The Annals of Mathematical Statistics
1969, Vol. 40, No. 4, 1196-1206

ASYMPTOTIC THEORY OF A CLASS OF TESTS FOR UNIFORMITY OF A CIRCULAR DISTRIBUTION[1]

By R. J. Beran[2]

The Johns Hopkins University

0. Summary. Let $(x_1, x_2, \cdots, x_n)$ be independent realizations of a random variable taking values on a circle C of unit circumference, and let

$$T_n = n^{-1} \int_0^1 [\sum_{j=1}^n f(x + x_j) - n]^2 \, dx,$$

where $f(x)$ is a probability density on C, $f \varepsilon L_2[0, 1]$, and the addition $x + x_j$ is performed modulo 1. T_n is used to test whether the observations are uniformly distributed on C. It includes as special cases several other statistics previously proposed for this purpose by Ajne, Rayleigh and Watson. The main results of the paper are the asymptotic distributions of T_n under fixed alternatives to uniformity and under sequences of local alternatives to uniformity. A characterization is found for those alternatives against which T_n, with specified $f(x)$, gives a consistent test. The approximate Bahadur slope of T_n is calculated from the asymptotic null distribution; however, an example indicates that this slope may not always reflect the power of T_n reliably. A Monte Carlo simulation for a special case of T_n suggests that a fair approximation to the power of T_n may be obtained from its mean and variance under the alternative.

1. Introduction. Suppose $(x_1, x_2, \cdots, x_n)$ are independent realizations of a random variable which is distributed about a circle C of unit circumference and has a density on C of the form

$$g(x \mid k) = 1 + k[f(x + a) - 1], \quad x \varepsilon [0, 1], \quad k \varepsilon [0, 1]. \tag{1.1}$$

Here a is an unknown location parameter, $f \varepsilon L_2[0, 1]$ is a density on C, $f(x) \not\equiv 1$, and the argument $x + a$ is to be interpreted modulo 1. Then $g(x \mid 0) = 1$ while $g(x \mid 1) = f(x + a)$. An argument similar to that in Beran [3] shows that for testing $H_0{:}k = 0$ (uniformity) versus $H_1{:}k > 0$, a locally (small k) most powerful invariant (under rotation) test is to reject H_0 when

$$T_n = n^{-1} \int_0^1 [\sum_{j=1}^n f(x + x_j) - n]^2 \, dx \tag{1.2}$$

is too large.

This result, which generalizes earlier work by Ajne [1], is the motivation for the present study of the limiting distributions of T_n under arbitrary fixed alternatives and under sequences of local alternatives to uniformity. It is worth noting

Received 12 February 1968; revised 31 January 1969.

[1] Research supported by the Office of Naval Research under contract NONR 4010(09) awarded to the Department of Statistics, The Johns Hopkins University. This paper, in whole or in part, may be reproduced for any purpose of the United States Government.

[2] Now at the University of California, Berkeley.

that T_n includes as special cases four statistics which have previously been proposed as tests for uniformity on the circle. The earliest of these, studied by Rayleigh [8], is the statistic

$$R_n^2 = n^{-1}[(\sum_{j=1}^n \cos 2\pi x_j)^2 + (\sum_{j=1}^n \sin 2\pi x_j)^2]. \tag{1.3}$$

The second is

$$U_n^2 = n \int_0^1 [F_n(x) - x - \int_0^1 [F_n(y) - y]\, dy]^2\, dx, \tag{1.4}$$

an analogue of the Cramér-von Mises statistic introduced by Watson in [10]. $F_n(x)$ is the sample distribution function cumulated from an arbitrary origin. The third statistic,

$$A_n = n^{-1} \int_0^1 [N(x) - n/2]^2\, dx, \tag{1.5}$$

was derived by Ajne [1] as a locally most powerful invariant test against alternatives with density $2p$ on one semi-circle, $2q$ on the complementary semi-circle, $p + q = 1$. $N(x)$ is the number of observations lying in the semi-circle $[x, x + \frac{1}{2})$. Also due to Watson [12] is the fourth statistic

$$S_n^2 = n^{-1} \int_0^1 [f_n(x) - 1]^2\, dx, \tag{1.5$'$}$$

where $f_n(x)$ is a consistent estimator of the true density generating the observations. For example,

$$f_n(x) = 1 + 2 \sum_{m=1}^{N(n))} \sum_{j=1}^n \cos [2\pi m(x - x_j)], \tag{1.6}$$

where the integer $N(n)$ is $o(n)$.

