Almost Sure Central Limit Theorems for Stationary Bootstrap Mean

Eunju Hwang* and Dong Wan Shin1,a

*Institute of Mathematical Sciences and Department of Statistics, Ewha Womans University

Abstract

Almost sure central limit theorems are established for a stationary bootstrap sample mean of strong mixing processes. Both weak and strong consistencies are obtained.

Keywords: Stationary bootstrap, almost sure central limit theorem.

1. Introduction

The almost sure central limit theorem (ASCLT) was first discovered independently by Brosamler (1988) and Schatte (1988) and has been further studied for independent or dependent random variables by several authors. Lacey and Philipp (1990), Berkes and Dehling (1994), and Berkes (1995) developed the ASCLT for independent and identically distributed random variables; however, Berkes and Dehling (1993) and Berkes and Csaki (2001) developed the ASCLT for independent but not identically distributed random variables. ASCLT were established by Peligrad and Sho (1995) for associated sequences, strong mixing and ρ-mixing sequences and by Matula (1998) for associated random variables. Lesigne (1999) studied the ASCLT for stationary sequences on an aperiodic measure preserving system, Dudzinski (2003) dealt with the ASCLT under a general weak dependence condition proposed by Doukhan and Louhichi (1999), and Chen and Lin (2008) obtained the ASCLT for functionals of absolutely regular processes. See Lahiri (2003) for CLT results of stationary bootstrapping for statistics other than the sample mean and the partial sum process.

This work establishes ASCLT for the stationary bootstrapping sample mean of stationary strong mixing sequences. Both weak and strong consistencies are established. Stationary bootstrapping, proposed by Politis and Romano (1994), is one of the most widely used block bootstrapping methods for dependent processes. We find several CLT results among the many results on stationary bootstrapping. Standard CLT versions of the stationary bootstrapping can be found in Politis and Romano (1994) and Hwang and Shin (2012), where the asymptotic validity of the stationary bootstrapping was shown in view of the weak and strong consistency, respectively. Functional CLTs for stationary bootstrapping are established under strong mixing by Paparoditis and Politis (2003) and Parker et al. (2006) in terms of unit root tests.

2. Stationary Bootstrapping

We describe the stationary bootstrap method for a sample mean. Let \(X_1, \ldots, X_n\) be the observed data set. First a new time series \(\{X_{ni}: i \geq 1\}\) is obtained by wrapping the data \(X_1, \ldots, X_n\) around
a circle, and relabelling them as $X_{n1}, X_{n2}, \ldots$ such that for each $i \geq 1$, $X_{ni} := X_j$ where $i = j \pmod{n}$. A stationary bootstrap sample $\{X_{n1}^*, \ldots, X_{nk}^*\}$ is then constructed by combining blocks $B(i, \ell) = \{X_{ni}, \ldots, X_{ni+\ell-1}\}$ consisting of $\ell$ observations with geometrically distributed block length $\ell$ and uniformly distributed starting point $i$ as described below.

Random variables $I_{n1}, I_{n2}, \ldots$ and $L_{n1}, L_{n2}, \ldots$ are generated such that conditionally on the observations $X_1, \ldots, X_n$, $\{I_{n1}, I_{n2}, \ldots\}$ are i.i.d. discrete uniform on $[1, \ldots, n]$, $P_n(I_{n1} = i) = 1/n$, $i = 1, \ldots, n$; $\{L_{n1}, L_{n2}, \ldots\}$ are i.i.d. random variables having the geometric distribution with a parameter $p \equiv p_n$ depending on $n$, $P_n(L_{n1} = \ell) = p(1-p)^{\ell-1}$, $\ell = 1, 2, \ldots$; and the collections $\{I_{n1}, I_{n2}, \ldots\}$ and $\{L_{n1}, L_{n2}, \ldots\}$ are independent. Here and in the following, $P_n$ and $E_n$ denote the conditional probability and the conditional expectation, respectively, given $X_1, \ldots, X_n$.

Now, a pseudo-time series $X_{n1}^*, \ldots, X_{nk}^*$ is generated by arranging elements in a series from blocks $B(I_{n1}, L_{n1}), \ldots, B(I_{nk}, L_{nk})$ and deleting the last $L_{n1} + \cdots + L_{nk} - n$ elements, where $\tau_n = \inf\{k \geq 1 : L_{n1} + \cdots + L_{nk} \geq n\}$.

We assume that $p = p_n$ goes to 0 as $n \to \infty$, and thus the expected block length $E_n(L_{n1}) = 1/p$ tends to $\infty$ as $n \to \infty$. Conditionally on $X_1, \ldots, X_n$, the process $X_{n1}^*, \ldots, X_{nk}^*$ is stationary.

