Approximations for weighted bootstrap processes
with an application

Lajos Horváth\textsuperscript{a,∗}, Piotr Kokoszka\textsuperscript{b}, Josef Steinebach\textsuperscript{c}

\textsuperscript{a} Department of Mathematics, University of Utah, Salt Lake City, UT 84112, USA
\textsuperscript{b} Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 3BX, UK
\textsuperscript{c} Fachbereich Mathematik und Informatik, Philipps-Universität, D-35032 Marburg, Germany

Received May 1999; received in revised form August 1999

Abstract

Let $\beta_n(t)$ denote the weighted (smooth) bootstrap process of an empirical process. We show that the order of the best Gaussian approximation for $\beta_n(t)$ is $n^{-1/2}\log n$ and we construct a sequence of approximating Brownian bridges achieving this rate. We also obtain an approximation for $\beta_n(t)$ using a suitably chosen Kiefer process. The result is applied to detect a possible change in the distribution of independent observations.

Keywords: Weighted bootstrap; Erdős–Rényi law; Brownian bridge; Best approximation; Change-point detection

1. Approximations for empirical and resampled processes

Let $X_1,X_2,...$ be a sequence of independent, identically distributed random variables with common distribution function $F$. The empirical distribution of $X_1,X_2,...,X_n$ is defined as

$$F_n(t) = \frac{1}{n} \sum_{1 \leq i \leq n} I\{X_i \leq t\}, \quad -\infty < t < \infty.$$ 

It is well-known that the empirical process $Z_n(t) = n^{1/2}(F_n(t) - F(t))$ converges in $\mathcal{D}[-\infty,\infty]$ to $B(F(t))$, where $\{B(t), 0 \leq t \leq 1\}$ stands for a Brownian bridge. The rate of convergence of $Z_n(t)$ to $B(F(t))$ is an important question in statistics as well as in probability and it was investigated by several authors. For a survey we refer to Csörgő and Révész (1981). The best result is due to Komlós et al. (1975):
Theorem 1.1. There exists a sequence of Brownian bridges \( \{B_n(t), \ 0 \leq t \leq 1\} \) such that
\[
P \left( \sup_{-1 < t < 1} |\sigma_n(t) - B_n(F(t))| > n^{-1/2}(c_1 \log n + x) \right) \leq c_2 \exp(-c_3 x)
\]
for all \( x > 0 \), where \( c_1, c_2 \) and \( c_3 \) are positive constants.

Komlós et al. (1975) also showed that the rate in Theorem 1.1 is the best possible one.

Theorem 1.2. If \( F \) is continuous, then for any sequence of Brownian bridges \( \{B_n^*(t), \ 0 \leq t \leq 1\} \) we have
\[
\lim_{n \to \infty} P \left( \sup_{-1 < t < 1} |\sigma_n(t) - B_n^*(F(t))| > \frac{1}{6} n^{-1/2} \log n \right) = 1.
\]

Several statistical procedures, including kernel density estimation and procedures based on the empirical characteristic function (see, e.g. Feuerverger and Mureika, 1977; Csörgő, 1981) can be described in terms of the empirical process \( \sigma_n(t) \) with the limit distributions given as functionals of \( B(F(t)) \). Thus the limits may depend on the unknown distribution \( F \). The bootstrap is a widely used technique to approximate these limiting distributions. For the bootstrapped version of Theorems 1.1 and 1.2 we refer to Csörgő et al. (1999) and Horváth and Steinebach (1999). One of the possible drawbacks of Efron’s (1979) original bootstrap is that some observations may be used more than once while others are not sampled at all. An alternative is the weighted (or smooth) bootstrap, which has also been shown to be computationally more efficient in several applications. The Bayesian bootstrap of Rubin (1981) utilizes exponential weights. Lo (1987) obtained approximations for \( \beta_n(t) \) when \( \epsilon_1 \) has an exponential distribution. Parzen et al. (1994) used standard normal weights to estimate the variance of an estimator based on an estimating function. Burke (1997a,b, 1998) utilizes standard normal weights to construct confidence bands for functionals of multivariate observations. For a survey of further results on weighted bootstrap we refer to Barbe and Bertrand (1995) and Shao and Tu (1996).

