Econometric Theory, **18**, 2002, 79–98. Printed in the United States of America.

ON THE JACKKNIFE-AFTER-BOOTSTRAP METHOD FOR DEPENDENT DATA AND ITS CONSISTENCY PROPERTIES

S.N. LAHIRI
Iowa State University

Motivated by Efron (1992, *Journal of the Royal Statistical Society, Series B* 54, 83–111), this paper proposes a version of the moving block jackknife as a method of estimating standard errors of block-bootstrap estimators under dependence. As in the case of independent and identically distributed (i.i.d.) observations, the proposed method merely regroups the values of a statistic from different bootstrap replicates to produce an estimate of its standard error. Consistency of the resulting jackknife standard error estimator is proved for block-bootstrap estimators of the bias and the variance of a large class of statistics. Consistency of Efron's method is also established in similar problems for i.i.d. data.

1. INTRODUCTION

In many complex problems, the bootstrap method can be used to estimate different population quantities related to an estimator, such as its bias and variance. However, just as with any other estimation procedure, the accuracy of estimators produced by the bootstrap method needs to be assessed. This can be accomplished using different approaches, such as by explicit estimation of the leading terms in the asymptotic expansion of the mean squared error (MSE) or by using a second level of bootstrapping. The problem with the first approach is that it requires a good deal of analytical calculation with the results being applicable to the specific estimator and its given functional. On the other hand, the approach involving the use of a second level of bootstrapping, though applicable more generally, is computationally too demanding and still not practically viable with the present day computing advances (cf. Efron, 1992). In his seminal paper, Efron (1992) suggested using the jackknife to (nonparametrically) estimate the standard errors of bootstrap estimators. When the data are independent and identically distributed (i.i.d.), he obtained an ingenious representation result for the jackknife point values and used it to show (cf. Efron,

This research was partially supported by NSF grant DMS 9505124. Address correspondence to: S.N. Lahiri, Department of Statistics, Iowa State University, Ames, IA 50011, USA.

© 2002 Cambridge University Press 0266-4666/02 $9.50

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

1992, p. 83) that "the jackknife estimate of standard error for a bootstrap quantity can be computed from the original bootstrap replications, with no further resampling required." Furthermore, the pseudovalues for the bootstrap quantities need no additional computation; they are readily obtained from the bootstrap computation itself. The computational advantage and the generality of application rendered the jackknife-after-bootstrap (JAB) method of Efron (1992) particularly attractive for deriving variance estimators of bootstrap quantities in the i.i.d. case.

In this paper a similar method (also referred to as the JAB) is proposed in the context of bootstrapping dependent data. It is well known that instead of resampling single observations, one needs to resample blocks of observations to obtain valid bootstrap approximations under weak dependence. Similarly, to obtain valid jackknife estimators of the standard error of a statistic, one needs to delete blocks of observations at a time in the dependent case. The proposed method utilizes the underlying principles of the block jackknife and the block-bootstrap methods in an efficient way to ensure a similar representation of the block-jackknife point values of block bootstrap quantities. Furthermore, it produces variance estimators of block-bootstrap quantities with a computational simplicity similar to that of its i.i.d. counterpart. This is of considerable importance for at least two reasons: first, it automatically provides an estimator of the standard error of bootstrap values under dependence, which are often analytically too complicated for explicit estimation, and second, it gives a computationally efficient method to evaluate these estimates.

A precise description of the JAB method for dependent data is given in the next section. A simple result (namely, Proposition 2.1) proved in Section 5 provides the theoretical justification for the method. As in the i.i.d. case, the present version of the JAB can be used for many different applications, such as estimating the standard errors of block-bootstrap quantities, empirical choice of optimal block sizes for the block bootstrap (cf. Lahiri, 1996b), and so on.

The main results of the paper concern theoretical validity of the JAB in the dependent and the independent cases. It follows from these results that not only does the JAB enjoy some remarkable computational properties, it also has some attractive theoretical properties. In Section 3, which follows, we establish the consistency of the JAB estimators of the standard errors of bootstrap quantities in different problems. To describe the results briefly, suppose $\hat{\theta}_n$ is an estimator that can be represented as a smooth function of sample means. Let $\hat{B}_n$ and $\hat{V}_n$, respectively, denote the block-bootstrap estimators of the bias and the variance of $\hat{\theta}_n$. The main results of the paper establish consistency of the JAB variance estimators of $\hat{B}_n$ and $\hat{V}_n$. More specifically, it is proved that for independent and also dependent data,

$$\widehat{\mathrm{Var}}_{JAB}(\hat{\phi}_n)/\mathrm{Var}(\hat{\phi}_n) \to_p 1 \tag{1.1}$$

as $n \to \infty$, for $\hat{\phi}_n = \hat{B}_n$ and $\hat{V}_n$, where $\widehat{\mathrm{Var}}_{JAB}(\hat{\phi}_n)$ denotes the JAB estimator of the variance of $\hat{\phi}_n$. For independent data, the JAB variance estimator of a

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

bootstrap estimator is defined using the standard jackknife variance estimator formula (cf. Efron, 1992). In the dependent case, however, the definition of $\widehat{\mathrm{Var}}_{JAB}(\hat{\phi}_n)$ is different from the usual block-jackknife variance estimator of $\hat{\phi}_n$. Indeed, we delete blocks of bootstrap blocks for defining the JAB variance estimator as compared to deleting blocks of observations for the block-jackknife variance estimator. This modification is a *distinguishing* feature of the JAB in the dependent case from its i.i.d. counterpart and is crucial for effectively *reusing* the values of an estimator from the original block-bootstrap computations. In the context of *bootstrapping* dependent data, "blocks of blocks" resampling has been considered by Politis and Romano (1992b) and Bühlmann and Künsch (1995), among others. To get some idea about finite sample properties of the JAB variance estimators in the dependent case, we also carry out a a small simulation study. These results are presented in Section 4.

Properties of jackknife estimators under independence have been studied by several authors: see Reeds (1978), Parr and Schucany (1982), Parr (1985), Shao (1989), Shao and Wu (1989), and the references therein. In an important paper, the use of jackknife to bootstrap estimates was first suggested by Beran (1984) for i.i.d. data in a different context. In his 1992 seminal paper, Efron formulated the JAB method for i.i.d. data and established its computational efficacy. Block-bootstrap and block-jackknife methods for dependent data have been put forward by several authors, notably by Hall (1985), Carlstein (1986), Künsch (1989), Liu and Singh (1992), and Politis and Romano (1992a). Properties of block-bootstrap methods have been studied by Lahiri (1991, 1996a), Hall, Horowitz, and Jing (1995), Hall and Horowitz (1996), Götze and Künsch (1996), and Li and Madala (1997), among others.

The rest of the paper is organized as follows. In Section 2, JAB is formulated for dependent data. Consistency of JAB variance estimators for the bias and the variance functionals for dependent and also independent data are established in Section 3. Section 4 reports results from a simulation in the dependent case. Finally, Section 5 gives proofs of the results from Sections 2 and 3.

2. THE JACKKNIFE-AFTER-BOOTSTRAP METHOD

We begin with a description of the block-bootstrap methods for dependent data. Assume that $X_1, X_2, \ldots$ is a stationary sequence of random vectors with values in $\mathbb{R}^d$. Given the data $\mathcal{X}_n = (X_1, \ldots, X_n)$, we define the blocks $\mathcal{B}_j = (X_j, \ldots, X_{j+\ell-1}), j = 1, \ldots, N$ where $N = n - \ell + 1$ and $1 \le \ell \le n$ denotes the block size. Let b be the largest integer not exceeding n/ℓ. Then the bootstrap blocks are obtained by selecting a random sample of size b from a suitable subset of the blocks $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$. For the moving block-bootstrap (MBB) method of Künsch (1989) and Liu and Singh (1992), one selects a random sample $\mathcal{B}_1^*, \ldots, \mathcal{B}_b^*$ from the whole collection $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$. On the other hand, for Carlstein's (1986) method, $\mathcal{B}_1^*, \ldots, \mathcal{B}_b^*$ are selected from the collection of nonoverlapping blocks $\{\mathcal{B}_{(i-1)\ell+1} : 1 \le i \le b\}$ only. Because each of the resampled blocks contains ℓ data values, arranging the components of $\mathcal{B}_1^*, \ldots, \mathcal{B}_b^*$

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

into a sequence yields $n_1 \equiv b \cdot \ell$ bootstrap samples $X_1^*, \ldots, X_{n_1}^*$. Note that here $n_1/n \to 1$ as $n \to \infty$. An alternative to this is to select $(b+1)$ bootstrap blocks (instead of b) and take the first n values out of $(b+1)\ell$ resampled values as the bootstrap sample. However, because $\ell/n \to 0$ as $n \to \infty$, the difference between the two alternatives is insignificant, and we will use the first one.

