DE GRUYTER Dependence Modeling 2023; 11: 20230103

Research Article

Juan José Fernández-Durán* and María Mercedes Gregorio-Domínguez

Test of bivariate independence based on angular probability integral transform with emphasis on circular-circular and circular-linear data

https://doi.org/10.1515/demo-2023-0103
received January 9, 2023; accepted September 18, 2023

Abstract: The probability integral transform of a continuous random variable X with distribution function F_X is a uniformly distributed random variable $U = F_X(X)$. We define the angular probability integral transform (APIT) as $\theta_U = 2\pi U = 2\pi F_X(X)$, which corresponds to a uniformly distributed angle on the unit circle. For circular (angular) random variables, the sum modulus 2π of absolutely continuous independent circular uniform random variables is a circular uniform random variable, that is, the circular uniform distribution is closed under summation modulus 2π, and it is a stable continuous distribution on the unit circle. If we consider the sum (difference) of the APITs of two random variables, X_1 and X_2, and test for the circular uniformity of their sum (difference) modulus 2π, this is equivalent to test of independence of the original variables. In this study, we used a flexible family of nonnegative trigonometric sums (NNTS) circular distributions, which include the uniform circular distribution as a member of the family, to evaluate the power of the proposed independence test by generating samples from NNTS alternative distributions that could be at a closer proximity with respect to the circular uniform null distribution.

Keywords: circular-circular dependence, circular-linear dependence, circular uniformity tests, copula, dependence measures

MSC 2020: 62H11, 62H05

1 Introduction

Testing for independence is one of the most important tasks in statistics, for example, when constructing the joint distribution of a set of random variables or considering the conditional dependence of one variable in terms of other variables as in regression models. According to Herwatz and Maxand [23], one can consider the following tests of independence: bivariate (pairwise), mutual, and groupwise independence tests. Hereinafter, we only consider absolutely continuous random variables. A random variable with a density function with support on an interval of the real line is a linear random variable, and one with support on the unit circle is a circular random variable. The distribution function of a circular random is a periodic function, and then, it has an arbitrary starting direction. In our case, we considered that all the circular random variables are defined on the interval $(0, 2\pi]$ with common starting direction equal to zero and are angles in radians measured all counterclockwise or clockwise. In this sense, if Θ is a circular random variable with starting direction zero and measured in radians counterclockwise,

*** Corresponding author: Juan José Fernández-Durán**, Business Administration, ITAM, CDMX, Mexico, e-mail: jfdez@itam.mx
María Mercedes Gregorio-Domínguez: Actuarial Science, ITAM, CDMX, Mexico

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

and

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

for k an integer number. The density function for an absolutely continuous circular random variable satisfies being a periodic function such that $f_\Theta(\theta) \ge 0$, $f_\Theta(\theta + 2k\pi) = f_\Theta(\theta)$, and $\int_0^{2\pi} f_\Theta(\theta)\mathrm{d}\theta = 1$. For two random variables, X_1 and X_2, bivariate (pairwise) independence tests have null hypothesis, $H_0 : F_{X_1,X_2}(x_1, x_2) = F_{X_1}(x_1)F_{X_2}(x_2)$, where $F_{X_1,X_2}(x_1, x_2) = P\{X_1 \le x_1, X_2 \le x_2\}$ is the bivariate joint distribution function and $F_{X_1}(x_1) = P\{X_1 \le x_1\}$ and $F_{X_2}(x_2) = P\{X_2 \le x_2\}$ are the corresponding marginal distribution functions. A mutual independence test for a set of random variables, $X_1, X_2, \ldots, X_d$ considers the null hypothesis $H_0 : F_{X_1,X_2,\ldots,X_d}(x_1, x_2, \ldots, x_d) = \prod_{k=1}^{d} F_{X_k}(x_k)$, where $F_{X_1,X_2,\ldots,X_d}(x_1, x_2, \ldots, x_d)$ is the joint distribution and $F_{X_k}(x_k)$ for $k = 1, 2, \ldots, d$ are the marginal univariate distribution functions. Finally, groupwise independence tests consider the independence between disjoint subsets of random variables. If the parametric functional form of $F_{X_1,X_2,\ldots,X_d}$ and F_{X_k} for $k = 1, 2, \ldots, d$ are specified, a likelihood ratio test of independence from a sample of random vectors $\underline{X}_i = (X_{i1}, X_{i2}, \ldots, X_{id})^\top$ of size n ($i = 1, 2, \ldots, n$), can be constructed. The most commonly used test for independence is the chi-squared test of independence for contingency tables, which is not adequate when dealing with absolutely continuous random variables. Alternatively to the likelihood ratio independence test, other tests for independence were developed by considering nonparametric (distribution free) methods, rank tests [24,31], and measures of dependence (association) derived from the empirical copula process [6,19–21,43,44]. The empirical copula, C_n, for a vector of d absolutely continuous linear random variables, $\underline{X}^\top = (X_1, X_2, \ldots, X_d)^\top$, and a sample of size n is defined as follows:

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

where $I()$ is an indicator function, which is equal to one if the condition in its argument is satisfied and zero otherwise, and $\hat{F}_1, \hat{F}_2, \ldots, \hat{F}_d$ are the empirical distribution functions of the random variables $X_1, X_2, \ldots, X_d$. A test of independence based on the distance between the empirical copula and independence copula for absolutely continuous linear random variables is implemented in the *R* package *copula* [25,32].

Among the nonparametric tests of independence, there is a family of tests based on some functional of the empirical independence process [5], which is defined as the distance between the empirical joint distribution function and product of the empirical univariate distribution functions. Historically, the most used functionals have been the Cramér-von Mises and Kolmogorov-Smirnov functionals [3,6,7]. For example, Hoeffding [24] considered the Cramér-von Mises functional to generate a rank test of independence between two random variables. Modern rank tests of independence have been developed by Kallenberg and Ledwina [29]. Kernel-based methods have also been used to estimate the empirical independence process, as in the study by Pfister et al. [40]. Mardia and Kent [34] used the general Rao score test to generate independence tests. Csörgö [5] developed independence tests based on the multivariate empirical characteristic function, and Einmahl and McKeague [8] developed the tests based on the empirical likelihood. The measures of dependence derived from entropy were defined by Joe [27] and from mutual information by Berrett and Samworth [2].

