Journal of the Korean Statistical Society (2001), **30**: *3, pp 407–418*

Validity of Blockwise Bootstrapped Empirical Process with Multivariate Stationary Sequences

Tae-Yoon Kim, Ki-Dong Shin and Gyu-Moon Song[1]

Abstract

Bühlmann (1994) established the validity of the block bootstrap proposed by Künsch when it is applied to p-dimensional α-mixing dependent sequence. But his result requires a rather restrictive condition on p in the sense that p is entangled with dependence structure. We address that such restriction on p (or complication of dependence structure with p) could be removed completely when the underlying dependence structure is replaced by more weakly dependent structure such as ϕ-mixing.

Keywords: Bootstrap, multivariate empirical process, ϕ-mixing sequences

1. Introduction

It is well known that Efron's bootstrap fails to perform as an useful nonparametric tool when the data exhibit dependence. To resolve this problem some alternatives are proposed. The most popular among them is the one proposed by Künsch (1989), the block bootstrap method. The underlying idea is very simple. Instead of selecting single observation X_i from the sample $\{X_1, \ldots, X_n\}$ with replacement which is what the original bootstrap do, the block bootstrap involves selecting k blocks of consecutive observations of length l. Künsch showed that the block bootstrap correctly estimates the sampling distribution as well as the asymptotic variance of the sample mean, when the sequence of data is assumed to satisfy a α-mixing.

Several results have been established to check the validity of blockwise bootstrap in more general setting since. Recently Radulović (1996), Naik-Nimbalkar and Rajarshi (1994) and Bühlman (1994) have established these types of results under various conditions on dependence structure and block size. Indeed the dependence structure assumed is mainly weakly mixings such as α or β-mixing while the conditions on the block size have been usually set to include the theoretically

[1]Department of Statistics, Keimyung University, Taegu, 704-701, Korea

optimal block size $l(n) = O(n^{1/3})$ (see Künsch). Most of results, however, are restricted to the univariate case. Extension to the multivariate case (or vector valued random variable) has been done by Bühlman (1994) under strong (α-) mixing, which addresses that a sufficient condition on dependence structure for asymptotic validity of block bootstrap is entangled with the dimension p in a undesirable manner. Roughly speaking it could be summarized that performance of the block bootstrap handling multivariate dependent sequence gets worse as the dimension goes up (see (2.3) below). In this paper we study this issue by considering vector valued uniform mixing which is closer to independence than α or β mixing. Our finding is that the disadvantages of the block bootstrap due to dimensionality could be completely removed when the underlying sequences are ϕ-mixing. Since ϕ-mixing could involve many time series model and a data analyst often uses statistics with some finite dimensional marginal in analysis of the time series data, our results may be useful.

This paper consists as follows. Section 2 establishes the asymptotic validity of the block bootstrap method under ϕ-mixing, which resolves dependence complication with dimension. Section 3 contains some technical lemmas which are essential for modification of Bühlman's result.

2. Preliminaries and statement of the main results

In order to establish the validity of the blockwise bootstrap for vector valued time series data, we will mainly follow the Bühlman's approach (1994) with some necessary modifications. Indeed we will consider the blockwise bootstrapped multivariate empirical process and show its weakly convergence to the right Gaussian process under ϕ-mixing. One advantage of this approach is that a large class of statistics can be handled altogether since they often could be written as statistical functionals of the empirical distribution function.

