SORT 43 (2) July-December 2019, 317-336 DOI: 10.2436/20.8080.02.90

A class of goodness-of-fit tests for circular distributions based on trigonometric moments

S. Rao Jammalamadaka[1], M. Dolores Jiménez-Gamero[2] and Simos G. Meintanis[3,4]

Abstract

We propose a class of goodness–of–fit test procedures for arbitrary parametric families of circular distributions with unknown parameters. The tests make use of the specific form of the characteristic function of the family being tested, and are shown to be consistent. We derive the asymptotic null distribution and suggest that the new method be implemented using a bootstrap resampling technique that approximates this distribution consistently. As an illustration, we then specialize this method to testing whether a given data set is from the von Mises distribution, a model that is commonly used and for which considerable theory has been developed. An extensive Monte Carlo study is carried out to compare the new tests with other existing omnibus tests for this model. An application involving five real data sets is provided in order to illustrate the new procedure.

MSC: 62H15, 62G20.

Keywords: Goodness-of-fit, Circular data, Empirical characteristic function, Maximum likelihood estimation, von Mises distribution.

1 Introduction

Let Θ be an arbitrary circular random variable with cumulative distribution function (CDF) F. Then on the basis of independent and identically distributed (i.i.d.) copies $\vartheta_1, \ldots, \vartheta_n$ of Θ we are interested in testing goodness–of–fit (GOF) of the composite null hypothesis,

$$\mathscr{H}_0 : F \in \mathscr{F}_{\boldsymbol{\beta}} \tag{1}$$

against general alternatives, where $\mathscr{F}_{\boldsymbol{\beta}} = \{F(\cdot;\boldsymbol{\beta}), \boldsymbol{\beta} \in \mathscr{B}\}$ denotes a parametric family of CDFs indexed by the parameter $\boldsymbol{\beta} \in \mathscr{B} \subset \mathbb{R}^p$.

[1] Department of Statistics and Applied Probability, University of California Santa Barbara, USA, sreenivas@ucsb.edu

[2] Department of Statistics and Operations Research, University of Sevilla, Spain, dolores@us.es

[3] Department of Economics, National and Kapodistrian University of Athens, Greece, simosmei@econ.uoa.gr

[4] Unit for Business Mathematics and Informatics, North–West University, South Africa.

Received: April 2019

Accepted: October 2019

A well-known class of GOF tests that have been discussed in the literature, is obtained by comparing a nonparametric estimator of the CDF of Θ with the corresponding parametric estimator of the same quantity reflecting the null hypothesis. To this end, denote by $\widehat{\beta}$ a consistent estimator of the parameter β, and write $F(\cdot;\widehat{\beta})$ for the CDF corresponding to (1) with estimated parameter. Also let

$$F_n(x) = \frac{\#\{j : \vartheta_j' s \leq x\}}{n},$$

be the empirical CDF. Then, based on a distance function Δ, the CDF–based test statistics may be formulated as

$$\Delta_n := \Delta(F_n(\cdot), F(\cdot;\widehat{\beta})), \tag{2}$$

and rejects the null hypothesis $\mathscr{H}_0$ stated in (1) for large values of Δ_n. The specific type of distance Δ_n adopted in (2) leads to different GOF methods, chief among these are the Kuiper (1960) and the Watson (1961) tests, which are a variation of the Kolmogorov–Smirnov and the Cramér–von Mises tests, respectively. Note that both tests are appropriately adapted from the case of testing a distribution on the real line to the case of testing for circular distributions; see e.g. Jammalamadaka and SenGupta (2001) §7.2.1.

In this paper we suggest a new class of GOF tests which is based on the characteristic function (CF) of circular distributions. Such CF-based GOF tests for distributions on the real-line have proved to be more convenient, and compete well with corresponding methods based on the CDF; see for instance the normality test proposed by Epps and Pulley (1983), the test for the Cauchy distribution of Gürtler and Henze (2000), and the tests for the stable distribution suggested by Matsui and Takemura (2008), and Meintanis (2005).

The remainder of the paper is organized as follows. In Section 2 we introduce the new GOF procedure for circular distributions and prove consistency of the corresponding test criteria. In Section 3 we derive the limit distribution of the test statistic under the null hypothesis. Given the highly non–trivial structure of this distribution, we investigate in Section 4 the consistency of an appropriate resampling version of our method. In Section 5 the particular case of testing for the von Mises distribution is studied in detail. The finite–sample properties of the test are illustrated by means of a Monte Carlo study in Section 6, while Section 7 provides an application. Section 8 includes a brief summary and discussion. The paper contains a Supplement that includes the necessary R scripts for the benefit of potential users. Technical assumptions and proofs are deferred to the Appendix.

2 Tests based on the characteristic function

In a somewhat similar spirit with the Kuiper and Watson tests that use a distance between CDFs, we propose to use a distance between CFs instead of the CDFs. To this end, write

$\varphi(r) = \mathbb{E}(e^{\mathrm{i}r\Theta}),\ r \in \mathbb{R}$, for the CF of Θ and define the empirical CF corresponding to $\vartheta_1, \ldots, \vartheta_n$, as

$$\varphi_n(r) = \frac{1}{n}\sum_{j=1}^{n} e^{\mathrm{i}r\vartheta_j}, \quad (\mathrm{i} = \sqrt{-1}). \tag{3}$$

Also write $\varphi(\cdot;\boldsymbol{\beta}) := \Re\varphi(r;\boldsymbol{\beta}) + \mathrm{i}\Im\varphi(r;\boldsymbol{\beta})$ for the CF under the null hypothesis, where $\Re(z)$ (resp. $\Im(z)$) denotes the real (resp. imaginary) part of a complex number. In this paper we consider CF–based test statistics in the form $\Delta\ (\varphi_n(\cdot),\ \varphi(\cdot;\widehat{\boldsymbol{\beta}}))$. As before, rejection is for large values of the test statistic.

Specifically we consider a Cramér–von Mises type distance. However, since for circular distributions the CF needs to be evaluated only at integer values (Jammalamadaka and SenGupta, 2001, §2.2), and taking into account further the symmetry property of the CF and the empirical CF, our test statistic can be formulated as

$$C_{n,p} = n\sum_{r=0}^{\infty} \left|\varphi_n(r) - \varphi(r;\widehat{\boldsymbol{\beta}})\right|^2 p(r), \tag{4}$$

where $p(\cdot)$ denotes a probability function over the non–negative integers.

By straightforward algebra we have from (4)

$$C_{n,p} = n\sum_{r=0}^{\infty} \left\{R_n(r;\widehat{\boldsymbol{\beta}}) + I_n(r;\widehat{\boldsymbol{\beta}})\right\} p(r),$$

with

$$R_n(r;\widehat{\boldsymbol{\beta}}) = \left\{\frac{1}{n}\sum_{j=1}^{n}\cos(r\vartheta_j) - \Re\varphi(r;\widehat{\boldsymbol{\beta}})\right\}^2$$

and

$$I_n(r;\widehat{\boldsymbol{\beta}}) = \left\{\frac{1}{n}\sum_{j=1}^{n}\sin(r\vartheta_j) - \Im\varphi(r;\widehat{\boldsymbol{\beta}})\right\}^2.$$

Because of the one–to–one correspondence between CFs and CDFs, it readily follows that the test based on $C_{n,p}$ is consistent against any fixed alternative to $\mathscr{H}_0$ provided that

$$p(r) > 0, \quad \forall \quad r \geq 0. \tag{5}$$

To see this, assume that the estimator $\widehat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ has a strong probability limit, say $\boldsymbol{\beta}^0$, even under alternatives, and that $\varphi(r;\boldsymbol{\beta})$ is continuous as a function of $\boldsymbol{\beta}$. Then since $\left|\varphi_n(r) - \varphi(r;\widehat{\boldsymbol{\beta}})\right|^2 \leq 4$, we have from (4),

$$\frac{C_{n,p}}{n} \longrightarrow \sum_{r=0}^{\infty} \left|\varphi(r) - \varphi(r;\boldsymbol{\beta}^0)\right|^2 p(r) \quad a.s. \quad \text{as } n \to \infty, \tag{6}$$

due to the strong consistency of the empirical CF (see Csörgő, 1981 and Marcus, 1981), and by invoking Lebesgue's dominated convergence theorem. In view of the uniqueness of the CF, the right–hand side of (6) is positive, unless $F(\cdot) = F(\cdot;\boldsymbol{\beta}^0)$, which shows the strong consistency of the test that rejects the null hypothesis $\mathscr{H}_0$ for large values of $C_{n,p}$ since, from Theorem 1 in next Section, $C_{n,p}$ is bounded in probability.

