STAT 443: Lab 11

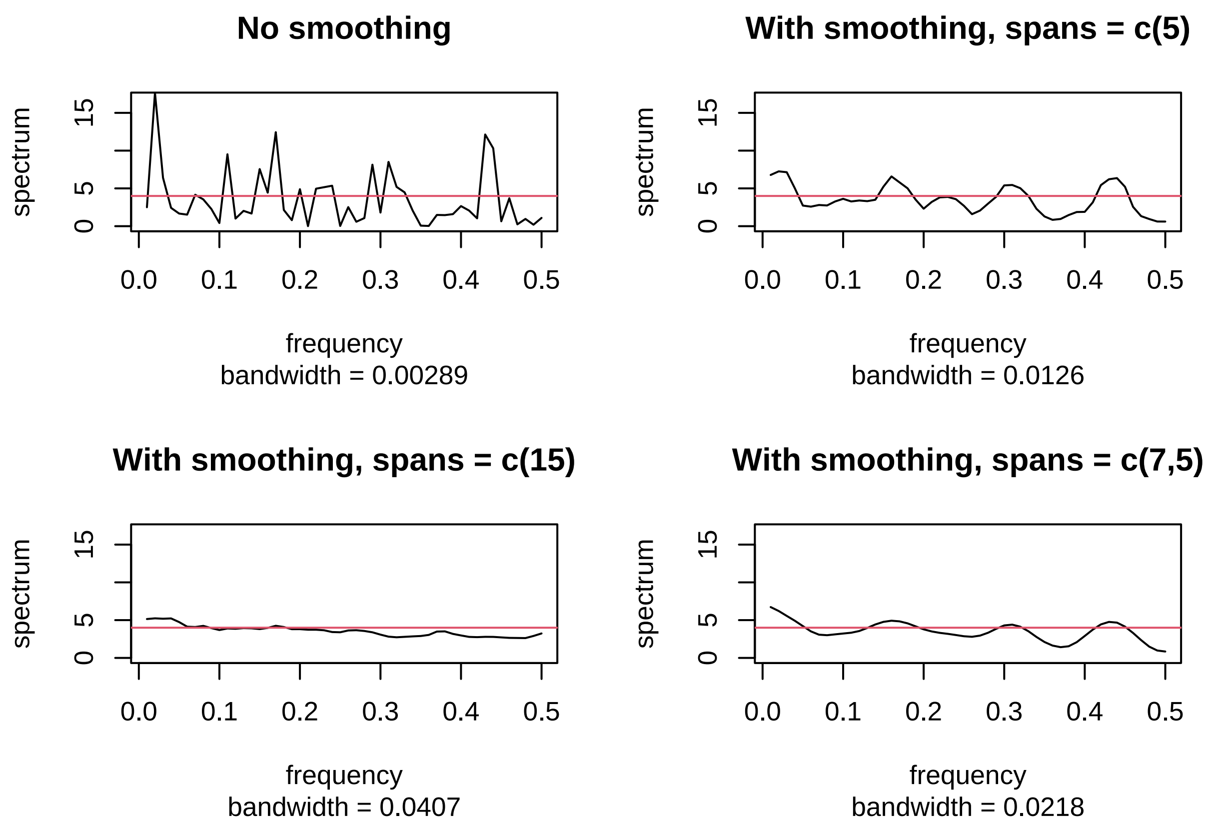
Saksham Sudershan (Student #31339427)

7 April, 2022

Problem 1

(a)

