Madelyn Caufield

1a. $E[r_t | F_{t-1}]$

$= E[p_t - p_{t-1}]$ by $r_t = p_t - p_{t-1}$

$= E[p_t | F_{t-1}] - E[p_{t-1} | F_{t-1}]$

$= p_{t-1} - p_{t-1}$, by $E[p_t | F_{t-1}] = p_{t-1}$ (MG) and $E[p_{t-1} | F_{t-1}] = p_{t-1}$

$= 0$

1b. i CC return is covariance stationary (mean, variance, covariance don't depend on time)

$E(x) = 0$ and $E(xt) = tE(x) = 0$

$\gamma - t \overset{iid}{\sim} N(0, \sigma^2)$

ii for any $t \neq s$ (say $t > s$)

$E[\varepsilon_t \varepsilon_s]$

$= E[E[\varepsilon_t \varepsilon_s | F_{t-1}]]$

$= E[\varepsilon_s E[\varepsilon_t | F_{t-1}]]$

$= E[\varepsilon_s \cdot 0]$

$= E[0]$

$= 0$

1c. i For an iid random process $\{X_k\}, k = 1, 2, \ldots$

$$\frac{1}{n} \sum_{k=1}^{n} X_k \longrightarrow EX_1$$

with probability 1.

A strong stationary process $\{X_t\}$ is one whose joint distributions for any set of times

$t_1, \ldots, t_k$ that $F_X(t_1, \ldots, t_k) = F_X(t_1 + \tau, \ldots, + \tau)$

An iid process always satisfies this since its joint distrib. at any set of times is the same.

ii Since $E[r_i^2] < \infty$, r_t is covariance stationary. If ε_t was infinite (student t with 1 degree of freedom), then it would not be covariance stationary.

2a. $E[Y_t] = c + \phi E[Y_{t-1}] + E[\varepsilon_t]$

$$= c + \phi E[Y_t]$$

$$E[Y_t] = \frac{c}{1-\phi} = \mu$$

2b. Since AR(1) has $|\phi| < 1$, it is covariance stationary. Therefore, $\text{Cov}(Y_{t-j} - \mu, \varepsilon_t) = \phi$ for any $j \geq 1$

2c. $\text{var}(Y_t) = E[(Y_t - \mu)^2] = E[(\phi(Y_{t-1} - \mu) + \varepsilon_t)^2]$

$$= \phi^2 E[(Y_{t-1} - \mu)^2] + 2E[(Y_{t-1} - \mu)\varepsilon_t] + E[\varepsilon_t^2]$$

$$= \phi \gamma_0 + \sigma^2$$

$$\gamma_0 = \sigma^2(1-\phi^2)^{-1}$$

2d. $\gamma_j = E[(Y_t - \mu)(Y_{t-j} - \mu)]$

$$= E[\phi(Y_{t-1})(Y_{t-j} - \mu)] + E[\varepsilon_t(Y_{t-j} - \mu)]$$

$$= \phi \gamma_{j-1}$$

2e. $\gamma_j = \phi^j \gamma_0 = \phi^j \sigma^2 (1-\phi^2)^{-1}$

2f. $\rho_j = \text{corr}(Y_t, Y_{t-j}) = \frac{\gamma_j}{\gamma_0} = \frac{\phi^j \sigma_Y^2}{\sigma^2 {}_Y} = \phi^j$

