then $a_i(x) > N$, hence by (4.33), $a(x) > N$, and $\phi_i(x) = \phi(x) = 1$ (a.e. $[\nu]$). Thus

$$\begin{aligned}\lim_{i\to\infty} \int (\phi_i(x) - \phi(x))\, d_i(x)\, \nu(dx) \\ = \lim_{i\to\infty} \int_{\{d_i(x)\leqslant 2N\}} (\phi_i(x) - \phi(x))\, d_i(x)\, \nu(dx) \\ = 0, \end{aligned} \tag{4.34}$$

where the limit and integral can be interchanged by the bounded convergence theorem (the integrand is essentially nonnegative by definition of $\phi_i$ and $d_i$ in (4.20)), and the limit of the integrand is zero a.e. $[\nu]$ by (2.9), (4.18), and (4.20). Thus (4.34) verifies (4.19), and the proof of Theorem 2.1 is complete.

## References


[1] Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[2] Brown, L. D., and Marden, J. I. (1989). Complete class results for hypothesis testing problem with simple null hypotheses. *Ann. Statist.* **17**.

[3] Farrell, R. H. (1976). *Techniques of Multivariate Calculation.* Springer-Verlag, New York/Berlin.

[4] Giri, N. (1968). On tests of equality of two covariance matrices. *Ann. Math. Statist.* **39** 275–277.

[5] James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.

[6] Kiefer, J., and Schwartz, R. (1965). Admissible Bayes character of $T^2$, $R^2$, and other fully invariant tests for classical multivariate normal problems. *Ann. Math. Statist.* **36** 747–770.

[7] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.

[8] Sugiura, N. Locally best invariant test for sphericity and the limiting distributions. *Ann. Math. Statist.* **43** 1312–1316.

[9] Wijsman, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. In *Proceedings, Fifth Berkeley Symp. Math. Statist. and Probab.* 1, pp. 389–400.

[10] Marshall, A. W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.


JOURNAL OF MULTIVARIATE ANALYSIS 27, 151–168 (1988)

# Invariance Principles for Changepoint Problems


Miklós Csörgő*

*Carleton University, Ottawa, Canada K1S 5B6*

AND

Lajos Horváth†

*Szeged University, Szeged, Hungary*


*Communicated by the Editors*

We study the asymptotic behaviour of $U$-statistics type processes which can be used for detecting a changepoint of a random sequence. Invariance principles are proved for these processes. © 1988 Academic Press, Inc.

## 1. Introduction

Let $X_1, ..., X_n$ be independent random variables. Suppose we want to test the null hypothesis

$H_0$. $X_i$, $1 \leqslant i \leqslant n$, have the same distribution

versus the alternative hypothesis that there is a changepoint in the sequence $X_1, ..., X_n$, namely that we have

$H_1$. There is a $\lambda \in (0, 1)$ such that $P\{X_1 \leqslant t\} = P\{X_2 \leqslant t\} = \cdots = P\{X_{[n\lambda]} \leqslant t\}$, $P\{X_{[n\lambda]+1} \leqslant t\} = \cdots = P\{X_n \leqslant t\}$, $-\infty < t < \infty$, and $P\{X_{[n\lambda]} \leqslant t_0\} \neq P\{X_{[n\lambda]+1} \leqslant t_0\}$ for some $t_0$.

The changepoint problem has been considerably studied in the literature from the parametric as well as the nonparametric point of view. Non-

Received October 28, 1987.

AMS 1980 subject classifications: Primary 60F17; Secondary 62F03.

Key words and phrases: $U$-statistics, Wiener process, weighted metrics, weak approximations, consistency.

* Research partially supported by a NSERC Canada grant.

† Research done while at Carleton University, supported by NSERC Canada grants of M. Csörgő, D. A. Dawson, and J. N. K. Rao, and by an EMR Canada grant of M. Csörgő.



0047-259X/88 $3.00

parametric results are summarized in Wolfe and Schechtman [15]. Recently Csörgő and Horváth [2] proposed statistics based on processes of linear rank statistics with quantile scores. In this paper we study tests for the changepoint problem which are based on processes of $U$-statistics. They are generalizations of Wilcoxon–Mann–Whitney type statistics.

Let $h(x, y)$ be a symmetric function and consider

$$Z_k = \sum_{1 \leqslant i < k} \sum_{k+1 \leqslant j \leqslant n} h(X_i, X_j), \qquad 1 \leqslant k < n. \tag{1.1}$$

We study $Z_k$ under the null hypothesis in Section 2, and under the alternative hypothesis in Section 3. Typical choices of $h$ are $xy$, $(x-y)^2/2$ (sample variancie), $|x-y|$ (Gini's mean difference), $\operatorname{sign}(x+y)$ (Wilcoxon's one-sample statistic) (cf. Serfling [13]). The case of $h(x, y) = \operatorname{sign}(x-y)$ has gained special attention in the literature. We cannot apply our results directly in this case, because $\operatorname{sign}(x-y)$ is not a symmetric function. However, $\operatorname{sign}(x-y) = -\operatorname{sign}(y-x)$ $(\operatorname{sign}(0)=0)$, i.e., $\operatorname{sign}(x-y)$ is an antisymmetric kernel. We show in Section 4 that our method can be also used in the case of an antisymmetric kernel.

## 2. Asymptotics under $H_0$

In Sections 2 and 3 we assume that $h$ is symmetric, i.e., $h(x, y) = h(y, x)$. Given $H_0$, $X_1, \ldots, X_n$ are i.i.d.r.v.'s. We assume

$$Eh^2(X_1, X_2) < \infty \tag{2.1}$$

and let $Eh(X_1, X_2) = \Theta$, $\tilde{h}(t) = E\{h(X_1, t) - \Theta\}$. Condition (2.1) implies that $E\tilde{h}^2(X_1) < \infty$ and we assume

$$0 < \sigma^2 = E\tilde{h}^2(X_1). \tag{2.2}$$

Here we investigate

$$U_k = Z_k - k(n-k)\,\Theta, \qquad 1 \leqslant k < n,$$

which can be expressed as

$$U_k = U_n^{(3)} - \{U_k^{(1)} + U_k^{(2)}\}, \tag{2.3}$$

where

$$U_k^{(1)} = \sum_{1 \leqslant i < j \leqslant k} h(X_i, X_j) - \binom{k}{2}\Theta,$$

$$U_k^{(2)} = \sum_{k+1 \leqslant i < j \leqslant n} h(X_i, X_j) - \binom{n-k}{2}\Theta,$$



and

$$U_n^{(3)} = \sum_{1 \leqslant i < j \leqslant n} h(X_i, X_j) - \binom{n}{2}\Theta.$$

The latter are nondegenerate $U$-statistics under the conditions (2.1) and (2.2). Thus while $U_k$ itself is not a $U$-statistic, in (2.3) we concluded that it can be expressed as a linear combination of $U$-statistics. Hence the basic idea of studying $U_k$ can be based on the projection of a $U$-statistic on the basic observations (cf. Chap. 5 of Serfling [3]).

