DATA 624 - Homework 6

OMER OZEREN

Table of Contents

Exercises 8.1, 8.2, 8.3, 8.5, 8.6 and 8.7 from the Hyndman online Forecasting book.

## Question 8.1

Figure 8.31 shows the ACFs for 36 random numbers, 360 random numbers and 1,000 random numbers.

### a) Explain the differences among these figures. Do they all indicate that the data are white noise?

A time series is white noise if the variables are independent and identically distributed with a mean of zero. This means that all variables have the same variance (sigma^2) and each value has a zero correlation with all other values in the series.The figures indicate that the data is white noise (residuals correlated mean=0,variance=constant).

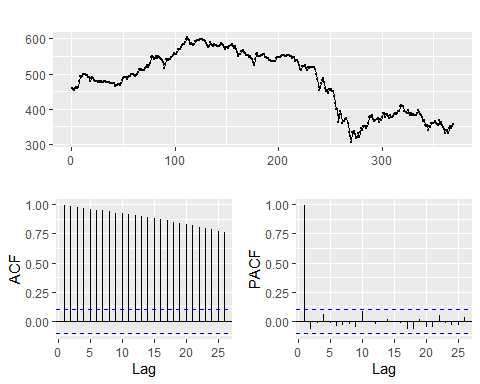
### B) Why are the critical values at different distances from the mean of zero? Why are the autocorrelations different in each figure when they each refer to white noise?

If the series is a weakly stationary time series (with mean ) with auto-covariances then a law of large numbers holds. The reason why critical values are different is that critical values differ depending upon the amount of data.

## Question 8.2

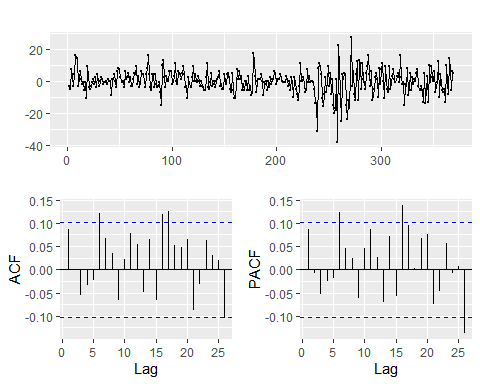
A classic example of a non-stationary series is the daily closing IBM stock price series (data set ibmclose). Use R to plot the daily closing prices for IBM stock and the ACF and PACF. Explain how each plot shows that the series is non-stationary and should be differenced.

ggtsdisplay(ibmclose)



The ACF plot indicates time series (IBM Closing Prices) is not stationary level because of very high autocorrelation values. By looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of the differenced series, we can verify the numbers of AR and/or MA terms that are needed.The ACF plot is merely a bar chart of the coefficients of correlation between a time series and lags of itself.The ACF graph in ibmclose data shows the observed correlation is well above the critical values.In stationary level series,The ACF will drop to zero quickly.The PACF plot is a plot of the partial correlation coefficients between the series and lags of itself.

ibmclose.diff <- diff(ibmclose)  
ggtsdisplay(ibmclose.diff)



In stationary level series(differenced),The ACF dropped to zero quickly.However, we still get some extreme values that greater than critical value (0.10).Our threshold is really important in this case because If we think that threshold for critical values needs to be 0.15, we can assume that there are no extreme values.

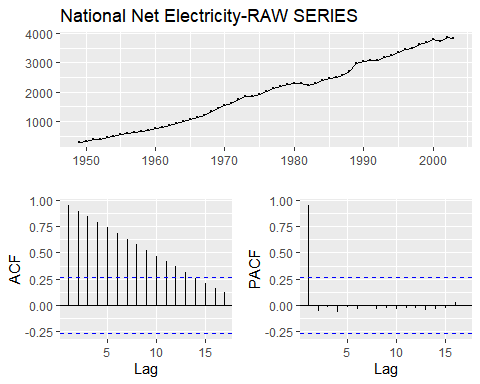
## Question 8.3

For the following series, find an appropriate Box-Cox transformation and order of differencing in order to obtain stationary data.

### A) usnetelec

**The trend in timeseries looks nearly linear.I wil check stationary status to determine the order of differencing.**

ggtsdisplay(usnetelec, main = "National Net Electricity-RAW SERIES")



#### Ljung-box test-RAW data series:

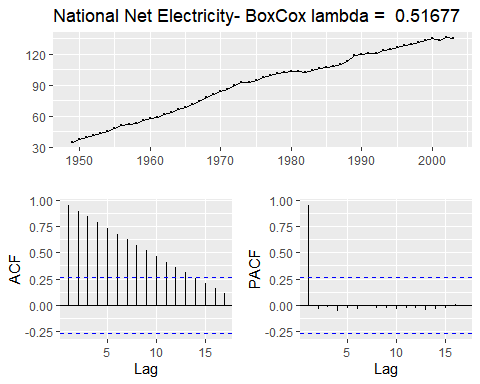
Box.test(usnetelec, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: usnetelec  
## X-squared = 52.549, df = 1, p-value = 4.196e-13

As we seet hat the p-value is low (p-value = 4.196e-13), the Null hypothesis is REJECTED (Autocorrelation exists).

#### Box-Cox transformation:

usnetelec\_lambda <- round(BoxCox.lambda(usnetelec),5)  
usnetelec\_BC <- BoxCox(usnetelec, usnetelec\_lambda)  
ggtsdisplay(usnetelec\_BC, main = paste("National Net Electricity- BoxCox lambda = ",  
 usnetelec\_lambda))



Time series data with Box-Cox transformation still gives non-stationary data (ADF plot)

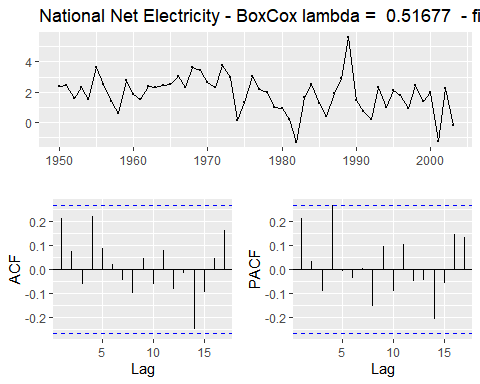
Box.test(usnetelec\_BC, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: usnetelec\_BC  
## X-squared = 52.191, df = 1, p-value = 5.035e-13

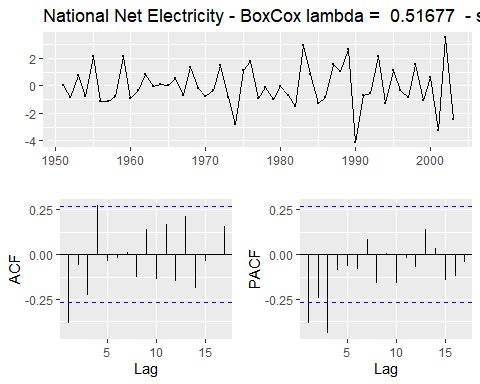
Because the p-value is still low (p-value = 5.035e-13), the Null hypothesis is REJECTED (Autocorrelation exists).

#### Differenced Box-Cox Data:

usnetelec\_BC\_ndiffs <- ndiffs(usnetelec\_BC)  
ggtsdisplay(diff(usnetelec\_BC), main=paste("National Net Electricity - BoxCox lambda = ",  
 usnetelec\_lambda," - first difference"))



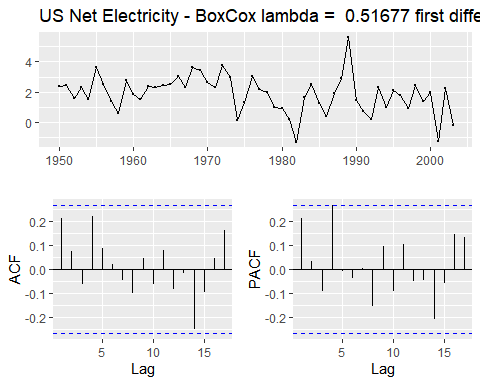
ggtsdisplay(diff(diff(usnetelec\_BC)), main=paste("National Net Electricity - BoxCox lambda = ",  
 usnetelec\_lambda," - second difference"))



I’m going to test two different lag of orders such as I(1),and (2)

#### Ljung-box test-Transformed and Differenced data series: Box-Cox order(1):

usnetelec\_BC %>% diff() -> usnetelec\_BC\_d1  
ggtsdisplay(usnetelec\_BC\_d1, main = paste("US Net Electricity - BoxCox lambda = ",  
 usnetelec\_lambda, "first difference"))



#Ljung-box test:  
Box.test(usnetelec\_BC\_d1, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: usnetelec\_BC\_d1  
## X-squared = 2.6132, df = 1, p-value = 0.106

I noticed that the p-value is greater than p-value threshold (0.05), we FAIL TO REJECT the null hypothesis, which states the data are independent (i.e., no serial correlation.)So, first-differenced usnetelec data can be thought of as a white noise series.

#### KPSS Test Transformed and Differenced data series: Box-Cox order(1):

kpss.test(usnetelec\_BC\_d1, "Level", lshort = F)

## Warning in kpss.test(usnetelec\_BC\_d1, "Level", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: usnetelec\_BC\_d1  
## KPSS Level = 0.33954, Truncation lag parameter = 10, p-value = 0.1

kpss.test(usnetelec\_BC\_d1, "Trend", lshort = F)

## Warning in kpss.test(usnetelec\_BC\_d1, "Trend", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: usnetelec\_BC\_d1  
## KPSS Trend = 0.074231, Truncation lag parameter = 10, p-value = 0.1

usnetelec\_BC\_d1 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 10 lags.   
##   
## Value of test-statistic is: 0.3395   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

usnetelec\_BC\_d1 %>% ur.kpss(type="tau", lags="long") %>% summary()

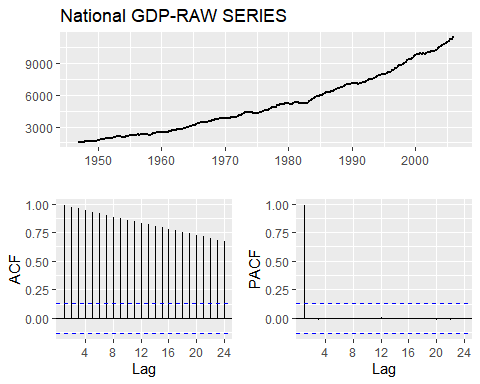
##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 10 lags.   
##   
## Value of test-statistic is: 0.0742   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

**The first differenced series ,order(1), data with Box-Cox transformation pass on Both KPSS tests which states that second-differencing is not required.**

### B) usgdp

**The trending timeseries usgdp looks nearly linear.I will check stationary status to determine the order of differencing.**

ggtsdisplay(usgdp, main = "National GDP-RAW SERIES")



#### Ljung-box test-RAW data series:

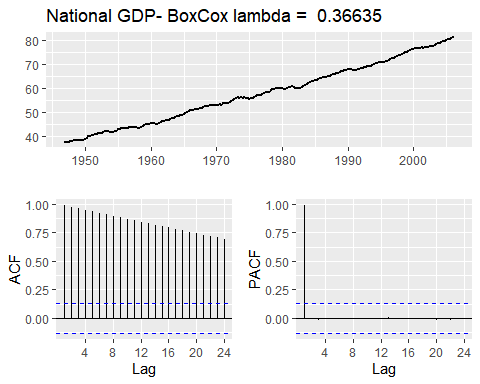
Box.test(usgdp, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: usgdp  
## X-squared = 233.07, df = 1, p-value < 2.2e-16

As we seet hat the p-value is low (p-value < 2.2e-16), the Null hypothesis is REJECTED

#### Box-Cox transformation:

usgdp\_lambda <- round(BoxCox.lambda(usgdp),5)  
usgdp\_BC <- BoxCox(usgdp, usgdp\_lambda)  
ggtsdisplay(usgdp\_BC, main = paste("National GDP- BoxCox lambda = ",  
 usgdp\_lambda))



Time series data with Box-Cox transformation still gives non-stationary data (ADF plot).The graph of the Box-Cox transformed data shows strong autocorrelation.