To show that R_n^2, U_n^2 and S_n^2 are special cases of T_n, as A_n clearly is, let $\{c_m\}$ denote the Fourier coefficients of $f(x)$ relative to the basis $\{e^{2\pi imx}; m = 0, \pm 1, \pm 2, \cdots\}$. Then, applying Parseval's theorem to (1.2) yields

$$T_n = n^{-1} \sum_{m \neq 0} |c_m|^2 \,|\sum_{j=1}^n e^{2\pi imx_j}|^2. \tag{1.7}$$

Similar Fourier analyses show that

$$\begin{aligned} R_n^2 &= \tfrac{1}{2} n^{-1} [|\sum_{j=1}^n e^{2\pi ix_j}|^2 + |\sum_{j=1}^n e^{-2\pi ix_j}|^2], \\ U_n^2 &= n^{-1} \sum_{m \neq 0} (2\pi m)^{-2} \,|\sum_{j=1}^n e^{2\pi imx_j}|^2, \\ S_n^2 &= n^{-1} \sum_{|m|=1}^{N(n)} |\sum_{j=1}^n e^{2\pi imx_j}|^2. \end{aligned} \tag{1.8}$$

Thus, T_n and U_n^2 generate equivalent tests provided that for some $\alpha > 0$, $\{|c_m^2| = \alpha m^{-2}; m = \pm 1, \pm 2 \cdots\}$. The phase of each Fourier component is left unspecified. In particular U_n^2 gives a test for uniformity which is most powerful invariant against local (small k) alternative densities of the form $g(x \mid k) = 1 + k(2x - 1)$. R_n^2 and S_n^2 may be treated analogously.

Another expression for T_n, derivable from (1.7), is

$$T_n = n^{-1} \sum_{j=1}^n \sum_{k=1}^n d(x_j - x_k), \tag{1.9}$$

where the function $d(x)$ is defined by

$$d(x) = \sum_{m\neq 0} |c_m|^2 e^{2\pi imx} = 2 \sum_{m=1}^{\infty} |c_m|^2 \cos 2\pi mx. \tag{1.10}$$

The series in (1.10) converge absolutely and uniformly, since $f \varepsilon L_2[0, 1]$. Thus, $d(x)$ is continuous, periodic with period 1, and symmetric with respect to 0 and $\frac{1}{2}$. Closed expressions for $d(x)$, often available, make (1.9) useful as a computational formula. The formula (1.9) is also a starting point in relating the present tests for uniformity to class of two-sample tests on the circle introduced by Schach [9].

2. Limiting distributions. The first topic of this section is the asymptotic distribution of T_n when the observations $(x_1, x_2, \cdots x_n)$ are generated by a random variable with distribution function $G(x)$, relative to an arbitrary origin on C. Consider the stochastic process

$$\xi_n(x) = n^{-\frac{1}{2}} \sum_{j=1}^{n} [f(x + x_j) - 1], \tag{2.1}$$

which is defined for $x \varepsilon [0, 1]$. For $y, y_1, y_2 \varepsilon [0, 1]$, let $\mu_n(y)$ and $B(y_1, y_2)$ denote, respectively, the mean and covariance kernel of $\xi_n(y)$. Clearly, $\mu_n(y) = n^{\frac{1}{2}}b(y)$, where $b(y) = \int_0^1 [f(y + x) - 1]\, dG(x)$, and

$$B(y_1, y_2) = \int_0^1 [f(y_1 + x) - 1][f(y_2 + x) - 1]\, dG(x) - b(y_1)b(y_2),$$

which is symmetric and does not depend upon n.

The distribution function $G(x)$ has the Fourier transforms

$$d_m = \int_0^1 e^{-2\pi imx}\, dG(x); \qquad m = 0, \pm 1, \pm 2, \cdots; \tag{2.2}$$

where $|d_m| \leqq 1$ for all m. Using Fubini's theorem, the Fourier coefficients of $b(x)$ are $\{c_m \bar{d}_m\}$ if $m \neq 0$ and 0 if $m = 0$; the $\{c_m\}$ denote here, as in the Introduction, the Fourier coefficients of $f(x)$. It is clear that $b(x) \varepsilon L_2[0, 1]$.

Further calculation and use of Fubini's theorem show that, relative to the double orthonormal basis $\{e^{2\pi i(my_1 - ly_2)}; m, l = 0, \pm 1, \pm 2, \cdots\}$, $B(y_1, y_2)$ has the Fourier coefficients $\{(d_{l-m} - d_l\bar{d}_m)\bar{c}_l c_m\}$. Consequently,

$$B(y_1, y_2) \varepsilon L_2([0, 1) \times [0, 1])$$

and, indeed, has finite trace.

Let $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq 0$ be the eigenvalues (including 0) of $B(y_1, y_2)$ as defined by the integral equation

$$\int_0^1 B(y_1, y_2)\varphi(y_2)\, dy_2 = \lambda\varphi(y_1). \tag{2.3}$$

Let $\varphi_1(x), \varphi_2(x), \cdots$ be a corresponding orthonormal sequence of eigenfunctions. Since $B(y_1, y_2)$ is a covariance kernel, the eigenvalues are non-negative. There are at most a countable number of eigenfunctions corresponding to the eigenvalue zero. Both $\sum_k \lambda_k^2$ and $\sum_k \lambda_k$ converge. The eigenfunctions form a complete orthonormal system in $L_2[0, 1]$.

