Lab 10: Spectral Density Estimation

STAT 443

Caden Hewlett

2024-03-22

Let $\{Z_t\}_{t\in\mathbb{Z}} \overset{\text{iid}}{\sim} \text{WN}(0,1)$..

In `R`, the default periodogram needs to be adjusted to `log = "no"` and uses $\omega/2\pi$ resulting in range scales from zero to 0.5 rather than 0 to π.

Importantly, as an estimate for spectral density, the periodogram **shoud be divided by** π to be consistent with the results in class.

Question 1

Here we will compare the spectral density of white noise with the periodograms obtained from simulated white noise samples.

Part a

Plot the spectral density function for $\{Z_t\}_{t\in\mathbb{Z}}$.

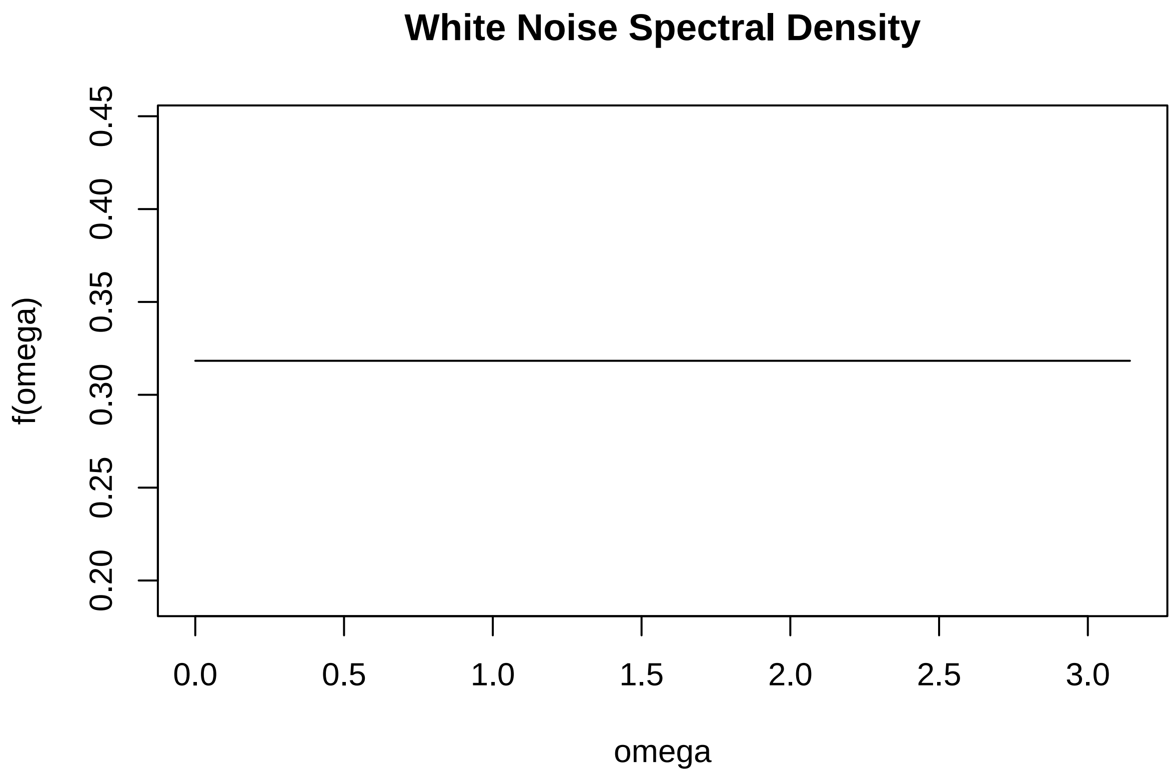
Recall, for the given pure WN process with variance 1, $\gamma(0) = 1$ and $\gamma(h) = 0 \;\forall h \neq 0 \in \mathbb{Z}$.

