Chapter 5

Estimation of the mean, autocovariance and autocorrelation

Let $\{X_t\}$ be a stationary time series with mean μ, autocovariance function $\gamma(\cdot)$, and spectral density $f(\cdot)$. Now consider the problem of estimating the mean μ, ACVF $\gamma(\cdot)$ and ACF $\rho(\cdot) = \frac{\gamma(\cdot)}{\gamma(0)}$ from observations of $X_1, X_2, \ldots, X_n$.

5.1 Asymptotic Normality

For a large class of strictly linear time series, estimators of the mean μ, ACVF $\gamma(.)$ and ACF $\rho(.)$ will satisfy a central limit theorem and, asymptotically, the distribution, appropriately rescaled, will be normal.

Definition 5.1 (Asymptotic Normality). *Let $Y_1, Y_2, \ldots$ be a sequence of random variables. They are said to be* asymptotically normal, *written $Y_n \sim AN(\mu_n, \sigma_n^2)$ if and only if $\mu_n = \mathbb{E}[Y_n]$ and $\sigma_n^2 = Var(Y_n)$ for each n and*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{Y_n - \mu_n}{\sigma_n} \leq x\right) = \Phi(x),$$

where $\Phi(x) = \mathbb{P}(Z \leq x)$, $Z \sim N(0,1)$.

Let $\underline{Y}_1, \underline{Y}_2, \ldots$ be a sequence of random k-vectors. The sequence is said to be asymptotically normal, written $\underline{Y}_n \sim AN(\underline{\mu}_n, \Sigma_n)$ if and only if $\underline{\mu}_n = \mathbb{E}[\underline{Y}_n]$ and $\Sigma_{n;i,j} = Cov(Y_{n,i}, Y_{n,j})$ (Σ_n is the covariance matrix of $\underline{Y}_n$ for each n) and

$$\underline{\lambda}'\underline{Y}_n \sim AN(\underline{\lambda}^t \underline{\mu}_n, \underline{\lambda}^t \Sigma_n \underline{\lambda}) \qquad \forall \underline{\lambda} \in \mathbb{R}^k.$$

5.2 Estimation of μ

The estimator $\overline{X}_n := \frac{1}{n}\sum_{j=1}^{n} X_j$ is the natural unbiased estimate of μ.

Theorem 5.2. *Let $\{X_t\}$ be a stationary time series with mean μ and autocovariance function $\gamma(\cdot)$ which satisfies $\lim_{h\to+\infty}|\gamma(h)| = 0$. Then*

$$Var(\overline{X}_n) \stackrel{n\to+\infty}{\longrightarrow} 0.$$

Suppose $\sum_{h=-\infty}^{\infty}|\gamma(h)| < +\infty$ and let f denote the spectral density. Then

$$n\,Var(\overline{X}_n) \stackrel{n\to+\infty}{\longrightarrow} \sum_{h=-\infty}^{\infty}\gamma(h) = 2\pi f(0)$$

Proof In all cases,

$$\begin{aligned} n\mathrm{Var}(\overline{X}_n) &= n\mathrm{Var}(\frac{1}{n}\sum_{j=1}^{n}X_j) = \frac{1}{n}\sum_{i,j=1}^{n}\mathrm{Cov}(X_i,X_j) = \frac{1}{n}\sum_{i,j=1}^{n}\gamma(i-j) = \gamma(0) + \frac{2}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\gamma(j-i) \\ &= \gamma(0) + \frac{2}{n}\sum_{i=1}^{n}\sum_{h=1}^{n-i}\gamma(h) = \gamma(0) + \frac{2}{n}\sum_{h=1}^{n}(n-h)\gamma(h). \end{aligned}$$

For the first statement, since $|\gamma(h)| \stackrel{|h|\to+\infty}{\longrightarrow} 0$, therefore

$$\left|\frac{1}{n}\sum_{h=1}^{n}(1-\frac{h}{n})\gamma(h)\right| \le \frac{1}{n}\sum_{h=1}^{n}|\gamma(h)| = \frac{1}{n}\sum_{h=1}^{\lceil n^{1/2}\rceil}|\gamma(h)| + \frac{1}{n}\sum_{h=\lceil n^{1/2}\rceil+1}^{n}|\gamma(h)| \le \frac{\gamma(0)}{n^{1/2}} + \sup_{h>n^{1/2}}|\gamma(h)| \stackrel{n\to+\infty}{\longrightarrow} 0$$

so that

$$\mathrm{Var}(\overline{X}_n) = \frac{1}{n}\gamma(0) + \frac{2}{n}\sum_{h=1}^{n}(1-\frac{h}{n})\gamma(h) \stackrel{n\to+\infty}{\longrightarrow} 0.$$