In order to state our results we define the Gaussian process $\Gamma$ by

$$\Gamma(t) = (1-t)\,W(t) + t\{W(1) - W(t)\}, \qquad 0 \leqslant t \leqslant 1, \tag{2.4}$$

where $\{W(t), 0 \leqslant t < \infty\}$ is a Wiener process.

THEOREM 2.1. *We assume that $H_0$ holds, and (2.1), (2.2) are satisfied. Then we can define a sequence of Gaussian processes $\{\Gamma_n(t), 0 \leqslant t \leqslant 1\}$ such that, as $n \to \infty$,*

$$\sup_{0 \leqslant t < 1} \left| \frac{n^{-3/2}}{\sigma} U_{[(n+1)t]} - \Gamma_n(t) \right| = o_P(1), \tag{2.5}$$

*where for each $n \geqslant 1$*

$$\{\Gamma_n(t), 0 \leqslant t \leqslant 1\} \overset{\mathcal{D}}{=} \{\Gamma(t), 0 \leqslant t \leqslant 1\}. \tag{2.6}$$

*Proof.* By Theorem 1 of Hall [6] we have

$$\max_{1 \leqslant k \leqslant n} \left| U_k^{(1)} - k \sum_{i=1}^{k} \tilde{h}(X_i) \right| = O_P(n), \tag{2.7}$$

$$\max_{1 \leqslant k \leqslant n} \left| U_k^{(2)} - (n-k) \sum_{i=k+1}^{n} \tilde{h}(X_i) \right| = O_P(n), \tag{2.8}$$

$$\left| U_n^{(3)} - n \sum_{i=1}^{n} \tilde{h}(X_i) \right| = O_P(n). \tag{2.9}$$

Hence

$$\max_{1 \leqslant k \leqslant n} \left| U_k - \left\{ (n-k) \sum_{i=1}^{k} \tilde{h}(X_i) + k \left( \sum_{i=1}^{n} \tilde{h}(X_i) - \sum_{i=1}^{k} \tilde{h}(X_i) \right) \right\} \right| = O_P(n). \tag{2.10}$$

Thus the result follows from Donsker's theorem (cf. Theorem 2.1.2 and Lemma 4.4.4 in Csörgő and Révész [3]).

One can say more about the weak convergence of $U_k$ if the existence of higher moments is assumed.

THEOREM 2.2. *We assume that $H_0$ holds,*

$$E\,|h(X_1, X_2)|^\nu < \infty \qquad \textit{for some} \quad \nu > 2, \tag{2.11}$$

*and* (2.2) *is satisfied. Then we can define a sequence of Gaussian processes $\{\Gamma_n(t), 0 \leqslant t \leqslant 1\}$ such that* (2.5) *holds,*

$$\sup_{1/(n+1) \leqslant t \leqslant n/(n+1)} \left| \frac{n^{-3/2}}{\sigma} U_{[(n+1)t]} - \Gamma_n(t) \right| \Big/ (t(1-t))^{1/2} = O_P(1), \tag{2.12}$$

*and we have* (2.6) *for each $n \geqslant 1$.*

*Proof.* First we note that by (2.11) we have $E\,|\tilde{h}(X_1)|^\nu < \infty$. We introduce

$$S_n^{(1)}(x) = \sigma^{-1} \sum_{1 \leqslant i \leqslant x} \tilde{h}(X_i), \qquad 1 \leqslant x \leqslant [n/2],$$

$$S_n^{(2)}(x) = \sigma^{-1} \sum_{n-x < i \leqslant n} \tilde{h}(X_i), \qquad 1 \leqslant x \leqslant n - [n/2],$$

and show that there exist two independent Wiener processes $\{W_n^{(1)}(x), 0 \leqslant x < \infty\}$ and $\{W_n^{(2)}(x), 0 \leqslant x < \infty\}$ such that

$$\sup_{1 \leqslant x \leqslant [n/2]} x^{-1/2}\,|S_n^{(1)}(x) - W_n^{(1)}(x)| = O_P(1), \tag{2.13}$$

$$\sup_{1 \leqslant x \leqslant n-[n/2]} x^{-1/2}\,|S_n^{(2)}(x) - W_n^{(2)}(x)| = O_P(1). \tag{2.14}$$

Using the Skorohod embedding scheme or the Komlós–Major–Tusnády approximation (cf. Theorem 2.2.4 and Theorem 2.6.3 in Csörgő and Révész [3]), we can define a sequence of Wiener processes $\{W_n^{(1)}(x), 0 \leqslant x < \infty\}$ so that

$$\max_{1 \leqslant k \leqslant [n/2]} k^{-1/2}\,|S_n^{(1)}(k) - W_n^{(1)}(k)| = O_P(1). \tag{2.15}$$

By Theorem 1.2.1 of Csörgő and Révész [3] we obtain

$$\begin{aligned} &\sup_{1 \leqslant x \leqslant [n/2]} x^{-1/2}\,|W_n^{(1)}(x) - W_n^{(1)}([x])| \\ &\quad \leqslant \sup_{1 \leqslant x \leqslant [n/2]} x^{-1/2} \sup_{0 \leqslant s \leqslant 1} |W_n^{(1)}([x]+s) - W_n^{(1)}([x])| = O_P(1). \end{aligned} \tag{2.16}$$

