The Annals of Statistics
1996, Vol. 24, No. 5, 2266–2279

LIMIT THEOREM FOR MAXIMUM OF STANDARDIZED *U*-STATISTICS WITH AN APPLICATION

BY LAJOS HORVÁTH[1] AND QI-MAN SHAO[2]

University of Utah and University of Oregon

We show that the maximally selected standardized U-statistic goes in distribution to an infinite sum of weighted chi-square random variables in the degenerate case. The result is applied to the detection of possible changes in the distribution of a sequence observation.

1. Introduction and results. Let $X_1, X_2, \dots, X_n$ be independent random variables. We want to test the null hypothesis H_0: $X_1, X_2, \dots, X_n$ are identically distributed against the alternative hypothesis that there is a change-point in the sequence $X_1, X_2, \dots, X_n$. Namely, H_A: there is an integer k^*, $1 \le k^* < n$, such that

$$P\{X_1 \le t\} = P\{X_2 \le t\} = \cdots = P\{X_{k^*} \le t\},$$

$$P\{X_{k^*+1} \le t\} = P\{X_{k^*+2} \le t\} = \cdots = P\{X_n \le t\} \quad \text{for all } t$$

and

$$P\{X_{k^*} \le t_0\} \ne P\{X_{k^*+1} \le t_0\} \quad \text{for some } t_0.$$

The change-point problem has been studied extensively in the literature. For a survey we refer to Brodsky and Darkhovsky (1993). Wolfe and Schechtman (1984) and Csörgő and Horváth (1987) suggested several tests based on the linear rank statistics with quantile scores and U-statistics. Csörgő and Horváth (1988) used U-statistics which are generalizations of Wilcoxon–Mann–Whitney-type statistics to detect a possible change-point. For surveys on U-statistics we refer to Serfling (1980), Lee (1990) and Koroljuk and Borovskich (1994).

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

$$U_{k,n} = \sum_{1 \le i \le k} \sum_{k < j \le n} h(X_i, X_j) - k(n-k)\theta \tag{1.1}$$

and

$$T_n = \max_{1 \le k < n} |U_{k,n}|/(\operatorname{Var} U_{k,n})^{1/2},$$

Received November 1995; revised January 1996.
[1]Work partially supported by a NSF-NATO East Europe Outreach Grant and a NSERC, Canada International Scientific Exchange Award.
[2]Work partially supported by a National University of Singapore Research Project.
AMS 1991 *subject classifications*. Primary 62G20; secondary 62G10, 60F05.
Key words and phrases. Degenerate U-statistic, limit distribution, standardized statistics, change-point analysis.

where $\theta = E_{H_0} h(X_1, X_2)$. For each k, $U_{k,n}$ compares the first k observations to the last $n-k$ using the kernel h, and T_n selects the maximum of the standardized U-statistics. According to Csörgő and Horváth (1988), we should reject H_0 for large values of T_n. The limit distribution of T_n under H_0 in case of nondegenerate kernels was obtained by Csörgő and Horváth (1988).

THEOREM A [Csörgő and Horváth (1988)]. *Let $\tilde{h}(t) = E\{h(X_1, t) - \theta\}$. Assume that H_0 holds, $E|h(X_1, X_2)|^\nu < \infty$ with some $\nu > 2$ and $\tau^2 = E\tilde{h}^2(X_1) > 0$. Then*

$$\lim_{n\to\infty} P\left\{a(\log n)\frac{1}{\tau}\max_{1\le k<n}\frac{|U_{k,n}|}{(k(n-k))^{1/2}} \le t + b(\log n)\right\} = \exp(-2e^{-t}) \tag{1.2}$$

for all t, where $a(x) = (2\log x)^{1/2}$ and $b(x) = 2\log x + \frac{1}{2}\log\log x - \frac{1}{2}\log\pi$.

Further results on the applications of U-statistics to change-point analysis can be found in Ferger and Stute (1992) and Ferger (1994a–c).

The main aim of this note is to give the limit distribution for maximally selected standardized U-statistics in the degenerate case, that is, $\tau = 0$. Of course, $\tau = 0$ means that the projections $\tilde{h}(X_i)$, $1 \le i \le n$, are zero with probability 1. If $\tau = 0$, then there are orthogonal eigenfunctions $\{\varphi_j(t), 1 \le j < \infty\}$ and eigenvalues $\{\lambda_j, 1 \le j < \infty\}$ such that [see, e.g., Serfling (1980)]

$$\lim_{K\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(h(x,y) - \theta - \sum_{1\le j\le K}\lambda_j\varphi_j(x)\varphi_j(y)\right)^2 dF(x)\,dF(y) = 0 \tag{1.3}$$

and

$$E\varphi_j(X_1)\varphi_k(X_1) = \begin{cases} 1, & \text{if } j = k,\\ 0, & \text{if } j \ne k,\end{cases} \tag{1.4}$$

where F denotes the common distribution of X_i under H_0. It follows from (1.3) and (1.4) that

$$E(h(X_1, X_2) - \theta)^2 = \sum_{1\le k<\infty}\lambda_k^2. \tag{1.5}$$

Let $\{N_i,\ 1 \le i < \infty\}$ be a sequence of independent, standard normal random variables and define

$$\xi = \left(\sum_{1\le i<\infty}\lambda_i^2 N_i^2\right)^{1/2}. \tag{1.6}$$

Now we can state the main result.

