The Annals of Probability
1988, Vol. 16, No. 3, 1291–1298

FUNCTIONAL LIMIT THEOREMS FOR U-PROCESSES[1]

BY DEBORAH NOLAN AND DAVID POLLARD

University of California, Berkeley and Yale University

A U-process is a collection of U-statistics indexed by a family of symmetric kernels. In this paper, two functional limit theorems are obtained for sequences of standardized U-processes. In one case the limit process is Gaussian; in the other, the limit process has finite dimensional distributions of infinite weighted sums of χ^2 random variables. Goodness-of-fit statistics provide examples.

1. Introduction. The U-process is a family of U-statistics treated as a stochastic process. Nolan and Pollard (1987) proved a rate of convergence result by extending comparable results for the empirical process; these extensions were in part suggested by strong law results for U-statistics. In this paper, we again borrow from classical limit theory to prove U-process analogues of the central limit theorem for empirical processes.

Let $\xi_1, \xi_2, \dots$ be independent observations taken from a distribution P on a set X, and $\mathscr{F}$ be a class of real-valued symmetric functions on $X \otimes X$. We define the U-process $\{S_n(f): f \in \mathscr{F}\}$ by

$$S_n(f) = \sum_{1 \le i \ne j \le n} f(\xi_i, \xi_j) \quad \text{for } f \text{ in } \mathscr{F}.$$

With a $[n(n-1)]^{-1}$ standardization, $S_n(f)$ would become a U-statistic in the sense adopted by Serfling (1980, Chapter 5).

Consider the standardized U-process

$$U_n(f) = n^{1/2}\left[\frac{S_n(f)}{n(n-1)} - P \otimes P(f)\right].$$

(We use linear functional notation for expectations.) For fixed f, g in $\mathscr{L}^2(P \otimes P)$, the sequence of random vectors $(U_n(f), U_n(g))$ has an asymptotic bivariate normal distribution with zero mean and covariance function

$$(1) \qquad c(f,g) = P[Pf(x,\cdot)Pg(x,\cdot)] - P \otimes P(f) P \otimes P(g)$$

[Serfling (1980), Section 5.5], and similarly for the higher finite dimensional distributions (fidis). This suggests that U_n, as a stochastic process indexed by $\mathscr{F}$, converges in distribution to a Gaussian process G with zero mean and covariance kernel c. Just as in the empirical central limit theorem [Pollard (1984), Section VII.5] G should have continuous sample paths in some sense, and the fidi convergence must be augmented by an equicontinuity condition to get

Received September 1986.
[1] Supported in part by NSF Grant DMS-85-03347.
AMS 1980 *subject classifications*. Primary 60F17; secondary 62E20.
Key words and phrases. U-statistics, empirical processes, functional limit theorems, equicontinuity, finite dimensional distributions, goodness-of-fit statistics.

convergence in distribution of the random sample paths. These notions will be made more precise in the next section, where a functional central limit theorem will be proved for U_n.

If $\mathscr{F}$ is degenerate, that is, if $Pf(x, \cdot) = 0$ for all $f \in \mathscr{F}$ and $x \in X$, then the covariance kernel $c(f, g)$ is identically 0. In this case the limit process G is degenerate, and another standardization for $S_n(f)$ is required for a nontrivial asymptotic distribution. For each f, the random variables $n^{-1}S_n(f)$ converge in distribution to an infinite weighted sum of independent χ^2 random variables [Serfling (1980), Section 5.5]. This suggests a functional limit theorem for $n^{-1}S_n(\cdot)$ in the degenerate cases. Such a theorem is stated and proved in the next section.

One application of U-process limit theory is found in goodness-of-fit statistics. For example, a Kolmogorov–Smirnov type statistic considered by Silverman (1983),

$$\sup_t \left| [n(n-1)]^{-1} \sum_{i \neq j} \{h(\xi_i, \xi_j) \le t\} - P \otimes P\{h \le t\} \right|, \tag{2}$$

is a functional of a U-process indexed by a class of indicator functions. (The indicator function is identified with its corresponding set.) Our results for U-processes give another representation for the limit distribution found by Silverman for this sequence of statistics.

Our limit theorems differ from those of Hall (1979), Mandelbaum and Taqqu (1984), and Dehling, Denker and Philipp (1984), all of whom treat a partial sum process similar to $\{S_{[nt]}f: 0 \le t \le 1\}$ as a sequence of processes indexed by t for fixed f. Our results are for the stochastic process S_n indexed by $\mathscr{F}$, a general collection of f.

