

# Forecasting Retail Sales of US Clothing and Clothing Accessory Stores

Andrew Chow

November 24, 2020

## Contents

1. Abstract . . . . .	2
2. Introduction . . . . .	2
3. Analysis . . . . .	3
3.1 Exploratory Data Analysis . . . . .	3
3.2 Data Transformations . . . . .	4
3.3 ACF and PACF Analysis . . . . .	10
3.4 Model Diagnostics . . . . .	15
3.5 Forecasting . . . . .	26
3.5 Spectral Analysis . . . . .	28
4. Conclusion . . . . .	30
5. References . . . . .	30
Appendix . . . . .	31

## 1. Abstract

The dataset to be analysed and discussed within this paper is the monthly sales (in millions of dollars) of clothing and clothing accessory stores from 1992 to 2019. Using this data set, our goal is to forecast the sales of clothing and clothing accessory stores 12 months into the future using seasonal autoregressive integrated moving average (SARIMA) models and Box-Jenkins methodology. Forecasting these values will help us gain insight into the fluctuations of the clothing and clothing accessory stores' sales due to inflation and also seasonal changes.

## 2. Introduction

This data set is for the monthly sales (in millions of dollars) of clothing and clothing accessory stores and can be used to assess the economic health among other things of the US clothing retail industry. We would like to analyze and forecast the observed fluctuations in values due to seasonal changes. The models created to forecast future values are trained from a monthly non-seasonally adjusted dataset. The data was collected from 1992 to 2019 by the U.S. Census Bureau and hosted by the Federal Reserve Economic Database (FRED) at [RSCCASN](#).

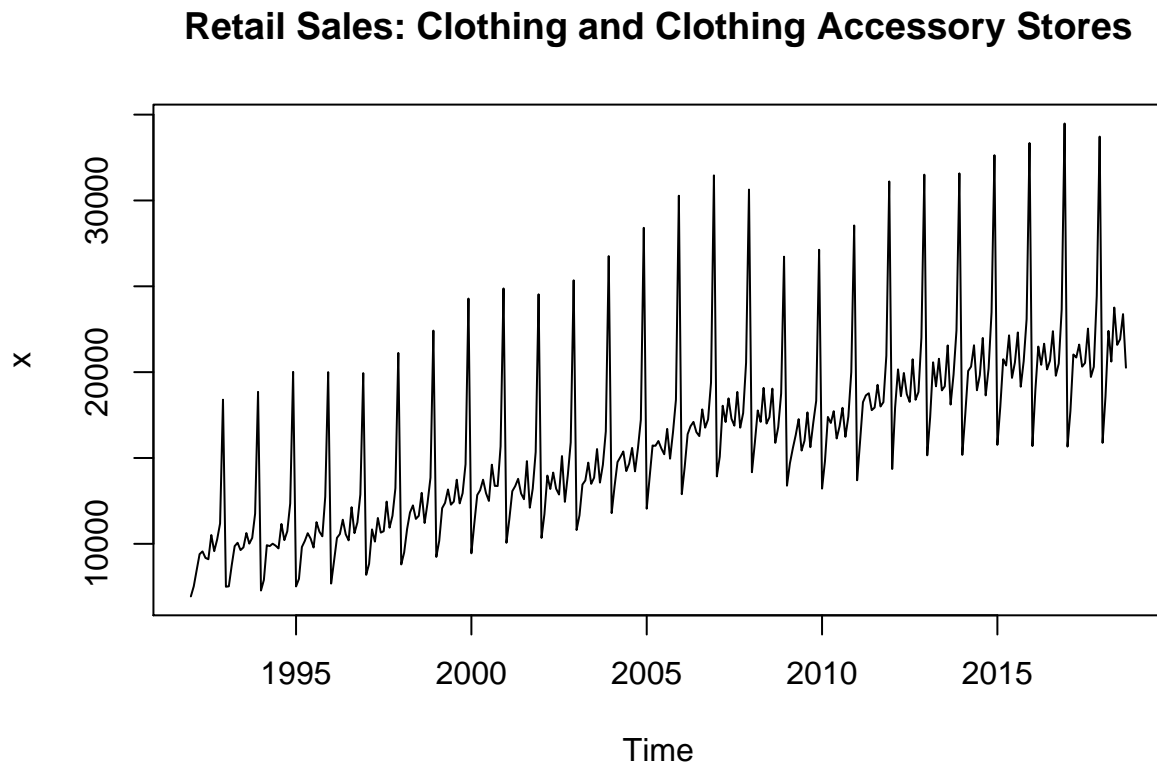
We begin our analysis and forecasting by loading the dataset and removing the last 12 values to compare our forecasts with. Then, we begin an exploratory analysis and difference the data to remove trend and seasonality. We hope to use our now stationary time series to plot the autocorrelations and partial autocorrelation functions to help identify a seasonal autoregressive integrated moving average (SARIMA) model to forecast after differencing. We then identify two potential models and conduct diagnostic checking on both to identify the best model that follows all of our assumptions. We then utilize our fitted model to forecast 12 future values. These predicted values are compared with the values we removed in the beginning in order to prove that the fitted model was successful in forecasting. Finally we perform a spectral analysis of the data and our model to test for periodic seasonal significances.

All statistical analysis was performed on 'RStudio' utilizing various software libraries from 'R'.

### 3. Analysis

#### 3.1 Exploratory Data Analysis

We begin by loading the time series data of “Retail Sales: Clothing and Clothing Accessory Stores” into RStudio and plotting it.

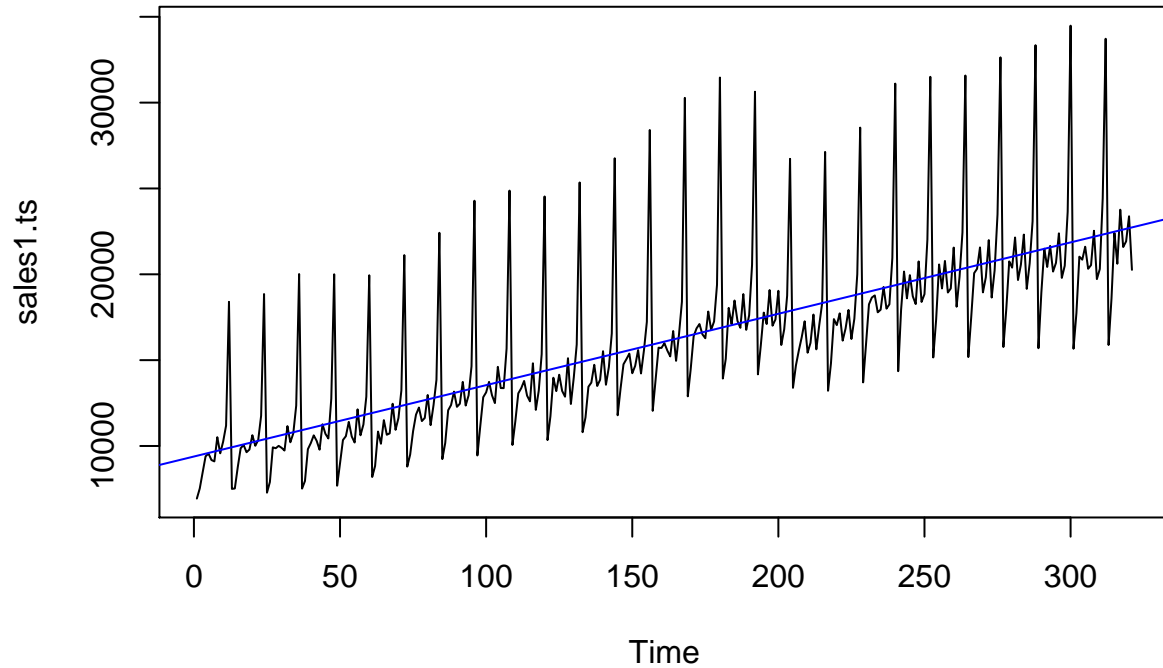


Firstly, from this plot we can see there is a positive trend moving upwards in sales. There is a small dip around 2008- 2010 which can be explained from the 2008 recession. There also seems to be a strong seasonal component and the variance changes slightly over time. We must transform and difference the data in order to make it stationary.

Also, we removed the last 12 values of the data to compare against the forecasted values later in our analysis.

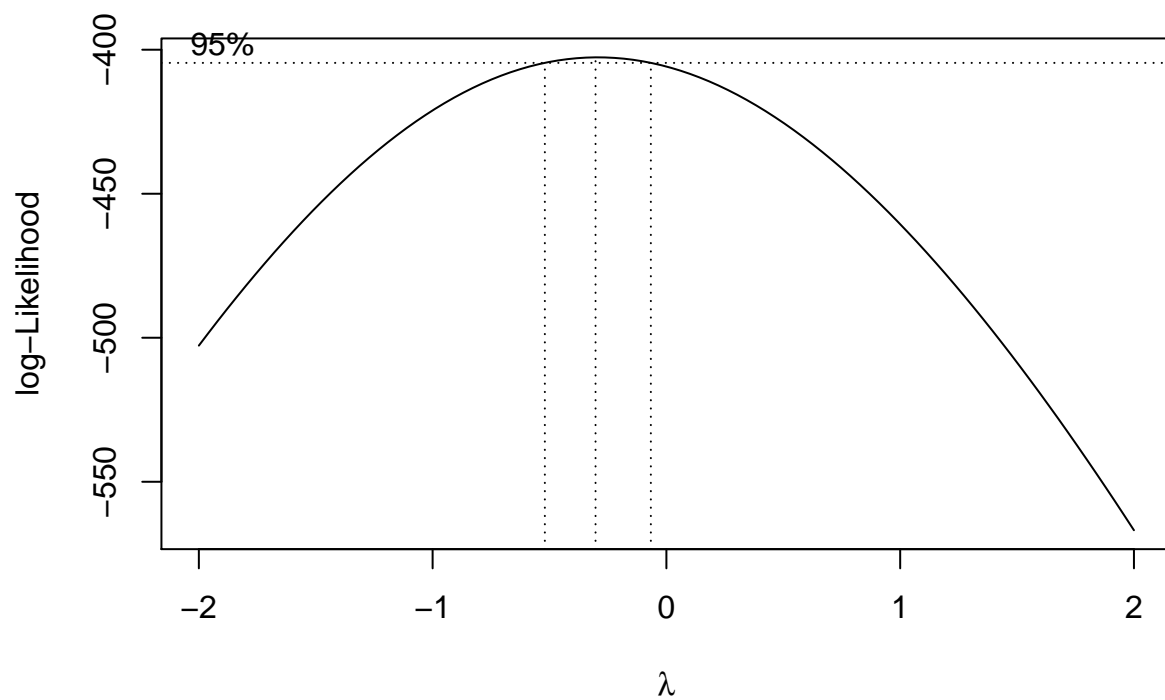
### 3.2 Data Transformations

#### Retail Sales: Clothing and Clothing Accessory Stores

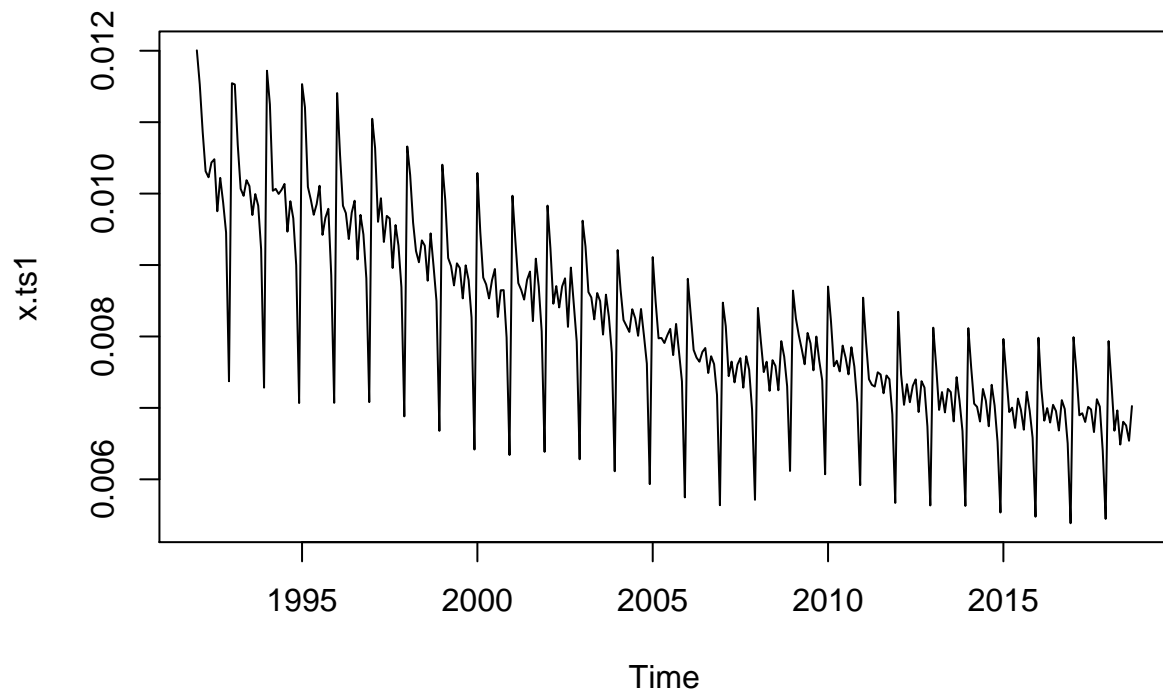


Once we fitted our blue line, a positive trend is clear. To remove trend, we have to difference at lag = 1 and we difference at lag = 12 to remove seasonality.

