INFERENCE FOR AUTOCORRELATIONS
UNDER WEAK ASSUMPTIONS

by

Joseph P. Romano
Lori A. Thombs

Technical Report No. 290
August 1993

Prepared Under the Auspices
of
National Science Foundation Grant
DMS 90-02502
Ingram Olkin, Project Director

Department of Statistics
Stanford University
Stanford, California
INFERENCE FOR AUTOCORRELATIONS
UNDER WEAK ASSUMPTIONS

by

Joseph P. Romano
Stanford University

Lori A. Thombs
University of South Carolina

Technical Report No. 290
August 1993

Prepared Under the Auspices
of
National Science Foundation Grant
DMS 90-02502
Ingram Olkin, Project Director

Also supported by National Science Foundation Grant DMS92-04864 and issued
as Technical Report No. 436, Department of Statistics, Stanford University

Department of Statistics
Stanford University
Stanford, California
Summary

In time series analysis, it is well known that if the process is pure white noise, then the sample autocorrelation estimates, scaled by the square root of the sample size, are asymptotically standard normal. This result is used extensively as a diagnostic check on the residuals of a fitted model, or as an initial test on the observed time series to determine if further model fitting is warranted. In this article, we show that this result can be quite misleading. Specifically, if the underlying process is assumed to be uncorrelated rather than independent, the asymptotic distribution is not necessarily standard normal. Although this distinction may appear superficial, the implications for making valid inference in time series modeling are broad. Usual procedures in time series analysis model correlation structure by fitting models whose estimated errors mimic an uncorrelated sequence. Therefore testing for the presence of zero autocorrelation using a result which assumes independence may lead to incorrect conclusions. Furthermore, there exist stationary time series with zero autocorrelation at all lags, yet are not independent, and so it is important to have valid procedures under general dependence structures. In this paper, we present general asymptotic theory for the estimated autocorrelation function and discuss testing for lack of correlation without the further assumption of independence. Appropriate resampling methods are proposed which can be used to approximate the sampling distribution of the autocorrelation estimates under weak assumptions.

KEY WORDS: Autocorrelation Function; Approximate confidence limit; Bootstrap; Stationary Time Series; Subsamples.
1. Introduction

Let $X_1, \ldots, X_n$ denote a realization of a real-valued stationary time series. Unless stated otherwise, we will assume the time series is stationary in the strict sense. This paper focuses on inference of the autocorrelation structure of the series. Define the autocovariance and autocorrelation for lag $k$ by,

$$R(k) = \text{Cov}(X_1, X_{1+k})$$

and

$$\rho(k) = \text{Corr}(X_1, X_{1+k}).$$

The usual estimates of $R(k)$ and $\rho(k)$ are

$$\hat{R}_n(k) = \frac{1}{n} \sum_{i=1}^{n-k} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)$$

and

$$\hat{\rho}_n(k) = \frac{\hat{R}_n(k)}{\hat{R}_n(0)},$$

where $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ is the sample mean. For convenience, we will also consider

$$\tilde{R}_n(k) = \frac{1}{n} \sum_{i=1}^{n-k} (X_i - \mu)(X_{i+k} - \mu)$$

where $\mu = E(X_i)$.

Inference for $\hat{\rho}_n(k)$ when the process $\{X_t\}$ is linear with i.i.d. innovations is well-known. In particular, if $\{X_t\}$ is the stationary process,

$$X_t - \mu = \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j},$$

where $\{\epsilon_t\}$ are independent with mean 0 satisfying $E(\epsilon_t^4) < \infty$ then the distribution of the first $k$ sample autocorrelations are asymptotically normal with mean $[\rho(1), \rho(2), \ldots, \rho(k)]$ and covariance matrix $n^{-1}W$, where the $(i,j)_{th}$ element of $W$ is given by

$$w_{ij} = \sum_{j=-\infty}^{\infty} \{\rho(k+i)\rho(k+j) + \rho(k-i)\rho(k+j) + 2\rho(i)\rho(j)\rho^2(k)$$

$$-2\rho(i)\rho(k)\rho(k+j) - 2\rho(j)\rho(k)\rho(k+i)\}.$$
These results are extensively in time series modeling. Perhaps the most widely-used application of Bartlett's (1946) formula (1.2) is when the process \( \{X_t\} \) is pure white noise (independent). In this case, \( \rho(k) = 0 \) for all \( k \neq 0 \), leading to the standard result, \( \hat{\rho}_n(k) \sim \text{Normal}(0, n^{-1}) \). A common diagnostic check on a fitted model is to apply this result to the estimated autocorrelation function of the residuals. Typically, \( \pm 2/n^{1/2} \) limits are superimposed on a graph of \( \hat{\rho}(k) \) versus \( k \). Using such a "check" of independence on possibly only uncorrelated quantities may lead to misleading results. Indeed, it will be seen that the asymptotic distribution of \( \hat{\rho}_n(k) \) cannot in general be approximated by the normal distribution with mean 0 and variance \( n^{-1} \) if we only assume the underlying process is uncorrelated. In particular, there exist uncorrelated processes for which the asymptotic variance of \( \hat{\rho}_n(k) \) is any nonnegative value.

Let us consider (1.2) in another setting. For the MA(1) process, \( X_t = \epsilon_t + \theta \epsilon_{t-1} \), the limiting variance of \( n^{1/2} \hat{\rho}_n(1) \) reduces to \( 1 - 3\rho^2(1) + 4\rho^4(1) \), and for lags greater than 1, \( n \text{Var}(\hat{\rho}_n(k)) \to 1 + 2\rho^2(1) \). To check if an observed time series follows a MA(1) model, one may use these results on the estimated autocorrelation function of the data. But if the innovations \( \{\epsilon_t\} \) were only uncorrelated, (instead of pure white noise,) the inference is incorrect. The MA(1) model may be "rejected" when in fact it is the true model. On the other hand, the standard inference which assumes i.i.d. may incorrectly "accept" a model as appropriate. More generally, inference for ARMA parameter estimates based on formulas derived under the assumption of i.i.d. innovations are invalid if the innovations are only assumed uncorrelated.