In the next section we investigate the large–sample behavior of $C_{n,p}$ under the null hypothesis. From now on, it will be assumed that (5) holds.

3 The limit null distribution of the CF test statistic

Let ℓ_p^2 denote the (separable) Hilbert space of all infinite sequences $z = (z_0, z_1, \ldots)$ of complex numbers such that $\sum_{r\geq 0} |z_r|^2 p(r) < \infty$, with the inner product defined as

$$\langle z, w\rangle_{\ell_p^2} = \sum_{r\geq 0} z_r \bar{w}_r p(r),$$

for $z = (z_0, z_1, \ldots)$, $w = (w_0, w_1, \ldots) \in \ell_p^2$, where for any complex number $x = a + \mathrm{i}b$, $\bar{x} = a - \mathrm{i}b$ stands for its complex conjugate. Let also $\|\cdot\|_{\ell_p^2}$ denote the norm in this space. With this notation our test statistic may be written as,

$$C_{n,p} = \|Z_n\|^2_{\ell_p^2}, \tag{7}$$

where $Z_n(r) = \sqrt{n}\{\varphi_n(r) - \varphi(r;\widehat{\boldsymbol{\beta}})\}$.

Also let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ and write

$$\nabla \Re\varphi(r;\boldsymbol{\beta}) = \left(\frac{\partial}{\partial\beta_1}\Re\varphi(r;\boldsymbol{\beta}), \ldots, \frac{\partial}{\partial\beta_p}\Re\varphi(r;\boldsymbol{\beta})\right)^\top,$$

$$\nabla \Im\varphi(r;\boldsymbol{\beta}) = \left(\frac{\partial}{\partial\beta_1}\Im\varphi(r;\boldsymbol{\beta}), \ldots, \frac{\partial}{\partial\beta_p}\Im\varphi(r;\boldsymbol{\beta})\right)^\top.$$

Next theorem shows convergence in distribution of $Z_n(\cdot)$ under Assumptions A, B and C stated in the Appendix.

Theorem 1 *Assume that $\vartheta_1, \ldots, \vartheta_n$, are i.i.d. copies of Θ and that Assumptions A, B and C are fulfilled. Then, under the null hypothesis $\mathscr{H}_0$, there is a centred Gaussian random element $\mathrm{Z}(\cdot)$ of ℓ_p^2 having covariance kernel*

$$K(r,s) = \mathbb{E}\{\Upsilon(r,\Theta;\boldsymbol{\beta})\overline{\Upsilon}(s,\Theta;\boldsymbol{\beta})\},$$

such that

$$Z_n \xrightarrow{\mathscr{L}} Z, \quad as \;\; n \to \infty,$$

where

$$\Upsilon(r,\Theta;\boldsymbol{\beta}) = \cos(r\Theta) - \Re\varphi(r;\boldsymbol{\beta}) - \nabla\Re\varphi(r;\boldsymbol{\beta})^\top L(\Theta;\boldsymbol{\beta})$$
$$+\,\mathrm{i}\left\{\sin(r\Theta) - \Im\varphi(r;\boldsymbol{\beta}) - \nabla\Im\varphi(r;\boldsymbol{\beta})^\top L(\Theta;\boldsymbol{\beta})\right\},$$

with $L(\Theta;\boldsymbol{\beta})$ defined in Assumption A.

In view of (7), the asymptotic null distribution of $C_{n,p}$ stated in next corollary is an immediate consequence of Theorem 1 and the Continuous Mapping Theorem.

Corollary 1 *Suppose that assumptions in Theorem 1 hold, then*

$$C_{n,p} \xrightarrow{\mathscr{L}} \|Z\|^2_{\ell^2_p},$$

where $Z(\cdot)$ is the Gaussian random element appearing in Theorem 1.

Remark 1 *The distribution of $\|Z\|^2_{\ell^2_p}$ is the same as that of $\sum_{j=1}^{\infty}\lambda_j N_j^2$, where $\lambda_1,\lambda_2,\ldots$ are the positive eigenvalues of the integral operator $f \mapsto Af$ on ℓ^2_p associated with the kernel $K(\cdot,\cdot)$ given in Theorem 1, i.e., $(Af)(r) = \sum_{s\geq 0}K(r,s)f(s)p(s)$, and $N_1,N_2,\ldots$ are i.i.d. standard normal random variables. In general, the calculation of those eigenvalues is a very difficult task.*

Remark 2 *Assumptions A, B and C in Theorem 1 are quite standard in the context of GOF testing. Specifically Assumption A refers to an asymptotic (Bahadur) representation of a given estimator of the parameter $\boldsymbol{\beta}$ and is satisfied by common estimators such as maximum likelihood and moment estimators. Assumptions B and C imply smoothness of the CF as a function of $\boldsymbol{\beta}$.*

Since our assumptions are relatively weak, our CF approach is quite general and may be applied for testing GOF for a wide spectrum of circular distributions. In Section 5 we will specialize to a CF–based GOF test for the von–Mises distribution, which is as popular for circular data as the Gaussian distribution is for linear data.

4 The parametric bootstrap

As pointed out in Remark 1, the asymptotic null distribution of the test statistic $C_{n,p}$ is complicated and depends on several unknown quantities in a highly complicated manner. There exists no feasible approximation of the distribution in Theorem 1 which will allow us to actually carry out the test. We study here a resampling method labelled "parametric bootstrap", which is a computer–assisted automatic procedure for performing this task. The parametric bootstrap estimates the null distribution of the test statistic $C_{n,p}$ by means of its conditional distribution, given the data, when the data come from $F(\cdot;\widehat{\beta})$. Although the exact bootstrap estimator is still difficult to derive, it can be approximated as outlined below within the (fairly general) setting considered in Section 3. Specifically, write for simplicity $C^{\mathrm{o}}_{n,p} := C_{n,p}(\vartheta_1,\dots,\vartheta_n;\widehat{\beta})$ for the test statistic based on the original observations. Then parametric bootstrap critical points are calculated in practice as follows:

(i) Generate i.i.d. observations, $\{\vartheta^*_j, 1 \le j \le n\}$ from $F(\cdot;\widehat{\beta})$.
(ii) Using the bootstrap observations $\{\vartheta^*_j, 1 \le j \le n\}$, obtain the bootstrap estimate $\widehat{\beta}^*$ of β.
(iii) Calculate the bootstrap test statistic, say $C^*_{n,p} := C_{n,p}(\vartheta^*_1,\dots,\vartheta^*_n;\widehat{\beta}^*)$.
(iv) Repeat steps (i) to (iii) a number of times, say B, and obtain $\{C^{*\mathrm{b}}_{n,p}\}^{\mathrm{B}}_{b=1}$.
(v) Calculate the critical point of a test of size α as the order $(1-\alpha)$ empirical quantile $C_{1-\alpha}$ of $\{C^{*\mathrm{b}}_{n,p}\}^{\mathrm{B}}_{b=1}$.