Dehling, Denker and Philipp (1987) have proved an almost sure invariance principle for the U-process in (2), and for higher order U-processes of the same type. They used a bracketing argument driven by moment bounds for degenerate U-statistics. To establish the more delicate functional limit theorem for $n^{-1}S_n$ we need an exponential bound for U-statistics. Unfortunately, we are not able to extend our inequality to higher order U-processes: we cannot extend Lemma 3 of Nolan and Pollard (1987) to symmetric functions of more than two arguments.

2. Main results. Treat the stochastic process U_n as a random element of the space $\mathscr{X}$ of all bounded real-valued functions on $\mathscr{F}$. There are various definitions for convergence in distribution of random elements $\{Z_n\}$ of $\mathscr{X}$ [see Dudley and Philipp (1983), Dudley (1985) or Pollard (1984) for some possibilities]. These approaches differ mainly in their treatment of the measurability difficulties associated with distributions on nonseparable spaces. They all require some sort of convergence of expectations $\mathbb{P}\Phi(Z_n)$ for all bounded real functions Φ on $\mathscr{X}$ that are continuous with respect to the supremum norm topology on $\mathscr{X}$. The limit processes typically have sample paths that concentrate on the set $C(\mathscr{F}, P \otimes P)$ of functions in $\mathscr{X}$ that are uniformly continuous for the $\mathscr{L}^2(P \otimes P)$ seminorm on $\mathscr{F}$. In this paper we omit mention of measurability conditions, with the understanding that some minor qualifications would be needed to make the theorems valid for any particular interpretation of convergence in distribution.

In familiar cases all definitions reduce to the same thing. For example, if $\mathscr{F}$ consists of all indicator functions $f_t(x, y) = \{h(x, y) \le t\}$ for $-\infty < t < \infty$, as in the goodness-of-fit example, then U_n can be identified with a random element of $D[-\infty, \infty]$. The convergence in distribution is then equivalent to the usual notion for $D[-\infty, \infty]$ under its uniform metric.

THEOREM 3. *Let $Z_1, Z_2, \ldots$ be random elements of $\mathscr{X}$. If*

(i) *$\mathscr{F}$ is a totally bounded subset of $\mathscr{L}^2(P \otimes P)$,*

(ii) *the fidis $(Z_n(f_1), \ldots, Z_n(f_k))$ converge in distribution for each $f_1, f_2, \ldots, f_k$,*

(iii) *for each $\varepsilon > 0$ and $\eta > 0$ there exists a $\delta > 0$ for which*

$$\limsup \mathbb{P}\Big\{ \sup_{[\delta]} |Z_n(f) - Z_n(g)| > \eta \Big\} < \varepsilon,$$

where $[\delta] = \{(f, g): P \otimes P(f - g)^2 < \delta^2\}$, then $\{Z_n\}$ converges in distribution to a process with sample paths in $C(\mathscr{F}, P \otimes P)$.

The proof of Theorem 3 is almost the same as the proof of Theorem VII.21 of Pollard (1984), which makes no essential use of the assumption that the random elements are empirical processes. The ideas behind the proof are due chiefly to Dudley; see Dudley (1984), Section 4, for details.

As for empirical processes, simple sufficient conditions for convergence in distribution of $\{U_n\}$ can be stated in terms of *random covering numbers*. Let $\mathscr{F}$ have envelope F, that is,

$$F(\cdot, \cdot) \ge |f(\cdot, \cdot)| \quad \text{if } f \in \mathscr{F};$$

let μ be a measure on $X \otimes X$ for which $\mu F^2 < \infty$. Define the covering number $N(\varepsilon, \mu, \mathscr{F}, F)$ as the smallest cardinality for a subclass $\mathscr{F}^*$ of $\mathscr{F}$ such that

$$\min_{\mathscr{F}^*} \mu|f - f^*|^2 \le \varepsilon^2 \mu F^2 \quad \text{for each } f \text{ in } \mathscr{F}.$$

(We omit the subscript 2 on N, which we added in our earlier paper, because we will use only covering numbers for $\mathscr{L}^2$ norms in this paper.) Define the *covering integral*

$$J(t, \mu, \mathscr{F}, F) = \int_0^t \log N(x, \mu, \mathscr{F}, F)\, dx.$$

Nolan and Pollard (1987) give methods for bounding N and J; we will refer to these methods as needed.

