

Approximations for weighted bootstrap processes with an application[☆]

Lajos Horváth^{a,*}, Piotr Kokoszka^b, Josef Steinebach^c

^a *Department of Mathematics, University of Utah, Salt Lake City, UT 84112, USA*

^b *Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 3BX, UK*

^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. © 2000 Elsevier Science B.V. All rights reserved

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, \dots be a sequence of independent, identically distributed random variables with common distribution function F . The empirical distribution of X_1, X_2, \dots, 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 $\alpha_n(t) = n^{1/2}(F_n(t) - F(t))$ 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 $\alpha_n(t)$ 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):

[☆] Partially supported by a NATO grant (CRG 960 503).

* Corresponding author. Tel.: +1-801-581-8159; fax: +1-801-581-4148.

E-mail address: horvath@math.utah.edu (L. Horváth)

Theorem 1.1. *There exists a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that*

$$P \left\{ \sup_{-\infty < t < \infty} |\alpha_n(t) - B_n(F(t))| > n^{-1/2}(c_1 \log n + x) \right\} \leq c_2 \exp(-c_3 x)$$

for all $x \geq 0$, 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*

$$\lim_{n \rightarrow \infty} P \left\{ \sup_{-\infty < t < \infty} |\alpha_n(t) - B_n^*(F(t))| \geq \frac{1}{6} n^{-1/2} \log n \right\} = 1.$$

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 $\alpha_n(t)$ 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) uses exponential weights. Lo (1987) obtained approximations for $\beta_n(t)$ when ε_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 Bertail (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 $\varepsilon_1, \varepsilon_2, \dots$ be a sequence of independent, identically distributed random variables with

$$E\varepsilon_1 = \mu, \quad \text{var } \varepsilon_1 = 1 \tag{1.1}$$

and

$$E \exp(t\varepsilon_1) < \infty \quad \text{if } |t| \leq t_0 \quad \text{with some } t_0 > 0. \tag{1.2}$$

We also assume that

$$\{X_i, 1 \leq i < \infty\} \quad \text{and} \quad \{\varepsilon_i, 1 \leq i < \infty\} \quad \text{are independent.} \tag{1.3}$$

The smooth bootstrap of α_n is

$$\beta_n(t) = n^{-1/2} \sum_{1 \leq i \leq n} (\varepsilon_i - \bar{\varepsilon}_n) I\{X_i \leq t\}, \quad -\infty < t < \infty,$$

where

$$\bar{\varepsilon}_n = \frac{1}{n} \sum_{1 \leq i \leq n} \varepsilon_i.$$

We note that $\beta_n(t)$ does not depend on μ .

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*

$$P \left\{ \sup_{-\infty < t < \infty} |\beta_n(t) - \hat{B}_n(F(t))| > n^{-1/2}(c_4 \log n + x) \right\} \leq c_5 \exp(-c_6 x)$$

for all $x \geq 0$, 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*

$$\lim_{n \rightarrow \infty} P \left\{ \sup_{-\infty < t < \infty} |\beta_n(t) - B_n^*(F(t))| \geq c_7 n^{-1/2} \log n \right\} = 1$$

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*

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

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

$$H_0: H_{(1)}(t) = H_{(2)}(t) = \dots = H_{(n)}(t) \text{ for all } t$$

against the alternative

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

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

$$H_k(t) = \frac{1}{k} \sum_{1 \leq i \leq k} I\{Y_i \leq t\} \quad \text{and} \quad H_k^*(t) = \frac{1}{n-k} \sum_{k < i \leq n} I\{Y_i \leq t\}.$$

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

$$R_n \xrightarrow{\mathcal{L}} \sup_{0 \leq x \leq 1} \sup_{-\infty < t < \infty} |\Gamma_H(t, x)|, \tag{2.1}$$

