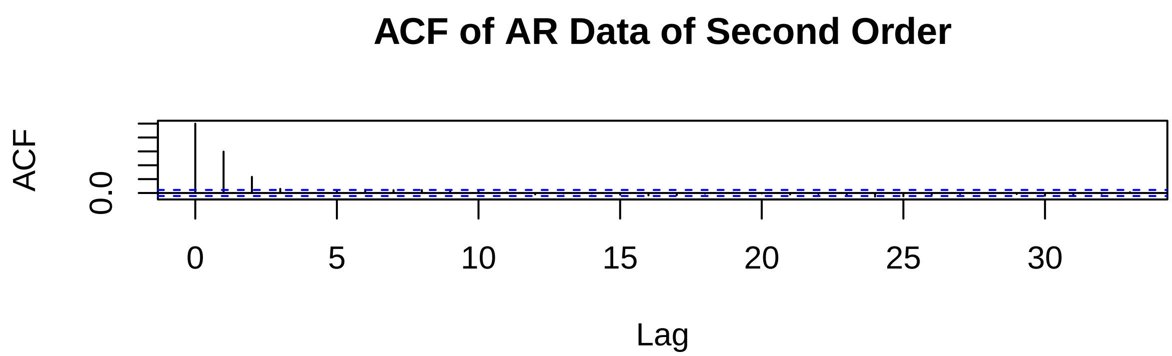
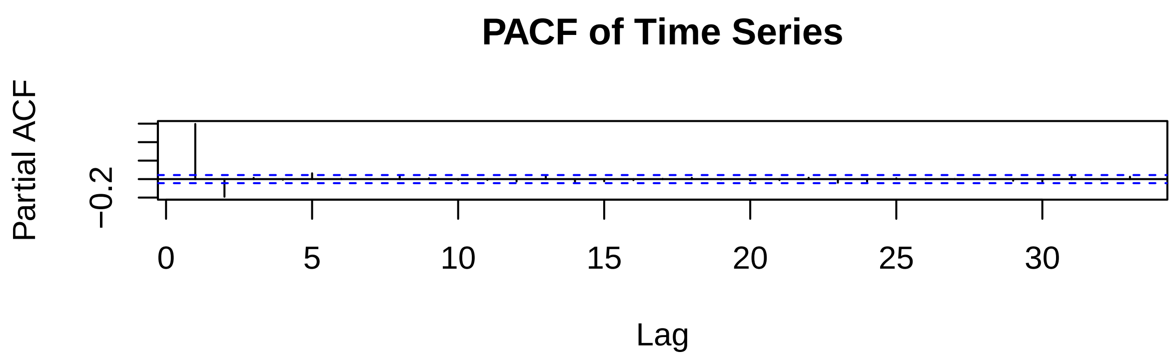
week-5.R

Ahmed

2023-04-16

```
set.seed(43)
data=arima.sim(list(order=c(2,0,0),ar=c(0.7,-0.2)),n=2000)
par(mfrow=c(2,1))
acf(data,main="ACF of AR Data of Second Order")
acf(data,type = "partial",main="PACF of Time Series")
```






```
arima(data,order=c(2,0,0),include.mean = FALSE)
```

```
##
## Call:
## arima(x = data, order = c(2, 0, 0), include.mean = FALSE)
##
## Coefficients:
```

```
##          ar1      ar2
##       0.7111  -0.1912
## s.e.  0.0219   0.0220
##
## sigma^2 estimated as 0.9985:  log likelihood = -2836.64,  aic = 5679.27
```

```
m=arima(data,order=c(2,0,0),include.mean = FALSE)
SSE=sum(resid(m)^2)
SSE
```

```
## [1] 1997.007
```

```
rm(list=ls(all=TRUE))
set.seed(500) #500  Seven kingdoms Kent, Essex, Sussex, Wessex, East Anglia, Mercia, and Northumbria.)
data = arima.sim( list(order = c(3,0,0), ar =c( 0.6, -0.1, .4)), n = 5000)

arima(data, order=c(2,0,0), include.mean=FALSE )
```

```
##
## Call:
## arima(x = data, order = c(2, 0, 0), include.mean = FALSE)
##
## Coefficients:
##          ar1     ar2
##       0.6836  0.1586
## s.e.  0.0140  0.0140
##
## sigma^2 estimated as 1.163:  log likelihood = -7473.05,  aic = 14952.1
```

```
arima(data, order=c(3,0,0), include.mean=FALSE )
```

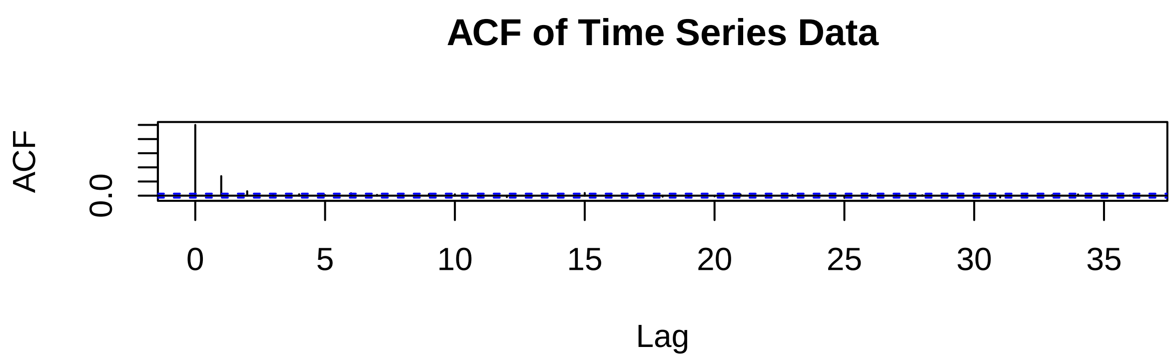
```
##
## Call:
## arima(x = data, order = c(3, 0, 0), include.mean = FALSE)
##
## Coefficients:
##          ar1      ar2     ar3
##       0.6195  -0.1164  0.4024
## s.e.  0.0129   0.0156  0.0130
##
## sigma^2 estimated as 0.975:  log likelihood = -7032.1,  aic = 14072.2
```