Now (2.15) and (2.16) imply (2.13). The proof of (2.14) is similar. Due to the independence of $S_n^{(1)}(x)$ and $S_n^{(2)}(x)$, the Wiener processes $W_n^{(1)}$ and



$W_n^{(2)}$ can be defined independently. Next we define the Wiener process $\{W_n(x), 0 \leqslant x \leqslant n\}$ by

$$W_n(x) = \begin{cases} W_n^{(1)}(x), & 0 \leqslant x \leqslant [n/2], \\ W_n^{(1)}(n) + W_n^{(2)}(n) - W_n^{(2)}(n-x), & [n/2] < x \leqslant n, \end{cases}$$

and conclude from (2.13) and (2.14) that

$$\begin{aligned} &\sup_{1/(n+1) \leqslant t \leqslant n/(n+1)} \Bigg| \left(1 - \frac{[(n+1)\,t]}{n}\right) \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \\ &\quad + \frac{[(n+1)t]}{n} \left( \sum_{i=1}^{n} \tilde{h}(X_i) - \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \right) \\ &\quad - \sigma\{(1-t)\,W_n((n+1)t) + t(W_n(n+1) - W_n((n+1)t))\} \Bigg| \Big/ (nt(1-t))^{1/2} \\ &= O_P(1). \end{aligned}$$

The latter in turn by (2.10) implies (2.12).

By the construction of the Wiener processes $W_n^{(1)}$ and $W_n^{(2)}$ we obtain

$$\begin{aligned} &\sup_{0 < t < 1} \Bigg| \left(1 - \frac{[(n+1)t]}{n}\right) \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \\ &\quad + \frac{[(n+1)t]}{n} \left( \sum_{i=1}^{n} \tilde{h}(X_i) - \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \right) \\ &\quad - \sigma\{(1-t)\,W_n((n+1)t) + t(W_n(n+1) - W_n((n+1)t))\} \Bigg| = o_P(n^{1/\nu}), \end{aligned}$$

resulting also in (2.5) via (2.10).

Let $Q^*$ be the class of functions $q\colon (0, 1) \to (0, \infty)$ which are monotone nondecreasing near 0 and monotone nonincreasing near one, and $\inf_{\delta \leqslant t \leqslant 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$. If $q \in Q^*$, we define the integral

$$I(q, c) = \int_0^1 (t(1-t))^{-1} \exp(-cq^2(t)/(t(1-t)))\,dt, \qquad c > 0.$$

This integral appears in the characterization of upper class functions of a Wiener process (cf., e.g., Csörgő *et al.* [1]).

COROLLARY 2.1. *We assume that $H_0$ holds, and* (2.2), (2.11) *are satisfied:*

(a) *If $q \in Q^*$, then*

$$\sup_{0<t<1} \left| \frac{n^{-3/2}}{\sigma} U_{[(n+1)t]} - \Gamma_n(t) \right| \Big/ q(t) = o_P(1) \tag{2.17}$$

*if and only if $I(q, c) < \infty$ for all $c > 0$.*

(b) *If $q \in Q^*$, then*

$$\frac{n^{-3/2}}{\sigma} \sup_{0<t<1} |U_{[(n+1)t]}|/q(t) \xrightarrow{D} \sup_{0<t<1} |\Gamma(t)|/q(t) \tag{2.18}$$

*if and only if $I(q, c) < \infty$ for some $c > 0$.*

*Proof.* First we note that $I(q, c) < \infty$ for some $c > 0$ implies (cf. Theorem 3.3 in Csörgő *et al.* [1])

$$\lim_{t \to 0} q(t)/t^{1/2} = \infty. \tag{2.19}$$

We have

$$\sup_{\delta \leqslant t \leqslant 1-\delta} \left| \frac{n^{-3/2}}{\sigma} U_{[(n+1)t]} - \Gamma_n(t) \right| \Big/ q(t) = o_P(1) \tag{2.20}$$

for all $\delta \in (0, \frac{1}{2})$ by Theorem 2.2. Also, by (2.12) and (2.19),

$$\sup_{1/(n+1) \leqslant t \leqslant \delta} \left| \frac{n^{-3/2}}{\sigma} U_{[(n+1)t]} - \Gamma_n(t) \right| \Big/ q(t) = O_P(1) \sup_{0<t\leqslant\delta} \frac{t^{1/2}}{q(t)} \xrightarrow{P} 0 \tag{2.21}$$

as $\delta \to 0$. Next

$$\begin{aligned} \sup_{0<t\leqslant 1/(n+1)} |\Gamma(t)|/q(t) &\leqslant \sup_{0<t\leqslant 1/(n+1)} |W(t)|/q(t) \\ &\quad + \sup_{0<t\leqslant 1/(n+1)} (t/q(t)) \sup_{0\leqslant t\leqslant 1/(n+1)} |W(1) - W(t)| \\ &= o_P(1) \end{aligned}$$

by (2.19) and Theorem 3. of Csörgő *et al.* [1]. One estimates near 1 in a similar way, and the "if" part of (a) is proven.

Assuming now (2.17), we must have

$$\sup_{0<t\leqslant 1/(n+1)} |\Gamma(t)|/q(t) = o_P(1) \tag{2.22}$$



and

$$\sup_{n/(n+1)\leqslant t<1} |\Gamma(t)|/q(t) = o_P(1). \tag{2.23}$$

It is easy to see that (2.22) and (2.33) imply

$$E\Gamma^2(t)/q^2(t) \to 0 \quad \text{as} \quad t \to 0 \quad \text{or} \quad t \to 1. \tag{2.24}$$

Consequently we have (2.22) if and only if

$$\sup_{0<t\leqslant 1/(n+1)} |W(t)|/q(t) = o_P(1). \tag{2.25}$$

Similarly, we have (2.23) if and only if

$$\sup_{n/(n+1)\leqslant t<1} |W(1) - W(t)|/q(t) = o_P(1), \tag{2.26}$$

which is equivalent to

$$\sup_{0<t\leqslant 1/(n+1)} |W(t)|/q(1-t) = o_P(1). \tag{2.27}$$

Now Theorem 3.4 of Csörgő *et al.* [1] combined with (2.25) and (2.27) results in the second part of (a).

