Australian & New Zealand Journal of Statistics

Aust. N. Z. J. Stat. **60**(1), 2018, 103–114 doi: 10.1111/anzs.12206

Consistency of a hybrid block bootstrap for distribution and variance estimation for sample quantiles of weakly dependent sequences

Todd A. Kuffner[1,*] **Stephen M. S. Lee**[2] **and G. A. Young**[3]

Washington University in St. Louis, The University of Hong Kong and Imperial College London

Summary

Consistency and optimality of block bootstrap schemes for distribution and variance estimation of smooth functionals of dependent data have been thoroughly investigated by Hall, Horowitz & Jing (1995), among others. However, for nonsmooth functionals, such as quantiles, much less is known. Existing results, due to Sun & Lahiri (2006), regarding strong consistency for distribution and variance estimation via the moving block bootstrap (MBB) require that $b \to \infty$, where $b = \lfloor n/\ell \rfloor$ is the number of resampled blocks to be pasted together to form the bootstrap data series, n is the available sample size, and ℓ is the block length. Here we show that, in fact, weak consistency holds for any b such that $1 \leqslant b = O(n/\ell)$. In other words we show that a hybrid between the subsampling bootstrap ($b = 1$) and MBB is consistent. Empirical results illustrate the performance of hybrid block bootstrap estimators for varying numbers of blocks.

Key words: asymptotics; nonsmooth functional; resampling; strong mixing; subsampling

1. Introduction

We consider a problem closely related to some of Peter's greatest contributions in bootstrap theory and methodology. Specifically we address the problem of choosing the number of blocks when using block bootstrap methods to estimate the distribution and variance of sample quantiles computed from weakly dependent sequences. Peter started thinking about a version of this problem in the mid-1980s. In Hall (1985) he proposed resampling spatial patterns, and introduced the core concepts that would eventually be implemented in the moving block bootstrap (MBB). Peter called this a 'tiled bootstrap', as it was based on the idea of partitioning an observation region into congruent, nonoverlapping tiles. He considered both 'fixed' and 'moving' tiles resampling schemes. Carlstein (1986) utilized similar ideas for the problem of variance estimation for a statistic computed from time series data. The key insight came from Hall (1985) and the subsequent papers on the MBB by Künsch (1989) and Liu & Singh (1992). Suppose that the observations are generated by a weakly dependent process. The dependence structure in the observations is preserved within blocks. Thus if the block length increases with sample size, then the dependence structure in the underlying process will be reproduced asymptotically by the MBB. Peter also played an important role in the ultimate resolution of

*Author to whom correspondence should be addressed.
[1]Department of Mathematics, Washington University in St. Louis, St. Louis, MO 63130, USA
[2]Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong City, Hong Kong
[3]Department of Mathematics, Imperial College London, London, SW7 2AZ, UK e-mail: kuffner@wustl.edu

© 2018 Australian Statistical Publishing Association Inc. Published by John Wiley & Sons Australia Pty Ltd.

this problem when dealing with smooth functionals of dependent data. Hall *et al.* (1995) is part of a series of papers which rigorously established the theory of optimal block bootstrap distribution and variance estimation for smooth functionals. Other work on this topic includes Bühlmann (1994), Naik-Nimbalkar & Rajarshi (1994), Götze & Künsch (1996), Lahiri (1992, 1996, 1999) and Bühlmann & Künsch (1999). Included among Peter's contributions to block bootstrap theory are Davison & Hall (1993), Hall & Horowitz (1996), Carlstein *et al.* (1998), Hall, Jing & Lahiri (1998) and Choi & Hall (2000). We also mention that, while Peter's intellectual contributions to the foundations of the MBB are undeniable, he did not seek to be recognized as its creator. Peter wrote an entire paper (Hall 2003) giving credit for the inspiration of the block bootstrap to P.C. Mahalanobis for his work on the analysis of jute production data in Bengal, India, during the 1930s and 1940s.

Sun (2004, 2007), Sun & Lahiri (2006) and Sharipov & Wendler (2013) investigated different aspects of the less well-studied setting of block bootstrap methods for nonsmooth functionals of dependent data. Of greatest relevance to our results, Sun & Lahiri (2006) showed the consistency, under mild moment conditions and polynomial mixing weak dependence, of MBB distribution and variance estimators for sample quantiles. On this topic, Hall & Jing (1996, p.132) offered only the comment that '… estimation of quantile variance is really a problem of curve estimation …', another topic on which Peter was an expert. Surprisingly, however, this particular setting of the block bootstrap is one in which Peter never published a paper – as far as we know.

The main contributions of the current paper are as follows. In Theorem 3 of Section 3, we prove that a 'hybrid' between MBB ($b = \lfloor n/\ell \rfloor$) and the subsampling ($b = 1$) bootstrap is consistent, under the assumption of a strongly mixing sequence with exponentially decaying coefficients. The proof of weak consistency, for $1 \leqslant b = O(n/\ell)$, covers both distribution estimation and variance estimation for sample quantiles. Our theoretical results are also validated empirically with three commonly-encountered models which satisfy our conditions.

2. Background

Let $\{Y_i\}_{i\in\mathbb{Z}}$ denote a strictly stationary process, where $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ is the set of all integers. The sequence $(Y_1, \ldots, Y_n)$ denotes a sample of size n from $\{Y_i\}_{i\in\mathbb{Z}}$. Suppose that the random variables $\{Y_i\}_{i\in\mathbb{Z}}$ are defined on a common probability space $(\Omega, \mathcal{F}, P)$, with common marginal distribution function F.

