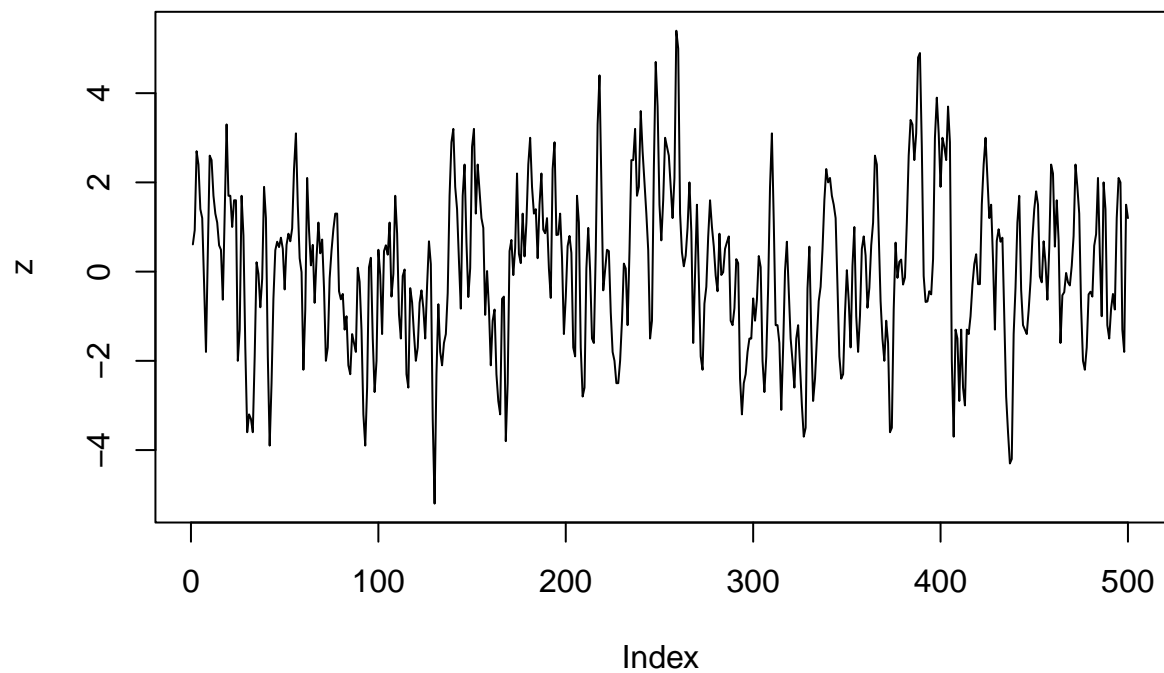# Time Series Analysis

## Anisha Mittal

### 2020-11-17

## 1. Read the given series.

**a)**

```r
z = scan("series.txt")
plot(z,type="l")
```



**Fitting an ARMA(1,1) model to our Time Series**

```r
Mdl1 <- arima(z,order = c(1,0,1))
Mdl1
```

```
##
## Call:
## arima(x = z, order = c(1, 0, 1))
##
## Coefficients:
##          ar1     ma1  intercept
##       0.4563  0.8041     0.0486
## s.e.  0.0427  0.0276     0.1505
##
## sigma^2 estimated as 1.034:  log likelihood = -718.9,  aic = 1443.8
```

```r
coeftest(Mdl1) #Z-test for coefficients of our model
```

```
