THE UNIVERSITY OF TEXAS AT AUSTIN
McCombs School of Business

STA 372.5 **Spring 2018**

HOMEWORK #9 - Due Wednesday, April 18

1. Consider the following regression model with AR(1) errors:

$$y_t = \gamma x_t + \varepsilon_t$$

where

$$\varepsilon_t - \mu = \beta(\varepsilon_{t-1} - \mu) + a_t \qquad\qquad a_t \text{ iid } N(0, \sigma_a^2).$$

The parameters of this model are estimated using the data set shown below (only the first six and last six observations are shown).

<u>Data</u>

```
First six rows of the data table are:
       Y     X
1 -0.165 0.039
2  0.358 0.737
3  0.575 0.948
4  0.254 0.307
5  0.087 0.825
6  0.379 0.736

 Last six rows of the data table are:
        Y     X
35 -0.417 0.201
36  0.056 0.606
37  0.091 0.254
38  0.334 0.886
39  0.028 0.007
40  0.035 0.239
```

The R commands and output to estimate μ, β, γ and σ_a are shown below. Note that the one- and two-step ahead forecasts and confidence intervals have been replaced with asterisks.

Using all available information, answer parts (a) and (b).

```
------------------------------------------------------------------------
y_time_series <- ts(data_table[,1])
result <- Arima(y_time_series, xreg=data_table$X, order=c(1, 0, 0))
print (result)
cat ("\n", "Forecasts and confidence intervals are:", "\n", "\n")
result_forecast <- forecast(result, xreg=c(0.5, 0.8), h=2)
print (result_forecast)
------------------------------------------------------------------------
```

```
Series: y_time_series
Regression with ARIMA(1,0,0) errors

Coefficients:
         ar1  intercept  data_table$X
      0.5461    -0.0693        0.4817
s.e.  0.1288     0.0772        0.0847

sigma^2 estimated as 0.03824:  log likelihood=9.9
AIC=-11.81   AICc=-10.66   BIC=-5.05

Forecasts and confidence intervals are:

   Point Forecast       Lo 80     Hi 80      Lo 95     Hi 95
41      *********   ********* *********  ********* *********
42      *********   ********* *********  ********* *********
```

(a) What is the one-month ahead forecast and 95% confidence interval for y_{41} given $x_{41} = 0.5$?

(b) What is the two-month ahead forecast and 95% confidence interval for y_{42} given $x_{41} = 0.5$ and $x_{42} = 0.8$?

2. Read the case "Harmon Foods, Inc." The data for this case is in the file STA372_Homework9_Question2.dat on the *Data sets* page of the Canvas class website.

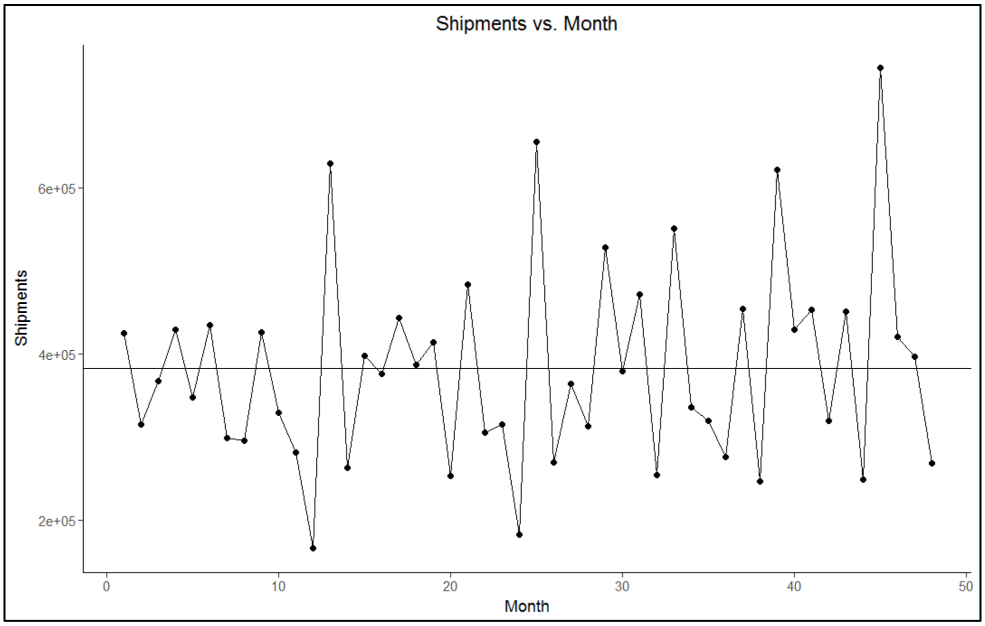
   The data are available for a 48-month period and the goal is to predict the sales of Treat cereal in months 49 and 50 given information through month t = 48, and to compute 80% confidence intervals for the predictions.

