Phillip Efthimion

Week 9 – Homework

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Please see the attached word document for the Homework.

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Please note all homework is due submitted on-line by 1PM CST (Dallas) on Saturday March 24th.

Please submit an email to cmaybin@smu.edu titled MSDS\_8390 - [Last Name] - Week 9 Homework. For example, my submission would be titled MSDS\_8390 - Maybin - Week 9 Homework. In the email should be the following attachments containing the answers to the questions below:

• 1 Word document: (Assignment 1) with answer sections filled out.

• Note: R code for SARIMA model analysis and determination

▪ “WK 9 – SARIMA\_Tour Class Example.txt” file has the R code – you can use as a basis….

▪ NepalTouristData2013.xlsx file has the source data. Note: You will need to remove the total row at the bottom of the data, and sometimes excel imports data as text (you may want to check/change format of data in excel before importing…)

Note: Files used during class have been loaded onto 2DS.

Regards,

Chad

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Homework Section

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**NOTE: Load your graphs and replies into the word document below…**

Assignment 1 – ARIMA: Identify best fit model and make prediction using Tourist data in R.

• Determine best SARIMA model for Tourist data (as per class example)

o Load the data into R

touristdata<- read.csv("/Users/Phillip/Documents/SMU/BusinessIntelligence/week9/NepalTouristData2013V2.csv", header = T)

touristdata<-as.numeric(touristdata)

data = c(touristdata$X1992, touristdata$X1993, touristdata$X1994,touristdata$X1995, touristdata$X1996, touristdata$X1997,touristdata$X1998, touristdata$X1999, touristdata$X2000, touristdata$X2001, touristdata$X2002, touristdata$X2003, touristdata$X2004, touristdata$X2005, touristdata$X2006, touristdata$X2007, touristdata$X2008, touristdata$X2009, touristdata$X2010, touristdata$X2011, touristdata$X2012, touristdata$X2013)

df<-as.data.frame(data)

data <-df$data

ts.data<-ts(data, frequency = 12, start = c(1992,1))

After the data is loaded into R we have to make some changes. First, we have to make sure the data is coded as numeric. Then, after doing that we need to convert it into a time series using the ‘tseries’ R package. Having the data as a time series allows us to perform our analysis.

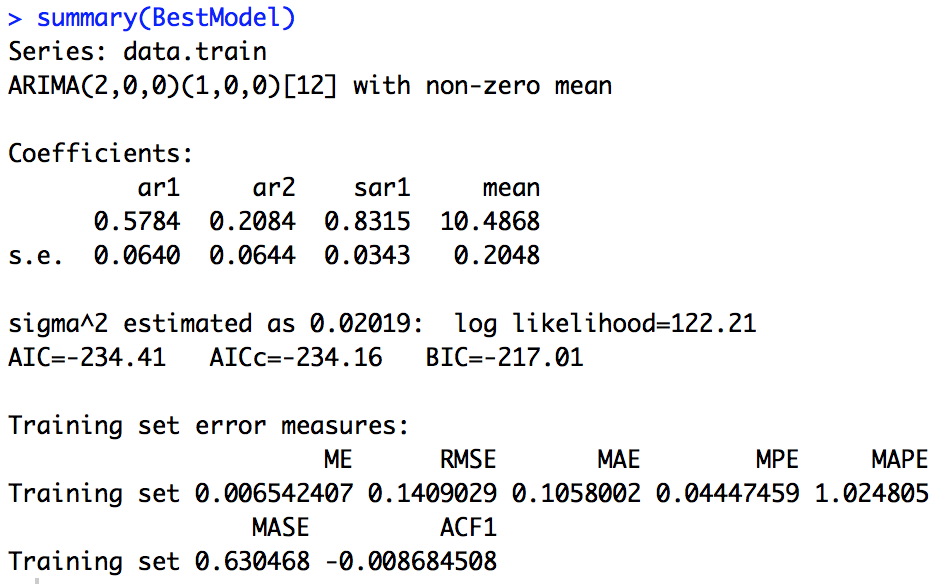
o Determine Best SARIMA model and for that model provide Coefficients and confidence interval. Note: you may need to transform (log, exponential or otherwise) the series. Also note review for unit root and stationarity of SARIMA model. Show graphical comparisons (i.e. graphs comparing prediction to test set, as per class) of the top three models reviewed.

HINT: You will probably have to transform the data and use several variations of the auto.arima function. See the documentation “13 - Time Series SARIMA Models - autoARIMA docs” loaded onto 2DS on options for this function.

While testing different models we checked which different test to use. First, we check having ic use the aic or bic as an indicator. They both recommended we use the same model: ARIMA(1,1,1)(0,1,1). However, there was minimal difference between the outputs. Looking at the diagnostics there was not much of a difference between the two models either.