THEOREM 1.1. *We assume that H_0 holds,*

$$E\tilde{h}^2(X_1) = 0 \quad \textit{and} \quad 0 < \sigma^2 = E(h(X_1, X_2) - \theta)^2 < \infty. \tag{1.7}$$

Then, as $n \to \infty$,

$$(2\log\log n)^{-1/2} \max_{1\le k<n} \frac{|U_{k,n}|}{(k(n-k))^{1/2}} \to_{\mathscr{D}} \xi. \tag{1.8}$$

It follows from (1.5) and (1.7) that ξ is finite with probability 1. Also, (1.8) can be rewritten as

$$(2\log\log n)^{-1/2} \max_{1\le k<n} \frac{|U_{k,n}|}{(\operatorname{Var} U_{k,n})^{1/2}} \to_{\mathscr{D}} \frac{\xi}{\sigma}.$$

It is interesting to note that we get completely different limit theorems in (1.2) and (1.8). The limit distribution in (1.2) is an extreme value distribution, while ξ^2 in (1.8) is a weighted sum of χ^2 random variables. We also note that ξ^2 is related to the usual limit of degenerate U-statistics [cf. Serfling (1980), Lee (1990) and Koroljuk and Borovskich (1994)].

The proof of Theorem 1.1 is presented in Section 3. An application of Theorem 1.1 to the intervals between coal-mining disasters is discussed in the next section.

2. An application. The time intervals between successive coal-mining disasters involving 10 or more men killed in British coal mines between 1875 and 1950 were analyzed by Maguire, Pearson and Wynn (1952). Later, Jarrett (1979) corrected several errors in the data given by Maguire, Pearson and Wynn (1952) and extended the data set to cover 191 disasters between 1851 and 1962. Jarrett (1979) concluded that the data had an exponential distribution; at the beginning of the observations the mean was 106 and it might have changed over time. Let $X_1, X_2, \dots, X_{190}$ denote the observations in Table 1 of Jarrett (1979). Now we apply Theorem 1.1 to test the null hypothesis H_0: $X_1, X_2, \dots, X_{190}$ are exponentially distributed with mean 106 against the alternative hypothesis that there is a change-point in the sequence. We use two different kernels for the test.

KERNEL 1. Let $h_1(x, y) = (x-106)(y-106)/(106^2)$. Under H_0, by Theorem 1.1 we have

$$\begin{aligned} T_{n,1} &:= (2\log\log n)^{-1/2} \max_{1\le k<n} \left| \sum_{1\le i\le k} \sum_{k<j\le n} h_1(X_i, X_j) \right| \Big/ (k(n-k))^{1/2} \\ &\to_{\mathscr{D}} |N|, \end{aligned} \tag{2.1}$$

where N is the standard normal random variable. A direct calculation shows that the value of $T_{n,1}$ for the coal-mine disasters is 137.627, so we reject H_0.

KERNEL 2. Let

$$h_2(x, y) = \int_{-\infty}^{\infty} (I\{x\le t\} - F(t))(I\{y \le t\} - F(t))\, dF(t),$$

where $F(t)$ stands for the common distribution function under H_0. It is easy to show that for $x \geq y$,

$$h_2(x, y) = \tfrac{1}{3}(F^3(x) + (1 - F(x))^3) - \tfrac{1}{2}(F^2(x) - F^2(y)).$$

The U-statistic generated by $h(x, y)$ above is related to the Cramér–von Mises statistic [cf. Lee (1990), page 190] and is distribution-free under the no-change null hypothesis. Moreover, $\lambda_i = (i\pi)^{-2}$ for $i = 1, 2, \dots$. Therefore, by Theorem 1.1 we have

$$\begin{aligned} T_{n,2} &:= (2\log\log n)^{-1/2} \max_{1\le k<n} \left| \sum_{1\le i\le k} \sum_{k<j\le n} h_2(X_i, X_j) \right| \Big/ (k(n-k))^{1/2} \\ &\to_{\mathscr{D}} \xi, \end{aligned} \tag{2.2}$$

where

$$\xi = \left(\sum_{1\le i<\infty} (i\pi)^{-4} N_i^2 \right)^{1/2}. \tag{2.3}$$

By applying Chebyshev's inequality to the moment generating function, one can easily verify that

$$P\left\{ \sum_{i=1}^{\infty} a_i N_i^2 \geq \sum_{i=1}^{\infty} \frac{a_i}{1 - ta_i} \right\} \leq \exp\left(-\frac{1}{2} \sum_{i=1}^{\infty} \left\{ \frac{ta_i}{1 - ta_i} - \log(1 - ta_i) \right\} \right) \tag{2.4}$$

for all $a_i \geq 0$ and $0 < t < \inf_{1\le i<\infty} 1/a_i$.

Applying (2.4) to ξ with $a_i = (i\pi)^{-4}$ and $t = 0.92\pi^4$, we get

$$P\{\xi \geq 0.3595\} \leq 0.001,$$

where ξ is given by (2.3). In the case of this data set $F(t) = 1 - \exp(-t/106)$ for $t \geq 0$ and the value of $T_{n,2}$ is 0.9075. Thus, we reject the null hypothesis at the 0.001 level of significance.

For further analysis of the coal-mine disasters, we refer to Cox and Lewis (1966), Worsley (1986) and Gombay and Horváth (1990).

