Time Series Analysis (STAT 560)

Homework 7

Bindu Paudel and Gena Ram Mahato

11/19/2020

5.12 Table B.3 contains data on chemical process viscosity.

a. Fit an ARIMA model to this time series, excluding the last 20 observations. Investigate model adequecy. Explain how this model would be used for forecasting.

Description of data

```
library(kableExtra)
setwd("~/Desktop/SDSU Fall 2020/Timeseries/Homework 7")
B.3 <- read.csv("B.3.csv", header=TRUE, sep=",")
kable(head(B.3), caption = "First six rows of Table B.3", booktabs=T, linesep ="")%>%
kable_styling(latex_option="hold_position")
```

Table 1: First six rows of Table B.3

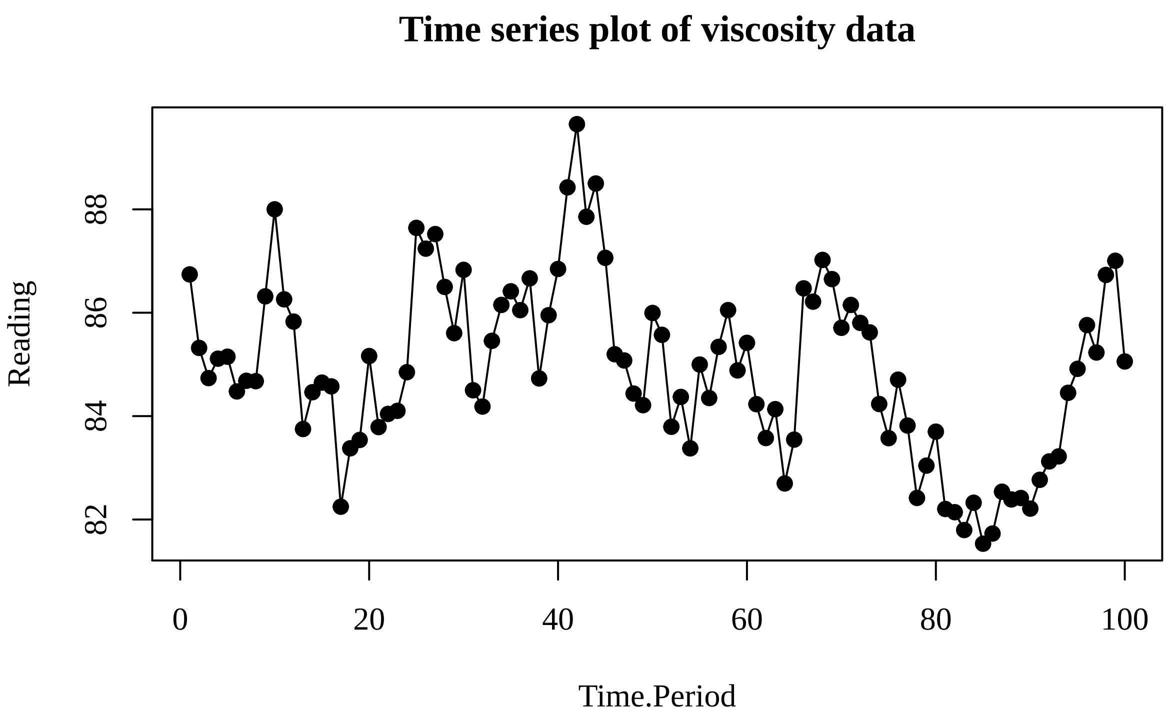
Time.Period	Reading
1	86.7418
2	85.3195
3	84.7355
4	85.1113
5	85.1487
6	84.4775

These are the first six rows of chemical process viscosity data. The data contains two variables time period and reading. And, there are 100 observations in data. For this question, we will identify a best fit ARIMA model excluding last 20 observations and will use the best ARIMA model to forecast last 20 observations.

Time series plot of the data

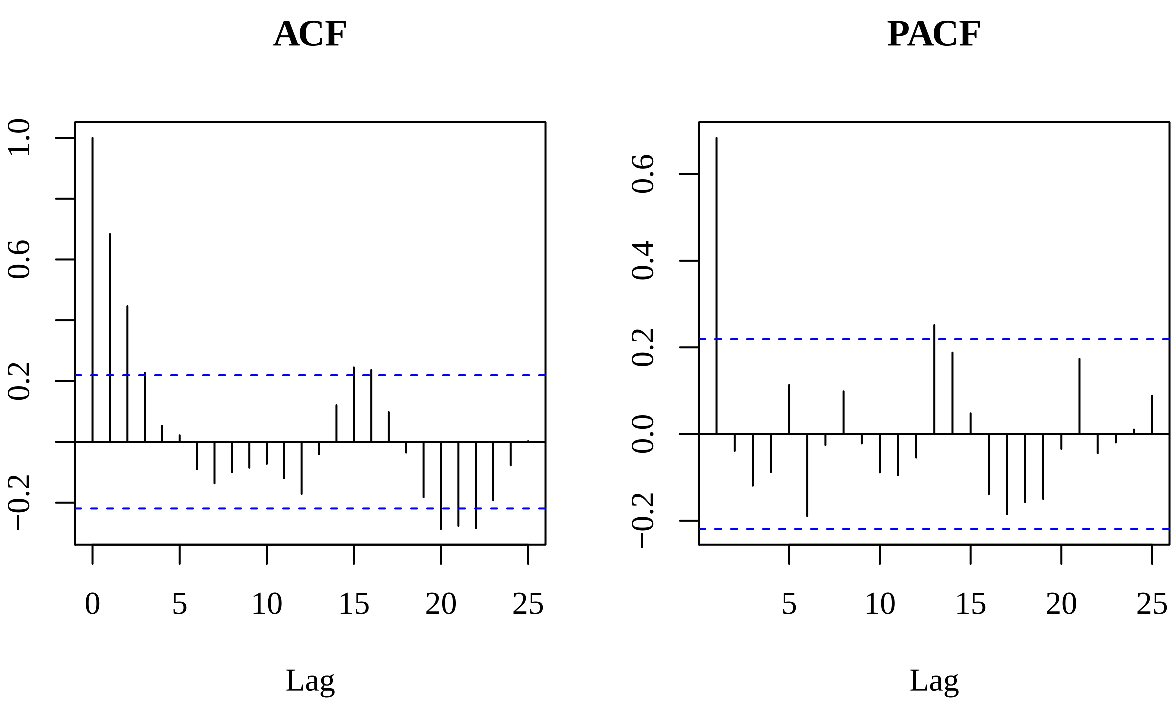
The plot below is the time series plot of viscosity data. The data shows slight fluctuation within a limited range. Moreover, looking into the time series plot, the data looks stationary. We will further confirm it by looking into ACF and PACF plots/values of the data.

```
par(family="Times")
plot(B.3, type = "l", pch = "19", col="black", main = "Time series plot of viscosity data")
points(B.3, pch = 19, col="black")
```



ACF and PACF plots of viscosity data

```
par("mar"=c(4,3,4,1))
par(family="Times")
par(mfrow=c(1,2),oma=c(0,0,0,0))
ACF <- acf(B.3[,2][1:80],lag.max=25,type="correlation",main="ACF")
PACF <- acf(B.3[,2][1:80], lag.max=25,type="partial",main="PACF")
```



Interpretation:

Looking into ACF and PACF plots of viscosity data, the data looks stationary. The ACF values quickly reduced to zero within the significant range. The PACF value cuts at lag 1 and then after it quickly reduced to zero, giving a hint of ARIMA(1, 0, 0) model. These plots shows that the data is stationary.