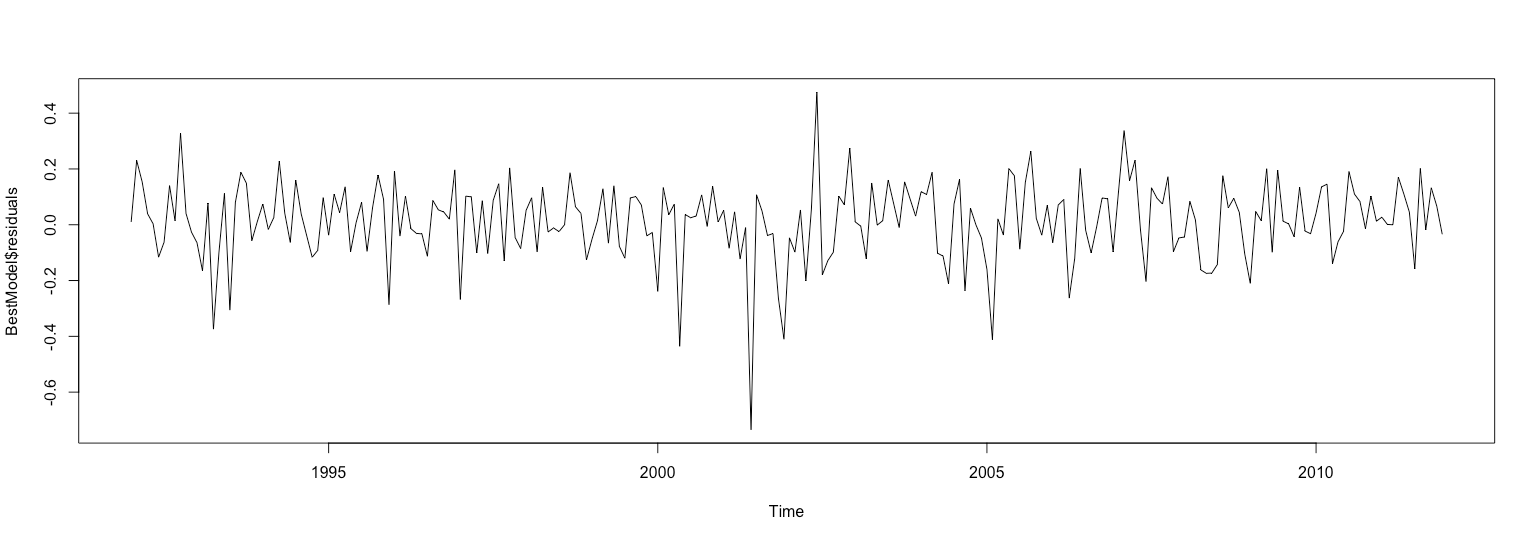
We now tested the model again to determine which type of test was best to use. We tested against ‘kpss’, ‘adf’, and ‘pp’. We determined which test was ‘best’ by looking at which had the lowest MAPE. This resulted in the ‘kpss’ test being determined best since it had the lowest MAPE value. This was our arima1 model. However, before we determine it was our nest we are going to perform a log transformation on the data to see if can get an even better model.

With the log transformation we first test whether to use ‘aic’ or ‘bic’ just as before. Interestingly, each gives us a different ARIMA model. Using AIC we get ARIMA(2,1,1)(1,0,0) with AIC: -238.9053 and our second arima models gives us ARIMA(0,1,1)(1,0,0) with BIC: -220.6738. Therefore, we are going to use BIC. The diagonstics are not different enough to provide us with any clear answers and the summary statistics provide us with similar MAPE. We will now determine which type of test (kpss, adf, or pp) we should use for our log transformed data. These tests determine that ‘pp’ is the best test for our purposes.

