Communications in Statistics - Theory and Methods

On the Gini Mean Difference Test for Circular Data

David D. Tung [a] & S. Rao Jammalamadaka [b]

[a] ALEKS Corporation, Irvine, California, USA

[b] Department of Statistics and Applied Probability, University of California, Santa Barbara, Santa Barbara, California, USA

Version of record first published: 25 Apr 2013.

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis
Taylor & Francis Group

On the Gini Mean Difference Test for Circular Data

DAVID D. TUNG[1] AND S. RAO JAMMALAMADAKA[2]

[1]ALEKS Corporation, Irvine, California, USA
[2]Department of Statistics and Applied Probability,
University of California, Santa Barbara, Santa Barbara, California, USA

We introduce a new test of isotropy or uniformity on the circle, based on the Gini mean difference of the sample arc-lengths and obtain both the exact and asymptotic distributions under the null hypothesis of circular uniformity. We also provide a table of upper percentile values of the exact distribution for small to moderate sample sizes. Illustrative examples in circular data analysis are also given. It is shown that a "generalized" Gini mean difference test has better asymptotic efficiency than a corresponding "generalized" Rao's test in the sense of Pitman asymptotic relative efficiency.

1. Introduction

In this article, we introduce a new test of uniformity on the circle based on the Gini mean difference of the sample arc-lengths. This test extends the Gini mean difference spacings test on the real line in Jammalamadaka and Goria (2004) to the circular case. These sample arc-lengths, which are the gaps between successive observations on the circumference of the circle, are analogous to sample spacings on the real line and provide a maximal-invariant under rotations so that all invariant tests have to be based on them. The Gini mean difference compares these arc-lengths between themselves and is very similar to the Rao's spacings test which has been used to test the uniformity of circular data, that compares the gaps to their expected length.

Observations representing directions in two dimensions can be modeled as random variables taking values on the circumference of the circle. We take this circle to be the circle with unit radius, and hence a circumference of length 2π. A circular probability distribution is one whose support is this circumference.

The simple goodness-of-fit problem on the circle consists of testing fit to a single fixed circular distribution for a given data set. In particular, consider a random

Address correspondence to David D. Tung, ALEKS Corporation, Irvine, CA 92618, USA; E-mail: david.deming.tung@gmail.com

sample of angular measurements $\theta_1, \theta_2, \ldots, \theta_n$ with circular distribution function F defined on the real line with the property that $F(x + 2\pi) - F(x) = 1$, for all $x \in \mathbb{R}$. We are interested in testing the null hypothesis

$$H_0 : F = F_0,$$

where F_0 is a completely specified distribution function.

Without loss of generality, if F is assumed to be continuous as we shall do, by way of the probability integral transform, the goodness-of-fit problem reduces to one of testing circular uniformity, i.e., testing the null hypothesis

$$H_0 : F(\theta) = \frac{\theta}{2\pi}, \quad \text{for } 0 \leq \theta < 2\pi,$$

Let $0 \leq \theta_{(1)} \leq \theta_{(2)} \leq \cdots \leq \theta_{(n)} < 2\pi$ denote the sample order statistics. The *sample arc-lengths* are defined by the random variables

$$D_k = \theta_{(k)} - \theta_{(k-1)}, \quad \text{for } k = 1, 2, \ldots, n, \tag{1.1}$$

where we take $\theta_{(0)} = \theta_{(n)} - 2\pi$ to make D_1 the natural gap between the first and last order statistics that straddle the origin. The sample arc-lengths $\{D_k\}$ represent the differences between successive observations on the circumference of the circle, and remain invariant under the choice of zero-direction or sense of rotation. Tests based on these sample arc-lengths are the focus here for testing the null hypothesis.

