 Gerar proposta $\mathbf{a}^* \sim q(\cdot | \mathbf{a}_{i-1})$

 Calcular a probabilidade de aceitação:

$$\alpha(\mathbf{a}_{i-1}, \mathbf{a}^*) = \min \left\{ 1, \frac{p(\mathbf{a}^*)q(\mathbf{a}_{i-1} | \mathbf{a}^*)}{p(\mathbf{a}_{i-1})q(\mathbf{a}^* | \mathbf{a}_{i-1})} \right\}$$

 Gerar $u \sim U(0, 1)$

 Atualizar \mathbf{a}_i :

$$\mathbf{a}_i = \begin{cases} \mathbf{a}^*, & \text{se } u \leq \alpha(\mathbf{a}_{i-1}, \mathbf{a}^*), \\ \mathbf{a}_{i-1}, & \text{caso contrário.} \end{cases}$$

end

Retornar $\{\mathbf{a}_i\}_{i=0}^N$

Dado o desafio de definir uma distribuição proposta eficiente, o Metropolis-Hastings Adaptativo (MHA) surge como uma alternativa para este problema. Este método utiliza uma distribuição proposta normal multivariada com passeio aleatório, $q(\mathbf{a}^*|\mathbf{a}_i) \sim N_d(\mathbf{a}_i, \Sigma_i)$, e adapta dinamicamente a matriz de covariância Σ_i com base nas amostras previamente geradas. Essa adaptação melhora a exploração do espaço paramétrico ao ajustar a matriz de covariância com base nas amostras anteriores, reduzindo a dependência de escolhas iniciais e melhorando a convergência (HAARIO; SAKSMAN; TAMMINEN, 2001). Na próxima seção, detalhamos a implementação do MHA.

4.2 ALGORITMO METROPOLIS-HASTINGS ADAPTATIVO

Neste trabalho, adotamos o algoritmo Metropolis-Hastings Adaptativo (MHA), proposto inicialmente por Haario, Saksman e Tamminen (2001) e posteriormente expandido por autores

como Andreiu e Thoms (2008), Roberts e Rosenthal (2009) e Vinhola (2011). A principal característica deste método é a capacidade de adaptar dinamicamente os parâmetros da distribuição proposta $q(\cdot|\cdot)$ ao longo das iterações, utilizando as informações acumuladas das amostras geradas. Essa adaptação reduz a sensibilidade à escolha inicial da proposta e melhora a eficiência na exploração do espaço paramétrico.

Seja $\pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T)$ a distribuição *a posteriori* do vetor de coeficientes \mathbf{a} , como definido na Seção 3. O Metropolis-Hastings Adaptativo (MHA) busca gerar amostras de $\pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T)$ por meio de uma cadeia de Markov $\{\mathbf{a}_i\}_{i \geq 0}$, cuja distribuição estacionária é $\pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T)$.

Diferentemente do Metropolis-Hastings, o MHA utiliza uma proposta adaptativa:

$$q(\mathbf{a}^*|\mathbf{a}_i) \sim N_d(\mathbf{a}_i, \lambda \Sigma_i),$$

em que Σ_t é a matriz de covariância estimada dinamicamente e λ é o fator de escala ajustado ao longo das iterações.

Observações

- O parâmetro c é calculado como:

$$c = \frac{\sqrt{2\pi}e^{\gamma^2/2}}{2\gamma}, \quad \text{com} \quad \gamma = -\Phi^{-1}\left(\frac{p^*}{2}\right).$$

Essa fórmula define automaticamente o tamanho do passo na adaptação, evitando a necessidade de ajustes manuais, como demonstrado por Garthwaite, Fan e Sisson (2016).

- Quando a matriz Σ é conhecida, então $\lambda = 2,38^2/d$, conforme demonstrado por Gelman, Roberts e Gilks (1996).
- O parâmetro t_0 define o número de iterações iniciais sem adaptação. Durante esse período, a matriz de covariância permanece constante ($\Sigma_i = \Sigma_0$), o que ajuda a estabilizar a cadeia antes do início da adaptação. Este parâmetro reflete a crença que se tem sobre a matriz inicial Σ_0 .
- O parâmetro $\varepsilon > 0$ é um termo de regularização adicionado à matriz de covariância, multiplicando a matriz identidade \mathcal{I}_d , de dimensão $d \times d$. Esse termo garante que Σ_i permaneça definida positiva em todas as iterações.

Algoritmo 2: Algoritmo de Metropolis-Hastings adaptativo

Inicializar $\mathbf{a}_0, \Sigma_0, \lambda_0$.

for $i = 1, 2, \dots, N$ **do**

Gerar $\eta \sim N_d(0, \Sigma_{i-1})$

Proposta $Y \leftarrow \mathbf{a}_{i-1} + \lambda_{i-1} \cdot \eta$

Calcular $\alpha(\mathbf{a}_{i-1}, Y)$

Gerar $u \sim U(0, 1)$

Atualizar \mathbf{a}_i :

$$\mathbf{a}_i = \begin{cases} Y, & \text{se } u \leq \alpha(\mathbf{a}_{i-1}, Y), \\ \mathbf{a}_{i-1}, & \text{caso contrário.} \end{cases}$$

Atualizar os parâmetros:

$$\Sigma_i = \begin{cases} \Sigma_0, & \text{se } i \leq t_0, \\ \lambda_{i-1} \cdot \text{Cov}(\mathbf{a}_0, \dots, \mathbf{a}_{i-1}) + \lambda_{i-1} \varepsilon \mathcal{I}_d, & \text{se } i > t_0 \end{cases}$$

$$\log(\lambda_i) = \log(\lambda_{i-1}) + c \cdot (p_{i-1} - p^*), \quad \text{com } p_{i-1} = \alpha(\mathbf{a}_{i-1}, Y)$$

end

Retornar $\{\mathbf{a}_i\}_{i=0}^N$

Algoritmo 2 é utilizado para estimar os coeficientes \mathbf{a} , incorporando os ajustes discutidos para assegurar uma convergência eficiente e robusta à distribuição objetivo.

4.3 ESTIMATIVA PONTUAL DOS COEFICIENTES

Nesta seção, descrevemos o procedimento detalhado utilizado para calcular as estimativas pontuais do vetor de coeficientes $\mathbf{a} = (a_0, a_1, \dots, a_n)$, que aproximam a função $\theta(t)$. Este processo é baseado nas amostras geradas pelo algoritmo Metropolis-Hastings Adaptativo (MHA), descrito anteriormente, e inclui etapas de burn-in, thinning e replicação.

Seja $\pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T)$ a distribuição *a posteriori* do vetor \mathbf{a} . Desejamos calcular a média *a posteriori* de cada coeficiente a_j , $j = 0, 1, \dots, n$, dada por:

$$\mathbb{E}[a_j|\mathbf{x}_1, \dots, \mathbf{x}_T] = \int a_j \pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T) d\mathbf{a}.$$

No entanto, devido à complexidade da distribuição $\pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T)$, essa integral não pode ser calculada analiticamente.

Utilizamos o método de Monte Carlo para aproximar esta integral. Dado um conjunto de amostras $\{\mathbf{a}_i; i = 1, 2, \dots, N\}$ extraídas de $\pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T)$, o valor esperado pode ser aproximado por:

$$\mathbb{E}[a_j|\mathbf{x}_1, \dots, \mathbf{x}_T] \approx \frac{1}{N} \sum_{i=1}^N a_{i,j}, \quad \text{quando } N \rightarrow \infty,$$

em que $a_{i,j}$ representa o j -ésimo elemento do vetor \mathbf{a}_i .

Etapas para Estimativa Pontual

1. **Execução do MHA:** Realizamos N iterações do MHA para gerar uma cadeia de Markov $\{\mathbf{a}_i\}_{i=0}^N$, cuja distribuição estacionária é $\pi(\mathbf{a}|\mathbf{x}_1, \dots, \mathbf{x}_T)$.
2. **Burn-in e Thinning:**
 - **Burn-in:** Removemos os $i_{\text{burn-in}}$ estados iniciais da cadeia para garantir que as amostras sejam extraídas da distribuição estacionária.
 - **Thinning:** Reduzimos a autocorrelação entre amostras subsequentes, selecionando uma a cada k estados. Seja \tilde{N} o número efetivo de amostras restantes após o burn-in e o thinning.
3. **Estimativa Pontual Inicial:** Após o burn-in e o thinning, calculamos a média das \tilde{N} amostras restantes para obter a estimativa pontual inicial:

$$\hat{a}_j^{(r)} = \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} a_{i,j}, \quad \text{para } j = 0, 1, \dots, n,$$

em que $a_{i,j}$ é o valor do coeficiente j correspondente à i -ésima amostra.

4. **Replicação do Procedimento:** Repetimos todo o procedimento acima (geração da cadeia, burn-in, thinning e cálculo da média) $R = 100$ às vezes, obtendo R estimativas pontuais $\{\hat{\mathbf{a}}^{(r)}; r = 1, 2, \dots, R\}$.
5. **Estimativa Pontual Final:** A estimativa pontual final para cada coeficiente a_j é obtida pela média das R replicações:

$$\hat{a}_j^{\text{final}} = \frac{1}{R} \sum_{r=1}^R \hat{a}_j^{(r)}, \quad \text{para } j = 0, 1, \dots, n. \quad (4.1)$$

Assim, o vetor final de estimativas é dado por:

$$\hat{\mathbf{a}}^{\text{final}} = (\hat{a}_0^{\text{final}}, \hat{a}_1^{\text{final}}, \dots, \hat{a}_n^{\text{final}}).$$

Com este procedimento, obtemos estimativas pontuais para os coeficientes de interesse, que são fundamentais para aproximar a função $\theta(t)$ em cópulas dinâmicas.

5 ESTUDO DE SIMULAÇÃO

Neste capítulo, analisaremos o desempenho do modelo proposto por meio de um estudo de simulação. O objetivo principal é avaliar a precisão da metodologia proposta, utilizando cópulas dinâmicas. Especificamente, testamos a capacidade do modelo para aproximar a função $\theta(t)$ com base em cenários simulados.

As famílias de cópulas dinâmicas que utilizamos neste estudo foram previamente introduzidas nas seções 3.1.3 e 3.1.4, considerando que o parâmetro da cópula é caracterizado por uma função $\theta(t)$, variando ao longo do tempo t . Para aproximar a função $\theta(t)$, utilizamos diferentes bases polinomiais (*Taylor*, *Bernstein* e *Chebyshev*) detalhadas no Capítulo 3. Em resumo:

$$\theta(t) = \psi^{-1} \left(\sum_{i=0}^n a_i \varphi_i \left(\frac{t-1}{T} \right) \right), \quad t = 1, 2, \dots, T, \quad (5.1)$$

em que $\varphi_i(v)$ é a base polinomial proposta. Mais detalhes sobre a definição de $\varphi_i(v)$ e $\psi(\cdot)$ podem ser vistos nas seções (3.1.3) e (3.1.4). Todas as funções, algoritmos e estimativas foram implementadas em Julia, utilizando o pacote `Copulas.jl` (LAVERNY; JIMENEZ, 2024).

