Ann. Inst. Statist. Math.
Vol. 54, No. 1, 29–44 (2002)


# A DATA DRIVEN SMOOTH TEST FOR CIRCULAR UNIFORMITY

M. BOGDAN[1], K. BOGDAN[1] AND A. FUTSCHIK[2]

[1] *Institute of Mathematics, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland*
[2] *Institute of Statistics, University of Wien, Universitätsstr. 5/9, A-1010 Vienna, Austria*

(Received December 14, 1999; revised October 16, 2000)

**Abstract.** We propose a new omnibus test for uniformity on the circle. The new test is based upon the idea of data driven smooth tests as presented in Ledwina (1994, *J. Amer. Statist. Assoc.*, **89**, 1000–1005). Our simulations indicate that the test performs very well for multifarious alternatives. In particular, it seems to outperform other known omnibus tests when testing against multimodal alternatives. We also investigate asymptotic properties of our test and we prove that it is consistent against every departure from uniformity.

*Key words and phrases*: Schwarz's BIC criterion, goodness-of-fit, score test, smooth test, Neyman's test, Monte Carlo simulations.

## 1. Introduction

In many fields of science, data occur that are measured in the form of angles or two-dimensional directions. Even events occurring over time can be viewed as angular data when the occurrence intensity is periodic. Statistical methods for handling circular data have been summarized e.g. in Upton and Fingleton (1989), Fisher (1993) and Mardia and Jupp (2000).

A classical question in this field is to test whether all angles/directions occur with the same probability. In the literature several tests for circular uniformity have been proposed. These tests are either omnibus tests that should have reasonable power against any alternative or especially constructed to have high power against some particular alternatives.

The most frequently used goodness of fit test seems to be the Rayleigh test which is the most powerful invariant test for uniformity against von Mises alternatives. (see e.g. Watson and Williams (1956)) However, as noted by Upton and Fingleton (1989), "the Rayleigh test is only sensible when the alternative to uniformity is some unimodal distribution". When no or little knowledge is available concerning the structure of possible alternatives, omnibus tests are usually chosen. Among them the most popular are Kuiper's (Kuiper (1960)) test, which is a rotation invariant analogue of the Kolmogorov-Smirnov test, Watson's (Watson (1961)) $U^2$ test, which is a circular adaptation of the linear Cramér-von-Mises test, and Ajne's test (Ajne (1968)) with its extensions proposed by Rao (1969) and Rothman (1972). Rayleigh, Watson's, Ajne's, Rao's as well as Rothman's tests belong to the general class of tests considered in Beran (1968, 1969). Beran's tests are locally most powerful rotation-invariant tests of uniformity against certain classes of alternatives. Following the terminology of Giné (1975) they are sometimes called "Sobolev" tests.

The classical tests of Kuiper, Watson and Ajne were compared in a simulation study by Stephens (1969). The study shows that while these classical tests are powerful against unimodal alternatives, their power can be prohibitively low against bimodal and four-peaked alternatives. In try to solve this problem Hermans and Rasson (1985) introduced a new class of Sobolev tests. These tests are extensions of Ajne's test. Results of simulations reported in Hermans and Rasson (1985) show that the new tests perform better than classical ones against bimodal and four-peaked alternatives at the price of sacrificing some power against unimodal alternatives. A review of the available tests for uniformity on a circle is given in Mardia and Jupp (2000).

For general, not necessarily circular observations, it is well known that classical goodness-of-fit tests that are based on the empirical distribution function (like the Kolmogorov-Smirnov or Cramér-von-Mises test) have a very low power against multimodal alternatives. A theoretical explanation of this phenomenon is e.g. given in Neuhaus (1976) and Milbrodt and Strasser (1990). As suggested by studies of Locke and Spurrier (1978) and Miller and Quesenberry (1979), a good alternative to those tests is Neyman's (1937) smooth test, which rejects the hypothesis that observations $y_1, \dots, y_n$ come from the uniform distribution on $[0,1]$ for large values of the statistic

$$\mathcal{N}_k = n \sum_{j=1}^{k} \left\{ n^{-1/2} \sum_{i=1}^{n} \phi_j(y_i) \right\}^2 ,$$

where $\phi_1, \phi_2, \dots$ are consecutive Legendre polynomials, orthonormal in $L^2[0,1]$.

Neyman's smooth test with a fixed number $k$ of components in the test statistic is consistent only against certain types of alternatives to uniformity. Trying to adapt this test to detect a wider range of alternatives, some authors proposed data driven procedures for the choice of $k$ (see e.g. Bickel and Ritov (1992), Eubank and LaRiccia (1992), Eubank *et al.* (1993), Ledwina (1994) and Fan (1996)). In particular, Ledwina (1994) introduced a new data driven smooth test for uniformity in which the number of components is chosen by the Schwarz selection rule. Ledwina's test is consistent against every alternative to uniformity (see Kallenberg and Ledwina (1995)) and has nice asymptotic optimality properties (see Inglot and Ledwina (1996) and Inglot *et al.* (1998)). Also, extensive simulation studies show that the test performs very well in comparison with classical tests for uniformity as well as with other, recently proposed, competitors.

In the present paper we follow Ledwina's approach and propose a new data driven smooth test for uniformity on the circle. In comparison with the original proposition of Ledwina (1994) the test is modified so as to obtain invariance of the statistic under rotations of the sample.

The structure of the paper is as follows. In Section 2 we introduce our new test. In Section 3 the asymptotic behavior of the test is investigated and its consistency is proved. We give new results concerning the behavior of the applied selection rule, which are more satisfactory than their analogues found in previous work. In Section 4 we provide approximate critical values. Furthermore the power against various alternatives is investigated in a simulation study. Compared to other popular procedures from the literature our proposed approach seems to be particularly suitable as an omnibus test. Finally, in Section 5 the practicalities of using our test are summarized.

