

APPROXIMATION FOR BOOTSTRAPPED EMPIRICAL PROCESSES

MIKLÓS CSÖRGŐ, LAJOS HORVÁTH, AND PIOTR KOKOSZKA

(Communicated by Wei Y. Loh)

In memory of Béla Szőkefalvi-Nagy

ABSTRACT. We obtain an approximation for the bootstrapped empirical process with the rate of the Komlós, Major and Tusnády approximation for empirical processes. The proof of the new approximation is based on the Poisson approximation for the uniform empirical distribution function and the Gaussian approximation for randomly stopped sums.

1. APPROXIMATIONS FOR BOOTSTRAPPED EMPIRICAL PROCESSES

Let X_1, X_2, \dots be independent, identically distributed random variables with distribution function F . The empirical distribution function of X_1, \dots, X_n is

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

Given the random sample, X_1, \dots, X_n , let X_1^*, \dots, X_m^* be conditionally independent random variables with common distribution function $F_n(t)$. Let

$$F_{m,n}(t) = \frac{1}{m} \sum_{1 \leq i \leq m} I\{X_i^* \leq t\}, \quad -\infty < t < \infty,$$

denote the empirical distribution function of the bootstrapped sample X_1^*, \dots, X_m^* . We say that $F_{m,n}(t)$ is the bootstrapped empirical distribution function. The empirical process of X_1, \dots, X_n is

$$\alpha_n(t) = n^{1/2}(F_n(t) - F(t)), \quad -\infty < t < \infty,$$

and the corresponding bootstrapped empirical process is

$$\alpha_{m,n}(t) = n^{1/2}(F_{m,n}(t) - F_n(t)), \quad -\infty < t < \infty.$$

The bootstrap is a widely used tool in statistics and, therefore, the properties of $\alpha_{m,n}(t)$ are of great interest in applied as well as in theoretical statistics. For the asymptotic properties of $\alpha_{m,n}(t)$ we refer to Shorack and Wellner [15]. In this note we are interested in studying the rate of convergence of $\alpha_{m,n}(t)$ to its

Received by the editors September 16, 1998.

1991 *Mathematics Subject Classification*. Primary 62G30; Secondary 62G07.

Key words and phrases. Bootstrap, empirical process, Poisson process, Brownian bridge, approximations.

The first author's research was supported by an NSERC Canada grant.

The second author's research was supported by NATO grant CRG 960 503.

limit in *sup* norm, as well as in that of the difference between the distributions of smooth functionals of $\alpha_n(t)$ and $\alpha_{m,n}(t)$. The first approximation for $\alpha_{m,n}(t)$ à la Komlós, Major and Tusnády [10, 11] was obtained by Csörgő, Csörgő and Horváth [1, Section 17], and it reads as follows.

Theorem 1.1. *If*

$$(1.1) \quad 0 < \liminf m/n \leq \limsup m/n < \infty,$$

then we can define Brownian bridges $\{B_{m,n}^(t), 0 \leq t \leq 1\}$ such that*

$$(1.2) \quad P\left\{ \sup_{-\infty < t < \infty} |\alpha_{m,n}(t) - B_{m,n}^*(F(t))| > A_1 n^{-1/4} (\log n)^{3/4} \right\} \leq A_2 n^{-\epsilon},$$

for all $\epsilon > 0$, where $A_1 = A_1(\epsilon)$ and A_2 is an absolute constant.

For a bootstrapped parallel to the asymptotic theory of weighted empirical processes in Csörgő, Csörgő, Horváth and Mason [2], we refer to Csörgő and Mason [6].

The rate of convergence in (1.2) is much weaker than the optimal $n^{-1/2} \log n$ in the Komlós, Major and Tusnády [10, 11] approximation for $\alpha_n(t)$. Our main result concludes that the rate in (1.2) can be improved via a different method of approximation. Moreover, the rate of this new approximation for $\alpha_{m,n}(t)$ coincides with that of the best possible approximation for $\alpha_n(t)$ by a sequence of Brownian bridges. Namely we prove the following result.