Let $\bar{X}_n = \sum_{i=1}^{n} X_i/n$ and $\bar{X}_n^* = \sum_{i=1}^{n} X_{ni}^*/n$. Obviously, $E_n(X_{n1}^*) = \bar{X}_n$.

### 3. Main Results

Let $\{X_t : t = 1, 2, \ldots\}$ be a stationary sequence of strong mixing random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ with strong mixing coefficient $\alpha_X(k)$, $k = 1, 2, \ldots$. We present the main results of ASCLTs for a triangular array of the stationary bootstrapping observations $\{X_{nk}^* : k = 1, 2, \ldots, n; i = 1, 2, \ldots, k; n \to \infty\}$.

Before giving the ASCLTs, we review the CLTs for stationary bootstrapping sample mean. Assume either

\begin{enumerate}[(i)]
  \item $E|X_1|^{6+\delta} < \infty$ for some $\delta > 0$; $\alpha_X(k) = O(k^{-r})$, for some $r > 3(6+\delta)/\delta$; $p = p_n \to 0$, $np_n \to \infty$.
  \item $E|X_1|^{q+\delta} < \infty$ for some $q > 2$, $\delta > 0$; $\alpha_X(k) = O(k^{-r})$ for some $r > q(q+\delta)/(2\delta)$; $p = c \cdot n^{-(\delta-2\delta)/(2+\delta)}$ for some $c > 0$, and $0 < \epsilon < \delta/2$.
\end{enumerate}

Under (i), Politis and Romano (1994) showed the weak consistency result that, conditionally on $X_1, \ldots, X_n$,

$$\sqrt{n} \left[ \bar{X}_n^* - \bar{X}_n \right] \xrightarrow{d} N\left(0, \sigma^2\right) \quad \text{(conditional probability)} \quad (3.1)$$

as $n \to \infty$, where $\sigma^2 = \text{Var}(X_1) + 2 \sum_{i=1}^{\infty} \text{Cov}(X_i, X_{i+1})$. Under some more restrictive condition (ii), Hwang and Shin (2012) obtained a strong consistency result that, conditionally on $X_1, \ldots, X_n$,

$$\sqrt{n} \left[ \bar{X}_n^* - \bar{X}_n \right] \xrightarrow{a.s.} N\left(0, \sigma^2\right) \quad (3.2)$$

The following theorem establishes two ASCLTs: the weak consistency under the condition (i) and the strong consistency under the stronger condition (ii).

**Theorem 1.** Let $X_1, X_2, \ldots$ be a strictly stationary process of strong mixing random variables. Then, conditionally on $X_1, \ldots, X_n$, as $n \to \infty$,

$$\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \left( \frac{\sqrt{k} \left[ \bar{X}_k^* - \bar{X}_k \right]}{\sigma} \right) \leq x \xrightarrow{a.s.} \Phi(x) \quad (3.3)$$
in (conditional) probability under (i) or almost surely under (ii), where $\Phi(x)$ is the standard normal distribution function and $I(\cdot)$ is the indicator function.

**Remark 1.** For processes with $E|X_1|^{q+\delta} < \infty$ for some $q > 6$, condition (ii) is uniformly more restrictive than condition (i) because $3(6+\delta)/\delta < q(q+\delta)/(2\delta)$, $n^{-(d-2\delta)/(2+\delta)} \to 0$, and $nm^{-(d-2\delta)/(2+\delta)} \to \infty$. Under this stronger condition (ii), we have a stronger result (3.3) than the result (3.3) in probability under the weaker condition (i).

### 4. Proofs

**Proof of Theorem 1 under condition (i):** Let $S_k^* = \sqrt{n}(X_k^* - \bar{X}_n)/\sigma$. Let $f$ be a bounded function with bounded continuous derivatives. According to Theorem 7.1 in Billingsley (1968) and in the same way as in Lacey and Philipp (1990), (3.3) in probability is equivalent to, conditionally on $X_1, \ldots, X_n$,

$$
\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} f(S_k^*) \xrightarrow{a.s.} E f(Z) \quad \text{in probability},
$$

where $Z$ denotes a standard normal random variable. Now it remains to show (4.1).