Let $\hat{\theta}_n = H(\bar{X}_n)$ be an estimator of the parameter $\theta = H(\mu)$, where $H : \mathbb{R}^d \to \mathbb{R}$ is a (smooth) function, $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ and $\mu = EX_1$. Then, the bootstrap version of $\hat{\theta}_n$ is given by $\theta_n^* = H(\bar{X}_n^*)$, where $\bar{X}_n^* = n_1^{-1} \sum_{i=1}^{n_1} X_i^*$. Write $T_n = \hat{\theta}_n - \theta$. Let $\phi_n = \phi(G_n)$ be a functional of the sampling distribution G_n of T_n. For example, $\phi_n = \int x dG_n$ gives the bias of $\hat{\theta}_n$ and $\phi_n = \int x^2 dG_n - (\int x dG_n)^2$ gives the variance of $\hat{\theta}_n$. The bootstrap estimator of ϕ_n is then given by

$$\hat{\phi}_n = \phi(\hat{G}_n),$$

where $\hat{G}_n$ is the conditional distribution of $T_n^* = H(\bar{X}_n^*) - H(\hat{\mu}_n)$, given the data $\mathcal{X}_n$, and $\hat{\mu}_n$ is the conditional expectation of $\bar{X}_n^*$ given $\mathcal{X}_n$. Under either of the block-bootstrap methods, $\hat{G}_n$ is known once the data have been observed, and hence, $\hat{\phi}_n$ can be calculated. However, an exact evaluation of $\hat{G}_n$ is difficult. In practice, $\hat{G}_n$ is approximated by Monte-Carlo simulations, by drawing a large number (say, K) replicates $({}_k\mathcal{B}_1^*, \ldots, {}_k\mathcal{B}_b^*), k = 1, \ldots, K$, calculating the bootstrap variable ${}_kT_n^*$ based on each replicate, and replacing $\hat{G}_n$ by the empirical distribution $\hat{G}_n^*$ of $\{{}_kT_n^*, k = 1, \ldots, K\}$.

Next we describe the block-jackknife method for dependent data. Here, too, one may use the overlapping or the nonoverlapping versions of the jackknife to derive bias and variance estimators. For brevity, we shall describe only the overlapping version (which deletes a block of size ℓ at time). Suppose that $\hat{\gamma}_n = g_n(\mathcal{X}_n)$ is an estimator of a parameter γ. For $i = 1, \ldots, N$, define the ith jackknife block-deleted point value $\hat{\gamma}_n^{(i)}$ by $\hat{\gamma}_n^{(i)} = g_{n-\ell}(\mathcal{X}_{n,i})$, where $\mathcal{X}_{n,i} \equiv \mathcal{X}_n \backslash \mathcal{B}_i$ denotes the set of data values after deleting the ith block. The block-jackknife estimate of the variance of $\hat{\gamma}_n$ is given by (cf. Liu and Singh, 1992; Künsch, 1989)

$$\widehat{\mathrm{Var}}_{BJ}(\hat{\gamma}_n) = (\ell(n-\ell)^{-1})N^{-1} \sum_{i=1}^{N} (\tilde{\gamma}_n^{(i)} - \hat{\gamma}_n)^2, \tag{2.1}$$

where $\tilde{\gamma}_n^{(i)} = \ell^{-1}(n\hat{\gamma}_n - (n-\ell)\hat{\gamma}_n^{(i)})$ is the ith jackknife pseudovalue of $\hat{\gamma}_n$. Note that if $\ell \equiv 1$, i.e., if we delete blocks of size one or a single observation at a time, the variance estimator given by (2.1) reduces to its i.i.d. counterpart:

$$\widehat{\mathrm{Var}}_{J}(\hat{\gamma}_n) = (n(n-1))^{-1} \sum_{i=1}^{n} (\tilde{\gamma}_n^{(i)} - \hat{\gamma}_n)^2. \tag{2.2}$$

Now we are in a position to describe the JAB for dependent data. In principle, the formulation of the JAB for dependent data proceeds in a similar path as in the i.i.d. case, in the sense that we apply (a version of the block-) jackknife method to a block-bootstrap estimator. However, a direct application of the block-

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

jackknife method to bootstrap estimators does not allow us to obtain the desired representation of the jackknife block-deleted point values in terms of the original bootstrap values. As a result, to achieve computational advantages similar to the i.i.d. case, the application of the jackknife method to the bootstrap quantities in the JAB requires some modifications under dependence. The key step in the formulation is that we delete blocks of bootstrap blocks instead of blocks of original data values to define the JAB estimator of the standard error of a block-bootstrap quantity.

For definiteness, first we consider the case of the overlapping block bootstrap. Let $\hat{\phi}_n$ be the block-bootstrap estimator of ϕ_n based on blocks of size ℓ. Let m be an integer between 1 and N, and let $M = N - m + 1$. Recall that here $N = n - \ell + 1$ gives the total number of observed blocks $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$ of length ℓ. For $i = 1, \ldots, M$, define the set $I_i = \{1, \ldots, N\} \backslash \{i, \ldots, i + m - 1\}$. To define the ith jackknife block-deleted point value $\hat{\phi}_n^{(i)}$, we need to resample b blocks from the reduced collection $\{\mathcal{B}_j : j \in I_i\}$ and apply the functional ϕ to the resulting block-bootstrap estimator $\hat{G}_{n,i}$ of G_n. Specifically, let $J_{i1}, \ldots, J_{ib}$ be a collection of b random variables such that conditional on $\mathcal{X}_n$, they are i.i.d. with the discrete uniform distribution on I_i, i.e.,

$$P_*(J_{i1} = j) = (N - m)^{-1} \quad \text{for all} \quad j \in I_i.$$

(Here and in the following discussion, P_* and E_*, respectively, denote the conditional probability and conditional expectation, given the data $\mathcal{X}_n$.) Then, the desired resampled blocks are given by $\{\mathcal{B}_j^{*i} \equiv \mathcal{B}_{J_{ij}} : j = 1, \ldots, b\}$. Let $\bar{X}_n^{*i}$ denote the bootstrap sample mean based on the $n_1 \equiv \ell \cdot b$ resampled data values in $\{\mathcal{B}_j^{*i} : j = 1, \ldots, b\}$ and let $\hat{\mu}_{n,i} = E_* \bar{X}_n^{*i} = (N - m)^{-1} \sum_{j \in I_i} U_j$, where $U_j \equiv (X_j + \cdots + X_{j+\ell-1})/\ell$ denotes the jth block average. Then, $\hat{G}_{n,i}$ is the conditional distribution (given $\mathcal{X}_n$) of $T_{n,i}^* \equiv (H(\bar{X}_n^{*i}) - H(\hat{\mu}_{n,i}))$. We define the ith jackknife block-deleted point value $\hat{\phi}_n^{(i)}$ by

$$\hat{\phi}_n^{(i)} = \phi(\hat{G}_{n,i}),$$

$i = 1, \ldots, M$. The JAB estimators for the variance of $\hat{\phi}_n$ are then defined using (2.1) as

$$\hat{Var}_{JAB}(\hat{\phi}_n) = (m(N - m)^{-1}) M^{-1} \sum_{i=1}^{M} (\tilde{\phi}_n^{(i)} - \hat{\phi}_n)^2, \tag{2.3}$$

where $\tilde{\phi}_n^{(i)} = m^{-1}(N\hat{\phi}_n - (N - m)\hat{\phi}_n^{(i)})$.

To implement the JAB in practice, one proceeds as follows. Suppose $({}_k\mathcal{B}_1^*, \ldots, {}_k\mathcal{B}_b^*)$, $k = 1, \ldots, K$ denote the K (conditionally independent) sets of bootstrap blocks drawn with replacement from $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$ for evaluating the *original* bootstrap quantity $\hat{\phi}_n$ by Monte-Carlo simulation (cf. the middle box in Figure 1, which follows). For $1 \le i \le M$, let I_i^* denote the set of all indices $k, 1 \le k \le K$, such that none of the resampled blocks ${}_k\mathcal{B}_1^*, \ldots, {}_k\mathcal{B}_b^*$ in the kth set equals $\mathcal{B}_i, \ldots, \mathcal{B}_{i+m-1}$. Then, it follows from Proposition 2.1, which follows, that for each $k \in I_i^*$, the blocks $\{{}_k\mathcal{B}_1^*, \ldots, {}_k\mathcal{B}_b^*\}$ behave as a random