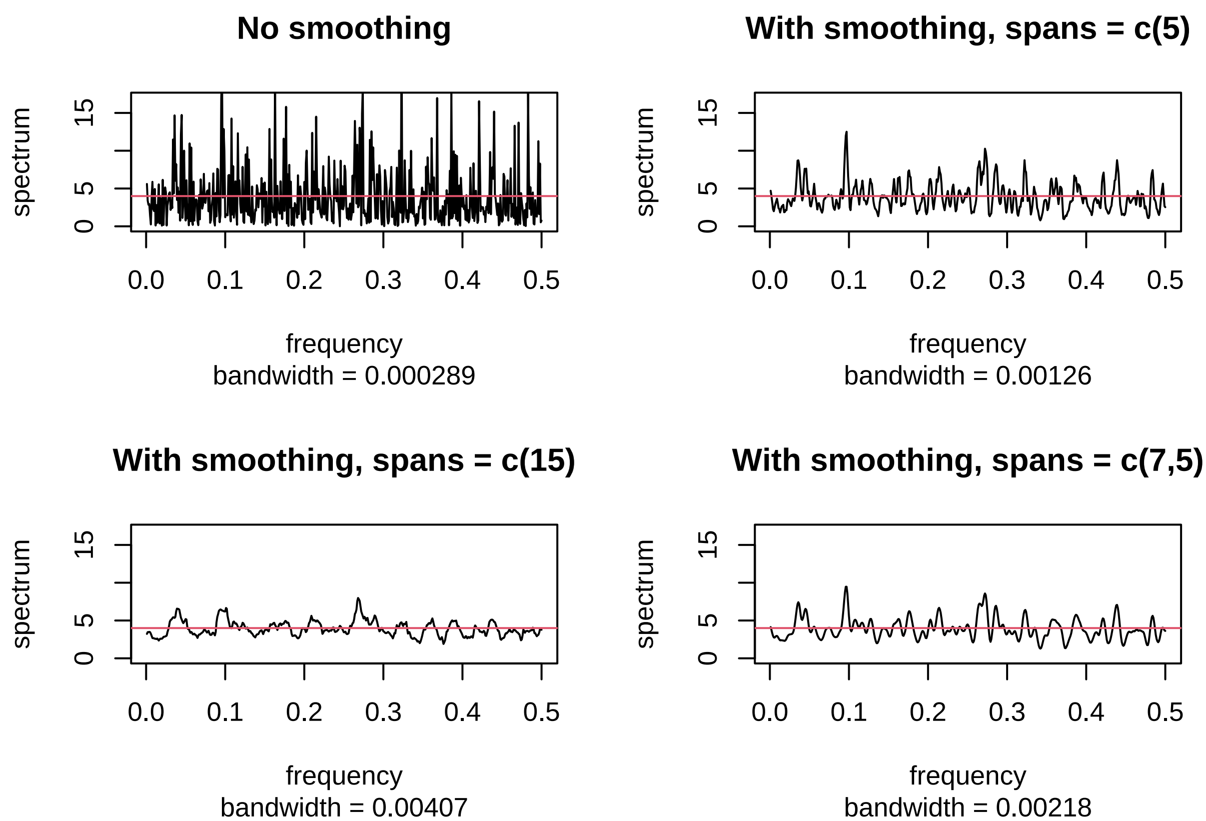
```r
# Creating sim data
set.seed(1234)
WN_1 <- arima.sim(model = list(), n = 100, sd = 2)

par(mfrow = c(2,2))

spec.pgram(WN_1, log = "no", ylim = c(0,17), main = "No smoothing")
abline(h = 4, col = 2)
# (i)
spec.pgram(WN_1, log = "no", ylim = c(0,17), spans = c(5), main = "With smoothing, spans = c(5)")
abline(h = 4, col = 2)
#(ii)
spec.pgram(WN_1, log = "no", ylim = c(0,17), spans = c(15), main = "With smoothing, spans = c(15)")
abline(h = 4, col = 2)
# (iii)
spec.pgram(WN_1, log = "no", ylim = c(0,17), spans = c(7, 5), main = "With smoothing, spans = c(7,5)")
abline(h = 4, col = 2)
```



The true spectrum is given by the red line. In the first graph, we can see that the sample periodogram has a lot more fluctuations as compared to the true spectrum, which is constant. The smoothing with spans = 5 and spans = (7,5) help preserve the most dominant frequencies present in our sample while also being quite similar to the true spectrum. With spans = 15, it is quite hard to see the dominant frequencies even though it does resemble our true spectrum well.