Under the null hypothesis of circular uniformity, the joint distribution of $\left(\frac{D_1}{2\pi}, \frac{D_2}{2\pi}, \ldots, \frac{D_n}{2\pi}\right)$ is a Dirichlet $(\alpha_1 = 1, \ldots, \alpha_{n-1} = 1; \alpha_n = 1)$ distribution on the unit $(n-1)$-simplex

$$\mathscr{S}_{n-1} = \left\{ (t_1, t_2, \ldots, t_{n-1}) \in \mathbb{R}^{n-1} : t_k \geq 0, k = 1, 2, \ldots, n-1, \sum_{k=1}^{n-1} t_k \leq 1 \right\}.$$

By a multivariate transformation, the sample arc-lengths $(D_1, D_2, \ldots, D_n)$ have probability density function

$$f_{D_1, D_2, \ldots, D_{n-1}}(d_1, d_2, \ldots, d_{n-1}) = \frac{(n-1)!}{(2\pi)^{n-1}} \cdot I\left(\bigcap_{k=1}^{n-1} (0 \leq d_k \leq 2\pi), \sum_{k=1}^{n-1} d_k \leq 2\pi \right). \tag{1.2}$$

Moreover, under the null hypothesis, these sample arc-lengths are exchangeable random variables and have the same distribution as the spacings from a random sample of $(n - 1)$ random variables from the Uniform distribution on the line segment $[0, 2\pi)$. This suggests that spacings tests on the real line, with some minor modifications, can be used for circular statistical inference. In fact, spacings tests are the only general class of goodness-of-fit tests that are directly applicable to both circular and linear data.

Most common among spacings tests are symmetric spacings tests, i.e., general test statistics of the form

$$V_n(g) = \frac{1}{n} \sum_{k=1}^{n} g(nD_k), \tag{1.3}$$

and

$$W_n(h) = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} h(nD_i, nD_j), \qquad (1.4)$$

where $g(\cdot)$ is a real-valued function satisfying some regularity conditions and $h : [0, \infty) \times [0, \infty) \to \mathbb{R}$ is a symmetric function satisfying some other regularity conditions. Test statistics of the form $V_n(g)$ are symmetric sum-functions of the sample spacings (e.g., cf. Pyke, 1965; Sethuraman and Rao, 1970; Rao and Sethuraman, 1975), and those of the form $W_n(h)$ are U-statistics of the sample spacings (cf. Tung and Jammalamadaka, 2012). Moreover, as these articles show, these symmetric spacings tests are known to have asymptotic Normal distributions under mild conditions.

Among spacings tests of the form $V_n(g)$, Rao's spacings test (cf. Rao, 1969, 1976) given by

$$J_n = \frac{1}{n} \sum_{k=1}^{n} \frac{|nD_k - 2\pi|}{2} = \frac{1}{2} \sum_{k=1}^{n} \left| D_k - \frac{2\pi}{n} \right| = \sum_{k=1}^{n} \left(D_k - \frac{2\pi}{n} \right)_+, \qquad (1.5)$$

is one of the more important tests and corresponds to taking $g(t) = |t - 2\pi|/2$. Large values of J_n indicate clustering of sample observations or evidence for directionality, and rejection of the null hypothesis of circular uniformity. Rao's test is a powerful statistic that can discriminate between uniform (isotropic) and concentrated (anisotropic) circular distributions, regardless of whether the distributions are unimodal or multimodal.

Under the null hypothesis of circular uniformity, the probability density function of J_n is

$$f_{J_n}(u) = \sum_{k=1}^{n-1} \binom{n}{k} \left(\frac{u}{2\pi} \right)^{n-k-1} \frac{\psi_k(nu) \cdot (n-1)! \cdot I[0 \le u \le 2\pi(1 - 1/n)]}{n^{k-1}(n-k-1)!}, \qquad (1.6)$$

where

$$\psi_k(x) = \frac{1}{2\pi(k-1)!} \sum_{j=0}^{\infty} (-1)^j \binom{k}{j} \left(\frac{x}{2\pi} - j \right)_+^{k-1}. \qquad (1.7)$$

Rao's test is one of the few spacings-type statistics for which both the exact and asymptotic distributions are known. A table of upper percentiles of the exact distribution for J_n was first given in Rao (1976), and extended tables of these critical values can be found in Russell and Levitin (1995). On the other hand, for almost all spacings tests, saddlepoint approximations to the null distribution, which give practically exact values, are available and have been studied in Gatto (2001) and Gatto and Jammalamadaka (1999).

