arXiv:2301.02991v2 [stat.ME] 1 Aug 2023

基于角度概率积分变换的二元独立性检验：重点关注圆周-圆周和圆周-线性数据

Fernández-Durán, J.J. 和 Gregorio-Domínguez, M.M. ITAM 电子邮箱：jfdez@itam.mx

摘要

连续随机变量 X 的概率积分变换（PIT） F_X 是一个均匀分布随机变量 $U = F_X$（X）。我们将角度概率积分变换（APIT）定义为 $\theta_U = 2\pi U = 2\pi F_X$（$X$），它对应于单位圆上的均匀分布角度。对于圆周（角度）随机变量，绝对连续且相互独立的圆形均匀随机变量之和仍为圆形均匀随机变量；也就是说，圆形均匀分布在求和运算下封闭，并且是单位圆上的稳定连续分布。若考虑两个随机变量 X_1 和 X_2的角度概率积分变换之和（差），并检验其和（差）的圆周均匀性，这等价于检验原始变量的独立性。本研究采用一类灵活的非负三角和（NNTS）圆分布，其中包含作为该分布族成员的圆形均匀分布，以评估所提出独立性检验的检验功效；我们通过从 NNTS 备择分布中生成样本来完成这一评估，这些备择分布与圆形均匀原假设分布的距离可能更近。

关键词：圆形-圆形依赖；圆形-线性依赖；圆周均匀性检验；Copula；依赖度量

1 引言

独立性检验是统计学中最重要的任务之一，例如，在构建一组随机变量的联合分布时，或在回归模型中考察一个变量相对于其他变量的条件依赖关系时。

根据 Herwatz 和 Maxand（2020年），可以考虑以下独立性检验：二元（成对）独立性检验、相互独立性检验和分组独立性检验。下文中，我们仅考虑绝对连续随机变量。密度函数的支撑集位于实数轴上某一区间的随机变量称为线性随机变量，而支撑集位于单位圆上的随机变量称为圆周随机变量。圆周随机变量的分布函数是周期函数，因此其起始方向可以任意选取。在本文中，我们假设所有圆形随机变量均定义在区间 $(0, 2\pi]$ 上，具有相同的起始方向，并且都是以弧度表示、沿逆时针或顺时针方向测量的角度。从这个意义上说，若 Θ 是一个起始方向为零、以弧度表示且沿逆时针方向测量的圆周随机变量，

$$F_\Theta(\theta) = P(0 < \Theta \leq \theta) \text{ for } 0 \leq \theta \leq 2\pi$$

and

$$F_\Theta(\theta + 2k\pi) = F_\Theta(\theta)$$

对于 k 整数。因此，绝对连续圆周随机变量的密度函数是一个周期函数，使得 $f_\Theta(\theta) \geq 0$, $f_\Theta(\theta + 2k\pi) = f_\Theta(\theta)$ 且 $\int_0^{2\pi} f_\Theta(\theta)d\theta = 1$。对于两个随机变量 X_1 和 X_2，二元（成对）独立性检验的原假设为 H_0：$F_{X_1,X_2}(x_1, x_2) = F_{X_1}(x_1)F_{X_2}(x_2)$，其中 $F_{X_1,X_2}(x_1, x_2) = P\{X_1 \leq x_1, X_2 \leq x_2\}$是二元联合分布函数，而 $F_{X_1}(x_1) = P\{X_1 \leq x_1\}$和 $F_{X_2}(x_2) = P\{X_2 \leq x_2\}$是相应的边缘分布函数。对于一组随机变量 $X_1, X_2, \ldots, X_d$，相互独立性检验考察原假设 H_0：$F_{X_1,X_2,\ldots,X_d}(x_1, x_2, \ldots, x_d) = \prod_{k=1}^{d} F_{X_k}(x_k)$，其中 $F_{X_1,X_2,\ldots,X_d}(x_1, x_2, \ldots, x_d)$ 是联合分布，且 $F_{X_k}(x_k)$ 对于 $k =$

1, 2, . . . , d 是边缘单变量分布函数。最后，分组独立性检验考察不相交随机变量子集之间的独立性。如果 $F_{X_1,X_2,\ldots,X_d}$ 和 F_{X_k} 对于 $k = 1, 2, \ldots, d$ 的参数函数形式已知，则可基于大小为 $n (i = 1, 2, \ldots, n)$ 的随机向量样本 $\underline{X}_i = (X_{i1}, X_{i2}, \ldots, X_{id})^\top$ 构造独立性的似然比检验。最常用的独立性检验是列联表的卡方独立性检验，但在处理绝对连续随机变量时并不适用。作为似然比独立性检验的替代方法，人们还通过考虑非参数（分布无关）方法、秩检验（见 Hoeffding，1948 以及 Kendall 和 Stuart，1951），以及由经验Copula过程导出的依赖性度量（关联）开发了其他独立性检验（Deheuvels，1981；Genest 和 Rémillard，2004；Genest 和 Verret，2005；Genest 等人，2019；Roy，2020 以及 Roy 等人，2020）。对于由 d 个绝对连续线性随机变量组成的向量 C_n， $\underline{X}^\top = (X_1, X_2, \ldots, X_d)^\top$，其样本量为 n，经验 Copula 定义如下：

$$C_n(u_1, u_2, \ldots, u_d) = \frac{1}{n} \sum_{j=1}^{n} I(\hat{F}_1(X_{i1}) \leq u_1, \hat{F}_2(X_{i2}) \leq u_2, \ldots, \hat{F}_d(X_{id}) \leq u_2) \tag{1}$$

其中， $I()$ 是一个指示函数：当其参数中的条件满足时取值为 1，否则取值为 0； F_1、 F_2、.....、 F_d 是随机变量 X_1、 X_2、.....、 X_d 的经验分布函数。R 软件包 $copula$ 实现了一种独立性检验，该检验基于绝对连续线性随机变量的经验Copula与独立Copula之间的距离（Hofert 等人，2022；Kojadinovic 和 Yan，2010）。

在非参数独立性检验中，有一类检验基于经验独立过程的某种泛函。经验独立过程定义为经验联合分布函数与各经验单变量分布函数乘积之间的距离。从历史上看，最常用的泛函是 Cramér-von Mises 泛函和 Kolmogorov-Smirnov 泛函（参见 Blum、Keifer 和 Rosenblatt，1961；DeWet，1980年；Deheuvels，1981）。例如，Hoeffding（1948）采用 Cramér-von Mises 泛函构造了

两个随机变量之间的独立性秩检验。Kallenberg 和 Ledwina（1999）提出了现代独立性秩检验。基于核的方法也被用于估计经验独立过程，例如 Pfister 等人（2018）的研究。Mardia 和 Kent（1991）采用一般的 Rao 评分检验构造独立性检验。Csörg˝o（1985）提出了基于多元经验特征函数的独立性检验，而 Einmahl 和 McKeague（2003）提出了基于经验似然的检验。Joe（1990）定义了由熵导出的依赖性度量，Berrett 和 Samworth（2019）则基于互信息定义了依赖性度量。

构造独立性检验时，可以利用多元联合分布的特性。例如，对于多元正态分布，可以通过检验相关矩阵是否为单位矩阵来检验独立性。一些两两检验包括 Pearson（1920）乘积矩相关系数检验、Kendall（1938）秩相关系数检验，以及 Spearman（1904年）秩相关系数检验。当然，对于一对高斯随机变量，拒绝零相关意味着拒绝两两独立；但是，对于包含两个以上高斯随机变量的集合，应用两两独立性（相关性）检验并不足以检验相互独立。Wilks 检验（Wilks，1935年）是针对多元高斯总体的最优独立性检验；对于向量 $\underline{X}_{D_1}$和 $\underline{X}_{D_2}$的双变量分组独立性检验，其中 $\underline{X} = \underline{X}_{D_1 \bigcup D_2} = (\underline{X}_{D_1}, \underline{X}_{D_2})^\top$，该检验针对样本量为 n 的样本，考虑以下检验统计量，

$$W = \frac{|\hat{\Sigma}_{D_1 \bigcup D_2}|}{|\hat{\Sigma}_{D_1}||\hat{\Sigma}_{D_2}|} \tag{2}$$

其中，$\hat{\Sigma}_{D_1 \bigcup D_2} = \sum_{j=1}^{n} (\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})^\top$是完整观测向量 $\underline{x} = \underline{x}_{D_1 \bigcup D_2}$ 的估计协方差矩阵，该向量划分为 $\hat{\Sigma}_{D_1}$ 和 $\hat{\Sigma}_{D_2}$，其中 $\hat{\Sigma}_{D_1} = \sum_{j=1}^{n} (\underline{x}_{D_1,j} - \bar{\underline{x}}_{D_1})(\underline{x}_{D_1,j} - \bar{\underline{x}}_{D_1})^\top$，且 $\hat{\Sigma}_{D_2} = \sum_{j=1}^{n} (\underline{x}_{D_2,j} - \bar{\underline{x}}_{D_2})(\underline{x}_{D_2,j} - \bar{\underline{x}}_{D_2})^\top$。因此，统计量 W 衡量的是 $\Sigma_{D_1 \bigcup D_2}$ 的行列式与 $\hat{\Sigma}_{D_2}$ 和 Σ_{D_2} 的行列式乘积之间的差距。在多元高斯总体中，等式关系成立的条件是