In this paper we consider the distribution of \( \hat{\rho}_n(k) \) when the process is a weakly dependent stationary sequence. This includes linear processes like (1.1), but with innovations that are only uncorrelated. Many nonlinear processes are included as well. In Section 2 we derive results for \( \hat{\rho}_n(1) \), and it is shown that the standard theory can be very misleading. In Section 3 the limiting joint distribution of the first \( k \) sample autocorrelations is derived, and examples with statistics that are functions of the \( \hat{\rho}_n(k) \) are presented. In Section 4 we propose the use of some nonparametric computer intensive approaches and apply them to some data. Concluding remarks are given in the last section.
2. The Asymptotic Distribution of the First Sample Autocorrelation

In this section we focus on results for \( \hat{\rho}_n(1) \) under weak dependence conditions for the underlying process. For a stationary time series \( X = \{X_n, n \in \mathbb{Z}\} \), define Rosenblatt’s \( \alpha \)-mixing coefficient by \( \alpha_X(j) = \sup_{A,B} |P(AB) - P(A)P(B)| \), where \( A \) and \( B \) vary over events in the \( \sigma \)-fields generated by \( \{X_n, n \leq 0\} \) and \( \{X_n, n \geq j\} \), respectively. The sequence \( X \) is said to be \( \alpha \)-mixing if \( \alpha_X(j) \to 0 \) as \( j \to \infty \).

**Theorem 2.1.** Suppose \( \{X_n, n \in \mathbb{Z}\} \) is a stationary process with mean \( \mu, R(0) > 0 \), and \( \rho(1) = 0 \).

(i). If, for some \( \delta > 0 \), \( E|X_i|^{4+\delta} < \infty \) and \( \sum_j [\alpha_X(j)]^{\delta/(2+\delta)} < \infty \), then \( n^{1/2} \hat{\rho}_n(1) \) is asymptotically normal with mean 0 and variance

\[
\tau^2 \equiv R^{-2}(0)[\text{Var}(X_1X_2) + 2 \sum_{i=1}^{\infty} \text{Cov}(X_1X_2, X_{i+1}X_{i+2})].
\]  

(ii). The same conclusion holds if the \( X_i \) are assumed essentially bounded with mixing coefficients satisfying \( \sum_j \alpha_X(j) < \infty \).

**Proof of Theorem 2.1.** Without loss of generality, assume \( \mu = 0 \). First, we consider \( n^{1/2} \tilde{R}_n(1) \), which is equal to \( n^{1/2} \sum_{i=1}^{n-1} Y_i/n \), where \( Y_i = X_i \cdot X_{i+1} \). The \( Y_i \) form a zero mean, stationary sequence with mixing coefficients satisfying \( \alpha_Y(j) \leq \alpha_X(j-1) \). Under the hypotheses of part (i), \( E|Y_i|^{2+\delta} < \infty \), and \( \sum_j [\alpha_Y(j)]^{\delta/(2+\delta)} < \infty \). Hence, we can apply Ibragimov’s (1962, Theorem 1.7) central limit theorem for stationary strongly mixing sequences to conclude \( n^{1/2} \tilde{R}_n(1) \) is asymptotically normal with mean 0 and variance given by \( \tau^2 \cdot [R(0)]^2 \), where \( \tau^2 \) is given by (2.1). Then,

\[
n^{1/2}[\tilde{R}_n(1) - \tilde{R}_n(1)] = n^{1/2} \tilde{X}_n^2 - n^{1/2} \tilde{X}_n[\sum_{i=1}^{n-1} (X_i + X_{i+1})/n].
\]  

But, by Ibragimov’s central limit theorem, \( n^{1/2} \tilde{X}_n \) is asymptotically normal (and so \( \tilde{X}_n \to 0 \) in probability as well ), so that (2.2) tends to 0 in probability. Thus, \( n^{1/2} \tilde{R}_n(1) \) and \( n^{1/2} \tilde{R}_n(1) \) have the same asymptotic distribution. By a similar argument, \( n^{1/2} [\tilde{R}_n(0) - R(0)] \) and \( n^{1/2} [\tilde{R}_n(0) - R(0)] \) have the same asymptotically normal distribution. But, all
that is needed now is that \( \hat{R}_n(0) \) tends to \( R(0) \) in probability. The result of Theorem 2.1 follows by Slutsky's theorem. The proof of (ii) is similar; see Corollary 9 of Carlstein (1986). A similar argument holds to prove the asymptotic normality of \( n^{1/2}[\hat{R}_n(0) - R(0)] \). In particular, \( \hat{R}_n(0) \) converges in probability to \( R(0) \).

Some technical remarks regarding Theorem 2.1 are in order.

(i) The assumption that \(|X_i|\) has \( 4 + 2\delta \) moments can be replaced by the assumptions that \( X_i \) has 4 moments, as long as \( E|X_1X_2|^{2+\delta} < \infty \).

(ii) If the underlying process is a strictly white noise sequence, i.e., a sequence of independent and identically distributed random variables, then only \( 2 + \delta \) moments are needed for the conclusion to hold. In this case, \( \tau^2 = 1 \).

(iii) The asymptotic joint distribution of the autocovariance estimates and autocorrelation estimates have been considered by many authors, though typically under the assumption that the underlying process is a linear process constructed from i.i.d. innovations. Hannan and Heyde (1972) relaxed this assumption so that innovations form a martingale difference sequence. Anderson (1991) further relaxed the moment conditions. In contrast, our conditions of weak dependence (and a moment condition) do not rely on any particular structure in the model generating the series. For example, our Example 2.2 below does not appear to be covered by previous results.

(iv) The conditions used throughout this paper can be relaxed further. In particular, in Appendix A we show that the assumption of strict stationarity can be weakened considerably.

Example 2.1. Uncorrelated Sequence, \( \tau^2 > 1 \). We now construct a sequence \( \{X_n\} \) satisfying the conditions of the theorem for which \( \tau^2 > 1 \). Let \( \{Z_i\} \) be any sequence of independent, identically distributed random variables having mean 0 and finite fourth moment. For fixed \( m \), let \( X_i = \prod_{k=0}^{m} Z_{i-k} \). Then, the \( X_i \) form an \( m \)-dependent sequence; that is, the mixing coefficients of the \( X_i \) sequence satisfies \( \alpha_X(j) = 0 \) if \( j > m \), and the conditions of Theorem 2.1(i) are satisfied. Moreover, the \( X_i \) form an uncorrelated sequence. To calculate \( \tau^2 \), note that for any \( i > 1 \), \( \text{Cov}(X_1X_2, X_{i+1}X_{i+2}) = 0 \). Also,
letting \( \sigma^2 \) denote the variance of \( Z_i \),

\[
\text{Var}(X_1 X_2) = E(X_1^2 X_2^2) = [E(Z_0^2)]^2 \cdot \prod_{k=1}^{m} E(Z_k^4) = \sigma^4 \cdot [E(Z_i^4)]^m.
\]

Since \( R(0) = (\sigma)^2(m+1) \), we have

\[
\tau^2 = [E(Z_i^4)/\sigma^4]^m.
\]

Note that it is always the case that \( \tau^2 > 1 \). In particular, if the \( Z_i \) are standard normal, then \( \tau^2 = 3^m \). In fact, \( \tau^2 \) can be made arbitrarily large even if \( m = 1 \), yielding a series which is one-dependent and uncorrelated.