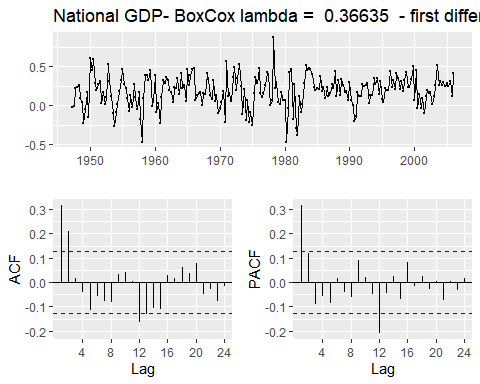
Box.test(usgdp\_BC, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: usgdp\_BC  
## X-squared = 233.76, df = 1, p-value < 2.2e-16

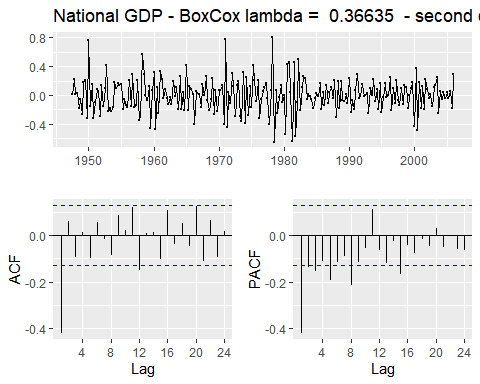
Because the p-value is still low (p-value < 2.2e-16), the Null hypothesis is REJECTED .(Autocorrelation exists)

#### Differenced Box-Cox Data:

usgdp\_BC\_ndiffs <- ndiffs(usgdp\_BC)  
ggtsdisplay(diff(usgdp\_BC), main=paste("National GDP- BoxCox lambda = ",  
 usgdp\_lambda," - first difference"))



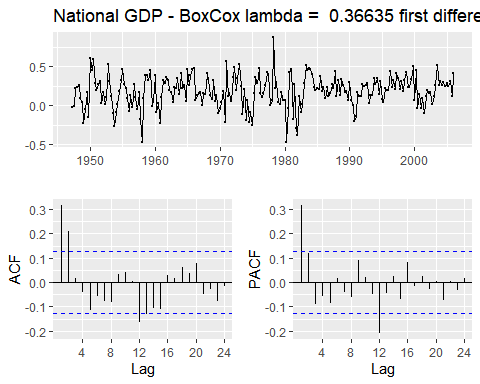
ggtsdisplay(diff(diff(usgdp\_BC)), main=paste("National GDP - BoxCox lambda = ",  
 usgdp\_lambda," - second difference"))



I’m going to test two different orders (1) and (2)

#### Ljung-box test-Transformed and Differenced data series: Box-Cox order(1):

usgdp\_BC %>% diff() -> usgdp\_BC\_d1  
ggtsdisplay(usgdp\_BC\_d1, main = paste("National GDP - BoxCox lambda = ",  
 usgdp\_lambda, "first difference"))



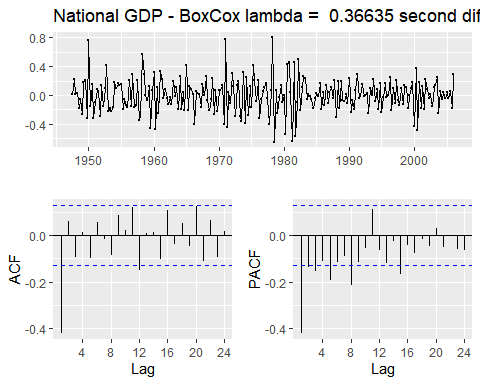
#Ljung-box test:  
Box.test(usgdp\_BC\_d1, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: usgdp\_BC\_d1  
## X-squared = 23.874, df = 1, p-value = 1.029e-06

Because the p-value is still low (p-value < 1.029e-06), the Null hypothesis is REJECTED (Autocorrelation exists).

#### Ljung-box test-Transformed and 2nd Differenced data series: Box-Cox order(2):

usgdp\_BC %>% diff() %>% diff() -> usgdp\_BC\_d2  
ggtsdisplay(usgdp\_BC\_d2, main = paste("National GDP - BoxCox lambda = ",  
 usgdp\_lambda, "second difference"))



#Ljung-box test:  
Box.test(usgdp\_BC\_d2, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: usgdp\_BC\_d2  
## X-squared = 41.715, df = 1, p-value = 1.056e-10

The test states that the p-value is less than p-value threshold (0.05), we REJECT the null hypothesis (Autocorrelation exists).

#### KPSS Test Transformed and Differenced data series: Box-Cox order(1):

kpss.test(usgdp\_BC\_d1, "Level", lshort = F)

## Warning in kpss.test(usgdp\_BC\_d1, "Level", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: usgdp\_BC\_d1  
## KPSS Level = 0.23949, Truncation lag parameter = 14, p-value = 0.1

kpss.test(usgdp\_BC\_d1, "Trend", lshort = F)

## Warning in kpss.test(usgdp\_BC\_d1, "Trend", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: usgdp\_BC\_d1  
## KPSS Trend = 0.034928, Truncation lag parameter = 14, p-value = 0.1

usgdp\_BC\_d1 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 14 lags.   
##   
## Value of test-statistic is: 0.2395   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

usgdp\_BC\_d1 %>% ur.kpss(type="tau", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 14 lags.   
##   
## Value of test-statistic is: 0.0349   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

#### KPSS Test Transformed and 2nd Differenced data series: Box-Cox order(2):

kpss.test(usgdp\_BC\_d2, "Level", lshort = F)

## Warning in kpss.test(usgdp\_BC\_d2, "Level", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: usgdp\_BC\_d2  
## KPSS Level = 0.035084, Truncation lag parameter = 14, p-value = 0.1

kpss.test(usgdp\_BC\_d2, "Trend", lshort = F)

## Warning in kpss.test(usgdp\_BC\_d2, "Trend", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: usgdp\_BC\_d2  
## KPSS Trend = 0.034844, Truncation lag parameter = 14, p-value = 0.1

usgdp\_BC\_d2 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 14 lags.   
##   
## Value of test-statistic is: 0.0351   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

usgdp\_BC\_d2 %>% ur.kpss(type="tau", lags="long") %>% summary()

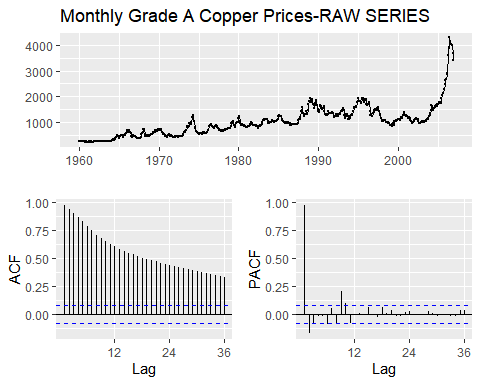
##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 14 lags.   
##   
## Value of test-statistic is: 0.0348   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

The KPSS test indicates that one differencing order(1) was sufficient to make the **Box-Cox transformed** data stationary.

### C) mcopper

**The trendin timeseries ‘mcopper’ looks nearly linear.I will check stationary status to determine the order of differencing.**

ggtsdisplay(mcopper, main = "Monthly Grade A Copper Prices-RAW SERIES")



#### Ljung-box test-RAW data series:

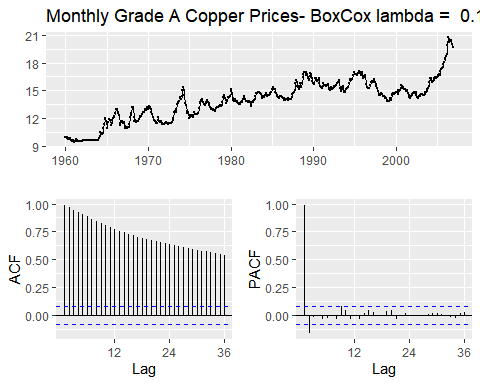
Box.test(mcopper, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: mcopper  
## X-squared = 539.54, df = 1, p-value < 2.2e-16

As we seet hat the p-value is low (p-value < 2.2e-16), the Null hypothesis is REJECTED (Autocorrelation exists).

#### Box-Cox transformation:

mcopper\_lambda <- round(BoxCox.lambda(mcopper),5)  
mcopper\_BC <- BoxCox(mcopper, mcopper\_lambda)  
ggtsdisplay(mcopper\_BC, main = paste("Monthly Grade A Copper Prices- BoxCox lambda = ",  
 mcopper\_lambda))



Time series data with Box-Cox transformation still gives non-stationary data(ADF plot)

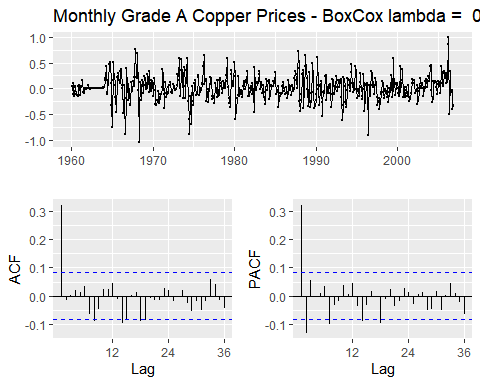
Box.test(mcopper\_BC, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: mcopper\_BC  
## X-squared = 550.42, df = 1, p-value < 2.2e-16

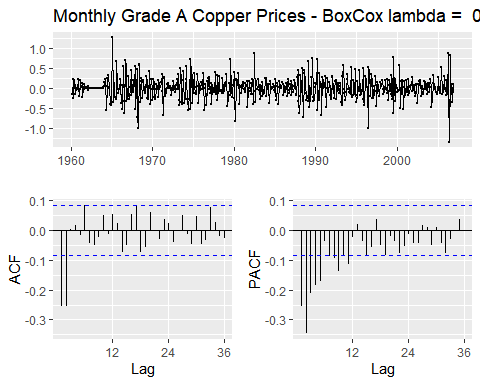
Because the p-value is still low (p-value < 2.2e-16), the Null hypothesis is REJECTED (Autocorrelation exists).

#### Differenced Box-Cox Data:

mcopper\_BC\_ndiffs <- ndiffs(mcopper\_BC)  
ggtsdisplay(diff(mcopper\_BC), main=paste("Monthly Grade A Copper Prices - BoxCox lambda = ",  
 mcopper\_lambda," - first difference"))



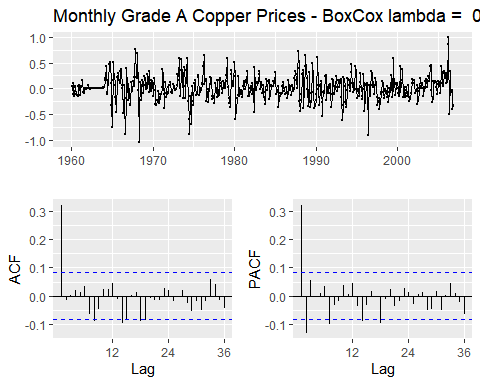
ggtsdisplay(diff(diff(mcopper\_BC)), main=paste("Monthly Grade A Copper Prices - BoxCox lambda = ",  
 mcopper\_lambda," - second difference"))



I’m going to test two different differencing : orders (1) and (2)

#### Ljung-box test-Transformed and Differenced data series: Box-Cox order(1):

mcopper\_BC %>% diff() -> mcopper\_BC\_d1  
ggtsdisplay(mcopper\_BC\_d1, main = paste("Monthly Grade A Copper Prices - BoxCox lambda = ",  
 mcopper\_lambda, "first difference"))



#Ljung-box test:  
Box.test(mcopper\_BC\_d1, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: mcopper\_BC\_d1  
## X-squared = 57.518, df = 1, p-value = 3.353e-14

The p-value is still low (p-value = 3.353e-14), the Null hypothesis is REJECTED (Autocorrelation exists).