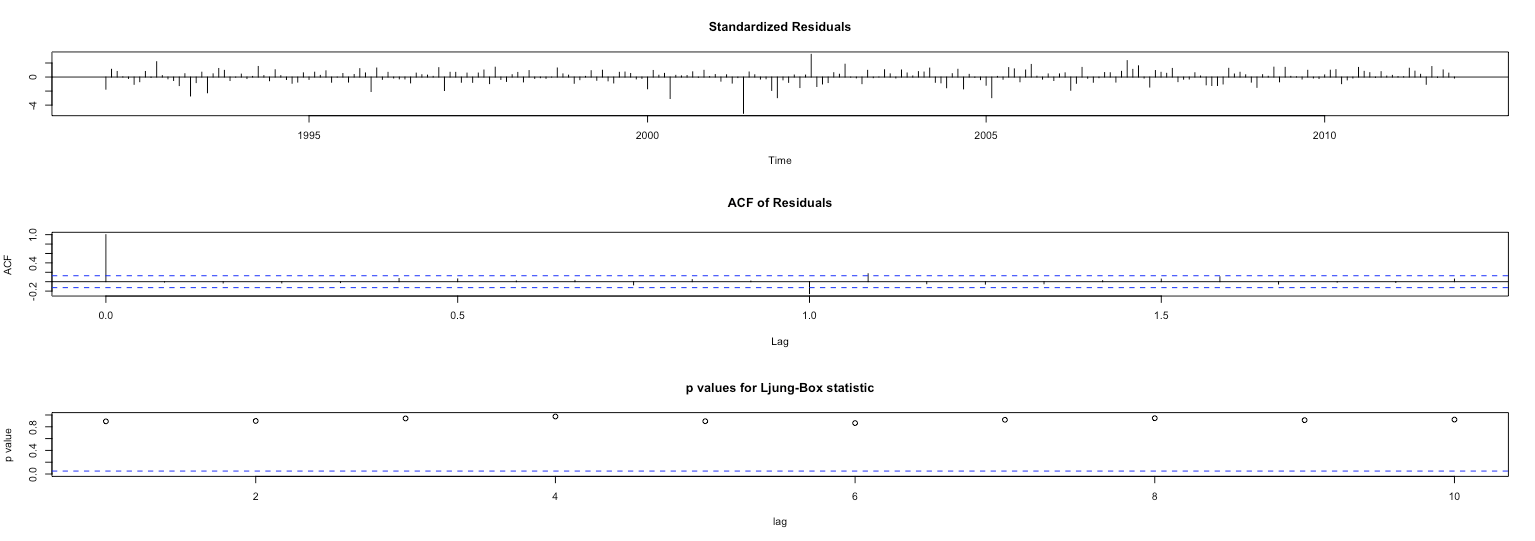


The MAPE for the log transformed data is much lower than with the non-transformed data. This model is much better for us. The ACF shows that there are significant points at lag(0) and lag(1). This model is normally distributed because performing a Jarque Bera Test provides us with a very low p-value of less than 0.0001.

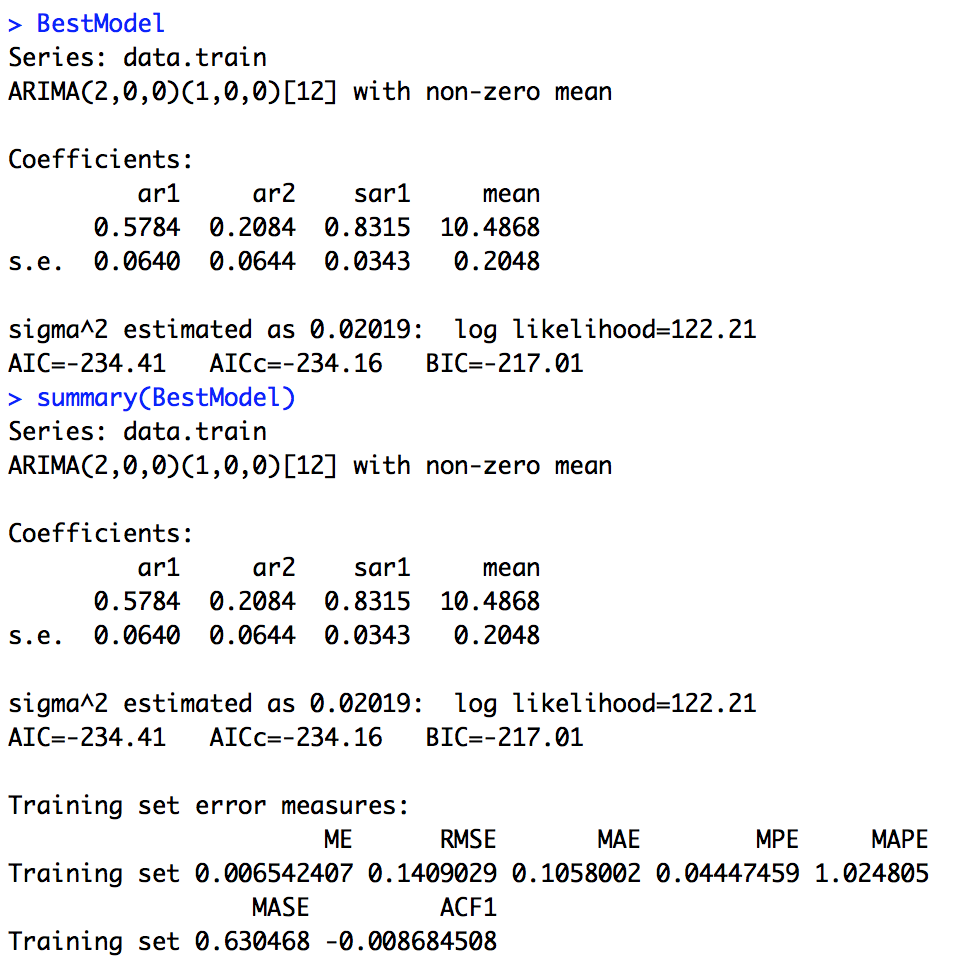
o Plot the SARIMA model chosen and place below.