In this paper we study the asymptotic properties of the weighted bootstrap for the empirical process which is defined as follows: Let \( \epsilon_1, \epsilon_2, \ldots \) be a sequence of independent, identically distributed random variables with
\[
E \epsilon_1 = \mu, \quad \text{var} \epsilon_1 = 1
\]
and
\[
E \exp(t \epsilon_1) < \infty \quad \text{if} \ |t| \leq t_0 \quad \text{with some} \ t_0 > 0.
\]
We also assume that
\[
\{X_i, \ 1 \leq i < \infty\} \quad \text{and} \quad \{\epsilon_i, 1 \leq i < \infty\} \quad \text{are independent.}
\]
The smooth bootstrap of \( \sigma_n \) is
\[
\beta_n(t) = n^{-1/2} \sum_{1 \leq i \leq n} (\epsilon_i - \bar{\epsilon}_n) 1\{X_i \leq t\}, \quad -\infty < t < \infty,
\]
where
\[
\bar{\epsilon}_n = \frac{1}{n} \sum_{1 \leq i \leq n} \epsilon_i.
\]
We note that \( \beta_n(t) \) does not depend on \( \mu \).

The main aim of this paper is to obtain the best possible approximation for \( \beta_n(t) \). We prove that the best possible rate is \( n^{-1/2} \log n \) and our construction reaches this rate. Thus we have a complete analogue of Theorems 1.1 and 1.2 for the smooth bootstrap.
Theorem 1.3. If (1.1)–(1.3) hold, then there exists a sequence of Brownian bridges \( \{ \hat{B}_n(t), 0 \leq t \leq 1 \} \) such that

\[
P \left\{ \sup_{-\infty < t < \infty} |\beta_n(t) - \hat{B}_n(F(t))| > n^{-1/2}(c_4 \log n + x) \right\} \leq c_5 \exp(-c_6x)
\]

for all \( x > 0 \), where \( c_4, c_5 \) and \( c_6 \) are positive constants.

The next result shows that the rate in Theorem 1.3 is optimal.

Theorem 1.4. If \( F \) is continuous and (1.1)–(1.3) hold, then for any sequence of Brownian bridges \( \{ B_n^*(t), 0 \leq t \leq 1 \} \) we have

\[
\lim_{n \to \infty} P \left\{ \sup_{-\infty < t < \infty} |\beta_n(t) - B_n^*(F(t))| \geq c_7n^{-1/2} \log n \right\} = 1
\]

with some \( c_7 > 0 \).

Since a sequence of Brownian bridges is used in Theorem 1.3, it does not immediately imply strong laws for \( \beta_n(t) \). Also, in some applications we need \( \beta_n(t) \) as a two-time parameter process of \( (t, n) \). The next result gives an almost sure approximation for \( \beta_n(t) \) based on a Kiefer process.

Theorem 1.5. There is a Kiefer process \( \{ K(t, x), 0 \leq t \leq 1, 0 \leq x \leq \infty \} \) such that

\[
\max_{1 \leq k \leq n} \sup_{-\infty < t < \infty} \left| \sum_{1 \leq i \leq k} (e_i - \bar{e}_n) I\{ X_i \leq t \} - K(F(t), k) \right| \overset{a.s.}{=} O\left(n^{1/4}(\log n)^{1/2}\right).
\]

The proofs are given in Section 3. First we show how the weighted bootstrap can be used in change-point analysis.

2. An application to change-point detection

In this section we discuss one possible application of the results established in Section 1.

Let \( Y_1, Y_2, \ldots, Y_n \) be independent random variables with distribution functions \( H_{(1)}(t), H_{(2)}(t), \ldots, H_{(n)}(t) \). We wish to test the null hypothesis

\( H_0: H_{(1)}(t) = H_{(2)}(t) = \cdots = H_{(n)}(t) \) for all \( t \)

against the alternative

\( H_1: \) there is \( k^* \), \( 1 \leq k^* < n \) such that \( H_{(1)}(t) = \cdots = H_{(k^*)}(t), H_{(k^*+1)}(t) = \cdots = H_{(n)}(t) \) for all \( t \) and \( H_{(k^*)}(t_0) \neq H_{(k^*+1)}(t_0) \) with some \( t_0 \).