THEOREM 1. *If, under the distribution function* $G(x)$, $b(x) \not\equiv 0$, *then the asymp-*

totic distribution of $n^{-\frac{1}{2}}[T_n - n\int_0^1 b^2(x)\,dx]$ *is normal with mean* 0 *and variance* $4\int_0^1\int_0^1 B(x, y)b(x)b(y)\,dx\,dy$. *Provided the variance does not vanish, the convergence to the asymptotic distribution function as* $n \to \infty$ *is uniform on the real line.*

PROOF. Since $T_n = \int_0^1 \xi_n^2(x)\,dx$, let

$$\begin{aligned} X_n &= n^{-\frac{1}{2}}[T_n - n\int_0^1 b^2(x)\,dx] \\ &= \int_0^1 [n^{-\frac{1}{2}}\xi_n(x) + b(x)][\xi_n(x) - n^{\frac{1}{2}}b(x)]\,dx, \\ Y_n &= \int_0^1 2b(x)[\xi_n(x) - n^{\frac{1}{2}}b(x)]\,dx. \end{aligned} \tag{2.4}$$

From the definition (2.1) of $\xi_n(x)$, Y_n can be expressed as a sum of normalized iid random variables, so that the asymptotic distribution of Y_n is normal with mean 0 and variance $4\int_0^1\int_0^1 B(x, y)b(x)b(y)$. The proof that X_n is asymptotically normal is completed by showing that as $n \to \infty$, $|X_n - Y_n| \to_p 0$. Using Fubini's theorem,

$$\begin{aligned} E|X_n - Y_n| &\leqq E\int_0^1 |(n^{-\frac{1}{2}}\xi_n(x) - b(x))(\xi_n(x) - n^{\frac{1}{2}}b(x))|\,dx \\ &= n^{-\frac{1}{2}}\int_0^1 B(x, x)\,dx. \end{aligned} \tag{2.5}$$

Since $B(x, y)$ has finite trace, Markov's inequality, applied to (2.5), proves that $|X_n - Y_n| \to_p 0$ as $n \to \infty$. The uniformity of the convergence in distribution is assured by Pólya's theorem (c.f. Pólya [7]). □

THEOREM 2. *If, under the distribution function* $G(x)$, $b(x) \equiv 0$, *then the asymptotic characteristic function of* T_n *is*

$$\Psi(t) = \prod_{k=1}^{\infty} (1 - 2\lambda_k it)^{-\frac{1}{2}}.$$

The corresponding asymptotic distribution function cannot be degenerate. The convergence to the asymptotic distribution function as $n \to \infty$ *is uniform on the real line.*

PROOF. The characteristic function is derived by the method used to prove Theorem 3 in [3]. Since the Fourier coefficients of $b(x)$ are $\{c_m \bar{d}_m\}$ if $m \neq 0$ and 0 if $m = 0$, the condition $b(x) \equiv 0$ implies

$$\sum_k \lambda_k \equiv \operatorname{tr} B(x, y) = \sum_m (1 - |d_m|^2)|c_m|^2 = \sum_{m\neq 0} |c_m|^2 \neq 0, \tag{2.6}$$

so that at least one of the $\{\lambda_k\}$ is non-zero. Therefore, the asymptotic distribution function of T_n is non-degenerate. By Lemma 1.1, it is also continuous, so that the last assertion of the theorem follows from Pólya's theorem. □

LEMMA 1.1. *If the* $\{\lambda_k\}$ *are not all zero, the distribution function* $H(x)$ *whose characteristic function is*

$$\Psi(t) = \prod_{k=1}^{\infty} (1 - 2\lambda_k it)^{-\frac{1}{2}}$$

is continuous and has on $(0, \infty)$ *a continuous density which does not vanish.*

PROOF. Assume an infinite number of the $\{\lambda_k\}$ do not vanish (otherwise the lemma is obvious) and denote the non-zero λ_k by $\nu_1 \geqq \nu_2 \geqq \cdots > 0$. Let $H_N(x)$ denote the distribution function corresponding to $\Psi_N(t) = \prod_{k=1}^N (1 - 2\nu_k it)^{-\frac{1}{2}}$. If $\nu > 0$ denotes min $(\nu_1, \nu_2, \nu_3, \nu_4)$, $0 < |\Psi_N(t)| < (1 + 4\nu^2 t^2)^{-1}$ for every N and $0 \leqq |\Psi(t)| < (1 + 4\nu^2t^2)^{-1}$. Since $\Psi(t)$ and all the $\{\Psi_N(t)\}$ belong to $L_1(-\infty, \infty)$, $H(x)$ and all the $\{H_N(x)\}$ are continuous and have continuous densities (c.f. Gnedenko [4]). Therefore, by Pólya's theorem, $H_N(x)$ converges to $H(x)$ uniformly on the real line. Since $H_N(x)$ is strictly monotone on $(0, \infty)$ for every N, so is $H(x)$. □

