Statistical Methods for Discrete Response, Time Series, and Panel Data (W271): Lab 3

K.C. Tobin, Weixing Sun, Winston Lin

August 6, 2017

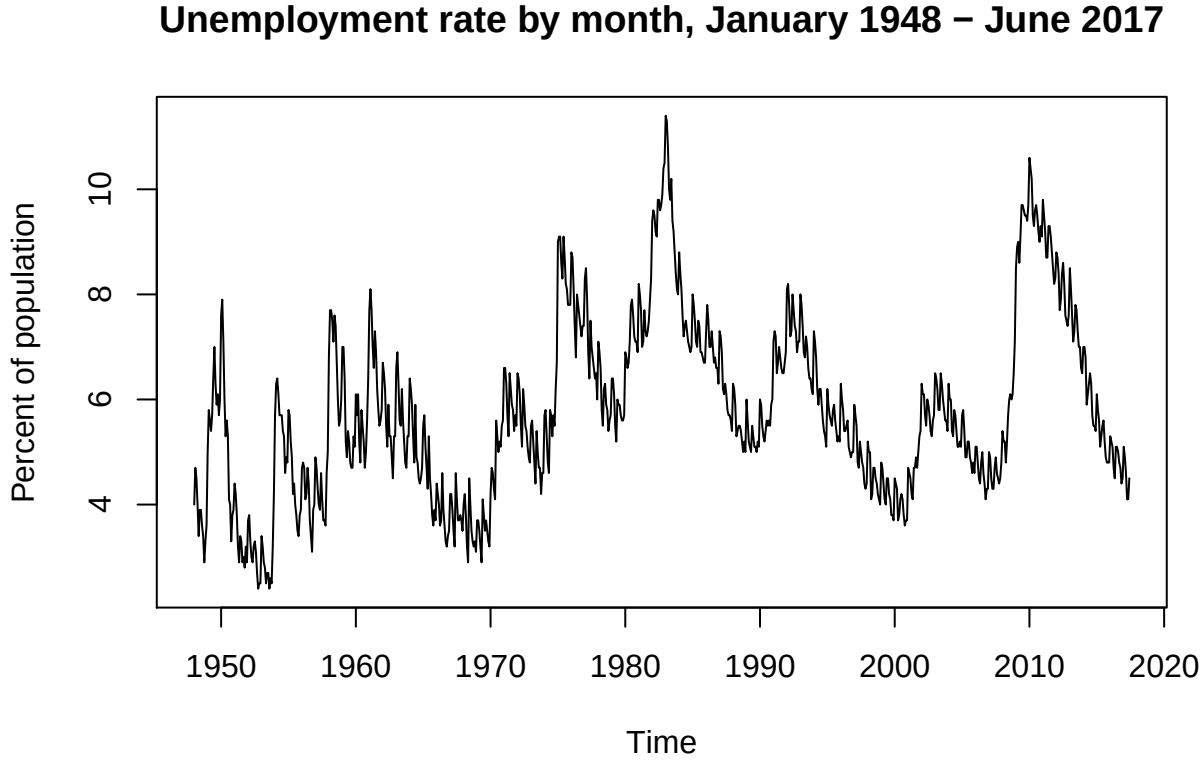
Question 1: EDA

During your EDA, you notice that your data exhibits both seasonality (different months ahve different heights) AND that there is a clear linear trend. How many order of non-seasonal and seasonal differencing would it take to make this time-series stationary in the mean? Why?

```
# EDA
unemp <- read.csv("UNRATENSA.csv")
summary(unemp)
```

```
##          DATE       UNRATENSA
##  1948-01-01:  1   Min.   : 2.400
##  1948-02-01:  1   1st Qu.: 4.700
##  1948-03-01:  1   Median : 5.600
##  1948-04-01:  1   Mean   : 5.801
##  1948-05-01:  1   3rd Qu.: 6.900
##  1948-06-01:  1   Max.   :11.400
##  (Other)   :828
```

```
unemp.ts = ts(unemp, start = c(1948, 1), frequency = 12)
plot.ts(unemp.ts[,2], main = 'Unemployment rate by month, January 1948 - June 2017', ylab = 'Percent of
```



Question 2: SARIMA

It is Dec. 31, 2016, and you work for a non-partisan think tank focusing on the state of the U.S. economy. You are interested in forecasting the unemployment rate through 2017 (and then 2020) to use it as a benchmark against the incoming administration's economic performance. Use the dataset UNRATENSA.csv and answer the following:

a. Build a SARIMA model using the unemployment data and produce a 1-year forecast and then a 4-year forecast. Because it is Dec. 31, 2016, leave out 2016 as your test data.

- How well does your model predict the unemployment rate up until June 2017?
- What does the unemployment rate look like at the end of 2020? How credible is this estimate?

b. Build a linear time regression and incorporate seasonal effects. Be sure to evaluate the residuals and assess this model on the bases of the assumptions of the classical linear model, and the produce a 1-year and 4-year forecast.

- How well does your model predict the unemployment rate up until June 2017?
- What does the unemployment rate look like at the end of 2020? How credible is this estimate?
- Compare this forecast to the one produced by the SARIMA model. What do you notice?

Question 3: VAR.

You also have data on automotive car sales.

- Use a VAR model to produce a 1-year forecast on both the unemployment rate and automotive sales for 2017 in the U.S.
- Compare the 1-year forecast for unemployment produced by the VAR and SARIMA models, examining both the accuracy AND variance of the forecast. Do you think the addition of the automotive sales data helps? Why or why not?