As to the proof of (b) we first note that (2.19) implies

$$\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)} \left| \frac{n^{-3/2}}{\sigma} U_{[(n+1)t]} - \Gamma_n(t) \right| \Big/ q(t) = o_P(1). \tag{2.28}$$

Hence it suffices to show that

$$\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)} |\Gamma(t)|/q(t) \xrightarrow{\mathscr{D}} \sup_{0<t<1} |\Gamma(t)|/q(t),$$

which follows immediately from Theorem 3.3 of Csörgő *et al.* [1]. The proof of the necessary part of (b) is similar to that of (a). Only here we have to use Theorem 3.3 of Csörgő *et al.* [1] instead of their Theorem 3.4.

*Remark* 2.1. The proof of the necessary part of Corollary 2.1(a) shows that if we have (2.17) with any sequence of Gaussian processes having the same distribution for each $n \geqslant 1$ as that of $\Gamma$, then $I(q, c)$ must be finite for all $c > 0$. This means that the necessary part does not depend on our construction.

The desirability of having weight functions $q$ around like in Corollary 2.1 is to make our statistical test more sensitive on the tails. A typical choice of

$q$ in (2.18) is the function $(t(1-t)\log\log(1/t(1-t)))^{1/2}$. The variance of $\Gamma(t)$ is $t(1-t)$, hence another choice of a weight function is $(t(1-t))^{1/2}$. However $I((t(1-t))^{1/2}, c)=\infty$ for every $c>0$, and hence we cannot apply Corollary 2.1. This case is studied in the next theorem. Let $a(y\cdot\log n)=(y+2\log\log n+\frac{1}{2}\log\log\log n-\frac{1}{2}\log\pi)(2\log\log n)^{-1/2}$, $-\infty<y<\infty$.

THEOREM 2.3. *We assume that $H_0$ holds, and* (2.2), (2.11) *are satisfied. Then*

$$\lim_{n\to\infty} P\left\{\sigma^{-1}\max_{1\leqslant k\leqslant n}\frac{U_k}{(k(n-k+1)n)^{1/2}}\leqslant a(y,\log n)\right\}=\exp(-\exp(-y)), \tag{2.29}$$

*and*

$$\lim_{n\to\infty} P\left\{\sigma^{-1}\max_{1\leqslant k\leqslant n}\frac{|U_k|}{(k(n-k+1)n)^{1/2}}\leqslant a(y,\log n)\right\}=\exp(-2\exp(-y)). \tag{2.30}$$

We note that it will also follow from the proof of this theorem that the same two limit statements hold for $(n^{-3/2}/\sigma)U_{[(n+1)t]}/(t(1-t))^{1/2}$, $0<t<1$. The proof will be based on the following lemma. Let $b(y,\log n)=(y+2\log\log n+\frac{1}{2}\log\log\log n-\frac{1}{2}\log(4\pi))(2\log\log n)^{-1/2}$, $-\infty<y<\infty$.

LEMMA. *Let $Y_1, Y_2, \ldots$ be i.i.d.r.v.'s with $EY_1=0$, $EY_1^2=1$, and $E|Y_1|^{2+\delta}<\infty$ for some $\delta>0$. Then*

$$\lim_{n\to\infty} P\left\{\max_{1\leqslant k\leqslant n} k^{-1/2}\sum_{i=1}^{k} Y_i\leqslant b(y,\log n)\right\}=\exp(-\exp(-y)) \tag{2.31}$$

*and*

$$\lim_{n\to\infty} P\left\{\max_{1\leqslant k\leqslant n} k^{-1/2}\left|\sum_{i=1}^{k} Y_i\right|\leqslant b(y,\log n)\right\}=\exp(-2\exp(-y)). \tag{2.32}$$

*Also, if $m_n\to\infty$ and $m_n/n\to 0$ $(n\to\infty)$, then*

$$\lim_{n\to\infty} P\left\{\max_{m_n\leqslant k\leqslant n} k^{-1/2}\sum_{i=1}^{k} Y_i\leqslant b(y,\log(n/m_n))\right\}=\exp(-\exp(-y)) \tag{2.33}$$



*and*

$$\lim_{n\to\infty} P\left\{\max_{m_n\leqslant k\leqslant n} k^{-1/2}\left|\sum_{i=1}^{k} Y_i\right|\leqslant b(y,\log(n/m_n))\right\}=\exp(-2\exp(-y)). \tag{2.34}$$

*Proof.* For the proof of (2.31) and (2.32) we refer to Darling and Erdös [4] and Shorack [14].

Of the two statements (2.33) and (2.34) we verify only (2.34). The proof of (2.33) is similar. First let $1\leqslant m_n\leqslant\log n$. Then by (2.32)

$$(2\log\log n)^{1/2}\max_{1\leqslant k\leqslant m_n} k^{-1/2}\left|\sum_{i=1}^{k} Y_i\right|-\log\log n\xrightarrow{P}-\infty,$$

and

$$\lim_{n\to\infty} P\left\{\max_{m_n\leqslant k\leqslant n} k^{-1/2}\left|\sum_{k=1}^{k} Y_i\right|\leqslant b(y,\log n)\right\}=\exp(-2\exp(-y)).$$

Observing now

$$\left|\left(\log\log\frac{n}{m_n}\right)^{1/2}-(\log\log n)^{1/2}\right|(\log\log n)^{1/2}=o(1)$$

and

$$\left|2\log\log\frac{n}{m_n}+\frac{1}{2}\log\log\lg\frac{n}{m_n}-\left(2\log\log n+\frac{1}{2}\log\log\log n\right)\right|=o(1)$$

we get (2.34). Similarly to the proof of Theorem 2.2, there is a Wiener process $W$ such that

$$\sup_{m_n\leqslant x\leqslant n} x^{-1/2}\left|\sum_{1\leqslant i\leqslant x} Y_i-W(x)\right|=o_P(m_n^{1/(2+\delta)-1/2})=o_P((\log n)^{-\delta/(2+\delta)}).$$

Let $\{V(t),\ -\infty<t<\infty\}$ be an Ornstein–Uhlenbeck process. Then we have

$$\sup_{m_n\leqslant x\leqslant n} x^{-1/2}|W(x)=\sup_{(1/2)\log m_n\leqslant t\leqslant(1/2)\log n}|V(t)|=\sup_{0\leqslant t\leqslant(1/2)\log(n/m_n)}|V(t)|,$$

and consequently by Darling and Erdös [4] we obtain (2.34). For the general $m_n$ sequence of the lemma we consider its subsequence with values in $[0,\log n]$ and that with values in $(\log n,\infty)$.

