Ukrainian Catholic University

Applied Sciences Faculty

Comparison of decision making golden rule and expected maximum value estimation, or how to make money on stocks?

Authors:

Iryna Popovych

Anastasiia Livochka

Supervisor:

Dr. Rostyslav Hryniv

December 17, 2018

Abstract

The problem of correctly guessing the maximum of a sequence is a sequential decision game. On each step, you see only one number, and you have to choose immediately whether to accept it as the maximum. You can choose a primitive strategy, and your chances to win will be near $\frac{1}{n}$. Can we do better? The problem is very interesting from the view of probability science; it firstly appeared in 1960, in an issue of Scientific America, and has been studied extensively in the fields of applied probability, statistics, and decision theory.

Furthermore, this problem has a lot of implications in real life. Deciding on buying a ticket on the Ryanair website, hiring a worker, getting married, selling or buying stocks are the problems that involve optimal stopping theory. Should you buy a ticket now, or wait for a little? It is the same problem as correctly guessing the maximum or minimum of a sequence. Haste makes waste. From the other side, waiting for too much leads to loosing the best option.

In the article, we considered a problem of selling stocks. Firstly, we applied the **decision making golden rule**, that increases your chances for choosing the maximum price up to **37%**. We took into account that golden rule approach maximizes the probability of getting top-1 maximum price, but in most cases, we obtain mediocre results, **9th rang**, on average. We used another approach, **expected maximum**, and obtained **9th rang** as well, on average, and **expected maximum*** returns on average **8th rang** . The article describes the approaches, simulations, and results we've got.

I. PROBLEM STATEMENT. TO SELL OR NOT TO SELL?

There are many studies about buying stocks, factors which influence price, and finding best moment to buy. People pay much attention and put their efforts to choose the best moment to buy a stock, but hardly ever think ahead, about selling it. But that's a mistake: the money is made when the sale is made. Understanding this can be the key to claiming your profits. Or, at least, cutting your losses.

Let's imagine that you are a stock holder who decided to sell some of your stocks. Sure, you main motivation is to gain as much profit as you can, which means to sell the stock in the moment of maximum stock price, considering current period.

The problem of guessing the moment of price maximum is a sequential decision process. In any time T you can choose: to sell or not to sell? Your expected winnings depend on your strategy. This problem reminds us a classical secretary problem, however, with some extensions. In secretary problem your aim is to hire the best employee, but you can interview only one at time and you should make your decision immediately, before current employee leave. The ranks of employees are independent of each other. In the problem of finding the best moment to sell the stocks we also can also consider market rules, there all things become complicated, but the answer should be more accurate. Our main aim is to find the best strategy for making money on stocks.

II. GOLDEN RULE APPROACH

Determining the best stock price

The problem of determining the best moment for selling stocks at a high price is a decision-making problem. You should guess the point where it becomes the most profitable. Let's discuss some strategies to gain the maximum profit. We denote by W the event of our win. So:

$$W = \{\text{GUESSING EXACTLY THE MAXIMUM PRICE ON STOCK}\}$$

What is the probability of getting W? Actually, it depends on you decision making strategy. We could use a primitive one. What if we always chose the first or the last day?

Let's denote such strategy by S. In such case our probability to win:

$$P(W|S) = \frac{1}{N}$$

Not impressing. That is why we decided to try using "A Golden Rule for Decision Making", proposed by Sardelis and Valahas [1].

This rule is about getting some information from observations before making the actual decision. In the beginning, we applied the golden rule to the data on Apple stock prices to see what results it will give us. What about the price data we had? We used APPL **daily opening stock prices** over the past 10 years (2008-2017). We divided every year into months to perform our estimations. Therefore, we had 120 time periods. The daily stock price **data spans** for every month vary from \$1.41 to \$26.63, which is the reason to assume that there are sharp price rises (and falls) during the chosen period. Our task was, starting from the beginning of each period, to determine the moment when the stock price is the highest among the whole month.

Out of all N days in month, the trader lets n days pass, so that $1 < n < N - 2$, ranks prices of those days them in order of desirability (from highest to lowest), and then among the next days prices selects the first one found with a higher rank, i.e., the first one that more expensive that before. The essential problem here: how many days trader should let to pass? Let's denote by S_n the strategy of observing stock prices exactly **n** days. This strategy wins if:

- The best price hasn't appeared in one of the first n days. Otherwise, we just passed it.
- The ranks, preceding the best, don't exceed those of the first n prices.

