The confidence intervals for the SOI series at the yearly cycle, $\omega = 1/12 = 40/480$, and the possible El Niño cycle of four years $\omega = 1/48 = 10/480$ can be computed in R as follows:

```
soi.per$spec[40]  # 0.97223; soi pgram at freq 1/12 = 40/480
soi.per$spec[10]  # 0.05372; soi pgram at freq 1/48 = 10/480
#  conf intervals  -  returned value:
U = qchisq(.025,2)    #  0.05063
L = qchisq(.975,2)    #  7.37775
2*soi.per$spec[10]/L  #  0.01456
2*soi.per$spec[10]/U  #  2.12220
2*soi.per$spec[40]/L  #  0.26355
2*soi.per$spec[40]/U  # 38.40108
```

The example above makes it clear that the periodogram as an estimator is susceptible to large uncertainties, and we need to find a way to reduce the variance. Not surprisingly, this result follows if we think about the periodogram, $I(\omega_j)$ as an estimator of the spectral density $f(\omega)$ and realize that it is the sum of squares of only two random variables for any sample size. The solution to this dilemma is suggested by the analogy with classical statistics where we look for independent random variables with the same variance and average the squares of these common variance observations. Independence and equality of variance do not hold in the time series case, but the covariance structure of the two adjacent estimators given in Example 4.9 suggests that for neighboring frequencies, these assumptions are approximately true.

4.5 Nonparametric Spectral Estimation

To continue the discussion that ended the previous section, we introduce a frequency band, $\mathcal{B}$, of $L << n$ contiguous fundamental frequencies, centered around frequency $\omega_j = j/n$, which is chosen close to a frequency of interest, ω. For frequencies of the form $\omega^* = \omega_j + k/n$, let

$$\mathcal{B} = \left\{\omega^* : \ \omega_j - \frac{m}{n} \le \omega^* \le \omega_j + \frac{m}{n}\right\}, \tag{4.44}$$

where

$$L = 2m + 1 \tag{4.45}$$

is an odd number, chosen such that the spectral values in the interval $\mathcal{B}$,

$$f(\omega_j + k/n), \quad k = -m, \ldots, 0, \ldots, m$$

are approximately equal to $f(\omega)$.

We now define an averaged (or smoothed) periodogram as the average of the periodogram values, say,

$$\bar{f}(\omega) = \frac{1}{L}\sum_{k=-m}^{m} I(\omega_j + k/n), \tag{4.46}$$

over the band $\mathcal{B}$. Under the assumption that the spectral density is fairly constant in the band $\mathcal{B}$, and in view of (4.41) we can show that under appropriate conditions,[11] for large n, the periodograms in (4.46) are approximately distributed as independent $f(\omega)\chi^2_2/2$ random variables, for $0 < \omega < 1/2$, as long as we keep L fairly small relative to n. This result is discussed formally in §C.2. Thus, under these conditions, $L\bar{f}(\omega)$ is the sum of L approximately independent $f(\omega)\chi^2_2/2$ random variables. It follows that, for large n,

$$\frac{2L\bar{f}(\omega)}{f(\omega)} \dot{\sim} \chi^2_{2L} \tag{4.47}$$

where $\dot{\sim}$ means *is approximately distributed as*.

In this scenario, where we smooth the periodogram by simple averaging, it seems reasonable to call the width of the frequency interval defined by (4.44),

$$B_w = \frac{L}{n}, \tag{4.48}$$

the bandwidth.[12] The concept of the bandwidth, however, becomes more complicated with the introduction of spectral estimators that smooth with unequal weights. Note (4.48) implies the degrees of freedom can be expressed as

$$2L = 2B_w n, \tag{4.49}$$

or twice the time-bandwidth product. The result (4.47) can be rearranged to obtain an approximate $100(1-\alpha)\%$ confidence interval of the form

$$\frac{2L\bar{f}(\omega)}{\chi^2_{2L}(1-\alpha/2)} \le f(\omega) \le \frac{2L\bar{f}(\omega)}{\chi^2_{2L}(\alpha/2)} \tag{4.50}$$

for the true spectrum, $f(\omega)$.

Many times, the visual impact of a spectral density plot will be improved by plotting the logarithm of the spectrum instead of the spectrum (the log

[11] The conditions, which are sufficient, are that x_t is a linear process, as described in Property 4.4, with $\sum_j \sqrt{|j|}\,|\psi_j| < \infty$, and w_t has a finite fourth moment.

[12] The bandwidth value used in R is based on Grenander (1951). The basic idea is that bandwidth can be related to the standard deviation of the weighting distribution. For the uniform distribution on the frequency range $-m/n$ to m/n, the standard deviation is $L/n\sqrt{12}$ (using a continuity correction). Consequently, in the case of (4.46), R will report a bandwidth of $L/n\sqrt{12}$, which amounts to dividing our definition by $\sqrt{12}$. Note that in the extreme case $L = n$, we would have $B_w = 1$ indicating that everything was used in the estimation; in this case, R would report a bandwidth of $1/\sqrt{12}$. There are many definitions of bandwidth and an excellent discussion may be found in Percival and Walden (1993, §6.7).

transformation is the variance stabilizing transformation in this situation). This phenomenon can occur when regions of the spectrum exist with peaks of interest much smaller than some of the main power components. For the log spectrum, we obtain an interval of the form

$$\big[\log \bar{f}(\omega) + \log 2L - \log \chi^2_{2L}(1-\alpha/2), \quad \log \bar{f}(\omega) + \log 2L - \log \chi^2_{2L}(\alpha/2)\big]. \tag{4.51}$$

We can also test hypotheses relating to the equality of spectra using the fact that the distributional result (4.47) implies that the ratio of spectra based on roughly independent samples will have an approximate $F_{2L,2L}$ distribution. The independent estimators can either be from different frequency bands or from different series.