Hence, for $\omega \in (0,\pi)$.

$$f(\omega) = \frac{1}{\pi}\Big(\gamma(0) + 2\sum_{h=1}^{\infty}\gamma(h)\cos(\omega h)\Big) = \frac{1}{\pi}\Big(1 + 2\sum_{h=1}^{\infty} 0 \times \cos(\omega h)\Big) = \frac{1}{\pi}$$

So the plot of $f(\omega)$ is simply the horizontal line $1/\pi$ in $(0,\pi)$.

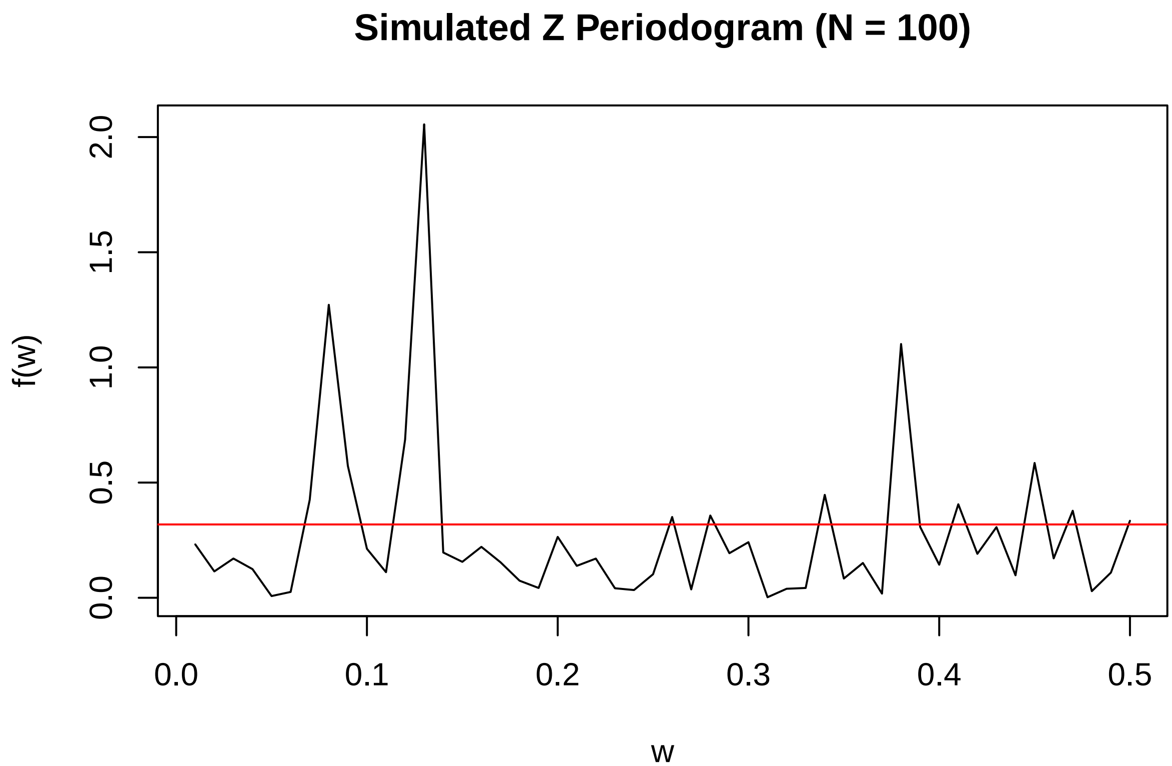
```
N = 100
plot(x = seq(from = 0, to = pi, length.out = N),
     y = rep(1/pi, times = N), type = 'l', main = "White Noise Spectral Density",
     xlab  = "omega", ylab = "f(omega)")
```



Part b

We simulate $N = 100$ pulls from a WN random variable.

```
set.seed(443)
WN = arima.sim(model = list(), n = N)
f_w_WN = spec.pgram(WN, log = "no", plot = F)
plot(y = f_w_WN$spec / pi, x = f_w_WN$freq, type = 'l',
     ylab = "f(w)", xlab = "w", main = "Simulated Z Periodogram (N = 100)")
abline(h = 1/pi, col = 'red')
```




```
print(c(mean(f_w_WN$spec / pi), 1/pi))
```

```
## [1] 0.2745045 0.3183099
```