##
## z test of coefficients:
##
##            Estimate Std. Error z value Pr(>|z|)
## ar1        0.456256   0.042740  10.675   <2e-16 ***
## ma1        0.804140   0.027620  29.115   <2e-16 ***
## intercept 0.048629   0.150537   0.323   0.7467
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Step 1: Plot the residuals**
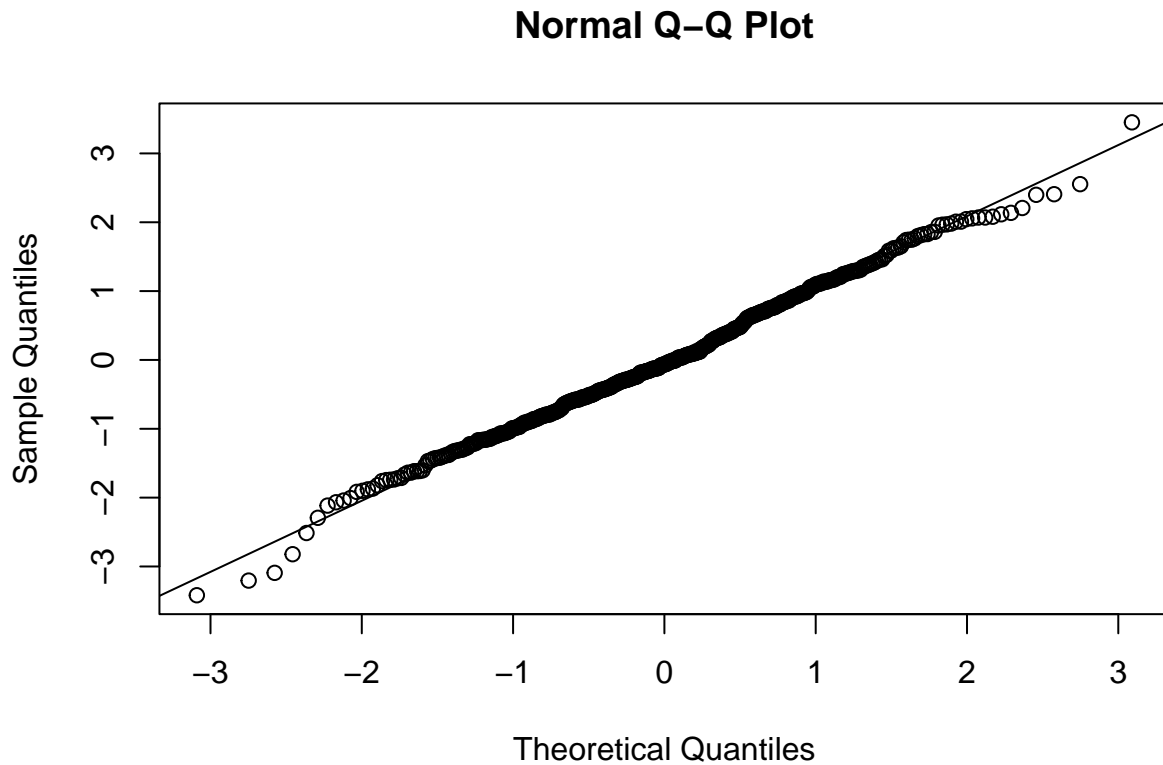
```r
plot(residuals(Mdl1))
```
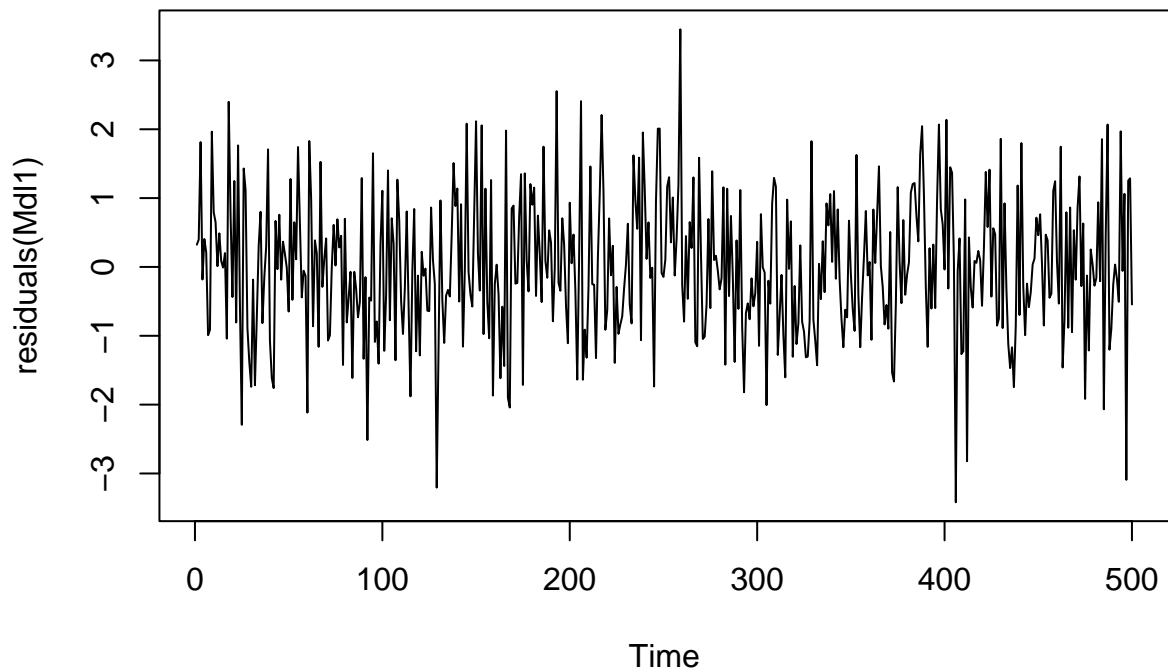
This is how the residuals look for our fitted ARMA(1,1) model.

**Step 2: qqplot of residuals to check their normality**

**Shapiro-Wilk test to test the normality of the residuals**

