The Annals of Mathematical Statistics
1972, Vol. 43, No. 2, 468-479

BAHADUR EFFICIENCIES OF SOME TESTS FOR UNIFORMITY ON THE CIRCLE[1]

BY J. S. RAO

Indiana University

1. Introduction and summary. In this paper we compare the asymptotic efficiencies of several tests that are available for testing uniformity on the circle. Since the problem of testing goodness of fit on the circle can be reduced to testing uniformity by a simple probability transformation, these comparisons are applicable also to the goodness of fit situation. The alternatives to uniformity considered here are the familiar circular normal distributions (CND's) with density

$$g(\alpha) = [2\pi I_0(\kappa)]^{-1} \exp[\kappa \cos \alpha], \quad -\pi \leqq \alpha < \pi . \tag{1.1}$$

$0 \leqq \kappa < \infty$ is a parameter of concentration, larger values of κ corresponding to more concentration towards the mean direction zero, and $I_0(\kappa)$ is the Bessel function of purely imaginary argument.

When $\kappa = 0$ (1.1) is the uniform density, so the null hypothesis is H_0: $\kappa = 0$. The tests compared here are

(i) Ajne's test A
(ii) Watson's test W
(iii) Rayleigh's test R
(iv) Ajne's test N
(v) Kuiper's test V
(vi) Spacings test U.

In subsequent sections each of these tests is briefly described and its Bahadur efficiency [4], [5] is computed, using large deviation results.

We compare the local slopes of the test statistics, i.e. the slopes in the neighborhood of the hypothesis. On the basis of these comparisons, we find that limiting efficiencies of the first three tests viz., Ajne's test A, Watson's W and Rayleigh's test based on R, are identical, while the other tests have lower asymptotic efficiencies. Further conclusions are given in Section 7. Finally in Section 8 a simple inequality between the Ajne's N and Kuiper's V, whose asymptotic performances are identical, is noted.

2. Some preliminaries. The concept of Bahadur efficiency by now is well

Received April 6, 1970; revised August 18, 1971.

[1] Research sponsored by the Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFOSR 71-2009. The United States Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation hereon.

known (see for example Bahadur [4], [5] and Gleser [7]). Throughout this paper we employ the notations of Bahadur [5]. The following three lemmas will be useful in the computation of the local slopes.

LEMMA 2.1 (Sethuraman). *Let $\mathscr{X}$ be a separable Banach space and $\mathscr{X}_1^*$, the space of all continuous linear functionals x^* on $\mathscr{X}$ with norm unity. Let $X_1(\omega)$, $X_2(\omega), \cdots$ be a sequence of random variables, defined on a probability space (Ω, S, P) with values in $\mathscr{X}$, which are independently and identically distributed with a common distribution $P(\cdot)$. Let*

$$(2.1) \qquad \int x^*(X(\omega))P(d\omega) = 0 \qquad \textit{for all } x^* \in \mathscr{X}_1^*$$

and let

$$(2.2) \qquad \int \exp(t\|X(\omega)\|)P(d\omega) < \infty \qquad \textit{for all } t .$$

Then for $\varepsilon > 0$,

$$(2.3) \qquad n^{-1} \log P\left\{\omega: \left\|\frac{X_1(\omega) + \cdots + X_n(\omega)}{n}\right\| \geqq \varepsilon\right\} \to \log \rho(\mathscr{X}_1^*, \varepsilon)$$

where

$$(2.4) \qquad \rho(\mathscr{X}_1^*, \varepsilon) = \sup_{x^* \in \mathscr{X}_1^*} \rho(x^*, \varepsilon)$$

and

$$(2.5) \qquad \rho(x^*, \varepsilon) = \max\{\rho_1(x^*, \varepsilon) = \min_{t \geqq 0} e^{-t\varepsilon} E(\exp[tx^*(X(\omega))]) , \quad \rho_2(x^*, \varepsilon) = \min_{t \leqq 0} e^{t\varepsilon} E(\exp[tx^*(X(\omega))])\} .$$

The following lemma deals with the behavior of $\rho(\mathscr{X}_1^*, \varepsilon)$ defined in (2.4), for sufficiently small ε. Notice that it is only the behavior of $\rho(\mathscr{X}_1^*, \varepsilon)$ for ε small that counts in the computation of the local slopes, since $b(\kappa)$ (see (30) of [5]) approaches zero as $\kappa \to 0$.