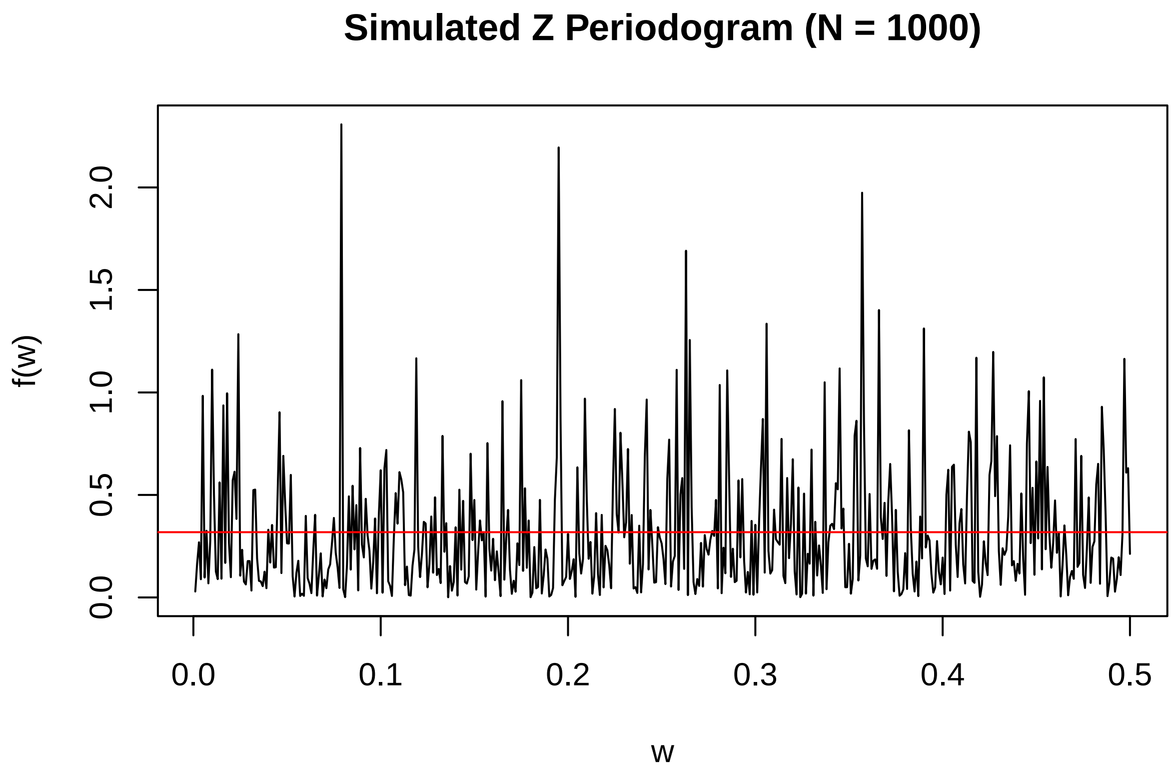
Comment on what you observe in regard to the true spectrum and its estimate here:

Unsurprisingly, the simulated values present a lot more noise than the true spectrum. There are occasionally large spikes in the simulated $f(\omega)$ values, but no particular pattern to these values across multiple seeds. However, it seems like the values in the periodogam oscillate about the true spectrum (red line), which makes sense. The mean observations seem rather far from the true mean.