When constructing tests of independence, one can take advantage of the characteristics of the multivariate joint distribution. For example, for the multivariate normal distribution, one can test for independence by testing for an identity correlation matrix. Some of these pairwise tests are the Pearson's [37] product moment correlation coefficient test, Kendall's [30] rank correlation coefficient test, and Spearman's [46] rank correlation coefficient test. Of course, for a pair of Gaussian random variables, rejecting null correlation implies rejecting pairwise independence, but applying pairwise independence (correlation) tests is not adequate to test for mutual independence for a set with more than two Gaussian random variables. The Wilks test [49] is an optimal test of independence for multivariate Gaussian populations and for the case of a bivariate groupwise independence test for the vectors $\underline{X}_{D_1}$ and $\underline{X}_{D_2}$ with $\underline{X} = \underline{X}^\top_{D_1 \cup D_2} = (\underline{X}_{D_1}, \underline{X}_{D_2})^\top$ considers the following test statistic for a sample of size n,

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

where $\hat{\Sigma}_{D_1 \cup D_2} = \sum_{j=1}^{n}(\underline{x}_j - \underline{\bar{x}})(\underline{x}_j - \underline{\bar{x}})^\top$ is the estimated covariance matrix of the complete vector of observations $\underline{x} = \underline{x}_{D_1 \cup D_2}$, which is partitioned into $\hat{\Sigma}_{D_1}$ and $\hat{\Sigma}_{D_2}$ with $\hat{\Sigma}_{D_1} = \sum_{j=1}^{n}(\underline{x}_{D_1,j} - \underline{\bar{x}}_{D_1})(\underline{x}_{D_1,j} - \underline{\bar{x}}_{D_1})^\top$ and $\hat{\Sigma}_{D_2} = \sum_{j=1}^{n}(\underline{x}_{D_2,j} - \underline{\bar{x}}_{D_2})(\underline{x}_{D_2,j} - \underline{\bar{x}}_{D_2})^\top$. The statistic W then measures the extent of the distance between the determinant of $\hat{\Sigma}_{D_1 \cup D_2}$ and the product of the determinants of $\hat{\Sigma}_{D_2}$ and $\hat{\Sigma}_{D_2}$. The equality relationship is satisfied in the multivariate Gaussian population under the null hypothesis of independence between $\underline{X}_{D_1}$ and $\underline{X}_{D_2}$. Asymptotically and under regularity conditions, $-n\ln(W)$ follows a chi-squared distribution with $\sharp D_1 \sharp D_2$ degrees of freedom, where $\sharp D_1$ and $\sharp D_2$ are the cardinalities of sets D_1 and D_2, respectively.

For the circular-circular (angular-angular) and circular-linear (angular-linear) cases, in which the objective is to test for bivariate independence between two circular random variables and one circular and one linear random variable, respectively, independence tests were developed by considering the specification of measures of dependence and studying their (asymptotic) distributions. By applying Kendall's tau and Spearman's rho general measures of dependence based on the concept of concordance, or the construction of distribution-free correlation coefficients based on ranks to a pair of circular random variables or a circular and a linear random variables, tests of independence were developed by Fisher and Lee [14–16] and reviewed by Fisher [17] and Mardia and Jupp [35].

The objective of this study is to develop a test of bivariate (pairwise) independence for two random variables by considering the angular probability transform of each variables, which correspond to circular uniform distributions on $(0, 2\pi]$, and an additional result of the theory of circular statistics ([17,26,35,38,47]). The test was evaluated using flexible nonnegative trigonometric sums (NNTS) distributions [10,11]. Although the proposed test of independence is a general one, it is especially suitable to test for the independence in the circular-circular and circular-linear cases. Thus, a measure of dependence was developed.

This article is divided into six sections, including the introduction. In Section 2, Johnson and Wehrly's [28] model is presented as a motivation for performing the test of bivariate independence for two random variables, and here, the theory of NNTS circular distributions is included. Section 3 presents the proposed bivariate independence test, a measure of dependence, and its application to simulated data to study the power of the test. The Section 4 includes a simulation study to evaluate the power of the proposed test in the linear-linear, circular-linear, and circular-circular cases. The Section 5 describes the application of the proposed independence test to real datasets. Finally, conclusions are presented in Section 6.

2 Bivariate Johnson and Wehrly model and NNTS family of circular densities

Sklar [45] theorem specifies that the joint cumulative distribution function of two continuous random variables, X_1 and X_2, with corresponding cumulative distribution functions $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$, can be expressed as follows:

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

where $C(u, v)$ is the copula (distribution) function that corresponds to a bivariate joint distribution of two identically distributed uniform random variables on the interval $[0, 1]$, $U = F_{X_1}(X_1)$ and $V = F_{X_2}(X_2)$. A linear-linear copula function must be an increasing function satisfying that $C(u, 1) = u$, $C(1, v) = v$, $C(0, v) = C(u, 0) = 0$ [36]. By differentiation of equation (3), one obtains the joint density function of X_1 and X_2, $f_{X_1,X_2}(x_1, x_2)$, as follows:

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

where $f_{X_1}(x_1)$ and $f_{X_2}(x_2)$ are the marginal density functions of X_1 and X_2, respectively. The function $c(u, v)$ is the copula density function defined as $c(u, v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}$. Johnson and Wehrly [28] and Wehrly and Johnson [48] proposed a large family of joint density functions for circular-circular (Θ_1 and Θ_2) and circular-linear (Θ and T) random vectors in the following way:

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

and

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

Fernández-Durán [9] identified the structure of the Johnson and Wehrly's model in terms of the theory of copula functions through Sklar's [45] theorem [36] satisfying

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

The function g must be the density function of an angular (circular) random variable on the interval $(0, 2\pi]$. For the case of circular-circular and circular-linear bivariate copulas, the function c has also to be periodic in their circular arguments to satisfy the periodicity of the density function of a circular random variable, and this is the reason that in the Johnson and Wehrly model the joining function g is the density of a circular random variable.

Johnson and Wehrly derived bivariate circular–circular and circular–linear models by considering conditional arguments. When function g corresponds to a uniform circular density on the circle, $g(\theta) = \frac{1}{2\pi}$, the joint density of the Johnson and Wehrly model corresponds to the independence case in which the joint density of the circular-circular (circular-linear) model is the product of the marginal univariate densities.

This property of the Johnson and Wehrly model motivated our independence test by approximating the circular density function g with a density function from NNTS family of densities [10,11], which includes uniform circular density as a particular case. In addition, by using the NNTS family of circular densities, it is possible to generate joining densities g that are in closer proximity to the circular uniform density as desired; this is explained below. Pewsey and Kato [39] developed a goodness-of-fit test for the circular-circular model of Wehrly and Johnson [48] in equation (5). Their test considers the independence between each circular random variable, Θ_1 and Θ_2, and the argument of the joining function g, $\Omega = 2\pi\{F_{\Theta_2}(\theta_2) \pm F_{\Theta_1}(\theta_1)\}(\text{mod } 2\pi)$. In a similar way to this article, their independence test is implemented by testing for the toroidal uniformity of the two bivariate random vectors, $(2\pi F_{\Theta_1}(\theta_1), 2\pi F_{\Omega}(\omega))^\top$ and $(2\pi F_{\Theta_2}(\theta_2), 2\pi F_{\Omega}(\omega))^\top$.

