CentER

Discussion Paper

No. 2004–112

GENERAL WEAK LAWS OF LARGE NUMBERS FOR BOOTSTRAP SAMPLE MEANS

By J.H.J. Einmahl, A. Rosalsky

October 2004

ISSN 0924-7815

General Weak Laws of Large Numbers for Bootstrap Sample Means

John H.J. Einmahl — *Tilburg University*

Andrew Rosalsky — *University of Florida*

19th October 2004

Abstract

For bootstrap sample means resulting from a sequence $\{X_n, n \geq 1\}$ of random variables, very general weak laws of large numbers are established. The random variables $\{X_n, n \geq 1\}$ do not need to be independent or identically distributed or to be of any particular dependence structure. In general, no moment conditions are imposed on the $\{X_n, n \geq 1\}$. Examples are provided which illustrate the sharpness of the main results.

AMS 2000 subject classifications. Primary 60F05, 62G09; secondary 62G20.

Key words and phrases. Bootstrap sample mean, weak law of large numbers, convergence in probability, almost certain convergence.

1 Introduction

In this paper very general weak laws of large numbers are obtained for bootstrap sample means. Bootstrap samples were introduced in Efron (1979) for a sequence of independent and identically distributed (i.i.d.) random variables. More generally, bootstrap samples are defined as follows. Consider a sequence of random variables $\{X_n, n \geq 1\}$ (not necessarily independent or identically distributed) defined on a probability space $(\Omega, \mathcal{F}, P)$ and let $\{m(n), n \geq 1\}$ be a sequence of positive integers. For $\omega \in \Omega$ and $n \geq 1$, let the random variables $\{\hat{X}_{nj}^{(\omega)}, 1 \leq j \leq m(n)\}$ result by sampling $m(n)$ times with replacement from the n observations $X_1(\omega), \cdots, X_n(\omega)$ such that for each of the $m(n)$ selections, each $X_i(\omega)$ has probability n^{-1} of being chosen. Hence for $\omega \in \Omega$ and $n \geq 1, \{\hat{X}_{nj}^{(\omega)}, 1 \leq j \leq m(n)\}$ are i.i.d. random variables uniformly distributed over $\{X_1(\omega), \cdots, X_n(\omega)\}$. For $\omega \in \Omega$ and $n \geq 1, \left\{\hat{X}_{nj}^{(\omega)}, 1 \leq j \leq m(n)\right\}$ is the so-called *bootstrap sample* from $X_1(\omega), \cdots, X_n(\omega)$ with *bootstrap sample size* $m(n)$.

The main classes of limit theorems of classical probability for partial sums of i.i.d. random variables have counterparts for bootstrap sample partial sums $\left\{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}, n \geq 1\right\}$. References to these bootstrap counterparts of the various classes of classical limit theorems are listed as follows:

- Weak Law of Large Numbers (Bickel and Freedman (1981), Athreya (1983), Athreya, Ghosh, Low, and Sen (1984), Csörgő (1992), Arenal-Gutiérrez, Matrán, and Cuesta-Albertos (1996a))
- Central Limit Theorem (Singh (1981), Bickel and Freedman (1981), Giné and Zinn (1989), Arcones and Giné (1989), Arenal-Gutiérrez and Matrán (1996))
- Strong Law of Large Numbers (Athreya (1983), Athreya et al. (1984), Mikosch (1994), Arenal-Gutiérrez, Matrán, and Cuesta-Albertos (1996b), Hu and Taylor (1997), Bozorgnia, Patterson, and Taylor (1997), Csörgő and Wu (2000))
- Law of the Iterated Logarithm (Mikosch (1994), Ahmed, Li, Rosalsky, and Volodin (2001))
- Complete Convergence Theorem (Li, Rosalsky, and Ahmed (1999), Csörgő and Wu (2000), Ahmed, Hu, and Volodin (2001), Csörgő (2003), Csörgő (2004))
- Large Deviation Principle (Li, Rosalsky, and Al-Mutairi (2002))
- Erdős-Rényi-Shepp Law (Li and Rosalsky (2002)).

In most of the above references, the $\{X_n, n \geq 1\}$ are assumed to be i.i.d. Discussions comparing the orders of convergence in the classical central limit theorem, strong law of large numbers, complete convergence theorem, and law of the iterated logarithm with the orders of convergence in their bootstrap counterparts, respectively, are given in Li et al. (1999) and in Ahmed, Li et al. (2001). A comprehensive survey of first-order limit laws for bootstrap sums was recently prepared by Csörgő and Rosalsky (2003).

In the current work, the main results, Theorems 1 and 2 will be presented in Section 3 and they establish general Weak Laws of Large Numbers (WLLNs) for bootstrap sample means $\left\{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}/m(n), n \geq 1\right\}$ from a sequence of random variables $\{X_n, n \geq 1\}$. An interesting and unusual feature of these theorems is that it is not assumed that the random variables $\{X_n, n \geq 1\}$ are independent or that they are identically distributed. Furthermore, in general, no moment conditions are imposed on the $\{X_n, n \geq 1\}$. The only other work on limit laws for bootstrap sample means that we are aware of without the assumptions that the $\{X_n, n \geq 1\}$ are identically distributed with a particular dependence structure is that of Li et al. (1999) and Ahmed, Li et al. (2001).