The main tool for verification of the equicontinuity condition, (iii) of Theorem 3, will be a maximal inequality for U-processes. Write T_n for the measure that places mass 1 on each of the pairs (ξ_i, ξ_j) for $1 \le i, j \le 2n$ with the exception of the $4n$ pairs for which $i = j$, $1 \le i \le 2n$; $i = j - n$, $1 \le i \le n$; $i = j + n$, $n + 1 \le i \le 2n$.

LEMMA 4 [Theorem 6, Nolan and Pollard (1987)]. *If $\mathscr{F}$ is a degenerate class with envelope F in $\mathscr{L}^2(P \otimes P)$, then there exists a universal constant C such that*

$$\mathbb{P} \sup_{\mathscr{F}} |S_n(f)| \le C\mathbb{P}\big(\theta_n + \tau_n J(\theta_n/\tau_n, T_n, \mathscr{F}, F)\big),$$

where

$$\tau_n = \left(T_n F^2\right)^{1/2} \quad \text{and} \quad \theta_n = \tfrac{1}{4} \sup_{\mathscr{F}} \left(T_n f^2\right)^{1/2}.$$

The next theorem adapts the usual technique for proving central limit theorems for sequences of U-statistics in order to prove convergence in distribution of $\{U_n\}$ to a Gaussian process. The argument depends on a decomposition of U_n into a sum of an empirical process plus a degenerate U-process. Because the fluctuations of the processes need to be bounded over infinite collections of pairs f-g, the usual $\mathscr{L}^2$ method for disposing of the degenerate part fails. Instead we need the maximal inequality from Lemma 4.

We write P_n for the empirical measure on the sample $(\xi_1, \dots, \xi_n)$ and $P\mathscr{F}$ for the class of functions $Pf(x, \cdot)$ on $\mathscr{X}$. We abbreviate $PF(x, \cdot)$ to PF.

THEOREM 5. *Let $\mathscr{F}$ be a class of functions with envelope F in $\mathscr{L}^2(P \otimes P)$. Suppose*

(i) $\sup_n \mathbb{P} J(1, T_n, \mathscr{F}, F)^2 < \infty$, $\sup_n \mathbb{P} J(1, P_n, P\mathscr{F}, PF)^2 < \infty$ *and* $J(1, P \otimes P, \mathscr{F}, F) < \infty$;

(ii) *for each $\eta > 0$ and $\varepsilon > 0$ there exists a $\gamma > 0$ such that*

$$\limsup \mathbb{P}\{J(\gamma, P_n, P\mathscr{F}, PF) > \eta\} < \varepsilon.$$

Then the $\{U_n\}$ converge weakly to a Gaussian process with 0 *mean, covariance kernel as in* (1) *and sample paths in $C(\mathscr{F}, P \otimes P)$.*

PROOF. Apply Theorem 3 with $Z_n = U_n$. Finiteness of $J(1, P \otimes P, \mathscr{F}, F)$ implies $\mathscr{F}$ is totally bounded. Classical central limit theory for U-statistics gives convergence of the fidis of U_n to the fidis of a Gaussian process with mean zero and covariance kernel as in (1).

For the equicontinuity condition (iii) of Theorem 3, we split $f - P \otimes P(f)$ into a degenerate part plus a function of one variable. Write $\tilde{f}$ for $f - Pf(x, \cdot) - Pf(\cdot, y) + P \otimes P(f)$. Then $U_n(f) = U_n(\tilde{f}) + 2\nu_n \otimes P(f)$ where ν_n is the empirical process $n^{1/2}(P_n - P)$. Standard empirical process methods [Pollard (1984), Lemma VII.15] and (ii), applied to the class $P\mathscr{F}$, establish convergence in distribution of $\nu_n \otimes P$ to the specified Gaussian process. To complete the proof we show that $\mathbb{P} \sup_{\mathscr{F}} |U_n(\tilde{f})| = O(n^{-1/2})$.

