Introduction to Statistical Time Series
WAYNE A. FULLER

CHAPTER 6

Estimation of the Mean and Autocorrelations

In this chapter we shall derive some large sample results for the sampling behavior of the estimated mean, covariances, and autocorrelations.

6.1. ESTIMATION OF THE MEAN

Consider a stationary time series X_t, with mean μ, which we desire to estimate. If it were possible to obtain a number of independent realizations, then the average of the realization averages would converge in mean square to μ as the number of realizations increased. That is, given m samples of n observations each,

$$\operatorname{Var}\left\{\frac{1}{m}\sum_{j=1}^{m}\bar{x}_{(j)}\right\} \le \frac{1}{m}\gamma_X(0),$$

where $\bar{x}_{(j)} = n^{-1}\sum_{t=1}^{n} X_{(j)t}$ is the mean of the n observations from the jth realization.

However, in many areas of application, it is difficult or impossible to obtain multiple realizations. For example, most economic time series constitute a single realization. Therefore, the question becomes whether or not we can use the average of a single realization to estimate the mean. Clearly, the sample mean $\bar{x}_n$ is an unbiased estimator for the mean of a covariance stationary time series. If the mean square error of the sample mean as an estimator of the population mean approaches zero as the number of observations included in the mean increases, we say that the time series is *ergodic* for the mean. We now investigate conditions under which the time series is ergodic for the mean. Theorem 6.1.1 demonstrates that the sample mean may be a consistent estimator for nonstationary time series if the nonstationarity is of a transient nature. The theorem follows Parzen (1962).

Theorem 6.1.1. Let $\{X_t: t \in (1, 2, \ldots)\}$ be a time series satisfying

$$\lim_{t\to\infty} E\{X_t\} = \mu ,$$
$$\lim_{n\to\infty} \operatorname{Cov}\{\bar{x}_n, X_n\} = 0 ,$$

where $\bar{x}_n = n^{-1} \sum_{t=1}^{n} X_t$. Then

$$\lim_{n\to\infty} E\{(\bar{x}_n - \mu)^2\} = 0 .$$

Proof. Now

$$E\{(\bar{x}_n - \mu)^2\} = \operatorname{Var}\{\bar{x}_n\} + \left(\frac{1}{n}\sum_{t=1}^{n} E\{X_t\} - \mu\right)^2$$

where the second term on the right converges to zero by Lemma 3.1.5. Furthermore,

$$\begin{aligned}
\operatorname{Var}\{\bar{x}_n\} &= \frac{1}{n^2}\sum_{t=1}^{n}\sum_{j=1}^{n} \operatorname{Cov}\{X_t, X_j\} \\
&= \frac{2}{n^2}\sum_{t=1}^{n}\sum_{j=1}^{t} \operatorname{Cov}\{X_t, X_j\} - \frac{1}{n^2}\sum_{t=1}^{n} \operatorname{Var}\{X_t\} \\
&= \frac{2}{n^2}\sum_{t=1}^{n} t \operatorname{Cov}\{\bar{x}_t, X_t\} - \frac{1}{n^2}\sum_{t=1}^{n} \operatorname{Var}\{X_t\} \\
&\le \frac{2}{n}\sum_{t=1}^{n} |\operatorname{Cov}\{\bar{x}_t, X_t\}| ,
\end{aligned}$$

which also converges to zero by Lemma 3.1.5. ▲

Corollary 6.1.1.1. Let $\{X_t\}$ be a stationary time series whose covariance function $\gamma(h)$ converges to zero as h gets large. Then

$$\lim_{n\to\infty} \operatorname{Var}\{\bar{x}_n\} = 0 .$$

Proof. By Lemma 3.1.5, the convergence of $\gamma(h)$ to zero implies that $\operatorname{Cov}\{\bar{x}_n, X_n\} = (1/n)\sum_{h=0}^{n-1} \gamma(h)$ converges to zero, and the result follows by Theorem 6.1.1. ▲

Corollary 6.1.1.2. A stationary time series with absolutely summable covariance function is ergodic for the mean. Futhermore,

$$\lim_{n\to\infty} n \operatorname{Var}\{\bar{x}_n\} = \sum_{h=-\infty}^{\infty} \gamma(h) .$$

Proof. The assumption of absolute summability implies

$$\lim_{h\to\infty} \gamma(h) = 0,$$

and ergodicity follows from Corollary 6.1.1.1. We have

$$n \operatorname{Var}\{\bar{x}_n\} = \frac{1}{n} \sum_{j=1}^{n} \sum_{t=1}^{n} \gamma(t-j) = \frac{1}{n} \sum_{h=-(n-1)}^{n-1} (n - |h|)\gamma(h),$$

and $n \operatorname{Var}\{\bar{x}_n\}$ converges to the stated result by Lemma 3.1.4. ▲

Theorem 6.1.2. If the spectral density of a stationary time series X_t is continuous, then

$$\lim_{n\to\infty} n \operatorname{Var}\{\bar{x}_n\} = 2\pi f(0), \tag{6.1.1}$$

where $f(0)$ is the spectral density of X_t evaluated at zero.

Proof. By Theorem 3.1.10 the Fourier series of a continuous periodic function is uniformly summable by the method of Cesàro. The autocovariances $\gamma(k)$ are equal to π times the a_k of that theorem. Therefore,

$$\begin{aligned} 2\pi f(0) &= \lim_{n\to\infty} 2\pi \left[\frac{a_0}{2} + \frac{1}{n} \sum_{r=2}^{n} \sum_{k=1}^{r-1} a_k \right] \\ &= \lim_{n\to\infty} \left[\gamma(0) + \frac{2}{n} \sum_{r=2}^{n} \sum_{k=1}^{r-1} \gamma(k) \right] \\ &= \lim_{n\to\infty} \left[\gamma(0) + \frac{2}{n} \sum_{r=1}^{n} (n-r)\gamma(r) \right] \\ &= \lim_{n\to\infty} n \operatorname{Var}\{\bar{x}_n\}. \end{aligned}$$

▲

Since the absolute summability of the covariance function implies that $f(\omega)$ is continuous, it follows that (6.1.1) holds for a time series with absolutely summable covariance function. Thus, for a wide class of time series, the sample mean has a variance that is declining at the rate n^{-1}. In large samples the variance is approximately the spectral density evaluated at zero multiplied by $2\pi n^{-1}$.

To investigate the efficiency of the sample mean as an estimator of μ, we write

$$Y_t = \mu + Z_t,$$

where Z_t is a time series with zero mean, and we define $\mathbf{V}$ to be the covariance matrix for n observations on Y_t. Thus,

$$\mathbf{V} = E\{\mathbf{z}\mathbf{z}'\},$$

where $\mathbf{z}$ is the column vector of n observations on Z_t. If the covariance matrix is known and nonsingular the best linear unbiased estimator of the mean is given by

$$\hat{\mu} = (\mathbf{1}'\mathbf{V}^{-1}\mathbf{1})^{-1}\mathbf{1}'\mathbf{V}^{-1}\mathbf{y}\,, \tag{6.1.2}$$

where $\mathbf{1}$ is a column vector composed of n ones and $\mathbf{y}$ is the vector of n observations on Y_t. The variance of $\hat{\mu}$ is

$$\operatorname{Var}\{\hat{\mu}\} = (\mathbf{1}'\mathbf{V}^{-1}\mathbf{1})^{-1}\,, \tag{6.1.3}$$

whereas the variance of $\bar{y}_n = n^{-1}\sum_{t=1}^{n} Y_t$ is

$$\operatorname{Var}\{\bar{y}_n\} = n^{-2}\mathbf{1}'\mathbf{V}\mathbf{1}\,. \tag{6.1.4}$$

Let Y_t be a pth order stationary autoregressive process defined by

$$(Y_t - \mu) + \sum_{j=1}^{p} \alpha_j(Y_{t-j} - \mu) = e_t\,, \tag{6.1.5}$$

where the e_t are uncorrelated $(0, \sigma^2)$ random variables, and the roots of the characteristic equation are less than one in absolute value. For known α_j, the estimator (6.1.2) can be constructed by transforming the observations into a sequence of uncorrelated constant variance observations. Using the Gram–Schmidt orthogonalization procedure, we obtain

$$\begin{aligned}
W_1 &= \delta_{11}Y_1\,,\\
W_2 &= \delta_{22}Y_2 - \delta_{21}Y_1\,,\\
&\ \vdots\\
W_p &= \delta_{pp}Y_p - \sum_{j=1}^{p-1} \delta_{p,p-j}Y_{p-j}\,,\\
W_t &= Y_t + \sum_{j=1}^{p} \alpha_j Y_{t-j}\,, \qquad t = p+1, p+2, \ldots, n\,,
\end{aligned} \tag{6.1.6}$$

where $\delta_{11} = \gamma^{-1/2}(0)\sigma$, $\delta_{22} = \{[1 - \rho^2(1)]\gamma(0)\}^{-1/2}\sigma$, $\delta_{21} = \rho(1)\{[1 - \rho^2(1)]\gamma(0)\}^{-1/2}\sigma$, and so forth. The expected values are

$$\begin{aligned}
E\{W_1\} &= \delta_{11}\mu\,,\\
E\{W_2\} &= (\delta_{22} - \delta_{21})\mu\,,
\end{aligned}$$

$$E\{W_p\} = \left(\delta_{pp} - \sum_{j=1}^{p-1} \delta_{p,p-j}\right)\mu ,$$

$$E\{W_t\} = \left(1 + \sum_{j=1}^{p} \alpha_j\right)\mu , \qquad t = p+1, p+2, \ldots, n .$$

In matrix notation we let $\mathbf{T}$ denote the transformation defined in (6.1.6). Then $\mathbf{T}'\mathbf{T}\sigma^{-2} = V^{-1}$, $E\{\mathbf{Ty}\} = \mathbf{T1}\mu$, and

$$\hat{\mu} = (\mathbf{1}'\mathbf{T}'\mathbf{T1})^{-1}\mathbf{1}'\mathbf{T}'\mathbf{Ty} . \tag{6.1.7}$$

With the aid of this transformation, we can demonstrate that the sample mean has the same asymptotic efficiency as the best linear unbiased estimator.

Theorem 6.1.3. Let Y_t be the stationary pth order autoregressive process defined in (6.1.5). Then

$$\lim_{n\to\infty} n \operatorname{Var}\{\bar{y}_n\} = \lim_{n\to\infty} n \operatorname{Var}\{\hat{\mu}\} ,$$

where $\hat{\mu}$ is defined in (6.1.2).

Proof. Without loss of generality, we let $\sigma^2 = 1$. The spectral density of Y_t is then

$$f_Y(\omega) = \frac{1}{2\pi}\left[\sum_{j=0}^{p} \alpha_j e^{-\iota\omega j} \sum_{j=0}^{p} \alpha_j e^{\iota\omega j}\right]^{-1} ,$$

where $\alpha_0 = 1$. With the exception of $2p^2$ terms in the upper left and lower right corners of $\mathbf{T}'\mathbf{T}$, the elements of $\mathbf{V}^{-1}$ are given by

$$v^{ir} = \begin{cases} \sum_{r=0}^{p-h} \alpha_r\alpha_{r+h} , & |i-r| = h \le p , \\ 0 & \text{otherwise} . \end{cases} \tag{6.1.8}$$

The values of the elements v^{ir} in (6.1.8) depend only on $|i-r|$, and we recognize $\sum_{r=0}^{p-|h|} \alpha_r\alpha_{r+|h|} = \gamma_m(h)$, say, as the covariance function of a pth order moving average. Therefore, for $n > 2p$,

$$\begin{aligned} \frac{1}{n}\mathbf{1}'V^{-1}\mathbf{1} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{r=1}^{n} v^{ir} \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{r=1}^{n} \gamma_m(i-r) + \frac{1}{n}\sum_{i=1}^{p}\sum_{r=1}^{p} [v^{ir} - \gamma_m(1-r)] \\ &\quad + \frac{1}{n}\sum_{i=n-p+1}^{n}\sum_{r=n-p+1}^{n} [v^{ir} - \gamma_m(i-r)] \end{aligned}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sum_{r=1}^{n} \gamma_m(i-r) + O(n^{-1}) .$$

It follows that

$$\lim_{n\to\infty} n \operatorname{Var}\{\hat{\mu}\} = \lim_{n\to\infty} \left[\frac{1}{n} \sum_{i=1}^{n} \sum_{r=1}^{n} \gamma_m(i-r) \right]^{-1} = \left| \sum_{j=0}^{p} \alpha_j \right|^{-2}$$
$$= 2\pi f_Y(0) = \lim_{n\to\infty} n \operatorname{Var}\{\bar{y}_n\} . \qquad \blacktriangle$$

Using the fact that a general class of spectral densities can be approximated by the spectral density of an autoregressive process (see Theorem 4.3.4), Grenander and Rosenblatt (1957) have shown that the mean and certain other linear estimators have the same asymptotic efficiency as the generalized least squares estimator for time series with spectral densities in that class.

6.2. ESTIMATORS OF THE AUTOCOVARIANCE AND AUTOCORRELATION FUNCTIONS

While the sample mean is a natural estimator to consider for the mean of a stationary time series, a number of estimators have been proposed for the covariance function. If the mean is known and, without loss of generality, taken to be zero, the estimator

$$\tilde{\gamma}(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} X_t X_{t+h} \qquad (6.2.1)$$

is seen to be the mean of $n-h$ observations from the time series, say,

$$Z_{th} = X_t X_{t+h} .$$

For stationary time series, $E\{Z_{th}\} = \gamma_X(h)$ for all t, and it follows that $\tilde{\gamma}(h)$ is an unbiased estimator of $\gamma(h)$. In most practical situations the mean is unknown and must be estimated. We list below two possible estimators of the covariance function when the mean is estimated. In both expressions h is taken to be greater than or equal to zero:

$$\gamma^{\dagger}(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (X_t - \bar{x}_n)(X_{t+h} - \bar{x}_n) , \qquad (6.2.2)$$

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_t - \bar{x}_n)(X_{t+h} - \bar{x}_n) . \qquad (6.2.3)$$

It is clear that these estimators differ by factors that become small at the rate n^{-1}. Unlike $\tilde{\gamma}(h)$, neither of the estimators is unbiased. The bias is given in Theorem 6.2.2 of this section.

The estimator $\hat{\gamma}(h)$ can be shown to have smaller mean square error than $\gamma^{\dagger}(h)$ for certain types of time series. This estimator also has the advantage of guaranteeing positive definiteness of the estimated covariance function. In most of our applications we shall use the estimator $\hat{\gamma}(h)$.

As one might expect, the variances of the estimated autocovariances are much more complicated than those of the mean. The theorem we present is due to Bartlett (1946). A result needed in the proof will be useful in later sections, and we state it as a lemma.