LEMMA 2.2. *For* $\varepsilon > 0$ *sufficiently small*

$$(2.6) \qquad \log \rho(\mathscr{X}_1^*, \varepsilon) = -\varepsilon^2/2\tau^2 + o(\varepsilon^2)$$

where

$$(2.7) \qquad \tau^2 = \sup_{x^* \in \mathscr{X}_1^*} \operatorname{Var} x^*(X(\omega)) .$$

PROOF. For any fixed $x^* \in \mathscr{X}_1^*$, we have from Lemma 2.1

$$\rho(x^*, \varepsilon) = \max\{\rho_1(x^*, \varepsilon), \rho_2(x^*, \varepsilon)\}$$

where $\rho_1(x^*, \varepsilon)$ and $\rho_2(x^*, \varepsilon)$ are as given in (2.5). Now let $\rho_1(x^*, \varepsilon)$ attain the minimum at the point $t = t_1$. Then for sufficiently small ε, in view of (2.1)

$$t_1 = \varepsilon/\sigma^2 + o(\varepsilon)$$

where $\sigma^2 = \operatorname{Var} x^*(X(\omega))$ (see for instance Bahadur [3]) and $o(\cdot)$ consists of terms in ε^2 and higher powers of ε, the coefficients of which involve the moments of $x^*(X(\omega))$. But since the moments of $x^*(X(\omega))$ are bounded by the corresponding moments of $||X(\omega)||$ which are all finite by (2.2), this $o(\cdot)$ holds uniformly for all x^* in $\mathscr{X}_1^*$. Now for ε small, if $\phi(t)$ denotes the moment generating function of $x^*(X(\omega))$,

$$\begin{aligned} \log \rho_1(x^*, \varepsilon) &= -\varepsilon t_1 + \log \phi(t_1) \\ &= -\varepsilon^2/\sigma^2 + \varepsilon^2/2\sigma^2 + o(\varepsilon^2) \\ &= -\varepsilon^2/2\sigma^2 + o(\varepsilon^2) \end{aligned} \tag{2.8}$$

where again $o(\cdot)$ is uniform in $x^* \in \mathscr{X}_1^*$ because of the reason given earlier. An expression similar to (2.8) is valid for the other term, namely $\rho_2(x^*, \varepsilon)$, and hence for their maximum $\rho(x^*, \varepsilon)$. Thus for ε sufficiently small

$$\log \rho(x^*, \varepsilon) = -\varepsilon^2/2\sigma^2 + o(\varepsilon^2)\,. \tag{2.9}$$

Now, since $o(\cdot)$ is uniform in $x^* \in \mathscr{X}_1^*$, by taking the supremum over $x^* \in \mathscr{X}_1^*$ on both sides of (2.9) we have the required result.

Three of the test statistics discussed in Section 3 viz., Ajne's A, Watson's and Rayleigh's tests are of the form

$$T_n = (2\pi n)^{-1} \textstyle\int_0^{2\pi} [\sum_{i=1}^n (f(\alpha + \alpha_i) - 1)]^2 d\alpha \tag{2.10}$$

where $(2\pi)^{-1} f \in L_2(0, 2\pi)$ is a density on unit circle, not uniform, and $\alpha_1, \cdots, \alpha_n$ are n independent random observations on the circumference. Let $\{c_m\}$ denote the Fourier coefficients of f relative to the basis $\{e^{im\alpha}, m = 0, \pm 1, \pm 2, \cdots\}$. Then we have

LEMMA 2.3. *Under the hypothesis of uniformity*, τ^2 *of* (2.7) *is given by*

$$\tau^2 = \max_m |c_m|^2\,. \tag{2.11}$$

Further, if $\{d_m\}$ *denote the Fourier coefficients of the alternative* (1.1), *then the local slope of* T_n *under the alternative, is given by*

$$c(\kappa) = \textstyle\sum_{m \neq 0} |c_m|^2 |d_m|^2 / \max_m |c_m|^2\,. \tag{2.12}$$

PROOF. Write

$$y_i(\alpha) = f(\alpha + \alpha_i) - 1 \tag{2.13}$$

and these will be treated as random variables in $L_2(0, 2\pi)$ · $\{T_n^{\frac{1}{2}}\}$ forms a standard test sequence. In order to obtain the slope function we need (P with subscript 'o' indicates probabilities under H_0 of uniformity)

$$\begin{aligned} P_0(T_n^{\frac{1}{2}} \geqq n^{\frac{1}{2}}\lambda) &= P_0\{(2\pi)^{-1} \textstyle\int_0^{2\pi} \bar{y}_n^2(\alpha) d\alpha \geqq \lambda^2\} \\ &= P_0\{||\bar{y}_n(\cdot)||^2 \geqq \lambda^2\} \end{aligned}$$

where

$$\bar{y}_n(\alpha) = n^{-1} \textstyle\sum_{i=1}^n y_i(\alpha) \quad \text{and} \quad ||h(\cdot)||^2 = (2\pi)^{-1} \int_0^{2\pi} h^2(\alpha) d\alpha$$

is the usual norm in $L_2(0, 2\pi)$. This approach allows us to use Lemmas 2.1

and 2.2 with $\mathscr{X}$ as $L_2(0, 2\pi)$ and $\mathscr{X}_1^*$, the space of all continuous linear functionals on $L_2(0, 2\pi)$ with norm unity. Corresponding to any element $h \in L_2(0, 2\pi)$ define the real-valued rv.

$$Z = (y(\cdot), h(\cdot)) = (2\pi)^{-1} \int_0^{2\pi} y(\alpha)h(\alpha)d\alpha . \tag{2.14}$$