In order to stabilize variance and keep it constant, we will need to perform a Box-Cox transformation.

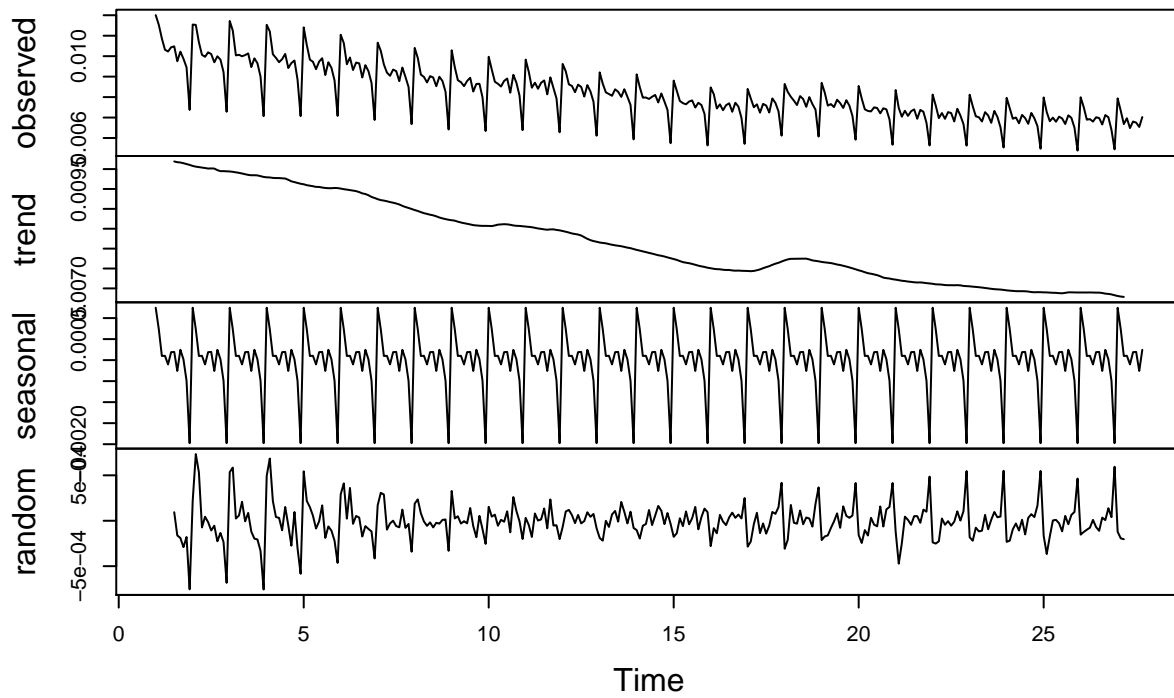


To find our optimal  $\lambda$  value we use the `boxcox()` function. From the plot we find that our optimal  $\lambda$  value is -0.3, but we approximate to -0.5 for the sake of easy analysis. Based on this  $\lambda$  value, we perform a  $1/\sqrt{Y}$  transform.

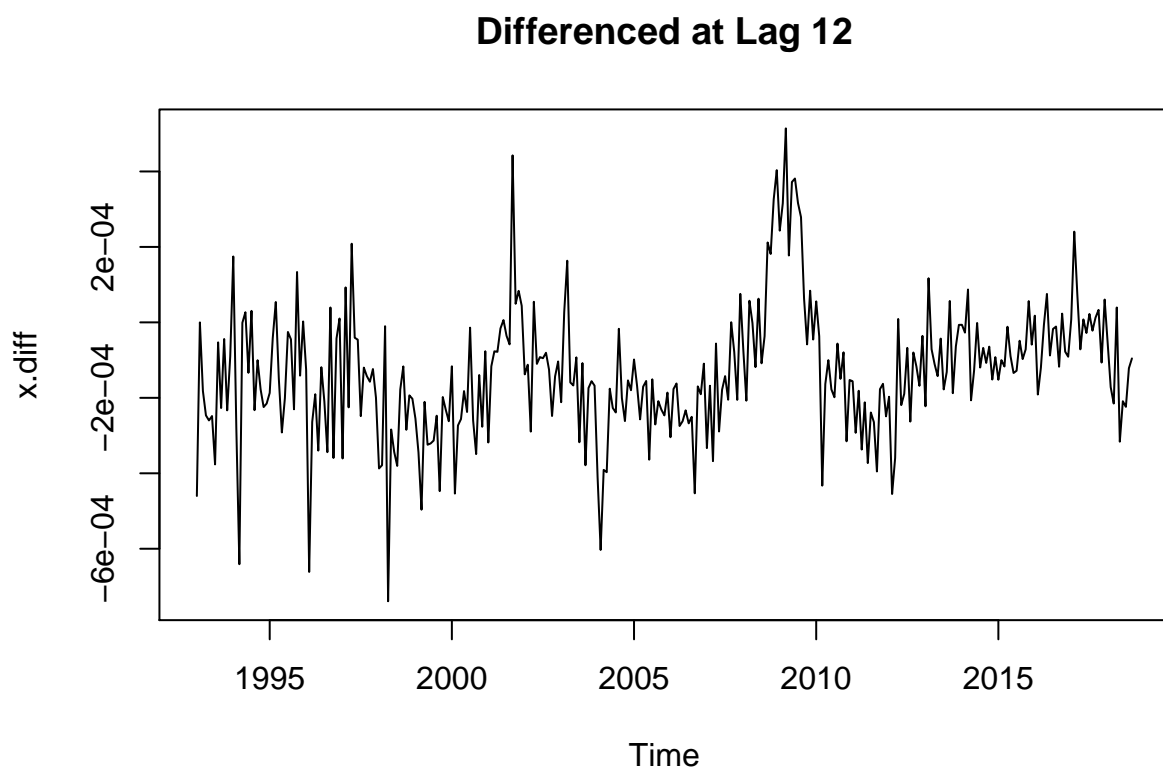


We find that the Variance of the pre-transformed original data was 27732604 and the new variance of our Box-Cox transformed data is 1.77098e-06, which is extremely close to zero showing that our Box-Cox transformation was succesful in stabilizing our variance.

## Decomposition of additive time series



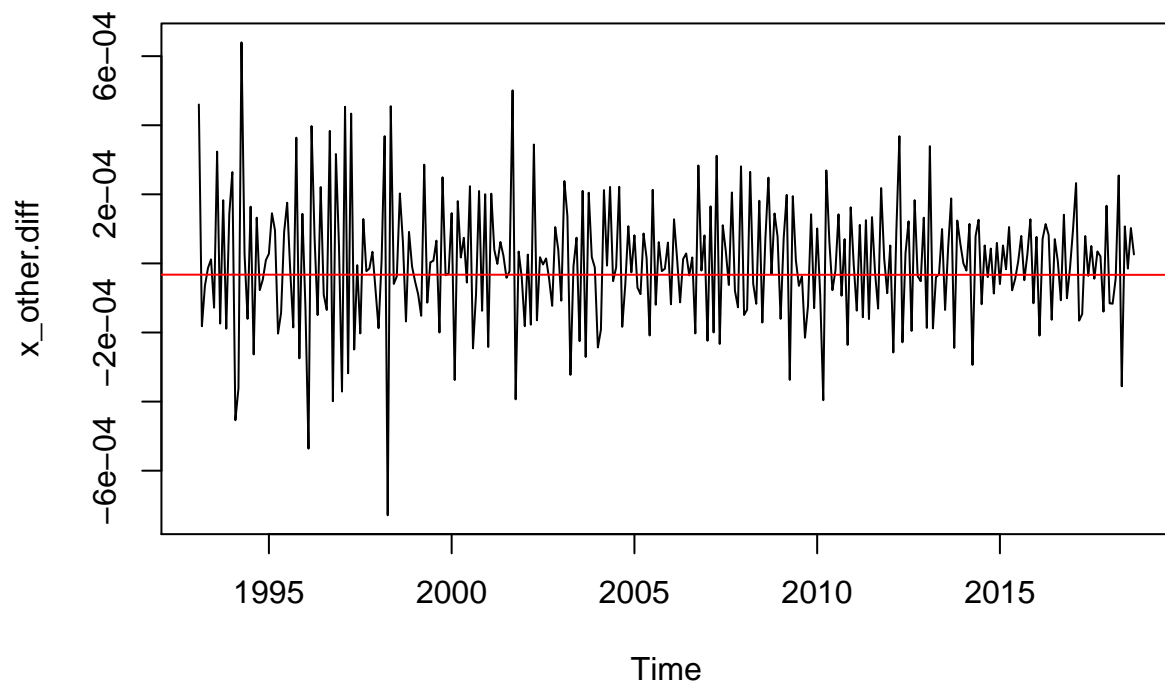
The Decomposition of our transformed data shows our trend and seasonal components. Thus, to remove them we will now difference at lag = 1 and lag = 12.



By differencing at lag 12, seasonality is clearly removed. We however still must remove the trend so we difference again at lag 1.

