The Annals of Statistics
1975, Vol. 3, No. 2, 401-412

LOCAL ASYMPTOTIC POWER OF QUADRATIC RANK TESTS FOR TREND[1]

BY RUDOLF BERAN

University of Toronto and University of California, Berkeley

A general class of quadratic rank tests for randomness versus trend is introduced and studied in this paper. Included within this class are the Cramér–von Mises two sample test, the Watson two-sample test, and their extensions to trend alternatives. Analytical study is made of the asymptotic power and efficiency of such tests in a neighborhood of the null hypothesis.

1. Introduction. Consider the problem of testing a set of observations $(X_1, X_2, \cdots, X_N)$ for randomness versus trend in location. Among the many tests available for this purpose are those based upon quadratic rank statistics. One such is the Kruskal–Wallis k-sample test and its extension (Hájek and Šidák (1967)) to other score functions; the test statistic in this case is

$$Q_N = [(N-1)/\sum_{j=1}^N a_N^2(j)] \sum_{i=1}^k n_i^{-1}[\sum_{j\in s_i} a_N(R_j)]^2 , \tag{1.1}$$

where $(R_1, R_2, \cdots, R_N)$ are the ranks of the observations and $\{s_1, s_2, \cdots, s_k\}$ is a partition of the set $\{1, 2, \cdots, N\}$ such that s_i contains n_i elements. Another example is the Cramér–von Mises two-sample test and its generalization (Hájek and Šidák (1967)) to trend alternatives; the test statistic in the latter case is

$$T_N = [\sum_{j=1}^N (c_j - \bar{c})^2]^{-1} \int_0^1 [\sum_{j=1}^N (c_j - \bar{c}) a_N(R_j, t)]^2 \, dt , \tag{1.2}$$

where the $\{c_i\}$ are constants and the function $a_N(i, t)$ is 0 if $i \leqq tN$, is $i - tN$ if $tN < i < tN + 1$, and is 1 if $i \geqq tN + 1$. A third example is Watson's (1962) two-sample test based upon the statistic

$$U_{m,n}^2 = (mn/N) \int_0^1 [F_m(x) - F_n(x) - \int_0^1 \{F_m(y) - F_n(y)\} \, dH_N(y)]^2 \, dH_N(x) , \tag{1.3}$$

where F_m, F_n, and H_N are the empirical distribution functions of the first, second, and combined samples respectively. A class of two-sample quadratic rank tests containing $U_{m,n}^2$ has been studied by Schach (1969) and by Beran (1969).

Each of the foregoing tests can be embedded (up to asymptotic equivalence under null hypothesis and contiguous trend alternatives) within the following general class of quadratic rank tests for trend. Let the $N-1$ row vectors $\{(c_r(1), c_r(2), \cdots, c_r(N));\ 1 \leqq r \leqq N-1\}$ be orthonormal contrasts and let

$$S_N = \sum_{r=1}^{N-1} \sum_{s=1}^{N-1} \alpha_{r,s}^2 [\sum_{j=1}^N c_r(j) a_s(R_j)]^2 , \tag{1.4}$$

Received September 1973; revised March 1974.

[1] This research was partially supported by National Research Council of Canada Grant A8724 and by National Science Foundation Grant GP-31091X.

AMS 1970 *subject classifications.* Primary 62.15; Secondary 62.70.

Key words and phrases. Local asymptotic power, quadratic rank tests, tests for trend.

where the score functions $\{a_s(\cdot); s \geqq 1\}$ are associated with a complete orthonormal basis for the nonconstant elements of $L_2[0, 1]$ and $\sum_{r,s=1}^{\infty} \alpha_{r,s}^2 < \infty$; detailed assumptions will be presented in Section 2 of this paper. Reject the hypothesis of randomness if the observed value of S_N exceeds a critical value. The k-sample statistic Q_N is asymptotically equivalent to a special case of S_N in which $\alpha_{r,1} = 1$ for $1 \leqq r \leqq k-1$ and $\alpha_{r,s} = 0$ otherwise. In Beran (1969), it was noted that $U_{m,n}^2$ is asymptotically equivalent to $(N/mn) \sum_{|s|=1}^{\infty} (2\pi s)^{-2} |\sum_{j=1}^{m} \times \exp(2\pi i s R_j/N)|^2$; truncation of the infinite sum after $N-1$ terms yields an asymptotically equivalent statistic of the form (1.4). The statistic T_N can similarly be associated with

$$[\textstyle\sum_{j=1}^{N} (c_j - \bar{c})^2]^{-1} \sum_{s=1}^{N-1} 2(\pi s)^{-2} [\sum_{j=1}^{N} (c_j - \bar{c}) \cos(\pi s R_j/N)]^2 .$$

The class of tests based upon S_N contains some interesting new rank tests for trend. For example, setting $\alpha_{r,s} = (rs)^{-1}$ for all $r, s \geqq 1$ gives a test which is asymptotically strictly unbiased against every regular trend alternative and has a tabulated asymptotic null distribution. One convenient choice for the orthonormal contrast vectors consists of the $N-1$ normalized nonconstant basis vectors of the finite Fourier transform of dimension N. In this case, the weights $\alpha_{r,s}^2$ can be assigned to the contrast frequencies according to their anticipated order of importance.

The asymptotic distribution of S_N under trend alternatives contiguous to the hypothesis of randomness is the same as the distribution of a linear combination of independent noncentral chi-square distributed random variables. Apart from the simplest special cases, the complexity of this distribution has deterred analytical power comparisons among quadratic rank tests. Section 2 of this paper provides assumptions under which the limiting distribution of S_N is as asserted above and presents some new results concerning the asymptotic power of S_N in a neighborhood of the null hypothesis. It is found that the squared component(s) of S_N associated with the largest of the $\{\alpha_{r,s}^2\}$ largely determines, at small significance levels, the local asymptotic power of the S_N-test. Moreover, this effect can still be pronounced at customary significance levels. Proofs of the theorems are sketched in Section 3.