If zeros are appended before computing the spectral estimators, we need to adjust the degrees of freedom and an approximation is to replace $2L$ by $2Ln/n'$. Hence, we define the adjusted degrees of freedom as

$$df = \frac{2Ln}{n'} \tag{4.52}$$

and use it instead of $2L$ in the confidence intervals (4.50) and (4.51). For example, (4.50) becomes

$$\frac{df\,\bar{f}(\omega)}{\chi^2_{df}(1-\alpha/2)} \le f(\omega) \le \frac{df\,\bar{f}(\omega)}{\chi^2_{df}(\alpha/2)}. \tag{4.53}$$

A number of assumptions are made in computing the approximate confidence intervals given above, which may not hold in practice. In such cases, it may be reasonable to employ resampling techniques such as one of the parametric bootstraps proposed by Hurvich and Zeger (1987) or a nonparametric local bootstrap proposed by Paparoditis and Politis (1999). To develop the bootstrap distributions, we assume that the contiguous DFTs in a frequency band of the form (4.44) all came from a time series with identical spectrum $f(\omega)$. This, in fact, is exactly the same assumption made in deriving the large-sample theory. We may then simply resample the L DFTs in the band, with replacement, calculating a spectral estimate from each bootstrap sample. The sampling distribution of the bootstrap estimators approximates the distribution of the nonparametric spectral estimator. For further details, including the theoretical properties of such estimators, see Paparoditis and Politis (1999).

Before proceeding further, we pause to consider computing the average periodograms for the SOI and Recruitment series, as shown in Figure 4.5.

Example 4.11 Averaged Periodogram for SOI and Recruitment

Generally, it is a good idea to try several bandwidths that seem to be compatible with the general overall shape of the spectrum, as suggested by the periodogram. The SOI and Recruitment series periodograms, previously computed in Figure 4.4, suggest the power in the lower El Niño frequency needs

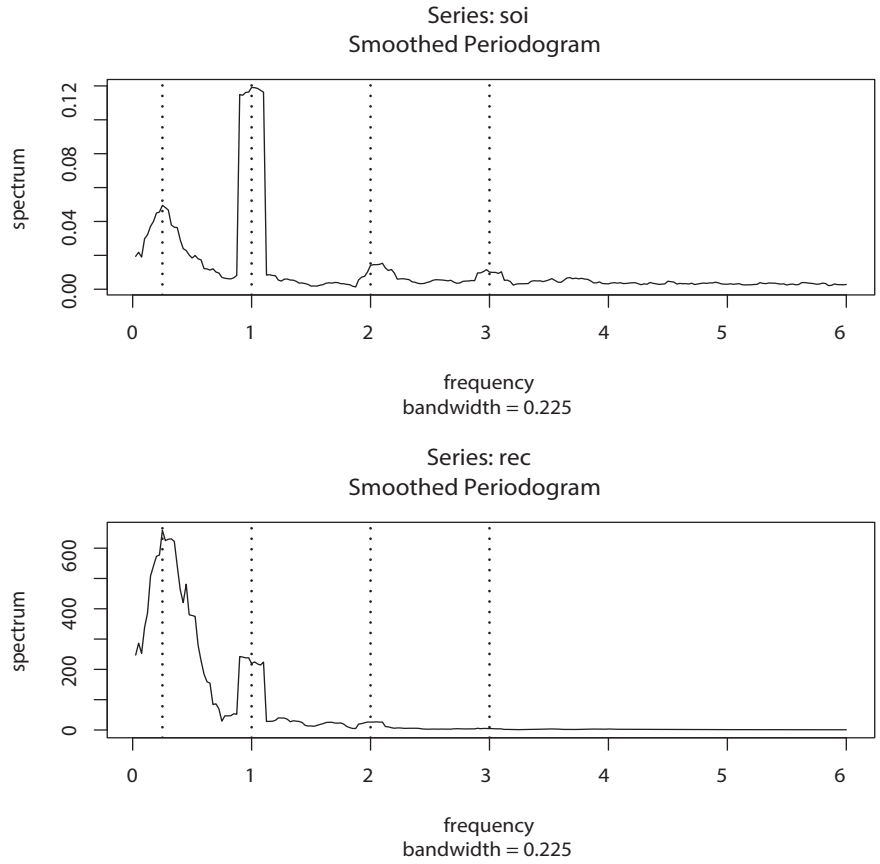

Fig. 4.5. The averaged periodogram of the SOI and Recruitment series $n = 453$, $n' = 480$, $L = 9$, $df = 17$, showing common peaks at the four year period, $\omega = \frac{1}{4}\Delta = 1/48$ cycles/month, the yearly period, $\omega = 1\Delta = 1/12$ cycles/month and some of its harmonics $\omega = k\Delta$ for $k = 2, 3$.

smoothing to identify the predominant overall period. Trying values of L leads to the choice $L = 9$ as a reasonable value, and the result is displayed in Figure 4.5.

The smoothed spectra shown in Figure 4.5 provide a sensible compromise between the noisy version, shown in Figure 4.4, and a more heavily smoothed spectrum, which might lose some of the peaks. An undesirable effect of averaging can be noticed at the yearly cycle, $\omega = 1\Delta$, where the narrow band peaks that appeared in the periodograms in Figure 4.4 have been flattened and spread out to nearby frequencies. We also notice, and have marked, the appearance of harmonics of the yearly cycle, that is, frequencies of the form $\omega = k\Delta$ for $k = 1, 2, \ldots$. Harmonics typically occur when a periodic component is present, but not in a sinusoidal fashion; see Example 4.12.

Figure 4.5 can be reproduced in R using the following commands. The basic call is to the function `mvspec`, which is available in `astsa`; alternately, use R's `spec.pgram`. To compute averaged periodograms, use the Daniell