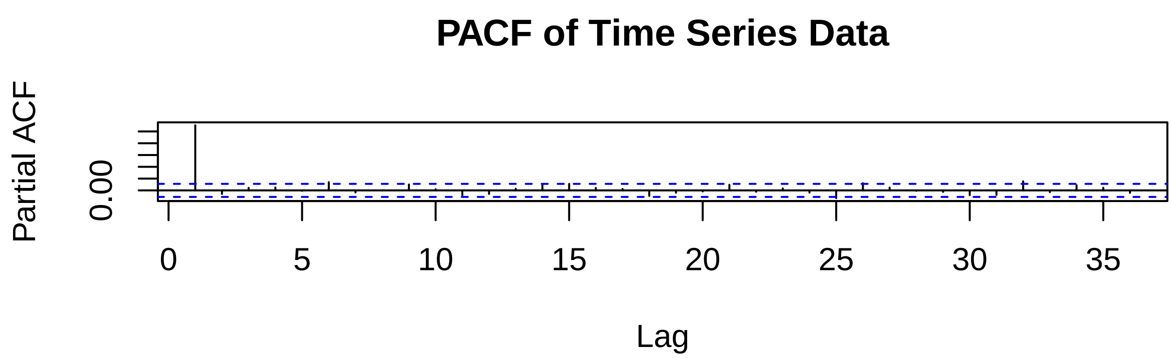
```
arima(data, order=c(4,0,0), include.mean=FALSE )
```

```
##
## Call:
## arima(x = data, order = c(4, 0, 0), include.mean = FALSE)
##
## Coefficients:
##          ar1      ar2     ar3     ar4
##       0.6188  -0.1162  0.4015  0.0015
```

```
## s.e.   0.0141   0.0156  0.0156  0.0142
##
## sigma^2 estimated as 0.9749:  log likelihood = -7032.09,  aic = 14074.19
```

```
rm(list=ls(all=TRUE))
set.seed(597) # Saint Augustine arrives in England
data = arima.sim( list(order = c(1,0,0), ar = .3), n = 5000)

par(mfrow=c(2,1))
acf(data, main="ACF of Time Series Data")
acf(data, type="partial", main="PACF of Time Series Data")
```



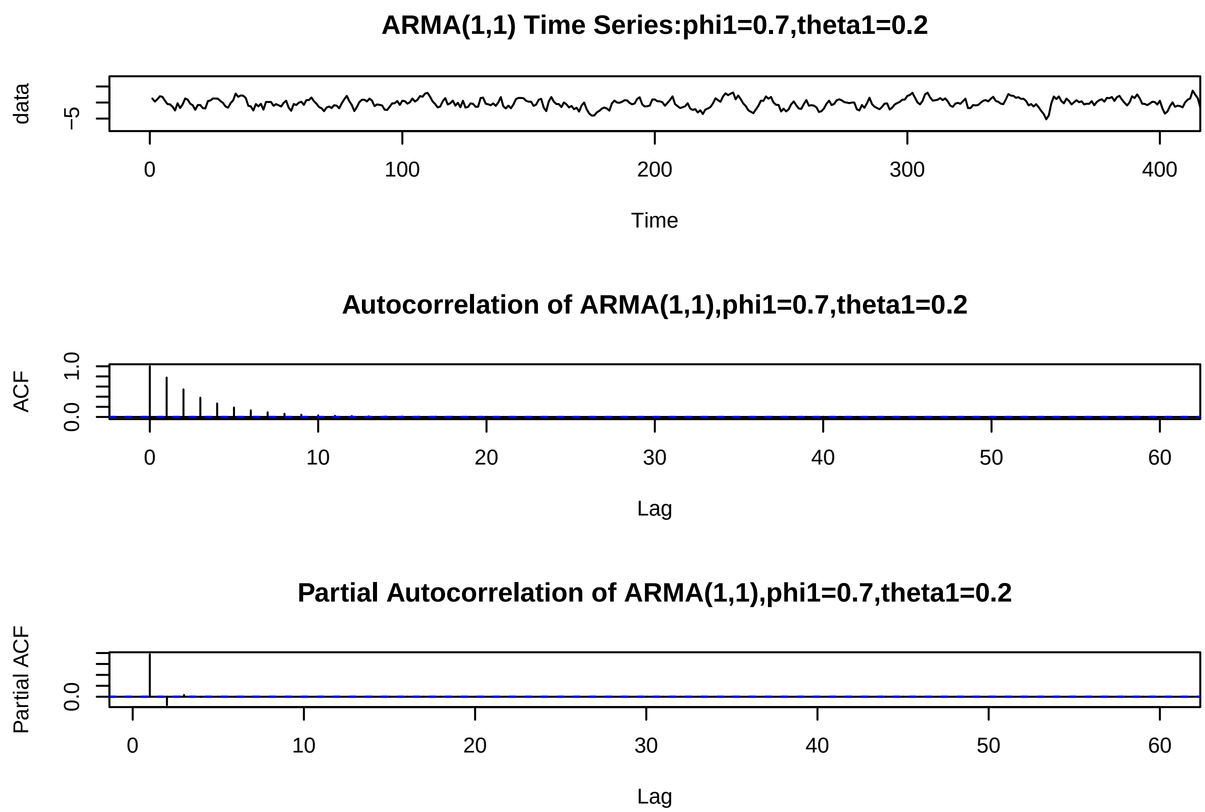



```
rm(list=ls(all=TRUE))
set.seed(597) # Saint Augustine arrives in England
data = arima.sim( list(order = c(1,0,0), ar = .3), n = 5000);
arima(data, order=c(1,0,0) );
```