在 $\underline{X}_{D_1}$ 与 $\underline{X}_{D_2}$ 相互独立的独立性原假设下。渐近地，在满足正则条件时， $-n\ln(W)$ 服从自由度为 $\sharp D_1 \sharp D_2$ 的卡方分布，其中 $\sharp D_1$ 和 $\sharp D_2$ 分别是集合 D_1 和 D_2 的基数。

对于圆周-圆周（角度-角度）和圆周-线性（角度-线性）情形，研究目标分别是检验两个圆形随机变量之间，以及一个圆形随机变量与一个线性随机变量之间的二元独立性。人们通过规定依赖性度量并研究其（渐近）分布，提出了独立性检验。将基于一致性概念的一般依赖性度量——肯德尔秩相关系数和斯皮尔曼秩相关系数——应用于一对圆形随机变量，或应用于一个圆形随机变量和一个线性随机变量；或者基于秩构造分布无关相关系数，由此，Fisher and Lee（1981、1982和1983）提出了独立性检验，随后Fisher（1993）以及Mardia和Jupp（2000）对其进行了修订。

本研究旨在通过考虑每个变量的角度概率变换来构建两个随机变量的二元（成对）独立性检验。该变换对应于（$0, 2\pi$]上的圆形均匀分布，并结合圆形统计学理论中的一项额外结果（参见Fisher，1993；Upton and Fingleton，1989年；Mardia和Jupp，2000年；以及Jammalamadaka and SenGupta，2001）。本检验采用灵活的非负三角级数（NNTS）分布进行了评估（Fernández-Durán，2004b、2007）。尽管所提独立性检验具有一般性，但尤其适用于检验圆周-圆周和圆周-线性情形下的独立性。因此，本文构建了一种依赖性度量。

本文共分为六个部分，包括引言。第二部分介绍 Johnson 和 Wehrly（1977）模型，作为对两个随机变量进行二元独立性检验的理论依据，并在此介绍 NNTS 圆分布理论。第三部分介绍所提二元独立性检验、依赖性度量，以及将其应用于模拟数据以研究检验效能的方法。本部分还阐述如何将二元独立性检验扩展为针对

两个不相交的随机变量子集的组间独立性检验。第四部分通过模拟研究评估所提检验在线性-线性、圆周-线性和圆周-圆周情形下的检验效能。第五部分介绍将所提独立性检验应用于真实数据集。最后，第六部分给出结论。

2 二元 Johnson 和 Wehrly 模型与 NNTS 族圆形密度

Sklar（1959）定理指出，两个连续随机变量 X_1 和 X_2的联合累积分布函数，其对应的累积分布函数分别为 $F_{X_1}(x_1)$ 和$F_{X_2}(x_2)$，可以表示为

$$F_{X_1,X_2}(x_1,x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2)) \tag{3}$$

其中，$C(u,v)$ 是 Copula（分布）函数，对应于区间 $(0,1]$ 上两个同分布均匀随机变量的二元联合分布，$U = F_{X_1}(X_1)$ 和$V = F_{X_2}(X_2$。线性-线性 Copula 函数必须是递增函数，并满足 $C(u,1) = u, C(1,v) = v, C(0,v) = C(u,0) = 0$（参见 Nelsen，1999）。对公式 3 求导，可得到 X_1 和 X_2 的联合密度函数 $f_{X_1,X_2}(x_1,x_2)$，即

$$f_{X_1,X_2}(x_1,x_2) = c(F_{X_1}(x_1), F_{X_2}(x_2)) f_{X_1}(x_1) f_{X_2}(x_2) \tag{4}$$

其中，$f_{X_1}(x_1)$ 和 $f_{X_2}(x_2)$ 分别是 X_1 和 X_2 的边际密度函数。函数 $c(u,v)$ 是 Copula 密度函数，定义为 $c(u,v) = {}^{\partial^2 C}{}_{(u,v)\,\partial u \partial v}$。Johnson 和 Wehrly（1977）以及 Wehrly 和 Johnson（1980）以如下方式提出了一大类适用于圆周-圆周（Θ_1 和 Θ_2）及圆周-线性（Θ 和 T）随机向量的联合密度函数:

$$f_{\Theta_1,\Theta_2}(\theta_1,\theta_2) = 2\pi g(2\pi(F_{\Theta_1}(\theta_1) \pm F_{\Theta_2}(\theta_2)) f_{\Theta_1}(\theta_1) f_{\Theta_2}(\theta_2) \tag{5}$$

and

$$f_{\Theta,T}(\theta,t) = 2\pi g(2\pi(F_{\Theta}(\theta) \pm F_T(t))f_{\Theta}(\theta)f_T(t). \tag{6}$$

Fernández-Durán（2004）通过 Sklar（1959）定理（见 Nelsen，1999），从 Copula函数理论的角度阐明了 Johnson-Wehrly模型的结构，该结构满足

$$c(u,v) = 2\pi g(2\pi(u \pm v)). \tag{7}$$

函数 g 必须是区间 $(0, 2\pi]$ 上角度（圆周）随机变量的密度函数。对于圆周-圆周和圆周-线性二元Copula，函数 c 在其圆周变量上也必须具有周期性，以满足圆周随机变量密度函数的周期性；这正是 Johnson 和 Wehrly 模型中连接函数 g 是圆周随机变量密度的原因。

Johnson 和 Wehrly 通过考虑条件变量，推导出了圆周-圆周和圆周-线性二元模型。当函数 g 对应于圆上的均匀圆周密度时，$g(\theta) = \frac{1}{2\pi}$，Johnson 和 Wehrly 模型的联合密度对应于独立情形，此时圆周-圆周（圆周-线性）模型的联合密度是边缘单变量密度的乘积。

Johnson 和 Wehrly模型的这一性质促使我们通过使用 NNTS 密度族中的密度函数近似圆周密度函数 g，从而构建独立性检验；该密度族包括均匀圆周密度这一特例（Fernández-Durán，2004b 和 2007）。此外，通过使用 NNTS 圆形密度族，可以生成所需的、与圆周均匀密度更为接近的连接函数 g；下文将对此进行解释。Pewsey 和 Kato（2016）针对方程 5 中 Wehrly 和 Johnson（1980）的圆周-圆周模型，开发了一种拟合优度检验。他们的检验考察了每个圆周随机变量 Θ_1 和 Θ_2 与连接函数 g 的辐角 $\Omega = 2\pi\{F_{\Theta_2}(\theta_2) \pm F_{\Theta_1}(\theta_1)\}(mod 2\pi)$ 之间的独立性。与本文类似，他们通过检验两个二元随机向量 $(2\pi F_{\Theta_1}(\theta_1)\ 2\pi F_{\Omega}(\omega))^{\top}$ 是否满足环面均匀性来实施独立性检验

和（$2\pi F_{\Theta_2}(\theta_2)$ $2\pi F_{\Omega}(\omega)$）$^{\top}$。

基于非负三角和（NNTS）的圆周密度函数，对于圆周（角度）随机变量 $\Theta \in (0, 2\pi]$（参见 Fernández-Durán，2004b）定义为

$$f_{\Theta}(\theta; M, \underline{c}) = \frac{1}{2\pi}\left|\left|\sum_{k=0}^{M} c_k e^{ik\theta}\right|\right|^2 = \frac{1}{2\pi}\sum_{k=0}^{M}\sum_{l=0}^{M} c_k \bar{c}_l e^{i(k-l)\theta} \tag{8}$$

其中 $i = \sqrt{-1}$、c_k 为复数 $c_k = c_{rk} + ic_{ck}$，适用于 $k = 0$、.....、M，而 $\bar{c}_k = c_{rk} - ic_{ck}$ 是 c_k 的共轭。为使密度函数 $f_{\Theta}(\theta; M, \underline{c})$ 有效且积分为 1，

$$\sum_{k=0}^{M} ||c_k||^2 = 1. \tag{9}$$

其中 $||c_k||^2 = c_{rk}^2 + c_{ck}^2$。参数向量 $\underline{c} = (c_0, c_1, \ldots, c_M)^{\top}$ 的参数空间是超球面表面的一个子集，因为 $\underline{c}$ 和 $-\underline{c}$ 会给出相同的 NNTS 密度，而 $\underline{c}$ 的共轭按逆序排列后，也会给出与 $\underline{c}$ 相同的 NNTS 密度。因此，为了确保参数向量 $\underline{c}$ 的可识别性，并依据公式 9，施加以下约束：$c_{c0} = 0$ 和$c_{r0} \geq 0$，即 c_0 是非负实数，且 $c_0^2 \geq ||c_M||^2$。共有 $2M$ 个自由参数 c，而 M 是一个附加参数，它决定了定义该密度的求和项总数，并且与该密度可能具有的最大众数数量有关。我们注意到，具有 $M = M_1$ 的 NNTS 模型嵌套于具有 $M = M_2$ 的 NNTS 模型中，其中 $M_2 > M_1$。区间 $(0, 2\pi]$ 上的圆周均匀密度对应于一个满足 $M = 0$，$f_{\Theta}(\theta; M = 0, \underline{c}) = \frac{1}{2\pi}$ 的 NNTS 密度，并且它嵌套于所有满足 $M > 0$ 的 NNTS 模型中。等价地，对于具有 $M > 0$ 的 NNTS 模型，当 c_0 趋近于 1 时，NNTS 密度收敛于圆周均匀密度；同样，对于具有 $\underline{c} = (1, 0, 0, \ldots, 0)^{\top}$ 的 NNTS 密度，即 $c_0 = 1$ 且 $\underline{c}$ 的其余元素均为零时，对应的就是圆周均匀密度。利用这一性质，可通过从参数 c_0 尽可能接近 1 的 NNTS 备择密度中生成样本，来评估所提出的独立性检验的检验功效。