## 2. A smooth test for uniformity on the circle

Let $X_1, \ldots, X_n$ be i.i.d. observations on a circle, i.e. $X_i = (X_{1i}, X_{2i})$ where $X_{1i}^2 + X_{2i}^2 = r^2$ for some constant $r > 0$. Let $\alpha_i \in [0, 2\pi)$ denote the angle between the vector $X_i$ and the vector $(1, 0)$. We want to test the null hypothesis $\mathcal{H}_0$ that $\alpha_1, \ldots, \alpha_n$ are uniformly distributed over the interval $[0, 2\pi)$.

The first step in constructing our smooth test of fit consists in embedding the null density into a larger exponential family. We shall build the exponential family upon the system of trigonometric functions.

Let $\boldsymbol{N}$ denote the set of natural numbers. For each $j \in \boldsymbol{N}$, let $b_{2j-1}(x) = \sqrt{2}\cos(jx)$ and $b_{2j}(x) = \sqrt{2}\sin(jx)$. The functions $b_j$ satisfy the conditions

$$\forall j \in \boldsymbol{N} \frac{1}{2\pi} \int_0^{2\pi} b_j(x)dx = 0, \tag{2.1}$$

$$\forall i, j \in \boldsymbol{N} \frac{1}{2\pi} \int_0^{2\pi} b_j(x) b_i(x) dx = \delta_{ij}, \tag{2.2}$$

where $\delta_{ij}$ is the Kronecker delta.

For $k = 1, 2, \ldots$ we consider the exponential family of densities

$$p_{2k}(x, \theta) = \exp\{\theta \circ b(x) - \psi_{2k}(\theta)\} \frac{1}{2\pi}, \qquad x \in [0, 2\pi], \tag{2.3}$$

where $\theta = (\theta_1, \ldots, \theta_{2k}) \in \boldsymbol{R}^{2k}$, $b(x) = (b_1(x), \ldots, b_{2k}(x))$, and

$$\psi_{2k}(\theta) = \log \int_0^{2\pi} \frac{1}{2\pi} \exp\{\theta \circ b(x)\} dx.$$

The symbol "$\circ$" denotes the inner product in $\boldsymbol{R}^{2k}$.

Observe that testing $\mathcal{H}_0$ within the exponential family (2.3) is equivalent to testing the parametric hypothesis $\mathcal{H} : \theta = 0$ against $\theta \neq 0$. The asymptotically optimal test statistic for this problem is the score statistic

$$N_{2k} = n \sum_{j=1}^{2k} (\bar{b}_j)^2, \tag{2.4}$$

where $\bar{b}_j = \frac{1}{n} \sum_{i=1}^n b_j(\alpha_i)$. By definition, the smooth test rejects the hypothesis $\mathcal{H}_0$ for large values of the statistic $N_{2k}$ (see e.g. Chapter 4.2 of Rayner and Best (1989)).

The following lemma states the rotational invariance of the smooth test.

LEMMA 2.1. *For each fixed $\beta \in \boldsymbol{R}$ and $k \in \boldsymbol{N}$ the smooth test statistic $N_{2k}$ is invariant under the following transformation of the vector of angles*:

$$\tilde{\alpha}_i = (\alpha_i + \beta) \bmod 2\pi, \qquad i = 1, \ldots, n. \tag{2.5}$$

PROOF. For $j = 1, \ldots, k$ we have that

$$\bar{b}_{2j-1}^2 + \bar{b}_{2j}^2 = 2 \left| \frac{1}{n} \sum_{l=1}^n \exp(\mathrm{i} j \alpha_l) \right|^2 = 2 \left| \frac{1}{n} \sum_{l=1}^n \exp(\mathrm{i} j \tilde{\alpha}_l) \right|^2. \qquad \square$$

*Remark* 2.1. For $k = 1$ the smooth test statistic $N_2$ coincides with the statistic used for Rayleigh's test.

*Remark* 2.2. Test statistics of the form $n(\bar{b}_{2j-1}^2 + \bar{b}_{2j}^2)$ were proposed by Beran (1968) for testing uniformity versus a symmetric $j$-modal alternative.

Before applying the smooth test we must decide on the number of components $k$ in the test statistic. However, the properties of the test strongly depend on the choice of $k$. For testing against unimodal alternatives the best solution is usually to take $k = 1$ (that is to use Rayleigh's test), while for testing against multimodal alternatives it is better to choose larger values of $k$. Since it is our aim to construct an omnibus test, we propose a data driven method for the choice of $k$, namely, the simplified version of the Schwarz selection rule considered in Inglot *et al.* (1997), Kallenberg and Ledwina (1997), Inglot (1999) and Inglot and Ledwina (2000).

To this end we note that the likelihood of the independent random variables $\alpha_1, \ldots, \alpha_n$, each having the density (2.3), is equal to $\exp\{n[\sum_{j=1}^{2k} \theta_j \bar{b}_j - \psi_{2k}(\theta)]\}(2\pi)^{-n}$. The Schwarz Bayesian information criterion for choosing the dimension of the exponential family of densities (2.3) recommends that model dimension $2k$ for which $n \sup_{\theta \in \boldsymbol{R}^{2k}} \{\sum_{j=1}^{2k} \theta_j \bar{b}_j - \psi_{2k}(\theta)\} - k \log n$ is maximal (see Schwarz (1978)). An analogous rule was used in Ledwina (1994) to choose the number of components in Neyman's smooth test $\mathcal{N}_k$ based on Legendre polynomials.

Later a simplified version of the rule has been proposed in the literature. The modification reduces the numerical complexity and simplifies asymptotic considerations (see Inglot and Ledwina (1996)). Observe that the likelihood ratio statistic $2n \sup_{\theta \in \boldsymbol{R}^{2k}} \{\sum_{j=1}^{2k} \theta_j \bar{b}_j - \psi_{2k}(\theta)\}$ for testing the parametric hypothesis $\mathcal{H}$ is locally asymptotically equivalent to the score statistic $N_{2k}$.