```r
qqnorm(residuals(Mdl1))
qqline(residuals(Mdl1))
```

## Normal Q–Q Plot



```r
shapiro.test(residuals(Mdl1))
```

```
## 
##  Shapiro-Wilk normality test
## 
## data:  residuals(Mdl1)
## W = 0.9958, p-value = 0.2035
```

By Q-Q Plot, we can see only a slight bit of deviation at the tails.

By Shapiro-wilk test, we get p-value greater than 5% significance level, therefore we will not reject the null hypothesis, i.e: we can infer that our data is normally distributed.

**Step 3: Residuals vs fitted values**

```r
plot(y=residuals(Mdl1), x = fitted(Mdl1) ,ylab = 'Residuals',
     xlab = 'Fitted Values', type = "p")
```
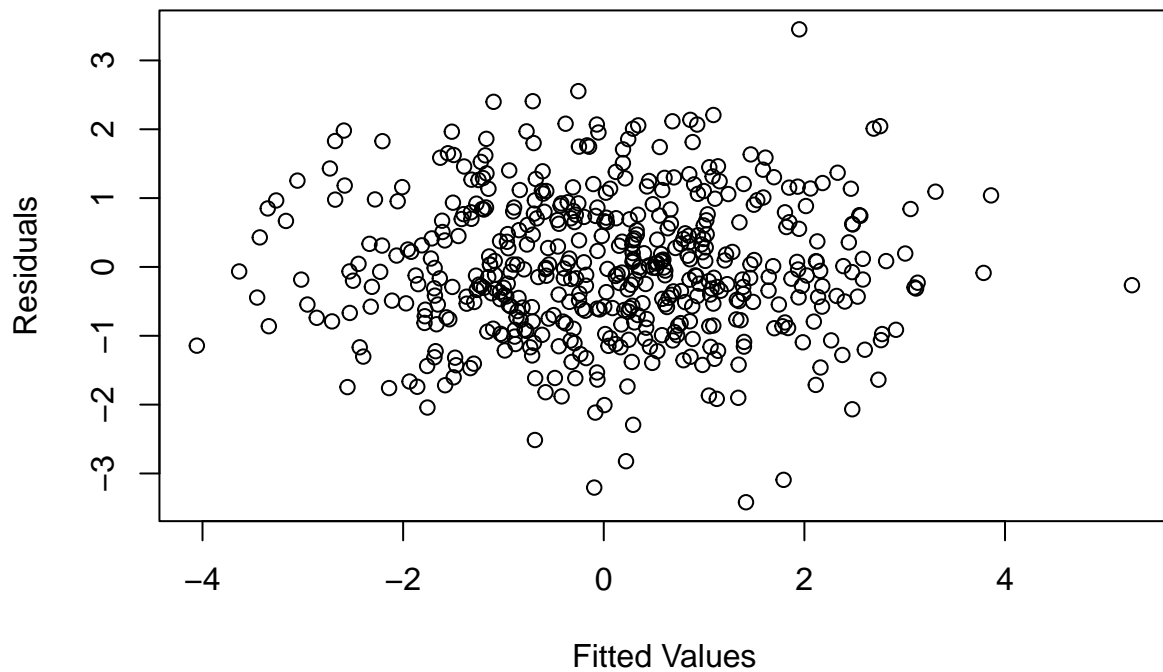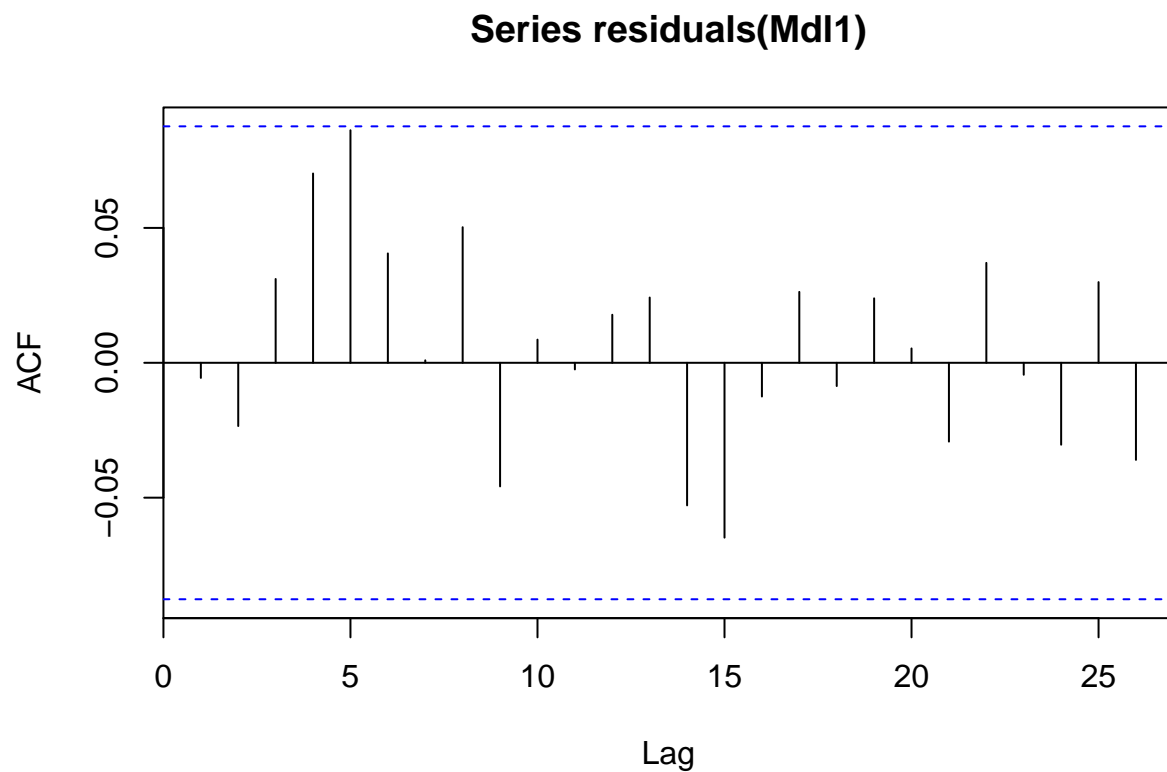
The above plot shows a bit of linear relation between the residuals and the actual data points with the fitted ARMA(1,1) Model.Its only when the fitted values are positive (>2) our plot is a bit scattered.

**Step 4: ACF of residuals**

**Ljung-Box test to check the auto correlation of the residuals**