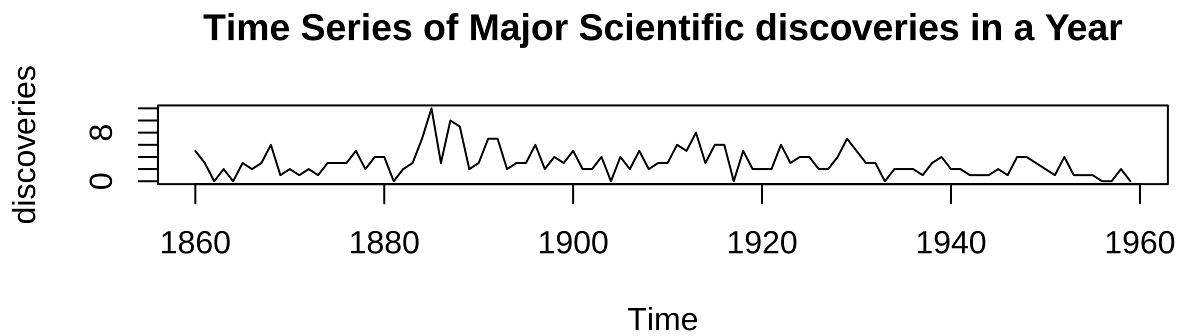
```
##
## Call:
## arima(x = data, order = c(1, 0, 0))
##
## Coefficients:
##          ar1  intercept
##       0.2762    -0.0084
## s.e.  0.0136     0.0197
```

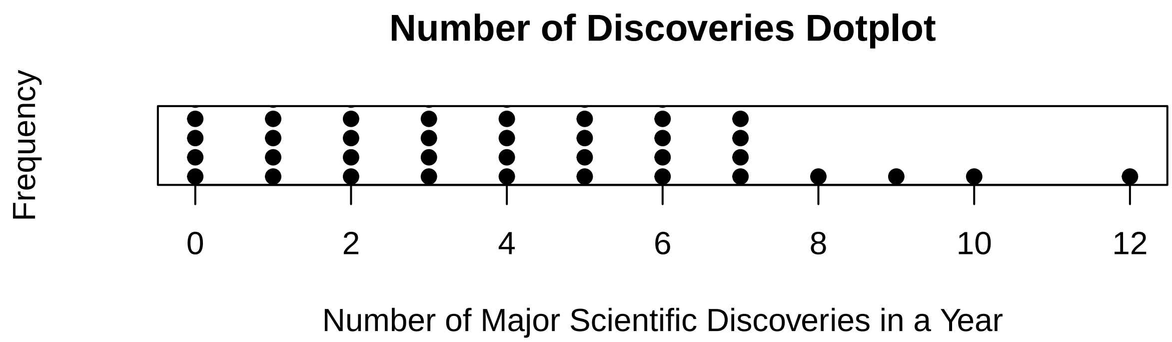
```
##
## sigma^2 estimated as 1.016:  log likelihood = -7134.7,  aic = 14275.4
```

```
#ARMA simulation
set.seed(500)
data=arima.sim(list(order=c(1,0,1),ar=0.7,ma=0.2),n=1000000)
par(mfcol=c(3,1))
plot(data,main="ARMA(1,1) Time Series:phi1=0.7,theta1=0.2",xlim=c(0,400))
acf(data,main="Autocorrelation of ARMA(1,1),phi1=0.7,theta1=0.2")
acf(data,type="partial",main="Partial Autocorrelation of ARMA(1,1),phi1=0.7,theta1=0.2")
```




```
par(mfcol=c(2,1))
plot(discoveries,main="Time Series of Major Scientific discoveries in a Year")
#for discrete data
stripchart(discoveries,method = "stack",offset = 0.5,at=0.15,pch=19,main="Number of Discoveries Dotplot
```



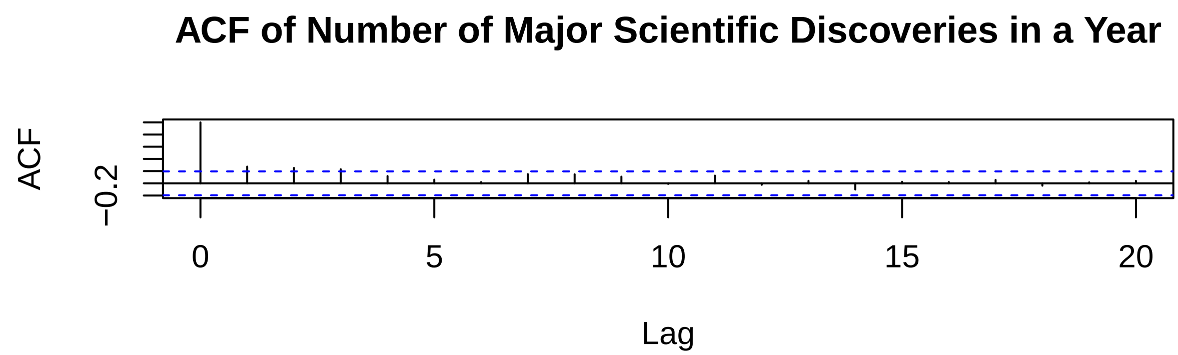


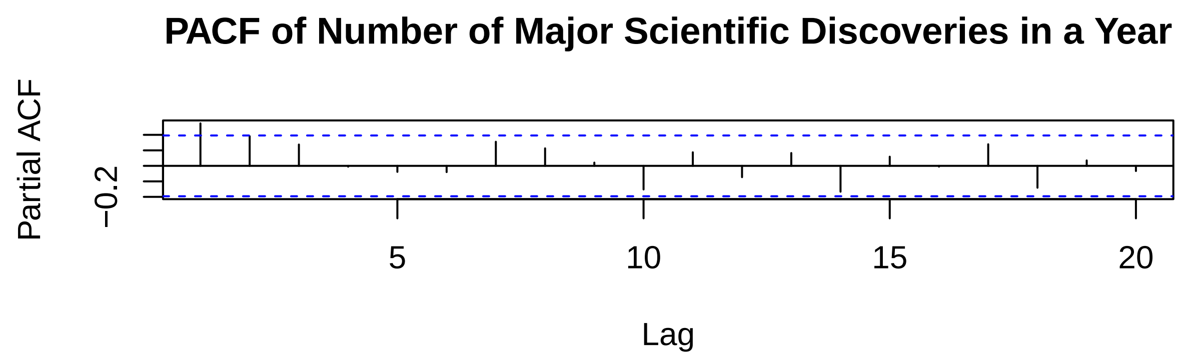

```
acf(discoveries,main="ACF of Number of Major Scientific Discoveries in a Year")
acf(discoveries,type = "partial",main="PACF of Number of Major Scientific Discoveries in a Year")
p=c(0:3)
q=c(0:3)
result_matrix <- matrix(0, nrow = length(p), ncol = length(q))
for (i in 0:length(p)) {
  for (j in 0:length(q)) {
    result_matrix[i, j] <- arima(discoveries, order=c(i,0,j))$aic

  }
}
result_matrix
```

```
##          [,1]     [,2]     [,3]     [,4]
## [1,] 440.1980 442.0428 442.6747 444.4217
## [2,] 442.0722 443.7021 441.6594 442.4198
## [3,] 443.5655 439.9263 441.2941 443.4179
## [4,] 445.4075 441.7127 439.8615 439.2558
```

```
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```