2.1. The MBB and subsampling bootstrap

The MBB (Künsch 1989) divides the sample $(Y_1, \ldots, Y_n)$ into overlapping blocks of size ℓ, $B_i = (Y_i, \ldots, Y_{i+\ell-1})$, yielding a set $\{B_1, \ldots, B_{n-\ell+1}\}$. Let $B_1^*, \ldots, B_b^*$ be a random sample drawn with replacement from the original blocks, where $b = \lfloor n/\ell \rfloor$ is the number of blocks that will be pasted together to form a pseudo-time series. The notation $\lfloor h \rfloor$ means the largest integer $\leqslant h$, and similarly $\lceil h \rceil$ is the smallest integer $\geqslant h$. That $B_1^*, \ldots, B_b^*$ is a random sample from $\{B_1, \ldots, B_{n-\ell+1}\}$ means that the sampled blocks are independently and identically distributed according to a discrete uniform distribution on $\{B_1, \ldots, B_{n-\ell+1}\}$. The observations in the ith resampled block, B_i^*, are $Y^*_{(i-1)\ell+1}, \ldots, Y^*_{i\ell}$, for $1 \leqslant i \leqslant b$. The resulting MBB sample is the concatenation of the resampled blocks, written as

© 2018 Australian Statistical Publishing Association Inc.

$$\underbrace{Y_1^*,\ldots,Y_\ell^*}_{B_1^*},\ \underbrace{Y_{\ell+1}^*,\ldots,Y_{2\ell}^*}_{B_2^*},\ \underbrace{Y_{2\ell+1}^*,\ldots,Y_{(b-1)\ell}^*}_{B_3^*,\ldots,B_{(b-1)}^*},\ \underbrace{Y_{(b-1)\ell+1}^*,\ldots,Y_{b\ell}^*}_{B_b^*}.$$

The overlapping blocks version of the subsampling bootstrap (Politis & Romano 1994) also divides the sample into the same overlapping blocks as the MBB. However, the subsampling bootstrap randomly draws only a single block, which is viewed as a partial observation of the original series. The subsampling bootstrap thus exactly retains the dependence structure in the original sample. Hence, the subsampling bootstrap is a special case of the MBB for which only a single block of length ℓ is resampled.

2.2. Some notation

The common marginal distribution function F is in particular the distribution function of Y_1, i.e. $F(x)=\Pr(Y_1\leqslant x)$, $x\in\mathbb{R}$. The corresponding quantile function $F^{-1}(p)$ is defined by

$$F^{-1}(p)=\inf\{x:F(x)\geqslant p\},\qquad 0<p<1.$$

The empirical distribution function of the sample $(Y_1,\ldots,Y_n)$, for $n\geqslant 1$, is denoted by F_n, and puts probability mass $1/n$ on each element,

$$F_n(x)=n^{-1}\sum_{i=1}^{n}\mathbf{1}(Y_i\leqslant x),\qquad x\in\mathbb{R},$$

where $\mathbf{1}(\text{Event})=1$ or 0 according to whether the Event occurs or does not occur, respectively. Define, for $\ell\in\{1,2,\ldots,n\}$, $b\in\{1,2,\ldots\}$ and $x\in\mathbb{R}$, $J_1,\ldots,J_b$ to be independent random indices uniformly drawn from the set $\{1,\ldots,n-\ell+1\}$,

$$U_i(x)=\ell^{-1}\sum_{t=i}^{i+\ell-1}\mathbf{1}\{Y_t\leqslant x\},\ i=1,\ldots,n-\ell+1,$$

and

$$U_i^*(x)=\ell^{-1}\sum_{t=J_i}^{J_i+\ell-1}\mathbf{1}\{Y_t\leqslant x\},\ i=1,\ldots,b.$$

The $U_i^*(x)$ just defined are the (conditionally i.i.d.) resampled block averages. Further define

$$\tilde{F}_n(x)=(n-\ell+1)^{-1}\sum_{i=1}^{n-\ell+1}U_i(x)\qquad\text{and}\qquad F_n^*(x)=b^{-1}\sum_{i=1}^{b}U_i^*(x).$$

The latter quantity is the hybrid block bootstrap empirical distribution function. Define, for $p\in(0,1)$,

$$\xi_p=F^{-1}(p),\qquad \hat{\xi}_n=F_n^{-1}(p),\qquad \tilde{\xi}_n=\tilde{F}_n^{-1}(p)\qquad\text{and}\qquad \xi_n^*=F_n^{*-1}(p).$$

Again, the last quantity, ξ_n^*, is the hybrid block bootstrap version of the sample quantile, $\hat{\xi}_n$. Assume that $f=F'$ is defined on a neighbourhood $\mathcal{N}_p$ of ξ_p, with

$$0<\inf_{x\in\mathcal{N}_p}f(x)\leqslant\sup_{x\in\mathcal{N}_p}f(x)<\infty.$$

© 2018 Australian Statistical Publishing Association Inc.

The hybrid block bootstrap version of the centered and scaled sample quantile, $\sqrt{n}(\hat{\xi}_n - \xi_p)$, is given by $\sqrt{b\ell}(\xi_n^* - \tilde{\xi}_n)$.

