

## Homework 1

1. Calculate the increase of memory of PCs over the last 30 years and check whether the fMRI analysis could have been done 20 years ago.

**Solution:**

From the Internet, I collect the history of computer memory and show it on the table

Time	Memory size
1970s	256KB
1982	256KB
1988-1990	2MB
1991-1995	16MB
1996-2000	256MB
2001	1GB
2004	4GB
2009	8GB
2014	16GB

We know that fMRI image observed every 2 sec, 12 horizontal slices of the brain's scan,  $91 \times 109 \times 91(x,y,z)$  data points of size 22 MB. And in the experiment, we have 19 volunteers and do fMRI image for each person 1400 times. Sum all, the fMRI image data in the experiment are 571GB in size and it is too big to deal with 20 years ago.

2. Prepare 2-5 slides explaining logistic regression

**Solution:**

The slide called LR is uploaded to Github

3. install R and run simple programs from Quantlet.de, make sure you have a Github (GH) account.

**Solution:**

I have installed R and Rstudio before the course in my PC. And I already a Github account named TangDexuan567.

After class, I tried some programs like CMB, SFM and so on

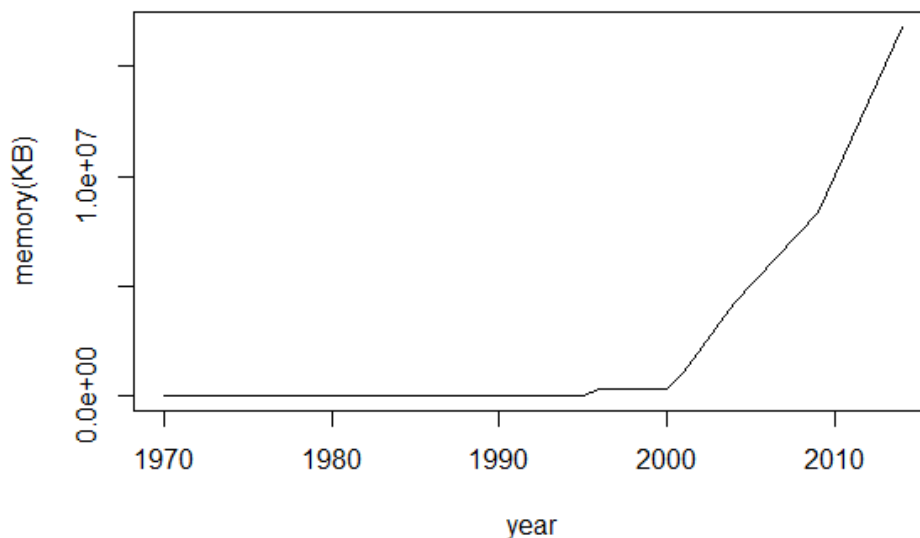
## Homework 2

1. make an R quantlet to solve HW #1 from unit 1 with R and show it on Github (GH) hint: use the CMB Qs for this work
2. use R with B-spline code to solve HW#1, any comments?

**Solution:**

I redo the HW1 and use the R code PC.R to plot the history of computer memory.

### The History of computer memory



3. Suppose you observe that in  $n=1000$  mails (in 1 week) you have about 2 scams. Use the LvB /Poisson cdf to calculate that you have 6 scam emails in 2 weeks. In Scammyland you have 5 scams on average, what is the probability to have no scam mail.

**Solution:**

(1)  $X$  denotes the number of scam emails you received.

Then  $X \sim \text{Poisson}(\lambda = 4)$

The probability of the event that you have received 6 scam emails:

$$P(X=6) = \frac{4^6}{6!} e^{-4} = 0.104.$$

(2) If  $\lambda = 5$ , then

The probability of the event that you have received no scam emails:

$$P(X=0) = e^{-5} = 0.0067.$$

## Homework 3

1. **Make an R quantlet on GH to produce hash code for the 2 sentences: "I learn a lot from this class when I am proper listening to the professor", "I do not learn a lot from this class when I am absent and playing on my Iphone". Compare the 2 hash sequences.**

### **Solution:**

The R code is as follows(It is also uploaded to Github named PHC.R):

```
rm(list=ls())
# call the library doing the hashes
library("digest")
# now do the hash code calculation
digest("I learn a lot from this class when I am proper listening to the professor",
"sha256")
digest("I do not learn a lot from this class when I am absent and playing on my
Iphone","sha256")
```

And the hash sequences for these two sentences are:

```
>digest("I learn a lot from this class when I am proper listening to the professor",
"sha256")
[1] "c16700de5a5c1961e279135f2be7dcf9c187cb6b21ac8032308c715e1ce9964c"
> digest("I do not learn a lot from this class when I am absent and playing on my
Iphone","sha256")
[1] "2533d529768409d1c09d50451d9125fdbaa6e5fd4efdeb45c04e3c68bcb3a63e"
```

We can see that It is totally different between these two hash sequences.

