White Noise and Random Walk

Son Nguyen

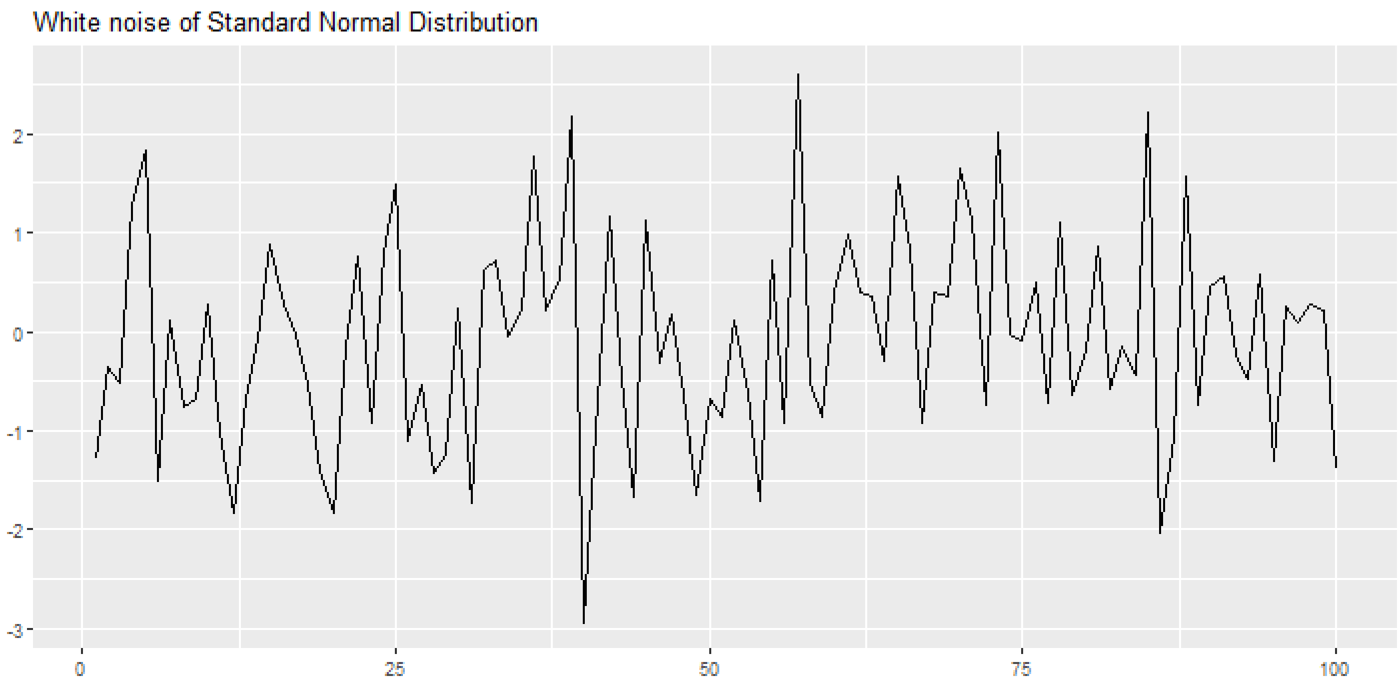
White Noise

- y_t is a white-noise process (series) if y_1, y_2,..., y_t are independent identical distributed (iid) zero mean random variables from a certain distribution (usually normal)

Example

noise from normal dist.

```
set.seed(30)
y ← ts(rnorm(100))
library(ggfortify)
autoplot(y) + ggtitle("White noise of Standard Normal Distribution")
```

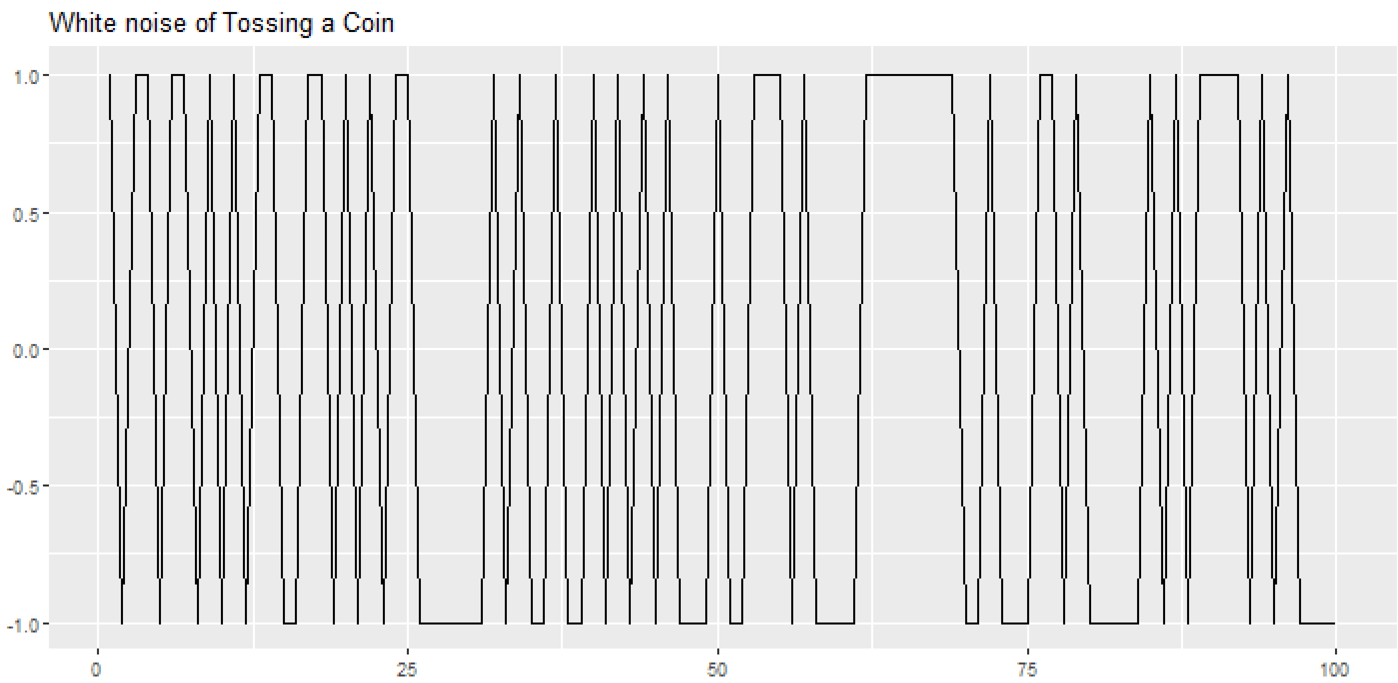


Example

a coin

tossing

```
set.seed(30)

y = sample(c(-1, 1), 100, replace = TRUE)

y <- ts(y)
library(ggfortify)
autoplot(y) + ggtitle("White noise of Tossing a Coin")
```