Let $\{\mathbf{X}_i = (X_{i1}, \ldots, X_{ip})'\}_{i \in \mathcal{Z}}$ be a stationary ϕ-mixing sequence of stochastic vectors with mixing coefficients $\phi(n)$ where the marginal distribution function of X_{ij} is denoted by $F^{(j)}$ and the $F^{(j)}$ is assumed to be continuous. Then by the continuous mapping theorem it suffices to consider the transformed empirical process of random vectors $\{\mathbf{Y}_i\}_{i=1}^n$ (instead of $\{\mathbf{X}_i\}_{i=1}^n$) given below;

$$Y_{ij} = F^{(j)}(X_{ij}), \quad j = 1, \ldots, p$$

$$\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{ip})', \quad i \in \mathcal{Z}$$

$$G(\mathbf{t}) = P[\mathbf{Y}_i \leq \mathbf{t}], \quad \mathbf{t} \in E^p$$

where $E^p = \{\mathbf{t}; \mathbf{0} \leq \mathbf{t} \leq \mathbf{1}\}$ is the p-dimensional unit cube. Note that $G^{(j)}(t) = t$ and $G(\mathbf{t}) = 0$ if at least one coordinate of $\mathbf{t}$ is 0 and we set $\mathbf{0} = (0, \ldots, 0)'$, $\mathbf{1} = (1, \ldots, 1)'$, and let $z^{(j)}$ denote the j-th component of a vector $\mathbf{z}$. Further note that $\mathbf{a} \leq \mathbf{b}$ means that $a^{(j)} \leq b^{(j)}$, $j = 1, \ldots, p$ and we write

$$|\mathbf{t} - \mathbf{s}| = \sup\left\{|t^{(j)} - s^{(j)}|; j = 1, \ldots, p\right\}.$$

Using the above definitions, the empirical distribution function of $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ is given by

$$G_n(\mathbf{t}) = n^{-1} \sum_{i=1}^{n} 1_{[\mathbf{Y}_i \leq \mathbf{t}]}, \quad \mathbf{t} \in E^p,$$

and accordingly the empirical process $\{W_n(\mathbf{t})\}_{\mathbf{t} \in E^p}$ is given by

$$W_n(\mathbf{t}) = n^{1/2}\left[G_n(\mathbf{t}) - G(\mathbf{t})\right], \quad \mathbf{t} \in E^p.$$

The weak convergence (denote it by $\Longrightarrow$) of the above empirical processes with respect to the Skorohod J_1-topology is studied by Yokoyama (1980). Indeed he showed if

$$\sum_n \phi(n) < \infty \tag{2.1}$$

then $W_n \Longrightarrow W$, where W is a Gaussian process with

$$E[W(\mathbf{t})] = 0$$

$$E[W(\mathbf{s})W(\mathbf{t})] = \sum_{k=-\infty}^{\infty} E\left[(1_{[\mathbf{Y}_0 \leq \mathbf{s}]} - G(\mathbf{s}))(1_{[\mathbf{Y}_k \leq \mathbf{t}]} - G(\mathbf{t}))\right].$$

To investigate the validity of the blockwise bootstrap, the bootstrapped empirical distribution function of $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ and the bootstrapped empirical process are defined as follows:

$$G_n^*(\mathbf{t}) = k^{-1} \sum_{i=1}^{k} l^{-1} \sum_{j=S_i+1}^{S_i+l} 1_{[\mathbf{Y}_j \leq \mathbf{t}]}, \quad \mathbf{t} \in E^p,$$

where $n = kl$, $l = l(n) = o(n)$ and $l(n) \to \infty$ as $n \to \infty$; S_i i.i.d.$\sim$ Uniform $(\{0, 1, \ldots, n-l\})$ and

$$W_n^*(\mathbf{t}) = n^{1/2}\left[G_n^*(\mathbf{t}) - \mu_n^*(\mathbf{t})\right], \quad \mathbf{t} \in E^p,$$

where

$$\mu_n^*(\mathbf{t}) = E^*\left[G_n^*(\mathbf{t})\right] = E^*\left(l^{-1} \sum_{j=S_i+1}^{S_i+l} 1_{[\mathbf{Y}_j \leq \mathbf{t}]}\right)$$

$$= (n-l+1)^{-1} \sum_{i=0}^{n-l} l^{-1} \sum_{j=i+1}^{i+l} 1_{[\mathbf{Y}_j \le \mathbf{t}]}.$$

Note E^* and Var^* and so on denote the moments under the conditional probability measure P^*, induced by the blockwise resampling scheme. Then the validity of block bootstrap under ϕ-mixing follows from;

Theorem 1. *Let $\{\mathbf{X}_i\}_{i\in\mathcal{Z}}$ be a stationary, uniform mixing sequence whose mixing coefficients satisfy (2.1). Assume that $\mathbf{X}_i$ has continuous marginal distributions and that*

$$l(n) = O\left(n^{1/2-\varepsilon}\right), \quad 0 < \varepsilon < 1/2. \tag{2.2}$$

Then $W_n^* \Longrightarrow W$ almost surely in the Skorohod J_1-topology.