The inferential consequences of this result should be clear. If \( \tau^2 \) is large, a test for zero correlation based on the standard normal approximation to the distribution of \( n^{1/2} \hat{\rho}_n(1) \) will lead to a high probability of incorrectly rejecting the hypothesis of zero correlation. Moreover, as \( \tau^2 \to \infty \), this rejection probability tends to one. That is, we can construct a one-dependent uncorrelated stationary sequence for which using the asymptotic standard normal approximation to the distribution of \( n^{1/2} \hat{\rho}_n(1) \) is arbitrarily poor.

Even though Theorem 2.1 only applies when \( \rho(1) \) is zero, the result has important implications for time series modeling. Consider the case of fitting autoregressive models. The usual formulae for standard errors of parameter estimates derived under the assumption of i.i.d. innovations or errors become invalid if the actual errors are only assumed uncorrelated. Indeed, for a simple autoregressive process of order one represented as \( X_t = \phi X_{t-1} + \epsilon_t \), the standard estimate of \( \phi \) is \( \hat{\rho}_n(1) \). The usual formula for the asymptotic variance of this estimate is \( 1 - \phi^2 \). The previous example shows that this formula can be way off even in the case \( \phi = 0 \).

**Example 2.2. Uncorrelated Sequence, \( \tau^2 < 1 \).** In this example, we construct uncorrelated sequences for which \( \tau^2 < 1 \). Define \( \{Z_i\} \) to be a sequence of independent, identically distributed mean 0 random variables having finite fourth moments. Also, assume \( E|Z_i^{-2}| < \infty \). Let \( X_i = Z_i/Z_{i+1} \) and so the \( X_i \) form a one-dependent, mean zero, uncorrelated stationary sequence. To calculate \( \tau^2 \), note again that all the covariance terms in (2.1) vanish. Here,

\[
\text{Var}(X_1 X_2) = \text{Var}(Z_1/Z_3) = E(Z_1^2)E(1/Z_1^2) = \text{Var}(X_1),
\]
so that

$$\tau^2 = [\text{Var}(X_1)]^{-1} = [E(1/Z_1^2) \cdot E(Z_1^2)]^{-1}. \quad (2.3)$$

By Jensen's inequality, it is always the case that $E(1/Z^2) \geq 1/E(Z^2)$, implying that $\tau^2 \leq 1$. It should be clear that the proper choice for the distribution of $Z_i$ can result in $\tau^2$ being arbitrarily near zero. An easy calculation shows that by letting $Z_i$ take on the four values $\pm 1$ and $\pm c$ with equal probabilities, the corresponding $\tau^2$ decreases from one to zero as $c$ increases from one to $\infty$. 
3. The Joint Asymptotic Distribution of the Sample Autocorrelations

Before deriving the joint asymptotic distribution of the autocorrelation estimates, we first deal with autocovariance estimates. In the following theorems, $\kappa(s, r, v)$ denotes the fourth joint cumulant of the distribution of $(X_j, X_{j+r}, X_{j+s}, X_{j+r+s+v})$; see (5.3.19) of Priestley (1981).

**Theorem 3.1.** Suppose $X_1, \ldots, X_n$ is a sample from a stationary mean zero process and $R(0) > 0$. Assume, for some $\delta > 0$, $E|X_i|^{4+\delta} < \infty$ and $\sum_j [\alpha_X(j)]^{\delta/(2+\delta)} < \infty$. Then, for any fixed nonnegative integer $k$,

$$n^{1/2} \{\hat{R}_n(0) - R(0), \hat{R}_n(1) - R(1), \ldots, \hat{R}_n(k) - R(k)\}$$

is asymptotically normal with mean 0 and covariance matrix $C$, where $C$ is a matrix of size $(k + 1) \times (k + 1)$ with the $(i + 1, j + 1)^{th}$ entry $c_{i+1, j+1}$ defined by

$$c_{i+1, j+1} \equiv \lim_{n \to \infty} \{nCov[\hat{R}_n(i), \hat{R}_n(j)]\}$$

$$\equiv \sum_{d = -\infty}^{\infty} \left[ R(d)R(d + j - i) + R(d + j)R(d - i) + \kappa(d, i, j - i) \right] \quad (3.1)$$

$$= \sum_{d = -\infty}^{\infty} Cov(X_d X_i, X_d X_{d+j}).$$

**Proof of Theorem 3.1.** As in the proof of Theorem 2.1, $\hat{R}(j)$ is an average of weakly dependent stationary random variables, so Ibragimov's central limit theorem still applies. As in the proof of Theorem 2.1, $\hat{R}_n(j)$ and $\check{R}_n(j)$ have the same limit laws (even if we consider several $j$'s jointly at a time). The joint asymptotic behavior follows by considering arbitrary linear combinations. The covariance calculation is easy; see (5.3.21) of Priestley (1981).

**Theorem 3.2.** Suppose $X_1, \ldots, X_n$ is a sample from a stationary mean zero process and $R(0) > 0$. Assume, for some $\delta > 0$, $E|X_i|^{4+\delta} < \infty$ and $\sum_j [\alpha_X(j)]^{\delta/(2+\delta)} < \infty$. Then, for any fixed nonnegative integer $k$,

$$n^{1/2} \{\hat{\rho}(0) - \rho(0), \hat{\rho}(1) - \rho(1), \ldots, \hat{\rho}(k) - \rho(k)\}$$
is asymptotically normal with mean 0 and covariance matrix \( T \), where \( T \) is a matrix of size \((k + 1) \times (k + 1)\) with the \((i + 1, j + 1)^{th}\) entry given by

\[
\tau_{i+1,j+1} \equiv \lim_{n \to \infty} \{ n \text{Cov} \{ \hat{\rho}_n(i), \hat{\rho}_n(j) \} \}
\]

\[
= \lim_{n \to \infty} nR^{-2}(0) \text{Cov} \{ \hat{\rho}_n(i) - \rho(i)\hat{\rho}_n(0), \hat{\rho}_n(j) - \rho(j)\hat{\rho}_n(0) \}
\]

\[
= R^{-2}(0) [c_{i+1,j+1} - \rho(i)c_{1,j+1} - \rho(j)c_{1,i+1} + \rho(i)\rho(j)c_{1,1}], \tag{3.2}
\]

where \( c_{i,j} \) is defined by (3.1). In matrix notation, \( T = DCD' \), where \( D \) is the \((k + 1) \times (k + 1)\) matrix

\[
D = R^{-1}(0) \begin{pmatrix}
1 - \rho(0) & 0 & 0 & \ldots & 0 \\
-\rho(1) & 1 & 0 & \ldots & 0 \\
-\rho(2) & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-\rho(k) & 0 & 0 & \ldots & 1
\end{pmatrix}.
\]