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

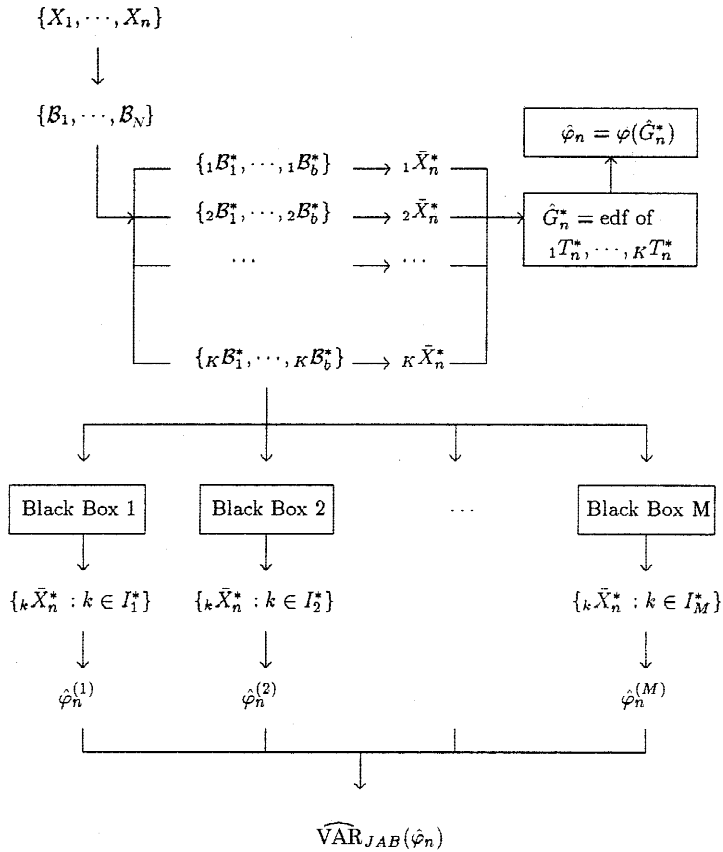


FIGURE 1. A schematic description of the MBB and the JAB in the dependent case. Here, $\{{}_k\mathcal{B}_1^*, \ldots, {}_k\mathcal{B}_b^*\}$, $k = 1, \ldots, K$ denote the resampled blocks, initially generated for Monte-Carlo evaluation of the MBB estimator $\hat{\phi}_n$, and similarly, $\{{}_k\bar{X}_n^* : k = 1, \ldots, K\}$ denote the values of the corresponding bootstrap statistic (the sample mean, in this case). The subcollection $\{{}_k\bar{X}_n^* : k \in I_j^*\}$, used for computing the jth jackknife block-deleted point value $\hat{\phi}_n^{(j)}$ for the JAB variance estimator is obtained by passing (*reusing*) these quantities through the jth black box for $j = 1, \ldots, M$. The JAB variance estimator is then obtained from $\hat{\phi}_n^{(j)}$'s using the formula (2.3).

sample of size b from the reduced collection $\{\mathcal{B}_j : j \in I_i\}$. Hence, for evaluating $\hat{\phi}_n^{(i)}$, we *extract* the desired subcollection $\{({}_k\mathcal{B}_1^*, \ldots, {}_k\mathcal{B}_b^*) : k \in I_i^*\}$ from the full collection $\{({}_k\mathcal{B}_1^*, \ldots, {}_k\mathcal{B}_b^*) : k = 1, \ldots, K\}$ (using "black box i" as, say, in Figures 1 and 2) and *reuse* the original bootstrap replicates *without additional resampling*. Define ${}_kT_{n,i}^* = (H({}_k\bar{X}_n^*) - H(\hat{\mu}_{n,i}))$ where ${}_k\bar{X}_n^*$ is the sample mean based on $\{{}_k\mathcal{B}_1^*, \ldots, {}_k\mathcal{B}_b^*\}$ and where $\hat{\mu}_{n,i}$ is as defined before. Then, the Monte-Carlo approximation to $\hat{\phi}_n^{(i)}$ is given by $\phi(\hat{G}_{n,i}^*)$ where $\hat{G}_{n,i}^*$ is the empirical distribution of $\{{}_kT_{n,i}^* : k \in I_i^*\}$. The JAB variance estimator is ob-

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

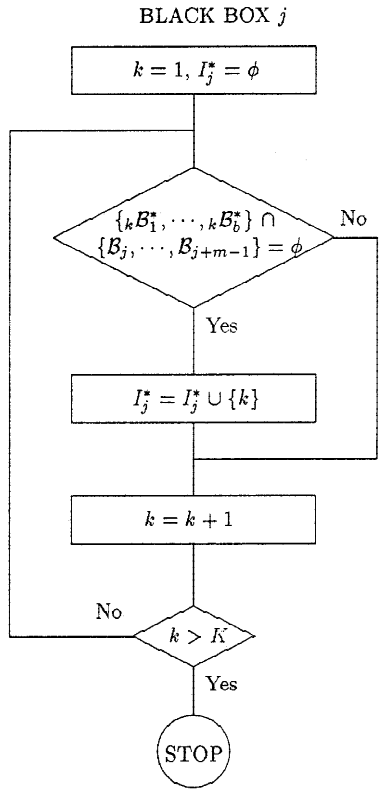


FIGURE 2. A flowchart summarizing the steps in the *j*th black box, used in the computation of the *j*th jackknife block-deleted point value $\hat{\phi}_n^{(j)}$.

tained from (2.3) by replacing $\hat{\phi}_n$ and $\hat{\phi}_n^{(i)}$'s with their Monte-Carlo values $\phi(\hat{G}_n^*)$ and $\phi(\hat{G}_{n,i}^*), 1 \le i \le M$, respectively.

The theoretical basis for this is contained in the following simple representation result. Let $1(S)$ denote the indicator function of a statement S such that $1(S) = 1$ or 0 according as S is true or false.

PROPOSITION 2.1. *Let $J_1, \ldots, J_b$ be i.i.d. random variables with the discrete uniform distribution on $\{1, \ldots, N\}$, and let $J_{i1}, \ldots, J_{ib}$ be i.i.d. random variables with the discrete uniform distribution on $I_i, 1 \le i \le N$. Write $\hat{p}_i = b^{-1} \sum_{j=1}^{b} 1(J_j \in I_i)$ for the proportion of random indices $J_1, \ldots, J_b$ that take values in the set I_i. Then, for any $i = 1, \ldots, M$, the conditional joint distribution of $(J_1, \ldots, J_b)$ given $\hat{p}_i = 1$ is the same as the joint distribution of $(J_{i1}, \ldots, J_{ib})$.*

Thus, it follows from the proposition that if the original resampled blocks $\{(\mathcal{B}_1^{*k}, \ldots, \mathcal{B}_b^{*k}) : k = 1, \ldots, K\}$ for the Monte-Carlo simulation of $\hat{\phi}_n$ are gen-

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

erated by the random indices $\{(J_1^k,\ldots,J_b^k): k = 1,\ldots,K\}$, then for each $k \in I_i^*, (J_1^k,\ldots,J_b^k)$ represents a random sample of size b from the reduced index set $I_i = \{1,\ldots,N\}\backslash\{i,\ldots,i+m-1\}$. Therefore, for n fixed, conditionally on $\mathcal{X}_n$, the empirical distribution $\hat{G}_{n,i}^*$ of $\{T_{n,i}^{*k}: k \in I_i^*\}$ converges to $\hat{G}_{n,i}$ almost surely in the supremum norm as $K \to \infty$, by the Glivenko–Cantelli theorem. Hence, if ϕ is a sufficiently smooth functional, then $\phi(\hat{G}_{n,i}^*)$ provides a reasonable approximation to the jackknife block-deleted point value $\hat{\phi}_n^{(i)}$ for large values of K.

Next we describe the JAB for the nonoverlapping version of the block bootstrap. In this case, the conceptual and theoretical details of the JAB are very similar. The only differences are that the random indices $J_1,\ldots,J_b$ generating the blocks for the original bootstrap quantity $\hat{\phi}_n$ are i.i.d. random vectors with the discrete uniform distribution on $I \equiv \{(j-1)\ell + 1 : 1 \le j \le b\}$ and the jackknife point values $\hat{\phi}_n^{(i)}$'s are obtained by using random samples $J_{i1},\ldots,J_{ib}$ from $I_i' \equiv I\backslash\{(j-1)\ell + 1 : i \le j \le i + m - 1\}$, for $1 \le i \le b - m + 1$. It is easy to see that an analogue of Proposition 2.1 holds for the nonoverlapping case, if we replace $\{1,\ldots,N\}$ and I_i by I and I_i', respectively. Hence, for implementing the JAB, one may again use the subset of the original bootstrap replicates having all indices lying in I_i' to calculate $\hat{\phi}^{(i)}, 1 \le i \le b - m + 1$.

