Advance Simulation Methods Project

Stock Price Forecasting

Qingmin Shi

1. Introduction:

The data I use in this simulation is the stock price provided by professor, 'price_updated.txt'. The last date of the data is 2016-12-02, and I tried to predict the stock's 80% confidence interval in next 10 days.

2. Methodology:

In general, the simulation is in a Geometric Brownian Motion framework, and with some modifications.

I first visualize price and daily returns to find some trend and features of that stock, then I use the analytical insights to construct the model that can capture as much information as possible. Finally, I quantify the information, and put them into my simulation model.

3. Analytic Process:

First of all, I plot the stock price, with 10 days moving average, and a relatively longer horizon, 30 days moving average in one chart.

The width of 10 days moving average line represents its deviations from 30 days moving average.

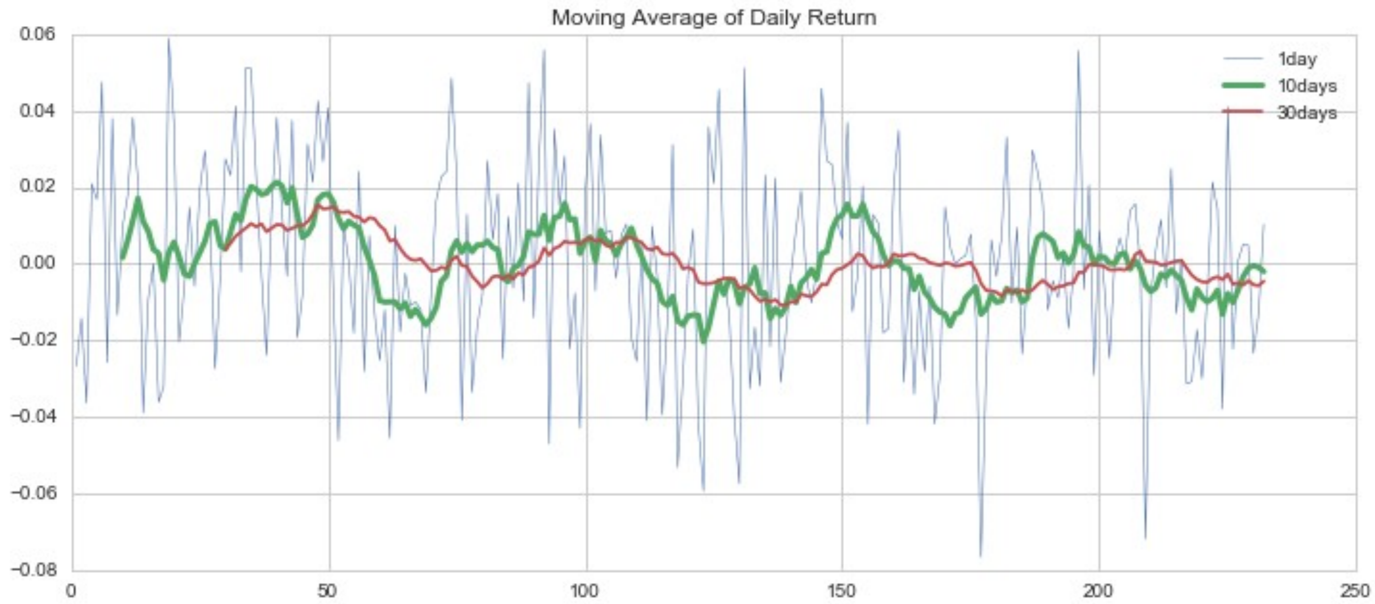
The blue line is the daily stock price, the red line is the 10 days moving average, and green line is 30 days moving average. Since I am going to predict 10 days' stock price, so the 10 days moving average might be a good indicator for the drift parameter in a Geometric Brownian Motion model. Then we should try to figure out the trend of the drift.

According to a classic technical indicator analysis, it is always important to explore the deviation of a short term moving average and a long term moving average. In this report, 10 days and 30 days moving average are short term and long term indicators respectively. So I play a small trick on the 10 days moving average line. The width of 10 days moving average line represents its deviations from 30 days moving average. We can conclude that, the wider the red line is, the more possibly the stock price converges to green line.

