- An extremely useful alternative approach touched on only briefly in this book is to express the series in terms of its Fourier components, which are sinusoidal waves of different frequencies (cf. Example 1.1.4). This approach is especially important in engineering applications such as signal processing and structural design. It is important, for example, to ensure that the resonant frequency of a structure does not coincide with a frequency at which the loading forces on the structure have a particularly large component.

1.4 Stationary Models and the Autocorrelation Function

We will explain this later

Loosely speaking, a time series $\{X_t, t = 0, \pm 1, \ldots\}$ is said to be stationary if it has statistical properties similar to those of the "time-shifted" series $\{X_{t+h}, t = 0, \pm 1, \ldots\}$, for each integer h. Restricting attention to those properties that depend only on the first- and second-order moments of $\{X_t\}$, we can make this idea precise with the following definitions.

As you read this definition, try to think (compute) the mean function for the "cos" example we had in class: $X_t = \cos(t/10) + \varepsilon_t$.

Definition 1.4.1 Let $\{X_t\}$ be a time series with $E(X_t^2) < \infty$. The **mean function** of $\{X_t\}$ is

$$\mu_X(t) = E(X_t).$$

This part you can ignore (I will explain in class)

The **covariance function** of $\{X_t\}$ is

$$\gamma_X(r, s) = \mathrm{Cov}(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))]$$

for all integers r and s.

Definition 1.4.2 $\{X_t\}$ is **(weakly) stationary** if

(i) $\mu_X(t)$ is independent of t,

and

(ii) $\gamma_X(t + h, t)$ is independent of t for each h.

In other words, the covariance function only depends on "h".

This will be one of the key definitions in the course

STRICT stationarity

Remark 1. Strict stationarity of a time series $\{X_t, t = 0, \pm 1, \ldots\}$ is defined by the condition that $(X_1, \ldots, X_n)$ and $(X_{1+h}, \ldots, X_{n+h})$ have the same joint distributions for all integers h and $n > 0$. It is easy to check that if $\{X_t\}$ is strictly stationary and $EX_t^2 < \infty$ for all t, then $\{X_t\}$ is also weakly stationary (Problem 1.3). Whenever we use the term *stationary* we shall mean weakly stationary as in Definition 1.4.2, unless we specifically indicate otherwise. □

You can skip this remark ~~for the~~ (I will go in detail over this in class).

Remark 2. In view of condition (ii), whenever we use the term covariance function with reference to a *stationary* time series $\{X_t\}$ we shall mean the function γ_X of *one* variable, defined by

$$\gamma_X(h) := \gamma_X(h, 0) = \gamma_X(t + h, t).$$

This is important: Restricting ourselves to weakly stationary models, really simplifies the task of computing the covariance function

The function $\gamma_X(\cdot)$ will be referred to as the autocovariance function and $\gamma_X(h)$ as its value at *lag* h. □

Definition 1.4.3

Let $\{X_t\}$ be a stationary time series. The **autocovariance function** (ACVF) of $\{X_t\}$ at lag h is

$$\gamma_X(h) = \text{Cov}(X_{t+h}, X_t).$$

The **autocorrelation function** (ACF) of $\{X_t\}$ at lag h is

$$\rho_X(h) \equiv \frac{\gamma_X(h)}{\gamma_X(0)} = \text{Cor}(X_{t+h}, X_t).$$

(this will translate into convenient estimation of the autocovariance function of an stationary process).

In the following examples we shall frequently use the easily verified **linearity property of covariances**, that if $EX^2 < \infty$, $EY^2 < \infty$, $EZ^2 < \infty$ and a, b, and c are any real constants, then

$$\text{Cov}(aX + bY + c, Z) = a\,\text{Cov}(X, Z) + b\,\text{Cov}(Y, Z).$$

Example 1.4.1 iid noise

If $\{X_t\}$ is iid noise and $E(X_t^2) = \sigma^2 < \infty$, then the first requirement of Definition 1.4.2 is obviously satisfied, since $E(X_t) = 0$ for all t. By the assumed independence,

$$\gamma_X(t + h, t) = \begin{cases} \sigma^2, & \text{if } h = 0, \\ 0, & \text{if } h \neq 0, \end{cases}$$

which does not depend on t. Hence iid noise with finite second moment is stationary. We shall use the notation

$$\{X_t\} \sim \text{IID}\left(0, \sigma^2\right)$$

to indicate that the random variables X_t are independent and identically distributed random variables, each with mean 0 and variance σ^2. □

This is kind on a trivial example, but you will see that i.i.d noise and white noise play a key role in the construction of stationary time series models.