Calculated values of ACF and PACF of the data

```r
acf <- as.data.frame(ACF$acf[2:26]) # Excluding 0 lag
pacf <- as.data.frame(PACF$acf[1:25])

K <- 1:25
cf <- as.data.frame(cbind("Lag"=K, "Sample ACF"=acf,"Sample PACF"=pacf))
colnames(cf) <- c("Lag", "Sample ACF", "Sample PACF")
kable(round(cf, 3),
caption = "Table showing Lag, Sample ACF and Sample PACF Values of the Data", booktabs=T,
linesep = "")%>%
kable_styling(latex_option="hold_position")
```

Table 2: Table showing Lag, Sample ACF and Sample PACF Values of the Data

Lag	Sample ACF	Sample PACF
1	0.683	0.683
2	0.446	-0.039
3	0.227	-0.119
4	0.053	-0.088
5	0.022	0.113
6	-0.090	-0.190
7	-0.137	-0.025
8	-0.100	0.099
9	-0.085	-0.022
10	-0.072	-0.089
11	-0.120	-0.095
12	-0.171	-0.054
13	-0.041	0.251
14	0.121	0.188
15	0.245	0.048
16	0.237	-0.139
17	0.098	-0.185
18	-0.035	-0.157
19	-0.182	-0.150
20	-0.287	-0.034
21	-0.277	0.174
22	-0.284	-0.044
23	-0.193	-0.020
24	-0.077	0.011
25	0.002	0.089

Interpretation:

The table above shows calculated ACF and PACF values of viscosity data using lag.max value of 25.

Fitting ARIMA(1, 0, 0) model

```
B.3.fit.2 <-arima(B.3[,2][1:80],order=c(1, 0, 0))
B.3.fit.2
```

```
##
## Call:
## arima(x = B.3[, 2][1:80], order = c(1, 0, 0))
##
## Coefficients:
##          ar1  intercept
##       0.6934    85.2721
## s.e.  0.0802     0.3756
##
## sigma^2 estimated as 1.121:  log likelihood = -118.42,  aic = 242.84
```

Interpretation:

This is the summary of ARIMA(1, 0, 0) model with its estimates. The AIC value of ARIMA(1, 0, 0) model is found to be 242.84.

Fitting ARIMA(2, 0, 0) model

```
B.3.fit.2 <-arima(B.3[,2][1:80],order=c(2, 0, 0))
B.3.fit.2
```

```
##
## Call:
## arima(x = B.3[, 2][1:80], order = c(2, 0, 0))
##
## Coefficients:
##          ar1      ar2  intercept
##       0.7195  -0.0378    85.2753
## s.e.  0.1120   0.1133     0.3627
##
## sigma^2 estimated as 1.12:  log likelihood = -118.36,  aic = 244.73
```

Interpretation:

This is the summary of ARIMA(2, 0, 0) model with its estimates. The AIC value of ARIMA(2, 0, 0) model is found to be 244.73 which is higher than ARIMA(1, 0, 0) model. Comparing both models, ARIMA(1, 0, 0) model makes a better fit to the data. This was also confirmed by ACF and PACF plots of the data above. However, we will use auto.arima funtion to double check.

Using auto.arima() function to double check

```
library(forecast)
B.3.fit.3 <-auto.arima(B.3[,2][1:80])
B.3.fit.3
```

```
## Series: B.3[, 2][1:80]
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##          ar1     mean
##       0.6934  85.2721
## s.e.  0.0802   0.3756
##
## sigma^2 estimated as 1.15:  log likelihood=-118.42
## AIC=242.84   AICc=243.16   BIC=249.99
```

Interpretation:

The auto.arima model also shows ARIMA(1, 0, 0) model as best model. The equation of fitted ARIMA(1, 0, 0) model is given below.

The δ for ARIMA(1, 0, 0) model is given by:

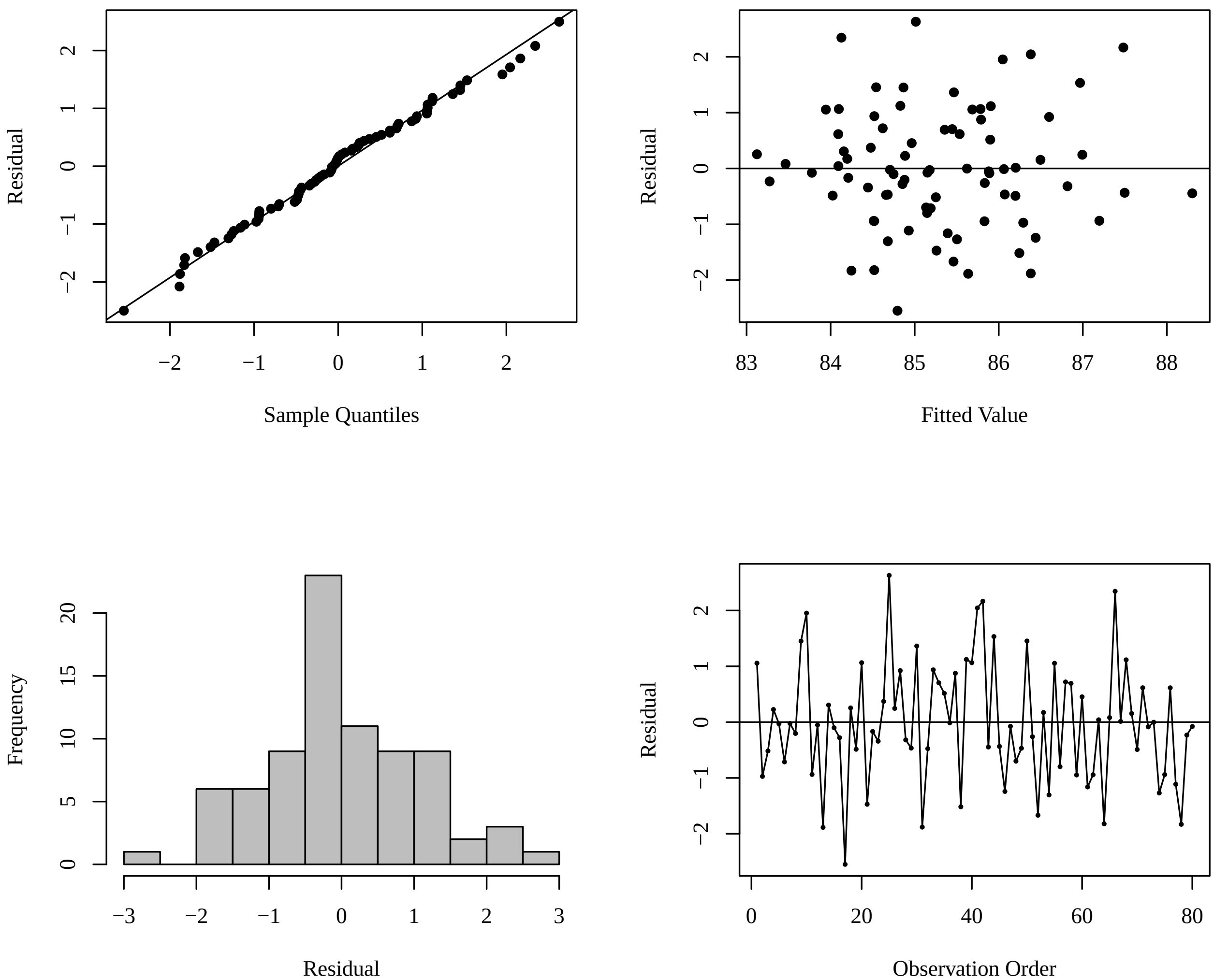
$$\delta = \mu * (1 - \phi_1) = 85.2721 * (1 - 0.6934) = 26.14$$