The WLLNs established in Theorems 1 and 2 take the form: for almost every $\omega \in \Omega$,

$$\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} - \frac{\sum_{i=1}^{n} X_i(\omega)}{n} \xrightarrow{P} 0.$$

These theorems differ substantially from the other WLLNs for bootstrap sample partial sums cited above. The WLLNs of Bickel and Freedman (1981), Athreya (1983), Athreya et al. (1984), and Csörgő (1992) assume that the $\{X_n, n \geq 1\}$ are i.i.d. with $E|X_1| < \infty$. The WLLN of Arenal-Gutiérrez et al. (1996a) assumes that the $\{X_n, n \geq 1\}$ are pairwise i.i.d. with $E|X_1| < \infty$.

The preliminaries needed prior to presenting the main results are consolidated into Section 2. In Section 4, examples are provided which illustrate the sharpness of the results.

2 Preliminaries

Some preliminaries are needed prior to presenting Theorems 1 and 2.

LEMMA 1. *Let $\{X_n, n \geq 1\}$ be a sequence of random variables and let $\{b_n, n \geq 1\}$ be a sequence of positive constants. If $b_n \uparrow \infty$, then*

$$\frac{\max_{1 \leq i \leq n} |X_i|}{b_n} \to 0 \quad a.c.$$

if and only if

$$\frac{X_n}{b_n} \to 0 \quad a.c.$$

PROOF. Necessity is obvious. To prove sufficiency, assume that $X_n/b_n \to 0$ a.c. Then for arbitrary $n \geq k \geq 2$,

$$\begin{aligned}
\frac{\max_{1 \leq i \leq n} |X_i|}{b_n} &\leq \frac{\max_{1 \leq i \leq k-1} |X_i|}{b_n} + \frac{\max_{k \leq i \leq n} |X_i|}{b_n} \\
&\leq \frac{\max_{1 \leq i \leq k-1} |X_i|}{b_n} + \max_{k \leq i \leq n} \frac{|X_i|}{b_i} \quad (\text{since } b_n \uparrow) \\
&\leq \frac{\max_{1 \leq i \leq k-1} |X_i|}{b_n} + \sup_{i \geq k} \frac{|X_i|}{b_i} \to 0 \quad \text{a.c.}
\end{aligned}$$

as first $n \to \infty$ and then $k \to \infty$ since $b_n \to \infty$ and $X_n/b_n \to 0$ a.c. □

The following two propositions provide WLLNs for double arrays of rowwise independent random variables. Proposition 1 is well known and may be found in Chow and Teicher (1997, p. 356). In Proposition 2, the random variables in each row of the array are identically distributed, but no such assumption is made concerning the random variables from different rows.

PROPOSITION 1. *Let* $\{Y_{nj}, 1 \leq j \leq m(n) \to \infty\}$ *be an array of rowwise independent random variables, and suppose that*

$$\sum_{j=1}^{m(n)} P\{|Y_{nj}| \geq \varepsilon\} \to 0 \text{ for all } \varepsilon > 0$$

and

$$\sum_{j=1}^{m(n)} Var(Y_{nj} I(|Y_{nj}| < 1)) \to 0.$$

Then the WLLN

$$\sum_{j=1}^{m(n)} Y_{nj} - \sum_{j=1}^{m(n)} E(Y_{nj} I(|Y_{nj}| < 1)) \xrightarrow{P} 0$$

obtains.

In Proposition 1, replacing Y_{nj} by $Y_{nj}/m(n), 1 \leq j \leq m(n), n \geq 1$ yields the following special case.

PROPOSITION 2. *Let* $\{Y_{nj}, 1 \leq j \leq m(n) \to \infty\}$ *be an array of rowwise i.i.d. random variables, and suppose that*

$$m(n) P\{|Y_{n1}| \geq \varepsilon m(n)\} \to 0 \text{ for all } \varepsilon > 0$$

and

$$\frac{Var(Y_{n1} I(|Y_{n1}| < m(n)))}{m(n)} \to 0.$$

Then the WLLN

$$\frac{\sum_{j=1}^{m(n)} Y_{nj}}{m(n)} - E\left(Y_{n1} I(|Y_{n1}| < m(n))\right) \xrightarrow{P} 0$$

obtains.

3 Mainstream

With the preliminaries accounted for, the theorems may now be established. Let $\{X_n, n \geq 1\}$ be a sequence of random variables and let $\{m(n), n \geq 1\}$ be a sequence of positive integers with $m(n) \uparrow \infty$. We will use

$$(1) \qquad \frac{X_n}{\sqrt{m(n)}} \to 0 \quad \text{almost certainly (a.c.).}$$

A sufficient condition for (1) is of course that

$$(2) \qquad \sum_{n=1}^{\infty} P\{|X_n| > \varepsilon\sqrt{m(n)}\} < \infty \text{ for all } \varepsilon > 0.$$

The conditions (1) and (2) are equivalent if the random variables $\{X_n, n \geq 1\}$ are pairwise independent, since the divergence half of the Borel-Cantelli lemma holds for a sequence of pairwise independent events (see Chung (1974, p. 76) or for an elegant and simpler proof based on the Schwarz inequality see Etemadi (1984)). No moment conditions in general are imposed on the random variables $\{X_n, n \geq 1\}$. However, if the $\{X_n, n \geq 1\}$ are pairwise i.i.d., then the assumption (1) is equivalent to

$$(3) \qquad \sum_{n=1}^{\infty} P\{|X_1| > \varepsilon\sqrt{m(n)}\} < \infty \text{ for all } \varepsilon > 0$$

which is in effect a moment type condition on X_1. If the $\{X_n, n \geq 1\}$ are identically distributed $\mathcal{L}_p$ random variables for some $p > 0$ (but not necessarily pairwise independent) and if $n^{2/p} = O(m(n))$, then (3) holds and hence so does (1). Moreover, we are *not* assuming that for almost every $\omega \in \Omega$ the bootstrap samples $\{\hat{X}_{nj}^{(\omega)}, 1 \leq j \leq m(n)\}, n \geq 1$ are independent.