Theorem 1.2. *We can define Brownian bridges $\{B_{m,n}(t), 0 \leq t \leq 1\}$ such that*

$$(1.3) \quad \begin{aligned} P\left\{ \sup_{-\infty < t < \infty} |\alpha_{m,n}(t) - B_{m,n}(F(t))| \geq m^{-1/2}(x + c_1 \log m) \right. \\ \left. + n^{-1/2}(y + c_4 \log n) \right\} \leq c_2 \exp(-c_3 x) + c_5 \exp(-c_6 y), \end{aligned}$$

for all $x, y > 0$, where c_1, c_2, \dots, c_6 are positive absolute constants.

If (1.1) holds and F is continuous, then Horváth and Steinebach [9] showed that there is a positive constant c_7 such that for any Brownian bridges $\{\hat{B}_{m,n}^*(t), 0 \leq t \leq 1\}$ we have

$$\lim_{n \rightarrow \infty} P\left\{ \sup_{-\infty < t < \infty} |\alpha_{m,n}(t) - \hat{B}_{m,n}^*(t)| > c_7 n^{-1/2} \log n \right\} = 1.$$

Thus we have the optimality of the approximation in (1.3) if n and m are proportional.

One of the immediate consequences of Theorem 1.2 is an upper bound for the convergence of distributions of smooth functionals of $\alpha_{m,n}(t)$.

Corollary 1.1. *If g is a Lipschitz functional of order 1 and $g(B(F(\cdot)))$ has a bounded density function, then*

$$(1.4) \quad \begin{aligned} \sup_{-\infty < x < \infty} |P\{g(\alpha_{m,n}(\cdot)) \leq x\} - P\{g(B(F(\cdot))) \leq x\}| \\ = O(n^{-1/2} \log n + m^{-1/2} \log m), \end{aligned}$$

$$(1.5) \quad \begin{aligned} \sup_{-\infty < x < \infty} |P\{g(\alpha_n(\cdot)) \leq x\} - P\{g(B(F(\cdot))) \leq x\}| \\ = O(n^{-1/2} \log n), \end{aligned}$$

and

$$(1.6) \quad \sup_{-\infty < x < \infty} |P\{g(\alpha_{m,n}(\cdot)) \leq x\} - P\{g(\alpha_n(\cdot)) \leq x\}| \\ = O(n^{-1/2} \log n + m^{-1/2} \log m),$$

where $\{B(t), 0 \leq t \leq 1\}$ stands for a Brownian bridge.

We note that (1.5) follows immediately from Komlós, Major and Tusnády [10, 11] and, using the same argument, we can derive (1.4) from Theorem 1.2 (cf. Csörgő and Révész [5]). Putting together (1.4) and (1.5), we get (1.6).

In general, whenever the Komlós, Major and Tusnády [10, 11] approximation was used in the proofs of limit theorems to derive weak convergence or rates for various weak approximations, we can now replace their approximation for $\alpha_n(t)$ with that of Theorem 1.2, and thus obtain the same results in the context of bootstrapped processes as well. For example we can use Theorem 1.2 to obtain limit results for bootstrapped density estimates. We note in passing that using (1.3) one can correct an error in the proof of Theorem 2.3 in Hall [8], where the best choice for the smoothing parameter is discussed in case of bootstrapped density estimators.

