arXiv:2601.09016v1 [math.ST] 13 Jan 2026

Sarmanov copula 的随机表示

Christopher Blier-Wong

多伦多大学统计科学系，加拿大，
christopher.blierwong@utoronto.ca

2026年1月15日

摘要

萨尔马诺夫 copula 通过扰动独立 copula，为构建多元分布提供了一种简单且易于处理的方法。它们能够为密度及许多受关注的泛函提供闭式表达式，因此对实际应用颇具吸引力。然而，为确保 Sarmanov copula 有效，依赖参数需满足复杂条件，这限制了其在高维场景中的应用。验证其 d-递增性质通常需要满足一组组合的不等式，使得直接构造十分困难。为规避这一问题，我们为二元 Sarmanov copula 建立了一种随机表示。我们证明，每个可容许的 Sarmanov copula 都可以表示为由潜在伯努利对索引的独立单变量分布的混合。该随机表示将验证 copula 有效性的问题转化为确保伯努利概率质量函数具有非负性的问题。该表示还可将经典 copula 族作为特殊情形重新得到，包括 Farlie–Gumbel–Morgenstern、Huang–Kotz 以及 Bairamov–Kotz–Bekçi copula。我们进一步推导出斯皮尔曼 rho 和肯德尔秩相关系数的精确的全局界。随后，我们将伯努利混合构造引入高维情形，从而得到一类新的多元 Sarmanov Copula；这类 copula 具有易于验证的参数约束和可扩展的模拟算法。最后，我们证明，二元 Sarmanov copula 的幂次版本可以通过分块极大次序统计量获得类似的随机表示。

关键词：萨尔马诺夫 copula、Farlie–Gumbel–Morgenstern copula、黄–科茨 copula、高维 copula、非对称 copula、多元关联度量、随机表示

1 引言

在构建多元统计模型时，将依赖性与边际行为分开处理通常更为方便。在许多情况下，参数模型或半参数模型能够很好地捕捉边际行为，尤其是在统计学家具备丰富领域知识的情况下 McNeil 等（2015）。相比之下，依赖性很少能由线性相关充分刻画，可能在支持集上发生显著变化，并且对单调边际变换较为敏感。Copula 为单独表示这种依赖结构与边缘分布提供了一个标准框架。它们能够将任意边缘分布组合成一个多元分布，并且可能产生具有可解释关联参数的Copula族。关于 Copula 理论及其应用的全面介绍，可参阅 Nelsen（2006）；Joe（2015）；Durante 和 Sempi（2015）；关于近期综述，可参阅 Größer 和 Okhrin（2021）；Genest 等（2024）。

Copula 还支持一种模块化的依赖建模方法，使得统计学家能够针对每个边缘分布，分别使用最合适的工具来设定或估计边缘分布。一旦边缘分布确定，统计学家便可以专注于选择能够捕捉相关依赖性质的Copula族，例如尾部行为、非对称性或可交换性。这种灵活性使 Copula 成为保险和金融领域的标准工具，因为严谨的风险管理和情景分析需要进行细致的多元建模 Cherubini 等人（2004）；Denuit等人（2006）；McNeil 等（2015）。

一个 d-Copula 是定义在 $[0,1]^d$ 上的分布函数，具有均匀单变量边缘分布。形式上，函数 $C\colon [0,1]^d \to [0,1]$ 是 Copula，当且仅当它满足以下三个条件：

(i) 有界性，即只要 $\boldsymbol{u}$ 的至少一个分量等于 0，就有 $C(\boldsymbol{u}) = 0$；

(ii) 均匀边缘分布，即对每个 j，$C(1, \ldots, 1, u_j, 1, \ldots, 1) = u_j$；以及

(iii) d-递增，即由坐标区间定义的 C 的所有矩形增量均非负。

后一条件的明确表述是，对于所有属于 $[0,1]^d$的向量 $\boldsymbol{a} \le \boldsymbol{b}$，

$$V_C([\boldsymbol{a}, \boldsymbol{b}]) = \sum_{i_1=0}^{1} \cdots \sum_{i_d=0}^{1} (-1)^{i_1+\cdots+i_d} C(u_{1i_1}, \ldots, u_{di_d}) \ge 0, \tag{1}$$

其中 $u_{j0} = a_j$ 和 $u_{j1} = b_j$。

Copula 可通过 Sklar 定理 Sklar（1959年）将依赖性与边缘分布分离。若 H 是一个具有单变量边缘分布 F_1，……，F_d 的 d 维分布函数，则存在一个 Copula C，使得

$$H(x_1, \ldots, x_d) = C\left(F_1(x_1), \ldots, F_d(x_d)\right), \qquad (x_1, \ldots, x_d) \in \mathbb{R}^d.$$

若边缘分布连续，则该 Copula 唯一。反之，对于任意 Copula C 以及任意一组单变量边缘分布 F_1，……，F_d，右侧均定义了一个具有这些边缘分布的有效多元分布。

对高维依赖模型的需求源于现代数据的规模，以及依赖性会对风险评估和决策产生重大影响这一事实。在金融和保险领域，投资组合往往包含数百或数千份保单，从而形成高维损失分布，其中尾部事件由众多风险因子的相互作用引发（McNeil 等（2015年））。同样，在气候、水文学和空气质量等环境科学领域，跨空间和时间的多元依赖性会影响联合极端事件和复合事件的风险（Ribeiro等人（2020年）；Naseri和Hummel（2022年））。

对高维依赖结构进行建模，通常需要在灵活性与可处理性之间做出权衡。例如，藤Copula具有很强的灵活性，但难以扩展。相反，较为简单的简约族虽然参数约束较为简单，却往往限制过多；例如，阿基米德Copula仅限于可交换的依赖结构。因此，我们需要能够兼顾灵活性、可解释性和可扩展性的替代模型。Sarmanov copula颇具吸引力，因为它们通过对独立 copula施加扰动，并且许多关注量都可以得到闭式表达式；然而，随着维度增加，其可容许条件会越来越难以满足，而且据我们所知，除最简单的形式外，目前尚无其他多元Sarmanov族。为此，我们提出二元 Sarmanov copula的随机表示。基于此，我们进一步将二元随机表示扩展至高维，提出一类新的多元Sarmanov Copula

并研究其性质。在随机表示下，参数约束易于验证，模拟也简单且可扩展。

Sarmanov copula 广泛应用于风险管理，因为它们在不牺牲解析可处理性的前提下提供了灵活的依赖结构，尤其是在与相位型分布或混合 Erlang 分布结合使用时。例如，参见 Hashorva 和 Ratovomirija （ 2015 年），作者在风险服从 Sarmanov 分布时推导了聚合公式和资本分配规则。他们证明，所得的聚合随机变量仍属于混合 Erlang 分布这一类。在 Ratovomirija 等人（2017年）的研究中，作者使用了 Sarmanov-混合 Erlang 风险，并研究了聚合风险和资本分配的性质。他们还研究了核的选择如何影响可实现依赖的范围。最近，Vernic 等人（2022年）对索赔频率与损失严重程度之间的依赖性进行了建模。他们表明，忽略这种依赖性会显著扭曲保险费。迄今为止，大多数研究仍集中于二元情形，因为其中的可容许性约束更易于处理。

一些最早且使用最广泛的 Copula 族是 Farlie–Gumbel–Morgenstern（FGM）Copula，由 Eyraud（1936年）、Gumbel（1960年）和 Morgenstern（1956年）提出。随后，Farlie（1960年）引入了一类广义二元分布，该分布推广了 FGM 构造，其形式由 $C(u_1,\ u_2) = u_1u_2(1+\theta A(u_1)B(u_2))$ 给出，其中核函数 A 和B 满足一定的正则性条件，并选择标量参数 θ 以确保 C 是一个 Copula。在 Sarmanov（1966年）中，作者引入了 Farlie copula 的一个略具一般性的表述，其密度形式为

$$c(u_1,u_2) = 1 + a\phi_1(u_1)\phi_2(u_2), \qquad (u_1,u_2)\in[0,1]^2,$$

其中，对于 $j=1,\ 2,\ \int_0^1 \phi_j(u)\,\mathrm{d}u = 0$，而 a 是为确保 c 的非负性而选取的标量参数；在常见正则性条件下，这会得到与 Farlie Copula 相同的一类 Copula。我们将这一族称为萨尔马诺夫 copula，其形式为

$$C(u_1,u_2) = u_1u_2 + ag_1(u_1)g_2(u_2), \qquad (u_1,u_2)\in[0,1]^2, \tag{2}$$

其中，核函数 g_1 和 g_2 满足 $g_1(0) = g_1(1) = 0$，且 $g_2(0) = g_2(1) = 0$；标量参数 a 的选取则需满足正则性条件，以确保 C 是一个 Copula。这类模型颇具吸引力，因为其密度和依赖性度量的表达式易于处理，而且可以将其解释为对独立性的扰动。在这类 Copula 中，核函数 g_1 和 g_2 控制依赖性的形态，而参数 a 控制其强度。

然而，要确保 (2) 是有效的 Copula，就必须验证 (1) 中的二阶递增性质。实际上，对于给定的核函数 g_1 和 g_2，要找出 a 的可容许范围颇具挑战性；提出新的 Sarmanov 族的大多数研究都将大量精力投入这项工作。可容许性一经确立，Sarmanov 延拓便易于处理，但直接验证可容许性通常是最具挑战性的一步。

经典的 Farlie–Gumbel–Morgenstern copula 对应于选择 $g_1(u) = u(1-u)$ 和 $g_2(u_2) = u_2(1-u_2)$，从而得到

$$C(u_1,u_2) = u_1u_2\left(1+\theta(1-u_1)(1-u_2)\right), \qquad (u_1,u_2)\in[0,1]^2,$$

参数 $\theta\in[-1,1]$ 取值 admissible 时。该族在解析上较为简便，且秩相关的表达式也很简单。例如，在二元情形中，Kendall's τ 和 Spearman'sρ 均为 θ 的线性函数，尽管可实现依赖的最大值仅限于 $|\rho|\le 1/3$ 和$|\tau|\le 2/9$。入门教材常用 FGM 族来说明 Copula 理论，因为其结构简单且具有闭式性质，易于分析，例如 Genest and Favre（2007年）所述。

在维度 d 下，多元 FGM Copula 的多项式形式为

$$C(\boldsymbol{u}) = \left(\prod_{m=1}^{d} u_m\right)\left(1 + \sum_{k=2}^{d} \sum_{1\le j_1<\cdots<j_k\le d} \theta_{j_1\cdots j_k}\bar{u}_{j_1}\cdots\bar{u}_{j_k}\right), \qquad \boldsymbol{u}\in[0,1]^d,$$

其中 $\bar{u}_j = 1-u_j$，参数向量 $\boldsymbol{\theta}$ 有 2^d-d-1 个分量。参数可容许集并非平凡：这些参数必须满足以下 2^d 个线性不等式

$$\sum_{k=2}^{d} \sum_{1\le j_1<\cdots<j_k\le d} \varepsilon_{j_1}\cdots\varepsilon_{j_k}\theta_{j_1\cdots j_k} \ge -1. \tag{3}$$

对所有符号向量 $\boldsymbol{\varepsilon}\in\{-1,1\}^d$ 均成立。

在 Huang 和 Kotz（1999年）中，作者提出了 FGM 族的多项式扩展，旨在提高可达到的正依赖程度，同时保持闭式可处理性。他们在第 2 节中考察了族（4），并在第 3 节中考察了族（5）：

$$C(u_1,u_2) = u_1u_2\left(1+a(1-u_1^p)(1-u_2^p)\right), \quad p>1; \tag{4}$$

$$C(u_1,u_2) = u_1u_2\left(1+a(1-u_1)^q(1-u_2)^q\right), \quad q>0. \tag{5}$$

我们分别将它们称为 Huang–Kotz I（HKI）族和 Huang–Kotz II（HKII）族。这两种延拓都是具有对称核 $g_1(u)=g_2(u)=u(1-u^p)$ 和 $g_1(u)=g_2(u)=u(1-u)^q$ 的 Sarmanov copula。他们的分析表明，在这类多项式扩展中，最大可达正相关可以略微超过经典 FGM 界，但负依赖仍然有限：Spearman 的 ρ 和 Kendall 的 τ 分别限制在 $-1/3$ 和 $-2/9$。式（4）和（5）中 a 的参数可容许区域，是通过直接验证二阶递增性推导得出的。Huang 和 Kotz（1999年）的工作大部分都致力于此。[1]

在高维情况下，验证可容许性变得愈发困难。d-copula 必须是d-递增的，而这一条件即使以最明确的形式表示，也涉及组合问题。对于多元 FGM copula，参数维数随 2^d-d-1 增长，可容许区域由式（3）中的 2^d 个线性约束所定义。对于一般 Sarmanov 族，尤其是具有非对称核或多个参数的族，可容许性问题尤为棘手；我们尚未发现任何关于构造或分析多元 Sarmanov copula 的文献，简单延拓 FGM copula 的情形除外。

Blier-Wong 等（2022）的作者建立了多元 FGM copula 与具有对称边际分布的多元伯努利分布之间的一一对应关系。依赖参数与潜在伯努利向量的混合中心矩相同。具体而言，若$\boldsymbol{I}=(I_1,\ldots\ldots,I_d)$ 满足 $\mathbb{P}(I_j=1)=1/2$，则

$$C(\boldsymbol{u}) = \mathbb{E}\left[\prod_{m=1}^{d} u_m\left(1+(-1)^{I_m}\bar{u}_m\right)\right],$$

是一个参数如下的 FGM copula

$$\theta_{j_1\cdots j_k} = (-2)^k\mathbb{E}\left[\prod_{n=1}^{k}\left(I_{j_n}-\frac{1}{2}\right)\right].$$