In the next section, we investigate the consistency properties of the JAB variance estimators. For the sake of brevity, we restrict our attention to the overlapping block-bootstrap estimators. Analogues of all the results presented here hold for the nonoverlapping version of the block-bootstrap also. It is worth mentioning that though the JAB variance estimator is consistent for the block bootstrap estimator $\hat{\phi}_n$ of different population quantities ϕ_n, the JAB estimator of the bias of $\hat{\phi}_n$ is typically inconsistent (cf. Lahiri, 1997). Because the bias of a block-bootstrap estimator can be of comparable order of magnitude to its standard deviation, for practical applications it is also very important to be able to estimate the bias part consistently. Lahiri (1996b, 1997) suggests some modification to the standard jackknife bias estimator and establishes consistency of the modified estimators for different $\phi's$. Let ℓ_0 be an integer such that $\ell/\ell_0 \to 0$ as $n \to \infty$. Then, the modified estimators of the $\text{Bias}(\hat{\phi}_n)$ are given by

$$\widehat{\text{Bias}}(\hat{\phi}_n) = M^{-1}\sum_{i=1}^{M}(\tilde{\phi}_n^{(i)} - \hat{\phi}_n(\ell_0)) \quad \text{and} \quad \widetilde{\text{Bias}}(\hat{\phi}_n) = (\hat{\phi}_n - \hat{\phi}_n(\ell_0)),$$

where $\hat{\phi}_n(\ell_0)$ is the block-bootstrap estimator of ϕ_n based on blocks of size ℓ_0. Though these estimators require computation of $\hat{\phi}_n$ for a second block length ℓ_0, the total amount of computation involved in evaluating both the JAB variance and the bias estimators is still negligible compared with the other alternatives, such as the "double block bootstrap" (which involves bootstrapping blocks of original bootstrap blocks). Thus, with either of the bias estimators, the JAB provides a computationally efficient and consistent method for estimating the *mean squared error* (MSE) of block-bootstrap quantities. As a consequence, it has potential applications in different inference problems as in the i.i.d. case (cf. Efron, 1992; Datta and McCormick, 1995).

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

3. CONSISTENCY

The main results of this section establish consistency of the JAB variance estimators for two smooth functionals $\phi(\cdot)$, namely, the bias and the variance of estimators, for dependent and independent data. Let $X_1, X_2, \ldots$ be a sequence of stationary random vectors in $\mathbb{R}^d$ with $EX_1 = \mu$. Let $\theta = H(\mu)$ be the parameter of interest, where $H: \mathbb{R}^d \to \mathbb{R}$ is a smooth function. We take $\hat{\theta}_n = H(\bar{X}_n)$ as an estimator of θ, where $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$ denotes the sample mean. The functionals of interest are the bias and the variance of $\hat{\theta}_n$, given by

$$\phi_{1n} = E\hat{\theta}_n - \theta \quad \text{and} \quad \phi_{2n} = \mathrm{Var}(\hat{\theta}_n). \tag{3.1}$$

Looking at suitable transformations of the data, the proposed framework allows us to consider the bias and the variance of estimators such as the sample variance, sample lag-covariances, the sample correlation coefficient, and sample versions of partial autocorrelations. Let $\hat{\phi}_{1n}$ and $\hat{\phi}_{2n}$ denote the corresponding block-bootstrap estimators of ϕ_{1n} and ϕ_{2n} based on blocks of size ℓ. Write $\widehat{\mathrm{Var}}_{JAB}(\hat{\phi}_{jn})$ for the JAB estimator of the variance of $\hat{\phi}_{jn}, j = 1,2$, obtained by deleting blocks of m bootstrap blocks at a time (cf. (2.3)).

For proving the consistency of the JAB variance estimators of ϕ_{1n} and ϕ_{2n}, we shall assume that the sequence $\{X_n\}$ is strongly mixing at a polynomial rate. For any $1 \le a \le b \le \infty$, let $\mathcal{F}_a^b$ denote the σ-field generated by $\{X_i : a \le i \le b\}$. Define the strong mixing coefficient of $\{X_n\}$ by

$$\alpha(n) \equiv \mathrm{Sup}\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+n}^\infty, k \ge 1\}.$$

Also, for $r \ge 1$ and $\delta > 0$, define $\Delta(r,\delta) = \sum_{n=1}^\infty n^{2r-1}\alpha(n)^{\delta/(2r+\delta)}$.

The following assumptions will be used for proving the results of this section.

(A.1) $H: \mathbb{R}^d \to \mathbb{R}$ is four times continuously differentiable, and for all $\nu \in \mathbf{Z}_+^d, |\nu| = 4$,

$$|D^\nu H(x)| \le C(1 + \|x\|^a), \qquad x \in \mathbb{R}^d$$

for some integer $a \ge 1$, and $C > 0$, where $\mathbf{Z}_+ = \{0,1,2,\ldots\}$, and for a vector $\nu = (\nu_1,\ldots,\nu_d)' \in \mathbf{Z}_+^d$, $|\nu| = \nu_1 + \cdots + \nu_d$ and D^ν denotes the νth-order partial differentiable operator $\partial^{\nu_1+\cdots+\nu_d}/\partial x_1^{\nu_1}\ldots\partial x_d^{\nu_d}$.

(A.2) There exists $\delta \in (0,1]$ such that

$$E\|X_1\|^{8+2a+\delta} < \infty \quad \text{and} \quad \Delta(4+a;\delta) < \infty.$$

(A.3) There exists $\delta \in (0,1]$ such that

$$E\|X_1\|^{12+4a+\delta} < \infty \quad \text{and} \quad \Delta(6+2a;\delta) < \infty.$$

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

Here (A.2) and (A.3) give, respectively, the moment and mixing conditions for proving the consistency of the JAB estimators of ϕ_{1n} and ϕ_{2n}.

THEOREM 3.1. *Assume that as* $n \to \infty$, $\ell^{-1} + n^{-1/2}\ell + m^{-1}\ell = o(1)$ *and that* $n^{-1}\ell^{-2}m^3 = O(1)$. *Then, for* $j = 1,2$,

$$\widehat{\mathrm{Var}}_{JAB}(\hat{\phi}_{jn})/\mathrm{Var}(\hat{\phi}_{jn}) \to_p 1 \quad as \quad n \to \infty,$$

provided Assumptions (A.1) and (A.(j + 1)) hold.

Theorem 3.1 shows that with appropriate choices of the resampling parameters ℓ and m, the JAB variance estimators of ϕ_{1n}, ϕ_{2n} are consistent. Consistency of the block-jackknife method for estimating the variance of a class of "well-behaved" estimators has been established by Künsch (1989) and Liu and Singh (1992). However, their results do not readily apply to our case. This is primarily due to the fact that the JAB estimators are defined by deleting blocks of blocks as opposed to deleting blocks of original observations in the case of the usual block-jackknife method.

Next we consider the i.i.d. case. In this case, the JAB variance estimators are defined by applying the i.i.d. jackknife method to i.i.d. bootstrap estimators as described in Section 2. The following theorem establishes the consistency of the JAB variance estimator of Efron(1992) in the i.i.d. case, for $\hat{\phi}_{jn}$'s.

THEOREM 3.2. *Assume that* $\{X_n\}_{n\geq 1}$ *are i.i.d., and that* $\widehat{\mathrm{Var}}_{JAB}(\hat{\phi}_{jn})$ *denote the i.i.d. JAB estimator of* $\mathrm{Var}(\hat{\phi}_{jn}), j = 1,2$, *defined using (2.3) with* $m \equiv 1$ *and* $\ell \equiv 1$ *for all* $n \geq 1$. *Then, for* $j = 1,2$,

$$\widehat{\mathrm{Var}}_{JAB}(\hat{\phi}_{jn})/\mathrm{Var}(\hat{\phi}_{jn}) \to_p 1 \quad \text{as} \quad n \to \infty,$$

provided (A.1) holds and $E\|X_1\|^{4+j(4+2a)} < \infty$.

