RESAMPLING TECHNIQUES FOR STATIONARY TIME-SERIES: SOME RECENT DEVELOPMENTS.

E. CARLSTEIN*

Abstract. A survey is given of resampling techniques for stationary time-series, including algorithms based on the jackknife, the bootstrap, the typical-value principle, and the subseries method. The techniques are classified as "model-based" or "model-free," according to whether or not the user must know the underlying dependence mechanism in the time-series. Some of the techniques discussed are new, and have not yet appeared elsewhere in the literature.

Key words. subsampling, jackknife, bootstrap, typical-values, subseries, nonparametric, dependence, mixing

AMS(MOS) subject classifications. 62G05, 62M10

---

*Department of Statistics, University of North Carolina, Chapel Hill, N.C. 27599.
Research supported by N.S.F. Grant #DMS-8902973, and by the Institute for Mathematics and its Applications with funds provided by the N.S.F.

1. Introduction. Resampling techniques enable us to address the following class of statistical problems: A sample series of n random variables $(Z_1, Z_2, \ldots, Z_n) =: \vec{Z}_n^0$ is observed from the strictly stationary sequence $\{Z_i: -\infty < i < +\infty\}$, and a statistic $t_n(\vec{Z}_n^0) =: t_n^0$ is computed from the observed data. The objective is to describe the sampling distribution of the statistic t_n^0, using only the data $\vec{Z}_n^0$ at hand.

The scenario is nonparametric, i.e., the marginal distribution F of Z_i is unknown to the statistician; also, the statistic defined by the function $t_n(\cdot)$: $\mathbf{R}^n \mapsto \mathbf{R}^1$ may be quite complicated (e.g., an adaptively defined statistic, or a robustified measure of location, dispersion, or correlation in the sample series). Therefore, direct analytic description of t_n^0's sampling distribution may be impossible. These analytic difficulties will be further exacerbated by nontrivial dependence in $\{Z_i\}$. Resampling techniques substitute nonparametric sample-based numerical computations in place of intractable mathematics. Moreover, resampling algorithms are "omnibus," i.e., they are phrased in terms of a general statistic $t_n(\cdot)$ so that each new situation does not require the development of a new procedure. In keeping with the spirit of omnibus nonparametric procedures, it is also desirable for resampling techniques to require only minimal technical assumptions.

Depending on the particular application, there are various different features of t_n^0's sampling distribution that one may wish to describe via resampling. For example, the goal might be to obtain point estimates of t_n^0's moments (e.g., variance, skewness, or bias). Or, the focus might be on estimating the percentiles of t_n^0's sampling distribution. In some cases, the sampling distribution of t_n^0 can be used in constructing confidence intervals on an unknown target parameter θ. As a diagnostic tool, one may want to determine whether the statistic t_n^0 has an approximately normal sampling distribution, and, if not, how it departs from normality. For each of these objectives there are appropriate resampling algorithms.

The fundamental strategy in resampling is to generate "replicates" of the statistic t from the available data $\vec{Z}_n^0$, and then use these replicates to model the true sampling distribution of t_n^0. The choice of a particular resampling algorithm for generating replicates depends upon the intended application (e.g., moment estimation, percentile estimation, confidence intervals, or diagnostics, as discussed above) and upon the structure in the original data $\vec{Z}_n^0$ (e.g., independence versus time-series versus regression).

Resampling algorithms for time-series are intuitively motivated by analogy to the established resampling algorithms for independent observations. Therefore, Section 2 reviews the jackknife, the bootstrap, and the typical-value principle -- three specific resampling algorithms for generating replicates from independent observations. When the original observations are serially dependent, these resampling algorithms must be appropriately modified in order to yield valid replicates of the statistic t. The modified resampling algorithms can be "model-based" (i.e., they can exploit an assumed dependence mechanism in $\{Z_i\}$) or they can be "model-free" (i.e., no knowledge of the dependence mechanism in $\{Z_i\}$ is needed). Section 3 surveys the model-based resampling algorithms for time-series, including the Markovian bootstrap and bootstrapping of residuals; Section 4 surveys the model-free approaches, including the blockwise jackknife, the blockwise bootstrap, the linked blockwise bootstrap, and the subseries method.

The survey given in Sections 2, 3, and 4 is expository, relying mostly on intuitive explanations of the resampling algorithms (for precise technical conditions the reader is directed to the original references). For each resampling algorithm, a natural question to ask is "Does it work?" Specifically, for which statistics $t_n(\cdot)$ and marginal distributions F does the resampling algorithm provide replicates which adequately model the desired feature of t_n^0's sampling distribution? In many cases, the answer to this question will be inextricably tied to the issue of t_n^0's asymptotic normality.

2. Resampling algorithms for independent observations. This review of the jackknife, the bootstrap, and the typical-value principle is meant to provide an intuitive foundation -- in the independent case -- for the time-series resampling techniques which will be discussed in Sections 3 and 4. Therefore, the focus of this Section is on the seminal works in resampling for independent observations. An exhaustive review of these resampling algorithms is not attempted here; indeed, there have been more than 400 publications on the bootstrap alone since its introduction just over a decade ago.