In next theorem we show that, under Assumptions A*, B* and C stated in the Appendix, this procedure provides a consistent estimator of the null distribution of the test statistic. With this aim, as in Section 2, we will assume that the estimator of $\widehat{\beta}$ has a strong probability limit, say β^0, even under alternatives. Let P_β denote the probability by assuming that the data come from $F(\cdot;\beta)$ and let $P_\star$ denote the bootstrap probability.

Theorem 2 *Assume that $\vartheta_1,\dots,\vartheta_n$, are i.i.d. copies of Θ and that Assumptions A*, B* and C are fulfilled. Then,*

$$\sup_x \left| P_*(C^*_{n,p} \le x) - P_{\beta^0}(C_{n,p} \le x) \right| \to 0 \quad a.s., \quad as \;\; n \to \infty.$$

Theorem 2 holds whether the null hypothesis is true or not. In particular, if $\mathscr{H}_0$ is true, then it states that the bootstrap distribution and the null distribution of $C_{n,p}$ are close. Thus the test Ψ^*, which rejects the null when $C^o_{n,p} > C_{1-\alpha}$, is asymptotically correct in the sense that $\lim_{n\to\infty} P(\Psi^* = 1) = \alpha$, when the null hypothesis is true. Also an immediate consequence of (6) and Theorem 2 is that the test Ψ^* is consistent, that is $P(\Psi^* = 1) \to 1$, as $n \to \infty$, whenever $F \notin \mathscr{F}_\beta$.

5 Tests for the von Mises distribution

5.1 Goodness-of-fit tests

For data distributed over the unit circle, the von Mises distribution (vMD), also called the Circular Normal distribution, is the pre-eminent model in circular data analysis when one has reason to believe the data might be symmetric and unimodal, much as the Normal distribution is on the real line. Sampling theory and inferential methods have been developed for this model, and as such it is a natural choice for our consideration. The density of the vMD with parameter vector $\boldsymbol{\beta} := (\mu, \kappa)$ is given by

$$f(\vartheta; \mu, \kappa) = \frac{1}{2\pi \mathscr{I}_0(\kappa)} e^{\kappa \cos(\vartheta - \mu)}, \; 0 \leq \vartheta < 2\pi, \tag{8}$$

where $\mathscr{I}_r(\cdot)$ denotes the modified Bessel function of the first kind of order r, and $0 \leq \mu < 2\pi$ and $\kappa \geq 0$ are location and concentration parameters, respectively.

Our CF–based test utilizes the CF corresponding to (8) which is given by

$$\varphi(r; \mu, \kappa) = e^{\mathrm{i} r \mu} A_r(\kappa), \; r \in \mathbb{Z}, \tag{9}$$

where $A_r(\kappa) = \mathscr{I}_r(\kappa)/\mathscr{I}_0(\kappa)$.

Specifically the test statistic figuring in (4) may readily be written as

$$C_{n,p} = n \sum_{r=0}^{\infty} |\widehat{\varphi}_n(r) - \varphi(r; 0, \hat{\kappa})|^2 \, p(r) = S_1 + S_2 - 2S_3, \tag{10}$$

with $\widehat{\varphi}_n(r)$ the empirical CF of $\widehat{\vartheta}_1, \ldots, \widehat{\vartheta}_n$,

$$S_1 = \frac{1}{n} \sum_{j,k=1}^{n} \mathscr{E}_1(\widehat{\vartheta}_j - \widehat{\vartheta}_k), \tag{11}$$

$$S_2 = n \mathscr{E}_2(\widehat{\kappa}), \tag{12}$$

and

$$S_3 = \sum_{j=1}^{n} \mathscr{E}_3(\widehat{\vartheta}_j; \widehat{\kappa}), \tag{13}$$

where $(\widehat{\mu}, \widehat{\kappa})$ is a consistent estimator of the parameter (μ, κ), and $\widehat{\vartheta}_j = \vartheta_j - \widehat{\mu}, \; j = 1, \ldots, n$. The series appearing in (11)-(13) are defined as

$$\mathscr{E}_1(\theta) = \sum_{r=0}^{\infty} \cos(\theta r) p(r),$$

$$\mathscr{E}_2(\kappa) = \sum_{r=0}^{\infty} A_r^2(\kappa) p(r),$$

and

$$\mathscr{E}_3(\theta;\kappa) = \sum_{r=0}^{\infty} \cos(\theta r) A_r(\kappa) p(r).$$

To proceed further note that all three series above may be viewed as expectations of corresponding quantities taken with respect to the law $p(r)$, and while these expectations are generally hard to obtain, they may be approximated by Monte Carlo by means of simulating i.i.d. variates from the law $p(r)$. In fact certain choices of $p(r)$ lead to closed form expressions, at least for the expectation in (11). Specifically if we let $p(r)$ be the Poisson law with parameter λ, we have

$$\mathscr{E}_1(\theta) = \cos(\lambda \sin\theta) e^{\lambda(\cos\theta - 1)}.$$

As for the calculation of S_2 and S_3 and since the corresponding series appearing in (12)-(13) converge rapidly, instead of Monte Carlo, we decided to approximate them by direct numerical computation of only a few terms. We have observed through simulations that summing up to $r = 100$ gives very accurate results. Strictly speaking this cut–off test is not universally consistent, but the practical effect on the power is negligible.

5.2 Estimation of parameters and a limit statistic

As for estimating parameters, we suggest the use of the maximum likelihood estimator (MLE) $\widehat{\beta} := (\widehat{\mu}, \widehat{\kappa})$ which is given by the following equations:

$$\frac{1}{n}\sum_{j=1}^{n} \sin(\vartheta_j - \widehat{\mu}) = 0, \;\; \frac{1}{n}\sum_{j=1}^{n} \cos(\vartheta_j - \widehat{\mu}) = A_1(\widehat{\kappa}). \tag{14}$$

It is well known that the MLE $\widehat{\mu}$ of μ satisfies $\widehat{\mu}(\vartheta_1 + a, \ldots, \vartheta_n + a) = \widehat{\mu}(\vartheta_1, \ldots, \vartheta_n) + a$, while the MLE $\widehat{\kappa}$ of κ satisfies $\widehat{\kappa}(\vartheta_1 + a, \ldots, \vartheta_n + a) = \widehat{\kappa}(\vartheta_1, \ldots, \vartheta_n)$, for each a, where the operations of addition in these equations are to be treated $\mathrm{mod}(2\pi)$ for circular data. Thus if one uses, instead of the original data $\vartheta_1, \ldots, \vartheta_n$, the centered data $\widehat{\vartheta}_j = \vartheta_j - \widehat{\mu}$, $j = 1, \ldots, n$, then the distribution of any test statistic that depends on $\widehat{\mu}$ via $\widehat{\vartheta}_j$, $j = 1, \ldots, n$, will not depend on the specific parameter–value of μ, and hence without loss of generality we can set $\mu = 0$. On the other hand, since the concentration parameter κ is a shape parameter, it cannot be standardized out. Consequently the distribution of such a test always depends on the value of this parameter. One way out is to use the limit null distribution for fixed κ along with a look–up table with a sufficiently dense grid on κ. This approach is suggested in Lockhart and Stephens (1985), and is fairly accurate for most of the parameter space if based on the MLE of κ, but as already mentioned in

Section 4 we will instead use the parametric bootstrap which consistently estimates the limit null distribution of any given test uniformly over κ.