Correlogram

- Autocorrelation lag with lag k is the the correlation between the time series y_t and y_{t-k}

$$\rho_k = \frac{\sum_{t=k+1}^{n}(Y_t-\bar{Y})(Y_{t-k}-\bar{Y})}{\sum_{t=1}^{n}(Y_t-\bar{Y})^2}$$

where:

- ρ_k: Autocorrelation at lag k
- Y_t: Value of the series at time t
- $\bar{Y}$: Mean of the series
- n: Number of observations

correlation between X, Y measures the «linear» relation

auto correlation ≈ correlate between (Y_t, Y_{t-1})

lag 1

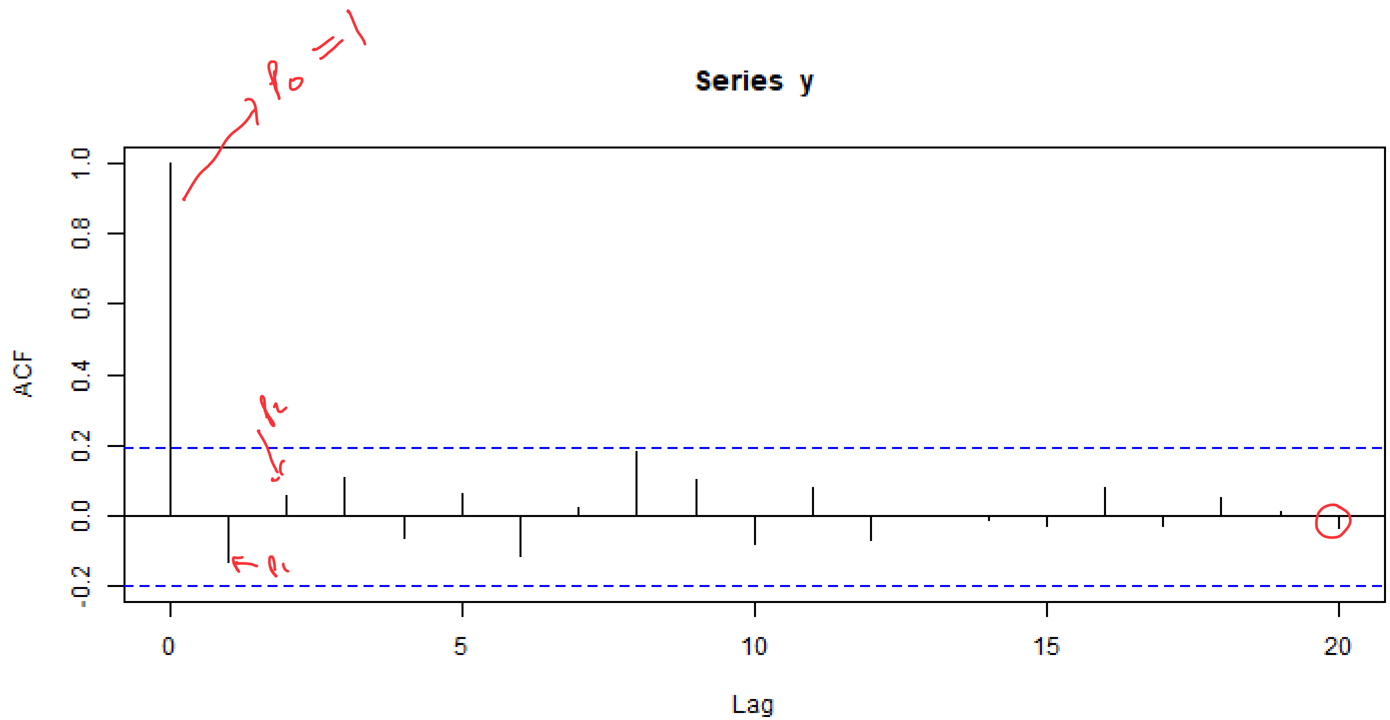
lag k : (Y_t, Y_{t-k})

- Autocorrelation lag with lag 0 is always 1
- The Correlogram is the plot of the autocorrelations for values of lag k = 0, 1, 2,...

Correlogram a white noise

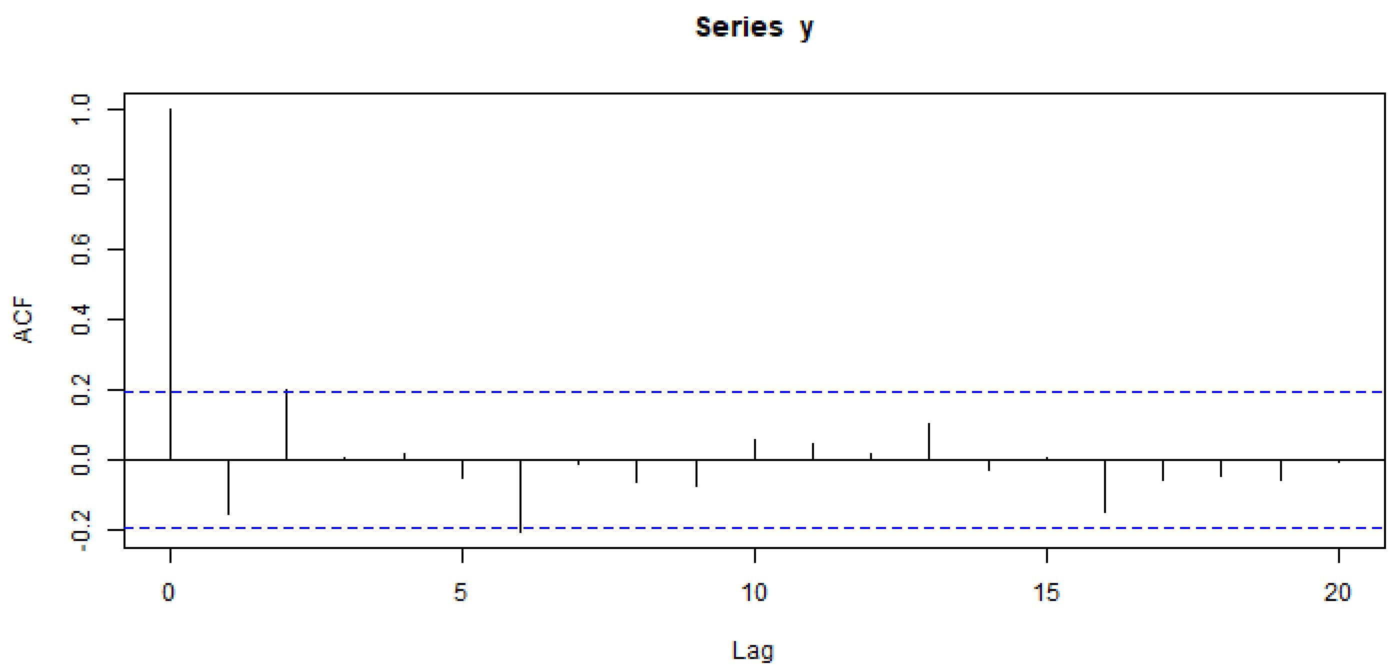
- Correlogram of a white noise

```r
# create a white-noise time series
y = ts(rnorm(100))

# plot its ACF or correlogram
acf(y)
```



