**Stat 510 Week 12 Homework**

**Submit answers in a Word document or a pdf to the designated ANGEL Drop Box in the Drop Boxes folder.**

1. A. What are the weighting coefficients for a Daniell kernel with *m* = 3?

The weights are:

coef[-3] = 0.1429

coef[-2] = 0.1429

coef[-1] = 0.1429

coef[ 0] = 0.1429

coef[ 1] = 0.1429

coef[ 2] = 0.1429

coef[ 3] = 0.1429

B. What are the weighting coefficients for a Daniell kernel that is convoluted (repeated) with *m* = 3 in both smoothings?

The weighting coefficients for a Daniell kernel that is convoluted (repeated) with m = 3 are:

coef[-6] = 0.02041

coef[-5] = 0.04082

coef[-4] = 0.06122

coef[-3] = 0.08163

coef[-2] = 0.10204

coef[-1] = 0.12245

coef[ 0] = 0.14286

coef[ 1] = 0.12245

coef[ 2] = 0.10204

coef[ 3] = 0.08163

coef[ 4] = 0.06122

coef[ 5] = 0.04082

coef[ 6] = 0.02041

C. What are the weighting coefficients for a modified Daniell kernel with *m* = 3?

The weighting coefficients for a modified Daniell kernel with m = 3 are:

