

THE UNIVERSITY OF TEXAS AT AUSTIN
McCombs School of Business

STA 372.5

Spring 2018

HOMEWORK #3 – Due Wednesday, February 14

1. Problem #2 on the 2017 midterm exam.
2. Problem #4 on the 2017 midterm exam.
3. The file STA372_Homework3_Question3.dat on the *Data sets* page of the Canvas class website contains 40 quarters of sales data for The Gap (in thousands of dollars). The columns in the dataset are (in order) *Time*, *Quarter* and *Sales*.

Hint: For help in writing the R script for this problem, see the R script used in class to estimate the exponential trend in Wal-Mart sales when there is a quadratic term in the exponent (i.e., see pages 30-32 of the lecture note “Deterministic Trends - Part 1”).

You seasonally adjusted The Gap’s sales in problem 1 of homework #2. You can use the same R script from this problem to plot and seasonally adjust *Sales* and $\log(\text{Sales})$ in the current problem. However, please note that this problem only uses observations for periods 1-40 in the seasonal adjustment process (not periods 1-44 as in homework #2).

- (a) What graphical evidence is there of trend and seasonality in The Gap’s sales data?
- (b) Compute and plot $\log(\text{Sales})$ vs. *Time*.
- (c) Use the *stl* package to seasonally adjust $\log(\text{Sales})$. Let $\log(A)$ represent the seasonally adjusted $\log(\text{Sales})$ values.
- (d) Plot $\log(A_t)$ vs. *Time*.
- (e) Estimate the coefficients in the model

$$\log(A_t) = \alpha + \beta_1 t + \beta_2 t^2 + \varepsilon_t$$

using the *lm* command in *R*. For notational purposes, we will let $w_t = \log(A_t)$ and use \hat{w}_t to denote the fitted values from the regression model.

- (f) Plot w_t and the fitted values \hat{w}_t on the same graph. The fitted values \hat{w}_t are the in-sample forecasts of $w_t = \log(A_t)$. Do these in-sample forecasts do a good job tracking the actual observations of $\log(A_t)$?
- (g) Plot the residuals from the regression in part (e) against *Time*. Is there evidence of autocorrelation in the plot?
- (h) Compute the autocorrelation function of the residuals from the regression in part (e). Are the residuals autocorrelated? What does this imply about information left in the residuals that will not be incorporated into the forecasts?

For the remainder of the problem, assume the residuals are independent (i.e. even though the residuals are not independent, it is not necessary for this problem to incorporate a lagged $\log(A_{t-1})$ term in the regression to account for the autocorrelation).

- (i) Compute the in-sample forecasts for The Gap's sales (not seasonally adjusted sales) for quarters 1-40 and the out-of-sample forecasts for sales for quarters 41-44.

Plot The Gap's sales for quarters 1-40 (connected by a solid line), and the in-sample forecasts for quarters 1-40 and out-of-sample forecasts for quarters 41-44 (connected by a dashed line) on the same graph. This plot is similar to the plot on page 29 of the lecture note "Deterministic Trends - Part 1".

Do the in-sample forecasts do a good job tracking sales in quarters 1-40? What does this imply about how good the out-of-sample forecasts of sales are likely to be?

- (j) Would you feel comfortable using your out-of-sample forecasts for quarters 41-44? Why or why not?
- (k) The actual values of The Gap sales in quarters 41-44 (in thousands of dollars) are 848688, 868514, 1.16E+06 and 1.52E+06. Did your model do a good job predicting these values?