The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 1011–1020

THE ASYMPTOTIC DISTRIBUTION OF SOME NON LINEAR FUNCTIONS OF THE TWO-SAMPLE RANK VECTOR[1]

BY SIEGFRIED SCHACH[2]

University of Minnesota and Stanford University

1. Introduction. Let $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ be independent samples from the same continuous distribution. Let R_{Ni} $(i = 1, 2, \cdots, m)$ be the rank of X_i in the combined sample $(N = m + n)$, and define

$Z_i = Z_{Ni} = 1$ if the ith element of the combined ordered sample is an X observation
$= 0$ otherwise.

Let $A^{(N)} = (a_{i,j}^{(N)})$ be a sequence of symmetric matrices. We find conditions under which a statistic of the form

$$S_N = \sum_{i=1}^{m} \sum_{j=1}^{m} a_{R_{Ni}, R_{Nj}}^{(N)} = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{i,j}^{(N)} Z_i Z_j \tag{1.1}$$

converges in distribution as $N \to \infty$ and $mn^{-1} \to \lambda$, $0 < \lambda < 1$.

Several examples of non-parametric test statistics of the form (1.1) can be found in the literature. In fact, any two-sided symmetric test based on a linear rank statistic is trivially equivalent to a test of the form $S_N \geqq C$. More interesting examples arise in the area of non-parametric statistics for circular distributions. Wheeler and Watson [6] proposed a two-sample non-parametric test for circular distributions of the form (1.1), which is related to the usual parametric test (likelihood ratio test for the class of v. Mises distributions) in much the same way as the Wilcoxon test is related to the two-sample t-test. The author [4] found that in detecting rotation alternatives for circular distributions a locally most powerful invariant test under a suitable group of transformations is based on a statistic of the form (1.1). Matthes and Truax [2] obtained a test statistic related to (1.1) when deriving locally most powerful invariant tests for two-sided shift alternatives.

2. Step function representation. Degenerate and non-degenerate limit functions. Let $h_N(\cdot, \cdot)$ be a step function on the unit square, which is constant on subsquares of the form $(i-1)/N < x \leqq i/N$, $(j-1)/N < y \leqq j/N$ and which is symmetric, i.e., $h_N(x, y) = h_N(y, x)$. If we set $h_N(i/N, j/N) = a_{i,j}^{(N)}$ then obviously every statistic of the form (1.1) can be written as

$$S_N = \sum_{i=1}^{m} \sum_{j=1}^{m} h_N(R_{Ni}/N, R_{Nj}/N) = \sum_{i=1}^{N} \sum_{j=1}^{N} h_N(i/N, j/N) Z_i Z_j \tag{2.1}$$

and vice versa. Since conditions on covergence can be stated more easily in terms of sequences $\{h_N\}$ we will use representation (2.1) in the sequel.

Assume that there exists a function $h(\cdot, \cdot)$ such that $h_N \to h$ in $L_2(I^2)$, i.e.,

Received 15 May 1968.
[1] This report was supported by the National Science Foundation (GP-6859) and by U. S. Public Health Service Grant (USPHS-GM-14554-02).
[2] Now at The Johns Hopkins University.

in the sense of convergence in the Hilbert space of square integrable functions on the unit square. By Fubini's theorem the function

$$g(x) = \int_0^1 h(x, y)\, dy \tag{2.2}$$

exists a.e. For convenience we assume that

$$\int_0^1 \int_0^1 h_N(x, y)\, dx\, dy = 0 \tag{2.3}$$

for each N (this can always be achieved by a suitable standardization of S_N). Then $\int_0^1 \int_0^1 h(x, y)\, dx\, dy = 0$.

DEFINITION. We say $h(\cdot\,, \cdot)$ is degenerate if $g(x) = \int_0^1 h(x, y)\, dy = 0$ a.e.

3. Asymptotic distribution of S_N if h is non-degenerate.

LEMMA 3.1. *If $h_N(\cdot\,, \cdot) \to h(\cdot\,, \cdot)$ in $L_2(I^2)$, then*

$$\int_0^1 h_N(\cdot\,, y)\, dy \to \int_0^1 h(\cdot\,, y)\, dy \tag{3.1}$$

in L_2 (i.e., *in the space of square integrable functions on* [0, 1]).