Following Csörgő and Horváth (1997, p. 152) we divide the data into two subsets, before and after the \( k \)th observation, and compute the corresponding empirical distribution functions

\[
H_k(t) = \frac{1}{k} \sum_{1 \leq i \leq k} I\{ Y_i \leq t \} \quad \text{and} \quad H^*_k(t) = \frac{1}{n-k} \sum_{k+1 \leq i \leq n} I\{ Y_i \leq t \}.
\]
We reject $H_0$, if
\[
R_n = \max_{1 \leq k < n} \frac{k(n-k)}{n^{3/2}} \sup_{-\infty < t < \infty} |H_k(t) - H_k^*(t)|
\]
is large. Csörgő and Horváth (1997, p. 153) obtained the limit distribution of $R_n$ under $H_0$. Let $H$ denote the common distribution function under $H_0$. If $H_0$ holds, then
\[
R_n \xrightarrow{D} \sup_{0 < x \leq 1} \sup_{-\infty < t < \infty} \left[ \Gamma_H(t,x) \right],
\]
where $\{\Gamma_H(t,x), 0 \leq x \leq 1, -\infty < t < \infty\}$ is a Gaussian process with $E \Gamma_H(t,x) = 0$ and $E \Gamma_H(t,x) \Gamma_H(s,y) = \{H(\min(t,s)) - H(t)H(s)\} \{\min(x, y) - xy\}$.

The Glivenko–Cantelli lemma yields that if $H_0$ holds and $k = \lfloor n\theta \rfloor$ with some $0 < \theta < 1$, then
\[
k^*(n-k^*) \frac{1}{n^2} [H_{k^*}(t_0) - H_{k^*}(t_0)] \xrightarrow{p} \theta(1 - \theta)[H_{k^*}(t_0) - H_{(k^*+1)}(t_0)],
\]
and therefore
\[
R_n \xrightarrow{p} \infty.
\]

We use the smooth bootstrap to approximate the distribution of $R_n$. Let $\bar{e}_1, \bar{e}_2, \ldots$ be independent identically distributed random variables satisfying (1.1) and (1.2). Define
\[
r(t,k) = \sum_{1 \leq i \leq k} (e_i - \bar{e}_k) I\{Y_i \leq t\}
\]
and
\[
r^*(t,k) = \sum_{k < i \leq n} (e_i - \bar{e}^*_k) I\{Y_i \leq t\}
\]
with
\[
\bar{e}_k = \frac{1}{k} \sum_{1 \leq i \leq k} e_i \quad \text{and} \quad \bar{e}^*_k = \frac{1}{n-k} \sum_{k < i \leq n} e_i.
\]

We show that $R_n$ and
\[
R^*_n = n^{-3/2} \max_{1 \leq k < n, -\infty < t < \infty} |(n-k)r(t,k) - kr^*(t,k)|
\]
have the same limit distributions.

**Theorem 2.1.** We assume that (1.1) and (1.2) are satisfied.

(i) If $H_0$ holds, then
\[
R^*_n \xrightarrow{D} \sup_{0 < x \leq 1} \sup_{-\infty < t < \infty} \left[ \Gamma_H(t,x) \right].
\]

(ii) If $H_0$ holds, then
\[
R^*_n = O_P(1).
\]

Theorem 2.1 is proved in Section 3.

In the remainder of this section we present the results of a small simulation study which compares the critical values of $R_n$ and $R^*_n$ for small $n$ and a data example which illustrates the procedure.
Table 1
Percentiles of $R_n$ and $R^*_n$ under $H_0$, $n = 25$

<table>
<thead>
<tr>
<th>Distribution $H$</th>
<th>Statistic</th>
<th>90.0%</th>
<th>95.0%</th>
<th>97.5%</th>
<th>99.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>$R_n$</td>
<td>0.66</td>
<td>0.72</td>
<td>0.76</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>$R^*_n$</td>
<td>0.60</td>
<td>0.66</td>
<td>0.71</td>
<td>0.78</td>
</tr>
<tr>
<td>Uniform</td>
<td>$R_n$</td>
<td>0.66</td>
<td>0.72</td>
<td>0.79</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>$R^*_n$</td>
<td>0.61</td>
<td>0.68</td>
<td>0.73</td>
<td>0.79</td>
</tr>
<tr>
<td>Cauchy</td>
<td>$R_n$</td>
<td>0.66</td>
<td>0.72</td>
<td>0.76</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>$R^*_n$</td>
<td>0.60</td>
<td>0.66</td>
<td>0.72</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Table 2
Percentiles of $R_n$ and $R^*_n$ under $H_0$, $n = 50$