The centering is about $\tilde{\xi}_n$, which is the sample quantile analogue of the centering constant used for statistics adhering to the smooth function model, such as the sample mean. For this to be well-defined, we need $\tilde{F}_n$ to be a valid distribution function. This follows from the fact that F_n^* is a valid distribution function for each set of resampled $\{Y_1^*,\ldots,Y_{bl}^*\}$, and that $\tilde{F}_n$ is simply the conditional expectation of F_n^*, given $(Y_1,\ldots,Y_n)$. Define $G_n(u) = \Pr(\sqrt{n}(\hat{\xi}_n - \xi_p) \leqslant u)$ and $\tau_n = \mathrm{var}\{\sqrt{n}(\hat{\xi}_n - \xi_p)\}$. It is our purpose to establish weak consistency of a hybrid block bootstrap for estimation of $G_n(u)$ and τ_n.

We now explain the weak dependence condition on the Y_i's. Weak dependence among the Y_i's means that the dependence between Y_i and Y_j decays in a specified way as $|i-j|$ increases. Conventionally, a strong mixing condition, which is defined with respect to the σ-algebras generated by sequences of the Y_i's, is used. Define $\mathcal{F}_k^t$ to be the σ-algebra generated by the random variables $Y_k, Y_{k+1},\ldots,Y_t$, $-\infty \leqslant k \leqslant t \leqslant \infty$. For $t \geqslant 1$, define

$$\alpha(t) = \sup_{k\in\mathbb{Z}} \sup_{A\in\mathcal{F}_{-\infty}^k, B\in\mathcal{F}_{k+t}^{\infty}} |Pr(A\cap B) - Pr(A)Pr(B)|.$$

The sequence $\{Y_i\}_{i\in\mathbb{Z}}$ is called strongly mixing or α-mixing under the condition that $\alpha(t)\to 0$ as $t\to\infty$. The mixing rate is called polynomial when $\alpha(t) = O(t^{-\beta})$, for some suitable $\beta\in(0,\infty)$. In Theorem 1, we assume that the mixing coefficient $\alpha(t)$ decays at an exponential rate. That is, for some $C>0$,

$$\alpha(t) = O(e^{-Ct}), \qquad t\to\infty.$$

Though this assumption is slightly stronger than the polynomial rate assumption in Sun & Lahiri (2006), it is sufficiently undemanding to include the practically important cases, such as a broad class of linear processes, as we illustrate with examples in Section 4.

2.3. Connections to existing results

The novelty of our theoretical results is that we allow b to vary, and hence our consistency result encompasses the subsampling bootstrap, the MBB, and everything in between. Earlier authors have considered narrower consistency results, regarding particular block bootstrap schemes. For example, Sun & Lahiri (2006, theorem 3.1) showed that, under a suitable polynomial mixing rate, with $\ell^{-1} = o(1)$ and $\ell = O(n^{1/2-\eta})$ for some $\eta\in(5/(2+4\beta), 1/2)$, the MBB is strongly consistent for estimating the distribution of the centered and scaled sample quantile, $\sqrt{n}(F_n^{-1}(p) - F^{-1}(p)) \equiv \sqrt{n}(\hat{\xi}_n - \xi_p)$. When the mixing coefficient satisfies $\alpha(t) \leqslant Cn^{-\beta}$ for some $C>0$ and $\beta>9.5$,

$$\sup_{x\in\mathbb{R}} |\Pr(\sqrt{n}(\xi_n^* - \tilde{\xi}_n) \leqslant x \,\big|\, Y_1,\ldots,Y_n) - \Pr(\sqrt{n}(\hat{\xi}_n - \xi_p) \leqslant x)| = o(1),$$

with probability 1. Under the same conditions, with the additional assumption that $\mathrm{E}(|Y_1|^\delta) < \infty$ for some $\delta>0$, theorem 3.2 of Sun & Lahiri (2006) established strong consistency of the MBB estimator of the asymptotic variance of $\sqrt{n}(\hat{\xi}_n - \xi_p)$. Somewhat less relevant to our results, the consistency properties of the circular block bootstrap and smoothed extended tapered block bootstrap were proven by, respectively, Sharipov & Wendler (2013) and

© 2018 Australian Statistical Publishing Association Inc.

Gregory, Lahiri & Nordman (2015). Our main result is a consistency theorem allowing the number of blocks b to be anything from 1 to $\lfloor n/\ell \rfloor$, and can be viewed as a quite general result about consistency of block bootstrap procedures for sample quantiles under a reasonably weak dependence assumption.

3. Main results

The following theorem establishes weak consistency of the block bootstrap for the distribution and variance of sample quantiles, based on any $b = O(n/\ell)$.

Theorem 1. *Suppose that $n^{-\delta}\ell \to \infty$, $\ell = O(n^{1/2-\delta})$ for any arbitrarily small $\delta \in (0, 1/2)$, and $1 \leqslant b = O(n/\ell)$. Assume that the mixing coefficient satisfies $\alpha(t) = O(e^{-Ct})$, $t \to \infty$, for some $C > 0$.*

(i) The equality

$$Pr\Big(\sqrt{b\ell}\big(\xi_n^* - \tilde{\xi}_n\big) \leqslant x \Big| Y_1, \ldots, Y_n\Big) = Pr\Big(\sqrt{n}\big(\hat{\xi}_n - \xi_p\big) \leqslant x\Big) + o_p(1)$$

holds for each fixed $x \in \mathbb{R}$.

(ii) If $\mathbf{E}(|Y_1|^\alpha) < \infty$ for some $\alpha > 0$, then

$$Var\big(\sqrt{b\ell}\,\xi_n^* \big| Y_1, \ldots, Y_n\big) = Var\big(\sqrt{n}\hat{\xi}_n\big) + o_p(1).$$

The proof of Theorem 1 is given in the Appendix.