THEOREM 1. *Let $\{X_n, n \geq 1\}$ be a sequence of random variables (which are not necessarily independent or identically distributed) and let $\{m(n), n \geq 1\}$ be a sequence of positive integers with $m(n) \uparrow \infty$. If (1) holds, then for almost every $\omega \in \Omega$ the bootstrap samples $\{\{\hat{X}_{nj}^{(\omega)}, 1 \leq j \leq m(n)\}, n \geq 1\}$ obey the WLLN*

$$(4) \qquad \frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} - \frac{\sum_{i=1}^{n} X_i(\omega)}{n} \xrightarrow{P} 0.$$

REMARK. It is interesting to observe that the sequence $\{\sum_{i=1}^n X_i/n, n \geq 1\}$ does not necessarily converge a.c. (see Example 3 below).

PROOF OF THEOREM 1. For almost every $\omega \in \Omega$ and all $\varepsilon > 0$, it follows from (1) and Lemma 1 that for all sufficiently large n (depending on ω and ε)

$$\left|\hat{X}_{n1}^{(\omega)}\right| \leq \max_{1 \leq i \leq n} |X_i(\omega)| < \varepsilon\sqrt{m(n)} \leq \varepsilon m(n) \tag{5}$$

implying

$$m(n)P\left\{\left|\hat{X}_{n1}^{(\omega)}\right| \geq \varepsilon m(n)\right\} = m(n)P\emptyset = 0.$$

Also for almost every $\omega \in \Omega$

$$\begin{aligned} & \frac{Var\left(\hat{X}_{n1}^{(\omega)} I\left(\left|\hat{X}_{n1}^{(\omega)}\right| < m(n)\right)\right)}{m(n)} \\ \leq \ & \frac{E\left(\left(\hat{X}_{n1}^{(\omega)}\right)^2 I\left(\left|\hat{X}_{n1}^{(\omega)}\right| < m(n)\right)\right)}{m(n)} \\ \leq \ & \frac{\max_{1 \leq i \leq n} X_i^2(\omega)}{m(n)} \\ \rightarrow \ & 0 \text{ (by (1) and Lemma 1).} \end{aligned}$$

Then by Proposition 2 we have for almost every $\omega \in \Omega$

$$\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} - E\left(\hat{X}_{n1}^{(\omega)} I\left(\left|\hat{X}_{n1}^{(\omega)}\right| < m(n)\right)\right) \xrightarrow{P} 0. \tag{6}$$

Now taking $\varepsilon = 1$ in (5), it follows that for almost every $\omega \in \Omega$ and all sufficiently large n

$$E\left(\hat{X}_{n1}^{(\omega)} I\left(\left|\hat{X}_{n1}^{(\omega)}\right| \geq m(n)\right)\right) = E(0) = 0. \tag{7}$$

Combining (6) and (7) yields for almost every $\omega \in \Omega$,

$$\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} - \frac{\sum_{i=1}^{n} X_i(\omega)}{n} = \frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} - E\hat{X}_{n1}^{(\omega)} \xrightarrow{P} 0$$

thereby proving (4). □

COROLLARY 1. *Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d. random variables and let $\{m(n), n \geq 1\}$ be a sequence of positive integers satisfying $m(n) \uparrow \infty$ and (3). Then there exists a random variable Y such that for almost every $\omega \in \Omega$*

$$\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} \xrightarrow{P} Y(\omega) \tag{8}$$

if and only if $E|X_1| < \infty$. In such a case, $Y = EX_1$ a.c.

PROOF. Note at the outset that since (3) and (1) are equivalent, it follows from Theorem 1 that (4) holds for almost every $\omega \in \Omega$. Suppose there exists a random variable Y such that (8) holds for almost every $\omega \in \Omega$. Then by subtracting the expression in (4) from that in (8), it follows that for almost every $\omega \in \Omega$

$$\frac{\sum_{i=1}^{n} X_i(\omega)}{n} \to Y(\omega).$$

Hence

$$\limsup_{n\to\infty} \frac{\left|\sum_{i=1}^{n} X_i\right|}{n} < \infty \quad \text{a.c.}$$

and this is equivalent to $E|X_1| < \infty$.

Conversely, if $E|X_1| < \infty$, then by the Kolmogorov Strong Law of Large Numbers (SLLN) we have for almost every $\omega \in \Omega$

$$\frac{\sum_{i=1}^{n} X_i(\omega)}{n} \to EX_1.$$

Adding this to the expression in (4) gives for almost every $\omega \in \Omega$

$$\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} \xrightarrow{P} EX_1$$

and so (8) holds with $Y(\omega) = EX_1,\ \omega \in \Omega$. The last statement is now clear. □

REMARKS. (i) Even without the assumption (3), the sufficiency half of Corollary 1 holds as was proved in Athreya (1983).
(ii) Corollary 1 is true if the i.i.d. hypothesis is replaced by pairwise i.i.d.

The following corollary is a direct application of Theorem 1 to finite population sampling.