4. A SIMULATION STUDY

In this section, we report the results of a small simulation study on finite sample performance of the JAB variance estimators. We considered four different time-series models, given by

(I) $X_{1,t} = 0.6X_{1,t-1} - 0.3X_{1,t-2} + 0.1X_{1,t-3} + \epsilon_t$

(II) $X_{2,t} = 0.2\epsilon_t + 0.3\epsilon_{t-1}$

(III) $X_{3,t} = 0.6X_{3,t-1} - 0.3X_{3t-2} + 0.1X_{3,t-3} + 0.2\epsilon_t + 0.3\epsilon_{t-1} + 0.1\epsilon_{t-2}$

(IV) $X_{4,t} = X_{3,t}^2 + X_{3,t}1(X_{3,t} < 0)$,

where $1(\cdot)$ denotes the indicator function and $X_{3,t}$ is as in model III. The error distribution $\{\epsilon_t\}$ in all cases was $N(0,1)$. Note that the marginal distribution of the $X_{i,t}$'s under the first three models is *symmetric* but that under model (IV) it is *asymmetric* (with a thicker right tail). For each of the models (I)–(IV), we

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

TABLE 1. Means and variances of MBB estimators of $\phi_{2n} = \mathrm{Var}(\bar{X}_n)$

Model	ϕ_{2n}	ℓ	$E(\hat{\phi}_{2n}(\ell))$	$\mathrm{Var}(\hat{\phi}_{2n}(\ell))$
I	0.01378	5	0.0117	6.0656e−06
	0.01378	10	0.0122	1.1960e−05
II	0.00129	5	0.0027	1.6732e−07
	0.00129	10	0.0019	1.5395e−07
III	0.00230	5	0.0057	9.9745e−07
	0.00230	10	0.0039	6.8011e−07
IV	0.02185	5	0.0192	6.1101e−05
	0.02185	10	0.0196	8.6124e−05

Note: Column 2 gives values of the parameter ϕ_{2n}; column 3 gives the block sizes $\ell = 5,10$ for the block-bootstrap methods. Columns 3 and 4, respectively, give the mean and the variances of the block-bootstrap estimators $\hat{\phi}_{2n}(\ell)$, evaluated using 10,000 simulation runs.

considered a sample size of $n = 100$ and applied the MBB with two different block lengths $\ell = 5,10$. Using the notation of Sections 2 and 3, we denote the MBB estimator of the variance functional $\phi_{2n} \equiv \mathrm{Var}(\bar{X}_n)$ as $\hat{\phi}_{2n}(\ell)$. Then, the *parameter of interest* here is given by $\mathrm{Var}(\hat{\phi}_{2n}(\ell))$, which is estimated by the JAB method. Table 1 gives the values of the parameter $\phi_{2n} \equiv \mathrm{Var}(\bar{X}_n)$ and (the mean and) the variance of its MBB estimators $\hat{\phi}_{2n}(\ell)$. The parameter ϕ_{2n} was evaluated using 100,000 simulation runs and the quantity $E(\hat{\phi}_{2n}(\ell))$ and $\mathrm{Var}(\hat{\phi}(\ell))$ were evaluated using 10,000 simulation runs, with $K = 1,000$ bootstrap replications for each block size and each of the 10,000 simulation runs. Thus, the target population parameters for the JAB method are given by the eight numbers in the last column of Table 1 for the two block sizes under models I–IV.

We summarize the performance of the JAB estimators of $\mathrm{Var}(\hat{\phi}_{2n}(\ell))$ in Table 2. For each model and each block size, we considered several values of the jackknife block length m, applying the formula $m = Cn^{1/3}\ell^{2/3}$ with $C = 0.1,0.5,1,2$. To appreciate why this may yield a reasonable choice of m, note that the JAB applies a version of the moving block-jackknife method to the block-bootstrap estimator $\hat{\phi}_{2n}(\ell)$ by deleting "blocks" of *m overlapping* blocks. Results of Künsch (1989) and Hall et al. (1995) show that the MSE-optimal block length for the block-jackknife variance estimator for a sample of size T is of the form $C \cdot T^{1/3}$. Because there are "$b \equiv [n/\ell]$" (the integer part of n/ℓ) "independent" or "distinct" blocks of length ℓ in a sample of size n, for a block-bootstrap estimator based on *nonoverlapping* blocks of length ℓ, the optimal block length for its jackknife variance estimator is of the form $C \cdot b^{1/3}$. Note that there are (approximately) $C \cdot b^{1/3}\ell$ *overlapping* blocks contained in $C \cdot b^{1/3}$ *nonoverlapping* blocks and vice versa. Hence, for a block-bootstrap estimator based on *overlapping* blocks of length ℓ, this heuristic argument suggests that one should delete $C \cdot b^{1/3}\ell$ overlapping blocks for its JAB variance estimator.

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

TABLE 2. Means, standard errors (SE's) and the mean squared errors (MSE's) of the JAB variance estimators $\widehat{\text{VAR}}_{JAB}(\hat{\phi}_{2n}(\ell))$ for different values of the jackknife parameter $m \equiv Cn^{1/3}\ell^{2/3}$

			C			
Model	ℓ		0.1	0.5	1.0	2.0
I	5	bias	2.472e−05	1.478e−05	0.786e−05	0.356e−05
		SE	1.472e−05	1.007e−05	7.382e−06	6.120e−06
		MSE	8.282e−10	3.202e−10	1.164e−10	5.014e−11
	10	bias	0.932e−05	0.760e−05	0.477e−05	0.205e−05
		SE	1.390e−05	1.439e−05	1.446e−05	1.472e−05
		MSE	2.802e−10	2.651e−10	2.319e−10	2.211e−10
II	5	bias	1.434e−06	8.104e−07	4.227e−07	1.889e−07
		SE	6.119e−07	3.836e−07	2.441e−07	1.698e−07
		MSE	2.432e−12	8.038e−13	2.382e−13	6.453e−14
	10	bias	3.233e−07	2.064e−07	1.138e−07	0.492e−07
		SE	2.408e−07	2.078e−07	1.836e−07	1.761e−07
		MSE	1.625e−13	8.579e−14	4.669e−14	3.345e−14
III	5	bias	6.112e−06	3.431e−06	1.814e−06	0.812e−06
		SE	2.870e−06	1.854e−06	1.288e−06	9.617e−07
		MSE	4.559e−11	1.520e−11	4.949e−12	1.584e−12
	10	bias	1.339e−06	0.886e−06	0.459e−06	0.171e−06
		SE	1.017e−06	9.110e−07	7.835e−07	7.168e−07
		MSE	2.830e−12	1.580e−12	8.250e−13	5.430e−13
IV	5	bias	4.95e−05	3.99e−05	1.923e−05	0.626e−05
		SE	1.276e−04	1.494e−04	1.332e−04	1.452e−04
		MSE	1.873e−08	2.393e−08	1.810e−08	2.113e−08
	10	bias	−1.188e−05	1.202e−05	0.546e−05	0.367e−05
		SE	1.013e−04	2.236e−04	1.893e−04	2.358e−04
		MSE	1.041e−08	5.013e−08	3.587e−08	5.562e−08

Note: The relevant target parameters are the variances of the MBB estimators $\phi_{2n}(\ell)$, given in column 4 of Table 1. The results are based on 3,000 simulation runs for each combination of ℓ and m under each model.

Because $C \cdot b^{1/3}$ is asymptotically equivalent to $C \cdot n^{1/3}\ell^{2/3}$, we choose the jackknife block size parameter m using the formula $m = Cn^{1/3}\ell^{2/3}$.

Table 2 gives the bias, the standard error (SE), and the MSE of the JAB estimators of $\text{Var}(\hat{\phi}_{2n}(\ell))$ for different choices of C and for both block sizes $\ell = 5,10$ under models I–IV. The results are based on 3,000 simulation runs in each case. For each simulation run, the number of bootstrap replicates was determined by the requirement that there were at least 300 bootstrap replications entering into the computation of every jackknife point value $\hat{\phi}_{2n}^{(i)}(\ell)$, $i = 1,\ldots,M$. We make the following observations from the results in Table 2.

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

(1) For models I–III, the MSE's of the JAB variance estimators decrease as the value of the jackknife parameter m (or C) is increased. The same pattern holds true for both the block sizes $\ell = 5$ and 10 under each of these models.
(2) For model IV (with asymmetric tails), the MSE has sinusoidal behavior at both block lengths, with the minimum MSE occurring at $C = 1$ for $\ell = 5$ and at $C = 0.1$ for $\ell = 10$.
(3) The time required for completing the simulation increases as the value of C increases. This was expected because the larger the value of C, the more bootstrap replicates were necessary to generate a given *minimum* number of replicates for computing $\hat{\phi}_{2n}^{(i)}$, $1 \le i \le M$. For both block sizes and for each of the four models, the simulation time for the case $C = 2$ is considerably (more than 1,000 times) longer than the rest of the C values considered.
(4) Writing $\hat{\sigma}_{C,\ell}^2$ for the JAB variance estimator of $\hat{\phi}_{2n}(\ell)$ under a given value C, from Table 2, we see that for all models and sample sizes, the MSE of $\hat{\sigma}_{1,\ell}^2$ (with $C = 1$) is *at most* 3.5 times the *minimum* MSE (over the set of C values considered) in each case. Furthermore, the MSE is within a much smaller range (namely, within only 1.5 times) of the optimal value under the symmetric models I–III, for moderately large value $\ell = 10$ of the block size. Therefore, taking into account the *computation time* (cf (3)), it appears that the value of $C = 1$ (or, equivalently, $m = n^{1/3}\ell^{2/3}$) yields a reasonable choice of the jackknife parameters for the JAB variance estimator in *all* cases considered here.