(b)

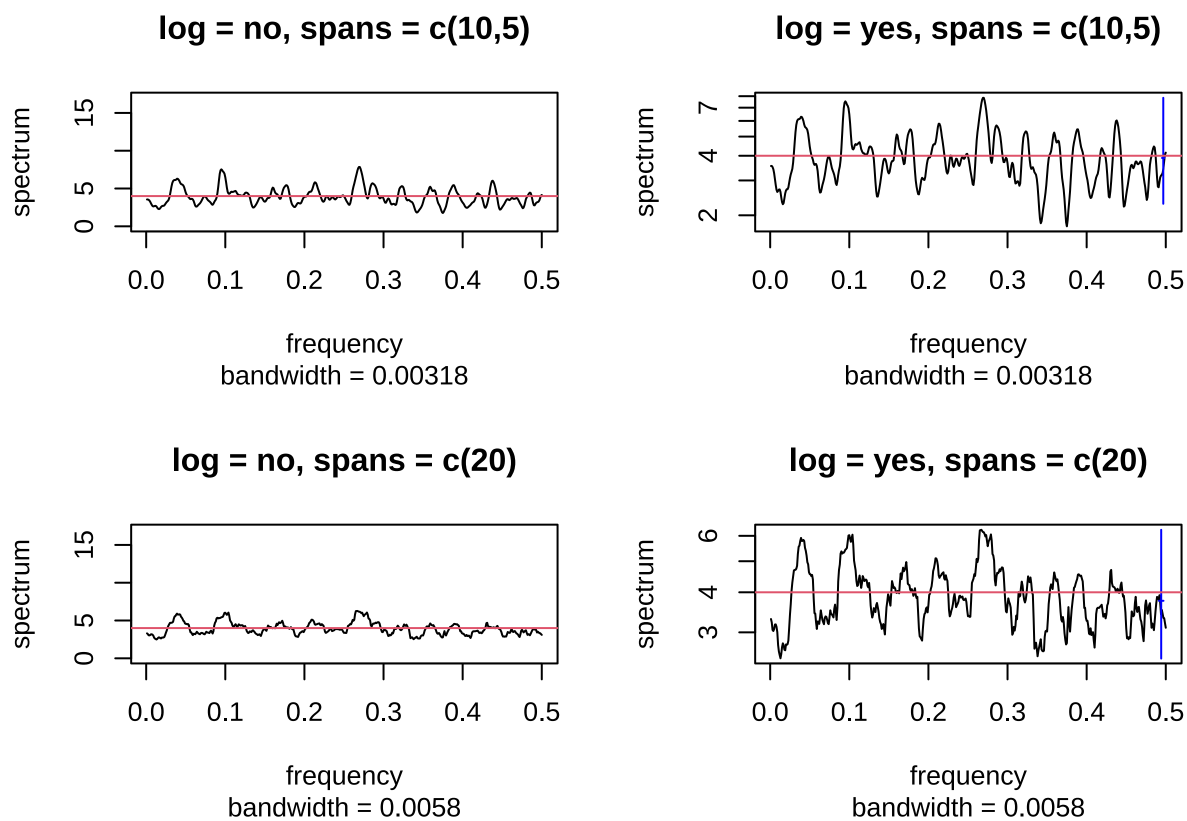
```
# Creating sim data
set.seed(1234)
WN_2 <- arima.sim(model = list(), n = 1000, sd = 2)

par(mfrow = c(2,2))

spec.pgram(WN_2, log = "no", ylim = c(0,17), main = "No smoothing")
abline(h = 4, col = 2)
# (i)
spec.pgram(WN_2, log = "no", ylim = c(0,17), spans = c(5), main = "With smoothing, spans = c(5)")
abline(h = 4, col = 2)
#(ii)
spec.pgram(WN_2, log = "no", ylim = c(0,17), spans = c(15), main = "With smoothing, spans = c(15)")
abline(h = 4, col = 2)
# (iii)
spec.pgram(WN_2, log = "no", ylim = c(0,17), spans = c(7, 5), main = "With smoothing, spans = c(7,5)")
abline(h = 4, col = 2)
```



The true spectrum is given by the red line. With more observation in our white noise, there is a lot more fluctuations compared to the previous graphs. This time however, choosing a larger m for spans helps due to the large number of observations. With spans = 15, it removes the large amount of fluctuations and allows us to still see the dominant frequencies while also being quite close to the true spectrum.