**Proof of Theorem 3.2.** It is easy to check that

\[
n^{1/2}[\hat{\rho}_n(j) - \rho(j)] =
\]

\[
R^{-1}(0) [n^{1/2}(\hat{\rho}_n(j) - R(j))] - \rho(j)R^{-1}(0)[n^{1/2}(\hat{\rho}_n(0) - R(0))] + o_P(1), \tag{3.3}
\]

where \( \hat{\rho}_n(j) = \hat{\rho}_n(j)/\hat{\rho}_n(0) \). Hence, the joint asymptotic normality as claimed follows immediately from Theorem 3.1 and the fact that the joint asymptotic distribution of the \( \hat{\rho}_n(j) \) is the same as that of the \( \hat{\rho}_n(j) \). The covariance calculation follows immediately from (3.3) and (3.1).

**Example 3.1.** *Moving Average Process of Order One.* Consider the following standard representation of a moving average process of order one:

\[
X_t = \epsilon_t + \theta \epsilon_{t-1}. \tag{3.4}
\]

Usually it is assumed that the \( \epsilon_t \)'s are an i.i.d. sequence with mean zero. Here, we consider the effect on the asymptotic distribution of \( n^{1/2}[\hat{\rho}_n(1) - \rho(1)] \) when the \( \epsilon_t \) form an uncorrelated sequence, but are one-dependent and strictly stationary. For the example, we recall the construction of Example 2.1 with \( m = 1 \). Let \( \epsilon_t = Z_t \cdot Z_{t-1} \), where the \( Z_t \) sequence is an i.i.d. sequence of mean 0 random variables with \( E(Z_t^2) = \sigma^2 \) and \( E(Z_t^4) = k\sigma^4 \). Since
\[ \hat{\rho}_n(1) = \hat{R}_n(1)/\hat{R}_n(0), \] we need terms \( c_{i,j} \) for \( i, j = 1, 2 \). First, we compute \( c_{1,1} \) defined by equation (3.1). In this case,

\[ c_{1,1} = E(X_0^4) + 2E(X_0^2X_1^2) + 2E(X_0^2X_2^2) - 5[E(X_0^2)]^2; \]

and the reader may check that (for \( \sigma = 1 \))

\[ E(X_1^4) = k^2(1 + \theta^4) + 6\theta^2 k, \]
\[ E(X_0^2X_1^2) = \theta^2 k^2 + (1 + \theta^4)k + \theta^2, \]
\[ E(X_0^2X_2^2) = k\theta^2 + 1 + \theta^2 + \theta^4, \]
\[ E(X_0^2) = 1 + \theta^2. \]

Combining these results gives

\[ c_{1,1} = k^2(1 + \theta^2)^2 + 2k(\theta^4 + 4\theta^2 + 1) - 3(1 + \theta^2)^2. \]

Similarly, one finds that

\[ c_{2,2} = k^2\theta^2 + k(1 + 4\theta^2 + \theta^4) - 2\theta^2 \]

and

\[ c_{1,2} = (\theta^3 + \theta)(k^2 + 4k - 3). \]

Thus, the asymptotic variance of \( n^{1/2}\hat{\rho}_n(1) \) is

\[ \tau_{2,2} = R^{-2}(0)[c_{2,2} - 2\rho(1)c_{1,2} + \rho^2(1)c_{1,1}] \]
\[ = (1 + \theta^2)^{-4} \cdot [k + \theta^2 + (2k + 2)\theta^4 + \theta^6 + k\theta^8]. \tag{3.5} \]

As we expect, in the case the \( Z_t \) sequence is i.i.d. standard normal, \( k = 3 \), and \( \tau_{2,2} = 1 \) if \( \theta = 1 \). If \( k = 3 \) and \( \theta = 2 \), then (3.5) reduces to 972/625. In contrast, the usual textbook formula (due to Bartlett (1946)) for the asymptotic variance of the first autocorrelation estimate in the case of an MA(1) model is \( 1 - 3\rho^2(1) + 4\rho^4(1) \), where \( \rho(1) = \theta/(1 + \theta^2) \). In fact, the formula \( 1 - 3\rho^2(1) + 4\rho^4(1) \) is identically equal to (3.5) in (and only in) the case \( k = 1 \), which is true only if \( Z_t^2 \) is constant with probability one. So this formula is correct only if the \( e_t \) sequence in (3.4) is an i.i.d. sequence, or if \( k = 1 \). In the case \( \theta = 1 \), the textbook formula reduces to 0.5, though the correct answer given by (3.5) is 1. Indeed,
our message here is that the usual inference can be completely misleading. Notice that
the textbook formula is never greater than one, but the true asymptotic variance given by
(3.5) can be arbitrarily large! Also, notice (3.5) is always strictly greater than the textbook
formula if \( k > 1 \).

**Example 3.2. Joint distribution for an \( m \)-dependent sequence.** To work out the joint
asymptotic distribution of the autocovariance estimates for Example 2.1, one can check
that \( c_{i+1,i+1} \) defined by (3.1) is always zero if \( i \) differs from \( j \). Otherwise,

\[
c_{i+1,i+1} = \sum_{d=-\infty}^{\infty} \text{Cov}(X_0 X_i, X_d X_{d+i}) = E(X_0^2 X_i^2).
\]

If \( i > m \), this reduces to \( (\sigma^4)^{m+1} \). For \( 0 \leq i \leq m \),

\[
c_{i+1,i+1} = (\sigma^4)^{m+1} \kappa^{m-i+1},
\]

where \( \kappa \) is defined by \( E(Z_i^4) = \kappa \sigma^4 \).