Example 1.4.2 White noise

If $\{X_t\}$ is a sequence of uncorrelated random variables, each with zero mean and variance σ^2, then clearly $\{X_t\}$ is stationary with the same covariance function as the iid noise in Example 1.4.1. Such a sequence is referred to as **white noise** (with mean 0 and variance σ^2). This is indicated by the notation

$$\{X_t\} \sim \text{WN}\left(0, \sigma^2\right).$$

Clearly, every IID$(0, \sigma^2)$ sequence is WN$(0, \sigma^2)$ but not conversely (~~see Problem 1.8 and the ARCH(1) process of Section~~ 10.3). □

Example 1.4.3 The random walk

What is the mean function of the random walk? Make sure you follow the derivation of the autocov. function.

If $\{S_t\}$ is the random walk defined in Example 1.3.3 with $\{X_t\}$ as in Example 1.4.1, then $ES_t = 0$, $E(S_t^2) = t\sigma^2 < \infty$ for all t, and, for $h \geq 0$,

$$\begin{aligned}\gamma_S(t+h, t) &= \mathrm{Cov}(S_{t+h}, S_t)\\ &= \mathrm{Cov}(S_t + X_{t+1} + \cdots + X_{t+h}, S_t)\\ &= \mathrm{Cov}(S_t, S_t)\\ &= t\sigma^2.\end{aligned}$$

Since $\gamma_S(t+h, t)$ depends on t, the series $\{S_t\}$ is *not* stationary. □

Example 1.4.4 First-order moving average or MA(1) process

Skip (we'll cover this next class)

Consider the series defined by the equation

$$X_t = Z_t + \theta Z_{t-1}, \quad t = 0, \pm 1, \ldots, \tag{1.4.1}$$

where $\{Z_t\} \sim \mathrm{WN}\left(0, \sigma^2\right)$ and θ is a real-valued constant. From (1.4.1) we see that $EX_t = 0$, $EX_t^2 = \sigma^2(1+\theta^2) < \infty$, and

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2\left(1+\theta^2\right), & \text{if } h = 0,\\ \sigma^2\theta, & \text{if } h = \pm 1,\\ 0, & \text{if } |h| > 1.\end{cases}$$

Thus the requirements of Definition 1.4.2 are satisfied, and $\{X_t\}$ is stationary. The autocorrelation function of $\{X_t\}$ is

$$\rho_X(h) = \begin{cases} 1, & \text{if } h = 0,\\ \theta/\left(1+\theta^2\right), & \text{if } h = \pm 1,\\ 0, & \text{if } |h| > 1.\end{cases}$$

□

Example 1.4.5 First-order autoregression or AR(1) process

Skip

Let us *assume* now that $\{X_t\}$ is a stationary series satisfying the equations

$$X_t = \phi X_{t-1} + Z_t, \quad t = 0, \pm 1, \ldots, \tag{1.4.2}$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$, $|\phi| < 1$, and Z_t is uncorrelated with X_s for each $s < t$. (We shall show in Section 2.2 that there is in fact exactly one such solution of (1.4.2).) By taking expectations on each side of (1.4.2) and using the fact that $EZ_t = 0$, we see at once that

$$EX_t = 0.$$

To find the autocorrelation function of $\{X_t\}$ we multiply each side of (1.4.2) by X_{t-h} $(h > 0)$ and then take expectations to get

$$\begin{aligned}\gamma_X(h) &= \mathrm{Cov}(X_t, X_{t-h})\\ &= \mathrm{Cov}(\phi X_{t-1}, X_{t-h}) + \mathrm{Cov}(Z_t, X_{t-h})\\ &= \phi\gamma_X(h-1) + 0 = \cdots = \phi^h\gamma_x(0).\end{aligned}$$