Throughout this Section, assume that the random variables $\{Z_i\colon -\infty<i<+\infty\}$ are independent, and that $t_m(\cdot)$, $m\geq 1$, is symmetric in its m arguments.

2.1 Jackknife. The jackknife algorithm generates replicates of the statistic t by deleting observations from the sample $\vec{Z}_n^0$, and then computing the statistic on the remaining data. Thus, the i^{th} "jackknife replicate" of the statistic t is

$$t_n^{(i)} := t_{n-1}(Z_1, Z_2, \ldots, Z_{i-1}, Z_{i+1}, \ldots, Z_n)$$

for $i \in \{1, 2, \ldots, n\}$.

To estimate $\mathbf{V}\{t_n^0\}$, the variance of t_n^0's sampling distribution, Tukey [65] proposed

$$\hat{\mathbf{V}}_J\{t_n^0\} := \sum_{i=1}^{n} \frac{(t_n^{(i)} - \overline{t}_n)^2}{n} \cdot (n-1) ,$$

where $\overline{t}_n := \sum_{i=1}^{n} t_n^{(i)}/n$. This "jackknife estimate of variance" $\hat{\mathbf{V}}_J\{t_n^0\}$ uses the variability amongst the jackknife replicates to model the true sampling variability of t_n^0. Since the $t_n^{(i)}$s share so many observations, they do not exhibit as much variability as would n independent realizations of t_n^0; the "extra" factor of $(n-1)$ accounts for this effect by inflating the variance estimate.

A similar approach can be used to estimate the bias $(\mathbf{E}\{t_n^0\} - \theta)$ of t_n^0. In fact, Quenouille [46, 47] originated jackknifing for this purpose. His bias estimate, $(\overline{t}_n - t_n^0) \cdot (n-1)$, uses the average of the jackknife replicates to model the true expectation of t_n^0.

In order for $\hat{\mathbf{V}}_J\{t_n^0\}$ to be an asymptotically unbiased and consistent estimator of $\mathbf{V}\{t_n^0\}$, it is necessary that the statistic t_n^0 have an asymptotically normal sampling distribution (van Zwet [66]). Asymptotic normality of the general statistic t_n^0 and its jackknife replicates has also been studied by Hartigan [30]. Note that the jackknife estimate of variance fails when t_n^0 is the sample median, even though t_n^0 is asymptotically normal (see Efron [20] for this example, as well as a thorough analysis of the jackknife). The jackknife method does allow deletion of more than one observation when computing the jackknife replicates; this is explored by Shao and Wu [52].

2.2 Typical-Values. Hartigan [29] introduced the typical-value principle for constructing nonparametric confidence intervals on an unknown parameter θ. A collection of random variables $\{V_1, V_2, \ldots, V_k\}$ are "typical-values for θ" if each of the $k+1$ intervals between the ordered random variables

$$-\infty \equiv V_{(0)} < V_{(1)} < V_{(2)} < \ldots < V_{(k)} < V_{(k+1)} \equiv +\infty$$

contains θ with equal probability. Given such a collection of typical-values, confidence intervals on θ may be constructed by taking the union of adjacent intervals; in particular, $\mathbf{P}\{V_{(1)} < \theta < V_{(k)}\} = 1 - 2/(k+1)$.

The connection between typical-values and resampling is this: Suppose that the statistic t estimates the unknown parameter θ. For any subset S of the data indices, i.e., $S \equiv \{i_1, i_2, \ldots, i_m\} \subseteq \{1, 2, \ldots, n\}$, compute the corresponding "subset replicate" of t:

$$t[S] := t_m(Z_{i_1}, Z_{i_2}, \ldots, Z_{i_m}).$$

In many situations, the collection of subset replicates $\{ t[S_1], t[S_2], \ldots, t[S_k] \}$ are actually typical-values for θ. This is true, for example, whenever F is a continuous symmetric distribution with center of symmetry θ, the statistic t is the sample mean, and the S_js are the $k=2^n-1$ possible non-empty subsets of $\{1, 2, \ldots, n\}$. Notice that t_n^0 need not be asymptotically normal in this case, e.g., if F is Cauchy. Valid typical-values can be obtained from certain smaller collections of subset replicates, where the S_js may be chosen deterministically or randomly (see also Efron [20]). Furthermore, typical-values can be obtained from subset replicates of statistics t other than the sample mean, e.g., the sample median (Efron [20]), M-estimates (Maritz [42]), and general asymptotically normal statistics (Hartigan [30]).

