Activity Solution: Properties of the Periodogram

The *periodogram* of a time series $x(1), \ldots, x(N)$ is a histogram over the frequencies $(0, \pi)$, the height of the histogram across $\omega_p \pm \frac{\pi}{N}$ being

$$I(\omega_p) = \frac{NR_p^2}{4\pi},$$

where

$$R_p = \left(a_p^2 + b_p^2\right)^{\frac{1}{2}}$$

with

$$\begin{aligned} a_p &= \frac{2}{N}\sum_{t=1}^{N} x(t)\cos(\omega_p t), \\ b_p &= \frac{2}{N}\sum_{t=1}^{N} x(t)\sin(\omega_p t), \end{aligned}$$

for $p = 1, \ldots, N/2 - 1$.

We consider a special case, and show that the periodogram is not consistent as an estimator of the underlying spectrum. The case we consider here is where $x(t) \sim N(0, \sigma^2)$, independently for each $t = 1, \ldots, N$. It will be helpful to recall that if $X_i \sim N(0, 1)$, for $i = 1, \ldots, n$ then $\sum_{i=1}^{n} X_i^2 \sim \chi_n^2$, with $E(\chi_n^2) = n$ and $\mathrm{Var}(\chi_n^2) = 2n$.

1. *Clickers question:* Find $E(a_p)$ and $E(b_p)$, for $p = 1, \ldots, N/2 - 1$.
   Now, assuming N is even and $p \neq N/2$,

$$\begin{aligned} E(a_p) &= E\left(\frac{2}{N}\sum_{t=1}^{N} x(t)\cos(\omega_p t)\right) \\ &= \frac{2}{N}\sum_{t=1}^{N} E(x(t))\cos(\omega_p t) \\ &= 0 \end{aligned}$$

   and similarly $E(b_p) = 0$.

2. *Clickers question:* Find $\text{Var}(a_p)$ and $\text{Var}(b_p)$, for $p = 1, \ldots, N/2 - 1$.
   Further,

$$\begin{aligned}
\text{Var}(a_p) &= \frac{4}{N^2}\text{Var}\left(\sum_{t=1}^{N} x(t)\cos(\omega_p t)\right) \\
&= \frac{4\sigma^2}{N^2}\sum_{t=1}^{N}\cos^2(\omega_p t) \\
&= \frac{2\sigma^2}{N}.
\end{aligned}$$

   Similarly

$$\text{Var}(b_p) = \frac{2\sigma^2}{N}.$$

3. *Clickers question:* Recalling that in our special case $x(t) \sim N(0, \sigma^2)$ for each t, write down the probability distribution for a_p here.
   Well for $p = 1, \ldots, N/2 - 1$, a_p is a linear combination of Normal variables, so is itself Normal, hence

$$a_p \sim N\left(0, \frac{2\sigma^2}{N}\right).$$

   The coefficients b_p have the same distribution.

4. Now show that a_p and b_q are uncorrelated, for all choices of $p, q = 1, \ldots, N/2 - 1$.
   Since the $x(t)$ are independent,

$$\begin{aligned}
\text{Cov}(a_p, b_q) &= \frac{4}{N^2}\text{Cov}\left(\sum_{t=1}^{N} x(t)\cos(\omega_p t), \sum_{t=1}^{N} x(t)\sin(\omega_q t)\right) \\
&= \frac{4\sigma^2}{N^2}\sum_{t=1}^{N}\cos(\omega_p t)\sin(\omega_q t) \\
&= 0,
\end{aligned}$$

   even when $p = q$.

5. *Clickers question:* Hence find the expectation and variance of $I(\omega_p)$. Comment on these values.