For the second statement, suppose that $\sum_{h=-\infty}^{\infty}|\gamma(h)| < \infty$. Then it follows firstly, that for any N $\sum_{|h|\le N}\left(1-\frac{|h|}{n}\right)\gamma(h) \stackrel{n\to+\infty}{\longrightarrow} \sum_{|h|\le N}\gamma(h)$ and secondly that

$$\left|\sum_{|h|\ge N+1}\left(1-\frac{|h|}{n}\right)\gamma(h)\right| \le \sum_{|h|\ge N+1}|\gamma(h)| \stackrel{N\to+\infty}{\longrightarrow} 0,$$

from which it follows that

$$n\mathrm{Var}(\overline{X}_n) \stackrel{n\to+\infty}{\longrightarrow} \sum_{h=-\infty}^{\infty}\gamma(h) = 2\pi f(0).$$

□

The sample average $\overline{X}$ is an unbiased estimator of μ and we also have *consistency*; directly from Chebyshev, $\mathrm{Var}(\overline{X}_n) \stackrel{n\to+\infty}{\longrightarrow} 0$ implies convergence in probability. When $\sum|\gamma(h)| < +\infty$, we also have the *rate* of convergence and $n\mathrm{Var}(\overline{X}_n) \stackrel{n\to+\infty}{\longrightarrow} 2\pi f(0)$. We can go further, and show that if $|\gamma(h)|$ decays quickly enough, then $\overline{X}_n$ is asymptotically normal.

Theorem 5.3. *Let $\{X_t\}$ be a strictly linear time series defined by*

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j} \qquad \{\epsilon_t\} \sim IID(0, \sigma^2)$$

where (ψ_j) satisfy $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ and $\sum_{j=-\infty}^{\infty} \psi_j \neq 0$. then

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sqrt{v}} \sim AN(0,1)$$

where

$$v = \sum_{h=-\infty}^{\infty} \gamma(h) = \sigma^2 \left(\sum_{j=-\infty}^{\infty} \psi_j \right)^2$$

and γ is the ACVF of $\{X_t\}$.

Since $\{X_t : t \in \mathbb{Z}\}$ is not an i.i.d. sequence, we first prove the result for m-dependent stationary processes, where for each t, $X_t \perp X_{t+n}$, $n \geq m+1$ and then extend to arbitrary linear stationary processes with well defined ACVF.

Theorem 5.4 (Central Limit Theorem for Strictly Stationary m-Dependent Sequences)**.** *Let $\{X_t\}$ be a strictly stationary m-dependent sequence of random variables. That is, $X_t \perp X_s$ for all s such that $|t-s| > m$. Let $\mu = 0$ and let γ denote the autocovariance function. Let $v_m = \gamma(0) + 2\sum_{j=1}^{m} \gamma(j)$ and suppose that $v_m \neq 0$. Then*

1. $\lim_{n\to+\infty} n\,Var(\overline{X}_n) = v_m$ *and*
2. $\overline{X}_n \sim AN\big(0, \frac{v_m}{n}\big)$.

Proof The first statement has already been dealt with. For the second statement (asymptotic normality), for each integer k such that $k > 2m$, let

$$Y_{nk} = \frac{1}{n^{1/2}} \left\{ (X_1 + \ldots + X_{k-m}) + \ldots + (X_{(r-1)k+1} + \ldots + X_{rk-m}) \right\}$$

where $r = \lfloor \frac{n}{k} \rfloor$. Since the series is m-dependent, removal of $X_{jk-m+1}, \ldots, X_{jk}$ for each j, splits it into independent pieces; $n^{1/2}Y_{nk}$ is the sum of r i.i.d. random variables, each with mean zero and variance:

$$\begin{aligned} R_{k-m} &= \mathrm{Var}(X_1 + \ldots + X_{k-m}) \\ &= \sum_{i=1}^{k-m} \sum_{j=1}^{k-m} \mathrm{Cov}(X_i, X_j) \\ &= (k-m)\gamma(0) + 2\sum_{i=1}^{k-m} \sum_{j=i+1}^{k-m} \gamma(j-i) \\ &= (k-m)\gamma(0) + 2\sum_{j=1}^{k-m} \gamma(j) \left(\sum_{i=1}^{k-m-j} 1 \right) \\ &= \sum_{|j|<k-m} (k-m-|j|)\gamma(j). \end{aligned}$$

By the Central Limit Theorem, therefore:

$$\frac{(n^{1/2}Y_{nk})}{(\lfloor \frac{n}{k} \rfloor R_{k-m})^{1/2}} \overset{n\to+\infty}{\longrightarrow_{(d)}} N(0,1)$$

which may be expressed as

$$Y_{nk} \overset{n\to+\infty}{\longrightarrow_{(d)}} N(0, \frac{1}{k}R_{k-m}) \overset{k\to+\infty}{\longrightarrow} N(0, v_m).$$

