Financial Risk Management Project

Valentina Righetti (VR 476263)
Letizia Bottacin (VR 473127)
Cesare Cottonaro (VR 464643)

January 15, 2023

1 Exercise 1.1

Since we performed the assigned exercises in R studio, we created the *csv file* that contains all the closing prices removing the prices that are simply repetition of previous day prices.

We remove the double data in excel using the formula $= if(b3 = b2; ""; b3)$, so if the price of that day is equal to the day before price, we will substitute it with an empty cell.

We filtered then column 1 excluding empty cells and copied the remaining data in our csv data source.

H	I
Raw Data	
Date	Close
01-gen-01	=SE(B3=B2;"";B3
02-gen-01	1785,86
03-gen-01	1875,56
04-gen-01	1855,82
05-gen-01	1807,13
08-gen-01	1804,04
09-gen-01	1810,98
10-gen-01	1828,50
11-gen-01	1847,37
12-gen-01	1835,53
15-gen-01	
16-gen-01	1847,13
17-gen-01	1851,16
18-gen-01	1876,94
19-gen-01	1869,38
22-gen-01	1869,89
23-gen-01	1894,25
24-gen-01	1899,71
25-gen-01	1890,27
26-gen-01	1886,80
29-gen-01	1899,75
30-gen-01	1913,13
31-gen-01	1902,55
01-feb-01	1913,11
02-feb-01	1879,69

Figure 1

Using R we compute the daily log return (natural logarithm of the ratio between the price at time t and the price at time t-1) and create a vector with these data. Finally, we plotted both the returns and ending prices.

```
#### Exercise 1.1 ####

data = read.csv('FinancialRisk.csv', sep = ';', header = T)
data = na.omit(data)
data$Date = as.Date(data$Date, format = '%d-%b-%y')

data$logClose = log(data$Close, exp(1))
logClose = data$logClose
logReturn = diff(logClose)
logReturn = c(NaN, logReturn)
data$logReturn = logReturn

ggplot(data, aes(x = Date, y = Close)) +
  geom_line(color ='#119dec') +
  ggtitle('Close Prices S&P 500')

ggplot(data, aes(x = Date, y = logReturn)) +
  geom_line(color = '#119dec') +
  ggtitle('LogReturns of S&P 500')
```

Figure 2

Figure 3: Time series of S&P 500 close price

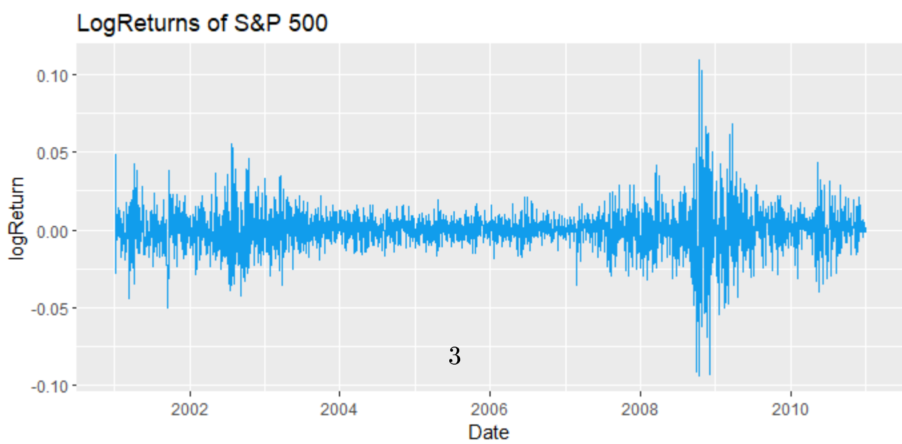

Figure 4: Time series of Log-Return S&P 500 close price

2 Exercise 1.3

We then compute the first 100th lags auto-correlation, plotting it with 2 different plots(the syntax is reported below).
We use the function $ggAcf()$ that automatically compute the auto correlation using the returns that we have previously computed and the function $plot()$ to have a more clear plot (the red lines are the acceptance intervals at $\alpha = 0.05$)