2.3 Bootstrap. Efron's [19] bootstrap algorithm uses F_n, the observed empirical distribution of $\vec{Z}_n^0$, in place of the unknown distribution F. Given F_n, which assigns mass $1/n$ at each Z_i $(1 \le i \le n)$, a "bootstrap sample" $(Z_1^*, Z_2^*, \ldots, Z_m^*) =: \vec{Z}_m^{0*}$ is generated by i.i.d. sampling from F_n. The corresponding "bootstrap replicate" of t is

$$t_m^{0*} := t_m(\vec{Z}_m^{0*}).$$

For fixed data $\vec{Z}_n^0$, it is possible to generate arbitrarily many bootstrap replicates of t, by repeatedly drawing bootstrap samples from F_n. The conditional distribution (i.e., given $\vec{Z}_n^0$) of these bootstrap replicates is used to model the true sampling distribution of t_n^0. For example, to estimate $\mathsf{V}\{t_n^0\}$, the "bootstrap estimate of variance" is

$$\hat{\mathsf{V}}_B\{t_n^0\} := \mathsf{E}\Big\{\Big(t_m^{0*} - \mathsf{E}\{t_m^{0*} | \vec{Z}_n^0\}\Big)^2 \Bigm| \vec{Z}_n^0\Big\}.$$

In many situations, the bootstrap replicates can actually be used to estimate the entire distribution function of t_n^0. To obtain a valid estimate, appropriate choices must be made for the bootstrap sample size $m \equiv m_n$ and for the standardizations $(a_n, b_n; a_m^*, b_m^*)$ of t_n^0 and t_m^{0*}. Specifically, the bootstrap estimate of $\mathsf{P}\{\tilde{t}_n^0 \le x\}$ is

$$\mathsf{P}\{\tilde{t}_m^{0*} \le x \mid \vec{Z}_n^0\},$$

where $\tilde{t}_n^0 := a_n(t_n^0 - b_n)$ and $\tilde{t}_m^{0*} := a_m^*(t_m^{0*} - b_m^*)$ are the standardized versions. This bootstrap estimate is said to be "strongly uniformly consistent" if

$$\sup_{x \in R} \Big| \mathsf{P}\{\tilde{t}_m^{0*} \le x \mid \vec{Z}_n^0\} - \mathsf{P}\{\tilde{t}_n^0 \le x\} \Big| \xrightarrow{a.s.} 0 \quad \text{as } n \to \infty.$$

When the statistic t is the sample mean, strong uniform consistency of the bootstrap has been established in several scenarios: Bickel and Freedman [6] deal with the asymptotically normal case, as does Singh [53] who actually shows that the bootstrap estimate can be second-order correct (i.e., strong uniform consistency holds with a rate factor of $\sqrt{n}$); Arcones and Giné [1] treat the asymptotically α-stable case ($0<\alpha<2$). The bootstrap is strongly uniformly consistent when t is a von Mises functional statistic, both in the asymptotically normal case (Bickel and Freedman [6]) and in the asymptotically non-normal case (Bretagnolle [10]). Bickel and Freedman [6] also show that the bootstrap is strongly uniformly consistent when t is the [asymptotically normal] sample median. Swanepoel [58] establishes strong uniform consistency for the bootstrap when t is the sample maximum [having an exponential limiting distribution]. Thus the bootstrap algorithm is valid in a wide variety of i.i.d. situations.

3. Model-based resampling algorithms for time-series. If the dependence mechanism in $\{Z_i\}$ is known, then that mechanism can be incorporated into the resampling algorithm and hence into the generated replicates of t. This Section discusses in detail two such model-based modifications of the bootstrap algorithm. [Note that model-based modifications of the bootstrap algorithm have also been studied in the case where $\{Z_i\}$ is a non-stationary AR(1) time-series and t is a least-squares estimator of the AR parameter (see Basawa, Mallik, McCormick, and Taylor [4] and Basawa, Mallik, McCormick, Reeves, and Taylor [5]).]

3.1 Bootstrapping residuals. Here it is assumed that the dependence structure in $\{Z_i\}$ satisfies

$$Z_i = g(Z_{i-1}, Z_{i-2}, \ldots, Z_{i-k}; \vec{\beta}) + \varepsilon_i ,$$

where $g(\cdot)$ is a known function, $\vec{\beta}$ is a vector of unknown parameters, and the additive unobservable errors $\{\varepsilon_i: -\infty<i<+\infty\}$ are i.i.d. with unknown distribution F_ε having mean zero.

The algorithm begins by calculating $\hat{\beta}$, a data-based estimate of $\vec{\beta}$. Using this $\hat{\beta}$, residuals

$$\hat{\varepsilon}_i := Z_i - g(Z_{i-1}, Z_{i-2}, \ldots, Z_{i-k}; \hat{\beta}) , \quad i\in\{k+1, k+2, \ldots, n\},$$

are computed. The residuals in turn are used to construct an estimate of F_ε, namely $\hat{F}_\varepsilon$ which puts mass $1/(n-k)$ on each centered residual $\hat{\varepsilon}_i - \sum_{j=k+1}^{n} \hat{\varepsilon}_j/(n-k)$, $i\in\{k+1, k+2, \ldots, n\}$. The bootstrap sample $\vec{Z}_m^{0*}$ is now generated using the estimated dependence mechanism, that is