5.1 CENÁRIO 1: AVALIAÇÃO DO VIÉS E EQM

Neste cenário, avaliamos o viés e o erro quadrático médio (EQM) associados à aproximação dos coeficientes $\mathbf{a} = (a_0, a_1, \dots, a_n)$ considerando as cópulas Gaussiana, Clayton e Gumbel.

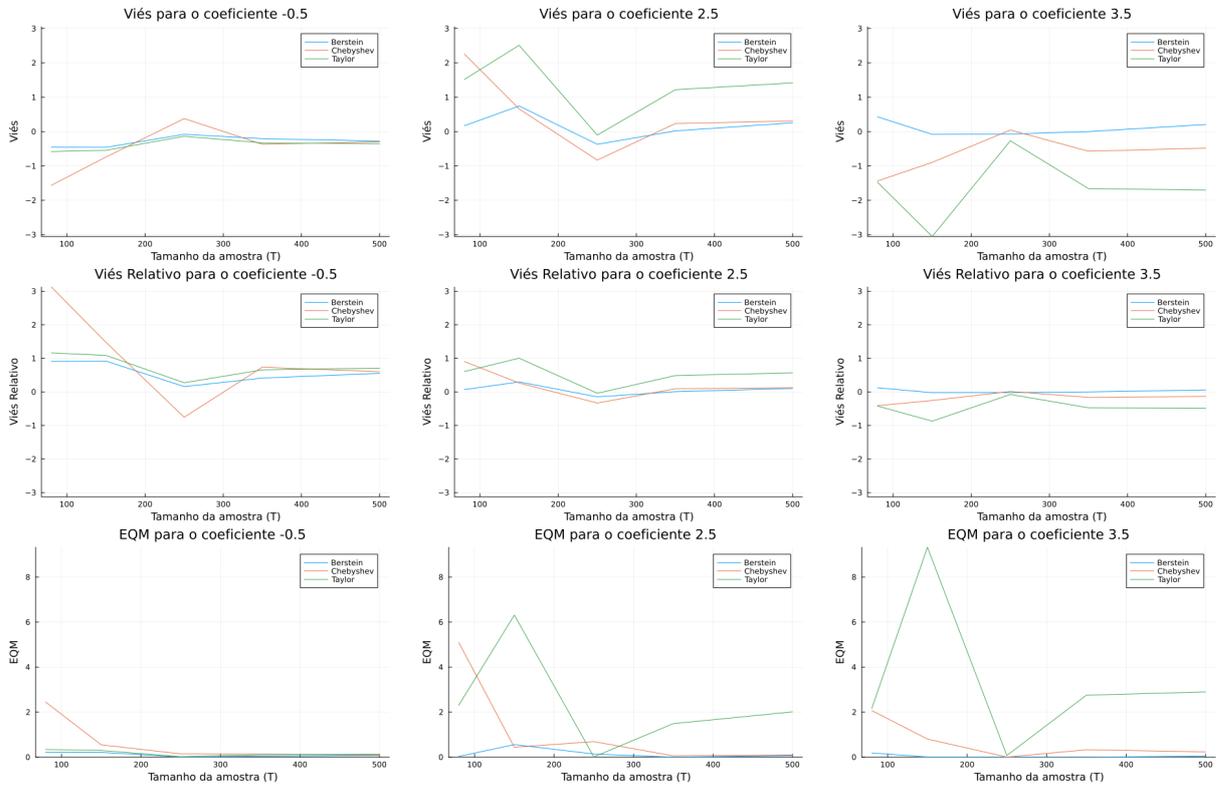
O procedimento consiste em gerar amostras para diferentes tamanhos de amostra $T \in \{80, 150, 250, 350, 500\}$ e calcular os coeficientes aproximados do vetor \mathbf{a} , por meio do Algoritmo 2 (Metropolis-Hastings Adaptativo) e repetir este processo em 100 réplicas. Para cada réplica, realizamos 100000 iterações, aplicando burn-in de 25% e thinning de $k = 3$, o que reduz a correlação entre as amostras extraídas da posteriori $\pi(\mathbf{a} | \mathbf{x}_1, \dots, \mathbf{x}_T)$.

A seguir, apresentamos as figuras, que mostram os gráficos do Viés, Viés Relativo e EQM dos coeficientes estimados para as cópulas Gaussiana, Clayton e Gumbel, considerando os polinômios de Taylor, Bernstein e Chebyshev, avaliando diferentes conjuntos de coeficientes $\mathbf{a} = (a_0, a_1, a_2)$ e tamanhos amostrais.

Cópula Gaussiana

Na figura 8, utilizamos o vetor de coeficientes $\mathbf{a} = (-0.5, 2.5, 3.5)$

Figura 8 – Viés, Viés Relativo e EQM para os coeficientes estimados com Cópula Gaussiana.



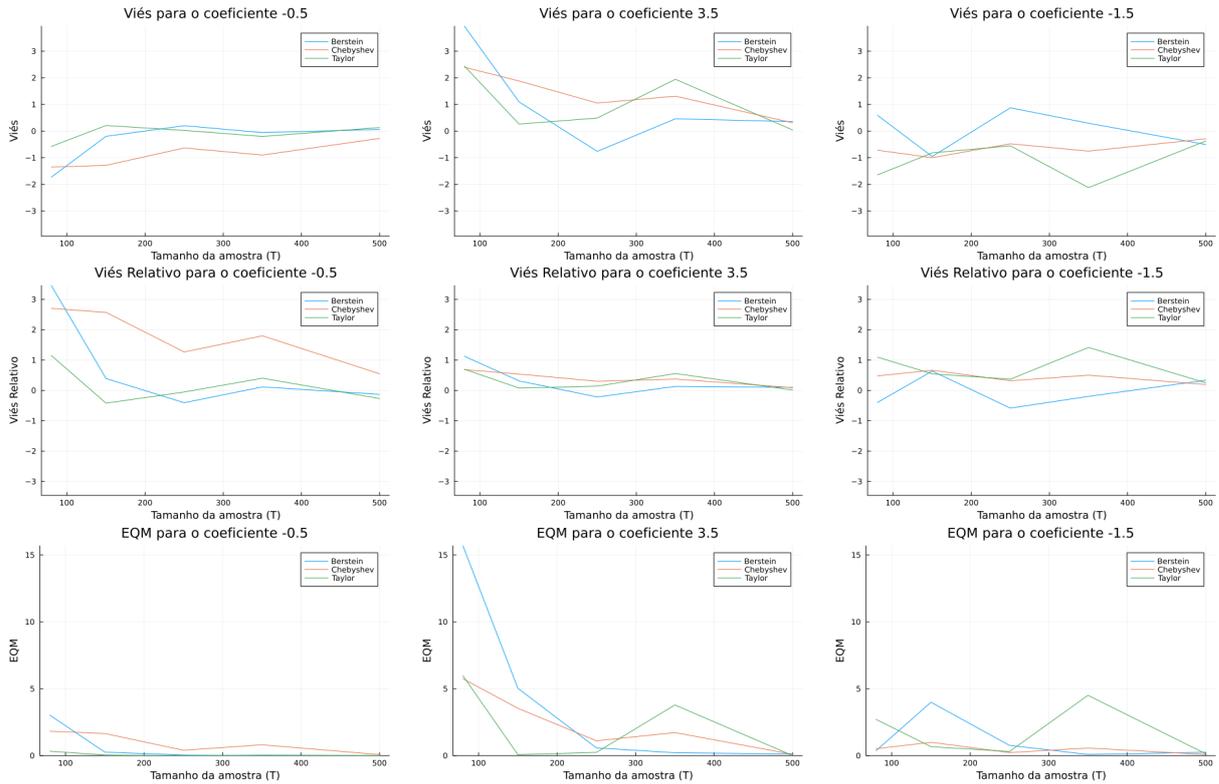
Fonte: Elaborado pelo autor (2024).

Considerando os polinômios de Bernstein e Chebyshev, os valores apresentam menor variação em termos de viés e viés relativo em todos os tamanhos amostrais, especialmente para $T \geq 250$, enquanto Taylor exibe maior variação e oscilações para $T \leq 250$, mas melhora consideravelmente à medida que T aumenta. Em relação ao erro quadrático médio, Bernstein e Chebyshev alcançam menores valores para $T \geq 300$, enquanto Taylor ainda apresenta picos significativos para os coeficientes a_1 e a_2 em diferentes tamanhos amostrais. Para $T \geq 300$, as diferenças entre os três métodos diminuem, embora a estabilidade dos métodos de Bernstein e Chebyshev continue sendo superiores.

Cópula Clayton

Na Figura 9, consideramos o vetor de coeficientes $\mathbf{a} = (-0.5, 3.5, -1.5)$.

Figura 9 – Viés, Viés Relativo e EQM para os coeficientes estimados com Cópula Clayton.



Fonte: Elaborado pelo autor (2024).

