The Annals of Statistics
1986, Vol. 14, No. 3, 1171–1179

THE USE OF SUBSERIES VALUES FOR ESTIMATING THE VARIANCE OF A GENERAL STATISTIC FROM A STATIONARY SEQUENCE[1]

BY EDWARD CARLSTEIN

University of North Carolina

Let $\{Z_i: -\infty < i < +\infty\}$ be a strictly stationary α-mixing sequence. Without specifying the dependence model giving rise to $\{Z_i\}$, and without specifying the marginal distribution of Z_i, we address the question of variance estimation for a general statistic $t_n = t_n(Z_1, \dots, Z_n)$. For estimating $\mathrm{Var}\{t_n\}$ from just the available data $(Z_1, \dots, Z_n)$, we propose computing subseries values: $t_m(Z_{i+1}, Z_{i+2}, \dots, Z_{i+m})$, $0 \le i < i + m \le n$. These subseries values are used as replicates to model the sampling variability of t_n. In particular, we use adjacent nonoverlapping subseries of length $m = m_n$, with $m_n \to \infty$ and $m_n/n \to 0$. Our variance estimator is just the usual sample variance computed amongst these subseries values (after appropriate standardization). This estimator is shown to be consistent under mild integrability conditions. We present optimal (i.e., minimum m.s.e.) choices of m_n for the special case where $t_n = \bar{Z}_n$ and $\{Z_i\}$ is a normal AR(1) sequence. A simulation study is conducted, showing that those same choices of m_n are effective when t_n is a robust estimator of location and $\{Z_i\}$ is subject to contamination.

1. Introduction. Consider this situation: a scientist is faced with data $\mathbf{Z}_n = (Z_1, \dots, Z_n)$ from a stationary sequence $\{Z_i: -\infty < i < +\infty\}$. He does not know what underlying dependence model (M) produced $\{Z_i\}$, nor does he know the distribution (F) of the Z_i's. The latter may include large-variance contamination. A statistic $t_n = t_n(\mathbf{Z}_n)$ is computed, e.g., a trimmed mean to estimate the level of the sequence, or a robust estimate of scale. In order to make any inferences from t_n, an estimate of the variance of t_n will be needed. Our objective is to provide a practicable and theoretically sound technique for calculating such a variance estimate—without assuming knowledge of M or F. This is accomplished by using (as replicates) "subseries values" of the statistic t computed on "subseries": $(Z_{i+1}, Z_{i+2}, \dots, Z_{i+m})$, $0 \le i < i + m \le n$. The literature contains no other procedure to address this question in its full generality. Furthermore, even if specific assumptions were made about M (e.g., autoregression) and F (e.g., joint normality), actual calculation of the theoretical variance of t_n in terms of the parameters of M and F may be intractable. This again points to the need for a nonparametric variance estimator for general statistics from dependent sequences.

The setting we address is more complex than the iid case due to the presence of M (be it known or unknown). Therefore the practical appeal of the bootstrap

Received November 1984; revised September 1985.

[1] Supported by NSF Grants MCS-8102725 and DMS-8400602.

AMS 1980 *subject classifications.* Primary 62G05; secondary 60G10.

Key words and phrases. Variance estimation, subseries, nonparametric, dependence, α-mixing, stationary.

and jackknife estimates of variance applies *a fortiori* to our variance estimator: it "can be applied to complicated situations where parametric modeling and/or theoretical analysis is hopeless" (Efron, 1982).

After presenting our basic notation and definitions in Section 2, we proceed in Section 3 to formally define our variance estimator and to discuss it in comparison with other variance estimators in the literature. Section 4 establishes conditions under which our estimator is consistent in the L_2 sense. This consistency result is combined with a distributional result from Carlstein (1986) to yield asymptotic normality for general statistics from α-mixing sequences—with the limiting distribution being free of the nuisance parameter σ^2. In Section 5 we determine analytically the optimal choices of m (subseries length) for a useful class of special cases. Finally, in Section 6, using the results of Section 5 as a guide, we apply the variance estimator (via simulations) to precisely the sort of situation described at the outset of this introduction.