We close this section with an interesting limit statistic resulting from $C_{n,p}$ appearing in (10). To this end notice that since $\varphi_n(0) = \varphi(0) = 1$, the first term in $C_{n,p}$ vanishes regardless of the distribution being tested, while the second term also vanishes on account of (14) since we employ the MLEs as estimators of μ and κ. Now write $C_{n,\lambda}$ for the criterion in (10) with $p(r)$ being the Poisson probability function, with parameter λ. Then we have

$$C_{n,\lambda} = e^{-\lambda}\left(|\widehat{\varphi}_n(2) - A_2(\widehat{\kappa})|^2 \frac{\lambda^2}{2} + \mathrm{o}(\lambda^2)\right), \ \lambda \to 0,$$

so that

$$\lim_{\lambda\to 0} \frac{2C_{n,\lambda}}{\lambda^2} = |\widehat{\varphi}_n(2) - A_2(\widehat{\kappa})|^2 := C_{n,0}. \tag{15}$$

Notice that the limit statistic $C_{n,0}$ only uses information on the CF of the underlying law as this is information is reflected on the corresponding empirical trigonometric moment of order $r = 2$.

On the other hand the test statistic $C_{n,\lambda}$ (and more generally $C_{n,p}$) uses an infinite weighted sum in which the empirical trigonometric moments of all integer orders $r \geq 0$ are accounted for. Thus the probability function $p(r)$ plays the role of a weight function that typically downweights the higher order terms which are known to be more prone to the periodic behavior intrinsically present in the empirical CF. A natural related question is whether there is some optimal choice for the probability function $p(\cdot)$. As asserted by Bugni et al. (2009) in a related context, the weight function cannot be selected empirically as this would require knowing how the true data-generation process differs from the parametric model. In this connection, and using the analogy with the choice of kernel in density estimation, prior experience has shown that the specific functional form of $p(\cdot)$ is not all that important. Carrying this analogy further, one suspects that the value of λ might have some sway over the results. Proper choice of λ however translates to a highly non–trivial analytic problem for which there are only a few results available in the literature; see Tenreiro (2009) and Meynaoui et al. (2019). This option is empirically investigated in the next section.

6 Finite-sample comparisons and simulations

This section summarizes the results of a simulation study, designed to evaluate the proposed GOF test for the vMD, and compare its performance with other existing tests. As competitors we include the Kuiper test and the Watson test for which there exist computationally convenient formulae; see for instance Section 7.2.1 of Jammalamadaka and SenGupta (2001). Specifically let $U_j = F(\vartheta_j; \widehat{\mu}, \widehat{\kappa})$ and write $U_{(j)}, \ j = 1, \ldots, n$, for the corresponding order statistics. Then we have

$$\mathrm{K} = \max_{1 \le j \le n} \left\{ U_{(j)} - \frac{j-1}{n} \right\} + \max_{1 \le j \le n} \left\{ \frac{j}{n} - U_{(j)} \right\}.$$

$$\mathrm{W} = \frac{1}{12n} + \sum_{j=1}^{n} \left(\left(U_{(j)} - \frac{2j-1}{2n} \right) - \left(\overline{U} - \frac{1}{2} \right) \right)^2,$$

where $\overline{U} = n^{-1} \sum_{j=1}^{n} U_j$.

We also include a test statistic based on the characterization of maximum entropy of the vMD suggested by Lund and Jammalamadaka (2000), denoted by E. These three criteria will be included in our Monte Carlo study. For our test statistic we took as $p(r)$ the probability function of a Poisson law with mean λ. This test is indexed by λ, and will be denoted by C_λ. We note that there exist alternative tests such as the conditional tests suggested by Lockhart (2012) (Lockhart, O'Reilly and Stephens, 2007, 2009), which we do not consider in our simulation study.

The simulated distributions are (i) the vMD, $vM(0,\kappa)$, (ii) mixtures of vMDs, $(1-\epsilon)vM(\mu_1,\kappa_1) + \epsilon\, vM(\mu_2,\kappa_2)$, $\epsilon \in (0,1)$, (iii) the generalized vMD, $GvM(\mu_1,\mu_2,\kappa_1,\kappa_2)$, with probability density function given by

$$f(\theta;\mu_1,\mu_2,\kappa_1,\kappa_2) = \frac{1}{2\pi G_0(\mu_1-\mu_2,\kappa_1,\kappa_2)} \exp\{\kappa_1 \cos(\theta-\mu_1) + \kappa_2 \cos(\theta-\mu_2)\},$$

where $G_0(\delta,\kappa_1,\kappa_2) = (1/2\pi)\int_0^{2\pi} \exp\{\kappa_1 \cos(\theta) + \kappa_2 \cos(\theta+\delta)\} d\theta$, (see Gatto and Jammalamadaka, 2007) and (iv) the wrapped normal distribution, $wn(\mu,\rho)$, with probability density function given by

$$f(\theta;\mu,\rho) = \frac{1}{2\pi} \left(1 + 2 \sum_{m=-\infty}^{\infty} \rho^{p^2} \cos\{p(\theta-\mu)\} \right),$$

(Jammalamadaka and SenGupta, 2001, Ch. 2). Table 1 displays the specific alternatives (ii) and (iii), while the densities of such alternatives jointly with the density of the closer vMD (in the sense that the parameters are chosen so that they minimize the Kullback-Leibler distance), are depicted in Figure 1. These alternatives exhibit either bimodality and/or asymmetry and/or heavier tails than the vMD. We also considered several instances of the family of wrapped normal distributions, which are known to possess densities that are quite close to those of the vMD. This fact can be graphically appreciated by looking at Figure 2, which displays the probability density function of a $wn(0,\rho)$ law for $\rho = 0.1(0.1)0.9$, together with the density of the closer vMD distribution (in the sense explained before). Looking at this figure it becomes evident that it is rather hard to discriminate between these distributions and the vMD, particularly for small and large values of ρ.

Table 1: *Alternatives (ii) and (iii).*

Alternative 1	$0.9vM(\pi,5)+0.1vM(\pi/2,5)$
Alternative 2	$0.8vM(\pi,5)+0.2vM(\pi/2,5)$
Alternative 3	$0.65vM(\pi,5)+0.35vM(\pi/2,5)$
Alternative 4	$0.5vM(\pi,5)+0.5vM(\pi/2,5)$
Alternative 5	$(2/3)vM(\pi,3)+(1/3)vM(0.62\pi,3)$
Alternative 6	$(1/3)vM(\pi,8)+(2/3)vM(\pi,0.1)$
Alternative 7	$GvM(0,0.5,1,0.6)$
Alternative 8	$GvM(0,0.5,1,0.2)$

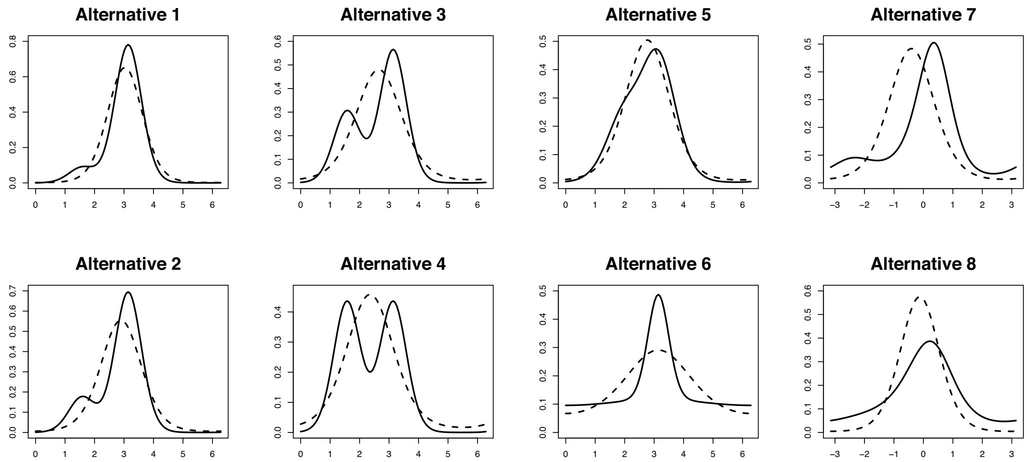

Figure 1: *Probability density function of alternatives in Table 1 (solid) and the probability density function of the closer vMD (dashed).*

All computations were performed using programs written in the R language. Specifically, we used the package CircStats for generating data from a vMD, and from mixtures of vMDs, and in order to calculate the MLEs of the parameters. Data from the generalized vMD were generated by the acceptance-rejection algorithm of von Neumann suggested in Gatto (2008). In all cases the p–values were approximated by using the parametric bootstrap algorithm given in Section 4 with B $= 1000$. For the benefit of potential users, we include the R codes necessary for calculating the new test statistics, in a Supplement.