```
acf(residuals(Mdl1))
```

## Series residuals(Mdl1)



```
Box.test(residuals(Mdl1),type='Ljung-Box',lag=1+1+10,fitdf=2)
```

```
##
##  Box-Ljung test
##
## data:  residuals(Mdl1)
## X-squared = 10.444, df = 10, p-value = 0.4024
```

From the ACF we can see lag 5 is almost near to the critical line but is not poking out of it.

By Ljung-Box-Pierce test, we get p-value greater than 5% significance level, Thus, we fail to reject the null hypothesis, i.e: we have no clear evidence that we should reject the residuals have no auto correlation and is not following white noise or is not dependent on it.

**Step 5: PACF of residuals**

```
pacf(residuals(Mdl1))
```

## Series  residuals(Mdl1)
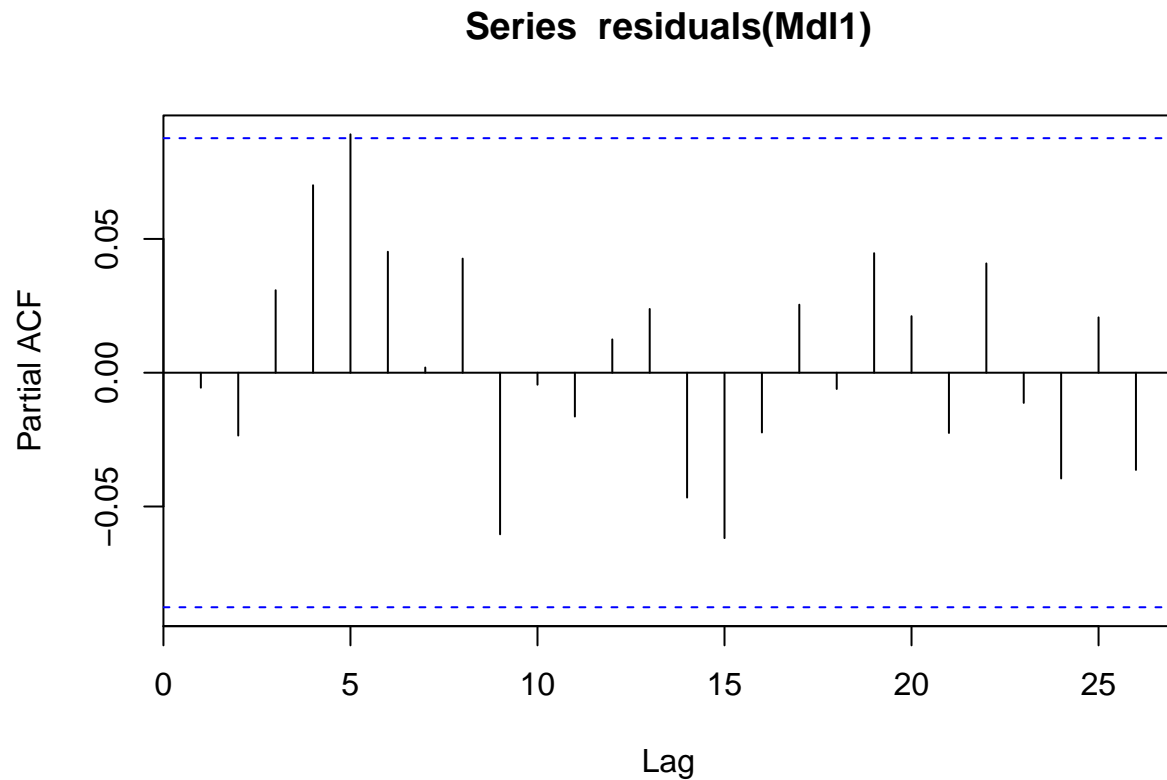


we can see lag 5 is poking out of the critical line, which is not too surprising since the expected number of phi hat greater than 0.05 is approximately more than 1.

## Exercise 2.

**a)**

**Fitting an AR(10) model to our Time Series**

