12 ESTIMATION OF THE SPECTRUM

In this chapter we discuss estimation of the spectrum. We begin with periodogram analysis, which is a useful technique to search for hidden periodicities. The smoothing of the sample spectrum and related concepts such as the lag and spectral windows are discussed. Empirical examples are presented to illustrate the procedures and results.

12.1 PERIODOGRAM ANALYSIS

12.1.1 The Periodogram

Given a time series of n observations, we can use the results in Section 10.3 to represent the n observations in the following Fourier representation:

$$Z_t = \sum_{k=0}^{[n/2]} (a_k \cos \omega_k t + b_k \sin \omega_k t) \tag{12.1.1}$$

where $\omega_k = 2\pi k/n$, $k = 0, 1, \ldots, [n/2]$, are Fourier frequencies, and

$$a_k = \begin{cases} \dfrac{1}{n}\sum_{t=1}^{n} Z_t \cos \omega_k t, & k = 0 \quad \text{and} \quad k = \dfrac{n}{2} \quad \text{if } n \text{ is even,} \\ \dfrac{2}{n}\sum_{t=1}^{n} Z_t \cos \omega_k t, & k = 1, 2, \ldots, \left[\dfrac{n-1}{2}\right], \end{cases}$$

and

$$b_k = \frac{2}{n}\sum_{t=1}^{n} Z_t \sin \omega_k t, \qquad k = 1, 2, \ldots, \left[\frac{n-1}{2}\right],$$

are Fourier coefficients. It is interesting to note the close relationship between the above Fourier representation and regression analysis. In fact, the Fourier

coefficients are essentially the standard regression coefficients in fitting the following regression model:

$$Z_t = \sum_{k=0}^{[n/2]} (a_k \cos \omega_k t + b_k \sin \omega_k t) + e_t, \tag{12.1.2}$$

where the ω_k are the same Fourier frequencies and the e_t are i.i.d. $N(0, \sigma^2)$. Thus, we can use the notion from regression analysis to regard the Parseval's relation

$$\sum_{t=1}^{n} Z_t^2 = \begin{cases} na_0^2 + \dfrac{n}{2}\displaystyle\sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2), & \text{if } n \text{ is odd,} \\ na_0^2 + \dfrac{n}{2}\displaystyle\sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2) + na_{n/2}^2, & \text{if } n \text{ is even,} \end{cases} \tag{12.1.3}$$

as an analysis of variance presented in Table 12.1.

The quantity $I(\omega_k)$ defined by

$$I(\omega_k) = \begin{cases} na_0^2, & k = 0, \\ \dfrac{n}{2}(a_k^2 + b_k^2), & k = 1, \ldots, (n-1)/2, \\ na_{n/2}^2, & k = \dfrac{n}{2} \text{ when } n \text{ is even,} \end{cases} \tag{12.1.4}$$

is called the periodogram. It was introduced by Schuster (1898) to search for a periodic component in a series.

Table 12.1 Analysis of variance table for periodogram analysis.

Source	Degrees of freedom	Sum of squares
Frequency ω_0 = Mean	1	na_0^2
Frequency $\omega_1 = 2\pi/n$	2	$\frac{n}{2}(a_1^2 + b_1^2)$
Frequency $\omega_2 = 4\pi/n$	2	$\frac{n}{2}(a_2^2 + b_2^2)$
⋮	⋮	⋮
Frequency $\omega_{(n-1)/2} = (n-1)\pi/n$	2	$\frac{n}{2}(a_{(n-1)/2}^2 + b_{(n-1)/2}^2)$
Frequency $\omega_{n/2} = \pi$ (exists only for even n)	1	$na_{n/2}^2$
Total	n	$\sum_{t=1}^{n} Z_t^2$

12.1.2 Sampling Properties of the Periodogram

Assume that $Z_1, Z_2, \ldots,$ and Z_n are i.i.d. $N(0, \sigma^2)$. Then,

$$E(a_k) = \frac{2}{n}\sum_{t=1}^{n} E(Z_t)\cos\omega_k t = 0,$$

and

$$\begin{aligned}\mathrm{Var}(a_k) &= \frac{4}{n^2}\sum_{t=1}^{n}\sigma^2(\cos\omega_k t)^2 \\ &= \frac{4\sigma^2}{n^2}\sum_{t=1}^{n}(\cos\omega_k t)^2 \\ &= \frac{4\sigma^2}{n^2}\cdot\frac{n}{2} = \frac{2\sigma^2}{n}.\end{aligned}$$

Hence, the a_k for $k = 1, 2, \ldots, (n-1)/2$, are i.i.d. $N(0, 2\sigma^2/n)$, and the $na_k^2/2\sigma^2$, $k = 1, 2, \ldots, (n-1)/2$, are i.i.d. chi-squares with 1 degree of freedom. Similarly, the $nb_k^2/2\sigma^2$, $k = 1, 2, \ldots, (n-1)/2$, are i.i.d. chi-squares with 1 degree of freedom. Furthermore, $na_k^2/2\sigma^2$ and $nb_j^2/2\sigma^2$ are independent for $k = 1, \ldots, (n-1)/2$ and $j = 1, 2, \ldots, (n-1)/2$, because, by the orthogonality property of the cosine and sine system, we have

$$\begin{aligned}\mathrm{Cov}(a_k, b_j) &= \frac{4}{n^2}E\left(\sum_{t=1}^{n} Z_t\cos\omega_k t\cdot\sum_{u=1}^{n} Z_u\sin\omega_j u\right) \\ &= \frac{4}{n^2}\sum_{t=1}^{n}\left[E(Z_t^2)\cos\omega_k t\cdot\sin\omega_j t\right] \\ &= \frac{4\sigma^2}{n^2}\sum_{t=1}^{n}\left[\cos\omega_k t\cdot\sin\omega_j t\right] \\ &= 0 \qquad \text{for any } k \text{ and } j. \end{aligned} \tag{12.1.5}$$

It follows that the periodogram ordinates

$$\frac{I(\omega_k)}{\sigma^2} = \frac{n}{2\sigma^2}(a_k^2 + b_k^2) \tag{12.1.6}$$

for $k = 1, 2, \ldots, (n-1)/2$ are i.i.d. chi-squares with 2 degrees of freedom. Clearly, by the same arguments, each of $I(0)/\sigma^2$ and $I(\pi)/\sigma^2$ (for even n) follows the chi-square distribution with 1 degree of freedom. With the adjustment for $I(\pi)$ in mind, we assume, without loss of generality, that the sample size n is odd in the following discussion.

Suppose that a time series can be represented by

$$Z_t = a_0 + a_k\cos\omega_k t + b_k\sin\omega_k t + e_t \tag{12.1.7}$$

where $\omega_k = 2\pi k/n$, $k \neq 0$ and the e_t are i.i.d. $N(0, \sigma^2)$. To examine the above assumed model, we can test

$$H_0: a_k = b_k = 0 \quad \text{vs.} \quad H_1: a_k \neq 0 \quad \text{or} \quad b_k \neq 0$$

using the test statistic

$$\begin{aligned}F &= \frac{\frac{n}{2}(a_k^2 + b_k^2)/2}{\frac{n}{2}\sum_{\substack{j=1\\ j\neq k}}^{[n/2]}(a_j^2 + b_j^2)/(n-3)} \\ &= \frac{(n-3)(a_k^2 + b_k^2)}{2\sum_{\substack{j=1\\ j\neq k}}^{[n/2]}(a_j^2 + b_j^2)}\end{aligned} \tag{12.1.8}$$

which follows the F-distribution $F(2, n-3)$ with 2 and $(n-3)$ degrees of freedom. It should be noted that since the postulated model contains a general mean, a_0, both the numerator and the denominator do not include the term $na_0^2/2$. In other words, they are sum of squares adjusted for the mean. In fact, since the periodogram at frequency zero reflects the sample mean and not the periodicity of the series, it is usually removed from consideration in the analysis. More generally, we can test whether a series contains multiple m periodic components by postulating the model

$$Z_t = a_0 + \sum_{i=1}^{m}(a_{k_i}\cos\omega_{k_i} t + b_{k_i}\sin\omega_{k_i} t) + e_t \tag{12.1.9}$$

where the e_t are i.i.d. $N(0, \sigma^2)$, $\omega_{k_i} = 2\pi k_i/n$, and the set $I = \{k_i: i = 1, \ldots, m\}$ is a subset of $\{k: k = 1, 2, \ldots, [n/2]\}$. The corresponding test statistic

$$F = \frac{(n-2m-1)\sum_{i=1}^{m}(a_{k_i}^2 + b_{k_i}^2)}{2m\sum_{j\in I}(a_j^2 + b_j^2)} \tag{12.1.10}$$

follows the F-distribution $F(2m, n-2m-1)$ with $2m$ and $(n-2m-1)$ degrees of freedom.