In this situation, the following trend could be 3 different scenarios. A) the red line is relatively wide now, and the price could probably go up. B) According to historical behavior, there still is possibility that the stock goes down or remain stable. I consider three different scenarios with different probability.

Scenario	UP	STAY	DOWN
Probability	40%	30%	30%

Then I plot the daily return to determine the possible values of drift in different scenarios,



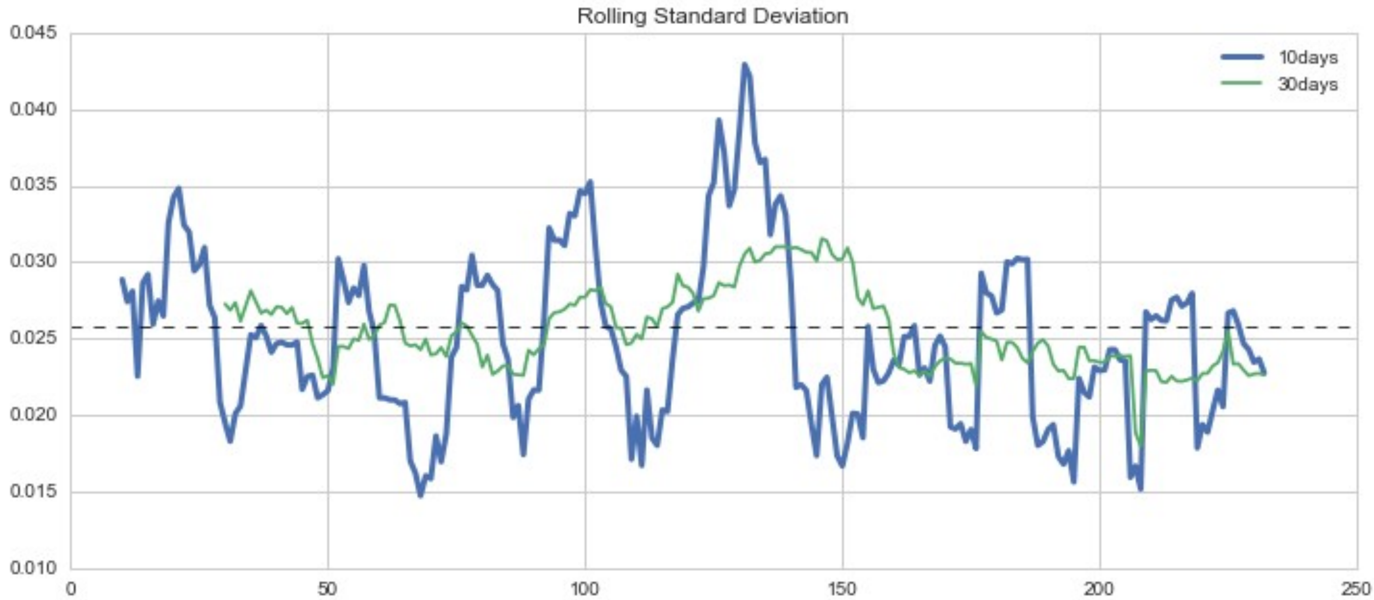
From the chart, we can clearly see the historical behavior of the return, and I make a reason deduction that, if the stock went up, the daily return could be around 0.006, with a range [0.003, 0.01], if the stock stay, the daily return could be around 0, with a range [-0.003, 0.003], and if the stock went down, the daily return could be around -0.008, with a range [-0.012, -0.003].

Scenario(mu)	Mean	Lower Bound	Upper Bound
UP	0.006	0.003	0.01
STAY	0	-0.003	0.003
DOWN	-0.008	-0.012	-0.003

Then I could use a uniform distributed variable to simulate the scenario, and within each scenario I use a truncated normal distribution to get the drift value.

Now, we could have drift parameter values, then I am going to get the variance parameter.

In the same way, a plot could provide some insights of how the variance could perform. So I plot the 10 days rolling standard deviation and 30 days rolling standard deviation.



The black dashed line is the value of historical standard deviation.