PROOF. Straightforward.

If we set

$$h'(x, y) = h(x, y) - g(x) - g(y), \tag{3.2}$$

then $h'(x, y)$ is symmetric and degenerate, since

$$\int_0^1 h'(x, y)\, dy = \int_0^1 h(x, y)\, dy - g(x) - \int_0^1 g(y)\, dy \equiv 0. \tag{3.3}$$

We define corresponding step functions

$$g_N(x) = \int_0^1 h_N(x, y)\, dy, \tag{3.4}$$

and

$$h_N'(x, y) = h_N(x, y) - g_N(x) - g_N(y). \tag{3.5}$$

h_N' is symmetric and degenerate. Now set

$$T_N = \sum_{i=1}^N \sum_{j=1}^N h_N'(i/N, j/N) Z_i Z_j\,, \tag{3.6}$$

then

$$\begin{aligned} S_N &= T_N + \sum_{i=1}^N \sum_{j=1}^N g_N(i/N) Z_i Z_j + \sum_{i=1}^N \sum_{j=1}^N g_N(j/N) Z_i Z_j \\ &= T_N + 2m \sum_{i=1}^N g_N(i/N) Z_i \\ &= T_N + 2m U_N \end{aligned} \tag{3.7}$$

where

$$U_N = \sum_{i=1}^N g_N(i/N) Z_i = \sum_{i=1}^m g_N(R_i/N). \tag{3.8}$$

THEOREM 3.2. *If $h_N(\cdot\,, \cdot) \to_{\text{m.s.}} h(\cdot\,, \cdot)$ then $N^{-\frac{1}{2}} U_N$ is asymptotically normal with*

$$\mu_N = N^{-\frac{1}{2}} E U_N = 0, \qquad \sigma^2 = \lambda(1 - \lambda) \int_0^1 g(x)^2\, dx. \tag{3.9}$$

PROOF. This result follows immediately from Theorem V.1.6. a of Hájek and Sidák [1] provided that

$$g_N(\cdot) \to g(\cdot) \text{ in m.s.}, \qquad \text{as} \quad N \to \infty.$$

But by Lemma 3.1 this condition is satisfied.

LEMMA 3.3. *If $h_N \to_{\text{m.s.}} h$, then $h_N' \to_{\text{m.s.}} h'$.*

PROOF. By Lemma 3.1 $g_N \to_{\text{m.s.}} g$. Since

$$h_N'(x, y) - h'(x, y) = h_N(x, y) - h(x, y) + g_N(x) - g(x) + g_N(y) - g(y)$$

the desired result follows.

LEMMA 3.4. *If $h_N \to_{\text{m.s.}} h$, $\int_0^1 h_N(x, x)^2\, dx$ is bounded and*

$$\int_0^1 h_N(x, y)\, dx = 0, \tag{3.10}$$

then for S_N defined by (2.1) we get

$$\begin{aligned} N^{-1}ES_N &= mN^{-1}(1 - (m-1)/(N-1)) \int_0^1 h_N(x, x)\, dx \\ &= \lambda(1-\lambda) \int_0^1 h_N(x, x)\, dx + o_N(1), \end{aligned} \tag{3.11}$$

$$N^{-2} \operatorname{Var} S_N = 2\lambda^2(1-\lambda)^2 \|h\|_{L_2}^2 + o_N(1). \tag{3.12}$$

PROOF. Obviously,

$$\begin{aligned} EZ_iZ_j &= m/N && \text{if } i = j \\ &= (m/N)(m-1)/(N-1) && \text{if } i \neq j. \end{aligned} \tag{3.13}$$

Hence

$$\begin{aligned} ES_N &= \textstyle\sum_{i=1}^N \sum_{j=1}^N h_N(i/N, j/N) EZ_iZ_j \\ &= [(m/N)((m - 1)/N - 1)] \textstyle\sum_{i=1}^N \sum_{j=1}^N h_N(i/N, j/N) \\ &\quad + (m/N)(1 - (m-1)/(N-1)) \textstyle\sum_{i=1}^N h_N(i/N, j/N) \\ &= N \cdot (m/N)(1 - (m-1)/(N-1)) \int_0^1 h_N(x, x)\, dx. \end{aligned} \tag{3.14}$$

Since $\int_0^1 h_N(x, x)\, dx$ is bounded $((\int_0^1 h_N(x, x)\, dx)^2 \leqq \int_0^1 h_N(x, x)^2\, dx)$ and since $m/N \to \lambda$ (3.11) follows.