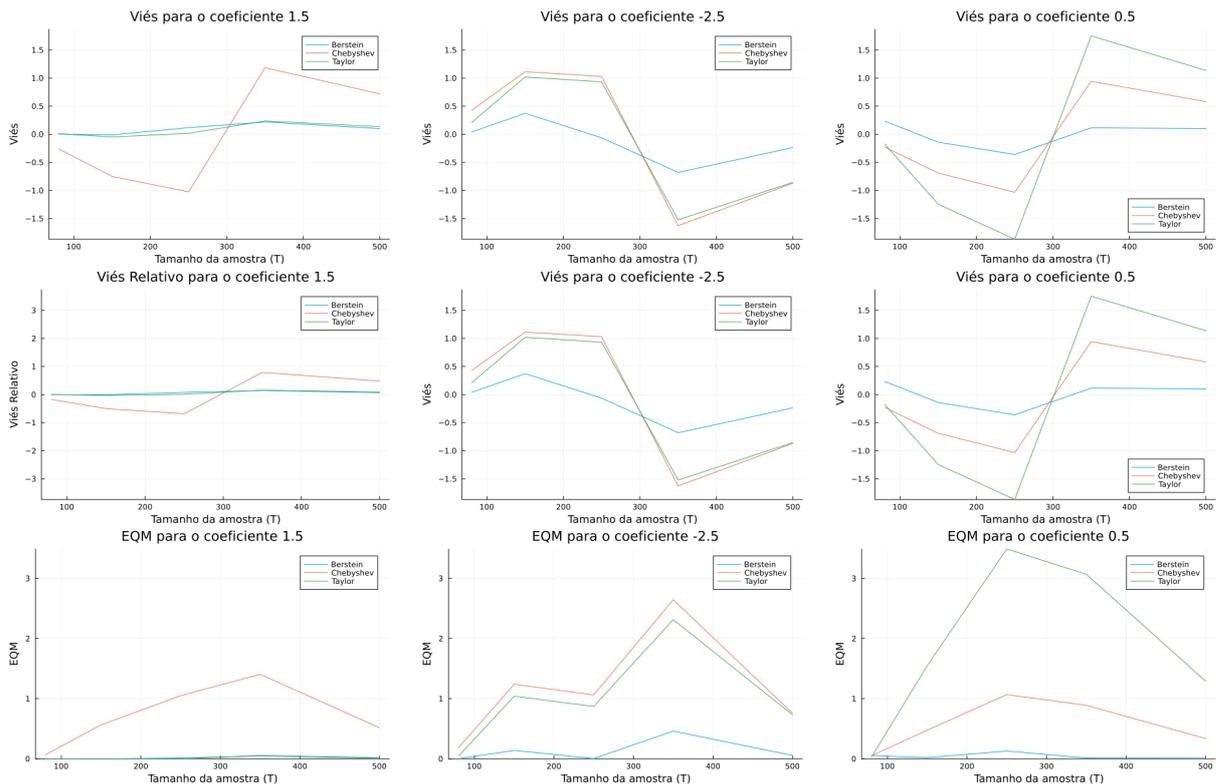
De modo geral, os polinômios de Bernstein e Chebyshev demonstram menor variação em termos de viés e viés relativo para todos os coeficientes analisados, especialmente para $T \geq 250$, onde convergem para valores próximos de zero. O método Taylor, por outro lado, apresenta maior sensibilidade para $T \leq 250$, com flutuações mais acentuadas, particularmente para a_1 e a_2 , embora melhore progressivamente conforme T aumenta. Em relação ao EQM, os métodos de Bernstein e Chebyshev mantêm valores mais baixos em todos os coeficientes e tamanhos de amostra, enquanto Taylor exibe picos elevados em $T \leq 250$, mas se torna competitivo para $T > 350$, reduzindo seus valores próximos de zero.

Cópula Gumbel

Na Figura 10, consideramos o vetor de coeficientes $\mathbf{a} = (1.5, -2.5, 0.5)$. Em geral, o método de Bernstein demonstra a menor variação entre os três, sendo o único a convergir sistematicamente para um viés próximo de zero para todos os coeficientes em $T \geq 80$. Chebyshev e Taylor, por outro lado, apresentam maior variação, especialmente para a_2 , com viés elevado em a_1 e a_2 . Em relação ao viés relativo, Bernstein mantém valores consistentemente baixos

para a_0 e a_1 em todos os tamanhos amostrais, enquanto Chebyshev e Taylor mostram picos elevados para a_2 , independentemente de T . No que se refere ao erro quadrático médio (EQM), Bernstein novamente se destaca com os menores valores em todas as amostras, enquanto Chebyshev e Taylor apresentam EQM mais alto, com Taylor exibindo oscilações mais pronunciadas.

Figura 10 – Viés, Viés Relativo e EQM para os coeficientes estimados com Cópula Gumbel.



Fonte: Elaborado pelo autor (2024).

5.2 CENÁRIO 2: AVALIAÇÃO DE MÉTRICAS PARA $\theta(t)$:

Neste cenário, em vez de utilizar coeficientes fixos, geramos valores reais para o parâmetro $\theta(t)$ a partir de uma função predefinida. Por exemplo, consideramos a função $\theta(t) = \sin\left(\frac{2\pi t}{T}\right)$, para $t = 1, 2, \dots, T$, o que proporciona uma variação contínua e suave do parâmetro ao longo do tempo.

Com os valores reais de $\theta(t)$ definidos, geramos amostras de cópulas dinâmicas correspondentes. Posteriormente, aplicamos os três métodos de aproximação — Taylor, Chebyshev e Bernstein — para estimar a função $\hat{\theta}(t)$. Neste caso, $\hat{\theta}(t)$ é construída com base nos coeficientes estimados $\hat{\mathbf{a}} = (\hat{a}_0, \hat{a}_1, \dots, \hat{a}_3)$.

Neste caso, não é possível avaliar a qualidade da aproximação utilizando métricas como Viés ou EQM, já que não há coeficientes reais para comparação. Em vez disso, avaliamos a proximidade entre as funções $\theta(t)$ e $\hat{\theta}(t)$ utilizando métricas definidas em um espaço métrico adequado.

Seja $\theta(t)$ uma função com domínio discreto $t \in \{1, 2, \dots, T\}$, tal que $\theta(t) \in \mathcal{D}$ e $\mathcal{D} \subseteq \mathbb{R}$. Definimos o conjunto \mathcal{X} como:

$$\mathcal{X} = \{(x_1, x_2, \dots, x_T) \mid x_t \in \mathcal{D}, \forall t = 1, 2, \dots, T\}.$$

Dado que $\mathcal{D} \subseteq \mathbb{R}$, cada $x_t \in \mathcal{D}$ pertence a \mathbb{R} , e por tanto $\mathcal{X} \subseteq \mathbb{R}^T$ em que \mathbb{R}^T é o espaço euclidiano T -dimensional. Consequentemente, como \mathbb{R}^T é um subespaço métrico então, \mathcal{X} é um subespaço métrico de \mathbb{R}^T herdando sua estrutura métrica.

No espaço \mathbb{R}^T , podem-se definir distintas métricas. Como \mathcal{X} é um subespaço de \mathbb{R}^T , essas métricas também são válidas em \mathcal{X} , (KREYSZIG, 1991). Algumas métricas são:

1. Distância Euclidiana (L^2):

$$d_2(x, y) = \sqrt{\sum_{t=1}^T (x_t - y_t)^2}, \quad \forall x, y \in \mathcal{X}.$$

2. Distância Média Quadrática (RMS):

$$d_{\text{RMS}}(x, y) = \sqrt{\frac{1}{T} \sum_{t=1}^T (x_t - y_t)^2}, \quad \forall x, y \in \mathcal{X}.$$

3. Distância Manhattan (L^1):

$$d_1(x, y) = \sum_{t=1}^T |x_t - y_t|, \quad \forall x, y \in \mathcal{X}.$$

4. Distância Máxima (L^∞):

$$d_\infty(x, y) = \max_{t=1, \dots, T} |x_t - y_t|, \quad \forall x, y \in \mathcal{X}.$$

Essas métricas fornecem diferentes formas de avaliar a proximidade entre $\theta(t)$ e $\hat{\theta}(t)$, considerando aspectos como a soma total das diferenças, os valores médios das discrepâncias ou os desvios máximos em pontos específicos. A escolha da métrica depende do objetivo da análise, sendo recomendável utilizar múltiplas métricas para uma avaliação mais completa.

Neste cenário, utilizamos a cópula Gaussiana para aproximar a função de dependência temporal $\theta(t)$. Para avaliar a qualidade do ajuste, comparamos $\theta(t)$ e $\hat{\theta}(t)$ utilizando a distância euclidiana entre as funções (L^2) e a raiz do erro quadrático médio (RMS). Além disso, calculamos os critérios de informação AIC e BIC para balancear a qualidade do ajuste com a complexidade dos modelos. Para a cópula Gaussiana, consideramos $\theta(t) = \sin\left(\frac{2\pi t}{T}\right)$.

A análise conjunta dessas métricas e critérios permite identificar o método mais adequado para ajustar a função $\theta(t)$, considerando precisão e parcimônia.

Tabela 4 – Métricas e Critérios de informação para a função $\theta(t)$ com cópula Gaussiana

		Grau do polinômio		
Método	Métrica	3	4	5
Bernstein	L^2	3.7137	3.7342	3.2394
	RMS	0.1660	0.1670	0.1448
	AIC	-442.9028	-441.5757	-441.9458
	BIC	-426.0443	-420.5026	-416.6582
Chebyshev	L^2	2.9250	3.5810	3.4142
	RMS	0.1308	0.1601	0.1526
	AIC	-389.9114	-423.4579	-440.9909
	BIC	-373.0530	-402.3849	-415.7033
Taylor	L^2	3.2932	3.0768	3.5262
	RMS	0.1472	0.1375	0.1577
	AIC	-374.1122	-413.0739	-407.4803
	BIC	-357.2538	-392.0008	-382.1926

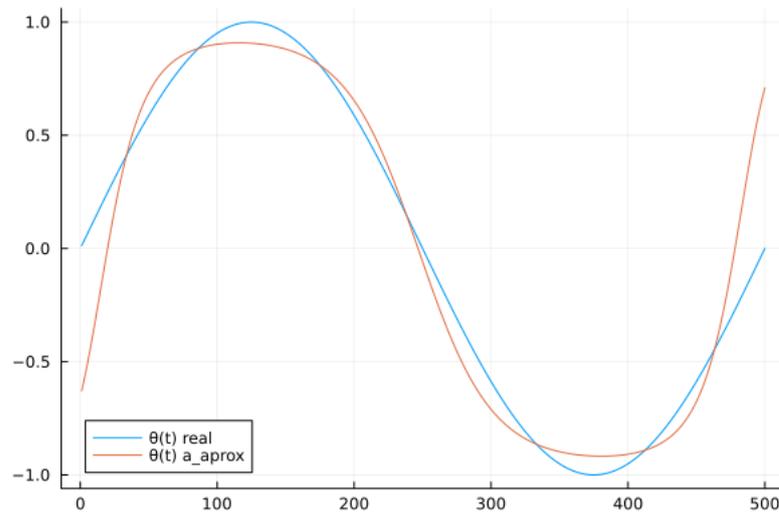
Fonte: Elaborado pelo autor (2024).

A Tabela 4 apresenta os valores das métricas L^2 , RMS e dos critérios AIC e BIC para os métodos de aproximação (Bernstein, Chebyshev e Taylor) na aproximação da função $\theta(t) = \sin\left(\frac{2\pi t}{T}\right)$ com cópula Gaussiana. O polinômio de Bernstein de grau 3 apresenta os menores valores de L^2 (3.2394), RMS (0.1448), AIC (-441.9458) e BIC (-416.6582), demonstrando maior precisão e parcimônia. Já os métodos de Chebyshev e Taylor apresentam desempenhos inferiores.

Note que, existe uma relação clara entre as métricas L^2 e RMS e os critérios de informação AIC e BIC, já que menores valores de L^2 e RMS, que indicam melhor qualidade de ajuste, geralmente resultam em menores valores de AIC e BIC. No entanto, AIC e BIC não dependem apenas dessas métricas, pois também incluem uma penalização baseada no número de parâmetros do modelo.

Assim, modelos com ajuste ligeiramente inferior podem ser preferidos sob BIC se apresentarem menos parâmetros, destacando a importância da parcimônia na escolha do método mais adequado.

Figura 11 – Aproximação da função seno com cópula Gaussiana



Fonte: Elaborado pelo autor (2024).