2. PROOF OF THEOREM 1.2

We can and shall assume without loss of generality that all random variables and processes introduced so far and later on in this paper can be defined on the same probability space (cf. Section A in Csörgő and Horváth [4]). It suffices (cf. Shorack and Wellner [15, p. 9] or Roussas [13, p. 244]) to consider the case when the observations are uniformly distributed on $[0, 1]$. Let U_1, U_2, \dots be independent, identically distributed random variables, uniform on $[0, 1]$. The empirical distribution function of U_1, \dots, U_n is

$$E_n(t) = \frac{1}{n} \sum_{1 \leq i \leq n} I\{U_i \leq t\}, \quad 0 \leq t \leq 1.$$

Given U_1, \dots, U_n , let $U_1^*, U_2^*, \dots, U_m^*$ be conditionally independent with common distribution function $E_n(t)$. The bootstrapped uniform empirical distribution function is

$$E_{m,n}(t) = \frac{1}{m} \sum_{1 \leq i \leq m} I\{U_i^* \leq t\}, \quad 0 \leq t \leq 1,$$

and the bootstrapped uniform empirical process is

$$e_{m,n}(t) = m^{1/2}(E_{m,n}(t) - E_n(t)), \quad 0 \leq t \leq 1.$$

Let $\{\xi_i, 1 \leq i < \infty\}$ and $\{\eta_i, 1 \leq i < \infty\}$ be two independent sequences of independent, identically distributed random variables, uniform on $[0, 1]$. The corresponding uniform empirical distribution functions are

$$L_n(t) = \frac{1}{n} \sum_{1 \leq i \leq n} I\{\xi_i \leq t\}, \quad 0 \leq t \leq 1,$$

and

$$V_m(t) = \frac{1}{m} \sum_{1 \leq i \leq m} I\{\eta_i \leq t\}, \quad 0 \leq t \leq 1.$$

Let

$$v_m(t) = m^{1/2}(V_m(t) - t), \quad 0 \leq t \leq 1.$$

Lemma 2.1. *For each n and m we have*

$$(2.1) \quad \{e_{m,n}(t), 0 \leq t \leq 1\} \stackrel{\mathcal{D}}{=} \{v_m(L_n(t)), 0 \leq t \leq 1\}.$$

Proof. We refer to Shorack [14].

In light of (2.1), it suffices to approximate $v_m(L_n(t))$.

Lemma 2.2. *We can define a sequence of Brownian bridges $\{\tilde{B}_m(t), 0 \leq t \leq 1\}$ such that*

$$(2.2) \quad \{\tilde{B}_m(t), 0 \leq t \leq 1, 1 \leq m < \infty\} \text{ and } \{\xi_i, 1 \leq i < \infty\}$$

are independent and

$$(2.3) \quad P\left\{ \sup_{0 \leq t \leq 1} |v_m(L_n(t)) - \tilde{B}_m(L_n(t))| > m^{-1/2}(x + c_1 \log m) \right\} \\ \leq c_2 \exp(-c_3 x)$$

for all $x > 0$.

Proof. Since $\{\xi_i, 1 \leq i < \infty\}$ and $\{\eta_i, 1 \leq i < \infty\}$ are independent, the Brownian bridges $\{\tilde{B}_m(t), 0 \leq t \leq 1\}$ that are constructed via Komlós, Major and Tusnády [10, 11] for $v_m(t)$ are such that (2.2) holds, and

$$(2.4) \quad P\left\{ \sup_{0 \leq t \leq 1} |v_m(t) - \tilde{B}_m(t)| > m^{-1/2}(x + c_1 \log m) \right\} \\ \leq c_2 \exp(-c_3 x)$$

for all $x > 0$. Observing that $0 \leq L_n(t) \leq 1$, (2.3) implies (2.4).

Lemma 2.3. *We can define a sequence of Poisson processes $\{N_n(x), 0 \leq x < \infty\}$ with $EN_n(x) = x$ such that*

$$(2.5) \quad \{\tilde{B}_m(t), 0 \leq t \leq 1, 1 \leq m < \infty\} \text{ and } \{N_n(x), 0 \leq x < \infty, 1 \leq n < \infty\}$$

are independent and

$$(2.6) \quad P\left\{ \sup_{0 \leq t \leq 1} |L_n(t) - N_n(nt)/N_n(n)| > \frac{1}{n}(x + c_{11} \log n) \right\} \\ \leq c_{12} \exp(-c_{13} x)$$

for all $x > 0$.