Remark 1. Optimal choice of (b, ℓ) is established, under the weaker assumption of a polynomial mixing rate, in Kuffner, Lee & Young (2017).

Remark 2. In the examples of Section 4, we examine coverages of two-sided bootstrap CIs with nominal levels 95% and 90% for ξ_p. The bootstrap confidence intervals are constructed as follows. Define q_s by $G_n(q_s) = s$. Then

$$(\hat{\xi}_n - n^{-1/2}q_{1-\alpha},\ \hat{\xi}_n - n^{-1/2}q_\alpha)$$

constitutes an exact $100(1-2\alpha)\%$ confidence interval for ξ_p. Define $G_n^*(u) = \Pr(\sqrt{b\ell}(\xi_n^* - \hat{\xi}_n) \leqslant u \big| Y_1, \ldots, Y_n)$, and q_s^* by $G_n^*(q_s) = s$. Then the bootstrap confidence interval of nominal coverage $100(1-2\alpha)\%$ is

$$(\hat{\xi}_n - n^{-1/2}q_{1-\alpha}^*,\ \hat{\xi}_n - n^{-1/2}q_\alpha^*).$$

We do not assume, *a priori*, that we have available a suitable variance estimator with which to studentize the sample quantile, which would permit the use of intervals based on percentiles of the t-distribution. It is for this reason that we have elected to use the foregoing simple percentile intervals instead.

We now present some empirical evidence supporting the theoretical results. Further theoretical developments are contained in forthcoming, more technical work (Kuffner *et al.* 2017) concerning the optimal rates of convergence of the bootstrap distribution estimator in this problem, for both polynomial and exponential mixing rates. Preliminary results indicate that for the exponential mixing rate context, the optimal ℓ and b are both of order $O(n^{1/3})$. For this reason, we have chosen to focus on setting ℓ as $n^{1/3}$ in the examples, but similar results are obtained when ℓ is specified otherwise: Table 2 provides an illustration with $\ell \approx n^{1/4}$.

© 2018 Australian Statistical Publishing Association Inc.

4. Examples

For each example, we simulated the mean squared errors (MSEs) of the hybrid block bootstrap estimators of τ_n, $G_n(u)$, for two values of u, for the case $p=1/2$.

The true reference values of τ_n and $G_n(\cdot)$ in each example were approximated using a massive simulation, with 5×10^6 replications. Recall that we use the term 'hybrid' to indicate that we are considering values of b in the interval $[1, \lfloor n/\ell \rfloor]$, where $b=1$ is the subsampling bootstrap, and $b=\lfloor n/\ell \rfloor$ is the standard MBB.

All entries reported in the tables are based on 20,000 replications, with 20,000 bootstrap samples drawn for each replication.

Example 1. First-order autoregressive model. Suppose that Y_t is generated by an *AR(1)* process

$$Y_t = 0.4Y_{(t-1)} + \epsilon_t,$$

with the ϵ_t i.i.d. $\mathbf{N}(0,1)$. In the validation study, to simulate from this model, Y_0 was first randomly sampled from the marginal $\mathbf{N}(0, 1.1905)$ distribution.

With a normal distribution for the innovations, and since $|0.4|<1$, the AR(1) model is strongly mixing, and the mixing coefficients decay exponentially (Bradley 1986, example 6.1). Table 1 gives results for $n=50$, with block length $\ell=4$, the closest integer to $n^{1/3}$, and for $1\leqslant b\leqslant 12$. Since $p=1/2$, we have $\xi_p=0$. The true values being estimated were computed as $\tau_n=3.39932$, $G_n(-0.5)=0.43349$, and $G_n(1.5)=0.82056$. The MSE of the variance and distribution estimators is smallest for a value of b which is small, but strictly greater than 1. The last row, $b=12$, corresponds to the MBB, which is seen to perform considerably worse than hybrid block bootstrap estimators. The confidence intervals have poor coverage accuracy; both the 95% and 90% bootstrap confidence intervals produce substantial undercoverage, for all values of b.

Example 2. ARMA(1,1). We generated samples from an ARMA(1,1) model

$$Y_t - 0.4Y_{(t-1)} = \epsilon_t + 0.3\epsilon_{(t-1)},$$

with ϵ_t i.i.d. $\mathbf{N}(0,1)$. This ARMA model satisfies the strong mixing condition with an exponential rate of decay for the mixing coefficients; see, for example, Lahiri (2003, example 6.1). The value of Y_0 was sampled randomly from the marginal distribution $\mathbf{N}(0, 1.5833)$, and ϵ_0 is sampled from $\mathbf{N}(0,1)$.

Again, since $p=1/2$, we have $\xi_p=0$. The true values being estimated were computed as $\tau_n=5.68256$, $G_n(-1.5)=0.28007$, and $G_n(1)=0.67978$. Table 2 reports simulation results for $n=200$, $\ell=4$, the closest integer to $n^{1/4}$, and selected values of b in $[1,50]$, where $b=50$ corresponds to the MBB. Table 3 contains results for $n=200$, $\ell=6$, closest integer to $n^{1/3}$, and values of b in $[1,33]$, where $b=33$ indicates the MBB. In both tables, there is still substantial undercoverage by the bootstrap confidence intervals, even with this increased sample size. Notice that the results for distribution and variance estimation do depend strongly on (b,ℓ). For instance, the subsampling bootstrap performs much better for variance estimation when $\ell=n^{1/3}$ rather than $\ell=n^{1/4}$, and this pattern holds for all b considered, though the difference in MSE is small for values of b near the middle of the interval $[1, \lfloor n/\ell \rfloor]$. For distribution estimation, when comparing the performance for $\ell=n^{1/3}$ and $\ell=n^{1/4}$, the differences are slight. Once again, however, we note that there is a substantial performance improvement over the standard MBB or subsampling bootstrap by choosing $b>1$.