**Example 3.3. Dependency structure for \( \hat{R}_n(i) \) and \( \hat{R}_n(j) \).** For a strictly white noise
sequence, the distribution of \( \hat{R}_n(i) \) and \( \hat{R}_n(j) \) are asymptotically independent if \( i \) differs
from \( j \). The point of this example is to show that, if the underlying sequence is just
uncorrelated and one-dependent, then \( \hat{R}_n(i) \) and \( \hat{R}_n(j) \) may be asymptotically dependent.
For the example, recall the construction of Example 2.2. Then, one can check that in the
covariance term for \( \hat{R}_n(i) \) and \( \hat{R}_n(j) \),

\[
c_{2,3} = \sum_{d=-\infty}^{\infty} \text{Cov}(X_0 X_1, X_d X_{d+2}) = E(X_0^2 X_1 X_2) = E(Z_0^2) / [E(Z_0)]^2,
\]

which clearly is nonzero.

**Example 3.4. MA(1) model with \( \epsilon_t = Z_t / Z_{t+1} \).** As in example 3.1, suppose the \( X_t \) satisfy
(3.4), but suppose \( \epsilon_t = Z_t / Z_{t+1} \). Based on example 2.2 (which only deals with the case
\( \theta = 0 \)), it is reasonable to expect that the asymptotic distribution of \( n^{1/2} [\hat{\rho}_n(1) - \rho(1)] \) is
normal with mean zero, but this time with asymptotic variance less than or equal to (and
sometimes less depending on the value of (2.3)) the textbook formula \( 1 - 3\rho^2(1) + 4\rho^4(1) \).
Hence, the usual formula yields confidence intervals for \( \rho(1) \) which are too big. This can,
for example, lead to the incorrect conclusion that $\rho(1) = 0$ when in fact $\rho(1)$ differs from zero.

**Example 3.5. Chi-squared test of correlation structure.** A simple test, proposed by Box and Pierce (1970), to test the correlation structure in a series is to compute

$$T_n \equiv n \sum_{i=1}^{\nu} [\hat{\rho}_n(i)]^2.$$ 

By simultaneously considering the estimated autocorrelations at lags 1 to $\nu$, $T_n$ attempts to detect general departures from zero correlation. Under the assumptions of Theorem 3.2 and the further assumption that $\rho(i) = 0$ for $i = 1, \ldots, \nu$, $n^{1/2}[\hat{\rho}_n(1), \ldots, \hat{\rho}_n(\nu)]$ converges in distribution to $(W_1, \ldots, W_{\nu})$, where $(W_1, \ldots, W_{\nu})$ is multivariate normal with mean vector 0 and covariance matrix with $(i, j)$ component given by $\tau_{i+1, j+1}$ defined by (3.2). It follows that $T_n$ converges weakly to the distribution of $\sum_{i=1}^{\nu} W_i^2$. If the underlying series is strictly white noise, then this limit law is the Chi-squared distribution with $\nu$ degrees of freedom. We now show this limit result does not hold if the underlying series is uncorrelated and even just one-dependent. In the situation of example 3.2 with $m < \nu$, $n^{1/2} \hat{\rho}_n(i)$ is asymptotically distributed as $\kappa^{(m-i+1)/2} Z_i$ for $1 \leq i \leq m$ and $Z_i$ for $i > m$, where $Z_1, \ldots, Z_{\nu}$ are i.i.d. standard normal. Hence, $T_n$ converges in distribution to a weighted sum of independent Chi-squared variables given by

$$\sum_{i=1}^{m} \kappa^{m-i+1} Z_i^2 + \sum_{j=m+1}^{\nu} Z_j^2.$$ 

Depending on the choice of $\kappa$, the usual critical value obtained from the Chi-squared distribution with $\nu$ degrees of freedom is wrong. Using this critical value would result in a probability of Type I error greater than the nominal level. In fact, as $\kappa$ gets large and increases to infinity, the resulting rejection probability increases to one.
4. Nonparametric Computer-intensive Approaches

In this section we propose the use of resampling and subsampling schemes in the construction of confidence intervals and testing for $\rho(k)$.

There are two approaches for resampling dependent data indexed by time. The model dependent methods fit a time series model to the data and then resample the residuals. The replicate is generated using these residuals and the estimated parameters of the fitted model. As an example, consider the AR($p$) model. The pseudo data would be generated by first setting $X_1^* = X_1, \ldots, X_p^* = X_p$ and then using the recursive formula,

$$X_j^* = \hat{\phi}_1 X_{j-1}^* + \cdots + \hat{\phi}_p X_{j-p}^* + \hat{\epsilon}_j^*, \quad j = p+1, \ldots, n. \tag{4.1}$$

In (4.1) the $\hat{\phi}_i$ are the estimated parameters from the observed time series and the $\hat{\epsilon}_j^*$ are iid random draws from the set of residuals. This method has proven useful when estimating the sampling distribution of the parameter estimates, (Efron and Tibshirani (1986),) and also in the forecast setting, (Thombs and Schucany (1990), Stine (1987), and Findley (1986).) However, a method that does not require knowledge of a model or assume the innovations are pure white noise is more suited in this setting of inference for $\rho(k)$.

The second resampling approach is a fully nonparametric one in which no structured time series model is assumed. We review two such resampling schemes, the moving blocks bootstrap and the stationary bootstrap. We also review an alternative subsamples approach, which is also computer-intensive.