For detailed reference of the Skorohod J_1-topology, reader might refer Billingsley (1968) and Bickel and Wichura (1971). Bühlman's condition for $\alpha-$ mixing (and hence ϕ-mixing) requires

$$\sum_{i=0}^{\infty}(i+1)^{8p+7}\phi^{1/2}(i) < \infty \text{ and } l(n) = (n^{1/2-\varepsilon}), \ \ 0 < \varepsilon < 1/2. \tag{2.3}$$

It is readily seen that dimension p is not a factor for ϕ-mixing in Theorem 1 while the condition (2.3) entangled with the dimension p results a quite restrictive use of bootstrap for multivariate ϕ-mixing variables. Note that α-mixing implies ϕ-mixing. Considering that most time series models in the statistical literature could fit to ϕ-mixing, our result Theorem 1 shows usefulness of the block bootstrap for vector valued time series data.

3. Proof and Technical Lemmas

As is done in Bühlman (1994), proof of Theorem 1 follows if one shows the convergence of the finite dimensional distributions and the tightness for W_n^*. To handle this under ϕ-mixing, all the Lemmas in Bühlman are to be established again under ϕ-mixing. However the only a few of them are reestablished here (the rest of them would be redundant if presented here again) and some Lemmas are added newly. In fact the main contribution of this paper comes from Lemma 3.5 and 3.8. For the complete proof of Theorem 1, refer Bühlman (1994).

Lemma 3.1. *Let $\{X_i\}_{i\in\mathcal{Z}}$ be a stationary, uniform mixing real valued sequence with mixing coefficients $\phi(n)$. Let $\mathcal{F}_a^b$ denote the σ-algebra $\sigma(\{X_i : a \le i \le b\})$.*

If $\xi \in \mathcal{F}_1^k$, $\eta \in \mathcal{F}_{k+n}^\infty$, *then*

$$|E(\xi\eta) - E(\xi)E(\eta)| \leq 2[\phi(n)]^{1/p}\|\xi\|_p\|\eta\|_q,$$

where $1 \leq p, q \leq \infty$ *with* $1/p + 1/q = 1$ *and* $\|\cdot\|_r = \{E(|\cdot|^r)\}^{1/r}$.

Lemma 3.2. *Suppose* $\{X_i\}$ *is a strictly stationary sequence of* ϕ*-mixing random variables. Assume* $EX_i = 0$, $|X_i| \leq C$ *and (2.1) holds. Then for any positive number* $g > 1$ *there exists a constant* K *such that*

$$E\left|\sum_{i=1}^{n} X_i\right|^{2g} < Kn^g, \quad n = 1, 2, \ldots.$$

Proof. The proof follows from arguing as in Doob (1953), pp. 225-226. □

For the rest of this section, we define some quantities that were used in Bühlman. Let $G(\mathbf{s},\mathbf{t}) = G(\mathbf{t}) - G(\mathbf{s})$, $I_j(\mathbf{s},\mathbf{t}) = 1_{[\mathbf{Y}_j \leq \mathbf{t}]} - 1_{[\mathbf{Y}_j \leq \mathbf{s}]}$ and $\mu_n^*(\mathbf{s},\mathbf{t}) = \mu_n^*(\mathbf{t}) - \mu_n^*(\mathbf{s})$. Define