© 2018 Australian Statistical Publishing Association Inc.

Table 1 AR(1) model. Sample size $n=50$, block length $\ell=4$.

b	MSE estimation τ_n	MSE estimation $G_n(-0.5)$	MSE estimation $G_n(1.5)$	Coverages (95%, 90%) CIs
1	1.97071	0.00983	0.01220	(80.1, 75.9)
2	1.85262	0.00633	0.00955	(81.9, 75.9)
3	1.88072	0.00627	0.00877	(82.0, 75.6)
4	2.01892	0.00704	0.00874	(82.6, 76.0)
5	2.14152	0.00819	0.00886	(82.2, 75.7)
6	2.27492	0.00914	0.00909	(82.4, 76.2)
7	2.33889	0.01027	0.00928	(82.4, 76.1)
8	2.56672	0.01145	0.00970	(81.9, 75.2)
9	2.65207	0.01227	0.01003	(82.1, 75.7)
10	2.72855	0.01304	0.01026	(82.1, 75.3)
11	2.82212	0.01381	0.01052	(82.0, 75.4)
12	3.04957	0.01448	0.01097	(81.5, 75.0)

Table 2 ARMA(1,1) model. Sample size $n=200$, block length $\ell=4$.

b	MSE estimation τ_n	MSE estimation $G_n(-1.5)$	MSE estimation $G_n(1.0)$	Coverages (95%, 90%) CIs
1	3.67013	0.00391	0.01465	(85.7, 79.1)
2	3.12188	0.00335	0.01084	(86.6, 80.1)
3	2.95750	0.00276	0.00816	(87.4, 80.8)
4	2.87318	0.00280	0.00696	(87.6, 80.9)
5	2.89054	0.00301	0.00630	(88.2, 81.6)
6	2.91291	0.00328	0.00593	(87.8, 81.4)
7	2.88793	0.00346	0.00561	(87.6, 81.0)
8	2.99171	0.00369	0.00553	(88.0, 81.3)
9	2.98520	0.00394	0.00539	(87.5, 80.9)
10	2.99115	0.00414	0.00532	(87.6, 81.0)
15	3.22651	0.00502	0.00528	(87.5, 81.1)
20	3.40073	0.00583	0.00544	(87.2, 80.5)
30	3.74392	0.00693	0.00597	(87.0, 80.5)
40	3.99044	0.00775	0.00647	(86.7, 80.0)
50	4.26749	0.00865	0.00696	(86.4, 79.7)

Example 3. Nonlinear ARMA(2,3). Let $\{X_t\}_{t\in\mathbb{Z}}$ be a sequence from the following specific ARMA(2,3) process. This sequence is strongly mixing with an exponential rate of decay for the mixing coefficients; see again Lahiri (2003, example 6.1),

$$X_t - 0.1X_{(t-1)} + 0.3X_{(t-2)} = \epsilon_t + 0.1\epsilon_{(t-1)} + 0.2\epsilon_{(t-2)} - 0.1\epsilon_{(t-3)}.$$

To simulate from this model, we initiate by generating X_0, X_{-1} from the marginal $\mathbf{N}(0, v^2)$ distribution, which has $v^2 = 1.0776$, with $\epsilon_0, \epsilon_{-1}, \epsilon_{-2}$ independent $\mathbf{N}(0,1)$. The nonlinear model we consider is the square transformation of the above ARMA process,

$$Y_t = X_t^2.$$

Any instantaneous Borel transformation, such as the square transformation above, preserves the strong mixing property and the mixing rate. Hence, Y_t is also strongly mixing with the same exponential rate as X_t; see Fan & Yao (2003, p. 69) or Davis & Mikosch (2009, p. 258). As before, we set $p = 1/2$, and so ξ_p satisfies

© 2018 Australian Statistical Publishing Association Inc.

Table 3 ARMA(1,1) model. Sample size $n=200$, block length $\ell=6$.

b	MSE estimation τ_n	MSE estimation $G_n(-1.5)$	MSE estimation $G_n(1.0)$	Coverages (95%, 90%) CIs
1	2.70465	0.00412	0.01054	(87.2, 80.9)
2	2.43701	0.00308	0.00711	(88.2, 81.8)
3	2.48573	0.00305	0.00583	(88.5, 82.2)
4	2.50317	0.00322	0.00519	(88.6, 82.2)
5	2.60016	0.00353	0.00494	(89.1, 83.1)
6	2.71171	0.00383	0.00480	(88.7, 82.6)
7	2.72484	0.00400	0.00468	(88.7, 82.4)
8	2.87785	0.00432	0.00477	(89.0, 82.6)
9	2.91608	0.00457	0.00478	(88.4, 82.2)
10	2.96183	0.00478	0.00480	(88.5, 82.1)
15	3.38465	0.00571	0.00508	(88.2, 82.2)
20	3.67521	0.00660	0.00548	(88.2, 81.4)
33	4.30900	0.00806	0.00644	(88.1, 81.6)

Table 4 Squared ARMA(2,3) model. Sample size $n=500$, block length $\ell=8$.