*Proof of Theorem* 2.3. Let $k_n^{(1)} = (\log n)^3$ and $k_n^{(2)} = n/(\log n)^2$, and consider

$$\begin{aligned}\max_{1\leqslant k\leqslant n} \frac{|U_k|}{(k(n-k+1)n)^{1/2}} &= \max_{1\leqslant k\leqslant k_n^{(1)}} \frac{|U_k|}{(k(n-k+1)n^{1/2}} \\ &\quad \vee \max_{k_n^{(1)}\leqslant k\leqslant k_n^{(2)}} \frac{|U_k|}{(k(n-k+1)n)^{1/2}} \\ &\quad \vee \max_{k_n^{(2)}\leqslant k\leqslant n/2} \frac{|U_k|}{(k(n-k+1)n)^{1/2}} \\ &\quad \vee \max_{n/2\leqslant k\leqslant n-k_n^{(2)}} \frac{|U_k|}{(k(n-k+1)n)^{1/2}} \\ &\quad \vee \max_{n-k_n^{(2)}\leqslant k\leqslant n-k_n^{(1)}} \frac{|U_k|}{(k(n-k+1)n)^{1/2}} \\ &\quad \vee \max_{n-k_n^{(1)}\leqslant k\leqslant n} \frac{|U_k|}{(k(n-k+1)n)^{1/2}} \\ &= A_n^{(1)} \vee \cdots \vee A_n^{(6)},\end{aligned}$$

where $a \vee b = \max(a, b)$. It is easy to see that

$$\begin{aligned}A_n^{(1)} &\leqslant \frac{2}{n} \max_{1\leqslant k\leqslant k_n^{(1)}} k^{-1/2} \left| \sum_{1\leqslant i\leqslant k} \sum_{k+1\leqslant j\leqslant n} \{h(X_i, X_j) - \tilde{h}(X_i)\} \right| \\ &\quad + \max_{1\leqslant k\leqslant k_n^{(1)}} k^{-1/2} \left| \sum_{i=1}^{k} \tilde{h}(X_i) \right| \\ &= A_n^{(1,1)} + A_n^{(1,2)}. \end{aligned} \tag{2.36}$$

First we note that by the definition of $\tilde{h}$ we have

$$E\left(n^{-1} \sum_{1\leqslant i\leqslant k} \sum_{k+1\leqslant j\leqslant n} \{h(X_i, X_j) - \tilde{h}(X_i)\}\right)^2 = O(k^2/n),$$

and so

$$\begin{aligned}P\{A_n^{(1,1)} > 1\} &\leqslant \sum_{k=1}^{k_n^{(1)}} P\left\{n^{-1} \left| \sum_{1\leqslant i\leqslant k} \sum_{k+1\leqslant j\leqslant n} \{h(X_i, X_j) - \tilde{h}(X_i)\} \right| > k^{1/2}\right\} \\ &= O(1)\, n^{-1} \sum_{k=1}^{k_n^{(1)}} 1/k = o(1). \end{aligned} \tag{2.37}$$



By Lemma we have

$$A_n^{(1,2)} = O_P((\log\log\log n)^{1/2}),$$

and thus by (2.36) and (2.37) we obtain

$$(2\log\log n)^{1/2} A_n^{(1)} - \sigma \log\log n \xrightarrow{P} -\infty. \tag{2.38}$$

By (2.10) we get

$$\begin{aligned}A_n^{(2)} = \max_{k_n^{(1)}\leqslant k\leqslant k_n^{(2)}} \Bigg| &\frac{n-k}{(n(n-k+1))} k^{-1/2} \sum_{i=1}^{k} \tilde{h}(X_i) \\ &+ \frac{k}{(n(n-k+1)^{1/2}} \sum_{i=k+1}^{n} \tilde{h}(X_i) \Bigg| + O_P(1/\log n). \end{aligned} \tag{2.39}$$

It is easy to verify that

$$\begin{aligned}&\max_{k_n^{(1)}\leqslant k\leqslant k_n^{(2)}} \left(\frac{k}{n}\right)^{1/2} \frac{1}{(n-k+1)^{1/2}} \left| \sum_{i=k+1}^{n} \tilde{h}(X_i) \right| \\ &\overset{\mathscr{D}}{=} \max_{n-k_n^{(2)}\leqslant m\leqslant n-k_n^{(1)}} \left(\frac{n-m}{n}\right)^{1/2} \frac{1}{(m+1)^{1/2}} \left| \sum_{i=1}^{m} \tilde{h}(X_i) \right| \\ &= O(1/\log n) \max_{n-k_n^{(2)}\leqslant m\leqslant n-k_n^{(1)}} \frac{1}{n^{1/2}} \left| \sum_{i=1}^{m} \tilde{h}(X_i) \right| \\ &= O(1/\log n) \max_{1\leqslant m\leqslant n} \frac{1}{n^{1/2}} \left| \sum_{i=1}^{m} \tilde{h}(X_i) \right| \\ &= O_P(1/\log n). \end{aligned} \tag{2.40}$$

Using the lemma we have

$$\begin{aligned}&\max_{k_n^{(1)}\leqslant k\leqslant k_n^{(2)}} \left| \frac{n-k}{(n(n-k+1))^{1/2}} - 1 \right| k^{-1/2} \left| \sum_{i=1}^{k} \tilde{h}(X_i) \right| \\ &\quad = O_P((\log\log n)^{1/2}/(\log n)^2),\end{aligned}$$

and hence (2.39) and (2.40) yield

$$A_n^{(2)} = \max_{k_n^{(1)}\leqslant k\leqslant k_n^{(2)}} k^{-1/2} \left| \sum_{i=1}^{k} \tilde{h}(X_i) \right| + O_P(1/\log n). \tag{2.41}$$

