Econometric_HW1_part5

Marjan Rezvani

9/13/2020

```
#install.packages('zoo')
library(zoo)
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
library(data.table)
```

Download data from yahoo/finance for SPY (SP500 Index) from 2018-12-31 to 2019-12-31.

I saved the data as csv and now I am going to read it into R.

```
setwd("/Users/marjanrezvani/Documents/Fall2020/eco_stat/assignment/")
spy <- read.csv(file = 'SPY.csv', header = TRUE)
head(spy)
```

```
##         Date   Open   High    Low  Close Adj.Close    Volume
## 1 2019-01-02 245.98 251.21 245.95 250.18  243.0259 126925200
## 2 2019-01-03 248.23 248.57 243.67 244.21  237.2266 144140700
## 3 2019-01-04 247.59 253.11 247.17 252.39  245.1727 142628800
## 4 2019-01-07 252.69 255.95 251.69 254.38  247.1058 103139100
## 5 2019-01-08 256.82 257.31 254.00 256.77  249.4274 102512600
## 6 2019-01-09 257.56 258.91 256.19 257.97  250.5931  95006600
```

I am going to calculate simple return as

$R(t) = \frac{p(t)-p(t-1)}{p(t-1)}$

Let's first add p(t-1) to the data

```
spy$lag_price <- shift(spy$Adj.Close)
head(spy)
```

```
##         Date   Open   High    Low  Close Adj.Close    Volume lag_price
## 1 2019-01-02 245.98 251.21 245.95 250.18  243.0259 126925200        NA
## 2 2019-01-03 248.23 248.57 243.67 244.21  237.2266 144140700  243.0259
```

```
## 3 2019-01-04 247.59 253.11 247.17 252.39  245.1727 142628800  237.2266
## 4 2019-01-07 252.69 255.95 251.69 254.38  247.1058 103139100  245.1727
## 5 2019-01-08 256.82 257.31 254.00 256.77  249.4274 102512600  247.1058
## 6 2019-01-09 257.56 258.91 256.19 257.97  250.5931  95006600  249.4274
```

```
spy$Ret <- (spy$Adj.Close - spy$lag_price)/spy$lag_price
head(spy)
```

```
##         Date   Open   High    Low  Close Adj.Close    Volume lag_price
## 1 2019-01-02 245.98 251.21 245.95 250.18  243.0259 126925200        NA
## 2 2019-01-03 248.23 248.57 243.67 244.21  237.2266 144140700  243.0259
## 3 2019-01-04 247.59 253.11 247.17 252.39  245.1727 142628800  237.2266
## 4 2019-01-07 252.69 255.95 251.69 254.38  247.1058 103139100  245.1727
## 5 2019-01-08 256.82 257.31 254.00 256.77  249.4274 102512600  247.1058
## 6 2019-01-09 257.56 258.91 256.19 257.97  250.5931  95006600  249.4274
##            Ret
## 1           NA
## 2 -0.023862833
## 3  0.033495718
## 4  0.007884672
## 5  0.009395268
## 6  0.004673604
```

```
spy <- na.omit(spy)
head(spy)
```

```
##         Date   Open   High    Low  Close Adj.Close    Volume lag_price
## 2 2019-01-03 248.23 248.57 243.67 244.21  237.2266 144140700  243.0259
## 3 2019-01-04 247.59 253.11 247.17 252.39  245.1727 142628800  237.2266
## 4 2019-01-07 252.69 255.95 251.69 254.38  247.1058 103139100  245.1727
## 5 2019-01-08 256.82 257.31 254.00 256.77  249.4274 102512600  247.1058
## 6 2019-01-09 257.56 258.91 256.19 257.97  250.5931  95006600  249.4274
## 7 2019-01-10 256.26 259.16 255.50 258.88  251.4771  96823900  250.5931
##            Ret
## 2 -0.023862833
## 3  0.033495718
## 4  0.007884672
## 5  0.009395268
## 6  0.004673604
## 7  0.003527523
```

let's calculate mean of return.