Apply Lemma 4 to the degenerate class $\tilde{\mathscr{F}} = \{\tilde{f} : f \in \mathscr{F}\}$ with envelope

$$G(x, y) = F(x, y) + PF(x, \cdot) + PF(\cdot, y) + P \otimes P(F).$$

Abbreviate $J(s, \tilde{T}_n, \tilde{\mathscr{F}}, G)$ to $J_n(s)$. Then

$$\mathbb{P} \sup |U_n(\tilde{f})| \le n^{-1/2} 2C\mathbb{P}[\theta_n + \tau_n J_n(\theta_n/\tau_n)],$$

where $\theta_n = \frac{1}{4}\sup(n^{-2}T_n \tilde{f}^2)^{1/2}$ and $\tau_n = (n^{-2}T_n G^2)^{1/2}$. The inequalities $\mathbb{P}\theta_n \le \frac{1}{4}(\mathbb{P}\tau_n^2)^{1/2} \le (P \otimes P(G^2))^{1/2}$ together with the Cauchy–Schwarz inequality bound the expectation in square brackets by

$$(P \otimes P(G^2))^{1/2}\Big[1 + \big(\mathbb{P}J_n(1)^2\big)^{1/2}\Big].$$

The covering number property (16) of Nolan and Pollard (1987) implies that

$$J_n(1) \le J(1, T_n, \mathscr{F}, F) + 2J(1, P_n, P\mathscr{F}, PF) + J(1, P \otimes P, \mathscr{F}, F),$$

which belongs to $\mathscr{L}^2(\mathbb{P})$ by (i). □

The Kolmogorov–Smirnov type statistic of (2) can be written as

$$\sup_t |[n(n-1)]^{-1} S_n f_t - P \otimes P(f_t)|, \tag{6}$$

where $f_t(x, y) = \{h(x, y) \le t\}$, $-\infty < t < \infty$. With a $n^{1/2}$ standardization we have a straightforward application of Theorem 5. Because the f_t are indicator functions that are increasing in t, the class is easily shown to meet the requirements of Theorem 5. [The class of all such f_t is Euclidean in the sense of Nolan and Pollard (1987).]

Now suppose $\mathscr{F}$ is a collection of degenerate functions. Let $\{w_\alpha\}$ be an orthonormal basis for $\mathscr{L}^2(P)$, and $\{W_\alpha\}$ be a sequence of independent $N(0,1)$ random variables. Write $\langle \cdot, \cdot \rangle$ for the usual inner product on $\mathscr{L}^2(P)$. Each f in $\mathscr{L}^2(P \otimes P)$ defines [Dunford and Schwartz (1963), Section XI.6] a Hilbert–Schmidt operator H_f on $\mathscr{L}^2(P)$ by $(H_f g)(x) = Pf(x, \cdot)g(\cdot)$. Define a process Q on $\mathscr{F}$ by

$$Q(f) = \sum_\alpha \langle H_f w_\alpha, w_\alpha \rangle (W_\alpha^2 - 1) + \sum_{\alpha \ne \beta} \langle H_f w_\alpha, w_\beta \rangle W_\alpha W_\beta.$$

The series can be shown to converge in the $\mathscr{L}^2$ sense for each f [Serfling (1980), page 196]. If the $\{w_\alpha\}$ happen to be the eigenfunctions of the operator H_f the cross-product terms disappear, leaving the infinite series representãtion given by Serfling.

THEOREM 7. *Let $\mathscr{F}$ be a degenerate class of functions with envelope F in $\mathscr{L}^2(P \otimes P)$. Suppose*

(i) $\sup_n \mathbb{P}J(1, T_n, \mathscr{F}, F)^2 < \infty$;

(ii) *for each $\alpha > 0$ and $\varepsilon > 0$ there exists a $\gamma > 0$ such that*

$$\limsup \mathbb{P}\{J(\gamma, T_n, \mathscr{F}, F) > \alpha\} < \varepsilon;$$

(iii) $\log N(\varepsilon, P_n \otimes P, \mathscr{F}, F) = o_p(n)$ *for each $\varepsilon > 0$.*

Then there is a version of Q with sample paths in $C(\mathscr{F}, P \otimes P)$ and $n^{-1}S_n$ converges in distribution to this Q.

PROOF. Apply Theorem 3 with $Z_n = n^{-1}S_n$. A truncation argument similar to Serfling's (1980, page 194) can be used to establish the fidi convergence of $n^{-1}S_n$ to Q.

