

ON THE SAMPLING WINDOW METHOD FOR LONG-RANGE DEPENDENT DATA

Peter Hall, Bing-Yi Jing and Soumendra Nath Lahiri

*Australian National University, Hong Kong University of Science and Technology
and Iowa State University*

Abstract: It is known that under conditions of long-range dependence, and for time series subordinated to Gaussian processes, the block bootstrap method produces invalid estimators of the distribution of the sample mean unless the limiting distribution is normal. In this paper we show that the sampling window method produces valid, consistent estimators for non-normal as well as normal limits. Additionally, we introduce a method for “studentizing” the sample mean of long-range dependent data, and show that sampling window approximations of its distribution are also valid. That result suggests that the sampling window method is useful for setting confidence intervals for a population mean in a particularly wide range of circumstances. This conclusion is supported by a small simulation study.

Key words and phrases: Block bootstrap, consistency, Gaussian processes, long-range dependence, sampling window method.

1. Introduction

Approximating sampling distributions of estimators based on dependent data remains a challenging problem in Statistics. When the dependence structure in the data can be explained by simple models driven by independent and identically distributed disturbances (e.g. autoregressive processes with independent innovations, or countable Markov chains satisfying certain geometric ergodicity conditions), the bootstrap method of Efron (1979) may be employed with only minor modifications. See for example Freedman (1984), Bose (1988), Athreya and Fuh (1992), Datta and McCormick (1992), and the references therein. For dependent data that do not admit such models, however, a more suitable resampling scheme is the “block bootstrap” method developed by Hall (1985), Carlstein (1986), Künsch (1989), and Liu and Singh (1992) in different contexts. When the data are weakly dependent, the block bootstrap provides accurate approximations for the unknown sampling distributions of many commonly used estimators. See, for example, Lahiri (1991, 1996), Götze and Künsch (1996), Hall, Horowitz and Jing (1996), and the references therein.

The situation is very different when the data are strongly dependent, however. For example, it has been shown by Lahiri (1993) that the block bootstrap

is not consistent when the observations come from certain long-range dependent processes. The main conclusion of that paper can be roughly summarized as “the block bootstrap fails to capture the limit law of the normalized sample mean \bar{X}_n of long-range dependent data, whenever \bar{X}_n has a non-normal limit law”. The main reason for this behaviour is that joining *independent* bootstrap blocks to define the bootstrap sample fails to reproduce the long-range dependence features of the time series, and so makes the block bootstrap ineffective.

In this note we discuss the sampling window method of Hall and Jing (1996) (see also Politis and Romano (1994) and Bickel, Götze and van Zwet (1997)) in the context of long-range dependent data. It is shown that under appropriate regularity conditions, the sampling window method produces consistent estimators of the distribution of the normalized sample mean in the cases of both normal and non-normal limit laws. Furthermore, we introduce a method for studentizing \bar{X}_n , and show that the sampling window method is also successful in approximating the distribution of the studentized statistic, again for both normal and non-normal limit laws. The ability of the sampling window approach to capture non-normal limit distributions under conditions of very-long-range dependence is particularly important, and it appears that this technique may be employed to construct valid confidence intervals in a particularly wide range of circumstances.

Section 2 first introduces the sampling window method and briefly reviews relevant asymptotic properties of long-range dependent processes. Then we state our main theoretical results. The problem of choosing window length is addressed through a numerical study in Section 3. Technical arguments are outlined in Section 4.

2. Methodology and Theoretical Properties

2.1. Asymptotic distribution of the mean

We assume that the observed data represent a realization of a stationary sequence $\mathbf{X}_n = \{X_1, \dots, X_n\}$, which may be represented as a function of a stationary Gaussian process $\mathbf{Z} = \{Z_n, -\infty < n < \infty\}$ in the following way: $X_i = G_1(Z_i)$, $i \geq 1$, where $G_1 : \mathcal{R} \rightarrow \mathcal{R}$ is a Borel-measurable function satisfying $EG_1(Z_1)^2 < \infty$. The sample mean $\bar{X}_n = n^{-1} \sum_{1 \leq i \leq n} X_i$ is a consistent estimator of the population mean, $\mu = E(X_1)$, although its rate of convergence may be slower than $n^{-1/2}$, and its asymptotic distribution may be non-normal. Following Taqqu (1975) we describe the limiting properties of \bar{X}_n in terms of the *Hermite rank* of $G = G_1 - \mu$. To this end, let $H_k(\cdot)$, for $k \geq 1$, denote the k th Hermite polynomial, $H_k(x) = (-1)^k \exp(x^2/2)(d^k/dx^k)(\exp(-x^2/2))$, $x \in \mathcal{R}$. Then, the Hermite rank q of $G(\cdot)$ is defined as

$$q = \inf\{k \geq 1 : E[H_k(Z_1)G(Z_1)] \neq 0\}.$$

The limit distribution of \bar{X}_n depends on q and on the asymptotic behaviour of the autocovariance function, $r(k) = \text{Cov}(Z_j, Z_{j+k})$, as follows. Suppose $r(\cdot)$ is regularly varying at infinity — that is,

$$r(k) = k^{-\alpha}L(k) \tag{2.1}$$

as $k \rightarrow \infty$, where $\alpha > 0$ and the function L is slowly varying (see Bingham, Goldie and Teugels (1987)). Let $d_n = \{n^{2-q\alpha}L^q(n)\}^{1/2}$, $A = 2\Gamma(\alpha) \cos(\alpha\pi/2)$ and $C_q = E\{H_q(Z_1)G(Z_1)\}/q!$.

Theorem 2.1. (Taqqu (1975, 1979), Dobrushin and Major (1979)) *Assume that r admits the representation at (2.1), and that G has Hermite rank q , where $0 < \alpha < q^{-1}$. Then, $n(\bar{X}_n - \mu)/d_n \rightarrow W_q$ in distribution, where W_q is defined in terms of a multiple Wiener-Itô integral with respect to the random spectral measure W of the Gaussian white-noise process as*

$$W_q = \frac{C_q}{A^{q/2}} \int \frac{\exp\{i(x_1 + \dots + x_q)\} - 1}{i(x_1 + \dots + x_q)} \prod_{k=1}^q |x_k|^{(\alpha-1)/2} dW(x_1) \dots dW(x_q). \tag{2.2}$$

When $q = 1$, W_q has a normal distribution with mean zero and variance $2C_q^2/\{(1-\alpha)(2-\alpha)\}$, but for $q \geq 2$ the distribution of W_q is non-normal (Taqqu (1975)). For details of the representation of W_q in (2.2), and the concept of a multiple Wiener-Itô integral with respect to the random spectral measure of a stationary process, see Dobrushin and Major (1979) and Dobrushin (1979) respectively.

