# 一类依赖范围大于 FGM 型联结函数的新 Copula

**Swaroop Georgy Zachariah**[a]，**Mohd. Arshad**[a1] 和 **Ashok Kumar Pathak**[ba] 印度理工学院印多尔分校数学系，印度[b] 旁遮普中央大学数学与统计系，印度巴廷达

arXiv:2304.02231v1 [math.ST] 5 Apr 2023

**摘要**

我们提出了一种具有正、负依赖性质的新型二元对称 copula。所提出的 Copula 具有简单的数学结构、相较于 FGM 联结函数及其推广形式更宽的依赖范围，且不存在下尾依赖和上尾依赖。所提出的 Copula 的斯皮尔曼秩相关系数最大范围为$[-0.5866, 0.5866]$，扩大了 FGM 联结函数及其各种推广形式的依赖范围。我们利用所提出的 Copula 构建了一种新的二元瑞利分布，并研究了其若干统计性质。通过分析一个真实数据集，说明所提出的二元分布在实际情境中的相关性。

关键词：联结函数、Kendall 秩相关系数、斯皮尔曼秩相关系数、瑞利分布、二元分布。

数学主题分类（*2020*年）：62H05, 62H10。

## 1 引言

联结函数在统计学、金融学、工程学和医学领域的相关数据集建模中发挥着重要作用。若有一个联结函数族，我们自然就能得到一组多元分布，并可按需指定边缘分布。联结函数的这一特性对依赖建模和模拟不可或缺的各个研究领域都很有用。Sklar（1959）证明，若 $X$ 和 $Y$ 是具有边缘累积分布 $G(x)$ 和 $H(y)$ 的两个随机变量，则存在一个函数 $C$，使得$F(x, y) = P(X \leq x, Y \leq y) = C(G(x) H(y)) \ \forall (x, y) \in \mathbb{R}^2$。若 $G(\cdot)$ 和 $H(\cdot)$ 连续，则 $C$ 唯一；否则，$C$ 在 Range($F$)×Range($G$) 上唯一确定。因此，每个联合累积分布函数都可以通过联结函数表示为相应边缘累积分布的函数。此外，若二元函数 $C : \mathcal{I}^2 \to \mathcal{I}$ 满足以下条件，则称其为二元 Copula：

$$C(u, 0) = C(0, v) = 0 \ \text{ and } \ C(u, 1) = u, C(1, v) = v, \ \forall u, v \in \mathcal{I} \tag{1.1}$$

$$C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0, \ \forall u_1 < u_2, v_1 < v_2 \in \mathcal{I} \tag{1.2}$$

[1]通讯作者。电子邮箱：arshad@iiti.ac.in（M. Arshad），ashokiitb09@gmail.com（Ashok KumarPathak），swaroopgeorgy@gmail.com（SwaroopGeorgy Zachariah）。

其中 $\mathcal{I} = [0,1]$ 和 $\mathcal{I}^2 = [0,1] \times [0,1]$。注意，式（1.2）称为 2-递增性质。关于联结函数的更多详情，参见 Nelsen（2006年）、Durante 和 Sempi（2016）、Mai 和 Scherer（2017）、Hofert 等（2018年）及其中的参考文献。

近年来，许多研究者使用 copula 函数构造多元分布。例如，Achcar 等人（2015年）通过 Farlie-Gumbel-Morgenstern（FGM）联结函数研究了二元广义指数分布。Kundu 和 Gupta（2017）使用高斯Copula讨论了二元 Birnbaum-Saunders 分布。Abd Elaal 和 Jarwan（2017）利用 FGM 和 Plackett copula 函数研究了二元广义指数分布，并讨论了各种估计方法。El-Sherpieny 等人（2018）使用 FGM 联结函数构造了二元威布尔分布。最近，Almetwally 和 Muhammed（2020年）使用 FGM 联结函数构造了二元 Fréchet 分布。

文献中有多种多样的联结函数，其中，FGM 联结函数因其数学结构简单且能够呈现正依赖和负依赖而备受关注（参见 Farlie（1960））。斯皮尔曼秩相关系数记为 $\rho_c$，是用于衡量联结函数所刻画的依赖结构的一种依赖度量。 $\rho_c$的取值范围为 -1 到 1。 $\rho_c$ 的正值表示正依赖，负值表示负依赖。对于 FGM 联结函数，斯皮尔曼秩相关系数的取值范围非常小，即 $\rho_c \in [-0.33, 0.33]$ （参见 Farlie（1960））。因此，FGM 联结函数不适合对具有较强依赖结构的数据进行建模。许多研究者尝试提出 FGM 型联结函数，以提高相关系数。Huang 和 Kotz（1999）提出了两个扩展FGM Copula，其分别具有 $\rho_c \in [-0.33, 0.375]$ 和 $\rho_c \in [-0.33, 0.391]$。Bairamov 和 Kotz（2002年）以及 Bekrizadeh 等人（2015年）也分别通过 $\rho_c \in [-0.48, 0.502]$和 $\rho_c \in [-0.50, 0.43]$ 扩展了 FGM 联结函数。最近，Chesneau（2022）提出了一种多项式正弦联结函数，能够呈现正依赖和负依赖，其具有 $\rho_c \in [-0.4927, 0.4927]$。在上述大多数研究中，可以看到，为扩大斯皮尔曼相关系数的取值范围而增加了参数，这导致数学结构变得复杂，估计未知参数的计算成本也更高。为克服这些缺点，我们提出一种简单的二元联结函数，无需向现有联结函数添加任何参数。所提出的 Copula 提高了文献中各种 FGM型Copula 的斯皮尔曼相关系数依赖范围。

本文结构安排如下：第 2 节和第 3 节将介绍一种新的二元 Copula，并研究所提出的 Copula 的一些依赖性度量。第 4 节将由所提出的 Copula 推导出一种新的二元瑞利分布。我们还将推导条件分布和乘积矩的表达式。此外，本文还将以所提出的二元瑞利分布为工具，对一组真实数据进行分析。