2. Results. Let $X_1, X_2, \cdots$ be a sequence of random variables, let $\{N_\nu; \nu \geqq 1\}$ denote a sequence of sample sizes, and let $R_{\nu,i}$ be the rank of X_i among $(X_1, X_2, \cdots, X_{N_\nu})$. Under the hypothesis of randomness H_ν, the observations $(X_1, X_2, \cdots, X_{N_\nu})$ will be assumed independent, identically distributed with common continuous distribution function F. Let l_2 denote the space of all double sequences $\{x_{r,s}; r, s \geqq 1\}$ for which $\sum_{r,s} x_{r,s}^2 < \infty$. In studying the asymptotic behavior of the test statistics

$$S_N = \textstyle\sum_{r=1}^{N_\nu - 1} \sum_{s=1}^{N_\nu - 1} \alpha_{r,s}^2 [\sum_{j=1}^{N_\nu} c_{r,\nu}(j) a_{s,\nu}(R_{\nu,j})]^2 , \tag{2.1}$$

the following assumptions will be made:

A. The double sequence of constants $\{\alpha_{r,s}; r, s \geqq 1\}$ belongs to l_2; $\max_{r,s}\alpha_{r,s}^2 = 1$.

B. The score functions $\{a_{s,\nu}(\cdot); s \geqq 1\}$ are associated with a complete orthonormal basis $\{\varphi_s(\cdot); s \geqq 1\}$ for the orthocomplement in $L_2[0, 1]$ of the constant function. Either

$$\text{(i)} \qquad a_{s,\nu}(j) = E_{H_\nu}[\varphi_s \cdot F(X_1) \mid R_{\nu,1} = j] , \qquad 1 \leqq j \leqq N_\nu$$

or

$$\text{(ii)} \qquad \begin{aligned} &\lim_{\nu\to\infty} \textstyle\int_0^1 [a_{s,\nu}(1 + [tN_\nu]) - \varphi_s(t)]^2\, dt = 0 \qquad \text{for} \quad s \geqq 1 , \\ &\lim_{\nu\to\infty} \sup_s |\textstyle\int_0^1 [a_{s,\nu}(1 + [tN_\nu])]^2\, dt - 1| = 0 , \\ &\lim_{\nu\to\infty} \sup_s |\textstyle\int_0^1 a_{s,\nu}(1 + [tN_\nu])\, dt| = 0 . \end{aligned}$$

C. The vectors $\{(c_{r,\nu}(1), c_{r,\nu}(2), \cdots, c_{r,\nu}(N_\nu)); 1 \leqq r \leqq N_\nu - 1\}$ form an orthonormal basis for the space of all N_ν-dimensional contrasts. Moreover,

$$\lim_{\nu\to\infty} \max_{1\leqq j\leqq N_\nu} c_{r,\nu}^2(j) = 0 \qquad \text{for} \quad 1 \leqq r \leqq N_\nu - 1 .$$

The score functions associated with the Cramér–von Mises and Watson tests satisfy B(ii). If all but a finite number of the $\{\alpha_{r,s}\}$ vanish, as is the case for the generalized Kruskal–Wallis test, the last two conditions in B(ii) become unnecessary. The normalized finite Fourier transform contrasts mentioned in the Introduction satisfy assumption C.

Let Z_ν be a stochastic double sequence whose (r, s)th element is $\alpha_{r,s} \sum_{j=1}^{N_\nu} \times c_{r,\nu}(j)a_{s,\nu}(R_{\nu,j})$ for $1 \leqq r, s \leqq N_\nu - 1$, and is zero ortherwise. Let $Z = \{\alpha_{r,s}Y_{r,s}; r, s \geqq 1\}$ where the $\{Y_{r,s}\}$ are independent $N(0, 1)$ random variables. Under assumption A, the stochastic double sequence Z belongs to l_2 with probability one.

THEOREM 1. *Suppose assumptions* A, B, C *are satisfied. Then, under* H_ν *as* $\nu \to \infty$, Z_ν *converges in distribution to* Z *in the* l_2*-topology.*

Suppose the distribution function F has density f and let $K_\nu(\theta)$ denote the following trend alternative to H_ν: under $K_\nu(\theta)$ the observations $(X_1, X_2, \cdots, X_{N_\nu})$ have joint density $\prod_{j=1}^{N_\nu} f(x_j - \theta d_\nu(j))$, where $\theta \neq 0$. Regarding f and the $\{d_\nu(j)\}$, the following assumptions will be made:

D. The density f is absolutely continuous and $\psi_F(u) = -f' \cdot F^{-1}(u)/f \cdot F^{-1}(u)$ belongs to $L_2[0, 1]$. The constants $\{d_\nu(j)\}$ are such that

$$\lim_{\nu\to\infty} \max_{1\leqq j\leqq N_\nu} (d_\nu(j) - \bar{d}_\nu)^2 = 0 , \qquad \lim_{\nu\to\infty} \sum_{j=1}^{N_\nu} (d_\nu(j) - \bar{d}_\nu)^2 < \infty , \qquad \text{and}$$
$$\lim_{\nu\to\infty} \sum_{j=1}^{N_\nu} c_{r,\nu}(j)d_\nu(j) = \beta_r \qquad \text{for every} \quad r \geqq 1 .$$

Let μ denote the double sequence whose (r, s)th element is $\mu_{r,s} = \beta_r \int_0^1 \varphi_s(u) \times \psi_F(u)\, du$ and let $Z(\theta\mu)$ denote the stochastic double sequence $\{\alpha_{r,s}(Y_{r,s} + \theta\mu_{r,s}); r, s \geqq 1\}$.