(c)

```r
# Creating sim data
set.seed(1234)
WN_2 <- arima.sim(model = list(), n = 1000, sd = 2)

par(mfrow = c(2,2))

spec.pgram(WN_2, log = "no", ylim = c(0,17), spans = c(10,5), main = "log = no, spans = c(10,5)")
abline(h = 4, col = 2)
#(i)
spec.pgram(WN_2, spans = c(10,5), main = "log = yes, spans = c(10,5)")
abline(h = 4, col = 2)
#(ii)
spec.pgram(WN_2, log = "no", ylim = c(0,17), spans = c(20), main = "log = no, spans = c(20)")
abline(h = 4, col = 2)
#(iii)
spec.pgram(WN_2, spans = c(20), main = "log = yes, spans = c(20)")
abline(h = 4, col = 2)
```



We can see here that using logs does not really make a difference, but allows us exaggerated fluctuations, making it a bit easier. Both the spans of (10,5) and (20) are helpful in making the data easy to interpret, however the averaging for a second time makes it comparatively easy to see the peaks and troughs in the data, while not overly simplifying it which we see in the graph with spans = (20). I would likely choose the logged graph with spans = (10,5).

(d)

We know that

$$\frac{2I(\omega)}{f(\omega)} \sim \chi_2^2$$

$$\Rightarrow log(2) + log(I(\omega)) - log(f(\omega)) \sim \chi_2^2$$

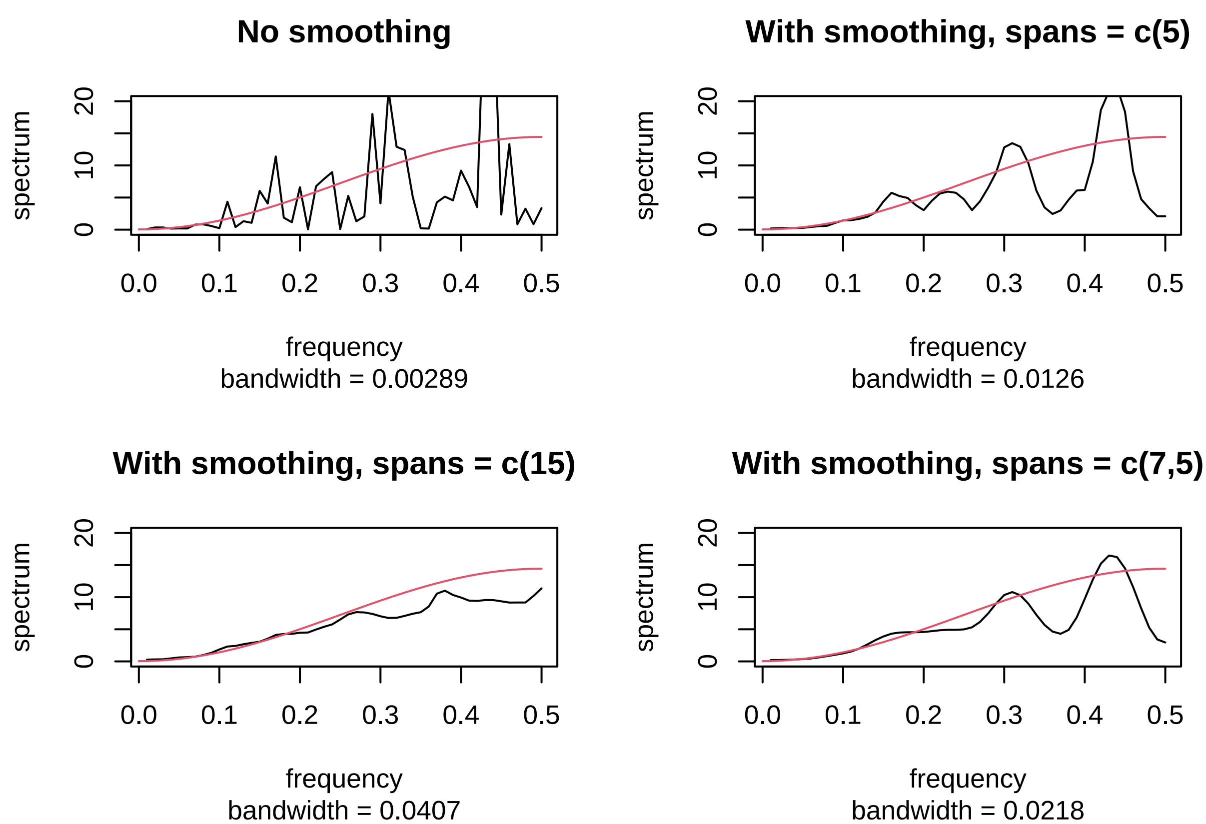
$$\Rightarrow log(I(\omega)) - log(f(\omega)) \sim \chi_2^2 - log(2)$$