Lemma 6.2.1. Let $\{\delta_j\}_{j=-\infty}^{\infty}$ and $\{c_j\}_{j=-\infty}^{\infty}$ be two absolutely summable sequences. Then, for fixed integers r, h, and d, $d \geq 0$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{s=1}^{n+d} \sum_{t=1}^{n} \delta_{s-t+r} c_{s-t+h} = \sum_{p=-\infty}^{\infty} \delta_{p+r} c_{p+h}$$
$$= \sum_{p=-\infty}^{\infty} \delta_p c_{p+h-r} .$$

Proof. Let $p = s - t$. Then

$$\begin{aligned}\frac{1}{n} \sum_{s=1}^{n+d} \sum_{t=1}^{n} \delta_{s-t+r} c_{s-t+h} &= \frac{1}{n} \sum_{p=0}^{n-1} \sum_{s=p+1}^{n} \delta_{p+r} c_{p+h} \\ &\quad + \frac{1}{n} \sum_{p=-(n-1)}^{-1} \sum_{s=1}^{n+p} \delta_{p+r} c_{p+h} \\ &\quad + \frac{1}{n} \sum_{s=n+1}^{n+d} \sum_{t=1}^{n} \delta_{s-t+r} c_{s-t+h} \\ &= \sum_{p=-(n-1)}^{n-1} \frac{(n-|p|)}{n} \delta_{p+r} c_{p+h} \\ &\quad + \frac{1}{n} \sum_{s=n+1}^{n+d} \sum_{t=1}^{n} \delta_{s-t+r} c_{s-t+h} .\end{aligned}$$

Now,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{s=n+1}^{n+d} \sum_{t=1}^{n} \delta_{s-t+r} c_{s-t+h} \leq \lim_{n\to\infty} \frac{1}{n} \sum_{p=1}^{n+d} d|\delta_{p+r}|\,|c_{p+h}| = 0$$

with the inequality resulting from the inclusion of additional terms and the introduction of absolute values. ▲

Theorem 6.2.1. Let the time series $\{X_t\}$ be defined by

$$X_t = \sum_{j=-\infty}^{\infty} \alpha_j e_{t-j} ,$$

where the sequence $\{\alpha_j\}$ is absolutely summable and the e_t are independent $(0, \sigma^2)$ random variables with $E\{e_t^4\} = \eta\sigma^4$. Then, for fixed h and q $(h \geq q \geq 0)$,

$$\lim_{n\to\infty} (n-q)\operatorname{Cov}\{\tilde{\gamma}(h), \tilde{\gamma}(q)\}$$

$$= (\eta - 3)\gamma(h)\gamma(q) + \sum_{p=-\infty}^{\infty} [\gamma(p)\gamma(p-h+q) + \gamma(p+q)\gamma(p-h)], \tag{6.2.4}$$

where $\tilde{\gamma}(h)$ is defined in (6.2.1).

Proof. Using

$$E\{e_t e_u e_v e_w\} = \begin{cases} \eta\sigma^4, & t = u = v = w, \\ \sigma^4 & \text{if subscripts are equal in pairs but not all equal} \\ 0 & \text{otherwise}, \end{cases}$$

it follows that

$$E\{X_t X_{t+h} X_{t+h+p} X_{t+h+p+q}\} = (\eta - 3)\sigma^4 \sum_{j=-\infty}^{\infty} \alpha_j \alpha_{j+h} \alpha_{j+h+p} \alpha_{j+h+p+q}$$
$$+ \gamma(h)\gamma(q) + \gamma(h+p)\gamma(p+q)$$
$$+ \gamma(h+p+q)\gamma(p). \tag{6.2.5}$$

Thus,

$$E\{\tilde{\gamma}(h)\tilde{\gamma}(q)\} - \gamma(h)\gamma(q)$$
$$= \frac{1}{(n-h)(n-q)} E\left\{\sum_{s=1}^{n-q} \sum_{t=1}^{n-h} X_t X_{t+h} X_s X_{s+q}\right\} - \gamma(h)\gamma(q)$$
$$= \frac{(\eta-3)\sigma^4}{(n-h)(n-q)} \sum_{s=1}^{n-q} \sum_{t=1}^{n-h} \sum_{j=-\infty}^{\infty} \alpha_j \alpha_{j+h} \alpha_{j+s-t} \alpha_{j+s-t+q}$$
$$+ \frac{1}{(n-h)(n-q)} \sum_{s=1}^{n-q} \sum_{t=1}^{n-h} [\gamma(s-t)\gamma(s-t-h+q)$$
$$+ \gamma(s-t+q)\gamma(s-t-h)]. \tag{6.2.6}$$

Applying Lemma 6.2.1 to equation (6.2.6), we have

$$\lim_{n\to\infty} (n-q)\operatorname{Cov}\{\tilde{\gamma}(h), \tilde{\gamma}(q)\}$$
$$= (\eta-3)\sigma^4 \sum_{j=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} \alpha_j \alpha_{j+h} \alpha_{j+p} \alpha_{j+p+q}$$
$$+ \sum_{p=-\infty}^{\infty} [\gamma(p)\gamma(p-h+q) + \gamma(p+q)\gamma(p-h)]. \qquad \blacktriangle$$

For normal time series $\eta = 3$, and we have

$$\operatorname{Cov}\{\tilde{\gamma}(h), \tilde{\gamma}(q)\} \doteq \frac{1}{n-q} \sum_{p=-\infty}^{\infty} [\gamma(p)\gamma(p-h+q) + \gamma(p+q)\gamma(p-h)] .$$

Corollary 6.2.1.1. Given a time series $\{e_t : t \in (0, \pm 1, \pm 2, \dots)\}$, where the e_t are normal independent $(0, \sigma^2)$ random variables, then for $h, q \geq 0$,

$$\operatorname{Cov}\{\tilde{\gamma}_e(h), \tilde{\gamma}_e(q)\} = \begin{cases} \dfrac{2\sigma^4}{n}, & h = q = 0, \\ \dfrac{\sigma^4}{n-h}, & h = q \neq 0, \\ 0 & \text{otherwise}. \end{cases}$$

Proof. Reserved for the reader. ▲

In Theorem 6.2.1 the estimator was constructed assuming the mean to be known. As we have mentioned, the estimation of the unknown mean introduces a bias into the estimated covariance. However, the variance formulas presented in Theorem 6.2.1 remain valid approximations for the estimator defined in equation (6.2.2).

Theorem 6.2.2. Given fixed $h \geq q \geq 0$ and a time series X_t satisfying the assumptions of Theorem 6.2.1,

$$E\{\hat{\gamma}(h) - \gamma(h)\} = -\frac{|h|}{n}\gamma(h) - \frac{n-|h|}{n}\operatorname{Var}\{\bar{x}_n\} + O(n^{-2})$$

and

$$\lim_{n\to\infty} \frac{n^2}{(n-h)} \operatorname{Cov}\{\hat{\gamma}(h), \hat{\gamma}(q)\}$$
$$= (\eta - 3)\gamma(h)\gamma(q) + \sum_{p=-\infty}^{\infty} [\gamma(p)\gamma(p-h+q) + \gamma(p+q)\gamma(p-h)] .$$

Proof. From the definition of $\hat{\gamma}(h)$, we obtain

$$\begin{aligned} \hat{\gamma}(h) &= \frac{1}{n}\left[\sum_{t=1}^{n-h} X_t X_{t+h} - \bar{x}_n \sum_{t=1}^{n-h} (X_t + X_{t+h}) + (n-h)\bar{x}_n^2\right] \\ &= \frac{n-h}{n}\tilde{\gamma}(h) + \frac{\bar{x}_n}{n}\left(\sum_{t=1}^{h} X_t + \sum_{t=n-h+1}^{n} X_t - 2h\bar{x}_n\right) - \frac{n-h}{n}\bar{x}_n^2 . \end{aligned}$$

Now

$$\left| E\left\{\bar{x}_n\left(\sum_{t=1}^{h} X_t + \sum_{t=n-h+1}^{n} X_t - 2h\bar{x}_n\right)\right\}\right| \leq \frac{2h}{n} \sum_{j=-\infty}^{\infty} |\gamma(j)| ,$$

and we have the first result. Since by an application of (6.2.5) we have $\operatorname{Var}\{\bar{x}_n^2\} = O(n^{-2})$, the second conclusion also follows. ▲

Using Theorems 6.2.1 and 6.2.2 and the results of Chapter 5, we can approximate the mean and variance of the estimated correlation function. If the mean is known, we consider the estimated autocorrelation

$$\tilde{r}(h) = [\tilde{\gamma}(0)]^{-1}\tilde{\gamma}(h),$$

and if the mean is unknown, the estimator

$$\hat{r}(h) = [\hat{\gamma}(0)]^{-1}\hat{\gamma}(h). \tag{6.2.7}$$

If the denominator of the estimator is zero, we define the estimator to be zero.

Theorem 6.2.3. Let the time series $\{X_t\}$ be defined by

$$X_t = \sum_{j=-\infty}^{\infty} \alpha_j e_{t-j},$$

where the sequence $\{\alpha_j\}$ is absolutely summable, and the e_t are independent $(0, \sigma^2)$ random variables with $E\{e_t^6\} = \eta\sigma^6$. Then for fixed h and q,

$$\begin{aligned}
\operatorname{Cov}\{\tilde{r}(h), \tilde{r}(q)\} &= [\gamma(0)]^{-2}[\operatorname{Cov}\{\tilde{\gamma}(h), \tilde{\gamma}(q)\} - \rho(h)\operatorname{Cov}\{\tilde{\gamma}(0), \tilde{\gamma}(q)\} \\
&\quad -\rho(q)\operatorname{Cov}\{\tilde{\gamma}(0), \tilde{\gamma}(h)\} + \rho(h)\rho(q)\operatorname{Var}\{\tilde{\gamma}(0)\}] \\
&\quad + O(n^{-2}) \\
&= \frac{1}{n}\sum_{p=-\infty}^{\infty} [\rho(p)\rho(p-h+q) + \rho(p+q)\rho(p-h) \\
&\quad -2\rho(q)\rho(p)\rho(p-h) - 2\rho(h)\rho(p)\rho(p-q) \\
&\quad +2\rho(h)\rho(q)\rho^2(p)] + O(n^{-2}),
\end{aligned} \tag{6.2.8}$$

$$\begin{aligned}
\operatorname{Cov}\{\hat{r}(h), \hat{r}(q)\} &= \operatorname{Cov}\{\tilde{r}(h), \tilde{r}(q)\} + O(n^{-2}), \\
E\{\tilde{r}(h) - \rho(h)\} &= [\gamma(0)]^{-2}[\rho(h)\operatorname{Var}\{\tilde{\gamma}(0)\} - \operatorname{Cov}\{\tilde{\gamma}(0), \tilde{\gamma}(h)\}] \\
&\quad + O(n^{-2}), \\
E\{\hat{r}(h)\} &= \frac{n-h}{n}\rho(h) - [\gamma(0)]^{-1}[1-\rho(h)]\operatorname{Var}\{\bar{x}_n\} \\
&\quad + [\gamma(0)]^{-2}[\rho(h)\operatorname{Var}\{\tilde{\gamma}(0)\} - \operatorname{Cov}\{\tilde{\gamma}(h), \tilde{\gamma}(0)\}] \\
&\quad + O(n^{-2}).
\end{aligned}$$

Proof. The estimated autocorrelations are bounded and are differentiable functions of the estimated covariances on a closed set containing the true parameter vector as an interior point. Furthermore, the derivatives are bounded on that set. Hence, the conditions of Theorem 5.4.3 are met with $\{\tilde{\gamma}(h), \tilde{\gamma}(q), \tilde{\gamma}(0)\}$

playing the role of $\{\mathbf{X}_n\}$ of that theorem. Since the function $\tilde{r}(h)$ is bounded, we take $\alpha = 1$.

Expanding $[\tilde{r}(h) - \rho(h)][\tilde{r}(q) - \rho(q)]$ through third order terms and using Theorem 5.4.1 to establish that the expected value of the third order moments is $O(n^{-2})$, we have the result (6.2.8). The remaining results are established in a similar manner. ▲

For the first order autoregressive time series $X_t = \rho X_{t-1} + e_t$ it is relatively easy to evaluate the variances of the estimated autocorrelations. We have, for $h > 0$,

$$\begin{aligned}\operatorname{Var}\{\hat{r}(h)\} &\doteq \frac{n-h}{n^2} \sum_{p=-\infty}^{\infty} (\rho^{|2p|} + \rho^{|p+h|+|p-h|} - 4\rho^{|h|}\rho^{|p|+|p-h|} + 2\rho^{|2h|}\rho^{|2p|}) \\ &= \frac{n-h}{n^2}\left[\frac{(1+\rho^2)(1-\rho^{2h})}{1-\rho^2} - 2h\rho^{2h}\right]. \end{aligned} \tag{6.2.9}$$

We note that for large h,

$$\operatorname{Var}\{\hat{r}(h)\} \doteq \frac{n-h}{n^2}\,\frac{1+\rho^2}{1-\rho^2}. \tag{6.2.10}$$

For a time series where the correlations approach zero rapidly, the variance of $\hat{r}(h)$ for large h can be approximated by the first term of (6.2.8). That is, for such a time series and for h such that $\rho(h) \doteq 0$, we have

$$\operatorname{Var}\{\hat{r}(h)\} \doteq \frac{1}{n} \sum_{p=-\infty}^{\infty} \rho^2(p). \tag{6.2.11}$$

We are particularly interested in the behavior of the estimated autocorrelations for a time series of independent random variables, since this is often a working hypothesis in time series analysis.

Corollary 6.2.3. Let $\{e_t\}$ be a sequence of independent $(0, \sigma^2)$ random variables with sixth moment $\eta\sigma^6$. Then, for $h \geqslant q > 0$,

$$E\{\hat{r}(h)\} = -\frac{n-h}{n(n-1)} + O(n^{-2}),$$

$$\operatorname{Cov}\{\hat{r}(h), \hat{r}(q)\} = \begin{cases} \dfrac{n-h}{n^2} + O(n^{-2}), & h = q \neq 0, \\ O(n^{-2}) & \text{otherwise}. \end{cases}$$

Proof. Omitted. ▲

For large n the bias in $\hat{r}(h)$ is negligible. However, it is easy to reduce the bias in the null case of independent random variables. It is suggested that the estimator

$$\hat{\rho}(h) = \hat{r}(h) + \frac{n-h}{(n-1)^2}\{1 - [\hat{r}(h)]^2\} \tag{6.2.12}$$

be used for hypothesis testing when only a small number of observations are available. For time series of the type specified in Corollary 6.2.3 and h, $q > 0$,

$$E\{\hat{\rho}(h)\} = O(n^{-2}),$$

$$\operatorname{Cov}\{\hat{\rho}(h), \hat{\rho}(q)\} = \begin{cases} \dfrac{n-h}{n^2} + O(n^{-2}), & h = q > 0, \\ O(n^{-2}) & \text{otherwise}. \end{cases}$$

In the next section we prove that the $\hat{r}(h)$ and $\hat{\rho}(h)$ are approximately normally distributed. The approximate distribution of the autocorrelations will be adequate for most purposes, but we mention one exact distributional result. A statistic closely related to the first order autocorrelation is the von Neumann ratio:[1]

$$d_v = \frac{\sum_{t=2}^{n} (X_t - X_{t-1})^2}{\sum_{t=1}^{n} (X_t - \bar{x}_n)^2}. \tag{6.2.13}$$

We see that

$$d_v = \frac{\sum_{t=2}^{n} (X_t - \bar{x}_n)^2 - 2\sum_{t=2}^{n} (X_t - \bar{x}_n)(X_{t-1} - \bar{x}_n) + \sum_{t=1}^{n-1} (X_t - \bar{x}_n)^2}{\sum_{t=1}^{n} (X_t - \bar{x}_n)^2}$$

$$= 2\left[1 - \hat{r}(1) - \frac{(X_1 - \bar{x}_n)^2 + (X_n - \bar{x}_n)^2}{2\sum_{t=1}^{n} (X_t - \bar{x}_n)^2}\right].$$

If the X_t are normal independent (μ, σ^2) random variables, it is possible to show that $E\{d_v\} = 2$. Therefore, $r_v = \frac{1}{2}(d_v - 2)$ is an unbiased estimator of zero in that case. von Neumann (1941) and Hart (1942) have given the exact distribution of r_v under the assumption that X_t is a sequence of normal independent (μ, σ^2) random variables. Tables of percentage points are given in Hart (1942) and in Anderson (1971, p. 345). Inspection of these tables demonstrates that the

[1] The ratio is sometimes defined with the multiplier $n/(n-1)$.

percentage points of $t_v = r_v(n+1)^{1/2}(1-r_v^2)^{-1/2}$ are approximately those of Student's t with $n+3$ degrees of freedom for n greater than 10.