12.1.3 Tests for Hidden Periodic Components

In practice, even if we believe that a time series contains a periodic component, the underlying frequency is often unknown. For example, we might test

$$H_0: \alpha = \beta = 0 \quad \text{vs.} \quad H_1: \alpha \neq 0 \quad \text{or} \quad \beta \neq 0$$

for the following model:

$$Z_t = \alpha\cos\omega t + \beta\sin\omega t + e_t \tag{12.1.11}$$

where $\{e_t\}$ is a Gaussian white noise series of i.i.d. $N(0, \sigma^2)$, but the frequency ω is unknown. Because the frequency ω is not known, the F-distribution and

the test statistic as discussed in the previous section are not directly applicable. However, the periodogram analysis can still be used. In fact, the original purpose of the periodogram was to search for "hidden periodicities." If the underlying model indeed contains a single periodic component at frequency ω, it is hoped that the periodogram $I(\omega_k)$ at the Fourier frequency ω_k closest to ω will be the maximum. Thus, we can search out the maximum periodogram ordinate and test whether this ordinate can be reasonably considered as the maximum in a random sample of $[n/2]$ i.i.d. random variables, each being a multiple of a chi-square distribution with 2 degrees of freedom. In this case, a natural test statistic will be

$$I^{(1)}(\omega_{(1)}) = \max\{I(\omega_k)\} \tag{12.1.12}$$

where we use $\omega_{(1)}$ to indicate the Fourier frequency with the maximum periodogram ordinate.

Now, under the null hypothesis H_0, the $\{Z_t\}$ are Gaussian white noise $N(0, \sigma^2)$. Hence, the periodogram ordinates $I(\omega_k)/\sigma^2$, $k = 1, 2, \ldots, [n/2]$, are i.i.d. chi-square random variables with 2 degrees of freedom, which have the probability density function

$$p(x) = \frac{1}{2}e^{-(x/2)}, \qquad 0 \le x < \infty. \tag{12.1.13}$$

Thus, for any $g \ge 0$, we have

$$\begin{aligned} P\left[\frac{I^{(1)}(\omega_{(1)})}{\sigma^2} > g\right] &= 1 - P\left[\frac{I^{(1)}(\omega_{(1)})}{\sigma^2} \le g\right] \\ &= 1 - P\left\{\frac{I(\omega_k)}{\sigma^2} \le g, \quad k = 1, 2, \ldots, \left[\frac{n}{2}\right]\right\} \\ &= 1 - \left\{\int_0^g \frac{1}{2}e^{-(x/2)}dx\right\}^{[n/2]} \\ &= 1 - \left(1 - e^{-(g/2)}\right)^{[n/2]} \end{aligned} \tag{12.1.14}$$

If σ^2 were known, we could use (12.1.14) to derive an exact test for the maximum ordinate.

However, in practice, σ^2 is often unknown and has to be estimated. To derive an unbiased estimator of σ^2, we note that

$$E\left[\sum_{k=1}^{[n/2]} I(\omega_k)\right] = \left[\frac{n}{2}\right]2\sigma^2.$$

Thus,

$$\hat{\sigma}^2 = \frac{1}{2\left[\frac{n}{2}\right]}\sum_{k=1}^{[n/2]} I(\omega_k) \tag{12.1.15}$$

is an unbiased estimator of σ^2, and this leads to the test statistics

$$V = \frac{I^{(1)}(\omega_{(1)})}{2\left[\frac{n}{2}\right]\sum_{k=1}^{[n/2]} I(\omega_k)}. \tag{12.1.16}$$

Using the fact that $I(\omega_k)$, $k = 1, 2, \ldots, [n/2]$ are independent, we have

$$\operatorname{Var}(\hat{\sigma}^2) \longrightarrow 0 \qquad \text{as } n \longrightarrow \infty. \tag{12.1.17}$$

It follows that $\hat{\sigma}^2$ is a consistent estimator of σ^2. Thus, for large samples, V can be approximated by the same distribution as $I^{(1)}(\omega_{(1)})/\sigma^2$, i.e., for any $g \ge 0$,

$$P(V > g) \simeq 1 - \left(1 - e^{-(g/2)}\right)^{[n/2]}. \tag{12.1.18}$$

An exact test for $I^{(1)}(\omega_{(1)})$ was derived by Fisher (1929), based on the following statistic:

$$T = \frac{I^{(1)}(\omega_{(1)})}{\sum_{k=1}^{[n/2]} I(\omega_k)}. \tag{12.1.19}$$

Under the null hypothesis of the Gaussian white noise process $N(0, \sigma^2)$ for Z_t, Fisher (1929) showed that

$$P(T > g) = \sum_{j=1}^{m}(-1)^{(j-1)}\binom{N}{j}(1 - jg)^{N-1} \tag{12.1.20}$$

where $N = [n/2]$, $g > 0$, and m is the largest integer less than $1/g$. Thus, for any given significance level α, we can use Equation (12.1.20) to find the critical value g_α such that

$$P(T > g_\alpha) = \alpha.$$

If the T value calculated from the series is larger than g_α, then we reject the null hypothesis and conclude that the series Z_t contains a periodic component. This test procedure is known as Fisher's test.

The distribution of T for the significance level $\alpha = .05$ as given by Fisher (1929) is shown in Table 12.2.

The third column in Table 12.2 is an approximation obtained by using only the first term in (12.1.20), i.e.,

$$P(T > g) \simeq N(1 - g)^{N-1}. \tag{12.1.21}$$

The approximation is very close to the exact result. Hence, for most practical purposes, we can use Equation (12.1.21) to derive the critical value g_α for the test.

It should be noted that under the null hypothesis H_0 in (12.1.11), a significant value of $I^{(1)}(\omega_{(1)})$ leads to the rejection of H_0. This implies that there exists a periodic component in the series at some frequency ω. This ω is not

Table 12.2 The critical values for $\alpha = .05$ for the ratio of the largest periodogram ordinate to the sum.

N^*	g_α (by exact formula)	g_α (by first term only)
5	.68377	.68377
10	.44495	.44495
15	.33462	.33463
20	.27040	.27046
25	.22805	.22813
30	.19784	.19794
35	.17513	.17525
40	.15738	.15752
45	.14310	.14324
50	.13135	.13149

*$N = (n-1)/2$ if n is odd and $N = (n/2-1)$ if n is even.

necessarily equal to $\omega_{(1)}$, since $\omega_{(1)}$ is chosen only from the Fourier frequencies and not from all the possible frequencies between 0 and π. However, Hartley (1949) has shown that the unknown ω with the maximum periodic component can be safely estimated by $\omega_{(1)}$, and the probability that $|\omega - \omega_{(1)}| > 2\pi/n$ is less than the significance level of the test.

Let $I^{(2)}(\omega_{(2)})$ be the second largest periodogram ordinate at Fourier frequency $\omega_{(2)}$. Whittle (1952) suggested extending Fisher's test for this second largest ordinate based on the following test statistic:

$$T_2 = \frac{I^{(2)}(\omega_{(2)})}{\left\{\sum_{k=1}^{[n/2]} I(\omega_k)\right\} - I^{(1)}(\omega_{(1)})} \tag{12.1.22}$$

where the distribution in (12.1.20) is used but with N being replaced by $(N-1)$. The procedure can be continued until an insignificant result is obtained. This leads to an estimate of m, the number of periodic components in the Equation (12.1.9).

Alternatively, we can use standard regression analysis and construct the usual regression test. For example, to examine whether a monthly time series is seasonal with period 12, we might postulate the model

$$Z_t = a_0 + \sum_{k=1}^{6} [a_k \cos(2\pi k/12) + b_k \sin(2\pi k/12)] + e_t \tag{12.1.23}$$

where the e_t are i.i.d. $N(0, \sigma^2)$. The test is then performed by the F ratio of the mean square for the six seasonal frequencies $2\pi k/12$, for $k = 1, 2, 3, 4, 5, 6$, to the residual mean square.

Before illustrating the procedure with an empirical series, we should note that we could fit a general sine-cosine model

$$Z_t = \mu + \sum_{j=1}^{m} (\alpha_j \cos \lambda_j t + \beta_j \sin \lambda_j t) + e_t \tag{12.1.24}$$

where the e_t are i.i.d. $N(0, \sigma^2)$. However, unless the λ_j are Fourier frequencies the components at two different frequencies λ_i and λ_j are not independent because the sine-cosine and complex exponential systems are complete and orthogonal only over the Fourier frequencies.