COROLLARY 2. *Let* $\{x_n, n \geq 1\}$ *be a sequence of real numbers and let* $\{m(n), n \geq 1\}$ *be a sequence of positive integers with* $m(n) \uparrow \infty$ *and* $x_n = o(\sqrt{m(n)})$. *For each* $n \geq 1$, *let* $X_{n1}, \cdots, X_{nm(n)}$ *be i.i.d. random variables where* X_{n1} *is uniformly distributed on* $\{x_1, \cdots, x_n\}$. *Then the WLLN*

$$\frac{\sum\limits_{j=1}^{m(n)} X_{nj}}{m(n)} - \frac{\sum\limits_{i=1}^{n} x_i}{n} \xrightarrow{P} 0$$

obtains.

Next we present a modification of Theorem 1. Here (1), the only condition of Theorem 1, is replaced by two conditions, which are weaker than (1), but the second of which is more complicated. In particular the second of the two conditions deals via the sample variance with the joint distribution of the $\{X_n, n \geq 1\}$, whereas (2), which implies (1), is a condition on the marginal distributions only. The two conditions are

$$\frac{X_n}{m(n)} \to 0 \quad \text{a.c.} \tag{9}$$

and

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{m(n)} \to 0 \quad \text{a.c.,} \tag{10}$$

where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $n \geq 1$. It is clear that (1) implies (9), and to see that (1) implies (10) when $m(n) \uparrow \infty$, note that

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{m(n)} \leq \frac{\sum_{i=1}^{n} X_i^2}{nm(n)} \leq \frac{\max\limits_{1 \leq i \leq n} X_i^2}{m(n)} \to 0 \quad \text{a.c.}$$

by Lemma 1.

THEOREM 2. *Let* $\{X_n, n \geq 1\}$ *be a sequence of random variables (which are not necessarily independent or identically distributed) and let* $\{m(n), n \geq 1\}$ *be a sequence of positive integers with* $m(n) \uparrow \infty$. *If (9) and (10) hold, then for almost every* $\omega \in \Omega$ *the bootstrap samples* $\{\{\hat{X}_{nj}^{(\omega)}, 1 \leq j \leq m(n)\}, n \geq 1\}$ *obey the WLLN (4).*

PROOF. The proof is similar to that of Theorem 1. Condition (9) is sufficient for (5) (without the symbols $\varepsilon\sqrt{m(n)} \leq$) and (7), and it also implies that for almost every $\omega \in \Omega$ and sufficiently large n

$$\begin{aligned}\frac{Var\left(\hat{X}_{n1}^{(\omega)}I\left(\left|\hat{X}_{n1}^{(\omega)}\right|<m(n)\right)\right)}{m(n)} &= \frac{Var\hat{X}_{n1}^{(\omega)}}{m(n)}\\ &= \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2}{m(n)}\\ &\to 0 \text{ (by (10))}.\end{aligned}$$

□

COROLLARY 3. Let $\{X_n, n \geq 1\}$ be a stationary sequence of random variables with $EX_1^2 < \infty$ and let $\{m(n), n \geq 1\}$ be a sequence of positive integers with $m(n) \uparrow \infty$. If (9) holds, then for almost every $\omega \in \Omega$ the bootstrap samples $\{\{\hat{X}_{nj}^{(\omega)}, 1 \leq j \leq m(n)\}, n \geq 1\}$ obey the WLLN (4) and, moreover, for almost every $\omega \in \Omega$

$$\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} \xrightarrow{P} E(X_1 | \mathcal{I})(\omega), \tag{11}$$

where $\mathcal{I}$ denotes the σ-algebra of invariant events for $\{X_n, n \geq 1\}$.

PROOF. To establish (4) it suffices to verify (10). Now $\{X_n^2, n \geq 1\}$ is also a stationary sequence and so by the pointwise ergodic theorem for stationary sequences (see Stout (1974, p. 181))

$$\frac{\sum_{i=1}^{n} X_i^2}{n} \to E(X_1^2 | \mathcal{J}) < \infty \quad \text{a.c.}$$

where $\mathcal{J}$ denotes the σ-algebra of invariant events for $\{X_n^2, n \geq 1\}$. Hence

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2}{m(n)} \leq \frac{\sum_{i=1}^{n} X_i^2}{nm(n)} \to 0 \quad \text{a.c.}$$

verifying (10). To prove (11), another application of the pointwise ergodic theorem for stationary sequences yields

$$\frac{\sum_{i=1}^{n} X_i}{n} \to E(X_1 | \mathcal{I}) \quad \text{a.c.}$$

which in conjunction with (4) yields (11). □

The final corollary is a further specialization of Corollary 3 to the celebrated autoregressive conditional heteroscedastic (ARCH) process, introduced in Engle (1982), frequently used in financial econometrics. We consider an ARCH process of first order, i.e. the X_n satisfy

$$X_n = \varepsilon_n \sqrt{\beta + \lambda X_{n-1}^2}\ ,\ n \geq 1$$

where $\beta > 0$ and $0 < \lambda < 1$, for some initial random variable X_0 independent of the i.i.d. standard normal innovations $\{\varepsilon_n, n \geq 1\}$. Furthermore let X_0 be such that $\{X_n, n \geq 1\}$ is a stationary sequence. Clearly the X_n have mean zero and are uncorrelated. It is well-known that for these values of the parameters $EX_1^2 < \infty$ and that, more precisely,

$$P\{|X_1| \geq x\} \sim cx^{-\alpha} \quad \text{as } x \to \infty \tag{12}$$

for some $\alpha > 2$ and $c > 0$. Actually $\alpha > 2$ is the root of the equation

$$(2\lambda)^{\alpha}\Gamma^2\left(\frac{\alpha+1}{2}\right) = \pi.$$

Note that $\lim_{\lambda\uparrow 1} \alpha = 2$. For this and more on ARCH processes, see Embrechts, Klüppelberg, and Mikosch (2001), Chapter 8.