The circular density function based on NNTS for a circular (angular) random variable $\Theta \in (0, 2\pi]$ [10] is defined as follows:

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

where $i = \sqrt{-1}$, c_k are complex numbers $c_k = c_{rk} + ic_{ck}$ for $k = 0, \ldots, M$ and $\bar{c}_k = c_{rk} - ic_{ck}$ is the conjugate of c_k. To have a valid density function, $f_{\Theta}(\theta; M, \underline{c})$, which integrates to one,

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

where $||c_k||^2 = c_{rk}^2 + c_{ck}^2$. The parameter space of the vector of parameters, $\underline{c} = (c_0, c_1, \ldots, c_M)^\top$, is a subset of the surface of a hypersphere since $\underline{c}$ and $-\underline{c}$ gives the same NNTS density and the conjugate of $\underline{c}$ written in reverse order also gives the same NNTS density as $\underline{c}$. Then, for identifiability of the parameter vector $\underline{c}$ and given equation (9), the following constraints are imposed: $c_{c0} = 0$ and $c_{r0} \geq 0$, i.e., c_0 is a nonnegative real number and $c_0^2 \geq ||c_M||^2$. There is a total of $2M$ free parameters c, and M is an additional parameter that determines the total number of terms in the sum defining the density and it is related to the maximum number of modes that the density can have. Note that an NNTS model with $M = M_1$ is nested on NNTS models with $M = M_2$ such that $M_2 > M_1$. The circular uniform density on $(0, 2\pi]$ corresponds to an NNTS density with $M = 0$, $f_{\Theta}(\theta; M = 0, \underline{c}) = \frac{1}{2\pi}$, and it is nested in all NNTS models with $M > 0$ or, for an NNTS model with $M > 0$, and as c_0 approaches the value of 1, the NNTS density converges to the circular uniform density or, equivalently, for an NNTS density with $\underline{c} = (1, 0, 0, \ldots, 0)^\top$, that is, with $c_0 = 1$ and the other elements of $\underline{c}$ equal to zero corresponds to the circular uniform density. It is this property that will be used to evaluate the power for the proposed independence test by generating samples from NNTS alternative densities with values of c_0 as close to one as desired. Fernández-Durán and Gregorio-Domínguez [12] developed an efficient algorithm based on optimization on manifolds to obtain the maximum likelihood estimates of the c parameters. This algorithm is included with other routines for the analysis of circular data based on NNTS models in the free *R* [42] package *CircNNTSR* [13].

3 Proposed test for bivariate independence

For absolutely continuous independent and identically distributed (i.i.d.) circular uniform random variables, $U_1, U_2, \ldots, U_d \sim U(0, 2\pi]$, consider $\sum_{k=1}^{d} U_k (\text{mod } 2\pi) \sim U(0, 2\pi]$, that is, the sum modulus 2π of i.i.d. circular uniform random variables is also circular and uniformly distributed. This is a consequence of the fact that the characteristic function of a circular uniform random variable, $\psi(t) = E(e^{it\Theta})$, is equal to one when $t = 0$ and equal to zero elsewhere ($t \neq 0$). Then, if the sum modulus 2π of circular uniform random variables is not circular uniformly distributed, this implies that the circular uniform random variables are not independent [35, p. 35]. The proposed test of independence is based on this result by considering the angular probability integral transform (APIT) of arbitrary (linear or circular) absolutely continuous random variables. Let $X_1, X_2, \ldots, X_d$ be d arbitrary absolutely continuous random variables. The angular probability transform of X_k is defined as the angular (circular) random variable $\text{APIT}(X_k) = 2\pi F_{X_k}(X_k)$, which is uniformly distributed on the unit circle. By considering the null hypothesis of mutual joint independence, $\text{APIT}(X_1)$, $\text{APIT}(X_2)$, $\ldots$, $\text{APIT}(X_d)$ are *i.i.d.* $U(0, 2\pi]$, then $\sum_{k=1}^{d} \pm \text{APIT}(X_k)(\text{mod } 2\pi) \sim U(0, 2\pi]$ is also circular and uniformly distributed. The proposed test for bivariate independence is based on testing for the circular uniformity of $(\text{APIT}(X_1) + \text{APIT}(X_2))(\text{mod } 2\pi)$ $((\text{APIT}(X_1) - \text{APIT}(X_2))(\text{mod } 2\pi))$ for absolutely continuous (circular or linear) random variables X_1 and X_2. Testing for bivariate independence is equivalent to testing for uniformity of the sum (difference) modulus 2π of the APITs. To test for circular uniformity of the sum (difference) modulus 2π of the angular integral transforms, we considered the tests of Rayleigh and Pycke [41]. The Rayleigh test considers an alternative unimodal circular density and has a test statistic for a sample $\theta_1, \theta_2, \ldots, \theta_n$, which is defined as follows [35]:

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

where $\bar{R}$ is the sample mean resultant length. The statistic T_{RT} asymptotically has a chi-squared distribution with two degrees of freedom. The p-values of the Rayleigh test used in this article correspond to the approximation given by Fisher [17, p. 70] that includes a correction for small sample sizes. The Pycke test considers an alternative multimodal density, and its test statistic for a sample $\theta_1, \theta_2, \ldots, \theta_n$ is defined as follows:

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

The critical values of the Pycke test are obtained via simulation.

The steps of the proposed independence test for two absolutely continuous random variables, X_1 and X_2, are as follows. First, the empirical distributions $\hat{F}_{X_1}$ and $\hat{F}_{X_2}$ were calculated from the observed values of each random variable. Second, the APITs were calculated by multiplying the pseudo-observations by 2π ($2\pi\hat{F}_{X_1}$ and $2\pi\hat{F}_{X_2}$). Third, the sum (difference) modulus 2π of the two APITs was calculated. Finally, the Rayleigh or Pycke test was applied to the vector of the observed values of the sum (difference) modulus 2π of the APITs. In the case of a positive association between the random variables, the proposed independence test must be applied to the difference modulus 2π of the APITs, and in the case of a negative association, it must be applied to the sum modulus 2π of the APITs. The case in which there is no prior indication regarding whether the association

Table 1: NNTS angular joining density circular-linear power study

Marginals	SS	c_0	$\hat{\lambda}_{c0} = 2(1-\hat{c}_0^2)$	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
				RT	PT	WT	ECT	RT	PT	WT	ECT	RT	PT	WT	ECT
$\Theta \sim \text{NNTS}(M=3)$	20	0.7	0.61	91	89	29	43	89	88	22	32	68	64	8	3
$X \sim \text{Exp}(1)$	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
$\Theta \sim \text{NNTS}(M=3)$	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
$\Theta \sim \text{NNTS}(M=3)$	20	0.7	0.52	81	75	17	28	73	66	13	14	46	46	6	6
$X \sim \text{Cauchy}(0,1)$	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