b	MSE estimation τ_n	MSE estimation $G_n(-1.5)$	MSE estimation $G_n(1.5)$	Coverages (95%, 90%) CIs
1	0.13558	0.00900	0.00077	(81.4, 77.8)
2	0.10000	0.00471	0.00070	(86.0, 81.7)
3	0.10240	0.00322	0.00082	(87.8, 83.1)
4	0.10694	0.00265	0.00091	(88.5, 84.0)
5	0.11384	0.00245	0.00102	(88.8, 83.5)
6	0.11941	0.00232	0.00113	(89.5, 84.7)
7	0.12359	0.00230	0.00117	(89.5, 84.4)
8	0.12816	0.00228	0.00127	(89.5, 84.0)
9	0.13368	0.00232	0.00133	(89.7, 84.4)
10	0.13457	0.00232	0.00139	(90.3, 85.1)
15	0.15984	0.00247	0.00172	(90.2, 84.6)
20	0.17711	0.00262	0.00195	(90.2, 84.5)
62	0.27639	0.00379	0.00331	(90.2, 84.7)

$$\Pr(Y_t = X_t^2 \leqslant \xi_p) = 1/2,$$

which yields $\xi_p = (0.675v)^2$. The computed true values were, respectively, $\tau_n = 1.38199$, $G_n(-1.5) = 0.09792$, and $G_n(1.5) = 0.89751$.

The results in Table 4 are for $n = 500$, $\ell = 8$, the closest integer to $n^{1/3}$, and b in $[1, 62]$, where $b = 62$ agrees with the standard MBB. As expected, with $n = 500$, the reported MSE's for variance estimation are considerably better than in the previous two examples, though the bootstrap confidence intervals still suffer from undercoverage. The primary observation, as before, is that the results depend on (b, ℓ). Small values of b, which are strictly greater than 1, offer improved MSE compared to the subsampling bootstrap and the MBB.

5. Concluding comments

We have given a proof that, for varying number of blocks $1 \leqslant b = O(n/\ell)$, the block bootstrap distribution and variance estimators for sample quantiles are weakly consistent

© 2018 Australian Statistical Publishing Association Inc.

under a strong mixing assumption with exponentially decaying mixing coefficients. This assumption holds in many models of practical importance, and we have illustrated, through examples, the effects of different choices of b and ℓ on MSE for both variance and distribution estimation for sample quantiles. There is clear evidence across the examples considered that there is substantial benefit from choosing a value of b somewhere between 1 and $\lfloor n/\ell \rfloor$. Our results open interesting avenues for future research. For such hybrid block bootstrap procedures, which are a compromise between the subsampling bootstrap and the standard MBB, our theoretical and empirical findings suggest that the performance depends on the choices of b and ℓ. Procedures for adaptive choice of the pair (ℓ, b) will be developed, but are beyond the scope of the present paper. Kuffner *et al.* (2017) provide a more detailed theoretical investigation of optimality in the choice of the number of blocks. More broadly, the problem studied here acts as a template for other related problems, involving estimation of quantities depending on local properties of the underlying distribution, where a hybrid block bootstrap of this kind might provide similar performance improvements.

Peter Hall was a giant of a man, in many respects. He combined the highest levels of scholarship with the deepest human warmth. Peter was an enormous influence in the scientific development of all three of us, and we dedicate this modest piece to his memory.

Appendix

Proof of Theorem 1. Denote by Φ and ϕ the standard normal distribution and density functions, respectively. Standard asymptotic properties (e.g. Lahiri & Sun 2009) of F_n for dependent data can be invoked to show that, for any $x \in \mathbb{R}$,

$$\Pr\Big(\sqrt{n}\big(\hat{\xi}_n - \xi_p\big) \leqslant x\Big) = \Phi\big(xf(\xi_p)\sigma(\xi_p)^{-1}\big) + o(1), \tag{1}$$

where $\sigma(x)^2 = \lim_{n\to\infty} n\mathrm{Var}\big(F_n(x)\big) = \sum_{t=-\infty}^{\infty} \mathrm{Cov}\ \big(\mathbf{1}\{Y_0 \leqslant x\}, \mathbf{1}\{Y_t \leqslant x\}\big)$.

Denote by $\hat{\kappa}_j(z)$ the jth conditional cumulant of $\sqrt{b\ell}\big(F_n^*(z) - \tilde{F}_n(z)\big)$ given $Y_1, \ldots, Y_n$, for any $z \in \mathbb{R}$. The fact that $\alpha(t) = O(e^{-Ct})$ enables us to establish that, for any arbitrarily small $\delta > 0$,

$$\hat{\kappa}_2(z) = \sigma(z)^2 + O_p\big(\ell^{-1} + \ell^{1/2} n^{-1/2}\big) = \sigma(z)^2 + o_p(1),$$
$$\hat{\kappa}_j(z) = O_p\Big((b\ell)^{-(j-2)/2}\ell^{\delta} + b^{-(j-2)/2}\ell^{1/2}n^{-1/2}\Big) = o_p(1),\ j \geqslant 3.$$

These results hold uniformly over $z \in \mathscr{N}_p$. It follows by standard Edgeworth expansion that, for any compact $\mathscr{K} \subset \mathbb{R}$,

$$\Pr\Big(\sqrt{b\ell}\big(F_n^*(z) - \tilde{F}_n(z)\big) \leqslant y \,\Big|\, Y_1, \ldots, Y_n\Big) = \Phi\big(y/\sigma(z)\big) + o_p(1), \tag{2}$$

uniformly over $(y, z) \in \mathscr{K} \times \mathscr{N}_p$.