We see that the log makes it such that the difference does not depend upon ω, which is in line with what we see in the previous part.

Problem 2

(a)

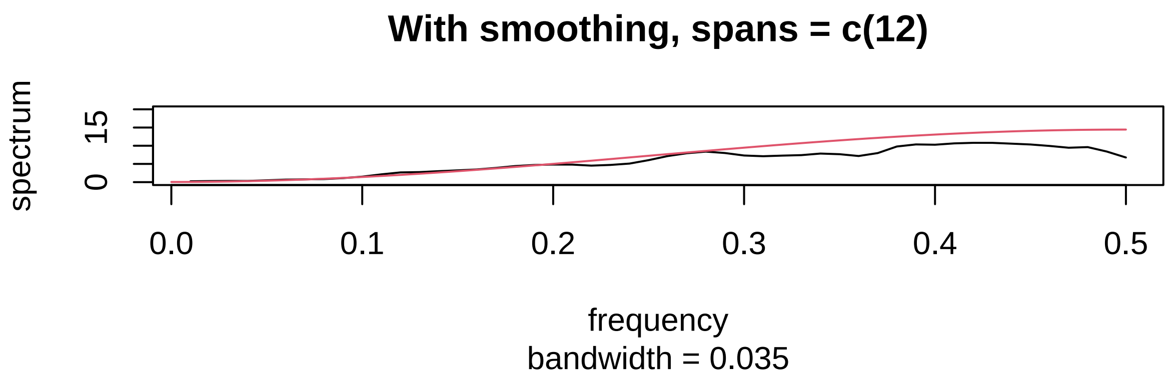
```r
set.seed(1234)
X <- arima.sim(model = list(ma = c(-0.9)), n = 100, mean=0, sd=2)

par(mfrow = c(2,2))

spec.pgram(X, log = "no", ylim = c(0, 20), main = "No smoothing")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)

spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(5), main = "With smoothing, spans = c(5)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)

spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(15), main = "With smoothing, spans = c(15)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)

spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(7, 5), main = "With smoothing, spans = c(7,5)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)
```

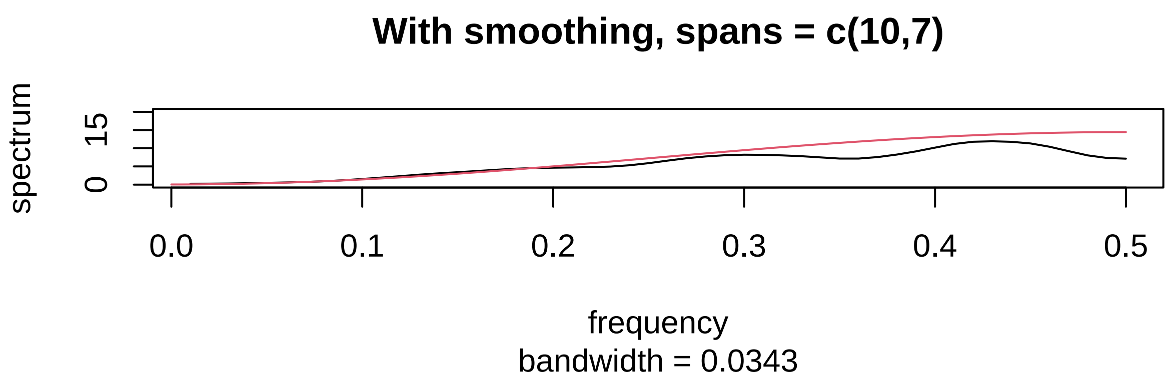


The periodogram with no smoothing has the most fluctuations, as would be expected. The periodograms with spans = 5 or spans = (7,5) give us estimates close to the true spectrum marked in red, still with fluctuations. Spans = 15 gives us a mostly consistent line with very little fluctuations but it is lower than the true spectrum.