Part c

Repeating the above with $N = 1000$.

```
set.seed(443)
WN = arima.sim(model = list(), n = 1000)
f_w_WN = spec.pgram(WN, log = "no", plot = F)
plot(y = f_w_WN$spec / pi, x = f_w_WN$freq, type = 'l',
     ylab = "f(w)", xlab = "w", main = "Simulated Z Periodogram (N = 1000)")
abline(h = 1/pi, col = 'red')
```



Below, we compare the mean with the true mean:

```
print(c(mean(f_w_WN$spec / pi), 1/pi))
```

```
## [1] 0.3137574 0.3183099
```

Comments: while the simulations still seem "noisy" as we increase N, the long-term average value of the periodogram seems to be getting closer and closer to the true value (as expressed by the comparison of means above.) In effect, we see the Law of Large Numbers at work.

Part D

(d) Repeat parts (b) and (c) several times. Comment on how the periodogram behaves as an estimator of the spectral density function based on what you have observed.

Comments: After repeating the previoous questions, it seems that the periodogram is a decent estimator of spectral density, but is quite reliant on the number of observations in terms of precision. It should be noted as well that in cases such as these where the true spectral density is some easily-defined function, the periodogram doesn't necessarily reflect this characteristic (i.e. it's not a horizontal line) due to random noise. However, certain other properties such as the mean value seem to be in-line with the true spectrum for N large. The approximation of spectral density via periodogram seems to be heavily reliant on the chosen number of observations N, which may imply that it is an asymptotically unbiased estimator of spectral density rather than purely unbiased (and not necessarily consistent either.)

Question 2

Part 1

Let $\{X_t\}_{t\in\mathbb{Z}}$ be defined by the following MA process, assuming that $\{Z_t\}_{t\in\mathbb{Z}}$ are *iid*

$$X_t = Z_t + 0.9Z_{t-1}$$

Directly,

$$\begin{aligned}
\gamma(0) &= \text{cov}(X_t, X_t)\\
\gamma(0) &= \text{cov}(Z_t + 0.9Z_{t-1}, Z_t + 0.9Z_{t-1})\\
\gamma(0) &= \text{cov}(Z_t, Z_t) + \text{cov}(0.9Z_{t-1}, 0.9Z_{t-1})\\
\gamma(0) &= \sigma^2 + (0.9)^2\sigma^2 \quad \text{note } \sigma^2 = 1\\
\gamma(0) &= (1 + 0.81) = \boxed{1.81}
\end{aligned}$$

Further,

$$\begin{aligned}
\gamma(1) &= \text{cov}(X_t, X_{t+1})\\
\gamma(0) &= \text{cov}(Z_t + 0.9Z_{t-1}, Z_{t+1} + 0.9Z_t)\\
\gamma(0) &= \text{cov}(Z_t, 0.9Z_t)\\
\gamma(0) &= 0.9(\sigma^2) = \boxed{0.9}
\end{aligned}$$

So, for $h \in \mathbb{Z}$,

$$\gamma(h) = \begin{cases} \gamma(-h), & h < 0\\ 1.81, & h = 0\\ 0.90, & h = 1\\ 0, & \text{otherwise} \end{cases}$$

