

Econometrics II - Problem 1

William Radaic Peron

September 20, 2020

Loading the database and creating dummy variables:

```
df <- read_excel("RS_USD.xlsx")

names(df)[names(df) == "R$/US$"] <- "p"

names(df)[names(df) == "Variação (em %)"] <- "delta"

names(df)[names(df) == "Data"] <- "date"

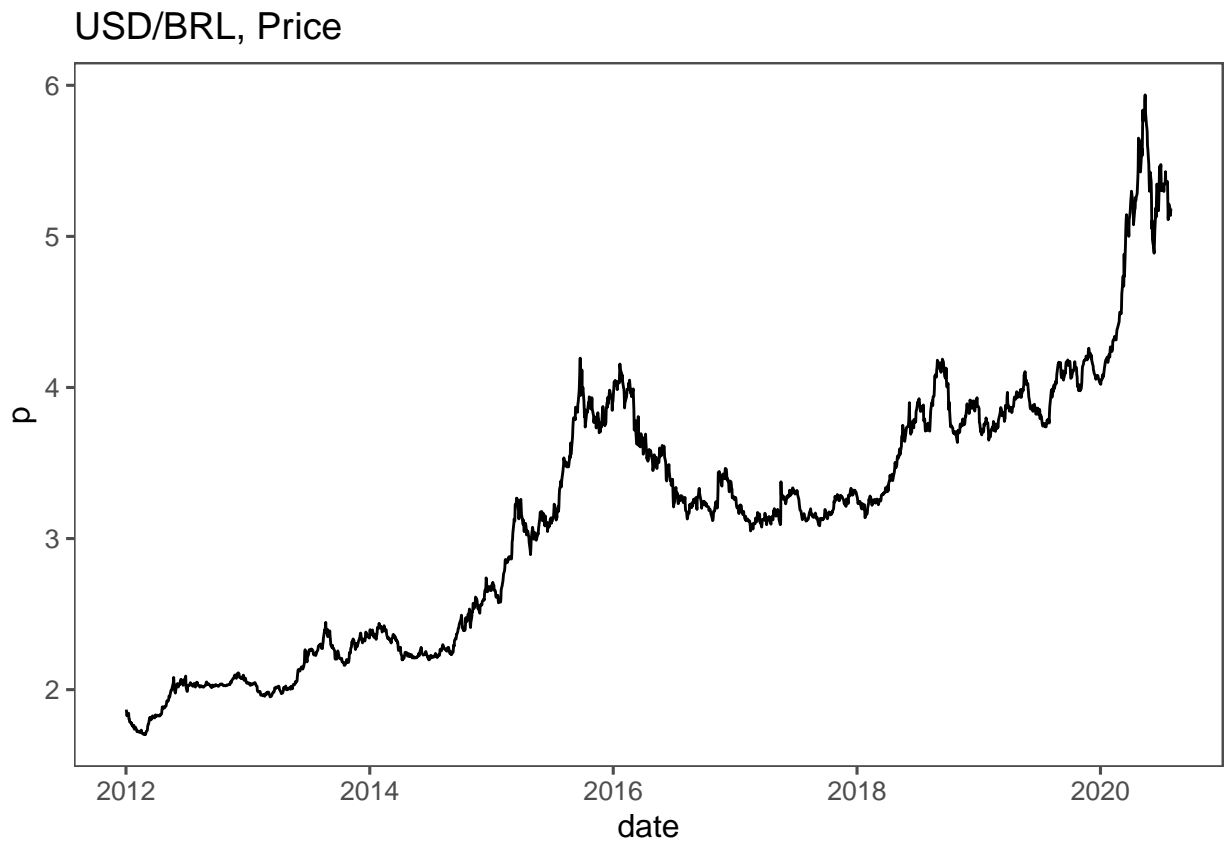
sign <- as.numeric(df$delta > 0)

count <- c(1:2153)

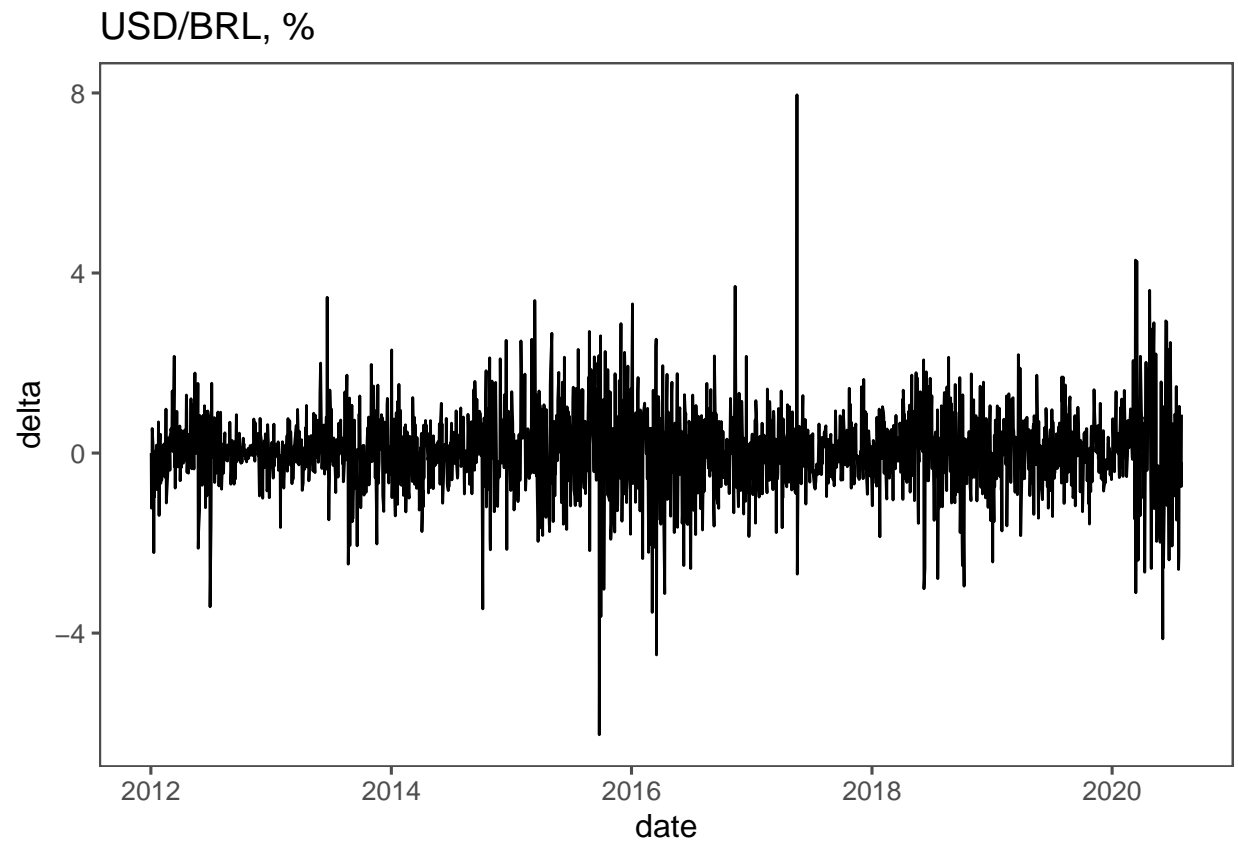
df <- data.frame(count, df, sign)
```

Before constructing our models, we need to check (intuitively) if the series at hand is *stationary* and *ergodic*. For this, we're going to plot the time series, its autocorrelations and partial autocorrelations.

```
pplot <- ggplot(data = df, aes(x = date, y = p)) + geom_line() +
  ggtitle("USD/BRL, Price") + theme_few()
pplot
```

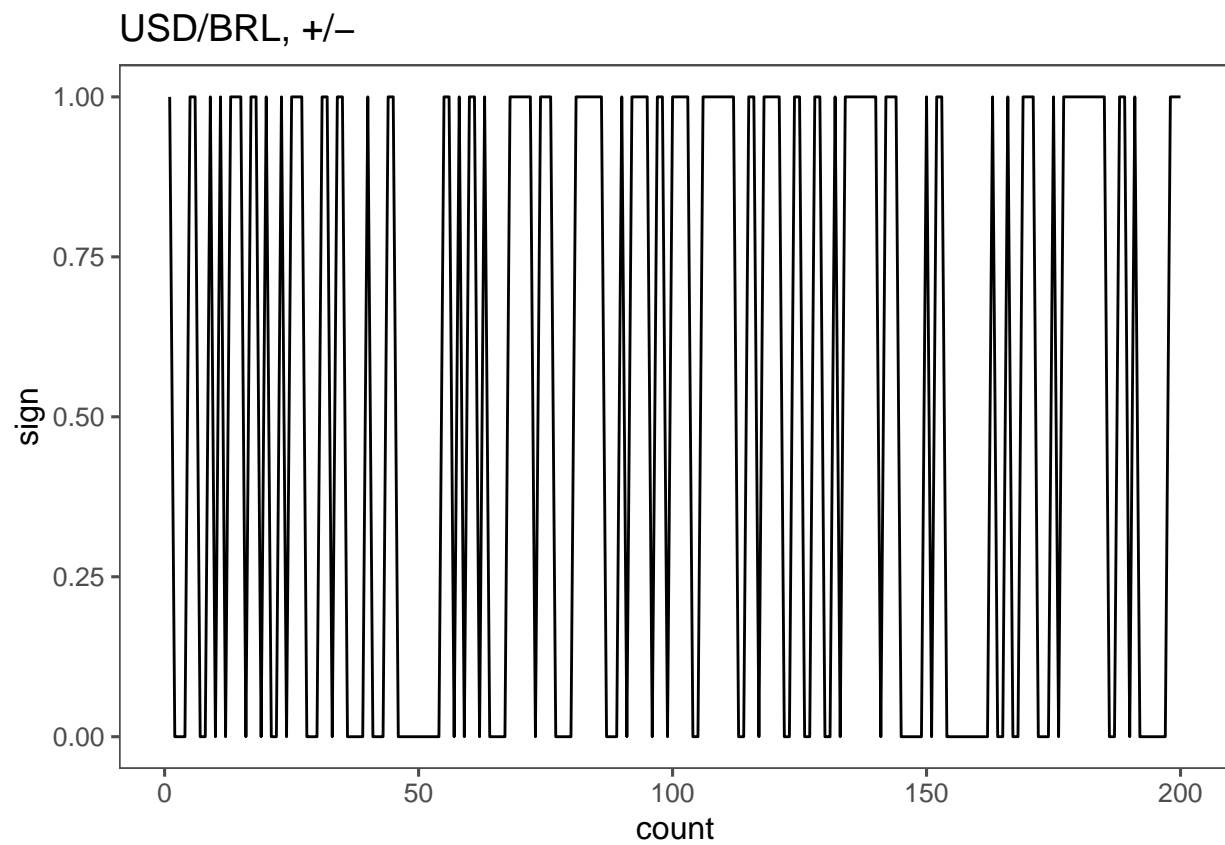


```
deltaplot <- ggplot(data = df, aes(x = date, y = delta)) + geom_line() +  
  ggtitle("USD/BRL, %") + theme_few()  
deltaplot
```



```
dummyplot <- ggplot(data = df, aes(x = count, y = sign)) + geom_line() +  
  ggtitle("USD/BRL, +/-") + xlim(1, 200) + theme_few()  
dummyplot
```