Another limiting distribution of interest is that of T_n under a sequence $\{G_n(x)\}$ of local alternatives to uniformity defined by

$$G_n(x) = (1 - \alpha/n^{\frac{1}{2}})x + \alpha/n^{\frac{1}{2}}G(x); \quad \alpha \,\varepsilon\, [0, 1], \quad x \,\varepsilon\, [0, 1] \tag{2.7}$$

For $y, y_1, y_2 \,\varepsilon\, [0, 1]$, let $\mu^*(y)$ and $B_n^*(y_1, y_2)$ denote, respectively, the mean and covariance kernel of $\xi_n(y)$ under $G_n(y)$; this time $\mu^*(y)$ is independent of n, while $B_n^*(y_1, y_2)$ is not. If the $\{d_m^{(n)}\}$ are the Fourier transforms of $G_n(y)$,

$$d_0^{(n)} = 1, \qquad d_m^{(n)} = (\alpha/n^{\frac{1}{2}})\, d_m \quad \text{if} \quad m \neq 0. \tag{2.8}$$

It follows, as previously, that the Fourier coefficients of $\mu^*(y)$ are $\{\alpha c_m \bar{d}_m\}$ if $m \neq 0$ and 0 if $m = 0$. The double Fourier coefficients of $B_n^*(y_1, y_2)$ are $\{(d_{l-m}^{(n)} - d_l^{(n)}\bar{d}_m^{(n)})\bar{c}_l c_m\}$.

THEOREM 3. *Under the sequence of alternatives with distribution functions* $\{G_n(x)\}$, *the asymptotic characteristic function of* T_n *is*

$$\Psi^*(t) = \prod_{k=1}^\infty [1 - 2\,|c_k|^2\, it]^{-1} \exp\,[2\alpha^2\, |c_k|^2\, |d_k|^2\, it/(1 - 2\,|c_k|^2\, it)].$$

The convergence to the asymptotic distribution function as $n \to \infty$ *is uniform on the real line.*

PROOF. Let

$$Z_{kn} = \int_0^1 \xi_n(x) e^{-2\pi ikx}\, dx. \tag{2.8'}$$

Under $G_n(x)$, $EZ_{kn} = \alpha c_k \bar{d}_k$, $E\,|Z_{kn} - EZ_{kn}|^2 = (1 - n^{-1}\alpha^2|d_k|^2)|c_k|^2$, and $E(Z_{kn} - EZ_{kn})\overline{(Z_{ln} - EZ_{ln})} = O(n^{-\frac{1}{2}})$ if $k \neq l$. Let T be the random variable with characteristic function $\Psi^*(t)$; T is a convolution of non-central chi-squared variables. Let S_N be the random variable whose characteristic function $\Psi_N^*(t)$ is the product of the first N factors in $\Psi^*(t)$, and let $S_{nN} = \sum_{|k|=1}^N |Z_{kn}|^2$. By monotone convergence,

$$\begin{aligned} E\,|T_n - S_{nN}| &= E \sum_{|k|=N+1}^\infty |Z_{kn}|^2 \\ &= \sum_{|k|=N+1}^\infty [(1 - n^{-1}\alpha^2|d_k|^2)|c_k|^2 + \alpha^2\, |c_k|^2\, |d_k|^2]. \end{aligned} \tag{2.9}$$

The sum on the right of (2.9) is bounded for all $n > 0$ and $\alpha \,\varepsilon\, [0, 1]$ by $2\sum_{|k|=N+1}^\infty |c_k|^2$, which tends to zero as $N \to \infty$. By Markov's inequality, $T_n - S_{nN} \to_p 0$ uniformly in n as $N \to \infty$. Clearly $S_N \to_L T$ as $N \to \infty$.

To complete the proof of the theorem, it is sufficient to show that for any $N > 0$, $S_{nN} \to_L S_N$ as $n \to \infty$. From the definition (2.8′) of Z_{kn}, S_{nN} can be expressed as a sum of squared sine and cosine transforms of $\xi_n(x)$. Each such transform can, in view of (2.1), be expressed as a sum of iid random variables. A routine characteriste function argument establishes the joint asymptotic normality and independence of the sine and cosine transforms, and the result $S_{nN} \rightarrow_L S_N$ as $n \rightarrow \infty$ follows. The uniformity of the convergence in distribution functions is proved, as for Theorem 2, by using a variant of Lemma 1.1. ▯