Clearly the distribution of $\hat{\rho}(1)[1-\hat{\rho}^2(1)]^{-1/2}(n+1)^{1/2}$, where $\hat{\rho}(1)$ is defined in (6.2.12), is close to the distribution of t_v and therefore may also be approximated by Student's t with $n+3$ degrees of freedom when the observations are independent normal random variables.

Kendall and Stuart (1966) and Anderson (1971) present discussions of the distributional theory of statistics such as d_v.

6.3. CENTRAL LIMIT THEOREMS FOR STATIONARY TIME SERIES

The results of this chapter have, so far, been concerned with the mean and variance of certain sample statistics computed from a single realization. In order to perform tests of hypotheses or set confidence limits for the underlying parameters, some distribution theory is required.

That the mean of a finite moving average is, in the limit, normally distributed is a simple extension of the central limit theorems of Section 5.3.

Proposition 6.3.1. Let $\{X_t: t \in (0, \pm 1, \pm 2, \ldots)\}$ be defined by

$$X_t = \mu + \sum_{j=0}^{m} b_j e_{t-j},$$

where $b_0 = 1$, $\sum_{j=0}^{m} b_j \neq 0$, and the e_t are uncorrelated $(0, \sigma^2)$ random variables. Assume $n^{1/2}\bar{e}_n$ converges in distribution to a $N(0, \sigma^2)$ random variable. Then,

$$n^{1/2}(\bar{x}_n - \mu) \xrightarrow{\mathscr{L}} N\left(0, \sigma^2\left[\sum_{j=0}^{m} b_j\right]^2\right).$$

Proof. We have

$$\begin{aligned} n^{1/2}(\bar{x}_n - \mu) &= n^{-1/2}\sum_{t=1}^{n}(X_t - \mu) \\ &= n^{-1/2}\sum_{t=1}^{n}(e_t + b_1 e_{t-1} + b_2 e_{t-2} + \cdots + b_m e_{t-m}) \\ &= n^{-1/2}\sum_{j=0}^{m} b_j \sum_{t=1}^{n} e_t + n^{-1/2}\sum_{s=1}^{m}\sum_{j=s}^{m} b_j e_{1-s} \\ &\quad - n^{-1/2}\sum_{s=0}^{m-1}\sum_{j=s+1}^{m} b_j e_{n-s}. \end{aligned}$$

Both $n^{-1/2}\sum_{s=0}^{m-1}\sum_{j=s+1}^{m} b_j e_{n-s}$ and $n^{-1/2}\sum_{s=1}^{m}\sum_{j=s}^{m} b_j e_{1-s}$ converge in probability to zero as n increases, by Chebyshev's inequality.

Therefore, the limiting distribution of $n^{1/2}(\bar{x}_n - \mu)$ is the limiting distribution of $n^{1/2} \sum_{j=0}^{m} b_j \bar{e}_n$, and the result follows. ▲

If $\sum_{j=0}^{m} b_j = 0$, then $\bar{x}_n - \mu$ has a variance that approaches zero at a rate faster than n^{-1}. The reader can verify this by considering, for example, the time series $X_t = \mu + e_t - e_{t-1}$. In such a case the theorem holds in the sense that $n^{1/2}(\bar{x}_n - \mu)$ is converging to the singular (zero variance) normal distribution.

Moving average time series of independent random variables are special cases of a more general class of time series called *m-dependent.*

Definition 6.3.1. The sequence of random variables $\{Z_t: t \in (0, \pm 1, \pm 2, \dots)\}$ is said to be m-dependent if $s - r > m$, where m is a positive integer, implies that the two sets

$$(\dots, Z_{r-2}, Z_{r-1}, Z_r), \qquad (Z_s, Z_{s+1}, Z_{s+2}, \dots)$$

are independent.

We give a theorem for such time series due to Hoeffding and Robbins (1948).

Theorem 6.3.1. Let $\{Z_t: t \in (1, 2, \dots)\}$ be a sequence of m-dependent random variables with $E\{Z_t\} = 0$, $E\{Z_t^2\} = \sigma_t^2 < \beta < \infty$, and $E\{|Z_t|^{2+2\delta}\} \le \beta^{2+2\delta}$ for some $\delta > 0$. Let the limit

$$\lim_{p \to \infty} p^{-1} \sum_{j=1}^{p} A_{t+j} = A,$$

$A \neq 0$, be uniform for $t = 1, 2, \dots$, where

$$A_t = E\{Z_{t+m}^2\} + 2 \sum_{j=1}^{m} E\{Z_{t+m-j} Z_{t+m}\}.$$

Then

$$n^{-1/2} \sum_{t=1}^{n} Z_t \xrightarrow{\mathscr{L}} N(0, A).$$

Proof. Fix an α, $0 < \alpha < 0.25 - \epsilon_\delta$, where

$$0.25 > \epsilon_\delta > \max\{0, 0.25(1 + 2\delta)^{-1}(1 - 2\delta)\}.$$

Let k be the largest integer less than n^α, and let p be the largest integer less than $k^{-1}n$. Define

$$Y_i = \sum_{j=1}^{k-m} Z_{(i-1)k+j}, \qquad i = 1, 2, \dots, p,$$

$$S_p = n^{-1/2} \sum_{i=1}^{p} Y_i.$$

Then the difference

$$D_n = n^{-1/2}\sum_{t=1}^{n} Z_t - S_p = n^{-1/2}\left[\sum_{i=1}^{p-1}\left(\sum_{j=1}^{m} Z_{ki-m+j}\right) + \sum_{t=pk-m+1}^{n} Z_t\right].$$

For n large enough so that $k > 2m$, the sums $\Sigma_{j=1}^{m} Z_{ki-m+j}$, $i = 1, 2, \ldots, p$, are independent. By the assumption on the moments, we have

$$\operatorname{Var}\left\{\sum_{j=1}^{m} Z_{ki-m+j}\right\} \le m^2\beta^2,$$

$$\operatorname{Var}\left\{\sum_{t=pk-m+1}^{n} Z_t\right\} \le (k+m)^2\beta^2.$$

It follows that

$$\operatorname{Var}\{D_n\} \le n^{-1}\beta^2\{m^2(p-1) + (k+m)^2\} = o(1),$$

and $n^{-1/2}\Sigma_{t=1}^{n} Z_t$ converges in mean square to S_p. Since Z_t is correlated only with $Z_{t-m}, Z_{t-m+1}, \ldots, Z_{t+m-1}, Z_{t+m}$, the addition of a Z_t to a sum containing the m preceding terms, $Z_{t-1}, Z_{t-2}, \ldots, Z_{t-m}$, increases the variance by the amount A_{t-m}. Therefore,

$$\operatorname{Var}\{Y_i\} = \operatorname{Var}\left\{\sum_{j=1}^{k-m} Z_{(i-1)k+j}\right\} = \operatorname{Var}\left\{\sum_{j=1}^{m} Z_{(i-1)k+j}\right\} + \sum_{j=1}^{k-2m} A_{(i-1)k+j}$$

and

$$\operatorname{Var}\{S_p\} = n^{-1}\sum_{i=1}^{p}\operatorname{Var}\left\{\sum_{j=1}^{m} Z_{(i-1)k+j}\right\} + n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{k-2m} A_{(i-1)k+j}.$$

Since $\operatorname{Var}\{\Sigma_{j=1}^{m} Z_{(i-1)k+j}\} \le m^2\beta^2$, and $p^{-1}\Sigma_{j=1}^{p} A_{t+j}$ converges uniformly, we have

$$\lim_{p\to\infty} n^{-1}\sum_{i=1}^{p}\operatorname{Var}\{Y_i\} = \lim_{p\to\infty} p^{-1}\sum_{i=1}^{p}\operatorname{Var}(k^{-1/2}Y_i) = A.$$

Now

$$E\{|k^{-1/2}Y_i|^{2+2\delta}\} = k^{1+\delta}E\left\{\left|k^{-1}\sum_{j=1}^{k-m} Z_{(i-1)k+j}\right|^{2+2\delta}\right\} \le k^{1+\delta}\beta^{2+2\delta}$$

and

$$\lim_{p\to\infty}\frac{\sum_{i=1}^{p} E\{|k^{-1/2}Y_i|^{2+2\delta}\}}{\left(\sum_{i=1}^{p} E\{|k^{-1/2}Y_i|^2\}\right)^{1+\delta}} \le \lim_{p\to\infty}\frac{pk^{1+\delta}\beta^{2+2\delta}}{(pA)^{1+\delta}} = 0.$$

Hence, $p^{-1/2}\sum_{i=1}^{p} k^{-1/2} Y_i$ satisfies the conditions of Liapounov's central limit theorem, and the result follows. ▲

For the mth order moving average, the A_t of Theorem 6.3.1 is the same for all t and is

$$\gamma(0) + 2\sum_{j=1}^{m} \gamma(j) = \sum_{j=0}^{m} b_j^2\sigma^2 + 2\sum_{j=1}^{m}\sum_{s=0}^{m-j} b_s b_{s+j}\sigma^2 = \left(\sum_{j=0}^{m} b_j\right)^2 \sigma^2 ,$$

which agrees with Proposition 6.3.1.

The results of Proposition 6.3.1 and Theorem 6.3.1 may be generalized to infinite moving averages of independent random variables. Our proofs follow those of Diananda (1953) and Anderson (1959, 1971). We first state a lemma required in the proofs of the primary theorems.

Lemma 6.3.1. Let the random variables ξ_n with distribution functions $F_{\xi_n}(z)$ be defined by

$$\xi_n = S_{kn} + D_{kn}$$

for $k = 1, 2, \ldots$ and $n = 1, 2, \ldots$. Let

$$\operatorname*{plim}_{k\to\infty} D_{kn} = 0$$

uniformly in n. Assume

$$F_{S_{kn}}(z) \xrightarrow{C} F_{\psi_k}(z) \qquad \text{as} \quad n \longrightarrow \infty$$

and

$$F_{\psi_k}(z) \xrightarrow{C} F_{\xi}(z) \qquad \text{as} \quad k \longrightarrow \infty .$$

Then

$$F_{\xi_n}(z) \xrightarrow{C} F_{\xi}(z) \qquad \text{as} \quad n \longrightarrow \infty .$$

Proof. Let z_0 be a point of continuity of $F_\xi(z)$, and fix $\delta > 0$. There is an $\epsilon > 0$ such that

$$|F_\xi(z) - F_\xi(z_0)| < 0.25\delta$$

for z in $[z_0 - \epsilon, z_0 + \epsilon]$, and $F_\xi(z)$ is continuous at $z_0 - \epsilon$ and at $z_0 + \epsilon$. By hypothesis, there exists a K_0 such that

$$|F_{\psi_k}(z_0 - \epsilon) - F_\xi(z_0 - \epsilon)| < 0.25\delta ,$$
$$|F_{\psi_k}(z_0 + \epsilon) - F_\xi(z_0 + \epsilon)| < 0.25\delta$$

for $k > K_0$. This means that

$$|F_{\psi_k}(z) - F_\xi(z_0)| < 0.5\delta$$

for $k > K_0$ and $z_0 - \epsilon \leq z \leq z_0 + \epsilon$.

Now there is a K_1 such that for $k \geq K_1$,

$$P\{|\xi_n - S_{kn}| > \epsilon\} < 0.25\delta$$

for all n. By the arguments used in the proof of Theorem 5.2.1, this implies that

$$F_{S_{kn}}(z - \epsilon) - 0.25\delta \leq F_{\xi_n}(z) \leq F_{S_{kn}}(z + \epsilon) + 0.25\delta$$

for all z.

Fix k at the maximum of K_0 and K_1. Let z_1 and z_2 be continuity points of $F_{\psi_k}(z)$ such that $z_0 - \epsilon \leq z_1 < z_0$ and $z_0 < z_2 \leq z_0 + \epsilon$. Then there exists an N such that for $n > N$,

$$|F_{S_{kn}}(z_1) - F_{\psi_k}(z_1)| < 0.25\delta ,$$
$$|F_{S_{kn}}(z_2) - F_{\psi_k}(z_2)| < 0.25\delta .$$

Therefore, for $n > N$,

$$\begin{aligned} F_\xi(z_0) - \delta &\leq F_{\psi_k}(z_1) - 0.5\delta \leq F_{S_{kn}}(z_1) - 0.25\delta \\ &\leq F_{\xi_n}(z_0) \leq F_{S_{kn}}(z_2) + 0.25\delta \leq F_{\psi_k}(z_2) + 0.5\delta \\ &\leq F_\xi(z_0) + \delta . \end{aligned}$$

▲

The following two lemmas are the convergence in probability and almost sure results analogous to the convergence in distribution result of Lemma 6.3.1.

Lemma 6.3.2. Let (ξ_n, S_{kn}, D_{kn}), $k = 1, 2, \ldots,$ $n = 1, 2, \ldots,$ be random variables defined on the probability space $(\Omega, \mathscr{A}, P)$ satisfying

$$\xi_n = S_{kn} + D_{kn} . \tag{6.3.1}$$

Assume:

(i) one has

$$\operatorname*{plim}_{k \to \infty} D_{kn} = 0 \qquad \text{uniformly in } n , \tag{6.3.2}$$

(ii) for every fixed k there is an $S_{k.}$ satisfying

$$\operatorname*{plim}_{n\to\infty} (S_{kn} - S_{k.}) = 0\,, \tag{6.3.3}$$

(iii) one has

$$\operatorname*{plim}_{k\to\infty} (S_{k.} - \xi) = 0\,. \tag{6.3.4}$$

Then

$$\operatorname*{plim}_{n\to\infty} (\xi_n - \xi) = 0\,.$$

Proof. Omitted. ▲

Lemma 6.3.3. Let the assumption (6.3.1) of Lemma 6.3.2 hold. Replace the assumptions (6.3.2), (6.3.3), and (6.3.4) with the following:

(i) one has

$$\lim_{k\to\infty} D_{kn} = 0 \quad \text{a.s., uniformly in } n\,; \tag{6.3.5}$$

(ii) for every fixed k, there is an $S_{k.}$ satisfying

$$\lim_{n\to\infty} (S_{kn} - S_{k}) = 0 \qquad \text{a.s.}\,; \tag{6.3.6}$$

(iii) one has

$$\lim_{k\to\infty} (S_{k.} - \xi) = 0 \qquad \text{a.s.} \tag{6.3.7}$$

Then

$$\lim_{n\to\infty} (\xi_n - \xi) = 0 \qquad \text{a.s.}$$

Proof. Omitted. ▲

We use Lemma 6.3.2. to prove that the sample mean of an infinite moving average converges in probability to the population mean under weak conditions on the e_t. For example, the conclusion holds for independently identically distributed e_t with zero mean. Corollary 6.1.1.2 used finite variance to obtain a stronger conclusion.

