

基于混合Copula函数的电子商务系统可靠性建模

梅倬铭^{1,2}, 杨剑锋³

¹贵州大学数学与统计学院, 贵州 贵阳

²贵大·贵安科创超级计算算力算法应用实验室, 贵州 贵阳

³南宁师范大学数学与统计学院, 广西 南宁

收稿日期: 2025年9月2日; 录用日期: 2025年9月16日; 发布日期: 2025年10月14日

摘要

随着电子商务系统业务规模的迅猛扩张, 系统在高并发访问、复杂交易流程及突发异常流量冲击下的稳健性面临严峻挑战。电子商务系统可靠性不仅依赖于单一组件的性能, 更受到模块间复杂非线性依赖与联合失效机制的深刻影响。传统可靠性建模方法在刻画尾部相关性、集群失效等极端风险时存在明显局限。针对该问题, 本文提出了一种基于混合Copula函数的电子商务系统可靠性模型, 利用期望最大化算法进行参数估计与优化。该模型中Copula函数用于从依赖结构中分离边缘分布, 实现对模块间关联系统的精准建模。利用蒙特卡洛方法随机生成符合实际业务场景的故障数据, 实验结果表明, 本文提出的模型能够精确地描述模块间的异质依赖结构, 在拟合优度与风险识别能力方面均显著优于单一Copula模型。本研究从方法论层面拓展了系统可靠性建模的理论工具, 为电商平台在高并发促销、异常流量管控等现实场景中的运维策略制定提供了科学依据。

关键词

电子商务系统, 高并发场景, 系统可靠性建模, 混合Copula函数

Reliability Modeling of E-Commerce Systems Based on Mixed Copula Functions

Zhuoming Mei^{1,2}, Jianfeng Yang³

¹School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

²Guizhou University and Gui'an Kechuang Supercomputing Power Algorithm Laboratory, Guiyang Guizhou

³School of Mathematics and Statistics, Nanning Normal University, Nanning Guangxi

Received: September 2, 2025; accepted: September 16, 2025; published: October 14, 2025

文章引用: 梅倬铭, 杨剑锋. 基于混合 Copula 函数的电子商务系统可靠性建模[J]. 电子商务评论, 2025, 14(10): 882-895. DOI: 10.12677/ecl.2025.14103220

Abstract

With the rapid expansion of e-commerce system operations, the robustness of these systems under high-concurrency access, complex transaction processes, and sudden abnormal traffic surges faces severe challenges. The reliability of e-commerce systems depends not only on the performance of individual components but is also profoundly influenced by complex nonlinear dependencies and joint failure mechanisms among modules. Traditional reliability modeling methods exhibit significant limitations in capturing extreme risks such as tail dependence and cluster failures. To address this issue, this paper proposes a reliability model for e-commerce systems based on mixed Copula functions, utilizing the Expectation-Maximization algorithm for parameter estimation and optimization. In this model, Copula functions are employed to separate marginal distributions from the dependency structure, enabling precise modeling of the interdependencies among modules. Using the Monte Carlo method to randomly generate failure data that aligns with real-world business scenarios, experimental results demonstrate that the proposed model accurately captures the heterogeneous dependency structures among modules and significantly outperforms single Copula models in terms of goodness-of-fit and risk identification capabilities. This study expands the theoretical toolkit for system reliability modeling from a methodological perspective and provides a scientific basis for formulating operational strategies in real-world scenarios such as high-concurrency promotions and abnormal traffic management in e-commerce platforms.

Keywords

E-Commerce Systems, High-Concurrency Scenarios, System Reliability Modeling, Mixed Copula Functions

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着数字经济的快速发展,电子商务平台已成为连接消费者与商家、推动线上交易与供应链协同的核心基础设施。近年来,尤其是在“双十一”“6·18”等大型促销活动期间,平台需在高并发访问、复杂交易逻辑与流量激增的背景下维持稳定运行。系统整体可靠性不仅依赖单个模块,还受模块间非线性相关与联合失效机制的深刻影响。因此,实现对电商平台内部依赖结构的精准建模与评估,对保障交易安全、用户体验与平台信誉具有核心意义。

目前,电子商务系统的可靠性建模主要分为分析法、仿真法与混合方法[1]。常用工具包括故障树分析[2]、马尔可夫模型和贝叶斯网络[3]等。然而,这类方法通常假设组件故障独立,或着重于线性转移机制,难以准确反映多源组件间的复杂依赖关系。针对上述挑战,Copula函数作为一种有效工具,应运而生。它能将边缘分布与依赖结构分离,从而分别拟合模块自身特性与模块间关联系统结构,显著增强模型的灵活性与表达能力。

为应对此类依赖结构的复杂性,Copula函数被引入可靠性分析领域。Copula能够将边缘分布与依赖结构有效分离,并根据Sklar定理构建联合分布,从而灵活选用不同族函数来描述多变量之间的相关性。已有大量研究成功应用Copula方法改善传统模型的性能。Li等[4]提出了针对阶段任务系统(PMS)的Copula模型,能够同时刻画组件在不同阶段的依赖结构变化,并选择最优Copula函数,显著提高了可靠

性分析精度。Jia 等[5]针对安全关键系统, 构建三态可靠性模型, 并利用 Copula 描述非线性相关性, 提出计算安全失效概率的公式。Teng 等[6]在多组件三态系统 Copula 模型下构建工作、安全失效、危险失效三态系统的可靠性索引表达式, 为 CPS 多状态系统建模提供参考。Xiong [4]等提出了多阶段任务系统可靠性分析的 Copula 模型, 通过建模不同任务阶段间组件失效的相关性, 突破了传统独立性假设的局限。Qie [7]等将 Copula 理论引入海洋工程土壤参数依赖关系建模, 通过 D-vine Copula 函数建立多变量联合分布, 考虑土壤参数空间变异性, 为海洋工程结构地震可靠性分析提供了新框架。Fang [8]等采用 Copula 函数建模双变量退化过程依赖关系, 结合汉密尔顿蒙特卡洛算法实现复杂退化系统可靠性评估。Shi [9]等将神经网络与 Copula 函数结合, 提高了数据分布描述精度和可靠性区间求解准确性。