[1]他们论文的标题是《Farlie–Gumbel–Morgenstern 分布的修正：一座难以攀登的高山》，这体现了作者推导这些可容许性约束的决心。

此外，令 $\boldsymbol{U}_{[j]} = (U_{1,\,[j]},\ \ldots\ldots,\ U_{d,\,[j]})$ 的分量独立同分布，且服从参数为 $U_{k,\,[j]} \sim$ 的贝塔分布 Beta$(2-j,\ 1+j)$，其中 $j \in \{0,1\}$，并定义 $\boldsymbol{U} = (1-\boldsymbol{I})\boldsymbol{U}_{[0]} + \boldsymbol{I}\boldsymbol{U}_{[1]}$。那么，$\boldsymbol{U}$ 具有均匀边缘分布和上述参数的 FGM copula。因此，FGM copula 是独立贝塔分布分量的潜在伯努利混合所对应的 Copula。该表示表明，$\boldsymbol{\theta}$ 的可容许性等价于多元伯努利概率质量函数在 $\{0,1\}^d$ 上的非负性，

$$f_{\boldsymbol{I}}(\boldsymbol{i}) = 2^{-d}\left(1 + \sum_{k=2}^{d}\sum_{j_1<\cdots<j_k} \theta_{j_1\cdots j_k}(-1)^{i_{j_1}+\cdots+i_{j_k}}\right) \geq 0, \qquad \boldsymbol{i} \in \{0,1\}^d,$$

并可通过从伯努利向量 $\boldsymbol{I}$ 和贝塔边缘分布中抽样，实现自然的模拟程序。

这种表示将验证 d-递增性转化为针对伯努利概率质量函数的一组有限非负约束。对于 $\boldsymbol{I}$ 的任意有效对称伯努利分布，都可立即得到可容许参数 $\boldsymbol{\theta}$。多元伯努利概率质量函数的约束已得到充分研究（Sharakhmetov 和 Ibragimov（2002）；Fontana 和 Semeraro（2018）），而各种结构化子类（可交换、马尔可夫、因子模型）则支持可扩展的推断和模拟。这种以伯努利分布为基础的视角也有助于分析随机序。后续研究利用了可交换情形下的这种关系（Blier-Wong 等（2024b））。

在 Blier-Wong 等人（2024a）中，作者提出推广 FGM 联结函数的伯努利混合表示。他们以 FGM 联结函数现有的随机表示为基础，将边缘分布从贝塔分布改为其他族，同时保留伯努利混合结构以及 Huang–Kotz I 族。具体而言，设 $\boldsymbol{I} = (I_1,\ \ldots\ldots,\ I_d)$ 是一个 d 维伯努利向量，其 $\mathbb{P}(I_m = 1) = \pi_m \in (0,\ 1)$（与 FGM 情形不同，边缘分布不必对称）。设 $\boldsymbol{U}_0 = (U_{0,\,1},\ \ldots\ldots,\ U_{0,\,d})$ 和 $\boldsymbol{U}_1 = (U_{1,\,1},\ \ldots\ldots,\ U_{1,\,d})$ 是由相互独立的均匀分布$(0,\ 1)$随机变量构成的向量，且与 $\boldsymbol{I}$相互独立。定义

$$\boldsymbol{U} \overset{\mathrm{d}}{=} \boldsymbol{U}_0^{\mathbf{1}-\boldsymbol{\pi}}\boldsymbol{U}_1^{\boldsymbol{I}} = \left(U_{0,1}^{1-\pi_1}U_{1,1}^{I_1}, \ldots, U_{0,d}^{1-\pi_d}U_{1,d}^{I_d}\right),$$

其中所有幂运算和乘积均按分量计算。此构造使 $\boldsymbol{U}$具有标准均匀边缘分布。对于 $d = 2$，$\boldsymbol{U}$ 的累积分布函数（cdf）为

$$C(u_1,u_2) = u_1u_2\left(1 + \theta_{12}\left(1 - u_1^{\frac{\pi_1}{1-\pi_1}}\right)\left(1 - u_2^{\frac{\pi_2}{1-\pi_2}}\right)\right), \qquad (u_1,u_2) \in [0,1]^2,$$

with

$$\theta_{12} = \mathbb{E}\left[\frac{I_1-\pi_1}{\pi_1}\frac{I_2-\pi_2}{\pi_2}\right].$$

在对称情形下，$\pi_1 = \pi_2$ 与式（4）中的 Huang–Kotz I copula 族相同。另请参阅 Cossette 等（2025），了解对该构造的高维分析和几何解释。

由于 FGM 和 Huang–Kotz I Copula 族都可以采用这种伯努利混合表示，因此一个自然的问题是：这种构造能否扩展到更一般的 Sarmanov copula，包括非对称核以及其他 Huang–Kotz 变体？本文的主要结果表明，答案确实是肯定的。我们为形如 $C(u_1,u_2) = u_1u_2 + ag_1(u_1)g_2(u_2)$ 的 Sarmanov copulas 构建了统一随机表示。主定理表明，在核函数 g_1 和 g_2满足温和正则性条件的情况下，任何二元 Sarmanov copula 都可以表示为两个边缘分布之间伯努利混合分布的累积分布函数，从而可容许性归结为对二元伯努利概率质量函数施加有限个非负性约束，而不再需要处理大量

不等式。这为构造和模拟非对称 Sarmanov 连接函数提供了一种构造性方法，并给出了一个统一的概率机制，将经典族（包括 FGM 和 Huang–Kotz 型扰动）作为特殊情形涵盖其中。随后，我们在高维情形中引入伯努利混合构造，得到具有广义 FGM 型子集展开的 Copula，从而支持可扩展的模拟和具有可解释性的依赖建模；同时，我们研究了关键子类中的极值依赖。最后，我们将该框架扩展到幂次扰动 $C(u_1, u_2) = u_1u_2(1 + ah_1(u_1)h_2(u_2))^r$，其中 $r > 0$，从而得到新的易处理族以及相应的依赖界和表示。

Sarmanov 连接函数族与文献中以不同名称出现的其他构造密切相关。本文研究的 Sarmanov copula 与 Farlie 连接函数类非常接近：Farlie Farlie (1960) 考虑了形如 $C(u_1, u_2) = u_1u_2(1 + ah_1(u_1)h_2(u_2))$ 的 copula。Farlie Copula 是 Sarmanov copula 的特例；在常见的正则性条件下，当 $uh_1(u)$ 和 $uh_2(u)$ 绝对连续且具有有界导数时，它涵盖了几乎所有 Sarmanov copula。Sarmanov 构造也由 Rodríguez-Lallena 和 Úbeda-Flores (2004) 重新发现。在常见的正则性条件下，经过参数化变换后，他们的定义与之相同。为清晰起见，本文始终将这一族称为 “Sarmanov” copula，即便 Farlie 表示法早于 Sarmanov 的研究。在附录 C 中，我们详细比较了这三种表述，并说明它们等价的确切条件。请注意，本文的主要结果适用于这三个族，因此为这些密切相关的构造提供了统一的随机表示。

本文其余部分的安排如下。第 2 节回顾 Sarmanov copula，并给出建立伯努利混合表示的主定理，阐释文献中现有 copula 的随机表示，并确定二元 Sarmanov 类中可达到的关联系数。第 3 节介绍高维中的伯努利混合构造，建立一类具有 FGM 型子集展开的多元Sarmanov Copula。第 4 节将该构造延拓至幂次双变量 Sarmanov 连接函数。第 5 节总结本文贡献并讨论延拓。附录 A 包含主要结果的证明；附录 B 提供 Sarmanov copula 及其随机表示的其他示例，而附录 C 将 Sarmanov copula 与文献中的相关构造进行比较。

2 Sarmanov copula 与随机表示

本节针对具有伯努利分布的广义二元 Sarmanov copula，构建其随机表示。我们将证明，Sarmanov copula 可以表示为具有适当边缘分布的独立随机变量的潜在伯努利混合所对应的 copula。该构造推广了上文讨论的 FGM 和 Huang–Kotz I 连接函数的伯努利混合表示。

2.1 Sarmanov copula

考虑对乘积 copula 的经典萨尔马诺夫扰动，

$$C(u_1, u_2) = u_1u_2 + ag_1(u_1)g_2(u_2), \qquad (u_1, u_2) \in [0, 1]^2,$$

其中核函数 $g_1, g_2 : [0, 1] \to \mathbb{R}$ 可以不同（因此可能不对称）。即使在二维中，确定使 C 成为 Copula 的 a 可容许取值，通常也要通过验证二增性质来实现。在绝对连续的情况下，这可归结为验证一个

在 $[0,1]^2$, 上的混合导数非负，但直接进行验证仍可能很繁琐，因为可容许区间取决于某些核导数的具体取值范围。

我们假设 g_1 和 g_2 在 $[0,1]$, 上绝对连续，满足 $g_1(0) = g_1(1) = 0$ 以及 $g_2(0) = g_2(1) = 0$，并且几乎处处存在导数 ϕ_1, $\phi_2 \in L^\infty([0,1])$。定义 $\phi_1(u) = g_1'(u)$ 和 $\phi_2(v) = g_2'(v)$（几乎处处），并定义常数

$$\Lambda_1 = \frac{1}{\operatorname{ess\,sup}_{u\in[0,1]} \phi_1(u)} \in (0,\infty), \qquad \lambda_1 = \frac{1}{\operatorname{ess\,inf}_{u\in[0,1]} \phi_1(u)} \in (-\infty,0),$$

$$\Lambda_2 = \frac{1}{\operatorname{ess\,sup}_{v\in[0,1]} \phi_2(v)} \in (0,\infty), \qquad \lambda_2 = \frac{1}{\operatorname{ess\,inf}_{v\in[0,1]} \phi_2(v)} \in (-\infty,0).$$

由于绝对连续性意味着 $\lambda_1 < 0 < \Lambda_1$ $0 = g_1(1) - g_1(0) = \int_0^1 \phi_1(u)\mathrm{d}u$, 因此 ϕ_1 必须同时取正值和负值，除非它恒等于零（此时 C 退化为独立 copula）。同样的推理也适用于 ϕ_2。在这些假设下，Sarmanov copula 的混合导数为

$$\frac{\partial^2}{\partial u_1 \partial u_2} C(u_1,u_2) = 1 + a\phi_1(u_1)\phi_2(u_2),$$

因此（在绝对连续的情形下）， C 是 Copula，当且仅当 $1 + a\phi_1(u_1)\phi_2(u_2) \geq 0$ 几乎处处成立。这给出了 a的一个明确可容许区间。不过，下面主定理的要点在于，我们无需直接分析这些不等式：随机表示会自动确保可容许性。

在下文中，我们将使用一个二元伯努利向量 (I_1, I_2)，其成功概率分别为 $\mathbb{P}(I_1 = 1) = \pi_1$ 和 $\mathbb{P}(I_2 = 1) = \pi_2$. 。其依赖结构将通过一个参数来指定，该参数为中心矩

$$\theta = \mathbb{E}\left[\frac{(I_1 - \pi_1)(I_2 - \pi_2)}{\pi_1\pi_2}\right] = \frac{\operatorname{Cov}(I_1, I_2)}{\pi_1\pi_2}. \tag{6}$$

由此可得，(I_1, I_2) 的概率质量函数为

$$\mathbb{P}(I_1 = i_1, I_2 = i_2) = \begin{cases} \pi_1\pi_2(1+\theta), & (i_1,i_2) = (1,1) \\ \pi_1(1-\pi_2) - \pi_1\pi_2\theta, & (i_1,i_2) = (1,0) \\ (1-\pi_1)\pi_2 - \pi_1\pi_2\theta, & (i_1,i_2) = (0,1) \\ (1-\pi_1)(1-\pi_2) + \pi_1\pi_2\theta, & (i_1,i_2) = (0,0) \end{cases}. \tag{7}$$

2.2 主结果：Sarmanov copula 的随机表示

现在我们可以陈述主定理了。该定理在上述正则性条件下，为一般的二元 Sarmanov copula 建立了伯努利混合表示。证明见附录 A。

定理 1. 假设 g_1 和 g_2 满足第*2.1*节中所述的绝对连续性和有界导数条件。定义缩放常数 $\pi_1 = \Lambda_1/(\Lambda_1 - \lambda_1) \in (0,1)$ 和 $\pi_2 = \Lambda_2/(\Lambda_2 - \lambda_2) \in (0,1)$。对每个边缘分布，在 $[0,1]$ 上定义两个辅助累积分布函数：

$$F_{U,[0]}(u) = u - \Lambda_1 g_1(u), \qquad F_{U,[1]}(u) = u - \lambda_1 g_1(u),$$

$$F_{V,[0]}(v) = v - \Lambda_2 g_2(v), \qquad F_{V,[1]}(v) = v - \lambda_2 g_2(v).$$

令 $(U_{[0]}, U_{[1]}, V_{[0]}, V_{[1]})$ 相互独立，其中 $U_{[i]} \sim F_{U,[i]}$ 和 $V_{[i]} \sim F_{V,[i]}$ 满足 $i \in \{0,1\}$。令 (I_1, I_2) 为一个与上述变量独立的二元伯努利向量，其参数化方式见(7)。考虑由下式定义的随机向量 (U_1, U_2)

$$U_1 = (1 - I_1)U_{[0]} + I_1 U_{[1]}, \qquad U_2 = (1 - I_2)V_{[0]} + I_2 V_{[1]}.$$

那么，

1. U_1 ~均匀分布(0，1) 和 U_2 ~均匀分布(0，1)。

2。(U_1, U_2) 的 ***Copula*** 满足(2)，其中 $a = \Lambda_1\Lambda_2\theta$， θ 由(6)给出。

定理 1 的一个直接推论是，确保 (2) 为有效 Copula 的 a 可容许范围，由参数为 $\theta = a/(\Lambda_1\Lambda_2)$ 的二元伯努利分布是否存在来确定。要构造有效的非对称 Sarmanov Copula $C(u_1, u_2) = u_1u_2 + ag_1(u_1)g_2(u_2)$，可以 (i) 计算 Λ_1, Λ_2, λ_1, λ_2，(ii) 计算 π_1, π_2 以及辅助累积分布函数 $F_{U,[0]}$, $F_{U,[1]}$, $F_{V,[0]}$, $F_{V,[1]}$，并且 (iii) 在可容许范围内选择 a，使得 $\theta = a/(\Lambda_1\Lambda_2)$ 通过 (7) 得到有效的二元伯努利概率质量函数。由此，验证二阶递增性质可转化为检查四个伯努利概率质量函数项的非负性； a 的可行性等价于参数为 $\theta = a/(\Lambda_1\Lambda_2)$ 的二元伯努利分布存在。伯努利概率质量函数的约束已得到充分研究；例如，参见 Fontana 和 Semeraro（2018）。这种方法为构造和模拟非对称 Sarmanov 连接函数提供了一种构造性方法。等价地，我们可直接得到

$$a \in \left[-\min\left(\Lambda_1\Lambda_2, \lambda_1\lambda_2\right), \min\left(\Lambda_1|\lambda_2|, \Lambda_2|\lambda_1|\right)\right].$$