```r
auto.arima(discoveries,d=0,approximation = FALSE)
```

```
## Series: discoveries
## ARIMA(1,0,1) with non-zero mean
##
## Coefficients:
##          ar1      ma1    mean
##       0.8353  -0.6243  3.0208
## s.e.  0.1379   0.1948  0.4728
##
## sigma^2 = 4.538:  log likelihood = -216.1
## AIC=440.2   AICc=440.62   BIC=450.62
```

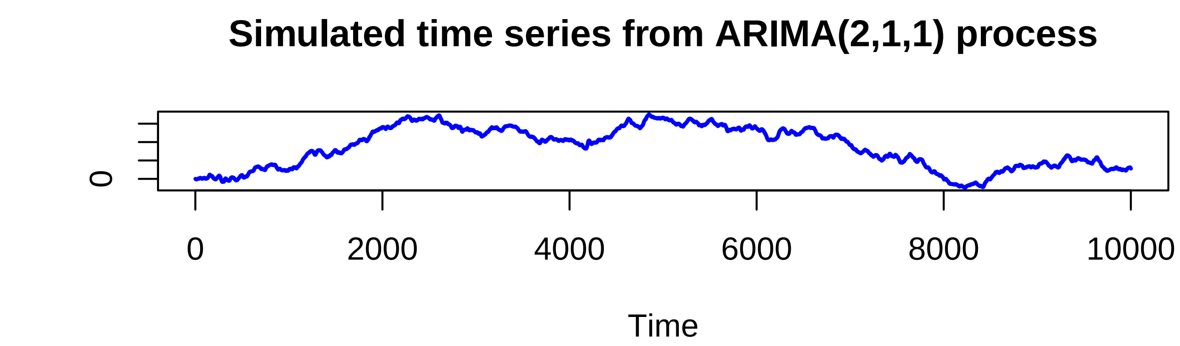
```r
auto.arima(discoveries,d=0,ic="bic",approximation = FALSE)
```

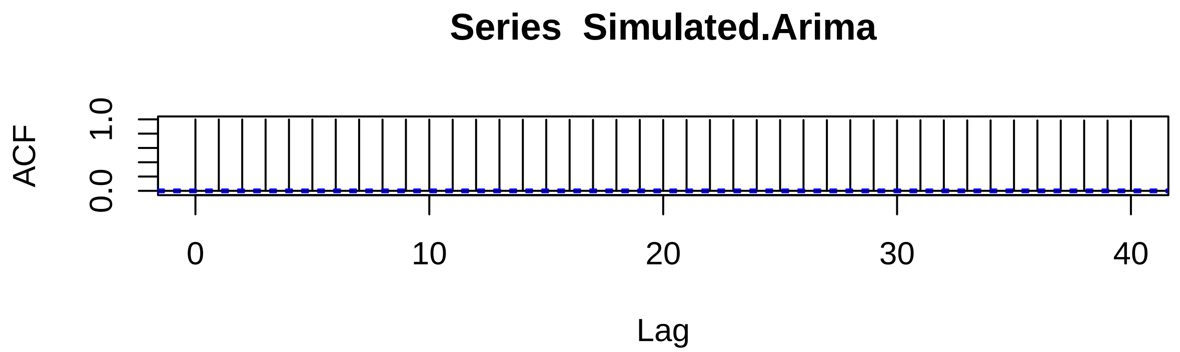
```
## Series: discoveries
## ARIMA(1,0,1) with non-zero mean
##
## Coefficients:
##          ar1      ma1    mean
##       0.8353  -0.6243  3.0208
## s.e.  0.1379   0.1948  0.4728
##
## sigma^2 = 4.538:  log likelihood = -216.1
## AIC=440.2   AICc=440.62   BIC=450.62
```

```r
data = arima.sim( n=1E4, list(ar=.5, ma=.2) )
auto.arima(data)
```

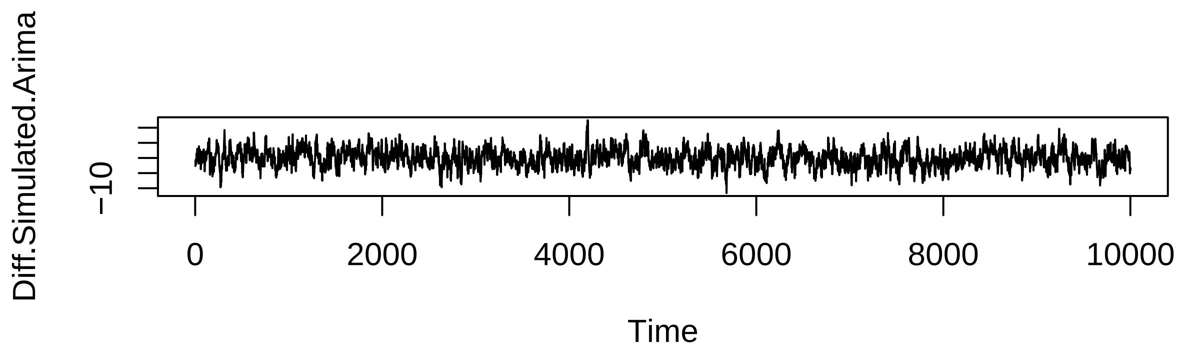
```
## Series: data
## ARIMA(1,0,1) with zero mean
##
## Coefficients:
##          ar1     ma1
##       0.5161  0.1909
## s.e.  0.0134  0.0154
##
## sigma^2 = 0.9889:  log likelihood = -14132.62
## AIC=28271.25   AICc=28271.25   BIC=28292.88
```

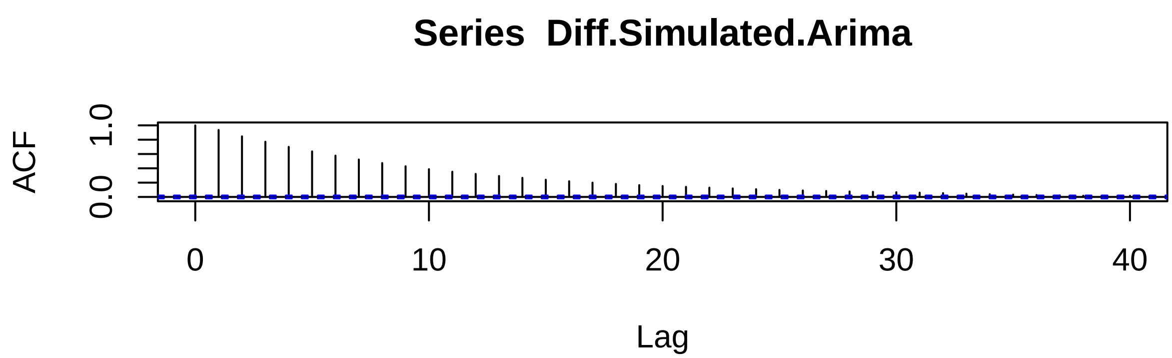
```r
#Arima(2,1,1) simulation
# parameters
phi=c(.7, .2)
beta=0.5
sigma=3
m=10000