Na Figura 11, observa-se a aproximação da função $\theta(t)$ com cópula gaussiana utilizando polinômios de Bernstein. A curva aproximada $\hat{\theta}(t)$ acompanha bem a forma geral da função real $\theta(t)$, com pequenas discrepâncias nas regiões próximas aos extremos. Essa visualização evidencia a eficácia do método em capturar a estrutura senoidal da função subjacente, embora as diferenças nas caudas sugiram uma certa ineficiência na captura da dependência em cenários extremos.

6 APLICAÇÃO

Neste capítulo, aplicamos a metodologia proposta utilizando dados reais. Primeiramente, realizamos uma análise detalhada das séries de taxas de câmbio selecionadas, ajustando funções marginais e uma função cópula para modelar a dependência entre elas. Em seguida, utilizamos esses resultados para calcular o Valor em Risco (VaR) de um portfólio simulado, combinando os ativos em diferentes proporções.

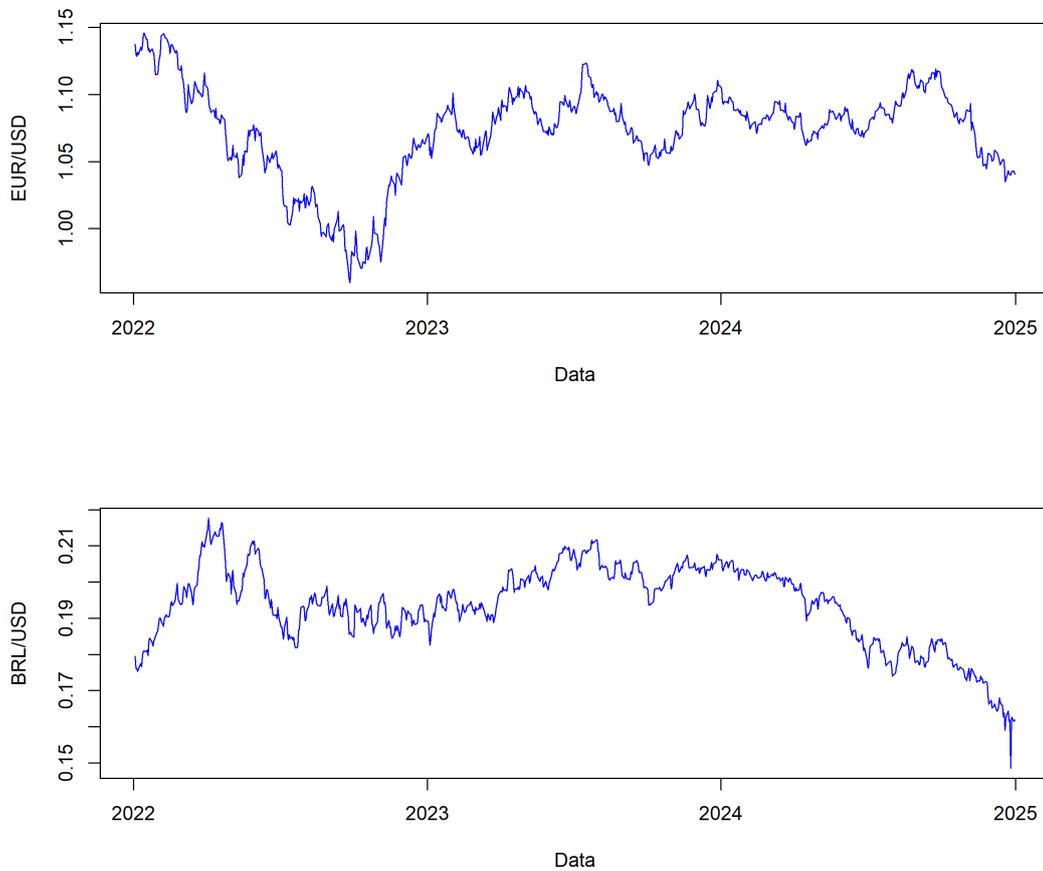
Estudos anteriores fundamentam a relevância de modelar dependências não lineares e assimétricas em mercados financeiros. Patton (2006b) analisou as taxas de câmbio DEM/USD (Marco Alemão em relação ao Dólar Americano) e JPY/USD (Iene Japonês em relação ao Dólar Americano), demonstrando como as cópulas condicionais capturam dependências dinâmicas e mudanças em cenários extremos, especialmente durante períodos de alta volatilidade. Giacomini e Härdle (2005) aplicaram séries como DEM/USD e GBP/USD (Libra Esterlina ao Dólar Americano) na estimação do VaR, destacando a eficácia das cópulas para mensurar riscos em portfólios financeiros. Berger (2013) ampliou essa abordagem ao explorar taxas de câmbio regionais como GBP/EUR (Libra Esterlina em relação ao Euro) e CZK/EUR (Coroa Checa em relação ao Euro), analisando os benefícios da diversificação dentro do mercado europeu.

Neste trabalho, utilizamos essas ideias e as aplicamos ao contexto específico das taxas de câmbio entre moedas de economias com características distintas: uma avançada (EUR/USD, Euro em relação ao Dólar Americano) e outra emergente (BRL/USD, Real Brasileiro em relação ao Dólar Americano), ambas cotadas em dólares. O objetivo é capturar as dinâmicas de dependência entre essas moedas e avaliar como a diversificação entre elas pode impactar o risco em um portfólio.

6.1 ANÁLISE DAS SÉRIES

Neste estudo, utilizamos as séries temporais das taxas de câmbio EUR/USD e BRL/USD, ambas precificadas em dólares. Os dados foram obtidos na plataforma Yahoo Finance (<<https://finance.yahoo.com>>) no período de 01/01/2022 a 31/12/2024, totalizando $T = 781$ observações.

Figura 12 – Retornos de EUR/USD e BRL/USD.



Fonte: Elaborado pelo autor (2024).

Na Figura 12, pode-se observar o comportamento dos retornos das séries EUR/USD e BRL/USD. Ambas apresentam variações ao longo do tempo, refletindo as flutuações nos mercados de câmbio. Nota-se que essas variações sugerem a presença de períodos de maior ou menor instabilidade, característica típica de séries financeiras.

Para realizar o ajuste da cópula e modelar a dependência entre as séries, utilizamos os log-retornos das taxas de câmbio EUR/USD e BRL/USD. Os log-retornos são calculados como:

$$x_t = \ln \left(\frac{P_t}{P_{t-1}} \right),$$

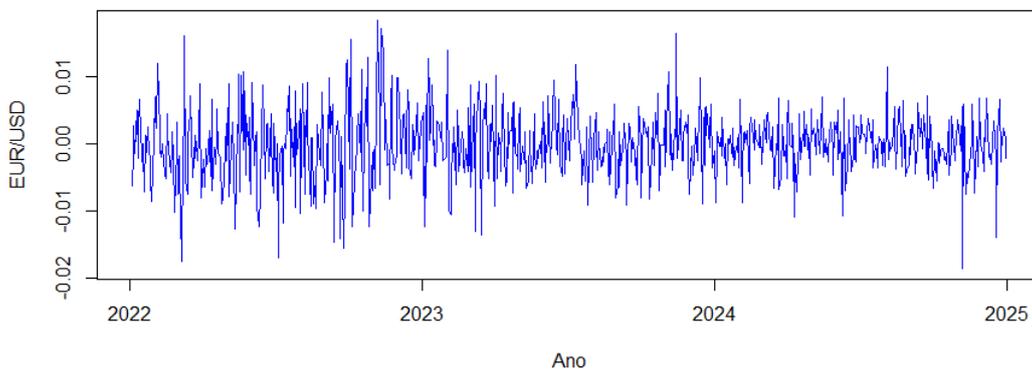
em que P_t representa o valor da taxa de câmbio no tempo t .

Antes de proceder ao ajuste do modelo, realizamos uma análise detalhada das séries de log-retornos, o que nos permite identificar características importantes dessas séries.

Série EUR/USD

Seja X_{1t} a série temporal que representa os log-retornos da cotação do Euro frente ao Dólar (EUR/USD).

Figura 13 – Log-retornos de EUR/USD



Fonte: Elaborado pelo autor (2024).

Nos log-retornos apresentados na Figura 13 observa-se a ocorrência de picos positivos e negativos frequentes, comportamento típico de séries financeiras (MORETTIN, 2017).

Tabela 5 – Estatísticas para a série X_{1t}

Estatística	Média	Desvio padrão	Assimetria	Curtose
Valor	-0.0001	0.0049	-0.0015	4.1608

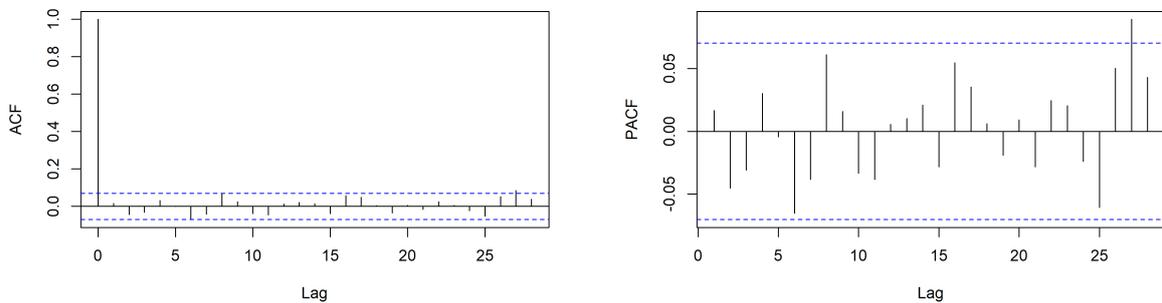
Fonte: Elaborado pelo autor (2024).

Na Tabela 5, apresentam-se algumas estatísticas descritivas da série X_{1t} . É superior à da distribuição normal, indicando que a série tem uma distribuição com caudas mais pesadas do que a distribuição normal. Já o coeficiente de assimetria indica que a série não apresenta assimetria, com média em torno de zero. Para aprofundar essa análise, foram calculadas as funções de autocorrelação (ACF) e autocorrelação parcial (PACF).

Os resultados das funções ACF e PACF, apresentados na Figura 14, mostram que as autocorrelações estão majoritariamente dentro dos intervalos de confiança, indicando ausência de dependência linear significativa nos log-retornos.