$$Z_i(\mathbf{s},\mathbf{t}) = \left[l^{-1}\sum_{j=i+1}^{i+l}\{I_j(\mathbf{s},\mathbf{t}) - G(\mathbf{s},\mathbf{t})\}\right]^2 - E\left|l^{-1}\sum_{j=i+1}^{i+l}\{I_j(\mathbf{s},\mathbf{t}) - G(\mathbf{s},\mathbf{t})\}\right|^2$$

$$V(\mathbf{s},\mathbf{t}) = [\mu_n^*(\mathbf{s},\mathbf{t}) - G(\mathbf{s},\mathbf{t})]^2 - E\left|\mu_n^*(\mathbf{s},\mathbf{t}) - G(\mathbf{s},\mathbf{t})\right|^2. \tag{3.1}$$

Further let

$$B = B(\mathbf{b}_0, \delta) = \{\mathbf{t}; \mathbf{b}_0 \leq \mathbf{t} \leq \mathbf{b}_0 + \delta\mathbf{1}\}$$

where we assume $\delta \in \mathcal{Q}$, $\mathbf{b}_0 \in \mathcal{Q}^p \cap E^p$ without loss of generality. Now we fix the hyperquardrangle B and consider the following grid (for a fixed n):

$$\mathbf{b}_n(\mathbf{i}) = \mathbf{b}_0 + \frac{\kappa}{6p}n^{-1/2}\mathbf{i}, \quad \mathbf{0} \leq \mathbf{i} \leq \mathbf{m}_n, \quad \kappa > 0,$$

where $\mathbf{m}_n = m_n\mathbf{1}$, $m_n = \left[6p\delta n^{1/2}/\kappa\right] - 1$ and for $\mathbf{i}$ with $i^{(j)} = m_n + 1$, we define $\mathbf{b}_n(\mathbf{i})^{(j)} = \mathbf{b}_0^{(j)} + \delta$, where $\mathbf{z}^{(j)}$ denotes the j-th component of a $p \times 1$ vector $\mathbf{z}$. Let $H_i(n)$, $(i = 1, \ldots, N(n) = O(n^p))$, denote a p-dimensional hyper-rectangle whose conner points are points from the grid of B. In the following we identify functions with hyper-rectangles H as arguments in the obvious way via the basic quantity $1_{[\mathbf{Y} \in H]}$. Let

$$D_j(H_i(n)) = l^{-1}\sum_{t=S_j+1}^{S_j+l} 1_{[\mathbf{Y}_t \in H_i(n)]}, \quad j = 1, \ldots, k.$$

Then

$$W_n^*(H_i(n)) = n^{1/2}k^{-1}\sum_{j=1}^{k}\{D_j(H_i(n)) - E^*\left[D_j(H_i(n))\right]\}$$

Note that $D_1, \ldots, D_k$ are i.i.d. under P^*. Let G be the distribution function with the corresponding probability measure P. Now we have the following two Lemmas:

Lemma 3.3. *If (2.1) and (2.2) hold, then*

$$Cov^*\left[W_n^*(\mathbf{s}), W_n^*(\mathbf{t})\right] = Cov\left[W(\mathbf{s}), W(\mathbf{t})\right] + \Delta_{\mathbf{s},\mathbf{t}}(n),$$

where $\Delta_{\mathbf{s},\mathbf{t}}(n) = o(1)$ *almost surely.*

Lemma 3.4. *Assume that (2.1) and (2.2) hold. Then there exists a sufficiently large positive integer* m *such that*

$$nk^{-1}E^*\left|D_1\left(H_i(n)\right) - E^*\left[D_1\left(H_i(n)\right)\right]\right|^2 \le G(H_i(n))^{1/2m}(const. + \Delta_i(n)),$$

where $\sup_{i\le N(n)}|\Delta_i(n)| = o(1)$ *almost surely.*

Proof of the above two Lemmas basically follow from Lemma 3.5 below. For the details of their proofs, see Lemmas 2 and 5 of Bühlman. □