We now have to show that the pieces that have been omitted do not contribute as $k \to +\infty$ and $r \to +\infty$. That is, it remains to show that

$$\lim_{k\to+\infty} \limsup_{n\to+\infty} \mathbb{P}\left(\left|n^{1/2}\overline{X}_n - Y_{nk}\right| > \epsilon\right) = 0 \qquad \forall \epsilon > 0.$$

To establish this,

$$(n^{1/2}\overline{X}_n - Y_{nk}) = \frac{1}{n^{1/2}} \sum_{j=1}^{r-1} (X_{jk-m+1} + \ldots + X_{jk}) + \frac{1}{n^{1/2}}(X_{rk-m+1} + \ldots + X_n).$$

The terms are independent for $k > m$. Therefore:

$$\mathrm{Var}(n^{1/2}\overline{X}_n - Y_{nk}) = \frac{1}{n}\left((\lfloor \frac{n}{k} \rfloor - 1)R_m + R_{h(n)}\right),$$

$$h(n) = n - k\lfloor \frac{n}{k} \rfloor + m \qquad 0 \le h(n) \le k + m.$$

Therefore:

$$\limsup_{n\to+\infty} \mathrm{Var}(n^{1/2}\overline{X}_n - Y_{nk}) = \frac{1}{k}R_m$$

and, as $k \to +\infty$, the result follows by Chebyshev. □

We therefore have asymptotic normality for the mean of an $MA(q)$ process for any $q < +\infty$. This is the key point for proving Theorem 5.3, since a stationary linear process can be approximated by a $2m+1$ dependent process.

Proof of Theorem 5.3 Let X_{tm} denote the $2m+1$ dependent approximation defined by:

$$X_{tm} = \mu + \sum_{j=-m}^{m} \psi_j \epsilon_{t-j} \qquad \{\epsilon_t\} \sim \mathrm{IID}(0, \sigma^2)$$

Set

$$Y_{nm} = \overline{X}_{nm} = \frac{1}{n}\sum_{t=1}^{n} X_{tm}.$$

It follows from the previous result that

$$\sqrt{n}(Y_{nm}-\mu) \to_d N\left(0, \sigma^2 \left(\sum_{j=-m}^{m} \psi_j\right)^2\right) \overset{m\to+\infty}{\longrightarrow} N\left(0, \sigma^2 \left(\sum_{j=-\infty}^{\infty} \psi_j\right)^2\right).$$

A direct computation, using $\sum_{j=-\infty}^{\infty} |\psi_j| < +\infty$ gives:

$$\begin{aligned}
\sup_{m_2\geq m_1} \mathrm{Var}\left(n^{1/2}(Y_{nm_2}-Y_{nm_1})\right) &= \sup_{m_2\geq m_1} n\mathrm{Var}\left(\frac{1}{n}\sum_{t=1}^{n}\sum_{|j|=m_1+1}^{m_2} \psi_j \epsilon_{t-j}\right) \\
&= \sup_{m_2\geq m_1+1} n\mathrm{Var}\left(\frac{1}{n}\sum_{s=m_1+2}^{m_2+n} \epsilon_s \sum_{t=s-m_2}^{s-m_1-1} \psi_{t-s} + \frac{1}{n}\sum_{s=1-m_2}^{n-m_1-1} \epsilon_s \sum_{t=s+m_1+1}^{s+m_2} \psi_{t-s}\right) \\
&= \frac{\sigma^2}{n}\left(\sum_{s=m_1+2}^{m_2+n} (\sum_{t=s-m_2}^{s-m_1-1} \psi_{t-s})^2 + \sum_{t=s-m_2}^{n-m_1-1} (\sum_{t=s+m_1+1}^{s+m_2} \psi_{t-s})^2\right) \\
&= \frac{\sigma^2(n+m_2-m_1-2)}{n}\left((\sum_{t=-m_2}^{-m_1-1} \psi_t)^2 + (\sum_{t=m_1+1}^{m_2} \psi_t)^2\right)
\end{aligned}$$

from which

$$\lim_{m_1\to+\infty} \sup_{m_2\geq m_1} \limsup_{n\to+\infty} \mathrm{Var}\left(n^{1/2}(Y_{nm_2}-Y_{nm_1})\right) = 0.$$

(We *first* let $n \to +\infty$.) From this, it now follows that

$$\sqrt{n}(\overline{X}_{mn}-\mu) \overset{n\to+\infty}{\longrightarrow} N(0, v_m) \overset{m\to+\infty}{\longrightarrow} N(0, v).$$

□

5.3 Estimation of $\gamma(\cdot)$

The estimators for γ and ρ which are used are:

$$\widehat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(X_t-\overline{X}_n)(X_{t+h}-\overline{X}_n), \quad 0 \leq h \leq n-1$$

and

$$\widehat{\rho}(h) = \frac{\widehat{\gamma}(h)}{\widehat{\gamma}(0)}$$

respectively. These estimators are clearly biased, but nevertheless are the estimators used, to ensure that the estimated covariance matrix

$$\widehat{\Gamma}_h = \begin{pmatrix} \widehat{\gamma}(0) & \dots & \widehat{\gamma}(h) \\ \vdots & & \\ \widehat{\gamma}(h) & \dots & \widehat{\gamma}(0) \end{pmatrix}$$

is non-negative definite. The estimators are asymptotically unbiased. The estimates $(\widehat{\gamma}(h))_{h=0}^{n}$ satisfy $\sum_h \gamma(h) = 0$ (exercise).