#### KPSS Test Transformed and Differenced data series: Box-Cox order(1):

kpss.test(mcopper\_BC\_d1, "Level", lshort = F)

## Warning in kpss.test(mcopper\_BC\_d1, "Level", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: mcopper\_BC\_d1  
## KPSS Level = 0.062912, Truncation lag parameter = 18, p-value = 0.1

kpss.test(mcopper\_BC\_d1, "Trend", lshort = F)

## Warning in kpss.test(mcopper\_BC\_d1, "Trend", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: mcopper\_BC\_d1  
## KPSS Trend = 0.057247, Truncation lag parameter = 18, p-value = 0.1

mcopper\_BC\_d1 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 18 lags.   
##   
## Value of test-statistic is: 0.0629   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

mcopper\_BC\_d1 %>% ur.kpss(type="tau", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 18 lags.   
##   
## Value of test-statistic is: 0.0572   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

#### KPSS Test Transformed and Differenced data series: Box-Cox order(2):

mcopper\_BC %>% diff() %>% diff() -> mcopper\_BC\_d2  
kpss.test(mcopper\_BC\_d2, "Level", lshort = F)

## Warning in kpss.test(mcopper\_BC\_d2, "Level", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: mcopper\_BC\_d2  
## KPSS Level = 0.024273, Truncation lag parameter = 18, p-value = 0.1

kpss.test(mcopper\_BC\_d2, "Trend", lshort = F)

## Warning in kpss.test(mcopper\_BC\_d2, "Trend", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: mcopper\_BC\_d2  
## KPSS Trend = 0.019782, Truncation lag parameter = 18, p-value = 0.1

mcopper\_BC\_d2 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 18 lags.   
##   
## Value of test-statistic is: 0.0243   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

mcopper\_BC\_d2 %>% ur.kpss(type="tau", lags="long") %>% summary()

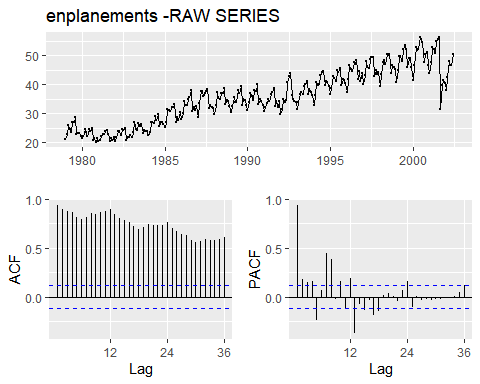
##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 18 lags.   
##   
## Value of test-statistic is: 0.0198   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

**The KPSS test indicates that one differencing ,order(1), should be used to make the mcopper data stationary**

### D) enplanements

**The graph of the enplanements raw data shows strong autocorrelation as well as seasonality.I will check stationary status to determine the order of differencing.**

ggtsdisplay(enplanements, main = "enplanements -RAW SERIES")



#### Ljung-box test-RAW data series:

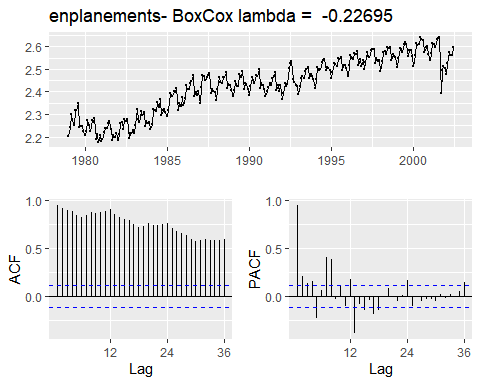
Box.test(enplanements, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: enplanements  
## X-squared = 249.22, df = 1, p-value < 2.2e-16

As we seet hat the p-value is low (p-value < 2.2e-16), the Null hypothesis is REJECTED (Autocorrelation exists).

#### Box-Cox transformation:

enplanements\_lambda <- round(BoxCox.lambda(enplanements),5)  
enplanements\_BC <- BoxCox(enplanements, enplanements\_lambda)  
ggtsdisplay(enplanements\_BC, main = paste("enplanements- BoxCox lambda = ",  
 enplanements\_lambda))



Time series data with Box-Cox transformation still gives non-stationary data(ADF plot).

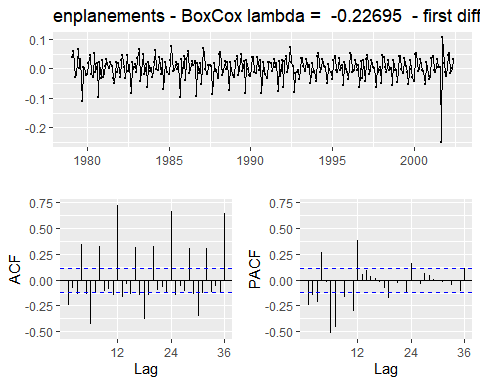
Box.test(enplanements\_BC, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: enplanements\_BC  
## X-squared = 253.7, df = 1, p-value < 2.2e-16

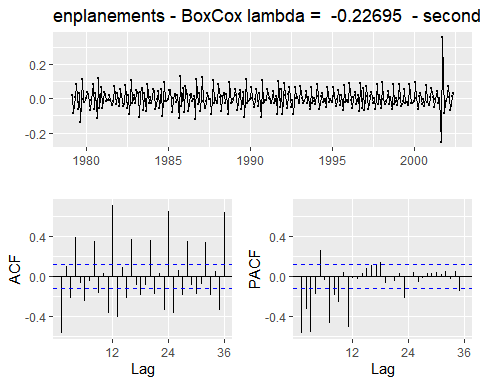
Because the p-value is still low (p-value < 2.2e-16), the Null hypothesis is REJECTED (Autocorrelation exists).

#### Differenced Box-Cox Data:

enplanements\_BC\_ndiffs <- ndiffs(enplanements\_BC)  
ggtsdisplay(diff(enplanements\_BC), main=paste("enplanements - BoxCox lambda = ",  
 enplanements\_lambda," - first difference"))



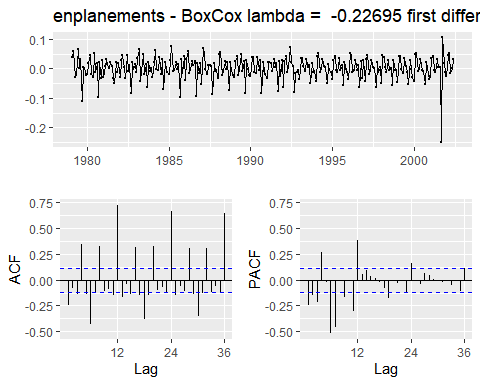
ggtsdisplay(diff(diff(enplanements\_BC)), main=paste("enplanements - BoxCox lambda = ",  
 enplanements\_lambda," - second difference"))



The graph of the Box-Cox transformed data still shows strong autocorrelation and seasonality I’m going to test two different differencing : orders (1) and (2)

#### Ljung-box test-Transformed and Differenced data series: Box-Cox order(1):

enplanements\_BC %>% diff() -> enplanements\_BC\_d1  
ggtsdisplay(enplanements\_BC\_d1, main = paste("enplanements - BoxCox lambda = ",  
 enplanements\_lambda, "first difference"))



#Ljung-box test:  
Box.test(enplanements\_BC\_d1, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: enplanements\_BC\_d1  
## X-squared = 16.55, df = 1, p-value = 4.739e-05

The p-value is still low (p-value = 4.739e-05), the Null hypothesis is REJECTED (Autocorrelation exists).

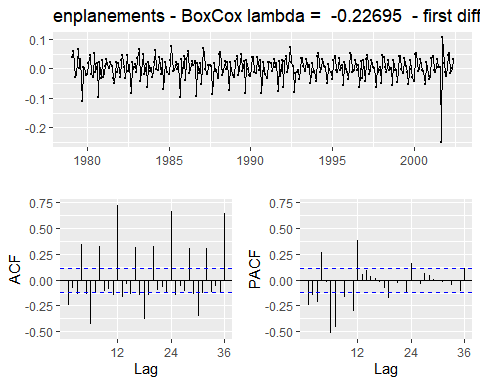
enplanements\_BC\_nsdiffs <- nsdiffs(enplanements\_BC)  
print(paste("Number of SEASONAL differences : enplanements\_BC:", enplanements\_BC\_nsdiffs))

## [1] "Number of SEASONAL differences : enplanements\_BC: 1"

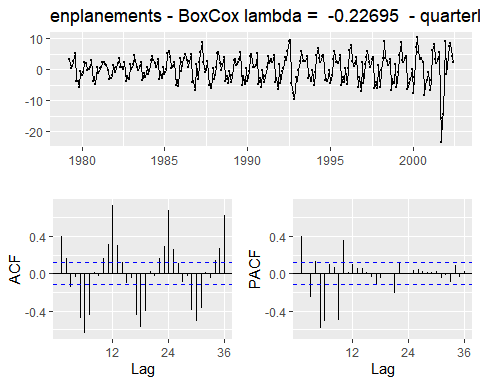
enplanements\_BC\_ndiffs <- ndiffs(enplanements\_BC)  
print(paste("Number of differences : enplanements\_BC:",   
 enplanements\_BC\_ndiffs))

## [1] "Number of differences : enplanements\_BC: 1"

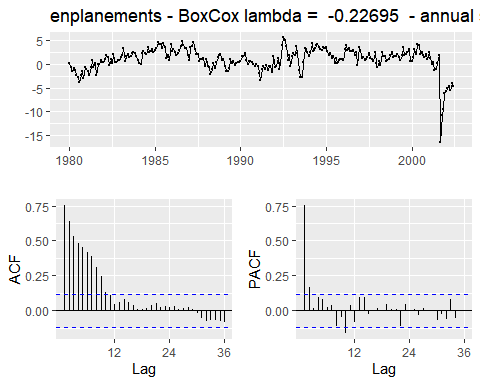
ggtsdisplay(diff(enplanements\_BC),   
 main=paste("enplanements - BoxCox lambda = ",  
 enplanements\_lambda," - first difference"))



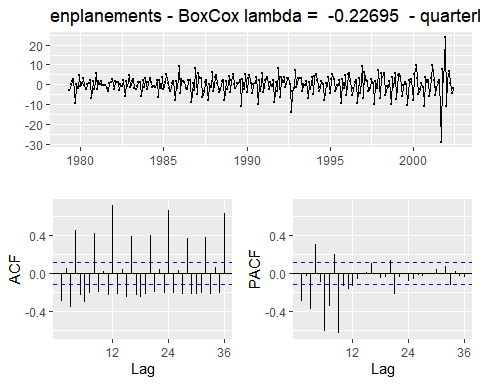
ggtsdisplay(diff(enplanements,lag=3),   
 main=paste("enplanements - BoxCox lambda = ",  
 enplanements\_lambda," - quarterly seasonal diff"))



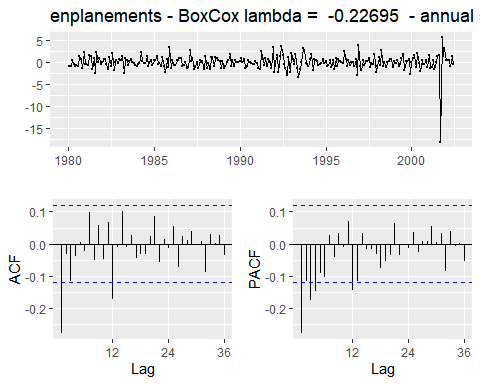
ggtsdisplay(diff(enplanements,lag=12),   
 main=paste("enplanements - BoxCox lambda = ",  
 enplanements\_lambda," - annual seasonal diff"))



ggtsdisplay(diff(diff(enplanements,lag=3),lag=1),   
 main=paste("enplanements - BoxCox lambda = ",  
 enplanements\_lambda," - quarterly seasonal diff + first diff"))



ggtsdisplay(diff(diff(enplanements,lag=12),lag=1),   
 main=paste("enplanements - BoxCox lambda = ",  
 enplanements\_lambda," - annual seasonal diff + first diff"))

 The result of Box-Cox transformation with annual seasonal differencing, looks like to provide the best result.