Copula 函数在电力系统、金融风险管理、结构可靠性等复杂系统中取得了良好效果。Myron [10]等建立基于可靠性框图与马尔可夫建模的电商与互联网信息系统模型, 分析了系统可用性与故障预测的可行路径。Mitrevski [11]研究评价了电子商务系统中服务性能与可靠性之间的关系, 并提出行为建模与可执行性分析方法来提高系统可管理性。Nabil [12]等在跨境电商供应链领域, 通过构建 Entropy Weight、SAW 与 ISM 混合模型, 识别出供应链中“恢复速度”“协作效率”等关键因素, 并据此提出策略提升系统韧性, 为系统可靠性建模提供了参考路径。Xia [13]等以 Copula 构建机械齿轮系统的可靠性分析模型, 采用粒子群算法与最大似然估计实现参数与权重估计, 案例验证表明其在样本量小且结构复杂的系统中具有高效与准确性。Wang [14]等将 Copula 应用于具有多失效模式的救援装置可靠性分析, 并构建时变依赖模型, 显著提升了系统建模的现实适应性。Erem [15]等研究通过对仿真微型逆变器系统应用多种 Copula 族函数, 展示了在多状态系统中依赖对可靠性影响的重要性。Zhou [13]等构建多种 Copula 族函数的混合 Copula 模型, 采用改进粒子群算法优化权重系数。

在此基础上, 本文将重点放在电子商务平台自身的运行特性, 利用仿真方法生成核心模块(交易处理、支付结算、推荐服务)在高并发条件下的故障数据, 并引入混合 Copula 模型对其进行系统可靠性建模。与传统案例不同, 本研究直接聚焦于电商环境, 通过模型参数估计和指标分析, 揭示关键模块在不同业务场景下的依赖关系与联合失效特征, 为平台架构优化与风险管理提供了新的量化工具。

2. 基于混合 Copula 函数的电子商务系统可靠性模型

在电子商务平台中, 核心模块(如交易处理、支付结算、推荐服务)通过网络与业务逻辑紧密交互, 形成高度耦合的系统结构。系统的整体可靠性不仅取决于单个模块的运行寿命与稳定性, 还受到模块之间非线性依赖关系的显著影响, 特别是在高并发与极端负载条件下, 常会出现交易失败与支付延迟等联合失效行为。传统方法虽然能够描述功能路径或系统状态转移, 但往往难以精确刻画电商平台中存在的尾部依赖与故障共爆现象[16]。为此, 本文引入混合 Copula 模型, 以增强对多源失效事件的联合概率建模能力, 更真实地反映电子商务平台在复杂依赖结构下的可靠性特征。

2.1. Copula 函数

Copula 函数作为一种概率分布函数, 能够有效划分多个随机变量间的依赖关系与其边缘分布。Copula 函数源自于 Sklar 在 1959 年提出的 Sklar 定理[17], 其定义如下: 对于一组随机变量 (x_i, x_j) 的联合分布 $H(x_i, x_j)$, 如果他们的边缘分布函数已知且分别为 F_1 和 F_2 , 则必然存在一个 Copula 函数 C , 使得:

$$H(x_i, x_j) = C(F_1(x_i), F_2(x_j)) \quad (1)$$

在上式中, 若其边缘分布函数 F_1 和 F_2 是连续的, 则联合分布具有唯一确定的 Copula 函数 C 。

设 $F(x)$ 是一个分布函数, 如果该分布函数的逆函数存在, 那么随机变量 $U = F(x)$ 服从 $[0, 1]$ 上的均匀分布, 有:

$$C(u, v) = H(F_1^{-1}(u), F_2^{-1}(v)) \quad (2)$$

其中 $C(u, v)$ 是随机变量 (u, v) 的一个 Copula 函数, F^{-1} 是 F 的逆函数。

由此可以得出, 对于高维随机变量 $(x_1, x_2, \dots, x_n)$ 的联合分布进行建模, 若其边缘分布 $F_1, F_2, \dots, F_n$ 均连续, 则存在唯一的 Copula 函数 C , 使得:

$$H(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)) \quad (3)$$

$$C(u_1, u_2, \dots, u_n) = H(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_n^{-1}(u_n)) \quad (4)$$

这一思想在电子商务平台的可靠性分析中尤为适用。不同模块的故障时间往往服从不同分布形式: 例如, 交易处理模块在高并发下更接近指数或 Gamma 分布, 支付结算模块可能呈现 Weibull 分布特征, 而推荐服务模块则表现出一定的长尾性。利用 Copula 函数可以先对每个模块的边际分布进行独立建模, 再通过依赖结构统一刻画其相关性, 从而避免强行假设所有模块服从相同分布的局限。

目前常见的 Copula 族函数包括 Clayton、Frank、Gumbel 等。假设 u, v 为两个存在关联的变量, θ 为 Copula 模型中的参数, 则相关族函数的具体表达式如表 1 所示。

Table 1. Common copula families functions

表 1. 常见的 Copula 族函数

Copula	$C(u, v \theta)$	$\theta \in R$
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$	$[0, \infty]$
Gumbel	$\exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\}$	$[1, \infty]$
Frank	$-\frac{1}{\theta} \ln \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right]$	$R \setminus \{0\}$
FGM	$uv + \theta uv(1-u)(1-v)$	$[-1, 1]$
AMH	$\frac{uv}{1 - \theta(1-u)(1-v)}$	$[-1, 1]$

2.2. 混合 Copula 函数模型构建

尽管单一 Copula 函数可能难以全面反映电商模块间的复杂依赖关系, 本文进一步构建混合 Copula 模型 $C_{mix}(\cdot)$ [18] 如下:

$$C_{mix}(F_1(x_1), F_2(x_2), \dots, F_n(x_n)) = \sum_{k=1}^n \lambda_k C_k(F_1(x_1), F_2(x_2), \dots, F_n(x_n); \theta_k) \quad (5)$$

其中 n 为 Copula 族函数数量, $C_k(\cdot)$ 为第 k 个 Copula 族函数, θ_k 为第 k 个 Copula 族函数的参数向量, λ_k 为第 k 个 Copula 族函数的权重参数, 满足 $0 \leq \lambda_k \leq 1, \sum_{k=1}^n \lambda_k = 1$ 。通过这种方式, 可以将不同 Copula 家族的优势结合起来, 更准确地刻画不同模块之间的尾部依赖和联合失效机制。

2.3. 混合 Copula 函数模型选择

对于表 1 中的常用 Copula 族函数, 并结合现有资料[19]。本文选择 Frank、Gumbel、Clayton 三种 Copula 函数构建混合模型, 原因如下:

(1) Clayton 适合描述下尾相关, 对应于电商场景中“低负载情况下的联合失效”, 如网络波动或轻量级异常导致交易处理模块与支付结算模块同时失效;