Now

$$\begin{aligned} E\left(I\left(\omega_p\right)\right) &= E\left(\frac{NR_p^2}{4\pi}\right) \\ &= E\left(\frac{N\left(a_p^2+b_p^2\right)}{4\pi}\right) \\ &= \frac{N}{4\pi}\times 2\times\frac{2\sigma^2}{N} \\ &= \frac{\sigma^2}{\pi}, \end{aligned}$$

the spectrum of a white noise process. So the periodogram is unbiased for the true spectral density here. However

$$\begin{aligned} \mathrm{Var}\left(I\left(\omega_p\right)\right) &= \frac{\sigma^4}{4\pi^2}\mathrm{Var}\left(\chi_2^2\right) \\ &= \frac{\sigma^4}{\pi^2}, \end{aligned}$$

a constant. So in this special case, and in fact in general, Var$(I\left(\omega_p\right)) \nrightarrow 0$ *as* $N\to\infty$*, so the periodogram is not a consistent estimator of the spectrum.*

6. *Clickers question:* What is the distribution of $I\left(\omega_p\right)$? How does this behave as N grows large?
   Now since

$$\begin{aligned} \frac{a_p}{\sqrt{2\sigma^2/N}} &\sim N\left(0,1\right), \\ \frac{b_p}{\sqrt{2\sigma^2/N}} &\sim N\left(0,1\right), \end{aligned}$$

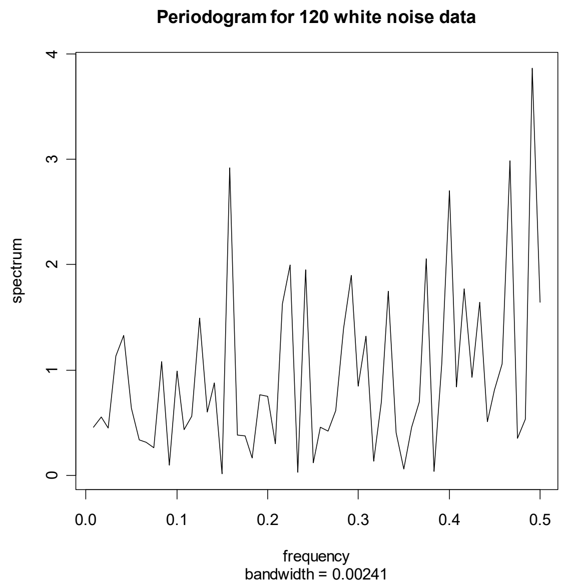
   we have

$$\frac{N\left(a_p^2+b_p^2\right)}{2\sigma^2}=\frac{2\pi I\left(\omega_p\right)}{\sigma^2}\sim\chi_2^2.$$

   Note that the distribution does not depend on N.

7. The vector $\mathbf{x}$ contains 120 independent observations simulated from

$N(0,1)$. Its periodogram is plotted below:



How is R defining frequency in the above? In theory, what should the periodogram look like?
R defines frequency as cycles per unit time, $f = \omega/2\pi$, *and defines the periodogram as* $\pi I(\omega)$. *So the horizontal axis is*

$$f_p = \frac{\omega_p}{2\pi} = \frac{p}{N}$$

for $p = 0, \ldots, N/2$. *Of course, the spectrum here is actually flat.*

8. How might you define a "significant" value of the periodogram in the above example?
Note that $\chi_2^2(0.95) = 6.0$, *so a "significant" value of* $2\pi I(\omega)$ *could be taken as one over 6. Now above there are plotted sixty values of* $\pi I(\omega)$, *of which we would expect about three to exceed 3.*

9. Re-cap what you have done during this activity. Why do you think you were asked to do this activity? What have you learned?
We explored how the periodogram will behave when the data are an independent sample from a Normal distribution. We found, as would be anticipated, that the expected values of the periodogram are all a

constant. Unexpectedly perhaps, we saw that the variance of the periodogram values is also a constant and does not depend on the sample size. In that sense the periodogram cannot be a consistent estimator of the spectral density, as its variance does not decrease as the sample size increases. We showed that, to a constant multiple, the periodogram values follow a Chi-squared distribution (on two degrees of freedom). Learning outcomes encountered in this activity include:

(a) *Define the periodogram for a time series.*

(b) *Explain why the (raw) periodogram is not a consistent estimator of the spectrum of a series.*