we can eitherr use mean command or summary command

```
mean(spy$Ret)
```

```
## [1] 0.001110172
```

```
summary(spy)
```

```
##          Date          Open            High            Low
##  2019-01-03:  1   Min.   :247.6   Min.   :248.6   Min.   :243.7
```

```
##  2019-01-04:  1   1st Qu.:282.1   1st Qu.:283.4   1st Qu.:280.4
##  2019-01-07:  1   Median :291.4   Median :292.9   Median :290.1
##  2019-01-08:  1   Mean   :291.0   Mean   :292.2   Mean   :289.7
##  2019-01-09:  1   3rd Qu.:299.8   3rd Qu.:300.8   3rd Qu.:298.5
##  2019-01-10:  1   Max.   :323.7   Max.   :323.8   Max.   :322.3
##  (Other)   :245
##      Close          Adj.Close         Volume            lag_price
##  Min.   :244.2   Min.   :237.2   Min.   : 20270000   Min.   :237.2
##  1st Qu.:281.8   1st Qu.:275.0   1st Qu.: 50871250   1st Qu.:274.6
##  Median :291.6   Median :285.3   Median : 64304000   Median :285.1
##  Mean   :291.2   Mean   :285.1   Mean   : 69805476   Mean   :284.8
##  3rd Qu.:299.9   3rd Qu.:294.4   3rd Qu.: 83246250   3rd Qu.:294.3
##  Max.   :322.9   Max.   :319.6   Max.   :178745400   Max.   :319.6
##
##       Ret
##  Min.   :-0.030073
##  1st Qu.:-0.002219
##  Median : 0.001221
##  Mean   : 0.001110
##  3rd Qu.: 0.005886
##  Max.   : 0.033496
##
```

```
spy$ret_lag1 <- shift(spy$Ret)

head(spy)
```

```
##         Date   Open   High    Low  Close Adj.Close    Volume lag_price
## 2 2019-01-03 248.23 248.57 243.67 244.21  237.2266 144140700  243.0259
## 3 2019-01-04 247.59 253.11 247.17 252.39  245.1727 142628800  237.2266
## 4 2019-01-07 252.69 255.95 251.69 254.38  247.1058 103139100  245.1727
## 5 2019-01-08 256.82 257.31 254.00 256.77  249.4274 102512600  247.1058
## 6 2019-01-09 257.56 258.91 256.19 257.97  250.5931  95006600  249.4274
## 7 2019-01-10 256.26 259.16 255.50 258.88  251.4771  96823900  250.5931
##            Ret     ret_lag1
## 2 -0.023862833           NA
## 3  0.033495718 -0.023862833
## 4  0.007884672  0.033495718
## 5  0.009395268  0.007884672
## 6  0.004673604  0.009395268
## 7  0.003527523  0.004673604
```

```
spy$ret_lag2 <- shift(spy$Ret,n = 2)
head(spy)
```

```
##         Date   Open   High    Low  Close Adj.Close    Volume lag_price
## 2 2019-01-03 248.23 248.57 243.67 244.21  237.2266 144140700  243.0259
## 3 2019-01-04 247.59 253.11 247.17 252.39  245.1727 142628800  237.2266
## 4 2019-01-07 252.69 255.95 251.69 254.38  247.1058 103139100  245.1727
## 5 2019-01-08 256.82 257.31 254.00 256.77  249.4274 102512600  247.1058
## 6 2019-01-09 257.56 258.91 256.19 257.97  250.5931  95006600  249.4274
## 7 2019-01-10 256.26 259.16 255.50 258.88  251.4771  96823900  250.5931
##            Ret     ret_lag1     ret_lag2
## 2 -0.023862833           NA           NA
## 3  0.033495718 -0.023862833           NA
```

```
## 4  0.007884672  0.033495718 -0.023862833
## 5  0.009395268  0.007884672  0.033495718
## 6  0.004673604  0.009395268  0.007884672
## 7  0.003527523  0.004673604  0.009395268
```

calculate mean return for days in which previous day return is positive

```
mean(subset(spy, ret_lag1 > 0)$Ret)
```

```
## [1] 0.001420169
```

calculate mean return for days in which two previous days returns are positive

```
mean(subset(spy, (ret_lag1 > 0) & (ret_lag2 >0))$Ret)
```

```
## [1] 0.002070532
```

calculate mean return for days in which previous day return is negative

```
mean(subset(spy, ret_lag1 < 0)$Ret)
```

```
## [1] 0.0009052065
```

calculate mean return for days in which two previous days returns are negative

```
mean(subset(spy, (ret_lag1 < 0) & (ret_lag2 < 0))$Ret)
```

```
## [1] 0.002273021
```