In the sequel, let $\underline{\gamma} = (\gamma(0), \gamma(1), \ldots, \gamma(h))'$, with similar notation for estimators of $\underline{\gamma}$. That is, $\underline{\widehat{\gamma}} = (\widehat{\gamma}(0), \ldots, \widehat{\gamma}(h))^t$. This section is devoted to the following result:

Theorem 5.5. *Let $\{X_t\}$ be a moving average process satisfying*

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j} \qquad \{\epsilon_t\} \sim IID(0, \sigma^2)$$

where $\sum_{j=-\infty}^{\infty} |\psi_j| < +\infty$ and $\mathbb{E}[\epsilon_t^4] = \eta\sigma^4 < +\infty$. Let γ be the autocovariance function of $\{X_t\}$. Then for any non negative integer h

$$\underline{\widehat{\gamma}} \sim AN\left(\widehat{\gamma}, \frac{1}{n} V\right)$$

where V is the covariance matrix with entries

$$v_{pq} = (\eta - 3)\gamma(p)\gamma(q) + \sum_{k=-\infty}^{\infty} (\gamma(k)\gamma(k - p + q) + \gamma(k + q)\gamma(k - p)).$$

Note that the 'noise' is i.i.d., but we do not require that it is Gaussian. If it is Gaussian, then (clearly) $\eta = 3$ and the first term in the expression for v_{pq} vanishes. We'll present the proof in stages.

If it is *known* that $\mu = 0$, then the estimator

$$\gamma^*(h) = \frac{1}{n} \sum_{j=1}^{n-h} X_j X_{j+j}$$

may be used for the ACVF. When trying to establish results, this is easier to work with; under conditions that $\overline{X}$ is asymptotically consistent, $\widehat{\gamma}$ and γ^* will have the same asymptotics. We'll use the notation

$\underline{\gamma}^* = (\gamma^*(0), \ldots, \gamma^*(h))$.

To establish that $\underline{\widehat{\gamma}}$ is asymptotically normal, we proceed in stages, firstly by considering $\underline{\gamma}^*$.

Theorem 5.6. *Let $\{X_t\}$ be a strictly linear time series with mean 0;*

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j} \qquad \{\epsilon_t\} \sim IID(0, \sigma^2)$$

satisfying $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ and $\mathbb{E}\left[\epsilon_t^4\right] = \eta\sigma^4 < \infty$. Let

$$\gamma^*(h) = \frac{1}{n} \sum_{t=1}^{n} X_t X_{t+h} \qquad h = 0, 1, 2, \ldots$$

Then

$$\lim_{n \to +\infty} n\, Cov(\gamma^*(p), \gamma^*(q)) = (\eta - 3)\gamma(p)\gamma(q) + \sum_{k=-\infty}^{\infty} (\gamma(k)\gamma(k - p + q) + \gamma(k + q)\gamma(k - p)).$$

Proof

$$\mathrm{Cov}(\gamma^*(p),\gamma^*(q)) = \mathrm{Cov}(\frac{1}{n}\sum_{j=1}^{n-p} X_jX_{j+q}, \frac{1}{n}\sum_{j=1}^{n-q} X_jX_{j+q}) = \frac{1}{n^2}\sum_{j=1}^{n-p}\sum_{k=1}^{n-q}\mathrm{Cov}(X_jX_{j+p}, X_kX_{k+q})$$

and

$$\begin{aligned}\mathrm{Cov}(X_jX_{j+p}, X_kX_{k+q}) &= \mathbb{E}[X_jX_{j+p}X_kX_{k+q}] - \mathrm{Cov}(X_j, X_{j+p})\mathrm{Cov}(X_k, X_{k+q})\\ &= \mathbb{E}[X_jX_{j+p}X_kX_{k+q}] - \gamma(p)\gamma(q)\end{aligned}$$

We'll use:

$$\mathbb{E}[X_jX_{j+p}X_kX_{k+q}] = \sum_{a_1,a_2,a_3,a_4} \psi_{a_1}\psi_{a_2}\psi_{a_3}\psi_{a_4}\mathbb{E}[\epsilon_{j-a_1}\epsilon_{j+p-a_2}\epsilon_{k-a_3}\epsilon_{k+q-a_4}].$$