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 distribution function of the limiting random variable is unknown. We note that the distribution of $\sup_{0 \leq x \leq 1} \sup_{-\infty < t < \infty} |\Gamma_H(t, x)|$ does not depend on H , if H is continuous, but even in this case it is unknown. The Glivenko–Cantelli lemma yields that if H_a holds and $k^* = [n\theta]$ with some $0 < \theta < 1$, then

$$\frac{k^*(n-k^*)}{n^2} |H_{k^*}(t_0) - H_{k^*}^*(t_0)| \xrightarrow{P} \theta(1-\theta) |H_{(k^*)}(t_0) - H_{(k^*+1)}(t_0)|,$$

and therefore

$$R_n \xrightarrow{P} \infty. \tag{2.2}$$

We use the smooth bootstrap to approximate the distribution of R_n . Let $\varepsilon_1, \varepsilon_2, \dots$ be independent identically distributed random variables satisfying (1.1) and (1.2). Define

$$r(t, k) = \sum_{1 \leq i \leq k} (\varepsilon_i - \bar{\varepsilon}_k) I\{Y_i \leq t\}$$

and

$$r^*(t, k) = \sum_{k < i \leq n} (\varepsilon_i - \bar{\varepsilon}_k^*) I\{Y_i \leq t\}$$

with

$$\bar{\varepsilon}_k = \frac{1}{k} \sum_{1 \leq i \leq k} \varepsilon_i \quad \text{and} \quad \bar{\varepsilon}_k^* = \frac{1}{n-k} \sum_{k < i \leq n} \varepsilon_i.$$

We show that R_n and

$$R_n^* = n^{-3/2} \max_{1 \leq k < n} \sup_{-\infty < t < \infty} |(n-k)r(t, k) - kr^*(t, k)|$$

have the same limit distributions.

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

(i) *If H_0 holds, then*

$$R_n^* \xrightarrow{\mathcal{L}} \sup_{0 \leq x \leq 1} \sup_{-\infty < t < \infty} |\Gamma_H(t, x)|. \tag{2.3}$$

(ii) *If H_a holds, then*

$$R_n^* = O_P(1). \tag{2.4}$$

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 , $n = 25$

Distribution H	Statistic	90.0%	95.0%	97.5%	99.0%
Gaussian	R_n	0.66	0.72	0.76	0.83
	R_n^*	0.60	0.66	0.71	0.78
Uniform	R_n	0.66	0.72	0.79	0.84
	R_n^*	0.61	0.68	0.73	0.79
Cauchy	R_n	0.66	0.72	0.76	0.84
	R_n^*	0.60	0.66	0.72	0.77

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

Distribution H	Statistic	90.0%	95.0%	97.5%	99.0%
Gaussian	R_n	0.69	0.75	0.81	0.86
	R_n^*	0.66	0.72	0.79	0.85
Uniform	R_n	0.69	0.74	0.80	0.87
	R_n^*	0.65	0.71	0.75	0.80
Cauchy	R_n	0.69	0.74	0.81	0.87
	R_n^*	0.65	0.71	0.77	0.83

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 ε_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 ε_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 ε_i in the definition of R_n (standard normal ε_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

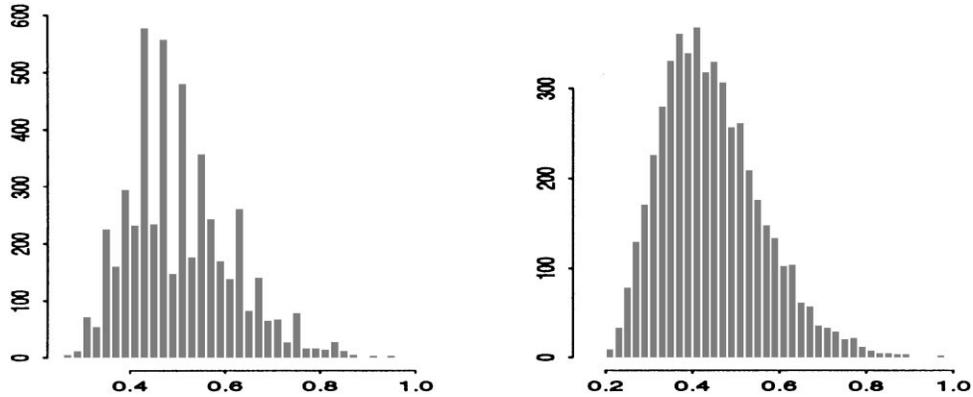


Fig. 1. Histograms of R_n (left) and R_n^* under H_0 , $n = 25$.