#### KPSS Test Transformed and annualy seasonal Differenced data series: Box-Cox order(1):

enplanements\_BC %>% diff(lag=12) %>% diff(lag=1) -> enplanements\_BC\_s1\_d1  
kpss.test(enplanements\_BC\_s1\_d1, "Level", lshort = F)

## Warning in kpss.test(enplanements\_BC\_s1\_d1, "Level", lshort = F): p-value  
## greater than printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: enplanements\_BC\_s1\_d1  
## KPSS Level = 0.074682, Truncation lag parameter = 15, p-value = 0.1

kpss.test(enplanements\_BC\_s1\_d1, "Trend", lshort = F)

## Warning in kpss.test(enplanements\_BC\_s1\_d1, "Trend", lshort = F): p-value  
## greater than printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: enplanements\_BC\_s1\_d1  
## KPSS Trend = 0.03852, Truncation lag parameter = 15, p-value = 0.1

enplanements\_BC\_s1\_d1 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 15 lags.   
##   
## Value of test-statistic is: 0.0747   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

enplanements\_BC\_s1\_d1 %>% ur.kpss(type="tau", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 15 lags.   
##   
## Value of test-statistic is: 0.0385   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

#### KPSS Test Transformed and Differenced data series: Box-Cox order(2):

enplanements\_BC %>% diff(lag=12) %>% diff(lag=2) -> enplanements\_BC\_s1\_d2  
kpss.test(enplanements\_BC\_s1\_d2, "Level", lshort = F)

## Warning in kpss.test(enplanements\_BC\_s1\_d2, "Level", lshort = F): p-value  
## greater than printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: enplanements\_BC\_s1\_d2  
## KPSS Level = 0.10006, Truncation lag parameter = 15, p-value = 0.1

kpss.test(enplanements\_BC\_s1\_d2, "Trend", lshort = F)

## Warning in kpss.test(enplanements\_BC\_s1\_d2, "Trend", lshort = F): p-value  
## greater than printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: enplanements\_BC\_s1\_d2  
## KPSS Trend = 0.040659, Truncation lag parameter = 15, p-value = 0.1

enplanements\_BC\_s1\_d2 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 15 lags.   
##   
## Value of test-statistic is: 0.1001   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

enplanements\_BC\_s1\_d2 %>% ur.kpss(type="tau", lags="long") %>% summary()

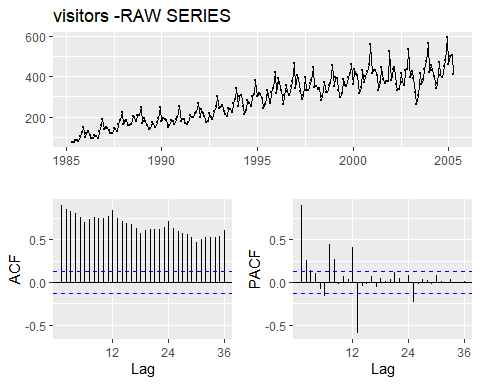
##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 15 lags.   
##   
## Value of test-statistic is: 0.0407   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

**The KPSS test indicates that annual seasonality and first differences,order(1), should be used to make the enplanements data stationary**

### E) visitors

**The graph of the visitors raw data shows strong autocorrelation as well as seasonality.I will check stationary status to determine the order of differencing.**

ggtsdisplay(visitors, main = "visitors -RAW SERIES")



#### Ljung-box test-RAW data series:

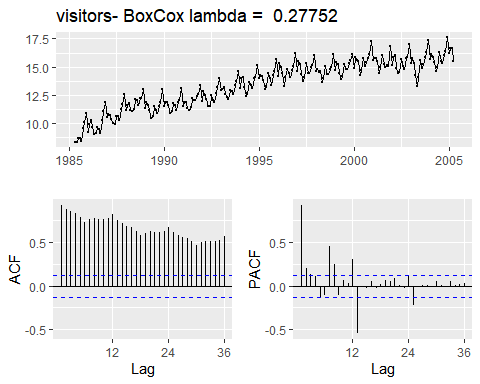
Box.test(visitors, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: visitors  
## X-squared = 195.64, df = 1, p-value < 2.2e-16

As we seet hat the p-value is low (p-value < 2.2e-16), the Null hypothesis is REJECTED (Autocorrelation exists).

#### Box-Cox transformation:

visitors\_lambda <- round(BoxCox.lambda(visitors),5)  
visitors\_BC <- BoxCox(visitors, visitors\_lambda)  
ggtsdisplay(visitors\_BC, main = paste("visitors- BoxCox lambda = ",  
 visitors\_lambda))



Time series data with Box-Cox transformation still gives non-stationary data(ADF plot).

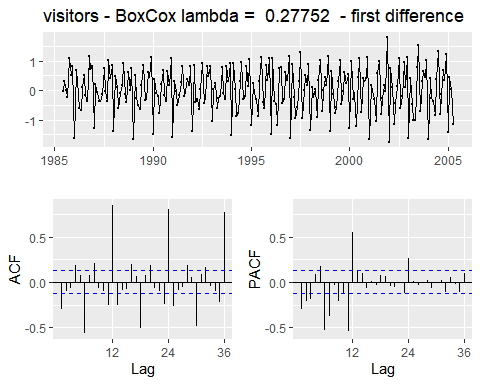
Box.test(visitors\_BC, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: visitors\_BC  
## X-squared = 205.74, df = 1, p-value < 2.2e-16

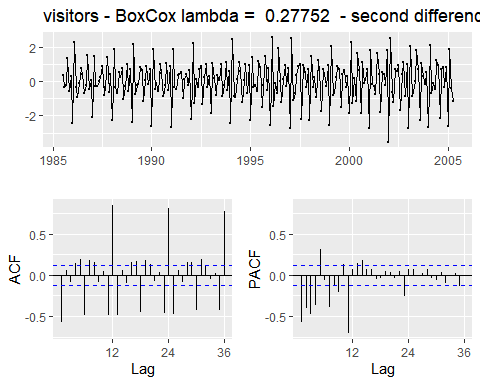
Because the p-value is still low (p-value < 2.2e-16), the Null hypothesis is REJECTED (Autocorrelation exists).

#### Differenced Box-Cox Data:

visitors\_BC\_ndiffs <- ndiffs(visitors\_BC)  
ggtsdisplay(diff(visitors\_BC), main=paste("visitors - BoxCox lambda = ",  
 visitors\_lambda," - first difference"))



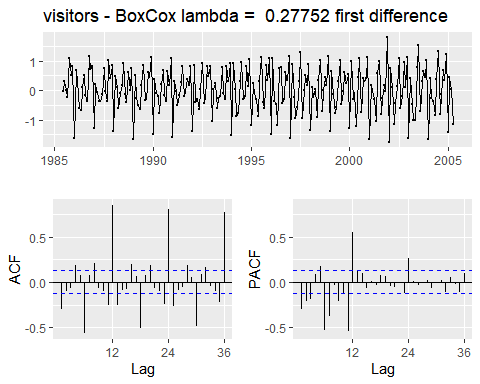
ggtsdisplay(diff(diff(visitors\_BC)), main=paste("visitors - BoxCox lambda = ",  
 visitors\_lambda," - second difference"))



The graph of the Box-Cox transformed data still shows strong autocorrelation and seasonality .I’m going to test two different differencing : orders (1) and (2)

#### Ljung-box test-Transformed and Differenced data series: Box-Cox order(1):

visitors\_BC %>% diff() -> visitors\_BC\_d1  
ggtsdisplay(visitors\_BC\_d1, main = paste("visitors - BoxCox lambda = ",  
 visitors\_lambda, "first difference"))



#Ljung-box test:  
Box.test(visitors\_BC\_d1, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: visitors\_BC\_d1  
## X-squared = 21.922, df = 1, p-value = 2.839e-06

The p-value is still low (p-value = 9.987e-05), the Null hypothesis is REJECTED (Autocorrelation exists).

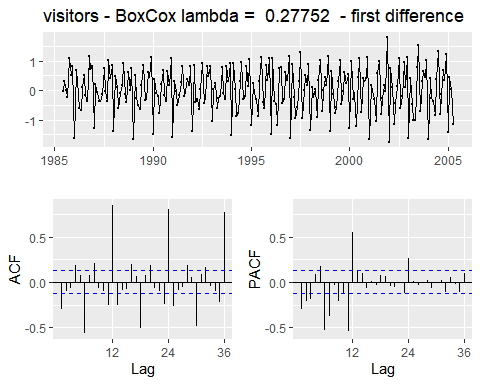
visitors\_BC\_nsdiffs <- nsdiffs(visitors\_BC)  
print(paste("Number of SEASONAL differences for visitors\_BC:", visitors\_BC\_nsdiffs))

## [1] "Number of SEASONAL differences for visitors\_BC: 1"

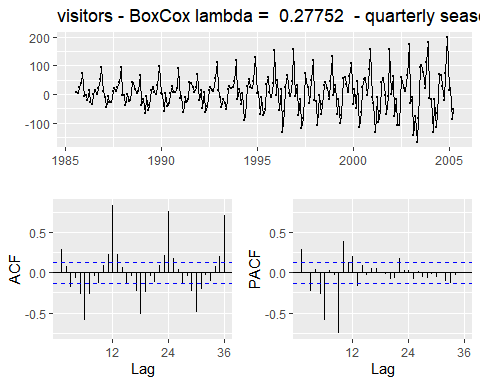
visitors\_BC\_ndiffs <- ndiffs(visitors\_BC)  
print(paste("Number of differences suggested visitors\_BC:",   
 visitors\_BC\_ndiffs))

## [1] "Number of differences suggested visitors\_BC: 1"

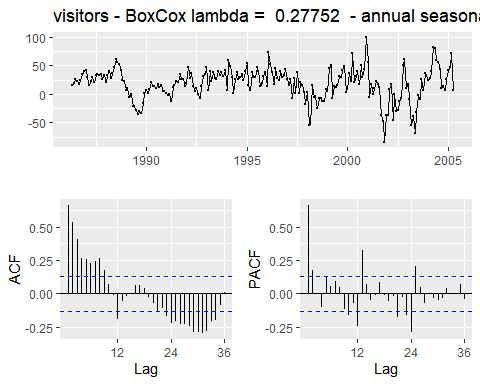
ggtsdisplay(diff(visitors\_BC),   
 main=paste("visitors - BoxCox lambda = ",  
 visitors\_lambda," - first difference"))



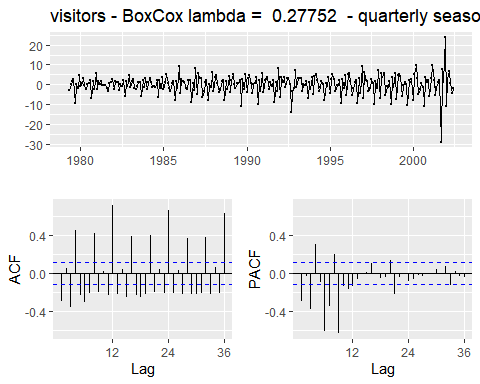
ggtsdisplay(diff(visitors,lag=3),   
 main=paste("visitors - BoxCox lambda = ",  
 visitors\_lambda," - quarterly seasonal diff"))



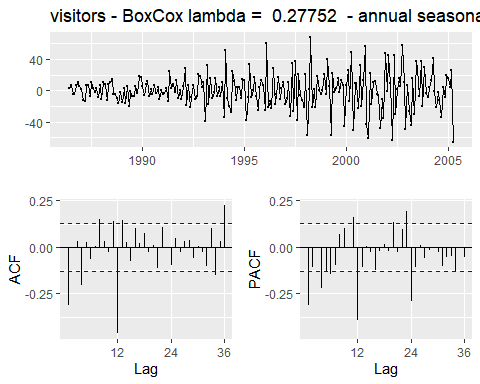
ggtsdisplay(diff(visitors,lag=12),   
 main=paste("visitors - BoxCox lambda = ",  
 visitors\_lambda," - annual seasonal diff"))



ggtsdisplay(diff(diff(enplanements,lag=3),lag=1),   
 main=paste("visitors - BoxCox lambda = ",  
 visitors\_lambda," - quarterly seasonal diff + first diff"))



ggtsdisplay(diff(diff(visitors,lag=12),lag=1),   
 main=paste("visitors - BoxCox lambda = ",  
 visitors\_lambda," - annual seasonal diff + first diff"))

 The result of Box-Cox transformation with annual seasonal differencing, looks like to provide the best result.Howevers some lags (1,12) are exremly high.