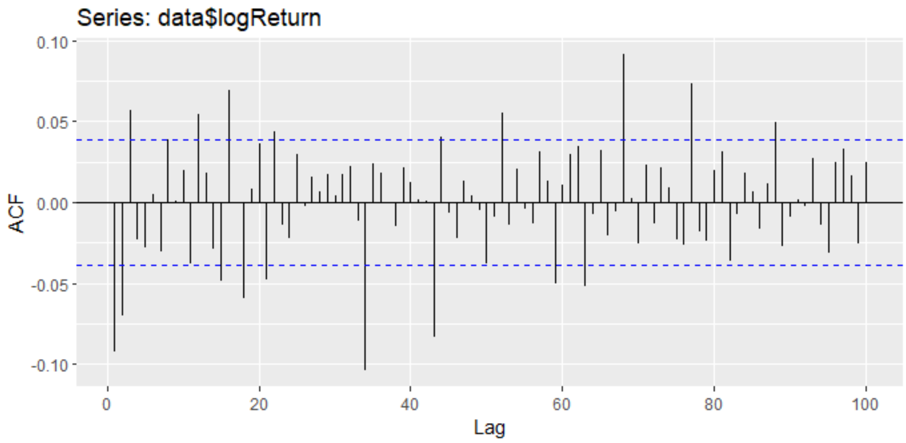

Figure 5

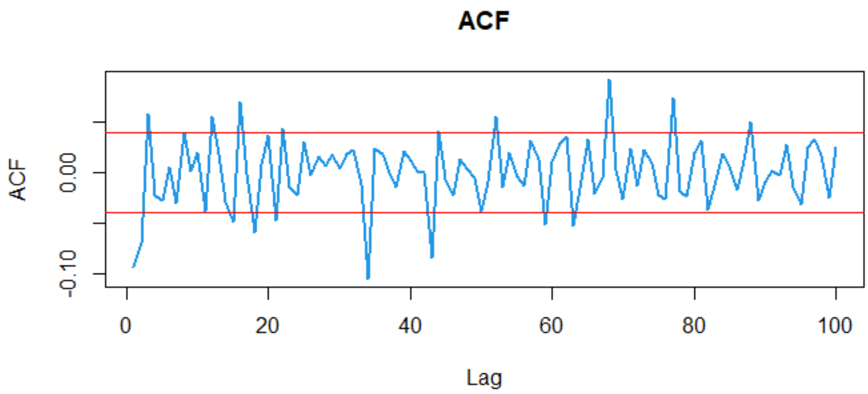

Figure 6

```
#### Exercise 1.3 ####

Correlation = c()

x = ggAcf(data$logReturn, plot = F, lag.max = 100)
names(x)
length(x$acf)
Correlation = x$acf[-1]
length(Correlation) #different values because R calculates the correct correlation coefficient with T-1 degrees of freedom
ggAcf(data$logReturn, lag.max = 100)

plot(Correlation, type = 'l', col = '4', lwd = '2.5', xlab = 'Lag', ylab = 'ACF', main = 'ACF')
```

Figure 7

3 Exercise 1.5

We firstly set the variance σ_0 as the variance of all the returns.
We than compute the variance for every t from 2 to 2514 and insert it in a vector whose length is 2514.
In order to do that, we create a *for cycle*, that goes from 3 ($t+1 = 2+1$ in the first lap) to T = 2514 and for every lap we compute the T+1 variance.

$$\sigma_{t+1}^2 + 0.94\sigma_t^2 + 0.06R_t^2 \ \ for\ t = 2, 3, \ldots, T$$

Finally, we compute the standard deviation by taking the square root of the vector and plot it. We plotted 2 types of graphs, one with only the Conditional Standard Deviation alone, other with Conditional Standard Deviation against the absolute values of the returns:

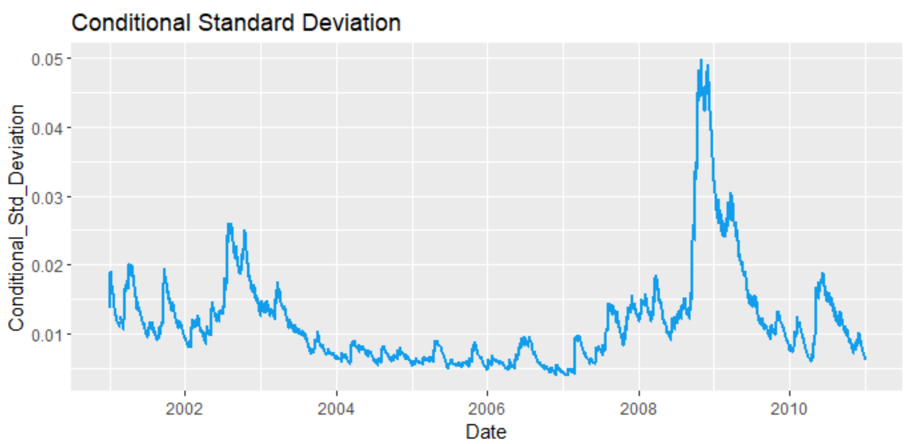

Figure 8: Conditional Standard Deviation calculated via Risk Metrics.

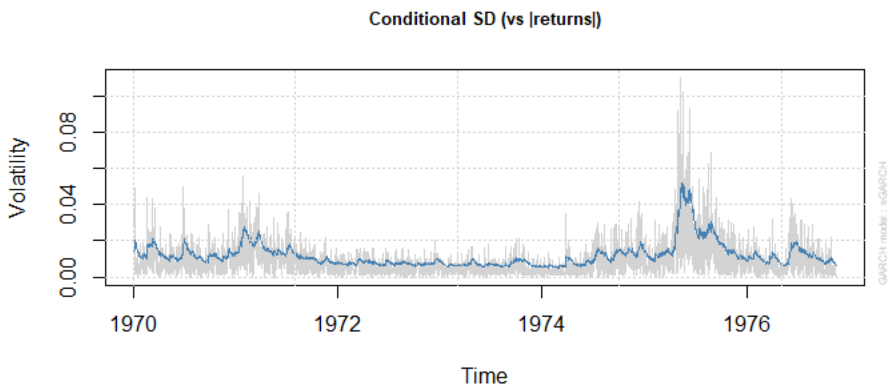

Figure 9: Conditional Standard Deviation calculated estimating a GARCH(1,1).

```
#### Exercise 1.5 ####

sigma0 = var(data$logReturn, na.rm = T)
lambda = 0.94
data$Conditional_Variance = NaN
data$Conditional_Variance[2] = sigma0
i = 1:1524

for (i in 3:2514){
  data$Conditional_Variance[i] = lambda * data$Conditional_Variance[i-1] + (1- lambda) * (data$logReturn[i-1])^2
}

data$Conditional_Std_Deviation = sqrt(data$Conditional_Variance)

ggplot(data, aes(x = Date, y = Conditional_Std_Deviation)) +
  geom_line(color = '#119dec', linetype = 1, linewidth = 0.8) +
  ggtitle('Conditional Standard Deviation')
```

Figure 10

4 Exercise 2.1

In order to computer the 1-day, 1% VaR on each day of October we use an historical simulation with a 250 days moving window.
Suppose that we are long 1$ of the S&P index each day, in order to predict a VAR of tomorrow return we perform an historical simulation making the following steps:

1. Define the returns of the past year by taking the $ln(\frac{P_t}{P_{t-1}})$ and the losses by taking the negative of the returns

Date	Close	Return	Loss
03-ott-86	233,71		
06-ott-86	234,78	0,46%	=-C4

Figure 11

2. On each day of October, we define the negative of the percentile of the 249 (250-1, N.of observations minus 1) days with a 99% confidence level.

$$VaR_t = -r_{(pm)}$$

   E.g. for the first day of October 87, we define the percentile of the distribution of the returns from 06-ott-86 to 30-sett-87, with 1% confidence level. In excel we use the formula $= -percentile(matrice, 0.01)$