2.2. Sampling window estimator of the distribution of the mean

Let $\mathbf{X}_n = \{X_1, \dots, X_n\}$ denote a stationary stochastic process, of which the observed data represent a realization; let $1 \leq \ell \leq n$ be the length of sampling window; and let $\mathcal{B}_i = (X_i, \dots, X_{i+\ell-1})$, for $1 \leq i \leq N = n - \ell + 1$, denote the i th among the blocks of length ℓ into which \mathbf{X}_n may be divided. We may regard the blocks as scaled-down replicates of the original data sequence \mathbf{X}_n . Let $S_{\ell i} = \sum_{i \leq j \leq i+\ell-1} X_j$ be the sum of the elements of \mathcal{B}_i , and let $T_{\ell i} = (S_{\ell i} - \ell \bar{X}_n)/d_\ell$ denote the analogue of the normalized sample mean $T_n = n(\bar{X}_n - \mu)/d_n$ for \mathcal{B}_i . The sampling window (SW) estimator $\hat{F}_n(x)$ of the distribution function $F_n(x)$ of T_n is defined as the proportion of the $T_{\ell i}$'s that do not exceed x :

$$\hat{F}_n(x) = N^{-1} \sum_{i=1}^N I\{(S_{\ell i} - \ell \bar{X}_n)/d_\ell \leq x\},$$

where $I(\cdot)$ denotes the indicator function.

2.3. Consistency of the sampling window estimator

Here it is convenient to describe the strength of dependence in terms of the spectral density, f , of $\{Z_n\}$, rather than its autocovariance. The two approaches are equivalent. Indeed, Abelian–Tauberian theorems (e.g. Zygmund (1968), Ch. 5, and Bingham, Goldie and Teugels (1987)) may be used to prove that if (2.1) holds then

$$f(x)/\{x^{\alpha-1}L(1/x)\} \rightarrow C(\alpha) \quad \text{as } x \rightarrow 0, \quad (2.3)$$

where $C(\alpha) > 0$ depends only on α ; and conversely, if (2.3) holds for a slowly varying function L then r admits the representation (2.1) with (without loss of generality) the same L .

Since f is symmetric the Fourier series of $\log f$ is a pure cosine series. Replacing each cosine function by the corresponding sine we obtain the *Harmonic conjugate* of $\log f$, which we denote by $\widetilde{\log f}$. We shall ask that $\widetilde{\log f}$ be continuous, which implies that the process $\{Z_n\}$ is completely regular (see Ibragimov and Rozanov (1978) for a definition). Note that while $\log f$, being unbounded in every neighbourhood of the origin, is not continuous on the circle $|x| \leq \pi$, an appropriately chosen branch of $\widetilde{\log f}$ can be continuous there.

Our first theorem describes consistency of the SW method under a condition similar to (2.3). (By way of comparison, the block bootstrap provides consistency only when $q = 1$.) We treat only the case $0 < \alpha < q^{-1}$, noting that the central limit holds (with a Normal limit) when $\alpha > q^{-1}$, and so result (2.4) is relatively trivial there.

Theorem 2.2. *Let $q \geq 1$ denote the Hermite rank of G . Suppose that $f(x) = |x|^{\alpha-1}L_1(|x|)$ for $0 < |x| \leq \pi$, where $0 < \alpha < q^{-1}$ and L_1 is slowly varying at 0 and of bounded variation on every closed subinterval of $(0, \pi]$; that a branch of $\widetilde{\log f}$ is continuous on $|x| \leq \pi$; and that $\ell^{-1} + n^{-(1-\epsilon)}\ell = o(1)$ for some $\epsilon > 0$. Then,*

$$\sup_{x \in \mathbb{R}} |\widehat{F}_n(x) - F_n(x)| \rightarrow 0 \quad (2.4)$$

in probability as $n \rightarrow \infty$.

When $q = 1$ or 2 the distribution of W_q is completely determined by its moments, which property we use to establish a version of Theorem 2.2 under weaker assumptions.

Theorem 2.3. *If $q = 1$ or 2, and $\ell^{-1} + n^{-(1-\epsilon)}\ell = o(1)$ for some $\epsilon > 0$, then (2.4) holds under the assumptions of Theorem 2.1.*

Arguments similar to those employed to prove Theorem 2.3 may be used to derive L^p consistency of the SW method *without* assuming that $\{X_i\}$ is a function of a Gaussian process. We need asymptotic independence of moments of shifted

partial sums (see (2.6) below), and that the limit distribution is completely determined by its moments. Let $\{X_j, -\infty < j < \infty\}$ denote a stationary stochastic process with all moments finite, mean μ and the property that $n(\bar{X}_n - \mu)/s_n$ converges in distribution to a proper random variable ζ , where $s_n^2 = n^2 \text{Var}(\bar{X}_n)$. Let $\ell = \ell(n) \leq n$ be a sequence of positive integers, and define $S_{\ell i}, T_{\ell i}$ and \hat{F}_n as in Section 2.2 but with d_ℓ replaced by s_ℓ . Write F for the distribution function of ζ .

Theorem 2.4. *Assume that the distribution of ζ is continuous and uniquely determined by its moments, all of which are finite. Suppose $\ell(n) \rightarrow \infty$,*

$$(\ell s_n)/(n s_\ell) \rightarrow 0, \tag{2.5}$$

and for each $\epsilon > 0$ and each pair of positive integers r_1, r_2 ,

$$\sup_{\epsilon n \leq i \leq n} \left| E \left[\{s_\ell^{-1}(S_{\ell i} - \ell\mu)\}^{r_1} \{s_\ell^{-1}(S_{\ell i} - \ell\mu)\}^{r_2} \right] - E(\zeta^{r_1})E(\zeta^{r_2}) \right| \rightarrow 0. \tag{2.6}$$

Then,

$$E \left\{ \sup_{-\infty < x < \infty} |\hat{F}_\ell(x) - F(x)| \right\} \rightarrow 0$$

as $n \rightarrow \infty$.

Condition (2.5) is of course equivalent to $\text{Var}(\bar{X}_n)/\text{Var}(\bar{X}_\ell) \rightarrow 0$ as $n \rightarrow \infty$, which is the standard assumption for the “ m -out-of- n bootstrap” (cf. Bickel, et al. (1997)). Since all moments of X_1 are finite, condition (2.6) is a relatively weak dependence assumption, and has the advantage over potential alternative conditions that it has a simple statistical interpretation — it asks that polynomials functions of normalised means, computed from well-separated parts of the stochastic process, be asymptotically uncorrelated.