此外，如果 $g_1 = g_2 = g$，那么 $\phi_1 = \phi_2 = \phi$，从而可恢复对称萨尔马诺夫类 $C(u_1, u_2) = u_1u_2 + ag(u_1)g(u_2)$。可容许范围缩减为

$$a \in \left[-\min\left(\lambda^2, \Lambda^2\right), \Lambda|\lambda|\right],$$

其中 $\Lambda^{-1} = \sup \phi$ 且 $\lambda^{-1} = \inf \phi$，并且同样适用伯努利混合构造。

定理 1 还提供了对萨尔马诺夫类的几何解释。对于固定的核函数 g_1， g_2（意味着固定的 π_1, π_2, Λ_1, Λ_2），可容许性归结为：为具有给定边缘分布的二元伯努利随机向量（I_1, I_2）选择一个分布。Copula 是通过随机表示所定义的仿射映射，将（I_1, I_2）的分布映射后得到的像。由于具有给定边缘分布的二元伯努利随机向量的弗雷歇类是一个恰好包含两个极点的线段，相应的 Sarmanov copula 族本身就是 Copula 凸集中的一条线段：每个可容许的 Copula 都可以表示为两个极值 Copula C^- 和 C^+ 的凸组合，这两个极值 Copula 由极值伯努利耦合得到，其混合权重唯一，且是 $\theta = a/(\Lambda_1\Lambda_2)$ 的仿射函数（等价地，也是 a 的仿射函数）。

现在，我们通过几个例子来阐释定理 1，推导出经典萨尔马诺夫 copula（如 FGM copula、Huang 和 Kotz（1999年）的黄–科茨 copula，以及 Bairamov 等人（2001年）的拜拉莫夫–科茨–贝克奇 copula）的已知随机表示。

例 1（FGM）。取 $g_1(u) = g_2(u) = u(1-u)$，使得 $C(u_1, u_2) = u_1u_2(1 + a(1-u_1)(1-u_2))$，这就是二元 ***FGM copula***。我们有 $\phi(u) = g'(u) = 1 - 2u$，因此 $\Lambda = 1$ 且 $\lambda = -1$。于是 $\pi = 1/2$，辅助累积分布函数变为

$$F_{U,[0]}(u) = u - u(1-u) = u^2, \qquad F_{U,[1]}(u) = u + u(1-u) = 2u - u^2.$$

这正是 ***Blier-Wong*** 等（*2022*）提出的随机表示，可容许范围变为 $-1 \le a \le 1$，与经典 ***FGM*** 可容许范围表述一致。我们还有 $U_{[0]}$ ~贝塔分布(2，1）以及 $U_{[1]}$ ~贝塔分布(1，2)。

例 2（*Huang–Kotz I*）。取 $g_1(u) = g_2(u) = u(1-u^p)$，使得 $C(u_1,u_2) = u_1u_2(1+a(1-u_1^p)(1-u_2^p))$。由于 $\sup \phi = 1$ 在 $u=0$ 处取得，我们有 $\Lambda = 1$，而 $\lambda = -1/p$ 在 $u=1$ 处取得。因此$\pi = \Lambda/(\Lambda-\lambda) = p/(p+1)$，辅助累积分布函数为

$$F_{[0]}(u) = u - u(1-u^p) = u^{p+1}, \qquad F_{[1]}(u) = u + u(1-u^p)/p = u\left(1 + \frac{1-u^p}{p}\right).$$

可容许范围为 $a \in [-\min(1/p^2, 1)\ 1/p]$。这正是 *Blier-Wong* 等人（*2024a*）给出的随机表示。或者，令 $V, W \sim$ 服从均匀分布(0,1) 且相互独立，则 $U_{[0]} \overset{\mathrm{d}}{=} V^{\frac{1}{p+1}}$，因此 $U_{[0]} \sim$服从贝塔分布$(\overline{\quad}$ $p+1, 1)$，且 $U_{[1]} \overset{\mathrm{d}}{=} WV^{\frac{1}{p+1}}$.

例 3（*Huang–Kotz II*）。取 $g_1(u) = g_2(u) = u(1-u)^q$，使得 $C(u_1,u_2) = u_1u_2(1+a(1-u_1)^q(1-u_2)^q)$。我们有 $\phi_1(u) = \phi_2(u) = (1-u)^{q-1}(1-(q+1)u)$。假设 $q > 1$，使得 ϕ 有界。我们有 $\sup \phi = 1$ 在 $u=0$ 处，因此 $\Lambda = 1$，并且

$$\lambda = -\left(\frac{q+1}{q-1}\right)^{q-1}.$$

因此 $\pi = \Lambda/(\Lambda-\lambda) = 1/(1-\lambda)$。辅助累积分布函数为

$$F_{[0]}(u) = u - \Lambda g(u) = u - u(1-u)^q = u\left(1-(1-u)^q\right),$$

$$F_{[1]}(u) = u - \lambda g(u) = u - \lambda u(1-u)^q = u\left(1-\lambda(1-u)^q\right).$$

在这一对称情形下，可容许范围为 $a \in [-1, |\lambda|]$。对于 *Huang–Kotz II* 族而言，该随机表示是一个新结果。另请注意，$U_{[0]}$ 具有一个简单的表示：如果V，$W \sim$均匀分布*(0, 1)*且相互独立，并且 $Z := 1 - V^{1/q} \sim$服从贝塔分布 *(1, q)*，那么 $U_{[0]} \overset{d}{=} \max(W, Z)$，因为 $\mathbb{P}(\max(W,Z) \le u) = u(1-(1-u)^q) = F_{[0]}(u)$。相比之下，$U_{[1]}$ 似乎无法通过初等变换得到同样简单的闭式表示。一个例外是情形 $q=2$。令 $W \sim$ 服从均匀分布*(0, 1)*，则有

$$U_{[1]} \overset{\mathrm{d}}{=} \frac{2}{3} + \sqrt[3]{\frac{W - \frac{8}{9}}{3}}.$$

例 4（Bairamov–Kotz–Bekçi 核）。取 $g_1(u) = u(1-u^{p_1})^{q_1}$ 和 $g_2(v) = v(1-v^{p_2})^{q_2}$，使得 $C(u_1,u_2) = u_1u_2(1+a(1-u_1^{p_1})^{q_1}(1-u_2^{p_2})^{q_2})$。该族由 *Bairamov* 等人（*2001*年）提出，是 *Huang–Kotz I* 族的扩展。对于 $k \in \{1,2\}$，我们有

$$\phi_k(t) = g_k'(t) = (1-t^{p_k})^{q_k-1}\left(1-(1+p_kq_k)\,t^{p_k}\right).$$

假设 $q_1 > 1$ 和 $q_2 > 1$，使得 ϕ_1 和 ϕ_2 有界。那么 上确界 $\phi_1 =$ 上确界 $\phi_2 = 1$，因此$\Lambda_1 = \Lambda_2 = 1$，并且

$$\lambda_k = -\frac{(1+p_kq_k)^{q_k-1}}{p_k^{q_k}\,(q_k-1)^{q_k-1}}, \qquad k \in \{1,2\}.$$

因此 $\pi_1 = 1/(1-\lambda_1)$ 和 $\pi_2 = 1/(1-\lambda_2)$，辅助累积分布函数为

$$F_{U,[0]}(u) = u\left(1-(1-u^{p_1})^{q_1}\right), \qquad F_{U,[1]}(u) = u\left(1-\lambda_1(1-u^{p_1})^{q_1}\right),$$

$$F_{V,[0]}(v) = v\left(1-(1-v^{p_2})^{q_2}\right), \qquad F_{V,[1]}(v) = v\left(1-\lambda_2(1-v^{p_2})^{q_2}\right).$$

可容许范围为

$$a \in \left[-\min\left(1, \frac{(1+p_1q_1)^{q_1-1}}{p_1^{q_1}\,(q_1-1)^{q_1-1}} \frac{(1+p_2q_2)^{q_2-1}}{p_2^{q_2}\,(q_2-1)^{q_2-1}}\right), \min\left(\frac{(1+p_1q_1)^{q_1-1}}{p_1^{q_1}\,(q_1-1)^{q_1-1}}, \frac{(1+p_2q_2)^{q_2-1}}{p_2^{q_2}\,(q_2-1)^{q_2-1}}\right)\right].$$

分布 $F_{U,[0]}$ 具有显式表示。设 $V, W \sim$ 均匀分布相互独立，并定义 $S := \left(1 - V^{1/q_1}\right)^{1/p_1}$。那么 $\mathbb{P}(S \le u) = 1 - (1 - u^{p_1})^{q_1}$ 对 $u \in [0,1]$成立，因此$U_{[0]}$ $d= \max(W, S)$，因为 $\mathbb{P}(\max(W, S) \le u) = u(1 - (1 - u^{p_1})^{q_1}) = F_{U,[0]}(u)$。 $V_{[0]}$也有类似的表示。相比之下，累积分布函数 $F_{U,[1]}$ 和 $F_{V,[1]}$ 似乎无法得到简单的表示。

2.3 极值斯皮尔曼 rho 和肯德尔秩相关系数

本小节量化了在式（2）的 Sarmanov copula 中能够达到的最大依赖程度，分别以斯皮尔曼 ρ_S 和肯德尔 τ衡量。我们确定了精确的全局界（遍历所有可容许核和所有可容许依赖参数），并找到了极值核；结果表明，这些核属于“二块”（棋盘格）类型。

二元斯皮尔曼秩相关系数为

$$\rho_S = \mathrm{Corr}(U_1, U_2) = 12 \int_{[0,1]^2} u_1 u_2 \,\mathrm{d}C(u_1, u_2) - 3 = 12 \int_{[0,1]^2} (C(u_1, u_2) - u_1 u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2. \tag{8}$$

肯德尔的 τ 定义为一致性概率与不一致性概率之差，即 $\tau = \mathbb{P}((X - X')(Y - Y') > 0) - \mathbb{P}((X - X')(Y - Y') < 0)$，其中 (X', Y') 是 (X, Y) 的一个独立副本。如果 C 具有密度 c，则

$$\tau = 4 \int_{[0,1]^2} C(u_1, u_2) c(u_1, u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 - 1.$$

注 1. 在 *Fredricks* 和 *Nelsen*（*2007*年）中，作者指出，斯皮尔曼 *rho* 与肯德尔秩相关系数满足恒等式 $3\tau = 2\rho_S$ 的 *Copula* 满足偏微分方程

$$\frac{\partial^2}{\partial u_1 \partial u_2} \log |C(u_1, u_2) - u_1 u_2| = 0,$$

只要 $C(u_1, u_2) \neq u_1 u_2$。*Sarmanov copula* 满足该偏微分方程；因此，找到一个一致性度量就足以确定另一个。

命题 1。设 $G_1 := \int_0^1 g_1$ 以及 $G_2 := \int_0^1 g_2$。则

$$\rho_S(C) = 12 \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 = 12 a G_1 G_2. \tag{9}$$

证明。我们有 $C(u_1, u_2) - u_1 u_2 = a g_1(u_1) g_2(u_2)$，应用 Fubini–Tonelli 定理可得

$$\int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 = a \left(\int_0^1 g_1(u_1)\,\mathrm{d}u_1\right)\left(\int_0^1 g_2(u_2)\,\mathrm{d}u_2\right) = a G_1 G_2,$$

由此可证得（9）。 □

由于 $\rho_S(C) = 12aG_1G_2$，求出 ρ_S 的极值可归结为理解：（i） a 的可取幅度；以及（ii）在 g_1 和 g_2 的斜率约束下，带符号“面积” G_1 和 G_2 的最大可能值。下述定理确定了第 2.1 节所述绝对连续性和有界导数条件下，所有形如（2）的萨尔马诺夫 copula 上 ρ_S 的全局极值。证明见附录 A。Shubina and Lee（2004年）此前已证明这一结果，但我们以更符合本研究设定的方式重新推导。我们不直接对所有可取的 Sarmanov核函数进行优化，而是通过其随机表示来考察该模型，并从所得结构中确定紧确界。

定理2。令 C 为式(2) 中的 ***Sarmanov copula***，其核函数 g_1、g_2 满足第*2.1*节所述的绝对连续性和有界导数条件。则

$$-\frac{3}{4} \le \rho_S(C) \le \frac{3}{4}. \tag{10}$$

这些界是紧的。

定理2中的极值由棋盘核取得。上述核函数可用一种方便的伯努利混合解释来说明。取相互独立的$U_{1,\ [1]} \sim$ 均匀分布$(0,\ 1/2)$、$U_{1,\ [0]}$ ~均匀分布$(1/2,\ 1)$，以及类似地取 U_2，并令 $(I_1,\ I_2)$ 为一个满足$\mathbb{P}(I_1 = 1) = \mathbb{P}(I_2 = 1) = 1/2$ 且 $\mathbb{P}(I_1 = 1,\ I_2 = 1) = (1+\theta)/4$的二元伯努利向量。则，$U_1 = (1 - I_1)U_{1,\ [0]} + I_1U_{1,\ [1]}$和 $U_2 = (1 - I_2)U_{2,\ [0]} + I_2U_{2,\ [1]}$ 具有均匀边缘分布，其 Copula 为

$$C_\theta(u_1, u_2) = u_1u_2 + \frac{\theta}{4}\left(1 - |1 - 2u_1|\right)\left(1 - |1 - 2u_2|\right),$$

且 $\rho_S(C_\theta) = 3\theta/4$。极值 $\rho_S = \pm 3/4$ 在 $\theta = \pm 1$处取得。

2.4 尾部依赖

对于二元copula C，定义下尾和上尾依赖系数（当极限存在时）为

$$\lambda_L(C) := \lim_{u\downarrow 0} \frac{C(u,u)}{u}, \qquad \lambda_U(C) := \lim_{u\uparrow 1} \frac{1 - 2u + C(u,u)}{1-u}.$$

