

An entropy-based test for goodness of fit of the von Mises distribution

Ulric J. Lund and S.R. Jammalamadaka

Western Washington University and University of California, Santa Barbara

Abstract

The maximum entropy characterization of the von Mises distribution on the circle is exploited to derive a consistent goodness of fit test statistic for the von Mises distribution. Monte Carlo simulation results are tabulated giving critical values of the test statistic for various sample sizes and values of the concentration parameter. A power analysis is presented for various alternative hypotheses, comparing the entropy statistic to two competing goodness of fit statistics. The entropy statistic is shown to compare favorably and may be more attractive, especially considering its ease of computation. Finally, it is cautioned that the maximum entropy statistic, which is derived under the assumption that the mean direction is well-defined, is not appropriate for differentiating between the von Mises and the uniform distribution on the circle.

Keywords: Circular statistics; Entropy; Goodness of fit; von Mises;

1 Introduction and Review

For a distribution function F , with density function f on R^1 , entropy is measured by the integral

$$H(f) = - \int_{-\infty}^{\infty} f(x) \log f(x) dx.$$

This may also be expressed as

$$H(f) = \int_0^1 \log \left\{ \frac{d}{dp} F^{-1}(p) \right\} dp,$$

for which a consistent estimate is given by Vasicek (1976):

$$H_{mn} = n^{-1} \sum_{i=1}^n \log \left\{ \frac{n}{2m} (x_{(i+m)} - x_{(i-m)}) \right\}. \quad (1.1)$$

In the above, $x_{(i)}$ indicates the i^{th} order statistic, and $2m$ is the size of the steps used for the spacings $x_{(i+m)} - x_{(i-m)}$, with the restriction that $m < n/2$. Consistency of H_{mn} , as $n \rightarrow \infty$, $m \rightarrow \infty$, and $m/n \rightarrow 0$, is proven in *Theorem 1* of Vasicek (1976).

For distribution functions defined on R^1 , it can be shown that for fixed first and second moments, the distribution with the maximum entropy is the normal distribution (Kagan, Linnik and Rao, 1973, p. 410). A test for normality, using this characterization, was first developed by Vasicek (1976). The intuition behind the test is that the maximum entropy for a normal distribution can be calculated analytically, and thus a sample's entropy should be close to this value if the sample comes from a normal distribution. The null hypothesis of normally distributed data is rejected for small sample entropies.

For various sample sizes and spacing widths, Vasicek (1976) tabulates the critical values of the test statistic and provides some power comparisons of the entropy statistic relative to various other goodness of fit tests like the Kolmogorov-Smirnov, Cramer-von Mises, Kuiper V, Watson U^2 , Anderson-Darling, and Shapiro-Wilk. None of the tests performed uniformly better for all of the alternative hypotheses examined: exponential, gamma, uniform, beta, cauchy. However, the entropy statistic had the highest power for three of the five alternatives.

2 Goodness of fit testing of the von Mises distribution

A statistic to test for goodness of fit of the von Mises distribution will now be derived. It is based on the von Mises' maximum entropy characterization on the circle, for fixed mean direction and circular variance.

Proposition 1 *For all probability distributions on the circle, $f(\theta)$, $0 \leq \theta < 2\pi$, with fixed mean direction μ_0 and circular variance $1 - \rho$, that is*

$$\begin{aligned} \int_0^{2\pi} \cos \theta f(\theta) d\theta &= \rho \cos \mu_0 \\ \int_0^{2\pi} \sin \theta f(\theta) d\theta &= \rho \sin \mu_0, \end{aligned} \tag{2.1}$$

the entropy of $f(\theta)$,

$$H(f) = - \int_{-\infty}^{\infty} f(\theta) \log f(\theta) d\theta$$

is maximized when $f(\theta)$ is the von Mises distribution.

A proof of this result can be found in Mardia (1972, pp. 65-66), and can also be derived quite easily via *Theorem 13.2.1* of Kagan, Linik and Rao (1973, p. 409).