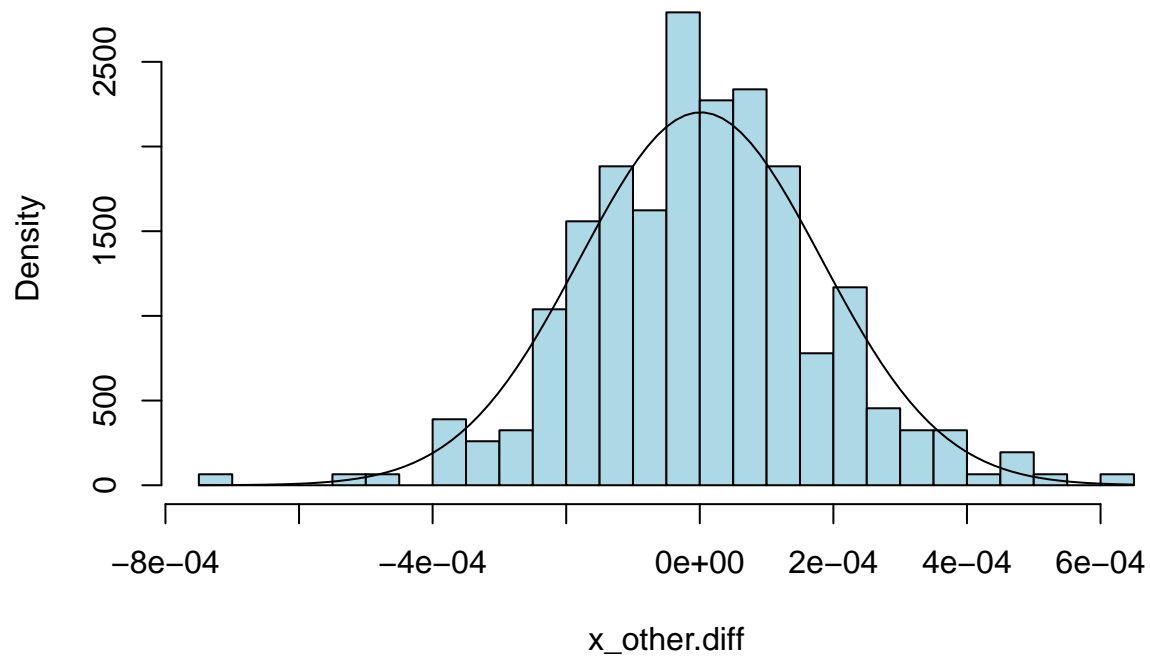


### Differenced at Lag 12 and Lag 1



When differenced again at lag 1, our variance is decreased to  $3.282317e-08$ , we can see that our mean shown by the red line is close to 0 and our data appears to be stationary.

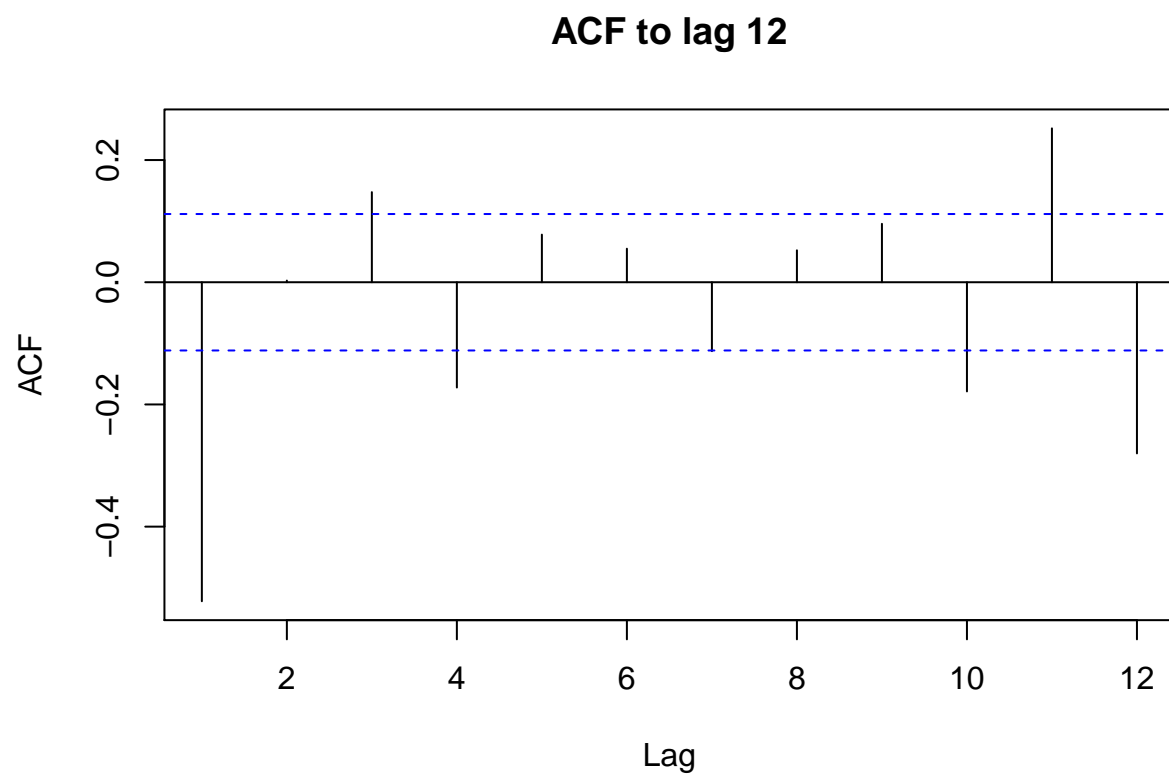
### Histogram: Differenced at lag 12 and lag 1



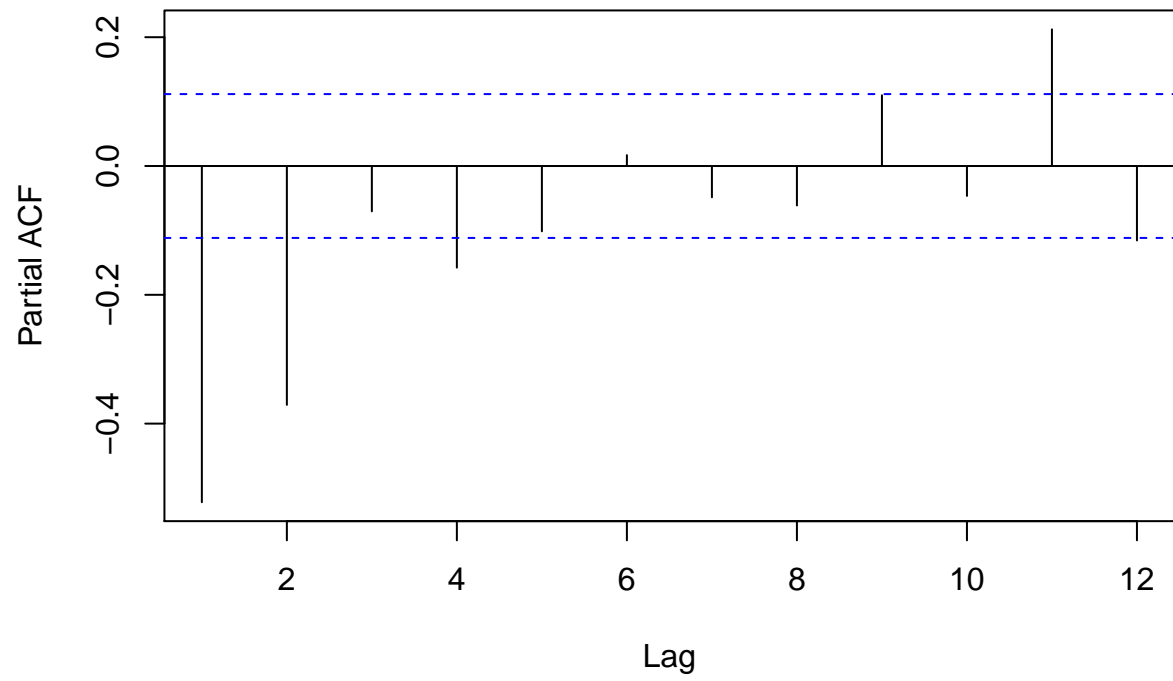
Looking at the histogram of our twice difference data, we notice it is symmetric and almost Gaussian with some apparent outliers.

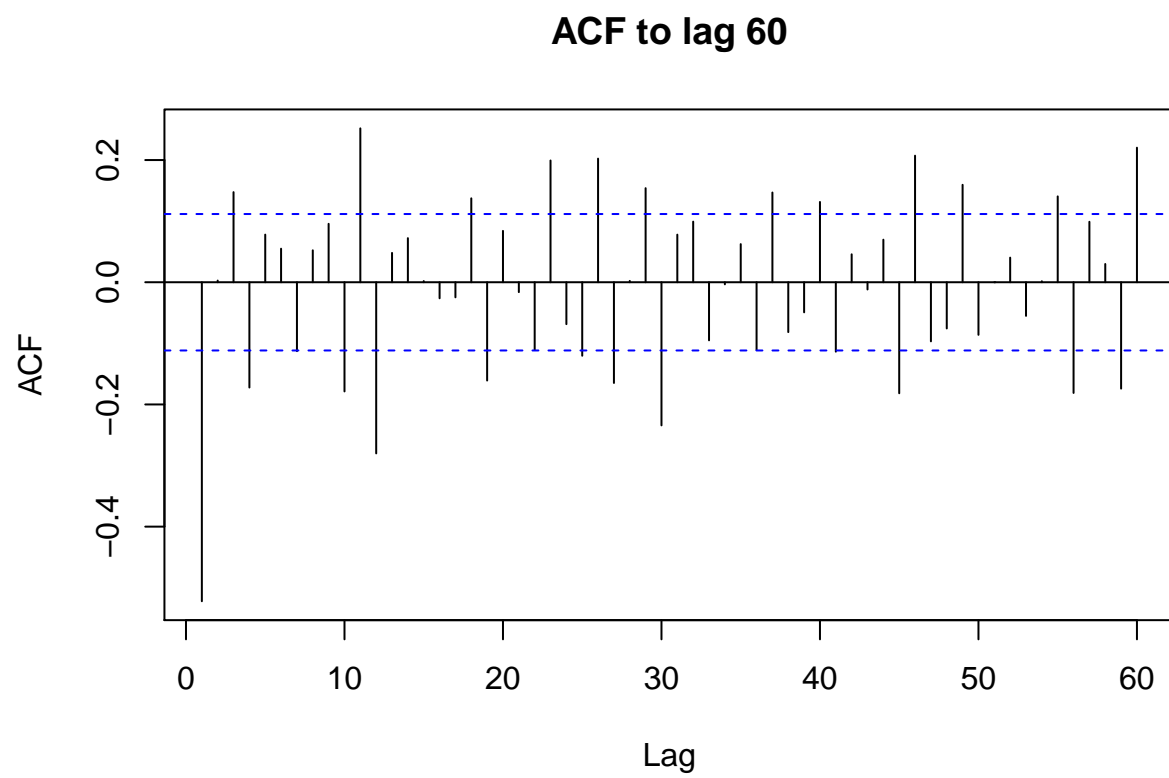
### 3.3 ACF and PACF Analysis

We can now use the ACF and PACF to fit possible SARIMA models because our data is stationary.

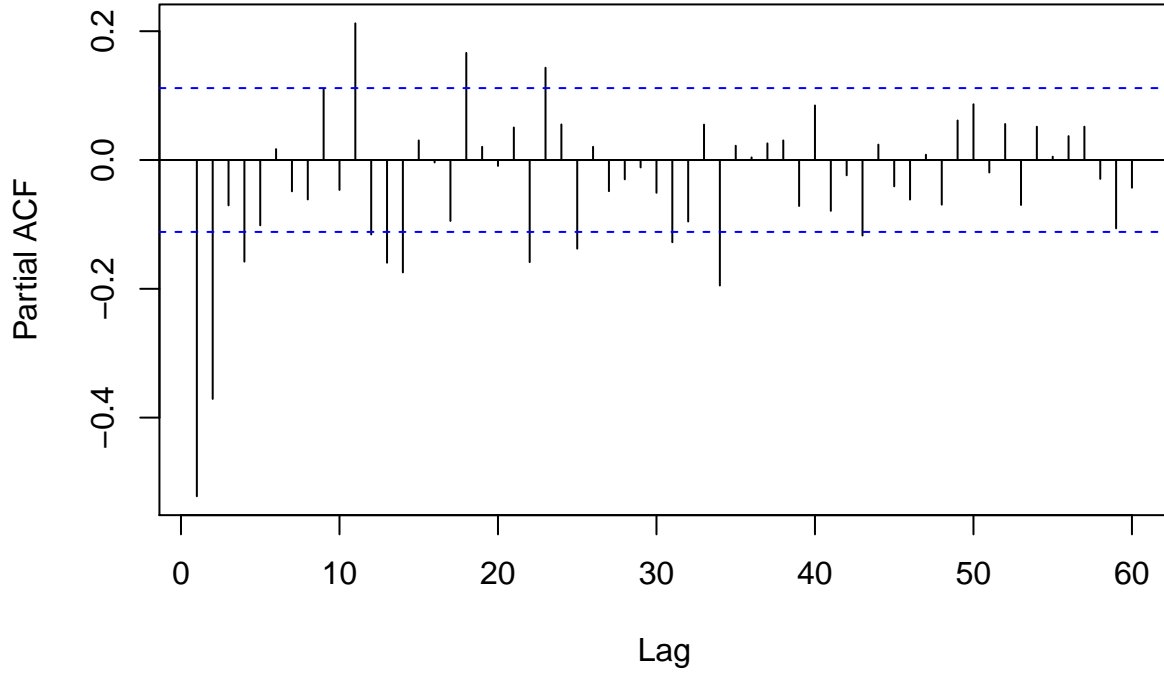


**PACF to lag 12**





## PACF to lag 60



We can identify the seasonal components by examining lags  $l = 12n, n \in N$ . The ACF cuts off after lag 12, so we can assume SMA(1). The PACF cuts off after lag 12, so we can assume SAR(1).