#### KPSS Test Transformed and annualy seasonal Differenced data series: Box-Cox order(1):

visitors\_BC %>% diff(lag=12) %>% diff(lag=1) -> visitors\_BC\_s1\_d1  
kpss.test(visitors\_BC\_s1\_d1, "Level", lshort = F)

## Warning in kpss.test(visitors\_BC\_s1\_d1, "Level", lshort = F): p-value greater  
## than printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: visitors\_BC\_s1\_d1  
## KPSS Level = 0.032281, Truncation lag parameter = 14, p-value = 0.1

kpss.test(visitors\_BC\_s1\_d1, "Trend", lshort = F)

## Warning in kpss.test(visitors\_BC\_s1\_d1, "Trend", lshort = F): p-value greater  
## than printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: visitors\_BC\_s1\_d1  
## KPSS Trend = 0.032061, Truncation lag parameter = 14, p-value = 0.1

visitors\_BC\_s1\_d1 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 14 lags.   
##   
## Value of test-statistic is: 0.0323   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

visitors\_BC\_s1\_d1 %>% ur.kpss(type="tau", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 14 lags.   
##   
## Value of test-statistic is: 0.0321   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

#### KPSS Test Transformed and Differenced data series: Box-Cox order(2):

visitors\_BC %>% diff(lag=12) %>% diff(lag=2) -> visitors\_BC\_s1\_d2  
kpss.test(visitors\_BC\_s1\_d2, "Level", lshort = F)

## Warning in kpss.test(visitors\_BC\_s1\_d2, "Level", lshort = F): p-value greater  
## than printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: visitors\_BC\_s1\_d2  
## KPSS Level = 0.047141, Truncation lag parameter = 14, p-value = 0.1

kpss.test(visitors\_BC\_s1\_d2, "Trend", lshort = F)

## Warning in kpss.test(visitors\_BC\_s1\_d2, "Trend", lshort = F): p-value greater  
## than printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: visitors\_BC\_s1\_d2  
## KPSS Trend = 0.035599, Truncation lag parameter = 14, p-value = 0.1

visitors\_BC\_s1\_d2 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 14 lags.   
##   
## Value of test-statistic is: 0.0471   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

visitors\_BC\_s1\_d2 %>% ur.kpss(type="tau", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 14 lags.   
##   
## Value of test-statistic is: 0.0356   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

**The KPSS test shows that annual seasonality with Box Cox transformation and first differences,order(1), should be used to make the visitors data stationary**

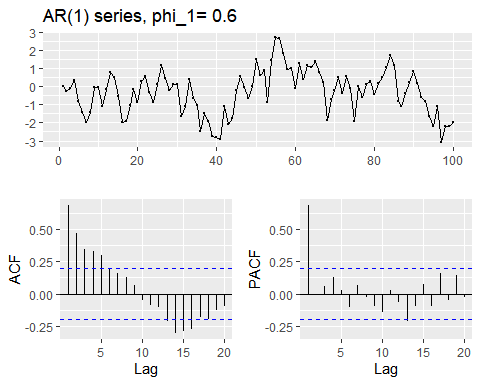
## Question 8.6

Use R to simulate and plot some data from simple ARIMA models.

### A)

Use the following R code to generate data from an AR(1) model with and . The process starts with .

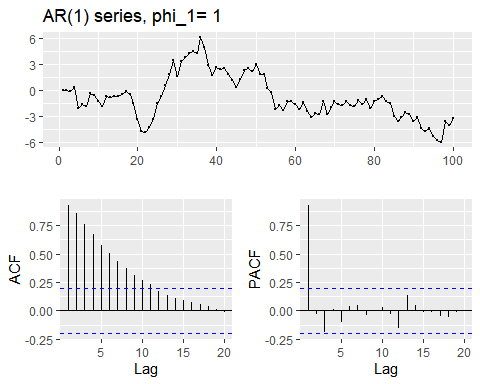
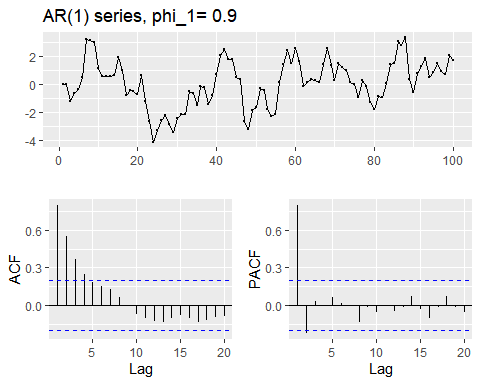
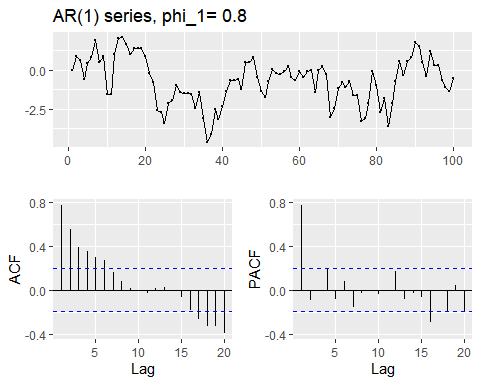
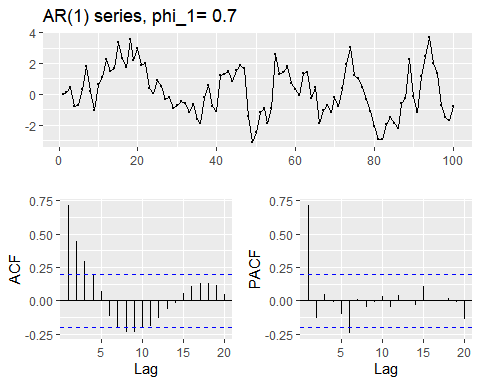
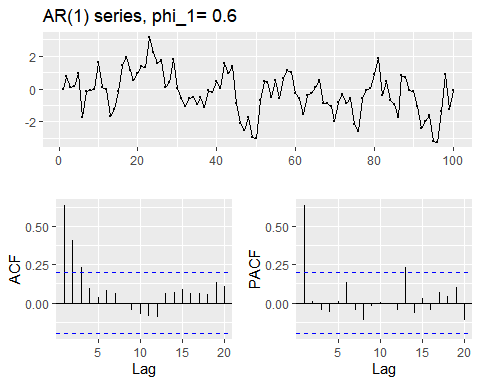
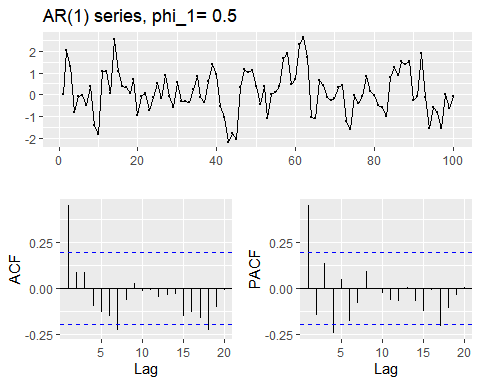
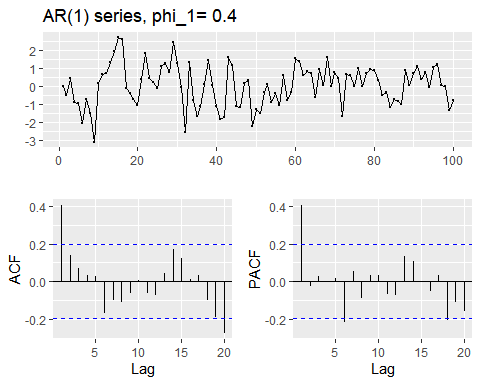
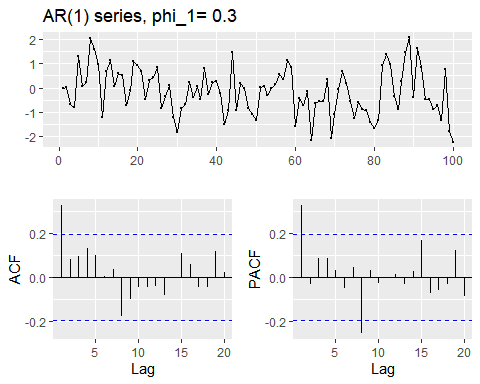
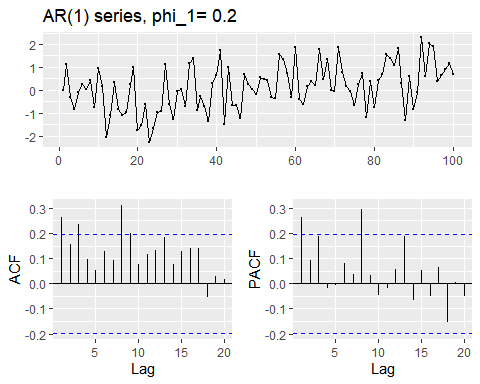
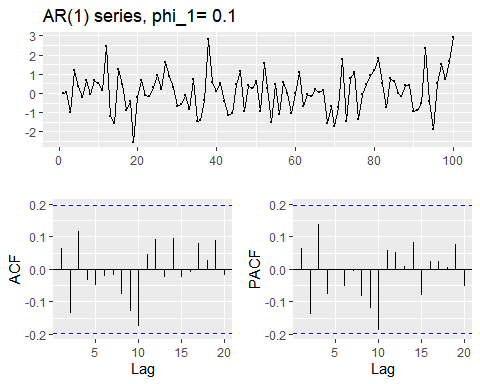
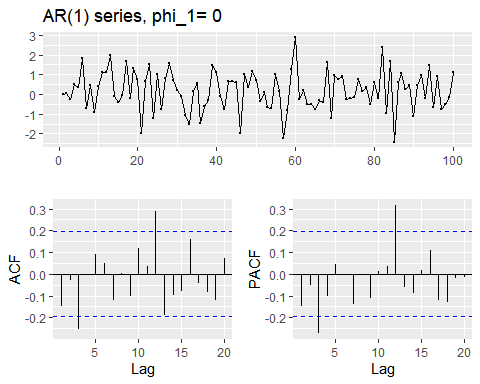
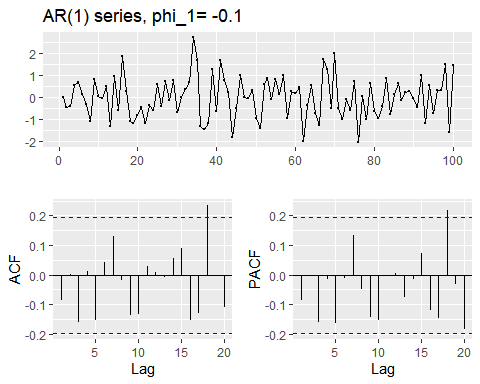
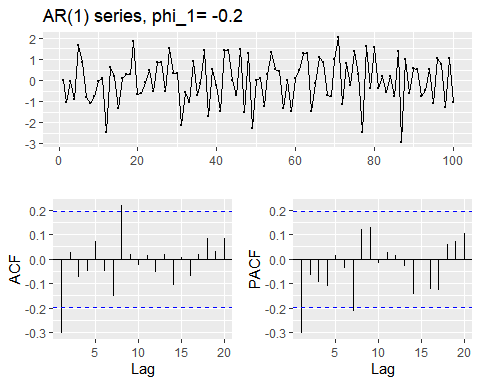
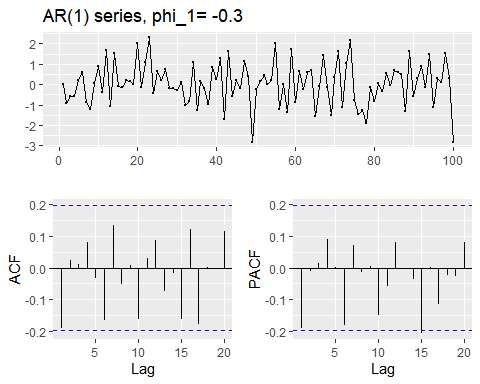
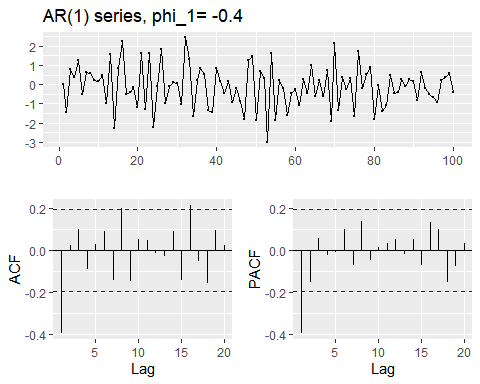
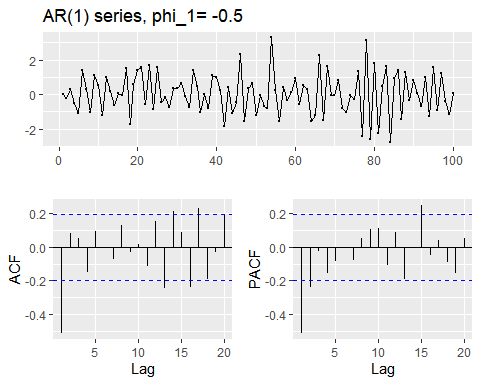
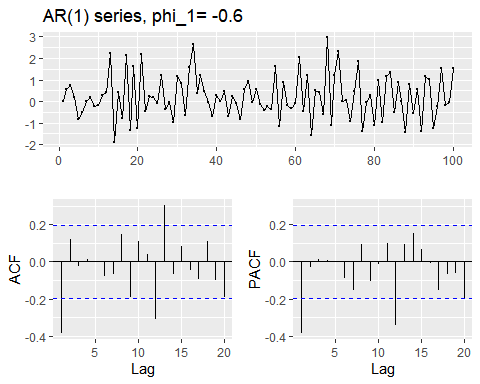
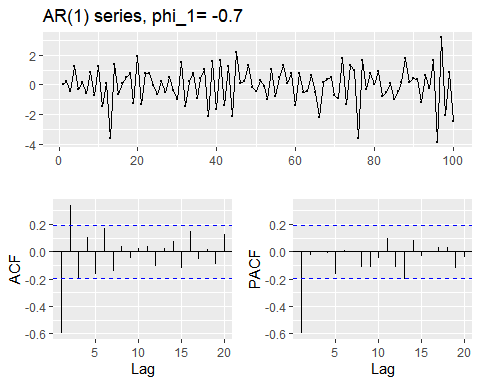
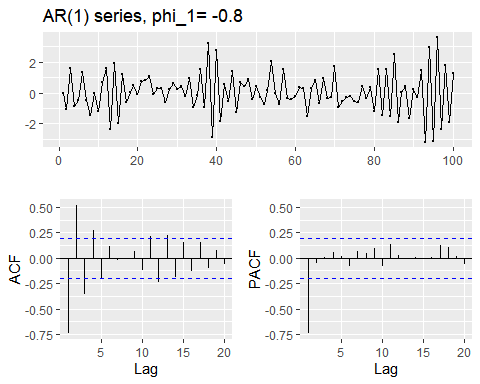
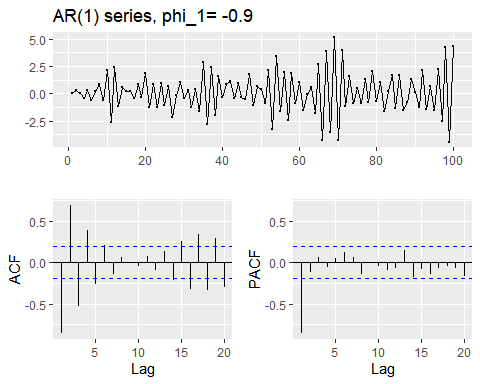
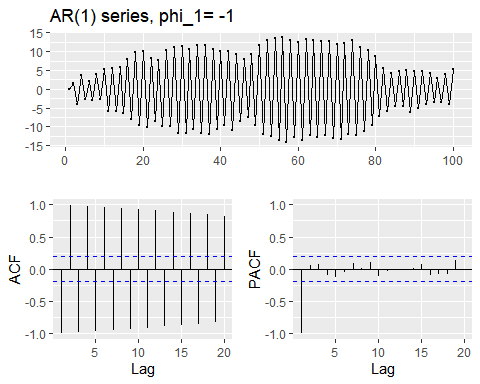
AR1 <- function(phi)  
{  
 y <- ts(numeric(100))  
 e <- rnorm(100)  
 for(i in 2:100)  
 y[i] <- phi\*y[i-1] + e[i]  
 return(y)  
}  
phi\_1 <- 0.60  
ggtsdisplay(AR1(phi\_1),main=paste("AR(1) series, phi\_1=",phi\_1))