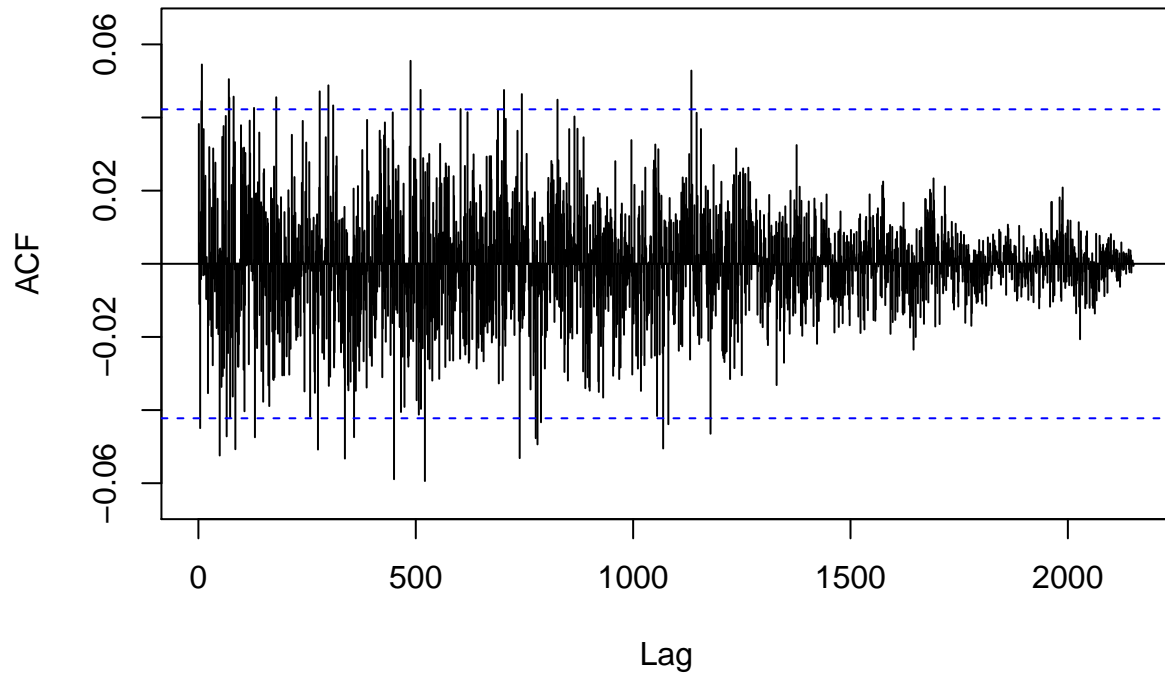
```
## Warning: Removed 1953 row(s) containing missing values (geom_path).
```



```
# For delta
```

```
acf_delta <- Acf(df$delta, lag.max = 5000)
```

Series df\$delta



```
acf_test_values <- acf_delta$acf/sd(acf_delta$acf)
```

```
head(data.frame(acf_test_values))
```

```
##   acf_test_values
## 1      37.9547672
## 2       1.4506537
## 3      -0.4173129
## 4       0.2125873
## 5      -1.7053782
## 6       0.5358210
```

```
facst <- ggAcf(df$delta, type = "correlation", lag.max = 20,
  plot = T) + theme_few()
fac1t <- ggAcf(df$delta, type = "correlation", lag.max = 5000,
  plot = T) + theme_few()
```

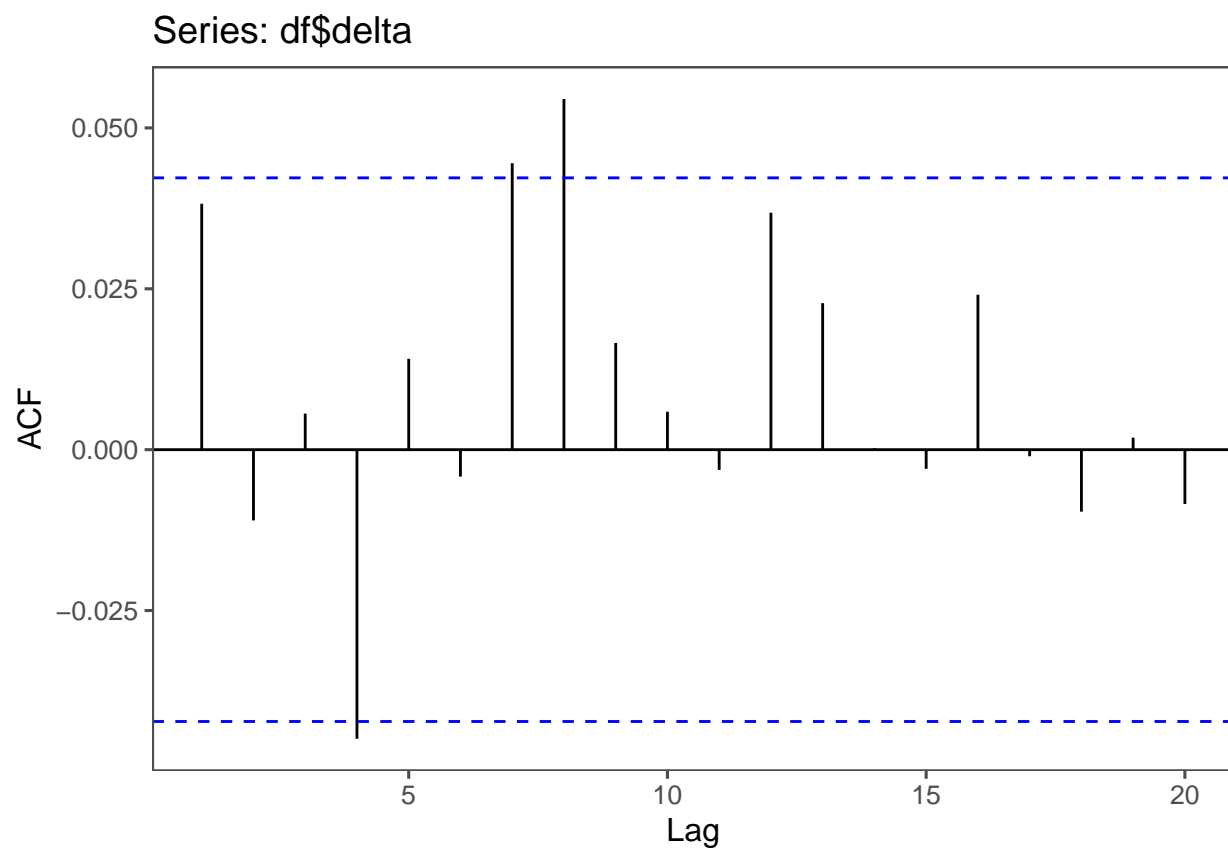
```
facpst <- ggPacf(df$delta, type = "correlation", lag.max = 100,
  plot = T) + theme_few()
```

```
## Warning: Ignoring unknown parameters: type
```

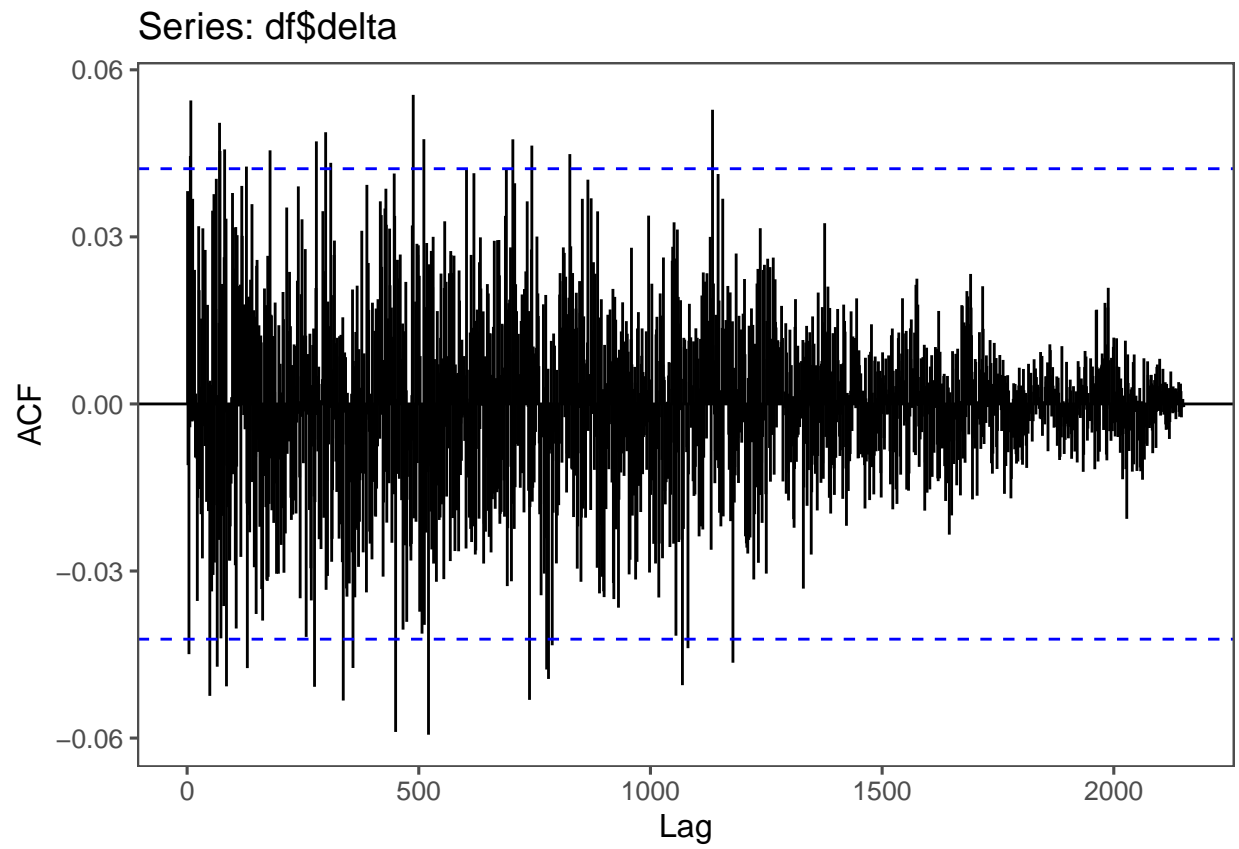
```
facplt <- ggPacf(df$delta, type = "correlation", lag.max = 5000,
  plot = T) + theme_few()
```

```
## Warning: Ignoring unknown parameters: type
```