By the lemma again,

$$(2\log\log k_n^{(2)})^{1/2} \max_{1\leqslant k\leqslant k_n^{(1)}} k^{-1/2} \left| \sum_{i=1}^{k} \tilde{h}(X_i) \right| - \sigma \log\log k_n^{(2)} \xrightarrow{P} -\infty,$$

and therefore,

$$\lim_{n\to\infty} P\left\{\frac{1}{\sigma}\max_{k_n^{(1)}\leqslant k\leqslant k_n^{(2)}} k^{-1/2}\left|\sum_{i=1}^{k}\tilde{h}(X_i)\right| \leqslant b(y,\log k_n^{(2)})\right\} = \exp(-2\exp(-y)). \tag{2.42}$$

Observing now that

$$|(2\log\log n)^{1/2}-(2\log\log k_n^{(1)})^{1/2}|\,(\log\log k_n^{(1)})^{1/2}=o(1)$$

and

$$|2\log\log n+\tfrac{1}{2}\log\log\log n-(2\log\log k_n^{(1)} + \tfrac{1}{2}\log\log\log k_n^{(1)})|=o(1),$$

(2.41) and (2.42) imply

$$\lim_{n\to\infty} P\left\{\frac{1}{\sigma}A_n^{(2)}\leqslant b(y,\log n)\right\}=\exp(-2\exp(-y)). \tag{2.43}$$

Towards estimating $A_n^{(3)}$, we first note that

$$\max_{k_n^{(2)}\leqslant k\leqslant n/2}\frac{1}{(n-k)^{1/2}}\left|\sum_{i=k+1}^{n}\tilde{h}(X_i)\right| \overset{\mathscr{D}}{=} \max_{n/2\leqslant m\leqslant n-k_n^{(2)}}\frac{1}{m^{1/2}}\left|\sum_{k=1}^{m}\tilde{h}(X_k)\right|=O_P(1).$$

Hence from (2.10) and (2.34) we obtain

$$\begin{aligned}A_n^{(3)}&\leqslant \max_{k_n^{(2)}\leqslant k\leqslant n/2}\frac{n-k}{(k(n-k+1)\,n)^{1/2}}\left|\sum_{i=1}^{k}\tilde{h}(X_i)\right| \\ &\quad+\max_{k_n^{(2)}\leqslant k\leqslant n/2}\frac{k}{(n(n-k+1))^{1/2}}\left|\sum_{i=k+1}^{n}\tilde{h}(X_i)\right|+O_P(\log n/n^{1/2}) \\ &=O_P((\log\log\log n)^{1/2}).\end{aligned}$$

This in turn implies

$$(2\log\log n)^{1/2}A_n^{(3)}-\sigma^{-1}\log\log n\xrightarrow{P}-\infty. \tag{2.44}$$



The estimation of the r.v.'s $A_n^{(4)}$, $A_n^{(5)}$, and $A_n^{(6)}$ is similar, resulting in the statements

$$(2\log\log n)^{1/2}A_n^{(i)}-\sigma^{-1}\log\log n\xrightarrow{P}-\infty,\qquad i=4,6, \tag{2.45}$$

$$A_n^{(5)}=\max_{n-k_n^{(2)}\leqslant k\leqslant n-k_n^{(1)}}\frac{1}{(n-k)^{1/2}}\left|\sum_{i=k+1}^{n}\tilde{h}(X_i)\right|+O_P(1/\log n), \tag{2.46}$$

and

$$\lim_{n\to\infty} P\left\{\frac{1}{\sigma}A_n^{(5)}\leqslant b(y,\log n)\right\}=\exp(-2\exp(-y)). \tag{2.47}$$

The events in (2.43) and (2.47) are asymptotically independent. Therefore the statement follows from (2.35), (2.38), (2.43), and (2.44)–(2.47).

## 3. Asymptotics under $H_1$

First we introduce some notations. Let

$$\theta=Eh(X_{[n\lambda]-1},X_{[n\lambda]}),\qquad \mu=Eh(X_{[n\lambda]+1},X_{[n\lambda]+2})$$
$$\tau=Eh(X_{[n\lambda]},X_{[n\lambda]+1}),$$

and we write $\log^+ x=\log(x\vee 1)$.

Theorem 3.1. *We assume that $H_1$ holds and*

$$E\,|h(X_{[n\lambda]-1},X_{[n\lambda]})|<\infty,\qquad E\,|h(X_{[n\lambda]+1},X_{[n\lambda]+2})|<\infty,$$
$$E\,|h(X_{[n\lambda]},X_{[n\lambda]+1})|\log^+(|h(X_{[n\lambda]},X_{[n\lambda]+1})|)<\infty \tag{3.1}$$

*are satisfied. Then*

$$\lim_{n\to\infty} Z_{[(n+1)t]}/n^2=\begin{cases}\theta t(\lambda-t)+t\tau(1-\lambda), & 0<t\leqslant\lambda,\\ \mu(t-\lambda)(1-t)+\tau\lambda(1-t), & \lambda\leqslant t<1,\end{cases} \tag{3.2}$$

*in probability.*

*Proof.* Let $1\leqslant[(n+1)t]\leqslant[n\lambda]$. Then

$$\begin{aligned}Z_{[(n+1)t]}&=\sum_{1\leqslant i<j\leqslant[n\lambda]}h(X_i,X_j)+\sum_{1\leqslant i\leqslant[n\lambda]}\ \sum_{[n\lambda]+1\leqslant j\leqslant n}h(X_i,X_j)\\ &\quad-\Bigg\{\sum_{1\leqslant i<j\leqslant[(n+1)t]}h(X_i,X_j)\\ &\quad+\sum_{[(n+1)t]+1\leqslant i<j\leqslant[n\lambda]}h(X_i,X_j)\\ &\quad+\sum_{[(n+1)t]+1\leqslant i\leqslant[n\lambda]}\ \sum_{[n\lambda]+1\leqslant j\leqslant n}h(X_i,X_j)\Bigg\}\\ &=R_n^{(1)}+R_n^{(2)}-\{R_n^{(3)}+\cdots+R_n^{(5)}\}.\end{aligned}$$

By Hoeffding [7] (cf. Theorem A in Section 5.4 of Serfling [13]) we get

$$R_n^{(1)}/n^2 \xrightarrow{\text{a.s.}} \lambda^2\theta/2, \qquad R_n^{(3)}/n^2 \xrightarrow{\text{a.s.}} t^2\theta/2,$$

$$R_n^{(4)}/n^2 \overset{\mathcal{D}}{=} \sum_{1 \leqslant i < j \leqslant [n\lambda]-[(n+1)t]} h(X_i, X_j) \xrightarrow{\text{a.s.}} (t-\lambda)^2\,\theta/2.$$

Now applying Sen [12] and condition (3.1) we obtain

$$R_n^{(2)}/n^2 \xrightarrow{P} \lambda(1-\lambda)\tau, \qquad R_n^{(5)}/n^2 \xrightarrow{P} (\lambda-t)(1-\lambda)\tau.$$

These observations clearly imply the first part of (3.2). The proof of its second part is similar.