### B)

Produce a time plot for the series. How does the plot change as you change ?

set.seed(12345678)  
results = numeric()  
phi\_seq <- seq(from=-1,to=1,by = 0.1)  
for (phi\_1 in phi\_seq) {  
 AR1\_data <- AR1(phi\_1)  
 AR1\_stats <- c(Phi\_1 = phi\_1, summary(AR1\_data),StDev=sd(AR1\_data),SdDiff=sd(diff(AR1\_data)))  
 AR1\_main <- paste("AR(1) series, phi\_1=",phi\_1)  
 ggtsdisplay(AR1\_data,main=AR1\_main)  
 results <- rbind(results,AR1\_stats)  
}



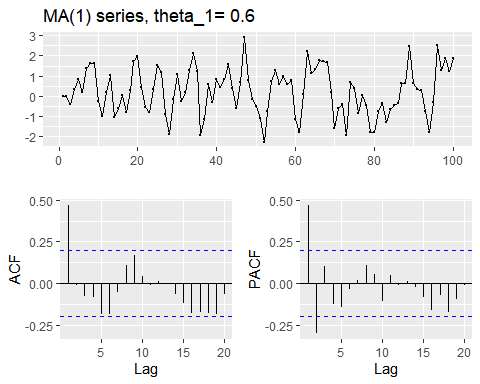
results

## Phi\_1 Min. 1st Qu. Median Mean 3rd Qu.  
## AR1\_stats -1.0 -14.062104 -9.6874291 0.809199676 -0.019024279 9.9235214  
## AR1\_stats -0.9 -4.443402 -1.1535149 0.140064912 0.095010441 1.1571075  
## AR1\_stats -0.8 -3.199613 -0.6358272 -0.016910471 0.042523719 0.7985269  
## AR1\_stats -0.7 -3.859066 -0.6950367 0.119200314 -0.033535258 0.7557429  
## AR1\_stats -0.6 -1.886435 -0.4584182 0.028648389 0.221188607 0.8664897  
## AR1\_stats -0.5 -2.766594 -0.8245263 0.031009976 0.042509026 0.9132246  
## AR1\_stats -0.4 -2.990606 -0.8940255 -0.064001374 -0.101253074 0.5582459  
## AR1\_stats -0.3 -2.819765 -0.5961379 0.041449779 0.039690417 0.6627684  
## AR1\_stats -0.2 -2.930321 -0.7490779 0.040460473 -0.021635716 0.8507102  
## AR1\_stats -0.1 -2.013560 -0.6376682 0.007097826 -0.005197204 0.6251735  
## AR1\_stats 0.0 -2.419691 -0.4837971 0.175597239 0.185282167 0.9042786  
## AR1\_stats 0.1 -2.534696 -0.6939627 0.119884378 0.091004011 0.6623034  
## AR1\_stats 0.2 -2.270066 -0.6670993 0.182851309 0.119082644 0.7244989  
## AR1\_stats 0.3 -2.228579 -0.8331157 -0.119490306 -0.146005580 0.5389209  
## AR1\_stats 0.4 -3.097211 -0.7838577 0.058539522 0.032670534 0.7902225  
## AR1\_stats 0.5 -2.180207 -0.4394042 0.010093219 0.137006940 0.8185993  
## AR1\_stats 0.6 -3.235804 -0.9358224 -0.123176998 -0.232467066 0.5582484  
## AR1\_stats 0.7 -3.114780 -0.9351570 0.076418353 0.186031378 1.4078848  
## AR1\_stats 0.8 -4.604911 -1.5648592 -0.634706373 -0.701113205 0.4273913  
## AR1\_stats 0.9 -4.107645 -0.8585411 0.263087693 0.126812428 1.4168184  
## AR1\_stats 1.0 -5.992400 -2.4326263 -1.277668824 -0.822016997 0.4688218  
## Max. StDev SdDiff  
## AR1\_stats 13.653510 9.5012316 19.0611000  
## AR1\_stats 5.128250 1.8824653 3.6133126  
## AR1\_stats 3.590143 1.3208697 2.4699870  
## AR1\_stats 3.165380 1.2467545 2.2241655  
## AR1\_stats 2.957920 1.0117494 1.6856881  
## AR1\_stats 3.286173 1.1927434 2.0849159  
## AR1\_stats 2.437646 1.0867314 1.8239736  
## AR1\_stats 2.298705 1.0318553 1.5735992  
## AR1\_stats 2.029409 1.0736376 1.7381326  
## AR1\_stats 2.704765 0.9346972 1.3756997  
## AR1\_stats 2.877745 0.9946989 1.5105604  
## AR1\_stats 2.896166 1.0347156 1.3928242  
## AR1\_stats 2.289869 1.0006973 1.2175162  
## AR1\_stats 2.065666 0.9489532 1.0876201  
## AR1\_stats 2.682038 1.1278613 1.2356588  
## AR1\_stats 2.618272 1.0019356 1.0561522  
## AR1\_stats 3.129357 1.2708693 1.0924445  
## AR1\_stats 3.664113 1.5523935 1.1768035  
## AR1\_stats 2.065503 1.4541925 0.9896369  
## AR1\_stats 3.365740 1.6728732 1.0653087  
## AR1\_stats 6.131975 2.5611879 0.9524565

### C)

Write your own code to generate data from an MA(1) model with and .

