CIS8695

Homework II

What to Turn In:

Please write your answers and paste required results/reports in **a MS Word file**. **Note: Please do NOT submit any other file format, because it will cause grading inconveniency. Failing to submit in correct file format will cause the loss of homework grades! You can use screenshots if it is convenient for you.**

Problems

Homework problems are all from the textbook. For your convenience, the pictures of the problems are also attached in the document at the end. Datasets are provided on iCollege.

Part 1. Look at Problem 17.6 in Chapter 17 of the Textbook. A picture of the problem is attached below. Complete (b), (c) and (d) of Problem 17.6.

Part 2. Use the same data as in the above Part 2. Use a **Holt-Winter's exponential smoothing** approach to make prediction.

- Fit a quadratic trend
- Fit an AR model to the residuals
- Fit a quadratic trend model to Sales (with Quarter and Quarter^2)

17.6 Souvenir Sales. Figure 17.17 shows a time plot of monthly sales for a souvenir shop at a beach resort town in Queensland, Australia, between 1995 and 2001 (Data are available in *SouvenirSales.csv*, source: Hyndman, R.J., Time Series Data Library, http://data.is/TSDLdemo. Accessed on 07/25/15.). The series is presented twice, in Australian dollars and in log-scale. Back in 2001, the store wanted to use the data to forecast sales for the next 12 months (year 2002). They hired an analyst to generate forecasts. The analyst first partitioned the data into training and validation sets, with the validation set containing the last 12 months of data (year 2001). She then fit a regression model to sales, using the training set.

a. Based on the two time plots, which predictors should be included in the regression model? What is the total number of predictors in the model?

b. Run a regression model with Sales (in Australian dollars) as the outcome variable, and with a linear trend and monthly seasonality. Remember to fit only the training data. Call this model A.

i. Examine the estimated coefficients: which month tends to have the highest average sales during the year? Why is this reasonable?

ii. The estimated trend coefficient in model A is 245.36. What does this mean?

c. Run a regression model with an exponential trend and multiplicative seasonality. Remember to fit only the training data. Call this model B.

i. Fitting a model to log(Sales) with a linear trend is equivalent to fitting a model to Sales (in dollars) with what type of trend?

ii. The estimated trend coefficient in model B is 0.02. What does this mean?

iii. Use this model to forecast the sales in February 2002.

d. Compare the two regression models (A and B) in terms of forecast performance. Which model is preferable for forecasting? Mention at least two reasons based on the information in the outputs.

e. Continuing with model B, create an ACF plot until lag 15 for the forecast errors. Now fit an AR model with lag 2 [ARIMA(2,0,0)] to the forecast errors.

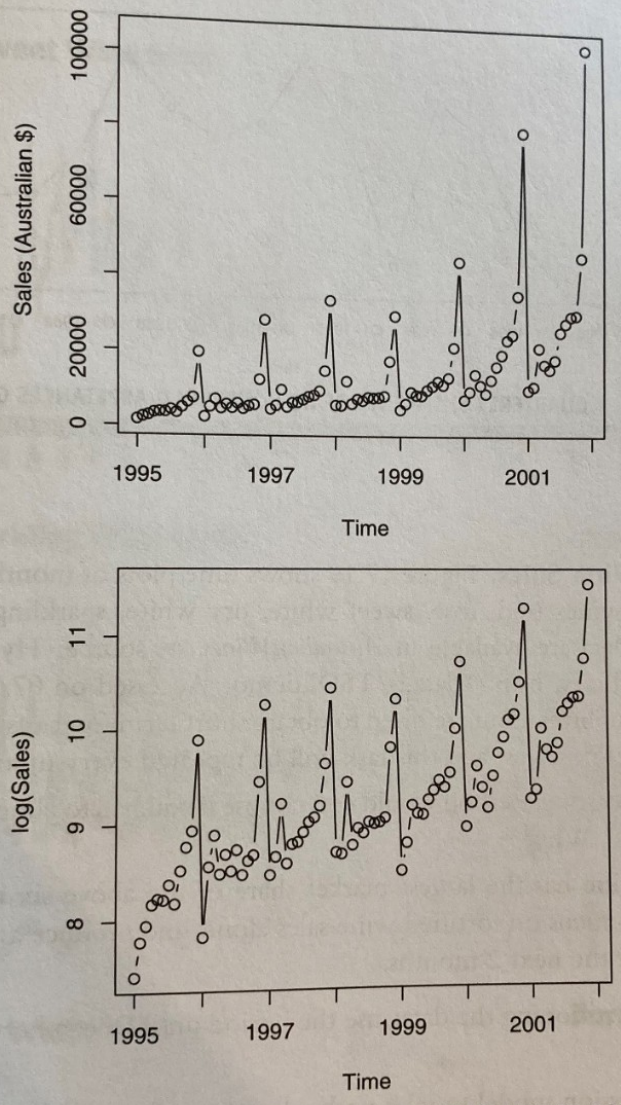

FIGURE 17.17 MONTHLY SALES AT AUSTRALIAN SOUVENIR SHOP IN DOLLARS (TOP) AND IN LOG-SCALE (BOTTOM)

i. Examining the ACF plot and the estimated coefficients of the AR(2) model (and their statistical significance), what can we learn about the forecasts that result from model B?

[illegible] information to compute an improved forecast for Jan-