5. PROOFS

We begin with some notation to be used in the rest of the paper. Let 1_A denote the indicator function of a set A and for any random vector W and $\nu \in \mathbf{Z}_+$, let $\chi_\nu(W)$ denote the νth cumulant of W. For $\alpha \in \mathbf{Z}_+^d$, set $\alpha! = \Pi_{j=1}^d \alpha_j$ and $c_\alpha = D^\alpha H(\mu)/\alpha!$. Let E_* and P_*, respectively, denote the conditional expectation and probability, given the data $\mathcal{X}_n$. For $j \ge 1$, define

$$U_j = (X_j + \cdots + X_{j+\ell-1})/\ell, \quad U_{1j} = U_j \ell^{1/2},$$

$$U_j^* = (X_j^* + \cdots + X_{j+\ell-1}^*)/\ell, \quad U_{1j}^* = U_j^* \ell^{1/2}.$$

For $i = 1, \ldots, M$, denote the resampled observations in the blocks $\{\mathcal{B}_j^{*i} : j = 1, \ldots, b\}$ by $X_1^{*i}, \ldots, X_{n_1}^{*i}$. Next define

$$U_j^{*i} = (X_j^{*i} + \cdots + X_{j+\ell-1}^{*i})/\ell, \quad U_{1j}^{*i} = U_j^{*i} \ell^{1/2}.$$

For notational simplicity, set $I_0 = \{1, \ldots, N\}$, $U_j^* = U_j^{*0}, U_{1j}^* = U_{1j}^{*0}$, and $\hat{\phi}_{jn} = \hat{\phi}_{jn}^{(0)}$. Also, in the proofs that follow, unless there is any chance of confusion, we set $\mu = 0$. Let $C, C(\cdot)$ denote generic constants, depending on their arguments (if any) but not on i, n, ℓ, and m. For notational simplicity, any such dependence on population quantities such as $\|\Sigma_\infty^{-1}\|$ or $E\|X_1\|^s$, which remain fixed throughout the proof of a theorem, will be suppressed. Also, unless otherwise specified, limits in the order symbols are taken letting $n \to \infty$.

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

Proof of Proposition 2.1. The proof follows from the fact that for any $j_1,\ldots,j_b \in I_i$, $P(J_1 = j_1,\ldots,J_b = j_b|\hat{p}_i = 1) = P(J_{i1} = j_1,\ldots,J_{ib} = j_b)$. ■

Proof of Theorem 3.1. First we consider ϕ_{1n}. Recall that in the steps that follow, for notational simplicity, we set $\mu = 0$. Then by Taylor's expansion,

$$\hat{\phi}_{1n}^{(i)} = E_*H(\bar{X}_n^{*i}) - H(\hat{\mu}_{n,i}) = b^{-1}\sum_{|\alpha|=2} c_\alpha E_*(U_1^{*i})^\alpha + \hat{R}_{11n}^{(i)}, \tag{5.1}$$

where the remainder term satisfies

$$\begin{aligned}
&\left|\hat{R}_{11n}^{(i)} - b^{-1}\left[-\sum_{|\alpha|=2} c_\alpha(\hat{\mu}_{n,i})^\alpha + \sum_{|\alpha|=2}\sum_{|\beta|=1} c_{\alpha,\beta}(\hat{\mu}_{n,i})^\beta E_*(U_1^{*i} - \hat{\mu}_{n,i})^\alpha\right]\right| \\
&\quad \le C\{b^{-2}(1 + \|\hat{\mu}_{n,i}\|^a)E_*\|U_1^{*i}\|^3 \\
&\qquad + (\|\hat{\mu}_{n,i}\|^4 + \|\hat{\mu}_{n,i}\|^{4+a}) + (E_*\|\bar{X}_n^{*i}\|^4 + E_*\|\bar{X}_n^{*i}\|^{4+a})\} \\
&\quad \equiv \hat{R}_{12n}^{(i)}, \quad \text{(say)}.
\end{aligned} \tag{5.2}$$

It is important to note that the constant C in (5.2) does not depend on i and n. Next, using the moment and mixing conditions given earlier, the cumulant expansion for moments, and arguments similar to those of Bulinskii and Zhurbenko (1976), one can show that

$$\begin{aligned}
&E\|\hat{\mu}_n\|^{8+2a} = O(n^{-(4+a)}), \\
&E\|\bar{X}_n\|^{8+2a} = O(n^{-(4+a)}), \quad \text{and} \\
&E(E_*\|U_1^*\|^{8+2a}) = E\|U_1\|^{8+2a} = O(\ell^{-(4+a)}).
\end{aligned} \tag{5.3}$$

Hence, writing $Y_i = \sum_{|\alpha|=2} c_\alpha U_i^\alpha, i = 1,\ldots,N$, and using Hölder's inequality, we get

$$\begin{aligned}
&\mathrm{Var}\left(b^{-1}\sum_{|\alpha|=2} c_\alpha E_*(U_1^*)^\alpha\right) \\
&\quad = b^{-2}N^{-1}\sum_{j=-\ell}^{\ell}\mathrm{Cov}(Y_1, Y_{1+j}) + O(b^{-2}n^{-2}\ell^2 EY_1^2) \\
&\qquad + O\left(b^{-2}n^{-1}\sum_{|j|\ge\ell}\alpha(j-\ell)^{\delta/(2+\delta)}\cdot(E|Y_1|^{2+\delta})^{2/(2+\delta)}\right) \\
&\quad = O(b^{-2}n^{-1}\ell^{-1})
\end{aligned}$$

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

and

$$
\begin{aligned}
E(\hat{R}_{11n}^{(0)})^2 \le\ & C[b^{-2}\{E\|\hat{\mu}_n\|^4 + E(\|\hat{\mu}_n\|^2(E_*\|U_1^*\|^4))\} \\
& + b^{-4}E\{E_*\|U_1^*\|^6(1+\|\hat{\mu}_n\|^{2a})\} \\
& + E\{\|\hat{\mu}_n\|^8(1+\|\hat{\mu}_n\|^{2a})\} \\
& + E\{E_*\|\overline{X}_n^*\|^8(1+\|\overline{X}_n^*\|^{2a})\}] \\
=\ & O(n^{-3}). \qquad \textbf{(5.4)}
\end{aligned}
$$

This, in particular, implies that

$$
\begin{aligned}
\mathrm{Var}(\hat{\phi}_{1n}) &= b^{-2}N^{-1}\left(\sum_{j=-\ell}^{\ell}\mathrm{Cov}(Y_1, Y_{1+j})\right)(1+o(1)) \\
&\equiv \sigma_{1n}^2(1+o(1)), \quad \text{say,} \qquad \textbf{(5.5)}
\end{aligned}
$$

where the leading term σ_{1n}^2 is of the order $O(b^{-2}n^{-1}\ell^{-1})$. Next note that by (5.1), for $i = 1, \ldots, M$,

$$
\begin{aligned}
\tilde{\phi}_{1n}^{(i)} &= m^{-1}(N\hat{\phi}_{1n} - (N-m)\hat{\phi}_{1n}^{(i)}) \\
&= (mb)^{-1}\sum_{j=i}^{i+m-1} Y_j + \hat{R}_{13n}^{(i)}, \qquad \textbf{(5.6)}
\end{aligned}
$$

where the remainder terms $\hat{R}_{13n}^{(i)}$'s admit the bound (cf. (5.2))

$$
\begin{aligned}
&\sum_{i=1}^{M}|\hat{R}_{13n}^{(i)}|^2 \\
&\quad \le (mb)^{-2}\sum_{i=1}^{M}\left|\sum_{|\alpha|=2} c_\alpha\{N(\hat{\mu}_n)^\alpha - (N-m)(\hat{\mu}_{n,i})^\alpha\}\right|^2 \\
&\qquad + (mb)^{-2}\sum_{i=1}^{M}\left|\sum_{|\alpha|=2}\sum_{|\beta|=1} c_{\alpha,\beta}\{N(\hat{\mu}_n)^\beta E_*(U_1^* - \hat{\mu}_n)^\alpha \right. \\
&\qquad\qquad \left. - (N-m)(\hat{\mu}_{n,i})^\beta E_*(U_1^{*i} - \hat{\mu}_{n,i})^\alpha\}\right| \\
&\qquad + N^2Mm^{-2}|\hat{R}_{12n}^{(0)}|^2 + (N-m)^2m^{-2}\sum_{i=1}^{M}|\hat{R}_{12n}^{(i)}|^2 \\
&\quad \le C(mb)^{-2}\sum_{i=1}^{M}\left[m^2\|\hat{\mu}_n\|^4 + (N-m)^{-2}\right. \\
&\qquad\qquad \left. \times\left\{\|\sum_{j=1}^{N}U_j\|^2\|\sum_{j=i}^{i+m-1}U_j\|^2 + \|\sum_{j=i}^{i+m-1}U_j\|^4\right\}\right] \\
&\qquad + C(mb)^{-2}\sum_{i=1}^{M}\left[\|\hat{\mu}_n - \hat{\mu}_{n,i}\|^2\left(\sum_{j=1}^{N}\|U_j\|^2\right)^2 + \|\hat{\mu}_{n,i}\|^2\left(\sum_{j=i}^{i+m-1}\|U_j\|^2\right)^2\right. \\
&\qquad\qquad \left. + N^2\|\hat{\mu}_n\|^6 + (N-m)^2\|\hat{\mu}_{n,i}\|^6\right] \\
&\qquad + N^2Mm^{-2}|\hat{R}_{12n}(0)|^2 + N^2m^{-2}\sum_{i=1}^{M}|R_{12n}^{(i)}|^2. \qquad \textbf{(5.7)}
\end{aligned}
$$