To assess the goodness of fit of a von Mises distribution for a set of data, the sample entropy will be compared to the entropy of a von Mises distribution, which is given by

$$\begin{aligned}
H(f) &= - \int_{-\infty}^{\infty} f(\theta) \log f(\theta) d\theta \\
&= - \int_{-\infty}^{\infty} \frac{\exp\{\kappa \cos(\theta - \mu)\}}{2\pi I_0(\kappa)} [-\log(2\pi I_0(\kappa)) + \kappa \cos(\theta - \mu)] \\
&= \log[2\pi I_0(\kappa)] - \kappa A(\kappa) \\
&= \log \left[\frac{2\pi I_0(\kappa)}{\exp(\kappa A(\kappa))} \right], \tag{2.2}
\end{aligned}$$

where $A(\kappa) = I_1(\kappa)/I_0(\kappa)$, and I_0 and I_1 are the modified Bessel functions of the first kind, orders 0 and 1 respectively.

Now, applying *Theorem 1* from Vasicek (1976), under the null hypothesis that a sample of observations x_1, \dots, x_n comes from a von Mises distribution, the statistic H_{mn} given in (1.1), is a consistent estimate of $H(f)$, defined in (2.2). That is,

$$H_{mn} \xrightarrow{p} \log \left[\frac{2\pi I_0(\kappa)}{\exp(\kappa A(\kappa))} \right],$$

as $n \rightarrow \infty$, $m \rightarrow \infty$, and $m/n \rightarrow 0$. Using the fact that

$$\begin{aligned}
H_{mn} &= n^{-1} \sum_{i=1}^n \log \left\{ \frac{n}{2m} (x_{(i+m)} - x_{(i-m)}) \right\} \\
&= \sum_{i=1}^n \log \left\{ \frac{n}{2m} (x_{(i+m)} - x_{(i-m)}) \right\}^{1/n} \\
&= \log \left\{ \frac{n}{2m} \prod_{i=1}^n [(x_{(i+m)} - x_{(i-m)})]^{1/n} \right\},
\end{aligned}$$

we use a consistent estimate $\hat{\kappa}$ of κ , to define

$$\begin{aligned}
K_{mn} &= \frac{\exp\{H_{mn}\} \exp\{\hat{\kappa} A(\hat{\kappa})\}}{I_0(\hat{\kappa})} \\
&= \frac{n \exp\{\hat{\kappa} A(\hat{\kappa})\}}{2 m I_0(\hat{\kappa})} \left[\prod_{i=1}^n (x_{(i+m)} - x_{(i-m)}) \right]^{1/n}. \quad (2.3)
\end{aligned}$$

Then, under the null hypothesis, we have

$$\begin{aligned}
K_{mn} &\xrightarrow{p} \frac{\exp\{H(f)\} \exp\{\kappa A(\kappa)\}}{I_0(\kappa)} \\
&= 2\pi.
\end{aligned}$$

That is,

$$K_{mn} \xrightarrow{p} 2\pi,$$

as $n \rightarrow \infty$, $m \rightarrow \infty$, and $m/n \rightarrow 0$. Again, samples coming from distributions that are not von Mises, will tend to have lower sample entropies, and thus smaller

values of K_{mn} . The null hypothesis is therefore rejected for sufficiently small values of K_{mn} . A discussion regarding the definition of spacings on the circle and the order statistics $x_{(i)}$ used in the calculation of (2.3) is presented in Section 2.1.

It is possible to derive the asymptotic distribution of H_{mn} when a simple null hypothesis is used, one in which μ and κ are specified. By taking $u_{(i)} = F_0(x_{(i)})$, and applying H_{mn} to $u_{(1)}, \dots, u_{(n)}$, one can employ the asymptotic theory for H_{mn} developed by Dudewicz and Van der Muelen (1981) for testing uniformity. However, even for a sample size of 100, Dudewicz et. al. concede that the asymptotic theory could not be validated by simulation results. Additionally, as will be remarked later in Section 2.2, one advantage of the entropy statistic presented in this paper, is that the distribution function does not need to be evaluated. This advantage would be lost, if the transformation $u_{(i)} = F_0(x_{(i)})$ was required.

