Homework 6

Jeffrey Yau

January 1, 2016

Instructions:

The weekly assignment serves two purposes: (1) Extend the materials taught in the asynchronized materials; some new concepts or techniques are introduced in the weekly assignment. (2) Ensure that you have learned the concepts, techniques, theories, statistical models covered in a specific week. Below are some guidelines:

- Submit 2 files. Missing one of the two files will result in a 50% reduction in grade.
  1. A report (in pdf format) detailing your answers and all the steps to arrive at your answers
  2. A well-documented R-script, jupyter notebook, or Rmd file detailing all of the codes used to arrive at your answers.
- Late submission will not receive any credit.
- All the steps used to arrive at your final answers need to be shown clearly. These steps are as important as the final answer.
- The final answer of each question needs to be very easy identified; the use of bold fonts, highlights, or circling will help.
- This is a group project. Form a group with 3 or 4 people.
- Although this is a group project, we encourage you to attempt all of the exercises before discussing with your teammates. Do not use the "division-of-labor" approach. Each of the students in a group is expected to make sufficient contribution to the lab. If any of your teammate does not make sufficient contribution, please contact your instructor.
- DO NOT copy and paste or even leverage on the solutions we gave to the students in previous semesters. Violation will be reported to the Director of the MIDS program and the Office that oversees UC Berkeley Academic Integrity. In any case, the lab has various subtle changes that make those answers not directly applicable.

Exercise 1:

a. Discuss the mean and variance functions and how the similarities and differences from those we studied in classical linear model

b. Define strict and weak statonarity

Exercise 2:

a. Generate a zero-drift random walk model using 500 simulation

b. Provide the descriptive statistics of the simulated realizations. The descriptive statistics should include the mean, standard deviation, 25th, 50th, and 75th quantiles, minimum, and maximum

c. Plot the time-series plot of the simulated realizations

d. Plot the autocorrelation graph

e. Plot the partial autocorrelation graph

Exercise 3:

a. Generate arandom walk with drift model using 500 simulation, with the drift = 0.5

b. Provide the descriptive statistics of the simulated realizations. The descriptive statistics should include the mean, standard deviation, 25th, 50th, and 75th quantiles, minimum, and maximum

c. Plot the time-series plot of the simulated realizations

d. Plot the autocorrelation graph

e. Plot the partial autocorrelation graph

Exercise 4:

Use the series from INJCJC.csv

a. Load the data and examine the basic structure of the data using str(), dim(), head(), and tail() functions

b. Convert the variables INJCJC into a time series object frequency=52, start=c(1990,1,1), end=c(2014,11,28). Examine the converted data series

c. Define a variable using the command INJCJC.time<-time(INJCJC)

d. Using the following command to examine the first 10 rows of the data. Change the parameter to examine different number of rows of data

head(cbind(INJCJC.time, INJCJC),10)

e1. Plot the time series plot of INJCJC. Remember that the graph must be well labelled.

e2. Plot the histogram of INJCJC. What is shown and not shown in a histogram? How do you decide the number of bins used?

e3. Plot the autocorrelation graph of INJCJC series

e4. Plot the partial autocorrelation graph of INJCJC series e5. Plot a 3x3 Scatterplot Matrix of correlation against lag values

f1. Generate two symmetric Moving Average Smoothers. Choose the number of moving average terms such that one of the smoothers is very smoother and the other one can trace through the dynamics of the series. Plot the smoothers and the original series in one graph.

f2. Generate two regression smoothers, one being a cubic trend regression and the other being a periodic regression. Plot the smoothers and the original series in one graph.

f3. Generate kernel smoothers. Choose the smoothing parametrs such that one of the smoothers is very smoother and the other one can trace through the dynamics of the series. Plot the smoothers and the original series in one graph.

f4. Generate two nearest neighborhood smoothers. Choose the smoothing parametrs such that one of the smoothers is very smoother and the other one can trace through the dynamics of the series. Plot the smoothers and the original series in one graph.

f5. Generate two LOWESS smoothers. Choose the smoothing parametrs such that one of the smoothers is very smoother and the other one can trace through the dynamics of the series. Plot the smoothers and the original series in one graph.

f5. Generate two spline smoothers. Choose the smoothing parametrs such that one of the smoothers is very smoother and the other one can trace through the dynamics of the series. Plot the smoothers and the original series in one graph.