Fernández-Durán 和 Gregorio-Domínguez（2010）开发了一种基于流形优化的高效算法，用于求取 c 个参数的最大似然估计。该算法已纳入

一系列基于 NNTS 模型分析圆形数据的其他程序中，并包含在免费的 R （R核心团队，2021）软件包 CircNNTSR 中（Fernández-Durán 和 Gregorio-Domínguez，2016）。

3 双变量独立性所提检验

对于绝对连续的独立同分布（i.i.d.）圆形均匀随机变量， $U_1, U_2, \ldots, U_d \sim U(0, 2\pi]$，考虑 $\sum_{k=1}^{d} U_k \sim U(0, 2\pi]$，即独立同分布的圆形均匀随机变量之和也服从圆形均匀分布。这是因为圆形均匀随机变量的特征函数 $\psi(t) = E(e^{it\Theta})$在 $t = 0$ 时等于一，在其他位置均等于零（$t \neq 0$）。因此，如果圆形均匀随机变量之和不服从圆形均匀分布，则这意味着这些圆形均匀随机变量并不独立（参见 Mardia 和 Jupp，2000年，第35页）。所提独立性检验基于这一结果，通过考察任意（线性或圆形）绝对连续随机变量的角度概率积分变换来进行。设 $X_1, X_2, \ldots, X_d$ 为 d 任意绝对连续随机变量。 X_k 的角度概率变换定义为角度（圆形）随机变量 $APIT(X_k) = 2\pi F_{X_k}(X_k)$，其在单位圆上服从均匀分布。考虑相互联合独立的原假设，即 $APIT(X_1)$,$APIT(X_2), \ldots, APIT(X_d)$为 $i.i.d.\ U(0, 2\pi]$，则 $\sum_{k=1}^{d} \pm APIT(X_k) \sim U(0, 2\pi]$ 也服从圆形均匀分布。所提二元独立性检验基于检验 $APIT(X_1) + APIT(X_2)$（$APIT(X_1) - APIT(X_2)$）是否具有圆周均匀性，其中 X_1 和 X_2为绝对连续（圆形或线性）随机变量。检验二元独立性等价于检验角度概率积分变换之和（差）是否服从均匀分布。为检验角度积分变换之和（差）的圆周均匀性，我们采用了瑞利检验和 Pycke 检验（Pycke，2010年）。瑞利检验考虑单峰圆形密度这一备择假设，其针对样本 $\theta_1, \theta_2, \ldots, \theta_n$ 的检验统计量定义如下：

定义如下（Mardia 和 Jupp，2000年）：如下：

$$T_{RT} = 2n\bar{R}^2 \tag{10}$$

¯其中，R 为样本平均合向量长度。统计量 T_{RT} 渐近服从自由度为 2 的卡方分布。本文所用 Rayleigh 检验的 p 值对应 Fisher（1993，第 70 页）给出的近似值，该近似值对小样本量进行了校正。Pycke 检验考虑了一种多峰密度的备择假设，其针对样本 $\theta_1, \theta_2, \ldots, \theta_n$ 的检验统计量定义如下：

$$T_{PT} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{2(\cos(\theta_i - \theta_j) - \sqrt{0.5})}{1.5 - (2\sqrt{0.5}\cos(\theta_i - \theta_j))} \right). \tag{11}$$

Pycke 检验的临界值通过模拟获得。

针对两个绝对连续随机变量，X_1 和 X_2，所提出的独立性检验步骤如下。首先，根据每个随机变量的观测值计算伪观测值（经验分布）ˆ ˆF_{X_1} 和F_{X_2}。其次，将伪观测值ˆ ˆ乘以 2π，计算角度概率积分变换（APIT），（2πF_{X_1} 和2πF_{X_2}）。第三，计算两个角度概率积分变换之和（差）的模 2π。最后，将瑞利检验或 Pycke 检验应用于 APIT 之和（差）的观测值向量。若随机变量之间呈正关联，则所提出的独立性检验必须应用于 APIT 之差；若呈负关联，则必须应用于 APIT 之和。若无法预先判断关联是正还是负，则可以计算两个检验统计量，一个针对 APIT 之和，另一个针对 APIT 之差，然后根据某种多重检验校正方法（例如 Bonferroni 校正法）对两个 p 值中的较小值进行调整（参见 Cinar 和 Viechtbauer，2022年；Goeman 和 Solari，2014年）。实际应用中，用户可以绘制 APIT 之和（差）的直方图，并根据观测到的众数数量决定采用瑞利检验还是 Pycke 检验。若

有疑问，则优先采用 Pycke 检验。还可以考虑 Ajne、Watson 和 Rao 提出的其他圆周均匀性综合检验（参见 Mardia和Jupp，2000年）。

用 $M = 1$ 拟合 APIT 总和（差值）的 NNTS 模型后，可定义以下依赖性度量：

$$\hat{\lambda}_{c_0} = \frac{M+1}{M}\left(1 - \hat{c}_0^2\right) = 2\left(1 - \hat{c}_0^2\right) \tag{12}$$

其中，校正项 $\frac{M+1}{M} = 2$ 源于这样一个事实：在 $M = 1$的情况下，零点附近集中度最高的 NNTS 密度具有参数向量$\underline{c}$，其每个分量的平方范数都等于$\frac{1}{M+1} = \frac{1}{2}$。这意味着 $\frac{1}{2} \leq c_0^2 \leq 1$，并且当 $M = 1$ 取值于区间 $[0, 1]$ 时，$\hat{}\,\lambda_{c_0}$ 也如此；取值接近零表示依赖性较低（独立），而取值越接近一，则表示所考察随机变量之间的依赖性越高。对于相互独立的随机变量，λ_{c_0} 的理论值等于零。依赖性度量λ_{c_0} 在圆周-圆周和圆周-线性情形下尤其有用。

4 模拟研究

本节通过模拟研究比较所提检验、Wilks 检验以及基于经验Copula的独立性检验的检验功效。我们使用已知的依赖结构参数和已知的边际密度，从不同的多元分布生成数据。样本量设为 20、50、100 和 200。对于给定的显著性水平 α （10%、5% 和 1%），通过以下方式获得各项检验的检验功效：从双变量密度中生成 100 个指定样本量的样本，计算每个样本对应的检验统计量，并统计检验统计量有多少次取值导致在给定的 α 下拒绝独立性原假设。因此，报告的检验功效取值范围为 0–100，可解释为百分比。圆周均匀性的瑞利检验通过

使用 $circular\ R$ package（Agostinelli 和 Lund，2017）进行。圆周均匀性的 Pycke 检验使用 $CircMLE\ R$ package（Fitak 和 Johnsen，2017；Landler 等，2019）进行。使用 R package $copula$ 计算经验 Copula 检验，并通过使用 R package$CircNNTSR$ 拟合 NNTS 模型（ $M = 1$ ），获得依赖性度量ˆ λ_{c_0} （Fernández-Durán 和 Gregorio-Domínguez，2016）。

4.1 圆周-线性模型

表1列出了所提出的独立性检验在采用瑞利（ART）和 Pycke（APT）圆周均匀性检验时的检验功效，并与 Wilks（WT）和经验Copula（ECT）独立性检验的检验功效进行了比较。模拟样本来自 Johnson 和 Wehrly 的圆周-线性模型，其中圆周边际密度为含有 $M = 3$ 的 NNTS 密度，该密度绘于图1的第一幅图中；模型包含三种不同的线性分布（指数分布、高斯分布和柯西分布），以及连接圆周密度 g，它对应于含有 $M = 3$ 的 NNTS 密度。参数 c_0 取五个不同的值（0.7、0.8、0.9、0.99 和 0.9999），以体现圆周随机变量与线性随机变量之间不同程度的关联。图1最后一幅图展示了这一情况，其中绘出了参数 c_0 取五个不同值时的角度连接函数。 $c_0 = 0.9999$ 的情形对应于近似圆周均匀密度，也对应于独立性原假设（参见图1最后一幅图）。总体而言，与 ECT 检验相比，所提出的 ART 和 APT 检验表现出显著更高的检验功效；只有在样本量较大且 c_0 取值较低（0.7、0.8 和 0.9）时，ECT 检验的功效才与之相近，这也代表了高度依赖的圆周随机变量和线性随机变量。与 ART、APT 和 ECT 检验相比，WT 的检验功效明显较低。