3a. Yes since $Y_{t-1} < 1$

3b. $E(Y_t) = E(5 - 0.55Y_{t-1} + \varepsilon_t)$

$= E(5) - E(0.55Y_{t-1}) + E(\varepsilon_t)$

$= 5 - 0.55(Y_{t-1}) + 0$

$\mu = \dfrac{5}{1+0.55} = 3.2258$

3c. $var(Y_t) = \dfrac{(1.2)^2}{1-(-0.55)^2} = \dfrac{1.44}{1-0.3025} = 2.064$

3d. $CF = 2.064(-0.55)^n$

ECON 432 Homework 5

Madelyn Caufield

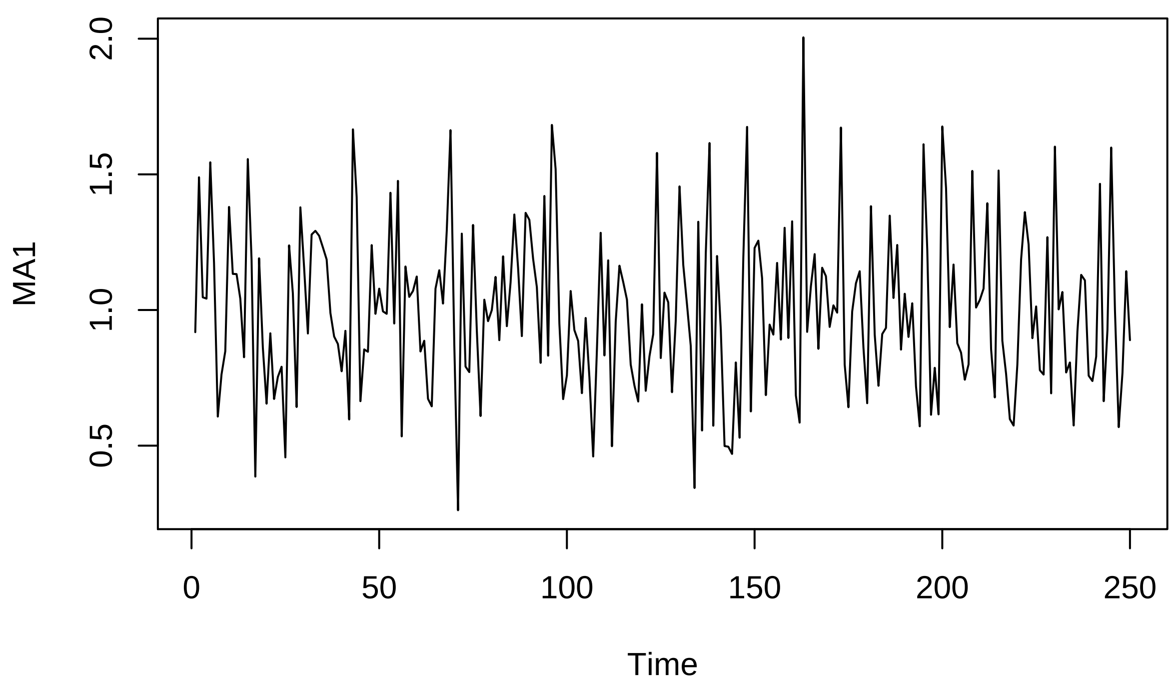
Mar 10,2021

Contents

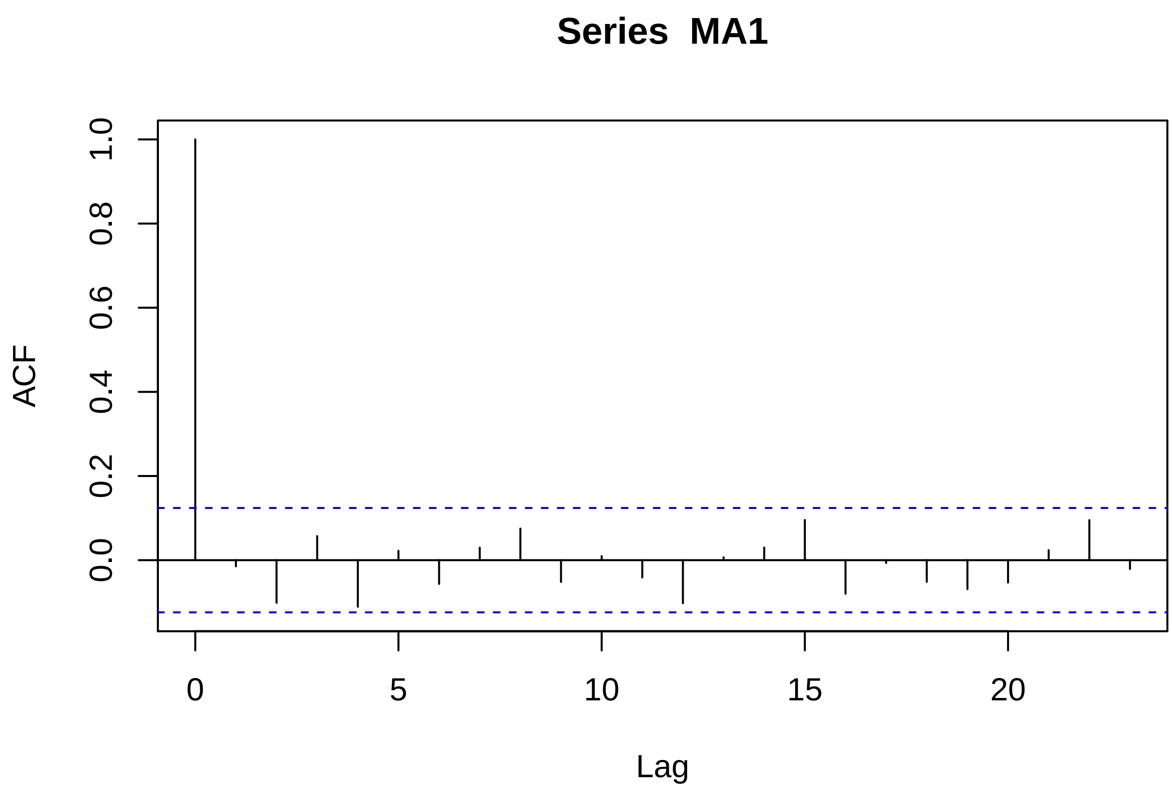
#Question 1

```
set.seed(123)
ls1 <- list(order = c(0, 0, 1), ma = 0.05)
mu <- 1

MA1 <- arima.sim(n = 250, model = ls1, sd = 0.31622777) + mu
plot(MA1)
```

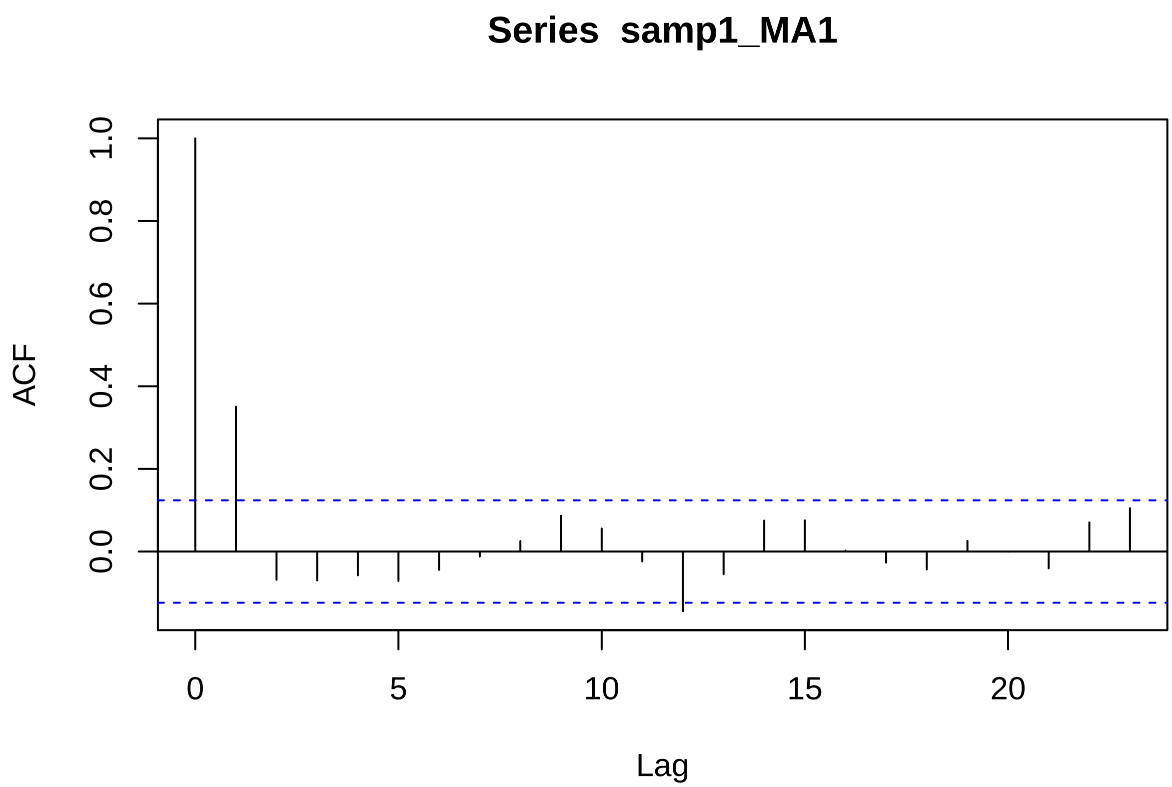


```r
acf(MA1)
```



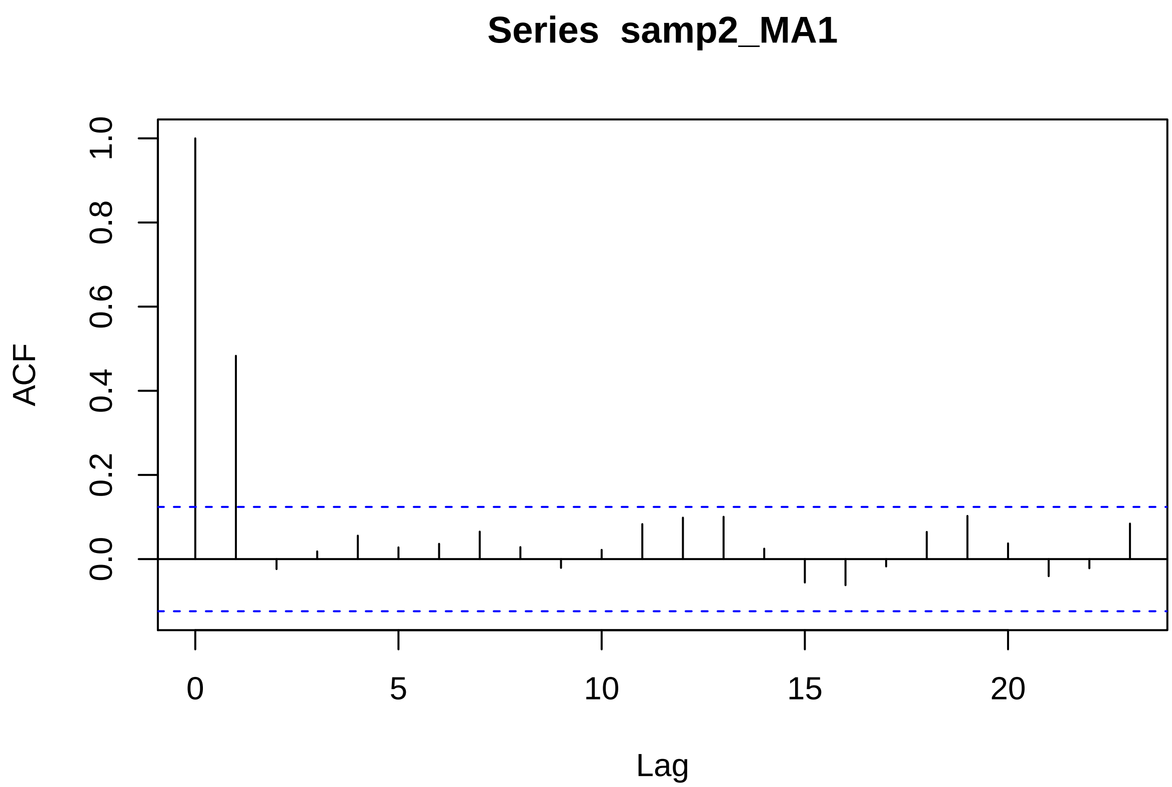

```r
ls1 <- list(order = c(0, 0, 1), ma = 0.5)

samp1_MA1 <- arima.sim(n = 250, model = ls1, sd = 0.31622777) + mu
acf(samp1_MA1)
```