Example 12.1 In this example we perform the periodogram analysis of the logarithms of the annual lynx pelt sales of the Hudson's Bay Company between 1857 and 1911 ($n = 55$), which we fitted an ARMA model earlier in Chapters 6 and 7. The results are given in Table 12.3 and plotted in Figure 12.1 in terms of both frequency and period. The periodogram is clearly

Table 12.3 Periodogram analysis of the lynx pelt sales.

k	Freq. (ω_k)	Period (P)	$I(\omega_k)$	F
1	0.11424	55.00000	0.20038	0.68914
2	0.22848	27.50000	0.00043	0.00144
3	0.34272	18.33333	0.29203	1.01667
4	0.45696	13.75000	0.13967	0.47652
5	0.57120	11.00000	0.25225	0.87353
6	0.68544	9.16667	5.87412	80.97115
7	0.79968	7.85714	0.43732	1.55269
8	0.91392	6.87500	0.10167	0.34516
9	1.02816	6.11111	0.07457	0.25226
10	1.14240	5.50000	0.06830	0.23086
11	1.25664	5.00000	0.07283	0.24632
12	1.37088	4.58333	0.04358	0.14683
13	1.48512	4.23077	0.04723	0.15921
14	1.59936	3.92857	0.01855	0.06230
15	1.71360	3.66667	0.01227	0.04117
16	1.82784	3.43750	0.02088	0.07014
17	1.94207	3.23529	0.00312	0.01046
18	2.05631	3.05556	0.03014	0.10137
19	2.17055	2.89474	0.00545	0.01827
20	2.28479	2.75000	0.00722	0.02421
21	2.39903	2.61905	0.00087	0.00292
22	2.51327	2.50000	0.00270	0.00905
23	2.62751	2.39130	0.02743	0.09223
24	2.74175	2.29167	0.00065	0.00218
25	2.85599	2.20000	0.00651	0.02183
26	2.97023	2.11538	0.01561	0.05240
27	3.08447	2.03704	0.00454	0.01522

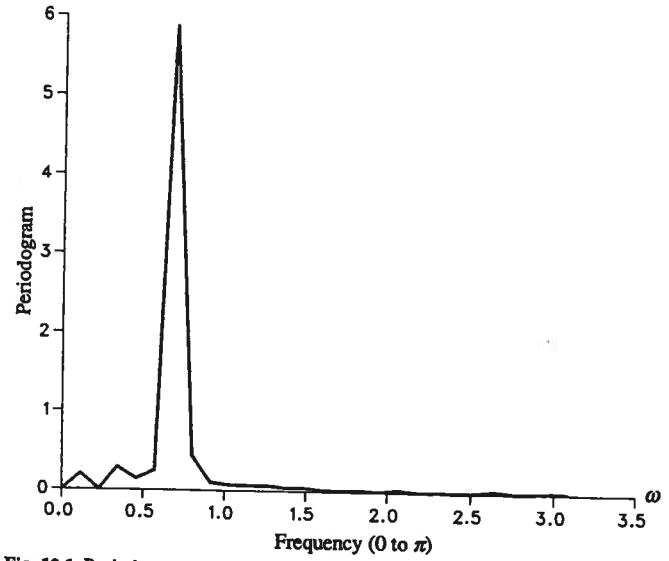

Fig. 12.1 Periodogram of the annual Canadian lynx pelt sales.

dominated by a very large peak at frequency $\omega_6 = 2\pi(6)/55 = .68544$. This frequency corresponds to a period of $P = 2\pi/\omega = 2\pi/\omega_6 = 9.1667$ years. It indicates that the data exhibit an approximate 9-year cycle. Also included in Table 12.3 are the values of the F statistics given in (12.1.8) to test the significance of the periodogram at each Fourier frequency. For significance level $\alpha = .05$, $F_{.05}(2, 52) = 3.19533$, and the periodogram is significant only at frequency $\omega_6 = .68544$.

For Fisher's exact test for the maximum periodogram, we have

$$T = \frac{I^{(1)}(\omega_{(1)})}{\sum_{k=1}^{[n/2]} I(\omega_k)} = \frac{I^{(1)}(\omega_6)}{\sum_{k=1}^{27} I(\omega_k)} = \frac{5.87412}{7.76032} = .756943.$$

From Table 12.2, $g_{.05} = .22805$ for $N = 25$, and $g_{.05} = .19784$ for $N = 30$. More precisely for $N = 27$, we can use the first term approximation given in (12.1.21). For $\alpha = .05$, we have

$$N(1-g)^{N-1} = 27(1-g)^{26} = .05$$

which gives $g = .21493$. Since $T = .756943 >> .21493$, the result is highly significant and we conclude that the series contains a periodic component at the frequency $\omega_6 = .68544$.

12.2 THE SAMPLE SPECTRUM

In this section we investigate the estimation of the spectrum of time series with absolutely summable autocovariances. From (11.1.3), the spectrum is given by

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\omega k} \tag{12.2.1a}$$

$$= \frac{1}{2\pi}\left(\gamma_0 + 2\sum_{k=1}^{\infty} \gamma_k \cos \omega k\right), \quad -\pi \le \omega \le \pi. \tag{12.2.1b}$$

Based on sample data, it is natural to estimate $f(\omega)$ by replacing the theoretical autocovariances γ_k by the sample autocovariances $\hat{\gamma}_k$. However, for a given time series of n observations, we can only calculate $\hat{\gamma}_k$ for $k = 0, 1, \ldots, (n-1)$. Hence, we estimate $f(\omega)$ by

$$\hat{f}(\omega) = \frac{1}{2\pi} \sum_{k=-(n-1)}^{(n-1)} \hat{\gamma}_k e^{-i\omega k} \tag{12.2.2a}$$

$$= \frac{1}{2\pi}\left(\hat{\gamma}_0 + 2\sum_{k=1}^{n-1} \hat{\gamma}_k \cos \omega k\right) \tag{12.2.2b}$$

and call it the sample spectrum. Because $\hat{\gamma}_k$ is asymptotically unbiased, as discussed in Section 2.5.2, we have

$$\lim_{n\to\infty} E(\hat{f}(\omega)) = f(\omega). \tag{12.2.3}$$

Thus, $\hat{f}(\omega)$ is asymptotically unbiased and looks very promising as a potential estimator for $f(\omega)$.

To further examine the properties of the sample spectrum, let us consider $f(\hat{\omega}_k)$ at the Fourier frequencies $\omega_k = 2\pi k/n$, $k = 1, \ldots, [n/2]$. At these Fourier frequencies, the sample spectrum and the periodogram are closely related. To see that, we note

$$\begin{aligned} I(\omega_k) &= \frac{n}{2}(a_k^2 + b_k^2) \\ &= \frac{n}{2}(a_k - ib_k)(a_k + ib_k) \\ &= \frac{n}{2}\left[\frac{2}{n}\sum_{t=1}^{n} Z_t(\cos \omega_k t - i \sin \omega_k t)\right]\left[\frac{2}{n}\sum_{t=1}^{n} Z_t(\cos \omega_k t + i \sin \omega_k t)\right] \end{aligned}$$

$$= \frac{2}{n}\left[\sum_{t=1}^{n} Z_t e^{-i\omega_k t}\right]\left[\sum_{t=1}^{n} Z_t e^{i\omega_k t}\right]$$
$$= \frac{2}{n}\left[\sum_{t=1}^{n} (Z_t - \bar{Z}) e^{-i\omega_k t}\right]\left[\sum_{t=1}^{n} (Z_t - \bar{Z}) e^{i\omega_k t}\right]$$
$$= \frac{2}{n}\sum_{t=1}^{n}\sum_{s=1}^{n} (Z_t - \bar{Z})(Z_s - \bar{Z}) e^{-i\omega_k (t-s)}, \qquad (12.2.4)$$

where we use the fact that $\sum_{t=1}^{n} e^{i\omega_k t} = \sum_{t=1}^{n} e^{-i\omega_k t} = 0$.

Because

$$\hat{\gamma}_j = \frac{1}{n}\sum_{t=1}^{n-j} (Z_t - \bar{Z})(Z_{t+j} - \bar{Z}),$$

letting $j = t - s$ in (12.2.4), we get

$$I(\omega_k) = 2\sum_{j=-(n-1)}^{n-1} \hat{\gamma}_j e^{-i\omega_k j} \qquad (12.2.5a)$$
$$= 2\left\{\hat{\gamma}_0 + 2\sum_{j=1}^{n-1} \hat{\gamma}_j \cos \omega_k j\right\}. \qquad (12.2.5b)$$

Hence, from (12.2.2b), we have

$$\hat{f}(\omega_k) = \frac{1}{4\pi} I(\omega_k),\ k = 1, 2, \ldots, [n/2], \qquad (12.2.6)$$

where we note that $\hat{f}(\omega_{n/2}) = I(\omega_{n/2})/2\pi = n a_{n/2}^2/2\pi$ if n is even.

It follows that if Z_t is a Gaussian white noise series with mean 0 and constant variance σ^2, then $\hat{f}(\omega_k)$, for $k = 1, 2, \ldots, (n-1)/2$, are distributed independently and identically as $(\sigma^2/4\pi)\chi^2(2) = (\sigma^2/2\pi)\chi^2(2)/2$, and we denote it as

$$\hat{f}(\omega_k) \sim \frac{\sigma^2}{2\pi}\frac{\chi^2(2)}{2} \qquad (12.2.7)$$

where $\chi^2(2)$ is the chi-square distribution with 2 degrees of freedom. Here, we note that $\sigma^2/2\pi$ in (12.2.7) is the spectrum of Z_t. In general, following the same arguments, it can be shown that if Z_t is a Gaussian process with the spectrum $f(\omega)$, then

$$\hat{f}(\omega_k) \sim f(\omega_k)\frac{\chi^2(2)}{2}. \qquad (12.2.8)$$

Now,

$$E(\hat{f}(\omega_k)) = E\left[f(\omega_k)\frac{\chi^2(2)}{2}\right] = f(\omega_k), \qquad (12.2.9)$$

and

$$\mathrm{Var}(\hat{f}(\omega_k)) = \mathrm{Var}[f(\omega_k)\frac{\chi^2(2)}{2}] = [f(\omega_k)]^2, \qquad (12.2$$

which do not depend on the sample size n. Thus, although the sample spec calculated at the Fourier frequencies is unbiased, it is an unsatisfactory mator since it is not consistent as the variance of $\hat{f}(\omega_k)$ does not tend to ze the sample size n tends to infinity. Furthermore, by (12.2.6) and the fact the periodogram ordinates are independent, as shown in (12.1.6), we hav

$$\mathrm{Cov}[\hat{f}(\omega_k), \hat{f}(\omega_j)] = 0 \qquad (12.2$$

for any two different Fourier frequencies ω_k and ω_j. More generally, ever other distinct non-Fourier frequencies ω and λ, as the sample size n incre and the mesh of the Fourier frequencies becomes finer and finer, the co ance between $\hat{f}(\omega)$ and $\hat{f}(\lambda)$ still tends to zero as $n \to \infty$. As a result, $\hat{f}($ unstable and tends to be very jagged, as illustrated in Figure 12.2. Note the pattern persists regardless of the size of the sample. The effort of cor ing these unsatisfactory properties leads to the smoothing of the periodo and the sample spectrum.