Theorem 6.3.2. Let the time series X_t be defined by

$$X_t = \mu + \sum_{j=-\infty}^{\infty} w_j e_{t-j}\,,$$

where $\sum_{j=-\infty}^{\infty} |w_j| < \infty$. Assume that $E\{|e_t|\} < M < \infty$ for all t and that

$$n^{-1} \sum_{t=1}^{n} e_t \xrightarrow{P} 0 .$$

Then $\bar{x}_n \xrightarrow{P} \mu$.

Proof. We apply Lemma 6.3.2. Let $\xi_n = \bar{x}_n - \mu$,

$$S_{kn} = n^{-1} \sum_{t=1}^{n} \sum_{j=-k}^{k} w_j e_{t-j} = n^{-1} \sum_{j=-k}^{k} w_j \sum_{t=1}^{n} e_{t-j} , \tag{6.3.8}$$

and

$$D_{kn} = n^{-1} \sum_{t=1}^{n} \sum_{|j|>k} w_j e_{t-j} .$$

By the assumption that the e_t are uncorrelated $(0, \sigma^2)$ random variables, $n^{-1} \sum_{t=1}^{n} e_{t-j}$ converges to zero in probability as $n \to \infty$. Hence, the sum (6.3.8) converges to zero $(=S_{k.})$ as $n \to \infty$ for every fixed k, and condition (ii) of Lemma 6.3.2 is satisfied. Condition (iii) is trivially satisfied for $\xi = 0$. Now

$$\begin{aligned} P\{|D_{kn}| > \epsilon\} &\le \epsilon^{-1} E\left\{ n^{-1} \sum_{t=1}^{n} \sum_{|j|>k} |w_j||e_{t-j}| \right\} \\ &\le M\epsilon^{-1} \sum_{|j|>k} |w_j| \end{aligned}$$

for some finite M, by Chebyshev's inequality. Hence, condition (i) of Lemma 6.3.2 is satisfied and the conclusion follows. ▲

Using Lemma 6.3.1, we show that if the e_t of an infinite moving average satisfy a central limit theorem, then the moving average also satisfies a central limit theorem.

Theorem 6.3.3. Let X_t be a covariance stationary time series satisfying

$$X_t = \sum_{j=0}^{\infty} \alpha_j e_{t-j} ,$$

where $\sum_{j=0}^{\infty} |\alpha_j| < \infty$, $\sum_{j=0}^{\infty} \alpha_j \neq 0$, and the e_t are uncorrelated $(0, \sigma^2)$ random variables. Assume

$$n^{1/2} \bar{e}_n \xrightarrow{\mathscr{L}} N(0, \sigma^2) .$$

Then

$$n^{-1/2} \sum_{t=1}^{n} X_t \xrightarrow{\mathscr{L}} N\left(0, \sum_{h=-\infty}^{\infty} \gamma_X(h)\right),$$

where

$$\sum_{h=-\infty}^{\infty} \gamma_X(h) = \left(\sum_{j=0}^{\infty} \alpha_j\right)^2 \sigma^2 .$$

Proof. Let

$$(Y_{tk}, W_{tk}) = \left(\sum_{j=0}^{k} \alpha_j e_{t-j}, \sum_{j=k+1}^{\infty} \alpha_j e_{t-j}\right),$$

and define the normalized sums

$$(S_{kn}, D_{kn}) = n^{-1/2} \sum_{t=1}^{n} (Y_{tk}, W_{tk}) .$$

For a fixed k, W_{tk} is a stationary time series such that

$$\gamma_W(h) = \sum_{j=k+1}^{\infty} \alpha_j \alpha_{j+|h|} \sigma^2 , \qquad h = 0, \pm 1, \pm 2, \ldots .$$

It follows that

$$\begin{aligned} \operatorname{Var}\{D_{kn}\} &= \frac{1}{n} \sum_{h=-(n-1)}^{n-1} (n - |h|) \gamma_W(h) \\ &\leq \sigma^2 \sum_{j=k+1}^{\infty} \alpha_j^2 + 2\sigma^2 \sum_{h=1}^{n-1} \sum_{j=k+1}^{\infty} |\alpha_j \alpha_{j+h}| \\ &\leq \sigma^2 \left(\sum_{j=k+1}^{\infty} |\alpha_j|\right)^2 . \end{aligned}$$

Therefore, by Chebyshev's inequality, D_{kn} converges to zero as $k \to \infty$ uniformly in n. For fixed k, Y_{tk} is a finite moving average of e_t that satisfies a central limit theorem. Therefore, by Proposition 6.3.1, as n tends to infinity,

$$S_{kn} \xrightarrow{\mathscr{L}} N\left[0, \left(\sum_{j=0}^{k} \alpha_j\right)^2 \sigma^2\right].$$

As k increases, the variance of the normal distribution converges to $(\sum_{j=0}^{\infty} \alpha_j)^2 \sigma^2$. The conclusion follows by Lemma 6.3.1. ▲

The stationary autoregressive moving average is a special case of an infinite moving average. Because the weights decline exponentially for the autoregressive moving average, the sample mean of Y_t can be expressed as a function of the sample mean of the e_t with remainder of order in probability n^{-1}.

Corollary 6.3.3. Let Y_t be a stationary finite order autoregressive moving average satisfying

$$\sum_{j=0}^{p} a_j(Y_{t-j} - \mu) = \sum_{i=0}^{q} b_i e_{t-i},$$

where $a_0 = 1$, $b_0 = 1$, the roots of the autoregressive characteristic equation are less than one in absolute value, and $\{e_t\}$ is a sequence of uncorrelated $(0, \sigma^2)$ random variables. Then

$$\bar{y}_n = \mu + n^{-1} \sum_{t=1}^{n} \left(\sum_{j=0}^{\infty} \upsilon_j \right) e_t + O_p(n^{-1}),$$

where the υ_j are defined in Theorem 2.7.1 and $\Sigma_{j=0}^{\infty} \upsilon_j = (\Sigma_{j=0}^{p} a_j)^{-1} \Sigma_{i=0}^{q} b_i$.

If $n^{1/2}\bar{e}_n$ converges in distribution to a $N(0, \sigma^2)$ random variable, then

$$n^{1/2}(\bar{y}_n - \mu) \xrightarrow{\mathscr{L}} N\left[0, \left(\sum_{j=0}^{\infty} \upsilon_j \right)^2 \sigma^2 \right].$$

Proof. By the representation of Theorem 2.7.1,

$$\begin{aligned} \sum_{t=1}^{n} (Y_t - \mu) = \sum_{t=1}^{n} \left(\sum_{j=0}^{\infty} \upsilon_j \right) e_t &+ \sum_{i=1}^{\infty} \sum_{j=i}^{n+i-1} \upsilon_j e_{1-i} \\ &- \sum_{i=0}^{n-1} \sum_{j=i}^{\infty} \upsilon_{j+1} e_{n-i}. \end{aligned}$$

Now, by Exercise 2.24, there is some M and $0 < \lambda < 1$ such that $|\upsilon_j| < M\lambda^j$ for all $j \geq 0$. Hence, there is a K such that

$$V\left\{ \sum_{i=1}^{\infty} \sum_{j=i}^{n+i-1} \upsilon_j e_{1-i} \right\} \leq \sum_{i=1}^{\infty} \lambda^{2i} K^2 \sigma^2 = (1 - \lambda^2)^{-1} \lambda^2 K^2 \sigma^2$$

and

$$V\left\{ \sum_{i=0}^{n-1} \sum_{j=i}^{\infty} \upsilon_{j+1} e_{n-i} \right\} \leq (1 - \lambda^2)^{-1} \lambda^2 K^2 \sigma^2$$

for all n. Dividing the variances by n^2, we obtain the first result. The limiting distribution result is an immediate consequence of the order in probability result and the assumption that $n^{1/2}\bar{e}_n$ converges in distribution. ▲

We now obtain a central limit theorem for a linear function of a realization of a stationary time series. We shall state the assumptions of this theorem in a manner that permits us to use the Lindeberg central limit theorem.

Theorem 6.3.4. Let $\{X_t: t \in T = (0, \pm 1, \pm 2, \ldots)\}$ be a time series defined by

$$X_t = \sum_{j=0}^{\infty} \alpha_j e_{t-j},$$

where $\sum_{j=0}^{\infty} |\alpha_j| < \infty$, and the e_t are independent $(0, \sigma^2)$ random variables with distribution functions $F_t(e)$ such that

$$\lim_{\delta \to \infty} \sup_{t \in T} \int_{|e| > \delta} e^2 \, dF_t(e) = 0 .$$

Furthermore, let $\{C_t\}_{t=1}^{\infty}$ be a sequence of fixed real numbers satisfying

(i) $\lim_{n \to \infty} \sum_{t=1}^{n} C_t^2 = \infty$,

(ii) $\lim_{n \to \infty} \left(\sum_{t=1}^{n} C_t^2 \right)^{-1} C_n^2 = 0$,

(iii) $\lim_{n \to \infty} \left(\sum_{t=1}^{n} C_t^2 \right)^{-1} \sum_{t=1}^{n-h} C_t C_{t+|h|} = g(h), \; h = 0, \pm 1, \pm 2, \ldots .$

Let $V = \sum_{h=-\infty}^{\infty} g(h) \gamma_X(h) \neq 0$. Then

$$\left(\sum_{t=1}^{n} C_t^2 \right)^{-1/2} \sum_{t=1}^{n} C_t X_t \xrightarrow{\mathscr{L}} N(0, V) .$$

Proof. Following the proof of Theorem 6.3.3, we set

$$(Y_{tk}, W_{tk}) = \left(\sum_{j=0}^{k} \alpha_j e_{t-j}, \sum_{j=k+1}^{\infty} \alpha_j e_{t-j} \right)$$

and note that

$$\begin{aligned} \operatorname{Var}\left\{ \left(\sum_{t=1}^{n} C_t^2 \right)^{-1/2} \sum_{t=1}^{n} C_t W_{tk} \right\} &= \left(\sum_{t=1}^{n} C_t^2 \right)^{-1} \sum_{t=1}^{n} \sum_{j=1}^{n} C_t C_j \gamma_w(t-j) \\ &\leqslant \sum_{h=-(n-1)}^{n-1} |\gamma_w(h)| , \end{aligned}$$

which, by the proof of Theorem 6.3.3, can be made arbitrarily small by choosing k sufficiently large.

For fixed k, Y_{tk} is a finite moving average, and, following the proof of Theorem

6.3.1, we have

$$\left(\sum_{t=1}^{n} C_t^2\right)^{-1/2}\left(\sum_{t=1}^{n} C_t Y_{tk}\right)=\left(\sum_{t=1}^{n} C_t^2\right)^{-1/2}\sum_{t=1}^{n}\sum_{j=0}^{k} b_t e_t + o_p(1),$$

where $b_t = \sum_{j=0}^{k} \alpha_j C_{t+j}$, and the random variables $b_t e_t$ are independent with mean zero and variance $b_t^2\sigma^2$. Let $V_n = \sum_{t=1}^{n} b_t^2\sigma^2$,

$$S_n = V_n^{-1/2}\sum_{t=1}^{n} b_t e_t,$$
$$M_n = \sup_{1\le t\le n} b_t^2,$$

and consider, for $\delta > 0$,

$$\lim_{n\to\infty} V_n^{-1}\sum_{t=1}^{n} b_t^2\int_{R_t} e^2\, dF_t(e) \le \lim_{n\to\infty}\sum_{t=1}^{n} V_n^{-1} b_t^2 \sup_{1\le t\le n}\int_{R_0} e^2\, dF_t(e),$$

where

$$R_t = \{e: |e| > V_n^{1/2}|b_t|^{-1}\delta\}$$

and

$$R_0 = \{e: |e| > M_n^{-1/2}V_n^{1/2}\delta\}.$$

Clearly, $R_t \subset R_0$ for all $t \le n$. Assumptions i and ii imply that

$$\lim_{n\to\infty}\sup_{1\le t\le n}\left(\sum_{j=1}^{n} C_j^2\right)^{-1} C_t^2 = 0.$$

Because $V = \lim_{n\to\infty}(\sum_{t=1}^{n} C_t^2)^{-1}\sum_{t=1}^{n} b_t^2\sigma^2 \ne 0$, we have

$$\lim_{n\to\infty}\sup_{1\le t\le n}\left(\sum_{r=1}^{n} b_r^2\right)^{-1} b_t = 0,$$

which, in turn, implies that $\lim_{n\to\infty} M_n^{-1/2}V_n^{-1/2} = \infty$.

By assumption, the supremum of the integral over R_0 goes to zero. Hence, the assumptions of the Lindeberg central limit theorem are met, and S_n converges in distribution to a normal random variable with zero mean and unit variance. Since

$$\lim_{n\to\infty}\left(\sum_{t=1}^{n} C_t^2\right)^{-1}\sum_{t=1}^{n} b_t^2 = g(0)\sum_{j=0}^{k}\alpha_j^2 + 2\sum_{h=1}^{k} g(h)\sum_{j=0}^{k-h}\alpha_j\alpha_{j+h}$$

and

$$\lim_{k\to\infty} \sigma^2\left\{ g(0)\sum_{j=0}^{k} \alpha_j^2 + 2\sum_{h=1}^{k} g(h)\sum_{j=0}^{k-h} \alpha_j\alpha_{j+h}\right\} = \sum_{h=-\infty}^{\infty} g(h)\gamma_X(h)\,,$$

the result follows by Lemma 6.3.1. ▲

Because a stationary finite order autoregressive moving average time series can be expressed as an infinite moving average with absolutely summable covariance function, the conclusions of Theorem 6.3.4 hold for such time series. Also, the condition that the e_t are independent $(0, \sigma^2)$ random variables can be replaced by the condition that the e_t are martingale differences with $E[e_t^2 \mid \mathscr{A}_{t-1}] = \sigma^2$ a.s. and bounded $2+\delta$ $(\delta > 0)$ moments, where $\mathscr{A}_{t-1}$ is the sigma-field generated by $e_1, \ldots, e_{t-1}$.

We now investigate the large sample properties of the estimated autocovariances and autocorrelations. We use Lemma 6.3.2 to demonstrate that the estimated autocovariances converge in probability to the true values. The assumptions and conclusions of Theorem 6.3.5 differ from those of Theorem 6.2.1. In Theorem 6.2.1, the existence of fourth moments enabled us to obtain variances of the estimated autocovariances. In Theorem 6.3.5, weaker assumptions are used to obtain convergence in probability of the sample autocovariances.