Correlogram a white noise

```
set.seed(30)
y = sample(c(-1, 1), 100, replace = TRUE)

y ← ts(y)

acf(y)
```

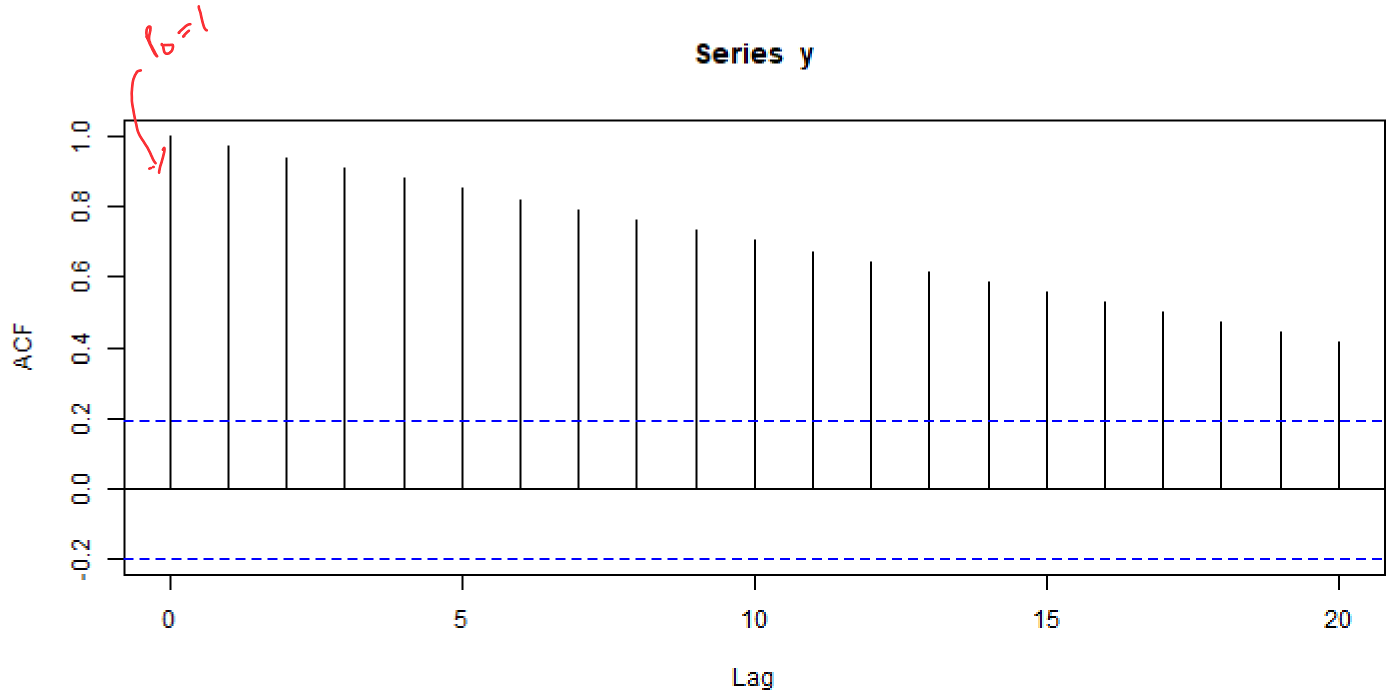


Correlogram a time series with trend

- Usually a trend in the data will show in the correlogram as a slow decay in the autocorrelation

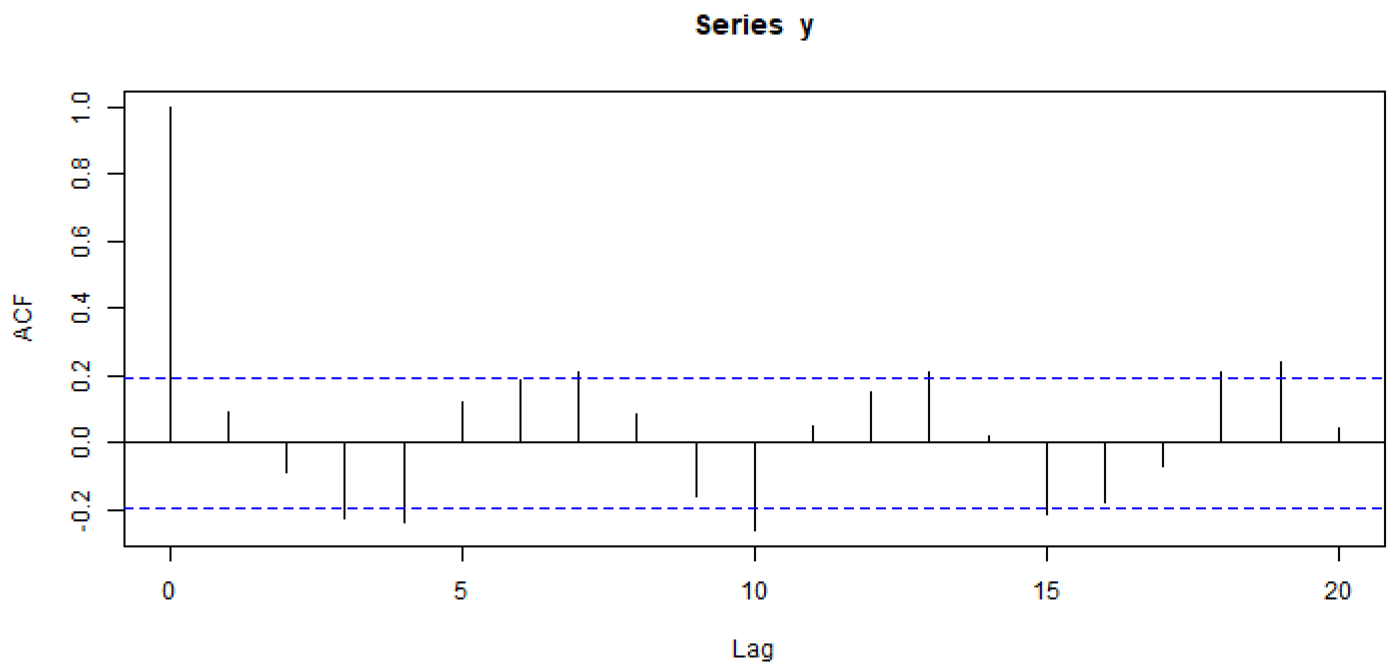
$y_t = \{1, 2, 3, \ldots 100\}$

```
y = ts(c(1:100))
acf(y)
```