2.4. The case of the studentized mean

Here we describe an empirical device for standardizing scale, replacing d_n by a function of the data but preserving consistency of the SW method for long-range dependent data. Let $m_1 = m_1(n), m_2 = m_2(n) \in [1, n]$ denote integers such that for some $\epsilon \in (0, 1)$,

$$m_1^2/m_2 \sim n \quad \text{and} \quad m_i = O(n^{1-\epsilon}) \quad \text{as} \quad n \rightarrow \infty, \tag{2.7}$$

where $a_n \sim b_n$ means that $a_n/b_n \rightarrow 1$ as $n \rightarrow \infty$. (For example, we could have $m_i \sim n^{\alpha_i}$ where $0 < \alpha_2 < 1$ and $\alpha_1 = \frac{1}{2}(1 + \alpha_2)$.) Define $d_m^{*2} = (n - m + 1)^{-1} \sum_{i=1}^{n-m+1} (S_{mi} - m\bar{X}_n)^2$ and $\hat{d}_n^2 = d_{m_1}^{*4}/d_{m_2}^{*2}$. Under the conditions of Theorem 2.5 below, \hat{d}_n is consistent for d_n in the sense that $\hat{d}_n/d_n \rightarrow 1$ in probability as $n \rightarrow \infty$. We use \hat{d}_n for studentizing the sample mean \bar{X}_n . Note that in

contrast to the case of short-range dependence, where the right scaling constant (viz. $n^{1/2}$) for asymptotic normality of the centered sample mean is known, the scaling constant $d_n = \{n^{2-q\alpha}L^q(n)\}^{1/2}$ in the long-range dependent case involves the unknown quantities α and $L(\cdot)$. Construction of the estimator \hat{d}_n here, with the two smoothing parameters m_1 and m_2 , yields a consistent estimator of d_n without requiring knowledge of α and $L(\cdot)$.

Next we define the SW estimator of the sampling distribution of the studentized sample mean $T_{1n} = n(\bar{X}_n - \mu)/\hat{d}_n$. Let $\hat{d}_\ell^{(i)}$ denote the version of \hat{d}_n for the i th block \mathcal{B}_i , defined using the smoothing parameters $m_1(\ell)$ and $m_2(\ell)$. (Note that the values $m_1(\ell)$ and $m_2(\ell)$ satisfy the relations ‘ $m_1(\ell)^2/m_2(\ell) \sim \ell$ ’ and ‘ $m_i = O(\ell^{1-\epsilon})$ for some $\epsilon > 0$ ’, which correspond to the length ℓ of the block \mathcal{B}_i as compared to (2.7). This in turn relates to the sample size n .) Let F_{1n} denote the distribution function of the studentized sample mean T_{1n} , and write

$$\hat{F}_{1n}(x) = (n - \ell + 1)^{-1} \sum_{i=1}^{n-\ell+1} I\{(S_{\ell i} - \ell\bar{X}_n)/\hat{d}_\ell^{(i)} \leq x\}$$

for the SW estimator of F_{1n} .

Theorem 2.5. *Assume the conditions of Theorem 2.1 for some $q \geq 1$, that m_1, m_2 satisfy (2.7), that $\ell^{-1} + n^{-(1-\epsilon)}\ell = o(1)$ for some $\epsilon > 0$, and that*

$$L^2(xy)/\{L(x^2)L(y^2)\} \rightarrow 1 \quad \text{as } x, y \rightarrow \infty. \tag{2.8}$$

Then, (a) $\hat{d}_n/d_n \rightarrow 1$ in probability as $n \rightarrow \infty$, (b) $T_{1n} \rightarrow W_q$ in distribution as $n \rightarrow \infty$, (c) for $q = 1, 2$,

$$\sup_{x \in \mathbb{R}} |\hat{F}_{1n}(x) - F_{1n}(x)| \rightarrow 0 \tag{2.9}$$

in probability as $n \rightarrow \infty$, and (d) if in addition the conditions of Theorem 2.2 hold, then (2.9) holds for all $q \geq 1$.

The accuracy of the SW approximation depends not only on ℓ but also on choice of the integers m_1, m_2 . Admissible values are $m_1 = n^{(1+\theta)/2}$ and $m_2 = n^\theta$, for $0 < \theta < 1$. It follows from the proof of Theorem 2.5 that the intermediate estimator d_m^* has smaller bias for large values of m . Hence, one may choose θ close to 1 in order to keep m_1 and m_2 large and to ensure that the effects described by Theorem 2.5 are more apparent.

3. Numerical Study

3.1. Algorithm for generating data

We generated stationary increments of a self-similar process with self-similarity parameter (or Hurst constant) $H = \frac{1}{2}(2 - \alpha)$, and took an appropriate transformation of these data to produce a realization of a long-range dependent process with Hermite rank q . The algorithm was as follows.

1. Generate a random sample $\mathbf{Z}_{n0} = \{Z_{10}, \dots, Z_{n0}\}$ of size n from the standard normal distribution.
2. Let $R = (r_{ij})$ denote the correlation matrix defined by

$$r_{ij} = \frac{1}{2}\{(k+1)^{2H} + (k-1)^{2H} - 2k^{2H}\}, \tag{3.1}$$

for $k = |j - i|$ and $\frac{1}{2} < H < 1$. Decompose R into $R = U^T U$ by Cholesky factorization.

3. Define $\mathbf{Z}_n \equiv \{Z_1, \dots, Z_n\} = U^T \mathbf{Z}_{n0}$. Then the process \mathbf{Z}_n is stationary and Gaussian with zero mean, unit variance and autocovariance at lag k given by

$$r(k) = \frac{1}{2}\{(k+1)^{2H} + (k-1)^{2H} - 2k^{2H}\} \sim Ck^{-\alpha} \text{ as } k \rightarrow \infty, \tag{3.2}$$

where $\alpha = 2 - 2H \in (0, 1)$. Therefore, $\mathbf{Z}_n \equiv \{Z_1, \dots, Z_n\}$ is a long-range dependent process. The $r(k)$'s are the autocovariances of the stationary increments of a self-similar process with self-similarity parameter H ; see e.g. Beran (1994), p. 50.

4. Define $X_i = H_q(Z_i)$, for $i = 1, \dots, n$, where H_q is the q th Hermite polynomial. Then $\mathbf{X}_n = \{X_1, \dots, X_n\}$ is a long-range dependent series with Hermite rank q . We report results derived by simulating from this process in the cases $q = 1, 2, 3$.

3.2. Coverage accuracy of the SW method

Let $\mathbf{X}_n = \{X_1, \dots, X_n\}$ be the long-range dependent time series with Hermite rank q , generated in Section 3.1. In the present section we analyse the finite-sample coverage probabilities of pointwise confidence intervals at a fixed point produced using the studentized sampling window approach from Section 2.4. Throughout, we take the nominal level to be 0.95. The empirical approximations to coverage probabilities reported here were derived by averaging over $B = 1000$ independent simulations.