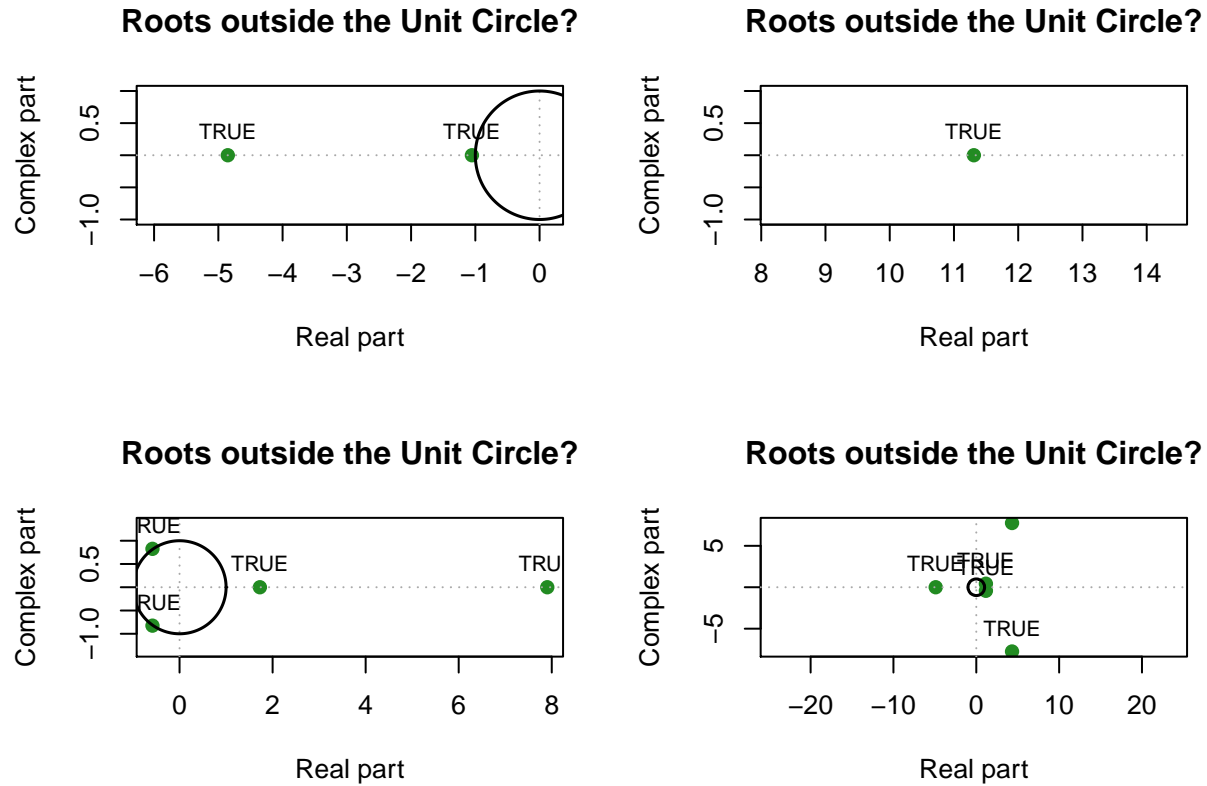
To identify the AR and MA orders, we will examine the lags 1 to 11. The ACF shows significant at lags 1, 3, 4 and 10 so we can assume MA(1), MA(3), MA(4), and MA(10). The PACF shows significance at lags 1, 2, and 4, so we assume AR(1), AR(2), and AR(4). We fit multiple models using our parameters from the ACF and PACF plots. We also consider models found by calculating AICc for a given model and choose one with the smallest AICc value. My first model is found using ACF and PACF, and my second model is obtained by using AICc.

We consider these two models and estimate the coefficients using ML estimation. For coefficients that have 0 in their confidence interval, we fix their value as 0 in order to maintain a lower AICc

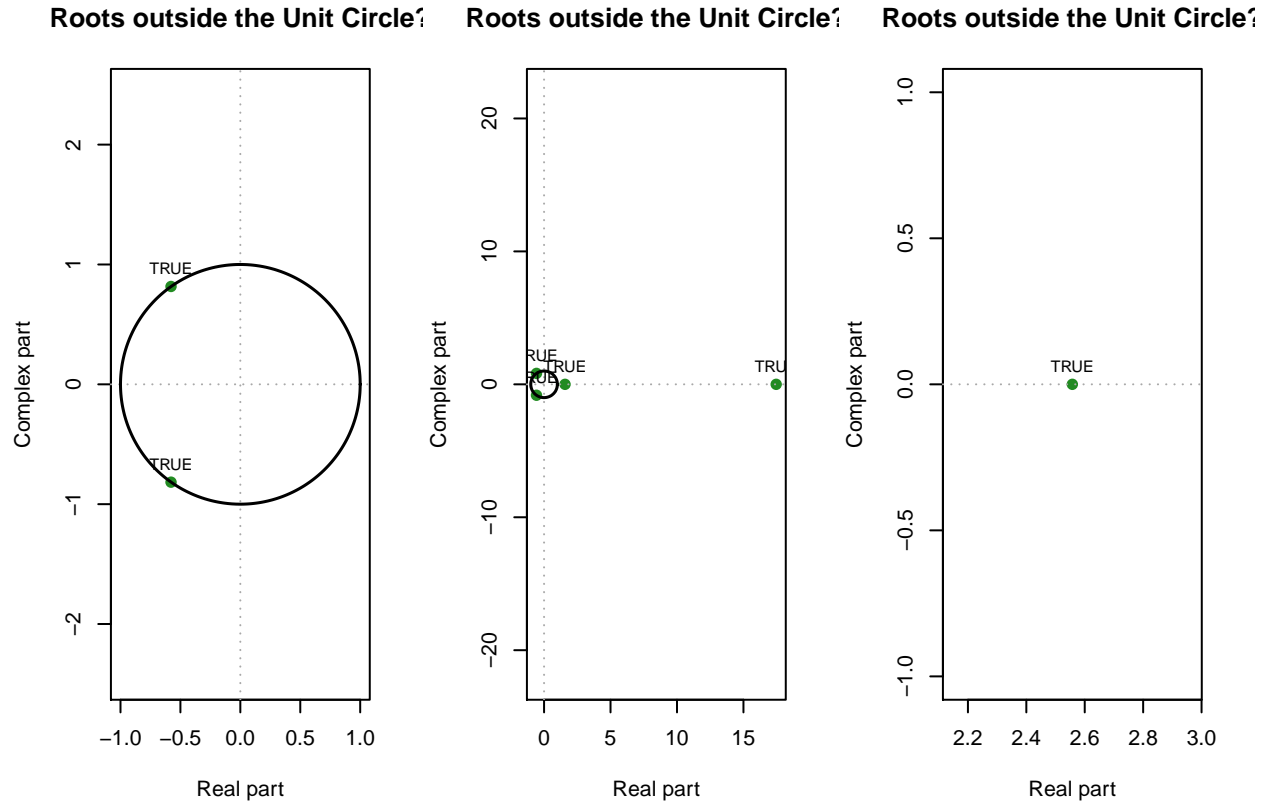
- 1. SARIMA (2, 1, 4)  $\times$  (1, 1, 5)<sub>12</sub>
  - AIC = -4669.807
  - $(1 + 1.1557_{0.0025}B + 0.9991_{0.0017}B^2)(1 - 0.8842_{0.0998}B^{12})(1 - B)(1 - B^{12})(1/\sqrt{X_t}) = (1 + 0.4356_{0.0582}B + 0.24821_{0.0547}B^2 - 0.6079_{0.0534}B^3 + 0.0718_{0.0587}B^4)(1 - 1.4259B^{12} + 0.5027_{0.1222}B^{24} + 0.0793_{0.1192}B^{36} - 0.01028_{0.1025}B^{48} + 0.0017_{0.0749}B^{60})Z_t$
- 2. SARIMA (2, 1, 4)  $\times$  (0, 1, 1)<sub>12</sub>
  - AIC = -4651.958
  - $(1 + 1.1555_{0.0033}B + 0.9988_{0.0719}B^2)(1 - B)(1 - B^{12})(1/\sqrt{X_t}) = (1 + 0.4474_{0.0616}B + 0.2493_{0.0543}B^2 - 0.6498_{0.0523}B^3 + 0.0363_{0.0594}B^4)(1 - 0.3910_{0.0482}B^{12})Z_t$

### 3.4 Model Diagnostics

We begin the diagnostic checking for both of our models by checking if they are causal and invertible. We check this by seeing if the roots of the polynomials are outside of the unit circle.



By plotting the roots of the AR, MA , SMA, and SAR separately we find that the roots for our Model 1 are all outside of the unit circle and conclude that our model is both causal and invertible.



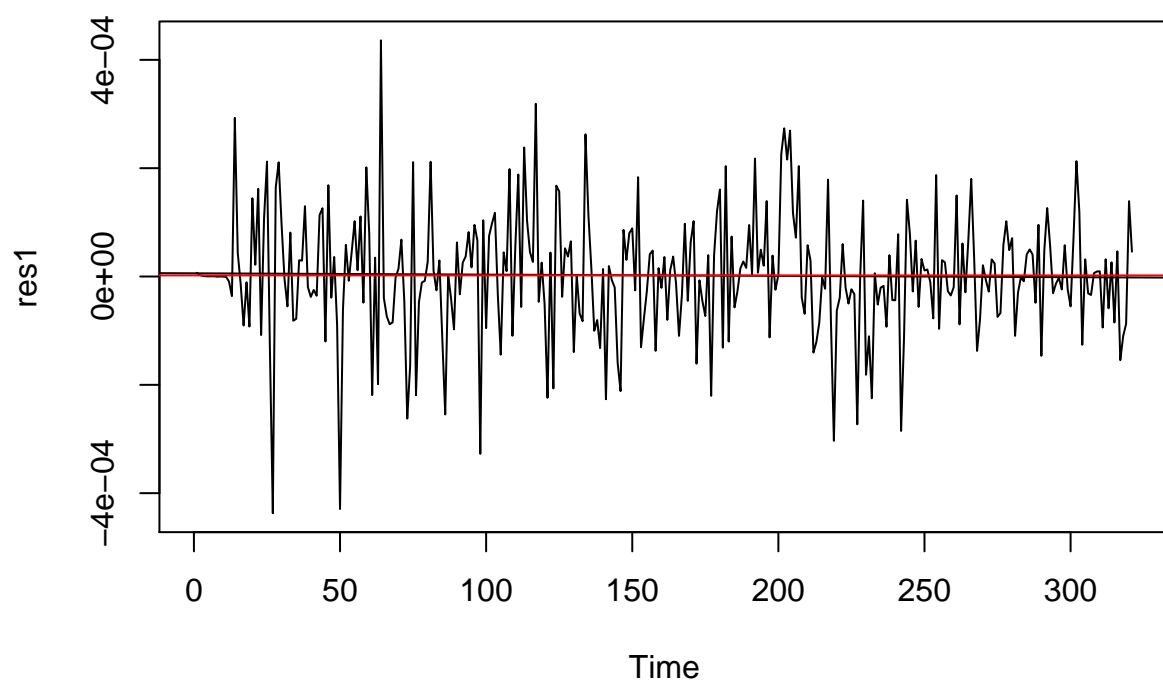
Similarly, by plotting the roots of the AR, MA , and SMA separately we find that the roots for our Model 2 are all outside of the unit circle and conclude that our model is both causal and invertible.