(2) Gumbel 擅长刻画上尾依赖, 适用于“极端高负载或秒杀场景”, 此时多个模块在峰值压力下同时表现退化;

(3) Frank 能捕捉对称依赖结构, 适合电商系统日常运行中模块间的平衡交互(如推荐服务模块与交易处理模块在常规流量下的协同)。

混合 Copula 模型通过结合上述三种依赖结构, 能够同时捕捉不同场景下的风险模式, 为刻画电子商务平台关键模块(交易处理、支付结算、推荐服务)在高并发与极端负载下的联合失效与尾部依赖, 本文采用由 Frank、Gumbel、Clayton 组成的高维混合 Copula 函数模型:

$$C_{mix}(F_1(x_1), F_2(x_2)) = \lambda_F C_F(F_1(x_1), F_2(x_2); \theta_F) + \lambda_G C_G(F_1(x_1), F_2(x_2); \theta_G) + \lambda_C C_C(F_1(x_1), F_2(x_2); \theta_C) \quad (6)$$

2.4. 混合 Copula 函数参数估计

混合 Copula 模型的参数估计采用两阶段方法: 第一阶段确定各模块边际分布, 第二阶段估计 Copula 函数的混合权重及依赖参数。

2.4.1. 边缘分布的非参数估计

在电商平台应用中, 不同模块(交易处理、支付结算、推荐服务)的故障数据往往分布差异显著, 难以事先假定具体分布形式。为此, 本文采用核密度估计方法对边际分布进行建模, 某随机变量 X 的边缘分布概率密度函数表达式如下:

$$f(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (7)$$

其中 n 表示观测数据对数; h 为平滑参数; $K(\cdot)$ 为满足正则条件的核函数; X_i 代表第 i 对观测数据值。

该方法通过平滑观测数据, 以灵活逼近真实分布, 避免因分布假设偏差带来的误差。由此得到的边际分布函数如下:

$$F(x) = \int_{-\infty}^x f(t) dt \quad (8)$$

2.4.2. 混合 Copula 函数的参数估计

在获得边际分布之后, 需要进一步估计混合 Copula 模型中的依赖参数与权重。本文采用分层优化策略进行估计: 权重参数通过贝叶斯模型平均方法确定[20], 而相依参数则采用最小化经验 - 模型差异准则进行估计[21]。

根据贝叶斯模型加权平均法, 对于二元混合 Copula 函数, 随机变量 X_2 已知时, 随机变量 X_1 的概率密度函数如下:

$$f(x_1|x_2) = \lambda_F c_F(F_1(x_1), F_2(x_2)) f(x_1) + \lambda_G c_G(F_1(x_1), F_2(x_2)) f(x_1) + \lambda_C c_C(F_1(x_1), F_2(x_2)) f(x_1) \quad (9)$$

其中 c_F, c_G, c_C 分别为二元 Frank、Gumbel、Clayton Copula 函数的概率密度函数。

由式(9), 可得关于二元混合 Copula 函数权重参数的对数似然函数:

$$\begin{aligned} L(\lambda_F, \lambda_G, \lambda_C) &= \log(f(x_1|x_2)) \\ &= \log\left\{\left[\lambda_F c_F(F_1(x_1), F_2(x_2)) + \lambda_G c_G(F_1(x_1), F_2(x_2))\right.\right. \\ &\quad \left.\left.+ \lambda_C c_C(F_1(x_1), F_2(x_2))\right] f(x_1)\right\} \end{aligned} \quad (10)$$

针对式(10)中权重参数 λ 的对数似然函数求解复杂问题。本文采用最大期望算法(Expectation-maximization Algorithm, EM)。该算法通过交替执行以下两个步骤实现参数优化:

(1) E 步(期望计算): 构建所有观测数据的的对数似然函数条件期望

$$Q(\lambda|\lambda^{(m)}) = E_{\lambda^{(m)}}[\log L(\lambda_F, \lambda_G, \lambda_C)|X_1, X_2] \quad (11)$$

(2) M 步(最大化): 通过对上一步中的权重参数进行求解 $\lambda_k^{(m+1)} = \arg \max_{\lambda} Q(\lambda|\lambda^{(m)})$, 可得权重更新公式

$$\lambda_{m+1} = \frac{1}{n} \sum_{i=1}^n \frac{\lambda_m c(F_1(X_{1i}), F_2(X_{2i})) f(X_{1i})}{\sum \lambda_m c(F_1(X_{1i}), F_2(X_{2i})) f(X_{1i})} \quad (12)$$

其中 λ_{m+1} 为第 $(m+1)$ 次迭代的权重参数; X_{1i}, X_{2i} 分别为随机变量 X_1, X_2 的样本值; $m(m=0, 1, 2, \dots)$ 为迭代次数。

根据最小化经验-模型差异准则, 定义目标混合 Copula 函数的相依参数应使得二元经验 Copula 函数 $\hat{C}(\cdot)$ 与混合 Copula 函数 $C_{mix}(\cdot)$ 的加权均方根偏差最小, 即:

$$O_{LS} = \sqrt{\frac{1}{n} \sum_{i=1}^n \{\hat{C}[F_1(X_{1i}), F_2(X_{2i})] - C_M[F_1(X_{1i}), F_2(X_{2i})]\}^2} \quad (13)$$

其中 $\hat{C}(\cdot)$ 为二元经验 Copula 函数, 其定义如下:

$$\hat{C}[F_1(X_{1i}), F_2(X_{2i})] = \frac{1}{n} \sum_{i=1}^n I_{[\hat{F}_1(X_{1i}) \leq F_1(X_{1i})]} I_{[\hat{F}_2(X_{2i}) \leq F_2(X_{2i})]} \quad (14)$$

其中 $I_{[\cdot]}$ 为示性函数, 其取值为 $I_{[\cdot]} = \begin{cases} 1, \hat{F}(x) \leq F(x) \\ 0, \hat{F}(x) > F(x) \end{cases}$ 。

为提高迭代过程的稳定性与避免陷入局部最优, 本文在权重更新过程中引入“温度系数平滑”策略[22]。其原理源于模拟退火思想, 通过在 E 步中对后验概率进行温度平滑处理, 使分布更趋均匀, 从而增加搜索空间覆盖性, 减少早期收敛引入的偏差。通过对权重进行如下变换:

$$\omega_k^{(t)} = \frac{(\omega_k^{(t)})^{\frac{1}{T}}}{\sum_{j=1}^k (\omega_j^{(t)})^{\frac{1}{T}}} \quad (15)$$