Experience with the case of short-range dependent data, reported by Hall and Jing (1996), suggests that we should take l considerably smaller than n . We employed $l = cn^{1/2}$, for $c = 1, 3, 6, 9$. This choice is based not on an assertion that the optimal value of l is of size $n^{1/2}$ (indeed, the best order of l is not known to us), but on intuition that the optimal size should be greater than that for the weakly dependent case, where $l \sim cn^d$ for $d \leq \frac{1}{3}$ is generally appropriate (Hall and Jing (1996)). In choosing m_1 and m_2 for the procedure mentioned in the last paragraph of Section 2, we took $\theta = 0.9$.

Results are summarized in Tables 1 to 4. In all tables the bracketed pairs (\cdot, \cdot) represent coverage probabilities of lower and upper 95% one-sided confidence intervals, respectively, and q denotes Hermite rank.

Table 1. $\ell = n^{1/2}$

	q	$\alpha = 0.1$	$\alpha = 0.5$	$\alpha = 0.9$
$n = 100$	1	(85.0, 85.6)	(89.4, 91.1)	(91.5, 92.9)
	2	(84.4, 85.0)	(89.6, 90.5)	(92.1, 93.1)
	3	(84.4, 84.1)	(90.6, 91.6)	(91.3, 93.9)
$n = 200$	1	(92.0, 94.0)	(94.5, 95.5)	(96.4, 97.0)
	2	(92.7, 93.7)	(95.4, 95.7)	(96.3, 97.6)
	3	(92.2, 91.7)	(95.8, 96.1)	(96.7, 97.4)
$n = 500$	1	(90.1, 91.5)	(93.6, 93.9)	(94.7, 95.0)
	2	(90.6, 90.4)	(94.6, 95.0)	(93.4, 96.0)
	3	(90.8, 89.4)	(92.4, 96.0)	(95.4, 95.8)
$n = 1000$	1	(96.0, 95.1)	(96.9, 96.0)	(97.1, 97.3)
	2	(96.2, 94.0)	(97.0, 97.0)	(96.8, 98.0)
	3	(95.4, 94.4)	(96.6, 96.6)	(97.6, 97.6)

Table 2. $\ell = 3n^{1/2}$

	q	$\alpha = 0.1$	$\alpha = 0.5$	$\alpha = 0.9$
$n = 100$	1	(82.9, 84.1)	(87.1, 89.5)	(90.1, 91.7)
	2	(83.0, 84.9)	(86.3, 88.8)	(89.0, 91.6)
	3	(82.5, 84.9)	(87.7, 89.2)	(90.2, 89.7)
$n = 200$	1	(85.2, 88.0)	(92.2, 91.6)	(93.4, 94.0)
	2	(86.9, 86.9)	(90.5, 91.6)	(92.6, 93.6)
	3	(87.6, 87.3)	(90.1, 92.0)	(92.9, 93.2)
$n = 500$	1	(90.8, 90.8)	(93.6, 94.2)	(94.2, 95.8)
	2	(85.6, 91.4)	(94.0, 93.4)	(94.6, 95.2)
	3	(88.5, 88.7)	(93.0, 93.6)	(95.1, 94.0)
$n = 1000$	1	(92.6, 93.8)	(94.2, 95.8)	(95.6, 95.8)
	2	(90.2, 89.2)	(95.2, 95.0)	(96.2, 96.2)
	3	(90.6, 91.6)	(94.4, 94.4)	(96.4, 96.4)

Table 3. $\ell = 6n^{1/2}$

	q	$\alpha = 0.1$	$\alpha = 0.5$	$\alpha = 0.9$
$n = 100$	1	(66.5, 70.5)	(70.8, 74.7)	(73.7, 77.7)
	2	(69.1, 68.6)	(74.1, 72.1)	(73.1, 78.0)
	3	(66.6, 69.9)	(70.4, 73.9)	(73.2, 76.0)
$n = 200$	1	(77.4, 80.5)	(83.8, 85.5)	(85.7, 87.5)
	2	(76.7, 79.5)	(81.1, 84.8)	(86.2, 87.4)
	3	(77.5, 80.1)	(83.0, 84.5)	(85.3, 85.9)
$n = 500$	1	(83.8, 84.6)	(88.2, 89.9)	(90.9, 91.3)
	2	(82.8, 84.4)	(89.6, 90.8)	(91.2, 92.4)
	3	(83.4, 84.7)	(90.6, 89.3)	(91.8, 92.8)
$n = 1000$	1	(87.4, 85.6)	(90.6, 91.2)	(92.2, 91.6)
	2	(85.4, 87.8)	(91.4, 92.8)	(93.8, 93.4)
	3	(88.4, 89.0)	(90.4, 92.0)	(93.2, 93.2)

Table 4. $\ell = 9n^{1/2}$

	q	$\alpha = 0.1$	$\alpha = 0.5$	$\alpha = 0.9$
$n = 100$	1	(54.5, 52.7)	(56.3, 58.1)	(59.5, 62.2)
	2	(54.1, 53.3)	(55.6, 58.1)	(59.8, 63.1)
	3	(53.6, 53.9)	(57.3, 58.1)	(58.9, 63.0)
$n = 200$	1	(66.6, 67.5)	(72.2, 73.4)	(72.4, 74.8)
	2	(66.4, 70.0)	(71.6, 73.3)	(74.0, 75.6)
	3	(66.4, 69.1)	(73.5, 72.3)	(73.6, 75.6)
$n = 500$	1	(78.2, 76.7)	(84.6, 85.2)	(85.3, 86.1)
	2	(78.8, 77.0)	(84.4, 84.2)	(86.6, 87.6)
	3	(78.6, 77.0)	(83.3, 83.2)	(87.1, 87.0)
$n = 1000$	1	(84.0, 84.7)	(89.0, 88.1)	(91.7, 90.8)
	2	(83.4, 83.4)	(88.4, 88.2)	(90.2, 89.8)
	3	(83.8, 86.8)	(87.6, 88.6)	(90.4, 89.4)

The following features are evident in our results.

1. As expected, optimal coverage accuracy (i.e. coverage accuracy with the “best” choice of l) improves with increasing sample size and with decreasing strength of dependence (i.e. with increasing α). Coverage inaccuracy is associated with under-coverage rather than over-coverage.
2. Surprisingly, for fixed choice of c in the formula $l = cn^{1/2}$, coverage accuracy of the SW method depends little on the Hermite rank, q .
3. There is little difference between coverage probabilities with $c = 1$ and 3 in the formula $l = cn^{1/2}$, but performance deteriorates rapidly as c increases to 6 and 9. Overall, our simulations suggest that l should be no more than 20% of sample size; this is in agreement with findings of Hall and Jing (1996).