3. Proof of Theorem 1.1. Throughout this section, we assume that without loss of generality, $\theta = 0$ and that the conditions of Theorem 1.1 are satisfied. The proof of Theorem 1.1 is based on the following lemmas.

LEMMA 3.1. *We have*

$$E \max_{1\le k<n} U_{k,n}^2 \leq 36 n^2 \sigma^2.$$

PROOF. Observing that

$$U_{k,n} = \sum_{1\le i<j\le n} h(X_i, X_j) - \sum_{1\le i<j\le k} h(X_i, X_j) - \sum_{k<i<j\le n} h(X_i, X_j),$$

we get

$$\max_{1\le k<n} U_{k,n}^2 \le \tfrac{9}{2}\max_{2\le k\le n}\biggl(\sum_{1\le i<j\le k} h(X_i, X_j)\biggr)^2 + \tfrac{9}{2}\max_{1\le k<n-1}\biggl(\sum_{k<i<j\le n} h(X_i, X_j)\biggr)^2.$$

It follows from (1.7) that $\{\sum_{1\le i<j\le k} h(X_i, X_j),\ \sigma(X_1, \dots, X_k), 2 \le k \le n\}$ is a martingale. Hence, by Doob's inequality [cf. Chow and Teicher (1988), page 247],

$$E\max_{2\le k\le n}\biggl(\sum_{1\le i<j\le k} h(X_i, X_j)\biggr)^2 \le 4E\biggl(\sum_{1\le i<j\le n} h(X_i, X_j)\biggr)^2.$$

Similar arguments give

$$E\max_{1\le k<n-1}\biggl(\sum_{k<i<j\le n} h(X_i, X_j)\biggr)^2 \le 4E\biggl(\sum_{1\le i<j\le n} h(X_i, X_j)\biggr)^2.$$

Hence,

$$E\max_{1\le k<n} U_{k,n}^2 \le 36E\biggl(\sum_{1\le i<j\le n} h(X_i, X_j)\biggr)^2 \le 36n^2\sigma^2,$$

as desired. □

LEMMA 3.2. *For any $x > 0$ we have*

$$P\biggl\{\max_{1\le k<n}\sum_{1\le i<\infty}\lambda_i^2\biggl(\sum_{k<j\le n}\varphi_i(X_j)\biggr)^2 \ge nx\biggr\} \le \frac{3\sigma^2}{x}.$$

PROOF. The lemma is obviously true if $x \le 3\sigma^2$. So, we assume $x > 3\sigma^2$. Let

$$Q_k = \sum_{1\le i<\infty}\lambda_i^2\biggl(\sum_{n-k<j\le n}\varphi_i(X_j)\biggr)^2 - k\sigma^2.$$

It is easy to see that $\{Q_k, \sigma(X_n, \dots, X_{n-k+1}),\ 1 \le k \le n\}$ is a martingale. Thus, by the martingale maximum inequality [cf. Chow and Teicher (1988), page 247],

$$\begin{aligned}
P\biggl\{\max_{1\le k<n}\sum_{1\le i<\infty}\lambda_i^2\biggl(\sum_{k<j\le n}\varphi_i(X_j)\biggr)^2 \ge nx\biggr\} &\le P\Bigl\{\max\nolimits_{1\le k\le n} Q_k \ge nx - n\sigma^2\Bigr\}\\
&\le P\biggl\{\max_{1\le k\le n} Q_k \ge \frac{2nx}{3}\biggr\}\\
&\le \frac{3E|Q_n|}{2nx} \le \frac{3\sigma^2}{x}.
\end{aligned}$$

□

LEMMA 3.3. *For any $x \ge 1$ we have*

$$P\left\{\max_{1\le k<n} \frac{|U_{k,n}|}{(k(n-k))^{1/2}} \ge x\right\} \le 10^5\sigma + 32\exp\left(-\frac{x^2}{512\sigma^{1/2}}\right)\log n.$$

PROOF. We assume that $0 < \sigma \le 10^{-5}$, since otherwise Lemma 3.3 is trivial. Let $m_1 = [\log_2(n/2)]$, where $[\cdot]$ denotes the integer part of the number. Clearly,

$$\begin{aligned}
&P\left\{\max_{1\le k\le n/2} |U_{k,n}|/(k(n-k))^{1/2} \ge x\right\}\\
&\quad\le P\left\{\max_{1\le k\le n/2} |U_{k,n}|/k^{1/2} \ge x(n/2)^{1/2}\right\}\\
&\quad\le P\left\{\max_{1\le \ell\le m_1}\max_{2^{\ell-1}\le k<2^\ell} |U_{k,n}|/2^{(\ell-1)/2} \ge x(n/2)^{1/2}\right\}\\
&\qquad + P\left\{\max_{2^{m_1}\le k\le n/2} |U_{k,n}| \ge x2^{m_1/2}(n/2)^{1/2}\right\}\\
&\quad\le P\left\{\max_{1\le \ell\le m_1}\max_{2^{\ell-1}\le k<2^\ell} |U_{k,n}|/2^{\ell/2} \ge xn^{1/2}/2\right\}\\
&\qquad + P\left\{\max_{1\le k\le n/2} |U_{k,n}| \ge nx/4\right\}.
\end{aligned}\tag{3.1}$$