(Continued)

Table 1: *Continued*

Marginals	SS	c_0	$\hat{\lambda}_{c0} = 2(1 - \hat{c}_0^2)$	α = 10%				α = 5%				α = 1%			
				RT	PT	WT	ECT	RT	PT	WT	ECT	RT	PT	WT	ECT
	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

The powers of the proposed test implemented using the Rayleigh (ART) and Pycke (APT) circular uniformity tests, the Wilks test (WT) and the empirical copula test (ECT) are compared when simulating 100 times samples of sizes 20, 50, 100, and 200 from a Johnson and Wehrly circular-linear density function constructed from an NNTS angular joining density with $M = 3$, an NNTS marginal density function with $M = 3$ (Figure 1) and, three different linear marginals (exponential, Gaussian, and Cauchy). The NNTS angular joining density is defined with five different values of the parameter c_0 (0.7, 0.8, 0.9, 0.99, and 0.9999). The case with $c_0 = 1$ corresponds to the null independence model.

is positive or negative, one can calculate the two test statistics, one for the sum modulus 2π and other for the difference modulus 2π of the APITs, and modify the minimum of the two p-values in accordance to some multiple testing correction such as Bonferroni procedure [4,22]. In practice, the user can plot the histogram of the sum (difference) modulus 2π of APITs and decide on the use of the Rayleigh or Pycke test in terms of the observed number of modes. In case of doubt, the Pycke test should be preferred. Other omnibus tests for circular uniformity can be considered as those of Ajne et al. [35].

Derived from the fitting of an NNTS model with $M = 1$ to the sum (difference) modulus 2π of APITs, the following measure of dependence can be defined as follows:

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

where the correction term $\frac{M+1}{M} = 2$ comes from the fact that the NNTS density, in the case $M = 1$, with the highest concentration around zero has a parameter vector $\underline{c}$ in which the squared norm of each of its components is equal to $\frac{1}{M+1} = \frac{1}{2}$. This implies that $\frac{1}{2} \le c_0^2 \le 1$ and $\hat{\lambda}_{c_0}$, for $M = 1$, takes values in the interval $[0, 1]$ with values close to zero, implying low dependence (independence) and values closer to one, further implying high dependence between the considered random variables. For independent random variables, the theoretical value of $\hat{\lambda}_{c_0}$ is equal to zero. The measure of dependence $\hat{\lambda}_{c_0}$ is particularly useful in the circular-circular and circular-linear cases.

4 Simulation study

In this section, we present a simulation study to compare the power of the proposed test with the Wilks test and a test of independence based on the empirical copula. We simulated the data from different multivariate distributions using known parameters that define the dependence structure and known marginal densities. We considered the sample sizes of 20, 50, 100, and 200. For a given significance level α (10, 5, and 1%), the powers of the tests were obtained by generating 100 samples of the specified sample size from the bivariate density, by calculating the values of the test statistics for each of the 100 samples, and determining the number of times that the test statistics considered a value that rejected the null hypothesis of independence at the given value of α. Thus, the reported powers of the tests considered values in the range of 0–100 and can be interpreted in terms of percentages. The Rayleigh test of circular uniformity was performed using the *circular R* package [1]. The Pycke test of circular uniformity was performed using the *CircMLE R* package [18,33]. The *R*

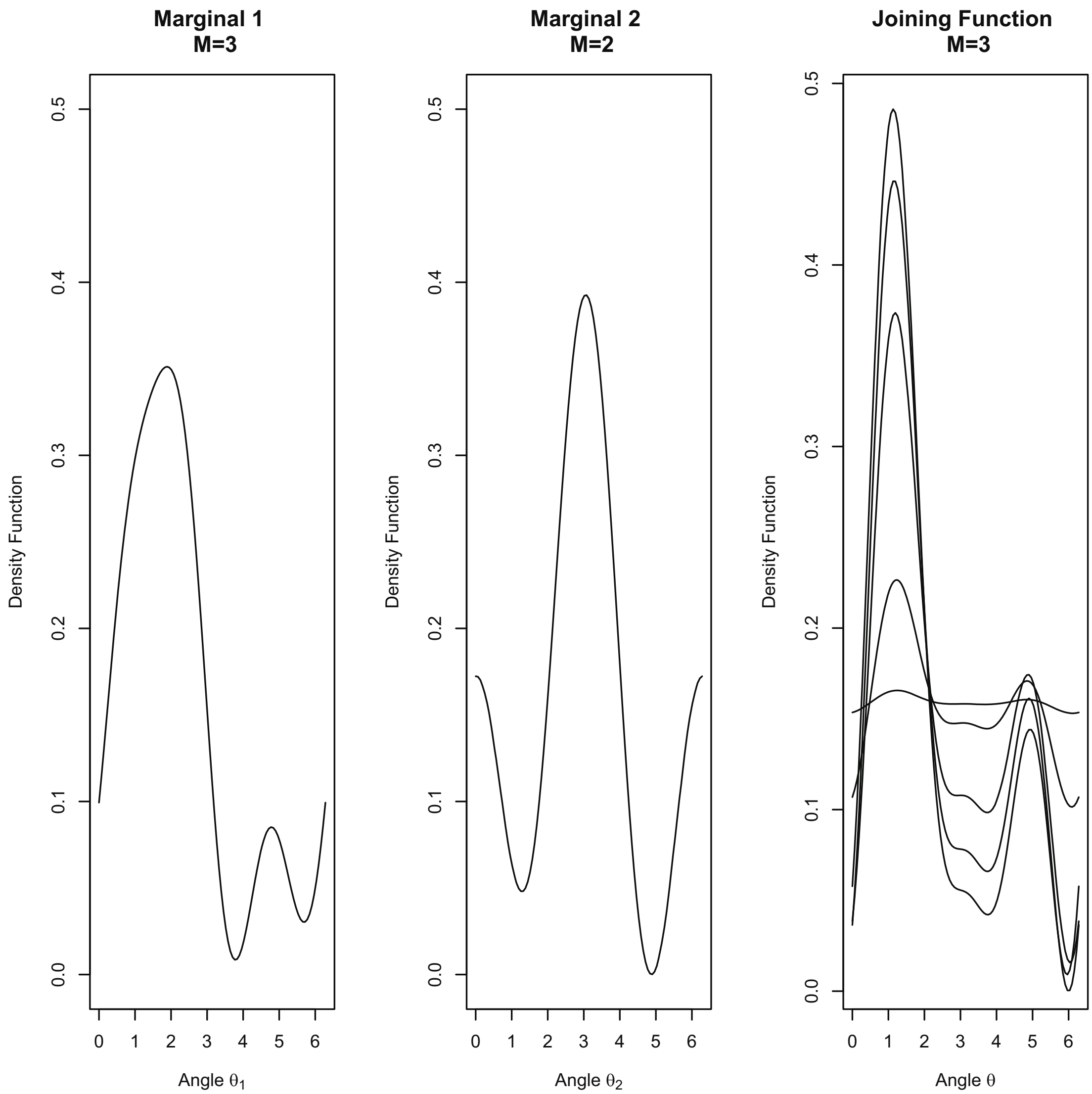

Figure 1: Circular-circular copula model: The first two plots show the marginal NNTS circular densities (M_1 = 3 and M_2 = 2) and the last third plot show the angular (circular) joining density for different values of the parameter c_0 (0.7, 0.8, 0.9, 0.99, and 0.9999). The case c_0 = 1 corresponds to the circular uniform density (null independence model).

package *copula* was used to calculate the empirical copula test, and the measure of dependence $\hat{\lambda}_{c_0}$ was obtained by fitting an NNTS model with $M = 1$ using the *R* package *CircNNTSR* [13].