<table>
<thead>
<tr>
<th>Distribution $H$</th>
<th>Statistic</th>
<th>90.0%</th>
<th>95.0%</th>
<th>97.5%</th>
<th>99.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>$R_n$</td>
<td>0.69</td>
<td>0.75</td>
<td>0.81</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>$R^*_n$</td>
<td>0.66</td>
<td>0.72</td>
<td>0.79</td>
<td>0.85</td>
</tr>
<tr>
<td>Uniform</td>
<td>$R_n$</td>
<td>0.69</td>
<td>0.74</td>
<td>0.80</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>$R^*_n$</td>
<td>0.65</td>
<td>0.71</td>
<td>0.75</td>
<td>0.80</td>
</tr>
<tr>
<td>Cauchy</td>
<td>$R_n$</td>
<td>0.69</td>
<td>0.74</td>
<td>0.81</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>$R^*_n$</td>
<td>0.65</td>
<td>0.71</td>
<td>0.77</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Tables 1 and 2 exhibit percentiles of the simulated values of $R_n$ and $R^*_n$ for $n = 25$ and $n = 50$. The simulations were done assuming that $H_0$ is true. Three underlying distributions $H$ were considered: standard Gaussian, standard uniform and standard Cauchy. At least $N=2,500$ repetitions were used for each entry. For each replication of $R_n$ a new sample from $H$ was generated. To obtain the $R^*_n$’s, one “original” sample from $H$ was generated, and based on it, $N$ Gaussian bootstrap samples (i.e. with standard normal $v_i$’s) were simulated.

The critical values of $R^*_n$ appear to be acceptable approximations of the critical values of $R_n$ for $n \geq 50$, even though they are systematically smaller. The approximation is worse if standard normal $v_i$’s are replaced by standard exponential or normalized uniform weights.

We note that for small $n$ the distributions of $R_n$ and $R^*_n$ look very different, even though they have similar quantiles. In both panels of Fig. 1, $H$ is standard Gaussian, $n = 25$ ($N = 5,000$ replications were used).

The difference is due to the absence of the smoothing effect of the $v_i$ in the definition of $R_n$ (standard normal $v_i$’s were used in the right panel).

2.1. A data example

We consider the spontaneous abortion data studied by Levin and Kline (1985). The data consist of 265 units comprising four karyotyped aborted conceptions. Each conception falls into one of five disjoint groups. The groups correspond to different chromosomal characteristics. The units roughly correspond to weeks, since, on average, four karyotyping procedures per week yielded conclusive results. We focus here on trisomic abortions, because for them the analysis of Levin and Kline is somewhat inconclusive. Using a CUSUM test, they test if during the period of time corresponding to the 265 units there has been a change in the proportion of trisomic abortions. Assuming binomial distribution, they obtain the $P$-value of 0.14. On the other hand, their test has power of only 0.39 against the likely alternative of a two-fold increase in the proportion of trisomic abortions.
over a time span of 20 units. Using a test based on local polynomial smoothing, Horváth and Kokoszka (1998) obtained asymptotic $P$-value smaller than 0.025. The trisomic abortions data are displayed in Fig. 2, which shows that the detection of a possible change in distribution is a challenging problem.

We used Theorem 2.1 to compute the $P$-value for the test of the null hypothesis that there has been no change in the proportion of trisomic abortions. The mean of the data is 0.883, so we assumed that under the null hypothesis, the data come from the binomial distribution with $p = 0.22$ (the value of $p$ does not effect the results). We generated a random sample of size $n = 265$ (the number of units) from this distribution and then repeating the Gaussian smooth bootstrap $N = 1000$ times we obtained $R_{n,1}^*, \ldots, R_{n,N}^*$ and used

$$L_N(x) = \frac{1}{N} \sum_{1 \leq j \leq N} I\{R_{n,j}^* \leq x\}$$

to approximate the distribution of $R_n$. The proportion of the $R_{n,j}^*$ which were greater than the observed value of $R_n$ was 0.034, which is our estimate of the $P$-value.
For comparison, we simulated $N = 500$ samples of length $n = 265$ from the binomial distribution with $p = 0.22$ and computed 500 values of $R_n$. The proportion of the $R_n$ which were greater than the observed value of $R_n$ was 0.016, which indicates that the test based on Gaussian bootstrap is fairly accurate.