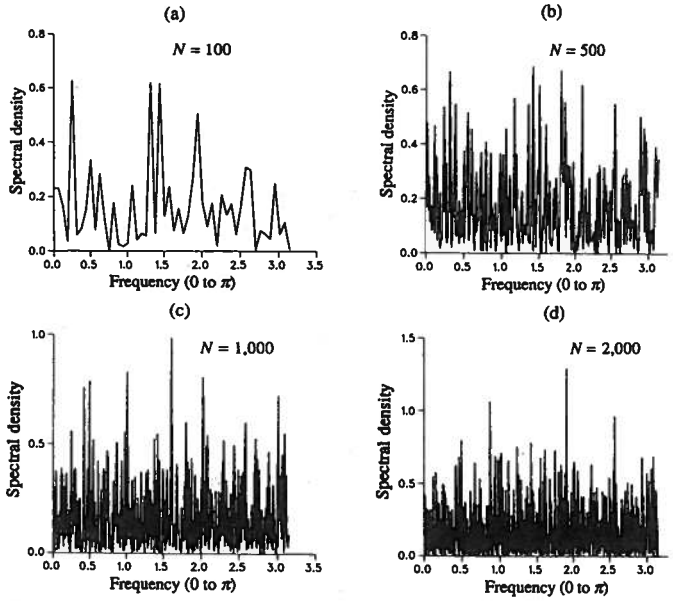

Fig. 12.2 Sample spectrum of a white noise process.

12.3 THE SMOOTHED SPECTRUM

12.3.1 Smoothing in the Frequency Domain—The Spectral Window

A natural way to reduce the variance of the sample spectrum is to smooth the sample spectrum locally in the neighborhood of the target frequency. In other words, the spectral estimator is the smoothed spectrum obtained from the following weighted average of m values to the right and left of a target frequency ω_k, i.e.,

$$\hat{f}_w(\omega_k) = \sum_{j=-m_n}^{m_n} W_n(\omega_j)\hat{f}(\omega_k - \omega_j) \tag{12.3.1}$$

where the $\omega_k = 2\pi k/n$, n being the sample size, $k = 0, \pm 1, \ldots, \pm[n/2]$, are Fourier frequencies, and $W_n(\omega_j)$ is the weighting function having properties that

$$\sum_{j=-m_n}^{m_n} W_n(\omega_j) = 1, \tag{12.3.2}$$

$$W_n(\omega_j) = W_n(-\omega_j), \tag{12.3.3}$$

and

$$\lim_{n\to\infty} \sum_{j=-m_n}^{m_n} W_n^2(\omega_j) = 0. \tag{12.3.4}$$

The weighting function $W_n(\omega_j)$ is called the spectral window because only some of the spectral ordinates are seen in the smoothing procedure. If $f(\omega)$ is flat and approximately constant within the window, we have

$$E[\hat{f}_w(\omega_k)] = \sum_{j=-m_n}^{m_n} W_n(\omega_j)E[\hat{f}(\omega_k - \omega_j)]$$

$$\simeq f(\omega_k) \sum_{j=-m_n}^{m_n} W_n(\omega_j) = f(\omega_k). \tag{12.3.5}$$

Also, by (12.2.10)

$$\mathrm{Var}[\hat{f}_w(\omega_k)] \simeq \sum_{j=-m_n}^{m_n} W_n^2(\omega_j)[f(\omega_k)]^2$$

$$\simeq [f(\omega_k)]^2 \sum_{j=-m_n}^{m_n} W_n^2(\omega_j), \tag{12.3.6}$$

where we also use the fact that the sample spectrum is independent at Fourier frequencies.

The property (12.3.4) of the spectral window implies that the variance the smoothed spectrum decreases as m_n increases. The value of m_n represe the number of frequencies used in the smoothing. This is directly related to width of the spectral window, also known as the bandwidth of the window. the bandwidth increases, more spectral ordinates are averaged, and hence resulting estimator becomes smoother, more stable, and has smaller varian However, unless $f(\omega)$ is really flat, the bias will also increase as the bandwi increases, because more and more spectral ordinates are used in the smoo ing. We are thus forced to compromise between variance reduction and bi a common dilemma with statistical estimators.

More generally, from (12.2.2a), the sample spectrum is defined at any f quency ω between $-\pi$ to π, and not only at the Fourier frequencies. Thus, can write the general smoothed spectrum in terms of the following integ form:

$$\hat{f}_W(\omega) = \int_{-\pi}^{\pi} W_n(\lambda)\hat{f}(\omega - \lambda)d\lambda, \tag{12.3}$$

$$= \int_{-\pi}^{\pi} W_n(\omega - \lambda)\hat{f}(\lambda)d\lambda, \tag{12.3}$$

where $W_n(\lambda)$ is the spectral window satisfying the conditions

$$\int_{-\pi}^{\pi} W_n(\lambda)d\lambda = 1, \tag{12.3}$$

$$W_n(\lambda) = W_n(-\lambda), \tag{12.3.}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \int_{-\pi}^{\pi} W_n^2(\lambda)d\lambda = 0. \tag{12.3.}$$

The spectral window is also known as the kernel in the literature.

If the spectrum is approximately constant within the bandwidth of spectral window, the condition (12.3.9) implies that

$$E(\hat{f}_w(\omega)) = f(\omega). \tag{12.3.}$$

For the variance, we note that (12.3.8) can be approximated by the sum

$$\hat{f}_w(\omega) \simeq \frac{2\pi}{n} \sum_{k=-[n/2]}^{[n/2]} W_n(\omega - \omega_k)\hat{f}(\omega_k) \tag{12.3.}$$

where $\omega_k = 2\pi k/n$. Thus,

$$\mathrm{Var}(\hat{f}_w(\omega)) \simeq \left(\frac{2\pi}{n}\right)^2 f^2(\omega) \sum_{k=-[n/2]}^{[n/2]} W_n^2(\omega - \omega_k)$$

$$\simeq \frac{2\pi}{n} f^2(\omega) \sum_{k=-[n/2]}^{[n/2]} \mathcal{W}_n^2(\omega - \omega_k) \frac{2\pi}{n}$$
$$\simeq \frac{2\pi}{n} f^2(\omega) \int_{-\pi}^{\pi} \mathcal{W}_n^2(\lambda) d\lambda. \qquad (12.3.14)$$

By condition (12.3.11), $\mathrm{Var}(\hat{f}_w(\omega)) \to 0$ as $n \to \infty$, and hence $\hat{f}_w(\omega)$ is a consistent estimator of $f(\omega)$.

Clearly, spectral ordinates at two different frequencies ω and λ are in general correlated because they may contain some common terms in the smoothing. However, because the sample spectral ordinates at different Fourier frequencies are independent, the covariance between smoothed spectral ordinates is easily seen to be proportional to the amount of overlap between the spectral windows centered at ω and λ. The covariance will be large if the spectral windows overlap considerably; the covariance will be small if the spectral windows overlap only slightly.

A few other remarks are in order. In the actual computation of the smoothed spectrum, the discrete form of (12.3.1) is used at the Fourier frequencies. Thus, the smoothing is essentially applied to the periodogram. Because the periodogram is periodic with period 2π, when the window covers frequencies that fail to lie entirely in the range between $-\pi$ and π, we can extend the periodogram and hence the sample spectrum using the periodic property. Equivalently, we can fold the weights back into the interval $-\pi$ to π. As the periodogram is also symmetric about frequency zero, calculation is only necessary for the frequency range between zero and π. Also, since the periodogram at frequency zero reflects the sample mean of Z_t and not the spectrum, it is not included in the smoothing, i.e., its weight is set to zero in the computation, and the value at ω_1 is used in its place.

12.3.2 Smoothing in the Time Domain—The Lag Window

Note that the spectrum $f(\omega)$ is the Fourier transform of the autocovariance function γ_k. Hence, from (12.2.2a), an alternative to spectrum smoothing is to apply a weighting function $W(k)$ to the sample autocovariances, i.e.,

$$\hat{f}_w(\omega) = \frac{1}{2\pi} \sum_{k=-(n-1)}^{n-1} W(k) \hat{\gamma}_k e^{-i\omega k}. \qquad (12.3.15)$$

Because the sample autocovariance function $\hat{\gamma}_k$ is symmetric, and $\hat{\gamma}_k$ is less reliable for larger k, the weighting function $W(k)$ should also be chosen to be symmetric, with its weights inversely proportional to the magnitude of k. Thus, we have

$$\hat{f}_w(\omega) = \frac{1}{2\pi} \sum_{k=-M}^{M} W_n(k) \hat{\gamma}_k e^{-i\omega k}, \qquad (12.3.16)$$

where the weighting function $W_n(k)$ is chosen to be an absolutely summab sequence

$$W_n(k) = W(k/M), \qquad (12.3.17$$

which is often derived from a bounded even continuous function $W(x)$ sati fying

$$\begin{aligned} &|W(x)| \le 1, \\ &W(0) = 1, \\ &W(x) = W(-x), \\ &W(x) = 0, \qquad |x| > 1. \end{aligned} \qquad (12.3.18$$

The value of M is the truncation point that depends on the sample size n. Th weighting function $W_n(k)$ for the autocovariances is called the lag window.