O teste de Ljung-Box, realizado para X_{1t} em até 30 defasagens, reforça essa observação, com p-valor de 0.3076. Assim, não rejeitamos a hipótese nula de ausência de autocorrelação

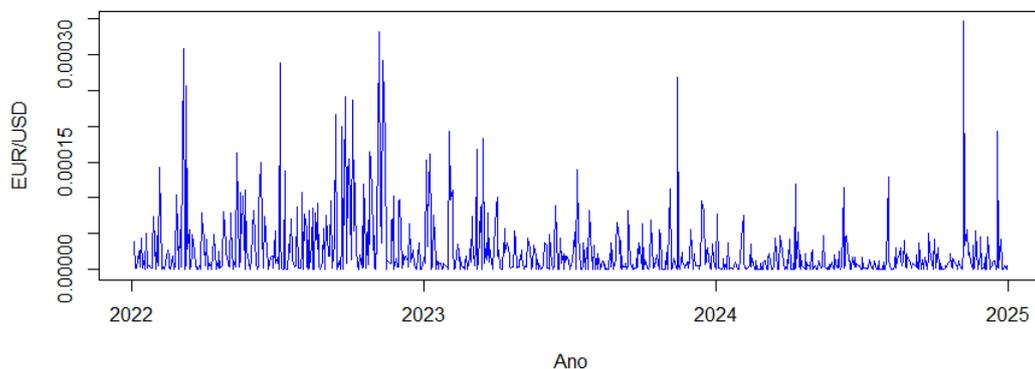
Figura 14 – Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) dos log-retornos de EUR/USD.



Fonte: Elaborado pelo autor (2024).

significativa. Portanto, os resultados indicam que não é necessário ajustar um modelo ARMA para essa série.

Figura 15 – Log-retornos ao quadrado de EUR/USD



Fonte: Elaborado pelo autor (2024).

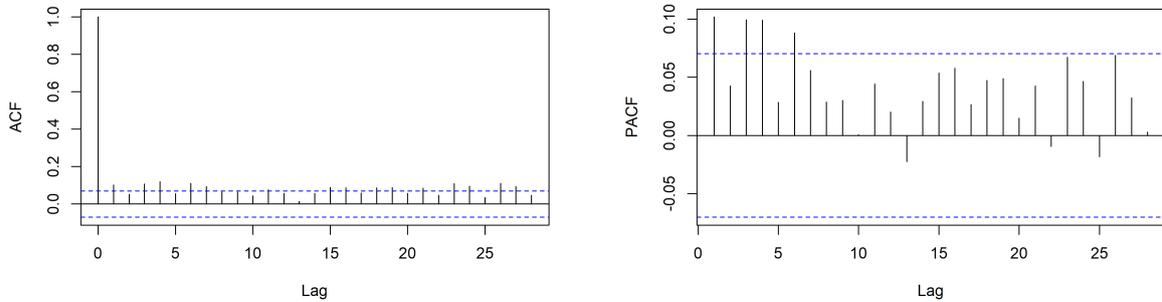
A análise dos log-retornos elevados ao quadrado X_{1t}^2 , indicam que os dados apresentam heterocedasticidade condicional. A Figura 15 apresenta picos de volatilidade em períodos específicos, indicando a presença de variação condicional ao longo do tempo.

Para investigar esses padrões, analisaremos as ACF e PACF de X_{1t}^2 para determinar a necessidade de modelagem da variância condicional por meio de um modelo GARCH.

Na Figura 16, é possível observar que a ACF apresenta valores que passam os limites de significância em diferentes defasagens, principalmente nos primeiros lags. Isso indica que existe dependência significativa na variância condicional.

A PACF, por sua vez, também mostra valores significativos nos primeiros lags, mas diminuem conforme a defasagem aumenta.

Figura 16 – Funções de autocorrelação e autocorrelação parcial dos log-retornos ao quadrado de EUR/USD.



Fonte: Elaborado pelo autor (2024).

Para modelar a variância condicional da série de log-retornos de EUR/USD X_{1t} , foi utilizado o modelo GARCH(1,1). Seja $\epsilon_{1t} = \sigma_{1t}Z_{1t}$, em que σ_{1t}^2 representa a variância condicional no tempo t , e Z_{1t} são resíduos padronizados. O modelo GARCH(1,1) é dado por:

$$\sigma_{1t}^2 = \omega + \alpha\epsilon_{1(t-1)}^2 + \beta\sigma_{1(t-1)}^2,$$

com $\omega > 0$, $0 < \alpha, \beta < 1$ e $\alpha + \beta < 1$.

A escolha do modelo baseou-se na análise dos resíduos padronizados, que não apresentaram autocorrelação nem heterocedasticidade, conforme os testes Ljung-Box e ARCH LM. Os parâmetros estimados são:

$$\hat{\sigma}_{1t}^2 = 0.020515\hat{\epsilon}_{1(t-1)}^2 + 0.978485\hat{\sigma}_{1(t-1)}^2.$$

Note que ω é omitido porque ele não é estatisticamente significativo para o modelo. Os resíduos padronizados estimados são:

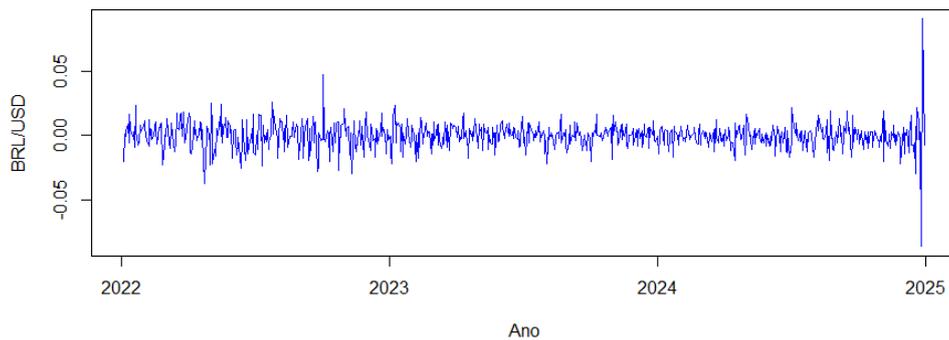
$$\hat{\epsilon}_{1t} = \hat{\sigma}_{1t}Z_{1t}, \quad \text{com } Z_{1t} \sim \text{GED}(1.44),$$

em que $\text{GED}(\cdot)$ denota a distribuição de Erro Generalizada com $\hat{\nu} = 1.44$ graus de liberdade (NELSON, 1991). A escolha dessa distribuição se deve à sua flexibilidade em modelar caudas mais leves ou mais pesadas, conforme observado nos dados. No caso da série X_{1t} , a curtose levemente elevada (4.1608) sugere uma distribuição com caudas mais pesadas do que a normal, justificando o uso da GED para capturar essa característica específica dos resíduos.

Série BRL/USD

Seja X_{2t} a série temporal que representa os log-retornos da cotação do Real Brasileiro frente ao Dólar (BRL/USD).

Figura 17 – Log-retornos de BRL/USD



Fonte: Elaborado pelo autor (2024).

A Figura 17 apresenta os log-retornos da cotação do Real Brasileiro frente ao Dólar Americano (BRL/USD). Observa-se que os valores flutuam em torno de zero, como esperado para séries de log-retornos. No final do período analisado, nota-se um aumento na magnitude das flutuações, o que pode indicar períodos de maior instabilidade nos retornos.

Tabela 6 – Estatísticas para a série X_{2t}

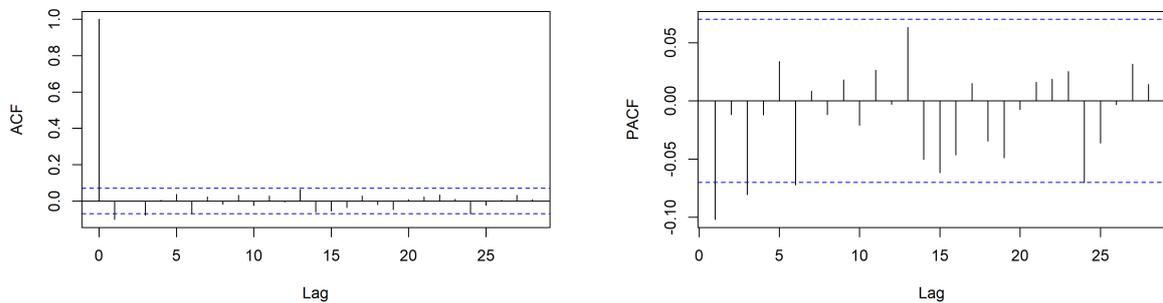
Estatística	Média	Desvio padrão	Assimetria	Curtose
Valor	-0.0001	0.0101	0.0312	17.6753

Fonte: Elaborado pelo autor (2024).

Na Tabela 6, apresentam-se algumas estatísticas descritivas da série X_{2t} . Observa-se que os dados apresentam curtose alta, indicando que a série tem uma distribuição com caudas pesadas. Já o coeficiente de assimetria indica que a série não apresenta assimetria, com média em torno de zero.

Na Figura 18, a ACF exhibe valores significativos em algumas defasagens iniciais, sugerindo a presença de dependência linear nos dados. A PACF, por sua vez, apresenta significância nos primeiros lags, indicando possível autocorrelação nas observações. O teste de Ljung-Box, realizado para X_{2t} em até 20 defasagens, reforça essa observação, com p-valor de 0.03045.

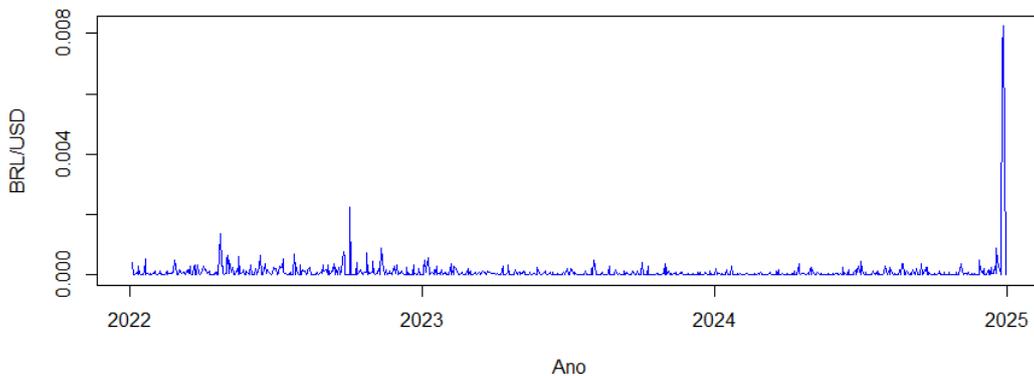
Figura 18 – Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) dos log-retornos de BRL/USD.



Fonte: Elaborado pelo autor (2024).

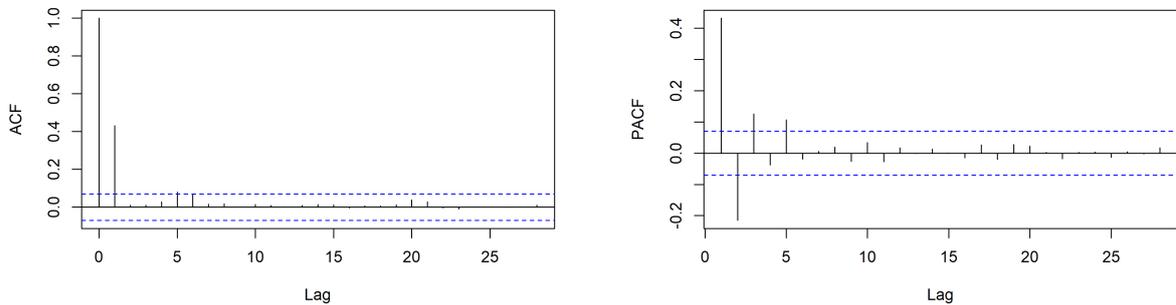
Na Figura 19, apresentamos os log-retornos ao quadrado da cotação do Real Brasileiro frente ao Dólar (BRL/USD). Observa-se a presença de picos de volatilidade em períodos específicos, especialmente no final do período analisado. Esses picos sugerem que a variância condicional da série não é constante ao longo do tempo, indicando a presença de heterocedasticidade condicional.