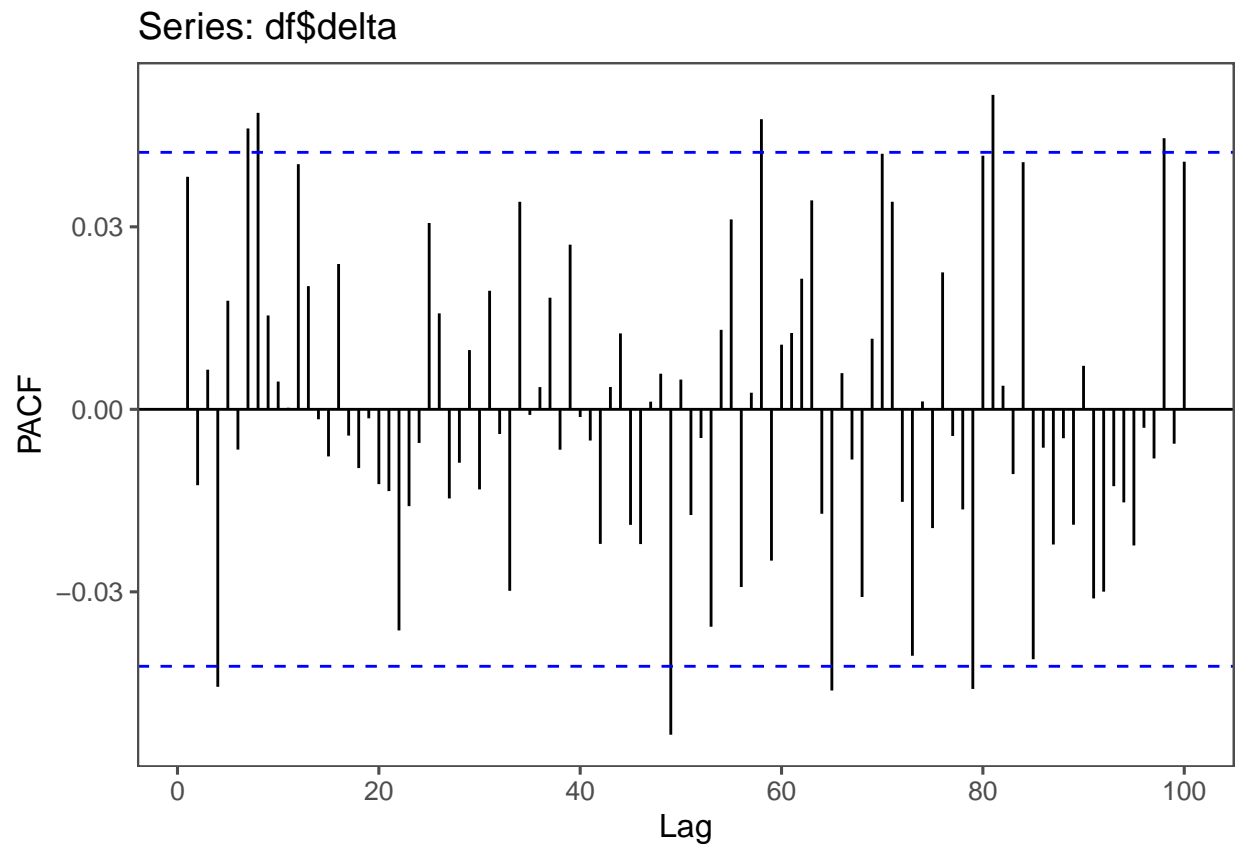
facst



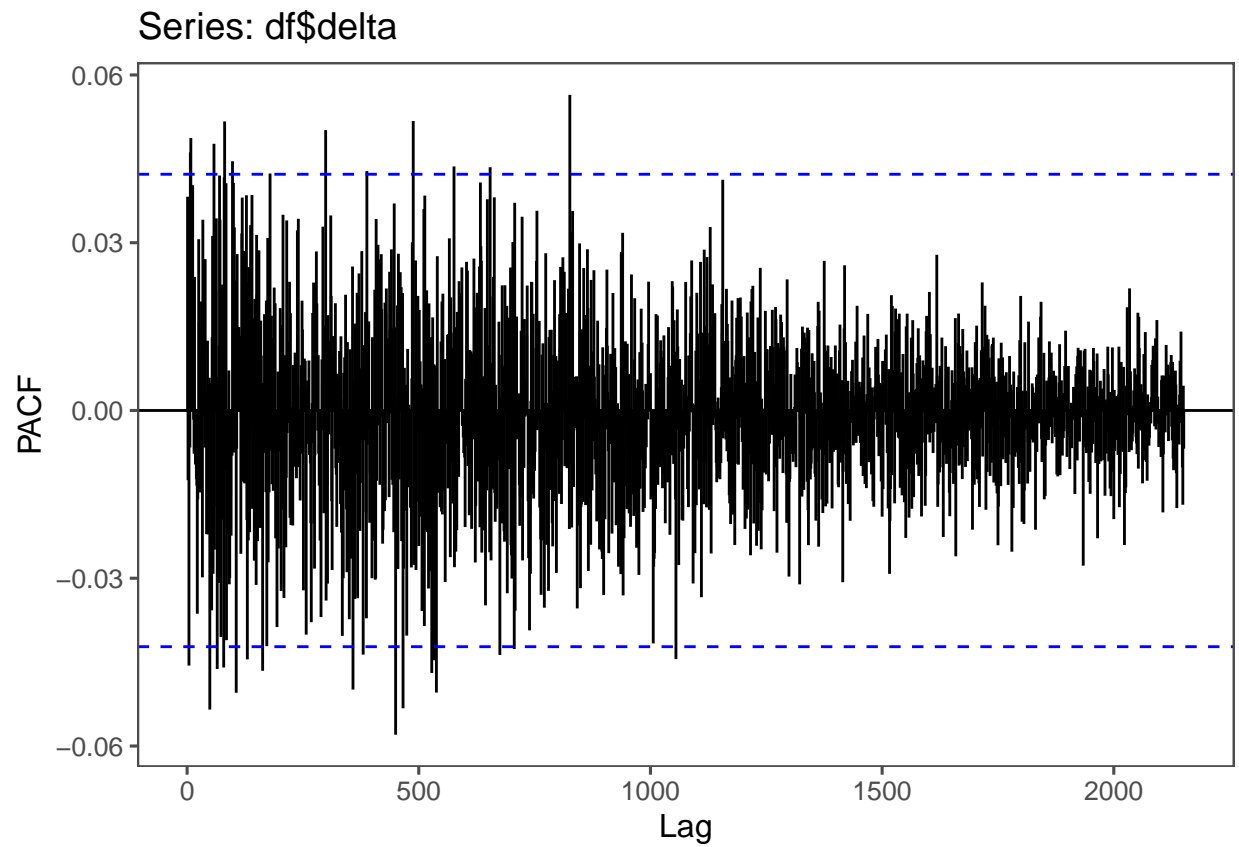
fac1t



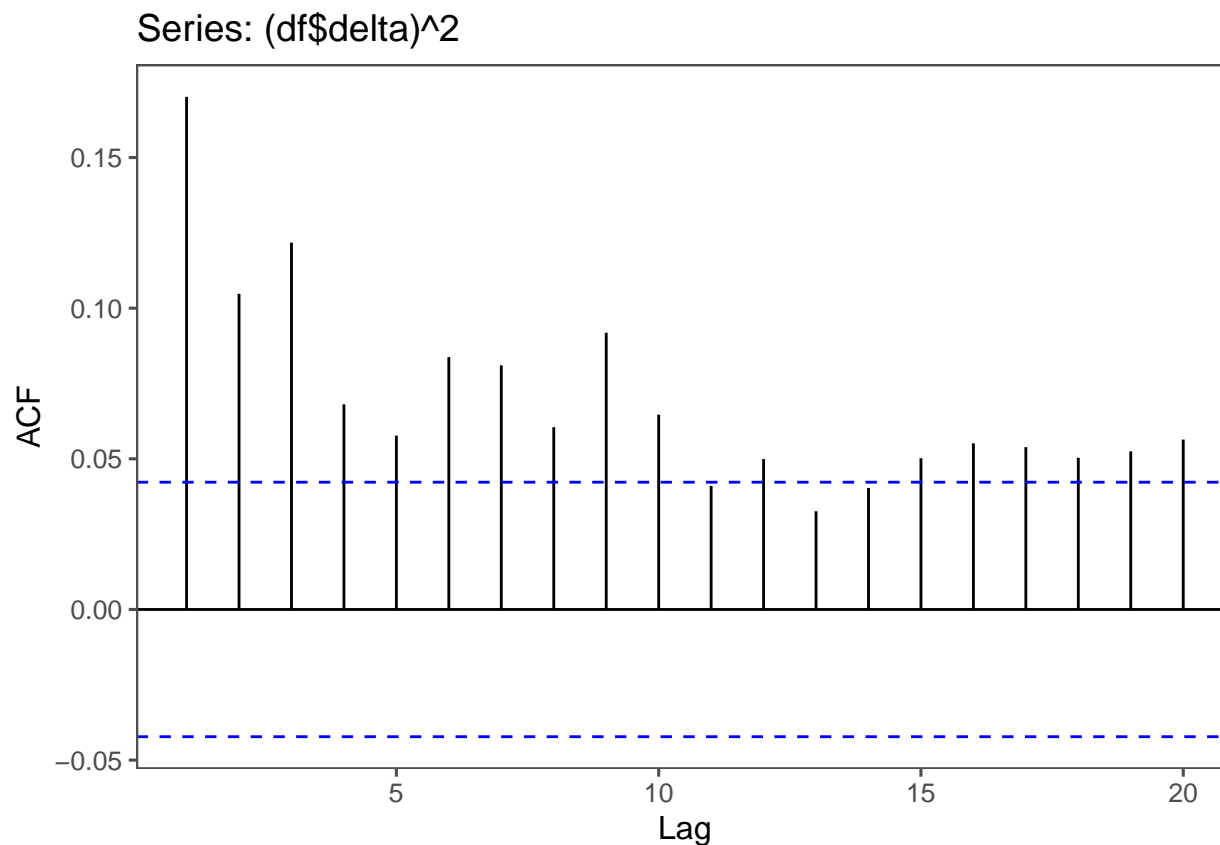
facpst



facplt



```
facst2 <- ggAcf((df$delta)^2, type = "correlation", lag.max = 20,  
  plot = T) + theme_few()  
facst2
```



Let's now create our first ARMA models (equivalent to ARIMA with 2nd argument = 0). We'll begin with the first hypothesis: $\mathbb{P}(+) = \mathbb{P}(-)$. Modelling this with an AR(1), we have:

$$Sign_{t+1} = \alpha + \beta Sign_t + \varepsilon, \quad \varepsilon \sim wn(0, \sigma^2)$$