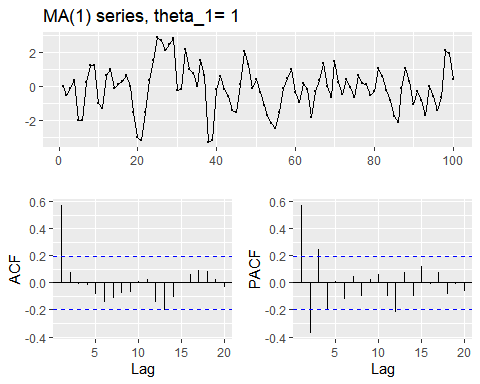
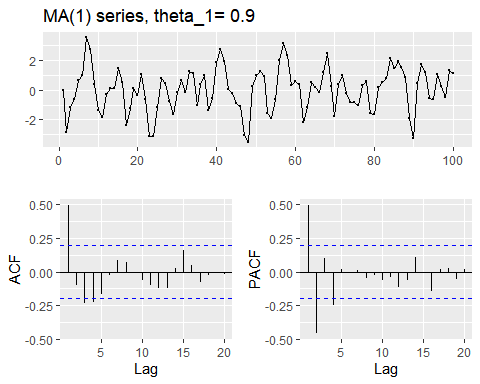
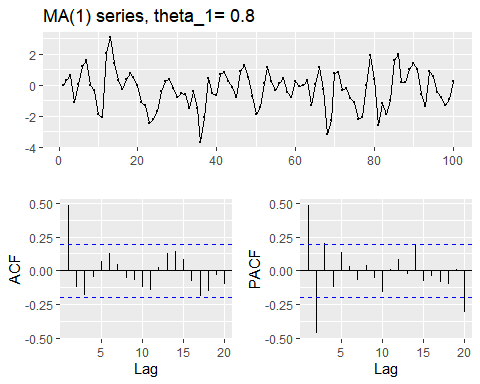
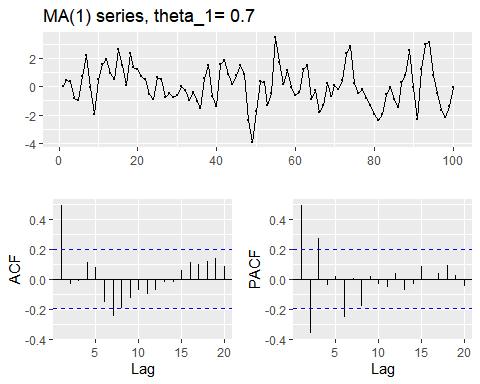
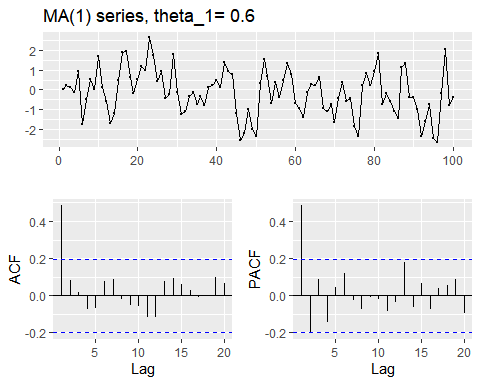
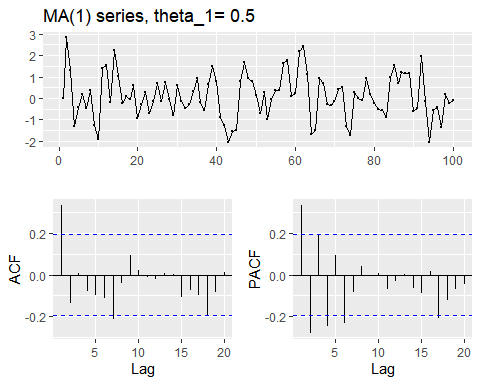
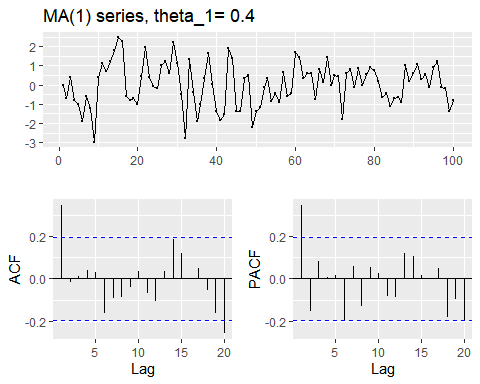
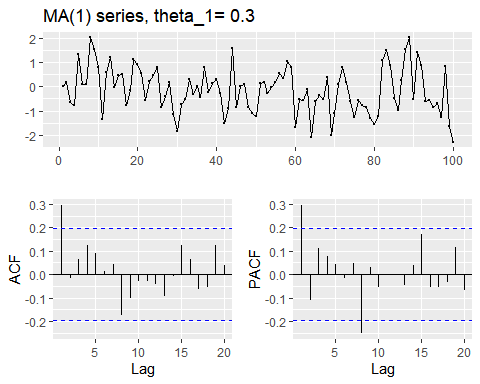
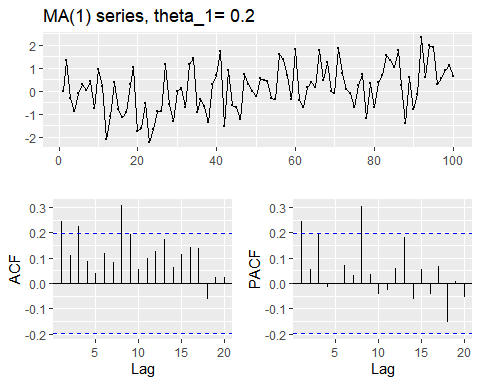
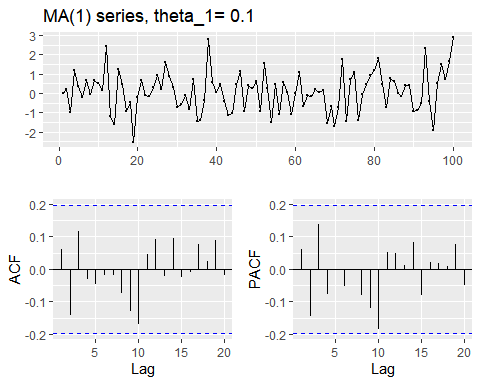
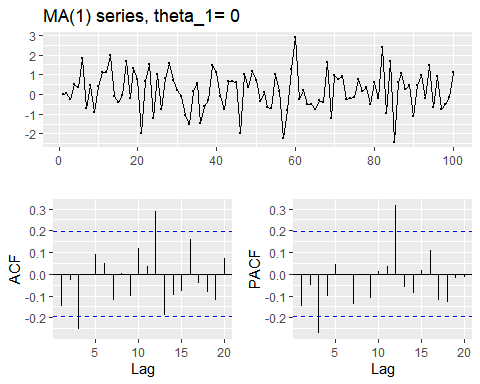
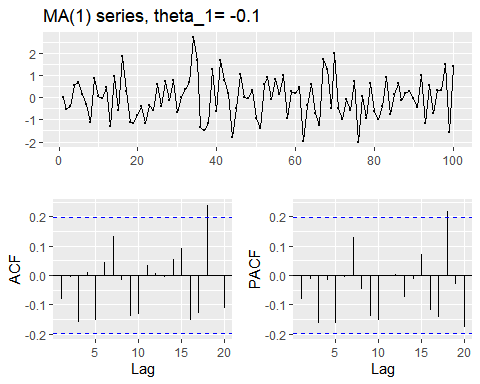
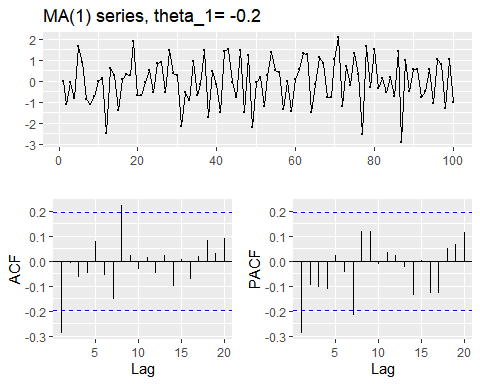
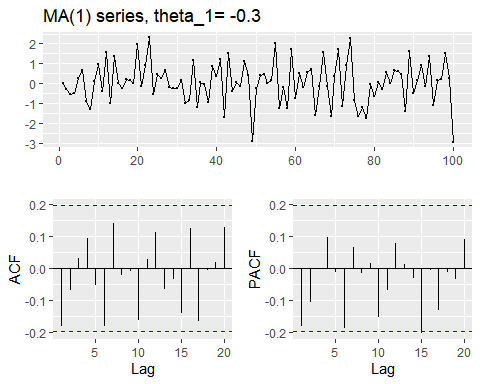
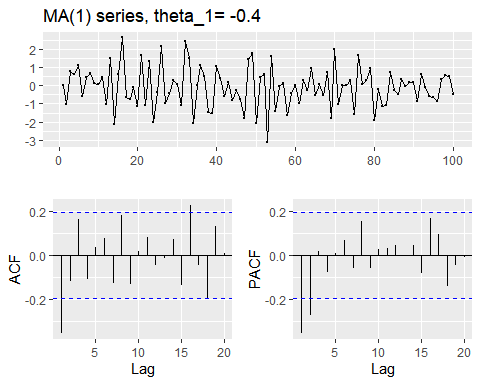
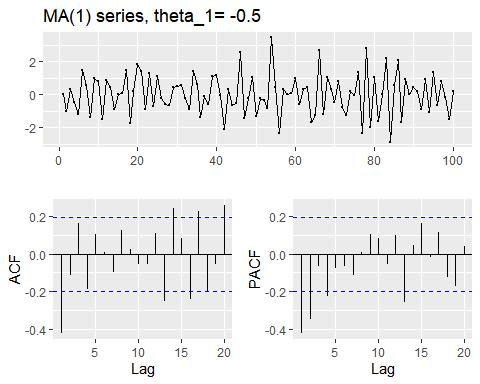
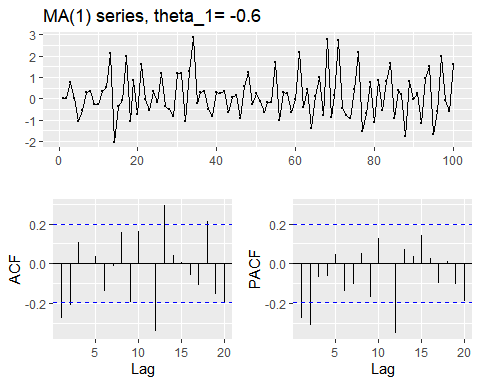
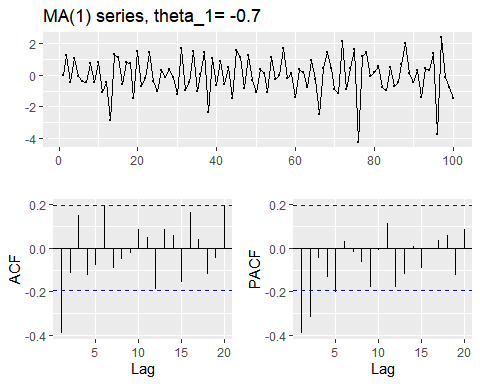
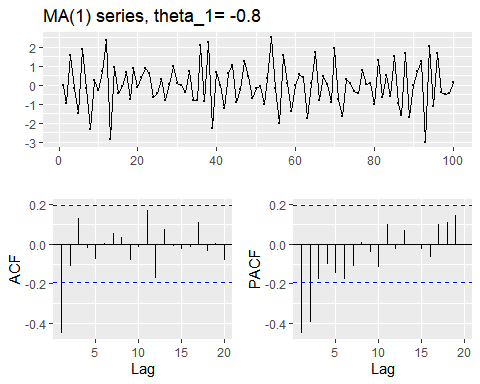
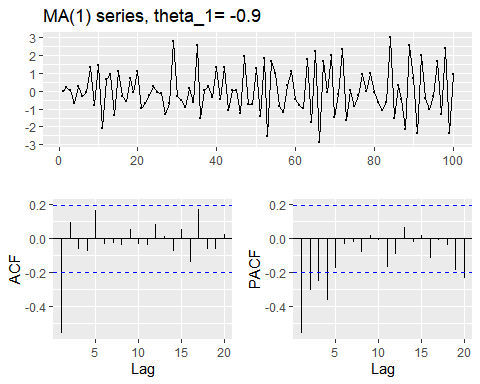
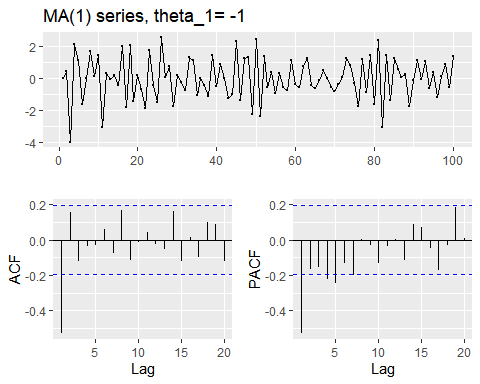
MA1 <- function(theta\_1)  
{  
 y <- ts(numeric(100))  
 e <- rnorm(100)  
 for(i in 2:100)  
 y[i] <- e[i] + theta\_1\*e[i-1]   
 return(y)  
}  
theta\_1 <- 0.60  
ggtsdisplay(MA1(theta\_1),main=paste("MA(1) series, theta\_1=",theta\_1))



### D)

Produce a time plot for the series. How does the plot change as you change ?

set.seed(12345678)  
results = numeric()  
theta\_seq <- seq(from=-1,to=1,by = 0.1)  
for (theta\_1 in theta\_seq) {  
 MA1\_data <- MA1(theta\_1)  
 MA1\_stats <- c(theta\_1 = theta\_1, summary(MA1\_data),StDev=sd(MA1\_data),SdDiff=sd(diff(MA1\_data)))  
 MA1\_main <- paste("MA(1) series, theta\_1=",theta\_1)  
 ggtsdisplay(MA1\_data,main=MA1\_main)  
 #print(MA1\_stats)  
 MA1\_stats\_results <- rbind(results,MA1\_stats)  
}



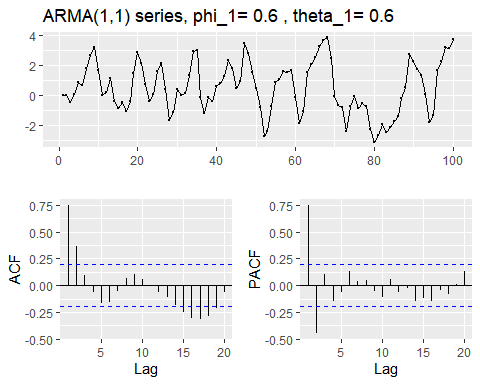
results

## numeric(0)

### E)

Generate data from an ARMA(1,1) model with , , and .