4.2 圆形-圆形模型

对于 Johnson 和 Wehrly 的圆周-圆周模型，我们采用了与圆周-线性模型相同的角度连接密度和其中一种边缘圆周密度。图1展示了边缘圆周密度的图形，它们分别对应含有 $M = 3$ 和$M = 2$ 的 NNTS 密度；同时还展示了 Johnson 和 Wehrly 圆周-圆周模型中的角度连接密度 g，其对应于含有 $M = 3$ 的 NNTS 密度，参数 c_0 取五个不同的值（0.7、0.8、0.9、0.99 和 0.9999）。与圆周-线性模型的结果类似，ART 和 APT 的检验功效高于 WT 和 ECT 检验。鉴于所涉及的所有圆周密度均具有多峰性，在所提出的两种独立性检验中，APT 的检验功效高于 ART；只有在最大样本量 200 时，两者的功效才相近。此外，对于高度依赖的圆周随机变量（c_0 = 0.7、0.8 或 0.9）以及较大的样本量 100 或 200，ECT 检验的功效与ˆ ART 和 APT 检验相近。表1和表2第四列所列的依赖度量 λ_{c_0} 的平均值，在 c_0 =0.7、0.8 和 0.9 时相近；而在 c_0=0.99 和 0.9999 时，这些值更小，这进一步反映出，当 c_0 接近 1 时，随机变量之间的关联非常弱。

4.3 线性-线性模型

表 3 和表 4 列出了从两个变量均为线性的二元分布中模拟样本时不同检验的检验功效。第一种情况下使用了高斯 copula，第二种情况下使用了 Frank copula。

4.3.1 二元高斯分布 Copula

二元高斯分布（正态）copula 对应于一个多元分布，其定义如下:

$$C(u_1, u_2) = \Phi_\Gamma(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) \tag{13}$$

其中， Φ_Γ 是具有零均值向量和相关矩阵 Γ的多元正态分布，而 Φ()是一元标准正态分布函数。将单位矩阵用作相关矩阵，可得到独立Copula $\Gamma = I$，即 $C(u_1, u_2) = u_1 u_2$ 。

表3比较了所提出的独立性检验在采用瑞利（ART）和 Pycke（APT）圆周均匀性检验时的检验功效，并与 Wilks（WT）和经验 Copula（ECT）独立性检验的检验功效进行比较。所用模拟样本来自具有高斯 copula 的二元线性-线性分布，其边缘分布分别采用 Herwatz 和 Maxand（2020年）研究中的三种不同情形：指数分布、高斯分布和柯西分布。对于高斯 copula，我们考虑了相关系数的五个不同取值 ρ （0、0.25、0.5、0.75 和 0.99）。当相关系数等于零时，对应于独立性原假设；报告的检验功效对应于各检验的尺寸，预期与显著性水平 α （10%、5% 和 1%）相近。总体而言，WT 和 ECT 的检验功效均高于所提出的 ART 和 APT 检验。不过，当样本量较大且相关系数较高（0.75 和 0.99）时，ART 和 APT 检验的检验功效与 WT 和 ECT 检验相近。正如预期的那样，对于指数分布和柯西边缘分布，与专为高斯边缘分布设计的情形相比，WT 的检验功效有所降低；也就是说，当边缘分布不是高斯分布时，WT 的检验功效会变差。对于高斯边缘分布，当样本量较大时，WT 的检验功效较高，且在某些情况下，其检验功效高于 ECT。就所有检验的尺寸而言，鉴于总共使用了100个样本，各检验的尺寸似乎都大致正确。表3的第四列给出了依赖度量 λ_{c_0} 的平均值，正如预期的那样，这些值会随着相关系数

的增大而增大。

4.3.2 Frank二元Copula

Frank二元Copula定义如下:

$$C(u,v) = -\frac{1}{\varphi}\ln\left(1 + \frac{(e^{-\varphi u}-1)(e^{-\varphi v}-1)}{e^{-\varphi}-1}\right) \tag{14}$$

其中 $u, v \in (0,1]$。参数 φ 的取值范围为区间 $(0,\infty)$，且在极限 $\varphi = 0$ 时，Frank copula 对应于独立Copula。在表4中，通过采用依赖参数 φ 的五个不同取值（0、5、10、15 和 50），将先前研究中的高斯 copula 替换为 Frank copula。对于当$\varphi = 0$ 时在极限下得到的独立情形，也计算了将该极限视为 $\varphi = 0$ 时的检验功效。总体而言，表4得出了与表3相同的结论，这进一步表明，与线性-线性情形相比，所提出的检验 ART 和 APT 的特性更适用于圆周 - 圆周和圆周 - 线性情形。

5 实际圆形-圆形和圆形-线性数据的应用

5.1 二元独立性检验

5.1.1 圆周-线性真实示例

图 2 展示了所研究真实案例的散点图。我们将所提出的独立性检验应用于风向（圆形变量）和臭氧浓度（线性变量）的圆周-线性数据。这些数据最初由 Johnson 和 Wehrly（1977年）分析，后来被收录为 Fisher（1993）中的数据集B.18。1975年4月18日至6月29日期间，在密尔沃基的一处气象站每隔四天于早上6点测量一次，共获得19次观测，

1975年。这些数据的散点图见图 2 左上图，其中展示了风向和臭氧浓度的取值；同时也显示，在考虑圆周随机变量的周期性时，圆周变量与线性变量之间可能存在正相关。对圆周变量和线性变量的角度概率变换之差应用 Pycke 检验和瑞利圆均匀性检验后，我们分别得到 p 值 0.0133 和 0.0077，因此在 5% 的显著性水平下，拒绝风向与臭氧浓度之间的独立性（均匀性）原假设。依赖性度量值 λ_{c_0} 计算为 0.4383。Fisher（1993）考虑了给定风向时臭氧浓度条件期望值的正弦波函数形式，并得出了相同结论。

Fisher（1993）分析的第二个示例是一个关于31只小蓝滨螺在从其正常栖息地移植后所行进方向和距离的数据集。图 2所示右上方图中包含了该数据集的散点图，进一步表明方向与行进距离之间可能存在负关联。对方向和距离的角度概率变换之和应用均匀性检验时，Pycke 检验的 p 值为0.0087，Rayleigh检验的 p 值为0.0096，这意味着我们拒绝独立性；这一结果与 Fisher（1993）在拟合具有非恒定离散度的、带有 von Mises误差的圆形-线性回归模型时所得的结果一致。依赖度量 $\hat{\lambda_{c_0}}$ 计算为0.2668。

5.1.2 圆形-圆形实际示例

圆形-圆形独立性检验中的第一个示例，对应于在密尔沃基的一处气象监测站测量的风向对。测量连续进行了21天，每天分别于6:00和12:00进行，最初收录于 Johnson 和 Wehrly（1977）。图 2 左下方的图中包含风向对的散点图，显示这两个角度之间可能存在正关联。Fisher（1993）

将该数据集列为 B.21 数据集。Fisher（1993）的主要结论是，基于圆形-圆形相关系数进行假设检验时，风向之间存在很强的正关联。对角度概率变换之差应用所提出的方法后，Pycke 检验的 p 值为 0.0148，Rayleigh检验的 p 值为 0.0075，这意味着在 5% 的显著性水平下，我们拒绝了两个角度相互独立的原假设^。依赖度量 λ_{c_0} 的计算结果为 1。

第二个例子涉及蛋白质中丙氨酸-丙氨酸-丙氨酸片段的 233 对二面角（ϕ，ψ）。Fernández-Durán（2007）最初使用双变量 NNTS 模型分析了这些数据，结果显示这两个角度之间存在依赖关系。这两个角度的散点图见图 2 右下图，但无法明显判断两者之间的关联是正还是负。对这两个角度的角度概率变换之差应用独立性检验后，瑞利圆周均匀性检验和 Pycke圆均匀性检验分别得到 0.0014 和 0.0064 的 p 值，因此明确拒绝独立性原假设，转而支持正关联。

6 结论

利用独立圆形均匀随机变量之和仍服从圆形均匀分布这一结果，我们提出了一种基于角度积分概率变换的通用独立性检验。我们展示了该检验的用法，尤其是在至少一个变量为角度（即圆周随机变量）时的应用；这进一步表明，在圆周情形下检验独立性等价于检验圆周均匀性。从本文所呈现的模拟研究可以清楚看出，所提检验特别适用于圆周-圆周和圆周-线性两种双变量随机变量对，其检验功效高于 Wilks 检验和经验 copula 独立性检验。我们通过从

NNTS 密度中模拟样本得出了这一结论，并对其与圆形均匀分布的接近程度进行了控制。尽管所提独立性检验可用于线性-线性情形，但其检验功效较低；随着样本量增加，其检验功效会趋近于常用独立性检验的检验功效。所提独立性检验具有实用性；当用于独立性检验的随机变量中至少有一个是圆周随机变量时，它表现出更优的性能。此外，我们还通过考察角度概率积分变换之和（差）相对于 $M = 1$ 的情形，提出了一种新的依赖性度量。

致谢

我们衷心感谢 Asociación Mexicana de Cultura, A.C. 提供的支持。

参考文献

[1] Agostinelli, C. 和 Lund, U.（2017）。R 软件包 'circular'：圆形统计学（版本 0.4-93）。网址：https://r-forge.r-project.org/projects/circular/

[2] Berrett, T.B. 和 Samworth, R.J.（2019）。基于互信息的非参数独立性检验，《Biometrika》，106（3），第 547–556 页。

[3] Blum, J.R.、Keifer, J. 和 Rosenblatt, M.（1961）。基于样本分布函数的分布无关独立性检验。《数学统计年刊》，32（2），第 485–498 页。