*Remark* 3.1. If we assume the existence of the second moments in Theorem 3.1, then we have an a.s. convergence in (3.2) by the moment inequalities of Grams and Serfling [5].

Theorem 3.1 can be used to study the consistency of tests based on the process $\{U_{[(n+1)t]}, 0 \leqslant t < 1\}$. For example, we conclude that rejecting $H_0$ vs $H_1$ when $\sup_{0 \leqslant t < 1} (n^{-3/2}/\sigma)\,|U_{[(n+1)t]}|$ is large, then the latter test is consistent except in the case of $\tau = \theta = \mu = 0$. The same can be said about the weighted versions of this test.

## 4. Antisymmetric Kernel

In this section we assume that $h$ is an antisymmetric kernel, i.e.,

$$h(x, y) = -h(y, x). \tag{4.1}$$

In this case $Eh(X_1, X_2) = 0$ and similarly to the symmetric case we let $\tilde{h}(t) = Eh(t, X_1)$. We assume

$$Eh^2(X_1, X_2) < \infty \quad \text{and} \quad 0 < \sigma^2 = E\tilde{h}^2(X_1). \tag{4.2}$$

Accordingly to Section 2 we now have $U_k = Z_k$, where $Z_k$ is defined by (1.1). It is easy to see that (2.3) remains true in the case of an antisymmetric kernel, with $\Theta$ taken to be zero, of course.

First we give an analog of Theorem 2.1.

Theorem 4.1. *We assume that $H_0$ holds, and* (4.1) *and* (4.2) *are satisfied.*



*Then we can define a sequence of Brownian bridges $\{B_n(t), 0 \leqslant t \leqslant 1\}$ such that, as $n \to \infty$,*

$$\sup_{0 \leqslant t \leqslant 1} \left| \frac{n^{-3/2}}{\sigma} U_{[(n+1)t]} - B_n(t) \right| = o_P(1) \tag{4.3}$$

*and for each $n \geqslant 0$, $EB_n(t) = 0$, $EB_n(t)\,B_n(s) = \min(t, s) - ts$.*

*Proof.* The proof is similar to that of Theorem 2.1. Instead of Theorem 1 of Hall [6] we use Theorem 2.1 of Janson and Wichura (1983), which gives

$$\max_{1 \leqslant k \leqslant n} \left| U_k^{(1)} - \sum_{i=1}^{k} (k-2i+1)\,\tilde{h}(X_i) \right| = O_P(n), \tag{4.4}$$

$$\max_{1 \leqslant k \leqslant n} \left| U_k^{(2)} - \sum_{i=k+1}^{n} (n+k-2i+1)\,\tilde{h}(X_i) \right| = O_P(n), \tag{4.5}$$

and

$$\left| U_n^{(3)} - \sum_{i=1}^{n} (n-2i+1)\,\tilde{h}(X_i) \right| = O_P(n). \tag{4.6}$$

By (4.4), (4.5), and (4.6) we have

$$\max_{1 \leqslant k \leqslant n} \left| U_k - \left\{ n \sum_{i=1}^{k} \tilde{h}(X_i) - k \sum_{i=1}^{n} \tilde{h}(X_i) \right\} \right| = O_P(n) \tag{4.7}$$

and hence Donsker's theorem implies Theorem 4.1.

Surprisingly, the limiting processes are different in Theorems 2.1 and 4.1. In the special case of $h(x, y) = \operatorname{sign}(x-y)$ (cumulative rank tests) Pettitt [9] (cf. also Pettitt [10]) indicate a proof of Theorem 4.1.

The following Theorem is an analog of Theorem 2.2.

Theorem 4.2. *We assume that $H_0$ holds,* (4.1) *and* (4.2) *are satisfied, and*

$$E|h(X_1, X_2)|^{\nu} < \infty \quad \textit{for some} \quad \nu > 2. \tag{4.8}$$

*Then we can define a sequence of Brownian bridges $\{B_n(t), 0 \leqslant t \leqslant 1\}$ such that* (4.3) *holds and*

$$\sup_{1/(n+1) \leqslant t \leqslant n/(n+1)} \left| \frac{n^{-3/2}}{\sigma} U_{[(n+1)t]} - B_n(t) \right| \Big/ (t(1-t))^{1/2} = O_P(1). \tag{4.9}$$

*Proof.* Using (4.4)–(4.6) with the Skorohod embedding scheme (or with the Komlós–Major–Tusnády approximation), the proof goes along the lines of the proof of Theorem 2.2.

The next results are direct consequences of Theorem 4.2. One can give detailed proofs using the methods of the proofs of Corollary 2.1 and Theorem 2.3. Let $\{B(t), 0 \leqslant t \leqslant 1\}$ be a Brownian bridge.

