CHAPTER 13

INTRODUCTION TO SPECTRAL ANALYSIS

Historically, spectral analysis began with the search for "hidden periodicities" in time series data. Chapter 3 discussed fitting cosine trends at various known frequencies to series with strong cyclical trends. In addition, the random cosine wave example in Chapter 2 on page 18, showed that it is possible for a stationary process to look very much like a deterministic cosine wave. We hinted in Chapter 3 that by using enough different frequencies with enough different amplitudes (and phases) we might be able to model nearly any stationary series.[†] This chapter pursues those ideas further with an introduction to spectral analysis. Previous to this chapter, we concentrated on analyzing the correlation properties of time series. Such analysis is often called *time domain* analysis. When we analyze frequency properties of time series, we say that we are working in the *frequency domain*. Frequency domain analysis or *spectral analysis* has been found to be especially useful in acoustics, communications engineering, geophysical science, and biomedical science, for example.

13.1 Introduction

Recall from Chapter 3 the cosine curve with equation[‡]

$$R\cos(2\pi ft + \Phi) \tag{13.1.1}$$

Remember that R (> 0) is the *amplitude*, f the *frequency*, and Φ the *phase* of the curve. Since the curve repeats itself exactly every $1/f$ time units, $1/f$ is called the *period* of the cosine wave.

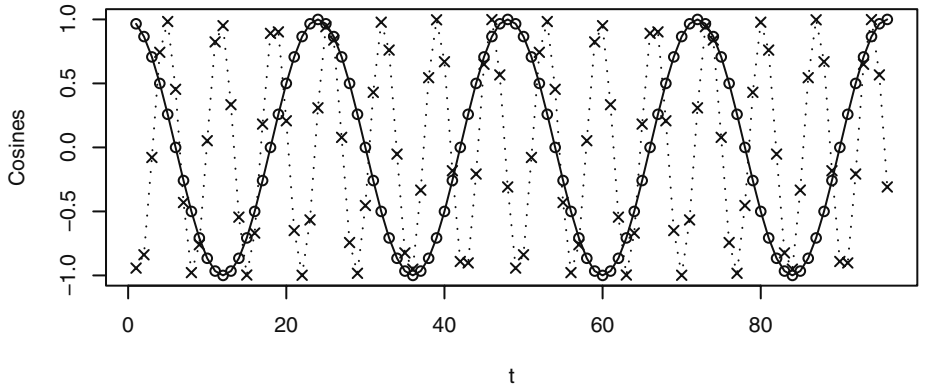
Exhibit 13.1 displays two discrete-time cosine curves with time running from 1 to 96. We would only see the discrete points, but the connecting line segments are added to help our eyes follow the pattern. The frequencies are 4/96 and 14/96, respectively. The lower-frequency curve has a phase of zero, but the higher-frequency curve is shifted by a phase of 0.6π.

Exhibit 13.2 shows the graph of a linear combination of the two cosine curves with a multiplier of 2 on the low-frequency curve and a multiplier of 3 on the higher-frequency curve and a phase of 0.6π; that is,

[†] See Exercise 2.25 on page 23, in particular.

[‡] In this chapter, we use notation slightly different from that in Chapter 3.

Exhibit 13.1 Cosine Curves with $n = 96$ and Two Frequencies and Phases

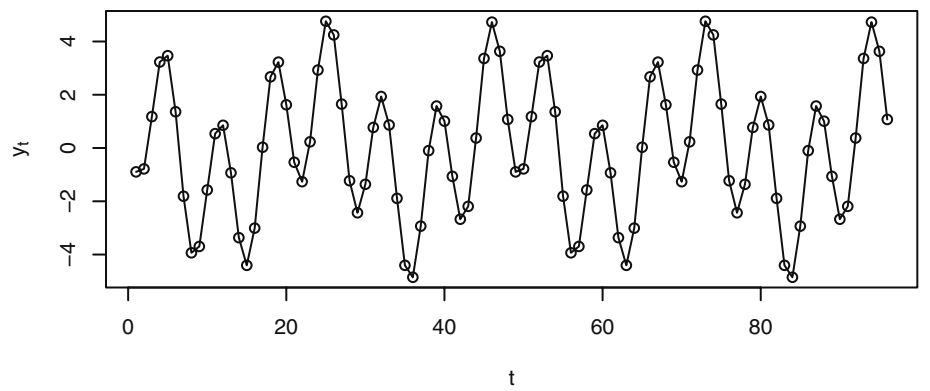



```
> win.graph(width=4.875,height=2.5,pointsize=8)
> t=1:96; cos1=cos(2*pi*t*4/96); cos2=cos(2*pi*(t*14/96+.3))
> plot(t,cos1, type='o', ylab='Cosines')
> lines(t,cos2,lty='dotted',type='o',pch=4)
```

$$Y_t = 2\cos\left(2\pi t\frac{4}{96}\right) + 3\cos\left[2\pi\left(t\frac{14}{96} + 0.3\right)\right] \tag{13.1.2}$$

Now the periodicity is somewhat hidden. Spectral analysis provides tools for discovering the "hidden" periodicities quite easily. Of course, there is nothing random in this time series.

Exhibit 13.2 Linear Combination of Two Cosine Curves




```
> y=2*cos1+3*cos2; plot(t,y,type='o',ylab=expression(y[t]))
```

As we saw earlier, Equation (13.1.1) is not convenient for estimation because the parameters R and Φ do not enter the expression linearly. Instead, we use a trigonometric identity to reparameterize Equation (13.1.1) as

$$R\cos(2\pi ft + \Phi) = A\cos(2\pi ft) + B\sin(2\pi ft) \tag{13.1.3}$$

where

$$R = \sqrt{A^2 + B^2}, \qquad \Phi = \operatorname{atan}(-B/A) \tag{13.1.4}$$

and, conversely,

$$A = R\cos(\Phi), \qquad B = -R\sin(\Phi) \tag{13.1.5}$$

Then, for a fixed frequency f, we can use $\cos(2\pi ft)$ and $\sin(2\pi ft)$ as predictor variables and fit the A's and B's from the data using ordinary least squares regression.

A general linear combination of m cosine curves with arbitrary amplitudes, frequencies, and phases could be written as[†]

$$Y_t = A_0 + \sum_{j=1}^{m} [A_j\cos(2\pi f_j t) + B_j\sin(2\pi f_j t)] \tag{13.1.6}$$

Ordinary least squares regression can be used to fit the A's and B's, but when the frequencies of interest are of a special form, the regressions are especially easy. Suppose that n is odd and write $n = 2k + 1$. Then the frequencies of the form $1/n$, $2/n$,…, k/n $(= 1/2 - 1/(2n))$ are called the *Fourier frequencies.* The cosine and sine predictor variables at these frequencies (and at $f = 0$) are known to be orthogonal,[‡] and the least squares estimates are simply

$$\hat{A}_0 = \bar{Y} \tag{13.1.7}$$

$$\hat{A}_j = \frac{2}{n}\sum_{t=1}^{n} Y_t\cos(2\pi tj/n) \quad \text{and} \quad \hat{B}_j = \frac{2}{n}\sum_{t=1}^{n} Y_t\sin(2\pi tj/n) \tag{13.1.8}$$

If the sample size is even, say $n = 2k$, Equations (13.1.7) and (13.1.8) still apply for $j = 1, 2,\ldots, k - 1$, but

$$\hat{A}_k = \frac{1}{n}\sum_{t=1}^{n} (-1)^t Y_t \text{ and } \hat{B}_k = 0 \tag{13.1.9}$$

Note that here $f_k = k/n = ½$.

If we were to apply these formulas to the series shown in Exhibit 13.2, we would obtain perfect results. That is, at frequency $f_4 = 4/96$, we obtain $\hat{A}_4 = 2$ and $\hat{B}_4 = 0$, and at frequency $f_{14} = 14/96$, we obtain $\hat{A}_{14} = -0.927051$ and $\hat{B}_{14} = -2.85317$. We would obtain estimates of zero for the regression coefficients at all other frequencies. These

[†] The A_0 term can be thought of as the coefficient of the cosine curve at zero frequency, which is identically one, and the B_0 can be thought of as the coefficient on the sine curve at frequency zero, which is identically zero and hence does not appear.

[‡] See Appendix J on page 349 for more information on the orthogonality properties of the cosines and sines.

results obtain because there is no randomness in this series and the cosine-sine fits are exact.

Note also that *any* series of any length n, whether deterministic or stochastic and with or without any true periodicities, can be fit perfectly by the model in Equation (13.1.6) by choosing $m = n/2$ if n is even and $m = (n - 1)/2$ if n is odd. There are then n parameters to adjust (estimate) to fit the series of length n.

13.2 The Periodogram

For odd sample sizes with $n = 2k + 1$, the *periodogram* I at frequency $f = j/n$ for $j = 1, 2, \ldots, k$, is defined to be

$$I\left(\frac{j}{n}\right) = \frac{n}{2}(\hat{A}_j^2 + \hat{B}_j^2) \tag{13.2.1}$$

If the sample size is even and $n = 2k$, Equations (13.1.7) and (13.1.8) still give the $\hat{A}$'s and $\hat{B}$'s and Equation (13.2.1) gives the periodogram for $j = 1, 2, \ldots, k - 1$. However, at the extreme frequency $f = k/n = ½$, Equations (13.1.9) apply and

$$I(\tfrac{1}{2}) = n(\hat{A}_k)^2 \tag{13.2.2}$$

Since the periodogram is proportional to the sum of squares of the regression coefficients associated with frequency $f = j/n$, *the height of the periodogram shows the relative strength of cosine-sine pairs at various frequencies in the overall behavior of the series*. Another interpretation is in terms of an analysis of variance. The periodogram $I(j/n)$ is the sum of squares with two degrees of freedom associated with the coefficient pair (A_j, B_j) at frequency j/n, so we have

$$\sum_{j=1}^{n} (Y_j - \overline{Y})^2 = \sum_{j=1}^{k} I\left(\frac{j}{n}\right) \tag{13.2.3}$$

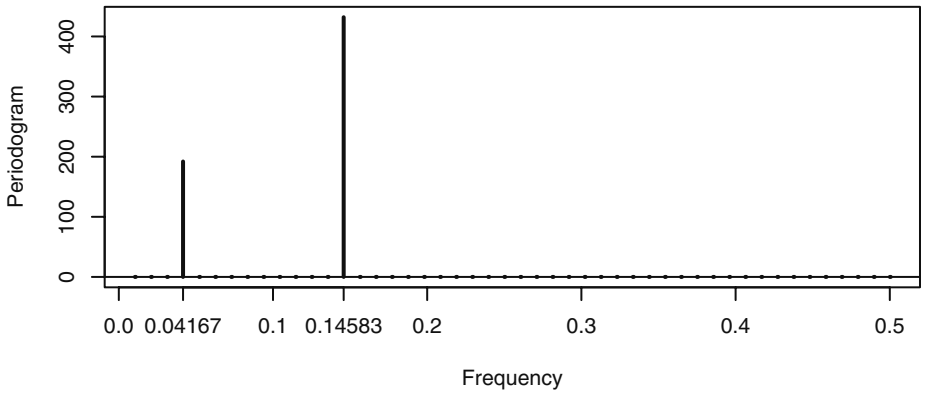
when $n = 2k + 1$ is odd. A similar result holds when n is even but there is a further term in the sum, $I(½)$, with one degree of freedom.