Theorem 6.3.5. Let the stationary time series Y_t be defined by

$$Y_t = \sum_{j=-\infty}^{\infty} w_j e_{t-j}\,,$$

where $\sum_{j=-\infty}^{\infty} |w_j| < \infty$ and $e_t \sim \mathrm{II}(0, \sigma^2)$. Then

$$\operatorname*{plim}_{n\to\infty} \tilde{\gamma}(h) = \operatorname*{plim}_{n\to\infty} \hat{\gamma}(h) = \gamma(h) = \sigma^2 \sum_{j=-\infty}^{\infty} w_j w_{j+h}\,.$$

Proof. Now

$$\begin{aligned}\tilde{\gamma}(h) &= (n-h)^{-1}\sum_{t=1}^{n-h}\sum_{j=-\infty}^{\infty} w_j e_{t-j} \sum_{i=-\infty}^{\infty} w_i e_{t-i+h} \\ &= (n-h)^{-1}\sum_{t=1}^{n-h}\left[\sum_{j=-\infty}^{\infty} w_j w_{j+h} e_{t-j}^2 + \sum_{i\neq j}\sum w_j w_{i+h} e_{t-j} e_{t-i}\right].\end{aligned}$$

Also,

$$\operatorname*{plim}_{n\to\infty} (n-h)^{-1}\sum_{t=1}^{n-h}\sum_{j=-\infty}^{\infty} w_j w_{j+h} e_{t-j}^2 = \gamma(h)$$

by Theorem 6.3.2.

To show that the term

$$(n-h)^{-1} \sum_{t=1}^{n-h} \sum_{i \neq j} \sum w_j w_{i+h} e_{t-j} e_{t-i}$$

converges to zero, we apply Lemma 6.3.2, letting

$$S_{kn} = (n-h)^{-1} \sum_{t=1}^{n-h} \sum_{\substack{|i| \leq k \\ i \neq j}} \sum_{|j| \leq k} w_j w_{i+h} e_{t-j} e_{t-i} ,$$

$$D_{kn}^{(1)} = (n-h)^{-1} \sum_{t=1}^{n-h} \sum_{\substack{|i| > k \\ i \neq j}} \sum_{|j| > k} w_j w_{i+h} e_{t-j} e_{t-i} ,$$

$$D_{kn}^{(2)} = (n-h)^{-1} \sum_{t=1}^{n-h} \sum_{\substack{|i| > k \\ i \neq j}} \sum_{|j| \leq k} w_j w_{i+h} e_{t-j} e_{t-i} ,$$

$$D_{kn}^{(3)} = (n-h)^{-1} \sum_{t=1}^{n-h} \sum_{\substack{|i| \leq k \\ i \neq j}} \sum_{|j| > k} w_j w_{i+h} e_{t-j} e_{t-i} ,$$

and

$$D_{kn} = D_{kn}^{(1)} + D_{kn}^{(2)} + D_{kn}^{(3)} .$$

For fixed i, j with $i \neq j$, using Chebyshev's inequality, it can be proved that

$$\operatorname*{plim}_{n \to \infty} (n-h)^{-1} \sum_{t=1}^{n-h} w_j w_{i+h} e_{t-j} e_{t-i} = 0 .$$

Hence, for any fixed k, $\operatorname{plim}_{n \to \infty} S_{kn} = 0$.

Now, by Chebyshev's inequality,

$$P\{|D_{kn}^{(1)}| > \epsilon\} \leq \epsilon^{-1} \Big(\sum_{\substack{|i| > k \\ i \neq j}} \sum_{|j| > k} |w_j| \, |w_{i+h}| \sigma^2 \Big) .$$

Therefore,

$$\operatorname*{plim}_{k \to \infty} D_{kn}^{(1)} = 0 \qquad \text{uniformly in } n .$$

In a similar manner, we can prove

$$\operatorname*{plim}_{k \to \infty} [D_{kn}^{(2)} + D_{kn}^{(3)}] = 0 \qquad \text{uniformly in } n ,$$

and it follows that $\tilde{\gamma}(h)$ converges in probability to $\gamma(h)$. The result for $\hat{\gamma}(h)$ follows because, by the proof of Theorem 6.2.2, $\tilde{\gamma}(h) - \hat{\gamma}(h) = O_p(n^{-1})$. ▲

Corollary 6.3.5. Let the stationary time series Y_t satisfy

$$Y_t = \sum_{j=-\infty}^{\infty} w_j e_{t-j},$$

where $\sum_{j=-\infty}^{\infty} |w_j| < \infty$,

$$E\{(e_t, e_t^2) \mid \mathscr{A}_{t-1}\} = (0, \sigma^2) \qquad \text{a.s.}$$
$$E\{|e_t|^{2+\delta} \mid \mathscr{A}_{t-1}\} < L < \infty \qquad \text{a.s.}$$

for some $\delta > 0$, and $\mathscr{A}_{t-1}$ is the sigma-field generated by $e_{t-1}, e_{t-2}, \ldots$. Then

$$\operatorname*{plim}_{n\to\infty} \tilde{\gamma}(h) = \operatorname*{plim}_{n\to\infty} \hat{\gamma}(h) = \gamma(h) = \sigma^2 \sum_{j=-\infty}^{\infty} w_j w_{j+h}.$$

Proof. The proof parallels that of Theorem 6.3.5. ▲

The vector of a finite number of sample autocovariances converges in distribution to a normal vector under mild assumptions.

Theorem 6.3.6. Let Y_t be the stationary time series defined by

$$Y_t = \sum_{j=-\infty}^{\infty} b_j e_{t-j},$$

where the e_t are independent $(0, \sigma^2)$ random variables with fourth moment $\eta\sigma^4$ and finite $4 + \delta$ absolute moment for some $\delta > 0$, and the b_j are absolutely summable. Let K be fixed. Then the limiting distribution of $n^{1/2}[\hat{\gamma}(0) - \gamma(0), \hat{\gamma}(1) - \gamma(1), \ldots, \hat{\gamma}(K) - \gamma(K)]'$ is multivariate normal with mean zero and covariance matrix $\mathbf{V}$ whose elements are defined by (6.2.4) of Theorem 6.2.1.

Proof. The estimated covariance, for $h = 0, 1, 2, \ldots, K$, is

$$\begin{aligned}\hat{\gamma}(h) &= n^{-1} \sum_{t=1}^{n-h} (Y_t - \bar{y}_n)(Y_{t+h} - \bar{y}_n) \\ &= n^{-1} \sum_{t=1}^{n-h} Y_t Y_{t+h} - n^{-1}\bar{y}_n \sum_{t=1}^{n-h} (Y_t + Y_{t+h}) + n^{-1}(n-h)\bar{y}_n^2 \qquad (6.3.9)\end{aligned}$$

and the last two terms, when multiplied by $n^{1/2}$, converge in probability to zero. Therefore, in investigating the limiting distribution of $n^{1/2}[\hat{\gamma}(h) - \gamma(h)]$ we need only consider the first term on the right of (6.3.9). Let

$$Y_t = X_{mt} + W_{mt},$$

where

$$(X_{mt}, W_{mt}) = \left(\sum_{j=-m}^{m} b_j e_{t-j}, \sum_{|j|>m} b_j e_{t-j} \right)$$
$$= \left(\sum_{j=-\infty}^{\infty} \alpha_j e_{t-j}, \sum_{j=-\infty}^{\infty} \beta_j e_{t-j} \right),$$

$\alpha_j = b_j$ if $|j| \le m$ and $\alpha_j = 0$ otherwise, and $\beta_j = b_j$ if $|j| \ge m+1$ and $\beta_j = 0$ otherwise. Then

$$n^{-1/2} \sum_{t=1}^{n-h} (X_{mt} + W_{mt})(X_{m,t+h} + W_{m,t+h})$$
$$= n^{-1/2} \left[\sum_{t=1}^{n-h} X_{mt} X_{m,t+h} + \sum_{t=1}^{n-h} X_{mt} W_{m,t+h} + \sum_{t=1}^{n-h} W_{mt} X_{m,t+h} + \sum_{t=1}^{n-h} W_{mt} W_{m,t+h} \right]$$
$$\stackrel{\text{(say)}}{=} S_{mhn} + D_{mn},$$

where $S_{mhn} = n^{-1/2} \sum_{t=1}^{n-h} X_{mt} X_{m,t+h}$. Following the proof of Theorem 6.2.1, we have

$$E\left\{ (n-h)^{-1}(n-q)^{-1} \sum_{s=1}^{n-q} \sum_{t=1}^{n-h} X_{mt} X_{m,t+h} W_{ms} W_{m,s+q} \right\} - \gamma_{X_m}(h)\gamma_{W_m}(q)$$
$$= (\eta - 3)\sigma^4 (n-h)^{-1}(n-q)^{-1} \sum_{s=1}^{n-q} \sum_{t=1}^{n-h} \sum_{j=-\infty}^{\infty} \alpha_j \alpha_{j+h} \beta_{j+s-t} \beta_{j+s-t+q}$$
$$+ (n-h)^{-1}(n-q)^{-1} \sum_{s=1}^{n-q} \sum_{t=1}^{n-h} \left[\sum_{j=-\infty}^{\infty} \alpha_j \beta_{j+s-t} \right.$$
$$\left. \times \sum_{i=-\infty}^{\infty} \alpha_{i+h} \beta_{i+q+s-t} \sum_{j=-\infty}^{\infty} \alpha_j \beta_{j+q+s-t} \sum_{i=-\infty}^{\infty} \alpha_{i+h} \beta_{i+s-t} \right] \sigma^4 .$$

Now,

$$\sum_{p=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} |\alpha_j| |\beta_{j+p}| \le \sum_{p=-\infty}^{\infty} |\beta_p| \sum_{j=-\infty}^{\infty} |\alpha_j| < L_m ,$$

where $L_m \to 0$ uniformly in n as $m \to \infty$ because $\sum_{|j|>m} b_j \to 0$ as $m \to \infty$. Therefore, $D_{mn} \to 0$ uniformly in n as $m \to \infty$.

Consider a linear combination of the errors in the estimated covariances of X_{mt},

$$S_{m,n} = n^{1/2} \sum_{h=0}^{K} \lambda_h \left[n^{-1} \sum_{t=1}^{n-h} X_{mt} X_{m,t+h} - \gamma_{X_m}(h) \right]$$
$$= n^{-1/2} \sum_{h=0}^{K} \sum_{t=1}^{n-h} \lambda_h [Z_{th} - E\{Z_{th}\}] - n^{-1/2} \sum_{h=0}^{K} h \lambda_h \gamma_{X_m}(h) ,$$

where the λ_h are arbitrary real numbers (not all zero) and

$$Z_{th} = X_{mt}X_{m,t+h}\,, \qquad h = 0, 1, 2, \ldots, K\,.$$

Now, Z_{th} is an $(m + h)$-dependent covariance stationary time series with mean $\gamma_{X_m}(h)$ and covariance function

$$\begin{aligned}\gamma_{Z_h}(s) = (\eta - 3)\sigma^4 \sum_{j=-\infty}^{\infty} \alpha_j\alpha_{j+h}\alpha_{j+s}\alpha_{j+s+h} \\ + \gamma^2_{X_m}(h) + \gamma^2_{X_m}(s) + \gamma_{X_m}(s+h)\gamma_{X_m}(s-h)\,,\end{aligned} \qquad (6.3.10)$$

where we have used (6.2.5). Thus, the weighted average of the Z_{th}'s,

$$U_t = \sum_{h=0}^{K} \lambda_h Z_{th} = \sum_{h=0}^{K} \lambda_h X_{mt}X_{m,t+h}\,,$$

is a stationary time series. Furthermore, the time series U_t is $(m + K)$-dependent, it has finite $2 + \delta/2$ moment, and

$$\lim_{n\to\infty} n^{-1/2} \sum_{h=0}^{K} h\lambda_h\gamma_{X_m}(h) = 0\,.$$

Therefore, by Theorem 6.3.1, S_{mn} converges in distribution to a normal random variable. Since the λ_h are arbitrary, the vector random variable

$$n^{1/2}[\hat{\gamma}_{X_m}(0) - \gamma_{X_m}(0), \hat{\gamma}_{X_m}(1) - \gamma_{X_m}(1), \ldots, \hat{\gamma}_{X_m}(K) - \gamma_{X_m}(K)]$$

converges in distribution to a multivariate normal by Theorem 5.3.3, where the covariance matrix is defined by (6.3.10). Because $E\{X_{mt}X_{m,t+h}\}$ converges to $E\{Y_tY_{t+h}\}$ as $m \to \infty$, the conditions of Lemma 6.3.1 are satisfied and we obtain the conclusion. ▲

Theorem 6.3.6 can be proven for e_t that are martingale differences. See Exercises 6.23 and 6.24, and Hannan and Heyde (1972).

Generally it is the estimated autocorrelations that are subjected to analysis, and hence their limiting distribution is of interest.

Corollary 6.3.6.1. Let the assumptions of Theorem 6.3.6 hold. Then the vector $n^{1/2}[\hat{r}(1) - \rho(1), \hat{r}(2) - \rho(2), \ldots, \hat{r}(K) - \rho(K)]'$ converges in distribution to a multivariate normal with mean zero and covariance matrix $\mathbf{G}$, where the hqth element of $\mathbf{G}$ is $\sum_{p=-\infty}^{\infty} [\rho(p)\rho(p-h+q) + \rho(p+q)\rho(p-h) - 2\rho(q)\rho(p)\rho(p-h) - 2\rho(h)\rho(p)\rho(p-q) + 2\rho(h)\rho(q)\rho^2(p)]$.

Proof. Since the $\hat{r}(h)$ are continuous differentiable functions of the $\hat{\gamma}(h)$, the result follows from Theorems 5.1.4, 6.2.3, and 6.3.5. ▲

Observe that if the original time series X_t is a sequence of independent identically distributed random variables with finite moments, the sample correlations will be nearly independent in large samples. Because of the importance of this result in the testing of time series for independence, we state it as a corollary.

Corollary 6.3.6.2. Let the time series e_t be a sequence of independent $(0, \sigma^2)$ random variables with uniformly bounded $4+\delta$ moments for some $\delta > 0$. Let $\hat{\rho}(h)$ be defined by (6.2.12), and let K be a fixed integer. Then $n(n-h)^{-1/2}\hat{\rho}(h)$, $h = 1, 2, \ldots, K$, converge in distribution to independent normal $(0, 1)$ random variables.

Proof. Omitted. ▲

Example 6.3.1. The quarterly seasonally adjusted United States unemployment rate from 1948 to 1972 is given in Table 6.3.1 and displayed in Figure 6.3.1. The mean unemployment rate is 4.77. The autocorrelation function estimated using (6.2.7) is given in Figure 6.3.2 and Table 6.3.2. This plot is sometimes called a correlogram.

To carry out statistical analyses we assume the time series can be treated as a stationary time series with finite sixth moment. If the original time series was a sequence of uncorrelated random variables, we would expect about 95% of the estimated correlations to fall between the lines plus and minus $1.96n^{-1}(n-h)^{1/2}$. Obviously unemployment is not an uncorrelated time series. Casual inspection of the correlogram might lead one to conclude that the time series contains a periodic component with a period of about 54 quarters, since the estimated correlations for h equal to 21 through 33 are negative and below the 1.96 sigma bounds for an uncorrelated time series. However, because the time series is highly correlated at small lags, the variance of the estimated correlations at large lags is much larger than the variance of correlations computed from a white noise sequence.