命题2。对于满足第*2.1*节中所述绝对连续和有界导数条件的任意形式为 (2) 的 ***Sarmanov copula*** C，有 $\lambda_L(C) = \lambda_U(C) = 0$，也就是说，尾部依赖参数恒等为零（与核函数无关）。

证明。令 $L_k := \|g_k'\|_\infty = \max(-1/\lambda_k, 1/\Lambda_k) < \infty$，其中 $k \in \{1,2\}$。由于 $g_k(0) = g_k(1) = 0$ 且 g_k 是绝对连续的，

$$|g_k(u)| = \left|\int_0^u g_k'(t)\,\mathrm{d}t\right| \le L_k u, \qquad |g_k(1-u)| = \left|\int_{1-u}^1 g_k'(t)\,\mathrm{d}t\right| \le L_k u, \qquad u \in [0,1].$$

利用 $C(u,u) = u^2 + ag_1(u)g_2(u)$，可得

$$0 \le \frac{C(u,u)}{u} = u + a\frac{g_1(u)g_2(u)}{u} \le u + |a|\frac{(L_1u)(L_2u)}{u} = (1 + |a|L_1L_2)u,$$

当 $u \downarrow 0$ 时趋于零，因此 $\lambda_L(C) = 0$。对于上尾，令 $u = 1 - u$，并注意到 $1 - 2u + C(u,\ u) = u^2 + ag_1(1-u)g_2(1-\overline{u})$，因此

$$0 \le \frac{1 - 2u + C(u,u)}{1-u} = \overline{u} + a\frac{g_1(1-\overline{u})g_2(1-\overline{u})}{\overline{u}} \le \overline{u} + |a|\frac{(L_1\overline{u})(L_2\overline{u})}{\overline{u}} = (1 + |a|L_1L_2)\overline{u},$$

当 $u \downarrow 0$ 趋于零时，因此 $\lambda_U(C) = 0$。 □

3 Sarmanov copula 的多元延拓

虽然第 2 节讨论的二元萨尔马诺夫扰动已经涵盖了广泛的非对称核，但现代应用通常需要高维依赖模型。原则上，可以写出一个 d 元扰动，然后直接验证 d 递增性质。然而，这在概念和技术上都颇具挑战。在实践中，这很快就会变得难以处理： d 递增约束是组合的（它们对应于所有 2^d 包含—排除增量的非负性），而一旦 $d \geq 3$，验证这些约束通常十分繁琐。Blier-Wong 等人（2022年、2024a）提出的随机表示的关键优势在于，无需再验证 d 递增性质：只需利用基于低维核的伯努利混合机制，将 $\boldsymbol{U}$ 构造为具有均匀边缘分布的随机向量；由此得到的联合累积分布函数便自动构成有效的 Copula。从 $d = 2$ 推广到一般的 d 在概念上并不复杂，只需将二元伯努利指标替换为 d 元伯努利随机向量。

3.1 一般多元构造

以下定理提出了 bivariate Sarmanov 类在 d 维空间中的 d 元延拓，该延拓基于一个 d 元伯努利混合。其原理是：在每个坐标中混合两个一维分量，从而构造一个具有标准均匀边缘分布的随机向量 $\boldsymbol{U}$，并利用潜在多元伯努利向量编码所有依赖性。由于 $\boldsymbol{U}$ 具有均匀边缘分布，其联合累积分布函数自动构成一个 Copula。

我们首先引入 π-校准的累积分布函数对的概念，该概念刻画了定理 1 中的辅助累积分布函数。

定义 1 （π-校准对）。 固定 $\pi \in$ （0,1）。将定义在 $[0,1]$ 上的两个累积分布函数 $F_{[0]}$ 和 $F_{[1]}$ 称为 π-校准的，如果它们的混合分布是均匀分布，即它们满足

$$(1-\pi)F_{[0]}(u) + \pi F_{[1]}(u) = u, \qquad \forall u \in [0,1]. \tag{11}$$

设 $I \sim$ Bernoulli(π) 与 $(U_{[0]}, U_{[1]})$ 相互独立，且 $U_{[0]} \sim F_{[0]}$ 和 $U_{[1]} \sim F_{[1]}$。如果 $(F_{[0]}, F_{[1]})$ 是一个 π-校准对，则混合分布 $(1-I)U_{[0]} + IU_{[1]}$ 在 $[0,1]$ 上服从均匀分布。

下文中，我们令 $\Pi(\boldsymbol{u}) = \prod_{m=1}^{d} u_m$ 表示 d 维独立 copula。令 $d \geq 2$，并记 $[d] := \{1, \ldots, d\}$。令 $\boldsymbol{I} = (I_1, \ldots, I_d) \in \{0,1\}^d$ 为一个多元伯努利向量，其 $\mathbb{P}(I_m = 1)$ =为： $\pi_m \in (0,1)$ $m \in [d]$.。对于每个 $m \in [d]$, 选取一个由 $[0,1]$, 支持的 π_m 校准对 $(F_{m,[0]}, F_{m,[1]})$，并令 $U_{m,[0]} \sim F_{m,[0]}$ 和 $U_{m,[1]} \sim F_{m,[1]}$.。假设集合 $\{U_{m,[0]}, U_{m,[1]} : m \in [d]\}$ 相互独立，且与 $\boldsymbol{I}$ 独立。定义

$$U_m := (1-I_m)U_{m,[0]} + I_m U_{m,[1]}, m \in [d], \tag{12}$$

并记 $\boldsymbol{U} = (U_1、\ldots\ldots、U_d)$。为后续使用，定义累积分布函数间隙和核因子

$$\Delta_m(u) := F_{m,[1]}(u) - F_{m,[0]}(u); \quad g_m(u) := \pi_m \Delta_m(u), \qquad u \in [0,1].$$

对于每个满足 $|S| \geq 2$ 的子集 $S \subseteq [d]$，定义归一化混合矩

$$\theta_S := \mathbb{E}\left[\prod_{m \in S} \frac{I_m - \pi_m}{\pi_m}\right]. \tag{13}$$

定理 3. 设 $\boldsymbol{U} = (U_1, \ldots, U_d)$ 如 (12) 中所定义。那么，$\boldsymbol{U}$ 具有 *Copula* C，它对所有 $\boldsymbol{u} = (u_1, \ldots, u_d) \in [0,1]^d$ 均满足，

$$C(\boldsymbol{u}) = \Pi(\boldsymbol{u}) + \sum_{\substack{S \subseteq [d] \\ |S| \geq 2}} \theta_S \left(\prod_{m \notin S} u_m \right) \left(\prod_{m \in S} g_m(u_m) \right). \tag{14}$$

证明。固定 $m \in [d]$。由于 $U_m = (1 - I_m) U_{m,\,[0]} + I_m U_{m,\,[1]}$ 且 I_m 与 $(U_{m,\,[0]},\ U_{m,\,[1]})$ 相互独立，根据 π_m-校准条件 (11)，我们有 $\mathbb{P}(U_m \leq u) = (1 - \pi_m) F_{m,\,[0]}(u) + \pi_m F_{m,\,[1]}(u) = u$。因此，对于每个 m，$U_m \sim$ 服从均匀分布 Unif(0，1)。对于 $u \in [0,1]^d$，在给定 $\boldsymbol{I}$的条件下，各坐标相互独立，因此

$$C(\boldsymbol{u}) = \mathbb{P}(\boldsymbol{U} \leq \boldsymbol{u}) = \mathbb{E}\left[\prod_{m=1}^{d} F_{m,[I_m]}(u_m) \right]. \tag{15}$$

令 $Z_m := (I_m - \pi_m)/\pi_m$，使得 $\mathbb{E}[Z_m] = 0$。由 (11) 以及 $\Delta_m = F_{m,\,[1]} - F_{m,\,[0]}$，我们有$F_{m,\,[0]}(u) = u - \pi_m \Delta_m(u)$，且 $F_{m,\,[1]}(u) = u + (1 - \pi_m)\Delta_m(u)$，因此，对于 $I_m \in \{0,1\}$，

$$F_{m,[I_m]}(u) = u + (I_m - \pi_m)\Delta_m(u) = u + g_m(u) Z_m. \tag{16}$$

代入（15）得

$$C(\boldsymbol{u}) = \mathbb{E}\left[\prod_{m=1}^{d} (u_m + g_m(u_m) Z_m) \right].$$

对 $S \subseteq [d]$ 上的乘积展开，得

$$\prod_{m=1}^{d} (u_m + g_m(u_m) Z_m) = \Pi(\boldsymbol{u}) + \sum_{\substack{S \subseteq [d] \\ |S| \geq 1}} \left(\prod_{m \notin S} u_m \right) \left(\prod_{m \in S} g_m(u_m) \right) \left(\prod_{m \in S} Z_m \right).$$

其中 $|S| = 0$项为 Π（$\boldsymbol{u}$），而 $|S| = 1$项均为零，因为 $\mathbb{E}[Z_m] = 0$。取期望得（14），其中系数 θ_S 由（13）定义。 □

表示式（14）给出了广义 FGM 型子集展开，其中依赖性参数（13）由伯努利向量 $\boldsymbol{I}$的中心化混合矩给出，推广了（6）中的双变量协方差型参数。当 $d = 2$时，（14）恰好化简为双变量 Sarmanov 形式（2）。对于 $d \geq 3$，一般会出现高阶交互项；只要参数 θ_S 与 $\boldsymbol{I}$的有效伯努利概率质量函数相容，且核函数 g_m 来自 π_m校准对，（14）中的 Copula 就自动有效。

3.2 三变量 Sarmanov copula

对于 $d = 3$，子集展开式变得格外清晰，并已展示出成对交互项和三元交互项的出现。以下是定理 3 的直接特例。

推论 1. 设$(I_1, I_2, I_3) \in \{0,1\}^3$ 为一个三元伯努利向量，满足 $\mathbb{P}(I_1 = 1) = \pi_1$, $\mathbb{P}(I_2 = 1) = \pi_2$，以及 $\mathbb{P}(I_3 = 1) = \pi_3$。按照定理 *3* 中的方法，通过(12) 构造$(U_1, U_2, U_3)$，其核函数为$g_1, g_2, g_3$。则$(U_1, U_2, U_3)$ 的 *Copula* 为

$$\begin{aligned} C(u_1, u_2, u_3) = u_1 u_2 u_3 &+ \theta_{12} g_1(u_1) g_2(u_2) u_3 + \theta_{13} g_1(u_1) u_2 g_3(u_3) + \theta_{23} u_1 g_2(u_2) g_3(u_3) \\ &+ \theta_{123} g_1(u_1) g_2(u_2) g_3(u_3), \end{aligned} \tag{17}$$

其中

$$\theta_{12}=\frac{\mathbb{E}\left[(I_1-\pi_1)(I_2-\pi_2)\right]}{\pi_1\pi_2},\qquad \theta_{13}=\frac{\mathbb{E}\left[(I_1-\pi_1)(I_3-\pi_3)\right]}{\pi_1\pi_3},$$
$$\theta_{23}=\frac{\mathbb{E}\left[(I_2-\pi_2)(I_3-\pi_3)\right]}{\pi_2\pi_3},\qquad \theta_{123}=\frac{\mathbb{E}\left[(I_1-\pi_1)(I_2-\pi_2)(I_3-\pi_3)\right]}{\pi_1\pi_2\pi_3}.$$

因此，无需直接检验3递增性质：(17) 是一个 ***Copula***，因为(U_1, U_2, U_3) 是一个具有标准均匀边缘分布的随机向量。

因此，我们可以通过在每个边缘分布中选择适当的π校准对，并为(I_1, I_2, I_3)选择适当的三元伯努利概率质量函数，构造出多种多样的三元 Sarmanov Copula。作为一个具体示例，我们现在将例 4 中的二元 BKB 族扩展为三元形式。设 $g_1(u)=u(1-u^{p_1})^{q_1}$、$g_2(u)=u(1-u^{p_2})^{q_2}$ 和 $g_3(u)=u(1-u^{p_3})^{q_3}$，则有

$$\begin{aligned}C(u_1,u_2,u_3)=u_1u_2u_3\big(1&+\theta_{12}(1-u_1^{p_1})^{q_1}(1-u_2^{p_2})^{q_2}+\theta_{13}(1-u_1^{p_1})^{q_1}(1-u_3^{p_3})^{q_3}\\&+\theta_{23}(1-u_2^{p_2})^{q_2}(1-u_3^{p_3})^{q_3}+\theta_{123}(1-u_1^{p_1})^{q_1}(1-u_2^{p_2})^{q_2}(1-u_3^{p_3})^{q_3}\big).\end{aligned}\tag{18}$$

参数 $\theta_{12},\theta_{13},\theta_{23},\theta_{123}$ 必须与(I_1, I_2, I_3)的有效三元伯努利概率质量函数相容，这一点利用 Teugels（1990年）的一些结果即可轻松验证。

3.3 可交换的 Sarmanov copula

许多应用要求依赖模型具有可交换性，即在对坐标进行排列时保持不变。在定理 3 的伯努利混合构造中，可以通过对 (i) 边缘分布 π 校准对和 (ii) 潜在伯努利指标向量施加简单的对称性约束，来实现可交换性（以及在需要时实现径向对称性）。

定义 2。 如果一个 d-***Copula*** C 满足

$$C(u_1,\dots,u_d)=C\left(u_{\sigma(1)},\dots,u_{\sigma(d)}\right)$$

对所有 σ 的排列 $[d]$均成立，

具有连续边缘分布的随机向量 $\boldsymbol{X}=(X_1,\dots,X_d)$ 若关于 $\boldsymbol{a}\in\mathbb{R}^d$ 满足 $\boldsymbol{X}-\boldsymbol{a}\overset{\mathrm{d}}{=}\boldsymbol{a}-\boldsymbol{X}$，则称其为径向对称。对于 ***Copula***，我们只需考虑 $\boldsymbol{a}=\frac{1}{2}\mathbf{1}$ 的情形，此时径向对称性可逐分量化简为 $\boldsymbol{U}\overset{\mathrm{d}}{=}\mathbf{1}-\boldsymbol{U}$。在二元情形下，若 C 是 (U_1,U_2) 的 ***Copula***，则关于 $(1/2,1/2)$ 的径向对称性等价于对所有 $(u_1,u_2)\in[0,1]^2$，都有 $C(u_1,u_2)=u_1+u_2-1+C(1-u_1,1-u_2)$；例如参见（*Nelsen，2006*年，第 *2.7* 节）。