4.1 Circular-linear models

Table 1 includes the powers of the proposed ART and APT, and the WT and ECT when simulating samples from the circular-linear model of Johnson and Wehrly with the circular marginal density being an NNTS density with $M = 3$, which is plotted in the first plot of Figure 1; three different linear marginals (exponential, Gaussian, and Cauchy); and a joining circular density that corresponds to an NNTS density with $M = 3$ with five different values of the parameter c_0 (0.7, 0.8, 0.9, 0.99, and 0.9999) to account for different degrees of association between the circular and linear random variables, as depicted in the last plot of Figure 1, which includes the plots of the angular joining functions for the five different values of parameter c_0. The case $c_0 = 0.9999$ corresponds to an almost circular uniform density and to the null hypothesis of independence

Table 2: NNTS angular joining density circular-circular power study

Marginals	SS	c_0	$\hat{\lambda}_{c0} = 2(1 - \hat{c}_0^2)$	α = 10%				α = 5%				α = 1%			
				RT	PT	WT	ECT	RT	PT	WT	ECT	RT	PT	WT	ECT
Θ_1 ~ NNTS(M_1 = 3)	20	0.7	0.63	87	80	16	38	78	73	11	29	59	61	1	11
Θ_2 ~ 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

The powers of the proposed test implemented using the ART and APT circular uniformity tests, the WT, and the empirical copula test (ECT) are compared when simulating 100 times samples of sizes 20, 50, 100, and 200 from a Johnson and Wehrly circular-circular density function constructed from an NNTS angular joining density with $M = 3$ and NNTS marginal density functions with $M_1 = 3$ and $M_2 = 2$. The NNTS angular joining density is defined with five different values of the dependence parameter c_0 (0.7, 0.8, 0.9, 0.99, and 0.9999). The case with $c_0 = 1$ corresponds to the null independence model. The plots of the NNTS angular joining density and NNTS circular marginal densities are shown in Figure 1.

(refer the last plot of Figure 1). In general, the proposed ART and APT demonstrated significantly larger powers when compared to the ECT, which was only similar for large sample sizes and low values of c_0 (0.7, 0.8, and 0.9), further representing highly dependent circular and linear random variables. The power of the WT was considerably lower when compared to that of the ART, APT, and ECT.

4.2 Circular-circular models

For Johnson and Wehrly's circular-circular model, we used the same angular joining density and one of the marginal circular densities as that used in the circular-linear model. Figure 1 depicts the plots of the marginal circular densities that correspond to NNTS densities with $M = 3$ and $M = 2$, and angular joining densities that correspond to NNTS densities with $M = 3$ for five different values of the parameter c_0 (0.7, 0.8, 0.9, 0.99, and 0.9999) for the circular-circular model of Johnson and Wehrly. Similar to the results in the circular-linear model, the ART and APT demonstrated larger power values when compared to the WT and ECT tests. Given the multimodality of all the circular densities involved, of the two proposed independence tests, APT exhibited a larger power when compared to ART, which was similar only for the largest sample size of 200. Moreover, for highly dependent circular random variables (c_0 = 0.7, 0.8, or 0.9) and large sample sizes of 100 or 200, the power of the ECT test was similar to those of the ART and APT. The average values of the dependence measure $\hat{\lambda}_{c_0}$ listed in the fourth column of Tables 1 and 2 assumed similar values when c_0 = 0.7, 0.8, and 0.9, but these

Table 3: Gaussian copula linear-linear power study

Marginals	SS	ρ	$\hat{\lambda}_{c0} = 2(1-\hat{c}_0^2)$	α = 10%				α = 5%				α = 1%			
				ART	APT	WT	ECT	ART	APT	WT	ECT	ART	APT	WT	ECT
X_1 ~ Exp(1)	20	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
X_2 ~ 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 ~ $N(0,1)$	20	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
X_2 ~ $N(0,1)$	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 ~ Cauchy(0, 1)	20	0.99	1	100	100	100	100	100	100	100	100	100	100	100	100
X_2 ~ 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

(Continued)

Table 3: *Continued*

Marginals	SS	ρ	$\hat{\lambda}_{c0} = 2(1 - \hat{c}_0^2)$	α = 10%				α = 5%				α = 1%			
				ART	APT	WT	ECT	ART	APT	WT	ECT	ART	APT	WT	ECT
	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

The powers of the proposed test implemented using the ART and APT circular uniformity tests, the WT and the ECT tests are compared when simulating 100 times samples of sizes 20, 50, 100, and 200 from a linear-linear density function constructed from a Gaussian copula and three different marginals (exponential, Gaussian, and Cauchy). The Gaussian copula is defined with an equicorrelated correlation matrix with five different common correlation values of 0, 0.25, 0.5, 0.75, and 0.99. The case with common correlation equal to zero corresponds to the null independence model.

values were smaller when c_0 = 0.99 and 0.9999, further reflecting the fact that values of c_0 near one are associated with random variables with a very weak association.

4.3 Linear-linear models

Tables 3 and 4 list the powers of different tests while simulating samples from a bivariate distribution in which both variables are linear. In the first case, a Gaussian copula was used, and in the second case, a Frank copula was used.

4.3.1 Bivariate Gaussian copula

The bivariate Gaussian (normal) copula correspond to a multivariate distribution, which is defined as follows:

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

where Φ_Γ is the multivariate normal distribution with a zero mean vector and correlation matrix Γ and $\Phi()$ is the univariate standard normal distribution function. By using an identity matrix as a correlation matrix, $\Gamma = I$, the independence copula, $C(u_1, u_2) = u_1 u_2$, is obtained.