2. **Make 3-5 slides (in PPTX) on the DSA (Digital Signature Algorithms)**

### **Solution:**

The slide called DSA is uploaded to Github.

3. **Make slide with R code where you create a JSON data set that you save and read again.**

### **Solution:**

The slide called JSON is uploaded to Github and R code used is named JSON.R.

4. **Download the CRIX data and make a plot of the time series, analyze its properties, i.e. fit ARMA, ARIMA etc. Is there a GARCH effect?**

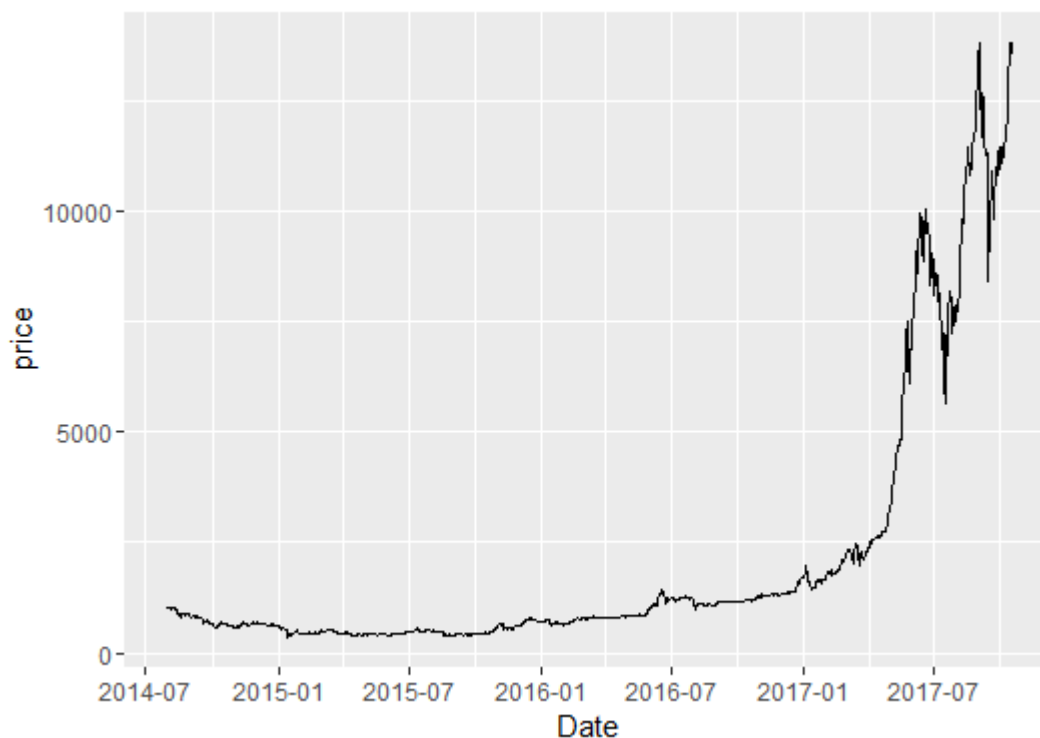
### **Solution:**

The whole R code used is named CRIX.R and uploaded to Github.

First I use the following R code to download the CRIX data and make a plot of the

time series:

```
rm(list=ls())
library(rjson)
library(zoo)
library(ggplot2)
#download CRIX file
Json_file = "http://crix.hu-berlin.de/data/crix.json"
Json_data = fromJSON(file=Json_file)
#transfer the JSON file to a standardized data frame format
crix_data_frame = as.data.frame(Json_data)
num1 = seq(from=1,to=2349,by=2)
num2 = seq(from=2,to=2350,by=2)
date = as.Date(t(crix_data_frame[,num1]))
price = as.numeric(t(crix_data_frame[,num2]))
graph = data.frame(time=date,price=price)
#plot the time series
p = ggplot(graph,aes(x=time,y=price))
p + geom_line(colour = 'black') + xlab('Date') + ylab('Price')
```