More recently, van Es (1992) considered the asymptotic distribution of a statistic very similar to Vasicek's estimate of entropy (1.1). Van Es shows that for a density f , defined on finite support and satisfying the Lipschitz condition, if $m, n \rightarrow \infty$ and $m = o(n^{1/2})$, the statistic defined by

$$V_{mn} = \frac{1}{n-m} \sum_{i=1}^{n-m} \log \left(\frac{n+1}{m} (X_{(i+m)} - X_{(i)}) \right) + \sum_{k=2m}^n \frac{1}{k} + \log(2m) - \log(n)$$

has the asymptotic distribution given by

$$n^{1/2} (V_{mn} - H(f)) \xrightarrow{D} N(0, \text{var}(\log(f(X))), \quad (2.4)$$

where $H(f)$ is the entropy of density f .

Unfortunately, even though it can be shown that the von Mises distribution is Lipschitz, simulations could not confirm the result in (2.4) for the von Mises

case. Also, as might be expected, simulations show that for finite n and m the distribution of the test statistic is highly dependent on the choice of m , and there is no guide as to what step size should be used for a fixed and finite n . These two confounding factors make the asymptotic theory inapplicable in practice. For this reason, the asymptotic distribution of the entropy statistic, K_{mn} (2.3), will not be pursued further at this time.

Instead, Monte Carlo methods are used to find critical values for the statistic K_{mn} . For each of various sample sizes and values of κ , 5000 samples were generated from a von Mises distribution. From each sample, K_{mn} was calculated. A condensed table of 95th and 99th percentiles are given in Table 1. The choice of step size in the tabulation of critical values was according to the step size that yielded the largest test statistic. Some smaller scale simulations were performed to compare powers using various step sizes, and it was usually found that the step size yielding the largest critical values also yielded the largest power. This phenomenon has also been documented in the papers by Gokhale (1983) and Dudewicz and Van der Muelen (1981).

2.1 Calculation of the test statistic

To calculate the statistic

$$K_{mn} = \frac{n \exp\{\hat{\kappa} A(\hat{\kappa})\}}{2 m I_0(\hat{\kappa})} \left[\prod_{i=1}^n (x_{(i+m)} - x_{(i-m)}) \right]^{1/n},$$

a consistent estimate for κ is needed, and the m -step spacings are required.

The maximum likelihood estimate $\hat{\kappa}$ of κ , is a consistent estimate, and is