The best fitted model:

$$y_t = 26.14 + 0.6934 y_{t-1} + \epsilon_t$$

4 in 1 residual plots of fitted ARIMA(1, 0, 0) model

```
par(family="Times")
par(mfrow=c(2,2),oma=c(0,0,0,0))
res.B.3 <-as.vector(residuals(B.3.fit.2))
fit.B.3 <-as.vector(fitted(B.3.fit.2))
qqnorm(res.B.3,datax=TRUE,pch=16,xlab='Residual',main='')
qqline(res.B.3,datax=TRUE)
plot(fit.B.3, res.B.3,pch=16, xlab='Fitted Value',
ylab='Residual')
abline(h=0)
hist(res.B.3,col="gray",xlab='Residual',main='')
plot(res.B.3,type="l",xlab='Observation Order',
ylab='Residual')
points(res.B.3,pch=16,cex=.5)
abline(h=0)
```



Interpretation:

QQ plot:

Normal probability plot is the plot that is used to observe the individual data and find out whether the data is normally distributed or not. From the above qq norm plot, we can see that most of the data are distributed close to qqnorm line. Hence we can say that the residuals are normally distributed.

Histogram of residuals:

Histogram are useful for large sample of data than small ones. Histogram also seems to have normal distribution, however it has a gap. Moreover, we can say that the residuals are normally distributed.

Residual vs fitted values:

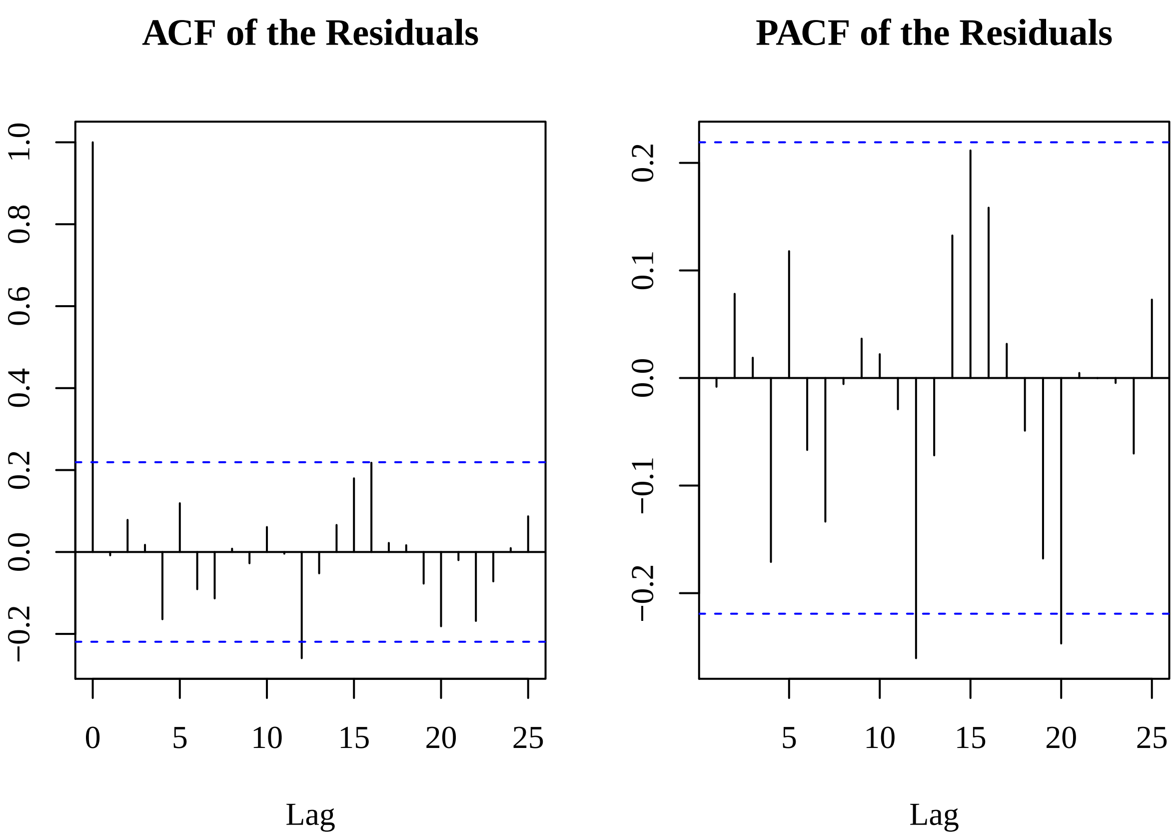
The residuals vs fitted value shows that the residuals are randomly distributed around the line and does no show any pattern. The residuals do not have perfect even distribution but looks good in case of distribution. Moreove, the residuals shows good fit of model.

Residual vs the order of the data:

This plot reveals how the data changes with time. Here, the plot indicates that variability of the observations are changing with time.

ACF and PACF plots of the residuals

```r
par("mar"=c(4,3,4,1))
par(family="Times")
par(mfrow=c(1,2),oma=c(0,0,0,0))
ACF <- acf(res.B.3,lag.max=25,type="correlation",main="ACF of the Residuals")
PACF <- acf(res.B.3, lag.max=25,type="partial",main="PACF of the Residuals")
```



Interpretation:

Looking into ACF and PACF plots of residuals, the residual looks stationary. The acf values quickly reduced to zero and is within the significant range. The PACF value also lies within a significant range. These plots shows that the residuals of fitted ARIMA(1, 0, 0) model are stationary.

Calculated ACF and PACF values of the residuals

```
acf <- as.data.frame(ACF$acf[2:26]) # Excluding 0 lag
pacf <- as.data.frame(PACF$acf[1:25])

K <- 1:25
cf <- as.data.frame(cbind("Lag"=K, "Sample ACF"=acf,"Sample PACF"=pacf))
colnames(cf) <- c("Lag", "Sample ACF", "Sample PACF")
kable(round(cf, 3),
caption = "Table showing Lag, Sample ACF and Sample PACF Values of Residuals", booktabs=T,
linesep = "")%>%
kable_styling(latex_option="hold_position")
```

Table 3: Table showing Lag, Sample ACF and Sample PACF Values of Residuals

Lag	Sample ACF	Sample PACF
1	-0.008	-0.008
2	0.078	0.078
3	0.018	0.019
4	-0.164	-0.171
5	0.119	0.118
6	-0.091	-0.067
7	-0.113	-0.134
8	0.008	-0.006
9	-0.027	0.037
10	0.061	0.022
11	-0.004	-0.029
12	-0.259	-0.260
13	-0.052	-0.072
14	0.066	0.132
15	0.180	0.212
16	0.218	0.158
17	0.022	0.032
18	0.017	-0.049
19	-0.077	-0.168
20	-0.181	-0.247
21	-0.020	0.005
22	-0.168	0.000
23	-0.072	-0.005
24	0.009	-0.070
25	0.087	0.073

Interpretation:

The table shows calculated ACF and PACF values of residuals of ARIMA(1, 0, 0) using lag.max value of 25.