For long series, the computation of a large number of regression coefficients might be intensive. Fortunately, quick, efficient numerical methods based on the fast Fourier transform (FFT) have been developed that make the computations feasible for very long time series.[†]

Exhibit 13.3 displays a graph of the periodogram for the time series in Exhibit 13.2. The heights show the presence and relative strengths of the two cosine-sine components quite clearly. Note also that the frequencies $4/96 \approx 0.04167$ and $14/96 \approx 0.14583$ have been marked on the frequency axis.

[†] Often based on the Cooley-Tukey FFT algorithm; see Gentleman and Sande (1966).

Exhibit 13.3 Periodogram of the Series in Exhibit 13.2




```
> periodogram(y); abline(h=0); axis(1,at=c(0.04167,.14583))
```

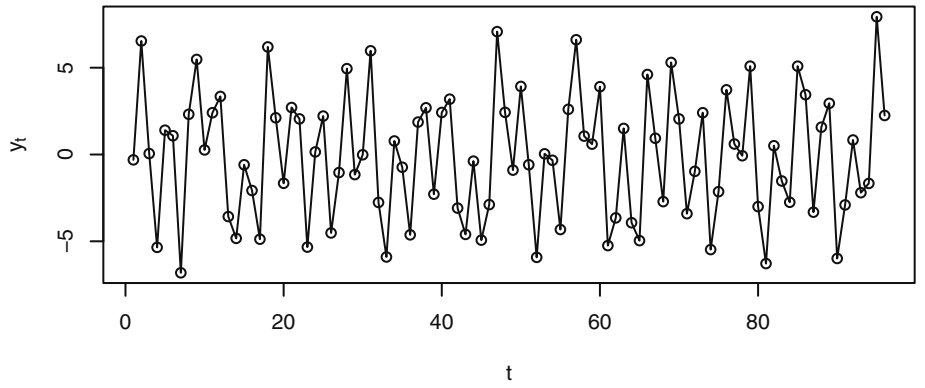
Does the periodogram work just as well when we do not know where or even if there are cosines in the series? What if the series contains additional "noise"? To illustrate, we generate a time series using randomness to select the frequencies, amplitudes, and phases and with additional additive white noise. The two frequencies are randomly chosen without replacement from among 1/96, 2/96,…, 47/96. The A's and B's are selected independently from normal distributions with means of zero and standard deviations of 2 for the first component and 3 for the second. Finally, a normal white noise series, $\{W_t\}$, with zero mean and standard deviation 1, is chosen independently of the A's and B's and added on. The model is[†]

$$Y_t = A_1\cos(2\pi f_1 t) + B_1\sin(2\pi f_1 t) + A_2\cos(2\pi f_2 t) + B_2\sin(2\pi f_2 t) + W_t \quad (13.2.4)$$

and Exhibit 13.4 displays a time series of length 96 simulated from this model. Once more, the periodicities are not obvious until we view the periodogram shown in Exhibit 13.5.

[†] This model is often described as a *signal plus noise* model. The *signal* could be deterministic (with unknown parameters) or stochastic.

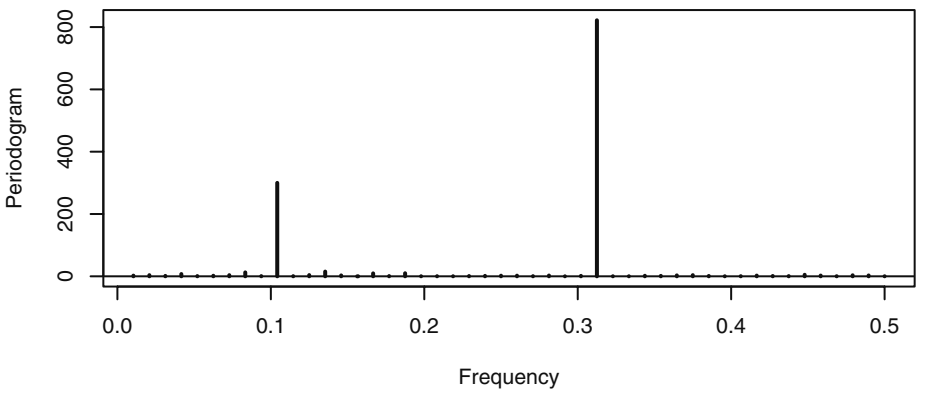
Exhibit 13.4 Time Series with "Hidden" Periodicities




```
> win.graph(width=4.875,height=2.5,pointsize=8)
> set.seed(134); t=1:96; integer=sample(48,2)
> freq1=integer[1]/96; freq2=integer[2]/96
> A1=rnorm(1,0,2); B1=rnorm(1,0,2)
> A2=rnorm(1,0,3); B2=rnorm(1,0,3); w=2*pi*t
> y=A1*cos(w*freq1)+B1*sin(w*freq1)+A2*cos(w*freq2)+
    B2*sin(w*freq2)+rnorm(96,0,1)
> plot(t,y,type='o',ylab=expression(y[t]))
```

The periodogram clearly shows that the series contains two cosine-sine pairs at frequencies of about 0.11 and 0.32 and that the higher-frequency component is much stronger. There are some other very small spikes in the periodogram, apparently caused by the additive white noise component. (When we checked the simulation in detail, we found that one frequency was chosen as $10/96 \approx 0.1042$ and the other was selected as $30/96 = 0.3125$.)

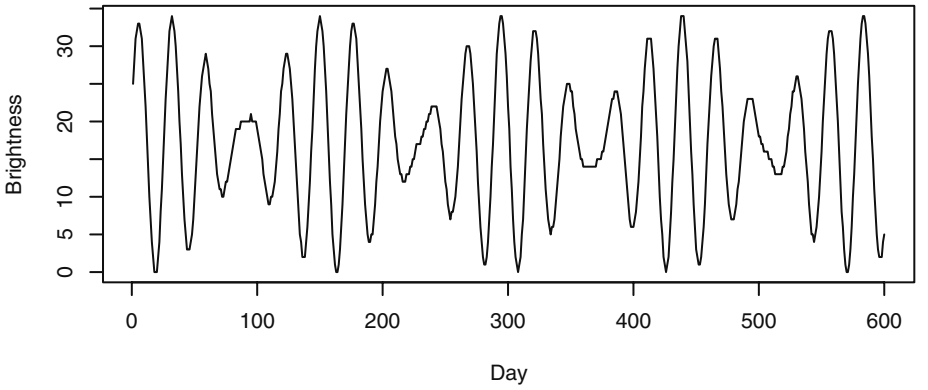
Exhibit 13.5 Periodogram of the Time Series Shown in Exhibit 13.4




```
> periodogram(y);abline(h=0)
```

Here is an example of the periodogram for a classic time series from Whittaker and Robinson (1924).[†] Exhibit 13.6 displays the time series plot of the brightness (magnitude) of a particular star at midnight on 600 consecutive nights.

Exhibit 13.6 Variable Star Brightness on 600 Consecutive Nights




```
> data(star)
> plot(star,xlab='Day',ylab='Brightness')
```

Exhibit 13.7 shows the periodogram for this time series. There are two very prominent peaks in the periodogram. When we inspect the actual numerical values, we find that the larger peak occurs at frequency $f = 21/600 = 0.035$. This frequency corresponds to a period of $600/21 \approx 28.57$, or nearly 29 days. The secondary peak occurs at $f = 25/600 \approx 0.04167$, which corresponds to a period of 24 days. The much more modest nonzero periodogram values near the major peak are likely caused by *leakage*.

The two sharp peaks suggest a model for this series with just two cosine-sine pairs with the appropriate frequencies or periods, namely

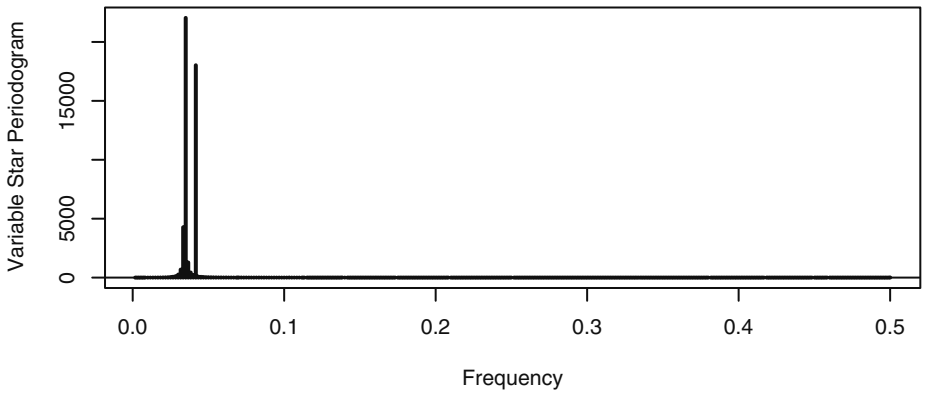
$$Y_t = \beta_0 + \beta_1 \cos(2\pi f_1 t) + \beta_2 \sin(2\pi f_1 t) + \beta_3 \cos(2\pi f_2 t) + \beta_4 \sin(2\pi f_2 t) + e_t \quad (13.2.5)$$

where $f_1 = 1/29$ and $f_2 = 1/24$. If we estimate this regression model as in Chapter 3, we obtain highly statistically significant regression coefficients for all five parameters and a multiple R-square value of 99.9%.

We will return to this time series in Section 14.5 on page 358, where we discuss more about leakage and tapering.

[†] An extensive analysis of this series appears throughout Bloomfield (2000).

Exhibit 13.7 Periodogram of the Variable Star Brightness Time Series




```
> periodogram(star,ylab='Variable Star Periodogram');abline(h=0)
```

Although the Fourier frequencies are special, we extend the definition of the periodogram to all frequencies in the interval 0 to ½ through the Equations (13.1.8) and (13.2.1). Thus we have for $0 \le f \le ½$

$$I(f) = \frac{n}{2}(\hat{A}_f^2 + \hat{B}_f^2) \tag{13.2.6}$$

where

$$\hat{A}_f = \frac{2}{n}\sum_{t=1}^{n} Y_t \cos(2\pi t f) \quad \text{and} \quad \hat{B}_f = \frac{2}{n}\sum_{t=1}^{n} Y_t \sin(2\pi t f) \tag{13.2.7}$$

When viewed in this way, the periodogram is often calculated at a grid of frequencies finer than the Fourier frequencies, and the plotted points are connected by line segments to display a somewhat smooth curve.