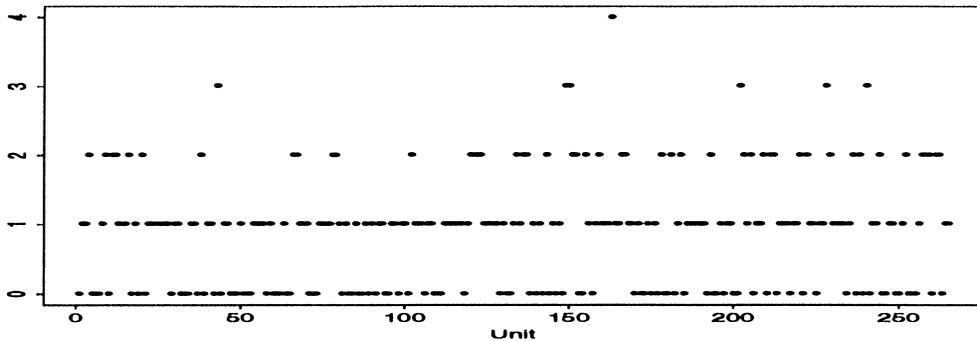


Fig. 2. Number of trisomic abortions per unit.

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 $R_{n,1}^*, \dots, R_{n,N}^*$ 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.

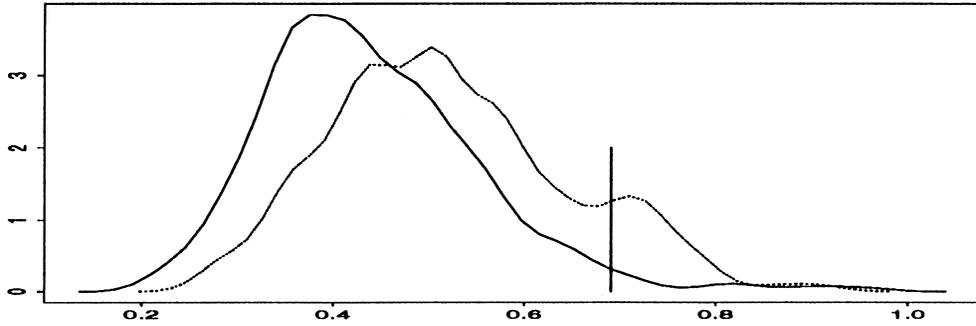


Fig. 3. Analysis of the abortion data.

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

$$\sum_{1 \leq i \leq n} (\varepsilon_i - \bar{\varepsilon}_n) I\{X_i \leq t\} = \sum_{1 \leq i \leq n} \varepsilon_i I\{X_i \leq t\} - F(t) \sum_{1 \leq i \leq n} \varepsilon_i + (F(t) - F_n(t)) \sum_{1 \leq i \leq n} \varepsilon_i. \tag{3.1}$$

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

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

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

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

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

$$P \left\{ \left| \sum_{1 \leq i \leq n} \varepsilon_i \right| \geq x \right\} \leq c_{12} \exp(-c_{13}x), \tag{3.4}$$

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 $\sup_t |F_n(t) - F(t)| \leq 1$ to obtain

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

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

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

for all $x \geq 0$, and therefore

$$\begin{aligned} P\left\{\sup_{-\infty < t < \infty} \left| n^{-1/2} \left(\sum_{1 \leq i \leq n} \varepsilon_i I\{X_i \leq t\} - F(t) \sum_{1 \leq i \leq n} \varepsilon_i \right) - (W_n(F(t)) - F(t)W_n(1)) \right| \right. \\ \left. > 2n^{-1/2}(c_{16} \log n + x) \right\} \leq c_{17} \exp(-c_{18}x). \end{aligned} \quad (3.7)$$