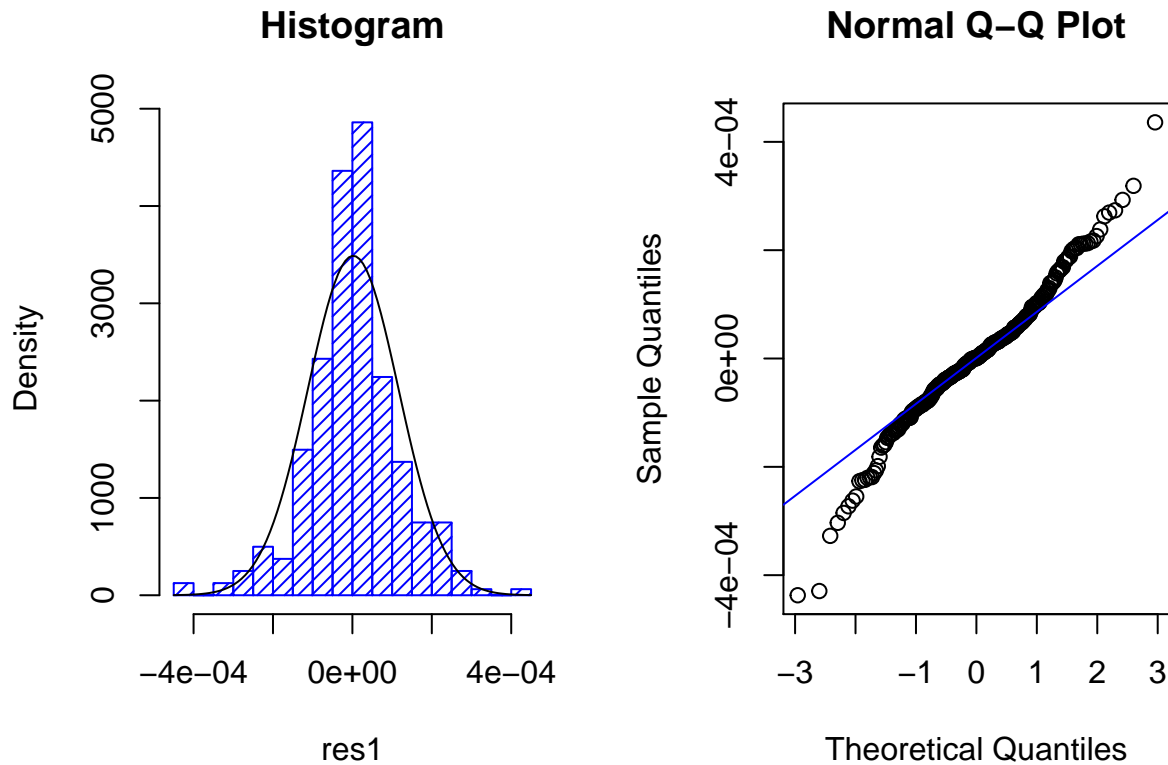
Now that we know both models are causal and invertible we can continue to analyse each models residuals separately. Beginning with our Model 1

$$\begin{aligned}
 & (1 + 1.1557_{0.0025}B + 0.9991_{0.0017}B^2)(1 - 0.8842_{0.0998}B^{12})(1 - B)(1 - B^{12})(1/\sqrt{X_t}) = \\
 & (1 + 0.4356_{0.0582}B + 0.2482_{0.0547}B^2 - 0.6079_{0.0534}B^3 + 0.0718_{0.0587}B^4) \\
 & *(1 - 1.4259B^{12} + 0.5027_{0.1222}B^{24} + 0.0793_{0.1192}B^{36} - 0.01028_{0.1025}B^{48} + 0.0017_{0.0749}B^{60})Z_t
 \end{aligned}$$



### Fitted Residuals





```
##
##  Shapiro-Wilk normality test
##
## data:  res1
## W = 0.97604, p-value = 3.44e-05
```

The plot of the residuals for Model 1 appears to resemble white noise. The histogram and Q-Q plot, however, appear to show that the residuals are skewed from normal gaussian. Performing the Shapiro-Wilk normality test confirms our observation of a non-gaussian distribution as it gives us a p-value =  $3.44e-05$ , which fails the test at significance level  $\alpha = 0.05$

To continue our diagnostics we will we also perform the Box-Pierce, Ljung-Box, and McLeod-Li tests on the residuals.

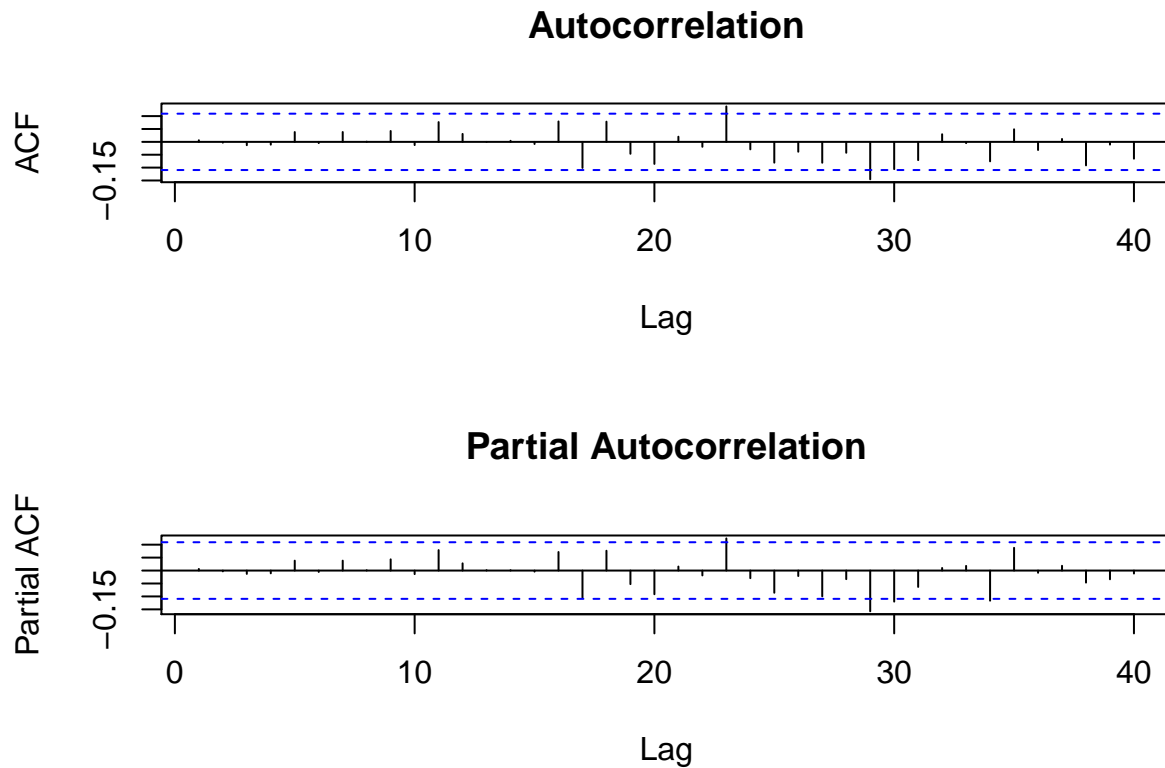
```
##
##  Box-Pierce test
##
## data:  res1
## X-squared = 11.192, df = 12, p-value = 0.5125

##
##  Box-Ljung test
##
## data:  res1
## X-squared = 11.79, df = 12, p-value = 0.4627
```

```
##
## Box-Ljung test
##
## data:  res1^2
## X-squared = 19.232, df = 18, p-value = 0.3777
```

The model passes all tests at the  $\alpha = 0.05$  significance level.

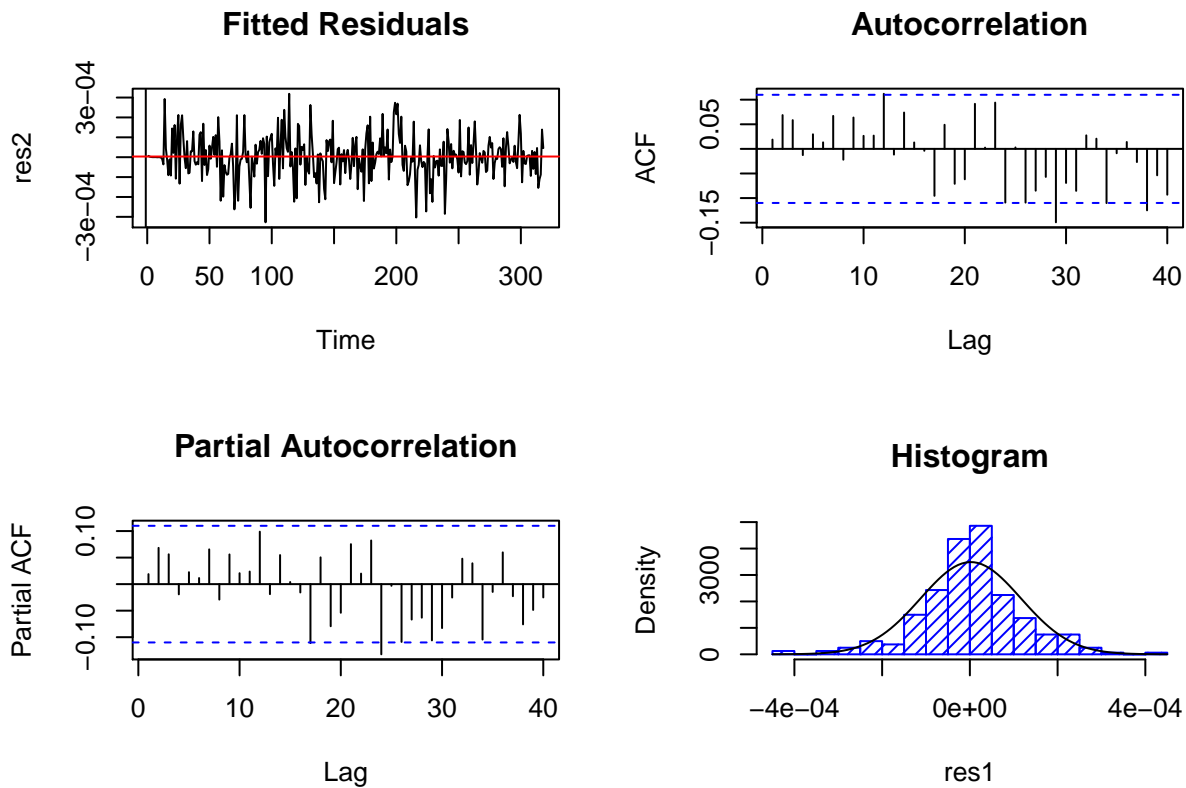
We also plot the ACF and PACF of the residuals to check if they resemble white noise.



Besides the ACFs and PACFs extending slightly over the confidence intervals at lags 23 and 29 respectively, they resemble white noise.