Table 1: Critical values for testing at the 5% (1%) significance level

κ	Sample Size									
	n = 20	n = 25	n = 30	n = 35	n = 40	n = 45	n = 50	n = 75	n = 100	
0.20	4.11(3.81)	4.39(4.12)	4.57(4.30)	4.69(4.47)	4.82(4.61)	4.92(4.71)	4.99(4.83)	5.27(5.14)	5.43(5.34)	
0.40	4.15(3.81)	4.40(4.12)	4.58(4.32)	4.71(4.49)	4.84(4.59)	4.95(4.77)	5.01(4.82)	5.28(5.14)	5.45(5.34)	
0.60	4.17(3.83)	4.43(4.16)	4.61(4.37)	4.76(4.53)	4.85(4.65)	4.97(4.75)	5.03(4.86)	5.31(5.16)	5.46(5.36)	
0.80	4.16(3.77)	4.44(4.14)	4.62(4.34)	4.78(4.52)	4.89(4.68)	4.99(4.79)	5.08(4.89)	5.33(5.19)	5.49(4.38)	
1.00	4.19(3.83)	4.44(4.16)	4.66(4.36)	4.80(4.51)	4.91(4.69)	5.00(4.81)	5.08(4.90)	5.35(5.19)	5.51(5.40)	
1.20	4.21(3.82)	4.44(4.13)	4.67(4.42)	4.81(4.54)	4.92(4.72)	5.00(4.76)	5.10(4.88)	5.38(5.24)	5.52(5.41)	
1.40	4.23(3.92)	4.46(4.15)	4.66(4.38)	4.82(4.54)	4.93(4.69)	5.04(4.83)	5.11(4.89)	5.39(5.24)	5.54(5.43)	
1.60	4.20(3.83)	4.46(4.13)	4.67(4.41)	4.82(4.54)	4.93(4.68)	5.02(4.80)	5.12(4.93)	5.40(5.27)	5.55(5.43)	
1.80	4.21(3.88)	4.48(4.13)	4.65(4.40)	4.79(4.54)	4.94(4.69)	5.04(4.84)	5.12(4.91)	5.40(5.24)	5.56(5.43)	
2.00	4.20(3.87)	4.44(4.13)	4.67(4.43)	4.81(4.54)	4.94(4.72)	5.02(4.77)	5.13(4.92)	5.41(5.25)	5.57(5.43)	
2.20	4.22(3.86)	4.45(4.11)	4.66(4.40)	4.82(4.55)	4.93(4.69)	5.03(4.79)	5.12(4.93)	5.40(5.24)	5.57(5.44)	
2.40	4.20(3.83)	4.45(4.14)	4.66(4.41)	4.82(4.58)	4.92(4.67)	5.03(4.82)	5.11(4.86)	5.41(5.25)	5.57(5.44)	
2.60	4.20(3.85)	4.48(4.17)	4.65(4.39)	4.82(4.55)	4.93(4.71)	5.03(4.82)	5.12(4.93)	5.40(5.22)	5.56(5.44)	
2.80	4.21(3.82)	4.45(4.14)	4.67(4.38)	4.80(4.55)	4.94(4.68)	5.03(4.83)	5.11(4.94)	5.39(5.23)	5.56(5.43)	
≥ 3.00	4.21(3.85)	4.46(4.16)	4.66(4.39)	4.81(4.57)	4.92(4.69)	5.02(4.81)	5.10(4.90)	5.38(5.23)	5.54(5.42)	
Step Size	m = 3	m = 4	m = 4	m = 4	m = 5	m = 5	m = 5	m = 7	m = 8	

defined implicitly by

$$A(\hat{\kappa}) = \frac{1}{n} \sum_{i=1}^n \cos(x_i - \hat{\mu}),$$

where $\hat{\mu}$ is the sample mean direction. Best and Fisher (1981) provide an algorithm to estimate $\hat{\kappa} = A^{-1}(\hat{\kappa})$:

$$A^{-1}(\kappa) = \begin{cases} 2\kappa + \kappa^3 + (5\kappa^5)/6 & 0 \leq \kappa < 0.53 \\ -0.4 + 1.39\kappa + 0.43/(1 - \kappa) & 0.53 \leq \kappa < 0.85 \\ 1/(\kappa^3 - 4\kappa^2 + 3\kappa) & \kappa \geq 0.85 \end{cases} ,$$

The calculation of $I_0(\hat{\kappa})$ can be carried out by some computer packages, but can also be done using a polynomial approximation given by Abramowitz & Stegun (1970). Letting $t = \kappa/3.75$,
 $0 \leq \kappa \leq 3.75$:

$$I_0(\kappa) = 1 + 3.5156229t^2 + 3.0899424t^4 + 1.2067492t^6 \\ + 0.2659732t^8 + 0.0360768t^{10} + 0.0045813t^{12}$$

$\kappa > 3.75$:

$$\kappa^{1/2} e^{-\kappa} I_0(\kappa) = 0.39894228 + 0.0132859t^{-1} + 0.00225319t^{-2} \\ - 0.00157565t^{-3} + 0.0091628t^{-4} - 0.02057706t^{-5} \\ + 0.02635537t^{-6} + 0.00392377t^{-8}.$$

Circular statistics are ideally invariant with respect to the choice of the zero direction. This will be kept in mind, when determining the m -step spaces $x_{(i+m)} - x_{(i-m)}$. There are two ways to define the spaces on the circle. One

method is to utilize the circularity of the observations, and to take $x_{(i+m)} = x_{(i+m \pmod n)}$ for values of $i + m$ larger than n . In the exploration of the sample entropy, H_{mn} (1.1), for von Mises data, this was the first method of spacings used, as it fully exploits the circularity of the data, and also makes H_{mn} invariant with respect to the zero location. Results, however, showed that the statistic behaved irregularly, and did not approximate the population entropy very well, especially for large values of κ .

