Midterm

FE542 Time Series in Finance

due by 11:55pm, Sunday April 1, 2018

This is a take home examination. Please write the report in a nice format and include all tables where needed to explain your conclusions in the text of the report. Submitting commented copied code directly from R and not formatted properly using R markdown or latex or word will earn you 0 for this exam. Please convert to a PDF and submit the report together with a zip file containing the code.

1. (50 points) Solve problems 2.1 and 2.2 on page 104 of the textbook (Tsay's *Financial Time Series* 3rd edition).

2. (100 points) For this problem download data for any equity you like. Please download data a.s.a.p. after you see this. Choose an equity that has a history of at least 20 years. Download daily data for the equity. A great concern for time series estimation is that one needs a large number of observations from the time series for a reliable estimation and forecast. On the other hand the data has to be stationary for a reliable estimation and in general it is believed that data stays stationary for only short periods of time.

   Here we will study the performance of various models and the effect of the data length on the estimation. To do this create new data vectors. Data 1 contains the last 6 months of the equity return. Data 2 contains the last 12 months. Data 3 the last 24 months. Data 4 the last 30 months. Data 5 the last 60 months. Data 6 the last 120 months and finally Data 7 the last 240 months.

   Repeat each of the following for all the data sets.

(a) Work with continuously compounded returns. Pay attention to the proper date order in the data you downloaded.

(b) Study unit root non-stationarity. Compare for all data periods.

(c) Build the best AR, the best MA and the best ARMA model after you remove non-stationarity (if needed) for each time period.

(d) Build a seasonal model for your data using whatever lag you feel approximates the data best.

(e) Using each of the models that you constructed forecast the next five observations. Pay attention to the holidays (i.e., you can't forecast Saturdays and Sundays return). You should have 4 models for each dataset. Do the best models in each category (AR, MA, etc.) depend to a particular time period length?

(f) 5 working days **after** the date you downloaded your equity data, go ahead and download the same equity data again. Test the performance of each model by comparing the forecasted returns (and prices) with the real observations collected after 5 working days. Calculate the squared error of your forecast for each model and the absolute value of your forecast. What was the better model? Alternately you could set aside the last 5 days of your data nad only use those for the evaluation of the forecast.

(g) Was there a difference in the time periods (i.e., is there a type of model that consistently outperformed the others for all the datasets under consideration)?

(h) Are the perceived differences between models due to chance or there is a model that performed significantly better? (Hint: Use the standard error of the forecast as given by the R output.)

3. (100 points) Please use the high frequency data provided in *"INTC-data.xls"* (minute data for Intel Corporation). The data is from the period June 19-June 30 2006. The data columns contain: Ticker, time stamp, Open, High, Low, Close, Volume, Nr. of trades, Average price all within the minute. You may use whichever column you wish as data input. These are two weeks (notice that there are only 10 days worth of data due to the weekend).

   (a) Divide the data into two parts: one (larger) part would be used to construct the models, and the later part would be used to evaluate

the model performance. It is up to you how to divide the data to evaluate the performance of the models but you need to decide on a model which you consider to be the best.

(b) Construct the best model you can with the material taught until now (and including seasonal time series). You should discuss a lot. Pay attention to the fact that the data contains a weekend.

(c) After constructing (what you consider to be) the best model, obtain the residuals and test if they are likely to be white noise. Use (and name) all the testing procedures for time series you learned for this purpose.

4. (50 points) For this problem you will study the change in behavior of random walk with parameters. Please decide on a number of observations to be generated which is large enough to actually observe a behavior. The problem looks at two random walks:

   - Regular random walk $r_t = r_{t-1} + a_t$, where $r_0 = 0$
   - Random walk with drift μ: $r_t = \mu + r_{t-1} + a_t$, where $r_0 = 0$

   Please include representative images and comment on the observed behavior. Please note that one model is a subset of the other one. The idea is to study the effect of changing parameters in the model and understand the change in behavior of the process as a result of changing these parameters. You have to come up with representative images and numerical measures that will allow you to illustrated what you perceive.

   (a) Study the effect of changing μ. To this end, experiment with $\mu = -1, 0, 1, 100$

   (b) Study the effect of changing the distribution of the white noise. Experiment with an i.i.d. noise distributed as σZ where Z is a standard normal random variable or a t distributed noise with say 5 degrees of freedom.

   (c) Study the effect of changing σ. Look at $\sigma = 0.1, 0.6, 1, 10$.