THEOREM 2. *Suppose assumptions* A, B, C, D *are satisfied. Then, under* $K_\nu(\theta)$ *as* $\nu \to \infty$, Z_ν *converges in distribution to* $Z(\theta\mu)$ *in the* l_2*-topology.*

Since the statistic S_ν is simply the l_2-norm of Z_ν, Theorems 1 and 2 imply that the limiting distributions of S_ν under H_ν and under $K_\nu(\theta)$ are the same as the distributions of $\sum_{r,s} \alpha_{r,s}^2 Y_{r,s}^2$ and $\sum_{r,s} \alpha_{r,s}^2(Y_{r,s} + \theta\mu_{r,s})^2$ respectively. By grouping together terms given equal weight, these two random variables can be re-written in the forms $S = \sum_{r,s} \sigma_{r,s}^2 \chi_{r,s}^2(n_{r,s})$ and $S(\theta b) = \sum_{r,s} \sigma_{r,s}^2 \chi_{r,s}^2(n_{r,s}, \theta^2 b_{r,s}^2)$, where the $\{\sigma_{r,s}^2\}$ are all distinct, $\sigma_{11}^2 = 1$, $\sigma_{r,s}^2 < 1$ for $(r, s) \neq (1, 1)$, and $\{\chi_{r,s}^2(n_{r,s})\}$, $\{\chi_{r,s}^2(n_{r,s}, \theta^2 b_{r,s}^2)\}$ are, respectively, double sequences of independent chi-square and noncentral chi-square random variables with the indicated degrees of freedom and noncentrality parameters. The symbol b represents the double sequence $\{b_{r,s}^2, r, s \geqq 1\}$. Because of assumptions A and D, $\sum_{r,s} n_{r,s}\sigma_{r,s}^2 < \infty$ and $\sum_{r,s} b_{r,s}^2 \sigma_{r,s}^2 < \infty$.

In several well-known cases, including the first three cited in the Introduction, the distribution function of S has been derived. If $\alpha_{r,s} = (rs)^{-1}$, the distribution function of S is not known explicitly, but has been tabulated by Blum, Kiefer and Rosenblatt (1961) in their study of tests for independence. For $S(\theta b)$ the situation is less fortunate, apart from trivial special cases. While series expansions and other approximations to the distribution of $S(\theta b)$ exist, they do not appear suitable for analytical study of the asymptotic power of the S_ν-tests. However, Durbin and Knott (1972) and Stephens (1973) have recently used such approximations to examine the distribution of $S(\theta b)$ numerically in several interesting cases.

To obtain some analytical results concerning the asymptotic power of an S_ν-test, we will study $D(x \mid S, b) = dP[S(\theta b) > x]/d\theta^2|_{\theta=0}$. Since the tail probability $P[S(\theta b) > x]$ depends on θ only through θ^2, the derivative $D(x \mid S, b)$ determines a local approximation to the asymptotic power function of an S_ν-test in a neighborhood of the null hypothesis. Comparisons with other tests can be made on the basis of this local asymptotic power.

The characteristic function of $S(\theta b)$ is

$$\varphi(t \mid \theta b) = \left[\prod_{r,s} (1 - 2\sigma_{r,s}^2 it)^{-n_{r,s}/2}\right] \times \exp\left[\theta^2 \sum_{r,s} b_{r,s}^2 \sigma_{r,s}^2 it(1 - 2\sigma_{r,s}^2 it)^{-1}\right]. \tag{2.2}$$

Let $S_{r,s}^* = \sigma_{r,s}^2 \chi_{r,s}^2(n_{r,s} + 2) + \sum_{(l,m)\neq(r,s)} \sigma_{l,m}^2 \chi_{l,m}^2(n_{l,m})$ and let $g(\cdot \mid S_{r,s}^*)$ be the density of the random variable $S_{r,s}^*$. Examination of $d\varphi(t \mid \theta b)/d\theta^2|_{\theta=0}$ and the inversion formula for characteristic functions and some special cases establishes the following basic relations:

$$D(x \mid S, b) \doteq \sum_{r,s} b_{r,s}^2 D_{r,s}(x \mid S), \tag{2.3}$$

where

$$\begin{aligned} D_{r,s}(x \mid S) &= \frac{\partial P[S(b) > x]}{\partial b_{r,s}^2}\bigg|_{b=0} \\ &= 2^{-1}\{P[S_{r,s}^* > x] - P[S > x]\} \\ &= \sigma_{r,s}^2 g(x \mid S_{r,s}^*). \end{aligned} \tag{2.4}$$

The convergence of the right side of (2.3) is assured by the convergence of $\sum_{r,s} b_{r,s}^2 \sigma_{r,s}^2$.

In an unpublished technical report, Withers (1970) has studied the local asymptotic power of the weighted Cramér–von Mises goodness-of-fit test. Unlike (2.3) and (2.4), his result is formulated in terms of the resolvent kernel corresponding to the weighted empirical distribution function process. It is interesting that for the usual Cramér–von Mises test under double exponential shift alternatives, Withers was also able to invert the characteristic function $\varphi(t/\theta b)$ analytically.

The next theorem describes asymptotic (in x) expansions for $P[S > x]$ and for the derivative $D(x \mid S, b)$. These results are a natural development of earlier work by Zolotarev (1961) and by Hoeffding (1964) on the distribution of S. Let $\xi = S - \chi_{11}^2(n_{11})$ and for every integer $k \geqq 0$, let

$$(2.5)\qquad \begin{aligned} A_k &= (-1)^k \binom{n_{11}/2 - 1}{k} E[\xi^k \exp(\xi/2)] \\ B_k &= (-1)^k \binom{n_{11}/2}{k} E[\xi^k \exp(\xi/2)] \\ h_k &= 2^{k+1} \binom{n_{11}/2 - 1}{k} k! \,, \end{aligned}$$

where $\binom{m}{k} = m(m-1) \cdots (m-k+1)/k!$ if $k \geqq 1$ and 1 if $k = 0$. Note that for every integer k, $E[\xi^k \exp(\xi/2)]$ is finite and may be calculated from the kth derivative of the Laplace transform of the distribution of ξ. In particular

$$(2.6)\qquad \begin{aligned} E[\exp(\xi/2)] &= \prod_{(r,s) \neq (1,1)} (1 - \sigma_{r,s}^2)^{-n_{r,s}/2} \\ E[\xi \exp(\xi/2)] &= E[\exp(\xi/2)] \sum_{(r,s) \neq (1,1)} n_{r,s} \sigma_{r,s}^2 (1 - \sigma_{r,s}^2)^{-1} \,. \end{aligned}$$