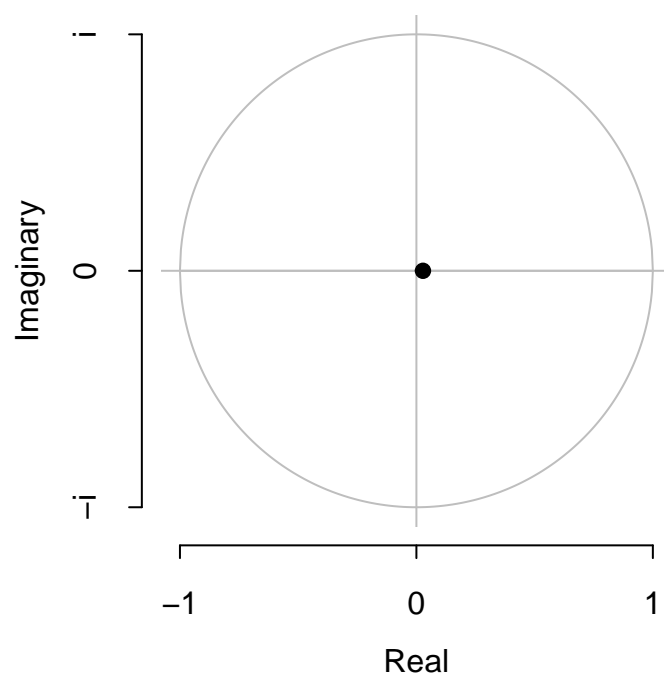
In R, we'll use the package *forecast* to construct this model:

```
AR1sign <- Arima(df$sign, order = c(1, 0, 0))
summary(AR1sign)

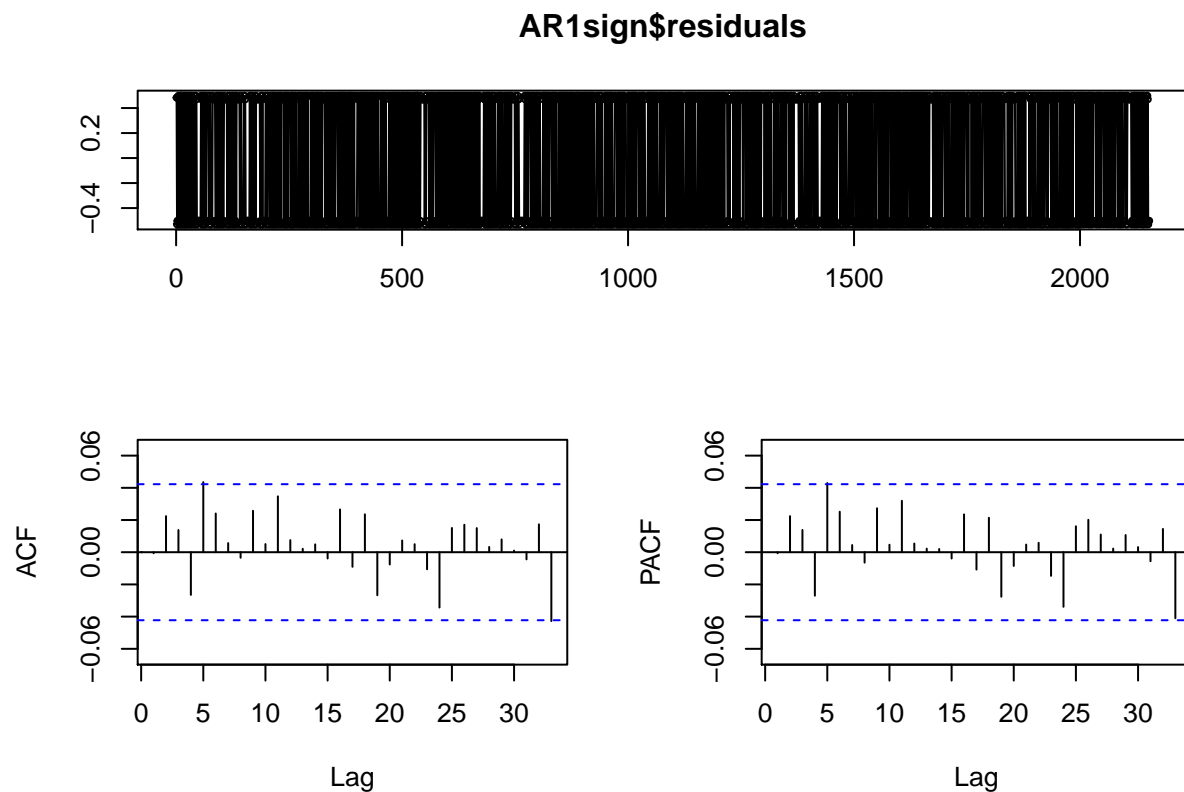
## Series: df$sign
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##      ar1      mean
##      0.0278  0.5165
## s.e.  0.0215  0.0111
##
## sigma^2 estimated as 0.2498:  log likelihood=-1560.63
## AIC=3127.26   AICc=3127.27   BIC=3144.28
##
## Training set error measures:
##              ME      RMSE      MAE  MPE  MAPE      MASE      ACF1
## Training set 2.157119e-05 0.4995356 0.4990724 -Inf  Inf  1.027755 -0.0006313563

plot(AR1sign)
```

Inverse AR roots



```
tsdisplay(AR1sign$residuals)
```



With the results of the summary, we can now apply a hypothesis test for our first question.¹

$$H_0 : \beta = 0$$

$$H_1 : \beta \neq 0$$

$$\frac{\hat{ar}_1 - ar_1}{s.e.(ar_1)};$$

```
AR1sign$coef[1]/sqrt(AR1sign$var.coef[1, 1])
```

```
##      ar1
## 1.287942
```

The second hypothesis in the problem refers to the delta of the variation:

$$\mathbb{E}(\Delta|+) \neq \mathbb{E}(\Delta|-).$$

$$\Delta_{t+1} = \alpha + \beta Sign_t + \varepsilon, \quad \varepsilon \sim wn(0, \sigma^2).$$

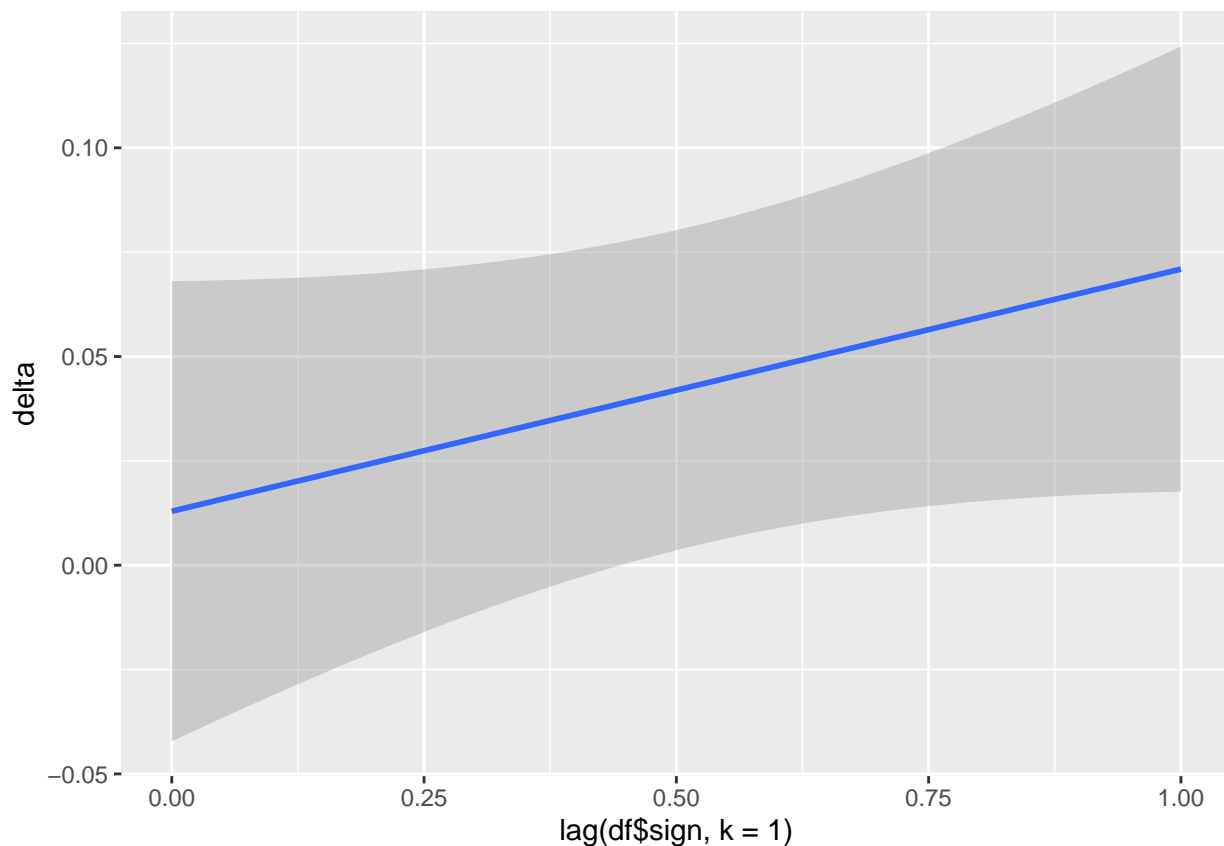
```
lmsignt <- lm(delta ~ lag(df$sign, k = 1), data = df)
summary(lmsignt)
```

```
##
## Call:
## lm(formula = delta ~ lag(df$sign, k = 1), data = df)
```

¹Testing β is equivalent to testing γ .