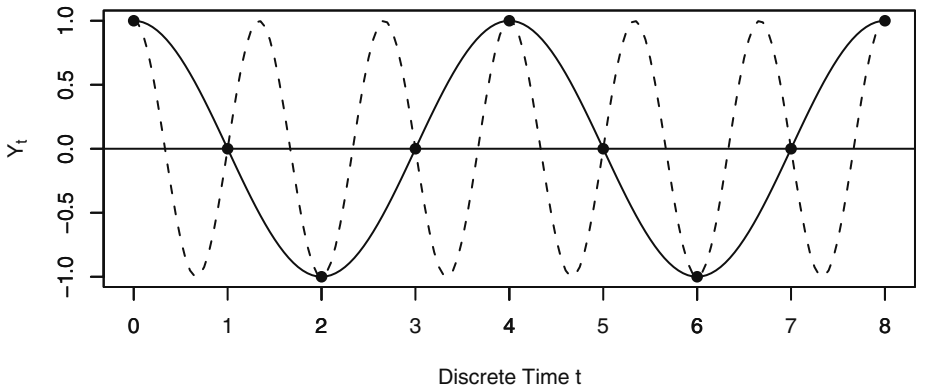
Why do we only consider positive frequencies? Because by the even and odd nature of cosines and sines, any cosine-sine curve with negative frequency, say $-f$, could just as well be expressed as a cosine-sine curve with frequency $+f$. No generality is lost by using positive frequencies.[†]

Secondly, why do we restrict frequencies to the interval from 0 to ½? Consider the graph shown in Exhibit 13.8. Here we have plotted two cosine curves, one with frequency $f = ¼$ and the one shown with dashed lines at frequency $f = ¾$. If we only observe the series at the discrete-time points 0, 1, 2, 3,…, the two series are identical. With discrete-time observations, we could never distinguish between these two curves. We say that the two frequencies ¼ and ¾ are **aliased** with one another. In general, each frequency f within the interval 0 to ½ will be aliased with each frequency of the form

[†] The definition of Equation (13.2.6) is often used for $-½ < f < +½$, but the resulting function is symmetric about zero and no new information is gained from the negative frequencies. Later in this chapter, we will use both positive and negative frequencies so that certain nice mathematical relationships hold.

$f + k(½)$ for any positive integer k, and it suffices to limit attention to frequencies within the interval from 0 to ½.

Exhibit 13.8 Illustration of Aliasing




```
> win.graph(width=4.875, height=2.5,pointsize=8)
> t=seq(0,8,by=.05)
> plot(t,cos(2*pi*t/4),axes=F,type='l',ylab=expression(Y[t]),
    xlab='Discrete Time t')
> axis(1,at=c(1,2,3,4,5,6,7));axis(1); axis(2); box()
> lines(t,cos(2*pi*t*3/4),lty='dashed',type='l'); abline(h=0)
> points(x=c(0:8),y=cos(2*pi*c(0:8)/4),pch=19)
```

13.3 The Spectral Representation and Spectral Distribution

Consider a time series represented as

$$Y_t = \sum_{j=1}^{m} [A_j \cos(2\pi f_j t) + B_j \sin(2\pi f_j t)] \tag{13.3.1}$$

where the frequencies $0 < f_1 < f_2 < \ldots < f_m < ½$ are fixed and A_j and B_j are independent normal random variables with zero means and $Var(A_j) = Var(B_j) = \sigma_j^2$. Then a straightforward calculation shows that $\{Y_t\}$ is stationary[†] with mean zero and

$$\gamma_k = \sum_{j=1}^{m} \sigma_j^2 \cos(2\pi k f_j) \tag{13.3.2}$$

In particular, the process variance, γ_0, is a sum of the variances due to each component at the various fixed frequencies:

[†] Compare this with Exercise 2.29 on page 24.

$$\gamma_0 = \sum_{j=1}^{m} \sigma_j^2 \tag{13.3.3}$$

If for $0 < f < ½$ we define two random step functions by

$$a(f) = \sum_{\{j|f_j \le f\}} A_j \quad \text{and} \quad b(f) = \sum_{\{j|f_j \le f\}} B_j \tag{13.3.4}$$

then we can write Equation (13.3.1) as

$$Y_t = \int_0^{½} \cos(2\pi ft)da(f) + \int_0^{½} \sin(2\pi ft)db(f) \tag{13.3.5}$$

It turns out that *any* zero-mean stationary process may be represented as in Equation (13.3.5).[†] It shows how stationary processes may be represented as linear combinations of infinitely many cosine-sine pairs over a continuous frequency band. In general, $a(f)$ and $b(f)$ are zero-mean stochastic processes indexed by frequency on $0 \le f \le ½$, each with uncorrelated[‡] increments, and the increments of $a(f)$ are uncorrelated with the increments of $b(f)$. Furthermore, we have

$$Var\left(\int_{f_1}^{f_2} da(f)\right) = Var\left(\int_{f_1}^{f_2} db(f)\right) = F(f_2) - F(f_1), \text{ say.} \tag{13.3.6}$$

Equation (13.3.5) is called the *spectral representation* of the process. The nondecreasing function $F(f)$ defined on $0 \le f \le ½$ is called the *spectral distribution function* of the process.

We say that the special process defined by Equation (13.3.1) has a *purely discrete* (or *line*) *spectrum* and, for $0 \le f \le ½$,

$$F(f) = \sum_{\{j|f_j \le f\}} \sigma_j^2 \tag{13.3.7}$$

Here the heights of the jumps in the spectral distribution give the variances associated with the various periodic components, and the positions of the jumps indicate the frequencies of the periodic components.

In general, a spectral distribution function has the properties

[†] The proof is beyond the scope of this book. See Cramér and Leadbetter (1967, pp. 128–138), for example. You do not need to understand stochastic Riemann-Stieltjes integrals to appreciate the rest of the discussion of spectral analysis.

[‡] Uncorrelated increments are usually called *orthogonal increments*.

$$\left.\begin{array}{ll} 1. & F \text{ is nondecreasing} \\ 2. & F \text{ is right continuous} \\ 3. & F(f) \geq 0 \text{ for all } f \\ 4. & \lim_{f \to 1/2} F(f) = Var(Y_t) = \gamma_0 \end{array}\right\} \tag{13.3.8}$$

If we consider the scaled spectral distribution function $F(f)/\gamma_0$, we have a function with the same mathematical properties as a cumulative distribution function (CDF) for a random variable on the interval 0 to ½ since now $F(½)/\gamma_0 = 1$.

We interpret the spectral distribution by saying that, for $0 \leq f_1 < f_2 \leq ½$, the integral

$$\int_{f_1}^{f_2} dF(f) \tag{13.3.9}$$

gives the portion of the (total) process variance $F(½) = \gamma_0$ that is attributable to frequencies in the range f_1 to f_2.

Sample Spectral Density

In spectral analysis, it is customary to first remove the sample mean from the series. For the remainder of this chapter, we assume that in the definition of the periodogram, Y_t represents deviations from its sample mean. Furthermore, for mathematical convenience, we now let various functions of frequency, such as the periodogram, be defined on the interval (–½,½]. In particular, we define the sample spectral density or sample spectrum as $\hat{S}(f) = ½I(f)$ for all frequencies in (–½,½) and $\hat{S}(½) = I(½)$. Using straightforward but somewhat tedious algebra, we can show that the sample spectral density can also be expressed as

$$\hat{S}(f) = \hat{\gamma}_0 + 2\sum_{k=1}^{n-1} \hat{\gamma}_k \cos(2\pi f k) \tag{13.3.10}$$

where $\hat{\gamma}_k$ is the sample or estimated covariance function at lag k ($k = 0, 1, 2, \ldots, n-1$) given by

$$\hat{\gamma}_k = \frac{1}{n} \sum_{t=k+1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y}) \tag{13.3.11}$$

In Fourier analysis terms, the sample spectral density is the (discrete-time) Fourier transform of the sample covariance function. From Fourier analysis theory, it follows that there is an inverse relationship, namely[†]

$$\hat{\gamma}_k = \int_{-½}^{½} \hat{S}(f) \cos(2\pi f k) df \tag{13.3.12}$$

[†] This may be proved using the orthogonality relationships shown in Appendix J on page 349.

In particular, notice that the total area under the sample spectral density is the sample variance of the time series.

$$\hat{\gamma}_0 = \int_{-½}^{½} \hat{S}(f)df = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \bar{Y})^2 \tag{13.3.13}$$

Since each can be obtained from the other, the sample spectral density and the sample covariance function contain the same information about the observed time series but it is expressed in different ways. For some purposes, one is more convenient or useful, and for other purposes the other is more convenient or useful.

13.4 The Spectral Density

For many processes, such as all stationary ARMA processes, the covariance functions decay rapidly with increasing lag.[†] When that is the case, it seems reasonable to consider the expression formed by replacing sample quantities in the sample spectral density of Equation (13.3.10) with the corresponding theoretical quantities. To be precise, if the covariance function γ_k is absolutely summable, we define the theoretical (or population) *spectral density* for $-½ < f \le ½$ as

$$S(f) = \gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k \cos(2\pi fk) \tag{13.4.1}$$

Once more, there is an inverse relationship, given by

$$\gamma_k = \int_{-½}^{½} S(f)\cos(2\pi fk)df \tag{13.4.2}$$

Mathematically, $S(f)$ is the (discrete-time) Fourier transform of the sequence $\ldots,\gamma_{-2},\gamma_{-1}$, $\gamma_0, \gamma_1, \gamma_2,\ldots$, and $\{\gamma_k\}$ is the inverse Fourier transform[‡] of the spectral density $S(f)$ defined on $-½ < f \le ½$.

A spectral density has all of the mathematical properties of a probability density function on the interval $(-½,½]$, with the exception that the total area is γ_0 rather than 1. Moreover, it can be shown that

---

[†] Of course, this is *not* the case for the processes defined in Equations (13.2.4) on page 323 and (13.3.1) on page 327. Those processes have discrete components in their spectra.

[‡] Notice that since $\gamma_k = \gamma_{-k}$ and the cosine function is also even, we could write

$$S(f) = \sum_{k=-\infty}^{\infty}\gamma_k e^{-2\pi ikf}$$

where $i = \sqrt{-1}$ is the imaginary unit for complex numbers. This looks more like a standard discrete-time Fourier transform. In a similar way, Equation (13.4.2) may be rewritten as $\gamma_k = \int_{-½}^{½} S(f)e^{2\pi ikf}df$.

$$F(f) = \int_0^f S(x)dx \qquad \text{for } 0 \le f \le \tfrac{1}{2} \tag{13.4.3}$$

Thus, twice the area under the spectral density between frequencies f_1 and f_2 with $0 \le f_1 < f_2 \le \frac{1}{2}$ is interpreted as the portion of the variance of the process that is attributable to cosine-sine pairs in that frequency interval that compose the process.

