**Stat-440/640 Regression and Time Series Analysis**
**Fall 2018**

Final Exam

Name _____

show work to ensure full credit and attach all the necessary printouts, tables, etc

1. (50 pts) Consider the time series data (motel.txt) (Total room nights occupied at hotels, motels, and guest houses in Victoria, Australia. Jan 1880 - Jun 1995.) Use R to answer the following questions:

   (a) Plot the time series data. Does the series exhibits constant seasonal variation?

   (b) When a time series displays increasing seasonal variation, it is common practice to apply a transformation to the data in order to produce a transformed series that displays constant seasonal variation. Make an appropriate transformation to the given series and plot the transformed. Does the plot exhibits constant seasonal variation?

   (c) Modeling seasonal variation: We consider the regression model for the transformed series.

$$\begin{aligned} y_t^\star &= Tr_t + Sn_t + \epsilon_t \\ &= \beta_0 + \beta_1 t + \beta_2 M_2 + \beta_3 M_3 + \cdots + \beta_{12} M_{12} + \epsilon_t \end{aligned}$$

   where $M_2, \cdots M_{12}$ are seasonal dummy variables. For example,

$$M_2 = \begin{cases} 1 & \text{if period } t \text{ is February} \\ 0 & \text{otherwise} \end{cases}$$

   Estimate the regression coefficients of the above model.

   Compute the point forecast and 95% prediction interval of $y$ for Jul 1995, Aug 1995 and Sep 1995.

(d) Handling first-order autocorrelation. Find the first-order autocorrelation and first-order Durbin-Watson test statistics by using both formula and r code. Test whether the error term in the model in part (c) are positively autocorrelated.

(e) When the error terms are autocorrelated, the following model is considered

$$\begin{aligned} y_t^\star &= Tr_t + Sn_t + \epsilon_t \\ &= \beta_0 + \beta_1 t + \beta_2 M_2 + \beta_3 M_3 + \cdots + \beta_{12} M_{12} + \epsilon_t \end{aligned}$$

where

$$\epsilon_t = \phi \epsilon_{t-1} + a_t$$

Here $a_t$ is a random shock with mean 0 that satisfies the constant variance, independence, and normality assumptions. ($a_t$ is often called a white noise) Use ARIMA in r to obtain point estimates $b_0, b_1, b_2, \cdots, b_k$, and $\hat{\phi}_t$ of the model parameters $\beta_0, \beta_1, \beta_2, \cdots, \beta_k$, and $\phi_t$, use the formula given in pages 310-311 to obtain the point forecast and 95% confidence interval for Jul 1995.

2. (35 pts) The data in Table 7.5 give quarterly sales of the popular game Oligopoly at the J-Mart variety store. Consider using the multiplicative decomposition method to forecast Oligopoly for year 4. Use Excel spreadsheet to do the following:

   (a) Compute appropriate four-period moving averages for these data.

   (b) Compute centered moving averages for these data.

   (c) Calculate $sn_t \times ir_t$ values for these data.

   (d) Calculate estimates of the seasonal factors for quarterly sales (i.e compute $sn_t$ values for these data).

   (e) Compute deseasonalized observation for these data.

   (f) Plot the deseasonalized observation versus time. From your data plot, what kind of trend appears to exist?

   (g) Assuming that a linear trend $Tr_t = \beta_0 + \beta_1 t$ describes the deseaonalized observations, compute the least squares estimates of $\beta_0$ and $\beta_1$.

   (h) Using estimated trend and seasonal factors, compute point forecasts of Oligopoly sales for each quarter of year 4.

3. (35 pts) Repeat problem # 2 using the additive decomposition method. Does this method seem more appropriate for these data than the multiplicative decomposition method? Justify the answer by showing graphs and SSE for both the cases.

4. (30 pts) The quarterly sales of Tiger Sports Drink for the last eight years are given in Table 8.2. (use R programming to do all following analysis)

   (a) Plot the time series data. Does the series indicate that seasonal pattern is increasing? If so multiplicative Holt - Winters might be employed to forecast future sales. Use $\alpha = 0.2$, $\beta = 0.1$, and $\gamma = 0.1$ and R to determine the initial states $l_0, b_0, sn_{-3}, sn_{-2}, sn_{-1}, sn_0$ and MSE.

   (b) Find the optimum values of $\alpha$, $\beta$, $\gamma$ and MSE. Obtain a point forecast and 95% prediction interval for the first quarter of year 9.

   (c) Now use the log transformed data of Tiger Sport data and additive Holt-Winters to answer part (a) and (b).

   (d) Which method appears to do better job? Additive or multiplicative Holt-Winters.

5. (50 pts) The data (hsales.txt) give monthly sales of new one-family houses, USA, from Jan 1987 through Dec 1995. Set up Excel spreadsheet of Holt's trend corrected exponential smoothing with $l_0$ and $b_0$ , where $l_0$ and $b_0$ are the intercept and slope of the simple linear regression based on the first 12 months of the data set.

   (a) Use this spreadsheet to find SSE, MSE, s, MAD, and MAPE when $\alpha = 0.2$ and $\beta = 0.1$

   (b) Use solver, starting with $\alpha = 0.2$ and $\beta = 0.1$ to find the values of $\alpha$ and $\beta$ that produce a minimum value of SSE.

   (c) Use $l_{108}$, $b_{108}$ and the optimal values of $\alpha$ and $\beta$ to find the point forecasts $y_{109}$, and $y_{110}$ and their 95% prediction intervals.