CSC425 – Time series analysis and forecasting
Homework 2

Submission instructions

Submit the homework at the Course Web page http:d2l.depaul.edu in the Dropbox page. Keep a copy of all your submissions!

1. Submit your answers in a **single pdf document**. Make sure to explain in detail your analyses, and include a short snippet of the code for the problem, relevant output and graphs.
2. If you have questions about the homework, email me **well before** the deadline.
3. Late assignments will have a 25% deduction.
4. Assignments submitted five days after the deadline will not be accepted.

Final Project Milestone

During week 3 we will be finalizing group formation, so make sure that you have posted to the final project dataset and project ideas forum, and then if you can settle on a set of group members, post the group to Milestone 2, the first group milestone, which is to list the group members and then have every member of your group reply to confirm membership. Note that any member that does not reply risks being dropped from the group and replaced with another class member.

Problem 1 [10 pts]
Continue your analysis of the weekly sales data for **ToothPaste**. Remember, it is in dataset "groceries.csv" which contains weekly units sold for three grocery items: ToothPaste (100ml container of toothpaste), PeanutButter (340g. jat of crunchy peanut butter), and Biscuits (200g., 10 finger package of shortbread cookies). The dataset contains the variable Date defined as the first day of the week for the sales period. Answer the following questions

a) Graph the lag plot of the ToothPaste series and its first lag. Interpret the graph. What does it tell you about autocorrelation for this series?
b) Compute and plot the first 15 lags of the autocorrelation function for ToothPaste weekly sales and discuss if the series shows evidence of serial correlation.
c) Use the Ljung Box test to hypothesis that ToothPaste weekly sales have significant serial correlation.
d) Discuss the importance of weak stationarity for time series analysis and describe a method to analyze whether a TS is stationary.

Problem 2 [10 pts]
Analyze the long-term behavior of the Intel stock price by investigating its autocorrelation behavior.

a) Graph the lag plot of the log-prices series and its first lag. Interpret the graph. What does it tell you about autocorrelation for this series?
b) Compute and plot the first 15 lags of the autocorrelation function of the series of log-prices and discuss if the series shows evidence of serial correlation

c) Explain the pattern that you see in the first 10 lags of the autocorrelation function. Does it appear that the series is stationary? Give a reason why or why not.
d) Compute and plot the first 15 lags of the ACF of the log returns (i.e. the difference of the logs). Interpret the result.
e) Use the Ljung Box test to hypothesis that the log-returns have significant serial correlation.

Problem 3 [10 pts]

Consider the following AR(1) time series process: $r_t = .9r_{t-1} + a_t$, where is a Gaussian (i.e. normally distributed) white noise series with zero mean and constant variance σ^2=0.5.

a) What is the mean of the time series r_t?
b) Determine if the AR(1) model is stationary. Explain.
c) Determine the overall variance of the series.

Problem 4 [10 pts]

Consider the following MA(1) time series process: $X_t = 5 + a_t - 0.5a_{t-1}$, where $\{a_t\}$ is a Gaussian white noise series with mean zero and constant variance σ^2=0.025.

a) What is the mean of the time series X_t?
b) What is the variance of the series?
c) Discuss if the MA (1) model is stationary. Explain.

Problem 5 [20 pts]

The Institute for Supply Management (ISM) (http://www.ism.ws/) is responsible for maintaining the Purchasing Managers Index (PMI). The index is derived from monthly surveys of private sector companies. Monthly data from 01/01/1980 to 12/01/2015 are saved in the NAPM.csv file. This problem asks you to apply an AR(p) autoregressive model to describe the dynamic behavior of the index.

a) Import the data and create a time series object for index using the ts() function where the starting date is first month of 1980.
b) Create a time plot of the data and determine whether the series is multiplicative or additive. If it is multiplicative, compute the log of the series before proceeding to the next parts of this question.
c) Create a decomposition of the series and analyze the series for trends and seasonality. Describe what you find in detail.
d) Analyze if the time series is serially correlated using the ACF plot and the Ljung Box test.
e) Fit an AR(2) model with the Arima function as described in class and write down the expression of the estimated AR(p) model. Explain the meanings of the coefficients in the model.
f) Examine the the significance of the model coefficients using the "coeftest" function from the "lmtest" package. Discuss which coefficients are significantly different from zero.

Problem 6 [10 pts]

Continue your investigation of the ToothPaste series.

a) From the ACF plot that you computed in problem 1, determine a degree for an MA model for the series. Explain your choice.

b) Fit an MA(#) model with for that degree using the Arima function as described in class and write down the expression of the estimated AR(p) model. Explain the meanings of the coefficients in the model.
c) Examine the the significance of the model coefficients using the "coeftest" function from the "lmtest" package. Discuss which coefficients are significantly different from zero.

"Reflection" Problem [5 pts]

Post a message on the discussion board reflecting on the topics in week 2 and on homework 2. Indicate the assignment in this module you found to be the easiest, the one you found to be the hardest, and why. A new thread called "Homework 2 Reflection" will be created.