The lag window and the spectral window are closely related. This follow from the fact that the autocovariance is the inverse Fourier transform of th spectrum. From (12.2.2a), the inverse Fourier transform of $\hat{f}(\lambda)$ is

$$\hat{\gamma}_k = \int_{-\pi}^{\pi} \hat{f}(\lambda) e^{i\lambda k} d\lambda \qquad \text{for } k = 0, \pm 1, \ldots, \pm(n-1). \qquad (12.3.19$$

Thus,

$$\begin{aligned} \hat{f}_{\mathcal{W}}(\omega) &= \frac{1}{2\pi} \sum_{k=-M}^{M} W_n(k) \hat{\gamma}_k e^{-i\omega k} \\ &= \frac{1}{2\pi} \sum_{k=-M}^{M} W_n(k) \int_{-\pi}^{\pi} \hat{f}(\lambda) e^{i\lambda k} e^{-i\omega k} d\lambda \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{k=-M}^{M} W_n(k) e^{-i(\omega-\lambda)k} \hat{f}(\lambda) d\lambda \\ &= \int_{-\pi}^{\pi} \mathcal{W}_n(\omega - \lambda) \hat{f}(\lambda) d\lambda \\ &= \int_{-\pi}^{\pi} \mathcal{W}_n(\lambda) \hat{f}(\omega - \lambda) d\lambda \end{aligned} \qquad (12.3.20$$

where

$$\mathcal{W}_n(\omega) = \frac{1}{2\pi} \sum_{k=-M}^{M} W_n(k) e^{-i\omega k} \qquad (12.3.21$$

is the spectral window. It is clear from (12.3.21) that the spectral window is th Fourier transform of the lag window, and that the lag window is the invers

Fourier transform of the spectral window, i.e.,

$$W_n(k) = \int_{-\pi}^{\pi} \mathcal{W}_n(\omega) e^{i\omega k} d\omega, \qquad k = 0, \pm 1, \ldots, \pm M. \tag{12.3.22}$$

Hence, the lag window and the spectral window form a Fourier transform pair, with one being uniquely determined by the other. Both the terms *lag window* and *spectral window* were introduced by Blackman and Tukey (1958). The *weighting function* was the standard term used in the earlier literature.

The variance expression given in (12.3.14) can also be written in terms of a lag window. By Parseval's relation

$$\sum_{k=-M}^{M} W_n^2(k) = 2\pi \int_{-\pi}^{\pi} \mathcal{W}_n^2(\omega) d\omega, \tag{12.3.23}$$

and hence

$$\mathrm{Var}[\hat{f}_w(\omega)] \simeq \frac{1}{n} f^2(\omega) \sum_{k=-M}^{M} W_n^2(k). \tag{12.3.24}$$

However, since the lag window is often derived from a bounded even continuous function as given in (12.3.18), we have

$$\begin{aligned} \sum_{k=-M}^{M} W_n^2(k) &= \sum_{k=-M}^{M} W^2(k/M) \\ &= M \sum_{k=-M}^{M} W^2(k/M) \cdot \frac{1}{M} \\ &\simeq M \int_{-1}^{1} W^2(x) dx. \end{aligned} \tag{12.3.25}$$

Thus,

$$\mathrm{Var}[\hat{f}_w(\omega)] \simeq \frac{M}{n} f^2(\omega) \int_{-1}^{1} W^2(x) dx. \tag{12.3.26}$$

This implies that the variance of the smoothed spectrum for a given lag window is related to the ratio M/n, which is the proportion of sample autocovariances for which the lag weights are nonzero.

12.3.3 Some Commonly Used Windows

With the above general background, we now introduce some lag and spectral windows commonly used in time series analysis.

The Rectangular Window The so-called rectangular or truncated l window

$$W_n^R(k) = \begin{cases} 1, & |k| \le M, \\ 0, & |k| > M, \end{cases} \tag{12.3.2}$$

where M is the truncation point less than $(n-1)$, is derived from the followi continuous function:

$$W(x) = \begin{cases} 1, & |x| \le 1, \\ 0, & |x| > 1. \end{cases} \tag{12.3.2}$$

From (12.3.21), the corresponding spectral window is given by

$$\begin{aligned} \mathcal{W}_n^R(\omega) &= \frac{1}{2\pi} \sum_{k=-M}^{M} W_n^R(k) e^{-i\omega k} \\ &= \frac{1}{2\pi} \sum_{k=-M}^{M} e^{-i\omega k} \\ &= \frac{1}{2\pi} (1 + 2 \sum_{k=1}^{M} \cos \omega k) \\ &= \frac{1}{2\pi} \left[1 + \frac{2 \cos(\omega(M+1)/2) \sin(\omega M/2)}{\sin(\omega/2)} \right], \qquad \text{(using (10.2.8)),} \\ &= \frac{1}{2\pi} \frac{\sin(\omega/2) + [\sin(\omega(M+1/2)) - \sin(\omega/2)]}{\sin(\omega/2)}, \qquad \text{(using (10.2.12)} \\ &= \frac{1}{2\pi} \frac{\sin(\omega(M+1/2))}{\sin(\omega/2)}. \end{aligned} \tag{12.3.2}$$

The rectangular lag and spectral windows are both shown in Figure 12.3.

Note that the spectral window as shown in Figure 12.3 has a main lobe $\omega = 0$ with a height of $(2M+1)/2\pi$, zeros at $\omega = \pm 2j\pi/(2M+1)$, and side pea (lobes) of decreasing magnitude approximately at $\omega = \pm(4j+1)\pi/(2M+1)$ f $j = 1, 2, \ldots$. As a result, the smoothed spectrum estimator using this windo may lead to negative values for some frequencies ω. Because the spectrum known to be a nonnegative function, this result is certainly not desirable.

For a given spectral window $\mathcal{W}_n(\omega)$ that attains its maximum at $\omega = 0$, o commonly used definition for the bandwidth of the window is the distance b tween the half-power points on the main lobe, as shown in Figure 12.4. That

$$\text{Bandwidth} = 2\omega_I, \tag{12.3.3}$$

where ω_I is such that $\mathcal{W}_n(\pm\omega_I) = \frac{1}{2}\mathcal{W}_n(0)$.

If ω_{II} is the first zero of $\mathcal{W}_n(\omega)$, the value of ω_I can be approximated by $\frac{1}{2}$ and hence the bandwidth is approximately equal to ω_{II}. In terms of the abo

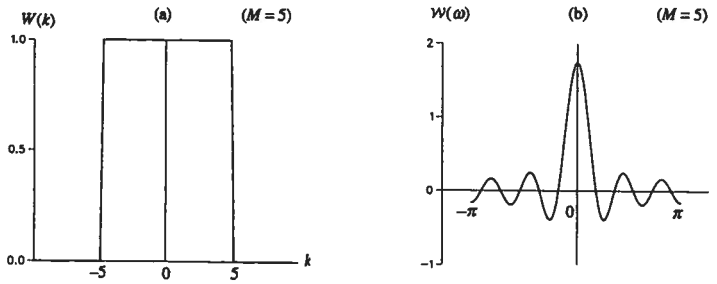

Fig. 12.3 The rectangular lag and spectral windows. (a) Rectangular lag window. (b) Rectangular spectral window.

rectangular spectral window $W_n^R(\omega)$, the bandwidth is approximately equal to $2\pi/(2M+1)$. In general, the bandwidth of the spectral window is inversely related to the truncation point M used in the lag window. Thus, as M increases, the bandwidth decreases, and hence the variance of the smoothed spectrum increases although the bias decreases, as discussed in Section 12.3.1. On the other hand, as M decreases, the bandwidth increases, and the variance decreases while the bias increases.