Lemma 3.5. *If (2.1) and (2.2) hold, then for any positive integer* $m > 1$,

$$E\left|\sum_{i=0}^{n-l} Z_i\right|^{2m} \le const.\ n^m l^{-m} \tag{3.2}$$

$$E\left|V\right|^{2m} \le const.\ n^{-2m} \tag{3.3}$$

where Z_i *and* V *are given in (3.1).*

Proof. Let $Q_j = I_j(\mathbf{s},\mathbf{t}) - G(\mathbf{s},\mathbf{t})$. Under the condition (2.1), it is easily seen that by Lemmas 3.1 and 3.2, we have

$$E\left|\sum_{j=1}^{l} Q_j\right|^2 \le \text{const. } l$$

and for any g

$$E\left|\sum_{j=1}^{l} Q_j\right|^{2g} \le \text{const.} l^g.$$

Using these results with Minkowski inequality, we have for any positive integer $m \geq 1$

$$E|Z_0|^m = E\left|\left[l^{-1}\sum_{j=1}^{l} Q_j\right]^2 - E\left|l^{-1}\sum_{j=1}^{l} Q_j\right|^2\right|^m \leq \text{const.} l^{-m}. \tag{3.4}$$

Recalling $kl = n$, rearrange the Z_i's into the following blocks, i.e.,

$$E\left|\sum_{i=0}^{n-l} Z_i\right|^{2m} = E\left|Z_0 + \sum_{j=0}^{k-2} Z_{jl+1} + \sum_{j=0}^{k-2} Z_{jl+2} + \cdots + \sum_{j=0}^{k-2} Z_{jl+l}\right|^{2m}$$

$$\leq \left\{\sum_{i=1}^{l-1}\left(E\left|\sum_{j=0}^{k-2} Z_{jl+i}\right|^{2m}\right)^{1/2m} + \left(E\left|\sum_{j=0}^{k-2} Z_{jl+l} + Z_0\right|^{2m}\right)^{1/2m}\right\}^{2m}. \tag{3.5}$$

Observe that with the given definition of Z_i, a sequence of blocks $\{W_j\}_{j\in Z} = l^2\{Z_{jl+i}\}_{j\in Z}$ for a fixed $i = 0, 1, \ldots, l$ is also a stationary and uniform mixing process, whose mixing coefficients satisfy

$$\phi_W(j) = \begin{cases} \phi_X(1), & \text{j=1} \\ \phi_X((j-1)l), & \text{j=2,}\ldots, \end{cases}. \tag{3.6}$$

Now an application of Lemma 3.1 with $p = 1$ and $q = \infty$ and (3.4) yields

$$Var(\sum_{j=0}^{k-2} W_j) \leq kEW_0^2 + kl^2 \sum_{i=1}^{k-1} \phi(li) E|W_0|$$

$$\leq \text{const.} kl^2 + kl^3 \sum_{i=1}^{k-1} \phi(li)$$

$$\leq \text{const.} kl^2 + kl^2 \sum_{i=1}^{\infty} \phi(i) \leq \text{const.} kl^2.$$

Using the above and arguing as in Doob (1953) (pp 225-226) with l being a fixed constant, we obtain

$$E(\sum_{j=0}^{k} W_j)^{2m} \leq \text{const.}(E|W_0|^{2m} + l^{2m})k^m$$

$$\leq \text{const.} l^{2m} k^m$$

for all $k = 1, 2, \cdots$. In the above inequality variance calculation and (3.4) is used (see Doob). Hence one has

$$E(\sum_{j=0}^{k} Z_j)^{2m} \leq \text{const.} l^{-2m} k^m. \tag{3.7}$$

Thus (3.5) is bounded by

$$\text{const. } \left(lk^{1/2} l^{-1}\right)^{2m} \leq \text{const. } k^m = \text{const. } n^m l^{-m}.$$

Thus verification of (3.2) is finished.

