Spectral Distribution and Density Functions

- We started with the basic model $X_t = R\cos(\omega t) + \epsilon_t$
  where $\omega$ is the 'dominant' frequency; $f = \omega/2\pi$ is the number of cycles per unit of time and $\lambda = 2\pi/\omega$ is the 'dominant' wavelength or period.

- This model can be generalized to

$$X_t = \sum_{j=1}^{k} R_j \cos(\omega_j t + \phi_j) + \epsilon_t$$

  which considers the existence of $k$-relevant frequencies $\omega_1, \omega_2, \ldots, \omega_k$.

- Given the trigonometric identity
\[
\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y), \text{ we have that }
\]
\[
X_t = \sum_{j=1}^{k} (a_j \cos(\omega_j t) + b_j \sin(\omega_j t)) + \epsilon_t
\]

with \(a_j = R_j \cos(\phi_j)\) and \(b_j = -R_j \sin(\phi_j)\).

- By making \(k \rightarrow \infty\), it can be shown that

\[
X_t = \int_0^{\pi} \cos(\omega t)du(\omega) + \int_0^{\pi} \sin(\omega t)dv(\omega)
\]

where \(u(\omega)\) and \(v(\omega)\) are continuous stochastic processes. This is the spectral representation of \(X_t\).

- The Wiener-Khintchine Theorem says that if \(\gamma(k)\) is the autocovariance function of \(X_t\), there must exist a
monotonically increasing function $F(\omega)$ such that

$$
\gamma(k) = \int_0^\pi \cos(\omega k) dF(\omega)
$$

- The function $F(\omega)$ is the spectral distribution function of the process $X_t$.
- Notice that for $k = 0$,

$$
\gamma(0) = \int_0^\pi dF(\omega) = F(\pi) = \sigma_x^2
$$

so all other variation in the process is for $0 < \omega < \pi$.
- We can redefine the spectral distribution function as:

$$
F^*(\omega) = \frac{F(\omega)}{\sigma_x^2}
$$

and so $F^*(\omega)$ is the proportion of variance accounted by $\omega$. 
• Also notice that $F^*(0) = 0$, $F^*(\pi) = 1$ and since $F(\omega)$ is monotonically increasing then $F^*(\omega)$ is a cumulative distribution function (CDF).

• The *Spectral Density function* is denoted by $f(\omega)$ and defined as

\[
  f(\omega) = \frac{dF(\omega)}{d\omega}; \quad 0 < \omega < \pi
\]

• This function is also known as the *power spectral function* or *spectrum*

• The existence of $f(\omega)$ is under the assumption that the spectral distribution function is differentiable everywhere (except in a set of measure zero).

• This spectral density gives us an alternative
representation for the covariance function

\[ \gamma(k) = \int_{0}^{\pi} \cos(\omega k) f(\omega) d\omega \]

This characterization is also known as Wold’s Theorem.

- If the spectrum has a ’peak’ at \( \omega_0 \), this implies that \( \omega_0 \) is an important frequency of the process \( X_t \).
- The spectrum or spectral density is a theoretical function of the process \( X_t \). In practice, the spectrum is usually unknown and we use the periodogram to estimate it.
- There is an inverse relationship between the \( f(\omega) \) and \( \gamma(k) \),

\[ f(\omega) = \frac{1}{\pi} \sum_{k=-\infty}^{\infty} \gamma(k) e^{-i\omega k} \]
so the spectrum is the Fourier transformation of the autocovariance function.

- From complex analysis, recall that
  
  \[ e^{-i\omega k} = \cos(\omega k) - i \sin(\omega k) \]

- This implies that
  
  \[ f(\omega) = \frac{1}{\pi} \left[ \gamma(0) + 2 \sum_{k=1}^{\infty} \gamma(k) \cos(\omega k) \right] \]

- The normalized spectral density \( f^*(\omega) \) is defined as:
  
  \[ f^*(\omega) = \frac{f(\omega)}{\sigma_x^2} = \frac{dF^*(\omega)}{d\omega} \]
• Then,

\[
f^*(\omega) = \frac{1}{\pi} \left[ 1 + 2 \sum_{k=1}^{\infty} \rho(k) \cos(\omega k) \right]
\]

so the normalized spectrum is the Fourier transform of the autocorrelation function (ACF).