The simplified version of the Schwarz rule recommends choosing the minimal $k$ for which $L(k) = N_{2k} - 2k \log n$ is maximal. Let

$$S = \inf\{k \in \boldsymbol{N} : L(k) \geq L(m), m \in \boldsymbol{N}\}.$$

(As usual, we put $\inf \emptyset = \infty$.) Then the data driven version of the smooth test for uniformity on the circle proposed in this paper rejects the null hypothesis for large values of the statistic

$$N_{2S} = n \sum_{j=1}^{2S} (\bar{b}_j)^2. \tag{2.6}$$

For completeness we define $N_{2S} = \infty$, if $S = \infty$. It turns out, however, that $S$ will be finite with probability one for all practically relevant simple sizes.

LEMMA 2.2. *Assume that the random variables $\alpha_1, \ldots, \alpha_n$ come from a continuous distribution on $[0, 2\pi)$ and that $n > 2$. Then*

$$P(S = \infty) = 0.$$

PROOF. We fix the sample $(\alpha_1, \ldots, \alpha_n)$ and apply the classical Wiener lemma (Katznelson (1976), p. 42) to the empirical measure $\mu = \frac{1}{n}(\delta_{\alpha_1} + \cdots + \delta_{\alpha_n})$. For its

Fourier coefficients $a_j = \int \exp(-\mathrm{i}jx)\mu(dx) = \frac{1}{n}\sum_{l=1}^{n} \exp(-\mathrm{i}j\alpha_l)$ we obtain

$$\frac{1}{2k+1}\sum_{j=-k}^{k} |a_j|^2 \to n\frac{1}{n^2} = \frac{1}{n} \quad \text{as} \quad k \to \infty,$$

provided $\alpha_1, \ldots, \alpha_n$ are distinct numbers. Since

$$N_{2k} = 2n\sum_{j=1}^{k} |a_j|^2 = n\left(\sum_{j=-k}^{k} |a_j|^2 - 1\right),$$

we have that $N_{2k}/(2k+1) \to 1$ as $k \to \infty$. In particular, $L(k) = N_{2k} - 2k\log n < L(1)$ for large $k$, provided that $\log n > 1$, i.e. $n \geq 3$. Thus $P(S < \infty) = 1$ for $n \geq 3$ under our assumption on the distribution of $\alpha_1, \ldots, \alpha_n$. Note that if $n = 1$ or 2 then $S = \infty$ by a similar reasoning. □

## 3. Asymptotic results

THEOREM 3.1. *Under $\mathcal{H}_0$ it holds that*

$$\lim_{n\to\infty} P(S = 1) = 1. \tag{3.1}$$

PROOF. To prove (3.1) assume that $S = k \geq 2$ and choose $l \in \boldsymbol{N}$ such that $k \in \{2^l, 2^l + 1, \ldots, 2^{l+1} - 1\}$. Now $S = k$ implies that $L(k) > L(1)$, i.e. $N_{2k} - 2k\log n > N_2 - 2\log n$. Since $N_{2k}$ is non-decreasing in $k$ it follows that $N_{2^{l+2}} - 2^{l+1}\log n > N_2 - 2\log n$. Thus for every $l \in \boldsymbol{N}$ we have

$$\begin{aligned} P(S \in \{2^l, 2^l + 1, \ldots, 2^{l+1} - 1\}) &\leq P(N_{2^{l+2}} > 2^l \log n) \\ &= P(N_{2^{l+2}} - 2^{l+2} > 2^l \log(n/e^4)) \\ &\leq \frac{\mathrm{Var}\ N_{2^{l+2}}}{(2^l \log(n/e^4))^2}, \end{aligned}$$

provided $n > e^4$. By a straightforward calculation we obtain that

$$n^2\,\mathrm{Cov}\{(\bar{b}_{2j-1})^2 + (\bar{b}_{2j})^2, (\bar{b}_{2p-1})^2 + (\bar{b}_{2p})^2\} = 4\frac{n-1}{n}\delta_{jp}, \qquad j, p \in \boldsymbol{N}.$$

Thus $\mathrm{Var}\,N_{2m} = 4m\frac{n-1}{n}$, $m \in \boldsymbol{N}$, and

$$P(S \in \{2^l, 2^l + 1, \ldots, 2^{l+1} - 1\}) \leq \frac{n-1}{n\log^2(n/e^4)}2^{-l+3}, \qquad n > e^4. \tag{3.2}$$

We finally obtain $P(S \in \{2, 3, \ldots\}) \leq 8/\log^2(n/e^4) \to 0$, as $n \to \infty$. □

The next proposition follows from standard results on score statistics (see e.g. Serfling (1980)).

PROPOSITION 3.1. *Under the null hypothesis $\mathcal{H}_0$ for every fixed $k \in \boldsymbol{N}$*

$$N_{2k} \xrightarrow{\mathcal{D}} \chi^2_{2k} \quad as \quad n \to \infty, \tag{3.3}$$

*where $\chi^2_{2k}$ is chi-squared distributed with $2k$ degrees of freedom.*

Proposition 3.1 together with (3.1) leads immediately to the asymptotic distribution of the data driven smooth test.