2. Definitions and notation. Let $\{Z_i(\omega)\colon -\infty < i < +\infty\}$ be a strictly stationary sequence of real-valued random variables (r.v.) defined on probability space (Ω, F, P). Let F_p^+ (F_q^-, respectively) be the σ-field generated by $\{Z_p(\omega), Z_{p+1}(\omega), \dots\}$ ($\{\dots, Z_{q-1}(\omega), Z_q(\omega)\}$, respectively).

For $N \geq 1$ denote: $\alpha(N) = \sup\{|P\{A \cap B\} - P\{A\}P\{B\}|\colon A \in F_N^+, B \in F_0^-\}$, and define *$\alpha$-mixing* to mean $\lim_{N\to\infty}\alpha(N) = 0$.

Let $t_n(z_1, \dots, z_n)$ be a function from $R^n \to R^1$, defined for each $n \geq 1$ so that $t_n(Z_1(\omega), \dots, Z_n(\omega))$ is F-measurable. Suppressing the argument ω of $Z_i(\cdot)$ from here on, we denote $\mathbf{Z}_n^i = (Z_{i+1}, Z_{i+2}, \dots, Z_{i+n})$ and $t_n^i = t_n(\mathbf{Z}_n^i)$; as a particular case: $\bar{Z}_n^i = \sum_{j=1}^n Z_{i+j}/n$.

For $B \geq 0$ denote: ${}_BX = X \cdot I\{|X| < B\}$ and ${}^BX = X - {}_BX$. Expectation, variance, and covariance will be denoted by E, V, and C, respectively.

Random variables $\{X_n\}$ will be said to be *uniformly integrable* (u.i.) iff: $\exists\ n_0$ s.t. $\lim_{A\to\infty}\sup_{n\geq n_0}E\{|{}^AX_n|\} = 0$. It will at times be convenient to use the equivalent condition: $\lim_{A\to\infty}\limsup_{n\to\infty}E\{|{}^AX_n|\} = 0$.

3. The variance estimator. Most variance estimation techniques for general statistics have been aimed at the special case where $\{Z_i\}$ is iid. Tukey's "jackknife," Hartigan's "typical values," and Efron's "bootstrap" [see Efron (1982) for descriptions] all make heavy use of exchangeability in their schemes for generating replicates of t. These techniques are based on the idea that by computing the statistic t on subsamples of the data $\mathbf{Z}_n^0$, we can gain insight about the sampling distribution of t_n^0. The bootstrap, for example, resamples data from the empirical distribution of $\mathbf{Z}_n^0$, and then recalculates the statistic t on each of these "bootstrap" samples. These replicates of t serve as an empirical approximation to the true sampling distribution of t_n^0. This approximation is sensible when $\{Z_i\}$ is iid; but when nontrivial dependence is present in $\{Z_i\}$, the true sampling distribution of t_n^0 depends on the *joint* distribution of $\mathbf{Z}_n^0$. Thus, the only subsamples that will yield valid replicates of t are those that preserve the dependence structure in $\{Z_i\}$. Therefore we shall focus on subsamples of the form $\{\mathbf{Z}_m^j\colon 0 \leq j \leq n - m, n \geq m \geq 1\}$.

[Recently, Freedman (1984) has considered applying the bootstrap to a linear model with autoregressive component; this approach assumes additive iid perturbations. Also, as he emphasizes, the bootstrap calculations assume that the user has correctly specified the form of the underlying autoregressive model.]