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.3285 -0.4706 -0.0060  0.4655  7.9442
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.01293    0.02811   0.460   0.646
## lag(df$sign, k = 1) 0.05802    0.03911   1.484   0.138
##
## Residual standard error: 0.9066 on 2150 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.001023, Adjusted R-squared:  0.000558
## F-statistic: 2.201 on 1 and 2150 DF, p-value: 0.1381
ggplot(df, aes(x = lag(df$sign, k = 1), y = delta)) + geom_smooth(method = "lm")

## Warning: Use of `df$sign` is discouraged. Use `sign` instead.
## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```



$$\Delta_{t+1} = \alpha + \beta_1 \Delta_t + \beta_2 \text{Sign}_t + \varepsilon, \quad \varepsilon \sim wn(0, \sigma^2)$$

```
AR1delta <- Arima(df$delta, order = c(1, 0, 0), xreg = lag(df$sign,
  k = 1))
```

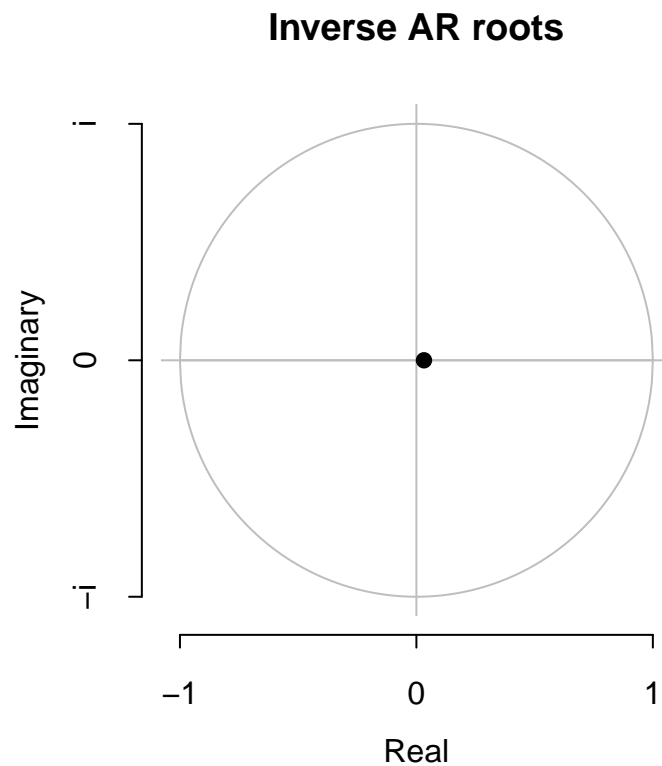
```
summary(AR1delta)
```

```
## Series: df$delta
## Regression with ARIMA(1,0,0) errors
##
## Coefficients:
##          ar1  intercept    xreg
##         0.0321    0.0343  0.0166
## s.e.  0.0306    0.0351  0.0556
##
## sigma^2 estimated as 0.8219:  log likelihood=-2840.96
## AIC=5689.93   AICc=5689.94   BIC=5712.62
##
## Training set error measures:
##              ME      RMSE      MAE MPE MAPE      MASE      ACF1
## Training set 1.147232e-05 0.9059341 0.643417 NaN  Inf  0.7252668 0.0003998294
```

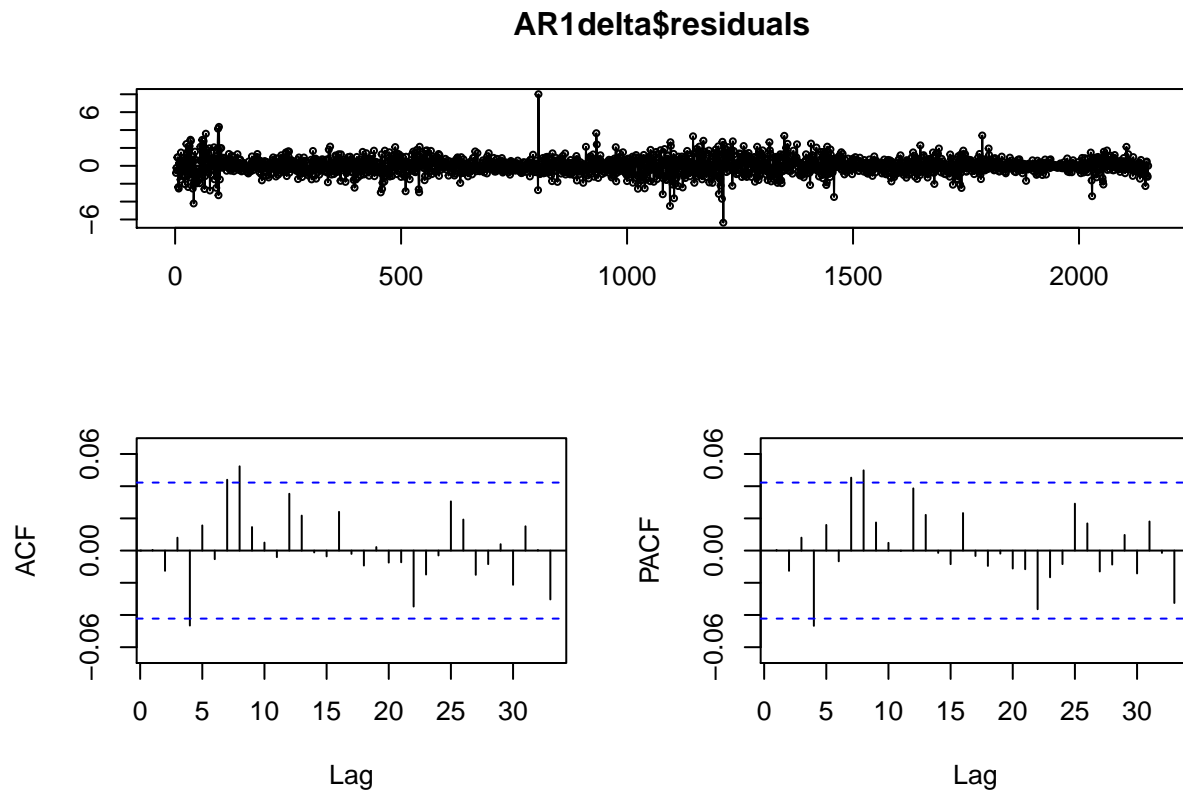
```
AR1delta$coef[1]/sqrt(AR1delta$var.coef[1, 1])
```

```
##      ar1
## 1.04747
```

```
plot(AR1delta)
```



```
tsdisplay(AR1delta$residuals)
```



The last hypothesis in the problem refers to the variance:

$$\mathbb{E}(\Delta_{t+1}^2 | \Delta_t).$$

$$\Delta_{t+1}^2 = \alpha + \beta \Delta_t^2 + \varepsilon, \quad \varepsilon \sim wn(0, \sigma^2)$$

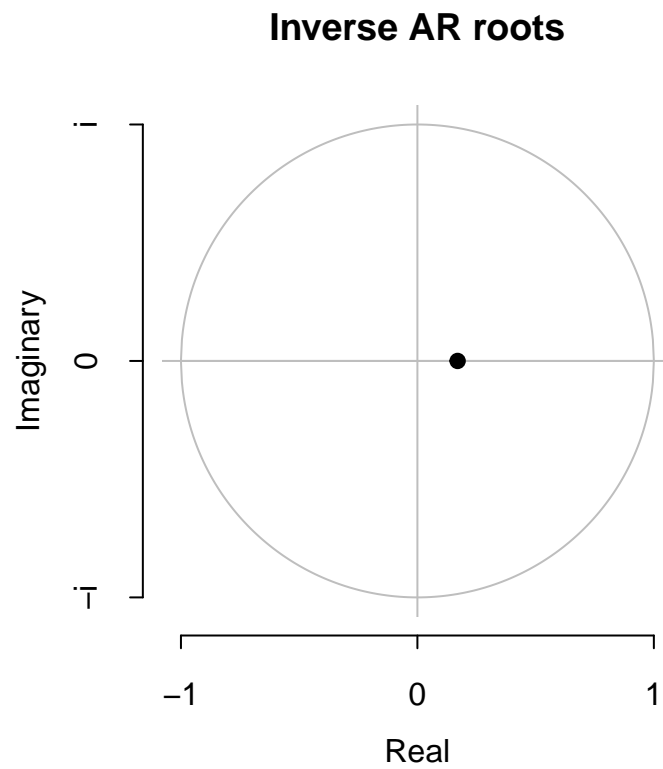
```
AR1var <- Arima((df$delta)^2, order = c(1, 0, 0))
summary(AR1var)
```

```
## Series: (df$delta)^2
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##          ar1      mean
##          0.1701  0.8238
## s.e.      0.0212  0.0582
##
## sigma^2 estimated as 5.026:  log likelihood=-4792.09
## AIC=9590.17   AICc=9590.18   BIC=9607.2
##
## Training set error measures:
##              ME      RMSE      MAE  MPE  MAPE      MASE      ACF1
## Training set -6.299133e-05 2.240784 0.9197335 -Inf  Inf  0.8595655 -0.01320252
```

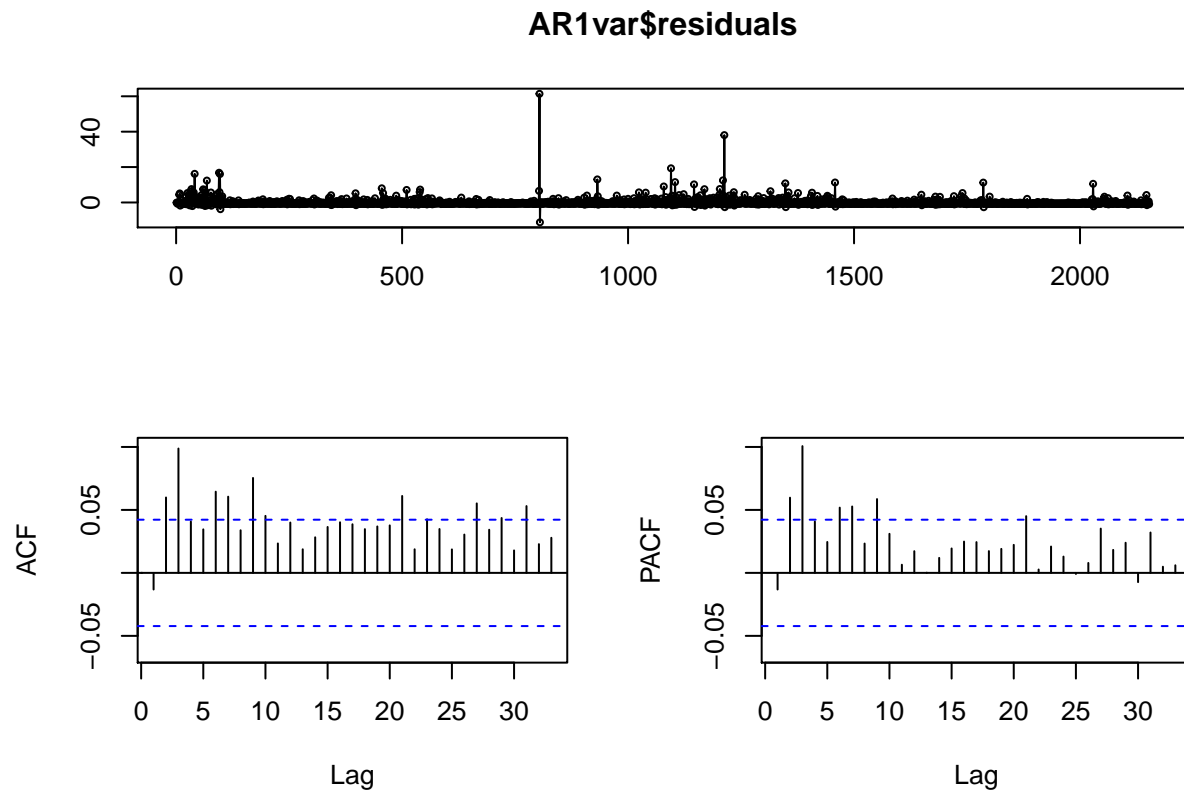
```
AR1var$coef[1]/sqrt(AR1var$var.coef[1, 1])
```

```
##      ar1  
## 8.011092
```

```
plot(AR1var)
```



```
tsdisplay(AR1var$residuals)
```