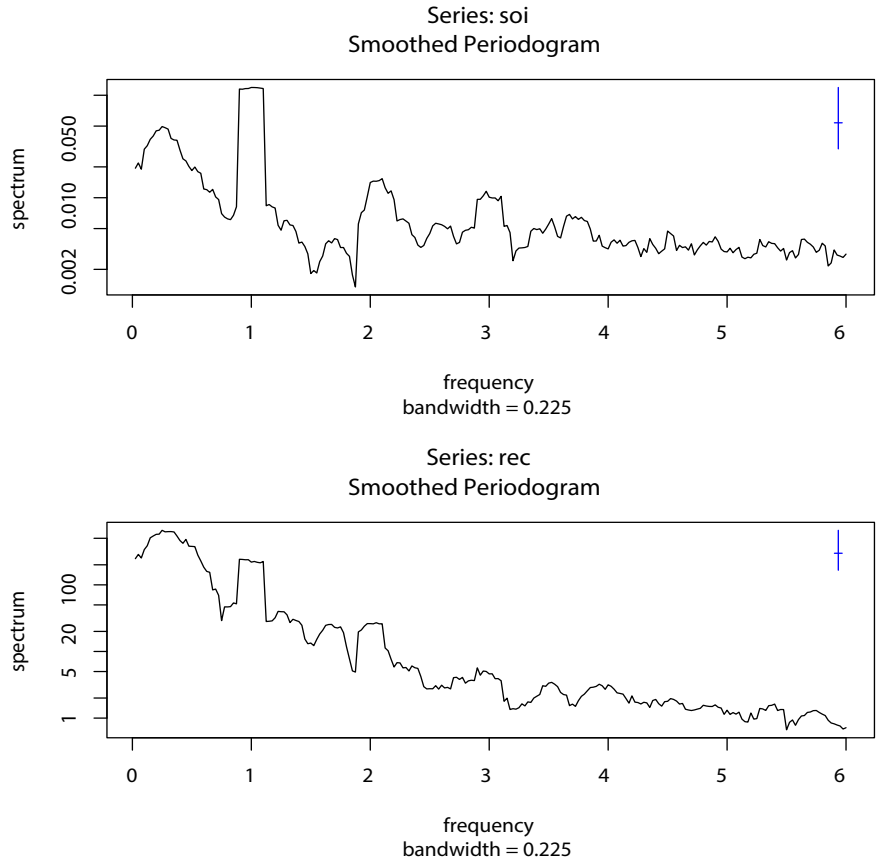

Fig. 4.6. Figure 4.5 with the average periodogram ordinates plotted on a $\log_{10}$ scale. The display in the upper right-hand corner represents a generic 95% confidence interval.

kernel, and specify m, where $L = 2m+1$ ($L = 9$ and $m = 4$ in this example). We will explain the kernel concept later in this section, specifically just prior to Example 4.13.

```
par(mfrow=c(2,1))
k = kernel("daniell", 4)
soi.ave = mvspec(soi, k, log="no")
abline(v=c(.25,1,2,3), lty=2)
# Repeat above lines using rec in place of soi on line 3
soi.ave$bandwidth       # = 0.225
```

The displayed bandwidth (.225) is adjusted for the fact that the frequency scale of the plot is in terms of cycles per year instead of cycles per month (the original unit of the data). Using (4.48), the bandwidth in terms of months is $9/480 = .01875$; the displayed value is simply converted to years, $.01875 \times 12 = .225$.

The adjusted degrees of freedom are $df = 2(9)(453)/480 \approx 17$. We can use this value for the 95% confidence intervals, with $\chi^2_{df}(.025) = 7.56$ and $\chi^2_{df}(.975) = 30.17$. Substituting into (4.53) gives the intervals in Table 4.1

Table 4.1. Confidence Intervals for the Spectra of the SOI and Recruitment Series

Series	ω	Period	Power	Lower	Upper
SOI	1/48	4 years	.05	.03	.11
	1/12	1 year	.12	.07	.27
Recruits	1/48	4 years	6.59	3.71	14.82
$\times 10^2$	1/12	1 year	2.19	1.24	4.93

for the two frequency bands identified as having the maximum power. To examine the two peak power possibilities, we may look at the 95% confidence intervals and see whether the lower limits are substantially larger than adjacent baseline spectral levels. For example, the El Niño frequency of 48 months has lower limits that exceed the values the spectrum would have if there were simply a smooth underlying spectral function without the peaks. The relative distribution of power over frequencies is different, with the SOI having less power at the lower frequency, relative to the seasonal periods, and the recruit series having relatively more power at the lower or El Niño frequency.

The entries in Table 4.1 for SOI can be obtained in R as follows:

```
df = soi.ave$df          #  df = 16.9875  (returned values)
U = qchisq(.025, df)     #  U = 7.555916
L = qchisq(.975, df)     #  L = 30.17425
soi.ave$spec[10]         #  0.0495202
soi.ave$spec[40]         #  0.1190800
# intervals
df*soi.ave$spec[10]/L    # 0.0278789
df*soi.ave$spec[10]/U    # 0.1113333
df*soi.ave$spec[40]/L    # 0.0670396
df*soi.ave$spec[40]/U    # 0.2677201
# repeat above commands with soi replaced by rec
```

Finally, Figure 4.6 shows the averaged periodograms in Figure 4.5 plotted on a $\log_{10}$ scale. This is the default plot in R, and these graphs can be obtained by removing the statement `log="no"`. Notice that the default plot also shows a generic confidence interval of the form (4.51) (with log replaced by $\log_{10}$) in the upper right-hand corner. To use it, imagine placing the tick mark on the averaged periodogram ordinate of interest; the resulting bar then constitutes an approximate 95% confidence interval for the spectrum at that frequency. We note that displaying the estimates on a log scale tends to emphasize the harmonic components.

Example 4.12 Harmonics

In the previous example, we saw that the spectra of the annual signals displayed minor peaks at the harmonics; that is, the signal spectra had a large peak at $\omega = 1\Delta = 1/12$ cycles/month (the one-year cycle) and minor peaks

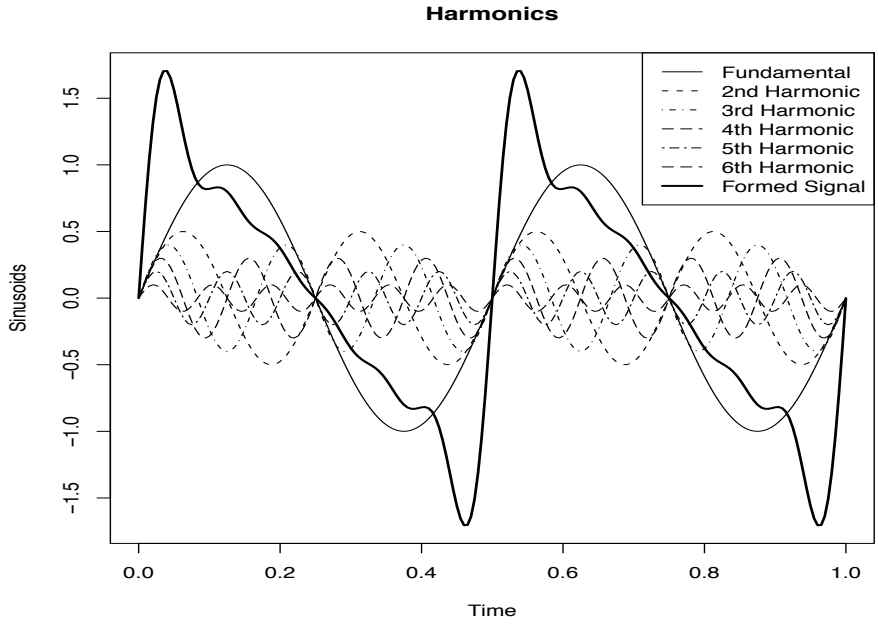

Fig. 4.7. A signal (thick solid line) formed by a fundamental sinusoid (thin solid line) oscillating at two cycles per unit time and its harmonics as specified in (4.54).