Now,

$$\mathbb{E}[\epsilon_s\epsilon_t\epsilon_u\epsilon_v] = \begin{cases}\eta\sigma^4 & s=t=u=v\\ \sigma^2 & s=t\neq u=v, \quad s=u\neq t=v \quad s=v\neq t=u\\ 0 & \text{otherwise}\end{cases}$$

so that:

$$\begin{aligned}\mathbb{E}[X_tX_{t+p}X_sX_{s+q}] &= \sum_{i,j,k,l}\psi_i\psi_{j+p}\psi_k\psi_{l+q}\mathbb{E}[\epsilon_{t-i}\epsilon_{t-j}\epsilon_{s-k}\epsilon_{s-l}]\\ &= \sum_{i,j,k,l}\psi_i\psi_{j+p}\psi_{k-t+s}\psi_{l-t+s+q}\mathbb{E}[\epsilon_{t-i}\epsilon_{t-j}\epsilon_{t-k}\epsilon_{t-l}]\\ &= \eta\sigma^4\sum_i \psi_i\psi_{i+p}\psi_{i-t+s}\psi_{i-t+s+q} + \sigma^4\sum_{i\neq k}\psi_i\psi_{i+p}\psi_{k-t+s}\psi_{k-t+s+q}\\ &\quad+\sigma^4\sum_{i\neq j}\psi_i\psi_{i-t+s}\psi_{j+p}\psi_{j-t+s+q} + \sigma^2\sum_{i\neq j}\psi_i\psi_{i-t+s+q}\psi_{j+p}\psi_{j-t+s}\\ &= (\eta-3)\sigma^4\sum_i\psi_i\psi_{i+p}\psi_{i-t+s+p}\psi_{i-t+s+q} + \gamma(p)\gamma(q) + \gamma(t-s)\gamma(p+t-s-q) + \gamma(t-s-q)\gamma(p+t-s)\end{aligned}$$

It follows that

$$\begin{aligned}\mathbb{E}\left[\gamma^*(p)\gamma^*(q)\right] &= \frac{1}{n^2}\mathbb{E}\left[\sum_{s=1}^{n}\sum_{t=1}^{n} X_tX_{t+p}X_sX_{s+q}\right]\\ &= \frac{1}{n^2}\sum_{s=1}^{n}\sum_{t=1}^{n}(\gamma(p)\gamma(q) + \gamma(s-t)\gamma(s-t-p+q) + \gamma(s-t+q)\gamma(s-t-p)\\ &\quad+(\eta-3)\sigma^2\sum_i\psi_i\psi_{i+p}\psi_{i+s-t}\psi_{i+s-t+q}\Big)\end{aligned}$$

Now set $k=s-t$, change the order of summation and subtract $\gamma(p)\gamma(q)$ from each side. It follows that

$$\mathrm{Cov}(\gamma^*(p),\gamma^*(q)) = \frac{1}{n}\sum_{|k|<n}\left(1-\frac{|k|}{n}\right)T_k,$$

where

$$T_k = \gamma(k)\gamma(k-p+q) + \gamma(k+q)\gamma(k-p) + (\eta-3)\sigma^4\sum_i\psi_i\psi_{i+p}\psi_{i+k}\psi_{i+k+q}.$$

Note that

$$\sum_k|T_k| \leq 2\sum_k|\gamma(k)|^2 + |\eta-3|\sigma^4(\sum_k|\psi_k|^2)^2 < +\infty$$

since $\sum_k|\gamma(k)| < +\infty$ and hence $\sum_k|\gamma(k)|^2 < +\infty$ and $\sum|\psi_i| < +\infty$ and hence $\sum|\psi_i|^2 < +\infty$. It follows (using a standard ϵ δ argument, noting that for any N, $\lim_{n\to+\infty}\sum_{|k|<n\wedge N}\left(1-\frac{|k|}{n}\right)T_k \to \sum_{|k|<N}T_k$ and that for any $\epsilon>0$, N may be chosen such that $\sup_n\left|\sum_{n\vee N\leq|k|<N}\left(1-\frac{|k|}{n}\right)\right| < \epsilon$ that $\sum_{k=-\infty}^{\infty}T_k$ is well defined and that

$$\lim_{n\to+\infty} n\mathrm{Cov}(\gamma^*(p),\gamma^*(q)) = \sum_{k=-\infty}^{\infty}T_k = (\eta-3)\gamma(p)\gamma(q) + \sum_{k=-\infty}^{\infty}(\gamma(k)\gamma(k-p+q) + \gamma(k+q)\gamma(k-p))$$

as required. □

The covariance structure of $\underline{\gamma}^*$ has now been established; the following results show asymptotic normality. Firstly, Theorem 5.7 proves the result for a MA(2m + 1) process where m is finite; Theorem 5.9 extends it to strictly linear processes.