Using Lemma 3.1, we get

$$P\left\{\max_{1\le k\le n/2} |U_{k,n}| \ge \frac{nx}{4}\right\} \le \frac{36n^2\sigma^2}{(nx/4)^2} = \frac{3^2 2^6\sigma^2}{x^2}.\tag{3.2}$$

For $2^{\ell-1} \le k < 2^\ell$, write

$$U_{k,n} = U_{k,2^\ell} + V_{k,\ell},$$

where $V_{k,\ell} = \sum_{1\le i\le k}\sum_{2^\ell<j\le n} h(X_i, X_j)$. Let

$$I_n = P\left\{\max_{1\le \ell\le m_1}\max_{1\le k<2^\ell} |V_{k,\ell}|/2^{\ell/2} \ge xn^{1/2}/2\right\}.$$

Applying Lemma 3.1 again, we obtain

$$\begin{aligned}
&P\left\{\max_{1\le \ell\le m_1}\max_{2^{\ell-1}\le k<2^\ell} \frac{|U_{k,n}|}{2^{\ell/2}} \ge \frac{xn^{1/2}}{2}\right\}\\
&\quad\le I_n + \sum_{1\le \ell\le m_1} P\left\{\max_{1\le k<2^\ell} |U_{k,2^\ell}| \ge \frac{x(n2^\ell)^{1/2}}{4}\right\}\\
&\quad\le I_n + \sum_{1\le \ell\le m_1} \frac{36\cdot 2^{2\ell}\sigma^2}{(x/4)^2 n2^\ell}\\
&\quad\le I_n + \frac{3^2 2^6\sigma^2}{x^2}.
\end{aligned}\tag{3.3}$$

Next we estimate the upper bound of I_n. Let

$$\mathscr{F}_\ell = \sigma(X_{2^\ell+1}, X_{2^\ell+2}, \dots, X_n),$$

$$\tau_\ell^2 = E(V_{2^\ell,\ell}^2 | \mathscr{F}_\ell)$$

and

$$J_n = P\Big\{ \bigcup_{1 \le \ell \le m_1} \Big\{ \tau_\ell^2 \ge \frac{\sigma n 2^\ell}{512} \Big\} \Big\}.$$

We note that conditionally on $\mathscr{F}_\ell$, $\{V_{k,\ell}, 1 \le k \le 2^\ell\}$ are partial sums of independent and identically distributed random variables with zero means. By Lévy's inequality [cf. Loéve (1977), page 260], we have

$$\begin{aligned}
I_n &\le J_n + \sum_{1 \le \ell \le m_1} P\Big\{ \max_{1 \le k < 2^\ell} |V_{k,\ell}| \ge \frac{x}{4}(n2^\ell)^{1/2},\ \tau_\ell^2 < \frac{\sigma n 2^\ell}{512} \Big\} \\
&= J_n + \sum_{1 \le \ell \le m_1} EP\Big\{ \max_{1 \le k < 2^\ell} |V_{k,\ell}| \ge \frac{x}{4}(n2^\ell)^{1/2},\ \tau_\ell^2 < \frac{\sigma n 2^\ell}{512} \Big| \mathscr{F}_\ell \Big\} \\
&\le J_n + 2 \sum_{1 \le \ell \le m_1} EP\Big\{ |V_{2^\ell,\ell}| \ge \frac{x}{4}(n2^\ell)^{1/2} - 2^{1/2}\tau_\ell,\ \tau_\ell^2 < \frac{\sigma n 2^\ell}{512} \Big| \mathscr{F}_\ell \Big\} \\
&\le J_n + 2 \sum_{1 \le \ell \le m_1} EP\Big\{ |V_{2^\ell,\ell}| \ge \frac{3x}{16}(n2^\ell)^{1/2},\ \tau_\ell^2 < \frac{\sigma n 2^\ell}{512} \Big| \mathscr{F}_\ell \Big\}.
\end{aligned} \tag{3.4}$$

It is easy to see that

$$\begin{aligned}
\tau_\ell^2 &= 2^\ell E\Big\{ \Big(\sum_{2^\ell < j \le n} h(X_1, X_j) \Big)^2 \Big| \mathscr{F}_\ell \Big\} \\
&= 2^\ell \sum_{1 \le i < \infty} \lambda_i^2 \Big(\sum_{2^\ell < j \le n} \varphi_i(X_j) \Big)^2.
\end{aligned}$$

Lemma 3.2 yields

$$\begin{aligned}
J_n &= P\Big\{ \max_{1 \le \ell \le m_1} \sum_{1 \le i < \infty} \lambda_i^2 \Big(\sum_{2^\ell < j \le n} \varphi_i(X_j) \Big)^2 \ge \frac{\sigma n}{512} \Big\} \\
&\le P\Big\{ \max_{1 \le k \le n} \sum_{1 \le i < \infty} \lambda_i^2 \Big(\sum_{k < j \le n} \varphi_i(X_j) \Big)^2 \ge \frac{\sigma n}{512} \Big\} \\
&\le 1536\sigma.
\end{aligned} \tag{3.5}$$

