Homework 6 - Week 8

Jeffrey Yau and Devesh Tiwari

October 15, 2016

Instructions:

The weekly assignment serves two purposes: (1) Review concepts, techniques, theories, statistical models covered during the week. (2) Extend the materials taught in the asynchronized lectures, assigned readings, and live sessions; some new concepts and/or techniques are introduced in the weekly assignment.

Below are specific instructions:

- **Due: 10/30/2016 (11:59pm PST)**
- Late submission will not be accepted.
- You may complete this assignment on your own or in a group of no more than 3 students.
- When working in a group, you are strongly encouraged to complete the assignment on your own before discussing your group mates. Do not use the "division-of-labor" approach to complete the assignment.
- The homework is designed as a quantitative analysis. The instructions and questions are designed to guide you through the analysis of data using regression techniques. As such, you should think of it as a quantitative case study and the result of the study is a report with a set of well-written codes that can be used to reproduce the results in the report.
- Submission:
  - Submit your own assignment via ISVC. Email submission will not be accepted and graded.
  - Submit 2 files:
    1. R-script or R markdown file
    2. A pdf file including the summary, the details of your analysis, and all the R codes used to produce the analysis
  - Fail to do so will automatically receive a 50% reduction in grade
  - Each group only needs to submit one set of files
  - Use the following file naming convensation; fail to do so will receive 10% reduction in the grade:
    - **SectionNumber_hw06_LastNameFirstInitial.fileExtension**
    - Examples:
      - Section1_hw06_YauJ.Rmd
      - Section1_hw06_YauJ.pdf
      - Section1_hw06_TiwariD_YauJ.Rmd
      - Section1_hw06_TiwariD_YauJ.pdf
  - Not following the naming conversation will receive a 20% reduction in grade.
- 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 homework has various subtle changes that make those answers not directly applicable.

Exercise 1:

a. Discuss the mean and variance functions and how do they differ from those we studied in classical linear models.

b. Define strict and weak statonarity

Exercise 2:

a. Generate a zero-drift random walk model using 100 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

f. As always, discuss your graphical presentations of the data

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

f. As always, discuss your graphical presentations of the data

Exercise 4:

Use the series from INJCJC4.csv

a. Load the data and examine the basic structure of the data

b. Convert the variables INJCJC4 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 INJCJC4.time<-time(INJCJC4)

d. Examine the first 10 rows of the data.

e1. Plot the time series plot of INJCJC4. Remember that the graph must be well labelled.

e2. Plot the histogram of INJCJC4. What is shown and not shown in a histogram? How do you decide the number of bins used?

e3. Plot the autocorrelation graph of INJCJC4 series

e4. Plot the partial autocorrelation graph of INJCJC4 series

e5. Plot a 3x3 Scatterplot Matrix of correlation against lag values. Make sure your plot is legible.

f. As always, discuss your graphical presentations of the data