```
Mdl2 <- arima(z,order = c(10,0,0))
Mdl2
```

```
##
## Call:
## arima(x = z, order = c(10, 0, 0))
##
## Coefficients:
##           ar1      ar2     ar3      ar4      ar5      ar6     ar7      ar8
##        1.2430  -1.0217  0.8647  -0.6293  0.5593  -0.4354  0.3221  -0.2091
## s.e.   0.0447   0.0715  0.0846   0.0921  0.0946   0.0947  0.0928   0.0858
##           ar9     ar10  intercept
##        0.0755  -0.0217     0.0515
## s.e.   0.0728   0.0455     0.1774
##
## sigma^2 estimated as 1.012:  log likelihood = -713.37,  aic = 1448.74
```

```
coeftest(Mdl2) #Z-test for coefficients of our model
```

```
## 
## z test of coefficients:
## 
##            Estimate Std. Error  z value  Pr(>|z|)
## ar1        1.242988   0.044721  27.7942 < 2.2e-16 ***
## ar2       -1.021732   0.071527 -14.2845 < 2.2e-16 ***
## ar3        0.864743   0.084606  10.2208 < 2.2e-16 ***
## ar4       -0.629305   0.092144  -6.8296 8.518e-12 ***
## ar5        0.559306   0.094624   5.9108 3.404e-09 ***
## ar6       -0.435447   0.094732  -4.5966 4.294e-06 ***
## ar7        0.322076   0.092797   3.4708  0.000519 ***
## ar8       -0.209059   0.085810  -2.4363  0.014838 *
## ar9        0.075462   0.072833   1.0361  0.300156
## ar10      -0.021671   0.045480  -0.4765  0.633724
## intercept  0.051515   0.177391   0.2904  0.771509
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Step 1: Plot the residuals**
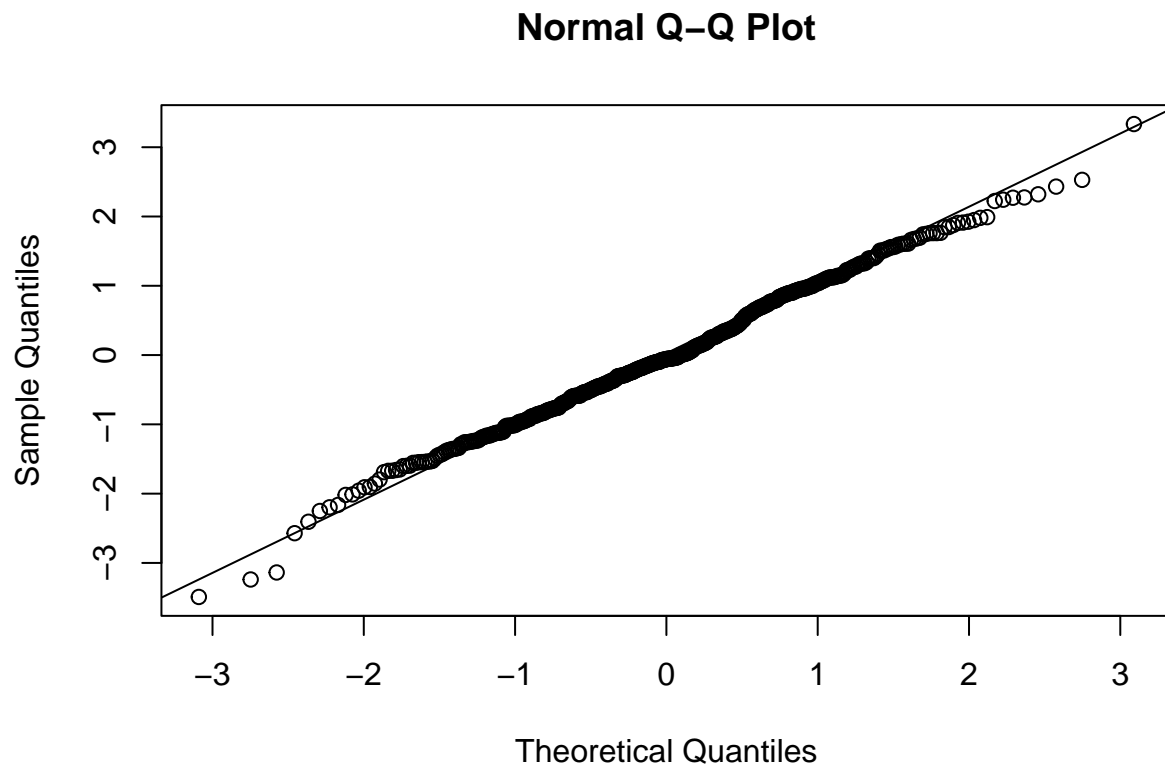
```
plot(residuals(Mdl2))
```
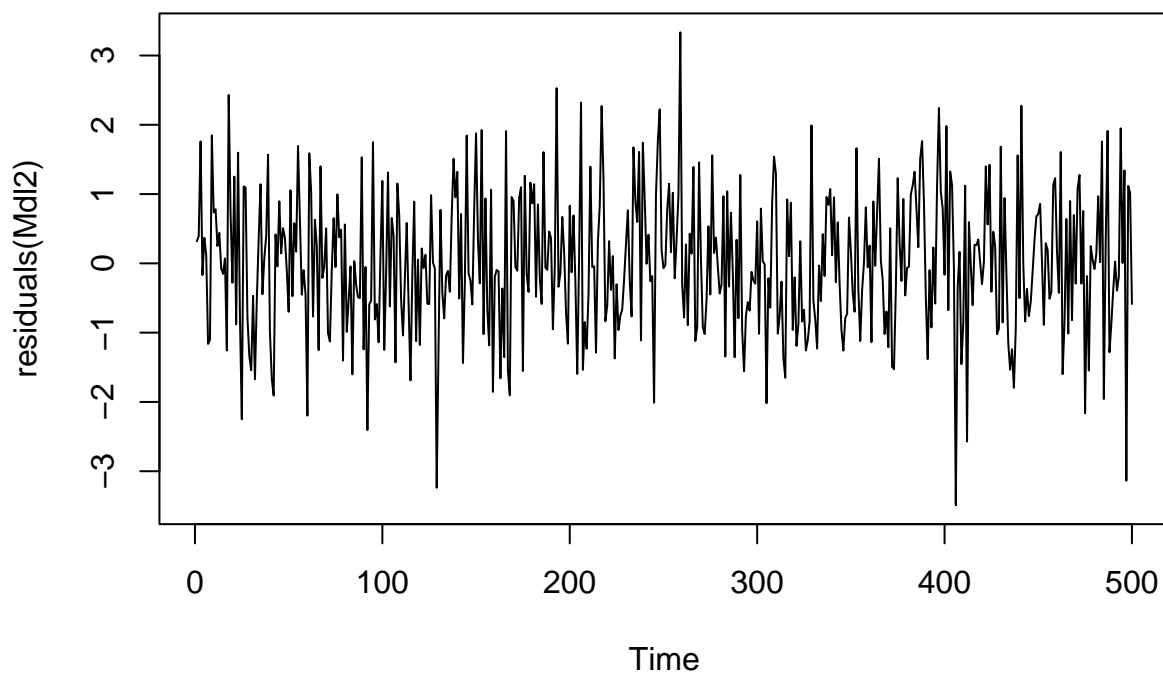
This is how the residuals look for our fitted AR(10) model.

**Step 2: qqplot of residuals to check their normality**

**Shapiro-Wilk test to test the normality of the residuals**

```
qqnorm(residuals(Mdl2))
qqline(residuals(Mdl2))
```

## Normal Q–Q Plot



```
shapiro.test(residuals(Mdl2))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(Mdl2)
## W = 0.99586, p-value = 0.2134
```

By Q-Q Plot, we can see a slight bit of deviation from the center and at the tails. But in comparison to ARMA(1,1) model, AR(10) model has more quantiles on the line at the tails.

By Shapiro-Wilk test, we get p-value greater than 5% significance level. Therefore we fail to reject the null hypothesis for the test. i.e: we can infer that our data is normally distributed.