3. We perform, these 2 steps for each day of October 87.

28-set-87	323,20	0,95%				
29-set-87	321,69	-0,47%				
30-set-87	321,83	0,04%				
01-ott-87	327,33	1,69%			=-PERCENTILE(C4:C253;J5)	
02-ott-87	328,07	0,23%			2,37%	
05-ott-87	328,08	0,00%			2,37%	
06-ott-87	319,22	-2,74%			2,37%	
07-ott-87	318,54	-0,21%			2,43%	
08-ott-87	314,16	-1,38%			2,43%	
09-ott-87	311,07	-0,99%			2,43%	
12-ott-87	309,39	-0,54%			2,43%	
13-ott-87	314,52	1,64%			2,43%	
14-ott-87	305,23	-3,00%			2,43%	
15-ott-87	298,08	-2,37%			2,58%	
16-ott-87	282,70	-5,30%			2,58%	
19-ott-87	224,84	-22,90%			2,71%	
20-ott-87	236,83	5,20%			2,87%	
21-ott-87	258,38	8,71%			2,87%	
22-ott-87	248,25	-4,00%			2,87%	
23-ott-87	248,22	-0,01%			3,51%	
26-ott-87	227,67	-8,64%			3,51%	
27-ott-87	233,19	2,40%			4,66%	
28-ott-87	233,28	0,04%			4,66%	
29-ott-87	244,77	4,81%			4,66%	
30-ott-87	251,79	2,83%			4,66%	

Figure 12

Finally we plot both losses and VaR for October 87.

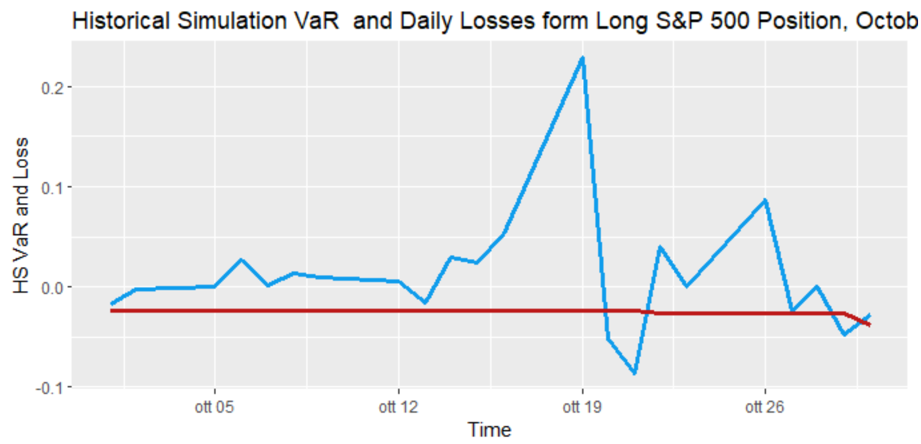

Figure 13

Suppose then we are short 1$ of the S&P index each day, in order to predict a VaR of tomorrow return we perform an historical simulation making the same steps as before with only 2 changes:

1. For the first day of October we define the negative percentile of the returns of the 249 (250-1, N.of observation minus 1) days with a 99% confidence level

2. For all the other days of October we define the negative percentile of the losses of the 249 (250-1, N.of observation minus 1) days with a 99% confidence level

01-ott-87	327,33	1,69%	-1,69%	=-PERCENTILE(D4:D253;J4)
02-ott-87	328,07	0,23%	-0,23%	-2,36%
05-ott-87	328,08	0,00%	0,00%	-2,36%
06-ott-87	319,22	-2,74%	2,74%	-2,36%
07-ott-87	318,54	-0,21%	0,21%	-2,36%
08-ott-87	314,16	-1,38%	1,38%	-2,36%
09-ott-87	311,07	-0,99%	0,99%	-2,36%
12-ott-87	309,39	-0,54%	0,54%	-2,36%
13-ott-87	314,52	1,64%	-1,64%	-2,36%
14-ott-87	305,23	-3,00%	3,00%	-2,36%
15-ott-87	298,08	-2,37%	2,37%	-2,36%
16-ott-87	282,70	-5,30%	5,30%	-2,36%
19-ott-87	224,84	-22,90%	22,90%	-2,36%
20-ott-87	236,83	5,20%	-5,20%	-2,36%
21-ott-87	258,38	8,71%	-8,71%	-2,42%
22-ott-87	248,25	-4,00%	4,00%	-2,64%
23-ott-87	248,22	-0,01%	0,01%	-2,64%
26-ott-87	227,67	-8,64%	8,64%	-2,64%
27-ott-87	233,19	2,40%	-2,40%	-2,64%
28-ott-87	233,28	0,04%	-0,04%	-2,64%
29-ott-87	244,77	4,81%	-4,81%	-2,64%
30-ott-87	251,79	2,83%	-2,83%	-3,85%