The second method is equivalent to Vasicek's (1976) definition of the spacings for linear data; that is, let

$$\begin{aligned} x_{(i+m)} &= x_{(n)}, \quad i + m > n \quad \text{and} \\ x_{(i-m)} &= x_{(1)}, \quad i - m < 1. \end{aligned}$$

However, under this definition of spacings, H_{mn} will not be invariant under rotations, because the arc between $x_{(0)}$ and $x_{(n)}$ is not being utilized, and will vary under different choices of the zero direction. The solution to this problem is to define $x_{(0)}$ and $x_{(n)}$ such that the largest gap between adjacent observations is between $x_{(0)}$ and $x_{(n)}$ (see Figure 1). Under this definition of the order statistics, the spacings will be invariant under rotations, and hence H_{mn} and K_{mn} will be as well. In terms of sufficiency, disregarding the largest gap can be reconciled with the fact that the length of this omitted gap could be obtained by subtracting the sum of the other arc lengths from 2π .

One final remark about the calculation of the test statistic, K_{mn} , is with regards to ties in the data set. As the statistic is a function of the product of the spacings between the observations, any ties among the observations will set the test statistic equal to zero. It is therefore essential that the data set does not contain ties, if the entropy statistic is to be used. Theoretically, as the data is assumed to be from a continuous distribution, ties should occur with probability zero.

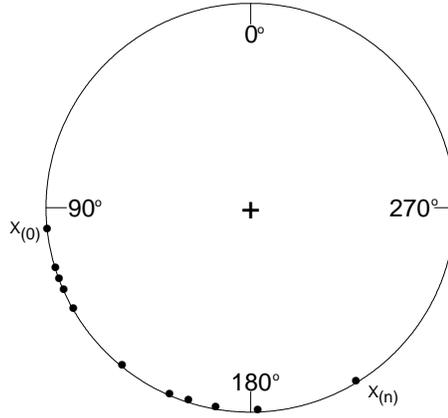


Figure 1: Defining the order statistics

2.2 Comparison of the entropy statistic with existing tests

The most commonly used statistic for testing goodness of fit of the von Mises distribution is the U^2 statistic, which was originally developed by Watson (1961), and is an invariant adaptation of the Cramer-von Mises statistic:

$$U^2 = n \int_0^{2\pi} \left[F_n(\theta) - F(\theta) - \int_0^{2\pi} \{F_n(\phi) - F(\phi)\} dF(\phi) \right]^2 dF(\theta).$$

Lockhart and Stephens (1985) use the asymptotic distribution of U^2 to tabulate critical values for various cases, including the case when neither the location nor shape parameters of the von Mises distribution are known. They also report that Monte Carlo studies validate their critical values for $n \geq 20$.

Bowman (1992) develops a density based, integrated squared error test statis-

tic. Letting

$$e(x) = E\{f_n(x)\},$$

where the expectation is taken under the null hypothesis, he defines the statistic

$$I = \int \{e(x) - f_n(x)\}^2 dx.$$

Bowman (1992) proposes using the kernel density estimate f_n , with the smoothing parameter given by Parzen's (1962) formula. Critical values are obtained by Monte Carlo simulations, and are given for various sample sizes and values of κ .

The sample version of Watson's (1971) statistic is given by

$$U^2 = \sum_{i=1}^n \left\{ z_{(i)} - \frac{2i-1}{2n} \right\}^2 - n \left(\bar{z} - \frac{1}{2} \right)^2 + \frac{1}{12n},$$

where $z_i = F(x_i; \hat{\mu}, \hat{\kappa})$. There is no further simplification for the von Mises distribution. Numerical methods can be used, otherwise tables are available as well. However, this can be time consuming, if done by hand for relatively large samples.