Time-Invariant Linear Filters

A *time-invariant linear filter* is defined by a sequence of absolutely summable constants $\ldots, c_{-1}, c_0, c_1, c_2, c_3, \ldots$. If $\{X_t\}$ is a time series, we use these constants to filter $\{X_t\}$ and produce a new time series $\{Y_t\}$ using the expression

$$Y_t = \sum_{j=-\infty}^{\infty} c_j X_{t-j} \tag{13.4.4}$$

If $c_k = 0$ for $k < 0$, we say that the filter is *causal*. In this case, the filtering at time t involves only present and past data values and can be carried out in "real time."

We have already seen many examples of time-invariant linear filters in previous chapters. Differencing (nonseasonal or seasonal) is an example. A combination of one seasonal difference with one nonseasonal difference is another example. Any moving average process can be considered as a linear filtering of a white noise sequence and in fact every general linear process defined by Equation (4.1.1) on page 55 is a linear filtering of white noise.

The expression on the right-hand side of Equation (13.4.4) is frequently called the (discrete-time) *convolution* of the two sequences $\{c_t\}$ and $\{X_t\}$. An extremely useful property of Fourier transforms is that the somewhat complicated operation of convolution in the time domain is transformed into the very simple operation of multiplication in the frequency domain.[†]

In particular, let $S_X(f)$ be the spectral density for the $\{X_t\}$ process and let $S_Y(f)$ be the spectral density for the $\{Y_t\}$ process. In addition, let

$$C(e^{-2\pi if}) = \sum_{j=-\infty}^{\infty} c_j e^{-2\pi ifj} \tag{13.4.5}$$

Then

$$\begin{aligned} Cov(Y_t, Y_{t-k}) &= Cov\left(\sum_{j=-\infty}^{\infty} c_j X_{t-j}, \sum_{s=-\infty}^{\infty} c_s X_{t-k-s}\right) \\ &= \sum_{j=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} c_j c_s Cov(X_{t-j}, X_{t-k-s}) \end{aligned}$$

[†] You may have already seen this with moment-generating functions. The density of the sum of two independent random variables, discrete or continuous, is the convolution of their respective densities, but the moment-generating function for the sum is the product of their respective moment-generating functions.

$$= \sum_{j=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} c_j c_s \int_{-½}^{½} e^{2\pi i(s+k-j)f} S_X(f)df$$

$$= \int_{-½}^{½} \left| \sum_{s=-\infty}^{\infty} c_s e^{-2\pi isf} \right|^2 e^{2\pi ifk} S_X(f)df$$

So

$$Cov(Y_t, Y_{t-k}) = \int_{-½}^{½} |C(e^{-2\pi if})|^2 S_X(f) e^{2\pi ifk} df \tag{13.4.6}$$

But

$$Cov(Y_t, Y_{t-k}) = \int_{-½}^{½} S_Y(f) e^{2\pi ifk} df \tag{13.4.7}$$

so we must have

$$S_Y(f) = |C(e^{-2\pi if})|^2 S_X(f) \tag{13.4.8}$$

This expression is invaluable for investigating the effect of time-invariant linear filters on spectra. In particular, it helps us find the form of the spectral densities for ARMA processes. The function $|C(e^{-2\pi if})|^2$ is often called the (*power*) *transfer function* of the filter.

13.5 Spectral Densities for ARMA Processes

White Noise

From Equation (13.4.1), it is easy to see that the theoretical spectral density for a white noise process is constant for all frequencies in $-½ < f \le ½$ and, in particular,

$$S(f) = \sigma_e^2 \tag{13.5.1}$$

All frequencies receive equal weight in the spectral representation of white noise. This is directly analogous to the spectrum of white light in physics—all colors (that is, all frequencies) enter equally in white light. Finally, we understand the origin of the name white noise!

MA(1) Spectral Density

An MA(1) process is a simple filtering of white noise with $c_0 = 1$ and $c_1 = -\theta$ and so

$$\begin{aligned} |C(e^{-2\pi if})|^2 &= (1 - \theta e^{2\pi if})(1 - \theta e^{-2\pi if}) \\ &= 1 + \theta^2 - \theta(e^{2\pi if} + e^{-2\pi if}) \\ &= 1 + \theta^2 - 2\theta\cos(2\pi f) \end{aligned} \tag{13.5.2}$$

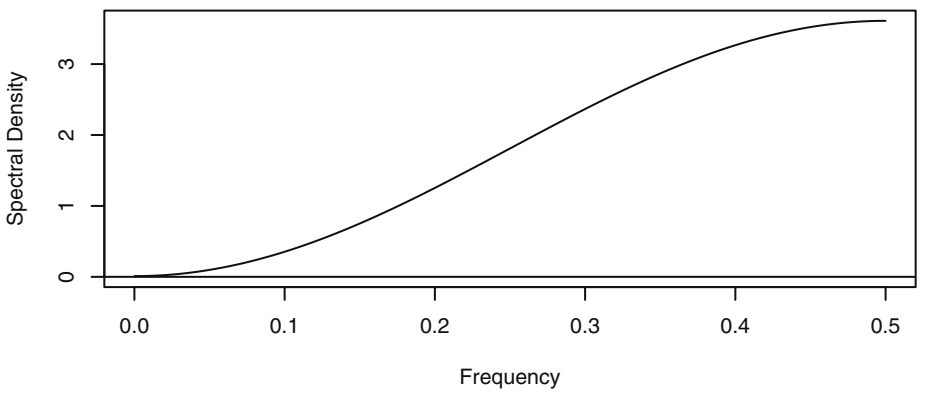
Thus

$$S(f) = [1 + \theta^2 - 2\theta\cos(2\pi f)]\sigma_e^2 \tag{13.5.3}$$

When $\theta > 0$, you can show that this spectral density is an increasing function of nonnegative frequency, while for $\theta < 0$ the function decreases.

Exhibit 13.9 displays the spectral density for an MA(1) process with $\theta = 0.9$.† Since spectral densities are symmetric about zero frequency, we will only plot them for positive frequencies. Recall that this MA(1) process has a relatively large negative correlation at lag 1 but all other correlations are zero. This is reflected in the spectrum. We see that the density is much stronger for higher frequencies than for low frequencies. The process has a tendency to oscillate back and forth across its mean level. This rapid oscillation is high-frequency behavior. We might say that the moving average suppresses the lower-frequency components of the white noise process. Researchers sometimes refer to this type of spectrum as a *blue spectrum* since it emphasizes the higher frequencies (that is, those with lower period or wavelength), which correspond to blue light in the spectrum of visible light.

Exhibit 13.9 Spectral Density of MA(1) Process with $\theta = 0.9$

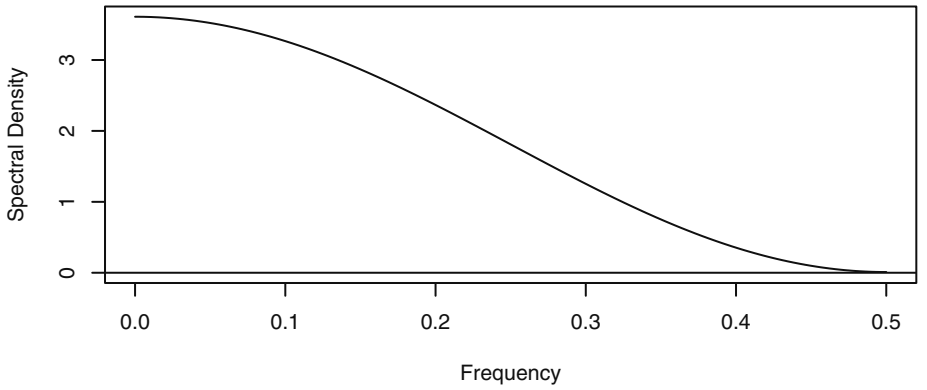



```
> win.graph(width=4.875,height=2.5,pointsize=8)
> theta=.9 # Reset theta for other MA(1) plots
> ARMAspec(model=list(ma=-theta))
```

Exhibit 13.10 displays the spectral density for an MA(1) process with $\theta = -0.9$. This process has positive correlation at lag 1 with all other correlations zero. Such a process will tend to change slowly from one time instance to the next. This is low-frequency behavior and is reflected in the shape of the spectrum. The density is much stronger for lower frequencies than for high frequencies. Researchers sometimes call this a *red spectrum*.

† In all of the plots of ARMA spectral densities that follow in this section, we take $\sigma_e^2 = 1$. This only affects the vertical scale of the graphs, not their shape.

Exhibit 13.10 Spectral Density of MA(1) Process with $\theta = -0.9$



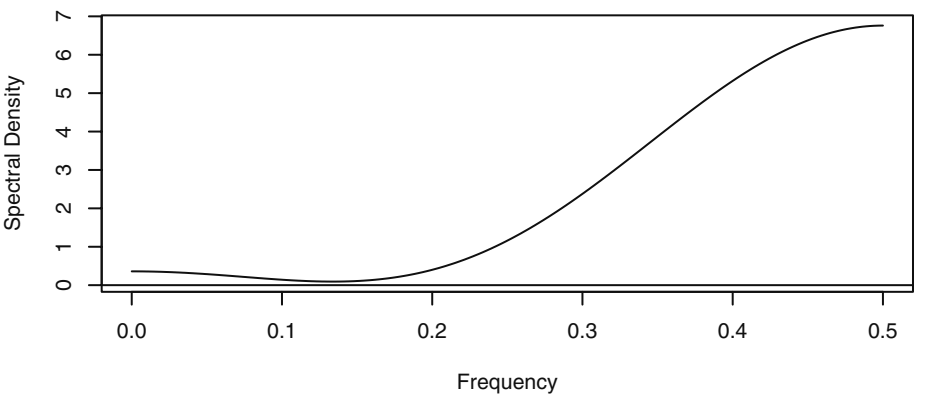
MA(2) Spectral Density

The spectral density for an MA(2) model may be obtained similarly. The algebra is a little longer, but the final expression is

$$S(f) = [1 + \theta_1^2 + \theta_2^2 - 2\theta_1(1 - \theta_2)\cos(2\pi f) - 2\theta_2\cos(4\pi f)]\sigma_e^2 \qquad (13.5.4)$$

Exhibit 13.11 shows a graph of such a density when $\theta_1 = 1$ and $\theta_2 = -0.6$. The frequencies between about 0.1 and 0.18 have especially small density and there is very little density below the frequency of 0.1. Higher frequencies enter into the picture gradually, with the strongest periodic components at the highest frequencies.