# 2 新的二元 Copula

考虑以下二元函数

$$C(u, v; \delta, \alpha) = uv + \delta \left(1 - e^{\alpha(u-u^2)}\right) \left(1 - e^{\alpha(v-v^2)}\right), \quad (u, v) \in \mathcal{I}^2, \tag{2.1}$$

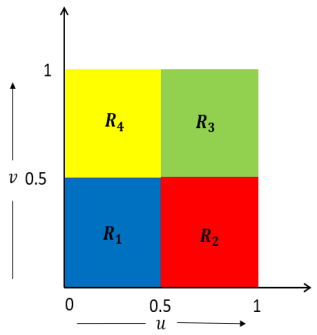
其中 $\alpha$ 和 $\delta$ 为实值参数。假设参数 $\delta$ 取决于 $\alpha$，且 $\alpha$在 $\mathbb{R}$中取任意值。式（2.1）中定义的二元函数满足式（1.1）中给出的二元 Copula 的边界条件。但我们需要找出参数$\delta$ 的取值范围，使该函数成为有效的二元 Copula，即式（2.1）中的函数满足式（1.2）中给出的 2-递增性质。Kim 等人（2011）证明，

绝对连续 Copula 的 2-递增性质等价于 Copula 密度 $c(u,\ v;\delta,\ \alpha)$ 非负，即

$$c(u,v;\delta,\alpha)=\frac{\partial^2 C(u,v)}{\partial u\partial v}=1+\delta\alpha^2 g(u)g(v)\geq 0, \tag{2.2}$$

其中 $g(t)\ =(1-2t)\,e^{\alpha(t-t^2)}\,t\ \in\mathcal{I}$。显然，式中的 Copula 密度非负性取决于函数 $g(\cdot)$的性质(2.2)以及参数 $\delta$ 的取值。为了求得 $\delta$ 的可行范围，我们将 $(u,v)$ 的定义域 $\mathcal{I}^2$ 划分为四个象限：

$$\begin{aligned}
R_1&=\left\{(u,v)\in\mathcal{I}^2:0\leq u\leq\frac{1}{2},0\leq v\leq\frac{1}{2}\right\},\\
R_2&=\left\{(u,v)\in\mathcal{I}^2:\frac{1}{2}<u\leq 1,0\leq v\leq\frac{1}{2}\right\},\\
R_3&=\left\{(u,v)\in\mathcal{I}^2:\frac{1}{2}<u\leq 1,\frac{1}{2}<v\leq 1\right\},\\
R_4&=\left\{(u,v)\in\mathcal{I}^2:0\leq u\leq\frac{1}{2},\frac{1}{2}<v\leq 1\right\}.
\end{aligned}$$




由于当 $t\in\left(0,\frac{1}{2}\right)$ 时 $g(t)$ 为正，而当 $t\in\left(\frac{1}{2},1\right)$ 时为负，因此乘积$g(u)g(v)$在 $R_1\cup R_3$ 上为正，在 $R_2\cup R_4$ 上为负。因此，联结函数密度 $c(u,v;\delta,\alpha)$ 非负，当且仅当

$$\delta\geq\frac{-1}{\alpha^2 g(u)g(v)},\quad (u,v)\in R_1\cup R_3,$$

and

$$\delta\leq\frac{-1}{\alpha^2 g(u)g(v)},\quad (u,v)\in R_2\cup R_4.$$

因此，联结函数密度 $c(u,v;\delta,\alpha)$ 非负，当且仅当

$$\frac{-1}{\alpha^2\sup\limits_{(u,v)\in R_1\cup R_3}\{g(u)g(v)\}}\leq\delta\leq\frac{-1}{\alpha^2\inf\limits_{(u,v)\in R_2\cup R_4}\{g(u)g(v)\}}. \tag{2.3}$$

由于函数 $g(\cdot)$ 的性质取决于 $\alpha\in\mathbb{R}$，我们考虑以下三种情形：**情形 I：** 当 $\alpha=0$ 时。

在这种情况下，联结函数密度 $c(u,\ v;\delta,\ 0)\ =1$，其为非负函数，因此对于 $\delta$ 的任意取值，2-递增性质均成立。此外，所提出的联结函数 (2.1) 退化为乘积联结函数。注意，乘积联结函数是一种众所周知的联结函数，它对应于两个随机变量的独立性。

**情形 II：** 当 $\alpha\leq 2,\ \alpha\neq 0$。在这种情况下，$g(t)$ 是 $\mathcal{I}$ 上的递减函数，且 $g(0)\ =1,g\left(\frac{1}{2}\right)=0$ 以及 $g(1)\ =-1$。由此可知，$g(t)$ 在 $[0,1]$ 上取值，其中 $t\in\left[0,\frac{1}{2}\right]$；而当 $t\in\left(\frac{1}{2},1\right]$ 时，取值为 $[-1,0)$。因此，乘积函数 $g(u)g(v)$ 在 $R_1\cup R_3$ 上有上界 1，并且上界 1 在 $(u,v)\in\{(0,0),(1,1)\}$ 处达到。因此，$\sup\{g(u)g(v)\}\ =1$。进一步地，对于 $(u,v)\ \in R_2,\ \ g(u)$ 取 $(u,v)\in R_1\cup R_3$ （$[-1,0)$）中的取值以及 $g(v)$ 的取值范围为 $[0,1]$。这意味着乘积函数 $g(u)g(v)$ 在 $R_2$ 上以下界 $-1$ 为界，且下界 $-1$ 在 $u=1,v=0$ 处取得。同样，对于 $(u,v)\ \in R_4$，乘积函数 $g(u)g(v)$ 以下界 $-1$为界，该下界在$u=0,v=1$ 处取得。因此，$\inf_{(u,v)\in R_2\cup R_4}\{g(u)g(v)\}\ =-1$。现在，将这些值代入不等式 (2.3)，