Table 3 compares the powers of the proposed independence test when using a ART and APT circular uniformity tests with respect to those of the WT and ECT independence tests when using simulated samples from a bivariate linear-linear distribution with a Gaussian copula and three different cases of marginal distributions following Herwatz and Maxand [23]: exponential, Gaussian, and Cauchy. For the Gaussian copula, we considered five different values of the correlation coefficient ρ (0, 0.25, 0.5, 0.75, and 0.99). When the correlation coefficient is equal to zero, it corresponds to the null independence hypothesis, and the reported powers correspond to the sizes of the tests expected to be similar to the significance levels α (10, 5, and 1%). In general terms, both WT and ECT have higher powers than the proposed ART and APT. However, for large sample sizes and high values of the correlation coefficient (0.75 and 0.99), the ART and APT have powers similar to those of the WT and ECT. As expected, for the exponential and Cauchy marginals, the power of the WT reduced when compared to the Gaussian marginal case for which it was designed; that is, the power of the WT deteriorates for marginals that are not Gaussian. In the case of Gaussian marginals, the WT for large sample sizes has high power, and in some cases, its power is larger than the ECT power. In terms of the sizes of all the tests, it appears that all tests have approximately the correct sizes, given that a total of 100 samples were

Table 4: Frank copula linear-linear power study

Marginals	SS	φ	$\hat{\lambda}_{c0} = 2(1 - \hat{c}_0^2)$	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
				ART	**APT**	**WT**	**ECT**	**ART**	**APT**	**WT**	**ECT**	**ART**	**APT**	**WT**	**ECT**
$X_1 \sim \text{Exp}(1)$	20	50	1	100	100	100	100	100	100	100	100	100	100	100	100
$X_2 \sim \text{Exp}(2)$	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
$X_1 \sim \text{Cauchy}(0,1)$	20	50	1	100	100	95	100	100	100	92	100	100	100	91	100
$X_2 \sim \text{Cauchy}(0,1)$	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

(Continued)

Table 4: *Continued*

Marginals	SS	φ	$\hat{\lambda}_{c0} = 2(1 - \hat{c}_0^2)$	α = 10%				α = 5%				α = 1%			
				ART	APT	WT	ECT	ART	APT	WT	ECT	ART	APT	WT	ECT
	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

The powers of the proposed test implemented using the ART and APT circular uniformity tests, the WT and the ECT are compared when simulating 100 times samples of sizes 20, 50, 100, and 200 from a linear-linear density function constructed from a Frank copula and three different marginals (exponential, Gaussian, and Cauchy). The Frank copula is defined with five different values of the dependence parameter φ (0, 5, 10, 15, and 50). The limit case with $\varphi = 0$ corresponds to the null independence model.

used. The fourth column of Table 3 includes the averages of the values of the dependence measure $\hat{\lambda}_{c_0}$ which, as expected, increase as the value of the correlation coefficient ρ increases.

4.3.2 Bivariate Frank copula

The Frank bivariate copula is defined as follows:

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

where $u, v \in (0, 1]$. Parameter φ assumes values in the interval $(0, \infty)$ and in the limit $\varphi = 0$, the Frank copula corresponds to the independence copula. In Table 4, the Gaussian copula of the previous study is replaced by the Frank copula using five different values of the dependence parameter φ (0, 5, 10, 15, and 50). For the case of independence obtained in the limit when $\varphi = 0$, the powers are also obtained when the limit is considered as $\varphi = 0$. In general, in Table 4, we obtained the same conclusions as those obtained in Table 3, further indicating that the characteristics of the proposed tests: ART and APT are more suitable for the circular-circular and circular-linear cases than to the linear-linear case.

5 Application to real circular-circular and circular-linear data

5.1 Test of bivariate independence

5.1.1 Circular-linear real examples

Figure 2 depicts the scatterplots of the considered real examples. We applied the proposed independence test to the circular-linear data on wind direction (circular variable) and ozone concentration (linear variable) originally analyzed by Johnson and Wehrly [28], and later included them as dataset B.18 in Fisher [17]. A total of 19 measurements were taken at a weather station in Milwaukee at 6 o'clock in the morning every fourth day starting on April 18 and ending on June 29, 1975. The scatterplot of this data is included in the top left plot of Figure 2, which presents the values of the wind direction and ozone concentration, further indicating a