According to the chart above, we could see that the historical standard deviation is not good candidate for the variance parameter of the Geometric Brownian Motion for two reasons, the first one is that historical standard deviation doesn't capture the recent trend of the volatility, the second one is that historical standard deviation can't reflect the jump.

In my model, I use a Poisson process to describe the jump of volatility. Lambda is the rate of jump, and I estimate it using recent historical data. From index 150 to last of above chart, we could see that nearly 9 jump happened out of 80 days. So a proper value for lambda is 9/80. Also, we could see that it is impossible for the volatility to jump three time during a 10 days' horizon. So I use a Poisson process with lambda being 9/80, and ignoring the situation that it jumps more than twice.

If there was no jump in next 10 days, my estimate of sigma is a truncated normal distribution with mean 0.0225 and range [0.02, 0.025]. If there was jump, I think there was 60% chance to jump down, then follows a mean 0.0175, range [0.015, 0.02] truncated normal distribution, and 40% chance to jump up, then follows a mean 0.028, range [0.025, 0.032] truncated normal distribution.

Scenario(sigma)	Mean	Lower Bound	Upper Bound
UP	0.028	0.025	0.032
STAY	0.0225	0.02	0.025
DOWN	0.0175	0.015	0.02

The Poisson process could give 5 scenarios for volatility parameter sigma:

Scenario	1	2	3	4	5
Jump Times	0	1	1	2	2
Crsp. Dist. Scn	STAY	UP *(60%)*	DOWN *(40%)*	UP, DOWN	DOWN, UP
Lambda			***9/80***		

So the method to generate volatility is as follows. First, Poisson process gives a number of jumps. Then, it would be "STAY" scenario if there was no jump, and it would be 60% chance of "UP" scenario and 40% chance of "DOWN" scenario if there was one jump. And if there were 2 jumps, the second jump is always opposite to the first jump, and probability is the same with former. Because, since observation 150^{th}, there had been no two jumps in the same way according the chart above.

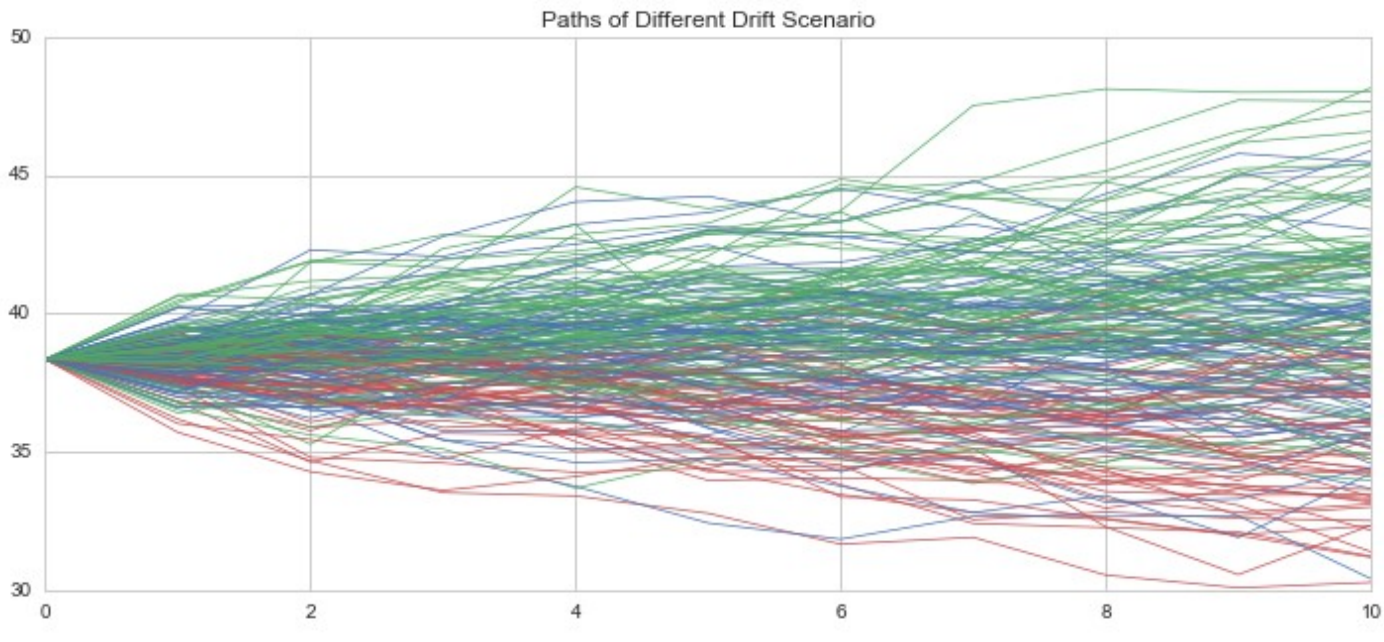
After I get the mu and sigma, I can simulate a Geometric Brownian Motion. To sum up, the following chart shows all random components of my model.