Let's denote by B_k the event that the best price appears in some day k. Here we abstract from the market rules and assume that all orderings of prices are equally likely. Then, considering the (1) condition above:

$$P(B_k) = \frac{1}{N}, \mathrm{k} > \mathrm{n}$$

Let R_k denote the event describer by (2) condition above. R_k occurs if the maximum of the first k - 1 days appears in the first n days, so:

$$P(R_k) = \frac{n}{k-1}$$

.

B_k and R_k are independent and probability of success of strategy S_n, denoted by $P_N(S_N)$ can be calculated as intersection of the conditions above:

$$P_N(S_N) = \sum_{k=n+1}^{N} P(B_k \cap R_k) = \sum_{k=n+1}^{N} P(B_k)P(R_k) =$$

$$\frac{n}{N}\sum_{k=n+1}^{N}\frac{1}{k-1}$$

As it was proved [1]:

$$P_N(S_N) \to max, \text{ as } n \to e^{-1}$$

$$\lim_{N\to\infty} P_N(S_N|n = e^{-1}) = e^{-1}$$

so we used n = e^{-1}, which is approximately 37%.

We estimated the best price using the golden rule, and compared it to the actual maximum price during month by calculating the difference between them. To make our results more informative, we normalized these differences, dividing them by sample span (At first, our key parameter was just the difference between maximum price and the price we got from the golden rule because these differences were distributed with $\mu = 7$. But, after that, we got to the point that our price data span at some periods could be 7, or even 4, which means, what we thought to be good estimated price could have been the worst one).

Here is what we've got (the golden-rule estimations are represented as a dot plot:

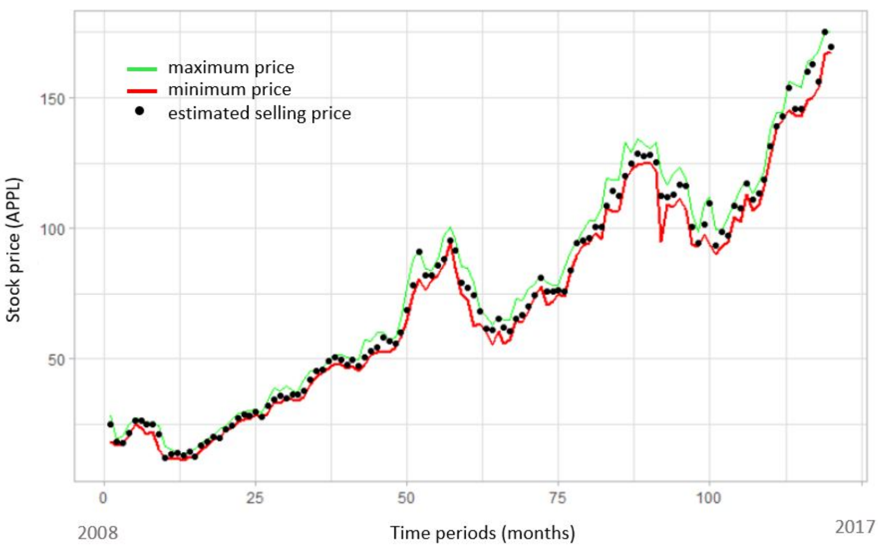

Figure 1: Golden rule estimation and real prices, monthly

The distribution of relative differences contained 0's (which means we obtained the real maximum with golden rule), and 1's (which means epic fail). In general, it wasn't as good as we expected it to be - we've got the distribution with mean = 0.51 and variance 0.07. On average, our estimated maximums are located close to the month average, not maximum.

III. EXPECTED MAXIMUM APPROACH

Actually, golden rule approach maximizes the probability of getting top-1 maximum price, but in most cases we obtain mediocre results. In looking for a price to sell people rather interested in always getting satisfactory price, that is close to maximum. The approach we suggest: calculate the expected maximum for given time frames and accept any price that appears on stock and is higher than or equal to expected maximum.

Let's denote by M the maximum that we get among month stock prices:

$$M = max(x_1, x_2, ...x_n), \ \textit{where n is the number of days}$$

The changes of stock's prices can be described by a random walk process. A random walk is a stochastic or random process, that describes a path that consists of a succession of random steps on some mathematical space. In our case stock prices travel the real line, for each month there is a starting point, first day price. Every day price has some probability p to increase by some percents, let's say u, and some probability q to decrease by some percents d.

Parameters estimation

So, we need the following parameters:

$$p = \{\text{probability to go up}\}$$

$$q = \{\text{probability to go down}\}$$

$$u = \{\text{going up value}\}$$

$$d = \{\text{going down value}\}$$

The values above one can calculate empirically from the previous data. For example, I want to know when in April should I sell my Apple stocks and I have March data. Firstly, let's find probabilities:

$$p = \frac{|N^+|}{n}, \text{ where } N^+ \text{ is set of days when price goes up}$$

$$q = 1 - p$$

So, probabilities could be calculated as relative frequencies, it is justified by the Law of Large Numbers.