   The variables in the Excel file are:

   Month: Month from 1 to 48 (four years of monthly data)
   Shipments: Shipments of Treat cereal in month t
   SeasonalIndex: Seasonal index for $Shipments_t$ in month t – The indices are computed using a multiplicative model and classical seasonal decomposition
   ConsumerPack: Consumer packs (coupons) in month t
   ConsumerPack_Lag1: Consumer packs lagged one month (i.e. $ConsumerPack_{t-1}$)
   ConsumerPack_Lag2: Consumer packs lagged two months (i.e. $ConsumerPack_{t-2}$)
   DealerAllowance: Dealer allowances in month t
   DealerAllowance_Lag1: Dealer allowances lagged one month (i.e. $DealerAllowance_{t-1}$)
   DealerAllowance_Lag2: Dealer allowances lagged two months (i.e. $DealerAllowance_{t-2}$)

   Harmon Foods will set $ConsumerPack_{49}$ = 300,000, $ConsumerPack_{50}$ = 200,000, $DealerAllowance_{49}$ = \$200,000, and $DealerAllowance_{50}$ = \$100,000.

   (a) Plot $Shipments_t$ vs. $Month_t$. Is there evidence of a trend and/or seasonality in $Shipments_t$?



(b) Seasonally adjust $Shipments_t$ using the seasonal indices in the data file. Let

$$ShipmentsSA_t = Shipments_t \,/\, SeasonalIndex_t$$

represent the seasonally adjusted values. Did the classical seasonal decomposition procedure do a good job removing the seasonal component of Treat cereal?

(c) Run the following regression assuming AR(2) errors:

$$\begin{aligned} ShipmentsSA_t = {} & \gamma_1 Month_t \\ & + \gamma_2 ConsumerPack_t + \gamma_3 ConsumerPack_{t\text{-}1} + \gamma_4 ConsumerPack_{t\text{-}2} \\ & + \gamma_5 DealerAllowance_t + \gamma_6 DealerAllowance_{t\text{-}1} + \gamma_7 DealerAllowance_{t\text{-}2} + \varepsilon_t \end{aligned}$$

where

$$\varepsilon_t - \mu = \beta_1(\varepsilon_{t\text{-}1} - \mu) + \beta_2(\varepsilon_{t\text{-}2} - \mu) + a_t \qquad a_t \text{ iid } N(0, \sigma_a^2).$$

To run this regression, use the following set of R commands (note that `data_table[,10]` contains *ShipmentsSA*).:

```
y_time_series <- ts(data_table[,10])
X <- cbind(data_table$Month, data_table$ConsumerPack, data_table$CP_lag1,
           data_table$CP_lag2, data_table$DealerAllowance, data_table$DA_lag1,
           data_table$DA_lag2)
colnames(X) <- c("Month", "ConsumerPack", "CP_lag1", "CP_lag2",
                 "DealerAllowance", "DA_lag1", "DA_lag2")
result <- Arima(y_time_series, xreg=X, order=c(2, 0, 0))
print (result)
```

Use the autocorrelation function to determine whether the residuals a_t from this regression are independent.

(d) Now run the regression

$$ShipmentsSA_t = \gamma_1 Month_t + \gamma_5 DealerAllowance_t + \gamma_6 DealerAllowance_{t\text{-}1} + \varepsilon_t$$

where

$$\varepsilon_t - \mu = \beta_1(\varepsilon_{t\text{-}1} - \mu) + a_t \qquad a_t \text{ iid } N(0, \sigma_a^2).$$

This is essentially the regression in part (c) with the variables removed that were not statistically significant.

To run this regression, use the following set of R commands:

```
X <- cbind(data_table$Month, data_table$DealerAllowance, data_table$DA_lag1)
colnames(X) <- c("Month", "DealerAllowance", "DA_lag1")
result <- Arima(y_time_series, xreg=X, order=c(1, 0, 0))
print (result)
```

Use the autocorrelation function to determine whether the residuals a_t from this regression are independent.

Should the model in part (c) or part (d) be used for forecasting purposes?

(e) Predict $ShipmentsSA_t$ (i.e. predict seasonally adjusted $Shipments_t$) in months $t = 49$ and 50, and compute the 80% analytical confidence intervals for the predictions.

It is fine to compute the forecasts and confidence intervals using R but be sure to know how to compute them by hand as well (as in Problem #1 on this homework).

(f) Using the results from part (e), predict $Shipments_t$ (not seasonally adjusted $ShipmentsSA_t$) in months $t = 49$ and 50, and compute the 80% analytical confidence intervals for the predictions.

You will need to compute these forecasts and confidence intervals by hand using the results from part (e).