Bartlett's Window Bartlett (1950) proposed the following lag window:

$$W_n^B(k) = \begin{cases} 1-|k|/M, & |k| \le M, \\ 0, & |k| > M, \end{cases} \tag{12.3.31}$$

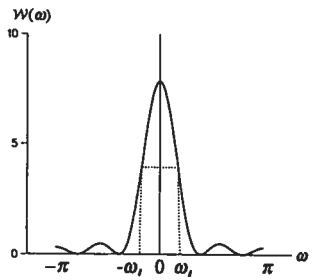

Fig. 12.4 Bandwidth of the spectral window.

based on the triangular function

$$W(x) = \begin{cases} 1-|x|, & |x| \le 1, \\ 0, & |x| > 1. \end{cases} \tag{12.3}$$

Hence, the window is also known as the triangular window. The correspon spectrum window is given by

$$\begin{aligned}
W_n^B(\omega) &= \frac{1}{2\pi}\sum_{k=-M}^{M}\left(1-\frac{|k|}{M}\right)e^{-i\omega k} \\
&= \frac{1}{2\pi M}\sum_{k=-M}^{M}(M-|k|)e^{-i\omega k} \\
&= \frac{1}{2\pi M}\sum_{j=0}^{M-1}\sum_{k=-j}^{j}e^{-i\omega k} \\
&= \frac{1}{2\pi M}\sum_{j=0}^{M-1}\frac{\sin(\omega(j+1/2))}{\sin(\omega/2)} \\
&= \frac{1}{2\pi M\sin(\omega/2)}\left[\sin(\omega/2)+\sum_{j=1}^{M-1}\sin(\omega(j+1/2))\right] \\
&= \frac{1}{2\pi M\sin(\omega/2)}\left\{\sin(\omega/2)+\sum_{j=1}^{M-1}[\sin(\omega j)\cos(\omega/2)+\cos(\omega j)\sin(\omega,\right. \\
&= \frac{1}{2\pi M\sin(\omega/2)}\left\{\sin(\omega/2)+\cos(\omega/2)\sum_{j=1}^{M-1}\sin(\omega j)\right. \\
&\quad \left.+\sin(\omega/2)\sum_{j=1}^{M-1}\cos(\omega j)\right\} \\
&= \frac{1}{2\pi M\sin(\omega/2)}\left\{\sin(\omega/2)+\frac{\cos(\omega/2)\sin(\omega M/2)\sin(\omega(M-1)/2)}{\sin(\omega/2)}\right. \\
&\quad \left.+\frac{\sin(\omega/2)\cos(\omega M/2)\sin(\omega(M-1)/2)}{\sin(\omega/2)}\right\}, \qquad \text{by (10.2.9) and (1} \\
&= \frac{1}{2\pi M\sin(\omega/2)}\left\{\sin(\omega/2)+\frac{\sin(\omega(M-1)/2)}{\sin(\omega/2)}[\cos(\omega/2)\sin(\omega M/2)\right. \\
&\quad \left.+\sin(\omega/2)\cos(\omega M/2)]\right\} \\
&= \frac{1}{2\pi M\sin(\omega/2)}\left\{\sin(\omega/2)+\frac{\sin(\omega(M-1)/2)}{\sin(\omega/2)}\sin(\omega(M+1)/2)\right\}
\end{aligned}$$

$$= \frac{1}{2\pi M[\sin(\omega/2)]^2}\{[\sin(\omega/2)]^2 + \sin(\omega(M-1)/2)\sin(\omega(M+1)/2)\}$$

$$= \frac{1}{2\pi M[\sin(\omega/2)]^2}\left\{\frac{1}{2}(1-\cos\omega) + \frac{1}{2}(\cos\omega - \cos\omega M)\right\}, \quad \text{by (10.2.12),}$$

$$= \frac{1}{2\pi M[\sin(\omega/2)]^2}\left\{\frac{1}{2}(1-\cos\omega M)\right\}$$

$$= \frac{1}{2\pi M}\left[\frac{\sin(\omega M/2)}{\sin(\omega/2)}\right]^2. \tag{12.3.33}$$

Bartlett's triangular lag window and its corresponding spectral window are shown in Figure 12.5. Since the spectral window $W_n^B(\omega)$ is nonnegative, the Bartlett's spectrum estimator is always nonnegative. Furthermore, a direct comparison of Equations (12.3.29) and (12.3.33) shows that the side lobes of the Bartlett window are smaller than the side lobes of the rectangular window. The effect of large side lobes is to allow $\hat{f}(\lambda)$ to make large contributions in the smoothing at frequencies distant from ω. Hence, the resulting spectrum estimate $\hat{f}_w(\omega)$ may reflect significant spectral components at other frequencies different from ω. This phenomenon is known as leakage. Because lobes are produced by the Fourier transform of sharp corners, the general principle of the selection of lag windows is to avoid functions with sharp corners.

The Blackman–Tukey Window Blackman and Tukey (1958) suggested the following lag window:

$$W_n^T(k) = \begin{cases} 1 - 2a + 2a\cos(\pi k/M), & |k| \le M, \\ 0, & |k| > M, \end{cases} \tag{12.3.34}$$

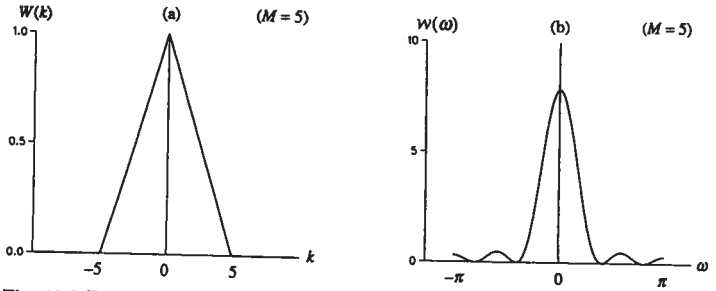

Fig. 12.5 Bartlett's lag and spectral windows. (a) Bartlett's triangular lag window. (b) Bartlett's spectral window.

based on the continuous weighting function,

$$W(x) = \begin{cases} 1 - 2a + 2a\cos\pi x, & |x| \le 1, \\ 0, & |x| > 1. \end{cases} \tag{12.3.35}$$

Again, M is the truncation point for the sample autocovariance function, and the constant a is chosen in the range $0 < a \le .25$ so that $W_n^T(k) \ge 0$ for all k. The corresponding spectral window can be derived as follows:

$$\begin{aligned} W_n^T(\omega) &= \frac{1}{2\pi}\sum_{k=-M}^{M} W_n^T(k)e^{-i\omega k} \\ &= \frac{1}{2\pi}\sum_{k=-M}^{M}[1 - 2a + 2a\cos(\pi k/M)]e^{-i\omega k} \\ &= \frac{1}{2\pi}\sum_{k=-M}^{M}[1 - 2a + a(e^{i\pi k/M} + e^{-i\pi k/M})]e^{-i\omega k} \\ &= a\frac{1}{2\pi}\sum_{k=-M}^{M} e^{-i(\omega-\pi/M)k} + (1-2a)\frac{1}{2\pi}\sum_{k=-M}^{M} e^{-i\omega k} \\ &\quad + a\frac{1}{2\pi}\sum_{k=-M}^{M} e^{-i(\omega+\pi/M)k} \\ &= \frac{a}{2\pi}\frac{\sin[(\omega-\pi/M)(M+1/2)]}{\sin[(\omega-\pi/M)/2]} \\ &\quad + \frac{(1-2a)}{2\pi}\frac{\sin[\omega(M+1/2)]}{\sin(\omega/2)} + \frac{a}{2\pi}\frac{\sin[(\omega+\pi/M)(M+1/2)]}{\sin[(\omega+\pi/M)/2]} \end{aligned}$$

where we have used (12.3.29). Thus, the Blackman–Tukey spectral window is a weighted linear combination of the spectral windows for the rectangular lag function at the frequencies $(\omega - \pi/M)$, ω, and $(\omega + \pi/M)$, i.e.,

$$W_n^T(\omega) = aW_n^R(\omega - \pi/M) + (1-2a)W_n^R(\omega) + aW_n^R(\omega + \pi/M) \tag{12.3.36}$$

where $W_n^R(\omega)$ is given by (12.3.29). As a result, the Blackman–Tukey spectrum estimator may also be negative at some frequencies ω.

Blackman and Tukey (1958) called the window in (12.3.34) with $a = .23$ as *hamming* after R.W. Hamming, one of Tukey's coauthors who had worked on the subject. The resulting window is

$$W_n(k) = \begin{cases} .54 + .46\cos(\pi k/M), & |k| \le M, \\ 0, & |k| > M, \end{cases} \tag{12.3.37}$$

which is also known as the Tukey–Hamming window in the literature. The window (12.3.34) with $a = .25$ is called as *hanning* after the Austrian meteorologist Julius von Hann, who was not directly related with the subject. The window is

given by

$$W_n(k) = \begin{cases} 1/2[1+\cos(\pi k/M)], & |k| \le M, \\ 0, & |k| > M, \end{cases} \tag{12.3.38}$$

which is also known as the Tukey–Hanning window or the Tukey window in the literature.

The Parzen Window Parzen (1961a) suggested the lag window

$$W_n^P(k) = \begin{cases} 1-6(k/M)^2+6(|k|/M)^3, & |k| \le M/2, \\ 2(1-|k|/M)^3, & M/2 < |k| \le M, \\ 0, & |k| > M, \end{cases} \tag{12.3.39}$$

which is based on the following continuous weighting function:

$$W(x) = \begin{cases} 1-6x^2+6|x|^3, & |x| \le 1/2, \\ 2(1-|x|)^3, & 1/2 < |x| \le 1, \\ 0, & |x| > 1. \end{cases} \tag{12.3.40}$$

The corresponding spectral window for an even value of M is given by

$$\begin{aligned} \mathcal{W}_n^P(\omega) &= \frac{1}{2\pi}\sum_{k=-M}^{M} W_n^P(k)\cos\omega k \\ &= \frac{1}{2\pi}\left\{\sum_{k=-M/2}^{M/2}[1-6(k/M)^2+6(|k|/M)^3]\cos\omega k \right. \\ &\quad \left. + 2\sum_{M/2<|k|\le M}(1-|k|/M)^3\cos\omega k\right\} \\ &= \frac{3}{8\pi M^3}\left[\frac{\sin(\omega M/4)}{1/2\sin(\omega/2)}\right]^4\{1-2/3[\sin(\omega/2)]^2\}, \end{aligned} \tag{12.3.41}$$

where we refer readers to a more detailed derivation in Parzen (1963). When M is large, (12.3.41) is approximated by

$$\mathcal{W}_n^P(\omega) \simeq \frac{3}{8\pi M^3}\left[\frac{\sin(\omega M/4)}{1/2\sin(\omega/2)}\right]^4. \tag{12.3.42}$$

Both the Tukey–Hanning window and the Parzen window are plotted in Figure 12.6. As can be seen from Figure 12.6(b) and (d), for the same truncation point M, the bandwidth of the Tukey–Hanning spectral window is narrower than the bandwidth of the Parzen window. Hence, the Tukey–Hanning estimator will have a smaller bias than the Parzen estimator. However, because $\mathcal{W}_n^P(\omega)$ is always nonnegative, the Parzen window will produce a nonnegative and smoother spectrum estimator.