COROLLARY 3.1. *Under the null hypothesis of uniformity, the asymptotic characteristic function of* T_n *is*

$$\Psi^*(t) = \prod_{k=1}^{\infty} [1 - 2|c_k|^2 it]^{-1}.$$

The convergence to the asymptotic distribution function as $n \rightarrow \infty$ *is uniform on the real line.*

When the non-vanishing $\{|c_m|^2; m > 0\}$ are all distinct, this last characteristic function may be inverted readily and it shows that, under uniformity,

$$\lim_{n\rightarrow\infty} P(T_n > x) = \sum_{m=1}^{\infty} a_m \exp[-x/2|c_m|^2], \tag{2.10}$$

where $a_m = \prod_{k\neq m} [1 - |c_k|^2|c_m|^{-2}]^{-1}$. The coefficients $\{a_m\}$ can be evaluated as a finite product of gamma functions (c.f. Whittaker and Watson [13], p. 238).

3. Consistency. The theorems of Section 2 enable us to determine the alternatives $G(x)$ on [0, 1] against which T_n, with given $f(x)$, yields a consistent test for uniformity. Let z_n be the exact α-level critical value for the test and let z be the α-level critical value relative to the asymptotic null distribution function $F(x)$ determined in Corollary 3.1. Using the continuity and monotonicity of $F(x)$, it is easily shown that $z_n \rightarrow z > 0$ as $n \rightarrow \infty$.

THEOREM 4. T_n *gives a consistent test for uniformity against an alternative with distribution function* $G(x)$ *if and only if* $b(x) \not\equiv 0$.

PROOF. Using (1.9), the continuity of $d(x)$, and the Helly-Bray theorem,

$$T_n/n = \int_0^1 \int_0^1 d(x-y)\, dF_n(x)\, dF_n(y) \rightarrow_{\text{a.s.}} \int_0^1 \int_0^1 d(x-y)\, dG(x)\, dG(y). \tag{3.1}$$

Moreover,

$$\int_0^1 \int_0^1 d(x-y)\, dG(x)\, dG(y) = \sum_{m\neq 0} |c_m|^2|d_m|^2 = \int_0^1 b^2(x)\, dx. \tag{3.2}$$

Therefore, if $b(x) \not\equiv 0$, the test generated by T_n is consistent.

Conversely, if $b(x) \equiv 0$, T_n has, by Theorem 2 and Lemma 1.1, an asymptotic non-null distribution function $H(x)$ which is continuous, strictly monotone on $(0, \infty)$, and to which the exact non-null distribution function $H^{(n)}(x)$ converges uniformly. Since $z_n \rightarrow z > 0$ as $n \rightarrow \infty$, $H^{(n)}(z_n) \rightarrow H(z) > 0$, and the test is not consistent. ▯

COROLLARY 4.1. T_n *gives a consistent test for uniformity against an alternative with distribution function* $G(x)$ *if and only if there exists at least one* $m \neq 0$ *such that both* c_m *and* d_m *do not vanish.*

As examples, consider the statistics U_n^2 and A_n discussed in Section 1. For U_n^2, c_m never vanishes, so that U_n^2 gives a consistent test for uniformity against all alternatives. On the other hand, for A_n, c_m is non-zero if and only if m is odd or zero. It is easily shown that if $G(x)$ possesses the symmetry $G(x + \frac{1}{2}) = G(x) + \frac{1}{2}$, where the addition $x + \frac{1}{2}$ is performed modulo 1, d_m vanishes whenever m is odd. Hence, A_n is not consistent against such alternatives.

4. Approximation of eigenfunctions and eigenvalues. Under many alternatives, the eigenfunctions and eigenvalues of $B(y_1, y_2)$ cannot be found analytically. This section describes a method for approximating them to any degree of accuracy. While the basic idea is by no means novel, it has seen little application in statistics outside the area of spectral analysis.

Let $\varphi(x) \varepsilon L_2[0, 1]$ be an eigenfunction of $B(y_1, y_2)$ and let the $\{a_m\}$ be its Fourier coefficients relative to the orthonormal basis $\{e^{2\pi imx}; m = 0, \pm 1, \pm 2, \cdots\}$. Substituting the Fourier coefficients of $B(y_1, y_2)$ computed in Section 2 into the integral equation (2.3) and using the completeness of the orthonormal basis, we find

THEOREM 5. *A function $\varphi(x)$ with Fourier coefficients $\{a_m\}$ is an eigenfunction of $B(y_1, y_2)$ and λ is the corresponding eigenvalue if and only if*

$$\sum_l a_l[d_{l-m} - d_l \overline{d_m}]c_m\overline{c_l} = \lambda a_m ; \quad m = 0, \pm 1, \pm 2, \cdots .$$

This reformulation of the problem suggests the following approximation technique. Consider the truncated kernel

$$B_N(y_1, y_2) = \sum_{|l| \leqq N} \sum_{|m| \leqq N} [d_{l-m} - d_l \overline{d_m}]c_m\overline{c_l}e^{2\pi i(my_1 - ly_2)}. \tag{4.1}$$