Figure 14

Then we plot the losses and the var of the short

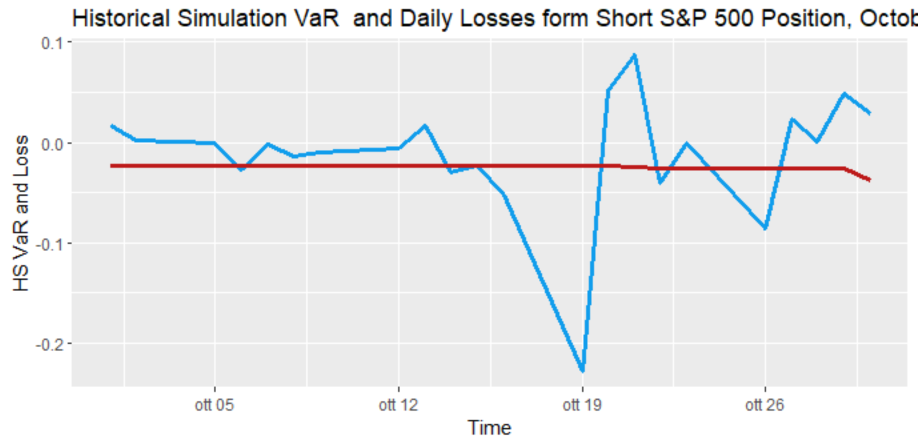

Figure 15

5 Exercise 3.1

In statistics we call moments a family of indecis, the most known are the first 4: mean, variance, skewness and kurtosis.
From the sample provided we have to compute:

- the first moment that is the mean,
- the second central moment that is the variance.

The mean is the sample average of the observation X that can be computed from the distribution of the returns from 06-ott-86 to 30-sett-87

$$\hat{E}(X) = \bar{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$$

We will use the mean formula in excel in order to compute it. The variance is the average squared deviation of the observation from the sample mean and we can compute be computed as:

$$v\hat{a}r(X) = \frac{1}{T-1}\sum_{t=1}^{T}(X_t - \bar{X})^2$$

We will use the variance formula in excel in order to compute the variance for both x and y for each set of data.
We have to define also the correlation between the 2 variables x and y

$$c\hat{o}rr(X,Y) = \frac{c\hat{o}rr(X,Y)}{s\hat{t}d(X)s\hat{t}d(Y)}$$

Data	I		II		III		IV	
	x	y	x	y	x	y	x	y
	10	8.04	10	9.14	10	7.46	8	6.58
	8	6.95	8	8.14	8	6.77	8	5.76
	13	7.58	13	8.74	13	12.74	8	7.71
	9	8.81	9	8.77	9	7.11	8	8.84
	11	8.33	11	9.26	11	7.81	8	8.47
	14	9.96	14	8.1	14	8.84	8	7.04
	6	7.24	6	6.13	6	6.08	8	5.25
	4	4.26	4	3.1	4	5.39	19	12.5
	12	10.84	12	9.13	12	8.15	8	5.56
	7	4.82	7	7.26	7	6.42	8	7.91
	5	5.68	5	4.74	5	5.73	8	6.89
Moments								
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	11.0	4.1	11.0	4.1	11.0	4.1	11.0	4.1
Correlation	0.82		0.82		0.82		0.82	
Regression								
a	3.00		3.00		3.00		3.00	
b	0.50		0.50		0.50		0.50	