可得参数 $\delta$ 的可行范围为

$$\frac{-1}{\alpha^2} \leq \delta \leq \frac{1}{\alpha^2}. \tag{2.4}$$

**情形 III：** 当 $\alpha > 2$。
令 $r_1 = \frac{1}{2} - \frac{1}{\sqrt{2\alpha}}$，并令 $r_2 = \frac{1}{2} + \frac{1}{\sqrt{2\alpha}}$。显然，$0 < r_1 < \frac{1}{2} < r_2 < 1$。可以观察到，$g(t)$ 在 $t \in [0, r_1]$ 上递增，在 $t \in (r_1, r_2)$ 上递减，并在 $t \in [r_2, 1]$ 上递增。此外，$g(t)$ 在 $t \in \left[0, \frac{1}{2}\right)$ 上取正值，在 $t \in \left(\frac{1}{2}, 1\right]$, 上取负值，且 $g\left(\frac{1}{2}\right) = 0$。此外, $g(t)$ 在 $t = r_1$ 处取得最大值，最大值为 $g(r_1) = \sqrt{\frac{2}{\alpha}} \exp\left\{\frac{\alpha}{4} - \frac{1}{2}\right\}$；$g(t)$ 在 $t = r_2$ 处取得最小值, 最小值为 $g(r_2) = -\sqrt{\frac{2}{\alpha}} \exp\left\{\frac{\alpha}{4} - \frac{1}{2}\right\}$。由于函数 $g(u)$ 和 $g(v)$ 在象限 $R_1$ 上为正，并在 $u = r_1, v = r_1$ 处取得最大值，因此乘积函数 $g(u)g(v)$ 的最大值为 $[g(r_1)]^2 = \frac{2}{\alpha} \exp\left\{\frac{\alpha}{2} - 1\right\}$，在 $(u, v) \in R_1$ 上取得。由于函数 $g(u)$ 和 $g(v)$ 在 $R_3$ 上为负，并在 $u = r_2, v = r_2$ 处取得最小值，因此乘积函数 $g(u)g(v)$ 为正，其最大值为 $[g(r_2)]^2 = \frac{2}{\alpha} \exp\left\{\frac{\alpha}{2} - 1\right\}$，在 $(u, v) \in R_3$ 上取得。因此，$\sup_{(u,v) \in R_1 \cup R_3} \{g(u)g(v)\} = \frac{2}{\alpha} \exp\left\{\frac{\alpha}{2} - 1\right\}$。类似地，我们发现，乘积函数 $g(u)g(v)$ 在象限 $R_2 \cup R_4$ 上的下确界等于 $g(r_1)g(r_2) = -\frac{2}{\alpha} \exp\left\{\frac{\alpha}{2} - 1\right\}$。现在，将这些值代入式 (2.3)，得到 $\delta$ 的可行范围为

$$-\frac{1}{2\alpha} \exp\left\{1 - \frac{\alpha}{2}\right\} \leq \delta \leq \frac{1}{2\alpha} \exp\left\{1 - \frac{\alpha}{2}\right\}. \tag{2.5}$$

因此，利用式 (2.4) 和式 (2.5) 中给出的参数 $\delta$ 的可行范围，我们提出如下二元 Copula

$$C(u, v; \delta, \alpha) = uv + \delta\left(1 - e^{\alpha(u-u^2)}\right)\left(1 - e^{\alpha(v-v^2)}\right), \quad (u, v) \in \mathcal{I}^2, \tag{2.6}$$

其中，$\alpha$ 是一个实值参数，且 $|\delta| \leq \delta^\star(\alpha)$。这里，

$$\delta^\star(\alpha) = \begin{cases} \frac{1}{\alpha^2}, & \text{if } \alpha \in (-\infty, 2] \setminus \{0\} \\ \frac{1}{2\alpha} \exp\left\{1 - \frac{\alpha}{2}\right\}, & \text{if } \alpha > 2. \end{cases}$$

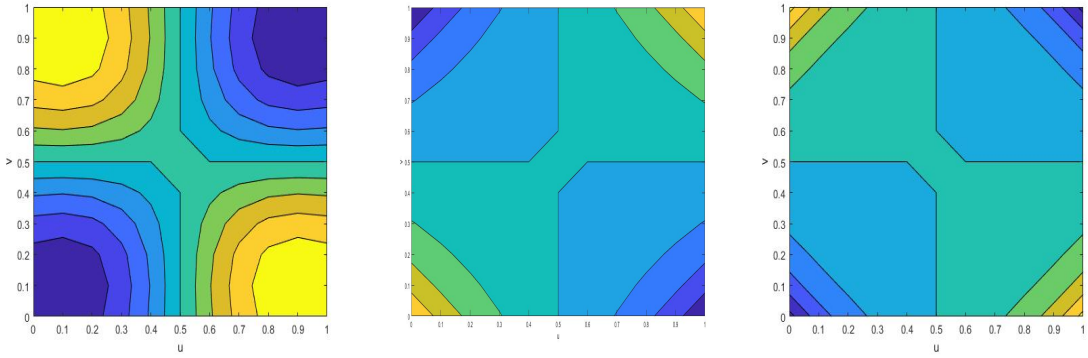
回顾一下，当 $\alpha = 0$ 时，我们所提出的 Copula 退化为乘积 Copula。此外，所提出的 Copula 的 Copula 密度 $c(u, v) = \frac{\partial^2 C}{\partial u \partial v}(u, v)$ 由下式给出

$$c(u, v) = 1 + \alpha^2 \delta (1 - 2u)(1 - 2v) \exp\left\{\alpha\left(u - u^2 + v - v^2\right)\right\}, \quad (u, v) \in \mathcal{I}^2, \tag{2.7}$$

其中 $\alpha \in \mathbb{R}$ 且 $|\delta| \leq \delta^\star(\alpha)$。对于参数 $\alpha$ 和 $\delta$ 的不同取值，Copula 密度的等高线图如图1所示。