THEOREM 3.2. *Under $\mathcal{H}_0$*

$$N_{2S} \xrightarrow{\mathcal{D}} \chi^2_2, \quad as \quad n \to \infty. \tag{3.4}$$

THEOREM 3.3. *The test based on $N_{2S}$ is consistent against every alternative to uniformity on $[0, 2\pi)$.*

PROOF. Let $\alpha_1, \ldots, \alpha_n$ be i.i.d. rv's distributed according to the probability measure $\mathcal{P}$ on $[0, 2\pi)$. Assume $\mathcal{P}$ is different from the uniform distribution. By the uniqueness of Fourier coefficients there exists a natural $K$ such that

$$a = E_{\mathcal{P}}(\sin K\alpha_1) \neq 0 \quad \text{or} \quad b = E_{\mathcal{P}}(\cos K\alpha_1) \neq 0. \tag{3.5}$$

We shall verify that

$$\lim_{n\to\infty} \mathcal{P}(S \geq K) = 1. \tag{3.6}$$

For $K = 1$ (3.6) holds trivially, so we assume that $K > 1$ and we let $j \in \{1, \ldots, K-1\}$. We have that

$$\begin{aligned} \mathcal{P}(S = j) &\leq \mathcal{P}(N_{2j} - 2j \log n \geq N_{2K} - 2K \log n) \\ &\leq \mathcal{P}(n(\bar{b}^2_{2K-1} + \bar{b}^2_{2K}) \leq 2(K-j) \log n) \to 0 \quad \text{as} \quad n \to \infty, \end{aligned}$$

because by (3.5) and the law of large numbers $\bar{b}^2_{2K-1} + \bar{b}^2_{2K} \to a^2 + b^2 > 0$ almost surely as $n \to \infty$. Thus $\lim_{n\to\infty} \mathcal{P}(S = j) = 0$ and (3.6) easily follows. Furthermore, for every constant $c$ it holds that $\mathcal{P}(N_{2K} > c) \geq \mathcal{P}(n(\bar{b}^2_{2K-1} + \bar{b}^2_{2K}) > c) \to 1$ as $n \to \infty$. Together with Theorem 3.2 this yields the consistency of the test. □

## 4. Simulation study

In this section we investigate the empirical behavior of the selection rule $S$ under uniformity and obtain approximate critical values. Then we compare the empirical power of the data driven smooth test with that of a range of other tests for uniformity on the circle. We include Rayleigh's test $R$, Kuiper's test $K$ and Watson's test $U^2$. We also provide simulation results for the test based on the statistic

$$T_{n,\infty} = \frac{n}{\pi} - \frac{1}{2n} \sum_{i,j=1}^{n} |\sin(\alpha_i - \alpha_j)|.$$

This test was introduced by Hermans and Rasson (1985) and is supposed to perform much better than classical tests for multimodal alternatives. Results for Ajne's test (Ajne (1968)) have not been included, since its performance is similar to that of Kuiper's and Watson's test (see e.g. Stephens (1969)).

Table 1. Counts of $\{S = s\}$ under $\mathcal{H}_0$, based on $10^6$ samples in each case.

| | $n = 30$ | $n = 50$ | $n = 100$ |
|---|---|---|---|
| 1 | 964086 | 979334 | 989669 |
| 2 | 30821 | 18739 | 9771 |
| 3 | 4163 | 1686 | 516 |
| 4 | 704 | 211 | 41 |
| 5 | 163 | 25 | 3 |
| 6 | 39 | 4 | 0 |
| 7 | 21 | 1 | 0 |
| 8 | 2 | 0 | 0 |
| 9 | 1 | 0 | 0 |
| 10 | 0 | 0 | 0 |

Table 2. Empirical critical values of $N_{2S}$ based on $10^6$ samples in each case.

| $\alpha$ | $n = 30$ | $n = 50$ | $n = 100$ | asymptotic |
|---|---|---|---|---|
| 0.01 | 13.56 | 12.22 | 11.53 | 9.21 |
| 0.05 | 7.88 | 6.90 | 6.39 | 5.99 |
| 0.1 | 5.38 | 5.00 | 4.79 | 4.61 |

### 4.1 *Empirical distribution of the selection rule under uniformity*

In Table 1 we present the empirical distribution of the selection rule $S$ under uniformity for three different sample sizes $n$. The results included in Table 1 are in accordance with the theoretical result (3.1) and indicate that the rule $S$ concentrates on 1 already for relatively small sample sizes. Furthermore, in our simulations we observed that the rule $S$ never took on values larger than 10 under the null hypothesis and for sample sizes $n \geq 30$. Therefore it seems appropriate to compute $S$ as $S = \text{argmax}_{1 \leq k \leq 10} L(k)$ in practice. This modification would not influence our empirical critical values. It also has negligible influence on the data driven smooth test's power for the alternatives considered in the simulations below.

### 4.2 *Critical values*

The distribution of the test statistic $N_{2S}$ under $\mathcal{H}_0$ converges to the chi-squared distribution relatively slowly. In Table 2 we give some empirical critical values for the data driven smooth test based upon $10^6$ samples.

A simple method to obtain approximate critical values is to fit a polynomial in inverse powers of $\sqrt{n}$ to the simulated critical values. For several goodness of fit tests, this approach has been used successfully by Stephens (1970). In our context and for $n$ between 30 and 100, it turns out that the asymptotic (chi-squared) critical values are fairly accurate when used with the modified statistic

$$\left(N_{2S} + \frac{23.8}{\sqrt{n}} - \frac{43.4}{n}\right)\left(1 + \frac{4.3}{\sqrt{n}}\right)^{-1}.$$

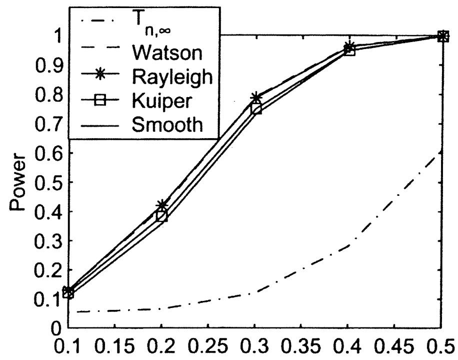


Fig. 1. Power functions for the wrapped Cauchy distribution with parameter $\rho$, $0.1 \leq \rho \leq 0.5$, $n = 50$, $\alpha = 0.05$.

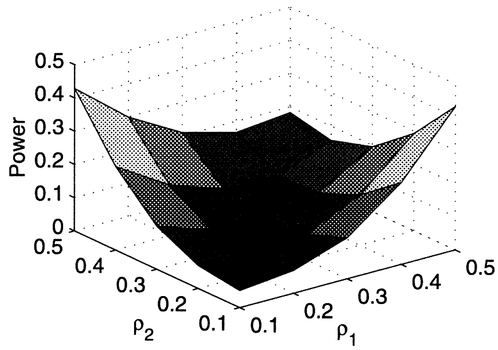


Fig. 2. Power of Kuiper test against $h_1$ as the function of parameters $\rho_1$ and $\rho_2$ $(n = 50, \alpha = 0.05)$.