A straightforward, but tedious, calculation (using (3.10) repeatedly) shows that

$$\begin{aligned} ES_N^2 &= 2N^2\lambda^2(1-\lambda)^2 \|h_N(\cdot\,, \cdot)\|_{L_2}^2 \\ &\quad + N^2\lambda^2(1-\lambda)^2 (\textstyle\int_0^1 h_N(x, x)\, dx)^2 + O(N). \end{aligned} \tag{3.15}$$

Hence

$$N^{-2} \operatorname{Var} S_N = 2\lambda^2(1-\lambda)^2 \|h\|_{L_2}^2 + o_N(1).$$

We are now able to prove the main result of this section.

THEOREM 3.5. *If $h_N \to_{\text{m.s.}} h$, h non-degenerate, $\int h_N(x, x)^2 dx$ bounded, then* *for* S_N *defined by* (2.1) *we have*

$$\lim_{N\to\infty} P[N^{-\frac{3}{2}}S_N \leqq x\sigma] = \int_{-\infty}^{x} (2\pi)^{-\frac{1}{2}} \exp(-t^2/2)\, dt \tag{3.16}$$

with

$$\sigma^2 = 4\lambda^3(1-\lambda)\|g\|_{L_2}^2, \tag{3.17}$$

where $g(\cdot)$ *is defined by* (2.2).

PROOF. By (3.7) we have

$$S_N = T_N + 2mU_N.$$

T_N, defined by (3.6), is based on the function h_N' defined by (3.5). By Lemma 3.3 $h_N' \to_{\text{m.s.}} h'$, and thus h_N' satisfies the assumptions of Lemma 3.4. Hence by (3.11) and (3.12),

$$N^{-3}ET_N^2 \to 0 \quad \text{as} \quad N \to \infty, \tag{3.18}$$

which implies that $N^{-\frac{3}{2}}T_N \to 0$ in probability. From Theorem 3.2 we obtain asymptotic normality of $N^{-\frac{3}{2}}mU_N$. Thus $N^{-\frac{3}{2}}S_N$ is asymptotically $N(0, \sigma^2)$, where σ^2 is defined by (3.17).

REMARK 3.6. Under the conditions of Theorem 3.6 the asymptotic distribution of S_N depends on h_N only through the "marginals" g_N; in fact, it depends only on $\|g\|^2 = \lim_{N\to\infty} \|g_N\|^2$.

4. Asymptotic distribution of S_N if h is degenerate. In this section we obtain the asymptotic distribution of S_N if h is degenerate, provided that certain conditions on h and on the approximating sequence h_N are satisfied.

We first obtain the limiting distribution for a particular sequence $\{h_N\}$ of approximating step functions (Theorem 4.7). Then we extend the results to more general approximating sequences (Theorem 4.10).

Let $h(\cdot, \cdot)$ be a symmetric element of $L_2(I^2)$. Then it is well-known that the relation $A: f \to g$ defined by

$$g(x) = \int_0^1 h(x, y)f(y)\, dy \tag{4.1}$$

is a compact operator which maps L_2 into L_2. The spectrum of A consists of a countable number of real eigenvalues λ_k, such that $\sum_{k=1}^{\infty} \lambda_k^2 < \infty$. Let $\{\varphi_k\}$ be a corresponding sequence of eigenfunctions. Then by the spectral theorem

$$h(x, y) = \sum_{k=1}^{\infty} \lambda_k\varphi_k(x)\varphi_k(y), \qquad \int_0^1 \varphi_k(x)\varphi_l(x)\, dx = \delta_{kl}. \tag{4.2}$$

LEMMA 4.1. *If h is degenerate, then* $\int_0^1 \varphi_k(x)\, dx = 0$ *for all* k.