Let $\{X_i^*,\ 1 \le i < \infty\}$ be an independent copy of $\{X_i,\ 1 \le i < \infty\}$. By the symmetrization inequality [cf. Loéve (1977), pages 257–258], we get

$$
\begin{aligned}
&EP\left\{|V_{2^\ell,\ell}| \ge \frac{3x}{16}(n2^\ell)^{1/2},\ \tau_\ell^2 \le \frac{\sigma n 2^\ell}{512}\Big|\mathscr{F}_\ell\right\}\\
&\quad\le 2EP\left\{\left|\sum_{1\le i\le 2^\ell}\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j))\right|\right.\\
&\qquad\qquad\left.\ge \frac{3x}{16}(n2^\ell)^{1/2}-2\tau_\ell,\ \tau_\ell^2 \le \frac{\sigma n2^\ell}{512}\Big|\mathscr{F}_\ell\right\}\\
&\quad\le 2EP\left\{\left|\sum_{1\le i\le 2^\ell}\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j))\right|\right.\\
&\qquad\qquad\left.\ge \frac{x}{8}(n2^\ell)^{1/2},\ \tau_\ell^2 \le \frac{\sigma n2^\ell}{512}\Big|\mathscr{F}_\ell\right\}\\
&\quad= 2P\left\{\left|\sum_{1\le i\le 2^\ell}\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j))\right| \ge \frac{x}{8}(n2^\ell)^{1/2},\ \tau_\ell^2 \le \frac{\sigma n2^\ell}{512}\right\}\\
&\quad\le 2P\left\{\frac{\left|\sum_{1\le i\le 2^\ell}\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j))\right|}{\left(\sum_{1\le i\le 2^\ell}\left(\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j))\right)^2\right)^{1/2}} \ge \frac{x}{16\sigma^{1/4}}\right\}\\
&\qquad + 2P\left\{\sum_{1\le i\le 2^\ell}\left(\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j))\right)^2 \ge 4\sigma^{1/2}n2^\ell,\ \tau_\ell^2 \le \frac{\sigma n2^\ell}{512}\right\}.
\end{aligned}
\tag{3.6}
$$

Since conditionally on $\mathscr{F}_\ell$, $\{W_i = \sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j)),\ 1 \le i \le 2^\ell\}$ is a sequence of independent and identically distributed symmetric random variables, we obtain [cf. Ledoux and Talagrand (1991), page 91]

$$
\begin{aligned}
&P\left\{\frac{|\sum_{1\le i\le 2^\ell}\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j))|}{(\sum_{1\le i\le 2^\ell}(\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*X_j)))^2)^{1/2}} \ge \frac{x}{16\sigma^{1/4}}\right\}\\
&\quad= EP\left\{\frac{|\sum_{1\le i\le 2^\ell}\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j))|}{(\sum_{1\le i\le 2^\ell}(\sum_{2^\ell<j\le n}(h(X_i,X_j)-h(X_i^*,X_j)))^2)^{1/2}} \ge \frac{x}{16\sigma^{1/4}}\Big|\mathscr{F}_\ell\right\}\\
&\quad\le 2E\left(\exp\left(-\frac{1}{2}\left(\frac{x}{16\sigma^{1/4}}\right)^2\right)\Big|\mathscr{F}_\ell\right)\\
&\quad= 2\exp\left(-\frac{x^2}{512\sigma^{1/2}}\right).
\end{aligned}
\tag{3.7}
$$

Introducing $W_{i,1} = \min(|W_i|, (n2^\ell)^{1/2})$, we can write

$$\begin{aligned}
&P\bigg\{\sum_{1\le i\le 2^\ell}\bigg(\sum_{2^\ell<j\le n}(h(X_i, X_j) - h(X_i^*, X_j))\bigg)^2 \ge 4\sigma^{1/2}n2^\ell,\ \tau_\ell^2 \le \frac{\sigma n2^\ell}{512}\bigg\}\\
&\qquad \le P\bigg\{\sum_{1\le i\le 2^\ell} W_{i,1}^2 \ge 4\sigma^{1/2}n2^\ell,\ \tau_\ell^2 \le \frac{\sigma n2^\ell}{512}\bigg\} \qquad (3.8)\\
&\qquad\quad + P\Big\{\max_{1\le i\le 2^\ell}|W_i| \ge (n2^\ell)^{1/2}\Big\}.
\end{aligned}$$

Using the Chebyshev inequality, we have

$$\begin{aligned}
&P\bigg\{\sum_{1\le i\le 2^\ell} W_{i,1}^2 \ge 4\sigma^{1/2}n2^\ell,\ \tau_\ell^2 \le \frac{\sigma n2^\ell}{512}\bigg\}\\
&\quad = EP\bigg\{\sum_{1\le i\le 2^\ell}(W_{i,1}^2 - E(W_{i,1}^2|\mathscr{F}_\ell))\\
&\qquad\qquad \ge 4\sigma^{1/2}n2^\ell - 2^\ell E(W_{1,1}^2|\mathscr{F}_\ell), \tau_\ell^2 \le \frac{\sigma\, n2^\ell}{512}\bigg|\mathscr{F}_\ell\bigg\}\\
&\quad \le EP\bigg\{\sum_{1\le i\le 2^\ell}(W_{i,1}^2 - E(W_{i,1}^2|\mathscr{F}_\ell)) \ge 4\sigma^{1/2}n2^\ell - \frac{\sigma n2^\ell}{512}\bigg|\mathscr{F}_\ell\bigg\}\\
&\quad \le EP\bigg\{\sum_{1\le i\le 2^\ell}(W_{i,1}^2 - E(W_{i,1}^2|\mathscr{F}_\ell)) \ge \sigma^{1/2}n2^\ell\bigg|\mathscr{F}_\ell\bigg\} \qquad (3.9)\\
&\quad \le E\bigg(\frac{2^\ell E(W_{1,1}^4|\mathscr{F}_\ell)}{(n\sigma^{1/2}2^\ell)^2}\bigg) = \frac{EW_{1,1}^4}{\sigma 2^\ell n^2}\\
&\quad = \frac{E\min\{(\sum_{2^\ell<j\le n}(h(X_1, X_j) - h(X_1^*, X_j)))^4, (n2^\ell)^2\}}{\sigma 2^\ell n^2}\\
&\quad \le \frac{E\min\{\max_{1\le k<n}(\sum_{k<j\le n}(h(X_1, X_j) - h(X_1^*, X_j)))^4, (n2^\ell)^2\}}{\sigma 2^\ell n^2}.
\end{aligned}$$