We face several competing considerations in designing a variance estimator based on $\{t_m^j\colon 0 \le j \le n - m,\ n \ge m \ge 1\}$. It is clear that the performance of such an estimator will depend upon how many representative subseries values t_m^j are used, how different the t_m^j's are from each other, and how accurately the t_m^j's model the behavior of t_n^0. For a particular value of m, one would not expect t_m^j and t_m^{j+1} to differ by much—especially in light of the dependence between $\mathbf{Z}_m^j$ and Z_{j+m+1}. Hence the collection of subseries values $\{t_m^j\colon 0 \le j \le n - m\}$ contains a great deal of redundancy that may not contribute information about t_n^0's sampling variability. The collection $\{t_m^{jm}\colon 0 \le j \le [n/m] - 1\}$, on the other hand, contains only nonoverlapping subseries values. If m grows with n, each t_m^{jm} will eventually behave as if it were independent of all but two of the other t_m^{jm}'s. Furthermore, if m remained fixed, a subseries value t_m^j would never be able to reflect the dependencies of lag $m + 1$ or greater. These arguments suggest the use of $\{t_{m_n}^{jm_n}\colon 0 \le j \le [n/m_n] - 1\}$, with $m_n \to \infty$ as $n \to \infty$.

Within this framework it seems reasonable to consider $m_n = [\beta n]$ $(0 < \beta < 1)$, since the corresponding $t_{m_n}^{jm_n}$'s are based on subseries of the same order of magnitude as t_n^0 itself. Unfortunately, only about $1/\beta$ disjoint $t_{m_n}^{jm_n}$'s of this form will ever be available. So an estimator based on such $t_{m_n}^{jm_n}$'s will never stabilize and home in on σ^2, even as $n \to \infty$. (Ironically, the bootstrap and typical-value methods use randomly selected subsets of the possible subsamples, since it is computationally impractical to use all the subsamples available.)

In light of these factors we propose the use of the subseries values $\{t_{m_n}^{jm_n}\colon 0 \le j \le k_n - 1\}$, where $\{m_n\colon n \ge 1\}$ are positive integers s.t. $m_n \to \infty$ and $m_n/n \to 0$ as $n \to \infty$, and $k_n = [n/m_n]$. Thus we obtain an increasing number (k_n) of subseries values, each of which is based on an ever-growing subseries $(\mathbf{Z}_{m_n}^{jm_n})$; and each $t_{m_n}^{jm_n}$ is becoming increasingly distant (m_n) from all but two of the other $t_{m_n}^{im_n}$'s.

From this point on we will assume the following set-up: $s_n^i := s_n(\mathbf{Z}_n^i)$ is a statistic that is wholly computable from the data $\mathbf{Z}_n^i$, and does not involve any unknown parameters. $t_n^i := (s_n^i - E\{s_n^0\})n^{1/2}$ is the correct theoretical standardization for s_n^i, in the sense that $\lim_{n\to\infty} E\{(t_n^0)^2\} =: \sigma^2 \in (0, \infty)$. The proposed estimator for σ^2 is simply

$$\hat{\sigma}_n^2 := m_n \sum_{i=0}^{k_n-1} \left(s_{m_n}^{im_n} - \bar{s}_n\right)^2 / k_n, \quad \text{where } \bar{s}_n := \sum_{i=0}^{k_n-1} s_{m_n}^{im_n}/k_n.$$

This is nothing more than the usual sample variance amongst the standardized subseries values $\{m_n^{1/2} s_{m_n}^{jm_n}\colon 0 \le j \le k_n - 1\}$.

4. L_2-consistency. In this section we work out some theory for subseries values. The first main result is a law of large numbers for these entities. This result is used to obtain consistency of $\hat{\sigma}_n^2$. Finally, we arrive at an asymptotic normality result for t_n^0 in which the limiting distribution is free of σ^2.

Let us begin with a useful truncation lemma:

Lemma 1. *Let X be F_q^+-measurable and Y be F_p^--measurable, $q > p$. Suppose* $\max\{E\{X^2\}, E\{Y^2\}\} \le C < \infty$. *Then for any* $A > 0$: $|C\{X, Y\}| \le 4A^2\alpha(q-p) + 3C^{1/2}((E\{({}^AX)^2\})^{1/2} + (E\{({}^AY)^2\})^{1/2})$.