PROOF. Straightforward.

If h is any element of $L_2(I^2)$, then the projection h_{NP} of h onto the space of step functions of order N is characterized by

$$h_{NP}(i/N, j/N) = N^2 \int_{I_{Ni}} \int_{I_{Nj}} h(x, y)\, dx\, dy, \qquad 1 \leqq i, j \leqq N, \tag{4.3}$$

where $I_{Ni} = \{x: (i-1)/N < x \leqq i/N\}$. h_{NP} is also the conditional expectation $E^{\mathfrak{F}_N}h$, where $\mathfrak{F}_N$ is the field generated by squares $I_{Ni} \times I_{Nj}$. Using the continuity of the projection operator it is easy to see that from the expansion (4.2) we get

$$h_{NP}(x, y) = \sum_{k=1}^{\infty} \lambda_k \varphi_{Nk}(x)\varphi_{Nk}(y) \tag{4.4}$$

where $\varphi_{Nk}(\cdot)$ is the projection of $\varphi_k(\cdot)$ onto the space of step functions of one variable, i.e., for $x \varepsilon I_i$ $(1 \leqq i \leqq N)$

$$\varphi_{Nk}(x) = N \int_{I_i} \varphi_k(u)\, du. \tag{4.5}$$

Set $\nu_k = |\lambda_k|^{\frac{1}{2}}$, $k = 1, 2, \cdots$, and

$$\eta_{Nk} = N^{-\frac{1}{2}} \nu_k \sum_{i=1}^{N} \varphi_{Nk}(i/N) Z_i, \qquad k = 1, 2, \cdots. \tag{4.6}$$

THEOREM 4.2. *Let $\{\varphi_k : k = 1, 2, \cdots, K\}$ be an orthonormal set in L_2 such that $\int_0^1 \varphi_k(x)\, dx = 0$. Define the step functions φ_{Nk} by (4.5). Then as $N \to \infty$ the joint distribution of $(\eta_{N1}, \eta_{N2}, \cdots, \eta_{NK})$ is asymptotically normal with mean vector 0 and covariance matrix (σ_{kl}), where $\sigma_{kl} = \lambda(1 - \lambda)\delta_{kl}\nu_k^2$.*

PROOF. It suffices to show that any linear combination $\xi_N = \sum_{k=1}^{K} t_k \eta_{Nk}$ is asymptotically normal with $E\xi_N = 0$, $\operatorname{Var} \xi_N = \lambda(1 - \lambda) \sum_{k=1}^{K} t_k^2 \nu_k^2$. From (4.6)

$$\xi_N = N^{-\frac{1}{2}} \sum_{k=1}^{K} t_k \nu_k \sum_{i=1}^{N} \varphi_{Nk}(i/N) Z_i = N^{-\frac{1}{2}} \sum_{i=1}^{N} \sum_{k=1}^{K} t_k \nu_k \varphi_{Nk}(i/N) Z_i. \tag{4.7}$$

Obviously $E\xi_N = 0$. Since $\varphi_{Nk} \to \varphi_k$ in m.s. for each k, we get $\sum_{k=1}^{K} t_k \nu_k \varphi_{Nk}(\cdot) \to \sum_{k=1}^{K} t_k \nu_k \varphi_k(\cdot)$ in L_2. Hence the theorem by Hájek and Sidák mentioned above implies asymptotic normality of ξ_N with

$$\mu = 0 \quad \text{and} \quad \sigma^2 = \lambda(1 - \lambda) \int_0^1 \left(\sum t_k \nu_k \varphi_k(x)\right)^2 dx.$$

Since $\{\varphi_k\}$ is orthonormal we get $\sigma^2 = \lambda(1 - \lambda) \sum_{k=1}^{K} t_k^2 \nu_k^2$ q.e.d.