命题 3. 设 $\boldsymbol{U}$ 按定理 *3* 中的方法构造。

1. 假设各边缘分布上的校准对相同，即，

$$\pi_m=\pi;\quad (F_{m,[0]},F_{m,[1]})=(F_{[0]},F_{[1]})$$

对于 $m=1$、.....、d，且伯努利指标向量 $\boldsymbol{I}$ 是可交换的。则 $\boldsymbol{U}$ 的 ***Copula*** 也是可交换的，并可写成约化形式

$$C(\boldsymbol{u})=\Pi(\boldsymbol{u})+\sum_{k=2}^{d}\theta_k\sum_{\substack{S\subseteq[d]\\|S|=k}}\left(\prod_{m\notin S}u_m\right)\left(\prod_{m\in S}g(u_m)\right),\tag{19}$$

其中 $g(u) = \pi(F_{[1]}(u) - F_{[0]}(u)$，且对于任意 $k = 2$、.....、d,

$$\theta_k = \mathbb{E}\left[\prod_{m=1}^{k} \frac{I_m - \pi}{\pi}\right], \tag{20}$$

这是良　定义良好（不依赖于 k 指标的选择），这由 $\boldsymbol{I}$的可交换性可知。

2. 进一步 此外，假设 $\pi = \frac{1}{2}$，且该对 $(F_{[0]}, F_{[1]})$ 满足反射恒等式

$$F_{[1]}(u) = 1 - F_{[0]}(1-u), \qquad u \in [0,1], \tag{21}$$

并且伯努利向量是回文对称的，即 $\boldsymbol{I} \overset{\mathrm{d}}{=} \mathbf{1} - \boldsymbol{I}$。则 $\boldsymbol{U}$ 关于 $1\,2\mathbf{1}$径向对称。令 $Z_m := 2I_m - 1 \in \{-1, 1\}$，有 $\theta_k = \mathbb{E}[Z_1 \cdots Z_k]$，且对于所有奇数 k，均有 $\theta_k = 0$。

证明。(i) 部分由定理 3 得出，并且根据给定假设，所有混合矩 $\theta_S = \mathbb{E}[\prod_{m\in S}(I_m - \pi)/\pi]$ 对 S 的依赖仅通过 $|S|$。

对于 (ii)，(21) 等价于 $U_{m,[1]} \overset{\mathrm{d}}{=} 1 - U_{m,[0]}$ 对每个 m都成立。因此

$$1 - U_m = (1 - I_m)(1 - U_{m,[0]}) + I_m(1 - U_{m,[1]}) \overset{\mathrm{d}}{=} (1 - I_m)U_{m,[1]} + I_m U_{m,[0]}.$$

记 $I'_m := 1 - I_m$，右侧等于 $(1 - I'_m)U_{m,[0]} + I'_m U_{m,[1]}$，也就是说，它与使用索引 I'_m 而非 I_m构造的 U_m 同分布。因此，$\mathbf{1} - \boldsymbol{U}$ 与定理 3 中使用索引向量 $\boldsymbol{I}' = \mathbf{1} - \boldsymbol{I}$ 的构造同分布，由 $\boldsymbol{I}' \overset{\mathrm{d}}{=} \boldsymbol{I}$ 可得结论 $\boldsymbol{U} \overset{\mathrm{d}}{=} \mathbf{1} - \boldsymbol{U}$。最后，当 $\pi = \frac{1}{2}$ 时，我们有 $\theta_k = \mathbb{E}[Z_1 \cdots Z_k]$；由对称性可知 $(Z_1, \ldots, Z_d) \overset{\mathrm{d}}{=} -(Z_1, \ldots, Z_d)$，因此 $\theta_k = (-1)^k \theta_k$，故当 k 为奇数时，$\theta_k = 0$。

□

注 2. 在定理 *1* 中，边缘伯努利成功概率为 $\pi_1 = \Lambda_1/(\Lambda_1 - \lambda_1)$和 $\pi_2 = \Lambda_2/(\Lambda_2 - \lambda_2)$。因此，$\pi_1 = \frac{1}{2}$ 当且仅当 $\Lambda_1 = -\lambda_1$，等价地，本质上确界$_{[0,1]}\, \phi_1 = -$本质下确界$_{[0,1]}\, \phi_1$；对 ϕ_2 也同样如此。一个方便的充分结构条件是核关于 $1/2$ 对称，

$$g_1(1-u) = g_1(u) \quad \textit{for all } u \in [0,1] \tag{22}$$

这迫使 $\phi_1(1-u) = -\phi_1(u)$，从而得到 $\Lambda_1 = -\lambda_1$。当（22）成立（且 $\Lambda_1 = -\lambda_1$）时，定理 *1* 中的辅助累积分布函数满足反射恒等式 $F_{U,\,[1]}(u) = 1 - F_{U,\,[0]}(1-u)$，因此两个混合分量关于 $1/2$ 镜像对称。这包括经典 *FGM* 核 $g(u) = u(1-u)$，但不包括 *HKI/HKII* 核，除非指数退化为 *FGM* 情形。

注 3. 若 $\boldsymbol{I}$ 与 $\mathbb{P}(I_m = 1) = \pi$可交换，则其分布由和 $S_d := \sum_{m=1}^{d} I_m$的分布刻画。令 $w_j := \mathbb{P}(S_d = j)$，则对于 $\boldsymbol{i} \in \{0,1\}^d$，有 $\mathbb{P}(\boldsymbol{I} = \boldsymbol{i}) = w_{|\boldsymbol{i}|}/\binom{d}{|\boldsymbol{i}|}$，因此可容许性约束简化为 $w_j \geq 0$ 和 $\sum_{j=0}^{d} w_j = 1$。在 $\pi = \frac{1}{2}$的对称情形下，附加条件 $\boldsymbol{I} \overset{\mathrm{d}}{=} \mathbf{1} - \boldsymbol{I}$等价于 $w_j = w_{d-j},\ j = 0, \ldots, d$，并意味着式 (19)中的所有奇数阶系数 θ_k 均为零。这正是 *Blier-Wong* 等（*2024b*）研究的可交换 *FGM copula* 所依据的潜在结构。极值可交换对称伯努利分布（因而也是超模序下的极值 *Copula*）对应于 S_d 的两点“射线”分布（例如参见 *Fontana* 等人（*2021*，命题 *4.5*））。我们将在第 *3.4* 小节讨论 *END/EPD* 界及其对 $\rho_d^{\pm}$ 的影响。

3.4 多元正交象限斯皮尔曼 rho 与极值依赖

令 C 为定理 3 中式（14）所示形式的 d元 Copula。定义核积分 ls

$$\kappa_m := \int_0^1 g_m(u)\,\mathrm{d}u, \qquad \kappa_S := \prod_{m\in S}\kappa_m.$$

上、下正交象限斯皮尔曼系数的平均值由 Nelsen（1996）提出，并由 Gijbels 等人（2021年）进一步研究。其定义如下:

$$\rho_d^+(C) = \frac{d+1}{2^d-(d+1)}\left(2^d\int_{[0,1]^d}\Pi(\boldsymbol{u})\,\mathrm{d}C(\boldsymbol{u})-1\right), \tag{23}$$

$$\rho_d^-(C) = \frac{d+1}{2^d-(d+1)}\left(2^d\int_{[0,1]^d}C(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}-1\right). \tag{24}$$

在本节引入的 d 元萨尔尔马诺夫 copula 类中，这些参数可以按以下方式计算出闭式表达。

命题 4. 设 C 由(14)给出。那么

$$\rho_d^-(C) = \frac{d+1}{2^d-(d+1)}\sum_{\substack{S\subseteq[d]\\|S|\geq 2}} 2^{|S|}\theta_S\kappa_S, \tag{25}$$

$$\rho_d^+(C) = \frac{d+1}{2^d-(d+1)}\sum_{\substack{S\subseteq[d]\\|S|\geq 2}} (-1)^{|S|}2^{|S|}\theta_S\kappa_S. \tag{26}$$

如果所有 od $\boldsymbol{d}$ 阶参数为零（即对于奇数 $|S|$，$\theta_S=0$ 为零），则 ρ_d^+ （C） $=\rho_d^-$ （C）。

证明。我们从 ρ_d^- 开始。由 (14) 和 Fubini–Tonelli 定理可得

$$\int_{[0,1]^d}C(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} = \int_{[0,1]^d}\Pi(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} + \sum_{|S|\geq 2}\theta_S\left(\prod_{m\notin S}\int_0^1 u_m\,\mathrm{d}u_m\right)\left(\prod_{m\in S}\int_0^1 g_m(u_m)\,\mathrm{d}u_m\right).$$

因此

$$\int_{[0,1]^d}C(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} = 2^{-d} + \sum_{|S|\geq 2}\theta_S 2^{-(d-|S|)}\kappa_S,$$

由此可得 (25)，因为

$$2^d\int_{[0,1]^d}C(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} - 1 = \sum_{|S|\geq 2}2^{|S|}\theta_S\kappa_S.$$

对于 ρ_d^+，使用定理 3 中的随机表示。给定 I_m 后，各分量相互独立，由 (16) 可得

$$\mathbb{E}[U_m \mid I_m] = \int_0^1 1-F_{m,[I_m]}(u)\,\mathrm{d}u = \int_0^1 1-u-g_m(u)Z_m\,\mathrm{d}u = \frac{1}{2}-\kappa_m\frac{I_m-\pi_m}{\pi_m}.$$

因此，通过展开并取期望，我们得到

$$\mathbb{E}\left[\prod_{m=1}^d U_m\right] = \mathbb{E}\left[\prod_{m=1}^d\left(\frac{1}{2}-\kappa_m\frac{I_m-\pi_m}{\pi_m}\right)\right] = 2^{-d}+\sum_{|S|\geq 2}(-1)^{|S|}2^{-(d-|S|)}\theta_S\kappa_S;$$

最后一个等式成立是因为对每个 m 都有 $\mathbb{E}[Z_m]=0$，因此所有一阶项的期望均为零。最后，根据 $\int \Pi(\boldsymbol{u})\,\mathrm{d}C(\boldsymbol{u}) = \mathbb{E}[\prod_m U_m]$ 可得（26）。 □

如果固定核函数 $(g_1、\ldots\ldots、g_d)$，等价地说，固定 π_m-校准对 $(F_{m,[0]},\ F_{m,[1]})_{m=1}^d$，我们便可以研究改变伯努利指标向量 $\boldsymbol{I}$ 对所诱导 Copula C 及依赖参数 $\rho_d^{\pm}$ 的影响。此处相关的随机序是超模序 $\preceq_{\mathrm{sm}}$（例如，参见 Müller 和 Stoyan（2002年）；Shaked 和 Shanthikumar（2007年））。设 $\boldsymbol{I}$ 和 $\boldsymbol{I}'$ 为具有相同边缘分布 $(\pi_1、\ldots\ldots、\pi_d)$ 的伯努利向量，并按照定理 3，使用相同的核函数但不同的指标向量构造 $(U_1、\ldots\ldots、U_d)$ 和 $(U_1'、\ldots\ldots、U_d')$。下面的定理改编自 Blier-Wong 等人（2022年，定理 4.2），且不附证明。

定理 4。 假设对于每个 $m\in[d]$，g_m 在 $(0,\ 1)$ 上不变号。令 $\boldsymbol{U}$ 和 $\boldsymbol{U}'$ 按照定理 *3* 构造，使用相同的核，但采用边缘分布相同的不同伯努利指标向量 $\boldsymbol{I}$ 和 $\boldsymbol{I}'$。若 $\boldsymbol{I}\preceq_{\mathrm{sm}}\boldsymbol{I}'$，则 $\boldsymbol{U}\preceq_{\mathrm{sm}}\boldsymbol{U}'$。因此，$\rho_d^+(C_{\boldsymbol{I}})\le\rho_d^+(C_{\boldsymbol{I}'})$ 且 $\rho_d^-(C_{\boldsymbol{I}})\le\rho_d^-(C_{\boldsymbol{I}'})$。在所有具有固定边缘分布的伯努利向量中，共单调耦合 $\boldsymbol{I}^+$ 在 $\preceq_{\mathrm{sm}}$ 中取最大值；因此，对于固定的核，$C_{\boldsymbol{I}^+}$ 是上超模界，并使 $\rho_d^{\pm}$ 达到最大。

前一定理假设核 g_m 在 (0，1) 上不变号。大多数标准的 π 校准对都满足这一条件，包括经典 FGM 核 $g(u)=u(1-u)$、HKI/HKII 核以及 BKB 核。我们在附录 B 中给出一个会变号的核的例子。

假设采用可交换对称设定，并固定一个共同核。则存在可交换向量 $\boldsymbol{I}^-$ 和 $\boldsymbol{I}^+$，使得对于所有可交换的 $\boldsymbol{I}$，都有 $\boldsymbol{I}^-\preceq_{\mathrm{sm}}\boldsymbol{I}\preceq_{\mathrm{sm}}\boldsymbol{I}^+$；在这一类 Copula 中，极值 Copula 被称为极端负依赖（END）Copula 和极端正依赖（EPD）Copula。在具有共同核积分 $\kappa:=\int_0^1 g(u)\,\mathrm{d}u$ 的可交换 Farlie 设定下，命题 4 中的闭式结果可简化为

$$\rho_d^-(C)=\frac{d+1}{2^d-(d+1)}\sum_{k=2}^{d}\binom{d}{k}(2\kappa)^k\theta_k,\qquad \rho_d^+(C)=\frac{d+1}{2^d-(d+1)}\sum_{k=2}^{d}\binom{d}{k}(-1)^k(2\kappa)^k\theta_k.$$

对于 EPD 和 END，当 k 为奇数时，有 $\theta_k=0$，因此在这两种情况下，$\rho_d^+(C)=\rho_d^-(C)$。

对于固定的 $\boldsymbol{I}$，命题 4 表明，$\rho_d^{\pm}$ 对核函数的依赖性仅通过乘积 $\kappa_S=\prod_{m\in S}\kappa_m$ 体现。因此，要使 $\rho_d^{\pm}$ 的绝对值最大，就应使每个边缘分布中的 $|\kappa_m|$ 最大。在对称伯努利分割构造（$\pi_m=1/2$）中，精确上界为 $|\kappa_m|\le 1/4$；在定理2的证明中讨论的棋盘核处可达到等号。