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

Therefore, by (5.5) and (5.6), it follows that

$$
\begin{aligned}
&|\widehat{\mathrm{Var}}_{JAB}(\hat{\phi}_{1n}) - \mathrm{Var}(\hat{\phi}_{1n})|/\mathrm{Var}(\hat{\phi}_{1n}) \\
&\quad \le C(b^2 n\ell)\Bigg|(m/N)M^{-1}\sum_{i=1}^{M}\Bigg[(bm)^{-1}\sum_{j=i}^{i+m-1}(Y_j - EY_j) + \hat{R}_{13n}^{(i)} \\
&\qquad\qquad - (\hat{\phi}_{1n} - b^{-1}EY_1)\Bigg]^2 - \sigma_{1n}^2\Bigg| + o(1) \\
&\quad \le C(n^3\ell^{-1})\Bigg[(m/N)|M^{-1}\sum_{i=1}^{M}(bm)^{-2}(Y_{mi}^2 - EY_{mi}^2)| \\
&\qquad\qquad + |(b^2 mN)^{-1}EY_{m1}^2 - \sigma_{1n}^2| \\
&\qquad\qquad + (m/N)M^{-1}\sum_{i=1}^{M}(\hat{R}_{13n}^{(i)})^2 + (m/N)(\hat{\phi}_{1n} - b^{-1}EY_1)^2\Bigg] \\
&\qquad + o(1) \\
&\quad = I_{11} + I_{12} + I_{13} + I_{14} + o(1), \quad \text{say,}
\end{aligned}
\tag{5.8}
$$

where $Y_{mi} = \sum_{j=i}^{i+m-1}(Y_j - EY_j)$.

Next we show that each of the terms $I_{11}, \ldots, I_{14}$ tends to zero in probability. To that end, define the partial sums $S_i = \sum_{j=(i-1)(\ell+m)+1}^{i(\ell+m)\wedge M}(Y_{mj}^2 - EY_{mj}^2)$, $i \ge 1$, and let $\sum_1$ and $\sum_2$, respectively, denote summation over all even and odd i's satisfying $(i-1)(\ell+m) < M$. Then, using the weak dependence of *alternate* S_i's, we have

$$
\begin{aligned}
EI_{11}^2 &\le C(\ell/m)^2 M^{-2}\left[E\left(\sum_1 S_i\right)^2 + E\left(\sum_2 S_i\right)^2\right] \\
&\le C(\ell/m)^2 M^{-2}(M/(\ell+m))\cdot\left(\sum_{j=1}^{\infty}\alpha(j)^{\delta/(8+\delta)}\right) \\
&\quad \cdot \max\left\{E|\sum_{j=1}^{k}Y_{mj}^2|^{2+\delta/4} : 1 \le k \le (\ell+m)\right\}^{8/(8+\delta)}.
\end{aligned}
\tag{5.9}
$$

Next, define $\widetilde{W}_{1j} = \sum_{|\alpha|=2} c_\alpha U_j^\alpha \cdot 1(\|U_j\| \le a_n)$ and $\widetilde{W}_{2j} = \sum_{|\alpha|=2} c_\alpha U_j^\alpha \cdot 1(\|U_j\| > a_n)$, $j = 1, \ldots, N$ where $a_n = \ell^{-1/2}[m\ell]^{1/a}$ with a as in Assumption (A.1). Let $W_{1j} = \widetilde{W}_{1j} - E\widetilde{W}_{1j}$ and $W_{2j} = \widetilde{W}_{2j} - E\widetilde{W}_{2j}$, $j = 1, \ldots, N$. Then, using arguments similar to those in the previous step, and setting $\delta_1 = \delta/4$ and $\delta_2 = 2\delta_1$, we have

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

$$\max\left\{E|\sum_{j=1}^{k} Y_{mj}^2|^{2+\delta_1} : 1 \le k \le (\ell+m)\right\}^{8/(8+\delta)}$$
$$\le C((\ell+m)^{2+\delta_1} E|Y_{m1}|^{4+2\delta_1})^{8/(8+\delta)}$$
$$\le C\left[m^{2+\delta_1}\left\{E\left|\sum_{j=1}^{m} W_{1j}\right|^{(8+\delta)/2} + E\left|\sum_{j=1}^{m} W_{2j}\right|^{(8+\delta)/2}\right\}\right]^{8/(8+\delta)}$$
$$\le Cm^2\left[(ma_n^2)^{\delta/2} E\left(\sum_{j=1}^{m} W_{1j}\right)^4 + m^{(8+\delta)/2} E|W_{21}|^{(8+\delta)/2}\right]^{8/(8+\delta)}$$
$$\le Cm^2[(ma_n^2)^{\delta/2}\cdot\Delta(2;\delta_2)\cdot(E|W_{11}|^{4+\delta_2})^{4/(4+\delta_2)}(m/\ell)^2$$
$$+ m^{(8+\delta)/2} E\|U_1\|^{8+\delta} 1(\|U_1\| > a_n)]^{8/(8+\delta)}$$
$$\le Cm^2[(ma_n^2)^{\delta/2}(E\|U_1\|^{8+2a})^{8/(8+a)}(m/\ell)^2$$
$$+ m^{(8+\delta)/2} a_n^{-(2a-\delta)}\ell^{-(4+a)}\cdot E\|U_{11}\|^{8+2a} 1(\|U_1\| > a_n)]^{8/(8+\delta)}$$
$$\le Cm^2[(ma_n^2)^{\delta/2}(m/\ell)^2\ell^{-4} + m^{(8+\delta)/2}\ell^{-(4+a)}a_n^{-(2a-\delta)}]^{8/(8+\delta)}. \qquad \textbf{(5.10)}$$

Hence, from (5.9) and (5.10), we get

$$EI_{11}^2 \le Cn^{-1}\ell^{-2}m^3[(\ell m)^{-(2a-\delta)/a}]^{8/(8+\delta)}. \qquad \textbf{(5.11)}$$

Next, using the fact that $\ell = o(m)$, it readily follows that $I_{12} = o(1)$. Also, using (5.3), (5.4), and (5.7), we have

$$EI_{13} \le C(n^2m/\ell)M^{-1}\sum_{i=1}^{M} E(\hat{R}_{13n}^{(i)})^2$$
$$\le C(n^2m/\ell)\left[(mb)^{-2}\left\{m^2n^{-2} + (E\|\hat{\mu}_n\|^4)^{1/2}\left(E\|\sum_{j=1}^{m} U_j\|^4\right)^{1/2}\right.\right.$$
$$\left. + n^{-2}E\|\sum_{j=1}^{m} U_j\|^4\right\}$$
$$+ (mb)^{-2}\{N^2(E\|\hat{\mu}_n - \hat{\mu}_{n,1}\|^6)^{1/3}(E\|U_1\|^6)^{2/3}$$
$$+ m^2(E\|\hat{\mu}_{n,1}\|^6)^{1/3}(E\|U_1\|^6)^{2/3} + n^2\cdot n^{-3}\}$$
$$\left. + m^{-2}n^2\{E(\hat{R}_{12n}^{(0)})^2 + E(\hat{R}_{12n}^{(1)})^2\}\right]$$
$$\le C(n^2m/\ell)[(mb)^{-2}\{m^2n^{-2}\} + (mb)^{-2}$$
$$\times\{n^2(m/n^2)\cdot\ell^{-2} + m^2n^{-1}\ell^{-2} + n^{-1}\}$$
$$+ m^{-2}n^2\{b^{-4}\ell^{-3}(1+n^{-a}) + n^{-4}(1+n^{-a})\}]$$
$$\le C(n^2m/\ell)[n^{-2}b^{-2} + m^{-1}b^{-2}\ell^{-2} + m^{-2}n^2n^{-3}b^{-1}]$$
$$\le C[\ell^{-1} + m^{-1}],$$

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

where we have used the facts $\|\hat{\mu}_n - \hat{\mu}_{n,i}\| \le C[(m/n)\|\hat{\mu}_n\| + n^{-1}\|\sum_{j=i}^{i+m-1} U_j\|]$, $E\|\sum_{j=1}^{m} U_j\|^4 = 0(m^{-2})$, and a moment bound similar to (5.3).