Set $f_1 = \max(f, 0)$. Obviously $f_1$ is a nonnegative bounded Lipschitz function. The CLT result in (3.1) implies that

$$
E_k f_1(\bar{S}_k^*) \xrightarrow{p} E f_1(Z) \quad \text{as } k \to \infty,
$$

and hence

$$
\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} E_k f_1(\bar{S}_k^*) \xrightarrow{p} E f_1(Z) \quad \text{as } n \to \infty. \tag{4.2}
$$

Consider a sequence $\{n_j : j = 1, 2, \ldots\}$ with $n_j = [\exp(j^{1/2})]$ the smallest integer greater than $\exp(j^{1/2})$ for some $\nu > 0$. Note that $\lim_{j \to \infty} \log n_j/\log n_{j+1} = 1$. By this fact and by the nonnegativity of $f_1$, we have

$$
\limsup_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} f_1(S_k^*) = \limsup_{j \to \infty} \frac{1}{\log n_j} \sum_{k=1}^{n_j} \frac{1}{k} f_1(\bar{S}_k^*) \quad \text{a.s.} \tag{4.3}
$$

and

$$
\liminf_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} f_1(S_k^*) = \liminf_{j \to \infty} \frac{1}{\log n_j} \sum_{k=1}^{n_j} \frac{1}{k} f_1(\bar{S}_k^*) \quad \text{a.s.} \tag{4.4}
$$

We note that for $k \neq l$,

$$
\{X_{i,j}^* : i = 1, \ldots, k\} \text{ and } \{X_{i,j}^* : j = 1, \ldots, l\} \text{ are (conditionally) independent} \tag{4.5}
$$

since $(I_{k,i}, L_{k,i})$ and $(I_{l,j}, L_{l,j})$ are independent. For any $\epsilon > 0$, by Chebyshev’s inequality,

$$
P_n \left[ \frac{1}{\log n_j} \sum_{k=1}^{n_j} \frac{1}{k} \left| f_1(S_k^*) - E_k f_1(S_k^*) \right| > \epsilon \right] \leq \frac{1}{\epsilon^2 (\log n_j)^2} \text{Var}_n \left[ \frac{1}{k} \left( f_1(S_k^*) - E_k f_1(S_k^*) \right) \right]
$$
which is bounded above a.s. by
\[
\frac{C}{e^2(\log n)^2} \sum_{k=1}^{n_j} \frac{1}{k^2} \leq \frac{C}{e^2(\log n)^2} \sum_{k=1}^{n_j} \frac{1}{k} \leq \frac{C}{e^2 \log n_j}
\]
because of boundedness of \( f_1 \) and (4.5). Hence we have
\[
\sum_{j=1}^{\infty} P^*_n \left[ \frac{1}{\log n_j} \sum_{k=1}^{n_j} \frac{1}{k} \left[ f_1 \left( S^*_k \right) - E^*_k f_1 \left( S^*_k \right) \right] > \epsilon \right] \leq \frac{C}{e^2} \sum_{j=1}^{\infty} \frac{1}{j^{1+\nu}} < \infty \quad \text{a.s.}
\]
thus, by the Borel-Cantelli Lemma, conditionally on \( X_1, \ldots, X_{n_j} \),
\[
\frac{1}{\log n_j} \sum_{k=1}^{n_j} \frac{1}{k} \left[ f_1 \left( S^*_k \right) - E^*_k f_1 \left( S^*_k \right) \right] \xrightarrow{a.s.} 0 \quad \text{a.s.} \quad (4.6)
\]
By (4.2)–(4.6) we obtain that, conditionally on \( X_1, \ldots, X_n \),
\[
\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} f_1 \left( S^*_k \right) \xrightarrow{a.s.} E f_1(Z) \quad \text{in probability.} \quad (4.7)
\]
In the same way, letting \( f_2 = \min(f_1, 0) \), we have
\[
\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} f_2 \left( S^*_k \right) \xrightarrow{a.s.} E f_2(Z) \quad \text{in probability.} \quad (4.8)
\]
By (4.7) and (4.8), we have the desired results of (4.1) and thus of (3.3) in conditional probability. \( \square \)

**Proof of Theorem 1 under condition (ii):** The CLT result in (3.2) implies that \( E^*_k f_i \left( S^*_k \right) \xrightarrow{a.s.} E f_i(Z) \) as \( k \to \infty \), for \( i = 1, 2 \), and hence
\[
\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} E^*_k f_i \left( S^*_k \right) \xrightarrow{a.s.} E f_i(Z) \quad \text{as } n \to \infty,
\]
Similarly to the proof of Theorem 1 under condition (i), (3.3) almost surely is equivalent to, conditionally on \( X_1, \ldots, X_n \),
\[
\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} f \left( S^*_k \right) \xrightarrow{a.s.} E f(Z) \quad \text{a.s.} \quad (4.9)
\]
and the remaining part in proving (4.9) is the same as in that of Theorem 1 under condition (i). \( \square \)

**References**
Almost Sure Central Limit Theorems for Stationary Bootstrap Mean


*Received March 11, 2013; Revised April 26, 2013; Accepted April 26, 2013*