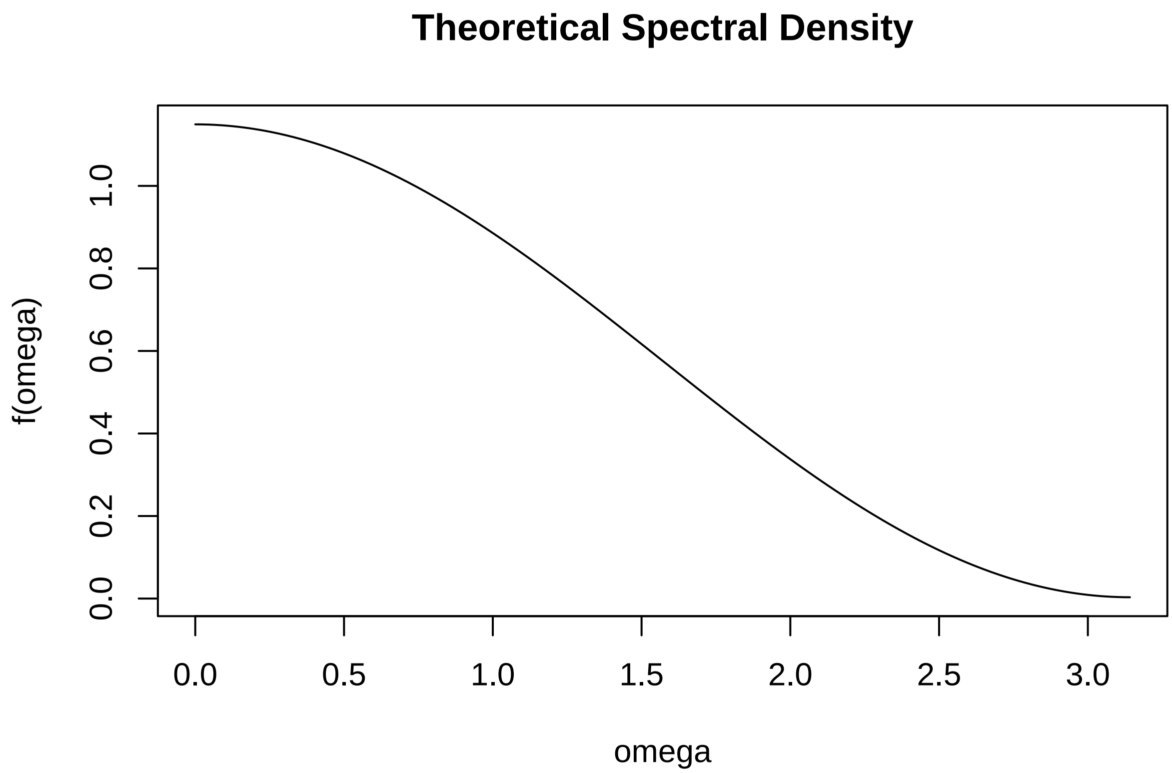
From this, we can find $f(\omega)$:

$$\begin{aligned}
f(\omega) &= \frac{1}{\pi}\Big(\gamma(0) + 2\sum_{h=1}^{\infty}\gamma(h)\cos(\omega h)\Big)\\
f(\omega) &= \frac{1}{\pi}\Big(\gamma(0) + 2\gamma(1)\cos(w \times 1) + \sum_{h=2}^{\infty}\gamma(h)\cos(\omega h)\Big)\\
f(\omega) &= \frac{1}{\pi}\Big(1.81 + 2(0.9)\cos(\omega) + 0\Big)\\
f(\omega) &= \boxed{\frac{1.81 + 1.8\cos(\omega)}{\pi}}
\end{aligned}$$

So, the theoretical plot of $f(w)$ looks like:

```
f = function(omega){
  (1/pi)*(1.81 + 1.8*cos(omega))
}
omega = seq(from = 0, to = pi, length.out = 10*N)

plot(omega, f(omega), main = "Theoretical Spectral Density", type = 'l')
```