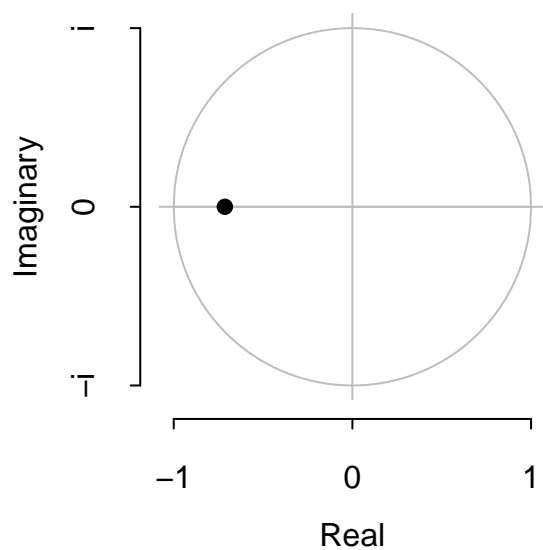
Now, let's run *auto.arima*.

```
aadelta <- auto.arima(df$delta, stepwise = F)
summary(aadelta)
```

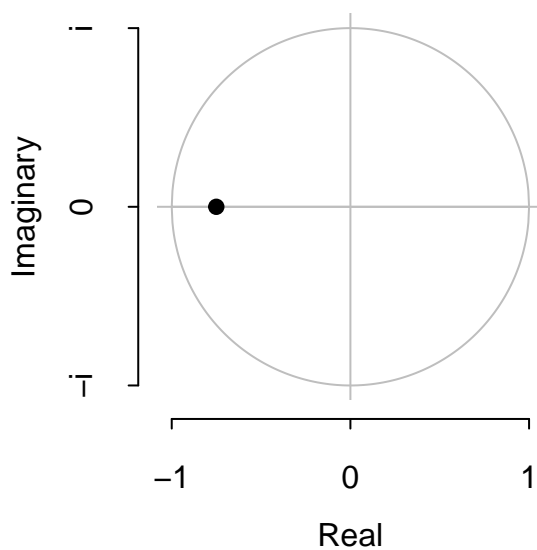
```
## Series: df$delta
## ARIMA(1,0,1) with non-zero mean
##
## Coefficients:
##      ar1      ma1      mean
##    -0.7138  0.7506  0.0433
## s.e.   0.1486  0.1399  0.0199
##
## sigma^2 estimated as 0.8208:  log likelihood=-2840.87
## AIC=5689.74   AICc=5689.76   BIC=5712.44
##
## Training set error measures:
##              ME      RMSE      MAE MPE MAPE  MASE      ACF1
## Training set 2.610156e-05 0.9053381 0.6434199 NaN  Inf  0.72527 0.003320553

plot(aadelta)
```

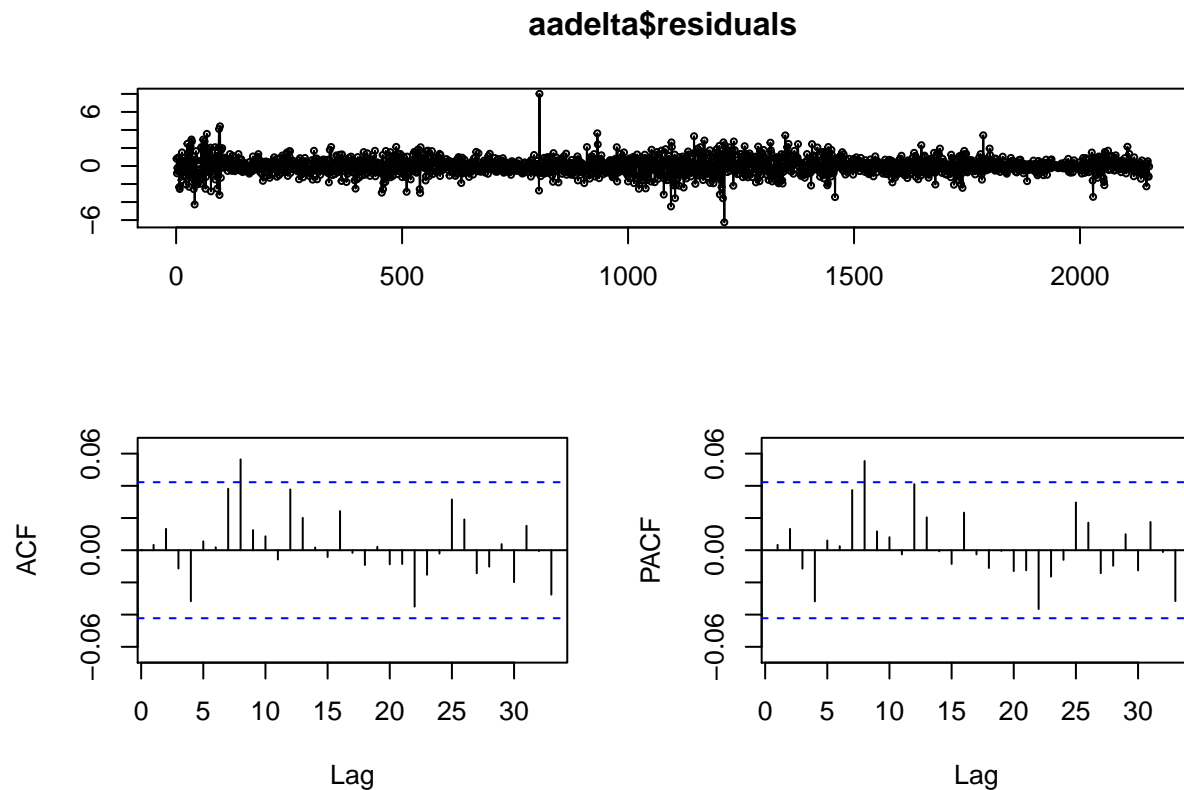
Inverse AR roots



Inverse MA roots



```
tsdisplay(aadelta$residuals)
```



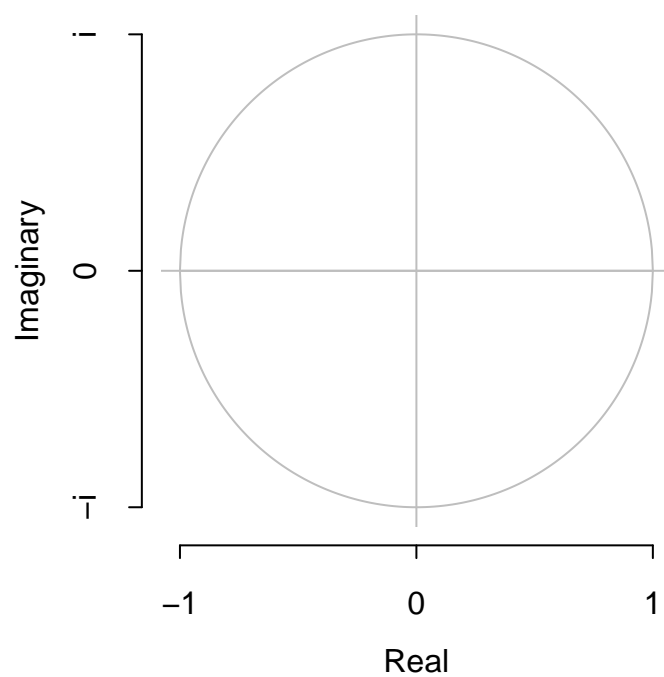
```
aassign <- auto.arima(df$sign, stepwise = F)
summary(aassign)
```

```
## Series: df$sign
## ARIMA(0,0,0) with non-zero mean
##
## Coefficients:
##      mean
##      0.5165
## s.e.  0.0108
##
## sigma^2 estimated as 0.2498:  log likelihood=-1561.46
## AIC=3126.91   AICc=3126.92   BIC=3138.26
##
## Training set error measures:
##              ME      RMSE      MAE  MPE MAPE      MASE      ACF1
## Training set -2.382602e-13 0.4997281 0.4994563 -Inf  Inf  1.028545 0.02773919
```

```
plot(aassign)
```

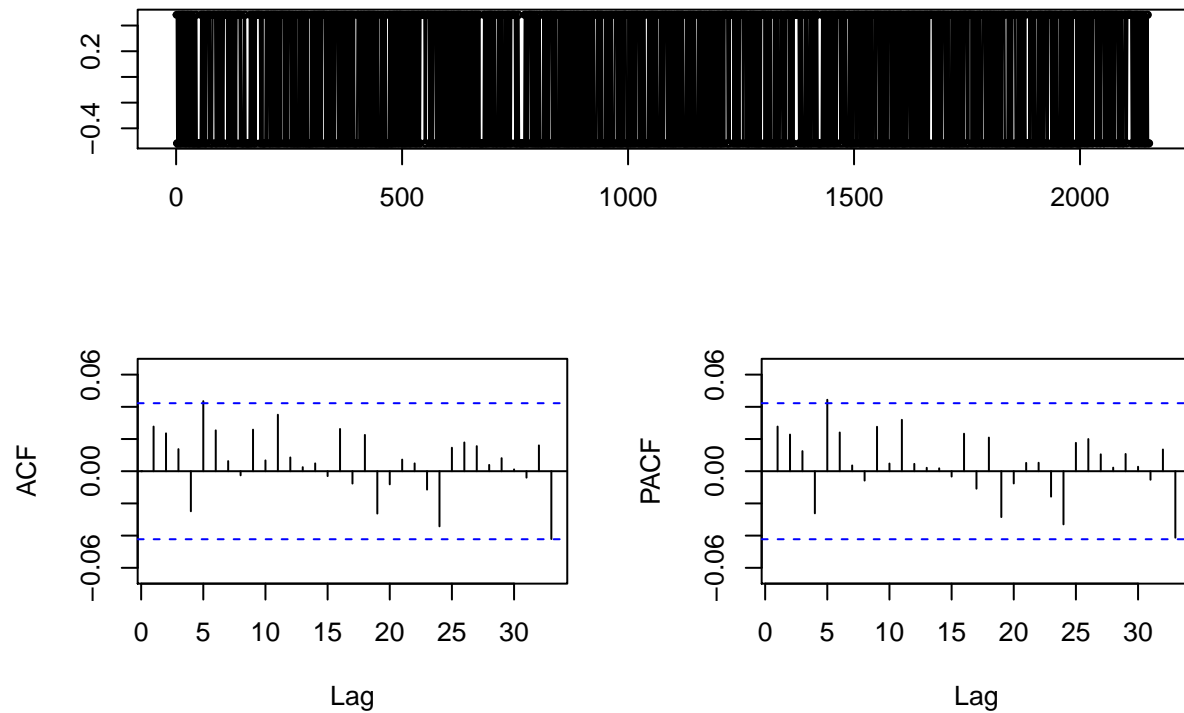
```
## Warning in plot.Arima(aassign): No roots to plot
```