coef[-3] = 0.08333

coef[-2] = 0.16667

coef[-1] = 0.16667

coef[ 0] = 0.16667

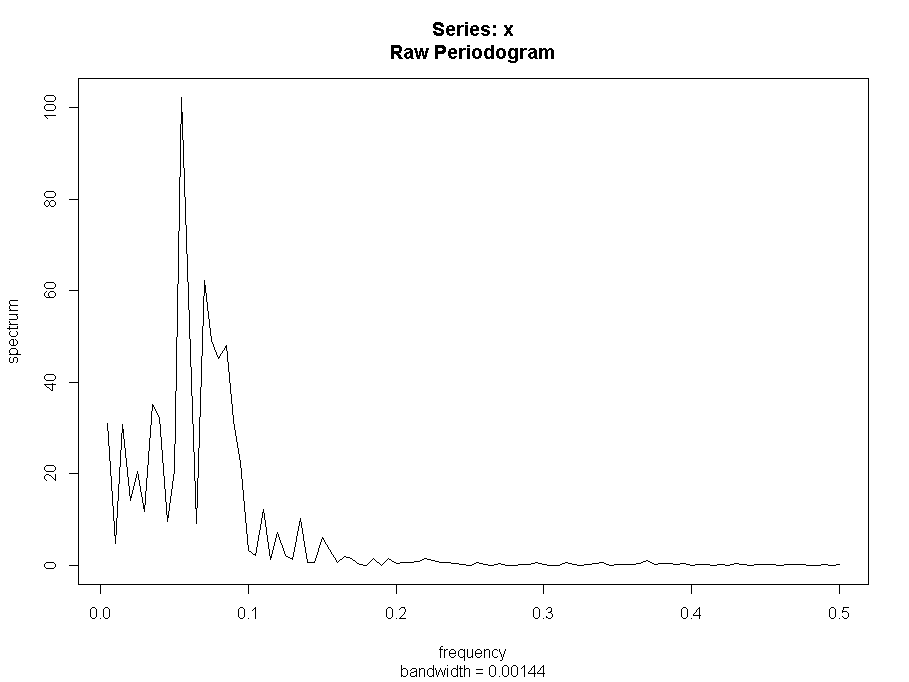
coef[ 1] = 0.16667

coef[ 2] = 0.16667

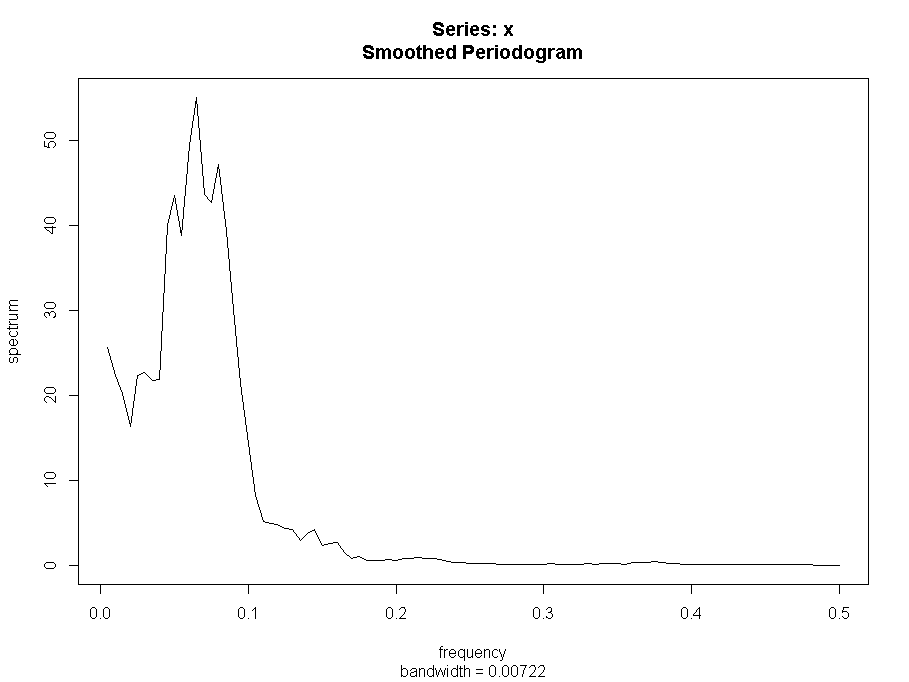
coef[ 3] = 0.08333

2. Use the dataset “week12sim.dat” from the Week 12 folder. The data are *n* = 200 values simulated from an AR(2) model (with no trend).

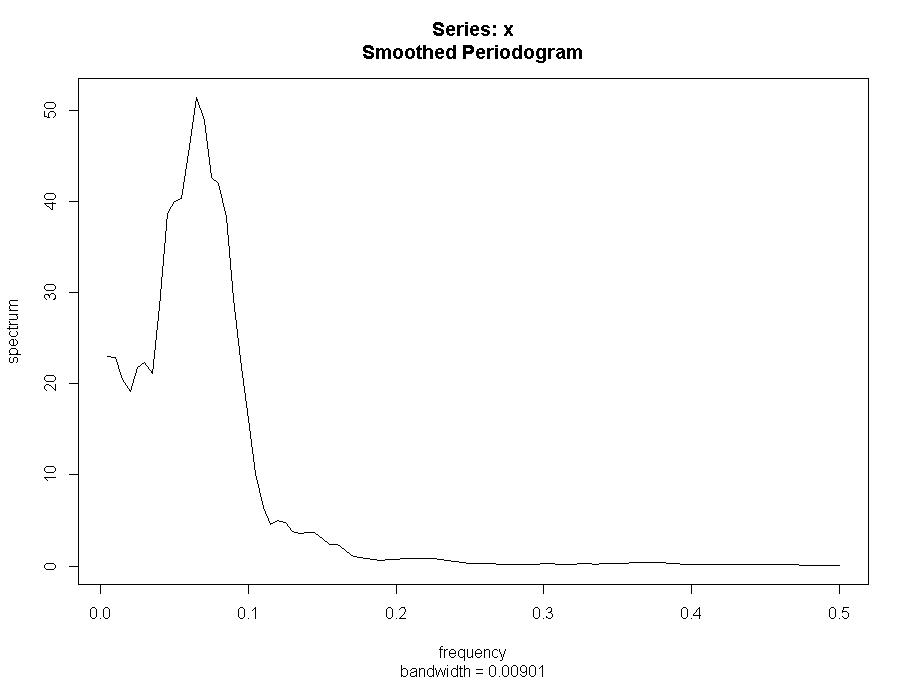
A. Use the spec.pgram command to determine the raw periodogram. Give the plot as the answer to this part.