$$Z_i^* := g(Z_{i-1}^*, Z_{i-2}^*, \ldots, Z_{i-k}^*; \hat{\beta}) + \varepsilon_i^*, \quad i\in\{1, 2, \ldots, m\},$$

where $\{\varepsilon_i^*: 1\leq i\leq m\}$ are i.i.d. from $\hat{F}_\varepsilon$ and $Z_i^* \equiv Z_{i+k}$ for $i\in\{-(k-1), -(k-2), \ldots, 0\}$. The corresponding bootstrap replicate of t is simply

$$t_m^{0*} := t_m(\vec{Z}_m^{0*}).$$

The literature has concentrated on the special case where $\{Z_i\}$ is autoregressive, i.e., $g(z_1, z_2, \ldots, z_k; (\beta_0, \beta_1, \beta_2, \ldots, \beta_k)) = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_k z_k$, and where t is a least-squares estimator of the autoregressive parameter having a limiting normal distribution. In this case, the bootstrap is strongly uniformly consistent (Freedman [25]) and furthermore can be second-order correct (Bose [7]); see Efron and Tibshirani [21] and Freedman and Peters [26] for further discussion and applications. Swanepoel and van Wyk [59] apply the bootstrap of autoregressive residuals to constructing a confidence interval on the power spectrum.

The related problem of obtaining prediction intervals for autoregressive time-series has been addressed via bootstrap algorithms by Stine [56] and by Thombs and Schucany [62, 63]; see also Findley [22]. Extensions to the case of ARMA time-series are considered by Thombs [60, 61].

3.2 Markovian bootstrap. Suppose now that $\{Z_i\}$ has first-order Markovian dependence structure with unknown transition density $f(z_1|z_0)$. The algorithm begins by computing $\hat{f}(\cdot|\cdot)$, a sample-based estimate of $f(\cdot|\cdot)$; specifically, this $\hat{f}(\cdot|\cdot)$ may be computed using kernel density estimation techniques. The bootstrap sample $\vec{Z}_m^{0*}$ is generated from the estimated dependence mechanism as follows: Select Z_1^* at random from F_n, and draw Z_i^* from the distribution with density $\hat{f}(\cdot|Z_{i-1}^*)$, $i\in\{2, 3, \ldots, m\}$. The corresponding bootstrap replicate of t is again $t_m^{0*} := t_m(\vec{Z}_m^{0*})$.

Rajarshi [48] introduced this bootstrap algorithm and established strong uniform consistency when t is the [asymptotically normal] sample mean. An extension of this algorithm to Markovian random fields is studied by Lele [38]. The case of Markov chains is considered by Kulperger and Prakasa Rao [34], Athreya and Fuh [2], and Basawa, Green, McCormick, and Taylor [3].

3.3 Other algorithms. Bose [9] studies the bootstrap for moving average models; related work can be found in Bose [8]. A bootstrap method for state space models is investigated by Stoffer and Wall [57]. Franke, Härdle, and Kreiss [24] apply the bootstrap to M-estimates of ARMA parameters. Solow [54] proposes a bootstrap algorithm which first estimates the pairwise correlations in the data and then transforms to approximate independence.

Several approaches to jackknifing time-series data have been suggested, beginning with Brillinger [11]. He suggests that, in computing the $i^{\underline{th}}$ jackknife replicate, the deleted observation Z_i should be treated as a "missing value" in the time-series and should be replaced via interpolation using, say, an ARMA model (Brillinger [12]); thus the complete time-series structure is retained by each replicate. When the statistic t is itself expressible as a function of ARMA residuals (e.g., $\sum_i (\hat{\varepsilon}_i)^2/n$), then Davis [18] considers applying the i.i.d. jackknife algorithm directly by deleting residuals. Gray, Watkins, and Adams [28] give an extension of the jackknife to piecewise continuous stochastic processes.

The present survey emphasizes the intuition of resampling algorithms in the time-domain. There are, however, several frequency-domain resampling algorithms: A bootstrap method designed for Gaussian time-series is studied by Stine [55], Ramos [50], and Hurvich and Zeger [32]; Hartigan [31] introduces a "perturbed periodogram" method; Hurvich, Simonoff, and Zeger [33] propose a bootstrap method for moving average data as well as a jackknife method; also see Franke and Härdle [23] and Thomson and Chave [64].

4. **Model-free resampling algorithms for time-series.** The drawback of the model-based techniques discussed in Section 3 is that they require the user to know the correct underlying dependence mechanism in $\{Z_i\}$ (this point is emphasized by Freedman [25] in the case of autoregression). The strength of the i.i.d. bootstrap, jackknife, and typical-value principles is that they are nonparametric (i.e., the marginal distribution of Z_i is unknown); therefore it is more appropriate to develop analogous resampling techniques for stationary data that are similarly free of assumptions regarding the dependence mechanism (i.e., the joint distribution of the Z_is is allowed to be unknown). In practice, the joint probability structure of $\{Z_i\}$ is more obscure than the marginal probability structure of Z_i, so it is unrealistic to assume that the former is known when the latter is unknown. These considerations motivate the model-free

algorithms discussed in this Section.