Observing that $\gamma(h) = \gamma(-h)$ and using Definition 1.4.3, we find that

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \phi^{|h|}, \quad h = 0, \pm 1, \ldots.$$

It follows from the linearity of the covariance function in each of its arguments and the fact that Z_t is uncorrelated with X_{t-1} that

$$\gamma_X(0) = \mathrm{Cov}(X_t, X_t) = \mathrm{Cov}(\phi X_{t-1} + Z_t, \phi X_{t-1} + Z_t) = \phi^2\gamma_X(0) + \sigma^2$$

and hence that $\gamma_X(0) = \sigma^2/\left(1 - \phi^2\right)$. □

1.4.1 The Sample Autocorrelation Function

Although we have just seen how to compute the autocorrelation function for a few simple time series models, in practical problems we do not start with a model, but with *observed data* $\{x_1, x_2, \ldots, x_n\}$. To assess the degree of dependence in the data and to select a model for the data that reflects this, one of the important tools we use is the **sample autocorrelation function** (sample ACF) of the data. If we believe that the data are realized values of a stationary time series $\{X_t\}$, then the sample ACF will provide us with an estimate of the ACF of $\{X_t\}$. This estimate may suggest which of the many possible stationary time series models is a suitable candidate for representing the dependence in the data. For example, a sample ACF that is close to zero for all nonzero lags suggests that an appropriate model for the data might be iid noise. The following definitions are natural *sample* analogues of those for the autocovariance and autocorrelation functions given earlier for stationary time series *models*.

Totally!

We will use this word a lot: "sample analogue"

This means that there will be some "population" object of interest and some analogous quantity in the "sample".

Definition 1.4.4

Let $x_1, \ldots, x_n$ be observations of a time series. The **sample mean** of $x_1, \ldots, x_n$ is

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t.$$

The **sample autocovariance function** is

$$\hat{\gamma}(h) := n^{-1}\sum_{t=1}^{n-|h|}(x_{t+|h|} - \bar{x})(x_t - \bar{x}), \quad -n < h < n.$$

The **sample autocorrelation function** is

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, \quad -n < h < n.$$

You can stop here

Remark 3. For $h \geq 0$, $\hat{\gamma}(h)$ is approximately equal to the sample covariance of the $n - h$ pairs of observations $(x_1, x_{1+h}), (x_2, x_{2+h}), \ldots, (x_{n-h}, x_n)$. The difference arises from use of the divisor n instead of $n - h$ and the subtraction of the *overall* mean, $\bar{x}$, from each factor of the summands. Use of the divisor n ensures that the sample covariance matrix $\hat{\Gamma}_n := [\hat{\gamma}(i - j)]_{i,j=1}^{n}$ is nonnegative definite (see Section 2.4.2). □

Remark 4. Like the sample covariance matrix defined in Remark 3, the sample correlation matrix $\hat{R}_n := [\hat{\rho}(i - j)]_{i,j=1}^{n}$ is nonnegative definite. Each of its diagonal elements is equal to 1, since $\hat{\rho}(0) = 1$. □