No AR or MA roots



```
tsdisplay(aassign$residuals)
```

aassign\$residuals

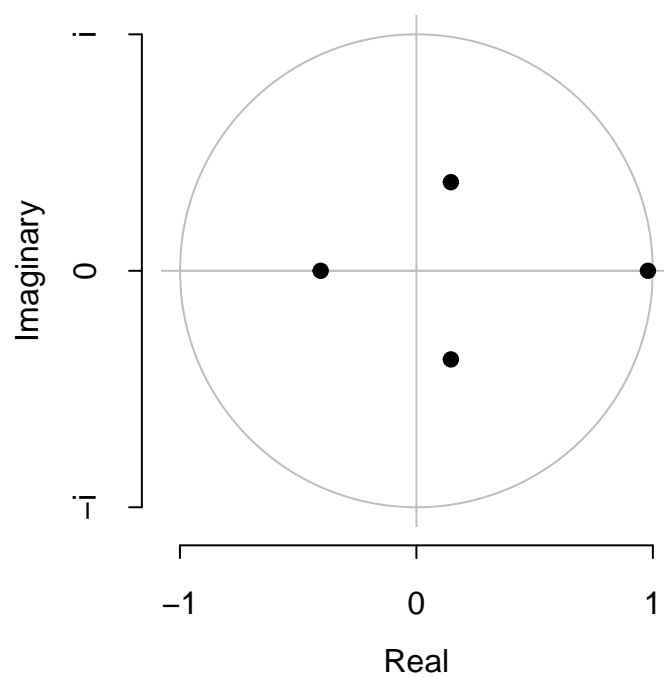


```
aavar <- auto.arima((df$delta)^2, stepwise = F)
summary(aavar)
```

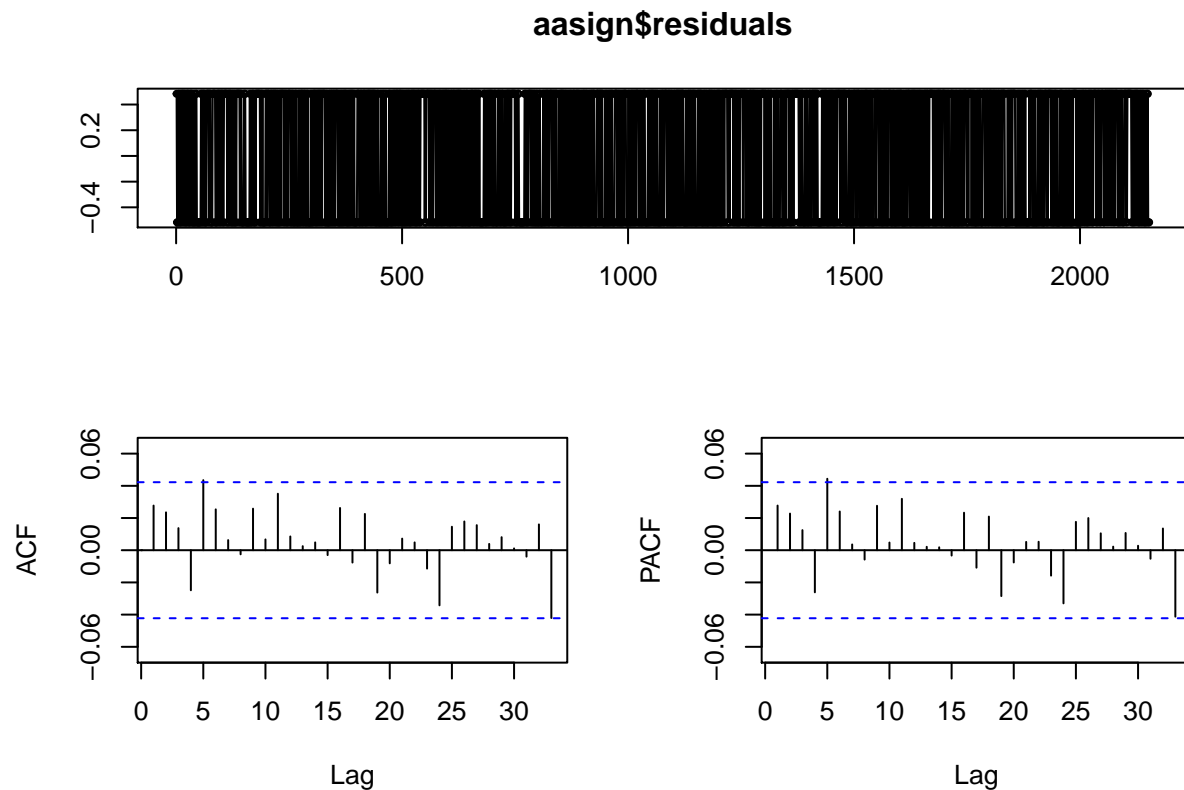
```
## Series: (df$delta)^2
## ARIMA(0,1,4)
##
## Coefficients:
##      ma1      ma2      ma3      ma4
##      -0.8662 -0.0671  0.0228 -0.0641
## s.e.    0.0215   0.0284  0.0280  0.0214
##
## sigma^2 estimated as 4.892:  log likelihood=-4761.22
## AIC=9532.45   AICc=9532.48   BIC=9560.82
##
## Training set error measures:
##              ME      RMSE      MAE  MPE  MAPE      MASE      ACF1
## Training set -0.02170361 2.209213 0.8904046 -Inf  Inf  0.8321553 0.0001189823

plot(aavar)
```

Inverse MA roots



```
tsdisplay(aassign$residuals)
```



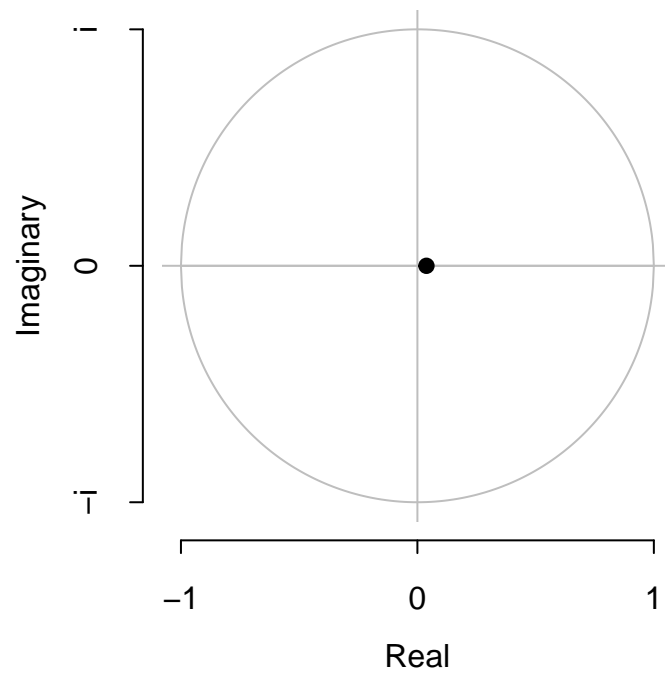
```
aardelta <- auto.arima(df$delta, max.q = 0, stepwise = F)
summary(aardelta)
```

```
## Series: df$delta
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##      ar1      mean
##      0.0382  0.0433
## s.e.  0.0215  0.0203
##
## sigma^2 estimated as 0.8215:  log likelihood=-2842.26
## AIC=5690.52   AICc=5690.53   BIC=5707.54
##
## Training set error measures:
```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
## Training set	-1.464203e-05	0.9059233	0.6434299	NaN	Inf	0.7252813	0.0004805619