In this model-free scenario, it is convenient to measure the strength of dependence in $\{Z_i\}$ by a "mixing coefficient"

$$\alpha(r) := \sup_{A\in\mathcal{F}\{\ldots,\, Z_{-1},\, Z_0\},\ B\in\mathcal{F}\{Z_r,\, Z_{r+1},\, \ldots\}} |\mathbf{P}\{A\cap B\} - \mathbf{P}\{A\}\mathbf{P}\{B\}|,$$

as introduced by Rosenblatt [51]. In order for the model-free resampling algorithms to be valid, it is typically assumed that $\alpha(r)\to 0$ at some appropriate rate as $r\to\infty$. Intuitively, this says that observations which are separated by a long time-lag behave approximately as if they were independent.

In the absence of assumptions about the dependence mechanism in $\{Z_i\}$, it is natural to focus attention on the "blocks" of sample data

$$\vec{Z}^{\,i}_{l} := (Z_{i+1}, Z_{i+2}, \ldots, Z_{i+l}), \quad 0\le i < i+\ell \le n.$$

These blocks automatically retain the correct dependence structure of $\{Z_i\}$. For asymptotic validity, it is usually required that $\ell\to\infty$ as $n\to\infty$, so that the blocks ultimately reflect the dependencies at all lags.

4.1 Blockwise jackknife. In its simplest form, the blockwise jackknife generates replicates of the statistic t by deleting blocks of ℓ observations from $\vec{Z}^{\,0}_{n}$, and then computing the statistic on the remaining data. Thus, the $i^{\underline{th}}$ "blockwise jackknife replicate" of t is

$$t_n^{<i>} := t_{n-l}(\vec{Z}^{\,0}_{n} \setminus \vec{Z}^{\,i}_{l})$$

for $i\in\{0, 1, \ldots, n-\ell\}$. The resulting estimate of $\mathbf{V}\{t_n^0\}$ is

$$\hat{\mathbf{V}}_{BJ}\{t_n^0\} := \sum_{i=0}^{n-l} \frac{(t_n^{<i>} - \overline{t}_n^{<\cdot>})^2}{n-\ell+1} \cdot c_{n,l}\ ,$$

where $\overline{t}_n^{<\cdot>} := \sum_{i=0}^{n-l} t_n^{<i>}/(n-\ell+1)$ and $c_{n,l}$ is an appropriate standardizing constant. This "blockwise jackknife estimate of variance" was proposed by Künsch [35]; he showed that $\hat{\mathbf{V}}_{BJ}\{t_n^0\}$ is consistent when t belongs to a certain class of asymptotically normal functional statistics (including the sample mean). A generalization of the blockwise jackknife is investigated by Politis and Romano [43].

4.2 Blockwise bootstrap. This method extends the bootstrap to dependent data by resampling the blocks. The algorithm is essentially as follows: For fixed ℓ, construct the "empirical ℓ-dimensional marginal distribution" $F_{l,n}$, i.e., the distribution putting mass $1/(n-\ell+1)$ on each sample block $\vec{Z}_l^i$, $i\in\{0, 1, \ldots, n-\ell\}$. Now, assuming $k:=n/\ell$ is an integer, generate k "bootstrap blocks" by i.i.d. random resampling of blocks from $F_{l,n}$. Denote these bootstrap blocks as

$$\vec{Z}_{l,j}^* \equiv (Z_{(j-1)l+1}^*, Z_{(j-1)l+2}^*, \ldots, Z_{jl}^*), \; j\in\{1, 2, \ldots, k\}.$$

The blockwise bootstrap sample $\vec{Z}_n^{0*}$ is then constructed by appending these blocks together, i.e.,

$$\vec{Z}_n^{0*} := (\vec{Z}_{l,1}^*, \vec{Z}_{l,2}^*, \ldots, \vec{Z}_{l,k}^*).$$

Thus $\vec{Z}_n^{0*}$ inherits the correct dependence structure -- at least within blocks. The corresponding "blockwise bootstrap replicate" of t is $t_n^{0*} := t_n(\vec{Z}_n^{0*})$.

This algorithm was introduced by Künsch [35] (see also Liu and Singh [41]). He shows that, when t is an asymptotically normal sample mean, the blockwise bootstrap is strongly uniformly consistent; second-order correctness is studied by Götze and Künsch [27] and by Lahiri [37]. A generalization of the blockwise bootstrap is considered by Politis and Romano [43] and by Politis, Romano, and Lai [44].

4.3 Linked blockwise bootstrap. The blockwise bootstrap sample $\vec{Z}_n^{0*}$ (obtained above) is not a good replicate of the original data $\vec{Z}_n^0$ in the following sense: The dependence structure near block "endpoints" is incorrect. For example, the bootstrap observations $\{Z_{jl}^*, Z_{jl+1}^*\}$ are adjacent in time but are [conditionally] independent! Graphically, the bootstrap sample $\vec{Z}_n^{0*}$ will exhibit anomalous behavior at the block endpoints. Künsch and Carlstein [36] propose the following modification of the blockwise bootstrap in order to correct this problem.