It follows from (3.9) that

$$\begin{aligned}
&\sum_{1\le \ell\le m_1} P\bigg\{\sum_{1\le i\le 2^\ell} W_{i,1}^2 \ge 4\sigma^{1/2}n2^\ell,\ \tau_\ell^2 \le \frac{\sigma}{512}n2^\ell\bigg\}\\
&\qquad \le \frac{1}{\sigma n^2}\sum_{1\le\ell\le m_1}\frac{1}{2^\ell}E\min(U_n^{*4}, (n2^\ell)^2)\\
&\qquad \le \frac{1}{\sigma n^2}\sum_{1\le\ell\le m_1}\frac{1}{2^\ell}EU_n^{*4}I\{|U_n^*| \le n^{1/2}2^{\ell/2}\} \qquad (3.10)\\
&\qquad\quad + \frac{1}{\sigma}\sum_{1\le\ell\le m_1}2^\ell P\{|U_n^*| \ge (n2^\ell)^{1/2}\},
\end{aligned}$$

where $U_n^* = \max_{1\le k<n} \sum_{k<j\le n}(h(X_1, X_j) - h(X_1^*, X_j))$. It is easy to see that

$$
\begin{aligned}
&\sum_{1\le \ell\le m_1} 2^{-\ell} EU_n^{*4} I\{|U_n^*| \le n^{1/2}2^{\ell/2}\} \\
&\quad = \sum_{1\le \ell\le m_1} 2^{-\ell} EU_n^{*4} I\{|U_n^*| \le n^{1/2}\} \\
&\qquad + \sum_{1\le \ell\le m_1} \sum_{1\le j\le \ell} 2^{-\ell} EU_n^{*4} I\{n^{1/2}2^{(j-1)/2} < |U_n^*| \le n^{1/2}2^{j/2}\} \\
&\quad \le nEU_n^{*2} + \sum_{1\le j\le m_1} \sum_{j\le \ell\le m_1} 2^{-\ell} EU_n^{*4} I\{n^{1/2}2^{(j-1)/2} < |U_n^*| \le n^{1/2}2^{j/2}\} \qquad (3.11)\\
&\quad \le 3nEU_n^{*2} = 3nE \max_{1\le k<n}\biggl(\sum_{k<j\le n}(h(X_1, X_j) - h(X_1^*, X_j))\biggr)^2 \\
&\quad = 3nE\biggl\{E\biggl(\max_{1\le k<n}\biggl(\sum_{k<j\le n}(h(X_1, X_j) - h(X_1^*, X_j))\biggr)^2\bigg| X_1, X_1^*\biggr)\biggr\} \\
&\quad \le 24n^2\sigma^2.
\end{aligned}
$$

Similarly, we have

$$
\sum_{1\le \ell\le m_1} 2^{\ell} P\{|U_n^*| \ge n^{1/2}2^{\ell/2}\} \le 48\sigma^2. \qquad (3.12)
$$

Therefore

$$
\begin{aligned}
\sum_{1\le \ell\le m_1} P\Bigl\{\max_{1\le i\le 2^\ell} |W_i| \ge (n2^\ell)^{1/2}\Bigr\} &\le \sum_{1\le \ell\le m_1} 2^\ell P\{|U_n^*| \ge (n2^\ell)^{1/2}\} \qquad (3.13)\\
&\le 48\sigma^2.
\end{aligned}
$$

Putting (3.8)–(3.13) together, we get

$$
\begin{aligned}
&\sum_{1\le \ell\le m_1} P\biggl\{\sum_{1\le i\le 2^\ell}\biggl(\sum_{2^\ell<j\le n}(h(X_i, X_j) - h(X_i^*, X_j))\biggr)^2 \ge 4\sigma^{1/2}n2^\ell,\ \tau_\ell^2 \le \frac{\sigma n 2^\ell}{512}\biggr\} \\
&\quad \le 120\sigma,
\end{aligned}
$$

which, together with (3.1)–(3.7), implies

$$
P\biggl\{\max_{1\le k\le n/2} \frac{|U_{k,n}|}{(k(n-k))^{1/2}} \ge x\biggr\} \le (10^4 + 10^3)\sigma + 16\exp\biggl(-\frac{x^2}{512\sigma^{1/2}}\biggr)\log n.
$$

By symmetry, we have

$$
P\biggl\{\max_{n/2\le k<n} \frac{|U_{k,n}|}{(k(n-k))^{1/2}} \ge x\biggr\} \le (10^4 + 10^3)\sigma + 16\exp\biggl(-\frac{x^2}{512\sigma}\biggr)\log n.
$$

This completes the proof of Lemma 3.3. □

Let $1 \leq M < \infty$ and define

$$h_M(x, y) = \sum_{1 \leq j \leq M} \lambda_j \varphi_j(x) \varphi_j(y), \qquad \xi_M = \Bigl(\sum_{1 \leq j \leq M} \lambda_j^2 N_j^2\Bigr)^{1/2}. \tag{3.14}$$

The corresponding U-statistics are

$$U_{k,n}^{(M)} = \sum_{1 \leq i \leq k} \sum_{k < j \leq n} h_M(X_i, X_j), \qquad 1 \leq k < n.$$

Our last lemma shows that Theorem 1.1 is true if the kernel is given by (3.14).