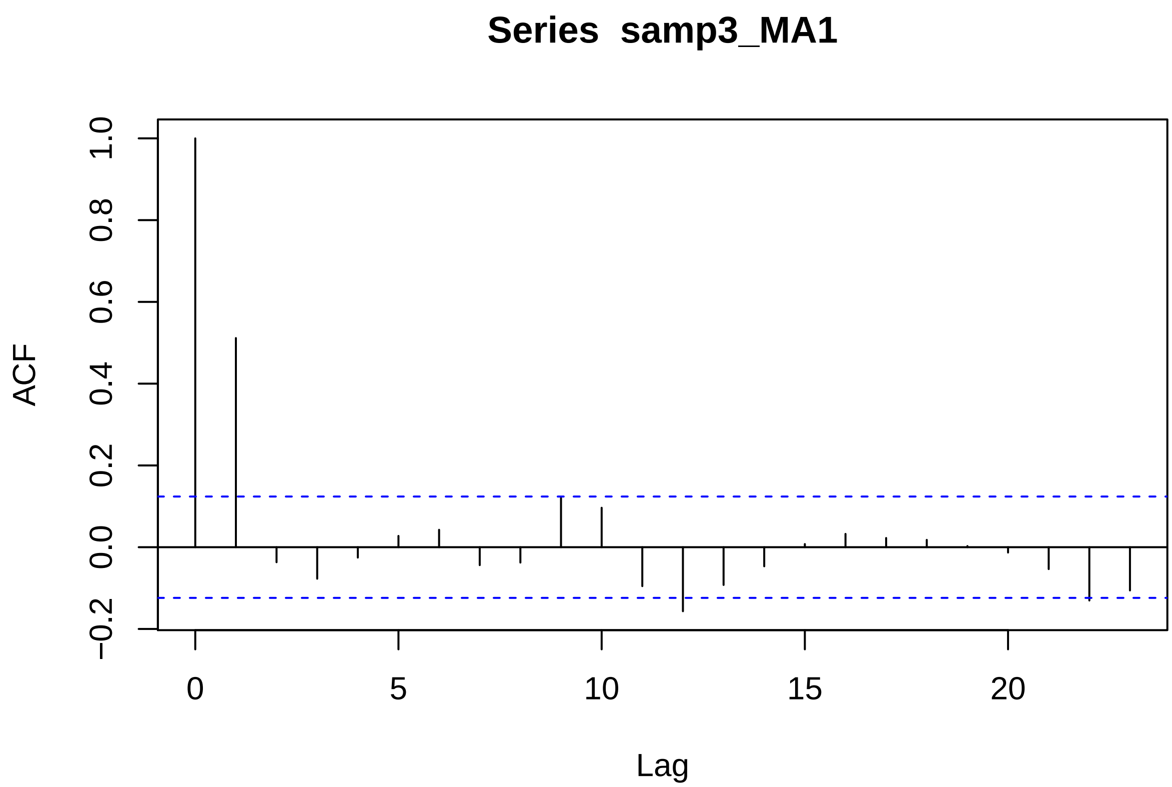

```r
ls1 <- list(order = c(0, 0, 1), ma = 0.9)

samp2_MA1 <- arima.sim(n = 250, model = ls1, sd = 0.31622777) + mu
acf(samp2_MA1)
```




```
ls1 <- list(order = c(0, 0, 1), ma = 0.99)

samp3_MA1 <- arima.sim(n = 250, model = ls1, sd = 0.31622777) + mu
acf(samp3_MA1)
```

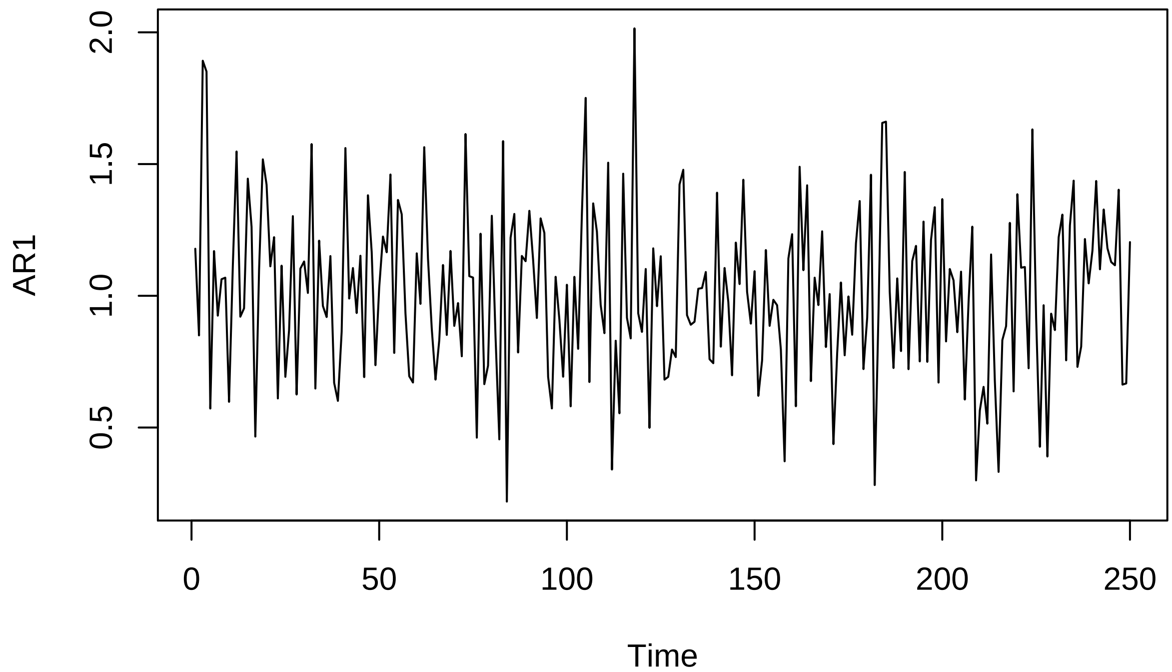


The theoretical ACF performed the best, with all spikes within the confidence linces, indicating 0 autocorrelation. The sample ACF's are mostly within the lines but all seem to have at least one long spike, indicating autocorrelation.