# 3 依赖关系度量

自从能够将依赖效应与边缘分布的影响分离开来以来，Copula 函数已被广泛用于刻画随机变量之间的依赖关系。随机变量之间的依赖关系可以通过若干广为人知的依赖度量来衡量，例如斯皮尔曼秩相关系数、基尼伽马系数、Kendall 秩相关系数、斯皮尔曼脚趾规则系数（见 Nelsen（2006年））以及 Blest 秩相关测度（见 Genest 和



(a) $\alpha = -3,\ \ \delta = -0.1$ (b) $\alpha = 2,\ \ \delta = 0.2$ (c) $\alpha = 3.8, \delta = -0.3$

图1：不同 $\alpha$ 和 $\delta$ 取值下联结函数密度 $c(u, v)$ 的等高线图。

Plante，2003）。这些度量通常基于观测值的秩，而非其实际取值。这些依赖度量的取值范围为 $-1$ 到 1。依赖度量为负值表示负依赖，为零表示独立性，为正值则表示正依赖。大多数此类度量在对随机变量施加严格递增变换后保持不变。由于一对随机变量 $X$ 和 $Y$ 的 copula 函数在对 $X$ 和 $Y$ 施加严格递增变换后保持不变，因此这些依赖度量都可以用联结函数表示。设 $C(u, v)$ 为与随机变量 $X$ 和 $Y$ 相关联的联结函数。以下是用联结函数表示的重要依赖度量：

- **斯皮尔曼秩相关系数**

$$\rho_c = 12 \int_0^1 \int_0^1 C(u, v) du dv - 3.$$

- **基尼伽马系数**

$$\gamma_c = 4 \left\{ \int_0^1 C(u, 1-u) du - \int_0^1 \left( u - C(u, u) \right) du \right\}.$$

- **Kendall 秩相关系数**

$$\tau_c = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1.$$

- **Blest 秩相关测度**

$$\eta_c = 24 \int_0^1 \int_0^1 (1-u)\, C(u, v) du dv - 2.$$

- **Spearman 脚距**

$$\phi_c = 6 \int_0^1 C(u, u) du - 2.$$

现在，我们将给出所提出的 Copula 函数 (2.6) 的各种依赖性度量表达式。

**命题 3.1.** 对于式 *(2.6)* 中定义的联结函数，斯皮尔曼秩相关系数 $\rho_C$ 和基尼伽马系数 $\gamma_C$ 为

$$\rho_C = \begin{cases} 12\delta\left[1-\sqrt{\frac{\pi}{|\alpha|}}e^{\alpha/4}erfi\left\{\frac{\sqrt{|\alpha|}}{2}\right\}\right]^2, & if\ \alpha<0 \\ 12\delta\left[1-\sqrt{\frac{\pi}{\alpha}}e^{\alpha/4}erf\left\{\frac{\sqrt{\alpha}}{2}\right\}\right]^2, & if\ \alpha\geq 0, \end{cases}$$

$$\gamma_C = \begin{cases} 8\delta\left[1-2\sqrt{\frac{\pi}{|\alpha|}}e^{\alpha/4}erfi\left\{\frac{\sqrt{|\alpha|}}{2}\right\}+\sqrt{\frac{\pi}{2|\alpha|}}e^{\alpha/2}erfi\left\{\sqrt{\frac{|\alpha|}{2}}\right\}\right], & if\ \alpha<0 \\ 8\delta\left[1-2\sqrt{\frac{\pi}{\alpha}}e^{\alpha/4}erf\left\{\frac{\sqrt{\alpha}}{2}\right\}+\sqrt{\frac{\pi}{2\alpha}}e^{\alpha/2}erf\left\{\sqrt{\frac{\alpha}{2}}\right\}\right], & if\ \alpha\geq 0, \end{cases}$$

其中，$erf(t) = \frac{2}{\sqrt{\pi}}\int_0^t e^{-z^2}dz$ 表示误差函数（参见 *Abramowitz and Stegun (1972)*），$erfi(t) = \frac{2}{\sqrt{\pi}}\int_0^t e^{z^2}dz$ 表示虚误差函数（参见 *Marcinowski and Sadowski (2020)*）。

**注 3.1**。 可以验证，在式（*2.6*）给出的联结函数下，其他依赖性度量的表达式满足以下关系。

$$\eta_c = \rho_c = \frac{3}{2}\tau_c, \quad and \quad \phi_c = \frac{3}{4}\gamma_c.$$

表 1 展示了新联结函数在联结函数参数 $\alpha$ 取不同值时，斯皮尔曼秩相关系数和基尼伽马系数的数值。由于新联结函数具有对称性，表 1 中仅列出了依赖度量的上界。下界为上界的相反数。从表 1 可以看出，当 $\alpha = 3.8$ 时，斯皮尔曼秩相关系数 $\rho_c \in [-0.5866, 0.5866]$，从而拓宽了斯皮尔曼秩相关系数这一依赖度量相对于广受欢迎的 FGM 联结函数及其各种推广形式的取值范围。