a) Scenario Randomness (UP, STAY, DOWN) b) Value Randomness within Each Scenario	**Drift/mu**	**Geometric Brownian Motion**
a) Poisson Process Randomness b) Scenario Randomness (UP, STAY, DOWN) c) Value Randomness within Each Scenario	**Volatility/sigma**	

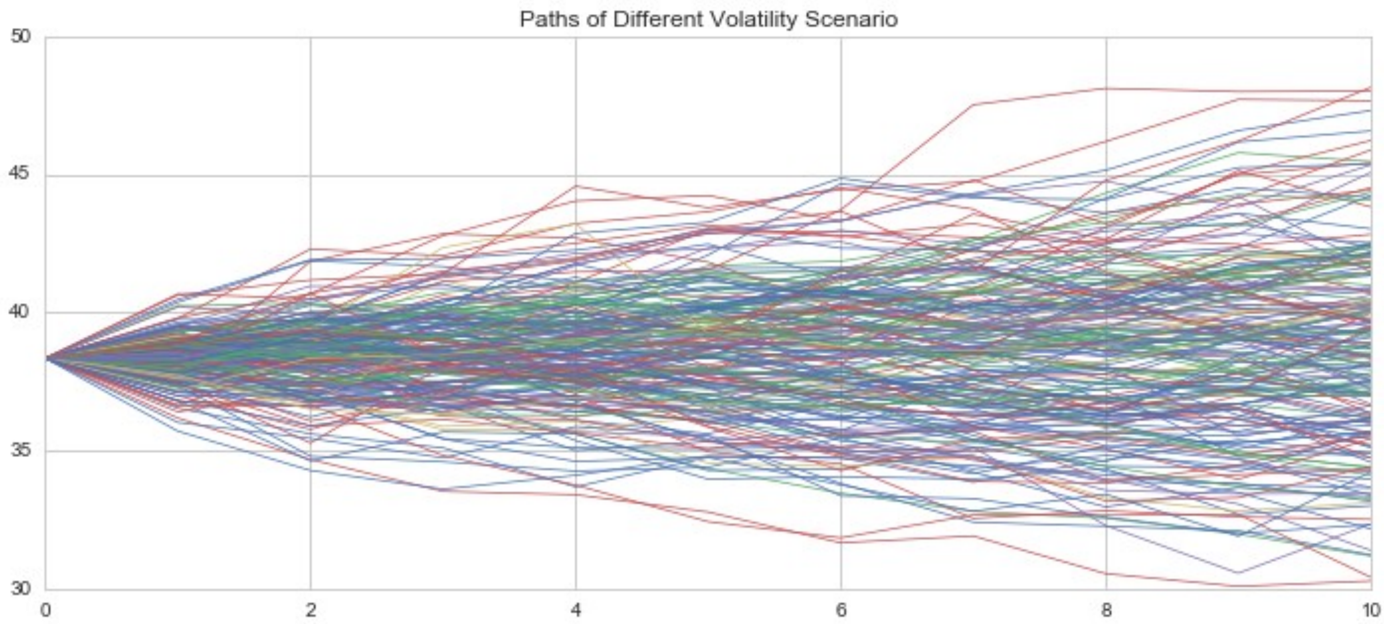
4. Simulation Results:

All the simulation methods and parameters have been listed above. Since there are 3 scenarios for drift and 5 scenarios for volatility, I simulate 200,000 paths to ensure there is a decent number of occurrences for each combination of drift and volatility.