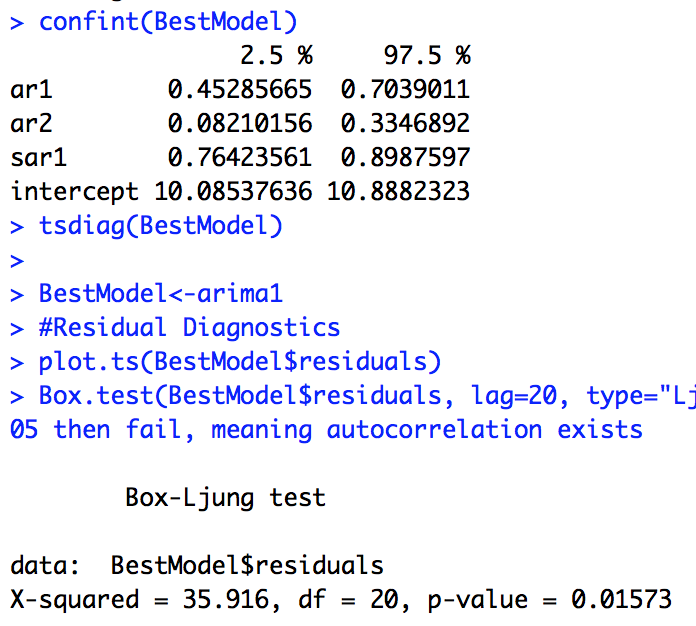


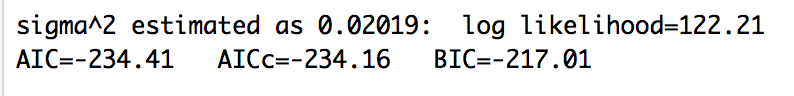
o Plot the Standardized Residuals, ACF of Residuals and p values for Ljung-Box statistic and provide a (brief) analysis for each.



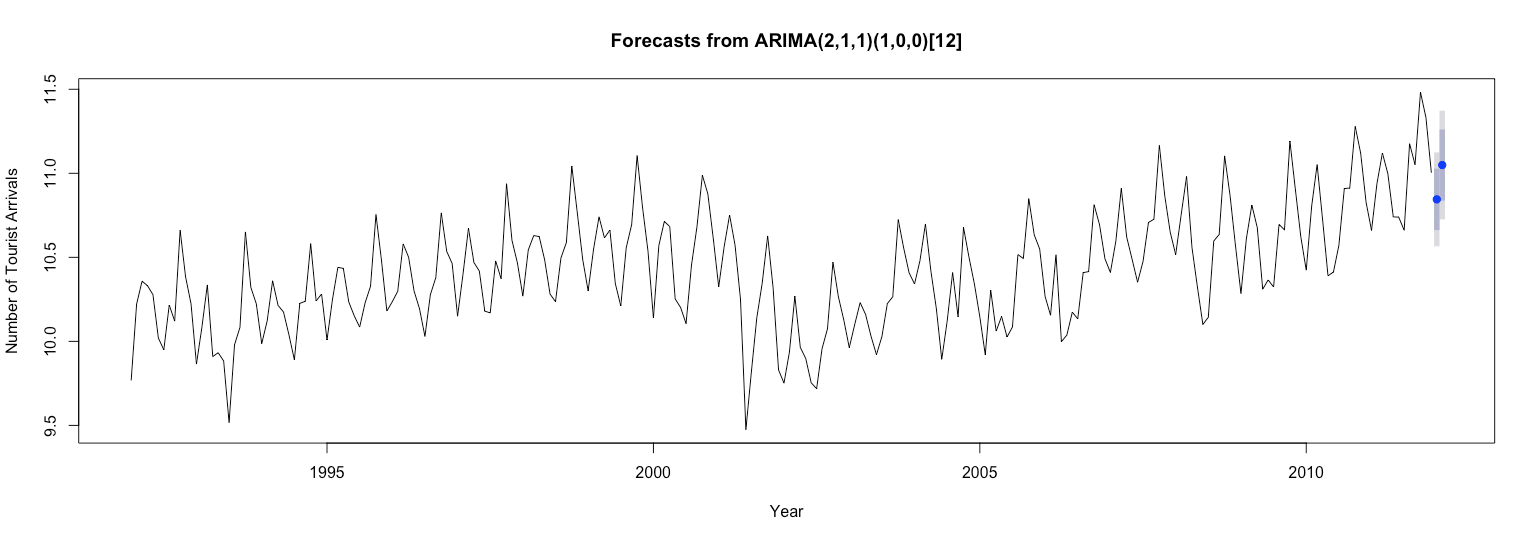
o Provide accuracy metrics for chosen model and place below.