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 $\{\varepsilon_i, 1 \leq i < \infty\}$. The times of the jumps of $N(t)$ are $0 = S(0) < S(1) < \dots$. It is well-known that

$$\{nF_n(t), 0 \leq t \leq 1\} \stackrel{\mathcal{D}}{=} \{N(tS(n+1)), 0 \leq t \leq 1\}$$

and therefore

$$\left\{ \sum_{1 \leq i \leq n} \varepsilon_i I\{X_i \leq t\}, 0 \leq t \leq 1 \right\} \stackrel{\mathcal{D}}{=} \{U(N(tS(n+1))), 0 \leq t \leq 1\}, \quad (3.8)$$

where

$$U(k) = \sum_{1 \leq i \leq k} \varepsilon_i.$$

Let $z(i, n) = i/S(n+1)$, $0 \leq i \leq n$, and $a(n) = [c \log[n^{1/3}]]/S(n+1)$, 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

$$\left\{ \sum_{1 \leq i \leq n} \varepsilon_i I\{X_i \leq t\}, 0 \leq t \leq 1, \tau(i, n), 0 \leq i \leq n, h(n) \right\} \stackrel{\mathcal{D}}{=} \{U(N(tS(n+1))), 0 \leq t \leq 1, z(i, n), 0 \leq i \leq n, a(n)\}. \quad (3.9)$$

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

$$\begin{aligned} & \frac{1}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} (U(N((z(i, n) + a(n))S(n+1))) - U(N(z(i, n)S(n+1)))) \\ &= \frac{1}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} (U(N(i + [c \log[n^{1/3}]])) - U(N(i))) \\ & \xrightarrow{\text{a.s.}} \alpha^*(c) \quad \text{as } n \rightarrow \infty. \end{aligned} \quad (3.10)$$

Erdős and Rényi (1970) also showed that there is a one-to-one correspondence between the function $\alpha^*(c)$ and the moment generating function of $U(N(1))$. Putting together (3.1), (3.5), (3.9) and (3.10) we get that

$$\frac{n^{1/2}}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} (\beta_n(\tau(i, n) + h(n)) - \beta_n(\tau(i, n))) \xrightarrow{P} \alpha^*(c) \tag{3.11}$$

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

$$\{B(t), 0 \leq t \leq 1, \tau(i, n), 0 \leq i \leq n, h(n)\} \stackrel{\mathcal{D}}{=} \{\hat{B}(t), 0 \leq t \leq 1, z(i, n), 0 \leq i \leq n, a(n)\}. \tag{3.12}$$

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

$$\begin{aligned} & \frac{n^{1/2}}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} (\hat{B}(z(i, n) + a(n)) - \hat{B}(z(i, n))) \\ &= \frac{n^{1/2}}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} (\hat{W}(z(i, n) + a(n)) - \hat{W}(z(i, n))) + o_p(1). \end{aligned} \tag{3.13}$$

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

$$\begin{aligned} & \frac{n^{1/2}}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} (\hat{W}(z(i, n) + a(n)) - \hat{W}(z(i, n))) \\ &= \frac{n^{1/2}}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} \left(\hat{W} \left(\frac{i}{n} + \frac{[c \log[n^{1/3}]]}{n} \right) - \hat{W} \left(\frac{i}{n} \right) \right) + o_p(1). \end{aligned} \tag{3.14}$$

The scale transformation of \hat{W} yields that

$$\begin{aligned} & \frac{n^{1/2}}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} \left(\hat{W} \left(\frac{i}{n} + \frac{[c \log[n^{1/3}]]}{n} \right) - \hat{W} \left(\frac{i}{n} \right) \right) \\ & \stackrel{\mathcal{D}}{=} \frac{1}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} (\hat{W}(i + [c \log[n^{1/3}]]) - \hat{W}(i)). \end{aligned} \tag{3.15}$$

Using again Erdős and Rényi (1970) we obtain that

$$\frac{1}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} (\hat{W}(i + [c \log[n^{1/3}]]) - \hat{W}(i)) \stackrel{\text{a.s.}}{\rightarrow} \left(\frac{2}{c} \right)^{1/2}. \tag{3.16}$$