[4] Cinar, O. 和 Viechtbauer, W.（2022）。用于合并独立和相关 p 值的 poolr 软件包。《统计软件杂志》，101，1–42。

[5] Csörg˝o, S.（1985）。基于经验特征函数的独立性检验。《多元分析杂志》，16，290–299 页。

[6] Deheuvels, P.（1981）。多元分布无关独立性检验的渐近分解，《多元分析杂志》，11 (1), pp. 102–113。

[7] DeWet, T.（1980）。独立性的克拉梅尔-冯·米塞斯检验。《多元分析杂志》，10,pp. 38–50。

[8] Einmahl, J.H. 和 McKeague, I.W.（2003）。基于经验似然的假设检验。《伯努利》，pp. 267–290。

[9] Fernández-Durán, J.J.（2004）。使用 Copula 对地面臭氧浓度建模。载于 Erickson, G.J. 和 Zhai, Y.（编）《第23届科学与工程中的贝叶斯推断和最大熵方法国际研讨会论文集》，会议于2003年8月3日至8日在怀俄明州杰克逊霍尔举行。AIP会议论文集，第707卷。纽约：美国物理联合会，pp. 406–413。

[10] Fernández-Durán, J.J.（2004b）。基于非负三角和的圆分布。《生物统计学》，60，499–503页。

[11] Fernández-Durán, J. J.（2007）。基于非负三角和的圆分布构建的圆周-线性和圆周-圆周数据模型。《生物统计学》，63（2），579–585页。

[12] Fernández-Durán, J.J. 和 Gregorio-Domínguez, M.M.（2010年）。使用流形上的类牛顿算法对非负三角和模型进行最大似然估计。《电子统计学杂志》，4，1402-10。

[13] Fernández-Durán, J. J. 和 Gregorio-Domínguez, M. M.（2016年）。CircNNTSR：一个用于使用非负三角和对圆形、多元圆形和球面数据进行统计分析的 R 软件包。《统计软件杂志》，70，第 1–19 页。

[14] Fisher, N.I. 和 Lee, A. J.（1981）。角度-线性关联的非参数度量，《Biometrika》，68，页码 629–636。

[15] Fisher, N.I. 和 Lee, A. J.（1982）。角度-角度关联的非参数度量，《Biometrika》，69，页码 315–321。

[16] Fisher, N.I. 和 Lee, A. J.（1983）。圆形数据相关系数，《Biometrika》，70，页码 327–332。

[17] Fisher, N.I.（1993）。圆形数据的统计分析。剑桥、纽约：剑桥大学出版社。

[18] Fitak, R. R. 和 Johnsen, S.（2017）。让动物定向数据分析回归完整周期：基于最大似然的模型方法。《实验生物学杂志》，220，页码 3878–3882。

[19] Genest, C. 和 Rémillard, B.（2004）。基于经验Copula过程的独立性与随机性检验，《Test》，13，335–369页。

[20] Genest, C. 和 Verret, F.（2005）。针对Copula模型的局部最强独立性秩检验，《非参数统计》，17，521–539页。

[21] Genest, C.、Nešlehová, J.G.、Rémillard, B. 和 Murphy, O.A.（2019）。任意分布中的独立性检验。Biometrika，106，47–68页。

[22] Goeman, J.J. 和 Solari, A.（2014）。基因组学中的多重假设检验，《医学统计学》，33，1946–1978。

[23] Herwatz, H. 和 Maxand, S.（2020年）。非参数独立性检验：综述与比较模拟研究，并应用于印度营养不良数据。《统计论文》，61，第 2175–2201 页。

[24] Hoeffding, W.（1948）。非参数独立性检验。《数学统计年鉴》，19（4），第546–557页。

[25]Hofert, M.、Kojadinovic, I.、Maechler, M. 和 Yan, J.（2022）。copula：基于 Copula 的多元依赖。R 软件包版本 1.1-1，网址：https://CRAN.R-project.org/package=copula

[26] Jammalamadaka, S.R. 和 SenGupta, A.（2001）。圆形统计学专题。新泽西州里弗埃奇：世界科技出版公司。

[27] Joe, H.（1990）。多元依赖的多元熵度量。《美国统计学会会刊》，84，第157–164页。

[28] Johnson, R. A. 和 Wehrly, T.（1977年）。角相关性与角线性相关的度量和模型。《皇家统计学会会刊》B辑，39（2），第222–229页。

[29] Kallenberg, W.C.M. 和 Ledwina, T.（1999）。用于独立性的数据驱动秩检验。《美国统计学会杂志》，94，页码 285–301。

[30] Kenda ll, M.G.（1938）。一种新的秩相关度量。Biometrika，页码 81–93 .

[31] Kendall, M.G. 和 Stuart, A.（1951）。《高等统计学理论》第 2 卷。推断与关系。纽约：Hafner 出版公司。

[32] Kojadinovic, I. 和 Yan, J.（2010年）。使用 Copula R 软件包对具有连续边缘分布的多变量分布建模。《统计软件杂志》，34（9）。

[33] Landler, L.、Ruxton, G. D. 和 Malkemper, E. P.（2019）。Hermans-Rasson检验：一种适用于生物学圆形统计学的强有力替代方案，可替代瑞利检验。BMC生态学，19：30。

[34] Mardia, K.V. 和 Kent, J.T.（1991）。拟合优度与独立性的 Rao得分检验。Biometrika，78（2），第 355–363 页。

[35] Mardia, K.V. 和 Jupp, P.E.（2000）。方向统计。奇切斯特、纽约：约翰·威立父子公司。

[36] Nelsen, R.（1999）。Copula导论。纽约：施普林格出版社。

[37] Pearson, K.（1920）。相关性历史札记。Biometrika，第 25–45 页。

[38] Pewsey, A.、Neuhäuser, M. 和 Ruxton, G.D.（2013）。R 中的圆形统计学，牛津大学出版社，英国牛津。

[39] Pewsey, A. 和 Kato, S.（2016）。Wehrly-Johnson 二元圆分布的参数自助法拟合优度检验。《统计与计算》，26，1307-1317。

[40] Pfister, N., Bühlmann, P., Schölkopf, J.N. 和 Peters, J.（2018）。用于联合独立性检验的基于核的方法，《皇家统计学会杂志》B辑，80（1），第 5–31 页。

[41] Pycke, J-R.（2010）。若干针对圆分布均匀性的检验，对多峰替代假设具有较高检验功效。《加拿大统计学杂志》，38，80–96。

[42] R核心团队（2021）。R：一种用于统计计算的语言和环境。奥地利维也纳：R统计计算基金会。网址：http://www.R-project.org/。

[43] Roy, A.（2020年）。若干用于检验具有任意概率分布的多个随机变量之间独立性的基于Copula的检验，《Stat》，9，1

[44] Roy, A.、Ghosh, A.K.、Goswami, A. 和 Murthy, C.A.（2020年）。若干新的、无分布假设的用于检验多个随机变量之间独立性的基于Copula的检验，Sankhya A，84，第556–596页。

[45] 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出版物》，8，第229–231页。

[46] Spearman, C.（1904年）。两个事物之间关联的证明与测量。《美国心理学杂志》，15（1），第72–101页。

[47] Upton, G.J.G. 和 Fingleton, B.（1989年）。《实例空间数据分析》第2卷（类别数据与方向数据）。奇切斯特、纽约：约翰·威立父子公司。

[48] Wehrly, T. 和 Johnson, R.A.（1980年）。角度观测依赖性的二元模型及相关马尔可夫过程。《Biometrika》，67（1），第255–256页。