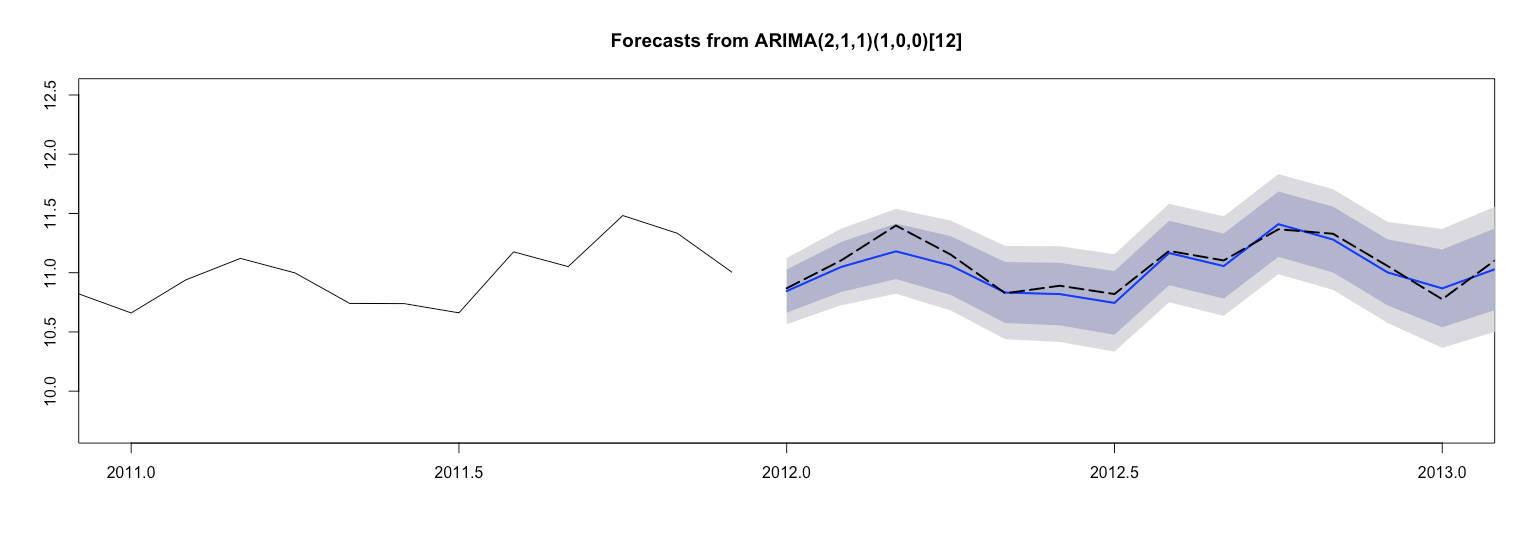






o Use the chosen SARIMA model to predict 2 periods forward and plot. Place output below.





Appendix

#Algorithm to analyze SARIMA

# Code adapted from Ani Katchova, Robert Nau and Shishir Shakya

rm(list = ls())

gc()

par(mfrow=c(1,1))

#Importing the data

touristdata<- read.csv("/Users/Phillip/Documents/SMU/BusinessIntelligence/week9/NepalTouristData2013V2.csv", header = T)

View(touristdata)

touristdata<-as.numeric(touristdata)

#touristdata <- as.numeric(touristdata)

#data = c(touristdata$V1, touristdata$V2, touristdata$V3, touristdata$V4, touristdata$V5,touristdata$V6, touristdata$V7, touristdata$V8, touristdata$V9,touristdata$V10, touristdata$V11,touristdata$V12, touristdata$V13,touristdata$V14, touristdata$V15, touristdata$V16, touristdata$V17, touristdata$V18, touristdata$V19, touristdata$V20, touristdata$V21)

data = c(touristdata$X1992, touristdata$X1993, touristdata$X1994,touristdata$X1995, touristdata$X1996, touristdata$X1997,touristdata$X1998, touristdata$X1999, touristdata$X2000, touristdata$X2001, touristdata$X2002, touristdata$X2003, touristdata$X2004, touristdata$X2005, touristdata$X2006, touristdata$X2007, touristdata$X2008, touristdata$X2009, touristdata$X2010, touristdata$X2011, touristdata$X2012, touristdata$X2013)

View(data)

#Let's make sure we have what we want...

df<-as.data.frame(data)

data <-df$data

data

#data<-log(data) #to log the data...should you need to...

ts.data<-ts(data, frequency = 12, start = c(1992,1))

View(ts.data)

plot(ts.data)

# the plot shows seasonality

#And let's make sure all the data is structured as we wanted

dim(as.matrix(ts.data))

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Split data into training and testing data set

data.train <- window(ts.data, start=c(1992,1), end=c(2011,12))

plot(data.train)

dim(as.matrix(data.train))

data.test <- window(ts.data, start=c(2012,1))

plot(data.test)

dim(as.matrix(data.test))

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Developing the SARIMA model

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

library(forecast)

# difference is using aic or bic. trsin using kpss

arima1 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="aic")

arima2 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="bic") #Note: BIC is usually better for larger samples...

summary(arima1)

confint(arima1)

# arima 1 & 2 both recommended to use the same model

# ARIMA(1,1,1)(0,1,1)[12] BIC: 4272.389 4247.858

# minimal difference

tsdiag(arima1)

# standardied residuals - white noise

# ACF - 1 large positive then nothing else special

# Ljun-Box p values - 1st 2 high but then close to 0. Stays above 0.01 until lag 10

tsdiag(arima2)

# not much of a difference from arim1

summary(arima1)

summary(arima2)

# not much of a difference from arima1

#Note that the model is the same, but the differing values of aic vs bic...