Proof. By Theorem 3.1.3 of Csörgő and Horváth [4, p. 139] and (2.2) we can find a sequence of Poisson processes $\{N_n(x), 0 < x < \infty\}$ such that (2.5) holds and

$$P\left\{ \sup_{0 \leq t \leq 1} \left| L_n(t) - t - \frac{1}{n}(N_n(nt) - tN_n(n)) \right| > \frac{1}{n}(x + c_{14} \log n) \right\} \\ \leq c_{15} \exp(-c_{16} x)$$

for all $x > 0$. Hence by the Dvoretzky, Kiefer and Wolfowitz inequality we have

$$(2.7) \quad P\left\{ \sup_{0 \leq t \leq 1} |N_n(nt) - tN_n(n)| > (nx)^{1/2} + x + c_{14} \log n \right\} \leq c_{17} \exp(-c_{18}x)$$

for all $x > 0$. By Theorem 2.6 of Petrov [12, p. 55] we conclude

$$(2.8) \quad P\{|N_n(n) - n| > n/2\} \leq \exp(-c_{19}x)$$

for all $0 \leq x \leq n$. Observing that

$$\begin{aligned} & \frac{1}{n} \{N_n(nt) - tN_n(n)\} - \left\{ \frac{N_n(nt)}{N_n(n)} - t \right\} \\ &= \frac{1}{nN_n(n)} \{N_n(nt) - tN_n(n)\} \{N_n(n) - n\}, \end{aligned}$$

Lemma 2.3 is established for $0 \leq x \leq n$. Since (2.6) clearly holds for $x > n$, the proof is complete.

Next we note that for all m and n ,

$$(2.9) \quad \{\tilde{B}_m(N_n(nt)/N_n(n)), 0 \leq t \leq 1\} \stackrel{\mathcal{D}}{=} \{B(N(nt)/N(n)), 0 \leq t \leq 1\},$$

where

$$(2.10) \quad \{B(t), 0 \leq t \leq 1\} \text{ is a Brownian bridge and } \{N(x), 0 \leq x < \infty\} \text{ is a Poisson process with } EN(x) = x \text{ and these two processes are independent of each other.}$$

In light of (2.9), it suffices to approximate $B(N(nt)/N(n))$.

Lemma 2.4. *If (2.10) holds, then we can define a sequence of Brownian bridges $\{\hat{B}_n(t), 0 \leq t \leq 1\}$ such that*

$$(2.11) \quad P\left\{ \sup_{0 \leq t \leq 1} |B(N(nt)/N(n)) - \hat{B}_n(t)| > n^{-1/2}(x + c_{20} \log n) \right\} \leq c_{21} \exp(-c_{22}x)$$

for all $x > 0$.

Proof. First we note that there is a Wiener process $\{W(x), 0 \leq x < \infty\}$, independent of $\{N(x), 0 \leq x < \infty\}$, such that

$$B(t) = W(t) - tW(1), 0 \leq t \leq 1.$$

Using the independence of $\{W(t), 0 \leq t < \infty\}$ and $\{N(x), 0 \leq x < \infty\}$ we get for each n that

$$(2.12) \quad \{B(N(nt)/N(n)), 0 \leq t \leq 1\} \stackrel{\mathcal{D}}{=} \left\{ \left(\frac{1}{N(n)} \right)^{1/2} \left(W(N(nt)) - \frac{N(nt)}{N(n)} W(N(n)) \right), 0 \leq t \leq 1 \right\}.$$

The process $\{W(N(x)), 0 \leq x < \infty\}$ has already appeared in Csörgő, Deheuvels and Horváth [3] (cf. also Chapter 2 in Csörgő and Horváth [4]). Now, as a special

case of Theorem 1.1 in Csörgő, Deheuvels and Horváth [3], we can construct a sequence of Wiener processes $\{W_n^*(t), 0 \leq t \leq 1\}$ such that

$$(2.13) \quad P\left\{\sup_{0 \leq t \leq 1} |n^{-1/2}W(N(nt)) - W_n^*(t)| > n^{-1/2}(x + c_{23} \log n)\right\} \\ \leq c_{24} \exp(-c_{25}x)$$

for all $x > 0$.