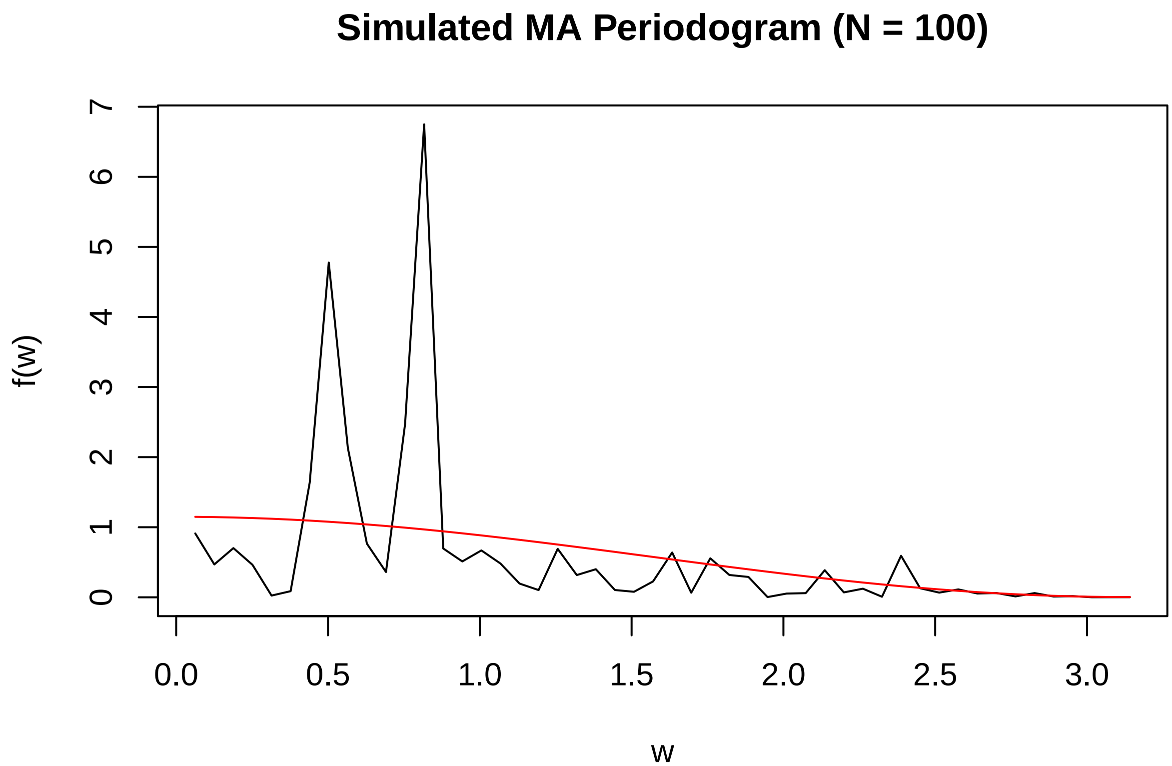


Part 2

Use the `arima.sim()` command to simulate a series of length 100. Use `spec.pgram()` to create and plot the periodogram for your sample. Comment on what you observe in regard to the true spectrum and its estimate here.

Since `R` uses $\omega/2\pi$

```
set.seed(443)
MA = arima.sim(model = list(ma = 1), n = 100, sd = 1)
f_w_MA = spec.pgram(MA, log = "no", plot = F)
plot(y = f_w_MA$spec / pi, x = f_w_MA$freq*(2*pi), type = 'l',
     ylab = "f(w)", xlab = "w", main = "Simulated MA Periodogram (N = 100)")
lines(x = f_w_MA$freq*(2*pi), y = f(f_w_MA$freq*(2*pi)), col = 'red')
```