Exhibit 13.11 Spectral Density of MA(2) Process with $\theta_1 = 1$ and $\theta_2 = -0.6$




```
> theta1=1; theta2=-0.6
> ARMAspec(model=list(ma=-c(theta1,theta2)))
```

AR(1) Spectral Density

To find the spectral density for AR models, we use Equation (13.4.8) "backwards." That is, we view the white noise process as being a linear filtering of the AR process. Recalling the spectral density of the MA(1) series, this gives

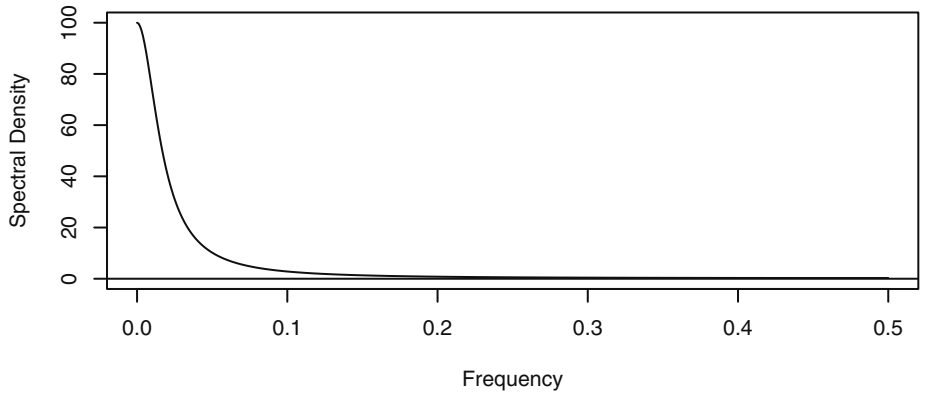
$$[1 + \phi^2 - 2\phi\cos(2\pi f)]S(f) = \sigma_e^2 \tag{13.5.5}$$

which we solve to obtain

$$S(f) = \frac{\sigma_e^2}{1 + \phi^2 - 2\phi\cos(2\pi f)} \tag{13.5.6}$$

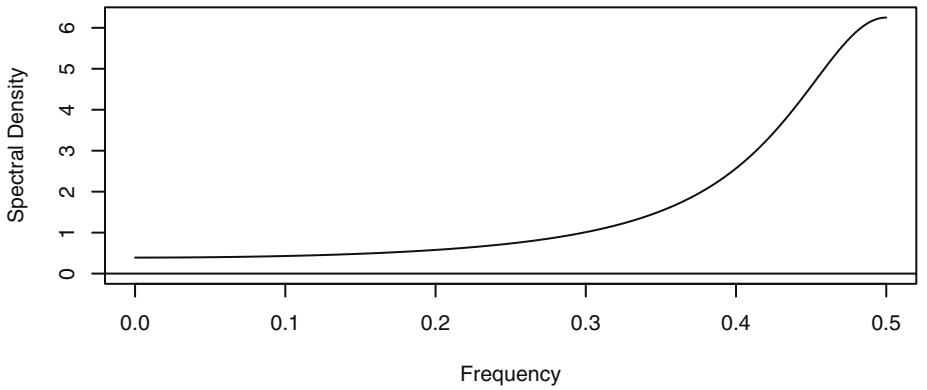
As the next two exhibits illustrate, this spectral density is a decreasing function of frequency when $\phi > 0$, while the spectral density increases for $\phi < 0$.

Exhibit 13.12 Spectral Density of an AR(1) Process with $\phi = 0.9$




```
> phi=0.9 # Reset value of phi for other AR(1) models
> ARMAspec(model=list(ar=phi))
```

Exhibit 13.13 Spectral Density of an AR(1) Process with $\phi = -0.6$



AR(2) Spectral Density

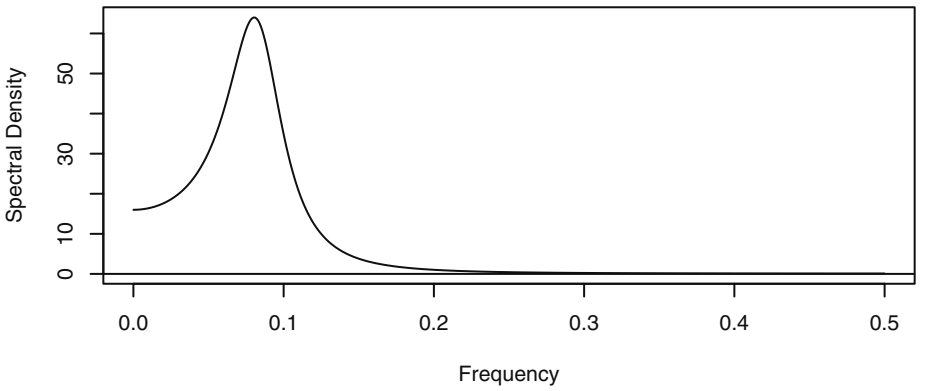
For the AR(2) spectral density, we again use Equation (13.4.8) backwards together with the MA(2) result to obtain

$$S(f) = \frac{\sigma_e^2}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2)\cos(2\pi f) - 2\phi_2\cos(4\pi f)} \tag{13.5.7}$$

Just as with the correlation properties, the spectral density for an AR(2) model can exhibit a variety of behaviors depending on the actual values of the two ϕ parameters.

Exhibits 13.14 and 13.15 display two AR(2) spectral densities that show very different behavior of peak in one case and trough in another.

Exhibit 13.14 Spectral Density of AR(2) Process: $\phi_1 = 1.5$ and $\phi_2 = -0.75$




```
> phi1=1.5; phi2=-.75
> # Reset values of phi1 & phi2 for other AR(2) models
> ARMAspec(model=list(ar=c(phi1,phi2)))
```

Jenkins and Watts (1968, p. 229), have noted that the different spectral shapes for an AR(2) spectrum are determined by the inequality

$$|\phi_1(1-\phi_2)| < |4\phi_2| \tag{13.5.8}$$

and the results are best summarized in the display in Exhibit 13.16. In this display, the dashed curve is the border between the regions of real roots and complex roots of the AR(2) characteristic equation. The solid curves are determined from the inequality given in Equation (13.5.8).

Exhibit 13.15 Spectral Density of AR(2) Process with $\phi_1 = 0.1$ and $\phi_2 = 0.4$

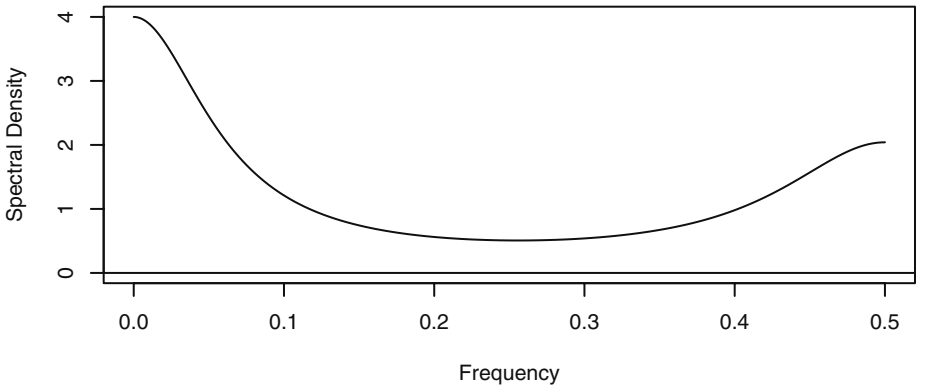
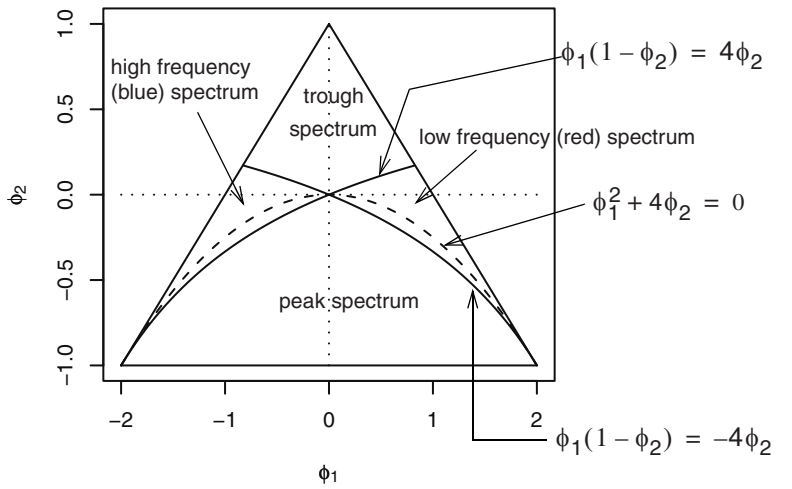


Exhibit 13.16 AR(2) Parameter Values for Various Spectral Density Shapes



Note that Jenkins and Watts also showed that the frequency f_0 at which the peak or trough occurs will satisfy

$$\cos(2\pi f_0) = -\frac{\phi_1(1-\phi_2)}{4\phi_2} \tag{13.5.9}$$

It is commonly thought that complex roots are associated with a peak spectrum. But notice that there is a small region of parameter values where the roots are complex but the spectrum is of either high or low frequency with no intermediate peak.

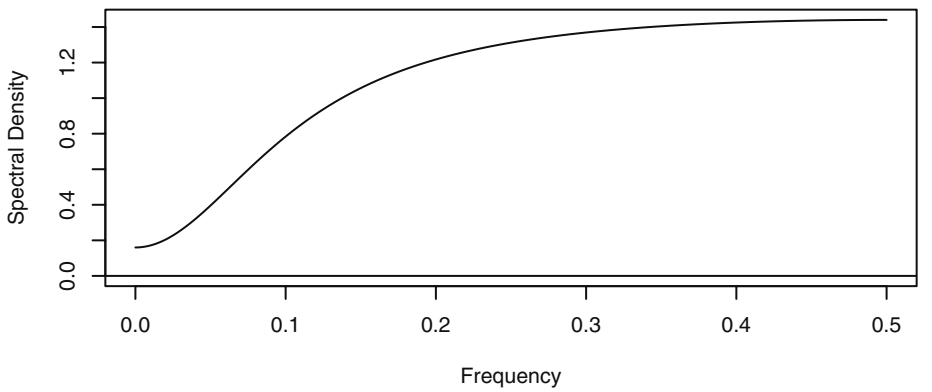
ARMA(1,1) Spectral Density

Combining what we know for MA(1) and AR(1) models, we can easily obtain the spectral density for the ARMA(1,1) mixed model

$$S(f) = \frac{1+\theta^2-2\theta\cos(2\pi f)}{1+\phi^2-2\phi\cos(2\pi f)}\sigma_e^2 \tag{13.5.10}$$

Exhibit 13.17 provides an example of the spectrum for an ARMA(1,1) model with $\phi = 0.5$ and $\theta = 0.8$.