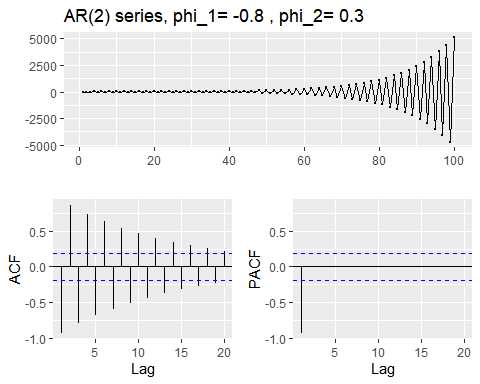
ARMA\_1\_1 <- function(phi\_1, theta\_1)  
{  
 y <- ts(numeric(100))  
 e <- rnorm(100)  
 for(i in 2:100)  
 y[i] <- phi\_1\*y[i-1] + theta\_1\*e[i-1] + e[i]  
 return(y)  
}  
phi\_1 <- 0.60  
theta\_1 <- 0.60  
ARMA\_1\_1\_result <- ARMA\_1\_1(phi\_1,theta\_1)  
ggtsdisplay(ARMA\_1\_1\_result,main=paste("ARMA(1,1) series, phi\_1=", phi\_1, ", theta\_1=",theta\_1))



### F)

Generate data from an AR(2) model with , , and .(Note that these parameters will give a **non-stationary** series.)

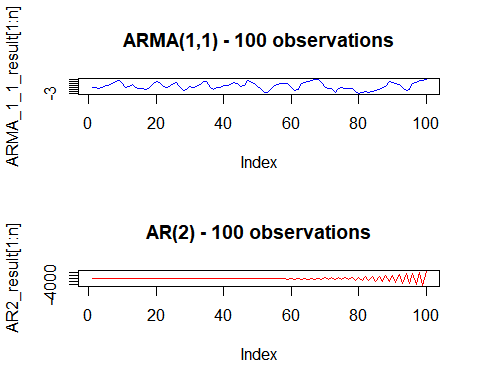
AR2 <- function(phi\_1, phi\_2)  
{  
 y <- ts(numeric(100))  
 e <- rnorm(100)  
 for(i in 3:100)  
 y[i] <- phi\_1\*y[i-1] + phi\_2\*y[i-2] + e[i]  
 return(y)  
}  
phi\_1 <- -0.80  
phi\_2 <- 0.30  
AR2\_result <- AR2(phi\_1,phi\_2)   
ggtsdisplay(AR2\_result,main=paste("AR(2) series, phi\_1=", phi\_1, ", phi\_2=",phi\_2))



### G)

Graph the latter two series and compare them.

n=100  
par(mfrow=c(2,1))  
plot(ARMA\_1\_1\_result[1:n], type="l", main="ARMA(1,1) - 100 observations",col="blue")  
plot(AR2\_result[1:n], type="l", main="AR(2) - 100 observations",col="red")



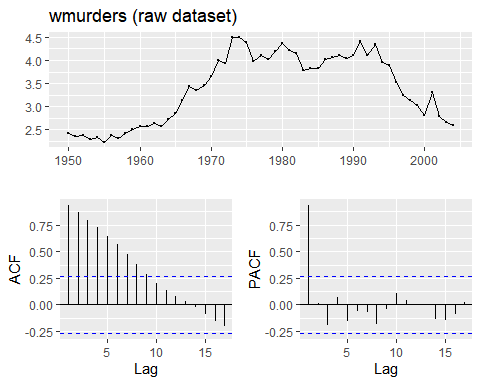
## Question 8.7

Consider wmurders, the number of women murdered each year (per 100,000 standard population) in the United States.

### A)

By studying appropriate graphs of the series in R, find an appropriate ARIMA(p,d,q) model for these data.

wmurders %>% ggtsdisplay(main="wmurders (raw dataset)")



The raw data series doesn’t indicates any trend (e.g., seasonality,extreme values.)There is a strong pattern in the ACF function which might cause also pattern in PACF

**(KPSS) test:**

library(tseries)  
kpss.test(wmurders, "Level", lshort = F)

## Warning in kpss.test(wmurders, "Level", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: wmurders  
## KPSS Level = 0.28623, Truncation lag parameter = 10, p-value = 0.1

kpss.test(wmurders, "Trend", lshort = F)

##   
## KPSS Test for Trend Stationarity  
##   
## data: wmurders  
## KPSS Trend = 0.1591, Truncation lag parameter = 10, p-value = 0.03909

wmurders %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 10 lags.   
##   
## Value of test-statistic is: 0.2862   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

wmurders %>% ur.kpss(type="tau", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 10 lags.   
##   
## Value of test-statistic is: 0.1591   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

Level passes the test, but Trend fails.

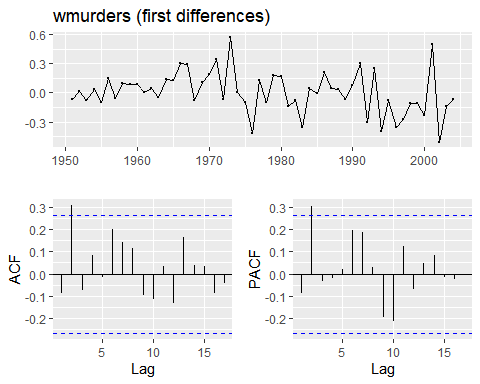
**ndiffs:**

ndiffs(wmurders)

## [1] 2

**I(1) first differences**

wmurders\_d1 <- wmurders %>% diff()   
wmurders\_d1 %>% ggtsdisplay(main="wmurders (first differences)")



**(KPSS) test:**

kpss.test(wmurders\_d1, "Level", lshort = F)

## Warning in kpss.test(wmurders\_d1, "Level", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: wmurders\_d1  
## KPSS Level = 0.33134, Truncation lag parameter = 10, p-value = 0.1

kpss.test(wmurders\_d1, "Trend", lshort = F)

## Warning in kpss.test(wmurders\_d1, "Trend", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: wmurders\_d1  
## KPSS Trend = 0.11631, Truncation lag parameter = 10, p-value = 0.1

wmurders\_d1 %>% ur.kpss(type="mu", lags="long") %>% summary()

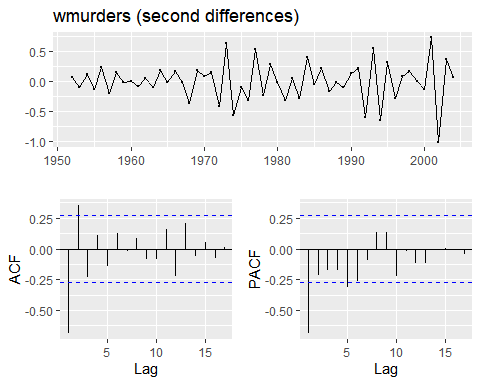
##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 10 lags.   
##   
## Value of test-statistic is: 0.3313   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

wmurders\_d1 %>% ur.kpss(type="tau", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 10 lags.   
##   
## Value of test-statistic is: 0.1163   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

**I(2) Second differences**

wmurders\_d2 <- wmurders %>% diff() %>% diff()  
wmurders\_d2 %>% ggtsdisplay(main="wmurders (second differences)")



**(KPSS) test:**

kpss.test(wmurders\_d2, "Level", lshort = F)

## Warning in kpss.test(wmurders\_d2, "Level", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: wmurders\_d2  
## KPSS Level = 0.12299, Truncation lag parameter = 10, p-value = 0.1

kpss.test(wmurders\_d2, "Trend", lshort = F)

## Warning in kpss.test(wmurders\_d2, "Trend", lshort = F): p-value greater than  
## printed p-value

##   
## KPSS Test for Trend Stationarity  
##   
## data: wmurders\_d2  
## KPSS Trend = 0.10938, Truncation lag parameter = 10, p-value = 0.1

wmurders\_d2 %>% ur.kpss(type="mu", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: mu with 10 lags.   
##   
## Value of test-statistic is: 0.123   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.347 0.463 0.574 0.739

wmurders\_d2 %>% ur.kpss(type="tau", lags="long") %>% summary()

##   
## #######################   
## # KPSS Unit Root Test #   
## #######################   
##   
## Test is of type: tau with 10 lags.   
##   
## Value of test-statistic is: 0.1094   
##   
## Critical value for a significance level of:   
## 10pct 5pct 2.5pct 1pct  
## critical values 0.119 0.146 0.176 0.216

The differecing I(2) passes the KPSS test.

The differecing term,d,is 2 so our model will be ARIMA(p,2,q).For the AR and MA terms,we need to look at the ACF and PASF plots.There is as spike in lag 1 so the model may be ARIMA(1,2,1)

### B)

Should you include a constant in the model? Explain.

WE should Not include constant in the model, because a constant would create a shift, and there is no shift visible in the raw data.

### C)

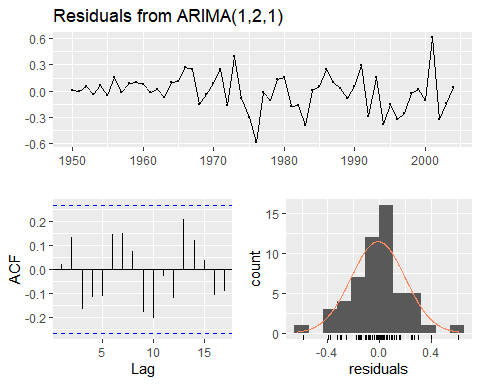
Write this model in terms of the backshift operator.

We omitted the constat term :

### D)

Fit the model using R and examine the residuals. Is the model satisfactory?

my\_arima <- Arima(wmurders, order=c(1,2,1))  
checkresiduals(my\_arima)



##   
## Ljung-Box test  
##   
## data: Residuals from ARIMA(1,2,1)  
## Q\* = 12.419, df = 8, p-value = 0.1335  
##   
## Model df: 2. Total lags used: 10

The residuals plot shows that it is normal distributed however p-values greater than 0.10. The model is satisfactory

## E)

Forecast three times ahead. Check your forecasts by hand to make sure that you know how they have been calculated.

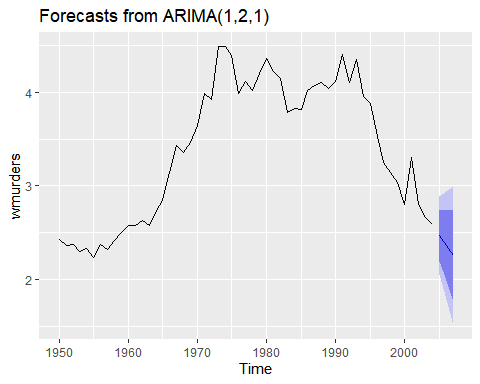
wmurders\_forecast <- forecast(my\_arima, h = 3)  
wmurders\_forecast

## Point Forecast Lo 80 Hi 80 Lo 95 Hi 95  
## 2005 2.470660 2.194836 2.746484 2.048824 2.892496  
## 2006 2.363106 1.986351 2.739862 1.786908 2.939304  
## 2007 2.252833 1.765391 2.740276 1.507354 2.998313

### F)

Create a plot of the series with forecasts and prediction intervals for the next three periods shown.

autoplot(wmurders\_forecast)



### G)

Does auto.arima() give the same model you have chosen? If not, which model do you think is better?

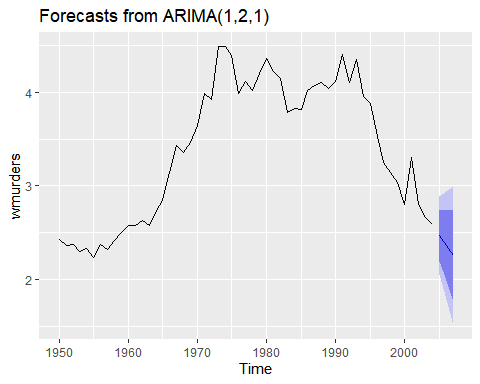
wmurders\_fit\_autoarima <- auto.arima(wmurders)  
wmurders\_fit\_autoarima

## Series: wmurders   
## ARIMA(1,2,1)   
##   
## Coefficients:  
## ar1 ma1  
## -0.2434 -0.8261  
## s.e. 0.1553 0.1143  
##   
## sigma^2 estimated as 0.04632: log likelihood=6.44  
## AIC=-6.88 AICc=-6.39 BIC=-0.97

wmurders\_forecast\_autoarima <- forecast(wmurders\_fit\_autoarima, h = 3)  
wmurders\_forecast\_autoarima

## Point Forecast Lo 80 Hi 80 Lo 95 Hi 95  
## 2005 2.470660 2.194836 2.746484 2.048824 2.892496  
## 2006 2.363106 1.986351 2.739862 1.786908 2.939304  
## 2007 2.252833 1.765391 2.740276 1.507354 2.998313

autoplot(wmurders\_forecast\_autoarima)



**The model selected is the same as chosen.**