Calculating Ψ's values of fitted ARIMA(1, 0, 0) model

General forecasting model

$$Eq(1): \hat{y}_{T+\tau}(T) = E[y_{T+\tau}|y_T, y_{T-1}, ...] = \mu + \sum_{i=\tau}^{\infty} \Psi_i \epsilon_{T+\tau-i}$$

Infinite MA representaion of $y_{T+\tau}$ can be written as,

$$Eq(2): y_{T+\tau} = (\Psi_0 + \Psi_1 B + \Psi_2 B^2 + \Psi_3 B^3.....)\epsilon_{T+\tau}$$

ARIMA model(1, 0, 0) can be written as:

$$Eq(3): (1 - \phi_1 B)y_{T+\tau} = \epsilon_{T+\tau}$$

Combining Eq(2) and Eq(3), we have

$$(\Psi_0 + \Psi_1 B + \Psi_2 B^2 + \Psi_3 B^3.....)(1 - \phi_1 B) = 1$$

Equating power of B, we have

$$\begin{gathered}
B^0 : \Psi_0 = 1 \\
B^1 : \Psi_1 - \Psi_0\phi_1 = 0 \rightarrow \Psi_1 = \Psi_0\phi_1 = 1 * 0.6934 = 0.6834 \\
B^2 : \Psi_2 - \Psi_1\phi_1 = 0 \rightarrow \Psi_2 = \Psi_1\phi_1 = 0.6934 * 0.6934 = 0.48 \\
B^3 : \Psi_3 - \Psi_2\phi_1 = 0 \rightarrow \Psi_3 = \Psi_2\phi_1 = 0.48 * 0.6934 = 0.33
\end{gathered}$$

b. Forecast the last 20 observations.

```
library(forecast)
B.3.fit.3 <-auto.arima(B.3[,2][1:80])
B.3.forecast<-as.array(forecast(B.3.fit.3,h=20))
kable(B.3.forecast,
caption = "Table showing point forecast, 80 percent and  95 percent prediction intervals",
booktabs=T,
linesep = "")%>%
kable_styling(latex_option="hold_position")
```

Table 4: Table showing point forecast, 80 percent and 95 percent prediction intervals

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
81	84.18154	82.80723	85.55585	82.07972	86.28337
82	84.51592	82.84356	86.18828	81.95827	87.07357
83	84.74777	82.94962	86.54592	81.99774	87.49780
84	84.90853	83.05294	86.76412	82.07064	87.74642
85	85.02000	83.13741	86.90258	82.14083	87.89917
86	85.09729	83.20186	86.99271	82.19848	87.99609
87	85.15088	83.24930	87.05245	82.24267	88.05908
88	85.18803	83.28352	87.09255	82.27532	88.10074
89	85.21380	83.30787	87.11973	82.29893	88.12867
90	85.23166	83.32505	87.13828	82.31575	88.14758
91	85.24405	83.33711	87.15099	82.32764	88.16046
92	85.25264	83.34554	87.15974	82.33599	88.16929
93	85.25860	83.35142	87.16577	82.34183	88.17536
94	85.26272	83.35552	87.16993	82.34590	88.17955
95	85.26559	83.35836	87.17281	82.34874	88.18244
96	85.26757	83.36034	87.17481	82.35071	88.18444
97	85.26895	83.36171	87.17619	82.35208	88.18582
98	85.26990	83.36266	87.17714	82.35303	88.18678
99	85.27057	83.36332	87.17781	82.35369	88.18744
100	85.27102	83.36378	87.17827	82.35415	88.18790

Interpretation:

These are the forecast values of last 20 observations of viscosity data. The table above shows point forecast, 80 percent and 95 percent prediction intervals.

c. Show how to obtain prediction intervals for the forecasts in part b above.

For ARIMA (1, 0, 0)

The infinte MA representation coefficients $\Psi_i = \phi^i$,

$$Var[e_T(\tau)] = \sigma^2 \sum_{i=0}^{\tau-1} \Psi_i^2 = \sigma^2 \sum_{i=0}^{\tau-1} \phi^{2i} = \sigma^2 \frac{1-\phi^{2\tau}}{1-\phi^2}$$

Thus the 100(1-α) PI for $y_{T+\tau}$ is,

$$\hat{y}_{T+\tau}(T) \pm Z_{\alpha/2} * \sqrt{Var[e_T(\tau)]}$$
$$\hat{y}_{T+\tau}(T) \pm Z_{\alpha/2} * \sigma\sqrt{1-\phi^{2\tau}/1-\phi^2}$$

5.33 Table B.15 presents data on the occurrence of violent crimes. Develop an appropriate ARIMA model and a procedure for forecasting for these data. Explain how prediction intervals would be computed.

```
B.15 <- read.csv("B.15.csv", header=TRUE, sep=",")
kable(head(B.15), caption = "First six rows of Table B.15", booktabs=T, linesep ="")%>%
kable_styling(latex_option="hold_position")
```