Finally, using (5.1) (with $i = 0$) and (5.4), it is easily checked that

$$I_{4n} = 0_p(m/N).$$

Hence, the theorem is proved for $\hat{\phi}_{1n}$.

To prove the result for $\hat{\phi}_{2n}$, note that a third-order Taylor's expansion of $H(\bar{X}_n^{*i})$ around $\mu(=0)$ yields

$$\hat{\phi}_{2n}^{(i)} = b^{-1}E_*\Big(\sum_{|\alpha|=1} c_\alpha (U_1^{*i})^\alpha\Big)^2 + \hat{R}_{21n}^{(i)}$$

for $i = 0,1,\ldots,M$, where

$$\Big|\hat{R}_{21n}^{(i)} - \Big[-b^{-1}\Big(\sum_{|\alpha|=1} c_\alpha \hat{\mu}_{n,i}^{\alpha}\Big)^2 + \mathrm{Cov}_*\Big(\sum_{|\alpha|=1} c_\alpha(\bar{X}_n^{*i})^\alpha\Big), \quad \sum_{|\beta|=2} c_\beta(\bar{X}_n^{*i})^\beta)\Big]\Big|$$
$$\le C(E_*\|\bar{X}_n^{*i}\|^4 + E_*\|\bar{X}_n^{*i}\|^{6+2a}).$$

Evidently, the leading term in the expansion of $\hat{\phi}_{2n}^{(i)}$ is of the same form (namely, the conditional expectation of a quadratic form in U_1^{*i}) as that of $\hat{\phi}_{1n}^{(i)}$ (cf. (5.1)), and so is the form of the remainder term $\hat{R}_{21n}^{(i)}$ as compared with $\hat{R}_{11n}^{(i)}$. Thus, retracing the proof for $\hat{\phi}_{1n}$, one can easily establish the result for $\hat{\phi}_{2n}$. ■

Proof of Theorem 3.2. Because the jackknife and the bootstrap for i.i.d. data correspond to block jackknife and block bootstrap with block lengths $\ell = m = 1$, we only outline the changes in the proof of Theorem 3.1 that are necessary for proving Theorem 3.2. For $i = 0,1,\ldots,M$ ($M = b \equiv n$ in this case), $\hat{\phi}_{1n}^{(i)}$ has the expansion

$$\hat{\phi}_{1n}^{(i)} = n^{-1}\sum_{|\alpha|=2} c_\alpha E_*(U_1^{*i})^\alpha$$
$$+ n^{-1}\sum_{|\beta|=1}\sum_{|\alpha|=2} c_{\alpha,\beta}(\hat{\mu}_{n,i})^\beta E(U_1^\alpha) + \hat{Q}_{11n}^{(i)},$$

where, using (5.2), we get

$$\Big|\hat{Q}_{11n}^{(i)} - n^{-1}\Big[-\sum_{|\alpha|=2} c_\alpha(\hat{\mu}_{n,i})^\alpha + \sum_{|\beta|=1}\sum_{|\alpha|=2} c_{\alpha,\beta}(\hat{\mu}_{n,i})^\beta (E_*(U_1^{*i})^\alpha - EU_1^\alpha)\Big]\Big|$$
$$\le C\{n^{-1}\|\hat{\mu}_{n,i}\|^3 + \hat{R}_{12n}^{(i)}\}$$
$$\equiv \hat{Q}_{12n}^{(i)}, \quad \text{say.}$$

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

Then, $\mathrm{Var}(\hat{\phi}_{1n}) = \sigma^2 n^{-3}(1 + o(1))$ for some $\sigma^2 > 0$, and by arguments similar to (5.4),

$$E(\hat{Q}_{11n}^{(0)})^2 = 0(n^{-4}).$$

The expansion for the pseudovalues $\hat{\phi}_{1n}^{(i)}$'s given by (5.6) continue to hold, if we set $m = 1$ and redefine

$$Y_j = \sum_{|\alpha|=2} c_\alpha U_j^\alpha + \left(\sum_{|\alpha|=2} c_{\alpha,\beta} E U_1^\alpha \right) U_1^\beta, \quad 1 \le j \le N.$$

Furthermore, the remainder terms in (5.6) satisfy the inequality (5.7) with $\hat{R}_{12n}^{(i)}$'s replaced by $\hat{Q}_{12n}^{(i)}$ and the second summation term replaced by

$$Cn^{-2} \sum_{i=1}^{n} \sum_{|\alpha|=2} \left\{ \|\hat{\mu}_n - \hat{\mu}_{n,i}\|^2 \left| \sum_{j=1}^{n} (U_j^\alpha - E U_j^\alpha) \right|^2 \right.$$
$$\left. + \|\hat{\mu}_{n,i}\|^2 |U_i^\alpha - E U_1^\alpha|^2 + n^2 \|\hat{\mu}_n\|^6 + n^2 \|\hat{\mu}_{n,i}\|^6 \right\}.$$

To complete the proof, it is now enough to check that (analogues of) $I_{11}, \ldots, I_{14}$ in (5.8) (with the preceding modification) are $o_p(1)$. That $I_{11} + I_{13} = o_p(1)$ follows using standard moment bounds for sums of independent random variables to I_{11} and I_{13}. Also, because of independence, $\sigma_{1n}^2 = n^{-3} E Y_1^2$, and hence $I_{12} \equiv 0$. Finally, $I_{14} = 0_p(n^{-1})$, by the same argument as in the dependent case. Hence, the theorem is proved for $\hat{\phi}_{1n}$. ■

The proof for $\hat{\phi}_{2n}$ is similar, and hence it is omitted.

REFERENCES

Beran, R. (1984) Jackknife approximations to bootstrap estimates. *Annals of Statistics* 12, 101–118.

Bühlmann, P. & H.R. Künsch (1995) Blockwise bootstrap for general parameters of a stationary time series. *Scandinavian Journal of Statistics* 22, 35–54.

Bulinskii, A.V. & I.G. Zhurbenko (1976) The central limit theorem for random fields. *Soviet Mathematics Doklady* 17, 14–17.

Carlstein, E. (1986) The use of subseries methods for estimating the variance of a general statistic from a stationary time series. *Annals of Statistics* 14, 1171–1179.

Datta, S. & W.P. McCormick (1995) Bootstrap inference for a first order autoregression with positive innovations. *Journal of the American Statistical Association* 90, 1289–1300.

Efron, B. (1992) Jackknife-after-bootstrap standard errors and influence functions (with discussion). *Journal of the Royal Statistical Society, Series B* 54, 83–111.

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

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

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

Hall, P., J.L. Horowitz, & B.-Y. Jing (1995) On blocking rules for the bootstrap with dependent data. *Biometrika* 82, 561–574.

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press

Künsch, H.R. (1989) The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* 17, 1217–1261.
Lahiri, S.N. (1991) Second order optimality of stationary bootstrap. *Statistics and Probability Letters* 11, 335–341.
Lahiri, S.N. (1996a) 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. (1996b) On Empirical Choice of the Optimal Block Size for Block Bootstrap Methods. Preprint. Department of Statistics, Iowa State University, Ames.
Lahiri, S.N. (1997) On inconsistency of the jackknife-after-bootstrap bias estimator for dependent data. *Journal of Multivariate Analysis* 63, 15–34.
Li, H. & G.S. Maddala (1997) Bootstrapping cointegrating regressions. *Journal of Econometrics* 80, 297–318.
Liu, R.Y. & K. Singh (1992) Moving blocks jackknife and bootstrap capture weak dependence. In R. Lepage & L. Billard (eds.), *Exploring the Limits of Bootstrap*, pp. 225–248. New York: Wiley.
Parr, W.C. (1985) Jackknifing differentiable statistical functionals. *Journal of the Royal Statistical Society, Series B* 47, 56–66.
Parr, W.C. & W.R. Schucany (1982) Jackknifing *L*-statistics with smooth weight functions. *Journal of the American Statistical Association* 80, 620–630.
Politis, D. & J.P. Romano (1992a) A circular block resampling procedure for stationary data. In R. Lepage & L. Billard (eds.), *Exploring the Limits of Bootstrap*, pp. 263–270. New York: Wiley.
Politis, D. & J.P. Romano (1992b) A general resampling scheme for triangular arrays of α-mixing random variables. *Annals of Statistics* 20, 1985–2007.
Reeds, J.A. (1978) Jackknifing maximum likelihood estimates. *Annals of Statistics* 6, 727–739.
Shao, J. (1989) The efficiency and consistency of approximations to the jackknife variance estimators. *Journal of the American Statistical Association* 84, 114–119.
Shao, J. & C.F.J. Wu (1989) A general theory for jackknife variance estimation. *Annals of Statistics* 17, 1176–1197.

https://doi.org/10.1017/S0266466602181059 Published online by Cambridge University Press