(b)

```
par(mfrow = c(2,1))

spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(12), main = "With smoothing, spans = c(12)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)

spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(10,7), main = "With smoothing, spans = c(10,7)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)
```



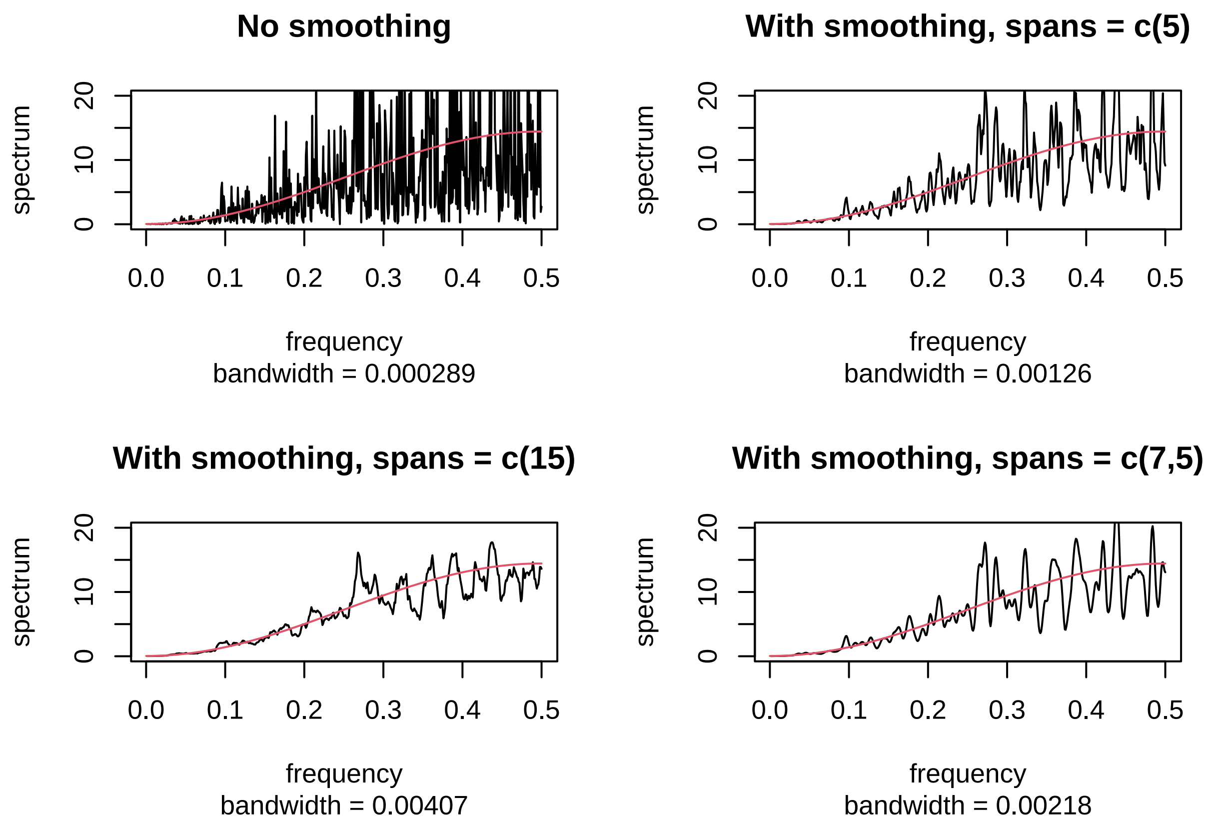


Smoothing with higher spans gives us a smoother however, it is consistently lower than the true spectrum. However, spans = 12 and spans = 10,7 give us decently smoothed periodograms which resemble the true spectrum quite closely.