Under the null hypothesis, J_n has an asymptotic Normal distribution, i.e. in the limit as $n \to \infty$,

$$\sqrt{n}\left(J_n - e^{-1} \right) = \sqrt{n} \left(\frac{1}{2} \sum_{k=1}^{n} \left| D_k - \frac{2\pi}{n} \right| - e^{-1} \right)$$

$$\xrightarrow{D} N_1 \left(0, 2e^{-1} - 5e^{-2} \right). \qquad (1.8)$$

We introduce the Gini mean difference arc-lengths test in the next section and obtain both its exact and asymptotic distributions under the null hypothesis. We also furnish a table of upper percentile values of the exact distribution. Section 3 contains examples of circular data analysis featuring Rao's test and the Gini mean difference test. Section 4 discusses the Pitman asymptotic relative efficiencies of a generalized Rao's test, and a generalized Gini mean difference test.

2. The Gini Mean Difference Arc-Lengths Test

Comparable to Rao's arc-lengths test is the Gini mean difference of the sample arc-lengths, i.e.,

$$G_n = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \frac{\left| nD_i - nD_j \right|}{2} = \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \left| nD_i - nD_j \right|, \qquad (2.1)$$

which corresponds to taking $h(u, v) = |u - v|/2$ in (1.4), and may also be rewritten as

$$G_n = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(nD_i - nD_j \right)_+. \qquad (2.2)$$

The statistic G_n, which compares these sample arc-lengths between themselves, is of the form $W_n(h)$ and an average over all pairs of absolute pairwise differences of the sample arc-lengths. The Gini mean difference spacings test was first proposed in Jammalamadaka and Goria (2004) for testing goodness-of-fit on the real line. There, under the goodness-of-fit null hypothesis (i.e., linear uniformity on [0, 1]), they derive both the exact and asymptotic distributions, and show that it has good performance based on Monte Carlo powers.

Under the null hypothesis of circular uniformity, the sample arc-lengths between successive observations should be approximately evenly spaced, about $(2\pi)/n$ apart, and G_n should be close to zero. Large values of G_n resulting from unusually large arc-lengths or unusually short arc-lengths between observations are evidence for directionality, and rejection of the null hypothesis of circular uniformity.

Here, we will adapt both the exact and asymptotic null distributions for the Gini mean difference spacings test on the real line to the case of the unit circle with circumference of length 2π.

Let $U_1, U_2, \ldots, U_{n-1}$ be independent Uniform([0, 1]) random variables, and let $\{X_k\} = \{(2\pi)U_k\}$ define $(n-1)$ independent Uniform([0, 2π)) random variables. We define the *uniform spacings* on the unit interval [0, 1] by the random variables

$$T_k = U_{(k)} - U_{(k-1)}, \quad \text{for } k = 1, 2, \ldots, n \qquad (2.3)$$

where $0 \equiv U_{(0)} \le U_{(1)} \le U_{(2)} \le \cdots \le U_{(n-1)} \le U_{(n)} \equiv 1$.

Under the null hypothesis, the sample arc-lengths $\{D_k\}$ are related to the uniform spacings $\{T_k\}$ by the relation

$$D_k \simeq X_{(k)} - X_{(k-1)} = (2\pi)[U_{(k)} - U_{(k-1)}] = (2\pi)T_k. \qquad (2.4)$$

Here, as elsewhere, we use $\simeq$ to denote the distributional equivalence of quantities on the left and right hand sides of the symbol. Since

$$\sum_{i=1}^{n}\sum_{j=1}^{n}|T_i - T_j| \simeq 2\sum_{k=1}^{n-1} U_k, \qquad (2.5)$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}|D_i - D_j| \simeq (2\pi)\sum_{i=1}^{n}\sum_{j=1}^{n}|T_i - T_j| \simeq 2\sum_{k=1}^{n-1} X_k. \qquad (2.6)$$

Thus, we have

$$G_n = \frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}|nD_i - nD_j| \simeq \frac{S_{n-1}}{n-1}, \qquad (2.7)$$

where $S_{n-1} = \sum_{k=1}^{n-1} X_k$ is the sum of $(n-1)$ independent Uniform$([0, 2\pi))$ random variables. The probability distribution of S_{n-1} is a variation of the classical Irwin-Hall Uniform sum distribution, which was first derived by P.S. Laplace in 1814 (cf. Wilks, 1962; Feller, 1971, Theorem 1, I.9). The probability density function of S_{n-1} has the form