Figura 19 – Log-retornos ao quadrado de BRL/USD



Na Figura 20, é possível observar que a ACF apresenta valores que passam os limites de significância em diferentes defasagens, especialmente nos primeiros lags. Isso sugere a presença de dependência significativa na variância condicional. A PACF, por sua vez, também mostra valores significativos nos primeiros lags, mas diminuem conforme a defasagem aumenta.

Figura 20 – Funções de autocorrelação e autocorrelação parcial dos log-retornos ao quadrado de BRL/USD.



Fonte: Elaborado pelo autor (2024).

Para modelar a variância condicional de X_{2t} , será utilizado um modelo $ARMA(1, 1) - ARCH(1)$, selecionado por garantir que os resíduos padronizados não apresentem autocorrelação nem heterocedasticidade, conforme indicado pelos testes de Ljung-Box e ARCH LM. O modelo é definido como:

$$X_{2t} = \phi X_{2(t-1)} + \eta \epsilon_{2(t-1)} + \epsilon_{2t},$$

em que ϕ representa o coeficiente autorregressivo e η o coeficiente da média móvel. A variância condicional segue um processo $ARCH(1)$:

$$\sigma_{2t}^2 = \omega + \alpha \epsilon_{2(t-1)}^2,$$

em que $\epsilon_{2t} = \sigma_{2t} Z_{2t}$, e $Z_{2t} \sim t(\nu)$. O modelo ajustado é:

$$\hat{X}_{2t} = -0.51840 \hat{X}_{2(t-1)} - 0.44233 \hat{\epsilon}_{2(t-1)} + \hat{\epsilon}_{2t},$$

e variância condicional estimada é:

$$\hat{\sigma}_{2t}^2 = 0.99900 \hat{\epsilon}_{2(t-1)}^2.$$

Note que ω é omitido porque não é estatisticamente significativo para o modelo. Os resíduos padronizados estimados são:

$$\hat{\epsilon}_{2t} = \hat{\sigma}_{2t} Z_{2t}, \quad Z_{2t} \sim t(3.45),$$

em que $t(\cdot)$ denota a distribuição t -Student com $\hat{\nu} = 3.45$ graus de liberdade. A escolha dessa distribuição se deve à presença de caudas pesadas na série, conforme evidenciado pela alta curtose (14.6753). Para obter as pseudo-observações necessárias para o ajuste da cópula,

aplicamos a função de distribuição acumulada da distribuição GED e t -Student ajustada para cada série. Assim, temos:

$$u_{1t} = GED(Z_{1t}; \nu_1), \quad u_{2t} = t(Z_{2t}; \nu_2),$$

em que $GED(Z_{1t}; \nu_1)$ é a função de distribuição acumulada do erro generalizada e $t(\cdot; \nu_2)$ é a função de distribuição acumulada t -Student. Neste contexto, $\hat{\nu}_1 = 1.44$ e $\hat{\nu}_2 = 3.45$ correspondem aos graus de liberdade ajustados para Z_{1t} e Z_{2t} , respectivamente.

Com as pseudo-observações $u_{1t}, u_{2t} \in [0, 1]$, obtemos as transformações marginais necessárias para a estimação de uma cópula dinâmica. Nesse modelo, o parâmetro $\theta(t)$ varia ao longo do tempo de forma funcional, permitindo capturar a dependência entre as séries X_{1t} (EUR/USD) e X_{2t} (BRL/USD) de maneira flexível e adaptativa.

Para isso, aplicamos os procedimentos descritos nos Capítulos 3 e 4, utilizando diferentes famílias de cópulas dinâmicas. A seleção do modelo que melhor se ajusta foi realizada com base nos critérios de informação de Akaike (AIC) e Bayessiano (BIC), permitindo identificar a cópula mais adequada para descrever a dependência entre as séries.

Nas Tabelas 7, 8 e 9, são apresentados os valores dos critérios AIC e BIC para diferentes cópulas e graus de polinômios, utilizando os métodos de Bernstein, Chebyshev e Taylor, respectivamente.

Tabela 7 – Critérios AIC e BIC para diferentes cópulas e graus de polinômios utilizando o método de Bernstein

		Grau do polinômio				
Cópula	Criterio	2	3	4	5	6
Gaussiana	AIC	-21.8291	-21.0901	-19.1988	-21.2956	-19.3721
	BIC	-7.8474	-2.4478	4.1039	6.6677	13.2519
t $\nu = 1.45$	AIC	113.8421	115.6280	117.1228	118.4073	120.2681
	BIC	127.8239	134.2703	140.4256	146.3707	152.8921
t $\nu = 3.45$	AIC	113.8423	115.6276	117.1228	118.4090	120.2733
	BIC	127.8241	134.2699	140.4257	146.3724	152.8974
Frank	AIC	-23.9809	-21.9535	-20.0773	-18.0773	-16.0773
	BIC	-9.9992	-3.3112	3.2255	9.8861	16.5466
Gumbel	AIC	2.26188	-2.8820	-0.3221	0.5570	0.2729
	BIC	16.2436	15.7602	22.9806	28.5205	32.8970

Fonte: Elaborado pelo autor (2024).

Tabela 8 – Critérios AIC e BIC para diferentes cópulas e graus de polinômios utilizando o método de Chebyshev

Cópula		Critério	Grau do polinômio				
			2	3	4	5	6
Gaussiana		AIC	-21.8299	-21.0406	-19.1825	-17.2391	-18.2184
		BIC	-7.8481	-2.39837	4.12035	10.7242	14.40553
t	$\nu = 1.45$	AIC	113.8419	115.6371	117.6246	118.8927	120.5804
		BIC	127.8237	134.2794	140.9275	146.8562	153.2044
t	$\nu = 3.45$	AIC	113.8420	115.6405	117.6228	118.8901	120.5594
		BIC	127.8238	134.2828	140.9256	146.8536	153.1834
Frank		AIC	-24.9036	-19.0760	-20.0773	-18.0773	-16.0773
		BIC	-10.9219	-0.4337	3.2255	9.8861	16.5466
Gumbel		AIC	2.7605	1.7592	1.0261	-0.5968	1.6658
		BIC	16.7422	20.4015	24.3289	27.3666	34.2899

Fonte: Elaborado pelo autor (2024).

Tabela 9 – Critérios AIC e BIC para diferentes cópulas e graus de polinômios utilizando o método de Taylor

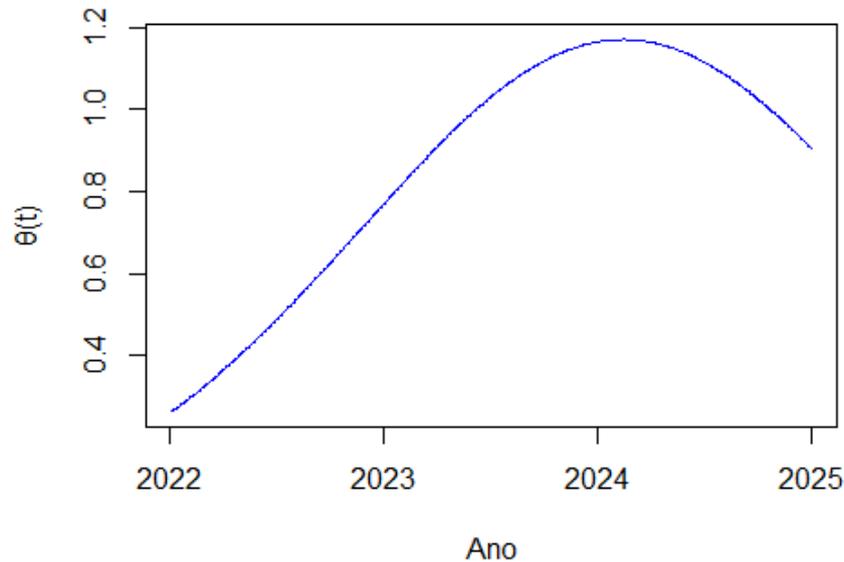
Cópula		Critério	Grau do polinômio				
			2	3	4	5	6
Gaussiana		AIC	-21.8291	-21.0444	-19.1672	-17.1605	-15.3526
		BIC	-7.8473	-2.4021	4.13560	10.8028	17.2713
t	$\nu = 1.45$	AIC	113.8420	115.6489	117.6391	119.4134	121.1146
		BIC	127.8237	134.2912	140.9419	147.3769	153.7386
t	$\nu = 3.45$	AIC	113.8420	115.6540	117.6274	119.4254	121.1078
		BIC	127.8237	134.2963	140.9302	147.3888	153.7319
Frank		AIC	-24.8877	-22.3915	-19.7024	-16.8085	-4.8805
		BIC	-10.9060	-3.74926	3.6004	11.1548	27.7434
Gumbel		AIC	0.4524	4.1251	6.1324	6.5528	10.3774
		BIC	14.4341	22.7674	29.4353	34.5163	43.0014

Fonte: Elaborado pelo autor (2024).

De acordo com os critérios de informação apresentados nas tabelas acima, selecionamos a cópula Frank para modelar a dependência entre as séries temporais, por apresentar os menores valores de AIC e BIC comparados aos demais modelos testados. O ajuste da função $\theta(t)$ é dado por:

$$\theta(t) = \exp\{h(v)\}, \quad h(v) = \sum_{i=0}^2 a_i T_2(v), \quad \text{e} \quad v = \frac{t-1}{781}, \quad t = 1, 2, \dots, 781,$$

em que, $T_2(\cdot)$ representa a base de Chebyshev, definida na Equação (2.3) e o vetor de coeficientes estimados é $\hat{\mathbf{a}} = (-2.8251, 4.2166, -1.4945)^\top$, obtidos utilizando o Algoritmo 2.

Figura 21 – $\theta(t)$ ajustada para a Cópulas Frank

Fonte: Elaborado pelo autor (2024).

Na Figura 21, observamos que a função $\theta(t)$ da cópula de Frank, varia ao longo do tempo, indicando que os valores de dependência entre as séries não é constante. Valores mais altos de $\theta(t)$ refletem períodos de maior dependência entre as séries, enquanto valores mais baixos sugerem menor associação. Essa dinâmica pode estar associada a mudanças nas condições de mercado ou eventos específicos que afetaram simultaneamente os ativos, evidenciando que a relação entre eles é sensível ao tempo.