Now let's find the estimator for u and d. Suppose I_1 is an indicator that in that day the price went up and I_2 the price went down.

$$I_1(x_i) = \begin{cases} 1 & \text{i} \in N^+ \\ 0 & \text{otherwise} \end{cases}$$

$$I_2(x_i) = 1 - I_1(x_i)$$

$$\hat{u} = \sqrt[\sum_i I_1(x_i)]{\prod_i (\frac{x_i}{x_{i-1}})^{I_1(x_i)}}, \ i > 1$$

$$\hat{d} = \sqrt[\sum_i I_2(x_i)]{\prod_i (\frac{x_i}{x_{i-1}})^{I_2(x_i)}} \ i > 1$$

It is important that we calculate u and d as relative values, so we go in different direction not by some value, but some **percentages**. Let's suppose $|N^+| = k$, so $|N^-| = n - k$ and we always go up by u, go down by value d. One can prove that $\hat{u}$ and $\hat{d}$ are unbiased estimators, i $>$ 1:

$$E[\hat{u}] = E(\sqrt[\sum_i I_1(x_i)]{\prod_i (\frac{x_i}{x_{i-1}})^{I_1(x_i)}}) = E(\sqrt[k]{u^k}) = u$$

$$E[\hat{d}] = E(\sqrt[\sum_i I_2(x_i)]{\prod_i (\frac{x_i}{x_{i-1}})^{I_2(x_i)}}) = E(\sqrt[n-k]{d^{n-k}}) = d$$

The estimators û and $\hat{\text{d}}$ are unbiased! Now one can easily calculate the estimates using, for example, R and previous data.

Calculating expected maximum

If one know the parameters of the random walk, expected maximum one can find via total rule for expectations. We applied a little trick: used $\log(u)$ and $\log(d)$ instead of u and d for simplifying calculations and considered M_g as the maximum growth rate over the whole month.

$$M_g = max(\log(\frac{x_1}{x_1}), \log(\frac{x_2}{x_1}), ... \log(\frac{x_n}{x_1}))$$

$$E[M_g] = \sum_{m=0}^{n\log(u)} m_g P(M_g = m_g)$$

We used such upper bound for maximum, because the price cannot growth in more times than u^n, by taking logarithm, one obtain nlog(u).

Probability that we will get a certain maximum can be described as following:

$$P(M_g = m_g) = \sum_{i=0}^{n} P(M_g \geq m_g + 1|N^+ = i)P(N^+ = i) - P(M_g \geq m_g|N^+ = i)P(N^+ = i)$$

We conditioned on having exactly i days in which price goes up. Distribution of N^+ is binomial, can be considered as number of successes in n number of trials:

$$N^+ \sim \beta(n, p)$$

$$P(N^+ = i) = \binom{n}{i} p^i q^{n-i}$$

Let's find $P(M_g \geq m_g|N^+ = i)$. Let's denote by v = $\{x_n\}$, the price in the end of the month. We know the number of days when the price goes up and, automatically, we know the number of days it goes down. So,

$$v = i\log(u) - log(d)(n - i) = i\log(\frac{u}{d}) + n\log(d)$$

We should consider two cases:

$$v \geq m_g, i \geq \frac{m - n\log(d)}{\log(\frac{u}{d})}$$

$$v < m_g, i < \frac{m - n\log(d)}{\log(\frac{u}{d})}$$

In the first case $P(M_g \geq m_g|N^+ = i) = 1$, for reaching v we have to cross the m_g.

In the second case it is not so easy. Here we applied a **reflection principle**[2]:

The number of passes from x_1 to x_n with touch or cross the $y = m_g$ is equal to the number of all passes from x_1 to $\acute{x}_n$, where $\acute{x}_n$ is symmetric reflection of x_n with respect to $y = m_g$