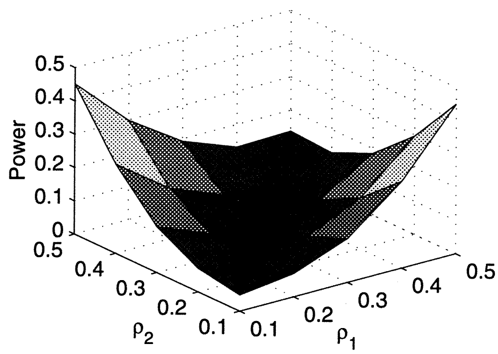


Fig. 3. Power of Watson test against $h_1$ as the function of parameters $\rho_1$ and $\rho_2$ $(n = 50, \alpha = 0.05)$.

## 4.3 *Power simulations*

In order to investigate the performance of the data driven smooth test we carried out simulations under the following alternatives.

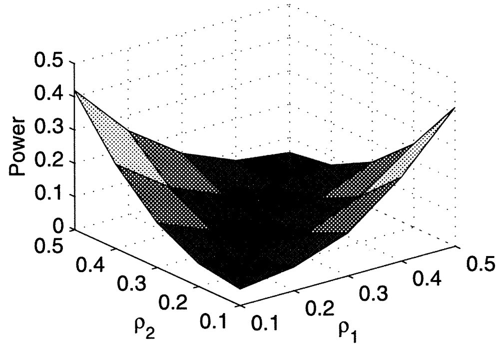


Fig. 4. Power of Rayleigh test against $h_1$ as the function of parameters $\rho_1$ and $\rho_2$ ($n = 50, \alpha = 0.05$).

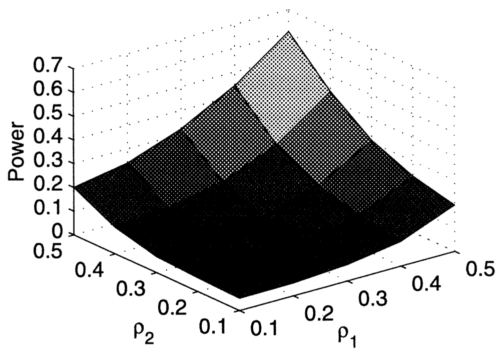


Fig. 5. Power of $T_{n,\infty}$ against $h_1$ as the function of parameters $\rho_1$ and $\rho_2$ ($n = 50, \alpha = 0.05$).

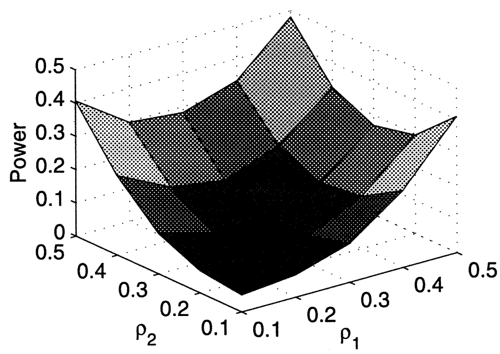


Fig. 6. Power of $N_{2S}$ against $h_1$ as the function of parameters $\rho_1$ and $\rho_2$ ($n = 50, \alpha = 0.05$).

(A) Wrapped Cauchy distribution with the density function

$$f(x; \rho, \mu) = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(x - \mu)}, \qquad x \in [0, 2\pi),$$

Table 3. Notation for alternative distributions.

| Notation | k | $\lambda_1,\ldots,\lambda_k$ | $\rho_1,\ldots,\rho_k$ | $\mu_1,\ldots,\mu_k$ |
|---|---|---|---|---|
| M(1) | 1 | 1 | 0.33 | 0 |
| M(2) | 2 | 0.5,0.5 | 0.42,0.42 | $0,\frac{\pi}{2}$ |
| M(3) | 2 | 0.25,0.75 | 0.42,0.42 | $0,\frac{\pi}{2}$ |
| M(4) | 2 | 0.5,0.5 | 0.6,0.6 | $0,\pi$ |
| M(5) | 2 | 0.5,0.5 | 0.23,0.69 | $0,\pi$ |
| M(6) | 2 | 0.25,0.75 | 0.54,0.54 | $0,\pi$ |
| M(7) | 3 | 0.5,0.2,0.3 | 0.75,0.75,0.75 | $0,\frac{2}{3}\pi,\frac{4}{3}\pi$ |
| M(8) | 3 | $\frac{1}{3},\frac{1}{3},\frac{1}{3}$ | 0.75,0.75,0.75 | $0,\frac{2}{3}\pi,\frac{4}{3}\pi$ |
| M(9) | 4 | 0.4,0.2,0.25,0.15 | 0.69,0.575,0.69,0.575 | $0,\frac{\pi}{4},\pi,\frac{7}{4}\pi$ |
| M(10) | 4 | $\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}$ | 0.84,0.84, 0.84,0.84 | $0,\frac{\pi}{2},\pi,\frac{3}{2}\pi$ |

where $\rho \in [0,1)$ and $\mu \in [0,2\pi)$.

(B) Mixtures of wrapped Cauchy distributions with densities given by $h(x;k,\lambda_1,\ldots,\lambda_k,\mu_1,\ldots,\mu_k,\rho_1,\ldots,\rho_k) = \sum_{i=1}^{k}\lambda_i f(x;\rho_i,\mu_i)$, where $\lambda_i \in [0,1]$ with $\sum_{i=1}^{k}\lambda_i = 1$.

(C) Distributions with density functions of the form $g_j(x) = 1+\rho\cos(jx)$, $x \in [0,2\pi)$. These alternatives have been considered for several parameter values of $\rho \in [-1,1]$ and for $1 \le j \le 6$.