set.seed(5)
Simulated.Arima=arima.sim(n=m,list(order = c(2,1,1), ar = phi, ma=beta))
plot(Simulated.Arima, ylab=' ',main='Simulated time series from ARIMA(2,1,1) process', col='blue', lwd=
acf(Simulated.Arima)
```





```
Diff.Simulated.Arima=diff(Simulated.Arima)
plot(Diff.Simulated.Arima)
acf(Diff.Simulated.Arima)
```

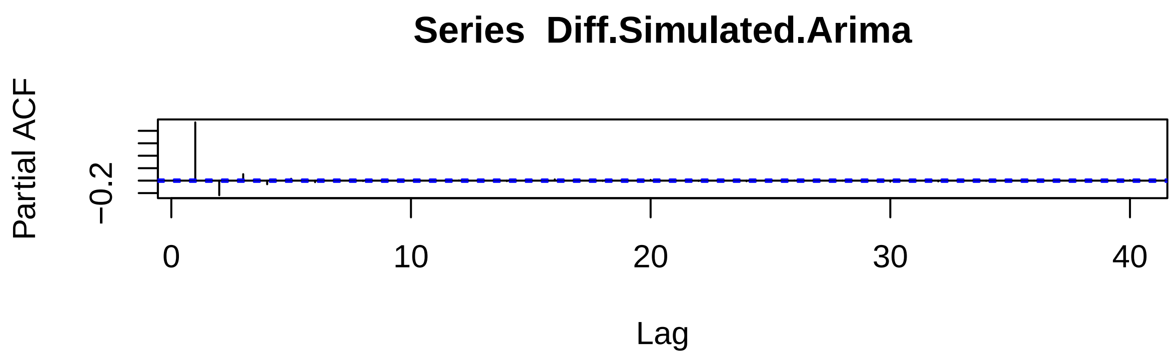





```
pacf(Diff.Simulated.Arima)
library(astsa)
```

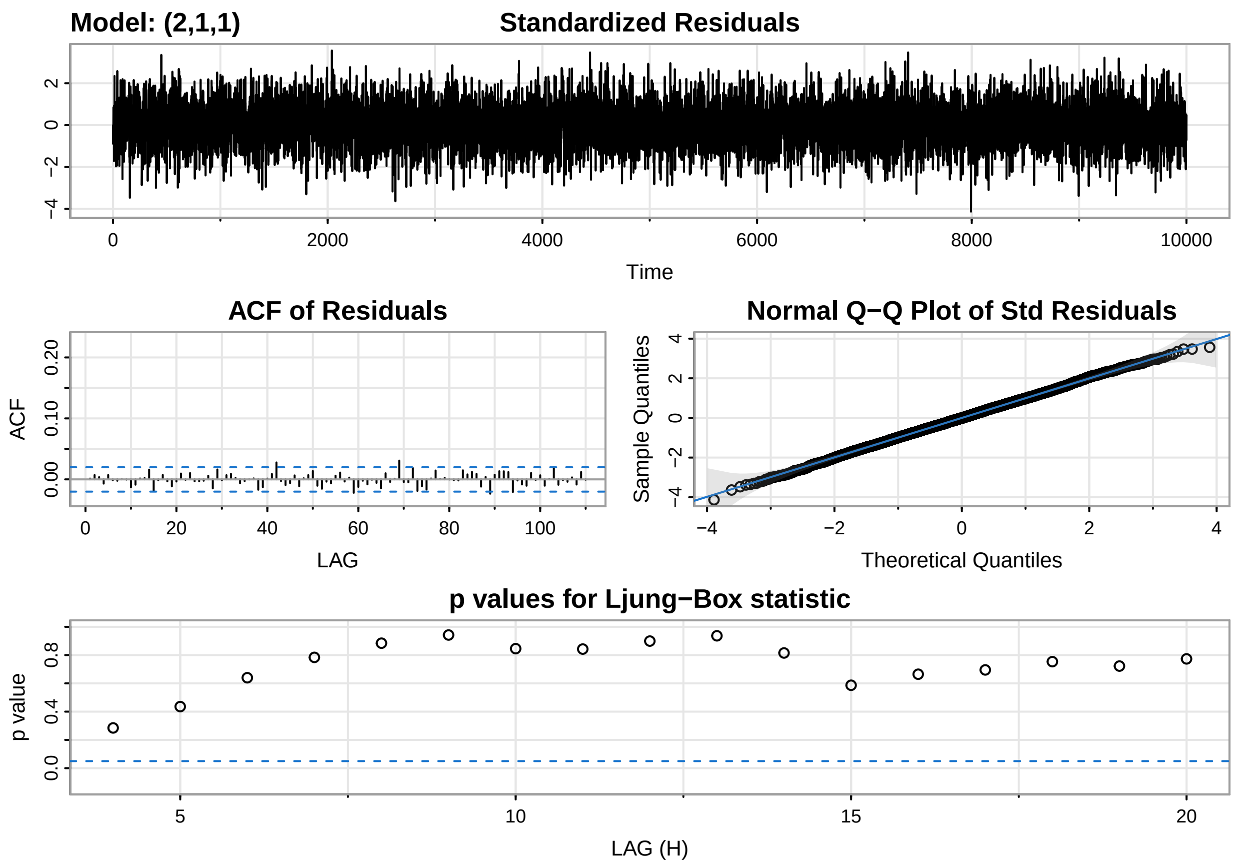
```
##
## Attaching package: 'astsa'