For the verification of (3.3) let $U = [\mu_n^*(\mathbf{s}, \mathbf{t}) - G(\mathbf{s}, \mathbf{t})]^2$. Then it follows that

$$E\,|V|^{2m} = E\,|U - EU|^{2m} = \sum_{r=0}^{2m} \binom{2m}{r} EU^r (EU)^{2m-r}. \tag{3.8}$$

Notice

$$\begin{aligned} EU^r &= E\left[(\mu_n^*(\mathbf{s}, \mathbf{t}) - G(\mathbf{s}, \mathbf{t}))^2\right]^r \\ &= E\left[(n - l + 1)^{-1} \sum_{i=0}^{n-l} l^{-1} \sum_{j=i+1}^{i+l} \{I_j(\mathbf{s}, \mathbf{t}) - G(\mathbf{s}, \mathbf{t})\}\right]^{2r}. \end{aligned}$$

To obtain an appropriate bound for EU^r, we may follow similar steps taken in the verification of (3.2). Indeed set $T_i = l^{-1} \sum_{j=i+1}^{i+l} \{I_j(\mathbf{s}, \mathbf{t}) - G(\mathbf{s}, \mathbf{t})\}$. Then we have by Lemma 3.2

$$E\,|T_0|^{2r} = E\left|l^{-1} \sum_{j=1}^{l} \{I_j(\mathbf{s}, \mathbf{t}) - G(\mathbf{s}, \mathbf{t})\}\right|^{2r} \leq \text{const.} l^{-r}$$

and hence similar argument leading to (3.7) yields

$$\begin{aligned} E|\sum_{i=0}^{n-l} T_i|^{2r} &\leq \left\{ \sum_{i=1}^{l-1} \left(E\left|\sum_{j=0}^{k-2} T_{jl+i}\right|^{2r} \right)^{1/2r} + \left(E\left|\sum_{j=0}^{k-1} T_{jl}\right|^{2r} \right)^{1/2r} \right\}^{2r} \\ &\leq \mathit{const.} (lk^{1/2} l^{-1/2})^{2r} = \mathit{const.} n^r. \end{aligned}$$

Now we have under (2.1) for any positive integer r

$$EU^r \leq \text{const}.n^{-r} \tag{3.9}$$

Now (3.3) follows from (3.8)-(3.9). □

Other Lemmas that need our attention for proper modifications of their proofs include the following three Lemmas.

Lemma 3.6. *Assume that (2.1) and (2.2) hold. Let* $\mu(c) = G(c) + \lambda(c)$, *where* λ *denotes the Lebesque measure in* E^p, *and* C *a Borel set in* E^p. *Then there exists an* $r = r(p) \in \mathcal{N}$ *such that*

$$E^*|W_n^*(H_i(n))|^{2r+2} \leq [(const. + \xi_i(n))\mu(H_i(n))]^\beta, \quad \beta > 1$$

where $\sup_{i \leq N(n)} |\xi_i(n)| \leq K$ *almost surely (K a constant).*

Proof. Let $\tau = nk^{-1}E^*|D_1(H_i(n)) - E^*[D_1(H_i(n))]|^2$. According to Bühlmann (1994) (see Lemma 7 there), it is easy to find that

$$E^*|W_n^*(H_i(n))|^{2r+2} \leq K_r \sum_{j=1}^{r+1} (ln^{-1})^{r+1-j}\tau^j.$$

Since $\lambda(H_i(n)) \geq \text{const}.n^{-p/2}$ there exists a $m_1 > p/2$ such that $ln^{-1} \leq [\lambda(H_i(n))]^{1/m_1}$ for n sufficiently large and $i \leq N(n)$. Therefore $(ln^{-1})^{m_1+s} \leq [\lambda(H_i(n))]^{(m_1+s)/m_1} = [\lambda(H_i(n))]^{\beta_1}$ ($\beta_1 > 1$) for $i \leq N(n)$, $s \geq 1$, n sufficiently large. On the other hand we have by Lemma 3.4 there exists a m_2 such that $\tau \leq [G(H_i(n))]^{1/2m_2}$ (const. + $\Delta_i(n)$). Thus $\tau^{2m_2+s} \leq$ [const. $\mu(H_i(n))]^{\beta_2}$ ($\beta_2 > 1$) for all $i \leq N(n)$ a.s., for $s \geq 1$. Take $r = m_1 + 2m_2$. This finishes the proof. □