The "linked blockwise bootstrap" still selects the first block $\vec{Z}_{l,1}^*$ at random from the empirical ℓ-dimensional marginal distribution $F_{l,n}$. Now look at the final observation Z_l^* in this first block, and identify its p "nearest neighbors" among the set of original observations $\{Z_1, Z_2, \ldots, Z_{n-l}\}$. Randomly select one of these p nearest neighbors. The selected observation -- say, Z_v -- is the "link." The second bootstrap block is then taken to be $\vec{Z}_{l,2}^* = \vec{Z}_l^v$, the block of original data immediately following the link. The $(j+1)^{\underline{th}}$ bootstrap block $\vec{Z}_{l,j+1}^*$ is similarly obtained by randomly linking to the final observation Z_{jl}^* from the $j^{\underline{th}}$ bootstrap block.

This linked blockwise bootstrap is still based on blocks; hence $\vec{Z}_n^{0*}$ still has exactly the correct dependence structure within blocks -- without requiring any knowledge of the underlying dependence mechanism. The linked blockwise bootstrap improves on the blockwise bootstrap by guaranteeing a more natural transition from one bootstrap block to the next.

4.4 Subseries. The most simplistic way to generate replicates of t from the blocks is by calculating the statistic on the individual blocks themselves. Thus, for fixed ℓ, the $i^{\underline{th}}$ "subseries replicate" of t is

$$t_l^i := t_l(\vec{Z}_l^i) \ , \ i \in \{0, 1, \ldots, n-\ell\}.$$

These subseries replicates can be used to construct typical-values, to estimate moments of t, and for diagnostics on t's sampling distribution.

When t estimates a parameter θ, and $k := n/\ell$ is an integer, then the random variables

$$V_i := (t_l^0 + t_l^{il})/2 \ , \ \ i \in \{1, 2, \ldots, k-1\}$$

can behave asymptotically like typical-values for θ (Carlstein [16]). This approach is valid for a large class of asymptotically normal statistics t, including the sample mean and sample percentiles (see also Carlstein [13]).

The $p^{\underline{th}}$ moment of t's sampling distribution can be estimated via the $p^{\underline{th}}$ empirical moment of the subseries replicates:

$$\sum_{i=0}^{k-1} (t_l^{il})^p / k.$$

This method is consistent in a broad range of situations, and does not generally require the statistic t to be asymptotically normal (Carlstein [14, 15]). For an extension of this technique to spatial processes, see Possolo [45].

The subseries replicates can also be used for diagnostics, e.g., to graphically assess non-normality or skewness in t's sampling distribution. This suggestion is theoretically justified by a strong uniform consistency result for the empirical distribution of the subseries replicates (Carlstein [17]).

4.5 Other algorithms. For statistics obtained from estimating-equations, a jackknife algorithm (Lele [39]) and a bootstrap algorithm (Lele [40]) have been developed. Venetoulias [67] introduces a technique for generating replicates in image data. Rajarshi [49] proposes a "direct" method for estimating the variance of t; although his method actually does not involve any resampling, it is in the same spirit as the model-free algorithms of this Section.

REFERENCES

[1] M. A. ARCONES AND E. GINÉ, *The bootstrap of the mean with arbitrary bootstrap sample size*, Ann. Inst. H. Poincaré, 25 (1989), pp. 457-481.

[2] K. B. ATHREYA AND C. D. FUH, *Bootstrapping Markov chains: countable case*, Tech. Rep. (1989), Dept. of Statist., Iowa State U., Ames.

[3] I. V. BASAWA, T. A. GREEN, W. P. MCCORMICK, AND R. L. TAYLOR, *Asymptotic bootstrap validity for finite Markov chains*, Tech. Rep. (1989), Dept. of Statist., U. of Georgia, Athens.

[4] I. V. BASAWA, A. K. MALLIK, W. P. MCCORMICK, AND R. L. TAYLOR, *Bootstrapping explosive autoregressive processes*, Ann. Statist., 17 (1989), pp. 1479-1486.

[5] I. V. BASAWA, A. K. MALLIK, W. P. MCCORMICK, J. H. REEVES, AND R. L. TAYLOR, *Bootstrapping unstable autoregressive processes*, Tech. Rep. 122 (1989), Dept. of Statist., U. of Georgia, Athens.

[6] P. J. BICKEL AND D. A. FREEDMAN, *Some asymptotic theory for the bootstrap*, Ann. Statist., 9 (1981), pp. 1196-1217.

[7] A. BOSE, *Edgeworth correction by bootstrap in autoregressions*, Ann. Statist., 16 (1988), pp. 1709-1722.

[8] ——————, *Bootstrap in time series models*, in Proceedings of the Winter Simulation Conference, M. Abrams, P. Haigh, and J. Comfort (ed.s), IEEE, San Francisco, 1988, pp. 486-490.