Next we show that for all $0 \leq x \leq n$,

$$(2.14) \quad P\left\{\sup_{0 \leq t \leq 1} \left| \frac{N(nt)}{N(n)} - t \right| (N(n))^{-1/2} |W(N(n))| \right. \\ \left. \geq n^{-1/2}(x + c_{26} \log n) \right\} \leq c_{27} \exp(-c_{28}x).$$

By the independence of $\{N(x), 0 \leq x < \infty\}$ and $\{W(x), 0 \leq x < \infty\}$, we get that $(N(n))^{-1/2}W(N(n))$ is a standard normal random variable and, therefore,

$$(2.15) \quad P\{(N(n))^{-1/2}|W(N(n))| > x^{1/2}\} \leq c_{29} \exp(-x/2)$$

for all $x > 0$.

Recalling (2.7) we obtain

$$(2.16) \quad P\left\{\sup_{0 \leq t \leq 1} |N(nt) - tN(n)| > (nx)^{1/2} + x + c_{14} \log n\right\} \\ \leq c_{17} \exp(-c_{18}x)$$

for all $x > 0$, and (2.8) yields that

$$(2.17) \quad P\{|N(n) - n| > n/2\} \leq \exp(-c_{19}x)$$

for all $0 \leq x \leq n$. Now (2.14) follows from (2.15)–(2.17).

Putting together (2.13) and (2.14) we arrive at

$$(2.18) \quad P\left\{\sup_{0 \leq t \leq 1} \left| \left(\frac{n}{N(n)} \right)^{1/2} \left\{ n^{-1/2} (W(N(nt)) - \frac{N(nt)}{N(n)} W(N(n)) \right. \right. \right. \\ \left. \left. \left. - B_n^*(t) \right\} \right| > n^{-1/2}(x + c_{30} \log n) \right\} \\ \leq c_{31} \exp(-c_{32}x),$$

for all $0 \leq x \leq n$, where for each n

$$B_n^*(t) = W_n^*(t) - tW_n^*(1), \quad 0 \leq t \leq 1,$$

is a Brownian bridge.

Next we show that

$$(2.19) \quad P\left\{\sup_{0 \leq t \leq 1} |B_n^*(t)| \left| \left(\frac{n}{N(n)} \right)^{1/2} - 1 \right| > n^{-1/2}(x + c_{33} \log n)\right\} \\ \leq c_{34} \exp(-c_{35}x)$$

for all $0 \leq x \leq n$. First we note that there is a constant c_{36} such that

$$(2.20) \quad P\left\{\sup_{0 \leq t \leq 1} |B_n^*(t)| \geq x^{1/2}\right\} \leq c_{36} \exp(-x/2)$$

for all $x > 0$. Again using Theorem 2.6 of Petrov [12, p. 55] we find constants c_{37} and c_{38} such that

$$(2.21) \quad P\left\{\left|\left(\frac{n}{N(n)}\right)^{1/2} - 1\right| \geq (x/n)^{1/2}\right\} \leq c_{37} \exp(-c_{38}x)$$

for all $0 \leq x \leq n$. The inequality in (2.19) follows from (2.20) and (2.21).

The approximation in (2.11) is an immediate consequence of (2.12), (2.18) and (2.19), if $0 \leq x \leq n$. Next we assume that $x > n$. Since B and \hat{B}_n are Brownian bridges, (2.20) yields

$$P\left\{\sup_{0 \leq t \leq 1} |\hat{B}_n(t)| \geq xn^{-1/2}\right\} \leq c_{36} \exp\left(-\frac{1}{2} \frac{x^2}{n}\right) \leq c_{36} \exp(-x/2)$$

and

$$\begin{aligned} P\left\{\sup_{0 \leq t \leq 1} |B(N(nt)/N(n))| \geq xn^{-1/2}\right\} &\leq P\left\{\sup_{0 \leq t \leq 1} |B(t)| \geq xn^{-1/2}\right\} \\ &\leq c_{36} \exp(-x/2) \end{aligned}$$

as well. This completes the proof of Lemma 2.4.