Lemma 3.7. *Assume that (2.1) and (2.2) hold. Then*

$$\max_{0 \leq \mathbf{i} \leq \mathbf{m}_n} n^{1/2} |\mu_n^*(\mathbf{b}_n(\mathbf{i}+\mathbf{1})) - \mu_n^*(\mathbf{b}_n(\mathbf{i}))| < \kappa/3 + o(1),$$

where the o(1)-term is almost surely.

Proof. Denote by $\mu_n^*(\mathbf{t},\mathbf{s}) = \mu_n^*(\mathbf{t}) - \mu_n^*(\mathbf{s})$. It is easy to see that:

$$\max_{0 \leq \mathbf{i} \leq \mathbf{m}_n} n^{1/2} |\mu_n^*(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))|$$

$$\leq \max_{0\leq \mathbf{i}\leq \mathbf{m}_n} n^{1/2} \left|\mu_n^*(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i})) - G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\right|$$
$$+ \max_{0\leq \mathbf{i}\leq \mathbf{m}_n} n^{1/2} \left|G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\right|$$
$$\leq \max_{0\leq \mathbf{i}\leq \mathbf{m}_n} n^{1/2} \left|\mu_n^*(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i})) - G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\right| + \kappa/3.$$

The last step holds since

$$\max_{0\leq \mathbf{i}\leq \mathbf{m}_n} n^{1/2} \left|G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\right| \leq \kappa/3. \tag{3.10}$$

Thus it is necessary to show that

$$\max_{0\leq \mathbf{i}\leq \mathbf{m}_n} n^{1/2} \left|\mu_n^*(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i})) - G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\right| = o(1) \quad a.s. \tag{3.11}$$

Consider

$$E\left[\mu_n^*(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i})) - G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\right]^{2m}$$
$$= E\left[(n-l+1)^{-1}\sum_{t=0}^{n-l} l^{-1} \sum_{j=t+1}^{t+l} I_j(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i})) - G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\right]^{2m}$$
$$\leq (n-l+1)^{-2m}\left(E\left|\sum_{t=0}^{n-l} S_t\right|^{2m}\right),$$

where $S_t = l^{-1}\sum_{j=t+1}^{t+l}\{I_j(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i})) - G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\}$.

To obtain an upper bound for $E\left|\sum_{t=0}^{n-l} S_t\right|^{2m}$, apply similar argument as in proof of Lemma 3.5. Indeed with (3.10) at hand, one may apply Lemma 3.8 to get

$$ES_0^{2m} \leq \text{const.}\, l^{-2m} l^{(1-\tau)m} \tag{3.12}$$

for some $0 < \tau < 1/2$. Using the blocking technique previously employed again, we have

$$E\left|\sum_{i=0}^{n-l} S_i\right|^{2m} = E\left|S_0 + \sum_{j=0}^{k-2} S_{jl+1} + \sum_{j=0}^{k-2} S_{jl+2} + \cdots + \sum_{j=0}^{k-2} S_{jl+l}\right|^{2m}$$
$$\leq \left\{\sum_{i=1}^{l-1}\left(E\left|\sum_{j=0}^{k-2} S_{jl+i}\right|^{2m}\right)^{1/2m} + \left(E\left|\sum_{j=0}^{k-2} S_{jl+l} + S_0\right|^{2m}\right)^{1/2m}\right\}^{2m}. \tag{3.13}$$