However because the shapiro test fails, we will test to see if removing the outlier residuals will help our model diagnostics and improve the normality assumption.

```
## Warning in abline(fit.A): only using the first two of 12 regression coefficients
```

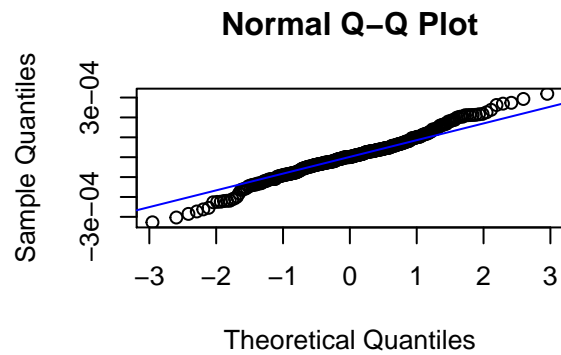


```
##
##  Shapiro-Wilk normality test
##
## data:  res2
## W = 0.98706, p-value = 0.006052

##
##  Box-Pierce test
##
## data:  res2
## X-squared = 15.898, df = 12, p-value = 0.196

##
##  Box-Ljung test
##
## data:  res2
## X-squared = 16.567, df = 12, p-value = 0.1666

##
##  Box-Ljung test
##
## data:  res2^2
## X-squared = 36.204, df = 18, p-value = 0.006642
```

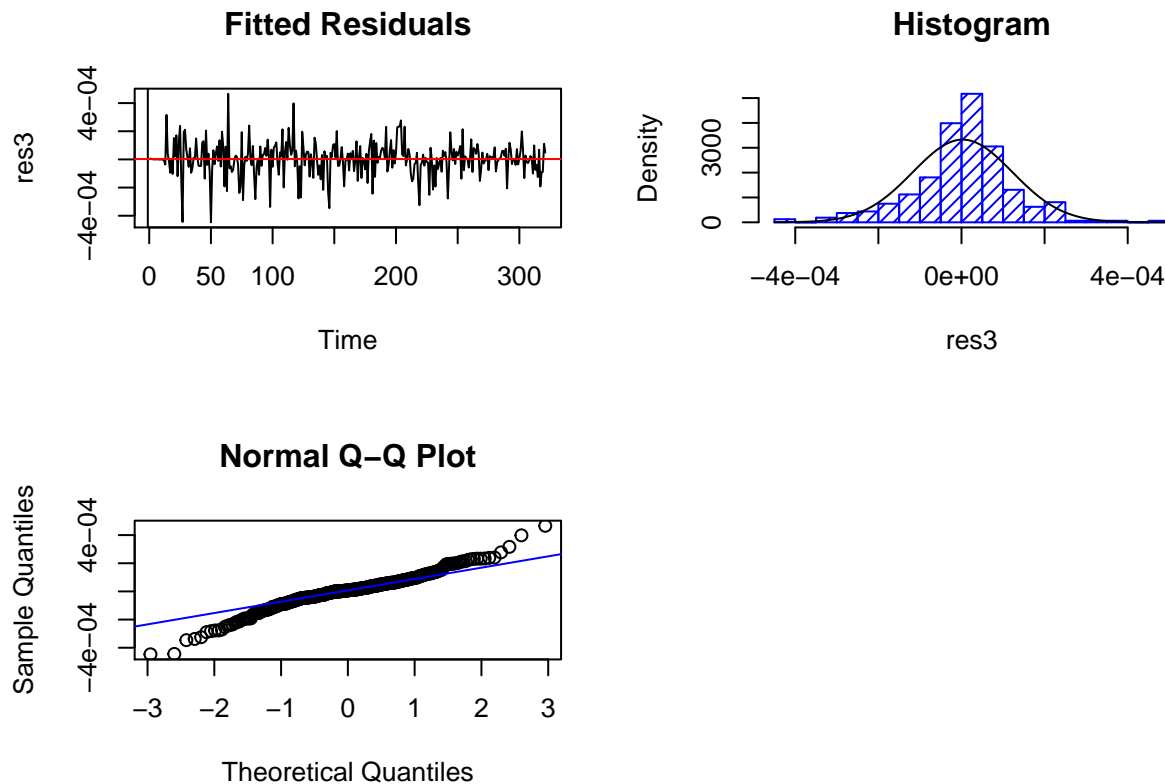


Removing the outliers results in the Box-Ljung test failing, thus we were wrong in taking away the outliers.

Moving forward, we will assess our Model 2:  $(1 + 1.1555_{0.0033}B + 0.9988_{0.0719}B^2)(1 - B)(1 - B^{12})(1/\sqrt{X_t}) = (1 + 0.4474_{0.0616}B + 0.2493_{0.0543}B^2 - 0.6498_{0.0523}B^3 + 0.0363_{0.0594}B^4)(1 - 0.3910_{0.0482}B^{12})Z_t$

```
## Warning in abline(fit.B): only using the first two of 7 regression coefficients
```

```
##
## Shapiro-Wilk normality test
##
## data:  res3
## W = 0.96332, p-value = 3.122e-07
```



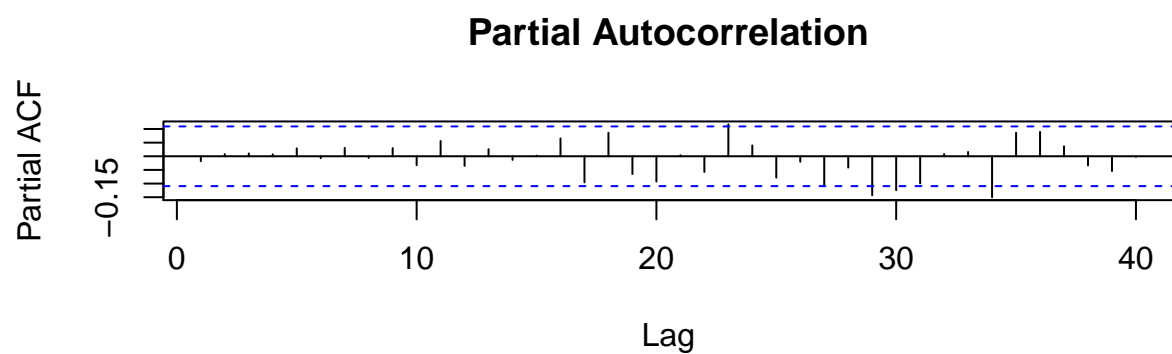
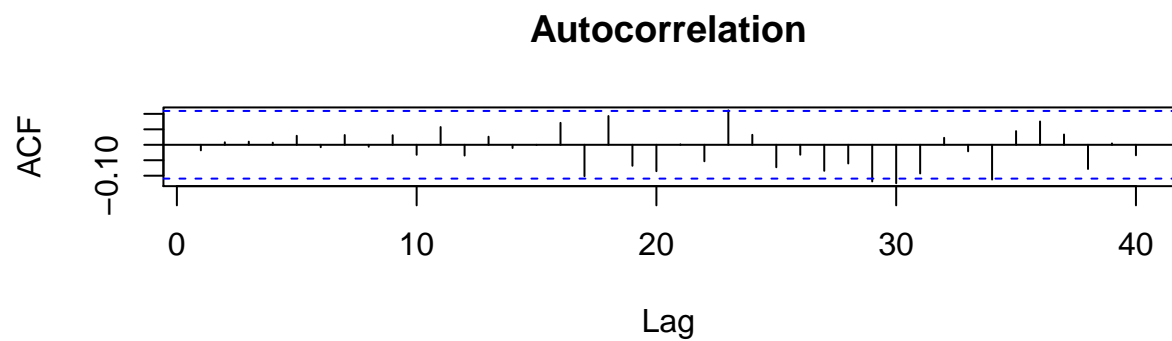
Once again, the plot of the residuals for Model 2 appears to resemble white noise. The histogram and Q-Q plot, however, appear to show that the residuals are skewed from normal gaussian. Performing the Shapiro-Wilk normality test confirms our observation of a non-gaussian distribution as it gives us a p-value =  $3.122e-07$ , which fails the test at significance level  $\alpha = 0.05$ .

```
##
## Box-Pierce test
##
## data: res3
## X-squared = 10.839, df = 12, p-value = 0.5428

##
## Box-Ljung test
##
## data: res3
## X-squared = 11.44, df = 12, p-value = 0.4916

##
## Box-Ljung test
##
## data: res3^2
## X-squared = 16.377, df = 18, p-value = 0.5663
```

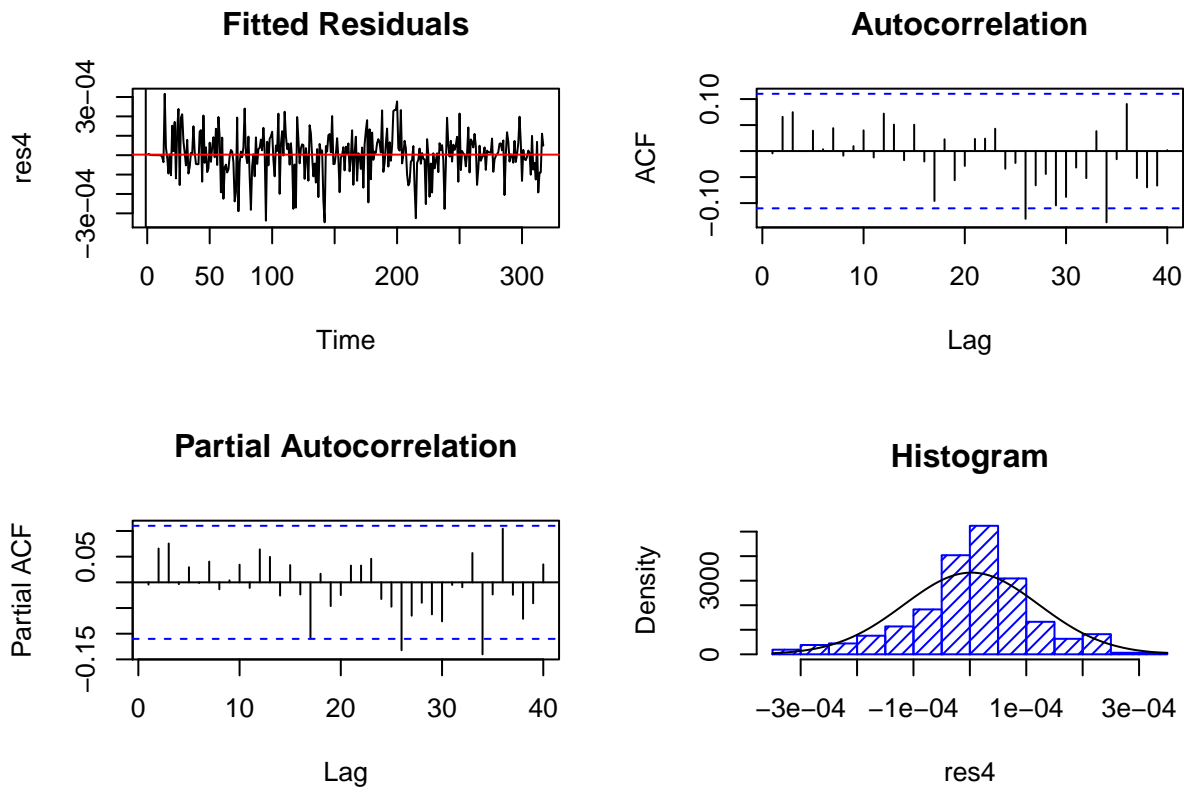
The model passes all tests at the  $\alpha = 0.05$  significance level. We also note that Model 2 performs better than Model 1 in passing the Box-Pierce test by having a higher p-value. This means that Model 2 fits a white noise process better.



Checking the acf and pacf, besides lag 29 and 30, it resembles a white noise process.