Proof. Writing $X = {}_AX + {}^AX$ we see that

$$\begin{aligned}|C\{X,Y\}| \le\; & |C\{{}_AX, {}_AY\}| + |E\{{}_AX \cdot {}^AY\}| + |E\{{}^AX \cdot {}_AY\}| + |E\{{}^AX \cdot {}^AY\}| \\ & + |E\{{}_AX\}E\{{}^AY\}| + |E\{{}^AX\}E\{{}_AY\}| + |E\{{}^AX\}E\{{}^AY\}|.\end{aligned}$$

The first term on the right-hand side is bounded above by $4A^2\alpha(q-p)$ [Theorem 17.2.1, Ibragimov and Linnik (1971)]. The required bounds on the other terms follow from the Schwarz inequality. □

Applying this lemma we can establish the following law of large numbers for subseries values from an α-mixing sequence.

Theorem 2. *Let $\{Z_i\}$ be α-mixing and let $f_n(\mathbf{Z}_n^i) = f_n^i$ be a statistic. Let $\{m_n\colon n \ge 1\}$ be s.t. $m_n \to \infty$ and $m_n/n \to 0$; let $k_n = [n/m_n]$. Define $\bar{f}_n = \sum_{i=0}^{k_n-1} f_{m_n}^{im_n}/k_n$. If*

(2a) $$\lim_{n\to\infty} E\{f_n^0\} = \phi \in R^1,$$

and

(2b) $$(f_n^0)^2 \text{ are u.i.,}$$

then

(2c) $$\bar{f}_n \to_{L_2} \phi \quad \text{as } n \to \infty.$$

Proof. By (2a) it suffices to show $\lim_{n\to\infty} V\{\bar{f}_n\} = 0$. Now

$$\begin{aligned}\tfrac{1}{2}V\{\bar{f}_n\}k_n^2 &\le \sum_{0 \le i \le j \le k_n - 1} \left|C\{f_{m_n}^{im_n}, f_{m_n}^{jm_n}\}\right| \\ &\le \left(2E\{(f_{m_n}^0)^2\} + \sum_{j=2}^{k_n-1} \left|C\{f_{m_n}^0, f_{m_n}^{jm_n}\}\right|\right)k_n.\end{aligned}$$

The idea here is that the covariance between nonadjacent $f_{m_n}^{jm_n}$'s is dropping off as the separation (m_n) increases. So, although there are order k_n of these terms, their average becomes negligible as $n \to \infty$.

Formally, we note first that [by (2b)] $E\{(f_n^0)^2\}$ are bounded uniformly in $n \ge n_0$ by $C < \infty$. Assume now that n is sufficiently large so that $m_n \ge n_0$. Then for each $j \in \{2, 3, \dots, k_n - 1\}$ we have

$$\left|C\{f_{m_n}^0, f_{m_n}^{jm_n}\}\right| \le 4A^2\alpha(m_n) + 6C^{1/2}\left(E\{({}^Af_{m_n}^0)^2\}\right)^{1/2} =: B(n, A) \quad \text{for any } A > 0$$

by Lemma 1. Hence: $\frac{1}{2}V\{\bar{f}_n\} \le 2C/k_n + B(n, A)$ for any $A > 0$. Now take $\lim_{A\to\infty}\limsup_{n\to\infty}(\cdot)$ of this last expression. □

We are ready to prove L_2-consistency of $\hat{\sigma}_n^2$. This result follows in part from Theorem 2, since $\hat{\sigma}_n^2$ is essentially a mean.

THEOREM 3. *Let $\{Z_\iota\}$ be α-mixing and let $\{m_n\}$ and k_n be as in Theorem 2. Let s_n^ι, t_n^ι, σ^2, $\hat{\sigma}_n^2$ be as defined in Section 3. If*

$$\text{(3a)} \qquad \left(t_n^0\right)^4 \textit{ are u.i.}$$

then

$$\text{(3b)} \qquad \hat{\sigma}_n^2 \to_{L_2} \sigma^2 \quad \textit{as } n \to \infty.$$

PROOF. Write $\hat{\sigma}_n^2 = \Sigma_n - (\bar{t}_n)^2$, where $\bar{t}_n = \Sigma_{\iota=0}^{k_n-1} t_{m_n}^{\iota m_n}/k_n$ *and* $\Sigma_n = \Sigma_{\iota=0}^{k_n-1}(t_{m_n}^{\iota m_n})^2/k_n$. Clearly we only need to show $\Sigma_n \to_{L_2} \sigma^2$ and $(\bar{t}_n)^2 \to_{L_2} 0$. The former follows from Theorem 2.