The uniform boundedness of $\mathbb{P}J(1, T_n, \mathscr{F}, F)^2$ implies $\mathscr{F}$ is totally bounded. Otherwise, for some $\varepsilon > 0$, for any k, we can find $f_1, \dots, f_k$ with

$$P \otimes P|f_i - f_j|^2 > \varepsilon^2 P \otimes P(F^2) \quad \text{for } i \neq j.$$

For each g in $\mathscr{L}^2(P \otimes P)$, the statistic $[n(n-1)]^{-1}T_n g^2$ converges almost surely to $4P \otimes P(g^2)$. Invoke this convergence finitely many times to deduce that, almost surely,

$$T_n|f_i - f_j|^2 > \varepsilon^2 T_n F^2 \quad \text{for } i \neq j,$$

eventually. When these inequalities hold, $N(\frac{1}{2}\varepsilon, T_n, \mathscr{F}, F) \geq k$. This contradicts (i) because, by Fatou's lemma,

$$\liminf \mathbb{P}J(1, T_n, \mathscr{F}, F) \geq \mathbb{P} \liminf \tfrac{1}{2}\varepsilon \log N\left(\tfrac{1}{2}\varepsilon, T_n, \mathscr{F}, F\right) \geq \tfrac{1}{2}\varepsilon \log k.$$

To establish (iii) of Theorem 3, apply the maximal inequality of Lemma 4 to the class of degenerate functions $\mathscr{F}(\delta) = \{f - g : (f, g) \in [\delta]\}$. Abbreviate $J(s, T_n, \mathscr{F}(\delta), 2F)$ to $J_n(s)$:

$$\mathbb{P}\left\{\sup_{[\delta]}|n^{-1}S_n| > \eta\right\} \leq \eta^{-1}2C\mathbb{P}\left[\theta_n + \tau_n J_n(\theta_n/\tau_n)\right], \tag{8}$$

where $\theta_n = \frac{1}{4}\sup_{[\delta]}(n^{-2}T_n(f-g)^2)^{1/2}$ and $\tau_n = (4n^{-2}T_n F^2)^{1/2}$. An application of a uniform law of large numbers will prove that

$$\theta_n \to \tfrac{1}{2}\sup_{[\delta]}\left(P \otimes P(f-g)^2\right)^{1/2} \leq \tfrac{1}{2}\delta \quad \text{almost surely.}$$

Write $\mathscr{F}_1$ for $\{f - g : f \in \mathscr{F}, g \in \mathscr{F}\}$ and $\mathscr{F}_2$ for $\{h^2 : h \in \mathscr{F}_1\}$. Then Theorem 7 of Nolan and Pollard (1987), with $n^{-2}T_n$ substituted for $[n(n-1)]^{-1}S_n$, implies

$$\sup_{\mathscr{F}_2}|n^{-2}T_n h - 4P \otimes P(h)| \to 0 \quad \text{almost surely.}$$

The requirements on $\mathscr{L}^1$ covering numbers for $\mathscr{F}_2$ that are needed for that theorem follow from their $\mathscr{L}^2$ analogues for $\mathscr{F}$. For example, our (i) implies that $\log N(\varepsilon, T_n, \mathscr{F}, F)$, and hence $\log N(\varepsilon, T_n, \mathscr{F}_1, 2F)$, has a uniformly bounded expectation for each fixed $\varepsilon > 0$. The Cauchy–Schwarz inequality converts this to a bound on $\mathscr{L}^1$ covering numbers of $\mathscr{F}_2$: If $h_1, h_2 \in \mathscr{F}_1$, then

$$T_n|h_1^2 - h_2^2| \leq \left(T_n|h_1 - h_2|^2\right)^{1/2}\left(T_n 16F^2\right)^{1/2}.$$

Similar reasoning takes care of the other two conditions of the theorem.

The sequence $\{\tau_n^2\}$ is bounded by a constant multiple of the reverse martingale $\{[n(n-1)]^{-1}S_n F^2\}$, which is necessarily uniformly integrable. Because $\theta_n \leq \tau_n$, it follows that the contribution of $\mathbb{P}\theta_n$ to (8) can be made as small as desired by an appropriate choice of δ. Split the other term into two parts according to whether τ_n exceeds some large constant M or not. Bound the contribution to the right-hand side of (8) by

$$2\eta^{-1}C\Big[\big(\mathbb{P}\tau_n^2\{\tau_n > M\}\mathbb{P}J_n(1)^2\big)^{1/2} + M\mathbb{P}J_n(\theta_n/\tau_n)\Big].$$

If M is large enough, the first term in the square brackets is appropriately small. For fixed $\alpha > 0$ and $\gamma > 0$ the other term is less than

$$2\eta^{-1}CM\big[\mathbb{P}J_n(1)\{\theta_n/\tau_n > \gamma\} + \mathbb{P}J_n(1)\{J_n(\gamma) > \alpha\} + \alpha\big]$$
$$\leq 2\eta^{-1}CM\Big[\big(\mathbb{P}J_n(1)^2\mathbb{P}\{\theta_n/\tau_n > \gamma\}\big)^{1/2} + \big(\mathbb{P}J_n(1)^2\mathbb{P}\{J_n(\gamma) > \alpha\}\big)^{1/2} + \alpha\Big].$$

