Residual Analysis

Srikar

2022-04-11

Introduction

In time series models, the residuals are equal to the difference between the observations and the corresponding fitted values:
#Residuals are useful in checking whether a model has adequately captured the information in the data. A good forecasting method will yield residuals with the following properties:

#The residuals are uncorrelated. If there are correlations between residuals, then there is information left in the residuals which should be used in computing forecasts.
#The residuals have zero mean. If the residuals have a mean other than zero, then the forecasts are biased.

Data Description

#The dataset contains the number of earthqukes from the year 1916 to 2015 that occured in Japan.

Objective

#1) Fit a suitable ARMA model
#2) Perform residual analysis for the fitted model

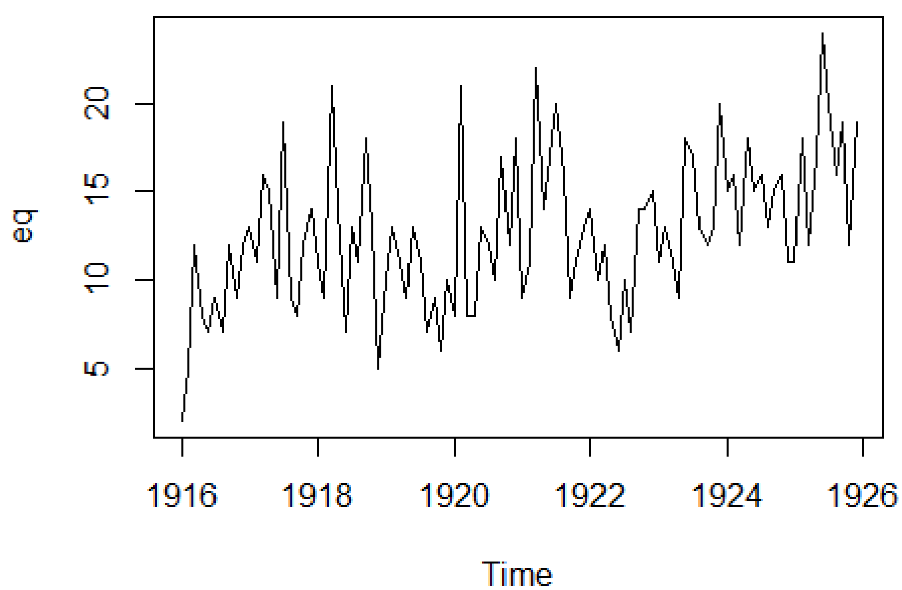
Analysis

```
library(tseries)

#1) Importing the dataset
```

```r
EQ <- read.delim("C:/Users/Srikar/Desktop/SS/R/Sem 6/Practical
8/earthquakes.txt")
head(EQ)
```

```
##   Year Quakes
## 1 1916      2
## 2 1917      5
## 3 1918     12
## 4 1919      8
## 5 1920      7
## 6 1921      9
```

```r
#Making it a time series object

eq=ts(EQ$Quakes,start = 1916,frequency = 10)

#2) Plotting the dataset

ts.plot(eq)
```




```r
adf.test(eq)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  eq
```

```
## Dickey-Fuller = -3.452, Lag order = 4, p-value = 0.04991
## alternative hypothesis: stationary
```

```
#We observe that it is stationary

#3) Fiiting an ARMA model

library(forecast)
```

```
## Warning: package 'forecast' was built under R version 3.6.3
```