Define constants $\{C_k;\ k \geqq 0\}$ through the recursion

$$(2.7)\qquad C_0 = A_0 \,, \qquad C_{k+1} = A_{k+1} + \sum_{i=0}^{k} i h_{k-i} C_i \,, \qquad k \geqq 0 \,.$$

For every $(r, s) \neq (1, 1)$, let $\xi_{r,s} = S_{r,s}^* - \chi_{11}^2(n_{11})$ and define $A_k(r, s)$ by replacing ξ with $\xi_{r,s}$ in the first line of (2.5). Note that $h_0 = 2$, $A_0 = B_0 = C_0 = E[\exp(\xi/2)]$, and $A_0(r, s) = A_0(1 - \sigma_{r,s}^2)^{-1}$. Let $p(x \mid m)$ denote the chi-square density with m degrees of freedom.

THEOREM 3. *The following expansions are valid for every integer* $m \geqq 1$:

$$(2.8)\qquad P[S > x] = \left[\sum_{k=0}^{m-1} C_k x^{-k} + O(x^{-m})\right] P[\chi_{11}^2(n_{11}) > x] \,,$$

and

$$(2.9)\qquad \begin{aligned} D(x \mid S, b) = {} & \left[b_{11}^2 \sum_{k=0}^{m-1} B_k x^{-k} + O(x^{-m})\right] p(x \mid n_{11} + 2) \\ & + \left[\sum_{(r,s) \neq (1,1)} b_{r,s}^2 \sigma_{r,s}^2 \sum_{k=0}^{m-1} A_k(r, s) x^{-k} \right. \\ & \left. + O(x^{-m})\right] p(x \mid n_{11}) \,. \end{aligned}$$

Expansion (2.8) provides a way to approximate c_α, the critical value for which $P[S > c_\alpha] = \alpha$. Under alternatives $K_\nu(\theta)$ for which $|\theta|$ is small, the asymptotic power of a level α S_ν-test is approximately $\alpha + \theta^2 D(c_\alpha \mid S, b)$. A calculation based on the first terms of (2.8), (2.9), and the well-known expansion

$$(2.10)\qquad P[\chi_{11}^2(n_{11}) > x] = \left[\sum_{k=0}^{m-1} h_k x^{-k} + O(x^{-m})\right] p(x \mid n_{11})$$

(see Whittaker and Watson (1927), page 159), shows that

$$\lim_{\alpha\to 0} 2(\alpha c_\alpha)^{-1} D(c_\alpha \mid S, b) = b_{11}^2 n_{11}^{-1} \tag{2.11}$$

and, if $b_{11}^2 = 0$,

$$\lim_{\alpha\to 0} 2\alpha^{-1} D(c_\alpha \mid S, b) = \sum_{(r,s)\neq(1,1)} b_{r,s}^2 \sigma_{r,s}^2 (1 - \sigma_{r,s}^2)^{-1} . \tag{2.12}$$

Thus, for small α, the local asymptotic power of a level α S_ν-test is largely determined by the value of b_{11}^2/n_{11}.

By grouping together equally weighted terms, the statistic S_ν can be written in the form $S_\nu = \sum_r \sum_s \sigma_{r,s}^2 T_{r,s,\nu}$. Another way of interpreting the information contained in expansion (2.9) is to compare the local asymptotic power of the S_ν-test with that of the simple tests based upon the individual components $\{T_{r,s,\nu};\, r, s \geqq 1\}$. From Theorems 1 and 2, the limiting distributions of $T_{r,s,\nu}$ under H_ν and under $K_\nu(\theta)$ are, respectively, the same as the distributions of $T_{r,s} = \chi_{r,s}^2(n_{r,s})$ and $T_{r,s}(\theta b) = \chi_{r,s}^2(n_{r,s}, \theta^2 b_{r,s})$. Let $d_\alpha(r, s)$ be the critical value for which $P[T_{r,s} > d_\alpha(r, s)] = \alpha$. The local asymptotic efficiency of S_ν relative to $T_{r,s,\nu}$ at level α under the alternative $K_\nu(\theta)$ is defined as

$$e[S, T_{r,s} \mid \alpha, K] = D(c_\alpha \mid S, b)/D(d_\alpha(r, s) \mid T_{r,s}, b) . \tag{2.13}$$

Hájek and Šidák (1967) have explored the relation between this concept and Pitman efficiency. Theorem 3 implies the following result.

THEOREM 4. *If* $b_{r,s}^2 \neq 0$, *then*

$$\lim_{\alpha\to 0} e[S, T_{r,s} \mid \alpha, K] = (b_{11}^2 n_{11}^{-1})/(b_{r,s}^2 n_{r,s}^{-1}) . \tag{2.14}$$

If $b_{r,s}^2 = 0$ *but* $b_{l,m}^2 \neq 0$ *for some* (l, m), *then*

$$\lim_{\alpha\to 0} e[S, T_{r,s} \mid \alpha, K] = \infty . \tag{2.15}$$

Since an S_ν-test is asymptotically strictly unbiased against every alternative $K_\nu(\theta)$ under which $b_{l,m}^2 \neq 0$ for some (l, m), the infinite limit in (2.15) is to be expected. The limit in (2.14) reflects the fact, noted earlier, that the local asymptotic power of an S_ν-test is largely determined, for small α, by the value of b_{11}^2/n_{11}. In particular, if $b_{11}^2 = 0$ and $b_{rs}^2 \neq 0$ for some $(r, s) \neq (1, 1)$, then $\lim_{\alpha\to 0} e[S, T_{r,s} \mid \alpha, K] = 0$; if $b_{11}^2 \neq 0$, then $\lim_{\alpha\to 0} e[S, T_{11} \mid \alpha, K] = 1$.