$$f_{S_{n-1}}(s) = \frac{I[0 < s < 2\pi(n-1)]}{(2\pi)^{n-1}(n-2)!}\sum_{k=0}^{n-1}\binom{n-1}{k}(-1)^k(s - 2\pi k)_+^{n-2}, \qquad (2.8)$$

and can be derived via the Fourier inversion formula and Cauchy's integral formula from complex analysis. The cumulative distribution function of S_{n-1} is

$$F_{S_{n-1}}(s) = \frac{I[0 < s < 2\pi(n-1)]}{(2\pi)^{n-1}(n-1)!}\sum_{k=0}^{n-1}\binom{n-1}{k}(-1)^k(s - 2\pi k)_+^{n-1}. \qquad (2.9)$$

Under the null hypothesis of circular uniformity, the probability density function of G_n is

$$f_{G_n}(y) = \frac{(n-1)\cdot I(0 < y < 2\pi)}{(2\pi)^{n-1}(n-2)!}\sum_{k=0}^{n-1}\binom{n-1}{k}(-1)^k\left[(n-1)y - 2\pi k\right]_+^{n-2}, \qquad (2.10)$$

with cumulative distribution function

$$F_{G_n}(y) = \frac{I(0 < y < 2\pi)}{(2\pi)^{n-1}(n-1)!}\sum_{k=0}^{n-1}\binom{n-1}{k}(-1)^k\left[(n-1)y - 2\pi k\right]_+^{n-1}, \qquad (2.11)$$

and characteristic function

$$\varphi_{G_n}(t) = \int_{-\infty}^{\infty} e^{ity}\, dF_{G_n}(y) = (n-1)^{n-1}\left(\frac{\exp\left(\frac{2\pi i t}{n-1}\right) - 1}{2\pi i t}\right)^{n-1}, \quad (i = \sqrt{-1}). \qquad (2.12)$$

Under the null hypothesis, G_n has an asymptotic Normal distribution which is applicable to large sample situations. From the classical Central Limit Theorem, in the limit as $n \to \infty$,

$$\sqrt{n}\left(\frac{1}{n-1}\sum_{k=1}^{n-1} U_k - \frac{1}{2}\right) \overset{D}{\to} N_1\left(0, \frac{1}{12}\right). \qquad (2.13)$$

Since $G_n \simeq \frac{S_{n-1}}{n-1} = \frac{(2\pi)}{n-1} \sum_{k=1}^{n-1} U_k$, we have in the limit as $n \to \infty$,

$$\sqrt{n}(G_n - \pi) = \sqrt{n}\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|nD_i - nD_j|}{2n(n-1)} - \pi\right) \xrightarrow{D} N_1\left(0, \frac{\pi^2}{3}\right). \qquad (2.14)$$

Let

$$\alpha = \mathbb{P}(G_n > y_\alpha) = 1 - F_{G_n}(y_\alpha) \qquad (2.15)$$

be the upper-tail probability corresponding to the critical value y_α of the test statistic G_n. In Table 1, we give the upper percentiles of the exact distribution function for the statistic G_n for testing the null hypothesis of circular uniformity. The table gives these critical values, which have been given in degrees for immediate applicability,

Table 1
Upper percentiles (in degrees) of the exact distribution for the Gini mean
difference arc-lengths test G_n

n	$\alpha = 0.01$	$\alpha = 0.02$	$\alpha = 0.03$	$\alpha = 0.04$	$\alpha = 0.05$	$\alpha = 0.10$
2	356.22	352.62	349.02	345.42	341.82	323.82
3	333.92	323.55	315.55	308.77	302.80	279.30
4	312.25	300.32	291.87	285.12	279.40	258.62
5	296.23	284.63	276.75	270.64	265.57	247.67
6	284.59	273.79	266.61	261.07	256.49	240.31
7	275.90	265.81	259.15	254.02	249.78	234.90
8	269.07	259.57	253.32	248.53	244.58	230.74
9	263.51	254.51	248.61	244.10	240.39	227.40
10	258.88	250.31	244.71	240.44	236.92	224.65
11	254.93	246.75	241.41	237.33	233.99	222.32
12	251.53	243.68	238.57	234.67	231.47	220.33
13	248.55	241.00	236.08	232.34	229.27	218.59
14	245.91	238.63	233.89	230.29	227.33	217.06
15	243.56	236.52	231.94	228.46	225.60	215.69
16	241.44	234.61	230.19	226.82	224.06	214.47
17	239.52	232.90	228.60	225.33	222.65	213.37
18	237.77	231.33	227.15	223.98	221.38	212.36
19	236.16	229.89	225.83	222.74	220.21	211.44
20	234.68	228.57	224.61	221.60	219.14	210.60
21	233.32	227.35	223.48	220.55	218.14	209.82
22	232.05	226.21	222.44	219.57	217.22	209.10
23	230.86	225.16	221.46	218.66	216.37	208.42
24	229.76	224.17	220.55	217.81	215.56	207.79
25	228.72	223.24	219.70	217.01	214.81	207.20
30	224.36	219.36	216.13	213.67	211.67	204.74
35	220.99	216.36	213.37	211.10	209.24	202.84
40	218.29	213.96	211.16	209.04	207.30	201.32
45	216.04	211.98	209.34	207.34	205.70	200.07
50	213.87	210.25	207.78	205.90	204.35	199.01