at its harmonics $\omega = k\Delta$ for $k = 2, 3, \ldots$ (two-, three-, and so on, cycles per year). This will often be the case because most signals are not perfect sinusoids (or perfectly cyclic). In this case, the harmonics are needed to capture the non-sinusoidal behavior of the signal. As an example, consider the signal formed in Figure 4.7 from a (fundamental) sinusoid oscillating at two cycles per unit time along with the second through sixth harmonics at decreasing amplitudes. In particular, the signal was formed as

$$\begin{aligned} x_t &= \sin(2\pi 2t) + .5\sin(2\pi 4t) + .4\sin(2\pi 6t) \\ &\quad + .3\sin(2\pi 8t) + .2\sin(2\pi 10t) + .1\sin(2\pi 12t) \end{aligned} \tag{4.54}$$

for $0 \le t \le 1$. Notice that the signal is non-sinusoidal in appearance and rises quickly then falls slowly.

A figure similar to Figure 4.7 can be generated in R as follows.

```
t = seq(0, 1, by=1/200)
amps = c(1, .5, .4, .3, .2, .1)
x = matrix(0, 201, 6)
for (j in 1:6) x[,j] = amps[j]*sin(2*pi*t*2*j)
x = ts(cbind(x, rowSums(x)), start=0, deltat=1/200)
ts.plot(x, lty=c(1:6, 1), lwd=c(rep(1,6), 2), ylab="Sinusoids")
names = c("Fundamental","2nd Harmonic","3rd Harmonic","4th Harmonic",
    "5th Harmonic", "6th Harmonic", "Formed Signal")
legend("topright", names, lty=c(1:6, 1), lwd=c(rep(1,6), 2))
```

Example 4.11 points out the necessity for having some relatively systematic procedure for deciding whether peaks are significant. The question of deciding

whether a single peak is significant usually rests on establishing what we might think of as a baseline level for the spectrum, defined rather loosely as the shape that one would expect to see if no spectral peaks were present. This profile can usually be guessed by looking at the overall shape of the spectrum that includes the peaks; usually, a kind of baseline level will be apparent, with the peaks seeming to emerge from this baseline level. If the lower confidence limit for the spectral value is still greater than the baseline level at some predetermined level of significance, we may claim that frequency value as a statistically significant peak. To be consistent with our stated indifference to the upper limits, we might use a one-sided confidence interval.

An important aspect of interpreting the significance of confidence intervals and tests involving spectra is that typically, more than one frequency will be of interest, so that we will potentially be interested in simultaneous statements about a whole collection of frequencies. For example, it would be unfair to claim in Table 4.1 the two frequencies of interest as being statistically significant and all other potential candidates as nonsignificant at the overall level of $\alpha = .05$. In this case, we follow the usual statistical approach, noting that if K statements $S_1, S_2, \ldots, S_k$ are made at significance level α, i.e., $P\{S_k\} = 1 - \alpha$, then the overall probability all statements are true satisfies the Bonferroni inequality

$$P\{\text{all } S_k \text{ true}\} \geq 1 - K\alpha. \tag{4.55}$$

For this reason, it is desirable to set the significance level for testing each frequency at α/K if there are K potential frequencies of interest. If, a priori, potentially $K = 10$ frequencies are of interest, setting $\alpha = .01$ would give an overall significance level of bound of .10.

The use of the confidence intervals and the necessity for smoothing requires that we make a decision about the bandwidth B_w over which the spectrum will be essentially constant. Taking too broad a band will tend to smooth out valid peaks in the data when the constant variance assumption is not met over the band. Taking too narrow a band will lead to confidence intervals so wide that peaks are no longer statistically significant. Thus, we note that there is a conflict here between variance properties or bandwidth stability, which can be improved by increasing B_w and resolution, which can be improved by decreasing B_w. A common approach is to try a number of different bandwidths and to look qualitatively at the spectral estimators for each case.

To address the problem of resolution, it should be evident that the flattening of the peaks in Figure 4.5 and Figure 4.6 was due to the fact that simple averaging was used in computing $\bar{f}(\omega)$ defined in (4.46). There is no particular reason to use simple averaging, and we might improve the estimator by employing a weighted average, say

$$\widehat{f}(\omega) = \sum_{k=-m}^{m} h_k \, I(\omega_j + k/n), \tag{4.56}$$

using the same definitions as in (4.46) but where the weights $h_k > 0$ satisfy

$$\sum_{k=-m}^{m} h_k = 1.$$

In particular, it seems reasonable that the resolution of the estimator will improve if we use weights that decrease as distance from the center weight h_0 increases; we will return to this idea shortly. To obtain the averaged periodogram, $\bar{f}(\omega)$, in (4.56), set $h_k = L^{-1}$, for all k, where $L = 2m + 1$. The asymptotic theory established for $\bar{f}(\omega)$ still holds for $\widehat{f}(\omega)$ provided that the weights satisfy the additional condition that if $m \to \infty$ as $n \to \infty$ but $m/n \to 0$, then

$$\sum_{k=-m}^{m} h_k^2 \to 0.$$

Under these conditions, as $n \to \infty$,

(i) $E\left(\widehat{f}(\omega)\right) \to f(\omega)$

(ii) $\left(\sum_{k=-m}^{m} h_k^2\right)^{-1} \operatorname{cov}\left(\widehat{f}(\omega), \widehat{f}(\lambda)\right) \to f^2(\omega) \quad \text{for } \omega = \lambda \neq 0, 1/2.$

In (ii), replace $f^2(\omega)$ by 0 if $\omega \neq \lambda$ and by $2f^2(\omega)$ if $\omega = \lambda = 0$ or $1/2$.