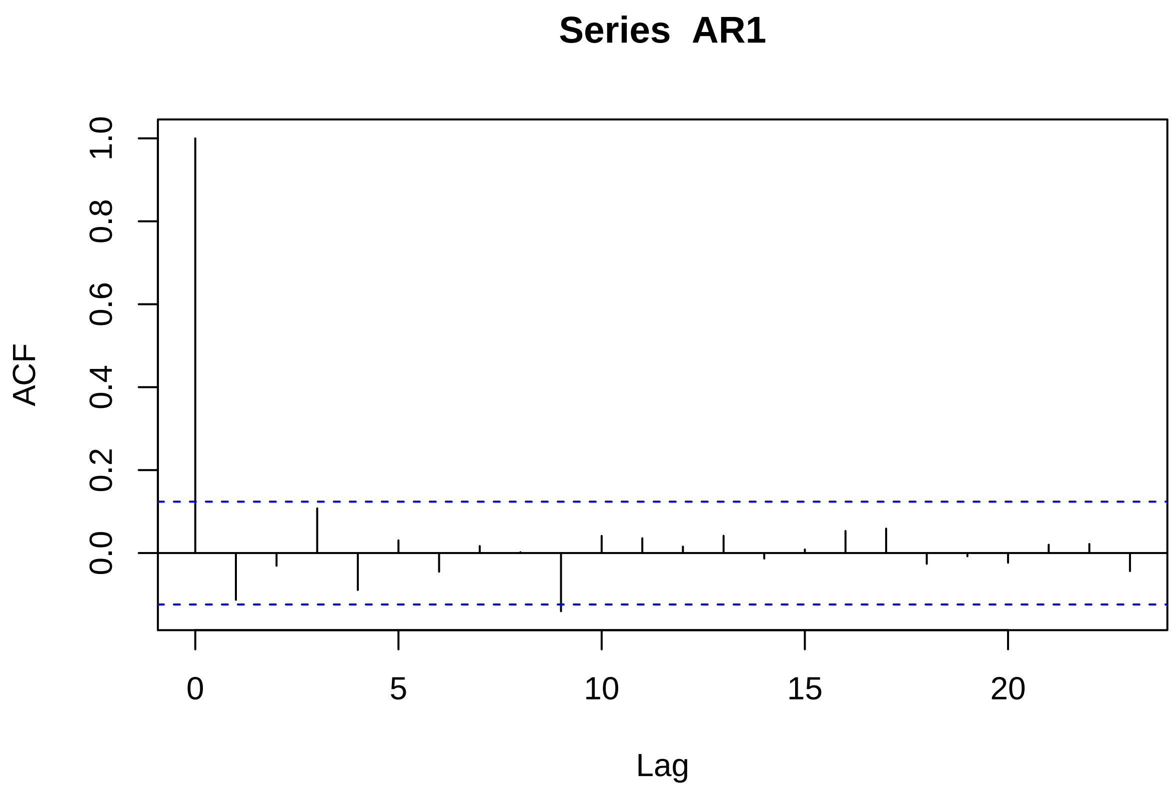
#Question 2

```
ls1 <- list(order = c(1, 0, 0), ar = -0.05)
mu <- 1

AR1 <- arima.sim(n = 250, model = ls1, sd = 0.31622777) + mu
plot(AR1)
```




```
acf(AR1)
```

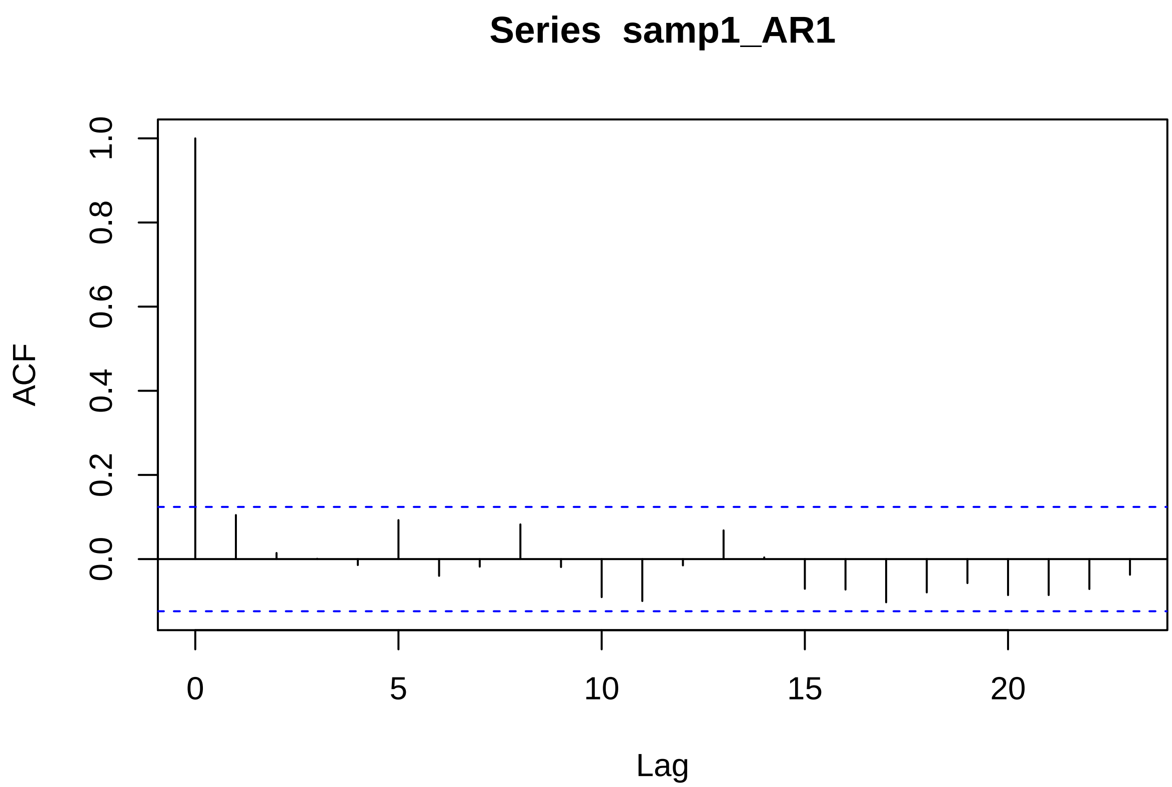



```
ls1 <- list(order = c(1, 0, 0), ar = 0)
mu <- 1

samp1_AR1 <- arima.sim(n = 250, model = ls1, sd = 0.31622777) + mu
```