表 1：不同 $\alpha$ 值下的斯皮尔曼秩相关系数和基尼伽马系数

| $\alpha$ | $\delta_{upper}$ | $\rho_{upper}$ | $\gamma_{upper}$ | $\alpha$ | $\delta_{upper}$ | $\rho_{upper}$ | $\gamma_{upper}$ |
|---|---|---|---|---|---|---|---|
| -3 | 0.1111 | 0.1899 | 0.1463 | 1.2 | 0.6944 | 0.4264 | 0.3473 |
| -2.7 | 0.1372 | 0.2003 | 0.1547 | 1.5 | 0.4444 | 0.4544 | 0.3718 |
| -2.4 | 0.1736 | 0.2113 | 0.1638 | 1.8 | 0.3086 | 0.4845 | 0.3984 |
| -2.1 | 0.2268 | 0.2231 | 0.1736 | 2 | 0.25 | 0.506 | 0.4174 |
| -1.8 | 0.3086 | 0.2358 | 0.1841 | 2.3 | 0.1871 | 0.5348 | 0.4434 |
| -1.5 | 0.4444 | 0.2493 | 0.1954 | 2.6 | 0.1425 | 0.5561 | 0.4634 |
| -1.2 | 0.6944 | 0.2638 | 0.2076 | 2.9 | 0.1099 | 0.571 | 0.4783 |
| -0.9 | 1.2346 | 0.2794 | 0.2207 | 3.2 | 0.0858 | 0.5805 | 0.4888 |
| -0.6 | 2.7778 | 0.2961 | 0.2349 | 3.5 | 0.0675 | 0.5855 | 0.4956 |
| -0.3 | 11.1111 | 0.314 | 0.2502 | 3.8 | 0.0535 | **0.5866** | 0.4992 |
| 0 | 0 | 0 | 0 | 4.1 | 0.0427 | 0.5845 | **0.5002** |
| 0.3 | 11.1111 | 0.3541 | 0.2845 | 4.4 | 0.0342 | 0.5798 | 0.499 |
| 0.6 | 2.7778 | 0.3764 | 0.3038 | 4.7 | 0.0276 | 0.5729 | 0.4959 |
| 0.9 | 1.2346 | 0.4005 | 0.3247 | 5 | 0.0223 | 0.5643 | 0.4912 |

为了继续讨论随机变量之间的依赖关系，文献中还提出了一些其他依赖性质。例如，象限依赖、二阶全正性（$TP_2$）以及尾部依赖系数。关于这些性质的更详细讨论，可参见 Lehmann（1966）、Barlow 和 Proschan（1975年）、Drouet Mari 和 Kotz（2001）、Nelsen（2006年）、Lai 和 Balakrishnan（2009）以及 Bhuyan 等人（2020年）。下面我们将讨论式（2.6）所给联结函数的这些依赖性质。

## 3.1 象限依赖

该 ndom 向量 $(X, Y)$ 称为正（负）象限依赖，如果

$$P(X \leq x, Y \leq y) \geq (\leq)\ P(X \leq x)P(Y \leq y), \quad \forall\ (x, y) \in \mathbb{R}^2.$$

Si类似地，若联结函数 $C(u, v)$ 满足正（负）象限依赖，则称其为正（负）象限依赖联结函数 f

$$C(u, v) \geq (\leq)\ uv, \quad \forall (u, v) \in \mathcal{I}^2.$$

可以验证，联结函数 (2.6) 的象限依赖仅取决于联结函数参数 $\delta$。证明很直接，故略去。因此，我们得到以下结果。

命题 *3.2. </style}*联结函数 *(2.6)* 具有正（负）象限依赖，当且仅当 $\delta \geq 0$ *(*$\delta \leq 0$*)*。

## 3.2 二阶全正性 ($TP_2$)

若二元函数 $g(x,\ y)$ 满足以下条件，则称其具有二阶全正性 ($TP_2$)：

$$g(x_1, y_1)g(x_2, y_2) - g(x_2, y_1)g(x_2, y_1) \geq 0, \quad \forall\ x_1 < x_2, y_1 < y_2.$$

$TP_2$ 性质是一种更强的依赖关系概念。如果 Copula 密度 $c(u,\ v)$ 具有 $TP_2$性质，那么相应的联结函数 $C(u,\ v)$ 具有随机递增（SI）、右尾递增（RTI）和正象限依赖（PQD）性质。更多详情请参见 Nelsen（2006年）、Karlin（1968）和 Joe（1997）。

Holland和Wang（1987）证明，函数 $g(x,\ y)$ 具有 $TP_2$ 性质 i f

$$\zeta_g(x, y) = \frac{\partial^2 \ln g(x, y)}{\partial x \partial y} \geq 0, \quad \forall (x, y) \in \mathbb{R}^2.$$

现在，利用 Holland和Wang（1987）的结果，我们将证明以下结论。

**命题 3.3**。*Copula* 密度（*2.7*）在 $\delta \geq 0$ 和 $\alpha \leq 2$ 成立时具有 $TP_2$ 性质。

证明。我们有：

$$\begin{aligned}\zeta_c(u, v) =& \frac{\partial^2 \ln c(u, v)}{\partial x \partial y} \\ =& \frac{\delta\alpha^2 \left[2 - \alpha(1 - 2u)\right] \left[2 - \alpha(1 - 2v)\right] \exp\left\{\alpha\left((u - u^2) + (v - v^2)\right)\right\}}{\left(1 + \delta\alpha^2 \left(1 - 2u\right) \left(1 - 2v\right) \exp\left\{\alpha\left((u - u^2) + (v - v^2)\right)\right\}\right)^2} \geq 0,\end{aligned}$$

对于每个 $\delta \geq 0$ 和 $\alpha \leq 2$。 □

### 3.3 尾部依赖系数

尾部依赖系数衡量$\mathcal{I}^2$的右上象限和左下象限中随机变量之间依赖关系的程度。用联结函数表示，上尾和下尾依赖系数分别记为 $\lambda_U$ 和 $\lambda_L$，其定义为

$$\lambda_L = \lim_{u\to 0^+} \frac{C(u,u)}{u}, \quad \lambda_U = \lim_{u\to 1^-} \frac{1-2u+C(u,u)}{1-u}. \tag{3.1}$$

已知 $0 \le \lambda_L \le 1$和 $0 \le \lambda_U \le 1$。如果 $\lambda_L \in (0,1]$, 我们称联结函数 $C(u,\ v)$ 具有下尾部依赖；如果 $\lambda_L = 0$，则称 $C(u,\ v)$ 不具有下尾部依赖。对于 $\lambda_U$ 也可以作出类似的解释（参见 Nelsen（2006），第214页）。现在，我们将证明以下结果。