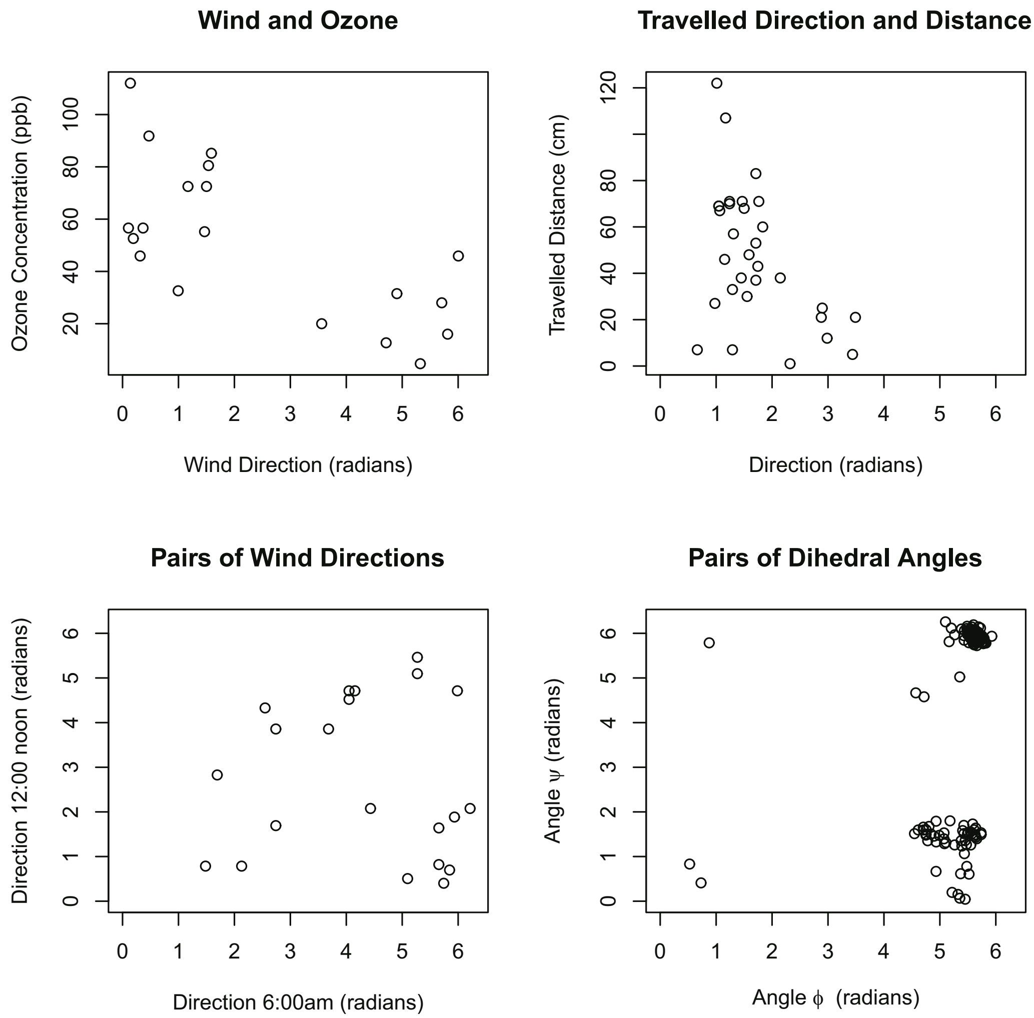

Figure 2: Scatterplots of the circular-linear (upper plots) and circular-circular (lower plots) real datasets. The upper left scatterplot shows the wind direction (relative to north) and ozone concentration (ppb) datapoints of the dataset of Johnson and Wehrly (1977). The upper right plot corresponds to the small blue periwinkles dataset on travelled distance (cm) and direction analyzed by Fisher (1993). The bottom left plot corresponds to the Johnson and Wehrly (1977) dataset on pairs of wind directions (relative to north) at 6:00 am and 12:00 noon in a weather monitoring station. At last, the bottom right includes the pairs of dihedral angles in segments alanine-alanine-alanine of proteins originally analyzed by Fernández-Durán (2007).

possible positive association between the circular and linear variables and considering the periodicity of the circular random variable. By applying the Pycke and Rayleigh circular uniformity tests to the difference modulus 2π of the angular probability transforms of the circular and linear variables, we obtained p-values of 0.0133 and 0.0077, respectively, thus rejecting the null hypothesis of independence (uniformity) at a 5% significance level for the wind direction and ozone concentration. The value of the dependence measure $\hat{\lambda}_{c_0}$ was calculated to be 0.4383. Fisher [17] reached the same conclusion by considering an expected sine-wave functional form for the conditional expected value of ozone concentration given the wind direction.

A second example analyzed by Fisher [17] is a dataset on the directions and distances travelled by 31 small blue periwinkles after undergoing transplantation from their normal place of living. The top-right plot depicted in Figure 2 includes the scatterplot for this dataset, further indicating a possible negative association between the direction and travelled distance. When applying the uniformity test to the sum modulus 2π of the angular probability transforms of the direction and distance, a p-value of 0.0087 for the Pycke test and a p-value of 0.0096 for the Rayleigh test were obtained, which rejected independence in accordance with the results obtained by Fisher [17] when fitting a circular-linear regression model with von Mises errors with non-constant dispersion. The dependence measure $\hat{\lambda}_{c_0}$ was calculated to be 0.2668.

5.1.2 Circular-circular real examples

The first example in the circular-circular test of independence corresponds to pairs of wind directions measured at a weather monitoring station at Milwaukee. The measurements were taken at 6:00 and 12:00 o'clock for 21 consecutive days and were originally included in Johnson and Wehrly [28]. The bottom-left plot depicted in Figure 2 includes a scatterplot of the pairs of wind directions, which indicates a possible positive association between the two angles. Fisher [17] listed this dataset as the B.21 dataset, and the main conclusion of Fisher [17] was that there exists a strong positive association between the wind directions when applying a hypothesis test based on a circular-circular correlation coefficient. When applying the proposed methodology to the difference modulus 2π of the angular probability transforms, a p-value of 0.0148 for the Pycke test and a p-value of 0.0075 for the Rayleigh test were obtained, which rejected the null hypothesis of independence between the two angles at a 5% significance level. The dependence measure $\hat{\lambda}_{c_0}$ was calculated to be one.

The second example corresponds to 233 pairs (ϕ, ψ) of dihedral angles in segments alanine-alanine-alanine of proteins that were originally analyzed by Fernández-Durán [11] using bivariate NNTS models demonstrating the dependence between the two angles. The scatterplot of these two angles is included in the bottom-right plot of Figure 2, and the type of association between the two angles, whether it is positive or negative is not evident. When applying the independence test to the difference modulus 2π of the angular probability transforms of the two angles, we obtained p-values of 0.0014 and 0.0064 for the Rayleigh and Pycke circular uniformity tests, respectively, thus clearly rejecting the null hypothesis of independence in favor of a positive association.

6 Conclusion

By using the result that the sum modulus 2π of independent circular uniform random variables is circular and uniformly distributed, a general test of independence based on the angular integral probability transform was developed. We demonstrated its use, particularly when at least one of the variables is an angle, that is a circular random variable, further implying that testing for independence in the circular case could be equivalent to testing for circular uniformity. From the simulation study presented in this article, it is clear that the proposed test is particularly useful for bivariate cases of circular-circular and circular-linear pairs of random variables with more power than the Wilks and empirical copula independence tests. We reached this conclusion by simulating samples from NNTS densities, in which the degree of closeness to the circular uniform distribution was under control. Although the proposed independence test can be applied to the linear-linear case, its power is smaller than that of the commonly used independence tests converging to the power of the common independence tests when the sample size increases. The proposed independence test can be put in practice and it demonstrated superior performance when at least one circular random variable is among the random variables used to test for independence. In addition, a new measure of dependence was introduced based on the fitting of an NNTS density by considering $M = 1$ to the sum (difference) modulus 2π of the APITs.

Acknowledgments: We express our sincere gratitude to the Asociación Mexicana de Cultura, A.C. for their support.

Funding information: No funding was received to assist with the preparation of this manuscript.

Conflict of interest: The authors have no conflict of interest to declare that are relevant to the content of this article.

References

[1] Agostinelli, C., & Lund, U. (2017). R package 'circular': Circular Statistics (version 0.4-93). https://r-forge.r-project.org/projects/circular/.

[2] Berrett, T. B., & Samworth, R. J. (2019). Nonparametric independence testing via mutual information. *Biometrika*, *106*(3), 547–556.

[3] Blum, J. R., Keifer, J., & Rosenblatt, M. (1961). Distribution Free Tests of Independence Based on the Sample Distribution Function. *Annals of Mathematical Statistics*, *32*(2), 485–498.

[4] Cinar, O., & Viechtbauer, W. (2022). The poolr Package for Combining Independent and Dependent p values. *Journal of Statistical Software*, *101*, 1–42.

[5] Csörgö, S. (1985). Testing for independence by the empirical characteristic function. *Journal of Multivariate Analysis*, *16*, 290–299.

[6] Deheuvels, P. (1981). An asymptotic decomposition for multivariate distribution-free tests of independence. *Journal of Multivariate Analysis*, *11*(1), 102–113.

[7] DeWet, T.(1980). Cramér-von Mises Tests for Independence. *Journal of Multivariate Analysis*, *10*, 38–50.

[8] Einmahl, J. H., & McKeague, I. W. (2003). Empirical likelihood based hypothesis testing. *Bernoulli*, *9*(2), 267–290.