for small to moderate sample sizes. If for a given sample size n and significance level α, the observed value of the test statistic G_n, say y_{obs}, is greater than the tabulated critical value y_α, i.e., $y_{obs} > y_\alpha$, then we reject the null hypothesis of circular uniformity.

Note that, under the null hypothesis, the so-called "p-value" or observed significance level can be calculated by

$$p = \mathbb{P}(G_n > y_{obs}) = 1 - F_{G_n}(y_{obs}). \tag{2.16}$$

Equivalently, the null hypothesis is rejected whenever $p < \alpha$.

3. Illustrative Examples

In this section, we present a couple of circular data analysis examples. We illustrate how the Gini mean difference test G_n compares with Rao's test J_n on two classical circular data sets.

Example 3.1 (Hospital Birth Times Data). Suppose one wants to know whether or not birth times at a hospital are uniformly distributed throughout the day. The alternative hypothesis is that there is a time (or times) when births are more frequent. Table 2 displays data for delivery times collected across several days. This data can be found in Russell and Levitin (1995).

These observed event times are modeled as realizations from a continuous circular distribution. The observations can be converted to angles around a circle in an obvious way, e.g., if we want the angular units in degrees, we use 1 hr. $= \frac{360\,\text{deg.}}{24} = 15°$ and 1 min. $= \frac{360\,\text{deg.}}{24\,\text{hr.}} \cdot \frac{1\,\text{hr.}}{60} = 0.25°$. Thus, 12:00 am $= 0°$, 6:00 am $= 90°$, 12:00 pm $= 180°$, 6 pm $= 270°$, 9:15 am $= 138.75°$, etc.

Table 2
Hospital Birth Times Data

k	Delivery Time	$\theta_{(k)}$	D_k
1	12:20 am	5	34
2	12:40 am	10	5
3	12:40 am	10	0
4	12:48 am	12	2
5	1:08 am	17	5
6	5:40 am	85	68
7	6:00 am	90	5
8	6:36 am	99	9
9	6:40 am	100	1
10	7:20 am	110	10
11	10:12 am	153	43
12	3:32 pm	233	80
13	3:40 pm	235	2
14	7:44 pm	296	61
15	10:04 pm	331	35

Rao's arc-lengths test statistic gives an observed value of $J_{15} = 177$ with a p-value between 0.01 and 0.05. At the 5% significance level, this is sufficient evidence to reject the null hypothesis of circular uniformity and conclude that there are times when births are more frequent.

On the other hand, the Gini mean difference arc-lengths test statistic gives an observed value of $G_{15} = 224.86$ with a p-value of 0.053. The results from the Gini test are borderline significant, and may possibly indicate there are times when births are more frequent.

Example 3.2 (Homing Pigeon Data). Thirteen homing pigeons were released one at a time in the Toggenburg Valley in Switzerland under sub-Alpine conditions. They did not appear to have adjusted quickly to the homing direction, but preferred to fly in the axis of the valley, indicating a somewhat bimodal distribution. The vanishing angles are arranged here in increasing order as follows:

$$20, 135, 145, 165, 170, 200, 300, 325, 335, 350, 350, 350, 355.$$

Do these homing pigeons have a preferred direction of flight? (This example can also be found in Jammalamadaka and SenGupta, 2001).