Theorem 5.7. *Let $\{X_t\}$ be the moving average process*

$$X_t = \sum_{j=-m}^{m} \psi_j \epsilon_{t-j} \qquad \{\epsilon_t\} \sim IID(0, \sigma^2),$$

where $\mathbb{E}[\epsilon_t^4] = \eta\sigma^2 < +\infty$. Let γ be the autocovariance function of X. Let $\underline{\gamma}^* = (\gamma^*(0), \ldots, \gamma^*(h))^t$ and $\underline{\gamma} = (\gamma(0), \ldots, \gamma(h))^t$. Then

$$\underline{\gamma}^* \sim AN\left(\underline{\gamma}, n^{-1}V\right)$$

where $V = (v_{pq})_{p,q=0,\ldots,h}$ is the covariance matrix with entries

$$v_{pq} = \sum_{k=-\infty}^{\infty} T_k = (\eta - 3)\gamma(p)\gamma(q) + \sum_{k=-\infty}^{\infty} \left(\gamma(k)\gamma(k-p+q) + \gamma(k+q)\gamma(k-p)\right).$$

Proof This follows directly from the central limit theorem (Theorem 5.3); consider the random $h+1$ vectors

$$\underline{Y}_t = (X_t X_t, X_t X_{t+1}, \ldots, X_t X_{t+h})^t$$

then $\underline{Y}_t$ is a strictly stationary $(2m+h)$ dependent sequence and, taking $X_t X_{t+i} = 0$ for $t + i \geq n$,

$$\frac{1}{n}\sum_{t=1}^{n} \underline{Y}_t = (\gamma^*(0), \ldots, \gamma^*(h)).$$

For any linear combination $\underline{\lambda}^t \underline{\gamma}^*$ such that $\underline{\lambda}^t V \underline{\lambda} > 0$, it follows that $\{\underline{\lambda}^t \underline{Y}_t\}$ satisfies the hypotheses of Theorem 5.3 and hence

$$\frac{\sqrt{n}(\underline{\lambda}^t \underline{\gamma}^* - \underline{\lambda}^t \underline{\gamma})}{\sqrt{\underline{\lambda}^t V \underline{\lambda}}} \longrightarrow_{(d)} N(0,1)$$

from which the result follows. □

Lemma 5.8. *Let $\{\underline{X}_n, n = 1, 2, \ldots\}$ and $\underline{Y}_{nj}, j = 1, 2, \ldots; n = 1, 2, \ldots$ be random k-vectors such that*

1. *$\underline{Y}_{nj} \to \underline{Y}_j$ for each $j = 1, 2, \ldots$*
2. *$\underline{Y}_j \to \underline{Y}$ as $j \to +\infty$ and*
3. *$\lim_{j\to+\infty} \lim_{n\to+\infty} \mathbb{P}\left(|\underline{X}_n - \underline{Y}_{nj}| > \epsilon\right) = 0$ for every $\epsilon > 0$.*

Then

$$\underline{X}_n \to Y \qquad n \to +\infty.$$

Proof Clear from the definitions. □

Theorem 5.9. *The result of Theorem 5.7 remains true for a process*

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j} \qquad \{\epsilon_j\} \sim IID(0, \sigma^2)$$

where $\sum_{j=-\infty}^{\infty} |\psi_j| < +\infty$ and $\mathbb{E}[\epsilon_t^4] = \eta\sigma^2 < +\infty$.

Proof The proof follows directly by applying Theorem 5.7 to the process

$$X_{tm} = \sum_{j=-m}^{m} \psi_j \epsilon_{t-j} \qquad \{\epsilon_j\} \sim IID(0, \sigma^2).$$

Let

$$\gamma_m^*(p) = \frac{1}{n} \sum_{t=1}^{n} X_{tm} X_{(t+p)m},$$

then

$$n^{1/2}(\underline{\gamma}_m^* - \gamma_m) \longrightarrow \underline{Y}_m$$

where γ_m is the autocovariance function of $\{X_{tm}\}$ and the vector notation is as in the previous theorem. Then $\underline{Y}_m \sim N(\underline{0}, V_m)$, where V_m is the covariance matrix from Theorem 5.7. As $m \to +\infty$, $V_m \to V$. The proof now follows by Lemma 5.8, provided it can be shown that

$$\lim_{m\to+\infty} \limsup_{n\to+\infty} \mathbb{P}\left(n^{1/2} \left|\gamma_m^*(p) - \gamma_m(p) - \gamma^*(p) + \gamma(p)\right| > \epsilon\right) = 0$$

for $p = 0, 1, \ldots, h$. This follows by Chebyshev; the probability is bounded by

$$\frac{n}{\epsilon^2} \mathrm{Var}\left(\gamma_m^*(p) - \gamma^*(p)\right) = \frac{1}{\epsilon^2}\left(n\mathrm{Var}(\gamma_m^*(p)) + n\mathrm{Var}(\gamma^*(p)) - 2n\mathrm{Cov}(\gamma_m^*(p), \gamma^*(p))\right).$$

Firstly,

$$\lim_{m\to+\infty} \lim_{n\to+\infty} n\mathrm{Var}(\gamma_m^*(p)) = \lim_{n\to+\infty} \mathrm{Var}(\gamma^*(p)) = v_{pp}$$

$$\lim_{m\to+\infty} \lim_{m\to+\infty} \lim_{n\to+\infty} n\mathrm{Cov}(\gamma_m^*(p), \gamma^*(p)) = v_{pp}$$

from which the result follows. □

Now, we have already established that $\overline{X}_n \to \mu$ in probability. Therefore, we should expect that $\widehat{\gamma} - \gamma^* \to 0$. The following proposition establishes that $\widehat{\gamma}$ is asymptotically normal, with the same asymptotic covariance structure as for γ^*.