LEMMA 3.4. *As $n \to \infty$, we have*

$$(2 \log\log n)^{-1/2} \max_{1 \leq k < n} \frac{|U_{k,n}^{(M)}|}{(k(n-k))^{1/2}} \to_{\mathscr{D}} \xi_M.$$

PROOF. Let

$$S_m(k) = \sum_{1 \leq i \leq k} \varphi_m(X_i), \qquad 1 \leq m \leq M.$$

Then, elementary calculations give

$$U_{k,n}^{(M)} = \sum_{1 \leq m \leq M} \lambda_m S_m(k)(S_m(n) - S_m(k)).$$

Let $a = n/(\log n)^2$ and write

$$\max_{1 \leq k < n} \frac{|U_{k,n}^{(M)}|}{(k(n-k))^{1/2}} = \max(T_1, T_2, T_3),$$

where $T_1 = \max_{1 \leq k < a} |U_{k,n}^{(M)}|/(k(n-k))^{1/2}$, $T_2 = \max_{a \leq k \leq n-a} |U_{k,n}^{(M)}|/(k(n-k))^{1/2}$ and $T_3 = \max_{n-a < k < n} |U_{k,n}^{(M)}|/(k(n-k))^{1/2}$. Since

$$\max_{a \leq k \leq n-a} |S_m(k)|/k^{1/2} = O_P((\log\log\log n)^{1/2})$$

and

$$\max_{a \leq k \leq n-a} |S_m(n) - S_m(k)|/(n-k)^{1/2} = O_P((\log\log\log n)^{1/2}),$$

we get immediately that

$$T_2 = O_P(\log\log\log n). \tag{3.15}$$

By the weak convergence of partial sums and the continuity of Brownian motion, we have

$$\max_{1 \leq k < a} \left| \frac{S_m(n) - S_m(k)}{(n-k)^{1/2}} - \frac{S_m(n-a) - S_m(a)}{n^{1/2}} \right| = o_P(1) \tag{3.16}$$

and similarly

$$\max_{n-a<k<n}\left|\frac{S_m(k)}{k^{1/2}}-\frac{S_m(n-a)-S_m(a)}{n^{1/2}}\right|=o_P(1). \tag{3.17}$$

Using the law of the iterated logarithm, we get

$$\max_{1\le k<a}|S_m(k)|/k^{1/2}=O_P((\log\log n)^{1/2}) \tag{3.18}$$

and

$$\max_{n-a<k<n}|S_m(n)-S_m(k)|/(n-k)^{1/2}=O_P((\log\log n)^{1/2}). \tag{3.19}$$

Putting (3.15)–(3.19) together, we obtain that

$$\left|\max_{1\le k<n}\frac{|U_{k,n}^{(M)}|}{(k(n-k))^{1/2}}-T_n^*\right|=O_P((\log\log n)^{1/2}), \tag{3.20}$$

where

$$\begin{aligned}T_n^*=\max\Biggl\{&\max_{1\le k<a}\left|\sum_{1\le m\le M}\lambda_m\frac{S_m(k)}{k^{1/2}}\,\frac{S_m(n-a)-S_m(a)}{n^{1/2}}\right|,\\&\max_{n-a<k<n}\left|\sum_{1\le m\le M}\lambda_m\frac{S_m(n)-S_m(k)}{(n-k)^{1/2}}\,\frac{S_m(n-a)-S_m(a)}{n^{1/2}}\right|\Biggr\}.\end{aligned}$$

Applying the multivariate Strassen's invariance principle, for each n we can define independent Brownian motions $W_{1,n},\dots,W_{M,n}$ such that

$$\begin{aligned}T_n^*=\max\Biggl\{&\max_{1\le k<a}\left|\sum_{1\le m\le M}\lambda_m\frac{W_{m,n}(k)}{k^{1/2}}\,\frac{W_{m,n}(n-a)-W_{m,n}(a)}{n^{1/2}}\right|,\\&\max_{n-a<k<n}\Biggl|\sum_{1\le m\le M}\lambda_m\frac{W_{m,n}(n)-W_{m,n}(k)}{(n-k)^{1/2}}\\&\qquad\times\frac{W_{m,n}(n-a)-W_{m,n}(a)}{n^{1/2}}\Biggr|\Biggr\}\\&+o((\log\log n)^{1/2}).\end{aligned} \tag{3.21}$$