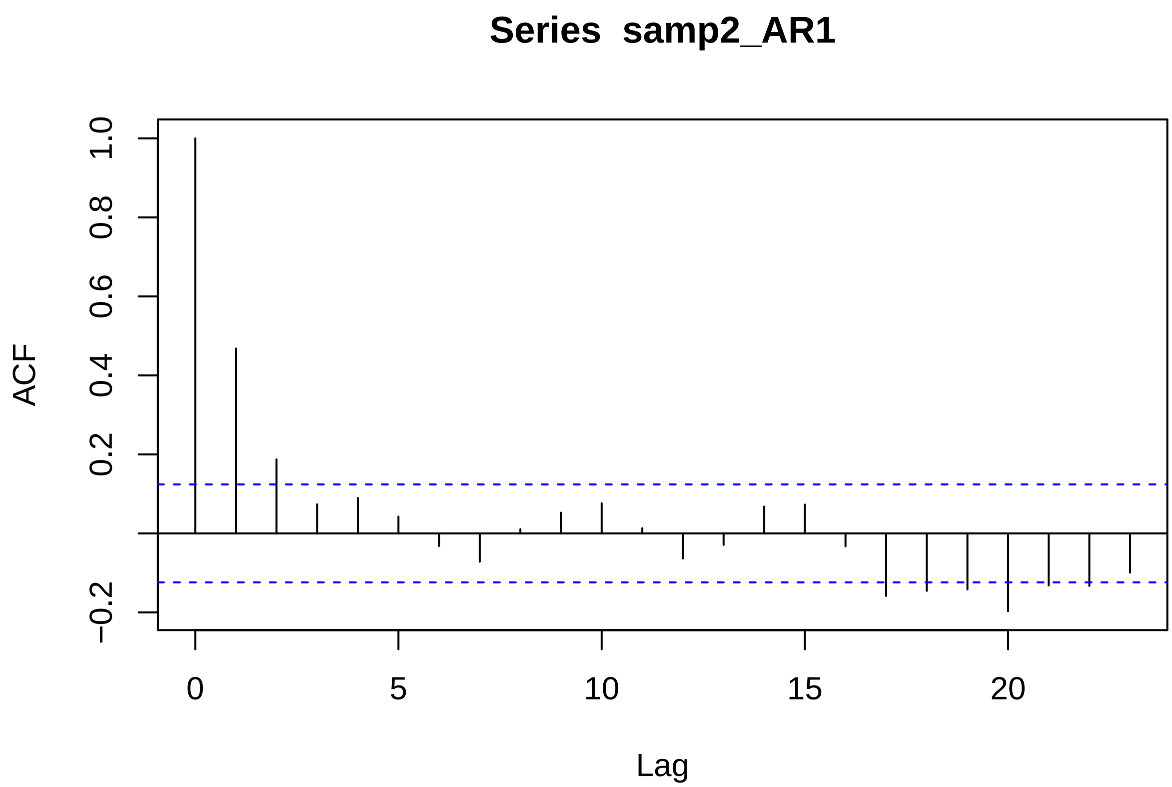
```
## Warning in min(Mod(polyroot(c(1, -model$ar)))): no non-missing arguments to min;
## returning Inf
```

```
acf(samp1_AR1)
```



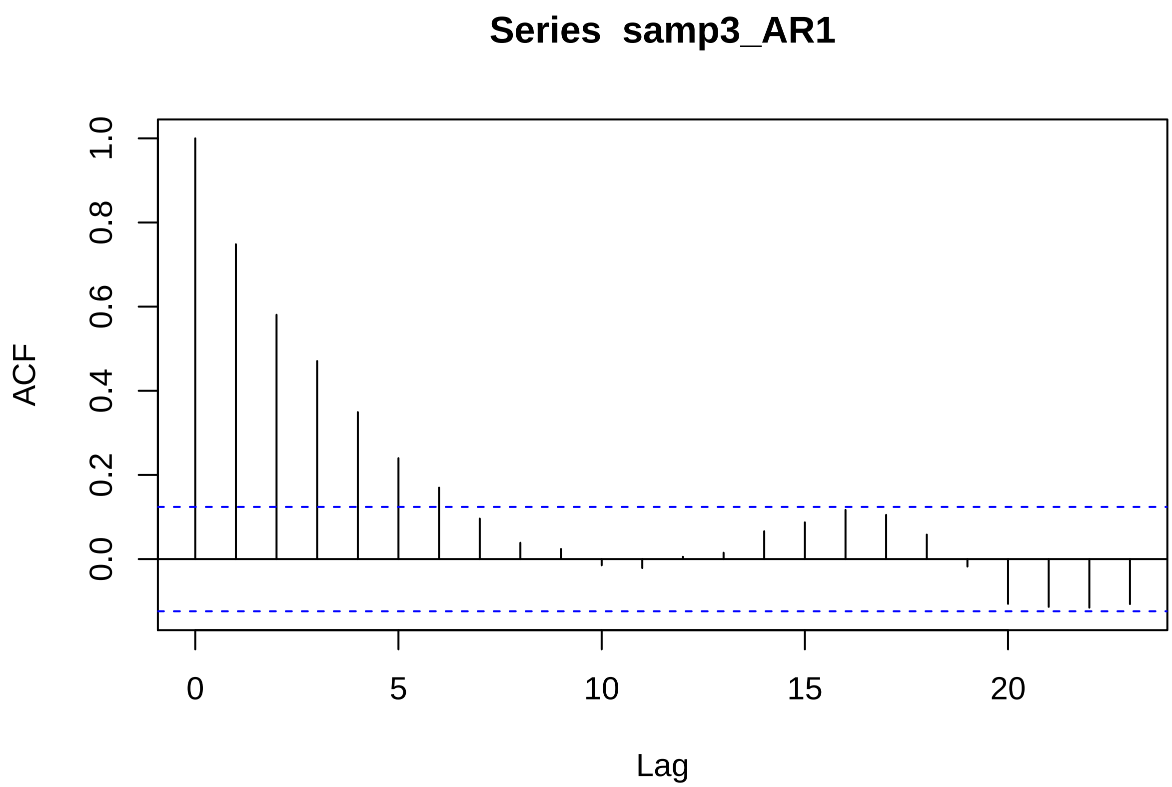

```r
ls1 <- list(order = c(1, 0, 0), ar = 0.5)
mu <- 1

samp2_AR1 <- arima.sim(n = 250, model = ls1, sd = 0.31622777) + mu
acf(samp2_AR1)
```