Within the class of all S_ν-tests, there exist tests which are asymptotically most powerful among symmetric two-sided tests for $H_\nu(\theta)$ at level α. For example, consider the level α test which rejects H_ν for sufficiently large values of

$$V_\nu = [\textstyle\sum_{j=1}^{N_\nu} (d_\nu(j) - \bar{d}_\nu)^2 \int_0^1 \psi_F^2(u)\, du]^{-1} [\sum_{j=1}^{N_\nu} (d_\nu(j) - \bar{d}_\nu) a_\nu(R_{\nu,j}, f)]^2 \tag{2.16}$$

where the score function $a_\nu(\cdot, f)$ is associated with ψ_F in the sense of assumption B(i). The limiting distributions of V_ν under H_ν and under $K_\nu(\theta)$ are, respectively, the same as the distributions of $V = \chi^2(1)$ and $V(\theta c) = \chi^2(1, \theta^2 c^2)$ where $c^2 = \lim_{\nu\to\infty} \sum_{j=1}^{N_\nu} (d_\nu(j) - \bar{d}_\nu)^2 \int_0^1 \psi_F^2(u)\, du$. An argument analogous to the proof of Theorem 4 establishes

$$\lim_{\alpha\to 0} e[S, V \mid \alpha, K] = (n_{11}^{-1} b_{11}^2)/c^2 . \tag{2.17}$$

From this, it may be seen that $\lim_{\alpha\to 0} e[S, V\,|\,\alpha, K] \leqq 1$ for every S; equality occurs if $n_{11} = 1$, $\varphi_1 = [\int_0^1 \psi_F^2(u)\,du]^{-\frac{1}{2}}\psi_F$ and $c_{1,\nu}(j) = [\sum_{j=1}^{N_\nu} (d_\nu(j) - \bar{d}_\nu)^2]^{-\frac{1}{2}} \times (d_\nu(j) - \bar{d}_\nu)$. Thus, any S_ν-test for which $n_{11} = 1$ and $T_{11,\nu} = V_\nu$ (or something asymptotically equivalent to V_ν) is locally asymptotically most powerful among symmetric two-sided tests for H_ν versus $K_\nu(\theta)$ at small levels α. It is interesting to note that the tests of Cramér–von Mises type are optimal in this sense when the underlying density is $f(x) = \pi^{-1}\operatorname{sech}(x)$; for then $\psi_F(u) = -\cos(\pi u)$.

Are the small α results of Theorem 4 and (2.17) adequate as approximations to local asymptotic efficiency when $.01 \leqq \alpha \leqq .10$? Numerical and analytical studies of special cases suggest that the accuracy of the approximation depends upon the noncentrality parameter sequence b involved; for some b, the small α results are reasonably trustworthy. We summarize the evidence available.

Durbin and Knott (1972) and Stephens (1973) have numerically approximated, for certain interesting alternatives, the asymptotic powers of the level .05 Cramér–von Mises, Anderson–Darling, and Watson goodness-of-fit tests. Since the asymptotic distribution theory of these tests is mathematically isomorphic to that of the S_ν-tests, their work is relevant to the present question. Their findings include the following points. For certain families of alternatives for which $b_{11}^2 \neq 0$, the asymptotic power of the level .05 goodness-of-fit tests is nearly the same as the asymptotic power of the corresponding first component tests; it is not closely linked to the asymptotic power of the corresponding second component tests. These properties are consistent with (2.14), in particular with the fact $\lim_{\alpha\to 0} e[S, T_{11}\,|\,\alpha, K] = 1$.

Analytical study of the rate of convergence as $\alpha \to 0$ can be based upon the following observation. Whenever the asymptotic distribution of S_ν under H_ν is known explicitly, it is possible to calculate $D(x\,|\,S, b)$ exactly for all x and hence find $e[S, T_{r,s}\,|\,\alpha, K]$ for all α. We demonstrate this computation in the case of Watson's two-sample statistic $U_{m,n}^2$. Let $\{(m_\nu, n_\nu);\ \nu \geqq 1\}$ be a sequence of sample sizes such that $\min(m_\nu, n_\nu) \to \infty$ as $\nu \to \infty$ and write U_ν^2 in place of U_{m_ν,n_ν}^2. According to Theorems 1 and 2, with suitable adjustments to notation, the limiting distributions of U_ν^2 under H_ν and under $K_\nu(\theta)$ are the same as the distributions of $U^2 = (4\pi^2)^{-1}\sum_{r=1}^\infty r^{-2}\chi_r^2(2)$ and $U^2(b) = (4\pi^2)^{-1}\sum_{r=1}^\infty r^{-2}\chi_r^2(2, b_r^2)$. Watson (1962) showed that

$$(2.18)\qquad P[U^2 > x] = 2\sum_{k=1}^\infty (-1)^{k-1}\exp(-2\pi^2k^2x) \qquad \text{for} \quad x \geqq 0\,.$$

From (2.3), $D(x|U^2, b) = \sum_r b_r^2 D_r(x|U^2)$, where $D_r(x|U^2) = \partial P[U^2(b) > x]/\partial b_r^2|_{b=0}$. A calculation using (2.18) and the middle line of (2.4) yields

$$(2.19)\qquad \begin{aligned} D_r(x\,|\,U^2) = {} & (-1)^{r-1}(2\pi^2r^2x - 3/4)\exp(-2\pi^2r^2x) \\ & + \sum_{k\neq r} (-1)^{k-1}k^2(r^2 - k^2)^{-1}\exp(-2\pi^2k^2x)\,, \end{aligned}$$

for every integer $r \geqq 1$ and $x \geqq 0$.