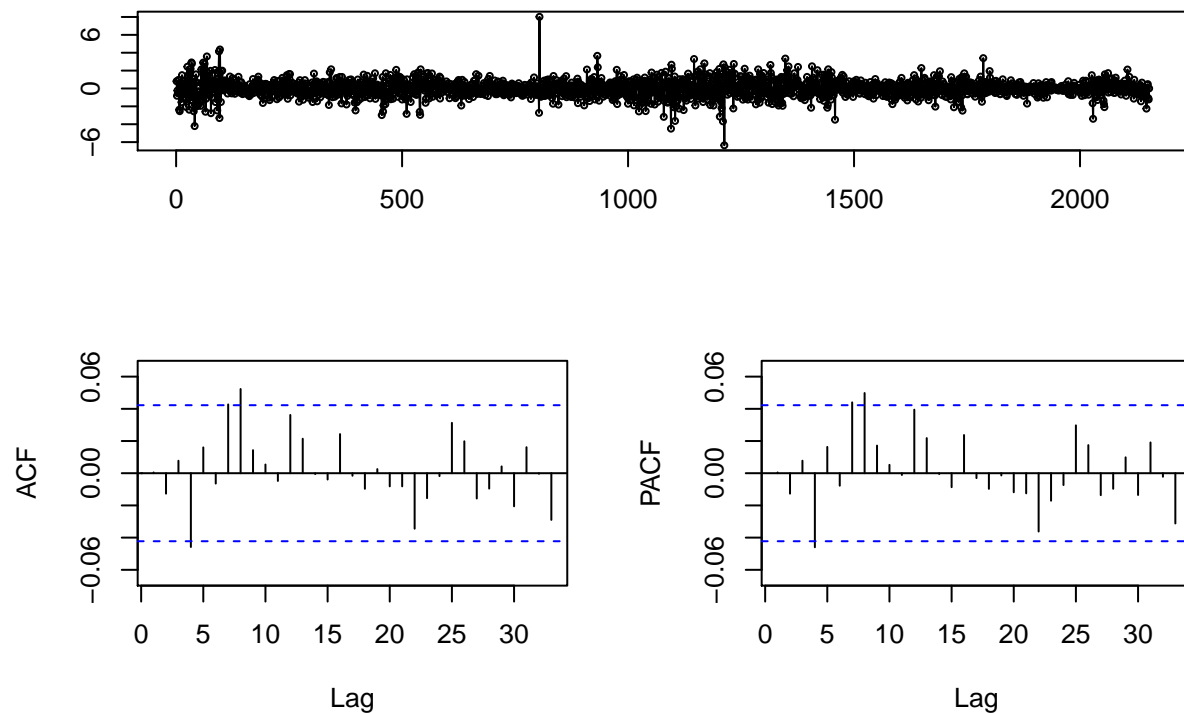
```
plot(aardelta)
```

Inverse AR roots



```
tsdisplay(aardelta$residuals)
```


aardelta\$residuals



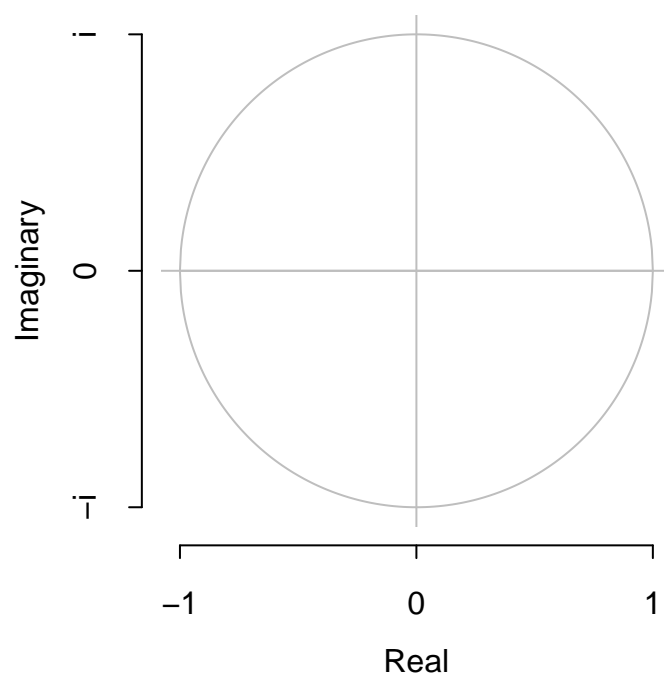
```
aarsign <- auto.arima(df$sign, max.q = 0, stepwise = F)
summary(aarsign)
```

```
## Series: df$sign
## ARIMA(0,0,0) with non-zero mean
##
## Coefficients:
##      mean
##      0.5165
## s.e.  0.0108
##
## sigma^2 estimated as 0.2498:  log likelihood=-1561.46
## AIC=3126.91   AICc=3126.92   BIC=3138.26
##
## Training set error measures:
##              ME      RMSE      MAE  MPE  MAPE      MASE      ACF1
## Training set -2.382602e-13 0.4997281 0.4994563 -Inf  Inf  1.028545 0.02773919
```

```
plot(aarsign)
```

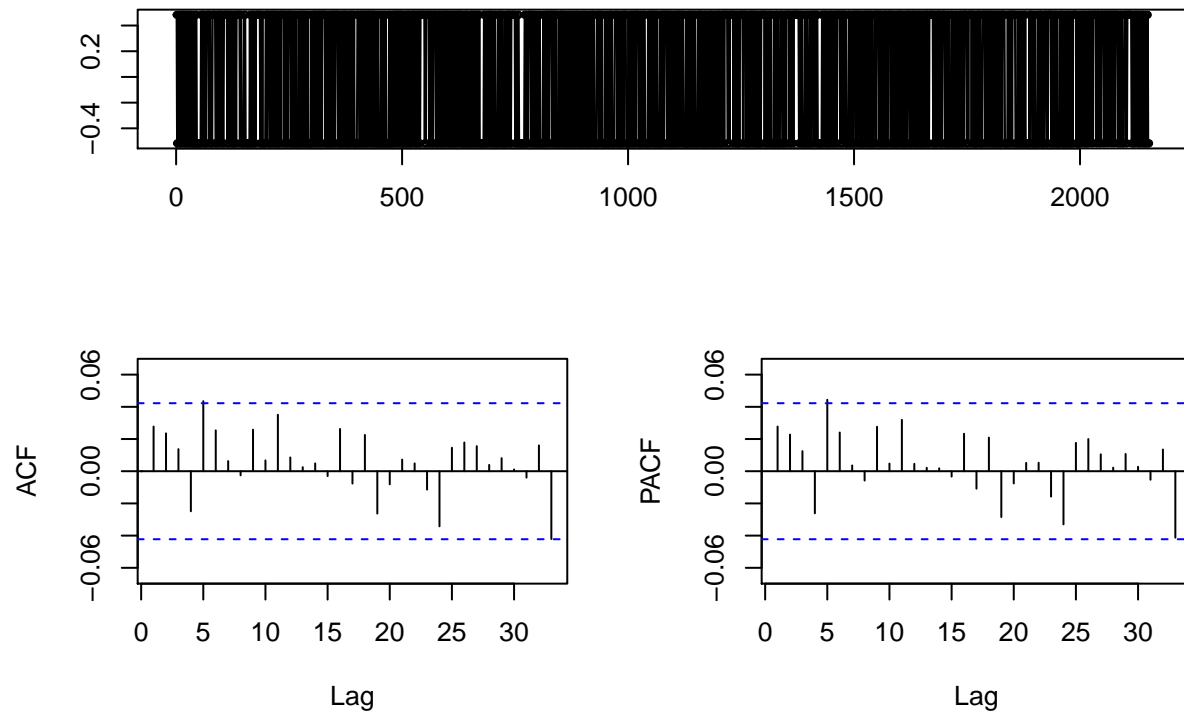
```
## Warning in plot.Arima(aarsign): No roots to plot
```

No AR or MA roots



```
tsdisplay(aarsign$residuals)
```

aarsign\$residuals

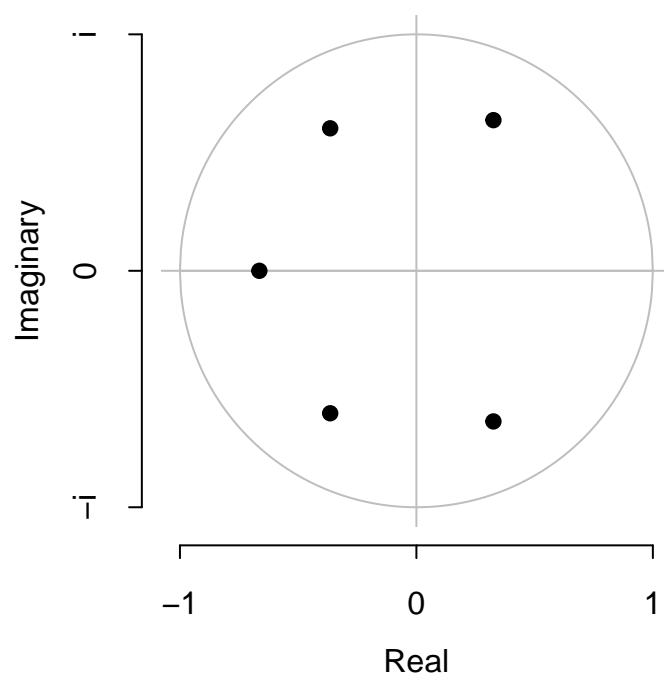


```
aarvar <- auto.arima((df$delta)^2, max.q = 0, stepwise = F)
summary(aarvar)
```

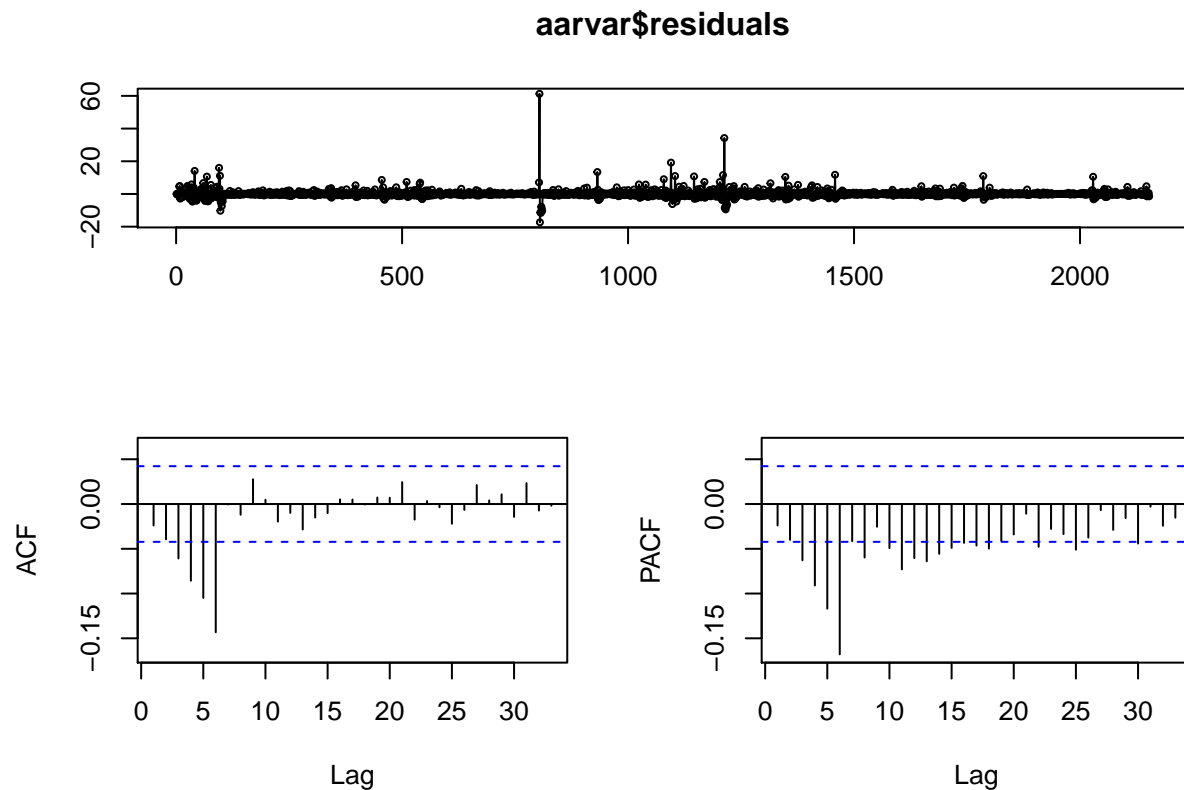
```
## Series: (df$delta)^2
## ARIMA(5,1,0)
##
## Coefficients:
##      ar1      ar2      ar3      ar4      ar5
##      -0.7420 -0.5845 -0.4042 -0.2873 -0.1687
## s.e.    0.0213  0.0259  0.0274  0.0258  0.0212
##
## sigma^2 estimated as 5.465:  log likelihood=-4878.95
## AIC=9769.89  AICc=9769.93  BIC=9803.94
##
## Training set error measures:
##              ME      RMSE      MAE  MPE  MAPE      MASE      ACF1
## Training set -3.234587e-05 2.334504 0.9170278 -Inf  Inf  0.8570369 -0.0239684

plot(aarvar)
```

Inverse AR roots



```
tsdisplay(aarvar$residuals)
```



$$\Delta_{t+1} = c + \beta \Delta_t + \varepsilon$$

```
AR1_2 <- Arima(df$delta, order = c(1, 0, 0))
```

```
summary(AR1_2)
```

```
## Series: df$delta
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##      ar1      mean
##      0.0382  0.0433
## s.e.  0.0215  0.0203
##
## sigma^2 estimated as 0.8215:  log likelihood=-2842.26
## AIC=5690.52   AICc=5690.53   BIC=5707.54
##
## Training set error measures:
##              ME      RMSE      MAE  MPE  MAPE      MASE      ACF1
## Training set -1.464203e-05 0.9059233 0.6434299 NaN   Inf  0.7252813 0.0004805619
```

```
confint(AR1_2, level = 0.95)
```

```
##              2.5 %      97.5 %
## ar1          -0.003988796 0.08042284
## intercept    0.003511958 0.08308440
```