## The following object is masked from 'package:forecast':
##
##     gas
```

```
sarima(Simulated.Arima,2,1,1,0,0,0)
```




```
## initial  value 1.092704
## iter   2 value 0.655083
## iter   3 value 0.576329
## iter   4 value 0.250793
## iter   5 value 0.124855
## iter   6 value 0.033738
## iter   7 value 0.013225
## iter   8 value 0.012554
## iter   9 value 0.012517
## iter  10 value 0.012292
## iter  11 value 0.012267
## iter  12 value 0.012258
## iter  13 value 0.012170
## iter  14 value 0.012069
## iter  15 value 0.011860
## iter  16 value 0.011703
## iter  17 value 0.011609
## iter  18 value 0.011601
## iter  19 value 0.011601
## iter  20 value 0.011601
## iter  20 value 0.011601
## iter  20 value 0.011601
## final  value 0.011601
## converged
## initial  value 0.011653
## iter   2 value 0.011653
```

```
## iter   3 value 0.011653
## iter   3 value 0.011653
## iter   3 value 0.011653
## final  value 0.011653
## converged
```




```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     xreg = constant, transform.pars = trans, fixed = fixed, optim.control = list(trace = trc,
##         REPORT = 1, reltol = tol))
##
## Coefficients:
##          ar1    ar2     ma1  constant
##       0.6876  0.204  0.5002    0.0280
## s.e.  0.0334  0.032  0.0301    0.1398
##
## sigma^2 estimated as 1.023:  log likelihood = -14305.92,  aic = 28621.83
##
## $degrees_of_freedom
## [1] 9996
##
## $ttable
##          Estimate     SE t.value p.value
```

```
## ar1        0.6876 0.0334 20.5786  0.0000
## ar2        0.2040 0.0320  6.3817  0.0000
## ma1        0.5002 0.0301 16.6139  0.0000
## constant   0.0280 0.1398  0.2001  0.8414
##
## $AIC
## [1] 2.862183
##
## $AICc
## [1] 2.862183
##
## $BIC
## [1] 2.865788
```

```
library(forecast)
auto.arima(Simulated.Arima)
```

```
## Series: Simulated.Arima
## ARIMA(4,2,0)
##
## Coefficients:
##          ar1      ar2     ar3      ar4
##       0.2279  -0.1633  0.0337  -0.0707
## s.e.  0.0100   0.0102  0.0102   0.0100
##
## sigma^2 = 1.064:  log likelihood = -14495.62
## AIC=29001.24   AICc=29001.25   BIC=29037.29
```

```
fit1<-arima(Diff.Simulated.Arima, order=c(4,0,0))
fit1
```

```
##
## Call:
## arima(x = Diff.Simulated.Arima, order = c(4, 0, 0))
##
## Coefficients:
##          ar1      ar2     ar3      ar4  intercept
##       1.1862  -0.3761  0.1733  -0.0581     0.0280
## s.e.  0.0100   0.0154  0.0154   0.0100     0.1353
##
## sigma^2 estimated as 1.025:  log likelihood = -14313.1,  aic = 28638.2
```

```
fit2<-arima(Diff.Simulated.Arima, order=c(2,0,1))
fit2
```

```
##
## Call:
## arima(x = Diff.Simulated.Arima, order = c(2, 0, 1))
##
## Coefficients:
##          ar1    ar2     ma1  intercept
##       0.6876  0.204  0.5002     0.0280
```

```
## s.e.  0.0334  0.032  0.0301     0.1398
##
## sigma^2 estimated as 1.023:  log likelihood = -14305.92,  aic = 28621.83
```

```
fit3<-arima(Simulated.Arima, order=c(2,1,1))
fit3
```

```
##
## Call:
## arima(x = Simulated.Arima, order = c(2, 1, 1))
##
## Coefficients:
##          ar1     ar2     ma1
##       0.6876  0.2039  0.5001
## s.e.  0.0334  0.0320  0.0301
##
## sigma^2 estimated as 1.023:  log likelihood = -14305.93,  aic = 28619.85
```

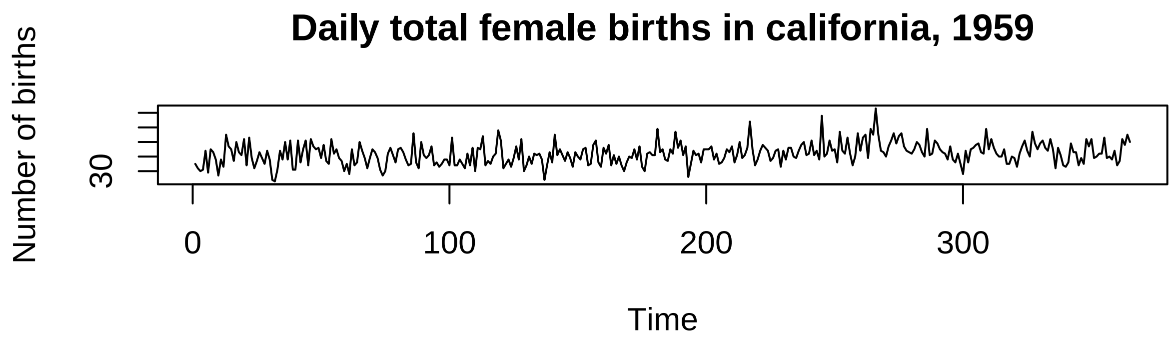
```
Box.test(Diff.Simulated.Arima)
```

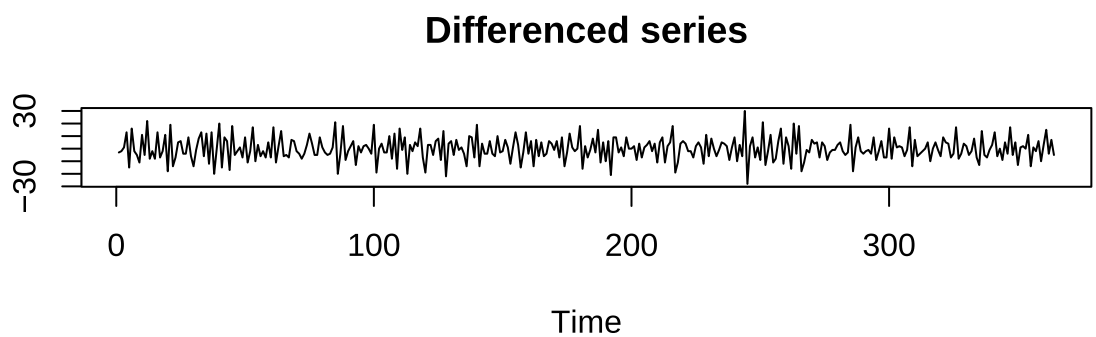
```
##
##  Box-Pierce test
##
## data:  Diff.Simulated.Arima
## X-squared = 8761.2, df = 1, p-value < 2.2e-16
```