So, if

$$2l - n = 2m - i\log(\frac{u}{d}) + n\log(d)$$

$$l = m - \frac{1}{2}i\log(\frac{u}{d}) + \frac{n}{2}(\log(d) + 1$$

$P(M_g \geq m_g | N^+ = i) = \binom{l}{n}/\binom{i}{n}$, as the relation of number of ways to v, where we cross or touch m_g to the whole number of ways to v. We can update and rewrite:

$$S_m = \sum_{i=0}^{n} P(M_g \geq m_g | N^+ = i)P(N^+ = i) = \sum_{i=0}^{\lceil \frac{m - n\log(d)}{\log(\frac{u}{d})} \rceil} \binom{l}{n} p^i q^{n-i} + \sum_{\lceil \frac{m - n\log(d)}{\log(\frac{u}{d})} \rceil}^{n} \binom{i}{n} p^i q^{n-i}$$

$$S_{m+1} = \sum_{i=0}^{n} P(M_g \geq m_g + 1 | N^+ = i)P(N^+ = i) =$$

$$= \sum_{i=0}^{\lceil \frac{m+1-n\log(d)}{\log(\frac{u}{d})} \rceil} \binom{l+1}{n} p^i q^{n-i} + \sum_{\lceil \frac{m+1-n\log(d)}{\log(\frac{u}{d})} \rceil}^{n} \binom{i}{n} p^i q^{n-i}$$

$$E[M_g] = \sum_{m_g=0}^{n\log(u)} m_g(S_{m+1} - S_m)$$

Actual expected maximum price one can express as compound interests:

$$E[M] = x_1 u^{E[M_g]/\log(u)}$$

where u is the interest rate and $E[M_g]/\log(u)$ is the length of growth period.

Experiment

Similar to the golden rule experiment, we used the Apple stock prices monthly data to test the results of this models. As it was discussed previously, we took one month (i.e. January 2008) to calculate the model parameters. At this point, we have our probabilities of going up and down, and the values by which the stock price increases or decreases on every next day. After that, we move forward to the next month. Going forward day by day, our task is to choose the right moment to sell the stock. For this, we need to find the expected maximum for the upcoming period. The way we did it is precisely described in the previous section. As an outcome, we designed a function, which takes 6 input parameters:

- start: the starting price (stock price on the first day in a month)

- p, q : probability that price increases or goes down on the next day
- u, d : values by which price changes
- n : length of the period (number of days in month)

The return value of this model is the expected maximum price for the following month. After we know the expected maximum price, it becomes easier for us to make the decision on selling stock. Our decision-making rule is simple: we choose to sell on the day i, if the i-th price is larger or equal to the evaluated expected maximum price.

We repeated this process and determined the selling day for each of 120 months we have, to collect some data and analyze it. Here are the results of expected maximum estimation (the estimations are represented as dots):

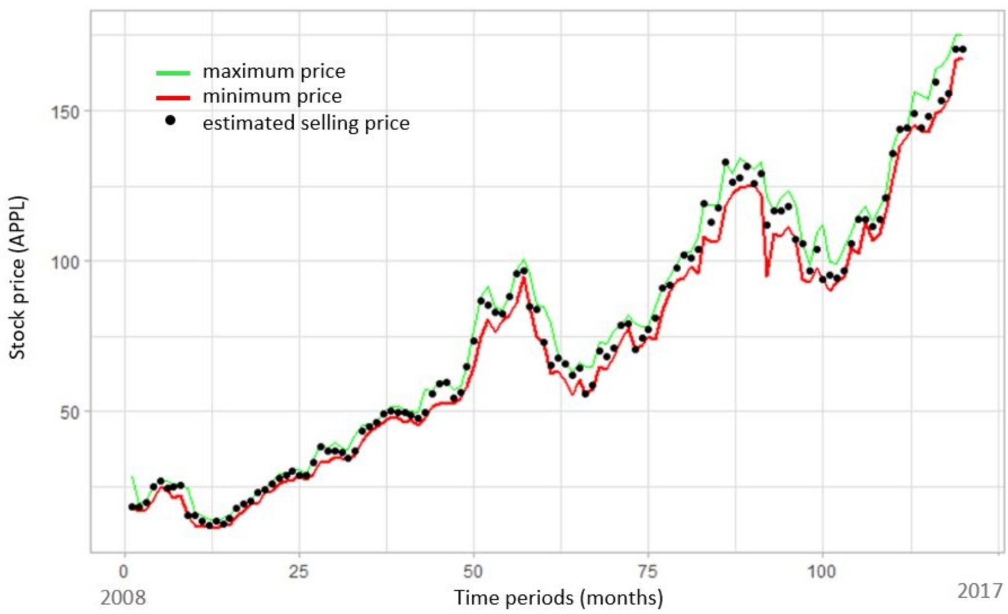

Figure 2: Expected maximum estimation and real prices, monthly

For this approach, we also found the distribution of relative differences, which was very similar to the golden rule one, with mean = 0.50 and variance 0.08. On average, our estimated maximums are located close to the month average.

We also tested one modification of this approach (mentioned above as expected maximum*), which showed better results. The main difference between these two is, that in the last one, we update our parameters frequently (in this case - every day). So, for every next day, we calculate new expected maximum, using updated information on u, d parameters that we obtain as we move forward.