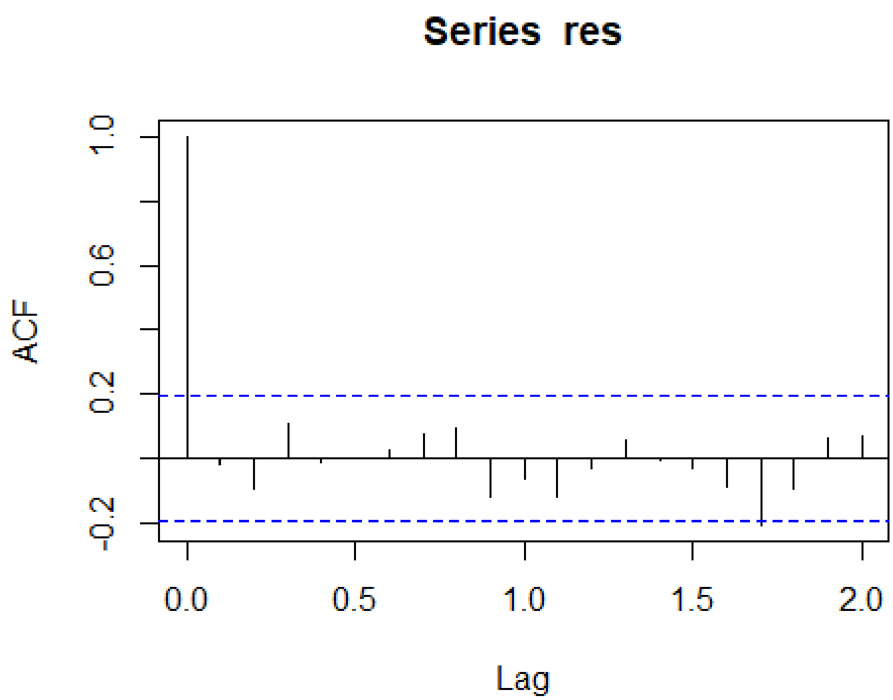
```
fit=auto.arima(eq,seasonal=FALSE)
fit
```

```
## Series: eq
## ARIMA(0,1,1)
##
## Coefficients:
##           ma1
##       -0.8092
## s.e.   0.0710
##
## sigma^2 estimated as 15.53:  log likelihood=-276.26
## AIC=556.53   AICc=556.65   BIC=561.72
```

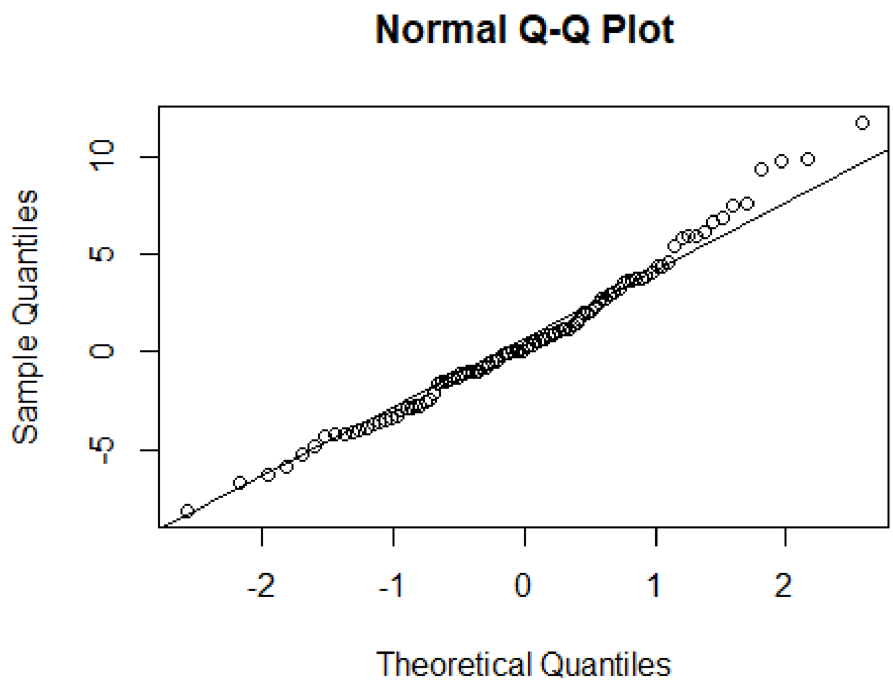
```
#The best fitted model is ARMA(0,1). The equation can be given as -0.8092a


#4) Fiiting the model

arim=arima(order=c(0,1,1),eq)
```

Residual Analysis

```
res=resid(fit)


#1.Checking for uncorrelated errors

acf(res)
```




```
#From the ACF plot, it is evident that none of the errors are autocorrelated.

#2.Checking if residuals are normally distributed

qqnorm(res)
qqline(res)
```




```
#We observe that some observations do not fall on the line.To confirm
normality,we can use the Shapiro-wilk test.

shapiro.test(res)

##
##  Shapiro-Wilk normality test
##
## data:  res
## W = 0.98273, p-value = 0.2156

#Since p-value is greater than 0.05, we accept the NULl Hypothesis and say
that it the errors are normally distributed.
```

Conclusion

The best suited ARMA model is an ARMA(0,1) and from the residual analysis, the assumption of no autocorrelation and normally distributed are satisfied.