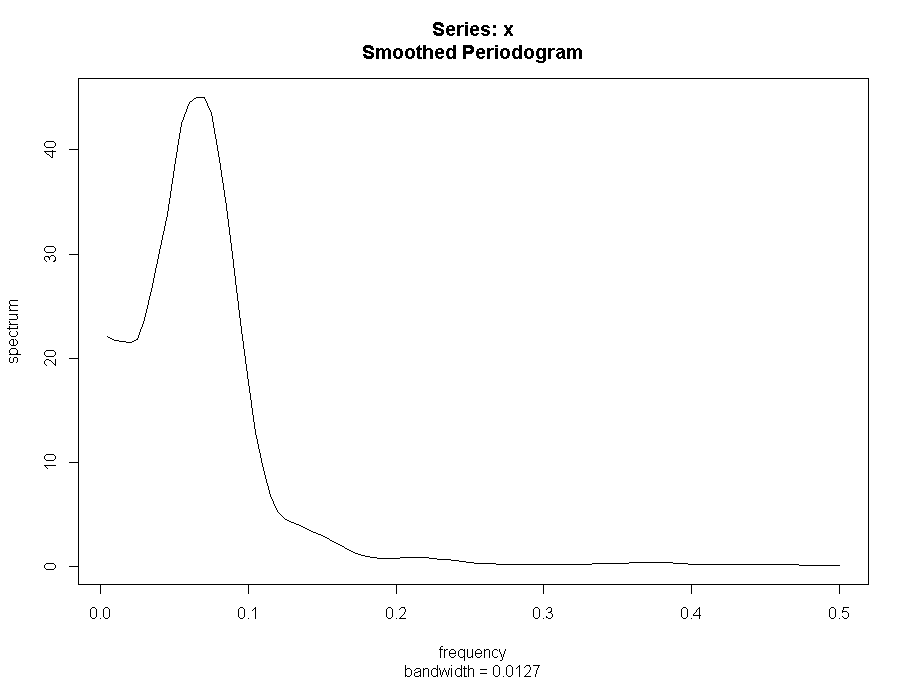
B. Smooth the periodogram using a Daniell kernel with m = 2. Give the plot as the answer to this part.



C. Smooth the periodogram using a modified Daniell kernel with m = 3 (so span length = 7). Give the plot as the answer to this part.



D. Smooth the periodogram using two passes of a modified Daniell kernel with m = 3 for each pass. (If you’re using the spans parameter of spec.pgram, it would be spans = c(7,7). Give the plot as the answer to this part.



E. Using the smoothed periodogram in part D, identify the approximate frequency at which the maximum spectral density occurs. This can be estimated visually, but also can be done be assigning a name to the spec.pgram result by doing something like specvalues = spec.pgram(……

After carrying out this command, use the command specvalues and examine the results.

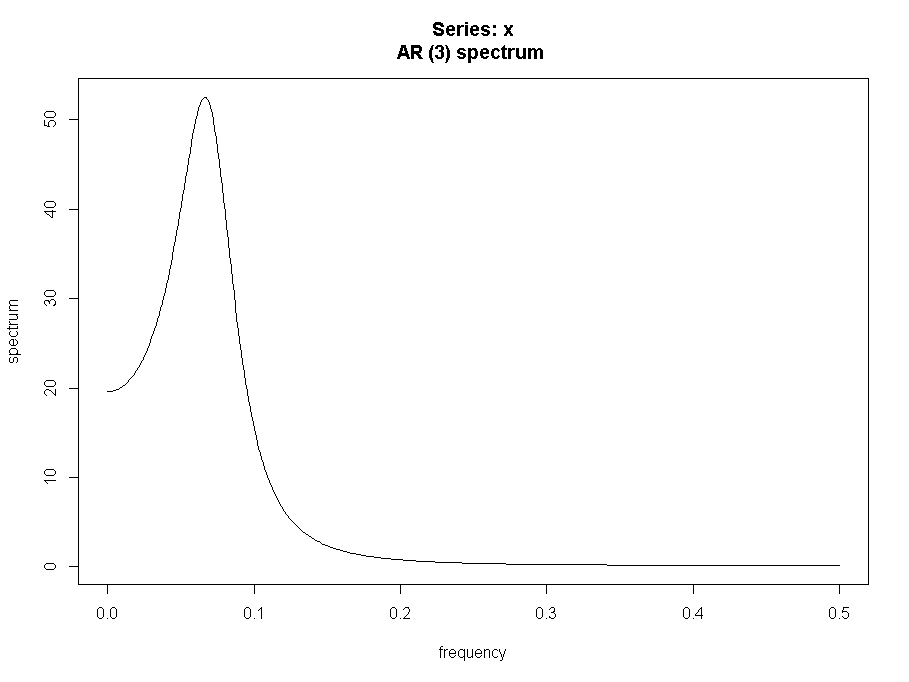
What is the frequency at which the peak occurs? What is the period associated with this frequency? Recall that period = 1/frequency.