Applied to U^2, Theorem 3 produces the following expansions, valid for every integer $m \geqq 1$:

$$
\begin{aligned}
P[U^2 > x] &= [2 + O(x^{-m})] \exp(-2\pi^2 x) \\
D(x \mid U^2, b) &= [(2\pi^2 x - 3/4)b_1^2 + \textstyle\sum_{r=2}^{\infty} (r^2 - 1)^{-1} b_r^2 \\
&\quad + O(x^{-m})] \exp(-2\pi^2 x) .
\end{aligned} \tag{2.20}
$$

It is interesting to note that the errors incurred in using (2.20) instead of the exact results (2.18) and (2.19) decrease exponentially in x. This phenomenon explains the vanishing of all higher order terms in the expansions (2.20).

Let c_α now denote the critical value for which $P[U^2 > c_\alpha] = \alpha$. Table 1 records values of $(2\pi^2 \alpha c_\alpha)^{-1} D_r(c_\alpha \mid U^2)$ calculated from (2.19) for $1 \leqq r \leqq 4$ and $\alpha = .10, .05, .01$ and the corresponding limits as $\alpha \to 0$. (A factor of $4\pi^2$ enters (2.11) in the case of U^2 because U^2 does not have the same normalization as S.) Since the rate of convergence is gradual, (2.11) will, at best, give only a first approximation to $D(c_\alpha \mid U^2, b)$.

By grouping together equally weighted terms, the statistic U_ν^2 can be written in the form $U_\nu^2 = (4\pi^2)^{-1} \sum_r r^{-2} T_{r,\nu}$. The limiting distributions of $T_{r,\nu}$ under H_ν and under $K_\nu(\theta)$ are the same as the distributions of $T_r = \chi_r^2(2)$ and $T_r(\theta b) = \chi_r^2(\theta^2 b_r^2)$, respectively. From (2.4),

$$
D_r(x \mid T_r) = (x/4) \exp(-x/2) \tag{2.21}
$$

for all $r \geqq 1$ and $x \geqq 0$. If $b_r^2 \neq 0$ and d_α is such that $P[\chi^2(2) > d_\alpha] = \alpha$, then from (2.3) and (2.13),

$$
e[U^2, T_r \mid \alpha, K] = \textstyle\sum_{k=1}^{\infty} (b_k^2/b_r^2) E_k[U^2, T_r \mid \alpha] , \tag{2.22}
$$

TABLE 1
Values of $(2\pi^2 \alpha c_\alpha)^{-1} D_r(c_\alpha \mid U^2)$

r	α = .10	.05	.01	0
1	.373	.397	.433	.500
2	.055	.045	.032	0
3	.021	.017	.012	0
4	.011	.009	.006	0

TABLE 2
Values of $E_k(U^2, T_r \mid \alpha)$

k	α = .10	.05	.01	0
1	.973	.980	.994	1.000
2	.143	.111	.073	0
3	.054	.042	.027	0
4	.029	.022	.015	0

where $E_k[U^2, T_r \mid \alpha] = D_k(c_\alpha \mid U^2)/D_r(d_\alpha \mid T_r)$; in the present example, $E_k[U^2, T_r \mid \alpha]$ is the same for all $r \geqq 1$.

Table 2 records values of $E_k[U^2, T_r \mid \alpha]$ for $1 \leqq k \leqq 4$ and $\alpha = .10, .05, .01$ and the corresponding limits as $\alpha \to 0$, calculated from (2.14). The rate of convergence is swift when $k = 1$ and slower when $k \geqq 2$. Thus, (2.14) approximates $e[U^2, T_r \mid \alpha, K]$ reasonably well for some noncentrality parameter sequences b, but not all.

The considerations above indicate that, in some instances, local asymptotic power of an S_ν-test (or its small α asymptote) provides an acceptable analytical approximation to asymptotic power. Does asymptotic power of an S_ν-test yield a reliable approximation to exact power for interesting sample sizes? Though some encouraging Monte Carlo results have been reported by Stephens (1973), we have no theoretical answer.

3. **Proofs.** This section contains proofs for the theorems of Section 2. To simplify notation, the subscript ν will often be dropped.

Proof of Theorem 1. Let $Z_{r,s,\nu}$ and $Z_{r,s}$ denote the (r, s)th elements in the double sequences Z_ν and Z, respectively. From Wichura (1971), necessary and sufficient conditions for the desired weak convergence in the l_2-topology are

(a) For every (L, M), the random matrix $\{Z_{r,s,\nu}; 1 \leqq r \leqq L, 1 \leqq s \leqq M\}$ converges in distribution under H_ν to the random matrix

$$\{Z_{r,s}; 1 \leqq r \leqq L, 1 \leqq s \leqq M\} \qquad \text{as} \quad \nu \to \infty .$$

(b) If $W_{L,M,\nu} = \sum_{r>L,s>M} Z^2_{r,s,\nu}$,

$$\lim_{\min(L,M)\to\infty} \limsup_\nu P_{H_\nu}[W_{L,M,\nu} > \varepsilon] = 0 \qquad \text{for every } \varepsilon > 0 .$$

To verify (a), consider the arbitrary linear combinations $K_{1,\nu} = \sum_{r=1}^L \times \sum_{s=1}^M \gamma_{r,s}[\sum_{j=1}^N c_r(j)a_s(R_j)]$ and $K_{2,\nu} = \sum_{r=1}^L \sum_{s=1}^M \gamma_{r,s}[\sum_{j=1}^N c_r(j)\varphi_s(U_j)]$. Application of Theorem V. 1.5a or Theorem V. 1.6a in Hájek and Šidák (1967), according to whether B(i) or B(ii) is assumed, yields $\lim_{\nu\to\infty} E|K_{1,\nu} - K_{2,\nu}| = 0$. The linear form $K_{2,\nu}$ can also be written as $K_{2,\nu} = \sum_{j=1}^N \mathbf{g}_j'\mathbf{Y}_j$, where $\mathbf{g}_j' = (\sum_{r=1}^L \gamma_{r,1}c_r(j), \cdots, \sum_{r=1}^L \gamma_{r,M}c_r(j))$ and $\mathbf{Y}_j' = (\varphi_1(U_j), \cdots, \varphi_M(U_j))$. If $||\cdot||$ denotes Euclidean distance, assumption C implies that $\lim_{\nu\to\infty} \max_{1\leq j\leq N_\nu} ||\mathbf{g}_j||^2 = 0$ and $\sum_{j=1}^{N_\nu} ||\mathbf{g}_j||^2 = \sum_{r=1}^L \sum_{s=1}^M \gamma^2_{r,s} < \infty$. It follows from the Lindberg–Feller theorem that the distribution of $K_{2,\nu}$ is asymptotically normal $(0, \sum_{r=1}^L \sum_{s=1}^M \gamma^2_{r,s})$ and consequently, condition (a) above is fulfilled (see also Beran (1970) regarding the central limit theorem used here).