Finding the eigenvalues and eigenfunctions of $B_N(y_1, y_2)$ is equivalent, in view of Theorem 5, to diagonalizing a finite-dimensional hermitian matrix $H_N = \{h_{ml}; |m|, |l| = 1, 2, \cdots N\}$, where $h_{ml} = \{d_{l-m} - d_l \overline{d_m}\}c_m\overline{c_l}$. Indeed, the use of real Fourier series reduces this problem to that of diagonalizing a real symmetric matrix. Let $\lambda_{1N} \geqq \lambda_{2N} \geqq \cdots \geqq \lambda_{2N,N}$ be the ordered eigenvalues of $B_N(y_1, y_2)$ and let $\varphi_{1N}(x), \varphi_{2N}(x), \cdots, \varphi_{2N,N}(x)$ be the corresponding eigenfunctions. The following theorem justifies the use of λ_{kN} and $\varphi_{kN}(x)$ as approximations to λ_k and $\varphi_k(x)$ respectively.

THEOREM 6. *As $n \to \infty$, $\lambda_{kN} \to \lambda_k$ and $\|\varphi_{kN} - \varphi_k\| \to 0$.*

PROOF. $B(y_1, y_2)$ defines a linear operator B mapping $L_2[0, 1]$ into $L_2[0, 1]$ through the relation

$$(Bh)(y_1) = \int_0^1 B(y_1, y_2)h(y_2)\, dy_2 ; \qquad h \varepsilon L_2[0, 1]. \tag{4.2}$$

B is self-adjoint, positive semi-definite and completely continuous with finite trace. Similarly, $B_N(y_1, y_2)$ defines a linear operator B_N mapping $L_2[0, 1]$ into a finite-dimensional subspace of $L_2[0, 1]$. B_N is self-adjoint, completely continuous with finite trace, and has a finite number of non-zero eigenvalues. A simple computation on the Fourier coefficients of $B(y_1, y_2)$ and $B_N(y_1, y_2)$ shows that $\|B - B_N\| \to 0$ as $N \to \infty$.

The eigenvalues of B have a variational characterization (c.f. Gould [5]): The rth eigenvalue of B is the minimum value which can be given by the adjunction of $r - 1$ linear constraints to the maximum of $(Bh, h)/(h, h)$. A parallel result holds for B_N. Suppose $h \varepsilon L_2[0, 1]$ is constrained to be orthogonal to the first $k - 1$ eigenfunctions of B, i.e., there are $k - 1$ linear constraints on h. Then, taking $\|h\| = 1$ without loss of generality,

$$(4.3) \quad \lambda_{kN} \leqq \max_h (B_N h, h) \leqq \max_h (Bh, h) + \max_h ((B_N - B)h, h) \leqq \lambda_k + \|B_N - B\|.$$

Reversing the roles of B and B_N, λ_k and λ_{kN} in the argument shows that $\lambda_k \leqq \lambda_{kN} + \|B - B_N\|$; therefore $\lambda_{kN} \to \lambda_k$ as $N \to \infty$.

Let $\{g_{mN}\}$ denote the Fourier coefficients of the normalized eigenfunction $\varphi_{kN}(x)$ relative to the basis $\{\varphi_m(x)\}$; $\sum_{m=1}^{\infty} g_{mN}^2 = 1$. Since

$$(4.4) \qquad \|B\varphi_{kN} - \lambda_k \varphi_{kN}\| \leqq \|B - B_N\| + |\lambda_{kN} - \lambda_k|,$$

which tends to zero as $N \to \infty$,

$$(4.5) \qquad 0 = \lim_{N\to\infty} \|\textstyle\sum_m \lambda_m g_{mN} \varphi_m - \lambda_k \sum_m g_{mN} \varphi_m\| = \lim_{N\to\infty} \sum_{\lambda_m \neq \lambda_k} (\lambda_m - \lambda_k)^2 g_{mN}^2 .$$

Suppose the eigenvalue λ_k has multiplicity one and let $c = \min_{m \neq k} (\lambda_m - \lambda_k)^2 > 0$. From (4.5),

$$(4.6) \qquad \textstyle\sum_{m \neq k} g_{mN}^2 \leqq c^{-1} \sum_{m \neq k} (\lambda_m - \lambda_k)^2 g_{mN}^2 \to 0$$

as $N \to \infty$, so that $\lim_{N\to\infty} g_{kN}^2 = 1$. Without loss of generality, we may assume the sign of $\varphi_k(x)$ is such that $\lim_{N\to\infty} g_{kN} = 1$, whereupon $\|\varphi_{kN} - \varphi_k\| \to 0$ as $N \to \infty$. If the multiplicity of λ_k is greater than one, only a minor change in the argument is required. □

5. Moments and approximate power. This section examines two simple approximations to the power of T_N; both depend only on the mean M_n and variance V_n of T_n. The first, motivated by Theorem 1, is a normal approximation. The other, suggested by Theorem 3, is a chi-square approximation compounded with the Wilson-Hilferty approximation to the chi-square (as in Grad and Solomon [6]):

$$(5.1) \quad P(T_n > x) = 1 - \Phi[((x/M_n)^{1/3} - (1 - \tfrac{2}{9} V_n M_n^{-2}))(\tfrac{2}{9} V_n M_n^{-2})^{-\frac{1}{2}}];$$

$\Phi(\cdot)$ is the $N(0, 1)$ distribution function.