**命题3.4**。 联结函数（*2.6*）不存在尾部依赖。

证明。 利用式（3.1），可得

$$\begin{aligned}
\lambda_L &= \lim_{u\to 0^+} \frac{u^2 + \delta\left(1-e^{\alpha(u-u^2)}\right)^2}{u} \\
&= \delta \lim_{u\to 0^+} \frac{1}{u}\left(1 - \sum_{n=0}^{\infty} \frac{\alpha^n u^n (1-u)^n}{n!}\right)^2 \\
&= \delta \lim_{u\to 0^+} \alpha^2 u (1-u)^2 \left(\sum_{n=2}^{\infty} \frac{\alpha^{n-1} u^{n-1}(1-u)^{n-1}}{n!}\right)^2 \\
&= 0.
\end{aligned}$$

$$\begin{aligned}
\lambda_U &= \lim_{u\to 1^-} \frac{1-2u+u^2 + \delta\left(1-e^{\alpha(u-u^2)}\right)^2}{1-u} \\
&= \lim_{u\to 1^-} \frac{(1-u)^2 + \delta\left(1-\sum_{n=0}^{\infty} \frac{\alpha^n u^n (1-u)^n}{n!}\right)^2}{(1-u)} \\
&= \delta \lim_{u\to 1^-} \alpha^2 u^2 (1-u) \left(\sum_{n=2}^{\infty} \frac{\alpha^{n-1} u^{n-1}(1-u)^{n-1}}{n!}\right)^2 \\
&= 0.
\end{aligned}$$

□

下一节将应用所提出的联结函数（2.6）构造一个二元瑞利分布。我们将研究新的二元瑞利分布的一些统计性质，并给出涉及该新分布的真实数据分析。

# 4 一个新的二元瑞利分布

瑞利分布是医学、工程学、粒子物理学和经济学中最常用的模型之一。若随机变量 $X$ 服从参数为 $\lambda$ 的瑞利分布，记作 $Rayleigh(\lambda)$，且其累积分布函数（CDF）为 $F(x;\ \lambda)\ =$

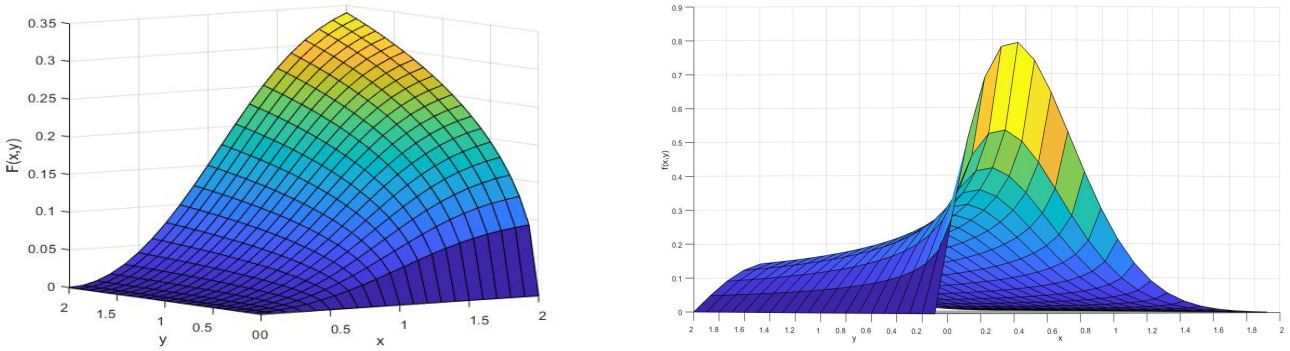
$1-e^{-x^2/2\lambda^2}$, $x>0$, $\lambda>0$, 其对应的概率密度函数（PDF）为 $f(x;\lambda)=\frac{x}{\lambda^2}e^{-x^2/2\lambda^2}$, $x>0$, $\lambda>0$。设 $X$ 和 $Y$ 是两个随机变量，分别服从 $Rayleigh(\lambda_1)$ 和 $Rayleigh(\lambda_2)$ 分布，而 $X$ 和 $Y$ 之间的依赖关系由联结函数 (2.6) 建模。那么， $X$ 和 $Y$ 的联合分布函数为

$$F(x,y;\Theta)=\left(1-e^{-x^2/2\lambda_1^2}-e^{-y^2/2\lambda_2^2}+e^{-\left(x^2/2\lambda_1^2+y^2/2\lambda_2^2\right)}\right)+\delta\left(1-e^{\alpha\left(e^{-x^2/2\lambda_1^2}-e^{-x^2/\lambda_1^2}\right)}\right)\left(1-e^{\alpha\left(e^{-y^2/2\lambda_2^2}-e^{-y^2/\lambda_2^2}\right)}\right), \tag{4.1}$$

其中 $x>0$, $y>0$, $\lambda_1>0$, $\lambda_2>0$, $\alpha\in\mathbb{R}$, $|\delta|\leq\delta^{\star}(\alpha)$ 和 $\Theta=(\lambda_1,\ \lambda_2,\ \alpha,\ \delta)$。若非负随机向量 $(X,\ Y)$ 的联合累积分布函数由式 (4.1) 给出，且记为 BRD$(\lambda_1,\ \lambda_2,\ \alpha,\ \delta)$，则称其服从参数为 $\lambda_1$, $\lambda_2$, $\alpha$ 和 $\delta$ 的二元瑞利分布。相应的联合密度函数为

$$f(x,y;\Theta)=\left(\frac{xy}{\lambda_1^2\lambda_2^2}e^{-\left(x^2/2\lambda_1^2+y^2/2\lambda_2^2\right)}\right)\left[1+\delta\alpha^2\left(2e^{-x^2/2\lambda_1^2}-1\right)\left(2e^{-y^2/2\lambda_2^2}-1\right)\left(\exp\left\{\alpha\left(e^{-x^2/2\lambda_1^2}-e^{-x^2/\lambda_1^2}+e^{-y^2/2\lambda_2^2}-e^{-y^2/\lambda_2^2}\right)\right\}\right)\right], \tag{4.2}$$