To obtain

$$\tau^2 = \sup_{x^* \in \mathscr{X}_1^*} \operatorname{Var} x^*(y) = \sup_{||h||=1} \operatorname{Var}(Z) \tag{2.15}$$

under uniformity, we use the standard Fourier expansion methods. From the definition (2.13) the stochastic process $\{y(\alpha), 0 \leqq \alpha < 2\pi\}$ has mean function zero and covariance kernel $\operatorname{Cov}(y(\alpha), y(\beta)) = K(\alpha, \beta)$ say, under uniformity. Write

$$y(\alpha) = \sum_m X_m e^{im\alpha} \tag{2.16}$$

with

$$X_m = (2\pi)^{-1} \int_0^{2\pi} y(\alpha)e^{-im\alpha}d\alpha .$$

Clearly X_m has mean zero and variance

$$(4\pi^2)^{-1} \int_0^{2\pi} \int_0^{2\pi} K(\alpha, \beta)e^{-im\alpha}e^{-im\beta}d\alpha d\beta = (2\pi)^{-1}\lambda_m$$

where $\{\lambda_m\}$ are the eigen values corresponding to $K(\alpha, \beta)$. The eigenvalues of $K(\alpha, \beta)$ under the hypothesis are in fact given by $\lambda_m = |c_m|^2$ (see Beran [6]). Further if $\{a_m\}$ denote the Fourier coeffiicients of h with respect to the same basis

$$\operatorname{Var}(Z) = \operatorname{Var}(\textstyle\sum_m a_m X_m) = \sum_m |a_m|^2\lambda_m = \sum_m |a_m|^2|c_m|^2 . \tag{2.17}$$

Maximization of (2.17) subject to $||h||^2 = \sum_m |a_m|^2 = 1$ occurs clearly when $a_m = 1$ corresponding to $\max_m |c_m|^2$ and hence

$$\tau^2 = \max_{\sum|a_m|^2=1} \textstyle\sum_m |a_m|^2|c_m|^2 = \max_m |c_m|^2 . \tag{2.18}$$

Now for the standard sequence $\{T_n^{\frac{1}{2}}\}$, from Lemma 2.2 and (2.11)

$$\begin{aligned} n^{-1} \log P_0\{T_n^{\frac{1}{2}} \geqq n^{\frac{1}{2}}\lambda\} &= n^{-1} \log P_0\{||\bar{y}_n(\cdot)||^2 \geqq \lambda^2\} \\ &\to \log \rho(\mathscr{X}_1^*, \lambda) = -\lambda^2/2\tau^2 + o(\lambda^2) \\ &= -\lambda^2/2 \max_m |c_m|^2 + o(\lambda^2) . \end{aligned} \tag{2.19}$$

Further it can be checked that (c.f. Beran [6])

$$T_n/n \to_{\text{a.s.}} \textstyle\sum_{m\neq 0} |c_m|^2|d_m|^2 \tag{2.20}$$

where $\{c_m\}$ and $\{d_m\}$ are respectively the Fourier coefficients of $y(\alpha)$ and $g(\alpha)$. Thus from (2.19) and (2.20) the local slope of the test sequence $\{T_n^{\frac{1}{2}}\}$ is given by (2.12).

3. Ajne's A, Watson's and Rayleigh's tests. Lemma 2.3 can now be applied to the following three special cases.

(i) *Ajne's test* A. Given n independent observations on the circumference of a unit circle, let $N(\alpha)$ denote the number of observations in the half circle $[\alpha, \alpha + \pi)$ taking say, the clockwise direction as positive. Then Ajne [2] proposed the statistic

$$A_n = (2\pi n)^{-1} \int_0^{2\pi} [N(\alpha) - n/2)]^2 d\alpha \tag{3.1}$$

for testing uniformity on the circle. Here corresponding to the ith observation α_i, if we define

$$Y_i(\alpha) = f(\alpha + \alpha_i) - 1 = \begin{cases} \tfrac{1}{2} & \text{if } \alpha \leqq \alpha_i < \alpha + \pi \\ -\tfrac{1}{2} & \text{if } \alpha + \pi \leqq \alpha_i < \alpha + 2\pi , \end{cases} \tag{3.2}$$

A_n becomes a special case of T_n defined in (2.10). $Y_i(\alpha)$ given in (3.2) have Fourier coefficients

$$c_m = 1/\pi i m \quad \text{if } m \text{ odd}$$
$$= 0 \quad \text{otherwise}$$

while the Fourier coefficients of (1.1) are given by

$$d_m = I_m(\kappa)/I_0(\kappa) , \qquad m = \pm 1, \pm 2, \cdots . \tag{3.3}$$

Therefore the local slope (see (2.12)) of $T_n^{(1)} = A_n^{\frac{1}{2}}$ is given by

$$c_1(\kappa) = 2 \sum_{m=0}^{\infty} I^2_{2m+1}(\kappa)/(2m + 1)^2 I_0^2(\kappa) \tag{3.4}$$