To try to fix for our normality assumption, we are going to take out the outliers of model 2 again.

```
## Warning in abline(fit.B): only using the first two of 7 regression coefficients
```



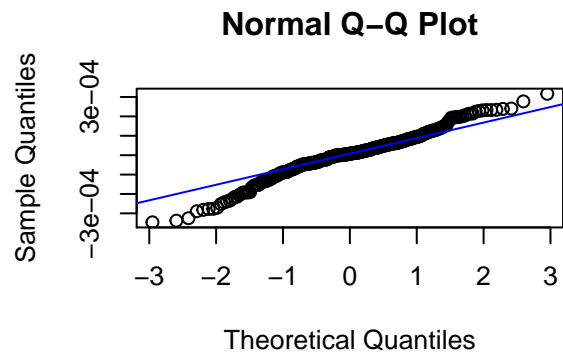
```
##
##  Shapiro-Wilk normality test
##
## data:  res4
## W = 0.97527, p-value = 2.829e-05

##
##  Box-Pierce test
##
## data:  res4
## X-squared = 11.443, df = 12, p-value = 0.4914

##
##  Box-Ljung test
##
## data:  res4
## X-squared = 11.911, df = 12, p-value = 0.4529

##
##  Box-Ljung test
##
## data:  res4^2
## X-squared = 40.128, df = 18, p-value = 0.002005
```



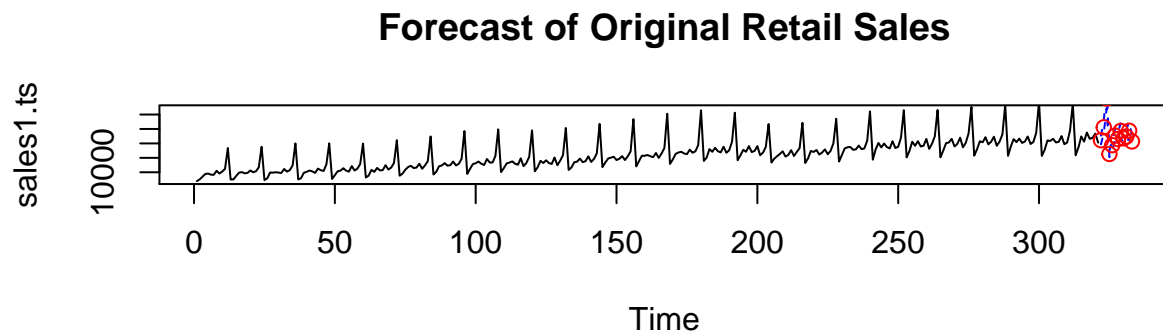
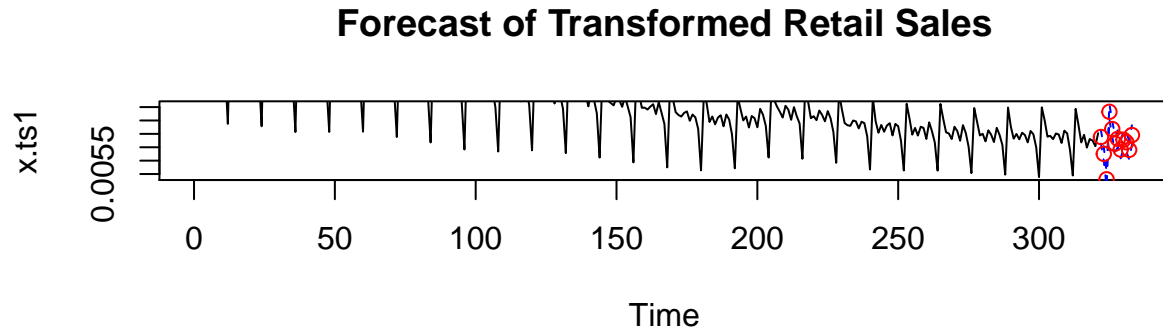


Once again, taking out the outliers does not help and has a negative affect shown by diagnostic testing by failing the Box-Ljung test.

We conclude that both models follow a non-gaussian white noise process and will choose to use Model 2 in our forecasting because it performed better under the Box-Pierce and Box-Ljung tests, and utilizing principle of parsimony we choose Model 2 because it has fewer parameters.

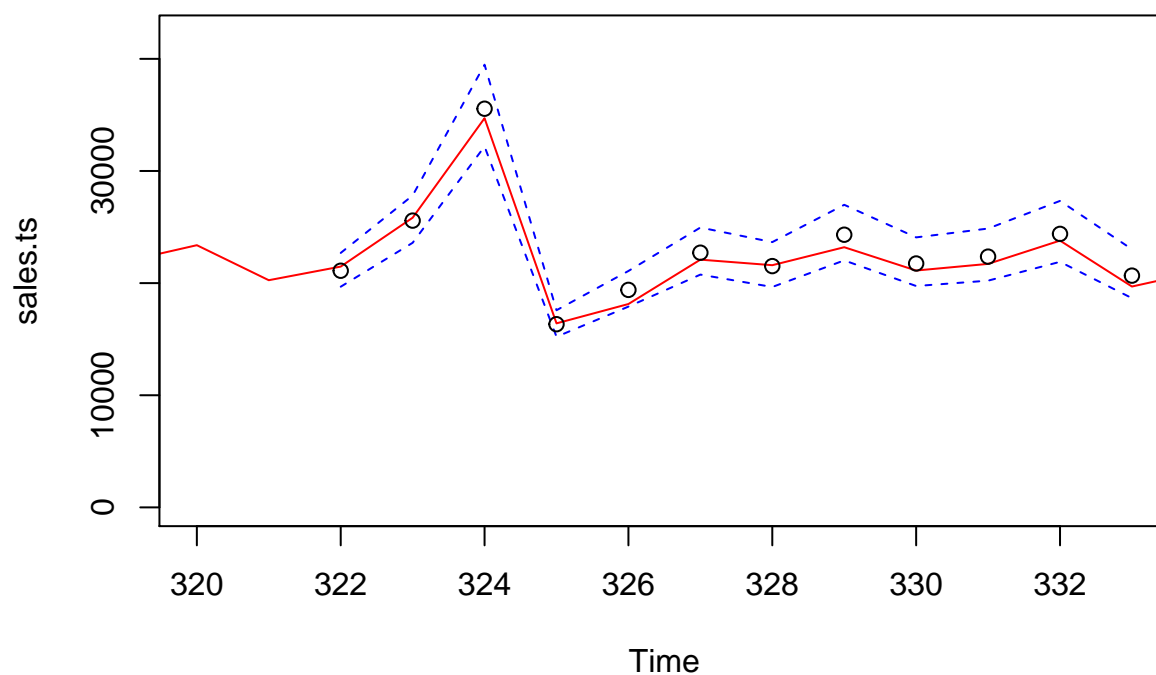
### 3.5 Forecasting

We selected our model and now we can begin forecasting.



The forecasted values are plotted in red and their confidence intervals are in blue. This becomes clearer once we zoom in.

### Zoomed Forecast of Original Retail Sales

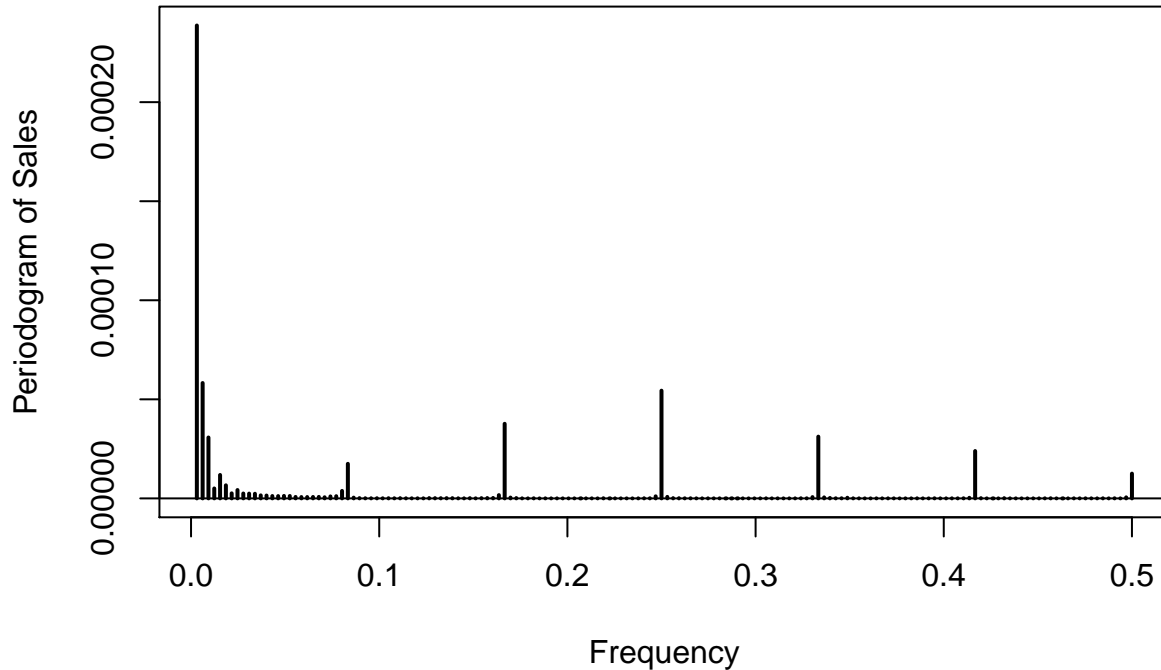


The original Data is in red and the predicted values are the black dots and the confidence interval is shown by the blue dotted lines. From this zoomed in plot, we can clearly see that the predicted values are very close to the actual observed values and are all within the confidence interval.

### 3.5 Spectral Analysis

Considering the Model:  $X_t = \mu + a\cos(\omega t) + b\sin(\omega t + Z_t, Z_t \sim WN(0, \sigma_Z^2))$

We can analyse the periodicity of our data and models using a Periodogram. First we will analyze the periodogram of the transformed data:



From this Periodogram, we clearly see that there are peaks and frequencies around approximately .8, .16, .25, .33, .41, and .5. Thus we can say that the transformed Sales data shows significant periodicity and seasonality.

Moving forward, we will analyze our Model 2 residuals using spectral analysis to test for periodicity.

Conducting a Fisher test and a Kolmogorov-Smirnov test to test for white noise and gaussian white noise within residuals:

```
## Loading required package: longitudinal

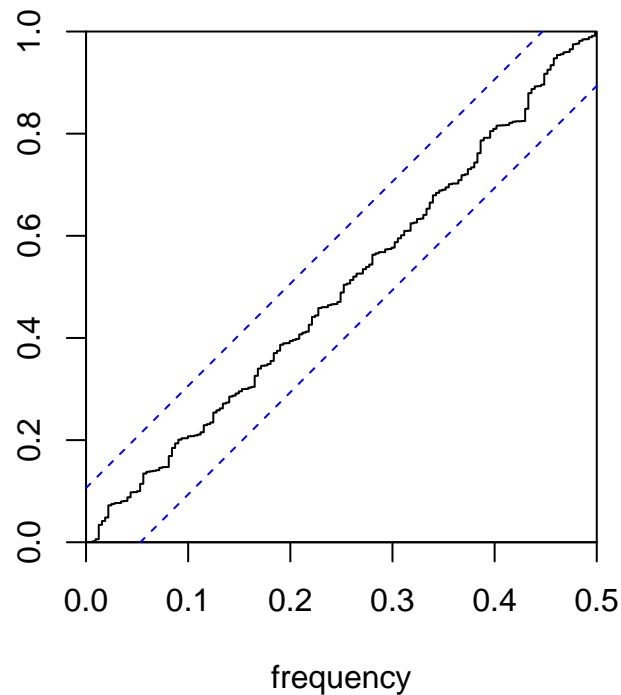
## Loading required package: corpcor

## Loading required package: fdrtool

##
## Attaching package: 'GeneCycle'

## The following object is masked from 'package:TSA':
##
##     periodogram
```

```
## The following object is masked from 'package:forecast':  
##  
##      is.constant  
  
## [1] 0.5441366
```



When conducting a Fesher test on the Model 2 residuals, we pass with a p-value of 0.5441366 which is far greater than the  $\alpha = 0.05$  level of significance. We can thus conclude that there is insufficient effidence to reject the hypothesis of following a white noise process.

It its interesting to note that the residuals of Model 2 pass the Kolmogorov-Smirnov test for Gaussian white noise, however we default to our earlier analysis using the Shapiro-Wilk test because the Kolmogorov-Smirnov test only considers the largest discrepancy between observed and hypothesized distributions. Thus, we can still conclude that the Model 2 residuals follow a non-gaussian white noise process.