Below we report the results of our simulations for the sample size $n = 50$ and significance level $\alpha = 0.05$.

Figure 1 shows the power functions of the considered tests against the wrapped Cauchy distribution. It shows that the power of the data driven smooth test against the wrapped Cauchy distribution is slightly lower than the power of the classical tests of Kuiper, Watson and Rayleigh and much higher than the power of the Hermans and Rasson test $T_{n,\infty}$.

Figures 2–6 show simulated power functions of the tests against a bimodal mixture of wrapped Cauchy distributions with the density function $h_1(x,\rho_1,\rho_2) = h(x;2,0.5,0.5,0,\pi,\rho_1,\rho_2)$.

In Fig. 2, 3 and 4 it can be observed that the classical tests perform very well if one of the parameters $\rho$ is large (close to 0.5) while the other one is small (close to 0.1), which corresponds to essentially unimodal alternatives. On the other hand the power of classical tests is very low for alternatives for which $\rho_1 = \rho_2$, which corresponds to symmetric bimodal alternatives.

On the contrary the test based on $T_{n,\infty}$ has a relatively low power if only one parameter is large ($\rho_1 = 0.5, \rho_2 = 0.1$) and performs very well against bimodal alternatives ($\rho_1 = \rho_2$).

The data driven smooth test $N_{2S}$ provides a good compromise between the classical tests and $T_{n,\infty}$. It is only slightly worse than classical tests for unimodal alternatives ($\rho_1 = 0.5, \rho_2 = 0.1$) and its power against bimodal alternatives is comparable to that of $T_{n,\infty}$.

In Table 3 we describe some further alternatives obtained as mixtures of wrapped Cauchy distributions. Simulated powers of the tests against these alternatives are given in Table 4. The values of parameters for the alternative distributions were chosen to obtain moderate power values.

In Table 5 we present the distribution of the selection rule $S$ under some of the

Table 4. Estimated power (in%) of $N_{2S}$, $K$, $U^2$, $R$ and $T_{n,\infty}$, based on 5000 samples in each case.

| alternatives | $N_{2S}$ | K | $U^2$ | R | $T_{n,\infty}$ |
|---|---|---|---|---|---|
| M(1) | 76 | 83 | 86 | 86 | 16 |
| M(2) | 66 | 75 | 77 | 79 | 59 |
| M(3) | 78 | 84 | 87 | 88 | 12 |
| M(4) | 85 | 39 | 31 | 6 | 92 |
| M(5) | 67 | 65 | 63 | 49 | 66 |
| M(6) | 79 | 78 | 77 | 66 | 75 |
| M(7) | 86 | 71 | 63 | 41 | 29 |
| M(8) | 76 | 26 | 15 | 5 | 6 |
| M(9) | 81 | 76 | 74 | 58 | 80 |
| M(10) | 76 | 25 | 12 | 5 | 62 |

Table 5. Counts of $\{S = s\}$ based on 5000 samples in each case.

| | alternatives | | | |
|---|---|---|---|---|
| $s$ | M(1) | M(4) | M(8) | M(10) |
| 1 | 4583 | 756 | 1128 | 1230 |
| 2 | 378 | 4082 | 1 | 5 |
| 3 | 32 | 86 | 3640 | 1 |
| 4 | 6 | 69 | 81 | 3629 |
| 5 | 1 | 5 | 4 | 65 |
| 6 | 0 | 2 | 41 | 11 |
| 7 | 0 | 0 | 4 | 4 |
| 8 | 0 | 0 | 0 | 44 |
| 9 | 0 | 0 | 1 | 11 |
| 10 | 0 | 0 | 0 | 0 |

considered alternatives.

The results reported in Table 4 show that again the classical tests perform very well for alternatives which are essentially unimodal. Instead their power for symmetric alternatives with more than one peak is prohibitively low (see M(4), M(8) and M(10)).

The test $T_{n,\infty}$ performs very well against a symmetric bimodal alternative M(4). However in all other cases its power is worse than the power of the data driven smooth test. The power of $T_{n,\infty}$ is particularly low for the unimodal alternative M(1) and the alternative with three evenly distributed peaks, M(8).

In Table 4 it can be observed that the data driven smooth test performs very well for a wide range of alternatives. It is much better than classical tests for symmetric alternatives with more than one peak. In other cases the power of the data driven smooth test is comparable to the power of classical tests.

Results reported in Table 5 show that the rule $S$ adapts very well to the data and chooses the number of components in the test statistic which is appropriate for detecting a particular type of alternative. For this reason the data driven smooth test performs so well against many different types of alternatives.

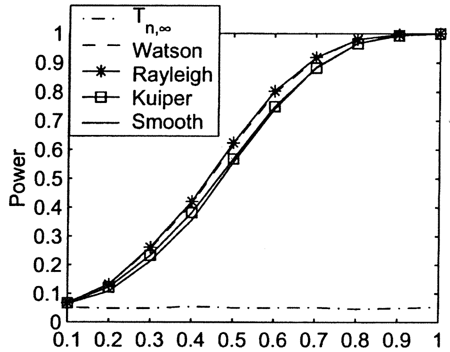


Fig. 7. Power functions for the alternative $g_1$ with parameter $\rho$, $0.1 \le \rho \le 1$, $n = 50$, $\alpha = 0.05$.

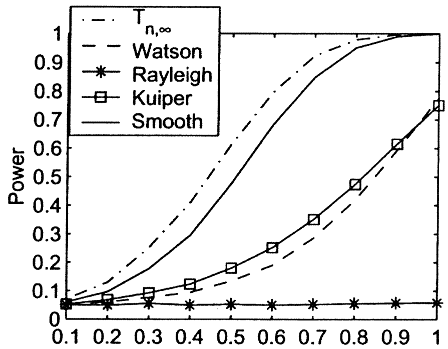


Fig. 8. Power functions for the alternative $g_2$ with parameter $\rho$, $0.1 \le \rho \le 1$, $n = 50$, $\alpha = 0.05$.