3.3. Exact distribution and its SW estimate

We calculate the exact distribution of studentized sample means by simulation. Figures 1 to 4, addressing the cases $n = 100, 200, 500, 1000$ respectively, compare the exact distribution with its sampling window estimator, also obtained by Monte Carlo simulation, for a specific sample. In all four figures we have taken $q = 2$, $\alpha = 0.5$ and $\ell = 3n^{1/2}$. In each figure the five thin solid lines are the sampling window estimates produced by five different, independent random samples. The thick line in each figure represents the “exact” distribution, computed by simulation. The exact distributions are well approximated by their sampling window estimates, and the latter vary only to a small extent, particularly for large n .

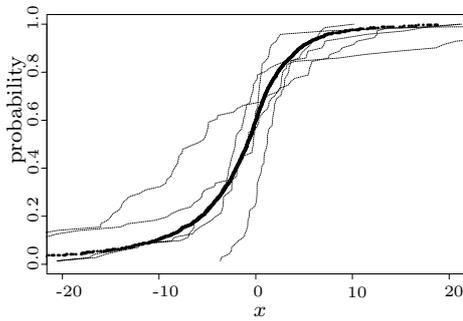


Figure 1. $n = 100$. (Thick line: 'Exact'; Thin lines: SW estimates)

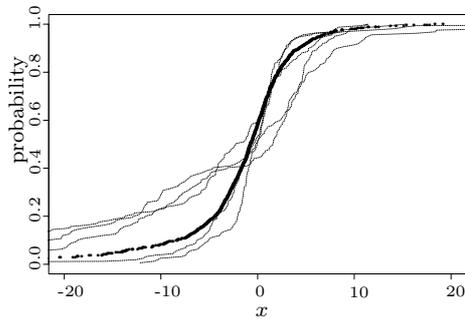


Figure 2. $n = 200$. (Thick line: 'Exact'; Thin lines: SW estimates)

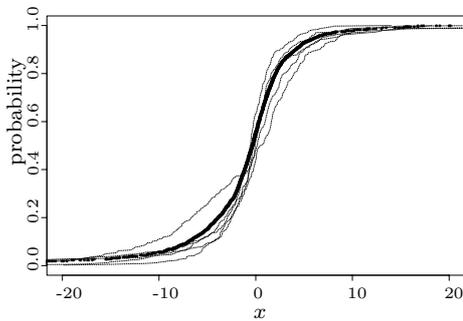


Figure 3. $n = 500$. (Thick line: 'Exact'; Thin lines: SW estimates)

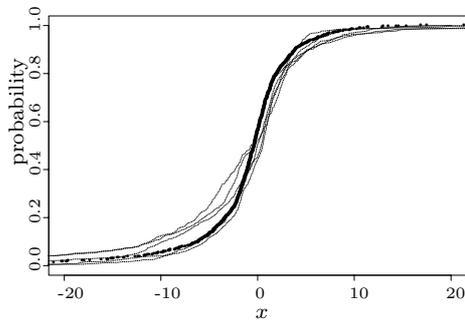


Figure 4. $n = 1000$. (Thick line: 'Exact'; Thin lines: SW estimates)

4. Proofs

We need an expression for the cross-product moments of Hermite polynomials in Gaussian variables. The result is essentially Lemma 3.2 of Taqqu (1977), but is recast here in a form that is more useful for proving our results.

Lemma 4.1. (Taqqu (1977)) *Let $\mathbf{Y} \equiv (Y_1, \dots, Y_m)^T$, for $m \geq 2$, be an m -dimensional Gaussian random vector with $E(Y_j) = 0$ and $E(Y_j^2) = 1$ for $1 \leq j \leq m$. Given $M \geq 1$, let $I_M = \{\mathbf{k} = (k_1, \dots, k_m)^T : k_1 + \dots + k_m = 2M, 0 \leq k_1, \dots, k_m \leq M\}$. Then, for any $\mathbf{k} \in I_M$,*

$$E\{H_{k_1}(Y_1) \cdots H_{k_m}(Y_m)\} = (k_1! \cdots k_m!) \sum_{\mathbf{k}, m, M} \prod_{j=1}^M E(Y_{t_j} Y_{s_j}),$$

where $\sum_{\mathbf{k}, m, M}$ extends over indices $(t_1, s_1), \dots, (t_M, s_M) \in \{1, \dots, m\}^2$ such that (a) $t_1 < t_2 < \dots < t_M$, (b) $t_1 < s_1, \dots, t_M < s_M$, and (c) for each $p = 1, \dots, m$,

exactly k_p elements in the array $\{t_1, s_1, \dots, t_M, s_M\}$ equal p . Furthermore,

$$E\{H_{k_1}(Y_1) \cdots H_{k_m}(Y_m)\} = 0 \quad \text{for all } (k_1, \dots, k_m) \notin \bigcup_{M=1}^{\infty} I_M.$$

Proof of Theorem 2.2. By Theorem 5.2.24 of Zygmund (1968), $r(k) \sim k^{-\alpha} L_1(1/k)$ as $k \rightarrow \infty$, and hence, by Theorem 2.1,

$$n(\bar{X}_n - \mu)/d_n \rightarrow W_q \tag{4.1}$$

in distribution, where the normalizing constants d_n are now defined by $d_n^2 = n^{2-q\alpha} L_1(1/n)^q$. Consequently, by the slow variation of $L_1(\cdot)$,

$$(\ell d_n)/(nd_\ell) = o(1). \tag{4.2}$$

In view of (4.1), this implies that $\ell(\bar{X}_n - \mu)/d_\ell = o_p(1)$. Hence, it is enough to show that

$$\sup_{x \in \mathbb{R}} |F_n^*(x) - F_n(x)| = o_p(1), \tag{4.3}$$

where $F_n^*(x) = N^{-1} \sum_{1 \leq i \leq N} I\{(S_{\ell i} - \ell\mu)/d_\ell \leq x\}$. Since the distribution of W_q is continuous, (4.3) holds provided $F_n^*(x) - F_n(x) = o_p(1)$ for each $x \in \mathbb{R}$.