For the von Mises case, Bowman (1992) defines the integrated squared error statistic as

$$\begin{aligned} I &= \frac{1}{n} M(0, A^{-1}(A(h))) + \frac{2}{n(n-1)} \sum_{i < j} M(x_i - x_j, A^{-1}[A(h)^2]) \\ &- \frac{2}{n} \sum_i M(x_i - \hat{\mu}, A^{-1}[A(h)^2 A(\hat{\kappa})]) + M(0, A^{-1}[A(h)^2 A(\hat{\kappa})^2]). \end{aligned}$$

with

$$h = A^{-1}[\exp\{(4/3)^{2/5} n^{-2/5} \log A(\hat{\kappa})\}],$$

and M is the von Mises density function. Clearly, the computation of K_{mn} is far simpler than that of I , as I requires numerous evaluations of A and A^{-1} .

Computationally, the entropy statistic, K_{mn} , is much more attractive than each of the two statistics above. Practitioners of statistics with limited computational resources should find K_{mn} an attractive alternative to U^2 and I .

However, this computational edge over U^2 and I would not be meaningful if the entropy statistic did not compete in terms of power. Power comparisons between the three statistics follows. For samples of size 25, 50, 75, and 100, Bowman (1992) performs Monte Carlo power comparisons between U^2 and I . The alternative hypotheses considered are mixtures of the von Mises distribution:

1. Bimodal

$$\frac{1}{2}VM(\pi, 5) + \frac{1}{2}VM\left(\frac{3\pi}{2}, 5\right)$$

2. Skewed

$$\frac{2}{3}VM(\pi, 3) + \frac{1}{3}VM(0.62\pi, 3)$$

3. Long-tailed

$$\frac{1}{3}VM(\pi, 8) + \frac{2}{3}VM(\pi, 0.1)$$

4. Half von Mises

$$\begin{cases} 2VM(\pi, 2) & \pi \leq \theta \leq 2\pi \\ 0 & 0 < x < \pi. \end{cases}$$

With 1000 iterations used for each case, his results are reproduced within Table 2.

Bowman (1992) shows that the integrated squared error statistic outperforms U^2 for the bimodal and skewed alternatives. For the long-tailed distribution, U^2

Table 2: Power comparison between U^2 , I and K_{mn} . Powers for I and U^2 are from Bowman (1992).

Distribution	Sample Size											
	$n = 25$			$n = 50$			$n = 75$			$n = 100$		
	I	U^2	$K_{4,25}$	I	U^2	$K_{5,50}$	I	U^2	$K_{7,75}$	I	U^2	$K_{8,100}$
1. Bimodal	0.74	0.58	0.63	0.97	0.91	0.90	1.00	0.99	0.97	1.00	1.00	0.99
2. Skewed	0.14	0.10	0.14	0.22	0.16	0.23	0.27	0.24	0.28	0.47	0.29	0.32
3. Long-tailed	0.14	0.38	0.16	0.24	0.56	0.35	0.40	0.75	0.49	0.51	0.87	0.69
4. Half	0.43	0.45	0.74	0.82	0.81	0.98	0.97	0.93	1.00	1.00	0.99	1.00

is far superior, while the two statistics are fairly equal for the half von Mises. Simulated power of the entropy statistic, K_{mn} , for the same alternatives used by Bowman (1992), is also given in Table 2. 5000 iterations have been used in this simulation. Under the bimodal alternative, K_{mn} beats U^2 for small samples, $n = 25$, and is only slightly lower than U^2 and I for the other sample sizes. For the skewed alternative, K_{mn} is uniformly better than U^2 , and again quite close to I . U^2 has a clear advantage under the long-tailed alternative, however, K_{mn} is uniformly more powerful than I . Under the half von Mises alternative K_{mn} outperformed the other two statistics, most decisively for the smaller sample sizes.

In addition to comparing K_{mn} to the results of Bowman (1992), the entropy statistic was compared to U^2 for two other alternatives, which closely resemble the von Mises distribution: the cardoid and triangular distributions. The results are given in Table 3. The entropy statistic clearly has the higher power for both alternatives and all sample sizes. Bowman's (1992) I statistic was not included in this comparison as it is not commonly employed, while U^2 is currently the statistic used most frequently.

Bowman (1992) comments that an entropy statistic for testing goodness of fit of the von Mises distribution changed too quickly as a function of κ , creating

Table 3: Power comparison between U^2 and K_{mn} , based on 5000 iterations.