• **Example 1:** White noise process. Suppose that \(X_t\) is a purely random process where \(E(X_t) = 0\) and \(Var(X_t) = \sigma^2\). The autocovariance function is \(\gamma(0) = \sigma^2\) and \(\gamma(k) = 0; k \neq 0\). Thus, the spectral density function is given by

\[
f(\omega) = \frac{\sigma^2}{\pi}
\]

• **Example 2:** Consider a first order autoregressive (AR) process

\[
X_t = \alpha X_{t-1} + \epsilon_t; \epsilon_t \sim N(0, \sigma^2)
\]
The autocovariance function of this process is given by

\[ \gamma(k) = \frac{\sigma^2 \alpha^{|k|}}{(1 - \alpha^2)} = \sigma_x^2 \alpha^{|k|}; \quad k = 0, \pm 1, \pm 2, \ldots \]

Then, the spectral density function is given by

\[ f(\omega) = \frac{\sigma^2}{\pi} \left( 1 + \sum_{k=1}^{\infty} \alpha^k e^{-i k \omega} + \sum_{k=1}^{\infty} \alpha e^{i k \omega} \right) \]

after some algebra, this gives

\[ f(\omega) = \sigma_z^2 / [\pi (1 - 2\alpha \cos(\omega) + \alpha^2)] \]

- **Example 3:** Define the sequence \( X_t \) by

\[ X_t = A \cos(\theta t) + B \sin(\theta t) + \epsilon_t \]

where \( \epsilon_t \) is white noise sequence with variance \( \sigma^2 \), \( A \) and
$B$ are independent random variables with mean zero and variance $\tau^2$. It can be shown that

$$E(X_t) = 0; Var(X_t) = \tau^2 + \sigma^2$$

Also, for $t \neq s$

$$cov(X_t, X_s) = \tau^2 \cos\{\theta(t - s)\}$$

Then $X_t$ is a stationary series with autocovariance

$$\gamma(k) = \begin{cases} 
\sigma^2 + \tau^2, & k = 0 \\
\tau^2 \cos(k\theta), & k \neq 0 
\end{cases}$$

The spectrum can be evaluated as

$$f(\omega) = \sigma^2 + \tau^2 + 2\tau^2 \sum_{k=1}^{\infty} \cos(k\theta)\cos(k\omega)$$
\[ = \sigma^2 + \tau^2 + 2\tau^2 \sum_{k=1}^{\infty} \left[ \cos\{k(\theta + \omega)\} + \cos\{k(\theta - \omega)\} \right] \]

If \( \theta = \omega \), then \( \cos\{k(\theta - \omega)\} = 1 \) for all \( k \) and the summation is infinite.

This means that the spectrum has a ’spike’ at \( \omega = \theta \).

The spectrum can only exist if we allow \( f(\omega) = \infty \) at isolated values of \( \omega \).
Periodogram revisited

- For $0 < \omega < \pi$, the periodogram is defined as

$$I(\omega) = \frac{n}{2}(\hat{a}^2 + \hat{b}^2)$$

$$= \frac{2}{n} \left[ \left( \sum_{t=1}^{n} x_t \cos(\omega t) \right)^2 + \left( \sum_{t=1}^{n} x_t \sin(\omega t) \right)^2 \right]$$

- If $\omega = 2\pi j/n; j < n/2$ is a Fourier frequency and since $\sum_t \cos(\omega t) = \sum_t \sin(\omega t) = 0$ then

$$I(\omega) = \frac{2}{n} \left[ \left( \sum_{t=1}^{n} (x_t - \bar{x}) \cos(\omega t) \right)^2 + \left( \sum_{t=1}^{n} (x_t - \bar{x}) \sin(\omega t) \right)^2 \right]$$