当 $T > 1$ 时, 权重分布趋于均衡, 有助于防止某一族函数过早占据主导地位; 当迭代逐步收敛时, 权重分布逐渐恢复真实结构。在此基础上本文选取 $T = 5$ 作为初始温度[23], 既兼顾了早期搜索的广度, 也能通过预实验验证该温度设置在电商仿真数据中的收敛速度与模型稳定性表现最佳。

3. 案例分析

为验证所提出的混合 Copula 模型的有效性, 本文在缺乏可获取的真实电商平台运行数据的情况下,

采用仿真方法生成故障样本数据。由于电商系统运行数据往往涉及用户隐私与商业机密,难以直接获得,因此本文参考了公开的电商平台运行指标(如高并发流量分布、交易响应时间统计特征等) [24] [25]。结合典型的业务场景(如秒杀活动、支付高峰期)来参数化仿真过程。具体而言,在数据生成过程中,设定交易处理模块在高并发条件下更易失效,支付结算模块受交易压力的联动影响较大,而推荐服务模块表现出一定的长尾失效特征。为增强实验的透明度与可复现性,仿真规则如下:当交易处理模块负载过高时,支付结算模块失效概率增加 30%,推荐服务模块失效概率增加 10%;当支付模块处于高负载时,交易处理模块失效概率增加 20%;各模块的边缘分布形态依据前期系统研究结果与公开指标确定:交易处理与推荐模块服从 Gamma 分布,支付模块服从指数分布。基于以上仿真设定,共生成 40 组样本数据如表 2 所示,该过程既保证了数据来源的合理性,也提高了实验结果的可信度和可复现性。

Table 2. E-commerce system simulated failure data
表 2. 电子商务系统仿真故障数据

t	F1	F2	F3	t	F1	F2	F3	t	F1	F2	F3
1	23	75	67	15	5	42	49	29	40	16	0
2	140	68	146	16	21	22	7	30	1	21	14
3	74	77	80	17	22	41	41	31	36	11	51
4	33	43	76	18	20	5	77	32	3	64	26
5	11	9	39	19	38	136	51	33	11	74	13
6	11	13	67	20	13	63	117	34	195	116	60
7	10	23	58	21	41	69	70	35	168	19	64
8	103	78	40	22	16	130	32	36	93	38	31
9	53	32	153	23	28	68	10	37	22	43	90
10	63	85	68	24	32	12	74	38	1	43	99
11	0	104	15	25	26	104	42	39	61	9	67
12	172	40	97	26	74	43	75	40	44	36	65
13	75	39	125	27	14	60	96				
14	5	58	45	28	55	104	55				

3.1. 混合 Copula 函数参数估计

在对电商平台各模块进行建模时,需要对模型效果进行合理评价。本文选择均方误差(MSE)、决定系数(R^2)和赤池信息准则(AIC)作为核心评估准则。

均方误差(MSE)是一个在多个领域中广泛应用的统计量,用于衡量模型预测值与真实值之间的平均平方偏差,其计算公式为:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{9}$$

其中, n 是样本数量, y_i 为第 i 个观测值, $\hat{y}_i$ 为第 i 个观测值的预测值。MSE 越小, 说明模型拟合精度越高。

决定系数(R^2)用于衡量模型解释变量方差的能力, 其计算公式为:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

其中, $\bar{y}$ 是观测值的平均值。 R^2 越接近 1, 表示模型解释能力越强。

赤池信息准则(AIC)是一种衡量统计模型拟合优良性的标准, 其计算公式为:

$$AIC = 2K - 2\ln(L) \quad (11)$$

其中, k 为模型参数个数, L 为模型的极大似然估计值。AIC 越小表示模型在精度与简洁度上表现更优。

这三个指标在本研究中共同用于比较单一 Copula 模型与混合模型的优劣, 为后续结果分析提供依据。

3.2. 模块边缘分布拟合与依赖结构

采用核密度估计与参数分布拟合, 对各模块的故障数据进行边缘分布建模, 选择最优分布及参数, 结果如表 3 所示。

Table 3. Fitting results of marginal distributions for each module

表 3. 各模块边缘分布拟合结果

模块	最优分布	参数估计结果	AIC
交易处理模块	Gamma 分布	$k = 0.819, loc = -5.01e - 27, \theta = 57.841$	371.07
支付结算模块	指数分布	$loc = 5.000, \theta = 48.325$	394.24
推荐服务模块	Gamma 分布	$k = 7.885, loc = -39.440, \theta = 12.776$	402.43

边缘分布的拟合结果表明, Gamma 分布能够较好地描述交易处理模块和推荐数据模块的失效特征, 而指数分布适合建模支付结算模块的失效规律。这一结果与电子商务系统各组件的实际运行特征相符。

通过计算各模块之间的伪观测计算值 Kendall's tau, 可得各模块间的依赖方向和关系, 结果如表 4 所示。

Table 4. Dependency relationships among modules

表 4. 各模块依赖关系

模块对	Kendall's tau	依赖方向	依赖强度
交易处理 - 支付结算	0.0571	正向依赖	弱
交易处理 - 推荐服务	0.2208	正向依赖	中等
支付结算 - 推荐服务	-0.0531	负向依赖	弱

结果显示, 交易处理模块与推荐推荐模块之间存在较强的正相关性, 表明在高并发条件下, 用户的行为链路(“浏览 - 推荐 - 下单 - 支付”)会同时对两个模块施加压力, 进而导致其性能联动退化, 提示平台在促销场景中需要对推荐与交易模块进行联合容量规划, 例如通过在高峰期引入弹性扩容或缓存机制来缓解联动风险。而支付结算模块与推荐服务模块之间则表现为弱负相关, 可能源于二者在资源占用上存在一定的竞争关系, 如数据库连接或通信通道, 提示运维人员在系统架构设计时, 应避免推荐与支付服务过度共享关键资源, 必要时可通过服务解耦或冗余资源分配来降低潜在冲突风险。

此外, 支付模块与交易模块虽然相关性较弱, 但仍在高负载条件下表现出一定的联动失效趋势。针对这一现象, 平台可在交易压力上升时, 预先对支付模块进行性能预警与优先资源调度, 以避免“交易完成却支付延迟”的关键用户体验风险。

3.3. 混合 Copula 模型参数估计与模型评价

为进一步刻画依赖结构, 采用 Clayton、Gumbel 和 Frank 三种 Copula 族函数及其混合模型 Mixed 对电子商务系统故障数据进行拟合。其中混合模型 Mixed 采用基于似然值的权重分配方法, 通过温度系数平滑($T = 5$)进行参数估计。各模型的参数估计结果和模型评价指标如表 5 所示。