Exhibit 13.17 Spectral Density of ARMA(1,1) with $\phi = 0.5$ and $\theta = 0.8$




```
> phi=0.5; theta=0.8
> ARMAspec(model=list(ar=phi,ma=-theta))
```

ARMA(p,q)

For the general ARMA(p,q) case, the spectral density may be expressed in terms of the AR and MA characteristic polynomials as

$$S(f) = \left|\frac{\theta(e^{-2\pi if})}{\phi(e^{-2\pi if})}\right|^2 \sigma_e^2 \tag{13.5.11}$$

This may be further expressed in terms of the reciprocal roots of these polynomials, but we will not pursue those expressions here. This type of spectral density is often referred to as a rational spectral density.

Seasonal ARMA Processes

Since seasonal ARMA processes are just special ARMA processes, all of our previous work will carry over here. Multiplicative seasonal models can be thought of as applying two linear filters consecutively. We will just give two examples.

Consider the process defined by the seasonal AR model

$$(1 - \phi B)(1 - \Phi B^{12})Y_t = e_t \tag{13.5.12}$$

Manipulating the two factors separately yields

$$S(f) = \frac{\sigma_e^2}{[1 + \phi^2 - 2\phi\cos(2\pi f)][1 + \Phi^2 - 2\Phi\cos(2\pi 12 f)]} \tag{13.5.13}$$

An example of this spectrum is shown in Exhibit 13.18, where $\phi = 0.5$, $\Phi = 0.9$, and $s = 12$. The seasonality is reflected in the many spikes of decreasing magnitude at frequencies of 0, 1/12, 2/12, 3/12, 4/12, 5/12, and 6/12.

As a second example, consider a seasonal MA process

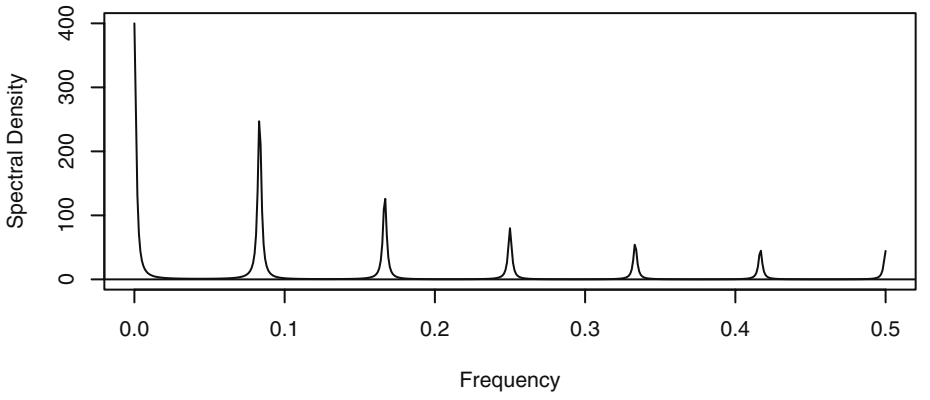
$$Y_t = (1 - \theta B)(1 - \Theta B^{12})e_t \tag{13.5.14}$$

The corresponding spectral density is given by

$$S(f) = [1 + \theta^2 - 2\theta\cos(2\pi f)][1 + \Theta^2 - 2\Theta\cos(2\pi 12 f)]\sigma_e^2 \tag{13.5.15}$$

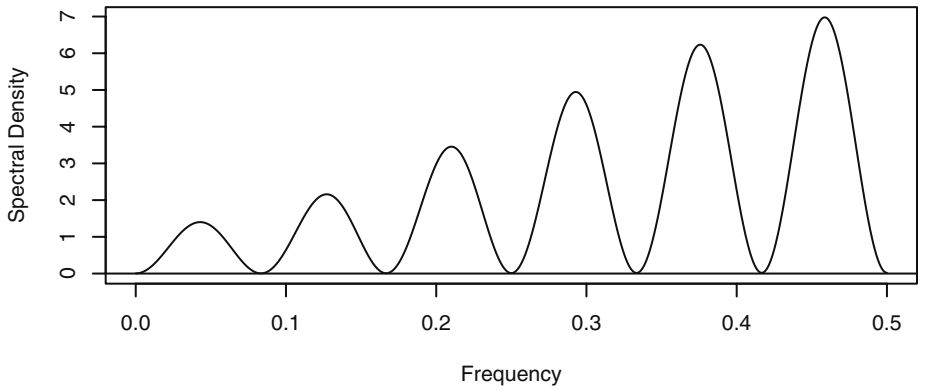
Exhibit 13.19 shows this spectral density for parameter values $\theta = 0.4$ and $\Theta = 0.9$.

Exhibit 13.18 Spectral Density of Seasonal AR with $\phi = 0.5$, $\Phi = 0.9$, $s = 12$




```
> phi=.5; PHI=.9
> ARMAspec(model=list(ar=phi,seasonal=list(sar=PHI,period=12)))
```

Exhibit 13.19 Spectral Density of Seasonal MA with $\theta = 0.4$, $\Theta = 0.9$, $s = 12$




```
> theta=.4; Theta=.9
> ARMAspec(model=list(ma=-theta,seasonal=list(sma=-Theta,
    period=12)))
```

13.6 Sampling Properties of the Sample Spectral Density

To introduce this section, we consider a time series with known properties. Suppose that we simulate an AR(1) model with $\phi = -0.6$ of length $n = 200$. Exhibit 13.13 on page 336, shows the theoretical spectral density for such a series. The sample spectral density for our simulated series is displayed in Exhibit 13.20, with the smooth theoretical spectral density shown as a dotted line. Even with a sample of size 200, the sample spectral density is extremely variable from one frequency point to the next. This is surely not an acceptable estimate of the theoretical spectrum for this process. We must investigate the sampling properties of the sample spectral density to understand the behavior that we see here.

To investigate the sampling properties of the sample spectral density, we begin with the simplest case, where the time series $\{Y_t\}$ is zero-mean normal white noise with variance γ_0. Recall that

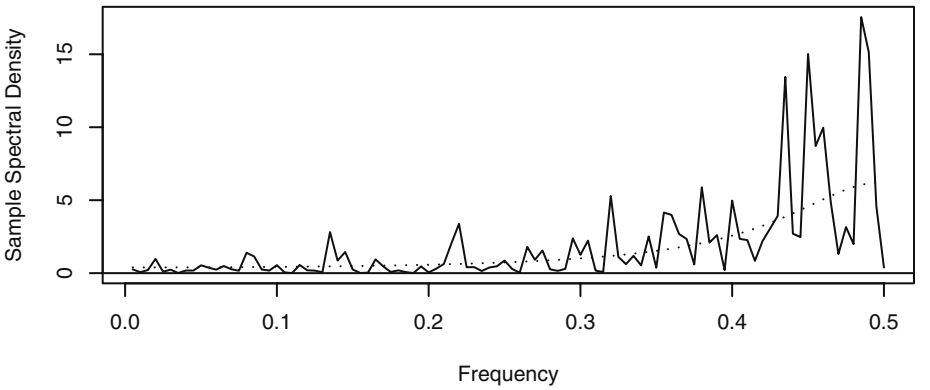
$$\hat{A}_f = \frac{2}{n}\sum_{t=1}^{n} Y_t \cos(2\pi tf) \quad \text{and} \quad \hat{B}_f = \frac{2}{n}\sum_{t=1}^{n} Y_t \sin(2\pi tf) \tag{13.6.1}$$

For now, consider only nonzero Fourier frequencies $f = j/n < ½$. Since $\hat{A}_f$ and $\hat{B}_f$ are linear functions of the time series $\{Y_t\}$, they each have a normal distribution. We can evaluate the means and variances using the orthogonality properties of the cosines and sines.[†] We find that $\hat{A}_f$ and $\hat{B}_f$ each have mean zero and variance $2\gamma_0/n$. We can also use the orthogonality properties to show that $\hat{A}_f$ and $\hat{B}_f$ are uncorrelated and thus indepen-

[†] See Appendix J on page 349.

dent since they are jointly bivariate normal. Similarly, it can be shown that for any two distinct Fourier frequencies f_1 and f_2, $\hat{A}_{f_1}$, $\hat{A}_{f_2}$, $\hat{B}_{f_1}$, and $\hat{B}_{f_2}$ are jointly independent.

Exhibit 13.20 Sample Spectral Density for a Simulated AR(1) Process




```
> win.graph(width=4.875,height=2.5,pointsize=8)
> set.seed(271435); n=200; phi=-0.6
> y=arima.sim(model=list(ar=phi),n=n)
> sp=spec(y,log='no',xlab='Frequency',
    ylab='Sample Spectral Density',sub='')
> lines(sp$freq,ARMAspec(model=list(ar=phi),freq=sp$freq,
    plot=F)$spec,lty='dotted'); abline(h=0)
```

Furthermore, we know that the square of a standard normal has a chi-square distribution with one degree of freedom and that the sum of independent chi-square variables is chi-square distributed with degrees of freedom added together. Since $S(f) = \gamma_0$, we have

$$\frac{n}{2\gamma_0}[(\hat{A}_f)^2 + (\hat{B}_f)^2] = \frac{2\hat{S}(f)}{S(f)} \tag{13.6.2}$$

has a chi-square distribution with two degrees of freedom.

Recall that a chi-square variable has a mean equal to its degrees of freedom and a variance equal to twice its degrees of freedom. With these facts, we quickly discover that

$$\hat{S}(f_1) \text{ and } \hat{S}(f_2) \text{ are independent for } f_1 \neq f_2 \tag{13.6.3}$$

$$E[\hat{S}(f)] = S(f) \tag{13.6.4}$$

and

$$Var[\hat{S}(f)] = S^2(f) \tag{13.6.5}$$

Equation (13.6.4) expresses the desirable fact that the sample spectral density is an unbiased estimator of the theoretical spectral density.

Unfortunately, Equation (13.6.5) shows that the variance in no way depends on the sample size n. Even in this simple case, *the sample spectral density is not a consistent estimator of the theoretical spectral density.* It does not get better (that is, have smaller variance) as the sample size increases. The reason the sample spectral density is inconsistent is basically this: Even if we only consider Fourier frequencies, $1/n$, $2/n$,…, we are trying to estimate more and more "parameters"; that is, $S(1/n)$, $S(2/n)$,… . As the sample size increases, there are not enough data points per parameter to produce consistent estimates.

The results expressed in Equations (13.6.3)–(13.6.5) in fact hold more generally. In the exercises, we ask you to argue that for any white noise—not necessarily normal—the mean result holds exactly and the $\hat{A}_f$ and $\hat{B}_f$ that make up $\hat{S}(f_1)$ and $\hat{S}(f_2)$ are at least uncorrelated for $f_1 \neq f_2$.