According to Theorem 2.1, if $z_N(1 - \alpha)$ denotes the $1 - \alpha$ quantile of $L_N(x)$, then by (2.1) and (2.3) we have $P(R_n \leq z_N(1 - \alpha)) \rightarrow 1 - \alpha$ as $n, N \rightarrow \infty$ under $H_0$. If $H_a$ holds, then (2.2) and (2.4) yield $P(R_n \leq z_N(1 - \alpha)) \rightarrow 0$ as $n, N \rightarrow \infty$, so the test is consistent. This is illustrated in Fig. 3, where the continuous line shows the estimated density of $R^*_n$ under the null described above, whereas the dotted line shows the estimated distribution of $R_n$ under the alternative of the twofold increase of $p$ over twenty units. The vertical line indicates the observed value of $R_n$.

3. Proofs

We can assume that $\mu = 0$.

**Proof of Theorem 1.3.** First we can write

$$\sum_{1 \leq i \leq n} (\tilde{e}_i - \bar{e}) I\{X_i \leq t\} = \sum_{1 \leq i \leq n} \tilde{e}_i I\{X_i \leq t\} - F(t) \sum_{1 \leq i \leq n} e_i + (F(t) - F_n(t)) \sum_{1 \leq i \leq n} e_i.$$  (3.1)

By the inequality of Dvoretzky et al. (1956) we have that

$$P\left\{ \sup_{-\infty < t < \infty} |F_n(t) - F(t)| > n^{-1/2}x^{1/2} \right\} \leq c_8 \exp(-2x)$$  (3.2)

for all $x \geq 0$ with some constant $c_8$. Using Petrov (1995, p. 55) there are constants $c_9, \ldots, c_{13}$ such that

$$P\left\{ \left| \sum_{1 \leq i \leq n} \tilde{e}_i \right| > (xn)^{1/2} \right\} \leq c_9 \exp(-c_{10}x)$$  (3.3)

for all $0 \leq x \leq c_{11}n$ and

$$P\left\{ \left| \sum_{1 \leq i \leq n} e_i \right| \geq x \right\} \leq c_{12} \exp(-c_{13}x).$$  (3.4)
if \( c_{11} n \leq x < \infty \). If \( 0 \leq x \leq c_{11} n \), then we use (3.2) and (3.3) and if \( c_{11} n \leq x < \infty \), then we apply (3.4) and the observation that \( \sup_{t \in (-\infty, \infty)} |F_n(t) - F(t)| \leq 1 \) to obtain

\[
P\left\{ \sup_{-\infty < t < \infty} \left| (F_n(t) - F(t)) \sum_{1 \leq i \leq n} e_i \right| > x \right\} \leq c_{14} \exp(-c_{15} x) \tag{3.5}
\]

for all \( x \geq 0 \). Horváth (2000) constructed a sequence of Wiener processes \( \{W_n(t), 0 \leq t < \infty\} \) such that

\[
P\left\{ \sup_{-\infty < t < \infty} \left| n^{-1/2} \sum_{1 \leq i \leq n} e_i I\{X_i \leq t\} - W_n(F(t)) \right| > n^{-1/2}(c_{16} \log n + x) \right\} \leq c_{17} \exp(-c_{18} x) \tag{3.6}
\]

for all \( x \geq 0 \), and therefore

\[
P\left\{ \sup_{-\infty < t < \infty} \left| n^{-1/2} \left( \sum_{1 \leq i \leq n} e_i I\{X_i \leq t\} - F(t) \sum_{1 \leq i \leq n} e_i \right) - (W_n(F(t)) - F(t)W_n(1)) \right| > 2n^{-1/2}(c_{16} \log n + x) \right\} \leq c_{17} \exp(-c_{18} x). \tag{3.7}
\]

Observing that \( B_n(t) = W_n(t) - tW_n(1) \) are Brownian bridges, Theorem 1.3 follows from (3.1), (3.5) and (3.7). \( \square \)