其中 $x>0$, $y>0$, $\lambda_1>0$, $\lambda_2>0$, $\alpha\in\mathbb{R}$, $|\delta|\leq\delta^{\star}(\alpha)$ 和 $\Theta=(\lambda_1,\ \lambda_2,\ \alpha,\ \delta)$。图 2 展示了 BRD 族的联合累积分布函数 (4.1) 和联合密度函数 (4.2) 的曲面图。这些图使用 Matlab R2021b 绘制。现在，我们将给出条件




(a) (b)

图 2：对于 $\lambda_1=3$、 $\lambda_2=2$、 $\delta=0.5$、$\alpha=3.8$，BRD 分布的 $F(x,\ y)$ 和 $f(x,\ y)$ 的曲面图。

分布和乘积矩。

**命题 4.1**。设 $(X,\ Y)\sim BRD(\lambda_1,\ \lambda_2,\ \alpha,\ \delta)$。则

*(i)* $X\sim Rayleigh(\lambda_1)$ *and* $Y\sim Rayleigh(\lambda_2)$,

(*ii*) $X$ 给定 $Y=y$ 时的条件密度函数为

$$f(x|y)=\left(\frac{x}{\lambda_1^2}e^{-x^2/2\lambda_1^2}\right)\left[1+\delta\alpha^2\left(2e^{-x^2/2\lambda_1^2}-1\right)\left(2e^{-y^2/2\lambda_2^2}-1\right)\left(e^{\alpha\left(e^{-x^2/2\lambda_1^2}-e^{-x^2/\lambda_1^2}+e^{-y^2/2\lambda_2^2}-e^{-y^2/\lambda_2^2}\right)}\right)\right],$$

*(iii)* $X$ 给定 $Y=y$ 的条件分布函数为

$$F(x|y)=\left(1-e^{-x^2/2\lambda_1^2}\right)+\delta\alpha\left(e^{\alpha\left(e^{-x^2/2\lambda_1^2}-e^{-x^2/\lambda_1^2}\right)}-1\right)\left(2e^{-y^2/2\lambda_2^2}-1\right)\left(e^{\alpha\left(e^{-y^2/2\lambda_2^2}-e^{-y^2/\lambda_2^2}\right)}\right),$$

其中 $x>0, y>0, \lambda_1>0, \lambda_2>0, \alpha\neq 0, |\delta|\leq\delta^{\star}$（$\alpha$）以及 $\Theta=$（$\lambda_1, \lambda_2, \alpha, \delta$）。

**命题 4.2**。 令 $(X,Y)\sim BRD(\lambda_1, \lambda_2, \alpha, \delta)$。则 $(r,s)$阶乘积矩可表示为

$$\begin{aligned}E(X^rY^s)=&\lambda_1^r\lambda_2^s 2^{(r+s)/2}\Gamma\left(1+r/2\right)\Gamma\left(1+s/2\right)\\&\left[1+\delta\alpha^2\left(\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}\sum_{t=0}^{k}\frac{(-1)^t}{\lambda_1^2}\binom{k}{t}\left(2\left(k+t+2\right)^{-\frac{r+2}{2}}-\left(k+t+1\right)^{-\frac{r+2}{2}}\right)\right)\right.\\&\left.\left(\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}\sum_{t=0}^{k}\frac{(-1)^t}{\lambda_2^2}\binom{k}{t}\left(2\left(k+t+2\right)^{-\frac{s+2}{2}}-\left(k+t+1\right)^{-\frac{s+2}{2}}\right)\right)\right],\end{aligned}$$

其中 $\Gamma(t)$ 表示通常的伽马函数。

## 4.1 真实数据应用

我们考虑 Meintanis（2007）报告的2004年至2006年欧洲足球协会联盟冠军联赛数据集。在该数据集中， $X$ 和 $Y$ 分别表示 A队和 B队首粒进球的时间（以分钟计）。为检验 $X$ 和 $Y$的边缘分布是否支持瑞利分布，我们进行 Kolmogorov-Smirnov（KS）单样本检验。ˆ KS 检验结果表明， $X$ 支持参数为 $\lambda_1=32.14599$的瑞利分布（p 值=0.934，KS 统计量值为=0.088515）。类似地， $Y$ 也支持参数为 $\lambda_2=28.172255$ 的瑞利分布（p 值=0.07727，KS 统计量值为=0.20968）。我们拟合所提出的二元瑞利分布，结果见表 2。我们将新的 BRD 模型与 Meintanis（2007）提出的 Marshall Olkin 二元指数分布（BMOED）、Mirhosseini 等（2015）提出的二元广义指数分布（BGED），以及 Pathak 和 Vellaisamy（2022）提出的二元广义瑞利分布（BGRD）进行比较。我们使用对数似然（LL）函数、赤池信息准则（AIC）和贝叶斯信息准则（BIC）作为比较准则。AIC 和 BIC 的公式如下:

$$\mathrm{AIC}=2k-2\ln L,\quad \text{and}\quad \mathrm{BIC}=k\ln n-2\ln L,$$

其中， $k$ 是模型中的参数个数， $n$ 是样本量， $L$ 是似然函数的最大值。从表 2 可以清楚地看出，对于欧洲足球协会联盟冠军联赛数据集，二元瑞利分布比 BGED、BMOED 和 BGRD 拟合得更好。

表 2：使用 UEFA 冠军联赛数据集得到的二元分布的极大似然估计、LL、AIC 和 BIC 值。

| 二元分布 | 极大似然估计 | LL | AIC | BIC |
|---|---|---|---|---|
| BGED | $\hat{}\alpha_1=0.0244,\ \hat{}\alpha_2=0.0304,\ \hat{}\theta=0.999$ | -340.5234 | 687.0468 | 691.8795 |
| BMOED | $\hat{\lambda}_1=0.012, \hat{\lambda}_2=0.014, \hat{\lambda}_3=0.022$ | -339.006 | 684.012 | 688.8448 |
| BGRD | $\hat{}b_1=0.000530,\ \hat{}b_2=0.000836,\ \hat{}\theta=0.40331$ | -331.879 | 664.589 | 672.6436 |
| BRD | $\hat{}\lambda_1=33.39429, \hat{}\lambda_2=28.08949, \hat{}\delta=10.39829, \hat{}\alpha=0.2871858$ | -327.256 | 664.512 | 668.9557 |