Note that, by lemmas 5.4 and 5.5 of Sun & Lahiri (2006), and Taylor expansion of F about ξ_p,

© 2018 Australian Statistical Publishing Association Inc.

$$
\begin{aligned}
&\{p-\tilde{F}_n(\tilde{\xi}_n+(b\ell)^{-1/2}x)\}\sigma(\tilde{\xi}_n+(b\ell)^{-1/2}x)^{-1}\\
&=\Big\{F_n(\tilde{\xi}_n)-F_n(\tilde{\xi}_n+(b\ell)^{-1/2}x)\Big\}\sigma(\xi_p)^{-1}\Big\{1+O_p\big((b\ell)^{-1/2}+n^{-1/2+\delta}\big)\Big\}\\
&\quad+O_p\big(n^{-1}\ell+n^{-1/2}(\log n)^{-2}\big)\\
&=-(b\ell)^{-1/2}xf(\xi_p)\sigma(\xi_p)^{-1}\\
&\quad+O_p\big(n^{-1}\ell+(b\ell)^{-1}+n^{-1/2+\delta}(b\ell)^{-1/2}+n^{-1/2}(\log n)^{-2}\big)\\
&=-(b\ell)^{-1/2}\{xf(\xi_p)\sigma(\xi_p)^{-1}+o_p(1)\}.
\end{aligned}
\tag{3}
$$

It follows from (2) and (3) that

$$
\begin{aligned}
&\Pr\Big(F_n^*(\tilde{\xi}_n+(b\ell)^{-1/2}x)\leqslant p\Big|Y_1,\ldots,Y_n\Big)\\
&=\Pr\Big(\sqrt{b\ell}\{F_n^*(\tilde{\xi}_n+(b\ell)^{-1/2}x)-\tilde{F}_n(\tilde{\xi}_n+(b\ell)^{-1/2}x)\}\\
&\qquad\leqslant\sqrt{b\ell}\{p-\tilde{F}_n(\tilde{\xi}_n+(b\ell)^{-1/2}x)\}\Big|Y_1,\ldots,Y_n\Big)\\
&=\Phi\Big(\sqrt{b\ell}\{p-\tilde{F}_n(\tilde{\xi}_n+(b\ell)^{-1/2}x)\}\sigma(\tilde{\xi}_n+(b\ell)^{-1/2}x)^{-1}\Big)+o_p(1)\\
&=\Phi\big(-xf(\xi_p)\sigma(\xi_p)^{-1}\big)+o_p(1).
\end{aligned}
\tag{4}
$$

Theorem 1(i) then follows by (1) and (4) and noting that

$$
\begin{aligned}
\Pr\Big(F_n^*(\tilde{\xi}_n+(b\ell)^{-1/2}x)>p\Big|Y_1,\ldots,Y_n\Big)&\leqslant\Pr\Big(\sqrt{b\ell}(\xi_n^*-\tilde{\xi}_n)\leqslant x\Big|Y_1,\ldots,Y_n\Big)\\
&\leqslant\Pr\Big(F_n^*(\tilde{\xi}_n+(b\ell)^{-1/2}x)\geqslant p\Big|Y_1,\ldots,Y_n\Big).
\end{aligned}
$$

For part (ii), it suffices, following the proof of theorem 3.2 in Sun & Lahiri (2006), to show that for some $\epsilon>0$,

$$
\begin{aligned}
&\int_1^{L_n}x^{1+\epsilon}\Pr\Big(\sqrt{b\ell}\{F_n^*(\tilde{\xi}_n+(b\ell)^{-1/2}x)-\tilde{F}_n(\tilde{\xi}_n+(b\ell)^{-1/2}x)\}\leqslant -Ax\Big|Y_1,\ldots,Y_n\Big)dx\\
&=O_p(1)
\end{aligned}
\tag{5}
$$

and

$$
\{(b\ell)^{1/2}n^{1/\alpha}\}^{1+\epsilon}\Pr\Big(F_n^*(\tilde{\xi}_n+(b\ell)^{-1/2}L_n)\leqslant p\Big|Y_1,\ldots,Y_n\Big)=O_p(1),
\tag{6}
$$

where $A=2^{-1}\inf_{x\in\mathcal{N}_p}f(x)$ and $L_n=A^{-1}(\alpha^{-1}+2^{-1})(1+\epsilon)\log n$. Using (2) and Markov's inequality, we have, for any $y>0$, that

$$
\begin{aligned}
&\Pr\Big(\sqrt{b\ell}\{F_n^*(\tilde{\xi}_n+(b\ell)^{-1/2}x)-\tilde{F}_n(\tilde{\xi}_n+(b\ell)^{-1/2}x)\}\leqslant -y\Big|Y_1,\ldots,Y_n\Big)\\
&\leqslant e^{-y}\mathbf{E}\Big(\exp\{-\sqrt{b\ell}(F_n^*-\tilde{F}_n)(\tilde{\xi}_n+(b\ell)^{-1/2}x)\}\Big|Y_1,\ldots,Y_n\Big)\\
&=O_p(e^{-y}),
\end{aligned}
\tag{7}
$$

with this result holding uniformly in $x\in[1,L_n]$. Thus, (5) follows immediately by setting $y=Ax$ in (7). On the other hand, setting $x=L_n$ and $y=\sqrt{b\ell}\{\tilde{F}_n(\tilde{\xi}_n+(b\ell)^{-1/2}L_n)-p\}$ in

© 2018 Australian Statistical Publishing Association Inc.