**Proof of Theorem 1.4.** We can assume without loss of generality that \( F(t) = t, 0 \leq t \leq 1 \). Let \( \{N(t), 0 \leq t < \infty\} \) be a homogeneous Poisson process with intensity parameter 1, independent of \( \{e_i, 1 \leq i < \infty\} \). The times of the jumps of \( N(t) \) are \( 0 = S(0) < S(1) < \cdots \). It is well-known that

\[\{nF_n(t), 0 \leq t \leq 1\} \overset{D}{=} \{N(tS(n + 1)), 0 \leq t \leq 1\}\]

and therefore

\[
\left\{ \sum_{1 \leq i \leq n} e_i I\{X_i \leq t\}, 0 \leq t \leq 1 \right\} \overset{D}{=} \{U(N(tS(n + 1))), 0 \leq t \leq 1\}, \tag{3.8}
\]

where

\[U(k) = \sum_{1 \leq i \leq k} e_i.\]

Let \( z(i, n) = i/S(n + 1), 0 \leq i \leq n, \) and \( a(n) = [c \log \lfloor n^{1/3} \rfloor]/S(n + 1) \), where \([x]\) denotes the integer part of \( x \). By (3.8) there are random variables \( \tau(i, n), 0 \leq i \leq n, \) and \( h(n) \) such that

\[
\left\{ \sum_{1 \leq i \leq n} e_i I\{X_i \leq t\}, 0 \leq t \leq 1, \tau(i, n), 0 \leq i \leq n, h(n) \right\}
\]

\[\overset{D}{=} \{U(N(tS(n + 1))), 0 \leq t \leq 1, z(i, n), 0 \leq i \leq n, a(n)\}. \tag{3.9}
\]

It follows from the Erdős and Rényi (1970) law of large numbers that for any \( c > 0 \)

\[
\frac{1}{[c \log \lfloor n^{1/3} \rfloor]} \max_{0 \leq i \leq \lfloor n^{1/3} \rfloor} \left( U(N((z(i, n) + a(n))S(n + 1)) - U(N(z(i, n))S(n + 1)) \right)
\]

\[= \frac{1}{[c \log \lfloor n^{1/3} \rfloor]} \max_{0 \leq i \leq \lfloor n^{1/3} \rfloor} \left( U(N(i + [c \log \lfloor n^{1/3} \rfloor])) - U(N(i)) \right)
\]

\[\overset{a.s.}{\to} x^*(c) \quad \text{as} \quad n \to \infty. \tag{3.10}
\]
Erdős and Rényi (1970) also showed that there is a one-to-one correspondence between the function \( z^*(c) \) and the moment generating function of \( U(N(1)) \). Putting together (3.1), (3.5), (3.9) and (3.10) we get that

\[
\frac{n^{1/2}}{[c \log [n^{1/3}]]} \max_{0 < i < n} (\beta(z(i, n) + h(n)) - \beta_n(z(i, n))) \xrightarrow{p} z^*(c)
\]  

(3.11)

for all \( c > 0 \). Let \( \{B(t), 0 \leq t \leq 1\} \) be a Brownian bridge. We can define a Brownian bridge \( \{\hat{B}(t), 0 \leq t \leq 1\} \) such that

\[
\{B(t), 0 \leq t \leq 1, \tau(i, n), 0 < i < n, h(n)\} \overset{d}{=} \{\hat{B}(t), 0 \leq t \leq 1, z(i, n), 0 < i < n, a(n)\}.
\]

(3.12)

There is a Wiener process \( \{\hat{W}(t), 0 \leq t \leq 1\} \) such that \( \hat{B}(t) = \hat{W}(t) - t \hat{W}(1), 0 \leq t \leq 1 \). Hence

\[
\frac{n^{1/2}}{[c \log [n^{1/3}]]} \max_{0 < i < [n^{1/3}]} (\hat{W}(z(i, n) + a(n)) - \hat{W}(z(i, n)))
\]

\[
= \frac{n^{1/2}}{[c \log [n^{1/3}]]} \max_{0 < i < [n^{1/3}]} \left( \hat{W}(z(i, n) + a(n)) - \hat{W}(z(i, n)) \right) + o_p(1).
\]