We tried a wide range of values for λ and observed that the power of the proposed test depends on the value of λ. Tables 2 and 3 report the results for those values of λ giving the greater, or closer to the greater power, in all tried alternatives. Table 2 displays the observed proportion of rejections in 1,000 Monte Carlo samples of size $n = 25$ under the null hypothesis and for the set of alternatives in Table 1. We also tried $n = 50$ and $n = 100$ yielding a quite similar picture (in the sense of comparison between tests, but

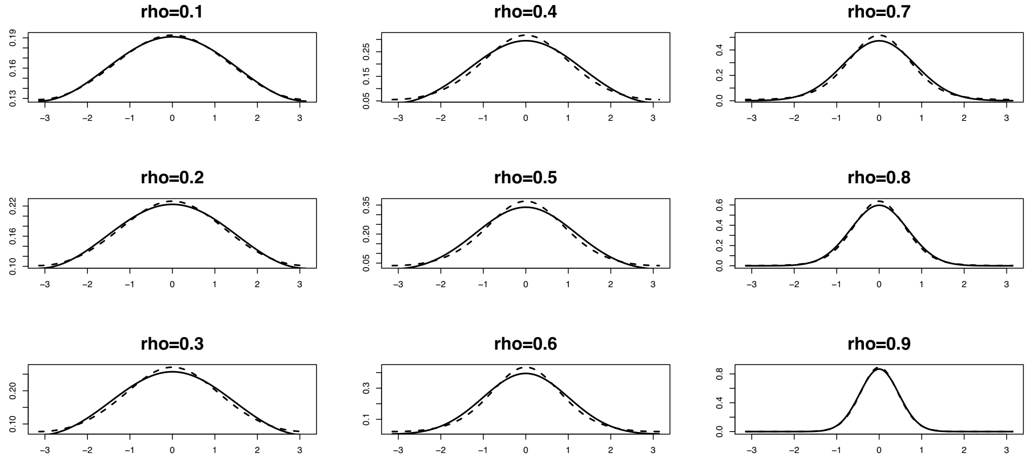

Figure 2: *Probability density function of a* $wn(0,\rho)$ *law for* $\rho = 0.1(0.1)0.9$ *(solid), and the probability density function of the closer vMD (dashed).*

Table 2: *Observed proportion of rejection in 1,000 Monte Carlo samples of size* $n = 25$.

Law	α	E	K	W	$C_{0.3}$	$C_{0.5}$	$C_{0.7}$	$C_{0.9}$	C_1
$vM(0,1)$	0.05	0.053	0.047	0.044	0.062	0.059	0.059	0.061	0.059
	0.10	0.105	0.093	0.089	0.115	0.117	0.117	0.122	0.122
$vM(0,5)$	0.05	0.054	0.048	0.047	0.033	0.032	0.036	0.040	0.043
	0.10	0.106	0.099	0.095	0.086	0.090	0.092	0.091	0.092
$vM(0,10)$	0.05	0.051	0.046	0.046	0.039	0.042	0.042	0.042	0.043
	0.10	0.103	0.090	0.092	0.093	0.096	0.095	0.095	0.098
Alt. 1	0.05	0.171	0.150	0.166	0.311	0.310	0.304	0.309	0.307
	0.10	0.267	0.235	0.272	0.450	0.451	0.445	0.443	0.437
Alt. 2	0.05	0.114	0.255	0.337	0.459	0.478	0.482	0.487	0.487
	0.10	0.197	0.422	0.470	0.631	0.634	0.645	0.635	0.627
Alt. 3	0.05	0.048	0.411	0.477	0.550	0.570	0.589	0.596	0.600
	0.10	0.097	0.547	0.620	0.720	0.737	0.747	0.749	0.742
Alt. 4	0.05	0.036	0.500	0.541	0.559	0.583	0.604	0.617	0.623
	0.10	0.059	0.627	0.688	0.719	0.739	0.741	0.750	0.751
Alt. 5	0.05	0.019	0.092	0.090	0.079	0.084	0.090	0.094	0.097
	0.10	0.056	0.151	0.163	0.176	0.184	0.195	0.209	0.211
Alt. 6	0.05	0.139	0.244	0.259	0.249	0.252	0.262	0.274	0.279
	0.10	0.243	0.358	0.397	0.379	0.390	0.397	0.410	0.409
Alt. 7	0.05	0.059	0.253	0.318	0.646	0.631	0.608	0.594	0.581
	0.10	0.102	0.381	0.465	0.774	0.757	0.737	0.721	0.713
Alt. 8	0.05	0.003	0.131	0.154	0.130	0.153	0.176	0.192	0.198
	0.10	0.007	0.212	0.244	0.267	0.305	0.320	0.329	0.337

Table 3: *Observed proportion of rejection in 1,000 Monte Carlo samples of size n from a wn$(0,\rho)$ law.*

n	ρ	α	E	K	W	$C_{0.3}$	$C_{0.5}$	$C_{0.7}$	$C_{0.9}$	C_1
50	0.3	0.05	0.060	0.052	0.059	0.081	0.081	0.078	0.072	0.072
		0.10	0.116	0.116	0.113	0.131	0.131	0.131	0.129	0.132
	0.4	0.05	0.053	0.053	0.053	0.072	0.072	0.073	0.070	0.069
		0.10	0.096	0.103	0.103	0.140	0.139	0.136	0.133	0.131
	0.5	0.05	0.041	0.072	0.072	0.099	0.096	0.096	0.097	0.095
		0.10	0.084	0.139	0.130	0.182	0.179	0.174	0.174	0.172
	0.6	0.05	0.035	0.069	0.072	0.089	0.091	0.087	0.090	0.088
		0.10	0.062	0.142	0.149	0.184	0.182	0.182	0.184	0.183
	0.7	0.05	0.019	0.079	0.092	0.098	0.098	0.098	0.103	0.103
		0.10	0.046	0.139	0.157	0.182	0.187	0.192	0.195	0.191
100	0.3	0.05	0.048	0.057	0.055	0.074	0.072	0.071	0.070	0.067
		0.10	0.092	0.114	0.109	0.144	0.143	0.140	0.138	0.139
	0.4	0.05	0.052	0.097	0.092	0.125	0.123	0.123	0.123	0.123
		0.10	0.102	0.149	0.175	0.212	0.210	0.211	0.208	0.203
	0.5	0.05	0.031	0.095	0.107	0.171	0.168	0.162	0.159	0.158
		0.10	0.067	0.162	0.194	0.272	0.269	0.264	0.262	0.261
	0.6	0.05	0.030	0.106	0.122	0.203	0.196	0.185	0.176	0.173
		0.10	0.049	0.185	0.195	0.316	0.310	0.302	0.283	0.279
	0.7	0.05	0.021	0.117	0.108	0.162	0.159	0.157	0.153	0.153
		0.10	0.040	0.190	0.193	0.285	0.284	0.275	0.262	0.254

with greater powers as the sample size increases), and therefore we omit those results. By contrast, and since the power for $n = 25$ is quite low we opted to present results for larger sample size for wrapped normal alternatives. Specifically Table 3 presents the results for wrapped normal alternatives for sample size $n = 50$ and $n = 100$, and $\rho = 0.3(0.1)0.7$.