(ii) *Watson's test.* Watson [16] proposed the statistic

$$W_n = n \int_0^{2\pi} \{F_n(\alpha) - \alpha/2\pi - \int_0^{2\pi} [F_n(\beta) - \beta/2\pi] d\beta/2\pi\}^2 d\alpha/2\pi \tag{3.5}$$

which can be used for testing uniformity on the circle since its value remains independent of the choice of the arbitrary point from which we begin cumulating the probability density or the masses corresponding to the sample points. This is a special case of T_n, with

$$Y_i(\alpha) = I_i(\alpha) - \alpha/2\pi - \int_0^{2\pi} [F_n(\beta) - \beta/2\pi] d\beta/2\pi$$

where $I_i(\alpha) = 1$ if $\alpha_i \leqq \alpha$ and $=0$ otherwise. The Fourier coefficients corresponding to this satisfy (c.f. Watson [16]) $|c_m|^2 = (4m^2\pi^2)^{-1}, m = \pm 1, \pm 2, \cdots$ while $\{d_m\}$ are again as given in (3.3). Thus the local slope of $T_n^{(2)} = W_n^{\frac{1}{2}}$ is given by (see equation (2.12))

$$c_2(\kappa) = 2 \sum_{m=1}^{\infty} I_m^2(\kappa)/m^2 I_0^2(\kappa) . \tag{3.6}$$

(iii) *Rayleigh's test.* Let us look at each observation α_i on the circumference as a unit vector with components $(\cos \alpha_i, \sin \alpha_i)$. Then the classical Rayleigh's test for uniformity is based on the length of the vector resultant

$$R_n^2 = (\textstyle\sum_1^n \cos \alpha_i)^2 + (\sum_1^n \sin \alpha_i)^2 . \tag{3.7}$$

The hypothesis of uniformity is rejected when R_n is too large. The exact null distribution of R_n is given by the pdf

$$f(r) = r \int_0^{\infty} J_0^n(t) J_0(rt) t\,dt \qquad \text{for } 0 \leqq r \leqq n . \tag{3.8}$$

Further discussion and critical points for testing can be found in Greenwood and Durand [8]. This is again a special case of T_n if we set

$$Y_i(\alpha) = 2^{\frac{1}{2}} \cos(\alpha - \alpha_1) .$$

The Fourier coefficients corresponding to this are

$$\begin{aligned} c_m &= 2^{-\frac{1}{2}} && \text{if} \quad m = \pm 1 \\ &= 0 && \text{otherwise} \end{aligned}$$

so that from (2.12), the local slope of $T_n^{(3)} = n^{-\frac{1}{2}} R_n$ is given by

$$c_3(\kappa) = 2I_1^2(\kappa)/I_0^2(\kappa) , \tag{3.9}$$

4. Ajne's test N. If $\mathrm{N}(\alpha)$ is as defined in Section 3 then Ajne [2] also proposes

$$\mathrm{N} = \sup_{0 \leqq \alpha < 2\pi} \mathrm{N}(\alpha) \tag{4.1}$$

which is the maximum number of points in any semicircle, for testing uniformity. We consider here the standard sequence

$$T_n^{(4)} = (\mathrm{N} - n/2)/n^{\frac{1}{2}} = n^{\frac{1}{2}}[\mathrm{N}/n - \tfrac{1}{2}] = \mathrm{N}_n^* , \quad \text{say} \tag{4.2}$$

Then we require the probability

$$P_0\{T_n^{(4)} \geqq n^{\frac{1}{2}}\lambda\} = P_0(\mathrm{N}/n - \tfrac{1}{2}) \geqq \lambda\} . \tag{4.3}$$

We shall obtain this by getting upper and lower bounds for this probability. Since for any fixed α

$$\mathrm{N} = \sup_\alpha \mathrm{N}(\alpha) \geqq \mathrm{N}(\alpha) , \tag{4.4}$$

$$n^{-1} \log P_0\{(\mathrm{N}/n - \tfrac{1}{2}) \geqq \lambda\} \geqq n^{-1} \log P_0\{(\mathrm{N}(\alpha)/n - \tfrac{1}{2}) \geqq \lambda\} .$$

But $\mathrm{N}(\alpha)$ under the hypothesis has a binomial distribution with parameter $\frac{1}{2}$. Therefore,

$$n^{-1} \log P_0\{\mathrm{N}/n - \tfrac{1}{2}) \geqq \lambda\} \geqq \log \rho^*(\tfrac{1}{2}, \lambda) \tag{4.5}$$

where $\rho^*(\frac{1}{2}, \lambda)$ is as defined in (24) of Sethuraman [13]. In order to get an upper bound for the probability in (4.3), for some $\partial > 0$ (to be chosen suitably later on), we divide the whole length of the circumference into $N(\partial) = [2\pi/\partial] + 1$ arcs of length ∂ each and define

$$\mathrm{N}(\alpha, \partial) = \text{number of observations in the arc} \quad [\alpha, \alpha + \pi + \partial) . \tag{4.6}$$