In order to show $\bar{t}_n \to_{L_4} 0$, note first that $\bar{t}_n \to_P 0$ by Theorem 2. Therefore, by the mean convergence criterion [see Chow and Teicher (1978), page 98], it suffices to establish that $(\bar{t}_n)^4$ are u.i. Now $(\bar{t}_n)^2 \le \Sigma_n$, so that for $A > 0$: $E\{(\bar{t}_n)^4 I\{(\bar{t}_n)^4 \ge A\}\} \le E\{(\Sigma_n)^2 I\{(\Sigma_n)^2 \ge A\}\}$. Hence we only need to show u.i. of $(\Sigma_n)^2$. But by (3a) we know that $E\{(\Sigma_n)^2\} < \infty$ when $m_n \ge n_0$; and $\Sigma_n \to_{L_2} \sigma^2$ as mentioned above. Thus the mean convergence criterion (converse) yields the required result. □

Notice that both Theorems 2 and 3 are logically independent of the question of convergence in distribution. These results give integrability conditions that guarantee L_2-consistency of estimators based on the subseries values from an α-mixing sequence—regardless of whether the t_n^0's (or $f_n{}^0$'s) are converging in distribution. Furthermore, we have not constrained the mixing coefficient α or the subseries length m_n in any way other than $\alpha(n) \to 0$, $m_n \to \infty$, $m_n/n \to 0$. In practice the L_2-consistency is desirable because it translates into shrinking variance and bias.

We can now combine the variance estimation result (Theorem 3) with the distributional results of Carlstein (1986), and obtain:

THEOREM 4. *Let $\{Z_i\}$ be α-mixing and let $\{m_n\}$, s_n^i, t_n^ι, $\hat{\sigma}_n^2$ be as defined in Theorem 3. If*

$$\text{(4a)} \qquad \lim_{r_n \ge u_n + v_n \ge v_n \to \infty} (r_n/v_n)^{1/2} C\left\{t_{r_n}^0, t_{v_n}^{u_n}\right\} = \sigma^2,$$

and

$$\text{(4b)} \qquad \limsup_{n \to \infty} E\left\{\left(t_n^0\right)^4\right\} = 3\sigma^4,$$

then (3b) *holds, and also*

$$\text{(4c)} \quad \left(t_{r_n}^0, t_{v_n}^{u_n}\right)/\hat{\sigma}_n \xrightarrow[v_n/r_n \to \rho^2,\ r_n \ge u_n + v_n \ge v_n \to \infty]{D} N_2(0,0,1,1,\rho) \quad \forall \rho^2 \in [0,1].$$

[*The generalized limit notation is the same as that defined in Carlstein* (1986).]

PROOF. We will begin by showing that $(t_{r_n}^0, t_{v_n}^{u_n})/\sigma \to_D N_2(0,0,1,1,\rho)$, via Theorem 4 of Carlstein (1986). Since $E\{t_n^0\} \equiv 0$, it suffices to observe that (4b) implies that $(t_n^0)^2$ are u.i.

Next we want to use Theorem 3 to conclude that (3b) holds. In light of (4a) with $u_n \equiv 0$ and $r_n = v_n = n$, it is enough to verify (3a). But (3a) follows directly from (4b) together with $t_n^0 \to_D N(0, \sigma^2)$ (established above). □

[Condition (4b) may of course be replaced by the less specific condition (3a).] The sample mean and sample fractile statistics are discussed as theoretical examples in Corollaries 8 and 10 (respectively) of Carlstein (1986).