To state more general results, suppose $\{Y_t\}$ is any linear process

$$Y_t = e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \cdots \tag{13.6.6}$$

where the e's are independent and identically distributed with zero mean and common variance. Suppose that the ψ-coefficients are absolutely summable, and let $f_1 \neq f_2$ be any frequencies in 0 to ½. Then it may be shown[†] that as the sample size increases without limit

$$\frac{2\hat{S}(f_1)}{S(f_1)} \text{ and } \frac{2\hat{S}(f_2)}{S(f_2)} \tag{13.6.7}$$

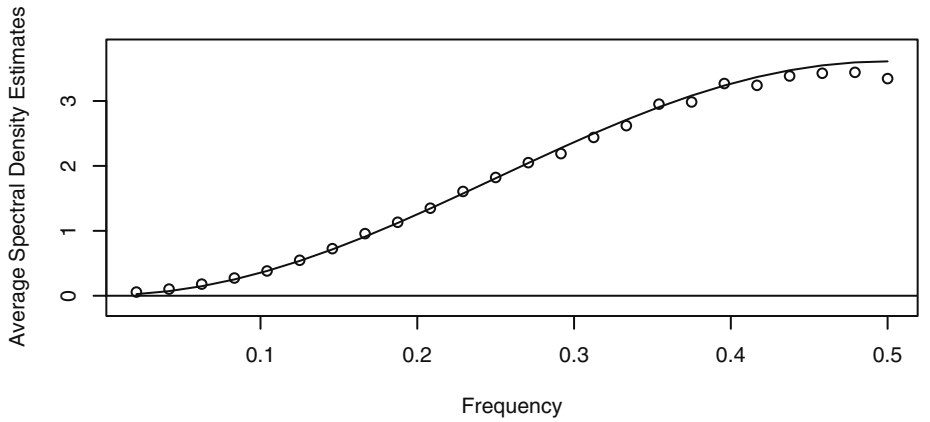
converge in distribution to independent chi-square random variables, each with two degrees of freedom.

To investigate the usefulness of approximations based on Equations (13.6.7), (13.6.4), and (13.6.5), we will display results from two simulations. We first simulated 1000 replications of an MA(1) time series with $\theta = 0.9$, each of length $n = 48$. The white noise series used to create each MA(1) series was selected independently from a t-distribution with five degrees of freedom scaled to unit variance. From the 1000 series, we calculated 1000 sample spectral densities.

Exhibit 13.21 shows the average of the 1000 sample spectral densities evaluated at the 24 Fourier frequencies associated with $n = 48$. The solid line is the theoretical spectral density. It appears that the sample spectral densities are unbiased to a useful approximation in this case.

[†] See, for example, Fuller (1996, pp. 360–361).

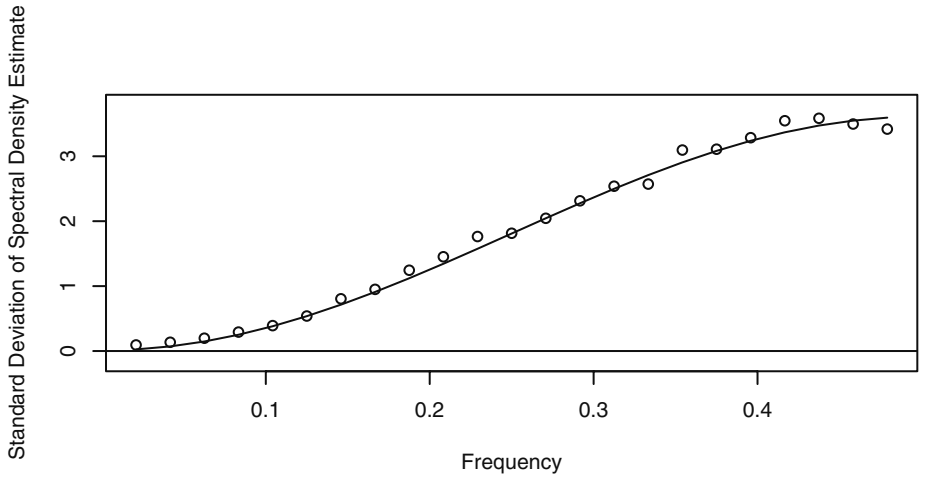
Exhibit 13.21 Average Sample Spectral Density: Simulated MA(1), $\theta = 0.9$, $n = 48$




```
For the extensive R code to produce Exhibits 13.21 through 13.26,
please see the Chapter 13 script file associated with this book.
```

Exhibit 13.22 plots the standard deviations of the sample spectral densities over the 1000 replications. According to Equation (13.6.5), we hope that they match the theoretical spectral density at the Fourier frequencies. Again the approximation seems to be quite acceptable.

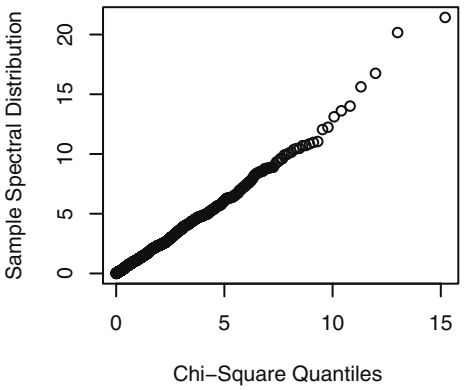
Exhibit 13.22 Standard Deviation of Sample Spectral Density: Simulated MA(1), $\theta = 0.9$, $n = 48$



To check on the shape of the sample spectral density distribution, we constructed a QQ plot comparing the observed quantiles with those of a chi-square distribution with

two degrees of freedom. Of course, we could do those for any of the Fourier frequencies. Exhibit 13.23 shows the results at the frequency 15/48. The agreement with the chi-square distribution appears to be acceptable.

Exhibit 13.23 QQ Plot of Spectral Distribution at $f = 15/48$



We repeated similar displays and calculations when the true model was an AR(2) with $\phi_1 = 1.5$, $\phi_2 = -0.75$, and $n = 96$. Here we used normal white noise. The results are displayed in Exhibits 13.24, 13.25, and 13.26. Once more the simulation results with $n = 96$ and 1000 replications seem to follow those suggested by limit theory quite remarkably.

Exhibit 13.24 Average Sample Spectral Density: Simulated AR(2), $\phi_1 = 1.5$, $\phi_2 = -0.75$, $n = 96$

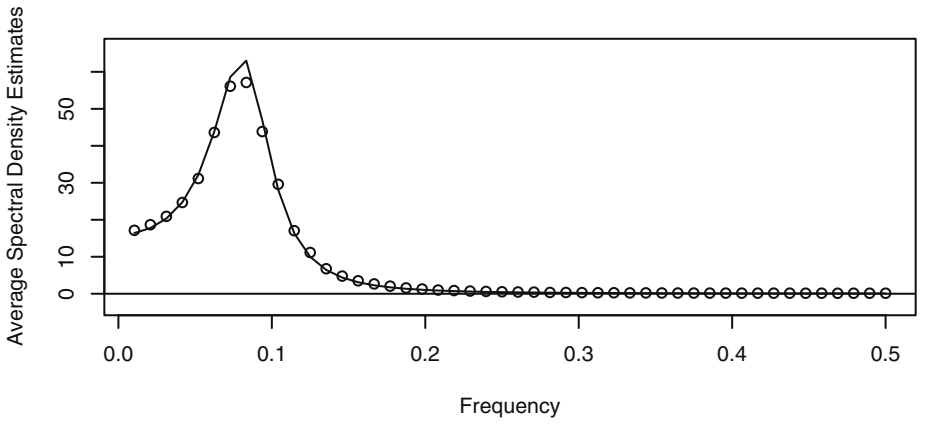


Exhibit 13.25 Standard Deviation of Sample Spectral Density: Simulated AR(2), $\phi_1 = 1.5$, $\phi_2 = -0.75$, $n = 96$

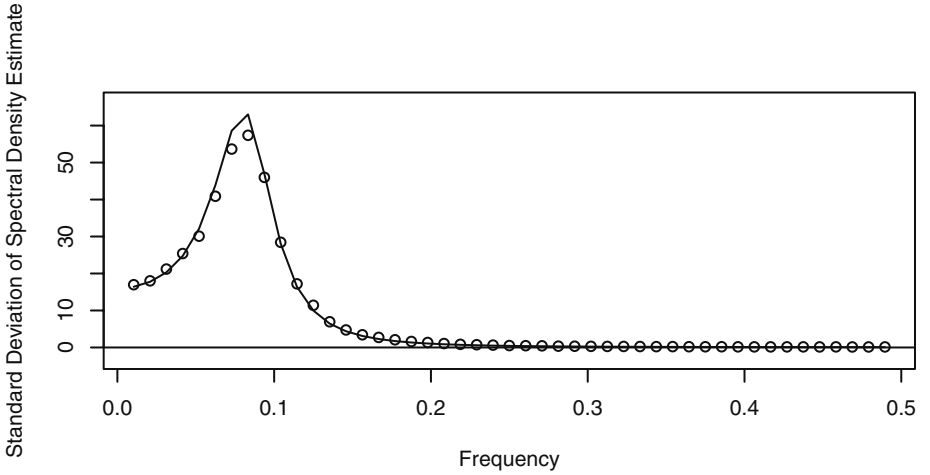
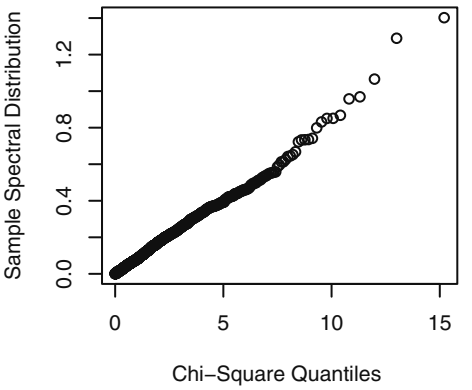


Exhibit 13.26 QQ Plot of Spectral Distribution at $f = 40/96$



Of course, none of these results tell us that the sample spectral density is an acceptable estimator of the underlying theoretical spectral density. The sample spectral density is quite generally approximately unbiased but also inconsistent, with way too much variability to be a useful estimator as it stands. The approximate independence at the Fourier frequencies also helps explain the extreme variability in the behavior of the sample spectral density.

13.7 Summary

The chapter introduces the ideas of modeling time series as linear combinations of sines and cosines—so-called spectral analysis. The periodogram was introduced as a tool for finding the contribution of the various frequencies in the spectral representation of the series. The ideas were then extended to modeling with a continuous range of frequencies. Spectral densities of the ARMA models were explored. Finally, the sampling properties of the sample spectral density were presented. Since the sample spectral density is not a consistent estimator of the theoretical spectral density, we must search further for an acceptable estimator. That is the subject of the next chapter.