Then, clearly

$$\mathrm{N}(\alpha, \partial) \geqq \mathrm{N}(\alpha) \tag{4.7}$$

for all α. Further since any α lies between $[r\partial, \overline{r+1}\,\partial)$ for some r, the corresponding $\mathrm{N}(\alpha) \leqq \mathrm{N}(r\partial, \partial)$ so that

$$\mathrm{N} = \sup_\alpha \mathrm{N}(\alpha) \leqq \max_r \mathrm{N}(r\partial, \partial) . \tag{4.8}$$

Thus

$$P_o\{[\mathrm{N}/n - \tfrac{1}{2}] \geqq \lambda\} \leqq P_o\{\max_r [\mathrm{N}(r\partial, \partial)/n - \tfrac{1}{2}] \geqq \lambda\}$$
$$\text{(4.9)} \qquad = P_o\{[\mathrm{N}(r\partial, \partial)/n - \tfrac{1}{2}] \geqq \lambda\,, \quad \text{for at least one} \quad r\}$$
$$\leqq \textstyle\sum_{r=1}^{N(\partial)} P_o\{\mathrm{N}[(r\partial, \partial)/n - \tfrac{1}{2}] \geqq \lambda\}\,.$$

Now $\mathrm{N}(\alpha, \partial)$ has a binomial distribution with parameter $(\frac{1}{2} + \partial)$ and all the terms in the summation (4.9) are equal so that

$$\text{(4.10)} \qquad P_o\{(\mathrm{N}/n - \tfrac{1}{2}) \geqq \lambda\} \leqq \mathrm{N}(\partial) \cdot P_o\{\mathrm{N}(r\partial, \partial)/n - \tfrac{1}{2} - \partial \geqq \lambda - \partial\}$$
$$= \mathrm{N}(\partial) \cdot \rho^{*n}(\tfrac{1}{2} + \partial, \lambda - \partial)$$

where $\rho^*(p, \lambda)$ is again as defined in (24) of Sethuraman [13]. Choosing $\partial = 1/n$ the term on the RHS of (4.10) is $2\pi n.\ \ \rho^{*n}(\frac{1}{2} + 1/n, \lambda - 1/n)$ so that

$$\text{(4.11)} \quad \limsup\, (1/n) \log P_o\{[N/n - \tfrac{1}{2}] \geqq \lambda\} \leqq \limsup \log \rho^*(\tfrac{1}{2} + 1/n, \lambda - 1/n)$$
$$= \log \rho^*(\tfrac{1}{2}, \lambda)\,.$$

Thus from (4.5) and (4.11) we have

$$\text{(4.12)} \qquad n^{-1} \log P_o\{(\mathrm{N}/n - \tfrac{1}{2}) \geqq \lambda\} \to \log \rho^*(\tfrac{1}{2}, \lambda) = -2\lambda^2 + o(\lambda^2)\,.$$

In order to get the probability limit to which $T_n^{(4)}/n^{\frac{1}{2}}$ converges under the CN alternatives, we observe that since $N(\alpha)$ is a binomial sum, for any fixed α

$$\text{(4.13)} \qquad \mathrm{N}(\alpha)/n \to_p p(\alpha) = P\{\alpha \leqq \theta < \alpha + \pi\}$$

where θ is a CN random variable with density (1.1). This convergence is uniform in α since the random variables involved in the summation $N(\alpha)$ are clearly uniformly bounded in n and α (refer Parzen [11]). Thus

$$\mathrm{N}/n = \sup \mathrm{N}(\alpha)/n \to_p \sup_{-\pi \leqq \alpha \leqq \pi} p(\alpha)\,.$$

For the CN density defined in (1.1), this supremum is attained when $\alpha = -\pi/2$ so that

$$\text{(4.14)} \qquad \mathrm{N}/n \to_p [2\pi I_o(\kappa)]^{-1} \textstyle\int_{-\pi/2}^{\pi/2} e^{\kappa \cos \alpha} d\alpha = \mathrm{A}(\kappa)$$

where $\mathrm{A}(\kappa)$ is the probability that a CN rv lies within $\pi/2$ from its mean direction. This can be expressed in the following series

$$\text{(4.15)} \qquad \mathrm{A}(\kappa) = [2\pi I_o(\kappa)]^{-1} \sum_{r=o}^{\infty} \frac{\kappa^r}{r!} B\left(\frac{r+1}{2}, \frac{1}{2}\right)\,.$$

Hence

$$T_n^{(4)}/n^{\frac{1}{2}} = (\mathrm{N}/n - \tfrac{1}{2}) \to_p [\mathrm{A}(\kappa) - \tfrac{1}{2}|$$

and the local slope of the sequence $\{T_n^{(4)}\}$ is given by

$$\text{(4.16)} \qquad c_4(\kappa) = [2\mathrm{A}(\kappa) - 1]^2\,.$$

We remark in passing, that the statistic N is similar to the Hodges' bivariate sign-test statistic [9], where for testing the equality of the bivariate df's, he proposes the maximum number of vector differences (differences between the observed vectors in the two populations) with positive projections on some line through the origin, as the direction of this line is varied.