- Expanding each square term and by the trigonometric
identities

\[ \left( \frac{n}{2} \right) I(\omega) = \sum_{t=1}^{n} (x_t - \bar{x})^2 + 2 \sum_{k=1}^{n-1} \sum_{t=k+1}^{n} (x_t - \bar{x})(x_{t-k} - \bar{x}) \cos(\omega k) \]

- This gives an alternative expression for the periodogram,

\[ I(\omega) = 2 \left( g_0 + 2 \sum_{k=1}^{n-1} g_k \cos(\omega k) \right) \]

- We also have a normalized periodogram

\[ I^*(\omega) = \frac{I(\omega)}{g_0} = 2 \left( 1 + 2 \sum_{k=1}^{n-1} \rho_k \cos(\omega k) \right); \quad \rho_k = g_k / g_0 \]

- The last two expressions justify the used of the periodogram as an estimate of the spectral density.

- What is the sampling distribution of \( I(\omega) \)?
• By definition, the periodogram satisfies the relation:

\[ nI(\omega) = A(\omega)^2 + B(\omega)^2 \]

where

\[ A(\omega) = \sum_{i=1}^{n} x_t \cos(\omega t); \quad B(\omega) = \sum_{i=1}^{n} x_t \sin(\omega t) \]

• To understand the sampling distribution of the periodogram, let's suppose \( x_t \) is a realization of a white noise process (i.i.d. \( X_t \sim N(0, \sigma^2) \)).

• What is the distribution of \( A(\omega) \) and \( B(\omega) \)?

• Linear combinations of normal variables are normal.

• In fact, \( E(A(\omega)) = E(B(\omega)) = 0 \)
• Additionally, by the trigonometric identities,

\[
\text{Var}(A(\omega)) = \sigma^2 \sum_{t=1}^{n} \cos^2(\omega t) = (n\sigma^2)/2
\]

\[
\text{Var}(B(\omega)) = \sigma^2 \sum_{t=1}^{n} \sin^2(\omega t) = (n\sigma^2)/2
\]

• Also,

\[
\text{Cov}(A(\omega), B(\omega)) = E \left[ \sum_{t=1}^{n} \sum_{s=1}^{n} X_t X_s \cos(\omega t) \sin(\omega s) \right]
\]

\[
= \sigma^2 \sum_{t=1}^{n} \cos(\omega t) \sin(\omega t) = 0
\]

• It follows that \( A(\omega) \sqrt{2/n\sigma^2} \) and \( A(\omega) \sqrt{2/n\sigma^2} \) are independent Normal random variables.
Therefore,

\[
2\{A(\omega)\}^2 + \{B(\omega)\}^2]/(n\sigma^2) \sim \chi^2_2
\]

so \(2I(\omega)/\sigma^2\) is a chi-square distribution with 2 degrees of freedom or

\[
I(\omega) \sim \sigma^2 \chi^2_2/2
\]

In particular,

\[
E(I(\omega)) = \sigma^2
\]

\[
V(I(\omega)) = \sigma^4
\]

Recall that if \(X_t\) is white noise, the spectrum \(f(\omega) = \sigma^2\) so in this case \(I(\omega)\) is unbiased but an inconsistent estimator of \(f(\omega)\)

In fact there is a theorem presented in Diggle’s book.
(page 97) that generalizes these results to the case of a Gaussian and stationary process.

- Let $X_t$ be a stationary and Gaussian process with spectrum $f(\omega)$. Let $x_t; t = 1 \ldots, n$ be a partial realization of this process and $I(\omega)$ the periodogram of $x_t$.

- Let $\omega_j = 2\pi j/n$ for $j < n/2$, then as $n \to \infty$
  1. $I(\omega_j) \sim f(\omega_j)\chi^2_2/2$
  2. $I(\omega_j)$ independent of $I(\omega_k)$ for all $j \neq k$

- As an example, consider $n = 200$ observations of a white noise process $N(0,1)$ and its corresponding periodogram $I(\omega)$ for $n = 50, n = 200$.