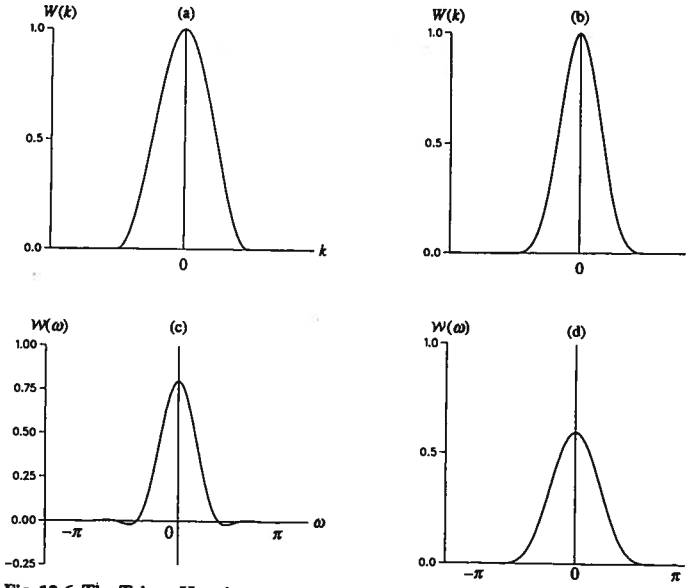

Fig. 12.6 The Tukey–Hanning and the Parzen windows. (a) Tukey–Hanning lag wind ($M = 5$); (b) Parzen lag window ($M = 5$); (c) Tukey–Hanning spectral window ($M =$ (d) Parzen spectral window ($M = 5$).

The above windows are some commonly used ones, particularly in ter of the commercially available software. There are many other windows int duced in the literature. Interested readers are referred to an excellent book Priestley (1981, Chapter 6) for more details.

The quality of a smoothed spectrum is determined by the shape of the w dow (i.e., the form of the weighting function) and the bandwidth of the w dow (or equivalently, the truncation point). The spectrum estimates produc for the same window shape and different bandwidths vary. Thus, in spectr smoothing, we are concerned not only about the design of spectral windo with desirable shapes (or "window carpentry" as John W. Tukey called it), t also about the bandwidth of the windows. The latter concern often is a mc crucial and difficult problem in time series analysis because for a given wind shape there is no single criterion for choosing the optimal bandwidth. Ad tionally, wider bandwidths produce smoother spectrums with smaller varian of the estimators; on the other hand, narrower bandwidths lead to smaller b and smudging and hence to higher resolution of the resulting spectrum es mates. A compromise is needed between high stability and high resolution.

ease the difficulty, the following steps are often suggested. First, choose a spectral window with an acceptable and desirable shape. Then initially calculate the spectral estimates using a large bandwidth, and then recalculate the estimates by using gradually smaller bandwidths until the required stability and resolution are achieved. This procedure is often referred to as "window closing."

Because the bandwidth of a spectral window is inversely related to the truncation point M used in the lag window, Jenkins and Watts (1968, Chapter 7) suggest choosing three values of M, say, M_1, M_2, and M_3 with $M_3 = 4M_1$. The sequences of spectral estimates are then computed and examined with these M values to see whether an intermediate value between M_1 and M_3, or a value smaller than M_1, or a value larger than M_3 should be chosen. The spectral bandwidth can also be determined by choosing a truncation point M such that $\hat{\gamma}(k)$ for $k > M$ are negligible. This choice of M has been found practical and was used successfully in a number of cases. Alternatively, a truncation point M can also be chosen as a function of n, the number of observations in a series. For a moderate size n, M can be chosen to be $M = [\frac{n}{10}]$.

In principle, if we wish to be able to distinguish several peaks of $f(\omega)$ at frequencies $\lambda_1, \lambda_2, \lambda_3, \ldots$, the bandwidth of the spectral window should not exceed the minimum interval between these adjacent peaks. Otherwise, the smoothed spectrum will blur these peaks. In fact, a commonly chosen bandwidth is equal to $\min_i |\lambda_{i+1} - \lambda_i|$.

Example 12.2 Figure 12.7 shows the sample spectrum of the lynx pelt sales data over the interval $(0, \pi)$. Also shown are the smoothed sample spectrum functions using the Bartlett window with truncation points $M = 2$, 5, and 10. The spectrum for $M = 10$ clearly is improper. To compare different estimates, Figure 12.8 shows the spectrum estimates using various windows with the same truncation point $M = 5$. The smoothed spectrum given by the rectangular window is clearly undesirable. The estimates using the Bartlett, Parzen, and Tukey–Hanning windows all lead to the periodic component at the frequency $\omega = .68544$, but the rectangular window does not lead to this periodic component.

12.3.4 Approximate Confidence Intervals for Spectral Ordinates

Recall that for a given sample $Z_1, Z_2, \ldots, Z_n$ from a process with the spectrum $f(\omega)$, the sample spectral ordinates $\hat{f}(\omega_k)$ at Fourier frequencies $\omega_k = 2\pi k/n$, with $\omega_k \neq 0$ and π, are independent and identically distributed as

$$\hat{f}(\omega_k) \sim f(\omega_k)\frac{\chi^2(2)}{2}. \tag{12.3.43}$$

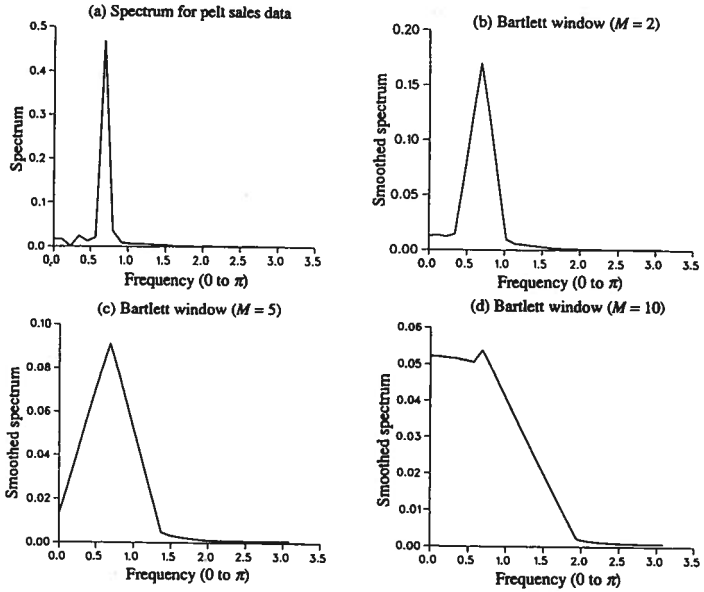

Fig. 12.7 Sample spectrum of the lynx pelt sales and Bartlett estimates for differen values of M.

Thus, if we smooth the sample spectrum with the simple $(2m+1)$ term movin average, i.e.,

$$\hat{f}_W(\omega_k) = \frac{1}{2m+1}\sum_{j=-m}^{m}\hat{f}(\omega_k - \omega_j), \tag{12.3.44}$$

the smoothed spectral ordinate $\hat{f}_W(\omega_k)$ will be distributed as

$$\hat{f}_W(\omega_k) \sim f(\omega_k)\frac{\chi^2(\mathrm{DF})}{\mathrm{DF}}, \tag{12.3.45}$$

where the degrees of freedom $\mathrm{DF} = (4m + 2)$ is simply the sum of the degrees of freedom of $(2m + 1)$ $\chi^2(2)$ random variables. In fact, this is the firs smoothing method introduced by Daniell (1946), one of the pioneers of spec trum estimation. The estimator in (12.3.44) is also known as Daniell's estima tor. However, this nice additive property no longer holds for $\hat{f}_W(\omega)$ if ω is not a Fourier frequency or if the sample spectrum is smoothed by a spectral windov with nonuniform weights. Thus, we can only approximate a general smoothe