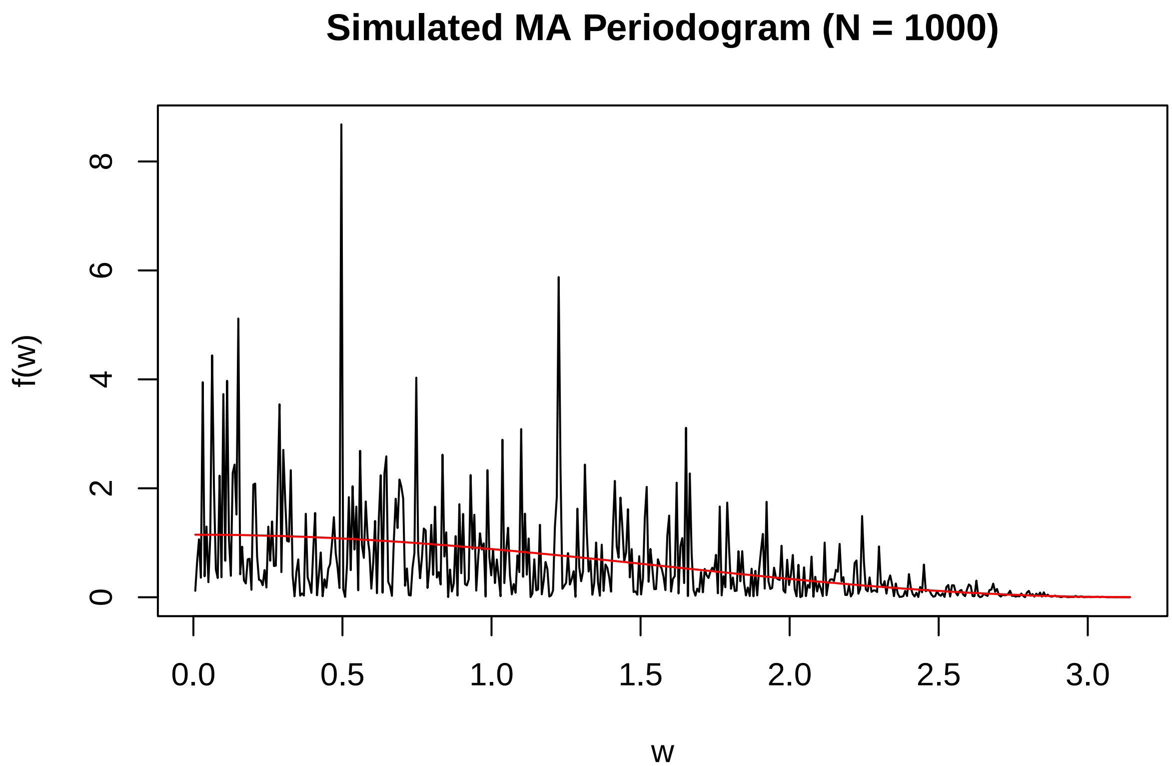
Comments we see the disribution of high and low frequencies roughly matches the theoretical spectral density; however, there is a lot of noise and we see a few more high-frequency values than we would expect under the theoretical model, as well as "spikes" in concentration of low-frequency values.

Part 3

Repeat…

Since `R` uses $\omega/2\pi$

```
set.seed(443)
MA = arima.sim(model = list(ma = 1), n = 1000,sd = 1)
f_w_MA = spec.pgram(MA, log = "no", plot = F)
plot(y = f_w_MA$spec / pi, x = f_w_MA$freq*(2*pi), type = 'l',
     ylab = "f(w)", xlab = "w", main = "Simulated MA Periodogram (N = 1000)")
lines(x = f_w_MA$freq*(2*pi), y = f(f_w_MA$freq*(2*pi)), col = 'red')
```



Comments while there is a lot more noise, the general pattern of the periodogram seems to follow the theoretical spectral density in terms of the distribution of high and low frequencies.

Part d

Repeat parts (b) and (c) several times. Comment on how the periodogram from R behaves as an estimator of the spectral density function based on what you have observed.

Comments After repeating the simulations many times, our conclusions are similar to those in part **a**; namely, that the general shape pattern of the periodogram seems to follow the theorized spectral density however with a lot more noise/error. Although the shape and mean are convergent as we increase N, it doesn't seem like the periodogram every exactly matches the spectral density due to the variance inherent in the simulation process. It's important to note that $f(\omega)$ is a theoretical mathematical function, and hence has no variance at all. The second we bring our values into the "real world" it is very difficult for them to match the theorized value. However, for a truly unbiased estimator, I would personally anticipate a better performance than what we're seeing as $N \uparrow$, which supports our observations in part **a** of the periodogram being potentially biased.