Now Lemma 3.1 with $p = 1$ and $q = \infty$, and (3.12) yields

$$Var(\sum_{j=0}^{k-2} S_j) \leq kES_0^2 + \text{const.}k \sum_{i=1}^{k-1} \phi(li)E|S_0|$$

$$\leq \text{const.}kl^{-1-\tau} + \text{const.}kl^{-1}l^{1-\tau} \sum_{i=1}^{k-1} \phi(li)$$

$$\leq \text{const.}kl^{-1-\tau} + \text{const.}kl^{-1-\tau} \sum_{i=1}^{\infty} \phi(i) \leq \text{const.}kl^{-1-\tau}.$$

Using the above and arguing as in Doob (1953) (pp 225-226) with l being a fixed constant, we obtain

$$E(\sum_{j=0}^{k} S_j)^{2m} \leq \text{const.}(E|S_0|^{2m} + l^{(-1-\tau)m})k^m$$

$$\leq \text{const.}l^{(-1-\tau)m}k^m. \tag{3.14}$$

Thus (3.13) is bounded by

$$\text{const.} \left(lk^{1/2}l^{(-1-\tau)/2}\right)^{2m} \leq \text{const.}\ l^{(1-\tau)m}k^m,$$

which in turn implies

$$(n-l+1)^{-2m}E\left|\sum_{t=0}^{n-l} S_t\right|^{2m}$$

$$\leq (n-l+1)^{-2m}\text{const.}l^m k^m l^{-\tau m} \leq \text{const.}n^{-m}l^{-\tau m}.$$

Now using the above, one can show that

$$P\left[\max_{\mathbf{0}\leq\mathbf{i}\leq\mathbf{m}_n} n^{1/2}\left|\mu_n^*(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i})) - G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\right| > \rho\right]$$

$$\leq m_n^p E\left|n^{1/2}\left(\mu_n^*(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i})) - G(\mathbf{b}_n(\mathbf{i}+\mathbf{1}), \mathbf{b}_n(\mathbf{i}))\right)\right|^{2m}$$

$$\leq m_n^p n^m\{\text{const.}n^{-m}l^{-\tau m} = O(m_n^p l^{-\tau m})$$

for a sufficiently large m. By Borel-Cantelli Lemma, (3.11) follows. □
The following lemma is due to Kim (1999).

Lemma 3.8. *Let Z_i be a stationary process and $\phi-$mixing with $E(Z_i) = 0$, $P(|Z_i| > c) = 0$, and $E|Z_i|^2$, $E|Z_i| \leq cl^{-\tau}$ for some $0 < \tau < 1/2$. Let m be any positive integer. If (2.1) holds, then for all $l \geq 1$,*

$$E(\sum_{i=1}^{l} Z_i)^{2m} \leq const.l^{(1-\tau)m}.$$

Acknowledgement

This work was supported by Korea Reserach Foundation (KRF-1999-015-DP0054).

References

Billingsley, P. (1968). *Convergence of Probability Measures,* Wiley, New York.

Bickel, P. J. and Wichura, M. J. (1971). "Convergence criteria for multiparameter stochastic processes and some applications," *The Annals of Mathematical Statistics* **42**, 1656-1670.

Bühlmann, P. (1994). "Blockwise bootstrapped empirical process for stationary sequences," *The Annals of Statistics* **22**, 995-1012.

Doob J. L. (1953). "Stochastic Processes", *New York, Wiley.*

Kim, T. Y. (1999). "On tail probabilities of Kolmogorov-Smirnov statistics based on uniform mixing processes," *Statistics and Probability Letters* **43**, 217-223.

Künsch, H. R. (1989). "The jackknife and the bootstrap for general stationary observations," *The Annals of Statistics* **17**, 1217-1241.

Naik-Nimbalkar, U. V. and Rajarshi, M. B. (1994). "Validity of blockwise bootstrap for empirical processes with stationary observations," *The Annals Statistics* **22**, 980-994.

Radulović, D. (1996). "The bootstrap for empirical processes based on stationary observations," *Stochastic Processes and their Applications* **65**, 259-279.

Yokoyama, R. (1980). "A note on weak convergence of empirical processes for $\phi-$mixing random variables," *Yokohama Mathematics* **28**, 37-40.