Note that

$$\begin{aligned} & E\{F_n^*(x) - F_n(x)\}^2 \\ & \leq (2\ell + 1)N^{-1}F_\ell(x) + \frac{2}{N} \sum_{i=\ell+1}^{N-1} |P\{S_{\ell i}/d_\ell \leq x, S_{\ell(i+1)}/d_\ell \leq x\} - \{F_\ell(x)\}^2|, \end{aligned}$$

where for simplicity of notation, we have set $\mu = 0$ in the last line. Now by Theorem 5.5.7 of Ibragimov and Rozanov (1978), the second term on the right hand side tends to zero. Hence, Theorem 2.2 is proved.

We shall give a relatively detailed proof of Theorem 2.4, and following that, an outline of the proof of Theorem 2.3, which is derived in the same way.

Proof of Theorem 2.4. Without loss of generality, let $\mu = 0$. Put $F_\ell^*(x) = (n - \ell + 1)^{-1} \sum_{i=1}^{n-\ell+1} I(S_{\ell i}/s_\ell \leq x)$. By hypothesis, $n\bar{X}_n/s_n \rightarrow \zeta$, and so by (2.5), for each $\epsilon > 0$,

$$P(s_\ell^{-1}\ell|\bar{X}_n| > \epsilon) \rightarrow 0. \tag{4.4}$$

Now,

$$\begin{aligned} & \sup_{-\infty < x < \infty} |\widehat{F}_\ell(x) - F(x)| \\ & \leq \sup_{-\infty < x < \infty} |F_\ell^*(x) - F(x)| + \sup_{-\infty < x < \infty} |F(x) - F(x + \epsilon)| + I(s_\ell^{-1}\ell|\bar{X}_n| > \epsilon). \end{aligned}$$

In view of the continuity of F , $\sup_x |F(x) - F(x + \epsilon)| \rightarrow 0$ as $\epsilon \rightarrow 0$, and so by (4.4), it suffices to prove that $E\{\sup_x |F_\ell^*(x) - F(x)|\} \rightarrow 0$. Again using the continuity of F , the latter result will follow if we show that for each fixed x , $E|F_\ell^*(x) - F(x)| \rightarrow 0$, or equivalently, that

$$E\{F_\ell^*(x) - F(x)\}^2 = (n - \ell + 1)^{-2} \sum_{i_1=1}^{n-\ell+1} \sum_{i_2=1}^{n-\ell+1} \{P(s_\ell^{-1}S_{\ell i_1} \leq x, s_\ell^{-1}S_{\ell i_2} \leq x) - F(x)^2\} \rightarrow 0.$$

For this is sufficient to prove that for each $\epsilon > 0$, and each x ,

$$\sup_{\epsilon n \leq i \leq n} |P(s_\ell^{-1}S_{\ell i} \leq x, s_\ell^{-1}S_{\ell i} \leq x) - F(x)^2| \rightarrow 0. \tag{4.5}$$

Our proof of (4.5) is by contradiction. If (4.5) fails then there exist $\epsilon, \delta > 0$, sequences $n_k \uparrow \infty$ and $i(n_k) \in [\epsilon n_k, n_k]$, and a real number x , such that

$$|P(S_{\ell(n_k),1}/s_{\ell(n_k)} \leq x, S_{\ell(n_k),i(n_k)} \leq x) - F(x)^2| \geq \delta \tag{4.6}$$

for all k . Since the distribution of ζ is completely determined by its moments (4.6) is contradicted by (2.6), proving the theorem.

A little additional notation is necessary for the proofs of Theorems 2.3 and 2.5. For any function $g : \mathbb{R} \rightarrow \mathbb{R}$, let $\|g\|_\infty = \sup\{|g(x)| : x \in \mathbb{R}\}$ denote the supremum norm of g . For $i \geq 1$ and $m \geq 1$, write

$$\tilde{S}_{mi} = \sum_{j=i}^{i+m-1} C_q H_q(Z_j),$$

where C_q is as in (2.2). We derive Theorem 2.3 by applying arguments in the proof of Theorem 2.4, with X_j replaced by $C_q H_q(Z_j)$ and \tilde{S}_{mi} playing the role of S_{mi} for the latter choice of X_j .

Proof of Theorem 2.3. Without loss of generality, $\mu = 0$. For $x \in \mathbb{R}$ define $\tilde{F}_n(x) = N^{-1} \sum_{1 \leq i \leq N} I(\tilde{S}_{\ell i}/d_\ell \leq x)$ and $\tilde{F}(x) = P(W_q \leq x)$. By Theorem 3.1 and Corollary 3.1 of Taqqu (1975), and by (4.3) above,

$$\begin{aligned} \|\hat{F}_n - F_n\|_\infty &\leq \|\hat{F}_n - F_n^*\|_\infty + \|F_n^* - \tilde{F}_n\|_\infty + \|\tilde{F}_n - \tilde{F}\|_\infty + \|F_n - \tilde{F}\|_\infty \\ &= \|\tilde{F}_n - \tilde{F}\|_\infty + o_p(1), \end{aligned} \tag{4.7}$$

where F_n^* is as defined in the proof of Theorem 2.2. Hence, it is enough to show that the first term in the last line of (4.7) tends to zero in probability, for which purpose we apply Theorem 2.4. Note that by (4.2), (2.5) holds. Also, by

Theorem 3 of Taqqu (1977), for any $p \geq 1$, $E(\tilde{S}_{\ell 1}/d_\ell)^p \rightarrow E(W_q)^p$ as $n \rightarrow \infty$. Hence, Theorem 2.3 will follow if we show that for positive integers a and b , and any $\epsilon > 0$,

$$\max_{\epsilon n \leq i \leq n} \left| E\{(\tilde{S}_{\ell 1})^a(\tilde{S}_{\ell i})^b\} - (E\tilde{S}_{\ell 1})^a(E\tilde{S}_{\ell 1})^b \right| = o(d_\ell^{a+b}). \tag{4.8}$$

Fix integers $a, b > 0$, and let $\epsilon n \leq i \leq n$. Then,

$$\begin{aligned} & \left| E\{(\tilde{S}_{\ell 1})^a(\tilde{S}_{\ell i})^b\} - (E\tilde{S}_{\ell 1})^a(E\tilde{S}_{\ell 1})^b \right| \\ & \leq \sum_{1 \leq j_1, \dots, j_a \leq \ell} \sum_{i \leq i_1, \dots, i_b \leq i + \ell - 1} \Delta(j_1, \dots, j_a; i_1, \dots, i_b), \end{aligned} \tag{4.9}$$

where $\Delta(j_1, \dots, j_a; i_1, \dots, i_b)$ is defined to equal the absolute value of

$$E\left[\left\{\prod_{k=1}^a H_q(Z_{j_k})\right\}\left\{\prod_{k=1}^b H_q(Z_{i_k})\right\}\right] - E\left\{\prod_{k=1}^a H_q(Z_{j_k})\right\}E\left\{\prod_{k=1}^b H_q(Z_{i_k})\right\}.$$