(3.13)

By the law of the iterated logarithm we have that \(|S(n + 1) - n|^\alpha = O((n \log \log n)^{1/2})\) and therefore, by Theorem 1.2.1 of Csörgő and Révész (1981),

\[
\frac{n^{1/2}}{[c \log [n^{1/3}]]} \max_{0 < i < [n^{1/3}]} \left( \hat{W}(z(i, n) + a(n)) - \hat{W}(z(i, n)) \right)
\]

\[
= \frac{n^{1/2}}{[c \log [n^{1/3}]]} \max_{0 < i < [n^{1/3}]} \left( \hat{W} \left( \frac{i}{n} + \frac{[c \log [n^{1/3}]]}{n} \right) - \hat{W} \left( \frac{i}{n} \right) \right) + o_p(1).
\]

(3.14)

The scale transformation of \( \hat{W} \) yields that

\[
\frac{n^{1/2}}{[c \log [n^{1/3}]]} \max_{0 < i < [n^{1/3}]} \left( \hat{W} \left( \frac{i}{n} + \frac{[c \log [n^{1/3}]]}{n} \right) - \hat{W} \left( \frac{i}{n} \right) \right)
\]

\[
= \frac{1}{[c \log [n^{1/3}]]} \max_{0 < i < [n^{1/3}]} (\hat{W}(i + [c \log [n^{1/3}]]) - \hat{W}(i)).
\]

(3.15)

Using again Erdős and Rényi (1970) we obtain that

\[
\frac{1}{[c \log [n^{1/3}]]} \max_{0 < i < [n^{1/3}]} (\hat{W}(i + [c \log [n^{1/3}]]) - \hat{W}(i)) \xrightarrow{a.s.} \left( \frac{2}{c} \right)^{1/2}.
\]

(3.16)

Putting together (3.12)–(3.16) we conclude that for any Brownian bridge \( \{B(t), 0 \leq t \leq 1\} \)

\[
\frac{n^{1/2}}{[c \log [n^{1/3}]]} \max_{0 < i < [n^{1/3}]} (B(\tau(i, n) + h(n)) - B(\tau(i, n))) \xrightarrow{p} \left( \frac{2}{c} \right)^{1/2}
\]

(3.17)

for all \( c > 0 \). If \( m(t) \) denotes the moment generating function of \( \varepsilon_i \), then the moment generating function of \( U(N(1)) \) is \( \exp(m(t) - 1) \). It is easy to see that \( \exp(m(t) - 1) \) cannot be the moment generating function of a normal random variable. Since there is a one-to-one correspondence between the Erdős–Rényi limits and moment generating functions, there is \( c^* > 0 \) such that \( z^*(c^*) \neq (2/c^*)^{1/2} \). If \( F(t) = t \), then Theorem 1.4 follows from (3.11) and (3.17) with some \( c^* \geq c^*[z^*(c^*) - (2/c^*)^{1/2}] \). Since \( F(t) \) is continuous, the case of uniformly distributed \( X_1, X_2, \ldots \) implies the general result. \( \square \)
Proof of Theorem 1.5. By Theorem 2.2 in Horváth (2000) there is a two-time parameter Wiener process \( \{W(t, x), 0 \leq t, x < \infty\} \) such that

\[
\sup_{-\infty < t < \infty} \left| \sum_{1 \leq i \leq n} \varepsilon_i I\{X_i \leq t\} - W(F(t), n) \right|^2 = O(n^{1.4} (\log n)^{1/2}).
\]

(3.18)

By (3.1) and (3.5) we have

\[
\sup_{-\infty < t < \infty} \left| \sum_{1 \leq i \leq n} \varepsilon_i I\{X_i \leq t\} - F(t) \sum_{1 \leq i \leq n} \varepsilon_i \right|^2 = O(\log n),
\]

and therefore (3.18) implies

\[
\sup_{-\infty < t < \infty} \left| n^{1/2} \beta_n(t) - (W(F(t), n) - F(t)W(1, n)) \right|^2 = O(n^{1.4} (\log n)^{1/2}).
\]

Observing that \( K(t, x) = W(t, x) - tW(1, x) \) is a Kiefer process, the proof of Theorem 1.5 is complete. \( \square \)