[9] ——————, *Bootstrap in moving average models*, Ann. Instit. Statist. Math., (1989), to appear.

[10] J. BRETAGNOLLE, *Lois limites du Bootstrap de certaines fonctionnelles*, Ann. Inst. H. Poincaré, 19 (1983), pp. 281-296.

[11] D. R. BRILLINGER, Discussion of *A generalized least-squares approach to linear functional relationships*, J. Royal Statist. Soc. (Series B), 28 (1966), p. 294.

[12] ——————, Discussion of *Influence functionals for time series*, Ann. Statist., 14 (1986), pp. 819-822.

[13] E. CARLSTEIN, *Asymptotic r ormality for a general statistic from a stationary sequence*, Ann. Probab., 14 (1986), pp. 1371-1379.

[14] ——————, *The use of subseries values for estimating the variance of a*

general statistic from a stationary sequence, Ann. Statist., 14 (1986), pp. 1171-1179.

[15] ————————, *Law of large numbers for the subseries values of a statistic from a stationary sequence*, Statist., 19 (1988), pp. 295-299.

[16] ————————, *Asymptotic multivariate normality for the subseries values of a general statistic from a stationary sequence -- with applications to nonparametric confidence intervals*, Probab. Math. Statist., 10 (1989), pp. 191-200.

[17] ————————, *Using subseries values to describe the distribution of a general statistic from a stationary sequence*, Tech. Rep. in preparation (1990), Dept. of Statist., U. of N. Carolina, Chapel Hill.

[18] W. DAVIS, *Robust interval estimation of the innovation variance of an ARMA model*, Ann. Statist., 5 (1977), pp. 700-708.

[19] B. EFRON, *Bootstrap methods: another look at the jackknife*, Ann. Statist., 7 (1979), pp. 1-26.

[20] ————————, *The Jackknife, the Bootstrap and Other Resampling Plans*, SIAM, Philadelphia, 1982.

[21] B. EFRON AND R. TIBSHIRANI, *Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy*, Statistical Science, 1 (1986), pp. 54-77.

[22] D. F. FINDLEY, *On bootstrap estimates of forecast mean square errors for autoregressive processes*, in Computer Science and Statistics: The Interface, D. M. Allen (ed.), Elsevier (North-Holland), Amsterdam, 1986, pp. 11-17.

[23] J. FRANKE AND W. HÄRDLE, *On bootstrapping kernel spectral estimates*, Ann. Statist., (1990), to appear.

[24] J. FRANKE, W. HÄRDLE, AND J. KREISS, *Applications of the bootstrap in time series analysis*, Instit. Math. Statist. Bulletin, 19 (1990), pp. 427-428.

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

[26] D. FREEDMAN AND S. PETERS, *Bootstrapping a regression equation: some empirical results*, J. Amer. Statist. Assoc., 79 (1984), pp. 97-106.

[27] F. GÖTZE AND H. R. KÜNSCH, *Blockwise bootstrap for dependent observation: higher order approximations for studentized statistics*, Instit. Math. Statist. Bulletin, 19 (1990), p. 443.

[28] H. L. GRAY, T. A. WATKINS, AND J. E. ADAMS, *On the jackknife statistic, its extensions, and its relation to* e_n*-transformations*, Ann. Math. Statist., 43 (1972), pp. 1-30.

[29] J. A. HARTIGAN, *Using subsample values as typical values*, J. Amer. Statist. Assoc., 64 (1969), pp. 1303-1317.

[30] ——————, *Necessary and sufficient conditions for asymptotic joint normality of a statistic and its subsample values*, Ann. Statist., 3 (1975), pp. 573-580.

[31] ——————, *Perturbed periodogram estimates of variance*, Internat. Statist. Rev., 58 (1990), pp. 1-7.

[32] C. M. HURVICH AND S. L. ZEGER, *Frequency domain bootstrap methods for time series*, Tech. Rep. 87-115 (1987), Statist. & Oper. Res., Grad. School of Busi., New York U., New York.

[33] C. M. HURVICH, J. S. SIMONOFF, AND S. L. ZEGER, *Variance estimation for sample autocovariances: direct and resampling approaches*, Austral. J. Statist., (1990), to appear.

[34] R. KULPERGER AND B. L. S. PRAKASA RAO, *Bootstrapping a finite state Markov chain*, Sankhya (Series A), 51 (1989), pp. 178-191.

[35] H. R. KÜNSCH, *The jackknife and the bootstrap for general stationary observations*, Ann. Statist., 17 (1989), pp. 1217-1241.

[36] H. R. KÜNSCH AND E. CARLSTEIN, *The linked blockwise bootstrap for serially dependent observations*, Tech. Rep. in preparation (1990), Dept. of Statist., U. of N. Carolina, Chapel Hill.

[37] S. N. LAHIRI, *Second order optimality of stationary bootstrap*, Tech. Rep. (1990), Dept. of Statist., Iowa State U., Ames.

[38] S. LELE, *Non-parametric bootstrap for spatial processes*, Tech. Rep. 671 (1988), Dept. of Biostatist., Johns Hopkins U., Baltimore.