The first few autocorrelations of the unemployment time series are in good agreement with those generated by the second order autoregressive process

$$X_t = 1.5356X_{t-1} - 0.6692X_{t-2} + e_t\,,$$

where the e_t are uncorrelated $(0, 0.1155)$ random variables. We shall discuss the estimation of the parameters of autoregressive time series in Chapter 8. However, the fact that the sample autocorrelations are consistent estimators of the population correlations permits us to obtain consistent estimators of the autoregressive parameters of a second order process from (2.5.7). The general agreement between the correlations for the second order autoregressive process and the sample correlations is clear from Table 6.3.2.

The roots of the second order autoregressive process are $0.768 \pm 0.282\iota$. We recall that the correlation function of such a second order process can be written as

$$\rho(h) = b_1 m_1^h + b_2 m_2^h\,, \qquad h = 0, 1, 2, \ldots,$$

Table 6.3.1. U.S. Unemployment Rate (Quarterly Seasonally Adjusted) 1948–1972

Year	Quarter	Rate	Year	Quarter	Rate	Year	Quarter	Rate
1948	1	3.73	1957	1	3.93	1966	1	3.87
	2	3.67		2	4.10		2	3.80
	3	3.77		3	4.23		3	3.77
	4	3.83		4	4.93		4	3.70
1949	1	4.67	1958	1	6.30	1967	1	3.77
	2	5.87		2	7.37		2	3.83
	3	6.70		3	7.33		3	3.83
	4	6.97		4	6.37		4	3.93
1950	1	6.40	1959	1	5.83	1968	1	3.73
	2	5.57		2	5.10		2	3.57
	3	4.63		3	5.27		3	3.53
	4	4.23		4	5.60		4	3.43
1951	1	3.50	1960	1	5.13	1969	1	3.37
	2	3.10		2	5.23		2	3.43
	3	3.17		3	5.53		3	3.60
	4	3.37		4	6.27		4	3.60
1952	1	3.07	1961	1	6.80	1970	1	4.17
	2	2.97		2	7.00		2	4.80
	3	3.23		3	6.77		3	5.17
	4	2.83		4	6.20		4	5.87
1953	1	2.70	1962	1	5.63	1971	1	5.93
	2	2.57		2	5.53		2	5.97
	3	2.73		3	5.57		3	5.97
	4	3.70		4	5.53		4	5.97
1954	1	5.27	1963	1	5.77	1972	1	5.83
	2	5.80		2	5.73		2	5.77
	3	5.97		3	5.50		3	5.53
	4	5.33		4	5.57		4	5.30
1955	1	4.73	1964	1	5.47			
	2	4.40		2	5.20			
	3	4.10		3	5.00			
	4	4.23		4	5.00			
1956	1	4.03	1965	1	4.90			
	2	4.20		2	4.67			
	3	4.13		3	4.37			
	4	4.13		4	4.10			

Sources. Business Statistics, 1971 Biennial Edition, pp. 68 and 233 and *Survey of Current Business*, January 1972 and January 1973. Quarterly data are the averages of monthly data.

where

$$(b_1, b_2) = (m_2 - m_1)^{-1}[m_2 - \rho(1), \rho(1) - m_1]\,.$$

For the unemployment time series the estimated parameters are $\hat{b}_1 = 0.500 -$

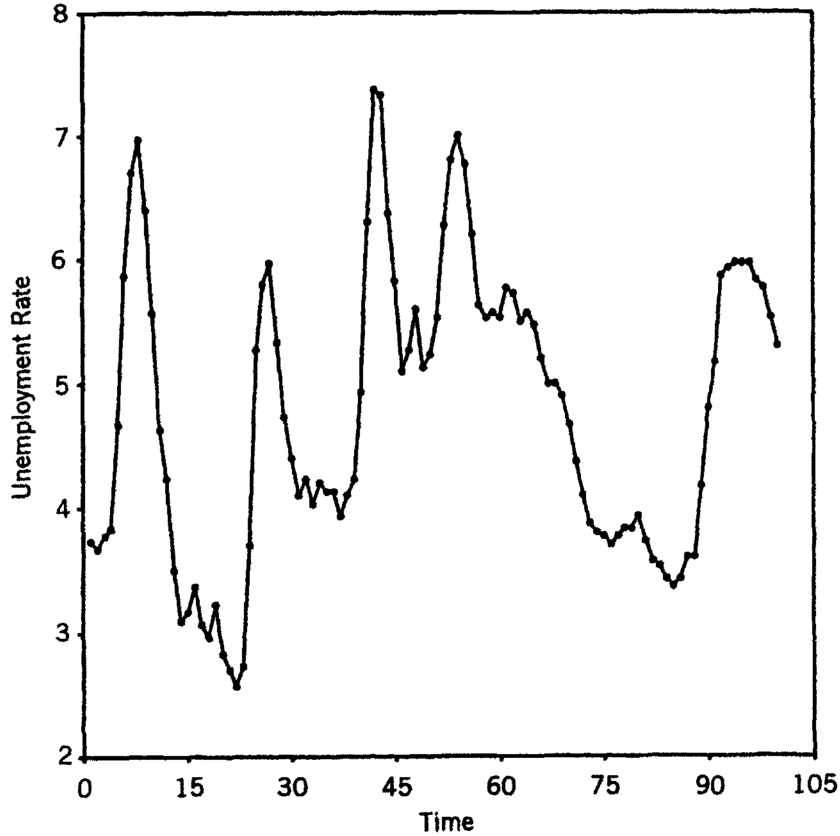

Figure 6.3.1. United States quarterly seasonally adjusted unemployment rate.

$0.270i$ and $\hat{b}_2 = \hat{b}_1^* = 0.500 + 0.270i$. Using these values, we can estimate the variance of the estimated autocorrelations for large lags using (6.2.11). We have

$$
\begin{aligned}
n\,\hat{\mathrm{V}}\mathrm{ar}\{\hat{r}(h)\} &\doteq \sum_{s=-\infty}^{\infty} \hat{\rho}^2(s) \\
&= \hat{b}_1^2 \frac{1+\hat{m}_1^2}{1-\hat{m}_1^2} + \hat{b}_2^2 \frac{1+\hat{m}_2^2}{1-\hat{m}_2^2} + 2\hat{b}_1\hat{b}_2 \frac{1+\hat{m}_1\hat{m}_2}{1-\hat{m}_1\hat{m}_2} \\
&= 4.812\,.
\end{aligned}
$$

Thus the estimated standard error of the estimated autocorrelations at large lags is about 0.22, and the observed correlations at lags near 27 could arise from such a process.

Given that the time series was generated by the second order autoregressive mechanism, the variance of the sample mean can be estimated by Corollary

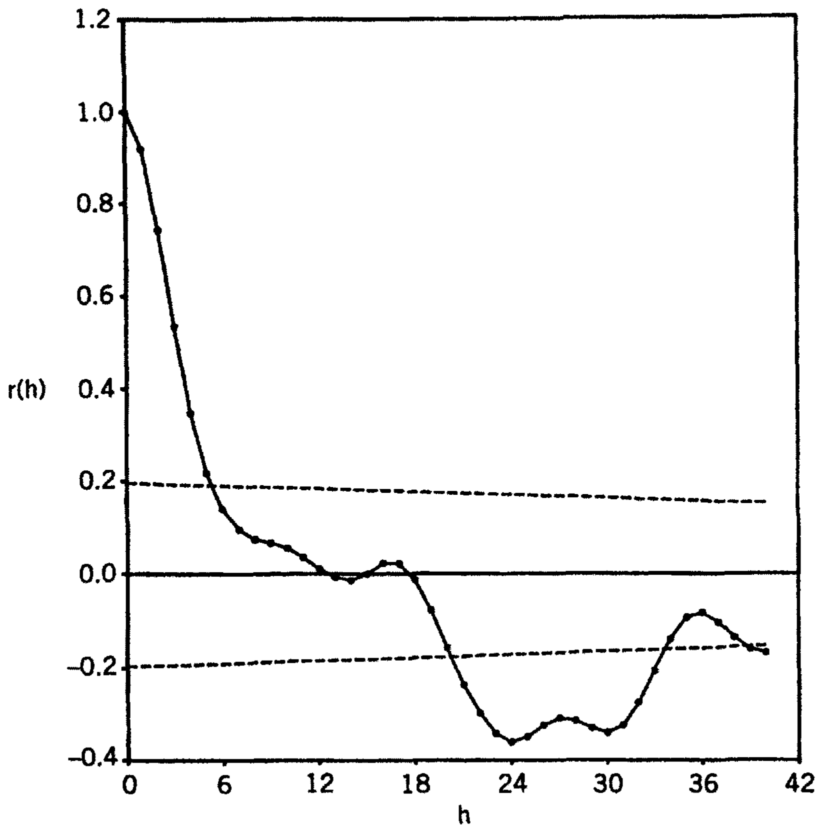

Figure 6.3.2. Correlogram of quarterly seasonally adjusted unemployment rate.

6.1.1.2. By that corollary

$$\begin{aligned} n\,\hat{\mathrm{V}}\mathrm{ar}\{\bar{x}_n\} &= \sum_{h=-\infty}^{\infty} \hat{\gamma}(h) = \hat{\gamma}(0) \sum_{h=-\infty}^{\infty} \hat{\rho}(h) \\ &= \frac{0.1155}{(1 - 1.5356 + 0.6692)^2} \\ &= 6.47\,. \end{aligned}$$

For this highly correlated process the variance of the mean is about five times that of an uncorrelated time series with the same variance. ▲▲

6.4. ESTIMATION OF THE CROSS COVARIANCES

In our Section 1.7 discussion of vector valued time series we introduced the $k \times k$

Table 6.3.2. Estimated Autocorrelations, Quarterly U.S. Seasonally Adjusted Unemployment Rate, 1948–72

Lag h	Estimated Correlations	$\|1.96n^{-1}(n-h)^{1/2}\|$	Correlations for Second Order Process
0	1.0000	—	1.0000
1	0.9200	0.1950	0.9200
2	0.7436	0.1940	0.7436
3	0.5348	0.1930	0.5262
4	0.3476	0.1920	0.3105
5	0.2165	0.1910	0.1247
6	0.1394	0.1900	−0.0164
7	0.0963	0.1890	−0.1085
8	0.0740	0.1880	−0.1557
9	0.0664	0.1870	−0.1665
10	0.0556	0.1859	−0.1515
11	0.0352	0.1849	−0.1212
12	0.0109	0.1839	−0.0848
13	−0.0064	0.1828	−0.0490
14	−0.0135	0.1818	−0.0186
15	0.0004	0.1807	0.0043
16	0.0229	0.1796	0.0190
17	0.0223	0.1786	0.0263
18	−0.0126	0.1775	0.0277
19	−0.0762	0.1764	0.0249
20	−0.1557	0.1753	0.0197
21	−0.2351	0.1742	0.0136
22	−0.2975	0.1731	0.0077
23	−0.3412	0.1720	0.0027
24	−0.3599	0.1709	−0.0010
25	−0.3483	0.1697	−0.0033
26	−0.3236	0.1686	−0.0044
27	−0.3090	0.1675	−0.0046
28	−0.3142	0.1663	−0.0041
29	−0.3299	0.1652	−0.0032
30	−0.3396	0.1640	−0.0022
31	−0.3235	0.1628	−0.0012
32	−0.2744	0.1616	−0.0004
33	−0.2058	0.1604	0.0002
34	−0.1378	0.1592	0.0006
35	−0.0922	0.1580	0.0007
36	−0.0816	0.1568	0.0008
37	−0.1027	0.1556	0.0007
38	−0.1340	0.1543	0.0005
39	−0.1590	0.1531	0.0004
40	−0.1669	0.1518	0.0002

covariance matrix

$$\boldsymbol{\Gamma}(h) = E\{(\mathbf{X}_t - \boldsymbol{\mu})(\mathbf{X}_{t+h} - \boldsymbol{\mu})'\},$$

where $\mathbf{X}_t$ is a stationary k-dimensional time series and $\boldsymbol{\mu} = E\{\mathbf{X}_t\}$. The jth diagonal element of the matrix is the autocovariance of X_{jt}, $\gamma_{jj}(h) = E\{(X_{jt} - \mu_j)(X_{j,t+h} - \mu_j)\}$, and the ijth element is the cross covariance between X_{it} and X_{jt},

$$\gamma_{ij}(h) = E\{(X_{it} - \mu_i)(X_{j,t+h} - \mu_j)\}.$$

Expressions for the estimated cross covariance analogous to those of (6.2.1) and (6.2.3) are

$$\tilde{\gamma}_{ij}(h) = \begin{cases} \dfrac{1}{n-h} \displaystyle\sum_{t=1}^{n-h} X_{it}X_{j,t+h}, & h = 0, 1, \ldots, n-1, \\ \dfrac{1}{n+h} \displaystyle\sum_{t=-h}^{n} X_{it}X_{j,t+h}, & h = -1, -2, \ldots, -(n-1), \end{cases} \tag{6.4.1}$$

for the means known and taken to be zero, and

$$\hat{\gamma}_{ij}(h) = \begin{cases} \dfrac{1}{n} \displaystyle\sum_{t=1}^{n-h} (X_{it} - \bar{x}_{in})(X_{j,t+h} - \bar{x}_{jn}), & h = 0, 1, \ldots, n-1, \\ \dfrac{1}{n} \displaystyle\sum_{t=-h}^{n} (X_{it} - \bar{x}_{in})(X_{j,t+h} - \bar{x}_{jn}), & h = -1, -2, \ldots, -(n-1), \end{cases} \tag{6.4.2}$$

where the unknown means are estimated by $\bar{x}_{in} = n^{-1} \sum_{t=1}^{n} X_{it}$. By (1.7.4), we can also write

$$\hat{\gamma}_{ij}(-h) = \hat{\gamma}_{ji}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_{jt} - \bar{x}_{jn})(X_{i,t+h} - \bar{x}_{in}), \qquad h = 0, 1, \ldots, n-1.$$

The corresponding estimators of the cross correlations are

$$\tilde{r}_{ij}(h) = [\tilde{\gamma}_{ii}(0)\tilde{\gamma}_{jj}(0)]^{-1/2}\tilde{\gamma}_{ij}(h) \tag{6.4.3}$$

and

$$\hat{r}_{ij}(h) = [\hat{\gamma}_{ii}(0)\hat{\gamma}_{jj}(0)]^{-1/2}\hat{\gamma}_{ij}(h) \tag{6.4.4}$$

By our earlier results (see Theorem 6.2.2) the estimation of the mean in estimator (6.4.2) introduces a bias that is $O(n^{-1})$ for time series with absolutely summable covariance function. The properties of the estimated cross covariances are analogous to the properties of the estimated autocovariance given in Theorems 6.2.1 and 6.2.3. To simplify the derivation, we only present the results for normal time series.