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

$$\frac{n^{1/2}}{[c \log[n^{1/3}]]} \max_{0 \leq i \leq [n^{1/3}]} (B(\tau(i, n) + h(n)) - B(\tau(i, n))) \xrightarrow{P} \left(\frac{2}{c} \right)^{1/2} \tag{3.17}$$

for all $c > 0$. If $m(t)$ denotes the moment generating function of ε_1 , 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 $\alpha^*(c^*) \neq (2/c^*)^{1/2}$. If $F(t) = t$, then Theorem 1.4 follows from (3.11) and (3.17) with some $c_7 \geq c^* |\alpha^*(c^*) - (2/c^*)^{1/2}|$. Since $F(t)$ is continuous, the case of uniformly distributed X_1, X_2, \dots 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

$$\sup_{-\infty < t < \infty} \left| \sum_{1 \leq i \leq n} \varepsilon_i I\{X_i \leq t\} - W(F(t), n) \right| \stackrel{\text{a.s.}}{=} O(n^{1/4}(\log n)^{1/2}). \quad (3.18)$$

By (3.1) and (3.5) we have

$$\sup_{-\infty < t < \infty} \left| n^{1/2} \beta_n(t) - \left(\sum_{1 \leq i \leq n} \varepsilon_i I\{X_i \leq t\} - F(t) \sum_{1 \leq i \leq n} \varepsilon_i \right) \right| \stackrel{\text{a.s.}}{=} O(\log n),$$

and therefore (3.18) implies

$$\sup_{-\infty < t < \infty} |n^{1/2} \beta_n(t) - (W(F(t), n) - F(t)W(1, n))| \stackrel{\text{a.s.}}{=} O(n^{1/4}(\log n)^{1/2}).$$

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

$$\begin{aligned} r(t, n) - r(t, k) &= r^*(t, k) + \bar{\varepsilon}_k^* \sum_{k < i \leq n} (I\{Y_i \leq t\} - H(t)) + \bar{\varepsilon} \sum_{1 \leq i \leq k} (I\{Y_i \leq t\} - H(t)) \\ &\quad - \bar{\varepsilon}_n \sum_{1 \leq i \leq n} (I\{Y_i \leq t\} - H(t)). \end{aligned}$$

By the law of the iterated logarithm we have

$$\begin{aligned} \max_{1 \leq k < n} \left| \bar{\varepsilon}_k^* \sum_{k < i \leq n} (I\{Y_i \leq t\} - H(t)) \right| &\leq \max_{1 \leq k < n} \frac{1}{(n-k)^{1/2}} \left| \sum_{k < i \leq n} \varepsilon_i \right| \max_{1 \leq k < n} \frac{1}{(n-k)^{1/2}} \\ &\quad \times \sup_{-\infty < t < \infty} \left| \sum_{k < i \leq n} (I\{Y_i \leq t\} - H(t)) \right| \\ &= O_p(\log \log n) \end{aligned}$$

and similarly

$$\max_{1 \leq k < n} \left| \bar{\varepsilon}_k \sum_{1 \leq i \leq k} (I\{Y_i \leq t\} - H(t)) \right| = O_p(\log \log n).$$

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

$$\sup_{-\infty < t < \infty} \left| \bar{\varepsilon}_n \sum_{1 \leq i \leq n} (I\{Y_i \leq t\} - H(t)) \right| = O_p(1).$$

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

$$\begin{aligned} n^{-3/2} \max_{1 \leq k < n} |(n-k)r(t, k) - kr^*(t, k) - \{(n-k)K^*(H(t), k) - k(K^*(H(t), n) - K^*(H(t), k))\}| \\ = O_p(n^{-1/4}(\log n)^{1/2}). \end{aligned}$$

Let

$$\Gamma_H(t, x) = n^{-1/2} \{K^*(H(t), nx) - xK^*(H(t), n)\}.$$

If is easy to check that $E\Gamma_H(t, x) = 0$, $E\Gamma_H(t, x)\Gamma_H(s, y) = (H(\min(t, s)) - H(t)H(s))(\min(x, y) - xy)$. Also, the modulus of continuity of Kiefer processes (cf. Csörgő and Révész, 1981) gives

$$\max_{0 \leq k \leq n} \sup_{|s| \leq 1/n} \sup_{-\infty < t < \infty} \left| \Gamma_H \left(t, \frac{k}{n} \right) - \Gamma_H \left(t, \frac{k}{n} + s \right) \right| = O_p(n^{-1/2}(\log n)^{1/2}),$$