The frequency is: 0.065

The period associated with this cycle = 1/0.065 = 15.38

F. Use spec.ar to determine the parametric estimate of the spectral density. Give the plot of the estimated spectral density. What AR order was used to create this estimate?

(Notice that it didn’t turn out to be the order used to simulate the data!)



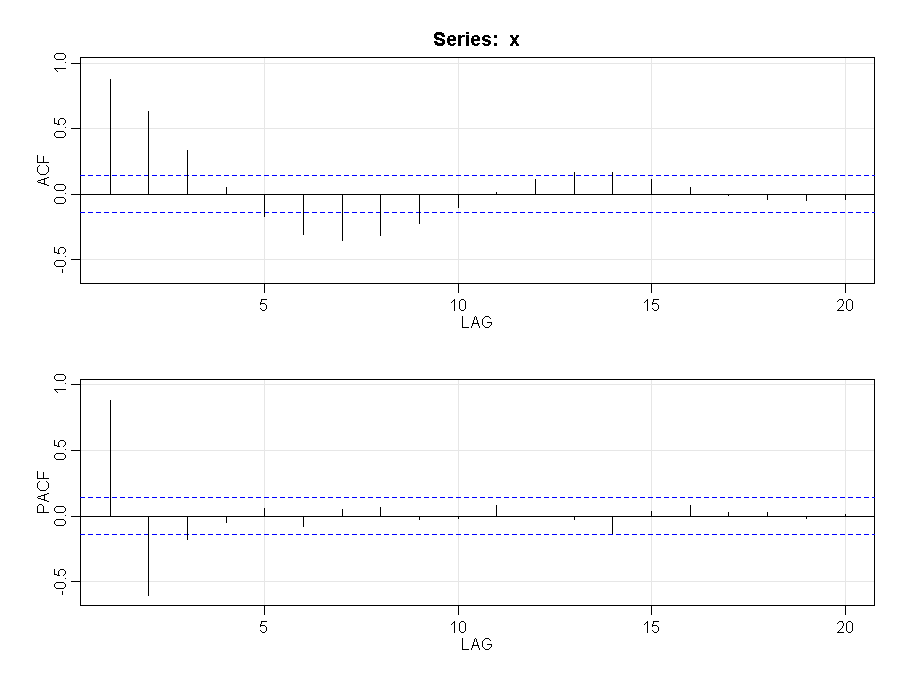
AR(3) was used.

G. Using the parametric estimate of the spectral density, estimate the approximate frequency at which the maximum spectral density occurs.

The frequency is: 0.066132265

The period associated with this cycle = 1/0.066132265 = 15.12

H. (review of ARIMA work) Determine the acf and pacf of the data series. Give the graphs and write a brief discussion of what model is suggested. You might load the astsa library and then acf2(..)



We see from the plots that:

* PACF has peaks at lags 1 and 2 and then cuts off. This points to an AR(2) model. It is a stretch but the lag 3 PACF value is much less than the lag 2 value. Nonetheless the lag 3 value is significant so we can also compare the AR(3) model with the AR(2) model
* The ACF shows a sinusoidal trailing behavior. This is in alignment with an AR(2) model

I. Use the sarima command to examine whether an AR(2) or an AR(3) is the better model for these data.

|  |  |
| --- | --- |
| AR(2) | AR(3) |
| sigma^2 estimated as 0.885  $AIC  [1] 0.9077873  $AICc  [1] 0.9188129  $BIC  [1] -0.04273797 | sigma^2 estimated as 0.8709  $AIC  [1] 0.9017207  $AICc  [1] 0.9132671  $BIC  [1] -0.03231294 |
|  |  |

From the above we can see that both the models perform equally well. While the AIC/BIC values are better for AR(3), the values have a negligible difference only.