Theorem 6.4.1. Let the bivariate stationary normal time series $\mathbf{X}_t$ be such that

$$\sum_{h=-\infty}^{\infty} |\gamma_{ii}(h)| < \infty\,, \qquad i = 1, 2\,.$$

Then

$$\lim_{n\to\infty} n \operatorname{Cov}\{\hat{\gamma}_{12}(h), \hat{\gamma}_{12}(q)\} = \sum_{p=-\infty}^{\infty} \gamma_{11}(p)\gamma_{22}(p+q-h) + \sum_{p=-\infty}^{\infty} \gamma_{12}(p+q)\gamma_{21}(p-h)\,.$$

Proof. Letting the mean vector be zero and $h, q \geq 0$, we have

$$E\{\hat{\gamma}_{12}(h)\hat{\gamma}_{12}(q)\} = E\left\{\frac{1}{n^2}\sum_{t=1}^{n-h} X_{1t}X_{2,t+h}\sum_{s=1}^{n-q} X_{1s}X_{2,s+q}\right\} + O(n^{-1})\,,$$

where the remainder term enters because the mean is estimated. Evaluating the expectation, we have

$$\begin{aligned}
&E\{[\hat{\gamma}_{12}(h) - \gamma_{12}(h)][\hat{\gamma}_{12}(q) - \gamma_{12}(q)]\} \\
&\quad = \frac{1}{n^2}\sum_{t=1}^{n-h}\sum_{s=1}^{n-q} \gamma_{11}(s-t)\gamma_{22}(s-t+q-h) \\
&\quad\quad + \frac{1}{n^2}\sum_{t=1}^{n-h}\sum_{s=1}^{n-q} \gamma_{12}(s+q-t)\gamma_{21}(s-t-h) + O(n^{-1})\,.
\end{aligned}$$

Using Lemma 6.2.1 to take the limit, we have the stated result. ▲

Since the cross correlations are simple functions of the cross covariances, we can obtain a similar expression for the covariance of the estimated cross correlations.

Corollary 6.4.1.1. Given the bivariate stationary normal time series of Theorem 6.4.1,

$$\begin{aligned}
&\lim_{n\to\infty} n \operatorname{Cov}\{\hat{r}_{12}(h), \hat{r}_{12}(q)\} \\
&\quad = \sum_{p=-\infty}^{\infty} \{\rho_{11}(p)\rho_{22}(p+q-h) + \rho_{12}(p+q)\rho_{21}(p-h) \\
&\qquad - \rho_{12}(h)[\rho_{11}(p)\rho_{21}(p+q) + \rho_{22}(p)\rho_{21}(p-q)] \\
&\qquad - \rho_{12}(q)[\rho_{11}(p)\rho_{21}(p+h) + \rho_{22}(p)\rho_{21}(p-q)] \\
&\qquad + \rho_{12}(h)\rho_{12}(q)[\tfrac{1}{2}\rho_{11}^2(p) + \rho_{12}^2(p) + \tfrac{1}{2}\rho_{22}^2(p)]\}\,.
\end{aligned}$$

Proof. By Theorem 5.5.1 we may use the first term in Taylor's series to obtain the leading term in the covariance expression. Evaluating $\operatorname{Cov}\{\hat{\gamma}_{12}(h), \hat{\gamma}_{12}(q)\}$, $\operatorname{Cov}\{\hat{\gamma}_{12}(h), \frac{1}{2}[\hat{\gamma}_{11}(0) + \hat{\gamma}_{22}(0)]\}$, $\operatorname{Cov}\{\hat{\gamma}_{12}(q), \frac{1}{2}[\hat{\gamma}_{11}(0) + \hat{\gamma}_{22}(0)]\}$, and $\operatorname{Var}\{\frac{1}{2}[\hat{\gamma}_{11}(0) + \hat{\gamma}_{22}(0)]\}$ by the methods of Theorem 6.2.1, we obtain the conclusion. ▲

Perhaps the most important aspect of these rather cumbersome results is that the covariances are decreasing at the rate n^{-1}. Also, certain special cases are of interest. One working hypothesis is that the two time series are uncorrelated. If X_{1t} is a sequence of independent normal random variables, we obtain a particularly simple result.

Corollary 6.4.1.2. Let $\mathbf{X}_t$ be a bivariate stationary normal time series satisfying

$$\sum_{h=-\infty}^{\infty} |\gamma_{22}(h)| < \infty ,$$

$$\gamma_{11}(h) = \begin{cases} \sigma_1^2 , & h = 0 , \\ 0 , & \text{otherwise} , \end{cases}$$

and

$$\gamma_{12}(h) = 0 , \qquad \text{all } h .$$

Then

$$\lim_{n\to\infty} n \operatorname{Cov}\{\hat{r}_{12}(h), \hat{r}_{12}(q)\} = \rho_{22}(q-h) .$$

In the null case, the variance of the estimated cross correlation is approximately n^{-1}, and the correlation between estimated cross correlations is the autocorrelation of X_{2t} multiplied by n^{-1}. If the two time series are independent and neither time series is autocorrelated, then the estimated cross correlations are uncorrelated with an approximate variance of n^{-1}.

It is possible to demonstrate that the sample covariances are consistent estimators under much weaker conditions.

Lemma 6.4.1. Let

$$(X_{1t}, X_{2t}) = \left(\sum_{j=0}^{\infty} \alpha_j e_{1,t-j}, \sum_{j=0}^{\infty} \beta_j e_{2,t-j} \right) ,$$

where $\{\alpha_j\}$ and $\{\beta_j\}$ are absolutely summable and $\{\mathbf{e}_t\} = \{(e_{1t}, e_{2t})'\}$ is a sequence of independent $(\mathbf{0}, \mathbf{\Sigma})$ random variables. Assume $E\{|e_{it}|^{2+\delta}\} < L$ for $i = 1, 2$ and some

$\delta > 0$, or that the $\mathbf{e}_t$ are identically distributed. Then

$$n^{-1} \sum_{t=1}^{n-h} X_{1t}, X_{2,t+h} \xrightarrow{P} \gamma_{X_1X_2}(h), \qquad h = 0, 1, \ldots .$$

Proof. Define

$$(Y_{1t}, Y_{2t}) = \left(\sum_{j=0}^{k} \alpha_j e_{1,t-j}, \sum_{j=0}^{k} \beta_j e_{2,t-j} \right),$$

$$(D_{1t}, D_{2t}) = \left(\sum_{j=k+1}^{\infty} \alpha_j e_{1,t-j}, \sum_{j=k+1}^{\infty} \beta_j e_{2,t-j} \right),$$

fix h, and consider

$$\frac{1}{n} \sum_{t=1}^{n-h} Y_{1t} Y_{2,t+h} = \frac{1}{n} \sum_{t=1}^{n-h} \sum_{j=0}^{k} \sum_{i=0}^{k} \alpha_j \beta_i e_{1,t-j} e_{2,t-i+h} .$$

If $j \neq i - h$,

$$\operatorname{Var}\left\{ \frac{1}{n} \sum_{t=1}^{n-h} \alpha_j \beta_i e_{1,t-j} e_{2,t-i+h} \right\} = \frac{n-h}{n^2} \alpha_j^2 \beta_i^2 \sigma_1^2 \sigma_2^2 ,$$

where σ_i^2 is the variance of e_{it}. If $j = i - h$ and $\sigma_{12} = E\{e_{1t} e_{2t}\}$, then

$$\frac{1}{n} \sum_{t=1}^{n-h} \alpha_j \beta_{j+h} e_{1,t-j} e_{2,t-j} \xrightarrow{P} \alpha_j \beta_{j+h} \sigma_{12}$$

by the weak law of large numbers. [See, for example, Chung (1968, p. 104).] Now

$$\left| \frac{1}{n} \sum_{t=1}^{n-h} X_{1t} X_{2,t+h} - \frac{1}{n} \sum_{t=1}^{n-h} Y_{1t} Y_{2,t+h} \right|$$
$$\leq \left| \frac{1}{n} \sum_{t=1}^{n-h} Y_{1t} D_{2t} \right| + \left| \frac{1}{n} \sum_{t=1}^{n-h} Y_{2t} D_{1t} \right| + \left| \frac{1}{n} \sum_{t=1}^{n-h} D_{1t} D_{2t} \right| .$$

By Chebyshev's inequality

$$P\left\{ \left| \frac{1}{n} \sum_{t=1}^{n-h} D_{1t}^2 \right| > \epsilon \right\} \leq \sum_{j=k+1}^{\infty} \alpha_j^2 \sigma_1^2 \epsilon^{-1} ,$$

and it follows that

$$\operatorname*{plim}_{k\to\infty} \frac{1}{n} \sum_{t=1}^{n-h} Y_{1t} D_{2t} = 0 ,$$

$$\operatorname*{plim}_{k\to\infty} \frac{1}{n} \sum_{t=1}^{n-h} D_{1t} D_{2t} = 0 ,$$

uniformly in n. Convergence in probability implies convergence in distribution, and by an application of Lemma 6.3.1 we have that $n^{-1}\sum_{t=1}^{n-h} X_{1t}X_{2,t+h}$ converges in distribution to the constant $\gamma_{X_1X_2}(h)$. The result follows by Lemma 5.2.1. ▲

Theorem 6.4.2. Let $\{e_{1t}\}$ and $\{X_t\}$ be independent time series, where $\{e_{1t}\}$ is a sequence of independent $(0, \sigma_1^2)$ random variables with uniformly bounded third moment and $\{X_t\}$ is defined by

$$X_t = \sum_{j=0}^{\infty} \alpha_j e_{2,t-j},$$

where $\sum_{j=0}^{\infty} |\alpha_j| < \infty$ and $\{e_{2t}\}$ is a sequence of independent $(0, \sigma_2^2)$ random variables with uniformly bounded third moment. Then, for fixed $h > 0$,

$$n^{1/2}\hat{r}_{12}(h) \xrightarrow{\mathscr{L}} N(0, 1).$$

Proof. We write

$$n^{1/2}\hat{r}_{12}(h) = \left[\left(n^{-1}\sum_{t=1}^{n} e_{1t}^2\right)^{1/2}\left(n^{-1}\sum_{t=1}^{n} X_t^2\right)^{1/2}\right]^{-1} n^{-1/2}\sum_{t=1}^{n-h} e_{1t}X_{t+h}$$
$$+ O_p(n^{-1/2})$$

and note that, by Lemma 6.4.1,

$$\operatorname{plim} n^{-1}\sum_{t=1}^{n} e_{1t}^2 = \sigma_1^2,$$
$$\operatorname{plim} n^{-1}\sum_{t=1}^{n} X_t^2 = \gamma_{XX}(0).$$

For fixed h, the time series

$$Z_{th} = e_{1t}X_{t+h}$$

is weakly stationary with

$$E\{Z_{th}\} = 0,$$
$$E\{Z_{th}^2\} = \sigma_1^2\gamma_{XX}(0),$$

and bounded third moment. Asymptotic normality follows by a modest extension of Theorem 6.3.3. ▲

Example 6.4.1. In Table 6.4.1 we present the sample autocorrelations and

Table 6.4.1. Sample Correlation Functions for Suspended Sediment in Des Moines River at Boone, Iowa and Saylorville, Iowa

h	Autocorrelation Boone $\hat{r}_{11}(h)$	Autocorrelation Saylorville $\hat{r}_{22}(h)$	Cross correlation Boone–Saylorville $\hat{r}_{12}(h)$
−12	0.20	0.10	0.10
−11	0.19	0.13	0.07
−10	0.22	0.16	0.08
−9	0.25	0.16	0.08
−8	0.29	0.16	0.13
−7	0.29	0.17	0.18
−6	0.29	0.15	0.21
−5	0.32	0.18	0.21
−4	0.39	0.27	0.23
−3	0.48	0.42	0.30
−2	0.62	0.60	0.40
−1	0.76	0.81	0.53
0	1.00	1.00	0.64
1	0.76	0.81	0.74
2	0.62	0.60	0.67
3	0.48	0.42	0.53
4	0.39	0.27	0.42
5	0.32	0.18	0.32
6	0.29	0.15	0.26
7	0.29	0.17	0.26
8	0.29	0.16	0.25
9	0.25	0.16	0.29
10	0.22	0.16	0.31
11	0.19	0.13	0.28
12	0.20	0.10	0.33

cross correlations for the bivariate time series $\mathbf{Y}_t = (Y_{1t}, Y_{2t})'$, where Y_{1t} is the logarithm of suspended sediment in the water of the Des Moines River at Boone, Iowa, and Y_{2t} is the logarithm of suspended sediment in the water at Saylorville, Iowa. Saylorville is approximately 48 miles downstream from Boone. The sample data were 205 daily observations collected from April to October 1973.

There are no large tributaries entering the Des Moines River between Boone and Saylorville, and a correlation between the readings at the two points is expected. Since Saylorville is some distance downstream, the correlation pattern should reflect the time required for water to move between the two points. In fact, the largest sample cross correlation is between the Saylorville reading at time $t + 1$ and the Boone reading at time t. Also, estimates of $\gamma_{12}(h)$, $h > 0$, are consistently larger than the estimates of $\gamma_{12}(-h)$.

The Boone time series was discussed in Section 4.5. There it was assumed that the time series Y_{1t} could be represented as the sum of the "true process" and a measurement error. The underlying true value X_{1t} was assumed to be a first order autoregressive process with parameter 0.81. If we define the time series

$$W_{3t} = Y_{1t} - 0.81Y_{1,t-1},$$
$$W_{4t} = Y_{2t} - 0.81Y_{2,t-1},$$

the transformed true process for Boone is a sequence of uncorrelated random variables, although the observed time series W_{3t} will show a small negative first order autocorrelation.

Table 6.4.2 contains the first few estimated correlations for $\mathbf{W}_t$. Note that the estimated cross correlation at zero is quite small. Under the null hypothesis that the cross correlations are zero and that the autocorrelations of W_{3t} and W_{4t} are zero

Table 6.4.2. Sample Correlation Functions for Transformed Suspended Sediment in Des Moines River at Boone, Iowa and Saylorville, Iowa

h	Autocorrelation W_{3t}	Autocorrelation W_{4t}	Cross Correlation W_{3t} with W_{4t}
−12	0.05	0.05	0.04
−11	−0.06	0.02	−0.07
−10	0.03	0.08	0.05
−9	0.01	0.01	0.11
−8	0.11	0.04	0.03
−7	0.05	0.07	0.08
−6	−0.03	−0.06	0.10
−5	0.04	−0.12	0.03
−4	−0.01	−0.07	−0.05
−3	−0.07	0.02	−0.01
−2	0.05	−0.04	−0.04
−1	−0.17	0.15	0.06
0	1.00	1.00	0.06
1	−0.17	0.15	0.41
2	0.05	−0.04	0.24
3	−0.07	0.02	−0.02
4	−0.01	−0.07	0.01
5	0.04	−0.12	0.04
6	−0.03	−0.06	−0.10
7	0.05	0.07	0.08
8	0.11	0.04	−0.08
9	0.01	0.01	0.05
10	0.03	0.08	0.16
11	−0.06	0.02	0.08
12	0.05	0.05	0.01

after a lag of two, the estimated variance of the sample cross correlations is

$$
\begin{aligned}
\hat{\mathrm{V}}\mathrm{ar}\{\hat{r}_{34}(h)\} &= \frac{1}{n} \sum_{p=-2}^{2} \hat{r}_{33}(p)\hat{r}_{44}(p) \\
&= \frac{1}{205} [1 + 2(-0.17)(0.15) + 2(0.05)(-0.04)] \\
&= 0.0046 .
\end{aligned}
$$

Under these hypotheses, the estimated standard error of the estimated cross correlation is 0.068.