5. Kuiper's test. Kuiper [10] proposed a variant of the Kolmogorov-Smirnov statistic for testing the hypothesis that the observations come from a population with specified df $F(x)$. If $F_n(x)$ denotes the empirical df, then Kuiper's statistic is

$$\text{(5.1)} \qquad \mathrm{V}_n = n^{\frac{1}{2}}\{\sup_x [F_n(x) - F(x)] - \inf_x [F_n(x) - F(x)]\}$$

which is specially suited for testing goodness of fit on the circle, as the statistic V_n is independent of the origin used for measurement of α. We take

$$\text{(5.2)} \qquad T_n^{(5)} = \mathrm{V}_n$$

as the standard sequence. Abrahamson [1] computed the exact slope of this test sequence. Under an alternative df $G(\alpha)$, it turns out that the local slope is

$$\text{(5.3)} \qquad c_5(G) = 4\{\sup_\alpha [G(\alpha) - F(\alpha)] - \inf_\alpha [G(\alpha) - F(\alpha)]\}^2 .$$

When $G(\alpha)$ is the CN alternative (1.1) and the null distribution $F(\alpha)$ is uniform, it is easy to check that $\sup_{-\pi \leqq \alpha \leqq \pi}[G(\alpha) - F(\alpha)]$ is attained at $\alpha = \pi/2$ and the $\inf_{-\pi \leqq \alpha \leqq \pi}[G(\alpha) - F(\alpha)]$ at $\alpha = -\pi/2$. Thus from (5.3) the local slope of $\{T_n^{(5)}\}$ is given by

$$\text{(5.4)} \qquad c_5(\kappa) = [2\mathrm{A}(\kappa) - 1]^2$$

where $\mathrm{A}(\kappa)$ is as defined in the preceding section. This slope is identical to that of Ajne's N.

6. The spacings test U_n. Let $\{D_i, i = 1, \cdots, n\}$ denote the lengths of the n arcs between successive sample points on the circumference (usually referred to as spacings). The author [12] has studied various tests based on these arc lengths for testing uniformity on the circle. From among the class of such spacings' tests, let us consider the following test based on the statistic

$$\text{(6.1)} \qquad \begin{aligned} \mathrm{U}_n &= \textstyle\sum_{i=1}^n \max [D_i - 2\pi/n, 0] \\ &= \tfrac{1}{2} \textstyle\sum_{i=1}^n |D_i - 2\pi/n| . \end{aligned}$$

In this case we take the test sequence

$$\text{(6.2)} \qquad T_n^{(6)} = n^{\frac{1}{2}}[\mathrm{U}_n - 2\pi/e]/2\pi[2e^{-1} - 5e^{-2}]^{\frac{1}{2}}$$

as the standard sequence. It can be shown (see for example Rao [12] and Sherman [15]) that this has an asymptotic normal distribution with mean zero and unit variance, so that this sequence of test statistics satisfies (43) of [5] with $a = 1$. Further, from Rao [12] it may be seen that under the alternatives (1.1)

$$
\begin{aligned}
U_n &\to_p \int_0^{2\pi} e^{-2\pi g}(\alpha)d\alpha \\
&= \int_0^{2\pi} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!}\left(\frac{e^{\kappa\cos\theta}}{I_0(\kappa)}\right)^j \\
&= 2\pi \sum_{j=0}^{\infty} (-1)^j I_0(j\kappa)/j!\, I_0^j(\kappa)\,.
\end{aligned} \tag{6.3}
$$

Thus the approximate slope of this standard sequence $\{T_n^{(6)}\}$ is given by

$$
c_6^{(a)}(\kappa) = \left[\sum_{j=0}^{\infty} \frac{(-1)^j}{j!}\frac{I_0(j\kappa)}{I_0^j(\kappa)} - e^{-1}\right]^2 / [2e^{-1} - 5e^{-2}]\,. \tag{6.4}
$$

7. Comparison of the local efficiencies. In this section we compare the limiting efficiencies of the six test statistics that have been considered in Sections 3 to 6 on the basis of the slopes given in (3.4), (3.6), (3.9), (4.16), (5.4) and (6.4) respectively.

The comparison of the limiting efficiencies is made easier by considering approximations for the slopes when κ is small, since in any case we let κ tend to zero for obtaining these efficiencies. Throughout this section the symbol $\sim$ is used to denote that the ratio of the 2 sides approaches one as κ tends to zero. Since the Bessel function $I_n(\kappa)$ has the expansion

$$
I_n(\kappa) = \frac{(\kappa/2)^n}{\Gamma(n+1)}\left\{1 + \frac{(\kappa/2)^2}{1\cdot(n+1)} + \frac{(\kappa/2)^4}{1\cdot 2(n+1)(n+2)} \cdots\right\}, \tag{7.1}
$$

