Time Series Analysis

Homework #6

1) Data set 1 contains information on measles outbreaks in two cities in the U.K. The data set is an autocorrelated time series of evenly spaced data taken from weekly medical surveys of the populations. Data sets of this kind are fairly tractable, with several well-established techniques to examine the behaviors contained in these evenly spaced time series.

    a. Plot the time series data, showing each city on a separate plot. Are these time series stationary? Why or why not?

    b. An epidemiologist thinks they notice a pattern in how often the outbreaks occur. Test their theory by plotting the autocorrelation function (ACF) for at least two of the U.K. cities. Which lag time (in weeks) shows the lowest magnitude in autocorrelation? How do you interpret this in terms of the outbreaks? Examining a wider range of lag times may reveal long-term periodicity; do you notice any patterns in the outbreaks?

    c. A lag *k* scatter plot can be a good visualization tool for time series autocorrelations. For each of the two cities, plot a lag *k* scatter plot for two relevant lag times: the lag time showing the least autocorrelation, and the longer lag time revealing the long-term periodicity. What differences do you notice for the low-autocorrelation lag time and the longer lag time?

2) Many common techniques for time series data, such as Fourier analysis, rely on even sampling of the data. However, evenly spaced data can be a luxury that astronomers do not have. Data set 2 contains radial velocity data that you measured for a star, hoping to find some periodicity corresponding to the presence of a planet. Unfortunately, you were only allotted a few nights of observing time here and there over the course of two years.

    a. Create a Lomb-Scargle periodogram for the unevenly spaced radial velocity data. The "lomb" package for R should prove useful. Can you conclude if there is a planet orbiting your star? If so, what is the orbital period of this newly discovered planet?

    Do you see more than one statistically significant (choose your own significance value α) feature in the periodogram? Discuss why these features might appear.

    b. You are thrilled to hear that you are granted more observing time. Data set 3 contains the same type of radial velocity data from the same instrument (so you can assume the instrument noise has not changed). Combine data sets 2 and 3, produce a new Lomb-Scargle periodogram, and determine whether your initial conclusions appear correct in light of your new data.