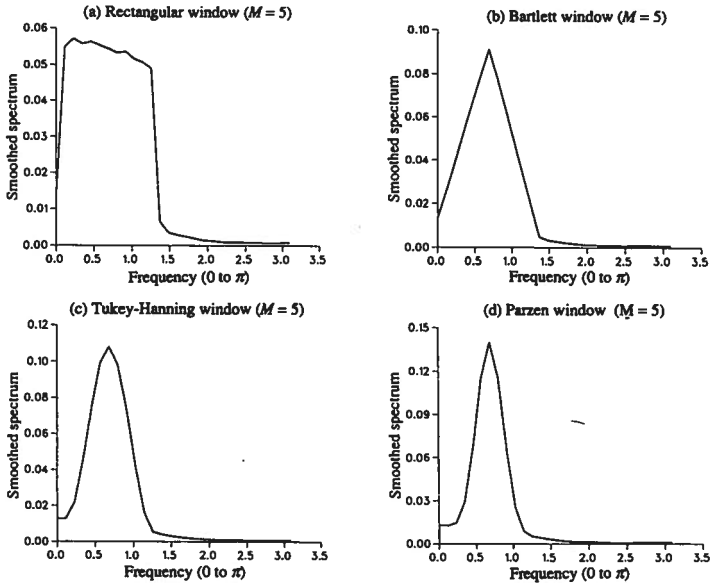

Fig. 12.8 Smoothed sample spectrum of the lynx pelt sales using different windows.

spectrum $\hat{f}_w(\omega)$ by

$$\hat{f}_W(\omega) \sim c\chi^2(\nu) \tag{12.3.46}$$

where c and ν are chosen such that

$$E[\hat{f}_W(\omega)] = E(c\chi^2(\nu)) = c\nu, \tag{12.3.47}$$

and

$$\text{Var}[\hat{f}_W(\omega)] = \text{Var}(c\chi^2(\nu)) = 2c^2\nu. \tag{12.3.48}$$

Now, for any given spectral window $W_n(\omega)$, from (12.3.12) and (12.3.14), we have

$$E[\hat{f}_w(\omega)] \simeq f(\omega)$$

and

$$\text{Var}[\hat{f}_w(\omega)] \simeq \frac{2\pi}{n}[f(\omega)]^2 \int_{-\pi}^{\pi} W_n^2(\omega)d\omega.$$

Hence,

$$c\nu = f(\omega),$$

and

$$2c^2\nu = \frac{2\pi}{n}f^2(\omega)\int_{-\pi}^{\pi} W_n^2(\omega)d\omega,$$

giving

$$\nu = \frac{n}{\pi\int_{-\pi}^{\pi} W_n^2(\omega)d\omega}, \tag{12.3.}$$

$$c = \frac{f(\omega)}{\nu}. \tag{12.3.}$$

It follows that

$$\hat{f}_W(\omega) \sim f(\omega)\frac{\chi^2(\nu)}{\nu} \tag{12.3.}$$

where ν is often referred to as equivalent degrees of freedom for a smooth spectrum. Using (12.3.26) or (12.3.23) and (12.3.25), we can also write equivalent degrees of freedom in terms of the lag window

$$\nu = \frac{2n}{M\int_{-1}^{1} W^2(x)dx}, \tag{12.3.}$$

where $W(x)$ is the continuous weighting function used in the associated window. The equivalent degrees of freedom ν for some commonly used windows are given in Table 12.4.

By (12.3.51), we have

$$P\left\{f(\omega)\frac{\chi^2_{1-\alpha/2}(\nu)}{\nu} \le \hat{f}_w(\omega) \le f(\omega)\frac{\chi^2_{\alpha/2}(\nu)}{\nu}\right\} = (1-\alpha),$$

where $\chi^2_\alpha(\nu)$ is the upper α% point of the chi-square distribution with ν degrees of freedom. Hence, a $(1-\alpha)100\%$ confidence interval for $f(\omega)$ is given

$$\frac{\nu\hat{f}_w(\omega)}{\chi^2_{\alpha/2}(\nu)} \le f(\omega) \le \frac{\nu\hat{f}_w(\omega)}{\chi^2_{1-\alpha/2}(\nu)}, \tag{12.3.5}$$

where ν is the equivalent degrees of freedom calculated by (12.3.49) (12.3.52).

Table 12.4 Equivalent degrees of freedom for various windows.

Window	ν
Rectangular	n/M
Bartlett	$3n/M$
Tukey–Hamming	$2.516n/M$
Tukey–Hanning	$8n/3M$
Parzen	$3.709n/M$

Recall that the asymptotic mean and variance of $\hat{f}_w(\omega)$ are proportional to $f(\omega)$ and $f^2(\omega)$, respectively. Hence, from the discussion in Section 4.3.2, the logarithmic transformation of the spectrum estimator $\ln\hat{f}_w(\omega)$ is often suggested. From (12.3.53), the $(1-\alpha)100\%$ confidence interval for $\ln f(\omega)$ is given by

$$\ln\hat{f}_w(\omega) + \ln\left[\frac{\nu}{\chi^2_{\alpha/2}(\nu)}\right] \le \ln f(\omega) \le \ln\hat{f}_w(\omega) + \ln\left[\frac{\nu}{\chi^2_{1-\alpha}(\nu)}\right]. \quad (12.3.54)$$

Note that the width of the interval for $f(\omega)$ as given in (12.3.53) is proportional to $\hat{f}_w(\omega)$ and hence varies with frequency. However, the width of the confidence interval for $\ln f(\omega)$ as given in (12.3.54) will be the same for any frequency ω.

Example 12.3 For illustration, we calculate the 95% confidence interval for the spectrum of the Canadian lynx pelt sales discussed in earlier examples using Parzen's window with $M = 5$. For $n = 55$ and $M = 5$, we have, from Table 12.4, $\nu = 3.709(55/5) = 40.79 \simeq 41$. Since $\chi^2_{.975}(41) = 25.22$ and $\chi^2_{.025}(41) = 60.56$, the 95% confidence interval for $f(\omega)$, from (12.3.53), becomes

$$.67\hat{f}_w(\omega) \le f(\omega) \le 1.62\hat{f}_w(\omega)$$

where $\hat{f}_w(\omega)$ is the estimated spectrum using Parzen's window with $M = 5$ given in Figure 12.8. Note that the maximum of $\hat{f}_w(\omega) = .14$ occurs at $\omega = .66$. Thus, the 95% confidence interval for $f(\omega)$ at $\omega = .66$ is given by

$$.094 \le f(\omega = .66) \le .227.$$

The confidence intervals for $f(\omega)$ at other frequencies ω can be calculated similarly and are shown in Figure 12.9.

12.4 ARMA SPECTRAL ESTIMATION

Based on a given time series $Z_1, Z_2, \ldots, Z_n$, we can approximate the unknown underlying process by an AR(p) model,

$$(1 - \phi_1 B - \cdots - \phi_p B^p)\dot{Z}_t = a_t \quad (12.4.1)$$

for some p. Let $\hat{\phi}_1, \ldots, \hat{\phi}_p$ and $\hat{\sigma}_a^2$ be the estimates of $\phi_1, \ldots, \phi_p$, and σ_a^2. A reasonable alternative to spectrum estimation is to substitute these parameter estimates in the theoretical expression for the spectrum of the AR(p) model discussed in Section 11.2.2, i.e.,

$$\hat{f}_A(\omega) = \frac{\hat{\sigma}_a^2}{2\pi}\frac{1}{\hat{\phi}_p(e^{-i\omega})\hat{\phi}_p(e^{i\omega})} \quad (12.4.2)$$

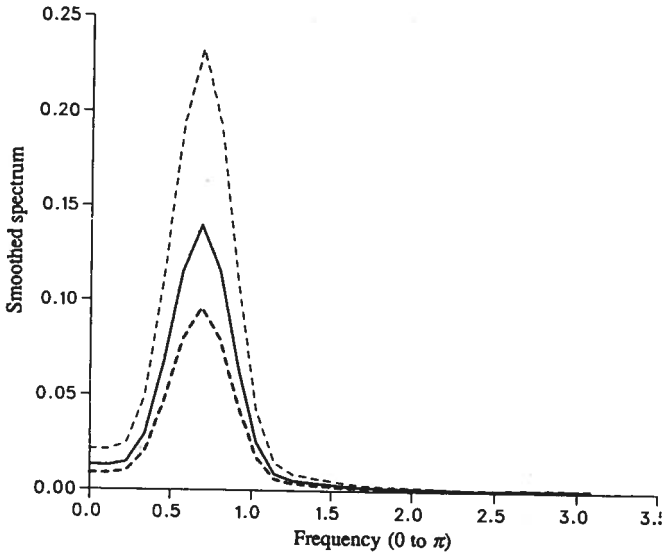

Fig. 12.9 The 95% confidence intervals for the spectrum of the logarithms of the dian lynx pelt sales using Parzen's window.

where $\hat{\phi}_p(e^{-i\omega}) = (1 - \hat{\phi}_1 e^{-i\omega} - \cdots - \hat{\phi}_p e^{-ip\omega})$. This method of spectrum timation through an AR approximation was suggested by Akaike (1969) Parzen (1974) and is often known as autoregressive spectral estimation.

For large n, Parzen (1974) showed that

$$\mathrm{Var}[\hat{f}_A(\omega)] \simeq \frac{2pf^2(\omega)}{n}. \quad (12$$

Thus, to control the variance, the order of p chosen to approximate the pr should not be too large. On the other hand, p should not be too small, an inadequate order of p leads to a poor approximation and hence ma crease the bias of the estimated spectrum. Consequently, similar to the cl of bandwidth, truncation point, and different windows, the determinati the order p in the autoregressive approximation becomes a crucial step i AR spectral estimation. Parzen (1977) suggested the use of CAT criteri choosing the optimal order p. If the AR order p is correctly determined, as totic simultaneous confidence bands for autoregressive spectra can be structed using the method by Newton and Pagano (1984).