The hypothesis of zero cross correlation is rejected by the estimates $\hat{r}_{34}(1)$, $\hat{r}_{34}(2)$, since they are several times as large as the estimated standard error. The two nonzero sample cross correlations suggest that the input–output model is more complicated than that of a simple integer delay. It might be a simple delay of over one day, or it is possible that the mixing action of moving water produces a more complicated lag structure. ▲▲

REFERENCES

Section 6.1. Grenander and Rosenblatt (1957), Hannan (1970), Parzen (1958, 1962).

Section 6.2. Anderson (1971), Bartlett (1946, 1966), Hart (1942), Kendall (1954), Kendall and Stuart (1966), Mariott and Pope (1954), von Neumann (1941, 1942).

Section 6.3. Anderson (1959, 1971), Anderson and Walker (1964), Diananda (1953), Eicker (1963), Hannan and Heyde (1972), Hoeffding and Robbins (1948), Moran (1947).

Section 6.4. Bartlett (1966), Box, Jenkins, and Reinsel (1994), Hannan (1970).

EXERCISES

1. Let $Y_t = \mu + X_t$, where $X_t = e_t + 0.4e_{t-1}$ and the e_t are uncorrelated $(0, \sigma^2)$ random variables. Compute the variance of $\bar{y}_n = n^{-1} \sum_{t=1}^{n} Y_t$. What is $\lim_{n\to\infty} n \operatorname{Var}\{\bar{y}_n\}$?

2. Let the time series $\{Y_t: t \in (1, 2, \ldots)\}$ be defined by

$$Y_t = \alpha + \rho Y_{t-1} + e_t ,$$

where Y_0 is fixed, $\{e_t: t \in (1, 2, \ldots)\}$ is a sequence of independent $(0, \sigma^2)$ random variables, and $|\rho| < 1$. Find $E\{Y_t\}$ and $\operatorname{Var}\{Y_t\}$. Show that Y_t satisfies the conditions of Theorem 6.1.1. What value does the sample mean of Y_t converge to?

3. Let $X_t = e_t + 0.5e_{t-1}$, where the e_t are normal independent $(0, \sigma^2)$ random variables. Letting

$$\tilde{\gamma}_X(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} X_t X_{t+h},$$
$$\tilde{r}_X(h) = \tilde{\gamma}_X(h)/\tilde{\gamma}_X(0),$$

find $\operatorname{Var}\{\tilde{\gamma}_X(h)\}$ for $h = 0, 1, 2, 3$; $\operatorname{Cov}\{\tilde{\gamma}_X(0), \tilde{\gamma}_X(h)\}$ for $h = 1, 2, 3$; $\operatorname{Var}\{\tilde{r}_X(h)\}$ for $h = 0, 1, 2, 3$; and $\operatorname{Cov}\{\tilde{r}_X(1), \tilde{r}_X(h)\}$ for $h = 2, 3, 4$.

4. Evaluate (6.2.4) for the first order autoregressive process

$$X_t = \rho X_{t-1} + e_t,$$

where $|\rho| < 1$ and the e_t are normal independent $(0, \sigma^2)$ random variables.

5. Given the finite moving average

$$X_t = \sum_{j=1}^{M} a_j e_{t-j},$$

where the e_t are normal independent $(0, \sigma^2)$ random variables, is there a distance $d = h - q$ such that $\tilde{\gamma}(h)$ and $\tilde{\gamma}(q)$ are uncorrelated?

6. Use the realization (10, 1, 10) and equation (6.2.2) to construct the estimated (3×3) covariance matrix for a realization of size 3. Show that the resulting matrix is not positive definite. Use the fact that

$$\frac{1}{n} \sum_{t=1}^{n-h} (X_t - \bar{x}_n)(X_{t+h} - \bar{x}_n) = \frac{1}{n} \sum_{j=1}^{2n-1} Z_{mj} Z_{m+h,j}$$

for $m = 0, 1, \ldots, n-1$, $h = 0, 1, \ldots, n-1$, where, for $j = 1, 2, \ldots, 2n-1$,

$$Z_{mj} = \begin{cases} X_{j-m} - \bar{x}_n, & j = m+1, m+2, \ldots, m+n, \\ 0 & \text{otherwise}, \end{cases}$$

to prove that (6.2.3.) yields an estimated covariance matrix that is always positive semidefinite.

7. Give the variance of $\bar{x}_n$, $n = 1, 2, \ldots$, for $\{X_t : t \in (1, 2, \ldots)\}$ defined by

$$X_t = \mu + e_t - e_{t-1},$$

where $\{e_t : t \in (0, 1, 2, \ldots)\}$ is a sequence of independent identically distributed $(0, \sigma^2)$ random variables. Do you think there is a function w_n such that $w_n(\bar{x}_n - \mu) \xrightarrow{\mathscr{L}} N(0, 1)$?

8. Prove the following result, which is used in Theorem 6.3.4. If the sequence $\{c_t\}$ is such that

$$\lim_{n\to\infty}\left(\sum_{t=1}^{n} c_t^2\right)^{-1} c_n^2 = 0,$$

then

$$\lim_{n\to\infty}\sup_{1\le t\le n}\left(\sum_{j=1}^{n} c_j^2\right)^{-1} c_t^2 = 0.$$

9. Prove Corollary 6.3.5.

10. The data in the accompanying table are the average weekly gross hours per production worker on the payrolls of manufacturing establishments (seasonally adjusted).

	Quarter			
Year	I	II	III	IV
1948	40.30	40.23	39.97	39.70
1949	39.23	38.77	39.23	39.30
1950	39.70	40.27	40.90	40.97
1951	40.90	40.93	40.43	40.37
1952	40.63	40.33	40.60	41.07
1953	41.00	40.87	40.27	39.80
1954	39.53	39.47	39.60	39.90
1955	40.47	40.73	40.60	40.93
1956	40.60	40.30	40.27	40.47
1957	40.37	39.97	39.80	39.13
1958	38.73	38.80	39.40	39.70
1959	40.23	40.53	40.20	40.03
1960	40.17	39.87	39.63	39.07
1961	39.27	39.70	39.87	40.40
1962	40.27	40.53	40.47	40.27
1963	40.37	40.40	40.50	40.57
1964	40.40	40.77	40.67	40.90
1965	41.27	41.10	41.03	41.30
1966	41.53	41.47	41.33	41.10
1967	40.60	40.43	40.67	40.70
1968	40.60	40.63	40.83	40.77
1969	40.57	40.73	40.63	40.53
1970	40.17	39.87	39.73	39.50
1971	39.80	39.39	39.77	40.07
1972	40.30	40.67	40.67	40.87

Sources: *Business Statistics*, 1971, pp. 74 and 237, and *Survey of Current Business*, Jan. 1972, Jan. 1973. The quarterly data are the averages of monthly data.

(a) Estimate the covariance function $\gamma(h)$, assuming the mean unknown.
(b) Estimate the correlation function $\rho(h)$, assuming the mean unknown.
(c) Using large sample theory, test the hypothesis H_0: $\rho(1) = 0$, assuming $\rho(h) = 0$, $h > 1$.

11. Using Hart's (1942) tables for the percentage points of d_v or Anderson's (1971) tables for r_v, obtain the percentage points for $t_v = r_v(n+1)^{1/2}(1 - r_v^2)^{-1/2}$ for $n = 10$ and 15. Compare these values with the percentage points of Student's t with 13 and 18 degrees of freedom.

12. Using the truncation argument of Theorem 6.3.3, complete the proof of Theorem 6.4.2 by showing that

$$(n-h)^{-1/2} \sum_{t=1}^{n-h} Z_{th}$$

converges in distribution to a normal random variable.

13. Denoting the data of Exercise 10 by X_{1t} and that of Table 6.3.1 by X_{2t}, compute the cross covariance and cross correlation functions. Define

$$Y_{1t} = X_{1t} - 1.53X_{1,t-1} - 0.66X_{1,t-2},$$
$$Y_{2t} = X_{2t} - 1.53X_{2,t-1} - 0.66X_{2,t-2}.$$

Compute the cross covariance and cross correlation functions for (Y_{1t}, Y_{2t}). Plot the cross correlation function of (Y_{1t}, Y_{2t}). Compute the variance of $\hat{r}_{Y_1 Y_2}(h)$ under the assumption that Y_{1t} is a sequence of uncorrelated random variables. Plot the standard error on your figure.

14. Let X_t be a time series with positive continuous spectral density $f_X(\omega)$.
(a) Show that, given $\epsilon > 0$, one may define two moving average time series W_{1t} and W_{2t} with spectral densities

$$f_{W_1}(\omega) = \frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{q_1} \beta_{1j} e^{-i\omega j}\right|^2,$$
$$f_{W_2}(\omega) = \frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{q_2} \beta_{2j} e^{-i\omega j}\right|^2$$

such that

$$f_X(\omega) - \epsilon \le f_{W_1}(\omega) \le f_X(\omega) \le f_{W_2}(\omega) \le f_X(\omega) + \epsilon.$$

(b) Let $\{a_t: t = 1, 2, \ldots\}$ be such that

$$\lim_{n\to\infty} \sum_{t=1}^{n} a_t^2 = \infty ,$$

$$\lim_{n\to\infty} \frac{\sum_{t=1}^{n} a_t a_{t+|h|}}{\sum_{t=1}^{n} a_t^2} = g(h) , \qquad h = 0, \pm 1, \pm 2, \ldots ,$$

$$\sum_{h=-\infty}^{\infty} |g(h)| < \infty .$$

Define $f_g(\omega) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} g(h) e^{-\iota\omega h}$. Show that

$$\lim_{n\to\infty} \operatorname{Var}\left\{ \left(\sum_{t=1}^{n} a_t^2 \right)^{-1/2} \sum_{t=1}^{n} a_t X_t \right\} = \int_{-\pi}^{\pi} f_X(\omega) f_g(\omega)\, d\omega .$$

See Exercise 4.15 and Grenander and Rosenblatt (1957, Chapter 7).

15. Prove Lemma 6.3.2.

16. Let $X_t = \sum_{j=0}^{q} b_j e_{t-j}$. Show that

$$\bar{x}_n = n^{-1}\left(\sum_{j=0}^{q} b_j \sum_{t=1}^{n} e_t + \sum_{s=1}^{q} \sum_{j=s}^{q} b_j e_{1-s} - \sum_{s=0}^{q-1} \sum_{j=s+1}^{q} b_j e_{n-s} \right).$$

17. Let Y_t be the stationary autoregressive moving average

$$\sum_{j=0}^{p} a_j (Y_{t-j} - \mu) = \sum_{i=0}^{q} b_i e_{t-i} .$$

Prove that if $\sum_{i=0}^{q} b_i = 0$, then $V\{\bar{y}_n\} = O(n^{-2})$.

18. Let X_t be a covariance stationary time series satisfying

$$X_t = \sum_{j=0}^{\infty} \alpha_j e_{t-j} ,$$

where $\sum_{j=0}^{\infty} |\alpha_j| < \infty$, $\sum_{j=0}^{\infty} \alpha_j \neq 0$, and the e_t are uncorrelated $(0, \sigma^2)$ random variables. Show that

$$\sum_{h=-\infty}^{\infty} \gamma_X(h) = \left(\sum_{j=0}^{\infty} \alpha_j \right)^2 \sigma^2 .$$

19. Let

$$Y_t = \sum_{j=1}^{\infty} c_j a_{t-j} \quad \text{and} \quad X_t = \sum_{i=1}^{\infty} b_i e_{t-i} ,$$

where $\sum_{j=1}^{\infty} |c_j| < \infty$, $\sum_{i=1}^{\infty} |b_i| < \infty$, and

$$\begin{pmatrix} a_t \\ e_t \end{pmatrix} \sim \mathrm{II}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{aa} & \sigma_{ae} \\ \sigma_{ea} & \sigma_{ee} \end{pmatrix} \right).$$

Prove

$$n^{-1} \sum_{t=1}^{n} Y_t X_t \xrightarrow{P} \sum_{j=1}^{\infty} c_j b_j \sigma_{ae} .$$

20. Let Y_t be a stationary time series with absolutely summable covariance function, and let

$$X_t = \sum_{j=0}^{\infty} \alpha_j Y_{t-j} ,$$

where $|\alpha_j| < M\lambda^j$ for some $M < \infty$ and $0 < \lambda < 1$. Show that

$$n^{-1} \sum_{t=1}^{n} \left(\sum_{j=0}^{t-1} \alpha_j Y_{t-j} \right)^2 = n^{-1} \sum_{t=1}^{n} X_t^2 + O_p(n^{-1}) .$$

21. Construct an alternative proof for Corollary 6.3.3 by averaging both sides of the defining equation to obtain

$$\left(\sum_{j=0}^{p} a_j \right)(\bar{y}_n - \mu) = \left(\sum_{i=0}^{q} b_i \right)\bar{e}_n + \sum_{i=0}^{q} b_i n^{-1} \sum_{j=1}^{i} (e_{j-i} - e_{n-i+j})$$
$$- \sum_{j=0}^{p} a_j n^{-1} \sum_{k=1}^{j} (Y_{k-j} - Y_{n-j+k}) .$$

22. For a moving average of order q, show that $n^{1/2}\hat{r}(h) \xrightarrow{\mathcal{L}} N[0, 1 + 2\sum_{k=1}^{q} \rho^2(k)]$ for $|h| > q$.

23. Prove the following.

Result 6.3.1. Let e_t be a sequence of random variables, let $\mathscr{A}_t$ be the sigma-field generated by $\{e_s : s \leq t\}$, and assume the e_t satisfy

(a) $E\{(e_t, e_t^2 - \sigma^2, e_t^4 - \eta\sigma^4) \mid \mathscr{A}_{t-1}\} = (0, 0, 0)$ a.s.,

(b) $\operatorname{plim}_{n\to\infty} n^{-1} \sum_{t=1}^{n} E\{e_t^3 e_{t-h} \mid \mathscr{A}_{t-1}\} = \lim_{n\to\infty} n^{-1} \sum_{t=1}^{n} E\{e_t^3 e_{t-h}\} = 0$ for every fixed $h > 0$,

(c) $E\{|e_t|^{4+2\delta}\} < M < \infty$ for some $\delta > 0$.

Then, for any fixed K,

$$n^{1/2}[\tilde{\gamma}_e(0)-\sigma^2,\ \tilde{\gamma}_e(1),\ldots,\tilde{\gamma}_e(K)]' \xrightarrow{\mathscr{L}} N[0,\sigma^4 \operatorname{diag}(\eta-1,1,\ldots,1)]\,.$$

24. Using Result 6.3.1 of Exercise 23, prove the following.

Result 6.3.2. Let $Y_t=\Sigma_{i=0}^{q}\beta_i e_{t-i}$, where the e_t satisfy the conditions of Result 6.3.1. Then, for any fixed K,

$$n^{1/2}[\tilde{\gamma}_Y(0)-\gamma_Y(0),\ \tilde{\gamma}_Y(1)-\gamma_Y(1),\ldots,\tilde{\gamma}_Y(K)-\gamma_Y(K)]' \xrightarrow{\mathscr{L}} N(\mathbf{0},\mathbf{V})\,,$$

where the elements of $\mathbf{V}$ are defined by (6.2.4) of Theorem 6.2.1.