We have already seen these results in the case of $\bar{f}(\omega)$, where the weights are constant, $h_k = L^{-1}$, in which case $\sum_{k=-m}^{m} h_k^2 = L^{-1}$. The distributional properties of (4.56) are more difficult now because $\widehat{f}(\omega)$ is a weighted linear combination of asymptotically independent χ^2 random variables. An approximation that seems to work well is to replace L by $\left(\sum_{k=-m}^{m} h_k^2\right)^{-1}$. That is, define

$$L_h = \left(\sum_{k=-m}^{m} h_k^2\right)^{-1} \tag{4.57}$$

and use the approximation[13]

$$\frac{2L_h \widehat{f}(\omega)}{f(\omega)} \dot{\sim} \chi^2_{2L_h}. \tag{4.58}$$

In analogy to (4.48), we will define the bandwidth in this case to be

$$B_w = \frac{L_h}{n}. \tag{4.59}$$

Using the approximation (4.58) we obtain an approximate $100(1-\alpha)\%$ confidence interval of the form

[13] The approximation proceeds as follows: If $\widehat{f} \dot{\sim} c\chi^2_\nu$, where c is a constant, then $E\widehat{f} \approx c\nu$ and $\operatorname{var}\widehat{f} \approx f^2 \sum_k h_k^2 \approx c^2 2\nu$. Solving, $c \approx f \sum_k h_k^2/2 = f/2L_h$ and $\nu \approx 2\left(\sum_k h_k^2\right)^{-1} = 2L_h$.

$$\frac{2L_h\widehat{f}(\omega)}{\chi^2_{2L_h}(1-\alpha/2)} \le f(\omega) \le \frac{2L_h\widehat{f}(\omega)}{\chi^2_{2L_h}(\alpha/2)} \tag{4.60}$$

for the true spectrum, $f(\omega)$. If the data are padded to n', then replace $2L_h$ in (4.60) with $df = 2L_h n/n'$ as in (4.52).

An easy way to generate the weights in R is by repeated use of the Daniell kernel. For example, with $m = 1$ and $L = 2m + 1 = 3$, the Daniell kernel has weights $\{h_k\} = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$; applying this kernel to a sequence of numbers, $\{u_t\}$, produces

$$\widehat{u}_t = \tfrac{1}{3}u_{t-1} + \tfrac{1}{3}u_t + \tfrac{1}{3}u_{t+1}.$$

We can apply the same kernel again to the $\widehat{u}_t$,

$$\widehat{\widehat{u}}_t = \tfrac{1}{3}\widehat{u}_{t-1} + \tfrac{1}{3}\widehat{u}_t + \tfrac{1}{3}\widehat{u}_{t+1},$$

which simplifies to

$$\widehat{\widehat{u}}_t = \tfrac{1}{9}u_{t-2} + \tfrac{2}{9}u_{t-1} + \tfrac{3}{9}u_t + \tfrac{2}{9}u_{t+1} + \tfrac{1}{9}u_{t+2}.$$

The modified Daniell kernel puts half weights at the end points, so with $m = 1$ the weights are $\{h_k\} = \{\frac{1}{4}, \frac{2}{4}, \frac{1}{4}\}$ and

$$\widehat{u}_t = \tfrac{1}{4}u_{t-1} + \tfrac{1}{2}u_t + \tfrac{1}{4}u_{t+1}.$$

Applying the same kernel again to $\widehat{u}_t$ yields

$$\widehat{\widehat{u}}_t = \tfrac{1}{16}u_{t-2} + \tfrac{4}{16}u_{t-1} + \tfrac{6}{16}u_t + \tfrac{4}{16}u_{t+1} + \tfrac{1}{16}u_{t+2}.$$

These coefficients can be obtained in R by issuing the `kernel` command. For example, `kernel("modified.daniell", c(1,1))` would produce the coefficients of the last example. It is also possible to use different values of m, e.g., try `kernel("modified.daniell", c(1,2))` or `kernel("daniell", c(5,3))`. The other kernels that are currently available in R are the Dirichlet kernel and the Fejér kernel, which we will discuss shortly.

Example 4.13 Smoothed Periodogram for SOI and Recruitment

In this example, we estimate the spectra of the SOI and Recruitment series using the smoothed periodogram estimate in (4.56). We used a modified Daniell kernel twice, with $m = 3$ both times. This yields $L_h = 1/\sum_{k=-m}^{m} h_k^2 = 9.232$, which is close to the value of $L = 9$ used in Example 4.11. In this case, the bandwidth is $B_w = 9.232/480 = .019$ and the modified degrees of freedom is $df = 2L_h 453/480 = 17.43$. The weights, h_k, can be obtained and graphed in R as follows:

```
kernel("modified.daniell", c(3,3))

   coef[-6] = 0.006944 = coef[ 6]
   coef[-5] = 0.027778 = coef[ 5]
   coef[-4] = 0.055556 = coef[ 4]
```