#If we want to change the model creation, we can vary the criteria...

arima1 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="bic")

arima2 <- auto.arima(data.train, trace=TRUE, test="adf", ic="bic") #Note: BIC is usually better for larger samples...

arima3 <- auto.arima(data.train, trace=TRUE, test="pp", ic="bic") #Note: BIC is usually better for larger samples...

#and let's check what the differences are...

summary(arima1)

summary(arima2)

summary(arima3)

#Lowest MAPE (and others) shows a bias towards kpss as the test to use...

BestModel <- auto.arima(data.train, seasonal = TRUE, trace=TRUE, test="kpss", ic="bic")

BestModel

summary(BestModel)

confint(BestModel)

tsdiag(BestModel)

BestModel<-arima1

#Residual Diagnostics

plot.ts(BestModel$residuals)

Box.test(BestModel$residuals, lag=20, type="Ljung-Box") #If p-val >0.05 then fail, meaning autocorrelation exists

acf(BestModel$residuals, lag.max = 24, main="ACF of the Model") #if spikes touch the bounds, this is also an indication of failing the autocorrelation test

#Box.test(BestModel$residuals^2, lag = 20, type = "Ljung-Box") #Test for GARCH effect, if pval >0.05 then have ARCH effect, and should consider a volatility model (i.e. GARCH, etc.)

library(tseries)

# p less than 0.5 normal dist

jarque.bera.test(BestModel$residuals) #if p-value >0.05 then the residuals have a normal distibution (not reject null Hypothesis of normality)

BestModel.forecast <- forecast(BestModel, h=24) #forecast 24 periods ahead

BestModel.forecast

plot(BestModel.forecast, xlab="Year", ylab="Number of Tourist Arrivals")

library(TSPred)

plotarimapred(data.test, BestModel, xlim=c(2011, 2013), range.percent = 0.05)

accuracy(BestModel.forecast, data.test)

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Residual Diagnostics

plot.ts(arima1$residuals)

Box.test(arima1$residuals, lag=20, type="Ljung-Box") #If p-val <0.05 then fail, meaning autocorrelation exists

acf(arima1$residuals, lag.max = 24, main="ACF of the Model") #if spikes touch the bounds, this is also an indication of failing the autocorrelation test

Box.test(arima1$residuals^2, lag = 20, type = "Ljung-Box") #Test for GARCH effect, if pval <0.05 then have ARCH effect, and should consider a volatility model (i.e. GARCH, etc.)

library(tseries)

jarque.bera.test(arima1$residuals) #if p-value >0.05 then the residuals have a normal distibution (not reject null Hypothesis of normality)

#library(forecast)

#install.packages('forecast')

#install.packages('forecast', dependencies = TRUE)

#arima1.forecast <- forecast.Arima(arima1, h=24) #forecast 24 periods ahead

#arima1.forecast <- forecast::auto.arima(data.train)

arima1.forecast <- forecast(arima1, h=24) #forecast 24 periods ahead

arima1.forecast

plot(arima1.forecast, xlab="Year", ylab="Number of Tourist Arrivals")

library(TSPred)

plotarimapred(data.test, arima1, xlim=c(2011, 2013), range.percent = 0.05)

accuracy(arima1.forecast, data.test)

#Does not look particularly great, so let's start trying different models...

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Different Model Combinations....

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

BestModel1 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="aic")

BestModel2 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="bic")

BestModel3 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="aicc")

BestModel4 <- auto.arima(data.train, trace=TRUE, test="adf", ic="aic")

BestModel5 <- auto.arima(data.train, trace=TRUE, test="adf", ic="bic")

BestModel6 <- auto.arima(data.train, trace=TRUE, test="adf", ic="aicc")

BestModel7 <- auto.arima(data.train, trace=TRUE, test="pp", ic="aic")

BestModel8 <- auto.arima(data.train, trace=TRUE, test="pp", ic="bic")

BestModel9 <- auto.arima(data.train, trace=TRUE, test="pp", ic="aicc")

modList <- list(BestModel1, BestModel2, BestModel3, BestModel4, BestModel5, BestModel6, BestModel7, BestModel8, BestModel9)

modList

#Amodel <- BestModel1

setwd("D:/My Document Storage/2015 SMU/Econometrics/Week 8\_Special Topics/Plots\_forecast")