4.1 The Moving Blocks Bootstrap

Let $X_1, \ldots, X_n$ be the observed record of data from a stationary time series. The moving blocks approach (Kunsch (1989) and Liu and Singh (1992)) resamples blocks from the observed time series and then concatenates them into one long vector to produce the replicate. Denote the block size by $b$, where the sample size $n = kb$. Let $B_i$ be the block of $b$ consecutive observations starting with $X_i$; that is, $B_i = (X_i, \ldots, X_{i+b-1})$, where $i = 1, \ldots, q$ and $q = n - b + 1$. Sampling with replacement $k$ times from the set $\{B_1, \ldots, B_q\}$ produces a set of blocks $(\xi_1, \ldots, \xi_k)$ which are then “glued together” to form the bootstrap replicate $X^* = (X_1^*, \ldots, X_n^*)$. 
This process is repeated a large number (say \( M \)) times and the statistic of interest is then recomputed from the \( i^{th} \) pseudo time series. In this setting, we have,

\[
\hat{\rho}^*_{n,i}(k) = \frac{\sum_{j=1}^{n-k}(X_j^{*i} - \overline{X}^{*i})(X_{j+k}^{*i} - \overline{X}^{*i})}{\sum_{j=1}^{n}(X_j^{*i} - \overline{X}^{*i})^2} \quad i = 1, \ldots, M. \tag{4.2}
\]

The idea behind the bootstrap is to estimate the true distribution of \( n^{1/2}[\hat{\rho}_n(k) - \rho(k)] \) by the sampling distribution (conditional on the data) of \( n^{1/2}[\hat{\rho}^*_{n,i}(k) - \hat{\rho}_n(k)] \). Let \( C_1 \) and \( C_2 \) denote percentiles of the bootstrap distribution of \( n^{1/2}[\hat{\rho}^*_{n,i}(k) - \hat{\rho}_n(k)] \) satisfying

\[
P^*(C_1 \leq n^{1/2}[\hat{\rho}^*_{n,i}(k) - \hat{\rho}_n(k)] \leq C_2) = 1 - \alpha. \tag{4.3}
\]

Here, \( P^* \) refers to the bootstrap resampling conditional on the original data; that is, \((C_1, C_2)\) contains \( 100(1 - \alpha)\% \) of the probability mass of the empirical distribution of the \( M \) values \( n^{1/2}[\hat{\rho}^*_{n,i}(k) - \hat{\rho}_n(k)] \). A \( 100(1 - \alpha)\% \) approximate confidence interval for \( \rho(k) \) is then given by

\[
[ \hat{\rho}_n(k) - C_2 / \sqrt{n} , \hat{\rho}_n(k) + C_1 / \sqrt{n} ]. \tag{4.4}
\]

Alternatively, we may use the bootstrap to construct intervals centered around the estimate \( \hat{\rho}_n(k) \) by approximating the distribution of \( n^{1/2}|\hat{\rho}_n(k) - \rho(k)| \) by the empirical distribution of the \( M \) values \( n^{1/2}|\hat{\rho}^*_{n,i} - \hat{\rho}_n(k)| \), and then inverting as in the construction of (4.4).

The mathematical properties of the interval (4.4) have been studied in Künsch (1989), Liu and Singh (1992) and Politis and Romano (1992a). In particular, it is known that (under mixing and moment conditions) the asymptotic probability that the interval (4.4) contains the true value \( \rho(k) \) is the nominal level \( 1 - \alpha \). Moreover, the bootstrap can be applied to estimate the joint distribution of several \( \hat{\rho}_n(k) \)'s at a time, leading to asymptotically valid simultaneous confidence sets for the \( \rho(k) \); see Politis and Romano (1992c).

### 4.2 The Stationary Bootstrap

A drawback of the moving blocks method is that it does not produce a pseudo data sequence \( X^* \) that is stationary. To see this, consider the case where \( n = 100 \) and \( b = 10 \). Clearly the joint distribution at a "piecing together point," say for \( (X^*_{10}, X^*_{11}) \), will not be
the same as the distribution of variates within the same sampled block. The stationary bootstrap, (Politis and Romano, 1991) corrects this problem and thus produces replicates that are stationary. In this method, blocks of random length are drawn, where the length of each block has a geometric distribution.

One begins the stationary bootstrap algorithm by choosing a single observation at random from the observed time series; that is, $X^*_1 = X_{I_1}$, where $I_1$ has the discrete Uniform $\{1, \ldots, N\}$ distribution. With probability $(1 - p)$ the next observation of the replicate is the next observation in the original data, $X^*_2 = X_{I_1+1}$. Otherwise, ( with probability $p$ ) this first block would only be of length 1 and a new starting point $I_2$ would determine the next value of $X^*$. In general, given that $X^*_t$ is determined by the $J^{th}$ observation $X_J$, the next value $X^*_{t+1}$ is equal to $X_{J+t}$ with probability $(1 - p)$ or picked at random from the original $n$ observations with probability $p$. The process continues until the pseudo time series has length $n$.

Choosing the design parameter $p$ is similar to choosing the blocksize $b$ in the moving blocks approach. Since the average blocksize is $1/p$, this suggests $b \approx 1/p$, but it has been shown that in the case of variance estimation, the stationary bootstrap is much less sensitive to the choice of $b$ than the moving blocks is to the choice of $b$.

Confidence intervals are constructed as outlined above in (4.1)-(4.4). The mathematical properties of the stationary bootstrap are developed in Politis and Romano (1991).

4.3 The Subsamples Method

Finally, a third method based on subsamples has been proposed by Politis and Romano (1992b.) As in the moving blocks bootstrap, fix a block size $b$. For each block $B_i, i = 1, \ldots, q$, recompute the statistic $\hat{\rho}_{b,i}(k)$ for the data set $B_i$. The subsampling approximation to the distribution of $n^{1/2}[\hat{\rho}_n(k) - \rho(k)]$ is the empirical distribution of the $q$ values $b^{1/2}[\hat{\rho}_{b,i}(k) - \hat{\rho}_n(k)]$. Consistency properties of this method are established under much weaker assumptions than the previous two methods; see Politis and Romano (1992b).

4.4 Examples and Simulations

We next consider some examples which illustrate how the usual inference which assumes independence can fail, and it is shown that the proposed bootstrap methods work
well. Only a portion of the simulations ran are reported. We have not attempted to optimize in any way the choice of block size or the choice of \( p \) in the stationary bootstrap method. Based on simulation experience, we applied what we thought were reasonable choices, though future choices will inevitably be data-dependent with the hope of choosing the parameters optimally. In spite of not choosing these design parameters in any optimal fashion, the simulations do support the use of computer-intensive methods for the construction of inferential procedures without having to assume a particular model structure.

First, we illustrate a special case of the situation described in Example 2.1. Let \( X_i = Z_i Z_{i-1} \), where \( \{Z_i\} \) is a sequence of independent, identically distributed standard normal random variables. The \( X_i \) sequence is uncorrelated, with \( \rho(k) = 0 \) for all \( k \), but the \( X_i \) are not independent. Moreover, the \( X_i \) are only one-dependent, so that observations separated by more than one time unit are independent of each other. In the simulations, we have used \( n = 1000 \) primarily to exhibit that the effects we observe are not small sample effects; that is, the usual methods are wrong even with large amounts of data. Moreover, the proposed computer-intensive methods are really only justified for \( n \) large, though large data sets are often available in time series studies. (Our basic philosophy is to develop valid methods that work without imposing model structure on the time series, and we are willing to settle for large sample methods as there are essentially no competing methods.)