ignoring terms involving Bessel functions of higher orders in (3.4),

$$
c_1(\kappa) \sim 2[I_1(\kappa)/I_0(\kappa)]^2 \sim \kappa^2/2\,. \tag{7.2}
$$

Similarly

$$
c_2(\kappa) \sim \kappa^2/2 \tag{7.3}
$$

and

$$
c_3(\kappa) \sim \kappa^2/2\,. \tag{7.4}
$$

For small κ, the expression for $A(\kappa)$ given in (4.15) can be approximated by

$$
A(\kappa) \sim \tfrac{1}{2} + \kappa/\pi
$$

so that

$$
c_4(\kappa) = c_5(\kappa) = [2A(\kappa) - 1]^2 \sim 4\kappa^2/\pi^2\,.
$$

Finally the probability limit of U_n given in (6.3) can be shown for small κ, to be

$$
2\pi e^{-1}[1 + (\kappa/2)^2] + o(\kappa^2)
$$

and therefore the approximate slope of U_n

$$
\begin{aligned}
c_6^{(a)}(\kappa) &\sim [e^{-1}(\kappa/2)^2]^2/(2e^{-1} - 5e^{-2}) \\
&= \kappa^4/16(2e - 5)\,.
\end{aligned}
$$

The following table gives the local efficiencies of the six test statistics studied in this paper. Here $L_{X,Y}$ denotes the local efficiency of the test sequence $\{X\}$ relative to the test sequence $\{Y\}$ when κ goes to zero.

TABLE 1
Table of local efficiencies

$L_{X,Y}$	Y → X	A	W	R	V	N	U
	A	1	1	1	$\pi^2/8$	$\pi^2/8$	∞
	W		1	1	$\pi^2/8$	$\pi^2/8$	∞
	R			1	$\pi^2/8$	$\pi^2/8$	∞
	V				1	1	∞
	N					1	∞
	U						1

Thus the three tests, Rayleigh's R, Watson's W and Ajne's A turn out to have the same limiting performance for testing uniformity against the CN alternatives. Kuiper's V and Ajne's N have also asymptotically identical performances but they do not fare as well as the first three. If the efficiency of R, W or A is taken to be unity, then V and N have an efficiency of 81%. That the asymptotic efficiency of the spacings test U is zero as compared with the other tests against the CN alternatives, is not altogether surprising in view of the results obtained in Sethuraman and Rao [14]. However one need not abandon the symmetric spacings tests because of this, since these local asymptotic Bahadur efficiencies seldom throw sufficient light on the relative powers of the tests in small samples with which most practical investigations are concerned. A modest simulation study was undertaken to assess the small sample performance of the spacings test U as compared to Rayleigh's test R. The Rayleigh's test was chosen for comparison since it is the likelihood ratio test for testing uniformity in circular normal populations and is by far the best test for the situation. 550 samples of size 10 each were generated from the circular normal distributions (CND's) with concentration parameter $\kappa = 1$ and $\kappa = 3$. Rayleigh's statistic R and the spacings statistic U were computed for each sample and compared with the corresponding 5% and 1% critical points, obtained from Greenwood and Durand [8] and Rao [12]. The proportion of samples which these two tests reject at these two levels of significance are shown in Table 2.

From these comparisons we find that the small sample power of the spacings test U as compared to that of Rayleigh's test, is not as bad as the limiting efficiency seems to indicate, even for observations from the CND. Besides its satisfactory small sample power, the spacings test U detects clustering of any sort and is a valid test for a much wider class of alternatives than Rayleigh's

TABLE 2

Monte Carlo powers of the tests R and U

CND *with* $\kappa = 1$ (*sample size* $n = 10$)			CND *with* $\kappa = 3$ (*sample size* $n = 10$)		
Test statistics	Level of significance		Test statistics	Level of significance	
	5%	1%		5%	1%
Rayleigh's test, R	0.4073	0.1891	Rayleigh's test, R	0.9854	0.9636
Spacings test, U	0.2636	0.0927	Spacings test, U	0.9636	0.8200

test. It may however be noted that Watson's test for example, is also valid for this wider class of alternatives and is likely to be preferable to the spacings test.

8. A simple inequality between Kuiper's V and Ajne's N. As remarked in Section 5, the local slopes of Kuiper's V and Ajne's N are identical. In this section we note that Kuiper's V_n defined in (5.1) is larger than Ajne's $N_n{}^*$ defined in (4.2) for all samples. If $F_n(\alpha)$ denotes the empirical df as measured from some point, define $F_n{}^*(\alpha)$ and $F^*(\alpha)$ on the interval $[0, 3\pi)$ as follows.