```
#Daily female birth

library(astsa)

# read data to R variable
birth.data<-read.csv("daily-total-female-births-in-cal.csv")

# pull out number of births column
number_of_births<-birth.data$Daily.total.female.births.in.California..1959

# use date format for dates
birth.data$Date <- as.Date(birth.data$Date, "%m/%d/%Y")

plot.ts(number_of_births,main='Daily total female births in california, 1959', ylab = 'Number of births

# Test for correlation
Box.test(number_of_births, lag = log(length(number_of_births)))
```

```
##
##  Box-Pierce test
##
## data:  number_of_births
## X-squared = 36.391, df = 5.8999, p-value = 2.088e-06
```

```r
# Plot the differenced data
plot.ts(diff(number_of_births), main='Differenced series', ylab = '')
```

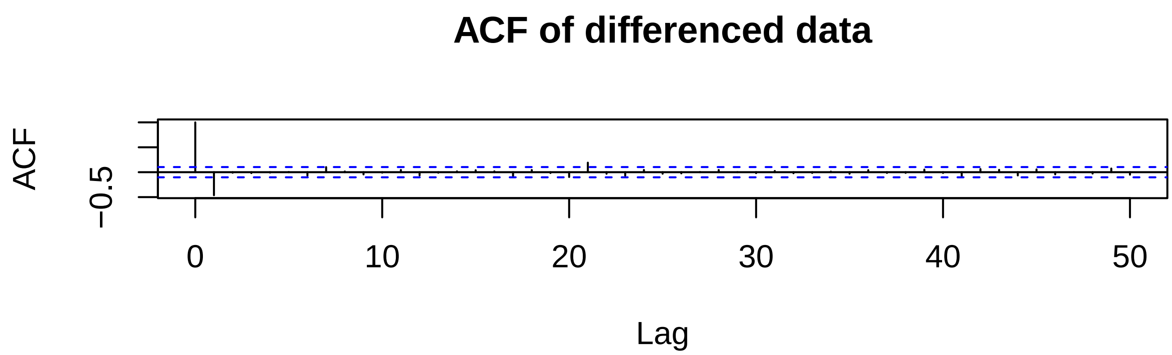


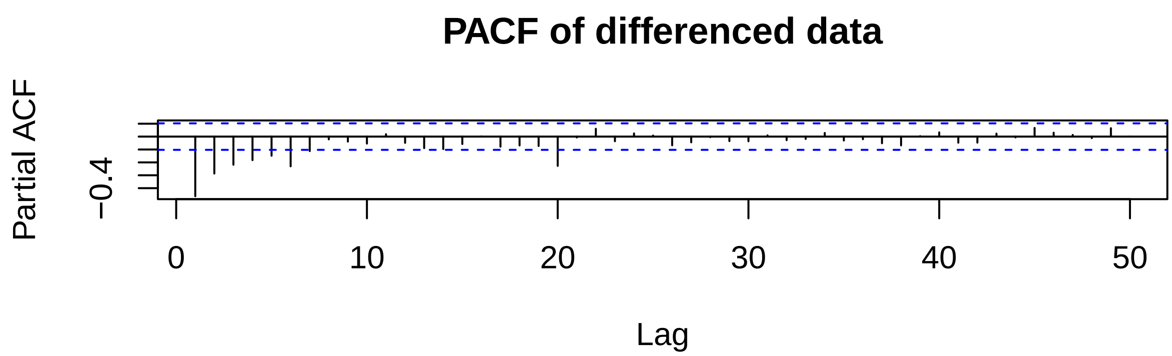



```r
# Test for correlation in the differenced data
Box.test(diff(number_of_births), lag = log(length(diff(number_of_births))))
```

```
##
##  Box-Pierce test
##
## data:  diff(number_of_births)
## X-squared = 78.094, df = 5.8972, p-value = 7.661e-15
```