6.2 MEDIDA DE RISCO (VAR)

De acordo com Morettin (2017), o *Valor em Risco* (VaR), traduzido do inglês *Value at Risk*, é uma medida amplamente utilizada em finanças para quantificar o risco de perda máxima esperada de uma carteira de ativos financeiros durante um horizonte temporal específico e com um nível de significância α . Esta medida busca responder à seguinte questão: *Qual é a perda máxima esperada, sob condições normais de mercado, com probabilidade α , em um determinado período de tempo?*

Para uma posição *comprada*, o VaR é definido como o menor valor l tal que a probabilidade de a perda $\Delta P(h)$ ser menor ou igual a l , sobre o horizonte h , seja igual ao nível de significância

α . Formalmente, tem-se:

$$VaR_{t+h}(\alpha) = \inf\{l \in \mathbb{R} : P(\Delta P(h) \leq l) = \alpha\},$$

em que $\Delta P(h) = P(t+h) - P(t)$ representa a variação do valor do ativo entre os instantes t e $t+h$. Considerando a função de distribuição acumulada (f.d.a.) $F_{\Delta P(h)}(\cdot)$, essa definição equivale a:

$$VaR_{t+h}(\alpha) = \inf\{l \in \mathbb{R} : F_{\Delta P(h)}(l) = \alpha\}.$$

O VaR pode ser calculado utilizando distribuições teóricas, como a normal ou t -Student, ou métodos empíricos ajustados aos dados históricos. No presente trabalho, será utilizada a definição para posição comprada, ajustando-se às características das séries financeiras modeladas neste estudo. Nesta seção, apresentamos a metodologia para calcular o VaR considerando dados reais e o ajuste de cópulas dinâmicas como ferramentas para modelar a dependência entre séries temporais financeiras.

Após obter os valores da função $\theta(t)$ da cópula apresentados na seção anterior, procedemos à estimativa do VaR utilizando as simulações geradas. Para isso, seguimos os passos descritos por Giacomini e Härdle (2005):

1. **Geração de dados simulados:** Utilizando os valores da função $\theta(t)$ ajustada, geramos 50000 amostras u_{1t} e u_{2t} a partir da cópula correspondente. Essas amostras são transformadas nos resíduos padronizados Z_{it} utilizando as distribuições marginais inversas de erro generalizado e t -Student. Os log-retornos são então reconstruídos da seguinte maneira:

$$X_{1t} = \epsilon_{1t}, \quad X_{2t} = \phi X_{2(t-1)} + \eta \epsilon_{2(t-1)} + \epsilon_{2t},$$

em que os resíduos são dados por, $\epsilon_{it} = \sigma_{it} Z_{it}$, $i = 1, 2$. Aqui, σ_{it} representa as volatilidades condicionais estimadas previamente pelos modelos ARMA-GARCH, ϕ é o coeficiente (AR) e η é o coeficiente da média móvel (MA).

2. **Construção do portfólio financeiro:** O valor do portfólio é calculado como:

$$V_t = \sum_{j=1}^2 w_j S_{j(t-1)} e^{X_{jt}},$$

em que w_1 e w_2 representam os pesos dos ativos, $S_{j(t-1)}$ são os preços no instante $t-1$.

As perdas simuladas L_{T+1} são obtidas como:

$$L_{T+1} = \sum_{j=1}^2 w_j S_{jT} (e^{X_{j(T+1)}} - 1),$$

onde S_{jT} é o preço no instante T , e $(e^{X_{j(T+1)}} - 1)$ representa o retorno percentual no período $T + 1$.

3. **Cálculo do VaR:** Com base na distribuição empírica das perdas simuladas L_t , o VaR é estimado utilizando o quantil correspondente ao nível de significância desejado α , para os valores 10%, 5% e 1% (correspondentes a níveis de confiança de 90%, 95% e 99%, respectivamente).

Os valores calculados do VaR para diferentes proporções de w_1 (associado ao EUR/USD) e w_2 (associado ao BRL/USD) estão apresentados na Tabela 10. Observa-se que, à medida que o peso w_1 aumenta, os valores absolutos do VaR crescem para todos os níveis de confiança (90%, 95% e 99%), refletindo um maior risco associado ao EUR/USD, moeda de uma economia avançada, em comparação ao BRL/USD, de uma economia emergente. Esse comportamento está de acordo com a literatura (PATTON, 2006b; BERGER, 2013; CECCHETTI; SCHOENHOLTZ, 2017), onde pares envolvendo moedas de economias desenvolvidas apresentam maior sensibilidade a eventos sistêmicos globais, enquanto moedas emergentes tendem a responder a fatores regionais específicos. Assim, a composição do portfólio afeta diretamente a exposição ao risco, destacando a importância de modelar a dependência dinâmica entre essas taxas de câmbio para estratégias de diversificação e controle de risco.

Tabela 10 – VaR utilizando Cópula Frank

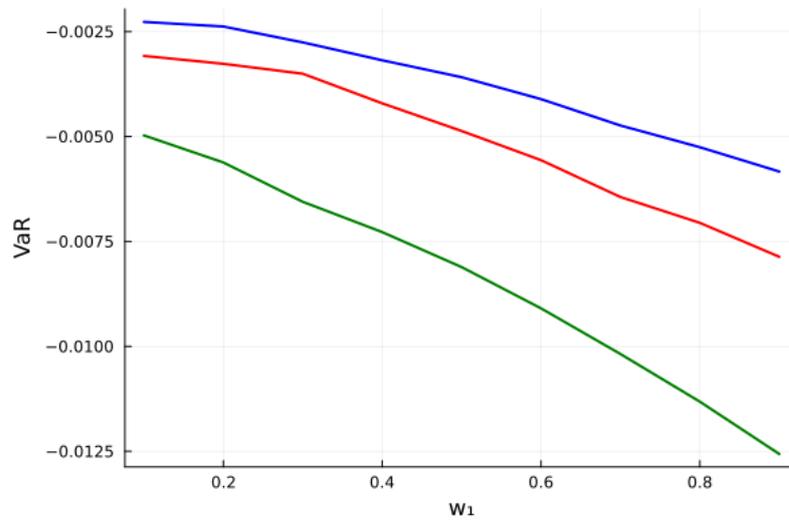
	$\times 10^{-2}$								
w_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$VaR_{90\%}$	-0.23	-0.24	-0.28	-0.32	-0.36	-0.41	-0.47	-0.53	-0.58
$VaR_{95\%}$	-0.31	-0.33	-0.35	-0.42	-0.49	-0.56	-0.64	-0.71	-0.79
$VaR_{99\%}$	-0.50	-0.56	-0.66	-0.73	-0.81	-0.91	-1.02	-1.13	-1.26

Fonte: Elaborado pelo autor (2024).

A Figura 22 mostra o comportamento do VaR estimado utilizando as cópulas Frank para diferentes proporções w_1 e w_2 . As linhas azul, vermelha e verde correspondem aos níveis de confiança de 90%, 95% e 99%, respectivamente.

Após aplicar a metodologia de cópulas para estimar o VaR de portfólios compostos por EUR/USD e BRL/USD, os resultados indicam que, o VaR estimado utilizando a cópula Frank

Figura 22 – VaR estimado com cópula Frank



Fonte: Elaborado pelo autor (2024).

mostram que o risco do portfólio varia de forma suave conforme os pesos w_1 e w_2 . A estrutura simétrica da cópula Frank permite capturar a dependência dinâmica entre EUR/USD e BRL/USD, sem priorizar exclusivamente eventos extremos.

Portfólios balanceados ($w_1 = w_2 = 0.5$) apresentaram menores valores de VaR, enquanto portfólios mais desbalanceados (w_1 próximo de 0.1 ou 0.9) indicaram maior exposição ao risco, especialmente em níveis de confiança mais elevados. Esses resultados destacam a importância de considerar a dependência dinâmica ao avaliar o risco de portfólios compostos por moedas de economias com características distintas.

7 CONCLUSÕES

Este trabalho apresentou uma abordagem para modelar a dependência dinâmica em séries temporais bivariadas utilizando cópulas. Diferente de métodos tradicionais, nos quais o parâmetro de dependência é fixo, aqui $\theta(t)$ foi tratado como uma função do tempo, permitindo capturar mudanças estruturais ao longo do horizonte analisado. A formulação teórica desse modelo foi apresentada no Capítulo 3, enquanto sua validação empírica foi realizada por meio de simulações no Capítulo 5.

Os resultados mostraram que a metodologia foi eficaz na recuperação dos coeficientes verdadeiros, demonstrando a capacidade das aproximações baseadas em polinômios de Bernstein e Chebyshev para estimar a função $\theta(t)$ com precisão. No entanto, observou-se que o desempenho desses métodos variou conforme a cópula utilizada, sendo que a aproximação por polinômios de Taylor apresentou resultados inferiores na maioria dos casos. Esse comportamento, evidenciado nos experimentos de simulação, indica que a escolha da base de funções impacta diretamente a qualidade da estimação, tornando essencial uma seleção criteriosa para capturar adequadamente a dependência dinâmica.

Além disso, a implementação do algoritmo Metropolis-Hastings Adaptativo para a estimação dos coeficientes mostrou-se precisa e bem-sucedida na maioria dos casos, o que permitiu capturar a evolução temporal da dependência entre as séries. No entanto, algumas dificuldades numéricas surgiram em certas cópulas, especialmente quando os valores iniciais escolhidos levavam à geração de matrizes não definidas positivas, o que impedia a obtenção de amostras viáveis.

A aplicação prática do modelo foi explorada no Capítulo 6 por meio da estimação do Valor em Risco (VaR), onde se verificou que considerar dependência dinâmica influencia diretamente na mensuração da medida de risco do portfólio financeiro. A análise empírica evidenciou que a escolha da cópula desempenha um papel crucial na quantificação desse risco, sendo a cópula Frank a mais adequada para modelar a relação entre EUR/USD e BRL/USD, conforme indicado pelos menores valores de AIC e BIC obtidos.

Trabalhos Futuros

Os resultados obtidos neste estudo abriram diversas possibilidades para futuras investigações e aprimoramentos da metodologia proposta. Algumas direções promissoras incluem:

Um primeiro ponto de interesse consiste na utilização de diferentes funções base para a aproximação de $\theta(t)$, como funções kernel, wavelets ou splines. A exploração dessas alternativas pode fornecer maior flexibilidade ao modelo, permitindo capturar padrões de dependência mais complexos.

Outra linha de pesquisa relevante é a implementação de um enfoque frequentista para a estimação da dependência dinâmica, explorando a obtenção de um estimador de máxima verossimilhança para os coeficientes e avaliando suas propriedades assintóticas. Essa abordagem permitiria comparar diretamente os resultados com aqueles obtidos pelo método bayesiano empregado neste trabalho.