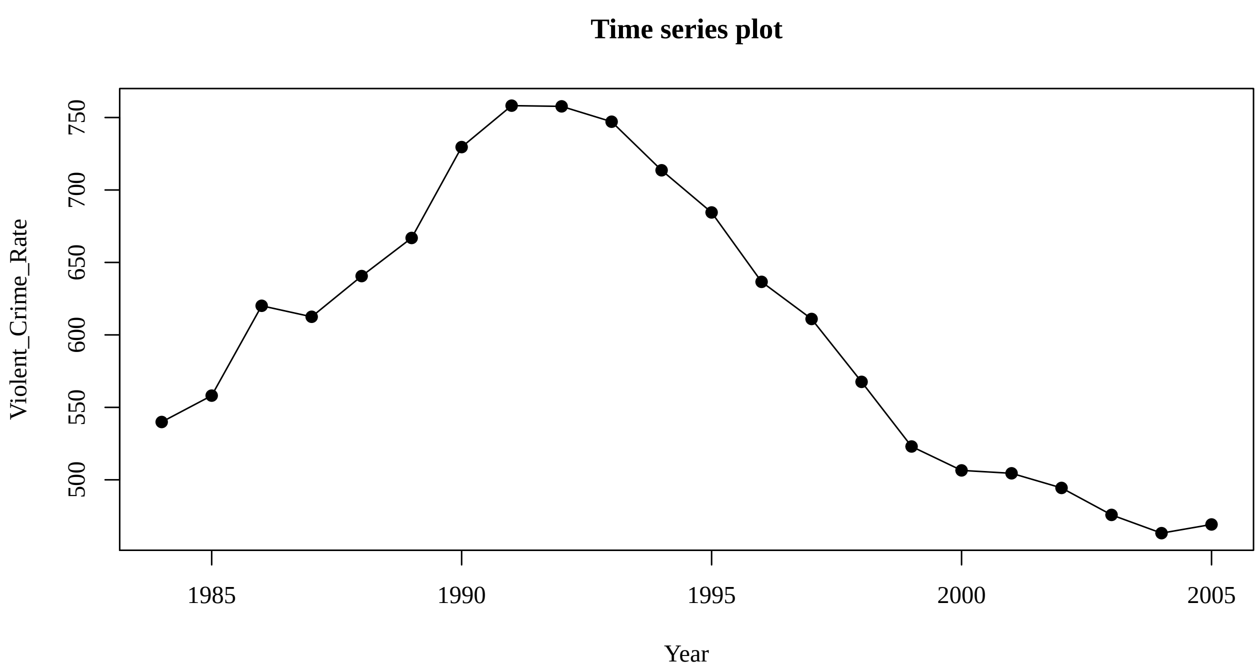
Table 5: First six rows of Table B.15

Year	Violent_Crime_Rate
1984	539.9
1985	558.1
1986	620.1
1987	612.5
1988	640.6
1989	666.9

These are the first six rows of data on the occurrence of violent crimes. The data contains two variables year and violent crime rate. And, there are 22 observations in data.

Time series plot of the violent crimes data

```
par(family="Times")
plot(B.15, type = "l", pch = "19", col="black", main="Time series plot")
points(B.15, pch = 19, col="black")
```

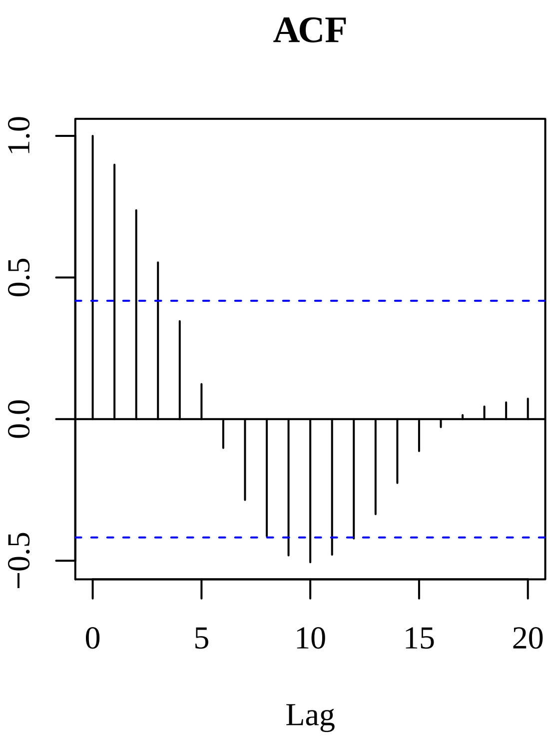


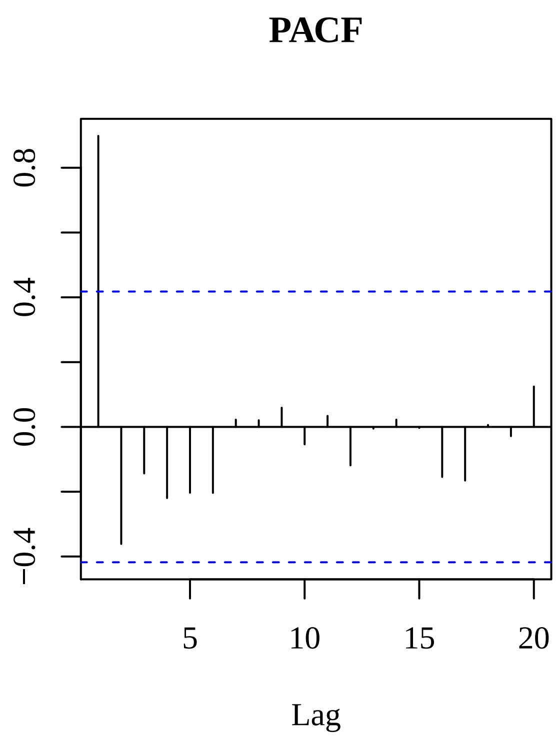
Interpretation:

The plot above is a time series plot of violent crimes data. The data shows a major trend of increasing and decreasing. Moreover, looking into the time series plot, the data does not look stationary. We will further confirm it by looking into ACF and PACF plots/values of the data.

ACF and PACF plots of the violent crimes data

```
par("mar"=c(4,3,4,1))
par(family="Times")
par(mfrow=c(1,2),oma=c(0,0,0,0))
ACF <- acf(B.15[,2],lag.max=20,type="correlation",main="ACF")
PACF <- acf(B.15[,2], lag.max=20,type="partial",main="PACF")
```





Interpretation:

Looking into the ACF values, the does not look stationary. The ACF value in the plot decreased to zero after lag 6 and are outside the significant range after being reduced to zero. So, the data is not stationary as it shows many variations.

Calculated ACF and PACF values of the violent crimes data

```
acf <- as.data.frame(ACF$acf[2:21]) # Excluding 0 lag
pacf <- as.data.frame(PACF$acf[1:20])

K <- 1:20
cf <- as.data.frame(cbind("Lag"=K, "Sample ACF"=acf,"Sample PACF"=pacf))
colnames(cf) <- c("Lag", "Sample ACF", "Sample PACF")
kable(round(cf, 3),
caption = "Table showing Lag, Sample ACF and Sample PACF Values", booktabs=T,
linesep = "")%>%
kable_styling(latex_option="hold_position")
```