Table 5. Parameter estimation and evaluation metrics for copula models
表 5. Copula 模型参数估计与评价指标

Copula 函数类型	权重 ω	参数估计结果 θ	MSE	R^2	AIC
Frank	-	$\theta_F = 0.2997$	0.001463	0.971091	-177.28
Gumbel	-	$\theta_G = 1.1480$	0.000865	0.982907	-180.56
Clayton	-	$\theta_C = 0.2626$	0.000977	0.980699	-177.17
Mixed	$\omega_F = 0.0263$	$\theta_F = 0.2098$	0.000735	0.986108	-188.83
	$\omega_G = 0.2072$	$\theta_G = 2.3036$			
	$\omega_C = 0.7665$	$\theta_C = 0.1868$			

从表 5 内容可知, 混合 Copula 模型在权重分配方面, Clayton 族函数在描述系统依赖结构中占主导地位, 反映了电子商务系统组件间主要表现为下尾相关的特征; 且在 MSE、 R^2 和 AIC 上均优于单一 Copula 模型, 在平衡拟合优度和模型复杂度方面表现最佳。

为了更直观地展示各 Copula 模型的拟合效果, 图 1~4 分别展示了 Frank、Gumbel、Clayton 三种单一 Copula 函数以及混合 Copula 模型的三维拟合结果。

图 1~3 分别显示了该模型在描述电商系统模块对对称依赖结构、上尾依赖关系和下尾相关性方面的特征。

(1) Frank 显示其对对称依赖关系的刻画能力有限, 在日常平稳流量下, 单一模型可能低估模块之间的协同压力。对于企业运维层面应避免依赖单一监控指标, 而应在日常运行中综合监控不同模块的耦合趋势;

(2) Gumbel 凸显出上尾依赖特征, 意味着在促销时模块可能同步退化。对于企业运维上需在高峰活动前进行容量冗余规划, 为交易支付模块预留额外资源池, 并在秒杀场景中提前启用弹性扩容策略;

(3) Clayton 主要捕捉下尾相关性, 表明在低负载异常场景下, 某些模块可能因共用资源或架构瓶颈而同时失效。对于企业运维需在非高峰时段也要保持故障注入与压测, 防止“低负载下的隐性耦合风险”被忽视。

图 4 中混合模型整合了 Frank、Gumbel、Clayton 的优势, 全面刻画了电商系统的多样依赖关系。不能仅针对单一业务场景(如高峰期), 而应构建分层次、全场景的监控体系, 展示了通过融合多种 Copula 函数的优势, 显示了混合模型的优越性: 混合模型能够更准确地描述电商平台复杂的依赖结构。右图的相对误差收敛曲线显示, 混合模型具有更快的收敛速度和更低的最终误差水平。