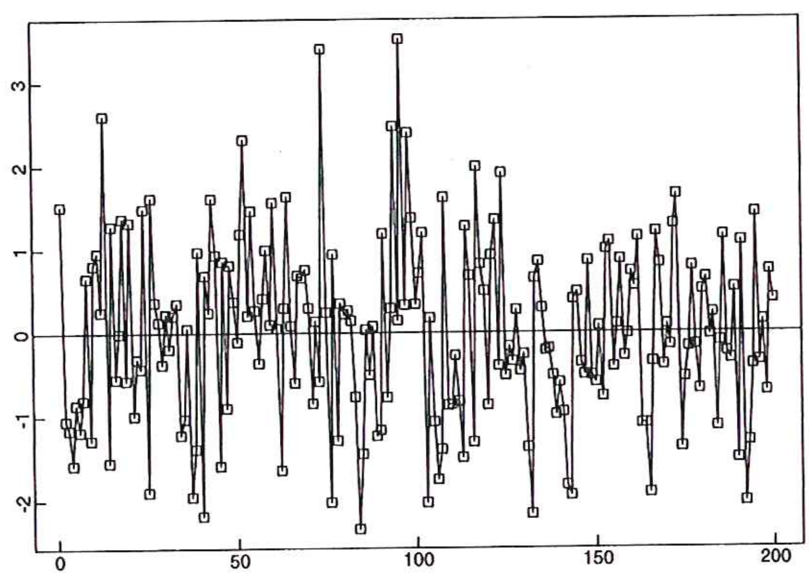

Figure 1-12
200 simulated values of iid N(0,1) noise.

Example 1.4.6 Figure 1.12 shows 200 simulated values of normally distributed iid (0, 1), denoted by IID N(0, 1), noise. Figure 1.13 shows the corresponding sample autocorrelation function at lags 0, 1, . . . , 40. Since $\rho(h) = 0$ for $h > 0$, one would also expect the corresponding sample autocorrelations to be near 0. It can be shown, in fact, that for iid noise with finite variance, the sample autocorrelations $\hat{\rho}(h)$, $h > 0$, are approximately IID N(0, $1/n$) for n large (see TSTM p. 222). Hence, approximately 95% of the sample autocorrelations should fall between the bounds $\pm 1.96/\sqrt{n}$ (since 1.96 is the .975 quantile of the standard normal distribution). Therefore, in Figure 1.13 we would expect roughly 40(.05) = 2 values to fall outside the bounds. To simulate 200 values of IID N(0, 1) noise using ITSM, select `File>Project>New>Univariate` then `Model>Simulate`. In the resulting dialog box, enter 200 for the required `Number of Observations`. (The remaining entries in the dialog box can be left as they are, since the model assumed by ITSM, until you enter another, is IID N(0, 1) noise. If you wish to reproduce exactly the same sequence at a later date, record the `Random Number Seed` for later use. By specifying different values for the random number seed you can generate independent realizations of your time series.) Click on `OK` and you will see the graph of your simulated series. To see its sample autocorrelation function together with the autocorrelation function of the model that generated it, click on the third yellow button at the top of the screen and you will see the two graphs superimposed (with the latter in red.) The horizontal lines on the graph are the bounds $\pm 1.96/\sqrt{n}$. □

Remark 5. The sample autocovariance and autocorrelation functions can be computed for *any* data set $\{x_1, \ldots, x_n\}$ and are not restricted to observations from a

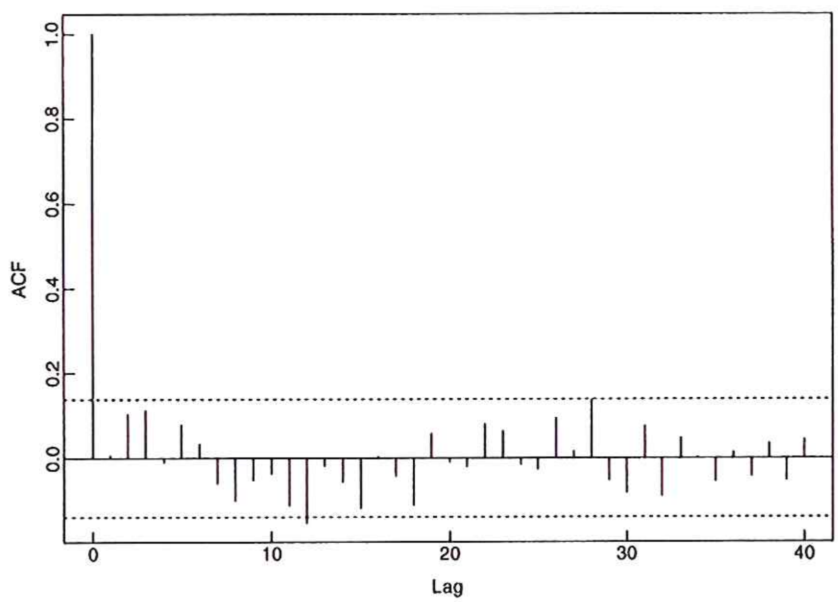

Figure 1-13 The sample autocorrelation function for the data of Figure 1.12 showing the bounds $\pm 1.96/\sqrt{n}$.