Finally, we put all this together to get the main result (the asymptotic distribution of $\widehat{\gamma}$).

Proof of Theorem 5.5 For $0 \leq p \leq h$, it follows directly from the definition that

$$\begin{aligned} \widehat{\gamma}(p) &= \frac{1}{n}\sum_{t=1}^{n-p} X_t X_{t+p} - \overline{X}_n \frac{1}{n}\left(\sum_{t=1}^{n-p} X_t + \sum_{t=1}^{n-p} X_t\right) + \left(1 - \frac{p}{n}\right)\overline{X}_n^2 \\ &= \gamma^*(p) - \frac{1}{n}\sum_{t=n-p+1}^{n} X_t X_{t+p} - \overline{X}_n \frac{1}{n}\left(\sum_{t=1}^{n-p} X_t + \sum_{t=1}^{n-p} X_{t+p} + \left(1 - \frac{p}{n}\overline{X}_n\right)\right). \end{aligned}$$

From this, it follows directly that

$$\sqrt{n}(\gamma^*(p) - \widehat{\gamma}(p)) = n^{1/2}\overline{X}_n\left(\frac{1}{n}\sum_{t=1}^{n-p} X_{t+p} + \frac{1}{n}\sum_{t=1}^{n-p} X_t + \left(1 - \frac{p}{n}\right)\overline{X}_n\right) + \frac{1}{\sqrt{n}}\sum_{t=n-p+1}^{n} X_t X_{t+p}.$$

Now,

$$\frac{1}{\sqrt{n}}\mathbb{E}\left[\left|\sum_{t=n-p+1}^{n} X_t X_{t+p}\right|\right] \leq \frac{1}{\sqrt{n}} p\gamma(0) \overset{n\to+\infty}{\longrightarrow} 0.$$

Furthermore,

$$n^{1/2}\overline{X}_n \sim \text{AN}\left(0, \sigma^2 \left(\sum_{j=-\infty}^{\infty} \psi_j\right)^2\right)$$

By the weak law of large numbers,

$$\left(\frac{1}{n}\sum_{t=1}^{n-p} X_{t+p} + \frac{1}{n}\sum_{t=1}^{n-p} X_t + \left(1 - \frac{p}{n}\right)\overline{X}_n\right) \longrightarrow 0$$

in probability, from which the result follows. □

5.4 Estimating the Autocorrelation Function $\rho(.)$

Recall that $\widehat{\rho}(h) = \frac{\widehat{\gamma}(h)}{\widehat{\gamma(0)}}$ so that $\widehat{\rho}(.) = g(\widehat{\gamma}(.))$ for a suitable function g. Asymptotic results for $\widehat{\rho}$ are therefore obtained by applying the delta method to asymptotic results for $\widehat{\gamma}$.

For the ACF $\rho(\cdot)$, the η term disappears.

Theorem 5.10. *Let $\underline{\rho} = (\rho(1), \ldots, \rho(h))^t$ and $\underline{\widehat{\rho}} = (\widehat{\rho}(1), \ldots, \widehat{\rho}(h))^t$. Let $\{X_t\}$ satisfy:*

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j \epsilon_j \qquad \{\epsilon_t\} \sim IID(0, \sigma^2)$$

where $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ and $\mathbb{E}\left[\epsilon_t^4\right] < \infty$, then

$$\underline{\widehat{\rho}} \sim AN(\underline{\rho}, n^{-1}W)$$

where $W = (w_{ij})_{i,j=1,\ldots,h}$ is the covariance matrix whose entries are given by:

$$\begin{aligned} w_{ij} = \sum_{k=-\infty}^{\infty} \{&\rho(k+i)\rho(k+j) + \rho(k-i)\rho(k+j) \\ &+ 2\rho(i)\rho(j)\rho^2(k) - 2\rho(i)\rho(k)\rho(k+j) - 2\rho(j)\rho(k)\rho(k+i)\}. \end{aligned} \tag{5.1}$$

Proof This follows from the Delta Method (see the course Statistics): Let $g : \mathbb{R}^{h+1} \to \mathbb{R}$ be defined by

$$g((x_0, \ldots, x_h)^t) = \left(\left(\frac{x_1}{x_0}\right), \ldots, \left(\frac{x_h}{x_0}\right)\right)^t.$$