Figure 4.1 shows a plot of the first 100 values taken from the larger set of \( n = 1000 \) generated data. The usual \( \pm 2/\sqrt{n} \) limits superimposed on the estimated autocorrelation function indicate that the series demonstrates a lack of correlation structure. For the data set, the first autocorrelation \( \hat{\rho}_{1000}(1) = -0.063 \). In fact the usual test of \( H_0 : \rho(1) = 0 \) results in a \( p \)-value of 0.045, again a misleading conclusion.

We applied all three resampling approaches to the problem of constructing confidence intervals for \( \rho(1) \) (and hence the problem of constructing a hypothesis test of \( \rho(1) = 0 \)). For the sake of concreteness, all intervals were based on estimating the distribution of \( n^{1/2} |\hat{\rho}_n(1) - \rho(1)| \), so that the resulting intervals are symmetric about \( \hat{\rho}_n(1) \); this was described in subsection 4.1 above.

In Figure 2, the stationary bootstrap approximation to the distribution of \( n^{1/2} |\hat{\rho}_n(1) - \rho(1)| \) is given for the data of Figure 1. The variance of the histogram in Figure 2 is 2.65, close to the true asymptotic variance for this situation of 3. Using this estimated
distribution results in a p-value of 0.22 for testing the hypothesis that \( \rho(1) = 0 \), and so we would accept this hypothesis for this data set (and make the correct decision). The subsampling distribution is displayed in Figure 3. The variance of this distribution is 1.96, and the corresponding p-value using this estimated distribution is 0.19. Similarly, the moving blocks distribution is given in Figure 4. It has an estimated variance of 2.83, and the corresponding p-value is 0.25.

In order to assess how well each of the three computer-intensive methods works over many data sets, as well as the method which uses a standard normal approximation to the distribution of \( n^{1/2} \hat{\rho}_n(1) \), 200 such data sets were generated. For each data set, a p-value for testing \( \rho(1) = 0 \) is computed for each of the four methods. If the estimated sampling distributions serve as good approximations to the true sampling distribution, the 200 p-values for each method should be uniformly distributed. Figure 5 displays a Q-Q plot of the 200 p-values versus uniform quantiles in the case when the p-values were constructed from the simple normal approximation. Obviously, the normal approximation is bad here. In fact, in 53 out of the 200 simulated data sets, the null hypothesis was rejected at the 5 percent level. Any reasonable test of uniformity would reject the hypothesis that the 200 p-values come from a uniform distribution.

An analogous plot is given in Figure 6 for the stationary bootstrap p-values, and we see that the method is doing a reasonable job here. At the 5 percent level, 14 out of the 200 simulations resulted in rejecting the null hypothesis, leading to an estimated rejection probability of 0.07 (which is close to the nominal level 0.05 and within binomial variation). In fact, a simple Chi-squared test of uniformity (with 20 cells) would accept the hypothesis of uniformity (and results in a p-value of 0.2). The corresponding Q-Q plots for the subsampling method and the moving blocks method are given in Figures 7 and 8, respectively. The estimated rejection probabilities at the 5 percent level are 0.095 and 0.055, respectively. The overall Chi-squared tests of uniformity result in p-values of 0.09 and 0.91, respectively. Thus, the moving blocks method performs admirably in this situation.

In Figure 9, a plot of the first 100 observations in a series of \( n = 1000 \) observations is displayed; here, the observations were generated by (3.4) with \( \theta = 0.5, \epsilon_t = Z_t Z_{t-1} \) and the \( Z_t \) are i.i.d. standard normal. The actual value of \( \rho(1) \) is \( \theta/(1 + \theta^2) = 0.4 \). In this case, the exact asymptotic variance of \( \hat{\rho}_n(1) \) given by (3.5) reduces to 1.5472. In contrast,
the textbook formula which assumes i.i.d. innovations gives a wrong asymptotic variance of 0.6224. For the data set of Figure 9, the observed value of \( \hat{\rho}_{1000}(1) \) is 0.35. The normal approximation based on an asymptotic variance of 0.6224 would lead one to doubt the hypothesis the true value of \( \rho(1) = 0.4 \).

In Figure 10, the stationary bootstrap approximation to the distribution of \( n^{1/2}[\hat{\rho}_n(1) - \rho(1)] \) is displayed. It has a variance of 1.72. A test of \( \rho = 0.4 \) would be accepted as the stationary bootstrap distribution leads to a p-value of 0.27. In Figures 11 and 12, the analogous plots of the subsampling approximation and the moving blocks approximation are displayed, respectively. The subsampling distribution estimates the asymptotic variance by 1.16, while the moving blocks distribution estimates the asymptotic variance by 1.9.
5. Conclusions

In this paper, we have derived the asymptotic distributions of the sample autocovariances and sample autocorrelations from an observed weakly dependent stationary time series. In particular, our limit results hold for weakly dependent stationary time series without imposing model structure such as the process is linear with i.i.d. innovations. Indeed, our results do apply to linear processes where the innovations are only assumed uncorrelated, and they apply to nonlinear processes as well. In general, we have found that certain basic limit results for $\hat{\rho}_n(k)$ found in many standard time series textbooks should not be applied unless the underlying assumptions behind them are met exactly. Otherwise, the resulting inferences can be completely misleading.

Our analysis has focused on the construction of tests and confidence regions for correlation parameters. For example, the often suggested method for testing lack of correlation is to apply a standard normal approximation to $n^{1/2} \hat{\rho}_n(k)$; we have demonstrated that such an approximation can be arbitrarily bad if we are only willing to assume the underlying process is uncorrelated. While many may claim that the distinction between an uncorrelated sequence and an i.i.d. sequence does not matter in practice, inferential tools such as asymptotic normality (even used just as guidelines) have important practical implications. Test for lack of correlation structure have important use in model building as well; indeed, such tests serve as diagnostic tools for examining the residuals of fitted models. Obviously, the formulae for asymptotic variances of ARMA parameters can be misleading as well.