# 5 结论与未来研究方向

本文提出了一种新的二元对称 copula，它同时表现出正向和负向依赖关系。该联结函数的主要特点是：(i) 数学结构简单；(ii) 与 FGM 联结函数及其推广形式相比，具有更宽的依赖范围；(iii) 不存在下尾依赖和上尾依赖。基于所提出的 Copula，我们构建了一种新的二元瑞利分布（BRD），并讨论了其若干统计性质。所提出的二元模型对一个真实数据集拟合得更好。由于本文仅考虑了二元联结函数的对称版本，因此其非对称版本仍是有待新研究者探索的未解决问题。


# 致谢

第一作者感谢印度理工学院印多尔分校为其开展研究提供经费支持。


## 利益冲突

作者声明不存在潜在利益冲突。

# 参考文献


Abd Elaal, M. K. and R. S. Jarwan (2017). 基于 copula 函数的二元广义指数分布推断。应用数学科学 *11*(24), 1155–1186. Abramowitz, M. and I. A. Stegun (1972). 数学函数手册：公式、图表和数学表。美国国家标准局应用数学系列第 *55* 卷。第十次印刷。ERIC. Achcar, J. A., F. A. Moala, M. H. Tarumoto, and L. F. Coladello (2015). 存在删失数据和协变量时，由 copula 函数导出的二元广义指数分布。运筹学 *35*(1), 165–186. Almetwally, E. M. and H. Z. Muhammed (2020). 关于二元 Fréchet 分布。统计学、应用与概率杂志 *9*(1), 1–21. Bairamov, I. and S. Kotz (2002). Huang-Kotz FGM 分布及其扩展的依赖结构与对称性。*Metrika 56*(1), 55–72. Barlow, R. E. and F. Proschan (1975). 可靠性与寿命试验的统计理论：概率模型。技术报告，佛罗里达州立大学塔拉哈西分校。Bekrizadeh, H., G. Parham, and M. Zadkarami (2015). 巴基斯坦统计学杂志，2015 年第 31 卷第 1 期，95–106：一种非对称广义 FGM copula 及其性质。巴基斯坦统计学杂志 *31*(1), 95–106. Bhuyan, P., S. Ghosh, P. Majumder, and M. Mitra (2020). 一种二元寿命分布与负相关性概念。*Stat 9*(1), e276. Chesneau, C. (2022). 关于一种简单多项式正弦 copula 的短论。亚洲数学与应用杂志 *2*, 1–14.

Drouet Mari, D. and S. Kotz (2001). 相关性与依赖关系。Imperial College Press. Durante, F. and C. Sempi (2016). *Copula* 理论原理，第 474 卷。CRC Press，博卡拉顿。El-Sherpieny, E., H. Z. Muhammed, and E. M. Almetwally (2018). FGM 二元 Weibull 分布。载于统计学、计算机科学与运筹学年度会议论文集（第 *53* 届），开罗大学统计研究院，第 55–77 页。Farlie, D. J. (1960). 一般二元分布中若干相关系数的表现。*Biometrika 47*(3/4), 307–323. Genest, C. and J.-F. Plante (2003). 关于 Blest 秩相关测度。《加拿大统计学杂志》*31*(1), 35–52. Hofert, M., I. Kojadinovic, M. Mächler, and J. Yan (2018). 使用 *R* 进行 *Copula* 建模的要素。施普林格。Holland, P. W. and Y. J. Wang (1987). 连续二元密度的相依函数。《统计学通讯：理论与方法》*16*(3), 863–876. Huang, J. S. and S. Kotz (1999). 对 Farlie-Gumbel-Morgenstern 分布的修正：一座难以逾越的高山。*Metrika 49*(2), 135–145. Joe, H. (1997). 多元模型与多元相依概念。CRC Press. Karlin, S. (1968). 全正性，第 1 卷。斯坦福大学出版社。Kim, J.-M., E. A. Sungur, T. Choi, and T.-Y. Heo (2011). 广义二元 Copula 及其性质。《模型辅助统计与应用》*6*(2), 127–136. Kundu, D. and R. C. Gupta (2017). 关于二元 Birnbaum–Saunders 分布。《美国数学与管理科学杂志》*36*(1), 21–33. Lai, C. D. and N. Balakrishnan (2009). 连续二元分布。施普林格。Lehmann, E. L. (1966). 若干相依概念。统计学年鉴 *37*(5), 1137–1153. Mai, J.-F. and M. Scherer (2017). *Copula* 模拟：随机模型、抽样算法与应用，第 6 卷。World Scientific. Marcinowski, J. and M. Sadowski (2020). 在受压杆形状优化问题中使用 erfi 函数。《国际应用力学与工程杂志》*25*(2), 75–87. Meintanis, S. G. (2007). Marshall–Olkin 分布的拟合优度检验及其应用。《统计规划与推断杂志》*137*(12), 3954–3963. Mirhosseini, S. M., M. Amini, D. Kundu, and A. Dolati (2015). 关于一种新的绝对连续二元广义指数分布。《统计方法与应用》*24*(1), 61–83. Nelsen, R. B. (2006). *Copula* 入门。施普林格科学与商业媒体。

Pathak, A. 和 P. Vellaisamy（2022）。双变量广义线性指数分布：性质与估计。*Communications in Statistics-Simulation and Computation 51*（9），5426–5446。

Sklar, M.（1959）。高维分布函数及其边缘分布。*Publ. inst. statist.univ. Paris 8*，229–231。