To verify condition (b) under assumption B(i), note that by Jensen's inequality for conditional expectations,

$$\begin{aligned} E_{H_\nu}[W_{L,M,\nu}] &\leqq \textstyle\sum_{r>L}^{N-1} \sum_{s>M}^{N-1} \alpha^2_{r,s} E_{H_\nu}[\sum_{j=1}^N c_r(j)\varphi_s(U_j)]^2 \\ &< \textstyle\sum_{r>L} \sum_{s>M} \alpha^2_{r,s} , \end{aligned} \tag{3.1}$$

which implies (b). Direct evaluation of $E_{H_\nu}[W_{L,M,\nu}]$ under assumption B(ii) yields a similar bound, again implying that (b) is satisfied.

PROOF OF THEOREM 2. It is sufficient to verify under $K_\nu(\theta)$ the counterparts to conditions (a) and (b) in the previous proof. Since the alternatives $K_\nu(\theta)$ are contiguous to the hypotheses H_ν, by Theorem VI. 2.1 of Hájek and Šidák (1967), fulfillment of (b) under $K_\nu(\theta)$ is evident. Let $Z_{r,s}(\theta\mu)$ denote the (r, s)th element of the double sequence $Z(\theta\mu)$. The usual contiguity argument shows that the random matrix $\{Z_{r,s,\nu};\ 1 \leqq r \leqq L,\ 1 \leqq s \leqq M\}$ converges in distribution under $K_\nu(\theta)$ to the random matrix $\{Z_{r,s}(\theta\mu);\ 1 \leqq r \leqq L,\ 1 \leqq s \leqq M\}$ as $\nu \to \infty$. This establishes (a) and completes the proof.

PROOF OF THEOREM 3. To justify expansion (2.8), it is sufficient to show that the relation

$$C_m = \lim_{x\to\infty} x^m \left[\frac{P[S > x]}{P[\chi^2_{11}(n_{11}) > x]} - \textstyle\sum_{k=0}^{m-1} C_k x^{-k}\right] \tag{3.2}$$

is satisfied for every integer $m \geqq 1$. Let $p_s(x)$ and $p(x \mid n_{11})$ denote the densities of S and of $\chi^2_{11}(n_{11})$ respectively. By Hoeffding (1964), slightly extended,

$$p_s(x) = [\textstyle\sum_{k=0}^{m-1} A_k x^{-k} + O(x^{-m})]p(x \mid n_{11}) \tag{3.3}$$

for every integer $m \geqq 1$, the $\{A_k\}$ being defined by (2.5). In particular, (3.3) implies the validity of (3.2) for $m = 1$. Assuming that (3.2) holds for every integer $m \leqq r$, we will show that it also holds for $m = r + 1$.

Write χ^2 in place of $\chi^2_{11}(n_{11})$. By Cauchy's mean value theorem, for every $x < y$ there exists $w \in (x, y)$ such that

$$\begin{aligned}\{x^{r+1}[P(S > x) - P(\chi^2 > x) \textstyle\sum_{i=0}^{r} C_i x^{-i}]& \\ - y^{r+1}[P(S > y) - P(\chi^2 > y) \textstyle\sum_{i=0}^{r} C_i y^{-i}]\}& \\ \div \{P(\chi^2 > x) - P(\chi^2 > y)\}& \\ = I_1(w) + I_2(w)\,,&\end{aligned} \tag{3.4}$$

where

$$\begin{aligned} I_1(w) &= (r + 1)w^r[P(S > w)/P(\chi^2 > w) - \textstyle\sum_{i=0}^{r} C_i w^{-i}] \\ &\qquad \times [P(\chi^2 > w)/p(w \mid n_{11})] \\ I_2(w) &= w^{r+1}[p_s(w)/p(w \mid n_{11}) - \textstyle\sum_{i=0}^{r} A_i w^{-i}] + \textstyle\sum_{i=0}^{r} (A_i - C_i)w^{r+1-i} \\ &\qquad + [P(\chi^2 > w)/p(w \mid n_{11})] \textstyle\sum_{i=0}^{r} iC_i w^{r-i}\,. \end{aligned} \tag{3.5}$$

A calculation using the inductive hypothesis, (2.10), Hoeffding's expansion (3.3), and the recursion defining the $\{C_k\}$ shows that $I_1(w) + I_2(w) = C_{r+1} + o(1)$. By first choosing x and y sufficiently large, then letting $y \to \infty$, we obtain (3.2) for $m = r + 1$.

Expansion (2.9) may be proved as follows. From (2.3) and (2.4), $D(x \mid S, b) = \sum_{r,s} b^2_{r,s}\sigma^2_{r,s} g(x \mid S^*_{r,s})$. By slight extension of Hoeffding (1964),

$$g(x \mid S^*_{11}) = [\textstyle\sum_{k=0}^{m-1} B_k x^{-k} + O(x^{-m})]p(x \mid n_{11} + 2) \tag{3.6}$$

and, for every $(r, s) \neq (1, 1)$,

$$g(x \mid S^*_{r,s}) = [\textstyle\sum_{k=0}^{m-1} A_k(r, s)x^{-k} + R_m(x \mid r, s)]p(x \mid n_{11})\,, \tag{3.7}$$

where the remainder terms $R_m(x \mid r, s)$ are individually $O(x^{-m})$. To complete the proof of (2.9), it remains to show that $\sup_{(r,s)\neq(1,1)} |R_m(x \mid r, s)| = O(x^{-m})$.