Regarding level, we conclude that the observed empirical rejection rates are reasonably close to the nominal values. In fact, for larger sample sizes (not displayed), we observed greater closeness. As for power, we observe that the power of the proposed test is comparable and most often greater than that of the tests based on the empirical CDF. On the other hand, the test based on the characterization of maximum entropy presents the poorest performance under the considered alternatives.

A natural question is which value of λ should be used in practical applications. Although the powers exhibited in the tables are quite close for the values of λ selected, it seems that $C_{0.5}$ has an intermediate behaviour in all tried cases, so we recommend $\lambda = 0.5$ as a compromise choice.

Another possibility is to choose λ by using some data-dependent method (see Cuparić, Milosević and Obradović, 2019, for a related approach). In this sense, Tenreiro (2019) has proposed a method for choosing the tuning parameter λ so that the power is maximized. It works as follows. Let $C_{n,\lambda}(\alpha)$ denote the upper α percentile of the null distribution of $C_{n,\lambda} = C_{n,\lambda}(\vartheta_1,\ldots,\vartheta_n)$. Assume that $\lambda \in \Lambda$, with Λ having a finite number of points. Then, reject H_0 if

$$\max_{\lambda\in\Lambda}\{C_{n,\lambda} - C_{n,\lambda}(u)\} > 0,$$

where u is chosen so that the test has level α. The key point is the way to determine u. In the context discussed in Tenreiro (2019), it is assumed that the exact null distribution of the test statistic can be calculated (or at least it can be approximated by simulation). Since this is not our case, we have adapted his procedure to calculate u to our setting as follows:

1. First, we must approximate the critical points $C_{n,\lambda}(u)$, $u \in (0,1)$, $\lambda \in \Lambda$. With this aim, we generate B_1 bootstrap samples and estimate $C_{n,\lambda}(u)$ by means of their bootstrap analogues, $C^*_{1,n,\lambda}(u)$, for $u \in \{1/B_1, 2/B_1, \ldots, (B_1-1)/B_1\} := U_{B_1}$, $\lambda \in \Lambda$.
2. Then, we must calibrate u so that the test has level α. For this purpose, we generate B_2 bootstrap samples, independently of those generated in the first step, and determine $u^* \in U_{B_2}$ such that
$$P_*\left(\max_{\lambda\in\Lambda}\{C^*_{n,\lambda} - C^*_{1,n,\lambda}(u^*)\} > 0,\right) \le \alpha.$$
3. Finally,
$$\text{reject } H_0 \text{ if } \quad \max_{\lambda\in\Lambda}\{C_{n,\lambda} - C^*_{1,n,\lambda}(u^*)\} > 0. \tag{16}$$

In addition to the determination of u, another delicate issue is the choice of the set Λ, which has a strong effect on the power of the resulting test. In order to study the practical behaviour of test (16), we repeated the experiment in Table 2 for $\Lambda = \Lambda_1$ and $\Lambda = \Lambda_2$, with $\Lambda_1 = \{0.1, 0.3, 0.5, 0.7, 0.9, 1, 2, 3, 4, 5, 7, 10\}$ and $\Lambda_2 = \{0.3, 0.5, 0.7, 0.9, 1, 2\}$, and $B_1 = B_2 = 1000$. Table 4 display the results obtained. Comparing the powers in that table with those in Table 2 we conclude that as Λ increases, the power of the test (16) decreases. This fact was also observed in the simulations in Tenreiro (2019). The power for $\Lambda = \Lambda_2$ is in most cases smaller than that obtained for $\lambda = 0.5$.

Table 4: *Observed proportion of rejection in 1,000 Monte Carlo samples of size $n = 25$, for $\alpha = 0.05$.*

	Alt1	Alt2	Alt3	Alt4	Alt5	Alt6	Alt7	Alt8
Λ_1	0.200	0.327	0.426	0.442	0.050	0.213	0.458	0.103
Λ_2	0.280	0.439	0.549	0.563	0.077	0.280	0.562	0.156

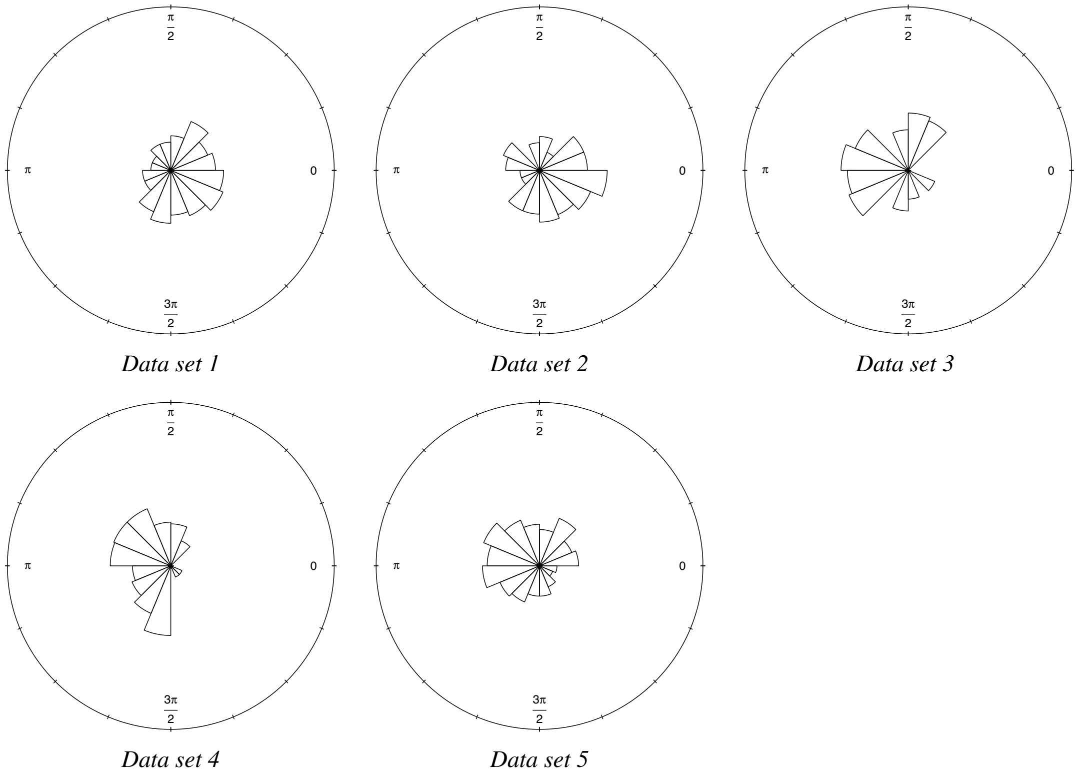

Figure 3: *Rose diagrams for the five real data sets.*

7 Real-data application

This section illustrates the proposed test on five real data sets. They come from a study by Taylor and Burns (2016) on the radial orientation of 2 species of mistletoes and 3 species of epiphytes, which the ecologists believe orient towards the direction of the availability of light and humidity. Specifically, Data Set 1 consists of $n = 67$ observations on *Peraxilla colensoi*, Data Set 2 consists of $n = 70$ observations on *Peraxilla tetrapetala*, Data Set 3 consists of $n = 65$ observations on *Asplenium flaccidum*, Data Set 4 consists of $n = 182$ observations on *Hymenophyllum multifidum*, and Data Set 5 consists of $n = 263$ observations on *Notogrammitis billardierei*. Taylor and Burns (2016) tested for uniformity in the five data sets and in all cases such hypothesis was rejected, indicating that the distribution of each of the studied species have certain orientation, as can be easily appreciated by looking at Figure 3, which displays the rose diagrams for each data set. So, it would be interesting to check if the data follow some distribution, such as the vMD. In fact, Taylor and Burns (2016) calculated certain confidence intervals based on the vMD. Table 5 reports the values of the maximum likelihood estimates by assuming a vMD, as well as the p-values for testing goodness-of-fit to that distri-