We have demonstrated that the use of computer-intensive methods allows one to construct inferential procedures that are (asymptotically) valid without having to impose much model structure. Hence, we believe one should not apply usual results (also based on asymptotic arguments) unless the assumptions used to derive such forumulae are true. Of course, refinements of the computer intensive methods are needed. For example, the choice of block size in the moving blocks method or subsampling method and the choice of mean block size in the stationary bootstrap method needs to be addressed. Even without optimal choice of these parameters, these methods offer the potential to construct correct inferential procedures under the weakest assumptions.

Acknowledgment: The authors wish to thank Michael Wolf, a Ph.D. student at Stanford University, for the computing and figures.
APPENDIX A

The goal of this appendix is to show how one may weaken the assumption of strict stationarity in order to derive limit distributions. An assumption of weak dependence will remain in force, but we first need to modify the definition of Rosenblatt’s strong mixing coefficients so that it applies to arbitrary (possibly nonstationary) sequences. So, given a random sequence \( \{Y_j\} \), let \( F_n^m \) be the \( \sigma \)-algebra generated by \( \{Y_i, n \leq i \leq m\} \), and define the corresponding mixing sequence by

\[
\alpha_Y(k) = \sup_n \sup_{A,B} |P(AB) - P(A)P(B)|,
\]

where \( A \) and \( B \) vary over the \( \sigma \)-fields \( F_{n-\infty}^n \) and \( F_{n+k}^\infty \), respectively.

**Theorem A.1.** Let \( \{Y_j\} \) be a mean 0 sequence satisfying \( E|Y_j|^r < \Delta \) for some \( r > 2 \) and mixing sequence \( \alpha_Y(\cdot) \) satisfying \( \alpha_Y(k) = O(k^\lambda) \) for all \( \lambda < -2r/(r - 2) \). For \( n = 1, 2, \ldots \) and \( d = 0, 1, \ldots, n - 1 \), define

\[
s_{n,d} = n^{-1} \sum_{i=1}^{n-d} Cov(Y_i, Y_{i+d}).
\]

Assume, for any fixed \( d \), \( s_{n,d} \) has a limit \( s_d \) as \( n \to \infty \). Set \( \bar{Y}_n = n^{-1} \sum_{i=1}^n Y_i \). Then, \( n^{1/2}\bar{Y}_n \) is asymptotically normal with mean 0 and variance \( s_0 + 2 \sum_{i=1}^{\infty} s_i \), and the series \( \sum_i s_i \) is absolutely summable.

Before proving the theorem, the following simple lemma is needed.

**Lemma A.1.** Suppose \( t_{n,d}, n = 1, 2, \ldots \) and \( d = 1, 2, \ldots, n - 1 \), is a triangular array of real numbers satisfying, for fixed \( d \), \( t_{n,d} \) has a limit \( t_d \), \( |t_{n,d}| \leq q_d \), and \( \sum_d q_d < \infty \). Then,

\[
\sum_{i=1}^{n-1} t_{n,d} \rightarrow \sum_{i=1}^{\infty} t_d,
\]

and the last series converges absolutely.

**Proof of Theorem A.1.** Note that

\[
nVar(\bar{Y}_n) = s_{n,0} + 2 \sum_{d=1}^{n-1} s_{n,d}.
\]
Also, for any \( i \),
\[ |\text{Cov}(Y_i, Y_{i+d})| \leq 10[\alpha_Y(d)]^{(r-2)/r}(E|Y_i|^r)^{(2/r)}, \]
by the well-known mixing inequality; see Ibragimov and Linnik (1971). Hence,
\[ |s_{n,d}| \leq 10\Delta^{(2/r)}[\alpha_Y(d)]^{(r-2)/r}. \]

But, the assumptions on the mixing coefficients force \( \sum_d \alpha_Y(d) < \infty \). Thus, the conditions of Lemma A.1 are all met with \( t_{n,d} = s_{n,d} \) and \( q_d = 10\Delta^{2/r}[\alpha_Y(d)]^{(r-2)/r} \). Hence, \( n\text{Var}(Y_n) \to s_0 + 2\sum_{d=1}^{\infty} s_d \). The asymptotic normality now follows as a corollary to Theorem 5.3 of Gallant and White (1988).

We now apply Theorem A.1 to prove asymptotic normality of \( \hat{R}_n(j) \) without the assumption of strict stationarity. However, in the context of estimating \( R(j) \), it makes sense to assume the process is weakly or second order stationary (or \( R(j) \) is not even properly defined), though this could be weakened further.

**Theorem A.2.** Let \( \{X_j\} \) be a mean 0 weakly stationary sequence with autocovariance at lag \( j \) denoted \( R(j) \). Assume \( E|X_iX_{i+j}|^r < \Delta \) for some \( r > 2 \) and all \( i \), and that \( \alpha_X(k) = O(k^\lambda) \) for all \( \lambda < -2r/(r - 2) \). Let
\[ s_{n,d} = n^{-1} \sum_{i=1}^{n-d-j} \text{Cov}(X_iX_{i+j}, X_{i+d}X_{i+d+j}) \]
and assume \( s_{n,d} \) has a limit \( s_d \). Then, \( n^{1/2} [\hat{R}_n(j) - R(j)] \) is asymptotically normal with mean 0 and variance \( s_0 + 2\sum_{d=1}^{\infty} s_d \).

**Proof of Theorem A.2.** Apply Theorem A.1 in the case \( Y_i = X_iX_{i+j} \).

If the process is fourth order stationary, \( s_{n,d} \) has a limit because the covariance between \( X_iX_{i+j} \) and \( X_{i+d}X_{i+d+j} \) does not depend on \( i \). By considering linear combinations, one could derive the asymptotic joint distribution of the first \( k \) autocovariances under similar conditions, and hence one could derive the asymptotic joint distribution of the first \( k \) autocorrelations as well.
References


Figure 1: $X(t) = Z(t) \times Z(t-1)$, $n = 1000$

Figure 2: Stationary Bootstrap distribution for normalized distribution of first autocorrelation, $p = 0.025$
Figure 3: Subsampling distribution for normalized distribution of first autocorrelation, $b = 40$

Figure 4: Moving blocks distribution for normalized distribution of first autocorrelation, $b = 40$
Figure 9: $X(t) = Z(t) \times Z(t-1) + 0.5 \times Z(t-1) \times Z(t-2)$, \(n = 1000\)

Figure 10: Stationary Bootstrap distribution for normalized distribution of first autocorrelation, \(p = 0.005\)
Figure 11: Subsampling distribution for normalized distribution of first autocorrelation, $b = 40$

Figure 12: Moving blocks distribution for normalized distribution of first autocorrelation, $b = 40$