Table 6: Table showing Lag, Sample ACF and Sample PACF Values

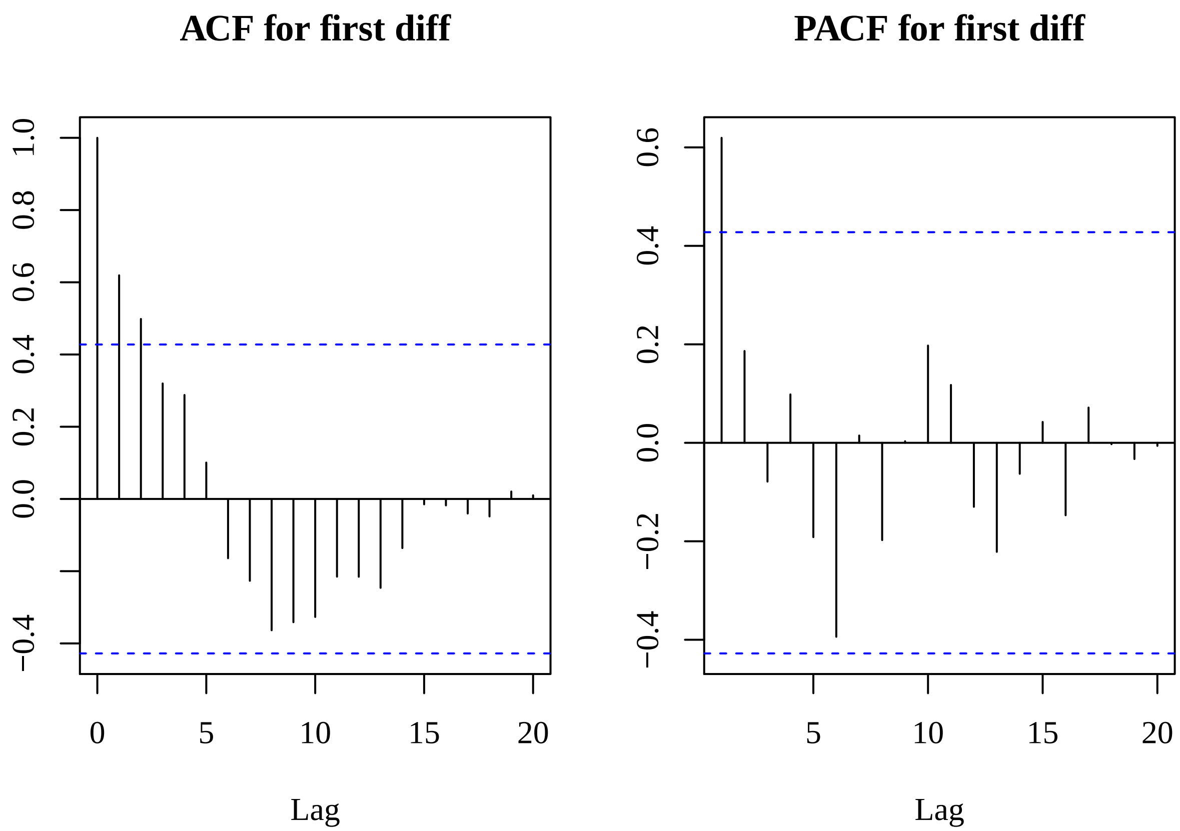
Lag	Sample ACF	Sample PACF
1	0.898	0.898
2	0.737	-0.361
3	0.553	-0.144
4	0.346	-0.220
5	0.124	-0.203
6	-0.102	-0.204
7	-0.285	0.022
8	-0.413	0.021
9	-0.481	0.059
10	-0.505	-0.054
11	-0.479	0.034
12	-0.422	-0.119
13	-0.336	-0.005
14	-0.225	0.022
15	-0.113	-0.003
16	-0.028	-0.155
17	0.014	-0.166
18	0.045	0.006
19	0.059	-0.028
20	0.072	0.125

Interpretation:

The above table shows calculated ACF and PACF values of violent crimes data using lag.max value of 20.

Since the time series is non stationary, we will take first difference of the data set and we will plot ACF and PACF of first differenced data set.

```
par("mar"=c(4,3,4,1))
par(mfrow=c(1,2))
par(family="Times")
first_diff <- diff(B.15[,2],lag=1)
acf(first_diff, lag.max = 20, type="correlation",main="ACF for first diff")
acf(first_diff, lag.max = 20, type="partial",main="PACF for first diff")
```



Interpretation:

The acf plot of first difference shows cut at lag 1 and 2 and has a mixture of exponential decay. However, plot of pacf shows that there is cut off at lag 1 only and there is a mixture of exponential decay in suggesting that time series can be fitted with ARIMA model. First we will try to fit the data set with ARIMA(1,1,0) model and then we will take second difference of the data set and fit ARIMA(1, 2, 0) model. At final, we will comapre the parameters of both models.

Fitting ARIMA (1,1,0) model

```
B.15.fit.1 <-arima(B.15[,2], order =c(1,1,0))
B.15.fit.1
```

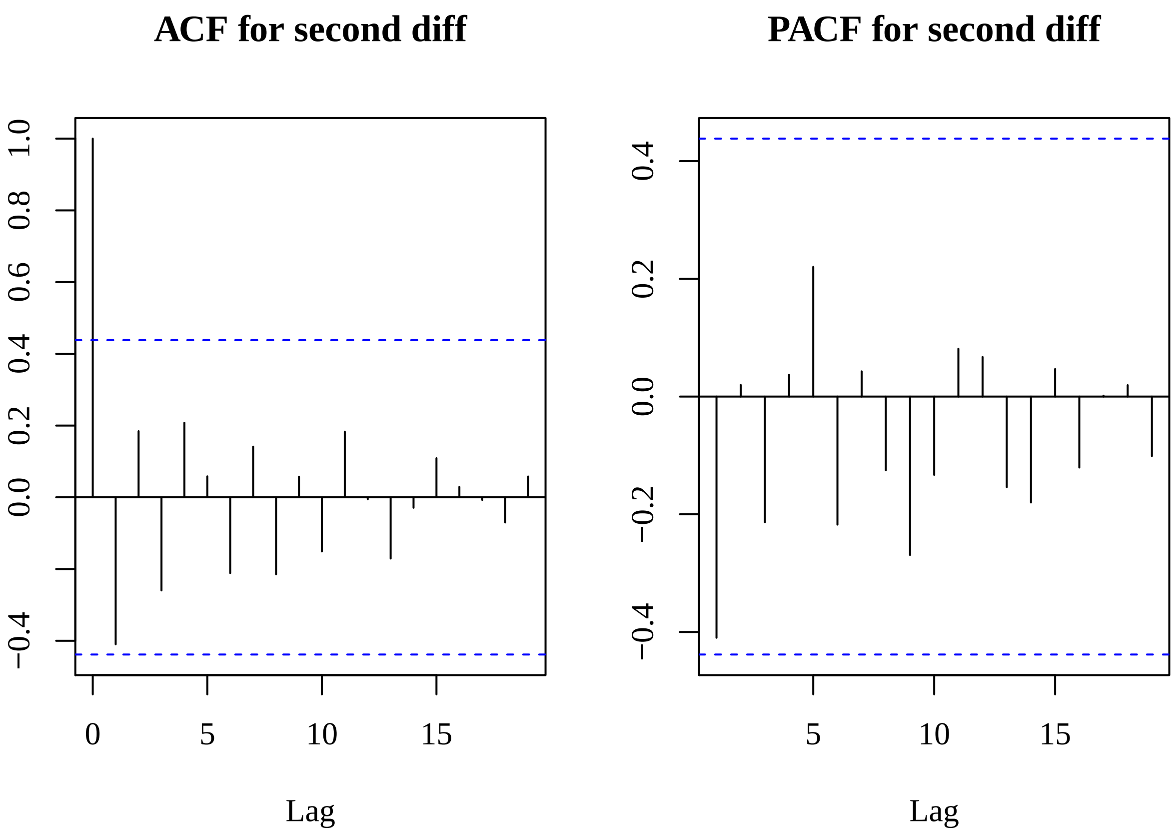
```
##
## Call:
## arima(x = B.15[, 2], order = c(1, 1, 0))
##
## Coefficients:
##          ar1
##       0.6116
## s.e.  0.1624
##
## sigma^2 estimated as 578.3:  log likelihood = -96.81,  aic = 197.63
```

Interpretation:

This is the summary of ARIMA(1, 1, 0) model with its estimates. The AIC value of ARIMA(1, 1, 0) model is found to be 197.63.