Table 5: *Maximum likelihood estimators of the parameters and p-values for the real data sets.*

	$\hat{\mu}$	$\hat{\kappa}$	K	W	$C_{0.5}$
1	2.5551	0.7700	0.5335	0.6410	0.4710
2	5.7677	0.8447	0.1505	0.2815	0.7555
3	2.8226	1.1120	0.0080	0.0050	0.0265
4	3.0454	1.2589	0.0080	0.0050	0.0220
5	2.5551	0.7699	0.8310	0.0060	0.0050

bution that resulted by applying the tests K, W and $C_{0.5}$. These three test criteria lean towards the null hypothesis for Data Set 1 and Data Set 2, and all of them suggest that the vMD is not a good model for Data Set 3 and Data Set 4. For Data Set 5, the tests W and $C_{0.5}$ reject that the vMD provides an adequate description of the data, while test K concludes in the opposite direction. From the power results in our simulations, we deduce that the vMD does not provide a satisfactory fit to Data Set 5.

8 Discussion

We suggest here a general class of GOF tests for circular distributions. The proposed test statistic may conveniently be expressed as a weighted L2–type distance between the empirical trigonometric moments and the corresponding theoretical quantities, and is shown to compete well with classical tests based on the CDF. Our method imposes minimal technical conditions is widely applicable for arbitrary distributions under test. Here however we focus specifically on GOF testing for the vMD because it is one of the most commonly used distributions in practice, and one would like to verify if this model fits a given data set before utilizing the various parametric tools that have been developed for this particular model.

A Appendix

All limits are understood to be taken as $n \to \infty$.

A.1 Technical assumptions

ASSUMPTION A. Under $\mathscr{H}_0$, if $\boldsymbol{\beta} \in \mathscr{B}$ denotes the true parameter value, then

$$\sqrt{n}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} L(\vartheta_j; \boldsymbol{\beta}) + o_P(1),$$

with $\mathbb{E}\{L(\Theta; \boldsymbol{\beta})\} = 0$ and $J(\boldsymbol{\beta}) = \mathbb{E}\{L(\Theta; \boldsymbol{\beta}) L(\Theta; \boldsymbol{\beta})^\top\} < \infty$.

ASSUMPTION B. $\frac{\partial}{\partial \boldsymbol{\beta}_k}\Re\varphi(r;\boldsymbol{\beta})$ and $\frac{\partial}{\partial \boldsymbol{\beta}_k}\Im\varphi(r;\boldsymbol{\beta})$, exist $\forall r \in \mathbb{N}_0$ and $1 \leq k \leq p$, and satisfy

$$\sum_{r\geq 0} \frac{\partial}{\partial \boldsymbol{\beta}_k}\Re\varphi(r;\boldsymbol{\beta})^2 p(r) < \infty,$$
$$\sum_{r\geq 0} \frac{\partial}{\partial \boldsymbol{\beta}_k}\Im\varphi(r;\boldsymbol{\beta})^2 p(r) < \infty.$$

Let $\|\cdot\|$ stand for the Euclidean norm.

ASSUMPTION C. For any $\varepsilon > 0$ there is a bounded neighborhood $\mathscr{N}_\varepsilon \subseteq \mathbb{R}^p$ of $\boldsymbol{\beta}$, such that if $\boldsymbol{\gamma} \in \mathscr{N}_\varepsilon$ then $\nabla\Re\varphi(r;\boldsymbol{\gamma})$ and $\nabla\Im\varphi(r;\boldsymbol{\gamma})$ exist and satisfy

$$\|\nabla\Re\varphi(r;\boldsymbol{\gamma}) - \nabla\Re\varphi(r;\boldsymbol{\beta})\| \leq \rho_{\Re}(r),\ \forall r \in \mathbb{N}_0, \quad \text{with} \quad \sum_{r\geq 0} \rho^2_{\Re}(r)p(r) < \varepsilon,$$
$$\|\nabla\Im\varphi(r;\boldsymbol{\gamma}) - \nabla\Im\varphi(r;\boldsymbol{\beta})\| \leq \rho_{\Im}(r),\ \forall r \in \mathbb{N}_0, \quad \text{with} \quad \sum_{r\geq 0} \rho^2_{\Im}(r)p(r) < \varepsilon.$$

Assumptions A^* and B^* below are a bit stronger than Assumptions A and B, respectively. They are required for the consistency of the parametric bootstrap null distribution estimator.

ASSUMPTION A^*. (a) There is a $\boldsymbol{\beta}^0 \in \mathscr{B}$ so that $\widehat{\boldsymbol{\beta}} \to \boldsymbol{\beta}^0$, a.s., $\boldsymbol{\beta}^0$ being the true parameter value if $\mathscr{H}_0$ is true,
(b)

$$\sqrt{n}\left(\widehat{\boldsymbol{\beta}}^* - \widehat{\boldsymbol{\beta}}\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} L(\vartheta^*_j;\widehat{\boldsymbol{\beta}}) + o_{P_*}(1),$$

with $\mathbb{E}_*\{L(\Theta^*;\widehat{\boldsymbol{\beta}})\} = 0$, $J(\widehat{\boldsymbol{\beta}}) = \mathbb{E}_*\{L(\Theta^*;\widehat{\boldsymbol{\beta}})L(\Theta^*;\widehat{\boldsymbol{\beta}})^\top\} \to J(\boldsymbol{\beta}^0) < \infty$, a.s.
(c) $\sup_{\boldsymbol{\beta}\in\mathscr{N}_0} \mathbb{E}_{\boldsymbol{\beta}}\left[\|L(\Theta;\boldsymbol{\beta})\|^2_{\ell^2_p} I\left\{\|L(\Theta;\boldsymbol{\beta})\|_{\ell^2_p} > \epsilon\sqrt{n}\right\}\right] \longrightarrow 0$, $\forall\epsilon > 0$, where $\mathscr{N}_0 \subseteq \mathscr{B}$ is an open neighborhood of $\boldsymbol{\beta}_0$, where $\mathbb{E}_{\boldsymbol{\beta}}$ stands for the expectation when data have CDF $F(x;\beta)$.

ASSUMPTION B^*. Assumption B holds true $\forall\boldsymbol{\beta}$ in an open neighborhood of $\boldsymbol{\beta}^0$, where $\boldsymbol{\beta}^0$ is as defined in Assumption A^*.