(3) and (7), we have

$$\begin{aligned}
&\left\{(b\ell)^{1/2}n^{1/\alpha}\right\}^{1+\epsilon}\Pr\left(F_n^*\left(\tilde{\xi}_n+(b\ell)^{-1/2}L_n\right)\leqslant p\,\middle|\,Y_1,\ldots,Y_n\right)\\
&=O_p\left\{n^{(1/2+1/\alpha)(1+\epsilon)}\exp\left(\sqrt{b\ell}\left\{p-\tilde{F}_n\left(\tilde{\xi}_n+(b\ell)^{-1/2}L_n\right)\right\}\right)\right\}\\
&=O_p\left\{n^{(1/2+1/\alpha)(1+\epsilon)}e^{-AL_n}\right\}=O_p(1),
\end{aligned}$$

which proves (6).

References

BRADLEY, R. (1986). Basic properties of strong mixing conditions. In *Dependence in Probability and Statistics*, eds. E. Eberlein and M. Taqqu, pp. 165–192. Boston: Birkhäuser.

BÜHLMANN, P. (1994). Blockwise bootstrapped empirical process for stationary sequences. *Annals of Statistics* **22**, 995–1012.

BÜHLMANN, P. & KÜNSCH, H.R. (1999). Block length selection in the bootstrap for time series. *Computational Statistics and Data Analysis* **31**, 295–310.

CARLSTEIN, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Annals of Statistics* **14**, 1171–1179.

CARLSTEIN, E., DO, K.A., HALL, P., HESTERBERG, T. & KÜNSCH, H.R. (1998). Matched-block bootstrap for dependent data. *Bernoulli* **4**, 305–328.

CHOI, E. & HALL, P. (2000). Bootstrap confidence regions computed from autoregressions of arbitrary order. *Journal of the Royal Statistical Society Series B* **62**, 461–477.

DAVIS, R.A. & MIKOSCH, T. (2009). Probabilistic properties of stochastic volatility models. In *Handbook of Financial Time Series*, eds. T.G. Andersen, R.A. Davis, J.P. Kreiß and T. Mikosch, pp. 255–267. Berlin: Springer.

DAVISON, A. & HALL, P. (1993). On studentizing and blocking methods for implementing the bootstrap with dependent data. *Australian & New Zealand Journal of Statistics* **35**, 215–224.

FAN, J. & YAO, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. New York: Springer.

GÖTZE, F. & KÜNSCH, H.R. (1996). Second-order correctness of the blockwise bootstrap for stationary observations. *Annals of Statistics* **24**, 1914–1933.

GREGORY, K.B., LAHIRI, S.N. & NORDMAN, D.J. (2015). A smooth block bootstrap for statistical functionals and time series. *Journal of Time Series Analysis* **36**, 442–461.

HALL, P. (1985). Resampling a coverage pattern. *Stochastic Processes and Their Applications* **20**, 231–246.

HALL, P. (2003). A short prehistory of the bootstrap. *Statistical Science* **18**, 158–167.

HALL, P. & HOROWITZ, J.L. (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. *Econometrica* **64**, 891–916.

HALL, P. & JING, B. (1996). On sample reuse methods for dependent data. *Journal of the Royal Statistical Society Series B* **58**, 727–737.

HALL, P., HOROWITZ, J.L. & JING, B.Y. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* **82**, 561–574.

HALL, P., JING, B.Y. & LAHIRI, S.N. (1998). On the sampling window method for long-range dependent data. *Statistica Sinica* **8**, 1189–1204.

KUFFNER, T.A., LEE, S.M. & YOUNG, G.A. (2017). Optimal block bootstrap distribution estimation for sample quantiles under weak dependence. Technical report.

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

LAHIRI, S.N. (1992). Edgeworth correction by moving block bootstrap for stationary and nonstationary data. In *Exploring the Limits of Bootstrap*, eds. R. LePage and L. Billard, pp. 263–270. New York: Wiley.

LAHIRI, S.N. (1996). On Edgeworth expansion and moving block bootstrap for studentized m-estimators in multiple linear regression models. *Journal of Multivariate Analysis* **56**, 42–59.

LAHIRI, S.N. (1999). Theoretical comparisons of block bootstrap methods. *Annals of Statistics* **27**, 386–404.

LAHIRI, S.N. (2003). *Resampling Methods for Dependent Data*. New York: Springer-Verlag.

LAHIRI, S.N. & SUN, S. (2009). A Berry-Esseen theorem for sample quantiles under weak dependence. *Annals of Applied Probability* **19**, 108–126.

© 2018 Australian Statistical Publishing Association Inc.

LIU, R. & SINGH, K. (1992). Moving blocks jackknife and bootstrap capture weak convergence. In *Exploring the Limits of Bootstrap*, eds. R. LePage and L. Billard, pp. 225–248. New York: Wiley.

NAIK-NIMBALKAR, U. & RAJARSHI, M. (1994). Validity of blockwise bootstrap for empirical processes with stationary observations. *Annals of Statistics* **22**, 980–994.

POLITIS, D. & ROMANO, J. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics* **22**, 2031–2050.

SHARIPOV, O.S. & WENDLER, M. (2013). Normal limits, nonnormal limits, and the bootstrap for quantiles of dependent data. *Statistics and Probability Letters* **83**, 1028–1035.

SUN, S. (2004). Bootstrapping the sample quantile based on weakly dependent observations (PhD Thesis). Iowa State University.

SUN, S. (2007). On the accuracy of bootstrapping sample quantiles of strongly mixing sequences. *Journal of the Australian Mathematical Society* **82**, 263–281.

SUN, S. & LAHIRI, S.N. (2006). Bootstrapping the sample quantile of a weakly dependent sequence. *Sankhyā* **68**, 130–166.

© 2018 Australian Statistical Publishing Association Inc.