Figure 16

In order to define the coefficient of the linear regression we use the following formulas:

$$a = E(r_t) - bE(r_t^m) \ ; \ b = \frac{cov(r_t, r_t^m)}{var(r_t^m)}$$

Firstly we have to compute the covariance between the dependent and the independent variable.

$$\hat{b} = \frac{c\hat{o}v(r_t, r_t^m)}{v\hat{a}r(r_t^m)} \ ; \ \hat{a} = \bar{r} - \hat{b}\bar{r}^m \ ; \ c\hat{o}v(X,Y) = \frac{1}{T-1}\sum_{t=1}^{T}(X_t - \bar{X})(Y_t - \bar{Y})$$

In excel we use the formula $covar()$

6 Exercise 3.4

In this exercise we used the excel function *LINEST* to estimate the one hundred *AR(1)* models in order to calculate their ϕ_1 coefficients. Then, we put on the x axis values from 0.76 to 1.16 in order to capture the bins in which the values fall in. And then we calculated the frequency of each value of the parameters inside each bin. So we plotted this result in the following histogram:

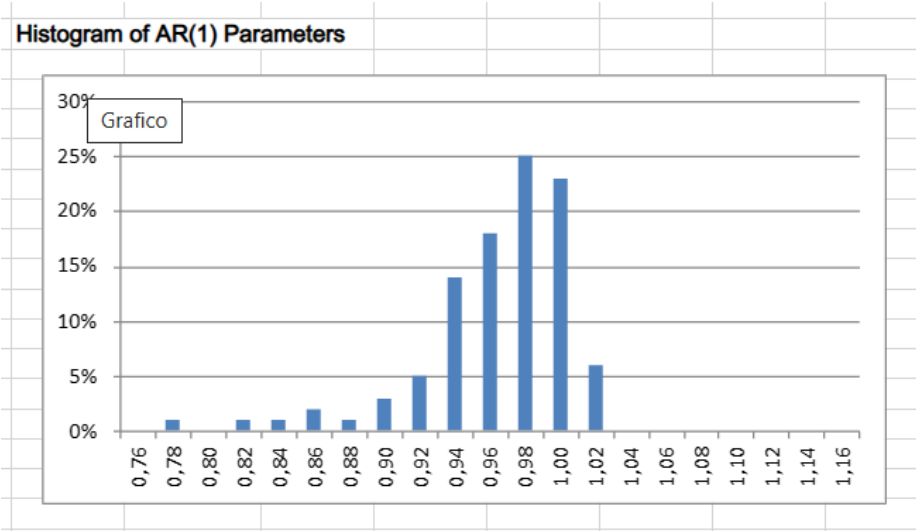

Figure 17

The histogram tells us that approximately the 23% of the estimated ϕ_1 are between 0.99 and 1.01, approximately 6% of ϕ_1's are between 1.01 and 1.03. So the *AR(1)* models that have parameters equal or higher than 1 show an explosive behavior, they cannot be used to make forecasts because they are substantially unpredictable. To be more specific we can say that the time series that show a coefficient greater than 1 have a unit root, so the series is not stationary, and its statistical features cannot be predicted according to an auto-regressive model of order 1.

7 Exercise 4.1

In this exercise we opted to fit the Garch(1,1) via the software R.
We simply used the *rugarch package*. This tool allows us to use the maximum likelihood estimation in order to obtain the estimates of α, β and ω, the parameters of the model. An important feature to take into account is the persistence, that is the sum of the 2 parameters α and β must be lower than 1 for the stationary condition, and both of them must be greater or at least equal to 0. If the persistent is greater then 1 the volatility blow up to infinity. In our case, what we got was a persistence lower than one, so our model is stationary and it is suppose to explain the volatility of the data and capture the non linear correlation among log-returns.
R delivered as estimates of the parameters:

$$\alpha = 7.996\,574 \times 10^{-2}, \beta = 9.112\,903 \times 10^{-1}, \omega = 1.262\,318 \times 10^{-6}$$

We plotted several plots in order to analyze if the Garch(1,1) model was a good model to explain the data. The first graph is only to show 1% VaR limits for the series, the latter ones are more important for our analysis.