A.2 Proofs

Proof of Theorem 1

By Taylor expansion,

$$\Re\varphi(r;\widehat{\boldsymbol{\beta}}) = \Re\varphi(r;\boldsymbol{\beta}) + \nabla\Re\varphi(r;\boldsymbol{\beta})^\top(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + g_{1n}(r).$$

From Assumptions A and C, it follows that

$$\|\sqrt{n}g_{1n}\|^2_{\ell^2_p} = o_P(1).$$

From Assumptions A and B, it follows that

$$\nabla\Re\varphi(r;\boldsymbol{\beta})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}) = \nabla\Re\varphi(r;\boldsymbol{\beta})^\top\frac{1}{n}\sum_{j=1}^{n}L(\vartheta_j;\boldsymbol{\beta}) + g_{2n}(r)$$

with

$$\|\sqrt{n}g_{2n}\|^2_{\ell^2_p} = o_P(1).$$

Analogous expansions hold for $\Im\varphi(r;\widehat{\boldsymbol{\beta}})$, so that if we let

$$Z_{0,n}(r) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\Upsilon(r,\vartheta_j;\boldsymbol{\beta}),$$

these expansions imply that

$$Z_n(r) = Z_{0,n}(r) + g_{3n}(r), \tag{17}$$

with

$$\|g_{3n}\|^2_{\ell^2_p} = o_P(1). \tag{18}$$

From Assumptions A and B, it follows that $\mathbb{E}_{\boldsymbol{\beta}}\left\{\|\Upsilon(\cdot,\Theta;\boldsymbol{\beta})\|^2_{\ell^2_p}\right\} < \infty$. Therefore, by applying the Central Limit Theorem in Hilbert spaces (van der Vaart and Wellner, 1996, p. 50), we get that

$$Z_{0,n} \xrightarrow{\mathscr{L}} Z, \tag{19}$$

and then the result follows from (17)–(19). ■

Proof of Theorem 2

Let $Z_n^*(r) = \sqrt{n}\{\varphi_n^*(r) - \varphi(r;\widehat{\boldsymbol{\beta}}^*)\}$, with $\widehat{\varphi}_n^*(r) = n^{-1}\sum_{j=1}^{n} e^{\mathrm{i}r\vartheta_j^*}$. Proceeding as in the proof of Theorem 1, we have that

$$Z_n^*(r) = Z_{0,n}^*(r) + g_n^*(r),$$

with $Z_{0,n}^*(r) = n^{-1/2}\sum_{j=1}^{n}\Upsilon(r,\vartheta_j^*;\widehat{\boldsymbol{\beta}})$,

$$\|g_n^*\|^2_{\ell^2_p} = o_{P_*}(1), \quad a.s.$$

To prove the result we derive the asymptotic distribution of $Z^*_{0,n}(r)$, showing that it coincides with the asymptotic distribution of $C_{n,p}$ when the data come from $F(\cdot;\beta^0)$. Notice that, for each n, the elements in the set $\{\Upsilon(\cdot,\vartheta_1^*;\widehat{\beta}),\ldots,\Upsilon(\cdot,\vartheta_n^*;\widehat{\beta})\}$ are independent and identically distributed random elements taking values in the separable Hilbert space ℓ^2_p, but their common distribution may vary with n. Because of this reason, in order to derive the asymptotic distribution of $Z^*_{0,n}(r)$, we apply Theorem 1.1 in Kundu, Majumdar and Mukherjee (2000). So we will prove that conditions (i)–(iii) in that theorem hold. For $k \geq 0$, let $e_k(j) = I(k=j)/\sqrt{p(k)}$. $\{e_k\}_{k\geq 0}$ is an orthonormal basis of ℓ^2_p.

Let $\mathscr{C}_n$ and $\mathscr{K}_n$ denote the covariance operator and the covariance kernel of $Z^*_{0,n}$, respectively. Let $\mathscr{C}_0$ and $\mathscr{K}_0$ denote the covariance operator and the covariance kernel of Z_0, respectively, where Z_0 stands for the random element figuring in Theorem 1 with $\beta = \beta^0$. Assumptions A* and C imply that

$$\langle \mathscr{C}_n e_k, e_r\rangle_{\ell^2_p} = \sqrt{p(k)p(r)}\mathscr{K}_n(k,r) \to \sqrt{p(k)p(r)}\mathscr{K}_0(k,r) = \langle \mathscr{C}_0 e_k, e_r\rangle_{\ell^2_p}, \quad a.s.,$$

Setting $a_{kr} = \langle \mathscr{C}_0 e_k, e_r\rangle_{\ell^2_p}$ in the aforementioned Theorem 1.1, this proves that condition (i) holds.

Assumptions A*, B* and C imply that

$$\sum_{k\geq 0}\langle \mathscr{C}_n e_k, e_k\rangle_{\ell^2_p} = \sum_{k\geq 0}\mathscr{K}_n(k,k)p(k) \to \sum_{k\geq 0}\mathscr{K}_0(k,k)p(k) = \mathbb{E}\left\{\|Z_0\|^2_{\ell^2_p}\right\} < \infty, \quad a.s.,$$

and thus condition (ii) holds. Finally, condition (iii) readily follows from Assumption A*. ∎

Acknowledgements

The authors thank the anonymous referees and the editor for their constructive comments and suggestions which helped to improve the presentation. MD Jiménez-Gamero has been partially supported by grant MTM2017-89422-P of the Spanish Ministry of Economy, Industry and Competitiveness, the State Agency of Investigation, and the European Regional Development Fund.

References

Bugni, F.A., Hall, P., Horowitz, J.L. and Neumann, G.R. (2009). Goodness-of-fit tests for functional data. *Econometrics Journal*, 12, S1–S18.

Cuparić, M., Milosević, B. and Obradović, M. (2019) New L2-type exponentiality tests. *SORT*, 43, 25–50.

Epps, T.W. and Pulley, L.B. (1983). A test for normality based on the empirical characteristic function. *Biometrika*, 70, 723–726.

Gatto, R. (2008). Some computational aspects of the generalized von Mises distribution. *Statistics and Computing*, 18, 321–331.

Gatto, R. and Jammalamadaka, S.R. (2007). The generalized von Mises distribution. *Statistical Methodology*, 4, 341–353.

Gürtler, N. and Henze, N. (2000). Goodness-of-fit tests for the Cauchy distribution based on the empirical characteristic function. *Annals of the Institute of Statistical Mathematics*, 52, 267–286.

Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics*. World Scientific, Singapore.

Kuiper, N.H. (1960). Tests concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A,* 63, 38–47.

Kundu, S., Majumdar, S. and Mukherjee, K. (2000). Central limit theorems revisited. *Statistics & Probability Letters*, 47, 265–275.

Lockhart, R.A. (2012). Conditional limit laws for goodness-of-fit tests. *Bernoulli*, 18, 857–882.

Lockhart, R.A., O'Reilly, F. and Stephens, M.A. (2007). The use of the Gibbs sampler to obtain conditional tests, with applications. *Biometrika*, 94, 992–998.

Lockhart, R.A., O'Reilly, F. and Stephens, M.A. (2009). Exact conditional tests and approximate bootstrap tests for the von Mises distribution. *Journal of Statistical Theory and Practice*, 3, 543–554.

Lockhart, R.A. and Stephens, M.A. (1985). Tests of fit for the von Mises distribution. *Biometrika*, 72, 647–652.

Lund, U. and Jammalamadaka, S.R. (2000). An entropy–based test for goodness–of–fit of the von Mises distribution. *Journal of Statistical Computation and Simulation*, 67, 319–332.

Matsui, M. and Takemura, A. (2008). Goodness-of-fit tests for symmetric stable distributions–Empirical characteristic function approach. *Test*, 17, 546–566.

Meintanis, S.G. (2005) Consistent tests for symmetric stability with finite mean based on the empirical characteristic function. *Journal of Statistical Planning and Inference*, 128, 373–380.

Meynaoui, A., Mélisande, A., Laurent-Bonneau, B. and Marrel, A. (2019) Adaptive tests of independence based on HSIC measures. https://arxiv.org/abs/1902.06441

Taylor, A. and Burns, K. (2016). Radial distributions of air plants: a comparison between epiphytes and mistletoes. *Ecology*, 97, 819–825.

Tenreiro, C. (2009). On the choice of the smoothing parameter for the BHEP goodness-of-fit test. *Computational Statistics & Data Analysis*, 53, 1038–1053.

Tenreiro, C. (2019). On the automatic selection of the tuning parameter appearing in certain families of goodness-of-fit tests. *Journal of Statistical Computation and Simulation*, 89, 1780–1797.

van der Vaart, A.W. and Wellner, J.A. (1996) *Weak Convergence and Empirical Processes*. Springer, New York, 1996.

Watson, G.S. (1961). Goodness-of-fit tests on the circle. *Biometrika*, 48, 109–114.