[49] Wilks, S.（1935年）。关于 k 组正态分布统计变量的独立性。《计量经济学》，3，第309–326页。

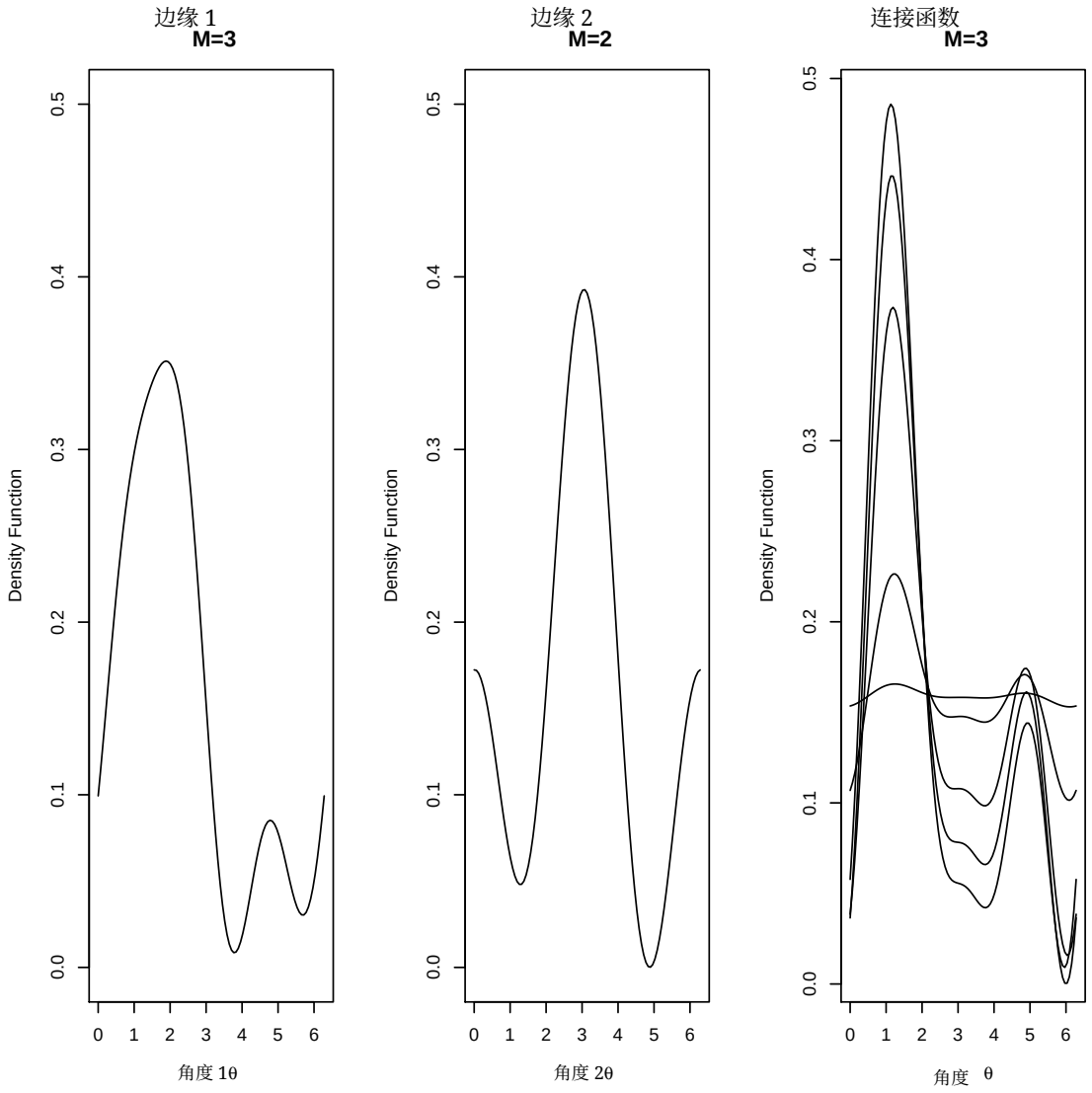

图1：圆周-圆周 Copula 模型：前两幅图显示边缘 NNTS 圆形密度（$M_1 = 3$ 和 $M_2 = 2$），第三幅图显示参数 c_0 取不同值（0.7、0.8、0.9、0.99 和 0.9999）时的角度（圆形）联合密度。$c_0 = 1$ 对应圆周均匀密度（零假设独立模型）。

边缘分布	SS	c_0	$\hat{\lambda}_{c0} = 2(1-\tilde{c}_0^2)$	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
				RT	PT	WT	ECT	RT	PT	WT	ECT	RT	PT	WT	ECT
$\Theta \sim NNTS(M=3)$	20	0.7	0.61	91	89	29	43	89	88	22	32	68	64	8	3
$X \sim Exp(1)$	20	0.8	0.66	88	85	26	48	83	79	22	39	67	64	6	5
	20	0.9	0.62	86	85	39	47	83	80	21	35	58	61	8	11
	20	0.99	0.22	34	35	18	11	28	26	10	9	10	7	3	2
	20	0.9999	0.15	13	10	12	10	3	5	7	7	1	1	1	2
	50	0.7	0.46	100	100	33	79	100	99	20	57	97	97	11	28
	50	0.8	0.51	100	100	33	86	100	100	18	61	99	98	6	29
	50	0.9	0.5	100	100	43	89	100	100	34	72	99	98	9	48
	50	0.99	0.15	71	81	21	40	64	70	13	20	40	41	4	3
	50	0.9999	0.04	12	10	7	10	6	6	3	4	1	2	0	2
	100	0.7	0.45	100	100	51	100	100	100	39	100	100	100	19	89
	100	0.8	0.49	100	100	47	100	100	100	30	99	100	100	16	94
	100	0.9	0.47	100	100	72	100	100	100	56	100	100	100	26	90
	100	0.99	0.12	90	95	31	55	82	92	22	41	66	80	7	18
	100	0.9999	0.02	10	7	11	9	5	4	4	4	1	0	1	2
	200	0.7	0.44	100	100	71	100	100	100	64	100	100	100	42	100
	200	0.8	0.49	100	100	65	100	100	100	58	100	100	100	31	100
	200	0.9	0.45	100	100	86	100	100	100	76	100	100	100	56	100
	200	0.99	0.11	100	100	39	97	100	100	31	79	99	100	13	55
	200	0.9999	0.01	11	14	10	12	9	7	3	8	3	1	0	0
$\Theta \sim NNTS(M=3)$	20	0.7	0.62	89	86	13	36	79	80	7	25	55	52	2	4
$X \sim N(0,1)$	20	0.8	0.63	89	85	17	39	78	76	8	23	53	52	6	6
	20	0.9	0.64	94	88	20	48	86	78	8	37	62	57	1	9
	20	0.99	0.32	38	30	17	23	28	19	11	19	11	10	1	8
	20	0.9999	0.16	13	18	18	14	7	8	10	12	1	2	3	3
	50	0.7	0.56	100	100	19	92	100	100	8	77	99	99	1	34
	50	0.8	0.59	100	100	12	89	100	100	7	78	99	99	2	33
	50	0.9	0.49	100	100	29	85	100	100	18	64	98	99	7	30
	50	0.99	0.13	59	63	21	27	52	55	13	11	30	33	5	5
	50	0.9999	0.04	14	12	8	10	5	6	4	5	2	2	1	1
	100	0.7	0.43	100	100	12	100	100	100	7	100	100	100	2	95
	100	0.8	0.5	100	100	14	100	100	100	10	100	100	100	0	97
	100	0.9	0.48	100	100	44	100	100	100	29	98	100	100	13	93
	100	0.99	0.13	91	97	21	64	86	91	14	36	66	76	4	23
	100	0.9999	0.02	10	14	7	13	5	7	4	3	1	1	1	2
	200	0.7	0.44	100	100	18	100	100	100	9	100	100	100	4	100
	200	0.8	0.48	100	100	16	100	100	100	10	100	100	100	3	100
	200	0.9	0.46	100	100	64	100	100	100	56	100	100	100	31	100
	200	0.99	0.11	100	100	30	92	100	100	15	81	97	100	8	50
	200	0.9999	0.01	11	7	9	4	6	3	8	1	0	0	3	0
$\Theta \sim NNTS(M=3)$	20	0.7	0.52	81	75	17	28	73	66	13	14	46	46	6	6
$X \sim Cauchy(0,1)$	20	0.8	0.58	82	84	15	40	78	72	9	22	53	50	1	5
	20	0.9	0.6	81	79	22	34	70	71	14	22	56	52	3	9
	20	0.99	0.26	34	37	12	21	22	25	8	9	7	10	1	4
	20	0.9999	0.15	12	16	12	15	6	10	10	6	2	6	3	0
	50	0.7	0.49	100	100	26	84	100	100	17	68	99	99	1	35
	50	0.8	0.5	100	100	15	89	100	100	8	80	99	98	0	35
	50	0.9	0.48	100	100	13	88	100	100	6	67	99	99	2	32
	50	0.99	0.13	62	69	13	28	52	56	8	19	26	26	2	5
	50	0.9999	0.05	12	12	15	9	4	5	9	3	1	2	3	1
	100	0.7	0.45	100	100	13	100	100	100	10	100	100	100	4	89
	100	0.8	0.49	100	100	14	100	100	100	10	99	100	100	3	89
	100	0.9	0.47	100	100	12	100	100	100	7	97	100	100	2	83
	100	0.99	0.12	97	94	12	52	88	94	10	34	74	84	3	15
	100	0.9999	0.02	9	14	7	7	4	5	6	5	0	1	2	0
	200	0.7	0.43	100	100	17	100	100	100	11	100	100	100	0	100
	200	0.8	0.48	100	100	17	100	100	100	13	100	100	100	1	100
	200	0.9	0.46	100	100	15	100	100	100	7	100	100	100	1	100
	200	0.99	0.1	100	100	13	92	99	100	5	84	97	100	3	56
	200	0.9999	0.01	14	13	10	7	7	5	8	2	2	1	1	1

表1：NNTS角度联合密度 g 圆周-线性检验功效研究：通过模拟从 Johnson 和 Wehrly 圆周-线性密度函数中抽取的样本，比较所提检验的检验功效。该密度函数由一个 NNTS 角度联合密度（ $M = 3$）、一个 NNTS边缘密度函数（$M = 3$ ，见图1）以及三种不同的线性边缘分布（指数分布、高斯分布和柯西分布）构成。分别采用瑞利（ART）和 Pycke（APT）圆周均匀性检验、Wilks（WT）检验及经验Copula（ECT）检验，每种情况下模拟100次，样本量为20、50、100和200。NNTS 角度联合密度的参数 c_0 取五个不同的值（0.7、0.8、0.9、0.99 和 0.9999）。 $c_0 = 1$ 对应于零假设独立模型。