COROLLARY 4. Let $\{X_n, n \geq 1\}$ be a stationary, first order ARCH process as defined above and let $\{m(n), n \geq 1\}$ be a nondecreasing sequence of positive integers such that $n^{\beta} = O(m(n))$ for some $\beta > 1/\alpha$. Then for almost every $\omega \in \Omega$ the bootstrap samples $\{\{\hat{X}_{nj}^{(\omega)}, 1 \leq j \leq m(n)\}, n \geq 1\}$ obey the WLLN (4) and, moreover, for almost every $\omega \in \Omega$

$$\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} \xrightarrow{P} 0. \tag{13}$$

PROOF. The condition (9) follows immediately from (12) and the Borel-Cantelli lemma. Hence (4) follows from Corollary 3. In order to show (13), note that, since the X_n are uncorrelated, it follows from the Chebychev inequality that $\frac{\sum_{i=1}^n X_i}{n} \xrightarrow{P} 0$. This in conjunction with the pointwise ergodic theorem applied to $\{X_n, n \geq 1\}$ yields $\frac{\sum_{i=1}^n X_i}{n} \to 0$ a.c. Combining this with (4) yields (13). □

4 Some Interesting Examples

In this section some examples are provided which illustrate the sharpness of the results. The first example shows that Theorem 1 can fail if the condition (1) is dispensed with.

EXAMPLE 1. Let $\{X_n,\ n \geq 1\}$ be a sequence of random variables (not necessarily independent), where X_n takes values in $[n-1, n+1]$, $n \geq 1$. Let $\{m(n),\ n \geq 1\}$ be a sequence of positive integers with $m(n) \uparrow \infty$ and $m(n) = O(n^2)$. Then (1) fails. Now for almost every $\omega \in \Omega$ and all $n \geq 1$,

$$E\hat{X}_{n1}^{(\omega)} = \bar{X}_n(\omega)$$

and

$$\begin{aligned} Var\hat{X}_{n1}^{(\omega)} &= \sum_{i=1}^{n}(X_i(\omega) - \bar{X}_n(\omega))^2 n^{-1} \\ &= \frac{\sum_{i=1}^{n} X_i^2(\omega)}{n} - \frac{\left(\sum_{i=1}^{n} X_i(\omega)\right)^2}{n^2} \\ &\geq \frac{\sum_{i=1}^{n}(i-1)^2}{n} - \frac{\left(\sum_{i=1}^{n}(i+1)\right)^2}{n^2} \\ &= \frac{n^2 - 24n - 25}{12}. \end{aligned} \tag{14}$$

We will first show that for almost every $\omega \in \Omega$,

$$\frac{\sum_{j=1}^{m(n)} (\hat{X}_{nj}^{(\omega)} - \bar{X}_n(\omega))}{\left(m(n)Var\hat{X}_{n1}^{(\omega)}\right)^{1/2}} \xrightarrow{d} N(0,1). \tag{15}$$

It suffices to verify that (see Chung (1974, p. 200)) the Liapounov condition

$$\frac{\sum_{j=1}^{m(n)} E|\hat{X}_{nj}^{(\omega)} - \bar{X}_n(\omega)|^3}{\left(m(n)Var\hat{X}_{n1}^{(\omega)}\right)^{3/2}} \to 0 \tag{16}$$

holds for almost every $\omega \in \Omega$. Now, by (14) it follows that for almost every $\omega \in \Omega$ and all large n,

$$\begin{aligned} \frac{\sum_{j=1}^{m(n)} E|\hat{X}_{nj}^{(\omega)} - \bar{X}_n(\omega)|^3}{\left(m(n)Var\hat{X}_{n1}^{(\omega)}\right)^{3/2}} &\leq 13^{3/2}\frac{m(n)E|\hat{X}_{n1}^{(\omega)} - \bar{X}_n(\omega)|^3}{(m(n))^{3/2}n^3} \\ &\leq \frac{13^{3/2}(n+1)^3}{(m(n))^{1/2}n^3} = o(1) \end{aligned}$$

proving (16) and hence (15). Then for almost every $\omega \in \Omega$ and arbitrary $\varepsilon > 0$ we have for all large n

$$\begin{aligned}
& P\left\{\left|\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} - \frac{\sum_{i=1}^{n} X_i(\omega)}{n}\right| \le \varepsilon\right\} \\
= \; & P\left\{\left|\frac{\sum_{j=1}^{m(n)} (\hat{X}_{nj}^{(\omega)} - \bar{X}_n(\omega))}{m(n)}\right| \le \varepsilon\right\} \\
\le \; & P\left\{\frac{\sum_{j=1}^{m(n)} (\hat{X}_{nj}^{(\omega)} - \bar{X}_n(\omega))}{\left(m(n) Var \hat{X}_{n1}^{(\omega)}\right)^{1/2}} \le \frac{(m(n))^{1/2}\varepsilon}{\left(Var \hat{X}_{n1}^{(\omega)}\right)^{1/2}}\right\}.
\end{aligned}$$

Since (14) and $m(n) = O(n^2)$ ensure that $(m(n))^{1/2}/(Var\hat{X}_{n1}^{(\omega)})^{1/2}$ is bounded above, it follows from (15) that the latter probability is, for large n, bounded away from 1. Thus for almost every $\omega \in \Omega$, (4) fails.