- The obtain the periodogram, I used the following commands:
x <- rnorm(200)
per <- spec.pgram(x, plot=FALSE)
plot(2*pi*per$freq, per$spec, type='l', xlab="omega", ylab="f(omega)")
per <- spec.pgram(x[1:50], plot=FALSE)
lines(2*pi*per$freq, per$spec, lty=2)
abline(h=qchisq(0.95,2)/2, lty=3)
• The solid line is the periodogram for all 200 observations and the dashed line is the periodogram only for the first 50 observations.

• The horizontal line is the .95% quantile of $\chi^2_2/2$ random variable.

• Notice that variability for the periodogram based on 50 observations is similar to the periodogram obtained with all 200 observations.

• Only a few values of $I(\omega)$ are greater than the 0.95 quantile. These values are scattered through the frequency range.

• The quantile value gives a valid test of significance of the $\chi^2_2/2$ distribution for a prespecified value of $\omega$. 
A Test for White Noise

- To test for white noise, the proposed test statistic is to use the *maximum periodogram ordinate*

\[ T = \max\{I_1, I_2, \ldots I_m\} \]

where \( I_j = I(2\pi j/n); j < n/2 \) and \( m \) is the largest integer less than \( n/2 \).

- We known that under the null hypothesis (i.e. \( X_t \) white noise) the periodogram ordinates \( I_j \) are a random sample with a scaled \( \chi^2_2 \) distribution.

- The distribution of \( I_j \) is

\[ G(u) = Pr[I_j \leq u] = 1 - \exp(-u/\sigma^2) \]

- Given the mutual independence of \( I_j \), under the white
noise hypothesis, the distribution function for $T$ is:

$$H(t) = G(t)^m = (1 - \exp(-u/\sigma^2))^m$$

- In practice, usually $\sigma^2$ is unknown. We can substitute and estimate the variance in $H(t)$,
  $$s^2 = \sum (x_i - \bar{x})^2/(n - 1),$$
  to obtain an approximate test.

- Fisher (1929) deduced the exact distribution for
  $$T_0 = T/\{\sum_{i=1}^{m} I_j/m\}$$
  under a white noise process:

  $$Pr[T_0 > mx] = \sum_{k=1}^{r} [m!/k!(m - r)!](-1)^{k-1}(1 - kx)^{m-1}$$

  where $r$ is the largest integer less than $x^{-1}$.

- Example For $n = 200$ observations following a $N(0, 1)$ distribution, I obtained a value of $t = 5.867294$ and
\[ s^2 = 1.037762 \]

1. For the approximate test, the p-value is 0.296048.
2. For Fisher’s test, \( t_0 = 5.619912 \) and the p-value is 0.2907733

```r
x <- rnorm(200)
I <- spec.pgram(x, plot=F)$spec
t <- max(I)
t0 <- max(I)/mean(I)
s2 <- var(x)
m <- length(I)
1-(1-exp(-t/s2))^m
# k!=gamma(k+1)
```
Tapering

- This is an option that is available within this function `spec.pgram`.

`spec.pgram(x, taper=0.2)`

- A *data taper* is a transformation of $x_t$ into a new series by multiplying it by constants and to reduce the effect of extreme observations,

$$y_t = c_t x_t; \quad t = 1, 2, \ldots n$$

- The sequence $c_t$ is chosen to be close to zero at the end sections of the series, but close to one towards the central part. $(0 < c_t \leq 1)$.

