Take-home Final

(Due on Wednesday, June 8)

Statistics 137
Spring Quarter, 2016

Please note that the take-home part is 20% of the final examination. This is due at the time of the Final Exam on June 8. You may work in a group (max group size=3) of registered students in the course. Only one report per group needs to be submitted. Please write down the names of the students in the group on submitted work.

[Data source: http://www.eia.gov/totalenergy/data/monthly/#consumption.]
Please find attached a data set on petroleum consumed (in trillion BTU) by the residential sector in the US (Jan, 1984 –Dec, 2015). Analyze the data using time series methods. You are free to use any number of methods that we have covered in this course. You may consider the following points in your analysis (with appropriate comments/explanations at each step):

- Explain the data, why it is a time series, why it is important to analyze it.
- Use graphical techniques to understand the nature of variation in the data.
- Determine if the series is stationary or not. You may need to transform, estimate the trend and the seasonal effects in order to carry out the analysis.
- Obtain the appropriate ACF, PACF plots and periodogram (and its smoothed version), and use these to make a preliminary identification of a time series model.
- Fit an ARIMA model obtained via preliminary identification, and examine the residuals and their properties.
- Select the final model using a model selection criterion such as AICC. [If you fit ARIMA(p,d,q), it may be enough to consider the 81 models with p=0,…,8 and q=0,…,8, where p is the AR order and q is the MA order. The R function auto.arima ('forecast package') may be used.]
- Plot the spectral density of the final model as well as the smoothed periodogram on the same graph.
- Perform the necessary diagnostics on the residuals of the final model including the ACF, PACF plots as well as the smoothed periodogram.
- Write down the final model, the estimated parameters and the standard errors.
- Refit the final model (i.e., use AR and MA orders of the final model, but not the parameter estimates) using all the data except for the year 2015 and use this model to forecast petroleum consumption for the 12 months in 2015. Plot the available observed and the forecasted values against time (12 months of 2015) on the same graph. [If you need to extrapolate the trend, often a linear extrapolation is reasonable.]

- Summarize your findings from the analysis and explain your conclusion. If you feel the analysis done by you can be improved, please provide a brief explanation.

Your report may include the following Sections:

- Introduction: Statement of the problem.
- Materials and Methods: Description of the data and the methods used in the analysis.
- Results: Explanation of the results of your analyses. You can cut and paste the relevant parts of your computer outputs and refer to them in explaining your results.
- Conclusion and Discussion: Highlight the main points and discuss them.