边缘分布	SS	c_0	$\lambda_{c0} = 2(1-\tilde{c}_0^2)$	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
				RT	PT	WT	ECT	RT	PT	WT	ECT	RT	PT	WT	ECT
$\Theta_1 \sim NNTS(M_1 = 3)$	20	0.7	0.63	87	80	16	38	78	73	11	29	59	61	1	11
$\Theta_2 \sim NNTS(M_2 = 2)$	20	0.8	0.6	84	81	13	41	72	74	9	28	57	57	3	10
	20	0.9	0.59	81	79	20	36	76	67	15	28	46	49	5	7
	20	0.99	0.25	34	33	19	15	26	23	13	12	10	8	7	4
	20	0.9999	0.13	11	13	12	10	7	7	7	10	1	0	1	1
	50	0.7	0.45	100	100	16	88	99	99	7	63	99	98	1	41
	50	0.8	0.5	100	100	15	91	100	100	5	70	98	98	0	38
	50	0.9	0.49	100	100	28	92	100	100	14	70	100	100	6	34
	50	0.99	0.14	70	73	25	36	55	63	16	17	36	36	7	6
	50	0.9999	0.05	17	14	9	12	6	6	4	3	0	1	3	1
	100	0.7	0.42	100	100	13	100	100	100	9	100	100	100	3	92
	100	0.8	0.47	100	100	17	100	100	100	9	100	100	100	2	93
	100	0.9	0.46	100	100	38	100	100	100	28	100	100	100	10	96
	100	0.99	0.12	90	95	32	60	83	92	23	41	64	85	10	15
	100	0.9999	0.02	9	14	17	8	4	7	7	1	3	1	1	0
	200	0.7	0.43	100	100	25	100	100	100	19	100	100	100	6	100
	200	0.8	0.48	100	100	18	100	100	100	12	100	100	100	6	100
	200	0.9	0.45	100	100	51	100	100	100	45	100	100	100	21	100
	200	0.99	0.1	100	100	42	86	100	100	31	80	99	99	15	40
	200	0.9999	0.01	8	10	12	9	4	6	6	4	0	0	2	1

表2：NNTS角度联合密度的圆周-圆周功效研究：通过模拟从由NNTS角度联合密度构造的Johnson 和 Wehrly圆形-圆形密度函数中抽取的样本（样本量为20、50、100和200，每种情况重复100次），比较了采用瑞利（ART）和Pycke（APT）圆周均匀性检验、Wilks（WT）检验及经验Copula（ECT）检验所实现的所提检验的检验功效。该NNTS角度联合密度由五种不同的依赖参数 c_0 （0.7、0.8、0.9、0.99和0.9999）定义。 $c_0 = 1$ 的情形对应于零假设独立模型。NNTS角度联合密度和NNTS圆形边缘密度的图示见图1。

边缘分布	SS	ρ	$\hat{\lambda}_{c0} = 2(1-\hat{c}_0^2)$	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
				ART	APT	WT	ECT	ART	APT	WT	ECT	ART	APT	WT	ECT
$X_1 \sim Exp(1)$	20	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
$X_2 \sim Exp(2)$	20	0.75	0.57	75	69	98	97	68	60	92	92	50	45	83	75
	20	0.5	0.25	38	36	64	60	24	25	51	47	9	7	40	24
	20	0.25	0.15	17	20	33	29	10	8	24	15	3	4	9	3
	20	0	0.14	10	14	12	10	7	9	8	5	0	0	2	3
	50	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
	50	0.75	0.48	99	98	100	100	99	96	100	100	91	89	100	100
	50	0.5	0.14	57	51	93	94	46	38	87	90	23	21	71	85
	50	0.25	0.06	18	17	52	42	10	7	42	28	2	1	14	15
	50	0	0.05	7	6	11	8	1	3	6	4	1	0	4	0
	100	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
	100	0.75	0.43	100	100	100	100	100	100	100	100	99	99	100	100
	100	0.5	0.11	79	76	97	99	72	61	96	99	50	43	94	98
	100	0.25	0.03	32	30	71	69	20	15	58	55	4	5	37	36
	100	0	0.02	8	12	13	10	4	3	10	5	1	1	4	1
	200	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
	200	0.75	0.4	100	100	100	100	100	100	100	100	100	100	100	100
	200	0.5	0.08	99	98	100	100	97	97	100	100	88	82	100	100
	200	0.25	0.02	33	27	90	94	21	24	86	90	10	7	63	74
	200	0	0.01	10	13	13	6	6	7	7	1	1	1	1	0
$X_1 \sim N(0, 1)$	20	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
$X_2 \sim N(0, 1)$	20	0.75	0.6	81	75	99	97	70	64	99	93	50	37	94	72
	20	0.5	0.26	27	21	73	66	19	13	65	54	7	4	42	25
	20	0.25	0.14	13	10	38	21	4	7	25	12	1	2	11	2
	20	0	0.15	13	8	15	7	4	6	8	5	1	1	2	2
	50	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
	50	0.75	0.49	99	99	100	100	98	97	100	100	96	92	100	100
	50	0.5	0.12	59	51	99	95	40	40	98	92	26	22	92	75
	50	0.25	0.05	19	15	58	45	8	11	47	32	1	1	28	13
	50	0	0.04	5	8	13	10	1	5	6	5	1	1	1	3
	100	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
	100	0.75	0.44	100	100	100	100	100	100	100	100	100	100	100	100
	100	0.5	0.1	81	75	100	100	69	66	100	100	47	44	99	100
	100	0.25	0.03	23	18	80	71	12	11	75	56	3	4	57	33
	100	0	0.02	4	6	9	5	1	5	4	1	0	0	2	0
	200	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
	200	0.75	0.4	100	100	100	100	100	100	100	100	100	100	100	100
	200	0.5	0.09	98	97	100	100	96	92	100	100	88	83	100	100
	200	0.25	0.02	41	28	98	93	22	19	93	90	9	8	83	79
	200	0	0.01	11	11	10	11	8	5	5	5	0	0	0	2
$X_1 \sim Cauchy(0, 1)$	20	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
$X_2 \sim Cauchy(0, 1)$	20	0.75	0.57	75	69	85	97	68	60	77	92	50	45	60	80
	20	0.5	0.25	38	36	42	60	24	25	35	47	9	7	23	26
	20	0.25	0.15	17	20	18	30	10	8	15	14	3	4	12	6
	20	0	0.14	10	14	16	10	7	9	13	5	0	0	7	4
	50	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
	50	0.75	0.48	99	98	93	100	99	96	90	100	91	89	81	100
	50	0.5	0.14	57	51	55	94	46	38	43	90	23	21	31	85
	50	0.25	0.06	18	17	18	42	10	7	15	28	2	1	9	15
	50	0	0.05	7	6	8	8	1	3	7	4	1	0	5	0
	100	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
	100	0.75	0.43	100	100	96	100	100	100	95	100	99	99	88	100
	100	0.5	0.11	79	76	63	99	72	61	53	99	50	43	40	98
	100	0.25	0.03	32	30	21	69	20	15	16	55	4	5	11	36
	100	0	0.02	8	12	10	10	4	3	6	5	1	1	2	1
	200	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
	200	0.75	0.4	100	100	96	100	100	100	95	100	100	100	89	100
	200	0.5	0.08	99	98	67	100	97	97	62	100	88	82	44	100
	200	0.25	0.02	33	27	19	94	21	24	13	90	10	7	10	74
	200	0	0.01	10	13	5	6	6	7	5	1	1	1	2	0

表3：高斯 Copula 线性-线性检验功效研究：从由高斯 Copula 构造的线性-线性密度函数中抽取样本量分别为 20、50、100 和 200 的样本，并进行 100 次模拟；该密度函数由三个不同的边缘分布（指数分布、高斯分布和柯西分布）构成。比较采用瑞利（ART）和 Pycke（APT）圆周均匀性检验、Wilks（WT）和经验Copula（ECT）检验实施的所提检验的检验功效。高斯 Copula 的相关矩阵为等相关矩阵，其中五个不同的共同相关值分别为 0、0.25、0.5、0.75 和 0.99。共同相关值为零的情形对应于零假设独立模型。