The following two charts shows paths of different scenario. The first one is paths of different drift scenario, and the second one is paths of different volatility scenario. For a better visualization quality, I randomly pick up 200 paths out of 200,000 paths.

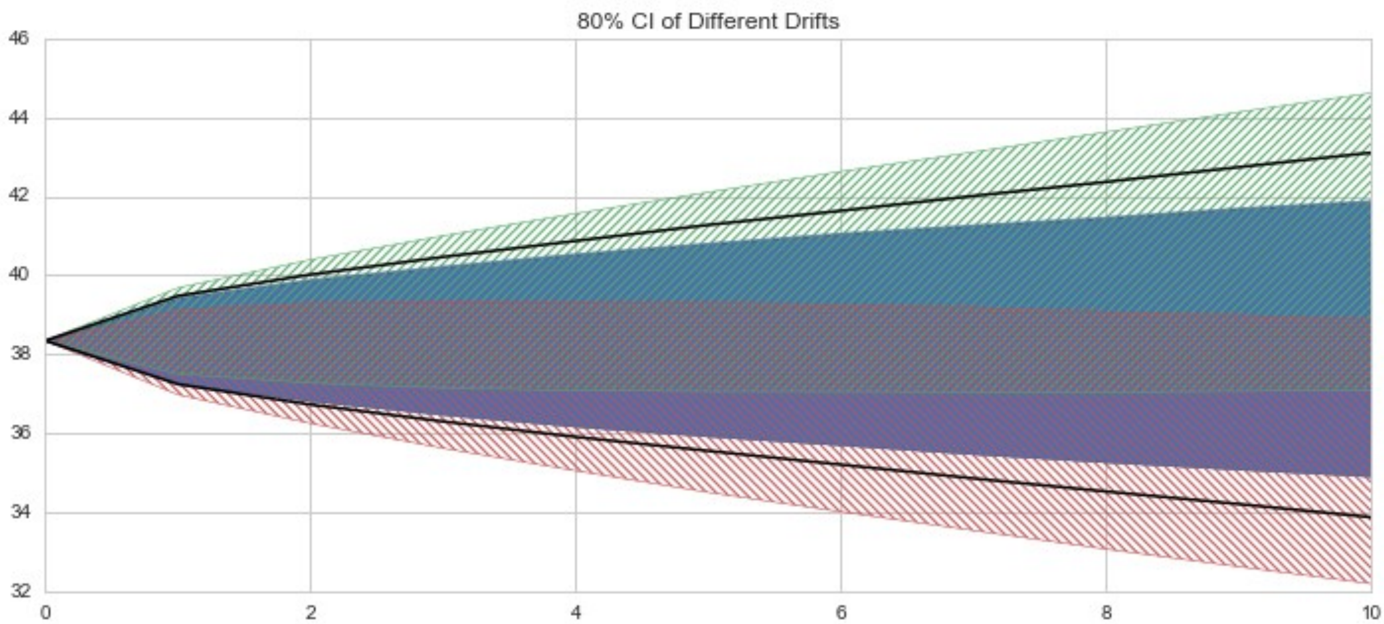


Green line: *go up*; **Blue line:** *stay*; **Red line:** *go down*.



Blue line: *stay*; **Green line:** *1 jump go down*; **Red line:** *1 jump go up*; **Purple line:** *2 jumps down and up*; **Yellow line:** *2 jumps up and down*

Then I have 200,000 paths for a 10 day's horizon, I get 10% and 90% quantile of those paths to construct an 80% confidence interval. I also plot the 80% confidence interval for 3 drift scenarios, representing the 80% interval if that scenario happened, e.g. stock price goes up.



Blue area: stay; green area: go up; red area: go down; area between black line is the total 80% CI

Above all, the 80% confidence interval prediction of this stock price over a 10 days' horizon is listed in the table below.

Days	**1**	**2**	**3**	**4**	**5**
Lower Bound	37.23	36.72	36.30	35.91	35.54
Upper Bound	39.48	40.01	40.45	40.87	41.25
Days	**6**	**7**	**8**	**9**	**10**
Lower Bound	35.19	34.85	34.52	34.21	33.88
Upper Bound	41.64	42.01	42.35	42.73	43.09