COROLLARY 4.1. *We assume that $H_0$ holds and* (4.1), (4.2), *and* (4.8) *are satisfied.*

(a) *If $q \in Q^*$, then*

$$\sup_{0<t<1} \left| \frac{n^{-3/2}}{\sigma} U_{[(n+1)t]} - B_n(t) \right| \Big/ q(t) = o_P(1)$$

*if and only if $I(q, c) < \infty$ for all $c > 0$.*

(b) *If $q \in Q^*$, then*

$$\frac{n^{-3/2}}{\sigma} \sup_{0<t<1} |U_{[(n+1)t]}|/q(t) \xrightarrow{\mathscr{D}} \sup_{0<t<1} |B(t)|/q(t)$$

*if and only if $I(q, c) < \infty$ for some $c > 0$.*

THEOREM 4.3. *We assume that $H_0$ holds and* (4.1), (4.2), *and* (4.8) *are satisfied. Then*

$$\lim_{n\to\infty} P\left\{\sigma^{-1} \max_{1\leqslant k\leqslant n} \frac{U_k}{(k(n-k+1)n)^{1/2}} \leqslant a(y, \log n)\right\} = \exp(-\exp(-y))$$

*and*

$$\lim_{n\to\infty} P\left\{\sigma^{-1} \max_{1\leqslant k\leqslant n} \frac{|U_k|}{(k(n-k+1)n)^{1/2}} \leqslant a(y, \log n)\right\} = \exp(-2\exp(-y)).$$

Now we assume that $X_1, ..., X_n$ have a continuous distribution function, and study the case of $h(x, y) = \operatorname{sign}(x-y)$. Under $H_0$, $E\operatorname{sign}(X_1 - X_2) = 0$ and $\sigma^2 = 1/12$. Then

$$U_k = Z_k = \sum_{1\leqslant i<k} \sum_{k+1\leqslant j\leqslant n} \operatorname{sign}(X_i - X_j)$$

is distribution free, and the results of the present section are applicable. By Theorem 4.1, $(12)^{1/2} n^{-3/2} U_{[(n+1)t]}$ converges weakly to a Brownian bridge in $D[0, 1]$. This result was obtained by Pettitt [9] using heuristic arguments.



Sen and Srivastava [11] also mention (without developing any properties) non-parametric tests as analogs to some parametric likelihood ratio procedures. In particular, they suggest rejecting $H_0$ for large values of

$$D_n = (12)^{1/2} \max_{1\leqslant k\leqslant n} |U_k|/(k(n-k+1)n)^{1/2}.$$

It follows from Theorem 4.3 that $D_n \to^P \infty$ even under $H_0$. This is the reason for them finding $D_n$ being superior to other statistics. We can, of course, use $D_n$ for testing $H_0$ with normalizing factors as given in Theorem 4.3. Naturally then further power studies are also needed in order to conclude any superiority properties.

## ACKNOWLEDGMENTS

We are very grateful to F. Lombard for his helpful comments and for pointing out an oversight in the first version of this paper. We also wish to thank Vera R. Huse for calling our attention to a slip in this version and for her careful reading of our manuscript.

## REFERENCES

[1] CSÖRGŐ, M., CSÖRGŐ, S., HORVÁTH, L., AND MASON, D. M. (1986). Weighted empirical and quantile processes. *Ann. Probab.* **14** 31–85.

[2] CSÖRGŐ, M. AND HORVÁTH L. (1987). Nonparametric tests for the changepoint problem. *J. Statist. Plann. Inference* **17** 1–9.

[3] CSÖRGŐ, M. AND RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York.

[4] DARLING, D. AND ERDŐS, P. (1956). A limit theorem for the maximum of normalized sums of independent random variables. *Duke Math. J.* **23** 143–155.

[5] GRAMS, W. F. AND SERFLING, R. J. (1973). Convergence rates for *U*-statistics and related statistics. *Ann. Statist.* **1** 153–160.

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

[7] HOEFFDING, W. (1961). *The Strong Law of Large Numbers for U-Statistics.* University of North Carolina Institute of Statistics Mimeo Series 302.

[8] JANSON, S. AND WICHURA, M. J. (1983). Invariance principles for stochastic area and related stochastic integrals. *Stochastic Process. Appl.* **16** 71–84.

[9] PETTITT, A. N. (1979). A non-parametric approach to the change-point problem. *Appl. Statist.* **28** 126–135.

[10] PETTITT, A. N. (1980). Some results on estimating a change-point using non-parametric type statistics. *J. Statist. Comput. Simul.* **11** 261–272.

[11] SEN, A. AND SRIVASTAVA M. S. (1975). On tests for detecting changes in mean. *Ann. Statist.* **3** 98–108.

[12] SEN, P. K. (1977). Almost sure convergence of generalized *U*-statistics. *Ann. Probab.* **5** 287–290.

[13] SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.
[14] SHORACK, G. R. (1979). Extension of the Darling and Erdős theorem on the maximum of normalized sums. *Ann. Probab.* 7 1092–1096.
[15] WOLFE, D. A. AND SCHECHTMAN, E. (1984). Nonparametric statistical procedures for the changepoint problem. *J. Statist. Plann. Inference* 9 389–396.

JOURNAL OF MULTIVARIATE ANALYSIS 27, 169–180 (1988)

# On the Area of the Circles Covered by a Random Walk

P. ERDÖS

*Mathematical Institute, Budapest, Hungary*

AND

P. RÉVÉSZ

*Mathematical Institute, Budapest, Hungary and Technical University, Vienna, Austria*

*Communicated by the Editors*

The area of the largest circle around the origin completely covered by a simple symmetric plane random walk is investigated. © 1988 Academic Press, Inc.

## 1. INTRODUCTION

Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random vectors taking values from $R^2$ with distribution

$$\mathbb{P}\{X_1=(0,1)\}=\mathbb{P}\{X_1=(0,-1)\}=\mathbb{P}\{X_1=(1,0)\}=\mathbb{P}\{X_1=(-1,0)\}=\tfrac{1}{4}$$

and let

$$S_0=0=(0,0) \qquad \text{and} \qquad S(n)=S_n=X_1+X_2+\cdots+X_n \qquad (n=1,2,\ldots),$$

i.e., $\{S_n\}$ is the simple symmetric random walk on the plane. Further let

$$\xi(x,n)=\#\{k: 0<k\leqslant n,\ S_k=x\}$$

$(n=1,2,\ldots;\ x=(i,j);\ i,j=0,\ \pm1,\ \pm2,\ldots)$ be the local time of the random walk. We say that the circle

$$Q(N)=\{x=(i,j): \|x\|=(i^2+j^2)^{1/2}\leqslant N\}$$

Received January 27, 1988.
AMS subject classifications: Primary 60J15; Secondary 60F15, 60G17.
Key words and phrases: random walk on the plane, strong laws, covered circle.



0047-259X/88 $3.00