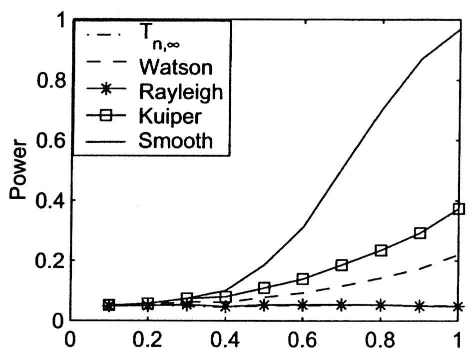


Fig. 9. Power functions for the alternative $g_3$ with parameter $\rho$, $0.1 \le \rho \le 1$, $n = 50$, $\alpha = 0.05$.

Figures 7–12 show the simulated power functions of the considered tests against the alternative distributions $g_j$ of type (C), $1 \le j \le 6$. Note that the alternative $g_j$ has $j$ evenly distributed peaks.

In Figs. 13 and 14 we display a sample from the uniform distribution as well as

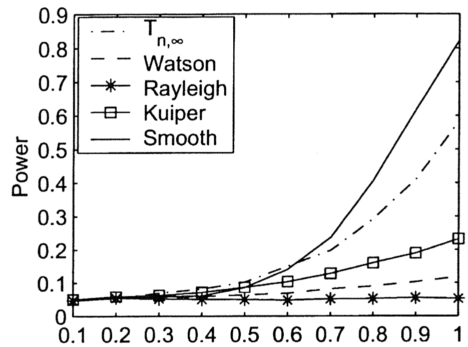


Fig. 10. Power functions for the alternative $g_4$ with parameter $\rho$, $0.1 \le \rho \le 1$, $n = 50$, $\alpha = 0.05$.

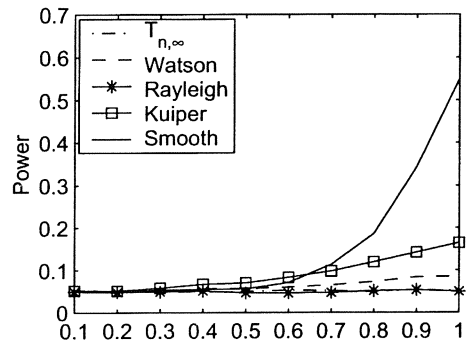


Fig. 11. Power functions for the alternative $g_5$ with parameter $\rho$, $0.1 \le \rho \le 1$, $n = 50$, $\alpha = 0.05$.

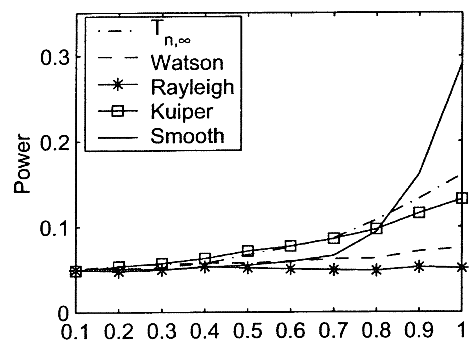


Fig. 12. Power functions for the alternative $g_6$ with parameter $\rho$, $0.1 \le \rho \le 1$, $n = 50$, $\alpha = 0.05$.

from the alternative $g_3$ ($\rho = 0.6$), an alternative for which the smooth test performs particularly well. Although the samples look quite similar, our smooth test decides correctly for both datasets at level $\alpha = 0.05$.

The results of our simulations indicate that the data driven smooth test based on

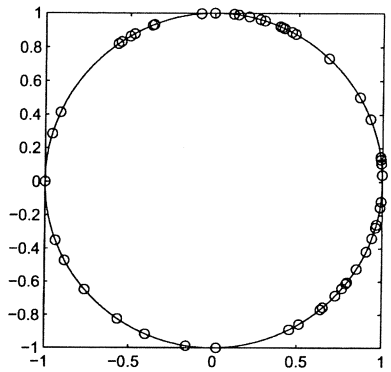

Fig. 13. A sample (size 50, $\rho = 0.6$) from the distribution $g_3$.

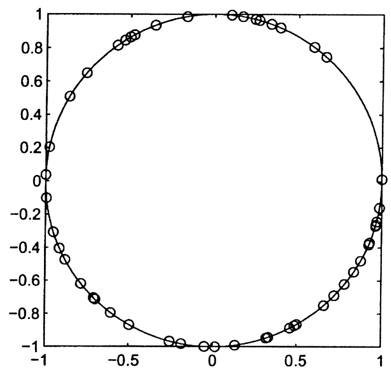

Fig. 14. A sample (size 50) from the uniform distribution.

$N_{2S}$ provides a good omnibus test for uniformity on the circle. For unimodal alternatives it is only slightly worse than the classical tests of Kuiper, Watson and Rayleigh and instead offers much higher power for symmetric alternatives with more than one peak. It also performs much better than the test $T_{n,\infty}$ by Hermans and Rasson which has been designed for multimodal alternatives. Note that the test $T_{n,\infty}$ fails completely for alternatives with an odd number of evenly distributed peaks (see the alternative M(8) and Fig. 9 and 11).

## 5. Practical remarks

Our proposed testing procedure may be summarized as follows.

1. Compute $L(k) = N_{2k} - 2k \log n$ (see (2.4)) for $1 \leq k < K$. In theory $K = \infty$, but simulations suggest that in practise it is sufficient to take $K = 10$ for $n \leq 100$.

2. Compute $S = \operatorname{argmax}_{1 \leq k < K} L(k)$.

3. Reject the hypothesis of uniformity for large values of the statistic $N_{2S} = n \sum_{j=1}^{2S} (\bar{b}_j)^2$.

Since the convergence of the distribution of $N_{2S}$ under $\mathcal{H}_0$ to the chi-square distribution with 2 degrees of freedom is rather slow, simulated critical values should be used, unless $n$ is very large. Approximate critical values may be obtained from Table 2. (Observe that the critical values are decreasing in $n$.)