Note, however, if $m^*(n) = [n^2 b_n]$, $n \ge 1$ where $1 \le b_n \uparrow \infty$, then $X_n/\sqrt{m^*(n)} \to 0$ a.c. and it follows from Theorem 1 that for almost every $\omega \in \Omega$

$$\frac{\sum_{j=1}^{m^*(n)} \hat{X}_{nj}^{(\omega)}}{m^*(n)} - \frac{\sum_{i=1}^{n} X_i(\omega)}{n} \xrightarrow{P} 0.$$

The following example shows even for a sequence of i.i.d. random variables $\{X_n, n \ge 1\}$ that under the hypotheses of Theorem 1, convergence in probability in the conclusion (4) of Theorem 1 cannot necessarily be replaced by a.c. convergence.

EXAMPLE 2. Let $\{X_n, n \ge 1\}$ be a sequence of i.i.d. nondegenerate random variables with

$$E(\exp\{cX_1^2\}) < \infty \quad \text{for all } c > 0, \tag{17}$$

and let $m(n) = [\log n] \vee 1,\ n \ge 1$. Since (17) and the Borel-Cantelli lemma ensure (1), by Theorem 1 the conclusion (4) holds for almost every $\omega \in \Omega$. Now by the Kolmogorov SLLN

$$\frac{\sum_{i=1}^{n} X_i(\omega)}{n} \to EX_1 \quad \text{for almost every } \omega \in \Omega.$$

Thus, if a.c. convergence prevails for the expression in (4) for almost every $\omega \in \Omega$, then

$$\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} \to EX_1 \quad \text{a.c. for almost every } \omega \in \Omega. \tag{18}$$

However, assuming that the bootstrap samples $\{\hat{X}_{nj}^{(\omega)}, 1 \le j \le m(n)\}$, $n \ge 1$ are independent for almost every $\omega \in \Omega$ and recalling that the $\{X_n, n \ge 1\}$ are nondegenerate, it follows from Theorem 4.1 of Csörgő and Wu (2000) that (18) fails. (Actually, it follows from that paper that an "unconditional" form of (18) fails. However, Csörgő and Wu (2000) also proved that the "conditional" and "unconditional" bootstrap SLLN are indeed one and the same in the sense that one of them holds if and only if the other holds.) Consequently, convergence in probability in (4) cannot be replaced by a.c. convergence.

REMARK. Further discussion of the relationship between "conditional" and "unconditional" bootstrap limit laws is in Csörgő and Rosalsky (2003).

The next example shows that the hypotheses of Theorem 1 can be satisfied (hence (4) prevails for almost every $\omega \in \Omega$) even when

(i) there does *not* exist a sequence of constants $\{C_n, n \ge 1\}$ with

$$\frac{\sum_{i=1}^{n} X_i}{n} - C_n \xrightarrow{P} 0, \tag{19}$$

(ii)

$$\limsup_{n\to\infty} \frac{\left|\sum_{i=1}^{n} X_i(\omega)\right|}{n} = \infty \quad \text{for almost every } \omega \in \Omega, \tag{20}$$

and

(iii) there does *not* exist a random variable Y such that

$$\frac{\sum_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} \xrightarrow{P} Y(\omega) \quad \text{for almost every } \omega \in \Omega. \tag{21}$$

EXAMPLE 3. Let $\{X_n, n \ge 1\}$ be a sequence of i.i.d. random variables with $nP\{|X_1| > n\} \nrightarrow 0$. Then there does not exist a sequence of constants $\{C_n, n \ge 1\}$ such that (19) holds (see, Chow and Teicher (1997, p. 128)), and $E|X_1| = \infty$ which ensures (20). Let $\{m(n), n \ge 1\}$ be a sequence of positive integers satisfying $m(n) \uparrow \infty$ and (3). Then (1) holds, and it then follows from Theorem 1 that (4) holds for almost every $\omega \in \Omega$. Corollary 1 and $E|X_1| = \infty$ ensure that there does not exist a random variable Y satisfying (21).

The next example is perhaps the simplest and it shows that condition (1) can prevail without any moment conditions on the $\{X_n, n \geq 1\}$ or conditions on the rate in which $m(n) \uparrow \infty$.

EXAMPLE 4. Let $X_n = A_n X, n \geq 1$ where X is any random variable and $\{A_n, n \geq 1\}$ is a sequence of uniformly bounded random variables. Let $\{m(n), n \geq 1\}$ be a sequence of positive integers with $m(n) \uparrow \infty$. Then (1) holds since the $\{A_n, n \geq 1\}$ are uniformly bounded whence by Theorem 1 the conclusion (4) prevails.

REMARK. In Example 4, if the $\{A_n, n \geq 1\}$ are i.i.d. bounded random variables, then by the Kolmogorov SLLN, for almost every $\omega \in \Omega$,

$$\frac{\sum_{i=1}^n X_i(\omega)}{n} = \frac{X(\omega)\sum_{i=1}^n A_i(\omega)}{n} \to (EA_1)X(\omega)$$

which when combined with (4) yields for almost every $\omega \in \Omega$

$$\frac{\sum\limits_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} \xrightarrow{P} (EA_1)X(\omega). \tag{22}$$

We note that by taking bootstrap samples from the sequence $\{A_n, n \geq 1\}$, (22) also follows directly from the bootstrap WLLN of Athreya (1983).