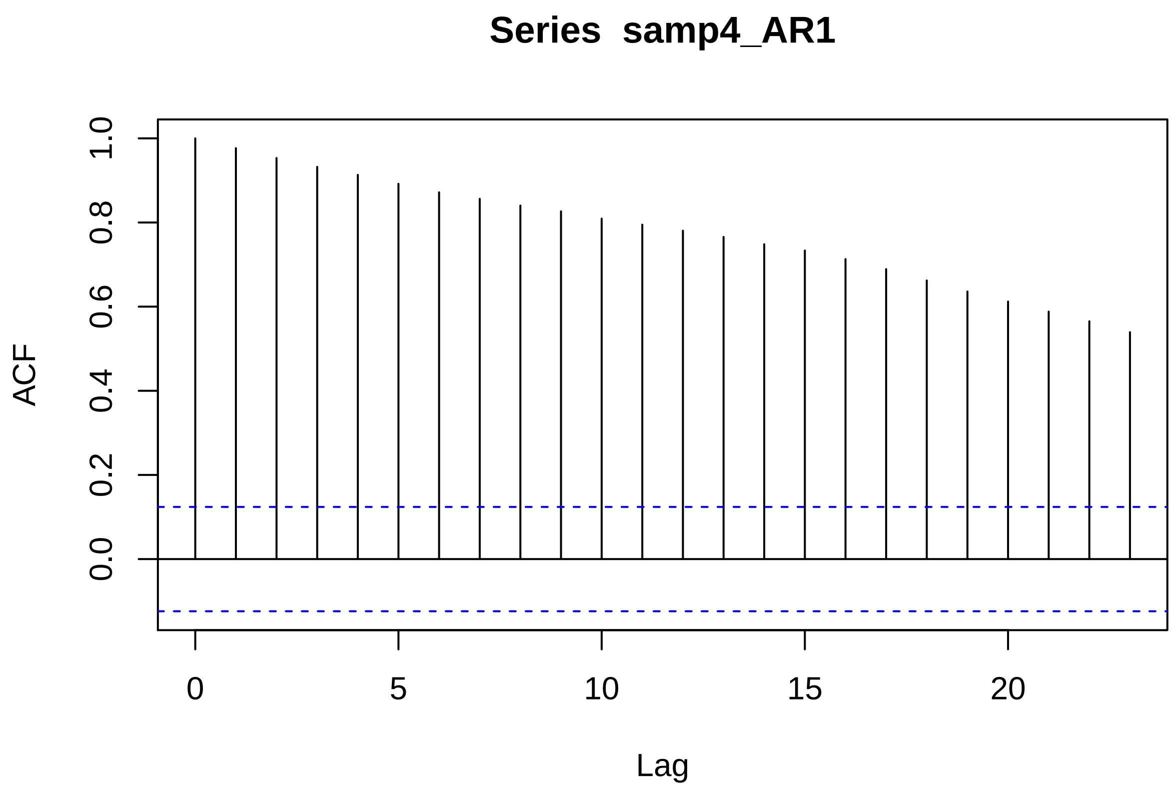

```
ls1 <- list(order = c(1, 0, 0), ar = 0.9)
mu <- 1

samp3_AR1 <- arima.sim(n = 250, model = ls1, sd = 0.31622777) + mu
acf(samp3_AR1)
```




```r
ls1 <- list(order = c(1, 0, 0), ar = 0.99)
mu <- 1

samp4_AR1 <- arima.sim(n = 250, model = ls1, sd = 0.31622777) + mu
acf(samp4_AR1)
```



The theoretical ACF for AR(1) looks good with all spikes within the confidence lines. As we increase phi, the ACF becomes increasingly significantly correlated. Sample 4, phi = 0.99, is is closest to nonstationary time series.

#Question 3

```r
setwd("/Users/madelyncaufield/Desktop/ECON432 Data Science for Financial Engineering/Homework 5")
```

```r
df <- read.csv("DAL.csv")
head(df)
```

```
##         Date  Open  High   Low Close Adj.Close   Volume
## 1 2015-01-01 49.92 50.01 46.25 46.97  42.27808 42590400
## 2 2015-01-08 47.48 48.50 44.47 45.31  40.78391 56548300
## 3 2015-01-15 45.09 49.60 44.61 49.18  44.26733 57643900
## 4 2015-01-22 49.93 51.06 48.28 48.46  43.61924 53395000
## 5 2015-01-29 48.67 50.35 44.81 46.96  42.26909 77611600
## 6 2015-02-05 46.93 46.98 43.71 44.98  40.48687 57113200
```

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
DAL.df = df %>%
          select(Date, Adj.Close)

head(DAL.df)
```

```
##         Date Adj.Close
## 1 2015-01-01  42.27808
## 2 2015-01-08  40.78391
## 3 2015-01-15  44.26733
## 4 2015-01-22  43.61924
## 5 2015-01-29  42.26909
## 6 2015-02-05  40.48687
```

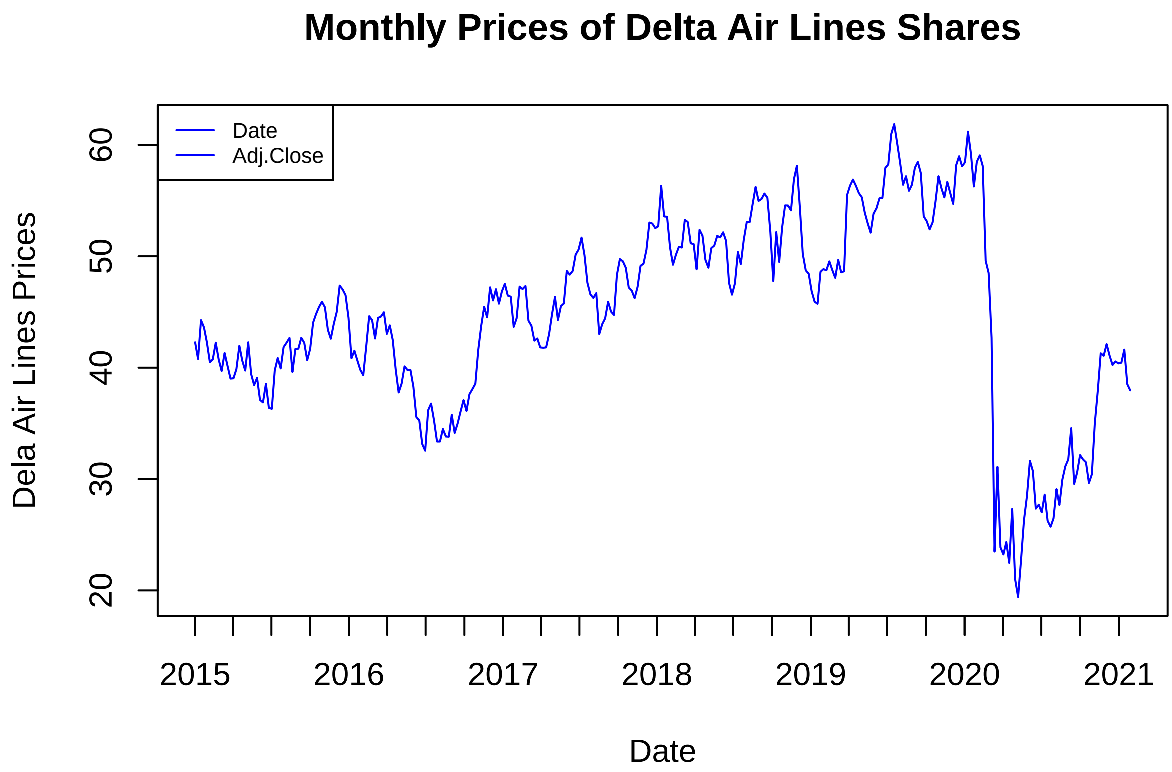
```
class(DAL.df$Date)
```

```
## [1] "character"
```

```
DAL.df$Date <- as.Date(DAL.df$Date,format = '%Y-%m-%d')
class(DAL.df$Date)
```

```
## [1] "Date"
```

```
plot(Adj.Close~Date, DAL.df, type='l',col="blue",xlab = 'Date',ylab='Dela Air Lines Prices',
     main='Monthly Prices of Delta Air Lines Shares')
T <- length(DAL.df$Date)
ticks <- seq(DAL.df$Date[1], DAL.df$Date[T], by = "quarters")
axis(1, at = ticks, labels = FALSE)
legend('topleft', names(DAL.df), col= "blue", lty=1, cex=.65)
```




```
DAL.df$cc.ret[2:T] <-  log(DAL.df$Adj.Close[2:T]/DAL.df$Adj.Close[1:(T-1)])
head(DAL.df$cc.ret)
```