Proof of Theorem 2.1. If \( H_0 \) holds, then we can write

\[
r(t, n) - r(t, k) = r^*(t, k) + \frac{\varepsilon^*_k}{n} \sum_{k < i \leq n} (I\{Y_i \leq t\} - H(t)) + \varepsilon \sum_{1 \leq i \leq k} (I\{Y_i \leq t\} - H(t)).
\]

By the law of the iterated logarithm we have

\[
\max_{1 \leq k < n} \left| \varepsilon^*_k \sum_{k < i \leq n} (I\{Y_i \leq t\} - H(t)) \right| \leq \max_{1 \leq k < n} \frac{1}{(n - k)^{1/2}} \left| \sum_{k < i \leq n} \varepsilon_i \right| \max_{1 \leq k < n} \frac{1}{(n - k)^{1/2}}
\]

\[
\times \sup_{-\infty < t < \infty} \left| \sum_{k < i \leq n} (I\{Y_i \leq t\} - H(t)) \right| = O_p(\log \log n)
\]

and similarly

\[
\max_{1 \leq k < n} \left| \varepsilon_k \sum_{1 \leq i \leq k} (I\{Y_i \leq t\} - H(t)) \right| = O_p(\log \log n).
\]

The central limit theorem and the weak convergence of the empirical process imply

\[
\sup_{-\infty < t < \infty} \left| \varepsilon_n \sum_{1 \leq i \leq n} (I\{Y_i \leq t\} - H(t)) \right| = O_p(1).
\]

Hence by Theorem 1.5 there is a Kiefer process \( \{K^*(t, x), 0 \leq t \leq 1, 0 \leq x < \infty\} \) such that

\[
n^{-1/2} \max_{1 \leq k < n} \left| (n - k) r(t, k) - k r^*(t, k) - k (K^*(H(t), k) - K^*(H(t), n) - K^*(H(t), k)) \right| = O_p(n^{-1/4} (\log n)^{1/2}).
\]

Let

\[
\Gamma_H(t, x) = n^{-1/2} \{K^*(H(t), nx) - xK^*(H(t), n)\}.
\]
If it is easy to check that \( E \mathcal{H}_{t; x} = 0 \), \( E \mathcal{H}_{t; x} \mathcal{H}_{s; y} = (H(t, s) - H(t)H(s))(\min(x, y) - xy) \). Also, the modulus of continuity of Kiefer processes (cf. Csörgő and Révesz, 1981) gives
\[
\max_{0 \leq k \leq n} \sup_{|x| \leq 1/n} \sup_{-\infty < t < \infty} \left| \mathcal{H}_{t; x} \right| = O_p(n^{-1/2}(\log n)^{1/2}),
\]
which completes the proof of (2.3). To prove (2.4) first we note that by Theorem 1.5
\[
n^{-3/2} \max_{1 \leq k \leq k^*} \sup_{-\infty < t < \infty} |(n - k)r(t, k)| = O_p(1).
\]
Next we write for all \( 1 \leq k \leq k^* \)
\[
r^*(t, k) = \sum_{k < i \leq k^*} \varepsilon_i I\{Y_i \leq t\} + \sum_{k^* < i \leq n} \varepsilon_i I\{Y_i \leq t\} - \varepsilon_k^* \sum_{k < i \leq n} I\{Y_i \leq t\}.
\]
Since \( Y_1, \ldots, Y_{k^*} \) are identically distributed, (3.18) yields
\[
\max_{1 \leq k \leq k^*} \sup_{-\infty < t < \infty} \left| \sum_{k < i \leq k^*} \varepsilon_i I\{Y_i \leq t\} \right| = O_p(n^{1/2})
\]
and similarly the weak convergence of partial sums gives
\[
\max_{1 \leq k \leq k^*} \left| \varepsilon_k^* \sum_{k < i \leq n} I\{Y_i \leq t\} \right| \leq \max_{1 \leq k \leq k^*} \left| \sum_{k < i \leq n} \varepsilon_i \right| = O_p(n^{1/2}).
\]
Thus we get
\[
n^{-3/2} \max_{1 \leq k \leq k^*} \sup_{-\infty < t < \infty} |(n - k)r(t, k) - kr^*(t, k)| = O_p(1)
\]
which completes the proof of (2.4).

References