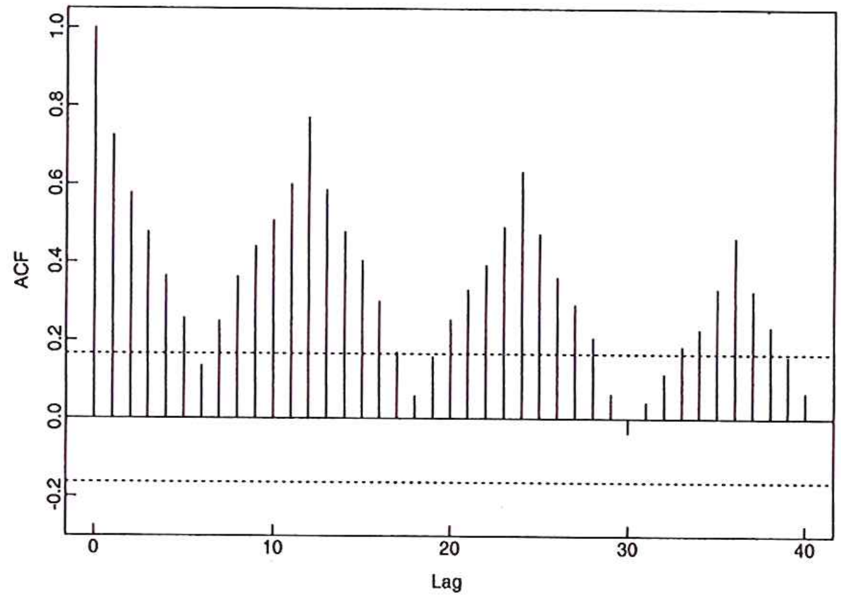

Figure 1-14 The sample autocorrelation function for the Australian red wine sales showing the bounds $\pm 1.96/\sqrt{n}$.

stationary time series. For data containing a trend, $|\hat{\rho}(h)|$ will exhibit slow decay as h increases, and for data with a substantial deterministic periodic component, $|\hat{\rho}(h)|$ will exhibit similar behavior with the same periodicity. (See the sample ACF of the Australian red wine sales in Figure 1.14 and Problem 1.9.) Thus $\hat{\rho}(\cdot)$ can be useful as an indicator of nonstationarity (see also Section 6.1). □

1.4.2 A Model for the Lake Huron Data

As noted earlier, an iid noise model for the residuals $\{y_1, \ldots, y_{98}\}$ obtained by fitting a straight line to the Lake Huron data in Example 1.3.5 appears to be inappropriate. This conclusion is confirmed by the sample ACF of the residuals (Figure 1.15), which has three of the first forty values well outside the bounds $\pm 1.96/\sqrt{98}$.

The roughly geometric decay of the first few sample autocorrelations (with $\hat{\rho}(h+1)/\hat{\rho}(h) \approx 0.7$) suggests that an AR(1) series (with $\phi \approx 0.7$) might provide a reasonable model for these residuals. (The form of the ACF for an AR(1) process was computed in Example 1.4.5.)

To explore the appropriateness of such a model, consider the points (y_1, y_2), $(y_2, y_3), \ldots, (y_{97}, y_{98})$ plotted in Figure 1.16. The graph does indeed suggest a linear relationship between y_t and y_{t-1}. Using simple least squares estimation to fit a straight line of the form $y_t = ay_{t-1}$, we obtain the model

$$Y_t = .791Y_{t-1} + Z_t, \tag{1.4.3}$$

where $\{Z_t\}$ is iid noise with variance $\sum_{t=2}^{98}(y_t - .791y_{t-1})^2/97 = .5024$. The sample ACF of the estimated noise sequence $z_t = y_t - .791y_{t-1}, t = 2, \ldots, 98$, is slightly