5. Optimal subseries length. The results of Section 4 gave an asymptotic justification for the use of $\hat{\sigma}_n^2$. In fact, the asymptotics held for an extremely large class of sequences $\{m_n\}$ of subseries lengths. In practice, however, the performance of $\hat{\sigma}_n^2$ (for fixed n) will be greatly influenced by the particular choice of m_n. Our intuition tells us that by increasing m_n we should reduce the bias of $\hat{\sigma}_n^2$, since our replicates $(t_{m_n}^{jm_n})^2$ will more closely resemble the large-N $(t_N^0)^2$ whose expectation is being estimated. Furthermore, as the dependence in $\{Z_i\}$ becomes stronger, we will need longer subseries in order for $t_{m_n}^{jm_n}$ to adequately model the dependence present in t_N^0. On the other hand, by decreasing m_n (i.e., increasing k_n) we expect to reduce the variance of $\hat{\sigma}_n^2$, since more replicates become available. This interaction of bias, variance, and dependence will yield an optimal (i.e., minimum m.s.e.) m_n, for a given statistic t and a fixed n. Unfortunately, but not surprisingly, we are unable to make statements regarding optimal subseries length that apply with the generality of the consistency results in Section 4. We can, however, make very precise statements in the following special case.

Let $\{Z_i\}$ be an AR(1) sequence: $Z_i = \phi Z_{i-1} + \varepsilon_i$, where $|\phi| < 1$ and $\{\varepsilon_i\}$ are iid $N(0,1)$. The statistic $s_n^0 = \bar{Z}_n^0$ has asymptotic variance $\sigma^2 = (1-\phi)^{-2}$, which is to be estimated by $\tilde{\sigma}_n^2 := m_n \sum_{i=0}^{k_n-1} (s_{m_n}^{im_n})^2 / k_n$ $(E\{s_n^0\} = 0)$. In this situation, we can explicitly calculate the effect of subseries length on bias and variance:

$$\sigma^2 - E\{\tilde{\sigma}_n^2\} = 2\phi e/a^2 c m_n = 2\phi/a^2 c m_n + o(1/m_n),$$

and

$$\begin{aligned}
V\{\tilde{\sigma}_n^2\} = 2\Big\{ & b^2/a + 2\phi m_n^{-1}\Big[3(\phi^2 - 2\phi - 4)/b \\
& + \big(b - 7d + 11\phi d + 3\phi(5\phi^2 - 4\phi - 5)/b\big)/a^2 \\
& + 3\big(3d + (\phi^3 + 2\phi^2 + 3)/b\big)/a + m_n^{-1}\Big(-3\phi f/bc \\
& + 6\big[(\phi^2 + 2\phi^3 + 8\phi^4 - 2db + g^2 b)/c \\
& - (5 - 11\phi + 4\phi^2 + 8\phi^3 + 5g - d)/a \\
& + (g^2 + 2\phi^4 - 3\phi^2 - f)/bc + 2(\phi^6 - g^2)/c^2 \\
& + (2d + 2g - 2\phi^4 + \phi^3 - 3\phi)/a^2 + 3g + 5a\big] \\
& - \phi e^2/a^3 + 6\phi\big[3(1 + 2\phi - \phi^2)/b + 2(d^2 - \phi^6)/bc \\
& + (3\phi^2 - 4\phi^3 + 2\phi^4 - 5d + 4d\phi)/a\big]/a^2 \\
& + \phi\big[1 - (1 - d^{2k_n})/k_n f\big] \\
& \quad \times \big[1 - d^4 + 2d^3 - 2d + 2e^3(\phi - d)/a\big]/fa^2 b\Big)\Big]\Big\}/k_n c^2 a \\
= {} & 2/a^4 k_n + o(1/k_n),
\end{aligned}$$

where $a = 1 - \phi$, $b = 1 + \phi$, $c = ab$, $d = \phi^{m_n}$, $e = 1 - d$, $f = 1 - d^2$, $g = d/\phi$.