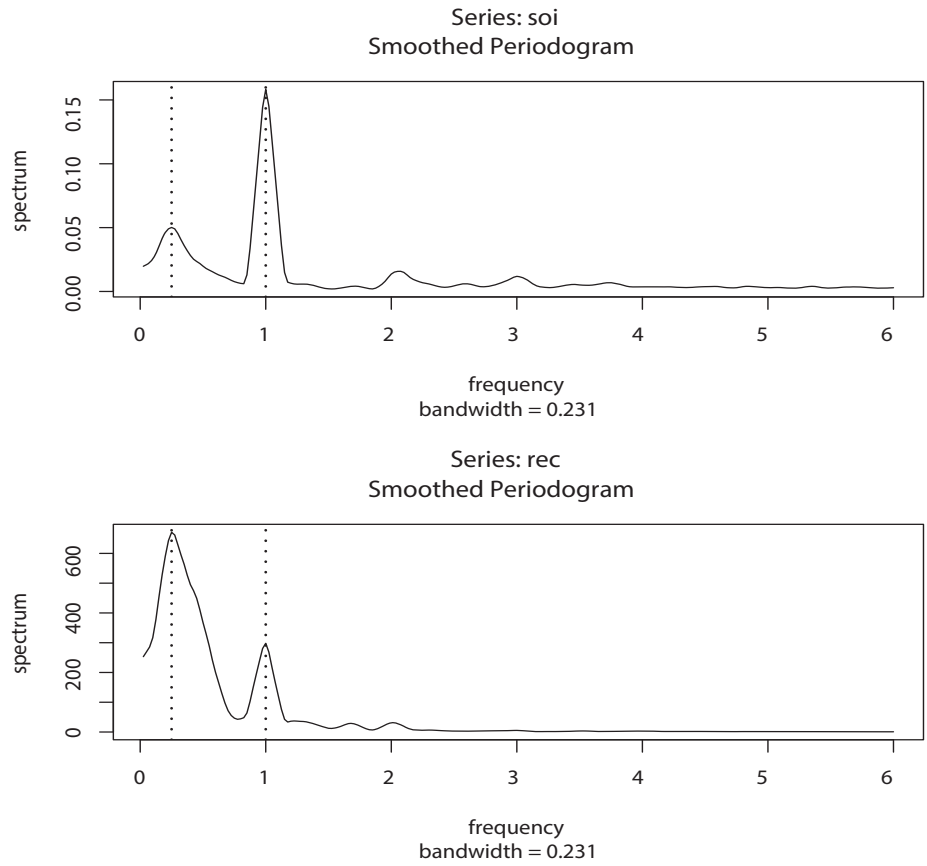

Fig. 4.8. Smoothed spectral estimates of the SOI and Recruitment series; see Example 4.13 for details.

```
coef[-3] = 0.083333 = coef[ 3]
coef[-2] = 0.111111 = coef[ 2]
coef[-1] = 0.138889 = coef[ 1]
coef[ 0] = 0.152778
```

```
plot(kernel("modified.daniell", c(3,3)))  # not shown
```

The resulting spectral estimates can be viewed in Figure 4.8 and we notice that the estimates more appealing than those in Figure 4.5. Figure 4.8 was generated in R as follows; we also show how to obtain df and B_w.

```
par(mfrow=c(2,1))
k = kernel("modified.daniell", c(3,3))
soi.smo = mvspec(soi, k, log="no")
abline(v=1, lty="dotted");  abline(v=1/4, lty="dotted")
## Repeat above lines with rec replacing soi in line 3
df = soi.smo$df             # df = 17.42618
soi.smo$bandwidth           # Bw = 0.2308103
```

Reissuing the `mvspec` commands with `log="no"` removed will result in a figure similar to Figure 4.6. Finally, we mention that the modified Daniell

kernel is used by default. For example, an easier way to obtain `soi.smo` is to issue the command:

```
soi.smo = mvspec(soi, spans=c(7,7))
```

Notice that `spans` is a vector of odd integers, given in terms of $L = 2m + 1$ instead of m. These values give the widths of the modified Daniell smoother to be used to smooth the periodogram.

We are now ready to briefly introduce the concept of *tapering*; a more detailed discussion may be found in Bloomfield (2000, §9.5). Suppose x_t is a mean-zero, stationary process with spectral density $f_x(\omega)$. If we replace the original series by the tapered series

$$y_t = h_t x_t, \tag{4.61}$$

for $t = 1, 2, \ldots, n$, use the modified DFT

$$d_y(\omega_j) = n^{-1/2} \sum_{t=1}^{n} h_t x_t e^{-2\pi i \omega_j t}, \tag{4.62}$$

and let $I_y(\omega_j) = |d_y(\omega_j)|^2$, we obtain (see Problem 4.15)

$$E[I_y(\omega_j)] = \int_{-1/2}^{1/2} W_n(\omega_j - \omega)\ f_x(\omega)\ d\omega \tag{4.63}$$

where

$$W_n(\omega) = |H_n(\omega)|^2 \tag{4.64}$$

and

$$H_n(\omega) = n^{-1/2} \sum_{t=1}^{n} h_t e^{-2\pi i \omega t}. \tag{4.65}$$

The value $W_n(\omega)$ is called a spectral window because, in view of (4.63), it is determining which part of the spectral density $f_x(\omega)$ is being "seen" by the estimator $I_y(\omega_j)$ on average. In the case that $h_t = 1$ for all t, $I_y(\omega_j) = I_x(\omega_j)$ is simply the periodogram of the data and the window is

$$W_n(\omega) = \frac{\sin^2(n\pi\omega)}{n \sin^2(\pi\omega)} \tag{4.66}$$

with $W_n(0) = n$, which is known as the Fejér or modified Bartlett kernel. If we consider the averaged periodogram in (4.46), namely

$$\bar{f}_x(\omega) = \frac{1}{L} \sum_{k=-m}^{m} I_x(\omega_j + k/n),$$

the window, $W_n(\omega)$, in (4.63) will take the form

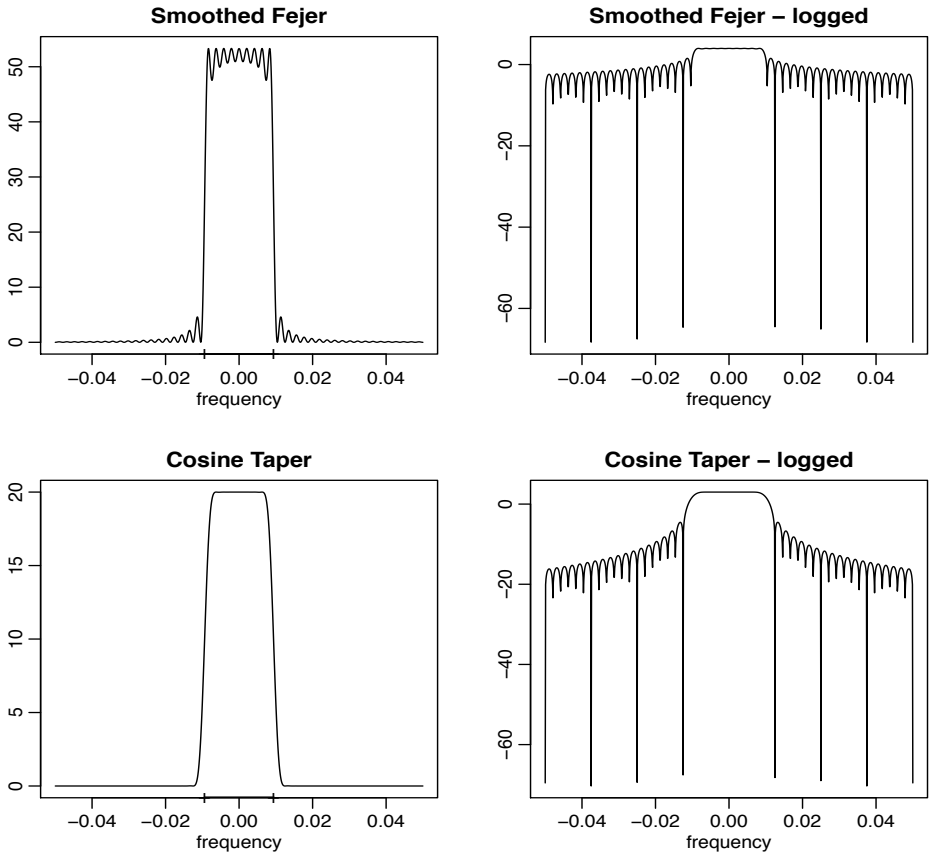

Fig. 4.9. Averaged Fejér window (top row) and the corresponding cosine taper window (bottom row) for $L = 9$, $n = 480$. The extra tic marks on the horizontal axis of the left-hand plots exhibit the predicted bandwidth, $B_w = 9/480 = .01875$.