[39] ——————, *Jackknifing linear estimating equations: asymptotic theory and applications in stochastic processes*, J. Roy. Statist. Soc. (Series B), (1990), to appear.

[40] ——————, *Resampling using estimating equations*, Tech. Rep. 706 (1990), Dept. of Biostatist., Johns Hopkins U., Baltimore.

[41] R. Y. LIU AND K. SINGH, *Moving blocks jackknife and bootstrap capture weak dependence*, Tech. Rep. (1988), Dept. of Statist., Rutgers U., New Brunswick.

[42] J. S. MARITZ, *A note on exact robust confidence intervals for location*, Biometrika, 66 (1979), pp. 163-166.

[43] D. N. POLITIS AND J. P. ROMANO, *A general resampling scheme for triangular arrays of α-mixing random variables with application to the problem of spectral density estimation*, Tech. Rep. 338 (1989), Dept. of Statist.,

Stanford U., Stanford.

[44] D. N. POLITIS, J. P. ROMANO, AND T. LAI, *Bootstrap confidence bands for spectra and cross-spectra*, Tech. Rep. 342 (1990), Dept. of Statist., Stanford U., Stanford.

[45] A. POSSOLO, *Subsampling a random field*, Tech. Rep. 78 (1986), Dept. of Statist., U. of Washington, Seattle.

[46] M. QUENOUILLE, *Approximate tests of correlation in time series*, J. Roy. Statist. Soc. (Series B), 11 (1949), pp. 18-84.

[47] ———————, *Notes on bias in estimation*, Biometrika, 43 (1956), pp. 353-360.

[48] M. B. RAJARSHI, *Bootstrap in Markov sequences based on estimates of transition density*, Ann. Instit. Statist. Math., (1989), to appear.

[49] ———————, *On estimation of the variance of a statistic obtained from dependent observations*, Tech. Rep. (1990), Dept. of Statist., U. of Poona, Pune.

[50] E. RAMOS, *Resampling methods for time series*, Ph.D. Thesis (1988), Dept. of Statist., Harvard U., Cambridge.

[51] M. ROSENBLATT, *A central limit theorem and a strong mixing condition*, Proc. Nat. Acad. Sci., 42 (1956), pp. 43-47.

[52] J. SHAO AND C. F. J. WU, *A general theory for jackknife variance estimation*, Ann. Statist., 17 (1989), pp. 1176-1197.

[53] K. SINGH, *On the asymptotic accuracy of Efron's bootstrap*, Ann. Statist., 9 (1981), pp. 1187-1195.

[54] A. R. SOLOW, *Bootstrapping correlated data*, Math. Geology, 17 (1985), pp. 769-775.

[55] R. A. STINE, *A frequency domain bootstrap for time series*, Tech. Rep. (1985), Dept. of Statist., Wharton School, U. of Pennsylvania, Philadelphia.

[56] ———————, *Estimating properties of autoregressive forecasts*, J. Amer. Statist. Assoc., 82 (1987), pp. 1072-1078.

[57] D. S. STOFFER AND K. D. WALL, *Bootstrapping state space models: Gaussian maximum likelihood estimation and the Kalman filter*, Tech. Rep. (1990), Dept. of Math. & Statist., U. of Pittsburgh, Pittsburgh.

[58] J. W. H. SWANEPOEL, *A note on proving that the (modified) bootstrap works*, Commun. Statist. Theory Meth., 15 (1986), pp. 3193-3203.

[59] J. W. H. SWANEPOEL AND J. W. J. VAN WYK, *The bootstrap applied to power spectral density function estimation*, Biometrika, 73 (1986), pp. 135-141.

[60] L. A. THOMBS, *Prediction for* $ARMA(p,q)$ *series using the bootstrap*, Instit. Math. Statist. Bulletin, 16 (1987), p. 225.

[61] ——————, *Model dependent resampling plans for time series*, Tech. Rep. (1987), Dept. of Statist., U. of S. Carolina, Columbia.

[62] L. A. THOMBS AND W. R. SCHUCANY, *Bootstrap prediction intervals for autoregression*, Proceedings of the Decision Sciences Institute, (1987), pp. 179-181.

[63] ——————, *Bootstrap prediction intervals for autoregression*, J. Amer. Statist. Assoc., 85 (1990), pp. 486-492.

[64] D. THOMSON AND A. CHAVE, *Jackknifed error estimates for spectra, coherences, and transfer functions*, Tech. Rep. (1988), AT&T Bell Laboratories, Murray Hill.

[65] J. TUKEY, *Bias and confidence in not quite large samples*, Ann. Math. Statist., 29 (1958), p. 614.

[66] W. VAN ZWET, *Second order asymptotics and the bootstrap*, Lect. Notes (1990), Dept. of Statist., U. of N. Carolina, Chapel Hill.

[67] A. VENETOULIAS, *Artificial replication in the analysis of image data*, Instit. Math. Statist. Bulletin, 19 (1990), p. 345.