If exactly one of qa and qb is odd then, using Lemma 4.1, it is easy to check that $\Delta(\mathbf{j}; \mathbf{i}) = 0$ for all $\mathbf{j} = (j_1, \dots, j_a)'$ and $\mathbf{i} = (i_1, \dots, i_b)'$. Hence, we need to consider the cases: (I) both qa and qb are even, and (II) both qa and qb are odd. First we treat case (I). Let $u = aq/2, v = bq/2, \lambda_k = j_k$ for $1 \leq k \leq a$, and $\lambda_{k+a} = i_k$ for $1 \leq k \leq b$. Then, by Lemma 4.1,

$$\Delta(\mathbf{j}; \mathbf{i}) = (q!)^{a+b} \left| \sum_{(a+b)} \prod_{k=1}^{u+v} r(\lambda_{t_k} - \lambda_{s_k}) - \left\{ \sum_{(a)} \prod_{k=1}^u r(j_{t_k} - j_{s_k}) \right\} \left\{ \sum_{(b)} \prod_{k=1}^v r(i_{t_k} - i_{s_k}) \right\} \right|,$$

where for any integer $k \geq 1$, we write $\sum_{(k)}$ for the summation $\sum_{\mathbf{q}, k, \langle kq/2 \rangle}$ of Lemma 4.1, with $\mathbf{q} = (q, \dots, q) \in \mathbb{R}^k$ and $\langle \cdot \rangle$ denoting the integer part function.

For any index $\{(t_1, s_1), \dots, (t_{u+v}, s_{u+v})\}$ under $\sum_{(a+b)}$ satisfying $\max\{s_k : 1 \leq k \leq u\} \leq a < t_{u+1}$, the term $\prod_{1 \leq k \leq u+v} r(\lambda_{t_k} - \lambda_{s_k})$ can be written as the product of the terms $\prod_{1 \leq k \leq u} r(j_{t_k} - j_{s_k})$ and $\prod_{u+1 \leq k \leq u+v} r(i_{t_k-a} - i_{s_k-a})$. Since each of the numbers $1, \dots, a + b$ appears exactly q times in the array $(t_1, s_1), \dots, (t_{u+v}, s_{u+v})$, and all the indices $(t_1, s_1), \dots, (t_u, s_u)$ are less than or equal to a whenever $s_u \leq a$, it follows that each of the numbers $1, \dots, a$ must appear q times in the array $(t_1, s_1), \dots, (t_u, s_u)$, and hence, each of the remaining numbers $a + 1, \dots, a + b$ must also appear exactly q times in the array $(t_{u+1}, s_{u+1}), \dots, (t_{u+v}, s_{u+v})$.

Therefore, $\prod_{1 \leq k \leq u+v} r(\lambda_{t_k} - \lambda_{s_k})$ can be written as the product of two terms, one coming from $\sum_{(a)}$ and the other from $\sum_{(b)}$. Conversely, each term in the cross-product of $\sum_{(a)} \prod_{1 \leq k \leq u} r(j_{t_k} - j_{s_k})$ and $\sum_{(b)} \prod_{1 \leq k \leq u} r(i_{t_k} - i_{s_k})$ corresponds to a term $\prod_{1 \leq k \leq u+v} r(\lambda_{t_k} - \lambda_{s_k})$ with $\max\{s_k : 1 \leq k \leq u\} \leq a < t_{u+1}$.

Next, write $\sum_{(a+b)} = \sum'_{(a+b)} + \sum''_{(a+b)}$, where $\sum''_{(a+b)}$ denotes summation over all indices $(t_1, s_1), \dots, (t_{u+v}, s_{u+v})$ under $\sum_{(a+b)}$ that satisfy $\max\{s_k : 1 \leq k \leq u\} \leq a < t_{u+1}$. Then, by the arguments above,

$$\Delta(\mathbf{j}, \mathbf{i}) = (q!)^{a+b} \left| \sum'_{(a+b)} \prod_{k=1}^{u+v} r(\lambda_{t_k} - \lambda_{s_k}) \right|. \quad (4.10)$$

Note that for any $(t_1, s_1), \dots, (t_{u+v}, s_{u+v})$ under $\sum'_{(a+b)}$, there exists k satisfying $1 \leq k \leq u+v$ such that $|\lambda_{t_k} - \lambda_{s_k}| = |j_m - i_p|$ for some $1 \leq m \leq a$ and $1 \leq p \leq b$. Now using (4.9), (4.10) and the decay rate of r , one can establish (4.8). This completes the proof when aq and bq are both even.

If both aq and bq are odd (case (II)) then a, b and q are odd integers. Furthermore, for any $1 \leq j_1, \dots, j_a \leq \ell$ and $i \leq i_1, \dots, i_b \leq i + \ell - 1$, by Lemma 4.1 and (4.9),

$$\Delta(\mathbf{j}; \mathbf{i}) = \left| E \left\{ \prod_{k=1}^a H_q(Z_{j_k}) \prod_{k=1}^b H_q(Z_{i_k}) \right\} \right| = (q!)^{a+b} \left| \sum_{(a+b)} \prod_{k=1}^{u+v} r(\lambda_{t_k} - \lambda_{s_k}) \right|,$$

where u, v and $\sum_{(a+b)}$ are as defined above. Fix any set of indices $(t_1, s_1), \dots, (t_{u+v}, s_{u+v})$ under $\sum_{(a+b)}$, and define the set I_1 [and I_2] consisting of all pairs (t_k, s_k) such that both t_k and s_k lie in the set $\{1, \dots, a\}$ [and in $\{a+1, \dots, a+b\}$, respectively]. Evidently, none of the numbers $\{1, \dots, a\}$ appears among the indices in I_2 , and similarly, none of $\{a+1, \dots, a+b\}$ appears in I_1 . Since each of $1, \dots, a+b$ must be repeated exactly q times in the array $t_1, s_1, \dots, t_{u+v}, s_{u+v}$, we must have $2|I_1| \leq aq$ and $2|I_2| \leq bq$, where $|J|$ denotes the cardinality of a set J . Noting that both aq and bq are odd we see that $|I_1| \leq (aq-1)/2$ and $|I_2| \leq (bq-1)/2$. Therefore, $|\{(t_1, s_1), \dots, (t_{u+v}, s_{u+v})\} \setminus (I_1 \cup I_2)| \geq 1$. Hence, there is at least one pair (t_k, s_k) such that $|\lambda_{t_k} - \lambda_{s_k}| = |i_m - j_p|$ for some $1 \leq m \leq b$ and $1 \leq p \leq a$. Now using the arguments that follow (4.10), one can establish (4.8) as in case (I). This completes the proof of Theorem 2.3.