The final example pertains to Corollary 1. It shows that if (3) fails, then the "necessity" half of Corollary 1 does not hold. In other words, we present an example where (3) fails, there exists a random variable Y such that

$$\frac{\sum\limits_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} \xrightarrow{P} Y(\omega) \quad \text{for almost every } \omega \in \Omega,$$

and $E|X_1| = \infty$.

EXAMPLE 5. Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d. random variables from the probability distribution that is symmetric around 0 and has distribution function

$$F(x) = \frac{\log 2}{-x\log(-x)} \quad \text{for } x \leq -2.$$

Set $m(n) = [\sqrt{n}],\ n \geq 1$. Then indeed (3) fails. Clearly also $E|X_1| = \infty$. We will show that nevertheless

$$\frac{\sum\limits_{j=1}^{m(n)} \hat{X}_{nj}^{(\omega)}}{m(n)} \xrightarrow{P} 0 \quad \text{for almost every } \omega \in \Omega. \tag{23}$$

The proof of (23) will be based on Proposition 2. First note that

$$\lim_{n\to\infty} \frac{n^{1/8}}{\log^{1/2} n} \sup_{n^{-3/4}\le t\le 1} \left| \frac{G_n(t)}{t} - 1 \right| = 0 \quad \text{a.c.},$$

where G_n is the empirical distribution function of the first n of a sequence of i.i.d. Uniform-(0,1) random variables; see Wellner (1978). This implies

$$\lim_{n\to\infty} \frac{n^{1/8}}{\log^{1/2} n} \sup_{-n^{1/2}\le x<\infty} \left| \frac{F_n(x)}{F(x)} - 1 \right| = 0 \quad \text{a.c.}, \tag{24}$$

where F_n is the empirical distribution function of the first n of a sequence of i.i.d. random variables with distribution function F. In particular we have

$$\limsup_{n\to\infty} \sup_{-n^{1/2}\le x<\infty} \frac{F_n(x)}{F(x)} < 2 \quad \text{a.c.} \tag{25}$$

We now check the two conditions of Proposition 2. Let Y_{n1} denote a random variable with distribution function F_n and let F_{n-} be the left-continuous version of F_n. First note that for $\varepsilon > 0$

$$m(n)P\{|Y_{n1}| \ge \varepsilon m(n)\} = [\sqrt{n}]\left(F_n(-\varepsilon[\sqrt{n}]) + 1 - F_{n-}(\varepsilon[\sqrt{n}])\right).$$

Because of symmetry of F, we only consider $[\sqrt{n}]F_n(-\varepsilon[\sqrt{n}])$, $\varepsilon \le 1$ and observe that in view of (25), for almost every $\omega \in \Omega$ and for large n this expression is bounded above by $2[\sqrt{n}]F(-\varepsilon[\sqrt{n}]) \to 0$. For the second condition consider

$$\frac{1}{m(n)} Var(Y_{n1}I(|Y_{n1}| < m(n))) \le \frac{1}{m(n)} E(Y_{n1}^2 I(|Y_{n1}| < m(n))) \le \frac{1}{[\sqrt{n}]} \int_{-[\sqrt{n}]}^{[\sqrt{n}]} x^2 dF_n(x).$$

Because of symmetry of F, we only consider

$$\frac{1}{[\sqrt{n}]} \int_{-[\sqrt{n}]}^{-2} x^2 dF_n(x) = \frac{1}{[\sqrt{n}]} x^2 F_n(x)\Big|_{-[\sqrt{n}]}^{-2} + \frac{2}{[\sqrt{n}]} \int_{-[\sqrt{n}]}^{-2} -xF_n(x)dx.$$

The right-hand side of this expression is in view of (25), for almost every $\omega \in \Omega$ and for large n bounded above by

$$\frac{4}{[\sqrt{n}]} + \frac{4}{[\sqrt{n}]} \int_{-[\sqrt{n}]}^{-2} -xF(x)dx \to 0.$$

So in order to prove (23), it remains to show that

$$E(Y_{n1}I(|Y_{n1}| < m(n))) = \int_{-[\sqrt{n}]}^{[\sqrt{n}]} x dF_n(x) \to 0 \quad \text{a.c.}$$

Since $\int_{-[\sqrt{n}]}^{[\sqrt{n}]} xdF(x) = 0$, we will show that

$$\int_{-[\sqrt{n}]}^{[\sqrt{n}]} xd(F_n(x) - F(x)) \to 0 \quad \text{a.c.}$$

Again, using symmetry of F we only consider

$$\begin{aligned}\int_{-[\sqrt{n}]}^{-2} xd(F_n(x) - F(x)) &= x(F_n(x) - F(x))|_{-[\sqrt{n}]}^{-2} - \int_{-[\sqrt{n}]}^{-2} (F_n(x) - F(x))dx \\ &= -2\left(F_n(-2) - \frac{1}{2}\right) + [\sqrt{n}]\left(F_n(-[\sqrt{n}]) - F(-[\sqrt{n}])\right) \\ &\quad - \int_{-[\sqrt{n}]}^{-2} (F_n(x) - F(x))dx.\end{aligned}$$

By the Glivenko-Cantelli theorem the first term on the right tends to 0 a.c. The absolute value of the second term is by (25) for almost every $\omega \in \Omega$, for large n, bounded by $[\sqrt{n}]F(-[\sqrt{n}]) \to 0$. The absolute value of the third term is by (24) bounded for almost every $\omega \in \Omega$, for large n, by

$$\frac{\log^{1/2} n}{n^{1/8}} \int_{-[\sqrt{n}]}^{-2} F(x)dx \to 0.$$

Hence we showed (23).