$$W_n(\omega) = \frac{1}{nL} \sum_{k=-m}^{m} \frac{\sin^2[n\pi(\omega + k/n)]}{\sin^2[\pi(\omega + k/n)]} \,. \tag{4.67}$$

Tapers generally have a shape that enhances the center of the data relative to the extremities, such as a cosine bell of the form

$$h_t = .5\left[1 + \cos\left(\frac{2\pi(t - \bar{t})}{n}\right)\right], \tag{4.68}$$

where $\bar{t} = (n + 1)/2$, favored by Blackman and Tukey (1959). In Figure 4.9, we have plotted the shapes of two windows, $W_n(\omega)$, for $n = 480$ and $L = 9$, when (i) $h_t \equiv 1$, in which case, (4.67) applies, and (ii) h_t is the cosine taper in (4.68). In both cases the predicted bandwidth should be $B_w = 9/480 = .01875$ cycles per point, which corresponds to the "width" of the windows shown in Figure 4.9. Both windows produce an integrated average spectrum over this band but the untapered window in the top panels shows considerable ripples

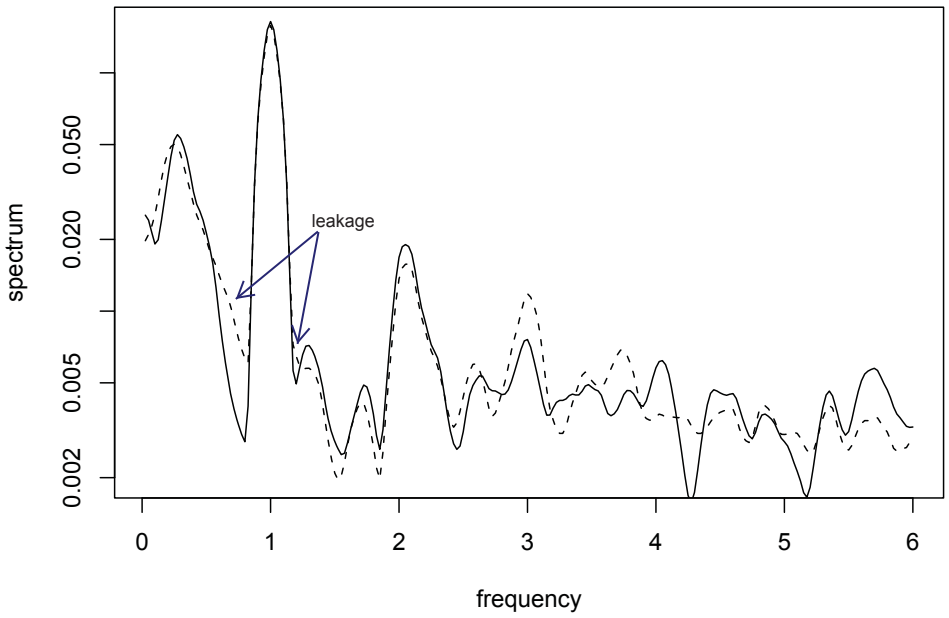

Fig. 4.10. Smoothed spectral estimates of the SOI without tapering (dashed line) and with full tapering (solid line); see Example 4.14 for details.

over the band and outside the band. The ripples outside the band are called sidelobes and tend to introduce frequencies from outside the interval that may contaminate the desired spectral estimate within the band. For example, a large dynamic range for the values in the spectrum introduces spectra in contiguous frequency intervals several orders of magnitude greater than the value in the interval of interest. This effect is sometimes called leakage. Figure 4.9 emphasizes the suppression of the sidelobes in the Fejér kernel when a cosine taper is used.

Example 4.14 The Effect of Tapering the SOI Series

In this example, we examine the effect of tapering on the estimate of the spectrum of the SOI series. The results for the Recruitment series are similar. Figure 4.10 shows two spectral estimates plotted on a log scale. The degree of smoothing here is the same as in Example 4.13. The dashed line in Figure 4.10 shows the estimate without any tapering and hence it is the same as the estimated spectrum displayed in the top of Figure 4.8. The solid line shows the result with full tapering. Notice that the tapered spectrum does a better job in separating the yearly cycle ($\omega = 1$) and the El Niño cycle ($\omega = 1/4$).

The following R session was used to generate Figure 4.10. We note that, by default, `mvspec` does not taper. For full tapering, we use the argument `taper=.5` to instruct `mvspec` to taper 50% of each end of the data; any value between 0 and .5 is acceptable.

```
s0 = mvspec(soi, spans=c(7,7), plot=FALSE)              # no taper
s50 = mvspec(soi, spans=c(7,7), taper=.5, plot=FALSE)   # full taper
```

```
plot(s0$freq, s0$spec, log="y", type="l", lty=2, ylab="spectrum",
    xlab="frequency")     # dashed line
lines(s50$freq, s50$spec)  # solid line
```

We close this section with a brief discussion of lag window estimators. First, consider the periodogram, $I(\omega_j)$, which was shown in (4.22) to be

$$I(\omega_j) = \sum_{|h|<n} \widehat{\gamma}(h) e^{-2\pi i \omega_j h}.$$

Thus, (4.56) can be written as

$$\begin{aligned} \widehat{f}(\omega) &= \sum_{|k|\le m} h_k \, I(\omega_j + k/n) \\ &= \sum_{|k|\le m} h_k \sum_{|h|<n} \widehat{\gamma}(h) e^{-2\pi i(\omega_j + k/n)h} \\ &= \sum_{|h|<n} g(h/n)\, \widehat{\gamma}(h) e^{-2\pi i \omega_j h}. \end{aligned} \tag{4.69}$$

where $g(h/n) = \sum_{|k|\le m} h_k \exp(-2\pi i k h/n)$. Equation (4.69) suggests estimators of the form

$$\widetilde{f}(\omega) = \sum_{|h|\le r} w(h/r)\, \widehat{\gamma}(h) e^{-2\pi i \omega h} \tag{4.70}$$

where $w(\cdot)$ is a weight function, called the lag window, that satisfies

(i) $w(0) = 1$
(ii) $|w(x)| \le 1$ and $w(x) = 0$ for $|x| > 1$,
(iii) $w(x) = w(-x)$.