Now we will take second difference of the data set we will plot acf and pacf of second differenced data set.

```
par("mar"=c(4,3,4,1))
par(mfrow=c(1,2))
par(family="Times")
second_diff <- diff(diff(B.15[,2],lag=1), lag = 1)
acf(second_diff, lag.max = 20, type="correlation",main="ACF for second diff")
acf(second_diff, lag.max = 20, type="partial",main="PACF for second diff")
```



Interpretation:

The acf plot of second difference shows cut at lag 1 and plot of pacf shows that there is no obvious cut off. However, at lag 1, partial afc is close to cut off and there is likely mixtire of mixture of exponential decay suggesting that time series can be fitted with ARIMA model. Therefore, we will try to fit a ARIMA model with AR(1) process and second difference i.e. fit ARIMA(1, 2, 0) model.

Fitting ARIMA (1,2,0) model

```
B.15.fit.2 <-arima(B.15[,2], order =c(1,2,0))
B.15.fit.2
```

```
##
## Call:
## arima(x = B.15[, 2], order = c(1, 2, 0))
##
## Coefficients:
##           ar1
##       -0.4533
## s.e.   0.2091
##
## sigma^2 estimated as 592.3:  log likelihood = -92.33,  aic = 188.67
```

Interpretation:

This is the summary of ARIMA(1, 2, 0) model with its estimates. The AIC value of ARIMA(1, 2, 0) model is found to be 188.67 which is less than ARIMA(1, 1, 0) model. Comparing both models, ARIMA(1, 2, 0) model makes a better fit to the data as it has lower AIC value. However, we will use auto.arima funtion to double check.

Using auto.arima to double check

```r
library(forecast)
B.15.fit <-auto.arima(B.15[,2])
B.15.fit
```

```
## Series: B.15[, 2]
## ARIMA(1,2,0)
##
## Coefficients:
##           ar1
##       -0.4533
## s.e.   0.2091
##
## sigma^2 estimated as 623.5:  log likelihood=-92.33
## AIC=188.67   AICc=189.37   BIC=190.66
```

Interpretation:

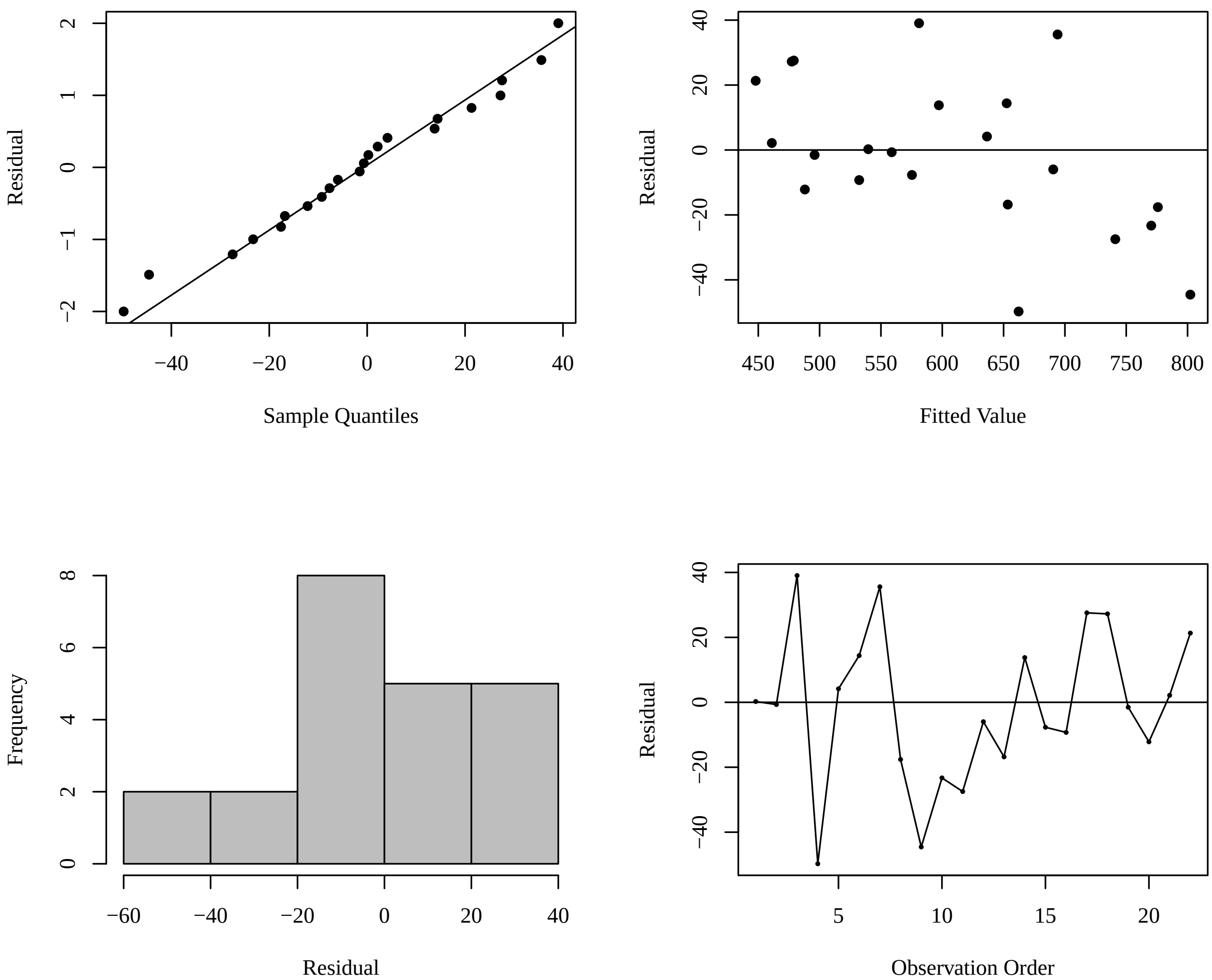
The auto.arima model also shows ARIMA(1, 2, 0) model as best model. So, we will use ARIMA(1, 2, 0) model for 1-10 step ahead forecasting. The equation of fitted ARIMA(1, 2, 0) model is given below.

Fitted model for ARIMA (1, 2, 0) is given by:

$$(1-B)^2(1+0.4533B)y_t = \epsilon_t$$

4 in 1 residual plots

```
par(family="Times")
par(mfrow=c(2,2),oma=c(0,0,0,0))
res.B.15 <-as.vector(residuals(B.15.fit))
fit.B.15 <-as.vector(fitted(B.15.fit))
qqnorm(res.B.15,datax=TRUE,pch=16,xlab='Residual',main='')
qqline(res.B.15,datax=TRUE)
plot(fit.B.15,res.B.15,pch=16, xlab='Fitted Value',
ylab='Residual')
abline(h=0)
hist(res.B.15,col="gray",xlab='Residual',main='')
plot(res.B.15,type="l",xlab='Observation Order', ylab='Residual')
points(res.B.15,pch=16,cex=.5)
abline(h=0)
```



Interpretation:

QQ plot:

Normal probability plot is the plot that is used to observe the individual data and find out whether the data is normally distributed or not. From the above qq norm plot, we can see that most of the data are distributed close to qqnorm line leaving few points. Hence we can say that the residuals are normally distributed.

Histogram of residuals:

Histogram are useful for large sample of data than small ones. Moreover, histogram also seems to have normal distribution, but does not show perfect normal distribution.