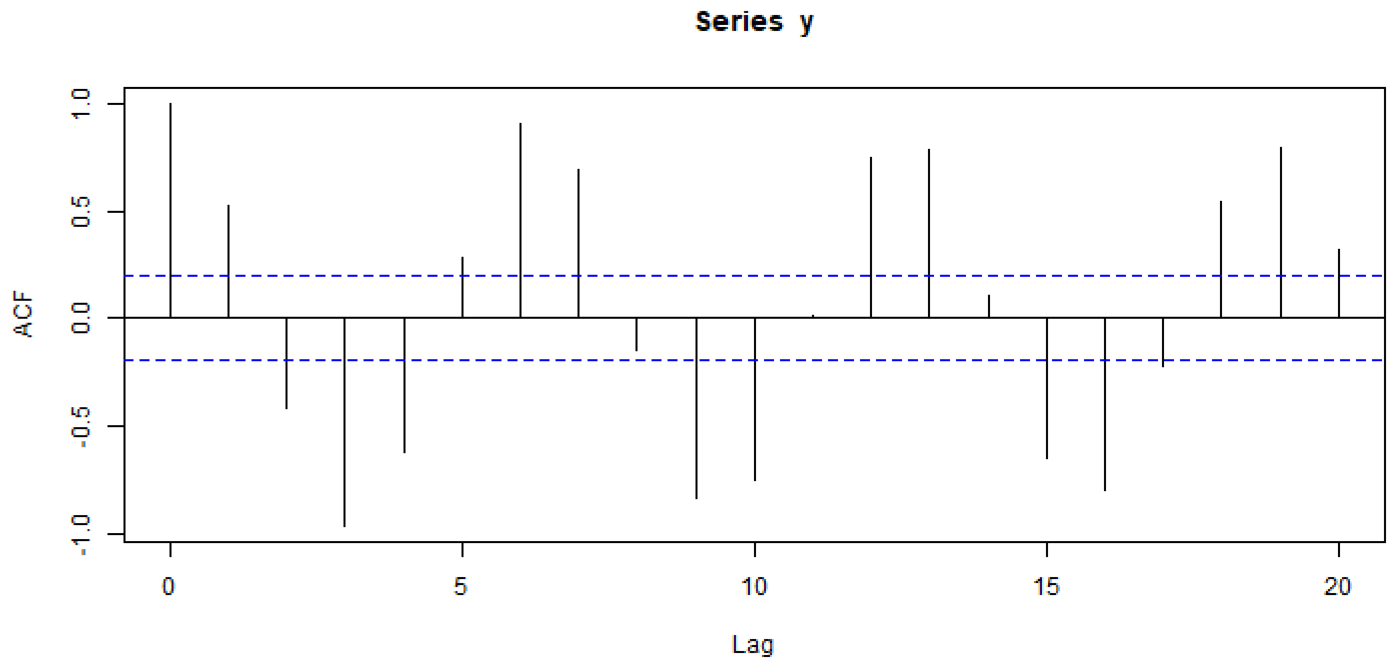
The Correlogram - Example

```
y = ts(cos(c(1:100))+rnorm(100))
acf(y)
```



ACF of a time series with seasonality

```
set.seed(30)
y = cos(1:100)
y <- ts(y)
acf(y)
```



Random Walk

- A time series y_t is called a random walk if

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

where ϵ_t is a white-noise

- A random walk can be written as

$$y_t = y_0 + \epsilon_1 + \epsilon_2 + \ldots + \epsilon_t$$

$$y_t = y_{t-1} + \varepsilon_t$$

$$= (y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

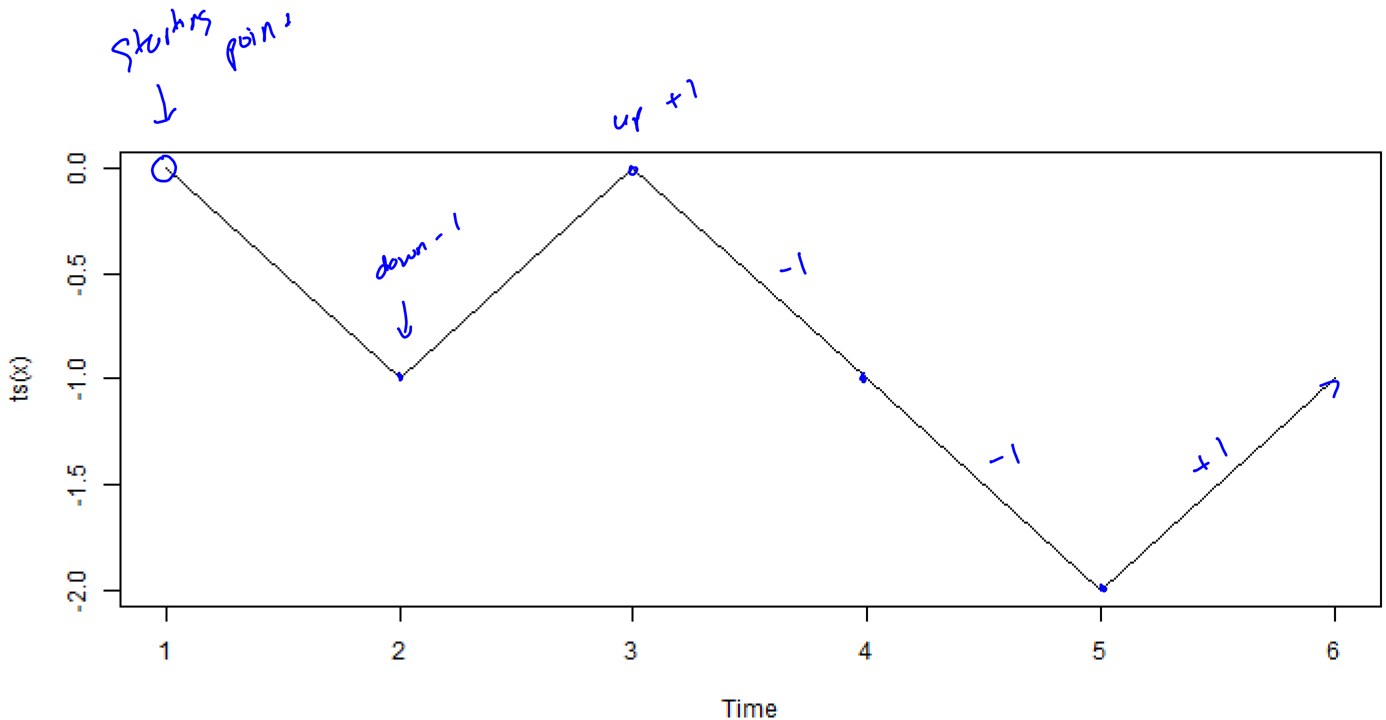
$$= y_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