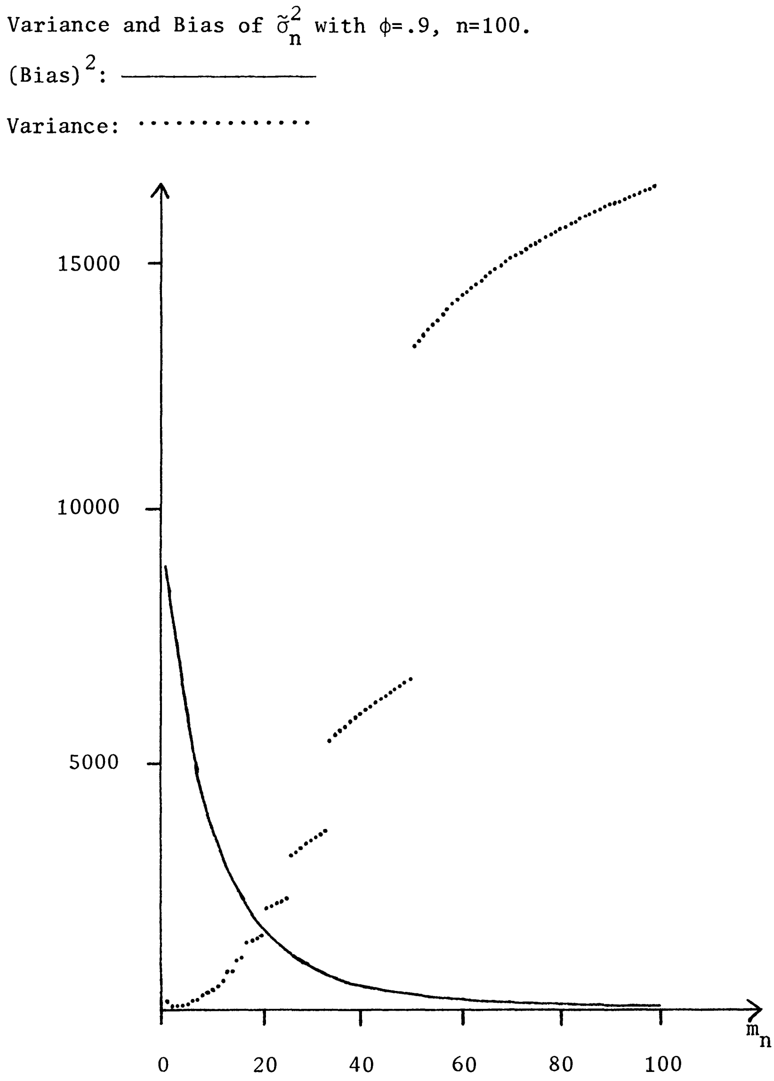

FIG. 1. *Variance and bias of $\tilde{\sigma}_n^2$ with $\phi = 0.9$, $n = 100$. (Bias)^2: —; Variance: ---.*

Figure 1 illustrates the influence of m_n and k_n on the bias and variance in the case $n = 100$, $\phi = 0.9$. The jumps in $V\{\tilde{\sigma}_n^2\}$ are due to abrupt changes in k_n. Notice that $V\{\tilde{\sigma}_n^2\}$ increases with m_n even while the number of replicates remains fixed.

Using just the first-order contributions from the bias and variance, we approximate

$$\text{m.s.e.}\{\tilde{\sigma}_n^2\} \approx (4\phi^2/a^4c^2)m_n^{-2} + (2/a^4)m_n/n.$$

Hence the optimal subseries length is approximately

$$m_n^* = (2\phi/c)^{2/3} n^{1/3},$$

with corresponding m.s.e $3(2\phi/c)^{2/3}/a^4 n^{2/3}$. Observe that longer subseries are required as the dependence becomes stronger.