```
## [1]          NA -0.03598113  0.08195913 -0.01474851 -0.03144228 -0.04307824
```

```
head(DAL.df)
```

```
##          Date Adj.Close      cc.ret
## 1 2015-01-01  42.27808          NA
## 2 2015-01-08  40.78391 -0.03598113
## 3 2015-01-15  44.26733  0.08195913
## 4 2015-01-22  43.61924 -0.01474851
## 5 2015-01-29  42.26909 -0.03144228
## 6 2015-02-05  40.48687 -0.04307824
```

```
DAL.df <- DAL.df[,c('Date','cc.ret')]
DAL.df <- DAL.df[-c(1),]

head(DAL.df)
```

```
##          Date       cc.ret
## 2 2015-01-08 -0.035981129
## 3 2015-01-15  0.081959128
## 4 2015-01-22 -0.014748505
```

```
## 5 2015-01-29 -0.031442275
## 6 2015-02-05 -0.043078240
## 7 2015-02-12  0.006426629
```

#3A

```r
Ret = DAL.df$cc.ret
Box.test(Ret, lag = 10, type = c("Ljung-Box"), fitdf = 0)
```

```
##
##  Box-Ljung test
##
## data:  Ret
## X-squared = 38.893, df = 10, p-value = 2.651e-05
```

The p-value is larger than 0.05, indicating that the residuals are independent.

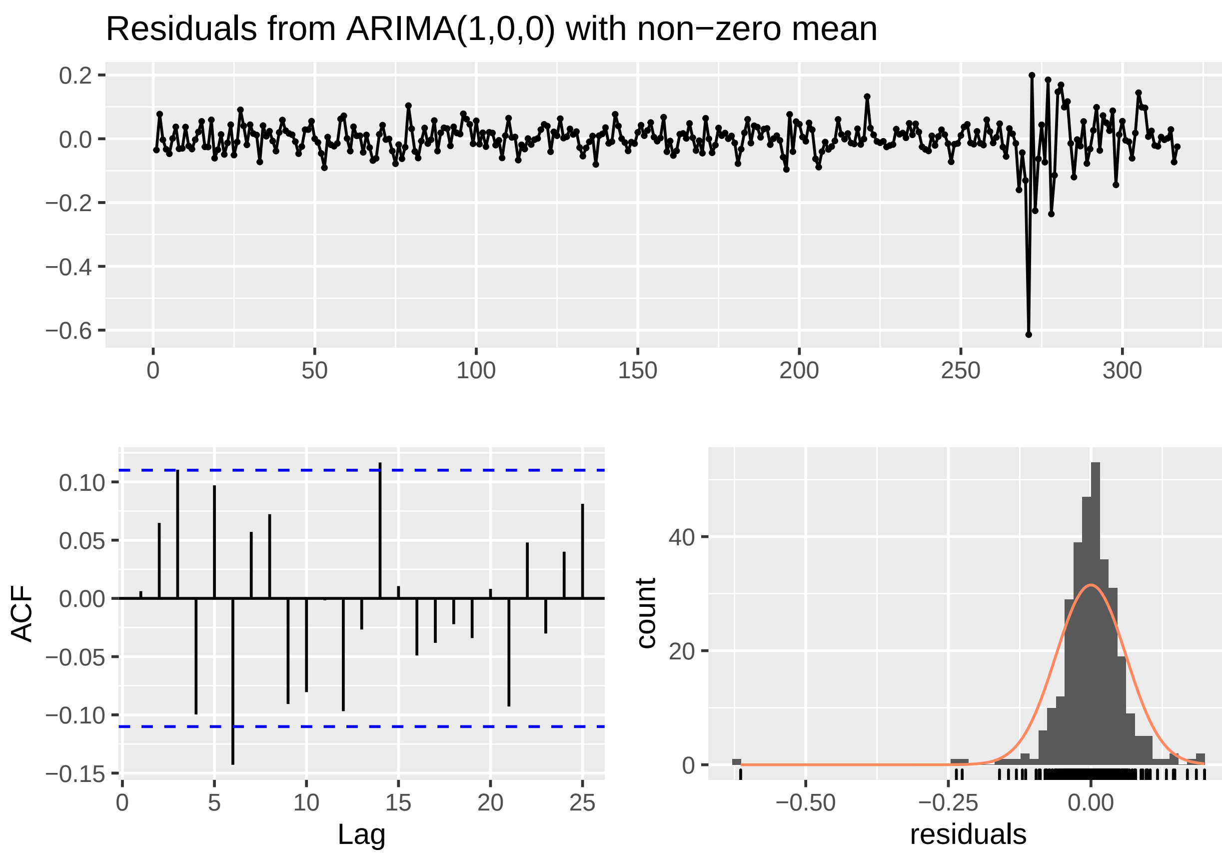
#3B

```r
Ret = DAL.df$cc.ret
AR1 = (arima(Ret, order = c(1,0,0)))
```

```r
library(forecast)
```

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

```r
checkresiduals(AR1)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(1,0,0) with non-zero mean
## Q* = 25.792, df = 8, p-value = 0.00114
##
## Model df: 2.   Total lags used: 10
```

The AR(1) model does not model the linear dependence of the cc return appropriately.

#3C

```
auto.arima(Ret, max.p=10, max.q=0, ic="aic", stepwise=F)
```

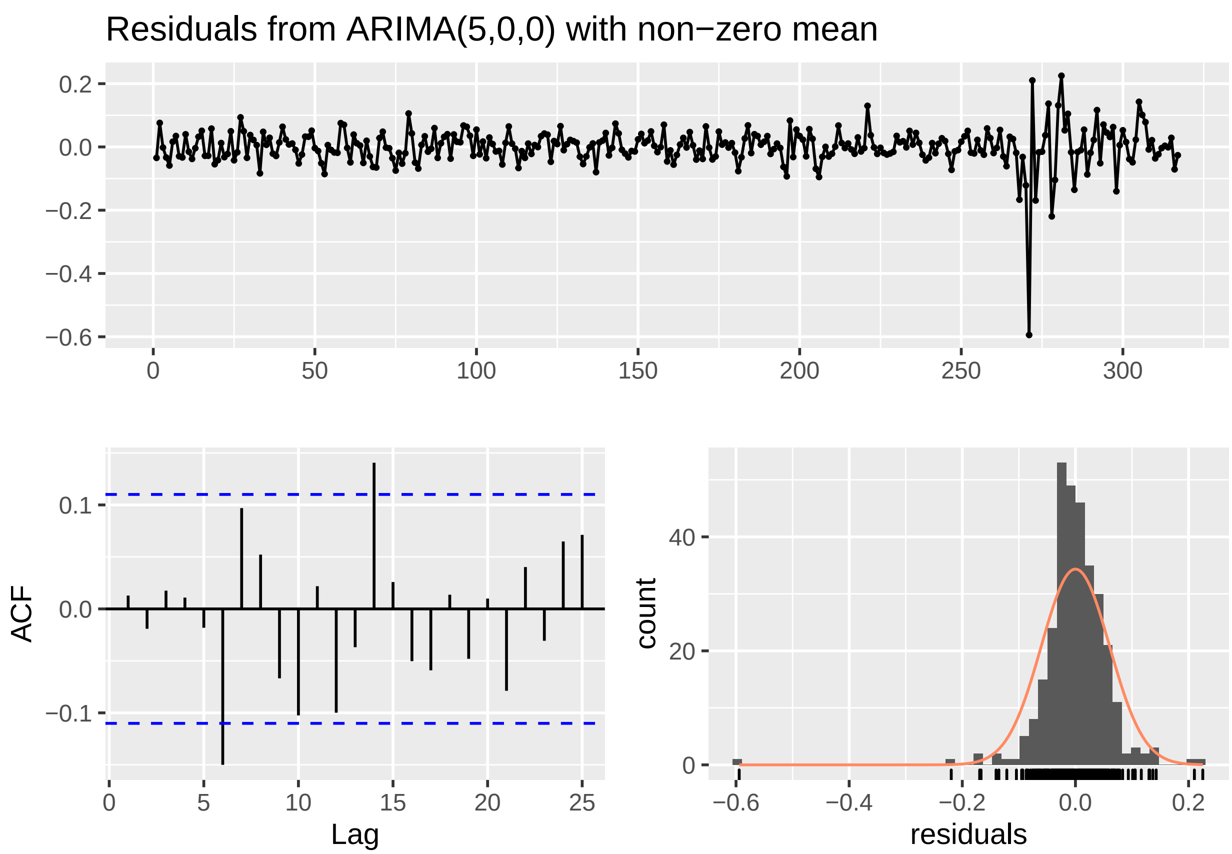
```
## Series: Ret
## ARIMA(5,0,0) with zero mean
##
## Coefficients:
##           ar1     ar2     ar3      ar4     ar5
##       -0.1134  0.0611  0.1122  -0.0931  0.0906
## s.e.   0.0559  0.0562  0.0558   0.0561  0.0559
##
## sigma^2 estimated as 0.00369:  log likelihood=440.57
## AIC=-869.14   AICc=-868.87   BIC=-846.58
```