COROLLARY 4.3. *If $h(\cdot, \cdot)$ is a degenerate function with a finite expansion*

$$h(x, y) = \sum_{k=1}^{K} \lambda_k \varphi_k(x) \varphi_k(y), \tag{4.8}$$

then the asymptotic distribution of

$$N^{-1} S_N = N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} h_{NP}(i/N, j/N) Z_i Z_j \tag{4.9}$$

exists and is equal to the distribution of $\lambda(1 - \lambda) \sum_{k=1}^{K} \lambda_k \chi_{1}^{2}{}_{(k)}$, where $\chi_1^2{}_{(k)}$, $k = 1, 2, \cdots, K$, are independent χ^2 random variables with 1 d.f.

PROOF. Because of (4.4) we have

$$\begin{aligned} N^{-1} S_N &= N^{-1} \sum_{k=1}^{K} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_k \varphi_{Nk}(i/N) \varphi_{Nk}(j/N) Z_i Z_j \\ &= \sum_{k=1}^{K} \operatorname{sgn}(\lambda_k) (\eta_{Nk})^2, \end{aligned} \tag{4.10}$$

and hence the result follows immediately from Theorem 4.2.

For h with an infinite expansion we use the theory of measures on separable Hilbert spaces in order to derive the limiting distribution of $N^{-1} S_N$, where S_N is based on a sequence $\{h_N\}$ defined by

$$h_N(x, y) = \sum_{k=1}^{N} \lambda_k \varphi_{Nk}(x) \varphi_{Nk}(y). \tag{4.11}$$

Let H be the Hilbert space of real sequences with finite sums of squares. Then we define a sequence $\{M_N\}$ of mappings from the space of the Z's to H by

$$M_N\colon (Z_{N1}, \cdots, Z_{NN}) \to (\eta_{N1}, \cdots, \eta_{NN}, 0, 0, \cdots) \tag{4.12}$$

where η_{Nk} is defined by (4.6). These mappings are measurable and hence they induce a sequence $\{\mu_N\}$ of probability measures on H.

The set of probability measures on H can be metrized in such a way that convergence in this metric is equivalent to weak convergence. With this metric the space of all probability measures on H is a complete separable metric space (Prokhorov [3]).

DEFINITION 4.4. We say that $h(\cdot, \cdot)$ is a trace-class function if

$$\sum_{k=1}^{\infty} |\lambda_k| < \infty. \tag{4.13}$$

From now on we always assume that $h(\cdot, \cdot)$ is a symmetric, degenerate, trace-class function.

THEOREM 4.5. *The sequence $\{\mu_N\}$ of probability measures is sequentially compact.*

PROOF. According to the corrected version of Prokhorov's Theorem 1.13 it suffices to show that

(i) $\sup_N \int_H \|x\|^2 \, d\mu_N < \infty$

(ii) $\lim_{L\to\infty} \sup_N \int_H \sum_{l=L}^{\infty} x_l^2 \, d\mu_N = 0,$

where $x = (x_1, x_2, \cdots)$ is a generic element of H. Since $x_l = 0$ a.e. (μ_N) for $l > N$, we have

$$\int_H \sum_{l=L}^{\infty} x_l^2 \, d\mu_N = \int_H \sum_{l=L}^{N} x_l^2 \, d\mu_N . \tag{4.14}$$

For $k \leqq N$ we have

$$\begin{aligned} \int_H x_k^2 \, d\mu_N &= E(\eta_{Nk})^2 = N^{-1}\nu_k^2 E[\textstyle\sum_{i=1}^{N} \sum_{j=1}^{N} \varphi_{Nk}(i/N)\varphi_{Nk}(j/N) Z_i Z_j] \\ &= \nu_k^2 m N^{-1}(1 - (m-1)/(N-1))N^{-1} \textstyle\sum_{i=1}^{N} \varphi_{Nk}(i/N)^2 \\ &\leqq \nu_k^2 \|\varphi_{Nk}\|^2 \leqq |\lambda_k|. \end{aligned} \tag{4.15}$$

Combining (4.14) and (4.15) we obtain

$$\int_H \sum_{l=L}^{\infty} x_l^2 \, d\mu_N \leqq \sum_{l=L}^{\infty} |\lambda_l|, \tag{4.16}$$

which shows that (i) and (ii) are satisfied. This completes our proof.