6. Application. Let $\{Z_i\}$ be an AR(1) sequence: $Z_i = \phi Z_{i-1} + \varepsilon_\iota$, where $|\phi| < 1$ and ε_ι are iid errors from the contaminated distribution $(1 - \pi)F_1(\cdot) + \pi F_\tau(\cdot)$ [$F_c(\cdot)$ denotes the c.d.f. of a $N(0, c^2)$ r.v.]. A scientist observes $\mathbf{Z}_n^0$, and, suspecting contamination, he computes a $\delta\%$ trimmed-mean (our s_n^0) to estimate $E\{Z_\iota\}$. In order to estimate the variance of s_n^0, he will apply $\hat{\sigma}_n^2$. [Gastwirth and Rubin (1975) give an expression for the asymptotic variance of s_n^0 in terms of an infinite sum of Hermite polynomials—assuming, however, that $\{Z_i\}$ is a normal sequence.]

We propose using the results of Section 5 simply as a guide in selecting an appropriate m_n: the scientist can calculate

$$\hat{\phi} = n \sum_{i=1}^{n-1} (Z_{i+1} - \bar{Z}_n^0)(Z_i - \bar{Z}_n^0)/(n-1) \sum_{i=1}^{n} (Z_i - \bar{Z}_n^0)^2$$

as a preliminary measure of the strength of dependence in $\{Z_i\}$. Based on this $\hat{\phi}$, he can now estimate m_n^*. Although the resulting choices of m_n are not in general going to be optimal, this is a realistic strategy, given the amount of information available to the practitioner.

The entire procedure described above was carried out on 200 realizations of $\mathbf{Z}_n^0$, with: $\pi = 0.3$, $\tau^2 = 10$, $\delta = 40\%$ (20% in each tail), $\phi = 0.2$ and 0.8, $n = 100$ and 1000. Table 1 shows that this procedure yields reasonable results. A balance between variance and bias is maintained, and m.s.e.-consistency is exhibited. Moreover, the quality of the performance of $\hat{\sigma}_n^2$ is not affected by the strength of dependence in $\{Z_i\}$.

TABLE 1

*Simulation study of $\hat{\sigma}_n^2$ as an estimator of the variance (σ^2) of a 40% trimmed mean (s_n^0).**

ϕ	$\sigma^{2\dagger}$	n	$E\{\hat{\sigma}_n^2\}^{\dagger\dagger}$	$V\{\hat{\sigma}_n^2\}$	$V\{\hat{\sigma}_n^2\}$ / m.s.e.$\{\hat{\sigma}_n^2\}$	m.s.e.$\{\hat{\sigma}_n^2\}$ / σ^4
0.2	3.3	100	4.5 (0.10)	1.95	0.57	0.31
		1000	4.0 (0.03)	0.16	0.25	0.06
0.8	88	100	50 (2.6)	1383	0.49	0.36
		1000	71 (1.1)	232	0.45	0.07

*The data are from an AR(1) sequence with 30% contamination. Subseries lengths (m_n) are based on m_n^*.

†Each $\sigma^2 = \lim_{N \to \infty} V\{N^{1/2} s_N^0\}$ was estimated empirically by 200 realizations of $N^{1/2} s_N^0$ with $N = 200$.

††Each row was estimated empirically by 200 realizations of $\hat{\sigma}_n^2$. An estimate of the standard deviation of $E\{\hat{\sigma}_n^2\}$ appears in parentheses.

Acknowledgments. I thank Professor John Hartigan, my thesis advisor at Yale University, for his guidance on this research. The suggestions of the referees and the Associate Editor were extremely constructive; in particular their comments led to the material of Sections 5 and 6.

REFERENCES

CARLSTEIN, E. (1986). Asymptotic normality for a general statistic from a stationary sequence. *Ann. Probab.* **14** 1371–1379.

CHOW, Y. S. and TEICHER, H. (1978). *Probability Theory*. Springer, New York.

EFRON, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Philadelphia.

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

GASTWIRTH, J. L. and RUBIN, H. (1975). The behavior of robust estimators on dependent data. *Ann. Statist.* **3** 1070–1100.

IBRAGIMOV, I. A. and LINNIK, YU. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, The Netherlands.

DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27514