对于 $d=3$，命题 4 给出

$$\rho_3^-(C)=4(\theta_{12}\kappa_1\kappa_2+\theta_{13}\kappa_1\kappa_3+\theta_{23}\kappa_2\kappa_3)+8\theta_{123}\kappa_1\kappa_2\kappa_3;$$

$$\rho_3^+(C)=4(\theta_{12}\kappa_1\kappa_2+\theta_{13}\kappa_1\kappa_3+\theta_{23}\kappa_2\kappa_3)-8\theta_{123}\kappa_1\kappa_2\kappa_3.$$

在可交换且对称情形（$\kappa_1=\kappa_2=\kappa_3=\kappa$）下，END 为（$\theta_2,\ \theta_3$）=（$-1/3$，0），而 EPD 为（$\theta_2,\ \theta_3$）=（1，0），因此

$$\rho_3^{\pm}(\mathrm{EPD})=12\kappa^2,\qquad \rho_3^{\pm}(\mathrm{END})=-4\kappa^2.$$

对于经典 FGM 核

$$g\ u\ =u\ 1-u\qquad \kappa=\int_0^1 u\ 1-u\quad u=1/6\qquad \rho_3^{\pm}\in(\,)\,(\,),\ 有\ (\,)\mathrm{d}\ ,$$

$[-1/9,1/3]$ $\kappa=1/4$ $\rho_3^{\pm}\in[-1/4,3/4]$

因此 。对于棋盘核，有 ，因此在可交换的 END–EPD 范围内。

4 Farlie Copula 的幂变换

受 Bekrizadeh 等人（2012 年）和 Bekrizadeh（2022 年）提出的 FGM 族幂次推广的启发，我们考虑通过对法利 copula 取幂所得到的 Copula。引入由下式隐式定义的归一化核 (h_1, h_2) 是很方便的

$$g_1(u_1) = u_1 h_1(u_1), \qquad g_2(u_2) = u_2 h_2(u_2), \qquad (u_1, u_2) \in [0,1]^2, \tag{27}$$

（必要时在 0 处按通常方式作连续延拓）。采用这一记号，(2) 等价于 Farlie (1960) 提出的 Farlie 型 Copula，其定义为

$$C(u_1, u_2) = u_1 u_2 \left(1 + a h_1(u_1) h_2(u_2)\right). \tag{28}$$

关于 Farlie 与 Sarmanov copula 之间的区别，请参阅附录 C。Bekrizadeh等人 (2012); Bekrizadeh (2022) 研究的幂变换族为

$$C_{a,r}(u_1, u_2) = u_1 u_2 \left(1 + a h_1(u_1) h_2(u_2)\right)^r, \qquad r \in \mathbb{N}. \tag{29}$$

直接验证 (29) 是二阶递增的，通常并不容易。本节的主要发现是，当 $r \in \mathbb{N}$ 时，幂 copula (29) 可通过一种简单的分块最大值构造得到：它是某个未作幂变换的（即 $r = 1$）Farlie 或 Sarmanov copula 的分量最大值的 Copula，其中使用了经过适当变换的核函数。核心思路是将定理 1 的伯努利混合表示与一种分块最大值构造相结合，从而得到所需的幂变换。

4.1 分块最大值与外幂变换

对于 Copula C 和 $r > 1$，定义外幂变换

$$C^{(r)}(u_1, u_2) := C\left(u_1^{1/r}, u_2^{1/r}\right)^r, \qquad (u_1, u_2) \in [0,1]^2, \tag{30}$$

只要右侧的 C 是 Copula，它就是 Copula；例如参见 Nelsen（2006）；Liebscher（2008）。以下命题广为人知（例如参见（Nelsen，2006，第 3.3.3 节）），但为完整起见，我们仍给出证明。

命题5 (分块最大值**)**。 设$(X_i, Y_i)_{i=1}^n$ 为具有连续边缘分布且 ***Copula*** 为 C的独立同分布随机变量。定义 $M_X :=$ 最大值$_{1 \le i \le n} X_i$和 $M_Y :=$ 最大值$_{1 \le i \le n} Y_i$。则(M_X, M_Y) 的 ***Copula*** 等于 (30)中的 $C^{(n)}$。

证明。 设 F_X, F_Y为 (X_1, Y_1) 的边缘分布，并令 $u_1 = F_X(x)$， $u_2 = F_Y(y)$。于是，

$$\mathbb{P}(M_X \le x, M_Y \le y) = \prod_{i=1}^n \mathbb{P}(X_i \le x, Y_i \le y) = C(u_1, u_2)^n.$$

边缘概率满足 $\mathbb{P}(M_X \le x) = u_1^n$ 且 $\mathbb{P}(M_Y \le y) = u_2^n$。因此，

$$\mathbb{P}(M_X \le x, M_Y \le y) = C(u_1, u_2)^n = C^{(n)}(u_1^n, u_2^n),$$

这就证明了该结论。 □

本节的主要发现是，外幂变换与分块最大值存在天然联系，这使我们能够结合伯努利混合表示与分块最大值表示，构造幂化的 Farlie copula。固定核函数 h_1，h_2 和 $a \in \mathbb{R}$，并考虑法利 copula

$$C_a(u_1, u_2) = u_1 u_2 (1 + a h_1(u_1) h_2(u_2)).$$

对于整数 $r \geq 1$，定义变换后的核函数 $h_1^{\langle r \rangle}(x) := h_1(x^r)$ 和 $h_2^{\langle r \rangle}(y) := h_2(y^r)$，以及基 Copula

$$C_a^{\langle r \rangle}(x, y) := xy(1 + a h_1(x^r) h_2(y^r)).$$

那么，其外幂变换为幂 copula

$$C_{a,r}(u_1, u_2) := C_a^{\langle r \rangle}\left(u_1^{1/r}, u_2^{1/r}\right)^r = u_1 u_2 (1 + a h_1(u_1) h_2(u_2))^r. \tag{31}$$

因此，$C_{a,\ r}$ 是对来自 $C_a^{\langle r \rangle}$ 的 r 个独立同分布样本中每个分量分别取 $k = 1$、.....、r 个观测值的最大值所构成的 Copula。

下文我们重点讨论整数幂 $r \in \mathbb{N}$，因为在这种情况下，可以通过分块最大值直接得到随机表示。不过，式 (30) 中的外幂变换对任意实数 $r > 1$ 都有定义。如果基 Copula 是最大无限可分的（例如，参见 Balkema 和 Resnick (1977)；Resnick (1987)；Genest 等 (2018)），那么这种幂构造也可推广到 $0 < r < 1$。

4.2 整数次幂的伯努利混合表示

现在我们证明，式 (31) 中的幂次 Farlie Copula 可以通过将 Farlie 表示与分块最大值相结合来构造。这将定理 3 扩展到幂次情形。

定理 5. 固定 $r \in \mathbb{N}$。设 $(X_k, Y_k)_{k=1}^r$ 为具有 *Copula* $C_a^{\langle r \rangle}(x, y) = xy(1 + a h_1(x^r) h_2(y^r))$ 的独立同分布随机变量。定义

$$U_1 := \left(\max_{1 \leq k \leq r} X_k\right)^r, \qquad U_2 := \left(\max_{1 \leq k \leq r} Y_k\right)^r.$$

则 (U_1, U_2) 的 *Copula* 为

$$C_{a,r}(u_1, u_2) = u_1 u_2 (1 + a h_1(u_1) h_2(u_2))^r, \qquad (u_1, u_2) \in [0, 1]^2.$$

证明。只需求出 (U_1, U_2) 的 Copula。设 (X, Y) 具有均匀边缘分布，且其 Copula 为

$$C_a^{\langle r \rangle}(x, y) = xy(1 + a h_1(x^r) h_2(y^r)).$$

取独立同分布的副本 $(X_k, Y_k)_{k=1}^r$，并令 $M_X = \max_{1 \leq k \leq r} X_k$、$M_Y = \max_{1 \leq k \leq r} Y_k$，再令 $U_1 = M_X^r$ 和 $U_2 = M_Y^r$。那么对于 $u \in [0, 1]$，

$$\mathbb{P}(U_1 \leq u) = \mathbb{P}(M_X^r \leq u) = \mathbb{P}(M_X \leq u^{1/r}) = \mathbb{P}(X \leq u^{1/r})^r = u,$$

因此 $U_1 \sim$ 均匀分布(0,1)（类似地，$U_2 \sim$ 均匀分布(0,1)）。对于 $(u_1, u_2) \in [0, 1]^2$，

$$\begin{aligned}
\mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) &= \mathbb{P}(M_X \leq u_1^{1/r}, M_Y \leq u_2^{1/r}) = \mathbb{P}(X \leq u_1^{1/r}, Y \leq u_2^{1/r})^r \\
&= (C_a^{\langle r \rangle}(u_1^{1/r}, u_2^{1/r}))^r = \left(u_1^{1/r} u_2^{1/r} (1 + a h_1(u_1) h_2(u_2))\right)^r \\
&= u_1 u_2 (1 + a h_1(u_1) h_2(u_2))^r = C_{a,r}(u_1, u_2),
\end{aligned}$$

如前所述。 □

注 *4.* 幂 *copula* $C_{a,r}(u_1, u_2) = u_1u_2(1 + ah_1(u_1)h_2(u_2))^r$ 的依赖参数可容许范围，与基 *copula* $C_{a,1}(u_1, u_2) = u_1u_2(1 + ah_1(u_1)h_2(u_2))$的可容许范围并不相同。定理 *5* 给出了一个充分条件，即若辅助 *copula* $C_a^{\langle r\rangle}(u_1, u_2) = u_1u_2(1 + ah_1(u_1^r)h_2(u_2^r))$ 是一个 *Copula*，则幂 *copula* $C_{a,r}$ 也是一个 *Copula*。因此，由变换后的核函数 $h_1(x^r)$ 和 $h_2(y^r)$ 得到的可容许范围为有效性提供了一个充分条件，而且两者的包含关系可能是严格的。

例如，如果 $h_1(x) = 1-x$，并且 $h_2(u) = 1-u$，那么 $g_1(x) = x(1-x^r)$ 以及 $\phi_1(x) = 1-(r+1)x^r$，因此下斜率界为 $-r$，从而得到一个随 a而变化的 r可容许范围。

5 结论

Sarmanov copula 是一类颇具吸引力的 Copula，因为它们在保持解析可处理性的同时，还能通过每个边缘分布中不同的核函数来容纳非对称性，并且通常可以为标准关联度量导出闭式表达式。然而，其实际应用受到可容许性的限制：即使在二元情形下，有效性也要求 copula 密度满足非负性约束；而在维度 d中， d-递增性质会变得具有组合性，对于大多数核函数而言很快就难以处理。本文的主要贡献是将这一可容许性问题重新表述为一个随机建模问题。通过将二元 Sarmanov copula 表示为简单一维核函数的伯努利索引混合，我们得到了一类通过构造即有效的 Copula，其可行性由潜在伯努利概率质量函数的存在性所刻画。依赖参数完全由一个潜在伯努利对承载，而每个边缘分布则由两个累积分布函数构成，其混合结果为均匀分布。该依赖参数与伯努利对的协方差成正比，而其可容许范围就是该协方差的弗雷歇–霍夫丁界。这为理解 Sarmanov copula 中依赖性的性质提供了诠释性认识，也为评估可容许性提供了一个简单准则。此外，这一框架还可用于构造和模拟非对称 Sarmanov copula，并在统一框架内重新得到若干经典的 Huang–Kotz 和 Bairamov–Kotz–Bekçi 规格及其可容许参数范围。

相同的思路可以自然地延拓到高维，我们证明， d 维中的精确随机表示会得到一个具有熟知 FGM 子集展开的 Copula，其中所有交互系数都是潜在伯努利随机向量的归一化中心混合矩。这一区分了边缘设计（通过单变量核因子实现）与依赖设计（通过伯努利分布实现），并绕过了对 d-递增性质的直接验证。在三维情形下，所得表述将成对交互与三方交互区分开来，并为基于正交象限的斯皮尔曼 rho 延拓导出闭式表达式。

最后，我们证明，形如 $u_1u_2(1 + ah_1(u_1)h_2(u_2))^r$ 的“幂化”萨尔马诺夫模型（其中r 是一个整数）可以通过未幂化的萨尔马诺夫 copula 的分块最大值来进行随机表示。

这项工作的若干自然后续方向值得探索。在方法上，潜在伯努利表示启发了利用给定伯努利指标时条件独立性的估计和检验程序（例如基于似然或 EM 类策略），并促使人们构建结构化伯努利模型（可交换、因子、稀疏或部分可交换），以控制高维情形下的参数增长。在理论上，将该表示推广到单次扰动之外，并刻画哪些基础构造允许具有最大无限可分性的分数阶分块最大值幂，将很有价值。从应用角度看，未来工作可以将这些 Copula 用于金融和保险领域的风险聚合、压力测试和情景生成。目前的应用主要集中于二元模型，但本文的 d元构造或可助力

更高维的应用。

致谢

作者感谢加拿大自然科学与工程研究委员会提供的经费支持（RGPIN-2025-06879）。

A 证明略

定理 *1* 的证明。 我们首先验证辅助函数 $F_{1,\,[0]}$，$F_{1,\,[1]}$，$F_{2,\,[0]}$，$F_{2,\,[1]}$ 是 $[0,1]$ 上有效的累积分布函数。随后，我们证明该随机表示具有均匀边缘分布，并给出所声称的 Sarmanov copula。注意，$g_1(0) = g_1(1) = 0$ 意味着 $F_{1,\,[i]}(0) = 0$ 且 $F_{1,\,[i]}(1) = 1$ 对于$i \in \{0,1\}$成立。由于 g_1 绝对连续，

$$F'_{1,[0]}(u) = 1 - \Lambda_1\phi_1(u) \geq 0 \quad \text{a.e.}, \qquad F'_{1,[1]}(u) = 1 - \lambda_1\phi_1(u) \geq 0 \quad \text{a.e.},$$