Let $(\cdot, \cdot)$ be the inner product on H. For every $f \varepsilon H$ and every probability measure μ we define the characteristic functional

$$\chi(f, \mu) = \int_H e^{i(f,x)} \, d\mu(x). \tag{4.17}$$

$\chi(\cdot, \mu)$ is continuous on H, and it determines μ uniquely.

THEOREM 4.6. *$\{\mu_N\}$ converges to the Gaussian measure with mean* 0 *and S-operator of the form $[S]_{k,l} = \lambda(1-\lambda)\delta_{k,l} |\lambda_k|$.*

PROOF. If $\mu_N \to \mu$ in the sense of weak convergence, then obviously $1 = \mu_N(H) \to \mu(H)$, so that any limit has to be a probability measure.

Let $f^{(L)} = (f_1, f_2, \cdots, f_L, 0, 0, \cdots)$, then by Theorem 4.2

$$\chi(f^{(L)}, \mu_N) \to \exp\{-\tfrac{1}{2}\lambda(1-\lambda) \textstyle\sum_{l=1}^{L} |\lambda_l| f_l^2\}, \qquad \text{as} \quad N \to \infty. \tag{4.18}$$

If μ is any limit measure of a suitably chosen subsequence, then by the definition of weak convergence we must have

$$\chi(f, \mu) = \exp\{-\tfrac{1}{2}\lambda(1 - \lambda) \textstyle\sum_{l=1}^{\infty} |\lambda_l| f_l^2\}. \tag{4.19}$$

The right hand side of this last equation characterizes μ uniquely. Hence any convergent subsequence of $\{\mu_N\}$ converges to the same limit, and since $\{\mu_N\}$ is sequentially compact, we have $\mu_N \to \mu$. (4.19) shows that μ is a Gaussian measure, which has mean 0 and S-operator of the desired form.

THEOREM 4.7. *If h is a symmetric trace-class function and if S_N is based on h_N defined by* (4.11), *then $N^{-1}S_N$ converges in distribution to the distribution with characteristic function*

$$\Psi(t) = \textstyle\prod_{k=1}^{\infty} (1 - 2i\lambda(1 - \lambda)\lambda_k t)^{-\frac{1}{2}}. \tag{4.20}$$

PROOF. Because of (4.14) we have

$$\begin{aligned} N^{-1}S_N &= N^{-1} \textstyle\sum_{k=1}^{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_k \varphi_{Nk}(i/N)\varphi_{Nk}(j/N) Z_i Z_j \\ &= \textstyle\sum_{k=1}^{N} \operatorname{sgn}(\lambda_k)(\eta_{Nk})^2. \end{aligned} \tag{4.21}$$

Let $\Psi_N(\cdot)$ be the characteristic function of the distribution of $N^{-1}S_N$. Then

$$\Psi_N(t) = E_{\mu_N} \exp\{it \textstyle\sum_{k=1}^{\infty} \delta_k x_k^2\}, \qquad \text{where} \quad \delta_k = \operatorname{sgn} \lambda_k. \tag{4.22}$$

By the definition of weak convergence

$$\Psi_N(t) = E_{\mu_N} \exp\{it \textstyle\sum_{k=1}^{\infty} \delta_k x_k^2\} \to E_\mu \exp\{it \sum_{k=1}^{\infty} \delta_k x_k^2\}. \tag{4.23}$$

But this limit is continuous in t, and hence $N^{-1}S_N$ converges in distribution. Furthermore

$$\Psi(t) = E_\mu \exp\{it \textstyle\sum_{k=1}^{\infty} \delta_k x_k^2\}. \tag{4.24}$$

We now evaluate $\Psi(t)$. From Theorem 4.6 it follows that for each finite K

$$\begin{aligned} E_{\mu_N} \exp\{it \textstyle\sum_{k=1}^{K} \delta_k x_k^2\} &\to \textstyle\prod_{k=1}^{K} (1 - 2i\lambda(1 - \lambda)\lambda_k t)^{-\frac{1}{2}} \\ &= E_\mu \exp\{it \textstyle\sum_{k=1}^{K} \delta_k x_k^2\}. \end{aligned} \tag{4.25}$$

By the dominated convergence theorem we can pass to the limit in K and get

$$\Psi(t) = E_\mu \exp\{it \textstyle\sum_{k=1}^{\infty} \delta_k x_k^2\} = \prod_{k=1}^{\infty} (1 - 2i\lambda(1 - \lambda)\lambda_k t^{-\frac{1}{2}}.$$

This finishes our proof.