Proof of Theorem 1.2. It follows from Lemmas 2.3 and 2.4 that there are Brownian bridges $\{\tilde{B}_{m,n}(t), 0 \leq t \leq 1\}$ such that

$$(2.22) \quad \begin{aligned} P\left\{\sup_{0 \leq t \leq 1} |e_{m,n}(t) - \tilde{B}_{m,n}| \geq m^{-1/2}(x + c_1 \log m) \right. \\ \left. + n^{-1/2}(y + c_4 \log n)\right\} \leq c_2 \exp(-c_3x) + c_5 \exp(-c_6y) \end{aligned}$$

for all x, y . This establishes (1.3) in the case of the uniform-[0, 1] random variables. As pointed out at the beginning of this section, this constitutes no loss of generality, and hence (2.22) implies the general case as well, as stated in Theorem 1.2.

REFERENCES

- [1] Csörgő, M., Csörgő, S. and Horváth, L. (1986). *An Asymptotic Theory for Empirical Reliability and Concentration Processes*. Springer-Verlag, Berlin. MR **88i**:62030
- [2] Csörgő, M., Csörgő, S., Horváth, L. and Mason, D. M. (1986). Weighted empirical and quantile processes. *Ann. Probab.* **14** 31–85. MR **87e**:60041
- [3] Csörgő, M., P. Deheuvels, and Horváth, L. (1987). An approximation of stopped sums with applications in queueing theory. *Adv. Appl. Prob.* **19** 674–690. MR **89c**:60045
- [4] Csörgő, M., and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*. Wiley, Chichester. MR **94c**:60060
- [5] Csörgő, M., and Révész, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press, New York. MR **84d**:60050
- [6] Csörgő, S., and Mason, D. M. (1989). Bootstrapping empirical functions. *Ann. Statist.* **17**, 1447–1471 MR **91a**:62044
- [7] Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution functions and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669. MR **18**:772i
- [8] Hall, P. (1991). On convergence rates of suprema. *Z. Wahrschein. Verw. Gebiete* **89** 447–455. MR **92i**:60101

- [9] Horváth, L. and Steinebach, J. (1999). On the best approximation for bootstrapped empirical processes. *Stat. Probab. Letters* **41** 117–122. CMP 99:06
- [10] Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent R.V.'s and the sample DF. I. *Z. Wahrschein. Verw. Gebiete* **32** 111–131. MR **51**:11605b
- [11] Komlós, J., Major, P. and Tusnády, G. (1976). An approximation of partial sums of independent R.V.'s and the sample DF. II. *Z. Wahrschein. Verw. Gebiete* **34** 33–58. MR **53**:6697
- [12] Petrov, V. V. (1995). *Limit Theorems of Probability Theory*. Clarendon Press, Oxford. MR **96h**:60048
- [13] Roussas, G. G. (1997). *A Course in Mathematical Statistics*. Second Ed. Academic Press, San Diego.
- [14] Shorack, G. R. (1982). Bootstrapping robust regression. *Comm. Statist. A- Theory Methods* **11** 961–972 MR **83g**:62033
- [15] Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York. MR **88e**:60002

DEPARTMENT OF MATHEMATICS AND STATISTICS, CARLETON UNIVERSITY, OTTAWA ONTARIO, CANADA K1S 5B6

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF UTAH, 155 SOUTH 1440 EAST, SALT LAKE CITY, UTAH 84112-0090

E-mail address: horvath@math-utah.edu

DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF LIVERPOOL, LIVERPOOL L69 3BX, UNITED KINGDOM