(8.1)
$$\begin{aligned} F_n{}^*(\alpha) &= F_n(\alpha) && \text{for } 0 \leqq \alpha < 2\pi \\ &= 1 + F_n(\alpha - 2\pi) && \text{for } 2\pi \leqq \alpha < 3\pi \end{aligned}$$

and

(8.2)
$$F^*(\alpha) = \alpha/2\pi \qquad \text{for } 0 \leqq \alpha < 3\pi .$$

Then

$$\begin{aligned} N(\alpha) &= \text{number of observations in } [\alpha, \alpha + \pi) \\ &= n[F_n{}^*(\alpha + \pi) - F_n{}^*(\alpha)] , \qquad \text{for } 0 \leqq \alpha < 2\pi \end{aligned}$$

and therefore

$$\begin{aligned} N^* &= n^{\frac{1}{2}}[N/n - \tfrac{1}{2}] \\ &= n^{\frac{1}{2}} \{\sup_{0 \leqq \alpha < 2\pi} [F_n{}^*(\alpha + \pi) - F_n{}^*(\alpha) - \tfrac{1}{2}]\} \\ &= n^{\frac{1}{2}} \{\sup_{0 \leqq \alpha < 2\pi} [F_n{}^*(\alpha + \pi) - F^*(\alpha + \pi) + F^*(\alpha) - F_n{}^*(\alpha)]\} \\ &\leqq n^{\frac{1}{2}} \{\sup_{0 \leqq \alpha < 2\pi} [F_n{}^*(\alpha + \pi) - F^*(\alpha + \pi)] \\ &\qquad - \inf_{0 \leqq \alpha < 2\pi} [F_n{}^*(\alpha) - F^*(\alpha)]\} \\ &= n^{\frac{1}{2}} \{\sup_{0 \leqq \alpha < 2\pi} [F_n(\alpha) - F(\alpha)] - \inf_{0 \leqq \alpha < 2\pi} [F_n(\alpha) - F(\alpha)]\} \\ &= V_n . \end{aligned}$$

Thus the value of Ajne's $N_n{}^*$ cannot exceed that of Kuiper's statistic V_n for any sample. From a consideration of the limiting distributions and other facts mentioned in Sections 4 and 5, the approximate slopes of V_n and $N_n{}^*$ are seen to be equal and the two statistics have asymptotic distributions which are of the same exponential order in the tails. But since $V_n \geqq N_n{}^*$, the limiting distributions must be quite dissimilar in their main parts. Thus in view of the fact that the tails and main parts of the sequence of distributions do not necessarily have the same limiting properties, one might doubt the reliability of the Bahadur comparison between V_n and N_n^*. However, as Abrahamson [1] puts it, the Bahadur efficiency concerns itself with how well the null hypothesis explains the sequences of test statistics when in fact the hypothesis is false and the statistics are growing roughly in proportion to $n^{\frac{1}{2}}$. Thus the fact that Kuiper's V_n exceeds Ajne's N_n^* in value, assures us in view of the equality of the slopes, that V_n "attains a smaller level of significance" than does N_n^* and is preferable to N_n^*.

Acknowledgments. I wish to thank Professor C. R. Rao of the Indian Statistical Institute for his encouragement during the course of this work, Professor J. Sethuraman for valuable discussions, and the referee for helpful suggestions.

REFERENCES

[1] ABRAHAMSON, I. G. (1967). Exact Bahdaur efficiencies for the Kolmogorov-Smirnov and Kuiper one and two-sample statistics. *Ann. Math. Statist.* **38** 1475-1490.

[2] AJNE, B. (1968). A simple test for uniformity of a circular distribution. *Biometrika* **55** 343-354.

[3] BAHADUR, R. R. (1960). On the asymptotic efficiency of tests and estimates. *Sankhyā* **22** 229-252.

[4] BAHADUR, R. R. (1960). Stochastic comparison of tests. *Ann. Math. Statist.* **31** 276-295.

[5] BAHADUR, R. R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38** 303-324.

[6] BERAN, R. J. (1969). Asymptotic theory of a class of tests for uniformity of a circular distribution. *Ann. Math. Statist.* **40** 1196-1206.

[7] GLESER, L. J. (1964). On a measure of test efficiency proposed by R. R. Bahadur. *Ann. Math. Statist.* **35** 1537-1544.

[8] GREENWOOD, J. A. and DURAND, D. (1955). The distribution of the length and components of the sum of *n* random unit vectors. *Ann. Math. Statist.* **26** 233-246.

[9] HODGES, J. L. (1955). A bivariate sign test. *Ann. Math. Statist.* **26** 523-527.

[10] KUIPER, N. H. (1960). Tests concerning random points on a circle. *Nederl. Akad. Wetensch. Proc. Ser. A.* **63** 38-47.

[11] PARZEN, E. (1954). On uniform convergence of families of sequences of random variables. *Univ. of California Publ. Statist.* **2** 23-54.

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

[13] SETHURAMAN, J. (1964). On the probability of large deviations of families of sample means. *Ann. Math. Statist.* **35** 1304-1316.

[14] SETHURAMAN, J. and RAO, J. S. (1969). Pitman efficiencies of tests based on spacings in *Nonparametric Techniques in Statistical Inference*, ed. M. L. Puri. Cambridge Univ. Press., 405-415.

[15] SHERMAN, B. (1950). A random variable related to the spacing of sample values. *Ann. Math. Statist.* **21** 339-361.

[16] WATSON, G. S. (1961). Goodness of fit tests on a circle. *Biometrika* 48 109-114.