Do ponto de vista teórico e metodológico, um estudo funcional mais aprofundado poderia investigar a relação entre as métricas apresentadas no Capítulo 5 para a escolha do grau e do tipo de polinômio. Uma análise detalhada poderia revelar padrões que auxiliem na seleção mais eficiente das funções base, além de explorar possíveis conexões com critérios de informação, como AIC e BIC.

No aspecto computacional, uma questão relevante a ser explorada é a escolha dos valores iniciais no algoritmo Metropolis-Hastings Adaptativo. Atualmente, esses valores são selecionados aleatoriamente, o que ocasionalmente resulta em problemas numéricos. Desenvolver estratégias mais sistemáticas para essa escolha pode melhorar a estabilidade e a eficiência do algoritmo.

Por fim, implementar a metodologia em diferentes classes de ativos e na composição de portfólios mais diversificados. Expandir a análise para ativos além de moedas poderia oferecer uma perspectiva mais ampla sobre a relevância da dependência dinâmica na mensuração de risco.

REFERÊNCIAS

- ABEGAZ, F.; GIJBELS, I.; VERAVERBEKE, N. Semiparametric estimation of conditional copulas. *Journal of Multivariate Analysis*, Elsevier, v. 110, p. 43–73, 2012.
- ANDRIEU, C.; THOMS, J. A tutorial on adaptive mcmc. *Statistics and computing*, Springer, v. 18, p. 343–373, 2008.
- AUSIN, M. C.; LOPES, H. F. Time-varying joint distribution through copulas. *Computational Statistics & Data Analysis*, Elsevier, v. 54, n. 11, p. 2383–2399, 2010.
- BERGER, T. Forecasting value-at-risk using time varying copulas and evt return distributions. *International Economics*, Elsevier, v. 133, p. 93–106, 2013.
- BERNSTEIN, S. Sur les équations du calcul des variations. In: *Annales scientifiques de l'École Normale Supérieure*. [S.l.: s.n.], 1912. v. 29, p. 431–485.
- BOYD, S.; VANDENBERGHE, L. *Convex optimization*. [S.l.]: Cambridge university press, 2004.
- BRAUN, V. *Dynamic copulas for finance: an application to portfolio risk calculation*. [S.l.]: BoD–Books on Demand, 2011. v. 166.
- BROOKS, S.; GELMAN, A.; JONES, G.; MENG, X.-L. *Handbook of markov chain monte carlo*. [S.l.]: CRC press, 2011.
- CECCHETTI, S. G.; SCHOENHOLTZ, K. L. *Money, banking, and financial markets*. [S.l.]: McGraw-Hill, 2017.
- CHERUBINI, U.; MULINACCI, S.; GOBBI, F.; ROMAGNOLI, S. *Dynamic copula methods in finance*. [S.l.]: John Wiley & Sons, 2011.
- COX, R. T. Probability, frequency and reasonable expectation. *American journal of physics*, American Association of Physics Teachers, v. 14, n. 1, p. 1–13, 1946.
- COX, R. T. *Algebra of probable inference*. [S.l.]: Johns Hopkins University Press, 2001.
- DEB, P.; TRIVEDI, P. K.; ZIMMER, D. M. Cost-offsets of prescription drug expenditures: data analysis via a copula-based bivariate dynamic hurdle model. *Health economics*, Wiley Online Library, v. 23, n. 10, p. 1242–1259, 2014.
- DENUIT, M.; DHAENE, J.; GOOVAERTS, M.; KAAS, R. *Actuarial theory for dependent risks: measures, orders and models*. [S.l.]: John Wiley & Sons, 2006.
- EMBRECHTS, P.; FREY, R.; MCNEIL, A. *Quantitative Risk Management*. 2011.
- EMBRECHTS, P.; HÖING, A.; JURI, A. Using copulae to bound the value-at-risk for functions of dependent risks. *Finance and Stochastics*, Springer, v. 7, p. 145–167, 2003.
- EMBRECHTS, P.; KLÜPPELBERG, C.; MIKOSCH, T. *Modelling extremal events: for insurance and finance*. [S.l.]: Springer Science & Business Media, 2013. v. 33.
- EMBRECHTS, P.; LINDSKOG, F.; MCNEIL, A. Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich*, Zurich, v. 14, p. 1–50, 2001.

- FAROUKI, R. T. The bernstein polynomial basis: A centennial retrospective. *Computer Aided Geometric Design*, Elsevier, v. 29, n. 6, p. 379–419, 2012.
- FREEDMAN, D. A. On the asymptotic behavior of bayes estimates in the discrete case ii. *The Annals of Mathematical Statistics*, JSTOR, v. 36, n. 2, p. 454–456, 1965.
- GARTHWAITE, P. H.; FAN, Y.; SISSON, S. A. Adaptive optimal scaling of metropolis–hastings algorithms using the robbins–monro process. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 45, n. 17, p. 5098–5111, 2016.
- GELMAN, A.; GILKS, W. R.; ROBERTS, G. O. Weak convergence and optimal scaling of random walk metropolis algorithms. *The annals of applied probability*, Institute of Mathematical Statistics, v. 7, n. 1, p. 110–120, 1997.
- GELMAN, A.; ROBERTS, G. O.; GILKS, W. R. Efficient metropolis jumping rules. *Bayesian statistics 5*, Oxford University PressOxford, v. 5, p. 599–608, 1996.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, n. 6, p. 721–741, 1984.
- GIACOMINI, E.; HÄRDLE, W. K. Value-at-risk calculations with time varying copulae. SFB 649 Discussion Paper 2005-004, 2005.
- HAARIO, H.; SAKSMAN, E.; TAMMINEN, J. An adaptive metropolis algorithm. *Bernoulli*, JSTOR, p. 223–242, 2001.
- HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. Oxford University Press, 1970.
- HOFF, P. D. *A first course in Bayesian statistical methods*. [S.l.]: Springer, 2009. v. 580.
- HYRŠ, M.; SCHWARZ, J. Elliptical and archimedean copulas in estimation of distribution algorithm. 2015.
- JAMMAZI, R.; TIWARI, A. K.; FERRER, R.; MOYA, P. Time-varying dependence between stock and government bond returns: International evidence with dynamic copulas. *The North American Journal of Economics and Finance*, Elsevier, v. 33, p. 74–93, 2015.
- KREYSZIG, E. *Introductory functional analysis with applications*. [S.l.]: John Wiley & Sons, 1991. v. 17.
- KRUPSKII, P.; JOE, H. Flexible copula models with dynamic dependence and application to financial data. *Econometrics and Statistics*, Elsevier, v. 16, p. 148–167, 2020.
- LANG, S. *Real and functional analysis*. [S.l.]: Springer Science & Business Media, 2012. v. 142.
- LAVERNY, O.; JIMENEZ, S. Copulas.jl: A fully distributions.jl-compliant copula package. *Journal of Open Source Software*, The Open Journal, v. 9, n. 94, p. 6189, 2024. Disponible em: <<https://doi.org/10.21105/joss.06189>>.
- MAI, J.-F.; SCHERER, M. *Financial engineering with copulas explained*. [S.l.]: Springer, 2014.

- MAI, J.-F.; SCHERER, M. *Simulating copulas: stochastic models, sampling algorithms, and applications*. [S.l.]: # N/A, 2017. v. 6.
- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. *The journal of chemical physics*, American Institute of Physics, v. 21, n. 6, p. 1087–1092, 1953.
- METROPOLIS, N.; ULAM, S. The monte carlo method. *Journal of the American statistical association*, Taylor & Francis, v. 44, n. 247, p. 335–341, 1949.
- MOIVRE, A. de. *Miscellanea Analytica de Seriebus et Quadraturis*. [S.l.]: J. Tonson, London, 1730.
- MORETTIN, P. A. *Econometria financeira: um curso em séries temporais financeiras*. [S.l.]: Editora Blucher, 2017.
- MORETTIN, P. A.; TOLOI, C. M. *Análise de séries temporais: modelos lineares univariados*. [S.l.]: Editora Blucher, 2018.
- NATANSON, I. P. *Constructive theory of functions*. [S.l.]: US Atomic Energy Commission, Office of Technical Information Extension, 1961. v. 1.
- NELSEN, R. B. *An introduction to copulas*. [S.l.]: Springer, 2006.
- NELSON, D. B. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the econometric society*, JSTOR, p. 347–370, 1991.
- PATTON, A. Copula methods for forecasting multivariate time series. *Handbook of economic forecasting*, Elsevier, v. 2, p. 899–960, 2013.
- PATTON, A. J. *Applications of copula theory in financial econometrics*. [S.l.]: University of California, San Diego, 2002.
- PATTON, A. J. Estimation of multivariate models for time series of possibly different lengths. *Journal of applied econometrics*, Wiley Online Library, v. 21, n. 2, p. 147–173, 2006.
- PATTON, A. J. Modelling asymmetric exchange rate dependence. *International economic review*, Wiley Online Library, v. 47, n. 2, p. 527–556, 2006.
- RAFTERY, A. E.; LEWIS, S. M. Implementing mcmc. *Markov chain Monte Carlo in practice*, p. 115–130, 1996.
- ROBERTS, G. O.; ROSENTHAL, J. S. Examples of adaptive mcmc. *Journal of computational and graphical statistics*, Taylor & Francis, v. 18, n. 2, p. 349–367, 2009.
- RÜSCHENDORF, L. Fréchet classes, risk bounds, and duality theory. In: *Mathematical Risk Analysis: Dependence, Risk Bounds, Optimal Allocations and Portfolios*. [S.l.]: Springer, 2012. p. 35–51.
- SALVADORI, G.; MICHELE, C. D. On the use of copulas in hydrology: theory and practice. *Journal of Hydrologic Engineering*, American Society of Civil Engineers, v. 12, n. 4, p. 369–380, 2007.
- SAVAGE, L. J. *The foundations of statistics* john wiley and sons new york. 1954.

SAVAGE, L. J. *The foundations of statistics*. [S.l.]: Courier Corporation, 1972.

SHERBERT, D. R.; BARTLE, R. *Introduction to real analysis*. [S.l.]: J. Wiley, 1992.

SKLAR, M. Fonctions de répartition à n dimensions et leurs marges. In: *Annales de l'ISUP*. [S.l.: s.n.], 1959. v. 8, n. 3, p. 229–231.

VIHOLA, M. On the stability and ergodicity of adaptive scaling metropolis algorithms. *Stochastic processes and their applications*, Elsevier, v. 121, n. 12, p. 2839–2860, 2011.

ZHANG, J.; GUEGAN, D. Pricing bivariate option under garch processes with time-varying copula. *Insurance: Mathematics and Economics*, Elsevier, v. 42, n. 3, p. 1095–1103, 2008.

ZHAO, X.; ZHOU, X. Estimation of medical costs by copula models with dynamic change of health status. *Insurance: Mathematics and Economics*, Elsevier, v. 51, n. 2, p. 480–491, 2012.