So far we have studied the asymptotic behavior of $N^{-1}S_N$ only for the particular approximating sequence $\{h_N\}$ defined by (4.11). Since this is our standard sequence we will denote it by $\{h_N^s\}$, i.e.,

$$h_N^s(x, y) = \textstyle\sum_{k=1}^{N} \lambda_k \varphi_{Nk}(x)\varphi_{Nk}(y). \tag{4.26}$$

We now find conditions under which the statistic $N^{-1}S_N$, based on an arbitrary approximating sequence $\{h_N\}$, converges to the same law as the particular one based on $\{h_N^s\}$. Let $\{h_N(\cdot, \cdot)\}$ be a sequence of symmetric step functions (i.e.,

h_N is measurable with respect to $\mathfrak{F}_N$), such that

(4.27) $$\int_0^1 \int_0^1 h_N(x, y)\,dx\,dy = 0 \qquad \text{for all} \quad N$$

and define

(4.28) $$g_N(x) = \int_0^1 h_N(x, y)\,dy.$$

The following lemmas are used in the proof of Theorem 4.10.

LEMMA 4.8. *Let the following conditions be satisfied: As* $N \to \infty$

(4.29) $$h_N \to 0 \qquad \textit{in } \text{m.s.},$$

(4.30) $$N\,\|g_N\|^2 = N \int_0^1 g_N(x)^2\,dx \to 0,$$

(4.31) $$\int_0^1 h_N(x, x)\,dx \to 0,$$

(4.32) $$\int_0^1 h_N(x, x)^2\,dx \leqq K^2 \qquad \textit{for some constant} \quad K.$$

Then for S_N *defined by* (2.1) *we have*

(4.33) $$N^{-1}S_N \to 0 \qquad \textit{in } \text{m.s.}$$

PROOF. Define

(4.34) $$h_N'(x, y) = h_N(x, y) - g_N(x) - g_N(y)$$

and let T_N be the statistic based on h_N', i.e.

(4.35) $$T_N = \sum_{i=1}^N \sum_{j=1}^N h_N'(i/N, j/N) Z_i Z_j\,.$$

By direct calculation we get

(4.36) $$N^{-2}E(S_N - T_N)^2 = 4(m/N)^3(N - m)(N - 1)^{-1}N \int_0^1 g_N(x)^2\,dx \to 0,$$

since $N \int_0^1 g_N(x)^2\,dx \to 0$.

By Lemma 3.3 $h_N' \to_{\text{m.s.}} h \equiv 0$. Also

(4.37) $$\left(\int_0^1 h_N'(x, x)^2\,dx\right)^{\frac{1}{2}} \leqq \left(\int_0^1 h_N(x, x)^2\,dx\right)^{\frac{1}{2}} + 2\left(\int_0^1 g_N(x)^2\,dx\right)^{\frac{1}{2}} \leqq K + 1$$

for N large enough. Since h_N' satisfies (3.10) we may apply Lemma 3.4 and obtain

(4.38) $$\begin{aligned} N^{-1}ET_N &= \lambda(1 - \lambda) \int_0^1 h_N'(x, x)\,dx + o_N(1) \\ &= \lambda(1 - \lambda) \int_0^1 h_N(x, x)\,dx + o_N(1) \to 0 \end{aligned}$$

by assumption (4.31), and

(4.39) $$N^{-2}ET_N^2 = 2\lambda^2(1 - \lambda)^2\,\|0\| + o_N(1) \to 0.$$

Thus $N^{-1}T_N \to 0$ in m.s., and by (4.36) we obtain $N^{-1}S_N \to 0$ in m.s.