- If $p$ is the proportion of observations to be tapered, $n$ is
total number of observations and \( m = np \), the *split cosine bell* taper is defined as:

\[
c_t = \begin{cases} 
0.5(1 - \cos(\pi(t - 0.5)/m)) & t = 1, \ldots, m \\
1 & t = m + 1, \ldots, n - m \\
0.5(1 - \cos(\pi(n - t - 0.5)/m)) & t = n - m + 1, \ldots, n 
\end{cases}
\]

**Smoothing the Periodogram**

- If we have the spectrum \( f(\omega) \) is a smooth function of \( \omega \), another periodogram based estimator of \( f(\omega) \) is:

\[
\hat{f}(\omega_j) = (2p + 1)^{-1} \sum_{l=-p}^{p} I(\omega_{j+l})
\]

- \( \hat{f}(\omega) \) is a simple moving average of \( I(\omega) \)
• If $X_t$ is a stationary random process with spectrum $f(\omega)$ for any Fourier frequency $\omega_j$ as $n \to \infty$
  
  - $\hat{f}(\omega_j) \sim f(\omega_j) \chi^2_{2(2p+1)}/(2(2p + 1))$
  
  - $\hat{f}(\omega_j)$ is independent of $\hat{f}(\omega_k)$ whenever $j - k \geq 2p + 1$

• A general version of this estimator is defined as

$$\hat{f}(\omega_j) = \sum_{l=-p}^{p} w_l I(\omega_{j+l})$$

with $\sum_{l=-p}^{p} w_l = 1$.

• The asymptotic distribution of $\hat{f}(\omega)$ is given by

$$\hat{f}(\omega) \sim f(\omega) \chi^2_{\nu}/\nu$$

but now the degrees of freedom are defined as
\[ \nu = 2 / \sum_{i=-p}^{p} w_i^2 \]

- Now recall that the periodogram can be expressed as

\[ I(\omega) = g_0 + 2 \sum_{k=1}^{n-1} g_k \cos(k\omega) \]

- A possible explanation of why \( I(\omega) \) is not such a great estimator of the spectrum is because \( g_k \) can be large when \( r_k \equiv 0 \), particularly for high values of \( k \).

- As we showed before, the variability of \( I(\omega) \) is not a function of the number of data points.

- Alternatively, we could use a truncated Periodogram
defined as

\[ I_K(\omega) = g_0 + 2 \sum_{k=1}^{K} g_k \cos(k\omega) \]

for a value of \( K \) that is less than \( n - 1 \).

- We also have a Lag window Periodogram,

\[ \hat{f}_\lambda(\omega) = g_0 + 2 \sum_{k=1}^{n-1} \lambda_k g_k \cos(k\omega) \]

where \( \lambda_k \) is a sequence of constants that needs to be specified by the user.

- Bartlett (1950) proposed that
\[ \lambda_k = \begin{cases} 
1 - k/S & k \leq S \\
0 & k > S 
\end{cases} \]

- Daniell (1946) proposed a sequence which corresponds to the “spans” option of spec.pgram in R/Splus.

\[ \lambda_k = \sin(\pi k/S)/(\pi k/S) \]

- Parzen (1961) proposed that

\[ \lambda_k = \begin{cases} 
1 - 6(k/S)^2 + 6(k/S)^3 & k \leq S/2 \\
2(1 - k/S)^3 & S/2 < k \leq S \\
0 & k > S 
\end{cases} \]

where large values of \( S \) correspond to less smoothing.
We consider again the CO2 data and we will look into different versions of the periodogram. Here is the R code.

```r
data(co2)
co2diff <- as.vector(diff(co2))
par(mfrow=c(2,2))
per<-spec.pgram(co2diff,taper=0,pad=0,detrend=F,
demean=F,plot=F)
lam<-1/per$freq
plam<-per$spec
i<-2<lam & lam<16
plot(lam[i],plam[i],type='l',ylab='periodogram')
mtext("Raw periodogram")
```
per<-spec.pgram(co2diff,spans=c(6),taper=0,pad=0,detrend=F,demean=F,plot=F)

per<-spec.pgram(co2diff,taper=0.3,pad=0, detrend=F, demean=F, plot=F)

per<-spec.pgram(co2diff,spans=c(6),taper=0.2,pad=0,detrend=F,demean=F,plot=F)
Raw periodogram

Smoothed

tapered with 0.2

smoothed and tapered