边缘分布	SS	φ	$\hat{\lambda}_{c0}=2(1-\tilde{c}_0^2)$	$\alpha=10\%$				$\alpha=5\%$				$\alpha=1\%$			
				ART	APT	WT	ECT	ART	APT	WT	ECT	ART	APT	WT	ECT
$X_1\sim Exp(1)$	20	50	1	100	100	100	100	100	100	100	100	100	100	100	100
$X_2\sim Exp(2)$	20	15	0.95	100	100	100	100	100	99	100	100	99	97	98	100
	20	10	0.83	97	95	98	100	96	90	97	99	88	78	92	98
	20	5	0.4	55	48	72	89	46	41	62	77	31	19	38	59
	20	0	0.14	14	10	4	6	4	2	1	3	1	0	0	0
	50	50	1	100	100	100	100	100	100	100	100	100	100	100	100
	50	15	0.99	100	100	100	100	100	100	100	100	100	100	100	100
	50	10	0.9	100	100	100	100	100	100	100	100	100	100	100	100
	50	5	0.34	97	89	95	100	94	84	94	99	81	67	83	99
	50	0	0.05	14	12	3	13	6	7	2	7	1	2	0	1
	100	50	1	100	100	100	100	100	100	100	100	100	100	100	100
	100	15	1	100	100	100	100	100	100	100	100	100	100	100	100
	100	10	0.89	100	100	100	100	100	100	100	100	100	100	100	100
	100	5	0.31	100	100	100	100	100	100	100	100	100	97	100	100
	100	0	0.02	9	9	3	9	4	4	2	4	1	0	2	0
	200	50	1	100	100	100	100	100	100	100	100	100	100	100	100
	200	15	1	100	100	100	100	100	100	100	100	100	100	100	100
	200	10	0.88	100	100	100	100	100	100	100	100	100	100	100	100
	200	5	0.29	100	100	100	100	100	100	100	100	100	100	100	100
	200	0	0.01	8	10	5	8	4	5	2	6	1	1	0	2
$X_1\sim N(0,1)$	20	50	1	100	100	100	100	100	100	100	100	100	100	100	100
$X_2\sim N(0,1)$	20	15	0.95	100	100	100	100	100	99	100	100	99	97	100	100
	20	10	0.83	97	95	100	100	96	90	99	99	88	78	99	97
	20	5	0.4	55	48	94	90	46	41	88	72	31	19	69	54
	20	0	0.14	14	10	10	6	4	2	5	3	1	0	1	0
	50	50	1	100	100	100	100	100	100	100	100	100	100	100	100
	50	15	0.99	100	100	100	100	100	100	100	100	100	100	100	100
	50	10	0.9	100	100	100	100	100	100	100	100	100	100	100	100
	50	5	0.34	97	89	99	100	94	84	99	99	81	67	99	99
	50	0	0.05	14	12	13	13	6	7	6	7	1	2	0	1
	100	50	1	100	100	100	100	100	100	100	100	100	100	100	100
	100	15	1	100	100	100	100	100	100	100	100	100	100	100	100
	100	10	0.89	100	100	100	100	100	100	100	100	100	100	100	100
	100	5	0.31	100	100	100	100	100	100	100	100	100	97	100	100
	100	0	0.02	9	9	5	9	4	4	3	4	1	0	0	0
	200	50	1	100	100	100	100	100	100	100	100	100	100	100	100
	200	15	1	100	100	100	100	100	100	100	100	100	100	100	100
	200	10	0.88	100	100	100	100	100	100	100	100	100	100	100	100
	200	5	0.29	100	100	100	100	100	100	100	100	100	100	100	100
	200	0	0.01	8	10	9	8	4	5	6	6	1	1	1	2
$X_1\sim Cauchy(0,1)$	20	50	1	100	100	95	100	100	100	92	100	100	100	91	100
$X_2\sim Cauchy(0,1)$	20	15	0.95	100	100	81	100	100	99	73	100	99	97	62	100
	20	10	0.83	97	95	71	100	96	90	64	99	88	78	47	97
	20	5	0.4	55	48	42	90	46	41	33	72	31	19	19	54
	20	0	0.14	14	10	12	6	4	2	7	3	1	0	2	0
	50	50	1	100	100	98	100	100	100	95	100	100	100	93	100
	50	15	0.99	100	100	81	100	100	100	72	100	100	100	62	100
	50	10	0.9	100	100	66	100	100	100	63	100	100	100	52	100
	50	5	0.34	97	89	43	100	94	84	38	99	81	67	25	99
	50	0	0.05	14	12	12	13	6	7	9	7	1	2	4	1
	100	50	1	100	100	92	100	100	100	89	100	100	100	81	100
	100	15	1	100	100	70	100	100	100	65	100	100	100	49	100
	100	10	0.89	100	100	64	100	100	100	54	100	100	100	35	100
	100	5	0.31	100	100	32	100	100	100	25	100	100	97	20	100
	100	0	0.02	9	9	7	9	4	4	5	4	1	0	5	0
	200	50	1	100	100	86	100	100	100	82	100	100	100	77	100
	200	15	1	100	100	62	100	100	100	55	100	100	100	43	100
	200	10	0.88	100	100	49	100	100	100	42	100	100	100	35	100
	200	5	0.29	100	100	30	100	100	100	26	100	100	100	14	100
	200	0	0.01	8	10	5	9	4	5	4	7	1	1	2	2

表 4：Frank copula 线性-线性检验功效研究：通过 Frank copula 和三个不同的边缘分布（指数分布、高斯分布和柯西分布）构造线性-线性密度函数，并据此模拟样本量为 20、50、100 和 200 的样本各 100 次，比较采用瑞利（ART）和 Pycke（APT）圆周均匀性检验、Wilks（WT）检验以及经验Copula（ECT）检验所实施的所提检验的检验功效。Frank copula 的依赖参数取五个不同值 φ (0、5、10、15 和 50)。当 $\varphi=0$ 时的极限情形对应于零假设独立模型。

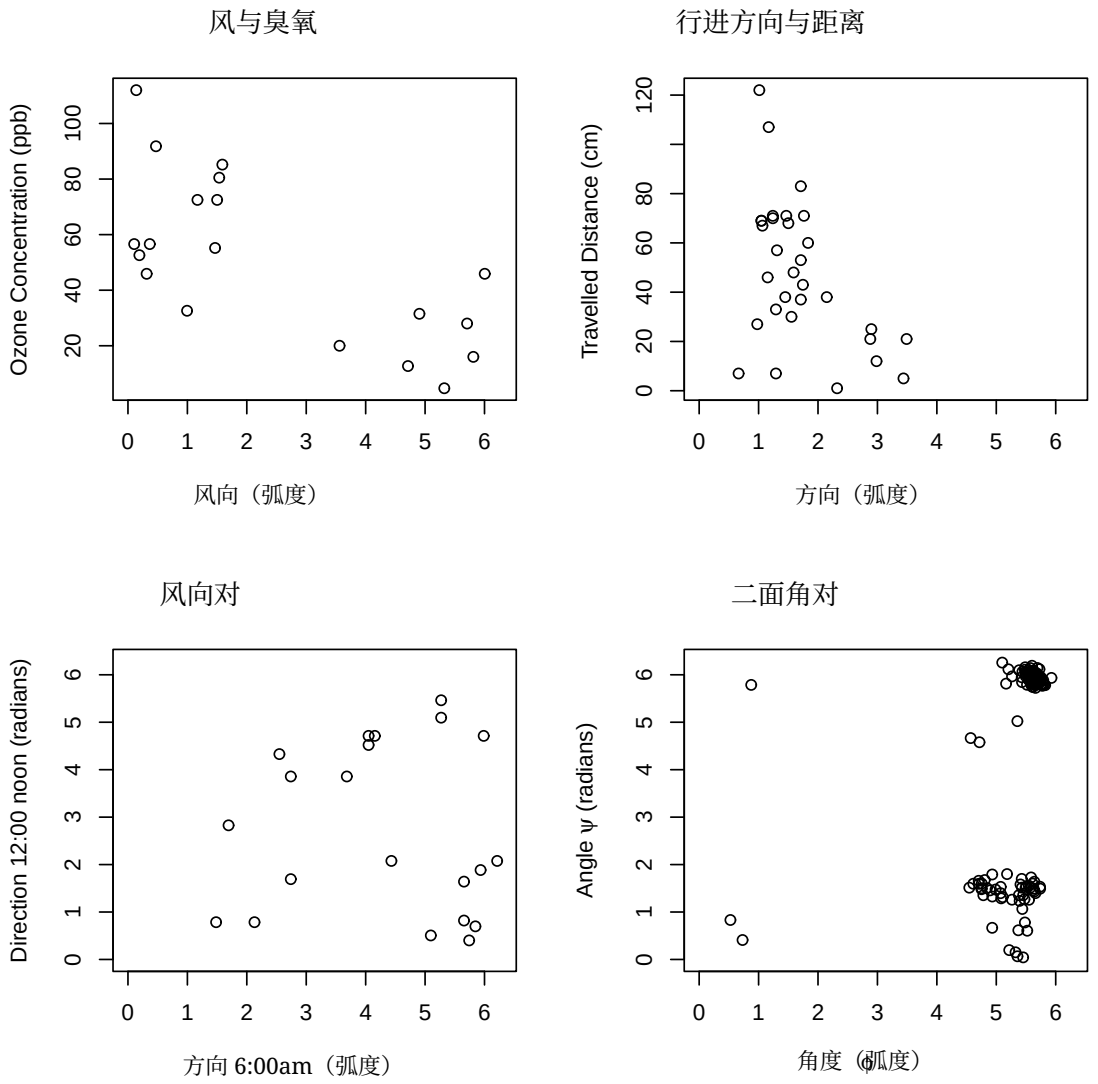

图 2：圆周-线性（上方图）和圆周-圆周（下方图）真实数据集的散点图。左上方的散点图展示了 Johnson 和 Wehrly（1977）数据集中风向（相对于正北方向）与臭氧浓度（ppb）的数据点。右上方的图对应 Fisher（1993）分析的小蓝滨螺数据集，内容为行进距离（cm）和方向。左下方的图对应 Johnson 和 Wehrly（1977）的数据集，展示了气象监测站在早上 6:00 和中午 12:00 记录的风向对（相对于正北方向）。最后，右下方展示了 Fernández-Durán（2007）最初分析的蛋白质丙氨酸-丙氨酸-丙氨酸片段中的二面角对。