CISC5352: Financial Data Analytics Quiz (3)

Stock data basic analytics

A) Retrieve stock data from web (60 points)

Go through your Lecture 3 to digest the stock retrieve, volatility calculation details via using Pandas and related materials. Then finish the following assignments.

1. Retrieve stock data for the following types of companies from Sept 01, 2006 to Sept 01, 2018 in Excel files[1]
   (a) IT: GOOGLE (GOOG), APPLE (AAPL), Microsoft (MSFT), Amazon (AMZN), "INTC'
   (b) Bank: 'JPM', 'BAC', 'HSBC', 'C' , 'GS'
   (c) Retail: WMT, TGT, JCP, KSS, HD
   (d) Pharmaceutical: PFE, MRK,JNJ, ABBV, BMY
   (e) Fashion: TPR, HM-B, GES,MC, TIF
2. Compare the stock price patterns of these companies during the years (need plots)
3. Calculate the days of up and down for each stock in each year.
4. Calculate the stock log returns for each stock and plot it

   The i^{th} day log return $u_i = \ln(\frac{S_i}{S_{i-1}})$, where S_i is stock close price at i^{th} day.
5. Compute the max, min, mean, median, standard deviation, skewness, kurtosis for close price and volume for each data set (need plots).

How to compute skewness and kurtosis for a data set (sample)?[2]

Given a sample $X : x_1, x_2 \cdots x_n$, its skewness and kurtosis can be calculated as follows

[1] Extra credits will be given for folks who retrieved them in one Excel file via programming.

[2] pandas has built-in function to calculate these measures, but you need to now its meaning

$$shewness(X) = \frac{n}{(n-1)(n-2)} \frac{\sum_{i=1}^{n}(x_i - \bar{X})^3}{s^3}$$

$$Kurtosis(X) = \frac{1}{n} \frac{\sum_{i=1}^{n}(x_i - \bar{X})^4}{s^4} - 3$$

where s is the standard deviation of X

6. Output all stock prices (close price) which are $>=$ 95% percentile and their corresponding volumes and dates for each data set.

7. Give your analysis about volatility performance of the stocks (need plots).

8. Calculate and visualize the recent five years S-ratios for the stocks.

   (a) S-ratio is defined as $\frac{E(u)-E(r)}{std(u)}$, where u is the stock log return in the time interval, $E(u)$ is the average return in the time interval, $E(r)$ is the average interest rate in the time interval (use 3% in your calculation), and $std(u)$ is the STD in the time interval.

B) Polish your software so that (20 points)

- user enter any stock symbol, your software can automatically show
  - 1. max, min, mean, median of the stock price in the last 1, 3, 5,10 years
  - 2. The plot of the stock close price in 1,3,5,10 years
  - 3. max, min, mean, median of the stock log return in the last 1, 3, 5,10 years
  - 4. A SIMPLE suggestion for customers about buy/sell/hold according to the data you get

Write your SVD (20 points)

Use eig function in numy to implement a SVD so that if you output U,S,V as linalg.svd(...). Your SVD signature is

```
def mySVD(X):
## INPUT X: NDarray
## OUTPUT:
## U,S,V :the same output as linalg.svd(X) (Ndarray type)
```

- You should verify if U,V are unitary matrices and your SVD really works.
- You don't need to use very big input matrix because such eig based implementation is not efficient in practice.

Pricing Barrier Options (30 points)

As an exotic option, the payoff of a barrier option is no longer dependent on the underlying asset price at expiration time. Instead, it has a special asset price S_b selected as a barrier value, which can be crossed or not. Here 'crossed' means the barrier value S_b above or below the current asset S. It is noted that the option holders has the right to exercise their European call or put option at the exercise price K, provided S does not cross S_b

- A knock-*out* (in) option is a European option whose contracts will be *canceled* (activated), if the barrier value is crossed,
- There are four main barrier options
  - Down-and-out options: the option will be void if $S < S_b$
  - Down-and-in options: the option will be activated if $S_b < S$
  - Up-and-out options: the option will be cancelled if $S_b < S$
  - Up-and-in options: the option will be activated if $S_b > S$

For some barrier options, there are analytic pricing solution available. Given *a down-and-out put* with strike price K, expiring time T and a barrier value S_b, we have the following results:

$$P = Ke^{-rT}\{N(d_4) - N(d_2) - a[N(d_7) - N(d_5)]\} - S\{N(d_3) - N(d_1) - b[N(d_8) - N(d_6)]\}$$

where we have the following parameters:

- $a = (S_b/S)^{-1+2r/\sigma^2}$, $b = (S_b/S)^{1+2r/\sigma^2}$
  - $d_1 = \frac{\log(S/K)+(r+\sigma^2/2)T}{\sigma\sqrt{T}}$
  - $d_2 = \frac{\log(S/K)+(r-\sigma^2/2)T}{\sigma\sqrt{T}}$
  - $d_3 = \frac{\log(S/S_b)+(r+\sigma^2/2)T}{\sigma\sqrt{T}}$
  - $d_4 = \frac{\log(S/S_b)+(r-\sigma^2/2)T}{\sigma\sqrt{T}}$

- $d_5 = \frac{\log(S/S_b)-(r-\sigma^2/2)T}{\sigma\sqrt{T}}$
- $d_6 = \frac{\log(S/S_b)-(r+\sigma^2/2)T}{\sigma\sqrt{T}}$
- $d_7 = \frac{\log(SK/S_b^2)-(r-\sigma^2/2)T}{\sigma\sqrt{T}}$
- $d_8 = \frac{\log(SK/S_b^2)-(r+\sigma^2/2)T}{\sigma\sqrt{T}}$

Complete the following assignment

- Price the following down-and-in put option with by using a sequence of barrier values: $S_b = [70, 60, 50, ...20, 10, 1]$ and compare the prices with its price under BS model.

  $S = 100$

  $K = 105$

  $r = 0.05$

  $T = 0.75$

  $\sigma = 0.4$

- Price the down-and-out-put option with by using a sequence of $\sigma : \sigma = [0.55, 0.5, 0.45, 0.4, 0.35..., 0.1]$ and visualize the relationship between volatility and such a barrier option price.

  $S = 100$

  $K = 105$

  $r = 0.05$

  $T = 0.75$

  $S_b = 60$

What should you turn in?

- 1. A folder contains your source files and related output.
- 2. Please name your folder as first-name_last-name_CISC5352_quiz_3. For example, John_Smith_CISC5352_quiz_3 if your name is John Smith.
- 3. Send the zipped file (.zip instead of ,rar) of your folder to Blackboard before 11:59 pm Oct 10, 2018