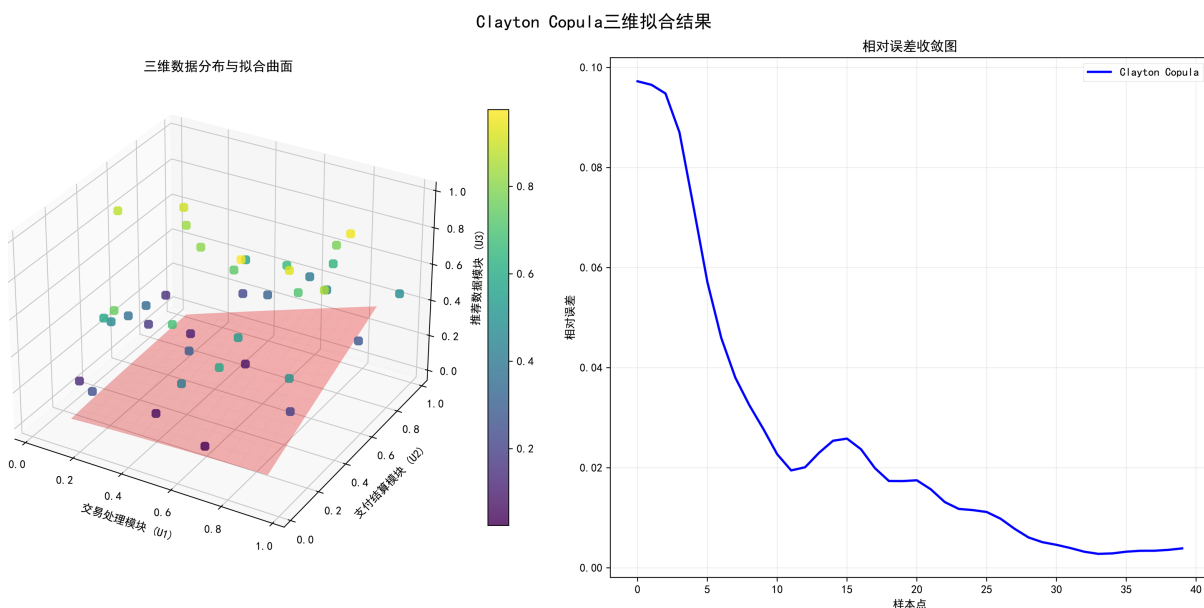


Figure 1. Fitting results of the 3-Dimensional Frank Copula

图 1. Frank Copula 三维拟合效果

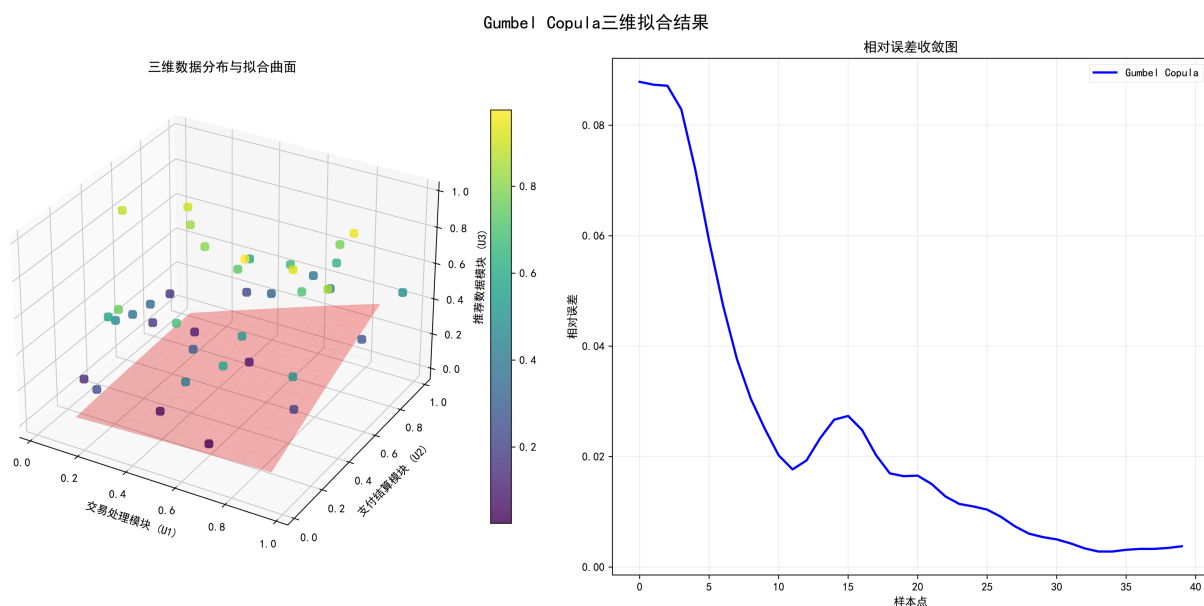


Figure 2. Fitting results of the 3-Dimensional Gumbel Copula

图 2. Gumbel Copula 三维拟合效果

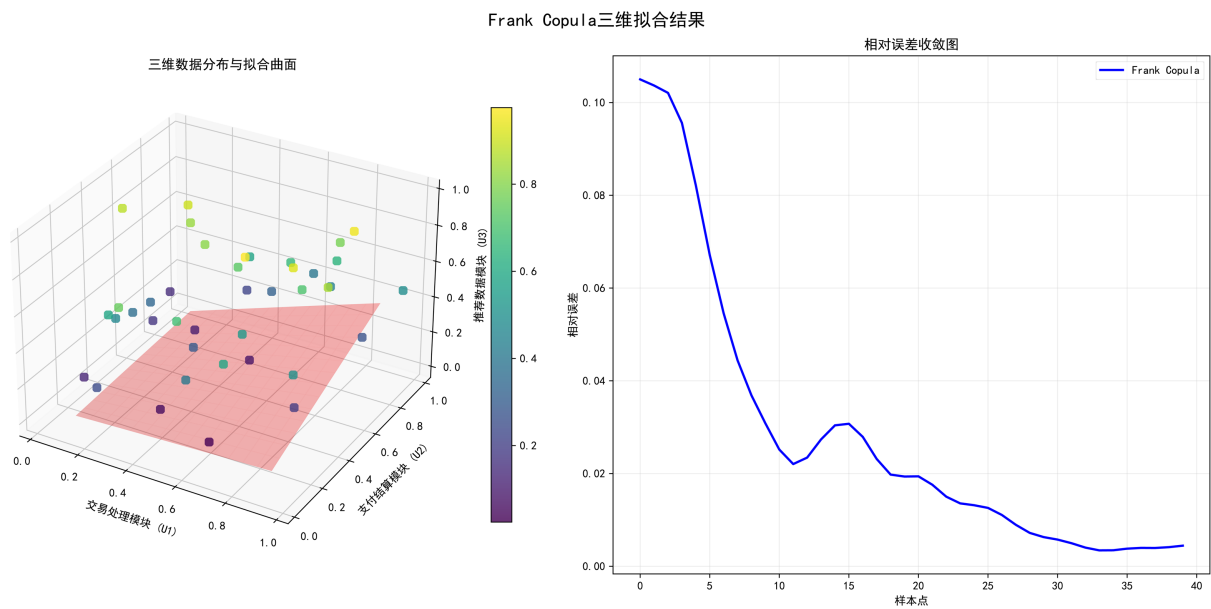


Figure 3. Fitting results of the 3-Dimensional Clayton Copula
图 3. Clayton Copula 三维拟合效果

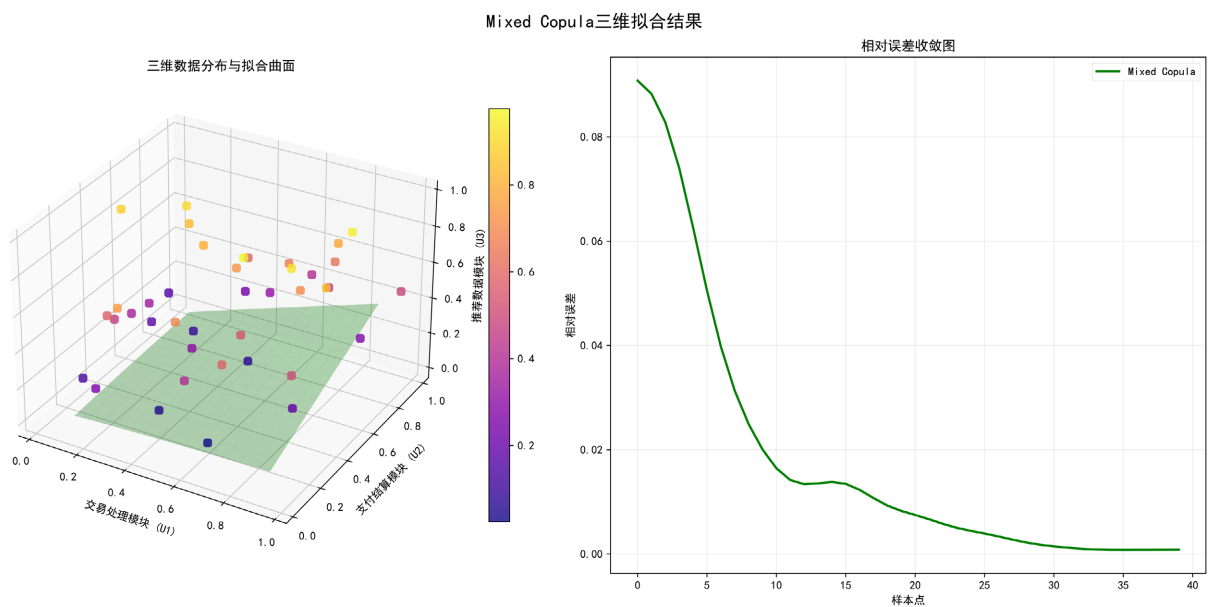


Figure 4. Fitting results of the 3-Dimensional Mixed Copula
图 4. Mixed Copula 三维拟合效果