**Step 3: Residuals vs fitted values**

```
plot(y=residuals(Mdl2), x = fitted(Mdl2) ,ylab = 'Residuals',
     xlab = 'Fitted Values', type = "p")
```
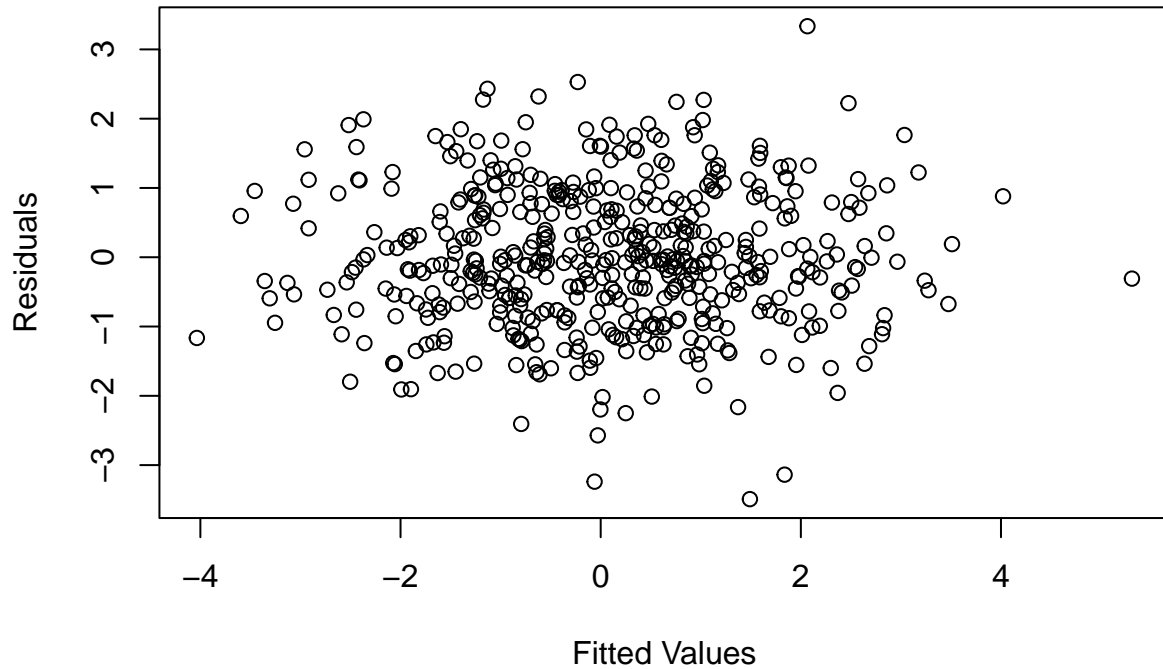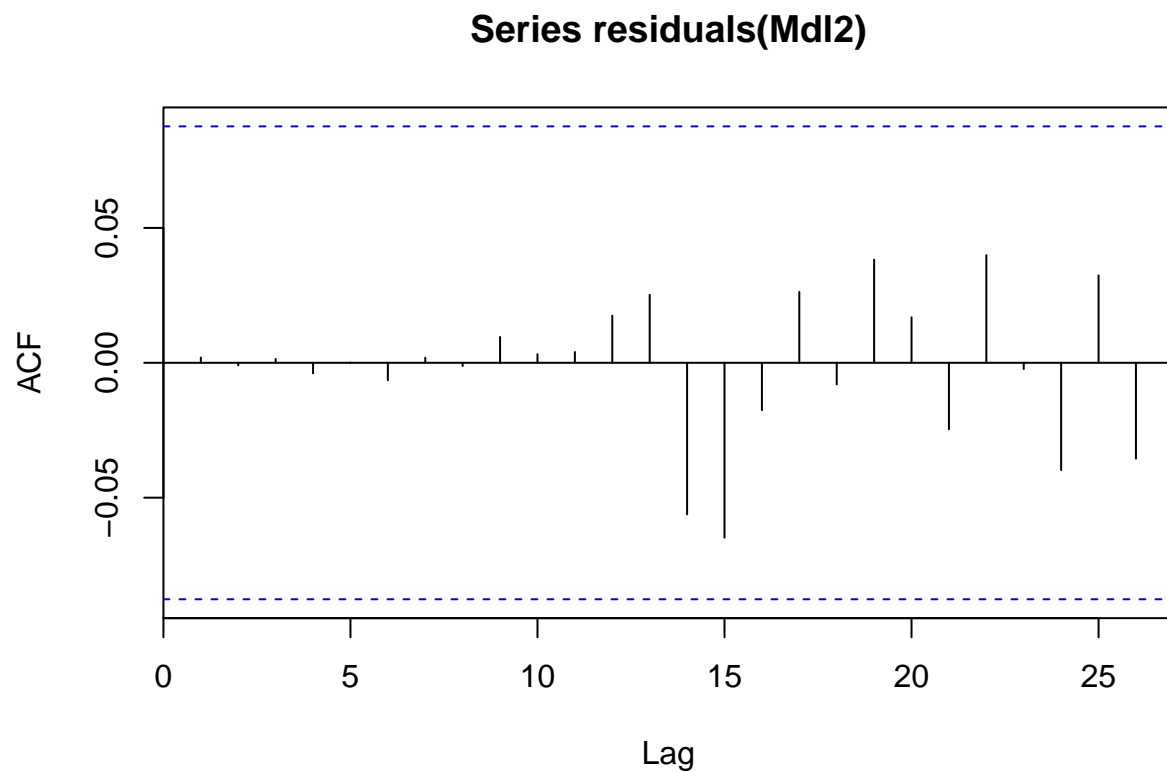


The above plot shows a bit of linear relation between the residuals and the actual data points with the fitted AR(10) Model.Its only when the fitted values are positive our plot is a bit scattered.