## 4. Conclusion

Our objective of this paper was to accurately forecast 12 months of the Sales of Retail Sales of US Clothing and Clothing Accessory Stores. From our analysis, we chose to use a SARIMA Model:

$$(1 + 1.1555_{0.0033}B + 0.9988_{0.0719}B^2)(1 - B)(1 - B^{12})(1/\sqrt{X_t}) = \\ (1 + 0.4474_{0.0616}B + 0.2493_{0.0543}B^2 - 0.6498_{0.0523}B^3 + 0.0363_{0.0594}B^4)(1 - 0.3910_{0.0482}B^{12})Z_t$$

which has one of the lowest AICc's and passes all the diagnostic tests except for the residuals following Gaussian-White noise so we found that it follows a non-gaussian white noise process. Our model was shown to be accurate in forecasting the Retail Sales (in millions) for US Clothing and Clothing Accessory Stores 12 months into the future. Our spectral analysis reinforced our earlier claims of seasonality and residuals following white noise.

## 5. References

- Brockwell, P.J., Davis, R.A. 2016. Introduction to Time Series and Forecasting.
- [Advance Retail Sales: Clothing and Clothing Accessory Stores](#)
- Dr. Raya Feldman, UCSB PSTAT 174/274, Time Series Lecture Notes
- Tamjid Islam, UCSB 4th Year Data Science Student

## Appendix

Below is the R code used to create this report.

```
#loading libraries
library(forecast)
library(MASS)
library(astsa)
library(tseries)
library(ggplot2)
library(TSA)

#loading data
sales <- read.csv("mycsvfile.csv")
sales.ts <- ts(sales$value)
sales1.ts <- sales.ts[-c(322:334)]
data <- sales.ts[322:334]
x <- ts(sales1.ts, start = c(1992,1,1), frequency = 12)
ts.plot(x, main = "Retail Sales: Clothing and Clothing Accessory Stores")

#finding acf and pacf of pre-transformed data
acf(x, lag.max= 50)
pacf(x, lag.max = 50)
var(x)

#Box Cox Transformation
t = 1:length(x)
fit = lm(x ~ t)
bcTransform = boxcox(x~t, plotit = TRUE)
lambda = bcTransform$x[which(bcTransform$y == max(bcTransform$y))]
lambda
x.bc = 1/(sqrt(x))
var(x.bc)
x.ts1 <- ts(x.bc, start=c(1992,1,1), frequency = 12)
ts.plot(x.ts1)
var(x.bc)

#decompostion
y <- ts(as.ts(x.bc), frequency = 12)
decomp <- decompose(y)
plot(decomp)

#acf and pacf of differenced data at frequency 1
sales1.ts <- sales.ts[-c(322:334)]
x.bc1 = 1/(sqrt(sales1.ts))
x.ts2 <- ts(x.bc1)
x.diff1 <-diff(x.ts2,12)
x.diffdiff<-diff(x.diff1, 1)
acf(x.diffdiff, lag.max = 12, main = "ACF to lag 12")
pacf(x.diffdiff, lag.max = 12, main = "PACF to lag 12")
acf(x.diffdiff, lag.max = 60, main = "ACF to lag 60")
pacf(x.diffdiff, lag.max = 60, main = "PACF to lag 60")
```

```

#Finding Lowest AIC Model
l <- list(q= c(0,1,4), p= c(1,2,4), Q= c(1,5), P=c(0,1))
combs <- do.call(expand.grid, l)
combs

for(row in 1:nrow(combs)){
  flag <- TRUE
  tryCatch({
    fit.ia <- sarima( xdata= x.bc,
                     p = combs[row,2], d= 1, q= combs[row,1],
                     P= combs[row,4], D=1, Q= combs[row,3], S=12, details =F)

    print(fit.ia)
    print(combs[row,2])
    print(combs[row,1])
    print(combs[row,4])
    print(combs[row,3])
  }, error = function(e){flag <- FALSE})
  if (!flag) next
}

fit.A <- arima(x.ts2, order=c(2,1,4),seasonal=list(order=c(1,1,5), period=12),fixed=c(NA,NA,NA,NA, NA, NA, NA, NA, NA, NA, NA, NA))
fitaicc <- AICc(arima(x.ts2, order=c(2,1,4), seasonal = list(order = c(1,1,5),period = 12),fixed=c(NA,NA,NA,NA, NA, NA, NA, NA, NA, NA, NA, NA)))
fit.B <- arima(x.ts2, order=c(2,1,4),seasonal=list(order=c(0,1,1), period=12),fixed=c(NA,NA,NA,NA, NA, NA, NA, NA, NA, NA, NA, NA))
fitbaicc <- AICc(arima(x.ts2, order=c(2,1,4),seasonal=list(order=c(0,1,1), period=12),fixed=c(NA,NA,NA,NA, NA, NA, NA, NA, NA, NA, NA, NA)))

#Roots of model 1

par(mfrow=c(1,4))
uc.check(pol_ = c(1,1.1557,0.1957),print_output = F)
uc.check(pol_ = (c(1,-.08842)),print_output = F)
uc.check(pol_ =(c(1,.4356, .24821, -.6079,0.0718)),print_output = F)
uc.check(pol_ =(c(1,-1.4259, .5027, .0793, -.01028,0.0017)),print_output = F)

#roots of model 2
par(mfrow=c(1,3))
uc.check(pol_ = c(1,1.1555,0.9988),print_output = F)
uc.check(pol_ =(c(1,.4474, .2493, -.6498,0.0363)),print_output = F)
uc.check(pol_ =(c(1,-.391)),print_output = F)

#diagnostics of Model 1
res1 <- residuals(fit.A)
mn = mean(res1)
variance = var(res1)
ts.plot(res1,main = "Fitted Residuals")
t = 1:length(res1)
fit.res1 = lm(res1~t)
abline(fit.res1)
abline(h = mean(res1), col = "red")
hist(res1,density = 20, breaks = 20, main = "Histogram", col = 'blue', prob = T)
curve( dnorm(x,mn,sqrt(variance)), add=TRUE )
# q-q plot
qqnorm(res1)

```



```

qqline(res1,col ="blue")
shapiro.test(res1)
Box.test(res1, lag = 18, type = c("Box-Pierce"), fitdf = 6)
Box.test(res1, lag = 18, type = c("Ljung-Box"), fitdf = 6)
Box.test(res1^2, lag = 18, type = c("Ljung-Box"), fitdf = 0)

# acf
acf(res1,main = "Autocorrelation", lag.max = 40)
# pacf
pacf(res1,main = "Partial Autocorrelation", lag.max = 40)

#stripping outliers of Model 1 Residuals
res2 <- res1[-which.min(res1)]
res2 <- res2[-which.min(res2)]
res2 <- res2[-which.max(res2)]

ts.plot(res2,main = "Fitted Residuals")
t = 1:length(res2)
fit.res1 = lm(res2~t)
abline(fit.A)
abline(h = mean(res2), col = "red")
acf(res2,main = "Autocorrelation", lag.max = 40)
pacf(res2,main = "Partial Autocorrelation", lag.max = 40)
hist(res1,density = 20, breaks = 20, main = "Histogram", col = 'blue', prob = T)
curve( dnorm(x,mn,sqrt(variance)), add=TRUE )
qqnorm(res2)
qqline(res2,col ="blue")
shapiro.test(res2)
Box.test(res2, lag = 18, type = c("Box-Pierce"), fitdf = 6)
Box.test(res2, lag = 18, type = c("Ljung-Box"), fitdf = 6)
Box.test(res2^2, lag = 18, type = c("Ljung-Box"), fitdf = 0)
acf(res2,main = "Autocorrelation", lag.max = 40)
pacf(res2,main = "Partial Autocorrelation", lag.max = 40)

#model 2 diagnostics
res3 <- residuals(fit.B)
mn = mean(res3)
variance = var(res3)
ts.plot(res3,main = "Fitted Residuals")
t = 1:length(res3)
fit.res1 = lm(res3~t)
abline(fit.B)
abline(h = mean(res3), col = "red")
hist(res3,density = 20, breaks = 20, main = "Histogram", col = 'blue', prob = T)
curve( dnorm(x,mn,sqrt(variance)), add=TRUE )
qqnorm(res3)
qqline(res3,col ="blue")
shapiro.test(res3)
Box.test(res3, lag = 18, type = c("Box-Pierce"), fitdf = 6)
Box.test(res3, lag = 18, type = c("Ljung-Box"), fitdf = 6)
Box.test(res3^2, lag = 18, type = c("Ljung-Box"), fitdf = 0)

```

```

acf(res3,main = "Autocorrelation", lag.max = 40)
pacf(res3,main = "Partial Autocorrelation", lag.max = 40)

#Model 2 without residual outliers diagnostic
res4 <- res3[-which.min(res3)]
res4 <- res4[-which.min(res4)]
res4 <- res4[-which.max(res4)]
res4 <- res4[-which.max(res4)]

ts.plot(res4,main = "Fitted Residuals")
t = 1:length(res4)
fit.res1 = lm(res4~t)
abline(fit.B)
abline(h = mean(res4), col = "red")
# acf
acf(res4,main = "Autocorrelation", lag.max = 40)
# pacf
pacf(res4,main = "Partial Autocorrelation", lag.max = 40)
hist(res4,density = 20, breaks = 20, main = "Histogram", col = 'blue', prob = T)
curve( dnorm(x,mn,sqrt(variance)), add=TRUE )
# q-q plot
qqnorm(res4)
qqline(res4,col ="blue")
shapiro.test(res4)
Box.test(res4, lag = 18, type = c("Box-Pierce"), fitdf = 6)
Box.test(res4, lag = 18, type = c("Ljung-Box"), fitdf = 6)
Box.test(res4^2, lag = 18, type = c("Ljung-Box"), fitdf = 0)

#Forecasting
train <- x.bc[1:309]
test <- x.bc[309:321]

pred.tr <- sarima.for(train, 12, 2, 1, 4, 0, 1, 1,
                      S = 12, no.constant = FALSE, plot.all = F)
points(309:321,test,col="blue")
legend("topleft", pch = 1,col=c("red","blue"),
legend=c("Forecast value", "True Value"))

ts.plot(train, xlim=c(0,321), ylab = "Monthly Unemployment in US Males")
points(309:320,pred.tr$pred,col="red")
points(309:321,test,col="blue")
legend("topleft", pch = 1,col=c("red","blue"),
legend=c("Forecast value", "True Value"))

x.ts1 <- ts(x.bc, start=c(1,1,1), frequency = 1)
fit.B <- arima(x.ts1, order=c(2,1,4),seasonal=list(order=c(0,1,1), period=12),fixed=c(NA,NA,NA,NA, NA, 0))

#forecast(fit.B)

```

```

pred.tr <- predict(fit.B, n.ahead = 12)
U.tr= pred.tr$pred + 2*pred.tr$se
L.tr= pred.tr$pred - 2*pred.tr$se
ts.plot(x.ts1, xlim=c(1,length(x.ts1)+12), ylim = c(min(x.ts1),max(U.tr)),main = "Forecast of Transformed Sales",
        lines(U.tr, col="blue", lty="dashed")
        lines(L.tr, col="blue", lty="dashed")
        points((length(x.ts1)+1):(length(x.ts1)+12), pred.tr$pred, col="red"))

pred.orig <- (pred.tr$pred)^-2
U= (U.tr)^-2
L= (L.tr)^-2
ts.plot(sales1.ts, xlim=c(1,length(sales1.ts)+12), ylim = c(min(sales1.ts),max(U)),main = "Forecast of Sales",
        lines(U, col="blue", lty="dashed")
        lines(L, col="blue", lty="dashed")
        points((length(sales1.ts)+1):(length(sales1.ts)+12), pred.orig, col="red"))

#Spectral analysis
periodogram(x.bc,ylab='Periodogram of Transformed Sales'); abline(h=0)
library("GeneCycle")
fisher.g.test(res3)
cpgram(res3,main="")

```