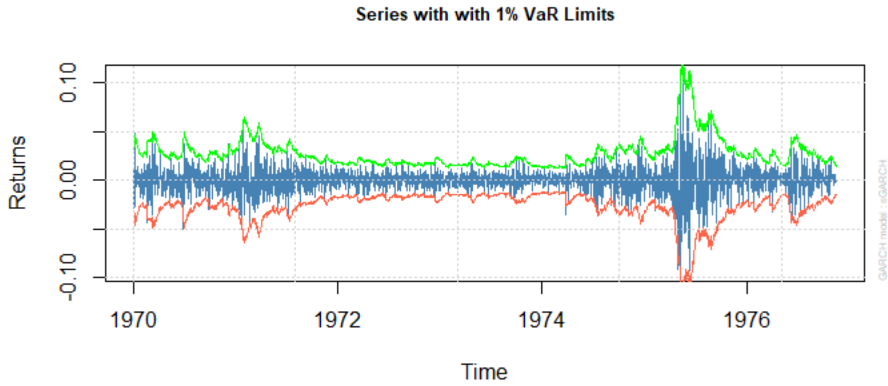

Figure 18

We plotted the empirical density of the residuals against a standard normal distribution, this 2 distributions should be more or less the same if the model is good for fitting the data. To better see this aspect we plotted a *Q-Q Plot* of the residuals, showing a distribution very close to the normal one.

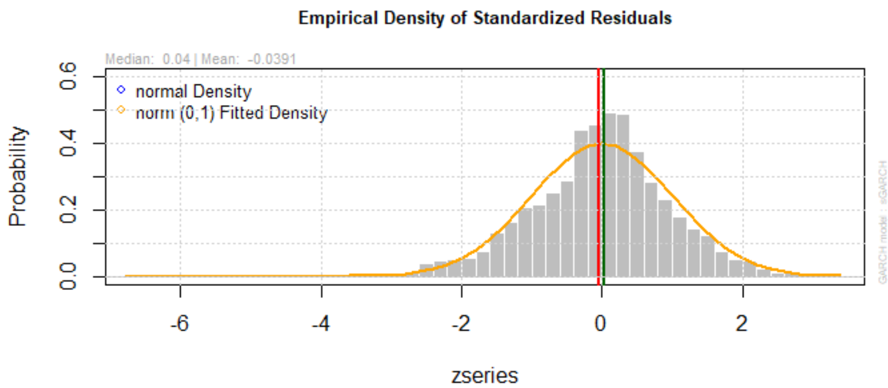

Figure 19

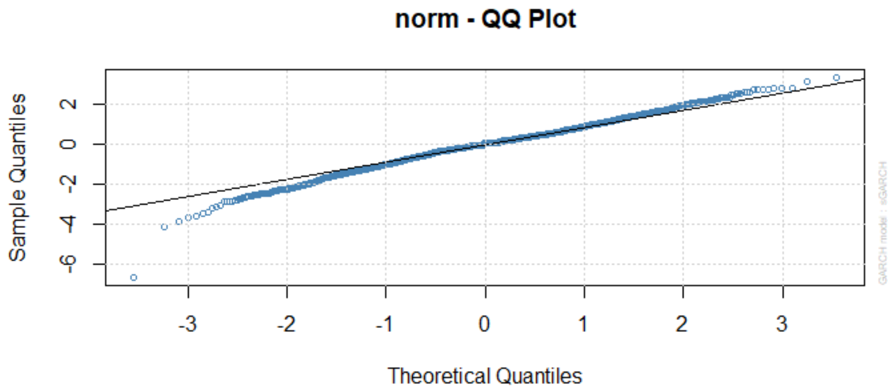

Figure 20

The ACF function of the residuals and squared residuals show that the they are uncorrelated, from the fact that most of the bars are inside the acceptance interval(the dashed lines), and the bars that are over the dashed lines ae very close to them, so we can state the are essentially 0.