EXERCISES

13.1 Find A and B so that $3\cos(2\pi ft + 0.4) = A\cos(2\pi ft) + B\sin(2\pi ft)$.

13.2 Find R and Φ so that $R\cos(2\pi ft + \Phi) = \cos(2\pi ft) + 3\sin(2\pi ft)$.

13.3 Consider the series displayed in Exhibit 13.2 on page 320.

(a) Verify that regressing the series on $\cos(2\pi ft)$ and $\sin(2\pi ft)$ for $f = 4/96$ provides perfect estimates of A and B.

(b) Use Equations (13.1.5) on page 321 to obtain the relationship between R, Φ, A and B for the cosine component at frequency $f = 14/96$. (For this component, the amplitude is 1 and the phase is 0.6π.)

(c) Verify that regressing the series on $\cos(2\pi ft)$ and $\sin(2\pi ft)$ for $f = 14/96$ provides perfect estimates of A and B.

(d) Verify that regressing the series on $\cos(2\pi ft)$ and $\sin(2\pi ft)$ for both $f = 4/96$ and $f = 14/96$ together provides perfect estimates of A_4, B_4, A_{14}, and B_{14}.

(e) Verify that regressing the series on $\cos(2\pi ft)$ and $\sin(2\pi ft)$ for $f = 3/96$ and $f = 13/96$ together provides perfect estimates of A_3, B_3, A_{13}, and B_{13}.

(f) Repeat part (d) but add a third pair of cosine-sine predictor variables at any other Fourier frequency. Verify that all of the regression coefficients are still estimated perfectly.

13.4 Generate or choose *any* series of length $n = 10$. Show that the series may be fit exactly by a linear combination of enough cosine-sine curves at the Fourier frequencies.

13.5 Simulate a signal + noise time series from the model in Equation (13.2.4) on page 323. Use the same parameter values used in Exhibit 13.4 on page 324.

(a) Plot the time series and look for the periodicities. Can you see them?

(b) Plot the periodogram for the simulated series. Are the periodicities clear now?

13.6 Show that the covariance function for the series defined by Equation (13.3.1) on page 327 is given by the expression in Equation (13.3.2).

13.7 Display the algebra that establishes Equation (13.3.10) on page 329.

13.8 Show that if $\{X_t\}$ and $\{Y_t\}$ are independent stationary series, then the spectral density of $\{X_t + Y_t\}$ is the sum of the spectral densities of $\{X_t\}$ and $\{Y_t\}$.

13.9 Show that when $\theta > 0$ the spectral density for an MA(1) process is an increasing function of frequency, while for $\theta < 0$ this function decreases.

13.10 Graph the theoretical spectral density for an MA(1) process with $\theta = 0.6$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.11 Graph the theoretical spectral density for an MA(1) process with $\theta = -0.8$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.12 Show that when $\phi > 0$ the spectral density for an AR(1) process is a decreasing function of frequency, while for $\phi < 0$ the spectral density increases.

13.13 Graph the theoretical spectral density for an AR(1) time series with $\phi = 0.7$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.14 Graph the theoretical spectral density for an AR(1) time series with $\phi = -0.4$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.15 Graph the theoretical spectral density for an MA(2) time series with $\theta_1 = -0.5$ and $\theta_2 = 0.9$. Interpret the implications of the shape of the spectrum on the possible time series plots of the series values.

13.16 Graph the theoretical spectral density for an MA(2) time series with $\theta_1 = 0.5$ and $\theta_2 = -0.9$. Interpret the implications of the shape of the spectrum on the possible time series plots of the series values.

13.17 Graph the theoretical spectral density for an AR(2) time series with $\phi_1 = -0.1$ and $\phi_2 = -0.9$. Interpret the implications of the shape of the spectrum on the possible time series plots of the series values.

13.18 Graph the theoretical spectral density for an AR(2) process with $\phi_1 = 1.8$ and $\phi_2 = -0.9$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.19 Graph the theoretical spectral density for an AR(2) process with $\phi_1 = -1$ and $\phi_2 = -0.8$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.20 Graph the theoretical spectral density for an AR(2) process with $\phi_1 = 0.5$ and $\phi_2 = 0.4$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.21 Graph the theoretical spectral density for an AR(2) process with $\phi_1 = 0$ and $\phi_2 = 0.8$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.22 Graph the theoretical spectral density for an AR(2) process with $\phi_1 = 0.8$ and $\phi_2 = -0.2$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.23 Graph the theoretical spectral density for an ARMA(1,1) time series with $\phi = 0.5$ and $\theta = 0.8$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.24 Graph the theoretical spectral density for an ARMA(1,1) process with $\phi = 0.95$ and $\theta = 0.8$. Interpret the implications of the shape of the spectrum on the possible plots of the time series values.

13.25 Let $\{X_t\}$ be a stationary time series and $\{Y_t\}$ be defined by $Y_t = (X_t + X_{t-1})/2$.

(a) Find the power transfer function for this linear filter.

(b) Is this a causal filter?

(c) Graph the power transfer function and describe the effect of using this filter. That is, what frequencies will be retained (emphasized) and what frequencies will be deemphasized (attenuated) by this filtering?

13.26 Let $\{X_t\}$ be a stationary time series and let $\{Y_t\}$ be defined by $Y_t = X_t - X_{t-1}$.

(a) Find the power transfer function for this linear filter.

(b) Is this a causal filter?

(c) Graph the power transfer function and describe the effect of using this filter. That is, what frequencies will be retained (emphasized) and what frequencies will be deemphasized (attenuated) by this filtering?

13.27 Let $\{X_t\}$ be a stationary time series and let $Y_t = (X_{t+1} + X_t + X_{t-1})/3$ define $\{Y_t\}$.

(a) Find the power transfer function for this linear filter.

(b) Is this a causal filter?

(c) Graph the power transfer function and describe the effect of using this filter. That is, what frequencies will be retained (emphasized) and what frequencies will be deemphasized (attenuated) by this filtering?

13.28 Let $\{X_t\}$ be a stationary time series and let $Y_t = (X_t + X_{t-1} + X_{t-2})/3$ define $\{Y_t\}$.

(a) Show that the power transfer function of this filter is the same as the power transfer function of the filter defined in Exercise 13.27.

(b) Is this a causal filter?

13.29 Let $\{X_t\}$ be a stationary time series and let $Y_t = X_t - X_{t-4}$ define $\{Y_t\}$.

(a) Find the power transfer function for this linear filter.

(b) Graph the power transfer function and describe the effect of using this filter. That is, what frequencies will be retained (emphasized) and what frequencies will be deemphasized (attenuated) by this filtering?

13.30 Let $\{X_t\}$ be a stationary time series and let $\{Y_t\}$ be defined by $Y_t = (X_{t+1} - 2X_t + X_{t-1})/3$.

(a) Find the power transfer function for this linear filter.

(b) Graph the power transfer function and describe the effect of using this filter. That is, what frequencies will be retained (emphasized) and what frequencies will be deemphasized (attenuated) by this filtering?

13.31 Suppose that $\{Y_t\}$ is a white noise process not necessarily normal. Use the orthogonality properties given in Appendix J to establish the following at the Fourier frequencies.

(a) The sample spectral density is an unbiased estimator of the theoretical spectral density.

(b) The variables $\hat{A}_{f_1}$ and $\hat{B}_{f_2}$ are uncorrelated for any Fourier frequencies f_1, f_2.

(c) If the Fourier frequencies $f_1 \neq f_2$, the variables $\hat{A}_{f_1}$ and $\hat{A}_{f_2}$ are uncorrelated.

13.32 Carry out a simulation analysis similar to those reported in Exhibits 13.21, 13.22, 13.23, and 13.24. Use an AR(2) model with $\phi_1 = 0.5$, $\phi_2 = -0.8$, and $n = 48$. Replicate the series 1000 times.

(a) Display the average sample spectral density by frequency and compare it with large sample theory.

(b) Display the standard deviation of the sample spectral density by frequency and compare it with large sample theory.

(c) Display the QQ plot of the appropriately scaled sample spectral density compared with large sample theory at several frequencies. Discuss your results.

13.33 Carry out a simulation analysis similar to those reported in Exhibits 13.21, 13.22, 13.23, and 13.24. Use an AR(2) model with $\phi_1 = -1$, $\phi_2 = -0.75$, and $n = 96$. Replicate the time series 1000 times.

(a) Display the average sample spectral density by frequency and compare it with the results predicted by large sample theory.

(b) Display the standard deviation of the sample spectral density by frequency and compare it with the results predicted by large sample theory.

(c) Display the QQ plot of the appropriately scaled sample spectral density and compare with the results predicted by large sample theory at several frequencies. Discuss your results.

13.34 Simulate a zero-mean, unit-variance, normal white noise time series of length $n = 1000$. Display the periodogram of the series, and comment on the results.

Appendix J: Orthogonality of Cosine and Sine Sequences

For $j, k = 0, 1, 2, \ldots, n/2$, we have

$$\sum_{t=1}^{n} \cos\left(2\pi \frac{j}{n} t\right) = 0 \qquad \text{if } j \neq 0 \tag{13.J.1}$$

$$\sum_{t=1}^{n} \sin\left(2\pi \frac{j}{n} t\right) = 0 \tag{13.J.2}$$

$$\sum_{t=1}^{n} \cos\left(2\pi \frac{j}{n} t\right) \sin\left(2\pi \frac{k}{n} t\right) = 0 \tag{13.J.3}$$

$$\sum_{t=1}^{n} \cos\left(2\pi\frac{j}{n}t\right)\cos\left(2\pi\frac{k}{n}t\right) = \begin{cases} \frac{n}{2} & \text{if } j = k \ (j \neq 0 \text{ or } n/2) \\ n & \text{if } j = k = 0 \\ 0 & \text{if } j \neq k \end{cases} \tag{13.J.4}$$

$$\sum_{t=1}^{n} \sin\left(2\pi\frac{j}{n}t\right)\sin\left(2\pi\frac{k}{n}t\right) = \begin{cases} \frac{n}{2} & \text{if } j = k \ (j \neq 0 \text{ or } n/2) \\ 0 & \text{if } j \neq k \end{cases} \tag{13.J.5}$$

These are most easily proved using DeMoivre's theorem

$$e^{-2\pi if} = \cos(2\pi f) - i\sin(2\pi f) \tag{13.J.6}$$

or, equivalently, Euler's formulas,

$$\cos(2\pi f) = \frac{e^{2\pi if} + e^{-2\pi if}}{2} \text{ and } \sin(2\pi f) = \frac{e^{2\pi if} - e^{-2\pi if}}{2i} \tag{13.J.7}$$

together with the result for the sum of a finite geometric series, namely

$$\sum_{j=1}^{n} r^j = \frac{r(1-r^n)}{1-r} \tag{13.J.8}$$

for real or complex $r \neq 1$.