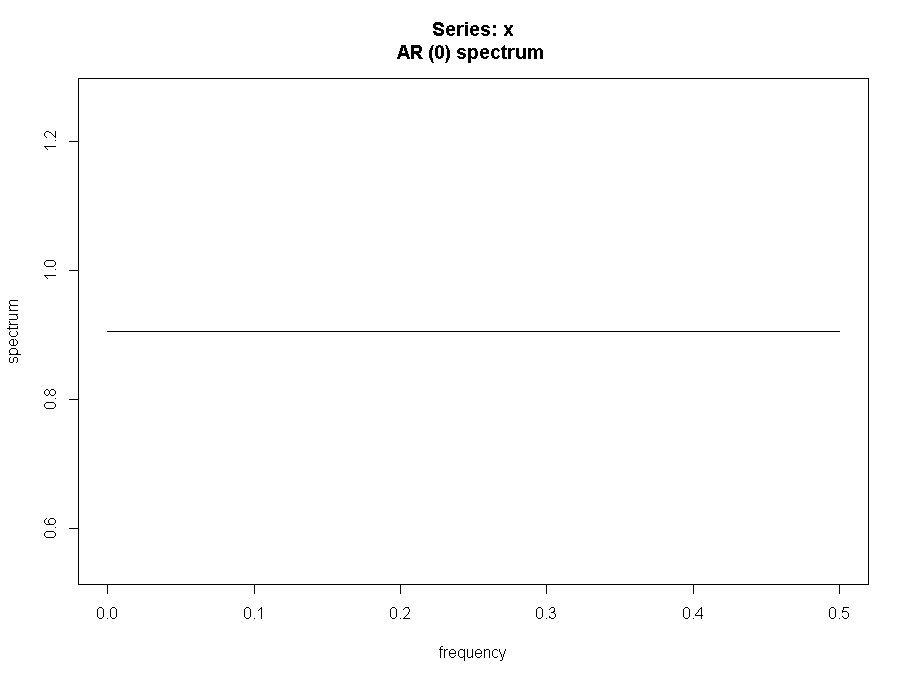
3. In R, use the command

x = rnorm(250,0,1)

This will generate a series of 250 random values from a normal distribution with mean 0 and variance 1. In other words, we’ve generated a white noise series.

Use spec.ar to determine a parametric estimate of the spectral density for the simulated data.

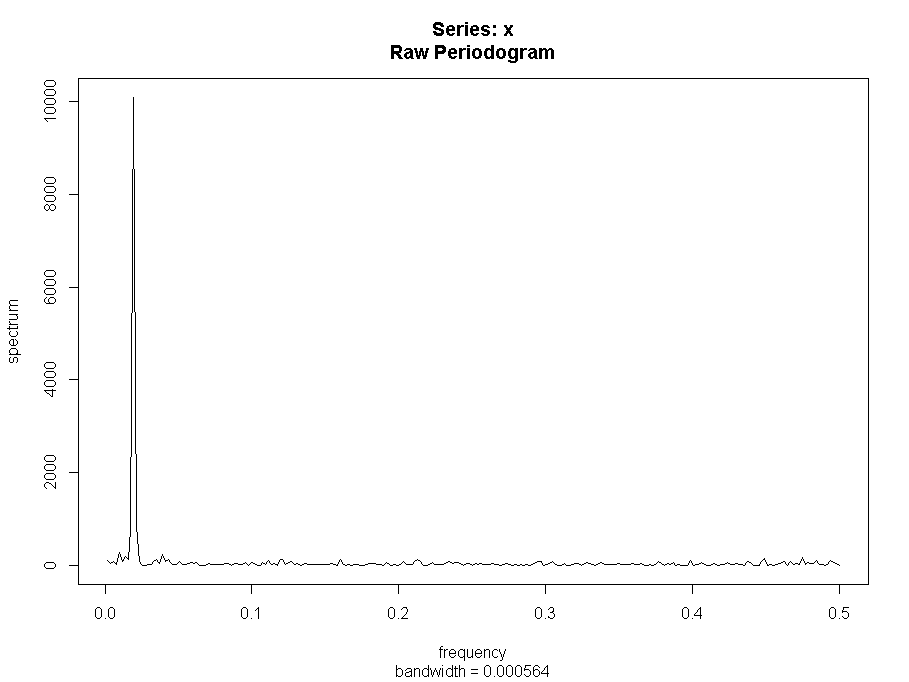
Give the graph and write a (very) brief description of its appearance.