$$\ldots$$

$$= y_0 + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t$$

Example

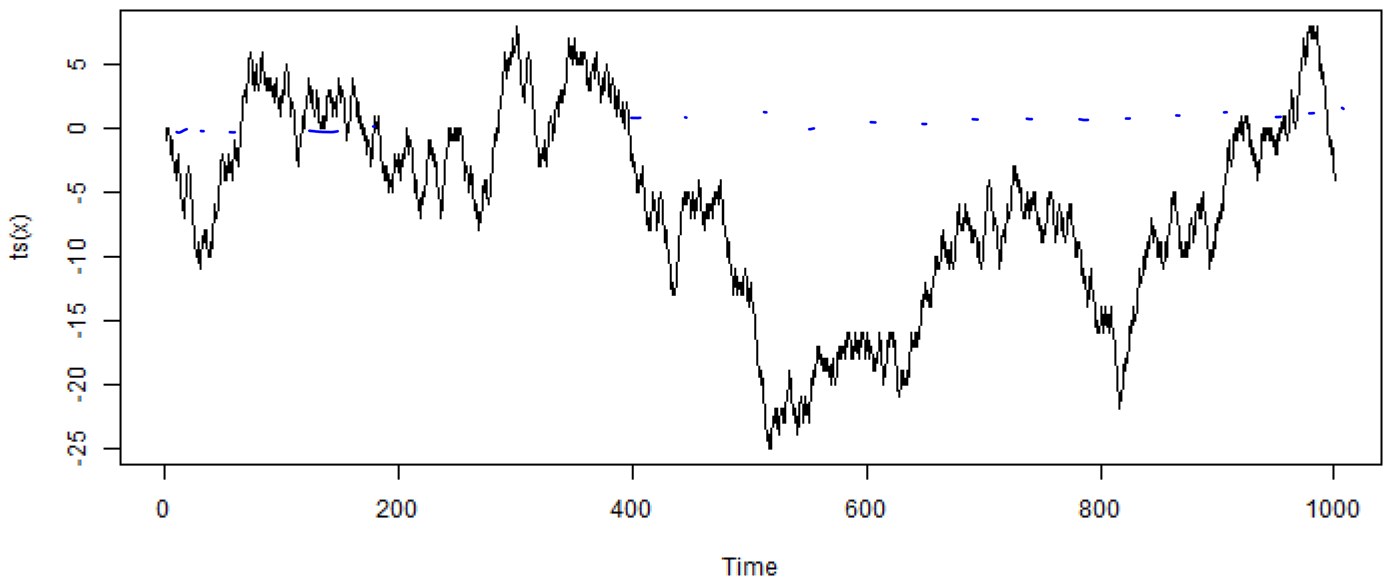
```
set.seed(1)
n ← 5
ct = sample(c(-1, 1), n, TRUE)
x ← cumsum(c(0,ct))
plot(ts(x))
```



Example

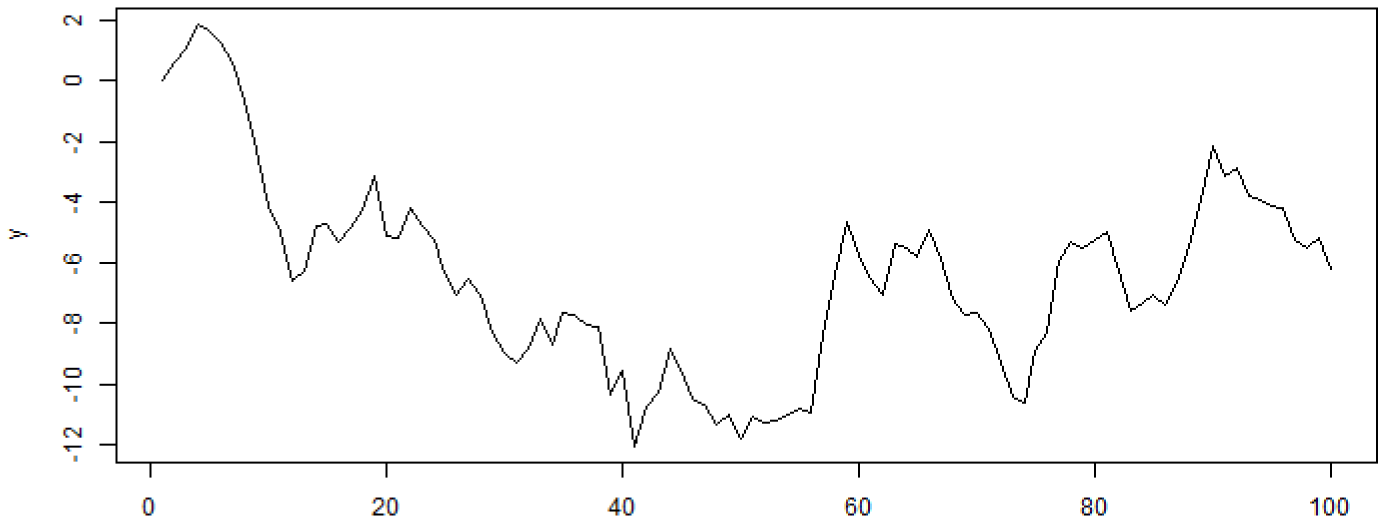
0 +1 -1

```
set.seed(1)
n ← 1000
ct = c(0, sample(c(-1, 1), n, TRUE))
x ← cumsum(ct)
plot(ts(x))
```



Example

```
set.seed(3000)
n = 100
c ← rnorm(n)
y_0 = 0
y = c(y_0, 2:n)
for (i in 2:n)
{
  y[i] = y[i-1]+c[i]
}
y = ts(y)
plot(y)
```