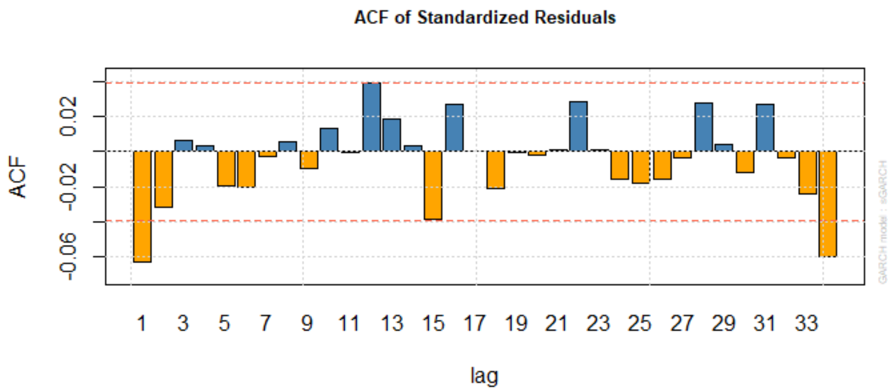

Figure 21

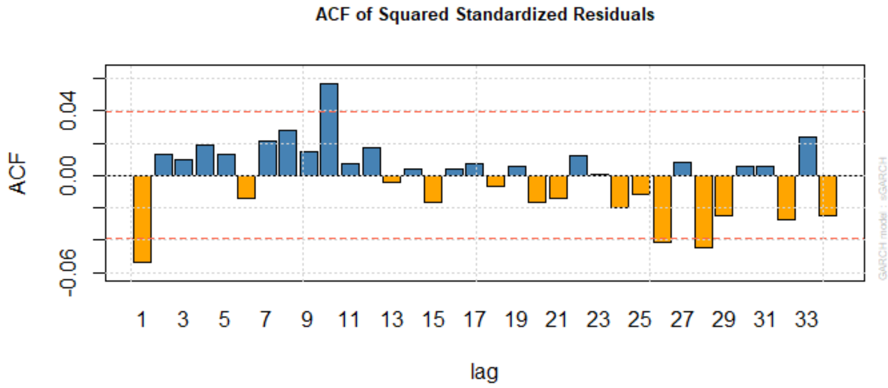

Figure 22

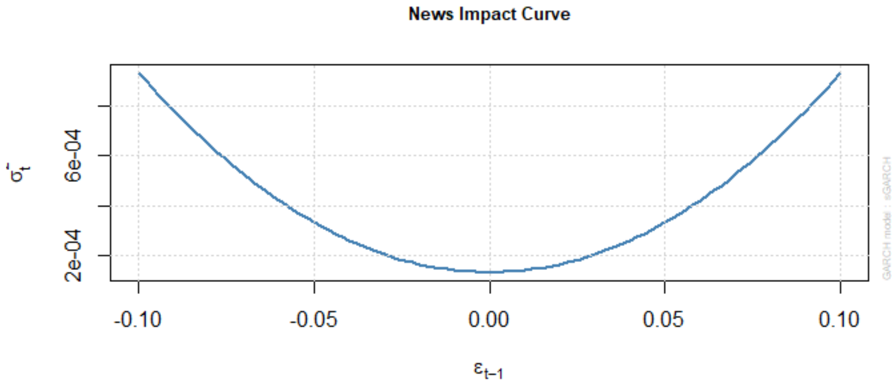

Figure 23: News-Impact curve of the Garch(1,1)

In comparison to the Garch(1,1) model, the Risk Metrics seen in Chapter 1 is a special case of the simple Garch model, forcing $\alpha = 1 - \lambda$ and $\beta = \lambda$, so that $\alpha + \beta = 1$, and further $\omega = 0$. So, the appear quite similar, but there is an important difference, in the simple Garch model we can define the unconditional variance as follows:

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta}$$

It is now clear that when $\alpha + \beta = 1$, as in Risk Metrics case, the long-run variance is not well defined in that model. An important oddity emerges: The Risk Metrics model ignores the fact that the long-run variance tends to be relatively stable over time.
The Garch model, in turn, implicitly relies on σ^2.