Proof of Theorem 2.5. Put $M = n - m + 1$ and $\tilde{d}_m^2 = M^{-1} \sum_{1 \leq i \leq M} (S_{mi} - m\mu)^2$. In view of (4.2), if m denotes either m_1 or m_2 then

$$\begin{aligned} E|\tilde{d}_m^2 - d_m^{*2}| &\leq 4M^{-1} \sum_{i=1}^M m \left[E(\bar{X} - \mu)^2 \{E(S_{mi} - m\mu)^2 + m^2 E(\bar{X} - \mu)^2\} \right]^{1/2} \\ &= 4mn^{-1} d_n (d_m^2 + m^2 n^{-2} d_n^2)^{1/2} = o(d_m^2). \end{aligned} \quad (4.11)$$

Next, write $\bar{d}_m^2 = M^{-1} \sum_{1 \leq i \leq M} \tilde{S}_{mi}^2$. By Corollary 3.1 of Taqqu (1975) and the Cauchy-Schwartz inequality, for $m = m_1, m_2$,

$$\begin{aligned}
 & E|\tilde{d}_m^2 - \bar{d}_m^2| \\
 & \leq E\left[\left\{M^{-1} \sum_{i=1}^M (S_{mi} - m\mu + \tilde{S}_{mi})^2\right\}^{1/2} \times \left\{M^{-1} \sum_{i=1}^M (S_{mi} - m\mu - \tilde{S}_{mi})^2\right\}^{1/2}\right] \\
 & \leq \{2E(S_{m1} - m\mu)^2 + 2E\tilde{S}_{m1}^2\}^{1/2} \{E(S_{m1} - m\mu - \tilde{S}_{m1})^2\}^{1/2} = o(d_m^2). \quad (4.12)
 \end{aligned}$$

By (4.8), for $m = m_1$ or m_2 ,

$$E(\bar{d}_m^2 - E\bar{d}_m^2)^2 = O\left[n^{-2} \sum_{j_1=1}^M \sum_{j_2=1}^M \left| \text{Cov}\{(\tilde{S}_{mj_1})^2, (\tilde{S}_{mj_2})^2\} \right|\right] = o(d_m^4). \quad (4.13)$$

It follows that $E\bar{d}_{m_i}^2 = d_{m_i}^2 \{1 + o(1)\}$ for $i = 1, 2$, whence $\hat{d}_n/d_n \rightarrow 1$ in probability. This proves part (a) of the theorem. Parts (b), (c) and (d) now follow by applying part (a) and Theorems 2.2 and 2.3. This completes the proof of Theorem 2.5.

Acknowledgement

Jing’s research was supported by a Hong Kong RGC Competitive Earmarked Research Grant HKUST6162/97P and Lahiri’s research was partially supported by NSF grant no. DMS 9505124. The authors thank a referee for some helpful comments.

References

Athreya, K. B. and Fuh, C. D. (1992). Bootstrapping Markov chains: Countable case. *J. Statist. Plann. Inference* **33**, 311-331.

Beran, J. (1994). *Statistical Methods for Long Memory Processes*. Chapman and Hall, London.

Bickel, P., Götze, F. and van Zwet, W. R. (1997). Resampling fewer than n observations: gains, losses and remedies for losses. *Statist. Sinica* **7**, 1-31.

Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*. Cambridge University Press, Cambridge, U.K.

Bose, A. (1988). Edgeworth corrections by bootstrap in autoregressions. *Ann. Statist.* **16**, 1709-1722.

Carlstein, E. (1986). The use of subseries methods for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist.* **14**, 1171-1179.

Datta, S. and McCormick, W. P. (1992). Bootstrap for a finite state Markov chain based on I.I.D. resampling. In *Exploring the Limits of the Bootstrap* (Edited by R. Lepage and L. Billard), 77-98. Wiley, New York.

Dobrushin, R. L. (1979). Gaussian and their subordinated self-similar random generalized fields. *Ann. Probab.* **7**, 1-28.

Dobrushin, R. L. and Major, P. (1979). Non-central limit theorems for non-linear functionals of Gaussian fields. *Z. Wahr. Verw. Gebiete* **50**, 27-52.

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

Freedman, D. (1984). On bootstrapping two stage least squares estimates in stationary linear models. *Ann. Statist.* **12**, 827 - 842.

- Götze, F. and Künsch, H. R. (1996). Second-order correctness of the blockwise bootstrap for stationary observations. *Ann. Statist.* **24**, 1914-1933.
- Hall, P. (1985). Resampling a coverage pattern. *Stoch. Proc. Appl.* **20**, 231-246.
- Hall, P., Horowitz, J. L. and Jing, B.-Y. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* **82**, 561-574.
- Hall, P. and Jing, B.-Y. (1996). On sample reuse methods for dependent data. *J. Roy. Statist. Soc. Ser. B* **58**, 727-737.
- Ibragimov, I. A. and Rozanov, Y. A. (1978). *Gaussian Random Processes*. Springer, New York.
- Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17**, 1217-1261.
- Lahiri, S. N. (1991). Second order optimality of stationary bootstrap. *Statist. Probab. Lett.* **11**, 335-341.
- Lahiri, S. N. (1993). On the Moving block bootstrap under long range dependence. *Statist. Probab. Lett.* **18**, 405-413.
- Lahiri, S. N. (1996). On edgeworth expansion and moving block bootstrap for studentized M -estimators in multiple linear regression models. *J. Multivariate Anal.* **56**, 42-59.
- Liu, R. Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In *Exploring the Limits of the Bootstrap* (Edited by R. Lepage and L. Billard), 225-248. Wiley, New York.
- Politis, D. and Romano, J. P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Ann. Statist.* **22**, 2031-2050.
- Taqqu, M. S. (1975). Weak convergence to fractional brownian motion and to the Rosenblatt process. *Z. Wahr. Verw. Gebiete* **31**, 287-302.
- Taqqu, M. S. (1977). Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long range dependence. *Z. Wahr. Verw. Gebiete* **40**, 203-238.
- Taqqu, M. S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahr. Verw. Gebiete* **50**, 53-83.
- Zygmund, A. (1968). *Trigonometric Series* Volumes I, II. Cambridge University Press, Cambridge, U.K.

Centre for Mathematics and its Applications, Australian National University, Canberra, ACT 0200, Australia.

E-mail: halpstat@durra.anu.edu.au

Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong.

E-mail: majing@uxmail.ust.hk

Department of Statistics, Iowa State University, Ames, Iowa 50011, U.S.A.

E-mail: snlahiri@iastate.edu

(Received June 1997; accepted December 1997)