Note that if $w(x) = 1$ for $|x| < 1$ and $r = n$, then $\widetilde{f}(\omega_j) = I(\omega_j)$, the periodogram. This result indicates the problem with the periodogram as an estimator of the spectral density is that it gives too much weight to the values of $\widehat{\gamma}(h)$ when h is large, and hence is unreliable [e.g, there is only one pair of observations used in the estimate $\widehat{\gamma}(n-1)$, and so on]. The smoothing window is defined to be

$$W(\omega) = \sum_{h=-r}^{r} w(h/r) e^{-2\pi i \omega h}, \tag{4.71}$$

and it determines which part of the periodogram will be used to form the estimate of $f(\omega)$. The asymptotic theory for $\widehat{f}(\omega)$ holds for $\widetilde{f}(\omega)$ under the same conditions and provided $r \to \infty$ as $n \to \infty$ but with $r/n \to 0$. We have

$$E\{\widetilde{f}(\omega)\} \to f(\omega), \tag{4.72}$$

$$\frac{n}{r}\mathrm{cov}\left(\widetilde{f}(\omega), \widetilde{f}(\lambda)\right) \to f^2(\omega) \int_{-1}^{1} w^2(x)dx \qquad \omega = \lambda \ne 0, 1/2. \tag{4.73}$$

In (4.73), replace $f^2(\omega)$ by 0 if $\omega \neq \lambda$ and by $2f^2(\omega)$ if $\omega = \lambda = 0$ or $1/2$.

Many authors have developed various windows and Brillinger (2001, Ch 3) and Brockwell and Davis (1991, Ch 10) are good sources of detailed information on this topic. We mention a few.

The rectangular lag window, which gives uniform weight in (4.70),

$$w(x) = 1, \quad |x| \leq 1,$$

corresponds to the Dirichlet smoothing window given by

$$W(\omega) = \frac{\sin(2\pi r + \pi)\omega}{\sin(\pi\omega)}. \tag{4.74}$$

This smoothing window takes on negative values, which may lead to estimates of the spectral density that are negative a various frequencies. Using (4.73) in this case, for large n we have

$$\text{var}\{\widetilde{f}(\omega)\} \approx \frac{2r}{n} f^2(\omega).$$

The Parzen lag window is defined to be

$$w(x) = \begin{cases} 1 - 6x + 6|x|^3 & |x| < 1/2, \\ 2(1 - |x|)^3 & 1/2 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

This leads to an approximate smoothing window of

$$W(\omega) = \frac{6}{\pi r^3} \frac{\sin^4(r\omega/4)}{\sin^4(\omega/2)}.$$

For large n, the variance of the estimator is approximately

$$\text{var}\{\widetilde{f}(\omega)\} \approx .539 f^2(\omega)/n.$$

The Tukey-Hanning lag window has the form

$$w(x) = \frac{1}{2}(1 + \cos(x)), \quad |x| \leq 1$$

which leads to the smoothing window

$$W(\omega) = \frac{1}{4} D_r(2\pi\omega - \pi/r) + \frac{1}{2} D_r(2\pi\omega) + \frac{1}{4} D_r(2\pi\omega + \pi/r)$$

where $D_r(\omega)$ is the Dirichlet kernel in (4.74). The approximate large sample variance of the estimator is

$$\text{var}\{\widetilde{f}(\omega)\} \approx \frac{3r}{4n} f^2(\omega).$$

The triangular lag window, also known as the Bartlett or Fejér window, given by

$$w(x) = 1 - |x|, \quad |x| \le 1$$

leads to the Fejér smoothing window:

$$W(\omega) = \frac{\sin^2(\pi r\omega)}{r\sin^2(\pi\omega)}.$$

In this case, (4.73) yields

$$\mathrm{var}\{\widetilde{f}(\omega)\} \approx \frac{2r}{3n} f^2(\omega).$$

The idealized rectangular smoothing window, also called the Daniell window, is given by

$$W(\omega) = \begin{cases} r & |\omega| \le 1/2r, \\ 0 & \text{otherwise}, \end{cases}$$

and leads to the sinc lag window, namely

$$w(x) = \frac{\sin(\pi x)}{\pi x}, \quad |x| \le 1.$$

From (4.73) we have

$$\mathrm{var}\{\widetilde{f}(\omega)\} \approx \frac{r}{n} f^2(\omega).$$

For lag window estimators, the width of the idealized rectangular window that leads to the same asymptotic variance as a given lag window estimator is sometimes called the equivalent bandwidth. For example, the bandwidth of the idealized rectangular window is $b_r = 1/r$ and the asymptotic variance is $\frac{1}{nb_r} f^2$. The asymptotic variance of the triangular window is $\frac{2r}{3n} f^2$, so setting $\frac{1}{nb_r} f^2 = \frac{2r}{3n} f^2$ and solving we get $b_r = 3/2r$ as the equivalent bandwidth.

4.6 Parametric Spectral Estimation

The methods of §4.5 lead to estimators generally referred to as nonparametric spectra because no assumption is made about the parametric form of the spectral density. In Property 4.3, we exhibited the spectrum of an ARMA process and we might consider basing a spectral estimator on this function, substituting the parameter estimates from an ARMA(p, q) fit on the data into the formula for the spectral density $f_x(\omega)$ given in (4.15). Such an estimator is called a parametric spectral estimator. For convenience, a parametric spectral estimator is obtained by fitting an AR(p) to the data, where the order p is determined by one of the model selection criteria, such as AIC, AICc, and BIC, defined in (2.19)-(2.21). Parametric autoregressive spectral estimators will often have superior resolution in problems when several closely spaced narrow