## Acknowledgements

The authors wish to thank Dr. Marcin Kotulski, Jacek Sachanbiński and Piotr Szajowski for performing simulations.

## References

Ajne, B. (1968). A simple test for uniformity of a circular distribution, *Biometrika*, **55**, 343–354.

Beran, R. J. (1968). Testing for uniformity on a compact homogeneous space, *J. Appl. Probab.*, **5**, 177–195.

Beran, R. J. (1969). Asymptotic theory of a class of tests for uniformity of a circular distribution, *Ann. Math. Statist.*, **40**, 1196–1206.

Bickel, P. J. and Ritov, Y. (1992). Testing for goodness of fit: A new approach, *Nonparametric Statistics and Related Topics* (ed. A.K.Md.E. Saleh), 51–57, North Holland, Amsterdam.

Eubank, R. L. and LaRiccia, V. N. (1992). Asymptotic comparison of Cramér von Mises and nonparametric function estimation techniques for testing goodness of fit, *Ann. Statist.*, **20**, 2071–2086.

Eubank, R. L., Hart, J. D. and LaRiccia, V. N. (1993). Testing goodness of fit via nonparametric function estimation techniques, *Comm. Statist. Theory Methods*, **22**, 3327–3354.

Fan, J. (1996). Tests of significance based on wavelet tresholding and Neyman's truncation, *J. Amer. Statist. Assoc.*, **91**, 674–688.

Fisher, N. I. (1993). *Statistical Analysis of Circular Data*, Cambridge University Press, Cambridge, Massachusetts.

Giné, M. E. (1975). Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms, *Ann. Statist.*, **3**, 1243–1266.

Hermans, M. and Rasson, J. P. (1985). A new Sobolev test for uniformity on the circle, *Biometrika*, **72**, 698–702.

Inglot, T. (1999). Generalized intermediate efficiency of goodness of fit tests, *Math. Methods Statist.*, **8**, 487–509.

Inglot, T. and Ledwina, T. (1996). Asymptotic optimality of data driven Neyman's test, *Ann. Statist.*, **24**, 1982–2019.

Inglot, T. and Ledwina, T. (2001). Intermediate approach to comparison of some goodness-of-fit tests, *Ann. Inst. Statist. Math.*, **53**(4), 810–834.

Inglot, T., Kallenberg, W. C. M. and Ledwina, T. (1997). Data driven smooth tests for composite hypotheses, *Ann. Statist.*, **25**, 1222–1250.

Inglot, T., Kallenberg, W. C. M. and Ledwina, T. (1998). Vanishing shortcoming of data driven Neyman's test, *Asymptotic Methods in Probability and Statistics*, 811–829, Elsevier, Amsterdam.

Kallenberg, W. C. M. and Ledwina, T. (1995). Consistency and Monte Carlo simulation of a data driven version of smooth goodness-of-fit tests, *Ann. Statist.*, **23**, 1594–1608.

Kallenberg, W. C. M. and Ledwina, T. (1997). Data driven smooth test when the hypothesis is composite, *J. Amer. Statist. Assoc.*, **92**, , 1094–1104.

Kallenberg, W. C. M. and Ledwina, T. (1997). Data driven smooth tests for composite hypothesis: Comparison of powers, *J. Statist. Comput. Simulation*, **59**, 101–121.

Katznelson, Y. (1976). *An Introduction to Harmonic Analysis*, Dover, New York.

Kuiper, N. H. (1960). Tests concerning random points on a circle, *Proc. Konink. Neder. Akad. Wetensch., Ser. A*, **63**, 38–47.

Ledwina, T. (1994). Data driven version of Neyman's smooth test of fit, *J. Amer. Statist. Assoc.*, **89**, 1000–1005.

Locke, C. and Spurrier, J. D. (1978). On tests of uniformity, *Comm. Statist. Theory Methods*, **7**, 241–258.

Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*, Wiley, New York.

Milbrodt, H. and Strasser, H. (1990). On the asymptotic power of the two-sided Kolmogorov-Smirnov test, *J. Statist. Plann. Inference*, **26**, 1–23.

Miller, F. L. and Quesenberry, C. P. (1979). Power studies of tests for uniformity II, *Comm. Statist. Simulation Comput.*, **8**, 271–290.

Neuhaus, G. (1976). Asymptotic power properties of the Cramér-von-Mises test under contiguous alternatives, *J. Multivariate Anal.*, **6**, 95–110.

Neyman, J. (1937). 'Smooth test' for goodness of fit, *Skandinavisk Aktuarietidskrift*, **20**, 149–199.

Rao, J. S. (1969). Some contributions to the analysis of circular data, Ph.D. thesis, Indian Statistical Institute, Calcutta.

Rayner, J. C. W. and Best, D. J. (1989). *Smooth Tests of Goodness of Fit*, Oxford University Press, New York.

Rothman, E. D. (1972). Tests for uniformity of a circular distribution, *Sankhyā Ser. A*, **34**, 22–32.

Schwarz, G. (1978). Estimating the dimension of a model, *Ann. Statist.*, **6**, 461–464.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistic*, Wiley, New York.

Stephens, M. A. (1969). A goodness-of-fit statistic for the circle, with some comparisons, *Biometrika*, **56**, 161–168.

Stephens, M. A. (1970). Use of the Kolmogorov-Smirnov, Cramér-von Mises and related statistics without extensive tables, *J. Roy. Statist. Soc. Ser. B*, **32**, 115–122.

Upton, G. J. G. and Fingleton, B. (1989). *Categorical and Directional Data*, Spatial Data Analysis by Example, Vol. 2, Wiley, New York.

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

Watson, G. S. and Williams, E. J. (1956). On the construction of significance tests on the circle and on the sphere, *Biometrika*, **43**, 344–352.