The observed value of Rao's arc-lengths test statistic is $J_{13} = 161.92$ with a p-value between 0.05 and 0.10 (cf. with the table of upper percentiles of the distribution for J_n in Rao, 1976). On the basis of Rao's arc-lengths test, there is not enough evidence to reject the hypothesis of circular uniformity at the 5% significance level.

On the other hand, the observed value of the Gini mean difference arc-lengths test statistic is $G_{13} = 231.67$ with an observed significance level or p-value of $p = 0.043$. Therefore, the results of the Gini mean difference arc-lengths test are significant at the 5% significance level and we can reject the null hypothesis of circular uniformity. On the basis of Gini mean difference arc-lengths test, there is sufficient evidence that the homing pigeons have a preferred direction of flight.

4. Asymptotic Relative Efficiencies

There are clearly many other spacings tests as well as other uniformity tests for circular data. There is also considerable literature on comparing their asymptotic efficiencies. For instance, Pitman asymptotic relative efficiencies (ARE's) for sum-functions of spacings have been discussed in Sethuraman and Rao (1970), while exact Bahadur efficiencies have been studied in Zhou and Jammalamadaka (1989). In this section, we discuss the Pitman ARE's of both the Gini mean difference test G_n and Rao's test J_n, as well as generalized versions of these statistics.

We define the generalized Rao's arc-lengths test

$$J_n(r) = \frac{1}{2n} \sum_{k=1}^{n} |nD_k - 2\pi|^r, \quad r > 0, \tag{4.1}$$

and the generalized Gini mean difference arc-lengths test

$$G_n(r) = \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} |nD_i - nD_j|^r, \quad r > 0. \tag{4.2}$$

For the special case $r = 1$, we have $J_n(1) = J_n$ and $G_n(1) = G_n$. Moreover, the special case of $J_n(2)$ corresponds to both $G_n(2)$ as well as the statistic $\frac{1}{n} \sum_{k=1}^{n} (nD_k)^2$, which is called the Greenwood statistic.

Broadly speaking, the Pitman ARE of one sequence of tests against another corresponds to the limit of the inverse ratio of sample sizes required for the two tests to attain the same power at a sequence of alternatives which converges to the null hypothesis. In order to study Pitman ARE's, one needs to obtain the asymptotic distribution of test statistics under a sequence of close alternatives, which converges to the null hypothesis. In the circular case, the alternative hypothesis can be specified by a sequence of distribution functions $\{F_n(x) : n \geq 1\}$ that converges to the Uniform $([0, 2\pi))$ distribution function, which corresponds to the null hypothesis, in the limit as $n \to \infty$.

For symmetric spacings tests, the appropriate sequence of close alternatives (cf. Sethuraman and Rao, 1970; and Rao and Sethuraman, 1975) is obtained by using the distribution function

$$F_n(x) = \frac{x}{2\pi} + \frac{L_n(x)}{n^{1/4}}, \quad \text{for } 0 \leq x < 2\pi, \tag{4.3}$$

where $L_n(0) = L_n(2\pi) = 0$. We further assume that $L_n(x)$ is twice differentiable on the unit interval $[0, 2\pi)$ and that there exists a function $L(x)$ which is twice continuously-differentiable with $L(0) = L(2\pi) = 0$ and

$$n^{1/4} \sup_{0 \leq x < 2\pi} |L_n(x) - L(x)| = o(1), \tag{4.4}$$

$$n^{1/4} \sup_{0 \leq x < 2\pi} |L_n'(x) - l(x)| = o(1), \tag{4.5}$$

$$n^{1/4} \sup_{0 \leq x < 2\pi} |L_n''(x) - l'(x)| = o(1), \tag{4.6}$$

where $l(x)$ and $l'(x)$ are, respectively, the first and second derivatives of $L(x)$.

The asymptotic Normal distributions of test statistics, under both the null hypothesis and the sequence of close alternatives, can be adapted to the circular case. However, such an adaptation is not really necessary for finding the Pitman ARE's of test statistics, because the Pitman ARE's in the linear case carry over nicely to the circular case without much painstaking effort. The Pitman ARE's of $J_n(r)$ and $G_n(r)$ were obtained in Tung and Jammalamadaka (2012) in the context of goodness-of-fit testing on the real line.