(c)

```
set.seed(1234)
X <- arima.sim(model = list(ma = c(-0.9)), n = 1000, mean=0, sd=2)

par(mfrow = c(2,2))

spec.pgram(X, log = "no", ylim = c(0, 20), main = "No smoothing")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)
```

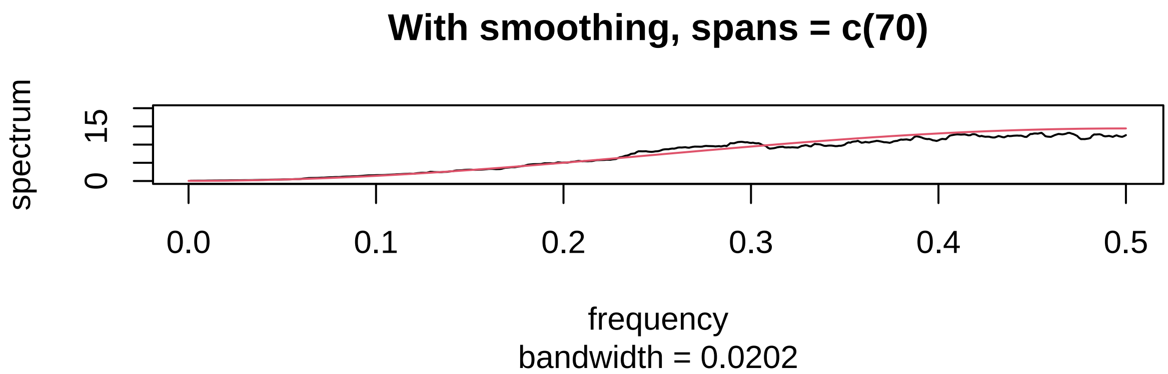
```
spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(5), main = "With smoothing, spans = c(5)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)

spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(15), main = "With smoothing, spans = c(15)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)

spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(7, 5), main = "With smoothing, spans = c(7,5)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)
```

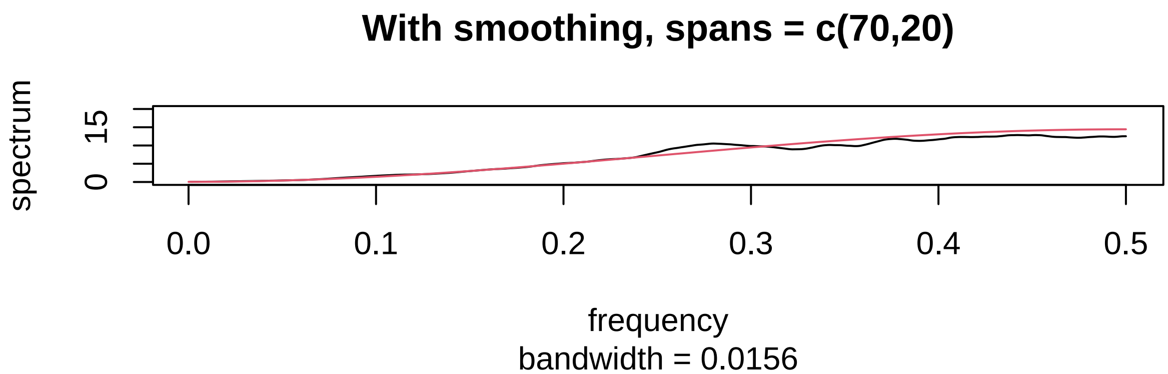


The true spectrum is given by the red line. With more observation in our sample, there is a lot more fluctuations compared to the previous graphs.The periodograms with spans = 5 or spans = (7,5) give us estimates which have a lot fluctuations but still seem close to the true spectrum marked in red. Spans = 15 gives us the least fluctuations while being very close to the true spectrum.

(d)

```
par(mfrow = c(2,1))

spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(70), main = "With smoothing, spans = c(70)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)

spec.pgram(X, log = "no", ylim = c(0, 20), spans = c(50,20), main = "With smoothing, spans = c(70,20)")
curve((7.24-7.2*cos(x*2*pi)), from = 0, to = 0.5, add = T, col = 2)
```





Smoothing with spans = c(50, 20) or spans = c(70) gives periodograms that are very close to the spectrum. As we can see, we needed more smoothing for a larger sample size, but the estimate performs better with a larger sample size, especially at higher frequencies