LEMMA 4.9. *For the particular sequence* $\{h_N^s\}$ *defined by* (4.26) *we have*

(4.40) $$h_N^s \to h \qquad \textit{in } \text{m.s. } \textit{as} \quad N \to \infty.$$

PROOF. For convenience we use the terminology of conditional expectations wrt $\mathfrak{F}_N$ instead of projections. Expectations are with respect to uniform measure on I^2. Set

(4.41) $$h_N(x, y) = \sum_{k=1}^{N} \lambda_k \varphi_k(x) \varphi_k(y)$$

then $h_N^s = E^{\mathfrak{F}_N} h_N$. By Jensen's inequality for conditional expectations,

$$E(h_N^s - E^{\mathfrak{F}_N} h)^2 = E[E^{\mathfrak{F}_N}(h_N - h)]^2 \leqq E[h_N - h]^2 \to 0$$

as $N \to \infty$. Since $E^{\mathfrak{F}_N} h \to h$ in m.s. the lemma follows.

We are now in a position to derive our final result:

THEOREM 4.10. *Let $h(\cdot, \cdot)$ be a symmetric trace-class function. Let $h_N \to h$ in $L_2(I^2)$ such that*

(4.42) $$h_N(x, x) - h_N^s(x, x) \to 0 \quad \text{in} \quad L_2[0, 1], \quad \text{as} \quad N \to \infty$$

(4.43) $$N \int_0^1 g_N(x)^2 \, dx \to 0 \quad \text{as} \quad N \to \infty.$$

Then, if S_N and S_N^s are the statistics corresponding to h_N and h_N^s, respectively, we have

(4.44) $$N^{-1}(S_N - S_N^s) \to 0 \quad \text{in m.s.}$$

PROOF. We apply Lemma 4.8 to $h_N' = h_N - h_N^s$. By Lemma 4.9 $h_N^s \to h$ in $L_2(I)$ hence $h_N' \to 0$ in $L_2(I^2)$. Since

(4.45) $$\int_0^1 h_N^s(x, y) \, dy = \sum_{k=1}^{N} \lambda_k \varphi_{Nk}(x) \int_0^1 \varphi_{Nk}(y) \, dy = 0$$

for all x and all N, we have

(4.46) $$g_N'(x) = \int_0^1 (h_N(x, y) - h_N^s(x, y)) \, dy = g_N(x),$$

and hence

(4.47) $$N \|g_N'\|^2 \to 0 \quad \text{by assumption (4.43).}$$

Finally, (4.31) and (4.32) are satisfied by (4.42).

REMARK 4.11. Condition (4.42) seems to be satisfied in all cases of practical interest (provided that $h(x, x)$ is defined in some natural way, e.g. by continuity), because we usually have $h_N(x, x) \to h(x, x)$ and $h_N^s(x, x) \to h(x, x)$ in $L_2[0, 1]$.

REMARK 4.12. The particular case where $h_N(x, y)$ depends only on the difference $x - y$ has been studied by the author in a previous paper [5]. In this case $h(x, y)$ is always degenerate, and the results are easier to obtain since in (4.2) we can use the Fourier expansion, which has uniformly bounded eigenfunctions.

Acknowledgment. I wish to thank the referee for several helpful suggestions. In particular, the elegant proof of Lemma 4.9 is due to him.

REFERENCES

[1] HÁJEK, J. and SIDÁK, S. (1967). *Theory of Rank Tests.* Academic Press, New York.

[2] MATTHES, T. K. and TRUAX, D. R. (1965). Optimal invariant rank tests for the k-sample problem. *Ann. Math. Statist.* **36** 1207–1222.

[3] PROKHOROV, Y. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theor. Prob. Appl.* **1** 158–214.

[4] SCHACH, S. (1968). On a class of non-parametric two-sample tests for circular distributions. Technical Report No. 9, Stanford Univ.

[5] SCHACH, S. (1969). The asymptotic distribution of a class of non-parametric test statistics. To appear in *Metrika.*

[6] WHEELER, S. and WATSON, G. S. (1964). A distribution-free two-sample test on a circle. *Biometrica* **51** 256–258.