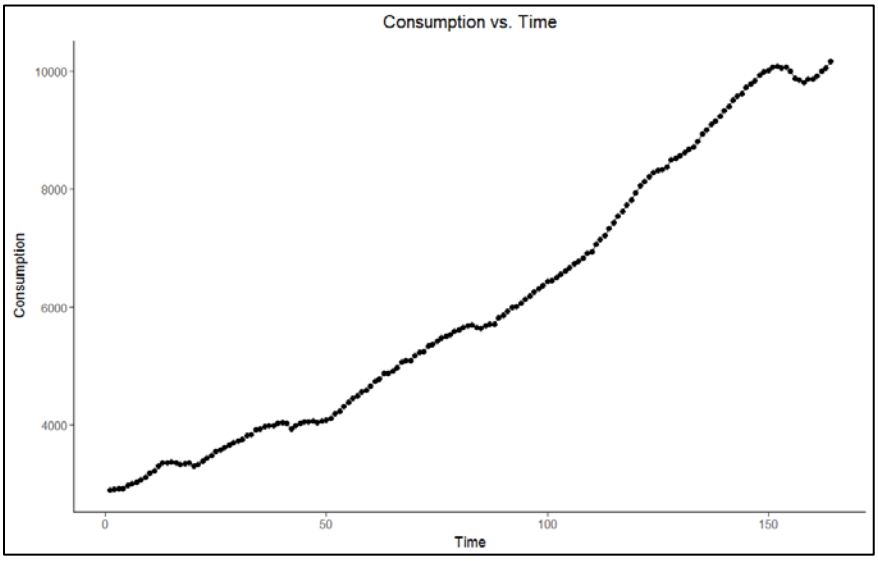
# THE UNIVERSITY OF TEXAS AT AUSTIN
## McCombs School of Business

**STA 372.5** **Spring 2018**

**HOMEWORK #10 – Due Wednesday, April 25**

1. The file STA372_Homework10_Question1.dat on the *Data sets* page of the Canvas class website contains quarterly U.S. consumption expenditures for the period 1970Q1 – 2010Q4 (i.e. 164 quarters of data). You are asked to forecast U.S. consumption for the first two quarters of 2011 and to provide 80% confidence intervals for each prediction. Note that there is no seasonality in the data so the original data rather than seasonally adjusted data will be analyzed.

   (a) Plot U.S. consumption expenditures (denoted $Consumption_t$) for the period 1970Q1 – 2010Q4. Is there a pattern in the data?




   (b) Let $W_t = Consumption_t - Consumption_{t-1}$ represent the first differences of $Consumption_t$.

   Compute and plot $W_t$. Are the first differences stationary? Why or why not?

(c) What empirical evidence (based on the PACF) is there that the AR(3) model

$$W_t - \mu = \beta_1(W_{t-1} - \mu) + \beta_2(W_{t-2} - \mu) + \beta_3(W_{t-3} - \mu) + \varepsilon_t \qquad \varepsilon_t \text{ iid } N(0, \sigma^2)$$

is the appropriate model to use for $W_t$?

(d) Estimate the parameters $\mu$, $\beta_1$, $\beta_2$, $\beta_3$ and $\sigma$ in the AR(3) model in part (c) using the *Arima* command in R. Be sure to use the *include.constant=TRUE* option. What are the estimates?

(e) Are the residuals from this model independent?

(f) Use the *auto.arima* command to select the appropriate model for $Consumption_t$. Set *stepwise=FALSE* so *auto.arima* does an exhaustive search of all models with $p + q \leq 5$.

Does it give the ARIMA(3, 1, 0) with drift model in part (c)?

Note that *auto.arima* returns *inf* (which represents infinity) for the ARIMA(1, 1, 4) model. This is *auto.arima*'s way of telling you that this model is not an appropriate one to consider in practice. Please see the next page for a more detailed explanation.

(g) Use the *auto.arima* command to select the appropriate model for $Consumption_t$ with *stepwise=TRUE* so that *auto.arima* only does a stepwise search through possible models (rather than an exhaustive search of all models with $p + q \leq 5$).

Does it give the same ARIMA(3, 1, 0) with drift model that the full search gives? Why are the results different?

(h) Using the ARIMA(3, 1, 0) with drift model obtained in part (c) and by *auto.arima* in part (f), what are the forecasts and 80% confidence intervals for U.S. consumption for the first two quarters of 2011 (i.e. in time periods $t$ = 165 and 166)?

**Why the ARIMA(1, 1, 4) model is not appropriate and *auto.arima* returns AICc = *inf***

You are <u>not</u> responsible for this explanation on the final exam. I have only included it in case you are interested.

An ARIMA(1, 1, 4) without drift model is

$$W_t = \beta_1 W_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \theta_4 \varepsilon_{t-4}.$$

Using *Arima* to estimate the parameters of this model gives:

```
------------------------------------------------------------------------
result <- Arima(y_time_series, order=c(1, 1, 4), include.drift=FALSE)
print (result)
------------------------------------------------------------------------

Series: y_time_series
ARIMA(1,1,4)

Coefficients:
         ar1      ma1     ma2     ma3      ma4
      0.9938  -0.6571  0.0016  0.0860  -0.2879
s.e.  0.0097   0.0794  0.0915  0.0961   0.0787

sigma^2 estimated as 1151:  log likelihood=-804.09
AIC=1620.18   AICc=1620.72   BIC=1638.74
```

Note that the estimate of the AR parameter $\beta_1$ is 0.994 so it is nearly one. The estimate of $\beta_1$ implies that $W_t$ is nearly nonstationary (if $\beta_1 = 1$, then $W_t$ is nonstationary). However, the augmented Dickey-Fuller test in the *auto.arima* command has already determined that $W_t$ is stationary because it determined that only one differencing step of $Consumption_t$ is required (if $W_t$ is nonstationary, then the augmented Dickey-Fuller test in *auto.arima* would have determined that $Consumption_t$ needed to be differenced twice). In addition, the plot of $W_t$ in part (b) shows that the $W_t$ values are not meandering so the graphical evidence also strongly suggests that $W_t$ is stationary.

Given the result of the augmented Dickey-Fuller test, *auto.arima* determines that the ARIMA(1, 1, 4) model with the estimate of $\beta_1$ of 0.994 is not the correct model. It therefore returns a value of *inf* for the AICc value even though the actual value of AICc for this model is 1620.72, as shown in the output above. In any case, the AICc of 1620.72 for this model is greater than the AICc for the ARIMA(3, 1, 0) with drift model that *auto.arima* selects in part (f).