REMARK. Observe that for this example, for almost every $\omega \in \Omega$, the sample path of sequential sample means shows "irregular" behavior due to $E|X_1| = \infty$, whereas the distribution of bootstrap sample means is in the limit degenerate at 0.

REFERENCES

1. S.E. Ahmed, T.-C. Hu, and A.I. Volodin, On the rate of convergence of bootstrapped means in a Banach space. *Int. J. Math. & Math. Sci.* **25** (2001), 629-635.
2. S.E. Ahmed, D. Li, A. Rosalsky, and A.I. Volodin, Almost sure lim sup behavior of bootstrapped means with applications to pairwise i.i.d. sequences and stationary ergodic sequences. *J. Statist. Plann. Inference* **98** (2001), 1-14.
3. M.A. Arcones and E. Giné, The bootstrap of the mean with arbitrary bootstrap sample size. *Ann. Inst. Henri Poincaré, Probab. Statist.* **25** (1989), 457-481.
4. E. Arenal-Gutiérrez and C. Matrán, A zero-one law approach to the central limit theorem for the weighted bootstrap mean. *Ann. Probab.* **24** (1996), 532-540.
5. E. Arenal-Gutiérrez, C. Matrán, and J.A. Cuesta-Albertos, Unconditional Glivenko-Cantelli-type theorems and weak laws of large numbers for bootstrap. *Statist. Probab. Lett.* **26** (1996a), 365-375.
6. E. Arenal-Gutiérrez, C. Matrán, and J.A. Cuesta-Albertos, On the unconditional strong law of large numbers for the bootstrap mean. *Statist. Probab. Lett.* **27** (1996b), 49-60.

7. K.B. Athreya, Strong law for the bootstrap. *Statist. Probab. Lett.* **1** (1983), 147-150.

8. K.B. Athreya, M. Ghosh, L.Y. Low, and P.K. Sen, Laws of large numbers for bootstrapped U-statistics. *J. Statist. Plann. Inference* **9** (1984), 185-194.

9. P.J. Bickel and D.A. Freedman, Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** (1981), 1196-1217.

10. A. Bozorgnia, R.F. Patterson, and R.L. Taylor, Chung type strong laws for arrays of random elements and bootstrapping. *Stochastic Anal. Appl.* **15** (1997), 651-669.

11. Y.S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, 3rd ed. Springer-Verlag, New York 1997.

12. K.L. Chung, *A Course in Probability Theory*, 2nd ed. Academic Press, New York 1974.

13. S. Csörgő, On the law of large numbers for the bootstrap mean. *Statist. Probab. Lett.* **14** (1992), 1-7.

14. S. Csörgő, Rates in the complete convergence of bootstrap means. *Statist. Probab. Lett.* **64** (2003), 359-368.

15. S. Csörgő, On the complete convergence of bootstrap means. In: *Asymptotic Methods in Stochastics: Festschrift for Miklós Csörgő* (L. Horváth, B. Szyszkowicz, eds.), Fields Institute Communications, Volume **44** (2004).

16. S. Csörgő and A. Rosalsky, A survey of limit laws for bootstrapped sums. *Int. J. Math. & Math. Sci.* **2003** (2003), 2835-2861.

17. S. Csörgő and W.B. Wu, Random graphs and the strong convergence of bootstrap means. Combin. Probab. Comput. **9** (2000), 315-347.

18. B. Efron, Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7** (1979), 1-26.

19. P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, New York 2001.

20. R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** (1982), 987-1007.

21. N. Etemadi, On the maximal inequalities for the average of pairwise i.i.d. random variables. *Comm. Statist. A - Theory Methods* **13** (1984), 2749-2756.

22. E. Giné and J. Zinn, Necessary conditions for the bootstrap of the mean. *Ann. Statist.* **17** (1989), 684-691.

23. T.-C. Hu and R.L. Taylor, On the strong law for arrays and for the bootstrap mean and variance. *Int. J. Math. & Math. Sci.* **20** (1997), 375-382.

24. D. Li and A. Rosalsky, Erdős-Rényi-Shepp laws for arrays with applications to bootstrapped sample means. *Pakistan J. Statist.* **18** (2002), 255-263.

25. D. Li, A. Rosalsky, and S.E. Ahmed, Complete convergence of bootstrapped means and moments of the supremum of normed bootstrapped sums. *Stochastic Anal. Appl.* **17** (1999), 799-814.

26. D. Li, A. Rosalsky, and D.K. Al-Mutairi, A large deviation principle for bootstrapped sample means. *Proc. Amer. Math. Soc.* **130** (2002), 2133-2138.

27. T. Mikosch, Almost sure convergence of bootstrapped means and U-statistics. *J. Statist. Plann. Inference* **41** (1994), 1-19.

28. K. Singh, On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** (1981), 1187-1195.

29. W.F. Stout, *Almost Sure Convergence*, Academic Press, New York 1974.

30. J.A. Wellner, Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. Verw. Gebiete* **45** (1978) 73-88.

Dept. of Econometrics & OR
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
Email: j.h.j.einmahl@uvt.nl

Dept. of Statistics
University of Florida
Box 118545
Gainesville, Florida 32611-8545
U.S.A.
Email: rosalsky@stat.ufl.edu