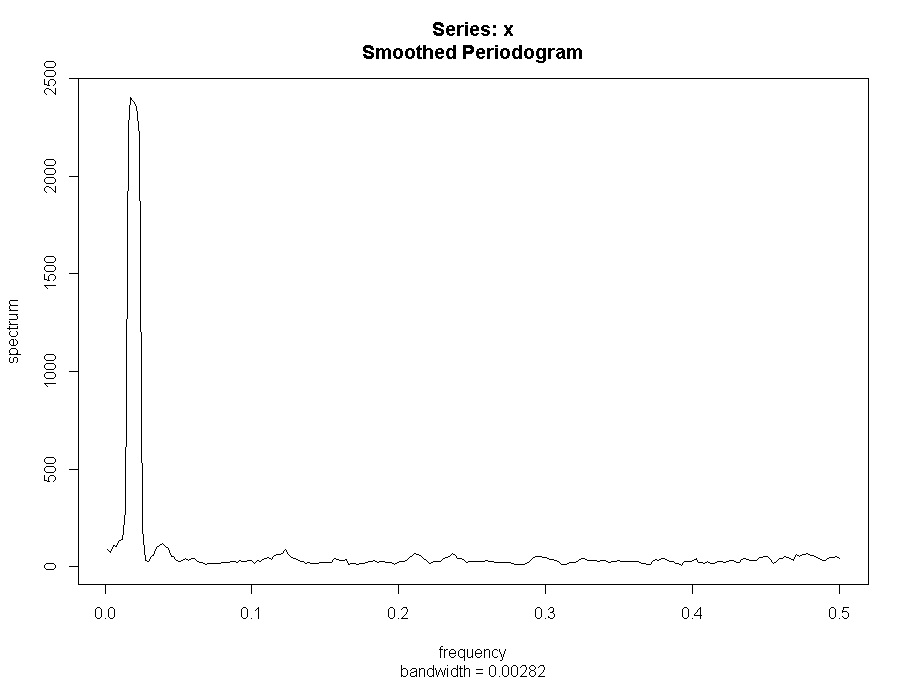
The graph above as expected shows that we are dealing with white noise with no autocorrelation - all autocorrelations are non-significant and so the series is random (white noise)

4. Use the dataset temp.dat in the Week 12 folder. It gives n = 508 weekly values of the average temperature in Los Angeles County. Determine a nonparametric estimate of the spectral density of the series. Give the graph, describe what smoothing you used and describe the location of the peak density.

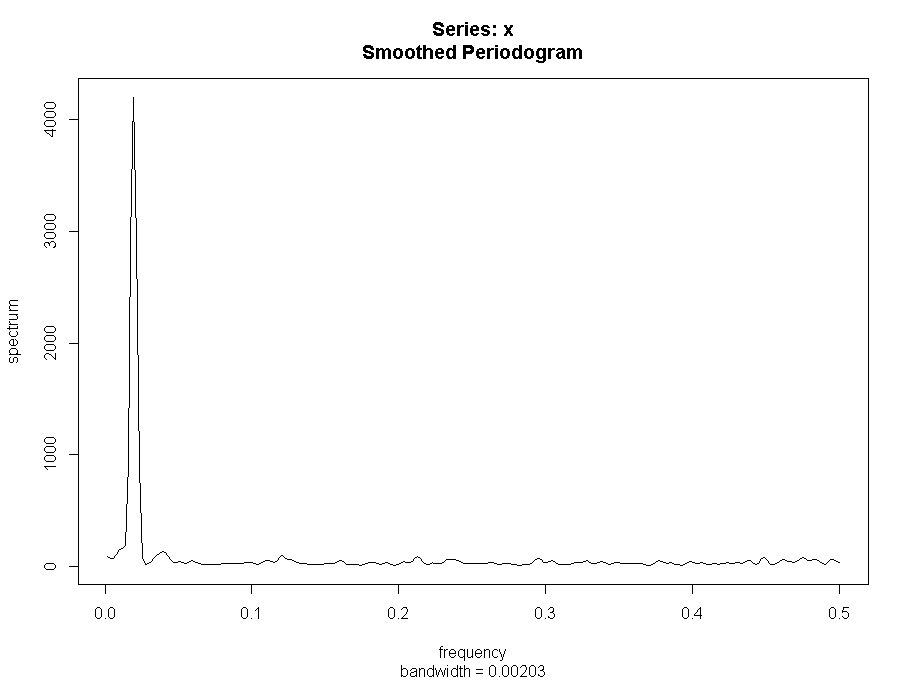
The raw periodogram:



The next plot is a smoothed periodogram using a Daniell kernel with m = 2



The next plot is a smoothed periodogram using two passes of a Daniell kernel with m = 2 on each pass



The peak density is at frequency is: 0.019531250

The period associated with this cycle = 1/ 0.019531250 = 51.2