Financial Statistics: Time Series, Forecasting, Mean Reversion, and High Frequency Data

FINM 33170 and STAT 33910 Winter 2021
HW 1, due Friday 22 January, 2021.

1. **A data experiment.**

    (a) Install R or R studio on your computer.

    (b) Read file "f657ab9f9ecd5e33.csv" (on the class home page) into R, as follows

    ```
    series <- read.csv("f657ab9f9ecd5e33.csv")
    ```

    The file is 18.9 MB, and represents at least most trades in OEF (an Exchange Traded Fund (ETF) for S&P 100) for the year 2017. To see what you have on your hands, you can try the commands

    ```
    dim(series)
    series[1,]
    ```

    To more easily operate on the file, you can now give the command attach(series) We shall now investigate whether we can obtain a signature plot of the type shown in class:

    ```
    pr<-PRICE[DATE==20170103]
    vol<-c(1:20)*0
    for (k in 1:20){
     dipr<-diff(pr,lag=k)
     vol1<-sum(dipr^2)/k
     vol[k]<-sqrt(vol1*250) #annualized square root scale
    }
    plot(vol)
    ```

    Submit the plot, and conclude whether you think the price is a semi-martingale (or has an equivalent martingale measure).

    (c) (Optional) Have a look at

    ```
    ti<-TIME_M[DATE==20170103]
    ```

    Can you determine whether prices are listed in correct time order? If not, can you fix this problem?

    Hint: try

```
ti2<-levels(ti)[ti]
ti3<-ti2
for (i in 1:length(ti2)){
if ( nchar(ti2[i])==17) { ti3[i]<-paste("0",ti2[i],sep="")}
}
hour<-as.numeric(substr(ti3,start=1,stop=2))
```

etc. You may also need the command "order".

Further work in this direction: download and try to use the R package "chron".