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 $1 \leq k \leq k^*$

$$r^*(t, k) = \sum_{k < i \leq k^*} \varepsilon_i I\{Y_i \leq t\} + \sum_{k^* < i \leq n} \varepsilon_i I\{Y_i \leq t\} - \bar{\varepsilon}_k^* \sum_{k < i \leq n} I\{Y_i \leq t\}.$$

Since Y_1, \dots, 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

$$\max_{1 \leq k \leq k^*} \left| \bar{\varepsilon}_k^* \sum_{k < i \leq n} I\{Y_i \leq t\} \right| \leq \max_{1 \leq k \leq k^*} \left| \sum_{k < i \leq n} \varepsilon_i \right| = O_p(n^{1/2}).$$

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). \square

References

- Barbe, P., Bertail, P., 1995. The Weighted Bootstrap, Lecture Notes in Statistics. Springer, Heidelberg, 98.
- Burke, M.D., 1997a. Nonparametric model checks for regression using a weighted bootstrap, preprint.
- Burke, M.D., 1997b. Multivariate test-of-fit and uniform confidence bands using a weighted bootstrap, preprint.
- Burke, M.D., 1998. A Gaussian bootstrap approach to estimation and tests. In: Asymptotic Methods in Probability and Statistics. Elsevier, Amsterdam.
- Csörgő, M., Horváth, L., 1997. Limit Theorems in Change-Point Analysis. Wiley, Chichester.
- Csörgő, M., Horváth, L., Kokoszka, P., 1999. Approximation for bootstrapped empirical processes. Proc. Amer. Math. Soc., to appear.
- Csörgő, M., Révész, P., 1981. Strong Approximations in Probability and Statistics. Academic Press, New York.
- Csörgő, S., 1981. Limit behaviour of the empirical characteristic function. Ann. Probab. 9, 130–144.
- Dvoretzky, A., Kiefer, J., Wolfowitz, J., 1956. Asymptotic minimax character of the sample distribution functions and of the classical multinomial estimator. Ann. Math. Statist. 27, 642–669.
- Efron, B., 1979. Bootstrap method: another look at the jackknife. Ann. Stat. 7, 1–26.
- Erdős, P., Rényi, A., 1970. On a new law of large numbers. J. Analyse Math. 23, 103–111.
- Feuerverger, A., Mureika, R.A., 1977. The empirical characteristic function and its applications. Ann. Statist. 5, 88–97.
- Horváth, L., 2000. Approximations for hybrids of empirical and partial sums processes. J. Stat. Planning Inf., to appear.
- Horváth, L., Kokoszka, P., 1998. Change-point detection with non-parametric regression, preprint.
- Horváth, L., Steinebach, J., 1999. On the best approximation for bootstrapped empirical processes. Statist. Probab. Lett. 41, 117–122.
- Komlós, J., Major, P., Tusnády, G., 1975. An approximation of partial sums of independent R.V.'s and the sample DF. I.Z. Wahrsch. Verw. Gebiete. 32 111–131.
- Levin, B., Kline, J., 1985. The CUSUM test of homogeneity with an application in spontaneous abortion epidemiology. Statist. Med. 4, 469–488.
- Lo, A.Y., 1987. A large sample study of the Bayesian bootstrap. Ann. Stat. 15, 360–375.

- Parzen, M.I., Wei, L.J., Ying, Z., 1994. A resampling method based on pivotal estimating functions. *Biometrika* 81, 341–350.
- Petrov, V.V., 1995. *Limit Theorems of Probability Theory*, Clarendon Press, Oxford.
- Rubin, D.B., 1981. The Bayesian bootstrap. *Ann. Stat.* 9, 130–134.
- Shao, J., Tu, D., 1996. *The Jackknife and the Bootstrap*. Springer, New York.