Question 3

The process:

$$X_t = -0.8X_{t-1} + Z_t$$

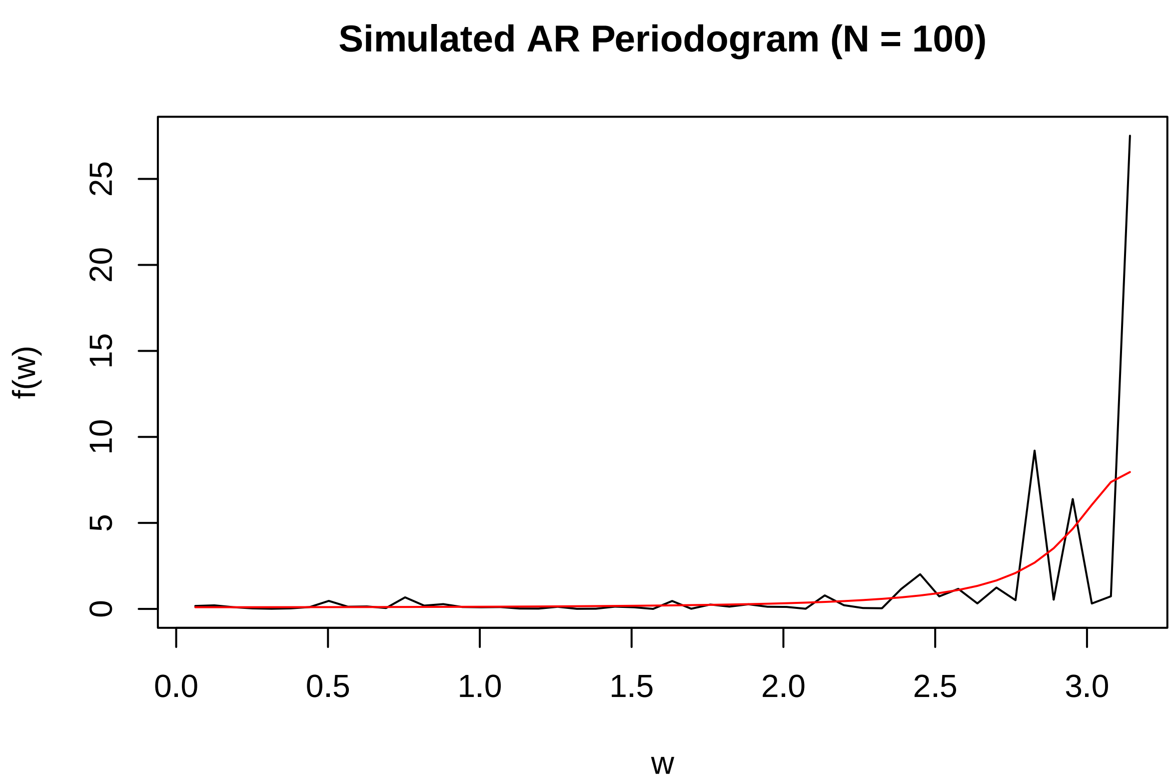
With spectral density:

$$f(\omega) = \frac{1}{\pi(1 + 1.6\cos(\omega) + 0.8^2)}$$

```
f <- function(omega){
  1 / (pi*(1 + 1.6*cos(omega) + 0.8^2))
}
```

Part b

```
set.seed(443)
AR = arima.sim(model = list(ar = -0.8), n = 100, sd = 1)
f_w_AR = spec.pgram(AR, log = "no", plot = F)
plot(y = f_w_AR$spec / pi, x = f_w_AR$freq*(2*pi), type = 'l',
     ylab = "f(w)", xlab = "w", main = "Simulated AR Periodogram (N = 100)")
lines(x = f_w_AR$freq*(2*pi), y = f(f_w_AR$freq*(2*pi)), col = 'red')
```



Part d

Comments After repeating the simulations many times, our conclusions are similar to those in part **a** and **b**; namely, that the general shape pattern of the periodogram seems to follow the theorized spectral density however with a lot more noise/error. Although the shape and mean are convergent as we increase N, it doesn't seem like the periodogram every exactly matches the spectral density due to the variance inherent in the simulation process. It's important to note that $f(\omega)$ is a theoretical mathematical function, and hence has no variance at all. The second we bring our values into the "real world" it is very difficult for them to match the theorized value. However, for a truly unbiased estimator, I would personally anticipate a better performance than what we're seeing as $N \uparrow$, which supports our observations in part **a** and **b** of the periodogram being potentially biased.