利用 $\phi_1 \leq 1/\Lambda_1$ 几乎处处成立以及 $\phi_1 \geq 1/\lambda_1$ 几乎处处成立，并且 $\lambda_1 < 0$ （因此 $\lambda_1\phi_1 \leq 1$ 几乎处处成立）。因此，$F_{1,\,[0]}$ 和 $F_{1,\,[1]}$均为非递减函数，且取值于 $[0,1]$。同样的论证也适用于 $F_{2,\,[0]}$ 和 $F_{2,\,[1]}$。接下来，我们验证该随机表示具有均匀边缘分布。对 I_1 进行条件化可得

$$\mathbb{P}(U_1 \leq u) = (1-\pi_1)F_{1,[0]}(u) + \pi_1 F_{1,[1]}(u) = u - \left(\Lambda_1(1-\pi_1) + \lambda_1\pi_1\right) g_1(u).$$

由于 $\pi_1 = \Lambda_1/(\Lambda_1 - \lambda_1)$，我们有 $(1-\pi_1)\Lambda_1 + \pi_1\lambda_1 = 0$，因此对所有 $u \in [0,1]$和 $U_1 \sim$均匀分布$(0,1)$，$\mathbb{P}(U_1 \leq u) = u$。对 $U_2 \sim$均匀分布$(0,1)$同样成立。最后，对于 $u_1, u_2 \in [0,1]$，在 (I_1, I_2) 条件下进行条件化，并利用给定 (I_1, I_2) 时 U 与 V 的条件独立性，Copula 为

$$C(u_1,u_2) = \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) = \mathbb{E}\left[F_{1,[I_1]}(u_1)F_{2,[I_2]}(u_2)\right],$$

其中

$$F_{1,[I_1]}(u) = u - g_1(u)\left((1-I_1)\Lambda_1 + I_1\lambda_1\right), \qquad F_{2,[I_2]}(u) = u - g_2(u)\left((1-I_2)\Lambda_2 + I_2\lambda_2\right).$$

展开乘积并重新组合各项可得

$$C(u_1,u_2) = u_1u_2 - u_1g_2(u_2)\mathbb{E}\left[(1-I_2)\Lambda_2 + I_2\lambda_2\right] - u_2g_1(u_1)\mathbb{E}\left[(1-I_1)\Lambda_1 + I_1\lambda_1\right] + g_1(u_1)g_2(u_2)\Xi,$$

其中 $\Xi := \mathbb{E}\left[\left((1-I_1)\Lambda_1 + I_1\lambda_1\right)\left((1-I_2)\Lambda_2 + I_2\lambda_2\right)\right]$。由于下式成立，两个线性项为零：

$$\mathbb{E}\left[(1-I_1)\Lambda_1 + I_1\lambda_1\right] = (1-\pi_1)\Lambda_1 + \pi_1\lambda_1 = 0$$

对于 I_2 也同样成立。此外，我们还有

$$(1-I_1)\Lambda_1 + I_1\lambda_1 = (\lambda_1 - \Lambda_1)(I_1 - \pi_1), \qquad (1-I_2)\Lambda_2 + I_2\lambda_2 = (\lambda_2 - \Lambda_2)(I_2 - \pi_2).$$

利用 $(\lambda_k - \Lambda_k)\,\pi_k = -\Lambda_k$（对 $k = 1,2$），可得 $\Xi = \Lambda_1\Lambda_2\theta$。因此，

$$C(u_1,u_2) = u_1u_2 + \Lambda_1\Lambda_2\theta g_1(u_1)g_2(u_2),$$

这就是所需的 Sarmanov 形式，其中 $a = \Lambda_1\Lambda_2\theta$。 □

定理2的证明。回忆一下，$\rho_S(C) = 12aG_1G_2$ 满足 $G_k := \int_0^1 g_k(u)\,\mathrm{d}u$。我们对 $|G_1|$ 进行界定；同样的论证也适用于 $|G_2|$。由 $g_1(0) = 0$ 且几乎处处有 $\phi_1 \le 1/\Lambda_1$，可得 $g_1(u) = \int_0^u \phi_1 \le u/\Lambda_1$。由 $g_1(1) = 0$ 且几乎处处有 $\phi_1 \ge 1/\lambda_1$，可得 $g_1(u) = -\int_u^1 \phi_1 \le -(1/\lambda_1)(1-u) = |1/\lambda_1|(1-u)$。这些包络在 u_0 处相交：$= \Lambda_1/(\Lambda_1 + |\lambda_1|)$，因此

$$G_1 \le \int_0^{u_0} \frac{u}{\Lambda_1}\,\mathrm{d}u + \int_{u_0}^1 \left|\frac{1}{\lambda_1}\right|(1-u)\,\mathrm{d}u = \frac{1}{2(\Lambda_1 + |\lambda_1|)}.$$

对 $-g_1$ 应用相同的界，得到 $|G_1| \le 1/(2(\Lambda_1 + |\lambda_1|))$。同样，

$$|G_2| \le \frac{1}{2(\Lambda_2 + |\lambda_2|)}.$$

现在我们求 a 的可容许范围。密度为 $c(u_1, u_2) = 1 + a\phi_1(u_1)\phi_2(u_2) \ge 0$，几乎处处成立。由于 $\phi_k \in [1/\lambda_k, 1/\Lambda_k]$ 几乎处处成立，$\phi_1\phi_2$ 的极值出现在各个角点。因此，若 $a \ge 0$，则 $1 + a/(\Lambda_1\lambda_2) \ge 0$ 且 $1 + a/(\lambda_1\Lambda_2) \ge 0$，所以 $a \le \min(\Lambda_1|\lambda_2|, |\lambda_1|\Lambda_2)$。若 $a \le 0$，记 $a = -|a|$，则 $1 - |a|/(\Lambda_1\Lambda_2) \ge 0$ 且 $1 - |a|/(\lambda_1\lambda_2) \ge 0$，所以 $|a| \le \min(\Lambda_1\Lambda_2, |\lambda_1\lambda_2|)$。

利用 $|\rho_S| = 12|a||G_1||G_2|$ 以及关于 $|G_1|, |G_2|$ 的界，我们得到

$$|\rho_S| \le \frac{3|a|}{(\Lambda_1 + |\lambda_1|)(\Lambda_2 + |\lambda_2|)}.$$

令 $x = |\lambda_1|$，$y = \Lambda_1$，$s = |\lambda_2|$ 和 $t = \Lambda_2$ 满足：x, y, s 和 t 均为正数。若 $a \ge 0$，则 $|a| \le \min(ys, xt)$，并且

$$|\rho_S| \le \frac{3\min(ys, xt)}{(x+y)(s+t)}.$$

由于 $(x+y)(s+t) = xs + yt + xt + ys \ge xt + ys + 2\sqrt{xsyt} = (\sqrt{xt} + \sqrt{ys})^2 \ge 4\min(xt, ys)$，我们得到 $|\rho_S| \le 3/4$。若 $a \le 0$，则 $|a| \le \min(yt, xs)$，因此

$$|\rho_S| \le \frac{3\min(xs, yt)}{(x+y)(s+t)}.$$

由于 $(x+y)(s+t) = xs + yt + xt + ys \ge xs + yt + 2\sqrt{xtys} = (\sqrt{xs} + \sqrt{yt})^2 \ge 4\min(xs, yt)$，我们再次得到 $|\rho_S| \le 3/4$。

最后，该界是锐的。设 $g_1^*(u) = g_2^*(u) = \min(u, 1-u)$，且 $\phi_k^*(u) = \mathbb{1}_{\{u\in[0,1/2)\}} - \mathbb{1}_{\{u\in(1/2,1]\}}$ 几乎处处成立。那么 $\Lambda_1 = \Lambda_2 = 1$，$\lambda_1 = \lambda_2 = -1$，并且 $G_1^* = G_2^* = \int_0^1 \min(u, 1-u)\,\mathrm{d}u = 1/4$。对于 $a = \pm 1$，Copula $C_\pm(u_1, u_2) = u_1u_2 \pm g_1^*(u_1)g_2^*(u_2)$ 的密度为 $1 \pm \phi_1^*(u_1)\phi_2^*(u_2)$，几乎处处成立，因此有效，且 $\rho_S(C_\pm) = 12(\pm 1)(1/4)^2 = \pm 3/4$。 □

B 示例核函数

我们在表1中汇集了文献中的若干已知 Sarmanov 核函数，以及它们对应的 $\kappa = \int_0^1 g(u)\,\mathrm{d}u$ 值、参数 Λ 和 λ，并列出相关参考文献。我们还介绍了若干满足定理2正则性条件的其他核函数。表中用 $B(\cdot,\cdot)$ 表示贝塔函数，用 Φ 表示标准正态累积分布函数。其中一些核函数沿用了 Fischer 和 Klein（2007年）的方法：他们选取适当单变量分布所对应的 $u \mapsto f(F^{-1}(u))$ 作为联结函数生成元，从而构造核函数。

#	$g(u)$	$\kappa=\int_0^1 g(u)\,\mathrm{d}u$	Λ	λ	参考
1	$u(1-u)$	$1/6$	1	-1	Morgenstern（1956年），Gumbel（1960年）
2	$u(1-u^p)$ $(p>0)$	$\dfrac{p}{2(p+2)}$	1	$-1/p$	Huang 和 Kotz（1999年）
3	$u(1-u)^q$ $(q>1)$	$\dfrac{1}{(q+1)(q+2)}$	1	$-\left(\dfrac{q+1}{q-1}\right)^{q-1}$	Huang 和 Kotz（1999年）
4	$u(1-u^p)^q \quad (p>0,\quad q>1)$	$\dfrac{1}{p}B\left(\dfrac{2}{p},\ q+1\right)$	1	$-\dfrac{(1+pq)^{q-1}}{p^q(q-1)^{q-1}}$	Bairamov 等人（2001年）
5	$\sin(\pi u)/\pi$	$\dfrac{2}{\pi^2}$	1	-1	Fischer 和 Klein（2007年）
6	$\sin^2(\pi u)/\pi$	$\dfrac{1}{2\pi}$	1	-1	Fischer 和 Klein（2007年）
7	$\begin{cases} u, & 0<u\le\frac{1}{2},\\ 1-u, & \frac{1}{2}<u\le 1, \end{cases}$	$1/4$	1	-1	Fischer 和 Klein（2007年）
8	$u^b(1-u)^a \quad (a>1,\quad b>1)$	$B(b+1,\ a+1)$	(32)	(33)	Lai 和 Xie（2000）
9	$\dfrac{u(u-1)}{2}$	$-\dfrac{1}{12}$	2	-2	Ting Lee (1996)
10	$\dfrac{u^{k+1}-u}{k+1}\quad k\in\mathbb{N}$	$-\dfrac{k}{2(k+1)(k+2)}$	$\dfrac{k+1}{k}$	$-(k+1)$	Ting Lee (1996)
11	$(1-e^{-u})-(1-e^{-1})u$	$\dfrac{3-e}{2e}$	e	$-\dfrac{e}{e-2}$	Ting Lee (1996)
12	(34)	$\dfrac{e^{2/3}}{\sqrt{3}}\left(\Phi\left(\dfrac{1}{\sqrt{3}}\right)-\dfrac{1}{2}\right)$	$\dfrac{1}{e-\frac{e^{2/3}}{\sqrt{3}}}$	$-\sqrt{3}e^{-2/3}$	Ting Lee (1996)
13	$u\dfrac{e^{2(1-u)}-1}{e^2-1}$	$\dfrac{e^2-5}{4(e^2-1)}$	1	$-\dfrac{e^2-1}{2}$	—
14	$u\dfrac{1-u}{1+u}$	$\dfrac{3}{2}-\log 4$	1	-2	—
15	$u(1-u)(1-2u)$	0	1	-2	—
16	$(1-u)$ 正弦$(\pi u)/\pi$	$\dfrac{1}{\pi^2}$	1	≈ -2.25 (35)	—
17	$u\dfrac{1-u}{1+u^2}$	$-1+\dfrac{\log 2}{2}+\dfrac{\pi}{4}$	1	-2	—
18	$u\dfrac{(1-u)^2}{1+u}$	$\dfrac{17}{6}-\log 16$	1	$\dfrac{1}{3\sqrt[3]{4}-5}$	—
19	$u(1-u)e^{-u}$	$-1+\dfrac{3}{e}$	1	$-e$	—
20	$\begin{cases} 3u, & 0<u\le\frac{1}{4},\\ 1-u, & \frac{1}{4}<u\le 1, \end{cases}$	$3/8$	$1/3$	-1	—

表1：所选核函数 $g(u)$及其对应的 κ、 Λ和 λ值。

表1中使用了以下表达式：

$$\Lambda = \frac{(a+b)^{a+b-2}\sqrt{a+b-1}}{\sqrt{ab}\left(b-\frac{\sqrt{ab}}{\sqrt{a+b-1}}\right)^{b-1}\left(a+\frac{\sqrt{ab}}{\sqrt{a+b-1}}\right)^{a-1}}; \tag{32}$$

$$\lambda = -\frac{(a+b)^{a+b-2}\sqrt{a+b-1}}{\sqrt{ab}\left(b+\frac{\sqrt{ab}}{\sqrt{a+b-1}}\right)^{b-1}\left(a-\frac{\sqrt{ab}}{\sqrt{a+b-1}}\right)^{a-1}}; \tag{33}$$

$$\frac{e^{2/3}}{\sqrt{3}}\left(\Phi\left(\sqrt{3}\Phi^{-1}(u)+\frac{2}{\sqrt{3}}\right)-u\right); \tag{34}$$

$$\lambda = -\pi\frac{\sqrt{y^2+4}}{y^2+2}, \tag{35}$$

其中 $y \in (0, \pi/2)$ 是 $y \sin y \tan y = 2$ 的解。数值上，(35) 给出 $\lambda \approx -2.2585010314$.