Let γ be the autocovariance of $\{X\}$. Then

$$\underline{\widehat{\rho}} = g(\underline{\widehat{\gamma}}) \sim \text{AN}\left(g(\underline{\gamma}), \frac{1}{n} DVD^t\right)$$

where D is the matrix of partial derivatives:

$$\begin{cases} \frac{\partial g_j}{\partial x_0} = -\frac{g_j}{x_0} & j = 1, \ldots, h \\ \frac{\partial g_j}{\partial x_k} = \frac{1}{x_0}\mathbf{1}_{x_k}(x_j) & (j,k) \in \{1, \ldots, h\}^2 \end{cases}$$

giving

$$D = \frac{1}{\gamma(0)} \begin{pmatrix} -\rho(1) & 1 & 0 & \dots & 0 \\ -\rho(2) & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ -\rho(h) & 0 & 0 & \dots & 1 \end{pmatrix}.$$

□

The expression (5.1) is called *Bartlett's formula.* It may be re-arranged to obtain the more convenient form:

$$w_{ij} = \sum_{k=1}^{\infty} \{\rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k)\} \times \{\rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k)\}. \tag{5.2}$$

The assumption $\mathbb{E}\left[\epsilon_t^4\right] < \infty$ is relaxed at the expense of a slightly stronger assumption on the sequence $\{\psi_j\}$.

Theorem 5.11. *If $\{X_t\}$ is a strictly linear time series where $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ and $\sum_{j=-\infty}^{\infty} \psi_j^2 |j| < \infty$, then*

$$\underline{\widehat{\rho}} \sim AN\left(\underline{\rho}, n^{-1}W\right)$$

where W is given by the previous theorem.

Proof Omitted. □

Using similar techniques, the asymptotic correlations between the estimators can be established;

$$\lim_{n\to+\infty} \text{Corr}\left(\widehat{\gamma}(i), \widehat{\gamma}(j)\right) = \frac{v_{ij}}{\sqrt{v_{ii}v_{jj}}} \qquad \lim_{n\to+\infty} \text{Corr}\left(\widehat{\rho}(i), \widehat{\rho}(j))\right) = \frac{w_{ij}}{\sqrt{w_{ii}w_{jj}}}.$$

5.5 The Ljung-Box Test

A very important situation is application to a series of residuals and deciding whether or not a 'white noise' model fits. As ever, $\rho(0) = 1$. If the series is white noise, then $\rho(1) = \rho(2) = \ldots = \rho(h) = 0$ for all $h \geq 1$. In this situation, the Bartlett formula (5.1) reduces to (for $1 \leq i \leq j < +\infty$):

$$w_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

so that $\widehat{\rho} \sim AN(0, \frac{1}{n}I)$. Hence, consider testing:

$$\begin{cases} H_0: & \text{The data are an observed random sample from WN}(0,\sigma^2) \\ H_1 & \text{The data are not an observed random sample from WN}(0,\sigma^2). \end{cases}$$

The test statistic

$$Q := n \sum_{j=1}^{h} \widehat{\rho}(j)^2 \stackrel{n \to +\infty}{\longrightarrow}_{(d)} \chi_h^2;$$

the null hypothesis is rejected for $Q > \chi^2_{h;\alpha}$ where $\chi^2_{h;\alpha}$ is the value such that $\mathbb{P}(X > \chi^2_{h,\alpha}) = \alpha$ for $X \sim \chi^2_h$.

The Ljung - Box test The Ljung–Box test (named after Greta M. Ljung and George E. P. Box) modifies the above test, to provide something that has greater accuracy for smaller n. They propose the test statistic

$$Q = n\,(n+2) \sum_{k=1}^{h} \frac{\widehat{\rho}^2(k)}{n-k}$$

where n is the sample size, $\widehat{\rho}(k)$ is the sample autocorrelation at lag k, and h is the number of lags being tested. Under the null hypothesis that the series is i.i.d. $N(0, \sigma^2)$, $Q \sim \chi^2_h$ as $n \to +\infty$.

The motivation is as follows: let $\epsilon_1, \epsilon_2, \ldots$ be i.i.d. $N(0, \sigma^2)$ and let

$$r_k = \frac{\sum_{j=1}^{n-k} \epsilon_j \epsilon_{j+k}}{\sum_{j=1}^{n} \epsilon_j^2}.$$

(so that $r_k = \widehat{\rho}(k)$). By scaling, we can take ϵ_j's i.i.d. $N(0,1)$. Clearly $\mathbb{E}[r_k] = 0$, so that $\mathbb{E}[r_k^2] = \mathrm{Var}(r_k)$. Some computation gives $\mathrm{Var}(r_k) = \frac{n-k)}{n(n+2)}$ so that the Ljung Box statistic has the same expectation as the χ^2 distribution.