Random Walk with drift

- A time series y_t is called a random walk if with drift

$$y_t = y_{t-1} + d + \epsilon_t,$$

where ϵ_t is a white-noise

- A random walk can be written as

$$y_t = y_0 + dt + \epsilon_1 + \epsilon_2 + \ldots + \epsilon_t$$

$$y_t = y_{t-1} + d + \varepsilon_t$$

$$= y_{t-2} + d + \varepsilon_{t-1} + d + \varepsilon_t$$

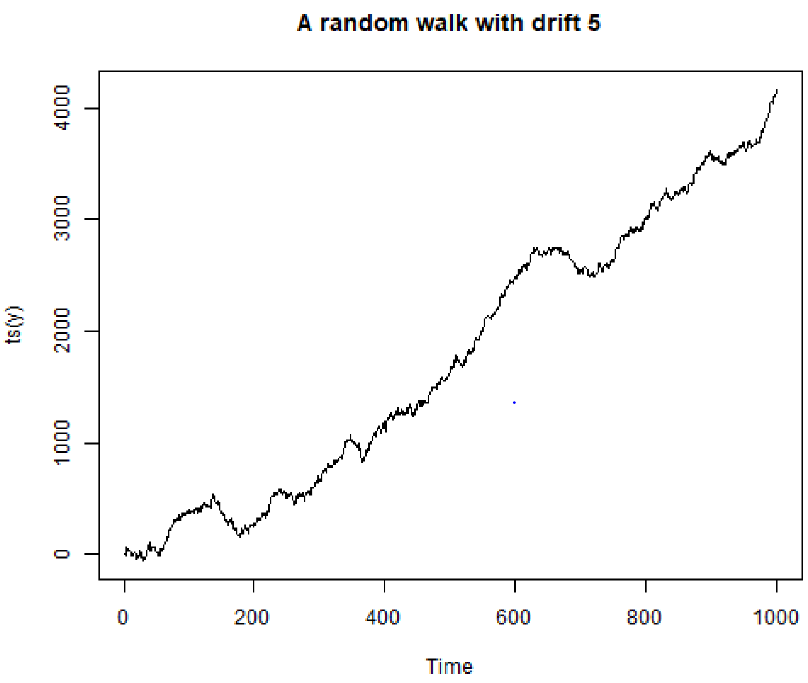
$$= y_{t-2} + 2d + \varepsilon_{t-1} + \varepsilon_t$$

$$\cdot \ \cdot$$

$$=$$

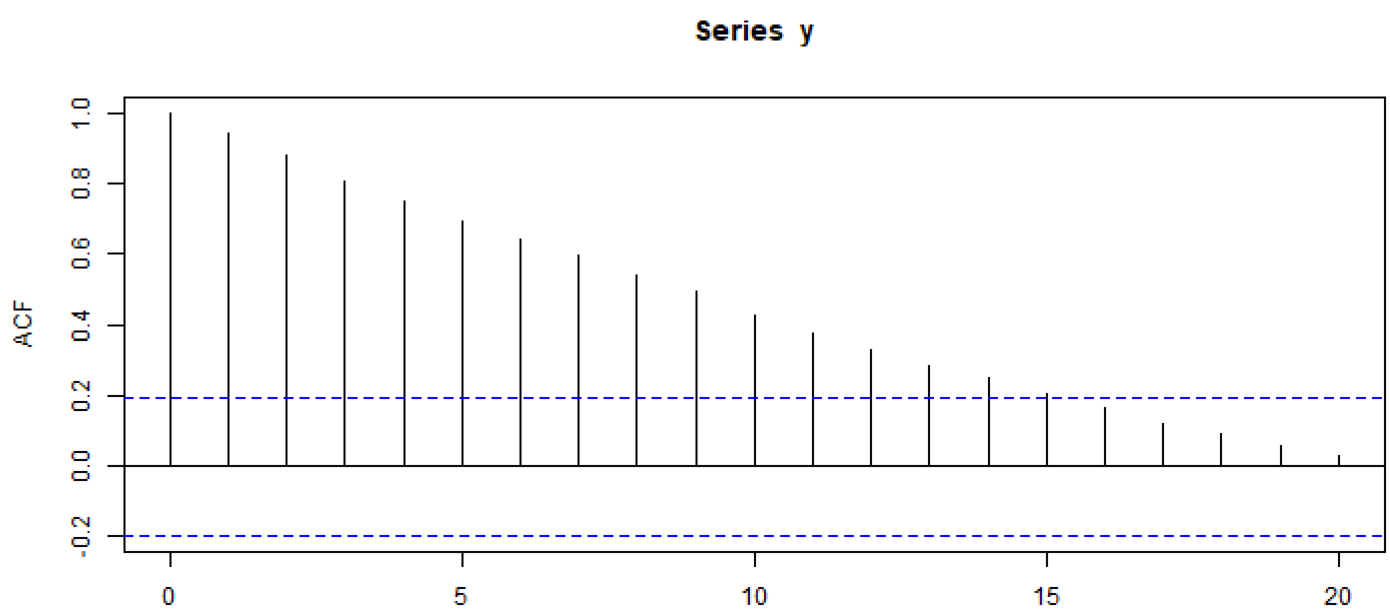
Example

```r
set.seed(30)
n = 1000
c <- rnorm(n, sd = 20)
y_0 = 0
drift = 5

y = c(y_0, 2:n)

for (i in 2:n)
{
  y[i] = drift  + y[i-1]+c[i]
}

library(ggfortify)
library(latex2exp)

plot(ts(y))
title(paste0("A random walk with drift ", drift))
```



The ACF of Random Walks

```
n = 100
error_mean = 0
c ← rnorm(n, mean = error_mean, sd = 30)
y_0 = 0
y = c(y_0, 2:n)

for (i in 2:n)
{
  y[i] = y[i-1]+c[i]
}

acf(y)
```



$\{y_t\}$: time series

d_t : differenced series

$d_t = y_{t+1} - y_t$

$$\underbrace{y_1 \quad y_2}_{d_1} \quad \underbrace{\quad y_3}_{d_2} \quad \cdots \quad \underbrace{y_t \quad y_{t+1}}_{d_t}$$

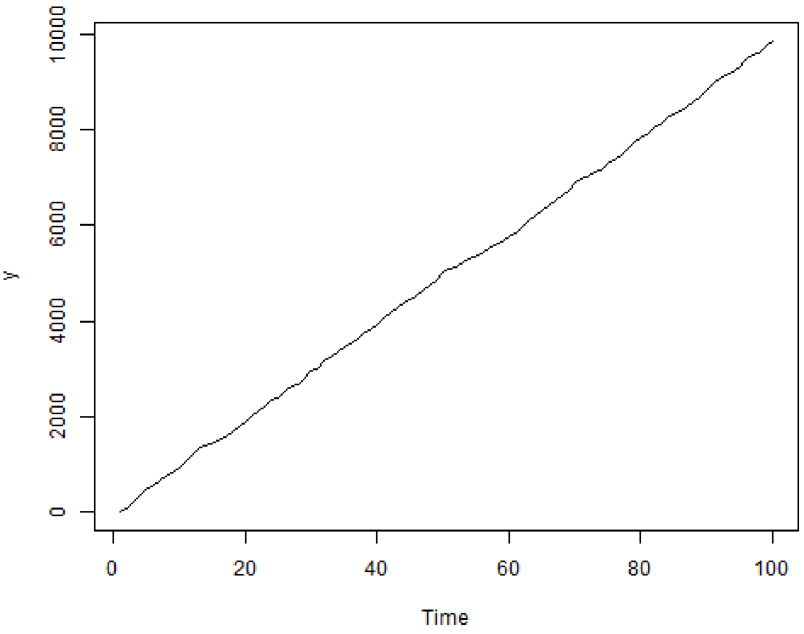
If y_t is a random walk:

$$y_{t+1} = y_t + \varepsilon_{t+1}$$

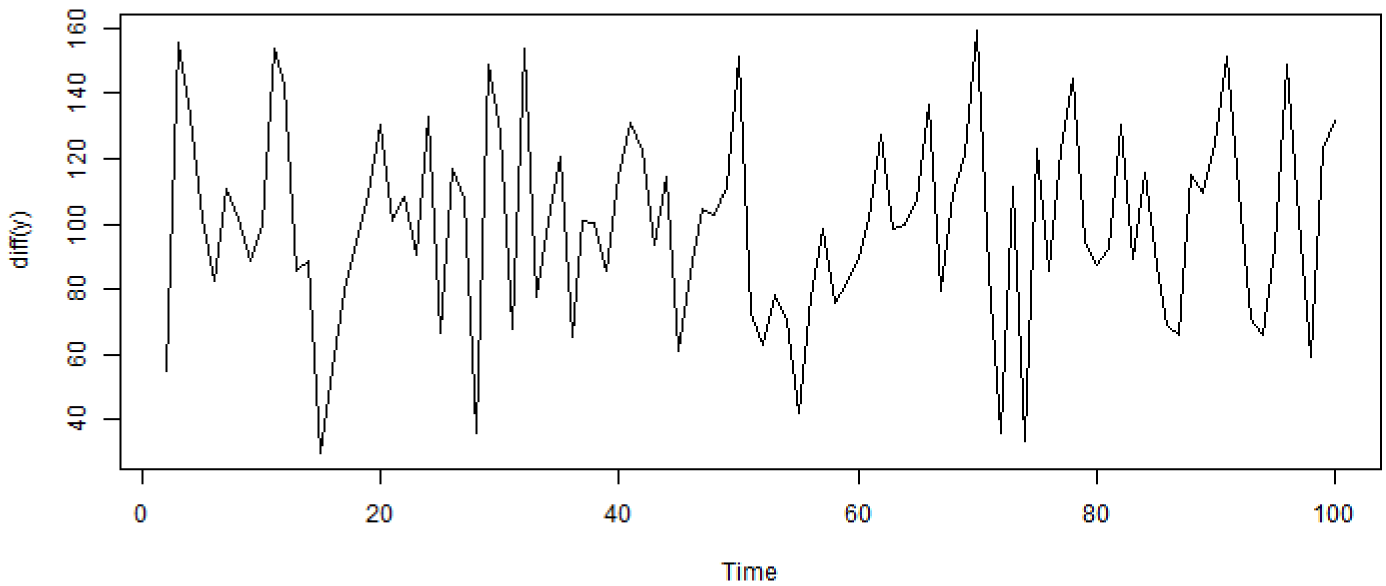
$$\Rightarrow \underbrace{y_{t+1} - y_t}_{d_t} = \varepsilon_{t+1} \Rightarrow \boxed{d_t} = \varepsilon_{t+1} \text{ or white noise}$$

Differencing Time Series

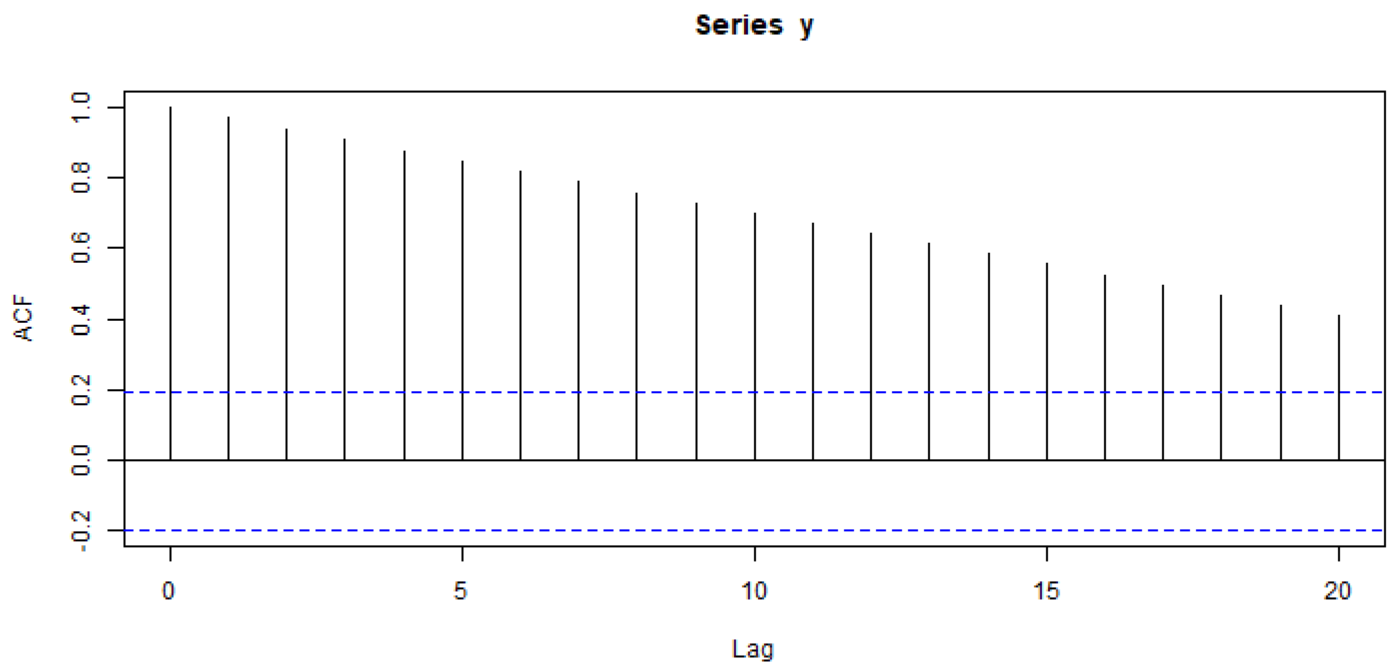
```
n = 100
error_mean = 0
drift = 100
c ← rnorm(n, mean = error_mean, sd = 30)
y_0 = 0
y = c(y_0, 2:n)

for (i in 2:n)
{
  y[i] = drift + y[i-1]+c[i]
}

y = ts(y)

plot(y)
```