#par(mfrow=c(1,1))

for (Amodel in modList){

i=0

arima1.forecast <- forecast(Amodel, h=24) #forecast 24 periods ahead

arima1.forecast

plot(arima1.forecast, xlab="Year", ylab="Number of Tourist Arrivals")

library(TSPred)

plotarimapred(data.test, Amodel, xlim=c(2011, 2013), range.percent = 0.05)

accuracy(arima1.forecast, data.test)

#plot as confirmation

pdf( paste0("Plot - ", Amodel,format(Sys.time(), "%a %b %d %H %M %S %Y"), " .pdf"),width=7,height=5)

plotarimapred(data.test, Amodel, xlim=c(2011, 2013), range.percent = 0.05)

accuracy(arima1.forecast, data.test)

dev.off()

}

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Manual Plotting...

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*

plotarimapred(data.test, BestModel1, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel2, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel3, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel4, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel5, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel6, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel7, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel8, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel9, xlim=c(2011, 2013), range.percent = 0.05)

# Homework 9 log transformation

#Algorithm to analyze SARIMA

# Code adapted from Ani Katchova, Robert Nau and Shishir Shakya

rm(list = ls())

gc()

par(mfrow=c(1,1))

#Importing the data

touristdata<- read.csv("/Users/Phillip/Documents/SMU/BusinessIntelligence/week9/NepalTouristData2013V2.csv", header = T)

View(touristdata)

touristdata<-as.numeric(touristdata)

#touristdata <- as.numeric(touristdata)

#data = c(touristdata$V1, touristdata$V2, touristdata$V3, touristdata$V4, touristdata$V5,touristdata$V6, touristdata$V7, touristdata$V8, touristdata$V9,touristdata$V10, touristdata$V11,touristdata$V12, touristdata$V13,touristdata$V14, touristdata$V15, touristdata$V16, touristdata$V17, touristdata$V18, touristdata$V19, touristdata$V20, touristdata$V21)

data = c(touristdata$X1992, touristdata$X1993, touristdata$X1994,touristdata$X1995, touristdata$X1996, touristdata$X1997,touristdata$X1998, touristdata$X1999, touristdata$X2000, touristdata$X2001, touristdata$X2002, touristdata$X2003, touristdata$X2004, touristdata$X2005, touristdata$X2006, touristdata$X2007, touristdata$X2008, touristdata$X2009, touristdata$X2010, touristdata$X2011, touristdata$X2012, touristdata$X2013)

View(data)

#Let's make sure we have what we want...

df<-as.data.frame(data)

data <-df$data

data

data<-log(data) #to log the data...should you need to...

ts.data<-ts(data, frequency = 12, start = c(1992,1))

View(ts.data)

plot(ts.data)

# the plot shows seasonality

#And let's make sure all the data is structured as we wanted

dim(as.matrix(ts.data))

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Split data into training and testing data set

data.train <- window(ts.data, start=c(1992,1), end=c(2011,12))

plot(data.train)

dim(as.matrix(data.train))

data.test <- window(ts.data, start=c(2012,1))

plot(data.test)

dim(as.matrix(data.test))

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Developing the SARIMA model

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

library(forecast)

# difference is using aic or bic. trsin using kpss

arima1 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="aic")

arima2 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="bic") #Note: BIC is usually better for larger samples...

summary(arima1)

confint(arima1)

summary(arima2)

confint(arima2)

# arima 1: ARIMA(2,1,1)(1,0,0)[12] BIC: -238.9053

# arima 2: ARIMA(0,1,1)(1,0,0)[12] BIC: -220.6738

# minimal difference, 2 better

tsdiag(arima1)

# standardied residuals - white noise

# ACF - 1 large positive then nothing else special

# Ljun-Box p values - high p-value all throughout

tsdiag(arima2)

# not much of a difference from arima1

# Ljun-Box p values - oscillating, high p-value all throughout

summary(arima1)

summary(arima2)

# not much of a difference from arima1 in MAPE

#Note that the model is the same, but the differing values of aic vs bic...

#If we want to change the model creation, we can vary the criteria...

arima1 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="bic")

arima2 <- auto.arima(data.train, trace=TRUE, test="adf", ic="bic") #Note: BIC is usually better for larger samples...

arima3 <- auto.arima(data.train, trace=TRUE, test="pp", ic="bic") #Note: BIC is usually better for larger samples...

#and let's check what the differences are...

summary(arima1)

summary(arima2)

summary(arima3)

#Lowest MAPE (and others) shows a bias towards 'pp' as the test to use...