IV. RESULTS

Except using relative differences between estimated maximums and real values of maximums, we also analyzed the ranks of estimated parameters in sorted monthly price sequences. For example, if the estimated maximum gets the first rank - we win, because the model produced the best result, real maximum price.

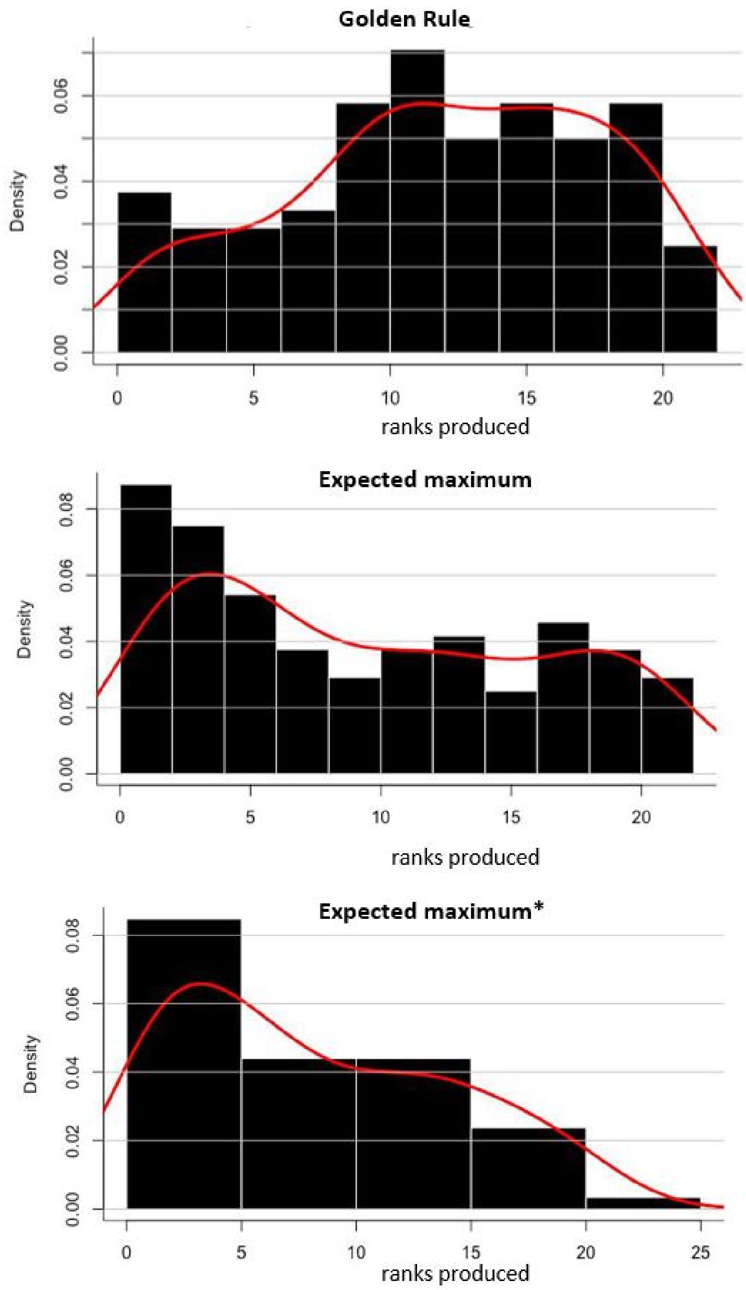

Figure 3: Ranks produced using different approaches

The mean values of the ranks produced are he following: 9.87 for Golden Rule, 9.55 for expected maximum, and 8.15 for expected maximum*. Even though means of distributions of the values themselves, and means of distributions of the ranks are very close to each other, from the visualization we can observe that the expected maximum estimation is better. The expected maximum produces much more values ranked 'first' or 'second', than the golden rule estimation.

What makes us extremely happy here, is the result of expected maximum approximation with changing parameters (3rd plot). This result is intuitive, because we obtained it by involving new and new data from the market in the model on every different day. This gave the opportunity to avoid wrong choices based on the data from previous month, and concentrate on following the trends of current time period. As one can see of the visualization, this method produced the greatest amount of top-ranked results, and it almost always avoids falling down into minimum instead of maximum.

[1] Dimitris A. Sardelis and Theodoros M. Valahas, *Decision Making: A Golden Rule* (The American Mathematical Monthly Vol. 106, No. 3 (Mar., 1999)), pp. 215-226

[2] Feller William , *An Introduction to Probability Theory and Its Applications* (Vol. 1, 3rd Edition)