**Step 4: ACF of residuals**

**Ljung-Box test to check the auto correlation of the residuals**

```
acf(residuals(Mdl2))
```

## Series residuals(Mdl2)



```
Box.test(residuals(Mdl2),type='Ljung-Box',lag=10+0+10,fitdf=10)
```

```
##
##  Box-Ljung test
##
## data:  residuals(Mdl2)
## X-squared = 5.8498, df = 10, p-value = 0.8277
```
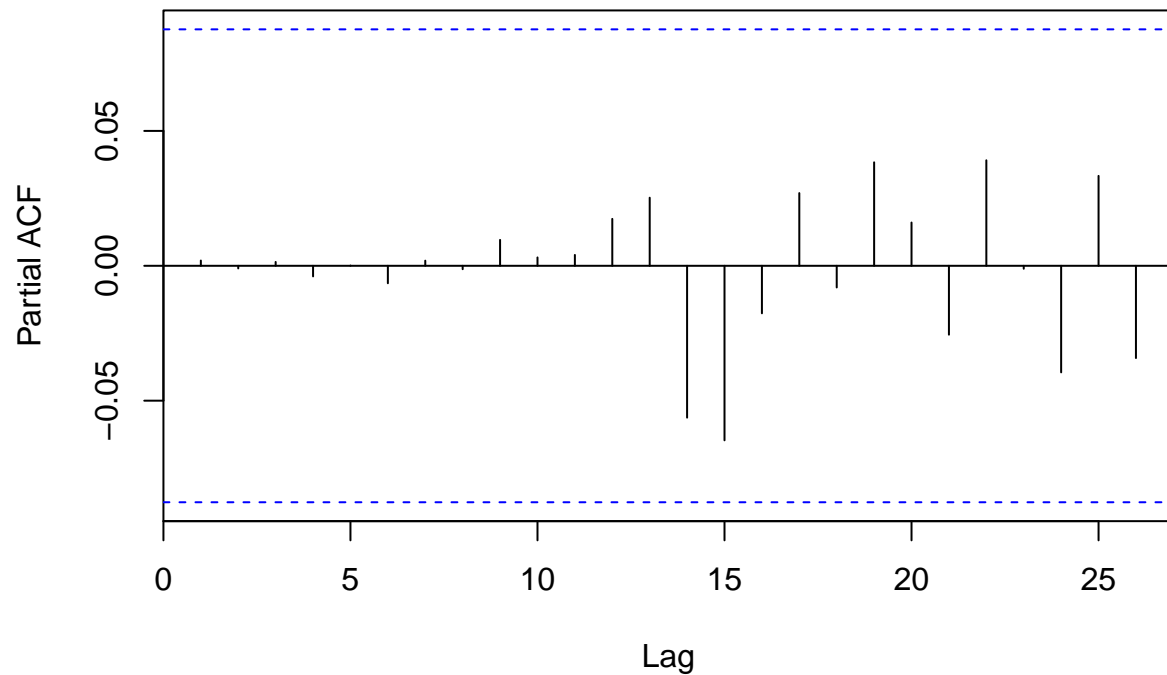
By the ACF plot we can see all are null and under the critical line, the value of q looks good.

By Ljung-Box-Pierce test we get p-value greater than 5% significance level. Therefore we fail to reject the null hypothesis for the test. we have no clear evidence that we should reject the residuals have no auto correlation and is not following white noise or is not dependent on it.

**Step 5: PACF of residuals**

```
pacf(residuals(Mdl2))
```

## Series residuals(Mdl2)



By the PACF plot we can see all are null and under the critical line, the value of p looks good.

**b)**

From the coeftest (z test of coefficients) it can be said that AR(1) to AR(7) are significant since there p-values are very small (almost near to 0), while AR(8) is significant at 5% significance level and AR(9), AR(10) and Intercept are not significant enough. By this we conclude that we have over fitted the value of p for the model.

**c)**

AR(10) model doesn't fit the series perfectly but it is still better than ARMA (1,1) model. Only a very small number of residual points from the tails shift towards the linear line when we move from ARMA (1,1) model to AR(10) model.