Then I used the function `auto.arima` in package `forecast` to fit the ARIMA model, and the code is as follows:

```
library(forecast)
zoo = zoo(price,order.by = date)
```

```
price_ts = ts(zoo)
```

```
fit = auto.arima(price_ts)  
fit
```

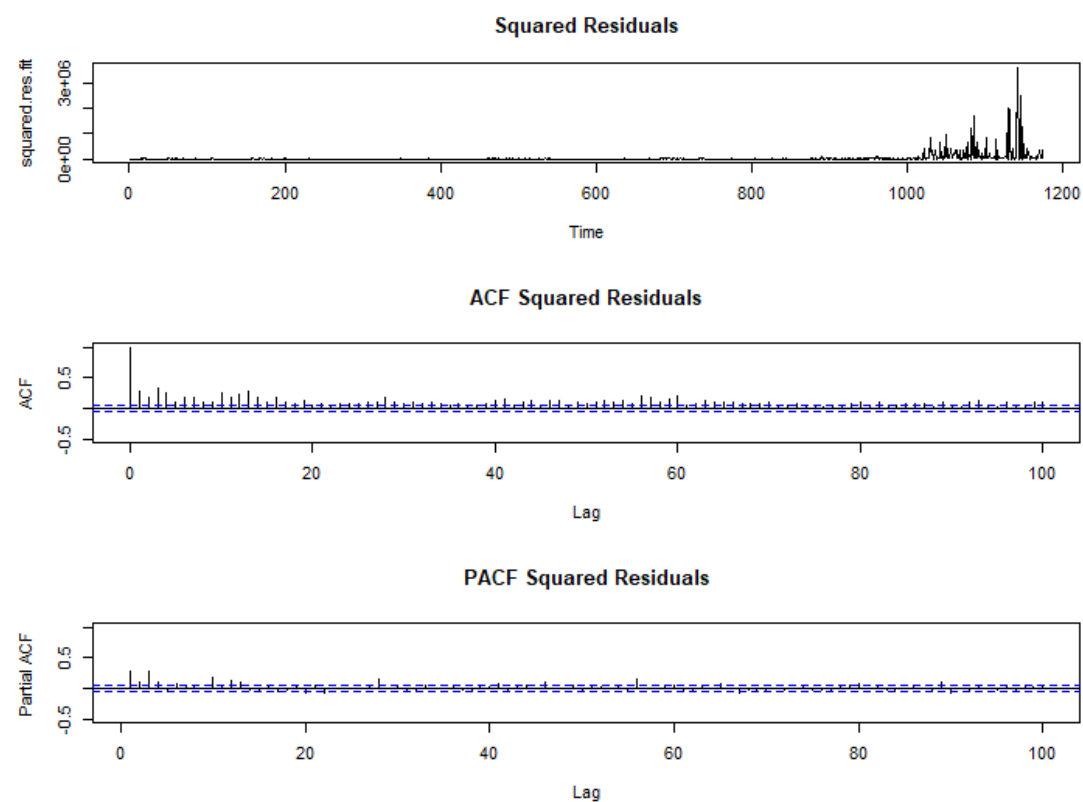
The results shows that the sequence is fitted ARIMA(5,2,0)

Now, we have to justify if there is a Garch effect.

Firstly, check if residual plot displays any cluster of volatility. Next, observe the squared residual plot. If there are clusters of volatility, ARCH/GARCH should be used to model the volatility of the series to reflect more recent changes and fluctuations in the series. Finally, ACF & PACF of squared residuals will help confirm if the residuals (noise term) are not independent and can be predicted. If the residuals are strict white noise, they are independent with zero mean, normally distributed, and ACF & PACF of squared residuals displays no significant lags.

Followings are the plots of squared residuals:

- Squared residuals plot shows cluster of volatility at some points in time
- ACF seems to die down
- PACF cuts off after lag 13 even though some remaining lags are significant



So it is convinced that there is an ARCH effect.

## Homework 4

(all this to be done on perfect PPTX slides)

1. improve the R quantlets on GH (from CRIX directory on quantlet.de) and make excellent graphics that follow Fig 3,4,5,6 of the „Econometrics of CRIX“ paper.
2. make your R code perfect as in the R examples on quantlet.de i.e. make sure that the code is „time independent“ by using actual dimensions of the data that you are collecting from [crix.hu-berlin.de](http://crix.hu-berlin.de) Recreate Fig 7 from „Econometrics of CRIX“.
4. redo as many figures as you can.

### Solution:

The slide is uploaded to Github and named HW4. The R code used in HW4 is named CRIX2.R