The ratio θ_n/τ_n is eventually less than $\frac{1}{2}\delta/(P \otimes P(F^2))^{1/2}$. Thus the last term can be made small by choosing α, then γ, then δ, appropriately. □

The study of directional data on circles and spheres provides nontrivial examples of degenerate limit processes. Let $\xi_1, \dots, \xi_n$ be observations taken from the uniform distribution on $(-\pi, \pi)$. Given ξ_i, the difference $\xi_i - \xi_j$ has, modulo 2π, a conditional distribution that is uniform on $(-\pi, \pi)$. Thus the collection $\mathscr{G}$ of centered indicator functions

$$g_t(x, y) = \{\cos(x - y) \leq t\} - P \otimes P\{\cos(x - y) \leq t\} \quad \text{for } -1 \leq t \leq 1$$

is degenerate. Goodness-of-fit statistics can be based on the process $\{n^{-1}S_n(g_t): -1 \leq t \leq 1\}$ [see Silverman (1978), Ripley (1976) and Mardia (1972), Section 7.2]. Theorem 7 gives the asymptotic distribution for the process. Just as for the f_t in (6), the g_t meet the conditions of the theorem. The functions

$$w_\alpha(x) = \sin(\alpha x) + \cos(\alpha x) \quad \text{for } \alpha = 0, \pm 1, \pm 2, \dots$$

form an orthonormal basis for $\mathscr{L}^2(P)$. For each t, the $\{w_\alpha\}$ are also a complete set of eigenfunctions of the Hilbert–Schmidt operator corresponding to g_t, for the eigenvalues $\lambda_0(t) = 0$ and

$$\lambda_\alpha(t) = (-1/\pi\alpha)[\sin(\alpha \arccos t) - 1 + \arccos t] \quad \text{for } \alpha = \pm 1, \pm 2, \dots.$$

Thus the limit process has the simplified representation

$$Q(g_t) = \sum_\alpha \lambda_\alpha(t)(W_\alpha^2 - 1).$$

REFERENCES

DEHLING, H., DENKER, M. and PHILIPP, W. (1984). Invariance principles for von Mises and U-statistics. *Z. Wahrsch. verw. Gebiete* **67** 139–167.

DEHLING, H., DENKER, M. and PHILIPP, W. (1987). The almost sure invariance principle for the empirical process of U-statistic structure. *Ann. Inst. H. Poincaré Probab. Statist.* **23** 121–134.

DUDLEY, R. (1984). A course on empirical processes. *École de Été de Probabilités de Saint Flour XII —1982. Lecture Notes in Math.* **1097** 1–142. Springer, New York.

DUDLEY, R. (1985). An extended Wichura theorem, definitions of Donsker class, and weighted empirical distributions. *Probability in Banach Spaces V. Lecture Notes in Math.* **1153** 141–178. Springer, New York.

DUDLEY, R. and PHILIPP, W. (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Z. Wahrsch. verw. Gebiete* **62** 509–552.

DUNFORD, N. and SCHWARTZ, J. T. (1963). *Linear Operators, Part II*. Wiley, New York.

HALL, P. (1979). On the invariance principle for *U*-statistics. *Stochastic Process. Appl.* **9** 163–174.

MANDELBAUM, A. and TAQQU, M. (1984). Invariance principles for symmetric statistics. *Ann. Statist.* **12** 483–496.

MARDIA, K. V. (1972). *Statistics of Directional Data*. Academic, London.

NOLAN, D. and POLLARD, D. (1987). *U*-processes: Rates of convergence. *Ann. Statist.* **15** 780–799.

POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

RIPLEY, B. D. (1976). The second-order analysis of stationary point processes. *J. Appl. Probab.* **13** 255–266.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

SILVERMAN, B. W. (1978). Distances on circles, toruses and spheres. *J. Appl. Probab.* **15** 136–143.

SILVERMAN, B. W. (1983). Convergence of a class of empirical distribution functions of dependent random variables. *Ann. Probab.* **11** 745–751.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

DEPARTMENT OF STATISTICS
YALE UNIVERSITY
NEW HAVEN, CONNECTICUT 06520