```r
# acf and pacf of the differenced data

acf(diff(number_of_births), main='ACF of differenced data', 50)
pacf(diff(number_of_births), main='PACF of differenced data', 50)
```





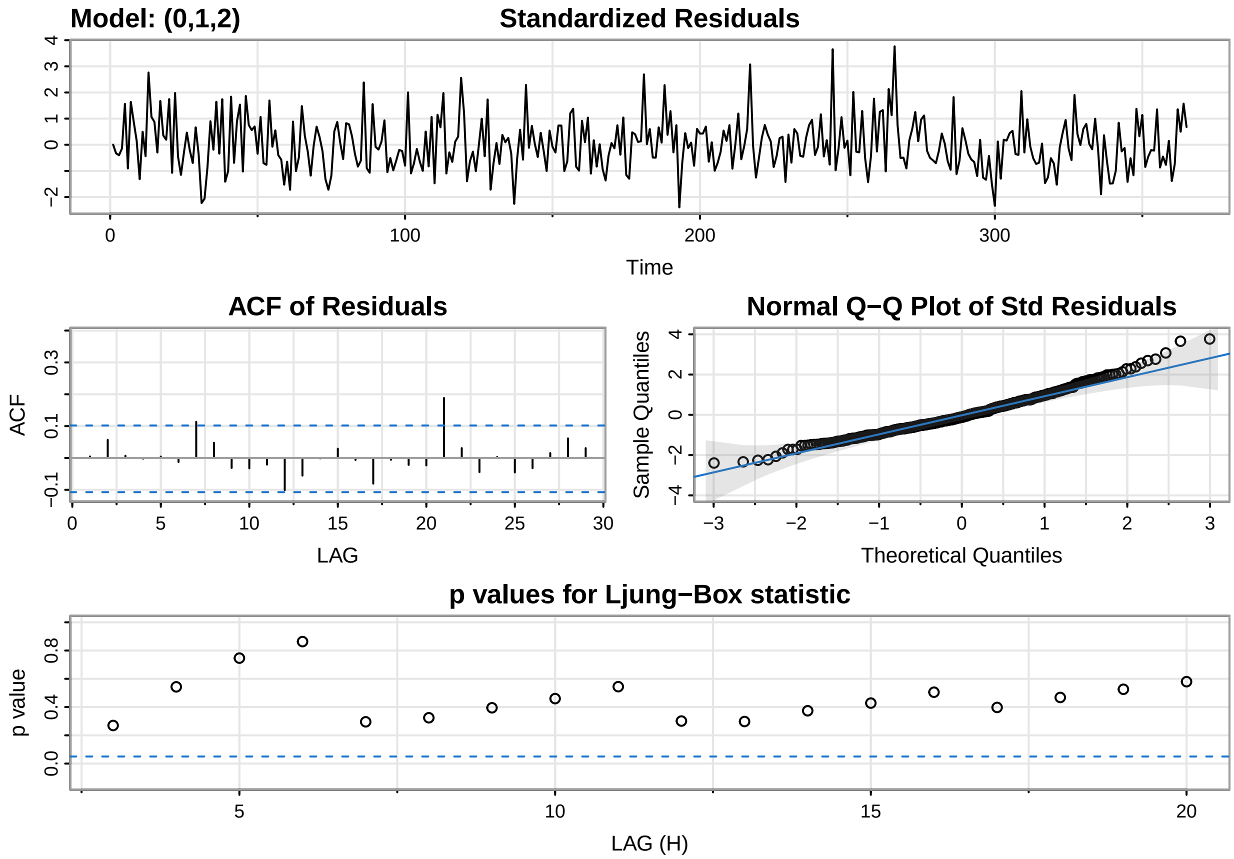

```
# Fit various ARIMA models


model1<-arima(number_of_births, order=c(0,1,1))
SSE1<-sum(model1$residuals^2)
model1.test<-Box.test(model1$residuals, lag = log(length(model1$residuals)))

model2<-arima(number_of_births, order=c(0,1,2))

SSE2<-sum(model2$residuals^2)

model2.test<-Box.test(model2$residuals, lag = log(length(model2$residuals)))


model3<-arima(number_of_births, order=c(7,1,1))
SSE3<-sum(model3$residuals^2)
model3.test<-Box.test(model3$residuals, lag = log(length(model3$residuals)))

model4<-arima(number_of_births, order=c(7,1,2))
SSE4<-sum(model4$residuals^2)
model4.test<-Box.test(model4$residuals, lag = log(length(model4$residuals)))

df<-data.frame(row.names=c('AIC', 'SSE', 'p-value'), c(model1$aic, SSE1, model1.test$p.value),
               c(model2$aic, SSE2, model2.test$p.value), c(model3$aic, SSE3, model3.test$p.value),
               c(model4$aic, SSE4, model4.test$p.value))
colnames(df)<-c('Arima(0,1,1)','Arima(0,1,2)', 'Arima(7,1,1)', 'Arima(7,1,2)')
```

```r
format(df, scientific=FALSE)
```

```
##          Arima(0,1,1)  Arima(0,1,2)  Arima(7,1,1)  Arima(7,1,2)
## AIC      2462.2207021  2459.5705306  2464.8827225  2466.6664136
## SSE     18148.4561632 17914.6513437 17584.3902548 17574.0578118
## p-value     0.5333604     0.9859227     0.9999899     0.9999929
```

```r
# Fit a SARIMA model

sarima(number_of_births, 0,1,2,0,0,0)
```

```
## initial  value 2.216721
## iter   2 value 2.047518
## iter   3 value 1.974780
## iter   4 value 1.966955
## iter   5 value 1.958906
## iter   6 value 1.952299
## iter   7 value 1.951439
## iter   8 value 1.950801
## iter   9 value 1.950797
## iter  10 value 1.950650
## iter  11 value 1.950646
## iter  12 value 1.950638
## iter  13 value 1.950635
## iter  13 value 1.950635
## iter  13 value 1.950635
## final  value 1.950635
## converged
## initial  value 1.950708
## iter   2 value 1.950564
## iter   3 value 1.950290
## iter   4 value 1.950196
## iter   5 value 1.950185
## iter   6 value 1.950185
## iter   7 value 1.950185
## iter   7 value 1.950185
## iter   7 value 1.950185
## final  value 1.950185
## converged
```




```
## $fit
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     xreg = constant, transform.pars = trans, fixed = fixed, optim.control = list(trace = trc,
##         REPORT = 1, reltol = tol))
##
## Coefficients:
##           ma1      ma2  constant
##       -0.8511  -0.1113     0.015
## s.e.   0.0496   0.0502     0.015
##
## sigma^2 estimated as 49.08:  log likelihood = -1226.36,  aic = 2460.72
##
## $degrees_of_freedom
## [1] 361
##
## $ttable
##          Estimate     SE  t.value p.value
## ma1       -0.8511 0.0496 -17.1448  0.0000
## ma2       -0.1113 0.0502  -2.2164  0.0273
## constant   0.0150 0.0150   1.0007  0.3176
##
## $AIC
## [1] 6.760225
##
```

```
## $AICc
## [1] 6.760408
##
## $BIC
## [1] 6.803051
```