Sethuraman and Rao (1970) showed that among spacings tests of the form $V_n(g)$, the most asymptotically efficient, i.e., the asymptotically locally most powerful test (ALMP) is the Greenwood statistic. Tung and Jammalamadaka (2012) investigated U-statistics based on spacings of the form $W_n(h)$ (see Equation (1.4)) and showed that among such tests, the ALMP test is the Gini mean squared difference test $G_n(2)$. However it turns out that this is algebraically the same as the Greenwood statistic and thus has the same efficiency.

Suppose the Pitman ARE of $J_n(2)$ and $G_n(2)$ is taken to be 1. The following Table 3, taken from Tung and Jammalamadaka (2012), lists the Pitman ARE of $J_n(r)$ and $G_n(r)$ with respect to various choices of $r > 0$. It is seen that the Pitman ARE's of $J_n(1)$ and $G_n(1)$ are 0.572654 and 0.75, respectively, thus the Gini mean difference test $G_n(1)$ is asymptotically more efficient than Rao's test $J_n(1)$.

Table 3
Pitman asymptotic relative efficiencies for $J_n(r)$ and $G_n(r)$

r	Generalized Rao	Generalized Gini
1	0.572654	0.75
3/2	0.892135	0.946889
2	1	1
5/2	0.93921	0.96137
3	0.818649	0.867857
4	0.550562	0.615384

Moreover, it is also seen that the generalized Gini mean difference test $G_n(r)$ is more Pitman efficient than the generalized Rao's test $J_n(r)$, except for the case $r = 2$, when both tests $G_n(2)$ and $J_n(2)$ correspond to the Greenwood statistic and have a Pitman ARE of 1.

5. Conclusion

We introduced a new test of uniformity on the circle based on the Gini mean difference of the sample arc-lengths, and obtained both its exact and asymptotic distributions under the null hypothesis. We provided a table of upper percentile values for this test, which will be useful to applied scientists employing it for circular data analysis. This new test extends the use of one by Jammalamadaka and Goria (2004) from the linear case to the circular case. On the basis of Pitman asymptotic relative efficiency, the generalized Gini mean difference test is asymptotically more efficient than the generalized Rao's test.

References

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. Vol. II. New York: John Wiley and Sons.

Gatto, R. (2001). Symbolic computation for approximating the distributions of some families of one and two-sample nonparametric test statistics. *Statist. Comput.* 11:449–455.

Gatto, R., Jammalamadaka, S. R. (1999). A conditional saddlepoint approximation for testing problems. *J. Amer. Statist. Assoc.* 94(446):533–541.

Jammalamadaka, S. R., Goria, M. N. (2004). A test of goodness of fit based on Gini's index of spacings. *Statist. Probab. Lett.* 68:177–187.

Jammalamadaka, S. R., SenGupta, A. (2001). *Topics in Circular Statistics*. Series on Multivariate Analysis, Vol. 5. Singapore: World Scientific.

Pyke, R. (1965). Spacings. *J. Roy. Statist. Soc. Ser. B (Methodol.)* 27(3):395–449.

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

Rao, J. S. (1976). Some tests based on arc-lengths for the circle. *Sankhya Ind. J. Statist. Ser. B* 38(4):329–338.

Rao, J. S., Sethuraman, J. (1975). Weak convergence of empirical distribution functions of random variables subject to perturbations and scale factors. *Ann. Statist.* 3(2):299–213.

Russell, G. S., Levitin, D. J. (1995). An expanded table of probability values for Rao's spacings test. *Commun. Statist. Simul. Computat.* 24:879–888.

Sethuraman, J., Rao, J. S. (1970). Pitman efficiencies of tests based in spacings. In: Puri, M. L., ed., *Nonparametric Techniques in Statistical Inference*. Cambridge: Cambridge University Press, pp. 405–415.

Tung, D. D., Jammalamadaka, S. R. (2012). *U*-statistics based on spacings. *J. Stat. Plan. Infer.* 142(3):673–684.

Wilks, S. S. (1962). *Mathematical Statistics*. New York: John Wiley and Sons.

Zhou, X., Jammalamadaka, S. R. (1989). Bahadur efficiencies of spacings tests for goodness of fit. *Ann. Instit. Statist. Math.* 41(3):541–553.