```
AR5 = (arima(x = Ret, order = c(5,0,0)))
AR5
```

```
##
## Call:
## arima(x = Ret, order = c(5, 0, 0))
##
## Coefficients:
##           ar1     ar2     ar3      ar4     ar5  intercept
##       -0.1134  0.0611  0.1122  -0.0931  0.0906    -0.0004
## s.e.   0.0559  0.0562  0.0558   0.0561  0.0559     0.0036
##
## sigma^2 estimated as 0.003632:  log likelihood = 440.57,  aic = -867.15
```

#3D

```
checkresiduals(AR5)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(5,0,0) with non-zero mean
## Q* = 16.61, df = 4, p-value = 0.002301
##
## Model df: 6.   Total lags used: 10
```

The AR(5) model does not model the linear dependence of the cc return appropriately.

#3E

```r
auto.arima(Ret, max.p=10, max.q=10, ic = "aic", stepwise = F)
```

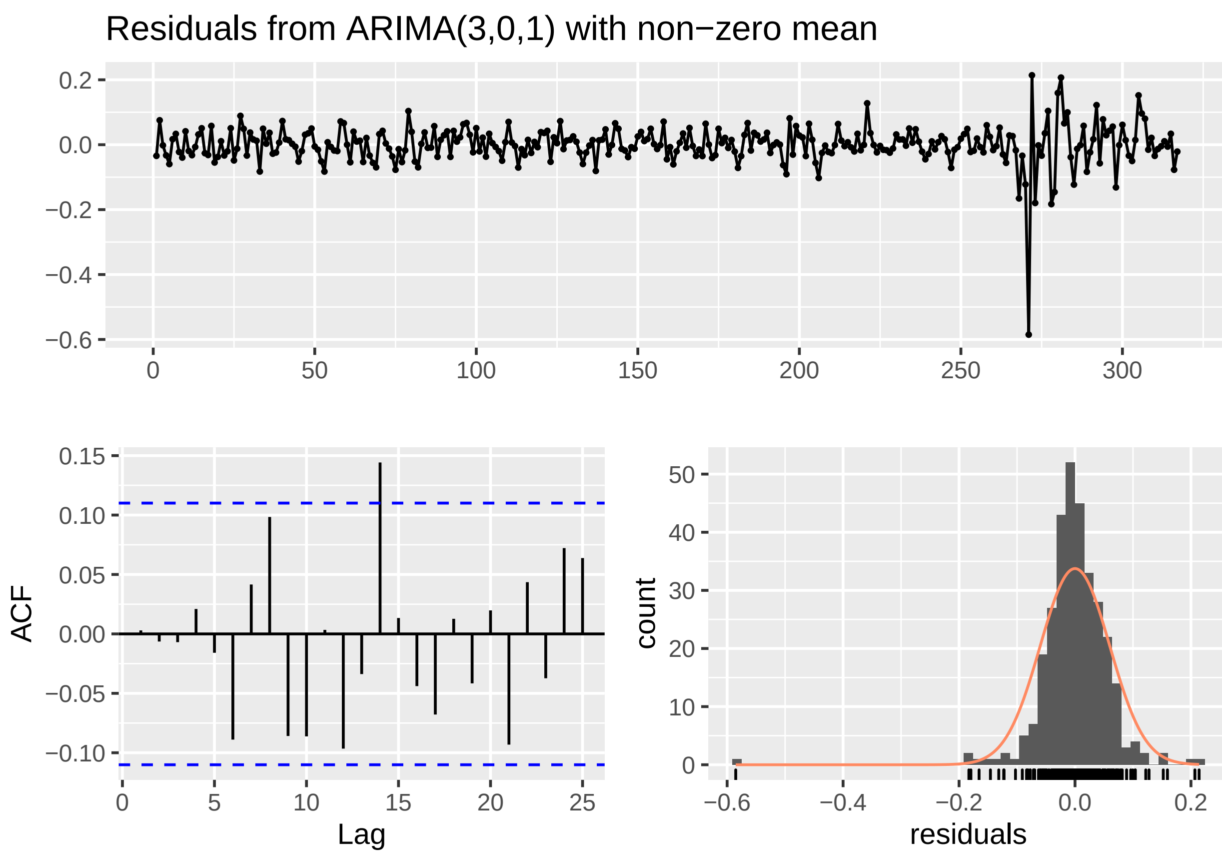
```
## Series: Ret
## ARIMA(3,0,1) with zero mean
##
## Coefficients:
##           ar1      ar2     ar3     ma1
##       -0.8801  -0.0305  0.1815  0.7828
## s.e.   0.0932   0.0752  0.0579  0.0803
##
## sigma^2 estimated as 0.003622:  log likelihood=442.98
## AIC=-875.96   AICc=-875.77   BIC=-857.16
```

```r
ARMA31 = (arima(x = Ret, order = c(3,0,1)))
ARMA31
```

```
##
## Call:
## arima(x = Ret, order = c(3, 0, 1))
##
## Coefficients:
##           ar1      ar2     ar3     ma1  intercept
##       -0.8799  -0.0305  0.1815  0.7826    -0.0003
## s.e.   0.0933   0.0752  0.0579  0.0804     0.0035
##
## sigma^2 estimated as 0.003576:  log likelihood = 442.98,  aic = -873.97
```

#3F

```r
checkresiduals(ARMA31)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(3,0,1) with non-zero mean
## Q* = 11.428, df = 5, p-value = 0.04352
##
## Model df: 5.   Total lags used: 10
```

The ARMA(3,1) model models the linear dependence of the cc return the most appropriately out of all the other models.There appears to be only one spike in the ACF spot of the fitted residuals.