Let $H_{r,s}(x)$, $h_{r,s}(x)$ denote, respectively, the distribution function and density of the random variable $\xi_{r,s} = S^*_{r,s} - \chi^2_{11}(n_{11})$. Following Hoeffding (1964), we express $R_m(x \mid r, s)$ as the sum of three terms,

$$(3.8)\qquad \begin{aligned} J_{1,r,s} &= \int_0^{\delta x} [(1-(y/x))^{n_{11}/2-1} - \textstyle\sum_{k=0}^{m-1} (-1)^k \binom{n_{11}/2-1}{k} (y/x)^k] \\ &\qquad \times \exp(y/2)\, dH_{r,s}(y)\,, \\ J_{2,r,s} &= \int_{\delta x}^{x} (1-(y/x))^{n_{11}/2-1} \exp(y/2)\, dH_{r,s}(y)\,, \\ J_{3,r,s} &= -\int_{\delta x}^{\infty} [\textstyle\sum_{k=0}^{m-1} (-1)^k \binom{n_{11}/2-1}{k} (y/x)^k] \exp(y/2)\, dH_{r,s}(y)\,, \end{aligned}$$

where $0 < \delta < 1$. On the one hand, $|J_{1,r,s}|$ and $|J_{3,r,s}|$ are bounded from above by a constant multiple of $x^{-m}E[\xi^m_{r,s} \exp(\xi_{r,s}/2)]$, and $\sup_{(r,s)\neq(1,1)} E[\xi^m_{r,s} \exp(\xi_{r,s}/2)]$ is finite. On the other hand, $|J_{2,r,s}|$ is bounded from above by a constant multiple of $x^{-m} \sup_{y \geqq \delta x} [y^{m+1} \exp(y/2) h_{r,s}(y)]$, which implies, after a short argument, that $\sup_{(r,s)\neq(1,1)} |J_{2,r,s}| = o(x^{-m})$.

Proof of Theorem 4. The definition of c_α through $P[S > c_\alpha] = \alpha$ and expansions (2.8), (2.10) imply that for small α

$$(3.9)\qquad \alpha \sim 2C_0 p(c_\alpha \mid n_{11})\,.$$

For each (r, s), let $\{e_\alpha(r, s); 0 < \alpha \leqq 1\}$ be a set of constants such that $\lim_{\alpha\to 0} [(n_{r,s} - n_{11}) \log(c_\alpha) + (c_\alpha - e_\alpha(r, s))] = 2 \log(L)$, where $L = 2^{(n_{r,s}-n_{11})/2} \times C_0 \Gamma(n_{r,s}/2)/\Gamma(n_{11}/2)$. From (2.10) and (3.9), $P[T_{r,s} > e_\alpha(r, s)] \sim 2p(e_\alpha(r, s) \mid n_{r,s}) \sim \alpha$.

Let $M_\alpha = \max[d_\alpha(r, s), e_\alpha(r, s)]$ and let $m_\alpha = \min[d_\alpha(r, s), e_\alpha(r, s)]$. By the mean value theorem, there exists $u_\alpha \in (m_\alpha, M_\alpha)$ such that

$$|P(T_{r,s} > m_\alpha)/P(T_{r,s} > M_\alpha) - 1| \geqq (M_\alpha - m_\alpha) p(u_\alpha \mid n_{r,s})/P(T_{r,s} > u_\alpha)\,.$$

Because of (2.10) and the previous paragraph, this in turn implies $\lim_{\alpha\to 0} [d_\alpha(r, s) - e_\alpha(r, s)] = 0$. Hence, $\lim_{\alpha\to 0} [d_\alpha(r, s)/c_\alpha] = 1$. The theorem now follows readily from (2.11) and (2.12).

Acknowledgment. I am grateful to Michael Stephens for helpful conversations concerning the asymptotic power of goodness-of-fit tests.

REFERENCES

[1] Beran, R. J. (1969). The derivation of nonparametric two-sample tests from tests for uniformity of a circular distribution. *Biometrika* **56** 561–570.

[2] Beran, R. J. (1970). Linear rank statistics under alternatives indexed by a vector parameter. *Ann. Math. Statist.* **41** 1896–1905.

[3] Blum, J. R., Kiefer, J. and Rosenblatt, M. (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.* **32** 485–492.

[4] Durbin, J. and Knott, M. (1972). Components of Cramér–von Mises statistics I. *J. Roy. Statist. Soc. Ser. B* **34** 290–307.

[5] Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[6] Hoeffding, W. (1964). On a theorem of V. M. Zolotarev. *Theor. Probability Appl.* **9** 89–92.

[7] Schach, Siegfried (1969). On a class of nonparametric two-sample tests for circular distributions. *Ann. Math. Statist.* **40** 1791–1800.

[8] STEPHENS, M. A. (1973). Components of goodness-of-fit statistics. Unpublished paper.
[9] WATSON, G. S. (1962). Goodness-of-fit tests on a circle II. *Biometrika* **49** 57-63.
[10] WHITTAKER, E. T. and WATSON, G. N. (1927). *A Course of Modern Analysis*. Cambridge Univ. Press.
[11] WICHURA, M. J. (1971). A note on the weak convergence of stochastic processes. *Ann. Math. Statist.* 42 1769-1772.
[12] WITHERS, CHRISTOPHER S. (1970). Power and efficiency of a class of goodness-of-fit tests. Technical Report, Dept. of Statistics, Stanford Univ.
[13] ZOLOTAREV, V. M. (1961). Concerning a certain probability problem. *Theor. Probability Appl.* **6** 201-204.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720