The moments of T_n required for these power approximations can be found exactly from its representation (1.7) as a quadratic form. Because of absolute convergence, term by term multiplication of the series with itself yields legitimate series representations of the powers of T_n. By calculating the expectation of the appropriate series (term by term in view of monotone convergence), any moment of T_n about the origin can be found. Hence

THEOREM 7. *If the observations are independent realizations of a random variable whose distribution function $G(x)$ has the Fourier transforms $\{d_m\}$, the first* *two moments of T_n are*

$$ET_n = \sum_{m\neq 0}|c_m|^2[1 + (n-1)|d_m|^2],$$

$$ET_n^2 = n^{-2}\sum_{m\neq 0}\sum_{l\neq 0}|c_m|^2|c_l|^2[n^2 + n^2(n-1)(|d_m|^2 + |d_l|^2) + n(n-1)(n-2)(n-3)|d_m|^2|d_l|^2 + 2n(n-1)(n-2) \cdot (R_e\{d_{m+l}\, d_{-m}\, d_{-l}\} + R_e\{d_{m-l}\, d_{-m}\, d_l\}) + n(n-1)(|d_{m+l}|^2 + |d_{m-l}|^2)].$$

The two approximations to the power of T_n were checked for a special case, Ajne's statistic A_n (1.5), by comparison with Monte Carlo simulations of A_n under several alternatives with densities belonging to the parametric family

$$g(x \mid p) = \begin{cases} 2p & \text{if } x \,\varepsilon\, [0, \tfrac{1}{2}) \\ 2q & \text{if } x \,\varepsilon\, [\tfrac{1}{2}, 1) \end{cases}; \quad p \,\varepsilon\, [0, 1], p + q = 1. \tag{5.2}$$

The Fourier coefficients $\{d_m\}$ of $g(x \mid p)$ are $\{2(q - p)(\pi i m)^{-1}\}$ if m is odd and zero otherwise. The Fourier coefficients $\{c_m\}$ of $N(x) - \frac{1}{2}n$ are $\{(\pi i m^{-1})\}$ if m is odd and zero otherwise (c.f. Watson [11]). By Theorem 7, therefore, the mean and variance of A_n under $g(x \mid p)$ are, respectively,

$$M_n = \tfrac{1}{4} + (n-1)(p-q)^2/12 \tag{5.3}$$

$$V_n = ((n-1)n^{-1})[\tfrac{1}{24} + \tfrac{1}{30}(n-2)(p-q)^2 - \tfrac{1}{72}(2n-3)(p-q)^4].$$

A_n was calculated from each simulated sample by means of the computational formula

$$A_n = \tfrac{1}{4}n - 2\sum_{i<j}\sum d_{ij} n^{-1}, \tag{5.4}$$

where d_{ij} is the shortest distance on the unit circle between observations x_i and x_j. Table 1, which reports the results, suggests that for alternatives which are not too distant, power approximation (5.1) is better than the simple normal approximation.

TABLE 1

Simulated Moments and Power of A_n versus Theoretical Moments and Fitted Approximate Power of A_n

	p	.5	.6	.7	.8	.9	1.0
Mean	Simulated	.241	.319	.502	.832	1.252	1.843
	Theoretical	.250	.313	.503	.820	1.263	1.833
Variance	Simulated	.0371	.0625	.117	.185	.216	.125
	Theoretical	.0396	.0616	.118	.182	.204	.121
Power	Simulated	.042	.098	.279	.608	.904	1.000
	Approx. (5.1)	.046	.093	.263	.594	.936	1.000
	Normal Approx.	.021	.080	.327	.649	.910	.999

Simulation Parameters: 2500 samples, sample size $n = 20$, asymptotic test size $\alpha = .05$.

6. Approximate Bahadur efficiency. In this section, the approximate Bahadur efficiency of T_n (c.f. Bahadur [2] for a definition) under an arbitrary alternative with distribution function $G(x)$ is found from the asymptotic null distribution of T_n. An example shows, however, that here, as in other instances noted in the literature, the approximate efficiency can be deceptive as a criterion of power.