Noting that for any constants $\alpha_1,\alpha_2,\dots,\alpha_M$

$$\left\{\sum_{1\le i\le M}\alpha_iW_{i,n}(t),\ 0\le t<\infty\right\}=_{\mathscr{D}}\{(\alpha_1^2+\cdots+\alpha_M^2)^{1/2}W(t),\ 0\le t<\infty\},$$

where $W(t)$ is a Brownian motion, we have

$$\max_{1\le k\le a}\left|\sum_{1\le i\le M}\alpha_i\frac{W_{i,n}(k)}{k^{1/2}}\right|\Big/(2\log\log n)^{1/2}\to_P(\alpha_1^2+\cdots+\alpha_M^2)^{1/2} \tag{3.22}$$

and similarly,

$$\max_{n-a\le k<n}\left|\sum_{1\le i\le M}\alpha_i\frac{W_{i,n}(n)-W_{i,n}(k)}{(n-k)^{1/2}}\right|\Big/(2\log\log n)^{1/2} \tag{3.23}$$
$$\to_P (\alpha_1^2+\cdots+\alpha_M^2)^{1/2}.$$

It is easy to see that $\{W_{i,n}(k),\ 1\le k<a\}$, $W_{i,n}(a)-W_{i,n}(n-a)$, $\{W_{i,n}(n)-W_{i,n}(k),\ n-a<k<n\}$, $1\le i\le M$, are independent and

$$\{n^{-1/2}(W_{i,n}(n-a)-W_{i,n}(a)),\ 1\le i\le M\}\to_{\mathscr{D}}\{N_i,\ 1\le i\le M\}.$$

Now Lemma 3.4 follows from (3.20)–(3.23). □

PROOF OF THEOREM 1.1. Let

$$\tilde h_M(x,y)=h(x,y)-h_M(x,y),\qquad \tilde U_{k,n}^{(M)}=\sum_{1\le i\le k}\ \sum_{k<j\le n}\tilde h_M(x,y)$$

and

$$\tilde\sigma_M^2=\sum_{M<i<\infty}\lambda_i^2.$$

Using Lemma 3.3, we get

$$P\left\{\max_{1\le k<n}\frac{|\tilde U_{k,n}^{(M)}|}{(k(n-k))^{1/2}}\ge(512\tilde\sigma_M^{1/2})^{1/2}(2\log\log n)^{1/2}\right\}\le 10^5\tilde\sigma_M+\frac{32}{\log n},$$

and therefore

$$\lim_{M\to\infty}\ \limsup_{n\to\infty}P\left\{\frac{1}{(2\log\log n)^{1/2}}\ \max_{1\le k<n}\frac{|\tilde U_{k,n}^{(M)}|}{(k(n-k))^{1/2}}>\varepsilon\right\}=0$$

for all $\varepsilon>0$. Now the result follows from Lemma 3.4. □

Acknowledgments. We thank three referees and an Associate Editor for their helpful comments and suggestions.

REFERENCES

BRODSKY, B. E. and DARKHOVSKY, B. S. (1993). *Nonparametric Methods in Change-Point Problems*. Kluwer, Dordrecht.

CHOW, Y. S. and TEICHER, H. (1988). *Probability Theory*, 2nd ed. Springer, New York.

COX, D. R. and LEWIS, P. A. W. (1966). *The Statistical Analysis of Series of Events*. Methuen, London.

CSÖRGŐ, M. and HORVÁTH, L. (1987). Nonparametric tests for the changepoint problem. *J. Statist. Plann. Inf.* **17** 1–9.

CSÖRGŐ, M. and HORVÁTH, L. (1988). Invariance principles for changepoint problems. *J. Multivariate Anal.* **27** 151–168.

FERGER, D. (1994a). Change-point estimators in case of small disorders. *J. Statist. Plann. Inf.* **40** 33–49.

FERGER, D. (1994b). An extension of the Csörgő–Horváth functional limit theorem and its applications to changepoint problems. *J. Multivariate Anal.* **51** 338–351.

FERGER, D. (1994c). Nonparametric detection of change-points for sequentially observed data. *Stochastic Process. Appl.* **51** 359–372.

FERGER, D. and STUTE, W. (1992). Convergence of changepoint estimators. *Stochastic Process. Appl.* **42** 345–351.

GOMBAY, E. and HORVÁTH, L. (1990). Asymptotic distributions of maximum likelihood tests for change in the mean. *Biometrika* **77** 411–414.

JARRETT, R. G. (1979). A note on the intervals between coal-mining disasters. *Biometrika* **66** 191–193.

KOROLJUK, V. S. and BOROVSKICH, YU. V. (1994). *Theory of U-Statistics.* Kluwer, Dordrecht.

LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces*. Springer, New York.

LEE, A. J. (1990). *U-Statistics.* Dekker, New York.

LOÉVE, M. (1977). *Probability Theory* **1**, 4th ed. Springer, New York.

MAGUIRE, B. A., PEARSON, E. S. and WYNN, A. H. A. (1952). The time intervals between industrial accidents. *Biometrika* **39** 168–180.

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

WOLFE, D. E. and SCHECHTMAN, E. (1984). Nonparametric statistical procedures for the change-point problem. *J. Statist. Plann. Inf.* **9** 389–396.

WORSLEY, K. J. (1986). Confidence regions and tests for a change-point in a sequence of exponential family random variables. *Biometrika* **73** 91–104.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
SALT LAKE CITY, UTAH 84112
E-MAIL: horvath@math.utah.edu

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OREGON
EUGENE, OREGON 97403
E-MAIL: shao@math.uoregon.edu