Distribution	Sample Size							
	$n = 25$		$n = 50$		$n = 75$		$n = 100$	
	$K_{4,25}$	U^2	$K_{5,50}$	U^2	$K_{7,75}$	U^2	$K_{8,100}$	U^2
1. Cardoid	0.11	0.08	0.15	0.11	0.21	0.14	0.24	0.18
2. Triangular	0.09	0.06	0.12	0.08	0.16	0.08	0.22	0.10

difficulty in the control of the level of the test. Table 4 shows the simulated level of K_{mn} , generating data from a von Mises distribution with various sample sizes and values for κ . The simulated levels are all quite close to the desired 0.05. It is possible that Bowman (1992) used the circular spacings, which could produce erratic results. Regardless, the control of type I error does not seem to be a problem in this paper.

Table 4: Simulated levels of K_{mn} , based on 5000 iterations ($\alpha = 0.05$).

κ	Sample Size			
	$n = 25$	$n = 50$	$n = 75$	$n = 100$
1	0.0496	0.0524	0.0472	0.0498
3	0.0526	0.0560	0.0546	0.0472
5	0.0532	0.0518	0.0468	0.0538

2.3 Concluding remarks

The entropy statistic in (2.3) is not suitable for testing the null hypothesis of a von Mises distribution versus the alternative of uniformity, since the uniform distribution is not within the class of distributions for which the von Mises distribution has the maximum entropy characterization - the mean direction of the uniform distribution is undefined. Since the von Mises distribution has the maximum entropy of all distributions with fixed mean direction and circular variance,

whereas the uniform distribution on the circle maximizes the entropy unconstrained, the entropy of the uniform distribution will exceed that of the von Mises distribution for all $\kappa > 0$. Therefore, the entropy statistic in (2.3) will be ineffective in distinguishing between uniform and von Mises distributions.

To test between the uniform and von Mises distributions, the Rayleigh test is recommended. Noting that the von Mises density function reduces to the uniform density when $\kappa = 0$, this test takes the null hypothesis to be $\kappa = 0$, and the alternative $\kappa > 0$. A description of this test procedure is given in Mardia (1972, p. 133), and Fisher (1993, p. 82).

References

- [1] Abramowitz, M., and Stegun, I.A., (1970), *Handbook of Mathematical Functions*, New York: National Bureau of Standards.
- [2] Best, D.J., and Fisher, N.I., (1981), The bias of the maximum likelihood estimators of the von Mises-Fisher concentration parameters, *Communications in Statistics - Simulations and Computations*, B10(5), 493-502.
- [3] Bowman, A.W., (1992), "Density based tests for goodness-of-fit", *Journal of Statistical Computation and Simulation*, 40, 1-13.
- [4] Dudewicz, E.J., and Van der Meulen, E.C., (1981), "Entropy-based tests of uniformity", *Journal of the American Statistical Association*, 76, 967-974.
- [5] Fisher, N.I., (1993), *Statistical Analysis of Circular Data*, Cambridge: Cambridge University Press.
- [6] Gokhale, D.V., (1983), "On entropy-based goodness-of-fit tests", *Computational Statistics and Data Analysis*, 1, 157-165.
- [7] Kagan, A.M., Linnik, Y.V., and Rao, C.R., (1973), *Characterization Problems in Mathematical Statistics*, New York: John Wiley and Sons.

- [8] Lockhart, R.A., and Stephens, M.A., (1985), "Tests of fit for the von Mises distribution", *Biometrika*, 72, 647-652.
- [9] Mardia, K.V., (1972), *Statistics of Directional Data*, London: Academic Press.
- [10] Van Es, B., (1992), "Estimating Functionals Related to a Density by a Class of Statistics Based on Spacings", *Scandinavian Journal of Statistics*, 19, 61-72.
- [11] Vasicek, O., (1976), "A test of normality based on sample entropy", *Journal of the Royal Statistical Society Series B*, 38, 54-59.
- [12] Watson, G.S., (1961), "Goodness-of-fit tests on a circle", *Biometrika*, 48, 109-114.