THEOREM 8. *The approximate slope of T_n under an alternative with distribution function $G(x)$ is*

$$s^*(G) = \sum_{m\neq 0} |c_m|^2|d_m|^2/\max_m |c_m|^2.$$

PROOF. From (3.1) and (3.2), as $n \to \infty$,

$$T_n n^{-1} \to_{a.s.} \sum_{m\neq 0} |c_m|^2|d_m|^2. \tag{6.1}$$

To complete the proof, it is sufficient to show that, if $F(x)$ is the asymptotic null distribution function

$$n^{-1} \log [1 - F(nt)] \to -t/(2 \max_m|c_m|^2) \tag{6.2}$$

as $n \to \infty$. Suppose that r_1 of the $\{|c_m|^2, m > 0\}$ equal ν_1, r_2 of them equal ν_2, and so forth, with $\nu_1 > \nu_2 > \cdots \geqq 0$. By Corollary (3.1), the asymptotic characteristic function of T_n under the null hypothesis is

$$\Psi^*(t) = \prod_{k=1}^{\infty} [1 - 2\nu_k it]^{-r_k}. \tag{6.3}$$

Therefore, (c.f. Zolotarev [14]),

$$\lim_{x\to\infty} (1 - F(x))/P[\chi^2(2r_1) > x/\nu_1] = \prod_{k=2}^{\infty} [1 - \nu_k/\nu_1]^{-r_k}. \tag{6.4}$$

By a well-known identity,

$$P[\chi^2(2r_1) > x/\nu_1] = \sum_{j=0}^{r_1-1} (x/2\nu_1)^j/j! \exp [-x/2\nu_1]. \tag{6.5}$$

For x sufficiently large, the exponential dominates in (6.5) and (6.2) follows. □

We give an example where $s^*(G)$ is not a reliable measure of power. Let K_n be the special case of T_n obtained when the Fourier coefficients $\{c_m\}$ of $f(x)$ are 1 if $|m| = 1, 2, \cdots 10$ and 0 if $|m| > 10$. The asymptotic null distribution of K_n is chi-square with 20 degrees of freedom, by Corollary 3.1. We compare K_n with Ajne's statistic A_n against the family of alternatives defined in (5.2). The approximate slope of A_n is $\frac{1}{12}\pi^2(p - q)^2 = .82(p - q)^2$, while the approximate slope of K_n is $2 \sum_{k=1}^{5} 4(p - q)^2\pi^{-2}(2k - 1)^{-2} = .96(p - q)^2$. Therefore, K_n

TABLE 2

Simulated Power of K_n versus Simulated Power of A_n

p	.5	.6	.7	.8	.9	1.0
Power of K_n	.050	.066	.134	.262	.468	.802
Power of A_n	.042	.098	.279	.608	.904	1.000

ought to be more efficient than A_n for every $p \neq \frac{1}{2}$, including p near $\frac{1}{2}$; but A_n is lmp invariant.

From 500 simulated samples of size 20, the power of K_n against alternatives of the form (5.2) was estimated; the asymptotic size of the test was set at .05. Table 2 compares the results with the power of A_n, found by simulation in Section 5. The power of K_n is notably less than that of A_n through the whole range of p.

7. Acknowledgment. Professor G. S. Watson suggested several of the topics discussed in this paper; for this and for his guidance in the research, I am grateful.

REFERENCES

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

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

[3] BERAN, R. J. (1968). Testing for uniformity on a compact homogeneous space. *J. Appl. Prob.* **5** 177–195.

[4] GNEDENKO, B. V. (1963). *The Theory of Probability.* Chelsea, New York.

[5] GOULD, S. H. (1966). *Variational Methods for Eigenvalue Problems.* Univ. of Toronto Press.

[6] GRAD, A. and SOLOMON, H. (1955). Distribution of quadratic forms and some applications. *Ann. Math. Statist.* **26** 464–477.

[7] PÓLYA, G. (1920). Uber den Grenzwertsatz der Wahrscheinlichtskeitrechnung und das Momentproblem. *Math. Zeitschrift* **8** 171–181.

[8] RAYLEIGH, LORD (1919). On the problems of random vibrations and flights in one, two, and three dimensions. *Phil. Mag.* (6), **37** 321–347.

[9] SCHACH, S. (1967). Non-parametric tests of location for circular distributions. Tech. Rep. No. 95, Dept. of Stat., Univ. of Minnesota.

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

[11] WATSON, G. S. (1967). Another test for uniformity of a circular distribution. *Biometrika* **54** 675–677.

[12] WATSON, G. S. (1967). Some problems in the statistics of directions. To appear in *Bull. Int. Statist. Inst.*

[13] WHITTAKER, E. T. and WATSON, G. N. (1967). *A Course of Modern Analysis.* Cambridge.

[14] ZOLOTAREV, V. M. (1961). Concerning a certain probability problem. *Theory Prob. Appl.* **6**, 201–204.