图 5 展示了电子商务系统可靠性建模的完整过程:

- (1) 左上原始数据分布子图: 可以看出业务数据本身存在明显的异质性, 提示企业运维应通过模块化监控而非统一阈值来设定告警标准;
- (2) 右上伪观测值变换分布子图: 经过标准化后的故障数据可以更准确识别模块间依赖, 企业运维上可借鉴此思路, 将多维监控指标进行归一化, 提升异常检测的敏感性;
- (3) 左下混合 Copula 理论分布曲面子图: 显示了模型预测与真实数据的吻合程度, 企业运维可基于该模型构建联合失效预警系统, 提前识别多模块共退化的风险;

(4) 右下不同 Copula 模型对比子图：直观展示了混合模型的优势，提示企业在运维策略上，应优先采用能兼顾多种依赖特性的综合风险指标取代单模块指标，用于高并发运维决策。

通过这一系列可视化分析可以看出，混合 Copula 模型不仅在数值指标上优于单一模型，在可视化效果上也更好地反映了电商平台各模块间的复杂依赖关系，特别是在高并发和极端负载条件下的联合失效特征。

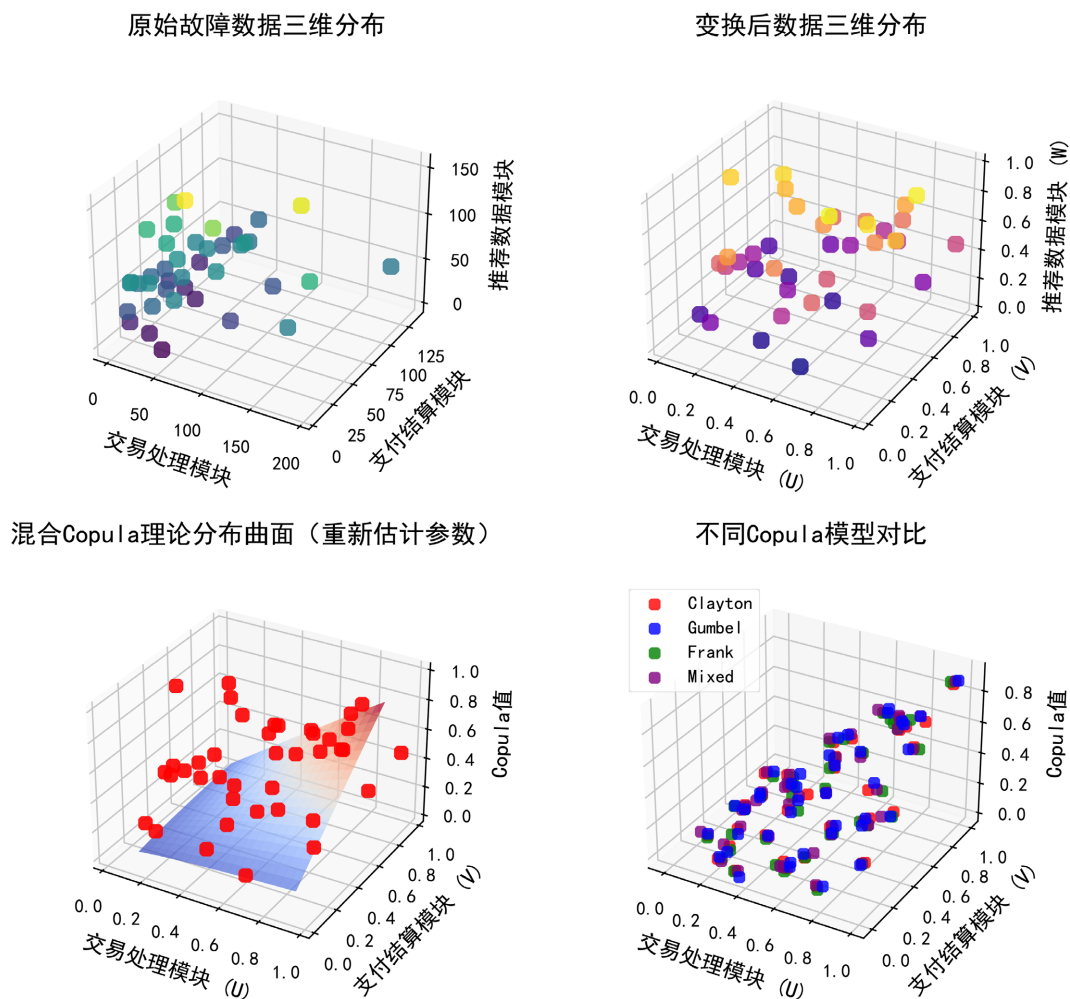


Figure 5. Comprehensive comparison of 3-Dimensional data distributions

图 5. 三维数据分布综合对比

4. 结语

总体而言，本文提出的混合 Copula 模型为电商系统可靠性研究提供了一种探索性建模方法。基于公开的电商平台运行指标进行仿真验证，模型能够较为准确地刻画关键模块间的非线性依赖结构，并为系统架构优化与风险管理提供新的分析视角。然而，本研究仍存在局限性：其一，实验数据主要来自仿真，未能完全反映真实电商系统的复杂环境；其二，模型仅选择了三个核心模块，未覆盖如存储、通信等其他重要电商平台中的重要模块；其三，未考虑动态依赖关系随时间演化的影响。因此，本文的结论应被理解为探索性结果，而非最终定论。未来工作可在获取真实业务数据、扩展多模块建模以及引入动态 Copula 方法等方向进一步深化研究。

基金项目

国家自然科学基金项目(编号: 72361008), 广西自然科学基金面上项目(编号: 2025GXNSFAA069686), 贵州省科技计划项目(编号: 黔科合支撑[2023]一般 268)。