BestModel <- auto.arima(data.train, seasonal = TRUE, trace=TRUE, test="pp", ic="bic")

BestModel

# ARIMA(2,0,0)(1,0,0)[12] with non-zero mean : -217.0097

# MAPE: 1.024805

summary(BestModel)

confint(BestModel)

tsdiag(BestModel)

BestModel<-arima1

#Residual Diagnostics

plot.ts(BestModel$residuals)

Box.test(BestModel$residuals, lag=20, type="Ljung-Box") #If p-val >0.05 then fail, meaning autocorrelation exists

acf(BestModel$residuals, lag.max = 24, main="ACF of the Model") #if spikes touch the bounds, this is also an indication of failing the autocorrelation test

# important lines at 0 (+) and lag 1 (-)

#Box.test(BestModel$residuals^2, lag = 20, type = "Ljung-Box") #Test for GARCH effect, if pval >0.05 then have ARCH effect, and should consider a volatility model (i.e. GARCH, etc.)

library(tseries)

# p less than 0.5 normal dist

jarque.bera.test(BestModel$residuals) #if p-value >0.05 then the residuals have a normal distibution (not reject null Hypothesis of normality)

# normal distribution

BestModel.forecast <- forecast(BestModel, h=2) #forecast 24 periods ahead

BestModel.forecast

plot(BestModel.forecast, xlab="Year", ylab="Number of Tourist Arrivals")

library(TSPred)

plotarimapred(data.test, BestModel, xlim=c(2011, 2013), range.percent = 0.05)

accuracy(BestModel.forecast, data.test)

# MAPE much lower than before

# test set lower than training set

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Residual Diagnostics

plot.ts(arima1$residuals)

Box.test(arima1$residuals, lag=20, type="Ljung-Box") #If p-val <0.05 then fail, meaning autocorrelation exists

acf(arima1$residuals, lag.max = 24, main="ACF of the Model") #if spikes touch the bounds, this is also an indication of failing the autocorrelation test

Box.test(arima1$residuals^2, lag = 20, type = "Ljung-Box") #Test for GARCH effect, if pval <0.05 then have ARCH effect, and should consider a volatility model (i.e. GARCH, etc.)

library(tseries)

jarque.bera.test(arima1$residuals) #if p-value >0.05 then the residuals have a normal distibution (not reject null Hypothesis of normality)

#library(forecast)

#install.packages('forecast')

#install.packages('forecast', dependencies = TRUE)

#arima1.forecast <- forecast.Arima(arima1, h=24) #forecast 24 periods ahead

#arima1.forecast <- forecast::auto.arima(data.train)

arima1.forecast <- forecast(arima1, h=2) #forecast 2 periods ahead

arima1.forecast

plot(arima1.forecast, xlab="Year", ylab="Number of Tourist Arrivals")

library(TSPred)

plotarimapred(data.test, arima1, xlim=c(2011, 2013), range.percent = 0.05)

accuracy(arima1.forecast, data.test)

# same accuracies

#Does not look particularly great, so let's start trying different models...

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Different Model Combinations....

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

BestModel1 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="aic")

BestModel2 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="bic")

BestModel3 <- auto.arima(data.train, trace=TRUE, test="kpss", ic="aicc")

BestModel4 <- auto.arima(data.train, trace=TRUE, test="adf", ic="aic")

BestModel5 <- auto.arima(data.train, trace=TRUE, test="adf", ic="bic")

BestModel6 <- auto.arima(data.train, trace=TRUE, test="adf", ic="aicc")

BestModel7 <- auto.arima(data.train, trace=TRUE, test="pp", ic="aic")

BestModel8 <- auto.arima(data.train, trace=TRUE, test="pp", ic="bic")

BestModel9 <- auto.arima(data.train, trace=TRUE, test="pp", ic="aicc")

modList <- list(BestModel1, BestModel2, BestModel3, BestModel4, BestModel5, BestModel6, BestModel7, BestModel8, BestModel9)

modList

#Amodel <- BestModel1

setwd("D:/My Document Storage/2015 SMU/Econometrics/Week 8\_Special Topics/Plots\_forecast")

#par(mfrow=c(1,1))

for (Amodel in modList){

i=0

arima1.forecast <- forecast(Amodel, h=24) #forecast 24 periods ahead

arima1.forecast

plot(arima1.forecast, xlab="Year", ylab="Number of Tourist Arrivals")

library(TSPred)

plotarimapred(data.test, Amodel, xlim=c(2011, 2013), range.percent = 0.05)

accuracy(arima1.forecast, data.test)

#plot as confirmation

pdf( paste0("Plot - ", Amodel,format(Sys.time(), "%a %b %d %H %M %S %Y"), " .pdf"),width=7,height=5)

plotarimapred(data.test, Amodel, xlim=c(2011, 2013), range.percent = 0.05)

accuracy(arima1.forecast, data.test)

dev.off()

}

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*

#Manual Plotting...

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*

plotarimapred(data.test, BestModel1, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel2, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel3, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel4, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel5, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel6, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel7, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel8, xlim=c(2011, 2013), range.percent = 0.05)

plotarimapred(data.test, BestModel9, xlim=c(2011, 2013), range.percent = 0.05)