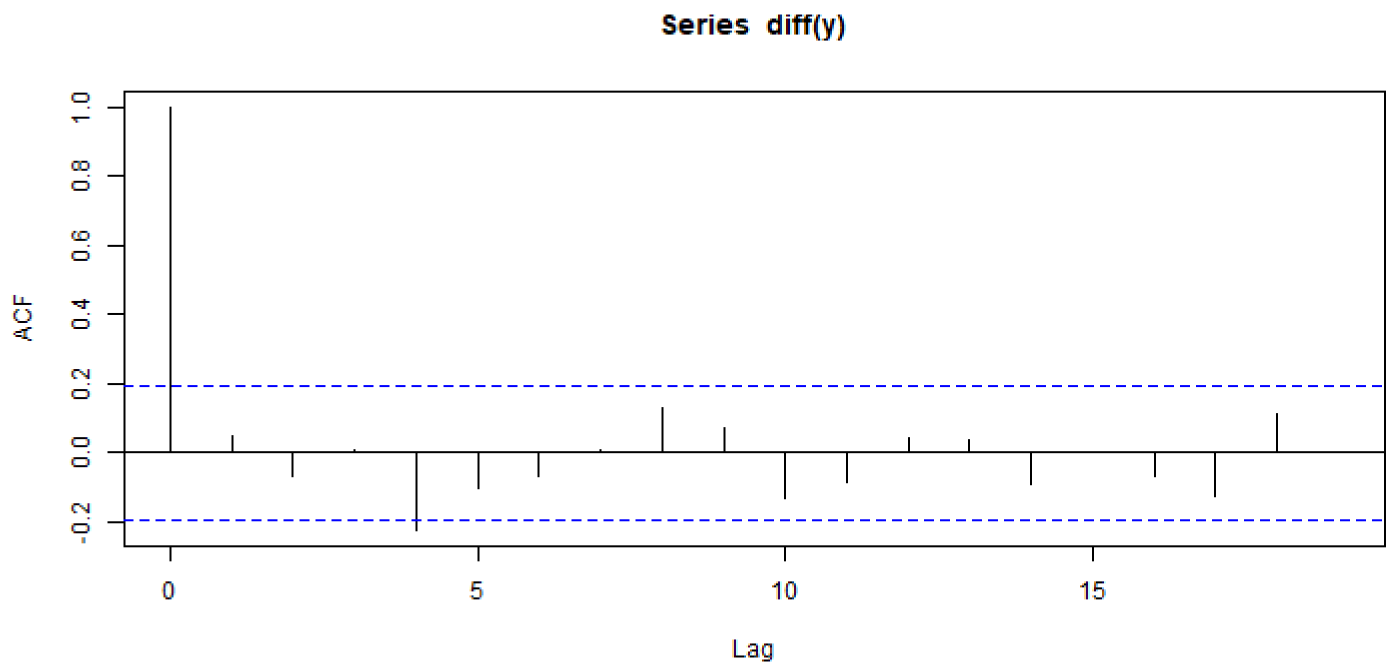
```
plot(diff(y))
```



```
acf(y)
```



```
acf(diff(y))
```



Quantmod Package

https://www.quantmod.com/

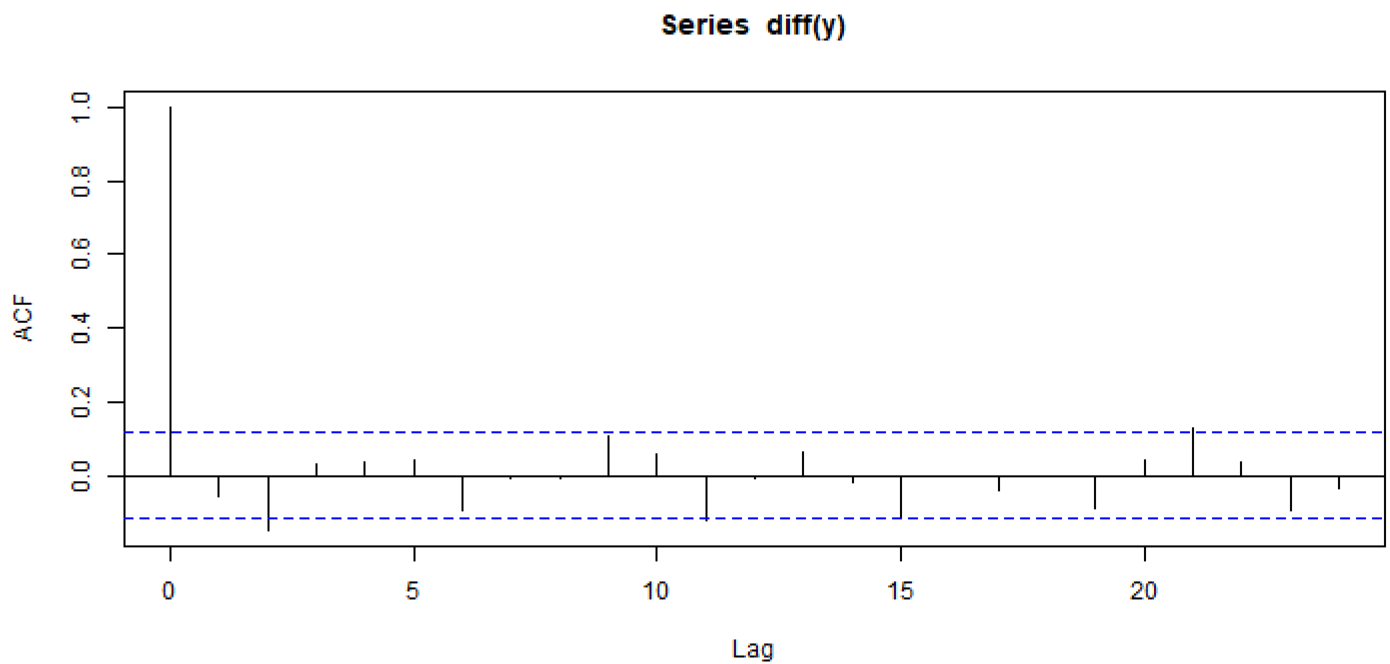
Random Walks and Stocks

```r
library(quantmod)
getSymbols('MSFT', src='yahoo')
```

```
## [1] "MSFT"
```

```r
y = Ad(MSFT[index(MSFT)>"2023-01-01",])

plot(y)
```

```
acf(diff(y), na.action = na.pass)
```



- The differencing series could be a white noise
- It is very reasonable to assume that the stock follows the random walk model.