参考文献

- [1] Huo, R., Zeng, S., Wang, Z., Shang, J., Chen, W., Huang, T., *et al.* (2022) A Comprehensive Survey on Blockchain in Industrial Internet of Things: Motivations, Research Progresses, and Future Challenges. *IEEE Communications Surveys & Tutorials*, **24**, 88-122. <https://doi.org/10.1109/comst.2022.3141490>
- [2] Uddin, M., Mo, H. and Dong, D. (2025) A Review on Modelling, Evaluation, and Optimization of Cyber-Physical System Reliability. arXiv: 2503.10924.
- [3] Kabir, S. (2017) An Overview of Fault Tree Analysis and Its Application in Model Based Dependability Analysis. *Expert Systems with Applications*, **77**, 114-135. <https://doi.org/10.1016/j.eswa.2017.01.058>
- [4] Li, X., Zhan, K., Xiong, X. and Vu, T.Q. (2023) A Copula-Based Reliability Model for Phased Mission Systems with Dependent Components. *Quality and Reliability Engineering International*, **39**, 1533-1547. <https://doi.org/10.1002/qre.3342>
- [5] Jia, X., Xing, L. and Li, G. (2018) Copula-based Reliability and Safety Analysis of Safety-Critical Systems with Dependent Failures. *Quality and Reliability Engineering International*, **34**, 928-938. <https://doi.org/10.1002/qre.2301>
- [6] Teng, F., Wang, X., Li, Y. and Hou, W. (2020) A Reliability Analysis Method of Multiple Components System Based on Copula Function. In: Xu, Z., Parizi, R., Hammoudeh, M. and Loyola-González, O., Eds., *Cyber Security Intelligence and Analytics*, Springer, 172-179. https://doi.org/10.1007/978-3-030-43306-2_25
- [7] Qie, X., Li, X., Tao, X., Xu, S. and Huang, L. (2025) Seismic Reliability Analysis of Reinforced Slope Considering Soil Parameter Dependence Structure. *Scientific Reports*, **15**, Article No. 8908. <https://doi.org/10.1038/s41598-025-92789-w>
- [8] Fang, G., Pan, R. and Hong, Y. (2020) Copula-Based Reliability Analysis of Degrading Systems with Dependent Failures. *Reliability Engineering & System Safety*, **193**, Article ID: 106618. <https://doi.org/10.1016/j.ress.2019.106618>
- [9] Yang, R., Sun, L., Li, H. and Yang, Y. (2024) A Neural Network Copula Function Approach for Solving Joint Basic Probability Assignment in Structural Reliability Analysis. *Quality and Reliability Engineering International*, **40**, 3096-3119. <https://doi.org/10.1002/qre.3568>
- [10] Hecht, M. (2001) Reliability/availability Modeling and Prediction for E-Commerce and Other Internet Information Systems. Annual Reliability and Maintainability Symposium. 2001 *Proceedings. International Symposium on Product Quality and Integrity* (Cat. No.01CH37179), Philadelphia, 22-25 January 2001, 176-182. <https://doi.org/10.1109/rams.2001.902463>
- [11] Mitrevski, P.J. and Hristoski, I.S. (2014) Behavioral-Based Performability Modeling and Evaluation of E-Commerce Systems. *Electronic Commerce Research and Applications*, **13**, 320-340. <https://doi.org/10.1016/j.elerap.2014.06.005>
- [12] Nabil, D.H., Al Amin, M. and Baldacci, R. (2024) Enhancing Resilience in Transnational E-Commerce Supply Chains: Critical Factors, Perspectives and Strategic Action Plan. *Heliyon*, **10**, e31274. <https://doi.org/10.1016/j.heliyon.2024.e31274>
- [13] ErDong, X., FuPing, Z., Kun-Chieh, W., ChunRong, W. and Hao, G. (2024) A Novel Reliability Analysis Methodology Based on IPSO-MCopula Model for Gears with Multiple Failure Modes. *Advances in Mechanical Engineering*, **16**, 1-13. <https://doi.org/10.1177/16878132241228194>
- [14] Wang, C., Xia, E., Wang, K. and Zhao, J. (2023) A Mixed-Copula-Based Integral Method for Reliability Analysis of a Novel Multi-Functional Rescue End-Effector. *Advances in Mechanical Engineering*, **15**, 1-17. <https://doi.org/10.1177/16878132221147463>
- [15] Erem Halilsoy, A. and Iscioglu, F. (2024) Reliability Analysis of (R, S)-Out-Of-N Multi-State Systems Using Copulas. *Engineering Computations*, **42**, 234-254. <https://doi.org/10.1108/ec-03-2024-0202>
- [16] Joe, H., Li, H. and Nikoloulopoulos, A.K. (2010) Tail Dependence Functions and Vine Copulas. *Journal of Multivariate Analysis*, **101**, 252-270. <https://doi.org/10.1016/j.jmva.2009.08.002>
- [17] Geenens, G. (2024) (Re-)reading Sklar (1959)—A Personal View on Sklar's Theorem. *Mathematics*, **12**, Article 380. <https://doi.org/10.3390/math12030380>
- [18] Boateng, M.A., Omari-Sasu, A.Y., Avuglah, R.K. and Frempong, N.K. (2022) A Mixture of Clayton, Gumbel, and Frank Copulas: A Complete Dependence Model. *Journal of Probability and Statistics*, **2022**, Article ID: 1422394. <https://doi.org/10.1155/2022/1422394>

-
- [19] Risca, G., Galimberti, S., Rebora, P., Cattoni, A., Valsecchi, M.G. and Capitoli, G. (2025) Archimedean Copulas: A Useful Approach in Biomedical Data—A Review with an Application in Pediatrics. *Stats*, **8**, Article 69. <https://doi.org/10.3390/stats8030069>
- [20] Saekow, S., Chiawkhun, P., Yamaka, W., Nakharutai, N. and Phetpradap, P. (2024) Estimation of Contagion: Bayesian Model Averaging on Tail Dependence of Mixture Copula. *Mathematics*, **12**, Article 3350. <https://doi.org/10.3390/math12213350>
- [21] Eguchi, S. and Kato, S. (2025) Minimum Copula Divergence for Robust Estimation. arXiv: 2502.16831.
- [22] Ueda, N. and Nakano, R. (1998) Deterministic Annealing EM Algorithm. *Neural Networks*, **11**, 271-282.
- [23] Lartigue, T., Durrleman, S. and Allasonnière, S. (2022) Deterministic Approximate EM Algorithm; Application to the Riemann Approximation EM and the Tempered Em. *Algorithms*, **15**, Article 78. <https://doi.org/10.3390/a15030078>
- [24] Poggi, N., Carrera, D., Gavalda, R., Ayguadé, E. and Torres, J. (2012) A Methodology for the Evaluation of High Response Time on E-Commerce Users and Sales. *Information Systems Frontiers*, **16**, 867-885. <https://doi.org/10.1007/s10796-012-9387-4>
- [25] Wei, Y. and Pan, X. (2025) The Analysis of Marketing Performance in E-Commerce Live Broadcast Platform Based on Big Data and Deep Learning. *Scientific Reports*, **15**, Article No. 15594. <https://doi.org/10.1038/s41598-025-00546-w>