[9] Fernández-Durán, J. J. (2004a). Modelling ground-level ozone concentration using copulas. In G. J., Erickson, & Y., Zhai (Eds.) *23rd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Proceedings of the Conference held 3-8 August, 2003 in Jackson Hole, Wyoming. AIP Conference Proceeding* (Vol. 707, pp. 406–413). New York: American Institute of Physics.

[10] Fernández-Durán, J. J. (2004b). Circular distributions based on nonnegative trigonometric sums. *Biometrics*, *60*, 499–503.

[11] Fernández-Durán, J. J. (2007). Models for circular-linear and circular-circular data constructed from circular distributions based on nonnegative trigonometric sums. *Biometrics*, *63*(2), 579–585.

[12] Fernández-Durán, J. J., & Gregorio-Domínguez, M. M. (2010). Maximum likelihood estimation of nonnegative trigonometric sums models using a Newton-like algorithm on manifolds. *Electronic Journal of Statistics*, *4*, 1402–10.

[13] Fernández-Durán, J. J., & Gregorio-Domínguez, M. M. (2016). CircNNTSR: An R package for the statistical analysis of circular, multivariatecircular, and spherical data using nonnegative trigonometric sums. *Journal of Statistical Software*, *70*, 1–19.

[14] Fisher, N. I., & Lee, A. J. (1981). Nonparametric measures of angular-linear association. *Biometrika*, *68*, 629–636.

[15] Fisher, N. I., & Lee, A. J. (1982). Nonparametric measures of angular-angular association. *Biometrika*, *69*, 315–321.

[16] Fisher, N. I., & Lee, A. J. (1983). A correlation coefficient for circular data. *Biometrika*, *70*, 327–332.

[17] Fisher, N. I.(1993). *Statistical analysis of circular data*. Cambridge, New York: Cambridge University Press.

[18] Fitak, R. R., & Johnsen, S. (2017). Bringing the analysis of animal orientation data full circle: Model-based approaches with maximum likelihood. *Journal of Experimental Biology*, *220*, 3878–3882.

[19] Genest, C., & Rémillard, B. (2004). Test of independence and randomness based on the empirical copula process. *Test*, *13*, 335–369.

[20] Genest, C., & Verret, F. (2005). Locally most powerful rank tests of independence for copula models. *Nonparametric Statistics*, *17*, 521–539.

[21] Genest, C., Nešlehová, J. G., Rémillard, B., & Murphy, O. A. (2019). Testing for independence in arbitrary distributions. *Biometrika*, *106*, 47–68.

[22] Goeman, J. J., & Solari, A. (2014). Multiple hypothesis testing in genomics. *Statistics in Medicine*, *33*, 1946–1978.

[23] Herwatz, H., & Maxand, S. (2020). Nonparametric tests for independence: A review and comparative simulation study with an application to malnutrition datain India. *Statistical Papers*, *61*, 2175–2201.

[24] Hoeffding, W. (1948). A non-parametric test of independence. *The Annals of Mathematical Statistics*, *19*(4), 546–557.

[25] Hofert, M., Kojadinovic, I., Maechler, M., & Yan, J. (2022). *Copula: Multivariate dependence with copulas*. R package version 1.1-1. https://CRAN.R-project.org/package=copula.

[26] Jammalamadaka, S. R., & SenGupta, A. (2001). *Topics in Circular Statistics*. River Edge, N.J.: World Scientific Publishing, Co.

[27] Joe, H. (1990). Multivariate entropy measures of multivariate dependence. *Journal of the American Statistical Association*, *84*, 157–164.

[28] Johnson, R. A., & Wehrly, T. (1977). Measures and models for angular correlation and angular-linear correlation. *Journal of the Royal Statistical Society, Series B*, *39*(2), 222–229.

[29] Kallenberg, W. C. M., & Ledwina, T. (1999). Data driven rank tests for independence. *Journal of the American Statistical Association*, *94*, 285–301.

[30] Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*(1/2), 81–93.

[31] Kendall, M. G., & Stuart, A. (1951). The advanced theory of statistics. *Inference and Relationship*(Vol. 2). New York: Hafner publishing Company.

[32] Kojadinovic, I., & Yan, J. (2010). Modeling multvariate distributions with continuous margins using the copula R package. *Journal of Statistical Software*, *34*(9), 1–20.

[33] Landler, L., Ruxton, G. D., & Malkemper, E. P. (2019). The Hermans-Rasson test as a powerful alternative to the Rayleigh test for circular statistics in biology. *BMC Ecology*, *19*, 30.

[34] Mardia, K. V., & Kent, J. T. (1991). Rao score tests for goodness of fit and independence. *Biometrika*, *78*(2), 355–363.

[35] Mardia, K. V., & Jupp, P. E. (2000). *Directional statistics*. Chichester, New York: John Wiley and Sons.

[36] Nelsen, R. (1999). *An introduction to copulas*. New York: Springer Verlag.

[37] Pearson, K. (1920). Notes on the history of correlation. *Biometrika*, *13*(1), 25–45.

[38] Pewsey, A., Neuhäuser, M., & Ruxton, G. D. (2013). *Circular statistics in R*. Oxford, U.K.: Oxford University Press.
[39] Pewsey, A., & Kato, S. (2016). Parametric bootstrap goodness-of-fit testing for Wehrly-Johnson bivariate circular distributions. *Statistics and Computing*, *26*, 1307–1317.
[40] Pfister, N., Bühlmann, P., Schölkopf, J. N., & Peters, J. (2018). Kernel-based tests for joint independence. *Journal of the Royal Statistical Society, Series B*, *80*(1), 5–31.
[41] Pycke, J-R. (2010). Some tests for uniformity of circular distributions powerful against multimodal alternatives. *The Canadian Journal of Statistics*, *38*, 80–96.
[42] R Core Team. (2021). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria*. http://www.R-project.org/.
[43] Roy, A. (2020). Some Copula-based tests of independence among several random variables having arbitrary probability distributions. *Stat*, *9*(1), e263.
[44] Roy, A., Ghosh, A. K., Goswami, A., & Murthy, C. A. (2020). Some new Copula based distribution-free tests of independence among several random variables. *Sankhya A*, *84*, 556–596.
[45] Sklar, A. (1959). Fonctions de Répartition à n Dimensions et Leurs Marges. *Publications de laInstitut de Statistique de laUniversité de Paris*, *8*, 229–231.
[46] Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, *15*(1), 72–101.
[47] Upton, G. J. G., & Fingleton, B. (1989). *Spatial data analysis by Example Vol. 2 (Categorical and Directional Data)*. Chichester, New York: John Wileyand Sons.
[48] Wehrly, T., & Johnson, R. A. (1980). Bivariate models for dependence of angular observations and a related Markov process. *Biometrika*, *67*(1), 255–256.
[49] Wilks, S. (1935). On the independence of k sets of normally distributed statistical variables. *Econometrica*, *3*, 309–326.