Residual vs fitted values:

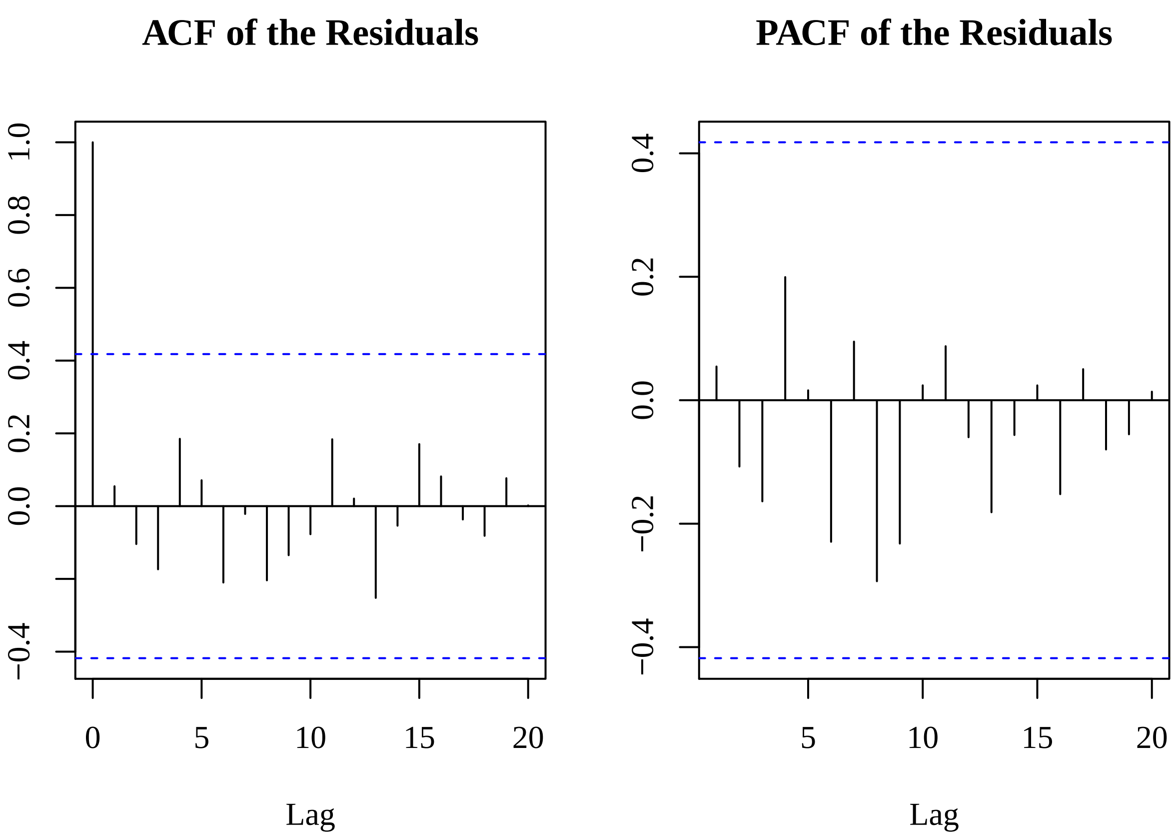
The residuals vs fitted value shows that the residuals are randomly distributed around the line and does no show any pattern. The residuals do not have perfect even distribution but looks good in case of distribution. Moreove, the residuals shows good fit of model.

Residual vs the order of the data:

This plot reveals how the data changes with time. Here, the plot indicates that variability of the observations are changing with time.

ACF and PACF plots of the residuals

```
par("mar"=c(4,3,4,1))
par(family="Times")
par(mfrow=c(1,2),oma=c(0,0,0,0))
ACF <- acf(res.B.15,lag.max=20,type="correlation",main="ACF of the Residuals")
PACF <- acf(res.B.15, lag.max=20,type="partial",main="PACF of the Residuals")
```



Interpretation:

Looking into ACF and PACF plots of residuals, the residual looks stationary. The acf values quickly reduced to zero and is within the significant range. The PACF value also lies within a significant range. These plots shows that the residuals of fitted ARIMA(1, 2, 0) model is stationary.

Calculated ACF and PACF values of the residuals

```r
acf <- as.data.frame(ACF$acf[2:21]) # Excluding 0 lag
pacf <- as.data.frame(PACF$acf[1:20])

K <- 1:20
cf <- as.data.frame(cbind("Lag"=K, "Sample ACF"=acf,"Sample PACF"=pacf))
colnames(cf) <- c("Lag", "Sample ACF", "Sample PACF")
kable(round(cf, 3),
caption = "Table showing Lag, Sample ACF and Sample PACF Values of Residuals", booktabs=T,
linesep = "")%>%
kable_styling(latex_option="hold_position")
```

Table 7: Table showing Lag, Sample ACF and Sample PACF Values of Residuals

Lag	Sample ACF	Sample PACF
1	0.055	0.055
2	-0.104	-0.107
3	-0.174	-0.164
4	0.185	0.199
5	0.071	0.016
6	-0.210	-0.229
7	-0.022	0.095
8	-0.204	-0.293
9	-0.135	-0.232
10	-0.077	0.024
11	0.184	0.087
12	0.021	-0.060
13	-0.252	-0.181
14	-0.054	-0.056
15	0.170	0.024
16	0.082	-0.152
17	-0.037	0.050
18	-0.082	-0.080
19	0.077	-0.055
20	0.002	0.014

Interpretation:

The table shows calculated ACF and PACF values of residuals of ARIMA(1, 2, 0) model using lag.max value of 20.

General forecasting model

$$\hat{y}_{T+\tau}(T) = E[y_{T+\tau}|y_T, y_{T-1}, ...] = \mu + \sum_{i=\tau}^{\infty} \Psi_i \epsilon_{T+\tau-i}$$

1-10 step ahead forecasts

```
library(forecast)
B.15.fit <-auto.arima(B.15[,2])
B.15.forecast<-as.array(forecast(B.15.fit,h=10))
kable(B.15.forecast,
caption = "Table showing point forecast, 80 percent and  95 percent prediction intervals",
booktabs=T,
linesep = "")%>%
kable_styling(latex_option="hold_position")
```

Table 8: Table showing point forecast, 80 percent and 95 percent prediction intervals

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
23	466.7685	434.76854	498.7685	417.82878	515.7082
24	468.1591	409.22121	527.0969	378.02140	558.2967
25	467.8171	373.55509	562.0790	323.65581	611.9783
26	468.2604	334.67715	601.8437	263.96243	672.5584
27	468.3478	290.70970	645.9859	196.67377	740.0218
28	468.5965	243.08759	694.1055	123.71037	813.4827
29	468.7721	191.67339	745.8709	44.98618	892.5581
30	468.9809	136.90442	801.0573	-38.88624	976.8480
31	469.1746	78.89640	859.4528	-127.70440	1066.0536
32	469.3751	17.85539	920.8949	-221.16472	1159.9150

Interpretation:

These are the 1-10 step ahead forecast values of violent crime data. The table above shows point forecast, 80 percent and 95 percent prediction intervals.

Prediction interval formula

The general formula for prediction interval is given below.

The 100(1-α) PI for $y_{T+\tau}$ is

$$\hat{y}_{T+\tau}(T) \pm Z_{\alpha/2} * \sqrt{Var[e_T(\tau)]}$$