特别值得关注的一个核是表1中的核 #15，其定义为 $g(u) = u(1-u)(1-2u)$。这也是 Hamadou 和 Longla (2025) 的核 Φ_2，他们将此核视为第二个标准化平移勒让德多项式的原函数，从而为经典 FGM 项提供了正交扩展。

此核的有向面积为零，即 $\kappa = 0$，因此对所有可容许的 a，它所诱导的 Sarmanov copula 的 Spearman's ρ_S 和 Kendall's τ 均等于零。然而，除非 $a = 0$，该 Copula 并非独立 copula。由此得到的 Copula 为

$$C_a(u_1, u_2) = u_1u_2 + au_1(1-u_1)(1-2u_1)u_2(1-u_2)(1-2u_2), \qquad (u_1, u_2) \in [0,1]^2$$

当 $a \in [-1, 2]$. 时。另请注意，$g(u)$ 在 $(0, 1)$ 上变号，在 $(0, 1/2)$ 上为正，在 $(1/2, 1)$ 上为负。这意味着定理 4 不适用于此核，因为相关分布$F_{1,[0]}$ 和 $F_{1,[1]}$ 在随机序下不可比较。

C Farlie 与 Sarmanov copula 的比较

我们比较三种经典的二元依赖结构，它们常以不同名称呈现。第一种是 Farlie (1960) 提出的 Farlie 型 copula 设定：

$$C_\alpha(u_1, u_2) = u_1u_2(1 + \alpha h_1(u_1)h_2(u_2)), \qquad (u_1, u_2) \in [0,1]^2,$$

由标量 α 和固定核函数 h_1，h_2 参数化。第二种是 Rodríguez-Lallena 和 Úbeda-Flores (2004) 提出的加性 copula 设定，定义如下：

$$C(u_1, u_2) = u_1u_2 + g_1(u_1)g_2(u_2), \qquad (u_1, u_2) \in [0,1]^2,$$

由函数 g_1，g_2 参数化，并满足 copula 可容许条件。第三种是 Sarmanov (1966) 和 Ting Lee (1996) 提出的 Sarmanov 构造，其密度定义为：

$$f(x_1, x_2) = f_1(x_1)f_2(x_2)(1 + \alpha_{12}\phi_1(x_1)\phi_2(x_2)),$$

其核函数 ϕ_1，ϕ_2 有界，并满足均值为零和正性约束。当边缘分布连续时，通过概率积分变换可导出一个 Copula。

这三种构造密切相关。在绝对连续的条件下，它们都可以得到如下形式的 copula 密度：

$$c(u_1, u_2) = 1 + \gamma\psi(u_1)\eta(u_2),$$

其中 γ 是一个标量，ψ、η 是适当的单变量因子。因此，这些族之间的差异不在于密度层面的因子分解本身，而在于：(i) 施加因子分解的层面（Copula 层面还是联合密度层面）；(ii) 对因子函数所要求的边界条件和正则性限制。

对于 Farlie Copula，假设 h_1 和 h_2 绝对连续，且 $g_1(u) = uh_1(u)$ 和 $g_2(u) = uh_2(u)$ 的导数本质有界。那么该 Copula 绝对连续，其密度为

$$c_\alpha(u_1, u_2) = 1 + \alpha\phi_1(u_1)\phi_2(u_2), \qquad \phi_1(u) = (uh_1(u))', \quad \phi_2(u) = (uh_2(u))'.$$

因此，密度层面的因子必定是 $u_1h_1(u_1)$ 和 $u_2h_2(u_2)$ 的导数。

对于加性设定，如果 g_1 和 g_2 绝对连续，那么 C 也绝对连续，且

$$c(u_1, u_2) = 1 + g_1'(u_1)g_2'(u_2) \qquad \text{for a.e. } (u_1, u_2) \in (0, 1)^2.$$

通过以下对应关系，每个 Farlie 型 Copula 都具有加性形式

$$g_1(u) = \alpha uh_1(u), \qquad g_2(u) = uh_2(u).$$

因此，乘法型 Farlie 参数化是加性参数化的特例。

存在一个部分逆命题，但它取决于原点处的边界正则性。固定 $\alpha \neq 0$ 并考虑 $C(u_1, u_2) = u_1u_2 + g_1(u_1)g_2(u_2)$。对于 $u_1, u_2 > 0$，可将 $u_1u_2 + g_1(u_1)g_2(u_2)$ 改写为$u_1u_2(1 + \alpha h_1(u_1)h_2(u_2))$，其中 $h_1(u_1) = g_1(u_1)/(\alpha u_1)$，且 $h_2(u_2) = g_2(u_2)/u_2$。为得到与 Farlie 型框架相容的核函数，必须排除 $u = 0$ 和$v = 0$ 处可能出现的不定式。若极限存在，可以定义

$$h_1(0) := \lim_{u\downarrow 0} \frac{g_1(u)}{u\alpha}, \qquad h_2(0) := \lim_{u\downarrow 0} \frac{g_2(u)}{u},$$

如果所得的延拓有界，并满足对 Farlie 核函数施加的正则性条件。当这些极限不存在，或除以 u 会破坏绝对连续性时，加性 Copula 仍可能是可容许的，但不属于本文定义的 Farlie 型类；不过，定理 1 的随机表示仍然适用。事实上，我们对 $uh_1(u)$ 和$uh_2(u)$施加条件，而不是直接对 h_1 和 h_2 施加条件，这使原点处可以有更一般的边界行为，并且与 Rodríguez-Lallena 和 Úbeda-Flores (2004) 的加性类一致。

对于 Sarmanov 构造，假设边缘分布连续，并定义 $U_1 = F_1(X_1)$ 和$U_2 = F_2(X_2)$。由此得到的 Copula 密度为

$$c(u_1, u_2) = 1 + \alpha\tilde{\phi}(u_1)\tilde{\psi}(u_2), \qquad \tilde{\phi}(u) = \phi_1(F_1^{-1}(u)), \quad \tilde{\psi}(v) = \phi_2(F_2^{-1}(v)),$$

根据零均值约束，$\int_0^1 \tilde{\phi}(u)\,\mathrm{d}u = \int_0^1 \tilde{\psi}(v)\mathrm{d}v = 0$。积分后得到加性 Copula 表示

$$C(u_1, u_2) = u_1u_2 + \alpha\Phi(u_1)\Psi(u_2), \qquad \Phi(u) = \int_0^u \tilde{\phi}(s)\,\mathrm{d}s, \quad \Psi(u) = \int_0^u \tilde{\psi}(t)\,\mathrm{d}t,$$

因此，在绝对连续性条件下，萨尔马诺夫 copula 与 Rodríguez-Lallena 和 Úbeda-Flores（2004年）的加法类一致，只是采用了等价的重新参数化。

因此，在绝对连续的条件下，这三个族与萨尔马诺夫 copula 具有完全相同的因子分解；只要满足边界条件和正则性条件，定理 1 便适用于这三种构造。它们的主要区别在于因子函数的解释、施加因子分解的层级，以及可容许性所要求的边界条件。

参考文献

Bairamov, I., Kotz, S., and Bekçi, M. (2001年). 新的广义 Farlie-Gumbel-Morgenstern 分布及次序统计量的伴随变量。《应用统计学杂志》, 28(5):521–536. Balkema, A. A. and Resnick, S. I. (1977). 最大无限可分性。《应用概率学杂志》, 14(2):309–319. Bekrizadeh, H. (2022年). 广义 FGM copula：性质与应用。《统计学通讯：模拟与计算》, pages 1–12. Bekrizadeh, H., Parham, G. A., and Zadkarmi, M. R. (2012). Farlie–Gumbel–Morgenstern copula 的新推广。《应用数学科学》, 6(71). Blier-Wong, C., Cossette, H., Legros, S., and Marceau, E. (2024a). 一种利用伯努利分布和 Coxian-2 分布构造高维 copula 的新方法。《多元分析杂志》, 201:105261. Blier-Wong, C., Cossette, H., and Marceau, E. (2022年). 利用多元伯努利随机变量对 FGM copula 进行随机表示。《计算统计学与数据分析》, 173. Blier-Wong, C., Cossette, H., and Marceau, E. (2024b). 可交换 FGM copula。《应用概率进展》, 56(1):205–234. Cherubini, U., Luciano, E., and Vecchiato, W. (2004年). 《金融中的 *Copula* 方法》。Wiley 金融丛书。John Wiley & Sons, 霍博肯, NJ. Cossette, H., Marceau, E., Mutti, A., and Semeraro, P. (2025年). 广义 FGM 依赖：几何表示与和的凸界。《统计学论文》, 66(7):149. Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2006年). 《相依风险精算理论》。John Wiley & Sons, 奇切斯特, 西萨塞克斯。Durante, F. and Sempi, C. (2015年). 《*Copula* 理论原理》。Chapman and Hall/CRC, 佛罗里达州博卡拉顿。Eyraud, H. (1936). 相关性测度原理。《里昂大学年刊。第 *3* 辑，科学。*A* 栏，数学与天文学》, pages 30–47. Farlie, D. J. (1960年). 若干相关系数在一般二元分布中的表现。《生物统计学》, 47(3 and 4):307–323. Fischer, M. and Klein, I. (2007年). 利用某些单变量分布构造广义 FGM copula。《*Metrika*》, 65(2):243–260. Fontana, R., Luciano, E., and Semeraro, P. (2021年). 信用风险中的模型风险。《数理金融》, 31(1):176–202. Fontana, R. and Semeraro, P. (2018年). 具有一组指定矩的多元伯努利分布表示。《多元分析杂志》, 168:290–303. Fredricks, G. A. and Nelsen, R. B. (2007年). 连续随机变量对的斯皮尔曼 rho 与肯德尔秩相关系数之间的关系。《统计规划与推断杂志》, 137(7):2143–2150.

Genest, C. 和 Favre, A.-C.（2007）。关于 Copula 建模，你一直想知道却不敢问的一切。《水文工程学杂志》，12(4):347–368。Genest, C.、Nešlehová, J. G. 和 Rivest, L.-P.（2018）。具有 ℓ_1-范数对称指数测度的多元最大无限可分 Copula 类。《伯努利》，24(4B)。Genest, C.、Okhrin, O. 和 Bodnar, T.（2024）。从 Abe Sklar 到当今的 Copula 建模。《多元分析杂志》，201:105278。Gijbels, I.、Kika, V. 和 Omelka, M.（2021）。关于多元关联度量的设定及其随维数增加的表现。《多元分析杂志》，182:104704。Größer, J. 和 Okhrin, O.（2021）。Copula：概述与最新进展。*WIREs* 计算统计学。Gumbel, E. J.（1960年）。二元指数分布。《美国统计学会杂志》，55(292):698–707。Hamadou, M.-A. 和 Longla, M.（2025年）。d 元 FGM Copula 的一种扩展及其应用。*arXiv:2509.07281*。Hashorva, E. 和 Ratovomirija, G.（2015年）。保险应用中的 Sarmanov 混合 Erlang 风险。《*ASTIN* 公报》：国际精算师协会期刊，45(1):175–205。Huang, J. S. 和 Kotz, S.（1999年）。Farlie-Gumbel-Morgenstern 分布的修正：一座难以攀登的高山。《*Metrika*》，49(2):135–145。Joe, H.（2015年）。《*Copula* 依赖建模》。《统计与应用概率专著》第134卷。CRC 出版社，佛罗里达州博卡拉顿。Lai, C. D. 和 Xie, M.（2000）。正象限相依二元分布的新族。《统计与概率快报》，46(4):359–364。Liebscher, E.（2008）。非对称多元 Copula 的构造。《多元分析杂志》，99(10):2234–2250。McNeil, A. J.、Frey, R. 和 Embrechts, P.（2015年）。《定量风险管理：概念、技术与工具》。普林斯顿金融丛书。普林斯顿大学出版社，新泽西州普林斯顿，修订版。Morgenstern, D.（1956年）。二维分布的简单示例。《数学与统计通讯》，8:234–235。Müller, A. 和 Stoyan, D.（2002年）。《随机模型与风险的比较方法》。威利出版社，奇切斯特，第1版。Naseri, K. 和 Hummel, M. A.（2022年）。一种基于贝叶斯 Copula 的非平稳框架，用于评估美国海岸线沿岸的复合洪水风险。《水文学杂志》，610:128005。Nelsen, R. B.（1996）。多元关联的非参数度量。载于数学统计学会讲义—专著系列，第223–232页。数学统计学会，加利福尼亚州海沃德。

Nelsen, R. B.（2006年）。《*Copula* 入门》。Springer 统计学丛书。Springer，纽约，第2版。Ratovomirija, G.、Tamraz, M. 和 Vernic, R.（2017年）。关于若干多元 Sarmanov 混合 Erlang 再保险风险：聚合与资本分配。《保险：数学与经济学》，74:197–209。Resnick, S. I.（1987）。《极值、正则变差与点过程》。Springer 运筹学与金融工程丛书。Springer，纽约，纽约州。Ribeiro, A. F. S.、Russo, A.、Gouveia, C. M.、Páscoa, P. 和 Zscheischler, J.（2020年）。利用嵌套 Copula 估算复合干热极端事件导致作物歉收的风险。《生物地球科学》，17:4815–4830。Rodríguez-Lallena, J. 和 Úbeda-Flores（2004年）。一种新型二元 Copula。《统计与概率快报》，66(3):315–325。Sarmanov, O. V.（1966年）。广义正态相关与二维 Fréchet 类。载于《科学院通报》，第168卷，第32–35页。俄罗斯科学院。Shaked, M. 和 Shanthikumar, J. G.（2007年）。《随机序》。Springer 统计学丛书。Springer，纽约。Sharakhmetov, S. 和 Ibragimov, R.（2002年）。二值随机变量联合分布的刻画及其应用。《多元分析杂志》，83(2):389–408。Shubina, M. 和 Lee, M.-L. T.（2004年）。关于 Sarmanov 族二元分布的最大可达相关性及其他依赖性度量。《统计通讯：理论与方法》，33(5):1031–1052。Sklar, A.（1959年）。《n 维分布函数及其边缘分布》。《巴黎大学统计研究所刊物》，8:229–231。Teugels, J. L.（1990年）。多元伯努利分布和二项分布的一些表示。《多元分析杂志》，32(2):256–268。Ting Lee, M.-L.（1996）。Sarmanov 族二元分布的性质与应用。《统计通讯：理论与方法》，25(6):1207–1222。Vernic, R.、Bolancé, C. 和 Alemany, R.（2022年）。用于刻画保险索赔频率与平均赔款额之间依赖性的 Sarmanov 分布。《保险：数学与经济学》，102:111–125。